

# Analyse d'inertie locale

## Résumé

La fiche décrit l'usage des analyses d'inertie locale dans plusieurs circonstances. D'un principe relativement simple, cette analyse peut donner des résultats semblables à ceux obtenus d'une analyse simple dans le cas de gradients multivariés. Dans d'autre cas on obtiendra des résultats complètement différents. Ce type d'analyse concentre l'attention sur la variabilité entre individus voisins et se place comme radicalement antinomique à la recherche des composantes cartographiables d'une enquête écologique à relevés régionalisés. Les exemples traités montrent comment structure locale et globale peuvent se superposer ou se mélanger dans une analyse simple.

## Plan

1 — ACP locale .....	2
1.1 — Préparation des données.....	2
1.2 — Test de Geary .....	4
1.3 — Maximiser la variance locale.....	5
1.4 — Produit scalaire local.....	8
1.5 — Le tableau faunistique.....	13
2 — Gradients biogéographiques.....	13
3 — Ordination locale.....	15
3.1 — Pavage de Voronoï .....	17
3.2 — Analyses simultanées .....	20
Références .....	25

D. Chessel et J. Thioulouse & S. Dolédec

# 1 — ACP locale

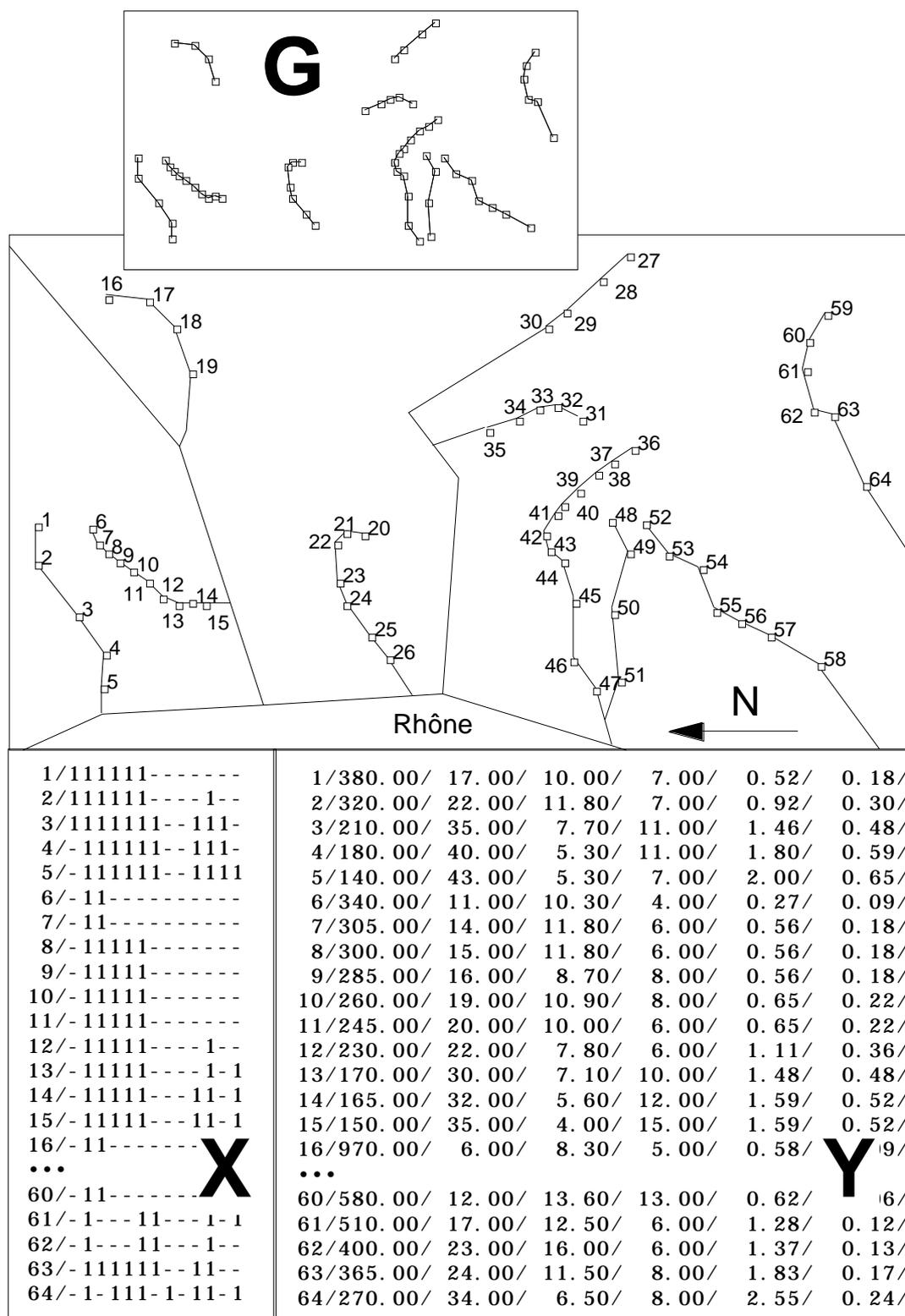
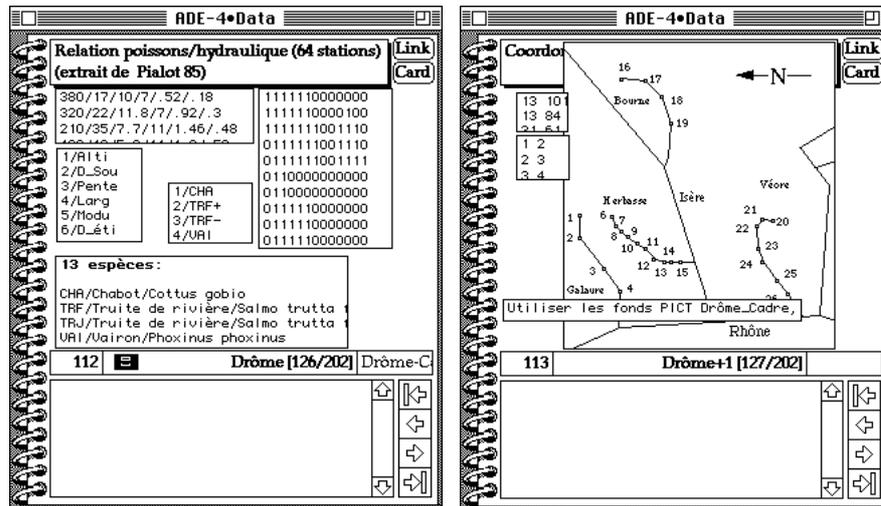


Figure 1 : Couple de tableaux écologiques et structure spatiale des stations dans un réseau hydrographique.

## 1.1 — Préparation des données

L'espace concret des mesures écologiques est ici effectivement pris en compte. On rencontre une énorme diversité de situations concrètes (placettes sur grille, transect, carte, prélèvements sur radiales, relevés le long des rivières, stratifications horizontales, verticales, temporelles, ...). La technique des graphes de voisinages permet d'enregistrer l'information spatiale avec un grande plasticité pratique qui a l'avantage de renvoyer à une théorie unifiée. Nous utiliserons un premier jeu hydrobiologique pour illustrer la généralité des techniques proposées. Il porte sur des stations dans un réseau hydrographiques. Créer un fichier de travail. Aller à la carte Drôme de la pile ADE•Data :

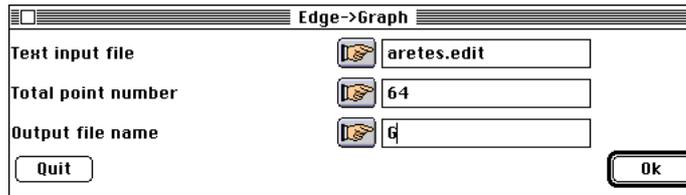


Envoyer les fichiers textes Mil.Txt, Code\_Var, Fau.car et Code\_Esp qui contiennent un couple classique de tableaux écologiques. Ces données sont extraites de la thèse de D. Pialot (1985<sup>1</sup>). Chacune des 76 stations donne un relevé de variables morphologiques (altitude, distance à la source, pente, largeur, module, débit d'étéage) et un relevé faunistique en présence-absence d'organismes de 13 catégories (TRF et TRJ désigne les truites adultes et juvéniles). Aller à la carte Drôme+1 de la pile ADE•Data :

Envoyer les fichiers textes xy.txt et aretes.txt. Passer les tableaux en binaire. Mil est fichier binaire 64-6, Fau est un fichier binaire 64-13 et XY est un fichier binaire 64-2. Dupliquer dans le dossier de travail les fichiers PICT Drôme-Cadre, Drôme-Carto et Drôme-Digi dans le dossier de travail. Vérifier le bon état du matériel en étiquetant le fond de carte avec Maps (option labels) et en listant les données (figure 1). Le dernier fichier aretes.txt sert enfin à décrire la structure de voisinage des stations. Vérifier que ce fichier texte contient une liste de couples :

Delta: ... Drôme: Aretes.txt	
1Δ	2↔
2Δ	3↔
3Δ	4↔
4Δ	5↔
5Δ	7↔
7Δ	8↔
8Δ	9↔
9Δ	10↔
10Δ	11↔
11Δ	12↔
12Δ	13↔
13Δ	14↔
14Δ	15↔

On y lit que la station 1 est voisine de la station 2, ... Chaque couple de points voisins est une arête du graphe de voisinage. Un graphe de voisinage est une matrice carrée ne comportant que des 1 (couples de points voisins) ou des 0 (couples de points non voisins). Dans NGUtil utiliser l'option Edge->Graph :



Le programme indique :

```
Neighbouring relationship
Input text file: aretes.edit
Neighbouring relationship in text file: G_G
It contains graph matrix (LEBART's M) with 64 rows and columns
Neighbouring weights in binary file: G_G.gpl
It contains 64 rows and 1 column
```

L'usage des graphes de voisinage en analyse des données a été introduit par Lebart (1969<sup>2</sup>). La figure 1 regroupe toutes les informations disponibles. On peut replacer les arêtes du graphe de voisinage sur le fond de carte par l'option Neighborhood Graphs du module Maps.

## 1.2 — Test de Geary

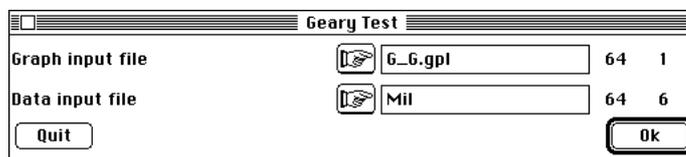
Deux questions distinctes interviennent dès qu'on manipule les structures spatiales. La première met en évidence ce qui se passe entre deux points voisins et renvoie aux notions de

**variance locale, inertie locale, co-inertie locale, indice de Geary.**

La seconde ignore ce qui se passe entre deux points voisins pour mettre en évidence ce qui varie le moins entre points voisins et renvoie aux notions de

**corrélation de voisinage, autocorrélation, autocorrélogramme, indice de Moran, composantes cartographiables, co-structure cartographiable.**

Les deux points de vue en univarié sont intimement liés. En multivarié par contre ils sont nettement distincts. La fiche Autocorrélation Spatiale Multidimensionnelle a présenté le second point de vue. On s'occupe ici du premier point exclusivement. En ce sens, les données qui servent d'illustration représentent la classe des problèmes que pose l'analyse des données hydrobiologiques quand elle aborde le continuum le long des rivières, tant au plan du milieu que des communautés. Pour un choix méthodologique conséquent on prendra connaissance des fiches ADE associées aux deux questions. Une connaissance élémentaire de la notion de variance locale est indispensable. Dans le module NGStat :



```
Geary Autocorrelation test
Access to neighbouring relationship: G_G.gpl
Data matrix: Mil
Point number: 64
```

N°	c observ.	Test N	Prob. x 10000	Test R	Prob. x 10000
1	1.009e-01	6.397e+00	0	6.244e+00	0
2	1.197e-01	6.264e+00	0	6.254e+00	0
3	2.397e-01	5.410e+00	0	5.143e+00	0
4	4.850e-01	3.664e+00	1	3.620e+00	1
5	1.597e-01	5.979e+00	0	5.898e+00	0
6	1.139e-01	6.305e+00	0	5.994e+00	0

