

ÉQUATION DIFFÉRENTIELLE STOCHASTIQUE BASÉE SUR UN POTENTIEL GAUSSIEN POUR DÉCRIRE LE DÉPLACEMENT EN ÉCOLOGIE

Pierre Gloaguen ¹, Sylvain Le Corff ² & Marie-Pierre Etienne ³

¹ *Écologie et Modèles pour l'Halieutique. IFREMER, rue de l'île d'Yeu, Nantes.
pierre.gloaguen@ifremer.fr*

² *CNRS, Département de Mathématiques Université Paris-Sud, Orsay.
sylvain.lecorff@math.u-psud.fr*

³ *UMR AgroParisTech/INRA 518, Team MORSE 16 rue Claude Bernard, Paris.
marie.etienne@agroparistech.fr*

Résumé court L'analyse à fine échelle des déplacements individuels en écologie apporte une information importante pour la définition de mesures de gestion adéquates dans un contexte spatialisé. Dans cette optique, l'analyse des trajectoires d'individus, obtenues grâce au système de télémétrie, est utile à la compréhension des mécanismes guidant le déplacement à différentes échelles d'espace et de temps. Un modèle en temps et en espace continu a été développé pour décrire les trajectoires d'individus. On fait l'hypothèse que la trajectoire d'un individu est solution d'une équation différentielle stochastique (EDS) dont la dérive suit un gradient spatial. Ce gradient est celui d'une surface de carte de potentiel, décrivant l'attractivité de l'environnement, tel que perçu par l'individu (donc par exemple par un gradient d'abondance de la ressource). Une méthode d'inférence statistique est développée pour estimer cette carte de potentiel à partir de données de télémétrie.

Cette méthode se base sur un algorithme de type MCEM. La fonction d'objectif est d'abord approchée via des méthodes de Monte Carlo. Cette étape nécessite la simulation du processus solution de l'EDS. On utilise un algorithme de simulation conditionnelle exacte, évitant ainsi une erreur due aux schémas discrétisés. La fonction approchée est ensuite maximisée grâce à l'algorithme CMA-ES.

La méthode est illustrée sur des données de navires de pêche exerçant en Manche Est.

Mots-clés. Modèle de mouvement, Equation différentielle stochastique, Simulation exacte, algorithme MCEM.

Short abstract The fine-scale analysis of individual movements in ecology provides important information for defining adequate management measures in a spatial context. In this context, the analysis of individual trajectories obtained through the telemetry system is useful for understanding the mechanisms guiding the movement at different scales of space and time. A continuous model (in space and time) is proposed to describe the trajectories of individuals. It is assumed that the trajectory of an individual is the solution of a stochastic differential equation (SDE), which drift is a spatial gradient. This

gradient is that of a potential surface describing the attractiveness of the environment as perceived by the individual (for example the gradient of abundance of the resource). A statistical inference method is developed to estimate the potential map from telemetry data. This method is based on a MCEM algorithm. The objective function is approximated via Monte Carlo methods. This step requires the conditional simulation of the process solution of the SDE. An exact simulation algorithm is used, avoiding errors due to discretization schemes. The approximate function is then maximized by the CMA-ES algorithm. The method is illustrated on a data set of fishing vessels performing in the Eastern Channel.

Key words. Movement Model, Stochastic differential equation, Exact simulation, MCEM algorithm.

Résumé En écologie, comprendre les mécanismes guidant le mouvement d'un individu est un fort enjeu pour définir de meilleures mesures de gestion. Les équations différentielles stochastiques (EDS) offrent un cadre général pour modéliser en temps continu le processus des positions d'un individu se déplaçant dans son environnement. Cependant, les modèles de mouvement existant se basant sur les EDS sont restreints aux EDS classiques (Harris et Blackwell, 2013), dont les lois de transition du processus solution sont analytiques (mouvement brownien, processus de Ornstein Ulhenbeck...). Ces modèles ont l'avantage d'être analytiques mais n'ont pas toujours une base écologique fondée. Nous proposons un nouveau cadre de modélisation-estimation pour décrire le mouvement d'un individu. Le processus bidimensionnel X_t des positions d'un individu est supposé être solution de l'EDS

$$dX_t = \nabla P(X_t)dt + \gamma dW_t$$

où $P(X_t)$ est une fonction de potentiel décrivant l'attractivité de l'environnement dans lequel se déplace l'individu, ∇ est l'opérateur gradient, et W_t est le mouvement Brownien standard. Ainsi, le déplacement de l'individu traduit l'attractivité de l'environnement tel qu'il est perçu par cet individu (Brillinger, 2010).

La carte de potentiel P est supposée être un mélange de distributions gaussiennes, soit, formellement:

$$\begin{aligned} \mathbb{R}^2 &\mapsto \mathbb{R} \\ x &\mapsto P(x) = \sum_{k=1}^K \pi_k \varphi_k(x) \end{aligned}$$

où

- K est le nombre de composantes de mélange (supposé connu)
- φ_k est la densité de probabilité d'une loi normale en dimension 2, de moyenne μ_k et de covariance C_k .

- π_k est le poids relatif de la k -ième composante $\pi_1 + \dots + \pi_K = 1$

L'ensemble des paramètres à estimer est donc

$$\theta = \{(\pi_k)_{k=1,\dots,K-1}, (\mu_k, C_k)_{k=1,\dots,K}, \gamma\}$$

En suivant les travaux de Beskos et al (2006), nous proposons un cadre d'estimation se basant sur la simulation exacte du processus X_t conditionnellement aux observations (les données GPS).

Nous sommes donc confrontés à un problème d'estimation avec données manquantes (les trajectoires entre chaque observation) qui peut se traiter grâce à l'algorithme EM. L'étape E est approchée par méthodes de Monte Carlo. La simulation du processus conditionnellement aux observations se fait de manière exacte (selon la vraie loi du processus sous θ), ce qui permet d'éviter les erreurs d'approximation des schémas discrétisés et de définir un estimateur sans biais de la vraisemblance des observations. L'étape M est un problème d'optimisation non convexe. On utilise l'algorithme CMA-ES, algorithme stochastique et ne nécessitant pas le gradient de la fonction d'objectif (Hansen et Ostermeier, 2001).

Nous montrerons une application de ce modèle à des données réelles.

Bibliographie

- [1] Harris, K. J. and Blackwell, P. G. (2013), Flexible continuous-time modelling for heterogeneous animal movement, *Ecological Modelling*, vol 255, pp 29-37.
- [2] Brillinger D. (2010), Modelling spatial trajectories, *Handbook of Spatial Statistics*, CRC Press: Boca Raton, FL.
- [3] Beskos, A. and Papaspiliopoulos, O. and Roberts, G.O. and Fearnhead, P. (2006), Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes, *Journal of Royal Statistical Society, Series B: Statistical Methodology*, vol 68, pp 333-382.
- [4] Hansen, N. and Ostermeier, A. (2001), Completely derandomized self-adaptation in evolution strategies, *Evolutionary Computation*, vol 9, pp 159-195