

ACOM et ordination de K ensembles de relevés

Résumé

On compare l'analyse de co-inertie multiple (ACOM) à la méthode STATIS en testant sa capacité à ordonner simultanément K ensembles de relevés. On utilise pour ce faire les données hydrobiologiques de la thèse de J. Verneaux (Cours d'eau de Franche-Comté (Massif du Jura). Recherches écologiques sur le réseau hydrographique du Doubs. Essai de biotypologie. Thèse d'état, Besançon. 1-257). Les trois exemples apportent une compréhension plus fine du rôle des deux méthodes. L'ACOM fait K analyses (coordonnées) et en trouve une moyenne. STATIS cherche une moyenne et en fait une analyse. Si chaque analyse est pertinente et mérite un examen approfondi, l'ACOM se justifie. Si chaque analyse est d'intérêt limité STATIS est plus clair. Dans la comparaison de peu d'études régionales comportant beaucoup de relevés, on pourra préférer l'ACOM. Dans la comparaison de nombreuses études régionales comportant un nombre limité de relevés, on devra s'en tenir à STATIS. Les deux méthodes sont centrées sur la notion de structures communes et les résultats communs qu'elles donnent avec des logiques différentes sont normalement cohérents.

Plan

1 — La structure géomorphologique.....	2
1.1 — Coinertie et variables communes.....	3
1.1 — Gradient amont-aval	7
2 — Positions typologiques des espèces	9
2.1 — La position spécifique de l'AFC.....	9
2.2 — A la recherche d'une rivière de référence.....	11
2.3 — Logique d'ACP.....	15
3 — Physico-chimie : Rv et coinertie.....	17
Références	21

D. Chessel & M. Hanafi

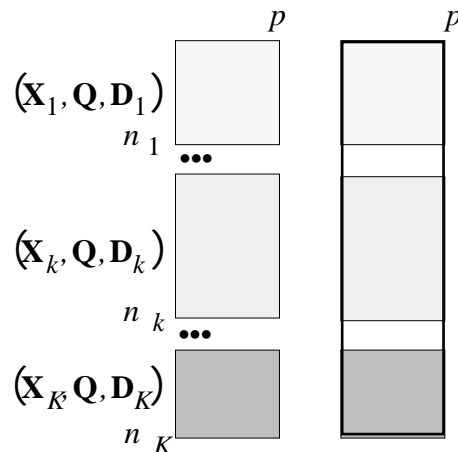
1 — La structure géomorphologique

Les données traitées sont strictement celles du tableau 1 de la fiche ACT-Statist 2 (fascicule 6). Les données d'origine ont été recodées avec précision. 92 stations sont réparties sur 12 rivières et supportent l'enregistrement de 5 variables géomorphologiques (Altitude, distance à la source, pente, section mouillée et débit moyen). Le tableau Geo comporte donc 92 lignes et 5 colonnes. Le fichier blolignum :

```

Input file: blolignum
Row: 12 Col: 1
 1 | 16.0000 |
 2 |  6.0000 |
 3 |  5.0000 |
 4 |  8.0000 |
 5 |  5.0000 |
 6 |  4.0000 |
 7 | 17.0000 |
 8 |  5.0000 |
 9 |  9.0000 |
10 |  8.0000 |
11 |  3.0000 |
12 |  6.0000 |
    
```

indique la répartition des 92 stations en 12 blocs. Dans cette configuration, il y a identité entre la présentation théorique et l'implantation pratique. On dispose de K tableaux ayant en commun les colonnes (5 variables), les lignes étant les stations réparties en 12 groupes (rivières). On dispose donc de K triplets statistiques $(\mathbf{X}_k, \mathbf{Q}, \mathbf{D}_k)$ ($1 \leq k \leq K$) :



On définit le K -tableaux :

InitKTab	
Matrix input file	<input type="text" value="Geo"/> 92 5
Row indicator (Default = 1 class)	<input type="text" value="blolignum"/> 12 1
Col indicator (Default = 1 class)	<input type="text"/>
Option: output file name	<input type="text"/>

Les fichiers auxiliaires sont en place :

```

Qualitative variables file: Geo!TLI
Number of rows: 92, variables: 3, categories: 121
Auxiliary ASCII output file Geo!TLI.label: labelling file
-----
    
```

Qualitative variables file: Geo!TCc
 Number of rows: 60, variables: 3, categories: 22
 Auxiliary ASCII output file Geo!TCc.label: labelling file

 Qualitative variables file: Geo!T4a
 Number of rows: 48, variables: 2, categories: 16
 Auxiliary ASCII output file Geo!T4a.label: labelling file

On utilise un centrage par colonnes dans chacun des tableaux :

CentringKtab	
---.ktp input file	<input type="text" value="Geo!.ktp"/>
Matrix input file	<input type="text" value="Geo"/> 92 5
Row weighting (default 1/nk)	<input type="text"/>
Column weighting (default 1/pj)	<input type="text"/>
Centring option (no default)	<input type="text" value="2"/>
Option: output file name	<input type="text" value="A"/>

Option CentringKtab
 Input file Geo. Number of rows: 92, columns: 5
 Each row has uniform weight (Sum inside block = 1)
 File A.ktpl contains the row weights
 It has 92 rows and 1 column
 Each column has uniform weight (Sum inside block = 1)
 File A.ktpc contains the column weights
 It has 5 rows and 1 column
 File A.ktta contains the (column) block-centred table
 It has 92 rows and 5 columns

Les ACP séparées définissant une structure unidimensionnelle (fiche Statis 2, p. 9) on garde un axe (fiche ACOM et valeur typologique, p. 3) dans l'analyse de coinertie multiple :

ACOM	
---.kttta input file	<input type="text" value="A.ktta"/> 92 5
Table weights (default = 1/Inertia)	<input type="text" value="2"/>
Selected axis number (default = 8)	<input type="text" value="1"/>

On remarque que l'option exécute les analyses séparées dans le cas où l'option Separate analysis n'a pas été sollicitée explicitement.

1.1 — Coinertie et variables communes

Le dépouillement suppose qu'on s'interroge sur le principe de coinertie (fiche ACOM introduction et fiche ACOM valeur typologique, figure 2 p. 9). Dans une analyse simple on fait naturellement deux analyses d'inertie fortement structurées par le schéma de dualité. Il n'est peut-être pas inutile de rappeler cet aspect dans une ACP centrée (figure 1).

Generalised Ktab-PCA analysis
 Input file A
 Row number: 92, Column number: 5, Block number: 12

```

Selected weighting option
  -> Uniform weights
vp level 1=9.7013e-01
File A.--C1 contains column scores (unit norm)
It has 5 rows and 1 columns
File :A.--C1
----- Minimum/Maximum:
Col.:  1 Mini = -1.1311  Maxi = 1.4826

```

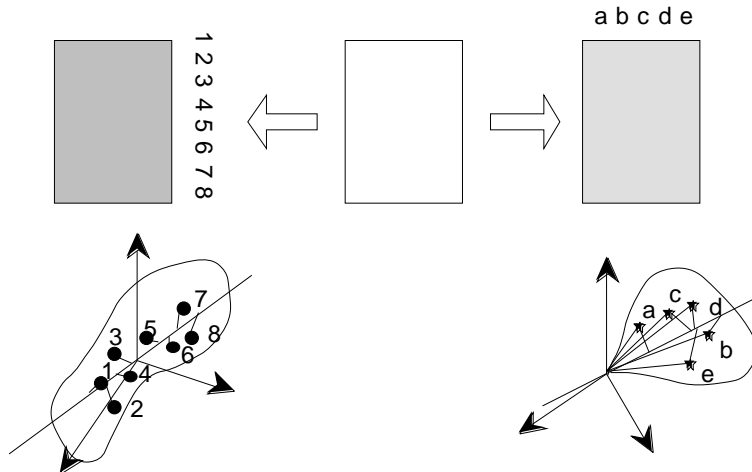
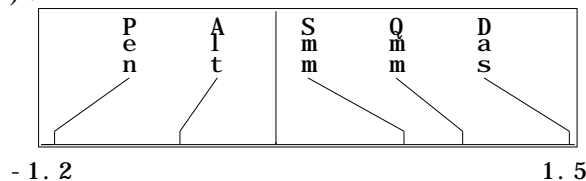


Figure 1 : schéma de principe d'une analyse d'inertie. Un tableau définit deux nuages. Chaque nuage définit ses éléments principaux. Pour une ACP classique centrée par colonne le nuage des lignes (à gauche) est centré et son premier élément principal est appelé axe principal. Le nuage des colonnes (à droite) n'est pas centré et son premier élément principal est appelé composante principale. Les n (par exemple 8) lignes sont projetées sur l'axe principal et les coordonnées des projections sont proportionnelles aux composantes dans l'autre espace de la composante principale. Les p (par exemple 5) colonnes sont projetées sur la composante principale et les coordonnées des projections sont proportionnelles aux composantes dans l'autre espace de l'axe principal. Le même principe d'inertie maximale conduit à maximiser une variance des coordonnées des projections (à gauche) et une somme de carrés de covariances (à droite).

