contact: raphaelle.sauzede@imev-mer.fr

machine learning et données BGC-ARgo: produits opérationnels et

R. Sauzède

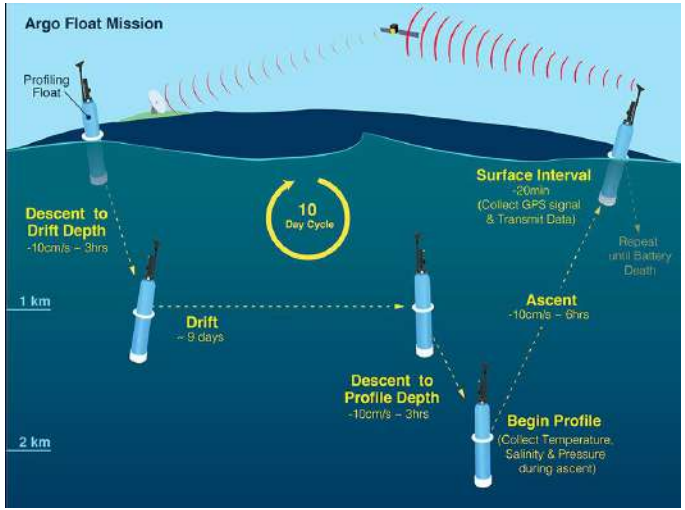
P.R. Renosh, L. Terrats, J. Uitz and H. Claustre

The Argo program

Prior to the turn of the 21st century, comprehensive *in-situ* ocean observations were difficult to obtain (mainly from ships and moored buoys):

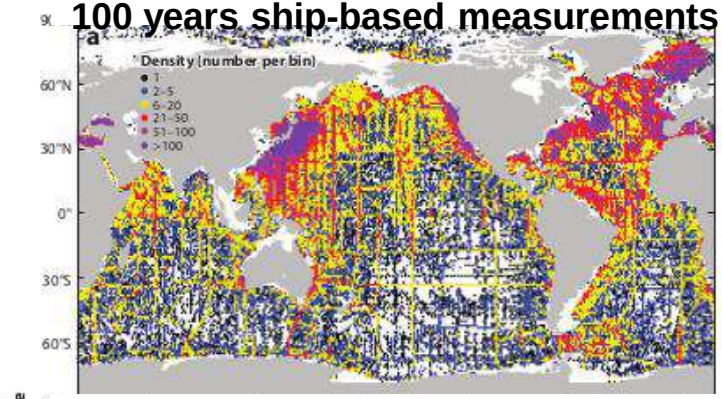
- bias to the Northern hemisphere
- large spatial gaps because of transect lines
- seasonal bias
- sparse data for high latitudes

