

Triplet d'analyses des correspondances

Résumé

La fiche décrit une procédure d'analyse à trois tableaux utilisant un tableau d'AFC comme lien entre deux tableaux quelconques. Cette analyse est appelée analyse RLQ (R pour tableau en mode R, Q pour tableau en mode Q, et L tableau du lien) est basée sur les propriétés d'optimisation de la corrélation lignes-colonnes en AFC. Dans l'exemple traité les tableaux R et Q sont des tableaux de variables qualitatives et constituent donc un cas particulier (Dolédec, S., Chessel, D., Ter Braak, C.J.F. & Champely, S. (1996) Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* : 3, 143-166).

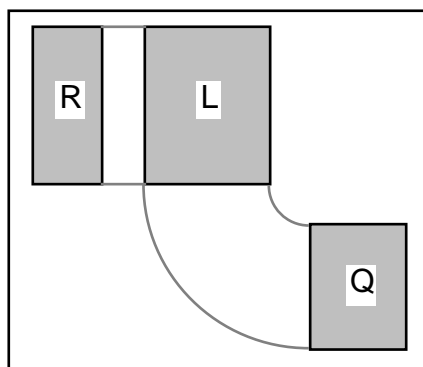
Plan

1 - Présentation du problème	2
1.1 - Identification des objets.....	2
1.2 - Tableaux de données.....	3
1.3 - La question posée.....	4
2 - L'analyse des correspondances centrale	5
2.1 - Corrélations du type xLy	5
2.2 - La corrélation intrinsèque d'une table de contingence	8
2.3 - La corrélation canonique d'une table de contingence.....	9
2.4 - Corrélations autour d'un tableau faunistique.....	11
3 - Analyse RLQ ou co-inertie généralisée.....	13
3.1 - Analyses séparées des tableaux marginaux.....	14
3.2 - Test de randomisation.....	15
3.3 - Analyse de co-inertie comme lien particulier.....	17
3.4 - Co-inertie généralisée	18
4 - Analyse RLQ et AFC sous double contraintes	30
Appendice 1 - Préparation des données	32
Appendice 2 - Analyses séparées	34
Appendice 3 - Analyses de co-inertie	42
Appendice 4 - Analyse de co-inertie généralisée.....	48
Appendice 5 - Dépouillement graphique	54
Appendice 6 - Alternatives.....	58
Références	62

D. Chessel, S. Dolédec & S. Champely

1 - Présentation du problème

Les données analysées forment un triplet de tableau. Le premier sera appelé **L** parce qu'il sert de lien aux deux autres.



Il a I lignes et J colonnes. I vaut 51 et désigne 51 relevés d'avifaune (Tatibouet et al., 1980)¹. **R** désigne le second tableau associé au précédent par les lignes. Il a I lignes et p colonnes. p vaut 11 et désigne 11 variables qualitatives décrivant l'environnement autour du relevé. **Q** est le troisième tableau associé au précédent par les colonnes. **L** est appariée dans l'ordre naturel à une seule ligne de **Q**. Comme **L** a J colonnes, **Q** a J lignes. J vaut 40 et désigne 40 espèces d'oiseaux nicheurs de l'agglomération lyonnaise. **Q** a q colonnes. q vaut 4 et désigne 4 variables qualitatives décrivant quelques traits biologiques des espèces rencontrées (Tatibouet, 1981)².

Le triplet de tableaux met donc en jeu quatre types d'objets, respectivement, les relevés (lignes de **L** et de **R**), les espèces (lignes de **Q** et colonnes de **R**), les variables environnementales (colonnes de **R**) et les variables biologiques (colonnes de **Q**).

1.1 - Identification des objets

Tableau 1 Code taxonomique des 40 espèces.

1 <i>Circus pygargus</i>	Busard cendré	21 <i>Sylvia borin</i>	Fauvette des jardins
2 <i>Falco tinnunculus</i>	Faucon crécerelle	22 <i>Sylvia atricapilla</i>	Fauvette à tête noire
3 <i>Larus ridibundus</i>	Mouette rieuse	23 <i>Phylloscopus trochilus</i>	Pouillot fitis
4 <i>Columba livia</i>	Pigeon domestique	24 <i>Phylloscopus collybita</i>	Pouillot véloce
5 <i>Streptopelia turtur</i>	Tourterelle des bois	25 <i>Farus caeruleus</i>	Mésange bleue
6 <i>Streptopelia decaocto</i>	Tourterelle turque	26 <i>Farus major</i>	Mésange charbonnière
7 <i>Cuculus canorus</i>	Coucou gris	27 <i>Inberiza caudata</i>	Bruant proyer
8 <i>Agus agus</i>	Martinet noir	28 <i>Inberiza ciris</i>	Bruant zizi
9 <i>Alauda arvensis</i>	Alouette des champs	29 <i>Passer domesticus</i>	Moineau domestique
10 <i>Hirundo rustica</i>	Hirondelle de cheminée	30 <i>Passer montanus</i>	Moineau friquet
11 <i>Helictes urbica</i>	Hirondelle de fenêtre	31 <i>Iringilla coelebs</i>	Pinson des arbres
12 <i>Motacilla flava</i>	Bergeronnette printanière	32 <i>Carduelis chloris</i>	Verdier
13 <i>Troglodytes troglodytes</i>	Troglodyte	33 <i>Carduelis carduelis</i>	Chardonneret élégant
14 <i>Susicola rubetra</i>	Traquet tairier	34 <i>Carduelis cannabina</i>	Linotte mélodieuse
15 <i>Susicola torquatus</i>	Traquet pâtre	35 <i>Serinus serinus</i>	Serin cini
16 <i>Phoenicurus ochruros</i>	Rougequeue noir	36 <i>Sturnus vulgaris</i>	Itourneau sansonnet
17 <i>Phoenicurus phoenicurus</i>	Rougequeue à front blanc	37 <i>Oriolus oriolus</i>	Loriot jaune
18 <i>Luscinia megarhynchos</i>	Rossignol philomèle	38 <i>Ficus picus</i>	Pie bavarde
19 <i>Turdus merula</i>	Merle noir	39 <i>Corvus monedula</i>	Chouca des tours
20 <i>Hippoboscus polyglotta</i>	Hypobosc polyglote	40 <i>Corvus corone</i>	Corneille noire

Les 40 espèces sont nommées dans le Tableau 1. Les 51 relevés occupent dans l'espace une position décrite dans la Fig. 1. Ils forment un double transect qui part de la zone périurbaine rurale, traverse les villages de l'est lyonnais, parvient au centre ville et ressort de l'agglomération vers le sud. Ils sont simplement numérotés de 1 à 51.

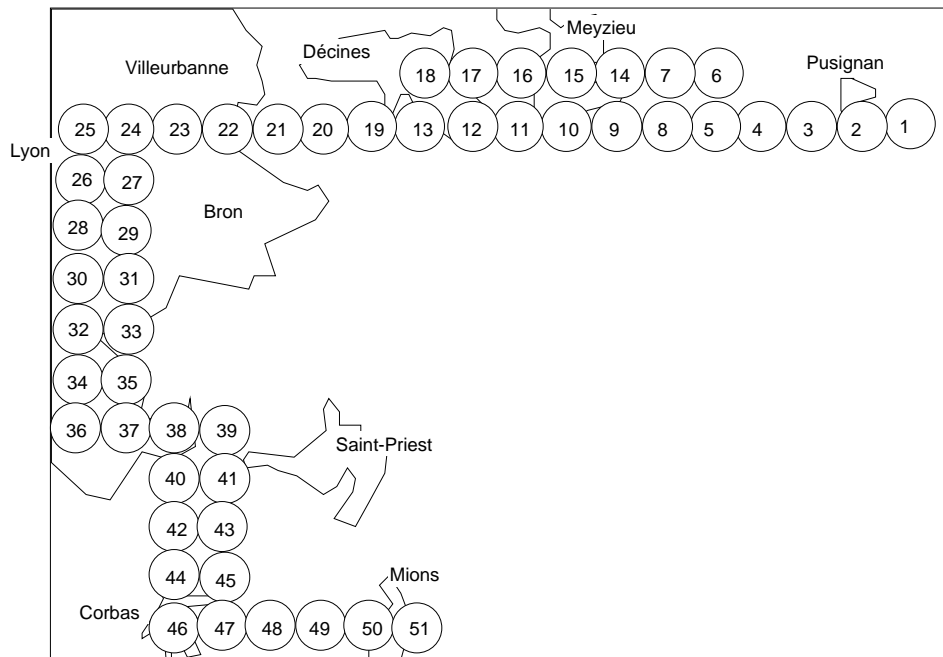


Figure 1 Disposition spatiale de 51 relevés d'avifaune.

Les 11 variables environnementales sont qualitatives. Elles décrivent le paysage autour du relevé. Les 10 premières codent la présence d'éléments caractéristiques (1 avec, 2 sans) :

- | | |
|------------------------------|--------------------|
| 1 Fermes-Villages | 6 Pelouses |
| 2 Petits immeubles | 7 Haies-arbustes |
| 3 Immeubles de + de 4 étages | 8 Feuillus |
| 4 Chantier/usines/gravières | 9 Conifères |
| 5 Champs | 10 Bruit permanent |

La dernière donne le recouvrement de la végétation naturelle avec le code :

- | | |
|--------------|--------------------------------------|
| 1- R = 100 % | (Zone de végétation naturelle) |
| 2- 97 | R < 100 (Périurbain agricole) |
| 3- 90 | R < 97 (Périurbain résidentiel) |
| 4- 75 | R < 90 (Banlieue ouverte) |
| 5- 50 | R < 75 (Banlieue fermée) |
| 6- 25 | R < 50 (Quartiers) |
| 7- 10 | R < 25 (Centre ville) |
| 8- 0 | R < 10 (Zone industrielle, carrière) |

Les quatre variables biologiques donnent respectivement :

- 1- Position trophique (1-insectivore, 2-granivore, 3-polyphage)
- 2- Site d'alimentation (1-sol, 2-air, 3-buissons ou frondaisons)
- 3- Site de nidification (1-sol, 2-bâtiment, 3-buissons, 4-frondaison)
- 4- Statut migratoire (1-sédentaire, 2-migrants)

On connaît également le poids moyen des individus de chaque espèce (en grammes).

1.2 - Tableaux de données

Le Tableau 2 contient les données traitées.

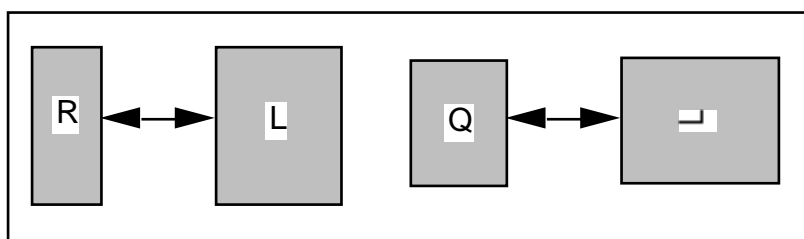
1.3 - La question posée

La question posée à partir de ces données dérive directement de leur structure très forte :

Y-a-t'il dans la variation des traits biologiques des espèces des correspondances avec la variation des caractéristiques de l'environnement ?

Tout tableau floro-faunistique appelle explicitement ou non de l'information externe sur ses deux marges. Sur la marge relevés, on a la problématique classique faune/milieu (indicateurs biologiques, déterminisme écologique). Sur la marge espèces, est virtuellement présente la grande quantité d'informations souvent disponible sur la biologie et l'écologie des groupes étudiés (Statzner et al., 1994)³.

On reconnaît alors dans les données deux problèmes de co-structures :



(1) Le premier relie la typologie de relevés induite par les variations de l'abondance des espèces avec la typologie des relevés induite par les variations des variables mésologiques. Des centaines d'articles d'écologie statistique sont consacrées depuis cinquante ans à cette question.

(2) Le second relie la typologie des espèces induites par la diversité de leurs modes de distribution entre les relevés et la typologie des mêmes espèces faite par la diversité des choix stratégiques mis en œuvre pour assurer leur pérennité. Entre les deux problèmes on reconnaît la partie commune, c'est-à-dire la dualité entre typologie d'espèces et typologie de relevés inhérente à tout tableau écologique.

En ignorant les tableaux marginaux, on peut hésiter sur la méthode à utiliser. En n'utilisant que l'un des deux, on peut hésiter encore plus, tant la présence de **R** à gauche de **L** rend dissymétrique les deux marges. En prenant en compte les deux, pour des raisons théoriques et pratiques, il est souhaitable d'utiliser dans le couple **R-L** comme dans le couple **Q-L^t** une méthode statistique cohérente (**L^t** est la matrice transposée de **L**).

De ce fait, il y a, sous-jacente à la question posée cinq analyses de base : les trois typologies simples (**L** : relevés-espèces, **R** : relevés-variables mésologiques, **Q** : espèces-variables biologiques) et les deux typologies croisées (**R-L** espèces-variables mésologiques par l'intermédiaire des relevés, qui fonde la théorie des profils écologiques; **L-Q** relevés-variables biologiques par le biais des espèces, qui fonde les théories modernes de l'habitat). Nous allons montrer qu'on peut faire ces cinq analyses simultanément et rendre compte de la cohérence du total.

*Tableau 2 Triplet de matrice **R** (à gauche), **L** (au centre) et **Q** (en bas). Le mode d'impression indique la disposition des matrices dans le modèle mathématique.*

On calcule les fréquences conjointes p_{ij} , les fréquences marginales $p_{i.}$ et $p_{.j}$:

$$p_{ij} = \frac{n_{ij}}{n} \quad p_{i.} = \frac{n_{i.}}{n} \quad p_{.j} = \frac{n_{.j}}{n}$$

On note \mathbf{P} le tableau des p_{ij} , \mathbf{D}_I et \mathbf{D}_J les matrices diagonales :

$$\mathbf{D}_I = \text{Diag}(p_{1.}, \dots, p_{I.}) \quad \mathbf{D}_J = \text{Diag}(p_{.1}, \dots, p_{.J})$$

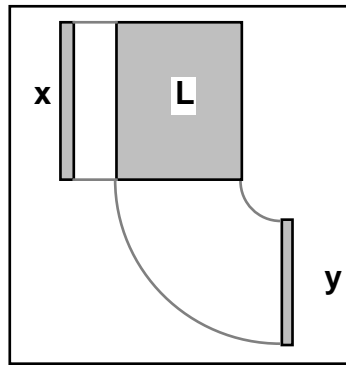
Soit alors \mathbf{Z} le tableau :

$$\mathbf{Z} = \mathbf{D}_I^{-1} \mathbf{P} \mathbf{D}_J^{-1} - \mathbf{1}_{IJ} \quad \mathbf{D}_I^{-1} = \text{Diag}(1/p_{1.}, \dots, 1/p_{I.}) \quad \mathbf{D}_J^{-1} = \text{Diag}(1/p_{.1}, \dots, 1/p_{.J})$$

Le terme général de \mathbf{Z} s'écrit simplement :

$$z_{ij} = \frac{p_{ij}}{p_{i.} p_{.j}} - 1 = \frac{p_{ij} - p_{i.} p_{.j}}{p_{i.} p_{.j}}$$

Considérons alors la situation :



où \mathbf{x} est un vecteur à I composantes et \mathbf{y} est un vecteur à J composantes.

Notons encore \mathbf{P}_0 le tableau :

$$\mathbf{P}_0 = \mathbf{P} - \mathbf{D}_I \mathbf{1}_{IJ} \mathbf{D}_J$$

Le terme général de \mathbf{P}_0 s'écrit simplement :

$$p_{ij0} = p_{ij} - p_{i.} p_{.j}$$

Le vecteur \mathbf{x} , code numérique des lignes a pour variance (associée à la pondération marginale des lignes) :

$$\text{Var}(\mathbf{x}) = \mathbf{x}^t \mathbf{D}_I \mathbf{x} - \mathbf{x}^t \mathbf{D}_I \mathbf{1}_{IJ} \mathbf{D}_I \mathbf{x}$$

Le vecteur \mathbf{y} , code numérique des colonnes a pour variance (associée à la pondération marginale des colonnes) :

$$\text{Var}(\mathbf{y}) = \mathbf{y}^t \mathbf{D}_J \mathbf{y} - \mathbf{y}^t \mathbf{D}_J \mathbf{1}_{JJ} \mathbf{D}_J \mathbf{y}$$

Les vecteurs \mathbf{x} et \mathbf{y} ont une covariance associée à \mathbf{L} :

$$\text{Cov}(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t \mathbf{P}_0 \mathbf{y}$$

La corrélation entre x et y à travers le tableau P vaut :

$$Cor(\mathbf{x}, \mathbf{y}) = \frac{Cov(\mathbf{x}, \mathbf{y})}{\sqrt{Var(\mathbf{x})}\sqrt{Var(\mathbf{y})}}$$

Examinons l'exemple suivant d'une remarquable simplicité apparente. Legay & Pontier (1983)⁴ ont noté l'âge et la fécondité (nombre de chatons produits dans l'année) pour 350 chattes domestiques (Tableau 3).

Tableau 3 Répartition de 350 chattes en fonction de l'âge (1 an à 8 ans et plus) et du nombre de chatons produits dans l'année 1 ou 2 (1.5), 3 ou 4 (3.5), ..., 13 ou 14 (13.5).

	0	1 ou 2	3 ou 4	5 ou 6	7 ou 8	9 ou 10	11 ou 12	13 et +
1 an	8	15	44	11	7	4	0	0
2 ans	6	12	36	21	11	6	1	1
3 ans	4	7	18	13	12	4	0	2
4 ans	2	8	7	3	7	5	1	0
5 ans	2	3	5	3	4	6	0	0
6 ans	1	0	5	3	2	2	0	1
7 ans et +	2	2	8	3	12	8	1	1

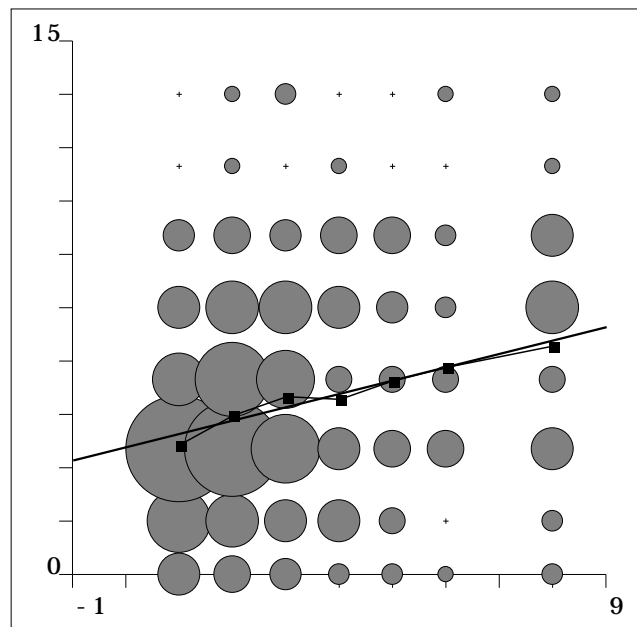


Figure 2 La liaison entre les deux variables est exprimée par le nuage bivarié élémentaire (surface des cercles proportionnelle au nombre d'observations), la courbe de régression (rapport de corrélation de 0.082) et la droite de régression (carré de corrélation de 0.0767). On y "voit" une liaison linéaire comme dans un cas d'école de statistique descriptive.

Il est question d'étudier la fécondité en fonction de l'âge, une augmentation simultanée étant un cas fréquent chez les mammifères. Utilisons les statistiques descriptives élémentaires sur la Fig. 2 où sont placées en abscisse les classes d'âge et en ordonnée les centres des classes de fécondité. A chaque case correspond un cercle de diamètre proportionnel à l'effectif. De plus la droite et la courbe de régression qui relie les points représentatifs des moyennes de la variable y (fécondité) sachant la valeur de la variable x (âge) sont représentées. Notons que l'on est directement dans la situation habituelle par l'écriture :

i	x_i	y_i	ψ_i
1	1	0	8/350
2	1	1.5	15/350
\vdots	\vdots	\vdots	\vdots
55	8	11.5	1/350
56	8	13.5	1/350

La variable x est considérée soit comme une variable quantitative (1 an, 2 ans, etc.) soit comme une variable qualitative (classe 1, classe 2, etc.). On obtient pour x une moyenne de 3.097 et une variance de 4.636. Pour y ces paramètres valent respectivement 4.747 et 8.599. La covariance entre x et y vaut 1.749, la corrélation est de 0.277 soit un carré de 0.0767. Le rapport de corrélation vaut :

$$\eta_{yx}^2 = 0.082 .$$

En dépit du caractère certainement erroné des hypothèses nécessaires, et en invoquant la notion de robustesse, on pourrait tester le coefficient de corrélation et dire qu'il est significatif au risque de 10^{-3} . La faible différence entre $r^2(x,y)$ et η_{yx}^2 et l'écart modéré entre ligne et droite de régression devrait témoigner, en première approche, d'une liaison linéaire entre les deux variables.

2.2 - La corrélation intrinsèque d'une table de contingence

L'exercice qui précède soulève une question des plus étonnantes. Si nous ne connaissions pas la nature des lignes et des colonnes du Tableau 3, mais que le descriptif se réduise à l'assertion : 350 chattes ont été classées selon deux critères. Le premier a 7 modalités et le second en a 8.

Peut-on dire que les deux variables qualitatives qui ont construit le tableaux sont corrélées ? La solution est bien connue. Si les deux variables ne sont pas corrélées, la quantité :

$$w = \frac{\sum_{i=1}^I \sum_{j=1}^J \left(n_{ij} - \frac{n_{i.}n_{.j}}{n} \right)^2}{\sum_{i=1}^I \sum_{j=1}^J \frac{n_{i.}n_{.j}}{n}} = \frac{\sum_{i=1}^I \sum_{j=1}^J (Obs. - Mod.)^2}{Mod.}$$

suit un χ^2 à $(I-1)(J-1)$ degrés de liberté. On trouve ici une valeur de 67.409 pour 42 degrés de liberté, valeur qui est dépassée sous l'hypothèse nulle avec une probabilité de 6×10^{-3} .

On pourra encore arguer du fait que les conditions d'approximation ne sont pas satisfaites. Un test de simulation peut rassurer. Il consiste à construire 1000 réalisations du modèle dans lequel 350 observations sont appariées au hasard.

```

number of random matching: 1000   Observed: 0.192598
Histogramm:   minimum = 0.057060, maximum = 0.242129
number of simulation X<0bs: 986 (frequency: 0.986000)
number of simulation X>=0bs: 14 (frequency: 0.014000)

```


- Les moyennes marginales sont nulles (les coordonnées sont centrées pour les distributions marginales);

- Les points représentatifs des moyennes conditionnelles sont alignés: droite et courbe de régression sont confondues. Ce serait également le cas si on étudiait la prédiction de y (fécondité) par x (âge), ce qui n'a pas de signification expérimentale mais qui renvoie à la plus ancienne définition mathématique connue de l'AFC, solution donnée par Hirschfeld (1935)⁶ à la question:

"Étant donnée une distribution discrète p_{ij} est-il possible d'introduire de nouvelles valeurs pour x_i et y_j de telle manière que les deux régressions soient linéaires?" (in Nishisato (1980)⁷: p.13).

- Les variances marginales de x et y , le rapport de corrélation de y en x (et celui de x en y), le carré du coefficient de corrélation de x et y sont tous égaux à $\frac{1}{2}$ (0.1123). La différence numérique avec le même calcul sur les données de base est bien mince mais la figure est totalement différente.

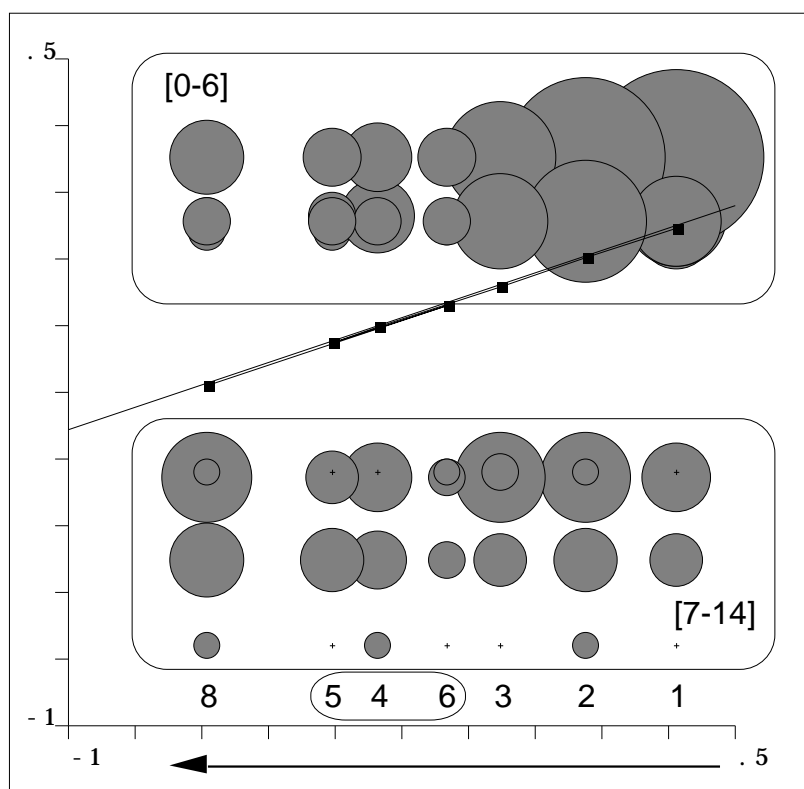


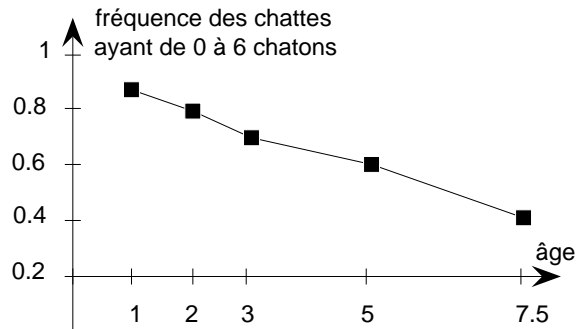
Figure 3 Graphe canonique d'une table de contingence vue par la première coordonnée de son analyse des correspondances. En abscisse, position des colonnes par la coordonnée de rang 1. En ordonnée position des lignes par la coordonnée de rang 1.