Le nombre de variables communes aux tableaux est p (5). On obtient dans A.--C1 un vecteur de \mathbb{R}^p , normé pour la pondération uniforme (moyenne des carrés des composantes égale à 1) :



Ce code à 5 valeurs est répété 12 fois pour des raisons pratiques :

```

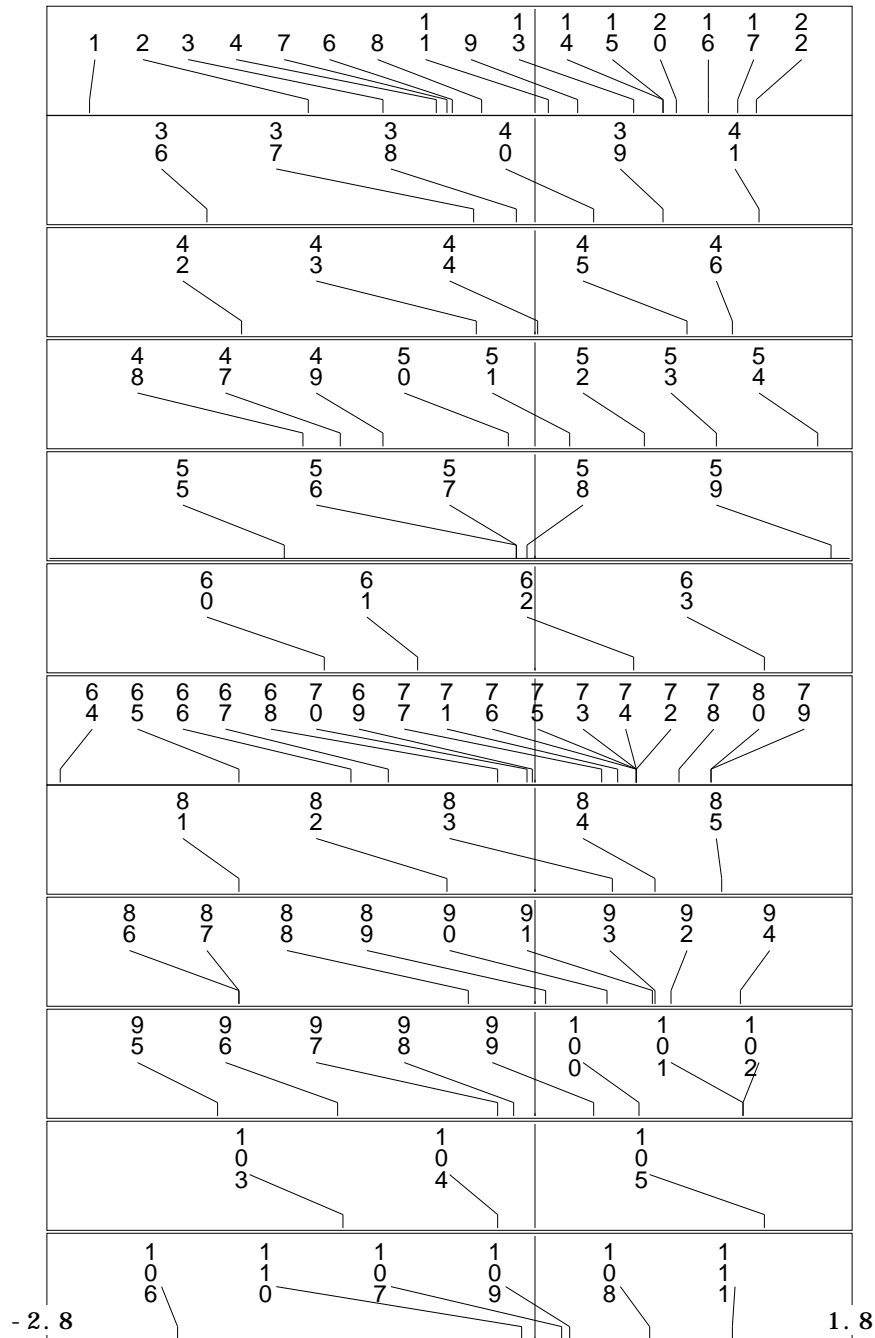
File A.--TC1rep contains column scores (unit norm)
Duplication of the previous file
It has 60 rows and 1 columns
It is to be used with --Tcc.label and --Tcc.cat files
File :A.--TC1rep
----- Minimum/Maximum:
Col.:  1 Mini = -1.1311  Maxi = 1.4826

```

En même temps, dans les K (12) espaces \mathbb{R}^{n_k} , on cherche des composantes principales (vecteurs centrés et réduits, codes numériques des stations d'une rivière) :

File A.--TL1 contains row scores (unit norm for each block)
 It has 92 rows and 1 columns
 File :A.--TL1

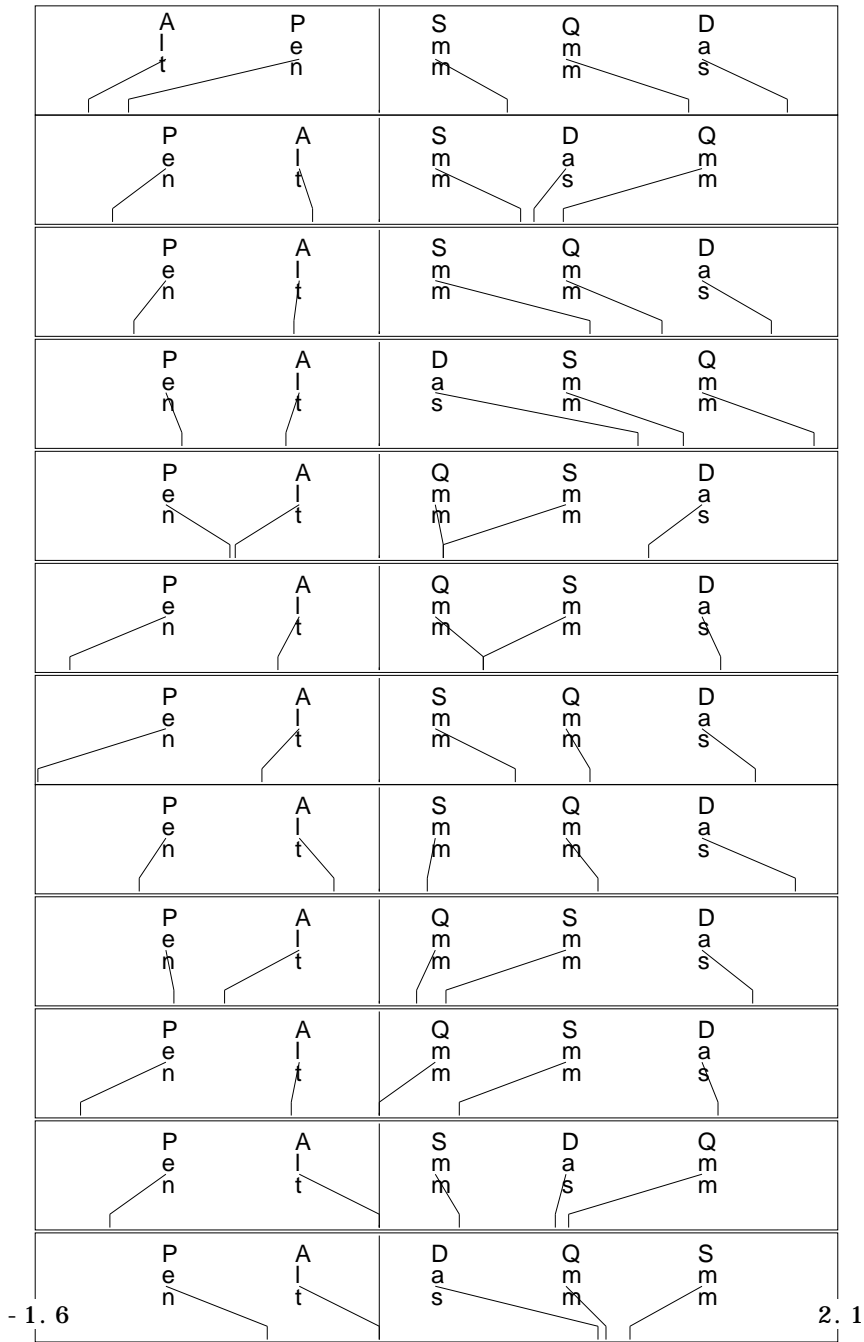
----- Minimum/Maximum:
 Col.: 1 Mini = -2.7258 Maxi = 1.7063



Ces codes numériques par rivière présentent dans chaque bloc une covariance avec chacune des variables, covariance qui dépend de la variance dans ce bloc de la variable et de la corrélation entre le code et cette variable :

