



**PatriNat**

**Centre d'expertise et de données sur la nature**

**OFB-CNRS-MNHN**

**Le MNHN recrute :**

**«Un(e) stagiaire de Master 2 « Vers une utilisation des données du SINP pour l'ajustement de modèles de distribution à espace d'états »**

## **PRÉAMBULE**

L'Unité Patrimoine Naturel (PatriNat) - co-habituée par le Muséum national d'Histoire naturelle (MNHN), l'Office française de la biodiversité (OFB) et le Centre national de la recherche scientifique (CNRS) – a pour mission de produire une expertise scientifique et technique en faveur de la conservation de la biodiversité.

Au sein du MNHN, PatriNat est rattachée à la Direction Générale Déléguée Recherche, Expertise, Valorisation et Enseignement du Muséum (DGD REVE), et a pour objectif de fournir, en lien fort avec les activités de recherche, une expertise scientifique sur la biodiversité et géodiversité de France métropolitaine et ultra-marine, sur les thématiques terrestres et marines, pour l'environnement passé et actuel. Cette expertise et l'ingénierie associée porte sur la connaissance du patrimoine naturel, dont les systèmes d'informations et à l'application de ces connaissances pour l'appui aux politiques et programmes de conservation de la biodiversité.

## **CONTEXTE**

L'érosion rapide de la biodiversité est une source d'inquiétude majeure. Son suivi le plus précis et exhaustif possible est d'une importance capitale pour documenter ce déclin mais aussi s'assurer de l'efficacité des mesures de conservation implémentées. Toutefois, le monitoring précis de l'ensemble des espèces au travers de dispositifs scientifiques protocolés, bien qu'idéal, est extrêmement coûteux en moyens humains. Par exemple certaines espèces discrètes sont particulièrement difficiles à suivre (Znidarsic et al., 2020). De plus, même pour les espèces moins discrètes, estimer correctement les variations d'abondances à large échelle demande de prendre en compte les différences spatio-temporelles de détectabilités mais aussi de présence réelle de l'espèce. Par exemple, le pic de migration de certaines espèces peut varier (Lehikoinen et al., 2019). Les conditions météorologiques peuvent également influencer nos capacités de détection. Dans le cas du développement de modèles de distribution d'espèces (Species distribution model, SDM, Feldman et al., 2021), cela peut entraîner de fortes erreurs d'estimation de la probabilité de présence d'une espèce (Kéry et al., 2006).

Chronologiquement, deux approches ont été préférentiellement utilisées pour contourner ces problèmes. La première a été le recours aux avis d'experts. Cette approche pose deux problèmes : d'une part ils

reposent sur la participation très active d'une communauté réduite, d'autre part ces avis ils ne se basent pas non plus sur des protocoles standardisés. Plus récemment, des travaux statistiques ont été menés à partir des bases de données issues des sciences participatives. Cependant, si les données sont très abondantes, les initiatives reposant sur un protocole standardisé sont souvent limitées à des communautés animales et des listes d'espèces circonscrites (Sauvages de ma rue, STOC-EPS par exemple). De plus ces protocoles sont souvent circonscrits à une période précise avec peu, voir une seule, session d'observation, et les données ne permettent souvent pas de différencier les absences réelles et des cas de non-détection. Enfin, ce dernier point n'est pas toujours traité dans un cadre statistique adéquat même si les données le permettent (Dickinson et al., 2010).

Les modèles intégrés à espace d'états (Isaac et al., 2020 et références associées) semblent prometteurs pour traiter les biais méthodologiques susmentionnés. Tout d'abord, ces modèles reposent sur une modélisation jointe i) de la probabilité de l'espèce d'être présente au point d'observation et ii) de la probabilité de la détecter si elle est présente. Des variables d'habitat et des coordonnées géographiques, notamment, peuvent être considérées comme variables explicatives. De plus, l'intégration simultanée de données issues de sources variées (différents protocoles de sciences participatives par exemple) permet de gagner en puissance, notamment en permettant d'identifier des observations multiples.

Ici, nous proposons d'appliquer ces modèles à certaines espèces dont les données d'observation sont bancarisées dans la plateforme nationale du SINP (Système d'information de l'inventaire du patrimoine naturel). Pour ce faire, nous faisons l'hypothèse qu'il est possible d'identifier au sein de la base de données des experts de groupe taxonomique (i.e. odonates, carabidés, oiseaux) dont l'absence de déclaration d'observation correspond à une non détection d'espèce. Une fois ces observateurs experts identifiés, et pour certaines espèces faisant l'objet de plusieurs passages annuels dans les mêmes mailles territoriales, l'ensemble des déclarations et non déclarations seront utilisées pour ajuster des SDM à espace d'états dans un cadre bayésien. L'inclusion d'effets aléatoires pourra permettre de corriger pour l'intensité des observations en fonction des localités et des protocoles d'observation. Les résultats obtenus via les modèles intégrés à espace d'états seront comparés à ceux de SDM alimentés des seules données de présence. La validité de cette approche ainsi que son application à d'autres espèces seront discutées.

