

Associations d'espèces

Pierre Legendre
Département de sciences biologiques
Université de Montréal

Université des Antilles et de la Guyane, Pointe-à-Pitre, Guadeloupe, 21 octobre 2008

Introduction

- Problème classique en écologie (Whittaker 1962) –

« Interspecific associations arise when two or more species co-occur either more or less frequently than expected due to chance alone. **Positive associations** between two species can occur when both select the same habitat or have the same environmental requirements. Conversely, **negative associations** can occur if the species have differing ecological requirements (Dale 1977).

Association, in either the positive or negative direction, can also occur as a direct consequence of **biotic interactions** such as mutualism, competition and predation. Although it is not possible to unambiguously infer the action of specific processes from the examination of patterns alone (Schluter 1984, Rejmanek and Lepš 1996), association analyses remain a valuable tool for ecologists. Such analyses can be used for generating hypotheses about the factors responsible for the patterns, and hence can be used for identifying particular patterns that may be worthy of further study.

A number of statistical tests have been utilized for detecting species associations. These include correlation analysis (Greig-Smith 1983, O'Connor and Aarssen 1987, Myster and Pickett 1992), analysis by contingency table (Greig-Smith 1983, Dale et al. 1991), variance tests (Schluter 1984, McCulloch 1985), and the use of cross-variograms (Rossi et al. 1992). [...] » (Roxburgh and Chesson 1998).

Roxburgh, S. H., and P. Chesson. 1998. A new method for detecting species associations with spatially autocorrelated data. *Ecology* 79: 2180-2192.

Whittaker, R. R. 1962. Classification of natural communities. *The Botanical Review* 28: 1-239.

Recherche d'associations statistiques d'espèces

Une association au sens statistique est un groupe récurrent d'espèces co-occurentes (données de présence-absence) ou corrélées (données d'abondance).

Méthode : recherche de la concordance entre les espèces (Legendre 2005)

1. Analyse de la corrélation entre les espèces : diviser les espèces en groupes d'espèces corrélées (plusieurs méthodes possibles).
2. Calculer un test global de concordance entre les espèces membres de chaque groupe. Y a-t-il des espèces concordantes au sein de ce groupe ?
3. Tester la contribution de chaque espèce à la concordance de son groupe.

Legendre, P. 2005. Species associations: the Kendall coefficient of concordance revisited. *Journal of Agricultural, Biological, and Environmental Statistics* 10: 226-245.

Petit exemple en R

```
# Script, analyse des données d'acariens du sol
# (70 carottes de sol × 35 espèces ; Borcard & Legendre 1994)
# [Exemple utilisé dans l'article de Legendre 2005]

library(vegan)

data(mite)

mite.hel <- decostand(mite, "hel") # Transformation de Hellinger

# Sous-matrice : 10 sites, 4 espèces (pas toutes corrélées)

mite.small <- mite.hel[c(4,9,14,22,31,34,45,53,61,69),
c(13,14,15,23)]
```

Table 1. Illustrative example. Upper panel: Hellinger-transformed abundances of four mite species at 10 sites selected along the long axis of Figure 5. The Hellinger transformation was computed for the full dataset (70 sites). Lower panel: the same data transformed into ranks (with ties); last column: sum of the ranks for each site.

	<i>Hellinger-transformed abundances</i>				
	<i>Species 13</i>	<i>Species 14</i>	<i>Species 15</i>	<i>Species 23</i>	
Site 4	.25087	.40538	.24380	.08362	
Site 9	.40324	.25503	.39303	.00000	
Site 14	.26577	.47620	.27267	.06097	
Site 22	.32350	.63337	.47003	.00000	
Site 31	.26312	.29089	.39223	.08771	
Site 34	.33675	.44836	.53727	.10153	
Site 45	.07956	.19487	.19487	.11251	
Site 53	.00000	.18570	.26261	.11744	
Site 61	.00000	.15430	.15430	.00000	
Site 69	.12769	.62987	.27584	.34578	
	<i>Ranks (species-wise)</i>				
	<i>Species 13</i>	<i>Species 14</i>	<i>Species 15</i>	<i>Species 23</i>	<i>Sum of ranks R_i</i>
Site 4	5	6	3	5	19.0
Site 9	10	4	8	2	24.0
Site 14	7	8	5	4	24.0
Site 22	8	10	9	2	29.0
Site 31	6	5	7	6	24.0
Site 34	9	7	10	7	33.0
Site 45	3	3	2	8	16.0
Site 53	1.5	2	4	9	16.5
Site 61	1.5	1	1	2	5.5
Site 69	4	9	6	10	29.0

```

out1 <- kendall.global(mite.small, nperm=999)

$Concordance_analysis
  Group.1
W          0.44160305
F          2.37252221
Prob.F     0.04403791
Chi2       15.89770992
Prob.perm  0.04410000

# Tests a posteriori

out2 <- kendall.post(mite.small, mult="holm", nperm=999)

$A_posteriori_tests
          Spec13   Spec14   Spec15   Spec23
Spearman.mean 0.3265678 0.3965503 0.4570402 -0.1681251
W.per.species 0.4949258 0.5474127 0.5927802  0.1239061
Prob          0.0880000 0.0290000 0.0060000  0.7170000
Corrected prob 0.1760000 0.0870000 0.0240000  0.7170000

