Vol. 59 – SUPPLEMENT 1

Bollettino di Geofisica
teorica ed applicata

An International Journal of Earth Sciences

IMDIS 2018
International Conference on Marine Data and Information Systems

5-7 November, 2018 - Barcelona (Spain)

Organised by
CSIC
jointly with IFREMER, OGS, IOC/IODE, in the frame of SeaDataCloud project (EU H2020 - Grant Agreement 730960).

Organizing Committee
Susana Diez Tagarro - CSIC, Spain; **Jordi Sorribas Cervantes** - CSIC, Spain; **Michele Fichaut** - IFREMER, France; **Vanessa Tosello** - IFREMER, France.

International Steering Committee
Vittorio Barale - JRC, EU; **Alexander Barth** - ULiège, Belgium; **Sergey Belov** - RIHMI, Russian Federation; **Simon Claus** - VLIZ, Belgium; **Susana Diez Tagarro** - CSIC, Spain; **Michèle Fichaut** - IFREMER, France; **Lotta Fyrberg** - SMHI, Sweden; **Alessandra Giorgetti** - OGS, Italy; **Mark Hebden** - BODC, United Kingdom; **Neil Holdsworth** - ICES, Denmark; **Athanasia Iona** - HCMR, Greece; **Adam Leadbetter** - MI, Ireland; **Damien Lecarpentier** - CSC, Finland; **Gilbert Maudire** - IFREMER, France; **Volodymyr Myroshnychenko** - METU-IMS, Turkey; **Leda Pecci** - ENEA, Italy; **Lesley Rickards** - BODC, United Kingdom; **Helge Sagen** - IMR, Norway; **Dick Schaap** - MARIS, The Netherlands; **Reiner Schlitzer** - AWI, Germany **Serge Scory** - RBINS-BMDC, Belgium; **Simona Simoncelli** - INGV, Italy; **Jordi Sorribas Cervantes** - CSIC, Spain; **Susanne Tamm** - BSH, Germany.

Istituto Nazionale di Oceanografia
e di Geofisica Sperimentale

ISSN 0006-6729

TABLE OF CONTENTS

Posters

Posters

Posters

Posters

SeaDataNet: What's up in the cloud?

Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), Michele.fichaut@ifremer.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
on behalf of the SeaDataCloud consortium

Access to marine data is of vital importance for marine research and a key issue for various studies, from the climate change prediction to off shore engineering. But, the marine observing systems are highly fragmented: for example, in the countries bordering the European seas, more than 600 scientific data collecting laboratories from governmental organisations and private industry are collecting marine data by using various sensors on board of research vessels, submarines, fixed and drifting platforms, airplanes and satellites, to measure physical, geophysical, geological, biological and chemical parameters, biological species etc. The collected data used to be neither easily accessible, nor standardized. They were not always validated and their security and availability not assured.

Therefore the standardised SeaDataNet infrastructure was established for managing the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems. It was implemented in several phases during the SeaDataNet project (2006-2011), grant agreement 026212, EU Sixth Framework Programme, the SeaDataNet 2 project (2011-2015), grant agreement 283607, EU Seventh Framework Programme and the current SeaDataCloud H2020 European project (under grant agreement 730960).

SeaDataNet contributes to build research excellence in Europe by networking and enhancing the currently existing infrastructures, which are the National Oceanographic Data Centres (NODC) or data focal points of 34 countries, active in data collection. The networking of these professional data centres, in a unique SeaDataNet virtual data management system provides integrated data sets of standardized quality on-line.

SeaDataCloud project, started in November 2016, aims at considerably advancing SeaDataNet Services and increasing their usage, adopting cloud and High Performance Computing technology for better performance, with the following objectives:

- improve discovery and access services for users and data providers;
- optimise connecting data providers and their data centres and data streams to the infrastructure;
- improve interoperability with other European and international networks to provide users overview and access to additional data sources and;
- develop a Virtual Research Environment with tools for analysing data and generating and publishing data products.

The presentation will give an overview of the overall setting and background of the SeaDataNet infrastructure, the partnership and relations with other projects and will insist more particularly on the new challenges and the technical progress that have been achieved since the start of SeaDataCloud project, involving different partners of the Technical Task Group:

- The new workflow (see Fig. 1) and new service components that are developed using the EUDAT cloud: replication of data from the data centres connected to SeaDataNet infrastructure, quality checks on data file formats, data file integrity and compliancy with metadata, duplicate detection… (MARIS, IFREMER, CINECA, DKRZ, GRNET, CSC, STFC);
- The expansion and extra services on the SeaDataNet common vocabularies (BODC);
- The implementation of the linked data concept in order to improve the connections between SeaDataNet catalogues and improve the user's search functions (MI);
- The development of a Sensor Web Enablement (SWE) service for facilitating the publication of real-time sensor data and their historical time-series through interoperable standards (52°North);
- The definition of standard data format for new data types such as flow cytometer data (CNRS-MIO, BODC, IFREMER);
- The transformation of SeaDataNet formats into INSPIRE data standards positively analysed (BODC, SYKE, OGS);
- The development of a Virtual Appliance (VA) to ease the installation of the SeaDataNet componant at data centres (ENEA);
- The specification of a Virtual Research Environment (VRE) offering users the possibility to work online with high capacity and performance for big data and to configure virtual

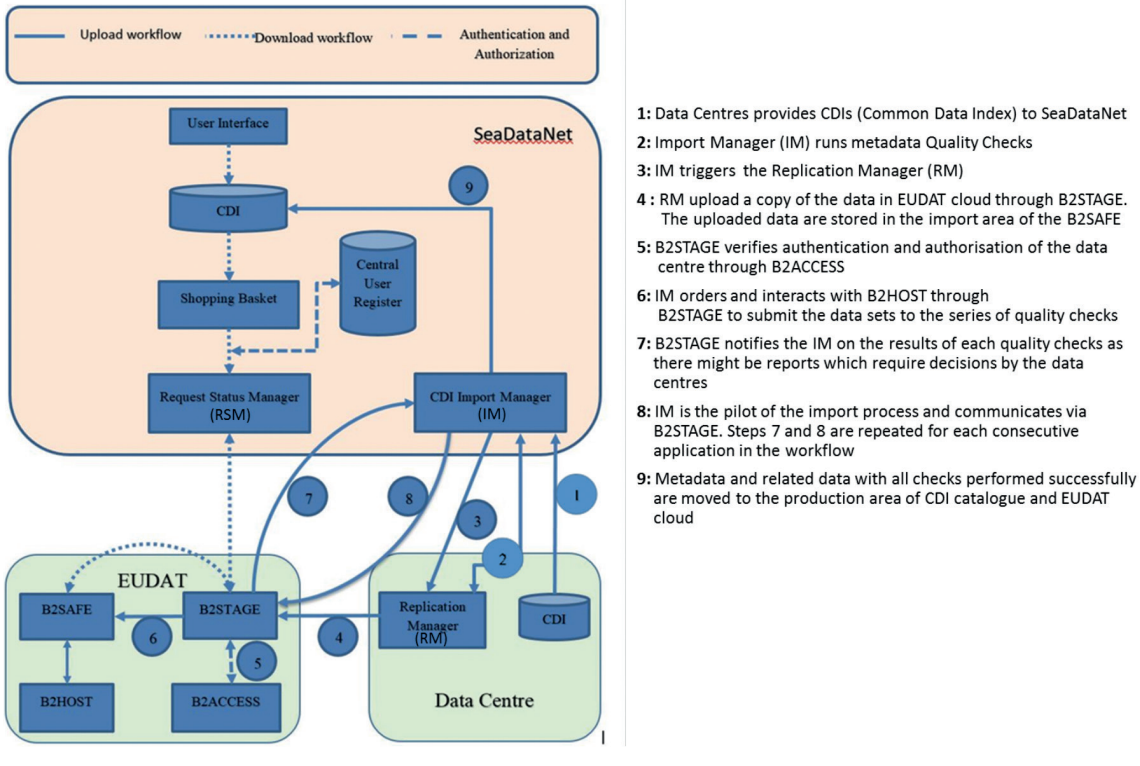


Fig. 1 - New workflow in SeaDataNet infrastructure, making use of EUDAT cloud and of the Replication Manager new SeaDataNet components.

workspaces for individuals or group with dedicated pools of data (MARIS, IFREMER, AWI, CINECA, DKRZ, GRNET, CSC, STFC, ULiège, VLIZ, SYKE, Deltares).

All these new developments are on-going, some of them are implemented in beta version others are already operational. The final objective is to provide benefits to SeaDataNet users and data managers, such as, for example:

Improvement of the data downloading service (speed-up the performance, expand data discovery and ease of use of data access and downloading),

Improvement of the data and metadata quality and coherence,

The easier installation of the SeaDataNet connection "package' in the data centres.

SESSION 1

Data services and tools in ocean science

- Services and tools for discovery, visualisation, dissemination
- Services and tools for data processing and data analysis
- Services and tools for data ingestion
- Services and tools for data management
- Virtual Research Environment (VRE), Cloud environment
- European Open Science Cloud (EOSC) and Blue Cloud pilot
- Open data for the marine domain

ORAL PRESENTATIONS

Bridging the gap between data and sensor information

Roland Koppe, Alfred Wegener Institute (Germany), Roland.Koppe@awi.de
Ana Macario, Alfred Wegener Institute (Germany), Ana.Macario@awi.de

Over the last 2 decades, the Alfred Wegener Institute (AWI) Helmholtz Center for Polar and Marine Research has been continuously committed to develop and sustain an eResearch infrastructure for coherent discovery, view, dissemination, and archival of scientific data and related information in polar and marine regions. Most of the data collected by scientists originates from research activities carried out on a wide variety of research platforms operated by AWI (vessels, aircraft, stations, buoys, moorings, in-situ ocean floor stations, drones, and ocean floor crawling systems). Archival and publishing in the information system PANGAEA along with DOI assignment to individual datasets is a typical end-of-line for most data producers.

In order to address the increasing heterogeneity of research platforms and respective devices and sensors along with varying project-driven requirements, we built a generic and cost-effective virtual research infrastructure, hereafter named O2A, intended to support the flow of sensor observation to archives. O2A components are modular, extensible and in compliance with OGC standards and FAIR principles, ensuring interoperability at international level and data re-use.

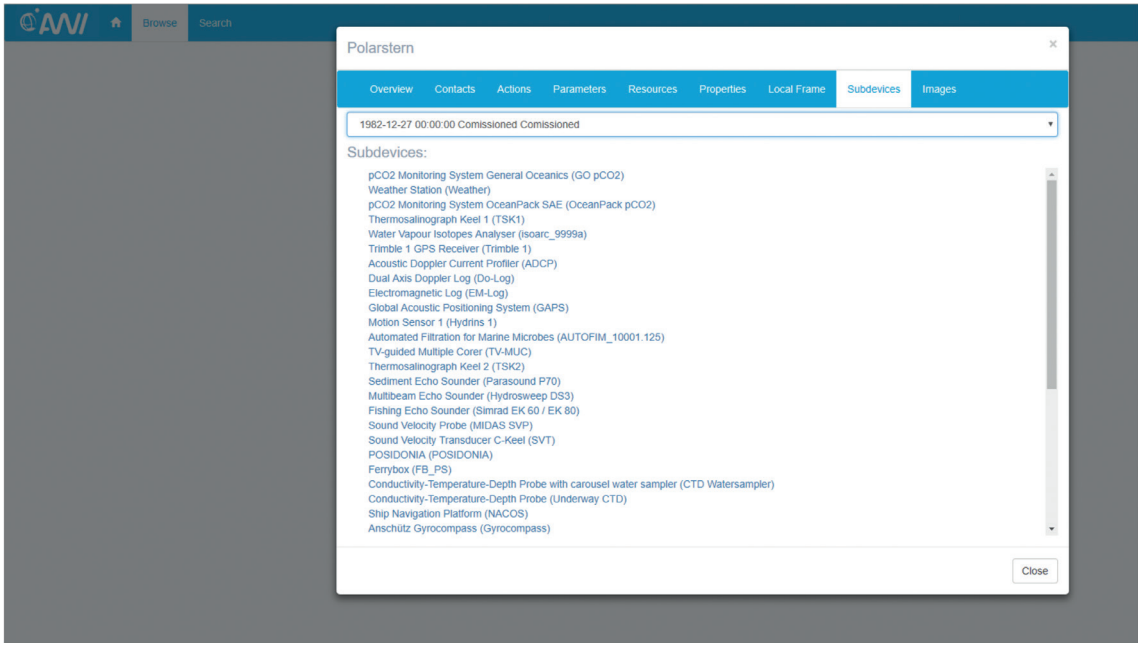


Fig. 1 - SENSOR https://sensor.awi.de, is one of the key components within O2A as a beginning-of-line for scientists. To date, over 1400 sensors are described.
In this figure we show the shipborne instruments for RV Polarstern.

The component SENSOR (http://sensor.awi.de), designed as repository for various types of information related to platforms, instruments and sensors, is intended to provide the following enhancements:

- Address specific requirements from scientists meant to support their scientific pipelines. Examples:
  - Integration of station lists from expeditions to action/event component (as part of the sensor provenance metadata);
  - Harmonization of various vocabularies with existing data acquisition systems on board of our vessels and stations (device names, types, parameters, units).
- In-house inventory and management of platforms/devices/sensors;
- Traceability of sensors in particular in regard to varying payloads for small vehicles (e.g. ROVs) and instrumentation of expeditions/landing parties and observatories.

SENSOR has been publicly available since 2015 (incl. bi-annual release packages) and will be adopted as repository for the international Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC; start planned for Autumn 2019). Given that platforms and sensors evolve in time (sensors are being calibrated, mounted/unmounted, etc.), it is clearly important to trace changes applied to these over time. We have intensively discussed the usability scenarios with AWI scientists in particular in regard to which types of actions/events associated with a sensor are relevant in the long run from data use perspective. While tracking changes to platforms/instruments/sensors is technically feasible, it became clear to us early on that in practice one has to decide which type of action/event (e.g. sensor calibration) warrants the creation of an individual version of a sensor (= a new PID). Pragmatically, we have opted to (a) narrow down the scenarios in which metadata versioning takes place (including minting PIDs/handles), and, (b) "force" certain changes to sensors to trigger the creation of a new version.

Having successfully integrated AWI Handle Server services with SENSOR, we are currently able to support:

- Sensor versioning, i.e. audit trail of changes;
- Minting of persistent identifiers, generating UUIDs in the handle syntax. Example: https://hdl.handle.net/10013/sensor.bf472f3a-d236-47ee-972e-0e67c86085eb;
- Automated generation of a sensor citation.

In this talk we will present our approach in detail and illustrate with concrete examples how sensor metadata enhances the quality of data archived in PANGAEA.

New SOCIB Data Catalog REST API

Juan Gabriel Fernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jfernandez@socib.es
Paz Rotllán, Sistema d'observació i predicció costaner de les Illes Balears (Spain), protllan@socib.es
Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Inmaculada Ruíz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), iruiz@socib.es
Miguel Charcos, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mcharcos@socib.es
Miquel Àngel Rújula, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mrujula@socib.es
Xisco Notario, Sistema d'observació i predicció costaner de les Illes Balears (Spain), xnotario@socib.es
Sonia Gómara, Sistema d'observació i predicció costaner de les Illes Balears (Spain), sgomara@socib.es
Miquel Gomila, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mgomila@socib.es
Joaquín Tintoré, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jtintore@socib.es

Introduction

API stands for Application Programming Interface. An API is basically a programmatic tool that is used to interact with a system. More specifically, the new SOCIB API offers the possibility of exploring our data catalog and retrieving the data it contains. This API is a REST API, which means that the requests are issued from a usual HTTP client, available with any programming language.

Regarding interoperability and discoverability, one of the key components that data providers should issue are REST APIs. In this line, SOCIB Data Centre has released in 2018 the Data Catalog API, a REST API that is meant to extend and enhance the capabilities of previous one (Data Discovery).

This new API eases the recovering of different data sources (i.e. observational data) coming from a wide range of platforms (e.g. oceanographical buoy, coastal stations, weather stations, autonomous underwater vehicles, drifter buoys...). Next versions will include also models, images, etc... to the list of available data sources.

The new SOCIB API is addressed to two different types of users: the IT (software developer) user and the data scientific user. For the IT expert, the use will be almost trivial. For scientific data experts, with no experience on this kind of tool, some training will be needed. A success story is the development of the SOCIB Data Catalog web that fully relies on the capacity of this API.

Main features

One of the main new features is the possibility of discovering and retrieving data without temporal restrictions, unlike Data Discovery API, that only lets you retrieve the latest 60 days of data, at most. Also important is that the data can be filtered with new sets of criteria and retrieved with several resampling methods.

Gridded data type is now also supported (currently only for HF Radar data). Another important breakthrough has been achieved introducing the concept of "data products", that let us provide grouped data in almost any way. The API also supports authentication via API keys, so the tracking

of users will now be possible.

A netCDF layer has been developed to act as a middleman between netCDF files (raw data format), queries and JSON formatted responses. This layer aggregates, resamples, prepares and formats the data from the netCDF files that match the query criteria.

This API has been built following the Open API 3.0 specification and a Swagger UI tool has been used to implement an interactive documentation. This is a major improvement since the API itself becomes interoperable.

Data model

The SOCIB API enables users to discover quickly and easily which data products are available at SOCIB. Data products are combinations of data sources grouped by a certain criteria (i.e. all belong to the same campaign or project).

At the core of this API a data model has been developed in order to support this functionality. The main concept in this data model is the so-called data-source, which represents the data produced during the deployment of an instrument. Also a dataset is used to store the data derived from several instruments of a platform (i.e. platforms products such as sea level). The following diagram represents this model in detail:

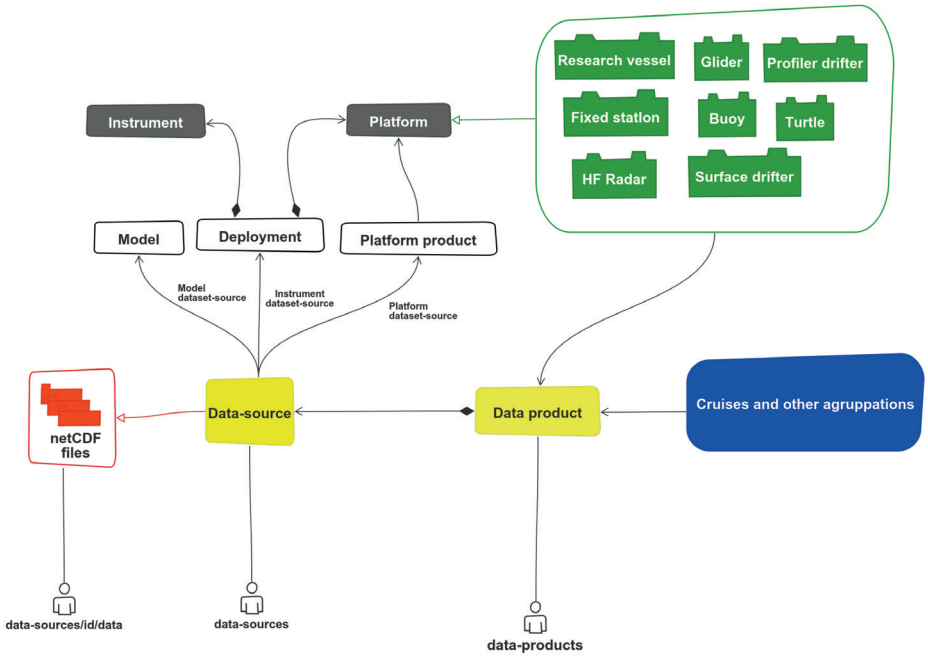


Fig. 1 - New SOCIB REST API data model.

Conclusions

Providing up-to-date tools for users to fetch data from data providers is a must regarding user uptake and engagement. Since the release of the first SOCIB API (Data Discovery), new technologies and methodologies arisen and led to the necessity of readdressing such first approach. The new SOCIB API, the Data Catalog API, is meant to fulfil late requirements in terms of web services and interoperability.

Australian Ocean Data Network utilises Amazon Web Service Batch processing for gridded data

Sebastien Mancini, Integrated Marine Observing System, University of Tasmania (Australia), sebastien.mancini@utas.edu.au

Roger Proctor, Integrated Marine Observing System, University of Tasmania (Australia), roger.proctor@utas.edu.au

Peter Blain, Integrated Marine Observing System, University of Tasmania (Australia), peter.blain@utas.edu.au

Kate Reid, Integrated Marine Observing System, University of Tasmania (Australia), kate.reid@utas.edu.au

The Australian Ocean Data Network (AODN) is an interoperable online network of marine and climate data resources. It is a collaboration between six major Australian Commonwealth agencies, the Integrated Marine Observing System (IMOS), and a growing list of universities, state government offices and others organisations in Australia, New-Zealand and the Pacific.

AODN data collections cover a large geographic area (from coast to open ocean, from equator to Antarctica), a wide range of observed parameters (physical, chemical, biological) and are obtained using a variety of platforms and technologies (e.g. ships, autonomous floats, gliders, moorings, satellites, animal tags, coastal radar. The end users include researchers, students, managers, policy makers, consultants, sailors and fishers. All data products contributed to the AODN are made freely and openly available to the public via the AODN Portal (https://portal.aodn.org.au/).

The AODN Portal provides access to the following different types of gridded dataset collections:

- Sea Surface Temperature (SST) products;
- Ocean Colour products;
- Satellite altimetry products;
- Coastal radar products;
- Climatology:
- Bathymetry.

Many of the gridded data collections are very large and therefore tools have been developed over the years to enable data products to be visualised through the AODN Portal as well as to be partitioned temporally and spatially to then be aggregated into a single file. Tools like netCDF Operator (NCO) or the netCDF Java library are widely used by the scientific community in order to subset and aggregate gridded data. The main issue for the AODN has been the management of jobs submitted by many different users through the AODN Portal. A single queue, no indication of the position of a particular job in the queue, loss of jobs during a restart of services and restriction on output files were some of the problems encountered over the past few years.

In 2016, the AODN relocated its services to the commercial cloud service Amazon Web Services (AWS). In addition to the generic advantages of cloud computing, AWS offers a growing array of complimentary services. These services include data durability, the automated deployment of auto scaling and load balancing architecture, which increase system reliability. One of the most recent service to be implemented in the AODN infrastructure is AWS Batch, a fully managed batch processing service.

AWS Batch dynamically provisions the optimal quantity and type of compute resources (e.g. CPU or memory optimised instances) based on the volume and specific resource requirements of the batch jobs submitted. AWS Batch offers many benefits like reduced operational complexities, time saving and reduced costs, the possibility to configure multiple queues and increased download size options.

An Open Geospatial Consortium (OGC) Web Processing Service (WPS) allows standardised requests and responses to the asynchronous processing service used to subset and aggregate gridded data into a single netCDF file for a variety of dataset collections.

AODN have also implemented customised dashboards in order to efficiently track jobs in real-time and to ensure rapid responses to queries that come to the AODN helpdesk. The content of data reports used by the AODN management team have been extended to include information, such as the most popular dataset collection, response time to provide download output, a confidential list of users and their affiliation.

eCUDO.pl - towards the Polish Oceanographic Data Committee

Marcin Wichorowski, Institute of Oceanology Polish Academy of Sciences (Poland), wichor@iopan.pl
Michał Piotrowski, Maritime Institute in Gdańsk (Poland), Michal.Piotrowski@im.gda.pl
Lena Szymanek, National Marine Fisheries Research Institute (Poland), lszymanek@mir.gdynia.pl
Urszula Pączek, Polish Geological Institute National Research Institute (Poland), upac@pig.gov.pl
Mirosława Ostrowska, Institute of Oceanology, Polish Academy of Sciences (Poland), ostra@iopan.pl
Michał Wójcik, Maritime Institute in Gdańsk (Poland), Michal.Wojcik@im.gda.pl

General overview

Institute of Oceanology, Polish Academy of Sciences, Institute of Meteorology and Water Management National Research Institute, Maritime Institute in Gdańsk, Polish Geological Institute National Research Institute, National Marine Fisheries Research Institute and University of Technology in Gdańsk - consortium of organizations engaged in research and exploitation of marine resources, established as POLMAR, for many years lead actions targeting harmonization, integration and coordinated provisioning of environmental data resources. Sharing of data is focusing on transfer of information and enabling access to detailed oceanographic data. This is fundamental activity for planning and performing of activities and investments in coastal and offshore areas. The consortium is leading activities towards deployment operational state of the system delivering these demanded data and products to the users. The consortium consists of most of the organizations involved in marine research and continuous acquisition of oceanographic data. For this reason the consortium is perspective Polish Oceanographic Data Committee.

System design

The project assumes obtaining data openness at the level of 5 stars according to the "5 Star Open Data" scale, because the project develops a system that gives users not only the ability to access current, open and searched data, but also integrates the resources of other open collections. The data will be described with metadata compliant with the standards applicable to geospatial data. The data will also be marked with unique URIs and in all appropriate places the metadata will contain relevant links to other resources available on the network.

The following design and architectural assumptions have been adopted:

- The central element of the platform is infrastructure based on a virtualization server connected to the data storage system. The use of such a solution allows for easy implementation of individual modules, as well as their replacement in the event of a need to conduct a complex software update or to remove errors in the operation of the system;
- The platform's infrastructure consists of a main server that provides required functionalities to external users, a single-signal server allowing one-time login to the platform and use of all its resources, spatial data server performing WMS, WFS and CSW services and a server processing data from non-federated external systems;

- The platform's environment (users, federated systems and non-federated external systems) can be distributed to any geographical location with communication via the Internet. For the sake of simplifying the model, it is assumed that each external system (both federated and un-federated) will consist of a server providing data, combined with a data warehouse;
- Users can access the system via any device (computer, laptop, tablet, mobile phone) with installed software supporting the display of websites using the http protocol or supporting WFS, WMS and CSW type services;
- All virtualized system components work on the CentOS operating system;
- All applications running under the virtualization server operate as part of the virtual Java machine provided by the Java runtime environment for the CentOS system. In the case of web applications and enterprise-class applications, they will be launched as part of the WildFly application server running in the Java virtual machine;
- The PostgreSQL server is used as the database server with the PostGIS extension providing support for geographic data processing (GDAL);
- The central element is eCUDOApp, an enterprise class application consisting of the EJB eCUDOEngine module and the eCUDOWeb web module communicating with each other using local EJB interfaces. All operations on the database will be performed from the level of the EJB module through the JDBC protocol (Java Data Base Connectivity);
- A GeoServer application embedded in a Jetty container running in a Java virtual machine is used as a GIS data server. Both the GeoServer application and the Jetty container are free software available under the Open Source license. The application through the eCUDOWeb module will issue WMS, WFS and CSW services in the Representational State Transfer (REST) architecture. As in the case of eCUDOEngine, communication with the database will be implemented through the JDBC protocol;
- The eCUDOLinkedDataParser application is launched directly in the Java virtual machine and in specified time intervals it will synchronize metadata from non-federated external systems using the http protocol (Hypertext Transfer Protocol) with metadata on the eCUDO platform via the eCUDOWeb module using the published services in REST and SOAP architecture (Simple Object Access Protocol). The application will be able to process data provided by external systems in the Linked Data format;
- The SSO server will contain the Keycloak web application acting as a central login point that allows the use of all platform elements after a single login. Keycloak is a free application available under the Open Source license. Login will be possible via the website by the http protocol as well as through authorization connected to REST and SOAP service calls.

The above elements, the CentOS system, the WildFly application server and the PostgreSQL database server together with the PostGIS extension, are free software available under the Open Source license. The data presented by the eCUDOWeb module will come from external systems, especially from federated systems. Each of the federated systems will already have an enterprise application for storing and sharing data. A module or a separate application serving as a proxy for the central eCUDO system should be prepared for each of these systems.

Summary

The eCUDO.pl project, when finished is intended to be well integrated with SeaDataCloud and EMODNet systems. Advanced serivces provided by clients (including data analysis services) extend availability of oceanographic data both to Polish and European organisations.

The integration platform of Roshydromet for data exchange within Russian and international projects

Alina Lazareva, Russian Research Institute of Hydrometeorological Information - WDC (Russia), alazareva@meteo.ru

Sergey Belov, Russian Research Institute of Hydrometeorological Information - WDC (Russia), belov@meteo.ru

In modern world every day we receive and process great amount of hydrometeorological information, obtained from different parts of the world for ensuring activity. Ocean data is used in different spheres of human activities. Many organizations handle collection, storage, accumulation, processing, and dissemination of information of global scale. Projects that are interested in these processes have similar but at the same time various options of registration and formalization of data and metadata of so called "information resources". It makes difficulties in the migration of information resources between systems. Information exchange can be done using local systems "adapters", providing transformation from one particular data format to another, but this is economically unprofitable, resource-intensive and simply inexpedient. For this reason mutual exchange of data between systems is almost not possible.

Creation of a platform for integration information resources can complement in solving interoperability problem between systems. There are several major problems of integration that such platform should solve: unification of structural presentation of information resources, actuality of information, monitoring of the data transmission as well access rights to information resources and services.

Such integration platform is an intermediate element in the process of data transferring between data source to destination system – end users, acting as some sort of "interpreter" between two interacting systems, originally speaking different "languages". Integration of heterogeneous information resources take place by using definite data and metadata formalization techniques.

All interaction in the Roshydromet data integration platform is based on the concept of "information resource". This concept describes metadata about the data as well as data source (database, file catalogues, web services, etc.). Comprehensive metadata descriptions helps to identify the location of the data source, access and retrieve data from the local data sources, transform to appropriate transport data format (point, profile or grid using netCDF). Roshydromet data integration platform has strict rule that metadata should be always accompanied by data. No incorporeal metadata descriptions allowed. Metadata is based on ISO-19115 standard, controlled code lists and Roshydromet unified parameter vocabulary and has transformation procedures to be exported as common ISO 19139 or WMO Core Profile metadata standards. Information resource can describe a complete data set (e.g. historical oceanographic cruises as a whole data store) or subsets (e.g. cruises in Arctic, Atlantic ocean, Pacific ocean, etc.) with multiple conditions, for instance, combined by vessel types, temporal extents, measurements, instruments, etc. Based on

such rules information resource is producing single or multiple instance descriptions that contains metadata for individual subset. Instance descriptions are derived from “parents” metadata records. Using metadata, data originators assign permissions - no conditions for data access and use or restricted access based on agreement model.

Integration platform contains a number of software components that provides web interfaces to administrators and operators in the system. It utilizes intelligent algorithms and set of rules to simplify metadata creation, update in automatic or semi-automatic mode based on data update frequency. Therefore metadata in Roshydromet integration platform is always adequate to data in term of spatial and temporal coverage.

At present moment developed integration platform directly or indirectly is used in national and international projects. Developed under aegis of ESIMO project (www.esimo.ru) it has been transferred to IODE OceanDataProject as in-kind contribution. Further it has been used for creation of Russian segment of WMO Information System, Informational and telecommunication network of Roshydromet for collecting geophysical data, state fund system of RIHMI-WDC. As a success story, ESIMO now has 36 institutions from 12 ministries and agencies connecting over 500 databases and data storages that producing 3300 information resources and around 30000 instances respectively. This integration platform is in new evolution phase, but even being a promising development in Roshydromet, it is not limited to hydrometeorology and can be used in other fields of human activity where there is a need to integrate various heterogeneous information.

Data driven Blue Growth - meeting ocean renewable energy sector needs with iMarDIS

Graham Worley, Bangor University (Wales), g.worley@bangor.ac.uk
David Mills, Bangor University (Wales), d.mills@bangor.ac.uk
Thomas Prebble, Bangor University (Wales), t.prebbles@bangor.ac.uk
Cathy Blakey, Bangor University (Wales), cathy.blakey @bangor.ac.uk
Gwyn Roberts, Bangor University (Wales), gwyn.roberts@bangor.ac.uk
Colin Jago, Bangor University (Wales), c.f.jago@bangor.ac.uk

Wales has significant ocean energy resources and the EU has recently committed €100m to drive Blue Growth by supporting the growing ocean renewables energy (ORE) sector in Wales. Funded by the European Regional Development Fund the SEACAMS2 project (http://www.seacams.ac.uk) undertakes a programme of collaborative research to support the ORE sector. SEACAMS2 specifically emphasises the de-risking of business decisions for tidal stream, tidal range and wave renewable energy companies in Wales. As part of SEACAMS2 a new integrated data management and information system (iMarDIS) is being developed that will secure existing data and employs state of the art systems for archiving data and making it available for re-use.

SEACAMS2 collaborative research requires observations that are multi-disciplinary, variable frequency in space (< 1m to > 10km) and time (seconds to years) and encompass intertidal, shallow coastal and offshore domains. Observations range from sea surface to seabed and sub-seabed using research vessels, moorings, seabed landers and UAVs. The resulting data sets are large, complex and require rapid processing for timely delivery to meet industry and other stakeholder needs. Furthermore, SEACAMS has a legal and contractual requirement to share the publicly funded data it has acquired.

The design of iMarDIS (www.imardis.org) is 'end user' driven. At the same time, it was necessary to ensure that iMarDIS: adopted existing UK MEDIN meta-data standards, could interface with the existing UK and European data management infrastructures, avoided replication of existing capability, was capable of ingesting and disseminating in real-time very high resolution data and could streamline the workflow from data collection to data management and curation.

A review of existing data services within the UK showed they would be unable meet the demands of SEACAMS end users either in terms of speed of data access, the high resolution of the data or in meeting our need to construct applications that would retrieve data on a machine-to-machine basis for further processing. Furthermore, it became apparent that there were other issues that prevented their adoption within the context of iMarDIS including difficulties in scaling the technology for concurrent access by many users or handling very large volumes of data and allowing new applications to be built upon the data archive.

Consequently, a concept for the overall iMarDIS infrastructure was developed that required new software that was cloud hosted. The system (Fig. 1) was conceived as a series of micro-services, each capable of operating semi-independently, and communicating with each other

through a RESTful JSON based Application Programming Interface layer (API). These APIs will eventually become public. There would be a data discovery and download portal and advanced services website communicating with a back-end set of micro-services. The micro-services would support a range of key functions including security and authentication, metadata management, file upload and download, tabular data manipulation, point cloud (or generic raster) data processing and time-series data manipulation.

Fig. 1 - An overview of the iMarDIS architecture showing the 'back end' microservices to the left of the Application Programming Interface and the applications domain on the right handside. The applications domain is where advanced service capability can be implemented either by iMarDIS or third party developers.

Breaking the problem down into micro-services allowed three developers to work in parallel on each service to reduce the time taken for infrastructure development. The architecture is based around Enterprise Java REST services implemented within the Amazon Web Services cloud. This is both scalable in terms of storage capacity and throughput.

An important feature of this design is that the iMarDIS hosted portal acts as a showcase for its capability in terms of data discovery and download and advanced services. Using published APIs external software developers could develop their own innovative applications reliant on the iMarDIS base infrastructure. This will encourage collaboration and innovation and promote sustainability of the infrastructure by increasing the diversity of its use.

To manage and collect metadata and upload data to the infrastructure a metadata manager application has been written and is in daily use. This also provides a level of orchestration of the interactions between the micro-services through the passing of messages to a series of queues. To date many of the back-end services have been written and are in an advanced state of testing. Work on the front-end application will be offered as an external contract in the near-term.

In parallel with the development of the micro-services a data librarian has been ingesting data and metadata into the system and around 4Tbytes of a potential 30Tbytes has now been catalogued. In order to facilitate data sharing a new licence agreement has been drawn up that allows commercial exploitation of the work but addresses the needs of researchers who ultimately are providing the data for SEACAMS.

As well as ensuring existing SEACAMS data are ingested and made available for re-use new in-situ data streams are beginning to be ingested and published to the iMarDIS website. For example, in response to user requirements a relocatable Coastal Observatory is being developed as part of the SEACAMS2 programme to improve understanding of key coastal processes. The first mooring of the network is currently being deployed in the Irish Sea and multi-variate data from the first of these buoys in already streaming live into the iMarDIS infrastructure.

Detailed knowledge of user requirements has been key to successful collaborative research programmes with industry. Building on this knowledge via further stakeholder engagement is critical to the design of iMarDIS. Consequently, a stakeholder workshop was recently held that brought together a cross-section of the community actively engaged in development of the ORE sector in Wales including commercial developers, consultants and regulators. The workshop report (https://www.imardis.org/workshop2018/Report-En.pdf) highlights the detailed outcomes from the successful workshop that are being used to inform the next stage of design of iMarDIS, further ensuring that the informatics infrastructure will meet stakeholder needs.