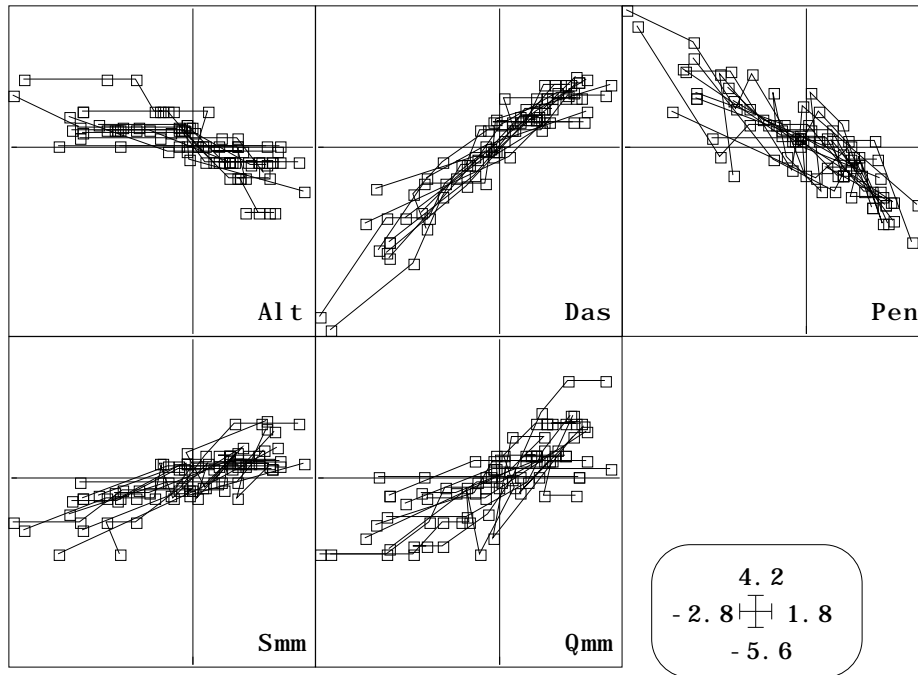
File A.--TCco contains column scores
 from projections of separate clouds
 It has 60 rows and 1 columns
 It is to be used with --TCc.label and --TCc.cat files
 File :A.--TCco

----- Minimum/Maximum:
 Col.: 1 Mini = -1.5876 Maxi = 2.0118



Cette figure n'est pas instructive. Cherchons à exprimer ces covariances avec CurveClass :

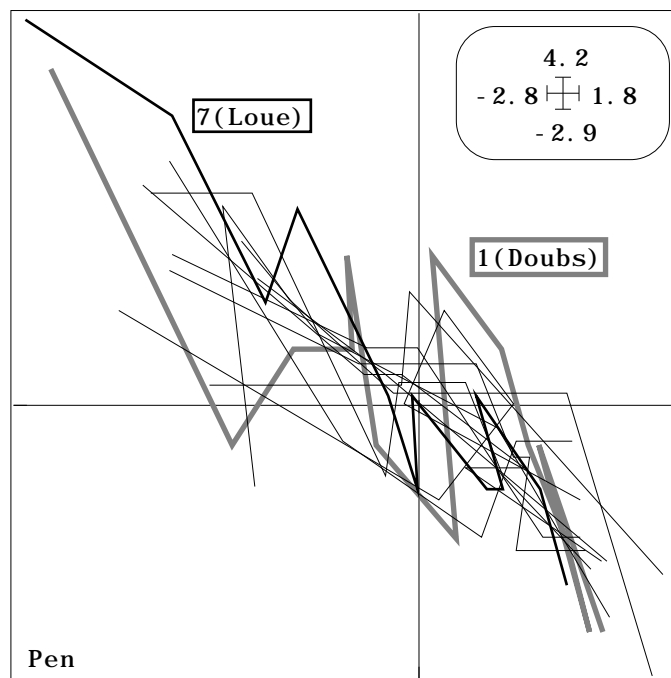
Lines	
X file (default = 1, 2, 3, ..., n)	<input type="text" value="A.--TL1"/> 92 1
X file column number (default = 1)	<input type="text"/>
Y file (no default)	<input type="text" value="A.ktta"/> 92 5
Categories file (.cat)	<input type="text" value="Geo!TLI.cat"/>
Categories file column number	<input type="text"/>
Variable label file (optional)	<input type="text" value="Label_Geo"/>



On montre clairement l'idée sous-jacente à cette analyse. On cherche à reproduire une typologie de variables, ce qui conduit à recalculer les covariances intra-blocs de ces variables avec des codes de synthèse. On perd logiquement dans cette opération toute idée de l'hétérogénéité intra-rivière.

1.1 — Gradient amont-aval

L'analyse de coinertie multiple cherche une typologie reproductible des descripteurs communs, ici peu de chose à savoir la croissance de la distance à la source, de la section et du débit et la décroissance de la pente et de l'altitude dans le gradient amont-aval. L'analyse n'est pas destinée à trouver ce résultat mais à bien comprendre son fonctionnement.

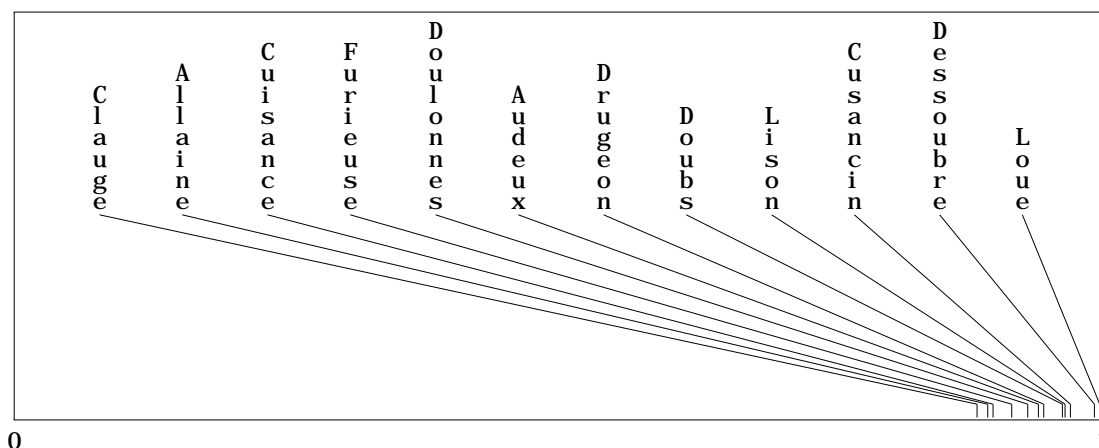


On notera cependant que l'expression de la simple covariance permet de représenter les écarts à un modèle linéaire simple, comme les ruptures de pente du système étudié (agrandissement d'une fenêtre du graphe précédent).

On retrouve numériquement la permanence de l'axe 1 des analyses séparées.

N	Iner	Var	Sca. Pro. 2	Cos2
1	1.841e+00	1.826e+00	1.685e+00	9.228e-01
2	6.544e-01	6.529e-01	5.793e-01	8.872e-01
3	1.479e+00	1.478e+00	1.446e+00	9.783e-01
4	1.722e+00	1.696e+00	1.345e+00	7.928e-01
5	5.390e-01	5.315e-01	4.685e-01	8.815e-01
6	1.047e+00	1.045e+00	9.769e-01	9.345e-01
7	1.431e+00	1.430e+00	1.410e+00	9.861e-01
8	1.211e+00	1.208e+00	1.121e+00	9.278e-01
9	9.088e-01	8.983e-01	7.502e-01	8.352e-01
10	9.361e-01	9.356e-01	7.495e-01	8.010e-01
11	6.239e-01	6.224e-01	5.360e-01	8.612e-01
12	7.464e-01	7.410e-01	5.741e-01	7.748e-01

L'expression graphique de la projection des axes 1 sur l'axe compromis (fiche Statis 2, figure 10, p. 13) trouve son équivalent dans la représentation des Cos2 de synthèse qu'on trouve dans le fichier A.--tab1 :



Il est même étonnant qu'à partir de principe théorique aussi différents on retrouve des images aussi proches. On notera donc que STATIS et ACOM sur tableaux ayant les mêmes variables (donc STATIS sur les VQ dans la terminologie de l'école de Montpellier) sont d'accord sur l'expression du compromis très simple sous-jacent au cas étudié mais que les deux méthodes indiquent de manière différente l'inégalité entre hétérogénéité intra-tableaux.

STATIS, avec le carré des normes (fiche STATIS 2, p. 13) insiste sur l'hétérogénéité des structures de covariation en sommant les carrés des covariances.

Les ACP séparées insistent sur l'hétérogénéité simple en sommant les variances et l'ACOM, si on utilise les carrés des produits scalaires (Sca. Pro. 2, sur le listing), tient compte des variances et de la structure commune (valeurs dans le fichier A.--tab)

Rappelons enfin, comme dans la fiche précédente, que l'ACOM indique l'hétérogénéité des objets spécifiques d'une étude (ici les stations) non directement mais

par l'effet produit par ces objets spécifiques sur les objets de la marge commune (ici les variables).

