

Inertie intra-classes et mesures de diversité

Résumé

Dans 25 stations on a mesuré 11 paramètres environnementaux en 10 points aléatoires. La variabilité stationnelle définit la diversité mésologique. Dans ces 25 stations on a capturé 1008 poissons de 23 espèces. Ceci permet de définir la diversité taxonomique. Ces poissons portent 3 mesures morphométriques qui définissent la diversité stationnelle de la taille des poissons d'une part et de la forme des poissons d'autre part. Une méthode générale de calcul de ces différentes composantes de la diversité est proposée. Les relations entre ces composantes est analysées. La diversité de la forme est seule reliée à la diversité mésologique de manière complexe.

Plan

1 — Données traitées.....	2
2 — Diversité faunistique.....	7
3 — Diversité morphométrique.....	10
4 — Diversité mésologique.....	12
4 — Relations entre type de diversité	14
5 — Explicitation.....	17

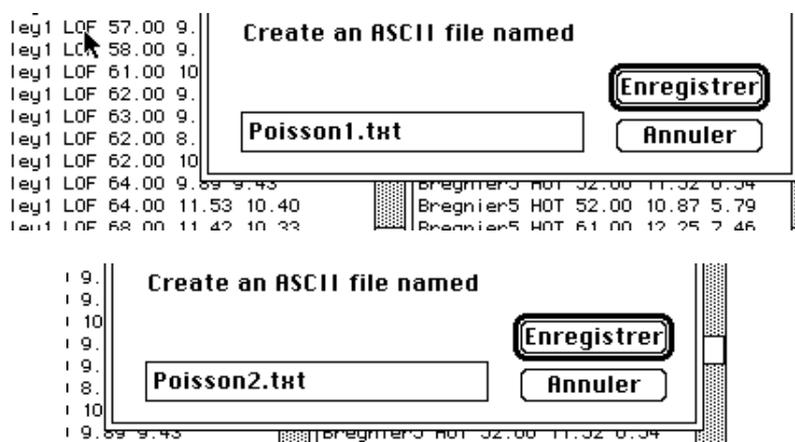
L. Hill, B. Statzner, J.M. Olivier & D. Chessel

1 — Données traitées

Elles sont stockées dans les cartes Diversité et Diversité+1 de la pile ADE-4•Data. La première contient les mesures biologiques.

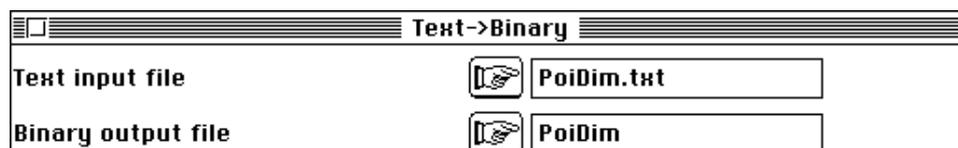
Data 147		Louize Hill (1995)	
Belley1	BAF	52.00	9.99 7.00
Belley1	LOF	40.00	5.98 5.27
Belley1	LOF	42.00	6.89 6.20
Belley1	LOF	42.00	7.75 6.68
Belley1	LOF	48.00	7.54 6.99
Belley1	LOF	49.00	7.73 7.58
Belley1	LOF	48.00	7.26 5.19
Belley1	LOF	51.00	7.46 6.17
Belley1	LOF	61.00	8.67 7.72
Belley1	LOF	52.00	7.83 5.90
Belley1	LOF	52.00	7.35 5.98
Belley1	LOF	55.00	10.07 8.06
Belley1	LOF	57.00	8.98 8.10
Belley1	LOF	57.00	9.19 8.40
Belley1	LOF	58.00	9.54 8.80
Belley1	LOF	61.00	10.96 10.00
Belley1	LOF	62.00	9.19 8.40
Belley1	LOF	63.00	9.31 9.80
Belley1	LOF	62.00	8.36 7.00
Belley1	LOF	62.00	10.88 10.71
Belley1	LOF	64.00	9.89 9.43
Belley1	LOF	64.00	11.53 10.40
Belley1	LOF	68.00	11.42 10.33
Belley1	ORB	132.00	30.00 18.54
Bregnier5	GOU	68.00	12.06 9.66
Bregnier5	GOU	71.00	12.99 10.32
Bregnier5	GOU	71.00	12.88 10.56
Bregnier5	GOU	72.00	13.00 10.45
Bregnier5	GOU	74.00	15.31 12.33
Bregnier5	GOU	73.00	16.10 12.45
Bregnier5	GOU	73.00	16.10 12.45
Bregnier5	HOT	39.00	8.58 4.82
Bregnier5	HOT	38.00	7.42 4.60
Bregnier5	HOT	41.00	8.36 5.10
Bregnier5	HOT	41.00	8.38 5.10
Bregnier5	HOT	44.00	8.54 5.74
Bregnier5	HOT	40.00	9.72 6.26
Bregnier5	HOT	40.00	9.43 5.71
Bregnier5	HOT	40.00	10.49 6.14
Bregnier5	HOT	40.00	9.65 5.48
Bregnier5	HOT	40.00	9.92 6.61
Bregnier5	HOT	40.00	11.33 6.19
Bregnier5	HOT	40.00	11.26 6.58
Bregnier5	HOT	52.00	10.87 6.82
Bregnier5	HOT	52.00	11.52 6.54
Bregnier5	HOT	52.00	10.87 5.79
Bregnier5	HOT	61.00	12.25 7.46
Bregnier5	HOT	63.00	15.50 9.22

Reconstituer un fichier Excel™ en collant bout à bout le contenu des deux champs.



Sélectionner la première colonne et la mettre dans un fichier PoiSta.txt, sélectionner la seconde colonne et la mettre dans un fichier PoiEsp.txt, sélectionner les trois dernières colonnes et les mettre dans un fichier PoiDim.txt. Sauvegarder en Texte et ajouter un retour chariot à la fin de chacun des trois fichiers avec l'éditeur de texte.

Passer en binaire les trois fichiers. Le premier ne pose pas de problème :



PoiDim est un fichier binaire à 1008 lignes et 3 colonnes. Les deux autres utilisent l'option nouvelle :

LabelToCateg	
Label input file	 PoiSta.txt
Output file name	 PoiSta

```

[ 1]Category:  1 Num:   26 Freq.:  0.02579  Belley1
[ 2]Category:  2 Num:  154 Freq.:   0.1528  Belley2
[ 3]Category:  3 Num:   7  Freq.: 0.006944  Belley3
[ 4]Category:  4 Num:  19  Freq.: 0.01885   Belley4
[ 5]Category:  5 Num:  18  Freq.: 0.01786   Belley5
[ 6]Category:  6 Num:   5  Freq.: 0.00496   Belley6
[ 7]Category:  7 Num:  77  Freq.: 0.07639   Belley7
[ 8]Category:  8 Num:  27  Freq.: 0.02679   Belley8
[ 9]Category:  9 Num:   3  Freq.: 0.002976   Belley9
[10]Category: 10 Num:  46  Freq.: 0.04563   Chautagne1
[11]Category: 11 Num:  62  Freq.: 0.06151   Chautagne2
[12]Category: 12 Num:   9  Freq.: 0.008929   Chautagne3
[13]Category: 13 Num:  10  Freq.: 0.009921   Chautagne4
[14]Category: 14 Num:  16  Freq.: 0.01587   Chautagne5
[15]Category: 15 Num: 106  Freq.:  0.1052   Chautagne6
[16]Category: 16 Num:  20  Freq.: 0.01984   Chautagne7
[17]Category: 17 Num:  24  Freq.: 0.02381   Chautagne8
[18]Category: 18 Num:  44  Freq.: 0.04365   Bregnier1
[19]Category: 19 Num:  43  Freq.: 0.04266   Bregnier2
[20]Category: 20 Num:  11  Freq.: 0.01091   Bregnier3
[21]Category: 21 Num:  40  Freq.: 0.03968   Bregnier4
[22]Category: 22 Num: 138  Freq.:  0.1369   Bregnier5
[23]Category: 23 Num:  45  Freq.: 0.04464   Bregnier6
[24]Category: 24 Num:  40  Freq.: 0.03968   Bregnier7
[25]Category: 25 Num:  18  Freq.: 0.01786   Bregnier8

