Stage de fin d'étude

Intégration de la variabilité temporelle dans la prédiction des communautés phytoplanctoniques à l'aide du deep learning.

Domaine de recherche Science des données, écologie numérique, télédétection

Employeur Sorbonne Université, Institut des sciences du calcul et des données

Lieu : LOV, 181 chemin du Lazaret Villefranche-sur-Mer, France

Type de contrat : 6 mois

Statut de l'emploi : Temps plein

Date de début : À voir en fonction de la disponibilité du candidat

Contexte Scientifique

Le **phytoplancton** regroupe les micro-organismes marins dérivant au gré des courants et capables de produire de la matière organique grâce à la photosynthèse.
Il joue un rôle essentiel dans le fonctionnement du système Terre : il contribue à environ **la moitié de la photosynthèse planétaire**, participe au **stockage du carbone** à l'échelle climatique via la **pompe biologique**, et constitue la **base des réseaux trophiques marins**.

Mieux comprendre la **répartition** et la **diversité** du phytoplancton, en lien avec les conditions physiques et biogéochimiques de son environnement, est donc fondamental. Cela permet d'estimer plus finement sa biomasse, son rôle dans les cycles biogéochimiques et d'anticiper les réorganisations induites par les changements climatiques et océaniques globaux.

Cependant, l'observation directe du phytoplancton à grande échelle demeure limitée, car elle repose sur des campagnes en mer et des prélèvements coûteux. Les données **HPLC** (High Performance Liquid Chromatography), qui permettent d'analyser les pigments photosynthétiques pour quantifier précisément les communautés phytoplanctoniques [1][2][3], offrent une référence fiable, mais restent **coûteuses et spatialement limitées**.

Pour étendre ces observations ponctuelles à l'échelle globale, les **satellites d'observation de la couleur de l'océan** constituent un outil clé. Les algorithmes empiriques actuels permettent d'estimer la composition du phytoplancton à partir des réflectances optiques [4][5][6], mais ils présentent encore une marge d'erreur importante et ne capturent pas toujours la complexité **spatiale et temporelle** des structures océaniques (fronts, tourbillons, etc.) [7].

Description du Stage

Ce stage s'inscrit dans la continuité d'une thèse centrée sur la prédiction des Phytoplankton Functional Types (PFTs) à partir de la couleur de l'eau et de variables satellitaires à l'aide de méthodes de deep learning [8].

L'objectif principal est d'améliorer la robustesse et la précision des prédictions en intégrant explicitement **la dimension temporelle des données au moyen de méthodes avancées de deep learning.** Concrètement, cela consiste à exploiter les images satellites des jours ou des semaines précédentes pour estimer la **quantité et la composition du phytoplancton à un instant t donné**. En effet, la croissance du phytoplancton est un **processus dynamique**, et son état actuel dépend fortement des états antérieurs du système.

Un pipeline opérationnel est déjà en place :

- les données satellitaires et in situ ont été extraites et prétraitées,
- des modèles de type MLP (Multi-Layer Perceptron) ont été développés et évalués,
- une première étude de la sensibilité, en termes de qualité des prédictions, aux variables d'entrée et à l'architecture du modèle a été menée,
- une seconde étude, sur l'intégration de la dimension spatiale dans les données d'entrée, est en cours.

Le stage consistera à poursuivre et à approfondir ces travaux, notamment en :

- explorant de **nouvelles architectures de deep learning** (par ex. RNN) permettant de mieux **capturer la structure temporelle des données ;**
- évaluant l'impact de cette intégration sur la qualité des prédictions de PFTs ;
- affinant l'interprétation des modèles pour mieux comprendre les liens entre conditions environnementales et composition des communautés phytoplanctoniques.

Ce travail contribuera directement à l'amélioration des produits satellitaires de distribution phytoplanctonique à l'échelle globale, utiles pour le suivi de la santé des écosystèmes marins et la gestion durable des ressources océaniques.