The next phase of development for iMarDIS includes development of the website that will act as the main presentation layer of the initiative, further enhancements of the micro-services, development of access libraries to aid end-users in connecting tools like MATLAB, R etc. directly to iMarDIS data feeds, as well as continued ingestion of SEACAMS legacy data and the publishing of discovery metadata to MEDIN and EMODnet.

webODV - a tool for the online analysis of marine data

Sebastian Mieruch, Alfred Wegener Institute (Germany), sebastian.mieruch@awi.de
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de

webODV – Ocean Data View online

Ocean Data View (ODV), Schlitzer R. 2002, is a widely used software package for the analysis, exploration and visualization of oceanographic and other environmental data. It plays a fundamental role in the SeaDataNet (SDN) community and is heavily used for data file and parameter aggregation as well as for data quality control by regional data coordinators. In the framework of the SeaDataCloud (SDC) project an online version of the ODV software is being developed called webODV. The online webODV tool will provide typical ODV functionality as a collection of modular web services. In the background, the modular web services make use of the full power of ODV running on the server side. The basic concept is to provide a user-friendly Browser interface which communicates with ODV on the server. On the server we run a special version of ODV that is equipped with a secure Websocket server and allows bi-directional communication with the client via encrypted Websocket messages.

webODV architecture

webODV will be implemented on EUDAT cloud infrastructure dedicated to SDC, and it is planned to develop a so called Virtual Research Environment (VRE), where users can work online on marine data. webODV will represent a major part of this VRE. As shown in Fig. 1 the main concept is to have a state of the art Browser interface offering different modular services, like data extraction, quality control and visualisation. These services communicate via the fast (and secure) Websocket protocol with the ODV instances on the server, which are responsible for the actual data processing.

webODV online quality control

As a core webODV service online data quality control of marine data will be implemented in the SDC infrastructure. Figure 2 shows a prototype of the online quality control service.

Data quality control is a fundamental requirement of any data system. Within SDC regional experts perform quality control on ocean profile data in regular intervals using the desktop ODV software. Here we present the first webODV prototype for online quality control. Users are able to perform visual quality control by inspecting scatterplots of specific variables like temperature and/or salinity. By clicking on single samples within a scatterplot, information about that sample and measurement station are provided. Quality control experts can zoom into the plots, flag single samples, entire stations, all stations visible in the zoom window or all samples in the collection.

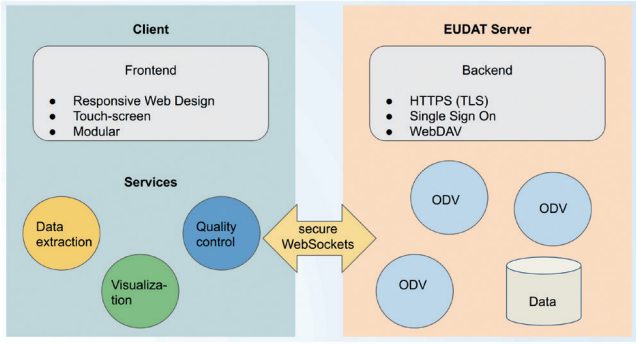


Fig. 1 - Concept.

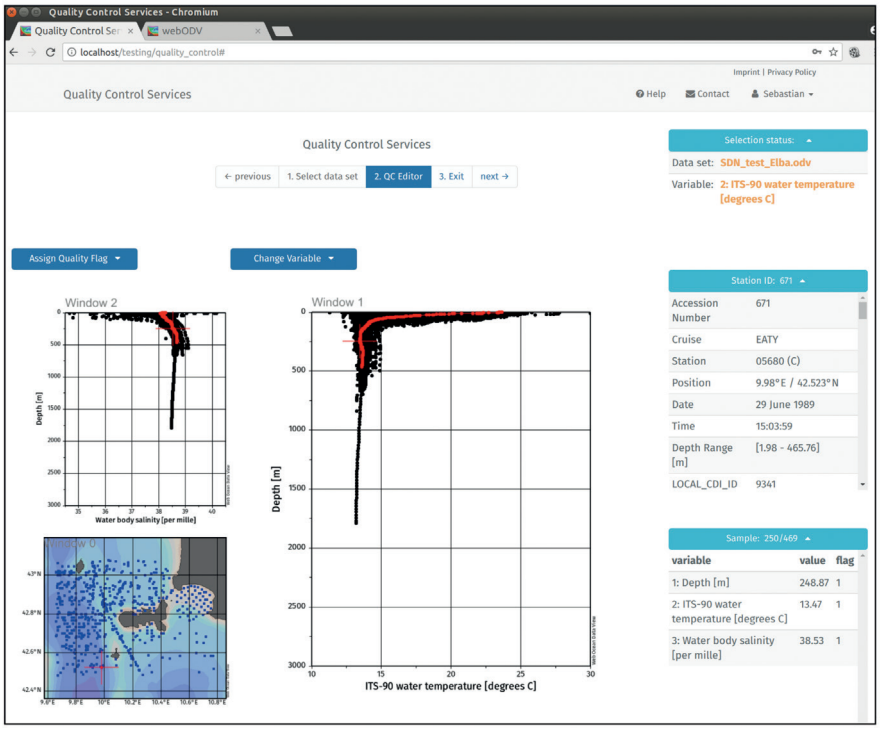


Fig. 2 - Quality Control service.

webODV other services

Another important webODV service developed for SDC is data extraction, allowing selection by cruise names, geographical domain as well as time windows. An operational example of a webODV data extraction service is available at https://webodv.awi.de.

Other planned webODV services comprise data import, data aggregation as well as data visualisation.

References

SCHLITZER R., *Interactive analysis and visualization of geoscience data with ocean data view*, Computers & geosciences 28 (2002) 1211-1218.

The BODC Parameter Usage Vocabulary (PUV) semantic model exposed

Gwenaëlle Moncoiffé, British Oceanographic Data Centre (United Kingdom), gmon@bodc.ac.uk
Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk

The BODC Parameter Usage Vocabulary started its existence in the 1980s as a simple list of terms. It has since evolved to constitute a fully structured controlled vocabulary schema, a meta-vocabulary whose building blocks are themselves controlled vocabularies. The logical semantic model that governs the construction of its text label enables us to express a complex array of scientific information related to the observation, in 8-byte codes. The BODB PUV has been published as a linked data SKOS vocabulary that makes it understood by both humans and machines but the underlying semantic model was not exposed, limiting its advantages to a closed community . In this presentation, we are attempting to introduce the different elements of the semantic model and publish it as linked data by describing it with the relevant ontology elements (properties and classes).

The three elements at the core of the semantic model are: the parameter entity, the object of interest and the matrix. The parameter entity can be modulated to be the property of interest itself or any statistical parameters derived from the statistical analysis of the property. The object of interest can be either a biological entity, a chemical entity or a physical object or phenomenon. Both these fields (the parameter entity and the object of interest) are mandatory for the semantic model The third element, which is the matrix, is the environment in which the object of interest is embedded or to which the property relates (e.g. a phase). This can be switched on and off depending on the type of parameter. It will be set to "not applicable" for example for many engineering parameters that do not relate to a measurement matrix. The matrix itself consists of four semantic elements, enabling us to define whether the measurement is applied on the whole matrix or on any of its sub-components or phases for example, analysis on a filtered sample or analysis on samples filtered through different filter sizes (size-fractionation). Additionally the model provides 3 optional fields to encompass key elements of the methodology at the sampling, analytical and data processing stage, as necessary.

Expressing the semantic model as linked data enables data aggregation but also integration of data across diverse disciplines and domains, where parts of these compound terminologies can be mapped to each other, boosting the semantic discovery.

Exposing the SeaDataNet metadata catalogues via SPARQL endpoints

Chris Wood, British Oceanographic Data Centre (United Kingdom), c.c.wood@gmail.com
Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk
Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie
Rob Thomas, Marine Institute (Ireland), rob.thomas@marine.ie

Many scientific disciplines have metadata catalogues managed by a designated responsible organisation. The content of these catalogues can vary by discipline, but may contain information about institutions who contribute data to central data repositories, projects that have been carried out within a particular field, or individual datasets. Within the European oceanographic science community, metadata catalogues have been managed under the SeaDataNet infrastructure. The five catalogues (EDMED: for datasets, EDMO: for organisations, EDMERP: for projects, EDIOS: for observing systems, and CSR: to describe cruise summary reports) have long been available online through individual web based search interfaces via the organisation that hosts the catalogue. However, such interfaces implicitly limit how the queries can be conducted, and how results can be viewed. Such limitations can be removed through the development of Application Programming Interfaces (APIs).

We have led the API development for all five catalogues, and have taken a Linked Data approach to the implementation. This approach requires the catalogues to be stored in triplestores, a form of graph database, with querying available using SPARQL, a query language for triplestores analogous to SQL for relational databases. The query interface is publicly available over HTTP, allowing the whole catalogue to be openly queryable. The true strength of the triplestore approach is the ability to conduct federated queries across the different SPARQL endpoints which simulates a join between tables in a traditional table-based database, but without the need for subsets of the data to be located within the same database. A sixth triplestore, supporting the NERC Vocabulary Server (NVS), underpins the content of the catalogues by providing lists of standardised terms that provide consistency and semantic harmonization across the catalogues. Federated queries can be used to determine the theme of a dataset, described in NVS terms, by simultaneously querying the SeaDataNet dataset triplestore and the NVS triplestore.

In this presentation, we will show the advantages that can be gained by both the catalogue publisher and the end user, the steps needed to setup both a new triplestore and the corresponding SPARQL endpoint and the ease by which this software stack can be implemented, as well as the options available to the publisher. We will finish with the lessons learnt in this project, and the future work that will be carried out to further enhance the service to the users of the SeaDataNet infrastructure.

The SeaDataCloud Virtual Research Environment: researching the sea from the cloud

Merret Buurman, Deutsches Klimarechenzentrum GmbH (Germany), buurman@dkrz.de
Peter Thijsse, Mariene Informatie Service 'MARIS' Bv (The Netherlands), peter@maris.nl
Sri Harsha Vathsavayi, Tieteen Tietotekniikan Keskus OY (Finland), sriharsha.vathsavayi@csc.fi
Sebastian Mieruch, Alfred Wegener Institut (Germany), sebastian.mieruch@awi.de
Gael Leblan, Institut Français de Recherche pour l'Exploitation de la Mer (France), Gael.Leblan@ifremer.fr
Giorgio Santinelli, Deltares (The Netherlands), Giorgio.Santinelli@deltares.nl
Alexander Barth, University of Liege (Belgium), a.barth@uliege.be , *et al.*

The SeaDataNet project offers a robust and state-of-the-art Pan-European infrastructure to harmonise metadata and data from marine data centres in Europe, and offers the technology to make these data accessible. The user was able to use the SeaDataNet infrastructure to download the data to local servers or local machines and work from there with the data there. As marine observation data continues to increase in size and number of datasets, transferring it over the network and processing it requires more and more efficient machines and network bandwidth. Downloading a large amount of large data sets, processing it on a laptop, and uploading the results to send them to colleagues, then receiving feedback and re-initiating the entire process, is cumbersome.

To make life easier for users and data holders, as part of the SeaDataCloud project, SeaDataNet is moving its unrestricted data to the cloud. Keeping the data centrally in high-performance data centres includes some other advantages than ease of access and download: The centralized ingestion process allows for standardized quality checks to be performed on the data before it is made available to users. Corrupt files, non-conforming formats, and duplicate observations which could lead to bias/artefacts in analyses, can be more easily detected. Loss of data is precluded by storing the data in several locations across Europe. Five major EUDAT data centres will form a cloud (DKRZ in Germany, CINECA in Italy, CSC in Finland, GRNET in Greece, and STFC in the United Kingdom) that hosts copies of the SeaDataNet datasets for highly available download.

But why only move the stored data to the cloud? Another aspect of working with large amounts of data is the resources needed for processing them. In general, the current trend is less downloading, more processing where the data is. While marine observation data continues to increase in size and number of datasets, transferring it over the network and processing it requires more and more efficient machines and network bandwidth. Downloading a large amount of large data sets, processing it on a laptop, and uploading the results to send them to colleagues, then receiving feedback and re-initiating the entire process, is cumbersome. And then, after all analyses are run (phew!), you see that a new version of the dataset had been published, correcting some important shortcoming!

To tackle these problems, the creation of the SeaDataNet Virtual Research Environment (VRE) is being developed in close collaboration of SeaDataNet developers and the 5 partners of

European Research Data Infrastructure EUDAT.

Virtual Research environments are web-based workspaces providing seamless access to all services researchers need to do their work and collaborate with their community. For SeaDataNet, this means that all the tasks that a researcher would usually do with the data, the entire workflow of data-driven science - finding data, accessing data, processing iteratively data with various tools, visualizing results, sharing results with colleagues, and publishing data - can be realized without having to download data to the desktop and using the local compute power, which might not be available to all users at the same rate.

The first use case that is being addressed is that of the SeaDataNet expert groups creating temperature and salinity climatologies products. Data from many cruises/stations/projects is aggregated to a global dataset, which is being reviewed manually to ensure its scientific quality, using the webODV tool. Then, geospatial interpolation algorithms are used to create a continuous field of temperature and salinity using the DIVA tool (Barth *et al*. 2014). These climatologies, separated by basin, are available on the Sextant catalogue for SDN products https://www.seadatanet.org/Products.

As part of the first VRE release this use case will be integrated and tested by the user groups. This use case represents a core workflow in the VRE and will be used as a template for enabling more use cases (e.g. EMODNET Bathymetry DTM processing, EMODNET Chemistry, bio-geochemistry QC). Of course the scientists are not restricted to predefined workflows, but encouraged to use and recombine the VRE's functionalities for custom analysis workflows.

The most common tools that the SeaDataNet community is using, notably webODV (Ocean Data View - online) and DIVA using Jupyter Notebook, are already included in the first prototype. Thanks to the extensible and extendable architecture (described in a second abstract), many more are to come. To serve expert users as well as more applied users, most services will come with an easy-to-use GUI as well as a command line interface. On top of this, a chat/forum-like communication channel for scientists will encourage more collaborative work style. A notifier mechanism, that warns users of new versions of data they are using, will prevent the use of outdated datasets. Finally, chaining several of the VRE's processing tools to an entire workflow that can easily be documented, rerun and reproduced is planned. In addition to the SDN CDI data from the SDN data cloud, other datasets relevant to the marine community will be available, for example SDN products and EMODNet data products.

As a sum-up, we can say that beyond easing the problem of resources, several other advantages are achieved by the VRE: The use of outdated data and software is prevented, slow and unnecessary downloads are reduced and jam-packed hard drives are avoided. Processing of data gets faster and multiple tasks can easily be run in parallel. Sharing input data, intermediate results or end results becomes easier. All this, facilitates collaborative science in a scientific world that relies more and more on team work, by internationally distributed teams, and open science.

Stay tuned and check out https://vre.seadatanet.org for updates on the progress of SeaDataNet's VRE!

References

Barth, A., Beckers, J.-M., Troupin, C., Alvera-Azcárate, A., and Vandenbulcke, L.: *divand-1.0: n-dimensional variational data analysis for ocean observations,* Geosci. Model Dev., 7, 225-241, doi:10.5194/gmd-7-225-2014, 2014.

EMODnet Central Portal data services

Paula Oset, Vlaams Instituut voor de Zee VZLU (Belgium), paula.oset.garcia@vliz.be
Simon Claus, Vlaams Instituut voor de Zee VZLU (Belgium), simon.claus@vliz.be
Francisco Hernandez, Vlaams Instituut voor de Zee VZLU (Belgium), francisco.hernandez@vliz.be
Bart Vanhoorne, Vlaams Instituut voor de Zee VZLU (Belgium), bart.vanhoorne@vliz.be
Filip Waumans, Vlaams Instituut voor de Zee VZLU (Belgium), filip.waumans@vliz.be
Jan-Bart Calewaert, EMODnet Secretariat (Belgium), janbart.calewaert@emodnet.eu
Pascal Derycke, EMODnet Secretariat (Belgium), pascal.derycke@emodnet.eu
Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@ogs.trieste.it
Helen Lillis, Joint Nature Conservation Committee (United Kingdom), helen.lillis@jncc.gov.uk
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com
Alessandro Pititto, Cogea S.r.l. (Italy), apititto@cogea.it
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
Henry Vallius, Geological Survey of Finland (Finland), henry.vallius@gtk.fi
Thierry Schmitt, Service Hydrographique et Océanographique de la Marine (France), thierry.schmitt@shom.fr

The European Marine Observation and Data Network (EMODnet) is a network of more than 160 organisations supported by the EU's integrated maritime policy. These organisations work together to assemble marine data and data products to make these resources freely available across seven different thematic portals: Bathymetry, Geology, Seabed habitats, Chemistry, Biology, Physics and Human Activities.

The EMODnet Central Portal centralises information about the EMODnet project and offers a single entry point for metadata, data and data products that are being made available by the individual thematic portals. The data services offered by EMODnet Central Portal comprise a geoviewer, a metadata catalogue, a query tool and documentation on how to access data using web services.

The map viewer of the Central Portal allows to display the data products developed by the thematic lots. It currently gives access to over 40 different data products, together with additional data layers and administrative units, all of them based in OGC web services. Each of the layers in the geoviewer is accompanied by detailed descriptions of the products, with links to the associated metadata records. Depending on the nature of the data product, the viewer allows to animate or apply filters on the layers. The EMODnet Central portal geoviewer is periodically updated as new data products are developed by the thematic lots.

The Central Portal query tool provides users with the possibility to query the data products elaborated by the different thematic projects and to obtain combined output (a table), with values for selected parameters and for a specific window in time and space. The query tool will be re-designed in the coming months to fully exploit the capabilities of the EMODnet thematic web

services. The concept and user requirements will be fine-tuned in collaboration with potential data users from the Marine Knowledge Expert Group.

The Central Portal hosts a Geonetwork-based metadata catalogue where the data products created by each thematic portal are described. In preparation for the Open Sea Lab competition (November 2017), the Central Portal compiled documentation and examples on how to access the different EMODnet web services.

The experience of the Open Sea Lab competition enabled the EMODnet network to fully realise the potential of the data services that are currently available, but it also revealed shortcomings that need to be tackled. In close cooperation with each of the thematic lots and the EMODnet Secretariat, the major goal of Central Portal will be to implement best practices to improve the data services inter-operabilty and improve user experience and uptake. This will be achieved by working in four main subjects: a) ensuring the INSPIRE compliance of data products metadata; b) making the EMODnet web services fully OGC compliant; c) providing data services documentation in a standardised approach, using community accepted repositories and making it accessible via metadata; and, d) building the above-mentioned re-design of the query tool.

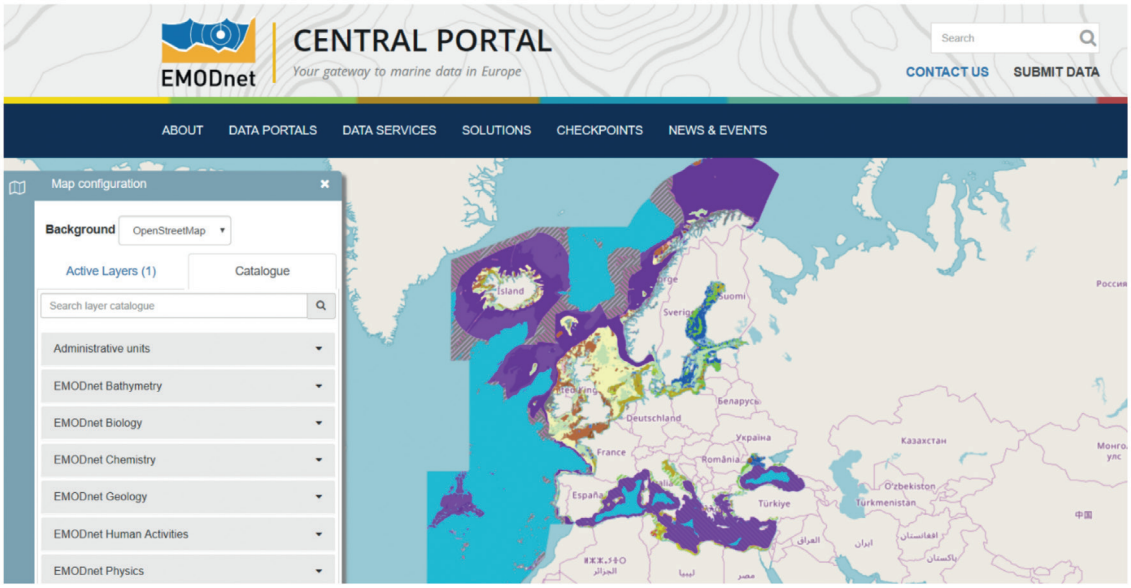


Fig. 1 - EMODnet Central Portal geoviewer displaying a data product from EMODnet Seabed Habitats.

EMODnet Physics: a horizontal platform serving blue growth

Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com
Paolo D'Angelo, ETT Solutions Ltd (Italy), paolo.dangelo@ettsolutions.com
Patrick Gorringe, EuroGOOS/Swedish Meteorological and Hydrological Institute (Sweden), Patrick.Gorringe@smhi.se
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
Peter Thijsse, Mariene Informatie Service 'MARIS' Bv (The Netherlands), peter@maris.nl
Sylvie Pouliquen, Institut Français de Recherche pour l'Exploitation de la Mer (France), Sylvie.Pouliquen@ifremer.fr

EMODnet - the European Marine Observation and Data network – is a long term marine data initiative from the European Commission Directorate-General for Maritime Affairs and Fisheries (DG MARE) involving and networking more than 150 organizations for assembling marine data, products, and metadata. The data infrastructure has been developed through a stepwise approach in 3 major phases by running 7 thematic portals, 6 regional check points and a Data Ingestion facility.

EMODnet Physics (www.emodnet-physics.eu) is a domain specific portal of portals aggregating data and metadata from several data portals. The concept of the portal is a federation, intended as 'alliance' (federation from latin foedus = alliance). This means that there is a mutual agreement between EMODnet Physics and the data providers, each contribution being visible in the portal. Interoperability is a key issue of the federated system: common vocabularies, compliance with ISO, OGC standards and adherence to INSPIRE Directive build coherent services for users, although individual components are technically different and managed by different organizations.

EMODnet Physics is developing a combined array of services and functionalities such as facility for viewing and downloading, dashboard reporting and machine-to-machine communication services, to obtain, free of charge data, meta-data and data products on the physical conditions of the ocean from many different distributed data sets (www.emodnet-physics.eu/map).

EMODnet Physics is providing Regional stakeholders and international networks with tools to serve their users and communities, e.g. Physics is powering and hosting the South Ocean Observing System (SOOS) data portal (http://www.soos.aq/data/soosmap) and SOOS is helping Physics to unlock and make available more valuable data, Euskoos – the Basque Operational Oceanography System enriched its data portal and data dissemination with the EMODnet Physics widgets (http://www.euskoos.eus/radar-eu/).

The acquisition of physical parameters is largely an automated process that allows the dissemination of near real time information. In particular EMODnet Physics is a stock-share portal strongly federated to the Copernicus Marine Environment Monitoring Service In Situ Thematic Assembly Center. Historical validated datasets are organized in collaboration with SeaDataNet and its network of National Oceanographic Data Centers.

The EMODnet Physics portal is currently providing easy access to data and products of: wave height and period; temperature and salinity of the water column; wind speed and direction; horizontal velocity of the water column; light attenuation; sea ice coverage and sea level trends (relative and absolute).

EMODnet Physics is continuously increasing the number and type of platforms in the system by unlocking and providing high quality data from a growing network. Lately EMODnet Physics started working on river runoff data, total suspended matter and underwater noise (acoustic pollution).

EMODnet Data Ingestion: 'Wake up your data'

Dick Schaap, Mariene Informatie Service (The Netherlands), dick@maris.nl
Sissy Iona, Hellenic Centre for Marine Research (Greece), sissy@hnodc.hcmr.gr
Francis Strobbe, Royal Belgian Institute of Natural Sciences (Belgium), fstrobbe@naturalsciences.be

The 'EMODnet Ingestion and safe-keeping of marine data' project, started mid-2016, 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 freely distribution and publication through EMODnet. The EMODnet Data Ingestion portal facilitates submission of their sleeping marine datasets for further processing, Open Data publishing and contributing to applications for society.

The activities are undertaken by a large European network that is geographically anchored in the countries bordering all European marine basins, and covers all EMODnet data themes. The EMODnet Data Ingestion members are national and regional marine and oceanographic data repositories and data management experts. The coordinators of the EMODnet thematic portals are also part of this new initiative. Moreover the data centres work together on pan-European and international scales in organisations such as IODE, ICES, EuroGeoSurveys, EuroGOOS, and IHO, and for pan-European marine data management infrastructures such as SeaDataCloud, EurOBIS and EGDI. The latter are feeding into several EMODnet thematic portals.

The emphasis of activities in the first year has been put towards developing the EMODnet Data Ingestion portal and its services for ingesting and publishing data sets, developing the pathways for processing and elaborating of data submissions, laying a basis for promotion and marketing activities, and making an initial inventory of potential data sources and their providers.

The EMODnet Data Ingestion portal been launched early February 2017. It encourages data providers to share marine data, gives marine data management guidance information, and provides a range of services such as:

- submission service for easy ingestion of marine data packages
- view submissions service to oversee submitted data sets 'as is'
- data wanted service to post requests for specific data types

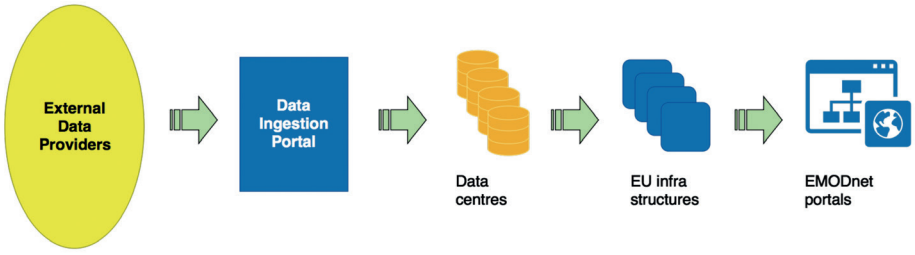


Fig. 1 - Flow of data sets from external data providers to EMODnet portals.

Submission forms with data packages are assigned to qualified data centres depending on the country of the data provider and the type of EMODnet theme. This group includes not only the EMODnet Ingestion consortium but also the groups of data centres who are involved in each of the EMODnet Thematic portals.

A distinction is made between 2 phases in the life cycle of a data submission:

- **Phase I:** from submission to publishing of the submitted datasets package '*as is*'
- **Phase II:** further elaboration of the data sets and integration (of subsets) in national, European and EMODnet thematic portals.

This split allows to publish already in an early stage the original data package with high quality metadata.

For operational oceanography a close cooperation takes place with EMODnet Physics. This aims at identifying and arranging inclusion of additional stations for Near Real Time (NRT) data exchange. The Data Ingestion portal explains how the NRT exchange is organised with EuroGOOS – Copernicus and guidance how to connect in practice. Furthermore a Sensor Web Enablement (SWE) pilot is set-up for Real Time data exchange. A client service to locate stations and to retrieve data streams in a time series viewer is hosted at the EMODnet Physics portal and 'advertised' at the EMODnet Data Ingestion portal.

Fig. 2 - Poster in the Brussels metro.

Promotion and outreach activities are equally important as technical developments. In the first year it has focused on establishing cooperation and synergy within the EMODnet community. A portfolio of promotional items has been developed, such as leaflets, posters, presentations, stickers, and a wonderful animation. These are part of the promotion and marketing strategy that was designed to reach out to potential data providers. In the second year this plan has been put into motion on full scale for a wider outreach and marketing to potential data providers in government, science and industry. This has so far resulted in many submissions and also in development of special use cases, such as for monitoring data from offshore renewable energy projects or minting DOIs for research data to support data citing for data submitters.

The Deep Sea Spy system: building a marine images annotation database from participative science

Catherine Borremans, Institut Français de Recherche pour l'Exploitation de la Mer (France), catherine.borremans@ifremer.fr

Marjolaine Matabos, Institut Français de Recherche pour l'Exploitation de la Mer (France), marjolaine.matabos@ifremer.fr

Introduction

Most of the current knowledge of deep-sea environments is based on punctual, at best yearly, oceanographic cruises. Since 2006, deep-sea ecosystems are continuously being monitored using video cameras deployed on deep-sea platforms. The acquisition of high-frequency video data from deep-sea observatories like EMSO-Azores (www.emso-fr.org/fr/EMSO-Azores) or Ocean networks Canada/NEPTUNE Canada (NorthEast Pacific Time-series Underwater NEtwork) provide for the first time information on species behaviour, feeding habits, growth, possibly reproduction and organisms' response to changes in environmental conditions. Cameras deployed on those observatories acquire hourly video data representing thousands of hours and Tera Bytes of footage that require 10 times more hours of viewing to extract useful information. Since their first deployment in 2006, more than 5 Tb of video data from both the Atlantic and the Pacific oceans were acquired that cannot possibly be analyzed by a few researchers. Only with the help of citizen scientists will be able to process the huge archive of imagery.

The annotation system

The main objective of the project was to build a web-based application for manual imagery processing that will help gather information of interest for scientists as well as raise awareness among the general public about deep-sea ecosystems.

In order to meet this goal the specific technical objectives were to:

- Develop an online image annotation program that will allow participants to simultaneously perform defined tasks on the extracted images.
- Organise the output information in a searchable database consistent with existing Ifremer and EMSO databases following ENVRI + standards.

Ifremer, in association with the company Noveltis (Labège, France), developed the web-based application linked with its structured database. This system was named «Deep Sea Spy» (www.deepseaspy.ifremer.fr).

The software is built as a game with dedicated missions. The goal of each mission is to annotate a series of images extracted from archived video sequences acquired with deep-sea observatories. The data obtained is related to the image (e.g. origin, date and position of acquisition, camera type), to the participant (e.g. age, country) and to the annotation (e.g. date, taxon, position/measurement/area in pixels). All this information is stored and exported in a central Oracle database compatible with international common vocabularies.

Fig. 1 - The Deep Sea Spy game annotation interface.

The query system

In order to achieve the final objective that is to answer scientific questions thanks to the data extracted from images researchers need a tool allowing data mining in the entire dataset collected through the annotation application. The «Deep Sea Spy Request» web interface was then developed to fulfil their requirements in terms of data selection by game mission, observatory, species and temporal criteria. This last parameter is particularly complex but crucial for dynamic studies that are based on observatories video acquisition. After query users can export annotation data results in a table suitable for further analyses. Ultimately the application will allow to perform quality checks such as outlier detection.

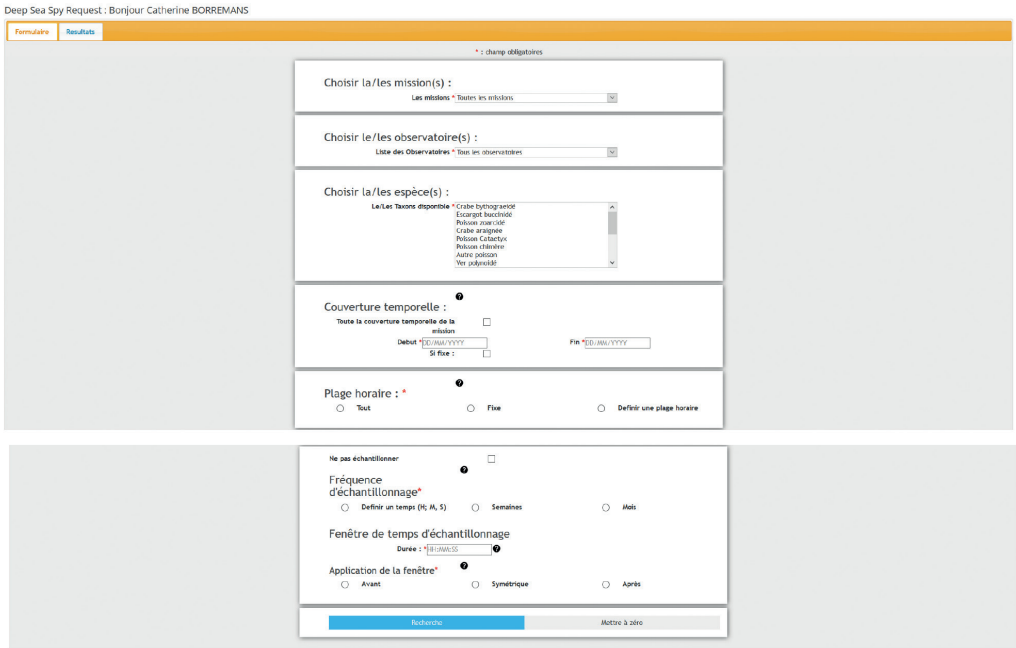


Fig. 2 - The Deep Sea Spy Request interface.

Perspectives

In addition to the challenge of processing such a big citizen dataset another approach will be to implement deep learning algorithms: annotations will help train computer programs for the automatic detection of animal species in the image (Kuminski *et al*., 2014). It is noteworthy that this would only be possible thanks to standardization of images data in one common database schema.

References

Kuminski E, George J, Wallin J, Shamir L. 2014. *Combining Human and Machine Learning for Morphological Analysis of Galaxy Images.* Publ Astron Soc Pacific [Internet]. [cited 2015 Dec 15]; 126:959–967. Available from: http://www.jstor.org/stable/info/10.1086/678977.

Cloud-based national on-line services to annotate and analyse underwater imagery

Roger Proctor, Integrated Marine Observing System (Australia), roger.proctor@utas.edu.au
Tim Langlois, School of Biological Sciences, University of Western Australia (Australia), timothy.langlois@uwa.edu.au
Ariell Friedman, Greybits Engineering (Australia), Ariell Friedman ariell@greybits.com.au
Sebastien Mancini, Integrated Marine Observing System (Australia), sebastien.mancini@utas.edu.au
Xavier Hoenner, Integrated Marine Observing System (Australia), xavier.hoenner@utas.edu.au
Brendan Davey, University of Tasmania (Australia), brendan.davey@utas.edu.au

Fish image annotation data and benthic habitat imagery is currently collected by various research, management and academic institutions globally (+100,000's hours of deployments) with varying degrees of standardisation and limited formal collaboration or data synthesis.

Two developing software tools have been brought together in the Australian Research Data Cloud to provide marine biologists with a powerful service for image annotation. SQUIDLE+ (http://squidle.org/) is an online platform designed for exploration, management and annotation of georeferenced images & video data. It provides a flexible annotation framework allowing users to work with their preferred annotation schemes. We have used *SQUIDLE+* to sample the habitat composition and complexity of images of the benthos collected using autonomous underwater vehicles (AUV) and stereo-Baited Remote Underwater Video (BRUV). *GlobalArchive* (http://globalarchive.org/) is designed to be a centralised repository of aquatic ecological survey data with design principles including ease of use, secure user access, flexible data import, and the collection of any sampling and image analysis information. To easily share and synthesise data we have implemented data sharing protocols, including Open Data and synthesis Collaborations, and a spatial map to explore global datasets and filter to create a synthesis.

We present a case study of how these national on-line services, developed within a domain-oriented research cloud, are being used to a) annotate habitat images collected using AUVs and b) synthesise fish annotation data sets from baited remote underwater stereo-BRUVs. These outputs are integrated into GlobalArchive and linked to an R virtual desktop analysis suite to offer an unprecedented capability to deliver marine biodiversity information of value to marine managers and scientists alike.

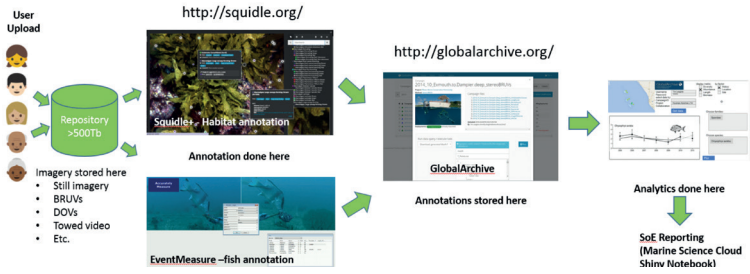


Fig. 1 - Cloud-based national on-line services to annotate and analyse fish and benthic habitat imagery.