```

LabelToCateg	
Label input file	 PoiEsp.txt
Output file name	 PoiEsp

```

[ 1]Category:  1 Num:   77 Freq.:  0.07639  BAF
[ 2]Category:  2 Num:  164 Freq.:  0.1627  LOF
[ 3]Category:  3 Num:   6  Freq.: 0.005952  OBR
[ 4]Category:  4 Num:  33  Freq.: 0.03274  TRF
[ 5]Category:  5 Num:  45  Freq.: 0.04464  BOU
[ 6]Category:  6 Num:   1  Freq.: 0.0009921 BRO
[ 7]Category:  7 Num: 177  Freq.:  0.1756  CHE
[ 8]Category:  8 Num: 200  Freq.:  0.1984  GOU
[ 9]Category:  9 Num:  13  Freq.:  0.0129  GRE
[10]Category: 10 Num:  80  Freq.: 0.07937  HOT
[11]Category: 11 Num:  42  Freq.: 0.04167  PER
[12]Category: 12 Num:  13  Freq.:  0.0129  SPI
[13]Category: 13 Num:  75  Freq.:  0.0744  VAN
[14]Category: 14 Num:  17  Freq.: 0.01687  VAI
[15]Category: 15 Num:   6  Freq.: 0.005952  TAN
[16]Category: 16 Num:  32  Freq.: 0.03175  ABL
[17]Category: 17 Num:   7  Freq.: 0.006944  GAR
[18]Category: 18 Num:   7  Freq.: 0.006944  PES
[19]Category: 19 Num:   2  Freq.: 0.001984  BLA
[20]Category: 20 Num:   6  Freq.: 0.005952  EPI
[21]Category: 21 Num:   3  Freq.: 0.002976  BRB
[22]Category: 22 Num:   1  Freq.: 0.0009921 BRE
[23]Category: 23 Num:   1  Freq.: 0.0009921 LOT

```

On obtient directement deux fichiers à une seule variable qualitative, avec la distribution marginale et le fichier des étiquettes (--.123) correspondant.

Pour obtenir le tableau faunistique :

CrossingCateg	
First ---.cat type file (Table H)	<input type="text" value="PoiSta.cat"/>
Second ---.cat type file (table Y)	<input type="text" value="PoiEsp.cat"/>
Weighting file (Default = none)	<input type="text"/>
Output file (XtDY)	<input type="text" value="SP"/>

File SPBB contains crossed Burt's matrix X'DY from coding matrix X and coding matrix Y.
 It has 25 rows (categories) and 23 columns (categories)
 Access for X: file PoiSta.cat
 Access for Y: file PoiEsp.cat

 Khi2 tests on Burt's crossed table SPBB

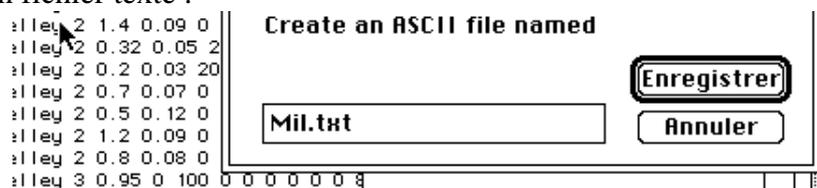
Var 1 versus var 1:Khi2 = 2.6417e+03 ddl = 528 Proba = 0.000000

	BAF	LOF	OBR	TRF	BOU	BRO	CHE	GOU	GRE	HOT	PER	SPI	VAN	VAI	TAN	ABL	GAR	PES	BLA	EPI	BRB	BRE	LOT
Belley1	1	22	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Belley2	5	2	0	3	2	1	46	6	6	27	14	2	40	0	0	0	0	0	0	0	0	0	0
Belley3	1	0	0	2	0	0	1	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0
Belley4	4	0	0	1	1	0	3	0	1	0	8	0	0	1	0	0	0	0	0	0	0	0	0
Belley5	1	15	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Belley6	0	1	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Belley7	2	0	0	0	6	0	20	12	0	13	0	0	2	4	2	14	1	1	0	0	0	0	0
Belley8	4	0	0	21	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
Belley9	0	0	0	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Chautagne1	1	40	0	0	0	0	0	4	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Chautagne2	9	39	0	0	0	0	10	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chautagne3	0	3	0	0	0	0	2	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chautagne4	1	2	0	0	0	0	0	4	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
Chautagne5	1	2	0	0	0	0	4	6	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0
Chautagne6	1	7	0	0	0	0	14	49	3	3	10	0	14	5	0	0	0	0	0	0	0	0	0
Chautagne7	1	15	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Chautagne8	3	11	4	0	0	0	2	0	0	2	1	1	0	0	0	0	0	0	0	0	0	0	0
Bregnier1	18	0	0	0	0	0	7	0	0	0	2	3	0	0	0	14	0	0	0	0	0	0	0
Bregnier2	19	0	0	1	0	0	4	14	1	0	0	4	0	0	0	0	0	0	0	0	0	0	0
Bregnier3	0	0	0	1	0	0	5	2	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0
Bregnier4	1	1	0	0	1	0	22	4	0	4	0	0	0	2	0	0	0	2	2	1	0	0	0
Bregnier5	1	4	0	0	27	0	12	53	0	18	0	0	13	0	0	0	4	1	0	3	2	0	0
Bregnier6	0	0	0	0	5	0	10	10	0	11	4	0	0	0	1	0	1	1	0	2	0	0	0
Bregnier7	0	0	0	0	3	0	8	18	0	1	0	0	5	0	0	0	1	1	0	0	1	1	1
Bregnier8	3	0	0	0	0	0	2	6	0	0	0	2	0	1	0	4	0	0	0	0	0	0	0

La deuxième carte contient les données morphométriques :

Data	148	Louize Hill (1995)									
Belley 1	0.4	0	0	0	15	85	0	8	10	0	30
Belley 1	0.39	4	0	0	10	90	0	8	0	0	0
Belley 1	0.38	0.78	0	0	10	90	0	8	0	0	0
Belley 1	0.39	0.81	0	0	5	95	0	12	0	0	0
Belley 1	0.3	1.11	0	0	5	95	0	15	0	0	0
Belley 1	0.34	1.16	0	0	5	95	0	10	0	0	0
Belley 1	0.36	1.13	0	0	5	95	0	10	0	0	0
Belley 1	0.1	0.38	0	0	5	75	20	40	0	0	0
Belley 1	0.18	0.42	0	0	5	95	0	13	0	0	0
Belley 1	0.35	0.65	0	0	5	95	0	15	0	0	0
Belley 2	0.15	0.52	0	0	40	60	0	9	5	0	10
Belley 2	0.32	0.48	0	0	10	60	30	45	20	0	8
Belley 2	0.3	0.04	100	0	0	0	0	0	0	0	70
Belley 2	1.4	0.09	0	0	0	100	90	0	0	0	0
Belley 2	0.32	0.05	20	0	20	40	20	60	0	50	
Belley 2	0.2	0.03	20	0	80	0	0	0	0	0	30
Belley 2	0.7	0.07	0	0	80	0	20	60	0	0	90
Belley 2	0.5	0.12	0	0	0	100	80	0	0	50	
Belley 2	1.2	0.09	0	0	0	100	105	0	0	40	

Obtenir un fichier texte :



L'ouvrir avec Excel™ et extraire la première colonne dans MilSta.txt . Conserver le reste. Rajouter des retours-chariots. Passer en binaire :



Mil a 250 lignes et 11 colonnes.



