Reference database discussion

Virginie Racapé, Catherine Schmechtig

Initial discussion (May 2020)

BGC Argo needs a formalized reference database for standardized handling/storage of shipboard data used in calibration/validation

Questions initially posed by Lynne Talley (Scripps):

- Is there a part of the ADMT that deals with the shipboard data sets that are ingested for calibration?
- Is there a group in place that takes care of detailed oversight (compilation, formatting, QC, etc) for expanding (ie BGC) Argo?

Yes, this exists for Core-Argo, Deep-Argo, but not for BGC!

Reference datasets

GLODAP a logical partnership (notes from Henry Bittig):

- Already an established pathway from CCHDO (and other centers) for BGC data
- Capacity for secondary quality control, standardized formatting
- Provides convenient, easy-to-use global products
- Actively improving their data handling for better automation, timely updates/releases of merged datasets

Additional details re GLODAP submissions (notes from Bob Key):

- Data centers (ie CCHDO) primary sources of data but many datasets also come direct from PI
- WOCE,CLIVAR,GO-SHIPS cruises considered the most important included in GLODAP, but are only a small percentage of the total cruises.
- Many cruises that go into GLODAP are very small (3-20 stations).
- GLODAP focuses on BGC + Tracer parameters. We let WOA deals with the T/S/O2-only data sets
- GLODAP will gather CTD data, but we have no CTD experts and do no formal processing

⇒ GLODAP sustainability ?
⇒ Develop common tools to work on these datasets

⇒ Complete GLODAP datasets with marginal seas data (Black sea, Mediterranean Sea, Baltic Sea)
⇒ How to deal with cruises that do not have a direct pathway to GLODAP ?
⇒ Biooptics (HPLC, POC, Radiometry)

Global Ocean Data Analysis Project (GLODAP)

- Yellow - data ingestion + primary QC
- Green - Secondary QC
- Blue - make and publish product
- Gray - System maintenance and coordination

Assumes ~100 new cruises for each annual update
Assumes no automated data upload system
Accounts for Bob Key's retirement
Minimal to no communication with data PIs

Previous funding has come from several EU projects
Current funding comes from H2020 EuroSea
Covers approximately 0.25 FTE (or 3PM)

Out of the approximately 15PM used for each annual update, one full person year is voluntary and/or in-kind effort
Future GLODAP effort will rely on more user support

What	PMs	Cost (k€/k$)
Ingestion and 1st QC	6	60
2nd QC	3	30
Make product and check	1	10
Update paper	2	20
Adjustment table maintenance	1	10
Communication/coordination	1	10
Meetings to agree on adjustments[a]	1	10
TOTAL	**15**	**150**

Additional reference datasets

Data acquired at deployment OR other cruises that do not have a direct pathway to GLODAP

⇒ Develop a common repository (ERDDAP server ? Where ?)

⇒ Development of an Best Practices/S.O.P document for shipboard data collection protocols

⇒ Deal with embargo and citation

⇒ Format the file (netcdf files ? controlled vocabulary ? Controlled Units ?)

⇒ Stringent QC in line with needs of BGC Argo

⇒ Encourage their diffusion into larger datasets (QC, Format) => TO GLODAP ?

Additional reference datasets

Synthetic datasets

✓ Built on other dataset (Ocean Color, profiles)
✓ Using IA
✓ Gridded or not
✓ Climatological

⇒ Encourage the development of such datasets
⇒ Encourage their use

SOCA

AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE
10.1002/2015JC011408

Key Points:
- A neural network is developed to infer the vertical distribution of the backscattering coefficient
- The neural network requires as input Argo T/S profiles and ocean color remote sensing products
- Validation of the method is highly satisfactory which supports its application to the global ocean

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

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]

[1]Sorbonne Universités, UPMC Univ Paris 06, CNRS, Observatoire Océanologique de Villefranche, Laboratoire d'Océanographie de Villefranche, Villefranche-Sur-Mer, France, [2]Laboratoire d'Océanologie et de Géosciences, UMR8187, ULCO/CNRS, Wimereux, France, [3]Plymouth Marine Laboratory, Plymouth, UK, [4]National Centre for Earth Observation, Plymouth, UK