→ critical for operational oceanography

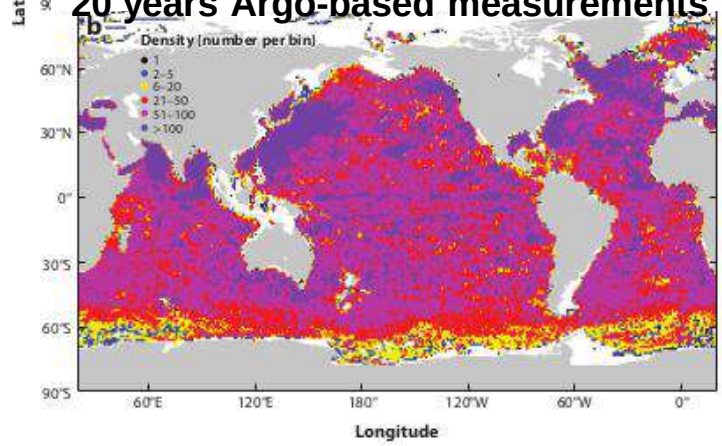


Wong et al., 2020

100 years ship-based measurements



20 years Argo-based measurements

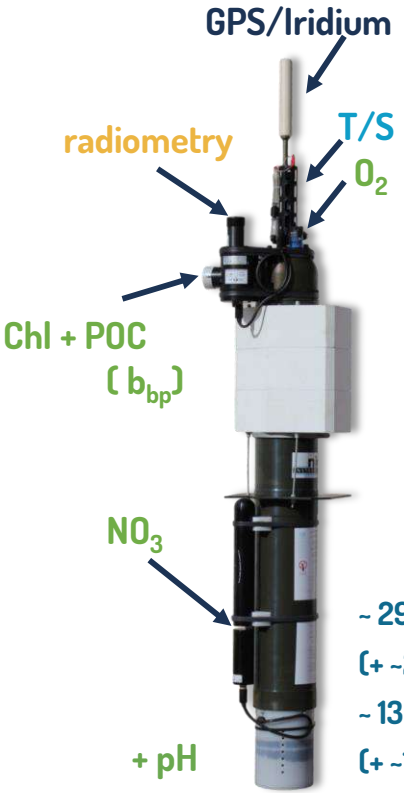


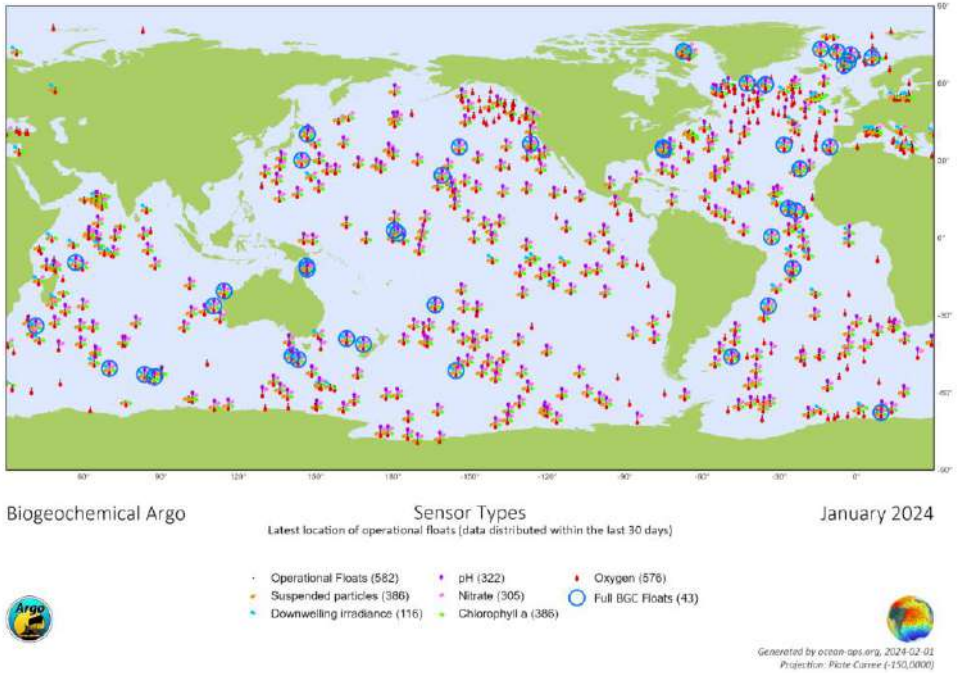
Johnson et al., 2021

Argo international scientific program begun in 1999

The BGC-ARgo program

BGC-Argo launched in 2016





(Claustre et al., 2020)

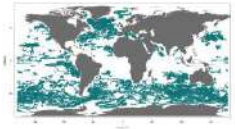
~ 295,000 profiles of O2
(+ ~25,000 / year)
~ 130,000 profiles of Chla + POC
(+ ~12,000 / year)

→ Take advantage of the power of ML and the growing amount of BGC-Argo profiling floats data with the main objective to finally provide useful and easy usable observation-based products

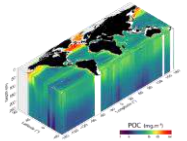
Observation-based products from BGC-Argo

LOV/IMEV develops 2 different type of observation-based products:

- BGC-Argo 'augmented' variables non or poorly measured from profiling floats



- BGC-Argo upscaling → 3D gap-filled gridded products



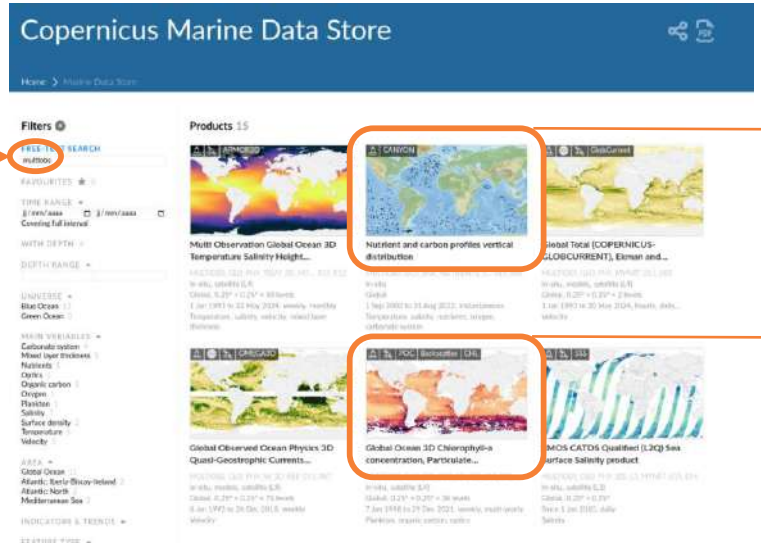
Some of these products are **operationally** delivered to the end-users through **Copernicus Marine Service**, some others are still in development

Copernicus Marine Service - MULTIOBS TAC

The Ocean Multi Observations TAC of Copernicus Marine Service (**MULTIOBS**)

