

Problème Pratique de Statistique - 23

Chili, Californie, Bourgogne, Provence : communautés d'oiseaux

Avec **data (ecomor)** dans la librairie **ade4**, on obtient un objet complexe qui pose des questions statistiques difficiles. C'est une transcription fidèle des données publiées dans Blondel et al. (1984). S'associent dans cet article des biogéographes ornithologues spécialistes des avifaunes des régions de climat méditerranéens.



Carte des régions méditerranéennes du globe : <http://www.imep-cnrs.com/ecologia/carted2.jpg>

ecomor est une liste de 6 composantes. L'ensemble est centré sur un cortège de 129 espèces d'oiseaux vivant soit en Europe, soit en Amérique du Nord, soit en Amérique du Sud. Ces cortèges sont totalement disjoints. Les éléments de la liste décrits ici.

\$forsub est un data.frame à 129 lignes et 6 colonnes. Il donne le lieu d'alimentation (*foraging substrate*) des espèces étudiées indiqué par des variables binaires (1 = oui, 0 = non). Ces variables sont "foliage" "ground" "twig" "bush" "trunk" "aerial".

foliage	alimentation dans le feuillage
ground	alimentation au sol
twig	alimentation sur les rameaux
bush	alimentation dans les buissons
trunk	alimentation sur les troncs
aerial	alimentation aérienne

\$diet est un data.frame à 129 lignes et 8 colonnes. Il donne le type d'alimentation (régime alimentaire, *diet type*) des espèces étudiées indiqué par des variables binaires (1 = oui, 0 = non). Ces variables sont :

Gr	granivorous	consommation de graines
Fr	frugivorous	baies, fruits, glands
Ne	nectarvorus	consommation de nectar
Fo	folivorous	consommation des feuilles
In	invertebrate feeder	
		insectes, araignées, myriapodes, isopodes, vers, serpents
Ca	carnivorous	consommation de la chair des petits vertébrés
Li	limnivorous	consommation des invertébrés d'eaux douces
Ch	carrion feeder	consommation de charognes

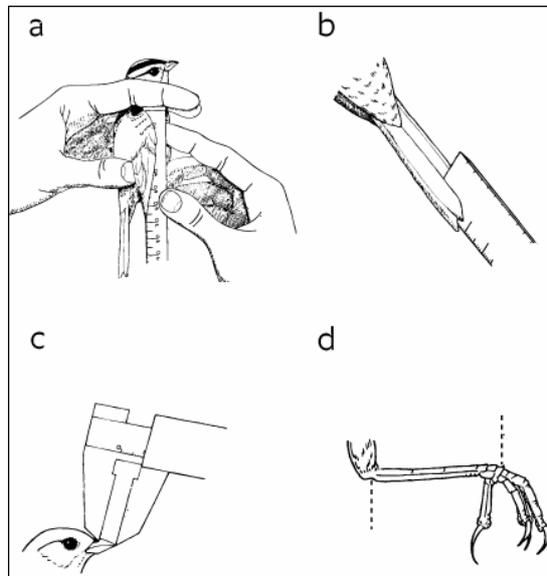
\$habitat est un data.frame à 129 lignes et 16 colonnes. Il donne le type d'habitat où est présente chacune des espèces étudiées, type indiqué par des variables binaires (1 = présent, 0 = absent). Ces variables sont

Bu1, Bu2, Bu3, Bu4, Ca1, Ca2, Ca3, Ca4, Ch1, Ch2, Ch3, Ch4, Pr1, Pr2, Pr3 et Pr4.

L'étude associe une zone témoin (Bu pour Bourgogne) et trois zones méditerranéennes (Ca la Californie, Ch le Chili et Pr la Provence). Dans chaque zone les auteurs ont défini 4 types d'habitat notés respectivement 2 pour une forêt de 1 à 7 ans, 3 pour une forêt de 9 à 20 ans, 4 pour une forêt de 40 à 70 ans, et 5 pour une forêt de 71 à 125 ans. Nous utiliserons simplement le code 1|2|3|4 pour désigner ces quatre types de milieu. Il s'agit des stades de successions du milieu forestier depuis le stade pionnier jusqu'à la forêt mûre.

\$morpho est un data.frame à 129 lignes et 8 variables morphologiques :

```
wingl longueur de l'aile en mm (Wing length)
taill longueur de la queue en mm (Tail length)
culm longueur du bec en mm (Culmen length)
bilh hauteur du bec en mm (Bill height)
billw largeur du bec en mm (Bill width)
tarsl hauteur du tarse en mm (Tarsus length)
midtl longueur de l'orteil médian en mm (Middle toe length)
weig poids en g (Weight)
```



<http://cm27personal.fal.buffalo.edu/birds/anatomy/molt/size.html> Measuring the wing chord (*wingl* = a), the tail between the central rectrices (*taill* = b), the bill from nares to tip (*culm* = c) and the tarsus (*tarsl* = d).