On peut noter que :

- Le codage ligne est sensiblement (à une transformation linéaire près) l'âge mesuré. Les rares inversions invite à regrouper pour assurer une précision constante par classe et on gardera les classes 1 an, 2 ans, 3 ans, 4 à 6 ans, 7 ans et plus.

- Le codage colonne impose deux groupes respectivement de 0 à 6 chatons et de 7 à 14 chatons. Il y a là un exemple remarquable d'une recherche de liaison doublement linéaire qui renvoie à une ordination sur une variable et à une partition sur une autre, image caractéristique de la pensée de Benzécri (1969)⁸ pour qui les modèles doivent émerger des données, et non l'inverse.

Ne retenons de l'analyse que la simplification du tableau qu'elle nous propose :



Il existe deux classes de chattes et la proportion de la première décroît continuellement avec l'âge : ce pourrait être alors le nombre de portées par an qui caractériserait correctement l'évolution de la fécondité. Pour affiner l'interprétation il conviendra d'analyser la liaison trivariée âge-nombre de portées-nombre de chatons, ce qui renvoie à l'analyse des correspondances multiples. Nous retiendrons que la première fonction de l'AFC est celle de la description d'une corrélation par double codage (*dual scaling*) et qu'il peut être utile de recoder numériquement des valeurs qualitatives pour apprécier correctement une relation bivariée.

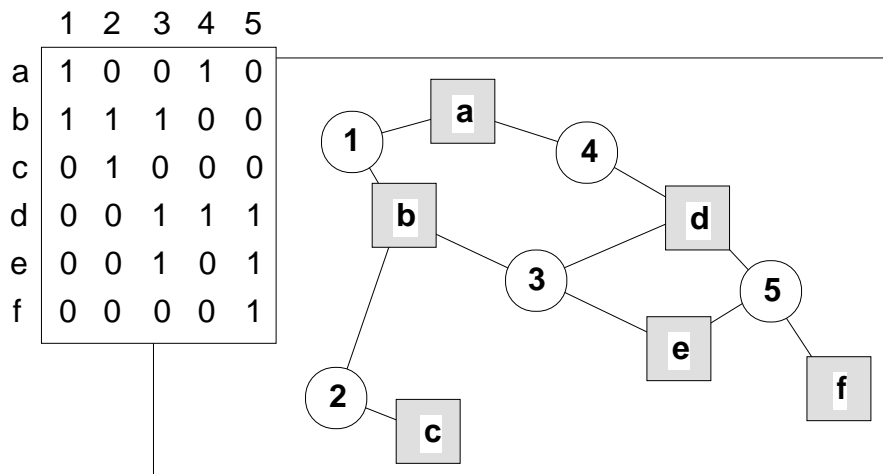


Figure 4 Deux ensembles de points de mesures sont reliés par un lien (ici un graphe de voisinage). Le tableau du lien définit une corrélation entre deux variables mesurées l'une dans les cercles et l'autre dans les carrés.

2.4 - Corrélation autour d'un tableau faunistique

Il importe de noter que, dans le cas général, le tableau central d'un triplet statistique n'est pas une table de contingence et que la notion inférentielle de Khi^2 d'une table de contingence n'est plus valide. Dans une case du tableau contenant l'entier m , m ne désigne plus m individus mais une intensité numérique du lien entre la ligne i et la colonne j qui définissent la case en question. Rien n'empêche de calculer la corrélation, mais pour en tester la signification il convient de changer de raisonnement. On supposera que le tableau \mathbf{L} est un lien fixé entre deux ensembles de mesures. Le cas le plus clair de cet état de lien fixe est représenté sur la Fig. 4.

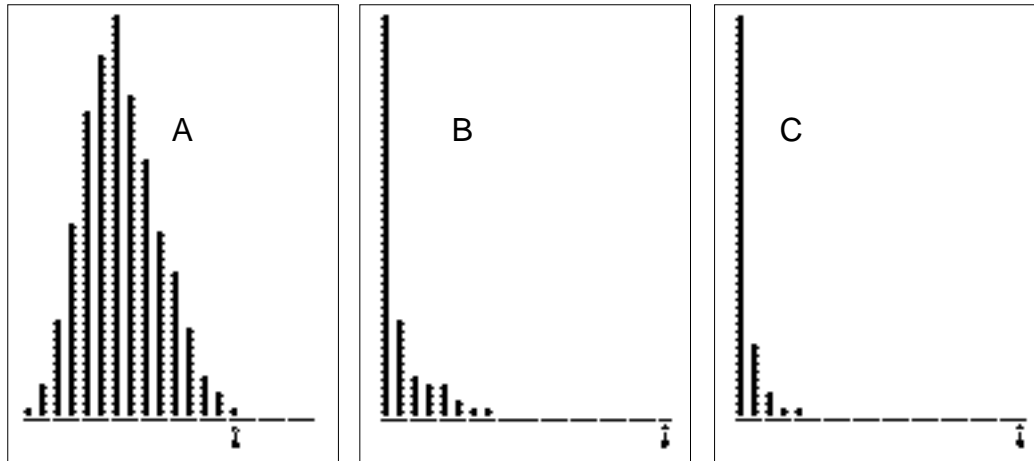


Figure 5 Trois tests de permutation pour mesurer la corrélation entre deux variables définies sur les marges d'une table de contingence. A - Simulation de l'indépendance des marges sans introduction des mesures, équivalent au test χ^2 . B - Permutation aléatoire des valeurs sur les marges, la table et les distributions marginales étant fixées (double indice de Geary et modèle R). C - Simulation de l'indépendance des marges avec introduction des mesures (équivalent au test du coefficient de corrélation de Pearson). Le test B sera étendu au cas doublement multivarié par le test sur la trace.

A partir du tableau **L** défini par une relation de voisinage entre deux ensembles de points on peut définir une corrélation spatiale qui étend la notion d'autocorrélation et l'indice de Geary (1954)⁹.

Le test non paramétrique portera sur l'espace des $I!$ permutations des mesures sur les lignes de **L** et celui des $J!$ permutations des mesures sur les colonnes. Par exemple sur l'exemple des chats qui précède on obtient :

```

number of random matching: 500   Observed: 0.076711
Histogramm:  minimum = 0.000000, maximum = 0.076711
number of simulation X<Obs: 500 (frequency: 1.000000)
number of simulation X>=Obs: 0 (frequency: 0.000000)

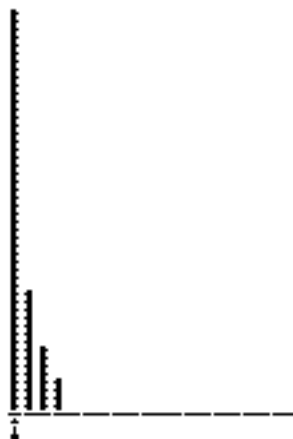
```

Le résultat est reporté dans la Fig. 5B qui résume les trois points de vue. Par exemple on note qu'il n'y a pas de corrélation significative entre le poids des oiseaux et le recouvrement de la végétation :

```

number of random matching: 500   Observed: 0.000096
Histogramm:  minimum = 0.000000, maximum = 0.023735
number of simulation X<Obs: 101 (frequency: 0.202000)
number of simulation X>=Obs: 399 (frequency: 0.798000)

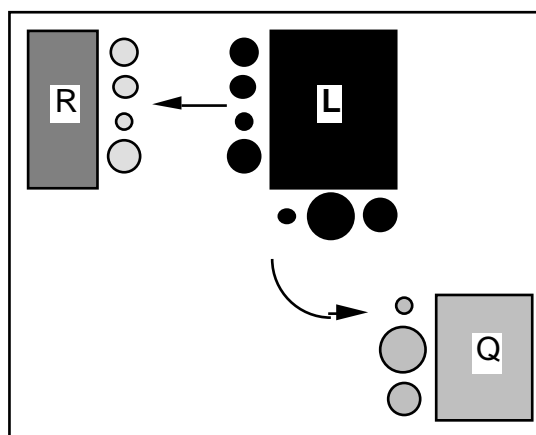
```



Ceci introduit à la nécessité d'un test de signification globale de la corrélation entre le tableau **R** et le tableau **Q** à travers le tableau **L**. Si **L** est fixé et si $Cov(\mathbf{x}, \mathbf{y}) = \mathbf{x}^t \mathbf{P}_0 \mathbf{y}$ mesure la covariance entre deux variables, alors on pourra mesurer la covariance entre deux tableaux par l'inertie du tableau $\mathbf{R}^t \mathbf{P}_0 \mathbf{Q}$. Ceci nécessite de définir une analyse de co-inertie généralisée qu'on appellera *RLQ-analysis* ou analyse RLQ¹⁰.

3 - Analyse RLQ ou co-inertie généralisée

Dans tout ce qui suit, le tableau **L** qui donne la matrice **P** elle-même transformée en **P**₀ est considérée comme un lien fixé définitivement entre les lignes de **R** et celles de **Q**. Cela signifie en particulier que les poids des lignes de **R** sont fixés comme étant les poids des lignes de **L** et que les poids des lignes de **Q** sont fixés comme étant les poids des colonnes de **L**. L'ensemble de la procédure de calcul se trouve dans les appendices.



Le lien impose donc ses pondérations aux deux tableaux marginaux. Le tableau **R** muni de la pondération **D**_I est centré en **R**₀. Le tableau **Q** muni de la pondération **D**_J est centré en **Q**₀. On ajoute une métrique quelconque de \mathbb{R}^p notée **D**_p et une métrique quelconque de \mathbb{R}^q notée **D**_q pour obtenir deux triplets (**R**₀, **D**_p, **D**_I) et (**Q**₀, **D**_q, **D**_J). Dans l'exemple étudié, le passage aux tableaux effectifs des analyses marginales modifie les dimensions, puisqu'il s'agit de variables qualitatives. *p* vaut 28 car les 11 variables de départ compte ensemble 28 modalités. *q* vaut 12 car les 4 variables de départ compte ensemble 12 modalités. Les analyses marginales sont des analyses des correspondances multiples pondérées par les marges du lien, ici le tableau faunistique à 51 relevés et 40 espèces. Les pondérations imposées des lignes se reportent sur les pondérations des colonnes des tableaux marginaux dans la logique habituelle de l'ACM.

Pour **R** on a :

```
Variable number 1 has 2 categories
[1]   Category:  1 Weight:  0.473
[2]   Category:  2 Weight:  0.527

Variable number 2 has 2 categories
[3]   Category:  1 Weight:  0.248
[4]   Category:  2 Weight:  0.752

Variable number 3 has 2 categories
[5]   Category:  1 Weight:  0.146
[6]   Category:  2 Weight:  0.854
..... etc.
Variable number 10 has 2 categories
```

```

[19] Category: 1 Weight: 0.15
[20] Category: 2 Weight: 0.85

Variable number 11 has 8 categories
[21] Category: 1 Weight: 0.0862
[22] Category: 2 Weight: 0.156
[23] Category: 3 Weight: 0.163
[24] Category: 4 Weight: 0.103
[25] Category: 5 Weight: 0.114
[26] Category: 6 Weight: 0.149
[27] Category: 7 Weight: 0.102
[28] Category: 8 Weight: 0.126

total inertia: 1.5455

```

Pour **Q** on trouve :

```

Variable number 1 has 3 categories
[1] Category: 1 Weight: 0.438
[2] Category: 2 Weight: 0.446
[3] Category: 3 Weight: 0.117
.....etc.
Variable number 4 has 2 categories
[11] Category: 1 Weight: 0.676
[12] Category: 2 Weight: 0.324

total inertia: 2

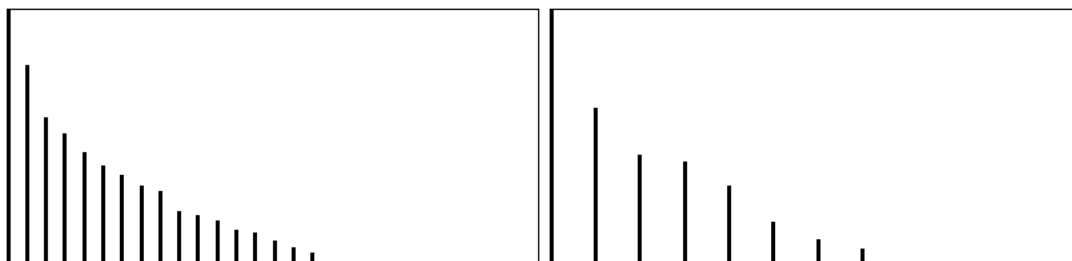
```

Dans ce type d'analyse, les inerties totales sont définies par les dimensions des tableaux, à savoir nombre de modalités/nombre de variables-1. On vérifie ainsi que :

$$\frac{m}{v} - 1 = \frac{28}{11} - 1 = 1.5455 \quad \text{et} \quad \frac{m}{v} - 1 = \frac{12}{4} - 1 = 2$$

3.1 - Analyses séparées des tableaux marginaux

Dans l'exemple étudié, les triplets $(\mathbf{R}_0, \mathbf{D}_p, \mathbf{D}_I)$ et $(\mathbf{Q}_0, \mathbf{D}_q, \mathbf{D}_J)$ fournissent deux analyses des correspondances multiples. On garde deux facteurs pour chaque analyse dont les graphes de valeurs propres respectifs sont les suivants :



La Fig. 6 résume le dépouillement de l'analyse du milieu. On reconnaît sans peine l'opposition ville-campagne sur le facteur 1 et la redondance des modalités *Pas de végétation* (11g) et *Usines-gravières* (4a) caractéristiques des zones industrielles péri-urbaines sur le facteur 2. La Fig. 7 permet d'interpréter l'analyse préliminaire des traits biologiques. On reconnaît, même sur un cortège faunistique réduit et un ensemble de variables élémentaire, la diversité des stratégies mises en œuvre. L'association insectivore-migrateur est la plus forte. Rien ne laisse encore deviner ce qui pourrait bien relier les Fig. 6 et 7, puisque les deux ensembles d'objets représentés (sites-modalités d'environnement, espèces-traits biologiques) sont sans point commun. Rien ne permet

d'ailleurs de dire qu'une relation est possible. L'analyse du tableau qui fait le lien est évidemment indispensable (cf. Thioulouse & Chessel, 1992)¹¹.

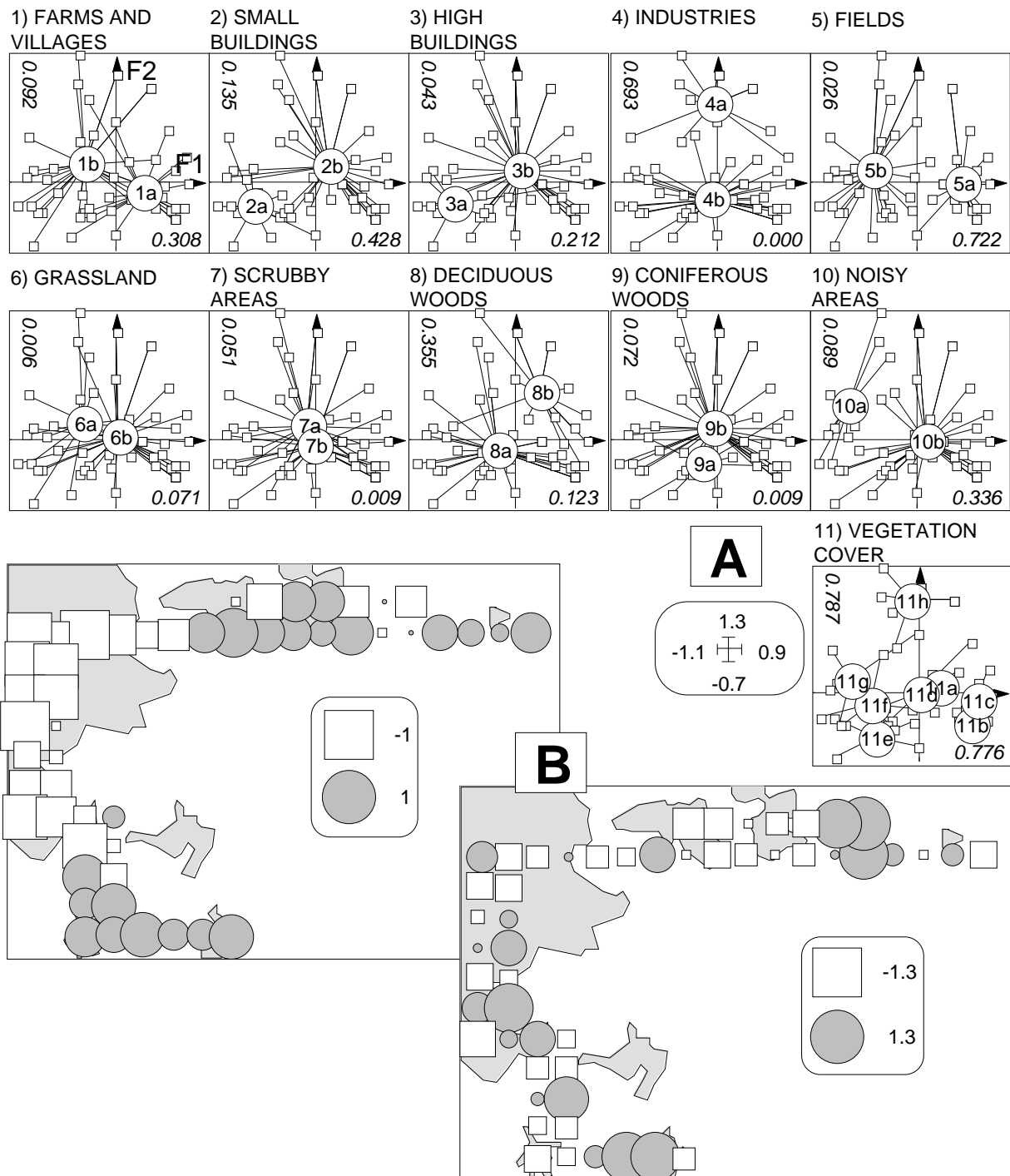


Figure 6 Dépeuplement de l'ACM sites-milieu. A - Plan factoriel 1-2 des lignes et moyennes par modalités (a: présence ; b absence; voir code pour la végétation (variable 11). B - Cartographie des deux premières coordonnées factorielles.

3.2 - Test de randomisation

Pour savoir si l'appariement modalités d'habitat-trait biologiques a une signification statistique, un test non paramétrique s'impose (Voir la revue de Potvin & Roff, 1993)¹².

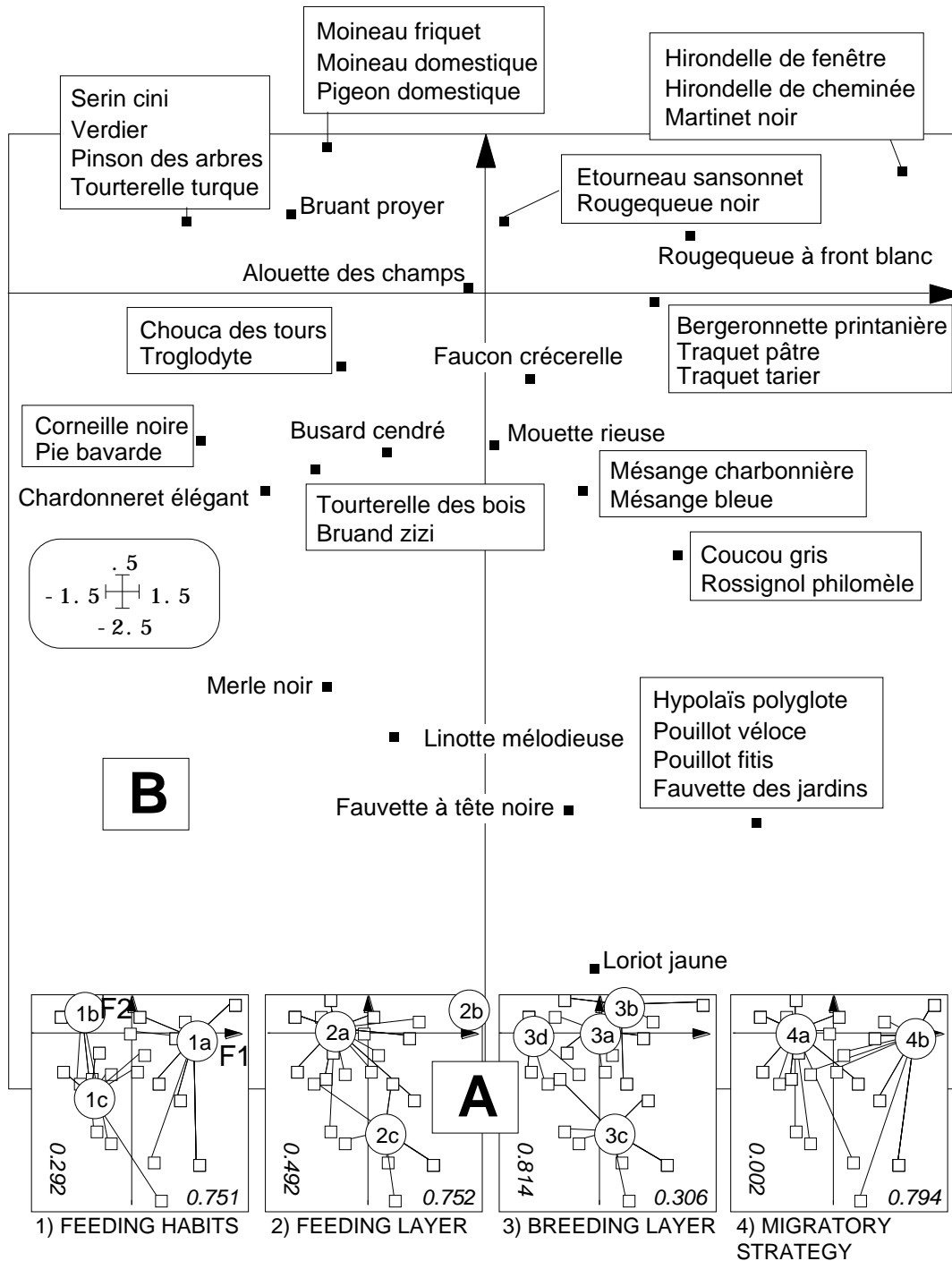


Figure 7 Dépeuplement de l'ACM espèces-stratégies. A - Plan factoriel 1-2 des lignes/espèces et moyennes par modalités (voir code d'identification). B - Plan factoriel 1-2 des lignes/espèces et moyenne par famille.

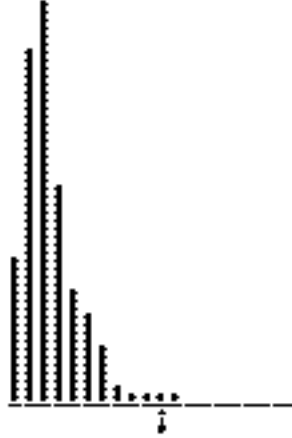
Pour une simulation on tire au sort une des $51!$ permutations possibles des lignes du tableau \mathbf{R} et une des $40!$ permutations des lignes de \mathbf{Q} . Chacun des tableaux est recentré par la pondération marginale de \mathbf{L} qui reste en place. On considère alors le triplet défini par $(\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, \mathbf{D}_p, \mathbf{D}_q)$ et la quantité définie par l'inertie totale de cette analyse, soit $z = \text{Trace}(\mathbf{Z}^t \mathbf{D}_p \mathbf{Z} \mathbf{D}_q)$ avec $\mathbf{Z} = \mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0$.

Cette quantité grandit avec l'intensité du lien entre les deux tableaux externes à travers le troisième et fait l'objet d'un test unilatéral à droite. Par exemple :

```

number of random matching: 500   Observed: 0.016542
Histogramm:  minimum = 0.002722, maximum = 0.028372
number of simulation X<0bs: 492 (frequency: 0.984000)
number of simulation X>=0bs: 8 (frequency: 0.016000)

```



Le choix de la trace est cohérent avec les articles classiques de Escoufier (1973)¹³, Robert & Escoufier (1976)¹⁴ ou récents de Kazi-Aoual et al.¹⁵ et Fraile et al. (1993)¹⁶.

On rejettera ici l'hypothèse nulle avec un risque inférieur à 5%. La liaison, si elle existe n'est cependant pas triviale et on demandera à l'analyse RLQ de préciser dans quels éléments elle s'exprime.