MilSta a 250 lignes et une colonne qualitative avec 25 modalités représentées 10 fois (10 points par station). Tester l'hétérogénéité inter-stationnelle :



variable 1 from Mil versus variable 1 from MilSta

Source	SS	d.f.	MS	F	Proba
Between	10.71	24	0.4463	7.867	0
Within	12.76	225	0.05673		
Total	23.47	249			

variable 2 from Mil versus variable 1 from MilSta

Source	SS	d.f.	MS	F	Proba

Between	3421	24	142.5	8.086	0
Within	3966	225	17.62		
Total	7386	249			

variable 3 from Mil versus variable 1 from MilSta

Source	SS	d.f.	MS	F	Proba
Between	2.701E+05	24	1.125E+04	20.04	0
Within	1.264E+05	225	561.6		
Total	3.965E+05	249			

...

variable 11 from Mil versus variable 1 from MilSta

Source	SS	d.f.	MS	F	Proba
Between	1.262E+04	24	526	5.309	0
Within	2.229E+04	225	99.07		
Total	3.491E+04	249			

variable 12 from Mil versus variable 1 from MilSta

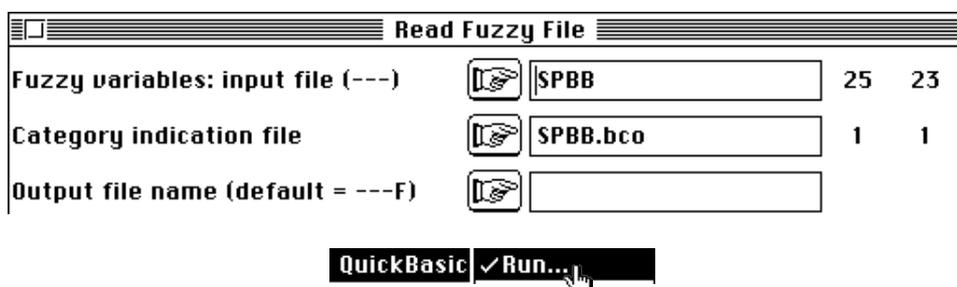
Source	SS	d.f.	MS	F	Proba
Between	5.881E+04	24	2450	10.56	0
Within	5.219E+04	225	232		
Total	1.11E+05	249			

Toutes les probabilités sont nulles. Faune et milieu varient fortement d'une station à l'autre. Il serait aisé de coupler les deux tableaux mais ce n'est pas du tout l'objectif. La question est : **peut-on dire que la variabilité des poissons d'une station (variabilité d'appartenance spécifique, variabilité de taille et de forme) est significativement liée à la variabilité des conditions météorologiques locales à l'intérieur d'une station (variabilité de la profondeur, de la vitesse, de la granulométrie, ...)** ?

Pour tester cette hypothèse nous disposons de 25 observations.

2 — Diversité faunistique

On mesure habituellement la variabilité des organismes par la diversité taxonomique. Utilisons l'indice de Simpson. Le tableau SPBB a 25 lignes (stations) et 23 colonnes (espèces). On peut le considérer comme un tableau à une seule variable floue avec 23 modalités. Le fichier SPBB.bco contient un seul nombre (23) qui convient pour :



```

FuzDivIndex
FuzFCMean
FuzRowMean
FuzRowScor
FuzSimpson
FuzzyScore1
FuzzyScore2
FuzzyTable
GNDEN
Histo1Categ

```

Simpson diversity index for fuzzy data

.fuz type file (Txt) **SPBBF.fuz**

Simpson diversity index for fuzzy data
 Input file : SPBBF

File SPBBF*simtot is binary
 Number of rows = 25 Number of columns = 1
 It contains Simpson Diversity Indices

La diversité spécifique mesurée par l'indice de Simpson ($1 - \sum_{i=1}^p \frac{2}{i}$, si i est l'abondance relative du taxon i dans le relevé et p le nombre total de taxons) est dans le fichier SPBBF*simtot. Faisons alors une expérience simple. Les 1008 poissons sont rangés par espèces dans PoiEsp (une seule variable qualitative). On obtient le tableau disjonctif complet associé par :

Categ -> Disj
 ---.cat type file **PoiEsp.cat**

PoiEsp01 a 1008 lignes et 23 colonnes (les indicatrices des noms d'espèces). Dans ce tableau, à la ligne k , le poisson k , s'il appartient à l'espèce j , donne des valeurs nulles sur toute les colonnes, à l'exception de la colonne k où on trouve 1. L'ACP centrée de ce tableau est exécutée :

Covariance matrix PCA

Matrix input file **PoiEsp01** 1008 23

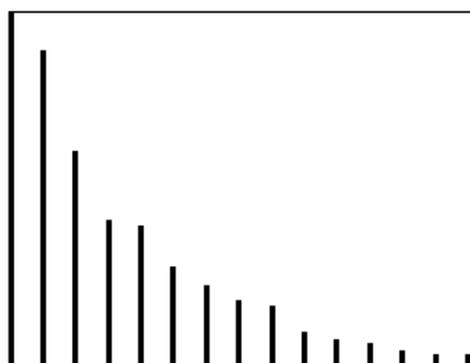
Row weights (default=1/n)

Column weights (default=1)

Option: file for row weighting

Option: file for column weighting

Les valeurs propres sont sans intérêt car toutes les variables sont non covariantes :



L'analyse est sans intérêt (il y a un facteur par espèce), mais on peut lancer l'association de cette analyse avec la variable qui trie les poissons par stations ;

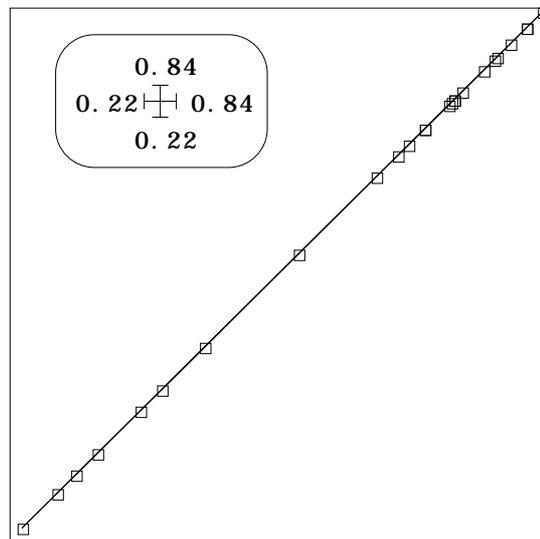
Initialize: LinkPrep	
Statistical triplet	<input type="text" value="PoiEsp01.cpta"/> 1008 23
Categories file (.cat)	<input type="text" value="PoiSta.cat"/>
Selected variable (default=1)	<input type="text"/>
Output file name	<input type="text" value="A"/>

Exécuter alors le nouveau module :

Within Parameters	
--.dis input file	<input type="text" value="A.dis"/>

Il calcule dans le fichier A.whdiv les inerties de chaque classe c'est à dire la somme des variances des indicatrices des espèces par stations. On trouve exactement l'indice de Simpson :

Lines	
X file (default = 1, 2, 3, ..., n)	<input type="text" value="SPBBF*simtot"/> 25 1
X file column number (default = 1)	<input type="text"/>
Y file (no default)	<input type="text" value="A.whdiv"/> 25 1



L'indice de Simpson est donc un cas particulier de l'inertie intra-classes pour les variables indicatrices des noms d'espèces et l'ACP centrée (ce qui est facile à démontrer). On peut détruire tous les fichiers associés à cette remarque.