Les estimations de la variabilité donnée par les couples de points voisins sont très inférieures aux variances estimées par tous les couples. Pour chacune des variables le rapport  $c$  est extrêmement significatif et il y a une forte autocorrélation spatiale (deux stations voisines se ressemblent). La même observation est vraie pour les données faunistiques :

```

Geary Autocorrelation test
Access to neighbouring relationship: G_G.gpl
Data matrix: Fau
Point number: 64

```

N°	c observ.	Test N	Prob. x10000	Test R	Prob. x10000
1	3.218e-01	4.825e+00	0	4.266e+00	0
2	7.485e-01	1.789e+00	368	1.628e+00	518
3	4.800e-01	3.700e+00	1	3.705e+00	1
4	5.147e-01	3.453e+00	3	3.590e+00	2
5	4.667e-01	3.795e+00	1	3.953e+00	0
6	3.271e-01	4.788e+00	0	4.880e+00	0
7	4.355e-01	4.017e+00	0	4.175e+00	0
8	7.778e-01	1.581e+00	569	1.277e+00	1007
9	4.080e-01	4.212e+00	0	3.152e+00	8
10	3.889e-01	4.348e+00	0	4.409e+00	0
11	2.928e-01	5.032e+00	0	5.275e+00	0
12	4.167e-01	4.150e+00	0	3.863e+00	1
13	5.867e-01	2.941e+00	16	2.945e+00	16

### 1.3 — Maximiser la variance locale

Dés que la plupart des variables présentent une forte structure spatiale, se pose la question de la synthèse multivariée. Si les indices de Geary ou certains d'entre eux sont significativement négatifs, on a affaire à la description de la structure locale (agrégativité ou variabilité entre voisins élevée). Si les indices de Geary sont significatifs positivement, la variance entre voisins est faible, la corrélation entre voisins est élevée. Si les indices de Geary ne sont pas significatifs, ce peut être par annulation des deux effets. Un effet d'autocorrélation positif (gradient) diminue la variabilité entre voisins, un effet d'autocorrélation négatif l'augmente. Au total, dans ce cas, il semble ne pas y avoir d'effets spatiaux parce qu'il y en a trop. Dans tous les cas, il est légitime de se demander si localement une variation d'une ou plusieurs variables s'accompagne systématiquement d'une variation de l'une ou plusieurs autres, dans le même sens ou en sens contraire.

Chercher les combinaisons de variables qui varient le plus entre voisins c'est faire de l'ordination locale. On s'appuie pour ce faire sur les propriétés habituelles des analyses linéaires, en substituant à l'optimisation de la variance celle de la variance locale. En effet, la relation fondamentale (Lebart 1969, op. cit.) consiste à considérer la variance locale comme une forme quadratique.

$$Var_{loc} = \left(\frac{1}{2}\right) \sum_{i \text{ voisin } j} p_i p_j (x_i - x_j)^2$$

La variance globale est simplement la variance locale du graphe complémentaire, celui qui définit deux points comme voisins si et seulement si ils ne sont pas voisins (un point n'est jamais voisin de lui-même).

**La question est de choisir des poids.** On peut imposer les poids de l'analyse simple dans le calcul de la variance locale ou au contraire utiliser le graphe de voisinage pour définir les poids de l'analyse simple. Cette dernière option a été privilégiée pour des raisons qui sont discutées dans la fiche "Autocorrélation spatiale multidimensionnelle". La variance locale utilisée est celle de Le Foll (1982<sup>3</sup>) :

$$VL(\mathbf{x}) = \left(\frac{1}{2}\right) \sum_{i,j} p_{ij} (x_i - x_j)^2 = \mathbf{x}^t (\mathbf{D} - \mathbf{P}) \mathbf{x}$$

où  $\mathbf{D}$  est la diagonale des poids de voisinage (nombre de voisins du point  $i$  sur le nombre total de couples de voisins).  $\mathbf{P}$  (table de contingence) est la matrice du graphe divisée par le nombre de couples de voisins ( $p_{ij}$  vaut  $1/m$  si  $i$  et  $j$  sont voisins, 0 sinon) et  $\mathbf{x}$  le vecteur des observations. La variance locale s'étend à la covariance locale :

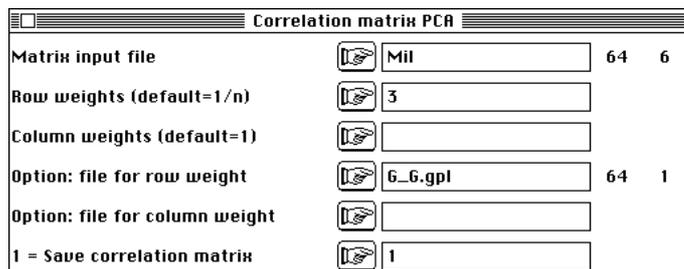
$$CL(\mathbf{x}, \mathbf{y}) = \left(\frac{1}{2}\right) \sum_{i,j} p_{ij} (x_i - x_j)(y_i - y_j) = \mathbf{x}^t (\mathbf{D} - \mathbf{P}) \mathbf{y}$$

et à l'inertie locale. De même que la variance s'étend à l'inertie, la variance locale s'étend à l'inertie locale (Carlier 1985<sup>4</sup>) :

$$IL = \left(\frac{1}{2}\right) \sum_{i,j} p_{ij} \|M_i - M_j\|_Q^2 = Trace(\mathbf{X}^t (\mathbf{D} - \mathbf{P}) \mathbf{X} \mathbf{Q})$$

En substituant aux notions de variance, covariance et inertie, les notions de variance locale, covariance locale et inertie locale, on généralise les analyses de base en focalisant l'attention sur ce qui se passe localement. Pour l'analyse de l'inertie locale des tableaux de l'exemple Drôme procéder à l'enchaînement suivant.

Exécuter l'ACP normée de Mil :



**Observer qu'il est impératif de donner à l'ACP la pondération associée au graphe (G\_G.gpl, fichier créé au moment de la lecture du graphe de voisinage).** Ceci veut dire qu'en tenant compte de la relation de voisinage un point est d'autant plus important qu'il a plus de voisins. On a gardé trois facteurs. L'analyse d'inertie locale, version ACP normée, est exécutée dans NGStat (option Geary Index Analysis) :



Le programme vérifie que la pondération du graphe de voisinage et la pondération de l'ACP sont les mêmes et affiche les valeurs propres. On obtient :

```
Geary Index linear analysis
Access to neighbouring relationship: G_G.gpl
Preliminary analysis: Mil.cnta
-----
File Mil_GA.xgcs contains local covariances
--- It has 6 rows and 6 columns
Num. Eigenval.  R.Iner.  R.Sum  | Num. Eigenval.  R.Iner.  R.Sum  |
01  +8.1347E-01 +0.7162 +0.7162  | 02  +3.0465E-01 +0.2682 +0.9845  |
03  +1.6666E-01 +0.1467 +1.1312  | 04  +6.9680E-02 +0.0614 +1.1926  |
05  +4.0725E-02 +0.0359 +1.2284  | 06  +1.8762E-02 +0.0165 +1.2450  |
```

File Mil\_GA.xgvp contains the eigenvalues  
 --- It has 6 rows and 1 column

File Mil\_GA.xgw1 contains weights of columns for row scores  
 --- It has 6 rows and 2 columns

File :Mil\_GA.xgw1

Col.	Mini	Maxi
1	-4.013e-01	7.016e-01
2	-3.733e-01	7.069e-01

Ce fichier contient les poids des variables (coefficients utilisés dans les combinaisons linéaires de variance locale maximale).

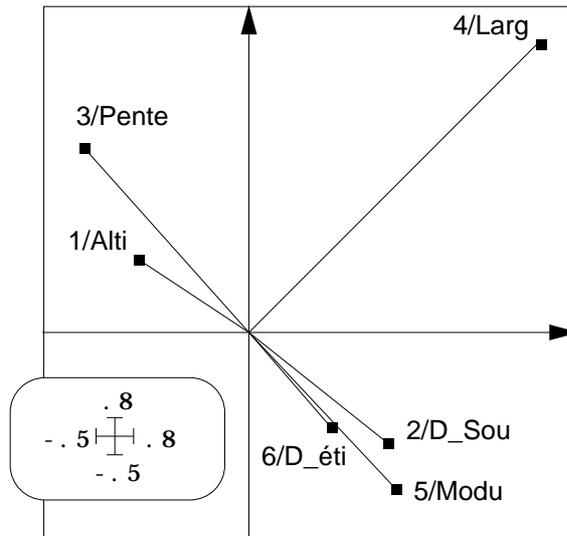


Figure 2 : poids canoniques des variables dans l'ACP locale de Mil (plan 1-2).

File Mil\_GA.xgli contains row scores  
 (linear combination having maximal local variance)  
 --- It has 64 rows and 2 columns

File :Mil\_GA.xgli

Col.	Mini	Maxi
1	-4.553e+00	3.915e+00
2	-2.297e+00	2.640e+00

Ce fichier contient les coordonnées des individus, c'est-à-dire les combinaisons linéaires de variance locale maximale. En maximisant la variance locale, on s'éloigne autant que peut se faire de la notion de corrélation spatiale, c'est-à-dire de *cartographiabilité* des résultats. On éloigne les points voisins, on mesure ce qui évolue entre voisins. Pour restituer l'intention utiliser **NeiScatter**. Avec les paramètres ci-dessous, on obtient la figure 3.

Neighbouring relationship		
HV coordinates file	<input type="button" value="..."/>	Mil_GA.xgli 64 2
H-axis column number (default = 1)	<input type="button" value="..."/>	<input type="text"/>
V-axis column number (default = 2)	<input type="button" value="..."/>	<input type="text"/>
Label file (or # for item numbers)	<input type="button" value="..."/>	# <input type="text"/>
Draw points (no = 2)	<input type="button" value="..."/>	<input type="text"/>
Neighbouring relationship	<input type="button" value="..."/>	G_G.gpl 64 1

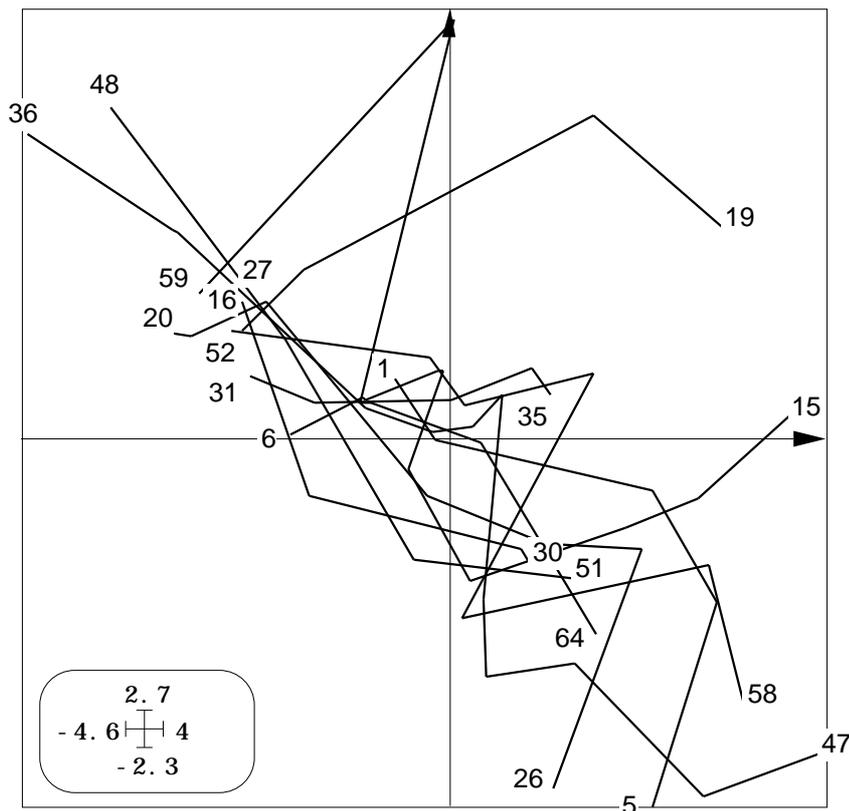


Figure 3 : carte factorielle (1-2) d'une ACP locale

Cette carte factorielle de la figure 3 se distingue radicalement d'une carte ordinaire d'ACP en ce que les deux coordonnées sont covariantes. Il faut comprendre qu'elles sont localement non covariantes. L'inertie locale projetée est représentée par la longueur des arêtes du graphe dont la somme des carrés est maximale. On optimise donc la lecture de l'évolution des paramètres le long de chacune des rivières, éventuellement l'originalité de cette évolution pour l'une ou l'autre de ces rivières, éventuellement l'évolution brutale des paramètres entre deux stations voisines (barrage). Il est donc clair que cette analyse très particulière est bien adaptée à la problématique hydrobiologique (*River continuum Concept*). Ces fonctions de variation locale optimale sont construites avec les poids des variables représentés dans la figure 2.