2 — Positions typologiques des espèces

2.1 — La position spécifique de l'AFC

Avec l'utilisation d'un tableau faunistique morcelé, on en arrive à des problèmes purement écologiques. Une question d'entrée se pose. Si on veut une AFC sur un tableau, doit-on chercher une méthode AFC sur K tableaux ? Disons le encore une fois : l'analyse des correspondances est une double analyse d'inertie doublement centrée. Cela n'est pas un privilège. Une ACP doublement centrée est de cette catégorie. Reprenons l'exemple récemment proposé par Prodon et Lebreton¹ de deux espèces A et B ayant un profil de distribution sur un transect de 10 relevés du type :

A	1	3	15	40	50	42	20	8	4	1
B	1	2	5	10	34	44	50	36	14	10

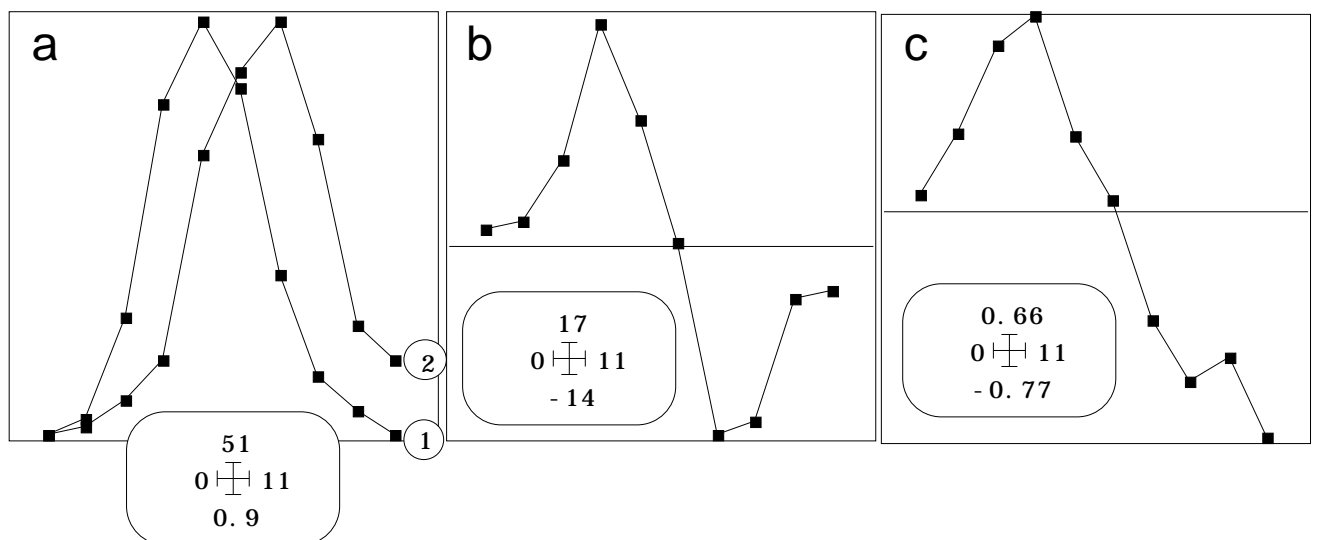


Figure 2. Gradient à deux espèces. a - Données. b - Coordonnées (relevés) d'une ACP à double centrage multiplicatif. c - Coordonnées (relevés) de l'AFC.

L'AFC est une méthode du gradient qui place les espèces au centre de gravité des relevé qui la contiennent. Ce n'est pas non plus un privilège. L'ACP du tableau en pourcentage par espèce et centrée par relevé en fait autant :

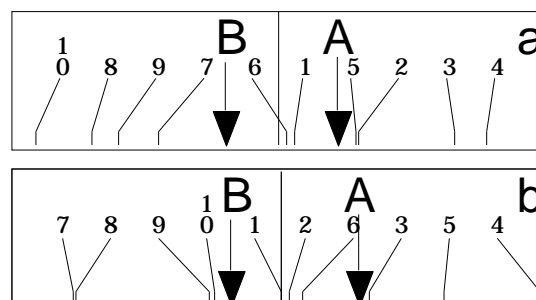


Figure 3 : Position des relevés sous contraintes pour maximiser la variance des moyennes par espèces. a - AFC. b - ACP en pourcentage par espèce centrée par relevé.

L'AFC est une méthode du gradient qui place les relevés au centre de gravité des espèces qu'ils contiennent. Ce n'est toujours pas un privilège. L'ACP du tableau en pourcentage par relevés et centrée par espèces en fait autant :

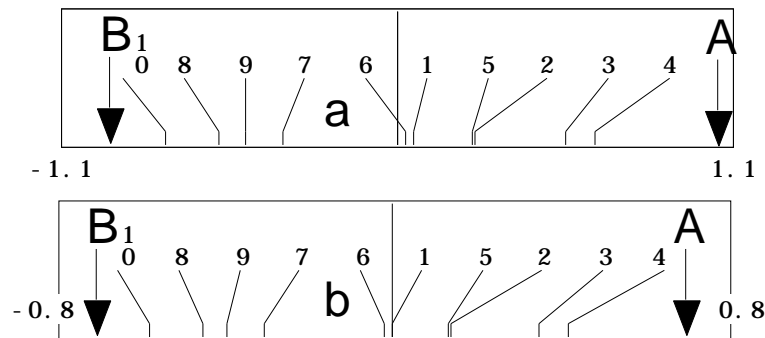
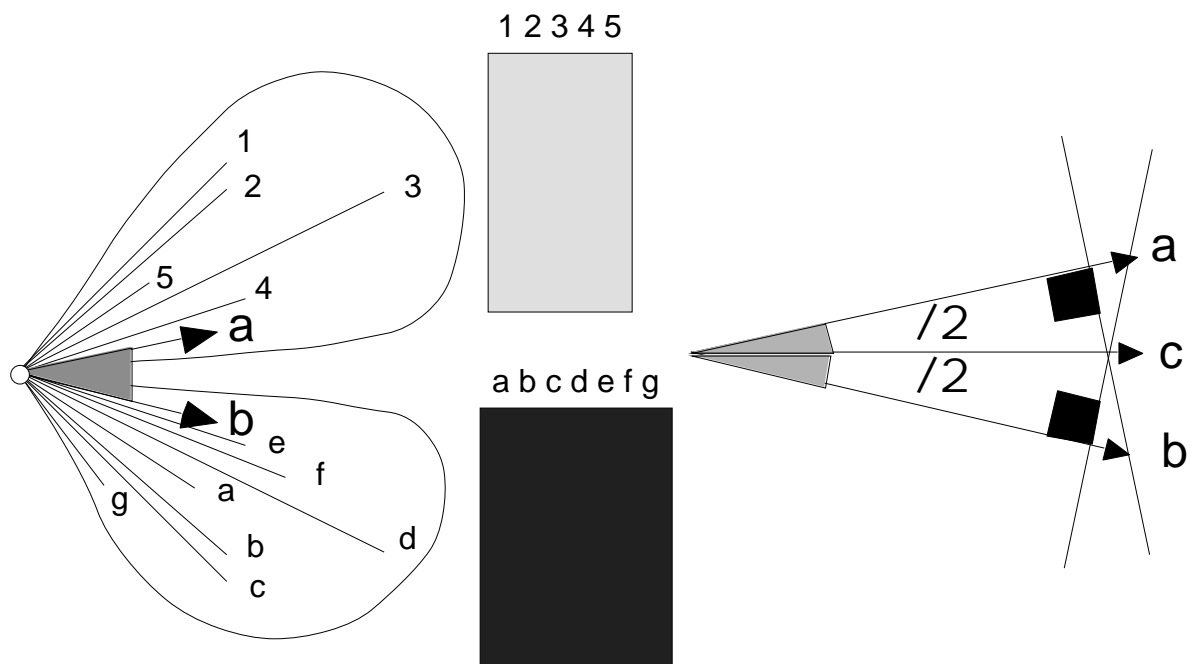


Figure 4 : Position des espèces sous contrainte pour maximiser la variance des moyennes par relevés.
a - AFC. b - ACP en pourcentage par relevé.

Le propre d'une analyse des correspondances est d'être une analyse canonique².

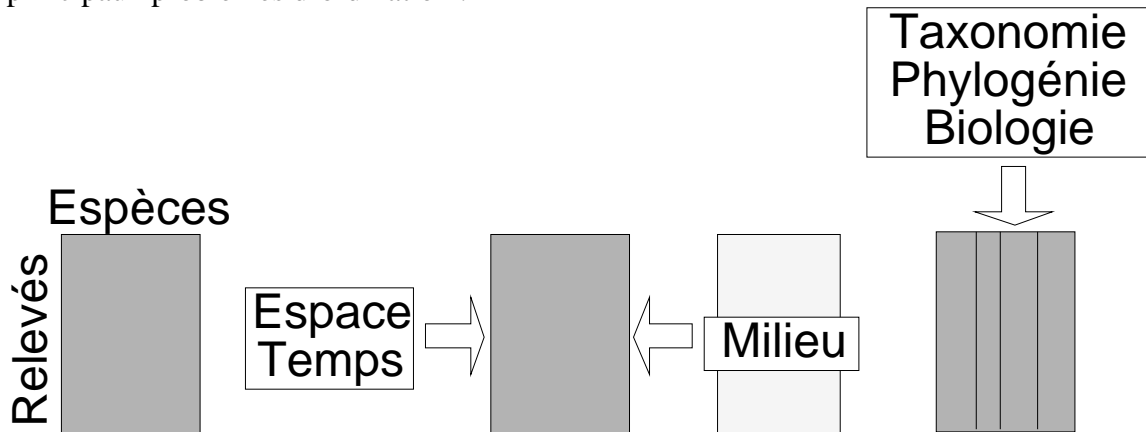
L'analyse canonique de deux tableaux X et Y ayant les mêmes lignes a une propriété très particulière dans l'espace des variables :



Trouver deux vecteurs unitaires **a** et **b**, le premier dans l'espace engendré par les colonnes de **X** et le second dans l'espace engendré par les colonnes de **Y**, faisant entre eux un angle minimum est toujours possible. A la solution, la bissectrice de l'angle défini par **a** et **b** porte un vecteur unitaire **c** dans la somme des deux espaces qui minimise sa distance aux deux vecteurs **a** et **b**. Dans l'AFC d'un tableau faunistique le vecteur **a** donne un code des espèces, le vecteur **b** donne un code des stations et le vecteur **c** donne un code des correspondances espèces-stations (les cases non nulles du tableau). Ce qui permet de symétriser non seulement position des espèces et des relevés sur un gradient mais encore diversité des relevés et amplitude de niche des espèces³.