MBARI media management - tools for turning video into data

Brian Schlining, Monterey Bay Aquarium Research Institute (United States of America), brian@mbari.org

Nancy Jacobsen Stout, Monterey Bay Aquarium Research Institute (United States of America), jana@mbari.org

Linda Kuhnz, Monterey Bay Aquarium Research Institute (United States of America), linda@mbari.org

Lonny Lundsten, Monterey Bay Aquarium Research Institute (United States of America), lonny@mbari.org

Kyra Schlining, Monterey Bay Aquarium Research Institute (United States of America), schlin@mbari.org

Susan von Thun, Monterey Bay Aquarium Research Institute (United States of America), svonthun@mbari.org

The Monterey Bay Aquarium Research Institute's (MBARI) Video Annotation and Reference System (VARS) has been used for annotating deep-sea video and images since 2004. The central tenants of VARS--using a controlled vocabulary and a centralized archive of annotations--have proven extremely effective for extracting quantitative and qualitative information from images and video. The VARS observation database has contributed to over 400 peer-reviewed publications.

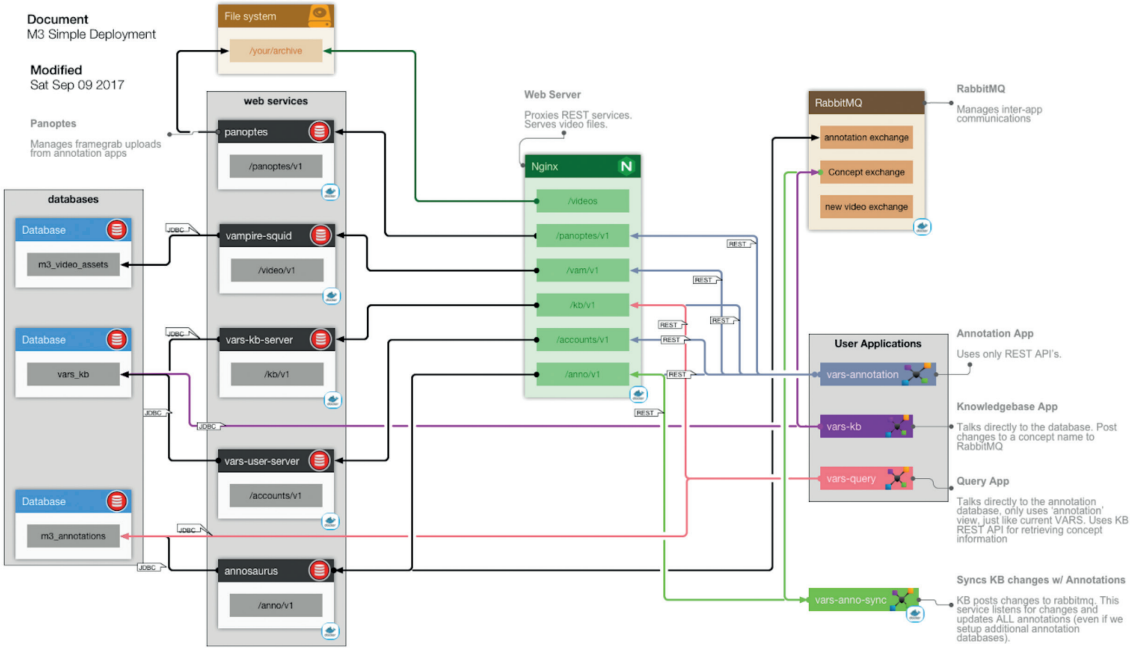


Fig. 1 - The MBARI Video Annotation and Reference System.

MBARI recently transitioned from videotape recordings to file-based formats. As part of this transition, MBARI has developed a toolkit of microservices for building video and image annotation applications. Each microservice communicates using standard web protocols and is relatively small and independently deployable. The design of the toolkit allows for the evolution and replacement of individual services over time as technologies and science requirements evolve and gives researchers tremendous flexibility in developing custom annotation applications. MBARI has created a suite of VARS applications on top of these services. These applications and microservices are open-source and available to external researchers for managing video assets and developing new, custom video analysis applications.

POSTERS

Enhancing ARGO floats data re-usability

Gianpaolo Coro, ISTI Consiglio Nazionale delle Ricerche (Italy), gianpaolo.coro@isti.cnr.it
Paolo Scarponi, ISTI Consiglio Nazionale delle Ricerche (Italy), scarponi@isti.cnr.it
Pasquale Pagano, ISTI Consiglio Nazionale delle Ricerche (Italy), pasquale.pagano@isti.cnr.it

Many research communities in a great variety of fields are interested in accessing collections of reliable environmental data. These data are typically used in environmental monitoring systems, data processing workflows, ecological models, societal and economical analyses, etc. Research communities need to carry out their studies in a fast and efficient manner and thus require data to be well structured, well described, and possibly represented in standard formats that allow direct access and usage. In this context, reducing data preparation and pre-processing time is crucial.

ARGO data have been long-used by marine science communities in global oceans observing systems. These data are collected using a large network of floats, monitored by the ARGO Information Center (AIC) and are sent to Global Data Assembly Centers (GDACs). The datasets are available for download on the official ARGO website (www.argo.ucsd.edu), as Network Common Data Format (NetCDF) *Point- feature* files and CSV files through FTP sites and online tools. However, these formats present many challenges from a technical point of view, especially in terms of re-usability. Every dataset has dimension ranging from 5MB to 3GB and contains measurements in time of different physical parameters recorded at different locations. Every file corresponds to one month and the overall repository time-span ranges from January 1999 to today. An overall CSV repository is available (ftp://ftp.ifremer.fr/ifremer/coriolis/co0547-bigdata-archive/) where a JSON file stores metadata about the parameters, e.g. the unit of measure, the full name, the reliability of the measurement, etc. Although accessing this unique endpoint is convenient, every dataset is not a standalone object and requires continuously parsing the JSON file to be fully understood. Further, managing a 3GB CSV file can be memory demanding, especially for processes that need to combine this dataset with other data.

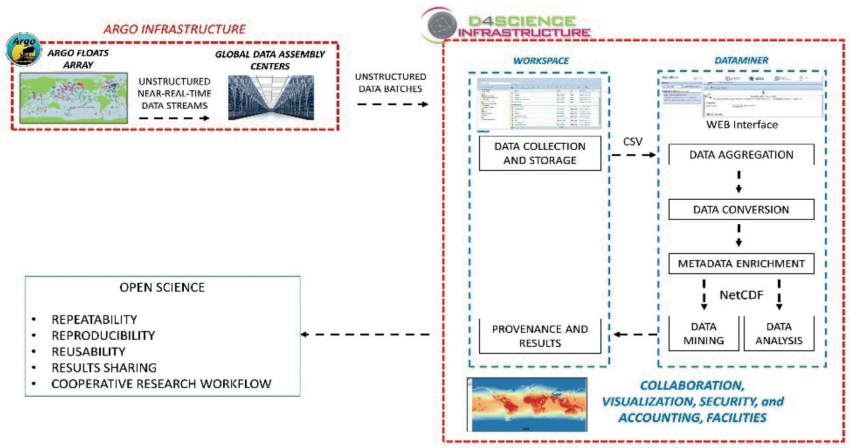


Fig. 1 - Conceptual schema of the ARGO-to-NetCDF conversion workflow.

In this paper, we present a workflow to convert ARGO observation data into a standard raster file. This workflow has been implemented in the context of a research e-Infrastructure with the aim to enhance the structure and re-usability of the ARGO data. We use an Open Science approach where all the standardized data are published in a Virtual Research Environment along with standardized metadata. The same conversion workflow is available as a Web service respecting the standard OGC Web Processing Service *(WPS)* and keeps track of the *provenance* of the data conversion process that allows reconstructing the processing history. This service was developed both to process all the historical ARGO data and to convert them as soon as new data are available. Our workflow transforms the ARGO unstructured data into NetCDF *Grid-feature* files. NetCDF is a self-describing, machine-independent data format meant to represent and store array-oriented data. It allows including information about the data in the file itself as attributes, and thus allows creating complex objects that do not need any external reference to be fully understood and thus re-usable and portable. The NetCDF Grid format assigns environmental variables values to a coordinate system with defined resolutions for longitude, latitude, depth/altitude, and time. Differently from the Point format, Grid is a raster format widely used by many communities and research institutions because many models/tools/libraries written in a large variety of programming languages are natively able to visualize, manipulate, and process this format. Our transformation workflow was implemented in R, thanks to availability of libraries to easily manipulate NetCDF files (e.g. ncdf4), and goes through the following steps:

1. Load a monthly-observation dataset in memory using R-specific importing functions optimized for large tables;
2. Represent metadata information, i.e. variable names and descriptions, units of measure, time instants, global parameters etc., in compliance with the Climate and Forecast (CF) standard vocabulary;
3. Represent all depth variables values in meters, using different conversion sub-routines depending on the original unit measure (e.g. pressures in dbar);
4. Generate a 3D grid with 10 logarithmic-divided depth intervals and 0.5° longitude-latitude resolution grids associated to each depth interval;
5. Clamp observation values to this 3D grid and associate values averages to each cell;
6. Create one NetCDF-CF file for every clamped variable.

This workflow is data-, memory-, and computing-intensive. Thus, in order to process the large amount of ARGO data we used a parallel processing and Cloud computing system named DataMiner offered by the D4Science e-Infrastructure (www.d4science.org). *DataMiner* (Coro *et al*., 2017) is an open source computational system that allows using Map-Reduce coupled with multi-core processing in scripts and programs written in a large variety of programming languages. An importing tool (SAI, Coro *et al*., 2016) facilitates both software integration and use of distributed computing, and automatically generates a Web interface for the integrated software. We configured our workflow in order (i) to parallelise point 4 on several cores of one (virtual) machine in the DataMiner system, and (ii) to make each input ARGO CSV file processed by one different machine in the Cloud. With respect to other Cloud computing platforms, DataMiner publishes and describes the hosted processes under WPS and produces an XML provenance information file for all the executed processes in the Prov-O ontological format. Our workflow was published as a WPS service through DataMiner. It accepts one ARGO CSV-file as input and produces one NetCDF file for each variable included in the input file. Internally, the workflow uses

the ARGO JSON file to retrieve metadata information. Another advantage of using DataMiner is that it interoperates with the other services of the D4Science e-infrastructure, i.e. (i) a distributed storage system for Big Data (the *Workspace*), (ii) data visualization, browsing, manipulation, and access services, (iii) social networking and collaborative spaces, (iv) security, authorization, and accounting services. D4Science supports Virtual Research Environments (VREs), i.e. online environments that foster collaboration between users and regulate users' access to data and services. We published our workflow in the *ScalableDataMining* VRE (accessible at services. d4science.org/web/scalabledatamining/), which grants free access to a DataMiner computing cluster made up of 20 Ubuntu machines with 32GB RAM, 16 virtual cores for single-machine parallelized processing, and 1TB of distributed and high-availability storage wherein the NetCDF files are uploaded after the computations. Through our workflow, the VRE users can process other ARGO data and optionally share the NetCDF files between them. Further, the VRE services allow accessing and visualising these files via OPeNDAP, and retrieving their metadata in ISO-19139 format (thredds.d4science Catalog, 2017). In the ScalableDataMining VRE our workflow required ~15 hours to process all the ~200GB of ARGO data and produced ~120GB of NetCDF files, with provenance information associated to each file. These files were eventually published in the VRE catalogue3 and made accessible through another VRE (the *BiodiversityLab*) that collects people interested in these data for various modelling applications.

Overall, our approach fosters re-usability of the data and goes towards Open Science and the free sharing of results and processes. It allows a user to add specific terms as variable- and global-attributes in order to connect a given dataset to domain-specific ontologies and thus to make it more understandable for certain communities of practice. This also allows using ARGO data in many experiments, for example to feed ecological models and geospatial interpolation services of other e-Infrastructures (e.g. the SeaDataNet DIVA service, www.seadatanet.org/Software/DIVA)

References

Coro G., Panichi G. and Pagano P. 2016. *A Web application to publish R scripts as-a-Service on a Cloud computing platform.* Bollettino di Geofisica Teorica ed Applicata. 57, 51-53. Proceedings of IMDIS 2016.

Coro G., Panichi G., Scarponi P. and Pagano P. 2017. *Cloud computing in a distributed e-infrastructure using the web processing service standard.* Concurrency and Computation: Practice and Experience, 29(18).

thredds.d4science Catalog, 2017, thredds.d4science.org/thredds/catalog/public/netcdf/Argo/catalog.html services.d4science.org/group/biodiversitylab/geo-visualisation.

MEDTRIX: a mapping platform for monitoring data in the Mediterranean Sea

Célia Fery, L'Oeil d'Andromède (France), celia.fery@andromede-ocean.com
Gwenaëlle Delaruelle, L'Oeil d'Andromède/ Andromède Océanologie (France), gwenaelle.delaruelle@andromede-ocean.com
Pierre Boissery, Agence de l'Eau RMC (France), pierre.BOISSERY@eaurmc.fr
Florian Holon, Andromède Océanologie (France), florian.holon@andromede-ocean.com

Since the Water Framework Directive (WFD, 2000/60/ EC) has been enforced, Member States have committed to achieving or maintaining good environmental status of water bodies by 2020 (European Parliament, 2000). The Marine Strategy Framework Directive (MSFD, 2008/56/CE) complements the concept of "good environmental status". The implementation of the Directives requires a prior inventory or an assessment of marine ecosystems through monitoring programs. Chemical, biological, and pressure measurements on spatial scales ranging from a square meter to tens of kilometres allow these networks to estimate the quality of coastal waters.

MEDTRIX is an online platform which groups together such data concerning the monitoring of Mediterranean coastal waters and ecosystems. It has been created in 2013 by the French Water Agency (Agence de l'eau Rhône-Méditerranée-Corse) and the company Andromède Oceanologie. Using Lizmap (3Liz company), this platform makes freely available high resolution spatialized data (between 0 and 80 meters deep, mapping at 1:10000 scale) all along the French Mediterranean coastline as well as some areas in Italy, Tunisia, Morocco and the Atlantic ocean. Users (marine experts, scientists, managers, technical agents, etc.) have access to marine habitat maps, ecological and chemical quality indicators, anthropic pressures (anchoring, aquatic farms etc.), biological data (fish assemblages, sea birds), physical data (temperature, noise) and geographical information (water activities, pollutions, socio-economic data etc.).

Around thirty different projects (12 project holders, 26 different structures) are available for consultation. They are gathered among six categories: monitoring network, coastal water status, workshop sites, coastal management, habitat mapping, and marine observatories (Fig. 1)

Each project has its own tools allowing the user : 1) to have an abstract of the selected project and methods used (the information tool) ; 2) to visualise the data with the ability to check and uncheck the different layers (the layer tool) ; 3) to choose a particular study zone (the locating tool) ; 4) to export the map to PDF (the printing tool) ; 5) to compare data between different parameters (percentage, index..) for each station at different years or between the stations (the comparing tool) and finally 6) to download the layers (the data download tool) (Fig. 2).

As the development of the platform has been a success, with over 1300 current users, further improvements are planned for the coming months.

These include:

- A new visual and graphic user interface, more dynamic. This homepage, will provide a detailed presentation of the platform, the latest news, the tools, the partners and collaborators

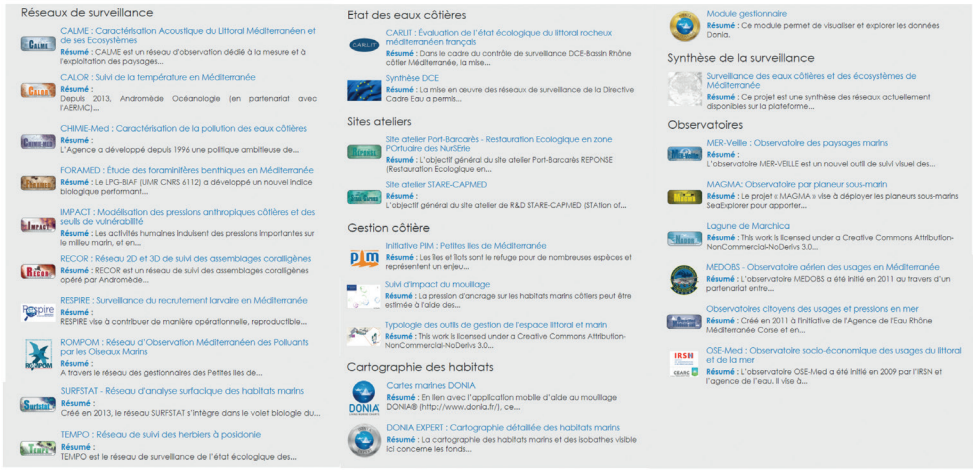


Fig. 1 - Projects currently available (currently in French but in English from November 2018).

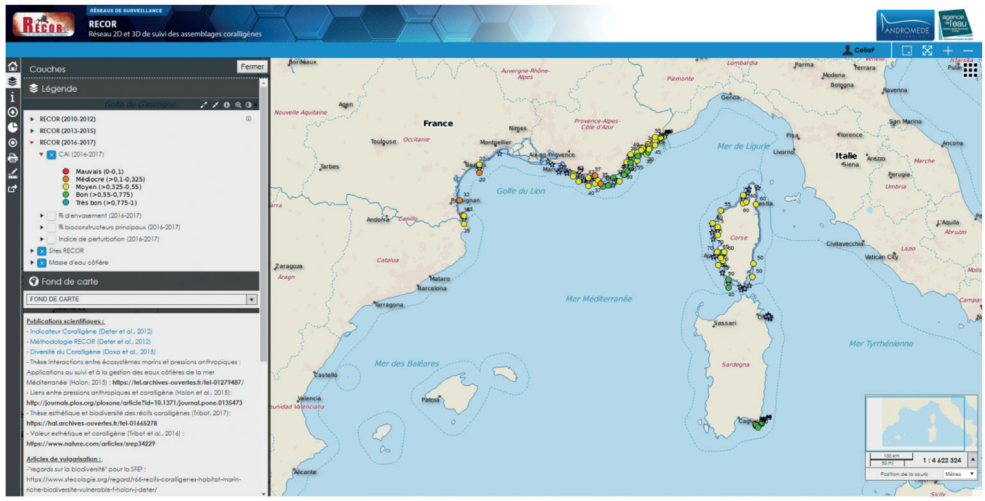


Fig. 2 - The project "RECOR" on coralligenous reefs, showing the ecological status of the stations: CAI (Coralligenous Assemblage Index). The different tools are also visible on the vertical left-hand menu bar.

and will display the different publications, mission reports and other documents cited and used on medrix.fr. It will also be possible to switch between several languages: English, Italian, and Spanish, for a better understanding by non-French speakers.

- New functions : 1) the data interoperability using WMS ; 2) the keyword search ; 3) an on-line agreement for downloading data by project ; 4) an observation tool, allowing users to report any misunderstanding, error of display or lack of data on the various projects of the platform; and 5) a special space to "build his/her own map". This new feature will allow users to build their own map in a dedicated project by combining the available layers in Medtrix. The user will be able to add a title and export his/her map with the printing tool.

SEANOE, a publisher of scientific data in the field of marine sciences

Fred Merceur, Institut Français de Recherche pour l'Exploitation de la Mer (France), Frederic.merceur@ifremer.fr

Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), Michele.Fichaut@ifremer.fr

SEANOE, a publisher of scientific data in the field of marine sciences

In 2015, Ifremer opened SEANOE[1] (SEA scieNtific Open data Edition), a publishing tool of marine scientific datasets.

Data published by SEANOE are freely available. They can be used in accordance with the terms of the Creative Commons license selected by the author of data. SEANOE contributes to Open Access / Open Science movement for a free access for everyone to all scientific data financed by public funds for the benefit of research.

An embargo limited to 2 years on a dataset is possible in order to restrict access to data of a publication under scientific review, for example.

Each data set published by SEANOE has a DOI, which enables it to be cited in a publication in a reliable and sustainable way.

Assessment of 3 years management of SEANOE

A majority of scientists is not used to publish and share the data associated to articles: the positive return to suggestion made to Ifremer authors of scientific papers to publish the data sets linked to their publications is less than 1%.

But the situation could change in the near future thanks to the increasing number of publishers (e.g : PLOS ONE, Elsevier) requiring that the data used in a publication is cited and freely accessible online. It will improve credibility and transparency of the articles.

Furthermore, to convince authors to publish datasets in SEANOE, Ifremer has developed a set of services such as:

- A pro-active survey of citation of DOIs set by SEANOE (bibliography)
- The automatic provision of download statistics to authors
- In addition to properly cite the dataset, the authors can ask users to cite some additional work(s) when using the dataset in a publication

If only a few[2] datasets have been published in SEANOE yet, there are some grounds for satisfaction:

- Most authors are happy with the provided service. They highlight the simplicity and the speed of the publication process and the quality of Landing Pages.

[1] http://www.seanoe.org.

[2] In April 2018, 300 datasets are available in SEANOE.

- The analysis of download statistics shows that datasets published in SEANOE have an international visibility, mainly through search engines such as Google.
- The number of SEANOE DOI citations is relatively high compared to the number of datasets.

Data replication from SEANOE to EMODnet Ingestion

The automated replication of the data published in SEANOE to EMODnet data ingestion portal (emodnet-ingestion.eu) is under study. It would allow reporting to Marine European data centers the existence and the availability of datasets published in SEANOE. Then the data centers will be able to ingest them into other international databases such as SeaDataNet and EMODnet thematic ones, if the dataset falls into their domains of interest

From a technical point a view, this should not be a problem. However, most of dataset published in SEANOE are available under the CC-BY creative commons license. This means that the authors ask to be cited when the dataset is used. This may not be compatible with systems such as SeaDataNet that do not handle metadata on authors, but only on originator organisations. So only dataset published with the CC0 (Public Domain) license may have to be duplicated in EMODnet Ingestion or a specific authorization for this replication may have to be provided by the authors. This kind of problems will be addressed in frame of this study in order to find an appropriate solution.

Fig. 1 - The Landing Page of a French dataset published in SEANOE.

Sharing survey data from the offshore renewables sector

Chelsea Bradbury, The Crown Estate (United Kingdom), Chelsea.bradbury@thecrownestate.co.uk

As The Crown Estate, we manage the seabed around England, Wales, and Northern Ireland out to international borders. These waters harbour some of the best wind resources worldwide, alongside the infrastructure of ports, transmission grids, a diverse supply chain and a skilled labour force. With 36 large offshore wind farms already contributing on average 6% of the nation's electricity demand, the sector is on track to becoming a vital source of power generation for the country.

Paramount to the success of these offshore renewable projects is good data. Each year the industry invests millions of pounds on site investigations, resource assessments and understanding potential environmental impacts. All of this data from customers of The Crown Estate is captured in the Marine Data Exchange (MDE); a system that was developed to store, manage and disseminate offshore survey data.

With the number of offshore renewable projects increasing, the size of survey data growing exponentially and data becoming increasingly sensitive in a competitive market, we have had to ensure that the MDE develops in tandem and that our processes for managing and sharing data benefit the industry without jeopardising any one of the projects.

With over 100TB of data stored on the MDE, The Crown Estate has worked with the offshore renewables industry to publish over half of all survey data.

This presentation discusses some of the lessons learnt of working with industry to share offshore survey data.

On-line visualization of the cal/val indices for the Mediterranean Monitoring and Forecasting Centre

Elena Zhuk, Marine Hydrophysical Institute of the Russian Academy of Science (Russia), alenixx@gmail.com

Georgios V. Kozyrakis, Foundation for Research and Technology - Hellas, Coastal & Marine Research Lab., Institute of Applied and Computational Mathematics (Greece), gkoz@iacm.forth.gr

George Galanis, Hellenic Naval Academy, Mathematical Modelling and Applications Lab., Section of Mathematics (Greece), gngalanis@gmail.com

George Zodiatis, Oceanography Centre, University of Cyprus, (Cyprus), Foundation for Research and Technology - Hellas, Coastal & Marine Research Lab., Institute of Applied and Computational Mathematics (Greece), oceanosgeos@gmail.com

Katerina Spanoudaki, Foundation for Research and Technology - Hellas, Coastal & Marine Research Lab., Institute of Applied and Computational Mathematics (Greece), kspanoudaki@gmail.com

Nikos Kampanis, Foundation for Research and Technology - Hellas, Coastal & Marine Research Lab., Institute of Applied and Computational Mathematics (Greece), kampanis@iacm.forth.gr

Dmitry Soloviev, Marine Hydrophysical Institute of the Russian Academy of Science (Russia), solmit@gmail.com

Data description

The current development allows the visualization of the statistical indices estimated between the observations of the NOAA ISD/ISH METAR stations (National Ocean and Atmospheric Administration Integrated Surface Data / Integrated Surface Hourly MEteorological Terminal Aviation Routine weather report) (NOAA 2001) and the ECMWF forecasting data (European Center for Medium Weather Forecast), as part of the calibration/validation (cal/val) of the surface forcing used by the Med MFC (Mediterranean Monitoring and Forecasting Center). Five well-established statistical indices were selected and implemented for the cal/val of the ECMWF data used by the Med-MFC:

(a) Bias, (b) RMSE (Root Mean Square Error), (c) the Nash-Sutcliffe Model Efficiency Coefficient (Nash & Sutcliffe, 1970), (d) the Correlation Coefficient and (e) the Precipitation Capture Rate. The aforementioned indices provide a good correlation estimate between the METAR observations and the ECMWF forecasting data in the Mediterranean and is a useful tool for the calibration/validation purposes of the Med-MFC (Kozyrakis *et al*. 2018).

Structure of the cal/val data access system

The METAR observations, the ECMWF forecasts and the cal/val indices estimations are archived in a dedicated FTP server (Fig. 1). For nowcast cal/val indices estimatations four files are generated for the last three days, on a daily basis. These files include hourly observations for each METAR station, ECMWF forecasting data and the statistical indices estimations for the

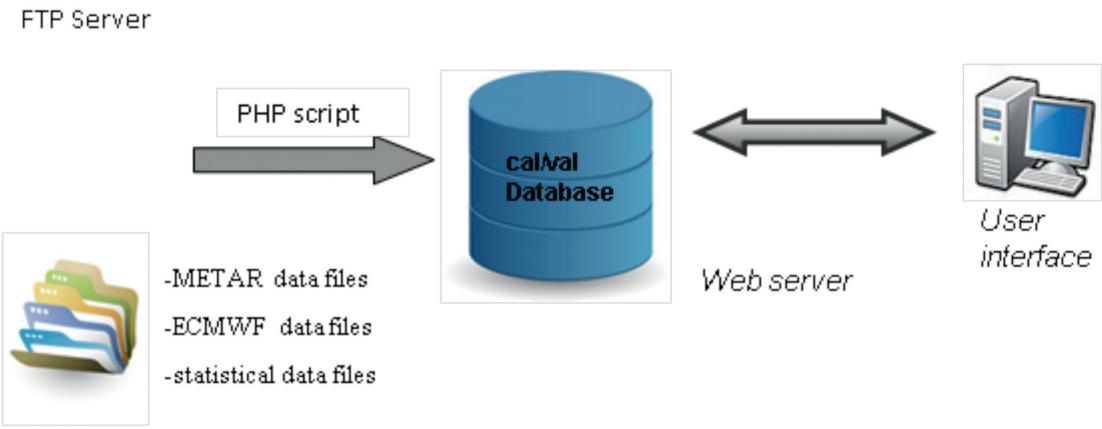


Fig. 1 - Data access schema.

following parameters:

- Two-meter a.s.l. Dewpoint Temperature (in oC).
- Mean Sea-Level Pressure (in bar).
- Precipitation (occurance and approximate duration).
- Two-meter a.s.l. Temperature (in oC).
- Ten-meter a.s.l. Wind Speed (in m/s).

The above mentioned data are processed by a PHP-script, which parses data and upload them to the cal/val database tables. The web server provides a user-friendly interface to access all the relevant cal/val data (Fig. 2).

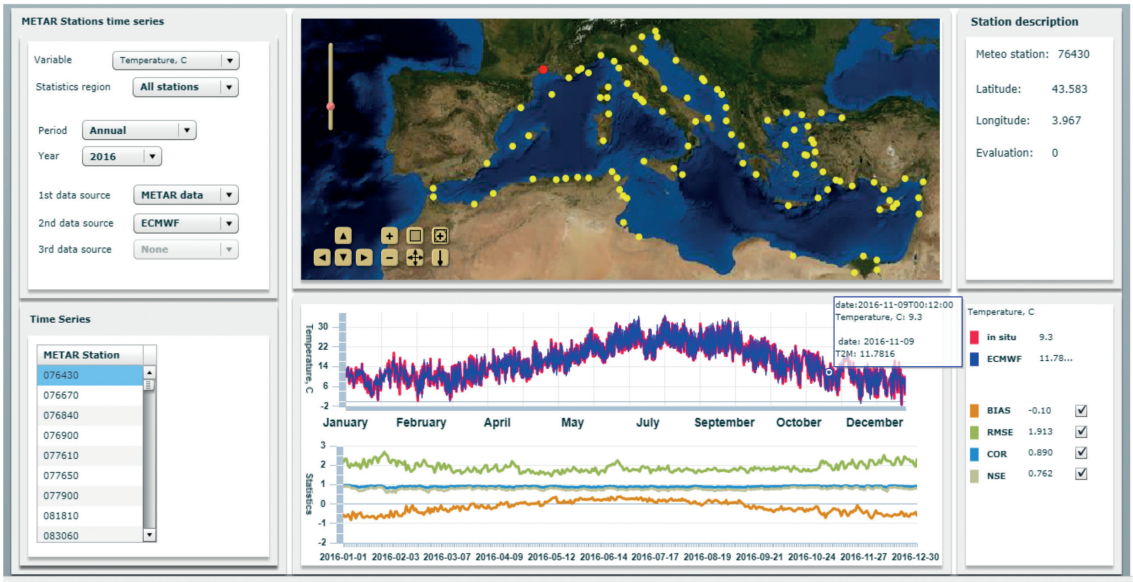


Fig. 2 - Data access interface, visualization of temperature (in-situ, forecast and statistics) for METAR station 076430.

cal/val data access interface

The cal/val data access interface provides access and visualisation of the METAR observations, the ECMWF forecasting data and the estimations of the five statistical indices for each METAR station, as well as, the mean average of all the METAR station's data for the years 2016 and 2017.

METAR observations and ECMWF forecasting data can be shown simulaneously as time-series plots, while the data for the statistical indices are shown in a separate time-series plot. The cal/val statistical indices for the year 2018 are accesible on a daily base, using the latest three days observations. The parameters at each METAR station can be accessed and presented as values and time-series plots with relevant cal/val information. The cal/val Med-MFC on-line system can be accessed at: http://orioncyprus.org/meteoStations/meteoStations.html

Conclusions

In this work the development of a user-friendly data access web interface dedicated to the cal/val indices of the Med-MFC is demonstrated. The software provides access and vizualization of both archived and operational data for the Mediterranean sea and can be used by all interested scientist to access the quality of the surface forcing used for the Med-MFC.

References

Georgios V. Kozyrakis, George Galanis, Katerina Spanoudaki , Nikolaos A. Kampanis, George Zodiatis, and Elena Zhuk (2018). *Long - Term validation of forecasting results based on in situ ground measurements for contributing to the cal/val of the Mediterranean Monitoring and Forecasting Centre (Med-MFC).* Geophysical Research Abstracts, Vol. 20, EGU2018-5341.

NOAA National Centers for Environmental Information (2001): *Global Surface Hourly.* NOAA National Centers for Environmental Information. [2016-2017].

J.E. Nash, J.V. Sutcliffe (1970). *River flow forecasting through conceptual models part I - A discussion of principles.* Journal of Hydrology, Volume 10, Issue 3.

QuinCe: An online tool for processing and quality control of surface ocean CO_2 measurements

Steve D. Jones, ICOS OTC/Bjerknes Climate Data Centre, University of Bergen (Norway), steve.jones@uib.no

Jonas Henriksen, ICOS OTC/Bjerknes Climate Data Centre, University of Bergen (Norway), jonas.henriksen@uib.no

Scientists collecting surface ocean CO_2 measurements must perform data reduction and quality control (QC) prior to further usage of the data. The tools for these tasks are typically developed by the scientists themselves, leading to dozens of different software programs in use across the community. Different interpretations of protocols lead to inconsistent data handling, and since these self-developed tools are rarely published there is little transparency and traceability for users of the data – increasing the overall uncertainty of these observations.

The Ocean Thematic Centre (OTC) of the pan-European Research Infrastructure Integrated Carbon Observation System (ICOS) is developing online software to provide data reduction and QC tools for all members measuring surface ocean CO_2. Scientists submit the raw data from their instruments, which will be processed to calculate surface fCO_2 using a fully open source set of algorithms based upon internationally agreed and defined criteria. A suite of automated QC tools, based on those developed for the Surface Ocean CO_2 Atlas (SOCAT), will identify basic issues such as out-of-range data, outliers, date/time issues etc., and flag them for later verification by the scientist. Plotting and mapping tools will allow detailed inspection of the data to locate further potential issues. Each data record will be assigned a WOCE flag stating the quality (Good, Questionable or Bad). Once this QC has been completed, the data can be exported in a variety of formats, and/or submitted directly to central data archives such as the ICOS Carbon Portal or SOCAT.

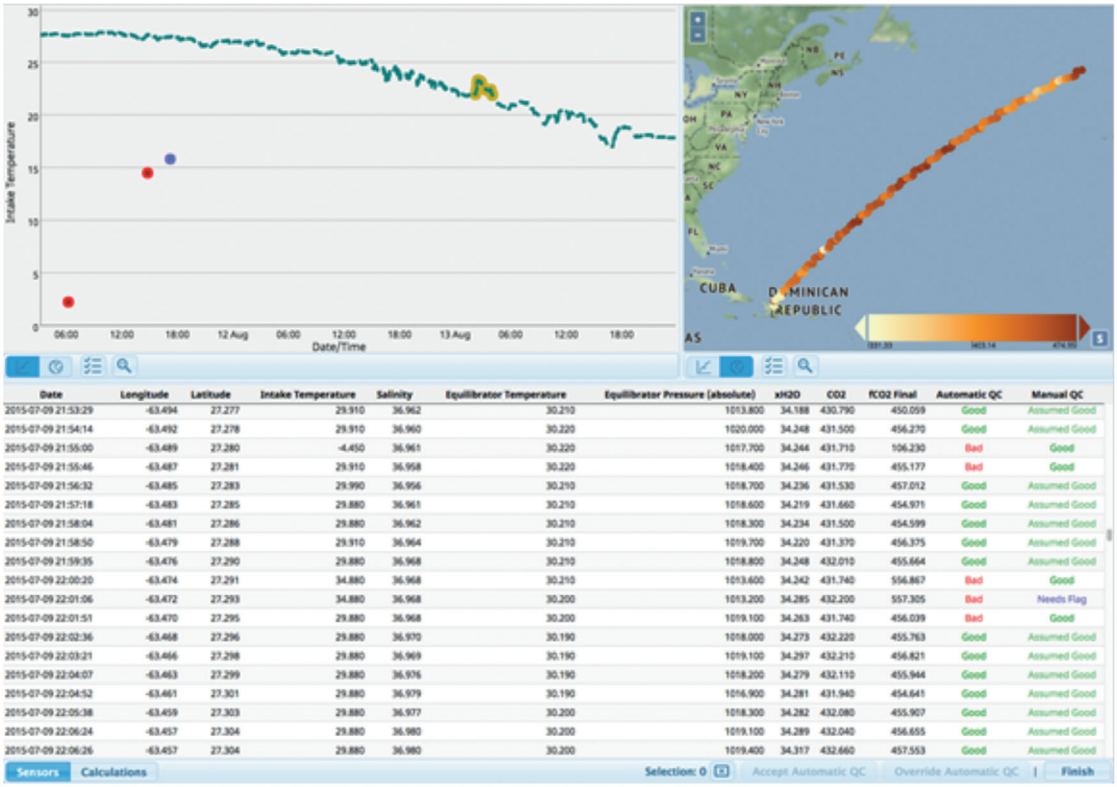
Metadata will be integrated into this process using internationally approved protocols and schemas which are currently being developed in conjunction with partner institutions and projects[1], thus rendering the handling of data processing, quality control and metadata from collection to publication fully operational within the sphere of ICOS.

