

OccurData

Généralités.....	2
OccurData : Array_to_Occur.....	3
OccurData : CANOCO-Classes.....	6
OccurData : CANOCO-Subspace.....	10
OccurData : CANOCO-Table.....	15
OccurData : Categ_to_Occur.....	18
OccurData : Occur_to_Array.....	20
OccurData : Read_Occur_File.....	21

Généralités

Ce module est construit pour le traitement des listes d'occurrences. Ces données très particulières sont des séries d'observations, largement indépendantes les unes des autres, de la présence d'une espèce dans certaines circonstances. Un exemple caractéristique est fourni par les exemplaires des musées d'histoire naturelle. Une plante dans un herbier ou un poisson dans un bocal indique la présence de l'espèce selon les conditions mentionnées sur l'étiquettes (qui, quand, comment, où, dans quel environnement).

La propriété essentielle de ces données est qu'il n'y a pas de relevés mais des observations de la présence d'un taxon. L'intérêt théorique de ce type de données est considérable dans la mesure où son approche retentit sur l'analyse des données écologiques classiques.

Le point d'entrée est l'option **Read_Occur_File** qui lit un fichier de chaîne de caractères. Sur chaque ligne figure simplement le nom du taxon correspondant à l'occurrence. Utiliser des codes à 50 caractères au plus. Les occurrences peuvent être pondérées. Par défaut, elles sont uniformément pondérées. Les listes d'occurrences sont accessibles par des fichiers du type ---.opl (équivalent du type ---.cat pour les variables qualitatives, ---.gpl pour les graphes de voisinage, ...) La présence des espèces rares est une perturbation forte dans l'analyse de ces données. L'option **Rare_Species** propose de modifier la liste en regroupant les occurrences d'espèces rares dans une catégorie artificielle, ce qui évite de modifier les tableaux associés à la liste d'occurrence.

Trois utilitaires sont d'intérêt mineur. **Occur_to_Array** fabrique un tableau à partir d'une liste d'occurrences classées par une variable qualitative externe. **Categ_to_Occur** fabrique une liste d'occurrences avec une variable qualitative pour l'utilisateur qui a déjà codé numériquement sa liste de taxa. **Array_to_Occur** fabrique une liste d'occurrences à partir d'un tableau classique pour contrôler les relations de ce programme avec d'autres modules d'ADE-4.

Une liste d'occurrences ne s'analyse pas en soit. Elle est croisée avec de l'information externe portant sur l'environnement de l'occurrence. Si cette information externe est une seule variable qualitative, le lien entre la liste et cette variable se fait par **CANOCO-Classes** qui est une analyse des correspondances utilisée comme analyse canonique ¹. Si cette information est un tableau de variables quantitatives, utiliser l'option **CANOCO-Table** qui est une analyse canonique des correspondances ² utilisée comme analyse canonique, ce qui est original. Dans les autres cas, utiliser le module Projectors pour définir des bases orthonormales d'un sous-espace de référence et l'option **CANOCO-Subspace** qui est une analyse canonique entre deux sous-espaces.

Dans la version actuelle, les trois options fournissent simplement les corrélations canoniques et les scores canoniques. Utiliser les fiches thématiques 4.9 (*Quand les données écologiques sont des occurrences d'espèces*) et 4.A (*Quand l'analyse canonique des correspondances est-elle une analyse canonique ?*) pour construire des aides à l'interprétation numériques ou graphiques adaptées.

Les illustrations de la fiche utilise la carte Occurrences contenant une liste de 1653 occurrences d'arbres extraites de l'Atlas de B.R. Ramesh & J.P. Pascal ³.



¹ Thioulouse, J. & Chessel, D. (1992) A method for reciprocal scaling of species tolerance and sample diversity. *Ecology* : 73, 670-680.

² Ter Braak, C.J.F. (1986) Canonical correspondence analysis : a new eigenvector technique for multivariate direct gradient analysis. *Ecology* : 67, 1167-1179.

³ Ramesh, B.R. & Pascal, J.P. (1998) Atlas of endemics of the Western Ghats (India). Institut Français de Pondichery. In press.

OccurData : Array_to_Occur



Utilitaire de transformation de données pour contrôles de la programmation.



La plupart des données écologiques se présentent sous la forme de tableaux. Le présent module traite des données écologiques sous la forme de listes d'occurrences d'espèces (données muséographiques, bibliographiques, ...). Pour vérifier les calculs des options de ce module, Array_to_Occur transforme un tableau floro-faunistique en liste d'occurrences. Son intérêt est donc purement technique.



L'option utilise une seule fenêtre de dialogue :

Array_to_Occur

Input file (array) DouPoi 30 27

Label file Poi_Label

Output file name LP01

Quit Ok

Nom du fichier binaire d'entrée. Ce fichier contient un tableau floro-faunistique avec relevés en lignes et taxa en colonnes.

Nom du fichier des étiquettes des espèces. Ce fichier texte contient autant de chaînes de caractères séparées par des retour-charriot qu'il y a de colonnes dans le tableau précédent.

Nom générique des fichiers de sortie (créations multiples).



