

Problème pratique de statistique n° pps056

# Structure des communautés aviaires de lacs du Nord-Est des Etats-Unis

Julien de Bortoli

Master aMIV option oMIV 2004-2205

## Table des matières

<b>1</b>	<b>Présentation</b>	<b>1</b>
<b>2</b>	<b>Les communautés aviaires</b>	<b>1</b>
<b>3</b>	<b>Informations géographiques</b>	<b>2</b>
<b>4</b>	<b>Variables de milieu</b>	<b>5</b>
	<b>Références</b>	<b>5</b>

## 1 Présentation

Dans le cadre de l'Environmental Monitoring and Assessment Program (EMAP) conduit pour appréhender les questions d'état écologique, en particulier les ressources aquatiques du Nord Est des Etats-Unis, plusieurs centaines de lacs naturels et artificiels ont été sélectionnés parmi plus de 10 000 lacs répertoriés par l'US Geological Survey (USGS). La sélection a été opérée à l'aide d'une méthode statistique respectant la diversité des conditions régionales (Larsen et al. 1994 [2]). Sur chacun des sites retenus ont été menés des inventaires faunistiques pour plusieurs compartiments biotiques (poissons, oiseaux, zooplancton, diatomées).

L'étude ne portera que sur les populations aviaires observées. Outre l'inventaire faunistique la base de données contient également des informations relatives au milieu. Les analyses ont pour but de démontrer la présence ou non d'une structure dans l'assemblage des espèces et l'existence un lien entre la composition des communautés et les variables décrivant le milieu.

Les données utilisées sont disponibles sur le site de l'United States Environmental Protection Agency sous forme de fichiers textes à l'adresse suivante :

<http://www.epa.gov/emap/html/dataI/surfwatr/data/nelakes/>

L'objet pps056 contient le tableau d'occurrence des espèces, le tableau des variables de milieu et le tableau des sites (avec leurs coordonnées et leur état d'appartenance).

```
load(url("http://pbil.univ-lyon1.fr/R/donnees/pps056.rda"))
names(pps056)
[1] "avifaune" "milieu" "sites"
```

## 2 Les communautés aviaires

L'échantillonnage aviaire a été réalisé sur 208 lacs entre 1991 et 1995 par l'Environmental Monitoring and Assessment Program (EMAP) de l'U.S. Environmental Protection Agency. Les

relevés ont été effectués en journée le long d'un transect suivant les rives du lac d'environ 10 mètres; les données ont été collectées tous les 200 mètres sur le périmètre du lac, excepté pour les lacs dont le périmètre est supérieur à 4,8 km pour lesquels 24 points ont été déterminés de manière à couvrir tous les habitats. A chaque "point stop", pendant 5 minutes, tous les individus vus ou entendus dans un rayon d'environ 100 mètres ont été enregistrés, avec reconnaissance au niveau spécifique. Des précisions supplémentaires sur les opérations sont données dans Baker et al. (1997) [1].

Un tableau de présence/absence a été construit à partir des relevés faunistiques disponibles. Sur les 208 plans d'eau échantillonnés 176 espèces ont été enregistrées. Afin de s'affranchir du phénomène des espèces rares (dont le poids biaise l'analyse) les espèces présentes dans moins de 11 sites sont retirées du jeu de données; la limite d'occurrence est fixée à 5%. Après sélection il reste 95 espèces. Aucun site n'est retiré car ils présentent tous un nombre suffisant d'espèces. Les espèces sont codées par concaténation de lettres du nom commun des espèces.


```
dim(pps056$avifaune)
[1] 208 95
names(pps056$avifaune)
 [1] "ALFL" "AMCR" "AMGO" "AMRE" "AMRO" "BAOR" "BBWA" "BCCH" "BEKI" "BHCO" "BKSW"
 [12] "BLDU" "BLJA" "BLWA" "BRCR" "BRSW" "BTBW" "BTGW" "BWA" "CAGO" "CAWA" "CEWA"
 [23] "CHSP" "CHSW" "COGR" "COLO" "COME" "CORA" "COSN" "COYE" "CSWA" "DCCO" "DEJU"
 [34] "DOWO" "EAKI" "EAPH" "EAWO" "EUST" "EVGR" "GBHE" "GCFL" "GCKI" "GRCA" "HAWO"
 [45] "HEGU" "HETH" "HOFL" "HOME" "HOSP" "HOWR" "KILL" "LEFL" "MADU" "MAWA" "MODO"
 [56] "NAWA" "NOCA" "NOFL" "NOPA" "NOWA" "OSFL" "OSPR" "OVEN" "PISI" "PIWA" "PIWO"
 [67] "PUFI" "RBGR" "RBGU" "RBNU" "RCKI" "REVI" "RODO" "RSTO" "RTHU" "RWBB" "SCTA"
 [78] "SOXP" "SOVI" "SPSA" "SWSP" "SWTH" "TRSW" "TUTI" "VEER" "WAVI" "WBNU" "WIWR"
 [89] "WODU" "WOTH" "WTSP" "YBFL" "YBSA" "YEWA" "YRWA"
```

La liste des codes et des noms scientifiques des espèces correspondants sont donnés dans le tableau 1.

