



## Proposition de stage de 6 mois

---

Estimation Bayésienne de la dynamique spatiale d'invasion  
en combinant observations opportunistes et standardisées

---

### Contexte scientifique

Les invasions biologiques sont l'une des principales causes de l'érosion récente de la biodiversité et représentent des coûts importants pour nos sociétés. La modélisation de la dynamique spatiale des populations est cruciale pour comprendre les déterminants des invasions biologiques aux changements globaux (réchauffement climatique, changement d'utilisation des sols). La dynamique spatiale peut être modélisée à partir de processus causaux, au travers des traits d'histoire de vie d'une espèce (e.g. fécondité, mortalité, taux de dispersion dans un paysage hétérogène) et de sa réponse aux variations environnementales (liée au concept de niche écologique), dont les paramètres sont en général ajustés en conditions contrôlées ou via des protocoles d'échantillonnage standardisés, ce qui limite fortement les applications. La plupart des observations de biodiversité accessibles à large échelle sont des présences seules (l'observation d'une espèce n'implique pas l'absence des autres) dont l'échantillonnage n'est pas standardisé. Les données de présence seule ont le potentiel de compléter les données standardisées en nous informant par exemple sur les taux de dispersion à large échelle spatiale et les voies d'invasion (e.g. routes, cours d'eau, voies de train). Cependant, l'intégration des données de présence seule complexifie la modélisation, car il faut notamment prendre en compte les biais d'échantillonnage spatio-temporels et l'hétérogénéité des sources de données. Le problème peut être surmonté en alliant un modèle latent de dynamique spatiale de population à un modèle d'observation tenant compte des biais d'échantillonnage afin d'estimer les paramètres du premier sur des données opportunistes (Louvrier et al., 2020, Botella et al., 2022). Cependant, la difficulté à définir un modèle suffisamment complexe tout en parvenant à faire converger les algorithmes d'estimation rend l'approche difficile à mettre en pratique.

Pour répondre à ce défi, un package R est en cours de développement pour faciliter l'estimation de modèles de distribution d'espèces Bayésiens dynamiques (dSDM) en temps et espace discrets intégrant données opportunistes et standardisées, et généralisant ainsi l'approche introduite dans Botella et al. (2022). La vraisemblance des observations spatio-temporelles y découle de la population latente par site spatial et du protocole d'échantillonnage, incluant les variations spatio-temporelles d'effort d'échantillonnage et de détectabilité entre les jeux de données. L'inférence Bayésienne permet notamment de contraindre la distribution a priori des paramètres au vu de la connaissance experte de l'espèce (e.g. fécondité, capacité de charge, aire d'introduction). Le stage visera à mettre l'approche dSDM à l'épreuve de la reconstruction de l'invasion de l'ambrosie à feuilles d'armoise (*Ambrosia artemisiifolia* L.) en France depuis la fin du XXème siècle en capitalisant sur des milliers de données opportunistes et sur des inventaires de présence-absence (Metcalf et al., 2023).

## Objectifs et étapes du stage

- En concertation avec des partenaires écologues (Anses, CIRAD), identifier la connaissance experte sur l'ambrosie à feuilles d'armoise (aire d'introduction, mécanismes de dispersion spatiale) et concevoir un dSDM adapté.
- Préparer les données de présence-seule (Pl@ntNet, iNaturalist, Fédération des Conservatoires Botaniques Nationaux) et de présence-absence (e.g. AWME, Metcalfe et al., 2023, 500 ENI Fried et al., 2023) d'*Ambrosia artemisiifolia*, ainsi que les variables environnementales (climatique, sol, utilisation du sol, réseau de routes) à partir de sources largement utilisées en écologie pour le dSDM.
- Estimer et analyser plusieurs instances du dSDM, correspondant à différentes hypothèses sur l'écologie de l'invasion (e.g. dépendance de la vitesse dispersion à la taille des axes routiers, pertinence de l'information a priori sur les traits pour expliquer les données) ou différents sous-ensemble des données (effet d'intégrer ou pas des données de présence absence). Le stagiaire utilisera un package R bientôt disponible qu'il n'aura pas à développer par lui-même. Valider et comparer les instances dSDM via des données indépendantes.
- Produire des visualisations cartographiques et temporelles de la dynamique passée de l'Ambrosie pour interpréter sa trajectoire d'invasion en France. Les valoriser par la production et la publication d'un notebook R (e.g. RMarkdown, Quarto) soulignant les fonctionnalités clés du package.
- Si le temps le permet, des simulations sous des scénarios alternatifs de gestion de l'invasion ou d'aménagement du territoire pourront être réalisées pour identifier les déterminants clés de l'invasion passée ou à venir.

**Déroulement du stage** : Le stage de 6 mois pourra débuter en février 2024, mais un début plus tardif est possible. Le ou la stagiaire sera affecté.e dans l'équipe Inria Zénith, sur le Campus Saint Priest de l'Université de Montpellier, et sera principalement encadré.e par Christophe Botella (Inria). L'équipe héberge [Pl@ntNet](#), est active sur l'intelligence artificielle pour la biodiversité, notamment via trois projets européens en cours, et plusieurs étudiants et chercheurs travaillent sur des sujets connexes au stage. Des échanges réguliers seront organisés avec des partenaires locaux du projet, Guillaume Fried (Anses, Laboratoire de la Santé des Végétaux, Montpellier) et Pierre Bonnet (CIRAD, AMAP, Montpellier), ainsi que les partenaires internationaux (Stellenbosch University, Afrique du Sud). Gratification minimale de stage en vigueur.

**Profil recherché** : Etudiant en Master 2, école d'ingénieur ou césure. Une formation principale en écologie/agronomie avec des connaissances en statistiques est attendue, ou une formation principale en probabilités/statistiques avec des connaissances en écologie. Une expérience de programmation en R est requise. Une compréhension des concepts probabilistes et statistiques, en particulier de l'inférence Bayésienne, sera utile.

**Contact** : Les candidatures devront être envoyées par mail [christophe.botella@inria.fr](mailto:christophe.botella@inria.fr) avec CV.

## Références

-Botella, C., Bonnet, P., Hui, C., Joly, A., & Richardson, D. M. (2022). Dynamic species distribution modeling reveals the pivotal role of human-mediated long-distance dispersal in plant invasion. *Biology*, 11(9), 1293.

- Brunel, S., Schrader, G., Brundu, G., & Fried, G. (2010). Emerging invasive alien plants for the Mediterranean Basin. *EPPO bulletin*, 40(2), 219-238.
- Louvrier, J., Papaix, J., Duchamp, C., & Gimenez, O. (2020). A mechanistic–statistical species distribution model to explain and forecast wolf (*Canis lupus*) colonization in South-Eastern France. *Spatial Statistics*, 36, 100428.
- Fried, G., Poinas, I., Henckel, L., Alignier, A. (2023). Major field margin vegetation types in France and their relationships with climate, agricultural landscapes and management intensity. *Botany Letters* <https://doi.org/10.1080/23818107.2023.2269243>
- Metcalfe, H., Bürger, J., Redwitz, C. V., Cirujeda, A., Fogliatto, S., Dostatny, D. F., ... & Fried, G. (2023). The utility of the 'Arable Weeds and Management in Europe'database: Challenges and opportunities of combining weed survey data at a European scale. *Weed Research*, 63(1), 1-11.