

Modèles linéaires généralisés à effets mixtes

Ce travail doit être mis en ligne sur Moodle avant le **vendredi 10 mars, 17h**.

Les chiffres entre parenthèses indiquent le nombre de points pour chaque question.

Total : 10 points.

Données

Les données de cet exercice proviennent de l'article :

Marrotte, R. R., & Bowman, J. (2021). Seven decades of southern range dynamics of Canada lynx. *Ecology and Evolution*, 11(9), 4644–4655. doi: <https://doi.org/10.1002/ece3.7364>

Nous voulons vérifier si la population de lynx du sud de l'Ontario est sensible à la connectivité à la forêt boréale et aux conditions d'enneigement. La table `spatial_df.csv` contient les colonnes suivantes :

Hex: Sampling unit or Hexagon ID

x = Easting in North America Lambert Conformal Conic of the sampling unit centroid

y = Northing in North America Lambert Conformal Conic of the sampling unit centroid

region: Region of the southern range of the Canada lynx. West, Central, East

OccNb: Number of years where Canada lynx are harvested within a sampling unit

Dist_boreal: Distance to the boreal forest in km

Road_dist_km: Distance to the nearest major road in km.

Snow: Average Winter Severity Index

BobcatP: Number of years where bobcats are harvested within a sampling unit

CoyoteP: Number of years where coyotes are harvested within a sampling unit

```
data = read.csv("../donnees/spatial_df.csv")
str(data)
```

```
## 'data.frame':   65 obs. of  10 variables:
## $ Hex          : chr  "H100" "H116" "H117" "H118" ...
## $ x            : num  1299659 1215422 1271580 1327737 1383895 ...
## $ y            : num  586392 635026 635026 635026 635026 ...
## $ region       : chr   "East" "East" "East" "East" ...
## $ OccNb        : int    0 0 0 0 0 0 0 0 0 0 ...
## $ Dist_boreal  : num    332 244 276 304 339 ...
## $ Road_dist_km: num    0.305 0.838 1.175 1.093 1.154 ...
## $ Snow         : num    497 644 624 554 454 ...
## $ BobcatP      : int     1 1 1 2 6 3 2 8 1 0 ...
## $ CoyoteP      : int     41 60 49 52 51 42 45 45 43 49 ...
```

1. Variations du nombre d'années avec succès de capture pour le lynx

- a. Préparation des données. Transformez la distance à la forêt boréale (*Dist_boreal*) de km en centaines de km et normalisez l'indice de sévérité hivernale moyenne (*Snow*) pour avoir 0 de moyenne et 1 d'écart-type. (1)
- b. Nous voulons savoir si la connectivité à la forêt boréale et les conditions de neige ont un impact sur le nombre d'années où le lynx du Canada est capturé dans une unité d'échantillonnage (*OccNb*), en contrôlant les différences entre les régions. Nous supposons que l'effort d'échantillonnage est identique dans les unités d'échantillonnage et sur des années différentes. Quel type de modèle proposeriez-vous et pourquoi ? (1)
- c. Exploration de données. Représentez le *OccNb* en fonction de *Dist_boreal* en donnant différentes couleurs aux unités d'échantillonnage dans la même région. Vous pouvez utiliser l'option *theme(aspect.ratio = 1)*. Ensuite, tracez le *OccNb* en fonction de *Snow* en donnant différentes couleurs aux unités d'échantillonnage dans la même région. Enfin, regroupez les données par région en utilisant les fonctions du package *dplyr* et calculez la moyenne et l'écart type de *OccNb*. Vous pouvez maintenant représenter la moyenne *OccNb* par région et sa variabilité (+ 1 SD) avec un bar plot. (1)
- d. Avec le modèle complet (fonction de *Dist_boreal*, *Snow*, leur interaction et avec l'effet aléatoire de la région sur l'ordonnée à l'origine), nous testerons si l'excès de zéros peut poser problème. Nous utiliserons la fonction *glmmTMB* testant les modèles de poisson et binomiaux négatifs, avec et sans versions avec excès de zéros. Nous évaluerons la surdispersion de ces modèles et justifierons le choix du meilleur modèle. (1)
- e. Testez l'ajustement de différents modèles selon l'AIC. Comparez le modèle final du point *d* avec d'autres versions du modèle : sans effet aléatoire, sans interaction, avec des variables explicatives uniques. Décrivez le meilleur modèle. (1)
- f. Tester les suppositions. Tester l'absence de surdispersion dans le meilleur modèle. (1)
- g. Calculer la probabilité d'avoir des zéros. (1)
- h. Selon la section des effets fixes du sommaire, définissez le nombre moyen d'années où le lynx du Canada est capturé dans une unité d'échantillonnage et l'effet *Dist_boreal* sur ce nombre. Ignorer la probabilité d'avoir des zéros dans le calcul. (1)
- i. Calculez les intervalles de confiance pour chacun des paramètres du modèle mixte. (1)
- j. Représentez le *OccNb* en fonction de la distance à la forêt boréale puis de *Snow*. Dans les deux graphiques, ajoutez une ligne indiquant les valeurs prédites par le meilleur modèle. Commentez les résultats des graphiques. (1)