Data provenance will be recorded in a number of ways. Data sets published through the ICOS Carbon Portal will include both the quality-controlled data and the original data files as they were uploaded to QuinCe. For other projects, these files will be available on demand. The metadata for each data set will include a note stating that processing was performed using QuinCe, with version numbers so the exact code used can be traced in future (all versions of QuinCe will be archived in a public version control system). Records of every quality control flag set by the automatic and manual quality control processes will be kept in QuinCe. These

[1] NOAA Ocean Acidification Project, NOAA PMEL, and NOAA NCEI OCADS.

can be made available as required, but protocols for standardising and automating such features have not yet been developed.

Unified tools like QuinCe will streamline data processing, helping to reduce delays in data publishing. This is becoming ever more important as demand for up-to-date climate data grows (e.g. the United Nations Sustainable Development Goals 14.3 [Ocean Acidification]). Planned features such as accepting near real time data streams transmitted directly from the instruments themselves and more sophisticated automatic QC checks will allow even greater time savings. Easy-to-use data inspection tools will reduce the effort required to perform additional manual QC. Data centres will be assured that data submitted via QuinCe will have undergone a minimum level of QC, with records of both the automatic and manual QC providing complete transparency and improved stewardship. QuinCe will also ensure automatic data conversion to formats required by data centres, reducing the incidence of data transmission and formatting errors and also reducing the workload of both the scientists and data managers.



Date	Longitude	Latitude	Intake Temperature	Salinity	Equilibrator Temperature	Equilibrator Pressure (absolute)	xH2O	CO2	fCO2 Final	Automatic QC	Manual QC
2015-07-09 21:53:29	-63.494	27.277	29.910	36.962	30.210	1013.800	34.188	430.790	450.059	Good	Assumed Good
2015-07-09 21:54:14	-63.492	27.278	29.910	36.960	30.220	1020.000	34.248	431.500	456.270	Good	Assumed Good
2015-07-09 21:55:00	-63.489	27.280	-4.450	36.961	30.220	1017.700	34.244	431.710	106.230	Bad	Good
2015-07-09 21:55:46	-63.487	27.281	29.910	36.958	30.220	1018.400	34.246	431.770	455.177	Bad	Good
2015-07-09 21:56:32	-63.485	27.283	29.990	36.956	30.210	1018.700	34.236	431.530	457.012	Good	Assumed Good
2015-07-09 21:57:18	-63.483	27.285	29.880	36.961	30.210	1018.600	34.219	431.660	454.971	Good	Assumed Good
2015-07-09 21:58:04	-63.481	27.286	29.880	36.962	30.210	1018.300	34.234	431.500	454.599	Good	Assumed Good
2015-07-09 21:58:50	-63.479	27.288	29.910	36.964	30.210	1019.700	34.220	431.370	456.375	Good	Assumed Good
2015-07-09 21:59:35	-63.476	27.290	29.880	36.968	30.210	1018.800	34.248	432.010	455.664	Good	Assumed Good
2015-07-09 22:00:20	-63.474	27.291	34.880	36.968	30.210	1013.600	34.242	431.740	556.867	Bad	Good
2015-07-09 22:01:06	-63.472	27.293	34.880	36.968	30.200	1013.200	34.285	432.200	557.305	Bad	Needs Flag
2015-07-09 22:01:51	-63.470	27.295	29.880	36.968	30.200	1019.100	34.263	431.740	456.039	Bad	Good
2015-07-09 22:02:36	-63.468	27.296	29.880	36.970	30.190	1018.000	34.273	432.220	455.763	Good	Assumed Good
2015-07-09 22:03:21	-63.466	27.298	29.880	36.969	30.190	1019.100	34.297	432.210	456.821	Good	Assumed Good
2015-07-09 22:04:07	-63.463	27.299	29.880	36.976	30.190	1018.200	34.279	432.110	455.944	Good	Assumed Good
2015-07-09 22:04:52	-63.461	27.301	29.880	36.979	30.190	1016.900	34.281	431.940	454.641	Good	Assumed Good
2015-07-09 22:05:38	-63.459	27.303	29.880	36.977	30.200	1018.300	34.282	432.080	455.907	Good	Assumed Good
2015-07-09 22:06:24	-63.457	27.304	29.880	36.980	30.200	1019.100	34.289	432.040	456.655	Good	Assumed Good
2015-07-09 22:06:26	-63.457	27.304	29.880	36.980	30.200	1019.400	34.317	432.660	457.553	Good	Assumed Good

Fig. 1 - Example of the Quality Control screen in the QuinCe software.

An integrated database for marine environment monitoring and management system at the Tongyoung bay in Korea

Sung-Dae Kim, Korea Institute of Ocean Science and Technology (Republic of Korea), sdkim@kiost.ac.kr
Sang-Hwa Choi, Korea Institute of Ocean Science and Technology (Republic of Korea), choish@kiost.ac.kr
Hyuk-Min Park, Korea Institute of Ocean Science and Technology (Republic of Korea), hyukmin@kiost.ac.kr

A marine environmental monitoring and management system to support efficient operation of the aquafarm at Tongyoung bay in Korea, was started to be developed by the Korea Institute of Ocean Science and Technology (KIOST) in 2017. The system consists of an intensive observation system, a data management system, a hybrid environment prediction system, and a facility management system based on artificial intelligence technology. During the 1st phase of the project (2017-2018), we are setting up a database system to manage marine data of the intensive observation and to archive all available environmental data of Tongyoung bay area. Several physical parameters and chemical parameters measured at the surface buoy and automatic vertical profiler are transmitted to the server through wireless network, processed and stored into DB system, and provided to the researchers via intranet in real-time mode. Observed data of biological parameters and biogeochemical parameters are being submitted to the database system by the related researchers in delayed mode. We also collected all oceanographic data produced by several organizations and satellites (Fig. 1). We manipulated all collected data in accordance with the metadata standard and quality control procedures which had been prepared for rasearch marined data of KIOST. An internet web site was established to support data retrieval and share collected data with researchers. To develop marine environment prediction system, a data driven model and a numerical dynamic model are under development. We are trying to build a data driven model based on LSTM (Long Short tem Memory) network of RNN (Recurrnet Neural Network) to simulate biological parameters based on physical parameters and chemical parameters. Tensorflow librarys was used to set up machine learning programs.

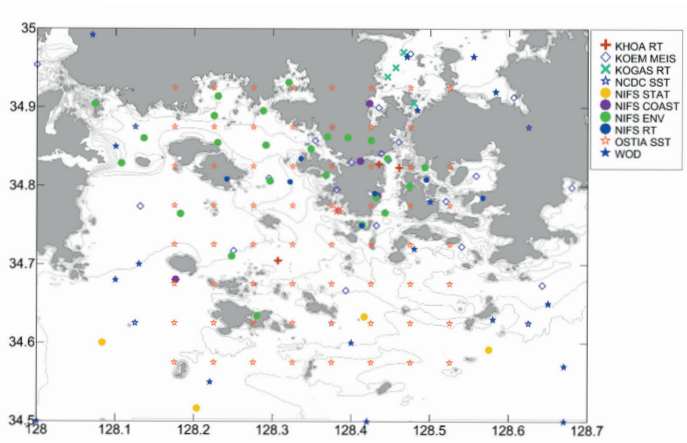


Fig. 1 - Observation stations near Tongyoung bay.

An ensemble model will be set up to combine the results of the data driven model and the numerical model and provide prediction information to the aquafarm operation system.

Automatic assessment of metadata quality in ISO 191xx

Susanne Feistel, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), susanne.feistel@io-warnemuende.de

Ulrike Kleeberg, Helmholtz-Zentrum Geesthacht (Germany), ulrike.kleeberg@hzg.de

Jörn Kohlus, Nationalpark Wattenmeer, (Germany), joern.kohlus@lkn.landsh.de

Rainer Lehfeldt, Federal Waterways Engineering and Research Institute (Germany), rainer.lehfeldt@baw.de

Carsten Schirnick, GEOMAR Helmholtz Centre for Ocean Research Kiel (Germany), cschirnick@geomar.de

Susanne Tamm, Federal Maritime and Hydrographic Agency (Germany), susanne.tamm@bsh.de

Sharing geoinformation and data across communities is becoming more important, which is enabled by web services. The Open Data initiative (e.g. data.gov) promoted by public agencies and research institutions enhances this development. However, poor data quality with insufficient quality information may be hindering the acceptance and usage of the data by the scientific community in the future. The mandatory metadata within ISO 191xx is not sufficient for a comprehensive quality assessment or long-term usability of any dataset. Quality information is mostly optional and not easily found or filtered. It might be stored within the metadata as information in LI_Lineage, DQ_DataQuality or even within the dataset itself, possibly coded. Ultimately, in most cases, quality information is not directly accessible for interested users.

On the other side, for the data creator the documentation of quality information can get time consuming and overwhelming, especially if required all at once in the reporting phase of a project. Quality Flag Schemes have been proposed by multiple organisations. In UNESCO 2013, Manuals and Guides 54, the most commonly used are presented and compared. Mutual conversion rules between them are provided at http://odv.awi.de/fileadmin/user_upload/odv/misc/ODV4_QualityFlagSets.pdf. Furthermore, quality information gets significantly more complex if more than one dataset is considered as in data products such as maps or modelling results.

We present a more practical approach to get a sufficient and standardized quality assessment (quality flag), dynamically generated from the entire quality information of the metadata XML file. While the ISO standard offers a complex array of optional fields to provide quality information, we recommend a manageable number of fields to be filled. As prepared in (Feistel *et al*. 2018, in prep.) there are specific information a user needs first and foremost to evaluate the usefulness of a dataset.

Our task group provides a set of tools to assess quality information recorded in the recommended fields of the ISO standard. The first is a web-based form to manually assess a single dataset, the second is a program for syntactically analysing multiple dynamically generated XML in ISO 19139. To make the quality assessment compatible and comparible between platforms and schemes, as well as machine-friendly, we propose a coded summary

string in the pattern of the scheme plus a flag, e.g. “SDN::1” to be put in the ISO field “DQ_ StandaloneQualityInformation”.

For data creators, metadata templates can be prepared in advance. This is especially useful for repeating data collection or creation procedures within projects, as well as long-term data series. Thus, the compilation of metadata can be simplified for individual scientists. The advantage of using an automatic assessment of metadata quality (ISO XML checker) is primarily in getting a standardized quality flag based on a common quality flag scheme. A time consuming individual control is not necessary. The quality control mechanism supports users in finding data relevant for their own work more efficiently.

Compact automated system for sea ice monitoring based on SAR data

Eduard Kazakov, Nansen International Environmental and Remote Sensing Centre (Russia), ekazakov@niersc.spb.ru

Vladimir Volkov, Nansen International Environmental and Remote Sensing Centre (Russia), vladimir.volkov@niersc.spb.ru

Denis Demchev, Nansen International Environmental and Remote Sensing Centre (Russia), demchev@gmail.com

The understanding of ice regime of the seas of the Arctic region is important for two reasons. First, changes in ice types distribution, variations in ice circulation patterns, and other regional sea ice processes are important indicators of climate change, especially in view of the Arctic Amplification phenomenon. Second, the polar and sub-polar regions of the Northern hemisphere are now attracting more and more investments, the volume of economic activity and transportation is growing, mining is intensified. In this regard, many companies are required to provide their activities with high-quality operational information on the state and dynamics of sea ice.

To perform this kind of work, there was a task to develop a comprehensive information system that would aggregate data flows, their processing and bringing to the consumer. At the same time, such a system should be compact, capable of rapid deployment for the tasks of a particular consumer within a certain area, as well as provide automated processing and presentation of data on standard protocols.

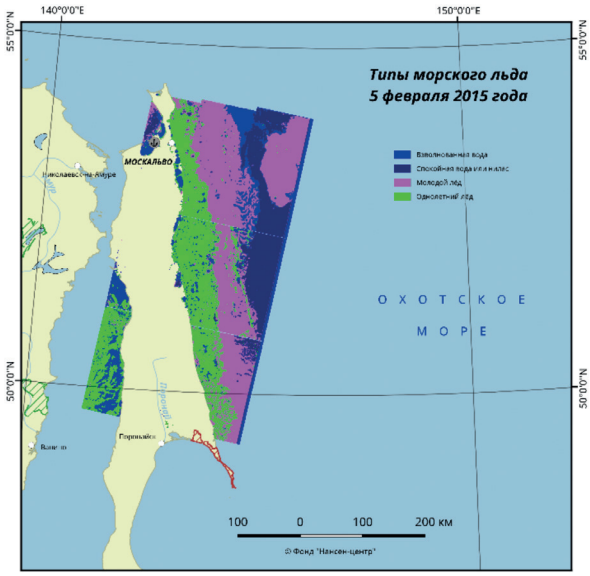


Fig. 1 - Example of automated ice classification in Sea of Okhotsk with 3 Sentinel-1A scenes. Blue areas - water, pink - young ice, green - one year ice.

The functional basis of the developed system consists of high-performance thematic algorithms for extracting information about sea ice from satellite radar data by Radarsat-2 and Sentinel-1 A/B spacecrafts.

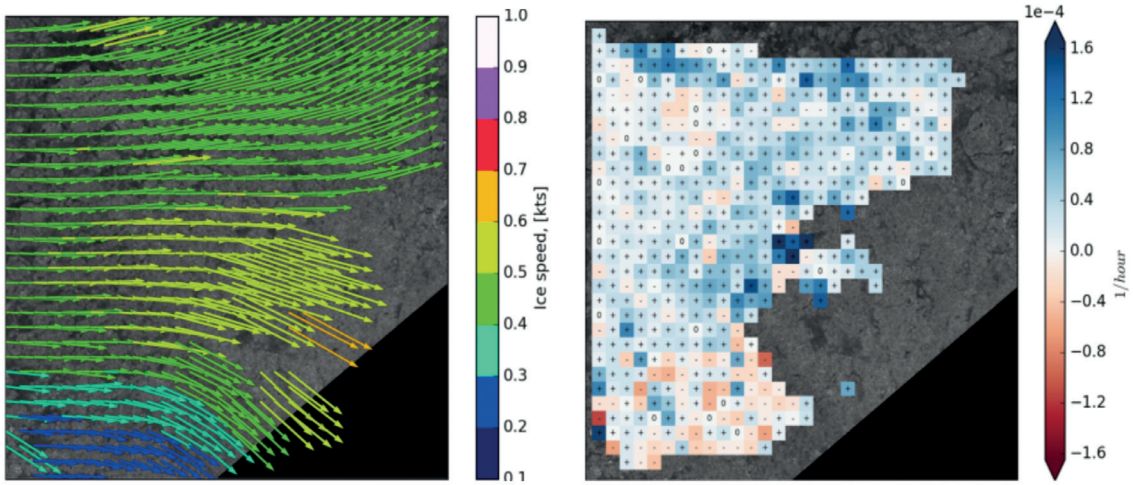


Fig. 2 - Left - example of retrieved ice drift vectors for two Sentinel-1A scenes.
Right - Calculated with ice drift vectors fields of ice divergence and convergence.

The algorithm of automated classification of sea ice by types is based on machine learning and involves preliminary training by the ice expert and, then, operation in automatic mode. The result of this algorithm is detailed (with a spatial resolution up to 500 meters) fields of ice distribution by standard types (Fig. 1). The algorithm for calculating ice drift fields with pairs of radar images allows to restore the detailed (with a distance between drift vectors up to 600 meters) characteristics of the movements of ice masses (Fig. 2, left). The highest (in comparison with other known algorithms) detalization of ice drift fields also allows to calculate accurate maps of divergence and convergence of the ice cover (Fig. 2, right), which are especially important for ensuring safe navigation.

A compact information system is a framework based on open source software designed for deployment to Virtual Dedicated server (VDS), which can be rented quickly when a task appears. The deployed system provides continuous scanning of a set of FTP-servers on which the consumer or we ourselves place satellite information manually or automatically. Any detected SAR dataset then loaded onto the VDS and is used to calculate classification maps, drift fields, and divergence/convergence of sea ice. Products are automatically published via OGC protocols (WMS, WFS) and also placed in file storage with FTP access. In addition, the framework includes a web-GIS interface for quick visual access to data and assessment of the current situation.

The coastMap Approach for visualization and dissemination of marine geodata

Linda Baldewein, Helmholtz-Zentrum Geesthacht (Germany), linda.baldewein@hzg.de
Ulrike Kleeberg, Helmholtz-Zentrum Geesthacht (Germany), ulrike.kleeberg@hzg.de
Marcus Lange, Helmholtz-Zentrum Geesthacht (Germany), marcus.lange@hzg.de
Dietmar Sauer, Helmholtz-Zentrum Geesthacht (Germany), dietmar.sauer@hzg.de

coastMap is a coastal data portal with a focus on the North Sea and offers campaign data, model analysis tools and thematic maps predominantly in the Biogeosciences. Important research topics are highlighted by Spotlights, an edited form of scientific information for non-experts. In order to address the large variety of formats of the accessible data and the different user groups of the portal (modeling and non-modeling scientists, policy makers and the interested public) the coastMap technical infrastructure is highly diverse.

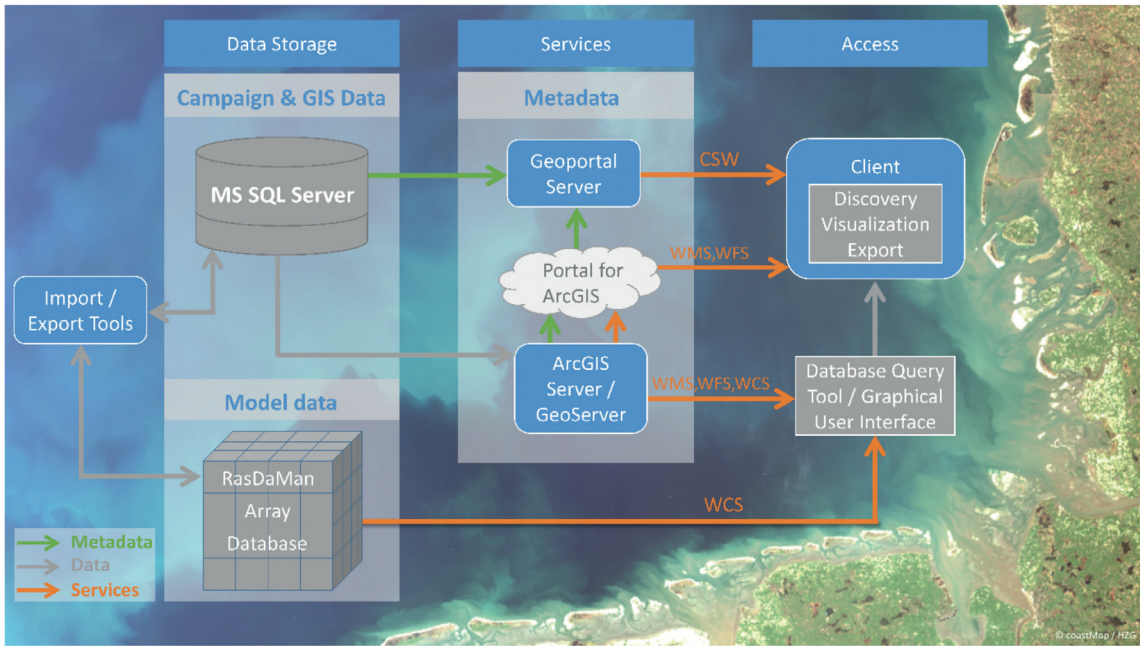


Fig. 1 - Overview of coastMap technical infrastructure including data, metadata and service flows.

In Fig. 1, the technical infrastructure of coastMap is schematically displayed. It is separated in data storage, services and access of the user to the information.

Marine data from research campaigns is highly complex in its description, which requires extensive metadata for useful access of the data. The data and its metadata are stored in the

relational database MS SQL Server allowing targeted queries. These queries can be sent using the Graphical User Interface (GUI) to search the data for particular information.

The biogeochemical model data is vast but comparatively simple in metadata and is stored in the RasDaMan Array Database. It is set up with a Web Coverage Service (WCS) that feeds into the GUI of the Model Analysis Tool.

The data for the thematic maps are stored in the Portal for ArcGIS cloud and its metadata on the MS SQL Server. The Geoportal Server publishes the metadata for all maps using a Catalog Service for the Web (CSW). Web Map Services (WMS) and Web Feature Services (WFS) convey the maps and data to the client. The Spotlights are conceptualized and published in the Portal for ArcGIS Cloud environment.

Using the GeoServer as an intermediary, the campaign data is made freely available on the web through Web Feature Services. The WFS request is assembled according to the query requests given by the user in the GUI. The data are visualized in maps and provided to the client as download links.

The coastMap Model Analysis Tool allows data analysis, management, discovery and visualization of large quantities of marine and atmospheric model data. The focus of the tool is set on creating timely responses for any user defined spatial and temporal subsetting of the data. The intermediary scripting enables data reformatting, so that the user receives the output in the desired format, such as netCDF files and file geodatabases. The front end is set up with an ArcGIS Server, which provides the user with a relatively simple GUI and a well-developed website design.

The architecture of the Model Analysis Tool follows the model-view-controller (MVC) software development paradigm and is displayed in Fig. 2. The high-resolution model data is stored in the RasDaMan multidimensional Array Database on the RasDaMan Server. For optimal

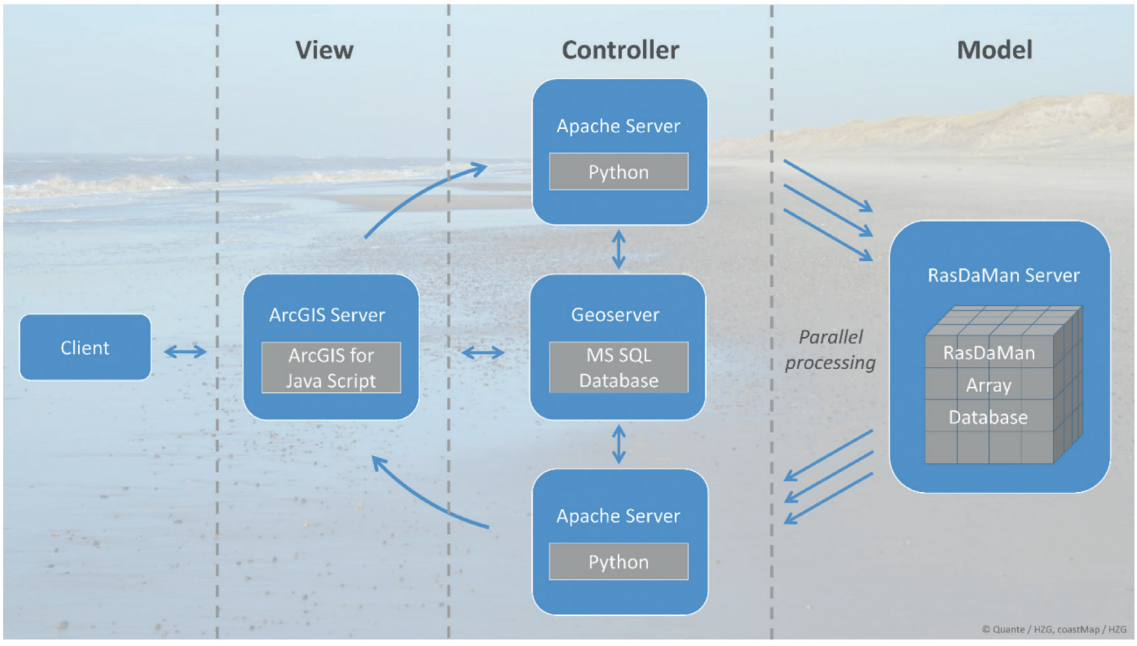


Fig. 2 - Architectural pattern of the coastMap Model Analysis Tool.

performance, the user requests for analysis and data access are split into parallel processes by Python scripts. These processes formulate rasql queries, which are received by the Web Coverage Processing Service (WCPS) of the RasDaMan Server. Further Python scripts recombine the query results, reformat them and deliver them to the user. The User Interface is configured to be derived from the model metadata that is stored in the MS SQL database and delivered by the Geoserver.

The coastal data portal coastMap has a fully developed infrastructure including a variety of services and tools required to publish all different kinds of marine data and data products to different user groups. It can serve as an example for other data portals struggling with the integration of the heterogeneous data landscape and the requirements of different stakeholders.

Real-time Web-GIS as part of Japan's MDA

Takanori Katsura, Japan Coast Guard (Japan), jodcint@jodc.go.jp
Takuma Fujiwara, Japan Coast Guard (Japan), jodcint@jodc.go.jp
Hiroyasu Furukawa, Japan Coast Guard (Japan), jodcint@jodc.go.jp

Information provision by Japan Coast Guard

Japan Coast Guard (JCG) have provided not only the navigational safety information for mariners but also the marine information which is expected to be used for various purposes. As part of Integration of Marine-related information under Ocean Policy, in 2010 JCG launched "Marine Information Clearing House (MICH)", which is an internet database service to provide a summary of the marine information and how to reach it. In 2011 JCG started the operation of "Marine Cadastre", which is a Web-GIS service for displaying and overlaying various marine-related information.

MDA in Japan

Japan's MDA (Maritime Domain Awareness) is defined as "An effective understanding of marine-related situations and circumstances through the efficient collection and sharing of a variety of marine-related information conducive to Japan's maritime security, marine safety, natural disaster countermeasures, the marine environment conservation, the promotion of marine industries, and the development of science and technology, while paying attention to the handling thereof". The Headquarters for Ocean Policy decided "Efforts to consolidate the Capability of

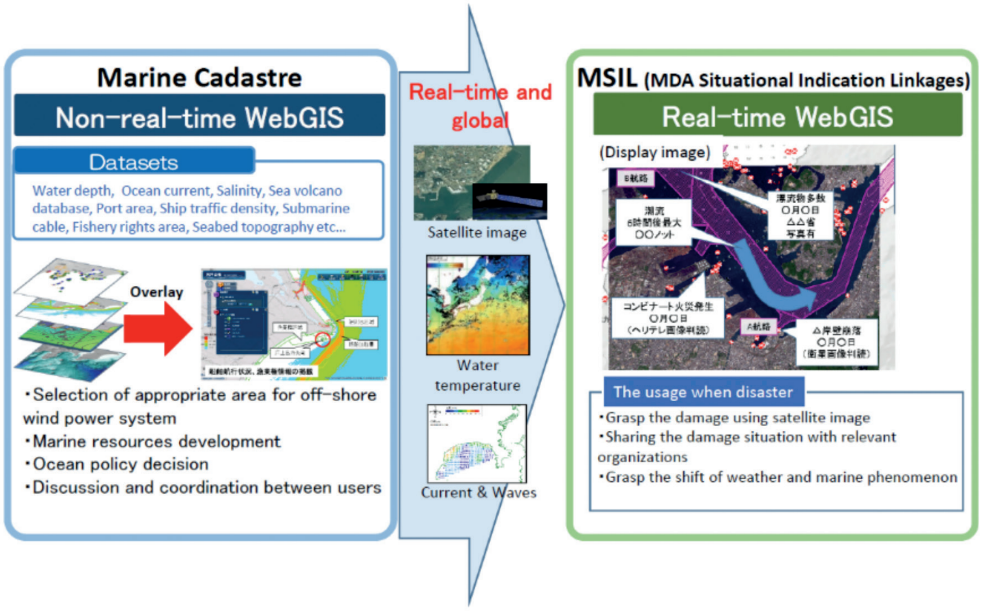


Fig. 1 - Concept image of MSIL(MDA Situational Indication Linkages).

MDA in Japan" in July 2016. To establish the framework for collection, sharing and provision of marine-related information, it was decided that JCG should develop a system for collection, sharing and provision of marine-related information related ministries and agencies possess.

Development of "MSIL (MDA Situational Indication Linkages)"

According to the decision about Japan's MDA by the Headquarters for Ocean Policy, JCG is developing "MDA Situational Indication Linkages", the wide-covered, real-time and user-friendly Web-GIS based on Marine Cadastre, targeting to start its service in 2019 spring. With the cooperation of related ministries and agencies, MSIL will cover wide range of information including real-time or near real-time information such as sea surface temperatures, surface currents, sea waves, wind, sea ice condition, weather maps (current & forecast), meteological satellite imagery and navigational warnings in addition to Marine Cadastre's 100+ kinds of static information. Furthermore user can display and overlay not only the infomation but also user's own data.

Application of elements of Big Data technology for storage, access and retrieval of metadata and Roshydromet data

Anastasia Gorbacheva, FGBU Russian Research Institute of Hydrometeorological Information - WDC, Obninsk (Russia), agorbacheva@meteo.ru

A huge array of data on the state of environment and the World Ocean are stored in the state fund of Roshydromet, which is in FGBU VNIIGMI-WDC. The Unified State Data Fund (EGFD) is an orderly, constantly updated set of documented information obtained as a result of activities of the Federal Service of Russia for Hydrometeorology and Environmental Monitoring. Access to hydro meteorological information of the state data fund is one of main problems that arises in RIHMI-WDC. Based on this, the task arises to speed up the process of providing and searching for metadata and data to them.

Processed information is inconvenient to store in traditional DBMS because of complex structure of information, size of databases and risk of system failure.

Increasingly, they use cloud-based In-Memory Data Grid (IMDG) architecture built on Big Data technology. This solution removes load from relational database and ensures the reliability of the system.

IMDG is a clustered key-value storage that is designed for highly loaded projects with large amounts of data and increased requirements for scalability, speed and reliability.

The task of technology is provide ultra-high availability of data by storing them in RAM in a distributed state. The data stored in the IMDG must be processed in parallel and always updated when new information is received at each node of the system. IMDG technology helps control the indexing of objects and organize a fast full-text search in the parameter dictionaries, paying special attention to a form of recording data and converting them into a readable form for user.

It also provides the ability to work with objects directly, and to use IMDG not only as a separate node, but also as a standalone storage. Using example of IMDG technology, a module for storing, searching and accessing an extensive catalog of metadata was developed. The module is designed to store results of indexing information from the component - Integrated Data Base (BID), the implementation of full-text fuzzy search and the issuance of metadata of information resources (with the connection “description of the array - description of data set sets”) in the IMDG environment (Jboss Infinispan) via REST-service. Due to this development, the load on database and the resource costs are much reduced, and the time for providing information is also increased. The advantage of using this development in a highly loaded system is the provision of an accelerated process of fuzzy and relevant search, using indexing of text and decoded fields of information resources.

The development of this work will be the connection to relational database systems for data access to data through or without metadata with the use of Big Data technology to give high-performance operational access to information from the state fund of data of RIHMI-WDC.

Tools to handle environmental concerns in marine seismic data exploration

Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), pdiviacco@inogs.it

Sebastien Mancini, Integrated Marine Observing System, University of Tasmania (Australia), sebastien.mancini@utas.edu.au

Alessandro Busato, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), busato@inogs.it

Xavier Hoenner, Integrated Marine Observing System, University of Tasmania (Australia), xavier.hoenner@utas.edu.au

Frank-Oliver Nitsche, Lamont-Doherty Earth Observatory of Columbia University (United States of America), fnitsche@ldeo.columbia.edu

Environmental concerns of marine seismic data exploration

Marine seismic data exploration is based on the analysis of the propagation, reflection and detection of acoustic perturbations in the sea and the underground. The source of theese perturbations are impulsive sources that generate sound in different ranges of frequencies. The generated pressure depends on the expected penetration of sound in the underground. In fact, since the propagation through rocks is affected by phenomena such as reflection, refraction, divergence and absorption, only a smaller amount of the energy produced returns to the surface to be detected by sensors. Therefore, higher pressures need to be produced whenever deeper underground areas need to be explored.

Seismic sources that can reach such energy can be harmful for marine life, in particular for some mammals. The impact on these species depends on the habits of these amimals. In the case of cetaceans the component of Oxygen and Nitrogen in the blood gasses depends on the time the animal spends at higher deeps and on the speed the animal returns to the surface. A behaviourally induced decompression sickness-like disease was proposed as a plausible causal mechanism for stranded cetaceans, although these findings remain scientifically controversial.

Mammals can be affected by seismic sources also in other ways, since the sound pressure disturbes communication

The need to integrate data

One very interesting point that emerges from these studies is that to understand the danger that seismic sources can exert on mammals, information on animal behaviour together with information on sea habitats, morphology and physics is needed. For example, there is a correlation between seabottom morphology and the possibility that cetaceans could be affected by decompression sickness, since when the depth of the sea is relatively small this is very unlikely to happen. At the same time it is important to understand the areas where these animals are present and when in the year they gather.

At the same time it is important to understand what seismic data is already available because often there is no actual need to acquire new data but only to reuse what has been already done in the past.

All this is particularly relevant in Antarctica where the protection of the environment is still one of the few cases where international legislation seems to be effective.

The Antarctic Seismic Data Library System

The Antarctic Seismic Data Library System (SDLS) works under the auspices of Scientific Committee on Antarctic Research (SCAR) and the Antarctic Treaty (ATCM XVI-12) to provide open access to all multichannel seismic-reflection data collected south of 60° S to study the structure of the earth's crust of Antarctica.The SDLS currently holds more than 300.000 Km of seismic data from all the institutions that acquired data in all the areas of Antarctica. The system offers a web portal where all data are georeferenced and can be accessed interactively using dedicated previewing systems. The SDLS system integrates data from the Global Multi-Resolution Topography (GMRT) which is a global compilation of multibeam datasets at around 100m resolution with additional grids and GEBCO 2014 as background that is hosted at Lamont-Doherty Earth Observatory of Columbia University.

Australian Ocean Data Network (AODN)

The Australian Ocean Data Network (AODN) is an interoperable online network of marine and climate data resources. AODN data collections cover a large geographic area from the equator to Antarctica and a wide range of observed parameters (physical, chemical, biological).

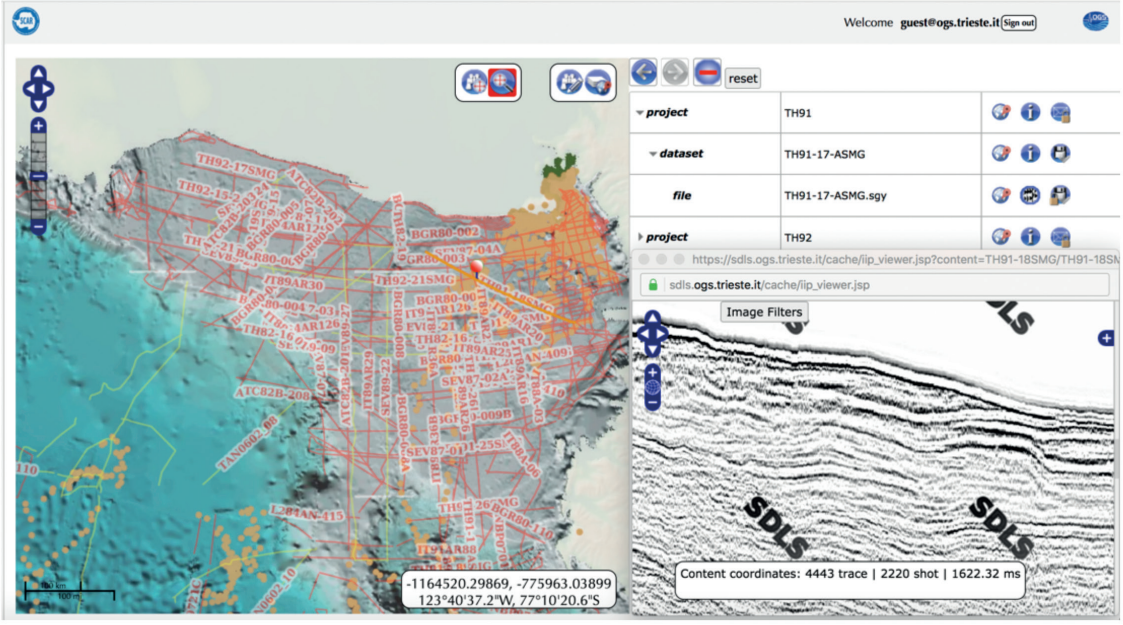


Fig. 1 - Snapshots of the SDLS portal that integrates data from the AODN web services (animals with tags mapped as yellow points) and seismic data (straight red lines) made available from SDLS. In the case of the Ross Sea, mapped in this example, it is very interesting to note how animals concentrate in specific areas, how seismic exploration is mostly interested in the continental platform.