3.3 - Analyse de co-inertie comme lien particulier

L'analyse de co-inertie, déjà étudiée et commentée dans la documentation ADE est un cas particulier de la situation qui nous occupe et s'exprime simplement dans la Fig. 8. Ceci n'est vrai qu'à condition que la pondération commune des deux tableaux soit fixée *a priori*. La pondération uniforme s'impose alors sauf cas particulier.

C'est pourquoi nous avons utilisé l'AFC des tableaux faunistiques en pourcentages par ligne pour conserver la pondération uniforme et les propriétés de l'AFC.

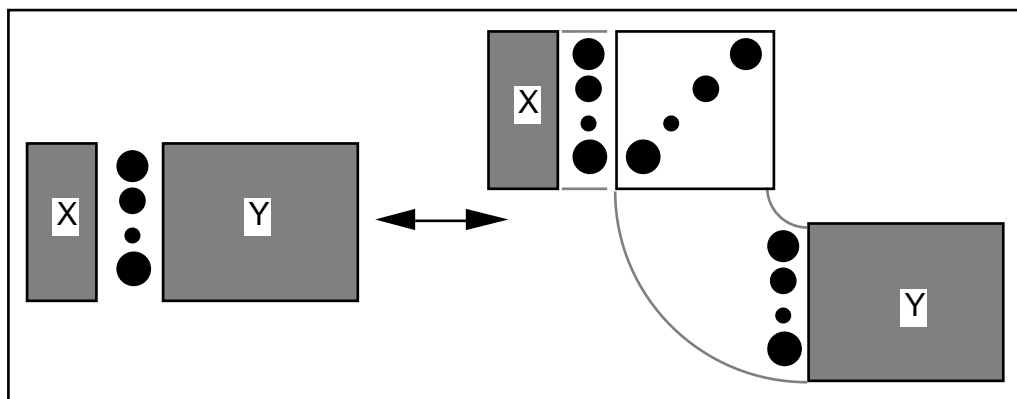


Figure 8 L'analyse de co-inertie ordinaire, à gauche, vue comme cas particulier de l'analyse d'un triplet de tableau, à droite.

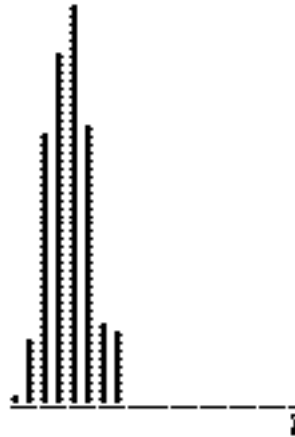
Ceci est illustré par un test de randomisation sur la trace totale de l'analyse de co-inertie de \mathbf{R} et \mathbf{L} d'une part et celle de l'analyse de co-inertie de \mathbf{Q} et \mathbf{L}^t d'une part. Dans

le premier cas on utilisera une ACM à pondération uniforme de \mathbf{R} contre une AFC sur pourcentage en ligne de \mathbf{L} dans le premier cas, une ACM à pondération uniforme de \mathbf{Q} contre une AFC sur pourcentage en ligne de \mathbf{L}^t dans le second. On trouve dans le premier cas :

```

number of random matching: 200   Observed: 0.157821
Histogramm:  minimum = 0.070700, maximum = 0.157821
number of simulation X<0bs: 200 (frequency: 1.000000)
number of simulation X>=0bs: 0 (frequency: 0.000000)

```

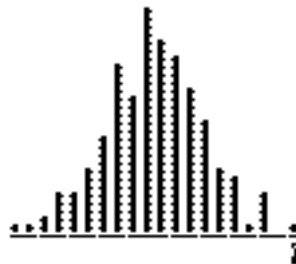


et dans le second :

```

number of random matching: 200   Observed: 0.397260
Histogramm:  minimum = 0.235525, maximum = 0.397260
number of simulation X<0bs: 200 (frequency: 1.000000)
number of simulation X>=0bs: 0 (frequency: 0.000000)

```



Les deux résultats sont significatifs, mais le lien entre variables de milieu et cortège faunistique est beaucoup plus net que le lien entre traits biologiques et mode de répartition. L'exemple montre également que le simple nombre de simulations défavorables ne suffit pas à rendre compte de toute l'information apportée par un test de randomisation. Tous les liens établis entre les trois tableaux étant significatifs, on cherche maintenant à les synthétiser dans une seule représentation.

3.4 - Co-inertie généralisée

On appelle analyse RLQ la diagonalisation du triplet $(\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, \mathbf{D}_p, \mathbf{D}_q)$. L'intérêt de cette approche dérive directement des théorèmes généraux du schéma de dualité, comme décrit, par exemple, par Escoufier (1987)¹⁷.

Les I lignes du tableau \mathbf{R}_0 forment un nuage de I points de \mathbb{R}^p muni du produit scalaire \mathbf{D}_p . Les I points sont pondérés par \mathbf{D}_I . Ce nuage a des axes d'inertie. Les J lignes du tableau \mathbf{Q}_0 forment un nuage de J points de \mathbb{R}^q muni du produit scalaire \mathbf{D}_q . Les J points sont pondérés par \mathbf{D}_J . Ce nuage a des axes d'inertie.

Faire une analyse du lien, c'est chercher un axe de R^p et un axe de R^q de manière à ce que les coordonnées des lignes de R_0 projetées sur le premier présente avec les coordonnées des lignes de Q_0 projetées sur le second une covariance du type xLy maximum (Fig. 9).

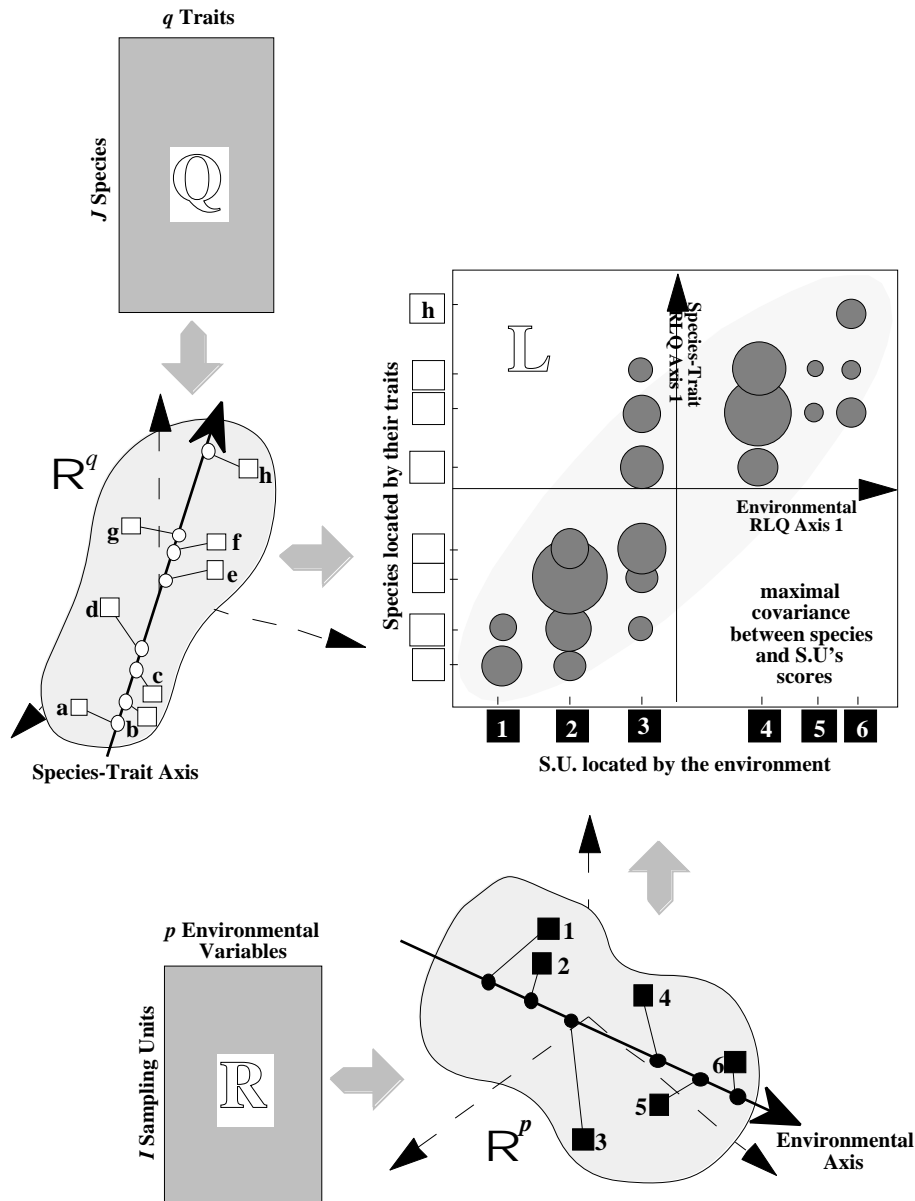
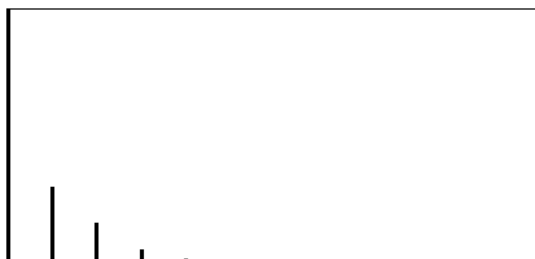


Figure 9 Schéma de principe de l'analyse RLQ.

Ces deux axes existent et sont, en général, uniques. Nous les appellerons les axes 1 de l'analyse RLQ. Perpendiculairement à chacun d'entre eux, il existe un axe de R^p et un axe de R^q maximisant à nouveau la covariance des coordonnées. Nous les appellerons les axes 2 de l'analyse RLQ. Le processus compte autant de couples d'axes que la plus petite des deux valeurs p et q .

Les valeurs propres sont des covariances au carrés, dès que les tableaux sont centrés pour les distributions marginales. On trouve :



Num.	Ei genval .	R. Iner.	R. Sum	Num.	Ei genval .	R. Iner.	R. Sum
01	+1. 0870E- 02	+0. 6571	+0. 6571	02	+3. 2239E- 03	+0. 1949	+0. 8520
03	+1. 6878E- 03	+0. 1020	+0. 9541	04	+5. 4385E- 04	+0. 0329	+0. 9869
05	+1. 2531E- 04	+0. 0076	+0. 9945	06	+5. 4464E- 05	+0. 0033	+0. 9978
07	+1. 9725E- 05	+0. 0012	+0. 9990	08	+1. 6637E- 05	+0. 0010	+1. 0000
09	+0. 0000E+00	+0. 0000	+1. 0000	10	+0. 0000E+00	+0. 0000	+1. 0000
11	+0. 0000E+00	+0. 0000	+1. 0000	12	+0. 0000E+00	+0. 0000	+1. 0000

p vaut 28 (28 modalités de variables environnementales) et q vaut 12 (12 modalités de variables de milieu). Le rang de \mathbf{Q}_0 est au plus de 8 (12 modalités - 4 variables) et il y a au moins 4 valeurs propres nulles (le rang de $\mathbf{Z} = \mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0$ ne peut excéder celui de \mathbf{Q}_0). Il y a une première valeur propre beaucoup plus grande que les autres. Les trois suivantes ont une décroissance régulière. On considérera que le premier axe exprime l'essentiel de la co-structure entre les trois tableaux. Deux facteurs sont conservés pour expliciter certaines aides à l'interprétation.

3.4.1 - Analyse RLQ comme analyse d'inertie

Comme analyse du triplet $(\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, \mathbf{D}_p, \mathbf{D}_q)$, l'analyse RLQ est une analyse d'inertie simple d'un tableau croisant les variables de \mathbf{R} (ici des modalités puisque les variables sont qualitatives) et celle de \mathbf{Q} (ici encore des modalités). On a donc directement deux cartes factorielles. Nous avons montré par ailleurs (Chessel & Mercier, 1993)¹⁸ et illustré (Mercier et al., 1992)¹⁹ que l'analyse de co-inertie entre un tableau floristique en présence-absence et un tableau de milieu en variables qualitatives était d'abord l'analyse des correspondances du tableau de profils écologiques initiée par Romane (1972)²⁰ et retrouvée vingt ans plus tard (Montana & Greigsmith, 1990)²¹. On étend ici cette proposition.

Considérons en effet une modalité de milieu et une modalité d'un trait biologique. On peut faire la somme de l'abondance de toutes les espèces porteuses de cette modalité ci dans tous les relevés porteurs de cette modalité là. On obtient le tableau $\mathbf{R}^t \mathbf{L} \mathbf{Q}$, où \mathbf{R} et \mathbf{Q} sont disjonctifs complets et \mathbf{L} est le tableau brut du lien (Tableau 4).

Les sommes marginales par lignes et par colonnes des sous-tableaux définis par les variables de chacun des tableaux de départ sont constantes. Ce tableau a les mêmes propriétés qu'un tableau de Burt croisé dont l'AFC définit l'analyse canonique sur variables qualitatives (Cazes, 1980)²² et dont nous avons montré qu'il s'agissait d'une analyse de co-inertie entre deux ACM. L'analyse des correspondances de ce tableau sera une analyse des correspondances internes (Cazes et al., 1988)²³ qui a des propriétés d'interprétation intéressantes. L'analyse des correspondances du tableau croisé donne :

Total inertia: 0. 0165417

Num.	Eigenval .	R. Iner.	R. Sum	Num.	Ei genval .	R. Iner.	R. Sum
01	+1. 0870E- 02	+0. 6571	+0. 6571	02	+3. 2239E- 03	+0. 1949	+0. 8520
03	+1. 6878E- 03	+0. 1020	+0. 9541	04	+5. 4385E- 04	+0. 0329	+0. 9869
11	+0. 0000E+00	+0. 0000	+1. 0000	12	+0. 0000E+00	+0. 0000	+1. 0000

Il est aisé de démontrer que l'analyse RLQ est exactement l'AFC de R^tLQ . De même c'est exactement l'analyse factorielle des correspondances internes de R^tLQ par rapport aux deux partitions des modalités entre les variables.

Ceci permet de préciser le rôle des variables dans cette analyse :

Tableau 4 Matrice croisée du type R^tLQ .

170	146	41	357	239	86	32	357	27	212	49	69	357	230	127	357
160	190	47	397	282	72	43	397	33	216	72	76	397	280	117	397
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
68	95	24	187	130	50	7	187	1	128	20	38	187	132	55	187
262	241	64	567	391	108	68	567	59	300	101	107	567	378	189	567
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
40	54	16	110	79	28	3	110	1	75	13	21	110	78	32	110
290	282	72	644	442	130	72	644	59	353	108	124	644	432	212	644
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
71	71	13	155	113	30	12	155	26	86	19	24	155	110	45	155
259	265	75	599	408	128	63	599	34	342	102	121	599	400	199	599
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
184	139	45	368	248	72	48	368	48	176	71	73	368	234	134	368
146	197	43	386	273	86	27	386	12	252	50	72	386	276	110	386
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
74	83	17	174	116	43	15	174	6	121	22	25	174	116	58	174
256	253	71	580	405	115	60	580	54	307	99	120	580	394	186	580
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
186	211	53	450	317	83	50	450	26	258	78	88	450	316	134	450
144	125	35	304	204	75	25	304	34	170	43	57	304	194	110	304
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
226	244	67	537	366	122	49	537	23	325	85	104	537	363	174	537
104	92	21	217	155	36	26	217	37	103	36	41	217	147	70	217
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
53	60	17	130	90	32	8	130	11	76	11	32	130	90	40	130
277	276	71	624	431	126	67	624	49	352	110	113	624	420	204	624
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
40	64	9	113	77	31	5	113	2	85	7	19	113	81	32	113
290	272	79	641	444	127	70	641	58	343	114	126	641	429	212	641
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754
30	24	11	65	44		21	65	9	13	28	15	65	44	21	65
54	45	19	118	80	19	19	118	12	51	30	25	118	77	41	118
69	42	12	123	77	36	10	123	14	75	11	23	123	71	52	123
33	37	8	78	55	16	7	78	2	50	11	15	78	56	22	78
34	41	11	86	59	23	4	86	1	58	10	17	86	60	26	86
42	55	15	112	77	31	4	112	2	73	12	25	112	76	36	112
23	45	9	77	59	18		77	1	56	6	14	77	57	20	77
45	47	3	95	70	15	10	95	19	52	13	11	95	69	26	95
330	336	88	754	521	158	75	754	60	428	121	145	754	510	244	754

Ei genvalue percentage

Bl oc 1 /-----/ . 004 / . 041 /
 Bl oc 2 /-----/ . 129 / . 007 /
 Bl oc 3 /-----/ . 062 / . 005 /
 Bl oc 4 /-----/ . 007 / . 154 /
 Bl oc 5 /-----/ . 159 / . 029 /
 Bl oc 6 /-----/ . 038 / . 000 /
 Bl oc 7 /-----/ . 002 / . 131 /

Bloc 8	/-----/	.080	/	.091	/
Bloc 9	/-----/	.012	/	.019	/
Bloc10	/-----/	.096	/	.006	/
Bloc11	/-----/	<u>.411</u>	/	<u>.515</u>	/
Eigenvalue percentage					
Bloc 1	/-----/	.086	/	.147	/
Bloc 2	/-----/	.280	/	.161	/
Bloc 3	/-----/	<u>.622</u>	/	<u>.670</u>	/
Bloc 4	/-----/	.012	/	.022	/

C'est donc la variable 11 (recouvrement de la végétation) et la variable 3 (site de nidification) qui font l'essentiel du lien entre les deux tableaux.

3.4.2 - Analyse RLQ comme analyse d'inertie partielle de R

L'analyse RLQ a de nombreuses aides à l'interprétation. En maximisant une covariance elle cherche à maximiser variances et corrélations, donc à faire partiellement une analyse d'inertie de **R**, de la même façon que les analyses de co-inertie sont partiellement des analyses d'inertie. On compare les deux points de vue en projetant les axes d'inertie de **R** sur les axes de l'analyse RLQ (Fig. 10). En gris, on voit les axes RLQ et en noir les 4 premiers axes de l'analyse simple de R. Il n'y a pas identité entre les deux, mais un point de vue original bien que contenant en partie l'axe 1 de l'analyse simple.

Comme en analyse de co-inertie on peut discuter de la valeur des axes RLQ (à droite valeurs cumulées). L'inertie projetée sur les axes RLQ (0.2208 et 0.1555) est assez loin de l'optimum atteint sur les axes d'inertie (0.2724 et 0.2135).

Num	Covaria.	Varian1	INER1	Num	Varian1	INER1
1	0.1043	<u>0.2208</u>	<u>0.2724</u>	1	0.2208	0.2724
2	0.05678	<u>0.1555</u>	<u>0.2135</u>	2	0.3763	0.4859

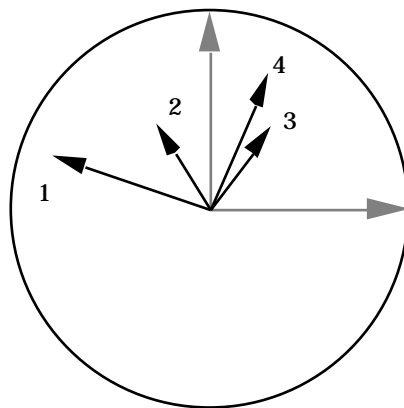


Figure 10 Projection des axes d'inertie de l'analyse de **R** sur les axes de l'analyse RLQ.

On peut également projeter le nuage des lignes de **R** sur les axes RLQ et cartographier la première coordonnée factorielle (Fig. 11 à comparer avec la Fig. 6B) :

Curieusement cette coordonnée RLQ est plus nettement caractéristique de la zone proprement urbaine que la première coordonnées de l'ACM de **R**. Le lien avec les variables biologiques décrit donc essentiellement la rupture ville/banlieue.

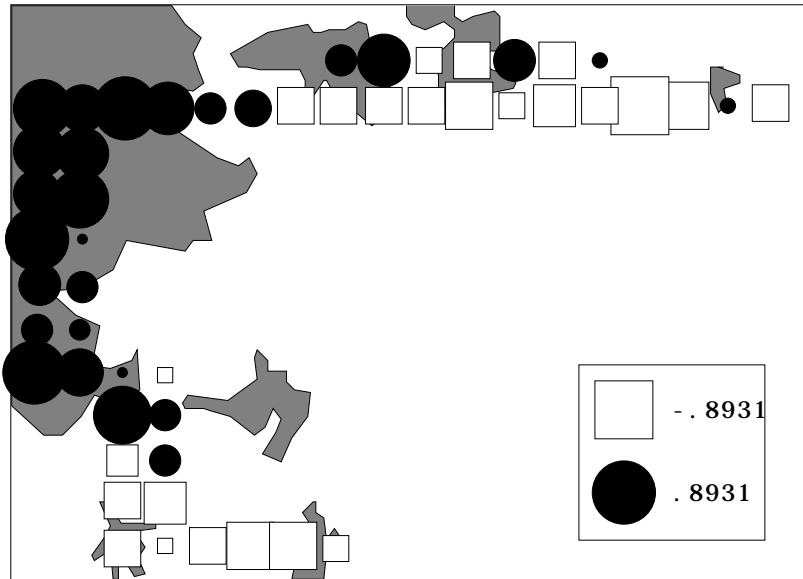


Figure 11 Cartographie de la première coordonnée factorielle de R dans l'analyse RLQ.

3.4.3 - Analyse RLQ comme analyse d'inertie partielle de Q

Comme en analyse de co-inertie on peut discuter de la valeur des axes RLQ (Fig. 12). Ici encore, l'inertie projetée sur les axes RLQ (0.3354 et 0.3717) peut être comparée à l'optimum (0.6508 et 0.4) :

Num	Covaria.	varian2	INER2	Num	Varian2	INER2
1	0.1043	0.3354	0.6508	1	0.3354	0.6508
2	0.05678	0.3717	0.4	2	0.7072	1.051

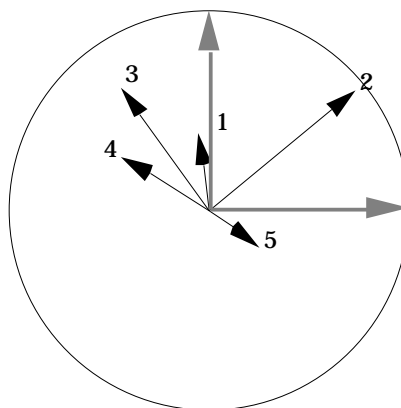
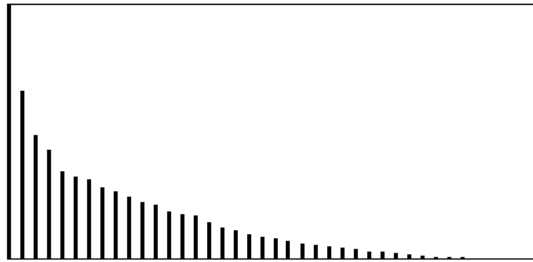


Figure 12 Projection des axes d'inertie de l'analyse de Q sur les axes de l'analyse RLQ.

Délibérément, l'axe 1 (insectivore/migrateur) de l'analyse simple (Fig. 7) est abandonné dans l'analyse RLQ. La direction choisie dans l'espace est franchement originale. Sur cette carte on s'est éloignée ainsi de l'analyse de Q (Fig. 13).

3.4.4 - Analyse RLQ comme analyse d'inertie partielle de L

La diagonalisation de l'AFC de L donne les valeurs propres :



En maximisant la covariance on maximise en partie la corrélation. Hors l'optimum de la corrélation est obtenue par la racine carrée de la première valeur propre de l'AFC (Williams, 1952, op. cit.). On peut comparer et voir que les coordonnées factorielles sur les premiers axes RLQ sont loin d'être optimaux en termes de coordonnées factorielles :

Num	Covari a.	Correl a.	AFC
1	0.1043	0.3831	0.6435
2	0.05678	0.2361	0.5220

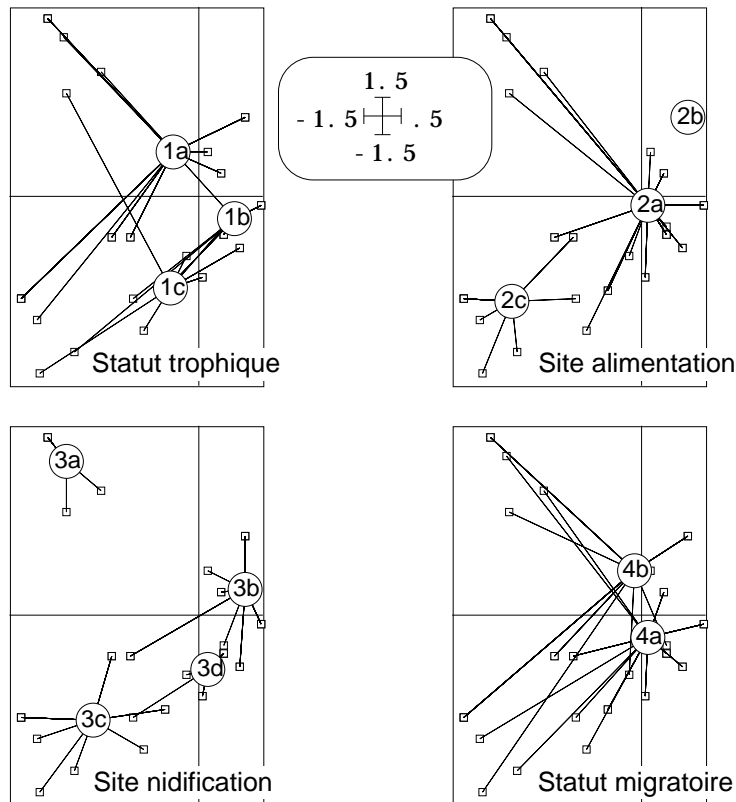


Figure 13 Plan 1-2 des coordonnées de espèces dans l'analyse RLQ munies de leurs traits biologiques (modalité à la moyenne pondérée des espèces présentant cette modalité).