Dans la figure 2, on note l'identification des paramètres communs à l'évolution le long de chacune des rivières (pente, débit, altitude, distance à la source, module) et la particularité de la variable largeur.

Remarquons encore la contradiction apparente qui existe entre l'affirmation d'une autocorrélation positive très forte chez toute les variables (à l'exception justement de la largeur), donc d'une variabilité faible entre voisins et la mise en évidence dans cette variabilité d'intérêt limité (en principe) d'une organisation inter-variables et inter-stations forte. On a ici un cas typique d'analyse intra-classes (en fait inter voisins, c'est-à-dire implicitement intra rivières) dont l'inertie serait quantitativement petite et qualitativement fondamentale.

## 1.4 — Produit scalaire local

La variance, la covariance et la corrélation sont les dérivés du produit scalaire associé à la pondération des individus. On peut manipuler ces notions directement avec l'option Diagonal Inner Product du module MatAlg. Par exemple, à la question quelles sont les

variances et les covariances des coordonnées de l'analyse locale, on obtient la réponse par :

Diagonal Inner product C=K'DY			
Input file for K matrix	<input type="text" value="Mil_GA.xgli"/>	64	2
Option for K matrix (default=none)	<input type="text"/>		
Input file for Y matrix	<input type="text" value="Mil_GA.xgli"/>	64	2
Option for Y matrix (default=none)	<input type="text"/>		
D inner product (default = 1/n)	<input type="text" value="2"/>		
Option: weight file	<input type="text" value="6_6.gpl"/>	64	1
Output file (default = Screen)	<input type="text"/>		

qui affiche :

```
-----
[ 1] 3.0136e+00 -9.7886e-01
[ 2] -9.7886e-01 9.5835e-01
-----
```

La variance de la première coordonnée de l'analyse locale vaut 3.01. La même question posée à propos de l'analyse simple (ACP normée) :

Diagonal Inner product C=K'DY			
Input file for K matrix	<input type="text" value="Mil.cnli"/>	64	3
Option for K matrix (default=none)	<input type="text"/>		
Input file for Y matrix	<input type="text" value="Mil.cnli"/>	64	3
Option for Y matrix (default=none)	<input type="text"/>		
D inner product (default = 1/n)	<input type="text" value="2"/>		
Option: weight file	<input type="text" value="6_6.gpl"/>	64	1
Output file (default = Screen)	<input type="text"/>		

```
-----
[ 1] 3.4823e+00 -1.3500e-08 -4.6652e-09
[ 2] -1.3500e-08 8.6869e-01 1.3741e-09
[ 3] -4.6652e-09 1.3741e-09 8.1669e-01
-----
```

La variance de la première coordonnée de l'analyse simple vaut 3.48. Or cette variance vaut la première valeur propre de l'analyse simple et est optimale. On a donc en maximisant la variance locale trouvé une variable de synthèse de forte variance totale. Pour les corrélations, on demande simplement :

Diagonal Inner product C=K'DY			
Input file for K matrix	<input type="text" value="Mil.cnli"/>	64	3
Option for K matrix (default=none)	<input type="text" value="3"/>		
Input file for Y matrix	<input type="text" value="Mil.cnli"/>	64	3
Option for Y matrix (default=none)	<input type="text" value="3"/>		
D inner product (default = 1/n)	<input type="text" value="2"/>		
Option: weight file	<input type="text" value="6_6.gpl"/>	64	1

Les coordonnées en ACP ne sont pas corrélées :

```
[ 1] 1.0000e+00 -7.7617e-09 -2.7664e-09
[ 2] -7.7617e-09 1.0000e+00 1.6314e-09
[ 3] -2.7664e-09 1.6314e-09 1.0000e+00
```

Par contre les coordonnées de l'analyse locale sont corrélées :

Diagonal Inner product C=H'DY			
Input file for H matrix	<input type="text" value="Mil_6A.xgli"/>	64	2
Option for H matrix (default=none)	<input type="text" value="3"/>		
Input file for Y matrix	<input type="text" value="Mil_6A.xgli"/>	64	2
Option for Y matrix (default=none)	<input type="text" value="3"/>		
D inner product (default = 1/n)	<input type="text" value="2"/>		
Option: weight file	<input type="text" value="6_6.gpl"/>	64	1
Output file (default = Screen)	<input type="text"/>		

```
[ 1] 1.0000e+00 -5.7599e-01
[ 2] -5.7599e-01 1.0000e+00
```

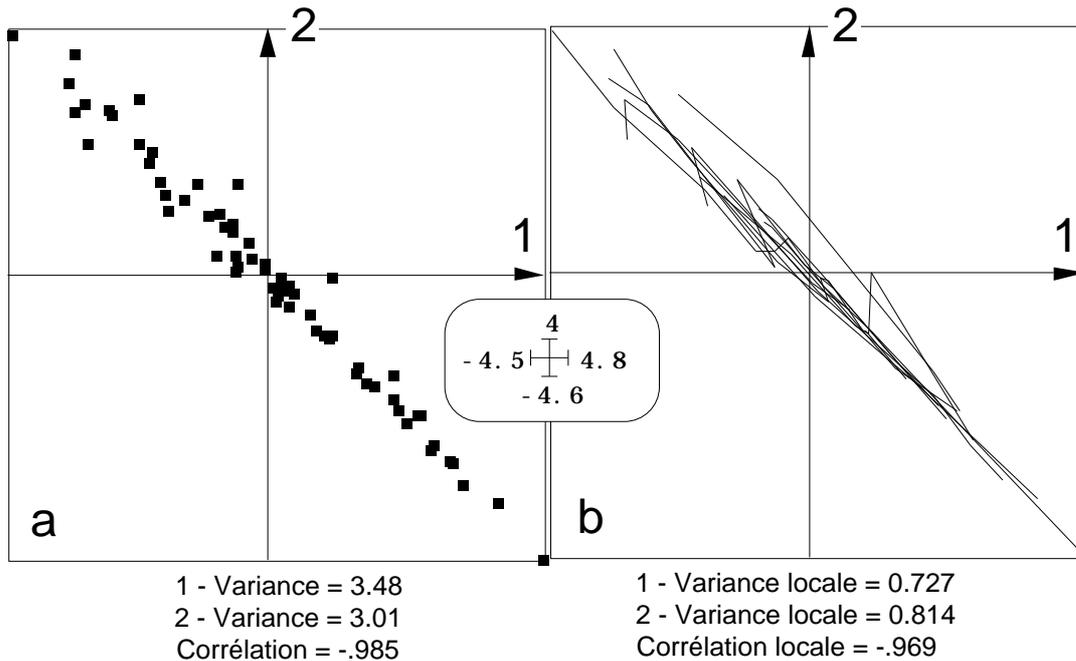


Figure 4 : comparaison des premiers facteurs d'une ACP normée et d'une ACP locale. 1- ACP normée ordinaire. 2 - ACP normée locale

Les mêmes calculs sont possibles avec la variance locale considérée comme produit scalaire avec l'option Local covariances du module NGStat. Quelles sont les variances et covariances locales des coordonnées de l'ACP locale :

Local covariances H[D-P]Y			
Input file for H matrix	<input type="text" value="Mil_6A.xgli"/>	64	2
Option for H matrix (default=none)	<input type="text"/>		
Input file for Y matrix	<input type="text" value="Mil_6A.xgli"/>	64	2
Option for Y matrix (default=none)	<input type="text"/>		
Graph input file	<input type="text" value="6_6.gpl"/>	64	1
Output file (default = Screen)	<input type="text"/>		

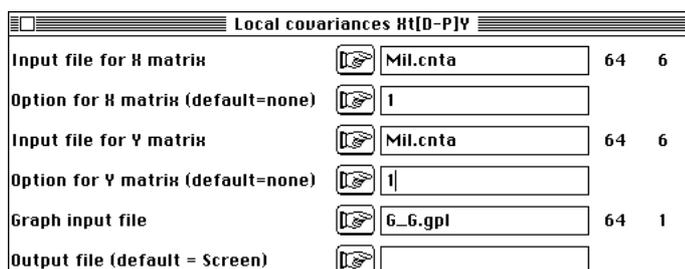
```
[ 1] 8.135e-01 | -6.373e-09 |
[ 2] -6.373e-09 | 3.046e-01 |
```

Ces variances locales sont les valeurs propres (les maxima possibles par combinaisons linéaires des variables normés avec des poids de somme de carrés unité) et les covariances locales sont nulles. Pour les coordonnées de l'ACP normée on a :

```
[ 1] 7.268e-01 | 7.693e-03 | -1.318e-01 |
[ 2] 7.693e-03 | 9.636e-02 | -3.232e-02 |
[ 3] -1.318e-01 | -3.232e-02 | 1.628e-01 |
```

Les covariances locales ne sont plus nulles mais la première variance locale est proche de l'optimum (0.727 pour un optimum de 0.814). Noter que les calculs des produits scalaires de voisinages sont beaucoup plus longs que ceux des produits scalaires ordinaires, parce qu'on utilise le graphe de voisinage. Le principe d'usage, par contre est le même.

La figure 4 identifie alors l'axe 1 d'inertie et l'axe 1 d'inertie locale. La question est alors : l'analyse locale est-elle la même que l'analyse ordinaire ? La réponse est certainement pas. La contrainte est forte. Mais, bien que la variabilité locale soit bien plus faible que la variabilité totale (autocorrélation) la structure de corrélation entre les variables dérive d'une corrélation locale. La matrice des corrélations locales est obtenue par :



La normalisation est faite avec la norme de Le Foll :

```
[ 1] 1.000e+00 | -7.854e-01 | 4.043e-01 | -4.398e-01 | -6.913e-01 | -5.901e-01 |
[ 2] -7.854e-01 | 1.000e+00 | -5.685e-01 | 4.676e-01 | 8.259e-01 | 4.912e-01 |
[ 3] 4.043e-01 | -5.685e-01 | 1.000e+00 | -3.395e-01 | -5.079e-01 | -2.807e-01 |
[ 4] -4.398e-01 | 4.676e-01 | -3.395e-01 | 1.000e+00 | 3.684e-01 | 2.401e-01 |
[ 5] -6.913e-01 | 8.259e-01 | -5.079e-01 | 3.684e-01 | 1.000e+00 | 5.452e-01 |
[ 6] -5.901e-01 | 4.912e-01 | -2.807e-01 | 2.401e-01 | 5.452e-01 | 1.000e+00 |
```

