

M. BEN SALEM

**Méthodes multidimensionnelles en analyse
discriminante : division en espèces des
poissons du genre *Trachurus***

Les cahiers de l'analyse des données, tome 20, n° 2 (1995),
p. 201-224

http://www.numdam.org/item?id=CAD_1995__20_2_201_0

© Les cahiers de l'analyse des données, Dunod, 1995, 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/>

MÉTHODES MULTIDIMENSIONNELLES EN ANALYSE DISCRIMINANTE: DIVISION EN ESPÈCES DES POISSONS DU GENRE *Trachurus*

[DISCR. POISSONS]

M. Ben SALEM*

0 Origine et objet du présent travail

Nous avons consacré notre thèse de doctorat à une "Révision du genre *Trachurus*". Au profane, nous rappellerons d'abord que, sous le nom de Chinchard ou de Saurel, les poissons de ce genre sont couramment vendus et consommés en France.

Puis, sans reprendre l'historique détaillé donné dans la thèse, nous précisons que plus de cent espèces et sous-espèces ont pu être créées, dont la plupart sont considérées aujourd'hui comme synonymes entre elles, ou de validité douteuse.

Dans le présent travail, on distinguera 14 espèces. Ce schéma taxinomique est généralement admis: à titre de comparaison, disons que le "Dictionnaire", publié à Moscou sous la direction de l'Académicien V. E. Sokolov, (1989), reconnaît 15 espèces ou sous-espèces; système qui coïncide avec le nôtre; à ceci près que l'espèce *T. picturatus* y est partagée en deux sous-espèces, *T. p. picturatus* (Bowdich, 1825) et *T. p. aleevi* Rytov et Razumovskaya; et que *T. mediterraneus* et *T. ponticus* y sont données pour être deux sous-espèces de l'espèce *T. mediterraneus*.

D'après les seuls noms, on voit que le genre *Trachurus* est répandu dans toutes les mers du globe. Grâce au patronage de Madame M.-L. BAUCHOT, Sous-Directeur au Muséum national d'Histoire naturelle de Paris, nous avons eu accès au matériel de base déposé dans les institutions du Monde entier. Nous avons examiné avec grand soin et mesuré 1193 spécimens, représentant 14 espèces suivant le bilan donné plus loin.

On peut présumer qu'à l'exception de O, toutes les espèces sont suffisamment représentées. D'autre part, J exceptée, chaque espèce est connue par au moins deux populations distinctes; le nombre pouvant aller jusqu'à 7 pour T. Nous ne donnerons pas ici le détail des origines géographiques, mais y ferons allusion occasionnellement dans la suite.

(*) Maître de conférences; Département des Sciences Biologiques;
Faculté des Sciences de Tunis.

Genre *Trachurus*, Rafinesque, 1810

- T : *T. trachurus trachurus* (Linnæus, 1758)
 P : *T. picturatus picturatus* (Bowdich, 1825)
 D : *T. declivis* (Jenyns, 1841)
 N : *T. novæ-zelandiæ* Richardson, 1842
 J : *T. japonicus* (Temminck et Schlegel, 1844)
 S : *T. symmetricus symmetricus* (Ayres, 1956)
 C : *T. trachurus capensis* Castelnau, 1861
 M : *T. mediterraneus mediterraneus* (Steindachner, 1868)
 U : *T. symmetricus murphyi* Nichols, 1920
 L : *T. lathamii* Nichols, 1920
 R : *T. treca* Cadenat, 1949
 X : *T. mediterraneus ponticus* (Aleev, 1956)
 I : *T. indicus* Nocrassov, 1966
 O : *T. delagoa* Nocrassov, 1970

ci-dessus: tableau des 14 espèces distinguées dans notre étude, avec les lettres qui servent à les désigner ;

ci-dessous: bilan des individus mesurés, par espèce .

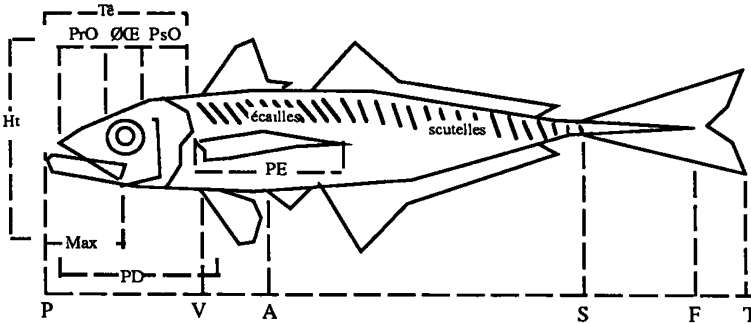
sp	C	D	I	J	L	M	N	O	P	R	S	T	U	X
n	51	42	65	91	99	102	201	18	65	112	156	96	46	49

Initialement, la distinction des espèces s'est fondée sur les caractères morphologiques directement observables. On dispose aujourd'hui de critères génético-biochimiques, sur lesquels la thèse s'étend longuement. L'objet du présent article est d'apprécier jusqu'à quel point une taxinomie précise peut être fondée sur le traitement statistique des données morphologiques macroscopiques.

De façon précise, les données morphologiques constituent deux groupes distincts: les variables morphométriques, ou mesures de longueurs prises sur le corps du poisson entre des repères déterminés de manière non équivoque; et variables méristiques, résultant de comptages généralement effectués sous la loupe binoculaire: nombre de rayons dans une nageoire, nombre d'écaillés le long d'une ligne etc..

Soumettre des mesures de longueurs à l'analyse des correspondances remonte aux recherches de L. BELLIER sur la taxinomie des micromammifères (1973): les règles de codage et d'interprétation sont ici bien établies. L'analyse des variables méristiques offre, *a priori*, des difficultés de principe; et nous ne croyons pas qu'elle soit d'une pratique classique. C'est pourtant seulement en conjuguant l'analyse des deux types de données que nous avons pu obtenir, sur un échantillon d'épreuve, le taux de discrimination de 93%; taux que nous croyons très satisfaisant, pour un ensemble de 14 espèces d'un même genre.

Dans la suite, nous considérerons d'abord séparément les deux types de variables (§§1 & 2); avant de les prendre en compte simultanément; suivant une méthode qui, elle non plus, n'est pas classique (§3).



1 Analyse des variables morphométriques

1.1 Description des variables

$$J1 = \{Tê \ Oe \ PrO \ PsO \ Max \ PD \ PV \ PE \ HE1 \ HE2 \ Ht \ VA \ AS \ SF \ FT\}$$

On a retenu un ensemble J1 de 15 mesures de longueurs que nous considérerons par groupes successifs, d'après le schéma publié ici.

{PV VA AS SF FT}: Suivant les principes de l'analyse des correspondances, il est souhaitable que les variables offrent un bilan exhaustif et homogène d'un aspect des objets décrits. On applique ce principe à des mesures de longueurs en divisant une dimension de l'objet en segments successifs délimités par des repères. Ici, suivant la longueur du corps, les repères sont: l'extrémité antérieure, la naissance de la nageoire ventrale, l'anus, l'origine, la fourche et l'extrémité de la queue.

{Tê PrO Oe PsO Max PD}: Pour la description de la tête, qui comporte des mesures longitudinales prises à des niveaux transverses différents, on n'a pas appliqué strictement le principe de la décomposition en segments consécutifs; ni, encore moins, tenté d'intégrer tous les repères de la tête à la segmentation longitudinale du corps. Prenons pour exemple le maxillaire, Max: le poisson étant prognathe, la lèvre inférieure précède nettement la lèvre supérieure: mais il nous a paru peu sûr de mesurer ce décalage pour l'introduire dans une analyse de correspondance, parmi des mesures dont la variation relative est, à la fois, moindre et mieux déterminée. En revanche, {PrO Oe PsO}, distance préorbitaire, diamètre horizontal de l'œil et distance postorbitaire constituent un bilan longitudinal de la tête; simultanément {Tê PD} offrent deux mesures du bloc antérieur du corps.

{PE HE1 HE2 Ht}: La longueur de la nageoire pectorale, PE, est comptée à part. Transversalement, on a trois mesures: Ht, Hauteur totale; HE1, HE2, Hauteur des écailles et des scutelles. Il faut savoir que le poisson a une ligne latérale de très fortes écailles, marquées sur le schéma par de simples traits; et que, sur la partie postérieure du corps, ces écailles ont une pointe centrale qui leur donne la forme d'un bouclier; en latin: scutum, d'où le nom de scutelle, petite plaque en bouclier.

1.2 Enchaînement des analyses de correspondance

Dans toutes les analyses, l'ensemble des colonnes principales est l'ensemble J1 des 15 mesures de longueur décrites au §1.1; il n'y a pas de colonne supplémentaire. Comme lignes, principales ou supplémentaires, on a, d'une part, des individus, et, d'autre part, des espèces; celles-ci définies, chacune, par cumul sur tout ou partie de l'ensemble de ses individus.

Dans une première analyse, le tableau principal, $I \times J1$, croise, avec J1, l'ensemble I des 1193 individus; en supplément, est l'ensemble Isp des espèces, dont chacune est définie par cumul de tous ses individus (i.e., cf. *supra*, §0: 51 pour spC, 42 pour spD, etc.).