- provides **global qualified ocean** multi-observations products
- products based on **observations** (satellite and in situ) and **data fusion techniques**
- Near Real Time, Multi Year products and Ocean Monitoring Indicators

https://marine.copernicus.eu/fr



In the search bar: MULTIOBS

BGC-Argo 'augmented' variables

(CANYON product)

BGC-Argo 3D gap-filled gridded products

(SOCA product)

Products covering:

→ Blue ocean

→ Green ocean

SOCA methods to estimate bio-optical properties

SOCA for Satellite Ocean-Color merged with Argo data to vertical distribution of bio-optical properties

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

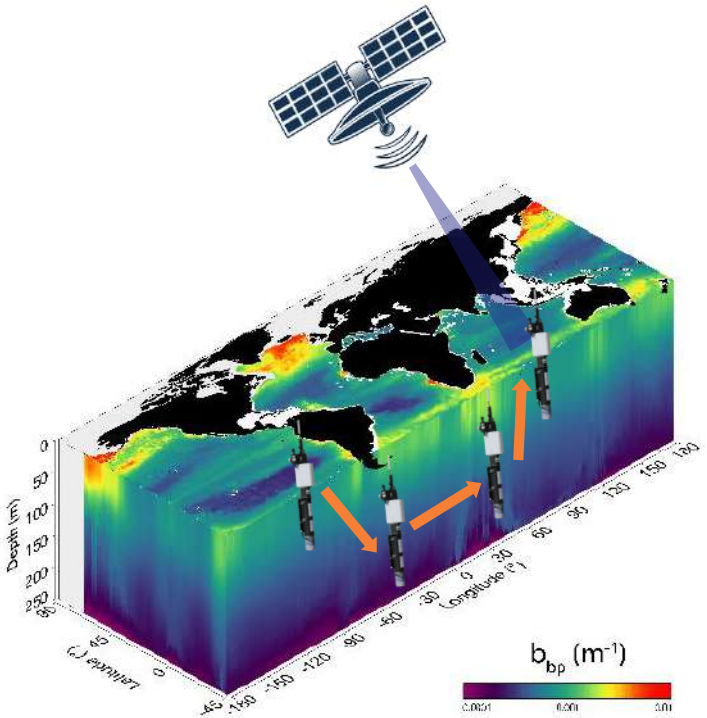
RESEARCH ARTICLE
10.1002/2015JC011408

A neural network-based method for merging ocean color and Argo data to extend surface bio-optical properties to depth: Retrieval of the particulate backscattering coefficient

Key Points:
- A neural network is developed to infer the vertical distribution of the backscattering coefficient

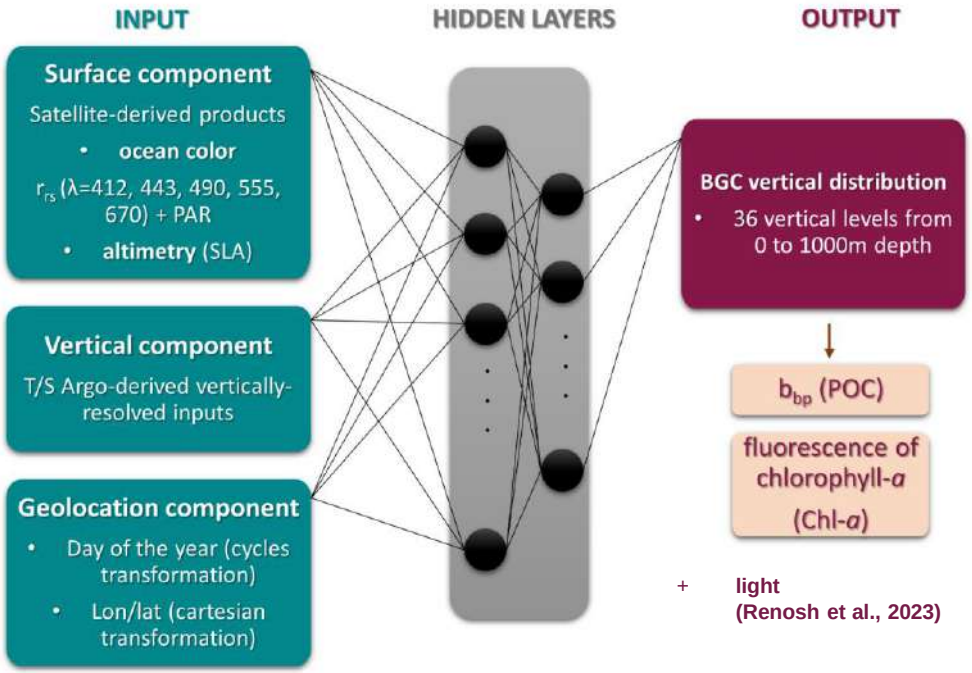
R. Sauzède[1], H. Claustre[1], J. Uitz[1], C. Jamet[2], G. Dall'Olmo[3,4], F. D'Ortenzio[1], B. Gentili[1], A. Poteau[1], and C. Schmechtig[1]

