3-year PhD position on

Possible futures of the Atlantic epipelagic ecosystems under climate change: A machine learning investigation

Host institution :
Institut Pierre Simon Laplace (IPSL)
Laboratoire des Sciences du Climat et de l'Environnement (LSCE)
Orme des Merisiers Bat. 714
91191 Gif-sur-Yvette cedex
France

Contacts : Marion Gehlen (marion.gehlen@lsce.ipsl.fr) & Mathieu Vrac (mathieu.vrac@lsce.ipsl.fr)

Summary
The marine environment will be profoundly modified by climate change. Changes in physical and chemical conditions will cascade on ecosystems with consequences that remain difficult to predict. Numerical models are key tools for the exploration of possible future states of marine ecosystems. The main objective of this thesis is to constrain projections of changes in plankton ecosystems in response to climate change by leveraging on extensive observational data sets and biogeochemical/ecological models. This project combines machine learning and numerical modeling with the goal to narrow down uncertainties in predictions of future ecosystem states. It seeks to identify novel causal relationships between environmental drivers and ecosystem dynamics in historical data. These relationships will be verified in an ensemble of biogeochemical-ecological model simulations of varying complexity, first over the period covered by observations and finally over the next 80 years. Beyond assessing the skill of numerical models to reproduce observed relationships, particular attention will be given to the screening of model output for novel emerging ones.

Context and objectives
Biogeochemical models are now integral parts of Earth system models. They allow quantifying changes in biogeochemical cycles in response to climate change and associated feedbacks to the Earth system. These models have gained in complexity over the past two decades to become simplified representations of first levels of the marine foodweb. Since they are embedded in ocean general circulation models, they allow exploring possible futures of marine ecosystems over the spectrum of spatial/temporal scale and abiotic/biotic interactions forming the pelagic seascape. The development of models went along with an important increase in observing capacities driven by technological innovations and the deployment of Earth observation networks. The potential of these novel observations to further our understanding of drivers of ecosystem variability and to disclose novel causal relationships, but also to drive the development of the next generation of marine biogeochemical/ecosystem components of Earth system models only starts to be exploited.

The main objective of this thesis is to constrain projections of changes in plankton ecosystems in response to climate change by leveraging on extensive observational data sets and biogeochemical/ecological models. To reach this goal, it relies and a step-wise approach combining methods known as "Knowledge-Discovery through Data mining" (KDD) – such as the δ-MAPS method (Fountalis et al, 2018 ; Bracco et al., 2018) – and numerical modeling.

The approach will consist in: (1) the implementation of a machine learning method (e.g., δ-MAPS) suited to the biogeochemical data; (2) the application of the method for the identification of patterns, as well as frequencies of variability in observations of physical and biogeochemical drivers of ecosystem dynamics, as well as proxies of ecosystem state (e.g. remote sensing derived plankton functional types, net primary production) and in situ data on community composition (e.g. species composition, biogeochemical functions) at monthly to multi-annual time scales; and (3) the verification of these relationships in output from biogeochemical/ecological models of varying complexity in simulations covering the historical period, as well as future high and low emission scenarios. The biogeochemical model PISCES will be used for this study (Aumont et al., 2015; Séférian et al., 2020; Kwiatkowski et al., 2020). Several versions of variable complexity have been developed from a common standard version (e.g. Kwiatkowski et al. 2018) which allows probing the effect of model complexity on model skill in an unified setting.

If successful, the use of KDD to derive driver-response relationships from large data sets and their verification in model output will allow to evaluate model skill beyond current approaches. Its application to output from a biogeochemical model of varying complexity should shed light on the link between model complexity and model skill in representing ecosystem dynamics. Finally, the application of the approach to future projections, should allow to identify the emergence of novel relationships in response to climate change and narrow down the uncertainty in projected ecosystem impacts.

Required:

- A Master or engineer diploma in science is needed;
- Excellent skills in programming. A past experience in machine learning would be desirable but is not mandatory;
- Good knowledge in statistics and/or in ecosystem modelling.

Starting date: fall 2021
Duration: 3-year PhD contract

To apply: Send CV + motivation letter + names and coordinates of 2 reference persons.

References

Aumont et al., 2015, 10.5194/gmd-8-2465-2015
Bracco et al., 2018, https://www.nature.com/articles/s41612-017-0006-4
Fountalis et al, 2018, https://doi.org/10.1007/s41109-018-0078-z
Kwiatkowski et al., 2018, https://doi.org/10.1002/2017GB005799
Kwiatkowski et al., 2020, https://www.biogeosciences.net/17/3439/2020/
Séférian et al., 2020, https://doi.org/10.1007/s40641-020-00160-0