L'analyse de $I \times J1$ ayant montré (cf. §1.3) une assez bonne discrimination entre les espèces, on poursuit en distinguant, au sein des individus, un échantillon de base Ib et un échantillon d'épreuve, Ie.

Une fois pour toutes, on réserve pour Ie un tiers des individus: ceux (au nombre de 398) qui ont un rang, i, dont le reste de la division par 3 est 1: $i \bmod 3 = 1$. Corrélativement, l'ensemble de base comprend les 795 individus définis par: $i \bmod 3 \neq 1$. L'ensemble, Isb, des lignes afférentes aux espèces est défini par cumul sur les individus de Ib rentrant dans chaque espèce: e.g., 34 pour spC, 28 pour spD; etc.

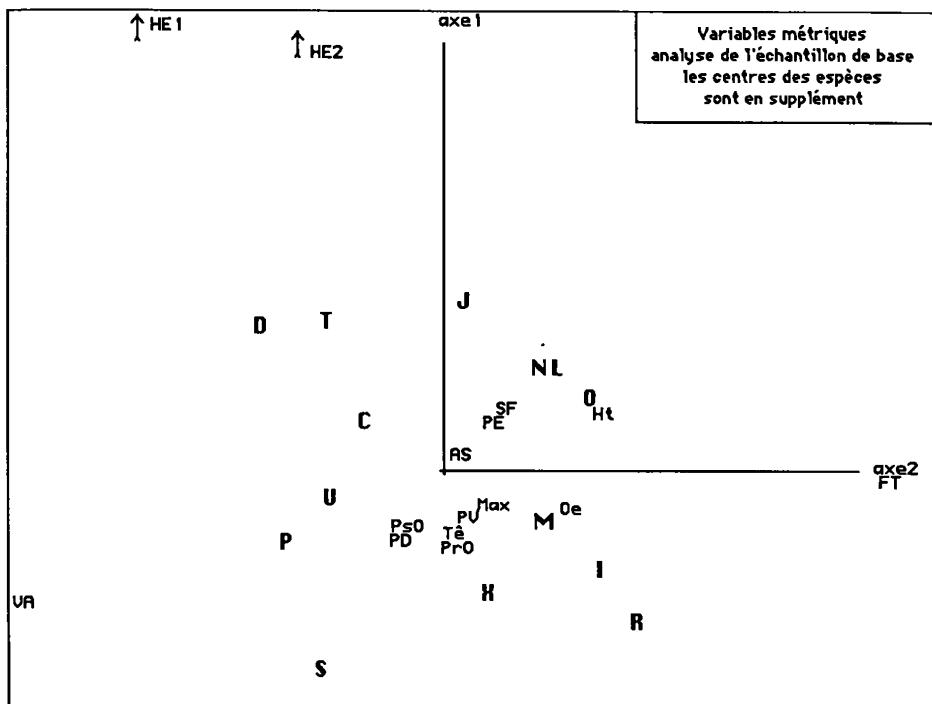
On effectue deux analyses. L'une où $Ib \times J1$ est en principal; $Ie \times J1$ étant en supplément, ainsi que $Isp \times J1$. L'autre, où $Isb \times J1$ est en principal; $Ib \times J1$ et $Ie \times J1$ étant en supplément.

Il y a, entre les résultats de toutes ces analyses, de telles similitudes, qu'on se bornera à présenter des graphiques issus de l'analyse de $Ib \times J1$; en décrivant les variantes afférentes aux autres analyses.

Le genre <i>Trachurus</i> : M. Ben Salem : analyse $I \times J1$												
rang :	1	2	3	4	5	6	7	8	9	10...	14	
lambda:	34	18	14	7	6	4	2	2	1	1...	0	e-4
taux :	3705	2006	1543	720	644	414	269	223	119	101...	43	e-4
cumul :	3705	5711	7254	7974	8618	9032	9301	9524	9642	9744	10000	e-4
Le genre <i>Trachurus</i> : M. Ben Salem: $Ib \times J1$												
trace : 9.210e-3												
rang :	1	2	3	4	5	6	7	8	9	10 ...	14	
lambda:	34	19	14	7	6	4	2	2	1	1 ...	0	e-4
taux :	3667	2047	1560	737	627	421	258	211	116	101 ...	40	e-4
cumul :	3667	5714	7274	8012	8639	9060	9319	9529	9645	9746	10000	e-4
Le genre <i>Trachurus</i> : M. Ben Salem: $Isb \times J1$												
trace : 5.073e-3												
rang :	1	2	3	4	5	6	7	8	9	10...		
lambda:	31	13	3	2	1	1	0	0	0	0	0	e-4
taux :	6078	2539	562	431	209	124	24	22	4	3	3	e-4
cumul :	6078	8617	9179	9610	9819	9943	9967	9989	9993	9996	9996	e-4

1.3 Valeurs propres et facteurs

Quant aux valeurs propres, il n'y a pas grande différence entre l'analyse de $I \times J1$ et celle de $Ib \times J1$; i.e., selon que l'ensemble principal comprend tous les individus ou seulement (2/3) de ceux-ci. De même, d'une analyse à l'autre, le nuage, J1, des variables change peu; non plus que celui des espèces, mises

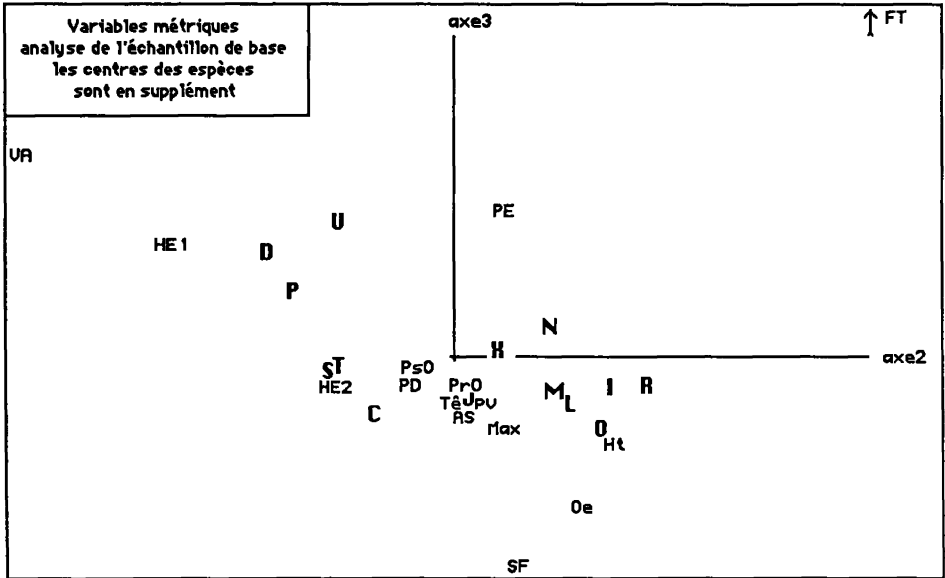


en supplément. Dans la suite du §1.3, nous nous bornerons donc à comparer les analyses $Ib \times J1$ et $Isb \times J1$.

Sur la suite des valeurs propres issues de $Isb \times J1$, il apparaît que le nuage des 14 espèces n'a que 6 dimensions principales: au-delà du rang 6, le taux d'inertie est au-dessous de $(1/400)$. Au contraire, pour le nuage des individus, le taux afférent au dernier facteur (de rang 14) est: $4\% \approx (1/250)$. La diversité des individus met en jeu un plus grand nombre de dimensions que celle des espèces. Le tableau des valeurs propres atteste que, sur les axes 1 et 2, l'inertie est du même ordre de grandeur pour le nuage Ib des individus ou pour celui, Isb , des espèces; corrélativement, la disposition de l'ensemble $J1$ des variables diffère peu.

L'axe 1 est créé par les deux variables HE1 HE2, hauteurs des écailles et des scutelles, très écartées sur ($F1 > 0$); d'une part, du fait de leur faible masse; et d'autre part, parce que la vie de l'animal est compatible avec d'importantes variations relatives de ces hauteurs. Ce déséquilibre entre petites et grandes dimensions, ne nuit pas à l'analyse, puisque, pour l'ensemble des individus ou des centres d'espèces, la distribution est régulière et quasi symétrique.

Sur l'axe 2, la principale contribution vient de VA (mesure prise au niveau du ventre), associé à HE1, ($F2 < 0$), opposé à FT (queue), suivi de Ht (hauteur totale), ($F2 > 0$).



À partir de l'axe 3 les facteurs issus de $I_b \times J_1$ et $I_{sb} \times J_1$ diffèrent, à la fois quant à l'ordre de grandeur de l'inertie et à l'interprétation.

Sur l'axe 3 issu de $I_b \times J_1$, FT (associé à VA et PE), s'oppose à SF: SF et FT sont les deux dimensions longitudinales qu'on a distinguées dans la queue; il est remarquable que FT, pointe au delà de la fourche, soit associée à PE, nageoire pectorale: les deux parties se ressemblent quant à la texture, et il n'est pas exclu que la conservation du spécimen les affecte dans le même sens. Quoiqu'il en soit de cette hypothèse (revue au §3.3), alors que les axes 1 et 2 ne sont pas modifiés si, au lieu des individus, les espèces sont en principal; l'axe 3 disparaît dans l'analyse de $(I_{sb} \times J_1)$. En effet, dans l'analyse de $I_b \times J_1$, l'inertie sur l'axe 3 du nuage (I_{sb} , mis en supplément,) des centres des espèces est faible: elle ne vaut que 17% de celle du nuage I_b des individus de base.

