

Méthodes De Couplage De Deux K-Tableaux Et Collections De Graphiques

J. Thioulouse, A. Siberchicot,
A. Julien-Laferrière, A.B. Dufour, S. Dray

`jean.thioulouse@univ-lyon1.fr`
Biométrie & Biologie Evolutive - CNRS - Université Lyon 1

June 26, 2013

Table of Contents

- 1 Introduction - Analyse de données écologiques avec ade4
- 2 Sorties des analyses - Coordonnées factorielles
- 3 Représentations graphiques - Utilisation d'adegraphics
- 4 Conclusion

Méthodes de couplage de deux K-tableaux

Collections de graphiques avec le package **adegraphics**

Rappels sur les méthodes d'analyse de données écologiques avec **ade4**

- Les méthodes d'analyse d'un tableau
- Les méthodes d'analyse d'un couple de tableaux
- Les méthodes d'analyse de K-tableaux
- Les méthodes d'analyse de **K couples de tableaux** (deux K-tableaux)

- Les méthodes d'analyse d'**un tableau**
ACP, AFC, ACM: une dizaine dans **ade4** (`dudi.xxx`)
 - En écologie: tableaux de dénombrements d'espèces
 - Abondances, % recouvrement, présence/absence, traits spécifiques
 - Données génétiques: micro-satellites, puces ADN, séquençage
 - Variables environnementales (milieu), données satellites (land cover)
- **Objectifs**: description de structures
 - Typologie d'individus (lignes = sites)
 - Corrélations entre variables (colonnes = espèces / variables)

- Les méthodes d'analyse d'**un couple de tableaux**
ACC, RDA, COIA, RLQ, pcaiv, etc. huit fonctions dans **ade4**
 - Un tableau de données d'espèces
 - Un tableau de données de variables du milieu (*mêmes sites*)
 - Plusieurs familles de méthodes:
 - Direct / indirect gradient analysis
 - Analyses sous contraintes (combinaisons linéaires des var. milieu)
 - Analyse de coinertie (covariances croisées)
- **Objectifs:** description des relations entre les deux tableaux
 - Prédiction d'un tableau par l'autre
 - Analyse de co-structures (typologies)
 - Test statistique (permutations)

- Les méthodes d'analyse de **K-tableaux**
STATIS, ATP, AFC Foucart
 - Une série de tableaux de données
En écologie: un tableau espèces x sites répété au cours du temps
 - Trois étapes: Interstructure / Compromis / Intrastructure
 - *Interstructure*: importance de chaque tableau, coeff. CL compromis
 - *Compromis*: analyse d'un tableau "moyen" pondéré
 - *Intrastructure*: projection des lignes et des colonnes de tous les tableaux initiaux dans l'analyse du compromis
- **Objectifs**: analyse de la stabilité des structures
 - Typologie des tableaux (interstructure)
 - Structures stables: tableau "moyen" (compromis)
 - Description des modifications de structures (intrastructure)

- Les méthodes d'analyse de **K couples de tableaux**
BGCOIA (coinertie inter-classes)
STATICO, COSTATIS (ATP / coinertie)
DO-ACT (double STATIS, Vivien & Sabatier 2004)
 - Une série de tableaux de données d'espèces
 - Une série de tableaux de données du milieu
 - Stabilité des relations espèces/milieu
- **Objectifs:** analyse de la stabilité des co-structures espèces/milieu
 - Typologie de co-structures (typologie de typologies)
 - Co-structures stables
 - Description des modifications de co-structures

Sorties des analyses (**ade4**): dataframes de coordonnées factorielles

- Les méthodes d'analyse d'**un tableau**
 - Tableau (n, p) , f axes conservés pour le calcul coordonnées
 - Coordonnées des individus (lignes): `xxx$li` (n, f)
 - Coordonnées des variables (colonnes): `xxx$co` (p, f)
- Les méthodes d'analyse d'**un couple de tableaux**
Analyse de coinertie:
 - Tableau espèces (n, p_1)
 - Tableau milieu (n, p_2)
 - Tableau croisé (p_1, p_2)
 - Coordonnées des colonnes du tableau espèces: `$li` (p_1, f)
 - Coordonnées des colonnes du tableau milieu: `$co` (p_2, f)
 - Coordonnées des lignes du tableau espèces: `$lX` (n, f)
 - Coordonnées des lignes du tableau milieu: `$lY` (n, f)

- Les méthodes d'analyse de **K-tableaux**

Analyse Triadique Partielle (ATP): t tableaux, (n, p)