Sauzède et al., 2016



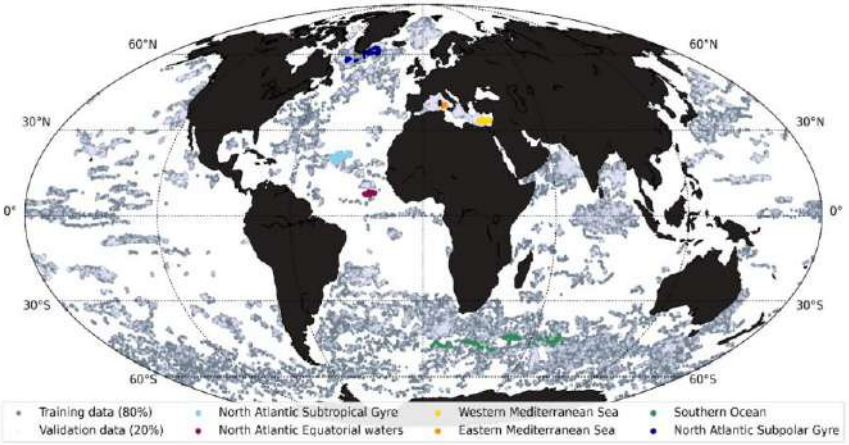
- Machine learning: neural network (Multi-Layer Perceptron)
- BGC-Argo data used as reference values to train and validate algorithm (~4,000 in 2016 and 55,000 profiles now)
- Concurrent profiles of T/S and bio-optical properties (b_{bp}+Chla) collected by BGC-Argo floats concomitant with satellite products
- Neural network (Multi-Layer Perceptron)
- Consequent updates from SOCA method published in 2016 to improve the performance (Sauzède et al., in prep.)

SOCA methods to estimate bio-optical properties



Sauzède et al., in prep.

~55,000 satellite/BGC-Argo matchups
→ **Database representative of the global ocean**

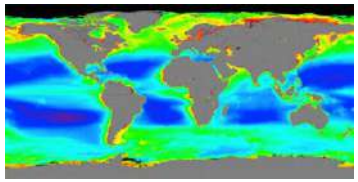


Global accuracies for SOCA retrieval:
b_{bp}: ~10%
Chla: ~30%

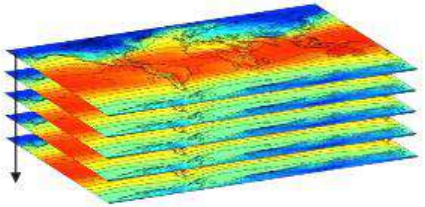
perational Copernicus marine service 4D gridded products

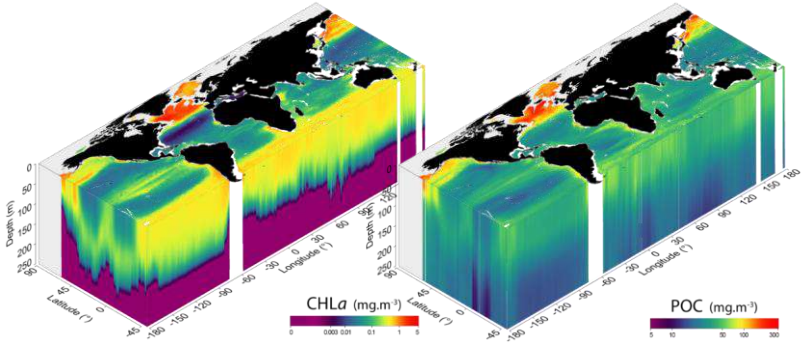
satellite fields

ARMOR3D fields (T/S)



+





Global 4D gridded POC (+b_{bp}) and Chl from Copernicus Marine Service :

- MULTIOBS_GLO_BIO_BGC_3D_REP_015_010 product from MULTIOBS TAC
- **Horizontal resolution:** ¼°
- **Vertical resolution:** 36 vertical depth levels from surface to 100 depth
- **Temporal resolution:** - Weekly fields from 1998 to 2022
  - Monthly climatological fields