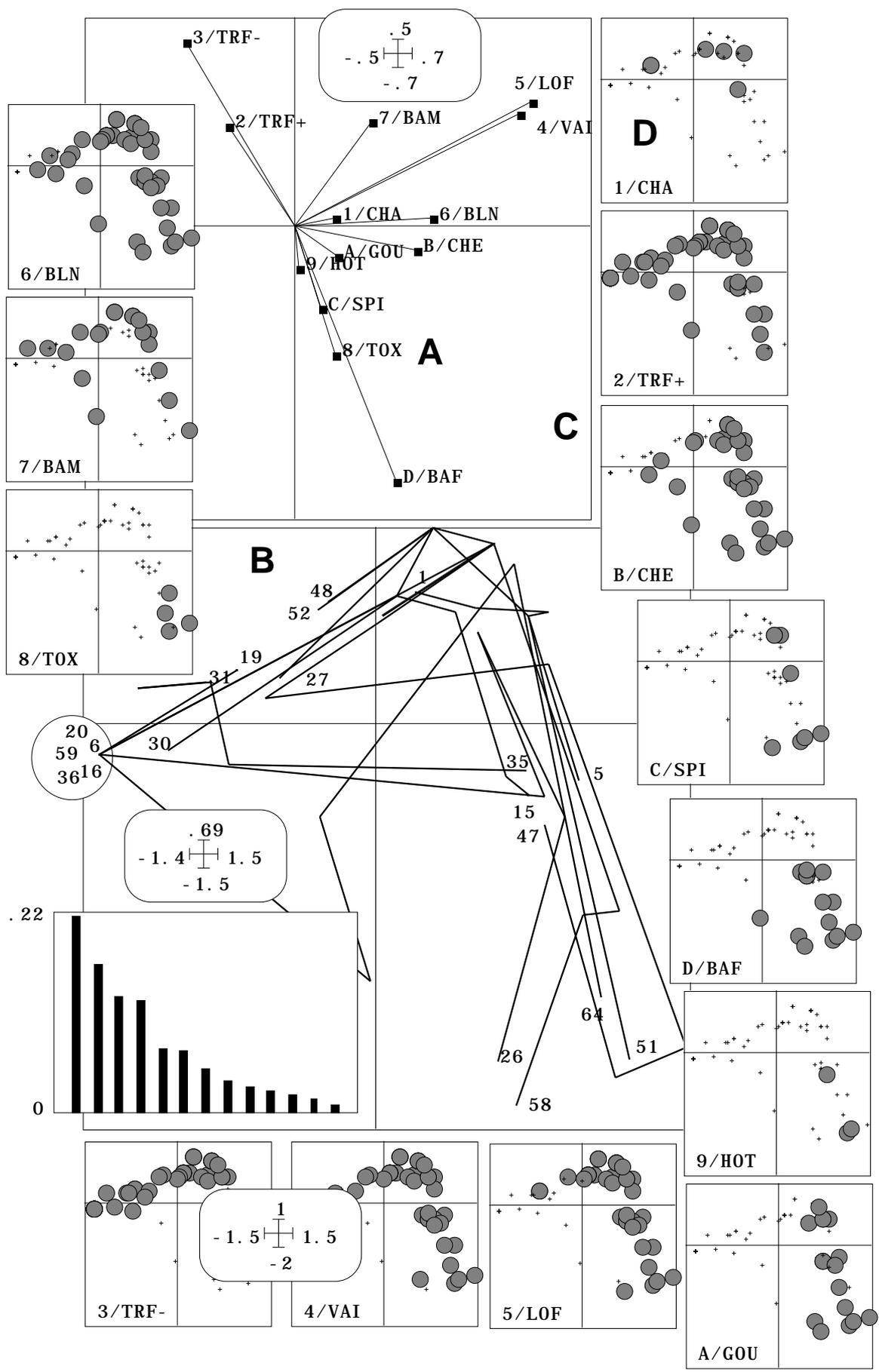
La matrice de corrélation ordinaire est :

```
[ 1] 1.000e+00 | -6.850e-01 | 3.913e-01 | -2.376e-01 | -3.408e-01 | -3.835e-01 |
[ 2] -6.850e-01 | 1.000e+00 | -6.905e-01 | 5.262e-01 | 8.371e-01 | 4.369e-01 |
[ 3] 3.913e-01 | -6.902e-01 | 1.000e+00 | -4.874e-01 | -6.570e-01 | -3.960e-01 |
[ 4] -2.376e-01 | 5.262e-01 | -4.874e-01 | 1.000e+00 | 4.878e-01 | 4.717e-01 |
[ 5] -3.408e-01 | 8.372e-01 | -6.570e-01 | 4.878e-01 | 1.000e+00 | 2.321e-01 |
[ 6] -3.835e-01 | 4.369e-01 | -3.960e-01 | 4.717e-01 | 2.321e-01 | 1.000e+00 |
```

On pourrait souhaiter faire une ACP normée locale, alors qu'on a fait une ACP sur matrice de covariances locales calculée sur des variables normalisées. Il s'agit là de raffinements théoriques. En général on trouvera de très fortes différences entre analyses locales et analyses ordinaires, sauf quand l'essentiel de la structure est locale. On retrouve le même phénomène dans deux situations opposées.

Quand l'essentiel de la structure est cartographiable, l'analyse ordinaire donne des cartes pratiquement optimales. Quand l'essentiel de la structure est locale l'analyse ordinaire restitue cette aspect des données (sans qu'on s'en aperçoive d'ailleurs). Dans ce cas l'identité entre analyse locale, globale et ordinaire prouve l'existence de structures en gradient multidimensionnel. Ici, il s'agit d'une évidence dans le cadre de l'expérience, mais ce n'est bien sur pas toujours le cas.

Figure 5 (page 12) : Dépouillement graphique d'une ACP locale. A - Graphe des valeurs propres. B - Carte f1-f2 des lignes et représentation des arêtes du graphe (maximisation de la variance locale). C - Poids canoniques des variables. D - Représentation de la présence-absence des taxons dans le plan F1-F2 de l'analyse locale.



## 1.5 — Le tableau faunistique

On enchaînera de même l'ACP centrée :

Covariance matrix PCA			
Matrix input file	<input type="text" value="Fau"/>	64	13
Row weights (default=1/n)	<input type="text" value="3"/>		
Column weights (default=1)	<input type="text"/>		
Option: file for row weight	<input type="text" value="6_6.gpl"/>	64	1
Option: file for column weight	<input type="text"/>		

et l'ACP centrée locale du tableau faunistique :

Geary Index Analysis			
Graph input file	<input type="text" value="6_6.gpl"/>	64	1
Input file	<input type="text" value="Fau.cpta"/>	64	13
<input type="button" value="Quit"/>		<input type="button" value="Ok"/>	

Utiliser l'option Neighborhood graphs du module Scatters pour exprimer la notion de variance locale sur une carte factorielle :

Neighbouring relationship			
HV coordinates file	<input type="text" value="Fau_GA.hgli"/>	64	2
H-axis column number (default = 1)	<input type="text"/>		
Y-axis column number (default = 2)	<input type="text"/>		
Label file (or # for item numbers)	<input type="text" value="#"/>		
Draw points (no = 2)	<input type="text" value="2"/>		
Neighbouring relationship	<input type="text" value="6_6.gpl"/>	64	1

On observera que la variabilité locale exprimée dans le plan 1-2 traduit une variance locale uniformément répartie sur l'espace étudié alors que le plan 3-4 exprime de la variance locale *locale* qui ne nous intéresse pas ici et qui est vraisemblablement de peu de signification. Le gradient faunistique est clairement exprimé par le dépouillement du plan 1-2 de l'analyse locale (figure 5).

On notera sur la figure 5 que la restitution du graphe de voisinage sur la carte factorielle rend cette dernière plus lisible comme expression d'une structure spatiale que la carte réelle effective. Il reste pour faire le bilan de la notion d'inertie locale à étendre cette notion à celle de co-inertie qui donnera les analyses de co-inertie locale.

## 2 — Gradients biogéographiques

Les relations entre analyses simples, analyses locales et analyses globales sont difficilement prévisibles dans les cas d'études concrets. Le matériel utilisé (végétation, invertébrés, oiseaux), les objectifs de l'étude, depuis la cartographie à vocation d'aménagement (sciences de l'environnement) jusqu'à la définition de la complexité (écologie du paysage) et les échelles de perception de l'espace de l'individu aux gradients biogéographiques proposent une multitude de cas possibles. Deux exemples utilisant l'avifaune nicheuse sont examinés ici, au titre d'illustrations méthodologiques. Le premier est celui de l'Atlas ornithologique Rhône-Alpes publié sous la direction de P. Lebreton<sup>5</sup>. Il propose une monographie de 60 districts naturels et la carte commentée de la répartition de 215 espèces.

On extrait arbitrairement 10 groupes d'espèces, les oiseaux aquatiques (13), les Corvidés (9), les Embéridés (6), les Fringillidés (11), les Ralidés et limicoles (11), les Motacillidés (7), les Paridés (8), les Rapaces (24), les Sylviinés (26) et les Turdidés (16) qui définissent un tableau croisant 60 districts et 131 espèces.

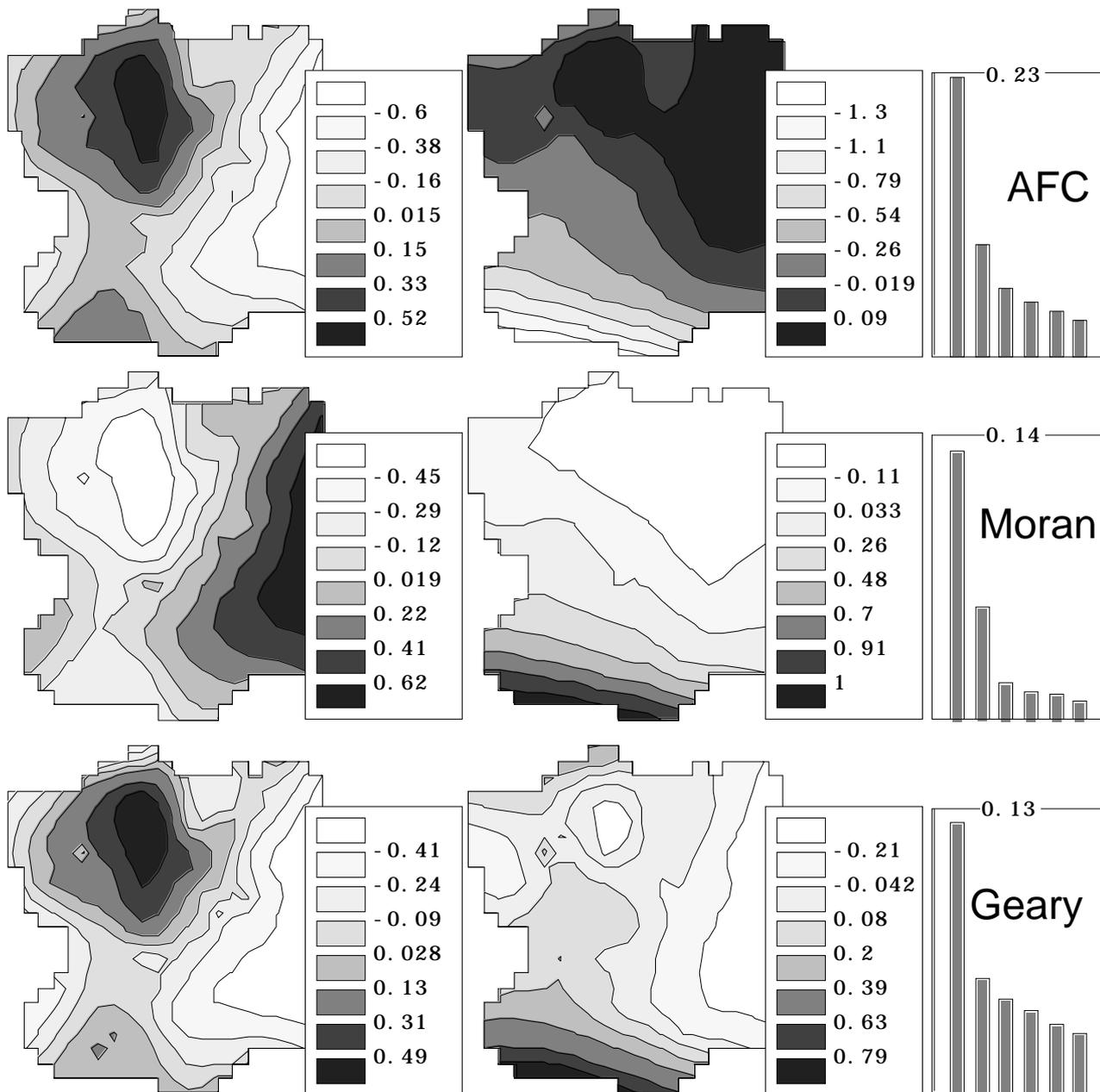
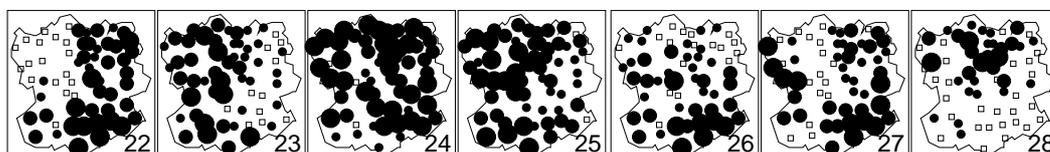


Figure 6 : Analyses des correspondances simple, locale (Geary) et globale (Moran). Cartes des deux premiers facteurs avec 9 voisins et graphe des 20 premières valeurs propres.

