

# Mesure de la diversité en zone alluviale

*Pierre Marle - Laboratoire d'Ecologie et de Biologie Aquatique*

*Master 2 - Qualité des Eaux, des Sols et Traitements - 11 Février 2019*

## Objectifs

1. Acquérir la pratique de technique de mesure et d'analyse de la diversité.
2. Utiliser des informations relatives à la biologie et à l'écologie des espèces afin de mener une analyse des annexes fluviales étudiées.
3. Lier les concept des processus fluviaux à la diversité.

Les données utilisées ici comme exemple concernent des trichoptères au stade adultes, capturés par piège lumineux dans différents sites de la zone alluviale du Haut-Rhône français. Ce jeux de données est tiré de l'article *Marle et al. (2016) Light-trapped caddisfly assemblages in two floodplain reaches of the French upper Rhone River [Trichoptera]. Ephemera18(1):41-59* et ont été acquises dans le cadre du suivi scientifique de la restauration hydrologique et écologique du Rhône (<http://restaurationrhone.univ-lyon1.fr/>).

## Chargement et présentation du jeu de données

```
setwd("C:/Users/marle/Dropbox/Cours")
Tr <- read.csv("tricho2.csv", header = TRUE, sep=";") #Attention sep=";" ou sep=","
attach(Tr)
names(Tr)
```

```
## [1] "site" "periode"
## [3] "type" "date"
## [5] "latitude" "longitude"
## [7] "annee" "Agapetus.ochripes"
## [9] "Agraylea.multipunctata" "Agraylea.sexmaculata"
## [11] "Agrypnia.varia" "Allogamus.kefes"
## [13] "Anabolia.nervosa" "Athripsodes.albifrons"
## [15] "Athripsodes.aterrimus" "Athripsodes.cinereus"
## [17] "Ceraclea.albimacula" "Ceraclea.aurea"
## [19] "Ceraclea.dissimilis" "Cyrnus.crenaticornis"
## [21] "Cyrnus.trimaculatus" "Goera.pilosa"
## [23] "Glossosoma.boltoni" "Glyphotaelius.pellucidus"
## [25] "Grammotaulius.nigropunctatus" "Halesus.radiatus"
## [27] "Halesus.tesselatus" "Hydropsyche.contubernalis"
## [29] "Hydropsyche.incognita" "Hydropsyche.exocellata"
## [31] "Hydropsyche.modesta" "Hydropsyche.pellucidula"
## [33] "Hydropsyche.siltalai" "Hydroptila.angulata"
## [35] "Hydroptila.forcipata" "Hydroptila.sparsa"
## [37] "Hydroptila.vectis" "Lepidostoma.hirtum"
## [39] "Leptocerus.tineiformis" "Limnephilus.affinis"
## [41] "Limnephilus.flavicornis" "Limnephilus.flavospinosus"
## [43] "Limnephilus.lunatus" "Limnephilus.rhombicus"
## [45] "Lype.phaeopa" "Lype.reducta"
```

```

## [47] "Mesophylax.impunctatus"      "Micropterna.testacea"
## [49] "Mystacides.azurea"           "Mystacides.longicornis"
## [51] "Oecetis.lacustris"          "Oecetis.notata"
## [53] "Orthotrichia.costalis"       "Oxyethira.flavicornis"
## [55] "Phryganea.grandis"           "Phryganea.bipunctata"
## [57] "Plectrocnemia.conspersa"     "Poycentropus.flavomaculatus"
## [59] "Potamophylax.latipennis"     "Psychomyia.pusilla"
## [61] "Rhyacophila.dorsalis"        "Rhyacophila.fasciata"
## [63] "Sericostoma.galeatum"        "Setodes.punctatus"
## [65] "Silo.nigricornis"            "Stenophylax.permistus"
## [67] "Trichostegia.minor"          "Tinodes.waeneri"

```

```
summary(site)
```

```

## BEAR.AM BEAR.AV CHAN.AV ENIL.AV FOUR.AV LUCE.AM LUIS.AV MOIR.AM  RHONE
##      6      9      15      6      15      9      9      15      9

```

```
summary(type)
```

```

## eupo para plesio river
##   30   33   21    9

```

```
summary(periode)
```

```

## P0 P1 P2 P3 P4 P5 P6 P7 P8
##  7  7  7 12 12 12 12 12 12

```

Ces données faunistiques sont issues d'un suivi bi-annuel dans le secteur de Belley de la zone alluviale du Rhône après restauration de trichoptères capturés la nuit par piégeage lumineux. Ces pièges lumineux sont installés sur un support ou suspendus à un arbre à quelques mètres du milieu étudié. En son centre, un tube néon produit de la lumière noire qui permet d'attirer les trichoptères par phototropisme. Le piège est constitué de 3 parois en plexiglass transparente situées autour de ce tube, visant à intercepter leur vol et à provoquer leur chute dans un bidon contenant 70% d'alcool. Un détecteur de luminosité permet d'allumer et d'atteindre le néon.

Chaque nuit de piégeage a eu lieu en lune noire afin de diminuer l'effet de la lune sur l'efficacité du piège. L'échantillonnage par piège lumineux se révèle efficace lorsque la température minimale ne descend pas en dessous de 10°C. Ainsi, la période d'échantillonnage s'étend de Mars/Avril à Octobre. Ce fichier contient 93 échantillonnages de trichoptères. Chaque site comprend entre 6 et 9 échantillonnages annuel (période de P0 à P8). Certains sites ont été suivi sur 2 années consécutives et comportent ainsi 15 échantillonnages.

Les adultes de trichoptères ont ensuite été triés puis les mâles ont été déterminés à l'espèce. Les femelles n'ont pas été prises en compte du fait qu'il soit difficile pour certains genre d'atteindre un tel niveau d'identification.

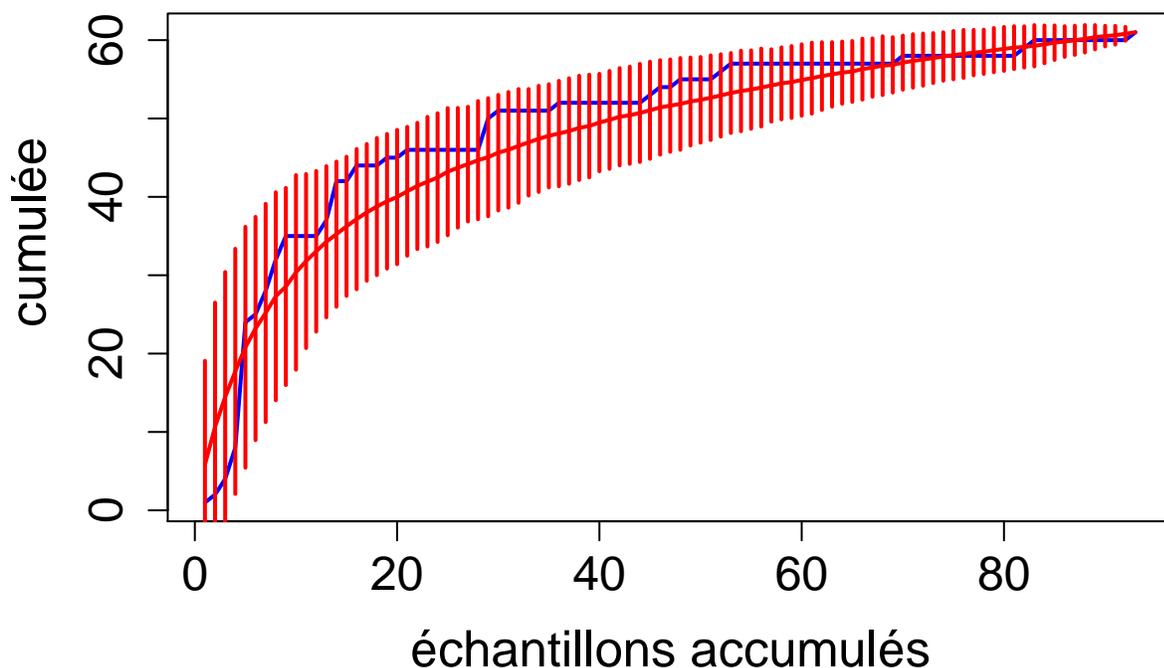
Trois grand types de milieux sont considérés:

- {eupo} - eupotamons -> stations connectés au Rhône en débit moyen (30 éch.)
- {para} - parapotamons -> stations déconnectés à l'amont mais connectés à l'aval en débit moyen (33 éch.)
- {plesio} - plesiopotamons -> stations déconnectés du Rhône en débit moyen (21 éch)
- {river} - chenal principal du Rhône (9 éch.)