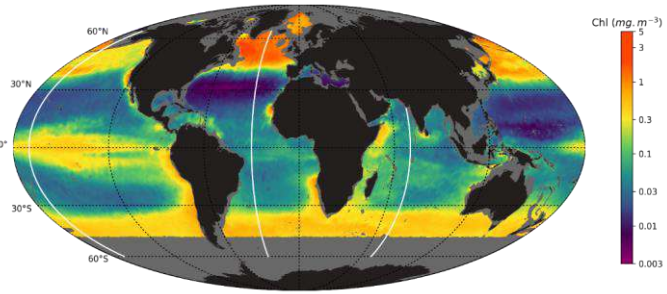
SCAN ME

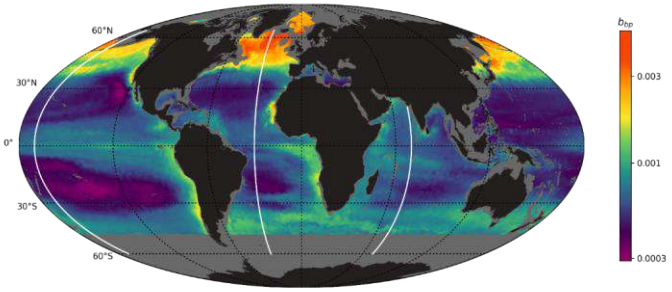
➔ Major update of the product in next November 2024 (Sauzède et al. in prep.)

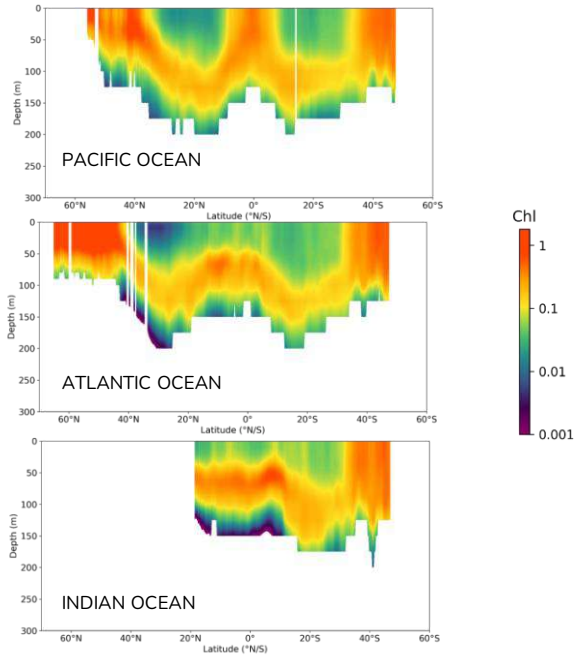
New insights from Observation-based products

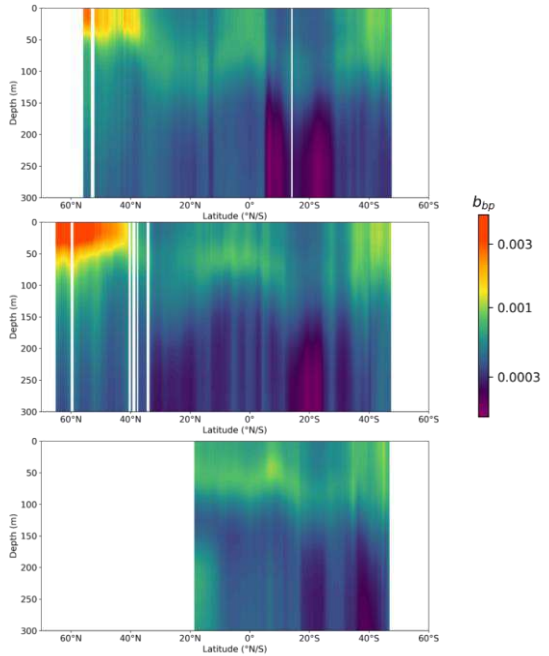
June climatology of the SOCA product:

Longitudinal transect at 170°W (Pacific) 30°W (Atlantic) and 70°E (Indian)





