Sujet de thèse

Vers une modélisation parcimonieuse de la biogéochimie océanique basée sur des indicateurs clés pour des projections décennales robustes en Méditerranée

Towards parsimonious data-driven modelling of the ocean biogeochemistry for robust decadal projections in the Mediterranean Sea

Contexte scientifique

Pour répondre aux questions scientifiques toujours plus nombreuses liées aux cycles biogéochimiques (cycle du carbone notamment) et à la dynamique du réseau planctonique dans l'océan, ou pour tenter de reproduire certaines caractéristiques biogéochimiques observées in situ, les modèles biogéochimiques (ce qui désigne ici aussi bien des modèles basés sur des classes de taille, des groupes fonctionnels, des traits ou des espèces) ont vu leur complexité augmenter quasi continûment ces dernières décennies. Cette complexification prend des formes différentes selon les modèles, avec un nombre de variables très conséquent pour les modèles basés sur la diversité du vivant incluant plusieurs dizaines de groupes planctoniques, ou une augmentation du nombre de processus représentés et/ou des formulations qui les représentent. En outre, le fait de modéliser la flexibilité de la stœchiométrie au sein des organismes planctoniques s'accompagne aussi d'un accroissement significatif du nombre de variables d'état. De plus, ces modèles biogéochimiques sont couplés à (ou forcés par) des modèles de circulation marine, voire intégrés à des chaînes de modèles incluant aussi une représentation explicite d'autres composantes du système terre. La réalisation de simulations multi-décennales (notamment des projections climatiques) à l'aide de ces modèles intégrés de grande complexité induit des difficultés de deux types :

(i) elle nécessite des ressources de calcul très conséquentes, ce qui peut s'avérer pénalisant pour quantifier des incertitudes (notamment via des approches ensemblistes), pour explorer une diversité de scénarios pour le 21ème siècle, ou pour toute autre application nécessitant de réaliser des simulations en nombre,

(ii) un niveau de complexité excessive peut s'avérer inutile selon l'objectif poursuivi, d'autant que l'empreinte carbone des simulations produites par ces modèles complexes n'est pas négligeable.

Dans ce contexte, l'identification du « juste niveau » de complexité, notamment pour la composante biogéochimique de ces modèles intégrés, devient un enjeu majeur. Ce niveau de complexité et la simplification qui en est attendue dépendent nécessairement de la question scientifique étudiée ou des indicateurs écologiques à estimer, et donc des variables et des échelles de temps d'intérêt, mais également de la quantité d'information contenue dans les observations disponibles. La littérature propose déjà un nombre significatif de méthodes de simplification de modèles dynamiques qui consistent soit à (i) décomposer le réseau trophique initial en sous-modules ayant peu d'interactions entre eux, (ii) agréger des variables d'état ou des processus de manière à obtenir un modèle réduit dont la structure peut potentiellement différer du modèle original, (iii) exploiter la présence potentielle d'échelles de temps différentes au sein des processus en présence soit de manière classique « manuelle », ou automatique, ce qui dans tous les cas transforme le système d'équations initial, (iv) simplifier les expressions mathématiques des équations d'origine, ou encore (v) identifier les variables et les processus qui gouvernent la dynamique du système et supprimer les variables ou les processus jugés peu influents, ce qui peut se faire à l'issue d'une étude de sensibilité, d'une optimisation, ou d'autres méthodes.

Objectifs de la thèse

Cette thèse s'inscrit dans le projet MEDIATION[1] qui a notamment pour objectif de mettre au point des développements méthodologiques destinés à (i) définir le niveau de complexité nécessaire des chaînes de modélisation pour évaluer la réponse des écosystèmes à divers scénarios, (ii) quantifier l'incertitude des modèles et des projections qu'elles fournissent, et (iii) diminuer les temps d'exécution pour pouvoir explorer un grand nombre de scénarios.

L'objectif principal de cette thèse est dans un premier temps d'explorer deux voies principales pour réduire les temps de calcul du modèle biogéochimique Eco3M-MED (Baklouti et al., 2021) lorsqu'il est utilisé dans une chaîne de modélisation telle que celle déjà mise en place pour la Méditerranée (e.g. Pagès et al., 2020) ou dans les versions enrichies d'autres maillons du système terre telles qu'envisagées dans le projet MEDIATION. Le ou les modèles plus parcimonieux issus de cette première étape devraient permettre d'estimer les incertitudes par exemple au moyen de méthodes ensemblistes (Garnier et al., 2016) sur quelques indicateurs clés qui seront identifiés dans le cadre de cette thèse.

