

F. GASSE

F. TÉKAÏA

## **La paléoécologie des diatomées : évolution des lacs de l'Afar central**

*Les cahiers de l'analyse des données*, tome 4, n° 1 (1979),  
p. 81-94

[http://www.numdam.org/item?id=CAD\\_1979\\_\\_4\\_1\\_81\\_0](http://www.numdam.org/item?id=CAD_1979__4_1_81_0)

© Les cahiers de l'analyse des données, Dunod, 1979, tous droits réservés.

L'accès aux archives de la revue « Les cahiers de l'analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

# LA PALÉOÉCOLOGIE DES DIATOMÉES ÉVOLUTION DES LACS DE L'AFAR CENTRAL [DIATOMÉES]

par F. Gasse <sup>(1)</sup> et F. Tékaïa <sup>(2)</sup>

La thèse de Madame Gasse porte sur : L'évolution des lacs de l'Afar central du Plio-pleistocène à l'Actuel et la reconstitution des paléomilieux lacustres à partir de l'étude des diatomées.

Son étude minutieuse des associations de diatomées dans les coupes géologiques rencontrées pour les différents épisodes lacustres et dans les différents bassins a fourni un volume de données quantitatives considérable. C'est une partie des échantillons qu'elle a analysés pour sa thèse que nous avons soumise à l'analyse factorielle des correspondances.

Nous proposons dans l'introduction de préciser les intérêts géologiques des diatomées, la région où ont eu lieu les prélèvements des échantillons, la méthode d'évaluation des fréquences des diverses espèces dans chaque échantillon ; nous présenterons ensuite les différentes analyses effectuées.

## 1 Introduction

### 1.1 Les diatomées ; leurs intérêts géologiques

#### 1.1.1 Définition des diatomées

Les diatomées ( ou *diatomophycees* ou *bacillariophycees*) sont des algues unicellulaires ou coloniales à plastes bruns ou jaunes, caractérisées avant tout par leurs parois cellulaires imprégnées de silice et formant une logette bivalve appelée frustule : les deux valves sont ajustées comme une boîte et un couvercle. Ces deux valves du frustule minéral se conservent dans le sédiment et c'est à partir de leur ornementation que se font les déterminations spécifiques. Les diatomées sont d'ailleurs bien connues des microscopistes amateurs, car elles sont souvent très belles et sont faciles à observer. Certains sédiments siliceux sont constitués quasi-exclusivement de frustules de diatomées, plus ou moins brisés : c'est la diatomite ou *kieselguhr*.

#### 1.1.2 La paléocologie des diatomées

Dans une large mesure, la paléontologie est inséparable de la paléocologie ; c'est-à-dire de l'étude des associations des espèces fossiles dans les gisements, témoins de milieux disparus. Parmi les organismes fossilisés dans les milieux lacustres, les diatomées représentent un outil précis pour reconstituer les paléoenvironnements. En effet, l'intérêt majeur des diatomées est d'être d'excellents indicateurs biologiques.

---

(1) Docteur-ès-sciences. E.N.S. de Jeunes Filles - Fontenay-aux-Roses

(2) Docteur 3<sup>o</sup> cycle, laboratoire de statistique. Université P.&M. Curie

Leurs associations sont des enregistreurs précis des conditions des milieux (profondeur des eaux, luminosité, températures, alcalinité, salinité, richesse en gaz dissous et en matière organique...). La paléoécologie des diatomées a permis, en Afar, de retracer l'évolution des milieux de sédimentation au cours du temps et de comparer, à une époque donnée, plusieurs bassins lacustres. Les associations des diatomées ont enregistré, au cours de chaque épisode lacustre, l'évolution de l'environnement (caractères écologiques des lacs). Cet environnement dépend des conditions locales (nature du substrat rocheux, géochimie des eaux d'alimentation...) mais aussi des conditions climatiques régionales.

### 1.1.3 Intérêt stratigraphique

Les diatomées se révèlent aussi être de très bons marqueurs chronologiques. L'évolution et la différenciation spécifique des diatomées a déjà conduit à établir de bonnes échelles stratigraphiques en particulier en Afrique, dans le bassin du Tchad par exemple. Certaines espèces rencontrées dans les sédiments anciens sont actuellement éteintes ou en voie d'extinction, et leur présence permet de caractériser certaines époques géologiques. *Stephanodiscus transylvanicus*, *S. rarusonnensis*, *S. niagarae*,... étaient très abondants en Afar central au Plio-pleistocène (4 à 0,7 millions d'années (MA) avant nous) et avaient disparu de cette région au Pleistocène moyen (0,7 à 0,2 MA avant nous).

Dans l'interprétation des analyses on trouvera les termes de Plio-pleistocène, Pleistocène moyen et aussi Pleistocène supérieur (de 200.000 à 10.000 ans avant nous) et Holocène (de 10.000 avant nous à l'Actuel).

Pour collaborer avec le spécialiste, le statisticien doit se familiariser avec quelques termes, définissant le mode de vie ou le milieu d'élection des diatomées :

planctonique = de pleine eau ; littoral , généralement benthique ;  
lié au fond ;

oligohalobe = vit dans un milieu de faible salinité (0,2‰ à 2‰)

mesohalobe = de 2‰ à 16‰ ; polyhalobe = de 16‰ à 35‰ ;

hyperhalobe = plus de 35‰ ...

Nous rappellerons le sens de ces termes quand ils reviendront dans le texte.

plaste : organite cellulaire qui sert pour la photosynthèse.

### 1.1.4 Domaine de l'étude

En Afar central (Ethiopie et Territoire de Djibouti), actuellement désertique, les phases climatiques humides du Quaternaire ont entraîné l'installation de lacs profonds occupant les nombreux fossés tectoniques effondrés dans d'épaisses séries basaltiques, datées de 4 à 1 MA.