Utiliser les cartes de données Doubs et Doubs+1. Le dialogue ci-dessus fournit :

Sorry: line number 8 is empty

Continue Quit

L'option interdit les lignes ou les colonnes de zéros. Supprimer la ligne 8 dans le tableau faunistique et dans le tableau de milieu :

Row-Col Selection

Input file DouPoi 30 27

Selection of rows (default = all) 1a7;9a30

Selection of columns (default = all)

Output file Poi

Row-Col Selection

Input file DouMil 30 11

Selection of rows (default = all) 1a7;9a30

Selection of columns (default = all)

Output file Mil

Array_to_Occur

Input file (array) Poi 29 27

Label file Poi_Label

Output file name LP01

```
Input file :DouPoi
Rows-samples: 30 Col-species: 27
Label input file :Poi_Label
Rows-labels: 27
Input file :Poi
```

Rows-samples: 29 Col-species: 27
Label input file :Poi_Label
Rows-labels: 27

Label output file (species names):LPOI_Spe.txt
Rows-occurrences: 375

Ce fichier texte n'est pas utile, il est un exemple d'entrée pour l'option
OccurData: Read_Occur_File

Output file (sample number):LPOI_Rel
Rows-occurrences: 375

On a un fichier binaire contenant une seule variable qualitative dont les lignes sont des
occurrences et les classes sont les numéros d'espèces.

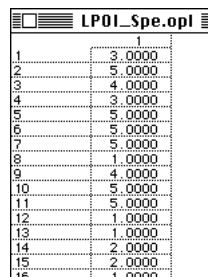
One categorical variable file: LPOI_Spe
Number of rows: 375, variable: 1, categories: 27

Description of categories:

Num	Label	Occur	F*10000
1	Truite	17	453
2	Vairon	20	533
3	Loche	24	640
4	Brochet	18	480
5	Chevaine	25	667
6	Goujon	20	533
7	Perche	15	400
8	Tanche	17	453
9	Vandoise	18	480
10	Rotengle	11	293
11	Gardon	18	480
12	Chabot	8	213
13	Ombre	8	213
14	Blageon	8	213
15	Barbeau	14	373
16	Toxostome	11	293
17	Spiralin	12	320
18	Perche_soleil	13	347
19	Carpe	12	320
20	Hotu	12	320
21	Bouvière	11	293
22	Ablette	14	373
23	Anguille	11	293
24	Grémille	12	320
25	Brème_comm.	9	240
26	Brème_bord.	10	267
27	Poisson_chat	7	187

Auxiliary ASCII output file LPOI_Spe.123: labels for 27 modalities
Occurrence weights in file LPOI_Spe.opl
Uniform weights = 1

Les occurrences sont pondérées par les valeurs du tableau faunistique :



1	3.0000
2	5.0000
3	4.0000
4	3.0000
5	5.0000
6	5.0000
7	5.0000
8	1.0000
9	4.0000
10	5.0000
11	5.0000
12	1.0000
13	1.0000
14	2.0000
15	2.0000
16	1.0000

On retrouve les poids des occurrences (cases non nulles) dans le tableau :

	Poi																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	5	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	5	5	5	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	4	5	5	0	0	0	0	0	1	0	0	1	2	2	0	0	0	0	1	0	0	0	0	0	0	0	0
5	0	2	3	2	0	0	0	0	5	2	0	0	2	4	4	0	0	2	0	3	0	0	0	5	0	0	0	
6	0	3	4	5	0	0	0	1	2	0	0	1	1	1	0	0	0	0	2	0	0	0	1	0	0	0	0	
7	0	5	4	5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
8	0	0	1	3	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	1	0	0	0	4	0	0	0	
9	0	1	4	4	0	0	0	2	2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
10	1	3	4	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Chaque occurrence définit le numéro du relevé comme une autre variable qualitative :

One categorical variable file: LPOI_Rel
 Number of rows: 375, variable: 1, categories: 29

Description of categories:

Num	Label	Occur	F*10000
1	R1	1	27
2	R2	3	80
3	R3	4	107
4	R4	8	213
5	R5	11	293
6	R6	10	267
7	R7	5	133
...			
21	R21	22	587
22	R22	3	80
23	R23	8	213
24	R24	8	213
25	R25	21	560
26	R26	22	587
27	R27	22	587
28	R28	26	693
29	R29	21	560

Auxiliary ASCII output file LPOI_Rel.123: labels for 29 modalities

Les étiquettes des relevés sont constituées par R1, R2, ... R29.



Cette option a essentiellement une valeur pédagogique. Elle permet d'aborder des tableaux comme des listes d'occurrences pour comprendre le rôle et la nature des méthodes dérivées de l'analyse des correspondances (fiche thématique 4.9 et 4.A). Les options OccurData: Occur_to_Array et OccurData: Array_to_Occur sont symétriques.

Si on enchaîne Occur_to_Array puis Array_to_Occur on retrouve le point de départ. Si on enchaîne Array_to_Occur puis Occur_to_Array on retrouve le point de départ à une transposition et une permutation des relevés près.

OccurData : CANOCO-Classes



Analyse canonique d'une liste d'occurrences face à une partition.



L'option étudie le lien entre une liste d'occurrences et une partition.



L'option utilise une seule fenêtre de dialogue :

Nom du fichier d'accès à la liste d'occurrences (type ---.opl).

Nom du fichier d'accès à une variable qualitative (type ---.cat).

Nom générique des fichiers de sortie.