Le fond de carte est digitalisé sous forme de 60 unités surfaciques dans Carto-2D. Un fichier de coordonnées des districts en découle qui permet de représenter les 131 cartes de répartition, dont on a ci-dessous un aperçu :



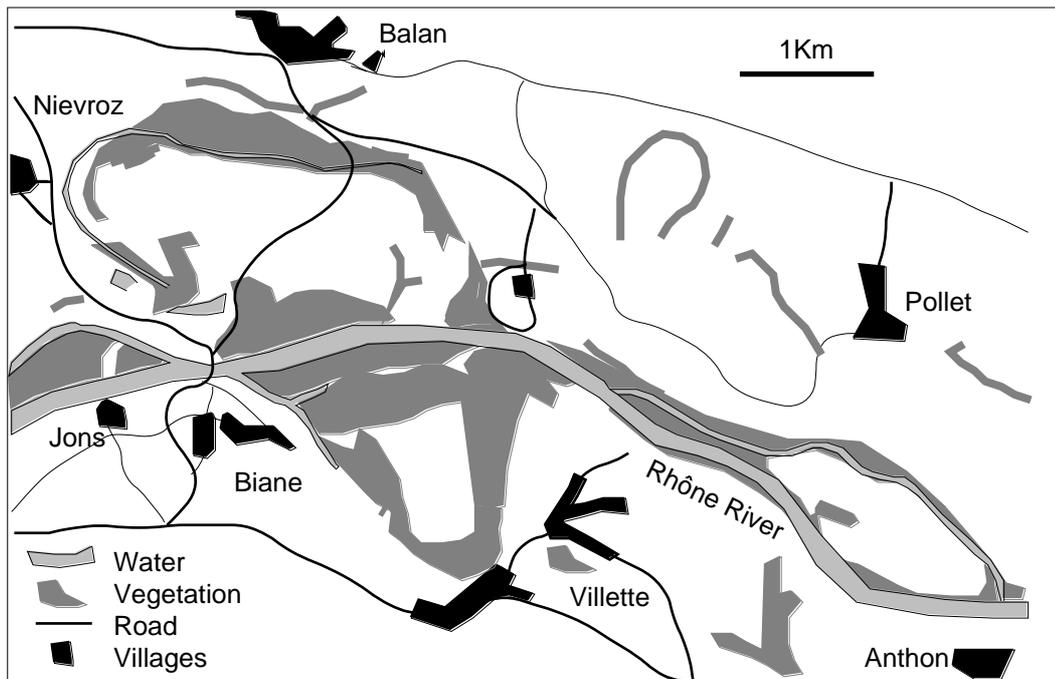
Le graphe de voisinage utilisé est celui de l'approche surfacique : deux districts ayant une frontière commune sont voisins. Les trois analyses donnent la figure 6.

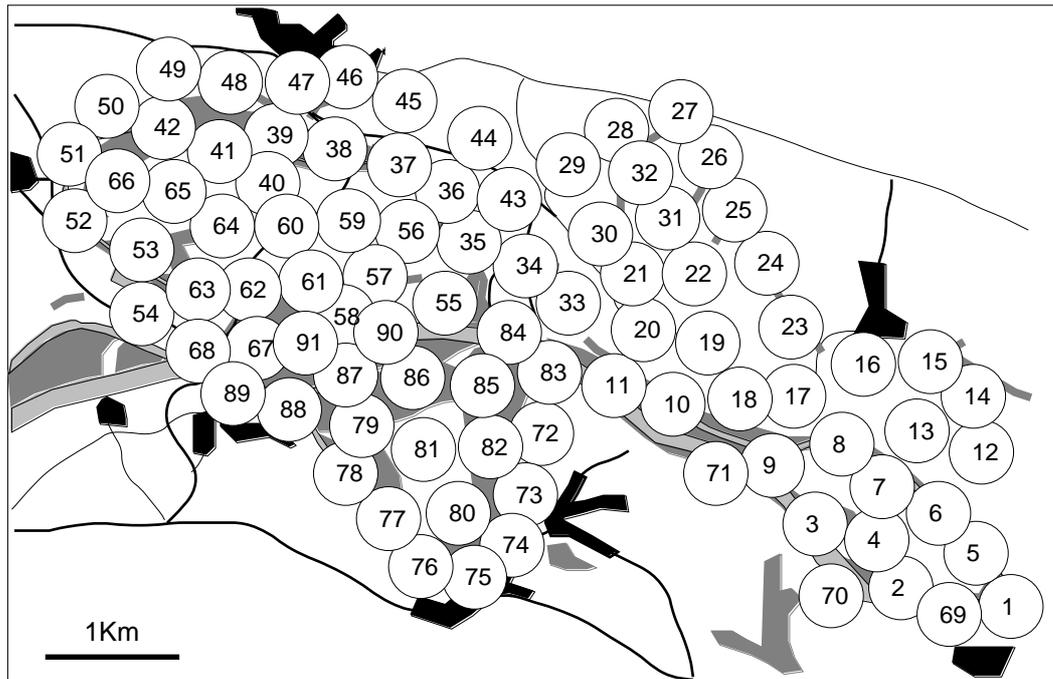
On remarquera l'identité presque parfaite entre les trois facteurs 1 en terme d'importance relative et d'expression de la structure spatiale (au signe près qui est aléatoire).

On reconnaît sans peine, sur les deux premières analyses, les deux gradients principaux en interaction, celui de l'altitude (F1) et celui de l'influence méditerranéenne (F2). Dans ce cas, l'analyse simple est identique à l'analyse globale, deux gradients constituant l'essentiel de l'ordination.

### 3 — Ordination locale

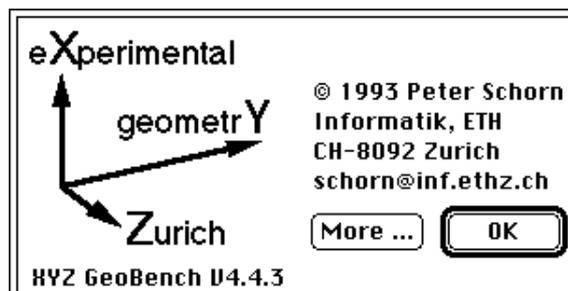
Ce n'est pas le cas de l'exemple suivant<sup>6</sup>. L'objectif est de relier complexité spatiale (horizontale et verticale) et diversité spécifique de l'avifaune sur plusieurs secteurs fonctionnels<sup>7</sup>. Le secteur de Jons est situé entre le confluent de l'Ain et la division du Rhône en deux bras artificiel. Il est hérité d'un style à méandres sur laquelle se superpose une bande de tressage<sup>8</sup>. Une carte élémentaire du secteur est dessinée :



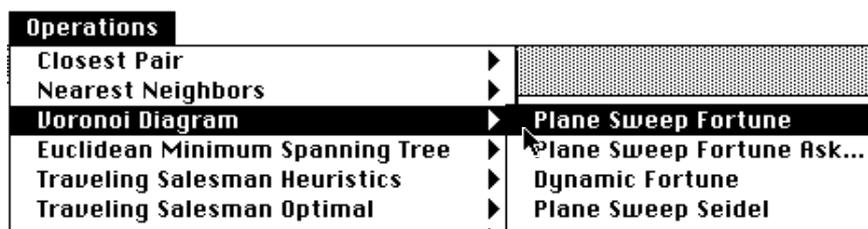


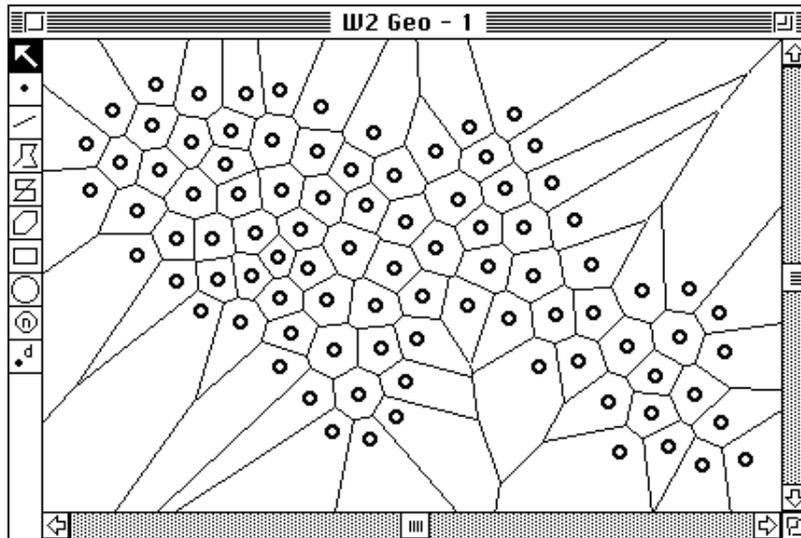
91 points sont régulièrement disposés sur la zone qui recouvre 2200 ha. Un relevé du milieu est exécuté et plusieurs relevés ornithologiques (Indice Ponctuel d'Abondance de Blondel<sup>9</sup>) sont enregistrés. L'illustration porte sur un ensemble de 64 espèces et une série de relevés. La note d'abondance d'une espèce dans un relevé est le nombre d'oiseaux contactés limité à 10 pour prévenir l'effet des troupes monospécifiques.

La distance entre deux points est de l'ordre de 500 m. Leur disposition régulière, compte tenu des contraintes de déplacement conduit à diviser l'espace étudié en 91 surfaces élémentaires. La figure du plan d'échantillonnage est importée dans le logiciel XYZ :



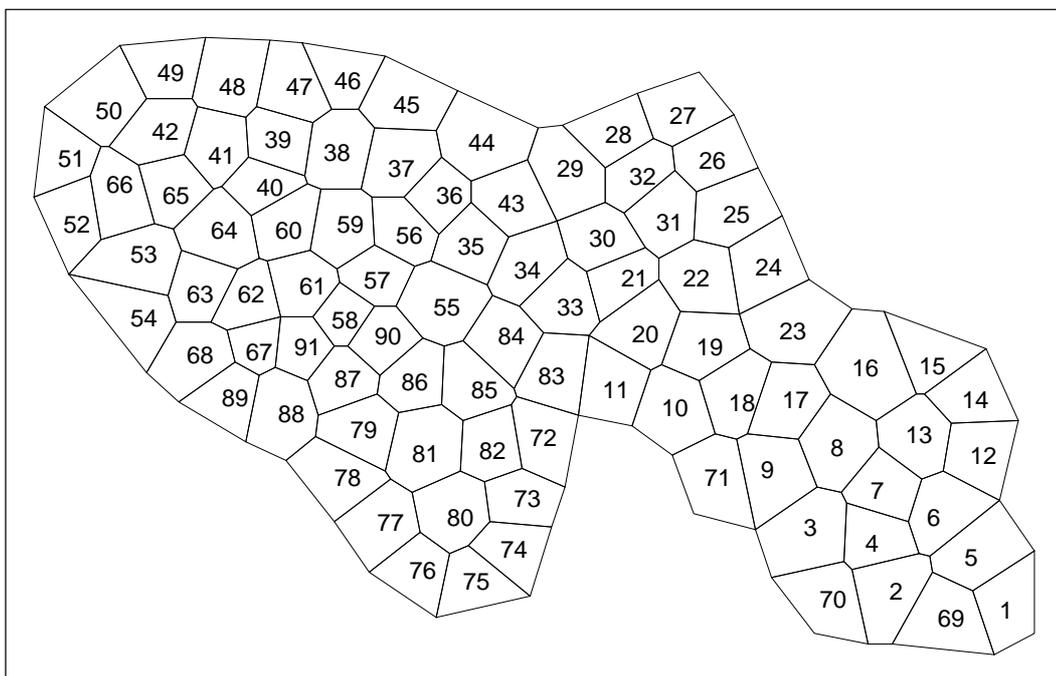
Le programme XYZ, d'une qualité extraordinaire, est en diffusion scientifique libre.