Au cours du Quaternaire, l'évolution climatique de cette région a été extrêmement complexe et de nombreuses phases lacustres se sont succédées. L'existence de ces anciens lacs est enregistrée par des sédiments affleurant sur le fond des bassins. Ces sédiments contiennent les restes d'organismes qui vivaient dans ces anciens lacs. Certains de ces organismes, les diatomées, sont des plus intéressants pour les études géologiques.

### 1.2 Evaluation de l'abondance des différentes espèces de diatomées dans chaque échantillon

Dans certains cas les sédiments riches en diatomées sont à l'air libre ; parfois il faut les prélever par carottage de fonds lacustres. Nous verrons au § 3 comment une même carotte peut révéler toute l'histoire d'un lac.

Lors de la préparation des lames de diatomées destinées à l'observation en microscopie optique, on part toujours d'un poids connu de sédiment sec (1g le plus souvent). Après traitement chimique, lavages et centrifugation, le culot de centrifugation est dilué dans un volume d'eau connu (généralement 200 ml). La suspension ainsi obtenue est homogénéisée par l'action des ultra-sons. Le volume de suspension prélevé à la micropipette automatique est fixé avec précision. La fréquence relative de chaque espèce rencontrée a été calculée après comptage de 500 à 1000 individus (chaque valve étant considérée comme un individu) répartis sur 4 lames. La valeur minimale du nombre d'individus nécessaires pour que le pourcentage soit représentatif a été déterminée expérimentalement. En traçant la courbe des valeurs de fréquences obtenues en fonction du nombre d'individus comptés, on constate que cette valeur oscille toujours largement tant que le nombre d'individus est inférieur à 500 environ et qu'elle se stabilise peu à peu ensuite. Par ailleurs, ces fréquences varient d'une lame à l'autre, ce qui prouve la non représentativité d'une seule préparation par rapport à l'échantillon. C'est pourquoi les comptages ont toujours été réalisés à partir de 4 lames. Les valves cassées ont été comptées comme 1/2 individu lorsque l'espèce est reconnaissable. Les échantillons dont les frustules sont très fragmentés n'ont pas donné lieu à un comptage.

Ainsi est finalement défini le pourcentage relatif de l'espèce  $i$  dans l'échantillon  $j$ , c'est ce que nous appelons l'abondance relative  $k(i,j)$ ; le total de chaque colonne  $j$  est donc constant. Les préparations étant toujours menées quantitativement, il est de plus possible de calculer le nombre  $k'(i,j)$  de l'espèce  $i$  contenu dans 1g de sédiment de l'échantillon  $j$ .

Grâce à cette procédure rigoureuse de dilution des sédiments, il est possible de calculer l'abondance absolue des diverses espèces i.e. le nombre d'individus (valves) de chaque espèce rapporté à 1g de sédiment. Toutefois dans les analyses factorielles on s'est attaché uniquement aux *abondances relatives* i.e. au rapport du nombre d'individus de l'espèce  $i$ , trouvée dans une masse de sédiment  $j$ , au nombre total de valves contenues dans cette masse. Plus précisément voici comment a procédé Madame Gasse, afin d'éviter les fluctuations d'échantillonnage : pour chaque sédiment  $j$ , elle a compté les diatomées étalés sur les plaques, jusqu'à atteindre le total de 1000 valves :  $k(i,j)$  est le nombre des valves de l'espèce  $i$  trouvées dans le sédiment  $j$ ; et  $k(j)$  total de la colonne afférant au relevé  $j$  vaut donc toujours 1000.

La plus grande critique que l'on puisse faire à cette méthode est le trop faible volume prélevé à la micropipette. Il est en effet très difficile d'obtenir une suspension homogène de frustules de diatomées, et le volume de 0.2 ml (volume de la goutte déposée sur la lamelle) peut ne pas être représentatif de l'échantillon. De plus, la répartition des diatomées sur la lamelle, après séchage est souvent assez irrégulière. Les résultats obtenus dans les conditions décrites ci-dessus paraissent toutefois reproductibles et ont donc été considérées comme valables.

## 2 Différentes analyses effectuées pour cette étude

L'étude considère un ensemble I de 336 espèces de diatomées définies par leur abondance dans un ensemble J de 291 échantillons ou relevés.

Nous avons envisagé deux points de vue :

- 1) en ne tenant compte que de la simple présence (notée 1) ou absence (notée 0) de l'espèce  $i$  dans l'échantillon  $j$  ;
- 2) en considérant le tableau brut  $k_{IJ}$  où  $k(i,j)$  désigne l'abondance de l'espèce  $i$  dans l'échantillon  $j$ .

Dans le premier cas, nous avons fait les analyses suivantes :

PRESABS I : l'analyse factorielle du tableau en 0 et 1 de dimension (336,291) a permis de dégager les groupements d'échantillons et d'espèces les plus particuliers.

PRESABS II : afin de faire éclater la masse d'échantillons et d'espèces se trouvant autour du centre de gravité dans l'analyse PRESABS I, nous avons mis les groupements déterminés par celle-ci en éléments supplémentaires. L'analyse du tableau ainsi obtenu a permis de dégager une structure restée invisible par PRESABS I.

CAH ESPECES : la classification ascendante hiérarchique de ce même tableau nous a permis de comparer les groupes obtenus par celle-ci avec ceux de l'analyse PRESABS I et PRESABS II.

ABONDANCE : dans le deuxième cas, nous avons analysé le tableau des données brutes.

Chacune de ces études a permis de dégager une partition de l'ensemble des échantillons, mais il est difficile de caractériser les groupements des échantillons par les espèces. Afin de pallier cette difficulté, nous avons étudié les tableaux de contingence  $k_{IQ}$  où I est l'ensemble des espèces et Q la partition déterminée soit par PRESABS soit par ABONDANCE.

Nous avons à cet effet effectué deux analyses : GROUPABOND et GROUPRES.

