Project information	
Project full title	EuroSea: Improving and Integrating European Ocean Observing and Forecasting Systems for Sustainable use of the Oceans
Project acronym	EuroSea
Grant agreement number	862626
Project start date and duration	1 November 2019, 50 months
Project website	https://www.eurosea.eu

Deliverable information	
Deliverable number	D3.17
Deliverable title	**EuroSea Data Integration**
Description	The objective of this deliverable is to help users looking for in situ marine data or datasets, in their choice of data infrastructures (CMEMS, EMODnet and SeaDataNet) best suited to their needs. These three data infrastructures are explained and their cooperations are analysed. This document could also be a help for these data infrastructures to better understand user needs and adapt their integrators to this need.
Work Package number	WP3
Work Package title	Network integration and Improvements
Lead beneficiary	IFREMER
Lead authors	Dominique Obaton, Cécile Liétard
Contributors	Antonio Novellino, Alessandra Giorgetti
Due date	31.08.2023
Submission date	31.08.2023
Comments	

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 862626.

https://doi.org/10.3289/eurosea_d3.17

Table of contents

Executive summary

This D3.17 “Data Integration” deliverable has been written in complementarity of the deliverables D3.13 “Data handbook” and D3.7 “Networks harmonisation recommendations”. It has been primarily written with the aim to be useful for users, looking for in situ data or datasets, in their choice of data infrastructures (CMEMS - Copernicus Marine Environment Monitoring Service – EMODnet – European Marine Observation and Data network - and SeaDataNet) best suited to their needs.

To start, this deliverable provides a description of these three major European data integrators and explains how to access to the data and what type of data it is possible to find. The cooperation between these three data infrastructures is also presented.

A recommendation about what type of metadata should be attached to the measurement is also included in this deliverable. Its objective is to encourage data infrastructures to harmonize their metadata, which would allow data marine users to switch more easily from one infrastructure to another one and thus extend access to more data.

This deliverable also presents two case studies, in which we put ourselves in the place of a in situ marine data user.

1 Glossary

Acronym	Meaning
ADCP	Acoustic Doppler Current Profiler
AODN	Australia Ocean Data Network
ASCII	American Standard Code for Information Interchange
BGC	BioGeoChemical
BODC	British Oceanographic Data Centre
CC-BY	Creative Commons Attribution
CDI	Common Data Index
CMEMS	Copernicus Marine Environment Monitoring Service
CORA	COriolis Ocean Dataset for Reanalysis
CSR	Cruise Summary Report
CSV	Comma-Separated Values
CT	Conductivity and Temperature
CTD	Conductivity Temperature and Depth
DG DEFIS	Directorate General for DEFense Industry and Space
DG GROW	Directorate General for Internal Market, Industry, Entrepreneurship and SMEs
DG MARE	Directorate General for Maritime Affairs and Fisheries
DG RTD	Directorate General for Research and Innovation
DOI	Digital Object Identifier
DTM	Digital Terrain Model
EDIOS	European Directory of the initial Ocean-observing Systems
EDMED	European Directory of Marine Environmental Data
EDMERP	European Directory of Marine Environmental Research Projects

EDMO	European Digital Media Observatory
EMODnet	European Marine Observation and Data Network
ERDDAP	Environment Research Division Data Access Protocol
EU	European Union
EurOBIS	European Ocean Biodiversity Information System
FTP	File Transfert Protocol
GDAC	Global Data Assembly Centre
GML	Geography Markup Language
GPS	Global Positioning System
GTS	Global Telecommunication System
HCMR	Hellenic Centre for Marine Research
ICES	International Council for the Exploration of the Sea
ID	Identifier
IFREMER	Institut Français de Recherche pour l'Exploitation de la Mer
IMR	Institute of Marine Research
INSTAC	In Situ Thematic Centre
IOBAS	Bulgarian Academy of Science
IOC	Intergovernmental Oceanographic Commission
ISO	International Organization for Standardization
IMO	International Marine Organisation
JSON	JavaScript Object Notation
MFC	Monitoring and Forecasting Centres
MOU	Memorandum of Understanding
MOTU	Motu is a high efficient and robust Web Server
NERC	Natural Environment Research Council
NETCDF	NETwork Common Data Form
NOAA	National Oceanic and Atmospheric Administration
NODCs	National Oceanographic Data Centres
NRT	Near Real Time
NVS	NERC Vocabulary Server
ODV	Ocean Data View
OGC	Open Geospatial Consortium
OPeNDAP	OPen source project for a Network Data Access Protocol
PAR	Photosynthetically Active Radiation
PIDoc	Product Information Document
QA	Quality Assurance
QC	Quality Control
QGIS	Free and Open Source Geographic Information System
RD	Radar
SBE	Sea Bird Sensor
SDN	SeaDataNet
SEANOE	SEA scieNtific Open data Edition
SMHI	Swedish Meteorological and Hydrological Institute
TAC	Thematic Centre
Tiff	Tag Image File Format
TG	Tide Gauge

UK	United Kingdom
VRE	Virtual Research Environment
WGS	World Geodetic System
WOD	World Ocean Database
WORMs	World Register of Marine Species
WMO	World Meteorological Organization
WMS	Web Map Service
XLS	Excel format

2 Introduction

Observations of the oceans have been carried out for decades by scientific organisations, during marine cruises or via fixed platforms at sea. This data has long remained the property of scientific organisations and it took several years before being able to have access and sometimes even only through scientific publications. In recent years, these organisations have worked to make this data accessible to as many users as possible.

In addition, some projects have been set up to create data integrators to make the link between producers and users. This is where SeaDataNet, EMODnet and the CMEMS (Copernicus Marine Environment Monitoring Service) were born. Indeed, the common objective of these three integrators is to make data accessible. The challenge today is to make this marine observation data as accessible as possible so that they can be inserted into various studies including numerical forecast models.

This deliverable D3.17 has been primarily written with the aim to be useful for users looking for in situ marine data and datasets, in their choice of data infrastructures best suited to their needs. We will mainly focus on in situ chemical, physical, and biogeochemical data.

This deliverable provides a description of the three major European data infrastructures. It explains how to access to the data and what type of data it is possible to find. We also discuss what type of metadata should be attached to the measurement made. Its objective is to encourage data infrastructures to harmonize their metadata, which would allow data marine users to switch more easily from one infrastructure to another one and thus extend access to more data.

This deliverable also presents two case studies, in which we put ourselves in the shoes of a user looking for in situ marine data. These case studies provide a better understanding of how these integrators work, their advantages and disadvantages. It will also allow infrastructures to identify points for improvement on their interface.

3 European Data integrators

3.1 Copernicus Marine Service

3.1.1 Objective and description of CMEMS

The Copernicus Marine Service (or Copernicus Marine Environment Monitoring Service (CMEMS)[1]) is funded by the EU Commission DG DEFIS and led by Mercator Ocean. It is one of the 6 operational services of Copernicus. The major objective of the CMEMS infrastructure has been to make, freely available, on a regular basis marine model, satellite and in situ data in the most homogeneous possible way through a catalogue and a common access. The access to data is done by regions and sub regions that formed products covering the European seas and the global ocean.

The CMEMS Service encompasses two kinds of production centres:

- Monitoring and Forecasting Centres (MFCs) charged to maintain and produce operational numerical models of the ocean.
- Thematic Assembly Centres (TAC) tasked with the collection of past and near real time (made available in few hours) ocean data, both in situ (water column) and satellite observations.

The **In Situ TAC**[2] is the component of the Copernicus Marine Service which ensures a consistent and reliable access to a range of in situ data for the purpose of assimilation and validation of ocean models, among others. Due to its characteristics (discrete data), it is difficult to access to in situ data by regions through the CMEMS website. However, it is possible to access to them, by platform, through the In Situ TAC website.

3.1.2 Where does the data come from?

In Situ observations are on-site local measurements of sea water properties. They are obtained using on-site sensors on board of a wide range of platforms from autonomous observatories at sea (e.g. floats, drifting buoys, gliders, moorings, etc.) and ships of opportunity or research ships. They allow to capture ocean variability at different scales.

Fifteen institutes have joined their expertise to provide an in situ service compliant with the Copernicus Marine Service requirements. These institutes are responsible for listing the data while relying, at European level on the principles of EuroGOOS and at international level on the principles of the IOC. The In Situ TAC is coordinated by IFREMER (France). The In Situ TAC is also constituted by seven Regional in situ data production units:

- Puertos Del Estado (Spain) for the Iberia-Biscay-Ireland region,
- HCMR (Greece) for the Mediterranean Sea,
- IOBAS (Bulgaria) for the Black Sea,
- IMR (Norway) for the Arctic,
- SMHI (Sweden) for the Baltic Sea,
- BSH (Germany) for the North West Shelves region,
- IFREMER (France) for the global ocean.

[1] https://marine.copernicus.eu/
[2] http://www.marineinsitu.eu/

In addition to these regional leaders, the In Situ TAC counts several products experts and other institutions for cross cutting activities to answer Mercator Ocean requests useful for the Copernicus marine service.

3.1.3 Access to Copernicus Marine Service and tools

The service is continuously monitored and availability, issues and changes in the products are displayed on the website. In addition, a manned service desk answers to any users requests through chat or mails.

Moreover, CMEMS proposes user **learning services** with training and workshops on and offline to help fully harness its services[3]. CMEMS allows also an access to any **resources**: media, documentation, as well as the CMEMS tool Jupyter Notebooks developed by the Copernicus Marine Service experts[4].

3.1.4 How to access the data and what type of data?

To access, visualise and download marine data from the CMEMS catalogue or from In Situ Ocean TAC Dashboard, it is imperative to first register as a Copernicus user[5], it is free and very fast. After completing the form, an email is sent to the user so that he can define a password.

On the **In Situ TAC "dashboard"[6],** you can explore, visualise and download the multi-source, multi-platform and heterogeneous data collected in near-real-time (within 24 hours) and delivered to the Copernicus Marine Service.

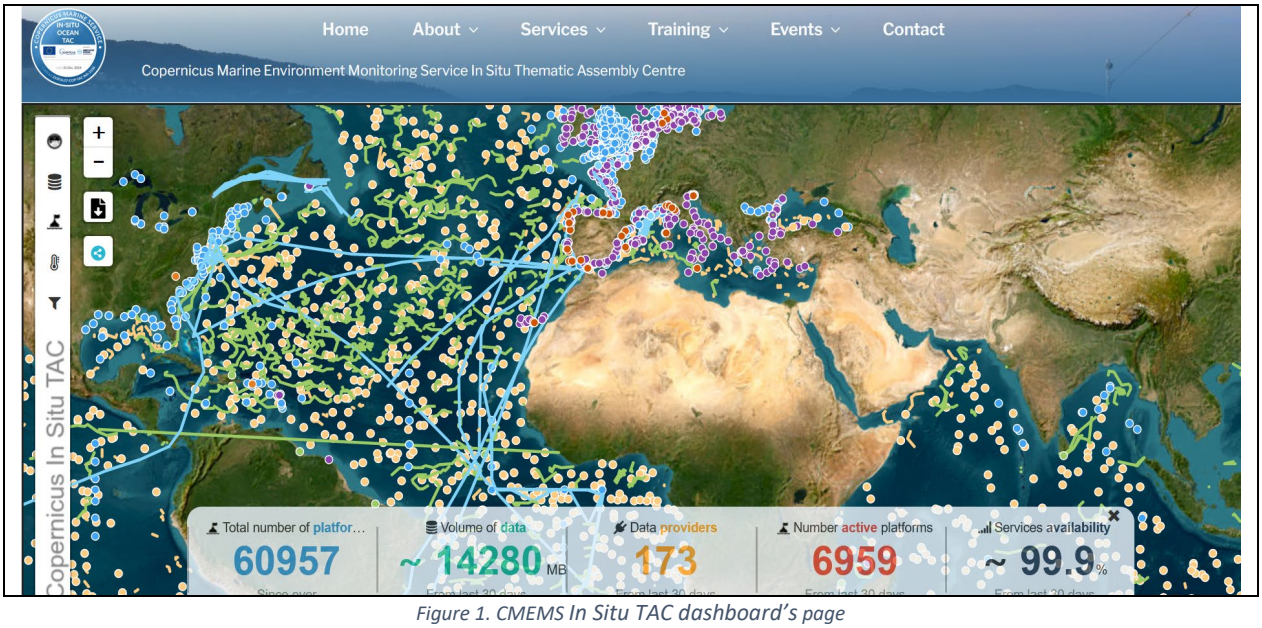


Figure 1. CMEMS In Situ TAC dashboard's page

[3] https://marine.copernicus.eu/services/user-learning-services
[4] https://marine.copernicus.eu/services/user-learning-services
[5] https://resources.marine.copernicus.eu/?option=com_sla
[6] http://www.marineinsitu.eu/dashboard/

The In Situ TAC “dashboard” provides a data access by platform. This interface allows you to realise a multi criteria search:

- By coverage: area and period,
- By platform: type and status,
- By data: parameters and data type,
- By « other »: provider, platform code and depth.