L'axe 4 issu de $I_b \times J_1$ est dominé par la contribution de PE, $CTR_4(PE)=644\%$; qui, sur cet axe, s'oppose à FT. Quand les espèces sont seules en principal, $(I_{sb} \times J_1)$, on trouve un axe 3 créé par PE.

Au §1.4.3, nous reviendrons sur la faible contribution qu'apporte, à la discrimination des espèces, l'axe 3 issu de $I_b \times J_1$.

1.4 Discrimination des espèces

1.4.1 Discrimination sans échantillon d'épreuve

À l'analyse de $I \times J_1$, est adjoint, en supplément, l'ensemble I_{sp} des espèces. On affecte chaque individu, i , à l'espèce du centre de laquelle il est le plus proche. Le bilan des affectations est donné par une matrice de confusion, où $k(>x, y) =$ nombre de poissons de l'espèce y affectés à x : e.g, on lit, à

I → Isp : confusions entre espèces du genre *Trachurus*

14	C	D	I	J	L	M	N	R	O	P	S	T	U	X
>C	37	1	0	1	3	2	7	0	0	0	0	3	12	0
>D	2	25	0	1	0	0	0	0	0	0	0	8	0	0
>I	0	0	46	0	2	1	1	21	0	0	0	0	0	1
>J	0	1	0	66	19	0	18	0	0	0	0	1	0	0
>L	2	0	0	5	32	0	38	0	6	0	0	0	0	0
>M	0	0	5	0	2	83	5	0	1	4	3	0	0	7
>N	1	1	0	14	22	1	117	0	1	0	0	2	0	0
>R	0	0	8	0	0	1	0	90	0	0	0	0	0	0
>O	0	1	0	1	9	0	4	0	10	0	0	0	0	0
>P	0	0	2	0	0	2	0	0	0	59	1	0	4	1
>S	0	0	0	0	0	2	0	1	0	1	144	0	0	2
>T	6	13	0	3	9	0	11	0	0	0	0	82	0	0
>U	3	0	0	0	0	0	0	0	0	1	0	0	30	0
>X	0	0	4	0	1	10	0	0	0	0	8	0	0	38

l'intersection de la ligne >L et de la colonne N que 38 des poissons de l'espèce N ont été affectés, à tort, à l'espèce L.

Au total, plus de 69% des affectations sont exactes. Ainsi qu'on l'a dit au §1.2, c'est après le relatif succès de cet essai de discrimination qu'on a décidé de poursuivre les analyses.

1.4.2 Discrimination avec échantillon d'épreuve

L'analyse discriminante montre toute son utilité si elle permet de reconnaître l'espèce d'un nouvel individu d'après l'ensemble, J1, de ses mesures. Afin de simuler cela, on a réservé un échantillon d'épreuve, Ie.

Si on affecte chaque individu d'épreuve ie, de Ie, à l'espèce du centre de laquelle il est le plus proche, les taux de confusions sont du même ordre que ceux obtenus au §1.4.1, par la même méthode; et cela, que l'ensemble Isb des espèces figure, dans l'analyse, en principal ou en supplémentaire.

Les résultats sont meilleurs si, pour conjecturer l'espèce d'un poisson de Ie, on passe par l'intermédiaire de l'échantillon de base, Ib. De façon précise, à un individu ie, on attribue l'espèce de l'individu ib (de Ib) duquel il est le plus proche. Ainsi, le taux des affectations exactes dépasse 78%. On voit toutefois, sur la matrice de confusion, que l'espèce O (*T. delagoa*) est mal reconnue: cette espèce est peu représentée dans le corpus; pourtant, au §3, il apparaîtra qu'on peut la discriminer en considérant l'ensemble des variables.

affectation dans l'espace de dimension 14 : Ie → Ib

14	C	D	I	J	L	M	N	R	O	P	S	T	U	X
C	15	0	0	0	0	0	1	0	0	0	0	1	0	0
D	0	8	1	0	0	0	1	0	0	0	0	4	0	0
I	0	0	15	0	0	0	2	0	5	0	0	0	0	0
J	0	0	0	24	1	0	4	0	0	0	0	1	0	0
L	1	0	0	0	19	1	11	0	1	0	0	0	0	0
M	0	0	0	0	0	29	0	0	0	1	1	0	0	3
N	1	1	0	5	6	0	50	0	2	0	0	2	0	0
R	0	0	5	0	0	0	0	32	0	0	0	0	0	0
O	0	0	0	0	3	0	2	0	2	0	0	0	0	0
P	0	0	0	0	0	0	0	0	0	21	0	0	0	0
S	0	0	0	0	0	0	0	0	0	0	52	0	0	0
T	3	4	1	0	3	0	2	0	0	0	0	19	0	0
U	1	0	0	0	0	0	0	0	0	0	0	0	15	0
X	0	0	1	0	0	1	0	1	0	0	2	0	0	11

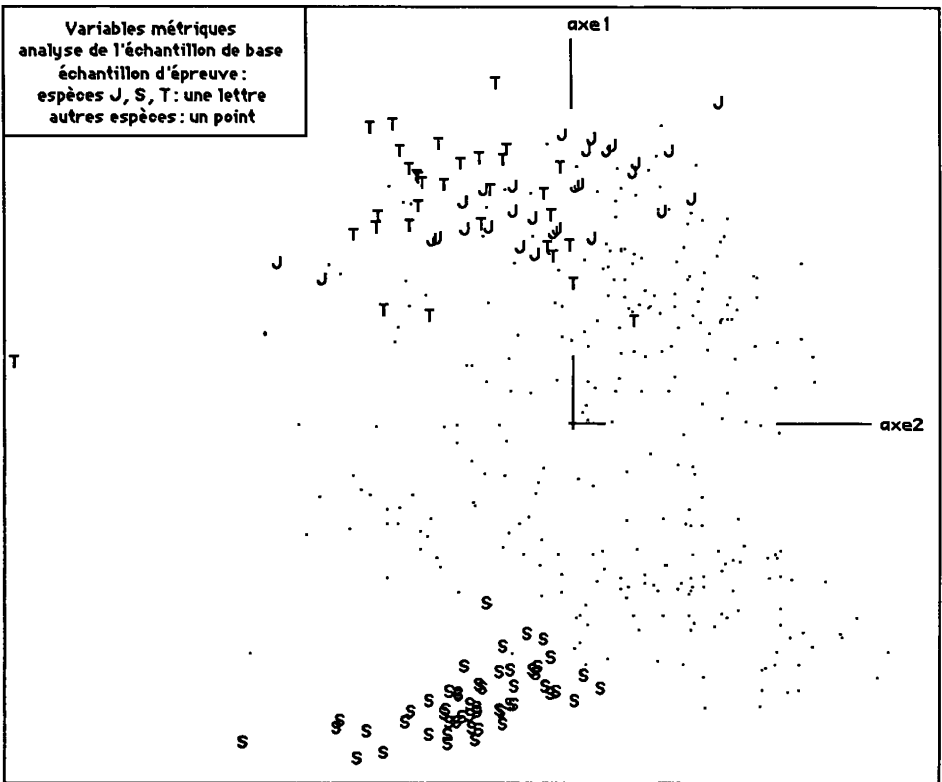
Le procédé admet des variantes, selon qu'on se place dans l'espace rapporté aux axes issus de l'analyse de $I_b \times J_1$ (individus en principal), ou de $I_{sb} \times J_1$ (espèces en principal); et selon le nombre de facteurs pris en compte pour calculer les distances et rechercher le plus proche voisin.

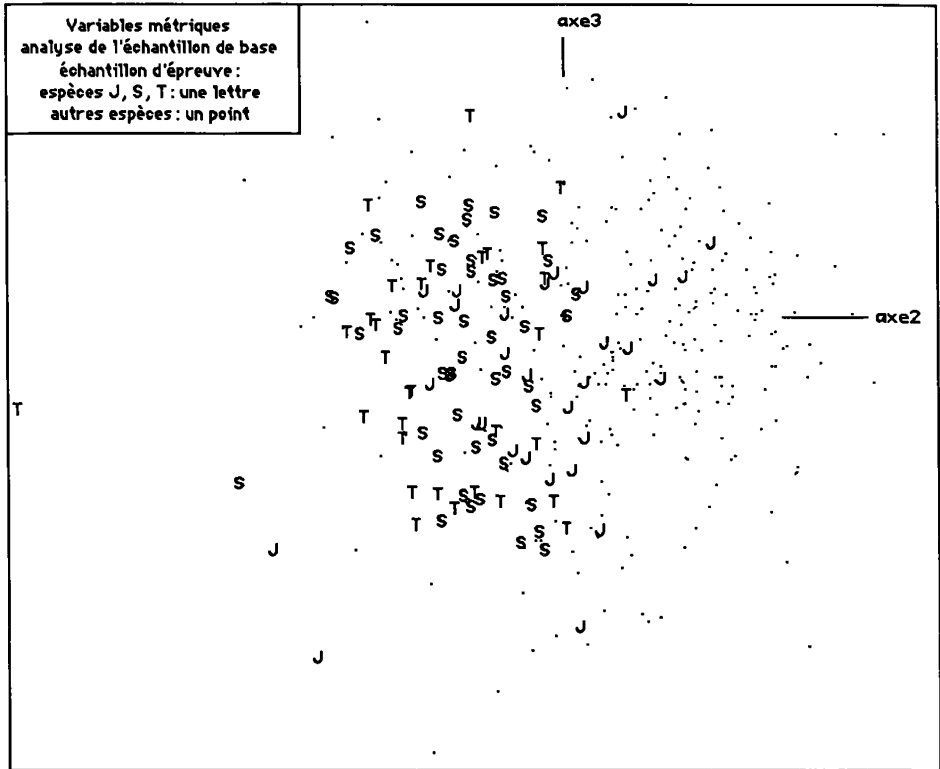
affectation dans l'espace de dimension 6 : $I_e \rightarrow I_b$