Utiliser la carte AviUrba. Transformer le tableau avifaunistique en tableau d'occurrences :

```
Input file :AUFau
Rows-samples: 51   Col-species: 40
Label input file :AUCode_OisFR
Rows-labels: 40
```

```
Label output file (species names):LIST_Spe.txt
Rows-occurrences: 492
```

```
Output file (sample number):LIST_Rel
Rows-occurrences: 492
One categorical variable file: LIST_Spe
Number of rows: 492, variable: 1, categories: 40
```

Description of categories:

Num	Label	Occur	F*10000
1	Coucou gris	8	163
2	Rougequeue noir	19	386
3	Rossignol philomèle	17	346
...			
35	Busard cendré	3	61
36	Bergeronnette printanière	2	41
37	Rougequeue à front blanc	4	81
38	Hypolaïs polyglote	9	183
39	Chouca des tours	2	41
40	Mouette rieuse	4	81

```
Auxiliary ASCII output file LIST_Spe.123: labels for 40 modalities
Occurrence weigths in file LIST_Spe.opl
Uniform weigths = 1
```

```
One categorical variable file: LIST_Rel
Number of rows: 492, variable: 1, categories: 51
```

Description of categories:

Num	Label	Occur	F*10000
1	R1	11	224

2	R2	12	244
...			
48	R48	10	203
49	R49	4	81
50	R50	11	224
51	R51	9	183

Auxiliary ASCII output file LIST_Rel.123: labels for 51 modalities

Le tableau 51 relevés - 40 espèces est transformé en une liste de 492 occurrences.

CANOCO-Classes	
Occurrence list	LIST_Spe.opl 492 1
Categories	LIST_Rel.cat
Output file name	LL

File LL contains XtdY species-classes table
It has 40 rows (categories) and 51 columns (categories)

Le tableau croisé est le transposé (espèces-lignes, relevés-colonnes) du tableau de départ).

Access for X: file LIST_Spe.cat
Access for Y: file LIST_Rel.cat
Access for D: file LIST_Spe.opl

Correspondence analysis as a canonical analysis
--- Number of rows: 40, columns: 51

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+4.1415E-01	+0.1557	+0.1557	02	+2.7249E-01	+0.1025	+0.2582
03	+2.0078E-01	+0.0755	+0.3337	04	+1.7658E-01	+0.0664	+0.4001
05	+1.4374E-01	+0.0541	+0.4542	06	+1.3354E-01	+0.0502	+0.5044
07	+1.3032E-01	+0.0490	+0.5534	08	+1.1691E-01	+0.0440	+0.5974
09	+1.1053E-01	+0.0416	+0.6389	10	+1.0046E-01	+0.0378	+0.6767
11	+9.2890E-02	+0.0349	+0.7116	12	+8.7307E-02	+0.0328	+0.7445
13	+7.6323E-02	+0.0287	+0.7732	14	+7.2505E-02	+0.0273	+0.8004
15	+7.0162E-02	+0.0264	+0.8268	16	+5.8857E-02	+0.0221	+0.8490
17	+5.0919E-02	+0.0191	+0.8681	18	+4.7711E-02	+0.0179	+0.8861
19	+3.9840E-02	+0.0150	+0.9010	20	+3.6752E-02	+0.0138	+0.9149
21	+3.2753E-02	+0.0123	+0.9272	22	+3.0082E-02	+0.0113	+0.9385
23	+2.5692E-02	+0.0097	+0.9482	24	+2.2164E-02	+0.0083	+0.9565
25	+2.0063E-02	+0.0075	+0.9640	26	+1.8282E-02	+0.0069	+0.9709
27	+1.5914E-02	+0.0060	+0.9769	28	+1.2878E-02	+0.0048	+0.9817
29	+1.1978E-02	+0.0045	+0.9862	30	+8.7436E-03	+0.0033	+0.9895
31	+7.9261E-03	+0.0030	+0.9925	32	+5.7175E-03	+0.0022	+0.9947
33	+3.7970E-03	+0.0014	+0.9961	34	+3.3457E-03	+0.0013	+0.9973
35	+3.1396E-03	+0.0012	+0.9985	36	+1.5571E-03	+0.0006	+0.9991
37	+1.3815E-03	+0.0005	+0.9996	38	+5.6946E-04	+0.0002	+0.9998
39	+4.1851E-04	+0.0002	+1.0000	40	+0.0000E+00	+0.0000	+1.0000

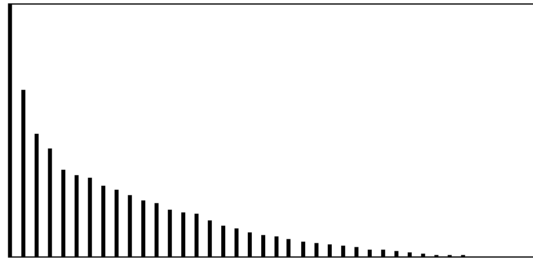
Canonical correlation coefficients

k= 1 rk = 6.4355e-01 rk2 = 4.1415e-01
k= 2 rk = 5.2201e-01 rk2 = 2.7249e-01
k= 3 rk = 4.4809e-01 rk2 = 2.0078e-01
k= 4 rk = 4.2021e-01 rk2 = 1.7658e-01
k= 5 rk = 3.7913e-01 rk2 = 1.4374e-01
k= 6 rk = 3.6544e-01 rk2 = 1.3354e-01

File LL.scor contains the canonical normed scores of the occurrences
It has 492 rows and 6 columns

Exécuter l'AFC du tableau de départ :

COrrespondence Analysis	
Data file	RUFau 51 40



```

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