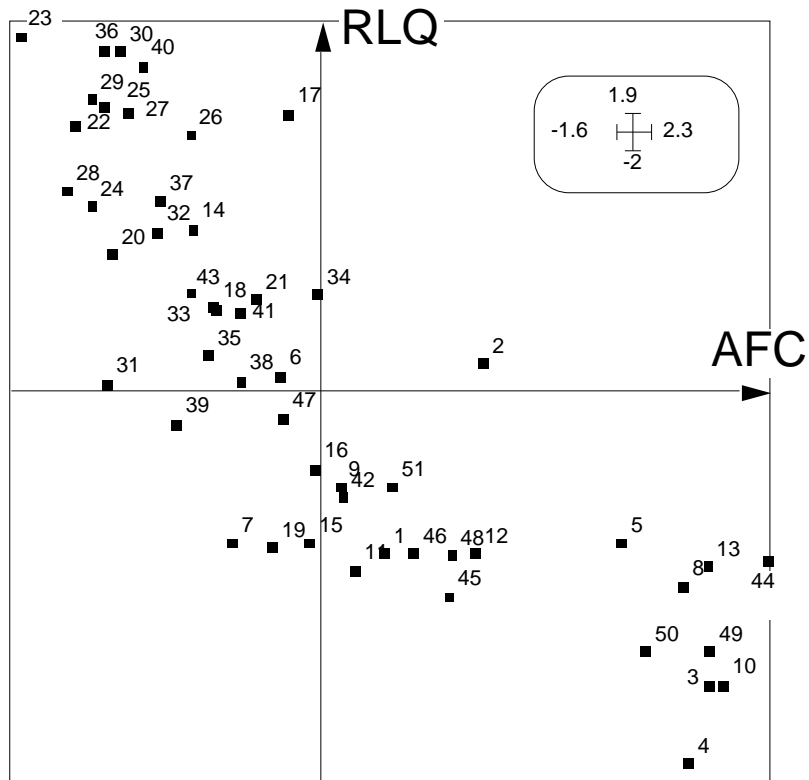


Figure 14 Graphe bivarié de la première coordonnée normalisée des lignes-relevés de **L** (AFC) et de la première coordonnée normalisée des lignes-relevés de **R** (analyse RLQ).

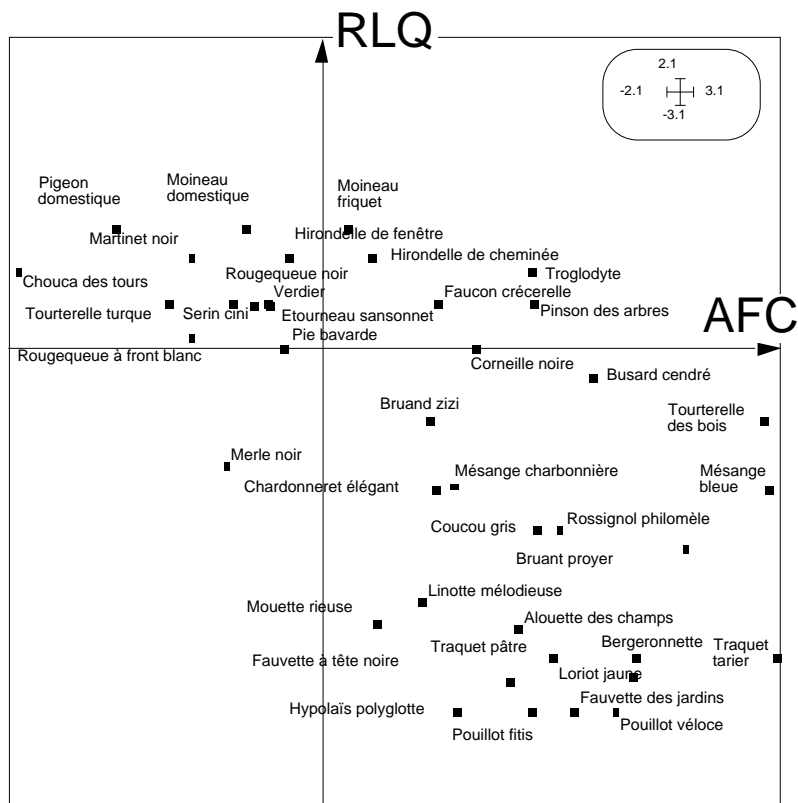


Figure 15 Graphe bivarié de la première coordonnée normalisée des colonnes-espèces de **L** (AFC) et de la première coordonnée normalisée des lignes-espèces de **Q** (analyse RLQ).

Pourtant, il y a bonne corrélation manifeste ($R = -0.817$ avec les poids marginaux) entre la première coordonnée normalisée des lignes de l'AFC de **L** et la première coordonnée normalisée des lignes de **R** sur l'axe 1 de l'analyse RLQ (Fig. 14). Cela signifie que l'on trouve dans le tableau **R** une combinaison de variable reconstituant à peu près l'ordination des relevés faite par le cortège faunistique et cela n'est pas surprenant. Le gradient urbain rural s'exprime tant par la faune que par les indications de milieu. Si on a perdu sur la corrélation canonique c'est que l'enregistrement des variables biologiques est plus sommaire. On n'y trouve pas de quoi reconstituer l'ordination des espèces faites par les relevés (ce qui est parfaitement cohérent avec les remarques faites au § 3.3). La corrélation ($R = -0.723$ avec les poids marginaux) entre la première coordonnée normalisée des colonnes de l'AFC de **L** et la première coordonnée normalisée des lignes de **Q** sur l'axe 1-RLQ est cependant non négligeable (Fig. 15).

3.4.5 - Choix de l'expression d'un lien triple

Aucun exemple n'épuise une présentation méthodologique. On a simplement rencontré le lien modalités de milieu - relevés (**R**), le lien relevés - espèces (**L**) et le lien espèces - modalités biologiques (**Q**). L'expression simultanée des trois est au cœur de l'analyse RLQ. Ici un facteur exprime l'essentiel. Pour faire une synthèse on choisira de mettre en évidence les trois relations d'averaging sous-jacentes à un triplet d'AFC. Donnons un code numérique des modalités du tableau **R** centré par variables et de variance moyenne (entre variables) unité. Cette opération est la base de l'ACM de **R**. Chaque relevé peut se placer à la moyenne des modalités de milieu qu'on y rencontre (Fig. 16).

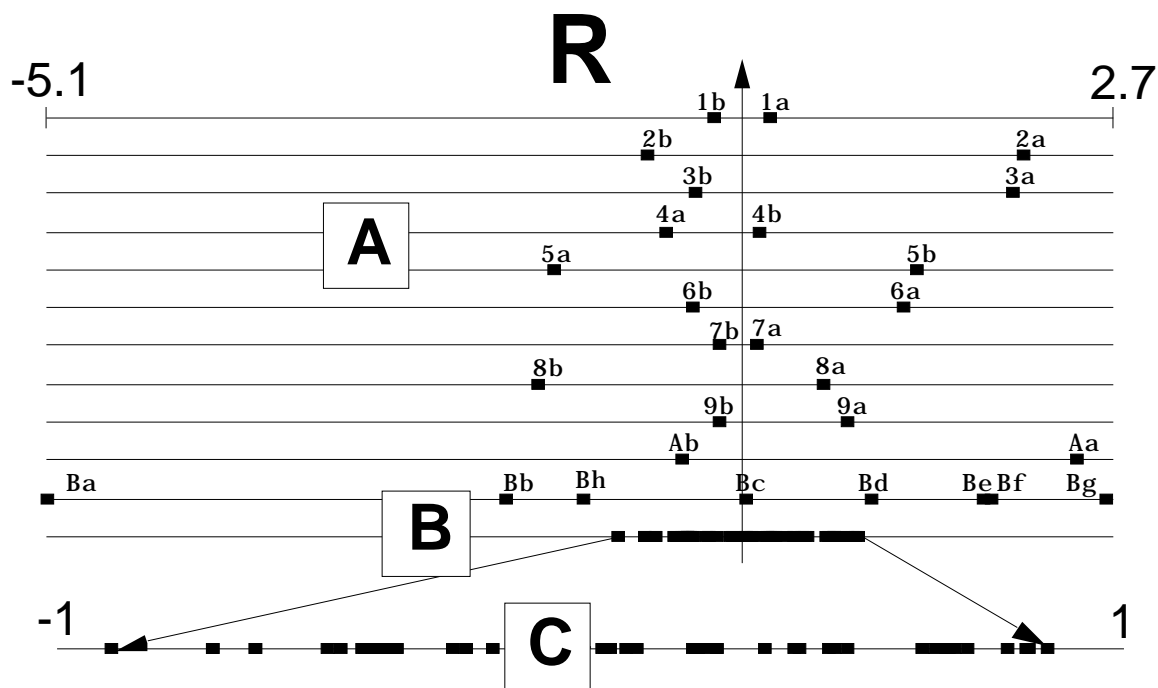


Figure 16 Première opération d'averaging. A - Positionnement des 28 modalités de milieu réparties entre les 11 variables. La moyenne des positions est nulle pour chaque variable. La variance des positions vaut en moyenne 1. B - Relevés à la moyenne des modalités présentes. C - Le résultat pour être lisible est dilaté sur l'intervalle $[-1, 1]$.

Les sites se retrouvent ainsi ordonnés par leurs caractéristiques mésologiques. De plus, chaque espèce est une distribution de fréquence inter-relevés. Ce point de vue du **gradient** (Whittaker, 1967)²⁴ est un fondement de l'écologie factorielle développé dans (Whittaker, 1973)²⁵.

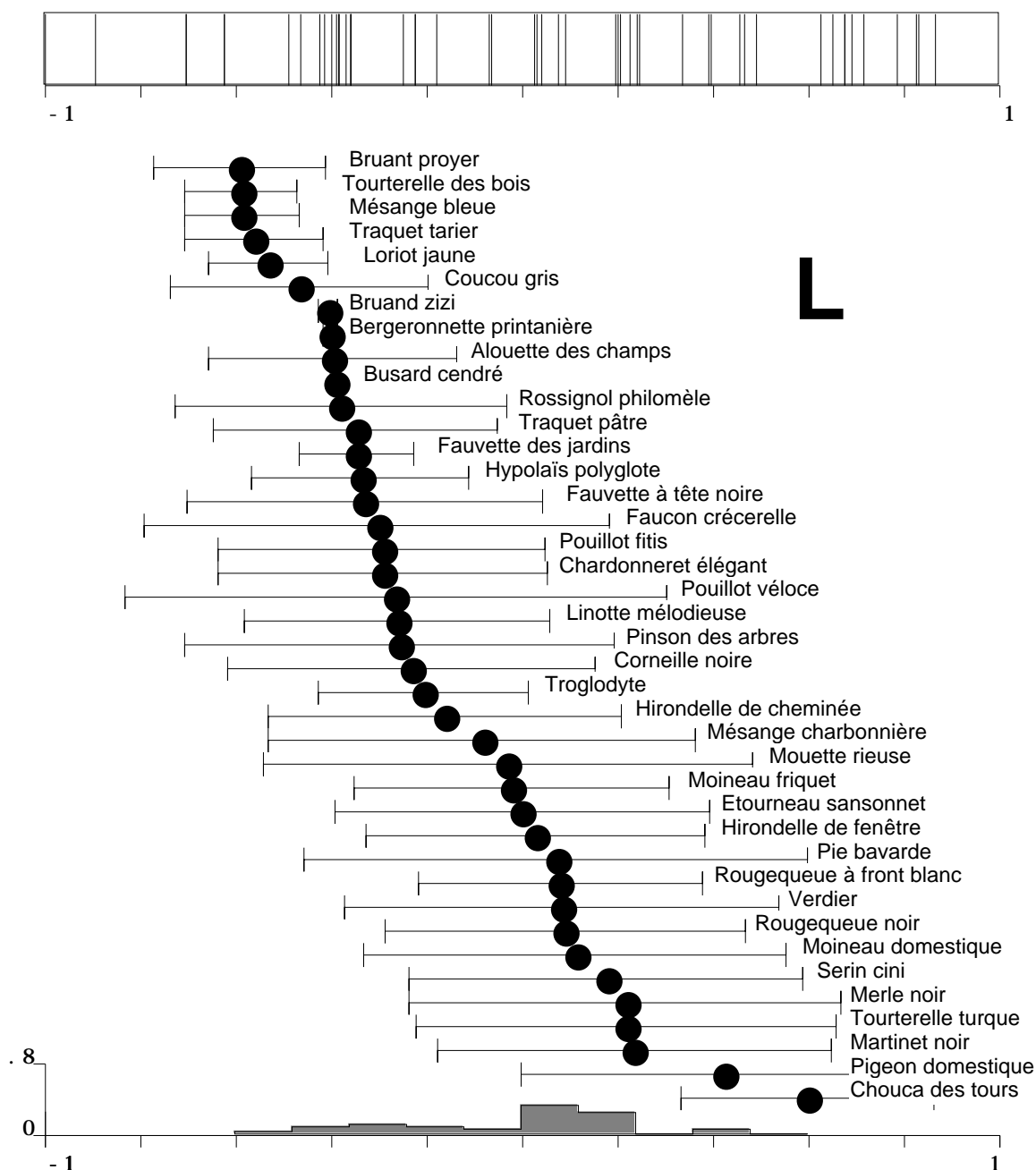


Figure 17 Reprise du code des relevés de la Fig. 16 et position par averaging ordonné des espèces le long du gradient. Chaque espèce est figurée par sa moyenne et son écart-type. En bas distribution d'abondance totale le long du gradient. Le code numérique qui positionne les espèces est centré pour la distribution marginale.

Ce point de vue justifie fondamentalement l'usage de l'AFC en écologie dans sa version simple (Hill, 1973)²⁶ et dans son extension aux variables instrumentales (Ter Braak, 1986)²⁷.

Sur le gradient les espèces sont positionnées à la moyenne conditionnelle des relevés qu'elles occupent et sont ainsi ordonnées (*Averaging species/samples*). En prenant l'ordre proposé par la Fig. 16, on obtient la Fig. 17.

Il s'agit en partie d'une pratique d'AFC sur **L** en partant d'une pratique d'ACM sur **R**. En s'arrêtant ici, on serait typiquement dans une situation d'analyse canonique des correspondances ou encore d'AFCVI^{28,29}. En bas de la Fig. 17 on trouve la distribution

des notes d'abondance de toutes les espèces confondues sur ce gradient. Chaque taxon y participe pour sa note totale.

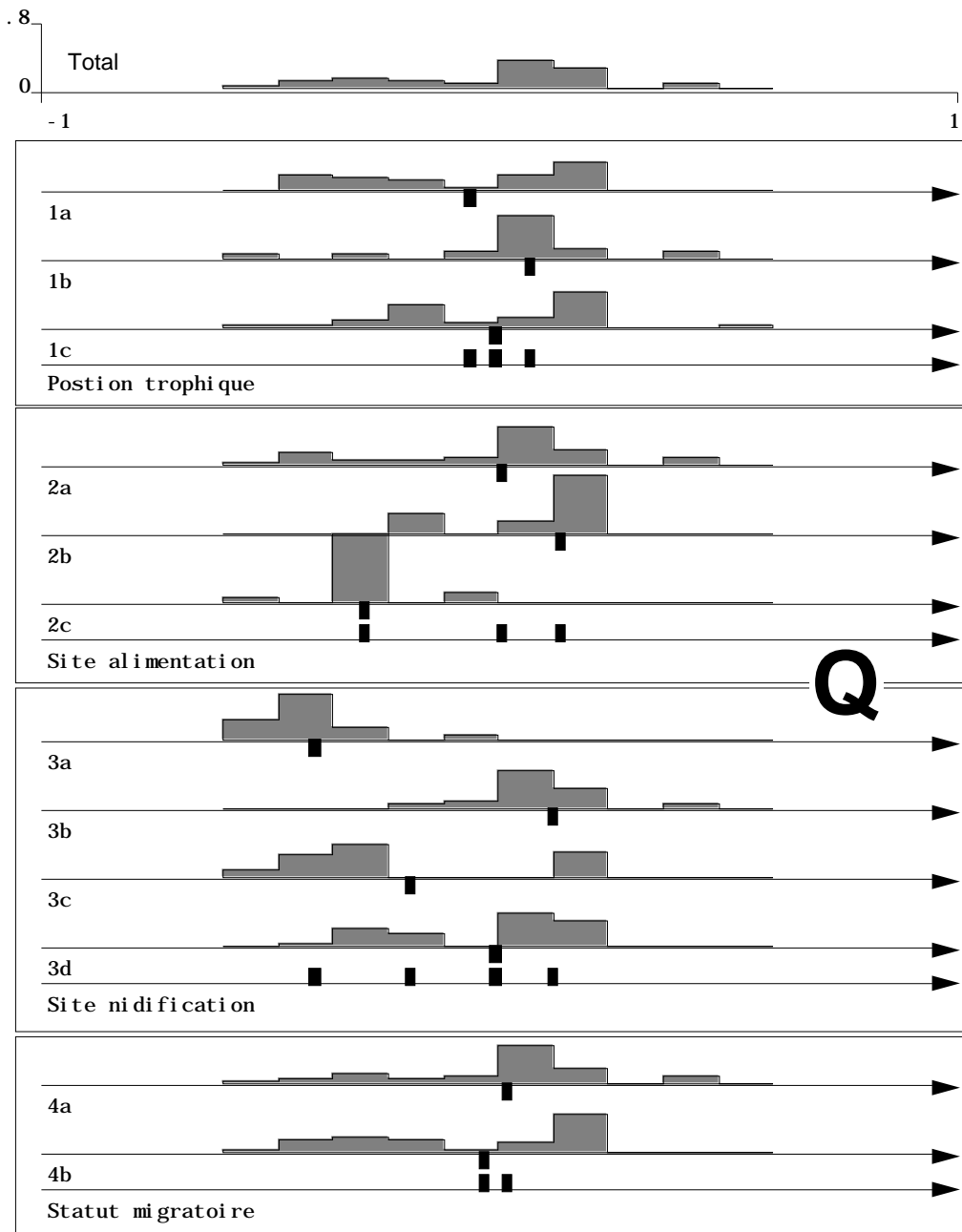


Figure 18 Reprise du code espèce et position moyenne de chaque modalité des traits biologiques à la moyenne pondérée des espèces qui les portent. L'analyse RLQ maximise la moyenne (entre variables biologiques) des variances des positions des modalités.

Les modalités de milieu ont positionné les relevés sur un gradient par averaging sur **R**. Les moyennes conditionnelles (averaging sur **L**) ont positionné sur ce gradient les espèces. Il reste alors à positionner par averaging sur **Q** les modalités biologiques (Fig. 18). C'est pourquoi on a choisi le nom d'analyse RLQ, car dans le cas de trois analyses des correspondances on exécute une suite cohérente de moyennages.

Chaque modalité biologique est portée par un ensemble d'espèces qui ont des poids (note marginale dans **L**) et des positions (représentées dans la Fig. 17). On peut donc définir la moyenne de la modalité sur le gradient, puis pour une variable la variance de

ces moyennes (variance inter), puis pour toutes les variables la variance de ces moyennes.

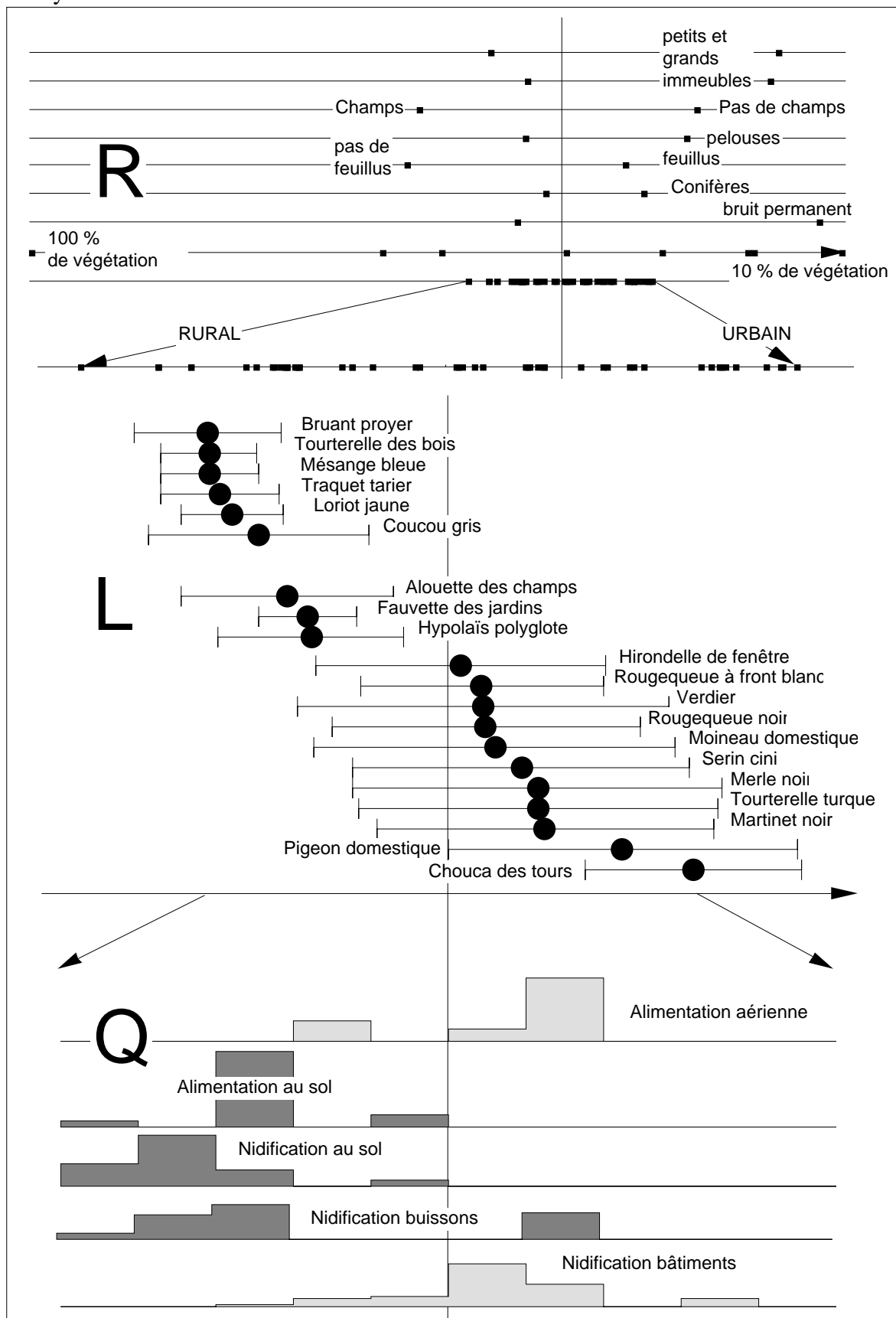


Figure 19 Résumé des averagings successifs d'une analyse RLQ reliant une AFC et deux ACM.

Plus la variance de ces moyennes sera grande, plus pertinent aura été le choix du code de milieu de début (Fig. 16) et plus clairement sera exprimée la corrélation quadruple modalités de milieu/relevés/espèces/modalités biologiques (Fig. 18). Il est aisé de démontrer, comme propriété particulière du schéma de départ que le choix optimum est fait par l'analyse RLQ.

La Fig. 19 résume ce qu'on apprend de cette façon. On aurait pu faire le chemin inverse. L'analyse ne privilégie pas un modèle de facteur limitant contre un modèle de milieu spécifique pour l'espace urbain, mais laisse les deux interprétations vraisemblables. Il n'est pas question, ici de faire des observations originales sur un matériel parfaitement connu, mais de voir si la procédure repère automatiquement les associations sur trois tableaux et de tester la cohérence de la méthode sur du matériel concret. Il semble difficile de résumer plus efficacement les données présentes dans le Tableau 2.

4 - Analyse RLQ et AFC sous double contraintes

En guise de conclusion provisoire, on notera les points suivants qui replacent l'analyse RLQ³⁰ dans l'ensemble des méthodes linéaires.

L'exemple traité est formé d'un triplet d'analyses des correspondances, une AFC centrale et deux ACM marginales. A l'intérieur du groupe de méthodes attachées à l'AFC, celle du tableau **R^tLQ** avec projection en individus supplémentaires des lignes de **R** et des lignes de **Q** retrouve l'essentiel de l'analyse RLQ du triplet.