Click on the selected platform to open its associated sheet and to see the platform’s metadata and have access to data. For each platform we find the following metadata:

- Availability date,
- Platform type,
- Number of platform/n°ID,
- List of parameters proposed,
- Source Institute with EDMO code,
- Coordinates (latitude/longitude),
- Last observation date…

On each sheet, you have two possibilities: to download data in NetCDF files or to visualise data on graphs.

The CMEMS also allows access to an **Ocean product catalogue**, where it is possible to download and/or visualise data by regions or sub regions (for gridded data). Products are accessible and downloadable via the Copernicus Marine data store[7]. Products within the CMEMS catalogue are organised by types (modelling, in situ, satellite), regions (global, Arctic, Baltic, EU North East Atlantic, EU South East Atlantic, Mediterranean and Black seas) and by variables (blue: physical, white: ice and green: biogeochemical). This is done for 2 blocks of products: the near real time ones and the multi years ones. All products are either at measurements points (in situ), along track (satellite) or gridded (model, satellite and few in situ). They are generic products made for any kind of applications. In order to look for in situ products only, please use the “INSITU” keyword on the catalogue input box.

The CMEMS also proposes Ocean Monitoring Indicators (OMIs). OMIs are free downloadable trends and data sets covering the past quarter of a century. These are key variables used to track the vital health signs of the ocean and changes in line with climate change[8].

[7] https://data.marine.copernicus.eu/products
[8] https://marine.copernicus.eu/access-data/ocean-monitoring-indicators

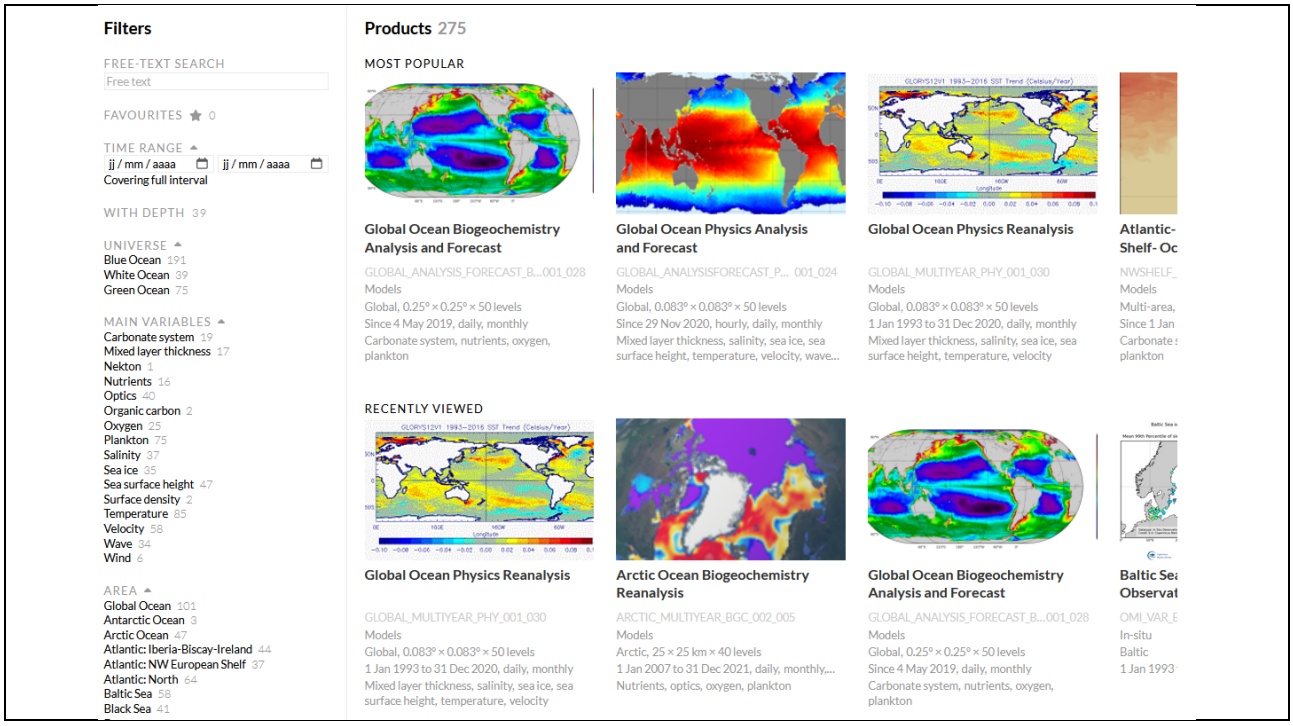


Figure 2. CMEMS catalogue to browse and access to the products

Product's pages (or product landing pages) contain:

- A description of the product, with an overview and a map of the data,
- Data access, download and mapping services,
    - Datasets with multiple protocol download choices (MOTU, OPeNDAP, ERDDAP, FTP and WMS) –only FTP is available for in situ products (non gridded),
    - Metadata,
- A contact access,
- An access to the documentations:
    - User manual (one per product),
    - Quality information document (one per product),
    - Licence,
    - How to cite,
    - Reference: DOI (attached to product).

To access the data, click on "Data access". Several ways to download data are proposed:

- To subset data in time and/or space, choose MOTU,
- If you use an OpenDAP client such as NetCDF4/xarray (Python), ferret, or MATLAB, choose OpenDAP or ERDDAP,
- To download raw files, choose FTP,
- To request maps from QGIS or similar tools, use the Web mapping Service (WMS).

For each CMEMS proposed protocol, a help is offered, in the form of a tutorial.

3.2 EMODnet

3.2.1 Objective and description of EMODnet

The European Marine Observation and Data network (EMODnet) is a long-term, marine-data initiative funded by the European Maritime and Fisheries Fund (managed by EU DG MARE). EMODnet nowadays connects a network of over 120 organizations, supported by the EU's Integrated Maritime Policy, who work together to observe the sea, process the data according to international standards and make that information freely available as interoperable data layers and in particular as pan-European data products.

EMODnet provides free access to marine data, metadata, and data products and services spanning seven broad disciplinary themes **or portals**: bathymetry, geology, physics, chemistry, biology, seabed habitats, and human activities**:**

- Bathymetry: to provide the best available harmonised Digital Terrain Model (DTM) for the European sea regions,
- Biology: to provide temporal and spatial distribution of marine species and species traits from European regional seas,
- Chemistry: to provide standardised harmonized validated data collections and reliable marine chemical data and products for all European marine regions,
- Geology: to bring together harmonised offshore data and consolidate the existing data products with higher resolution and more contents,
- Human activities: to make available information on the geographical position, spatial extent and attributes of a wide array of marine and maritime human activities throughout Europe
- Physics: to provide a single point of access to in situ ocean physics data, data products and metadata built with common standards,
- Seabed Habitats: to provide access to seabed habitat data in Europe.

EMODnet turns marine data into maps, digital terrain models, time series and statistics, dynamic plots, map viewers, and other applications ready to support researchers, industries, and policymakers when tackling large societal challenges. In addition to the central portal and thematic portals and services, there is an EMODnet Ingestion portal, aiming at the uptake of data submissions from third parties that are not yet connected to the major European data infrastructures.

The program is developed towards seven discipline-based themes or portals:

- Bathymetry: to provide the best available harmonised Digital Terrain Model (DTM) for the European sea regions,
- Biology: to provide temporal and spatial distribution of marine species and species traits from European regional seas,
- Chemistry: to provide standardised harmonized validated data collections and reliable marine chemical data and products for all European marine regions,
- Geology: to bring together harmonised offshore data and consolidate the existing data products with higher resolution and more contents,
- Human activities: to make available information on the geographical position, spatial extent and attributes of a wide array of marine and maritime human activities throughout Europe
- Physics: to provide a single point of access to in situ ocean physics data, data products and metadata built with common standards,

- Seabed Habitats: to provide access to seabed habitat data in Europe.

3.2.2 Where does the data come from?

EMODnet is made up of over the organizations involved within the networks that bring together marine data, products and metadata. As anticipated the program is developed by thematic-projects and each project involve a network of institutes to collect, harmonize, qualify, process the data and make products available in the EMODnet central portal.

As an example, EMODnet Physics is coordinated by ETT and has a core team SMHI, IFREMER, ICES, MARIS and ETT and an extended team including OCEANOPS, OGS, CMCC, CNR ISMAR, CISC BEC, CoLAB+, AZTI, CTN, VLIZ, and experts from ISPRA, INGV, UNIGE etc.

EMODnet Chemistry is coordinated by OGS and is organized in consortium partners and subcontractors

All the EMODnet projects connect and use in situ data, anyhow the data processing and data validation process naturally differs between disciplines.

3.2.3 Access to EMODnet services and tools

EMODnet proposes a common **Map viewer**[9] which offers the users easy access to discover, visualise and download marine data across the seven thematic portals from one place, powered by a searchable unified catalogue. Data and products are also consumable by OGC compliant services (WMS, WFS) and are INSPIRE compliant.

EMODnet also proposes a collection of **tools, applications and software**[10] to support marine data users. Best practices, information and general guidelines on different data types, found on EMODnet are also available.

In addition, the **EMODnet Data Ingestion portal** was created to allow marine data managers to easily integrate and publish their data as open data. EMODnet Ingestion and safe-keeping of marine data is a trans-thematic platform that seeks to identify and reach out to organisations from research, public, and private sectors who are holding marine datasets and who are not yet connected and contributing to the existing marine data management infrastructures which are driving EMODnet. Those potential data providers should be motivated and supported to release their datasets for safekeeping and subsequent free distribution and publication through EMODnet.

3.2.4 How to access the data and what type of data?

No need to register to access marine data via EMODnet's Map viewer portal.

[9] https://emodnet.ec.europa.eu/geoviewer/
[10] https://emodnet.ec.europa.eu/en/tools-guidelines

EMODnet proposes several levels of processing for the data (L0 to L4) and data product (L5 and L6)[11]:

- L0: Raw data,
- L1: Full resolution data reconstructed with calibration coefficients, geo- and time-referenced,
- L2: Geo- and time-referenced processed (derived) data with a minimum QC,
- L3: Delayed mode data with further QC,
- L4: Collated data from different measurements, samples and/or sources that have been integrated in a data system by means of standardisation and/or categorisation, and subset or otherwise selected or derived to fulfil a specific requirement
- L5: Model or analysis output that uses data of level 2 and/or 3 as input,
- L6: Derived information from multi-variable model or analysis that has Level 5 data products and/or Level 2-3 data as input.

EMODnet is not a raw data provider. Indeed, EMODnet mainly offers data with Level L4 and products with Levels L5 and L6.

The main portal provides access to all EMODnet data via a Map viewer and a general catalogue of products.

- EMODnet Map viewer,
- EMODnet data product catalogue,

EMODnet Map viewer[12] is a cartographical tool that allows to select data from the seven discipline-based themes directly on a map. Each data layer selected can be downloaded independently from different formats: .JSON, .CVS, shapefile, .GML.

[11] file:///C:/Users/clietard/Downloads/EMODnet_2021_Portfolio_FINAL-1.pdf
[12] https://emodnet.ec.europa.eu/geoviewer/

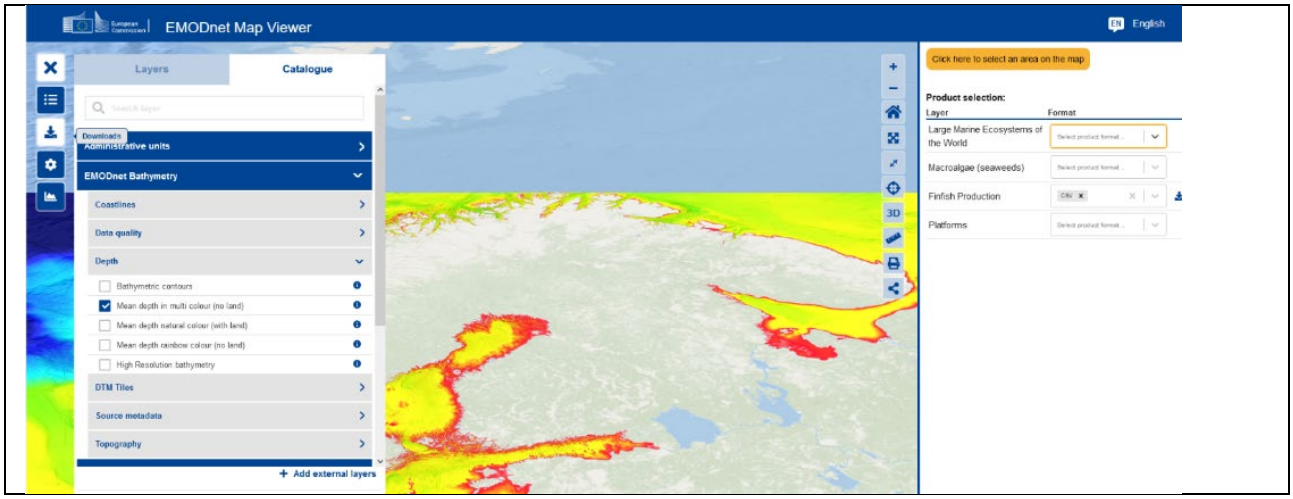


Figure 2. EMODnet Map viewer

Ocean physics data from in situ platforms is offered under the EMODnet Physics section.