```

Table 2. Results of (a) the overall and (b) the *a posteriori* tests of concordance among the mite species. P = permutational probability, based upon 9,999 random permutations. (c) Complementary Spearman correlation coefficients (r) with results of one-tailed tests of significance, and partial concordance statistics \bar{r}_j and W_j for each species j described in Section 5. * Reject H_0 at $\alpha = .05$.

(a) Overall test of the W statistic. H_0 : The four species are not concordant with one another							
Kendall's W =							
Friedman's chi-square =		.44160	15.89771	$P = .0448^*$			Reject H_0
(b) A posteriori tests		H_0 : This species is not concordant with the other three					
Species 13		$P = .0766$					Do not reject H_0
Species 14		$P = .0240^*$					Reject H_0
Species 15		$P = .0051^*$					Reject H_0
Species 23		$P = .7070$					Do not reject H_0
(c) Spearman correlation table		$H_0: r = 0; H_1: r > 0$ (one-tailed test)					
		Species 13	Species 14	Species 15	Species 23	\bar{r}_j	W_j
Species 13	r	1.0000	.5593	.8389	-.4185	.32657	.49493
	P	—	.0464	.0012	.8856		
Species 14	r	.5593	1.0000	.6242	.0061	.39655	.54741
	P	.0464	—	.0269	.4933		
Species 15	r	.8389	.6242	1.0000	-.0920	.45704	.59278
	P	.0012	.0269	—	.5998		
Species 23	r	-.4185	.0061	-.0920	1.0000	-.16813	.12391
	P	.8856	.4933	.5998	—		

Détails de la méthode

1. Analyse de la corrélation entre les espèces : diviser les espèces en groupes d'espèces corrélées. Plusieurs méthodes possibles.

- Analyse en composantes principales : peut distinguer 2 groupes.
- Groupement agglomératif basé sur la matrice des corrélations (Pearson, Spearman) entre les espèces : la méthode la plus courante.
- Partitionnement K -means des espèces centrées-réduites.

2. Calculer un test global de concordance entre les espèces de chaque groupe. Y a-t-il des espèces concordantes au sein de ce groupe ?

Calcul du coefficient de concordance de Kendall (W) :

	Espèce 1 (rangs des abondances)	Espèce 2	Espèce 3	Somme des rangs de chaque ligne
Site 1	1	1	6	8
Site 2	6	5	3	14
Site 3	3	6	2	11
Site 4	2	4	5	11
Site 5	5	2	4	11
Site 6	4	3	1	8

W est la variance des sommes de rangs divisée par la plus grande valeur que peut prendre cette somme. Cette valeur maximum serait obtenue s'il y avait accord total entre les juges (espèces) quant au classement des objets (sites).

$$W = \frac{12 \sum_{i=1}^n (R_i - \bar{R})^2}{p^2 (n^3 - n)}$$

W peut aussi être calculé à partir de la moyenne \bar{r} des coefficients de corrélation de Spearman entre toutes les espèces du groupe :

$$W = \frac{(p-1)\bar{r} + 1}{p}$$

On transforme W en une statistique khi-carré avant le test.

$$\chi^2 = p(n-1)W$$

Test paramétrique (déconseillé) ou test par permutations. On laisse tomber les groupes qui ne sont pas significativement concordants.

3. Tester la contribution de chaque espèce à la concordance de son groupe.

- Calculer \bar{r}_j , la moyenne des corrélations de Spearman entre l'espèce j et toutes les autres espèces membres du même groupe.
- On peut en dériver une statistique W_j mesurant la contribution de l'espèce j à la concordance de son groupe :

$$W_j = ((p_{\text{gr}} - 1)\bar{r}_j + 1) / p_{\text{gr}}$$

- Test par permutations : on ne permute que les rangs de l'espèce j dans le tableau avant de calculer les valeurs de ces statistiques sous permutation. Les statistiques \bar{r}_j ou W_j sont équivalentes pour un test par permutations. On peut donc utiliser l'une ou l'autre.