Ceci se passe exactement comme pour l'AFC du tableau des profils écologiques avec projections en individus supplémentaires des lignes des deux tableaux faune et milieu qui redonne l'essentiel de l'analyse de co-inertie (Mercier et al., op.cit.). Ce point de vue a déjà été utilisé (Ch. Mullon, comm. pers.). Mais il s'agit d'un cas particulier et l'analyse RLQ est apte à prendre en compte une AFC centrale et deux tableaux marginaux quelconques. La question des triplets de tableaux a déjà été abordé dans la littérature de la théorie de l'analyse des données. Takane & Shibayama (1991)³¹ aborde la question des tableaux **L** de notations juges-produits sur lesquels on dispose d'informations **R** sur les juges (sexe, âge, niveau d'éducation) et **Q** sur les produits. La question est posée en termes de reconstitution de données du type :

$$\mathbf{L} = \mathbf{RMQ}^t + \mathbf{BQ}^t + \mathbf{RC} + \mathbf{E}$$

où **R**, **L** et **Q** sont les tableaux initiaux, **M**, **B** et **C** sont des tableaux de coefficients à estimer et **E** est le tableau des résidus. Ce point de vue conduit à l'introduction des projecteurs associés aux tableaux marginaux dans l'espace des lignes et celui des colonnes de **L** et à la décomposition :

$$\mathbf{L} = \pi_{\mathbf{R}}\mathbf{L}\pi_{\mathbf{Q}} + \pi_{\mathbf{R}}\mathbf{L}\pi_{\mathbf{Q}} + \pi_{\mathbf{R}}\mathbf{L}\pi_{\mathbf{Q}} + \pi_{\mathbf{R}}\mathbf{L}\pi_{\mathbf{Q}}$$

Les quatre termes sont orthogonaux pour la trace et confère à cette décomposition une canonicité indiscutable. Si **R** et **Q** sont simplement deux variables qualitatives on retrouve dans le quatrième terme la double intra-classe (Cazes et al. op. cit.) et dans le premier la double inter-classes qui est exactement celle de l'analyse RLQ.

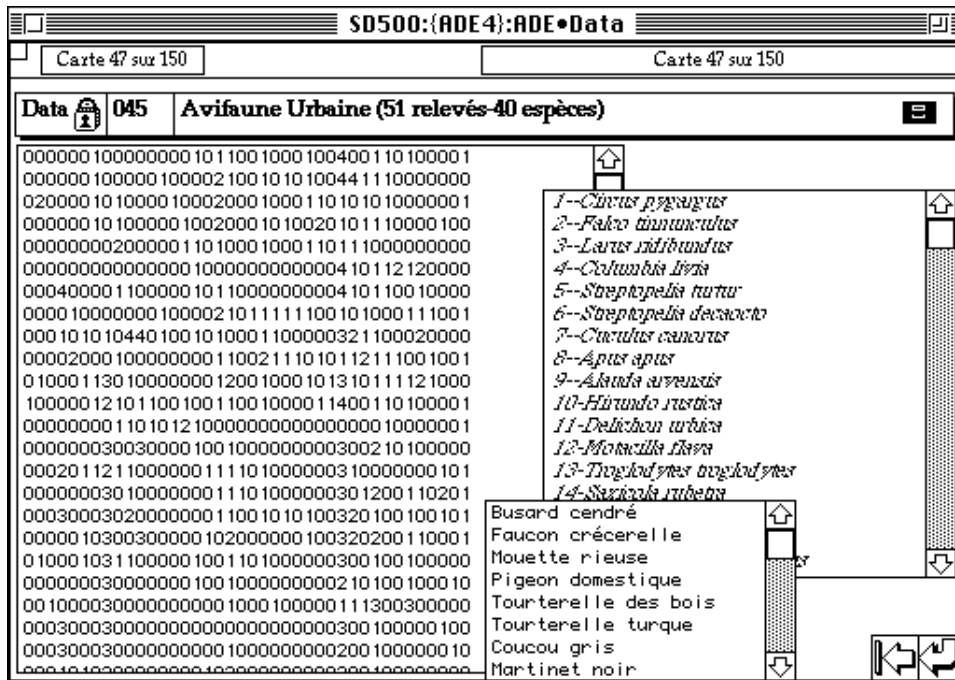
Si **L** est un tableau d'AFC avec **R** et **Q** quelconques on trouve par l'analyse de $\pi_{\mathbf{R}}\mathbf{L}\pi_{\mathbf{Q}}$ l'analyse des correspondances sous double contraintes linéaires de

Böckenholt & Böckenholt (1990)³². L'analyse RLQ est une option sans inversion de norme de cette dernière. Cela signifie que les analyses inter-classes et les analyses discriminantes, les analyses de co-inertie et les ACP sur variables instrumentales, les analyses RLQ et les AFC sous double contraintes entretiennent respectivement les mêmes relations. Nous avons préféré systématiquement les premières aux secondes pour des questions de stabilité numérique vu la présence en général en écologie du grand nombre de descripteurs.

Mathématiquement l'AFC sous double contraintes est équivalente à la diagonalisation du schéma $\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, (\mathbf{Q}_0^t \mathbf{D}_J \mathbf{Q}_0)^-, (\mathbf{R}_0^t \mathbf{D}_I \mathbf{R}_0)^-$ alors que l'analyse RLQ est celle du schéma $(\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, \mathbf{D}_p, \mathbf{D}_q)$. Toutes les discussions ouvertes sur le rôle de la norme de Mahalanobis sont maintenant étendues aux deux marges.

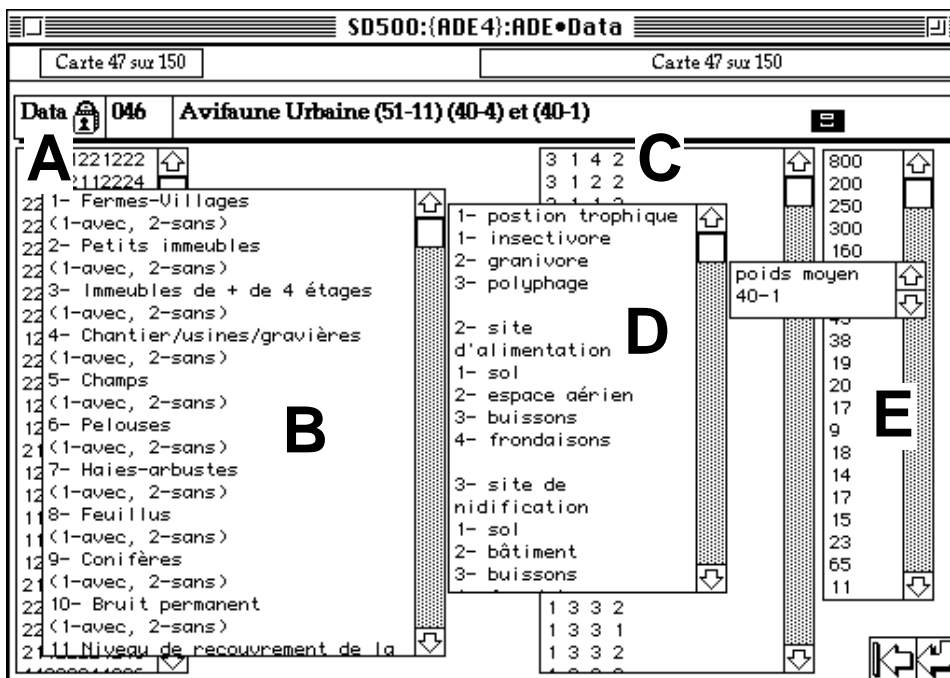
Appendice 1 - Préparation des données

Aller à la carte «AviUrba» de la pile ADE•Data :



Faire un fichier AUFau.Car avec le champ de données et un fichier d'étiquettes AUCode_0i sFR.

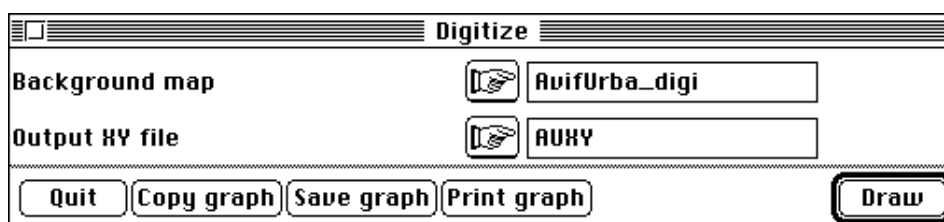
Aller à la carte «AviUrba+1» de la pile ADE•Data :



Faire un fichier AUMi1.Car avec le champ A et un fichier d'étiquettes AUCode_VarMi1 avec le champ B. Faire un fichier AUEsp.txt avec le champ C et un fichier d'étiquettes AUCode_VarEsp avec le champ D. Faire un fichier AUpoi ds.txt avec le champ E.

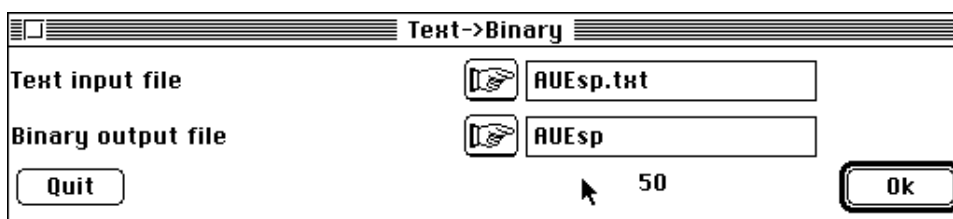
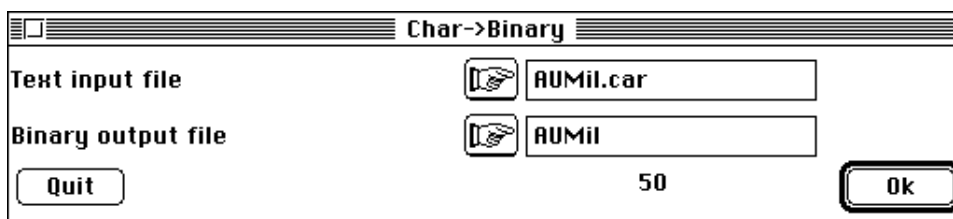
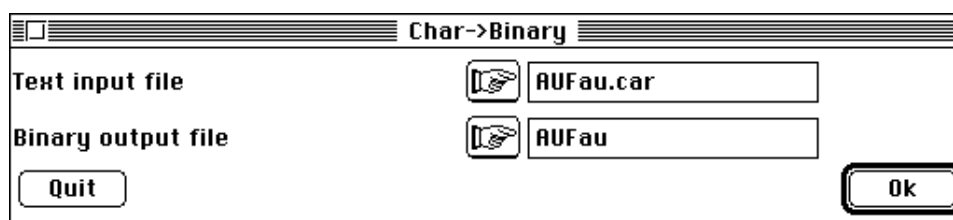
Copier le fond de carte AvifUrba_carto et le fond de carte pour digitalisation AvifUrba_Digi à l'aide de la commande **Copy files** du menu **Data Folder**.

Digitaliser l'emplacement des relevés à l'aide de l'option **Digitalize** du module **Digit** :



On dispose alors d'un fichier AUXY (51-2) contenant les coordonnées géographiques des relevés et permettant de faire une représentation spatiale des coordonnées factorielles.

Transformer en binaire (module **TextToBin**) les fichiers de données suivants en utilisant l'option adéquate :

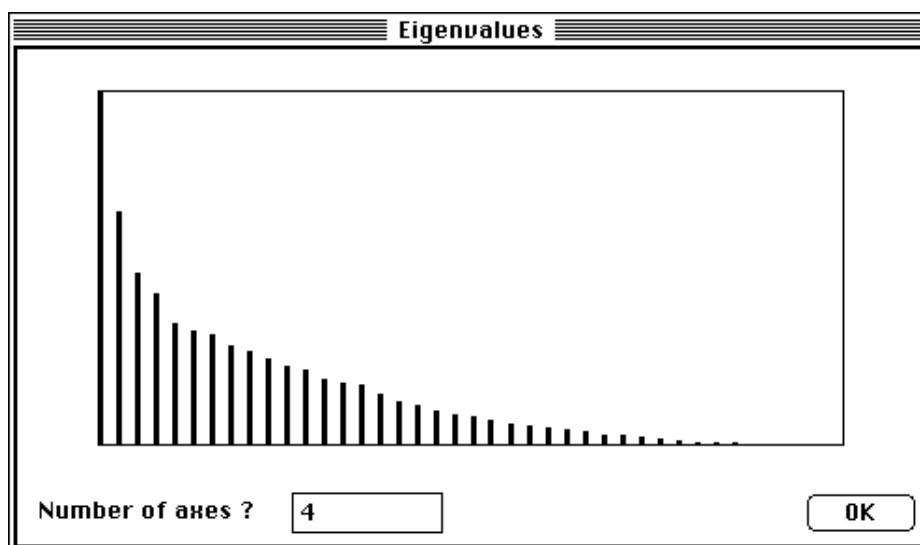
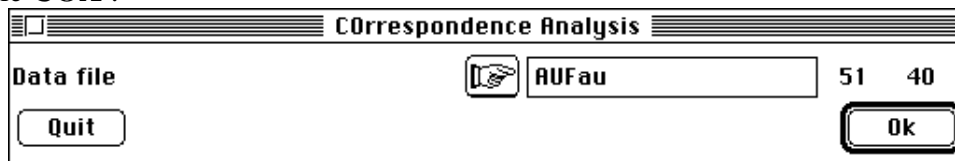


On obtient en sortie un fichier **L** qui est le fichier AUFau (51 lignes-relevés / 40 colonnes-espèces), un fichier **R** qui est le fichier AUMi 1 (51 lignes-relevés / 11 colonnes-variables de milieu), et un fichier **Q** qui est le fichier AUEsp (40 lignes-espèces / 4 colonnes-trait biologiques).

AUFau (51-40) est un tableau d'abondance espèces d'oiseaux-relevés. AUMi 1 (51-11) est un tableau de variables de milieu qualitatives. AUEsp (51-11) est un tableau des traits biologiques des espèces d'oiseaux également codés qualitativement.

Appendice 2 - Analyses séparées

Faire l'AFC de AUFau (51-40) en utilisant l'option **CO**rrespondence Analysis du module **COA** :



```
fc/COA: Correspondance analysis
Input file: AUFau
Number of rows: 51, columns: 40
File AUFau.fcpl contains the edge distribution of rows
It has 51 rows and 1 column
File AUFau.fcpc contains the edge distribution of columns
It has 40 rows and 1 column
File AUFau.fcta contains the doubly centred table DI-1*P*DJ-1 -II*1J'
It has 51 rows and 40 columns
File AUFau.fcma contains:
    the number of rows: 51
    the number of columns: 40
    the total number: 754
```

```
-----
DiagoRC: General program for two diagonal inner product analysis
Input file: AUFau.fcta
--- Number of rows: 51, columns: 40
-----
Total inertia: 2.65918
-----
```

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+4.1415E-01	+0.1557	+0.1557	02	+2.7249E-01	+0.1025	+0.2582
03	+2.0078E-01	+0.0755	+0.3337	04	+1.7658E-01	+0.0664	+0.4001
05	+1.4374E-01	+0.0541	+0.4542	06	+1.3354E-01	+0.0502	+0.5044
07	+1.3032E-01	+0.0490	+0.5534	08	+1.1691E-01	+0.0440	+0.5974
09	+1.1053E-01	+0.0416	+0.6389	10	+1.0046E-01	+0.0378	+0.6767
11	+9.2890E-02	+0.0349	+0.7116	12	+8.7307E-02	+0.0328	+0.7445
13	+7.6323E-02	+0.0287	+0.7732	14	+7.2505E-02	+0.0273	+0.8004
15	+7.0162E-02	+0.0264	+0.8268	16	+5.8857E-02	+0.0221	+0.8490
17	+5.0919E-02	+0.0191	+0.8681	18	+4.7711E-02	+0.0179	+0.8861
19	+3.9840E-02	+0.0150	+0.9010	20	+3.6752E-02	+0.0138	+0.9149
21	+3.2753E-02	+0.0123	+0.9272	22	+3.0082E-02	+0.0113	+0.9385

```

23 +2.5692E-02 +0.0097 +0.9482 | 24 +2.2164E-02 +0.0083 +0.9565 |
25 +2.0063E-02 +0.0075 +0.9640 | 26 +1.8282E-02 +0.0069 +0.9709 |
...
37 +1.3815E-03 +0.0005 +0.9996 | 38 +5.6946E-04 +0.0002 +0.9998 |
39 +4.1851E-04 +0.0002 +1.0000 | 40 +0.0000E+00 +0.0000 +1.0000 |

```

File AUFau.fcvp contains the eigenvalues and relative inertia for each axis

--- It has 40 rows and 2 columns

File AUFau.fcco contains the column scores

--- It has 40 rows and 4 columns

..... etc.

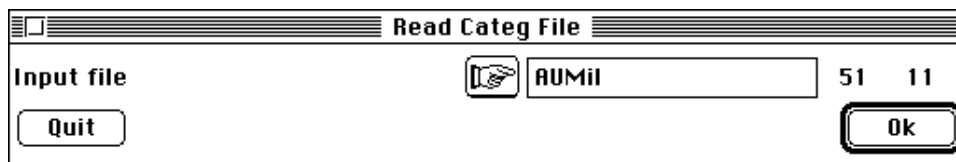
File AUFau.fcli contains the row scores

--- It has 51 rows and 4 columns

..... etc.

La racine carrée de la première valeur propre de cette analyse (0.414) donne l'optimum de corrélation espèces/relevés (0.643: voir p. 26 de cette fiche).

Le tableau AUMi1 (51-11) comme tout tableau de variables qualitatives doit être préalablement lu par l'option **Read Categ File** du module **CategVar** :



Categorical variables: file AUMi1

Rows: 51, Variables: 11, Categories: 28, Missing data: 0

Description of categories:

Variable number 1 has 2 categories

[1]Category: 1 Num: 22 Freq.: 0.431
[2]Category: 2 Num: 29 Freq.: 0.569

..... etc.

Variable number 11 has 8 categories

[21]Category: 1 Num: 4 Freq.: 0.0784
[22]Category: 2 Num: 7 Freq.: 0.137
[23]Category: 3 Num: 7 Freq.: 0.137
[24]Category: 4 Num: 5 Freq.: 0.098
[25]Category: 5 Num: 5 Freq.: 0.098
[26]Category: 6 Num: 9 Freq.: 0.176
[27]Category: 7 Num: 6 Freq.: 0.118
[28]Category: 8 Num: 8 Freq.: 0.157

Auxiliary binary output file AUMi1Moda: Indicator vector of modalities

It contains variable number for each modality

It has 28 rows (modalities) and one column

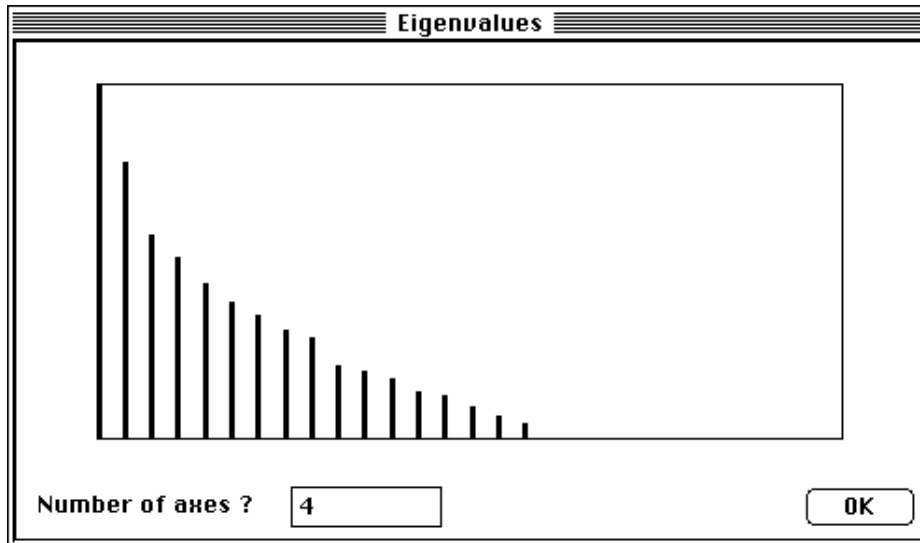
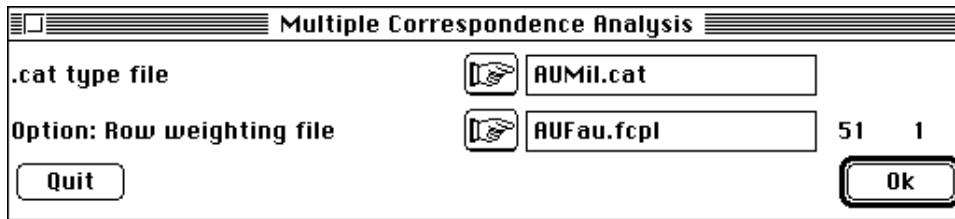
Auxiliary ASCII output file AUMi1.123: labels (two characters) for 28 modalities

It contains one label for each modality

It has 28 rows (modalities) and labels 1a, 1b, ..., 2a, 2b, ...

Variable number 1, 2, ..., A, ..., Z, +, Modality number a, b, ..., z, +

Enchaîner avec l'analyse des correspondances multiples de AUMi1 en utilisant l'option **Multiple Correspondence Analysis** du module **MCA** et en prenant garde conformément aux indications de la p. 14 d'utiliser le poids des lignes du tableau faunistique comme poids des lignes :



cm/MCA: Multiple correspondence analysis

Row weights from file AUFau.fcpl

File AUMil.cmpl contains the row weights
It has 51 rows and 1 column

File AUMil.cmpc contains the column weights $(1/V) * DM$
It has 28 rows and 1 column

Marginal distributions by variable:

Variable number 1 has 2 categories

[1] Category: 1 Weight: 0.473
[2] Category: 2 Weight: 0.527

..... etc.

Variable number 11 has 8 categories

[21] Category: 1 Weight: 0.0862
[22] Category: 2 Weight: 0.156
[23] Category: 3 Weight: 0.163
[24] Category: 4 Weight: 0.103
[25] Category: 5 Weight: 0.114
[26] Category: 6 Weight: 0.149
[27] Category: 7 Weight: 0.102
[28] Category: 8 Weight: 0.126

File AUMil.cmta contains the tabled processed by MCA

It has 51 rows and 28 columns (categories)

DiagoRC: General program for two diagonal inner product analysis

Input file: AUMil.cmta

--- Number of rows: 51, columns: 28

Total inertia: 1.54545

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+2.7245E-01	+0.1763	+0.1763	02	+2.1347E-01	+0.1381	+0.3144
03	+1.5694E-01	+0.1015	+0.4160	04	+1.3953E-01	+0.0903	+0.5063
05	+1.1955E-01	+0.0774	+0.5836	06	+1.0488E-01	+0.0679	+0.6515
07	+9.5740E-02	+0.0619	+0.7134	08	+8.4135E-02	+0.0544	+0.7679
09	+7.8036E-02	+0.0505	+0.8184	10	+5.6533E-02	+0.0366	+0.8549
11	+5.2744E-02	+0.0341	+0.8891	12	+4.6863E-02	+0.0303	+0.9194
13	+3.5938E-02	+0.0233	+0.9426	14	+3.3622E-02	+0.0218	+0.9644
15	+2.4592E-02	+0.0159	+0.9803	16	+1.7814E-02	+0.0115	+0.9918
17	+1.2613E-02	+0.0082	+1.0000	18	+0.0000E+00	+0.0000	+1.0000
19	+0.0000E+00	+0.0000	+1.0000	20	+0.0000E+00	+0.0000	+1.0000
21	+0.0000E+00	+0.0000	+1.0000	22	+0.0000E+00	+0.0000	+1.0000
23	+0.0000E+00	+0.0000	+1.0000	24	+0.0000E+00	+0.0000	+1.0000
25	+0.0000E+00	+0.0000	+1.0000	26	+0.0000E+00	+0.0000	+1.0000
27	+0.0000E+00	+0.0000	+1.0000	28	+0.0000E+00	+0.0000	+1.0000

File AUMil.cmvp contains the eigenvalues and relative inertia for each axis

--- It has 28 rows and 2 columns

File AUMil.cmco contains the column scores

--- It has 28 rows and 4 columns

..... etc.

File AUMil.cmlr contains the row scores

--- It has 51 rows and 4 columns

..... etc.

CorRatioMCA: Correlation ratios after a MCA

Title of the analysis: AUMil.cm

Number of rows: 51, columns: 11

Variable : 1

> Categ= 1	Weight=	0.473	-0.585	0.320	-0.416	-0.270
> Categ= 2	Weight=	0.527	0.526	-0.288	0.374	0.243
-----> r=		0.308	0.092	0.156	0.065	

Variable : 2

> Categ= 1	Weight=	0.248	1.139	0.639	-0.284	0.027
> Categ= 2	Weight=	0.752	-0.376	-0.211	0.094	-0.009
-----> r=		0.428	0.135	0.027	0.000	

..... etc.