On retiendra qu'on a un procédé universel de calcul de la diversité pour tout type d'analyse et tout type de variables sous la forme de l'inertie intra-classe (module Within Parameters de Discrimin).

On a donc une alternative à l'indice de Simpson si on utilise l'ACM à la place de l'ACP. L'analyse elle-même n'a aucun intérêt, comme dans le cas précédent, mais pas pour la même raison. Ici toutes les valeurs propres sont égales :

Multiple Correspondence Analysis

.cat type file

Option: Row weighting file

Eigenvalues

Num.	Eigenval.	%Iner.	R.Sum	Num.	Eigenval.	%Iner.	R.Sum
01	+1.0000E+00	+0.0455	+0.0455	02	+1.0000E+00	+0.0455	+0.0909
03	+1.0000E+00	+0.0455	+0.1364	04	+1.0000E+00	+0.0455	+0.1818
05	+1.0000E+00	+0.0455	+0.2273	06	+1.0000E+00	+0.0455	+0.2727
07	+1.0000E+00	+0.0455	+0.3182	08	+1.0000E+00	+0.0455	+0.3636
09	+1.0000E+00	+0.0455	+0.4091	10	+1.0000E+00	+0.0455	+0.4545
11	+1.0000E+00	+0.0455	+0.5000	12	+1.0000E+00	+0.0455	+0.5455
13	+1.0000E+00	+0.0455	+0.5909	14	+1.0000E+00	+0.0455	+0.6364
15	+1.0000E+00	+0.0455	+0.6818	16	+1.0000E+00	+0.0455	+0.7273
17	+1.0000E+00	+0.0455	+0.7727	18	+1.0000E+00	+0.0455	+0.8182
19	+1.0000E+00	+0.0455	+0.8636	20	+1.0000E+00	+0.0455	+0.9091
21	+1.0000E+00	+0.0455	+0.9545	22	+1.0000E+00	+0.0455	+1.0000
23	+0.0000E+00	+0.0000	+1.0000				

Ceci permet cependant de calculer les inertie intra-classes :

Initialize: LinkPrep

Statistical triplet 1008 23

Categories file (.cat)

Selected variable (default=1)

Output file name

Within Parameters

--.dis input file

Le fichier Provi.whdiv contient le nouvel indice de diversité. Sa valeur analytique est:

$$I_{ACM} = \frac{p}{i=1} \frac{i(1-i)}{p_i}$$

où i est l'abondance relative du taxon i dans le relevé, p le nombre total de taxons et p_i est l'abondance relative du taxon i dans l'ensemble de tous les relevés. Il faudra discuter de la signification de ce nouvel indice de diversité. Les résultats obtenus sont stockés dans un fichier Excel que nous appellerons Div. Notons que l'analyse inter-classe associée à l'indice de Simpson est l'ACP centrée du tableau faunistique alors que la même analyse associée au nouvel indice est l'AFC du tableau faunistique. La question mérite donc d'être approfondie.

3 — Diversité morphométrique

D'un poisson à l'autre le nom d'espèce change. La variabilité dans une station est la diversité faunistique. D'un poisson à l'autre les mesures morphométriques changent. Nous utilisons la même notion d'inertie intra-classes pour calculer la diversité morphométrique. Deux analyses sont possibles. La plus simple est l'ACP normée :

Correlation matrix PCA	
Matrix input file	PoiDim 1008 3

La corrélation entre les trois mesures est écrasante (voir la fiche Analyse discriminante) :

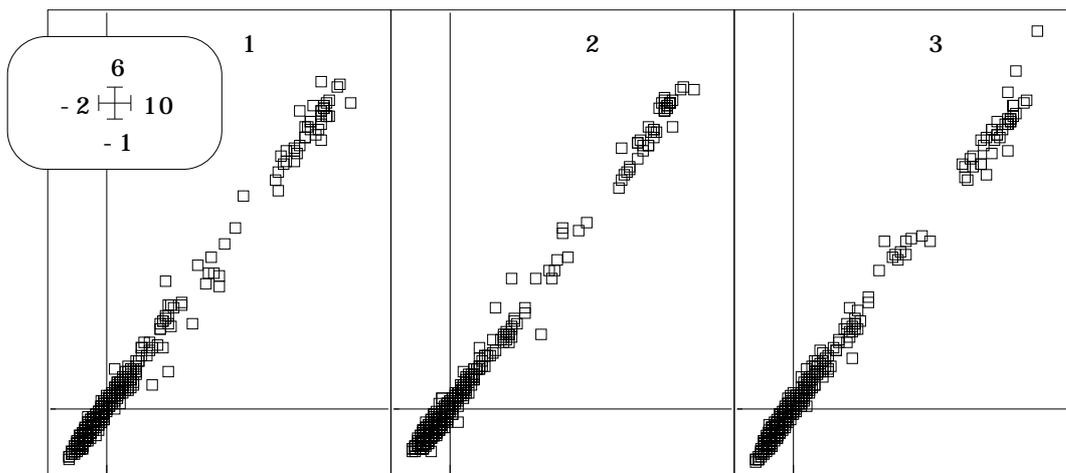
```

Correlation matrix -----
[ 1] 1000
[ 2]  990 1000
[ 3]  989  989 1000
-----

```

Il n'y a qu'un facteur qui regroupe 99.3% de la variabilité. On a un graphe canonique simpliste :

Lines	
X file (default = 1, 2, 3, ..., n)	PoiDim.cnli 1008 1
X file column number (default = 1)	
Y file (no default)	PoiDim.cnta 1008 3



En utilisant cette analyse on calculera essentiellement une variabilité de la taille.

Initialize: LinkPrep	
Statistical triplet	PoiDim.cnta 1008 3
Categories file (.cat)	PoiSta.cat
Selected variable (default=1)	
Output file name	B

Within Parameters	
--.dis input file	B.dis

Le fichier B.whdiv contient un indice de diversité de la taille des poissons qui est ajouté au fichier Div. Ce qui nous intéresse essentiellement est la forme. Pour éliminer l'effet taille nous utilisons une ACP non centrée :

Le résultat est sans nuance :

```
input file: PoiDim
Option 1 = No action (non centred PCA)      Inertia = 12963.9
Option 2 = Centred table (overall centred PCA) Inertia = 9001.65
Option 3 = Centred (zero mean) columns      Inertia = 6190.06
Option 4 = Centred (zero mean) rows        Inertia = 5169.82
Option 5 = Additive model                   Inertia = 2358.23
Option 6 = Multiplicative model             Inertia = 18.3429
```

Il reste normalement deux dimensions :

Num.	Eigenval.	R. Iner.	R. Sum	Num.	Eigenval.	R. Iner.	R. Sum
01	+1.6206E+01	+0.8835	+0.8835	02	+2.1371E+00	+0.1165	+1.0000
03	+0.0000E+00	+0.0000	+1.0000				

Nos pouvons calculer alors la diversité de la forme par le même procédé :

Les valeurs contenues dans C.whdiv sont un indice de diversité de la forme et sont gardées dans Div.