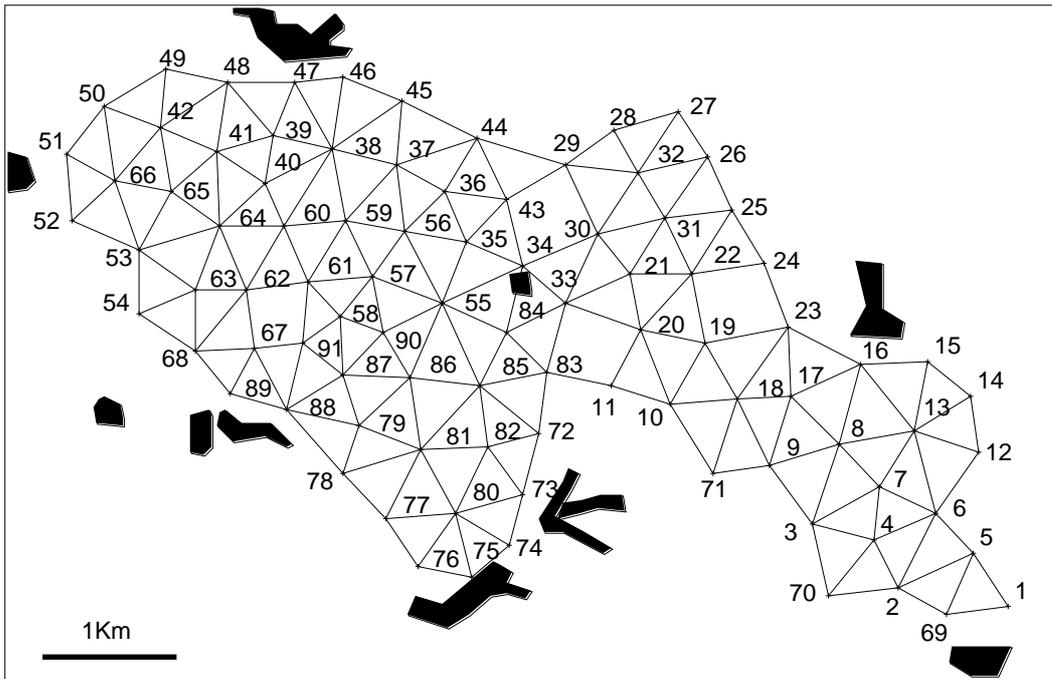
On trace le pavage de Voronoï<sup>10</sup> qui est exporté au format PICT dans Claris™ Draw où un contour est ajouté manuellement. Le résultat est exporté en PICT dans Carto-2D et digitalisé. Le résultat est exporté en TEXT dans ADE-4 qui définit le graphe de voisinage associé dit Graphe de Delaunay.

### 3.1 — Pavage de Voronoï

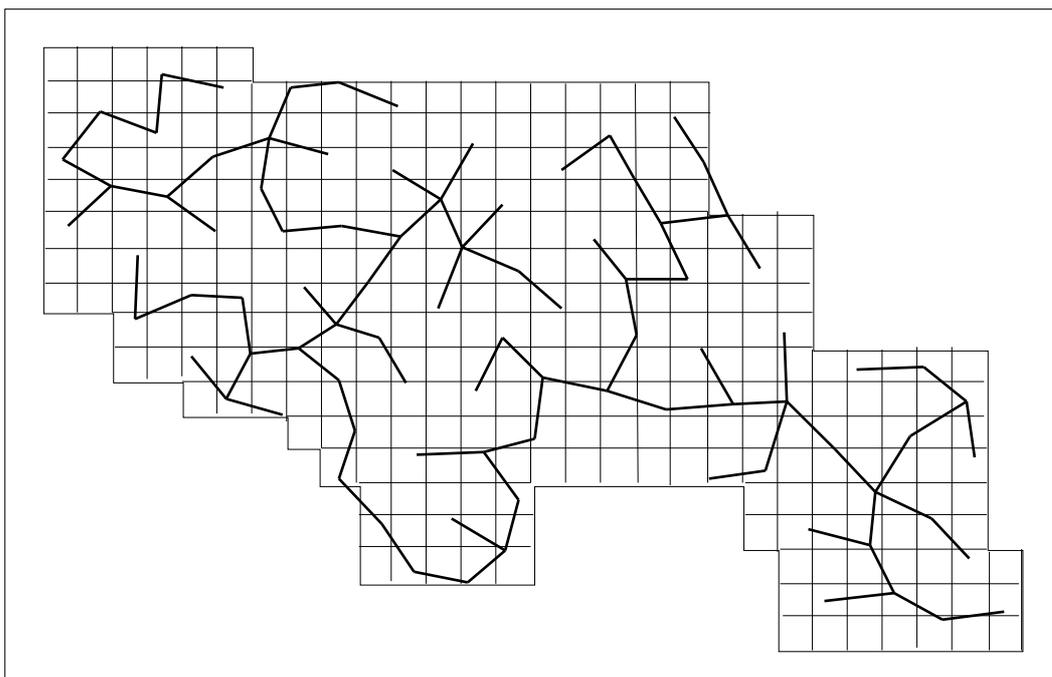


*Pavage de Voronoï*

Le module Digit permet enfin de préparer le grille de tracé des courbes de niveaux. On y a ajouté le graphe de l'arbre de longueur minimale<sup>11</sup>. L'espace concret, le plan d'échantillonnage, le pavage de Voronoï, le graphe de Delaunay, l'arbre de longueur minimal (module Distances, option Minimal Spanning Tree), le quadrillage et son contour sont ainsi manipulés à la même échelle et les représentations sont superposables.



*Graphe de Delaunay*



*Grille de régression locale et arbre de longueur minimale*

Les analyses qui suivent utilisent le graphe de Delaunay et la pondération associée. La figure 7 et le tableau 1 représente les données brutes.

Tableau 1 : Données traitées (60 espèces — 91 relevés) - pour 0, A pour 10.

-----1-----1-----1-----146-----1-----

---3-----111-----2-----1-----1-----4-----2-1-----12-1-

2-----7-----6-2-----5-----12-----1-----112-1

-----1-----4-1-----21-1-----1-1-----111-----1-21-11-1-----1

---1-----1-----1-1-----1-----1-----1-----1-----1-----

-----1-1-1-----11-1-----

-----21-----1-----11-----1-----1-----

---11-----1112--221-2-122-221--121-----11-----1-----1-1-----1-1-----1-----

-----1-----2-----2-----

---1-21-1-----111121--111-3112-1--24-----1-----7-----2-----1--

-----111-----2-----11-----1-----1-----462-----4-26-1

-----2211-2-5-----1-----

4A9-----1-----1-----1-----A4-1A-A---AAA-1-----1-1--2-4

11-----4-----1-----1-----1-----1-----1-----1-----1--1--1

--1-----15-----1--1-----2-1-----21--1--1-----

---1-----2-----1-1-1-114-3--1-111-----111-11-----1-111-1-----

-----111--1-----1-1-21232222-1111-11-----1--1-----11--11111--111-1--1

---1-----1-----1-1-----862--A--21-----A1-----1-----

-----1-----1-----2--1-----1-----

-1-11-121-----1--212-112-----1-11--11-----2--1-----1-----1--31-1-2-31--11-1

-1-----1-----1-----11-----

-----36214-11-22-3341-2-1-42-----1211--1--1-11-----1--1-----

--4-----58-----1--2--8-----13--51-1-----3-----332-----

-----1-121-112-----2-111-----11-----12-----1

211-2--41-----2-1--1-----13-----112--1-122--11241-1--1111131-4

111--12-----11--1-1-12-----1--1-----2112-123-----123-1-112-2--221--1131-3143111-

-11-3322123--2212-312-212314133--211322-121--1-12423112-1-----221211-1322111-1-12212-1111

-----2--2-----1-1-----1-----

-----1-----2--1-----

-2224223432--131141222322513221--2325132-12--2344446652-142-234-132132143524-32131442332432

-----1-----1-----12-----1-----1--1--1--122--2--2-----12--112--1-2--

-----1-11-----

-----122311-1-----2-----

-----1-2411--3--21-2-----1--1--3-----14-----1-----

-----111-----

-----1--11--1--1121--11--111--111-1--3111--1-----1-2--1-----1-----

-----21--1-----1-----1-----1-----2-1-----

---21--1-----11-----1-----1

22213436264-2442233534423222224--4424454441213223335443-532-232-44424354355422432135343-242

-432213334--31-211213313223221--2334532431--22-2812642-21--241-24223143-353322143143331133

---11-11111--1-----11131-----1-121-11-----1--1-----11--1-21-----11--3333--1-

-12-11-2-----2-1--1111-----22--1-----7--1-----52-4--6-112--6-3-111--

-1-----11-----1--1-----1--1-----6--11-----123-----

-12-111112--11-22-2-12221114-2--11112--124--2-113-1323-----241--2211-1-2--23632--1-111-

1611223-21--32-22-112232213121--223121123--2632223331-3-1-2-1--32542148235136242214131141

---1-11-----1--1-1--11--1-----211-----1-----1-----211111--1--1--

---1-112-----1-1--21--1-2--21111-----1-1--211-1--12--2-1--111-11231--11--1--

---1--1--1--1141-21--1-1-11-3--11-111-----1-111-----1--1-1-----1-----

---11213-----1-----112-11-----3-12-----1-----12--1-1--1--12-411--422--

--1-----11--1-----2-----2-114--11--2-----1--132112--12-1-----1

---A1-----3-1-----

12565321511387787AA2A62321712432-3232416213312214541-22142A56-15-1--2-42323253-3124-2--21-2

-A-----23A--12A--AA--4-1-----1-----A-----4--4-----314--1--1-2-2--

-----A-A-----1-----23-A-----1-----

1-----2-1-1--1--1221132131-2-1--112-1--1--23-2--112--121-12--23231222212--4-133--1-2-1