⇒ Si certaines espèces ont des valeurs de \bar{r}_j négatives, cela indique que ces espèces sont négativement corrélées à plusieurs des espèces du groupe. Le groupe est donc trop englobant. Il faut alors diviser ce groupe à l'aide du dendrogramme et reprendre l'analyse.

=====

Simulations numériques (article de 2005)

- L'erreur de type I est correcte : lorsque les espèces ne sont pas corrélées, le test rejette H_0 dans une proportion des cas égale au seuil de signification (e.g. 5%).
- Lorsqu'il y a des espèces corrélées, la puissance du test augmente avec le nombre d'espèces intercorrélées incluses dans l'analyse.

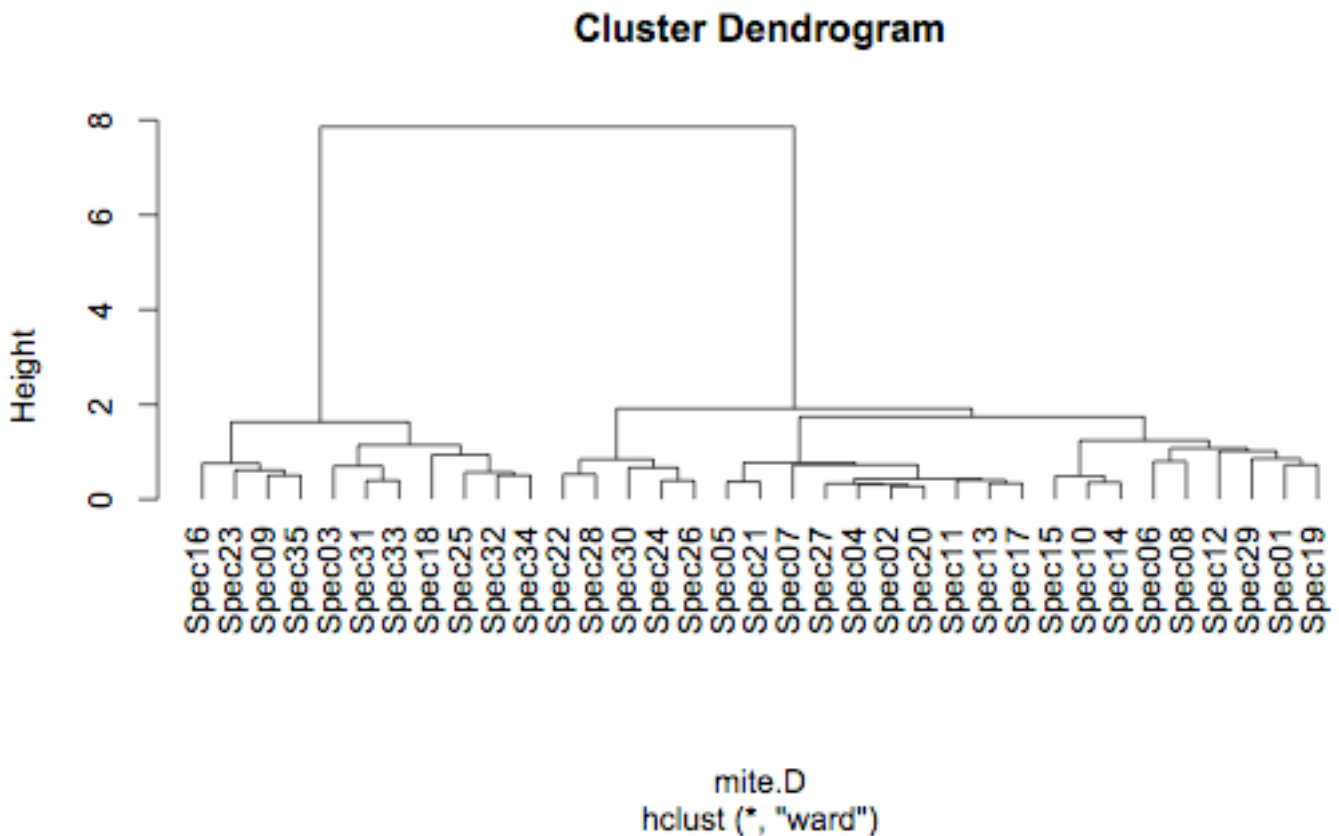
Exemple complet en R : les acariens du sol

```
# Script, analyse des données d'acariens du sol
(70 carottes de sol × 35 espèces ; Borcard & Legendre 1994)
[Exemple utilisé dans l'article de Legendre 2005]

library(vegan)

data(mite)

mite.hel <- decostand(mite, "hel") # Transformation de Hellinger
# Matrice de corrélation de Spearman entre les espèces
mite.cor = cor(mite.hel, method="spearman")
# Groupement agglomératif de Ward (classification ascendante)
mite.D = as.dist(1 - mite.cor)
mite.clust = hclust(mite.D, "ward")
plot(mite.clust, hang=-1)
```



```
# Couper le dendrogramme ; séparer les espèces en 2 groupes
```

```
mite.2 = cutree(mite.clust, k=2)
```

```
Spec01 Spec02 Spec03 Spec04 Spec05 Spec06 Spec07 Spec08 Spec09 Spec10
      1      1      2      1      1      1      1      1      2      1
Spec11 Spec12 Spec13 Spec14 Spec15 Spec16 Spec17 Spec18 Spec19 Spec20
      1      1      1      1      1      2      1      2      1      1
Spec21 Spec22 Spec23 Spec24 Spec25 Spec26 Spec27 Spec28 Spec29 Spec30
      1      1      2      1      2      1      1      1      1      1
Spec31 Spec32 Spec33 Spec34 Spec35
      2      2      2      2      2
```

```
# Test global de chaque groupe
```

```
out.2 <- kendall.global(mite.hel,group=mite.2,mult="holm")
```

```
$Concordance_analysis
```