14	C	D	I	J	L	M	N	R	O	P	S	T	U	X
C>	8	0	0	0	1	1	3	0	0	0	0	2	2	0
D>	0	6	0	0	0	0	2	0	1	0	0	5	0	0

Prendre les 14 facteurs issus de l'analyse de $I_b \times J_1$ équivaut à considérer exactement la distance distributionnelle; avec les 13 facteurs issus de $I_{sb} \times J_1$, on se place dans un sous-espace de codimension 1 de l'espace des profils sur J_1 , et la discrimination est peu altérée. Mais si l'on revient à l'interprétation, $I_{sb} \times J_1$ ne donne que 6 facteurs significatifs (quant à la dispersion des espèces); et avec 6 facteurs, le taux de reconnaissance est faible: comme on le voit sur le bilan d'affectation des espèces C (*T. capensis*) et D (*T. declivis*).

1.4.3 Distribution des espèces dans l'espace engendré par les axes de rang 1 à 3





Afin d'apprécier la distribution des espèces dans l'espace engendré par les premiers axes factoriels, on considère le nuage Ie dans les plans (2×1) et (2×3). Sur les graphiques publiés, l'échelle est la même sur tous les axes; les poissons des espèces {T, J, S}, qui s'opposent sur l'axe 1, sont marqués par leur lettre; et les autres par un point.

Dans le plan (2×1), du côté ($F1 > 0$), les aires des espèces T et J empiètent; et la séparation n'est pas nette entre {J, T} et les autres espèces. Du côté ($F1 < 0$), la séparation de l'espèce S est à peu près parfaite.

Dans le plan (2×3), il apparaît qu'en projection sur l'axe 3, chacune des espèces {T J S} a une dispersion quasi égale à celle de l'ensemble du nuage Ie.

Ceci nous rappelle l'hypothèse, suggérée au §1.3, que l'opposition exprimée par $F3$ contribue peu à la définition des formes spécifiques. Plus précisément, reprenons la méthode du §1.4.2 pour affecter Ie à Ib; mais en éliminant le facteur $F3$ du calcul de la distance. Ainsi, le nombre total des erreurs augmente un peu; mais la discrimination de certaines espèces s'améliore.

Ce qui confirme que $F3$ n'a qu'un rôle mineur.

affectation dans l'espace de dimension 13 (sans F3): Ie → Ib

14	C	D	I	J	L	M	N	R	O	P	S	T	U	X
C>	9	0	0	1	0	0	1	0	0	0	0	2	1	0
D>	0	8	1	0	0	0	1	0	0	0	0	3	0	0
I>	0	0	14	0	0	2	0	6	0	0	0	0	0	0
J>	0	0	0	27	1	0	1	0	0	0	0	1	0	0
L>	1	0	0	2	17	1	10	0	1	0	0	1	0	0
M>	0	0	0	0	0	25	0	1	0	2	1	0	0	5
N>	1	0	0	7	7	0	50	0	1	0	0	1	0	0
R>	0	0	2	0	0	0	0	35	0	0	0	0	0	0
O>	0	0	0	0	2	0	2	0	3	0	0	0	0	0
P>	0	0	0	0	0	2	0	0	0	18	1	0	0	0
S>	0	0	0	0	0	1	0	0	0	0	51	0	0	0
T>	2	2	1	0	2	0	0	0	0	0	0	25	0	0
U>	0	0	0	0	0	1	0	0	0	0	0	0	15	0
X>	0	0	1	0	0	2	0	0	0	0	3	0	0	10

Ci-dessus : matrice de confusion entre espèces décrites par les variables morphométriques

2 Analyse des variables méristiques

2.1 Description des variables

$$J2 = \{SC \ LAC \ LA \ LP \ BRS \ BRI \ P \ D \ A\}$$

Sans prétendre reproduire toutes les précisions que donne la thèse, nous décrivons l'ensemble J2 des 9 variables retenues, en renvoyant, s'il y a lieu, au schéma du §1.1.

SC : rang du rayon de la seconde nageoire dorsale sous lequel apparaît la première scutelle de la ligne latérale principale.

LAC : rang du rayon de la seconde nageoire dorsale sous lequel se termine la ligne latérale accessoire, ou commissure supra-temporale.

LA : nombre d'écaillés de la ligne latérale principale; ces écaillés (cf. schéma) se trouvent uniquement à la partie antérieure de la ligne.

LP : nombre de scutelles de la ligne latérale principale; ces scutelles (cf. schéma) se trouvent uniquement à la partie postérieure de la ligne.

BRS : nombre des branchiospines sur la partie supérieure du premier axe branchial gauche.

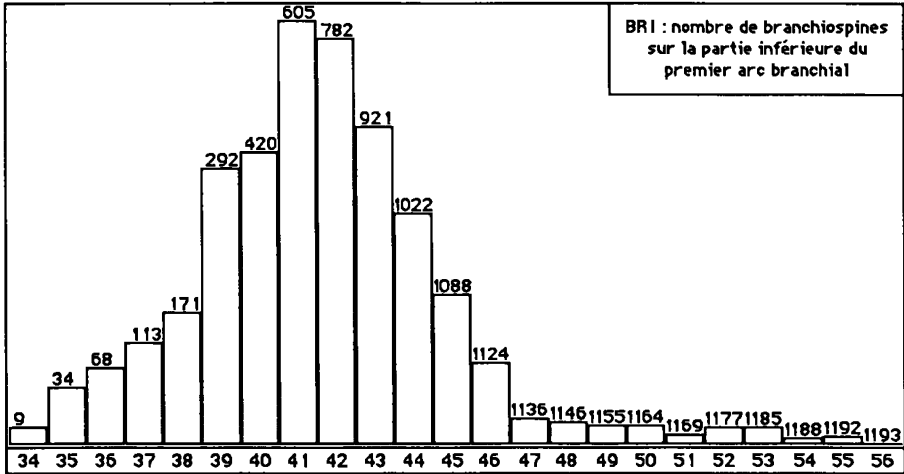
BRI : nombre des branchiospines sur la partie inférieure du premier axe branchial gauche.

P : nombre de rayons mous à la nageoire pectorale.

D : nombre de rayons mous à la nageoire dorsale.

A : nombre de rayons mous à la nageoire anale.

Au sein de l'ensemble J1 des variables morphométriques, considéré au §1, la viabilité même du poisson impose des relations assez étroites (la hauteur des écaillés et scutelles, faisant seule exception; son rapport à la longueur totale du corps étant très étalé sur l'ensemble des espèces du genre *Trachurus*; cf. §1.3). On ne peut être certain, *a priori*, qu'il en soit de même, pour des variables résultant de dénombrements. C'est pourquoi, avant toute analyse, on doit soumettre au lecteur les histogrammes de certaines variables de J2, avec leur distribution sur l'ensemble des espèces.



Aucune autre variable méristique ne présente une telle bimodalité. Le cas de BRI, unimodale avec une distribution s'étendant à droite, est typique.