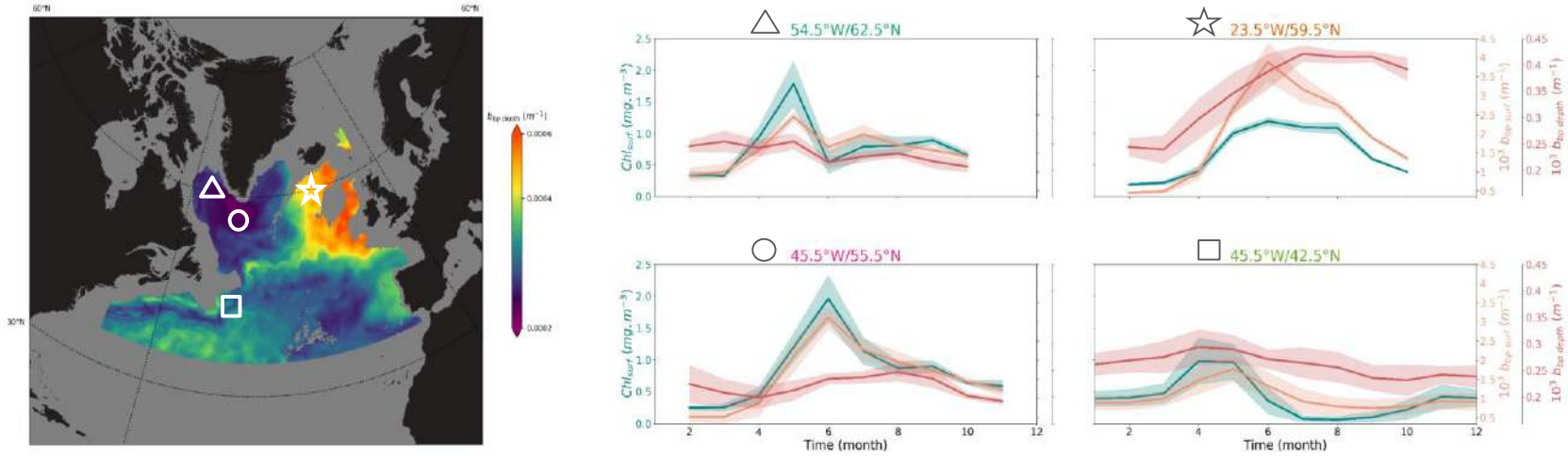




Sauzède et al., in prep.

New insights from Observation-based products

Regional focus: carbon export in the North Atlantic

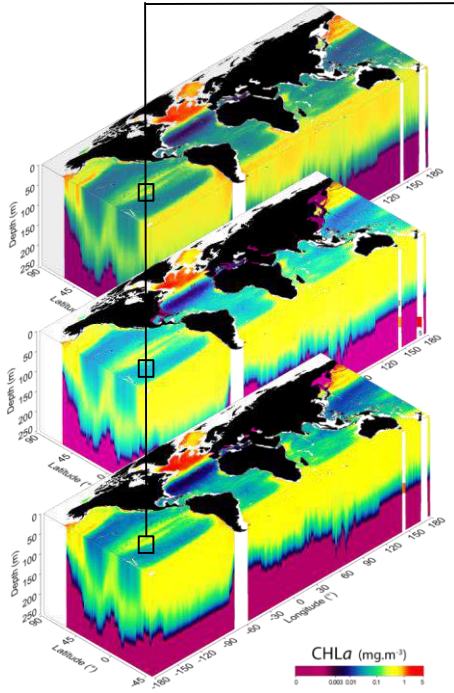


Sauzède et al., in prep.

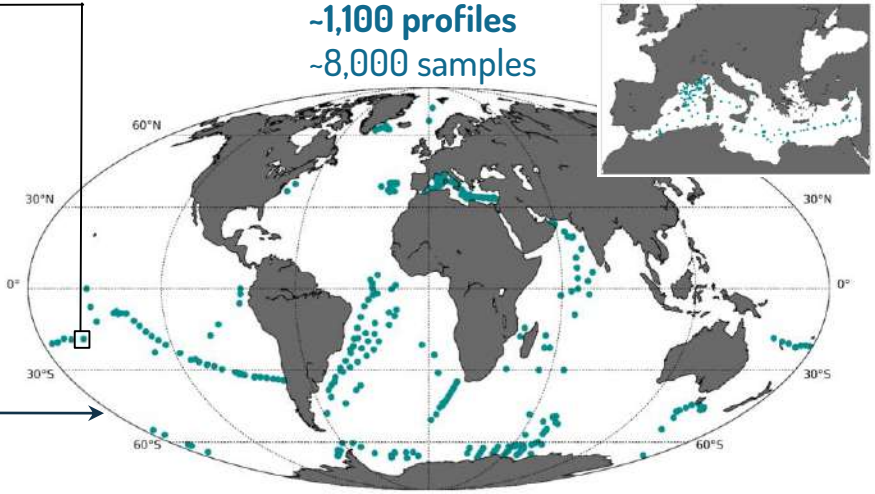
- **Gradient from North to South** reveals varying bloom onset across regions
- **Irminger Sea** exhibits a more prolonged bloom characterized by high phytoplankton biomass and the high b_{bp} signature in the surface layer suggests significant coccolithophore concentration → sustained and notably high b_{bp} signature in deeper layers
- **SOCA** enables the visualization of **the scale** of higher export (deep b_{bp} signature) at the regional level

SOCA-based workflow to evaluate BGC-ARgo dataset

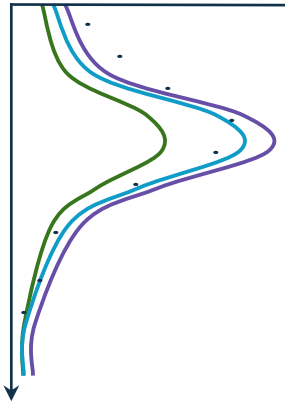
A **machine learning-based workflow** to evaluate the **improvement** of CHLA quality using new RT slopes.



