

# Profils écologiques et analyse de co-inertie

## Résumé

La fiche illustre les calculs de profils écologiques et l'exacte similitude entre AFC des profils écologiques (Romane, 1972, Un exemple d'utilisation de l'analyse factorielle des correspondances en écologie végétale. In : Grundfagen und Methoden in der Pflanzensoziologie. Maarel, E. van der & Tüxen, R. (Eds.) Dr. W. Junk b.v., The Hague. pp. 155-162.) et analyse de co-inertie d'un tableau de variables environnementales qualitatives et d'un tableau de végétation en présence-absence. Cette notion est étendue à des données de végétation quantitatives. La co-structure entre une AFC et une ACM est alors la version complète de l'AFC des tableaux de profils écologiques. L'illustration proposée est celle de Mercier et al. (1992, Complete correspondence analysis of an ecological profile data table: a central ordination method. *Acta Oecologica* : 13, 25-44).

## Plan

1 - Introduction .....	2
2 - Les données traitées.....	2
3 - Analyses séparées.....	4
4 - Profils écologiques.....	8
5 - Analyse du couple milieu-végétation.....	16
6 - Conclusions.....	27
Références .....	28

S. Dolédec, D. Chessel & P. Mercier

# 1 - Introduction

L'analyse des relations espèces-milieu est un problème central de l'écologie animale ou végétale qui s'intéresse aux liens existant entre les deux compartiments fondamentaux de l'écosystème, le biotope et la biocénose. La non mobilité des individus fait du matériel végétal un support privilégié pour une approche exploratoire d'une telle problématique. Nous disposons alors de deux tableaux, un tableaux descriptif de la végétation (données quantitatives ou en présence-absence) et un tableau descriptif du milieu (données quantitatives et/ou qualitatives). Le point commun entre ces deux jeux de données est représenté par les stations où sont récoltées les variables abiotiques et biotiques. Le problème peut alors être posé en terme d'analyse de la co-structure entre ces deux tableaux ou encore de la covariation entre les structures associées à ces deux tableaux. Quand le nombre des relevé est réduit on peut utiliser l'analyse de co-inertie.

## 2 - Les données traitées

La méthode est appliquée à un jeu de données acquises par Batterink & Wijffels (rapport non publié) sur des prairies dunaires de l'île Terschelling (Pays-Bas). Ce couple de tableaux est proposé par Ter Braak (1987)<sup>1</sup> dans le cadre de l'utilisation de l'Analyse Canonique des Correspondances.

### 2.1 - Le tableau de milieu

Les variables de milieu sont considérées par Ter Braak comme quantitatives, la cinquième étant décomposée en quatre indicatrices. Nous les considérons ici comme qualitatives (Tableau 1).

Tableau 1 Définition des variables

Article de Ter Braak (1987)	Données traitées dans cette fiche
1 - Epaisseur de l'horizon A1: variable quantitative (valeurs en cm)	1 - Epaisseur de l'horizon A1: variable qualitative à 4 modalités (mise en classes selon le code 1: 0 cm / 2: 1 à 2 cm / 3: 3 à 4 cm / 4: 5 à 9 cm)
2 - Humidité du sol : variable quantitative (5 classes à valeur ordinale)	2 - Humidité du sol : variable qualitative à 4 modalités (la classe 4 n'étant pas représentée)
3 - Quantité d'engrais apportée : variable quantitative (5 classes à valeur ordinale)	3 - Quantité d'engrais apportée : variable qualitative à 5 modalités
4 - Utilisation agricole : variable quantitative (3 classes à valeur ordinale. 1: pré de fauche / 2: pâture de fauche / 3: pâture)	4 - Utilisation agricole : variable qualitative à 3 modalités
5 à 8 - Type de gestion : 4 variables quantitatives prises en compte de façon indépendante (0 ou 1)	5 - Type de gestion : variable qualitative à 4 modalités 1: exploitation agricole standard 2: exploitation agricole biodynamique 3: exploitation agricole "hobby farming" 4: gestion naturelle

Aller à la carte «DuneCJT87» de la pile **ADE.Data** (Fig. 1) pour obtenir les données du Tableau 2.

Par défaut, le fichier contenant les données de milieu sous forme de chaîne de caractères est intitulé **Mi1.Car** et le fichier des étiquettes des variables est appelé **Cod\_Var**.

Tableau 2 Variables de milieu.

Environmental variables	
V1 Thickness A1	31211112222233224431
V2 Soil moisture	11112112122334444444
V3 Quantity of manuring	35342321155234422242
V4 Agricultural use	12231231122122313231
V5 Type of management	31332224411311344414

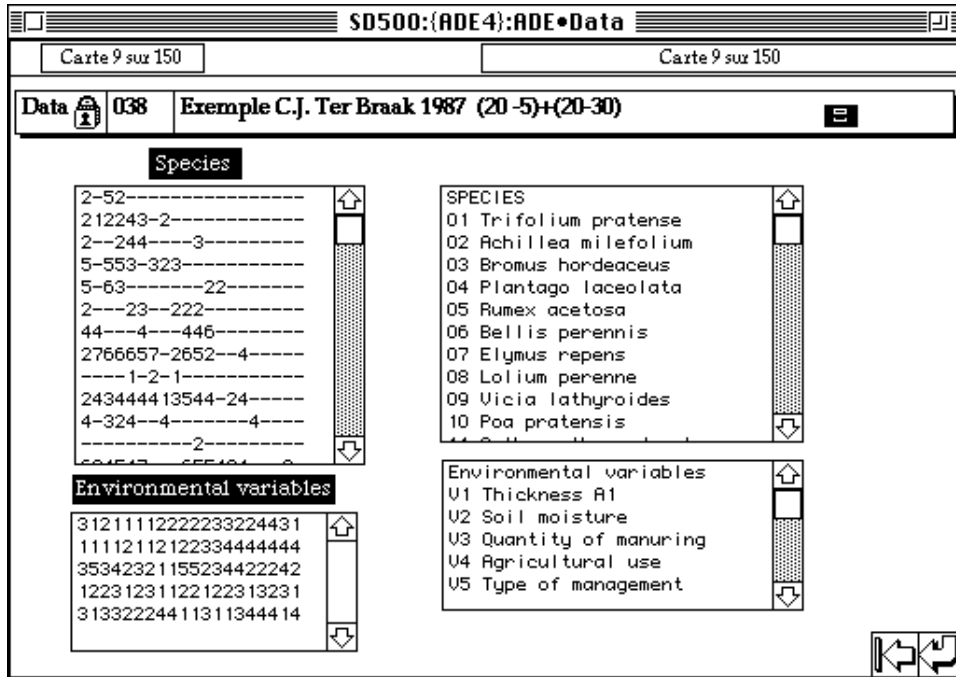


Figure 1 Les données traitées contenues sont contenues dans la pile ADE•Data.

	1	2	3	4	5
1	3	1	3	1	3
2	1	1	5	2	1
3	2	1	3	2	3
4	1	1	4	3	3
5	1	2	2	1	2
6	1	1	3	2	2
7	1	1	2	3	2
8	2	2	1	1	4
9	2	1	1	1	4
10	2	2	5	2	1
11	2	2	5	2	1
12	2	3	2	1	3
13	3	3	3	2	1
14	3	4	4	2	1
15	2	4	4	3	3
16	2	4	2	1	4
17	4	4	2	3	4
18	4	4	2	2	4
19	3	4	4	3	1
20	1	4	2	1	4

Transformer **Mi1.Car** en un fichier binaire **Mi1Tr** (5-20) à l'aide de l'option **Char->Binary** du module **TextToBin**. Puis transposer celui-ci en un fichier **Mi1** (20-5) à l'aide de l'option **Transpose** du module **FilesUtil**. Avec la commande **Edit with** et Excel™ comme logiciel d'accès on obtient le tableau ci-dessus.

## 2.2 - Le tableau floristique

Le tableau utilisé est celui de Ter Braak avec comme seule modification la numérotation des relevés qui correspond ici au numéro d'ordre dans le tableau proposé par l'auteur. Utiliser la carte «DuneCJT87» de la pile **ADE.Data** (Fig. 1) pour obtenir les données du Tableau 3. Par défaut, le fichier contenant les données de milieu sous forme de chaîne de caractères est intitulé Veg.Car et le fichier des étiquettes des variables est appelé Cod\_Esp.

Tableau 3 Données floristiques.

SPECIES	
01	<i>Trifolium pratense</i> 2- 52-----
02	<i>Achillea millefolium</i> 212243- 2-----
03	<i>Bromus hordeaceus</i> 2- - 244- - - 3-----
04	<i>Plantago lanceolata</i> 5- 553- 323-----
05	<i>Rumex acetosa</i> 5- 63----- 22-----
06	<i>Bellis perennis</i> 2- - - 23- - 222-----
07	<i>Elymus repens</i> 44- - - 4- - 446-----
08	<i>Lolium perenne</i> 2766657- 2652- - 4-----
09	<i>Vicia lathyroides</i> - - - - 1- 2- 1-----
10	<i>Poa pratensis</i> 243444413544- 24-----
11	<i>Anthoxanthum odoratum</i> 4- 324- - 4----- 4-----
12	<i>Cirsium arvense</i> ----- 2-----
13	<i>Poa trivialis</i> 624547- - - 655494- - - 2-
14	<i>Trifolium repens</i> 2- 52653- 2213322261- -
15	<i>Leontodon autumnalis</i> 3- 3335525222223622- 2
16	<i>Brachythecium rutabulum</i> 2- 622- 4- 62224- 23- 444
17	<i>Juncus bofonius</i> - - - 2----- 443-----
18	<i>Sagina procumbens</i> ----- 2- - - 524223- - -
19	<i>Alopecurus geniculatus</i> ----- 2- - - 723855- - - 4-
20	<i>Hypochaeris radicata</i> ----- 22----- 5-----
21	<i>Aira praecox</i> ----- 2----- 3-----
22	<i>Salix repens</i> ----- 2----- 3- - 5
23	<i>Agrostis stolonifera</i> ----- 483454- 4475
24	<i>Juncus articulatus</i> ----- 4- 4- - 334
25	<i>Chenopodium album</i> ----- 1-----
26	<i>Empetrum nigrum</i> ----- 2-----
27	<i>Ranunculus flammula</i> ----- 22- 2224
28	<i>Eleocharis palustris</i> ----- 4- 4584
29	<i>Calliargonella cuspidata</i> ----- 4- 33
30	<i>Potentilla palustris</i> ----- 22--

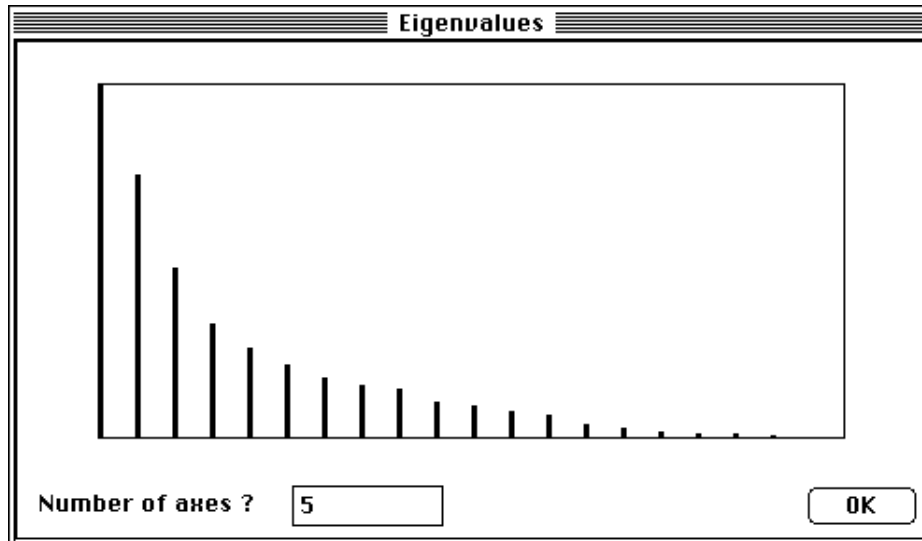
Transformer Veg.Car en un fichier binaire VegTr (30-20) à l'aide de l'option **Char->Binary** du module **TextToBin** après avoir pris soin de remplacer les tirets par des 0. Puis transposer celui-ci en un fichier Veg (20-30) à l'aide de l'option **Transpose** du module **FilesUtil**.