### 3 Informations géographiques

La base de données contient des informations relatives aux caractéristiques physiques des plans d'eau ainsi que de leur bassin versant. Les données topographiques sont issues de cartes digitalisées (échelle 1 :24000 ou 1 :100000). La zone d'étude recouvre 8 états du Nord-Est des Etats-Unis (Connecticut, Massachusetts, Maine, New Hampshire, New Jersey, New York, Rhode Island, Vermont).

```
dim(pps056$sites)
[1] 208 3
names(pps056$sites)
[1] "STATE" "LON_DD" "LAT_DD"
levels(pps056$sites$STAT)
[1] "CT" "MA" "ME" "NH" "NJ" "NY" "RI" "VT"
```

Pour avoir un fond de carte comportant les 8 états concernés, utiliser les ressources de .

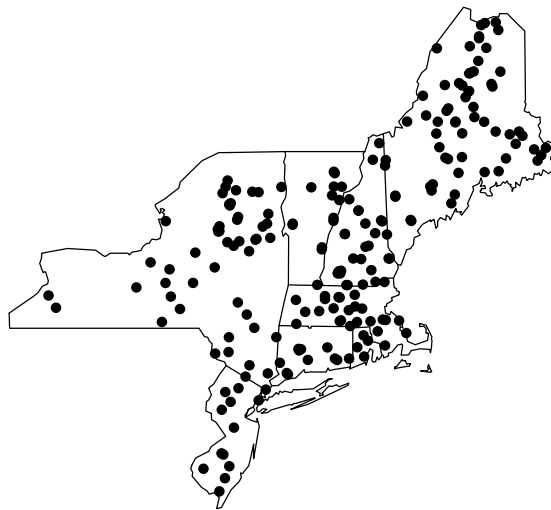
```
library(maps)
map("state")$names[1:10]
 [1] "alabama" "arizona" "arkansas"
 [4] "california" "colorado" "connecticut"
 [7] "delaware" "district of columbia" "florida"
[10] "georgia"
```

Code	Nom	Code	Nom	Code	Nom
AMCR	<i>Corvus brachyrhynchos</i>	DOWO	<i>Picoides pubescens</i>	PIWO	<i>Dryocopus pileatus</i>
AMGO	<i>Carduelis tristis</i>	EAKI	<i>Tyrannus tyrannus</i>	PUIFI	<i>Carpodacus purpureus</i>
AMRE	<i>Setophaga ruticilla</i>	EAPH	<i>Sayornis phoebe</i>	RBGR	<i>Pheucticus ludovicianus</i>
AMRO	<i>Turdus migratorius</i>	EAWO	<i>Contopus virens</i>	RBGU	<i>Larus delawarensis</i>
BAOR	<i>Icterus galbula</i>	EUST	<i>Sturnus vulgaris</i>	RBNU	<i>Sitta canadensis</i>
BBWA	<i>Dendroica castanea</i>	EVGR	<i>Coccothraustes vespertinus</i>	RCKI	<i>Regulus calendula</i>
BCCH	<i>Parus atricapillus</i>	GBHE	<i>Ardea herodias</i>	REVI	<i>Vireo olivaceus</i>
BEKI	<i>Ceryle alcyon</i>	GCFL	<i>Myiarchus crinitus</i>	RODO	<i>Columba livia</i>
BHCO	<i>Molothrus ater</i>	GCKI	<i>Regulus satrapa</i>	RSTO	<i>Pipilo erythrophthalmus</i>
BKSW	<i>Riparia riparia</i>	GRCA	<i>Dumetella carolinensis</i>	RTHU	<i>Archilochus colubris</i>
BLDU	<i>Anas rubripes</i>	HAWO	<i>Picoides villosus</i>	RWBB	<i>Agelaius phoeniceus</i>
BLJA	<i>Cyanocitta cristata</i>	HEGU	<i>Larus argentatus</i>	SCTA	<i>Piranga olivacea</i>
BLWA	<i>Dendroica fusca</i>	HETH	<i>Catharus guttatus</i>	SOSP	<i>Melospiza melodia</i>
BRCR	<i>Certhia americana</i>	HOFI	<i>Carpodacus mexicanus</i>	SOVI	<i>Vireo solitarius</i>
BRSW	<i>Hirundo rustica</i>	HOME	<i>Lophodytes cucullatus</i>	SPSA	<i>Actitis macularia</i>
BTBW	<i>Dendroica caerulescens</i>	HOSP	<i>Passer domesticus</i>	SWSP	<i>Melospiza georgiana</i>
BTGW	<i>Dendroica virens</i>	HOWR	<i>Troglodytes aedon</i>	SWTH	<i>Catharus ustulatus</i>
BWWA	<i>Mniotilta varia</i>	KILL	<i>Charadrius vociferus</i>	TRSW	<i>Tachycineta bicolor</i>
CAGO	<i>Branta canadensis</i>	LEFL	<i>Empidonax minimus</i>	TUTI	<i>Parus bicolor</i>
CAWA	<i>Wilsonia canadensis</i>	MADU	<i>Anas platyrhynchos</i>	VEER	<i>Catharus fuscescens</i>
CEWA	<i>Bombycilla cedrorum</i>	MAWA	<i>Dendroica magnolia</i>	WAVI	<i>Vireo gilvus</i>
CHSP	<i>Spizella passerina</i>	MODO	<i>Zenaida macroura</i>	WBNU	<i>Sitta carolinensis</i>
CHSW	<i>Chaetura pelagica</i>	NAWA	<i>Vermivora ruficapilla</i>	WIWR	<i>Troglodytes troglodytes</i>
COGR	<i>Quiscalus quiscula</i>	NOCA	<i>Cardinalis cardinalis</i>	WODU	<i>Aix sponsa</i>
COLO	<i>Gavia immer</i>	NOFL	<i>Colaptes auratus</i>	WOTH	<i>Hylocichla mustelina</i>
COME	<i>Mergus merganser</i>	NOPA	<i>Parula americana</i>	WTSP	<i>Zonotrichia albicollis</i>
CORA	<i>Corvus corax</i>	NOWA	<i>Seiurus noveboracensis</i>	YBFL	<i>Empidonax flaviventris</i>
COSN	<i>Gallinago gallinago</i>	OSFL	<i>Contopus borealis</i>	YBSA	<i>Sphyrapicus varius</i>
COYE	<i>Geothlypis trichas</i>	OSPR	<i>Pandion haliaetus</i>	YEWA	<i>Dendroica petechia</i>
CSWA	<i>Dendroica pensylvanica</i>	OVEN	<i>Seiurus aurocapillus</i>	YRWA	<i>Dendroica coronata</i>
DCCO	<i>Phalacrocorax auritus</i>	PISI	<i>Carduelis pinus</i>		