4 — Diversité mésologique

Le tableau de variables de milieu contient des unités hétérogènes et relève de l'ACP normée :

Correlation matrix PCA			
Matrix input file	<input type="button" value="👉"/>	Mil	250 11
Row weights (default=1/n)	<input type="button" value="👉"/>		
Column weights (default=1)	<input type="button" value="👉"/>		
Option: file for row weighting	<input type="button" value="👉"/>		
Option: file for column weighting	<input type="button" value="👉"/>		
1 = Save correlation matrix	<input type="button" value="👉"/>	1	

La structure de corrélation n'est pas caricaturale, comme dans le cas précédent. La redondance entre les variables est globalement faible. On peut regarder la carte des variables :

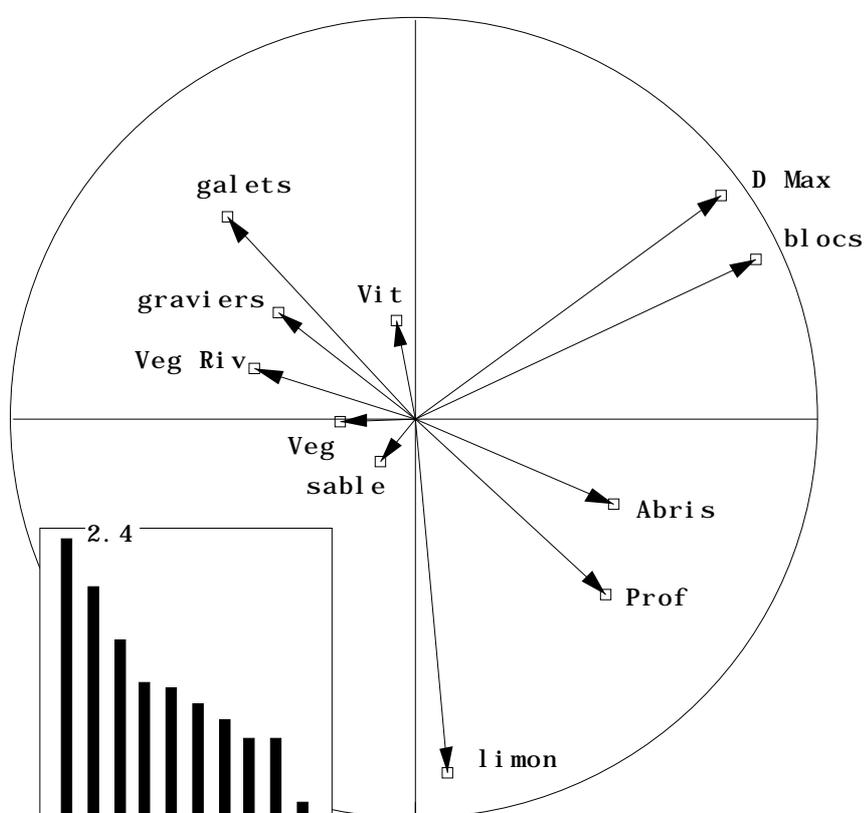


Figure 1 : Carte 1-2 des colonnes de l'ACP normée de Mil et graphe des valeurs propres.

On reconnaît (Figure 1) la redondance D Max (diamètre maximum) et blocs et l'opposition galets limon-profondeur. Mais rien n'indique de fortes corrélations entre variables. La typologie de points (Figure 2) est d'abord une typologie inter-stationnelle. Mais cet aspect est mineur du point de vue de la diversité. On utilise la même procédure :

Initialize: LinkPrep			
Statistical triplet	<input type="button" value="👉"/>	Mil.cnta	250 11
Categories file (.cat)	<input type="button" value="👉"/>	MilSta.cat	
Selected variable (default=1)	<input type="button" value="👉"/>		
Output file name	<input type="button" value="👉"/>	W	



L'inertie totale du nuage est de 11, l'inertie intra-classe est de 6.36 et la diversité intra-stationnelle du milieu est dans le fichier W.whdiv. Ces valeurs sont ajoutées au tableau Div. On se souviendra que le fichier W.whvar contient les composantes de cette diversité, c'est-à-dire les variances stationnelles de chaque variable et que W.whmoy contient les valeurs des moyennes par classes et par variables.

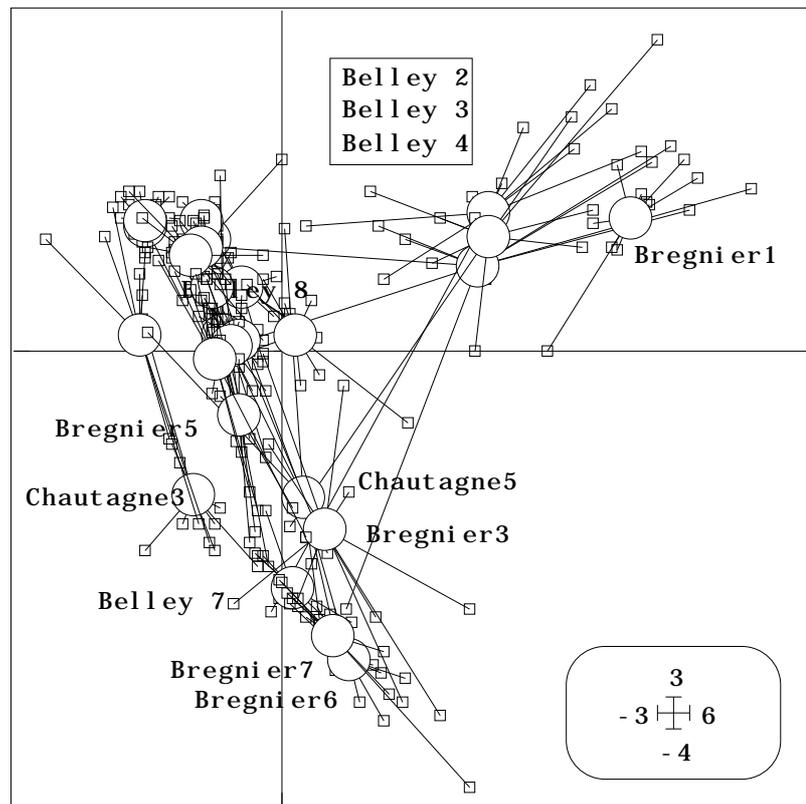


Figure 2 : Carte 1-2 des lignes de l'ACP normée de Mil et représentation des stations. L'analyse simple a essentiellement des onctions d'analyse inter-classe.

4 — Relations entre type de diversité

Les résultats des calculs de diversité sont édités dans le tableau 1. On teste la relation entre la cinquième colonne (diversité mésologique) et les quatre autres (diversité biologique).

On note une relation faible et non significative entre diversité mésologique et indice de Simpson. De même, aucune liaison avec l'indice associé à l'ACM qui présente en plus une valeur aberrante. Ce dernier est sans doute très sensible aux espèces rares. La relation diversité faunistique - diversité mésologique est donc sans signification statistique (Figure 3).

La relation entre diversité mésologique et diversité de la taille est assurément sans signification. La distribution de cette diversité entre quelques valeurs élevées et la plupart des autres quasiment nulles est de plus très défavorable. La seule des quatre mesures (Figure 4) à être liée à la variable explicative est la quatrième (diversité de la forme).