Ces remarques préliminaires veulent souligner un fait d'importance : toute introduction d'une dissymétrie des marges d'un tableau faunistique par introduction

d'information complémentaire, en détruisant la situation d'analyse canonique ne permet plus de parler spécifiquement d'analyse des correspondances. C'est le cas des principaux problèmes d'ordination :



Par exemple, l'AFCVI ou Analyse Canonique des Correspondances a toutes les propriétés d'une analyse de correspondances sauf celle qui lui est spécifique dans l'analyse canonique —le terme analyse canonique des correspondances est malheureux—. On pourra toujours trouver une ACPVI qui aura pratiquement les mêmes propriétés. Si on tient à mettre les espèces au centre de gravité des relevés et les relevés au centre de gravité des espèces (à la corrélation canonique près) il ne faut rien ajouter aux données. Si on veut seulement les espèces au centre de gravité des relevés la version ACP sur les profils de distribution ou la version AFC seront des variantes sur les contraintes. C'est pourquoi la version AFC de la PLS n'est pas franchement meilleure que la version ACP de la PLS⁴.

Ceci permet d'éviter des querelles stériles sur les vertus particulières d'une version ACP ou d'une version AFC de la coinertie multiple. Ce sont des variantes, qui seront souvent voisines. STATIS comme l'ACOM sont des méthodes dissymétriques. Dans l'ordination de K tableaux faunistiques avec le même cortège faunistique, STATIS procède de la démarche espèces -> relevés et exprime la structure des groupes de relevés, en particulier leur diversité β , tandis que l'ACOM procède de la stratégie relevés -> espèces et cherche à reconstituer une typologie commune des taxons.

2.2 — A la recherche d'une rivière de référence

On part du tableau décrit dans la fiche Statis 2 (92 stations et 19 taxons). Chaque taxon a un mode de distribution dans chaque rivière. Nous voulons comparer les procédures STATIS et ACOM dans la version AFC. Le tableau Poi est transposé pour que les blocs de relevés soient juxtaposés :

Transpose	
Input file	<input type="text" value="Poi"/> 92 19
Output file	<input type="text" value="Poi.tr"/>

Initier l'analyse des correspondances floues :

Read Fuzzy File			
Fuzzy variables: input file (---)		Poitr	19 92
Category indication file		blolignum	12 1
Output file name (default = ---F)		B	

Input file: Poitr
 Row number: 19, column number: 92
 Number of modalities for each variable: file blolignum
 Row number: 12
 Missing data: 105

105 fois une espèce est complètement absente d'une rivière. Exécuter l'analyse :

Fuzzy Correspondence Analysis			
.fuz type file		B.fuz	
Option: Row weighting file			

Initier un K-tableaux :

FuzzyToKTab			
.flta type file		B.flta	19 92

Faire les analyses séparées :

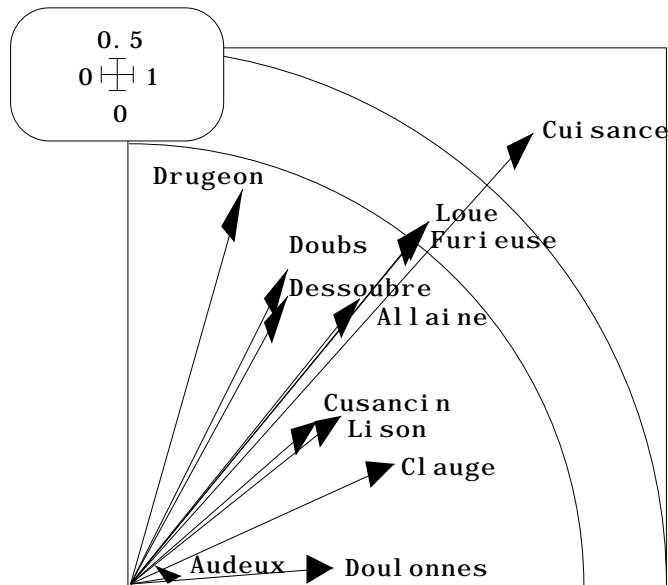
Separate analyses			
---.ktt input file		B.ktta	92 19

Garder deux facteurs dans l'ACOM avec réduction à une constante de l'inertie de tous les tableaux :

ACOM			
---.ktt input file		B.ktta	92 19
Table weights (default = 1/Inertia)			
Selected axis number (default = 8)		2	

Exécuter STATIS :

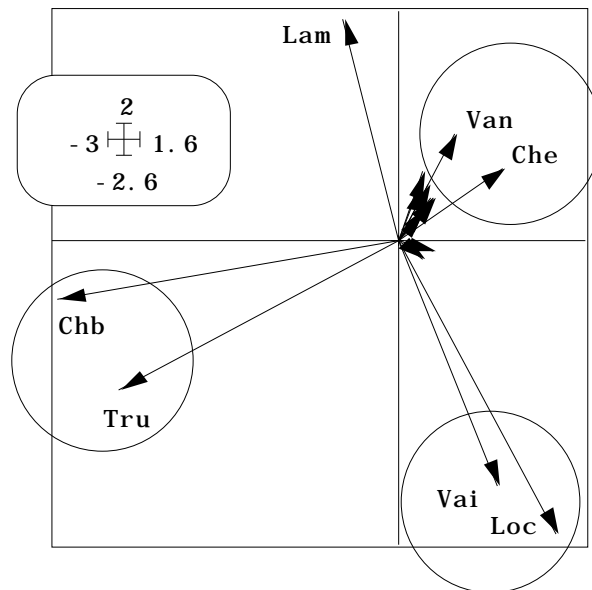
Operator averaging			
---.ktt input file		B.ktta	92 19
1 = Option COUV (default = RV)			



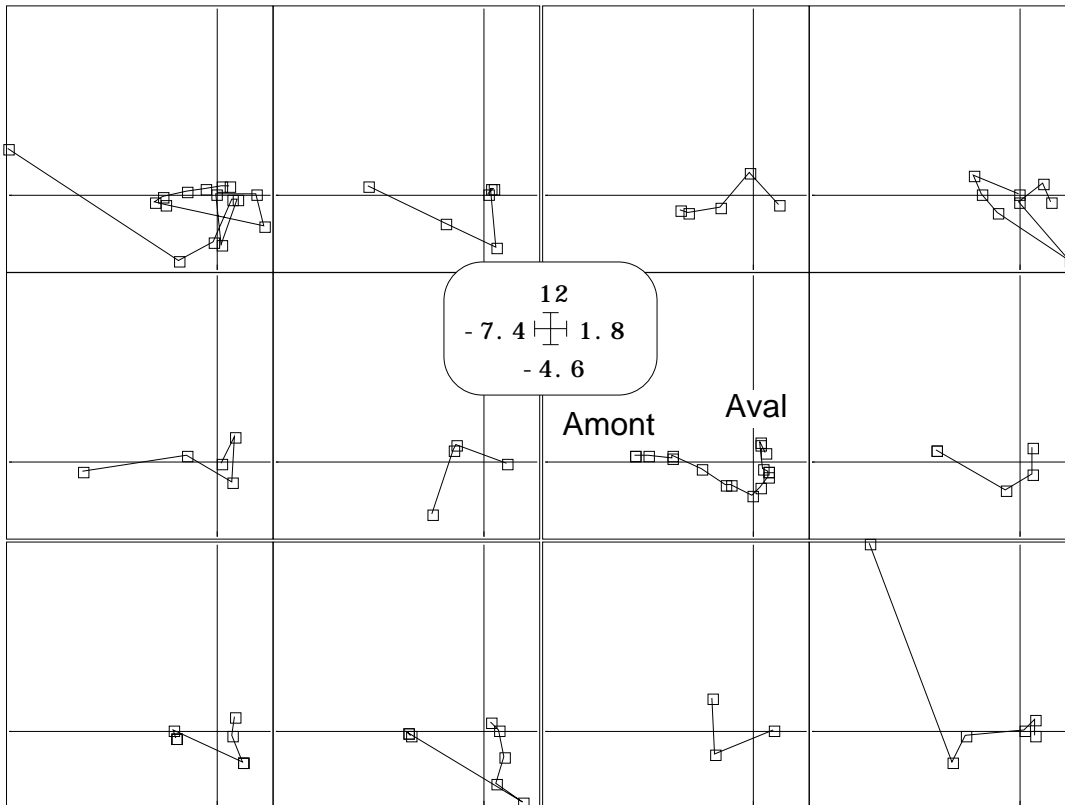
Le fichier B.--tab donne la Cuisance comme rivière de référence, l'Audeux comme rivière sans intérêt. STATIS donne des indications voisines :