TAB. 1 – Liste des codes espèces et des noms scientifiques correspondants



```
w = c("connecticut", "massachusetts:martha's vineyard", "massachusetts:main")
w = c(w, "massachusetts:nantucket", "maine", "new hampshire", "new jersey")
w = c(w, "new york:manhattan", "new york:main", "new york:staten island")
w = c(w, "new york:long island", "rhode island", "vermont")
map("state", w)
points(pps056$sites[, 2:3], pch = 19)
```



## 4 Variables de milieu

Par ailleurs sont renseignées des données caractérisant l'occupation du bassin versant (BV) ; diverses couches cartographiques ont permis la mesure d'indicateurs sur l'occupation de l'espace en terme d'activités humaine, industrielle et agricole. D'autres informations ont été collectées dans diverses bases américaines et dans des documents non publiés. Les variables contenues dans le tableau 2 ont été retenues pour l'étude. Certaines variables (AG\_TOT, FOR\_TOT, HOUDENKM, MINE\_TOT,

Code	Description
AG_TOT	Pourcentage du BV utilisé par l'agriculture
BAR_TOT	Pourcentage du BV occupé par la toundra
ELEV	Altitude du lac
FOR_TOT	Pourcentage de forêts dans le BV
KM_SEA	Distance à la mer
HOUDENKM	Densité d'habitations dans le BV (maisons/km <sup>2</sup> )
MINE_TOT	Pourcentage du BV utilisé par l'industrie minière
POPDENKM	Densité de population (habitants/km <sup>2</sup> )
PRECIP_M	Précipitation annuelle moyenne (m)
RD_DEN	Densité du réseau routier dans le BV
URB_TOT	Pourcentage du BV urbanisé
WETL_TOT	Pourcentage du BV en zones humides

TAB. 2 – Définition des variables environnementales

POPDENKM, RD\_DEN, URB\_TOT et WETL\_TOT) doivent subir une transformation logarithmique pour diminuer les écarts observés dans les distributions.

## Références

- [1] J. R. Baker, D. V. Peck, and D. W. Sutton, editors. *Environmental Monitoring and Assessment Program Surface Waters : Field Operations Manual for Lakes*. U.S. Environmental Protection Agency, Washington., 1997.
- [2] D. P. Larsen, K. W. Thornton, N. S. Urquhart, and S. G. Paulsen. The role of sample surveys for monitoring the condition of the nation's lakes. *Environmental Monitoring and Assessment*, 32 :101–134, 1994.