## 2 Analyses de correspondances entre espèces et relevés

Dans les études phytosociologiques ou écologiques, les analyses en présence-absence donnent souvent des résultats plus clairs que les analyses en abondance relative, dominées par les espèces qui prolifèrent le plus, mais n'ont pas de valeur discriminante appréciable. Toutefois dans la présente étude, il apparaîtra (cf § 2.3) que les données en abondance certes très difficiles à recueillir (cf § 1.2) sont à peu près irremplaçables ; ceci est sans doute dû au caractère particulier des relevés étudiés ici : ce ne sont pas des relevés phytosociologiques contemporains, effectués en cueillant un jour déterminé d'une certaine saison, les végétaux présents ; ce sont des moyennes cumulantes même sur les plus faibles épaisseurs qu'on puisse distinguer, toutes les saisons de multiples années successives !

2.1 Analyses en présence-absence : Nous l'avons dit : I = ensemble de 336 espèces ; J = ensemble de 291 relevés ; mais dans la présente analyse  $k(i,j) = 1$  si l'espèce i est présente dans le relevé j ; et zéro sinon.

Une première analyse PRESABS I du tableau complet 336 x 291 a permis de dégager 5 classes interprétables de relevés, associées à des espèces qui indiquent les caractères des milieux. Mais au centre du graphique la majorité des relevés se groupe en une classe compacte, certainement hétérogène, mais dont la première analyse n'indique pas la structure.

On a donc fait une deuxième analyse PRESABS II en mettant en éléments supplémentaires les 67 échantillons constituant les 5 classes distinguées d'après l'analyse PRESABS I, ainsi que les 72 espèces associées à ces classes sur les plans de l'analyse factorielle. Ainsi il devient possible de distinguer de nouvelles classes de relevés ; à condition de reporter sur les graphiques l'origine des échantillons, le mode de vie des espèces (planctonique = pleine eau ; ou littoral), la salinité qu'elles affectionnent : pour cela au lieu de représenter seulement les individus i ou j par un numéro on les représente par un sigle (carré, triangle, étoile, etc.) symbolisant des informations supplémentaires.

Dans cet article, nous ne donnerons aucun détail sur les analyses PRESABS : il suffit d'indiquer au statisticien le principe des analyses successives et de leur interprétation.

## 2.2 L'analyse factorielle du tableau des abondances relatives

C'est le tableau construit par comptage, comme nous l'avons expliqué au § 1.2.

$k(i,j)$  = abondance de l'espèce  $i$  dans l'échantillon  $j$ .

Les valeurs propres et pourcentages d'inertie décroissent lentement:

$$\lambda_1 = 0,832 ; \lambda_2 = 0,761 ; \dots ; \lambda_6 = 0,571 ;$$

$$\tau_1 = 5,39\% ; \tau_2 = 4,9\% ; \dots ; \tau_6 = 3,9\% ;$$

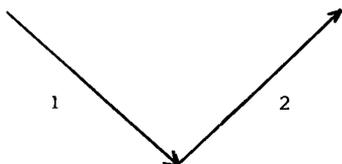
le pourcentage cumulé des six facteurs est de 27,4%. Au vu des plans croisant deux à deux les 4 premiers facteurs issus de l'analyse ABONDANCE, on a déterminé 9 classes de relevés ; les résultats ainsi obtenus s'accordent avec ceux de la classification automatique effectuée en tenant compte des 7 premiers facteurs (cf § 4.2).

Le présent exposé est destiné à des statisticiens non spécialistes des diatomées ; sans prétendre donner tout le détail des résultats, on en suivra les grandes lignes illustrées de quelques exemples.

### 2.2.1 Le plan 1×2 de l'analyse en abondance

Considérons d'abord le plan 1×2, où les 9 classes retenues n'empiètent guère, même s'il est impossible de les délimiter sans le secours des facteurs 3 et 4.

Le nuage des échantillons se présente suivant deux branches en V (cf figure).



La direction 1 porte les associations littorales en les présentant par ordre décroissant suivant le degré de salinité.

La direction 2 porte les associations planctoniques (de pleine eau).

Du groupe 1 au groupe 8, on observe un axe de salinité décroissante, quel que soit l'âge des sédiments.

Par exemple :

Les groupes 1 et 2 sont voisins, pourtant 1 est d'âge pleistocène, alors que 2 est d'âge holocène. La séparation entre les groupes 1 et 2, sera expliquée par le facteur 3 (cf *infra*).

Le groupe 4 est d'âge pleistocène supérieur alors que le groupe 8 est principalement d'âge holocène.

Quelques échantillons appartenant à différentes sections holocènes mais marquant toujours des stades provisoires de salinité élevée (par exemple 1505) présentent, malgré leur âge récent, de fortes analogies floristiques au point de vue quantitatif avec les flores plio-pleistocènes du Dikhil et d'Asa Gegalou.

Sur la deuxième branche, celle des associations planctoniques, considérons en détail une classe, la classe 7 subdivisée en classes 7a, 7b et 7c, afin de suivre sur une carotte l'histoire d'un lac : le lac Abhé.

Les premiers échantillons 7a représentent la base de la carotte, comprise entre 45.95 et 33.65 mètres au dessous du niveau actuel. Cette partie de la carotte correspond au premier épisode lacustre et présente une certaine homogénéité, bien que la teneur en frustules de diatomées y soit très variable. Elle est surtout caractérisée par : *Stephanodiscus astrea v. intermedia* et *S. astrea v. minutula*. Ce sont des espèces planctoniques. La constance et l'abondance des argiles entre 50 et 33.65 mètres constituent le caractère essentiel de ce premier épisode lacustre. Ce fait oppose ce premier épisode à la phase lacustre suivante, dont la liste d'espèces de diatomées n'est pas fondamentalement différente, mais dont la sédimentation est purement biogénique.

Le groupe 7b représente le second épisode lacustre observé dans la carotte du lac Abhé (entre 29.89 et 21.43 m). Il est enregistré par une diatomite très pure (sédiment orné exclusivement de diatomées) encadrée par des argiles calcaires marquant les phases de transgression et de régression. La phase d'extension maximale du lac est très originale par sa sédimentation purement biogénique et se caractérise par une flore très exigeante en matière organique : nombreux *Nitzschia*.

