Framework Service Contract EEA/DIS/R0/20/001 Lot 1 for Services supporting the European Environment Agency's (EEA) cross-cutting coordination of the Copernicus programme's in situ data activities – Observational data

Inventory of existing European Biogeochemical observations

Lőrinc Mészáros
Ghada El Serafy
Jun She
Helene Frigstad
Andrew Tyler
Georg Umgiesser
Steve Groom

Issue:
Date: 28-10-2021

[DOCUMENT RELEASE]

	Name(s)	**Affiliation**
Coordinated by:	Lőrinc Mészáros, Ghada El Serafy	Deltares
Contributions:	Jun She Helene Frigstad Andrew Tyler Georg Umgiesser Steve Groom	DMI NIVA University of Stirling ISMAR CNR PML
Approval:		

Prepared for:	European Environment Agency (EEA)
Represented by: (Project Manager)	Henrik Steen Andersen
Contract No.	EEA/DIS/R0/20/001 Lot 1

[Change Record]

Version	**Date**	**Changes**
1.0	16-3-2021	First draft release
1.1	9-4-2021	Revised version based on comments from EuroGOOS Office
1.2	22-4-2021	Second revision based on comments from EuroGOOS Office
1.3	8-10-2021	Inclusion of Regional Conventions

Table of Contents

Executive Summary

This task report was prepared within the EEA framework contract 'Support to Copernicus In Situ Data Coordination' (COINS) and specifically targets Task 5.1, **collecting and reporting information on existing BGC observations, including data repository and availability**. The BioGeoChemical (BGC) in situ observation data inventory covers coastal areas, transitional waters and open oceans (at global and regional level). In the report the timeliness of data delivery is evaluated with specific attention given to near-real time delivery, as that is part of the requirements of Copernicus Monitoring Forecasting Centres. Nevertheless, the report does not specifically target data for the operational use (i.e in near-real time and for forecasting purposes), instead it includes any BGC in situ data, irrespective to the update frequency and timeliness.

Within this task BGC in situ data sources have been investigated and an inventory of ~50 data portals was prepared. **It should be noted that the prepared inventory is not complete and other data sources may have been missed.** The reviewed BGC in situ data sources cover data aggregators (parties who collect, harmonize and aggregate the observations), international bodies, European marine research infrastructures and research projects. Organizations included in the inventory are national institutes, international programmes, Regional Sea Conventions, pan-European projects and initiatives, and global monitoring efforts. These actors provide a wide range of BGC in situ observations in European marine waters obtained from fixed moorings, drifting buoys, profiling floats (BGC-Argo floats), bottom landers, gliders, autonomous vehicles and FerryBoxes.

In the report major BGC in situ observation networks and data repositories have been inventoried, complemented by a survey which was filled by representatives of EuroGOOS ROOSes, JERICO-RI and Jerico-S3 projects, ICOS Marine, HELCOM and the Black Sea Regional Sea Convention. They were contacted with an online survey to report on the existing BGC observations in their respective programmes. The online survey questions can be found in the Appendix. Further complementary feedback was collected via email exchanges and through one-to-one interviews. The present inventory accompanied with a data catalogue aims to help EEA to better formulate long term in situ data coordination strategies in European regional seas, and especially in the coastal areas.

The main findings of the report include the following:

- The availability of BGC in situ data is still significantly less than the availability of physics data. Moreover, the investigated BGC in situ datasets were unbalanced with much more chemistry data (e.g. chlorophyll or oxygen) and less biological data (e.g., phytoplankton biomass and diversity).
- While a wealth of data portals was discovered that contain BGC in situ data, these repositories often contain overlapping data sources. The quantity of available BGC in situ data is therefore not in positive correlation with the number of existing data portals.
- The discovery and quantification of how many coastal data is present in the databases was not easy as data is labelled by regions and thematic (not by its coastal or offshore nature). In addition, gaps were identified in the availability of coastal BGC in situ data in several regional seas. Examples are the southern Mediterranean, the Black Sea and the Arctic Sea.
- Regional Sea Conventions use heterogenous approaches for data collection management and dissemination. Their data management practices and infrastructure have varying level of maturity. While their efforts for sharing the data underlying their assessment reports are very useful, these databases are sparse in terms of spatial and temporal coverage or present aggregate values. In addition, public accessibility is not guaranteed in Regional Sea Conventions that involve non-EU Contracting Parties. Based on our evaluation, currently not

all BGC data available in the data bases of Regional Sea Conventions is also accessible through EU data aggregators. It is recommended to map missing datasets and enhance the data sharing between Regional Sea Conventions and data aggregators.

- Data access and availability were satisfactory in most cases as instant online data download was provided. Nevertheless, sometimes registration or (manually processed) request via email was needed, which hinders data sharing. Data formats were largely homogenous due to the continued data management efforts of EuroGOOS Task Teams, CMEMS INSTAC, EMODnet, and SeaDataNet, which maintain common standard formats and quality control for the BGC EOVs.
- The data update frequency used by data aggregators is highly dependent on the quality control procedures of the data providers. Data availability can be considered in two ways: (1) sensor data with near-real time availability of raw (only automatically quality controlled) data, and (2) sample data with a longer time delay (typically yearly) due to manual (laboratory) quality control.

Priority issues to be addressed:

- A significant increase in the number of BGC in situ measurements in European seas is required to satisfy the growing need for BGC data in support of modelling (validation and data assimilation);
- In several regional seas existing monitoring programmes need to be further developed to better assist model validation and data assimilation, especially covering transitional water systems, coastal waters, and non-EU countries;
- Expand BGC monitoring capacities from chemistry towards biological parameters through coordinated monitoring programmes that make use of the full range of available BGC monitoring technics, including moorings, profiling floats, gliders, drifting buoys, and ship-based systems;
- Expand the data management practices to all BGC in situ data sources;
- Guarantee free and unrestricted access to BGC in situ data, including Regional Sea Convention data;
- Guarantee near real time (less than 1 day) data delivery to BGC variables that are used for data assimilation in CMEMS MFCs (chlorophyll-a, dissolved oxygen, pCO2, particulate matter, and pH).
- Labelling the coastal datasets in the metadata catalogues would facilitate the monitoring of the progress in the expansion of coastal BGC in situ observation networks. This would, however, require a shared definition of coastal systems.
- Better use of Regional Sea Convention data could be achieved by further developing and harmonizing their data management practices and infrastructure. Further study should be conducted to identify data records of the Regional Sea Conventions that are currently missing from the European data aggregators.

BGC in situ data inventory per coordinating organization

Only around 10% of in situ measurements are biogeochemical (BGC) observations[1]. The basic BGC variables that are measured by in situ sensors are the nine BGC EOVs (see a full list of BGC EOVs here: https://www.goosocean.org/index.php?option=com_content&view=article&id=170&Itemid=114) dissolved oxygen, nutrients, inorganic carbon, transient tracers, particulate matter, Nitrous Oxide, Stable Carbon Isotopes, Dissolved Organic Carbon (DOC), and Ocean Colour. These basic BGC variables are the most feasible (sustained observation) and the ones with the highest scientific impact. Nevertheless, this inventory is not restricted to the BGC EOVs as proper characterisation of biogeochemical processes, and validation of numerical models simulating these require further variables such as phytoplankton biomass and diversity. This chapter is structured according to BGC data sources classified as:

(1) data aggregators who collect, harmonize and aggregate the observations,
(2) Regional Sea Conventions, which collect, maintain, and process BGC in situ data to support periodic assessment reports on the State of Environments in their respective sea basins,
(3) European marine research infrastructures who aim to integrate and harmonize the observing systems, and
(4) other BGC in situ data sources from independent EU or national research projects.

Apart from research infrastructures and projects that are specifically targeting the coastal zone, such as DANUBIUS-RI, JERICO-RI, COSYNA, or H2020 ODYSSEA, the report focuses on the regional level (European regional seas) and not on the European coastal zone. The scientific definition of coastal zone is still debated, and the data platforms do not label their data as coastal or open ocean, instead the classification is done based on the region and the thematic group of variables. Nevertheless, BGC in situ data on the regional level is crucial for the CMEMS Monitoring and Forecasting Centres (for model validation and data assimilation) who provide operational marine forecasts as well as for validation of satellite products. Model outputs of CMEMS MFCs serve as boundary and initial conditions to high resolution coastal models. Consequently, the BGC in situ data availability in the open ocean has high impact on the coastal modelling capabilities.

[1] Fact sheet on Copernicus in situ data requirements: https://insitu.copernicus.eu/FactSheets/CMEMS/

Data aggregators

Intermediate or end-users of BGC in situ data often obtain the required information through so-called data repositories or data aggregators, and not directly from the data providers. Data aggregators address the issue of fragmented data sources and standardization; hence, their data products are more accessible to the users. Data aggregators such, as CMEMS INSTAC, EMODnet (https://emodnet.eu/) and SeaDataNet (https://www.seadatanet.org/), are in contact with many data producers (e.g. national or regional data centres) for BGC observations. In the following sub-sections, we review the major European data aggregators and their BGC in situ portfolios.

CMEMS In Situ TAC

The Copernicus Marine Environment Monitoring Service is the marine component of the Copernicus Programme of the European Union. CMEMS collects BGC in situ observations through its In Situ Thematic Assembly Centre (INSTAC) and makes them accessible to CMEMS The CMEMS INSTAC collects near-real-time in situ data from data providers (national and international networks), carries out quality control in a homogeneous manner, and distributes them in near real time. CMEMS In Situ TAC provides data access to member state coastal in situ observations which include fixed platforms and moorings, HFRs, gliders and ship-based observational systems. The following types of in situ observing systems are aggregated by INSTAC: BGC-Argo floats, research vessels, surface moorings, FerryBoxes, and gliders. The information of all the observing platforms can be monitored in the INSTAC dashboard http://www.marineinsitu.eu/dashboard/ (see *Figure 1*).

The list[2] of BGC parameters include:

- Dissolved oxygen
- Oxygen saturation
- Dissolved inorganic carbon
- Dissolved organic carbon
- CO2 partial pressure
- CO2 fugacity
- Chlorophyll-a
- CDOM
- Turbidity
- Nitrate
- Nitrite
- Phosphate
- Silicate
- Ammonium
- Dissolved nitrogen
- Dissolved organic nitrogen
- Total alkalinity
- pH

[2] https://archimer.ifremer.fr/doc/00422/53381/72333.pdf

It should be noted that having only a limited number of BGC variables is a conscious choice of CMEMS in order to ensure the rigorous quality check.

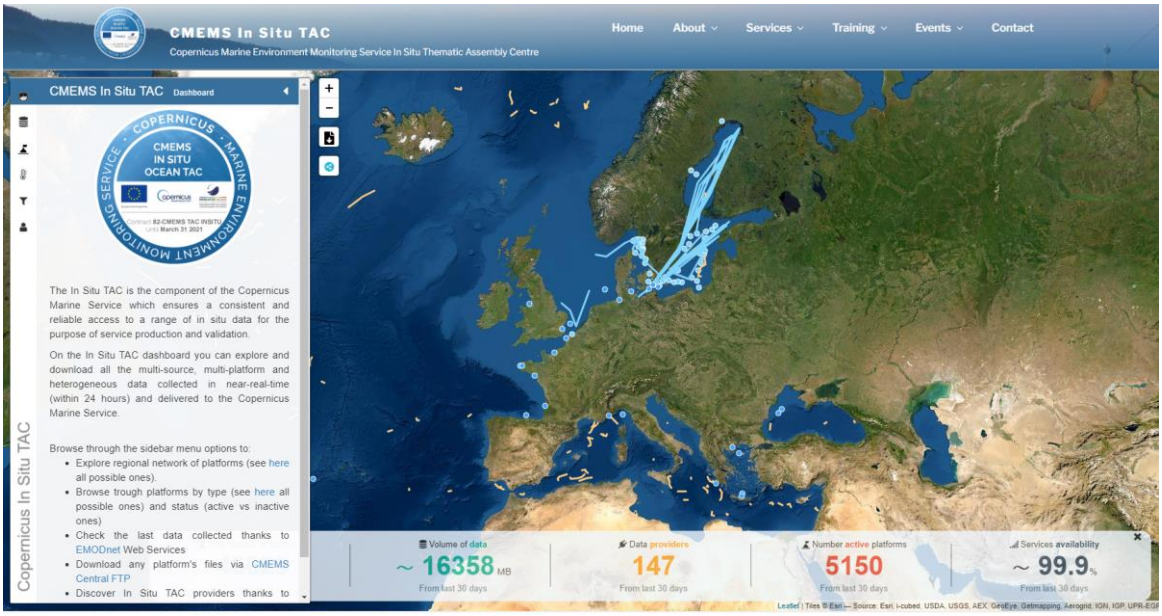
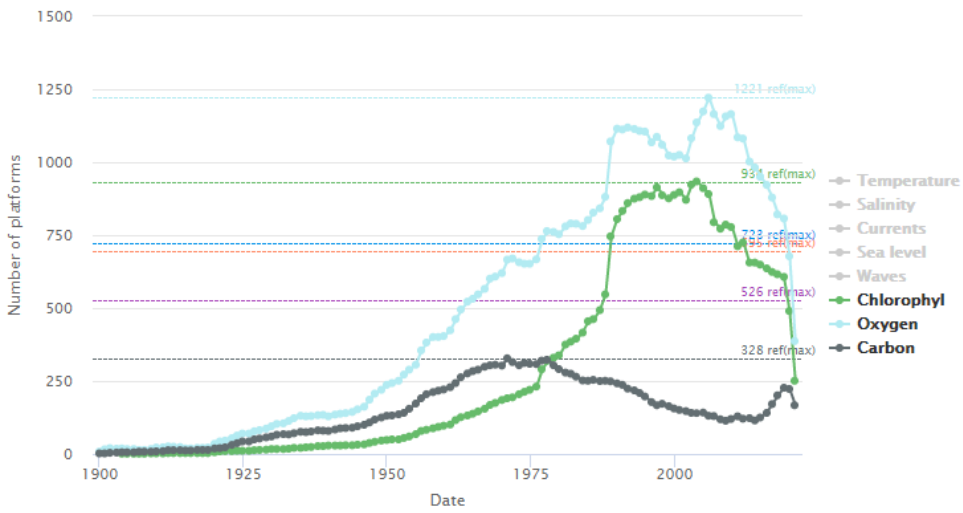
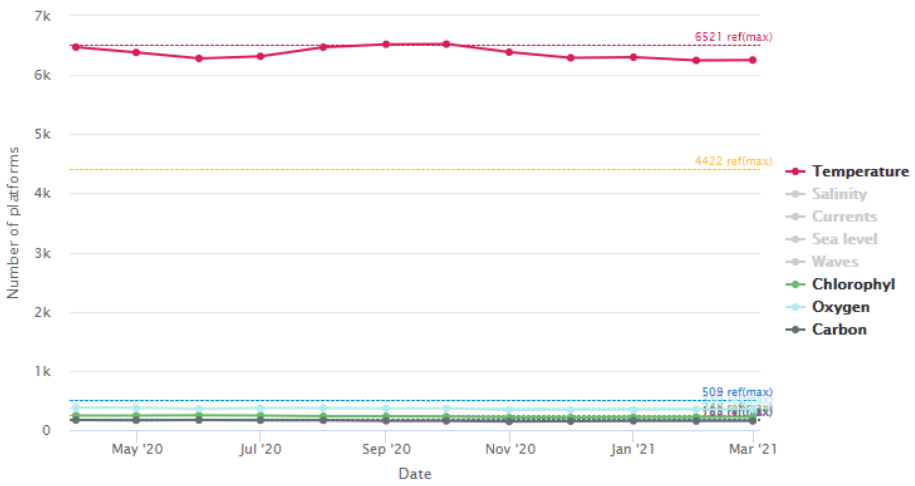


Figure 1. BGC in situ data in the CMEMS In Situ TAC data dashboard. Source: http://www.marineinsitu.eu/dashboard/

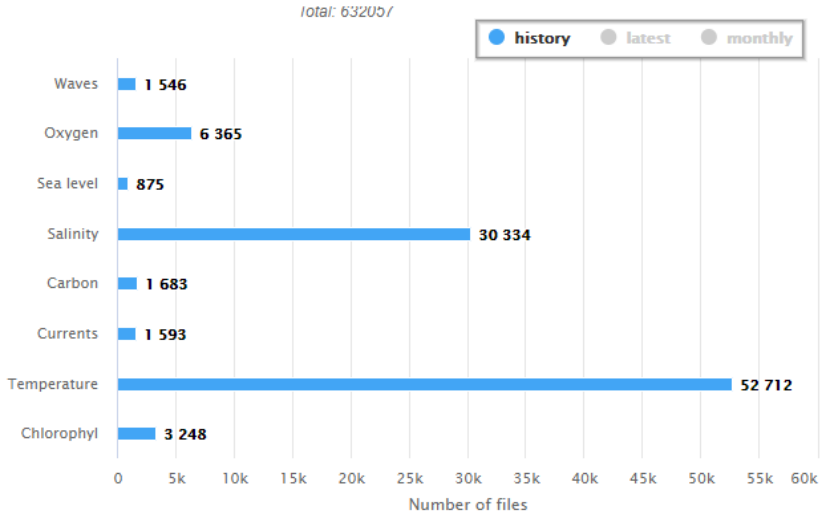
As depicted in Figure 2 (a), on a global level the number of BGC platforms has increased over the past decades, but comparing to physical platforms measuring temperature for instance, the proportion is still less than 10%, as shown in Figure 2 (b). Moreover, looking at the historical number of files per group, BGC variables account for less than 2% of the total. Regarding data accessibility, some datasets cannot be redistributed due to commercial interest, however, these can still be used by CMEMS for product improvement without disseminating the original data.



Number of platforms over time (a)



Number of platforms over the last year (b)



Number of files over the entire historical interval (c)

Figure 2. Comparison of physical and BGC observations in CMEMS INS TAC based on key performance indicators. Source: http://www.marineinsitu.eu/monitoring/

EMODNET Chemistry, Biology and Physics

The European Marine Observation and Data Network (EMODnet) facilitates that European marine data will become easily accessible, interoperable, and free of restrictions on use. EMODnet is covering European coastal waters, shelf seas and surrounding ocean basins to facilitate the monitoring of coastal and offshore activities. The EMODnet Physics, Chemistry and Biology portals provide easy access to marine chemical and biological data, and standardised harmonized validated data collections.