## Richesse et diversité alpha

Tout d'abord, on cherche à vérifier la qualité des échantillonnages effectués en construisant une courbe d'accumulation.

```
collector <- specaccum(Tr[,8:ncol(Tr)], method="collector")
plot(collector,col="blue",lwd=2,xlab="échantillons accumulés",ylab="richesse
      cumulée",cex.axis=1.5,cex.lab=1.6)
random <- specaccum(Tr[,8:ncol(Tr)], method="random")
plot(random,add=TRUE, col="red",lwd=2,xlab="",ylab="",cex.axis=1.5,cex.lab=1.6)
```



```
specpool(Tr[,8:ncol(Tr)])
```

```
## Species      chao chao.se    jack1 jack1.se    jack2      boot boot.se
## All         61 74.9319 10.29582 73.86022 3.834114 80.77349 66.92823 2.352567
##           n
## All 93
```

On voit ici comment les richesses par échantillonnage s'accumulent et conduisent à la richesse totale du secteur formant un plateau autour de 62-63 espèces. Ce principe est à la base de techniques d'extrapolation ou d'estimation de la richesse. En consultant la fonction `specpool()`, on a une estimation de la richesse totale à partir du tableau actuel.

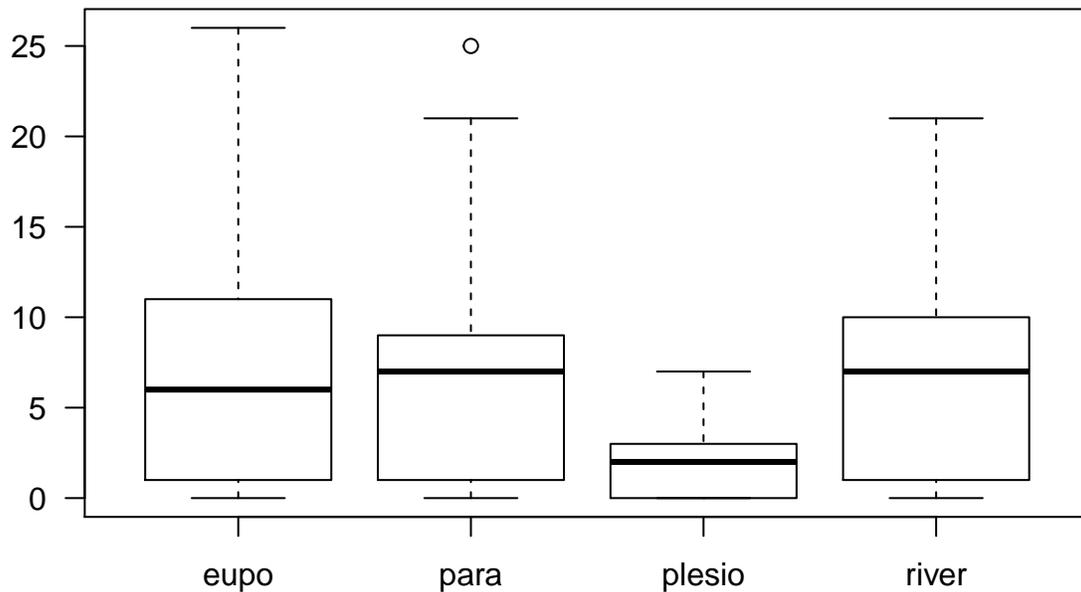
L'échantillonnage étant bon (environ 90% des espèces totales ont été capturées), entrons maintenant dans l'analyse des données. `Specnumber()` calcule le nombre de cellules non nulles par ligne ou par colonnes. Il s'agit donc ici de la richesse par échantillons. Elle varie de 0 à 26.

```
ri1 <- specnumber(Tr[,-c(1:7)]) # calcul des richesses
min(ri1) ; max(ri1) # min et max des richesses
```

```
## [1] 0
```

```
## [1] 26
```

```
boxplot(ri1~type,las=1) # boxplot
```



```
pairwise.wilcox.test(ri1,type,p.adj="none",exact=F) # test de Wilcoxon
```

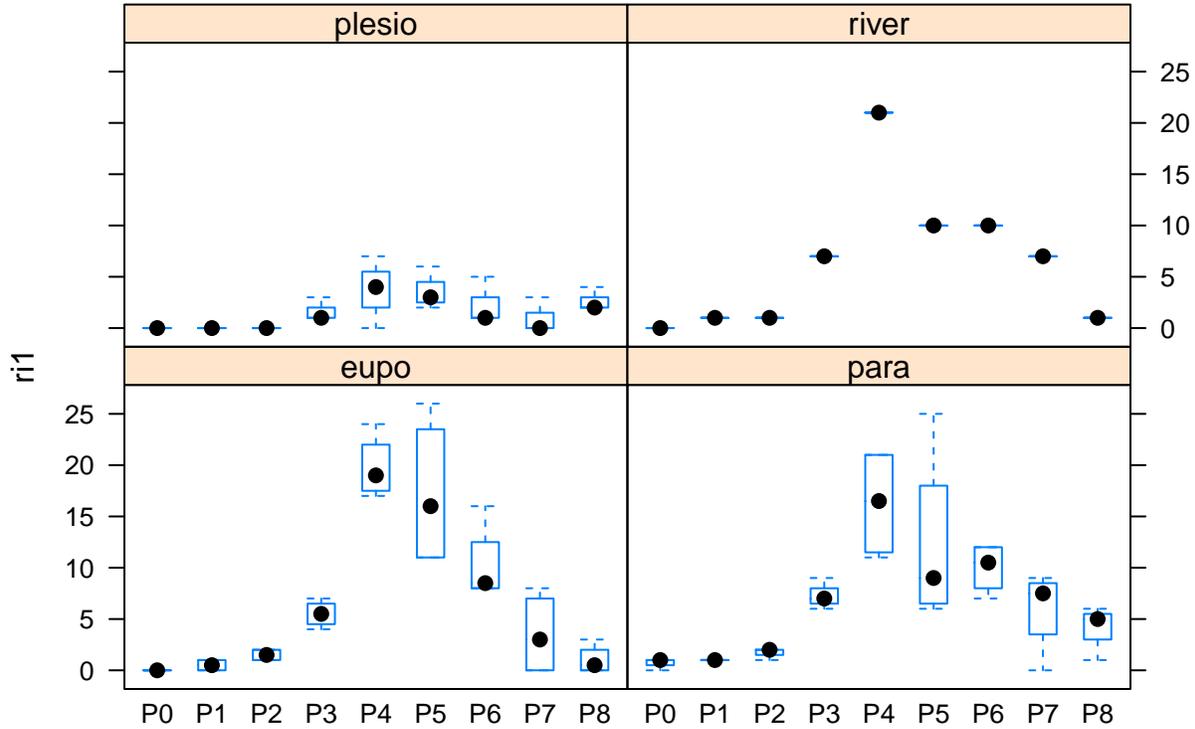
```
##
## Pairwise comparisons using Wilcoxon rank sum test
##
## data: ri1 and type
##
##      eupo   para   plesio
## para 0.74524 -      -
## plesio 0.01692 0.00078 -
## river 0.91981 0.69874 0.09842
##
## P value adjustment method: none
```

Si l'on regarde la médiane, le type plesiopotamon semble avoir une richesse légèrement inférieure aux autres types de milieu. La richesse du type plesiopotamon est significativement différente des deux autres types eupotamon et parapotamon mais légèrement non significative concernant la richesse du type river.

Ce graphique cache le fait que ces types de milieux ont été échantillonnés à différentes périodes de l'année. Le

package “lattice” permet de réaliser des graphiques conditionnées par un ou plusieurs facteurs, ici “type”.

```
bwplot(ri1~periode|type) #aussi possible avec /site
```



Le maximum de richesse est atteint en période P4 et P5, ces périodes correspondent aux dates d'échantillonnages des mois de Juin et de Juillet. Concernant le type plesio, on remarque un léger rebond de la richesse lors de la période P8 (Octobre).

## Diversité beta

La diversité beta peut être calculée selon le rapport entre la richesse gamma et la moyenne des richesses alpha mesurées au sein d'un secteur. C'est la mesure dite de Whittaker.

Dans notre cas, le calcul de la diversité beta sera une mesure de dissimilarité entre les assemblages de trichoptères récoltés dans nos échantillons par site ou type de milieu.

