

Marine Biogeochemical data quality control in EOSC Fair-Ease cloud

Catherine Schmechtig¹ (catherine.schmechtig@imev-mer.fr), Virginie Racapé² (virginie.racape@pokapok.org), Erwan Bodéré³ (Erwan.Bodere@ifremer.fr), Thierry Carval³ (Thierry.Carval@ifremer.fr), Fabio Conversano⁴ (fabio.conversano@szn.it), Jérôme Detoc³ (Jerome.Detoc@ifremer.fr), Delphine Dobler⁵ (delphine.dobler@euro-argo.eu), Alessandra Giorgetti⁶ (agiorgetti@ogs.it), Claire Gourcuff⁵ (claire.gourcuff@euro-argo.eu), Marie Jossé⁷ (Marie.Josse@data-terra.org), Sebastian Mieruch⁸ (sebastian.mieruch@awi.de), Marco Miralto⁴ (marco.miralto@szn.it), Catalina Reyes⁶ (nreyessuarez@ogs.it), Raphaëlle Sauzède¹ (raphaelle.sauzede@imev-mer.fr), Reiner Schlitzer⁸ (reiner.schlitzer@awi.de), Alban Sizun² (Alban.Sizun@pokapok.org), Clément Weber² (Clement.Weber@pokapok.org)

¹Sorbonne Université (France)

²POKaPOK (France)

³Ifremer (France)

⁴SZN-Stazione Zoologica Anton Dohrn (Italy)

⁵Euro-Argo ERIC (Europe)

⁶Istituto Nazionale di Oceanografia e Geofisica Sperimentale - OGS (Italy)

⁷CNRS Data-Terra (France)

⁸Alfred Wegener Institute (Germany)

The observation of marine biogeochemical (BGC) properties is fundamental to address scientific processes regarding the health of marine ecosystems (e.g., ocean acidification, oxygen minimum zone, biological carbon pump, phytoplankton communities, etc.) and needs for ocean resource management. BGC sensors have been deployed through various autonomous platforms (floats, gliders, sea mammals, moorings, etc.) by observing networks under GOOS-OCG international coordination (Global Ocean Observing System - Observations Coordination Group), leading to a dramatic increase of BGC observations at global scale for last decades. There are currently more than 900,000 BGC profiles measured either by a BGC-Argo float, a glider or a sea-mammal throughout the global ocean, including European marginal seas (source: Copernicus Marine Service - <https://doi.org/10.48670/moi-00036>, Dec. 2nd, 2022): 18% are already in delayed mode status (usable by scientists in total confidence), 10% are automatically adjusted in real time whereas the remainder needs to be adjusted before it can be used. Thus, BGC data adjustment and validation represent a major challenge to significantly increase the volume of high quality BGC data available for the scientific community.

At present, the BGC-Argo science team is a major contributor on an international level to calibrate, validate and trigger alerts on in situ BGC data at global level. In recent years, BGC-Argo sensors have diversified (oxygen, nitrate, chlorophyll, suspended particles, pH) and methods of data quality assessment and control, validation and adjustment have become more complex. Most of the methods are available as open source tools, available on various public github repositories. These methods require an efficient access to external data: gridded climatologies, model outputs (meteorologic, oceanographic), discrete in-situ data, satellite data, etc. Moreover, applying these methods requires combining the data in space and in time using extraction and colocation. Thus, a massive, high-performance, distributed data infrastructure that would combine in situ, satellite and models data would definitely help the data scientist community. Today, softwares development and metadata standards are specific to the Argo format. However, methods are essentially sensor-dependent and not platform-dependent, meaning that it is applicable to BGC sensors deployed on gliders, sea-mammals or other platforms.

As part of FAIR-EASE, a Horizon Europe project whose objective is to customize and operate distributed and integrated services for observation and modeling of the Earth system, a demonstrator is being developed for contributing to the overall improvement of BGC data quality, through software standardization and easy cloud development. The aim of the demonstrator presented here is to provide a single and efficient access to three services: qualification/calibration/validation of BGC data through a web portal. Tools deployed for the calibration focus on ocean BGC observations essentially measured by sensors deployed on BGC-Argo floats, gliders or sea mammals.

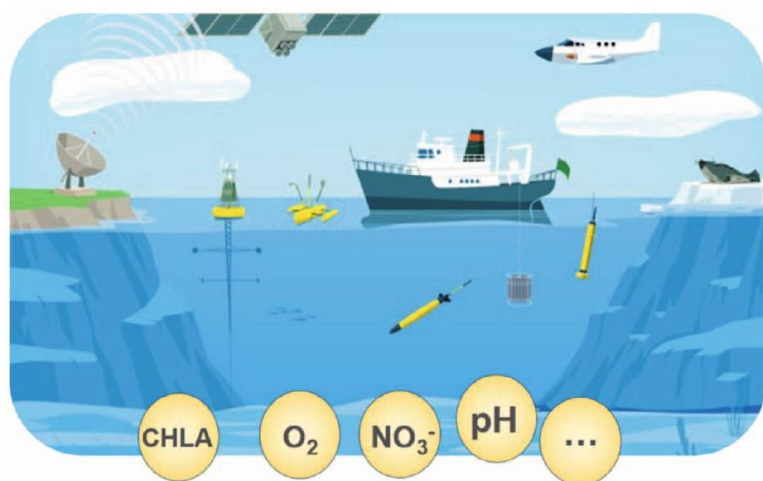


Figure 1 BGC in situ ocean observations.

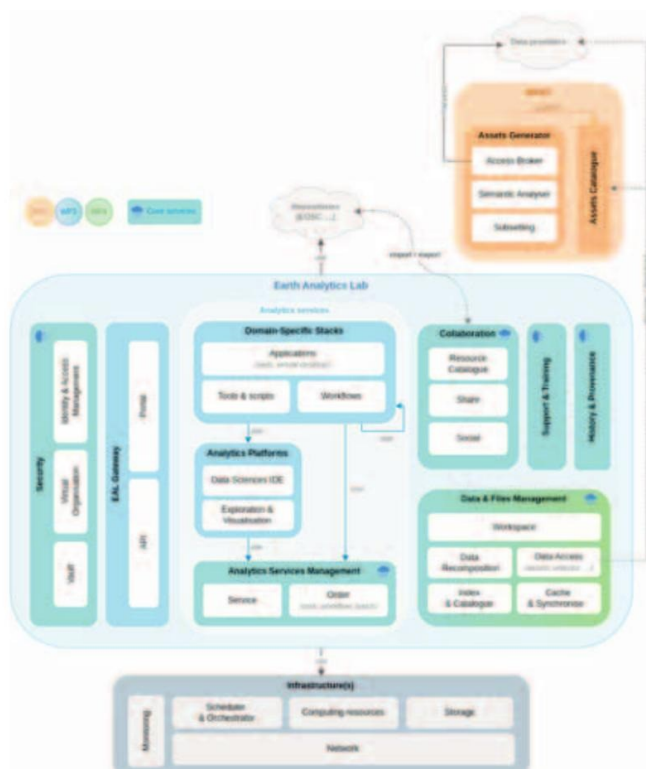


Figure 2 EOSC FAIR-EASE infra for BGC data scientists.

We shall present the scientific challenge to quality control, calibrate and validate in situ BGC Ocean data, the solutions and the technical infrastructure for European data scientists. The infrastructure is part of the European Open Science cloud, developed within the European project FAIR-EASE.