BGC data in EMODNET Physics and Chemistry portal:

- Chlorophyll-a
- Dissolved Inorganic Nitrogen
- Dissolved oxygen
- Phosphate

- Silicate

BGC data in EMODNET Biology portal:

- Bulk chemistry (e.g. pH, TCO2)
- Dissolved gases
- Nutrients (e.g. DIN, phosphate, silicate)
- Other inorganic chemistry
- Primary production
- Phytoplankton biomass
- Chlorophyll-a

BGC in situ data can be online instantly accessed through the EMODnet physics data portal (https://map.emodnet-physics.eu/), the EMODnet Chemistry Viewing and Downloading service https://ec.oceanbrowser.net/emodnet/ (see *Figure 3*) or the CDI Data Discovery and Access Service https://emodnet-chemistry.maris.nl/search . The EMODnet Biology data can be accessed through the https://www.emodnet-biology.eu/portal/. Both provide online instantly accessible data. Regarding timeliness, currently the time-lag of EMODnet BGC data can be more than 6 months. For example, in October 2021, the most recent phosphate data in the Baltic Sea are from December 2020. Exceptions are the ARGO floats, whose data is accessible with a smaller delay (around a month). This is different in the EMODnet Physics portal where data is accessible with a timeliness of less than two days.

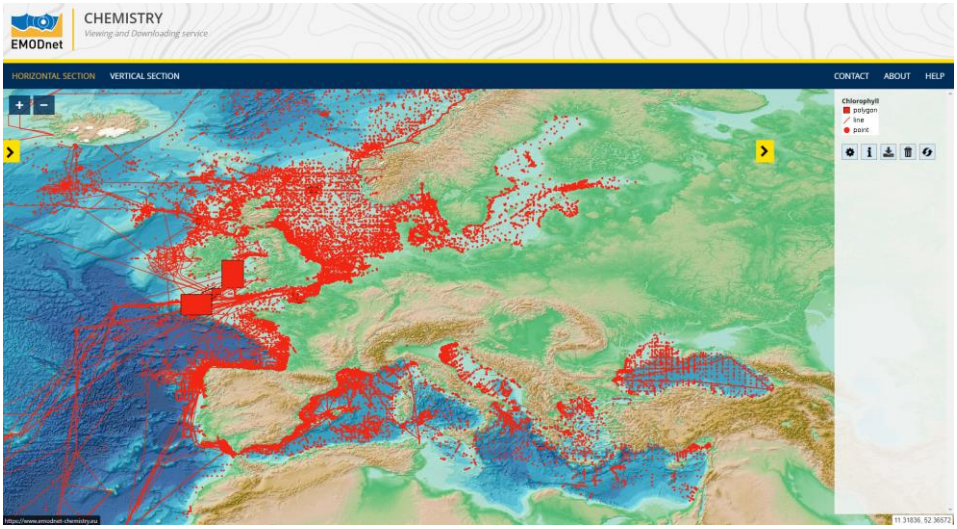


Figure 3. EMODNET Chemistry Viewing and downloading Service. Example of Chlorophyll-a measurements.

EMODnet increasingly covers the coast and visualizes coastal sections[3], showing concentration plots for a given time and space window and along the coast. The vertical section can be chosen by drawing a transect, defining bathymetry or distance to coast. Predefined coastal sections are also included. Current EMODnet Chemistry coastal data and products[4] also include eutrophication, contaminants, marine litter and ocean acidification products.

[3] Giorgetti, A., Ocean and Coastal Management (2018), https://doi.org/10.1016/j.ocecoaman.2018.03.016

[4] EMODnet Chemistry coastal data and products: https://www.emodnet-chemistry.eu/repository/EMODnet_Coastal_16062020.pptx

Further efforts of EMODnet to better cover the coastal regions include joint workshops with CMEMS, EU Commission Directorate General (DG MARE and DG GROW) representatives and the European Environment Agency.

SeaDataNet

SeaDataNet is a pan-European infrastructure for ocean and marine data management. It provides aggregated datasets based on data from external data sources. The SeaDataNet infrastructure links already more than 110 national oceanographic data centres and marine data centres from 35 countries riparian to all European seas. One of the main activities of SeaDataNet is to develop marine, coastal and riverine environmental data management systems, therefore also serving coastal zone management.

Datasets are available in **ODV4 ASCII** (Ocean Data View) and **NetCDF** (CF) SeaDataNet formats and online accessible upon approval via the SeaDataNet Data Access Services portal (https://cdi.seadatanet.org/search), depicted in *4*. SeaDataNet provides practically all biological and chemical oceanography variables that are measured in situ in European waters, such as alkalinity, acidity and pH of the water column, ammonium and ammonia concentration, chlorophyll pigment concentrations, concentration of organic matter, nitrate concentration, carbon, nitrogen and phosphorus, carbonate system, dissolved gases, other nutrients, other biological measurements, other inorganic chemical measurements, other organic chemical measurements, phytoplankton and microphytobenthos, pigments, and others. Regarding timeliness, all SeaDataNet data are accessible in delayed mode. Currently the time-lag of BGC data is, on average, more than 15 months. For example, in March 2021, only 1/3 of the Baltic Sea data from 2019 is accessible, being the latest observations from December 2019.

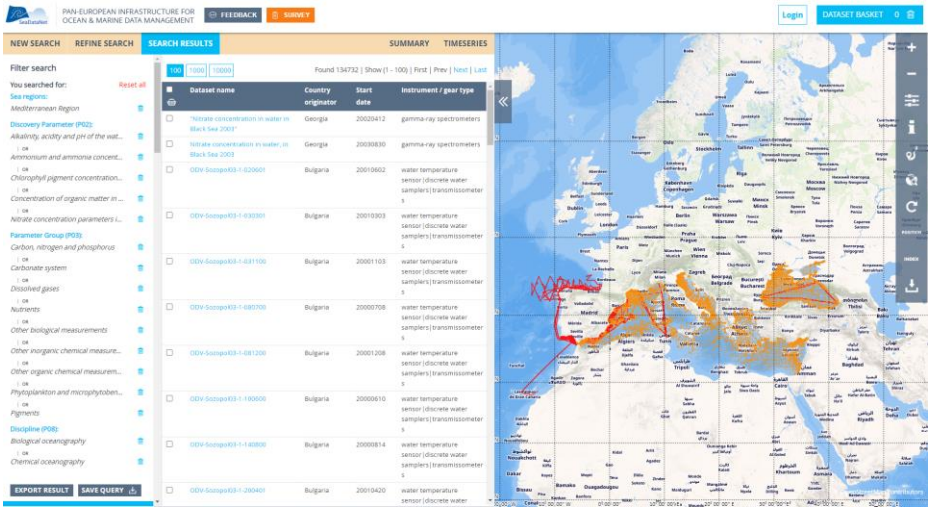


Figure 4. SeaDataNet DATA ACCESS SERVICES - BGC measurement results for the Mediterranean Sea. Source: https://cdi.seadatanet.org/search

EuroGOOS ROOS

EuroGOOS is one of the UNESCO-IOC Global Ocean Observing System (GOOS) Regional Alliance. Five Regional Sea Operational Oceanographic Systems (ROOSs) operate within EuroGOOS: the Arctic (Arctic ROOS), the Baltic (BOOS), the North West Shelf (NOOS), the Ireland-Biscay-Iberian area (IBI-

ROOS) and the Mediterranean (MONGOOS). Cooperation and data sharing within all partners and countries in these regions form the basis of EuroGOOS work. In addition to ROOSs, the EuroGOOS platform Task Teams are operational networks of observing promoting synergy and technological collaboration among European ocean observing infrastructures. Task Team members exchange open-source tools, collaborate in areas of common interest, and jointly make European data available to the EuroGOOS ROOS regional data portals, which in turn are feeding data to the above-described pan-European portals - CMEMS INSTAC, EMODnet and SeaDataNet

BGC observations shared in ROOSs are mainly o are generated from moorings, FerryBox, Argo floats, research vessels, and gliders. Measuring parameters are mainly Chlorophyll fluorescence and dissolved oxygen.

The data portals of the six ROOSes are the following:

- Arctic ROOS: https://arctic.emodnet-physics.eu/ l (see *Figure 5*)
- Baltic ROOS, BOOS: http://www.boos.org/boos-stations/ (see *Figure 6*)
- North-West Shelf ROOS, NOOS: http://nwsportal.bsh.de/ (see *Figure 7*)
- Ireland-Biscay-Iberia ROOS, IBI-ROOS: http://www.ibi-roos.eu/Access-to-data (see *Figure 8*)
- Mediterranean ROOS, MONGOOS: http://www.mongoos.eu/data-center (see *Figure 9*)

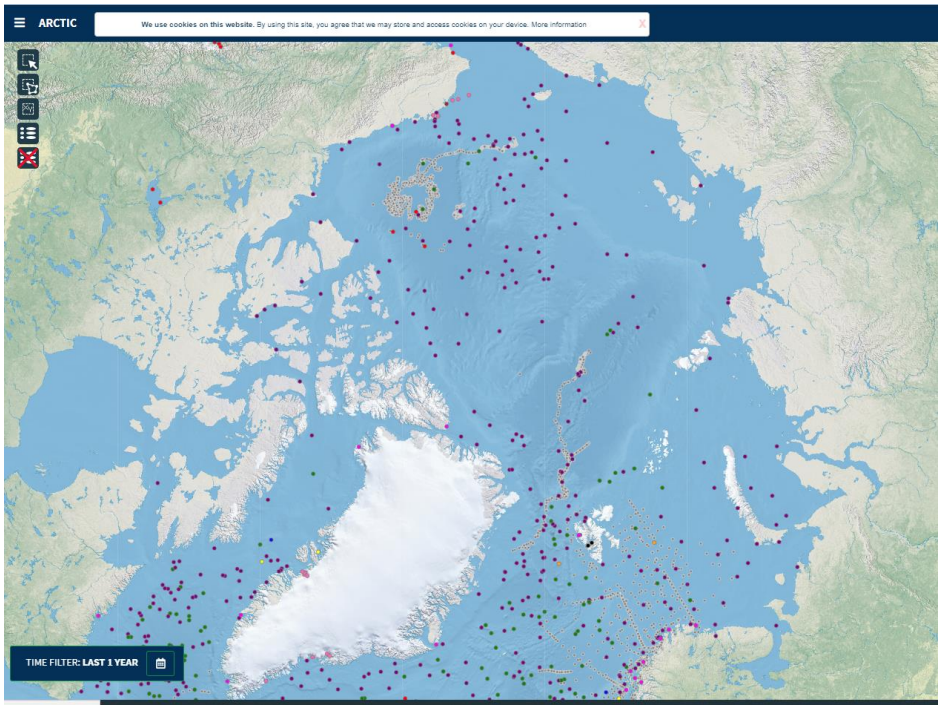


Figure 5. ARCTIC ROOS Data Portal

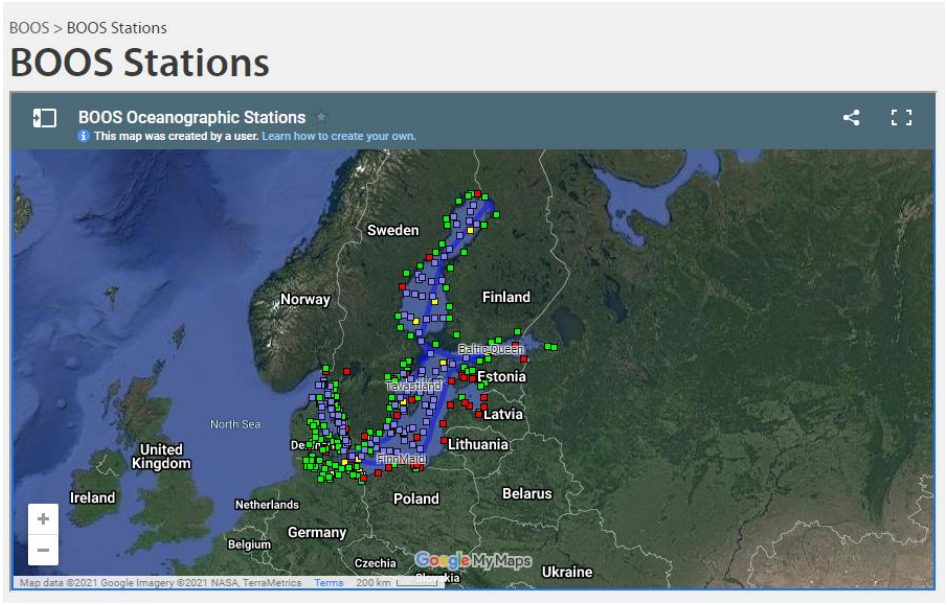


Figure 6. BOOS Data Portal

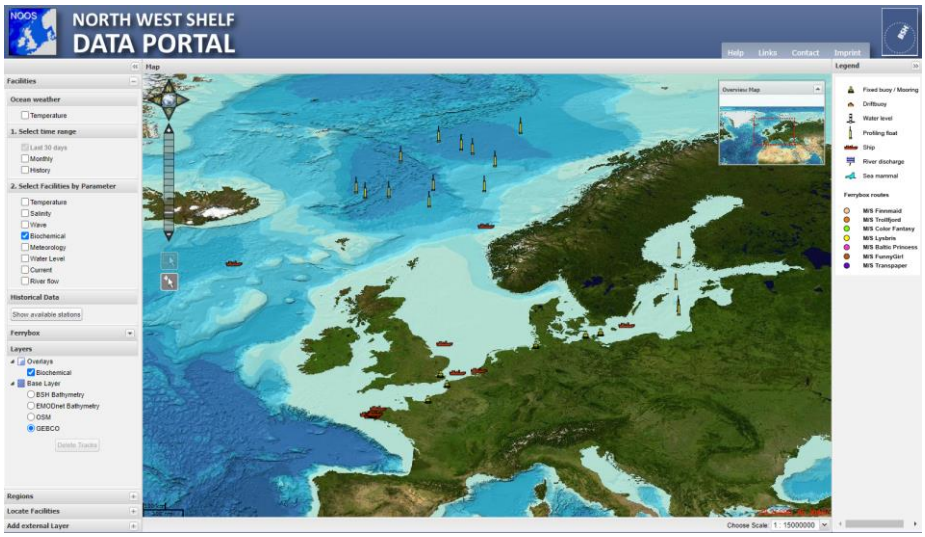


Figure 7. NOOS Data Portal

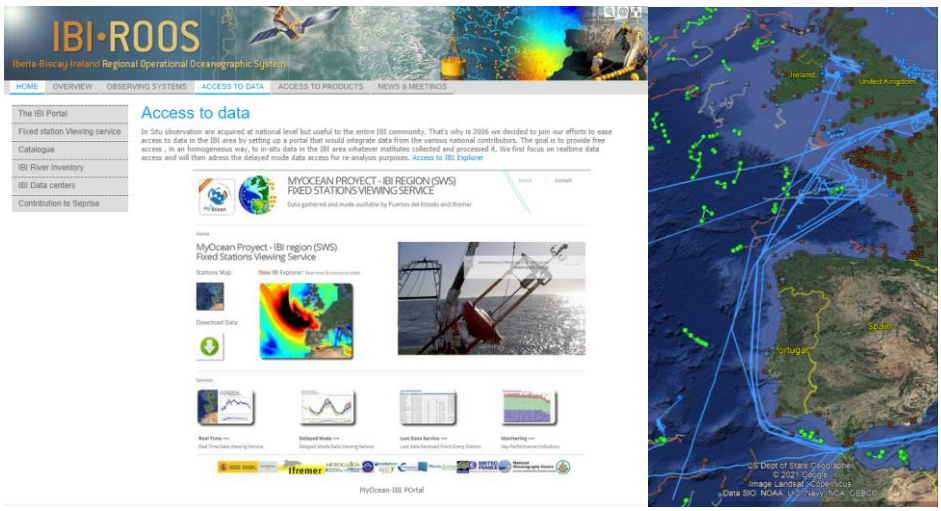


Figure 8. IBI-ROOS Data Portal

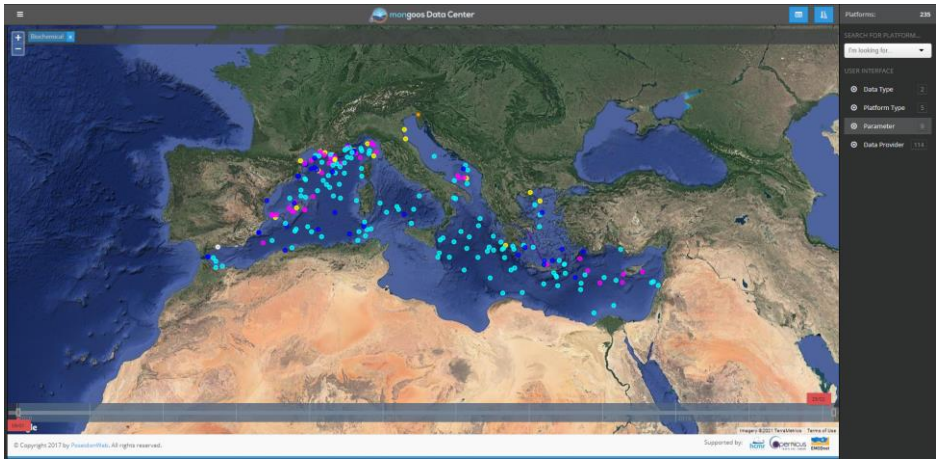


Figure 9. MONGOOS Data Portal.

ICES – International Council for the Exploration of the Sea

The ICES oceanographic database (https://ecosystemdata.ices.dk/) holds a wealth of oceanographic data from 1877 to present. All data are quality controlled according to documented guidelines and visually inspected by experienced staff to further improve the quality of the data. One of the science priorities of ICES is to monitor and explore the seas and oceans in order to track changes in the environment and ecosystems. This includes estuarine, coastal and shelf processes since a main task of ICES is to develop, test and apply indicators to assess the dependence of coastal communities on marine industries. The topic of coastal zone is also reflected in thematic workshops and working groups such as the ICES Working Group for Marine Planning and Coastal Zone Management.

Core parameters held in the ICES oceanographic database are available for download:

- Temperature
- Salinity
- Oxygen
- Phosphate, Total Phosphorus
- Silicate
- Nitrate, Nitrite, Ammonium, Total Nitrogen
- Hydrogen Sulphide
- pH, Alkalinity
- Chlorophyll a
- Secchi depths

The time-lag of ICES BGC data is mostly one year. Major data coverage is northern European Regional Seas, primarily covering the open ocean but also including coastal zones (*Figure 10*). Currently, coastal zone BGC data have not been widely used for model validation in the CMEMS Baltic MFC. More BGC observations in the coastal zone will be important for assessing model product quality in the coastal waters.