Après une régression brutale et passagère, le lac remonte à nouveau ; une flore planctonique à *Melosira* moins exigeante vis-à-vis de la matière organique et de la température s'installe ; elle est représentée par le groupe 7c.

### 2.2.2 Le plan 1 x 3

Du côté  $F_1 > 0$  et  $F_3 < 0$  on distingue nettement un groupe d'échantillons relatifs au lac Abhé et correspondant à la deuxième phase lacustre du Pleistocène supérieur. Ils caractérisent la régression extrêmement brutale du lac. Ces échantillons se détachent de la suite d'échantillons qui caractérisent la transgression de la même phase lacustre (augmentation de la masse des eaux).

On remarque aussi que le troisième axe oppose les deux groupes 1 et 2 qui sont voisins sur le premier plan factoriel.

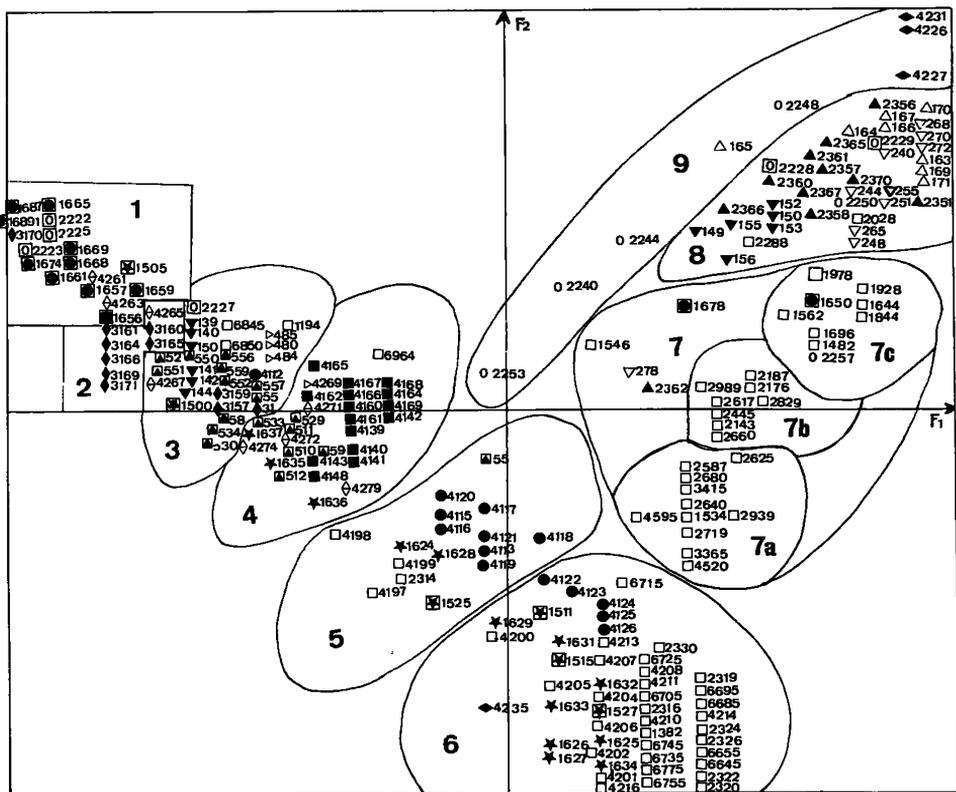
La séparation entre les classes 1 et 2 s'explique par le stock d'espèces relatif aux lacs Dikhil et Afrera qui est distinct, dû vraisemblablement aux faciès chimiques des eaux qui sont différents. En effet, au Dikhil et à Asa Gegalou, on rencontre de nombreuses espèces des eaux sulfatées calciques et leur régression conduit à la précipitation du gypse ( $\text{CaSO}_4$ ) en épaisse dalle, ce qui témoigne des caractères chimiques des eaux de l'ancien lac. Par contre à Afrera, la plupart des espèces caractéristiques vivent actuellement dans des eaux chlorurées sodiques. Aujourd'hui, le lac Afrera contient 159 g/l de sel et cette salinité est essentiellement due aux ions  $\text{Cl}^-$ ,  $\text{Na}^+$ .

### 2.3 Comparaison des analyses factorielles en présence et en abondance

L'analyse en présence-absence permet de définir une étude floristique (étude de l'ensemble des espèces qui composent la flore) et de dégager des groupements à espèces caractéristiques, différentielles ou exclusives. La présence de ces espèces exclusives peut s'expliquer :

- soit par des facteurs écologiques très originaux (par exemple 6860, 1194... sur le plan (1x3) PRESABS qui correspondent à une phase aride
- soit par la répartition stratigraphique très précise de ces espèces (par exemple certaines espèces du plan (1x2) de PRESABS) qui n'ont existé en Afar que pendant une période bien précise.

Par contre, d'autres groupes ne contiennent aucune espèce différentielle mais ont cependant une originalité due à l'abondance de certaines espèces.



9 groupes de relevés déterminés d'après les 4 premiers facteurs de l'analyse en abondance. Les sigles indiquent les lieux de prélèvements : (■) Dikhil, (□) Asa Gegalou, (■) Allomboulto, (◆) Afréra, (▲) Loggha et Saha, (◇) Gabilemma, (★) Asela, (●) Dobi Nord, (□) Abhé, (▲) Tikibie Deri, (▼) Erolé, (○) Maydalala, (◆) Bodo Le dabba, (▶) Boralé Ouest, (△) Erolé ancien, (■) Boralé Est, (▽) Malhadlou.

Les nombres indiquent les numéros de relevés.

Les relevés se regroupent suivant les degrés de salinité et le mode de vie des diatomées qui les définissent.

