Data-driven MOdeling and sampling to MOnitor PARticle origins in deep sediment traps (Biological Carbon Pump)

MOMOPAR

Research teams : LEMAR – LOPS/IUEM (Université Bretagne Occidentale) ; IMT Atlantique
Techonople Brest Iroise, 29280 Plouzané, France
Supervisor : Laurent Mémery (LEMAR) , laurent.memery@univ-brest.fr ; (33) 2 98 49 88 97
Co-supervisors : Ronan Fablet (IMT) : ronan.fablet@imt-atlantique.fr; Jonathan Gula (LOPS) : Jonathan.gula@univ-brest.fr

Scientific context

The Carbon Pump represents the processes that regulate the absorption and storage of atmospheric CO_2 in the deep ocean. CP plays a major role in climate and biogeochemical cycles (C, O_2, nutrients, ..). The biological part of this pump reduces unperturbed atmospheric CO_2 by 35 to 50%. It is driven by photosynthesis at the ocean surface, which creates particles that are exported by gravity in the deep ocean (export production). These particles differ in size, composition and vertical velocity, and are partially remineralized in the ocean.
Export production is historically observed with sediment traps from the upper thermocline down to several thousand meters. The interpretation of these data is based on a very strong hypothesis: the ocean is 1D vertical, which means that the trajectories of the particles, produced by the local primary production, are vertical. In fact, the ocean circulation is three-dimensional, and is further characterized by energetic structures at different scales: the origin of the particles that reach a deep sediment trap is thus spread over a domain that depends on the horizontal ocean circulation at all scales. In addition, due to vertical velocities constrained by frontal dynamics or topographic effects, the travel time of "identical" particles (same sinking velocities) can vary (which regulate the variability of the production date of the particles collected by the traps, potentially important during transient events such as blooms).
International programs, such as the American program EXPORTS or the French program APERO, focus on the fate of carbon in the mesopelagic layer (200-2000 m) and on the estimation of carbon stored in the deep ocean. They are based on ambitious cruises that aim to characterize the carbon flux in the deep ocean, i.e. sinking particle dynamics, from regional to small-scales. In addition, these cruises take place near historical sediment trap moorings (more than 20 years of temporal coverage), such as the British PAP station in the northeast Atlantic. It is therefore crucial to estimate the domain that defines the "funnel" feeding the sediment traps, as this domain must be sampled during the cruise to obtain a consistent picture of carbon fluxes.
In this context, the general objective of this PhD is to investigate data-driven and learning-based framework to investigate the monitoring and sampling of particle fluxes in the deep ocean from the synergies between available in situ, satellite-derived, simulation and reanalysis datasets.

Scientific issues

- *Q1. Can we learn a data-driven model for the reconstruction of particle fluxes in the deep ocean using multimodal observation data (e.g., satellite-derived sea surface fields, in situ data from BioGeoChemical (BGC) Argo floats) ?*

  To address the considered general objective, we aim to explore and develop data-driven and learning-based frameworks, especially deep learning schemes [1]. Formally, the considered issues will be stated as the prediction or reconstruction of the sea surface distribution of the origin of particles collected in sediment traps in the deep ocean (typically 2000 m depth) given some input data, such as sea surface fields as well as vertical profiles. Following recent studies, this reconstruction problem will be stated as a learning-based regression issue using deep neural networks [1]. Both standard neural networks, such as CNN (Convolutional Neural Network), and neural schemes associated with variational formulations [2] will be considered. The latter appear well suited to deal with irregularly-sampled and multimodal data and may deliver a greater interpretability in terms of underlying physical processes.

- *Q2. To which extent sea surface conditions constrain carbon fluxes to the deep ocean ?*
  Satellite observations (altimetry, sea surface temperature, ocean color) provide synoptic images of surface patterns in terms of dynamics and biological activity at meso and sub-meso scales. Since much of the physical and biological forcing (primary production) of the ocean takes place at the surface, and since satellite data are global, synoptic, and almost continuous, it is arguable whether it would not be possible to constrain the deep ocean solely by surface observations. Specifically, one of the hypotheses that several ongoing programs want to test is that satellite data can constrain not only the flux of carbon exported from the surface layer, but also this flux of carbon at deeper levels. There are many hypotheses and simplifications behind this assertion, mainly in terms of ecosystem functioning in the meso-pelagic layer, which modifies particle characteristics and size spectra. This is the reason why, in a first step, the question that is addressed here is simpler and easier to deal with. Is it possible to deduce particle distributions at depth from the information provided by satellites assuming that the particles are conservative (only physical processes are taken into account)? To address this question, we will apply the methodology developed in Q1 while considering as input data only sea surface information.

  For calibration purposes (training phase), we will exploit re-analyses datasets (e.g., CMEMS products) and high-resolution numerical simulations (< 2km resolution) using an OSSE (Observing System Simulation Experiment) approach, i.e. using simulating the satellite-derived-like observations from the re-analyzed fields. The learnt model(s) will be applied to satellite-derived sea surface fields, re-analyzed fields as well as high-resolution numerical simulations. A quantitative and qualitative statistical analysis will be carried out to diagnose the differences and similarities between the resulting predictions of particles' origins at sea surface