- La première voie a pour objectif d'élaborer une ou plusieurs versions simplifiées du modèle Eco3M-MED, dont l'une au moins permettrait de conserver les fondements mécanistes du modèle et ses capacités explicatives. L'objectif est de se baser sur une ou des méthodes existantes, sans exclure d'y apporter des modifications pour la rendre mieux adaptée à l'objectif poursuivi. Dans le meilleur des cas, la méthode adoptée sera suffisamment

1 1Le projet MEDIATION « Développements méthodologiques pour un jumeau numérique robuste et efficace de l'océan » est porté par L. Debreu (INRIA), F. Dumas (SHOM) et P. Marchesiello (IRD) et est financé par le PPR-OCEAN pour 5 ans (2022-2027).

générique pour pouvoir bénéficier à d'autres modèles biogéochimiques, et notamment les autres modèles mis en jeu dans le projet MEDIATION.

- La seconde voie explorera la possibilité d'optimiser le temps de calcul lié au transport, par le modèle hydrodynamique, des variables biogéochimiques associées à un même organisme. Dans les modèles à stoechiométrie variable actuels, ces concentrations sont advectées de manière indépendante par le modèle hydrodynamique, ce qui a pour conséquence d'induire une inconsistance dès lors que le schéma d'advection n'est pas linéaire, en plus d'accroître inutilement les coûts de calcul. L'objectif est donc de proposer, à partir de pistes déjà ébauchées, un schéma d'advection des concentrations associées à un même organisme qui soit de même ordre que le schéma initial mais qui combine gains de consistance et de temps de calcul.

Mise en œuvre et outils

Les recherches réalisées dans le cadre de cette thèse utiliseront le modèle biogéochimique Eco3M-Med sous différentes configurations, i.e. couplé à un modèle hydrodynamique (NEMO ou CROCO selon les configurations) dans une configuration 1D verticale et dans une configuration 3D à l'échelle du bassin méditerranéen. En amont de la simplification du modèle, une première étape consistera en l'identification d'indicateurs clés observables à partir de données de terrain, et en l'évaluation du modèle d'origine sur ces indicateurs afin d'y apporter les modifications nécessaires le cas échéant. La programmation de la méthode de simplification et du schéma d'advection se feront en Fortran 90-95 alors que le langage Python sera utilisé pour le post-traitement des simulations.

Attendus et perspectives

Cette thèse devrait permettre d'obtenir un ou des modèles significativement plus rapides dans leur exécution tout en conservant, au moins pour l'un d'entre eux, leur rationalité biogéochimique et les bases mécanistes du modèle d'origine. Les performances de ce modèle seront évaluées à l'aide de données acquises en Méditerranée sur une période d'une à deux décennies sur la période 2000-2020 et comparées à celles du modèle d'origine. Si celles-ci sont suffisantes, ce modèle devrait être utilisé à terme pour les projections climatiques à l'échelle du bassin méditerranéen envisagées dans le projet MEDIATION et incluant une quantification des incertitudes.

Contexte de travail

Le(la) candidat(e) sera encadré(e) par M Baklouti (AMU/MIO) et P Brasseur (CNRS/IGE). Il/elle sera principalement localisé au MIO (campus de Luminy, Marseille), et effectuera des séjours à l'IGE (Grenoble). Une étroite collaboration est envisagée avec E Arnaud, A Vidard, E Blayo, L Debreu (INRIA/AIRSEA) pour le premier volet et avec P Marchesiello (IRD/LEGOS) et F Dumas (SHOM) pour le second.
La thèse débutera au 1er octobre 2023.

Profil recherché

Le ou la candidate retenu(e) aura bénéficié d'une formation en océanographie (ou en sciences de l'environnement, le cas échéant), et/ou aura une bonne maîtrise de méthodes numériques (réduction de modèles, analyse de sensibilité, résolution d'EDPs). Des connaissances/expériences en modélisation numérique 3D en océanographie (ou en sciences de l'environnement), et une bonne pratique des langages Fortran et Python ainsi que de l'environnement linux sont également attendues.

Contacts

melika.baklouti@univ-amu.fr, pierre.brasseur@univ-grenoble-alpes.fr

Références

Baklouti M, Pagès R, Alekseenko E, Guyennon A, Grégori G (2021) On the benefits of using cell quotas in addition to intracellular elemental ratios in flexible-stoichiometry Plankton functional type models. Application to the Mediterranean Sea, Progress in Oceanography, 197, 102634.

Garnier F., Brankart J.-M., Brasseur P. and Cosme E., 2016: Stochastic parameterizations of biogeochemical uncertainties in a 1/4° NEMO/PISCES model for probabilistic comparisons with ocean color data, J. Mar. Systems, 155, 59-72, https://doi.org/10.1016/j.jmarsys.2015.10.012

Pagès R, Baklouti M, Barrier N, Ayache M, Sevault F, Somot S, Moutin T (2020). Projected Effects of Climate-Induced Changes in Hydrodynamics on the Biogeochemistry of the Mediterranean Sea Under the RCP 8.5 Regional Climate Scenario. Frontiers in Marine Science, 7, 10.3389/fmars.2020.563615