C'est par exemple le cas du groupe 6 (ABONDANCE), plan (1x2)) qui se définit par le très fort pourcentage de *Stephanodiscus astrea v. intermedia* ou *minutula* présents dans de très nombreux échantillons de tous les âges.

Ils ne sont par contre abondants que dans le plancton des lacs holocènes. Ce groupe 6 ne peut donc être mis clairement en évidence qu'en tenant compte de l'abondance des espèces citées. De plus, si l'on considère une coupe géologique donnée, représentant l'ensemble d'un épisode lacustre, on constate que très généralement, les échantillons correspondants sont tous très bien groupés si l'on ne considère que la simple présence ou absence des espèces. Ceci s'explique par le fait que chaque phase lacustre dans un bassin donné se caractérise par tout un lot d'espèces, dont le pourcentage relatif change considérablement, parallèlement aux fluctuations écologiques : transgression, optimum lacustre, régression. Il est intéressant de remarquer que ces fluctuations écologiques apparaissent nettement lorsqu'on tient compte de l'abondance relative.

Par exemple : les échantillons 4113 à 4124 (relatifs au lac Dobi Nord d'âge holocène supérieur) s'échelonnent progressivement le long de l'axe de salinité (ABONDANCE plan (1x2)) marquant ainsi les phases de transgression et de régression ; ceci correspond à l'approfondissement progressif, parallèlement à une dilution des eaux du lac en question.

Par contre, ces détails ne ressortent pas clairement de l'analyse (PRESABS, plan 1x2)), même si l'on y remarque deux échantillons qui se détachent du groupe.

Autre exemple : les échantillons (139 à 155), bien que subdivisés en deux sous-groupes, comme nous le voyons sur le plan (1x2) de ABONDANCE ne le sont pas aussi clairement sur le plan (1x2) de PRESABS. Ces échantillons appartiennent tous à la même coupe et possèdent à peu près le même stock floristique. Par contre, ils correspondent à deux phases lacustres superposées (2° et 3° phases lacustres du Pleistocène supérieur), d'où un changement très important dans les fréquences des espèces.

Les espèces littorales d'eau douce ou légèrement salée sont très abondantes entre 139 et 147 (*Nitzschia amphibia*, *Epithemia argus*, *Epithemia zebra*). Par contre, elles sont présentes mais peu abondantes entre 149 et 155. Elles sont alors relayées par l'espèce planctonique *Melosira granulata*.

### Conclusion

Les analyses en présence-absence montrent des groupements bien définis les uns par rapport aux autres, ainsi que des groupements isolés les uns des autres, on arrive à mettre en évidence cet isolement en mettant certains éléments apparaissant dans une première analyse en éléments supplémentaires, etc.) mais il est souvent plus difficile de comprendre les facteurs qui les déterminent.

Par contre, avec les analyses en ABONDANCE, il se dégage non seulement des groupements bien définis, mais en plus, c'est la dynamique de l'évolution lacustre liée aux fluctuations climatiques qui apparaît très clairement entre les différents groupes.

### 3 Les analyses en données groupées

Les travaux de Briane, Lazare et Salanon (cf Cahiers Vol III n° 2) ont montré l'intérêt des analyses en données groupées dans lesquelles, au lieu de considérer des relevés individuels, on prend des classes de relevés définies en cumulant les vecteurs de description des relevés individuels. Chez Briane et Coll, les groupes de relevés sont constitués automatiquement par un algorithme tel que celui de E. Diday : l'objectif principal est de diminuer les dimensions du tableau soumis à l'analyse factorielle, et de découvrir aussi la structure du tableau initial

1000 x 1000. Mis en terme écologique le succès de la méthode s'explique parce qu'à une classe de relevés correspond un type de milieu apte à recevoir la vie, c'est-à-dire ce qu'on appelle une *niche* ; or le nombre de *niches* ne peut pas être très élevé, à la différence du nombre des espèces, sur lesquelles pour cette raison il serait imprudent d'effectuer des groupements de données.

Dans la présente étude, le volume des tableaux a permis de faire une analyse factorielle directe ; toutefois le groupement des données conserve, on le verra, son intérêt. Mais ici il a été effectué au vu des résultats de l'analyse factorielle, et non automatiquement. Notons encore qu'un groupe de relevés est ici caractéristique non seulement d'une niche (e.g. le milieu aqueux fortement minéralisé en  $\text{Ca SO}_4$ ), mais d'une niche à une époque donnée : car à chaque époque des espèces différentes sont susceptible de coloniser une niche donnée.

### 3.1 Analyse GROUPABOND

Afin de dégager le rapport entre les espèces et les différents groupes dégagés par l'analyse ABONDANCE, nous avons considéré le tableau  $k_{IQ}$  où I est l'ensemble des espèces et Q est la partition en 9 classes de l'ensemble des relevés J, réalisée d'après l'analyse factorielle ABONDANCE.

Nous avons défini ce tableau comme suit :

Soit i une espèce de I et q une classe de Q

$$k(i,q) = \sum \{t(i,j) \mid j \in q\}$$

où  $t(i,j)$  = abondance relative de i dans j

= somme des abondances de l'espèce i dans les relevés

appartenant à q

= somme des fréquences de l'espèce i dans les observations

appartenant à q.

Pour plus de précision, nous aurions dû pondérer chaque  $k(i,q)$  par le nombre de relevés appartenant à la classe q.

*Interprétation des résultats de l'analyse factorielle*

*Inertie et pourcentages relatifs aux 7 premiers facteurs*

	F1	F2	F3	F4	F5	F6	F7
valeur propre	0.8	0.717	0.59	0.56	0.39	0.378	0.259
Pourcentage	20.4	18.1	14.9	14.1	10.	9.5	6.5

Regardons le plan factoriel  $1 \times 2$  qui explique 38.5% de l'inertie totale.