```

```

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

```

```

-----
Total inertia: 2.65918

```

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+4.1415E-01	+0.1557	+0.1557	02	+2.7249E-01	+0.1025	+0.2582
03	+2.0078E-01	+0.0755	+0.3337	04	+1.7658E-01	+0.0664	+0.4001
05	+1.4374E-01	+0.0541	+0.4542	06	+1.3354E-01	+0.0502	+0.5044
07	+1.3032E-01	+0.0490	+0.5534	08	+1.1691E-01	+0.0440	+0.5974
09	+1.1053E-01	+0.0416	+0.6389	10	+1.0046E-01	+0.0378	+0.6767
11	+9.2890E-02	+0.0349	+0.7116	12	+8.7307E-02	+0.0328	+0.7445
13	+7.6323E-02	+0.0287	+0.7732	14	+7.2505E-02	+0.0273	+0.8004
15	+7.0162E-02	+0.0264	+0.8268	16	+5.8857E-02	+0.0221	+0.8490
17	+5.0919E-02	+0.0191	+0.8681	18	+4.7711E-02	+0.0179	+0.8861
19	+3.9840E-02	+0.0150	+0.9010	20	+3.6752E-02	+0.0138	+0.9149
21	+3.2753E-02	+0.0123	+0.9272	22	+3.0082E-02	+0.0113	+0.9385
23	+2.5692E-02	+0.0097	+0.9482	24	+2.2164E-02	+0.0083	+0.9565
25	+2.0063E-02	+0.0075	+0.9640	26	+1.8282E-02	+0.0069	+0.9709
27	+1.5914E-02	+0.0060	+0.9769	28	+1.2878E-02	+0.0048	+0.9817
29	+1.1978E-02	+0.0045	+0.9862	30	+8.7436E-03	+0.0033	+0.9895
31	+7.9261E-03	+0.0030	+0.9925	32	+5.7175E-03	+0.0022	+0.9947
33	+3.7970E-03	+0.0014	+0.9961	34	+3.3457E-03	+0.0013	+0.9973
35	+3.1396E-03	+0.0012	+0.9985	36	+1.5571E-03	+0.0006	+0.9991
37	+1.3815E-03	+0.0005	+0.9996	38	+5.6946E-04	+0.0002	+0.9998
39	+4.1851E-04	+0.0002	+1.0000	40	+0.0000E+00	+0.0000	+1.0000

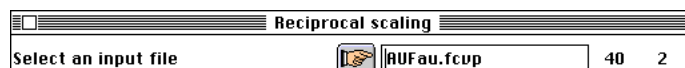
On obtient les mêmes valeurs propres.



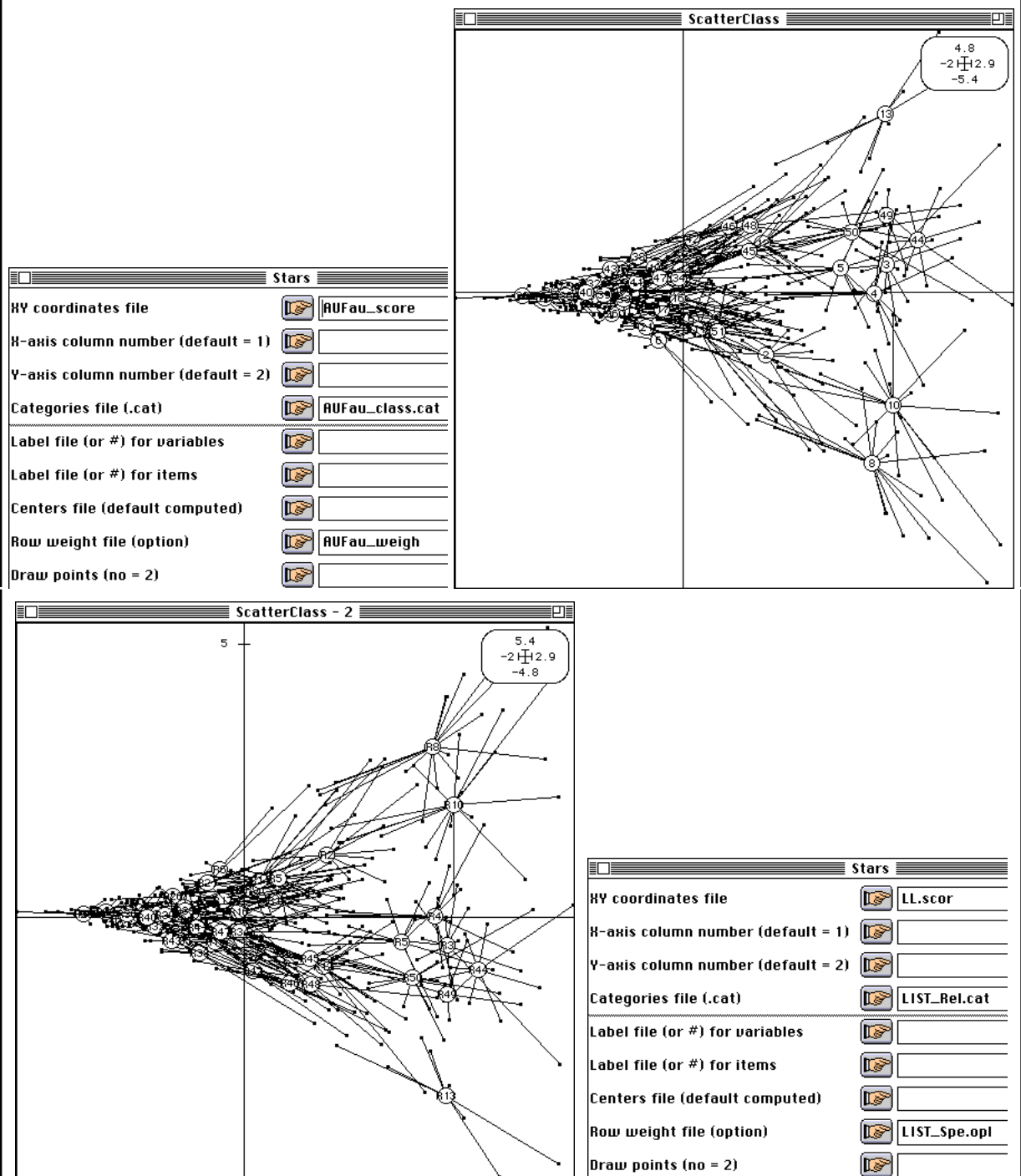
L'AFC du tableau est considérée comme une double analyse d'inertie. Les valeurs propres (inerties projetées) sont affichées pour sélectionner le nombre d'axes. La présente option est une analyse canonique entre les indicatrices des espèces et les indicatrices des relevés. Les valeurs propres sont des carrés de corrélation canonique. On garde systématiquement les 6 premières. C'est le même calcul vu dans deux logiques différentes. Une procédure ne suffit pas à définir ce qu'on pense en dépit de ce que croit un pragmatisme particulièrement réducteur comme dans ¹.