Variable : 11

> Categ= 1	Weight=	0.086	-0.379	-0.093	1.356	2.141
> Categ= 2	Weight=	0.156	-1.025	0.698	0.746	-0.431
> Categ= 3	Weight=	0.163	-1.114	0.164	-0.553	-0.848
> Categ= 4	Weight=	0.103	-0.089	-0.049	-0.866	1.027
> Categ= 5	Weight=	0.114	0.795	1.066	-1.456	0.373
> Categ= 6	Weight=	0.149	0.935	0.277	0.197	-0.195
> Categ= 7	Weight=	0.102	1.374	-0.194	1.057	-0.737
> Categ= 8	Weight=	0.126	0.113	-2.110	-0.197	-0.186
-----> r=		0.776	0.787	0.740	0.732	

File AUMil.cmrc contains the correlation ratios between the categorical variables and the factor scores

It has 11 rows and 4 columns

Utiliser l'option **Stars** du module **ScatterClass** pour réaliser la Fig. 6A :

Stars			
HY coordinates file		AUMil.cml	51 4
H-axis column number (default = 1)			
Y-axis column number (default = 2)			
Categories file (.cat)		AUMil.cat	
Variable label file (or #)		AUCode_VarMil	
Row weights file (option)		AUMil.cmpl	51 1
Draw points (no = 2)			
<input type="button" value="Quit"/> <input type="button" value="Copy graph"/> <input type="button" value="Save graph"/> <input type="button" value="Print graph"/>			<input type="button" value="Draw"/>

L'utilisation du fichier AUCode_VarMil requiert quelques manipulations préalables sous un éditeur de texte (**BBEdit Lite** par exemple). On ne gardera que les étiquettes de variables et on éliminera les étiquettes des modalités. Les rapports de corrélation (chiffres en italiques sur la Fig. 6A) sont ajoutés sous ClarisDraw™. Pour représenter les coordonnées factorielles des relevés (Fig. 6B), utiliser l'option **Values** du module **Maps** :

Values			
Background map (Pict file)		AuifUrba_carto	
HY file		AUHY	51 2
Label file (or #)		#	
Input data file		AUMil.cml	51 4
Rescale between 0 and 1 (yes=1)			
Same scale for all col. (no=1)			
<input type="button" value="Quit"/> <input type="button" value="Copy graph"/> <input type="button" value="Save graph"/> <input type="button" value="Print graph"/>			<input type="button" value="Draw"/>

Le tableau AUEsp (40-4) comme tout tableau de variables qualitatives doit être préalablement lu par l'option **Read Categ File** du module **CategVar** :

Read Categ File			
Input file		AUEsp	40 4
<input type="button" value="Quit"/>			<input type="button" value="Ok"/>

```

Categorical variables: file AUEsp
Rows: 40, Variables: 4, Categories: 12, Missing data: 0
Description of categories:
-----
Variable number 1 has 3 categories
-----
[ 1]Category:  1 Num:   19 Freq.:  0.475
[ 2]Category:  2 Num:   12 Freq.:  0.3
[ 3]Category:  3 Num:    9 Freq.:  0.225
..... etc.
Variable number 4 has 2 categories

```

```
-----
[ 11]Category:   1 Num:   23 Freq.:   0.575
[ 12]Category:   2 Num:   17 Freq.:   0.425
-----
```

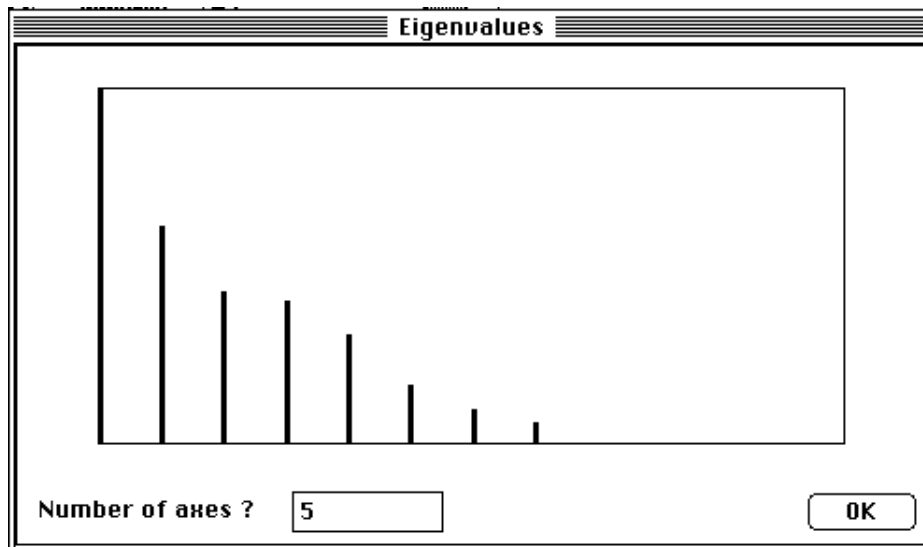
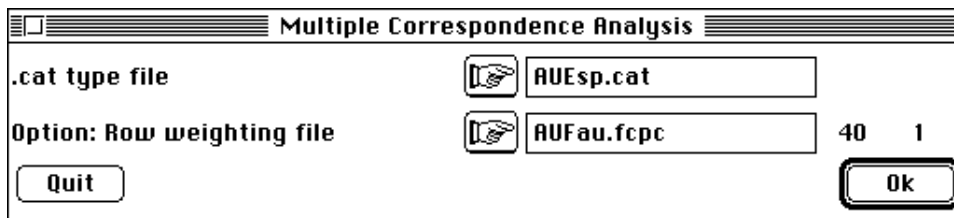
Auxiliary binary output file AUEspModa: Indicator vector of modalities

It contains variable number for each modality
It has 12 rows (modalities) and one column

Auxiliary ASCII output file AUEsp.123: labels (two characters) for 12 modalities

It contains one label for each modality
It has 12 rows (modalities) and labels 1a, 1b, ..., 2a, 2b, ...
Variable number 1, 2, ..., A, ..., Z, +, Modality number a, b, ..., z, +

Enchaîner avec l'analyse des correspondances multiples de AUEsp en utilisant l'option **Multiple Correspondence Analysis** du module **MCA** et en prenant garde conformément aux indications de la p. 14 d'utiliser le poids des colonnes du tableau faunistique comme poids des lignes :



cm/MCA: Multiple correspondence analysis
Row weights from file AUFau.fcpc

File AUEsp.cmpl contains the row weights
It has 40 rows and 1 column

File AUEsp.cmpc contains the column weights $(1/V) * DM$
It has 12 rows and 1 column

Marginal distributions by variable:

```
-----
Variable number 1 has 3 categories
-----
```

```
[1] Category:   1 Weight:   0.438
```

```
[2] Category: 2 Weight: 0.446
[3] Category: 3 Weight: 0.117
..... etc.
Variable number 4 has 2 categories
```

```
-----
[11] Category: 1 Weight: 0.676
[12] Category: 2 Weight: 0.324
-----
```

File AUEsp.cmta contains the tabled processed by MCA
It has 40 rows and 12 columns (categories)

```
-----
DiagoRC: General program for two diagonal inner product analysis
Input file: AUEsp.cmta
--- Number of rows: 40, columns: 12
-----
```

Total inertia: 2

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+6.5078E-01	+0.3254	+0.3254	02	+3.9996E-01	+0.2000	+0.5254
03	+2.7875E-01	+0.1394	+0.6647	04	+2.6107E-01	+0.1305	+0.7953
05	+1.9953E-01	+0.0998	+0.8950	06	+1.0739E-01	+0.0537	+0.9487
07	+6.2354E-02	+0.0312	+0.9799	08	+4.0158E-02	+0.0201	+1.0000
09	+0.0000E+00	+0.0000	+1.0000	10	+0.0000E+00	+0.0000	+1.0000
11	+0.0000E+00	+0.0000	+1.0000	12	+0.0000E+00	+0.0000	+1.0000

File AUEsp.cmvp contains the eigenvalues and relative inertia for each axis

--- It has 12 rows and 2 columns

File AUEsp.cmco contains the column scores

--- It has 12 rows and 5 columns

..... etc.

File AUEsp.cmli contains the row scores

--- It has 40 rows and 5 columns

..... etc.

CorRatioMCA: Correlation ratios after a MCA

Title of the analysis: AUEsp.cm

Number of rows: 40, columns: 4

Variable : 1

```
> Categ= 1 Weight= 0.438 0.981 -0.131 0.169 -0.249 -0.128
> Categ= 2 Weight= 0.446 -0.793 0.462 -0.365 -0.288 0.053
> Categ= 3 Weight= 0.117 -0.651 -1.274 0.762 2.031 0.277
-----> r= 0.751 0.292 0.140 0.546 0.017
```

..... etc.

Variable : 4

```
> Categ= 1 Weight= 0.676 -0.616 0.033 -0.030 -0.123 0.111
> Categ= 2 Weight= 0.324 1.288 -0.069 0.062 0.258 -0.232
-----> r= 0.794 0.002 0.002 0.032 0.026
```

File AUEsp.cmrc contains the correlation ratios between the categorical variables and the factor scores

It has 4 rows and 5 columns

Utiliser l'option **Stars** du module **ScatterClass** pour réaliser la Fig. 7A :

Stars			
XY coordinates file		AUEsp.cml	40 5
X-axis column number (default = 1)			
Y-axis column number (default = 2)			
Categories file (.cat)		AUEsp.cat	
Variable label file (or #)		AUCode_VarEsp	
Row weights file (option)		AUEsp.cmpl	40 1
Draw points (no = 2)			
<input type="button" value="Quit"/> <input type="button" value="Copy graph"/> <input type="button" value="Save graph"/> <input type="button" value="Print graph"/>			<input type="button" value="Draw"/>

L'utilisation du fichier AUCode_VarEsp requiert quelques manipulations préalables sous un éditeur de texte (**BEdit Lite** par exemple). On ne gardera que les étiquettes de variables et on éliminera les étiquettes des modalités. Les rapports de corrélation (chiffres en italiques sur la Fig. 7A) sont ajoutés sous ClarisDraw™.

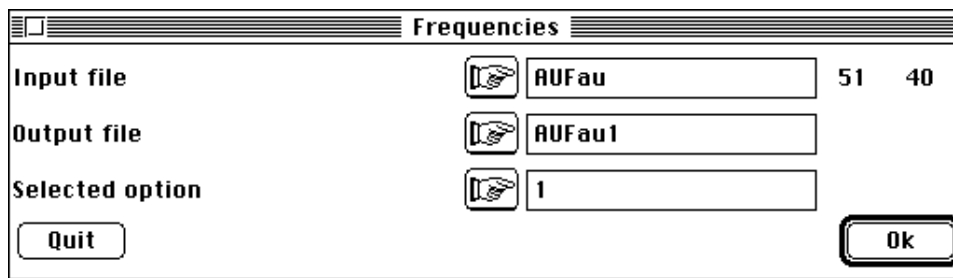
Pour représenter les coordonnées factorielles des espèces (carrés sur la Fig. 7B) utiliser l'option **Labels** du module **Scatters** :

Labels			
XY coordinates file		AUEsp.cml	40 5
X-axis column number (default = 1)			
Y-axis column number (default = 2)			
Label file (or # for item numbers)		AUCode_OisFR	
Draw vectors from origin (yes = 1)			
Draw unit circle (yes = 1)			
Draw points (no = 2)			
Constrain H/V ratio (yes = 1)			
<input type="button" value="Quit"/> <input type="button" value="Copy graph"/> <input type="button" value="Save graph"/> <input type="button" value="Print graph"/>			<input type="button" value="Draw"/>

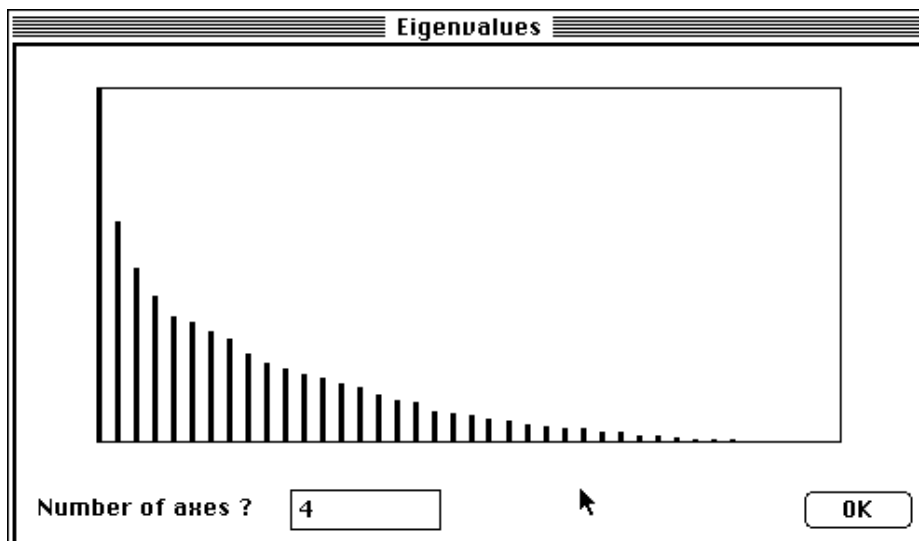
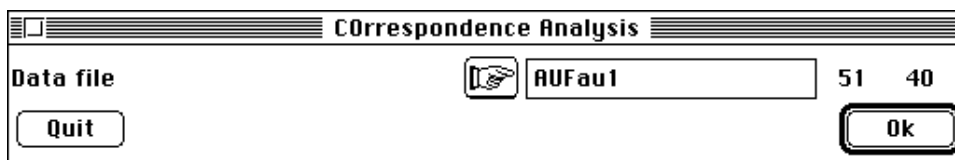
Appendice 3 - Analyses de co-inertie

L'analyse de co-inertie peut être vue comme un cas particulier de l'analyse RLQ (Fig. 8). On l'illustre par un test de randomisation sur l'analyse de co-inertie de AUMi1 (51-11) et AUFau (51-40) d'une part et l'analyse de co-inertie de AUEsp (40-4) et le transposé de AUFau (40-51) d'autre part.

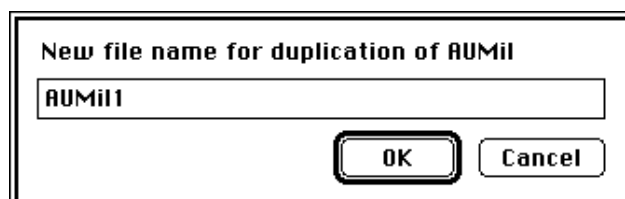
Utiliser l'option **Frequencies** du module **Bin->Bin** pour transformer le tableau AUFau (51-40) en un tableau de pourcentages en ligne :



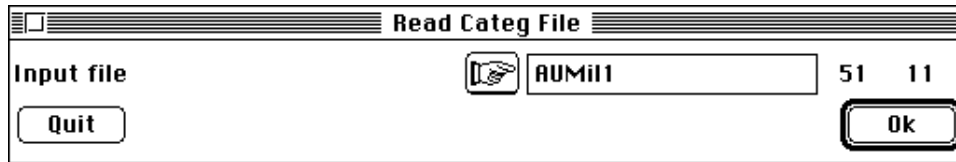
Faire l'AFC de AUFau1 (51-40) avec l'option **COrrrespondence Analysis** du module **COA** :



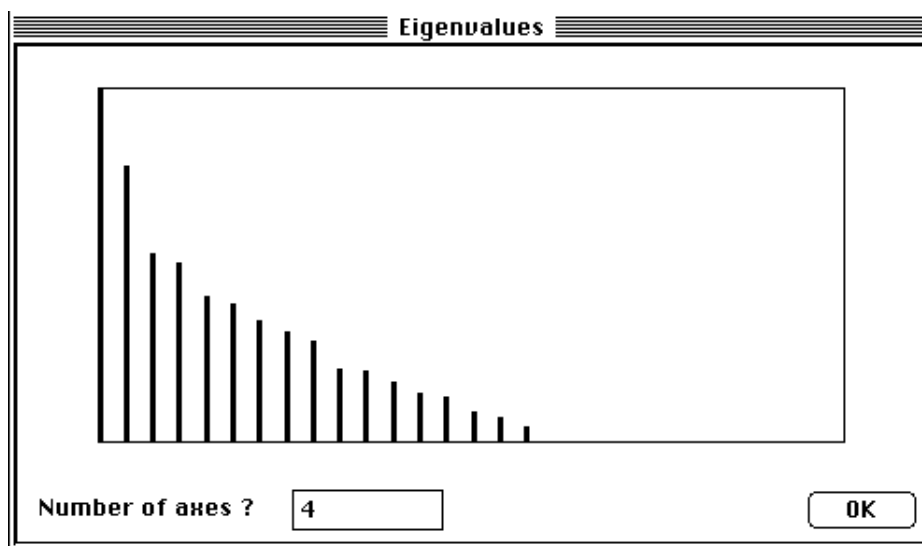
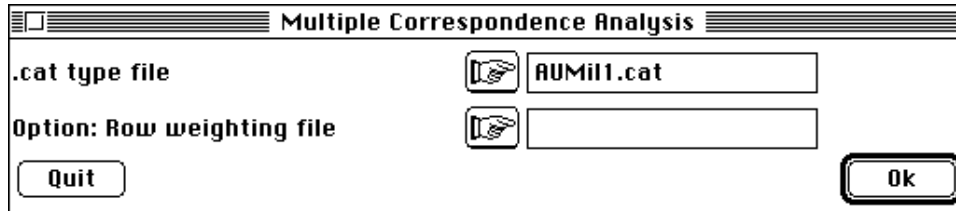
A l'aide de l'option **Duplicate** et du fichier AUMi1 (51-11) faire un nouveau fichier AUMi11 (51-11) :



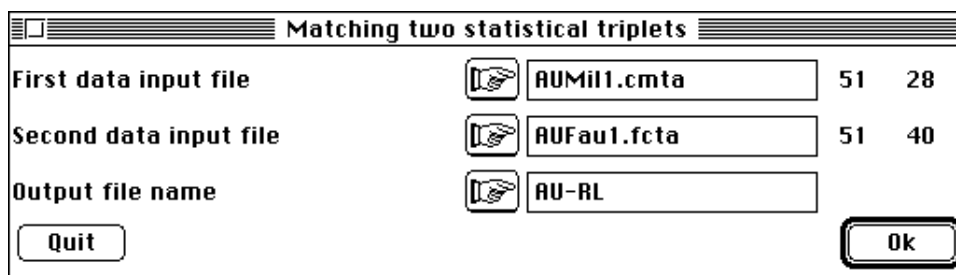
Relire ce fichier à l'aide de l'option **Read Categ File** du module **CategVar** :



Enchaîner avec l'analyse des correspondances multiples de AUMi11 en utilisant l'option **Multiple Correspondence Analysis** du module **MCA** et en conservant la pondération uniforme :



Préparer l'analyse de co-inertie entre "R" et "L" en utilisant l'option **Matching two statistical triplets** du module **CoInertia** :



```
First input statistical triplet: table AUMi11.cmta
Number of rows: 51, columns: 28
Second input statistical triplet: table AUFau1.fcta
Number of rows: 51, columns: 40
```

```
File AU-RL.<>ta contains the crossed table YtDnX
It has 40 rows and 28 columns
```

```
File AU-RL.<>pl contains the weight of each row
It has 40 rows and 1 column
```

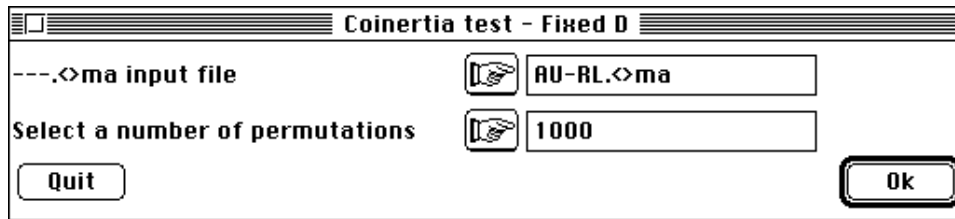
File AU-RL.<>pc contains the weight of each column
It has 28 rows and 1 column

File AU-RL.<>ma contains the name of the input statistical triplet

- 1 ---> AUMi11.cmta
- 2 ---> AUFau1.fcta

total inertia: 0.157821

Enchaîner avec le test en prenant l'option **Coinertia test - Fixed D** (cf. Fig. 8, pondération fixée *a priori*).



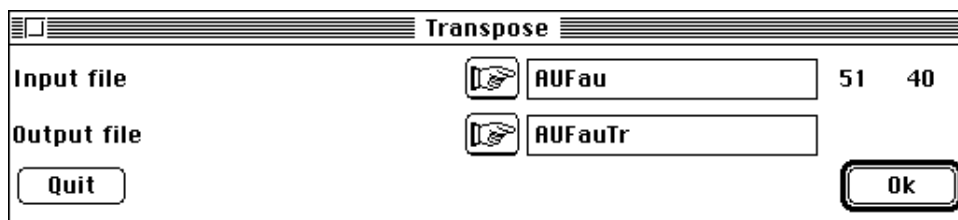
number of random matching: 1000 Observed: 0.157821
Histogramm: minimum = 0.071851, maximum = 0.157821
number of simulation X<Obs: 1000 (frequency: 1.000000)
number of simulation X>=Obs: 0 (frequency: 0.000000)

```

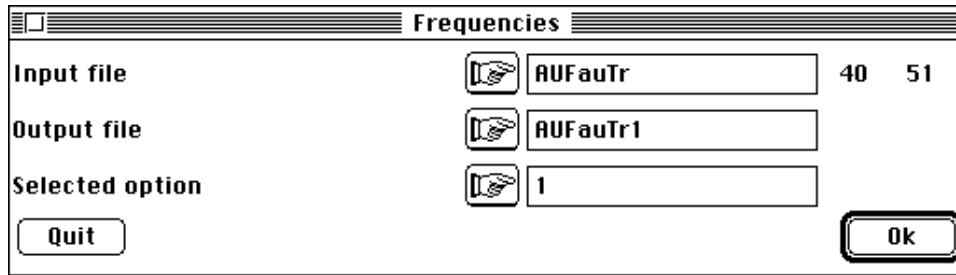
****
*****
*****
*****
*****
*****
*****
*****
*****
*****
*****
*****
*
•->

```

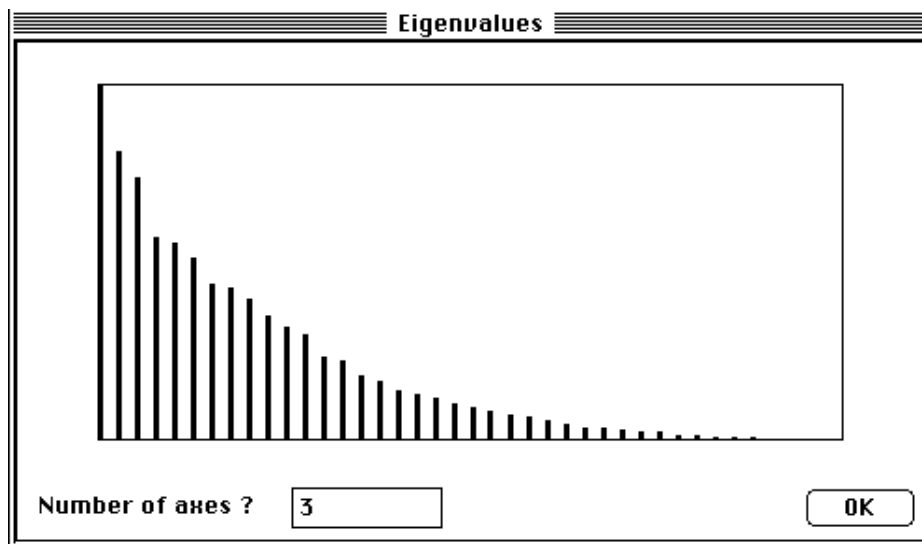
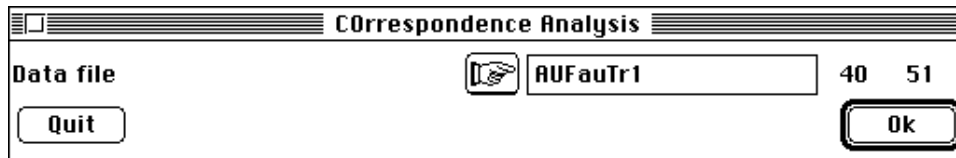
Une analyse de co-inertie peut être également réalisée avec le tableau AUEsp (40-4) et le transposé de AUFau (40-51). Utiliser l'option **Transpose** du module **FilesUtil** pour transformer le tableau AUFau (51-40) en un tableau AUFauTr (40-51) :



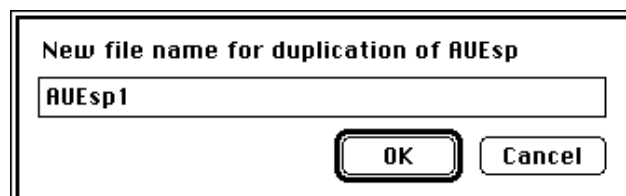
Utiliser l'option **Frequencies** du module **Bin->Bin** pour transformer le tableau AUFauTr (51-40) en un tableau de pourcentages en ligne :



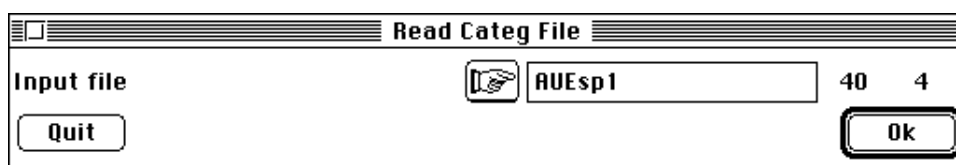
Faire l'AFC de AUFauTr1 (40-51) avec l'option **CORrespondence Analysis** du module **COA** :