Number	Rows	Weights	NS norm2	Cos2	
1	16	1.843e-01	3.966e-01	0.108	Doubs
2	6	2.189e-01	1.187e-01	0.244	Drugeon
3	5	2.943e-01	3.140e-02	0.228	Dessoubre
4	8	2.415e-01	1.747e-01	0.193	Allaine
5	5	1.717e-01	1.287e-03	0.026	Audeux
6	4	2.503e-01	6.211e-02	0.175	Cusancin
7	17	3.230e-01	3.387e-01	0.497	Loue
8	5	3.487e-01	4.424e-02	0.572	Lison
9	9	3.633e-01	2.036e-01	0.509	Furiouse
10	8	3.796e-01	5.428e-01	0.814	Cuisance
11	3	3.238e-01	2.945e-02	0.421	Doules
12	6	2.725e-01	3.043e-01	0.388	Clauge

La typologie de référence est dans le fichier B.--C1 :



Elle est triangulaire (Truite+Chabot, Vairon+Loche, Autres). La position excentrée de la Lamproie (taxon rare) est dans l'ordre des défauts connus de l'AFC. Chaque rivière cherche à la reconstituer à partir d'une carte relevés normalisée (fichier B.TL1) :



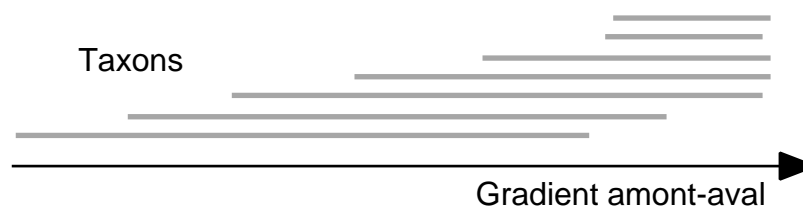
Cette approche n'est pas bonne et a les défauts caractéristiques de la représentation "position des relevés de variance 1" pour optimiser la variance des positions moyennes par espèces : les bornes sont définies par des points lointains de poids très faibles. Derrière cet artefact regrettable on reconnaît la Loue et son gradient amont-aval régulier, identique à celui de STATIS (fiche Statis 2, figure 18, p. 22).

La présence des espèces rares et des stations pauvres est négligeable dans un grand tableau faunistique. Ceux qui sont ici manipulés par rivière n'ont pas les dimensions suffisantes pour éliminer ces effets. L'ACOM ne fonctionnera bien qu'avec des analyses séparées de bonne qualité. L'ACOM peut être considérée comme une moyenne d'analyses alors que STATIS peut être considérée comme l'analyse d'une moyenne.

Comme toujours, la relation méthodes-données est affaire de résonance entre principes de lecture et structure interne des objets à lire. Un principe de lecture n'est jamais uniformément valable. Les données ici en cause n'ont pas une structure typique d'AFC, c'est-à-dire du type :



Au contraire, la richesse croît sans cesse, car la structure se résume par un modèle :



Éliminer cette augmentation de la richesse, comme un vulgaire artifice d'échantillonnage, est ici illégitime.

2.3 — Logique d'ACP

On comprend qu'il y a instabilité de l'approche par analyse séparée des différentes rivières. Cela peut provenir, soit des conditions numériques (quelques stations pour 19 espèces), soit de la logique AFC utilisée. Il est certain qu'il faut avoir de bonnes raisons pour dire qu'une station a comme poids dans l'analyse ce qu'on y trouve. Dans la thèse de J. Verneaux, il y a une volonté certaine de retrouver un modèle typologique de référence à l'intérieur d'un système perturbé par la pollution. Cela a justifié ce choix. On peut ici regarder comment la logique d'ACP fonctionne dans l'ACOM.

InitKtab			
Matrix input file		Poi	92 19
Row indicator (Default = 1 class)		blolignum	12 1
Col indicator (Default = 1 class)			
Option: output file name		cj	

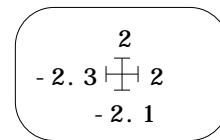
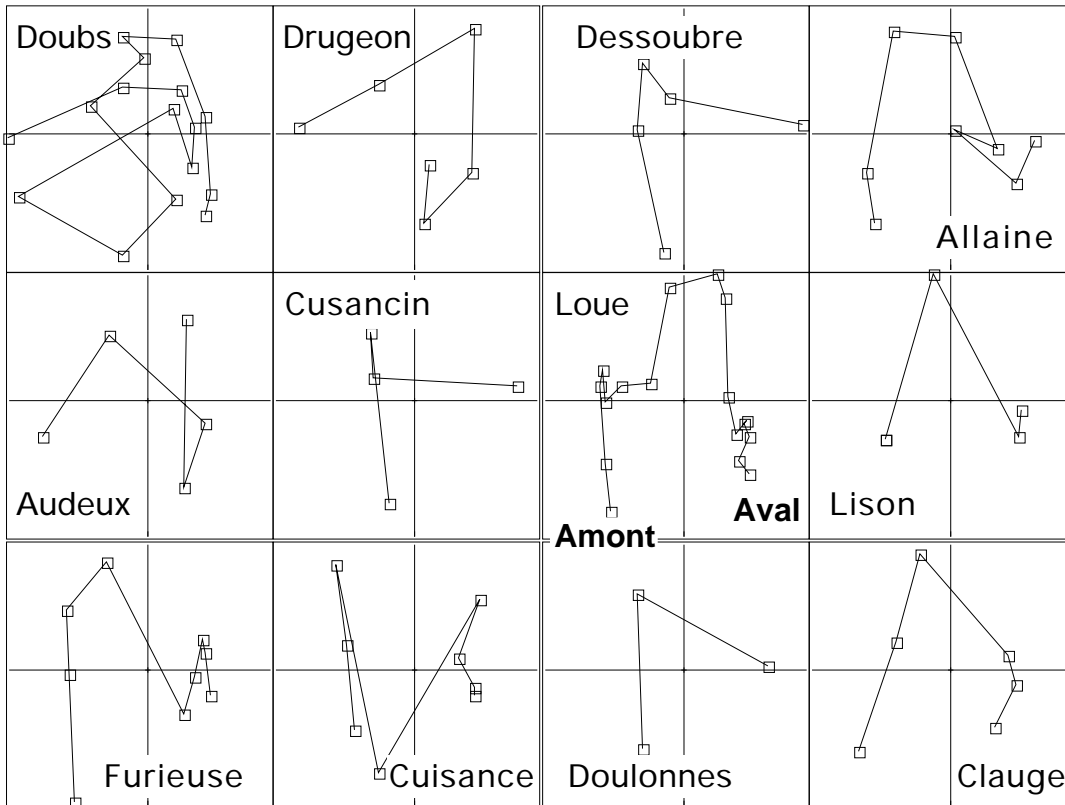
CentringKtab			
---.ktp input file		C!.ktp	
Matrix input file		Poi	92 19
Row weighting (default 1/nk)			
Column weighting (default 1/pj)			
Centring option (no default)		2	
Option: output file name			

ACOM			
---.kttta input file		Poi.kttta	92 19
Table weights (default = 1/Inertia)			
Selected axis number (default = 8)		2	

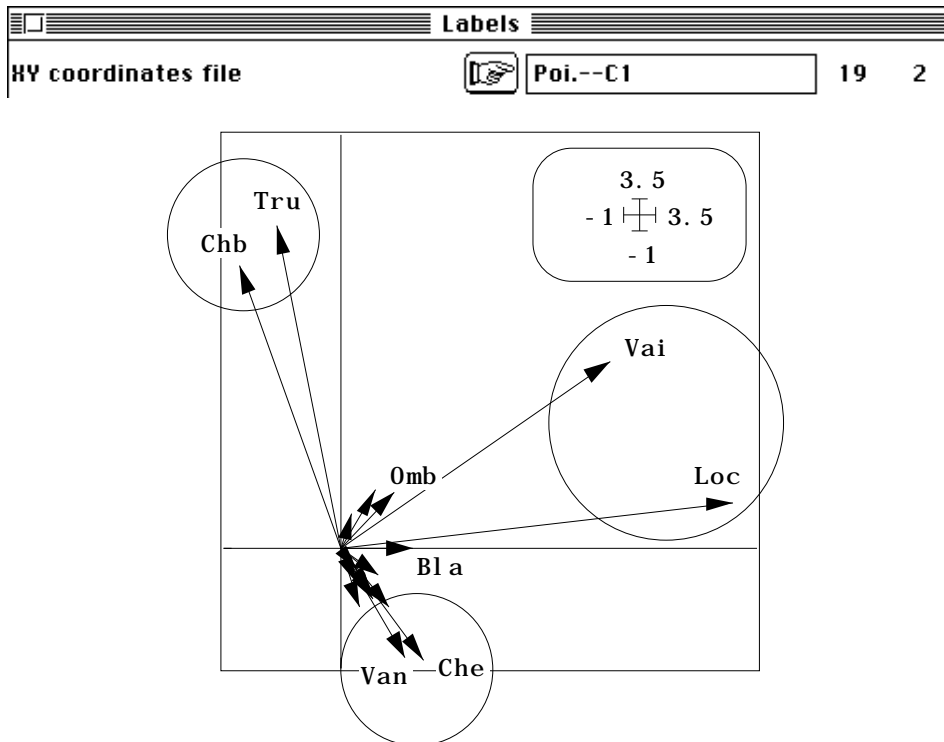
La carte des stations ne présente aucune anomalies.