\$taxo est un data.frame à 129 lignes et 3 facteurs qui donne la position taxonomique de chaque espèce (facteurs emboîtés) :

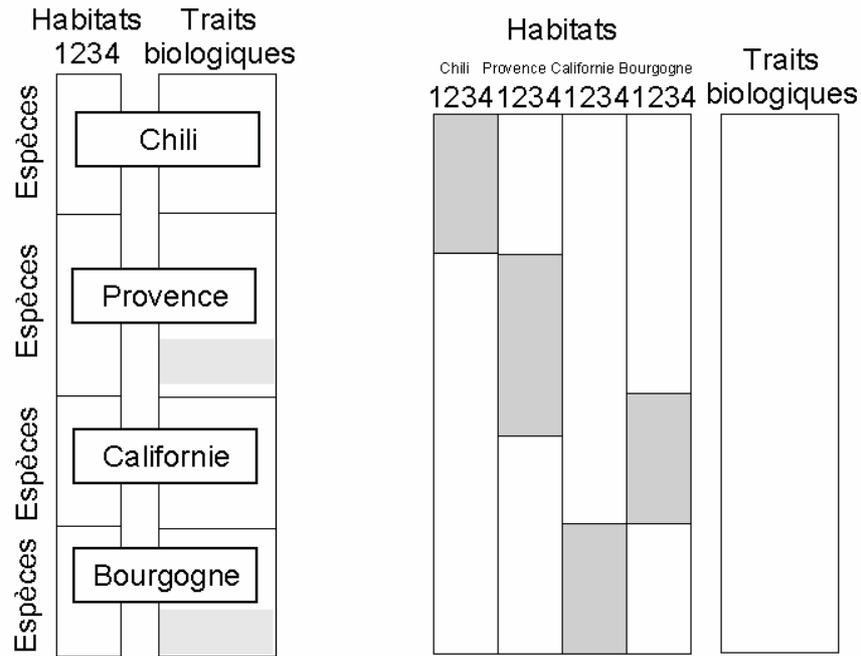
Genus	genre	86 genres différents
Family	famille	35 familles différentes
Order	ordre	7 ordres différents

\$labels est une matrice à 129 lignes et 2 colonnes de chaînes de caractères. Le nom des lignes est le nom des lignes de tous les tableaux, la première colonne contient les noms latins et la seconde colonne une abbréviation à 7 caractères du précédent :

ecomor\$labels

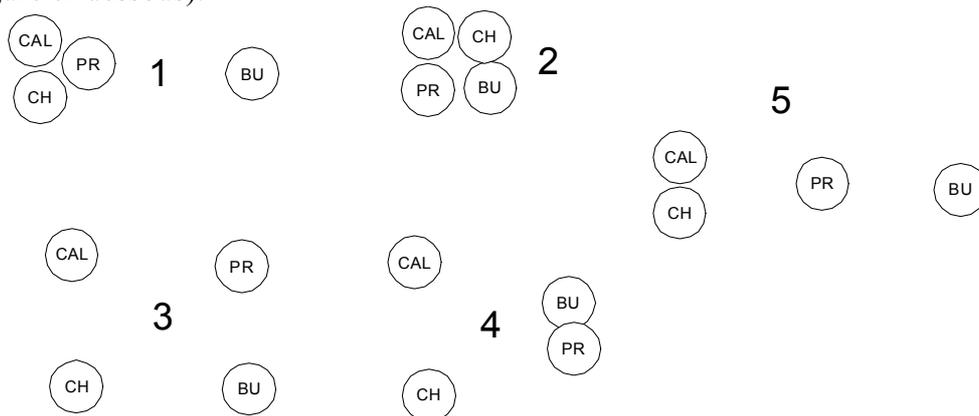
	latin	abbr
E033	"Archilochus alexandri"	"Arc ale"
E034	"Calypte anna"	"Cal ann"
E035	"Calypte costae"	"Cal cos" ...

La seule différence entre cette présentation et les données publiées s'explique par une petite imprécision dans l'article :



A gauche, dans l'article d'origine, seules quelques espèces européennes apparaissent dans deux régions. Leur profil entre type de milieu n'était présent qu'à un seul exemplaire. C'était une erreur que J. Blondel a corrigé en me communiquant les profils Pr1-4 et Bu1-4 de ces taxons : ces profils peuvent être différents. On a donc gardé le présentation de droite.