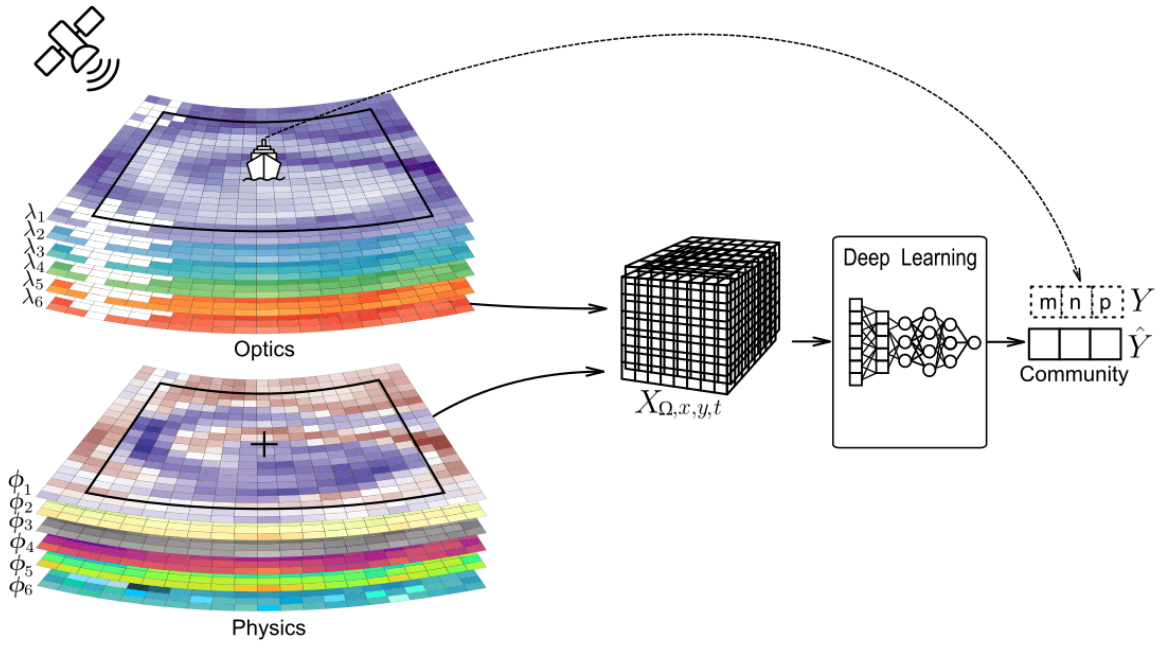

Fig. 1 : Schéma de l'objectif du stage. Intégration de la variabilité temporelle permettant la capture complète du contexte des structures océaniques caractérisant l'environnement où se développent les communautés phytoplanctoniques (PFTs).

Objectifs du stage

- Compréhension de la problématique **biologique**.
- Étude des relations entre les **propriétés optiques de l'eau et les PFTs**.
- Prise en main et optimisation du pipeline existant, de la sélection et du traitement des données jusqu'à leur exploitation dans un modèle de **deep learning**.
- Analyse de la **variabilité spatio-temporelle** : quantifier l'impact de l'intégration des dimensions spatiale et temporelle sur la prédiction des PFTs.
- Mise en forme et valorisation des résultats pour une utilisation dans **des travaux de recherche** à publier.

Compétences Recherchées

- Intérêt marqué pour la **recherche** et les **sciences marines**.
- Maîtrise du langage **Python**.
- Connaissances en **deep learning** (PyTorch).
- Des notions en **écologie**, **biologie marine** ou **océanographie** seraient un atout.

Les Atouts du Stage

- Encadrement personnalisé : travail en étroite collaboration avec une doctorante, disponible pour accompagner et vulgariser les concepts et les méthodes.
- Localisation : Laboratoire en bord de mer et possibilité d'avoir accès à un logement étudiant du laboratoire.
- Vie étudiante dynamique : nombreuses activités organisées au sein du laboratoire.
- Formation continue : accès à des conférences hebdomadaires sur divers sujets en océanographie, permettant de découvrir divers aspects du domaine et d'élargir ses connaissances.

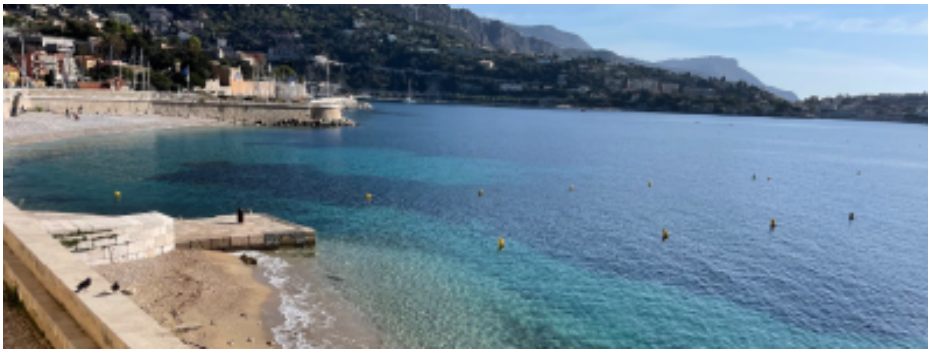

Fig. 2 : La jetée du Laboratoire d'océanographie de Villefranche.