### 3 - Analyses séparées

Ces analyses préalables sont indispensables à la mise en oeuvre pratique du couplage et s'intègrent dans la procédure générale. Chaque tableau est préalablement traité selon une AFC pour le tableau floristique et une ACM pour le tableau de milieu. Les résultats de ces analyses seront de plus utilement comparés à ceux issus du couplage.

#### 3.1 - Analyse factorielle des correspondances du tableau descriptif de la végétation

Le tableau floristique est typiquement un tableau d'analyse factorielle des correspondances. Exécuter l'option **C**orrespondence Analysis du module COA.



Conserver 5 axes. On obtient le listing suivant.

```
fc/COA: Correspondance analysis
Input file: Veg
Number of rows: 20, columns: 30
File Veg.fcpl contains the edge distribution of rows
It has 20 rows and 1 column
File Veg.fcpc contains the edge distribution of columns
It has 30 rows and 1 column
File Veg.fcta contains the doubly centred table DI-1*P*DJ-1 -II*1J'
It has 20 rows and 30 columns
File Veg.fcma contains:
    the number of rows: 20
    the number of columns: 30
    the total number: 684
```

```
-----
DiagoRC: General program for two diagonal inner product analysis
Input file: Veg.fcta
--- Number of rows: 20, columns: 30
-----
```

Total inertia: 2.1138

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+5.3739E-01	+0.2542	+0.2542	02	+4.0136E-01	+0.1899	+0.4441
03	+2.6014E-01	+0.1231	+0.5672	04	+1.7509E-01	+0.0828	+0.6500
05	+1.3856E-01	+0.0655	+0.7156	06	+1.1081E-01	+0.0524	+0.7680
07	+9.1126E-02	+0.0431	+0.8111	08	+8.0880E-02	+0.0383	+0.8494
09	+7.3777E-02	+0.0349	+0.8843	10	+5.6470E-02	+0.0267	+0.9110
11	+4.8512E-02	+0.0229	+0.9339	12	+4.1345E-02	+0.0196	+0.9535
13	+3.4951E-02	+0.0165	+0.9700	14	+2.0678E-02	+0.0098	+0.9798
15	+1.4814E-02	+0.0070	+0.9868	16	+9.2327E-03	+0.0044	+0.9912
17	+8.2409E-03	+0.0039	+0.9951	18	+6.9510E-03	+0.0033	+0.9984
19	+3.4654E-03	+0.0016	+1.0000	20	+0.0000E+00	+0.0000	+1.0000

```
File Veg.fcvp contains the eigenvalues and relative inertia for each
axis
--- It has 30 rows and 2 columns
```

```
File Veg.fcco contains the column scores
--- It has 30 rows and 5 columns
```

.....  
File Veg.fcpl contains the row scores  
--- It has 20 rows and 5 columns  
.....

### 3.2 - Analyse des correspondances multiples du tableau descriptif du milieu

Le tableau de milieu contenant des variables qualitative relève typiquement d'une analyse des correspondances multiples. Exécuter l'option **Read Categ File** du module **CategVar** sur le fichier **Mil**. On obtient le listing suivant :

Categorical variables: file Mil  
Rows: 20, Variables: 5, Categories: 20, Missing data: 0

Description of categories:

-----  
Variable number 1 has 4 categories  
-----

[ 1]Category:	1	Num:	6	Freq.:	0.3
[ 2]Category:	2	Num:	8	Freq.:	0.4
[ 3]Category:	3	Num:	4	Freq.:	0.2
[ 4]Category:	4	Num:	2	Freq.:	0.1

-----  
Variable number 5 has 4 categories  
-----

[ 17]Category:	1	Num:	6	Freq.:	0.3
[ 18]Category:	2	Num:	3	Freq.:	0.15
[ 19]Category:	3	Num:	5	Freq.:	0.25
[ 20]Category:	4	Num:	6	Freq.:	0.3

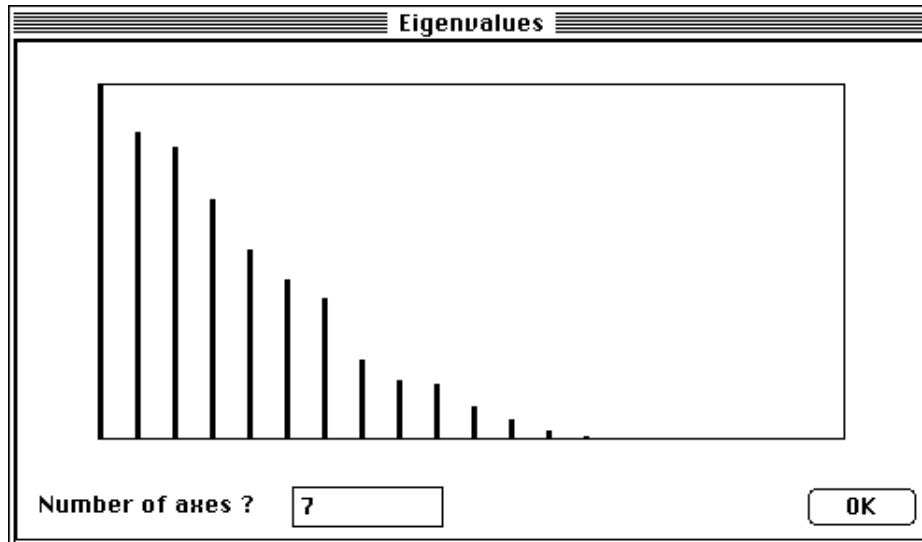
-----  
Auxiliary binary output file MilModa: Indicator vector of modalities  
It contains variable number for each modality  
It has 20 rows (modalities) and one column

Auxiliary ASCII output file Mil.123: labels (two characters) for 20 modalities  
It contains one label for each modality  
It has 20 rows (modalities) and labels 1a, 1b, ..., 2a, 2b, ...  
Variable number 1, 2, ..., A, ..., Z, +, Modality number a, b, ..., z, +

Enchaîner avec l'option **Multiple Correspondence Analysis** du module **MCA** en utilisant le poids de lignes de l'AFC du tableau végétation (cette option est utilisée pour la suite des analyses) :



Conserver les 7 premiers axes (2ème rupture de pente sur le graphe des valeurs propres) :



On obtient le listing suivant :

cm/MCA: Multiple correspondence analysis

Row weights from file Veg.fcpl

File Mil.cmpl contains the row weights  
It has 20 rows and 1 column

File Mil.cmpc contains the column weights (1/V)\*DM  
It has 20 rows and 1 column

Marginal distributions by variable:

-----  
Variable number 1 has 4 categories

-----  
[1] Category: 1 Weight: 0.301  
[2] Category: 2 Weight: 0.42  
.....  
[19] Category: 3 Weight: 0.311  
[20] Category: 4 Weight: 0.219  
-----

File Mil.cmta contains the table processed by MCA  
It has 20 rows and 20 columns (categories)

-----  
DiagoRC: General program for two diagonal inner product analysis

Input file: Mil.cmta

--- Number of rows: 20, columns: 20

-----  
Total inertia: 3

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+5.4832E-01	+0.1828	+0.1828	02	+4.7569E-01	+0.1586	+0.3413
03	+4.5038E-01	+0.1501	+0.4915	04	+3.7122E-01	+0.1237	+0.6152
05	+2.9309E-01	+0.0977	+0.7129	06	+2.4586E-01	+0.0820	+0.7949
07	+2.1712E-01	+0.0724	+0.8672	08	+1.2222E-01	+0.0407	+0.9080
.....	.....	.....	.....	.....	.....	.....	.....
19	+0.0000E+00	+0.0000	+1.0000	20	+0.0000E+00	+0.0000	+1.0000

File Mil.cmpv contains the eigenvalues and relative inertia for each axis  
--- It has 20 rows and 2 columns

File Mil.cmco contains the column scores

--- It has 20 rows and 7 columns

File Mil.cml contains the row scores

--- It has 20 rows and 7 columns

-----  
CorRatioMCA: Correlation ratios after a MCA

Title of the analysis: Mil.cm

Number of rows: 20, columns: 5

Variable : 1

> Categ= 1 Weight= 0.301 0.385 0.168 0.875 -0.968 0.023 -0.128 -0.206

> Categ= 4 Weight= 0.069 1.598 -0.605 -1.455 -0.502 -1.536 1.469 1.293  
-----> r= 0.315 0.345 0.441 0.468 0.396 0.161 0.463

Variable : 2

> Categ= 1 Weight= 0.364 -0.007 0.546 0.844 0.019 0.334 0.650 0.106

> Categ= 4 Weight= 0.314 0.658 0.174 -1.121 -0.175 -0.184 0.056 -0.130  
-----> r= 0.278 0.484 0.661 0.312 0.239 0.429 0.023

Variable : 3

> Categ= 1 Weight= 0.060 0.786 -1.478 0.313 1.590 1.858 1.320 -1.673

> Categ= 5 Weight= 0.151 -1.625 -1.233 -0.325 -0.636 0.343 0.025 0.446  
-----> r= 0.834 0.869 0.421 0.359 0.545 0.325 0.236

Variable : 4

> Categ= 1 Weight= 0.338 0.649 -0.472 0.433 0.662 0.060 -0.485 -0.426

> Categ= 3 Weight= 0.247 0.550 0.907 -0.569 -0.717 0.613 -0.148 0.227  
-----> r= 0.521 0.289 0.143 0.280 0.165 0.182 0.092

Variable : 5

> Categ= 1 Weight= 0.298 -1.189 -0.159 -0.706 -0.237 -0.245 -0.093 -0.427

0.765

> Categ= 4 Weight= 0.219 1.264 -0.918 -0.502 0.373 -0.105 0.673 -0.205  
-----> r= 0.795 0.391 0.586 0.437 0.121 0.132 0.272

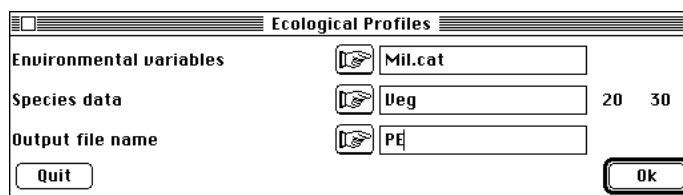
File Mil.cmrc contains the correlation ratios between  
the categorical variables and the factor scores  
It has 5 rows and 7 columns

En cas d'hésitation, on peut voir la différence entre fréquences et poids des variables  
qualitatives dans la fiche ACM.