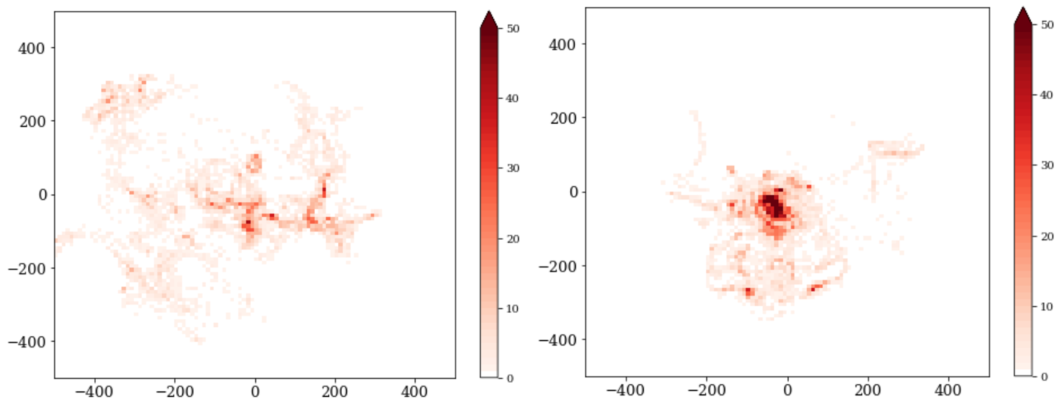
- *Q3. Can we determine optimal sampling strategies for given observation resources (e.g., BGC Argo floats and deep gliders, with satellite data) so that we best monitor particle fluxes in the deep ocean ?*
  Because of the complexity of the problem and the different scales involved, there is no assurance sea surface conditions alone constrain carbon fluxes to the deep ocean (Q2), which would be a result in itself. In this case, a diagnosis explaining this failure will have to be made. One of the reasons would certainly come from a loss of the surface signal with depth, i.e. a partial decoupling between surface and deep ocean dynamics. A next step will be to add more information on the state of the deep ocean, information coming from BGC Argo autonomous floats or deep gliders, equipped with UVPs (which give the distributions of large particles), which can be simulated in the 3D circulation model. This is of key importance in the design of a survey capable of adequately sampling the particle funnel. Using variational frameworks developed in Q1, we aim here to determine some optimal sampling strategies for vertical profiles as given by BGC Argo profilers or deep gliders. Formally, within the considered variational framework, this issue may be stated as the learning of some observation operator

under sparsity constraint. As case-study, we will design OSSEs using re-analyses or high-resolution numerical simulations as in the training phase of Q2.

<u>Methodology and workplan</u>

As described in the previous section, this PhD will rely on a general learning-based and data-driven methodology (in situ, satellite, 3D models). As such, for the targeted questions (in particular Q1 and Q2), the acquisition of all these data will be an important initial task.

We will benefit from preliminary results obtained in the framework of the PhD of Lu Wang supervised by J. Gula and L. Memery. These results suggest a correlation between the dispersion and surface distribution structures of the particle source, and the small-scale dynamics and presence of eddies in the water column. Using a high-resolution (~2 km) model of the North Atlantic, these results were obtained from backward simulations of thousands of particles in the region of the PAP station, considering traps at different depths (1000, 2000, 3000 and 4000 m), with different sinking velocities of the particles (from 20 to 200 m/day). The results of such simulations in different dynamical field configurations will be used as the basis of the statistical modeling (learning). The algorithm will link the structure of the field at the ocean surface (EKE, presence of eddies with their characteristics - cyclone vs anticyclone, size, lifetime, etc...), which can be obtained by altimetry, and the source distribution of the particles reaching the sediment traps.



Probability density distribution of the particle source feeding a sediment trap at 2000m (central position) for a sinking velocity of 20m/day (Lu Wang thesis). Left: winter 2006 situation; right: winter 2008 situation. The difference in distribution comes from the different eddy dynamics during these two periods. The thesis seeks to link the origin of the particles at the surface to the dynamics at the surface and in the water column.

The proposed workplan is as follows:

- M1-3 Review of the state-of-the-art
- M1-6 Data collection and design of the considered OSSE-like experiments, including the creation of different training and test datasets comprising sea surface conditions and simulated distribution maps of particle origins using backward Lagrangian simulations.

Different datasets will be issued from high-resolution numerical simulation, Mercator renalyses and satellite-derived data.

- M2-18 Focus on Q2 and associated methodological developments in Q1
- M18-30 Focus on Q3 and associated methodological developments in Q1
- M30-36 Preparation of the PhD defense, with the writing of scientific papers

Results and perspectives

The three main outcomes expected from this study are :

- a data-driven method to relate surface conditions (from satellite observations) to the source distribution of particles reaching sediment traps at given depths at the PAP station
- a better description and understanding of the impact of small- and medium-scale ocean circulation on the fate of sinking particles in the deep ocean
- assistance in the design of surveys capable of adequately sampling the particle funnel, considering, if necessary, also BioArgos floats and deep gliders (OSSE : Observing System Simulation Experiment).

This approach could then be :

- generalized in different regions of the ocean, more precisely in long-term observation sites with sediment traps, such as DYFAMED in the Mediterranean or US PAP in the North Pacific ;
- applied to establish the evolution of the funnel over time at these stations using re-analyses, such as those from Mercator, and to link the source of the particles to the primary production at the surface ;
- exploited to characterize how biogeochemical models and associated assimilation products differ or converge in the way they relate sea surface conditions and deep carbon fluxes.

It can also be pointed out that, if the approach is not conclusive, even taking into account the additional information in the water column, this study will have rejected the basic hypothesis (surface observations may constrain carbon fluxes at depth), although the processes have been greatly simplified (no interaction between particles, no modification due to biological activity and chemical processes). Although disappointing, this will also be a very strong result as such.

[1] LeCun et al. Deep learning, Nature, 2015.
[2] Fablet et al. Learning Variational Data Assimilation Models and Solvers, arXiv 2020.
[3] Siegel et al., Prediction of the Export and Fate of Global Ocean Net Primary Production: The EXPORTS Science Plan. Front. Mar. Sci. 2016