En conclusion, la diversité mésologique est un prédicteur essentiellement de la variabilité morphologique des poissons. Aucune manipulation préliminaire n'ayant précédé les calculs explicités ici, le seuil de signification de 1 pour cent est garanti. La diversité morphométrique de la forme des poissons croit avec la diversité du milieu. La liaison mérite d'être examinée en détail.

	Simpson	IACM	Taille	Forme	Milieu
Belley1	0.275	14.343	0.309	2.450	0.689
Belley2	0.799	19.532	0.378	8.856	11.437
Belley3	0.816	22.613	1.295	13.693	7.796
Belley4	0.742	18.246	0.660	17.875	7.084
Belley5	0.296	11.958	0.022	1.876	1.665
Belley6	0.320	1.895	7.705	33.347	7.881
Belley7	0.836	21.355	13.747	32.292	1.793
Belley8	0.370	9.900	1.089	5.077	3.349
Belley9	0.444	9.588	13.902	36.116	6.878
Chautagne1	0.235	2.637	0.028	1.724	0.677
Chautagne2	0.553	4.133	0.038	2.932	3.426
Chautagne3	0.642	3.595	0.044	2.953	7.303
Chautagne4	0.700	15.823	0.050	1.634	7.320
Chautagne5	0.766	19.482	7.769	45.717	11.909
Chautagne6	0.734	11.142	0.202	4.534	2.647
Chautagne7	0.395	2.581	0.015	1.784	0.173
Chautagne8	0.729	31.743	5.969	12.480	2.260
Bregnier1	0.699	16.728	1.676	7.175	0.818
Bregnier2	0.680	13.813	0.295	2.631	3.615
Bregnier3	0.727	32.455	7.184	57.404	11.611
Bregnier4	0.668	41.703	0.201	4.728	1.877
Bregnier5	0.778	21.473	0.122	9.142	2.865
Bregnier6	0.818	25.382	10.699	54.648	3.031
Bregnier7	0.733	69.840	1.231	14.738	3.078
Bregnier8	0.784	19.714	0.104	3.459	4.828

Tableau 1 : Cinq mesures de diversité stationnelle, respectivement la diversité faunistique vue par l'indice associé à l'ACP (Simpson) et à l'ACM, la diversité morphométrique associée à la taille des poissons (ACP normée) et à la forme des poissons (ACP à double centrage multiplicatif) et diversité mésologique (ACP normée).

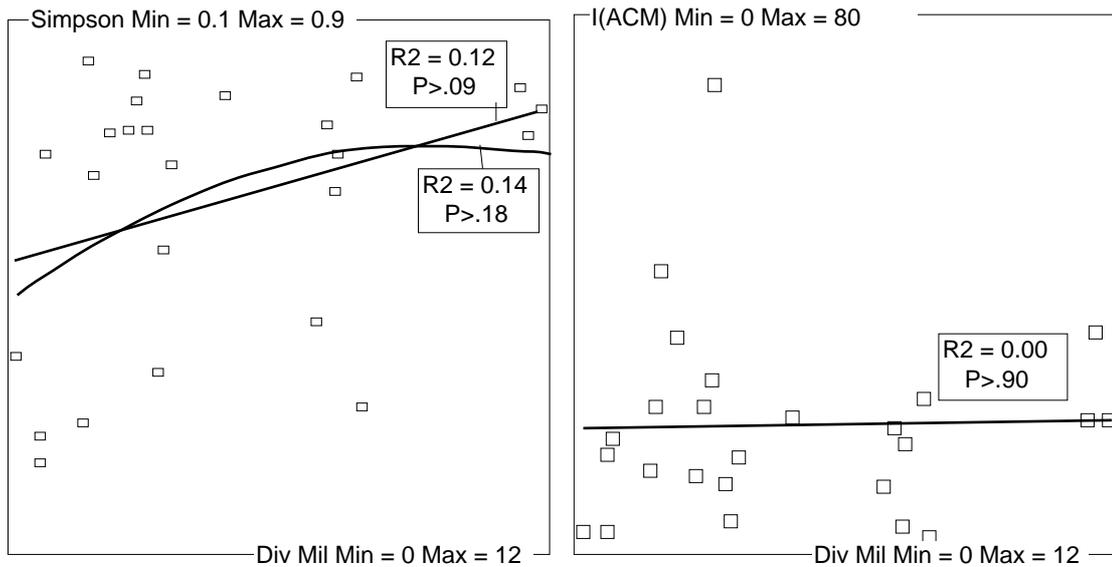


Figure 3 : relations entre diversité faunistique et mésologique.

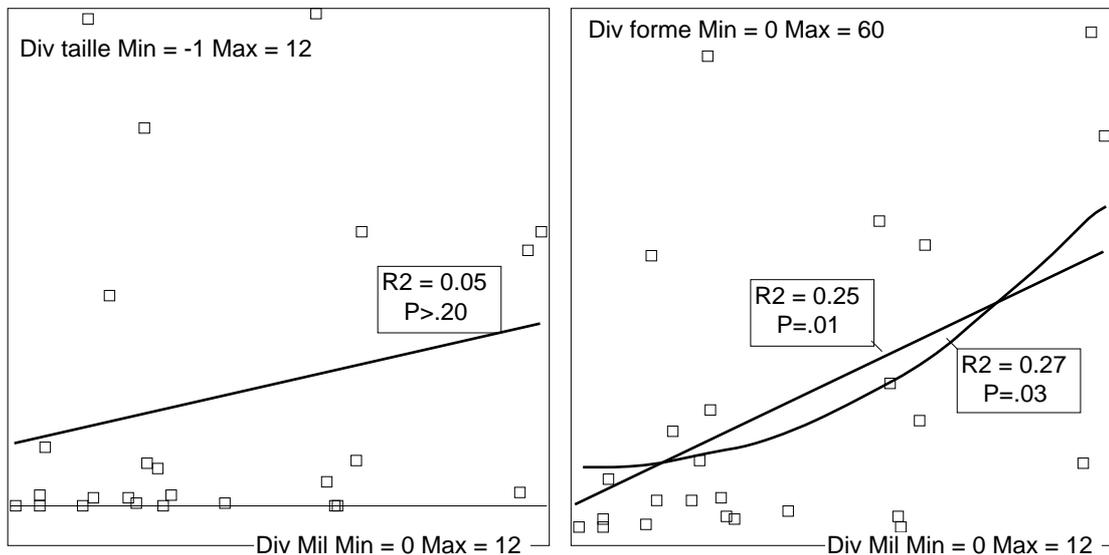


Figure 4 : relations entre diversités morphométrique et mésologique.

5 — Explicitation

On peut se demander alors quelles variables ont une diversité qui ont le plus grand rôle sur la diversité de la forme des poissons. Extraire la diversité de la forme (colonne 4 de Div) dans un fichier Y. 25 points pour 11 variables explicatives impose d'utiliser la régression PLS (voir fiche associée). Dans LinearReg :

Initialize			
Explanatory variables	<input type="button" value="👉"/>	<input type="text" value="W.whvar"/>	25 12
Dependent variables	<input type="button" value="👉"/>	<input type="text" value="Y"/>	25 1
Option: row weighting	<input type="button" value="👉"/>	<input type="text"/>	
Output file name	<input type="button" value="👉"/>	<input type="text" value="Z"/>	

New TEXT file Z.reg contains the parameters:
 ----> Explanatory variables: W.whvar [25][12]
 ----> Dependant variable file: Y [25][1]
 ----> Row weighting file: Uniform weighting

PLS -> Randomization Test	
Input file	 Z.reg
Number of permutations	 10000

PLS1 - Permutation test

 Explanatory variable file: W.whvar
 It has 25 rows and 12 columns

Step	Nrepet	X>Xobs	Frequencie
1	10000	57	5.700e-03
2	10000	2857	2.857e-01
3	10000	9952	9.952e-01
4	10000	9994	9.994e-01

Le seuil observé de signification de la régression PLS est de 5.7 pour mille avec un seul niveau. Le test de permutation confirme une relation indiscutable.