## 4 - Profils écologiques

### 4.1 - Profils écologiques bruts et tests de signification

Utiliser EcolTools : Ecological Profiles :



New TEXT file PE.pepa contains the parameters:  
----> Environmental variables: Mil [20 samples][5 var]  
----> Species data : Veg [20 samples][30 species]



```

Ecological profiles in file: PE
----> Columns = Species (30)
----> Rows = categories (20)
----> Values = frequency of the species in the samples of the category
Models of ecological profiles in file: PE.pemo
----> Columns = Species (30)
----> Rows = categories (20)
----> Values = expected frequency (independence model)

```

```

Variable: 1, categories: 4
*-----*-----*-----*-----*-----*-----*-----*-----*-----*-----*
|Spec| 1 | 2 | 3 | 4 | Tot| Freq|Entro|MutIn|DF| Khi2 | Proba|DTest|ITest|
*-----*-----*-----*-----*-----*-----*-----*-----*-----*-----*
| 1 | 1 | 1 | 1 | 0 | 3 | 0.150 | 0.610 | 0.035 | 3 | 0.72 | 0.869 | -0.19 | -0.05 |
| 2 | 4 | 2 | 1 | 0 | 7 | 0.350 | 0.934 | 0.172 | 3 | 4.25 | 0.234 | -1.79 | -1.40 |
| 3 | 3 | 1 | 1 | 0 | 5 | 0.250 | 0.811 | 0.132 | 3 | 3.33 | 0.343 | -1.33 | -1.39 |
| 4 | 3 | 3 | 1 | 0 | 7 | 0.350 | 0.934 | 0.090 | 3 | 1.87 | 0.604 | -1.31 | -0.32 |
| 5 | 1 | 2 | 2 | 0 | 5 | 0.250 | 0.811 | 0.092 | 3 | 2.22 | 0.531 | 0.27 | 0.60 |
| 6 | 2 | 3 | 1 | 0 | 6 | 0.300 | 0.881 | 0.062 | 3 | 1.15 | 0.768 | -0.81 | 0.30 |
| 7 | 2 | 3 | 1 | 0 | 6 | 0.300 | 0.881 | 0.062 | 3 | 1.15 | 0.768 | -0.81 | 0.30 |
| 8 | 5 | 6 | 1 | 0 | 12 | 0.600 | 0.971 | 0.289 | 3 | 7.15 | 0.066 | -2.45 | -0.14 |
| 9 | 2 | 1 | 0 | 0 | 3 | 0.150 | 0.610 | 0.117 | 3 | 2.68 | 0.446 | -1.49 | -1.01 |
| 10 | 5 | 7 | 2 | 0 | 14 | 0.700 | 0.881 | 0.269 | 3 | 7.10 | 0.067 | -2.22 | 0.45 |
| 11 | 2 | 3 | 1 | 0 | 6 | 0.300 | 0.881 | 0.062 | 3 | 1.15 | 0.768 | -0.81 | 0.30 |
| 12 | 0 | 1 | 0 | 0 | 1 | 0.050 | 0.286 | 0.069 | 3 | 1.58 | 0.668 | -0.11 | 1.02 |
| 13 | 4 | 5 | 4 | 0 | 13 | 0.650 | 0.934 | 0.277 | 3 | 5.90 | 0.115 | -0.63 | 0.32 |
| 14 | 4 | 7 | 3 | 2 | 16 | 0.800 | 0.722 | 0.067 | 3 | 1.51 | 0.684 | 0.81 | 0.77 |
| 15 | 5 | 8 | 3 | 2 | 18 | 0.900 | 0.469 | 0.112 | 3 | 2.41 | 0.495 | 0.15 | 0.80 |
| 16 | 4 | 7 | 3 | 1 | 15 | 0.750 | 0.811 | 0.056 | 3 | 1.56 | 0.674 | -0.27 | 0.99 |
| 17 | 1 | 1 | 2 | 0 | 4 | 0.200 | 0.722 | 0.109 | 3 | 3.07 | 0.381 | 0.35 | 0.09 |
| 18 | 1 | 4 | 2 | 0 | 7 | 0.350 | 0.934 | 0.139 | 3 | 3.15 | 0.370 | 0.15 | 1.48 |
| 19 | 1 | 4 | 3 | 0 | 8 | 0.400 | 0.971 | 0.214 | 3 | 5.07 | 0.165 | 0.57 | 1.54 |
| 20 | 1 | 2 | 0 | 0 | 3 | 0.150 | 0.610 | 0.090 | 3 | 1.70 | 0.641 | -0.84 | 0.43 |
| 21 | 0 | 2 | 0 | 0 | 2 | 0.100 | 0.469 | 0.144 | 3 | 3.33 | 0.343 | -0.15 | 1.49 |
| 22 | 1 | 2 | 0 | 0 | 3 | 0.150 | 0.610 | 0.090 | 3 | 1.70 | 0.641 | -0.84 | 0.43 |
| 23 | 1 | 4 | 3 | 2 | 10 | 0.500 | 1.000 | 0.243 | 3 | 5.67 | 0.127 | 2.31 | 1.03 |
| 24 | 1 | 2 | 1 | 1 | 5 | 0.250 | 0.811 | 0.030 | 3 | 0.89 | 0.830 | 0.80 | 0.20 |
| 25 | 0 | 0 | 1 | 0 | 1 | 0.050 | 0.286 | 0.124 | 3 | 4.21 | 0.238 | 0.95 | 0.24 |
| 26 | 0 | 1 | 0 | 0 | 1 | 0.050 | 0.286 | 0.069 | 3 | 1.58 | 0.668 | -0.11 | 1.02 |
| 27 | 1 | 1 | 2 | 2 | 6 | 0.300 | 0.881 | 0.269 | 3 | 7.10 | 0.067 | 2.22 | -0.45 |
| 28 | 1 | 1 | 1 | 2 | 5 | 0.250 | 0.811 | 0.237 | 3 | 6.89 | 0.074 | 1.87 | -0.60 |
| 29 | 1 | 0 | 1 | 1 | 3 | 0.150 | 0.610 | 0.153 | 3 | 3.66 | 0.300 | 1.10 | -1.01 |
| 30 | 0 | 0 | 0 | 2 | 2 | 0.100 | 0.469 | 0.469 | 3 | 20.00 | 0.000 | 2.93 | -0.80 |
*-----*-----*-----*-----*-----*-----*-----*-----*-----*-----*
| Ref | 6 | 8 | 4 | 2 |
*-----*-----*-----*-----*-----*

```

```

Mutual information (mean over all species): 1.4469e-01
Entropy (Shannon index): 1.8464e+00

```

Sur le listing (résultats pour chacune des variables) on trouve par exemple pour la variable 1 (en bas) le mode de répartition des 20 relevés entre les 4 modalités. Ainsi, il y a 6 relevés portant la modalité 1, 8 relevés portant la modalité 2, 4 relevés portant la modalité 3 et 2 relevés portant la modalité 4 de la variable 1.

En colonnes les profils écologiques bruts des espèces, c'est-à-dire les nombres de relevés portant la modalité de milieu  $i$  ( $1 \leq i \leq 4$ ) pour lesquels l'espèce  $j$  ( $1 \leq j \leq 30$ ) est présente. Par exemple, l'espèce 8 est présente 12 fois, 5 fois (sur 6 possible) en modalité 1, 6 fois (sur 8) en modalité 2, 1 fois (sur 4) en modalité 3 et est absente (sur 2 cas possibles) en modalité 4 pour la variable 1.

Le tableau PE a 20 lignes-modalités et 30 colonnes-espèces : il contient les profils écologiques bruts. Noter que les profils bruts ne tiennent compte que de la présence-absence des taxons. On peut lister tableau PE avec la commande **Edi**t **with** en utilisant Excel™ comme logiciel d'accès :

PE-t																									
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1	1	1	1	0	3	0	0	0	0	0	0	2	1	0	1	1	1	0	0	3	0				
2	4	2	1	0	5	2	0	0	1	1	3	1	1	3	3	1	1	2	3	1					
3	3	1	1	0	3	2	0	0	0	1	2	1	1	2	2	1	1	2	2	0					
4	3	3	1	0	5	2	0	0	2	2	2	1	0	4	1	2	0	2	3	2					
5	1	2	2	0	3	0	2	0	0	1	3	1	0	2	2	1	1	0	4	0					
6	2	3	1	0	3	3	0	0	1	1	2	0	2	3	3	0	2	2	1	1					
7	2	3	1	0	3	2	1	0	0	1	2	0	3	2	4	0	3	1	2	0					
8	5	6	1	0	7	3	1	1	1	3	3	2	3	4	5	3	3	3	5	1					
9	2	1	0	0	2	1	0	0	1	2	0	0	0	2	0	1	0	2	0	1					
10	5	7	2	0	7	4	1	2	2	3	3	3	3	5	6	3	4	3	5	2					
11	2	3	1	0	3	2	0	1	1	2	2	1	0	4	1	1	0	1	3	2					
12	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	1	0	0	0					
13	4	5	4	0	5	3	2	3	0	2	4	4	3	3	7	3	6	2	5	0					
14	4	7	3	2	6	3	2	5	1	6	4	3	2	5	7	4	4	3	5	4					
15	5	8	3	2	6	4	2	6	2	7	4	3	2	7	7	4	4	3	5	6					
16	4	7	3	1	5	3	2	5	1	6	3	3	2	6	5	4	4	2	5	4					
17	1	1	2	0	1	0	2	1	0	1	1	2	0	1	2	1	2	0	2	0					
18	1	4	2	0	1	1	2	3	0	3	1	2	1	2	3	2	3	1	2	1					
19	1	4	3	0	1	2	2	3	0	1	2	3	2	1	5	2	5	1	2	0					
20	1	2	0	0	1	1	0	1	1	2	0	0	0	2	0	1	0	1	0	2					
21	0	2	0	0	0	1	0	1	1	1	0	0	0	2	0	0	0	0	0	2					
22	1	2	0	0	1	0	0	2	1	2	0	0	0	3	0	0	0	0	0	3					
23	1	4	3	2	0	2	2	6	0	4	1	3	2	2	5	3	5	0	2	3					
24	1	2	1	1	0	0	1	4	0	3	0	2	0	2	1	2	1	0	2	2					
25	0	0	1	0	0	0	0	1	0	0	0	1	0	0	1	0	1	0	0	0					
26	0	1	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0	1	0	0					
27	1	1	2	2	0	0	0	6	0	3	0	3	0	1	2	3	2	0	1	3					
28	1	1	1	2	0	0	0	5	0	3	0	2	0	1	1	3	1	0	1	3					
29	1	0	1	1	0	0	0	3	0	2	0	1	0	1	0	2	1	0	0	2					
30	0	0	0	2	0	0	0	2	0	2	0	0	0	0	1	1	0	0	0	2					
32	6	8	4	2	7	4	2	7	2	7	4	4	3	7	8	5	6	3	5	6					

Transposer dans le tableur pour éditer les espèces en lignes. Il est clair que les profils bruts n'ont de sens que comparés au mode de répartition des relevés entre les modalités, l'écart entre profil brut et profil marginal définissant les préférences des espèces. La théorie des profils écologiques (Godron, 1968<sup>2</sup> ; Gounot, 1969<sup>3</sup> ; Daget & Godron, 1982<sup>4</sup>) vise à définir le niveau de signification statistique de ces préférences.

