

## OFFRE de thèse en statistique appliquées à l'écologie

### Intitulé de la thèse

Nouveaux modèles de Semi Markov caché multi-chaînes pour les dynamiques de métapopulation avec population partiellement observable

### Date de début et durée

La thèse pourra commencer entre le 01/10/2021 et le 01/12/2021, et pour une durée de 3 ans.

### Mots-clés

Statistique des processus, modèle à variable latentes, statistique computationnelle, dynamiques de métapopulation, dormance, adventices

### Présentation

La distribution d'une espèce dépend de sa dynamique locale et de sa dynamique régionale, via l'extinction et la colonisation. Le cadre de la **métapopulation** (Levins et al. 1969) permet de modéliser ces deux processus et de décrire la dynamique spatio-temporelle d'une espèce dans un réseau de patchs connectés. Ce modèle est très utilisé en écologie, cependant il n'est pas adapté pour les organismes (plantes ou animaux) qui ont la capacité de stopper temporairement leur développement par un processus appelé **dormance**, car ce stade n'est pas modélisé. Utiliser un modèle de métapopulation peut alors conduire à une fausse conclusion d'extinction de l'espèce. Mieux modéliser la dynamique des espèces avec dormance est donc un enjeu pour une meilleure compréhension de la part relative des processus pilotant leur dynamique, ce qui permettra une meilleure gestion ou conservation de ces espèces. Au-delà des espèces avec dormance, la question de savoir inclure des **stades non observés** dans un modèle de métapopulation se pose plus généralement pour toutes les espèces dont l'un des stades est cryptique. Il y a donc un **fort intérêt en écologie pour des modèles de dynamique sur patchs pour ce que l'on appellera ici des populations partiellement observables**. Elaborer un tel modèle demande un compromis entre la richesse de représentation du modèle et la complexité de l'estimation des paramètres du modèle et de l'inférence des stades non observés. Pour bien comprendre l'importance de la sous-population dormante ou cryptique dans la dynamique de l'espèce, il est crucial de bien représenter la durée du temps de dormance ou du temps de séjour dans le stade cryptique. Par ailleurs, les dépendances spatiales entre les dynamiques au sein de chaque patch, auxquelles s'ajoutent l'aspect partiellement observable, font que les méthodes statistiques d'estimation classiques, atteignent leurs limites au-delà de quelques patchs. Un cadre naturel pour modéliser des dynamiques sur patch avec populations partiellement observées est celui des modèles **graphiques dynamiques avec variables latentes**. Ainsi le modèle de chaîne de Markov cachée (**HMM** pour Hidden Markov Model en anglais) a déjà été utilisé pour modéliser la dynamique d'une espèce adventice sur un réseau de patchs organisés sur un cercle (Manna et al. 2017), ou encore sur un réseau de parcelles (Le Coz et al. 2019). Dans ces modèles, la modélisation du temps de séjour dans un état de la banque de graines est contrainte par l'hypothèse markovienne inappropriée car les graines peuvent survivre plusieurs années jusqu'au retour de conditions favorables. Le projet de thèse se place dans le cadre plus riche, mais plus complexe à estimer, des modèles semi-Markoviens cachés (HSMM, Barbu et Limnios 2008). De plus, les travaux pré-cités reposent sur un couplage spécifique des dynamiques des différents patchs dans un modèle spatio-temporel. Ce cas particulier n'est pas adapté à toutes les dynamiques de métapopulation (typiquement celles où le stade non observé est mobile, contrairement aux graines). L'originalité de ce travail de thèse sera de proposer de **nouvelles structures** de couplage spatial des chaînes cachées dans un HSMM multi-chaînes, pertinentes pour l'étude des métapopulations avec population partiellement observable et **d'apporter à la fois les éléments théoriques et opérationnels** pour l'estimation des modèles correspondants. Si l'objectif est de travailler sur la construction d'un cadre générique, il sera illustré sur un cas d'intérêt particulier pour l'INRAE et le CEFE : les **espèces adventices des cultures**. L'enjeu de l'agroécologie est de combiner protection de la biodiversité et production agricole à la hauteur des besoins. Dans ce contexte, les adventices des cultures sont à la fois des acteurs clés du

maintien de la biodiversité (relations plantes/pollinisateurs, maintien de populations d'auxiliaires des cultures) et un frein à la production des cultures avec lesquelles elles entrent en compétition. Afin de mieux appréhender cet équilibre, il est nécessaire de comprendre leurs stratégies de survie locale et par dispersion. L'état caché est ici la banque de graines, stade difficilement observable car l'échantillonnage des graines dans le sol est coûteux et empreint de biais méthodologiques.

### Encadrement et unité d'accueil

Les encadrants de la thèse sont

- Nathalie Peyrard : unité MIAT, INRAE Toulouse ([nathalie.peyrard@inrae.fr](mailto:nathalie.peyrard@inrae.fr))
- Nikolas Limnios : laboratoire LMAC, UTC Compiègne ([nikolaos.limnios@utc.fr](mailto:nikolaos.limnios@utc.fr))
- Pierre-Olivier Cheptou : laboratoire CEFE, CNRS Montpellier ([pierre-olivier.cheptou@cefe.cnrs.fr](mailto:pierre-olivier.cheptou@cefe.cnrs.fr))

La thèse se déroulera dans l'équipe SCIDyn, de l'unité MIAT du centre INRAE de Toulouse (<https://miat.inrae.fr/site/SCIDYN>). Des visites seront prévues au CEFE (<https://www.cefe.cnrs.fr/fr/>) et au LMAC (<http://lmac.utc.fr/>).

Le/la doctorant(e) s'inscrira à l'école doctorale MITT de Toulouse.

### Profil recherché

Compétences en statistiques (statistique des processus, statistique computationnelle). Goût pour les applications et la programmation. En fonction du profil de la personne recrutée, le focus de la thèse se portera plus sur les aspect théoriques ou algorithmiques de l'estimation des HSMM multi-chaines.

### Comment candidater ?

Envoyer un CV et une lettre de motivation à Nathalie Peyrard, Nikolas Limnios et Pierre-Olivier Cheptou avant le 31 aout 2012.

### Références

V.S. Barbu, N. Limnios, Semi-Markov Chains and Hidden Semi-Markov Models. Toward Applications. Their use in Reliability and DNA Analysis. Vol. 191. Lecture Notes in Statistics. Springer, New York, 2008.

S. Le Coz, P.-O. Cheptou, N. Peyrard, A spatial Markovian framework for estimating regional and local dynamics of annual plants with dormant stage. *Theoretical Population Biology*, vol 127, pp 120 - 132, 2019

R. Levins, D. Vagaggini, P. Zarattini, and G. Mura. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological Society of America*, 15(3), pp 237–240, 1969

F. Manna, R. Pradel, R. Choquet, H. Freville H, P.-O. Cheptou. Disentangling the role of seed bank and dispersal in plant metapopulation dynamics using patch occupancy surveys. *Ecology*, vol 98 (10), pp 2662-2672, 2017