

Sujet Stage M2

Approche hybride physique-IA pour améliorer la compréhension de la dynamique des écosystèmes planctoniques

Encadrant : Olivier Pannekoucke (CERFACS, HDR), Elodie Martinez et Thomas Gorgues (LOPS), M. Sourisseau (DYNECO), R. Fablet (Institut Mines Télécom-Atlantique)

Contact : olivier.pannekoucke@cerfacs.fr

Lieu : CERFACS (Toulouse), 1^{er} semestre 2021

Sujet :

Le phytoplancton (micro algue peuplant la couche supérieure éclairée de l'océan) alimente la chaîne alimentaire océanique et régule les niveaux de dioxyde de carbone océanique et atmosphérique grâce à la fixation photosynthétique du carbone. Tandis que les cycles saisonniers et interannuels de la biomasse phytoplanctonique sont maintenant relativement bien caractérisés, de larges incertitudes persistent quant à sa variabilité décennale. L'indisponibilité des observations à l'échelle globale sur une série temporelle continue de plus de deux décennies a conduit la communauté scientifique à s'appuyer fortement sur des simulations couplées de modèles océaniques physique-biogéochimique pour étudier la variabilité décennale du phytoplancton. Cependant ceux-ci divergent dans la reproduction des observations décennales, en particulier les changements de régime du phytoplancton (Henson et al., 2009). L'une des raisons est la difficulté à appréhender clairement selon les régions les relations entre 1) les organismes phytoplanctoniques et la dynamique océanique (pourvoyeuse de sels nutritifs vers la surface et impactant sur la physiologie des organismes. Effet « bottom-up »), et 2) entre les organismes phytoplanctoniques et zooplanctoniques (qui sont le niveau supérieur de la chaîne alimentaire. Effet « top-down »). Bien que les oscillations climatiques, et notamment décennales, peuvent avoir un impact différent sur les biomasses phytoplanctonique et zooplanctonique (Martinez et al., 2016), les liens de cause à effet entre ces deux compartiments restent encore largement méconnus, et non transposables en équation à intégrer dans les modèles biogéochimiques.

L'essor du deep learning et des réseaux convolutionnels apparaît donc comme une opportunité pour comprendre la dynamique des écosystèmes planctoniques. Ainsi, de récents travaux ont permis d'initier cette approche en reconstruisant la biomasse phytoplanctonique uniquement à partir de variables physiques (Martinez et al., 2020; in rev). Prendre en compte l'effet top-down n'a en revanche jamais été fait. Bien qu'il soit nécessaire de construire des architectures de réseaux capable de mener cet apprentissage, il n'existe toujours pas de méthode pour construire une architecture adaptée à ce problème. Par conséquent, il est difficile de capitaliser les connaissances physiques de ces systèmes pour améliorer la découverte des processus inconnus.

L'objectif de ce stage est d'explorer le potentiel d'une méthode de génération automatique de réseaux de neurones (Pannekoucke and Fablet, 2020 ; Pannekoucke, 2020; github repository) qui permet de traduire un système d'équations aux dérivées partielles en un réseau, qu'on peut ensuite entraîner sur des données environnementales et de communautés phytoplanctoniques et zooplanctoniques. Cette approche permet de focaliser l'apprentissage de la dynamique sur les processus réellement inconnus, rendant plus efficient l'utilisation

des ressources numériques (données, calculs) et d'avoir accès aux équations régissant les interactions entre la biologie et la physique.

Le stage se déroulera au CERFACS et bénéficiera d'un environnement scientifique stimulant avec des interactions avec des biogéochimiques du LOPS (Brest) et écologue de DYNECO (Brest), et numéricien de l'Institut Mines Télécom-Atlantique et Météo-France.

Outils de développements : python, git, tensorflow, keras

Prérequis : machine learning, aucune connaissance en écologie marine n'est requise pour ce stage.

Références :

- Henson S.A., D. Raitsos, J.P. Dunne, A. McQuatters-Gollop (2009). Decadal variability in biogeochemical models: Comparison with a 50-year ocean colour dataset. *Geophysical Research Letters*, 36, L21061.
- Martinez, E., T. Gorgues, , and R. Fablet (2020). Reconstructing Global Chlorophyll-a Variations Using a Non-linear Statistical Approach. *Front. Mar. Sci.*, 7:464. doi: 10.3389/fmars.2020.00464.
- Martinez, E., A. Brini, T. Gorgues, , and R. Fablet (**in revision**). Neural network approaches to reconstruct phytoplankton time-series in the global ocean. *Remote Sensing*
- Martinez, E., D. E. Raitsos, and D. Antoine (2016). Warmer, deeper, and greener mixed layers in the North Atlantic subpolar gyre over the last 50 years. *Global Change Biology*, vol. 22, no. 2, pp. 604–612, Dec. 2016, doi: 10.1111/gcb.13100.
- Pannekoucke, O., and R. Fablet (2020). PDE-NetGen 1.0: from symbolic PDE representations of physical processes to trainable neural network representations. *Geoscientific Model Development* (to appear), doi: <https://doi.org/10.5194/gmd-2020-35>.
- Pannekoucke, O., “opannekoucke/pdenetgen: pde-netgen-GMD.” Zenodo, 2020, doi: 10.5281/ZENODO.3891101.

<https://github.com/opannekoucke/pdenetgen>