Documents requis pour la candidature

1. Lettre de motivation indiquant vos intérêts de recherche et votre motivation pour le sujet.

2. Curriculum vitae à jour.

3. Noms et coordonnées d'une à deux personnes pouvant fournir des références.

4. Bulletin de M1 et un détail des unités d'enseignement (UE) actuellement suivies en M2.

Les candidats intéressés peuvent contacter enza.labourdette@imev-mer.fr pour plus d'informations et doivent envoyer les documents requis pour la candidature par e-mail à la même adresse avec pour objet « Stage Candidature 2026».

Encadrement

Encadrante Principale

- ***Enza Labourdette***, doctorante, LOV. Enza est une doctorante de troisième année ayant une double formation en mathématiques appliquées et en océanographie. Elle a conçu et développé tout le code de notre projet.

Encadrants Secondaires

- ***Jean-Olivier Irisson***, professeur, LOV. J-O est un écologue computationnel qui étudie le plancton à l'aide de méthodes d'imagerie quantitative et de science des données, en particulier les réseaux convolutifs. Il s'intéresse particulièrement à l'influence des processus à méso et subméso-échelle (100 m-100 km))sur le plancton.

- ***Raphaëlle Sauzède***, ingénieure de recherche, IMEV. Raphaëlle calcule les produits océaniques mondiaux en fusionnant les informations satellitaires et in situ grâce à des réseaux profonds. Elle possède également une expertise sur les données HPLC, l'imagerie satellitaire et la distribution du phytoplancton.

- ***Lokmane Abbas-Turki***, professeur associé, LPSM. Lokmane est l'un des premiers universitaires à avoir programmé sur des GPUs et utilise cette compétence pour résoudre des problèmes statistiques complexes, en particulier dans le domaine financier.

Bibliographie

[1] Vidussi, F., H. Claustre, B. B. Manca, A. Luchetta, and J. C. Marty (2001), Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter, *J. Geophys. Res.*, **106**(C9), 19,939–19,956, doi:10.1029/1999JC000308.

[2] Uitz, J., H. Claustre, A. Morel, and S. B. Hooker (2006), Vertical distribution of phytoplankton communities in open ocean: An assessment based on surface chlorophyll, *J. Geophys. Res.*, 111, C08005, doi:10.1029/2005JC003207.

[3] Brewin, R. J. W., S. Sathyendranath, T. Hirata, S. J. Lavender, R. Barciela, and N. J. Hardman-Mountford(2010), A three-component model of phytoplankton size class for the Atlantic Ocean, *Ecol. Modell.*, **221**, 1472–1483, doi:10.1016/j.ecolmodel.2010.02.014.

[4] Brewin, RJW, Sathyendranath, S, M ller, D, Brockmann, C, Deschamps, P-Y, Devred, E, Doerffer, R, Fomferra, N,Franz, B, Grant, M, Groom, S, Horseman, A, Hu, C, Krasemann, H, Lee, ZP, Maritorena, S, Mélin, F, Peters, M,Platt, T, Regner, P, Smyth, T, Steinmetz, F, Swinton, J Werdell, J, White, GN, III (2015a) The Ocean Colour ClimateChange Initiative: III. A round-robin comparison on in-water bio-optical algorithms. Remote Sensing ofEnvironment, 162: 271-294. http://dx.doi.org/10.1016/j.rse.2013.09.016

[5] Morel, A.: Optical properties of oceanic case 1 waters revisited. In Ocean Optics XIII (Vol. 2963,pp. 108-115). International Society for Optics and Photonics, February 1997.

[6] O'Reilly, J.E., Maritorena, S., Siegel, D.A., O'Brien, M.C., Toole, D., Mitchell, B.G., Kahru, M.,Chavez, F.P., Strutton, P., Cota, G.F. and Hooker, S.B.: Ocean color chlorophyll-a algorithms forSeaWiFS, OC2, and OC4: Version 4. SeaWiFS postlaunch calibration and validation analyses, Part,3, 9-23, 2000.

[7] Xi H., Losa S. N., Mangin A., Soppa M. A., Garnesson P., Demaria J., Liu Y., Hembise Fantond'Andon O., Bracher A. (2020) Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. Remote Sensing of Environment, 240, 111704,

[8] Labourdette, E., Irisson, J. O., Sauzede, R., and Abbas Turki, L.: How can convolutional neural networks help account for the impact of mesoscale ocean structures on phytoplankton distributions ?, One Ocean Science Congress 2025, Nice, France, 3–6 Jun 2025, OOS2025-1073, https://doi.org/10.5194/oos2025-1073, 2025.