EMODnet also offers a **data product catalogue**[13]. The products are created from a collection of aggregated and/or harmonized data and made available in the form of maps. The data product catalogue interface allows to refine searches by criteria, among other:

- Type of resources (dataset, series, ERDDAP, service…),
- Available actions (viewable or downloadable),
- Topics (ocean, environment, economy…),
- Keywords (ocean, North Sea…),
- Contact for the data (name or organization),
- Years,
- Resolution or scale (1 m, 3000 m…),
- Formats (.OGC, .GML, .TIFF…),
- Representation types (vector, grid, raster)…

13 https://emodnet.ec.europa.eu/geonetwork/srv/eng/catalog.search#/search?resultType=details&sortBy=sortDate&from=1&to=20

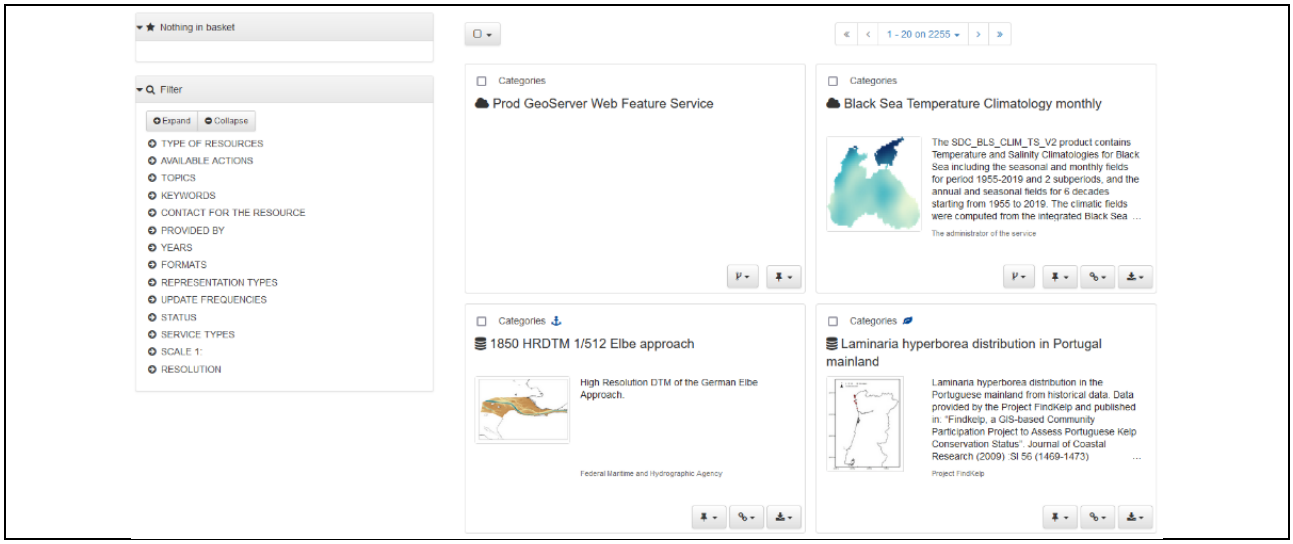


Figure 3. EMODnet product catalogue

The catalogue provides the users with:

- A description of the product,
- Period to which the product refers,
- The reference system information,
- Point of contact for the various contributors,
- DOI,
- Legal constraints,
- Map.

Direct download may be possible or not according the theme and the dataset product.

3.3 SeaDataNet

3.3.1 Objective and description of SeaDataNet

SeaDataNet[14] is a pan-European infrastructure, funded by the EU Commission, DG RTD. The major objective of this infrastructure has been to gather and standardise national in situ marine research data (mainly cruise data) from all European countries before making them freely available in a homogeneous way to users.

To achieve this goal, the SeaDataNet infrastructure created a single virtual data centre, the SeaDataNet portal, allowing a user looking for in situ physical and/or chemical data, metadata and products, to connect to any national marine in situ data from European countries through a single user interface. To be visible on the SeaDataNet portal, the data have to be in the "Seadatanet" format. Tools are made available to producers to help them standardise their data (standardized metadata created for the occasion and tools to implement them) and make them available. This portal was set up in cooperation with National Oceanographic Data Centres (NODCs) and data focal points from 35 countries bordering the European seas.

[14] https://www.seadatanet.org/

Over 115 marine data centres (NODCs) are today engaged within SeaDataNet and provide discovery and access to in situ data for all European researchers. The portal gives access to the tools and reference codes database (metadata) that have been set up and used to homogenise the data provided to users. SeaDataNet only provides delayed mode data. It ensures the data consistency within a data set and the quality and errors of the data are clearly stated to the user. The access to the data is done by platform.

3.3.2 Where does the data come from?

At present the SeaDataNet CDI service has more than 115 connected data centres, and provides metadata and access to more than 2.3 Million data sets, originating from more than 650 organisations in Europe, covering physical, geological, chemical, biological and geophysical data, and acquired in European waters and global oceans.

3.3.3 Access to SeaDataNet tools

In order to render interoperable node platforms constituting the SeaDataNet data centres, SeaDataNet has developed and implemented standards:

- Common vocabulary for metadata[15],
- Common protocol for data control and metadata[16],
- Common file formats[17].

It was also necessary to define common catalogues, rules for making data available and developing common software, made available to all producers wishing to publish their data in the SeaDataNet catalogue[18].

The section "**Look-up vocabularies**" gives a direct access to the BODC webservices by giving access to the NERC Vocabulary Server (NSV).

The NSV gives data managers the means to access lists of controlled terms to describe data, thus saving the time and costs associated with unraveling the meaning of a given data set.

In order to provide users common means for analysing and presenting data and data products, SeaDataNet has designed an overall system architecture, and developing common software tools for data centres and users (in the section "**Download software**"). Among these tools are:

- SeaDataNet file format convertor,
- Analysing and visualising of data sets,
- Interpolation and variational analysis of data sets,
- Tool for the generation of spatial objects from vessel navigation during observations…

The section "**Access metadata catalogues**" allows to access or contribute to the SeaDataNet metadata catalogues of Marine organisations (EDMO), datasets (EDMED), projects (EDMERP), observing systems

[15] https://www.seadatanet.org/Standards/Common-Vocabularies
[16] https://www.seadatanet.org/content/download/596/file/SeaDataNet_QC_procedures_V2_%28May_2010%29.pdf
[17] https://www.seadatanet.org/Standards/Data-Transport-Formats
[18] https://www.seadatanet.org/Software

(EDIOS), Research cruises (CSR) and data description (CDI). Each of the directories has its own origin, however as part of SeaDataNet all directories have been harmonised in use of syntax, semantics and tools.

The SeaDataCloud "**Virtual Research Environment**" (VRE)[19] was a pilot application which offers a collaborative environment to perform data-driven research in the cloud. The SeaDataCloud VRE is maintained but is no longer developed. A Marine-ID[20] is necessary to login to test the beta version.

SeaDataNet has established a brokering service ("**Discover international data**") with a user-friendly search interface to discover marine dataset collections which are managed at major international marine data portals (Australia Ocean Data Network (AODN), NOAA National Centers for Environmental Information and NOAA World Ocean Database (WOD).

SeaDataNet uses the SEANOE (SEA scieNtific Open data Edition) service to facilitate scientists to publish their research data in the field of marine sciences as citable resources (**Publish your marine data**). Using SEANOE, SeaDataNet enables to get a DOI (Digital Object Identifier) on data making them citable resources.

3.3.4 How to access the data and what type of data?

The SeaDataNet portal is achieved by developing, implementing and operating the Common Data Index (CDI) service that gives users a highly detailed insight in the availability and geographical spreading of marine data across the different data centres across Europe. The CDI provides an ISO19115 - ISO19139 based index (metadata base) to individual data sets (such as samples, timeseries, profiles, trajectories, etc).

SeaDataNet proposes two types of data access: "**Search data**" and "**Access products**".

"**Search data**": allows to search and download data from SeaDataNet CDI catalogue, with user own criteria. It is possible to access the platform while being logged in with a Marine-ID or not. In this latter case, access to certain options, such as downloading data will not be available.

The data search interface allows to realise a search with a large choice of criteria, among other:

- Keywords,
- GPS coordinates,
- Sea regions,
- Date,
- Parameters,
- Disciplines,
- Platform type,
- Instrument type,
- Water depth,
- Cruise or station name…

Once the datasets have been identified and added in the dataset basket, a request is sent to the CDI service. Acceptance takes a few minutes then files are available via the CDI user interface. A first email is sent to the

[19] https://vre.seadatanet.org/marine_id_login
[20] https://www.marine-id.org/

user to confirm the request and a second to notify him of the availability of the data. According to data type, files format received can be CFPOINT (= NetCDF), ODV, MedAtlas or .Tiff.

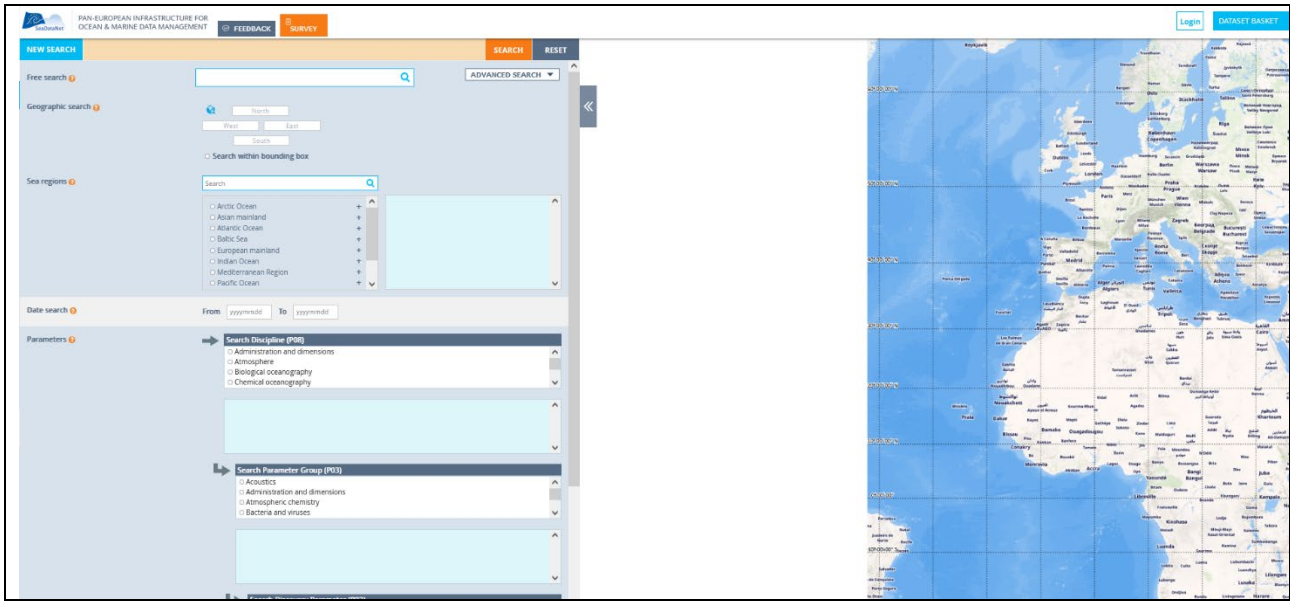


Figure 6. SeaDataNet search data portal

"**Access products**": allows to download the SeaDataNet aggregated datasets and climatologies based on the aggregated datasets and data from external data sources such as the COriolis Ocean Dataset for Reanalysis (CORA) and the World Ocean Database (WOD) for all the European sea basins (Mediterranean Sea, Baltic Sea, North Atlantic Ocean, Arctic Ocean, Black Sea, North Sea) and the Global Ocean. Each product is described in a Product Information Document (PIDoc) that can be accessed from the product's landing page. Downloaded files are NetCDF files.

Click on the User Manual tab on the top-right corner of the page for instructions.