La détermination des espèces caractéristiques des groupes considérés est faite directement sur le tableau de contingence  $k_{IQ}$ , alors que le mode de vie, l'âge et la salinité des espèces se font à partir des résultats donnés par ce premier plan factoriel, où apparaissent les groupes de relevés (de G1 à G9) avec les espèces dont les numéros peuvent être remplacés par des symboles de leurs caractéristiques.

### 3.2 Analyse GROUPRES

Nous avons vu qu'il est relativement difficile d'interpréter simultanément les observations et les espèces en PRESABS. Afin de dégager le rapport entre les espèces et les groupes d'observations considérés

d'après les analyses PRESABS I et II, nous avons examiné le tableau de contingence  $k_{IK}$  où I est l'ensemble des espèces et K la partition considérée d'après les analyses PRESABS ; le calcul de  $k_{IK}$  se faisant comme on l'a expliqué pour l'analyse GROUPABOND .

Nous avons fait deux analyses :

GROUPRES I correspondant à PRESABS I

GROUPRES II correspondant à PRESABS II.

Dans la première analyse, nous avons pris en considération 13 groupes, dans la seconde seulement les 7 premiers ; les autres groupes étant mis en éléments supplémentaires.

Le dépouillement du tableau de contingence présenté ci-dessus ne fournit pas de résultats très intéressants du point de vue écologique . Ceci s'explique par le fait que dans chacun des groupes les espèces ayant les plus fortes présences sont souvent extrêmement banales et se situent près du centre de gravité. Ce ne sont pas elles qui permettent de caractériser les groupes. Par contre, d'autres, bien que présentes dans quelques échantillons seulement, ont une grande valeur écologique et apparaissent par contre clairement autour des points correspondant au groupe qu'elles caractérisent.

#### 4 La classification ascendante hiérarchique

Un tableau de correspondance  $k_{IJ}$  peut se prêter à de multiples classifications ascendantes hiérarchiques. On sait en effet que l'algorithme de CAH agrège d'abord les individus deux à deux, puis les paires obtenues entre elles et avec les individus isolés restants ; etc. jusqu'à constituer les classes majeures. Mais il existe de nombreux critères d'agrégation : i.e. de nombreuses règles pour décider de ce qu'on appellera les deux individus les plus proches... Ou : les deux classes les plus proches, etc. Ici pour la comparaison avec l'analyse factorielle qui étudie l'inertie des ensembles de points et utilise la métrique du  $\chi^2$ , on a appliqué la méthode d'agrégation suivant la variance avec pour métrique la distance du  $\chi^2$  : cette méthode est d'ailleurs, indépendamment des comparaisons avec l'analyse des correspondances, généralement la plus satisfaisante. Cependant on peut calculer les distances soit d'après le tableau de correspondance  $k_{IJ}$  lui-même, soit après analyse factorielle, en filtrant en quelque sorte les informations contenues dans  $k_{IJ}$  et ne retenant que celles correspondant aux premiers facteurs extraits. De plus, alors que l'analyse factorielle représente simultanément les deux ensembles I et J, la classification automatique porte soit sur l'un ou l'autre de ces ensembles. Ces préliminaires permettront de situer les quelques classifications que nous rapportons ci-dessous.

##### 4.1 Classification des espèces d'après les relevés en présence-absence:

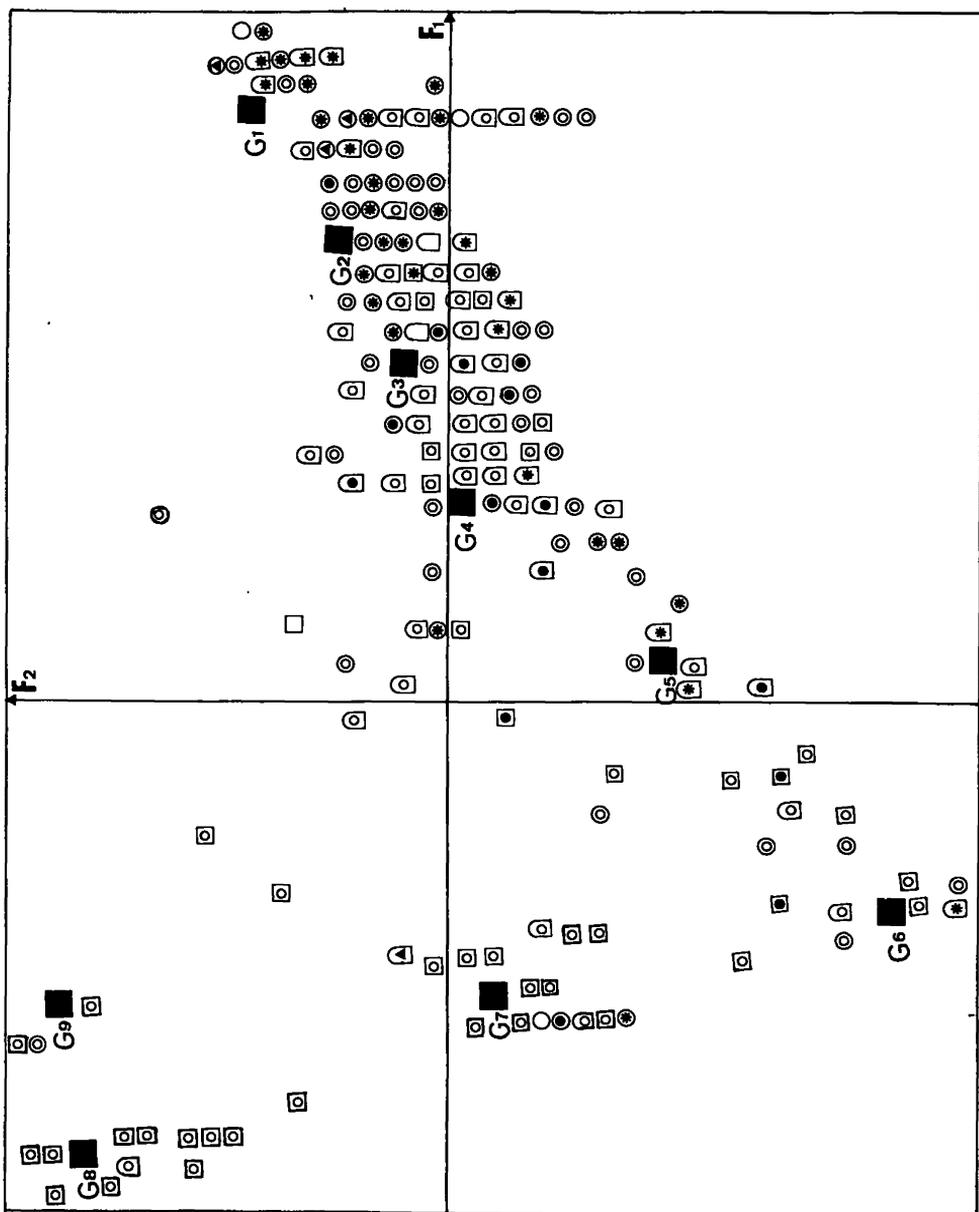
Au même tableau soumis à l'analyse factorielle, PRESABS I, nous avons appliqué la méthode d'agrégation suivant la variance avec pour métrique la distance du  $\chi^2$ . Nous essaierons de comparer les résultats obtenus par cette méthode avec ceux de PRESABS I et PRESABS II.