A particularly interesting observation platform that AODN makes available is animal equipped with satellite tags. Since 2008, more than 300 satellite tags have been deployed and around 140 000 profiles have been collected with a wide of range of data including behavioural and physical data such as depth, temperature, salinity and movement effort of indivual marine animals. These satellite tags allow to understand the presence and migration path of mammals in the antarctic area so that during seismic survey plannig it could be possible to understand whether any harm could be possible to mammals and in case plan mitigation.

Data integration

The SDLS, AODN adopt and extend existing standards and open-source software. They offer OGC-compliant Web Map Services and Web Feature Services, so that on each side ther contents offered by the other side can be added and correlations be made in order to understand on one side the possible harm to marine species, and on the other the possibile existence of already available data.

On board cloud system

Sabri Mutlu, TÜBITAK Marmara Araştırma Merkezi (Turkey), sabri.mutlu@tubitak.gov.tr
Pamir Talazan, TÜBITAK Marmara Araştırma Merkezi (Turkey), pamir.talazan@tubitak.gov.tr

Each day as we (individuals or corporations) are getting more benefits from cloud services we rely on them more in our private, corporate or scientific tasks. Therefore, these services become indispensable for any kind of data management solution. However, to sustain this way of dealing data there must be quick and uninterrupted connection which cannot be always the case on board due to difficulties on rough sea conditions. To overcome this problem and ensure data integrity, a private cloud-based service is installed on the server in our research vessel named RV TUBITAK MARMARA. This on board cloud system provides service as a buffer zone between the vessel and the institution to avoid any data conflicts due to communication loss. It also allows the scientists in the different laboratories on board to gather and analyze data concurrently. Moreover, the on board cloud system uploads the most recent sea data to the office by synchronizing the two private cloud-based services when there is robust connection. In this way, the data is backed up in at least three different computers.

A local network is set up to have connection between the main server and all computers on board. The 3G mobile network is firstly connected to the main server and then the network is distributed to whole computers in laboratories by a network switch. To build a private cloud-based service, ownCloud which is client-server software is installed and configured on the server and other computers for file hosting service. This software works like well-known web cloud storage services e.g. Google Drive, Dropbox and so on. The best part to run ownCloud is that its openness avoids enforced quotas on storage space or the number of connected clients, instead having hard limits (like on storage space or number of users) defined only by the physical capabilities of the server. Thus, huge size of data can be shared with scientists, technicians and even ship crew easily. In other words, no one has to deal with the data sharing, storing and backing-up processes after the file is saved in an ownCloud folder.

To sum up, building a private cloud-service on board avoids any data conflict during offline, allows to gather and analyze data during cruise, creates a two-way communication between office and vessel efficiently and takes backup of sea data automatically.

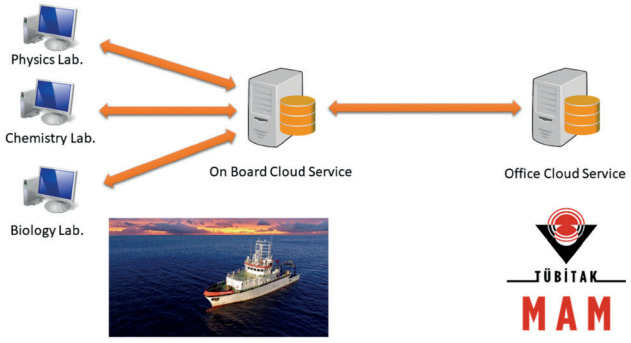


Fig. 1 - Private cloud-based services.

SatBałtyk - the sense of the Baltic environment

Mirosława Ostrowska, Institute of Oceanology Polish Academy of Sciences (Poland), ostra@iopan.pl
Mirosław Darecki, Institute of Oceanology Polish Academy of Sciences (Poland), darecki@iopan.pl

Design of the SatBałtyk System

The current state of knowledge enables supporting decisions undertaken by authorities governing activities in sea and coastal areas by generally accessible, cross-border complex information on state of the marine environment.

Up to date and detailed information, essential for assessing the current state of the environment, constituting the foundation for such decisions, will be highly reliable if:

- satellite data are used for the day-to-day monitoring of large sea areas;
- environmental information is updated and expanded by use of ecohydrodynamic models;
- the information from both sources is supported by in situ measurements from continuous monitoring systems (buoys, shore stations, drones) and research vessels.

Easy accessible and cost effective system operationally exploiting different sources of data and information on the Baltic Sea as well as expert knowledge was developed and launched in 2015 by SatBałtyk Scientific Consortium in frame of European Funds (Innovative Economy Programme). Variety of services provided by organisations contributing to project (ie the Institute

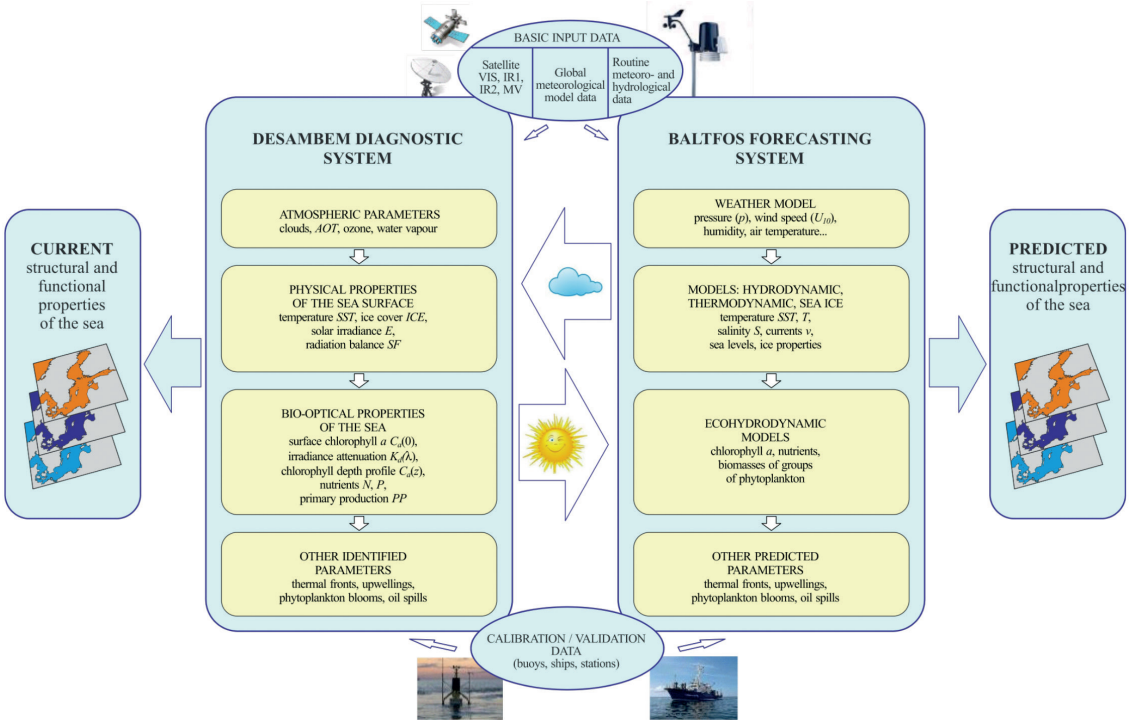


Fig. 1 - Deployment diagram of the SatBaltyk system.

of Oceanology PAN in Sopot - the project coordinator, the University of Gdańsk, the Pomeranian Academy in Słupsk and the University of Szczecin) are integrated on the unified platform and streamlined to the endusers.

To meet a present-day requirements for the continuous monitoring of the marine environment in SatBałtyk System two independent but cooperating subsystems were used: the diagnostic ones named the DESAMBEM, consists of several bio-optical algorithms, and the prognostic ones referred to as BALTFOS. The environmental parameters determined with the aid of these two systems complement one another: BALTFOS assimilates empirical data obtained from satellite information using the DESAMBEM algorithms, while at the same time filling in gaps in the DESAMBEM data when the satellite retrieval could not be made because the relevant areas were covered by clouds.

SatBałtyk Product Portal

Up to one hundred parameters currently available in the system have been divided into eight groups: 1. Atmosphere, meteorology, 2. Hydrology, 3. Ocean optics, 4. Radiation budget, 5. Sea water components, 6. Phytoplankton, photosynthesis, 7. Coastal zone, 8. Hazards. The values and maps of spatial distributions of these parameters are available in near real time to users on the website http://satbaltyk.iopan.gda.pl/ (Fig. 2).

Among the othes functionalities the system diagnoses and monitors various kinds of risk occurring in the Baltic Sea region. By supplying information on the regions where dangerous changes to the environment have been diagnosed, it will improve the capability of institutions to take decisions for preventing environmental disasters, mitigating their effects should they occur, tracking the changes and forecasting their possible development.

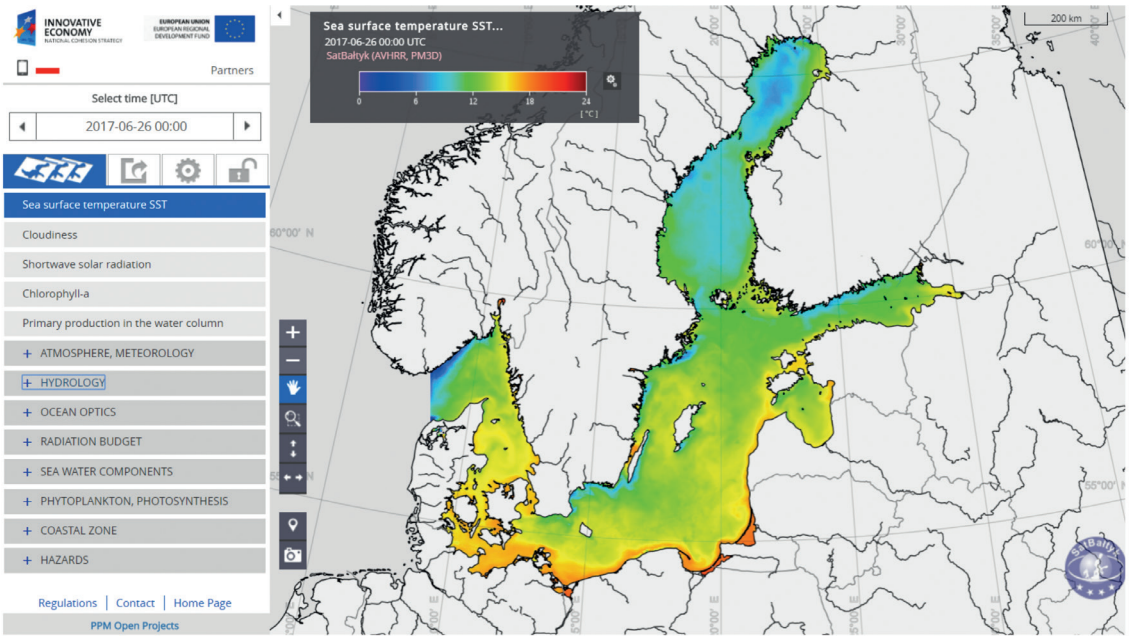


Fig. 2 - SatBaltyk Product Portal.

Significant feature of System SatBałtyk is facilitation of close cooperation between the project's designers and end users, focusing on identification of the individual needs of various groups of customers, tailor-made information can be provided to particular services, as support for their actions to protect the environment, improve security, manage maritime resources and plan military or rescue operations at sea and in the coastal zone.

Awareness of new data sources, access to them and efficiency of data exchange between institutions in the Consortium are to be increased. Reporting on the state of implementation of obligations arising out of legally binding regulations is to be supported by relevant governmental institutions. The system will become a significant link in the exchange of data and information among organisations and international observation systems of the Earth (Copernicus, GEO/GEOSS), the ocean (GOOS, EuroGOOS) and the Baltic Sea (BOOS). An independent data source, it may be of assistance in the arbitration of disputes relating to the influence on the Baltic Sea environment of users at different levels (individual/regional/national). The system's data resources, embracing long-term observations, can be used for developing adaptive strategies to climate change.

On the design of e-services for DANIBIUS RI

Vicente Gracia, Universitat Politècnica de Catalunya (Spain), vicente.gracia@upc.edu
Manuel A. Moreno, Universidad de Sevilla (Spain), mmoreno@gie.us.es
Joaquim Sospedra, Universitat Politècnica de Catalunya (Spain), joaquim.sospedra@upc.edu
Antonio Torralba, Universidad de Sevilla (Spain), torralba@us.es
Antonio Bejarano, Autoridad portuaria de Sevilla (Spain), abejarano@APsevilla.com
Agustín Sánchez-Arcilla, Universitat Politècnica de Catalunya (Spain), agustin.arcilla@upc.edu

The DANUBIUS Research Infrastructure e-services

DANUBIUS-RI will be a pan-European distributed research infrastructure devoted to interdisciplinary studies of large river–sea systems. It will enable and support research addressing the conflicts between society's demands, environmental change and environmental protection in river–sea systems worldwide. It will cover a wide range of disciplines working in the river-sea continuum such as biology, chemestry, geology, oceanography among many others. Because of that, DANUBIUS-RI will requiere the construction of different e-services to satisfy such scientific demands.

The main goal of this paper is to present the methodology used to define the e-services of DANUBIUS-RI. Two different river-sea continuums, the Ebro-Llobregat and the Guadalquivir, are used to analyse commonalities and specificities from the environmental observational and modelling perspective. The study analyses the existing river and coastal monitoring networks and the numerical modelling efforts to extract a common framework to be used in a future e-infrastructure service.

Review of existing e-services in RI

A total of 14 RIs has been reviewed covering the Energy, Environment and Social Science and Humanities disciplines. Most of the reviewed e-services have been tailor-made to cover the following needs: (i) Repository of documents; (ii) Remote experiment accesses; (iii) Data access; (iv) Education and training; (v) Networking and (vi) Modelling analysis.

A summary of the e-services and covered needs is given in Figure 1. Data is by far the most common product offered. The associated services range from simply downloading tools from web portals to more sophisticated environments in which the users are able to analyse data at different levels of complexity.The repository of public and internal RI documents requires Data storage and exchange services which are commonly provided through file hosting services or dedicated systems. The Database necessitates the same type of services and also a harmonization plus standardization effort that is achieved through dedicated Metadata services.

Data analysis is the function demanding more services. It requires the implementation of Data storage and Transfer services and depending on the degree of complexity offered, it may be accompanied by a computing capacity service such as HPC, the use of Modelling tools (requiring specific environments), a remote control access to perform experiments using RI facilities

		e-SERVICES											
		Blog	Social Media	Discussion Forums	Data storage	Data Transfer	Computing capacity	Modelling tools	Remote Control Systems	Dashboard	Metadata services	Data viewer	Wiki
NEEDS	Repository				X	X							
	Database				X	X					X		
	Data Analysis					X	X	X	X	X		X	
	Education		X	X					X	X		X	X
	Networking	X	X	X		X							

Fig. 1 - Common e-services and needs of reviewed RIs.

(physical or virtual), a dashboard space where the users can upload data and analyse the data generating their own products (maps or any other results) and a data viewer services commonly offering spatial and or temporal graphical capabilities.

The DANUBIUS e-service architecture

The main component parts of DANUBIUS-RI are: the Hub, which provides central direction and services: the Nodes, which provide expertise in key areas; the Supersites which are natural sites of scientific and societal importance for observation, analysis and research; the Data Centre which provides the data portal and the Technology Transfer Office in charge of the intellectual property generated by DANUBIUS-RI and generate a financial return to the RI.

Figure 2 shows the initial proposed structure for DANUBIUS-RI.

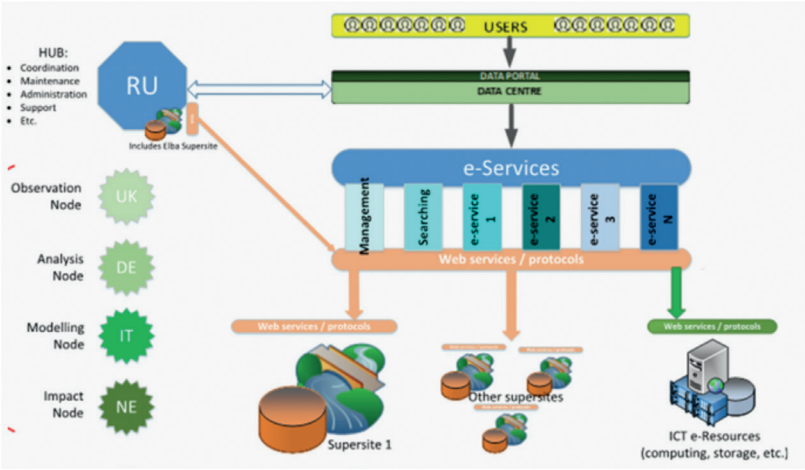


Fig. 2 - Initial proposed structure for e-services within DANUBIUS-RI.

The design of the e-services has also required the definition of the data handling and protocols within DANUBIUS-RI. This comprises the analysis of supported types and formats of digital data; the metadata description and formalizing documentation procedures; the review of processing data tools, equipment needed and data standardization and the interconnectivity design Data Centre sites storage capacities transfer requirements and protocols options and technologies.

The final paper will show the details of the approach used to identify and design the architecture of the e-services within DANUBIUS-RI.

Experiences from developing a Marine Spatial Management Tool for Norwegian Sea areas

Geir Arne Håland Nordhus, Norwegian Mapping Authority (Norway), norgei@kartverket.no

The Marine Spatial Managment Tool is the result of a governmental initiative based on the need for a more coherent and uniform geographic information content, suitable for underpinning tasks attached to marine spatial planning and marine management in norwergian sea areas :

- More effective updates of the management plans
- Better overview over political decisions and actions related to marine management
- Contribute to more transparancy, openness and increased involvment from the stakeholders

The purpose of the Marine Spatial Managment Tool is to support the marine spatial planning process with updated and reliable geospatial information.

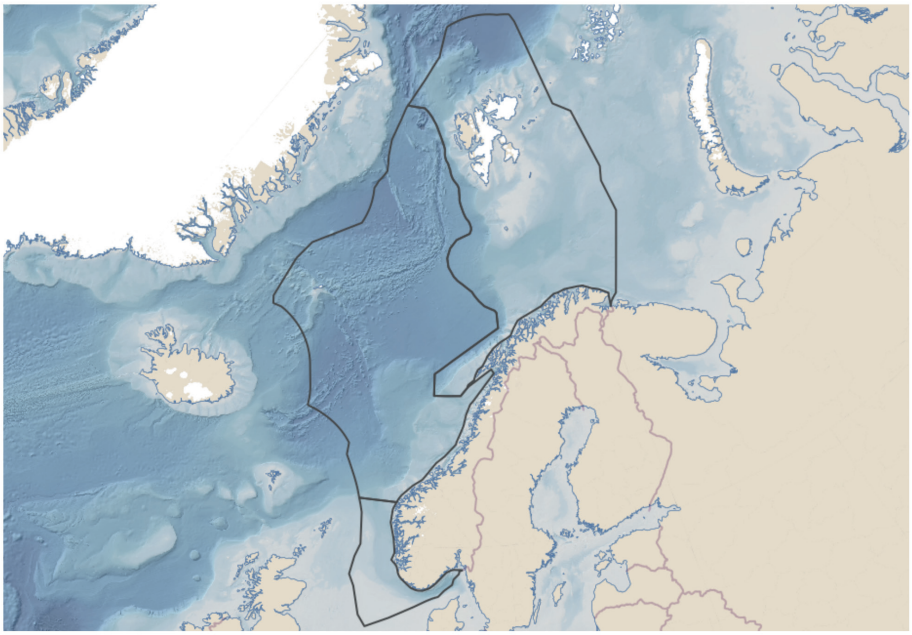


Fig. 1 - The Norwegian Government has developed integrated marine management plans for all Norwegian sea areas.

The presentation will focus on the experiences from the Marine Spatial Management Tool project with emphasis on a) the use of web services based on standards from Open Geospatial Consortium (OGC) through benefitial use of the national geoportal and infrastructures, b) challenges related to working cross-sectoral through an intergovernmental cooperation and c) cartography issues. In the end of the session there will be a live demonstration of the tool.

Data visualization and processing with Jupyter Notebook in SeaDataCloud Virtual Research Environment

Jani Ruohola, Finnish Environment Institute SYKE (Finland), jani.ruohola@ymparisto.fi
Sri Harsha Vathsavayi, CSC IT Center for Science (Finland), sriharsha.vathsavayi@csc.fi
Seppo Kaitala, Finnish Environment Institute SYKE (Finland), seppo.kaitala@ymparisto.fi
Christopher Ariyo, CSC IT Center for Science (Finland), chris.ariyo@csc.fi

As reproducibility is an essential part of scientific research, the failure to reproduce experiments can lead to unreliable results and false scientific findings. Currently, irreproducible research is a problem across all domains of science[1]. We are facing a similar replication crisis in marine research as well. With the complex and rapidly changing nature of computer environments and software libraries, the computational reproducibility is getting significant attention from the domain scientists. In order to reproduce experiments, we need to capture the data, software environment and the whole workflow. This study demonstrates the use of notebooks and virtual research environments (VREs) to create reproducible experiments. We also demonstrate how Jupyter Notebooks coupled with user's personal data storage solutions (like EUDAT's B2DROP[2]) can open new possibilities for doing marine research.

An example workflow for reproducible research in SeaDataCloud VRE (SDC-VRE) is presented, which is a work-in-progress service and aims at integrating different tools, such as Jupyter Notebooks, B2DROP and webODV[3], to offer a common workplace for all marine researchers to facilitate the marine research within and without institutional boundaries. Data exists in the user's B2DROP account, which is integrated with the VRE. Data processing, analysis and visualization is conducted within Jupyter Notebook (Fig. 1) and the results can be distributed for example as netCDF files. The computational environment exists within the VRE in a separate Docker container, which offers OS-level virtualization and user-space isolation. Moreover, the data and notebooks can be securely shared with the research groups via B2DROP for collaboration and for reproducing the workflow.

The purpose of this workflow is to allow easily setting up and distributing different computational environments, without the so-called dependency hell. Docker containers allow the lightweight distribution of the environments and research data, while Jupyter Notebooks document the whole workflow, with rich-text annotations and figures. The versatility of Jupyter Notebooks also offer the possibility for clear and easily distributed tutorials of different computational methods. In the era of increasing importance of data-driven research, these tools are invaluable. However,

[1] https://www.nature.com/news/1-500-scientists-lift-the-lid-on-reproducibility-1.19970

[2] https://eudat.eu/services/b2drop.

[3] https://webodv.awi.de/.

they offer many possibilities also for more traditional hypothesis-driven experimental research, since the documentation of the computational part is as important as the documentation of the experimental methods.

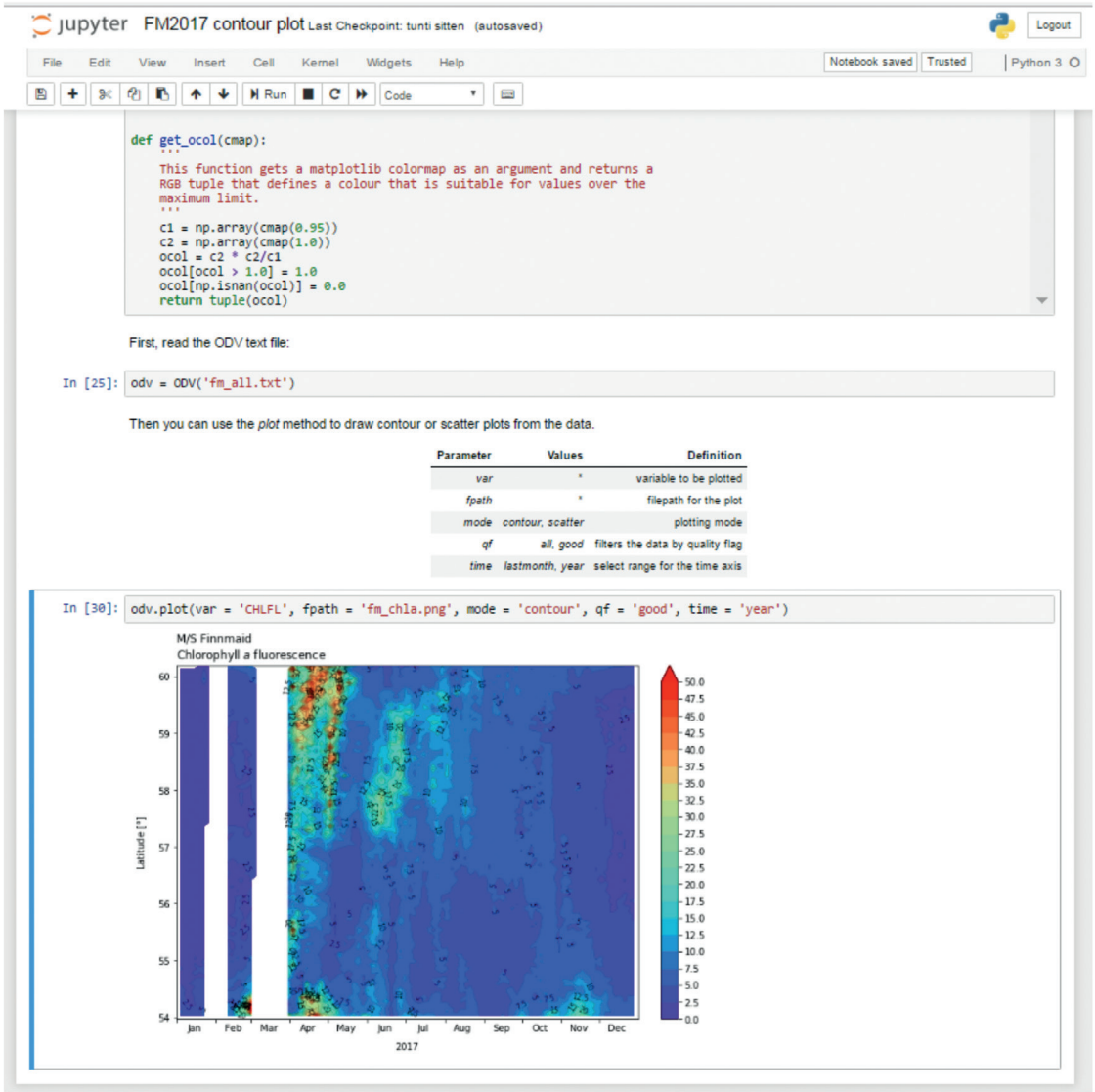


Fig. 1 - Contour plot of chlorophyll a fluorescence for a FerryBox route interpolated and visualized with Jupyter running inside the SDC-VRE. Inline figures, Markdown and LaTeX support offer an efficient method for documentation of the workflow.

Preparation of oceanographic data for international projects

Alexander Mikheev, ALL-Russian Research Institute of Hydrometeorological Information - WDC (Russia), amiheev@meteo.ru

Evgenii Viazilov, ALL-Russian Research Institute of Hydrometeorological Information - WDC (Russia), vjaz@meteo.ru

RIHMI-WDC participates in international Pan-European projects Sea Search, Black Sea Scene, SeaDataNet, EMODNet Chemistry, EMODNet Ingestion. These projects are part of the EU infrastructure and EMODnet thematic portals. The European system integration of oceanographic data, including the chemistry of the ocean, brings together data from National Oceanographic Data Centres and leading marine research organizations in Europe (http://www.seadatanet.org). The raw oceanographic data in RIHMI - WDC are presented in the national data storage format (Ocean PVM). Each cruise is stored in a separate file, consisting of a set of records of the following types:

- 0 record – general information about the cruise.
- 1 record – station information.
- 2 record – meteorological data.
- 3 record – hydrophysical and hydrochemical data.
- 4 record – pollution data.

To transfer data to the SeaDataNet portal, it is necessary to first convert the data from the Ocean PVM format to the SeaDataNet ODV transport format and prepare XML files for the common data index (CDI) metadata. For these purposes, the NIS_OPVM program developed by RIHMI-WDC is used (Fig. 1). To checking the received data, used the tool developed by the portal SeaDataNet - program Octopus.

Based on the approaches proposed by the SeaDataNet project, a number of thematic portals EMODNET on biology, chemistry, ocean Geology, bathymetry were created. Their overall concept - creating a common data index (metadata) and supporting data providers in ODV or NetCDF formats. For all time participation in projects RIHMI - WDC has transferred more than

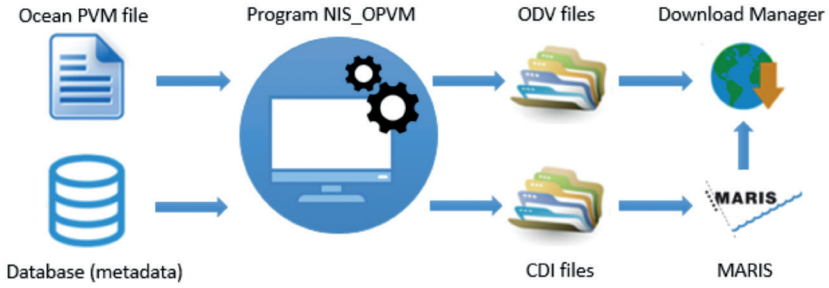


Fig. 1 - The scheme for transferring data to the SeaDataNet portal.

135 thousand stations for the Black sea, the Baltic sea, the Mediterranean sea and the North-Eastern part of the Atlantic Ocean.

The EMODnet Data Ingestion portal (https://www.emodnet-ingestion.eu/about) is a component of the present marine data management infrastructure , designed to download data to the EMODNet data portals. One of the important tasks of the EMODNet Ingestion project is the organization of automatic interaction of information systems. This is a new step in the development of international data exchange and data integration, which also means reaching a new level of inform to potential consumer. Within the framework of this project, RIHMI-WDC provides 11 open data sets with initial and climatic data. Most of the data sets are information resources of the Unified state system of information on the world ocean (ESIMO), which RNODC represents in international cooperation within the framework of the IOC UNESCO (Oceanographic data portal - ODP), World data center – B (WDC - B) "Oceanography". Several of the proposed data sets, regularly updated, which requires extra work. These are the TESAK operational data, obtained via the Global Telecommunication Network, the forecast data of the Russian Hydrometeorological Center, the trajectories of tropical cyclones for each month. If put this data through the portal EMODNet Ingestion, it will require a great deal of routine work. There are three options for automatically presenting these data sets on the EMODNet Physics Portal (Fig. 2):

- Include EMODNet links to visualization results prepared by resource authors.
- Development of web services or APIs for automatic acquisition of data for visualization on the EMODNet portal.
- Creating a map services in the WMS standard (OGC).

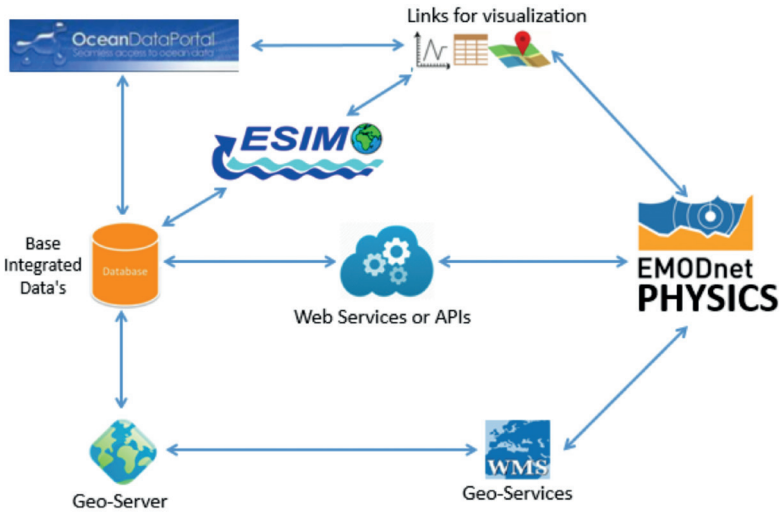


Fig. 2 - Options for presenting data in the EMODNet Physics portal.

RIHMI-WDC presented seven data sets to the EMODNet Ingestion project at the moment.

Automation of preparation CDI and ODV files has allowed to reduce essentially labor input of work on data integration. Increasing the level of automation in the provision of quickly updated data sets on the EMODNet Physics portal simplifies the process of data assimilation and interaction of European and Russian information systems.

"See The Sea" – new opportunities for distributed collaboration aimed at solution of oceanographic problems using remote sensing

Evgeny Loupian, Space Research Institute of RAS (Russia), evgeny@d902.iki.rssi.ru
Olga Lavrova, Space Research Institute of RAS (Russia), olavrova@iki.rssi.ru
Alexandr Kashnizky, Space Research Institute of RAS (Russia), kashnizky@gmail.com
Ivan Uvarov, Space Research Institute of RAS (Russia), uvarov@d902.iki.rssi.ru

General description

The "See the Sea" (STS) is an information system developed by the Space Research Institute of the Russian Academy of Sciences (IKI RAS, Moscow, Russia), to study various processes and features in the ocean and sea using diverse satellite data. The STS is the toolkit to work with remote sensing data and to analyze results. Its key advantage is in an ability to perform a complex analysis of data varying in physical character, spatial resolution and units of measurement. The system also provides access to long-term (over 20 years) distributed archives of satellite data for the entire Eurasia from 1995 to the present, including Synthetic Aperture Radar (SAR): Envisat ASAR, ERS-2 SAR, Sentinel-1A, -1B; optical systems: MODIS Terra/Aqua, MERIS Envisat, TM/ETM+/OLI Landsat-5/7/8 satellites series, MSI Sentinel-2A, -2B, OLCI Sentinel-3 and hyperspectral data from Hyperion and HICO systems. Currently, the volume of the archive is over 2 PB and it is updated with real time data daily. In addition, the multi-year weather data are available on-line. The STS is an open system capable to incorporate any required data (such as altimeter data, buoy data, in situ measured data, and other).

The STS capabilities

The main goal of the system is to develop the toolkit for efficient work with diverse data involved into analysis of various processes on the ocean surface. First of all, the STS satellite service allows to perform data transformation and analysis without any preliminary data download, but by accessing the data "on the fly" in the STS archives. It is important that users of the system have no need to install any special software to their PCs. They can also avoid downloading large volume of data for processing, which is especially significant when data processing from new sensors, such as Sentinel 1 & 2 (one MSI Sentinel-2 image may take up to 7 GB). Using the STS is also possible to analyze data provided by the system using a standard Internet browser. The STS automatically downloads complete satellite data as soon as they appear on appropriate servers, for example for Sentinel 1 & 2 data in the Sentinels Scientific Data Hub (SSDH) (https://scihub.copernicus.eu/). The received data are transferred into the UTM/WGS84 projection. From the resulting file, a number of quick view files are generated containing double-thinned georeferenced data with a pyramid of resolutions. These datasets accompanied by metadata are then transferred to the Center for Collective Use of IKI RAS ("IKI Monitoring") archives

where they are stored and from where delivered on request to various information systems. For displaying data in cartographic interfaces of various information systems (including the STS), the online data display unit compiles from the archived data information products (layers) using either the basic resolution or quick look files depending on the scale requested. The required data, so called virtual data products, are actually being prepared on-line when user requests it from the band data stored in the archives. The database stores rules for their formation in a special table structure. The rules allow building of arbitrary band combinations with different normalizations, as well as transformations according to formulas written in a meta-syntax and pixel by pixel transformations. Such approach significantly reduces storage requirements for the set of data products and enables easy addition of new products and modifications of existing ones without extra processing or enlarging physically stored data.

The STS provides a comprehensive toolkit for data processing and analysis; and has capabilities of cartographic web-interface similar to desktop GIS applications. This toolkit is also based on the technology developed by the IKI RAS. Image Algebra, Classification, Image Color Enhancement, Indexing, and some other tools are presently available in the STS. The Image Color Enhancement allows to correct image histograms and improve display characteristics of particular elements. The Image Algebra implements arithmetic, logic and various mathematical transformations according to formulas defined by the user directly via the interface. The Indexing is intended for calculation of various predefined indices. With the Classification one can identify data types and image peculiarities using various supervised and unsupervised classification methods and applying masks. It should be noted, this tool has a high flexibility: it is possible to classify simultaneously data from different sensors, for example, Landsat and Sentinel-2A, taken at different time, data of any bands by user choice, and more.