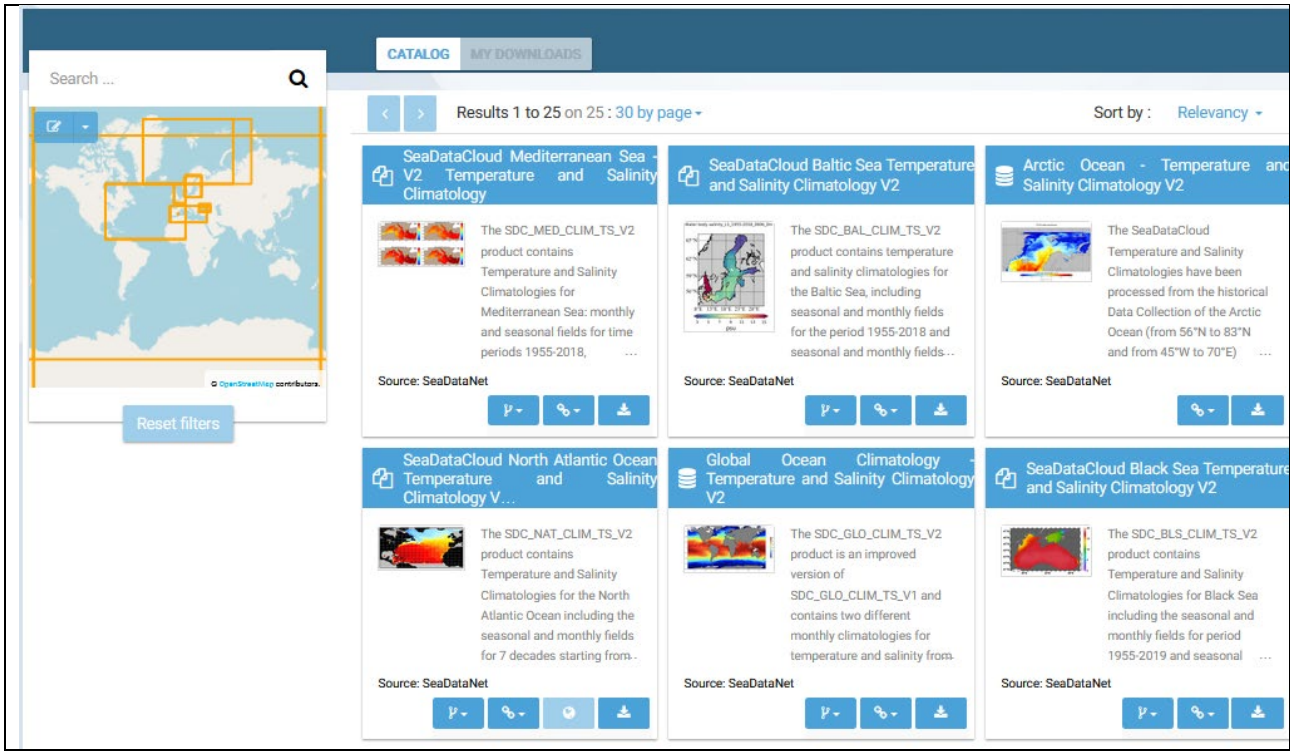


Figure 7. SeaDataNet product catalogue

4 Links between these European integrators

These three integrators all have the same objective of distributing open access marine data and data products free of charge. Their goal is also to provide the best possible quality. They are, however, somewhat different.

The Copernicus Marine Service focuses mainly on the fields of physics and biogeochemistry. In addition to in situ data, the catalogue proposes satellite and model products on the global and EU oceans and seas. These generic products are by regions or sub regions for any kind of dada type and not by platforms (for the in situ) on the central portal.

The EMODnet integrator takes into account any marine in situ fields (physics, biology, chemistry, geology, bathymetry, seabed habitat and human activities). Its central portal gives access to in situ marine data and bespoke products (gridded) for the Marine Strategy Framework Directive among others.

The SeaDataNet integrator also offers a large span of marine fields: physics, biology, chemistry, geology and bathymetry mainly coming from the EU NODCs. SeaDataNet targets both providers and users: the providers with references and tools for them to harmonise their data before uploading it on SeaDatNet, the users to access EU national data in a homogeneous and harmonised way.

4.1 Cooperation between EMODnet and the CMEMS

The Copernicus Marine Service and EMODnet are collaborating at coordination and operational levels for specific thematic areas. In 2016, DG DEFIS (formerly DG GROW) and DG MARE signed a Memorandum of Understanding (MoU) to closely collaborate about their respective marine activities developed under the Copernicus Marine Service and EMODnet Physics. The main aim is to emphasize the complementarity of their respective marine data offers (available under marine.copernicus.eu and emodnet.eu portals resp.) and explore synergies for stronger collaboration for the benefit of the marine user community.

In 2019, another MoU was set up between the Copernicus Marine Service and EMODnet Chemistry.

4.2 Cooperation between EMODnet and SeaDataNet

In 2008, the SeaDataNet members have been closely involved in some discipline-based themes of EMODnet planning and implementation. SeaDataNet infrastructure and standards have been adopted for developing and supporting the EMODnet portals for physics, chemistry and bathymetry, while also contributing to biology and geology.

The EMODnet portals make full use of the SeaDataNet infrastructure for managing additional thematic measurement data as gathered from data partners, connecting new data providers to the SeaDataNet data discovery and access service, and for harvesting large data collections from the SeaDataNet infrastructure which serve as input for the production and publication of thematic data products with a European coverage.

Physics uses the SDN P09, P02, P01 vocabularies for parameters, P06 for units, EDMO for institutions, and re-publish the Temperature and Salinity climatology that are developed and maintained under SeaDataNet. EMODnet Physics and SeaDatanet works together to reduce gaps between operational data and research quality validated datasets.

For the BGC data, the quality controls found on the SeaDataNet integrator are carried out by EMODnet Chemistry. In addition, SeaDataNet is an active network of data centres in EMODnet Data Ingestion, charged with seeking elaboration of ingested data sets to common standards and uptake of these resulting data sets in the SeaDataNet infrastructure[21].

4.3 Cooperation between CMEMS and SeaDataNet

A MoU between the CMEMS and SeaDataNet was signed in 2018 with the aim of setting up a cooperation in strategy, research, and operation, where mutually beneficial and desirable. Seven cooperation areas have been identified:

- Development, maintenance, adoption and promotion of standards,
- Development, adoption and promotion of QA-QC methodologies,
- Metadata and data exchange from SeaDataNet to CMEMS INSTAC,
- Metadata and data exchange from CMEMS to SeaDataNet,
- Development of products,

[21] https://newsletter.seadatanet.org/report/13#157

- Distribution of products,
- Teaming up in relevant projects.

5 The metadata in these European integrators

We will begin this chapter with an analysis of the deliverable D3.7 which addresses the metadata essential to understanding the marine data.

We will also analyze what metadata is offered by the networks but also by the three European integrators compared to what is offered in D3.7.

5.1 Analysis of the Deliverable 3.7: the metadata

After discussion with the experts of the different networks through the EuroSea project, it appears that certain metadata are mandatory to be able to use the data by user. Likewise, if they were filled in, or in a particular format, would allow optimal use of the data by users. Indeed, the use of standardised metadata will allow an easier access to data:

- The platform identification unique number,
- The date/time in ISO format and UTC: YYYY-MM-DDThh:mm:ssZ,
- The geographical position in WGS84,
- The depth of data
- The parameter measured, parameter definition CF standard name,
- The technical characteristics of instrument,
- The unit of the parameters.
- The quality control flag,
- Data in Netcdf CF format,
- Political access: license CC-BY 4.0,
- Dataset citation.

For most of ocean observing networks considered in the EuroSea project, ocean platforms are identified using the **WMO** (World Meteorological Organization) **number** reporting on the Global Telecommunication System (GTS). Considered platform types are profiling floats (Argo), autonomous vehicles (i.e. gliders, saildrones, drifting buoys), fixed ocean observing platform (moored buoys, rigs, offshore platforms and any deep-sea platforms), ocean reference sites and marine animals, which are equipped with instrumentation for measuring geo-physical variables. The WMO numbers are allocated to ocean platform depending upon deployment area, and platform type (i.e. drifting buoys, moored buoys, ocean reference sites, and profiling floats).

There are some exceptions:

- The HF radar network identifies its platform by an EDIOS code. This permits to consider the site (area covered by the monitoring) with the total currents and each single antenna with the radial currents. The EDIOS codes are managed by the SeaDataNet project,

- Tide gauge network is a very old international network with many sub-groups and different objectives. The network covers from tsunami detection to sea level changes, hence needs more effort to complete harmonisation,
- For the vessels: an IMO code or and ICES code, if IMO code does not exist, is attached to the ship as host of various instruments on board and a WMO characterises each platform type on board, e.g. ferry box, glider, mooring, Argo …
- For the augmented observatories (biology): the WMO is not relevant and there is the equivalent for biology, namely a WoRMS code that registers all the marine species.

The **ISO 8601** is an International Organization for Standardization (ISO) standard that specifies the numerical representation of **date and time**, based on the Gregorian calendar and the 24-hour time system, respectively. Often presented in the following form: YYYY-MM-DDThh:mm:ssz

The Global Positioning System (GPS) uses the **World Geodetic System (WGS84)** as its reference coordinate system. The WGS84 is a standard for GPS.

A common language concerning the **parameter definition** and the **technical characteristics of the instrument** would also be a plus.

Controlled vocabularies are used by data creators and data managers to standardise information (metadata) in database and data files. They are an important prerequisite towards consistency and interoperability and greatly facilitate searching for data in web portals. The vocabulary services are technically managed and hosted by the BODC by means of the NERC Vocabulary Server (NVS2.0) and populated by SeaDataNet (SDN)[22]. For **parameter definition**, it can only be advised to use the P01 Parameter Usage Vocabulary[23], which has underpinned BODC data management systems since the 1990s. The P01 Parameter Usage Vocabulary is a controlled vocabulary for labelling data channels and data value fields (i.e. recorded parameters, observations and measurement variables in environmental datasets). It could also be recommended to preferably use the **CF standard name** to complete the value of the variable attribute "standard_name" in a netCDF file.

The **technical characteristics of the instrument** can be defined by using the SDN L22 code[24]. The SDN L22 code is a controlled vocabulary containing names and descriptions of instruments and sensors used in the oceanographic and related domains.

We could also recommend to use the P06 BODC-approved data storage units vocabulary[25] to define the **units** of the parameters of the different variables.

For a perfect understanding of the data, it is imperative to have information on the **quality control** carried out on the data in the metadata.

NetCDF (network Common Data Form) is a “self-documenting” data **format** for storing multidimensional (variable) scientific data. It is independent of the hardware architecture and allows the creation, access and

[22] https://vocab.nerc.ac.uk/collection/
[23] https://vocab.seadatanet.org/v_bodc_vocab_v2/search.asp?lib=P01
[24] https://vocab.seadatanet.org/v_bodc_vocab_v2/search.asp?lib=L22
[25] https://vocab.seadatanet.org/v_bodc_vocab_v2/search.asp?lib=P06

sharing of scientific data stored in the form of tables. A header describes the layout of the data in the rest of the file, and in particular the data tables. This header also contains an arbitrary list of metadata in the form of name/value attributes. Data arrays are linear and are stored in a simple way to allow efficient access to a subset of an array.

The **political access** concerns the opening or closing of the data to users or groups of users. As stated in the D3.7 “networks harmonisation recommendations”, the best is to have the data as opened as possible and as closed as necessary. This is what is followed by the EuroSea networks which have adopted a Creative Common licence (CC-BY 4.0), largely opened for usage with reference to the creator. The 2 networks, vessels and augmented observatories, that haven’t yet chosen a licence are in line with the idea of an open licence which grant the creator and that is readable by human and machine as the Creative Commons ones.

To **identify the data** is a crucial point and very often essential to motivate creators to give access to their data to a network, a data centre or a data integrator. Two options could be considered: either a DOI is allocated to each platform deployment (glider, Eulerian observations) or at a level of the full network dataset (DOI for the Argo GDAC dataset for example). The difficult point is to identify to which granularity a DOI should be attached to a data or a dataset. A possibility would be to have a DOI associated to a platform and a deployment site (i.e. to the creator). This means each time the platform is used in another site, the DOI is new which is relevant for any EuroSea platforms even if not with the same frequency. Then an additional DOI could characterise the network dataset or a community or few datasets (European, international outside Europe) or communities. This is already the case for the Argo network to which is associated a unique DOI referring to the full dataset.

If this is feasible, i.e. to have one DOI attached to the measured data and another one (of few other ones) characterising the full network, we will need then to have a link between the 2 levels and then the DOIs. This tree structure could be extended to data integrators as EMODNet, SeaDataNet and Copernicus with the objective than when a user downloads or uses a data, the information is “recorded” at the data integrator level to the networks or the creators depending of needs. We are talking about citation attached to the data or dataset associated to a scientific paper or to the data itself (which can be the case for operational measurements) to have the creator clearly stated.