Le fichier LL.scor a le même contenu que le fichier AUFau_score créé par :



Les deux fichiers ne sont cependant pas identiques, car dans l'option COA: Reciprocal scaling les occurrences sont triées par relevés alors que dans la présente option elles sont triées dans l'ordre d'apparition des taxons dans la liste. Le contenu est le même mais pas l'ordre de présentation. Ceci apparaît clairement avec ScatterClass: Stars :



Le signe des coordonnées du second score sont inversés mais cela n'a aucune signification. L'information est la même. **Bien noter qu'utiliser l'AFC sur un tableau faunistique c'est le considérer comme une liste d'occurrences.**



- 1 Palmer, M. (1993) Putting things in even better order: the advantages of canonical correspondence analysis. *Ecology* : 74, 2215-2230.

OccurData : CANOCO-Subspace



Analyse canonique d'une liste d'occurrences face à un sous-espace vectoriel.





L'option étudie le lien entre une liste d'occurrences et un ensemble de variables orthonormées. L'option étend l'analyse canonique des correspondances à des situations très particulières. Par exemple la "Canonical trend surface analysis" ¹ devient dans l'exemple ci-dessous une "Canonical correspondence trend surface analysis"





L'option utilise une seule fenêtre de dialogue :


CANOCO-Subspace


Occurrence list 

Subspace file 

Output file name 

 Nom du fichier d'accès à la liste d'occurrences (type ---.opl).


 Nom du fichier de la base orthonormée (type xxx.@ob). La pondération associée à la liste (---.opl) doit être identique à la pondération associée à la base orthonormée (xxx.@pl)


 Nom générique des fichiers de sortie.




Utiliser la carte Occurrences. Lire la liste :

Read_Occur_File

Input file  List.edit

Option: file for row weight 

Output file name  LL

One categorical variable file: LL
Number of rows: 1653, variable: 1, categories: 249

Description of categories:

Num	Label	Occur	F*10000
1	actiboud	11	67
2	acticaca	9	54
3	acticaob	1	6
4	actimala	14	85
5	actisali	1	6
6	actitadu	1	6
...			
242	syzyzeel	1	6
243	tabegamb	8	48
244	termtrav	5	30
245	tricapio	4	24
246	turpmala	11	67
247	vateindi	16	97
248	veprbilo	1	6
249	verntrav	19	115

Auxiliary ASCII output file LL.123: labels for 249 modalities
Occurrence weights in file LL.opl
Uniform weights = 1

Regrouper les espèces ayant moins de 7 occurrences :

Rare_Species		
Occurrence list	<input type="text" value="LL.opl"/>	1653 1
Rare species level	<input type="text" value="7"/>	
Output file name	<input type="text" value="LL7"/>	

Label output file (species names):LL7.edit
 Rows-occurrences: 1653
 One categorical variable file: LL7
 Number of rows: 1653, variable: 1, categories: 91

Description of categories:

Num	Label	Occur	F*10000
1	actiboud	11	67
2	acticaca	9	54
3	Rare_Species	444	2686
4	actimala	14	85
5	aglaelbo	11	67
...			
85	syzybent	10	60
86	syzylaet	15	91
87	syzymund	24	145
88	tabegamb	8	48
89	turpmala	11	67
90	vateindi	16	97
91	vertrav	19	115

Auxiliary ASCII output file LL7.123: labels for 91 modalities
 Occurrence weights in file LL7.opl
 weights from file LL.opl

Créer une base orthonormée à partir des polynômes des coordonnées avec la nouvelle option Projectors: XY->Orthonormal Basis :

HY->Orthonormal Basis		
HY file	<input type="text" value="HY"/>	1653 2
Option: row weight	<input type="text"/>	
Option: output file name	<input type="text" value="HYpoly"/>	
Maximal degree (default = 2)	<input type="text" value="4"/>	

Orthonormalization: subspace generated by spatial (XY) variables

Explanatory variable file: XY
 It has 1653 rows and 2 columns

Orthonormal basis: XYpoly.@ob
 It has 1653 rows and 14 columns
 Row weight file: XYpoly.@pl
 Uniform row weight = 0.000605