The STS satellite service provides a powerful toolkit not only for access to satellite data and various information products derived from their processing, but also provides an opportunity for various types of specialized analyses of information. An integrated analysis of data of different physical nature, spatial resolution, dimension and time of acquisition is also available. The STS offers a comprehensive description of different phenomena and processes, their quantitative and qualitative assessment as well as tools to investigate conditions of appearance and development of these phenomena. Studies of spatial and temporal characteristics of distributions of different phenomena are also provided by STS.

Examples of oceanographic tasks solved by means of the STS

1. ***Determination of zones of ecological risk affected by anthropogenic pollution of the sea surface.*** In particular idetification of areas of oil pollutions discharges from ships. For the Black, Baltic and Caspian seas, the main areas of pollution were evaluated and mapped.
2. ***Study of river discharges.*** Terrigenous sediments enter the sea with river plumes of mesoscale structures with sharp boundaries. Not only salinity and temperature in a plume are different from surrounding seawater, but also turbidity in a plume is greater due to increased content of suspended particulates and dissolved organic matter. Every watershed forms a specific individual range of particulate solids. It helps in identifying boundaries of and individual outflow plume. True-color composites of visible satellite data are best suitable for discrimination of sea areas with different optical properties, associated with a concentration of suspended matter. With the Classification tool of the

STS, we have identified and estimated areas with different turbidity. The STS provides methods of a combined analysis of satellite and meteorological data, to reveal an influence of meteorological parameters such as wind and precipitation on formation and distribution of river plumes.

3. ***Study of mesoscale and submesoscale hydrodynamic processes.*** On the basis of long-term series of satellite and meteorological data, hydrodynamic processes of different spatial scales such as vortices, internal waves and fronts were studied. The STS helped to identify areas of their permanent manifestations and mechanisms of their generation. The latest satellite data of high spatial and temporal resolution allow estimating not only spatial but also dynamic characteristics of these processes. The joint analysis of the various satellite data provided in the STS made it possible, for example, to assess the velocity of propagation of these hydrodynamic processes.

The development of the See the Sea information system is supported by theme “Monitoring”, registration no. 01.20.0.2.00164. The use of the STS for a study of the state and pollution of seas was carried out with the financial support of the Russian Science Foundation, grant 14-17-00555. The information system is functioning with use of the data from the “IKI Monitoring” Center for Collective Use.

Data systems supporting marine protected area management in Parks & Wildlife Finland

Lasse Kurvinen, Parks & Wildlife Finland (Finland), lasse.kurvinen@metsa.fi

Parks & Wildlife Finland is part of the state of enterprise Metsähallitus and is, among other tasks, responsible for the management of National Parks and other state owned protected areas, including several marine protected areas (MPAs). Parks & Wildlife Finland manage over 50% of Finland´s MPA areal, including National Parks, a UNESCO World Heritage site, N2000-, HELCOM MPA- and Ramsar-sites, as well as other protected areas, such as seal protection areas.

In order to effectively manage these areas several kinds of data are needed and the data used should be of good quality, up to date and easily accessible.

At Parks & Wildlife Finland we have in the recent years moved towards a concentrated (spatial) data management system, to support the management of various types of protected areas. The system consists of several sub-parts which are integrated in one framework. The main part (SASS) is meant for planning and assessment of protected areas. Here one can make e.g. management plans, N2000-status evaluations etc. The other parts contain information on species (LajiGis), habitats and biotopes (SAKTI), basic information of protected areas (SATJ) and cultural heritage sites and recreational structures (PAVE). These are built in way as to make it possible for the information to move between the different parts, so that when you e.g. start making management plans for an MPA, you can get all the species, biotope etc information for the specific area from the other parts.

The sub-parts are also more than simple storages for data, as these can also be used for various kinds of planning and monitoring tasks. You can e.g. make plans for eradicating occurrences of invasive species or management plans for maintaining cultural biotopes.

At the moment the system is open only for the environmental administration in Finland, but in the future, there are plans to open at least some of the data, if not parts of the systems, for the bigger public as well.

ODYSSEA: A novel, interoperable platform for products and services in the Mediterranean Sea - system architecture and design

Nicolas Granier, Collecte Localisation Satellites (France), ngranier@cls.fr
Sylvan Marty, Collecte Localisation Satellites (France), smarty@cls.fr
Simon Keeble, Blue Lobster (United Kingdom), simon@bluelobster.co.uk
Carlos Figueiredo, Edisoft (Portugal), carlos.figueiredo@edisoft.pt
Eduard Huguet, GTD System & Software Engineering (Spain), eduard.huguet@gtd.eu
Adelio JR Silva, Hidromod (Portugal), adelio@hidromod.com
Ghada El Serafy, Deltares (The Netherlands), Ghada.ElSerafy@deltares.nl
Georgios Sylaios, Democritus University of Thrace (Greece), gsylaios@env.duth.gr

Introduction

ODYSSEA (http://odysseaplatform.eu/) is an EU H2020-funded project aiming to develop, operate and demonstrate an interoperable and cost-effective platform that fully integrates networks of observing and forecasting systems across the Mediterranean basin, addressing both the open sea and the coastal zone.

The project involves 28 partners (universities, research centres, international organizations, NGOs and private companies) working systematically towards the development of the ODYSSEA platform to:

- Integrate marine data from existing databases maintained by Earth Observing facilities, scientific networks, agencies, public authorities, and institutions of Mediterranean EU and non-EU countries (CMEMS, GEOSS, GOOS, EMODNet, ESFRI, Lifewatch, Med-OBIS, GBIF, AquaMaps, Marine IBA e-atlas, MAPAMED and others with marine and maritime links);
- Receive and process novel newly produced datasets (through models, remote sensing and on-line sensors) from nine prototype Observatories, established along the continental shelf of the Mediterranean Sea, especially at areas with intense human activity but also at marine protected zones;
- Transform marine data into meaningful information, ultimately developing, testing, validating and disseminating marine data products and services to end-users and stakeholders from a diverse spectrum of marine and maritime sectors (mariculture, shipping, oil and gas exploitation, port management, civil protection, etc).
- Stimulate Blue Growth throughout the Mediterranean basin, creating businesses, advancing science and supporting the societal use of digital information.

When operational by 2021, the final platform will provide easy discovery and access to marine data and derived products to a variety of users to improve knowledge and decision-making capabilities in the Mediterranean.

ODYSSEA Platform Novelties

As a platform, the ODYSSEA system will encompass and offer to its users a series of novelties classified as service-level, platform-level and information-based novelties. In terms of service-level innovations, ODYSSEA platform will provide the capability to search and retrieve marine data, but also to deliver products and services built upon the existing datasets. In terms of information-based novelties, ODYSSEA platform will aggregate physical, chemical, geological, biological, biodiversity and fisheries datasets through ontology approach and semantic information fusion to provide functionalities and services.

The set-up of the platform will allow the use of agile methods enabling short loops between user's feedback and the development team. The platform will be fully decentralised, event-driven and easily deployable on a cloud system for big data processing (e.g., DIAS) thanks to docker and kubernetes technologies.

ODYSSEA Platform Architecture

The ODYSSEA platform is designed to be expandable, scalable, flexible, transferable and interoperable. The main elements and components of ODYSSEA platform are illustrated in Fig. 1, involving: a) the Data Collection Component; b) the Data Pre-processing Component; c) the Data Storage Component; d) the Post-processing Component; e) the Product Factory Component; f) the Authentication and User Management Component; g) the Cataloguing Component; h) the Data Output and User-Interface Component.

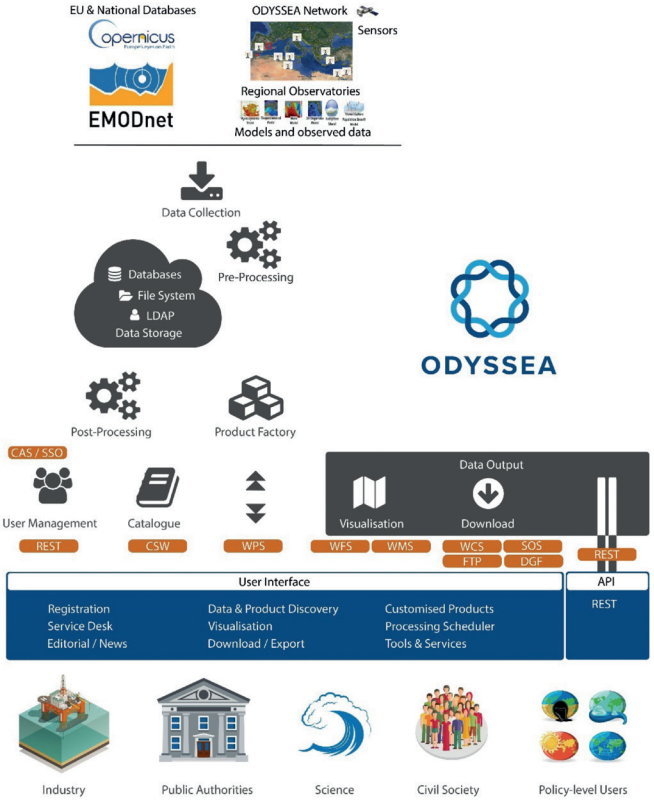


Fig. 1 - ODYSSEA Platform Architecture.

Tools for discovering and managing oceanographic data of Spanish Institute of Oceanography

Olvido Tello-Antón, Spanish Institute of Oceanography (Spain), olvido.tello@ieo.es
Luis Miguel Agudo Bravo, Spanish Institute of Oceanography (Spain), luismi.agudo@ieo.es
Jon Garrido, Bilbomática (Spain), jon.garrido@gmail.com

The Spanish Institute of Oceanography (IEO) is a Public Research Organization created in 1914, currently attached to the Ministry of Economy and Competitiveness. The IEO focuses its efforts on scientific knowledge of the oceans, the sustainability of fisheries resources and the marine environment.

Due to the large amount of information collected over the 100 years of IEO history, there is a clear need to organize, standardize, integrate and relate the different databases and information.

In 1997 began the first initiative to organize the Spanish Institute of Oceanography spatial information in a Geographical Information System (GIS). In 2010 began the development of Spatial Data Infrastructure of IEO in order to facilitate the access to the spatial data of IEO. This infrastructure involves a metadata catalogue, data viewers with interactive tools and Web Services of IDEO.

The data, metadata and Web services of IEO are following INSPIRE directive (2007/2/EC) standards.

In order to facilitate access to the Spatial Data Infrastructure of IEO, a Geoportal was developed. http://www.geo-ideo.ieo.es/geoportalideo/

Fig. 1 - IEO Geoportal.

The Viewer Base integrated on the Spatial Data Infrastructure of IEO has useful and advanced tools to discovering and managing the oceanographic data of IEO.

These tools allow display, select simbology and consult spatial and alphanumeric data. Besides allow insert new data, 3D displays, make several charts (bar, pie, scatter, etc.).

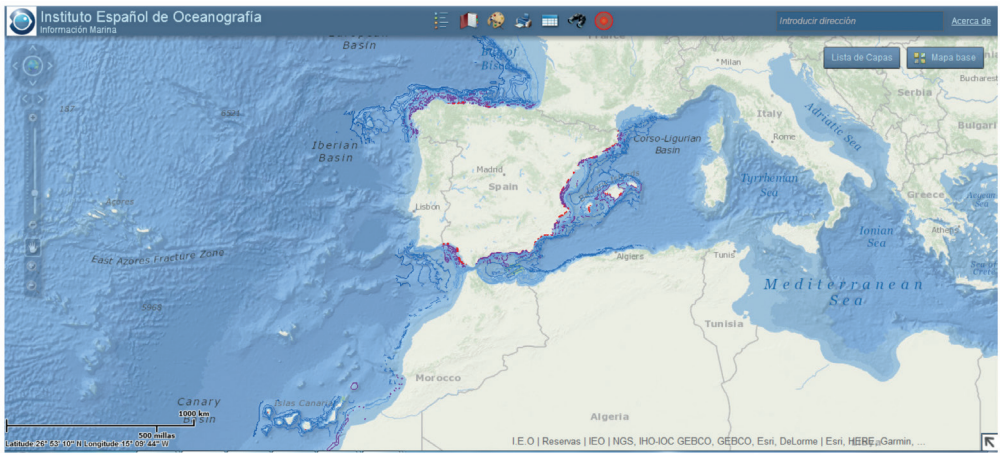


Fig. 2 - IEO Base Viewer.

Currently the IEO is participating in several European initiatives, especially in marine spatial planning projects and in several lots of EMODNET initiative (biology, chemistry, physics and bathymetry) with the data accessible from Data Portal EMODNET http://www.emodnet.eu/portals, among others. From GIS department it is working on EMODNET High Resolution Seabed Mapping http://www.emodnet-bathymetry.eu/. The main aim of this project is the building a high resolution DTM for European Seas. The currently project will allows to download a DTM with 100 meters of resolution.

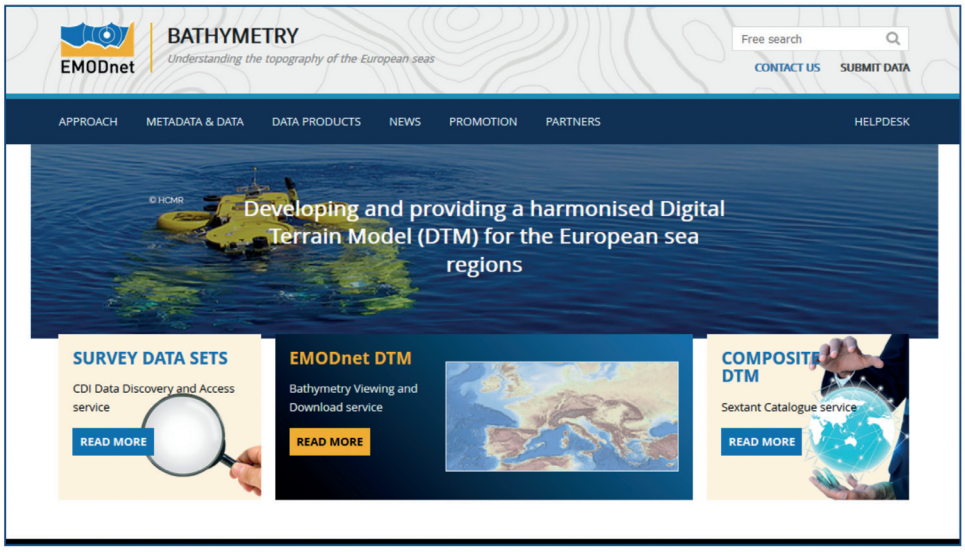


Fig. 3 - EMODNET Bathymetry Portal.

Integrated search and analysis of multidisciplinary marine data with GeRDI

Ingo Thomsen, Software Engineering, Kiel University (Germany), int@informatik.uni-kiel.de
Whilhelm Hasselbring, Software Engineering, Kiel University (Germany), wha@informatik.uni-kiel.de
Jörn Schmidt, Economics, Kiel University (Germany), jschmidt@economics.uni-kiel.de
Martin Quaas, Economics, Kiel University (Germany), quaas@economics.uni-kiel.de

An exemplary research question: "How marine fisheries impact on global food security up to 2050"

Multidisciplinary research usually requires data from more than one data repository that has to be retrieved and analyzed. Fig. 1 outlines the dataflow addressing such an exemplary research question: data from multiple discipline-specific repositories was aggregated and analyzed. The results were published as part of a WWF report (Quaas *et al*., 2016).

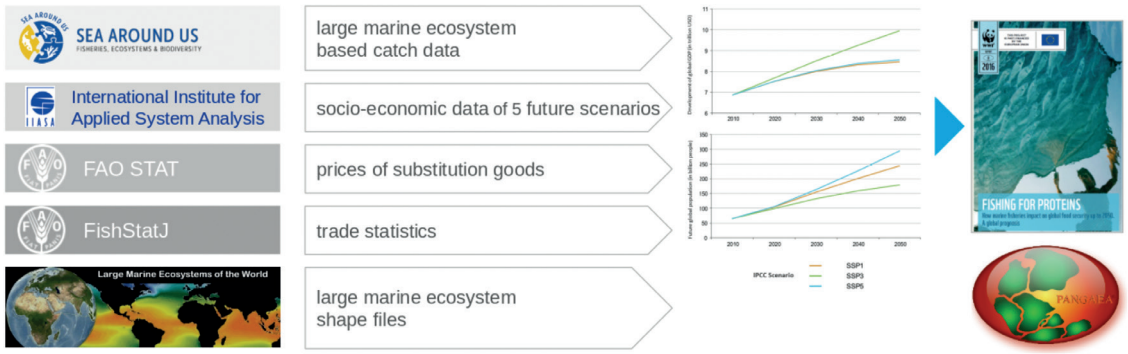


Fig. 1 - Dataflow for the creation of the WWF report.

In this example, parameters for bio-economic fishery models are statistically estimated using catch and price data from three main sources: Sea Around Us (Sea Around Us, 2018), the FAO database FAOStat (FAOSTAT, 2018) and the FishStatJ fishery databases (FishStat, 2018). Information on the area of Exclusive Economic Zones (EEZ) and Large Marine Ecosystems (LMEs) are taken from the LME database (NOAA, 2018). The model is finally based on scenarios for total expenditures for protein-rich food, and the availability of protein-rich food other than wild capture fish using GDP and population data derived from model output from IIASA using the Shared Socioeconomic Pathways from IPCC (O'Neill *et al*., 2014).

Development of a generic research data infrastructure - driven by research questions

The GeRDI project (www.gerdi-project.de, Grunzke *et al*., 2017) focuses on the development of a sustainable Generic Research Data Infrastructure. Its goal is to enable scientists to search,

use and re-use external research data. In the current pilot phase, the software development is driven by research questions – including the exemplary one above. These questions originate from participating communities in various research disciplines – marine sciences, but also digital humanities, bioinformatics, and others.

The GeRDI services are implemented in a modular manner as microservices (Tavares de Sousa *et al*., 2018) as outlined in Fig. 2. They communicate through well-defined protocols. Software and protocols are published as open-source. This offers the potential to "plug-in" and to replace parts with your own specialized services.

GeRDI offers an integrated web-interface to search repositories (for instance with ocean related data) whose metadata was previously harvested - preferably employing an open protocol (OAI-PMH, 2018). An established metadata scheme (DataCite, Metadata Working Group, 2017) was adapted as a base for the search index. The faceted search – filtering using time, categories, etc. – can also be based on geolocation with the help of an interactive map. The results can be saved as a set of bookmarks that offer the download links and/or instructions for accessing the data. This set can be saved, modified and re-used, which supports repeatability and sharing of research workflows and experiments.

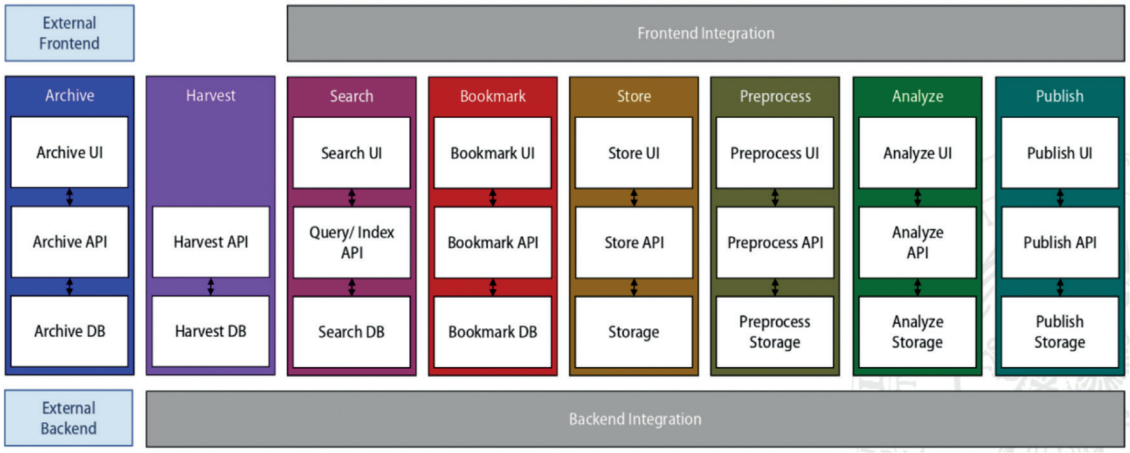


Fig. 2 - GeRDI Microservice-based software architecture [8].

The first three services in Fig. 2 are based upon metadata: (1) Harvest existing repositories, (2) generic keyword-based Search, and (3) a persistent Bookmark of selected data sets. The example in Figure 1 indicates that the research results are published in the PANGAEA repository (PANGAEA, 2018). This leads to the extended services in GeRDI which handle the actual data (in addition to the metadata): Store bookmarked data sets (locally or in a cloud), Preprocess the data as preparation and then Analyze it ("gaining new insights") and finally Publish it. Archive represents the data repositories that offer upload facilities, thus closing the research data cycle.

References

Quaas M.F., Hoffmann J., Kamin K., Kleemann L. and Schacht K., 2016. *"Fishing for Proteins. How marine fisheries impact on global food security up to 2050. A global prognosis"*, WWF, Hamburg.

Sea Around Us, University of British Columbia, www.seaaroundus.org (retrieved 2018).

FAOSTAT - Food and Agriculture Organization Corporate Statistical Database, www.faostat.org (retrieved 2018).

FishStat - software for fishery statistical time series. FAO Fisheries and Aquaculture Department, www.fao.org/fishery/statistics/software/fishstatj/ (retrieved 2018).

NOAA - National Oceanic and Atmospheric Administration, www.st.nmfs.noaa.gov/ecosystems/lme (retrieved 2018).

O'Neill B.C., Kriegler E., *et al.* (2014). *"A new scenario framework for climate change research. The concept of shared socioeconomic pathways."* Climate Change 122(3):387–400.

Grunzke R., Adolph T., *et al.* (2017), *"Challenges in Creating a Sustainable Generic Research Data Infrastructure."* Softwaretechnik-Trends, 37 (2). pp. 74-77.

Tavares de Sousa N., Weber T., *et. al.* (2018), *"Designing a Generic Research Data Infrastructure Architecture with Continuous Software Engineering"*, In: 3rd Workshop CSE 2018, pp. 85-88.

"Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH)", www.openarchives.org/pmh/ (retrieved 2018).

DataCite Metadata Working Group, "DataCite Metadata Schema 4.1", www.schema.datacite.org/meta/kernel-4.1/ (retrieved 2017).

PANGAEA - Data Publisher for Earth & Environmental Science, www.pangaea.de (retrieved 2018).

The EVER-EST Virtual Research Environment (VRE): outcomes and solutions for Earth Science

Valentina Grande, ISMAR Consiglio Nazionale delle Ricerche (Italy), valentina.grande@bo.ismar.cnr.it
Giorgio Castellan, ISMAR Consiglio Nazionale delle Ricerche (Italy), giorgio.castellan@bo.ismar.cnr.it
Francesco De Leo, ISMAR Consiglio Nazionale delle Ricerche (Italy), francescodeleo64@yahoo.it
Sergio Ferraresi, Meteorological Enviromental Earth Observation MEEO S.r.l. (Italy), ferraresi@meeo.it
Federica Foglini, ISMAR Consiglio Nazionale delle Ricerche (Italy), federica.foglini@bo.ismar.cnr.it
Jose Manuel Gómez, Expert System (Spain), jmgomez@expertsystem.com
Simone Mantovani, Meteorological Enviromental Earth Observation MEEO S.r.l. (Italy), mantovani@meeo.it
Fulvio Marelli, Terradue (Italy), fulviomarelli@me.com
Raul Palma, Poznań Supercomputing and Networking Center (Poland), rpalma@man.poznan.pl

The EVER-EST project developed a Virtual Research Environment (VRE) to manage the full research lifecycle in Earth Science: from discovery and access of data input, to the coding necessary to extract the information, till to the sharing of methodologies and results.

The logic behind EVER-EST is to put the scientist at the center and to bridge the technological and knowledge gap and barriers for open science, application of EOSC principles and digital innovation. The EVER-EST VRE enables FAIR services to improve findability, accessibility, interoperability and reusability of research data, processes and results in a web environment. EVER-EST ensures to its users the functionalities which are needed to search, access and process Earth Science data: but most importantly it provides services to re-use, preserve and share among researchers both data and scientific processes. This is enabled by the adoption – for the first time in Earth Science – of the Research Object paradigm and related technologies.

The Reasearch Object (RO) aim to account, describe and share everything about the research, including how those things are related. The RO model takes the central point to encapsulate all the resources relevant to the scientific work (data, live code, workflows, results, documents) in a single information unit [http://www.researchobject.org].

The RO model has been adapted to the Earth Science according to the user needs provided by the four different Virtual Research Communities (VRCs) engaged in the EVER-EST project (Land Monitoring, Supersites, Natural Hazard and Sea Monitoring). Nevertheless the four communities use the VRE for different goals (e. g. change detection on land, monitoring, risk assessment, marine habitat mapping), with different data, processes and results, working with the same facilities led to a cross-fertilisation process between VRCs, generating new knowledge.

The CNR-ISMAR represents the Sea Monitoring community and, in this view, developed case studies providing practical methods, procedures and protocols to support coherent and widely accepted interpretation of Good Environmental Status (GES) in the Marine Strategy Framework

Directive (MSFD). In this context, we present the Research Objects implemented so far, focusing on methodologies and results related to benthic habitat mapping such as Cold Water Corals habitat suitability models and seafloor roughness extractions using the EVER-EST VRE platform.

The Sea Monitoring portal provides the main user web interface to create and share Earth Science ROs, to discover data, to access, to process and visualize services rely on OGC standards (OpenSearch, Web Coverage Service, Web Processing Service, Web Map Service), to manage Research Objects, and finally, to execute remote workflow implemented via Taverna [https://taverna.incubator.apache.org]. Moreover, the VRE provides different user interfaces such as: *Collaboration spheres*, for the visualization of correlation between similar objects (e.g., users, Research Objects) based on collaborative filtering and versatile keyword content-based recommendations; *RoHub*, the reference platform for Research Object management supporting the preservation and lifecycle management of scientific investigations, research campaigns and operational processes; and *Jupyter Notebook* a web-based application suitable for capturing the whole computation process: developing, documenting, and executing code, as well as communicating the results.

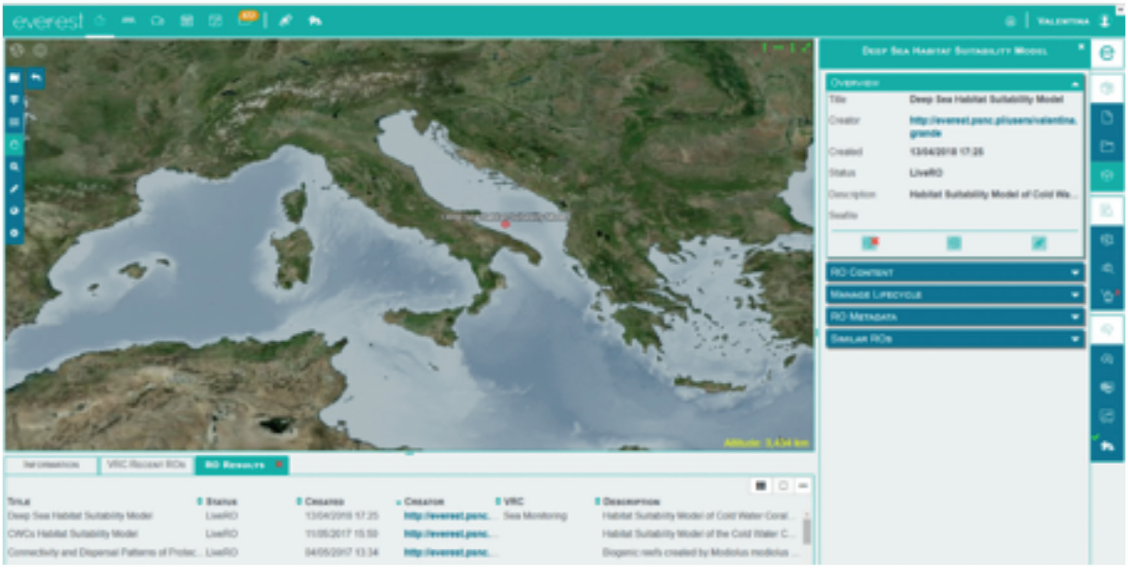


Fig. 1 - Overview of the EVER-EST VRE (https://vre.ever-est.eu/demovrc/).

Data integration of Scuba benthic communities' photography surveys with physico-chemical multiparameter sensor platform as a means for more efficient coastal ecosystems studies

Dimitar Berov, Di Ni Mar Ltd, IBER-BAS (Bulgaria), dimitar.berov@gmail.com
Nikolay Berov, Di Ni Mar Ltd (Bulgaria), nberov@gmail.com

Classical benthic ecology surveys based on in-situ sampling and laboratory analyses serve as basis for modern coastal ecosystems monitoring programs and are the basis for the EU Water Framework Directive and Marine Strategy Framework Directive monitoring programs. Such studies require expertise in benthic ecosystems and the taxonomy and biology of the flora and fauna in the coastal zone, and are time-consuming and slow to obtain results. In these studies, environmental data on physical and chemical parameters of the marine environment is usually added post-factum in the data analysis and it usually comes from separate surveys, often from a different location and time.

Modern advances in digital underwater photography and photogrammetry allow a rapid and more efficient survey of benthic communities, especially in the case of macroalgal and zoobenthic communities on hard bottom. The usage of photography as a non-destructive method for sampling allows the surveyors to gather significant amount of data in the limited time available underwater (see Berov *et al*. 2016, 2018), which could then be quickly processed with image analysis software (CPCE, PhotoQuad etc.) and analyzed in statistical and geographical information software packages (e.g. Primer-E, R, SPSS, Arc GIS), thus significantly shortening the time needed for sampling, data analysis and results generation. Recently developed multiparameter sensor platforms are compact enough to be used in surveys conducted from small research vessels and by scuba divers, opening up the opportunities for measuring relevant environmental parameters (e.g. depth, temperature, light intensity, pH, salinity, oxygen contents, chlorophyll-a, and DOM concentration) in the exact locations of the benthic ecology surveys, thus putting the obtained biological results in the correct oceanographical context. The purpose of this work was to develop and test a system integrating data from photogrammetry benthic surveys, GPS navigation data and physico-chemical parameters measurements, with possibilities for results exports in statistical packages, GIS, and oceanographical data visualization programs (Ocean Data View).

The system is based on a high-resolution full-frame photo camera (Canon 5D Mark III in an Ikelite housing), equipped with strobes (Ikelite DS-161), a multiparameter probe (MpX NKE Instrumentation with sensors measuring depth, temperature, salinity, oxygen concentration, pH, chl-a fluoresence), a sensor module for PAR (NKE sPAR logger), and a surface-towed GPS buoy. The different components of the system can be used modularly, allowing several configurations

of usage: for (1) georeferenced scuba divers photo surveys of benthic ecosystems combined with measurements of physico-chemical parameteres, and (2) autonomous deployment of the photo- and multiparameter sensor platform from a boat for surveys of the benthos covering larger areas and at depths higher than recommended for repetitive scuba dives.

The integration of the data is carried out in a MS Access SQL database. The georeferening of the photos is based on time-synchronization of the data logs from the GPS with the EXIFF time stamp of each photo, with different options for photo-GPS delay compensation. Depth readings for each photo are extracted from the CSV-format logs of the multiparameter platform, or alternatively – from the logs of the diving computer used by the diver operating the camera. The georeferencing and integration of the data logs from the multiparameter platform is again based on time synchronization of the data logs. Various modules for metadata attributes for each survey are also included in the database. The system allows data exports in tabular format for import in ArcGIS (georeferenced photos, physico-chemical data, metadata), Ocean Data View (physico-chemical data), as well as statistical packages (e.g. MS Excel, R, Primer-E).

Initial tests of the system in specific case studies in the Black Sea coastal zone are currently underway. Our first results show significant improvement of the efficiency of coastal benthic mapping surveys, water quality monitoring and environmental impact assessment studies in comparison with previous surveys based on separate and non-integrated instruments and data sources. Data output is used in surveys of the good ecological state and modelling of the distribution of the hard bottom infralittoral habitats along the Bulgarian Black Sea coast (MSFD Descriptor 1.6), monitoring of the recovery of a coastal area after the completion of a waste water treatement plant, and possible environmental impacts from Mytilus galloprovincialis black mussel farms.

References

Berov D., Hiebaum G., Vasilev V., Karamfilov V., 2016. *An optimized method for scuba digital photography surveys of infralittoral benthic habitats. A case study from the SW Black Sea Cystoseira-dominated macroalgal communities.* Underw. Technol. 34, 11–20. https://doi.org/https://doi.org/10.3723/ut.34.011

Berov D., Todorova V., Dimitrov L., Rinde E., Karamfilov V., 2018. *Distribution and abundance of phytobenthic communities: Implications for connectivity and ecosystem functioning in a Black Sea Marine Protected Area.* Estuar. Coast. Shelf Sci. 200. https://doi.org/10.1016/j.ecss.2017.11.020

Real-time lossless and lossy compression of MBES water column

Jordi Portell, Dept. Quantum Physics and Astrophysics, Institute of Cosmos Sciences - ICCUB, University of Barcelona - IEEC-UB (Spain) and DAPCOM Data Services (Spain), jportell@fqa.ub.edu
David Amblas, Scott Polar Research Institute, University of Cambridge (United Kingdom) and CRG Marine Geosciences, Dept. of Earth and Ocean Dynamics, University of Barcelona (Spain), damblas@ub.edu

Multibeam echosounders are excellent for the efficient and accurate mapping of the seafloor, also allowing to collect water column acoustic images [1]. The latter, however, leads to vast amounts of data, for which existing compression algorithms do not provide good enough results. In this work we present a new data compression solution, FAPEC, inherited from the Gaia space astrometry mission.

FAPEC is a highly optimized and versatile two-stage data compression system [2]. The first stage can be adapted to the specific kind of data being handled, aiming at the output of small signed values. These are then compressed by the second stage, an outlier-resilient entropy coder which can outperform an optimum Huffman coder. We have implemented a tailored pre-processing stage for Kongsberg MBES water column data. It arranges backscatter samples in a two-dimensional matrix and codes the differences between neighbour values. It provides lossless compression, allowing to recover the original data file without any loss. The user can also select lossy compression with different quality levels, allowing to achieve a larger reduction in data volume at the cost of some degradation in the image quality. This is achieved by a non-biased quantization of backscatter data samples, reducing the number of grey shades in the reconstructed image but keeping the full image resolution. It is a very interesting option which still allows detecting fish shoals, gas seeps or sunken structures, for example. FAPEC, including this tailored water column stage, is fully operational and provides block-based operation, multi-threading and encryption. Its ANSI C implementation makes it portable, including ARM-based computing platforms such as those present in AUVs.

Test setup and results obtained

We have tested FAPEC on Kongsberg EM302, EM710 and EM2040 water column files. EM2040 files, kindly provided by Kongsberg, were collected in a multibeam survey inside the harbour of Barcelona, presenting several structural elements and shoaling fish. EM302 data, kindly provided by Fugro, were collected in the Gulf of Mexico at depths around 1000m and present gas seeps. For comparison, lossless compression with gzip, bzip2, 7-zip and Zstandard has also been tested. Lossy FAPEC compression has been tested with option 1 (128 levels or shades of grey) to 5 (just 8 shades of grey, including black and white).

As can be seen in the left panel of Fig. 1, FAPEC provides the best lossless compression ratios on all this variety of sonar models and scenes. Furthermore, lossy compression achieves excellent ratios, as seen in the right panel. Besides this, FAPEC compresses at a speed only comparable

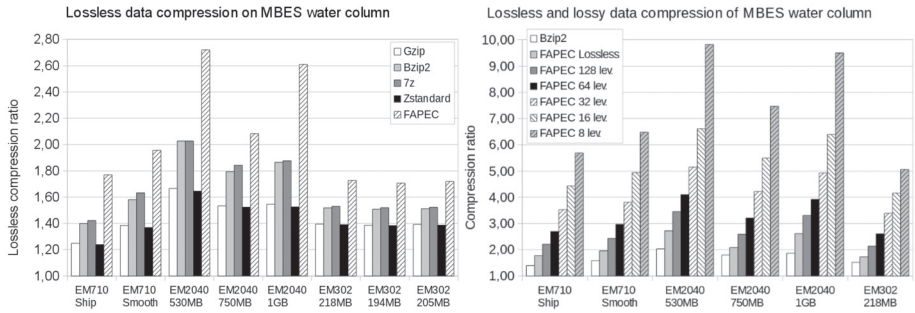


Fig. 1 - Lossless and lossy compression ratios of FAPEC and other solutions on Kongsberg MBES water column data.

to that of Zstandard (which gives much worse ratios), exceeding 60 MB/s even in single-thread mode on a typical desktop computer (Intel® CoreTM i5 3.1 GHz). For comparison, gzip just reaches 20 MB/s, whereas bzip2 and specially 7-zip are even slower. Compressed file sizes with 32 shades of grey or less are very similar (even smaller) than raw bathymetry files (Fig. 2). Together with the high compression speed, it means that continuous water column acquisition is finally feasible with FAPEC.