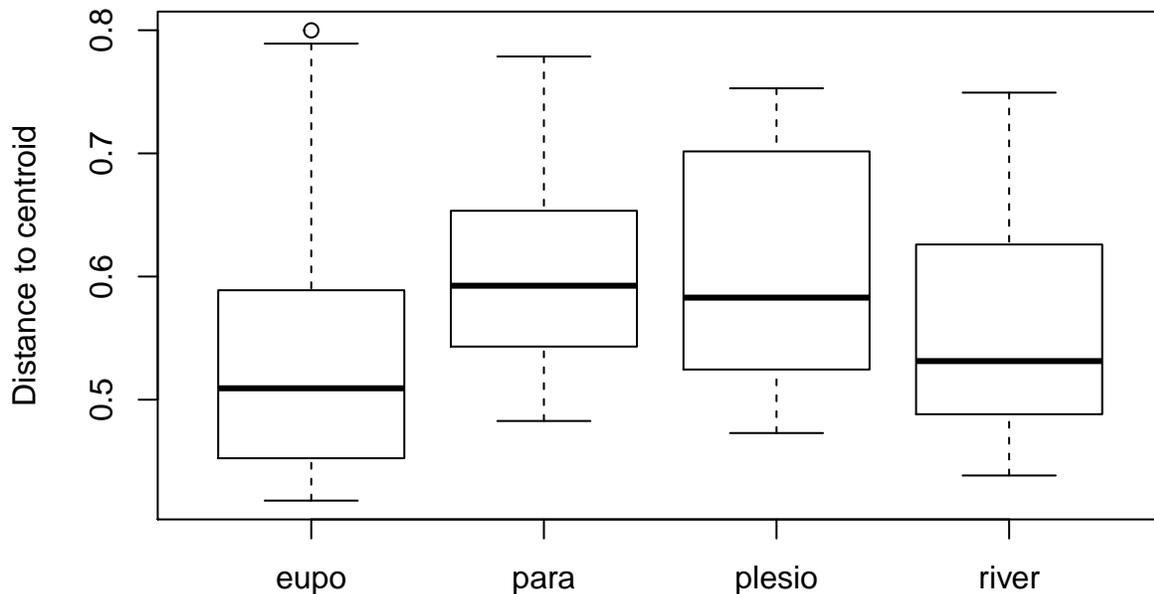
Tout d'abord, nous allons transformer notre tableau d'abondance en tableau de Présence/Absence.

```
Tr_PA <- ifelse(Tr[,-c(1:7)]>0,1,0) # transformation en données P/A
Tr_PA <- cbind.data.frame(type=Tr$type, Tr_PA)

dist1 <- vegdist(Tr_PA[,-1], "jaccard", binary = T) # matrice dissimilarité

disp<-betadisper(dist1,Tr_PA$type) # inter-sites

## missing observations due to 'd' removed
boxplot(disp)
```



```
anova(disp)

## Analysis of Variance Table
##
## Response: Distances
##          Df Sum Sq Mean Sq F value Pr(>F)
## Groups    3  0.07919  0.026396   2.6127 0.05756 .
## Residuals 74  0.74761  0.010103
```

```
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(dis)
```

```
## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##  
## Fit: aov(formula = distances ~ group, data = df)  
##  
## $group  
##          diff          lwr          upr          p adj  
## para-eupo  0.069721921 -0.002497303  0.14194115  0.0623654  
## plesio-eupo 0.067314376 -0.020364028  0.15499278  0.1907504  
## river-eupo  0.018109651 -0.090328554  0.12654786  0.9715182  
## plesio-para -0.002407545 -0.085075786  0.08026070  0.9998377  
## river-para  -0.051612270 -0.156041097  0.05281656  0.5664707  
## river-plesio -0.049204725 -0.164864856  0.06645541  0.6795525
```

L'objet `dist1` est une matrice de dissimilarité. Cette matrice est carrée et symétrique. Une telle matrice est limitée sous cette forme.

La fonction `betadisper` compare les valeurs de distance entre groupe de sites\*années (en fonction des rownames du tableau).

Dans notre cas, il existe un changement faiblement non significatif ( $p=0.058$ ) de diversité beta pour les 4 milieux considérés (voir résultat de l'anova), la diversité beta semble être plus élevée pour les parapotamons et les plésiopotamons.

Le Test de Tukey ne montre aucune différence significative entre nos groupes même si le boxplot montre que la déconnexion des milieux annexes semble entrainer leur diversification.

**Question 1: Que montre-t-on ici ? Quels processus entraineraient une plus grande diversité dans les parapotamons et les plésiopotamons ? Et à l'inverse, quel processus entraineraient une plus faible diversité dans les eupotamon et dans le cours principal ?**

## Qualité écologique d'un site

Il arrive parfois que la diversité à prendre en compte est au-delà des données disponibles, dans ce cas il est nécessaire d'identifier des espèces indicatrices représentant la diversité de groupes représentatifs d'une diversité d'habitats reconnus en zone alluviale.

Dufresne et Legendre (1997) proposent une définition de la valeur indicatrice d'une espèce pour les groupes d'une typologie, comme étant le produit de a) le rapport de l'abondance moyenne de l'espèce dans les sites du groupe considéré et de son abondance moyenne dans l'ensemble des relevés, et de b) la fréquence de l'espace dans les relevés du groupe considéré. Le calcul de cette valeur incorpore donc une composante liée à l'abondance des espèces et une liée à la fréquence des espèces. Le calcul de ces valeurs pour une typologie données s'accompagne de tests de permutation permettant d'accompagner les valeurs indicatrices d'une significativité. Le tout est réalisé par la fonction `multipatt` du package `indicspecies`.

```
vi <- multipatt(Tr[, -c(1:7)], as.factor(Tr$type), control= how(nperm=999))
summary(vi)
```

```
##
## Multilevel pattern analysis
## -----
##
## Association function: IndVal.g
## Significance level (alpha): 0.05
##
## Total number of species: 61
## Selected number of species: 12
## Number of species associated to 1 group: 6
## Number of species associated to 2 groups: 4
## Number of species associated to 3 groups: 2
##
## List of species associated to each combination:
##
## Group eupo #sps. 2
##               stat p.value
## Agapetus.ochripes 0.600  0.023 *
## Lype.phaeopa      0.436  0.020 *
##
## Group para #sps. 1
##               stat p.value
## Mesophylax.impunctatus 0.571  0.009 **
##
## Group river #sps. 3
##               stat p.value
## Oxyethira.flavicornis 0.565  0.002 **
## Ceraclea.aurea       0.507  0.026 *
## Agraylea.multipunctata 0.471  0.012 *
##
## Group eupo+para #sps. 1
##               stat p.value
## Mystacides.azurea 0.598  0.022 *
##
## Group eupo+river #sps. 1
##               stat p.value
## Psychomyia.pusilla 0.757  0.005 **
##
```

```

## Group para+plesio #sps. 1
##                stat p.value
## Limnephilus.lunatus 0.43  0.046 *
##
## Group para+river #sps. 1
##                stat p.value
## Silo.nigricornis 0.488  0.049 *
##
## Group eupo+para+river #sps. 2
##                stat p.value
## Hydroptila.sparsa 0.612  0.011 *
## Goera.pilosa      0.589  0.028 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

On remarque a) que certains groupes peuvent être mieux caractérisés que d'autres et b) que les valeurs indicatrices sont également calculées pour des ensembles de groupes.

**Question 2: Il peut être commun de se demander si ce partitionnement de sites (plésio, para, eupo, river) en zone alluviale convient le mieux étant donné que le groupe plésio ne contient aucune espèces indicatrices ? Le type plésio serait encore à sub-diviser ou bien à regrouper avec le type para ?**

Si l'on regarde attentivement dans le tableau, on remarque que ces espèces sont relativement courantes (sauf pour *Agraylea multipunctata*). Mais la qualité écologique d'un site se reflète aussi par sa capacité à accueillir des espèces rares. Essayons déjà de trouver les espèces peu courantes dans notre jeu de données.

Pour cela, nous allons utiliser la fonction `restrictedness` du package `funrar`. Cette fonction mesure la rareté d'une espèce à l'échelle régionale. Chaque espèce aura une valeur de rareté par rapport au pool d'espèces échantillonnées. Cette valeur de `restrictedness` est comprise entre 0 (espèce présente dans tous les sites) et 1 (espèce très rare).

Après avoir réordonné le tableau (de 1 -> 0), on peut identifier les espèces les plus rares dans notre cortège mais il reste difficile de leur attribuer des "classes" de protection pour guider certaines politiques de conservation par exemple. Une multitude de méthodes d'attribution de classes peuvent être utilisées, ici nous allons donc utiliser la méthode rapide par la normalisation de ces valeurs de rareté c'est à dire que nos valeurs de rareté vont être normalisées entre -1 et 1 permettant de diviser notre pool espèces en 2 groupes (rares vs communes). Ensuite, on appliquera un code couleur qui identifie sur le graphe en bleu les espèces dont la valeur normalisée est supérieure à 0 que l'on considérera comme espèces rares.