- Interstructure: $RV.coo(t, f)$
- Compromis (individus) : $li(n, f)$
- Compromis (variables) : $co(p, f)$
- Intrastructure (proj. lignes des tableaux initiaux): $Tli(t * n, f)$
- Intrastructure (proj. colonnes des tableaux initiaux): $Tco(t * p, f)$

- Les méthodes d'analyse de **K couples de tableaux**

Analyse STATICO (ATP des tableaux croisés espèces x milieu):

t tableaux espèces (n, p_1), t tableaux milieu (n, p_2)

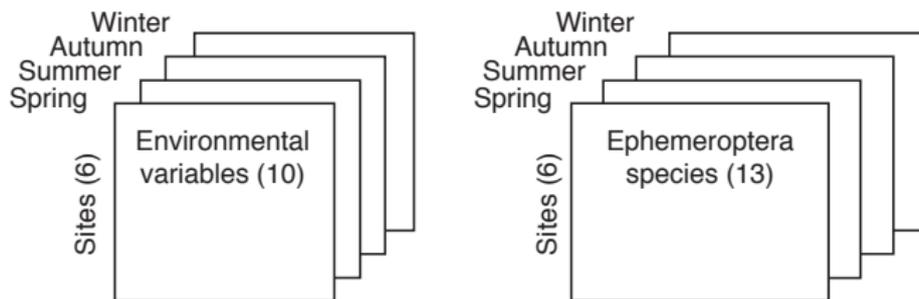
- Interstructure: $\$RV.coo(t, f)$
- Compromis (espèces) : $\$li(p_1, f)$
- Compromis (milieu) : $\$co(p_2, f)$
- Intrastructure (proj. des col. tableaux d'espèces): $\$Tli(t * p_1, f)$
- Intrastructure (proj. des col. tableaux de milieu): $\$Tco(t * p_2, f)$
- Intrastructure (proj. des lig. tableaux d'espèces): $\$supIX(t * n, f)$
- Intrastructure (proj. des lig. tableaux de milieu): $\$supIY(t * n, f)$

- Tableaux de coordonnées factorielles \rightarrow représentations graphiques
- Exemple: jeu de données réel de petites dimensions
- Méaudret: 4 couples de tableaux espèces \times milieu (un par saison)
- Méthode STATICO (couplage de deux K-tableaux)
- Anciennes fonctions graphiques d'**ade4** vs. **adegraphics**

Graphiques: exemple

Exemple: données du Méaudret (Autrans), package **ade4**

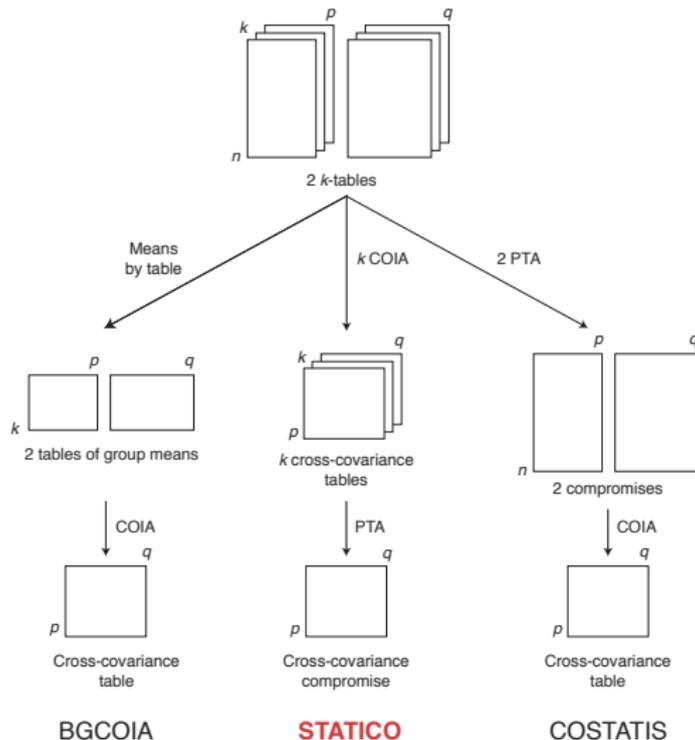
- 6 stations réparties le long d'un cours d'eau, échantillonnées 4 fois
- un K-tableau milieu: 10 variables physico-chimiques
- un K-tableau 13 espèces éphéméroptères.



Jeu de données du Méaudret, package **ade4**.