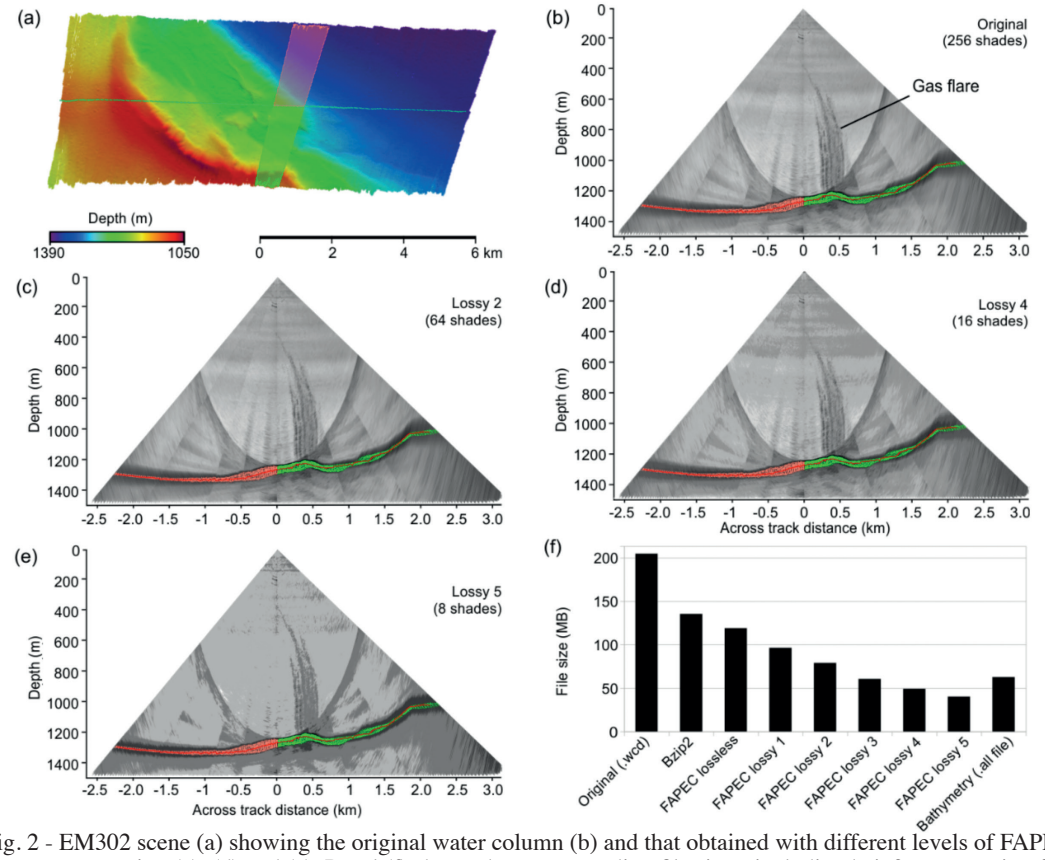


Fig. 2 - EM302 scene (a) showing the original water column (b) and that obtained with different levels of FAPEC lossy compression (c), (d) and (e). Panel (f) shows the corresponding file sizes, including bzip2 compression for comparison. Features in the water column such as gas seeps can still be detected even with low-quality compression levels. Actually, such levels seem to provide an even clearer evidence of the gas-related features.

Conclusions and forthcoming work

We present a new data compression solution for Kongsberg MBES water column files, which could also be extended to other manufacturers. It provides significantly better lossless ratios and speeds than any other existing solution, which makes it very interesting for massive data archiving or transfer. Lossy compression provides excellent ratios still with very good quality, meaning that it could be used in real-time to allow continuous water column acquisition. The block-based operation of FAPEC provides resilience in front of data corruption, minimizing data loss in such case. It also allows for quick detection of features in the water column from variations in the ratios.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 658358 (D. Amblas), MINECO (Spain) through ESP2016-80079-C2-1-R (MINECO/FEDER, UE) and ESP2014-55996-C2-1-R (MINECO/FEDER, UE), and MDM-2014-0369 of ICCUB (Unidad de Excelencia 'María de Maeztu').

References

[1] K. Colbo, T. Ross, C. Brown, T. Weber, 2014. *A review of oceanographic applications of water column data from multibeam echosounders,* Estuarine, Coastal and Shelf Science 145, 41-56.

[2] J. Portell, R. Iudica, E. García-Berro, A.G. Villafranca, G. Artigues, 2008. *FAPEC, a versatile and efficient data compressor for space missions,* International Journal of Remote Sensing 39 (7), 2022-2042.

Development of ecosystem-based maritime spatial planning decision support system for the marine protected areas designation process

Maija Viska, Latvian Institute of Aquatic Ecology (Latvia), maija.viska@lhei.lv
Solvita Strake, Latvian Institute of Aquatic Ecology (Latvia), solvita.strake@lhei.lv
Ingrida Purina, Latvian Institute of Aquatic Ecology (Latvia), ingrida.purina@lhei.lv

The Baltic Sea is one of the most human-influenced sea areas in the world and is in need for ecosystem-based maritime spatial planning (MSP). There are some separate examples in the Baltic Sea, but common platform for maritime spatial planning would facilitate the work with local stakeholders and cross boarder collaboration. In the sea everything is connected, it is a scene of a wide range of activities from natural processes to human activities such as wind farms, shipping, fishery, aquaculture and recreation. BONUS BASMATI project will develop the platform for an innovative decision support system that can facilitate between different actors and stakeholders in MSP and different stages of the planning process. The goal for this project is to ensure broad and easy access to data and information through an innovative web-based multi-channel decision support system and advanced tools for impact assessment and stakeholder involvement in the decision making process. This decision support system will be developed trough several case studies. The case studies were selected to represent marine spatial planning issues and will identify, collect, produce and supply data and maps concerning marine ecosystem services and human maritime activities.

Fig. 1 - Elements of BONUS BASMATI decision support system for ecosystem-based maritime spatial planning.

This research focuses on the case study for creation of tools facilitating the identification of new marine protected areas (MPAs) as well as the justification of the allocation of existing ones. MPAs will be based on most valuable benthic habitats. During the identification process of most valuable habitats are defined quality criteria for sea-bottom habitats which determine the territory with greatest ecological and socioeconomic value. Research area covers the south-eastern Baltic Sea area - Latvian territorial and exclusive economic zone waters.

MPAs are an important management instrument to achieve marine biodiversity conservation targets however establishment of these areas can be controversial as they can impose limitations to human activities in the sea and can have associated negative impacts on economic sectors. MPAs planing process should ensure the balance between the ecological conservation and socioeconomic requirements. Biological principles can be used as primary design criteria, but there should be included relevant socioeconomic aspects to ensure community support and meeting socioeconomic needs.

Decision support system will be linked with existing data systems of ICES, HELCOM data systems and EMODnet. Through the development of this system will be created or acquired data on ecosystem services and human uses to find the most suitable use for marine space. The goal of current study is to support the MPAs designation process considering not only conservation goals, but also including social and economic issues. This tool will evaluate suitable and efficient locations for MPAs using spatial multi-criteria analysis of alternative sea use options and utilizing the Analytical Hierarchy Process. Developed platform will present information as interactive maps, graphics and tables that will help involve interaction with stakeholders.

Acknowledgement. This research was funded by the BONUS BASMATI project that received funding from BONUS (Art 185), funded jointly by the EU and Innovation Fund Denmark, Swedish Research Council Formas, Academy of Finland, Latvian Ministry of Education and Science, and Forschungszentrum Jülich GmbH (Germany).

REACT-ESRI customizable application template

Luis Miguel Agudo Bravo, Spanish Institute of Oceanography (Spain), luismi.agudo@ieo.es
Jon Garrido, Bilbomática (Spain), jon.garrido@gmail.com
Olvido Tello-Antón, Spanish Institute of Oceanography (Spain), olvido.tello@ieo.es

React.js is a JavaScript library for building user interfaces. It is the view layer for web applications.

At the heart of all React applications are components. A component is a self-contained module that renders some output. We can write interface elements like a button, input field, map or widgets as a React component.

Unlike many of its predecessors, React operates not directly on the browser's Document Object Model (DOM) immediately, but on a virtual DOM. That is, rather than manipulating the document in a browser after changes to our data (which can be quite slow) it resolves changes on a DOM built and run entirely in memory. After the virtual DOM has been updated, React intelligently determines what changes to make to the actual browser's DOM.

The React Virtual DOM exists entirely in-memory and is a representation of the web browser's DOM. Because of this, when a React component is written, not is written directly to the DOM, but is generated a virtual component that React will turn into the DOM.

Using this library in the client side and the API for Javascript 4.6 from ESRI as map library, the Spanish Institute of Oceanography (IEO) has developed a new customizable web application template. Each React component or "widget" has a JSON file in order to configure style, behaviors and default init parameters.

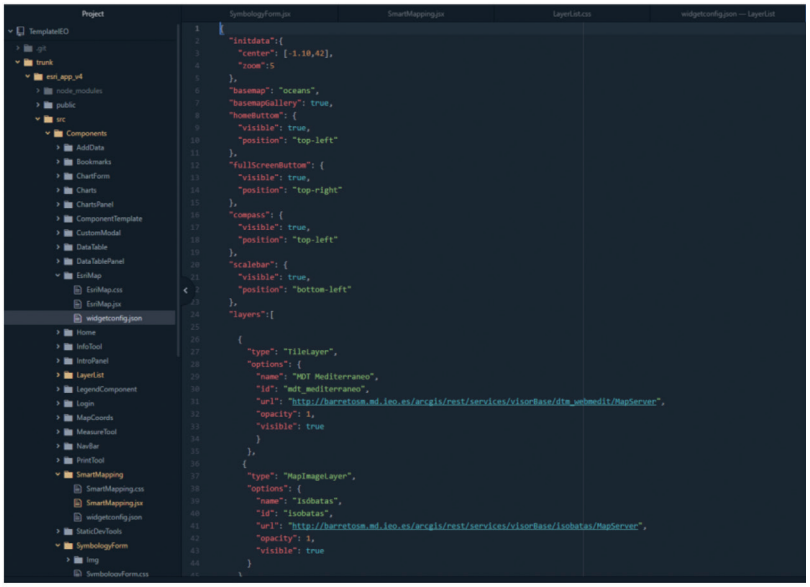


Fig. 1 - Example widget's JSON configuration file.

Also there are logic in the server side. In this case .NET is used to handle users and roles and ArcGIS Server is used to generate map and geoprocessing services.

The Viewer Base integrated on the Spatial Data Infrastructure of IEO has useful and advanced tools to discovering and managing the oceanographic data of IEO. This viewer has been updated with this new template

These tools allow display, select simbology and consult spatial and alphanumeric data. Besides allow insert new data, 3D displays, make several charts (bar, pie, scatter, etc.).

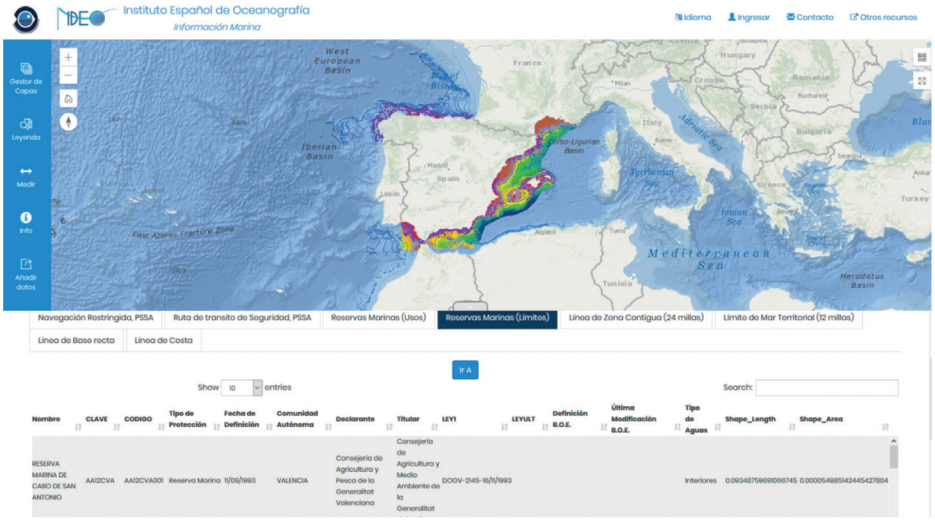


Fig. 2 - IEO Base Viewer.

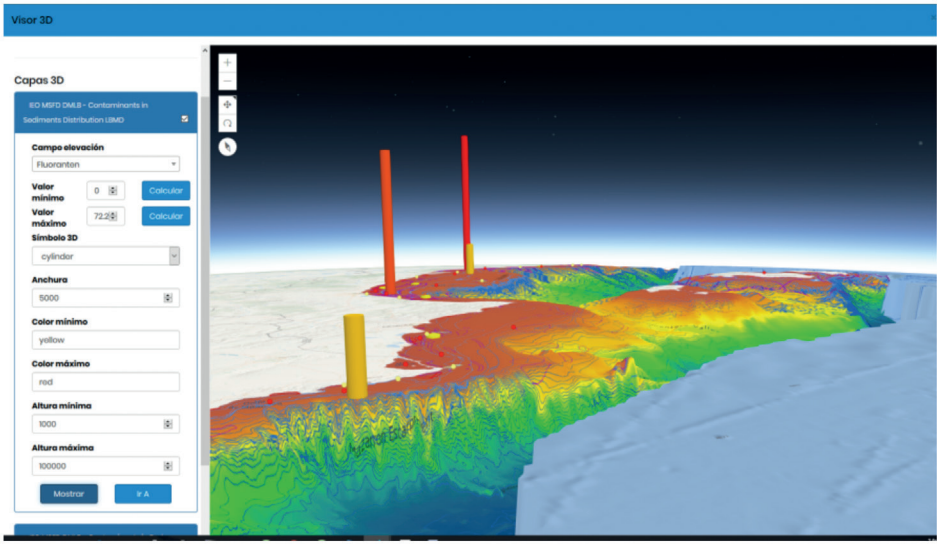


Fig. 3 - IEO Base Viewer- 3D Tool.

This template is available to download in order to deploy in a new production environment. Also it is easy develop new widgets (React components). In this way is possible to scale this application generating new analysis tools or modify existing tools.

Rosetta Stone service: a success story of standards, controlled vocabularies and communication

Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk
Gwenaelle Moncoiffe, British Oceanographic Data Centre (United Kingdom), gmon@bodc.ac.uk
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
Enrico Boldrini, National Research Council of Italy, Institute of Atmospheric Pollution Research (Italy), enrico.boldrini@cnr.it
Fabrizio Papeschi, National Research Council of Italy, Institute of Atmospheric Pollution Research (Italy), fabrizio.papeschi@cnr.it
Stefano Nativi, National Research Council of Italy, Institute of Atmospheric Pollution Research (Italy), stefano.nativi@cnr.it

The marine community across Europe, US and Australia, appears to be well connected in terms of communication and standards applied. The use of controlled vocabularies for data mark-up, that are based on W3C's Simple Knowledge Organisation System (SKOS) and are exposed as Linked data is a great success achieved by real human communication, enabled by collaborative projects like the Ocean Data Interoperability Platform (ODIP). The scene seems very promising to move to the next level of global integration. But is it enough?

European marine datasets served by SeaDataNet are marked up with standardised terms originating from controlled vocabularies hosted by the NERC Vocabulary Server (NVS) in terms of observable properties, instruments, platforms, organisations they belong to, disciplines they are related with and many more. The Australian and US originated datasets, apply the same principles, in terms of standards, but each continent uses their own national Vocabulary services, the US National Centers for Environmental Information Vocabularies (NCEI) and IMOS Australian Ocean Data Network (AODN) Vocabularies respectively. Part of ODIP's objectives is to interconnect marine communities worldwide, enabling global users to access data from regional data providers in EU, US and Australia, but is hindered by the different vocabularies, dialects and terminologies that co-exist.

In this presentation, we introduce the Rosetta Stone service as a translation service among the above mentioned controlled vocabularies, implemented via mappings stored in the NVS SPARQL endpoint. We then demonstrate how global data integration becomes feasible when collaboration, communication and standards are present.

Specifically Rosetta Stone was successfully experimented to semantically enhance the ODIP broker discovery capabilities. For example, ODIP users can now search the ODIP prototype 1+ portal using terms from a community vocabulary of choice: the ODIP broker engine leverages Rosetta Stone translation service capabilities to obtain translated (as well as related) query terms ready to be submitted and obtain results from all the heterogeneous ODIP data sources.

ABACUS – a cloud-based tool for standardising marine biological data recording

Ross Griffin, Ocean Ecology Limited (United Kingdom), ross.griffin@ocean-ecology.com
Paul Williams, Peninsula Data Solutions (United Kingdom), paul.williams@peninsuladatasolutions.co.uk
Elliot Carter, Peninsula Data Solutions (United Kingdom), elliot.carter@peninsuladatasolutions.co.uk

Many thousands of marine biological samples are collected and analysed on an annual basis to satisfy statutory monitoring commitments (e.g. Water Framework Directive (WFD), Habitats Directive) and conditions of marine licences granted for marine activities. These include seabed samples that undergo macrobenthic, particle size distribution and chemical analysis, water samples analysed to monitor planktonic communities and scientific trawl samples to assess fish and other mobile species. Despite the requirement for these analytical processes to be conducted by laboratories participating in recognised quality control schemes (e.g. the NE Atlantic Marine Biological Quality Control (NMBAQC) scheme), there remains fundamental issues surrounding the recording of non-standardised marine biological data. These issues stem from inter-analyst and inter-laboratory variability in sample analysis methodologies, recording practices, species naming, the use of taxonomic qualifiers and so on. This is thought to be the root cause for wide spread mis-interpretation of trends in marine biological communities which, in some cases, can have serious consequences for both Statutory Nature Conservation Bodies (SNCBs) and private sector organisations legally obligated to conduct robust marine ecological monitoring.

To address this ever-apparent issue, 'ABACUS' (v1.0) has been developed as a web-based platform for marine scientists to record, quality assure, store and export standardised marine biological data in line with internationally recognised data standards (e.g. MEDIN, GEMINI, ISO). The system has been developed with a 'live link' to key web resources including the World Register of Marine Species (WoRMS) ensuring use of the most up to date nomenclature as well as 'single click' exports to facilitate rapid data ingestion with key data archiving centres (e.g. DASSH).

Development Process

ABACUS has been developed using the latest Microsoft technologies (ASP.NET CORE, MVC, C#, Microsoft SSQL Server Database) and is encrypted using industry standard SSL and HTTPS. As part of the development of v1.0, a number of demonstration versions were tested by a team of taxonomists during analysis of hundreds of marine biological samples at Ocean Ecology Limited's (OEL) laboratory (UK).

Species Recording

Samples can be tracked through key analysis stages including log in, elutriation, extraction, identification, biometric measurements and biomass. Quality Control stages are available

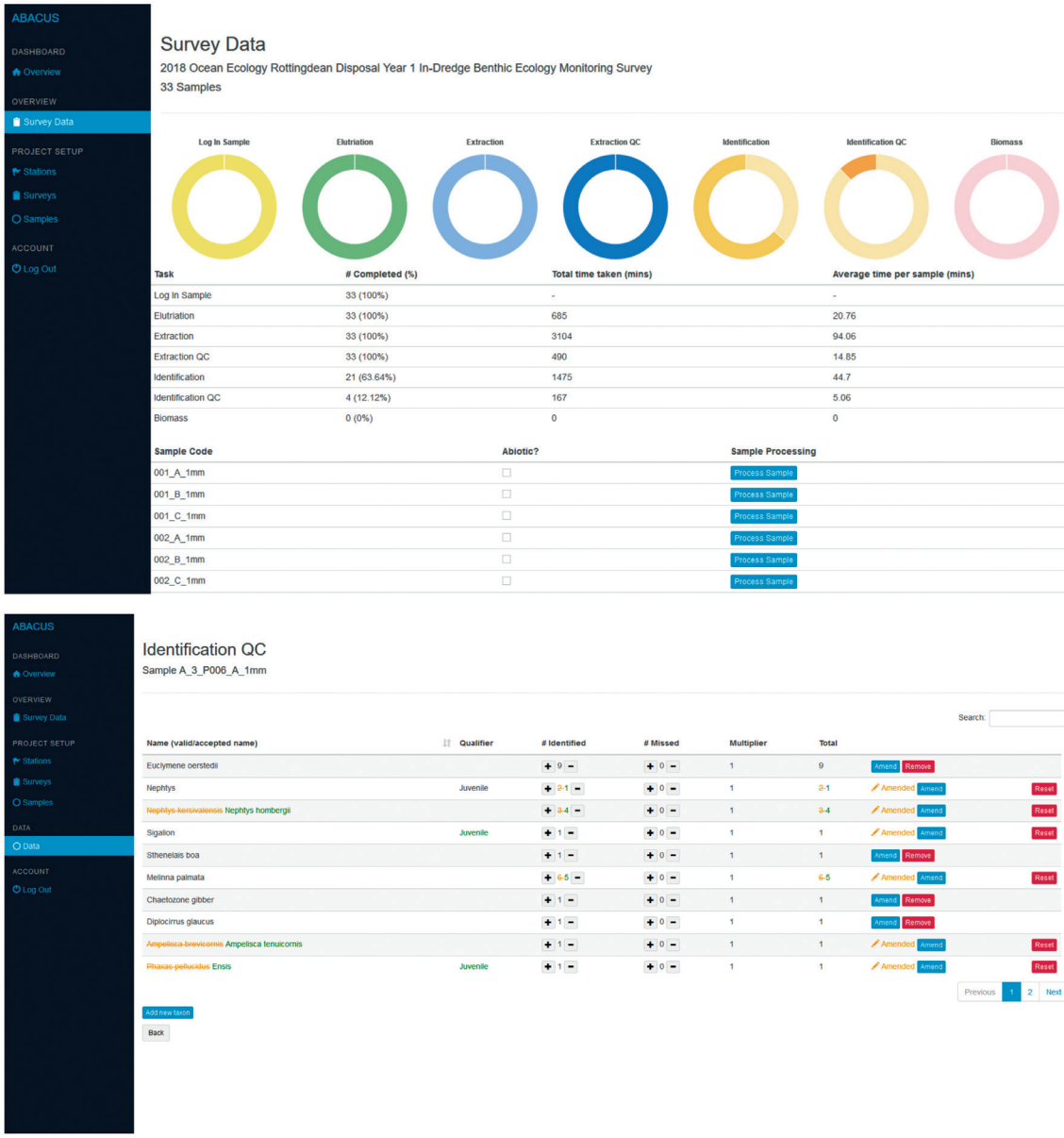


Fig. 1 - Top: Sample analysis overview page for an example project involving the analysis of 33 macrobenthic samples. Bottom: Identification QC page showing taxa entered by the primary analyst and corrections made by QC analyst.

for extraction and identification to improve quality or for the supervision and training of less experienced analysts. During the identification analysis stage, a web service provides a direct link to the WoRMS database. Typing a few characters of valid or scientific name will automatically return a list of matching taxa. Selecting one of the taxa will retrieve a collection of data from the WoRMS database including classification data, authority, AphiaIDs and other attribute data (e.g. AMBI groups) as well as other species information including Species Directory Codes (SDCs).

Data Exports

Marine Environmental Data and Information Network (MEDIN) compliant data export functionality is provided as standard, which has been funded and validated by MEDIN. Fully MEDIN compliant exports can be created and downloaded with just a few clicks, skipping error-prone and slow manual processes. Users are also able to customise data exports based on data type (e.g. count or biomass data) and subsets of samples and/or sampling events.

Users

Being cloud based, analysts can sign into ABACUS via a web browser from any device with an internet connection. Through an admin dashboard selected users can manage access permissions of others as well as creating user profiles for external partners. When signed in, user activity can be recorded providing a full audit trail of all data recording, quality control actions (e.g. amending a species name) and data exports.

Data harvesting - machine to machine

Lotta Fyrberg, Swedish Meteorological and Hydrological Institute (Sweden), Lotta.Fyrberg@smhi.se

Nils Nexelius, Swedish Meteorological and Hydrological Institute (Sweden), Nils.Nexelius@smhi.se

Arnold Andreasson, Swedish Meteorological and Hydrological Institute (Sweden), Arnold.Andreasson@smhi.se

Lisa Sundqvist, Swedish Meteorological and Hydrological Institute (Sweden), Lisa.Sundqvist@smhi.se

Users of data will always need access to highest quality, largest collection, and most recent version of data. Advances in technology can meet these demands via technical solutions. One way forward is to grant open access using machine to machine interfaces. SMHI has developed the SHARKdata.se system to handle these data flows. The data are currently being harvested in the DwC-A format, by portals such as EMODnet-Biology (and hence to EurOBIS/OBIS), and by the Swedish LifeWatch system.

There is now an ongoing study in co-operation with ICES Data centre. The aim of this study is to automate the national reporting of biological data (phytoplankton, zooplankton, phytobenthos and zoobenthos) from OSPAR and HELCOM regions, to the Marine Environmental data portal (DOME).

This will result in the following:

I. Reduce the amount of manual work (and human errors);
II. Ensure the most recent version of data in DOME;
III. Automated quality controls through DATSU REST API on data;
IV. Higher quality of data.

All technology used is open source (MIT license) and hence open for everyone to download and build their own system. For data users it is possible to use R and Python to set up any type of analysis on data from the Swedish National Oceanographic Data Centre. Examples are published on the server SHARKdata.se

Oceanids command and control (C2) data system - marine autonomous systems data for vehicle piloting, scientific data users, operational data assimilation

Justin Buck, National Oceanography Centre (United Kingdom), juck@bodc.ac.uk
Alexander Phillips, National Oceanography Centre (United Kingdom), abp@noc.ac.uk
Alvaro Lorenzo, National Oceanography Centre (United Kingdom), allore@noc.ac.uk
Alexandra Kokkinaki, National Oceanography Centre (United Kingdom), alexk@bodc.ac.uk
Malcolm Hearn, National Oceanography Centre (United Kingdom), mahe@bodc.ac.uk
Thomas Gardner, National Oceanography Centre (United Kingdom), thogar@bodc.ac.uk
Kay Thorne, National Oceanography Centre (United Kingdom), klt@bodc.ac.uk
Scientific Team: Oceanids command and control system development team, National Oceanography Centre (United Kingdom) and Scottish Association for Marine Science (United Kingdom)

Abstract

This paper describes the development of the data system that is a component of the Oceanids command and control (C2) system. Oceanids C2 will be an integrated platform that will combine piloting, data processing, data analytics, and data delivery for the fleet of autonomous platforms hosted by the UK National Oceanographic Centre.

The data system is built using standard ontologies for structure and the NERC vocabulary server for terms to maximise metadata interchangeability. Data delivery will include dissemination to the Ocean Glider Network (for inclusion in Copernicus and EMODNet) in addition to the projects deploying autonomous vehicles.

Introduction

The National Oceanography Centre (NOC) operate a fleet of approximately 36 autonomous marine platforms including submarine gliders, autonomous underwater vehicles, and autonomous surface vehicles. Each platform effectivity has the capability to observe the ocean and collect data akin to a small research vessel. This is creating a growth in data volumes and complexity while the amount of resource available to manage data remains static. The OceanIds Command and Control (C2) project aims to solve these issues by fully automating the data archival, processing and dissemination. This is as part of a combined planning, piloting data processing, visulisation and delivery platform as described in Fig. 1.

C2 data system

The data architecture being implemented jointly by NOC and the Scottish Association for Marine Science (SAMS) includes a single Application Programming Interface (API) gateway to handle authentication, forwarding and delivery of both metadata and data. Technicians and principle

investigators will enter expedition data prior to deployment of vehicles enabling automated data processing when vehicles are deployed. The system will support automated metadata acquisition from platforms as this technology moves towards operational implementation.

The metadata exposure to the web builds on a prototype developed by the European Commission supported SenseOCEAN project and is via open standards including World Wide Web Consortium (W3C) RDF/XML and the use of the Semantic Sensor Network ontology and Open Geospatial Consortium (OGC) SensorML standard. Metadata exposure via SSN and SensorML is achieved using a database build using prexesting ontologies and terms from the the NERC vocabulary server 2.0 (NVS2) as showing in Fig. 2. Data will be delivered in the marine domain Everyone's Glider Observatory (EGO) format and OGC Observations and Measurements. Additional formats will be served by implementation of endpoints such as the NOAA ERDDAP tool.

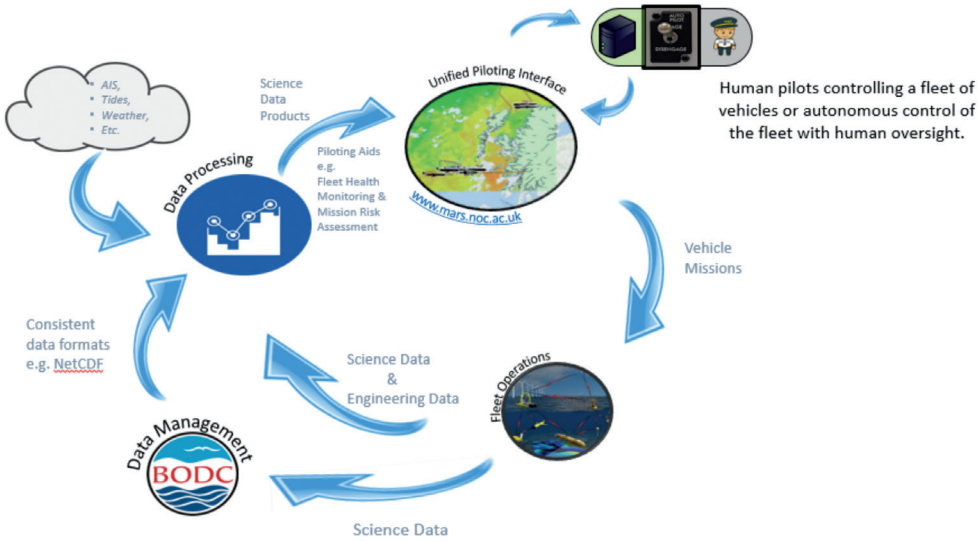


Fig. 1 - Flow chart showing the integrated Oceanids command and control system.

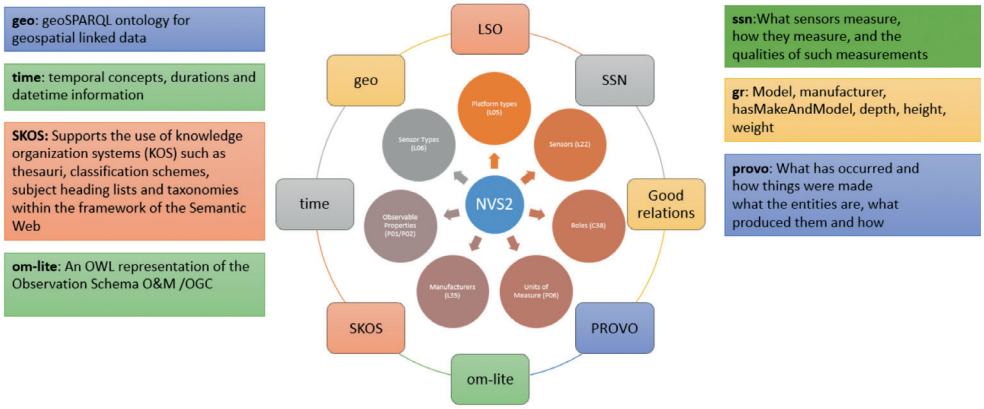


Fig. 2 - The fusion of ontologies and vocabularies from the NERC vocabulary server that enable the delivery of standardised metadata in both W3C SSN and OGC SensorML standards.

Data delivery

This standardised data delivery via the API gateway enables timely near-real-time data to be served to Oceanids users, BODC users, operational users and big data systems. The use of open standards will also enable web interfaces to be rapidly built on the API gateway and delivery to European research infrastructures that include aligned reference models for data infrastructure. It is intended for the C2 data system to contribute data and metadata directly to the Ocean Glider network which enabling the data to be included in Copernicus and EMODNet data products.

The SeaDataCloud Virtual Research Environment: the technical perspective

Merret Buurman, German Climate Computing Centre (DKRZ) (Germany), buurman@dkrz.de
Sri Harsha Vathsavayi, CSC IT Center for Science (Finland), sriharsha.vathsavayi@csc.fi
Peter Thijsse, Mariene Informatie Service 'MARIS' BV (The Netherlands), peter@maris.nl
Alexander Barth, GHER, Université de Liège (Belgium), a.barth@uliege.be
Sebastian Mieruch, Alfred Wegener Institut (Germany), sebastian.mieruch@awi.de
Nikolaos Evangelou, Greek Research and Technology Network S.A. (Greece), nikosev@admin.grnet.gr
Marco Rorro, Consorzio Interuniversitario CINECA (Italy), m.rorro@cineca.it
Adrian Coveney, Science and Technology Facilities Council, (United Kingdom), adrian.coveney@stfc.ac.uk

Virtual Research Environments (VREs) allow researchers to execute complex workflows of data-driven experiments in the cloud, thus reducing data transfers and leveraging remote computational facility. The VRE developed by SeaDataNet and EUDAT will allow users to use various processing services on data from a variety of sources. To accommodate this diversity, clear interactions between services are defined. Docker containers make it possible to run very diverse services together in one environment, and to update and migrate services without interfering with other services. Also, this modularity allows to easily scale and extend the services.

With the growth of data size and variety, and processing complexity, Virtual Research Environments (VREs) are getting more and more popular. VREs, also called Science Gateways or Virtual Labs, allow researchers to compose and execute complex workflows of data-driven experiments including finding data, accessing data, processing iteratively data with various tools, visualizing results, sharing relevant findings with colleagues, replicating experiments/workflows and publishing insights online, thus reducing data transfers and leveraging remote computational facility.

SeaDataNet is the primary entry point for many marine scientists to find and access data. After the recent move of SeaDataNet's data to the cloud, offering processing capacities in the cloud too is a logical next step. That is why SeaDataNet is now developing a Virtual Research Environment for their users. It is being developed in close collaboration between the SeaDataNet consortium and the EUDAT research infrastructure, represented by the scientific data and computing centres CSC (Finland), DKRZ (Germany), CINECA (Italy), GRNET (Greece) and STFC (United Kingdom).

This abstract aims to show the more technology-curious public a view behind the curtains of how the VRE is realized and what advantages this setup brings.

The SeaDataNet VRE will allow its users to perform resource-intensive analyses on the cloud. Applications that are commonly used in the marine sciences, such as AWI's OceanDataView (Schlitzer, 2018), Jupyter Notebooks and University of Liège's DIVA (Data-Interpolating Variational Analysis (Barth *et al*., 2014) to create gridded climatologies from in situ observations), will be available as VRE services. Additional services such as notifying users of outdated datasets,

enabling users to communicate while working on a common task, and chaining services will be added for the users' convenience.

Many of the applications that will be included in the VRE already exist, they are written in different languages, have different user interfaces, different architectures and different dependencies. Thus, the VRE has to be implemented in a way that allows applications that are quite different to be integrated and interoperable, and to add more applications later on. To fulfill these requirements, SeaDataNet has come up with a flexible architecture that is described in the next sections.

To accommodate the very diverse applications that already exist, to easily integrate new ones, and to avoid dependency conflicts, all the applications will run inside Docker containers. This approach will also improve environment portability, scalability, security and scientific reproducibility. Application containers will run on the EUDAT's computing service B2HOST, which provides scalable processing capability, fast access to storage volumes, and scheduling of containers execution. Every application exposes its own service interface, which could be a GUI, a command line interface in a Jupyter Notebook and/or a REST API.