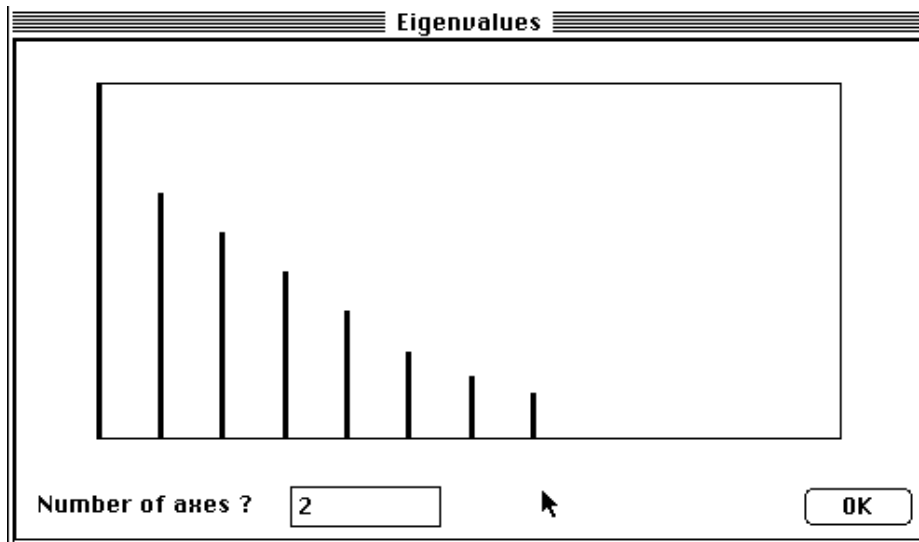
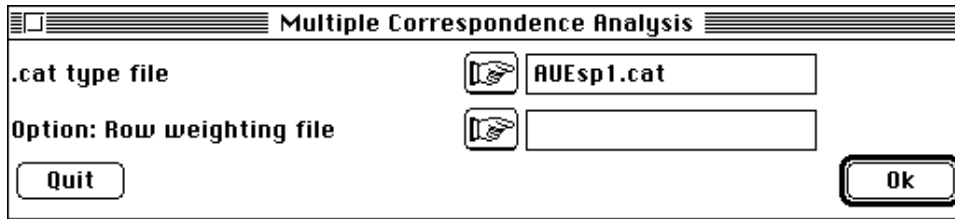
A l'aide de l'option **Duplicate** et du fichier AUEsp (40-4) faire un nouveau fichier AUEsp1 (40-4) :



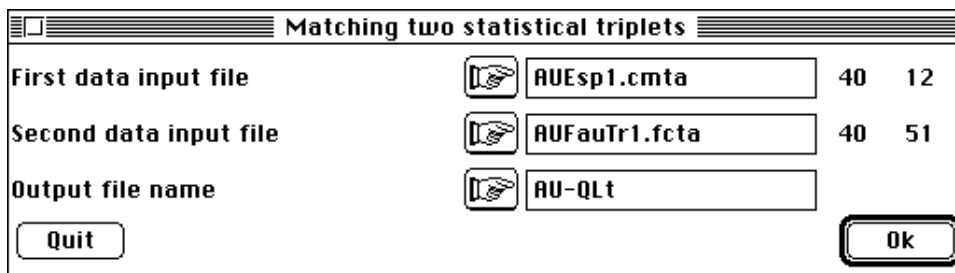
Relire ce fichier à l'aide de l'option **Read Categ File** du module **CategVar** :



Enchaîner avec l'analyse des correspondances multiples de AUEsp1 en utilisant l'option **Multiple Correspondence Analysis** du module **MCA** et en conservant la pondération uniforme :



Préparer l'analyse de co-inertie "L" et "Q" en utilisant l'option **Matching two statistical triplets** du module **CoInertia** :



```
First input statistical triplet: table AUEsp1.cmta
Number of rows: 40, columns: 12
Second input statistical triplet: table AUFauTr1.fcta
Number of rows: 40, columns: 51
```

```
File AU-QLt.<>ta contains the crossed table YtDnX
It has 51 rows and 12 columns
```

```
File AU-QLt.<>pl contains the weight of each row
It has 51 rows and 1 column
```

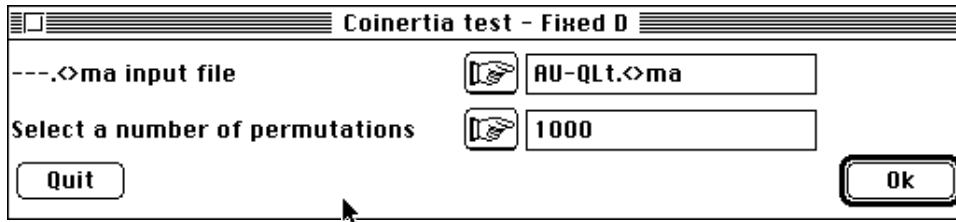
```
File AU-QLt.<>pc contains the weight of each column
It has 12 rows and 1 column
```

```
File AU-QLt.<>ma contains the name of the input statistical
triplet
```

```
1 ---> AUEsp1.cmta
2 ---> AUFauTr1.fcta
```

```
total inertia: 0.397260
```

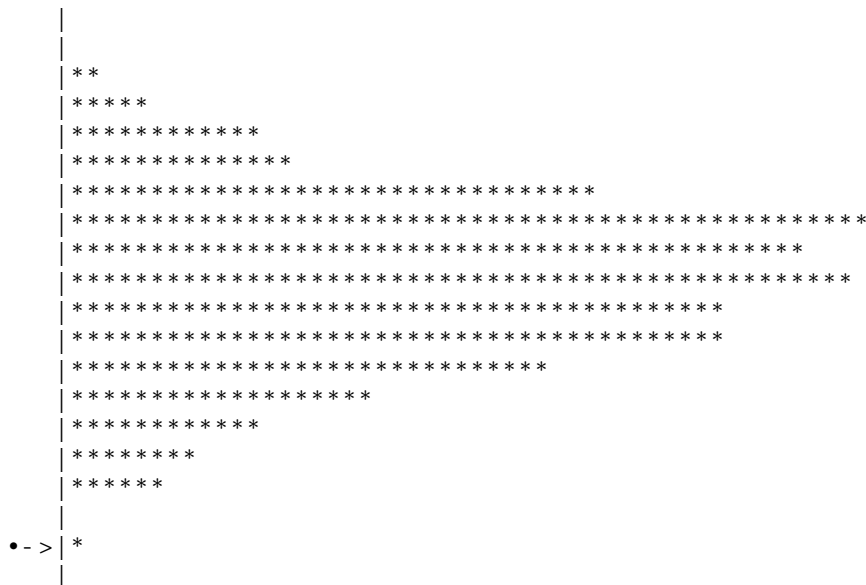
Enchaîner avec le test en prenant l'option **Coinertia test - Fixed D** (cf. Fig. 8, pondération fixée *a priori*).



```

number of random matching: 1000   Observed: 0.397260
Histogramm:  minimum = 0.229770, maximum = 0.413053
number of simulation X<Obs: 996 (frequency: 0.996000)
number of simulation X>=Obs: 4 (frequency: 0.004000)

```



Appendice 4 - Analyse de co-inertie généralisée

Utiliser l'option **Prepare RLQ analysis** du module **RLQ** :

RLQ analysis of a set of three tables

R: first table AUMil.cmta (rows: 51, columns: 28)

L: link table AUFau.fcta (rows: 51, columns: 40)

Q: second table AUEsp.cmta (rows: 40, columns: 12)

File AURLQ.&&ta contains the crossed table RtpY
It has 28 rows and 12 columns

File AURLQ.&&pl contains the weight of each row
It has 28 rows and 1 column

File AURLQ.&&pc contains the weight of each column
It has 12 rows and 1 column

File AURLQ.&&ma contains the name of the input files

1 ---> AUMil.cmta

2 ---> AUFau.fcta

3 ---> AUEsp.cmta

total inertia: 0.016542

Cette opération prépare le triplet $(\mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0, \mathbf{D}_p, \mathbf{D}_q)$. On peut ensuite tester si l'appariement modalités d'habitat/traits biologiques a une signification statistique à l'aide d'un test non paramétrique.

Utiliser l'option **RLQ test - Fixed L** du module **RLQ** qui indique que l'on effectue un test sur l'inertie totale ($z = \text{Trace}(\mathbf{Z}^t \mathbf{D}_p \mathbf{Z} \mathbf{D}_q)$) avec $\mathbf{Z} = \mathbf{R}_0^t \mathbf{P}_0 \mathbf{Q}_0$ de l'analyse RLQ dans lequel les tableaux **R** et **Q** sont permutés au hasard tandis que le tableau central **L** reste fixé :

number of random matching: 1000 Observed: 0.016542

Histogramm: minimum = 0.002344, maximum = 0.036618

number of simulation X<Obs: 986 (frequency: 0.986000)

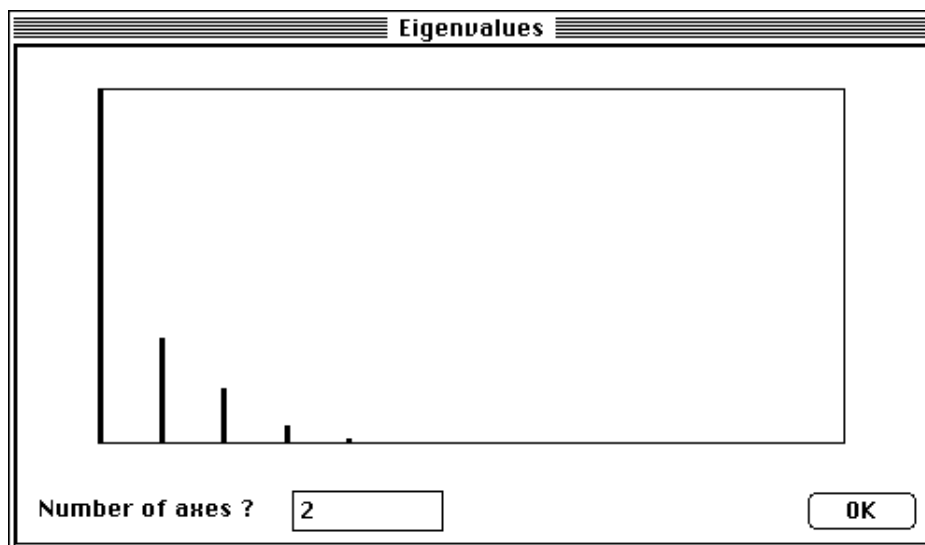
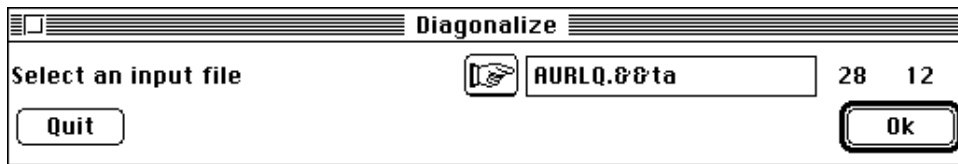
number of simulation X>=Obs: 14 (frequency: 0.014000)


```

*****
*****
*****
*****
*****
*****
*****
*****
*****
****
***
**
**
**
*
  ->

```

Diagonaliser le triplet à l'aide de l'option **Diagonalize** du module **RLQ** :



```

DiagoRC: General program for two diagonal inner product analysis
Input file: AURLQ.&&ta
--- Number of rows: 28, columns: 12

```

```
-----
Total inertia: 0.0165417
-----
```

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+1.0870E-02	+0.6571	+0.6571	02	+3.2239E-03	+0.1949	+0.8520
03	+1.6878E-03	+0.1020	+0.9541	04	+5.4385E-04	+0.0329	+0.9869
05	+1.2531E-04	+0.0076	+0.9945	06	+5.4464E-05	+0.0033	+0.9978
07	+1.9725E-05	+0.0012	+0.9990	08	+1.6637E-05	+0.0010	+1.0000
09	+0.0000E+00	+0.0000	+1.0000	10	+0.0000E+00	+0.0000	+1.0000
11	+0.0000E+00	+0.0000	+1.0000	12	+0.0000E+00	+0.0000	+1.0000

File AURLQ.&&vp contains the eigenvalues and relative inertia for each axis

--- It has 12 rows and 2 columns

File AURLQ.&&co contains the column scores

--- It has 12 rows and 2 columns

..... etc.

File AURLQ.&&li contains the row scores

--- It has 28 rows and 2 columns

..... etc.

La totalité des fichiers de coordonnées s'obtient par l'option **Coinertia analysis** du module **RLQ** :



RLQ analysis between two statistical triplets with COA matching

1 ---> AUMi1.cmta (rows: 51, col: 28, axes: 4, inertia: 1.545450)

2 ---> AUEsp.cmta (rows: 40, col: 12, axes: 5, inertia: 2.000000)

Coinertia: 0.016542, RV coefficient: 0.039488

AURLQ.&&w1 is a binary file with 28 rows and 2 columns

It contains the canonical weights of the variables of table 1

File : AURLQ.&&w1

----- Minimum/Maximum:

Col.: 1 Mini = -5.0635 Maxi = 2.6631

Col.: 2 Mini = -3.0706 Maxi = 5.3753

AURLQ.&&li is a binary file with 28 rows and 2 columns

It contains the coordinates of the columns of table 1

File : AURLQ.&&li

----- Minimum/Maximum:

Col.: 1 Mini = -0.52791 Maxi = 0.27765

Col.: 2 Mini = -0.17435 Maxi = 0.30521

AURLQ.&&w2 is a binary file with 12 rows and 2 columns

It contains the canonical weights of the variables of table 2

File : AURLQ.&&w2

----- Minimum/Maximum:

Col.: 1 Mini = -3.9073 Maxi = 1.2462

Col.: 2 Mini = -4.5377 Maxi = 2.394

AURLQ.&&co is a binary file with 12 rows and 2 columns

It contains the coordinates of the columns of table 2

File : AURLQ.&&co

----- Minimum/Maximum:

Col.: 1 Mini = -0.40737 Maxi = 0.12993

Col.: 2 Mini = -0.25765 Maxi = 0.13593

AURLQ.&&l1 is a binary file with 51 rows and 2 columns

It contains the coordinates of the rows (table 1)

File : AURLQ.&&l1

----- Minimum/Maximum:

Col.: 1 Mini = -0.89311 Maxi = 0.86358

Col.: 2 Mini = -1.0026 Maxi = 0.81373

AURLQ.&&l2 is a binary file with 40 rows and 2 columns

It contains the coordinates of the rows (table 2)

File : AURLQ.&&l2

```
----- Minimum/Maximum:
Col.: 1 Mini = -1.4134 Maxi = 0.47917
Col.: 2 Mini = -1.4194 Maxi = 1.4016
```

AURLQ.&&m1 is a binary file with 51 rows and 2 columns
It contains the normalized coordinates of the rows (table 1)
File : AURLQ.&&m1

```
----- Minimum/Maximum:
Col.: 1 Mini = -1.9009 Maxi = 1.838
Col.: 2 Mini = -2.5422 Maxi = 2.0632
```

AURLQ.&&m2 is a binary file with 40 rows and 2 columns
It contains the normalized coordinates of the rows (table 2)
File : AURLQ.&&m2

```
----- Minimum/Maximum:
Col.: 1 Mini = -2.4404 Maxi = 0.82733
Col.: 2 Mini = -2.3281 Maxi = 2.2989
```

AURLQ.&&a1 is a binary file with 4 rows and 2 columns
It contains the coordinates of the projections of inertia
axes onto the co-inertia axes (table 1)
In earlier version of ADE this file uses suffix cp1
File : AURLQ.&&a1

```
----- Minimum/Maximum:
Col.: 1 Mini = -0.30063 Maxi = 0.78678
Col.: 2 Mini = 0.2701 Maxi = 0.67721
```

AURLQ.&&a2 is a binary file with 5 rows and 2 columns
It contains the coordinates of the projections of inertia
axes onto the co-inertia axes (table 2)
In earlier version of ADE this file uses suffix cp2
File : AURLQ.&&a2

```
----- Minimum/Maximum:
Col.: 1 Mini = -0.43813 Maxi = 0.73702
Col.: 2 Mini = -0.61303 Maxi = 0.2645
```

```
-----
```

Num	Covaria.	Varian1	varian2	Correla.	INER1	INER2
1	0.1043	0.2208	0.3354	0.3831	0.2724	0.6508
2	0.05678	0.1555	0.3717	0.2361	0.2135	0.4

```
-----
```

```
-----
```

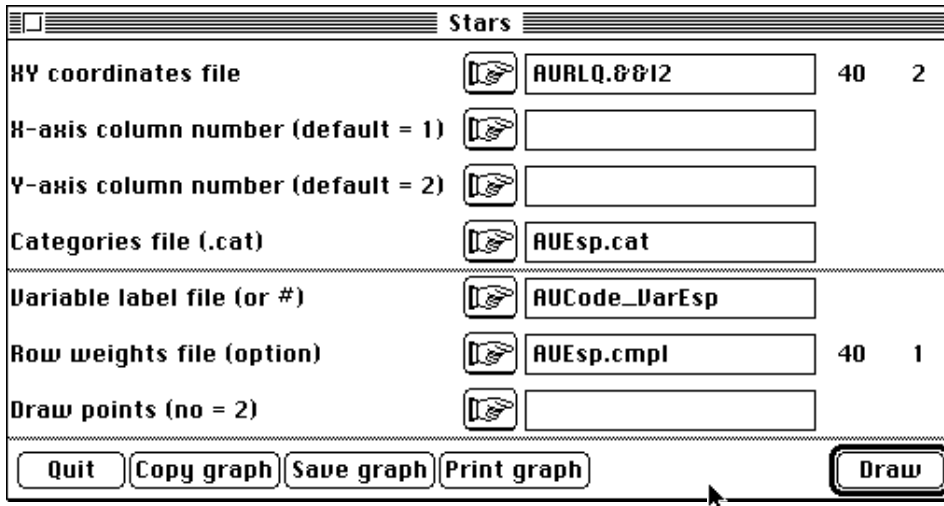
Num	Varian1	Varian2	INER2	INER2	
1	0.2208	0.3354	0.2724	0.6508	 Valeurs cumulées
2	0.3763	0.7072	0.4859	1.051	

```
-----
```

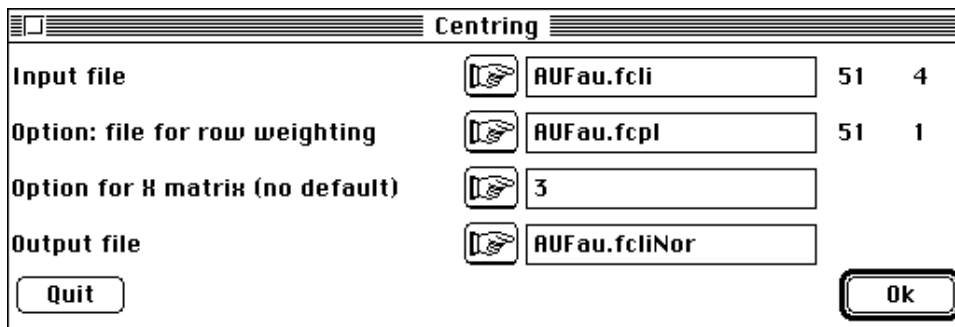
On peut ensuite réaliser un certain nombre de représentations graphiques. Par exemple, utiliser le fichier AURLQ.&&a1 et le module **Scatters** (option **Labels**) pour réaliser la Fig. 10.

De la même façon, Utiliser le fichier AURLQ.&&a2 et le module **Scatters** (option **Labels**) pour réaliser la Fig. 12.

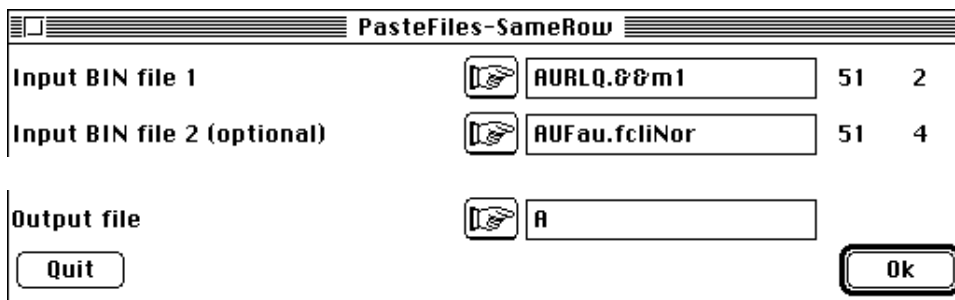
Les coordonnées des lignes de **R** (AUMi1) sont dans AURLQ.&&l1 et permettent à l'aide de l'option **Values** du module **Maps** de réaliser la Fig. 11. Les coordonnées des lignes de **Q** (AUEsp) sont dans AURLQ.&&l2 et permettent à l'aide de l'option **Stars** du module **ScatterClass** de réaliser la Fig. 13 :



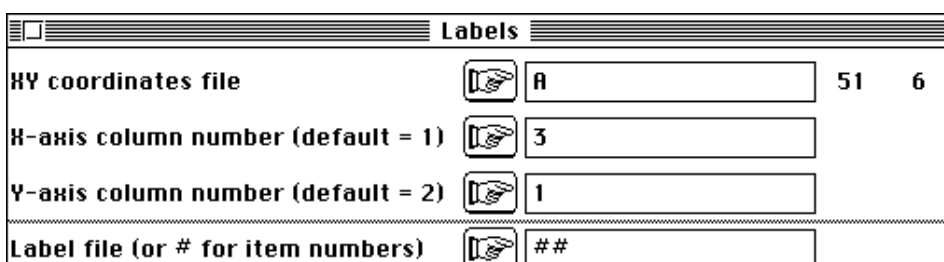
Pour comparer les résultats de l'analyse RLQ avec l'analyse du tableau faunistique, commencer par normaliser les coordonnées factorielles des lignes de AUFau (51-40) avec la pondération *ad hoc*. Utiliser l'option **Centring** du module **Bin->Bin** :



Coller ensemble le fichier AUFau.fcliNor et le fichier AURLQ.&&1 qui contient les coordonnées normalisée de **R** (AUMi 1) dans l'analyse RLQ. Pour cela utiliser l'option **PasteFiles - SameRow** du module **FilesUtil** :



La Fig. 14 est obtenue par l'option **Labels** du module **Scatters** :



Normaliser les coordonnées factorielles des colonnes de AUFau (51-40) avec la pondération *ad hoc*. Utiliser l'option **Centring** du module **Bin->Bin** :

Field	Value	Dimensions
Input file	AUFau.fcco	40 4
Option: file for row weighting	AUFau.fcpc	40 1
Option for K matrix (no default)	3	
Output file	AUFau.fccoNor	

Coller ensemble le fichier AUFau.fccoNor et le fichier AURLQ.&&m2 qui contient les coordonnées normalisée de **Q** (AUEsp) dans l'analyse RLQ. Pour cela utiliser l'option **PasteFiles - SameRow** du module **FilesUtil** :

Field	Value	Dimensions
Input BIN file 1	AURLQ.&&m2	40 2
Input BIN file 2 (optional)	AUFau.fccoNor	40 4
Output file	B	

La Fig. 15 est obtenue par l'option **Labels** du module **Scatters** :

Field	Value	Dimensions
HY coordinates file	B	40 6
X-axis column number (default = 1)	3	
Y-axis column number (default = 2)	1	
Label file (or # for item numbers)	AUCode_OisFR	

Appendice 5 - Dépouillement graphique

Le code numérique des modalités du tableau **R** (AuMi1) centré par variable et de variance moyenne (entre variables) unité est dans AURLQ.&&w1. Sélectionner la première colonne de ce fichier avec l'option **Row-Col Selection** du module **FilesUtil** :

Row-Col Selection

Input file: 28 2

Selection of rows (default = all):

Selection of columns (default = all):

Output file:

Pour tracer la Fig. 16A, utiliser l'option **Labels** du module **Graph1D** :

Labels

Data file (no default): 28 1

Rows label file (default = #):

Variable label file (or #):

Vertical (1) or horizontal (2) graphs:

Min/Max

Min. abscissa: Horiz. graphs:

Max. abscissa: Vert. graphs:

Min. ordinate: Nb. grad. X:

Max. ordinate: Nb. grad. Y:

Window height: G factor:

Window width:

Square drawing

Draw frame

Scale box

Row & col. selection

Col. selection:

Row selection method: File
 Keyboard

Row selection file (.cat):

Selection col. number:

Le paramétrage ci-dessus peut être modifié à loisir. Remarquer l'usage du fichier de sélection AuMi1Moda.cat. Le graphique résultant peut être retravaillé sous ClarisDraw™.