Graphiques: méthode STATICO

Méthode STATICO, package **ade4**



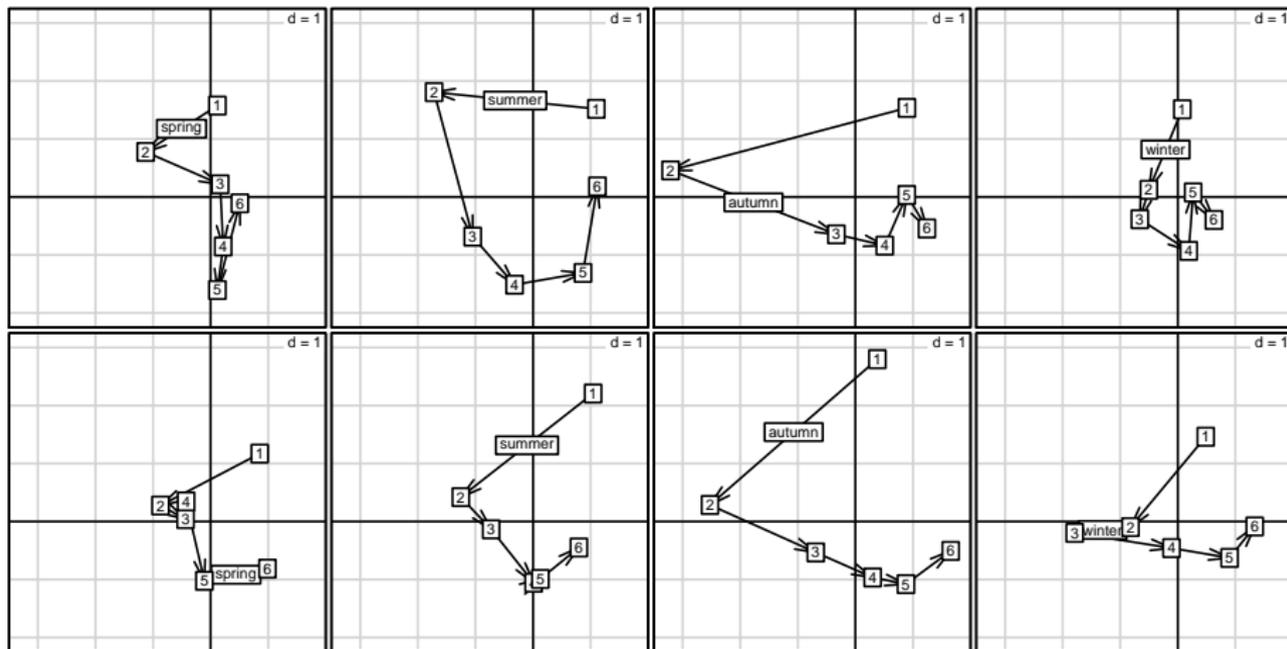
- Méthode STATICO (couplage de deux K-tableaux)
- Intrastructure, "Trajectoire des individus"
- Collection de graphiques obtenue par juxtaposition des projections des lignes (stations) de tous les tableaux des deux séries dans l'analyse du compromis.
- Calculs:

```
data(meau)
wit1 <- withinpca(meau$env, meau$design$season, scal = "total")
spepca <- dudi.pca(meau$spe, scale = FALSE)
wit2 <- wca(spepca, meau$design$season)
kta1 <- ktab.within(wit1) # premier k-tableau
kta2 <- ktab.within(wit2) # second k-tableau
kcoi <- ktab.match2ktabs(kta1, kta2) # K-tableau croisés
statico1 <- pta(kcoi) # Analyse triadique partielle
```

Graphiques des trajectoires: anciennes fonctions d'**ade4**

```
par(mfrow=c(2,2))
for (i in levels(reorder(meau$design$season, rep(c(1, 2, 3, 4), each = 6)))) {
  s.traject(statico1$supIX[meau$design$season == i, ], label = i, ,
  xlim = c(-3.5, 2), ylim = c(-2, 3))
  s.label(statico1$supIX[meau$design$season == i, ], label = sites, ,
  xlim = c(-3.5, 2), ylim = c(-2, 3), add.plot = T)
}
for (i in levels(reorder(meau$design$season, rep(c(1, 2, 3, 4), each = 6)))) {
  s.traject(statico1$supIY[meau$design$season == i, ], label = i, ,
  xlim = c(-3.5, 2), ylim = c(-2, 3))
  s.label(statico1$supIY[meau$design$season == i, ], label = sites, ,
  xlim = c(-3.5, 2), ylim = c(-2, 3), add.plot = T)
}
```