CANYON-B

frontiers in Marine Science | Ocean Observation

SECTION ABOUT ARTICLES RESEARCH TOPICS FOR AUTHORS EDITORIAL BOARD ARTICLE ALERTS

‹ Articles

ORIGINAL RESEARCH ARTICLE
Front. Mar. Sci., 20 September 2018 | https://doi.org/10.3389/fmars.2018.00328

An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO_2 Variables and Nutrient Concentrations From T, S, and O_2 Data Using Bayesian Neural Networks

Henry C. Bittig[1*], Tobias Steinhoff[2], Hervé Claustre[1], Björn Fiedler[2], Nancy L. Williams[3,4], Raphaëlle Sauzède[5], Arne Körtzinger[2,6] and Jean-Pierre Gattuso[1,7]

[1]Laboratoire d'Océanographie de Villefranche, CNRS, UMR 7093, Sorbonne Universités, Villefranche-sur-Mer, France

Additional reference datasets

BGC-Argo datasets

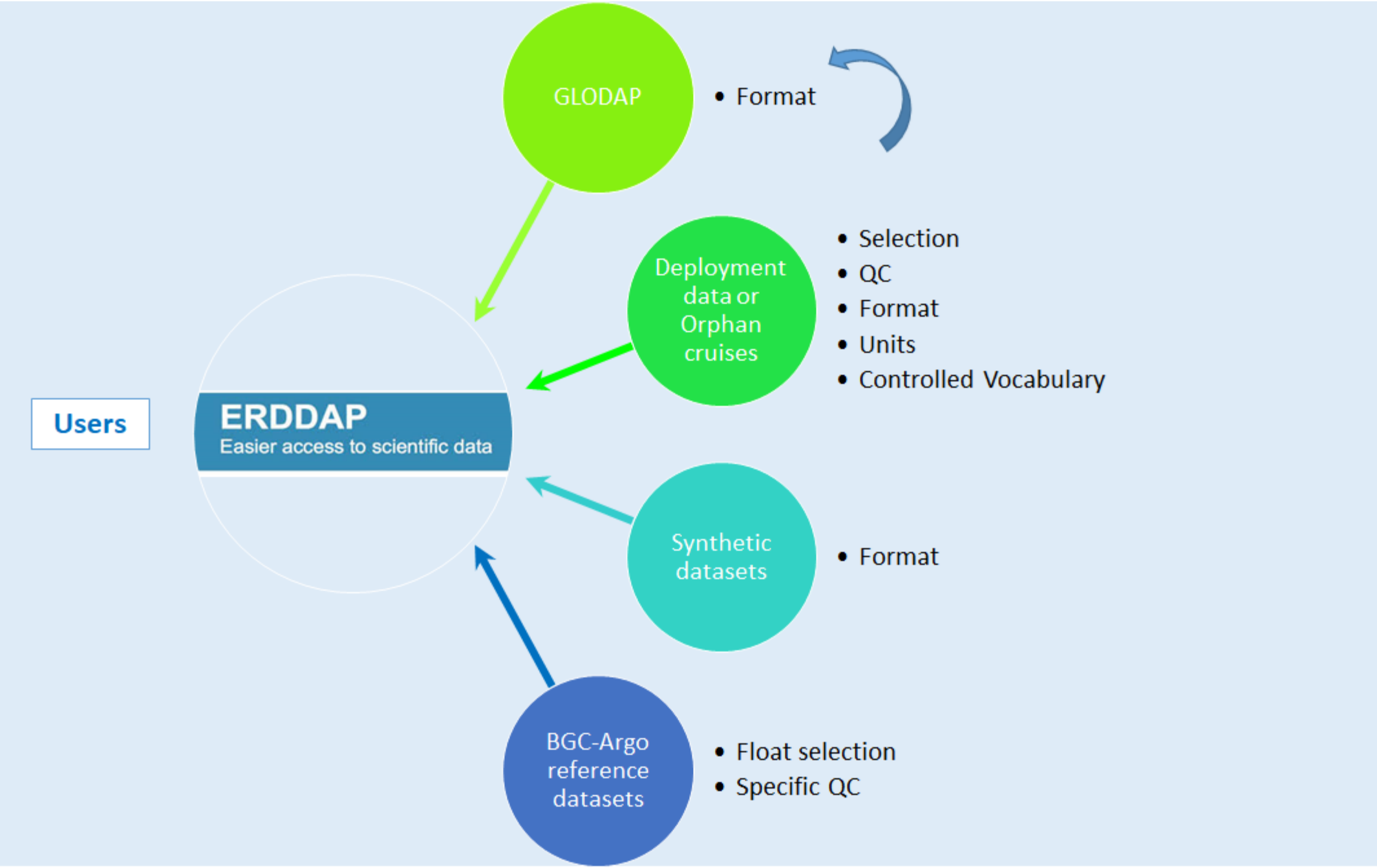
=> We should develop approved QC methods to select floats /profiles that could be considered as our BGC-Argo reference dataset