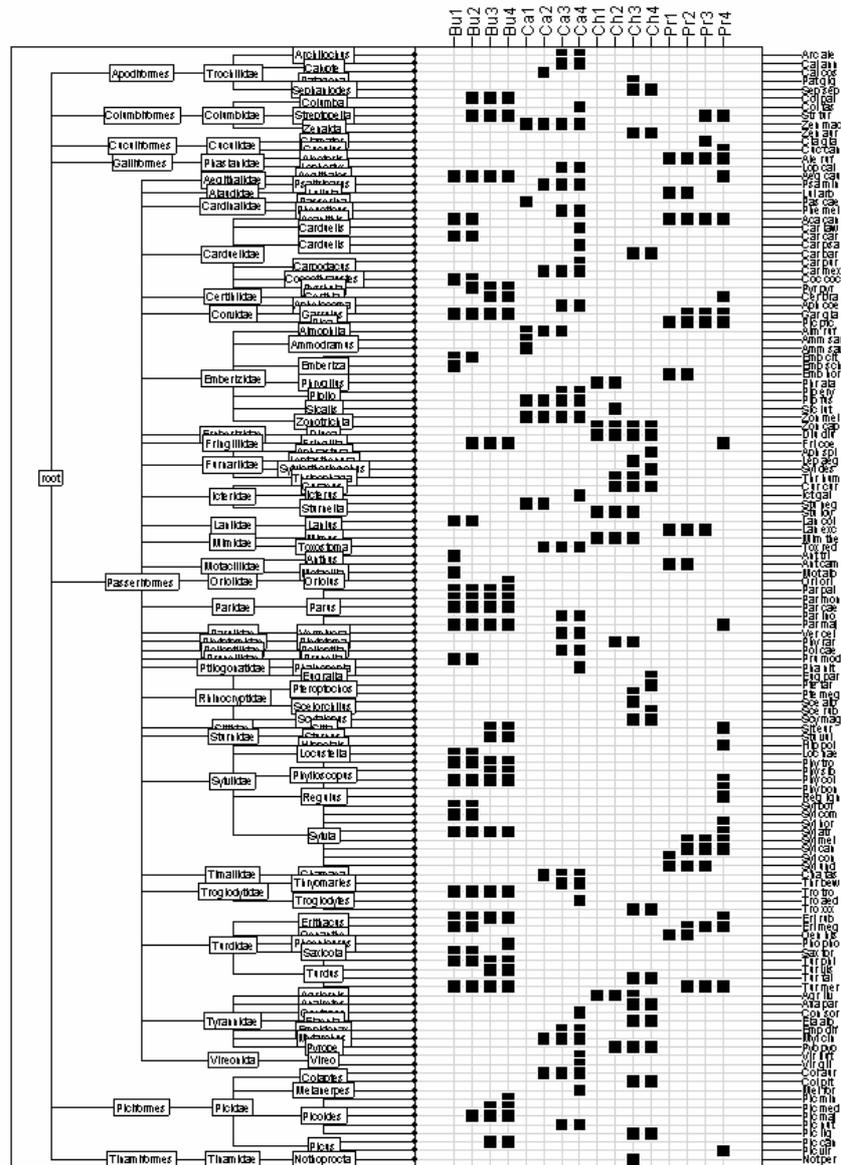
La perturbation est petite mais a grande signification car elle conduit à deux problèmes différents. La forme de gauche est un ensemble de quatre couples de tableaux écologiques qui demande la comparaison de quatre évolutions du contenu en traits biologiques entre les quatre types d'habitat. La forme de droite est un couple unique qui envisage la variation en traits biologiques soit entre régions (analyse inter), soit entre strates d'une même région (analyse intra), soit entre tous les éléments (analyse simple). Or la situation est encore compliquée par le schéma des hypothèses de l'article (figure ci-dessous).



1 - modèle de la convergence bioclimatique, 2/3 - absence de convergence, 4 - modèle phylogénétique, 5 - convergence bioclimatique ajustée (CAL Californie, CH Chili, PR Provence, BU Bourgogne).

On y voit une typologie de régions soit faible (2) soit forte (3) soit organisée (1, 4, 5) mais sans référence au gradient environnemental commun. Or comment s'intéresser à la variation entre régions sans tenir compte que chacune d'entre elles est elle-même une structure (et bien connue pour être forte) ? La question est donc très ouverte.

Les objectifs de l'analyse sont donc de comparer des cortèges d'avifaune entre régions (et/ou entre habitats) par le biais non pas des noms d'espèces mais des traits biologiques des espèces, par exemple la taxonomie, vue comme première approche de la phylogénie :



On trouvera des indications dans Pavoine et al. (2004).

Blondel, J., F. Vuilleumier, L. F. Marcus, and E. Terouanne. 1984. Is there ecomorphological convergence among mediterranean bird communities of Chile, California, and France. Pages 141-213 in M. K. Hecht, B. Wallace, and R. J. MacIntyre, editors. *Evolutionary Biology*. Vol. 18. Plenum Press, New York.

Pavoine, S., A.-B. Dufour, and D. Chessel. 2004. From dissimilarities among species to dissimilarities among communities: A Double Principal Coordinate Analysis. *Journal of Theoretical Biology* **228**:523-537.