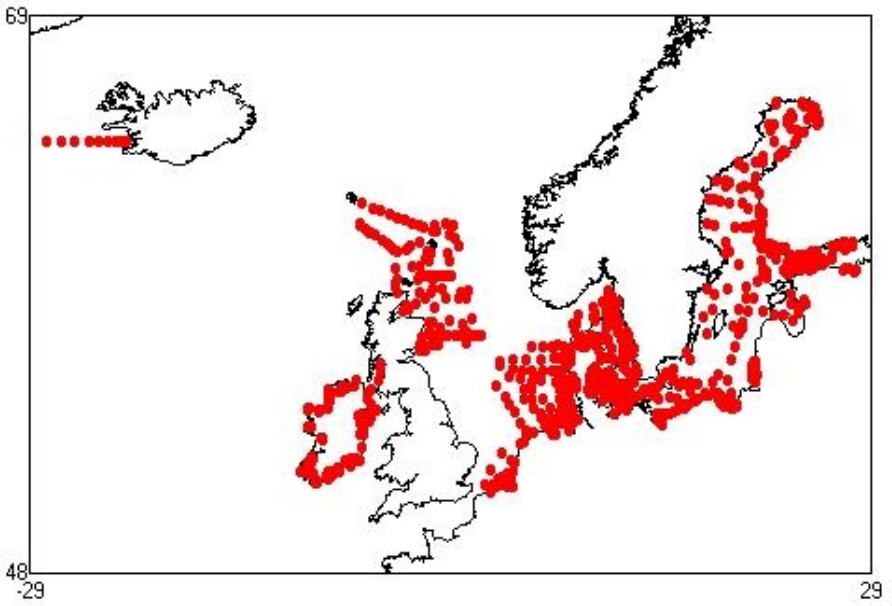


Figure 10. ICES phosphate observation stations during 2019/01/01 – 2020/03/01

Regional Sea Conventions

The Regional Sea Conventions (RSCs) are prominent drivers of coordinated monitoring collaboration among EU member states (and third countries in the same region) in the European regional sea basins. The Regional Sea Conventions perform joint assessments of the State of the Environment (SoE) in their sea basin, in order to support their contracting parties in the implementation of the Marine Strategy Framework Directive (MSFD). While in many cases national monitoring data is operationally ingested by data aggregators, there are additional monitoring data available that are not yet incorporated in those databases. Moreover, the Regional Sea Conventions often produce combined and enhanced datasets which support their assessments. Consequently, the BGC in situ data gathered and produced by Regional Sea Conventions provide an additional source of information to the marine and coastal community. While their data management practices are largely heterogenous, with the exception of OSPAR and HELCOM who commission their data management to ICES and therefore follow similar protocols, they all provide quality controlled and standardized datasets.

In Europe there are four Regional Sea Conventions (see figure below). The Regional Sea Conventions consist of national governments as contracting parties, together with the European Commission (except for the Bucharest convention[5]). The four Regional Sea Conventions are the following:

- OSPAR Convention (OSPAR): The Convention for the Protection of the Marine Environment in the North-East Atlantic of 1992
- Helsinki Convention (HELCOM): The Convention on the Protection of the Marine Environment in the Baltic Sea Area of 1992
- Barcelona Convention (UNEP-MAP): The Convention for the Protection of Marine Environment and the Coastal Region of the Mediterranean of 1995
- Bucharest Convention (Black Sea Commission): The Convention for the Protection of the Black Sea of 1992

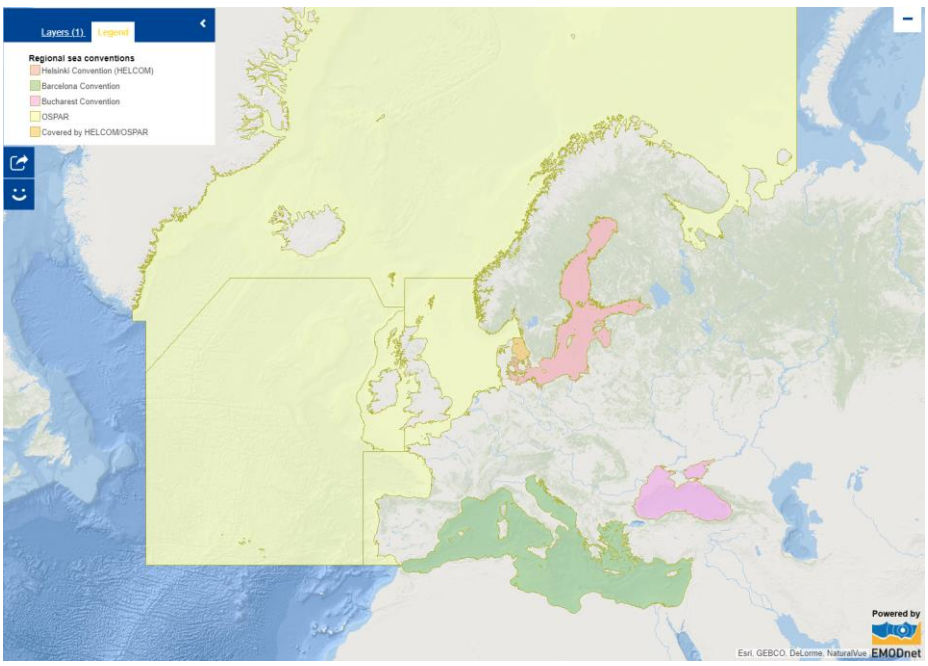


Figure 11. Regional Sea Conventions. Source: https://ec.europa.eu/maritimeaffairs/atlas/maritime_atlas/

[5] https://ec.europa.eu/environment/marine/international-cooperation/regional-sea-conventions/index_en.htm

OSPAR Convention (OSPAR)

OSPAR is the Convention for the Protection of the Marine Environment in the North-East Atlantic by which 15 member states and the EU cooperate. Within the framework of OSPAR the BGC in situ data is gathered by member states across the different thematic work areas of the Convention. Although in principle OSPAR's Joint Assessment and Monitoring Programme should be the guiding framework, it is currently too high level to be guiding the monitoring.

OSPAR Data and Information System (ODIMS)

OSPAR data are managed in a centralised fashion by ICES, and the collected data and information can be accessed via the OSPAR Data and Information System (ODIMS) online tool, see *Figure 12*. The objective of the ODIMS tool is to ensure that data is readily accessible for OSPAR assessments, but it also aims to make OSPAR data findable, accessible and usable to any other user. This objective is included in the OSPAR Data Policy, as stated in the OSPAR Rules of Procedure: *"OSPAR is committed to making as much information as possible publicly available, consistent with achieving other similarly important goals of public policy. The framework for this is set out in Article 9 of the OSPAR Convention."* [6]

Data in ODIMS are public and licensed according to Creative Commons Zero, which is essentially a "no rights reserved" policy that enables data users to freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law. ODIMS has REST API accessibility and adheres to data standards: Geoserver and Geoportal provide the OGC mapping and metadata webservices. Datasets in ODIMS can be queried based on different attributes: per submissions/data deliveries (e.g. for a thematic assessment report), per data manager organization, per OSPAR Committees, or per data themes. A data submission includes spatial data, the originally reported data which has been quality controlled as tabulated data (usually in MS Excel format) and metadata (in xml format). Data themes are not based on scientific classification but rather on policy questions. The following data themes are included:

- Biological Diversity and Ecosystems
- Cross cutting issues, applicable to more than one area of OSPAR work
- Environmental Impacts of Human Activity
- Hazardous Substances
- Hazardous Substances/Eutrophication
- Offshore Industry
- OSPAR Reference files
- Radioactive Substances

The BGC in situ data used for assessments are transferred to ODIMS. A recent comprehensive BGC in situ data collection can be found under IA2017 submission, which is the OSPAR Intermediate Assessment 2017 (an update of the QSR 2010, the next QSR is in 2023[7]). This assessment covers aspects that require BGC in situ data such as:

- ocean acidification,
- biodiversity status (e.g. phytoplankton and zooplankton biomass, abundance and communities),

[6] https://odims.ospar.org/en/data_policy/

[7] https://oap.ospar.org/en/ospar-assessments/quality-status-reports/

- eutrophication (e.g. nutrients, chlorophyll-a, dissolved oxygen),
- or contaminants (e.g. PCB, PAHs, PBDEs, mercury cadmium, lead).

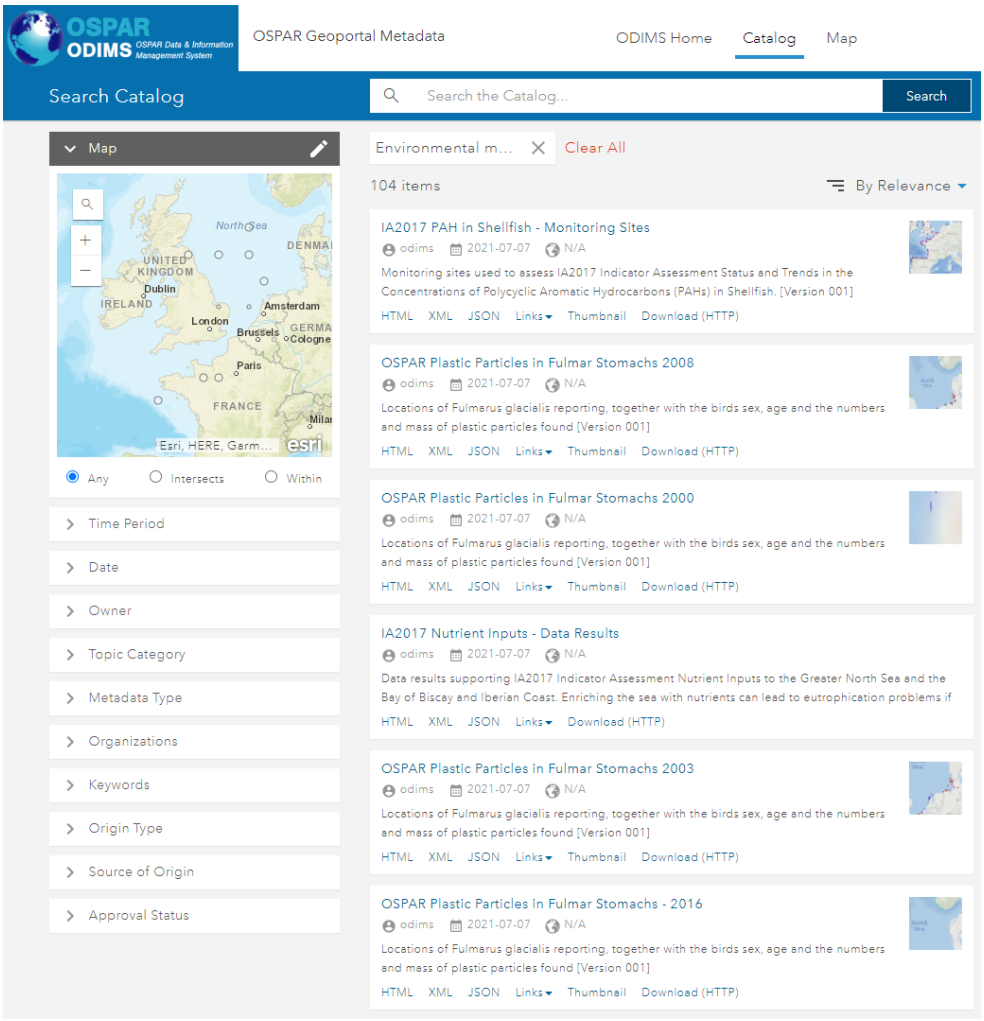


Figure 12. OSPAR Data and Information System (ODIMS) metadata catalogue: https://odims.ospar.org/geoportal/

Cooperation with ICES

OSPAR, similarly to HELCOM (see next section), cooperates with ICES for data management,. Based on the Memorandum of Understanding between the two entities, ICES serves as data centre for data collected under the Co-ordinated Environmental Monitoring Programme (CEMP) under the OSPAR Joint Assessment and Monitoring Programme (JAMP). Therefore, the data collected under the OSPAR Coordinated Environmental Monitoring Programmes for contaminants in biota, sediment and water are quality controlled and hosted at ICES (see figure below). These data are assessed annually by the OSPAR working group on Monitoring and on Trends and Effects of Substances in the Marine Environment (MIME).

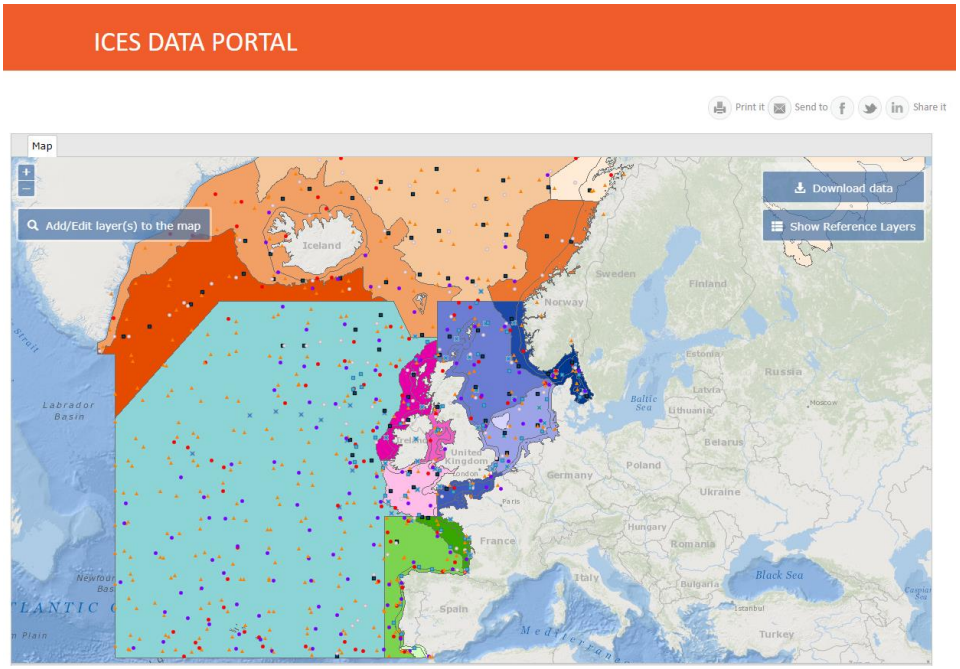


Figure 13. ICES DATA PORTAL – Data selected for OSPAR subregions: https://data.ices.dk/view-map?area=33

The biogeochemical data monitored at OSPAR is submitted to ICES DOME. DOME is the marine environment data portal used by OSPAR, HELCOM, AMAP and Expert Groups in the management of chemical and biological data for regional marine assessments. The DOME Browse (https://dome.ices.dk/browse/) provides users with access to the complete original data files, as uploaded to DOME.

The following OSPAR generated BGC in situ variables can be accessed via DOME:

- Hydrocarbons
- Major organic and inorganic constituents
- Metals
- Nutrients
- Pesticides, herbicides, insecticides
- Pharmaceuticals
- PAHs
- others

The files are zipped and include the requested data file in CSV format, and the data disclaimer. The Data file structure follows the Environmental Reporting Format 3.2, with all record types included in one file. Data uploaded to DOME undergoes quality control procedures by the national submitters and by ICES Data Screening Utility (DATSU). However, as data requirements change over time, there could be inconsistencies in the way data are reported. Another portal in DOME, the DOME Views page (https://dome.ices.dk/views/index.aspx) provides users with data from the Marine Environment Database for contaminants in biota, in sediment, in seawater, as well as phytoplankton community data. The downloaded file includes the requested data file in the CSV format, a PDF document describing the data output fields, and the data disclaimer. The user must accept ICES data policy before downloading data. This data policy is more restrictive than the one of ODIMS, e.g.:

- data sources must be duly acknowledged;
- data Users must respect any and all restrictions on the use or reproduction of data, such as restrictions on use for commercial purposes;
- data Users are obliged to inform ICES of any suspected problems in the data.

The full ICES Data Policy can be accessed on this link: http://ices.dk/data/guidelines-and-policy/Pages/default.aspx. Further details about the dataset and related services can be found at the ICES metadata catalogue. With the above ICES applications downloading data is limited to the first 5000 rows. The complete set of data can be requested via email.

In general, it can be concluded that OSPAR provides findable, and accessible public datasets and has protocols and ongoing partnership (with ICES) to guarantee data management best practices. We should note that the update frequency of OSPAR BGC in situ data is irregular making these datasets not fit for near-real time operational purposes. It is mostly coupled to the preparation of Quality Status Reports (QSR). Many data in this dataset are also available through main European data aggregators, such as EMODnet. Further study should be conducted to identify data records that are missing from the data aggregators.

Helsinki Convention (HELCOM)

HELCOM is the Baltic Marine Environment Protection Commission, also known as the Helsinki Commission (HELCOM). Their BGC in situ data covers the Baltic Sea and earliest data records date back as far as the late 1870s. The monitoring efforts are financed by member states. Until 1992 the monitoring of coastal waters was a national obligation and only the assessment of coastal data had to be reported to the HELCOM. Since 1992, it is also an obligation to conduct monitoring of the coastal waters.

The following BGC variables are monitored:

- Chlorophyll-a,
- Alkalinity,
- Hydrogen Sulphide,
- Total Nitrogen,
- pH,
- Total Phosphorus,
- Ammonium,
- Silicate,
- Nitrite,
- Nitrate,
- Phosphate,
- Salinity,
- Dissolved Oxygen,
- Temperature

A comprehensive database of fixed and non-fixed HELCOM stations (more than 1200 stations) can be accessed through ICES (see Figure below): https://ocean.ices.dk/Helcom/Helcom.aspx?Mode=1. Data

is online instantly accessible in CSV format. ICES currently holds a contract with HELCOM for managing all "at sea" observations collected as part of the HELCOM's COMBINE programme.

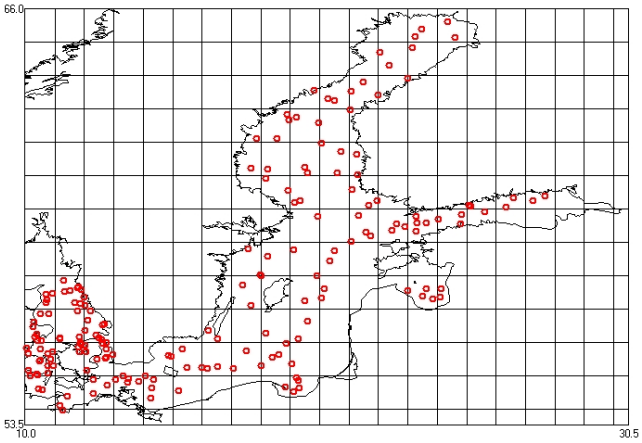


Figure 14. HELCOM stations (fixed and non-fixed) in the Baltic Sea. Source: https://ocean.ices.dk/Helcom/Helcom.aspx?Mode=1

Additional monitoring data can be accessed through the HELCOM Map and Data Service (HELCOM MADS) and from their thematic databases:

- HELCOM COMBINE database - COMBINE monitoring programme
- DOME database - hazardous substances in water, sediment and biota (see previous section)
- HELCOM Biodiversity database - macrospecies observation
- HELCOM PLC Database - waterborne nutrient and contaminant discharges

The HELCOM Map and Data Service provides direct access to all HELCOM geospatial datasets either by ArcGIS Rest interface or OGC WMS Standard. It also includes the output data of assessment reports such as the 'State of the Baltic Sea' reports.