Suggested way Forward

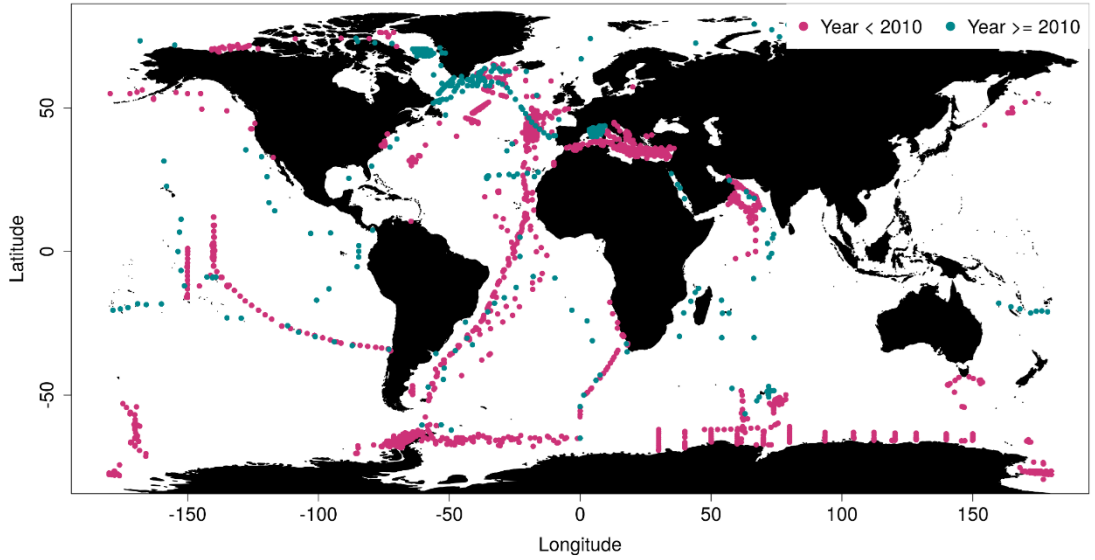
⇒ Set up a BGC-Working group on this topic (format, QC, Units, vocabulary, Citation, Embargo)