La liste LL7 avec 91 catégories taxonomiques engendre un sous-espace d'indicatrices centrées de dimension 90. La base orthonormée engendre un sous-espace de dimension 14. Ces deux sous-espace sont dans \mathbb{R}^{1653} . L'option cherche les scores canoniques, vecteurs normés bissecteurs des deux sous-espaces :

CANOCO-Subspace		
Occurrence list	<input type="text" value="LL7.opl"/>	1653 1
Subspace file	<input type="text" value="HYpoly.@ob"/>	1653 14
Output file name	<input type="text" value="LL7SPAT"/>	

--- Number of rows: 91, columns: 14

Num.	Eigenval.	R.Iner.	R.Sum	Num.	Eigenval.	R.Iner.	R.Sum
01	+3.6806E-01	+0.2269	+0.2269	02	+2.5653E-01	+0.1582	+0.3851
03	+2.0399E-01	+0.1258	+0.5109	04	+1.8071E-01	+0.1114	+0.6223

05	+1.2716E-01	+0.0784	+0.7007	06	+9.2741E-02	+0.0572	+0.7578
07	+8.6972E-02	+0.0536	+0.8115	08	+7.1249E-02	+0.0439	+0.8554
09	+5.3531E-02	+0.0330	+0.8884	10	+5.1375E-02	+0.0317	+0.9201
11	+4.1119E-02	+0.0254	+0.9454	12	+3.5184E-02	+0.0217	+0.9671
13	+3.0429E-02	+0.0188	+0.9859	14	+2.2912E-02	+0.0141	+1.0000

 Canonical correlation coefficients

k= 1 rk = 6.0668e-01 rk2 = 3.6806e-01

k= 2 rk = 5.0649e-01 rk2 = 2.5653e-01

k= 3 rk = 4.5165e-01 rk2 = 2.0399e-01

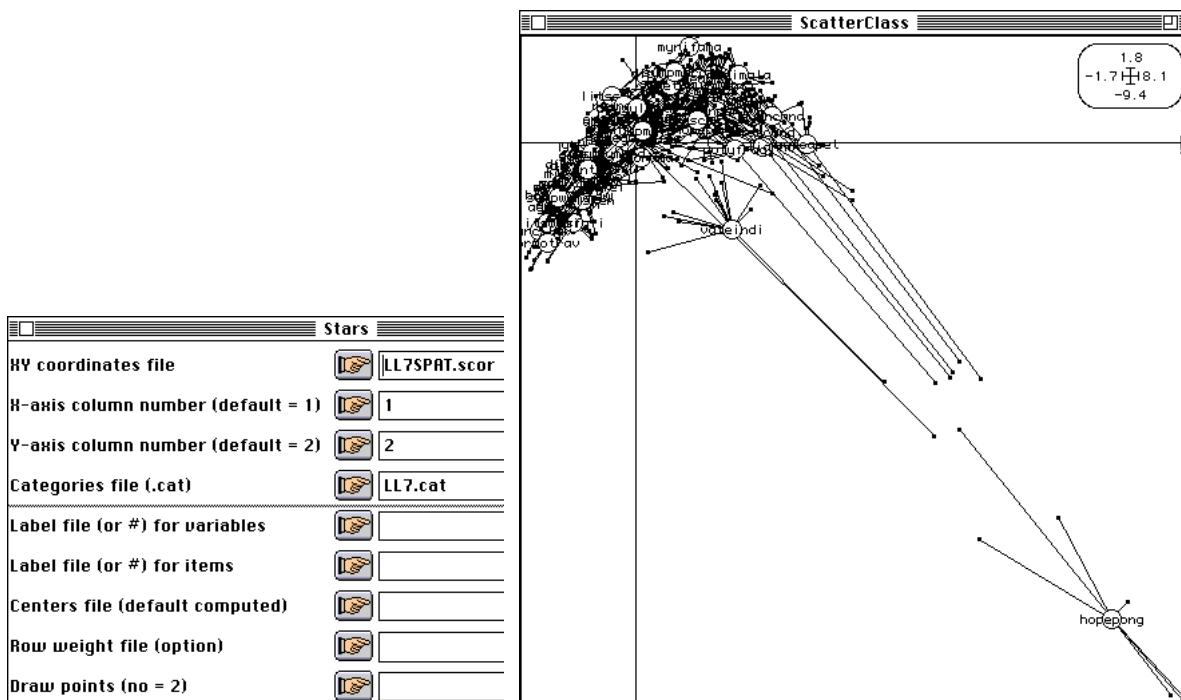
k= 4 rk = 4.2510e-01 rk2 = 1.8071e-01

k= 5 rk = 3.5659e-01 rk2 = 1.2716e-01

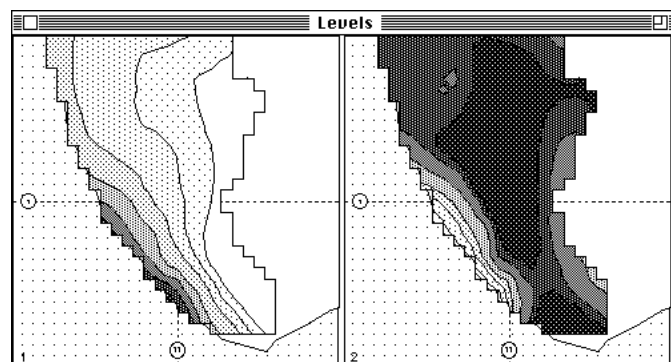
k= 6 rk = 3.0453e-01 rk2 = 9.2741e-02

File LL7SPAT.scor contains the canonical normed scores of the occurrences

It has 1653 rows and 6 columns



Le résultat est, en première lecture, décevant. Les corrélations canoniques sont modestes et les deux premières sont dominées par la position d'un taxon hopepong (n° 42) qui ne présente que 9 occurrences. On isole ses coordonnées :