```

Tr_rn <- as.matrix(Tr[,-c(1:7)]) # transformation en matrice
rownames(Tr_rn) <- paste(Tr$site, Tr$annee, Tr$periode, sep="_")

ri <- restrictedness(Tr_rn)
ri_df <- as.data.frame(ri)

ri_df <- ri_df[order(ri_df$Ri),] # ordonner Ri du + grand au plus petit
# conversion en facteur pour retenir l'ordre établi plus haut
ri_df$species <- factor(ri_df$species , levels = ri_df$species )

# normalisation des indices de rareté entre -1 et 1
ri_df$group <- round((ri_df$Ri - mean(ri_df$Ri))/sd(ri_df$Ri), 2)
# ajout de la colonne Communes/Rares
ri_df$group <- ifelse(ri_df$group < 0, "Communes", "Rares")

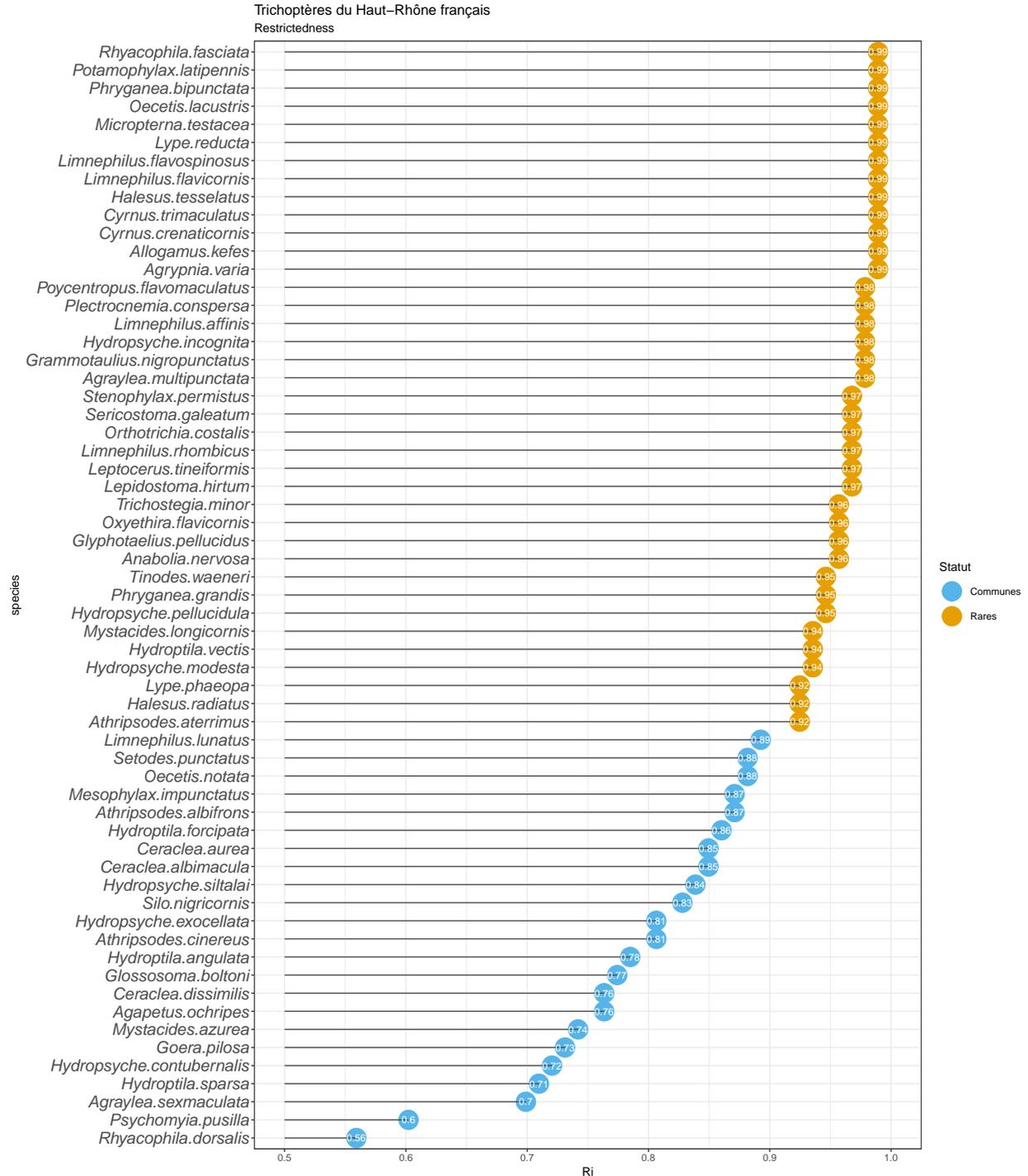
theme_set(theme_bw())

```

```

ggplot(ri_df, aes(x=species, y=Ri, label=round(Ri,2))) +
  geom_point(stat='identity', aes(col=ri_df$group), size=8) +
  geom_segment(aes(y = 0.5,
                  x = species,
                  yend = Ri,
                  xend = species),
              color = "grey30") +
  scale_color_manual(name="Statut",
                    labels = c("Communes", "Rares"),
                    values = c("Communes"="#56B4E9", "Rares"="#E69F00")) +
  geom_text(color="white", size=3) +
  labs(title="Trichoptères du Haut-Rhône français",
       subtitle="Restrictedness", size=4) +
  ylim(0.5, 1) +
  coord_flip()+
  theme(axis.text.y=element_text(size=15, face = "italic"))

```



On remarque que les espèces rares constituent 62 % du cortège total.

**Question 3 : Que penser de ce résultats ? Notre cortège d'espèce répond-t-il bien à la diversité des habitats disponibles en zone alluviale ? Selon vous, quels sont les sites qui accueilleraient potentiellement le plus d'espèces considérées ici comme rares ?**

Après avoir extrait les liste des espèces rares et communes, c'est le moment d'utiliser une boucle dans r ! Cette boucle va permettre de répéter la même opération qui est de compter et de coller dans un tableau le nombre d'espèces qui appartient à chacune de nos 2 catégories (espèces rares vs communes) et pour cela

dans chacun site. Ensuite, après une bref manipulations pour construire les étiquettes du graphique, nous construirons un barplot pour une meilleure visualisation.

```
Tr_sites <- aggregate(.~site, Tr[,c(1,8:ncol(Tr))], sum) # somme des sp par site

# extraction des espèces
rare_sp <- ri_df[ri_df$group=="Rares",][1]
rare_sp <- as.vector(rare_sp[,1])
comm_sp <- ri_df[ri_df$group=="Communes",][1]
comm_sp <- as.vector(comm_sp[,1])

# Transposer Tr en format long
Tr_sites_long <- gather(Tr_sites, species, abundance, Agapetus.ochripes:Tinodes.waeneri)

# suppression des lignes dont l'abondance = 0
Tr_sites_long = Tr_sites_long[which(Tr_sites_long$abundance > 0), ]

# Création d'une matrice de 2 lignes et 9 colonnes
mat <- as.data.frame(matrix(nrow=9, ncol=4))
colnames(mat) <- c("site", "type", "rare_sp", "comm_sp")

# extraction des noms des sites
sites <- levels(as.factor(Tr_sites_long$site))

# on remplit la matrice avec ce que l'on connaît déjà
mat[, "site"] <- sites
mat[, "type"] <- c("plesio", "para", "eupo", "eupo", "para",
                  "eupo", "para", "plesio", "river")

# calcul par site du nombre d'espèces rares et communes
for (i in 1:length(sites)){
  name <- sites[i]
  df <- as.vector(Tr_sites_long[which(Tr_sites_long==name), "species"])

  mat[i, "rare_sp"] <- length(df[df %in% rare_sp])
  mat[i, "comm_sp"] <- length(df[df %in% comm_sp])
}