	Group.1	Group.2
W	0.309787	0.291189
F	10.323050	4.108130
Prob.F	1.1771e-85	4.6766e-22
Corrected prob.F	2.3543e-85	4.6766e-22
Chi2	513.007306	221.012312
Prob.perm	0.001000	0.001000
Corrected prob.perm	0.002000	0.002000

```
# Tests a posteriori de la contribution de chaque espèce à la
concordance de son groupe
```

```
out.post.2 <- kendall.post(mite.hel, group=mite.2, mult="holm", nperm=999)
```

```
$A_posteriori_tests_Group[[1]]
```

	Spec01	Spec02	Spec04	Spec05	Spec06	Spec07	Spec08	Spec10
Spearman.mean	0.1851177	0.4258111	0.3590580	0.2505486	0.1802160	0.2833298	0.09248024	0.2444656
W.per.species	0.2190711	0.4497357	0.3857640	0.2817757	0.2143736	0.3131911	0.13029357	0.2759462
Prob	0.0020000	0.0010000	0.0010000	0.0010000	0.0040000	0.0010000	0.09400000	0.0010000
Corrected prob	0.0350000	0.0350000	0.0350000	0.0350000	0.0360000	0.0350000	0.17000000	0.0350000
	Spec11	Spec12	Spec13	Spec14	Spec15	Spec17	Spec19	Spec20
Spearman.mean	0.4138494	0.1263751	0.4177343	0.3301159	0.2185421	0.4212160	0.2574779	0.4180699
W.per.species	0.4382723	0.1627761	0.4419954	0.3580278	0.2511028	0.4453320	0.2884163	0.4423170
Prob	0.0010000	0.0360000	0.0010000	0.0010000	0.0010000	0.0010000	0.0010000	0.0010000
Corrected prob	0.0350000	0.1700000	0.0350000	0.0350000	0.0350000	0.0350000	0.0350000	0.0350000
	Spec21	Spec22	Spec24	Spec26	Spec27	Spec28	Spec29	Spec30
Spearman.mean	0.3623428	0.1250230	0.2188216	0.3016159	0.4217606	0.2577037	0.1108022	0.2301430
W.per.species	0.3889118	0.1614804	0.2513707	0.3307153	0.4458539	0.2886327	0.1478521	0.2622203
Prob	0.0010000	0.0340000	0.0010000	0.0010000	0.0010000	0.0010000	0.0470000	0.0020000
Corrected prob	0.0350000	0.1700000	0.0350000	0.0350000	0.0350000	0.0350000	0.1700000	0.0350000

\$A_posteriori_tests_Group[[2]]

	Spec03	Spec09	Spec16	Spec18	Spec23	Spec25	Spec31	Spec32
Spearman.mean	0.1222579	0.2712078	0.1906408	0.1375601	0.1342409	0.3342345	0.3446561	0.1833099
W.per.species	0.2020527	0.3374616	0.2642189	0.2159637	0.2129463	0.3947586	0.4042328	0.2575544
Prob	0.0420000	0.0010000	0.0060000	0.0230000	0.0230000	0.0010000	0.0010000	0.0030000
Corrected prob	0.1700000	0.0350000	0.0480000	0.1610000	0.1610000	0.0350000	0.0350000	0.0350000

	Spec33	Spec34	Spec35
Spearman.mean	0.3188922	0.1764232	0.2498877
W.per.species	0.3808111	0.2512938	0.3180797
Prob	0.0010000	0.0030000	0.0010000
Corrected prob	0.0350000	0.0350000	0.0350000

ACP des espèces centrées réduites