PLS -> Modelling	
Input file	 Z.reg
Number of components (no default)	 1

PLS1 - Modelling

 Explanatory variable file: W.whvar
 It has 25 rows and 11 columns

Dependent variable file: Y
 It has 25 rows and 1 columns

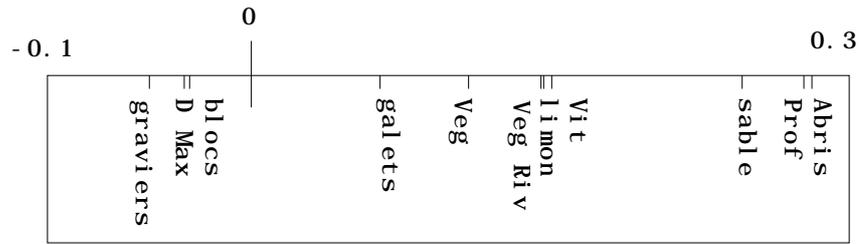
Step	Variance	Col: 1 Explained	Ratio	Exp. Sum
1	1.000e+00	5.685e-01	5.685e-01	5.685e-01

File Z.PLSw1 has 11 rows and 1 columns
 It contains coefficients

File : Z.PLSw1

Col.	Mini	Maxi
1	-5.027e-02	2.817e-01

Le poids des variables n'est pas homogène du tout. Il apparaît donc que prédire la diversité de la forme par la somme des variances intra-classes (inertie intra) n'est pas optimum. Trois variables jouent un rôle important (Abris, profondeur et sable).



File Z.PLSmod has 25 rows and 1 columns
It contains components models

File : Z.PLSmod

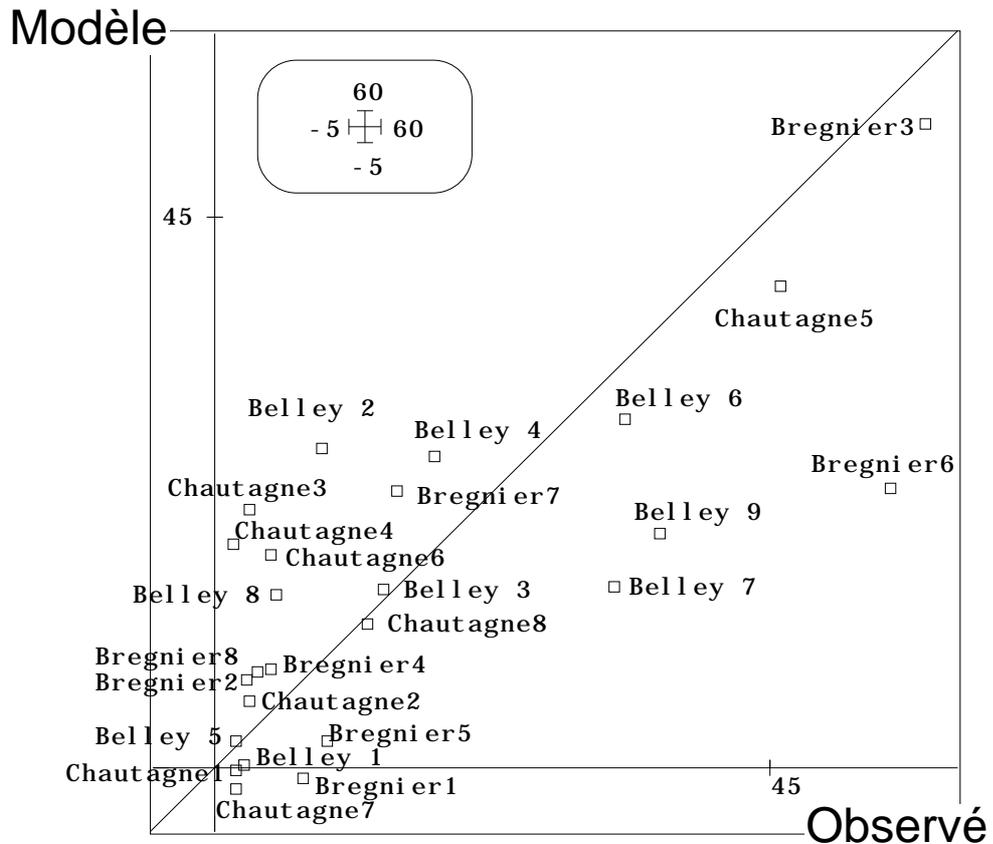
Col.	Mini	Maxi
1	-1.565e+00	5.245e+01

File Z.PLSres has 25 rows and 1 columns
It contains residuals

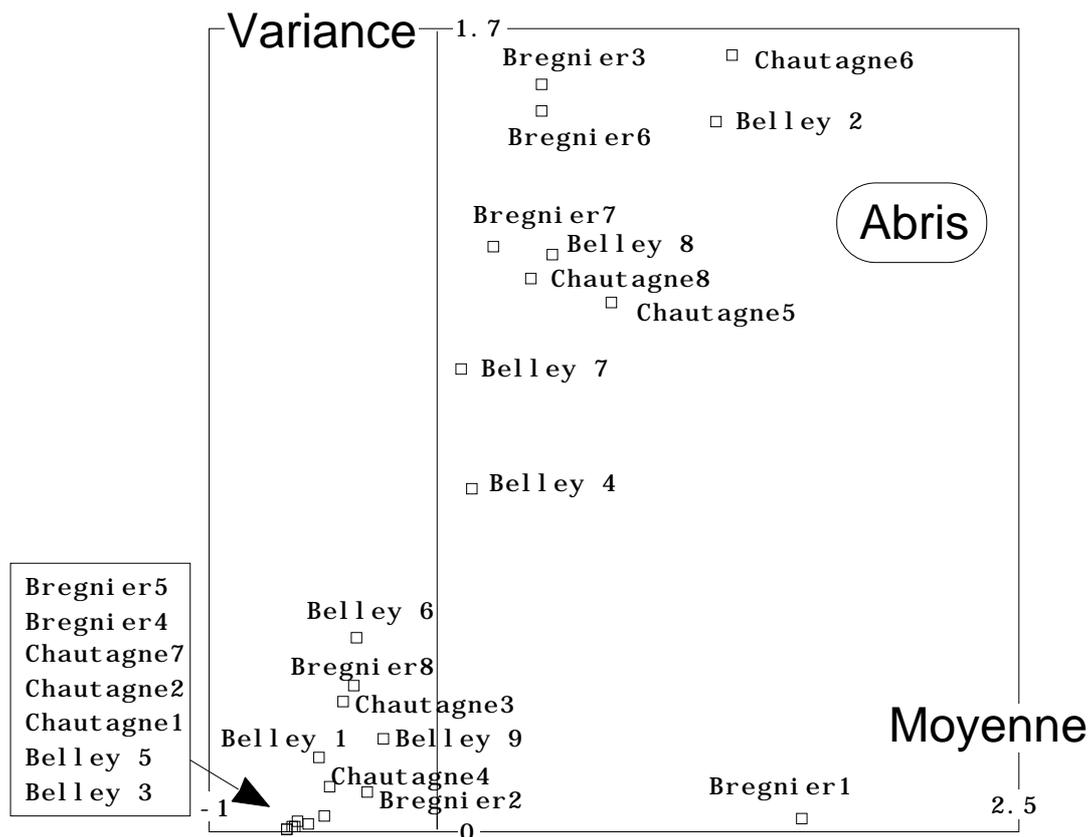
File : Z.PLSres

Col.	Mini	Maxi
1	1.456e+00	5.522e+01

On peut représenter les observations et le modèle estimé :