Bien que l'algorithme de classification procède par voie ascendante, nous allons examiner les classes en commençant par le sommet S de l'arbre.

S est la réunion des deux classes ou noeuds N1 et N2.

$$S = N1 \cup N2$$

Le noeud N1 est constitué par l'ensemble des espèces *Stephanodiscus*



Représentation des 9 groupes (■) de relevés déterminés par l'analyse GROUPABOND : chacun est entouré par les espèces caractéristiques. Chacune est définie par son mode de vie (□ planctonique (vivant dans des eaux profondes), ○ littorale, ◻ indifférente (planctonique ou littorale)) et sa salinité (● euryhalobe (sans exigence stricte), ○ oligohalobe (de salinité comprise entre 0.2 et 2‰), \* strictement mésosalobe (de salinité comprise entre 2‰ et 16‰), ● mésosalobe et oligohalobe, ▲ polyhalobe ou hyperhalobe (de salinité comprise entre 16‰ et 35‰ ou supérieure à 35‰).

*actrea v. spinulosa* (31), *S. carnonensis* (33), *S. transylvanicus* (32), *S. niagarae* (35), *S. cyclotalloides n. sp.* (36), *Rhopalodia rhopala* (37), *Melosira italica v. tenuissima* (16).

Ce noeud se distingue du reste des espèces et correspond au groupe 1 de l'analyse factorielle. Il est constitué d'espèces planctoniques caractéristiques du Plio-pleistocène (les plus anciennes dans notre étude). Il représente l'association à *Stephanodiscus* archaïque.

Quant à N2 = N3 ∪ N4, ce noeud contient trop d'espèces pour qu'on puisse voir ce qu'il représente globalement. Néanmoins, on voit que N3 est constitué d'espèces planctoniques (i.e. vivant en pleine eau) alors que N4 est constitué en majorité d'espèces littorales. Il n'est pas utile dans cet article de faire un inventaire détaillé.

La plupart des classes d'espèces caractérisent divers milieux déjà reconnus par l'analyse factorielle. Toutefois on remarque une classe qui n'a pas de signification stratigraphique contrairement aux autres groupes. On y rencontre des espèces présentes au cours des différents épisodes lacustres, indépendamment de leur âge : mais elles ont des caractères communs intéressants à souligner : il s'agit pour la plupart d'espèces littorales épiphytes qui vivaient donc vers les bordures marécageuses et riches en plantes supérieures des anciens lacs. On peut dégager de cette observation la remarque suivante :

Les épisodes lacustres sont bien caractérisés par leur flore d'eau profonde ; leur flore littorale est par contre beaucoup moins significative au point de vue stratigraphique.

La classification automatique hiérarchique a confirmé les résultats des analyses factorielles PRESABS I et PRESABS II ; en plus, elle a fait ressortir plus clairement et plus complètement les classes d'espèces.

#### 4.2 Classification des relevés d'après l'abondance des espèces

L'espace factoriel de dimension 7 déterminé par l'analyse factorielle ABONDANCE explique 31% de l'inertie totale du nuage.

Nous avons effectué une classification ascendante hiérarchique des observations sur les sept premiers facteurs déterminés par cette analyse. Les noeuds déterminés par cette classification correspondent généralement aux classes que nous avons définies d'après les quatre premiers facteurs de l'analyse factorielle ABONDANCE.

C'est ainsi qu'un noeud regroupe des échantillons des coupes de Boralé Est, de Gabilema et du golfe de Logghia, correspondant aux groupes G4, G5 et à une partie de G3.

Un autre rassemble la plus grande partie des échantillons relatifs au golfe de Logghia et du bassin de Boralé Ouest. Il correspond donc au groupe G3. Un autre rassemble les échantillons du paléolac du Dikhil et du paléolac d'Asa Gegalou. Il correspond strictement au groupe G1. Etc.

En somme nous constatons que la partition construite à vue d'après les quatre premiers facteurs de l'analyse ABONDANCE n'est pas significativement différente de la partition donnée par la classification automatique sur les sept premiers facteurs de la même analyse.

#### 4.3 Essai de classification des espèces d'après leur abondance dans les relevés

Nous avons effectué une classification des espèces d'après leurs sept premiers facteurs déterminés par l'analyse ABONDANCE. Les résultats obtenus sont beaucoup moins satisfaisants que ceux de la classification des relevés d'après le même tableau (§ 4.2). Les classes d'espèces sont

en général difficiles à interpréter comme telles, sauf pour certaines classes spécifiques (par exemple la classe des *Stephanodiscus*) qui correspondent bien aux noeuds des échantillons déterminés plus haut. Néanmoins nous retrouvons l'opposition entre les espèces planctoniques et les espèces littorales. Donc, si l'on se fonde sur le tableau des abondances, une classification satisfaisante des espèces est le mieux vue en décrivant celles-ci d'après des relevés déjà eux-mêmes classés au préalable en quelques groupes.

