

Modèles linéaires généralisés à effets mixtes 2

Ce laboratoire doit être remis le 12 février à 17h sur Moodle.

Données

Le jeu de données `aiv_ducks.csv` contient une partie des données de l'étude de Papp et al. (2017) sur la présence de la grippe aviaire (AIV pour *avian influenza virus*) dans les populations de différentes espèces de canards dans l'est du Canada.

Papp, Z., Clark, R.G., Parmley, E.J., Leighton, F.A., Waldner, C., Soos, C. (2017) The ecology of avian influenza viruses in wild dabbling ducks (*Anas* spp.) in Canada. PLoS ONE 12: e0176297. <https://doi.org/10.1371/journal.pone.0176297>.

```
aiv <- read.csv("../donnees/aiv_ducks.csv")
str(aiv)
```

```
## 'data.frame':    8967 obs. of  10 variables:
## $ Species       : Factor w/ 7 levels "ABDU","AGWT",...: 5 5 5 5 5 5 4 5 5 5 ...
## $ Age           : Factor w/ 2 levels "AHY","HY": 2 2 2 1 1 1 1 2 1 2 ...
## $ Sex           : Factor w/ 2 levels "F","M": 2 1 1 2 2 2 1 2 2 2 ...
## $ AIV           : int  1 0 1 1 1 0 1 1 1 0 ...
## $ Site          : Factor w/ 72 levels "Allisary","Amherst Point",...: 2 70 70 67 67 67 1 70 67 1
## $ Latitude      : num  45.8 46 46 46 46 ...
## $ Longitude     : num  -64.2 -64.3 -64.3 -64.3 -64.3 ...
## $ Year          : int  2005 2005 2005 2005 2005 2005 2005 2005 2005 2005 ...
## $ Temperature   : num  18.6 17.6 17.6 17.6 17.6 ...
## $ Population_Density: num  1.2 1.16 1.16 1.16 1.16 ...
```

Voici la description des champs de données:

- *Species*: Code d'espèce (ABDU = canard noir, AGWT = sarcelle à ailes vertes, AMWI = canard d'Amérique, BWTE = sarcelle à ailes bleues, MALL = canard mallard, MBDH = hybride canard noir / mallard, NOPI = canard pilet)
- *Age*: Âge (HY = année d'éclosion, AHY = après l'année d'éclosion)
- *Sex*: Sexe (F/M)
- *AIV*: Présence (1) ou absence (0) du virus de la grippe aviaire
- *Site*: Site d'échantillonnage
- *Latitude* et *Longitude*: Coordonnées géographiques du site
- *Year*: Année d'échantillonnage
- *Temperature*: Température moyenne dans les 2 semaines précédant l'échantillonnage
- *Population_Density*: Densité de population de canards (toutes espèces confondues) estimée pour le site et l'année

1. Ajustement du modèle

- a) Estimez les paramètres d'un modèle complet visant à prédire la présence/absence de l'AIV, incluant: les effets fixes de l'âge et du sexe des canards, de la température et de la densité de population du site; ainsi que les effets aléatoires de l'espèce, du site, de l'année et de l'interaction site x année (cette dernière est notée (1 | Site:Year) dans le modèle). Faut-il vérifier s'il y a surdispersion pour ce modèle?
- b) Quelle est la raison d'inclure chacun des effets aléatoires du modèle en (a)? Vérifiez si ces effets aléatoires suivent une distribution normale.

- c) Produisez un graphique de la probabilité de présence de l'AIV prédite par le modèle en (a) en fonction de la température pour chacune des quatre catégories d'âge et de sexe (HY/F, HY/M, AHY/F, AHY/M), en fixant la densité de population à sa valeur moyenne.
- d) À partir du modèle complet, utilisez l'AIC pour déterminer s'il est préférable d'inclure ou non chacun des effets suivants: température, densité de population, ainsi que l'effet aléatoire pour l'interaction site x année.
- e) Les auteurs de l'étude originale ont déterminé un effet significatif de la densité de population en ajustant un modèle avec effets aléatoires du site et de l'année, mais sans leur interaction. Pour quelle raison les conclusions de votre modèle pourraient-elles différer de ce résultat?

2. Prédiction du modèle

- a) Ajoutez au jeu de données original des colonnes représentant la prédiction de la probabilité de présence de l'AIV (1) en fonction seulement des effets fixes du modèle; (2) en fonction des effets fixes et aléatoires. Utilisez le meilleur modèle identifié dans la partie précédente.
- b) Pour chaque type de prédictions obtenues (effets fixes; effets fixes et aléatoires), déterminez la probabilité moyenne prédite de présence de l'AIV pour les observations avec $AIV = 1$, ainsi que la probabilité moyenne prédite de présence pour les observations avec $AIV = 0$. D'après vos résultats, les effets fixes du modèle permettent-ils de bien départager les cas de présence et d'absence? Qu'en est-il des effets aléatoires?
- c) Groupez le jeu de données par site et année, puis calculez la moyenne de la longitude, de la latitude et de la probabilité d'AIV prédite par le modèle complet pour chaque combinaison site-année. En utilisant ces variables, produisez une carte des sites avec leur probabilité d'AIV prédite pour chaque année. (Vous pouvez utiliser les facettes dans *ggplot2* pour séparer le graphique en panneaux pour chaque année.)