# on bascule la matrice de résultats en format long
mat_long <- melt(mat, id.vars=c("site", "type"))

# colonnes d'étiquettes
# Calculer la somme cumulée des espèces pour chaque site
mat_long$variable = factor(mat_long$variable, levels=c("comm_sp", "rare_sp"))
mat_long <- as.data.frame(mat_long[order(mat_long[,1], mat_long[,3]),])

mat_long <- ddply(mat_long, "site",
                  transform, label_ypos=cumsum(value))

# on construit le graphe
p<-ggplot(data=mat_long, aes(x=site, y=value, fill=factor(variable,
                  levels=c("rare_sp", "comm_sp")))) +
```

```

geom_bar(stat="identity", color="black")
p <- p + geom_text(aes(y=label_ypos, label=value), vjust=1.6,
                    color="white", size=3.5)
p <- p + scale_fill_grey() + theme_minimal()
p <- p + scale_fill_manual(values=c("#E69F00", "#56B4E9"))

```

```

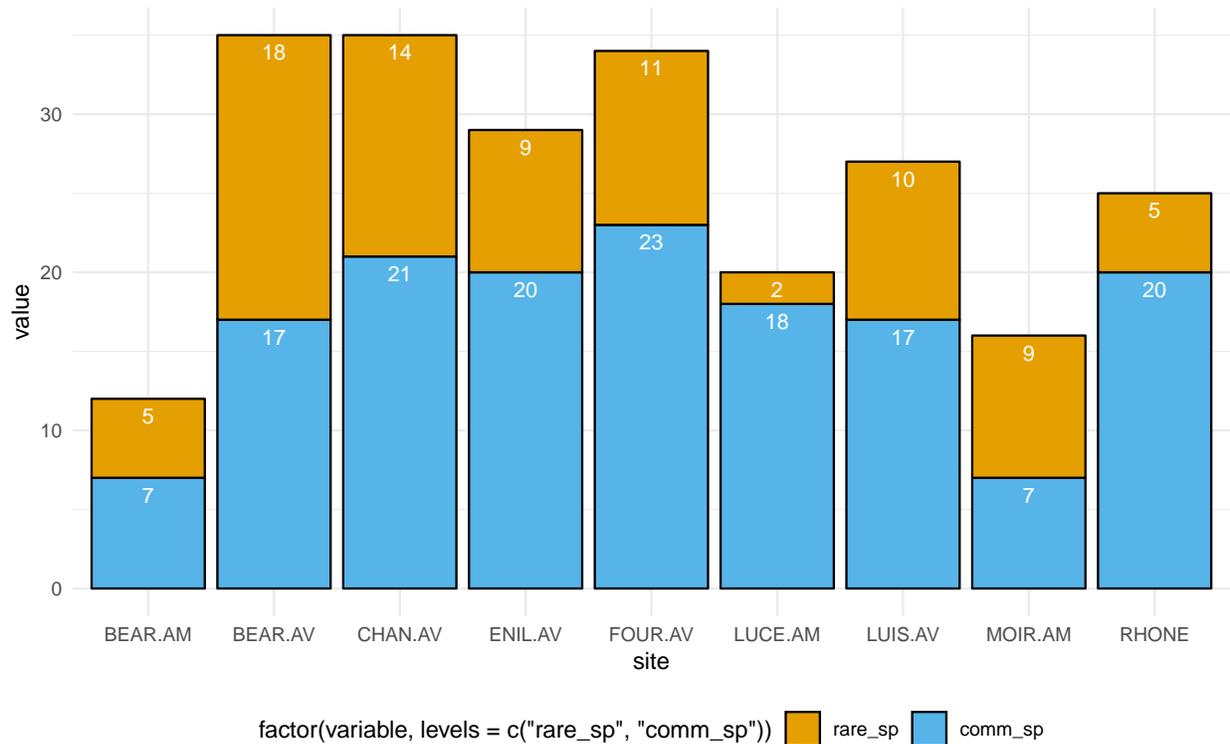
## Scale for 'fill' is already present. Adding another scale for 'fill',
## which will replace the existing scale.

```

```

p <- p + theme(legend.position="bottom")
p

```



En plus d'avoir le plus grand nombre d'espèces total (35), le site BEAR\_AV contient le plus grand nombre d'espèces que nous avons considérées comme rares (18). Au contraire BEAR\_AM est le mauvaise élève et avec son plus petit nombre d'espèces (12). LUCE\_AM obtient le plus petit nombre d'espèces rares (seulement 2).

## Quelle est l'origine de la diversité en milieu alluvial ?

Ici, on s'intéresse à la recherche de relation entre la diversité inter-milieux et des caractéristiques mésologiques dont on peut disposer.

### Diversité beta et distances géographiques

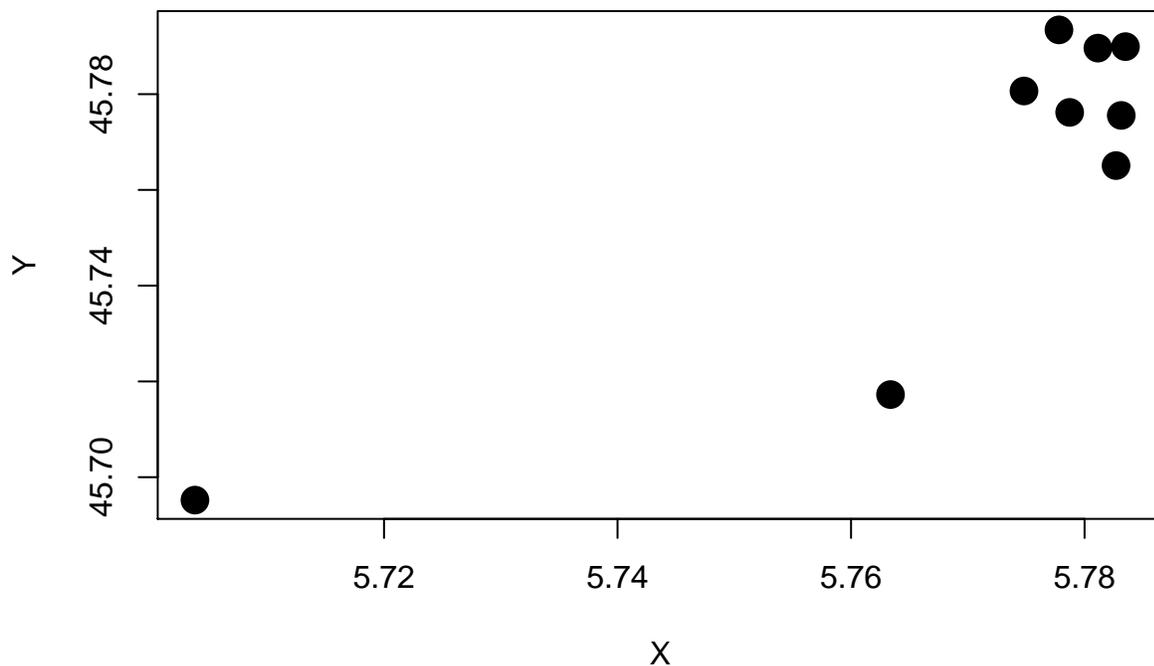
Les trichoptères ayant un stade aérien libre, regardons si la dissimilarité faunistique entre site augmente avec leur distance spatiale..

Les distances géographiques (spatdist) et "faunistiques" (dist1 = diversité beta) sont elles corrélées ?

Commençons d'abord par produire un fichier additionnant les échantillons par station (Tr\_sites) puis transformons le en tableau de données Présence/Absence (Tr\_sites\_PA).