5.2 Metadata provided by integrators

Metadata/Integrators	CMEMS	EMODnet Map viewer	SeaDataNet
Platform n°/WMO	WMO	WMO	No
Date ISO	Yes	Yes	Yes
Coordinates WGS84	Yes	Yes	Yes
Depth	Yes	Yes	Yes
Parameter definition	Yes, P09	Only parameter name*	Yes, P09
Instrument used	Yes	No	Yes, L22
Unit of the parameter	Yes	Yes	Yes, P06
Quality control	Yes	No**	Yes
Format of data (netCDF)	Yes	Yes***	Yes
Access policy	Yes	Yes	Yes
Dataset citation	Yes	No	No

* The mapViewer is presenting themes and parameters with natural names (Temperature, Salinity etc) anyhow dataset and products applies standard conventions (controlled vocabularies)

** all the EMODnet projects implements QC/QF: some are applied at data source, some are applied at product level, etc. the QC/QF procedure is usually linked into the metadata.

*** all the data and products are made available via the central ERDDAP and are available in different formats, ranging from csv to netCFD.

5.3 Metadata provided by EuroSea Networks

Metadata/EuroSea networks	Argo network	Glider network	ASV (drifting buoys) network	Vessel network	Eulerian network	Tide gauge network	HF radar network	Augmented observatories network
Platform n°/WMO	Yes	Yes	Yes	No	Yes	No	Yes	No
Date ISO	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Coordinates WGS84	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Depth	Yes	Yes	Yes/surface	Yes	Yes	No	Yes	No
Parameter definition	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Instrument used	Yes	Yes	Yes	Yes	No	No	Yes	No
Unit of the parameter	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Quality control	Yes	Yes	Yes	Yes	Yes	No	Yes	No
Format of data (netCDF)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
Political Access	Yes	No	Yes	Yes	Yes	Yes	Yes	No
Dataset citation	Yes	No	No	No	Yes	No	Yes	No

5.4 What network find in what European Data Integrator?

EuroSea Network/European Data Integrator	Argo network	Glider network	ASV (drifting buoys) network	Vessel network	Eulerian network	Tide gauge network	HF radar network	Augmented observatories network
CMEMS	Argo profiler	No	No	Research vessel,	Mooring buoys	Tide gauge	HF radar	No
EMODnet Physics	Argo profiler	Gliders	Drifting buoys	No	Mooring buoys	Tide gauge	HF radar	No
SeaDataNet	Argo profiler	Gliders	Drifting buoys ASV	Research vessel Ship of Opportunity	Moored surface buoy Mooring	Sea level recorder	HF radar	Only seals's abundance

6 Example of use case: a moored platform

In this use case study, a user is looking for all in situ marine physical and chemical data for the moored buoy station E1-M3A[26] (WMO n°61277), situated in Mediterranean Sea and managed by the Hellenic Center for Marine Research (HCMR).

The E1-M3A station is located 24 nautical miles north of the island of Crete anchored at a depth of 1,400 meters and it has been part of the POSEIDON network since 2007. The mooring is currently the most developed physical-biogeochemical observing site of the POSEIDON system. The system integrates sensors for physical data (temperature, salinity, turbidity, current speed and direction) in the upper 500 m, as well as optical and chemical observations (dissolved oxygen, chlorophyll-a, PAR, nitrate) in the upper 100 m of the water column. A recent addition of surface pH and pCO2 sensors has further expanded the biochemical component of the station.

The surface buoy is equipped with sub-surface (1,5 m depth) sensors for temperature and conductivity (Aanderaa 3211) as well as wave and meteorological sensors (wind speed and direction, air temperature, atmospheric pressure, humidity, wave height and direction). For the measurement of conductivity and temperature at deep layers, two CT (250 and 350 m) and two CTD (250 and 350 m) instruments (SBE-37 IM) are attached on the central mooring line. Mooring line 2 is hosting four CTD probes (SBE-16) at 40, 65, 90 and 115 m. Apart of temperature, conductivity and pressure sensors, each probe is equipped with a transmissometer (Wetlabs C-Star), a chlorophyll fluorometer (Wet-labs WetStar), a PAR (Photosynthetically Active Radiation) sensor (Li-Cor LI-193-SA) and a dissolved oxygen sensor (SBE-23B). The mooring line host the RDI 75 kHz Long Ranger ADCP. This device is at 500 m depth, measuring the current profile from that depth up to surface[27].

This E1-M3A platform is listed on the OceanSites and DBCP worldwide systems. They offer one NetCDF file per year from 2007 to 2023, except for the year 2012 which is missing[28].
Each file offers the following parameters:

- Sea water temperature,
- Turbidity,
- Salinity
- Dissolved Oxygen,
- Chl – Fluorescence,
- Current direction,
- Current speed,
- Wave direction,
- Wave height,
- Wave period,
- Wind direction,
- Wind speed,
- Air temperature,

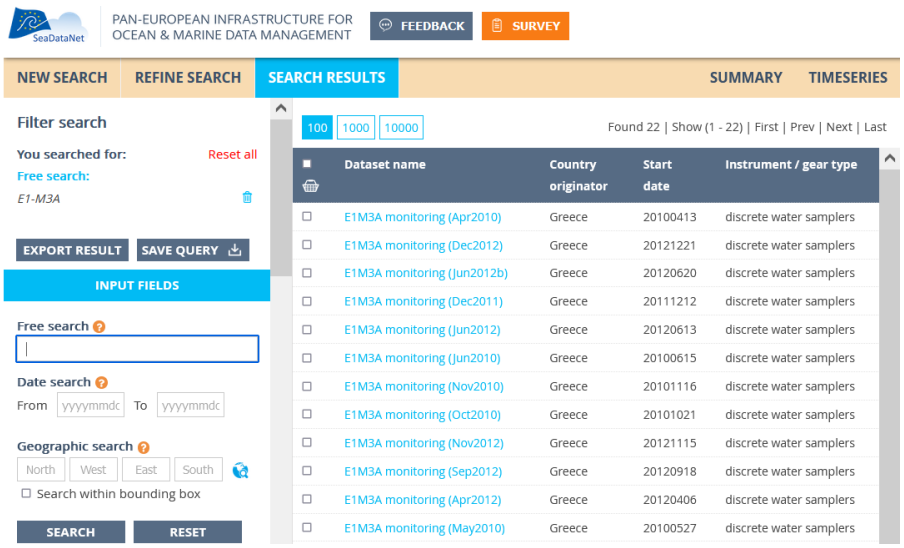
[26] https://poseidon.hcmr.gr/components/observing-components/buoys
[27] https://hal.science/hal-00298332v1/file/os-3-229-2007.pdf
[28] https://dods.ndbc.noaa.gov/thredds/catalog/data/oceansites/DATA/E1M3A/catalog.html

- Air pressure,
- Water depth,
- Sea pressure,
- pH.

The objective of this following study is to compare the data files and metadata of this platform ingested in the different European integrators.

Seadatanet, via the search data interface

We entered the name of the station E1-M3A in "Free search", without any other criteria[29]. The search result gives 22 results, with data available from March 2010 to May 2013, one file corresponds to one month:



Date available:
2010: March, April, May, June, October, November and December
2011: October, November and December
2012: March, April, May, June, August, September, November and December.
2013: April and May
(with June and November 2012 with 2 files)

Depending on the files, here is the associated metadata:
Abstract: POSEIDON site E1-M3A monitoring data
Platform_type: Self-propelled small boat or vessel
Coordinates: not uniform, depends on the file
Lat: 35.7836, Long: 24.917
Lat: 35.7836, Long: 24.9291
Lat: 35.717, Long: 25.0166
Lat: 35.7836, Long: 24.9291

[29] https://cdi.seadatanet.org/search

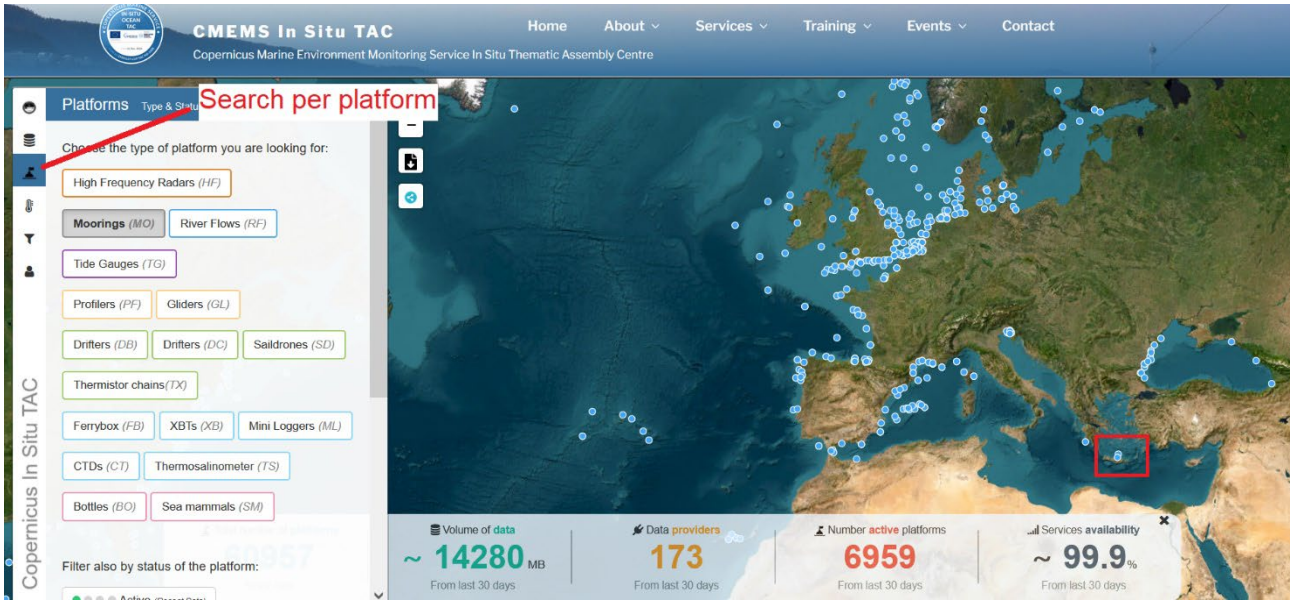
Files available from March 2010 to Mai 2013
Parameters: not uniform, depends on the file

- Vertical spatial coordinates,
- Ammonium and ammonia concentration parameters in water bodies,
- Chlorophyll pigment concentrations in water bodies,
- Nitrate concentration parameters in the water column,
- Phosphate concentration parameters in the water column,
- Phaeopigment concentrations in the water column,
- Silicate concentration parameters in the water column.

Each file is downloadable separately, with the proposed format ODV (Ocean Data View) only. ODV is an ASCII output format to handle profile, time series and trajectory data held. It needs to be imported directly into the ODV visualisation and analysis package. The software is freely available for non-commercial, non-military research and for teaching purposes. It is becoming an extremely popular choice within the marine community in which to visualise data.

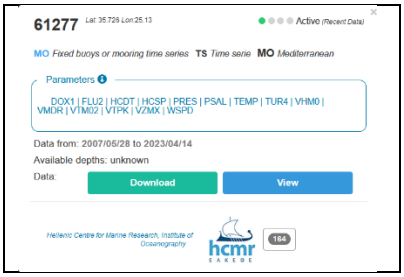
CMEMS, via the dashboard interface

On the CMEMS InSitu TAC[30], the search cannot be done by keyword, so we will use the search by platform and by status. Choose "Mooring". The map shows the available moorings. The station is easily identifiable by zooming in on the area and after opening the platform file. The platform is identified via its WMO number.



Click on the identified station on the map to open the following pop up and obtain all information on the station:

[30] http://www.marineinsitu.eu/dashboard/



This pop up allows either to visualise the data on graph, or to download them directly.
The metadata associated to these stations are:
Platform code: 61277 (WMO number)
Type of platform: Surface fixed buoy
Data availability: from 28/05/2007 to 14/04/2023
Area: Mediterranean
Type of data: time serie
Coordinates: Lat: 35.7263
Long: 25.1307
Parameters: DOX1 | FLU2 | HCDT | HCSP | PRES | PSAL | TEMP | TUR4 | VHM0 | VMDR | VTM02 | VTPK | VZMX | WSPD (using the P09 (MEDATLAS Parameter Usage Vocabulary)) Other parameters are available in the NetCDF files but not mentioned in the sheet.
Institution: Hellenic Center for Marine Research (HCMR)
Last observation date: 2023-04-14T06:00:00Z
Data download: no possibility to filter by variable. The choice is made by period: data of the last 30 days, the last 5 years or all the data. The files are almost complete (example: missing temperature data from January 2016 to June 2016, missing dissolved oxygen data from July 2015 to August 2016…).
All files downloaded are NetDCF files.