⇒ Start to fill ERDDAP server with reference datasets

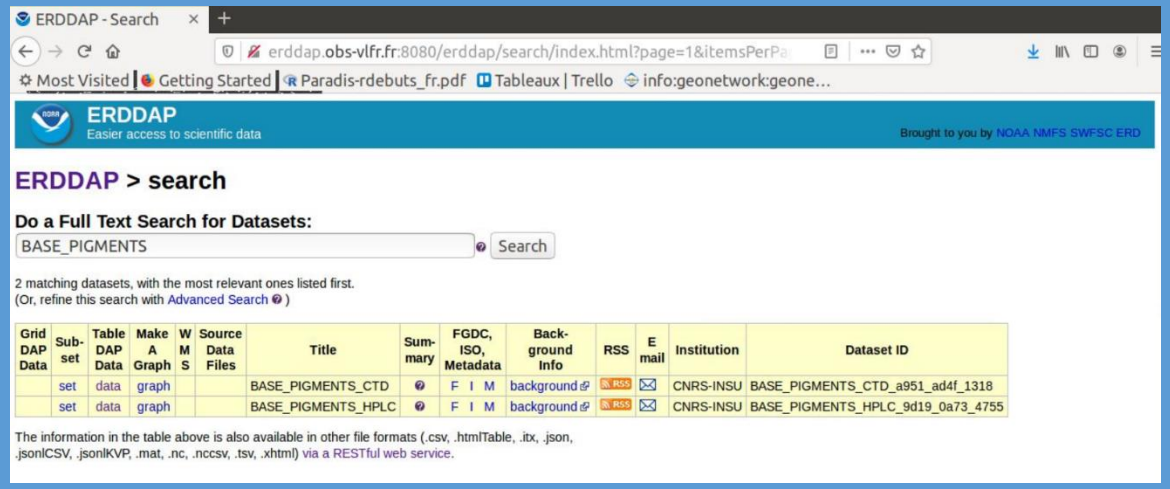
⇒ Test it and ask for feedback for next BGC-Admt


GLODAP
• Format
Deployment data or Orphan cruises
• Selection
• QC
• Format
• Units
• Controlled Vocabulary
Users
ERDDAP
Easier access to scientific data
Synthetic datasets
• Format
BGC-Argo reference datasets
• Float selection
• Specific QC

CHLA



HPLC data MAREDAT data + LOV Data



Pigments and Fluorimeters

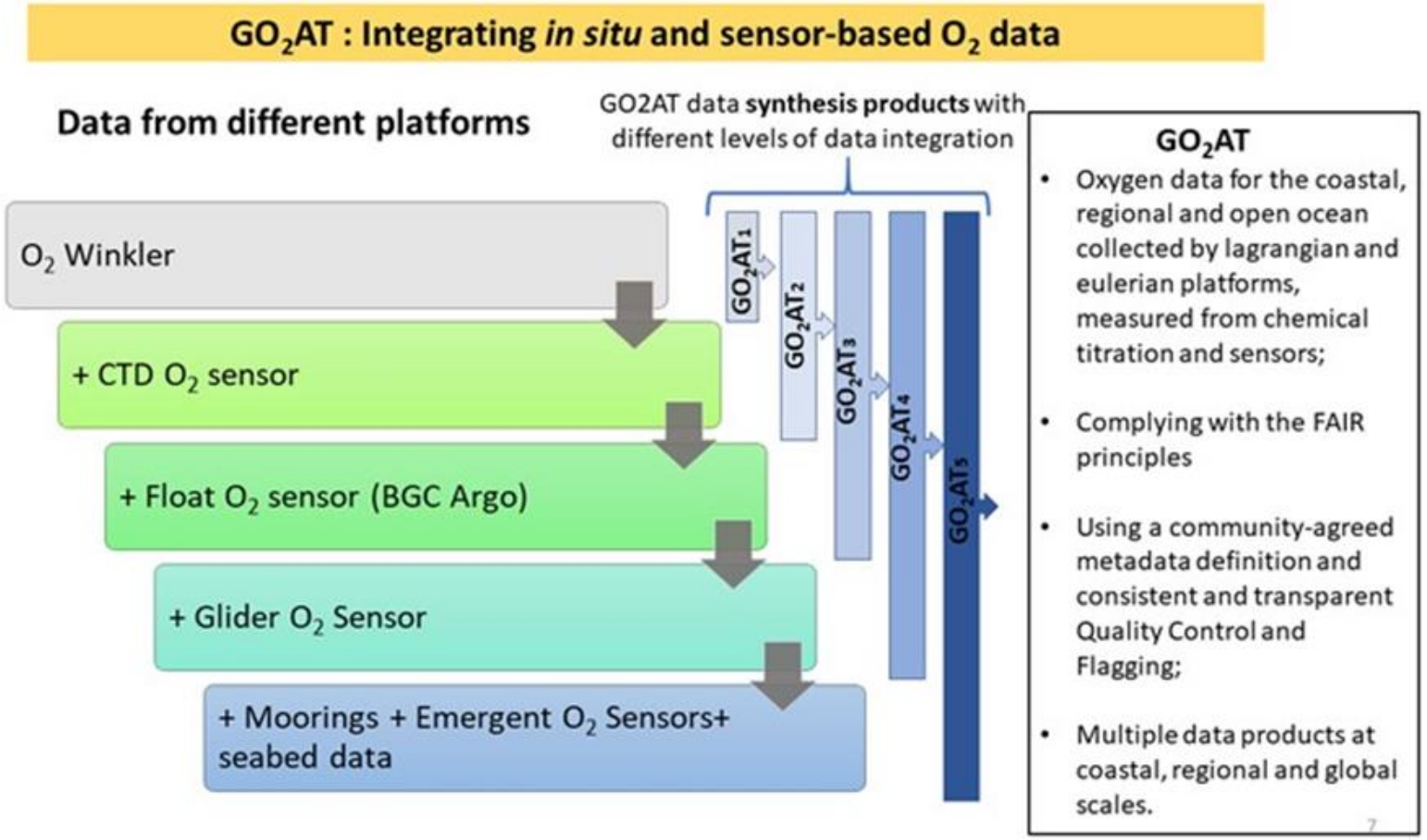
- All pigments (more than 25 types) analysed by the french service national of analysis of pigments (SAPIGH)
- All the relative CTD, Fluorescence, Characterisitics of fluorometers…