```
Tr_sites_PA <- ifelse(Tr_sites[, -c(1:7)] > 0, 1, 0) # transformation en données P/A
```

```
coord <- read.table("gps.txt", header=TRUE)
coord <- coord[order(rownames(coord)), ] # coordonnées GPS
plot(coord, pch=16, cex=2) # plot des coordonnées GPS
```



```
spatdist <- vegdist(coord, method="euclidean") # matrice de distances géographiques
dist1 <- vegdist(Tr_sites_PA, "jaccard", binary = T) # matrice de dissimilarité invertébrés
mantel(spatdist, dist1)
```

```
##
## Mantel statistic based on Pearson's product-moment correlation
##
## Call:
## mantel(xdis = spatdist, ydis = dist1)
##
## Mantel statistic r: -0.1186
##      Significance: 0.492
##
## Upper quantiles of permutations (null model):
##   90%   95% 97.5%   99%
## 0.402 0.505 0.760 0.784
## Permutation: free
## Number of permutations: 999
```

La réponse est non ! Le test de Mantel qui teste la corrélation entre deux matrices de distance indique que la corrélation est faible ( $r=-0.11$ ) et non-significative ( $p=0.48$ ).

## Richesse et hydrologie

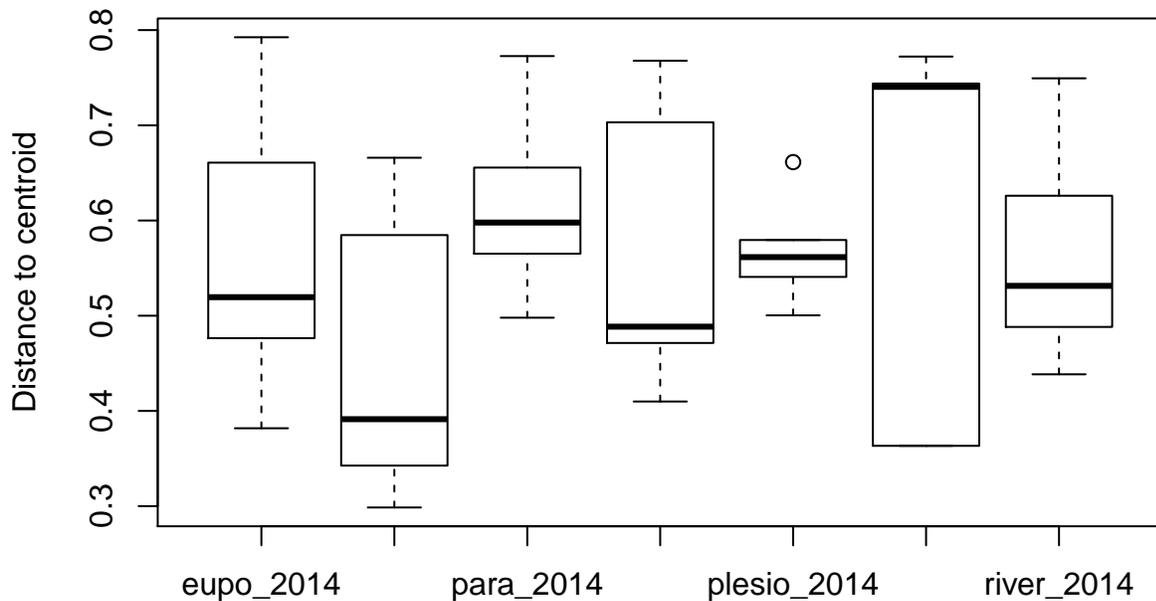
Les échantillonnages ont eu lieu sur deux années hydrologiquement différentes. On remarque sur le boxplot que la médiane peut-être très différente d'une année à l'autre sur un même type de milieu. D'ailleurs le test d'analyse de variance montre une différence presque significative ( $p=0.06$ ) entre nos sites\*années.

```
Tr_PA <- ifelse(Tr[,-c(1:7)]>0,1,0) # transformation en données P/A
Tr_PA <- cbind.data.frame(type_annee=paste(Tr$type,Tr$annee,sep="_"), Tr_PA)

dist1 <- vegdist(Tr_PA[,-1], "jaccard", binary = T) # matrice dissimilarité

disp<-betadisper(dist1,Tr_PA$type_annee) # inter-sites

## missing observations due to 'd' removed
boxplot(disp)
```



```
anova(dis)
```

```
## Analysis of Variance Table
##
## Response: Distances
##           Df Sum Sq Mean Sq F value Pr(>F)
## Groups      6 0.1755  0.029250  2.1285 0.06049 .
## Residuals  71 0.9757  0.013742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Avant de commencer, est-il problématique de mesurer des richesses sur des échantillons dont les abondances sont très différentes ?

Le fait de prélever plus d'individus tendrait mécaniquement à faire augmenter la richesse, non ? Vérifions cette hypothèse en logant les abondances afin de mettre en évidence une relation linéaire.

On mène cette exploration à l'échelle des échantillons mais aussi à l'échelle des stations\*années.

```
# somme des abondances par ligne (/échantillons)
ab <- apply(Tr[,-c(1:7)],1,sum)

# somme des abondances par sp par site/année
Tr_sites_annee <- aggregate(.~site+annee, Tr[,c(1,8:ncol(Tr))], sum)

# somme des abondances par sites années
ab_site <- apply(Tr_sites_annee[,-c(1:2)],1,sum)
```

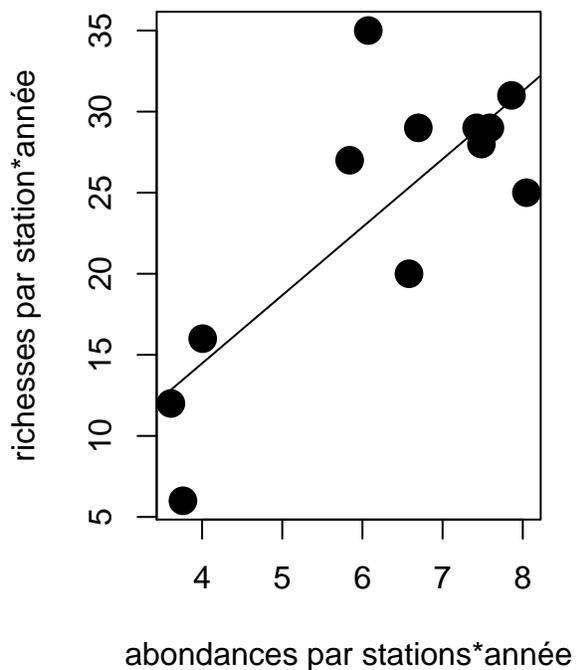
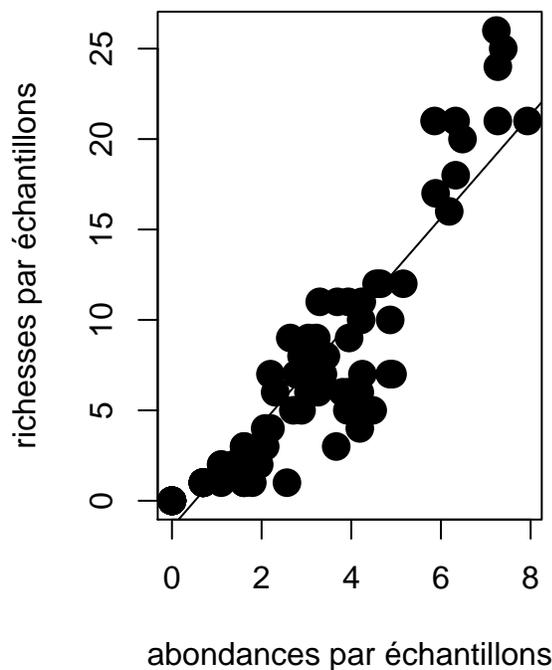
```

# calcul de la richesse par site année
ri2 <- specnumber(Tr_sites_annee[,-c(1:2)])

#ri1 a déjà été calculé précédemment

par(mfrow=c(1,2))
plot(log(1+ab), ri1, pch=16,cex=2,xlab="abondances par échantillons",
     ylab="richesses par échantillons",cex.axis=1,cex.lab=1)
abline(lm(ri1~log(1+ab))) # on ajoute la ligne de tendance
plot(log(1+ab_site), ri2, pch=16,cex=2,xlab="abondances par stations*année",
     ylab="richesses par station*année",cex.axis=1,cex.lab=1)
abline(lm(ri2~log(1+ab_site))) # on ajoute la ligne de tendance

```



Hypothèse vérifiée ! On discutera des conséquences de cette relation souvent ignorée ! Elle a conduit au développement de la richesse dite “raréfiée”: une mesure de la richesse standardisée par rapport à une abondance commune entre tous les échantillons comparés. Calculons la par station\*années (toujours avec vegan).