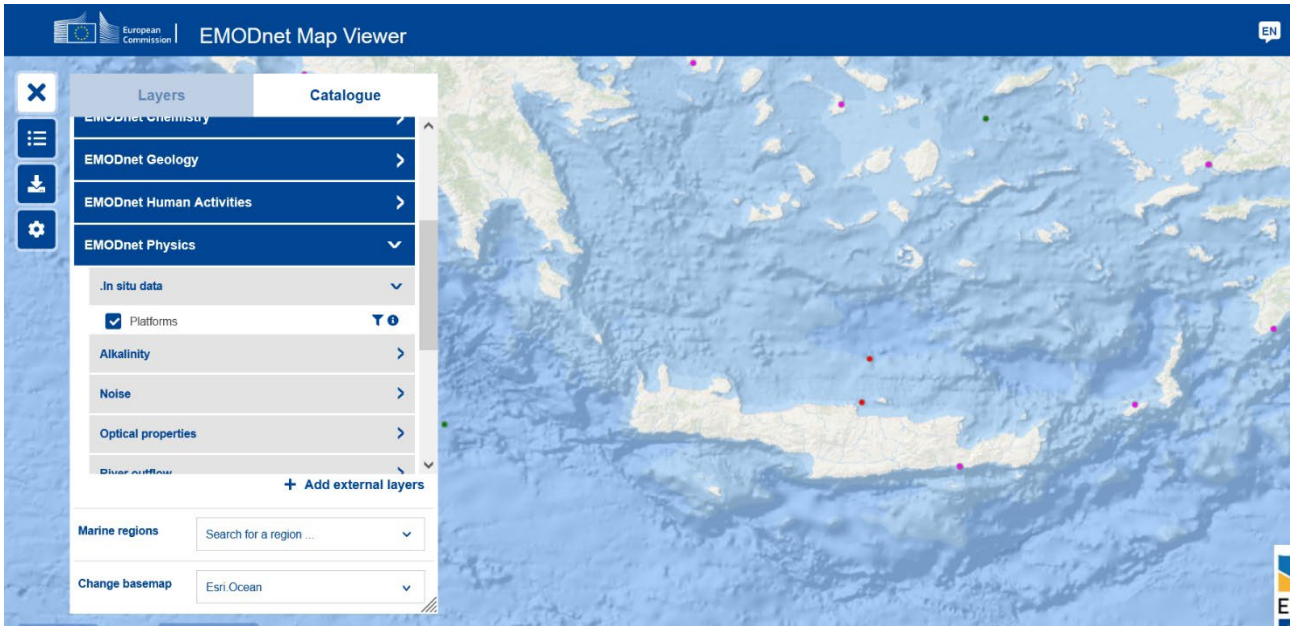
EMODnet: via the Map viewer

On the EMODnet Map viewer, the search must also be done by platform[31].

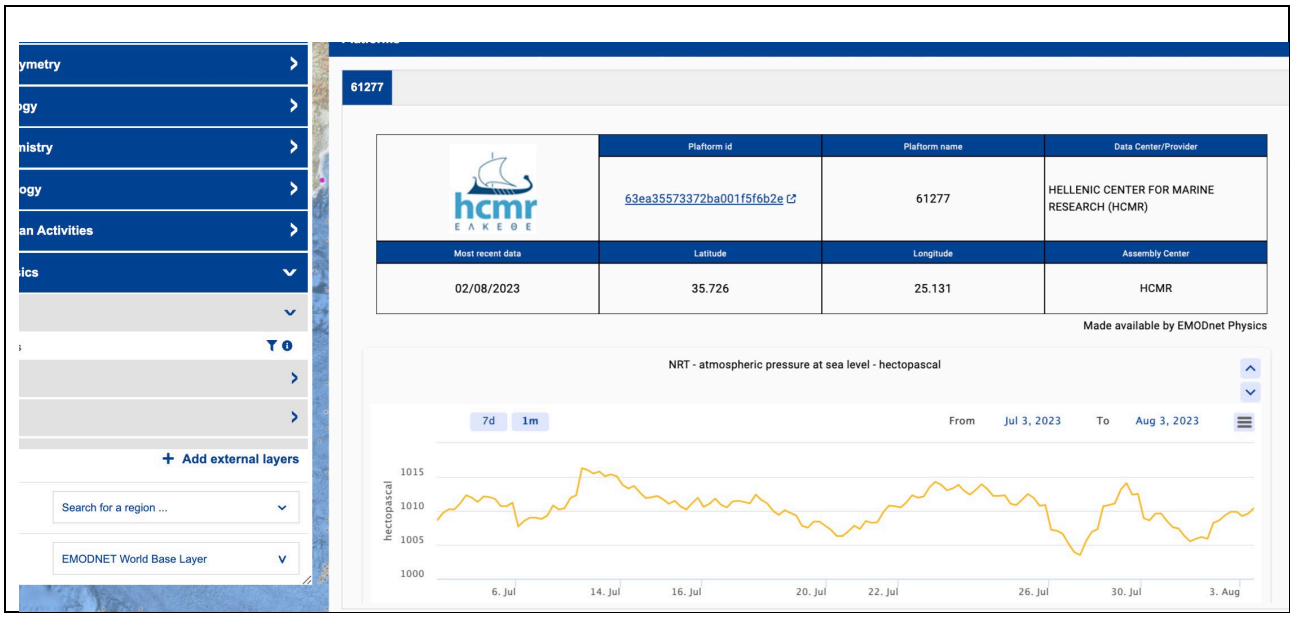
- Select “Catalogue”,
- Select “EMODnet Physics”,
- Select “In Situ Data”,
- Select “Platform”.

All platforms available in EMODnet Physic are displayed on the map. The station is easily identifiable by zooming in on the area and after opening the platform file. The platform is identified via its WMO number.

[31] https://emodnet.ec.europa.eu/geoviewer/

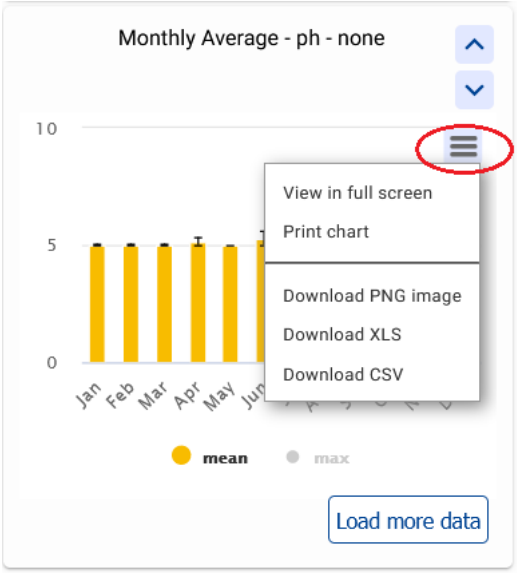


Click on the identified station on the map to open the following pop up and obtain all information on the station:



The metadata associated to these stations are:
Platform ID: 7271 (click on this number in the pop up to be redirected to a fullscreen page)
Platform name/WMO n°: 61277
Data centre/Producer: HCMR (Hellenic Center for Marine Research, Institute of Oceanography – Greece)
Last date: 14/04/2023
Coordinates: Lat: 35.726
Long: 25.131
EMODnet gives access to charts by variables. For each chart, it is possible to download the associated data. The interface gives the possibility to load more data (one year or more by specifying the period in the

calendar). For each platform the data availability may be different, for this specific platform data is available from 2007 to 2023. The platform page also provides the user with monthly averages (trends) for the selected parameter.



The data can be downloaded in .XLS or .CSV format. The graph is downloadable in .PNG image format.
Available charts and data associated:

- NRT and monthly average - atmospheric pressure at sea level – hectopascal,
- NRT and monthly average - dissolved oxygen - ml/l
- NRT and monthly average - fluorescence - milligram/m3
- NRT and monthly average - Gust wind from direction relative true north – degree
- NRT and monthly average - gust wind speed - meter/second
- NRT and monthly average - current to direction relative true north – degree
- NRT and monthly average - horizontal current speed - meter/second
- NRT and monthly average - practical salinity – psu
- NRT and monthly average - sea temperature - degree_Celsius
- NRT and monthly average – turbidity – ntu
- NRT and monthly average - spectral signif. wave height – meter
- NRT and monthly average - Mean wave direction from (Mdir) – degree
- NRT and monthly average - Spectral moments (0,2) wave period (Tm02) – second
- NRT and monthly average - wave spectrum peak period – second
- NRT and monthly average - Maximum zero crossing wave height (Hmax) – meter
- NRT and monthly average - wind from direction relative true north – degree
- NRT and monthly average - horizontal wind speed - meter/second
- Monthly Average - air temperature in dry bulb - degree_Celsius
- Monthly Average - ph – none

Conclusion:
The three integrators are serving different stakeholder and this may lead to some differences in the way the data are made available. This case study shows that there are some disparities but also some similarities between these three European integrators.

To do a search, for example, the SeaDataNet integrator allows to use keywords, which allows, in our case, to find instantly files associated to the platform “E1-M3A”. A search with the WMO number of the platform does not work in SeaDataNet, it is then necessary to use the name of the platform.

For EMODnet (Physics) and CMEMS integrators, the search is done by using the visualisation tool and by selecting the platform on the map. This last search seems less precise, could be a source of errors and could become very tedious, in the case that there are several platforms in the same area. Indeed, to obtain information on the platforms, you have to click on them, and open the associated sheet.

The new EMODnet Central portal does not allow to search by WMO number, whereas the CMEMS allows it. This possibility makes the search easier. For the CMEMS integrator, the name of the platform “E1-M3A” is only mentioned in the NetCDF data file. For the EMODnet integrator, the “E1-M3A” name is not mentioned in the downloaded files. This is the WMO number only that is used to identify the platform.

Regarding data availability, the integrator SeaDataNet only offers a small amount of data, only 3 years, while the CMEMS and EMODnet offer almost 16 years of complete data.

As far as the data format are concerned, here again, we do not find the same things.
EMODnet offers charts, with downloadable data in .CSV format, if the user needs other formats (netcdf) he has to use the advanced platform page downloading feature or the ERDDAP interface. Anyhow these tools need the user to know the portal and are difficult to find at a first access.
SeaDataNet and CMEMS offer NetCDF data files as a primary format for downloadable file.

Concerning parameters, the CMEMS integrator provides many parameters measured by the “E1-M3A” platform during the period between 2007 and 2023. EMODnet (Physics) offers the same parameters with the same duration as the CMEMS (except the water depth and the sea pressure). SeaDataNet offers specific chemistry data (phosphate, silicate, nitrate, ammonium and phaeopigment) only between March 2010 to March 2013. This chemical data is the only one in this integrator. SeaDataNet does not provide other parameters (yet).

7 Example of use case: study of restricted marine area

In this second study, we focused on a very specific area: the Strait of Gibraltar. In this case, our user wishes to recover a maximum of in situ physical and chemical data relating to this area. Let's compare what we can find in the three integrators.

Approximate search area:

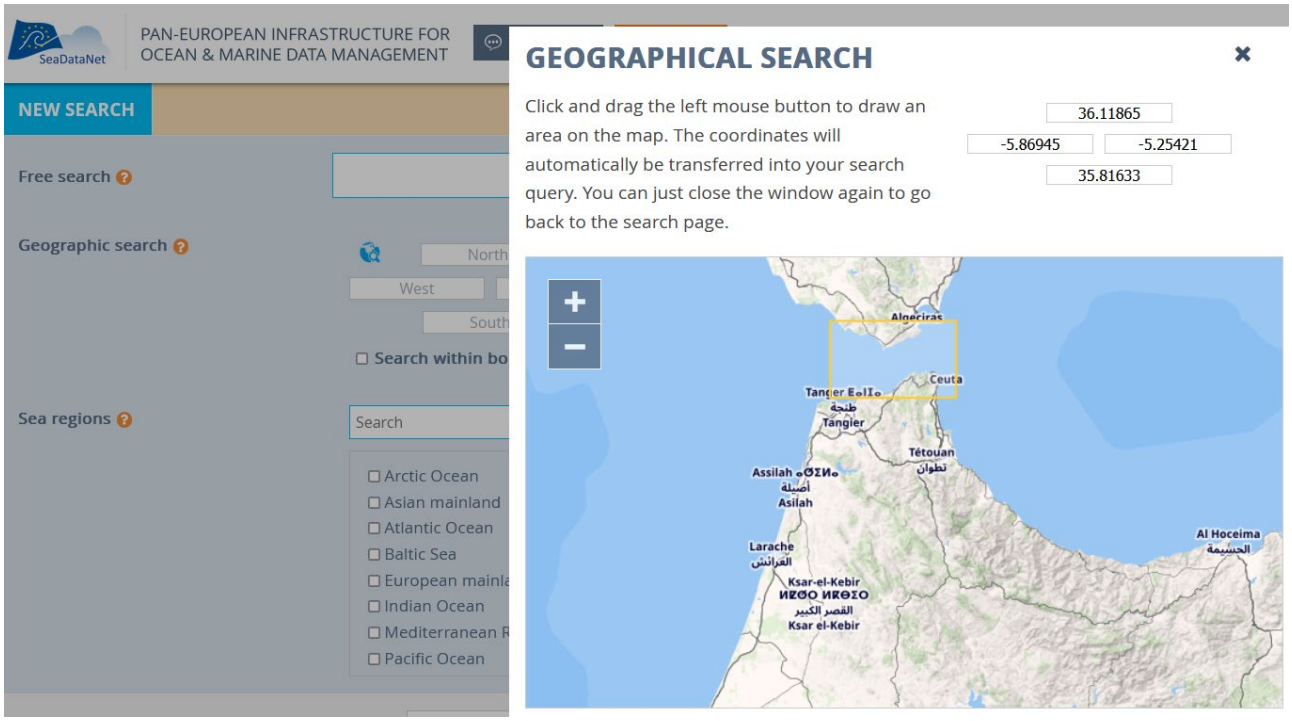
Latitude: between 36.1 to 35.8 degrees

Longitude: between -5.8 to -5.3 degrees

SeaDataNet, via the search data interface

The search is done through the geographic search button by drawing an area on the map. To access the data, you must first log in.