P \ SP	C	D	I	J	L	M	N	O	P	R	S	T	U	X
18	3	2	4	1	18	14	15	3		2	1	18		1
19	7	16	17	41	62	42	113	11	1	33	13	53	3	15
20	28	22	36	45	18	40	66	3	11	65	69	22	12	26
21	13	2	8	4	1	6	7	1	32	12	67	3	28	7
22									19		6		3	
23									2					

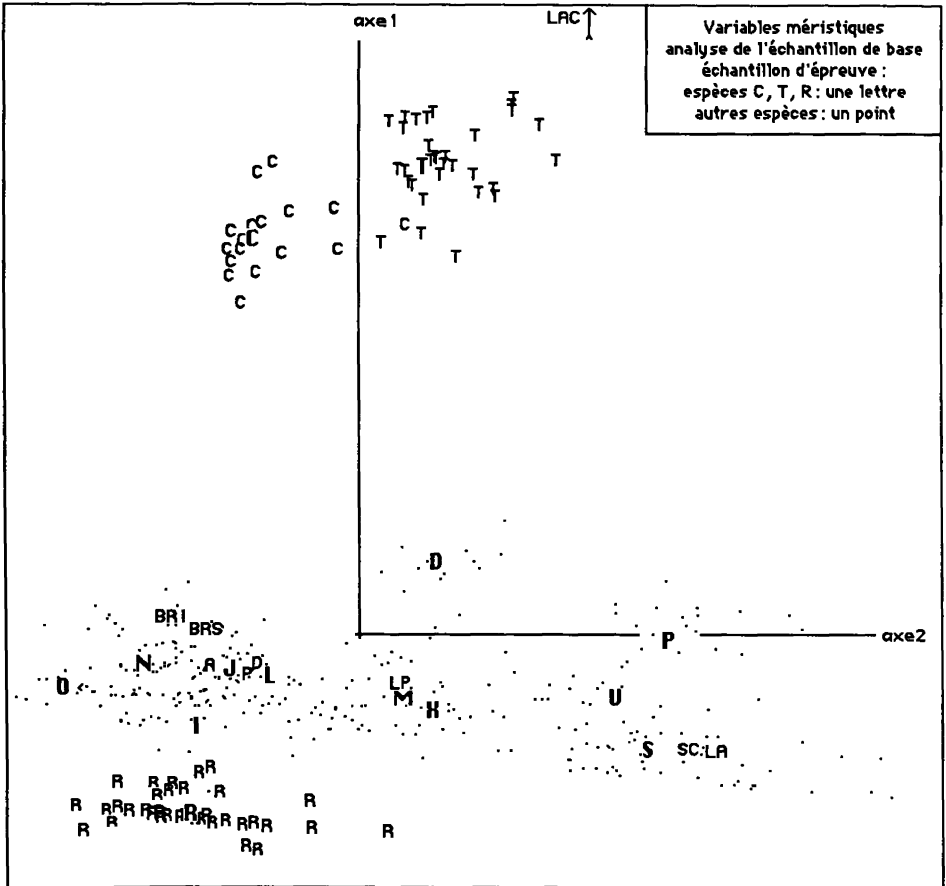
Corrélativement, il n'y a pas de partage net des espèces: mais le tableau ci-joint montre que la variable P, nombre de rayons mous à la nageoire pectorale, (qui ne prend que les six valeurs de 18 à 23,) contribue notablement à caractériser l'espèce *T. picturatus* (P).

2.2 Analyse de correspondance

L'hétérogénéité des données, la présence de seuils nets dans certaines distributions, suggéreraient de découper les variables en modalités; de préférence par codage barycentrique afin de tirer le meilleur parti des informations.

Mais comme, d'autre part, plusieurs dénombrements étant faits suivant des lignes, les rapports mutuels de ceux-ci, au sein d'un profil, peuvent avoir valeur taxinomique, on pouvait tenter d'analyser les variables méristiques sans codage préalable. C'est cette analyse que nous présentons ici; et, puisqu'elle a bien servi à la discrimination des espèces, nous n'en prendrons pas d'autre.

De façon précise, on analyse le tableau I x J₂, 1193 x 9; avec pour ensemble principal d'individus le même échantillon de base, Ib qu'au §1; le reste constituant l'échantillon d'épreuve, Ie. Et à l'analyse, on adjoint, en lignes supplémentaires, l'ensemble Isb des cumuls par espèce effectués sur Ib.



Le genre *Trachurus*: M. Ben Salem : variables méristiques.

trace :	2.819e-2								
rang :	1	2	3	4	5	6	7	8	
lambda :	217	44	9	6	2	2	1	1	e-4
taux :	7681	1570	303	227	84	63	48	24	e-4
cumul :	7681	9251	9554	9781	9865	9928	9976	10000	e-4

Les facteurs 1 et 2, nettement séparés des suivants quant au taux d'inertie, sont créés, respectivement, par les variables LAC et LA, dont on a remarqué la bimodalité. Vers ($F1 > 0$), les espèces {C, T}, sans mélange avec les individus des autres espèces, sont associées à LAC; et l'on a, suivant l'axe 2, une discrimination imparfaite entre C et T. Vers ($F1 < 0$), les individus de l'espèce R, marqués de cette lettre, sont presque parfaitement séparés de ceux des autres espèces, marqués d'un point. Vers ($F2 > 0$), se détachent les trois espèces {S, U, P}; suivies de {M, X}, avec D vers ($F1 > 0$). Ces résultats de l'analyse factorielle concordent bien avec l'examen préalable des données.

affectation dans l'espace de dimension 8 : Ie → Ib

	C	D	I	J	L	M	N	R	O	P	S	T	U	X
14														
C>	16	0	0	0	0	0	0	0	0	0	0	1	0	0
D>	0	14	0	0	0	0	0	0	0	0	0	0	0	0
I>	0	0	16	2	0	1	2	1	0	0	0	0	0	0
J>	0	0	0	22	3	0	5	0	0	0	0	0	0	0
L>	0	0	0	6	20	4	3	0	0	0	0	0	0	0
M>	0	0	0	1	1	23	2	0	0	0	1	0	0	6
N>	0	0	1	5	1	2	57	0	1	0	0	0	0	0
R>	0	0	1	0	0	0	0	36	0	0	0	0	0	0
O>	0	0	0	1	0	0	2	0	4	0	0	0	0	0
P>	0	0	0	0	0	0	0	0	0	20	1	0	0	0
S>	0	0	0	0	0	0	0	0	0	3	48	0	1	0
T>	0	0	0	0	0	0	0	0	0	0	0	32	0	0
U>	0	0	0	0	0	1	0	0	0	0	3	0	11	1
X>	0	0	0	0	0	9	0	0	0	0	1	0	0	6

2.3 Discrimination avec échantillon d'épreuve

Pour la discrimination, on se place dans l'espace des profils, rapporté aux 8 axes extraits de l'analyse du tableau Ib x J2. Il faut prendre garde que le choix de Ib n'a, sur la distance, qu'une influence restreinte, par le biais des pondérations des variables; et l'analyse factorielle ne sert qu'à donner une forme simple au calcul de la distance du χ^2 .

En attribuant à chaque individu, ie, de l'échantillon d'épreuve, Ie, l'espèce de celui des individus de base dont il est le plus proche, on obtient une matrice de confusion dont le total des termes extradiagonaux est de 73; soit un taux d'erreur de (73/398), inférieur à 18,4%. En particulier, la discrimination de D est parfaite; il n'y a, pour C et T, qu'une seule confusion, un individu de C étant affecté à T. Mais X est imparfaitement séparé de M; et, comme au §1.4.2, la discrimination de O est médiocre.

3 Analyse conjointe des variables morphométriques et méristiques

affectation de Ie à Ib, dans l'espace de dimension (14+8)

	C	D	I	J	L	M	N	R	O	P	S	T	U	X
14														
C>	15	0	0	0	0	0	0	0	0	0	0	2	0	0
D>	0	13	1	0	0	0	0	0	0	0	0	0	0	0
I>	0	0	20	0	0	1	0	1	0	0	0	0	0	0
J>	0	0	0	28	1	0	1	0	0	0	0	0	0	0
L>	0	0	0	1	26	0	6	0	0	0	0	0	0	0
M>	0	0	0	0	0	30	2	0	0	0	1	0	0	1
N>	0	0	0	0	2	0	65	0	0	0	0	0	0	0
R>	0	0	0	0	0	0	0	37	0	0	0	0	0	0
O>	0	0	0	0	0	0	1	0	6	0	0	0	0	0
P>	0	0	0	0	0	0	0	0	0	21	0	0	0	0
S>	0	0	0	0	0	0	0	0	0	1	51	0	0	0
T>	1	0	0	0	0	0	0	0	0	0	0	31	0	0
U>	0	0	0	0	0	0	0	0	0	0	0	0	16	0
X>	0	0	1	0	0	3	0	0	0	0	0	0	0	12

3.1 Discrimination sur échantillon d'épreuve après juxtaposition des tableaux de facteurs issus de l'analyse des deux groupes de variables

Les deux groupes, J1 et J2, de variables, considérés aux §§1 et 2, diffèrent quant à leur nature et aussi quant à leur ordre de grandeur. C'est pourquoi on n'a pas entrepris, d'abord, d'analyser un tableau hétérogène comprenant l'ensemble des données.

Mais, désirant prendre en compte les deux types d'information afin de parfaire la discrimination des espèces, on a juxtaposé en un même tableau, non les deux groupes de variables, mais les deux blocs de facteurs, respectivement 14 et 8, issus de l'analyse de $Ib \times J1$ et $Ib \times J2$.

En termes géométriques, ceci revient à construire un espace euclidien où la distance, au carré, est la somme des carrés des distances du χ^2 afférentes aux deux tableaux. La discrimination repose exclusivement sur les distances, sans qu'il soit besoin de faire choix de masses (comme en CAH, cf. *infra*)

Avec 27 erreurs, seulement, sur 398 affectations, le taux de succès dépasse 93,2%; pourcentage qui dépasse de beaucoup ceux obtenus aux §§1.4.2 & 2.3, en ne prenant en compte qu'un seul des deux groupes de variables, J1 ou J2. En particulier, 6 des 7 individus d'épreuve rentrant dans l'espèce O sont correctement affectés.