Par chance, la compagnie qui gère le barrage en amont vous fournit les durées de connection de vos sites en fonction des crues survenues en 2014 et 2015.

```

min(ab_site) # abondance minimale dans tous nos sites

```

```
## [1] 36
```

```

# calcul de la richesse raréfiée par site*année + dataframe
rirar1<-as.data.frame(rarefy(Tr_sites_annee[,-c(1:2)],35,se=FALSE,MARGIN=1))
# noms de lignes

```

```

rownames(rirar1) <- paste(Tr_sites_annee$site, Tr_sites_annee$annee, sep="_")
colnames(rirar1) <- "x"

# importation des données de connection hydrologique
connecc<- read.table("connection.txt",header=TRUE)

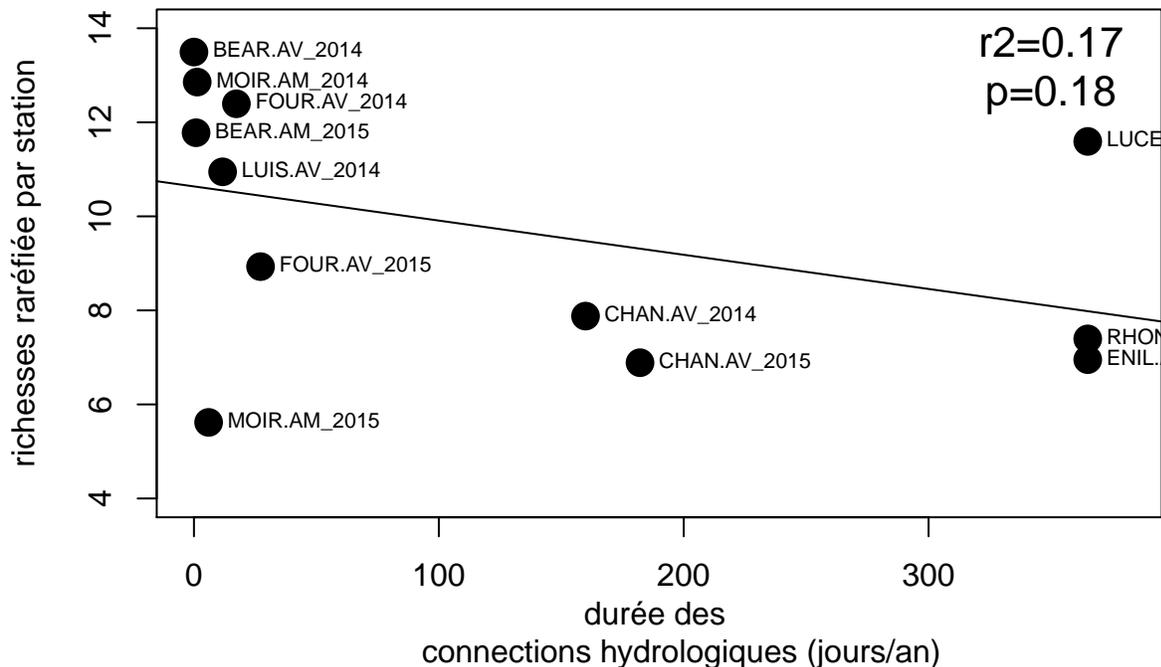
# on assemble les deux tableaux
df2 <- merge(as.data.frame(rirar1),connecc,by="row.names",all.x=TRUE)

# calcul du R2
r2 <- summary(lm(df2$x~df2$sum_of_connection))$r.squared

plot(df2$sum_of_connection, df2$x, pch=16,cex=2,xlab="durée des
      connections hydrologiques (jours/an)"
      ,ylab="richesses raréfiée par station",cex.axis=1,cex.lab=1,
      ylim=c(4,14), xlim=c(0,380)) # on fixe les limites des axes pour la visibilité
text(df2$sum_of_connection, df2$x,labels=df2$Row.names, cex= 0.7, pos=4) # étiquettes
abline(lm(df2$x~df2$sum_of_connection)) # on ajoute la courbe de tendance
cor.test(df2$sum_of_connection, df2$x) # un test de corrélation pour la p-value

##
## Pearson's product-moment correlation
##
## data: df2$sum_of_connection and df2$x
## t = -1.453, df = 10, p-value = 0.1769
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.7997833 0.2056568
## sample estimates:
## cor
## -0.4175238
text(350,13.2,"r2=0.17\np=0.18",cex=1.3)

```



Les résultats produits (rirar1) comprennent pour chaque site\*année, le nombre moyen de taxons attendu dans un tirage aléatoire de 35 individus (< minimum d'abondances observés). Cette richesse rarifiée peut être accompagnée d'une erreur standard (si se=TRUE).

**Question 4: Pourquoi la richesse rarifiée tend à être plus élevée dans les milieux déconnectés ? Quelles hypothèses peut-on apporter afin d'expliquer une telle tendance ?**

## Phénologie

La phénologie devient alors une science dès lors que l'on commence à relever les dates auxquelles certaines phénomènes naturels se reproduisent chaque année ! Avez-vous déjà remarqué des éclosions massives d'insectes en bord de rivière ? ou bien une forte intensité de gobages de poissons un soir d'été ? La phénologie souhaite comprendre quelles relations existent entre ces périodes d'activités (émergences) et les variations saisonnières (hydrologie, température etc.).

Les échantillonnages effectués sur le Haut-Rhône sont en réalité des pseudo-répétitions liées géographiquement et écologiquement. Mais ces échantillonnages répétés relèvent également de la saison car ils sont liés temporellement. Ces données constituent une nouvelle pierre pour l'édifice de la connaissance des cycles phénologiques des espèces souvent très différents de l'une à l'autre. Par leur lien à la température de l'eau, température de l'air, à l'hygrométrie, à la lumière et à leur phénologie annuelle (e.g. diapause estivale, reproduction au printemps, été, automne), chaque espèce présente une spécificité. On considère alors que la capture d'un certain nombre d'individus à un stade biologique particulier (ici le stade adulte) au cours d'une année est du à la phénologie de l'espèce.

Nous allons maintenant construire pour chaque espèce un graphique décrivant sa période de capture (et donc sa période de vol) en 2014 et 2015. Dans un soucis de quantité suffisante de données, nous ne considérerons ici seulement les espèces communes.

```

# On transforme tout d'abord la table d'abondances en abondance relative
Tr_prop <- cbind.data.frame(Tr[,c(1:7)], prop.table(as.matrix(Tr[, -c(1:7)]), 2))

#colSums(Tr_prop[, -c(1:7)]) # on vérifie si la somme d'une colonne = 1

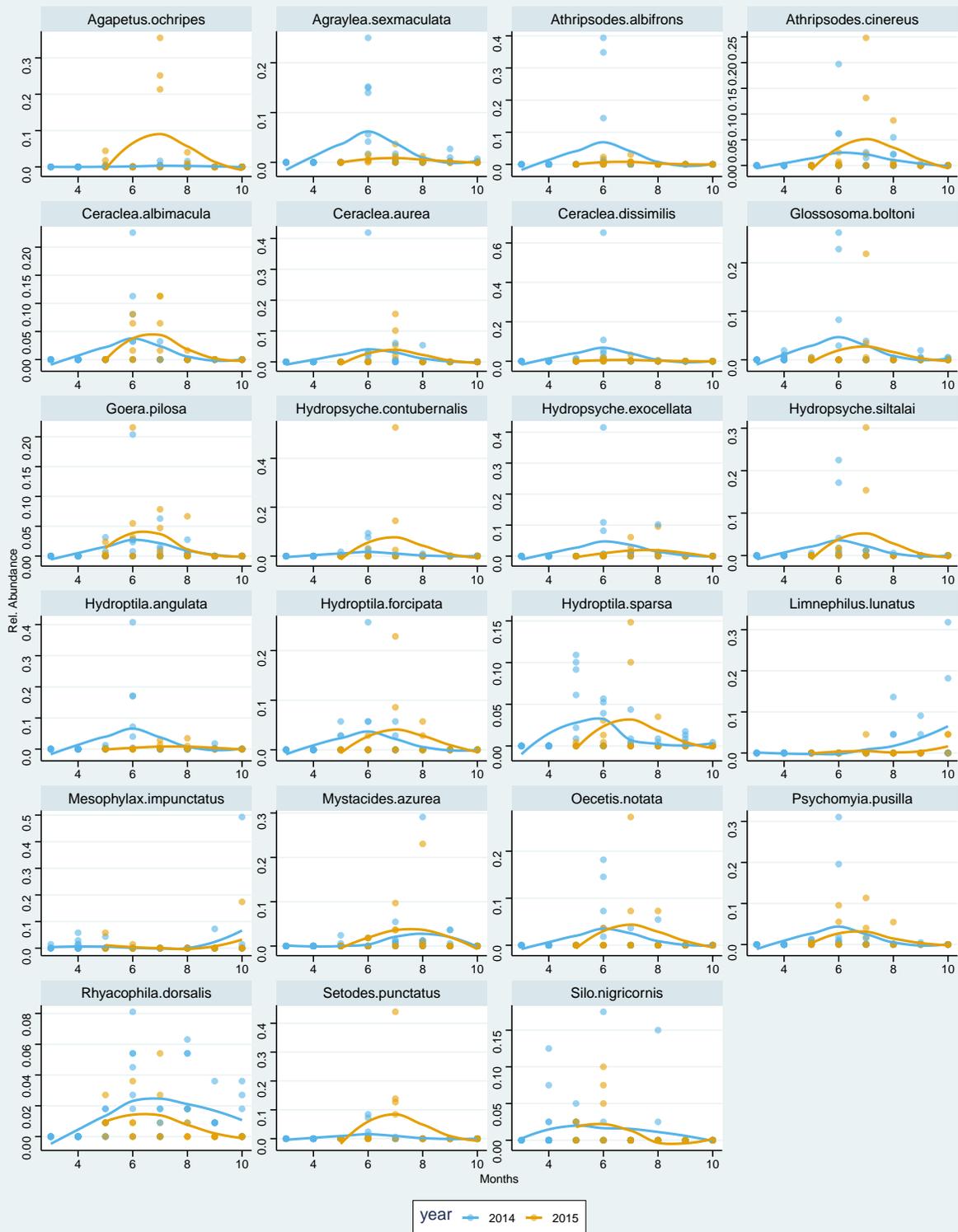
# On bascule en format long en sélectionnant les espèces communes
Tr_prop_long <- cbind.data.frame(Tr_prop[,c(1:7)], Tr_prop[, comm_sp])
Tr_prop_long <- gather(Tr_prop_long[, -c(5,6,7)], species, abundance,
                      comm_sp[1]:comm_sp[length(comm_sp)])

#paramétrage des colonnes temporelles
Tr_prop_long$date <- as.Date(Tr_prop_long$date, "%d.%m.%y")
Tr_prop_long$date <- as.POSIXct(Tr_prop_long$date, "%d.%m.%y")
# une colonne avec l'année
Tr_prop_long$year <- format(Tr_prop_long$date, format="%Y")
# une colonne avec le mois
Tr_prop_long$month <- strftime(Tr_prop_long$date, "%m")

library('ggthemes') # meilleur design pour le graphe qui suit

ggplot(data = Tr_prop_long, mapping = aes(x = as.numeric(month), y = abundance,
                                          color=year, fill=year)) +
geom_point(alpha=6/10, shape=21, size=2) +
  geom_smooth(method = 'loess', span = .9, se = FALSE) +
  xlab("Months") +
  ylab("Rel. Abundance") + scale_x_continuous(limits=c(3,10)) +
  #scale_x_date(labels = date_format("%b"), breaks = date_breaks("4 months")) +
  facet_wrap(~ species, ncol=4, scales="free")+ theme(axis.line=element_line()) +
  theme_stata() +
  scale_fill_manual(values=c("#56B4E9", "#E69F00")) +
  scale_color_manual(values=c("#56B4E9", "#E69F00"))