-----3122-----1-----111-----

-----2-----1-1-----1--1-1--1-----1-----1-----111-----

-----1-----1-----2--352--1--11-----

-----1-A-----A1-----1-----4--1-----1

-----1--2--1-----1-1-----13-----

-----2-----1-----11-----1-----

-----1211-1-----11-1-1-----1-----

-----11-----3-1--1-----11-----

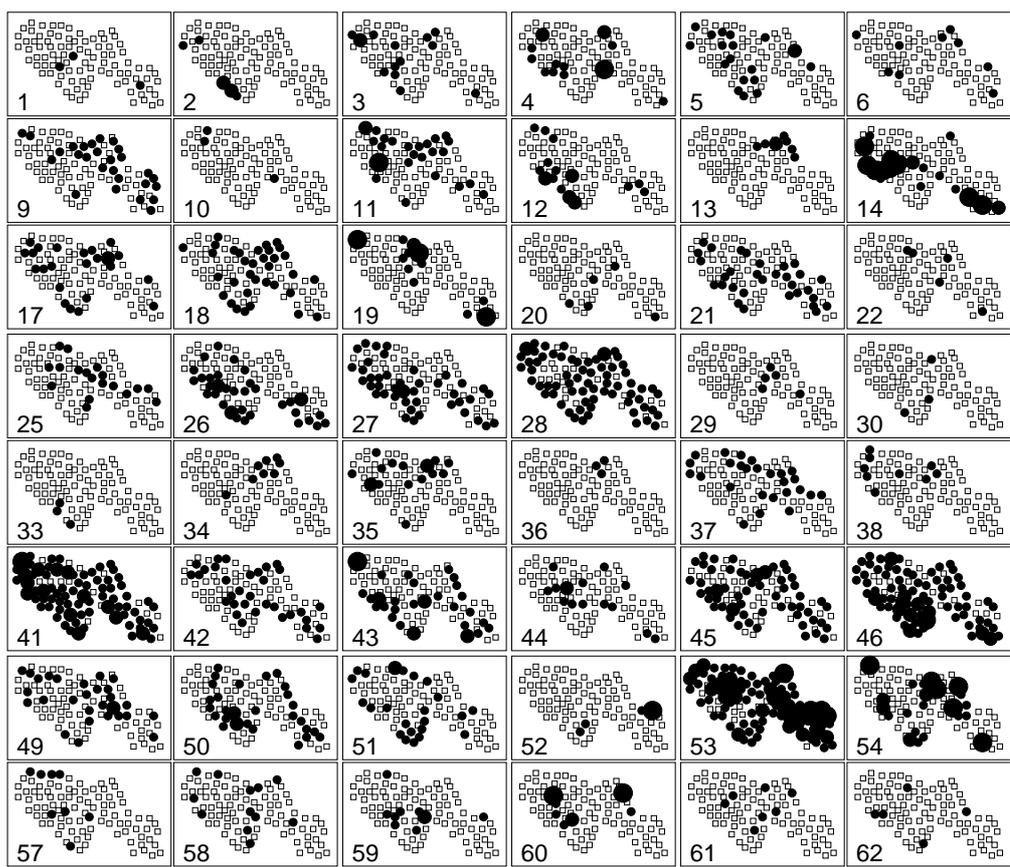


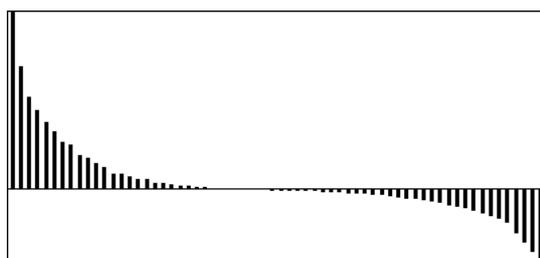
Figure 7 : Cartographie partielle des données (absence : petit carré, abondance entre 1 et 10 : cercle noir).

### 3.2 — Analyses simultanées

L'analyse des correspondances simples utilise la pondération de voisinage, puis définit les analyses globale et locale :

Row weighted COA			
Data file	<input type="button" value="..."/>	Rvi_Jons	91 64
Row weight (default 1/n)	<input type="button" value="..."/>	R_G.gpl	91 1
Generic output file	<input type="button" value="..."/>	RVIw	

Moran Index Analysis			
Input file	<input type="button" value="..."/>	RVIw.fcta	91 64
Graph input file	<input type="button" value="..."/>	R_G.gpl	91 1



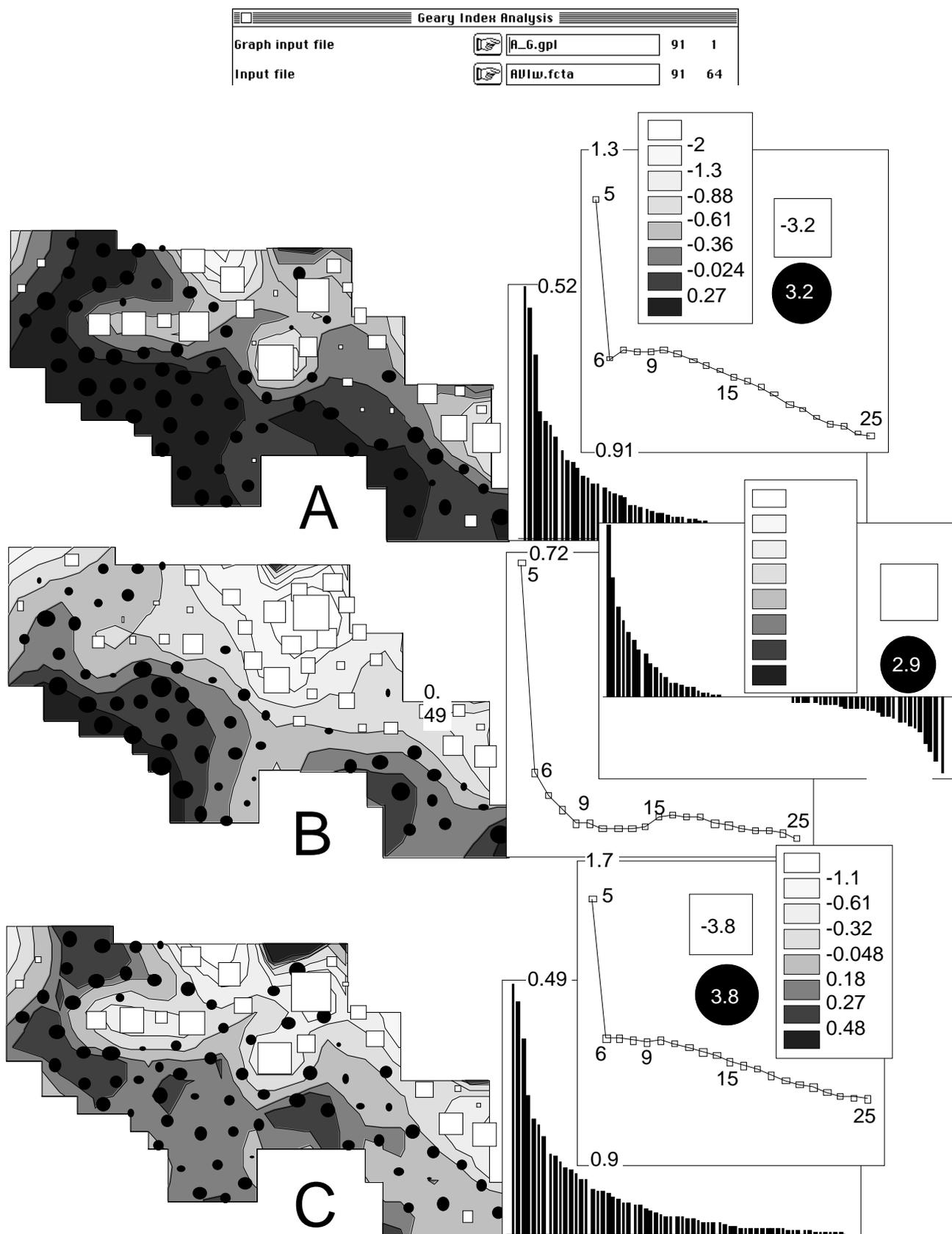


Figure 8 : A—Analyse des correspondances. B—Analyse des correspondances globale (Moran). C—Analyse des correspondances locale (Geary). Dans chaque cas, graphe des valeurs propres, cartographie du premier facteur (courbes de niveaux à 9 voisins) et graphe de l'erreur de lissage en fonction du nombre de voisins. L'analyse simple réunit les composantes globales et locales.

Dans le module Distances on peut en profiter pour diagonaliser l'opérateur de Moran, ceci par curiosité :



La figure 9 donne les cartes des 9 premiers vecteurs propres.

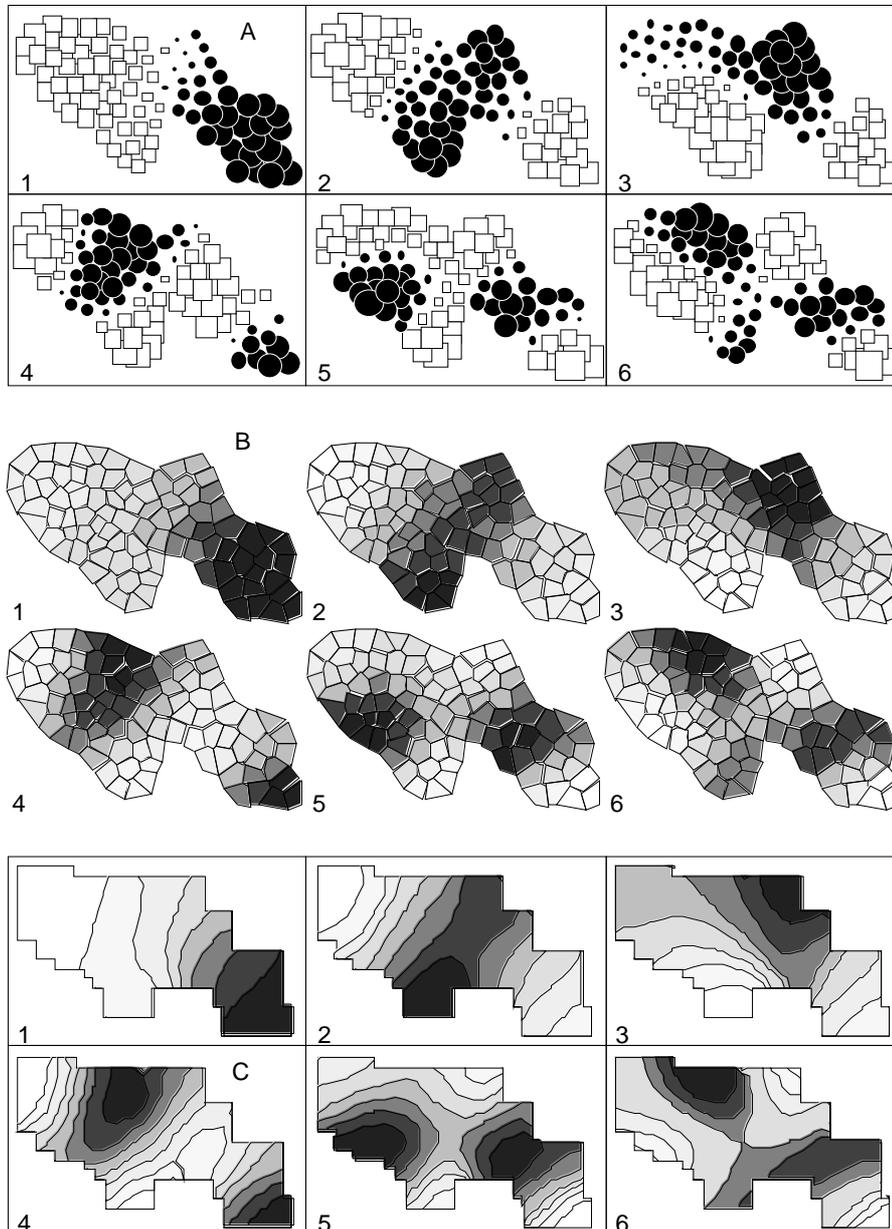


Figure 9 : Cartographie des 6 premiers vecteurs propres de l'opérateur de Moran. A—par cercles et carrés, B— par surfaces, C— par courbes de niveaux. D— Erreur de lissage par régression locale. Remarquer la croissance de l'erreur avec le rang du vecteur et avec le nombre de voisins dès que ce paramètre dépasse 5.

La figure 8 compare les trois analyses, principalement leur facteur 1. Il y a une valeur propre marquée dans l'analyse globale et c'est pourquoi la comparaison n'est pertinente qu'à ce niveau. Il y a identité de la répartition de l'inertie ordinaire et de

l'inertie de l'analyse locale. Ici c'est l'analyse locale qui est identique à l'analyse ordinaire, dans le cas précédent c'était l'analyse globale.