W: Mer de corail; X: Baie australienne; Y: Mer de Tasmanie;

S	NW	NX	NY	NZ	xx
NW>	13	2	4	0	1
NX>	2	28	3	0	1
NY>	6	3	3	0	0
NZ>	0	1	0	0	0

Sans entrer dans le détail des variantes régionales, considérons les affectations obtenues pour l'espèce la plus nombreuse, N = *T. novæ-zelandiæ*. L'échantillon d'épreuve compte 67 individus de diverses origines. L'espèce a été bien reconnue pour 65 d'entre eux. Le tableau ci-joint signale que des 20 individus, NW, originaires de la Mer de corail, un seul a une affectation en dehors de l'espèce (colonne xx); dans 13 cas, l'individu, dans l'échantillon de base, le plus proche appartient à la même population: NW \rightarrow NW; il y a, respectivement, 2 et 4 affectations à des individus de NX et de NY. Il apparaît qu'un tiers des individus a été rattaché à une population autre que la sienne.

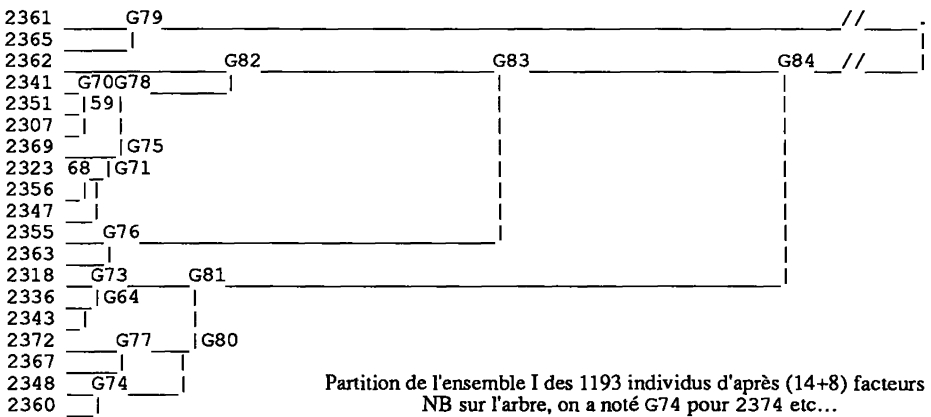
3.2 Classification de l'ensemble des individus après juxtaposition des tableaux de facteurs issus de l'analyse des deux groupes de variable

L'analyse discriminante prend en compte explicitement la subdivision de l'échantillon de base Ib en 14 espèces. L'exactitude de la discrimination suggère l'hypothèse que, dans l'espace, les sous-nuages afférents aux espèces sont nettement séparés. S'il en est ainsi, une classification automatique des individus, effectuée sans prendre en compte leur appartenance à des espèces, devrait retrouver celles-ci comme des classes d'une partition définie par les nœuds les plus hauts.

De façon précise, on fait une CAH avec pour critère d'agrégation la variance de la partition; en utilisant la même distance qu'au §3.1. Ces choix faits, il reste à attribuer aux individus des masses: dans la CAH décrite ci-après, on a donné à chaque individu une masse proportionnelle au total du bloc J1 des variables morphométriques. Les résultats ainsi obtenus sont un peu plus satisfaisants que si l'on donne aux individus la même masse, ou une masse proportionnelle au total du bloc J2 des variables méristiques. Ce qu'on expliquera par le fait que la pondération suivant J1, accroît l'importance

	C	T	D	J	L	O	N	I	R	M	X	S	U	P
2361	49	11	0	0	0	0	0	0	0	0	0	0	0	0
2365	2	85	0	0	0	0	0	0	0	0	0	0	0	0
2362	0	0	41	0	0	0	0	0	0	0	0	0	0	0
2341 \	0	0	0	35	0	0	5	0	0	0	0	0	0	0
2351 \	0	0	0	39	17	0	5	0	0	0	0	0	0	0
2307 /	0	0	0	4	29	0	5	0	0	0	0	0	0	0
2369	0	0	0	13	47	1	4	1	0	1	0	0	0	0
2323 \	0	0	0	0	0	17	3	0	0	0	0	0	0	0
2356 \	0	0	0	0	2	0	111	0	0	0	0	0	0	0
2347 /	0	0	0	0	4	0	67	0	0	0	0	0	0	0
2355	0	0	1	0	0	0	0	62	0	2	0	0	0	0
2363	0	0	0	0	0	0	1	3	112	0	0	0	0	0
2318	0	0	0	0	0	0	0	0	0	50	2	0	0	0
2336 \	0	0	0	0	0	0	0	0	0	7	36	4	0	0
2343 /	0	0	0	0	0	0	0	0	0	37	8	0	0	0
2372	0	0	0	0	0	0	0	0	0	3	0	2	0	63
2367	0	0	0	0	0	0	0	0	0	0	0	0	44	2
2348	0	0	0	0	0	0	0	0	0	2	2	47	2	0
2360	0	0	0	0	0	0	0	0	0	0	1	103	0	0

relative des individus de grande taille; qui sont ceux dont la forme spécifique est le plus parfaitement développée.



Partition de l'ensemble I des 1193 individus d'après (14+8) facteurs NB sur l'arbre, on a noté G74 pour 2374 etc...

Pour construire la partition, dont le tableau donne le bilan par classes, on a gardé les nœuds 2373 et au dessus (13 nœuds) et subdivisé certaines des 14 classes de la partition ainsi définie, quand cela semblait bon pour montrer des classes assimilables à des espèces. Sur le tableau, les colonnes ont été rangées de telle sorte qu'apparaissent, au mieux, comme des blocs, les groupes d'espèces contiguës, dont le découpage est suggéré, par des 'X'; et la hiérarchie des subdivisions est rappelée, en marge, par des traits \ ou /.

Le genre *Trachurus*: M. Ben Salem : k1 : J2 repondéré pour égaler J1
 trace :1.811e-2
 rang : 1 2 3 4 5 6 7 8 9 10 11
 lambda: 110 30 10 7 5 4 3 2 2 2 1 e-4
 taux :6076 1654 564 391 251 216 157 125 104 90 60 e-4
 cumul :6076 7730 8294 8685 8936 9151 9308 9433 9538 9627 9688 e-4

3.3 Analyse factorielle globale et classification des individus

La valeur taxinomique de la CAH considérée au §3.2 incite à en retrouver les résultats par l'analyse d'un seul tableau de correspondance, contenant toutes les données, de J1 ou de J2. On a dit que les deux blocs de variables ne sont pas du même ordre de grandeur: pour rétablir l'équilibre, il suffit de multiplier l'un des deux, ou les deux, par un coefficient de telle sorte que, pour chaque individu, ils aient même total.

De façon précise, nous avons expérimenté trois procédés: en construisant des tableaux k1, k2 et k3. On part du tableau k, dont l'ensemble des colonnes est $J1 \cup J2$. Notons, respectivement, tot1(i) et tot2(j) les totaux, pour l'individu i, des variables des bloc J1 et J2:

$$\text{tot1}(i) = \sum\{k(i, j1) \mid j1 \in J1\} \quad ; \quad \text{tot2}(i) = \sum\{k(i, j2) \mid j2 \in J2\} \quad .$$

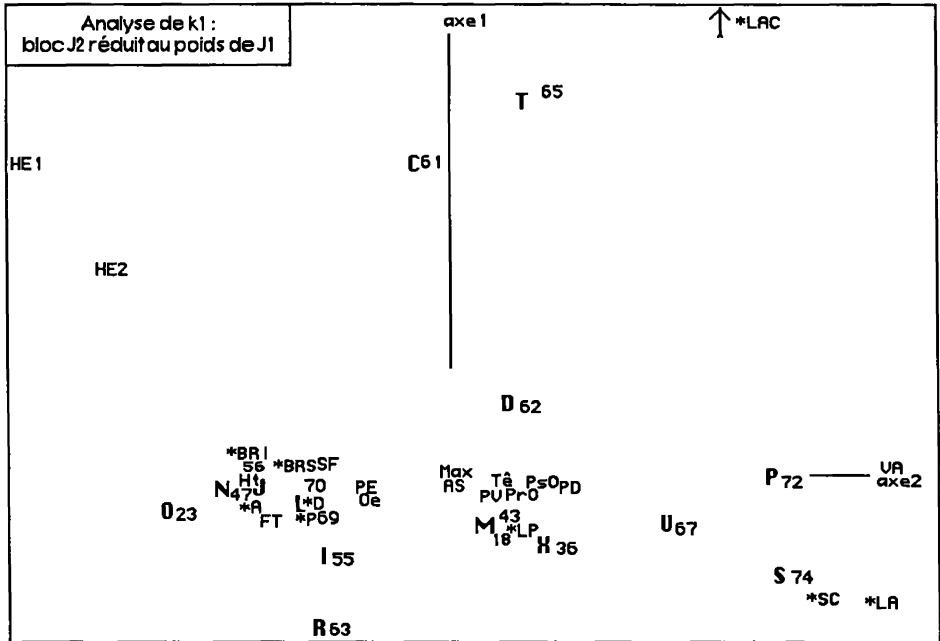
Le tableau k1 est défini en laissant tel quel les colonnes du bloc J1 et multipliant celles du bloc J2 par (tot1(i)/tot2(i)). Le tableau k2 est défini en laissant tel quel le bloc J2 et multipliant le bloc J1 par (tot2(i)/tot1(i)). Enfin, pour obtenir le tableau k3, on divise les colonnes du bloc J1 par tot1(i); et celles du bloc J2 par tot2(i).

