
Améliorer la pertinence et l'efficacité des modèles statistiques en écologie : extension des fonctions sigmoïdes dans le cadre de l'étude de la distribution de la biodiversité

Ugoline Godeau* and Frédéric Gosselin*¹

¹Ecosystèmes forestiers (UR EFNO) – Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement – Domaine des Barres, F-45290 Nogent-sur-Vernisson, France

Résumé

La modélisation statistique de la biodiversité est un enjeu fort pour mieux éclairer les choix de gestion ou suivre l'évolution de la biodiversité par exemple avec des indicateurs. Des outils développés récemment permettent de modéliser la biodiversité spécifique de manière hiérarchique, c'est-à-dire à la fois au niveau espèce, groupe écologique et communauté d'espèces, permettant de partager l'information environnementale entre espèces tout en prenant en compte les corrélations entre espèces. Ces outils utilisent le plus souvent des données de présence-absence d'espèces et utilisent un modèle statistique avec fonction de lien probit qui permet une simplification numérique. S'il est relativement bien établi que le choix entre une fonction de lien logit, probit voire cloglog n'a d'effet que dans des cas où les données sont très nombreuses, nous voulons ici attirer l'attention sur d'autres formes de fonctions de lien, avec asymptotes estimées, qui peuvent grandement changer les résultats. Nous expliquons pourquoi tel devrait être le cas et montrons par un travail de simulation qu'utiliser les fonctions de lien "classiques" peut entraîner des biais importants dans l'estimation de l'importance des effets. Nous illustrerons ce résultat sur un cas réel. Ces résultats ont une portée beaucoup plus générale que celui de notre cadre d'étude.

Mots-Clés: logistique, sigmoïde, GLM, régression non linéaire, inférence, biais, fonction de lien

*Intervenant