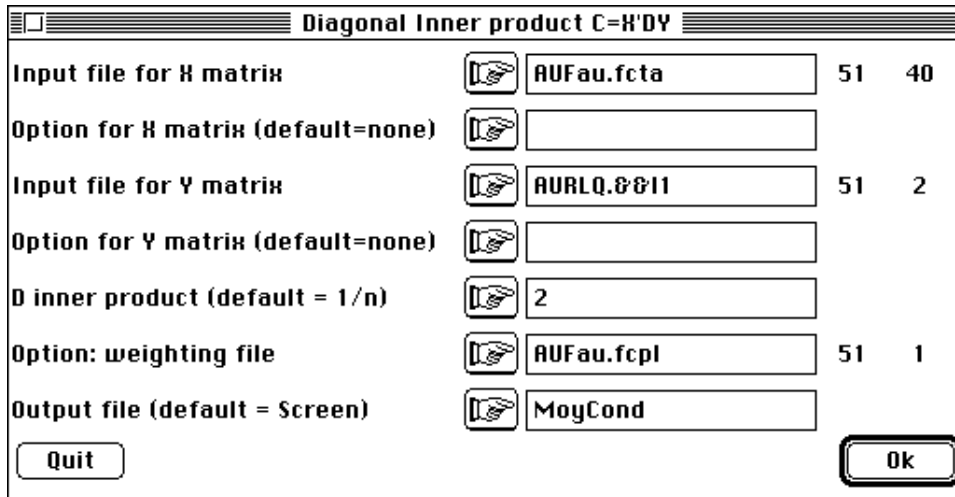
Pour tracer la Fig. 16B, utiliser également l'option **Labels** du module **Graph1D** :

Choisir les mêmes bornes de fenêtres logique et graphique que précédemment :

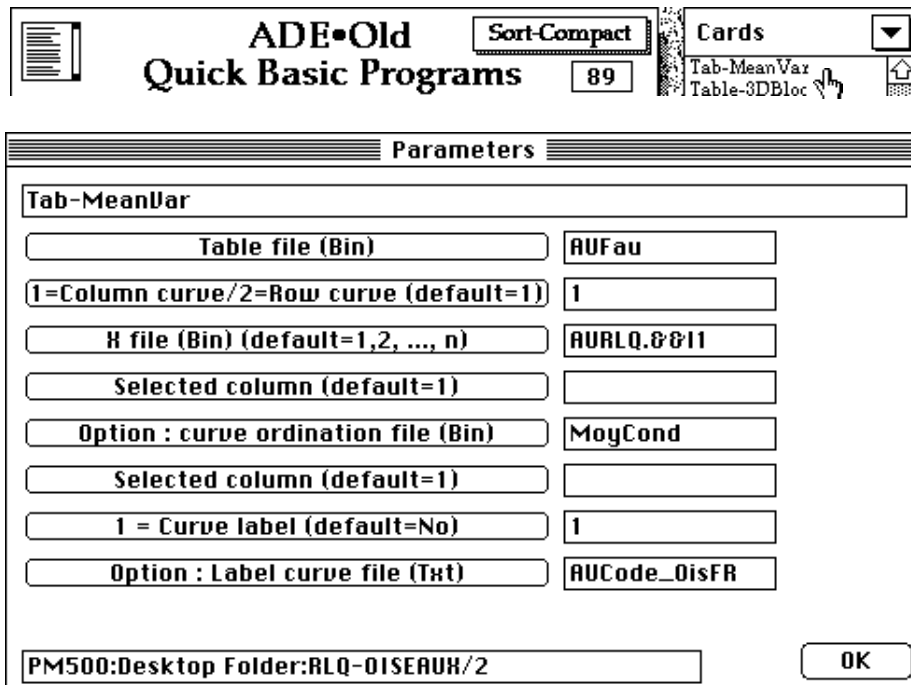
Dilater les bornes du graphiques sur l'intervalle [-1,1] pour obtenir la Fig. 16C :

L'obtention de la Fig. 17 nécessite au préalable le calcul de la position des espèces sur le gradient de milieu précédent à la moyenne conditionnelle des relevés qu'elles occupent.

Pour cela, utiliser l'option **Diagonal Inner Product $C=X'DY$** du module **MatAlg** :



Utiliser le programme Basic **TabMeanVar** via la pile **ADE•Old** pour tracer la Fig. 17 :



Rectangle window width :	400
Rectangle window height :	400
Mini H-axis:	-1
Maxi H-axis:	1
Mini Y-axis:	-1
Maxi Y-axis:	1

Noter que l'histogramme en bas de la Fig. 17 est repris de la Fig. 18. Utiliser le programme Basic **Histo1Categ** via la pile **ADE•Old** pour réaliser la Fig. 18 :



Le paramétrage suivant indique que l'on calcule le profil d'abondance des espèces par modalité par variable :

Categorical scoring diagram	
--.cat type file (Txt)	AUEsp.cat
Option : Row scoring file (Bin)	MoyCond
Selected column (default=1)	
Option : variable labels (Txt)	
Option : Row weighting file (Bin)	AUEsp.cmpl
Class number for histogram	10
Window width :	400
Window height :	600
Horizontal window number :	1
Vertical window number :	20
Mini H-axis:	-1
Maxi H-axis:	1
Mini Y-axis:	0
Maxi Y-axis:	.8

La Fig. 19 est tirée des illustrations précédentes sous ClarisDraw™.

Appendice 6 - Alternatives

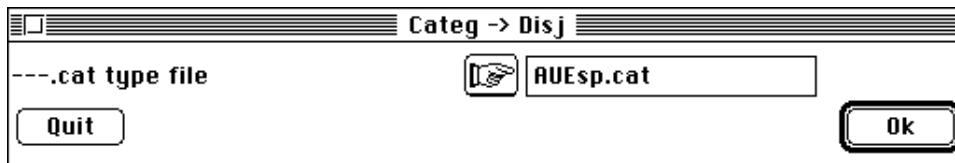
Du fait du type de codage qualitatif utilisé pour les tableaux \mathbf{R} (AUMi1) et \mathbf{Q} (AUEsp), on aurait pu ici utiliser une autre procédure que l'analyse générale RLQ en faisant l'AFC du tableau $\mathbf{R}^t\mathbf{LQ}$.

Commencer par transformer les deux tableaux qualitatifs en deux tableaux disjonctifs complets. Pour cela, utiliser l'option **Categ -> Disj** du module **CategVar** :



File AUMi101 contains dummy coded variable file
It has 51 rows and 28 columns (categories)

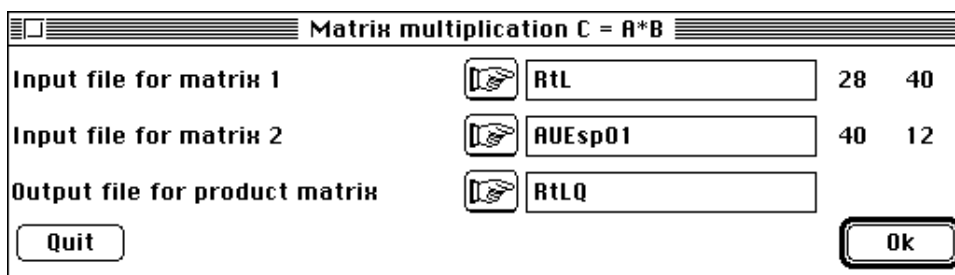
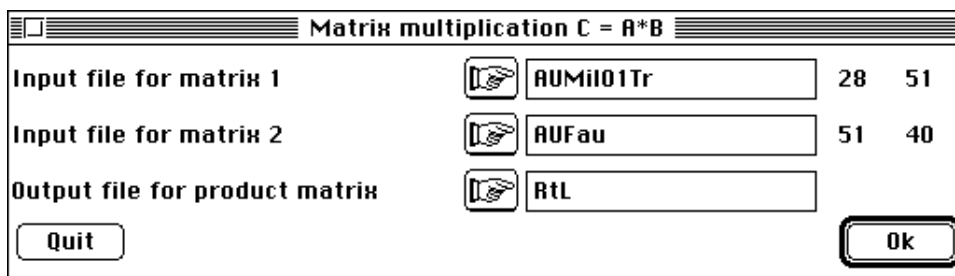
File AUMi101.num contains the number of modalities for each variable
It has 11 rows (variables) and 1 column



File AUEsp01 contains dummy coded variable file
It has 40 rows and 12 columns (categories)

File AUEsp01.num contains the number of modalities for each variable
It has 4 rows (variables) and 1 column

Transposer le fichier AUMi101 (51-28) en un fichier AUMi101Tr (28-51). Enchaîner avec l'option **Matrix multiplication C = A*B** du module **MatAlg** successivement :

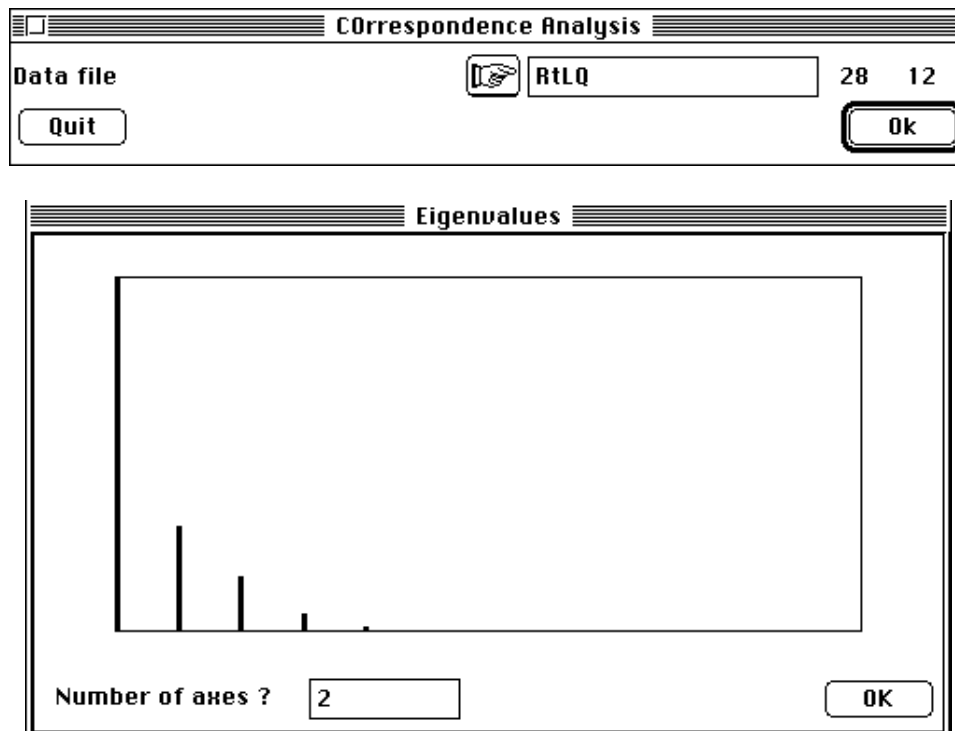


On obtient un fichier $\mathbf{R}^t\mathbf{LQ}$ (28-12) qui contient pour une modalité de milieu (ligne) et une modalité de trait biologique (colonne) la somme d'abondance des espèces

porteuses de ce couple. On peut lister ce fichier à l'aide de la commande **Edit with** et le logiciel d'accès BBEdit Lite :

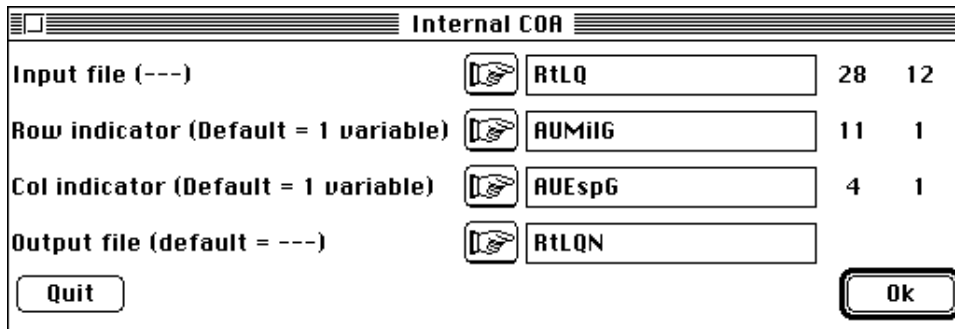
170	146	41	239	86	32	27	212	49	69	230	127
160	190	47	282	72	43	33	216	72	76	280	117
68	95	24	130	50	7	1	128	20	38	132	55
262	241	64	391	108	68	59	300	101	107	378	189
40	54	16	79	28	3	1	75	13	21	78	32
290	282	72	442	130	72	59	353	108	124	432	212
71	71	13	113	30	12	26	86	19	24	110	45
259	265	75	408	128	63	34	342	102	121	400	199
184	139	45	248	72	48	48	176	71	73	234	134
146	197	43	273	86	27	12	252	50	72	276	110
74	83	17	116	43	15	6	121	22	25	116	58
256	253	71	405	115	60	54	307	99	120	394	186
186	211	53	317	83	50	26	258	78	88	316	134
144	125	35	204	75	25	34	170	43	57	194	110
226	244	67	366	122	49	23	325	85	104	363	174
104	92	21	155	36	26	37	103	36	41	147	70
53	60	17	90	32	8	11	76	11	32	90	40
277	276	71	431	126	67	49	352	110	113	420	204
40	64	9	77	31	5	2	85	7	19	81	32
290	272	79	444	127	70	58	343	114	126	429	212
30	24	11	44	0	21	9	13	28	15	44	21
54	45	19	80	19	19	12	51	30	25	77	41
69	42	12	77	36	10	14	75	11	23	71	52
33	37	8	55	16	7	2	50	11	15	56	22
34	41	11	59	23	4	1	58	10	17	60	26
42	55	15	77	31	4	2	73	12	25	76	36
23	45	9	59	18	0	1	56	6	14	57	20
45	47	3	70	15	10	19	52	13	11	69	26

Faire l'AFC de ce tableau (option **CO**rrespondence **A**nalysis du module **COA**) :

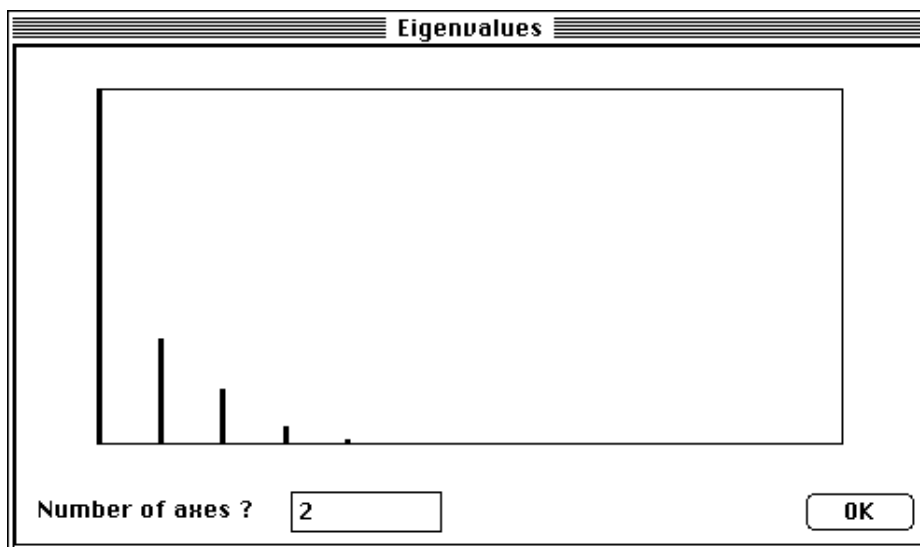


Le graphe des valeurs propres est identique à celui de l'analyse RLQ. De plus, l'inertie totale de cette analyse est égale à l'inertie totale de l'analyse RLQ (0.0165).

L'analyse RLQ est également l'analyse des correspondances internes du tableau R^tLQ . Pour le voir, utiliser l'option **Internal COA** du module **COA** :



On obtient un graphe des valeurs propres identique :



Le listing suivant montre que l'inertie totale est encore égale à celle de l'analyse RLQ. Par ailleurs cette analyse permet de mesurer le rôle des variables dans ce jeu de données.

```

Internal correspondence analysis
File RtLQnou.wwpl contains the row weights
It has 28 rows and 1 column
File RtLQnou.wwpc contains the column weights
It has 12 rows and 1 column
Row bloc number : 11
---> Row number in a bloc in file : AUMiIG
Col bloc number : 4
---> Row number in a bloc in file : AUEspG
File RtLQnou.wwil contains class indicator
It has 28 rows and 1 column
File RtLQnou.wwic contains class indicator
It has 12 rows and 1 column
File RtLQnou.wwta contains the doubly centered matrix
It has 28 rows and 12 column
-----
DiagoRC: General program for two diagonal inner product analysis
Input file: RtLQnou.wwta
--- Number of rows: 28, columns: 12
-----
Total inertia: 0.0165417

```

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+1.0870E-02	+0.6571	+0.6571	02	+3.2239E-03	+0.1949	+0.8520
03	+1.6878E-03	+0.1020	+0.9541	04	+5.4385E-04	+0.0329	+0.9869
05	+1.2531E-04	+0.0076	+0.9945	06	+5.4464E-05	+0.0033	+0.9978
07	+1.9725E-05	+0.0012	+0.9990	08	+1.6637E-05	+0.0010	+1.0000
09	+0.0000E+00	+0.0000	+1.0000	10	+0.0000E+00	+0.0000	+1.0000
11	+0.0000E+00	+0.0000	+1.0000	12	+0.0000E+00	+0.0000	+1.0000

File RtLQNou.wvvp contains the eigenvalues and relative inertia for each axis

--- It has 12 rows and 2 columns

File RtLQNou.wvco contains the column scores

--- It has 12 rows and 2 columns

..... etc.

File RtLQNou.wvli contains the row scores

--- It has 28 rows and 2 columns

..... etc.

Bl oc:	1	0.0909	0.000483	0.00146
Bl oc:	2	0.0909	0.0154	0.000248
Bl oc:	3	0.0909	0.00736	0.000182
Bl oc:	4	0.0909	0.000837	0.00547
Bl oc:	5	0.0909	0.0191	0.00104
Bl oc:	6	0.0909	0.00457	1.47E-05
Bl oc:	7	0.0909	0.000186	0.00465
Bl oc:	8	0.0909	0.00961	0.00324
Bl oc:	9	0.0909	0.00139	0.000686
Bl oc:	10	0.0909	0.0115	0.000204
Bl oc:	11	0.0909	0.0492	0.0183
-----			0.0109	0.00322

Ei genval ue percentage

Bl oc:	1	0.00404	0.0411
Bl oc:	2	0.129	0.00698
Bl oc:	3	0.0615	0.00513
Bl oc:	4	0.007	0.154
Bl oc:	5	0.159	0.0294
Bl oc:	6	0.0382	0.000414
Bl oc:	7	0.00155	0.131
Bl oc:	8	0.0804	0.0913
Bl oc:	9	0.0116	0.0193
Bl oc:	10	0.096	0.00575
Bl oc:	11	0.411	0.515

Bl oc:	1	0.25	0.00375	0.00189
Bl oc:	2	0.25	0.0122	0.00208
Bl oc:	3	0.25	0.027	0.00864
Bl oc:	4	0.25	0.00052	0.000287
-----			0.0109	0.00322

Ei genval ue percentage

Bl oc:	1	0.0863	0.147
Bl oc:	2	0.28	0.161
Bl oc:	3	0.622	0.67
Bl oc:	4	0.012	0.0223

Références

- ¹ Tatibouet, F., Chessel, D., Broyer, J. & Lebreton, J.D. (1980) Etude des peuplements d'oiseaux nicheurs de la zone urbaine de Lyon. Rapport final du Contrat Ecologie urbaine n° 237-01-78-00314. Ministère de l'Environnement. 106-156.
- ² Tatibouet, F. (1981) Approche écologique d'un établissement humain (environnement et structure). Exemple de la communauté urbaine de Lyon. Thèse de 3^o cycle, Université Lyon 1, 135 p.
- ³ Statzner, B., Resh, V. & Dolédec, S. (1994) Ecology of the Upper Rhône River: a test of habitat template theories. Special issue of *Freshwater Biology* (Editeurs invités). 31, 253-554
- ⁴ Legay, J.M. & Pontier, D. (1985) Relation âge-fécondité dans les populations de Chats domestiques, *Felis catus*. *Mammalia* : 49, 3, 395-402.
- ⁵ Williams, E.J. (1952) Use of scores for the analysis of association in contingency tables. *Biometrika* : 39, 274-289.
- ⁶ Hirschfeld, H.O. (1935) A connection between correlation and contingency. *Proceedings of the Cambridge Philosophical Society, Mathematical and Physical Sciences* : 31, 520-524.
- ⁷ Nishisato, S. (1980) *Analysis of categorical data : dual scaling and its applications*. University of Toronto Press, London. 1-276.
- ⁸ Benzecri, J.P. (1969) Statistical analysis as a tool to make patterns emerge from data. In : *Methodologies of pattern recognition*. Watanabe, S. (Ed.) Academic Press, New-York. 35-60.
- ⁹ Geary, R.C. (1954) The contiguity ratio and statistical mapping. *The incorporated Statistician* : 5, 3, 115-145.
- ¹⁰ Dolédec S., Chessel D. & Champely S. (1994) Using species distribution for matching species traits to environmental variables: a new three-table ordination method. Communication orale, *6th international congress of ecology (INTECOL) on ecological progress to meet the challenge of environmental change*, 20-26 août 1994, Manchester (Royaume Uni).
- ¹¹ Thioulouse, J. & Chessel, D. (1992) A method for reciprocal scaling of species tolerance and sample diversity. *Ecology* : 73, 670-680.
- ¹² Potvin, C. & Roff, D.A. (1993) Distribution-free and robust statistical methods: viable alternatives to parametric statistics? *Ecology* : 74, 1617-1628.
- ¹³ Escoufier, Y. (1973) Le traitement des variables vectorielles. *Biometrics* : 29, 750-760.
- ¹⁴ Robert, P. & Escoufier, Y. (1976) A unifying tool for linear multivariate statistical methods : the RV coefficient. *Applied Statistics* : 25, 3, 257-265.
- ¹⁵ Kazi-Aoual, F., Hitier, S., Sabatier, R. & Lebreton, J.D. (1994) Refined approximations to permutation tests for multivariate inference. *Computational Statistics and Data Analysis* : in press.

- ¹⁶ Fraile, L., Escoufier, Y. & Raibaut, A. (1993) Analyse des correspondances de données planifiées : Etude de la chémotaxie de la larve infestante d'un parasite. *Biometrics* : 49, 1142-1153.
- ¹⁷ Escoufier, Y. (1987) The duality diagram : a means of better practical applications. In : *Development in numerical ecology*. Legendre, P. & Legendre, L. (Eds.) NATO advanced Institute , Serie G .Springer Verlag, Berlin. 139-156.
- ¹⁸ Chessel, D. & Mercier, P. (1993) Couplage de triplets statistiques et liaisons espèces-environnement. In : *Biométrie et Environnement*. Lebreton, J.D. & Asselain, B. (Eds.) Masson, Paris. 15-44.
- ¹⁹ Mercier, P, Chessel, D. & Dolédec, S. (1992) Complete correspondence analysis of an ecological profile data table: a central ordination method. *Acta Œcologica* : 13, 25-44.
- ²⁰ Romane, F. (1972) Utilisation de l'analyse multivariable en Phytoécologie. *Investigación pesquera* : 36, 131-139.
- ²¹ Montana, C. & Greig-Smith, P. (1990) Correspondence analysis of species by environmental variable matrices. *Journal of Vegetation Science* : 1, 453-460.
- ²² Cazes, P. (1980) L'analyse de certains tableaux rectangulaires décomposé en blocs : généralisation des propriétés rencontrées dans l'étude des correspondances multiples. I. Définitions et applications à l'analyse canonique des variables qualitatives. II Questionnaires : variantes des codages et nouveaux calculs de contributions. *Les Cahiers de l'Analyse des Données* : 5, 145-161 & 387-406.
- ²³ Cazes, P., Chessel, D. & Doledec, S. (1988) L'analyse des correspondances internes d'un tableau partitionné : son usage en hydrobiologie. *Revue de Statistique Appliquée* : 36, 39-54.
- ²⁴ Whittaker, R.H. (1967) Gradient analysis of vegetation. *Biological Reviews* : 42, 207-264.
- ²⁵ Whittaker, R.H. (1973) *Ordination and classification of communities*. Hanbook of vegetation science. Part V. Dr. W. Junk b.v., The Hague. 1-737.
- ²⁶ Hill, M.O. (1973) Reciprocal averaging : an eigenvector method of ordination. *Journal of Ecology* : 61, 237-249.
- ²⁷ Ter Braak, C.J.F. (1986) Canonical correspondence analysis : a new eigenvector technique for multivariate direct gradient analysis. *Ecology* : 69, 69-77.
- ²⁸ Chessel, D., Lebreton, J.D. & Yoccoz, N. (1987) Propriétés de l'analyse canonique des correspondances. Une utilisation en hydrobiologie. *Revue de Statistique Appliquée* : 35, 4, 55-72.
- ²⁹ Lebreton, J.D., Sabatier, R., Banco, G. & Bacou, A.M. (1991) Principal component and correspondence analyses with respect to instrumental variables : an overview of their role in studies of structure-activity and species- environment relationships. In : *Applied Multivariate Analysis in SAR and Environmental Studies*. Devillers, J. & Karcher, W. (Eds.) Kluwer Academic Publishers. 85-114.

³⁰ Dolédec S., Chessel D., ter Braak C.J.F. & Champely S. (sous presse) Matching species traits to environmental variables: a new three-table ordination method (Environmental and Ecological Statistics).

³¹ Takane, Y. & Shibayama, T. (1991) Principal component analysis with external information on both subjects and variables. *Psychometrika* : 56, 97-120.

³² Böckenholt, U. & Böckenholt, I. (1990) Canonical analysis on contingency tables with linear constraints. *Psychometrika* : 55. 633-639.