

# Amastacomod - Analyse Mathématique et Statistique du Couplage Modèles-Données

Proposition d'un sujet de thèse de doctorat en Mathématiques - Statistique au LMBA

Pierre Ailliot & Emmanuel Frénod  
Laboratoire de Mathématiques Bretagne Atlantique  
UMR CNRS 6205

**Mots clés :** Agriculture Intelligente ; Analyse de Stabilité ; Assimilation de Données ; Base de Données ; Couplage Modèles-Données ; Environnement ; Intelligence Artificielle ; Intelligence Artificielle Territoriale ; Flux de Données ; Méthodes Data Driven ; Sélection de Modèles ; Simulation ; Tests

L'assimilation de données issues d'un système que l'on veut piloter (par exemple un territoire, un système environnemental ou un élevage) est essentielle. Elle permet en effet d'avoir une idée précise, à tout instant, de l'état dans lequel le système se trouve ; pour ensuite définir comment agir sur celui-ci pour l'orienter dans la direction voulue (voir **M. Nodet & A. Rousseau (2013)**).

Il est bien connu (depuis l'utilisation du filtre de Kalman pour localiser les sondes Apollo lors de leur voyage vers la lune) que pour réaliser correctement l'Assimilation des Données, il est indispensable de disposer d'un modèle mathématique ou statistique décrivant la dynamique du système à piloter. En effet, les données remontant du système étant généralement peu fiables, elles sont utilisées non pas de manière brute, mais comme un correcteur de l'évolution de la solution du modèle. Les méthodes d'assimilation de données sont utilisées dans ce contexte afin de trouver un compromis optimal au sens mathématiques entre le modèle et les observations disponibles. Le poids relatif donné au modèle et aux observations est contrôlé par des paramètres qui décrivent l'erreur du modèle et des observations (voir **P. Tando, P. Ailliot & al. (2018)**). Dans le cas des systèmes considérés pour les applications dans ce projet, vu leur complexité, les modèles mathématiques utilisés pour décrire l'évolution du système sont de haut niveau (voir **H. Flourent, E. Frénod & V. Sincholle (En cours de rédaction)** ou **E. Frénod (2017)** par exemple). Leurs propriétés (en particulier de stabilité), dont la connaissance est importante pour leur bonne utilisation, ne sont pas immédiatement accessibles. Leur établissement demande de mettre en œuvre une démarche de long terme impliquant de l'analyse mathématiques une approche statistique à base de simulation. Par exemple, une telle démarche est décrite dans **P. Ailliot, E. Frénod & V. Monbet (2006)**.

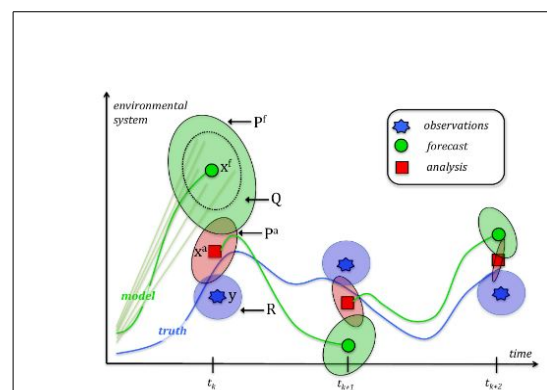


Fig 1. Principe de l'assimilation de données séquentielle

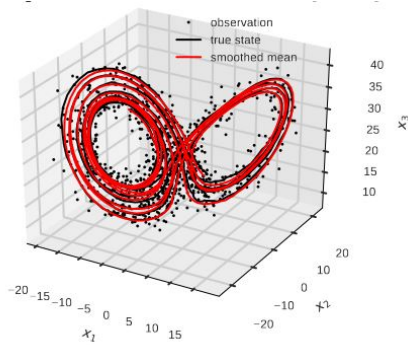


Fig 2. Assimilation dans le modèle Lorenz 63

Avec l'augmentation rapide de la quantité de données disponibles, il devient parfois possible de remplacer ou de compléter l'information apportée par le modèle mathématique en utilisant des méthodes d'apprentissage automatique (voir **R. Lguensat, P. Tandeo, P. Ailliot et al.** (2017)). La combinaison de modèles mathématiques et d'approches statistiques orientées données peut permettre d'améliorer la capacité prédictive du modèle et donc le résultat de l'assimilation de données (cf **T.T.T. Chau, P. Ailliot & V. Monbet** (2018)) tout en en réduisant la complexité numérique.

Le but de la thèse de doctorat proposée ici est de contribuer à la formalisation de la démarche mentionnée ci-dessus et à son application dans les applications ciblées. Parmi les objectifs, il y aura

- l'identification des modèles utilisés pour les applications ciblées ;
- leur compréhension ;
- la recherche de données liées aux applications ; leurs mise en base de donnée ; l'organisation de la remontée de flux de données ;
- la discussion de la pertinence des modèles pour l'assimilation de données par la construction d'un environnement de test ; leur évolution, voire leur remise en cause ;
- l'identification des propriétés à établir pour chacun d'entre eux
- l'identification de ce qui est établissable par une analyse mathématique
- identifier ce que peut apporter une couplage avec une approche statistique en utilisant de la simulation ;
- Appliquer cette démarche à certains de modèles pour certaines applications.

## Références

**P. Ailliot, E. Frénod & V. Monbet** (2006) Long term object drift in the ocean with tide and wind. *Multiscale Modelling and Simulation*, Vol 5, No 2, pp 514--531.

**T.T.T. Chau, P. Ailliot & V. Monbet** (2018). A non-parametric algorithm for reconstruction and estimation in nonlinear time series with observational errors. *En révision pour computational statistic & data analysis*.

**H. Flourent, E. Frénod** (En cours de rédaction) Existence, Uniqueness and Qualitative Properties of a PDE involved in an Artificial Intelligence for the Breeding.

**H. Flourent, E. Frénod & V. Sincholle** (En cours de rédaction) An Innovating Statistical Learning Tool Based on Partial Differential Equations, Intending Livestock Data Assimilation.

**E. Frénod** (2017) A PDE-like Toy-Model of Territory Working. *In Book "Understanding Interactions in Complex Systems - Toward a Science of Interaction"*, Cambridge Scholar Publishing, pp 37–47.

**R. Lguensat, P. Tandeo, P. Ailliot, M. Pulido & R. Fablet (2017).** The analog data assimilation. *Monthly Weather Review*, 145(10), 4093-4107.

**M. Nodet & A. Rousseau (2013)** Modélisation mathématique et assimilation de données pour les sciences de l'environnement <https://hal.inria.fr/hal-00825510/document>

**P. Tandeo, P. Ailliot, M. Bocquet, A. Carrassi, T. Miyoshi, M. Pulido, Y. Zhen (2018).** Joint Estimation of Model and Observation Error Covariance Matrices in Data Assimilation: a Review. *arXiv preprint arXiv:1807.11221*.