Le listing indique en outre que (pour la variable 2) :

Variable: 2, categories: 4

```

*-----*-----*-----*-----*-----*-----*-----*-----*-----*
|Spec| 1 | 2 | 3 | 4 | Tot| Freq|Entro|MutIn|DF| Khi2 |Proba|DTest|ITest|
*-----*-----*-----*-----*-----*-----*-----*-----*-----*
| 1| 3| 0| 0| 0| 3|0.150|0.610|0.265| 3| 6.55|0.086|-2.07|-1.90|
| 2| 5| 2| 0| 0| 7|0.350|0.934|0.432| 3| 9.32|0.025|-2.90|-0.85|
...

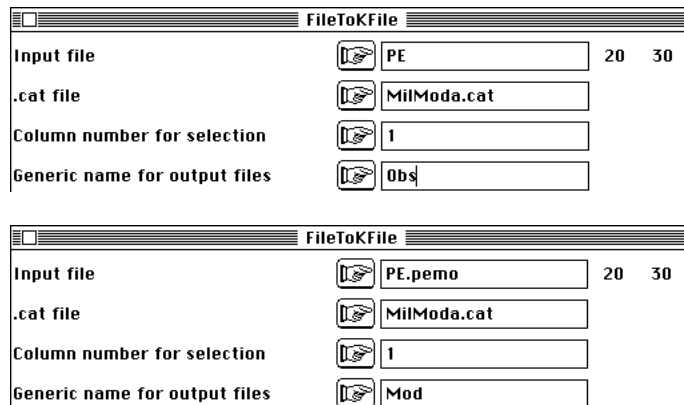
```

- l'espèce 2 est présente 7 fois (Tot), soit une fréquence (Freq) de 35% ;
- l'information mutuelle espèce-facteur (Daget & Godron, op. cit., p. 50) vaut 0.432
- le nombre de degrés de libertés (DF) de la table de contingence (2 lignes présence-absence et 4 colonnes-modalités) vaut 3 ;
- le Khi2 de cette table est 9.32, valeur qui est dépassée sous l'hypothèse nulle avec une probabilité de 2.5% ;
- le test du gradient direct (D) (Gros & Chessel, 1982<sup>5</sup> ; M'hamdi, 1984<sup>6</sup>, p. 106) vaut -2.90, pour une loi normale sous l'hypothèse nulle et indique que la moyenne des valeurs des relevés où l'espèce est présente est inférieure à la moyenne attendue (profil significatif à gauche) ;
- le test du gradient inverse (I) vaut -0.85, pour une loi normale sous l'hypothèse nulle. Le test D porte sur la moyenne (repérage des profils décalés à gauche et à droite), le test I porte sur la variance (détection des profils concentrés au milieu ou en U). Les deux n'ont de sens que pour des variables à modalités ordonnées.

La table de contingence correspondant à l'espèce 2 est déduite du tableau PE est donc significative, l'espèce 2 étant sureprésentée dans la première modalité :

					Total
Présence	5	2	0	0	7
Absence	2	2	2	7	13
Total	7	4	2	7	20

L'abondance des tests pose, à l'évidence, le problème de la synthèse de ces profils dont bon nombre sont significatifs dans les enquêtes phyto-écologiques. Pour représenter ces profils, on pourra toujours séparer simplement les profils observés et calculés par variables (FilesUtil) :



Par exemple :

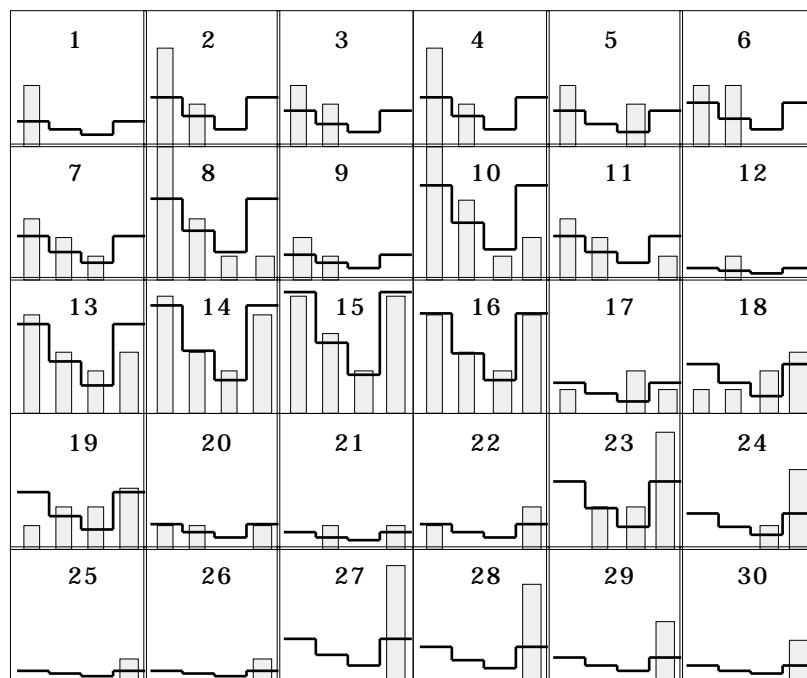


Figure 2 Profils écologiques de 30 espèces sur la variable humidité.

La question des centrages de ces profils offre trois possibilités :

(1) **Centrage de type ACP décentrée** : le profil de l'espèce 2 pour la variable 2 est (5/7,2/7,0,0), le profil marginal est (7/20,4/20,2/20,7/20), le profil centré est la différence entre les deux (.3643,.0857,-.10,-.35). Il est possible, pour un facteur limitant

que tous les profils soient identiques en étant significatifs. Par exemple dans le cas d'une pollution qui limite la fréquence de toutes les espèces.

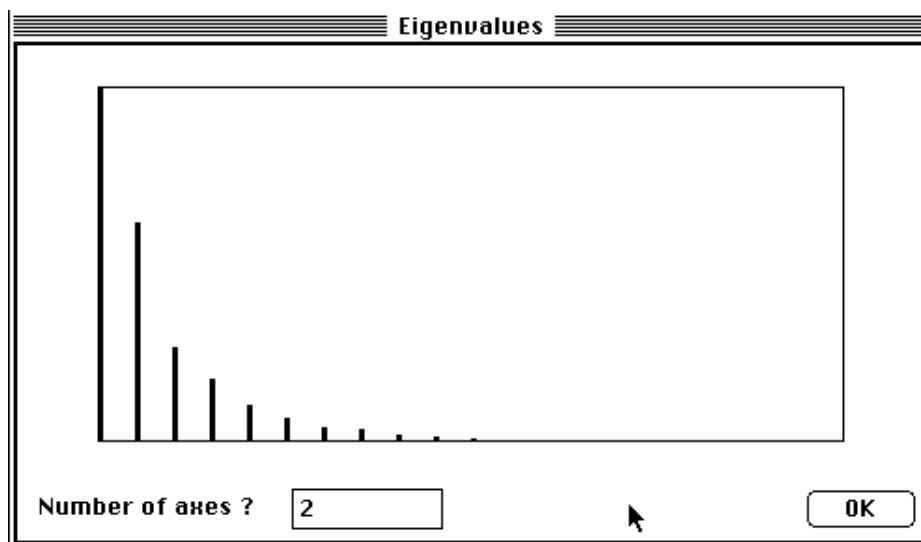
(2) **Centrage de type AFC** : en additionnant toutes les présences de toutes les espèces on trouve 197, soit par classes les fréquences (71/197, 42/197, 22/197 et 62/197). Pour la végétation les deux systèmes sont voisins.

(3) La fréquence de l'espèce sur tous les relevés est de 7/20, donc les écarts par classes sont  $5/7-7/20 = .3643$ ,  $2/7-7/20 = .15$ ,  $-.35$  et  $-.35$ .

## 4.2 - Analyse factorielle des correspondances du tableau de profils écologiques

Romane <sup>7</sup> a eu le premier l'idée de faire la synthèse des profils écologiques par l'AFC du tableau des profils bruts. Cette idée est reprise sans nouveauté dans Montana & Greig-Smith (1990)<sup>8</sup>.

Exécuter l'option **CORrespondence Analysis** du module **COA** sur le fichier PE et conserver deux axes.



Le total affiché pour le fichier PE vaut 985 soit 5 fois (5 variables) 197 (le nombre total de présences toutes espèces confondues). A l'aide de l'option **Labels** du module **Scatters**, on obtient les cartes (Fig. 3A, 3B) respectivement des espèces (PE. fcl i) et des modalités (PE. fcco). Le facteur 1 est un gradient profondeur-humidité du sol. On a observé que ce type d'analyse fait correctement la synthèse des profils mais perd la dimension relevés. Or les relevés sont présents dans le tableau de végétation et le tableau de milieu qui génèrent ces profils. Il est important de tenir compte du seul usage des présences-absences.

Par conséquent, transformer le tableau VegTr (30-20) en présence-absence. Pour cela utiliser l'option **0-1 Reducing** du module **Bin->Bin** :



Le fichier *veg01* contient 30 lignes-espèces et 20 colonnes-relevés avec la notation 1 si l'espèce est présente, 0 sinon. Les colonnes de ce tableau sont donc potentiellement des colonnes supplémentaires dans l'AFC du tableau PE (30 espèces et 20 modalités). La pratique des éléments supplémentaires a été introduite en phyto-écologie par Bonin & Roux (1978)<sup>9</sup>.

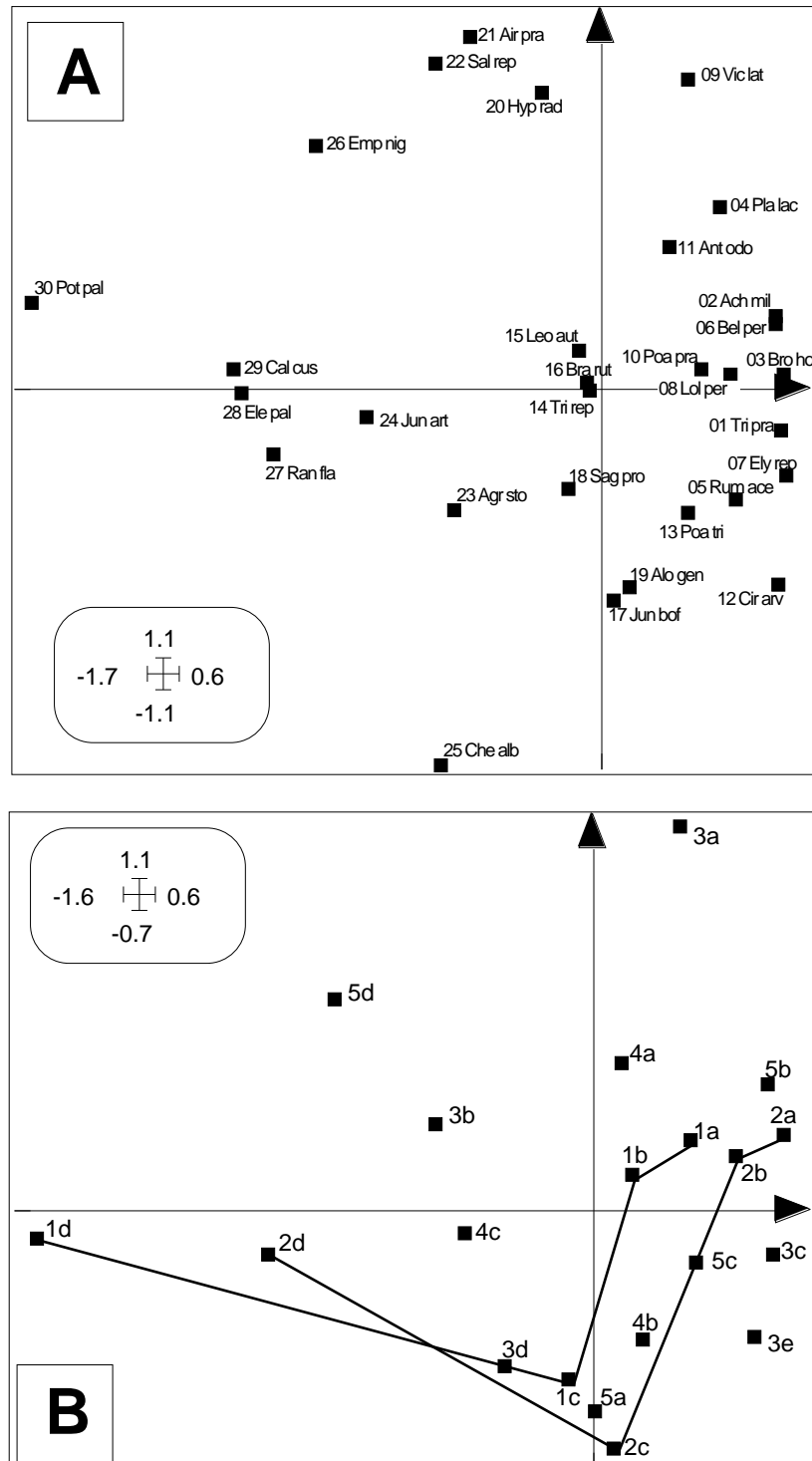


Figure 3 Analyse factorielle des correspondances du tableau de profils écologiques. A - Carte des espèces (identifiées par une chaîne de caractères). B - Carte des modalités de variables de milieu (le chiffre identifie la variable ; la lettre identifie la modalité).