##### 5 Conclusions de l'étude

Les résultats présentés ici prouvent l'intérêt des analyses en présence-absence PRESABS et en abondance ABONDANCE dans l'étude des flores de diatomées fossiles.

Les analyses en présence-absence ont l'avantage de tenir compte de certaines espèces intéressantes bien que peu fréquentes. Elles permettent de caractériser chaque épisode lacustre par une liste d'espèces : GROU-PRES. Elles sont particulièrement intéressantes du point de vue biostratigraphique.

Les analyses en abondance relative ABONDANCE donnent une classification écologique des échantillons, particulièrement significative du point de vue évolution climatique. Cette évolution apparaît beaucoup plus clairement qu'avec les analyses en présence-absence.

Ces analyses ont été faites sur un ensemble de données déjà bien dépouillées par les méthodes classiques utilisées en géologie. Les résultats sont parfaitement concordants avec ce qui avait été déduit antérieurement après de mûres et longues réflexions.

L'application directe de ces méthodes aux données futures entraînera donc un gain de temps considérable pour les interprétations écologiques et stratigraphiques. L'ensemble de ces résultats apporte des faits d'une importance capitale en méthodologie pour l'étude des diatomées. En effet, contrairement à ce que pensent encore certains auteurs, il apparaît indispensable, pour aboutir à des informations précises, d'une part d'établir la liste complète des espèces contenues dans chaque échantillon (y compris les espèces rares), d'autre part de calculer les fréquences de chacune des espèces.

De plus, ces résultats permettent de préciser l'importance de certaines espèces qui peuvent maintenant être considérées comme de bons indicateurs et de classer de façon plus rigoureuse les échantillons.

Le but des travaux de Madame Gasse est d'établir une échelle biostratigraphique pour le continent africain. Une telle échelle chronologique sera fort utile pour l'élaboration des cartes géologiques par exemple. Elle ne pourra être construite de façon fiable et sans facteurs subjectifs qu'à partir de telles analyses. Enfin les changements floristiques au cours du temps sont en grande partie liés à l'évolution climatique. La comparaison, par ces méthodes d'analyse, des flores des diatomées d'un grand nombre de régions permettra de mieux comprendre cette évolution climatique qui est un problème d'actualité.

##### 5.1 Description du milieu ; suggestions d'analyses complémentaires

Les tableaux analysés dans cet article contiennent des informations *cénologiques* : i.e. concernant la présence commune dans un milieu (dont le sédiment est le témoin) de telles ou telles espèces en plus ou moins grande abondance. Il est possible d'introduire aussi des variables diverses décrivant ces milieux : ce qui est proprement l'objet de l'écologie. Voici d'abord comment on procéderait ici pour constituer le tableau des données. A propos du milieu, (du sédiment j) on peut considérer un certain nombre de questions  $q \in Q$ , la réponse à chacune de celles-ci étant

codée suivant un ensemble  $M_q$  de modalités (avec  $M = \cup \{M_q | q \in Q\}$ , ensemble de modalités de réponses à toutes ces questions) : par exemple on aura la question  $q_1$  : niveau stratigraphique ; ou  $q_2$  : taux de diatomées dans ce sédiment ;  $q_3$  : taux de carbone organique dans ce sédiment ;  $q_4$  : taux de carbonates etc. D'où un tableau  $k_{MJ}$  : avec  $k(m, j) = 1$  si la modalité  $m$  (e.g. taux de diatomées dans le sédiment inférieur à 20%) convient au sédiment  $j$ , et zéro sinon. Ce tableau peut être placé en dessous de celui des abondances  $k_{IJ}$  ; à chaque modalité  $m$ , correspond une ligne. Dans l'analyse du tableau  $k_{IJ}$ , il est possible de faire figurer les modalités  $m$  en éléments supplémentaires ; mais on peut aussi considérer l'ensemble  $I \cup M$  de toutes les lignes comme éléments principaux : dans ce cas il convient de pondérer le sous-tableau  $k_{MJ}$  afin que sa contribution ne disparaisse pas vis-à-vis de celle de  $k_{IJ}$  (dont chaque colonne, on l'a dit, a pour total 1000) ; le choix de coefficients de pondération se faisant e.g. suivant le programme de A.W. Hamrouni et Y. Grelet (Cahiers, Vol II n° 3). De même les informations relatives aux espèces pourraient fournir des colonnes supplémentaires : ici, il faut seulement prendre garde que, à la différence des colonnes  $j$ , les lignes  $i$  du tableau  $k_{IS}$  (des abondances relatives) sont de poids très inégal : si donc on adjoint e.g. un bloc de colonnes  $S$  correspondant à des informations logiques en 0,1 ( $k(i, s) = 1$  si l'espèce  $i$  possède la modalité  $s$ , et zéro sinon), il faudra encore multiplier dans le tableau  $k_{IS}$  chaque ligne afférente à une espèce  $i$  par le total  $k(i)$  des abondances relatives de cette espèce (total de la ligne  $i$  dans  $k_{IJ}$ ), afin que la modalité  $s$  ne soit pas attirée par les espèces rares qui la possèdent.

## BIBLIOGRAPHIE

- F. GASSE : L'évolution des lacs de l'Afar Central (Ethiopie et T.F.A.I.) du Plio-pleistocène à l'actuel.  
Reconstitution des paléomilieus lacustres à partir de l'étude des diatomées. Volume I et II.  
Thèse d'état, Paris VI ; 1975.
- F. TEKAIA : Analyse factorielle des correspondances. Application à une étude typologique des diatomées. Volume I.  
Thèse de 3° cycle ; université Pierre et Marie Curie (PARIS VI). 1978.