

# Analyse des données oiseaux UBx Bordeaux Métropole

Frédéric Barraquand

Projet de statistique pour données environnementales / UE 4TMS817U – 2024

## 1 Contexte

L’objectif de la partie biodiversité de cette UE est double:

- vous faire découvrir la mesure de la (bio)diversité
- vous faire attaquer un problème d’analyse de données “non-balisé”, pour lequel on ne connaît pas encore la “bonne” réponse.

Depuis 2017, nous<sup>1</sup> nous intéressons à suivre la biodiversité à Bordeaux et alentours. Des comptages des différentes espèces d’oiseaux sont notamment réalisés par mon collègue Frédéric Revers (INRAE – unité BioGeCo), et ce sont ces données que nous allons analyser conjointement avec des données d’usage des sols produites par le bureau de télédétection I-sea (et traitées par la suite par mes soins).

Une diversité parfois insoupçonnée d’espèces sauvages sont observables en milieu urbain. De nombreuses études ont été réalisées sur les oiseaux et ont largement confirmé ce que nous savions dès les années 2000, à savoir que plusieurs courbes diversité =  $f(\text{imperméabilisation})$  sont possibles, en particulier chez les oiseaux, avec soit un déclin continu de la diversité à mesure que le pourcentage d’imperméabilisation des sols, soit un pic de diversité en milieu péri-urbain. Si ce pic existe, nous tenterons la seconde semaine de trouver son origine.

Les données oiseaux ont été collectées en 2018-2019 et 2022-2023 (le programme continue jusqu’en 2026). J’ai analysé les données 2018-2022 donc nous avons déjà une petite idée, à voir si elle se confirme avec la nouvelle année incluse.

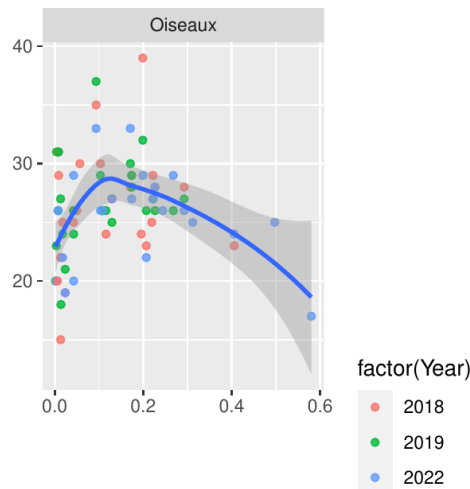


Figure 1: Nombre d’espèces en fonction du pourcentage d’imperméabilisation [fast screen grab of a larger image, axes labels will be added later on]

Les analyses faites ici concernent la mesure de la biodiversité, mais fondamentalement il s’agit d’un problème général et rien n’empêche d’appliquer le même genre de méthodes aux thématiques suivantes

<sup>1</sup>un consortium de chercheurs et associations naturalistes

- la mesure de la diversité des personnages dans un film
- la mesure de la diversité des produits dans un grand magasin
- la mesure de la diversité des auteurs dans une bibliothèque
- ...

## 2 Quelques références

Vous pourrez vous appuyer sur

- Les cours sur la mesure de la diversité d'Eric Marcon [version pdf](#) et [version html](#) (il y a là-dedans bien plus que ce dont vous avez besoin, mais si vous voulez des détails ils y seront).
- Un [article wikipedia fort bien fait et assez synthétique](#)
- Je joins également au dossier un article dans BioScience de McKinney qui résume ce que nous savons sur l'effet de l'urbanisation sur la biodiversité (il en existe de plus récents mais le contenu est assez similaire).
- et évidemment mes explications à l'oral...

## 3 Description des données

Les oiseaux sont comptés à la vue et au chant dans un rayon de 100m (cf. protocole), 2 fois par an et par site. Certains sites sont nos témoins de l'évolution du paysage, et sont donc ré-échantillonnés plusieurs années. D'autres sites sont fait une seule année. Il convient donc de travailler sur une combinaison année  $\times$  site.

### Liste des fichiers et métadonnées

- Fichier des comptages oiseaux `Oiseaux_up_to_2023.csv`
- Fichier décrivant l'usage des sols `LandUsePer_BM_2023_cartoISea.csv` aka MOS ou mode d'usage des sols, fraction de chaque usage dans des cercles concentriques de rayon 350 à 2km. On pourra commencer par 500m.
- Fichier décrivant les types d'usage des sols `themes_mos_gir.txt`
- Un fichier décrivant les caractéristiques des différents sites `BiodiverCite_sites.csv`, que vous pouvez coupler en dplyr avec l'usage des sols.
- Un pdf décrivant le protocole
- (pour la seconde semaine) Fichier avec les caractéristiques des espèces

**Qu'est-ce qu'un sol imperméabilisé, pour les besoins de notre étude?** On considérera que la surface imperméabilisée est celle dénotée comme "artificialisée" dans le MOS, ce qui regroupe les routes, le bâti et quelques autres éléments construits. Des distinctions plus subtiles sont possibles avec un usage des sols plus détaillés, mais nous resterons simple.

## 4 Questions

Principale

1. Quelle est la forme de la relation biodiversité vs imperméabilisation du sol?

La seconde semaine nous éluciderons les raisons profondes derrière cette relation.

Si vous avez le temps, des questions additionnelles sont traitables

1. la biodiversité réagit-elle à d'autres catégories d'usage des sols? Quelles précautions prendre pour incorporer des variables multiples?
2. la forme de la relation est-elle similaire si on considère non plus l'imperméabilisation du sol mais plutôt la distance au centre ville, sachant que ces deux métriques sont corrélées? On pourra considérer Pey-Berland le centre ville et utiliser le package `sf`.

## 5 Rendu du projet

Le rendu du projet sera un rapport écrit sur les analyses, avec les codes associés. Il peut être rédigé en format pdf (e.g. Latex) classique + code en R, ou directement en format Rmarkdown (rendu en Rmd + pdf compilé). Il n'y a pas de limite sur le nombre de pages car les graphes peuvent prendre pas mal d'espace. Mettez en début de rapport un résumé synthétisant l'ensemble de vos résultats ( $\approx$  250 à 500 mots).

Il n'y a pas de bons ou mauvais résultats. Outre vous faire découvrir des données environnementales, la philosophie de l'exercice est de vous placer dans la position de l'analyste qui découvre un jeu de données nouveau avec des questions nouvelles, e.g., dans une collaboration avec des biologistes. C'est une posture extrêmement fréquente pour un.e statisticien.n.e. Il y a de multiples choix méthodologiques pour chaque question, et j'attends simplement que ces choix soient raisonnablement justifiés.

*A bien noter qu'une seconde semaine de TDs est prévue (15 Avril) pour ce projet!*

Pour simplifier le travail de rédaction, la structure du rapport pourra être séquentielle, avec d'abord les analyses de la première semaine, suivies de celles de la seconde semaine. Essayez néanmoins de m'envoyer un brouillon de rapport sur la première partie à la fin de la première semaine, afin de vous simplifier la vie par la suite.