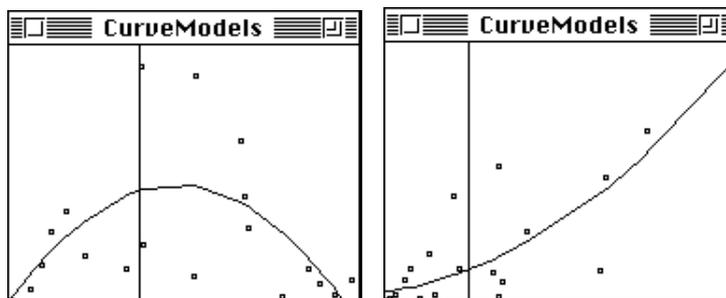


L'autocorrélation des résidus dans la partie droite du graphique montre qu'on n'a pas un modèle optimal. En outre on sait qu'on doit s'attendre à de fortes liaisons moyenne/variance :



La quantité d'abris est déjà par elle-même une mesure de la variabilité du milieu. Il faut noter également que la relation moyenne-variance est complexe :

Polynomials			
X file (default = 1, 2, 3, ..., n)	<input type="button" value="..."/>	<input type="text" value="W.whmoy"/>	25 11
X file column number (default = 1)	<input type="button" value="..."/>	<input type="text" value="6"/>	
Y file (no default)	<input type="button" value="..."/>	<input type="text" value="W.whvar"/>	25 11
Order of polynomial (default = 1) ?	<input type="button" value="..."/>	<input type="text" value="2"/>	



A gauche la relation moyenne (abscisse) contre variance (ordonnée) pour la variable blocs. Dans les milieux artificiels, quand on atteint 100 %, il est clair que la diversité de l'habitat redevient nulle. On a la courbe typique de la valeur indicatrice d'un taxon qui est nulle pour une espèce rare, optimale pour une espèce présente une fois sur deux et à nouveau nulle pour une espèce très fréquente. La diversité est maximum pour les valeurs intermédiaires. A droite la même relation pour la variable Galets montre la relation standard. Pour la variable abris (ci-dessus), la moyenne est une mesure d'hétérogénéité locale de l'habitat et la variance est une mesure d'hétérogénéité globale de la station. La station Brégnier1 montre 50% d'abris (valeur élevée) uniformément en

chaque point (variance très faible). La variance de la variable abris est donc une mesure de variabilité de l'hétérogénéité.

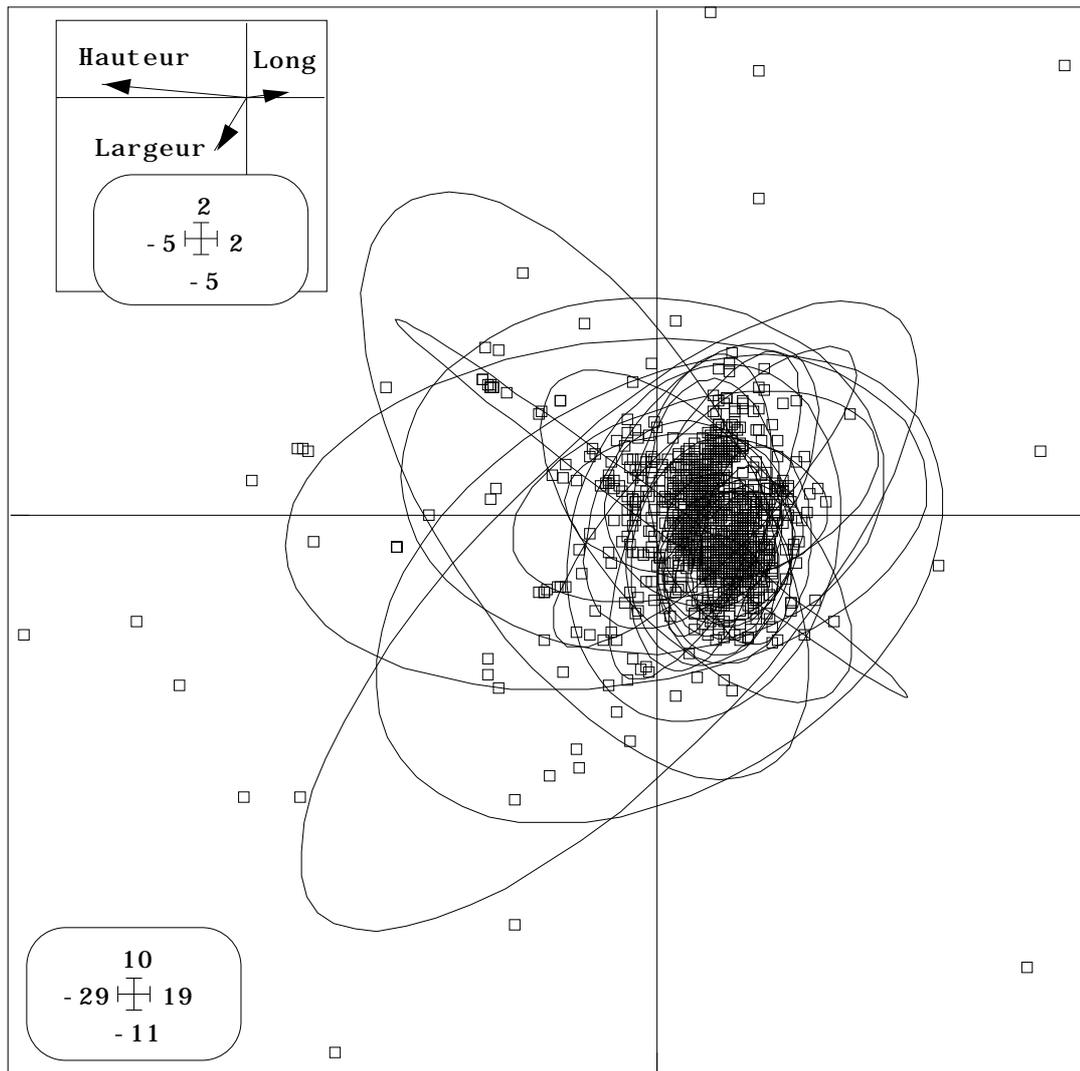


Figure 5 : Plan 1-2 de l'ACP à double centrage multiplicatif. Le nuage est dans le plan et toute la variabilité est représentée. Carte des variables associées. Représentation des stations par des ellipses d'inertie. C'est la variabilité de la taille de ces ellipses qui caractérise ce plan, donc la diversité de la forme.

Si on exécute une régression PLS sur le lien diversité de la forme et moyennes des variables de milieu on a encore une relation significative (2 pour mille sur 10000 simulations). Mais les variables ne jouent pas le même rôle :