## ÉTAPES DU STAGE

Au sein de l'équipe « données » le/la stagiaire aura pour mission :

1. En collaboration avec des experts du MNHN, d'affiner les critères d'identification et de sélection dans les bases de données des observateurs « experts ». A priori, Les observateurs sélectionnés seront des personnes (ou des groupes de personnes) responsables d'une large part des observations pour un groupe taxonomique particulier (amphibiens, odonates...). Une fois ceux-ci identifiés, les espèces pour lesquelles des mailles territoriales ont été visitées plusieurs fois par an seront sélectionnées comme candidates. La déclaration régulière d'espèces communes pourra servir de proxy pour estimer l'exhaustivité des déclarations des experts.
2. Pour chaque espèce sélectionnée, de compiler un jeu de données en présence-absence à partir de l'ensemble des données d'observation retenues. Les analyses seront a priori limitées à 3 ou 4 espèces et les données idéalement agrégées aux mailles 5 km par 5 km.
3. De modéliser la distribution des espèces à l'aide de SDM basés uniquement sur des présences et de SDM à espace d'états développés dans un cadre bayésien.
4. De comparer les résultats des deux types de modèles. Les résultats provenant d'études robustes préexistantes sur la distribution ou l'abondance globale des espèces considérées serviront de base de comparaison. A défaut une comparaison avec des atlas issus d'avis d'expert sera réalisée. Des comparaisons de performances prédictives effectuées grâce à une validation croisée pourront également être considérées.

## CONDITIONS ADMINISTRATIVES REQUISES

Être titulaire au minimum d'un diplôme de niveau «Master 1» en écologie ou biostatistique

## COMPÉTENCES RECHERCHEES

- Intérêt pour l'étude de la distribution des espèces
- Intérêt pour la modélisation et les biostatistiques
- Expérience préalable du logiciel R

## MODALITÉS DE RECRUTEMENT

Type de contrat : stage

Durée : 6 mois (négociable)

Rémunération : selon réglementation en vigueur

Date de prise de fonction : Janvier 2023 (négociable)

Lieu : Campus Muséum Brunoy, 4 avenue du petit Château, 91800 Brunoy

## PROCÉDURE DE RECRUTEMENT

Le dossier de candidature sera composé d'une lettre de motivation, d'un CV détaillé et de tous travaux que vous jugerez utile de joindre (rapport de stage de M1 par exemple). Il est à envoyer par courriel à Guillaume BAL ([guillaume.bal@mnhn.fr](mailto:guillaume.bal@mnhn.fr)) et Tiphaine OUISSE ([tiphaine.ouisse@mnhn.fr](mailto:tiphaine.ouisse@mnhn.fr)) et sera examiné dès réception.

## RÉFÉRENCES

Dickinson et al., 2010. Citizen science as an ecological research tool : challenges and benefits. *Annual review of Ecology and Systematics* 41(1):149-172 <https://doi.org/10.1146/annurev-ecolsys-102209-144636>

Feldman et al., 2021. Trends and gaps in the use of citizen science derived data as input for species distribution models: a quantitative review. *PLoS ONE* 16(3): e0234587. <https://doi.org/10.1371/journal.pone.0234587>

Isaac et al., 2020. Data Integration for Large-Scale Models of Species Distributions. *Trends in Ecology & Evolution* 35, 56–67. <https://doi.org/10.1016/j.tree.2019.08.006>

Kéry et al., 2006. How biased are estimates of extinction probability in revisitation studies ? *Journal of Ecology* (94) 980-986. <https://doi.org/10.1111/j.1365-2745.2006.01151.x>

Lehikoinen et al., 2019. Phenology of the avian spring migratory passage in Europe and North America : asymmetric advancement in time and increase in duration. *Ecological indicators* (101) 985-991 <https://doi.org/10.1016/j.ecolind.2019.01.083>

Znidarsic et al., 2020. Using visualization and machine learning methods to monitor low detectability species - the least bittern as a case study. *Ecological informatics* (55) 101014 <https://www.sciencedirect.com/science/article/pii/S1574954119303255>