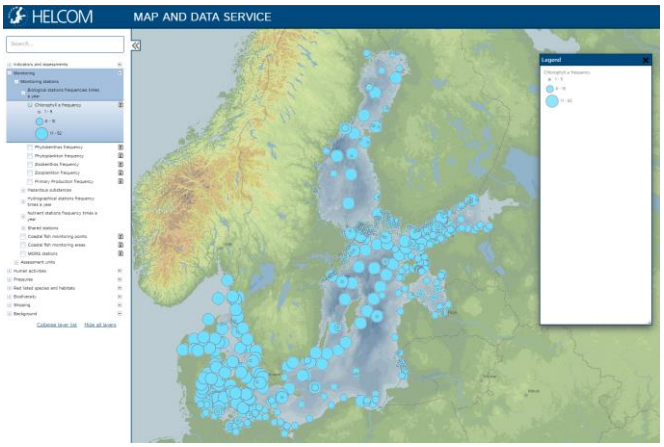


Figure 15. HELCOM Map and Data Service. Yearly frequency of chlorophyll-a observations. Source: https://maps.helcom.fi/website/mapservice/

Barcelona Convention (UNEP-MAP)

The Mediterranean Action Plan (MAP) of the United Nations Environment Programme (UNEP) was the first UNEP initiative under the Regional Seas Programme. MAP assists the Contracting Parties to formulate their national marine environmental policies and to assess and control pollution. MAP

established standardized monitoring and assessment methodologies and tools to deliver assessment reports. These assessments address various aspects, such as the state of the marine and coastal environment, environmental degradation, pressures and impacts on ecosystems, as well as providing an inventory of pollution loads discharging into the Mediterranean Sea. The findings and data of the assessments are assembled into a shared knowledge base for the Contacting Parties to the Barcelona Convention and its Protocols.

The mandate of MAP includes the establishment of a "coordinated programme for research, monitoring, and exchange of information and assessment of the state of pollution and of protection measures"[8]. MAP relies on several components, including MED POL, REMPEC, Plan Bleu/RAC, SPA/RAC, and INFO/RAC. From these components MED POL (Programme for the Assessment and Control of Marine Pollution in the Mediterranean) is responsible for the assessment of the status and trends in the quality of the marine and coastal environment, while INFO/RAC (Information and Communication Regional Activity Centre) helps in the collection and sharing of information to strengthen MAP information management and communication capabilities, and the establishment of a Shared Environmental Information System (SEIS), as described in the see section below. In order to fulfill these goals, the Integrated Monitoring and Assessment Programme of the Mediterranean Sea and Coast (IMAP) was launched. IMAP covers 11 Ecological Objectives and currently 27 Common Indicators. The Ecological Objectives that are most connected to biogeochemical in situ data are EO1-Biodiversity, EO5-Eutrophication or EO9-Pollution. Based on the reported national data and using the IMAP Common Indicators, the Mediterranean Quality Status Reports (MED QSRs) and State of Environment Reports are produced periodically. MED QSRs are produced every six-year (e.g. in 2017, and upcoming MED QSR in 2023).

Gathering data for these Common Indicators requires standardized monitoring and tools. For this reason, the national data reporting is facilitated by the InfoMAP system. The objective of this system is to (1) harmonize data structure and models; (2) create a common catalogue of resources; (3) create a common platform to view, query and analyze data; and (3) produce tools to support data and information dissemination.

The InfoMAP system includes the:

- InfoMAPNode Spatial Data Infrastructure (SDI), an entry point to InfoMAP Spatial Data and Metadata catalogue,
- Data Centre for harmonisation and standardisation of the management of data flows,
- MEDPOL Info System for managing, sharing and preserving pollution monitoring data,
- IMAP Pilot Info System, providing a web-based GIS access to geographic information origination from monitoring programs under IMAP (see detailed explanation below).

Pilot IMAP Compatible Data and Information System (IMAP (Pilot) Info System)

The IMAP Pilot Info System was developed to facilitate access to data and knowledge for decision-makers and the general public. The system is accessible via the following link: http://imappilot.info-rac.org/app/#/ (see figure below). It is fully operational and able to collect data. Currently, only 11 selected IMAP Common Indicators are available in the system:

[8] https://wedocs.unep.org/bitstream/handle/20.500.11822/5251/75wg2_5_mapphasei_eng.pdf

1. *Common Indicator 1*: Habitat distributional range (EO1) to also consider habitat extent as a relevant attribute
2. *Common Indicator 2*: Condition of the habitat's typical species and communities (EO1)
3. *Common Indicator 6*: Trends in abundance, temporal occurrence, and spatial distribution of non-indigenous species, particularly invasive, non-indigenous species, notably in risk areas (EO2, in relation to the main vectors and pathways of spreading of such species)
4. *Common Indicator 13*: Concentration of key nutrients in water column (EO5);
5. *Common Indicator 14*: Chlorophyll-a concentration in water column (EO5)
6. *Common Indicator 15*: Location and extent of the habitats impacted directly by hydrographic alterations (EO7) to also feed the assessment of EO1 on habitat extent
7. *Common Indicator 16*: Length of coastline subject to physical disturbance due to the influence of man-made structures (EO8);
8. *Common Indicator 17*: Concentration of key harmful contaminants measured in the relevant matrix (EO9, related to biota, sediment, seawater)
9. *Common Indicator 21*: Percentage of intestinal enterococci concentration measurements within established standards (EO9)
10. *Common Indicator 22*: Trends in the amount of litter washed ashore and/or deposited on coastlines (EO10);
11. *Common Indicator 23*: Trends in the amount of litter in the water column including microplastics and on the seafloor (EO10);

In the future, the 16 remaining Common Indicators will be added and the selected 11 will be further developed. While the IMAP Pilot Info System is accessible to the general public, login is needed to access the reserved sections. The data reporting is standardized using Excel spreadsheets.

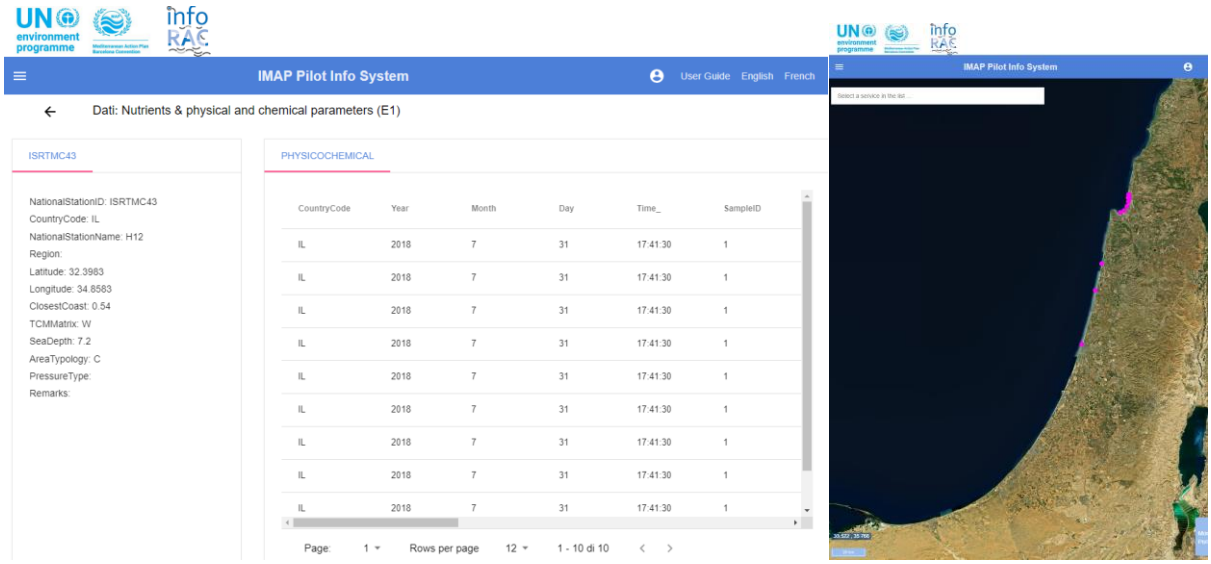


Figure 16. IMAP Pilot Info System. Nutrients, physical and chemical parameters in Israel. Source: http://imappilot.info-rac.org/app/#/

MEDPOL Info System

As mentioned above another important element of the InfoMAP system is the MED POL Info System (http://www.info-rac.org/en/infomap-system/medpol-info-system), which is focusing on pollution monitoring data. Pollution monitoring data is not included in the objective of this report, strictly

speaking, however, due to its importance as model input data, it was decided to include it. The MEDPOL platform contains (1) the national baseline budget database of pollution sources, e.g. pollutant source per country or economic activity (see figure below), and (2) the monitoring database containing the measurements of the pollutants. Pollutants included in the database are:

- BOD4
- Cd
- COD
- Hg
- Pb
- PCB / PCT
- PCDD / PCDF
- Total N
- Total P
- VOC

Data can be downloaded in CSV format. For the national baseline budget database the reported pollutants masses (total kg/y) are only available for assessment years 2003 and 2007. Unfortunately, the contents of the 'Monitoring Data' section are available only for authenticated users (no public access). Registration requests must be sent to mpis@info-rac.org. The database includes both land-based and coastal pollution.

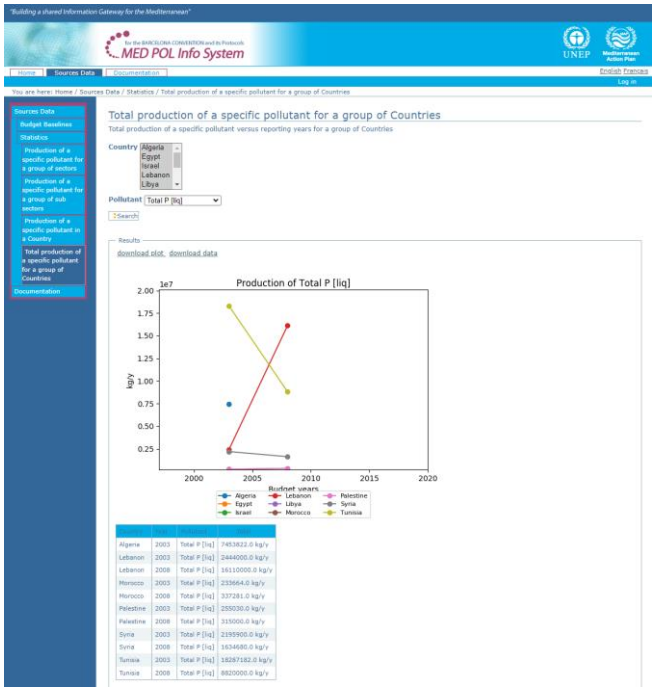


Figure 17. MEDPOL Info System. Total production of a specific pollutant (phosphorus) for a group of countries, Source: http://www.info-rac.org/en/infomap-system/medpol-info-system

EEA Mediterranean database

Apart from the InfoMAP system, there are other forms of databases that include Mediterranean data used for the Barcelona Convention. An example is the EEA Mediterranean database: https://www.eea.europa.eu/data-and-maps/data/meddb. The EEA Mediterranean database (2014

version) includes datasets that were used to derive the Horizon 2020 Mediterranean Initiative indicators. The reported indicators mostly include land-based pollution sources. The database contains national deliveries and UN database, as well as data used for the Reporting Obligation 'Horizon 2020 initiative for a cleaner Mediterranean Sea (SoE)'. The data base is accessible under EEA standard re-use policy: re-use for commercial or non-commercial purposes is permitted free of charge, provided that the source is acknowledged[9]. The data format is Microsoft Excel spreadsheet.

ENI SEIS II South

In order to support Barcelona Convention reporting obligations in the Southern Mediterranean and to implement the SEIS principles, BGC data is collected from southern Mediterranean countries in different thematic topics, including water, within the framework of the ENI SEIS II South project. The Shared Enviromental Information System (SEIS) principles are (1) managing data as near as possible to the source, (2) providing readily available and easily accessible data, (3) producing once and using many times. These principles are implemented in the southern European Neighbourhood Policy (ENP) countries: Algeria, Egypt, Israel, Jordan, Lebanon, Libya, Morocco, Palestine, and Tunisia (see figure below). The data can be accessed via https://eni-seis.eionet.europa.eu/south/areas-of-work/data-and-statistics. Submissions in the Water theme include various indicators such as BOD, TN, TP load discharged from treated and untreated wastewater to the coast. These indicators can be queried per region and year, moreover, there is a quality control flag (validated or non-validated). The data can be downloaded in CSV format. It should be mentioned that the number of submission and the data coverage (both spatial and temporal) is poor.

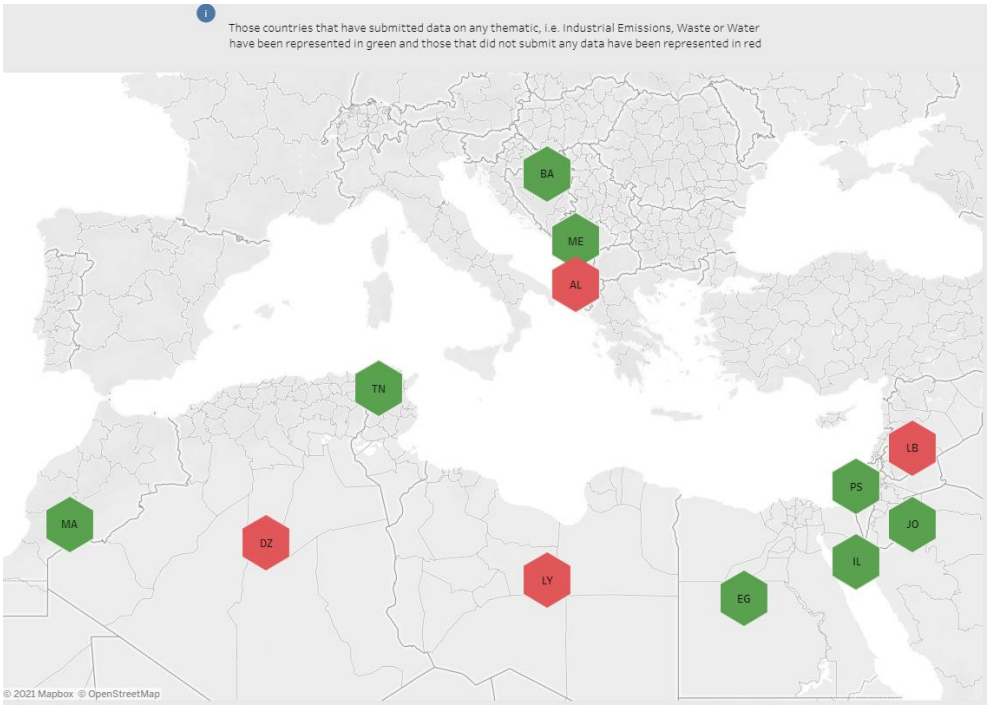


Figure 18. ENI SEIS II South data submissions. Green indicates submission in any of the thematics and red indicates no submission. Source: https://eni-seis.eionet.europa.eu/south/areas-of-work/data-and-statistics

[9] https://www.eea.europa.eu/legal/copyright

The reported indicators were also used for the updated Horizon Assessment report[10] (see figure below) which reviews the progress achieved and challenges ahead in the Union for the Mediterranean (UfM) Horizon 2020 initiative for a cleaner Mediterranean (H2020). The ENI SEIS II South reported indicators were complemented by data reported by EU Member States as well as other data sources, e.g. UN data and H2020 ODYSSEA (see section on ODYSSEA H2020 Project) for a more holistic assessment.

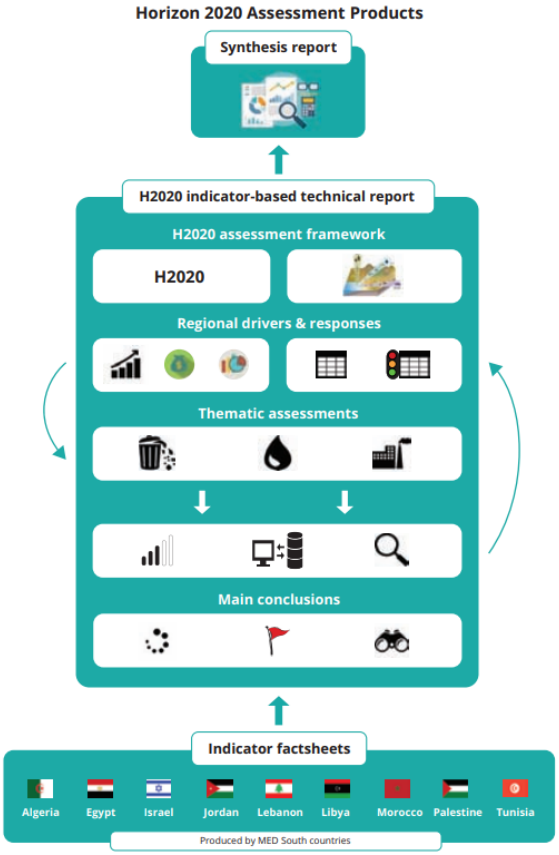


Figure 19. Use of pollution data reported by MED South Countries for the Horizon 2020 Assessment Products. Source: https://www.eea.europa.eu/publications/towards-a-cleaner-mediterranean

Bucharest Convention (Black Sea Commission)

One of the aims of The Convention on the Protection of the Black Sea Against Pollution (Bucharest Convention) is to provide a framework for scientific and technical co-operation and monitoring activities between the 6 member states (Bulgaria, Georgia, Romania, Russian Federation, Turkey, Ukraine). This regional environmental monitoring in the Black Sea is carried out in framework of the Black Sea Integrated Monitoring and Assessment Programme (BSIMAP), implemented by the Black Sea Commission (BSC) since 2001. BSIMAP covers the main transboundary environmental topics in the Black Sea region: eutrophication, water pollution and water quality, biodiversity change and habitats. BSIMAP is based on national monitoring programs financed by the Black Sea states. BSIMAP supports the State of Environment (SoE) of the Black Sea reports by provisioning quality assured data for scientifically-based and validated indicators to the Contracting Parties. The choice of parameters to monitor is related to the main environmental problems recognised in the Black Sea region and these are re-evaluated every 5 years.