```



On remarque des différences intéressantes de phénologie entre espèces qui peuvent être groupées selon: les espèces à émergences printanière-automnales (e.g. *Mesophylax impunctatus*), printanières-estivales (e.g. *Goera pilosa*, *Hydropsyche contubernalis*, *Setodes punctatus*) et uniquement automnales (e.g. *Limnephilus*

*lunatus*). Il est difficile de classer *Rhyacophila dorsalis* car il semble avoir une période d'émergence très étalée cependant ces maxima surviennent plutôt en période estivale. Ce type de graphique peut permettre d'adapter un échantillonnage afin d'améliorer le ratio couts-bénéfices d'une étude en diminuant la fréquence d'échantillonnage et en privilégiant certaines périodes.

**Question 5: Quelles sont les périodes de captures à privilégier avec ce type d'échantillonnage ?**

Ces cycles phénologiques sont intimement liées au biotope où vit l'espèce. Par exemple, *Mesophylax impunctatus* est une espèce typique des milieux annexes à l'état larvaire, elle effectue cependant une diapause estivale dans des milieux à hygrométrie plus élevée notamment dans des grottes. Le champ d'étude de la phénologie en milieu aquatique consiste à étudier les facteurs qui ont influencés son évolution. On remarque des différences phénologiques notables d'une année à l'autre. A ce stade des données nous pouvons seulement émettre des hypothèses quant aux effets 1) des crues dans certains habitats et 2) des variations de températures inter-annuelles (année "chaude" vs année "fraîche") voire de l'hygrométrie et de la luminosité. Ce travail relèverait de la modélisation sur des données recouvrant une période bien plus grande dans le but d'intégrer un maximum d'épisodes de crues et de chroniques de températures, hygrométrie, luminosité à mettre en adéquation avec les émergences d'espèces observées mais ce n'est pas l'objectif aujourd'hui.

## Synthèse

A partir de données type espèces x sites/dates, “mesurer de la diversité” recouvre une grande “diversité” d’actions possibles:

- Tout d’abord, on teste la qualité de notre échantillonnage en utilisant les richesses cumulées par ajout d’échantillons. Cette technique permet par extrapolation une estimation de la richesse totale (qui peut être vu comme une estimation de la diversité régionale).
- On s’intéresse à la diversité interne des échantillons, c’est à dire la richesse qui est la mesure la plus simple de cette diversité. Ces mesures de richesses peuvent être 1) regroupées par stations et par grands types de milieu afin d’effectuer une comparaison inter-sites mais aussi 2) peut être suivie au cours du temps afin de mieux percevoir les changements de richesses liées aux conditions environnementales et à la phénologie des espèces.
- Nous avons étendus notre échelle d’étude en s’intéressant à leur dissimilarités (= leur distance) c’est-à-dire à la diversité beta de l’ensemble des relevés. Parmi une multitude de mesures de distances possibles, a été explorée la distance de Jaccard sur des données en présence/absence.
- La richesse et la diversité des sites ne sont pas les seuls critères pouvant être pris en compte pour estimer la valeur écologique d’un site. L’aspect “espèces indicatrices” a été abordée mais celui aussi des “espèces rares” que peuvent abriter un site d’étude pouvant faire augmenter sa valeur écologique.
- Les distances et richesses calculées se prêtent à de nombreuses formes de représentations et d’analyses en fonction des objectifs de l’étude. Celles-ci peuvent être utilisées à titre de comparaison inter-sites ou inter-milieus mais elles peuvent être aussi mises en relation avec des variables environnementales (hydrologie) et spatiales (distances géographiques).
- Enfin, la diversité écologique d’une zone alluviale est fortement liée à la diversité des cycles phénologiques que représente un pool d’espèces. Nous avons démontré qu’une mesure de la diversité n’est en aucun cas figée dans le temps et que le fait d’obtenir des informations à l’échelle spécifique permet 1) d’expliquer les variations de richesse et de diversité d’un secteur, 2) d’adapter les futurs échantillonnages et 3) d’augmenter les connaissances biologiques des espèces.

De plus en plus d’études incorporent la fonction des espèces dans les communautés (diversité fonctionnelle) ou leurs distances phylogénétiques (diversité phylogénétique). Les traits biologiques et écologiques disponibles sous forme de base de données pour un nombre croissant de groupes taxonomiques permettent de calculer des distances fonctionnelles entre taxons qui peuvent être mises en relation avec des variables environnementales.

Cette partie de l’analyse des communautés non abordée ici a un réel intérêt écologique du fait de la prise en compte de la proximité des espèces entre elles. En effet, deux espèces du même genre sont de toute évidence plus proches fonctionnellement et phylogénétiquement que deux espèces de familles différentes. Ainsi, vous avez pu découvrir que la diversité pouvait être mesurée et estimée de nombreuses façons mais beaucoup d’autres peuvent être trouvées dans la littérature, je vous laisse le choix d’appliquer celle qui conviendra aux objectifs de vos travaux.

Sur le plan pratique et pour conclure, il convient de souligner que la mesure de la diversité est ici réduite à un seul groupe biologique: les trichoptères. Les hypothèses sur lesquelles reposent l’étude de ces organismes n’est peut être pas applicable aux autres groupes (éphéméroptères, gastéropodes etc.), ceci constituant l’une des principales limites de son application. Quoiqu’il en soit et pour se rassurer, il faut savoir que la diversité de ce groupe est particulièrement menacée en zone alluviale et également qu’un inventaire de la biodiversité totale d’une grande rivière de plaine y compris dans sa dimension latérale est un objectif considéré comme hors d’atteinte !