Dans k1, pour chaque i, chacun des deux blocs J1 et J2 a pour total le total qu'a le bloc J1, dans le tableau initial; dans k2, les deux blocs ont pour total le total qu'a le bloc J2, dans le tableau initial; enfin, dans k3, les deux blocs ont chacun pour total 1:

$$\forall i \in I: \sum\{k3(i, j1) \mid j1 \in J1\} = \sum\{k3(i, j2) \mid j2 \in J2\} = 1 \quad .$$

Il importe de noter que, dans chacun des trois cas, parce que les deux blocs ont même total dans chaque ligne, le carré de la distance distributionnelle entre deux éléments i et i', n'est autre que la demi-somme des carrés des distances afférentes aux deux blocs J1 et J2. Certes, les facteurs issus de l'analyse du tableau complet, $I \times (J1 \cup J2)$, ne sont pas ceux issus des tableaux partiels: $I \times J1$ et $I \times J2$; mais si l'on fait la somme des carrés des facteurs issus du tableau complet, on trouve exactement, la moitié de la somme des deux totaux obtenus avec les facteurs issus des tableaux partiels.

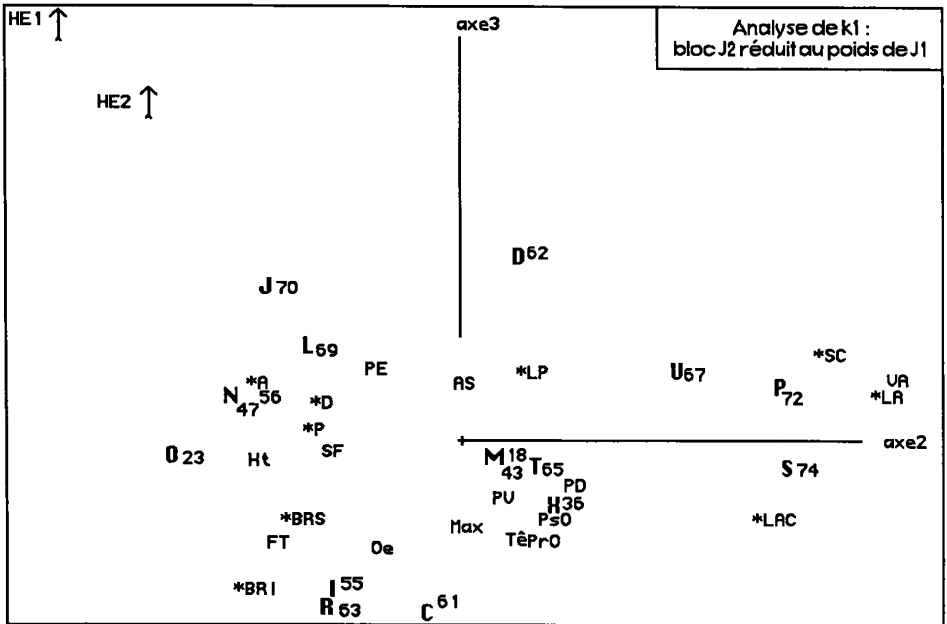
On retrouve donc, avec k1, le procédé de juxtaposition auquel on a eu recours aux §§3.1 et 3.2. Le premier bloc de colonnes, J1, est bien, dans k1, identique au tableau des variables morphométriques, considéré au §1; le deuxième bloc, n'est pas exactement identique au tableau analysé au §2, car les lignes ont été repondérées; mais chacune a conservé son profil; et la



distance du χ^2 n'est pas sensiblement modifiée, parce que le profil de marge, qui en fournit les coefficients, ne l'est pas.

En effet, la CAH de l'ensemble I des 1193 spécimens, issue de l'analyse de k1, est bien celle considérée au §3.2. L'analyse factorielle permet d'en montrer le nuage des centres de classes, e.g., dans les plans 2 × 1 et 2 × 3; avec les variables et les centres des espèces. Sur les graphiques, les espèces sont marquées d'une lettre grosse. Les centres des classes de la CAH, ayant toutes des numéros commençant par 23, sont notées par les deux chiffres qui suivent, e.g. 61 pour 2361; et, pour plus de clarté, on a fondu certaines subdivisions dont le profil spécifique est le même: e.g., au lieu de 2348 et 2360, où prédomine l'espèce **S**, on a placé seulement leur réunion, 2374, notée 74. Enfin, les sigles des variables du bloc J2 sont précédés d'une étoile '*'.

Dans le plan 2 × 1, on a une distribution des espèces semblable à celle déjà vue au §2.2. Là comme ici, l'axe 1 est dominé par LAC (noté *LAC); et l'axe 2, par *LA (associé à VA). Comme on pouvait l'attendre, les classes où prédomine une espèce tombent à proximité du centre de cette espèce: par exemple, le nombre 74 (pour 2374, cf. *supra*) se lit à côté de la lettre **S**; les nombres 61 et 65, à côté de **C** et **T** (;en fait , **C** et 61 sont presque rigoureusement superposés et 61 a été déplacé sur le graphique); **D** est avec 62; **R** avec 63; **I** avec 55.



L'axe 3 est créé par HE1, associé à HE2 et opposé à BRI. Quant aux espèces, on a sur ($F3 > 0$), D (62) suivi de J et L (70 et 69); et sur ($F3 < 0$), R, I et C: respectivement 63, 55, 61.

L'axe 4 (non représenté) est créé par FT, suivi de PE et opposé à SF. On retrouve l'axe 3 du §1.3; avec la question: l'association entre FT, pointe de la queue au-delà de la fourche, et PE, nageoire pectorale, parties fragiles qui se ressemblent quant à la texture, s'explique-t-elle par la conservation du spécimen, qui les affecte dans le même sens? Nous avons ici le moyen d'étayer une autre hypothèse, vraie, au moins partiellement.

Il s'agit d'un facteur de forme lié à la taille, sans relevance pour la distinction des espèces: du côté de ($F4 > 0$), sont les spécimens de petite taille; du côté de ($F4 < 0$), avec des pointes développées, ceux de grande taille (relativement à leur espèce).

Considérons, en effet, le tableau croisant la partition de l'ensemble I des spécimens retenue au §3.2, avec un découpage en classes de la variable LT, longueur totale; variable qui, sur le schéma illustrant le §1.1, correspond au segment horizontal PT, qu'en vue des analyses de profils on a découpé en cinq segments consécutifs: {PV VA AS SF FT}.

Dans ce tableau, chaque ligne donne la répartition, quant à LT, des spécimens d'une classe. Prenons, par exemple, la 6-ème ligne, afférente à la

LTsup	88	127	166	204	243	282	318	357	395	425	468	497	540	630
1 2361	11	7	3	13	14	3	2	1	2			1	3	
2 2365	19	23	16	14	4	4	4	2	1					
3 2362	1	5	1	9	7	4	3	3	1	3	2	2		
4 2341	1	8	14	4	4	2	6	1						
5 2351	3	24	29	3	1	1								
6 2307	29	9												
7 2369	1	13	17	12	11	7	5							
8 2323	5	6	2	5	2									
9 2356	2	79	25	4					2	1				
10 2347		5	25	28	7	3		3						
11 2355	2	13	23	24	3									
12 2363	3	31	40	23	11	3	2	3						
13 2318	35	13	4											
14 2336		5	33	4	3			1						1
15 2343	2	5	14	7	9	3	4						1	
16 2372		5	25	22	8	3	4	1						
17 2367	2	8	3	3		1	2	6	15		4	1	1	
18 2348	9	39	3	2										
19 2360		5	23	57	14	5								

classe 2307: celle-ci compte 29 spécimens dont la longueur est ≤ 88 mm; et 9 autres, dont la longueur est comprise entre 89 et 127 mm; au-delà, les autres modalités de LT sont vides. Au contraire, sur la 17-ème ligne, on lit que la classe 2367 contient, entre autres, 15 poissons dont la longueur est comprise entre 358 et 395 mm; taille que dépassent très peu de spécimens de notre échantillon.

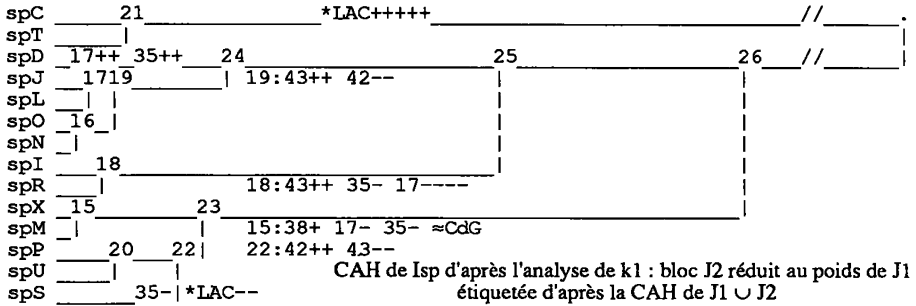
En consultant le listage FACOR, on distingue quelques classes dont les centres s'écartent le plus de l'origine sur l'axe 4; quelques dichotomies fortement corrélées avec cet axe.