[10] https://www.eea.europa.eu/publications/towards-a-cleaner-mediterranean

The data from the BSIMAP reports are loaded into the Black Sea Information System (BSIS). The latest access point to the BSIS is under the name of Black Sea Water Quality Database (http://blackseadb.org/), although it is currently undergoing development, see further details below. BSIS should serve as official entry point to data in the Black Sea. The main sources of data for the BSIS are the following:

- National monitoring Programs,
- Black Sea regional component of monitoring program,
- Scientific surveys and projects,
- Black Sea scientific conference,
- Relevant scientific publications.

Currently BSIS is functioning in delayed mode and does not possess modern GIS interface. Nevertheless, users can query and download data from BSIS. The BSIS needs further improvement as its components were developed under different projects and in different time periods, using various computer hardware and software. The latest improvement of the database is implemented in the EMBLAS project (https://emblasproject.org/), see also details below.

Regional Database on Pollution

The Regional Database on Pollution (**Error! Hyperlink reference not valid.**) is one of the components of the Black Sea Information System. It is hosted and maintained by UKRSces, a Ukrainian Institute. It contains data on pollutants in water, sediments and biota collected in countries in process of implementation of the BSIMAP and annually reported to the BSC. The data on nutrients constitute the major part of the database but other BGC variables are also available. Data can be downloaded per data provider (country or organization), location (region, user defined area or specific stations) or by parameter groups. The parameter groups are:

- Hydrochemistry (e.g. Alk, BOD5, H2S, O2, pH, Secchi depth, TOC, TSS)
- Nutrients (e.g. NH4, NO2, NO3, N total, P total PO4, SiO4)
- (Heavy) Metals (e.g. Al, As, Cd, Cr, Fe, Hg, Pb, Zn)
- Pesticides
- PCBs, PAHs
- Phenols
- Detergents
- Petroleum hydrocarbons (e.g. TPHs)
- Radionuclides
- Photosynthetic pigments (e.g. Chlorophyll-a)

It requires a Login and Password (“Public”) to enter the database. Selected datasets can be exported in .xls (MS Excel) format.

Figure 20. Regional Database on Pollution (http://rdbp.sea.gov.ua/)

Black Sea Water Quality Database

The latest Black Sea State Of Environment Report is primarily based on data from the joint **Black Sea Water Quality Database** (http://blackseadb.org/), see figure below. The Black Sea Water Quality Database was developed within the framework of the EU/UNDP EMBLAS-II and EMBLAS-Plus projects (https://emblasproject.org/). The datasets are collected through the Excel-based Data Collection Templates used for the Joint Black Sea Surveys and National Monitoring Studies in Georgia, Russian Federation and Ukraine. One of the main objectives of the EMBLAS project is to build national capacities and skills in use of the up-to-date monitoring and analytical techniques aligned with the MSFD and WFD principles and methodologies and the Black Sea Integrated Monitoring and Assessment Programme (BSIMAP), including environmental data sharing and assessment. The updated Black Sea Database merges previous databases, including the Regional Database on Pollution (see above):

- Emblas
- UkrSCES Historical data (monitoring)
- UkrSCES Historical data (cruises)
- ONU Histroical data (monitoring)
- IMB historical data (monitoring)
- NEA historical data
- Shirshov Institute of Oceanology, Russian Academy of Sciences (SIO-RAS) historical data
- Sustainable Ocean Initiative (SOI) historical data
- ANEMONE

In general, the Black Sea was relatively well monitored until the 1990s but there is a considerable gap in later years. While, the Black Sea Water Quality Database is a useful source of data, its coverage is irregular in time and space. The data portal does not require authentication; however, some content is restricted[11]. Selected datasets can be exported in .xls (MS Excel) format and the metadata standards follow ISO-19115 standard. This joint database includes data in water, sediment, and biota and

[11] https://catalogue.odis.org/view/907

contains a wide range of data on physicochemical, chemical and biological parameters. The observed BGC in situ parameters are:

- Biodiversity - water column
    - Phytoplankton
    - Zooplankton (Macrozooplankton, Microzooplankton, Mesozooplankton)
    - Microbial communities (biodegradation, taxonomy)
- Biodiversity - seabed
    - Macrozoobenthos
    - Meiobenthos
    - Macrophytobenthos
    - Microbial communities
- Biodiversity (marine mammals, fish, birds species)
- Eutrophication
    - Nutrients (nitrate, nitrite, ammonia, total N, total P, organic N, organic P, silicon, phosphate)
    - Chlorophyll-a
    - General physico-chemical parameters (dissolved oxygen, BOD, Total Suspended Solids, hydrogen sulphide, pH, salinity, alkalinity)
- Contaminants (e.g. Organochlorine pesticides (OCPs), Polycyclic Aromatic Hydrocarbons (PAHs), Organic pollution, Trace metals, PCBs, Other parameters)

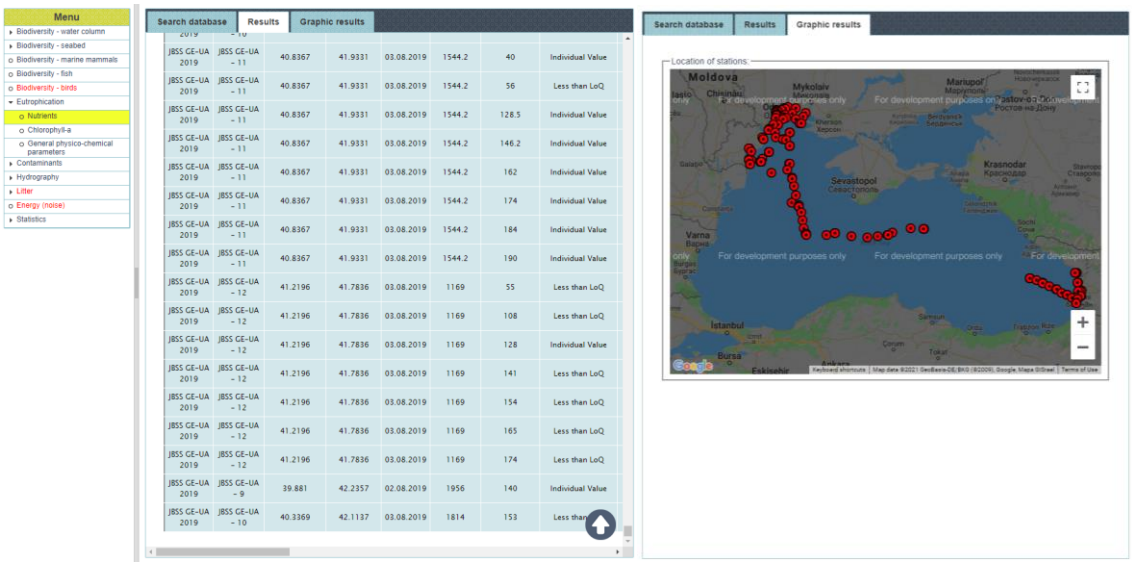


Figure 21. Black Sea Database (http://blackseadb.org/). Tabular (left) and geographic (right) result of the query.

European Marine Research Infrastructures

Data aggregators heavily rely on data coming from European marine research infrastructures. Within this task representatives of several major marine research infrastructures have been contacted with a survey to report on their BGC in situ observations. Representatives of DANUBIUS-RI, JERICO-RI, and ICOS Ocean Thematic Centre contributed to this section, nevertheless a description of the BGC in situ data in EuroArgo and EMSO is also included. The research infrastructures are listed in the order of relevance to the coast: Danubius-RI (transitional river-delta-sea systems), Jerico-RI (coastal), EuroArgo (partly coastal but more regional and open ocean), EMSO and ICOS (mostly open ocean).

Danubius-RI

The International Centre for Advanced Studies on River – Delta – Sea Systems DANUBIUS-RI is a pan-European research infrastructure supporting research on large river-sea systems. Twelve DANUBIUS-RI distributed supersites (see Figure 22) will provide natural laboratories for observation, covering systems from river source to transitional waters and coastal seas, and measuring physical, chemical and biological variables.

DANUBIUS-RI offers a Data portal to integrate existing data across sectors and disciplines, supplemented by new data from Earth observation, in situ sensors, dedicated sampling campaigns and models. The data portal provides access to a DANUBIUS-RI metadata base covering digital data from these digital sources of data. Near real time data flows from autonomous stations deploying in situ sensors and from satellite data (Sentinel-2 and Sentinel-3) provides data on BGC. This will be supplemented with data from sampling campaigns across river, estuarine/delta/lagoon and coastal cruises. Modelling is used to fill the observation gaps and test what-if scenarios to support science, policy and coastal management. DANUBIUS-RI also provides a set of harmonized regulations, methods, procedures and standards, collectively called the Commons. The DANUBIUS-RI Commons, managed by the observation, analysis and modelling nodes will ensure comparability of data across the river-sea continuum and across all supersites.

The Danubius Data Centre (https://danubius-ri.eu/architecture.html) is responsible for data availability and processing, and it constitutes the single point of access for the users to all services. These services should include the distribution of collected data, aggregated data from local and distributed storages, and search functionalities to access the data, among others.

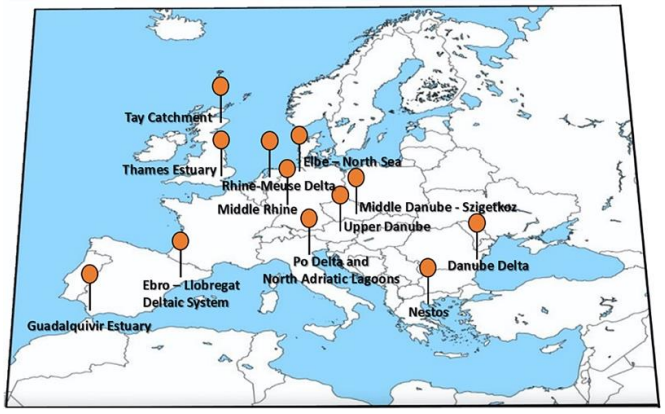


Figure 22. DANUBIUS-RI super sites (source: https://www.danubius-ri.eu/architecture.html)

DANUBIUS-RI target BGC variables include:

- Chlorophyll-a,
- Algal size classes
- Primary production
- Turbidity
- Total suspended matter
- Total suspended sediment
- Grain Size Distribution of Suspended and Bedload Sediment
- Colour dissolved organic matter
- Total Organic Carbon
- Dissolved Organic Carbon
- Temperature
- Conductivity and Salinity
- pH
- Total Phosphorus
- Soluble Reactive Phosphorus
- Ammonium
- Silicate
- Nitrite
- Nitrate
- Phosphate
- Dissolved Oxygen
- Temperature

JERICO-RI

JERICO-RI is the Joint European Research Infrastructure of Coastal Observatories, implemented through the JERICO-FP7, JERICO-NEXT and JERICO-S3 projects. JERICO-RI integrates several observing platform types such as fixed buoys, piles, moorings, drifters, Ferryboxes, gliders, and the associated technologies dedicated to the monitoring of the European coastal waters. This integrated observing system of systems monitors physical, chemical and biological processes characterizing the European coastal waters and aims at a better integration of marine biology with physical and chemical oceanology.

Major user-driven improvements will be realised in terms of observing the complexity of coastal seas and continuous observation of the biology, access to facilities, data and services, best practices and performance indicators, innovative monitoring strategies, cooperation with other European RIs (EuroARGO, EMSO, AQUACOSM, DANUBIUS, ICOS, EMBRC, LIFEWATCH) and international scientific communities, industry and other stakeholders, and aligning strategy with COPERNICUS/CMEMS, EMODNET and GEO/GEOSS.

BGC variables monitored within the JERICO-RI are:

- Alkalinity, acidity and pH of the water column
- Chlorophyll pigment concentrations in water bodies
- Dissolved oxygen parameters in the water column
- Electrical conductivity of the water column

- Light absorption in the water column
- Raw suspended particulate material concentration sensor output
- Variable fluorescence parameters
- Phytoplankton biomass and diversity
- Zooplankton biomass and diversity
- Fish abundance and distribution
- Marine turtles, birds, mammals abundance and distribution
- Hard coral cover and composition
- Seagrass cover and composition
- Macroalgal canopy cover and composition
- Benthic invertebrate abundance and distribution

BGC in situ data is instantly accessible through the JERICO-RI Data Access page: https://www.jerico-ri.eu/data-access/ or the JERICO-RI Catalogue (https://www.jerico-ri.eu/jerico-ri-catalogue/) (see Figure 23), which contains a structured overview of involved coastal platforms and developed services. Items in the catalogue are all searchable via metadata search. The JERICO-RI Biology Catalogue is already provided through EMODnet. Moreover, the JERICO-RI Survey on Coastal platforms will be available through the Chemistry Catalogue.

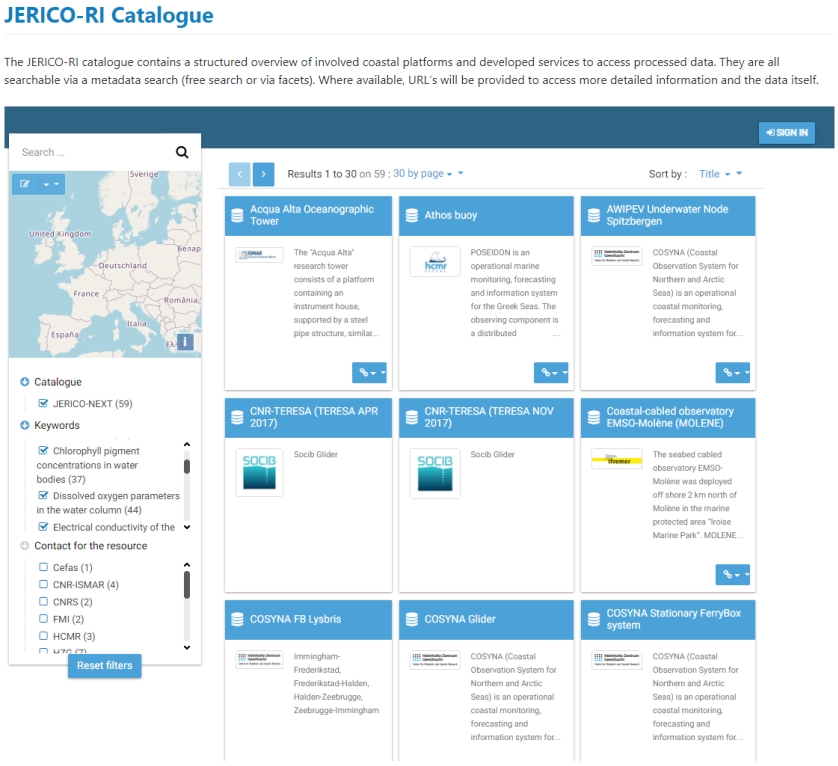


Figure 23. JERICO-RI Catalogue. Source: https://www.jerico-ri.eu/jerico-ri-catalogue/

Euro-ARGO

The Argo network is a global array of more than 3500 autonomous instruments, deployed over the world ocean, reporting subsurface ocean properties via satellite transmission links to data centres. The Euro-Argo ERIC (European Research Infrastructure Consortium) allows active coordination and strengthening of the European contribution to the international Argo programme.

BGC variables monitored by Argo floats are:

- CDOM,
- pH, nitrate,
- PAR
- Dissolved oxygen
- Chlorophyll-a

Data is accessible at: **https://dataselection.euro-argo.eu/** in ASCII and netCDF format, see *Figure 24*.

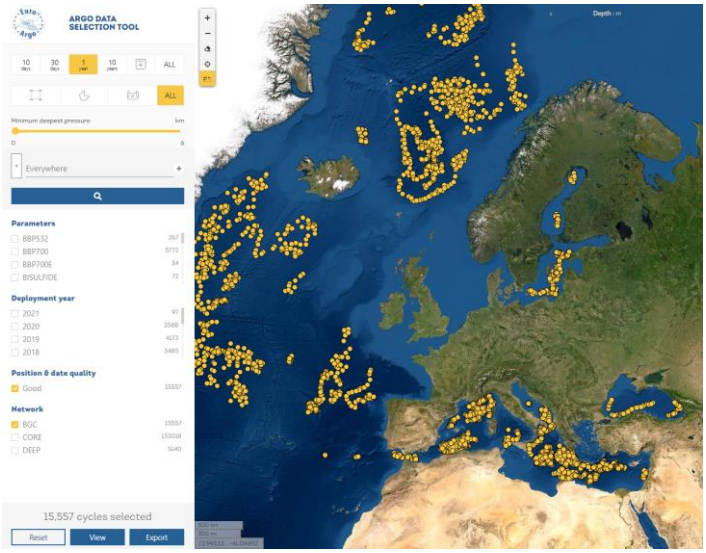


Figure 24. BGC ARGO floats in the last 1 year. Source: https://dataselection.euro-argo.eu/

EMSO

EMSO is the European Multidisciplinary Seafloor and water column Observatory (EMSO). EMSO consists of regional facilities placed at key sites around Europe, from Northeast to the Atlantic, through the Mediterranean, to the Black Sea. Observatories are platforms equipped with multiple sensors constantly measuring biogeochemical and physical parameters.

Monitored BGC variables within EMSO are:

- Chlorophyll-a,
- dissolved oxygen,
- turbidity

The data is online instantly accessible via https://data.emso.eu/ (see *Figure 25*) in CSV or netCDF format.

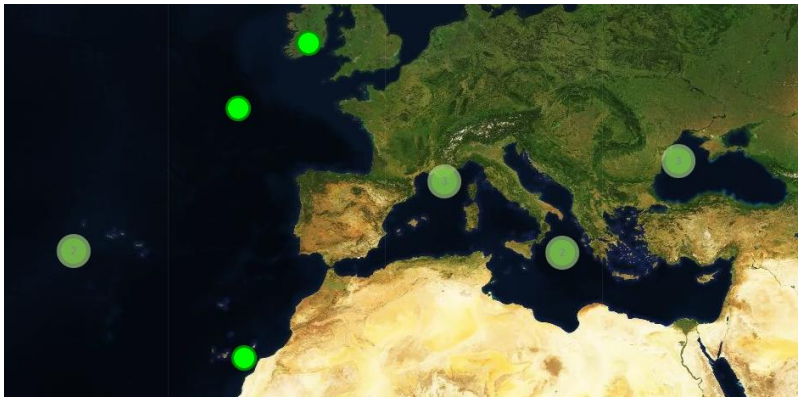


Figure 25. Location of EMSO facilities. Source: https://data.emso.eu/

ICOS Ocean Thematic Centre

The Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (OTC) is a European research infrastructure providing long-term oceanic observations on the global carbon cycle and climate-relevant gas emissions. ICOS OTC currently coordinates twenty-one ocean stations from seven countries monitoring carbon uptake and fluxes in the North Atlantic, Nordic Seas, Baltic, and the Mediterranean Sea. Measuring methods include sampling from research vessels, moorings, buoys, and commercial vessels that have been equipped with carbonate system sensors.