Graphiques avec ade4

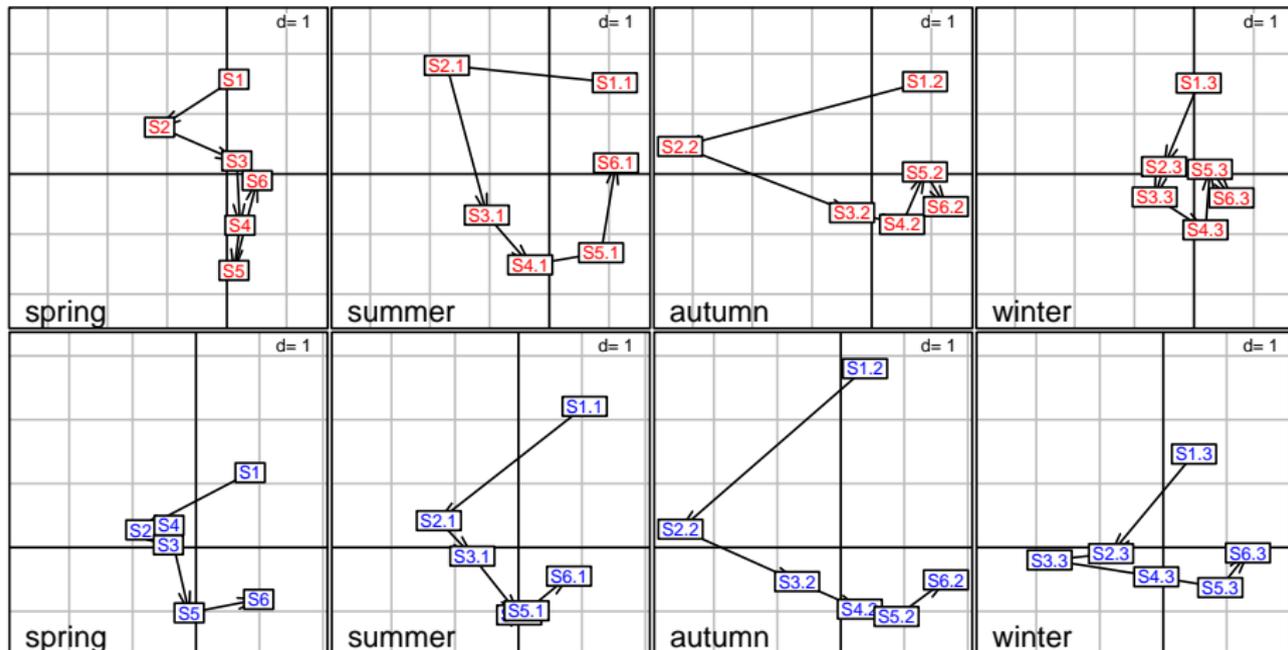


Intrastructure des stations pour les variables du milieu (en haut)
et pour les espèces (en bas), package **ade4**.

Nouvelles fonctions du package **adegraphics**

```
st1 <- s.traject(statico1$supIX, facets = statico1$supTI[ ,1],
  plabels.cex = 0, psub.cex = 0, plot = FALSE)
sla1 <- s.label(statico1$supIX, facets = statico1$supTI[ ,1],
  psub.cex = 1.5, plabel.col = "red", plot = FALSE)
st2 <- s.traject(statico1$supIY, facets = statico1$supTI[ ,1],
  plabels.cex = 0, psub.cex = 0, plot = FALSE)
sla2 <- s.label(statico1$supIY, facets = statico1$supTI[ ,1],
  psub.cex = 1.5, plabel.col = "blue", plot = FALSE)
pos1 <- rbind(c(0,0,.25,1),c(.25,0,.5,1),c(.5,0,.75,1),c(.75,0,1,1))
st1@positions <- st2@positions <- pos1
sla1@positions <- sla2@positions <- pos1
s1 <- superpose(st1, sla1)
s2 <- superpose(st2, sla2)
ADEgS(list(s1, s2), layout = c(2, 1))
```

Graphiques avec adeggraphics



Intrastructure des stations pour les variables du milieu (en haut)
et pour les espèces (en bas), package **adeggraphics**.

- Package **adegraphics**: Collections de graphiques Lattice Classes S4
- Beaucoup d'autres avantages (objets graphiques modifiables)
- Facilité d'utilisation et de programmation
- Méthodes d'analyse de données: couplage de 2 K-tableaux
- Stages de formation **ade4**: Novembre/Décembre 2013
- Ouvrage en cours de rédaction: Collection Use R! Springer