La quasi-totalité des poissons de l'espèce S sont compris dans les deux classes 2348 et 2360, subdivisions de la classe 2374 qui s'opposent, principalement suivant la direction de l'axe 4. Il apparaît que l'espèce S ne contient pas de poisson de très grande taille; mais les plus développés (dont la taille est moyenne, relativement à notre échantillon du genre *Trachurus*) sont dans 2360, classe qui ne contient qu'un seul poisson étranger (soit: 1%). Dans 2348 ($F_4 \gg 0$), au contraire, la plupart des poissons sont ≤ 127 mm; il y en a 6 ($\approx 10\%$) qui ne sont pas de l'espèce S: on dira que tout en étant de forme similaire, les petits poissons de 2348 sont moins caractérisés que ceux, développés, de la classe 2360.

L'espèce N s'identifie, à peu près, à la réunion des classes 2347 et 2356. Dans 2347, ($F_4 < 0$), la taille est, en moyenne plus forte que dans 2356.

L'espèce U s'identifie à la seule classe 2367 ($F_4 < 0$), où il y a des poissons de toutes tailles; mais, particulièrement, (cf. *supra*) 21 spécimens dont la longueur dépasse 357 mm.

La réunion des espèces J et L fait l'essentiel des classes 2369 et 2370; cette dernière subdivisée en {2341, 2351, 2307}. Les poissons de L prédominent dans 2369 et 2307: ces deux classes ont, quant à la taille, des inventaires bien différents; dans 2307, ($F_4 \gg 0$), on n'a que de très petits poissons.



3.4 Classification des espèces et classification des variables d'après l'analyse factorielle globale

3.4.1 Classification des espèces

À l'analyse du tableau k1 de description des 1193 spécimens par les deux groupes de variables, morphométriques et méristiques, J1 et J2, on a adjoint en supplément l'ensemble Isp des profils des espèces définis par cumul de leurs individus. Après les représentations spatiales de Isp dans les deux plans 2 × 1 et 2 × 3, on considère ici une classification ascendante hiérarchique.

On retrouve ainsi sous une autre forme, ce qui a déjà été vu. Au sommet de la hiérarchie, se séparent les deux espèces {C, T} agrégées entre elles à un bas niveau. Le reste, i.e. la classe 26, est au niveau des nœuds 26 et 25, scindé en trois classes principales, {23, {24, 18}} :

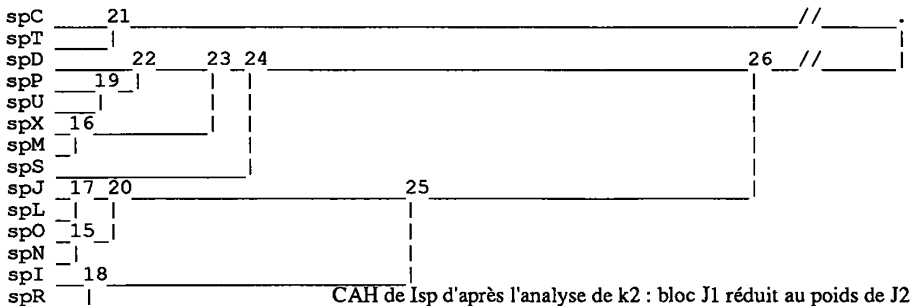
$$23 = \{(X, M), (P, U), S\} ; 24 = \{D, (J, L), (O, N)\} ; 18 = \{R, I\} .$$

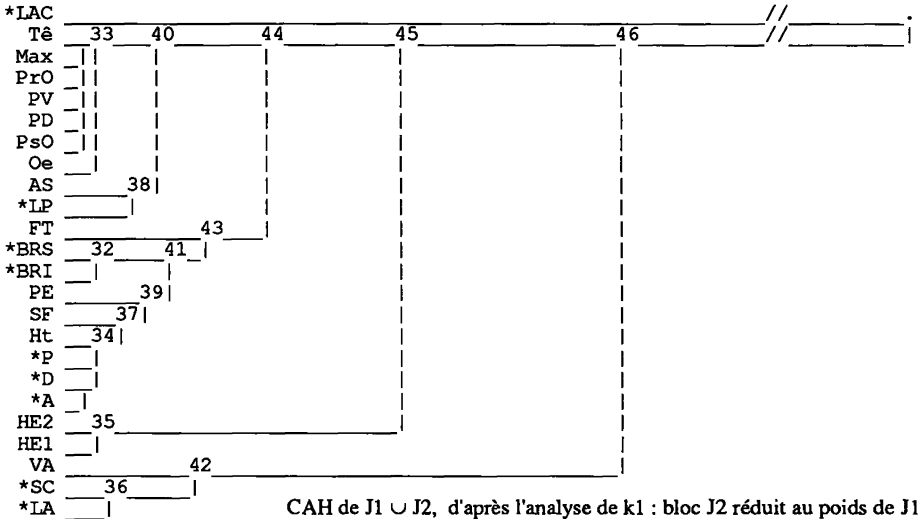
au sein desquelles on a distingué, entre parenthèses, des couples d'espèces agrégées entre elles, et dont la proximité est déjà apparue, aussi bien sur les graphiques plans du §3.2 que dans le dénombrement des confusions, au §3.1.

Comme terme de comparaison, on publie la CAH de Isp obtenue en procédant comme avec k1, mais à partir du tableau k2, défini au §3.2. Quant à la partition en 4 classes définie par les 3 nœuds les plus hauts :

$$\{CT\} \{IR\} \{(JL) (ON)\} \{(PU) (XM) S\} ; D \text{ incertæ sedis};$$

il n'y a qu'une seule discordance: la place de D.





3.4.2 Classification des variables

Au sommet de la hiérarchie, se sépare *LAC, qui crée l'axe 1.

Puis la classe 42, {VA, *SC, *LA}, qui comprend les variables créant l'axe 2; lesquelles expriment le développement de la partie ventrale.

On ne s'étonnera de trouver agrégées, à l'écart du reste, {HE1 HE2} dimensions d'écaillés et scutelles dont on dit qu'elles n'imposent pas de très forte contrainte à l'harmonie des autres variables.

Restent les classes 43 et 44; celle-ci étant subdivisée en 33 et 38. Dans 33 sont les dimensions de la tête; dans 38, celles du segment postventral et précaudal du corps.

Dans 43, enfin, on trouve tout ce qui, queue, nageoire et rayons, fait la parure du poisson aux yeux de l'amateur d'aquarium.

C'est d'après la CAH des variables qu'on a étiqueté la CAH des espèces.

4 Perspectives et conclusions

Selon l'usage ancien de la taxinomie, une espèce est définie par un individu particulier sur lequel la description en a été faite: le type. Il est clair que les individus d'une espèce ne peuvent être tous identiques au type: ils se conforment à celui-ci de plus ou moins près. On admet aujourd'hui, selon les vues synthétiques de SIMPSON, qu'une espèce est un ensemble de populations interfécondes; dont chacune a en partage un patrimoine génétique dont les caractères comportent des allèles inégalement représentés quant à la fréquence. De cette diversité essentielle résulte un étalement de formes, qui ne peut être adéquatement saisi que sur un ensemble assez nombreux d'individus.

Le procédé statistique utilisé, au §3.1, pour la discrimination des espèces s'accorde avec une telle conception. Plutôt que par un individu typique, ou même par un individu moyen fictif, c'est par l'ensemble de ses représentants dans l'échantillon de base que chaque espèce est décrite. Et tout individu de l'échantillon d'épreuve est comparé à chacun des individus de l'échantillon de base.

On rejoint ainsi l'intuition du spécialiste, qui, devant un cas nouveau, sans s'engager dans le tracé hiérarchique d'une clef, évoque les cas dont il a conservé l'image et s'arrête à celui qui convient. Naguères, la taxinomie mathématique renvoyait à des formules dont l'application était laborieuse; aujourd'hui, l'outil de calcul peut, quasi instantanément, rapporter tout cas nouveau à un corpus, ou collection, gardé en mémoire. La méthodologie aura pour rôle de guider le choix des mesures et l'appréciation, par l'analyse de corpus, du pouvoir discriminant de celles-ci. Telle est la voie que nous avons explorée ici.

Même si le problème de la discrimination entre des espèces déjà constituées a été, croyons-nous, bien résolu par nos calculs de distance entre formes, on peut se demander dans quelle mesure la CAH des individus et celle des espèces suffisent à fonder la taxinomie du genre *Trachurus*. Certes, la science, classique, du milieu et la biologie moléculaire contemporaine sont, l'une et l'autre, indispensables à l'intelligence de la spéciation; mais les classifications des §§3.2 et 3.4 montrent que l'essentiel de la taxinomie s'impose à qui a considéré mûrement les formes.

Reste à poursuivre nos investigations en considérant d'autres genres de poissons; et des groupes plus étendus: la famille des carangidés, où est compris le genre *Trachurus*, semble toute désignée pour cela.

Référence bibliographique

Dictionnaire des noms des animaux en cinq langues, POISSONS, publié sous la direction de l'Académicien V. E. SOKOLOV; Moscou, 1989.

L. BELLIER : "L'étude de la forme du squelette de la tête chez le *Praomys*, petit rongeur africain", in *L'Analyse des Données*, 1, *La Taxinomie*; J.-P. BENZÉCRI et coll.; Paris, 1973.

ejusdem, ibidem: "Sur le genre *Cricetomys*", rats géants africains de la famille des cricétidés.