Figure 26. ICOS OTC station network. Source: https://www.icos-cp.eu/observations/carbon-portal

Apart from the physical variables, ICOS OTC stations monitor the following BGC variables:

- Dissolved oxygen
- Dissolved nutrient
- Partial pressure of carbon dioxide (pCO2)
- Dissolved organic carbon (DOC)
- Dissolved inorganic carbon (DIC)
- pH
- Total alkalinity

ICOS OTC monitoring programme is ongoing and near real time data is online through the ICOS Carbon Portal (see *Figure 26*): https://www.icos-cp.eu/observations/carbon-portal. The Carbon Portal is a single access portal for all ICOS data products where FAIR data management principles are followed. **Monitoring funds generally run for a 5-year period, but currently the sustainability of long-term funding is a major problem.** Each institute that is a member of the ICOS Ocean Thematic Centre coordinates their observations before submitting for quality control and into a data repository (ICOS Carbon Portal and/or SOCAT).

Other BGC in situ observations sources

Apart from the above described well known data aggregators, research infrastructures, regional conventions and international observations programmes, there are numerous monitoring campaigns or time-limited research projects that produce and disseminate BGC in situ observations in the European waters. In this section we include some examples of additional BGC in situ data sources.

CORIOLIS

CORIOLIS coordinates the French contribution to the WMO-IOC Joint Technical Commission for Oceanography and Marine Meteorology (JCOMM). It maintains in situ networks in operation at global and national scale. More specifically, CORIOLIS coordinates the acquisition of in situ measurements with the objective to sustain and enhance the Argo network; acquire in situ data from ships (VOS and research vessels); maintain the SO-SSS network, for collecting in situ surface data during transit of research or voluntary merchant vessels; maintain PIRATA observation system, for collecting in situ data from a moored array in Tropical Atlantic; deploy surface drifters and moored buoys as a contribution to DBCP and OceanSITES; operate and maintain the French Glider fleet; and operate and maintain high frequency coastal networks (http://www.coriolis.eu.org/).

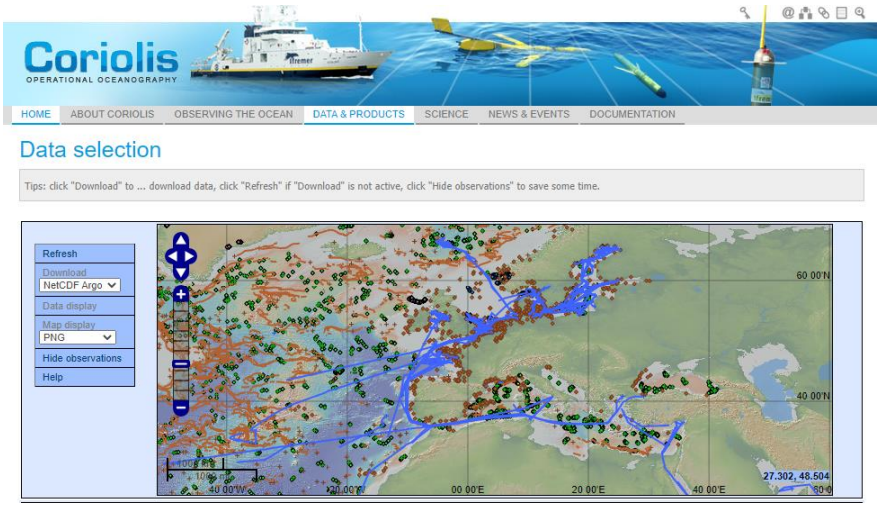


Figure 27. CORIOLIS Data Portal (http://www.coriolis.eu.org/Data-Products/Data-Delivery/Data-selection)

CORIOLIS data can be accessed through the Coriolis Data Center via various sites:
http://www.coriolis.eu.org/Data-Products/Data-Delivery/Data-selection (see *Figure 27*)
http://www.ifremer.fr/co-dataSelection/?theme=myocean
http://www.ifremer.fr/co-argoFloats/ (ARGO data selection)
http://www.ifremer.fr/co-dataSelection/?theme=gosud (GOSUD data selection)
http://www.ifremer.fr/co-en/ (Eulerian networks and fixed buoys)

CORIOLIS has a specific coastal component called Coastal Coriolis (https://www.coriolis-cotier.org/), which covers the North Western Mediterranean Sea, Bay of Biscay, and the English Channel. Coastal Coriolis is including High-Frequency coastal networks aiming to observe the coastal environment from the coast to the open ocean including the continental shelf. Coriolis Coastal monitors BGC variables,

namely oxygen, chlorophyll, fluorescence and turbidity, using four main operational platform types (see *Figure 28*):

- Eulerian platforms: Coastal OceAn observing SysTem – High Frequency, SMATCH buoys, HF SOMLIT stations , MOOSE network, Meteo-France moorings),
- measurements from vessels of opportunity (FerryBox, RECOPESCA environment network),
- coastal profilers (ARVOR-C and ARVOR-Cm),
- drifting systems (surface drifters and sub-surface Gliders).

Quality controlled Coastal Coriolis datasets are online instantly accessible as ASCII or NETCDF-Argo format (with only good quality codes or all quality codes) via https://data.coriolis-cotier.org/, see *Figure 28*.

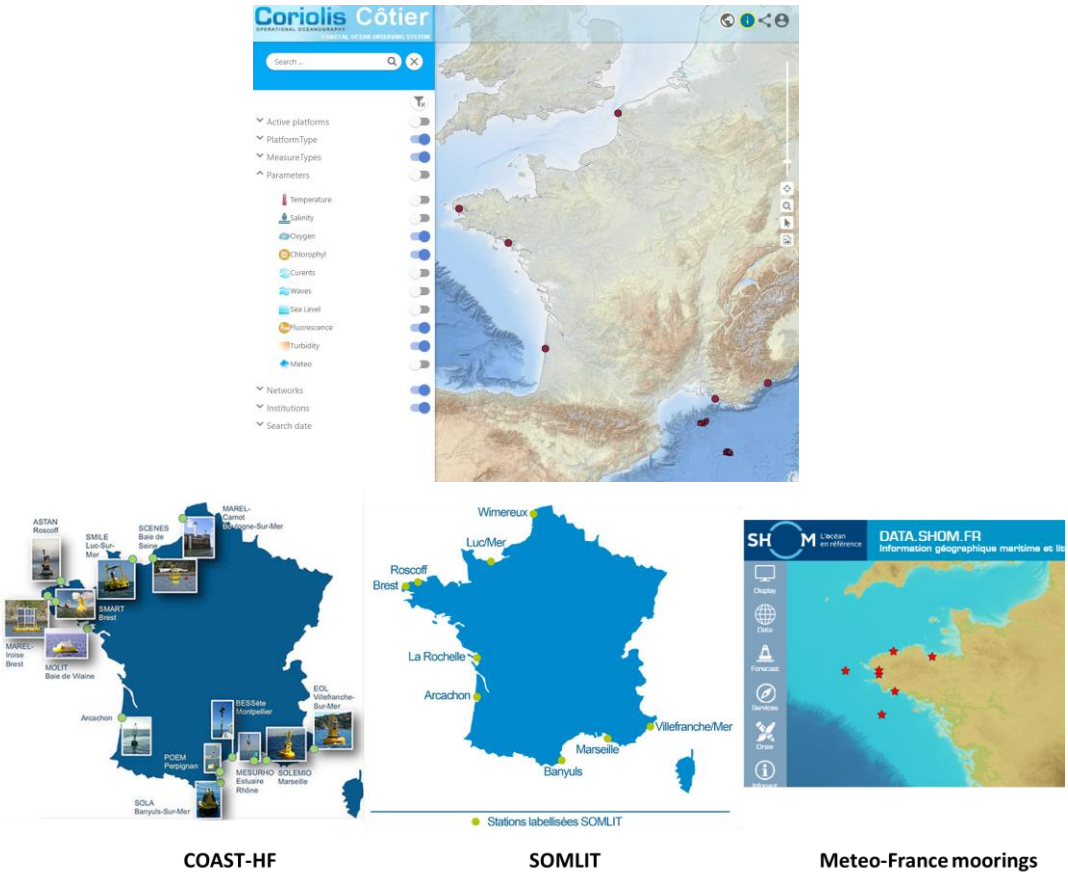


Figure 28. Coastal Coriolis Data portal (https://data.coriolis-cotier.org/) (top), and Coastal Coriolis platforms (bottom).

OceanSITES

OceanSITES is a worldwide system of long-term, open-ocean reference stations (see *Figure 29*) measuring dozens of variables and monitoring the full depth of the ocean from air-sea interactions down to the seafloor. It is a network of stations or observatories measuring many aspects of the ocean's surface and water column using, where possible, automated systems with advanced sensors and telecommunications systems, yielding high time resolution, often in real-time. Measurements cover biogeochemistry, and parameters relevant to the carbon cycle, ocean acidification, the ecosystem, and geophysics.

OceanSITES uses NetCDF (Network Common Data Form) format. The data can be accessed via FTP (ftp://ftp.ifremer.fr/ifremer/oceansites) or OPeNDAP/THREDDs (http://tds0.ifremer.fr/thredds/CORIOLIS-OCEANSITES-GDAC-OBS/CORIOLIS-OCEANSITES-GDAC-OBS.html).

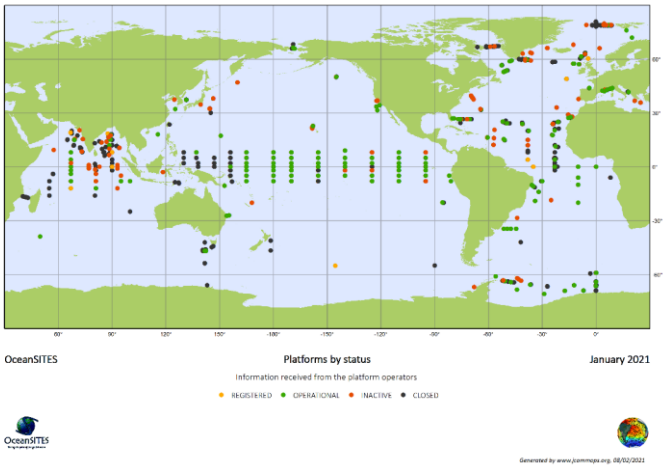


Figure 29. OceanSITES stations. Source: https://www.ocean-ops.org/board/?t=oceansites

NBSC (North Sea Biogeochemical Climatology)

Apart from data aggregation, which helps to provide centralized access to in situ data, another way of combining and fusing observations is deriving climatologies. A good example is the North Sea Biogeochemical Climatology (NSBC), depicted in *Figure 30*.

The NSBC has been constructed by bringing together in situ observations from various data centers providing a suite of biogeochemical parameters for the wider North Sea Region. The NSBC data was derived from the NOWESP project, World Ocean Database 2013, Uwe Brockmann. DOD, PANGAEA, ICES, and EMODNet. The data set consists of gridded fields of climatological mean values of biogeochemical parameters for the time period of 1960-2014. Data processing steps include formatting, elimination of duplicates, quality controls, vertical, spatial and temporal interpolation.

The Biogeochemical Climatology includes the following BC variables:

- ammonium,
- chlorophyll-a,
- nitrate
- nitrite,
- phosphate,
- oxygen
- silicate

Unrestricted data access to NSBC is provided via Live Access Server, HTTP, and OPeNDAP:

- https://icdc.cen.uni-hamburg.de/las/getUI.do?dsid=id-015565262b&catid=1D89D87646D19EABC2F3AC9C1F8D3CAC&varid=ammonium_mean-id-015565262b&plot=XY_zoomable_image&view=xy&auto=true (Live Access Server)
- https://icdc.cen.uni-hamburg.de/thredds/catalog/ftpthredds/nsbc/catalog.html (HTTP)
- https://icdc.cen.uni-hamburg.de/thredds/aggregationNsbcCatalog.html (OPeNDAP)

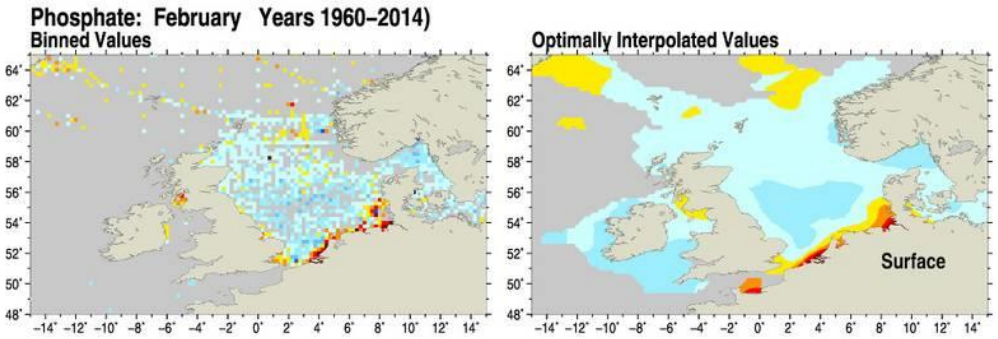


Figure 30. NSBC -Phosphate climatology. Souce: https://icdc.cen.uni-hamburg.de/en/nsbc.html

ODYSSEA H2020 Project

The European BGC in situ monitoring network is expanding not only through national funds but through European R&D projects such as the ones in the H2020 programme. These efforts are crucial to address localized knowledge gaps. A good example is the H2020 ODYSSEA projects in the Mediterranean, which primarily targets data poor regions such as the Eastern and Southern Mediterranean.

ODYSSEA fills selected high priority coastal gaps around the Mediterranean (Spain. Italy, Greece, Turkey, Israel, Egypt, Algeria, Tunisia, Morocco – see *Figure 31*) through a network of coastal observatories where in situ sensors are deployed. Two types of data collection systems are deployed in ODYSSEA Observatories: fixed (moored) systems and mobile systems. Fixed systems include surface monitoring systems in Greece, Turkey, Israel, Algeria, Tunisia, Morocco, and Egypt; as well as a benthic monitoring system in Greece. To reduce costs, to overcome technical problems with cable technology, and to ensure active participation of end-users on ODYSSEA platform, existing facilities (onshore and offshore), such as oil and gas terminals and rigs, mariculture installations, ports and harbours, will be used to deploy static sensors. This further increases the uptake of BGC in situ data by industry.

Data from mobile systems as Unmanned Underwater Vehicles (UUVs) operating along the shelf zone and the bordering troughs of observatories are integrated in the Marinomica platform (https://marinomica.com/).

Figure 31. ODYSSEA coastal observatories and trajectories of completed glider missions (Morocco – left, Thracian Sea - right).

The autonomous system used in ODYSSEA is Alseamar's SEAEXPLORER glider with payloads to measure temperature, salinity, pH, dissolved oxygen, chlorophyll-a, turbidity, CDOM. Two gliders and three payloads have been provided to the ODYSSEA observatories. Successfully completed glider missions took place in the Thracian Sea (Greece), and in Morocco. The next observatories to receive gliders will be Israel and Tunisia.

All in situ data gathered in ODYSSEA will be shared with the community through the CMEMS In Situ TAC and MONGOOS. Operational data from ODYSSEA deployed fixed moorings are already available via the CMEMS Mediterranean In situ TAC (managed by HCMR, Greece). Metadata from the glider missions will be shared with MONGOOS in the first half of 2021, and the actual data will be shared after the end of the project (November 2021).

Efforts in transitional waters - CERTO

European coastal BGC in situ monitoring capacity is the least developed in transitional waters. This is due to the high gradient of ecosystems and environmental variables (e.g. salinity or turbidity) in transitional waters. The CERTO project (Copernicus Evolution - Research for harmonised and Transitional water Observation) will try to address this issue and provide a harmonized capability to monitor water quality from lakes, through deltas, coastal waters and to the open ocean. CERTO currently focuses on 6 transitional water bodies (see *Figure 32*): Curonian lagoon, Russia/Lithuania; Elbe estuary, Germany; Razelm-Sinoe lagoon, Romania; Tagus estuary, Portugal; Tamar estuary, UK; and Venice lagoon, Italy. Field measurements will be made available on the project portal: https://certo-project.org/Home

Figure 32. CERTO case study areas.

Other relevant data sources in the Black Sea

Apart from the Black Sea Convention data (see Section on the Bucharest Convention (Black Sea Commission)), there are other useful sources of BGC in situ data in the Black Sea. Nowadays, the INSTAC of CMEMS, EMODNET and SeaDataNet, provide the most comprehensive datasets including near-real time delivery. Nevertheless, further study should be conducted to identify data records that are available at the Regional Sea Conventions but are missing from the data aggregators. These data assembly programs guarantee Black Sea data availability, operational reliability, efficiency, time consistency, and space consistency. Despite the efforts, a harmonized, sustainable, and cost-effective Black Sea Observing System is still under development[12].