Accept and close the windows. Click on “Search”. The search result gives 2302 results, with data from 1943 to 2023. Majority of data come from Spain (1369 results) and France (574 results). The parameters measured are essentially the temperature and the salinity of the water column (1592 results), but we can also find measurements of current (80 results), sea level (353 results) and chemical parameters such as nutrients (189 results), dissolved gases (342 results) and pigments (211 results). The SeaDataNet integrator provides data from several types of platforms: research, naval, fishing vessels, vessels of opportunity, moorings, coastal and onshore structures, drifting subsurface profiling floats and unknown platforms.

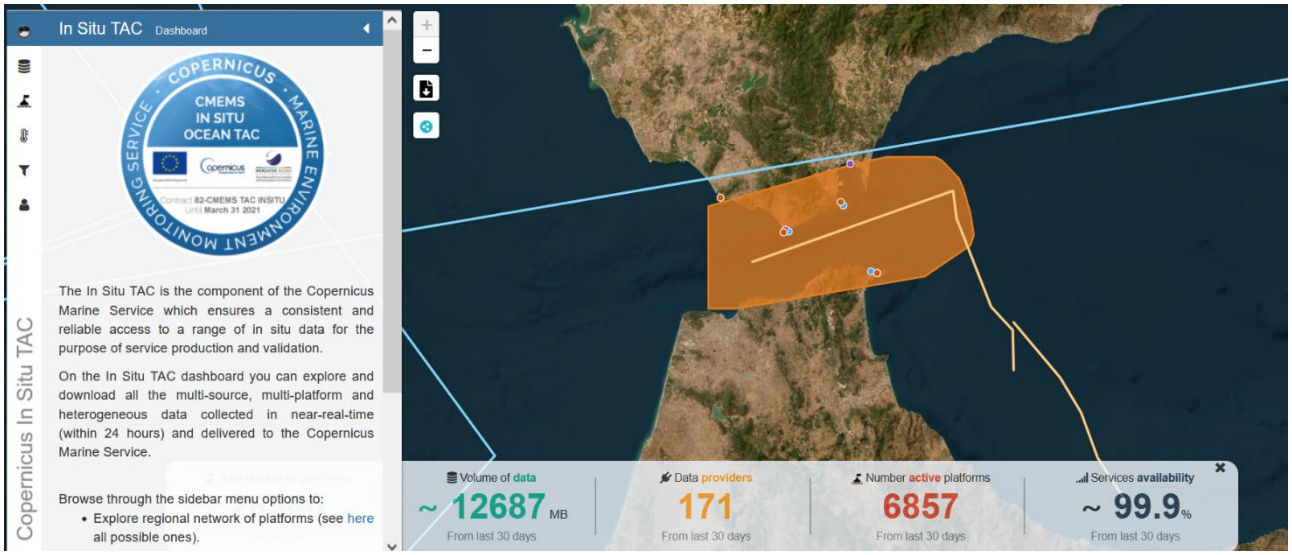


It is possible to refine the search on these 2302 results, which is particularly practical, according to:

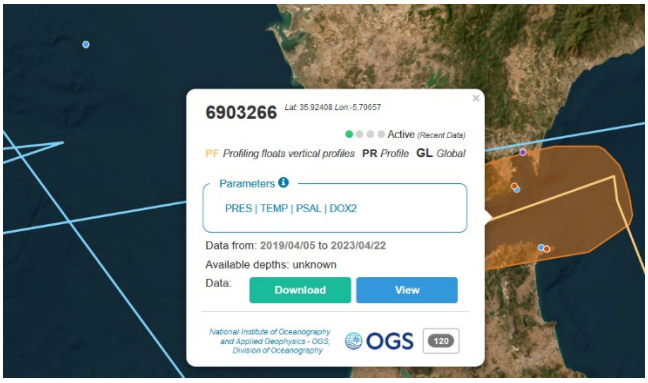
- Discipline (physical, biology, chemistry, geology, human activities…),
- Parameter group (water column temperature and salinity, current, nutrient…),
- Measuring area type (point, curve and surface),
- Date,
- Instrument type,
- Platform type,
- Data originator and data originator country…

CMEMS, via the dashboard interface

On the CMEMS InSitu TAC[32], searching by marine area is done using the map and zooming in on the selected area. To access the data, you must first log in.

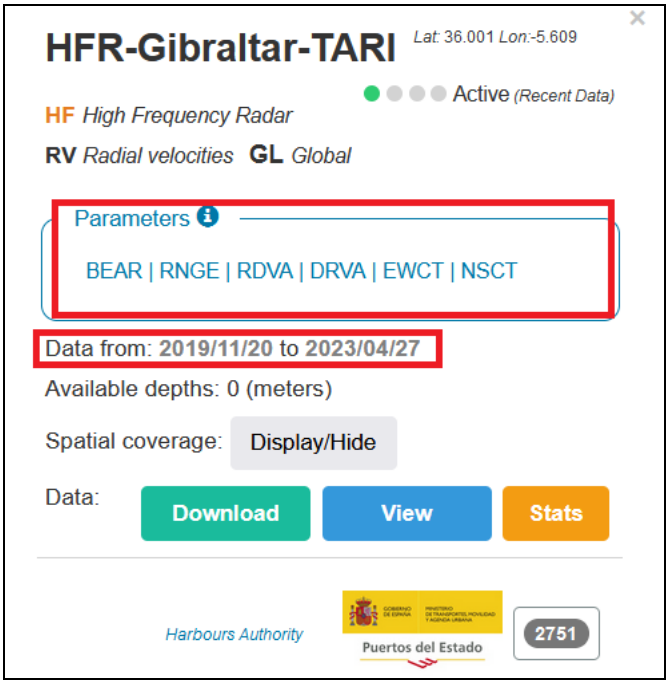


In the CMEMS integrator and for this selected area, the data that can be retrieved seems quite few. Indeed, the integrator lists only 11 platforms and essentially coastal platforms. Most platforms provide data between 2019 and 2023. Only one fixed buoy, situated in coastal area provides data between 1985 and 2023. The proposed measured parameters are temperature, salinity, current, velocity, dissolved oxygen and wave direction… Only one platform (a profiling floats vertical profiles) provides dissolved oxygen parameter between 2019 and 2023.
To obtain data, you must select each platform and download the associated data.



The "view" button gives access to a page which allows you to visualise on a graph the parameters of the platform over different periods to be selected (last available, last 7 days, last 30 days, and select month-year) and the metadata associated to the selected data. The "download" button gives access to downloading all the data from the platform, in NetCDF format. Just click on the chosen file to trigger the download.
For each platform, present on the map, the list of parameters and the dates of availability are indicated.

[32] http://www.marineinsitu.eu/dashboard/



For the HF radar platforms, another button "Stats" is available. It gives performances metrics for the HF radar: spatial vs temporal availability.

EMODnet: via the Map viewer

For the EMODnet integrator, the search for in situ data will be done by using the Map viewer[33] and by selecting the EMODnet Physic catalog which lists the in situ data via the platforms. No need to register.

On the EMODnet Map viewer:

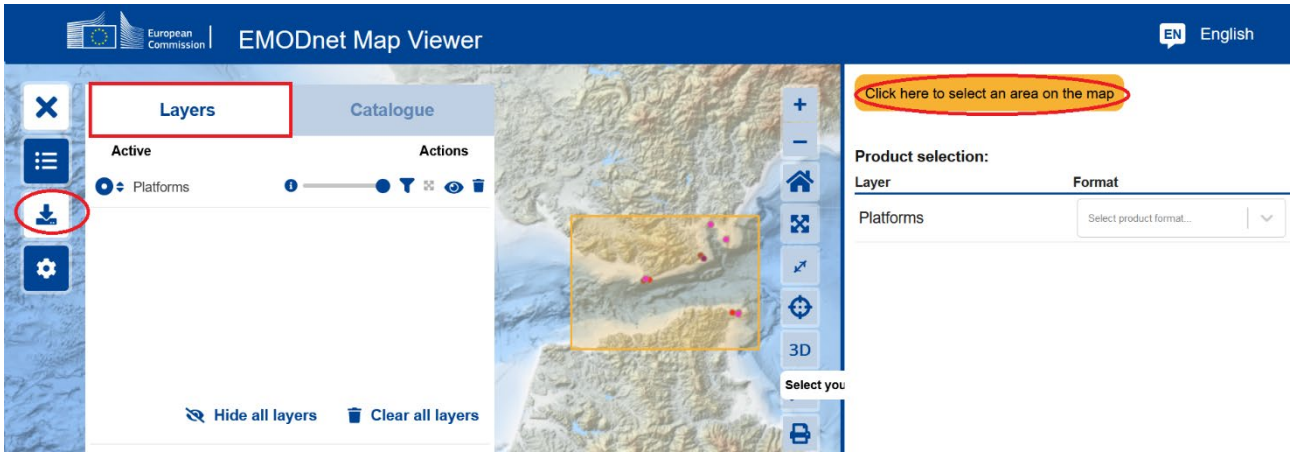
- Select "Catalogue",
- Select "EMODnet Physics",
- Select "In Situ Data",
- Select "Platform".

All platform available in EMODnet are displayed on the map. Select "Layers" then, click on "Download" and click on the button "Click here to select an area on the map" in order to select the study area of the use case.

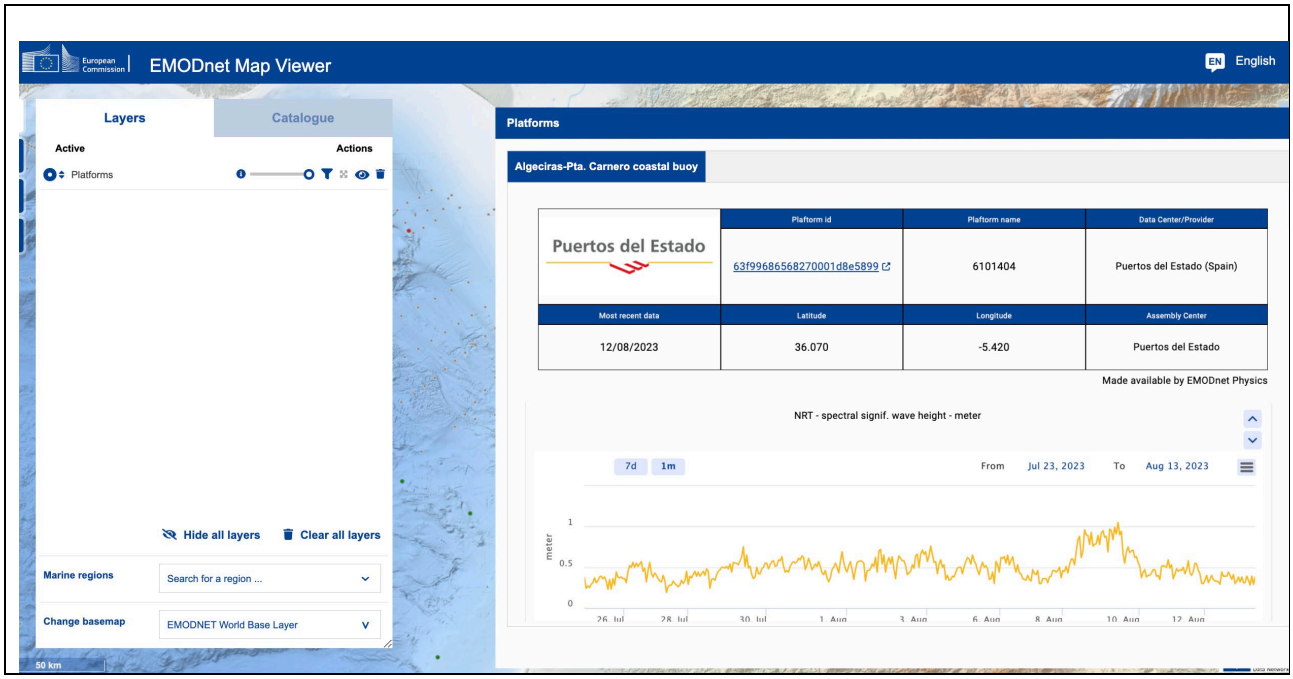
As a default EMODnet shows 9 platforms (which are the platform that delivered at least one dataset during the past 7 days). Anyhow if the user selects layers and the apply a larger integration time (all times) the number of available datasets increase (11 stations, 2 ships, a number of CTD casts).