Trajectories			
HY coordinates file		Poi.--TL1	92 2

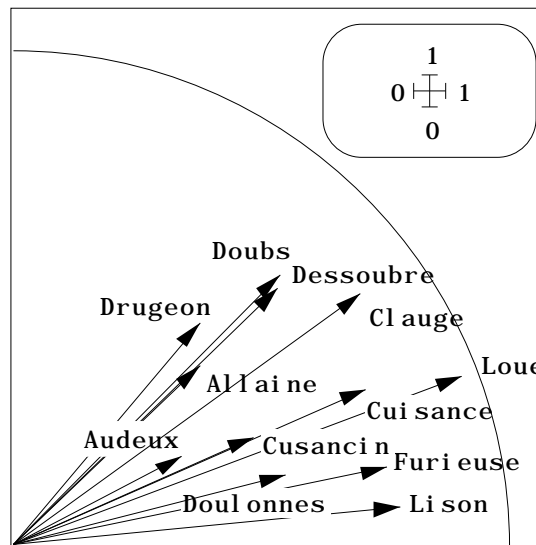
La variance des coordonnées est unitaire sur chaque axe et pour chaque rivière :



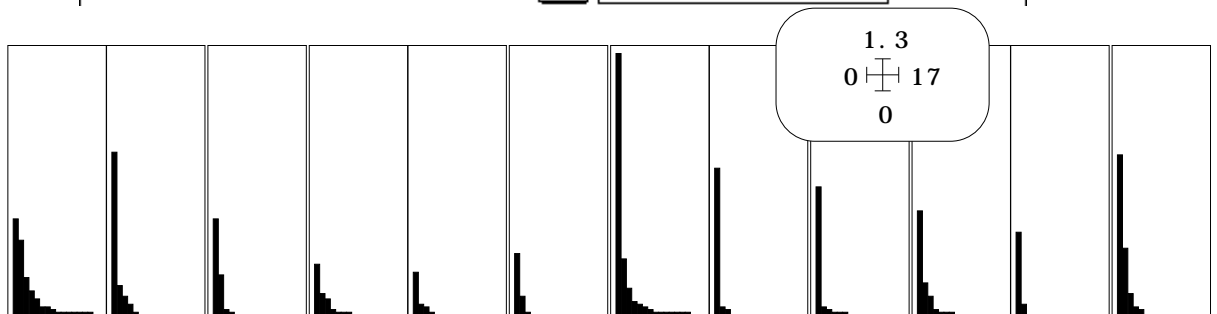
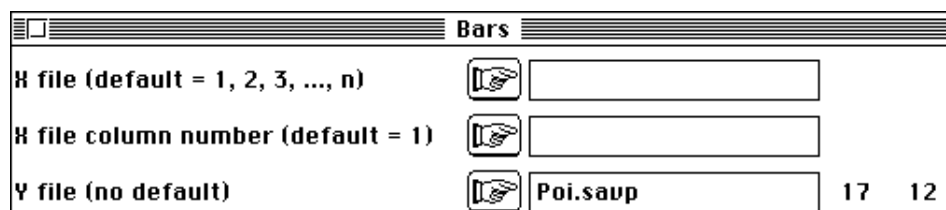
La carte de synthèse des espèces est celle de STATIS (en bas à gauche, effet *taille*) :



La reproduction de la structure par rivière est notée par :



La rivière de référence est encore la Loue, mais la tendance est à une bonne reproductibilité moyenne. Les axes 1 des ACP séparées sont souvent très marqués :



La structure de l'abondance est reproductible et se combine avec la position typologique des espèces pour donner à cette analyse une valeur très proche de celle de STATIS. Logique sur l'ensemble, dans la construction d'un modèle théorique d'ordination, l'AFC montre la faiblesse de la définition de zones spatiales effectives sur chaque rivière. Les échecs sont parfois signifiant en analyse des données. Toute l'étude porte sur la zone où le nombre d'espèces croit dans le continuum : cette structure forte⁵ minimisée par la version AFC de cette analyse est un invariant et à ce titre donne à l'analyse multitableau sa stabilité et sa pertinence. Si on ne peut réduire la structure faunistique à cet élément et éliminer la question du remplacement des espèces⁶, seule la prise en compte des deux aspects est satisfaisante.

3 — Physico-chimie : Rv et coinertie

Le troisième jeu de descripteurs, celui des paramètres physiques et chimiques de l'eau va nous permettre une dernière comparaison des deux méthodes. La procédure est toujours aussi simple.

InitKtab			
Matrix input file		Mil	92 13
Row indicator (Default = 1 class)		blolignum	12 1
Col indicator (Default = 1 class)			
Option: output file name			

CentringKtab			
---.ktp input file		Mil!.ktp	
Matrix input file		Mil	92 13
Row weighting (default 1/nk)			
Column weighting (default 1/pj)			
Centring option (no default)		2	
Option: output file name		D	

ACOM			
---.kttta input file		D.kttta	92 13
Table weights (default = 1/Inertia)			
Selected axis number (default = 8)		2	

Operator averaging			
---.kttta input file		D.kttta	92 13
1 = Option COUV (default = RV)			

On teste ici l'option RV (corrélation vectorielle) alors que la présentation dans la fiche Statis 2 (p. 23 et suivantes) utilisait l'option COVV (covariance vectorielle). La différence est sensible :

Number	Rows	Weights	NS norm2	Cos2	
1	16	3.726e-01	1.405e+00	0.293	Doubs
2	6	3.359e-01	3.997e-01	0.314	Drueon
3	5	1.611e-01	9.009e-02	0.009	Dessoubre
4	8	2.400e-01	7.901e+00	0.503	Allaine
5	5	2.854e-01	1.025e-01	0.137	Audeux
6	4	2.657e-01	4.435e-01	0.108	Cusancin
7	17	3.708e-01	7.328e-01	0.309	Loue
8	5	3.029e-01	6.330e-02	0.169	Lison
9	9	2.105e-01	7.611e+00	0.279	Furieuse
10	8	3.844e-01	1.673e-01	0.571	Cuisance
11	3	1.968e-01	4.044e-01	0.072	Doulonnes
12	6	2.296e-01	1.426e+00	0.091	Clauge

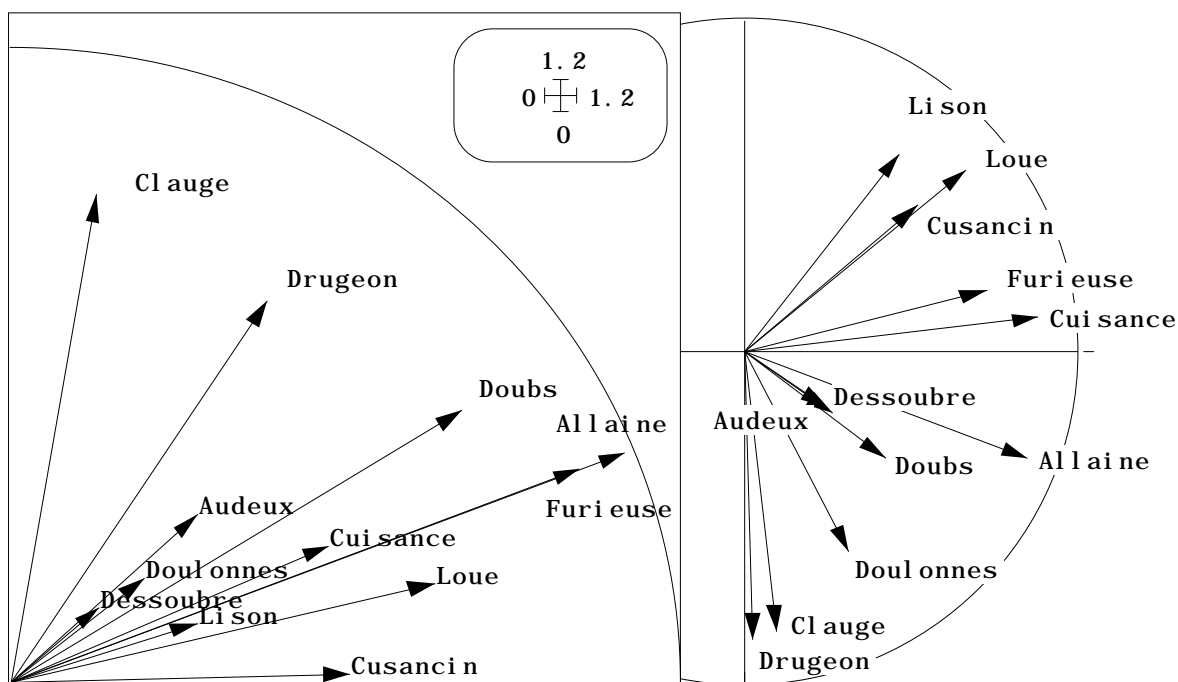


Figure 5 : A gauche, mesure de costructure vue par l'ACOM. Allaine, Furieuse et Doubs (points de pollution) ne présentent pas la même structure que Clauge et Drugeon. A droite, projections des axes 1 dans STATIS. L'éventail des positions indique la variété des structures physico-chimiques.

L'interprétation se fait à partir des remarques précédentes. Typologie de rivière (figure 5), typologie de variables (figure 6) et typologie de stations dans une rivière (figure 7) sont approchées par des moyens différents et, sans se contredire pas. Ici STATIS l'emporte par la lecture directe de la variabilité intra-rivière, car on n'utilise pas la représentation simultanée des variables, propriété plus pertinente dans l'ACOM.

