



Stage de M1 ou M2 Statistiques pour l'environnement

**Sujet : Modélisation d'arrondis pour des données de comptages.
Application au comptage de petits cétacés (dauphins).**

- **Encadrants :**
 - Sophie Donnet (CR INRA, MIA Paris) <https://www6.inra.fr/mia-paris/Equipes/Membres/Sophie-Donnet>
 - Isabelle Albert (CR INRA, MIA Paris) <https://www6.inra.fr/mia-paris/Equipes/Membres/Isabelle-Albert>
 - Matthieu Authier (Pelagis, La Rochelle) <http://www.observatoire-pelagis.cnrs.fr/1-equipe/article/authier-matthieu>
- **Durée :** 6 mois.
- **Début :** libre
- **Laboratoire :**
 - Equipe Morse, MIA Paris, Unité Mixte de recherche INRA - AgroParisTech 16 rue Claude Bernard, 75005 Paris, <https://www6.inra.fr/mia-paris>
 - Observatoire PELAGIS - UMS 3462, Université de La Rochelle / CNRS, 5 allée de l'Océan, 17 000 La Rochelle <http://www.observatoire-pelagis.cnrs.fr/>
- **Localisation :** Paris ou La Rochelle
- **Contact :** sophie.donnet@agroparistech.fr, isabelle.albert@inra.fr

Contexte scientifique Le stage est motivé par le jeu de données collectées depuis 2004 sur les campagnes bateau MEGASCOPE dans l'Atlantique Nord-Est. Lors de ces campagnes opportunistes, les observateurs à bord de bateaux ou d'avions dénombrent les groupes de mammifères marins observés. De ces comptages seront déduites les estimations d'abondance des différentes espèces. Une étude rapide des données montre que les observateurs ne peuvent dénombrer exactement le nombre d'animaux dans les grands groupes ; ils fournissent alors un comptage arrondi à la dizaine ou la centaine. Ne pas prendre en compte ces arrondis a pour effet de biaiser les estimations d'abondance. La modélisation de données arrondies (heaping data) est déjà considérée dans la littérature . Durant ce stage, il s'agit notamment de modéliser différents comportements d'arrondis pour les observateurs, puis d'étudier les propriétés statistiques des estimateurs d'abondance sous ces diverses façons d'arrondir les comptages de données groupées .

Objectifs L'objectif de ce stage est d'adapter les modèles de la littérature aux données de MEGASCOPE et de proposer une méthode d'inférence (de préférence bayésienne) de ces modèles. La pertinence, précision et stabilité de ces méthodes seront testées sur données simulées. Les conséquences en matière d'estimations d'abondance des divers stocks seront formulées à partir de données réelles.

Compétences attendues Au delà des compétence classiques attendues pour tout stagiaire (curiosité, motivation, rigueur, aptitude à lire et comprendre l'anglais, autonomie et goût du travail en équipe), le candidat à ce stage devra réunir les compétences académiques suivantes :

- solides connaissances en statistiques et probabilités, donc avoir suivi et validé la partie académique d'un cursus de type M1 ou M2 statistiques et/ou/ modélisation statistique et/ou probabilités,
- goût pour le développement d'outils algorithmiques,
- maîtriser (et aimer manipuler) un environnement de calcul scientifique, R (de préférence) ou Python.