

Workshop: Modélisation, analyse statistique, et applications en écologie et climatologie

19 et 20 mars

Abstracts

Aurélie Fischer (LPSM, Université Paris Cité)

Apprentissage statistique en sciences du climat : exemple des ondes internes de gravité

Dans cet exposé, nous considérerons l'application de méthodes d'apprentissage statistique en sciences du climat : l'objectif est d'améliorer la connaissance et la description de processus physiques de petite échelle. Pour tenir compte des effets de ces processus de petite échelle, qui ne sont pas explicitement décrits dans les modèles de climat, il peut être utile d'intégrer, grâce à l'apprentissage statistique, les informations précises qui peuvent être fournies par des observations de ces processus. Les processus de petite échelle auxquels nous nous intéresserons sont les ondes internes de gravité, ondes dues au phénomène de gravité et à un contraste de densité sur la verticale. Les ondes de gravité jouent en effet un rôle crucial dans la circulation atmosphérique au-dessus de 15-20 km.

Les observations dont nous disposons sont des mesures par ballons superpressurisés stratosphériques, obtenues dans le cadre de la campagne Stratéole 2, projet franco-américain du Centre national d'études spatiales. Le comportement quasi-Lagrangien des ballons permet d'accéder à des estimations précises de flux de quantité de mouvement associés aux ondes de gravité dans la basse stratosphère. Les variables explicatives décrivant l'écoulement à grande échelle sont quant à elles fournies par les données de réanalyse ERA5 provenant du Centre Européen pour les Prévisions Météorologiques à Moyen Terme.

Adeline Leclercq Samson (LJK, Université Grenoble Alpes)

Some statistical models to quantify the effect of climate change on whales in Greenland

Human activities have a profound impact on marine ecology in Greenland. In this presentation, I will focus on a study of the impact of these anthropogenic activities on narwhals, one of the endemic species of whales in Greenland. I will present different stochastic models to analyze the data from this study: point process with memory, stochastic Langevin diffusions, stochastic diffusions in a constraint domain. I will discuss more in detail the diffusion processes which can be multidimensional, hypoelliptic (with a degenerate noise) and partially observed. I will discuss the question of parameter es-

timation when only discrete observations are available. New estimation methods involve numerical schemes called splitting schemes, that prove to be numerically very stable.

Stéphane Robin (LPSM, Sorbonne Université)

Improving the variational inference of latent variable models in ecology

Latent variable models are a basic tool in ecology, particularly in community ecology. In particular, we will present the log-normal Poisson model (PLM), which provides a general framework for the joint distribution of several species in different environments, and the stochastic block model (SBM), which can be used to search for structure in an observed network of interactions between species.

The usual tool for inferring latent variable models is the EM algorithm, which requires at least partial knowledge of the conditional distribution of latent variables conditional to observed variables. For many models, including PLN and SBM, this step is unfeasible, and so we generally resort to a so-called variational approximation of this conditional distribution, resulting in approximate inference.

While variational algorithms are computationally efficient, they do not come with desirable guarantees such as consistency or asymptotic normality of estimators. We will present two different approaches based on Monte-Carlo approximations, which, starting from variational inference, lead to estimators with the usual guarantees. The presentation will be illustrated by examples from evolution and ecology. This is a joint work with Sophie Donnet (INRAE) and Julien Stoehr (Université Paris Dauphine).

Julien Worms (LMV, UVSQ)

Estimation de probabilités de records pour données climatiques non stationnaires

Les dernières années ont vu l'inflation inquiétante de survenue de records climatiques sur la planète et au niveau local. Dans la littérature sur les records, ce sont souvent uniquement les propriétés probabilistes de ceux-ci qui sont étudiées, et des techniques ou tests pour évaluer à quel point les données s'éloignent d'un cadre stationnaire. Mais rares sont les solutions au problème de l'estimation proprement dite des probabilités de records, qui plus est dans un cadre non gaussien et/ou non-stationnaire, qui s'impose lorsque l'on parle de records climatiques. Dans cet exposé, nous proposerons des solutions semi-paramétriques pour un tel cadre, adaptées à l'estimation des probabilités de record d'une variable climatique, et ne nécessitant pas l'observation de nombreux records. Nous présenterons des applications d'une part dans un contexte de données de températures issues de GCM à l'échelle du globe, d'autre part pour des données observationnelles pour lesquelles nous nous intéresserons au concept de record spatio-temporel. Ce sera l'occasion de discuter des avantages et des limites des techniques introduites.