Utiliser l'option **Supplementary columns** du module **DDUtil** :

Supplementary columns			
Initial analysis (up)		PE.fcup	20 2
Supplementary column file		Veg01	30 20
Output file name		PEsc	
<input type="button" value="Quit"/>		<input type="button" value="Ok"/>	

Title of the analysis: PE.fcta  
 Number of rows: 30, columns: 20

File PEsc contains the coordinates of the projected columns  
 It has 20 rows and 2 columns

File : PEsc  
 ----- Minimum/Maximum:  
 Col.: 1 Mini = -1.6914 Maxi = 0.86189  
 Col.: 2 Mini = -0.94715 Maxi = 1.3029

En outre les relevés sont les lignes du tableau de milieu (Mi 1). Comme l'ACM de ce tableau est l'AFC du tableau disjonctif associé, utiliser l'option **Categ->Disj** du module **CategVar** :

Categ -> Disj	
---.cat type file	Mil.cat
<input type="button" value="Quit"/>	<input type="button" value="Ok"/>

File Mil01 contains dummy coded variable file  
 It has 20 rows and 20 columns (categories)

File Mil01.num contains the number of modalities for each variable  
 It has 5 rows (variables) and 1 column

Le fichier Mil01 contient 20 lignes-relevés et 20 colonnes-modalités avec la notation 1 si le relevé est porteur de la modalité, 0 sinon. Ce tableau est potentiellement un tableau de lignes supplémentaires dans l'AFC de PE.

Utiliser l'option **Supplementary rows** du module **DDUtil** :

Supplementary rows			
Initial analysis (up)		PE.fcup	20 2
Supplementary row file		Mil01	20 20
Output file name		PEsr	
<input type="button" value="Quit"/>		<input type="button" value="Ok"/>	

Title of the analysis: PE.fcta  
 Number of rows: 30, columns: 20

File PEsr contains the coordinates of the projected rows  
 It has 20 rows and 2 columns

File : PEsr  
 ----- Minimum/Maximum:

Col. : 1 Mini = -1.7622 Maxi = 0.87155  
 Col. : 2 Mini = -1.2187 Maxi = 1.363

On a donc deux positions des relevés, l'une venant du tableau de milieu (PEsr) et l'autre venant du tableau de végétation (PEsc).

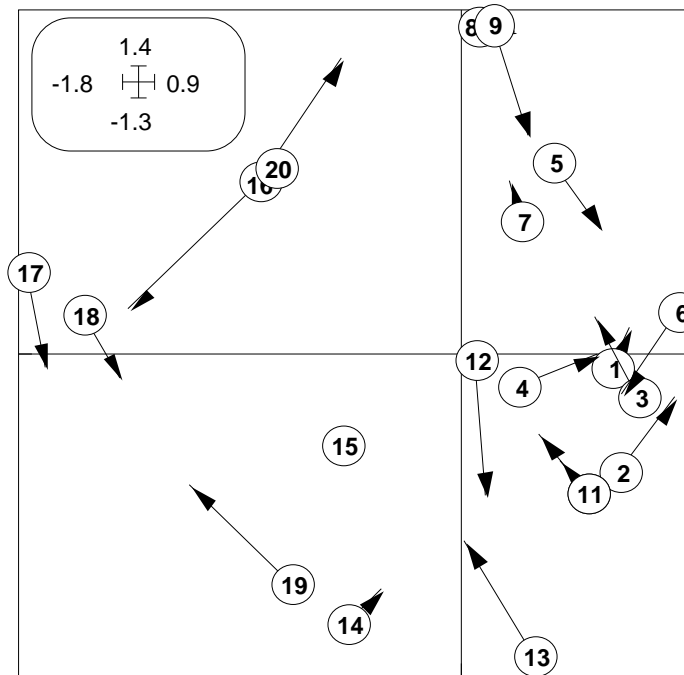
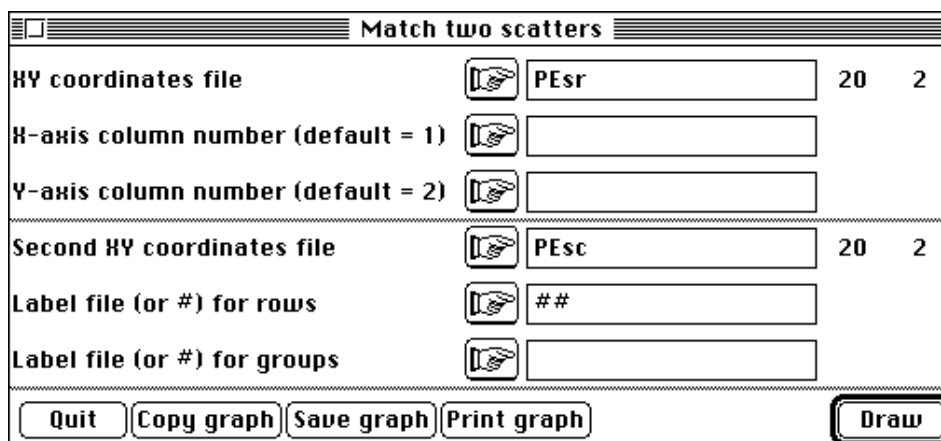


Figure 4 Superposition des relevés dans l'AFC d'un tableau de profils écologiques. A l'origine des flèches (étiquettes) on trouve la position des relevés comme lignes supplémentaires (milieu). A l'extrémité des flèches on trouve la position des relevés comme colonnes supplémentaires (végétation). A gauche les relevés à sol profond et humide.

Superposer les deux systèmes avec l'option **Match two scatters** du module **Scatters** (Fig. 4) :



En fait, on vient de faire ici une analyse de co-inertie entre le tableau de végétation en présence-absence et le tableau des variables de milieu en qualitatif.

### 4.3 - Comparaison avec le couplage AFC-ACM

Partir du tableau Veg01 (30 lignes-espèces/20 colonnes-relevés). Le transposer par l'option **Transpose** du module **FilesUtil** en un tableau VegPA (20 lignes-relevés/30

colonnes-espèces en présence-absence). Exécuter le module **COA** (option **Correspondence Analysis**) sur VegPA qui annonce un total de 197 (nombre total de présences d'espèces) et garder sans raison 5 facteurs. Les valeurs propres sont .5221, .3677, .2400, .1886 et .1538.

Reprendre le tableau Mi1 (20 relevés/5 variables qualitatives) et faire son analyse des correspondances multiples à l'aide du module **MCA** en prenant soin de déclarer comme fichiers d'entrée, outre Mi1.cat, le fichier des poids VegPA.fcpl, afin de rendre compatible les deux analyses et garder sans raison 5 facteurs. Les valeurs propres sont .5364, .4903, .4387, .3797 et .3035. Coupler Mi1.cmpa et VegPA.fcpa par l'option **Match two statistical triplets** du module **CoInertia** dans une analyse de nom générique Cou puis l'option **Coinertia analysis**. Comparer alors les fichiers :

- Cou<>co, coordonnées des espèces dans l'analyse de co-inertie
- PE.fcoco, positions des espèces dans l'AFC du tableau de profils écologiques (PE)
- Cou<>li, coordonnées des modalités dans l'analyse de co-inertie
- PE.fcli, positions des modalités dans l'AFC de PE
- Cou<>l1, coordonnées des relevés (milieu) dans l'analyse de co-inertie
- PEsrl, positions des relevés (milieu) en supplémentaires dans l'AFC de PE
- Cou<>l2, coordonnées des relevés (végétation) dans l'analyse de co-inertie
- PEscl, position des relevés (végétation) en supplémentaires dans l'AFC de PE

On vérifie qu'ils contiennent deux à deux exactement la même information (au signe près éventuellement, ce qui est sans importance). La démonstration mathématique correspondante est dans l'article de Chessel & Mercier (1993)<sup>10</sup>.

## 5 - Analyse du couple milieu-végétation

La notion de profils écologiques peut être étendue de manière optimale aux données quantitatives sur la végétation (AFC sur le tableau non réduit en présence-absence).

### 5.1 - Calcul et test de la co-structure

Coupler les deux analyses séparées précédemment réalisées avec l'option **Matching two statistical triplets** du module **CoInertia** :

Matching two statistical triplets			
First data input file		Mil.cmta	20 20
Second data input file		Veg.fcta	20 30
Output file name		MV	
<input type="button" value="Quit"/>		<input type="button" value="Ok"/>	

First input statistical triplet: table Mil.cmta  
 Number of rows: 20, columns: 20  
 Second input statistical triplet: table Veg.fcta  
 Number of rows: 20, columns: 30

File MV.<>ta contains the crossed table YtDnX  
 It has 30 rows and 20 columns

File MV.<>pl contains the weight of each row  
 It has 30 rows and 1 column

File MV.<>pc contains the weight of each column



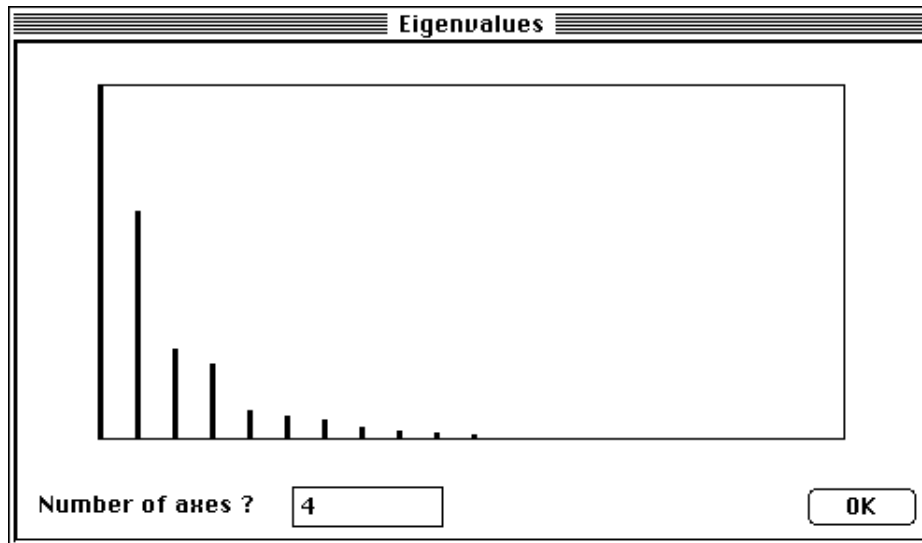
It has 20 rows and 1 column

File MV.<>ma contains the name of the input statistical triplet  
 1 ---> Mil.cmta  
 2 ---> Veg.fcta  
 total inertia: 0.482025

Dans le cas de ce couplage (AFC/ACM), le tableau croisé MV.<>ta est un tableau de profils écologiques de terme général  $q_{ij}$  qui peut être obtenu par la formule suivante :

$$q_{ij} = \frac{\% \text{ de l'abondance de l'espèce } i \text{ dans la modalité } j}{\% \text{ de l'abondance totale dans la modalité } j} - 1$$

Exécuter l'option **Coinertia analysis** du module **CoInertia** :



DiagoRC: General program for two diagonal inner product analysis  
 Input file: MV.<>ta  
 --- Number of rows: 30, columns: 20  
 -----  
 Total inertia: 0.482025  
 -----

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+1.9992E-01	+0.4147	+0.4147	02	+1.2909E-01	+0.2678	+0.6826
03	+5.1461E-02	+0.1068	+0.7893	04	+4.3083E-02	+0.0894	+0.8787
05	+1.6067E-02	+0.0333	+0.9120	06	+1.3061E-02	+0.0271	+0.9391
07	+1.0983E-02	+0.0228	+0.9619	08	+7.1845E-03	+0.0149	+0.9768
09	+4.2557E-03	+0.0088	+0.9856	10	+3.2887E-03	+0.0068	+0.9925
11	+2.1854E-03	+0.0045	+0.9970	12	+8.3616E-04	+0.0017	+0.9987
13	+4.7913E-04	+0.0010	+0.9997	14	+9.9501E-05	+0.0002	+0.9999
15	+2.8314E-05	+0.0001	+1.0000	16	+0.0000E+00	+0.0000	+1.0000
17	+0.0000E+00	+0.0000	+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

File MV.<>vp contains the eigenvalues and relative inertia for each axis  
 --- It has 20 rows and 2 columns

File MV.<>co contains the column scores  
 --- It has 20 rows and 4 columns

.....