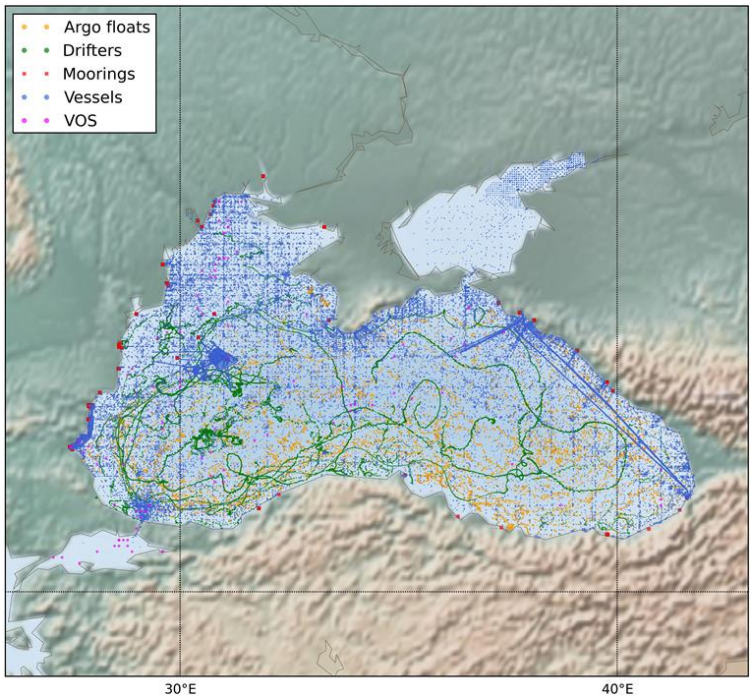


Figure 33. Black Sea in situ data sources in the past 30 years (1990–2018). Source: https://www.frontiersin.org/articles/10.3389/fmars.2019.00315/full

Other well-known Black Sea data sources are the PERSEUS database and the Black Sea Oceanographic Data Inventory (parts of these databases may already be included in EMODnet, but additional study is recommended to investigate data records missing from aggregators):

The **PERSEUS Oceanographic Mediterranean and Black Sea Data Management** (https://isramar.ocean.org.il/perseus_data/Default.aspx) is a Cast Data Base with vertical profiles of physical, chemical and biological data acquired with Bottle casts (Rosette), CTD casts, and Argo floats. Data can be downloaded in two forms ODV generic text format and MS ACCESS database. Data download is available after login for registered users only. Three data availability flags are used according to SeaDataNet Data Access Restriction Policies (L08) vocabulary: (1) unrestricted (data exported from public available databases MEDATLAS 2002; MATER; WODB0; CORIOLIS; ICES); (2) by negotiation (data can be obtained on a case-by-case basis through negotiation with data provider); and (3) organization (datasets are available to PERSEUS partners only).

[12] https://www.frontiersin.org/articles/10.3389/fmars.2019.00315/full

The **Black Sea Oceanographic Data Inventory** (http://sfp1.ims.metu.edu.tr/texts/inventory.htm) first created in the framework of the NATO Science for Stability TU-Black Sea Project provided by eight institutes from six countries: Ukraine; Turkey; Bulgaria; Romania; Russia; USA. New datasets were added to the Inventory for period 1996-2000 in the framework of the NATO Science for Peace ODBMS Project. The data Inventory is developed as a standalone Windows application. Apart from physics the inventory covers chemistry (24 variables) and biology (14 variables). The Inventory contains information since 1933 until 2000 (11 scientific institutions, 36 research vessels, 25275 stations, 51 variables).

Other projects

COSYNA

Within the Joint European Research Infrastructure of Coastal Observatories (JERICO-RI), one of the in situ monitoring infrastructures that is operated, is COSYNA. COSYNA is a Coastal Observing System for Northern and Arctic Seas. COSYNA's mission to operate an integrated observing system suitable for investigating the environmental state and variability of coastal areas. The observations encompass various measurements taken on-site from fixed stations, buoys, gliders, FerryBoxes, and research vessels. Observations are ongoing since 2010. Data is online instantly accessible (in some cases upon approval).

Monitored BGC variables on COSYNA are:

- Turbidity
- Chlorophyll-a
- CDOM
- Dissolved oxygen
- Dissolved nutrient
- pH
- CO2 in sea water

Observations are available through the Data Portal CODM http://codm.hzg.de/codm/ depicted in *Figure 34*. COSYNA is financed and coordinated by the Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research GmbH. The coastal platform of ACROSS will extend COSYNA to the land side (estuary) and to more offshore regions of the North Sea.

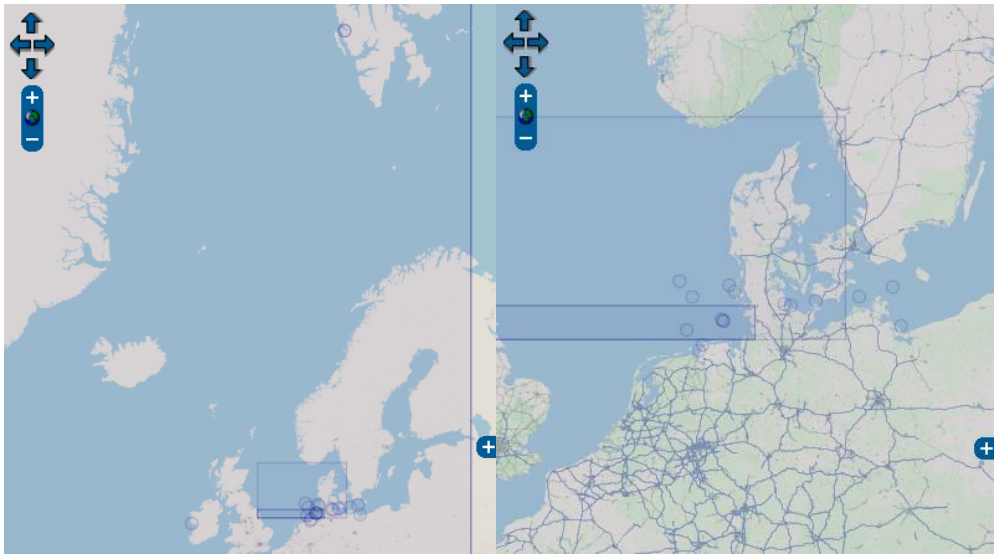


Figure 34. COSYNA data coverage in the CODM Data Portal (http://codm.hzg.de/codm/)

FERRYBOX

Within the JERICO-RI and also coordinated by the EuroGOOS Ferry Box Task Team, BGC data from FerryBox is also produced. Observations are ongoing and available since 2008. The FerryBox data is online accessible upon approval in two data bases the FerryBox Database (http://ferrydata.hzg.de/) (see *Figure 35*) and the Data portal CODM (http://codm.hzg.de/codm/). All FerryBox data from HZG and some other institutions are stored in the FerryBox database and can be accessed via internet. The data are initially checked automatically and stored in this database. All data are available for free and can be exported in ASCII (tab separated tables) or NetCDF format.

The BGC ferrybox data includes:

- turbidity,
- chlorophyll-a,
- dissolved oxygen,
- pCO2,
- pH
- total alkalinity.

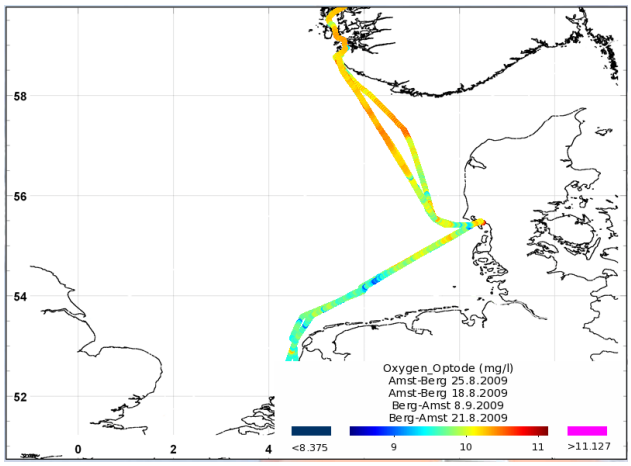


Figure 35. Example Ferrybox data retrieved from the FerryBox Database (http://ferrydata.hzg.de/)

Other BGC data providers

In addition to the above describe organizations, programmes and projects here we list additional BGC international and European data sources that are often used by the European marine research community:

- NOAA World Ocean Atlas 2018 (WOA2018)
  https://www.nodc.noaa.gov/OC5/woa18/woa18data.html
- NOAA World Ocean Data Base 2013
  https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html
- German Oceanographic Data Centre - GeoSeaPortal
  https://www.geoseaportal.de/
- PANGAEA Data Publisher
  http://www.pangaea.de/
- Surface Ocean CO_2 Atlas (SOCAT)
  https://www.socat.info/index.php/data-access/
- Global Ocean Data Analysis Project (GLODAP)
  https://www.glodap.info/index.php/data-access/
- Plymouth Marine Laboratory - Western Channel Observatory
  www.westernchannelobservatory.org.uk
- National Oceanographic Data Center (NODC) OGC
  https://nodc.inogs.it/data/access
- North-West European shelf programme (NOWESP) Research Data Base
  https://wiki.cen.uni-hamburg.de/ifm/ECOHAM/DATA_NOWESP
- NATO TU-Black Sea Data Inventory
  http://sfp1.ims.metu.edu.tr/TU-BlackSea/inventry/all.htm
- BATS (Bermuda Atlantic Time-Series) dataset
  http://bats.bios.edu/bats-data/
- Norwegian Marine Data Centre
  http://metadata.nmdc.no/UserInterface/#/
- Mediterranean Data Archaeology and Rescue (MEDAR) database
  http://medar.ieo.es/
- Rijkswaterstaat Waterinfo
  https://waterinfo.rws.nl/

Inventory of BGC in situ data used in CMEMS MFCs

Arguably, one of the main active users of BGC in situ data in the European marine landscape are currently the CMEMS Monitoring and Forecasting Systems (MFCs). They use BGC in situ data in near real time or in delayed mode for data assimilation and model calibration/validation for their hindcast and forecast products. The following section provides an overview on the data BGC in situ repositories that are currently used by the CMEMS MFCs.

GLO MFC

- **BGC-Argo floats datasets, for Total chlorophyll, nitrates, dissolved oxygen and pH**
- **BATS[13] (Bermuda Atlantic Time-Series) dataset, for Total chlorophyll and nitrates**

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-GLO-QUID-001-028.pdf

ARC MFC

Nutrients are validated against in situ observations from:

- **the Institute of Marine Research (IMR) accessible through the Norwegian Marine Data Centre: http://www.imr.no/forskning/forskningsdata/infrastruktur/viewdataset.html?dataset_id=104**
- **GLODAPv2_2019,**
- **ICES oceanographic database**
- **CLIVAR datasets**

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-ARC-QUID-002-005.pdf

BAL MFC

- **CTD data from monitoring stations and moorings (all parameters, except pCO2)**
- **CTD data from the ICES portal (https://www.ices.dk/data/data-portals/Pages/default.aspx) (all parameters, except pCO2)**
- **measurements from ferryboxes (dissolved oxygen, pCO2)**

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-BAL-QUID-003-007.pdf

NWS MFC

- **North Sea Biogeochemical Climatology (NSBC)[14] for nutrients and oxygen. The North Sea Biogeochemical Climatology is an integrated and spatially/temporally interpolated dataset based on:**
  - **NOWESP project[15]**
  - **World Ocean Database 2013[16]**
  - **Uwe Brockmann (Institute for Geology of the University of Hamburg, Department of Biogeochemistry and Marine Chemistry)**
  - **DOD (German Oceanographic Data Centre)**
  - **PANGAEA (PANGAEA Data Publisher for Earth & Environmental Science Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI), Bremerhaven, Germany)**
  - **ICES (International Council for Exploration of the Sea)**
  - **EMODNet (European Marine Observation and Data Network)**
- **SOCAT[17] gridded coastal fields for surface pCO2**

[13] Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., & Knap, A. H. (2001). Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. Deep Sea Research Part II: Topical Studies in Oceanography, 48(8-9), 1405-1447.

[14] Hinrichs,Iris; Gouretski,Viktor; Paetsch,Johannes; Emeis, Kay; Stammer, Detlef (2017). North Sea Biogeochemical Climatology (Version 1.1).

[15] Laane, R.W.P.M.; van Leussen, W.; Radach, G.; Berlamont, J.; Sündermann, J.; van Raaphorst, W. and Colijn, F. North-West European shelf programme (NOWESP): An Overview., 1996

[16] Boyer, T.P., J. I. Antonov, O. K. Baranova, C. Coleman, H. E. Garcia, A. Grodsky, D. R. Johnson, R. A. Locarnini, A. V. Mishonov, T.D. O'Brien, C.R. Paver, J.R. Reagan, D. Seidov, I. V. Smolyar, and M. M. Zweng, 2013: World Ocean Database 2013, NOAA Atlas NESDIS 72, S. Levitus, Ed., A. Mishonov, Technical Ed.; Silver Spring, MD, 209 pp., doi.org/10.7289/V5NZ85MT

[17] Bakker, D. C. E., et al. (2016) A multi-decade record of high quality fCO2 data in version 3 of the Surface Ocean CO2 Atlas (SOCAT). Earth System Science Data 8: 383-413. doi:10.5194/essd-8-383-2016.

- GLODAP[18,19,20] gridded field for surface pH
- Western Channel Observatory (www.westernchannelobservatory.org.uk) for surface chlorophyll and nitrate
- In situ measurements from the database of the International Council for the Exploration of the Sea (ICES) for all assessed variables except pCO2

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-NWS-QUID-004-011.pdf

IBI MFC

- World Ocean Atlas 2018 (WOA2018) (https://www.nodc.noaa.gov/OC5/woa18) for nutrients (Nitrate, phosphate, silicate, oxygen)
- SOCATv3 2020[5] for spCO2
- NOAA NCEI v2020[21] for spCO2
- GLODAP v2 2020[22] for various variables
- WOD update (https://www.nodc.noaa.gov/OC5/WOD/wod_updates.html) for various variables
- EMODNET (https://emodnet-chemistry.maris.nl/search) for various variables
- ICES (http://ices.dk/marine-data/dataset-collections/Pages/default.aspx) for various variables

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-IBI-QUID-005-003.pdf

MED MFC

- ISPRA dataset (for chlorophyll)
- OGS National Oceanographic Data Center (NODC-OGS) dataset (nitrate, phosphate, dissolved oxygen)
- Climatology from World Ocean Atlas 2013 historical dataset (nitrate, phosphate)
- Climatology from historical datasets (Alkalinity, DIC)

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-MED-QUID-006-008.pdf

BS MFC

The following datasets are used for the forecast product:

- In situ ARGO observations obtained from the Coriolis (Brest, France) website, see http://www.coriolis.eu.org/
- Chlorophyll provided from BGC Argo fluorescence using a correction adapted for the
- Black Sea (high content of CDOM, anoxia)

The datasets[23] used to validate the hindcast product are the following:

- Black Sea TU Ocean Base in the frame of the NATO TU-Black Sea project. This database contains data for 116 variables from 271 data sets, including 8,364,731 data values for 26,035 stations (http://sfp1.ims.metu.edu.tr/TU-BlackSea/inventry/all.htm).
- The MEDAR database, see http://medar.ieo.es/
- The PANGEA database, see http://www.wdc-mare.org
- The World Ocean data base (WOD)
- Black Sea Commission data (BSC, 382 data points).
- The R/V KNORR and R/V Endeavor cruise data (2001-2003-2005)

Source: https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-BS-QUID-007-005.pdf

[18] Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishi, M., Perez, F. F., and Suzuki, T. Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, NDP-P093. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2, 2015.

[19] Lauvset, S. K, R. M. Key, A. Olsen, S. van Heuven, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Pérez, T. Suzuki and S. Watelet. A new global interior ocean mapped climatology: the 1°x1° GLODAP version 2, Earth Syst. Sci. Data, 8, 325–340, 2016, doi:10.5194/essd-8-325-2016

[20] Olsen, A., R. M. Key, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Pérez and T. Suzuki. The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean, Earth Syst. Sci. Data, 8, 297–323, 2016, doi:10.5194/essd-8-297-2016

[21] Landschützer, Peter; Gruber, Nicolas; Bakker, Dorothee C. E. (2020). An observation-based global monthly gridded sea surface pCO2 product from 1982 onward and its monthly climatology (NCEI Accession 0160558). Version 5.5. NOAA National Centers for Environmental Information. Dataset. https://doi.org/10.7289/V5Z899N6.

[22] Olsen, A., Lange, N., Key, R. M., Tanhua, T., Bittig, H. C., Kozyr, A., Àlvarez, M., Azetsu-Scott, K., Becker, S., Brown, P. J., Carter, B. R., Cotrim da Cunha, L., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Jeansson, E., Jutterström, S., Landa, C. S., Lauvset, S. K., Michaelis, P., Murata, A., Pérez, F. F., Pfeil, B., Schirnick, C., teinfeldt, R., Suzuki, T., Tilbrook, B., Velo, A., Wanninkhof, R., and Woosley, R. J.: GLODAPv2.2020 – the second update of GLODAPv2, Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-165, in review, 2020.

[23] Joassin, P., 2011. Mathematical modeling of biogeochemical processes associated to a coccolithophorid (emiliania huxleyi) bloom - study of the seasonal and long-term variability of biogeochemical properties in the Black Sea using a data interpolating variational analysis (DIVA). Ph.D. thesis, Université de Liège.

Synthesis

After having reviewed a large number of existing BGC in situ data repositories we can conclude that there has been a long-standing effort among the marine biogeochemistry observing community to increase the BGC EOV data accessibility in standard quality-controlled and inter-comparable formats. However, BGC data availability is still lagging behind physics data availability, which hinders the validation and further improvement of (coastal) biogeochemical models. Moreover, BGC products are still predominantly chemistry focused, carbon-focused in many cases (e.g. SOCAT and GLODAP). Consequently, there is an urgent need to expand BGC data availability and data management practices in general, but especially beyond chemical variables towards biological elements (e.g. phytoplankton biomass or diversity). The first point is already being addressed with the BGC-Argo float system in the open ocean but for coastal areas, initiatives as European research infrastructures, or other national monitoring programmes can play an important role.

This report presents an inherently uncomplete, yet informative, inventory of existing European Biogeochemical in situ observations in transitional waters, coastal areas and open ocean (regional or global level). The listed ~50 data portals, which contain BGC in situ observations provide a good overview of existing BGC in situ observations in European waters. While the quantity of data portals suggests that existing European BGCs data is abundant, we must emphasize the significant overlap between repositories. The European BGC observation network relies on national and international programmes and research campaigns/projects that feed into national data portals. These are in most cases integrated and aggregated to regional observatories, such as the EuroGOOS ROOS, which then feed into the main European in situ data portals CMEMS, EMODnet, and SeaDataNet. Not even mentioning the downstream applications, which harvest their data from these aggregators and display it in various web-viewers. While the appearance of a single data source in many platforms is positive for the data accessibility, it does not increase further the data coverage. The coverage of BGC in situ data is still heterogenous with major gaps in transitional and coastal areas and in certain parts of regional seas, such as the Southern Mediterranean and the Arctic Ocean and Black Sea. Regional Sea Conventions help addressing these gaps by being the drivers, facilitators, and users of BGC in situ data collected by Contracting Parties (member states). While their data management practices are heterogenous, exceptions are OSPAR and HELCOM that commission their data management to ICES, they all provide quality controlled and standardized datasets. In general, the spatial and temporal coverage, update frequency, as well as the accessibility (restrictions) of Regional Sea Convention data could be improved in most cases, however, we should bear in mind that the purpose of these databases is to help their thematic assessment reports rather than serving Copernicus MFCs and other BGC in situ data users. It is recommended to conduct further assessment in order to check which data records of the Regional Sea Conventions are currently missing from the European data aggregators.

