Postdoctoral fellowship (2 years) in Machine Learning for Computational Oceanography

Keywords: 3D ocean dynamics, inverse problems, data challenge, large scale Machine Learning, Gaussian Processes, kernel methods, hyperparameterization,

Context

IMT Atlantique is looking for a postdoctoral fellow for 2 years starting as soon as possible before December 2025. The position is based on the Brest campus of the school. The fellow will join the Mathematical and Engineering Department of IMT Atlantique, and conduct his research wihtin the Odyssey Team Project studying Ocean Dynamics using data driven approaches, namely ML and AI techniques. The project is a collaboration with the Laboratoire d'Océanographie Physique et Spatiale (LOPS) of Univ. Brest.

Covering more than 70 % of the surface of the Earth, the oceans play key roles for the regulation of the Earth climate (e.g., climate change) as well as for human societies (e.g., marine resources and maritime activities). Despite ever-increasing development of simulation and observation capabilities leading to ocean big data, our ability to understand, reconstruct and forecast ocean dynamics remains limited.

Altimetry data and other geophysical measurements allow an improved understanding of ocean dynamics thanks to the various types of data acquired at the surface and inside of the ocean by different in situ sensors and satellite missions. A large number of different sensors exist, ranging from in situ floats to satellite data. However, even with all these acquisitions, the observation coverage of the 3D ocean remains very sparse, making prediction tasks challenging. In such a setting, it is also crucial to be able to produce uncertainties associated to the predictions, or even propose different evolution scenarios in ambiguous configurations.

In this context, a national initiative within the Programme Prioritaire de Recherche (PPR) Océan et Climat aims at building data challenges centered around the exploitation of such datasets, and easy benchmarking of AI-based solutions [1]. One of those challenges to be launched by the end of 2025 will be centered around the probabilistic short term forecasting of oceanic variables at the global scale in the 3D ocean from sparse measurements. AI-based methods for this type of problems are starting to emerge but are not so mature as e.g. in weather forecasting, and in particular are usually deterministic [2, 3].

Thus, the objective of this postdoctoral project is to develop new AI native models to learn and propagate ocean dynamics, with an emphasis on:

- Handling incomplete and noisy data
- A probabilistic formulation of the problem, i.e. not just learning a mean value over space and time of the parameters of interest, but also to generate and propagate uncertainties in the relevant oceanic variables over space in time.
- A method that scales to global oceanic states ($\approx 10^6$ variables or more)

To this end, we propose to merge Gaussian Process (GP) regression [4], which is the basis of most operational techniques for sparse oceanic data interpolation and forecast (Optimal Interpolation) and Machine/Deep Learning techniques. These kernel-based methods take advantage of the closed form Gaussian solution of the GP providing

uncertainties at little cost. Despite these appealing properties, their performance has been modest because of limited expressivity due to ad-hoc choices of rigid kernels and associated parameters. We propose to enhance this class of probabilistic methods by leveraging flexible AI-based hyperparameterizations of these kernel methods to train data-driven priors for interpolation/forecasting from complete (e.g. simulation) data.

Candidate

The candidate should hold a PhD in signal or image processing, machine learning, remote sensing or related fields. We are looking for strong candidates with the following skills:

- Machine/Deep learning, signal and image processing, applied mathematics, numerical methods
- Programming in Python (Numpy, scipy, matplotlib...), especially Pytorch
- Inverse problems in imaging, Bayesian Modeling, Kernel Methods
- Curiosity for or experience in applications to quantitative oceanography will be appreciated.

Contact

Lucas Drumetz, Associate Professor, IMT-Atlantique, Lab-STICC, Odyssey Team Project
Florian Sévellec, Director of Research CNRS, LOPS, Odyssey Team Project

Please send a CV and a cover letter to lucas.drumetz@imt-atlantique.fr and florian.sevellec@univ-brest.fr.

References

[1] J. E. Johnson, Q. Febvre, A. Gorbunova, S. Metref, M. Ballarotta, J. Le Sommer, *et al.*, "Oceanbench: The sea surface height edition," *Advances in Neural Information Processing Systems*, vol. 36, pp. 78275–78295, 2023.

[2] A. E. Aouni, Q. Gaudel, C. Regnier, S. Van Gennip, M. Drevillon, Y. Drillet, and J.-M. Lellouche, "Glonet: Mercator's end-to-end neural forecasting system," *arXiv preprint arXiv:2412.05454*, 2024.

[3] R. Fablet, B. Chapron, L. Drumetz, E. Mémin, O. Pannekoucke, and F. Rousseau, "Learning variational data assimilation models and solvers," *Journal of Advances in Modeling Earth Systems*, vol. 13, no. 10, p. e2021MS002572, 2021.

[4] C. E. Rasmussen, "Gaussian processes in machine learning," in *Summer school on machine learning*, pp. 63–71, Springer, 2003.