File MV.<>li contains the row scores

```

--- It has 30 rows and 4 columns
.....
-----
coinertia analysis between two statistical triplets
  1 ---> Mil.cmta (rows: 20, col: 20, axes: 7, inertia: 3.000000)
  2 ---> Veg.fcta (rows: 20, col: 30, axes: 5, inertia: 2.113800)
Coinertia: 0.482025, RV coefficient: 0.590059

MV.<>w1 is a binary file with 20 rows and 4 columns
It contains the canonical weights of the variables of table 1
.....

MV.<>co is a binary file with 20 rows and 4 columns
It contains the coordinates of the columns of table 1
In earlier version of ADE its name is MV.<>c1
.....

MV.<>w2 is a binary file with 30 rows and 4 columns
It contains the canonical weights of the variables of table 2
.....

MV.<>li is a binary file with 30 rows and 4 columns
It contains the coordinates of the columns of table 2
In earlier version of ADE its name is MV.<>c2
.....

MV.<>l1 is a binary file with 20 rows and 4 columns
It contains the coordinates of the rows (table 1)
.....

MV.<>l2 is a binary file with 20 rows and 4 columns
It contains the coordinates of the rows (table 2)
.....

MV.<>m1 is a binary file with 20 rows and 4 columns
It contains the normalized coordinates of the rows (table 1)
In earlier version of ADE this file name uses suffix bl1
.....

MV.<>m2 is a binary file with 20 rows and 4 columns
It contains the normalized coordinates of the rows (table 2)
In earlier version of ADE this file uses suffix bl2
.....

MV.<>a1 is a binary file with 7 rows and 4 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
.....

MV.<>a2 is a binary file with 5 rows and 4 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
.....

```

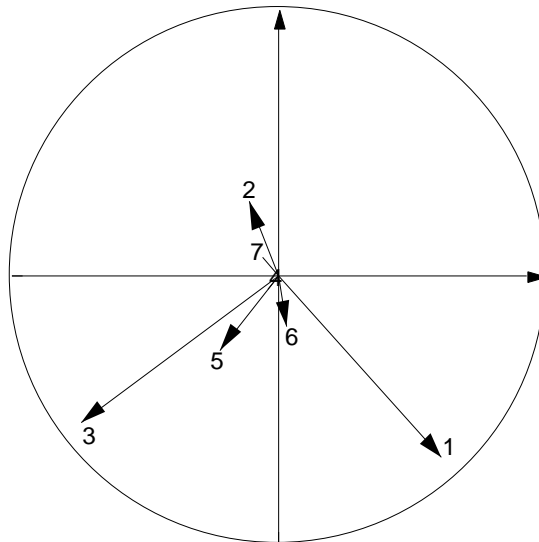


Figure 5 Projection des axes d'inertie de l'analyse préliminaire (7 facteurs) sur le plan 1-2 de co-inertie .

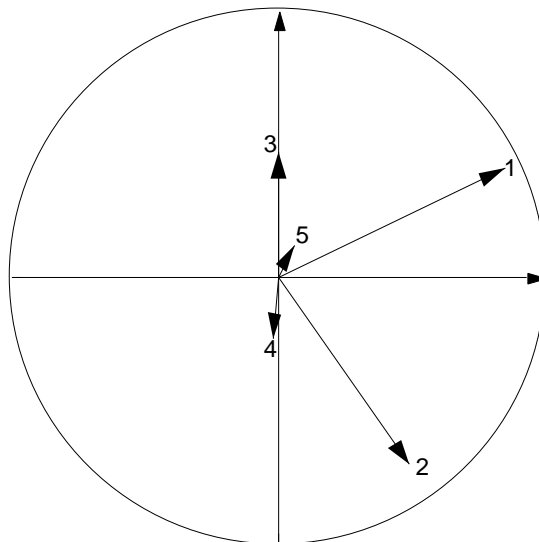


Figure 6 Projection des axes d'inertie de l'analyse préliminaire (5 facteurs) sur le plan 1-2 de co-inertie.

MV. <>a1 place les axes de l'ACM de  $M_i 1$  sur les axes de co-inertie de  $M_i 1$  (Fig. 5).

MV. <>1 1 place les lignes de  $M_i 1$  à la moyenne des modalités qu'ils portent lorsque ces modalités sont positionnées par MV. <>w1.

MV. <>a2 place les axes de l'AFC de  $V_{eg}$  sur les axes de co-inertie de  $V_{eg}$  (Fig. 6).

On peut noter sur les Fig. 5 et 6, la mise en correspondance automatique des axes 1 (respectivement 3) de  $M_i 1$  et 2 (respectivement 1) de  $V_{eg}$  ou encore du plan d'inertie 1-3 de  $M_i 1$  et du plan d'inertie 2-1 de  $V_{eg}$  avec les rotations qui s'imposent.

MV. <>1 2 place les lignes de  $V_{eg}$  à la moyenne des espèces qu'elles contiennent lorsque ces espèces sont positionnées par MV. <>w2.

MV. <>co contient exactement les coordonnées des modalités de  $M_i 1$  dans l'AFC du tableau des profils écologiques du couple milieu-végétation.

MV. <>l1 contient exactement les coordonnées des espèces de Veg dans l'AFC du tableau des profils écologiques du couple milieu-végétation.

MV. <>m1 contient les coordonnées des lignes de Mi 1 de variance 1.

MV. <>m2 contient les coordonnées des lignes de Veg de variance 1.

On commencera toujours par la lecture des paramètres fondamentaux :

Num	Covari a.	Vari an1	vari an2	Correl a.	INER1	INER2
1	0.4471	0.4763	0.4935	0.9223	0.5483	0.5374
2	0.3593	0.4699	0.3609	0.8725	0.4757	0.4014
3	0.2268	0.431	0.1817	0.8107	0.4504	0.2601
4	0.2076	0.333	0.2318	0.7471	0.3712	0.1751

Covari a. : Covariances entre coordonnées dans MV. <>l1 et MV. <>l2 = racines carrées des valeurs propres = quantités optimisées sous contraintes de normalisation dans MV. <>w1 et MV. <>w2.

Vari an1 = variances des coordonnées de MV. <>l1 = inertie projetée sur les axes de co-inertie.

Vari an2 = variances des coordonnées de MV. <>l2 = inertie projetée sur les axes de co-inertie.

INER1 = inertie projetée sur les axes d'inertie de Mi 1 (comparer avec Vari an1).

INER2 = inertie projetée sur les axes d'inertie de Veg (comparer avec Vari an2)

Correl ati on = corrélation entre coordonnées de MV. <>l1 et MV. <>l2.

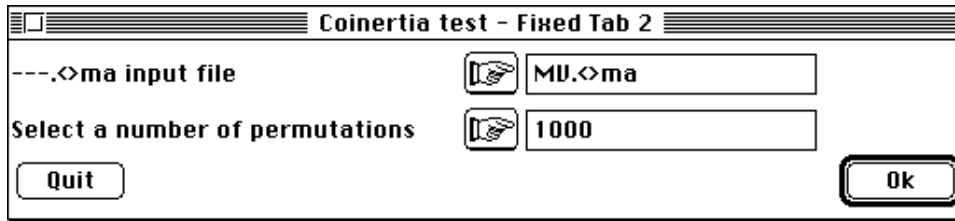
Les inerties projetées sur le plan 1-2 de co-inertie de Mi 1 valent 0.476 et 0.469. Elles sont dans la fourchette des inerties projetées sur les axes 1-2-3 (0.548, 0.475, 0.45). On sait déjà que le plan 1-2 de co-inertie de Mi 1 est à peu près le plan 1-3 d'inertie (Fig. 5). Les inerties projetées sur le plan 1-2 de co-inertie de Veg valent 0.493 et 0.36. Elles sont inférieures aux inerties projetées sur les axes 1-2 (0.537, 0.401). On sait également que le plan 1-2 de co-inertie est voisin du plan 1-2 d'inertie (Fig. 6).

On constate qu'à partir du facteur 3, l'inertie projetée sur les axes de co-inertie de Veg (Vari an2 = 0.1817) diminue pour approcher l'inertie projetée sur l'axe d'inertie de Veg (INER2 = 0.1751). Le fait que la valeur de Vari an2 réaugmente au facteur 4 souligne qu'en analyse de co-inertie une sorte de compensation s'établit entre corrélation et inertie projetées suivant le principe :

$$\text{Maximum} = \text{Covariance} = \text{Correlation} * (\text{Vari an1})^{1/2} * (\text{Vari an2})^{1/2}$$

L'analyse recherche donc le meilleur compromis alliant corrélation et structure des deux tableaux. On ne discutera ici que du plan 1-2. Les corrélations sont très bonnes (0.922, 0.872). Une bonne partie des structures des deux tableaux sont donc en correspondance.

Finalement on peut tester la relation globale entre milieu et végétation à l'aide de l'option **Coinertia test - Fixed Tab 2** du module **CoInertia** qui signifie que l'on permute et recentre le tableau 1 (Mi 1) en gardant fixe la pondération (AFC) du tableau 2 (Veg) :

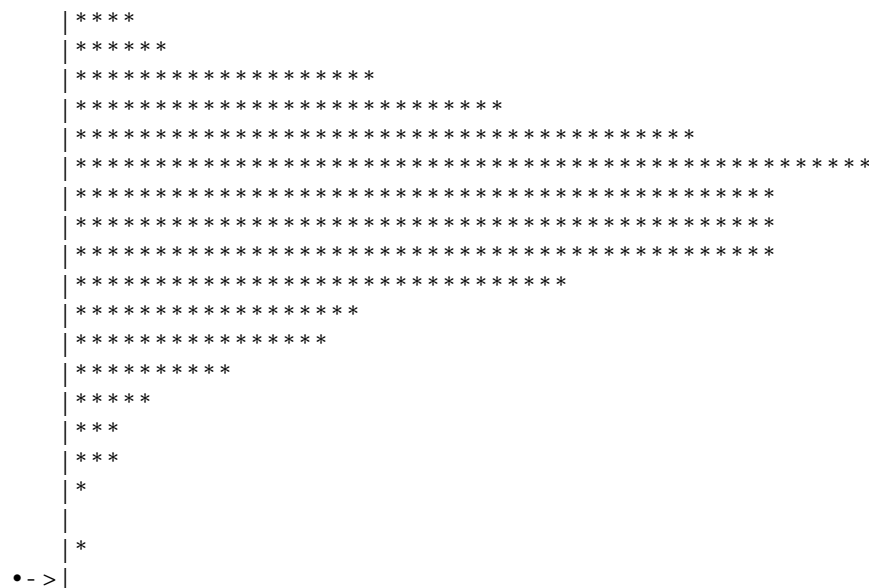


Warning: non uniform weighting. The results from permutations are valid only if the row weights come from the fixed table.

```

-----
The fixed table is table 2: Veg.fcta
number of random matching: 1000   Observed: 0.482025
Histogramm:  minimum = 0.242448, maximum = 0.482025
number of simulation X<Obs: 1000 (frequency: 1.000000)
number of simulation X>=Obs: 0 (frequency: 0.000000)

```



Cet histogramme indique qu'aucune valeur résultant des simulations n'est supérieure à la valeur observée d'inertie totale ce qui souligne la bonne liaison entre les deux tableaux déjà suggérée par les bonnes valeurs de corrélation et permet de poursuivre l'interprétation.