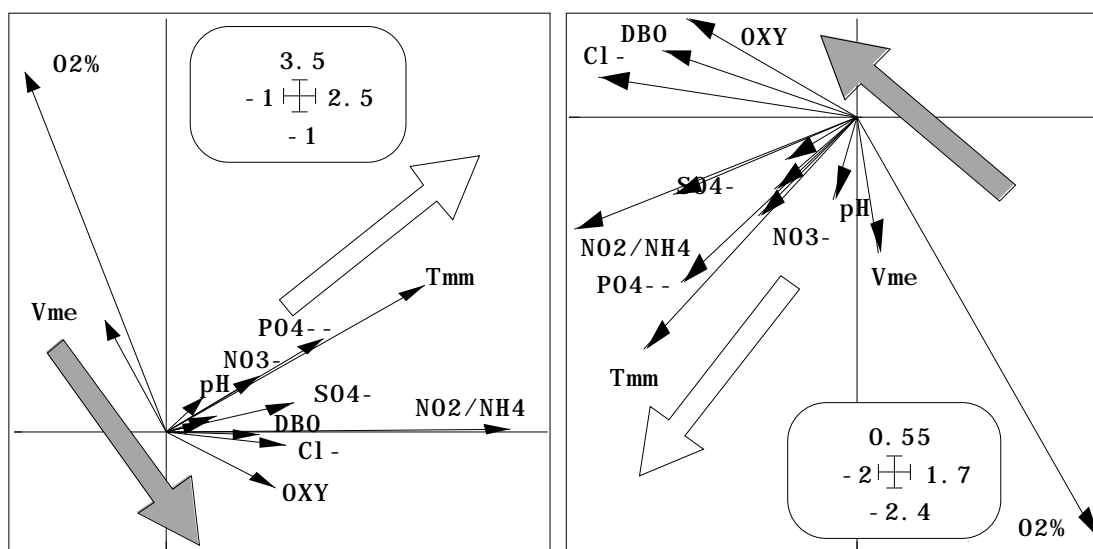


Figure 6 : A gauche, typologie de référence des variables dans l'ACOM. Flèche grise : indicateurs de pollution, flèche blanche augmentation de la température et la charge normale dans le gradient amont-aval. L'axe 1 regroupe les deux types d'évolution (pollution massive ou charge diffuse qui voit augmenter les mêmes variables de manières différentes). A droite, position des variables dans le compromis de STATIS. C'est très exactement la même structure. Moyenne de structure ou structure de moyenne conduite au même résultat.

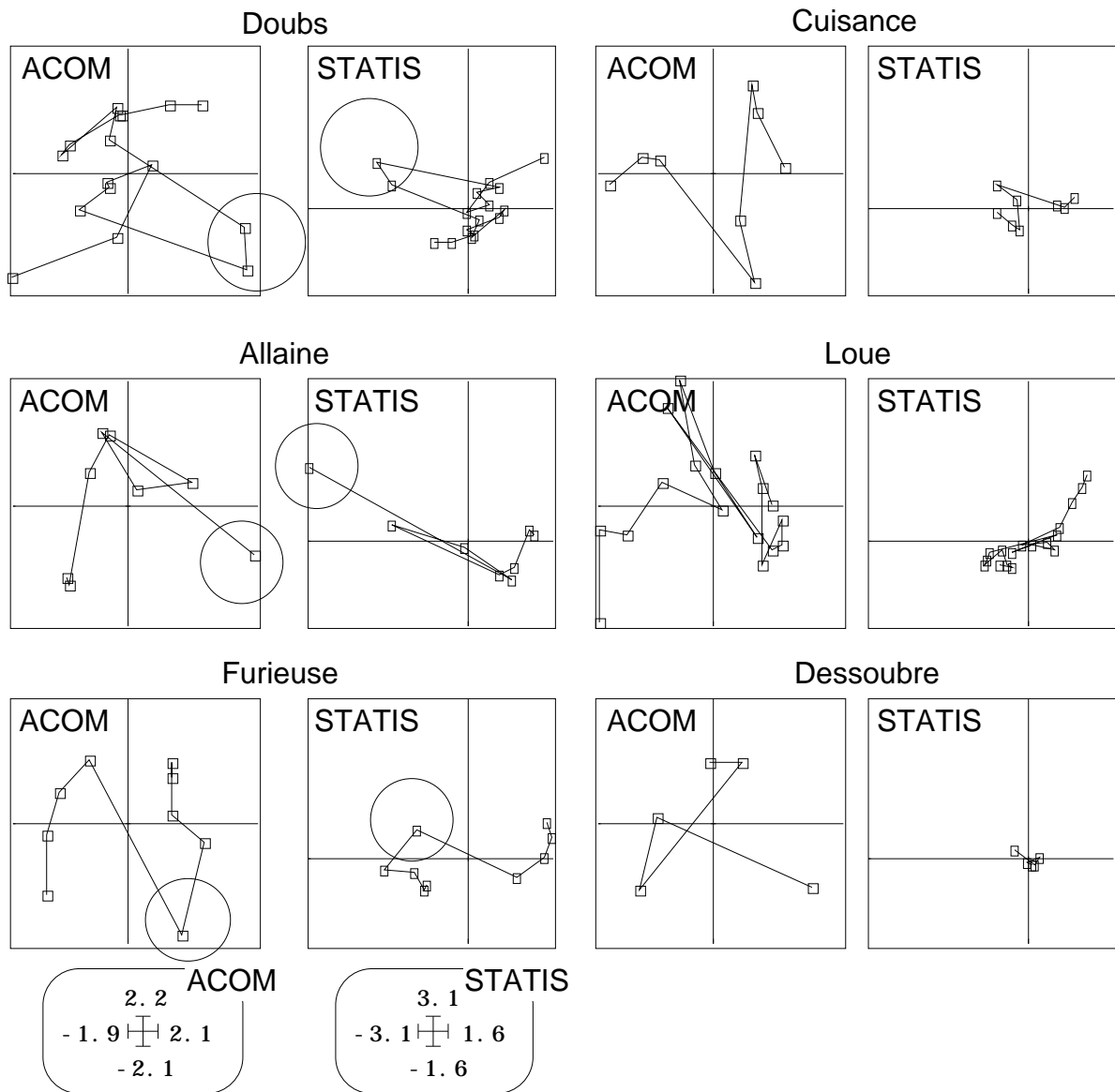


Figure 7 : sélection de quelques trajectoires des stations d'une rivière. A gauche, trois structures fortes, les cercles désignent les points de pollution. Les cartes lignes comme les cartes colonnes sont inversées mais de même signification. A droite quelques structures faibles. Les cartes de l'ACOM étant normalisées par rivière (par définition), la variabilité des positions n'a pas une signification constante. Les cartes de STATIS ne posent pas ce problème.

L'utilisation du coefficient R_v modifie sensiblement les résultats obtenus précédemment. Les variations d'inertie éliminées du calcul du compromis sont donc pertinentes. L'option R_v est proposée par défaut.

Dans les trois cas, l'ACOM n'a pas apportée d'indication qui aurait échappée à STATIS. La reproduction d'une typologie d'espèces ou de variables avec plusieurs tableaux comptant beaucoup plus de descripteurs que de stations n'est pas au centre de la problématique de l'ACOM. Nous n'avons pas mis en évidence non plus de contradiction. L'ACOM, version élaborée de l'AFMULT, partage avec STATIS une communauté d'objectif et s'en sépare par des différences de moyens mis en œuvre. Globalement, si on veut insister sur la comparaison des configurations des objets spécifiques (ici les stations par rivières), on utilisera d'abord STATIS. Si on veut insister sur la comparaison des configurations des objets communs (ici les espèces par rivière) on préférera, pour commencer, l'ACOM.

Références

¹ Prodon, R. & Lebreton, J.D. (1994) Analyses multivariées des relations espèces-milieu : structure et interprétation écologique. *Vie Milieu* : 44, 69-91.

² Esteve, J. (1978) Les méthodes d'ordination : éléments pour une discussion. In : *Biométrie et Écologie*. Legay, J.M. & Tomassone, R. (Eds.) Société Française de Biométrie, Paris. 223-250.

³ Thioulouse, J. & Chessel, D. (1992) A method for reciprocal scaling of species tolerance and sample diversity. *Ecology* : 73, 670-680.

⁴ Ter Braak, C.J.F., Juggins, S., Birks, H.J.B. & Voet, H. Van der. (1993) Weighted averaging partial least squares regression (WA-PLS): definition and comparison with other methods for species-environment calibration. In : *Multivariate Environmental Statistics*. Patil, G.P. & Rao, C.R. (Eds.) Elsevier Science Publishers. 525-560.

Ter Braak, C.J.T. & Juggins, S. (1993) Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. *Hydrobiologia* : 269/270: 485-502.

⁵ Vannote, R.L., Minshall, G.W., Cummins, K.W., Sedell, J.R. & Cushing, C.E. (1980) The River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences* : 37, 130-137.

⁶ Statzner, B. & Higler, B. (1985) Questions and comments on the River Continuum Concept. *Canadian Journal of Fisheries and Aquatic Sciences* : 42, 1038-1044.