=> Can be easily transfer to another ERDDAP server

DOXY

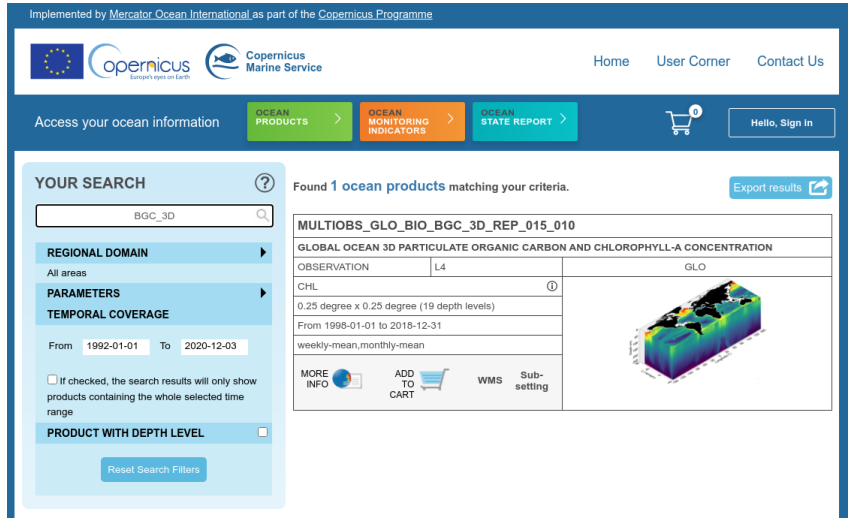
Effort lead by Veronique Garcon and Marilaure Gregoire

Scheme of integration of oxygen data sets from different platforms and sensors in different levels oxygen data synthesis products



GO$_2$NE Community White paper, 2020)

BBP

BBP700 climatological product =>Raphaelle's presentation

- Based on SOCA-BBP 2020 improved method used to deliver the POC product for the european service **CMEMS** (Copernicus Marine Environment Monitorng Service)
- Presentation of CMEMS products based on BGC-Argo data for end-users during the 19th ADMT 2018, La Jolla, CA: https://github.com/Sauzede/Presentations
- Products available at: → https://marine.copernicus.eu/ (in "access to products": search BGC_3D)
- **Resolution** of CMEMS 3D BGC product:
  - Global ocean
  - Black sea and coastal areas **not represented**
  - 0.25° x 0.25° horizontal resolution
  - 19 depth levels for the vertical resolution
  - Weekly products from 1998 to 2018
  - Monthly climatological products