## 5.2 - Averaging : espèces -> relevés -> modalités

Les aides graphiques à l'interprétation sont en partie celles utilisées pour les analyses classiques en ACP, AFC, ou ACM (graphe des valeurs propres, cartes factorielles des variables et des relevés). Il est cependant important de noter que les tableaux issus de l'analyse couplée s'inscrivent dans le schéma suivant :

```

MV. <>w1 ----- m. p. -----> MV. <>l 1 ----- m. p. -----> MV. <>l i
MV. <>co <----- m. p. ----- MV. <>l 2 <----- m. p. ----- MV. <>w2

```

où m. p. désigne l'opération de **moyennage pondérée**. Dans MV. <>w2 on trouve des codes des espèces centrés et de variance unité (Fig. 7A). Les relevés peuvent être positionnés à la moyenne des espèces qu'ils contiennent (ces moyennes sont dans MV. <>l 2) (Fig. 7B).

Si on cherche une variance maximum de ces positions moyenne des espèces, on fera simplement l'AFC de Veg. Ce n'est pas le cas, mais, puisque les relevés sont positionnés par des codes numériques, on peut placer les modalités à la moyenne des relevés qui les portent. Ces moyennes sont dans MV. <>co (Fig. 7C). Pour faire le tracé de la Fig. 7C, utiliser l'option **Labels** du module **Scatters** avec les paramètres :

Sélectionner les sous-graphiques à l'aide de l'option **Row & Col. selection** :

Si on cherche à maximiser la variance des positions des modalités on fera l'ACM de Mi 1. Ce n'est pas le cas, ici on part d'une variance unité des espèces pour obtenir une variance optimale des modalités par un double averaging : on fait donc en partie l'ACM de Mi 1 et en partie l'AFC de Veg.

La Fig. 7 constituée des espèces positionnées par MV. <>w2 (Fig. 7A), des relevés positionnés par MV. <>12 (Fig. 7B) et des modalités par MV. <>co (Fig. 7C) exprime donc un double averaging optimal.

### 5.3 - Averaging : modalités -> relevés -> espèces

La démarche est identique pour la première ligne du schéma du paragraphe précédent. Dans MV. <>w1 on trouve des codes de modalités centrés par variable et de variance moyenne par variable égale à l'unité. Les relevés peuvent être positionnés à la moyenne des modalités qu'ils portent (ces moyennes sont dans MV. <>1 1).

Si on cherche une variance maximum de ces positions on fait simplement l'ACM de Mi 1. Ce n'est pas le cas, mais puisque les relevés sont positionnés par des codes

numériques, on peut placer les espèces à la moyenne de leur distribution de fréquences. Ces moyennes sont dans MV. <>1 i .

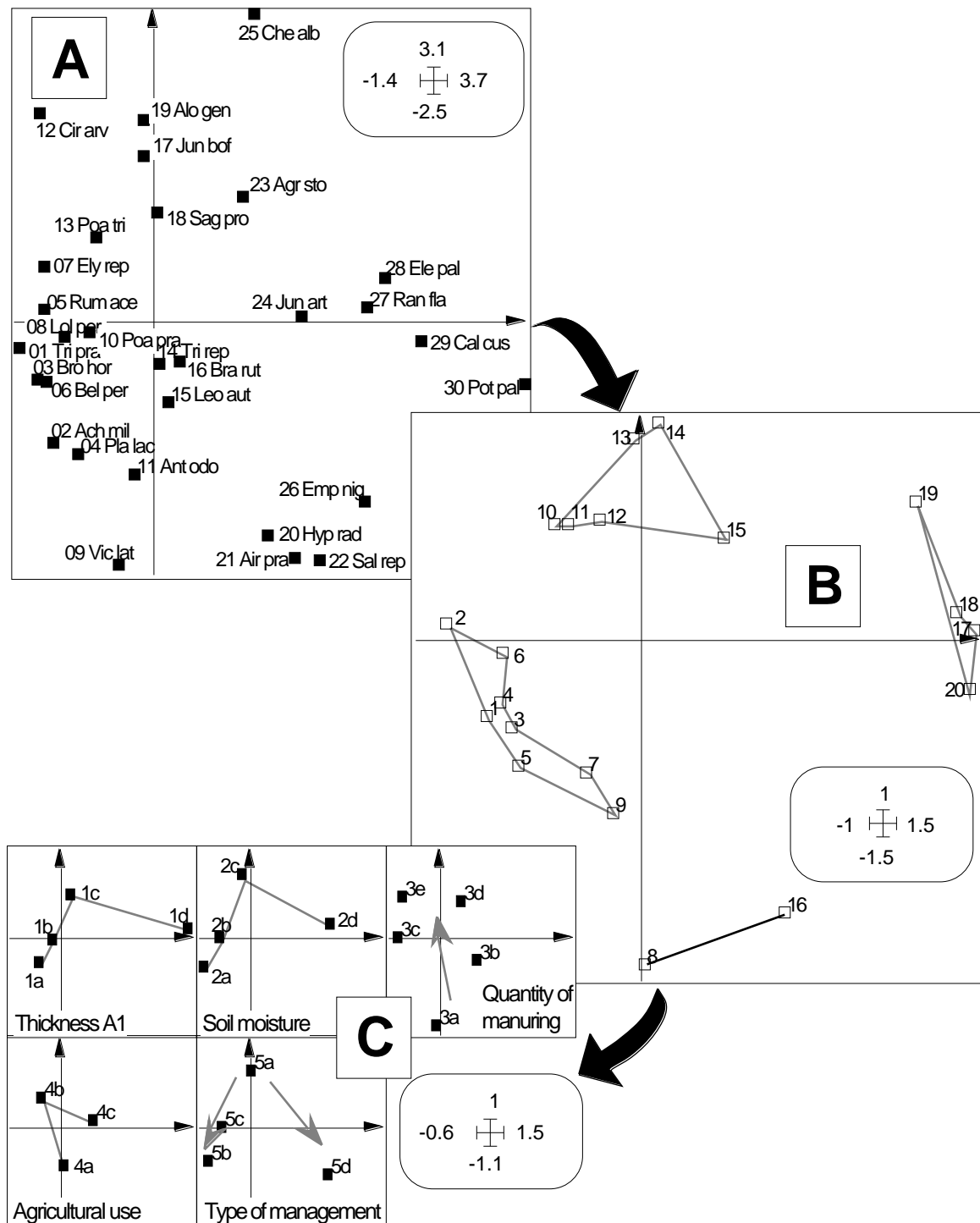


Figure 7 Cartes factorielles dans le plan 1-2 de l'analyse de co-inertie relatives au tableau floristique. A - Carte des espèces (identifiées par une code et un numéro) selon MV. <>w2. B- Carte des relevés selon MV. <>1 2. C - Cartes des modalités de variables de milieu selon MV. <>co.

Si on cherchait à maximiser la variance des positions des espèces, on ferait l'AFC de Veg ou l'AFCVI de Veg sur Mi 1. Ce n'est pas le cas. On part d'une variance moyenne unité des modalités pour obtenir une variance optimale des espèces par un double

averaging : on fait donc en partie l'ACM de Mi 1 et en partie l'AFC de Veg ou l'AFCVI de Veg sur Mi 1.

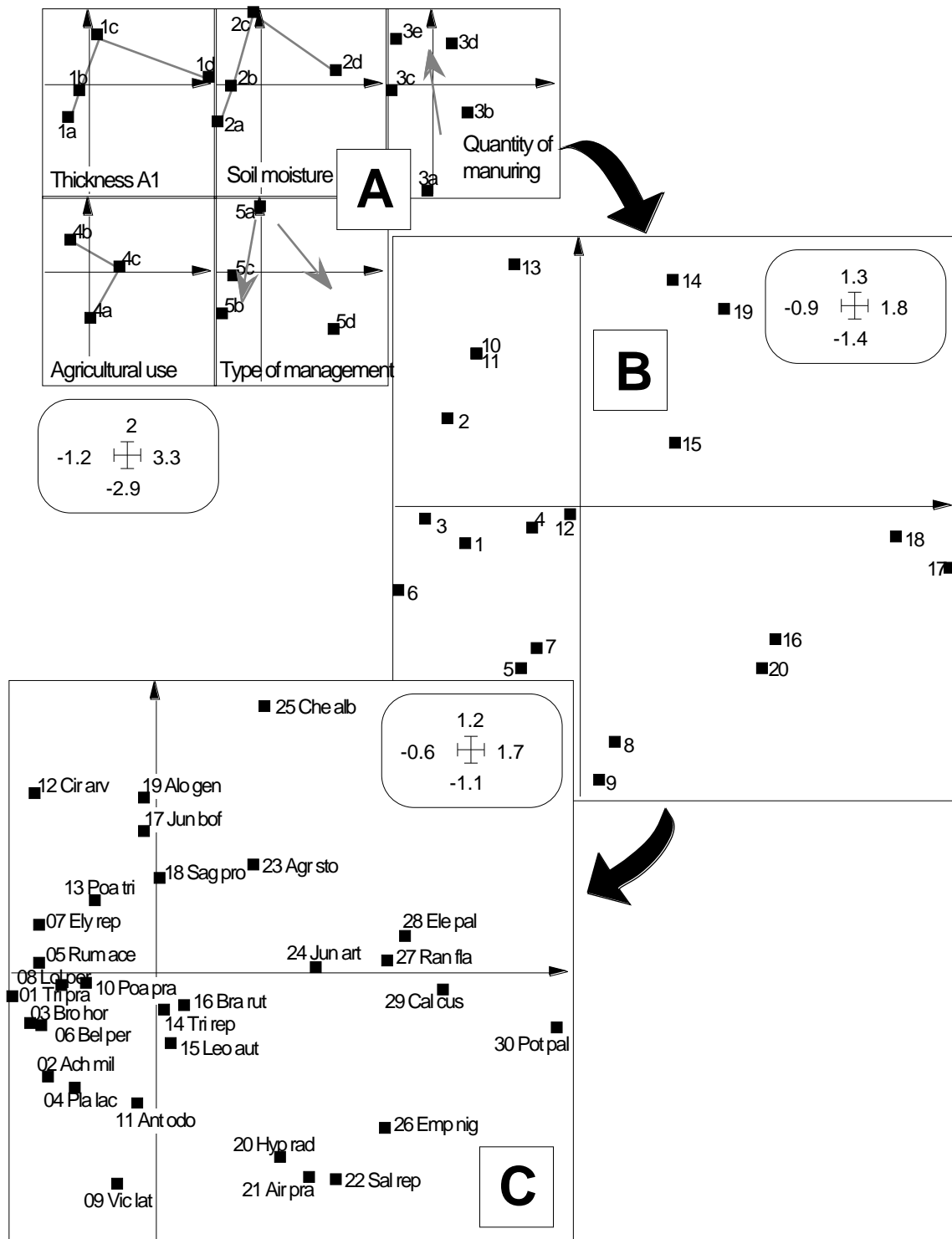


Figure 8 Cartes factorielles dans le plan 1-2 de l'analyse de co-inertie et relatives au tableau milieu. A - Cartes des modalités des variables de milieu selon MV.  $\langle \rangle w1$ . B - Carte des relevés selon MV.  $\langle \rangle I 1$ . C - Carte des espèces selon MV.  $\langle \rangle I i$ .

La Fig. 8 constituée des modalités positionnées par MV.  $\langle \rangle w1$  (Fig. 8A), des relevés positionnés par MV.  $\langle \rangle I 1$  (Fig. 8B) et des espèces par MV.  $\langle \rangle I i$  (Fig. 8C) exprime donc également un double averaging optimal.