MODEL A
MODEL B
...
MODEL X



~1,100 profiles
~8,000 samples

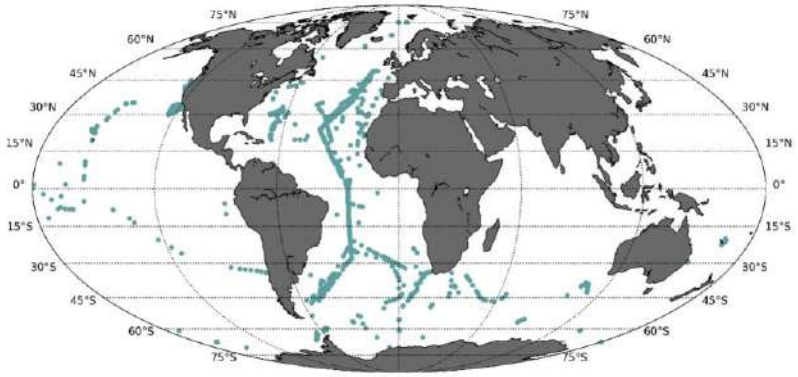


Global (or regional) statistics:
R^2, bias, MAPD, RMSD...
→ Best statistics = Best model
= CHLA_ADJUSTED more representative of HPLC
= Best adjustment

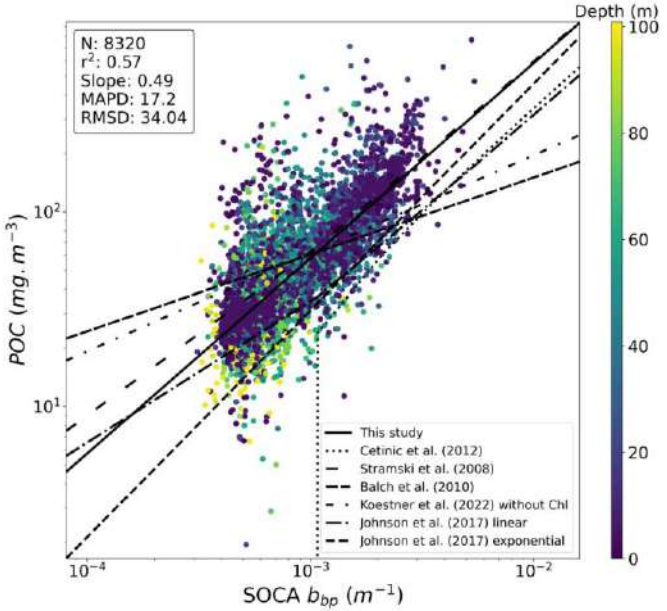
SOCA TO REFINE BBP/POC RELATIONSHIPS

→ By confronting SOCA-derived $b_{bp}(700)$ and POC in situ measurements matched with satellite data, a new **$b_{bp}(700)$/POC relationship** has been developed:

$$POC = 38687 * b_{bp}^{0.95}$$



Use of the ~**8,000** POC stations (Evers-King et al., 2017)



CMEMS QUID documentation

Applications and perspectives

In addition to the new scientific insights that offer these 4D-BGC observation-based products, they represent a most valuable source of data useful:

- For data assimilation, initialization/validation of **biogeochemical models**
- For the **quality control** of BGC-Argo float observations

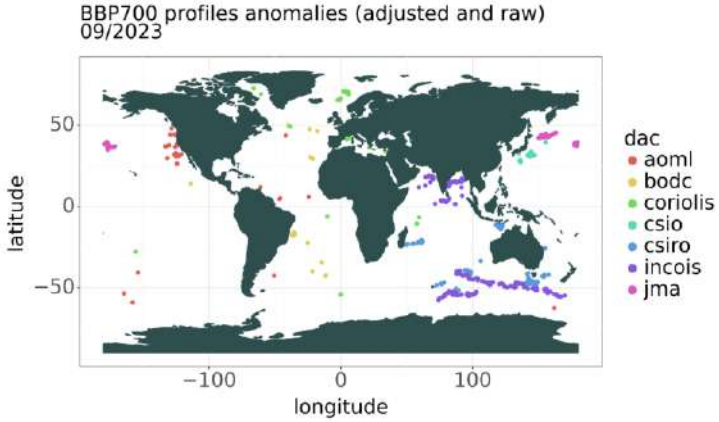
→ e.g. audit for b_{bp} released since June 2021 to help flagging anomalous profiles in the BGC-Argo database by comparison between BGC-Argo and SOCA-reference data

- Helping in the BGC-Argo **data management**

→ e.g. NPQ correction from SOCA-light estimations for floats non equipped with a radiometer or development of a new map of slopes of fluo/Chla at a global scale