Dans les stratégies d'analyses canoniques, le contrôle préalable des variables est très important. On a ici l'équivalent des aberrations fréquentes rencontrées en AFC (perturbation par les taxa rares et les relevés légers). Il vaudra préciser un espace sur lequel les occurrences ont une densité relativement régulière pour voir cette opération fonctionner avec satisfaction.



- 1 Wartenberg, D.E. (1985a) Canonical trend surface analysis: a method for describing geographic pattern. *Systematic Zoology* : 34(3), 259-279.

OccurData : CANOCO-Table



Analyse canonique d'une liste d'occurrences face à un tableau de données quantitatives.



L'option étudie le lien entre une liste d'occurrences et un ensemble de variables quantitatives (implicitement soumises à une ACP normée). L'option étend l'analyse canonique des correspondances à une situation qu'on peut considérer comme normale pour cette analyse alors que son emploi sur une liste d'occurrences dérivée d'un tableau doit être reconsidéré au niveau de la signification statistique.



L'option utilise une seule fenêtre de dialogue :

Nom du fichier d'accès à la liste d'occurrences (type ---.opl).

Nom du fichier du tableau de variables.

Nom générique des fichiers de sortie.



Utiliser l'exemple de l'option précédente (OccurData: CANOCO-Subspace). Assembler les quatre variables dans un tableau (FilesUtil: PasteFiles-SameRow) :

```
--- Number of rows: 91, columns: 4
Num. Eigenval.  R.Iner.  R.Sum  | Num. Eigenval.  R.Iner.  R.Sum  |
01  +4.0322E-01  +0.4223  +0.4223  | 02  +2.6364E-01  +0.2761  +0.6984  |
03  +1.6901E-01  +0.1770  +0.8754  | 04  +1.1902E-01  +0.1246  +1.0000  |
```

```
-----
Canonical correlation coefficients
k= 1 rk = 6.3500e-01  rk2 = 4.0322e-01
k= 2 rk = 5.1345e-01  rk2 = 2.6364e-01
k= 3 rk = 4.1110e-01  rk2 = 1.6901e-01
k= 4 rk = 3.4499e-01  rk2 = 1.1902e-01
File LL7MIL.scor contains the canonical normed scores of the occurrences
It has 1653 rows and 4 columns
```

Calculer les corrélations entre scores canoniques et variables initiales (MatAlg: Diagonal Inner product $C=X'DY$) :

Centring

Input file 1653 4

Option: file for row weight

Option for H matrix (no default)

Output file

Values

HV coordinates file

H-axis column number (default = 1)

V-axis column number (default = 2)

G values file

Label file (or #) for variables

Label file (or #) for items

Label file (or #) for groups

Dot if G = 0 (yes = 1)

Constrain H/U ratio (yes = 1)

Scatters

L'interprétation ne pose pas de problèmes particuliers.

Nous avons utilisé CANOCO pour le nom de cette méthode pour ouvrir le débat. Le nom se rapporte à la logique de la méthode car c'est une **analyse canonique** qui se pense comme telle. En fonction de l'origine du tableau ou du sous-espace bien des pratiques potentielles de dépouillement sont à explorer. Cette analyse canonique se fait sans relevé et on voit les occurrences deux fois.

Dans la version habituelle de la CCA (CANOCO classique), l'occurrence est vue trois fois. La première se fait par le nom du taxon, la seconde par le numéro du relevé et la troisième par le milieu du relevé. L'analyse canonique porte sur le lien entre la première position et la troisième. Le sous-espace engendré par les variables de milieu est dans le sous-espace engendré par les indicatrices des relevés. Toute la question est donc de savoir si l'effet du milieu est significatif sachant l'effet du relevé. En l'absence d'un test qui prend en compte cet aspect, on ne peut l'utiliser que comme AFCVI et non comme analyse canonique. Le débat est loin d'être clos.

OccurData : Categ_to_Occur



Utilitaire technique.



L'option est l'inverse de Read_Occur_File. Elle permet, si on dispose d'une variable qualitative à m modalités et d'un fichier d'étiquettes à m lignes de refaire un fichier d'occurrences avec le nouveau code



L'option utilise une seule fenêtre de dialogue :

Nom du fichier d'accès à la variable qualitative. Le fichier ne doit contenir qu'une seule variable qualitative.

Nom du fichier des étiquettes associées aux modalités de la variable qualitative d'entrée (par exemple, nom de famille pour les espèces figurant dans une liste d'occurrences).

Nom du fichier de sortie (création). L'extension .txt est automatiquement ajoutée au nom de ce fichier.