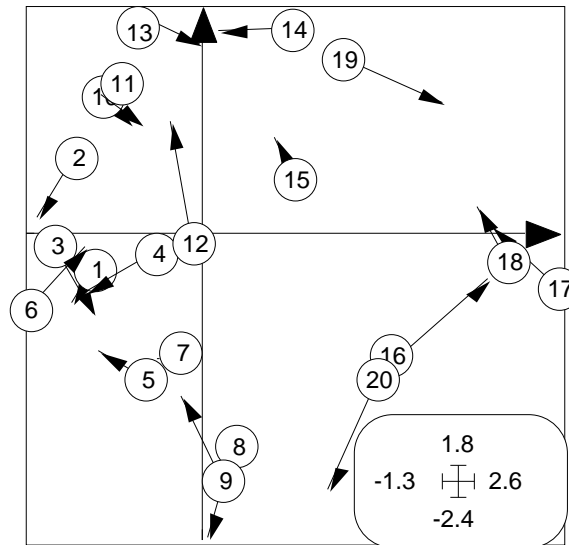


Figure 9 Confrontation des coordonnées des relevés sur les plans F1-F2 de l'analyse de co-inertie. Le début de chaque vecteur est positionné par les coordonnées associées au tableau de milieu (MV. <>m1) et l'extrémité par les coordonnées associées au tableau floristique (MV. <>m2).

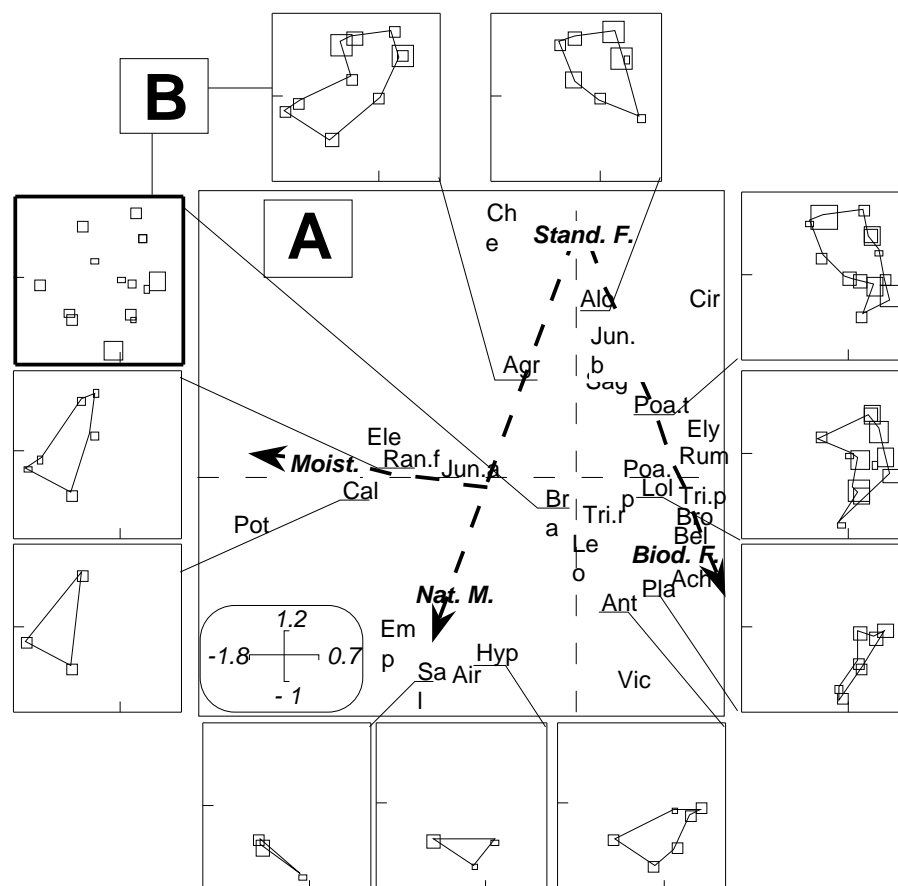


Figure 10 Graphiques de la distribution des espèces sur le plan des relevés associé au tableau milieu. A - Cartes des espèces positionnées à la moyenne des coordonnées des relevés qui les contiennent (MV. <>Ii). B - Cartes par carrés de la distribution de quelques espèces sur le plan des relevés (MV. <>I1, les carrés sont de dimension proportionnelle à l'abondance de l'espèce dans le relevé).

## 5.4 - Double dilatation

On observe immédiatement (et on démontre) que les positions relatives obtenues par  $MV. \langle w1 \rangle$  et  $MV. \langle co \rangle$  ou  $MV. \langle w2 \rangle$  et  $MV. \langle l1 \rangle$  sont les mêmes. En fait on obtient un système à partir de l'autre en multipliant ou en divisant par la racine carrée de la valeur propre correspondante. Comme la symétrie est totale on se contentera d'un des systèmes de représentation. Ici l'usage des deux systèmes explicitent le fonctionnement de l'analyse.

## 5.5 - Coordonnées normalisées

Parmi les propriétés de l'analyse, la plus importante est la maximisation de la covariance des coordonnées.  $MV. \langle w1 \rangle$  positionne les modalités et  $MV. \langle l1 \rangle$  les relevés à la moyenne des modalités portées.  $MV. \langle w2 \rangle$  positionne les espèces et  $MV. \langle l2 \rangle$  les relevés à la moyenne des espèces qu'ils contiennent. Axe par axe la covariance des coordonnées des deux systèmes est optimale.

Maximiser la covariance c'est maximiser simultanément les deux variances et la corrélation. L'analyse est donc *pro parte* une analyse canonique. Cette propriété s'exprime dans les paramètres calculés (cf. ci-dessus) et graphiquement par superposition des plans des coordonnées normalisées.

Sélectionner l'option **Match two scatters** du module **Scatters** (Fig. 9) :

Field	Value	Width	Height
HY coordinates file	MU.<m1	20	4
H-axis column number (default = 1)			
Y-axis column number (default = 2)			
Second HY coordinates file	MU.<m2	20	4
Label file (or #) for rows	##		
Label file (or #) for groups			

Buttons: Quit, Copy graph, Save graph, Print graph, Draw

On pourra ainsi discuter des disparités éventuelles de la position d'un relevé dans les tableaux floristique et environnemental.

## 5.6 - Amplitude des espèces

Cette discussion peut être basée sur les représentations graphiques de la distribution des espèces sur le plan des relevés associé au tableau milieu (Fig. 10).

## 5.7 - Rôle des variables

Exactement symétrique de la précédente, cette discussion peut être basée sur les représentations graphiques de la distribution des modalités sur le plan des relevés associé au tableau de végétation. Dans la Fig. 11 on a utilisé l'option **Stars** du module **Scatterclass**. La lecture complète du tableau de milieu est ici assurée de manière optimale :

Stars			
XY coordinates file		MU.<>12	20 4
X-axis column number (default = 1)			
Y-axis column number (default = 2)			
Categories file (.cat)		Mil.cat	
Variable label file (or #)		Code_Var	
Row weights file (option)		Veg.fcpl	20 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"/>

On retrouve la carte donnée par MV. <>co à condition de bien utiliser le fichier des poids Mil. cpl ou Veg. fcpl.

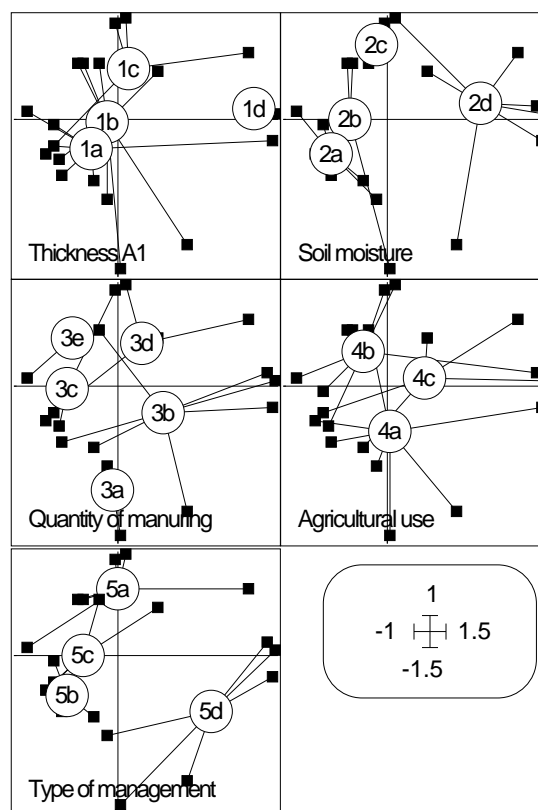


Figure 11 Les chiffres désignent les cinq variables de milieu. Dans chaque fenêtre, les relevés (carrés) sont positionnés par MV. <>1 2 (coordonnées définies par Veg) et les modalités de chacune des variables sont exprimées par une étoile reliant tous les relevés porteurs d'une même modalité au centre de gravité correspondant.

## 6 - Conclusions

On peut donc appeler analyse de co-inertie entre une AFC (relevés-espèces) et une ACM (relevés-milieu) avec toutes ses aides à l'interprétation : analyse des correspondances complète du tableau de profils écologiques (Mercier et al.<sup>11</sup>).

On préférera dans ADE la procédure de co-inertie plutôt que l'analyse du tableau des profils. On possède ainsi l'équivalent strict de la théorie des profils écologiques pour les variables de milieu quantitatives avec le couplage AFC-ACP (complètement compatible avec la théorie des gradients écologiques). Pour des données où la variation d'abondance des taxons est significative (pollution par exemple) on pourra de plus utiliser les couplages ACP-ACP (analyse inter-batterie de Tucker, 1958)<sup>12</sup> ou ACP-ACM où les tests associés sont des analyses de variance.

## Références

- <sup>1</sup> Ter Braak, C.J.F. (1987) The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* : 69, 69-77.
- <sup>2</sup> Godron, M. (1968) Quelques applications de la notion de fréquence en écologie végétale (recouvrement, information mutuelle entre espèces et facteurs écologiques, échantillonnage). *Acta Œcologica, Œcologia Plantarum* : 3, 185-212.
- <sup>3</sup> Gounot, M. (1969) Méthodes d'étude quantitative de la végétation. Masson, Paris. 1-314.
- <sup>4</sup> Daget, Ph. & Godron, M. (1982) Analyse fréquentielle de l'écologie des espèces dans les communautés. Masson, Paris. 1-163.
- <sup>5</sup> Gros, G. & Chessel, D. (1982) Un test exact de comparaison de deux moyennes de variables discrètes : utilisation en toxicologie. *Biométrie-praximétrie* : 12, 3-4, 117-130.
- <sup>6</sup> M'hamdi, A. (1984) Contribution à la description statistique des relations espace-espèces-milieu en phyto-écologie. Thèse de docteur-ingénieur, Université Lyon 1. 1-148.
- <sup>7</sup> Romane, F. (1972) Un exemple d'utilisation de l'analyse factorielle des correspondances en écologie végétale. In : *Grundfagen und Methoden in der Pflanzensoziologie*. Maarel, E. van der & Tüxen, R. (Eds.) Dr. W. Junk b.v., The Hague. pp. 155-162.
- <sup>8</sup> Montana, C. & Greig-Smith, P. (1990) Correspondence analysis of species by environmental variable matrices. *Journal of Vegetation Science* : 1, 453-460.
- <sup>9</sup> Bonin, G. & Roux, M. (1978) Utilisation de l'analyse factorielle des correspondances dans l'étude phyto-écologique de quelques pelouses de l'Apennin lucano-calabrais. *Acta Œcologica, Œcologia Plantarum* : 13, 121-138.
- <sup>10</sup> 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.
- <sup>11</sup> 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.
- <sup>12</sup> Tucker, L.R. . (1958) An inter-battery method of factor analysis. *Psychometrika* : 23, 2, 111-136.