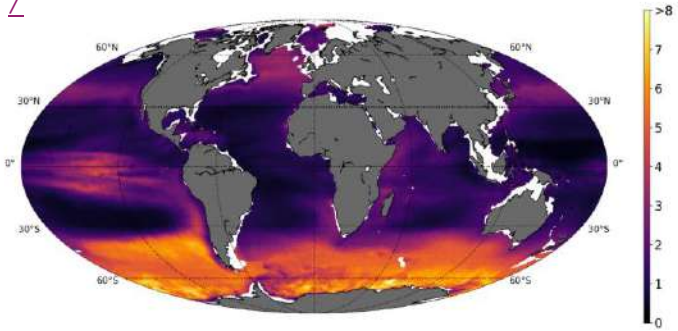
Ongoing development of SOCA for other properties measured or estimated from BGC-Argo

→ SOCA-PFT and SOCA-Carbon



→ ~1,000 profiles (1%) flagged as anomalous

ftp://ftp.mbari.org/pub/BGC_argo_audits/BBP700/



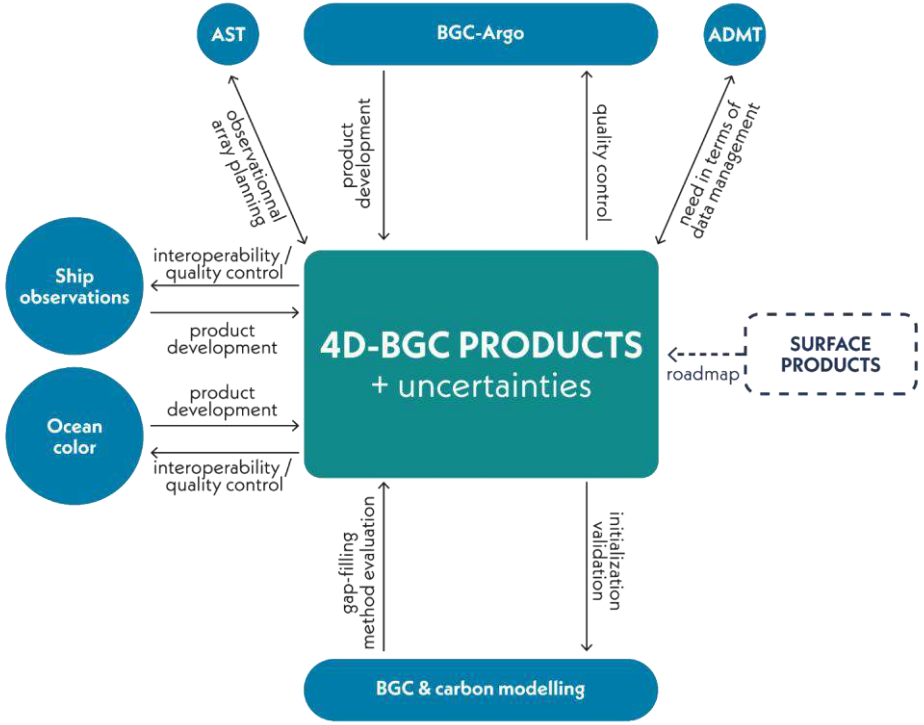
Fluo/Chla slope for the Chla RT adjustment in BGC-Argo

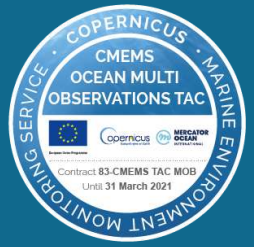
SCOR WG #168 4D-BGC: Coordinating the Development of Gridded Four-Dimensional Data Products from Biogeochemical Argo Observations

1. **Establish connections** among 4D-BGC product developers, observational communities and data synthesis efforts, and end-user communities
2. **Compile an inventory** of 4D-BGC products, and suggests relevant applications
3. **Synthesize available estimates**, variabilities, and trends of key biogeochemical processes that can be refined by 4D-BGC products, and identify actions to refine quantifications
4. **Develop recommendations** to create, distribute, and dynamically update 4D-BGC products, as well as strategies to estimate uncertainties
5. **Build capacity** within the oceanographic community, especially among early career researchers and within underrepresented groups

SCOR WG #168 4D-BGC: Coordinating the Development of Gridded Four-Dimensional Data Products from Biogeochemical Argo Observations

1. **Establish connections** among 4D-BGC product developers, observational communities and data synthesis efforts, and end-user communities
2. **Compile an inventory** of 4D-BGC products, and suggests relevant applications
3. **Synthesize available estimates**, variabilities, and trends of key biogeochemical processes that can be refined by 4D-BGC products, and identify actions to refine quantifications
4. **Develop recommendations** to create, distribute, and dynamically update 4D-BGC products, as well as strategies to estimate uncertainties
5. **Build capacity** within the oceanographic community, especially among early career researchers and within underrepresented groups





Thank you
Questions?

contact: raphaelle.sauzede@imev-mer.fr

LOV LABORATOIRE D'OCÉANOGRAPHIE DE VILLEFRANCHE
IMEV INSTITUT DE LA MER DE VILLEFRANCHE
cnrs
SORBONNE UNIVERSITÉ
REFINE
biogeochemical Argo