Another issue in the presented BGC in situ data inventory is the representation of coastal observations (compared to open ocean observations). Most data platforms in this report use a classification based on regions (e.g. European regional seas), thematic (e.g. chemistry or biology), and parameters. Only few of the portals allow the selection of coastal BGC in situ data, also noting that the scientific definition of “coastal zone” is still very much debated. Of course, exceptions are the research infrastructures, projects and monitoring programmes that are specifically targeting the coastal zone, such as DANUBIUS-RI, JERICO-RI, COSYNA, Coastal Coriolis or H2020 ODYSSEA. Therefore, we can only

assess the coastal BGC in situ data coverage of the portals visually. Looking at the coastal in situ observation availability of a commonly used BGC parameter such as chlorophyll-a in the EMODnet Chemistry or Physics portal, we can observe that major spatial gaps exist. Not all coastal zones of the European Regional Seas are monitored equally. The coastal zones of the southern North West Shelf region, or certain parts of the Northern Mediterranean region (e.g. Spanish, French, Italian coast) are well covered, while others have major coastal observation gaps, such as the Southern Mediterranean or the eastern Black Sea. An identification of coastal observation gaps and their prioritization should therefore be conducted, based on the ecosystem and socio-economic importance of the coastal zones. A good example of important coastal areas are river outlets such as the Po or the Danube estuary. Finally, it is also important to make sure that CMEMS MFCs make use of all the available coastal BGC in situ data. For instance, coastal zone BGC data have not been widely used for model validation in the CMEMS Baltic MFC, even though HELCOM makes it available through ICES. Consequently, not only more BGC observations in the coastal zone will be important for assessing model product quality in the coastal waters but also a better usage of existing observations.

Data availability was in most cases satisfactory, via online instant access, although there are several platforms where registration is necessary, or data is granted upon request via email. Even these minor data restrictions should be eliminated to further promote data sharing. The data management of BGC in situ data on European level seems to be well harmonized thanks to the EuroGOOS Task Teams, the CMEMS INSTAC, EMODnet, SeaDataNet, which maintain standard formats and quality control for the BGC EOVs. Timeliness of BGC data in data aggregator portals such as EMODnet or SeaDataNet should be improved as it is currently accessible in delay mode (with a delay of more than 6-12 months in some cases).

The attached BGC in situ data catalogue lists all data portals in this report with the corresponding links and a list of available BGC variables.

BGC in situ data catalogue

Data source	Link / Reference	BGC variables*
	DATA AGGREGATORS	
CMEMS In Situ TAC	http://www.marineinsitu.eu/dashboard/	Dissolved oxygen, Oxygen saturation, Dissolved inorganic carbon, Dissolved organic carbon, CO2 partial pressure, CO2 fugacity, Chlorophyll-a, CDOM, Turbidity, Nitrate Nitrite, Phosphate, Silicate, Ammonium, Dissolved nitrogen, Dissolved organic nitrogen, Total alkalinity, pH
EMODNET Chemistry	https://ec.oceanbrowser.net/emodnet/	Chlorophyll-a, Dissolved Inorganic Nitrogen, Dissolved oxygen, Phosphate, Silicate
EMODNET Physics	https://portal.emodnet-physics.eu/	Chlorophyll-a, Dissolved Inorganic Nitrogen, Dissolved oxygen, Phosphate, Silicate
EMODNET Biology	https://www.emodnet-biology.eu/portal/	Bulk chemistry (e.g. pH, TCO2), Dissolved gases, Nutrients, Other inorganic chemistry, Primary production, Phytoplankton biomass, chlorophyll-a
SeaDataNet	https://cdi.seadatanet.org/search	Alkalinity, acidity and pH of the water column, ammonium and ammonia concentration, Chlorophyll pigment concentrations, Concentration of organic matter, Nitrate concentration, Carbon, nitrogen and phosphorus, Carbonate system, Dissolved gases, Nutrients, Other biological measurements, Other inorganic chemical measurements, Other organic chemical measurements, Phytoplankton and microphytobenthos, and Pigments, and others
EuroGOOS ROOSes	Arctic ROOS http://webprod1.nodc.no/arctic-roos/arctic-roos.html Baltic ROOS, BOOS http://www.boos.org/boos-stations/ North-West Shelf ROOS, NOOS http://nwsportal.bsh.de/ Ireland-Biscay-Iberia ROOS, IBI-ROOS http://www.ibi-roos.eu/Access-to-data Mediterranean ROOS, MONGOOS http://www.mongoos.eu/data-center	Chlorophyll-a, Alkalinity, Hydrogen Sulphide, Total Nitrogen, pH, Total Phosphorus, Ammonium, Silicate, Nitrite, Nitrate, Phosphate, Salinity, Dissolved Oxygen, Temperature, and others
ICES portal	https://www.ices.dk/data/data-portals/Pages/default.aspx	BGC variables: Oxygen, Phosphate, Total Phosphorus, Silicate, Nitrate, Nitrite, Ammonium, Total Nitrogen, Hydrogen Sulphide, pH, Alkalinity Chlorophyll a, Secchi depths
	REGIONAL SEA CONVENTIONS	
OSPAR Convention	ODIMS https://odims.ospar.org/geoportal/ ICES Data Portal - OSPAR https://data.ices.dk/view-map?area=33 ICES DOME https://dome.ices.dk/browse/	Phytoplankton and zooplankton biomass, abundance and communities, chlorophyll-a, dissolved oxygen, Hydrocarbons, Major organic and inorganic constituents, Metals, Nutrients, Pesticides, herbicides, insecticides, Pharmaceuticals, PAHs, and others
HELCOM Convention	ICES – HELCOM https://ocean.ices.dk/Helcom/Helcom.aspx?Mode=1 HELCOM Map and Data Service https://maps.helcom.fi/website/mapservice/	Chlorophyll-a, Alkalinity, Hydrogen Sulphide, Total Nitrogen, pH, Total Phosphorus, Ammonium, Silicate, Nitrite, Nitrate, Phosphate, Salinity, Dissolved Oxygen, Temperature
Barcelona Convention	IMAP Pilot Info System http://imappilot.info-rac.org/app/#/ MEDPOL Info System	Habitat distribution, composition, communities, non-indigenous species, concentration of key nutrients in water column, Chlorophyll-a

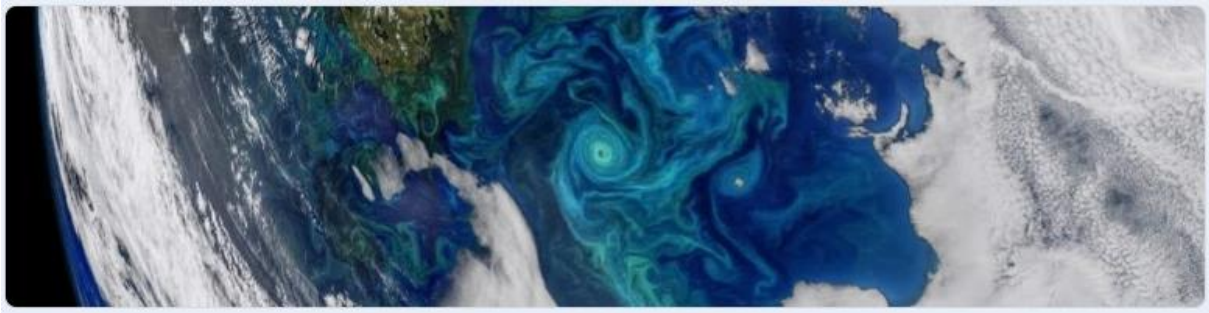
	http://www.info-rac.org/en/infomap-system/medpol-info-system EEA Mediterranean database https://www.eea.europa.eu/data-and-maps/data/meddb ENI SEIS II South https://eni-seis.eionet.europa.eu/south/areas-of-work/data-and-statistics	concentration in water column, Concentration of key harmful contaminants related to biota, sediment, seawater: BOD4, Cd, COD, Hg, Pb, PCB, PCT, PCDD, PCDF, Total N, Total P, VOC
Bucharest Convention	Regional Database on Pollution http://rdbp.sea.gov.ua/ Black Sea Database http://blackseadb.org/ PERSEUS Oceanographic Mediterranean and Black Sea Data Management https://isramar.ocean.org.il/perseus_data/Default.aspx Black Sea Oceanographic Data Inventory http://sfp1.ims.metu.edu.tr/texts/inventory.htm	Hydrochemistry, Nutrients, (Heavy) Metals, Pesticides, PCBs, PAHs, Phenols, Detergents, Petroleum hydrocarbons, Radionuclides, Photosynthetic pigments (e.g. chlorophyll-a), Phytoplankton, Zooplankton, Microbial communities, Macrozoobenthos, Meiobenthos, Macrophytobenthos, Microbial communities
RESEARCH INFRASTUCTURE		
JERICO-RI	https://www.jerico-ri.eu/data-access/	Alkalinity, acidity and pH, Chlorophyll, Dissolved oxygen, Electrical conductivity, Light absorption, suspended particulate material. Variable fluorescence parameters, Phytoplankton biomass and diversity, Zooplankton biomass and diversity, Fish abundance and distribution, Marine turtles, birds, mammals abundance and distribution, Hard coral cover and composition, Seagrass cover and composition • Macroalgal canopy cover and composition • Benthic invertebrate abundance and distribution
EuroARGO	https://dataselection.euro-argo.eu/	CDOM, pH, nitrate, PAR, Dissolved oxygen, CHLA
EMSO	https://data.emso.eu/	Chlorophyll-a, dissolved oxygen, turbidity
ICOS Carbon Portal	https://www.icos-cp.eu/observations/carbon-portal	Dissolved oxygen, Dissolved nutrient, Partial pressure of carbon dioxide (pCO2), Dissolved organic carbon (DOC), Dissolved inorganic carbon (DIC), pH, Total alkalinity
OTHER ORGANIZATIONS AND PROJECTS		
OceanSITES	http://tds0.ifremer.fr/thredds/CORIOLIS-OCEANSITES-GDAC-OBS/CORIOLIS-OCEANSITES-GDAC-OBS.html or ftp://ftp.ifremer.fr/ifremer/oceansites	biogeochemistry, and parameters relevant to the carbon cycle, ocean acidification, the ecosystem, and geophysics
Coriolis Data Center	http://www.coriolis.eu.org/Data-Products/Data-Delivery/Data-selection http://www.ifremer.fr/co-dataSelection/?theme=myocean http://www.ifremer.fr/co-argoFloats/ (ARGO data selection) http://www.ifremer.fr/co-dataSelection/?theme=gosud (GOSUD data selection) http://www.ifremer.fr/co-en/ (Eulerian networks and fixed buoys)	Chlorophyll-a, oxygen, CDOM, pH, nitrate, PAR, etc.
COSYNA Data	http://codm.hzg.de/codm/	Turbidity, Chlorophyll-a, CDOM, Dissolved oxygen, Dissolved nutrient, pH, CO_2 in sea water
FerryBox	www.ferrybox.org http://ferrydata.hzg.de/	Turbidity, chlorophyll-a, dissolved oxygen, pCO2, pH, total alkalinity.

BATS (Bermuda Atlantic Time-Series) dataset	http://bats.bios.edu/bats-data/	Dissolved oxygen, Florescence, total CO2, Alkalinity. Nitrate, nitrite, phosphate, silicate, Dissolved Organic Carbon, Dissolved Organic Nitrogen, Particulate Organic Carbon, Particulate Organic Nitrogen, Particulate silicate, Phytoplankton pigments, Chlorophyll-a, primary production, carbon flux, nitrogen flux
Norwegian Marine Data Centre	http://metadata.nmdc.no/UserInterface/#/	Various BGC variables, e.g. dissolved inorganic nutrients (nitrite nitrate phosphate, silicic acid), chlorophyll-a, oxygen, marine CO2 system (total alkalinity, total inorganic carbon), Carbonate, Phytoplankton abundance, Phytoplankton biomass, Zooplankton abundance, pH, turbidity, and others
Institute of Marine Research (IMR), Norway	http://www.imr.no/forskning/forskningsdata/infrastruktur/viewdataset.html?dataset_id=104	Nutrients, oxygen, chlorophyll-a
UH Integrated Climate Data Center - North Sea Biogeochemical Climatology (NSBC)	https://icdc.cen.uni-hamburg.de/las/getUI.do?dsid=id-015565262b&catid=1D89D87646D19EABC2F3AC9C1F8D3CAC&varid=ammonium_mean-id-015565262b&plot=XY_zoomable_image&view=xy&auto=true or https://icdc.cen.uni-hamburg.de/thredds/catalog/ftpthredds/nsbc/catalog.html or https://icdc.cen.uni-hamburg.de/thredds/aggregationNsbcCatalog.html	ammonium, chlorophyll-a, nitrate(+nitrite), phosphate, oxygen and silicate
NOWESP Research Data Base	https://wiki.cen.uni-hamburg.de/ifm/ECOHAM/DATA_NOWESP	phosphate, nitrate, nitrite, ammonium, silicate, suspended particulate matter and chlorophyll
World Ocean Data Base 2013	https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html	Oxygen, phosphate, silicate, nitrate and nitrite, pH, chlorophyll-a, plankton, alkalinity, partial pressure of carbon dioxide, Dissolved Inorganic Carbon
German Oceanographic Data Centre - GeoSeaPortal	https://www.geoseaportal.de/	Chlorophyll, oxygen saturation, turbidity,
PANGAEA Data Publisher	http://www.pangaea.de/	Various BGC variables from CTD, floats, profiling float, fluorometry, bottle, oxygen probe, etc.
SOCAT	https://www.socat.info/index.php/data-access/	Carbon system: xCO2, pCO2, fCO2
GLODAP	https://www.glodap.info/index.php/data-access/	oxygen, phosphate, nitrate, silicate, dissolved inorganic carbon, total alkalinity, pH, CFC-11, CFC-12, CFC-113 and CCl4
PML - Western Channel Observatory	www.westernchannelobservatory.org.uk	Fluorescence, turbidity, oxygen, CDOM, phytoplankton, pigments, nitrate, nitrite, phosphate, silicate and ammonium
World Ocean Atlas 2018 (WOA2018)	https://www.nodc.noaa.gov/OC5/woa18/woa18data.html	Dissolved oxygen, oxygen saturation, silicate, phosphate, nitrate
NODC OGC	https://nodc.inogs.it/data/access	Carbon, nitrogen and phosphorus, Carbonate system, Nutrients, Optical properties, Other biological measurements, Other inorganic chemical measurements, Other organic chemical measurements, Phytoplankton and microphytobenthos

NATO TU-Black Sea Data Inventory	http://sfp1.ims.metu.edu.tr/TU-BlackSea/inventry/all.htm	Fluorescence, light attenuation, CTD, Secchi depth, alkalinity, hydrogen sulphide, chlorophyll-a, nitrites, NO2, NO3, dissolved oxygen, pH, phosphates, Particulate Organic Carbon, Particulate Organic Nitrogen, Particulate Phosphates, Silicates, Total Organic Carbon, Phytoplankton, primary production
MEDAR database	http://medar.ieo.es/	biogeochemical profiles
Rijkswaterstaat Waterinfo	https://waterinfo.rws.nl/	Phytoplankton types and biomass, nutrients, ammonium, BOD, COD, chloride, phosphate, TP, nitrate, nitrite, TN, silicate, hydrogen carbonate, turbidity, dissolved oxygen, chlorophyll, primary production, and others

** the list of BGC variables may deviate from the actual offer of the data portals*

Appendix – Online survey questions



Section 1 of 3

Inventory of existing biogeochemical

Dear Madam, Sir,

This online form is in relation to the European Environmental Agency (EEA) framework contract 'Support to Copernicus In Situ Data Coordination' (COINS). Within this project we are collecting and reporting information on existing biogeochemical (BGC) observations, including data repository and availability.

We would like to achieve this by requesting regional or pan-European initiatives and networks to fill in this online form on existing biogeochemical (BGC) observations. The compiled inventory and synthesis of commonalities/gaps will help EEA to better formulate long term in-situ data coordination strategies, which is in the interest of these initiatives.

If you agree to the objectives of such inventory, please fill in this online form.

Thank you very much in advance for your cooperation.

Kind Regards,
Deltares, Danish Meteorological Institute (DMI), Institute of Marine Sciences (ISMAR) at the National Research Council (CNR), the Norwegian Institute for Water Research (NIVA), University of Stirling, and the Plymouth Marine Laboratory (PML)

After section 1 Continue to next section

Section 2 of 3

In-situ observations

Description of your existing biogeochemical observations (variables, geographical location, time period, data accessibility).

Biogeochemical variables *

Please tick all variables that are being measured

- [ ] Turbidity
- [ ] Secchi depth
- [ ] Chlorophyll-a
- [ ] Dissolved oxygen
- [ ] Dissolved nutrient
- [ ] Ammonium
- [ ] Partial pressure of carbon dioxide (pCO2)
- [ ] Pariculate matter
- [ ] Dissolved organic carbon (DOC)
- [ ] Dissolved inorganic carbon (DIC)
- [ ] Particulate organic carbon (POC)
- [ ] pH
- [ ] Phytoplankton diversity and biomass
- [ ] Zooplankton diversity and biomass
- [ ] Total alkalinity
- [ ] Other...

Geographical study area *

Please describe the geographical location of the study area

Short-answer text

Time period *

Pleas specify the start and end dates as precise as possbile but at least with years. If it is ongoing please use the word 'ongoing'. E.g. 2017-present

Short-answer text

Data repository *

Please specify the URL of the data repository

Short-answer text

Data accessibility *

Please specify how your data can be accessed

1. Online instantly accessible
2. Online accessible upon approval
3. Offline application
4. Other...

If other, please explain

Short-answer text

After section 2 Continue to next section

Section 3 of 3

Background information

Description (optional)

Email address

Your email address will only be used to forward a copy of your answers and the results. If you do not fill in your email address, then the survey is anonymous.

Short-answer text

Your institute

Please specify your institute

Short-answer text

Relevant Regional Sea *

The classification is according to the Regional Operational Oceanographic Systems of EuroGOOS https://eurogoos.eu/regional-operational-oceanographic-systems/

- [ ] Arctic
- [ ] Baltic
- [ ] North-West Shelf
- [ ] Ireland-Biscay-Iberia
- [ ] Mediterranean
- [ ] Other...

Project name

Please describe the project to which the in-situ observation gathering (if any)

Short-answer text

Funding source

Please describe the funding source of the in-situ observation gathering (if any)

Short-answer text

Coordinating organization *

Please specify which organization coordinates the observation gathering

Short-answer text