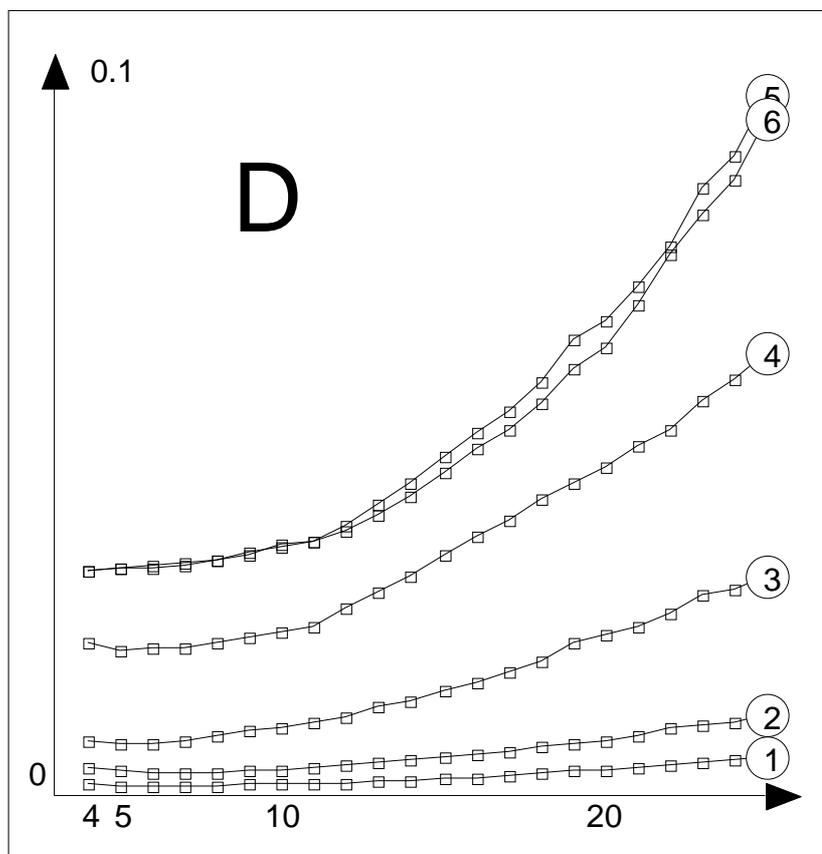


Figure 9 - partie D.

Quelle que soit l'origine de la structure, elle apparaît dans l'analyse générale. C'est pourquoi (Cf. Chapitre 1 de la fiche autocorrélation spatiale multidimensionnelle), les représentations spatio-temporelles des coordonnées factorielles sont des méthodes efficaces sans être optimisées de ce point de vue.

Mais ici, on identifie l'essentiel de l'origine de cette structure quant à l'échelle en cause. La carte A de la figure 8 est un compromis de la carte B (séparation de la plaine alluviale en deux zones, en gros, ouverte et fermée) et de la carte C (variation forte entre un point sans végétation haute et les lisières des formations boisées qui l'entoure).

Exprimée à deux échelles (entre deux stations voisines d'une part, à l'échelle de la zone étudiée d'autre part) la relation avifaune végétation classique s'exprime deux fois.

La première fait l'essentiel de l'analyse locale, la seconde, en l'occurrence beaucoup moins importante, n'apparaît que dans l'analyse globale. L'analyse globale est cachée dans l'analyse ordinaire. Noter combien l'erreur de lissage par régression locale est similaire dans les cadres A et C, tout comme le graphe des valeurs propres, et combien elle est originale dans le cas B où elle est minimum pour 9 voisins. Les 3 cartes ont été tracées avec 9 voisins pour comparaison.

Les trois analyses définissent des codes numériques des espèces de variance unité pour la pondération marginale qui maximise respectivement la variance des moyennes par relevés (averaging : analyse classique), la covariance locale (donc simultanément la variance et la corrélation locale) et l'autocovariance (donc simultanément

l'autocorrélation et la variance). En ce sens, les trois analyses poursuivent un même objectif en mettant ou non une priorité d'échelle. La comparaison de ces codes pour les analyses globales et locales est singulièrement instructive (figure 10). On pouvait s'attendre à une certaine identité vue l'importance de la relation avifaune-végétation.

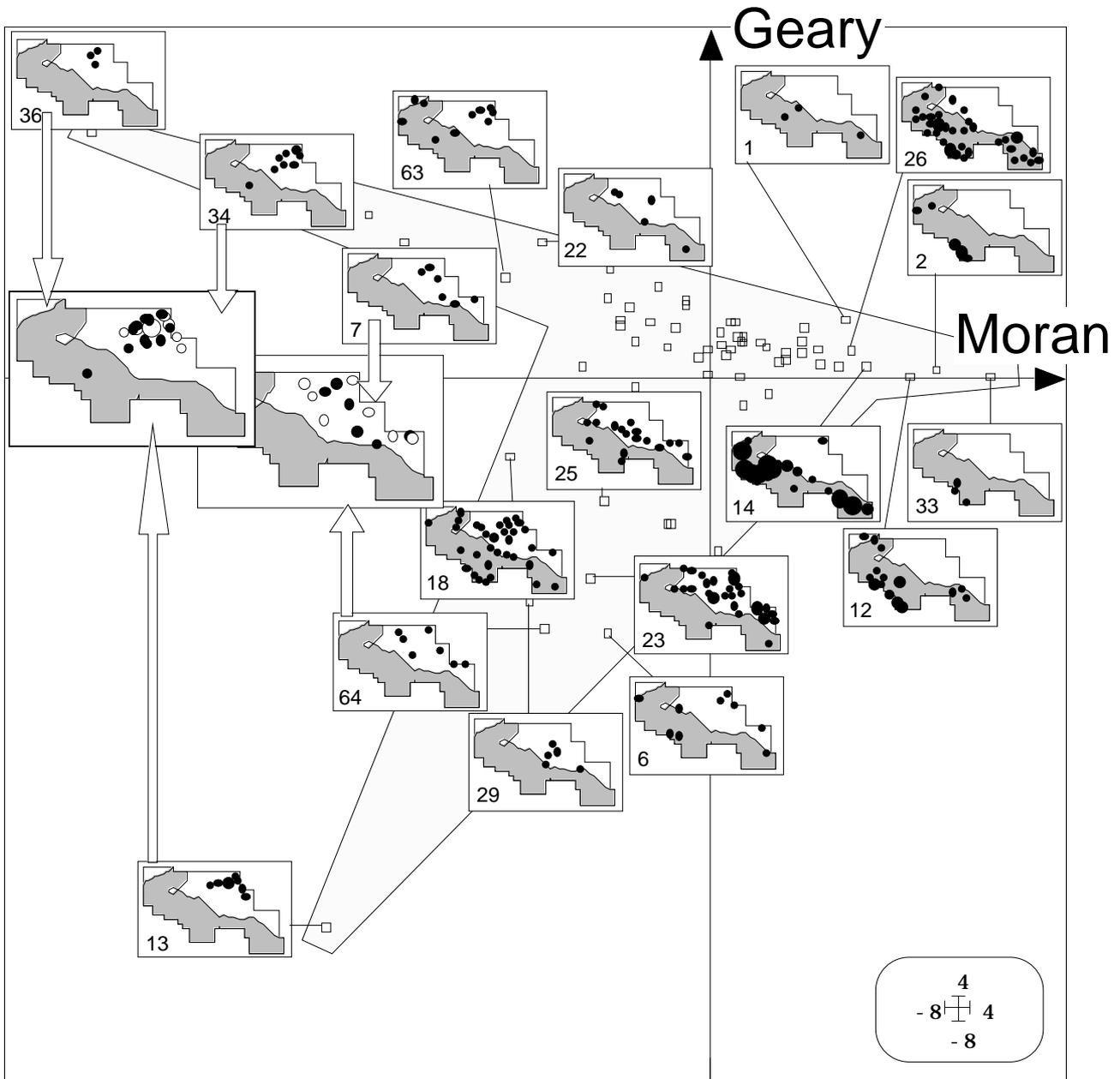


Figure 10 : Codage des espèces. En abscisse score du facteur 1 de l'analyse globale. En ordonnée, score du facteur 1 de l'analyse locale. La structure locale est essentiellement décrite dans les milieux ouverts.

L'analyse globale ordonne les espèces en fonction de leur mode de répartition au niveau de la zone d'étude en séparant les espèces de la partie sud (fleuve, végétation riveraine, côtière boisée) et celle de la partie nord (plaine de l'Ain ouverte). La totalité des indicatrices de la partie boisée ont des scores voisins de 0 pour l'analyse locale, laquelle sépare des espèces dont la répartition est entièrement nord. La structure locale principale est donc une structure de milieux ouverts. On peut ainsi caractériser des relations interspécifiques qui s'expriment dans une partie de la zone d'étude. Les associations négatives d'espèces ainsi mises en évidence ont été illustrées par superposition des cartes de répartition d'espèces ordonnées aux deux extrémités de

l'ordination locale. On peut penser soit à des différences de biotope de nidification (ecological pattern), soit à des relations comportementales (sociological pattern). Par exemple, la carte de l'espèce 7 (Faucon crécerelle, *Falco tinnunculus* L.) est imbriquée avec celle de l'espèce 64 (Bruant proyer, *Emberiza calendra* L.) espèce commune des paysages découverts dont on peut penser que la présence est plus discrète en présence d'un prédateur. Par contre deux espèces inféodées aux roselières en bordure de zones humides (36, Rousserolle turdoïde, *Acrocephalus arundinaceus* L. et 34, Locustelle luscinoïde, *Locustella luscinoïdes* L.) sont associées et opposées à 13 (Courlis cendré, *Numenius arquata* L.) qui partage avec les premières la même zone d'implantation mais rarement les mêmes stations.

Dans une analyse locale, deux espèces ne peuvent être prises en compte en opposition que si elles sont présentes dans des stations voisines (donc géographiquement proches) en s'excluant mutuellement. Elles sont prises en considération avec le même signe si elles apparaissent dans les mêmes stations en étant absentes dans les stations voisines. Ainsi, l'analyse locale supporte ici l'hypothèse d'un partage horizontal de l'espace en milieu ouvert qui serait le pendant de la séparation des niches entre strates de la végétation en milieu fermé. L'identité des résultats entre analyse locale et globale, qui conduit systématiquement à l'identité locale = globale = classique, n'est donc pas un artifice mathématique mais la preuve que la structure est d'ordre écologique (environmental pattern). Les outils graphiques jouent, dans tous les cas, un rôle important qui doit faire l'objet d'un contrôle stricte.

## Références

- <sup>1</sup> Pialot, D. (1985) Analyse des données de milieu en hydrobiologie. Apport des techniques d'analyses multivariées. Thèse de 3<sup>o</sup> cycle, Université Lyon 1. 1-233.
- <sup>2</sup> Lebart, L. (1969) Analyse statistique de la contiguïté. *Publication de l'Institut de Statistiques de l'Université de Paris* : 28, 81-112.
- <sup>3</sup> Le Foll, Y. (1982) Pondération des distances en analyse factorielle. *Statistique et Analyse des données* : 7, 13-31.
- <sup>4</sup> Carlier, A. . (1985) Applications de l'analyse factorielle des évolutions et de l'analyse intra-périodes. *Statistique et Analyse des Données* : 10, 27-53.
- <sup>5</sup> Lebreton, Ph. . (1977) *Les oiseaux nicheurs rhônalpins. Atlas ornithologique Rhône-Alpes*. Centre Ornithologique Rhône-Alpes, Université Lyon 1, 69621 Villeurbanne. Direction de la Protection de la Nature, Ministère de la Qualité de la Vie. 1-354.
- <sup>6</sup> Bournaud M. (1990) *Peuplements d'oiseaux et propriétés des écosystèmes de la plaine du Rhône : descripteurs de fonctionnement global et gestion des berges*. Rapport sur Programme SRETIE, Ministère de l'environnement, 1-135.
- <sup>7</sup> Amoros C., Roux A.L., Reygrobellet J.L., Bravard J.P. & Pautou G. (1987) Methodological research applied to the ecological management of fluvial hydrosystems. *Regulated Rivers*, 1, 17-36.
- <sup>8</sup> Bravard J.P., Amoros C. & Pautou G. (1986) Impact of civil engineering works on the succession of communities in a fluvial system. A methodological approach and predictive approach to a section of the Upper Rhône River. *Oikos*, 47, 92-111.
- <sup>9</sup> Blondel J., Ferry C. & Frochot B. (1970) La méthode des Indices Ponctuels d'Abondance (I.P.A.) ou des relevés d'avifaune par "stations d'écoute". *Alauda*, 38. 89-110.

<sup>10</sup> Chassery, J.M. & Montanvert, A. (1991) Géométrie discrète en analyse d'images. Hermès, Paris. 1-358.

<sup>11</sup> Kevin, V. & Whitney, M. (1972) Algorithm 422. Minimal Spanning Tree [H]. *Communications of the Association for Computing Machinery* : 15, 4, 273-274.