Utiliser la carte Ghat. Exécuter l'opération décrite dans OccurData : Read_Occur_File et lancer la présente option :

```
Input file :Occur4.cat
Lines: 3953  Var: 1  Categories: 269
Label input file :CodeFamille.txt
Lines: 269
```

```
Label output file :OccurFami.txt
Lines: 3953
```

On peut alors relire le nouveau fichier d'occurrences :

```
One categorical variable file: OccurFami
Number of rows: 3953, variable: 1, categories: 41
```

Description of categories:

Num	Label	Occur	F*10000
1	LAURACEAE	467	1181
2	MELIACEAE	164	415
3	OLACACEAE	2	5
4	EUPHORBIACEAE	364	921
5	ARALIACEAE	14	35
6	MYRSINACEAE	6	15
7	MORACEAE	166	420
8	RUTACEAE	42	106
9	CHRYSOBALANACEAE	7	18
10	ARECACEAE	34	86
11	RHIZOPHORACEAE	6	15
12	ANACARDIACEAE	333	842
13	RUBIACEAE	261	660
14	CLUSIACEAE	131	331
15	FLACOURTIACEAE	177	448
16	OLEACEAE	10	25
17	FABACEAE (Caesalpinioideae)	116	293
18	EBENACEAE	280	708
19	DIPTEROCARPACEAE	233	589
20	ELAEOCARPACEAE	59	149
21	MYRTACEAE	159	402
22	CELASTRACEAE	48	121
23	ANNONACEAE	224	567
24	THEACEAE	20	51

25	MYRISTICACEAE	149	377
26	STERCULIACEAE	70	177
27	FABACEAE (Mimosoideae)	2	5
28	SAPOTACEAE	78	197
29	MALVACEAE	3	8
30	CORNACEAE	28	71
31	MELASTOMATACEAE	56	142
32	MAGNOLIACEAE	5	13
33	FABACEAE (Faboideae)	19	48
34	SAPINDACEAE	31	78
35	PITTOSPORACEAE	4	10
36	ERICACEAE	1	3
37	SYMPLOCACEAE	88	223
38	APOCYNACEAE	37	94
39	COMBRETACEAE	7	18
40	STAPHYLEACEAE	28	71
41	ASTERACEAE	24	61

 Auxiliary ASCII output file OccurFami.123: labels for 41 modalities



ATTENTION : Les options du module fonctionnent avec des codes comportant au plus 50 caractères par étiquettes.

OccurData : Occur_to_Array



Utilitaire de manipulation de données.



L'option sert pour croiser une liste d'occurrences et une variables qualitative. Elle donne un tableau avec les espèces en lignes, les classes en colonnes et la somme des poids des occurrences par espèce et par classe dans chaque cellule du tableau.



L'option utilise une seule fenêtre de dialogue :

Nom du fichier des occurrences (---.opl).

Nom du fichier de la variable de tri (---.cat).

Nom du fichier de sortie (création).



Utiliser la carte Doubs et Doubs+1. Exécuter les indications de la fiche de OccurData : Array_to_Occur.

File Provi contains XtDY species-classes table
 It has 27 rows (categories) and 29 columns (categories)
 Access for X: file LPOI_Spe.cat
 Access for Y: file LPOI_Rel.cat
 Access for D: file LPOI_Spe.opl

Vérifier que le résultat est le transposé du tableau Poi :

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

	Poi																											
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	
1	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	5	4	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	5	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	4	5	5	0	0	0	0	0	1	0	0	1	2	2	0	0	0	0	0	1	0	0	0	0	0	0	0
5	0	2	3	2	0	0	0	5	2	0	0	2	4	4	0	2	0	3	0	0	0	0	0	5	0	0	0	0
6	0	3	4	5	0	0	0	1	2	0	0	1	1	1	0	0	0	0	2	0	0	0	1	0	0	0	0	0
7	0	5	4	5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



Les options OccurData: Occur_to_Array et OccurData: Array_to_Occur sont strictement symétriques. Elles s'enchaînent en donnant le même résultat (à une transposition près).

OccurData : Read_Occur_File



Utilitaire de lecture d'un fichier liste d'occurrences (point d'entrée du module).



Une liste d'occurrence est un fichier texte comportant une ligne par occurrence et une étiquette d'espèce par ligne. Elle est lue et transformée en variable qualitative par la présente option.



L'option utilise une seule fenêtre de dialogue :

Read_Occur_File

Input file

Option: file for row weight

Output file name

Nom du fichier d'entrée contenant la liste d'occurrences.

Nom du fichier binaire contenant un poids de chaque occurrence. Par défaut, ce poids vaut 1 (pondération unitaire). Si ce paramètre est utilisé, le programme vérifie que le fichier contient une seule colonne et un nombre de lignes égal au nombre d'étiquettes dans le fichier précédent.

Nom du fichier binaire de sortie (création).



Utiliser la carte de données Ghat :

```
One categorical variable file: Occur4
Number of rows: 3953, variable: 1, categories: 269
```

Description of categories:

Num	Label	Occur	F*10000
1	acbo	36	91
2	accac	12	30
3	accao	2	5
4	acma	35	89
5	acsa	1	3
6	acta	1	3
7	agba	13	33
8	agca	1	3
9	agel	25	63
10	agex	17	43
11	agin	1	3
...			
266	tuma	28	71
267	vain	116	293
268	vebi	10	25
269	vetr	24	61

```
-----
Auxiliary ASCII output file Occur4.123: labels for 269 modalities
Occurrence weights in file Occur4.opl
Uniform weights = 1
-----
```



Les fichiers XXX.opl définissent les objets "liste d'occurrences" dans toutes les options du module.



Le fichier XXX.123 contient les étiquettes trouvées dans la liste.



ATTENTION : Les options de ce module fonctionnent avec des codes comportant au plus 50 caractères par étiquettes.