[33] https://emodnet.ec.europa.eu/geoviewer/

Although the mapviewer offers a download feature, this is only indented to download the shape file/tail/ positions of the stations (depends on the selected themes).
In case of the in situ platform, this feature is not operational yet and user has to download data by browsing the single stations.



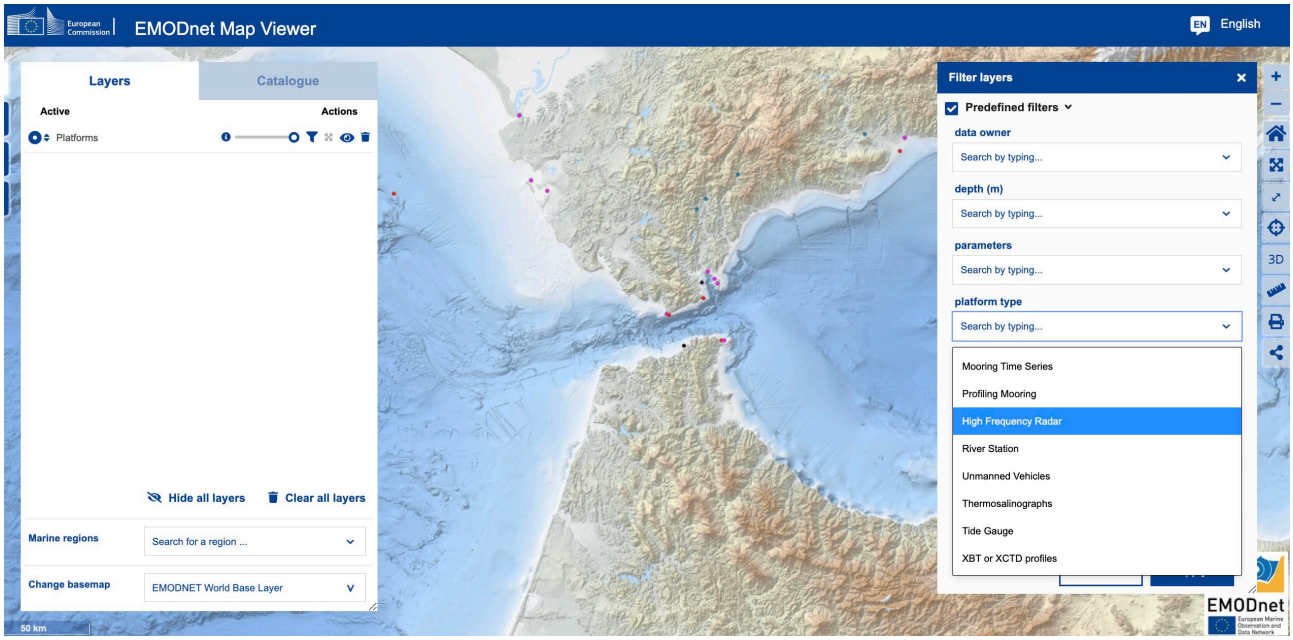
To do this, click on the identified platform on the map to open the following pop up and obtain metadata and data from the platform page:



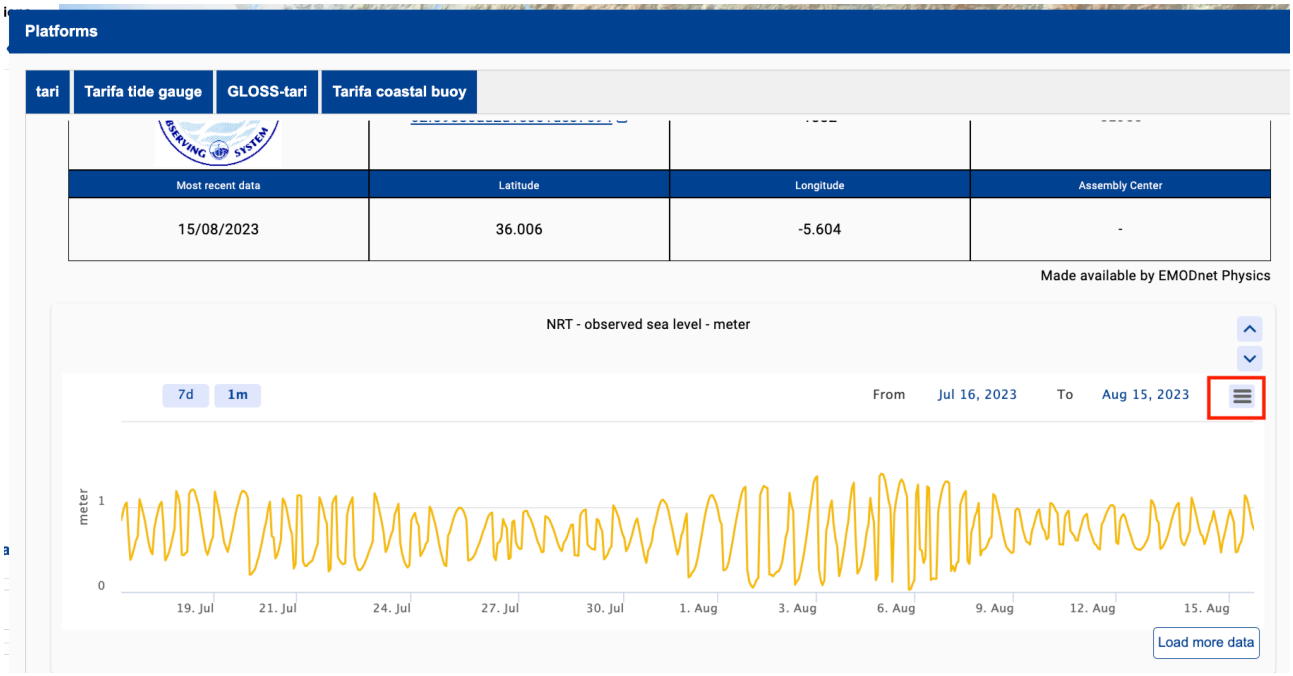
For most platforms, sheet is composed as follows:

- A summary table with platform information:
  - Platform ID,
  - Platform name,
  - Latitude,
  - Longitude,
  - Last date,
  - Data centre/producer.
- Charts.

Going back to the default view, of the 9 platforms shown in the study area, there are 3 fixed buoys, 3 tige gauges and 3 radars.
The user can subset the platform type by using the filtering feature.



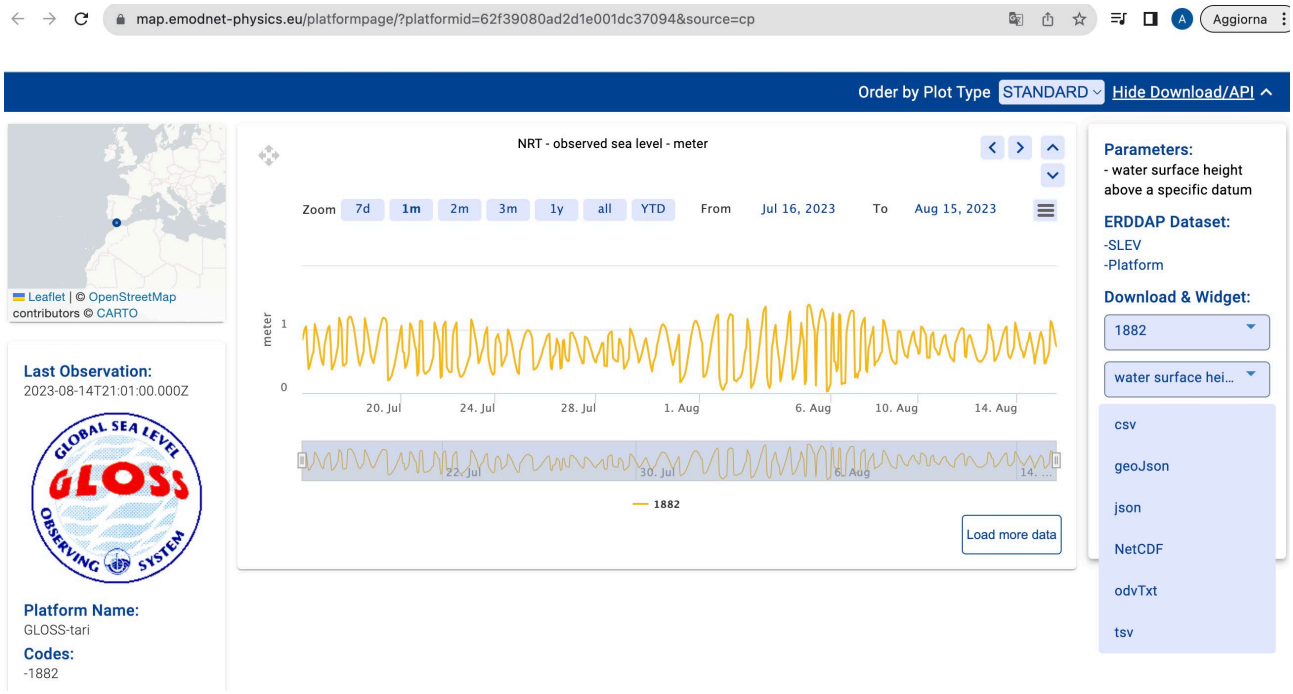
The three fixed buoys give access to several graphs, exclusively in physical domain and downloadable in .PNG, .XLS or .CSV format directly from the chart:



- NRT and monthly average – Spectral signif. Wave height – meter
- NRT and monthly average – Aver. Height highest 1/3 wave - meter
- NRT and monthly average – Maximum zero crossing wave height (Hmax) - meter
- NRT and monthly average – Average zero crossing wave height (Hzm) - meter
- NRT and monthly average – Average height highest 1/10 wave (H1/10) - meter
- NRT and monthly average – Wave spectrum peak period - second
- NRT and monthly average – Aver. period highest 1/3 wave - second
- NRT and monthly average – Aver. Zero crossing wave period – second

- NRT and monthly average – Average period highest 1/10 wave (T1/10) - second
- NRT and monthly average – Spectral moments (0,2) wave period (Tm02) - second
- NRT and monthly average – Mean wave direction from (Mdir) - degree
- NRT and monthly average – Sea temperature – degree_Celsius

While other formats area available from the full platform page (that is opened by clicking on the link next to the provider logo)



Two tige gauges give access to only one graph:

- NRT – Observed sea level - meter

The last tige gauge displays “No data available”.

The three radars each present a first graph but without associated legend, which does not allow the user to know which parameter is measured (or presented) on this graph.
The other measured parameters by the radar are:

- Cumulative – current magnitude and direction (m/s),
- Windrose - currents direction and magnitude frequency,
- NSCT - south-north current component (m/s),
- EWCT - west-east current component (m/s),

Conclusion

Again, this case study shows the disparities between the three European integrators, but essentially between SeaDataNet and the two others, EMODnet and CMEMS. It also shows some similarities between the EMODnet and the CMEMS integrators.

In this case study, the integrator SeaDataNet offers a large amount of data, both physical and chemical, coming mainly from marine cruises but also from fixed platforms (mooring, radar, etc.). Apart from the radar platforms, we do not find the same data in the SeaDataNet integrator and in the EMODnet and the CMEMS integrators. The CMEMS and EMODnet integrators, meanwhile, offer almost the same thing: the same platforms. The CMEMS integrator, however, offers chemical data that is not found in the EMODnet integrator, which is natural as the selected theme is Physics and no Chemistry (which is offering gridded data and not in situ data). The integrator EMODnet offers turnkey graphs, which are quite interesting.

8 Conclusion

This Data Integration deliverable, allows providing an overview of the 3 main EU data integrators namely CMEMS, EMODnet and SeaDataNet. This needs to be completed and enhanced to have clearer view of the complementarities of the 3 integrators before by largely displayed on websites and to all users.

The links and ingestions of the data networks involved in EuroSea has been described through the 3 deliverables due on this task. This could be a basis to improve them in the future.

In addition, this deliverable through the two use case studies made possible to highlight the differences, the similarities and the complementarities that exist between the 3 EU integrators. First, if we consider only the in situ data, the CMEMS and the EMODnet integrators are very similar.

We found, with a few exceptions, the same results, as the data come from the same platforms. What differs, is the presentation of the data and the access to (and download of) the data.

The SeaDataNet integrator really stands out from the two others. In the two use cases, we only found very few data identical to those found in the EMODnet and CMEMS integrators.

It is in total complementarity with these two integrators.