The application containers (i.e., services) interface with the data layer in the backend, with the user interfaces in the frontend, and with the controller.

The frontend will be developed as a responsive JavaScript layer, based on state-of-the-art JavaScript libraries. It will be responsible for authentication and authorization of the users, and for routing them to the interfaces of the services. Services that are common to all processing services can also be accessed from here, such as dataset management component, process chaining (workflows), user communication, and version notification.

For the time being, the main data backend of the VRE will be the EUDAT service B2DROP. B2DROP allows users to upload, store and securely share datasets. It will be accessible to all the

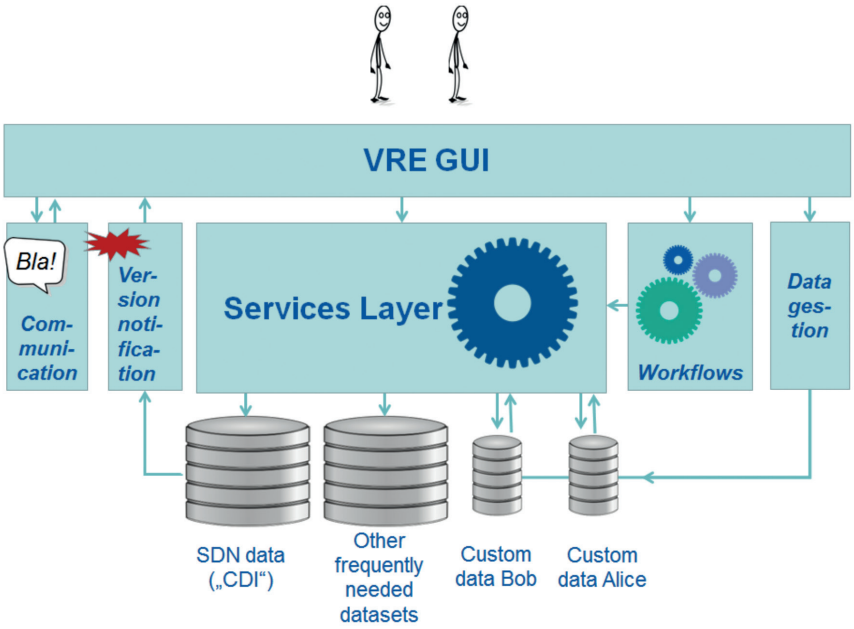


Fig. 1 - Architecture of the Virtual Research Environment.

applications by the WebDAV protocol, allowing the applications to interact with the B2DROP content as a file system.

Behind the frontend, a controller makes up the heart of the VRE and glues all the components together. It sees to the mounting of the user's private data from EUDAT's nextcloud-based B2DROP service, mounting other data and volumes, launches the correct service containers on-demand and it is responsible for all inter-container communications.

Summarising, the SeaDataNet VRE will allow users to use a variety of processing services on data from a variety of sources. To accommodate this diversity, clear interactions between services are defined. Docker containers make it possible to run very diverse processing services together in one environment, and to update and migrate services without interfering with other services. Also, this modularity will allow to scale the services when necessary and easily add new services or extend the existing services.

References

Barth, A., Beckers, J.-M., Troupin, C., Alvera-Azcárate, A., and Vandenbulcke, L., 2014. divand-1.0: n-dimensional variational data analysis for ocean observations, Geosci. Model Dev., 7, 225-241, doi:10.5194/gmd-7-225-2014.

Schlitzer, R., 2018. Ocean Data View, https://odv.awi.de.

https://www.eudat.eu/services/userdoc/b2host.

DECISION SUPPORT SYSTEM for the monitoring and management of the Romanian littoral's bathing areas

Razvan Mateescu, National Institute for Marine Research and Development "Grigore Antipa" (Romania), razvan_doru@yahoo.com

Elena Vlasceanu, National Institute for Marine Research and Development "Grigore Antipa" (Romania), evlasceanu74@yahoo.com

Dragos Niculescu, National Institute for Marine Research and Development "Grigore Antipa" (Romania), n.dragos8@gmail.com

The Competence Centre COSMOMAR is part of NIMRD/National Institute for Marine Research and Development "Grigore Antipa", from Constanta, Romania, being mainly involved in EO research and applied technology, fundamental for the understanding, protection and management of coastal and marine environment of western Black Sea coast/basin.

The Earth Observation (EO) activities performing in COSMOMAR were extended with the intention to act in the direction of meeting coastal users' necessities, thus enhancing certain relevant tools that improve the access to data and information and also allow stakeholders to perform different analyses and run scenarios for specific actions, such as precise mapping and evaluation of water mass circulation, upsetting the coastal ecosystems resources, in order to increase the efficiency of their conservation measures.

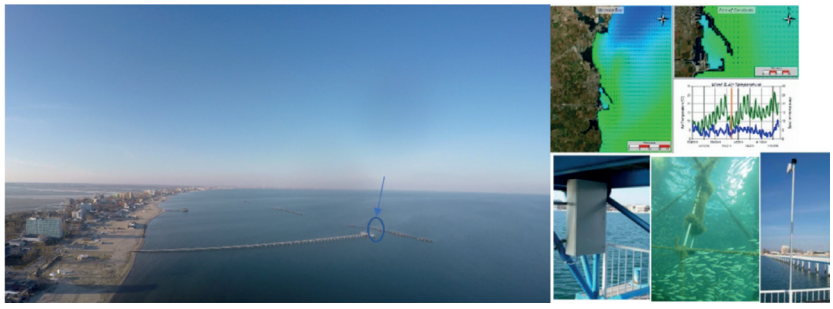


Fig. 1 - Oceanographic station of Mamaya Bay/ central unit of Romanian littoral.

The same purposes were followed in the implementation of a dynamic web-based mobile-friendly decision support system to enhance the management, monitoring and forecasting of the water quality (iSWIM, http://iswim.rmri.ro) for the Romanian marine and coastal zone, integrating numerical models (downscaled from Copernicus Marine Environment Monitoring Service - CMEMS - local solution) with in-situ measured data and CMEMS remote sensing products. The operational system also integrates in near real time specific oceanographic data from a coastal station as a base of documentation of water quality of marine bathing areas on central transitional unit of Romanian littoral. In this specific application, the coastal station data provide an important input of hydrometric, physical, chemical and biological parameters for the description of the coastal/marine waters' ecological status. The system is developed for public health and touristic/ navigation purposes and provides data and information through a monitoring-modeling service for bathing areas of the Romanian touristic littoral.

MyWay : a real time vessel tracking and cruise events system

Stefano Ferriani, Agenzia Nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (Italy), stefano.feriani@enea.it

Leda Pecci, Agenzia Nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (Italy), leda.pecci@enea.it

ENEA is a member of the National Programme of Antarctic Research (PNRA) funded by the Italian Ministry of Education, University and Research (MIUR). The government programme has instituted 32 national oceanographic campaigns in Antarctica and has developed many scientific and technological activities in which research centres and universities have participated.

The scientific personnel on board belong to different research institutes and organisations. The teams are involved in projects related to various disciplines such as meteorology, geology, physical, chemical oceanography, geological and marine biodiversity.

The MyWay system, used in the Antarctic cruises since 2014, is a set of hardware and software that, in an integrated way, ensures all the functions necessary for carrying out the activities of the Navigation Laboratory:

- acquisition and real-time display of vessel position data and weather and sea conditions;
- storing the data acquired in a geographic database;
- visualization of the cartography of the operative area with the information useful for navigation:
- trace of the ship;
- ice maps;
- satellite images;
- position of the scientific activities carried out and planned;
- automatic and periodic sending of the vessel's position by e-mail to addresses definable by the system itself;
- automatic and periodic sending of the vessel's position to the Ship Position Reporting System (SPRS) of the COMNAP.

The block diagram of MyWay is shown in Figure 1, which highlights the subsystems that compose it: measuring, processing and visualization.

The system architecture was designed with the intention of separating and rendering its parts independent, delegating to the local network the task of making functional connections. Following this approach, the data generated by the measuring instruments are injected into the LAN, making them available to all the devices that access the network, regardless of the presence or operations of the other components of MyWay. The heart of the system (processing) consists of a java web application that acquires data from the LAN, displays them in real time and stores them in a geographic database. A series of all-in-one personal computers (visualization) allows you to access and view information concerning navigation. The "core" of the processing section consists of a virtual machine with Linux Ubuntu 16.04.4 LTS operating system hosted on

VMWare vSphere platform. The MyWay three-tier web application runs on it. The application, built with Java technology, has an architecture based on the MVC model, implemented thanks to the Struts framework. Made entirely with open-source software, from a functional point of view, MyWay can be considered a web-GIS associated with a nautical satellite navigator. Being a web application, both its control and access to data and information, it can be performed by any network client equipped with a browser with JavaScript support.

The measuring section of the system has the task of acquiring the navigation data and making them available, in UDP protocol, on the local Ethernet network of the ship. For this reason, with the exception of the Motion Reference Unit (MRU) that interfaces directly with the LAN, all the other instruments are connected to two Miniplex that acquire the signals and send them with UDP protocol on Ethernet. Observing Fig. 1 we can see that the most critical equipment (Miniplex and GPS-Gyro) are redundant for safety reasons.

To view the information related to navigation, six personal all-in-one computers have been set up, conveniently located in the ship's premises:

- bridge command;
- secretariat corridor;
- canteen room;
- user room;
- aft cabins corridor;
- laboratory 16 Lidar.

Thanks to the installation on each of them of the VNC remote control software, it was possible to manage all the devices in a centralized manner, verifying their correct operation and setting the displays in relation to the current operating context. This solution has proved particularly useful for the management of the bridge terminal which, compared to the others, played a particularly critical role.

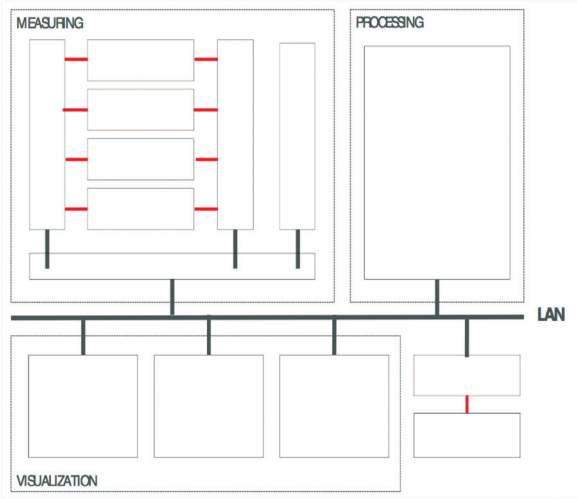


Fig. 1 - Block diagram of the MyWay navigation system.

MyWay gives an overview of projects in which the teams are involved, the Principal Investigators, the observations, the instruments used and the stations, it may be considered a useful support because it comprises a map of the ice, bathymetry, weather data which can be useful for rescheduling the planned routes.

EMODnet Atlantic checkpoint: data adequacy to EU challenges

Jacques Populus, Institut Français de Recherche pour l'Exploitation de la Mer (France), jacques.populus@ifremer.fr

Eric Moussat, Institut Français de Recherche pour l'Exploitation de la Mer (France), eric.moussat@ifremer.fr

Mickael Vasquez, Institut Français de Recherche pour l'Exploitation de la Mer (France), mickael.vasquez@ifremer.fr

Erwann Quimbert, Institut Français de Recherche pour l'Exploitation de la Mer (France), erwann.quimbert@ifremer.fr

Frédérique Blanc, Collecte Localisation Satellites (France), fblanc@cls.fr

Laurent Soudarin, Collecte Localisation Satellites (France), lsoudarin@cls.fr

Introduction

After a time when observations of the sea have been made for specific purposes, e.g. for specific national purposes or to demonstrate a technological capability, the European Commission has now moved to a new paradigm where data are collected once and used them for as many purposes as possible. This means relying preferably on users rather than on producers to assess existing data sets and data sources and promote recommendations for a better satisfaction of their needs. The EMODNET Atlantic checkpoint (http://www.emodnet-atlantic.eu/) was designed to evaluate the fitness-for-use of current observations and data assembly programs towards 11 marine applications and prioritizing the needs to optimize monitoring systems at the scale of the North Atlantic Ocean.

Methodology

The methodology adopted from the Medsea checkpoint documents the fitness for use of the existing data by providing indicators of adequacy to the challenge products. The assessment criteria and the development of checkpoint information and indicators are derived from the ISO standards for geographical information (ISO19131 Data product Specification, ISO19157 Data Quality and ISO 19115 Metadata). The fitness for use of input data is evaluated in function of the Data appropriateness, and of the availability conditions.

The infrastructure initially set up by the Medsea Checkpoint is based on standardized catalogues of products and input datasets including the specifications and assessment results of the challenge experts exploited to provide feedback and recommendations to their commissioners in the Data Adequacy Report.

To provide a consistent overview of what is needed or available and to reveal the potential synergies among users of the same variable, the results have been classified according to the SeaDataNet vocabulary initially designed for marine data (https://www.seadatanet.org/Standards/Common-Vocabularies) which offers three levels of granularity from finer to the coarser: the Parameter Usage Vocabulary (P01) list for Characteristics, the P02 list for the "Category of

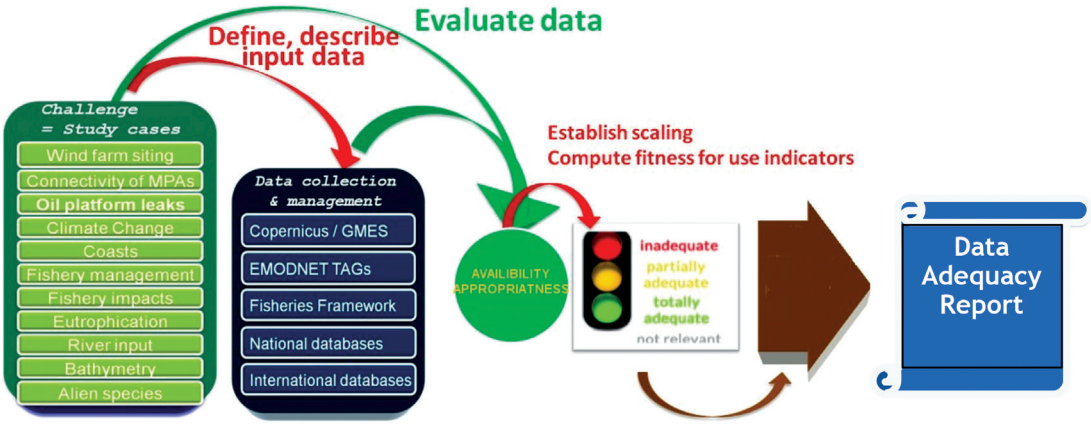


Fig. 1 - Checkpoint concept.

characteristics" and the P03 list for the "Groups of categories of characteristics".

The issues encountered were classified into the following types:

- Data gaps due to lack of appropriate spatial or temporal resolution or coverage or attributes
- Assembly needs which should be covered by the creation or update of Thematic Assembly Centres (TACs), which concerns data scattered but taken care of by a number of providers;
- Availability restrictions due to policy, lack of information on data quality and technical issues.

Decision makers and data providers will refer to the Data Adequacy Report to find the details of the specifications which have not been satisfied.

Data gaps

The main categories of characteristics (P02s) suffering from gaps are:

Air: wind speed and direction

Biology-Biota: fish abundance in water bodies is ill-known, especially for Eel and Salmon, two key species for river environments; phytoplankton generic abundance; invasive species monitoring parameters which suffer both of lack of completeness and of reliability

Fresh water: too few rivers are instrumented for a number of parameters allowing better modeling of the fate of marine water: water flow, temperature, salinity, oxygen, nitrates, phosphates, etc.

Marine water: horizontal velocity of the water column (currents) which generally needs at least 1 km resolution to properly address phenomena at the coast, whether for renewable energies or larvae dispersion computations; sea level change cannot be properly assessed at the coast due to the lack of density of tidal gauges; in situ chlorophyll, dissolved oxygen, salinity and temperature, along with chemicals such as nitrates and phosphates, are all suffering from a low density and lack of co-located measurements, which severely limits the ability to deal with eutrophication.

Seabed-riverbed: the composition of the seabed is ill-known. In terms of sediment and lithology, a scale of 1:1 000 000 would be necessary to address windfarm siting while the EU broad-scale seabed habitats map should be extended to the whole basin, while particular focus on the coastal zone.

Assembly needs

Besides true data gaps, a lot of data suffer from a lack of assembly, i.e. the data exist somewhere but are scattered, not catered for, or not homogeneous. This is mainly the cases in Biology for the threatened or declining species and for Human activities: Marine Traffic (AIS, VMS including vessels of length less than 12m), Marine Protected Areas monitoring parameters , fishing by-catch in numbers and discards in weight.

Availibility restrictions

High resolution VMS and ERS (logbook) data are missing from most countries to policy issues. Costs limit the use of AIS data. Open and more affordable marine traffic data would benefit to a large number of applications. Information on the completeness, the temporal validity, the accuracy of the datasets is a recurrent need. The extreme size of data files has been an obstacle to download and process data and metadata from Copernicus and EMODnet bathymetry highlighting the need of offering cloud computing services to end users.

Conclusions and perspectives

The lessons learned through this experience are discussed. Beyond data adequacy the Checkpoints can provide much needed information for optimizing future European ocean observing systems and data management.

SESSION 2

Technical developments for marine information and data management

- Standards for data interoperability
- Quality assurance
- Data citation, DOI, PID and data versioning
- Monitoring and tuning
- Best practice

ORAL PRESENTATIONS

The place of Schema.org in Linked Ocean Data

Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie
Rob Thomas, Marine Institute (Ireland), rob.thomas@marine.ie
Adam Shepherd, Woods Hole Oceanographic Institution (United States of America), ashepherd@whoi.edu
Doug Fils, Consortium for Ocean Leadership (United States of America), dfils@oceanleadership.org
Kevin O'Brien, National Oceanic and Atmospheric Administration (United States of America), kevin.m.o'brien@noaa.gov

Introduction

The use of Linked Open Data, that is publishing data on the World Wide Web using Web addresses to name things and using structured data standards (Berners-Lee, 2006), in marine data and information systems has been well documented over recent years (Leadbetter, *et al*. 2013; Leadbetter, 2015; Leadbetter *et al*. 2016). However, much of the work has been focused on the use of highly domain-specific Linked Data models (vocabularies, thesauri and ontologies), such as the NERC Vocabulary Server (Leadbetter, Lowry and Clements, 2014), Ocean Data Ontology[1] and GeoLink[2]. However, it has remained difficult to build a comprehensive, global search tool for marine data. In parallel, the major search engine companies have been driving a more generalised ontology known as Schema.org (Guha, Brickley & Macbeth, 2016). Schema.org patterns are well-known to both web developers and web users as they facilitate the display of enhanced search results through "Rich Snippets." Rich Snippets are related pieces of information displayed alongside a search result (for example when searching for a film or movie the running time; description; aggregated; reviews; related news items – see Fig. 1).

Schema.org for the Oceanographic Domain

Over the past year activity has been focussed on delivering dataset descriptions in Schema.org compliant forms. This has been done in a number of projects with slightly different approaches, but each activity has been done in liaison with Google, who are developing a dataset portal from the encoded metadata.

As part of Ireland's Digital Ocean Programme (Leadbetter, O'Grady and Burke, 2016), the Erddap data server software package published by NOAA (Simons, 2017) has been extended to output Schema.org Dataset metadata from its dataset catalogue landing pages. Additionally, the Drupal based Marine Institute data catalogue, which has also been developed under the Digital Ocean programme, has the same Schema.org dataset output driven from its internal database. Additionally, the Biological and Chemical Oceanography Data Management Office (BCO-DMO) is also embedding Schema.org dataset metadata from their own Drupal content management system.

[1] http://ocean-data.org/schema/.

[2] http://schema.geolink.org/.

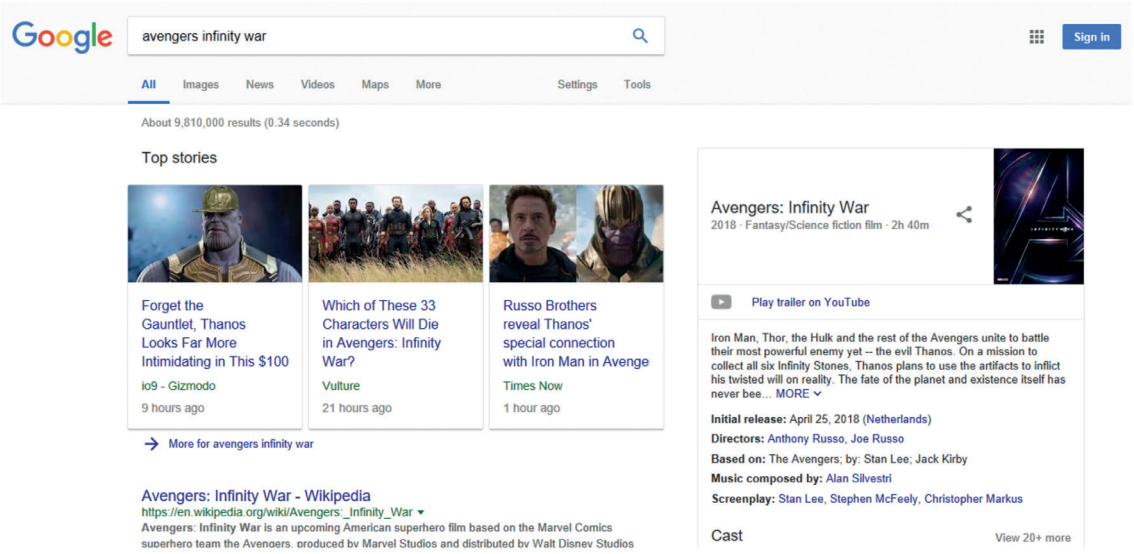


Fig. 1 - Rich Snippets in search powered by Schema.org. The enhanced search results are shown in the news items and the panel to the right hand side of the search page.

The EarthCube Architecture Project 418 is harvesting JSON-LD embedded in web pages across national data repositories in the U.S., looking for datasets that are described using Schema.org and other external vocabularies.

Under the Linked Data developments for the SeaDataNet catalogue structure, the various catalogues have been mapped through to Schema.org representations from the other Linked Data flavours in which they have been represented. Variously, this includes the Schema.org dataset pattern for the EDMED dataset and CDI data directory; Schema.org organisation for the EDMO organisation directory; and Schema.org Event for the Cruise Summary Report directory. Similarly, to address the issue of unique Linked Data identifiers for individual people, a Schema.org person profile of the IODE's OceanExpert repository has been developed under the Ocean Data Interoperability Platform and SeaDataCloud.

Conclusions

While the development of Linked Ocean Data has been valuable in gaining a richer understanding of our datasets and their structure, it would now seem that the time is right to invest in mapping these efforts to the more general Schema.org patterns. This is not without its issues, as there are patterns which have not yet been fully developed in Schema.org that are valuable in Linked Ocean Data. The Schema.org developers have called the vocabulary a core vocabulary which should be extended where richer semantics are required. As well, the referencing of standard vocabularies, such as those in the NERC Vocabulary Server, for naming parameters continues to be vital for interoperability but is not well defined in Schema.org. However, by agreeing to vocabulary referencing and by pursuing developments with the Schema.org governance bodies, these hurdles can be overcome and a truly global Linked Ocean Data profile can be achieved. In the face of continued difficulties in building successful marine data discovery systems, such a globally linked profile will enhance general discoverability through its inherent connections with well-known and widely used search tools.

References

Berners-Lee T. 2006. Design issues: Linked Data. URL http://www.w3.org/DesignIssues/LinkedData.html.

Guha R.V., Brickley D., Macbeth S. 2016. *Schema. org: evolution of structured data on the web.* Communications of the ACM, 59(2), 44-51.

Leadbetter A. 2015. *Linked Ocean Data.* Studies on the Semantic Web volume 20.: The Semantic Web in Earth and Space Science. Current Status and Future Directions, 11-31.

Leadbetter A., Arko R., Chandler C., Shepherd A., Lowry R. 2013. *Linked Data: An oceanographic perspective*. Journal of Ocean Technology, 8(3), 7-12.

Leadbetter A., Cheatham M., Shepherd A., Thomas R. 2016. *Linked Ocean Data 2.0.* Oceanographic and Marine Cross-Domain Data Management for Sustainable Development, 69-99.

Leadbetter A.M., Lowry R.K., Clements D.O. 2014. *Putting meaning into NETMAR–the open service network for marine environmental data*. International Journal of Digital Earth, 7(10), 811-828.

Leadbetter A., O'Grady E., Burke N. 2016. *Ireland's Integrated Digital Ocean*. Bollettino di Geofisica teorica ed applicata 57(supplement), 224-226.

Simons R. 2017. ERDDAP. https://coastwatch.pfeg.noaa.gov/erddap Monterey, CA: NOAA/NMFS/SWFSC/ERD.

Establishing "Best practice" data workflows in marine research at GEOMAR, Kiel

Hela Mehrtens, Helmholtz Centre for ocean Research Kiel (Germany), hmehrtens@geomar.de

Pina Springer, Helmholtz Centre for ocean Research Kiel (Germany), pspringer@geomar.de

Claas Faber, Helmholtz Centre for ocean Research Kiel (Germany), cfaber@geomar.de

Lisa Paglialonga, Helmholtz Centre for ocean Research Kiel (Germany), lpaglialonga@geomar.de

Carsten Schirnick, Helmholtz Centre for ocean Research Kiel (Germany), cschirnick@geomar.de

Background

The joint GEOMAR Data Management Team is a cooperation of GEOMAR Helmholtz Centre for Ocean Research Kiel and several large-scale research projects including collaboration with other marine research institutions. The coalition has established a permanent data management team and infrastructure in Kiel for marine research supporting the entire lifecycle of research data from description through storage and sharing for collaboration to publication. The infrastructure is continuously improved by extensions developed in close cooperation with scientists and data centres. In order to improve the findability and reusability of data, we are on the way to establish and describe reliable and reusable data workflows.

The role of data management in marine science

Marine research is often based on observations made at sea. Research cruises are therefore the starting point of the data workflow. Information on the acquired data (metadata) have to be preserved and collected from the vessels, deployed devices navigation and control systems. The information about research actions on a cruise is used to feed our institutional data information system (OSIS) and forwarded to a central world data centre for publication (PANGAEA) (Fig. 1). This allows to document the existence of data while data analysis is still ongoing. Given a data policy which sets timelines for periods between data sampling and data publication, scientists are supported in tracking their data flow. They are reminded to share their data in reusable formats within their working group and to start the data publication process in parallel or well before publication of their scientific results.

Who is needed to establish a data workflow

A close collaboration between scientists and data manager is necessary to establish a reliable workflow from cruise planning to published and reusable data. Each scientific community has to agree on a set of metadata relevant for their field of study and to discuss them with the data manager to make these metadata visible and searchable in the data repositories. The question of how to describe parameter, methods, quality levels, versions on the one hand and authorship, titles of datasets and references in a consistent way have to be answered along existing datasets and with different groups of interest.

Tools and services to support the data workflow

- data management plan (DMP) tool
- metadata tool
- application of persistent identifiers
- versioning systems

Data management plans describe the expected data and how and when they will be handled, stored and made available. The idea behind the latest tools is not only to assist scientists in writing the plan but to keep it up to date and keep track on the data deliverables and their success, thus leading to a comprehensive data management record.

Metadata tools can serve both as the instrument to collect the necessary metadata and also as an information system, allowing the search and dissemination of metadata and data. It is important to experience that only the metadata description efforts allow a successful search and retrieval later on.

Persistent identifiers (PID) assigned to data and metadata in their different levels (raw data, calibrated data, data products) help to make the data workflow reproducible, especially if the data handling can be documented by adding calibration routines or analysis scripts. Not only data and code but also samples (IGSN), platforms, devices and their sensors, or scientists will be made unambiguously identifiable via PIDs which allows to link back to enhanced descriptions and relations when necessary. Last but not least it allows the citation in papers to foster scientific credit.

Versioning systems play an important role in the data lifecycle. They are used to link data with scripts and documentation, allow structured sharing and collaboration among researchers and provide a convenient way for backup and restoring. One example of such a workflow is a Git project for processing AUV bathymetry data: The data is processed and at the same time documented in a Jupyter Notebook. Data and notebooks are versioned locally while on the ship. After the cruise, the repository is synchronized with the GEOMAR GitLab server, thereby automatically transfering data, scripts and documentation while preserving the processing history.

Results

Workflows for marine data publication are now fairly established at GEOMAR with respect to physical oceanography, sediment cores and underway bathymetry. The workflows for ocean modeling and (mesocosm) experiments are available but still need improvement. One major challenge is to find committed researchers or technical staff with a longterm perspective and standing in the community to ensure the continuity of the data workflow.

We also observe an ongoing need to train both scientists and data managers on existing and new systems and to explore new ways to include data output of new instrumentations. One of the essential data management experiences is the need for on-site personal contact and training of researchers during the entire data lifecycle.

Technologies for a FAIRer use of Ocean Best Practices

Pier Luigi Buttigieg, Alfred Wegener Institute (Germany), pbuttigi@mpi-bremen.de
Pauline Simpson, IOC-UNESCO Central Caribbean Marine Institute (Cayman Islands), psimpson@reefresearch.org
Jay Pearlman, Institute of Electrical and Electronics Engineers (France), jay.pearlman@ieee.org
Peter Pissierssens, UNESCO-IOC International Oceanographic Data and Information Exchange (Belgium), p.pissierssens@unesco.org
Scott Caltagirone, Element 84 (United States of America), scott@element84.com
Mark Bushnell, National Oceanic and Atmospheric Administration (United States of America) mark.bushnell@noaa.gov
Juliet Hermes, South African Environmental Observation Network (South Africa), juliet@saeon.ac.za
Emma Heslop, IOC-UNESCO Global Ocean Observing System (France), e.heslop@unesco.org
Johannes Karstensen, Helmholtz Centre for ocean Research Kiel (Germany), jkarstensen@geomar.de
Frank Muller-Karger, University South Florida, (USA), carib@usf.edu
Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Francoise Pearlman, IInstitute of Electrical and Electronics Engineers (United States of America) francoise.pearlman@ieee.org

Introduction

The publication and dissemination of best practices in ocean observing is pivotal for multiple aspects of modern marine science, including cross-disciplinary interoperability, improved reproducibility of observations and analyses, and training of new practitioners. Often, best practices are not published in a scientific journal and may not even be formally documented, residing solely within the minds of individuals who pass the information along through direct instruction. Naturally, documenting best practices is essential to accelerate high-quality marine science; however, documentation in a drawer has little impact. To enhance the application and development of best practices, we must leverage contemporary document handling technologies to make best practices discoverable, accessible, and interlinked, echoing the logic of the FAIR data principles (Wilkinson *et al*., 2016).

Implementation of an Advanced Repository for Ocean Best Practices

The future of best practice management rests upon resources which enable the efficient, targeted discovery and access of documented methodologies with innovative, community-tuned search functionality. Such systems would depend on a trusted and stable repository serving as a focal point for the harmonization of both reporting standards and technologies. Funded by AtlantOS (EU H2020), ODIP, and NSF, we are currently enhancing the IOC-UNESCO/IODE OceanBestPractices Repository (OBP-R) https://www.oceanbestpractices.net/ (Pearlman *et al*., 2017). In addition to the OBP-R's existing full text and metadata-driven functionality, we are enabling the discoverability of content through granular indexing via text-mining and

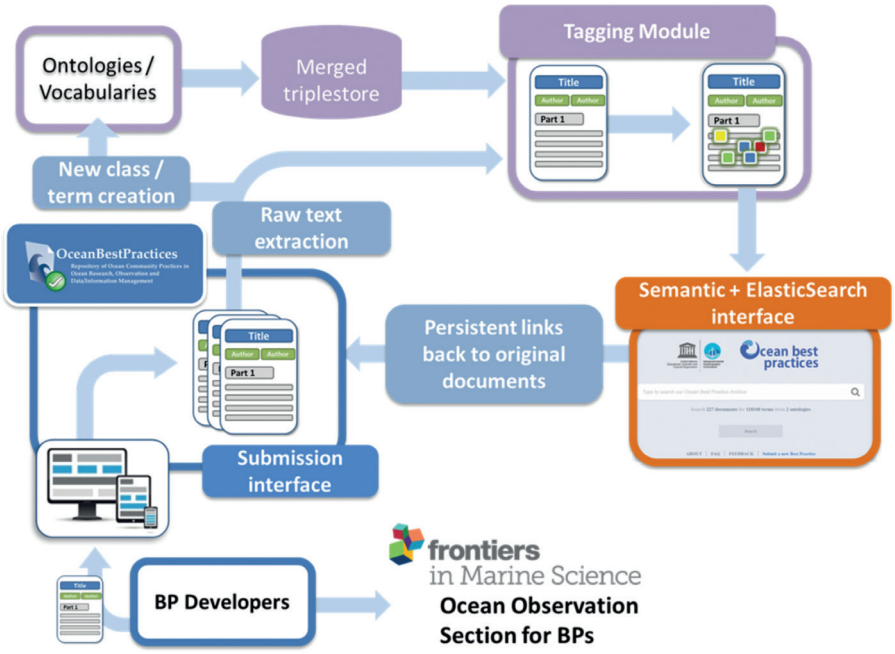


Fig. 1 - Simplified schematic of a semantically enhanced Best Practice Repository.

ontology-based semantic search tools. Our solution operates in two directions: 1) using expert knowledge represented in community ontologies (e.g. (Buttigieg *et al*., 2016) to enhance access to and mobilisation of best practices) while 2) using the content of the best practices to enrich the ontologies with ocean-relevant material and spur new research initiatives in ocean-oriented artificial intelligence.

A new, semantically enhanced ocean best practices portal has been incrementally implemented since April 2018 and is progressing through beta testing. In this contribution, we will describe the current state and future directions of our prototypical system. In the next year, we aim to integrate ontologies and thesauri covering more ocean-relevant themes, including sensors and societal goals such as the UN Sustainable Development Goals. We will develop closer technical ties to the academic journal associated with the OBP-R, ensuring that best practice developers are not only discoverable, but also credited and recognised in the literature. Further, we will expose the open source code driving each of the system's modules (Fig 1), inviting the community to offer new or revised modules to further pursue a future of Fairness for ocean best practices.

References

Wilkinson M.D., Dumontier M., Aalbersberg IJJ, Appleton G., Axton M., Baak A. *et al.,* 2016. *The FAIR Guiding Principles for scientific data management and stewardship*. Sci Data. 2016;3: 160018.

Pearlman J., Buttigieg P.L., Simpson P., Muñoz Mas C., Heslop E. and Hermes J., 2017. *Accessing Existing and Emerging Best Practices for Ocean Observation - a new approach for end-to-end management of Best Practices*. IEEE/MTS OCEANS'17 Proceedings – Anchorage, pp.1-7.

Buttigieg P.L., Pafilis, Lewis S.E., Schildhauer M.P., Walls R.L., Mungall C.J., 2016. *The environment ontology in 2016: bridging domains with increased scope, semantic density, and interoperation.* J Biomed Semantics. 2016;7: 57.

Leveraging FAIR principles to enhance SOCIB corporate data management system

Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Miguel Charcos, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mcharcos@socib.es
Sonia Gómara, Sistema d'observació i predicció costaner de les Illes Balears (Spain), sgomara@socib.es
Miquel Gomila, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mgomila@socib.es
Francisco Notario, Sistema d'observació i predicció costaner de les Illes Balears (Spain), xnotario@socib.es
Paz Rotllán, Sistema d'observació i predicció costaner de les Illes Balears (Spain), protllan@socib.es
Inmaculada Ruíz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), iruiz@socib.es
Miquel Àngel Rújula, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mrujula@socib.es
Juan Gabriel Fernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jfernandez@socib.es
Joaquín Tintoré, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jtintore@socib.es

Introduction

Nowadays, Operational Oceanography has reached a sufficient level of maturity to provide extremely helpful scientific information. It can be used to understand global processes such as climate change or mega-scale meteorology and oceanography phenomena. However, there is a clear need to harmonize and standardize methodologies and also make oceanographic data globally available and interoperable to be reused. FAIR principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals (Wilkinson D. *et al*., 2016). The following of such principles provides an excellent guidance for any organization to deliver standardized and interoperable data on a worldwide scale. This paper presents the efforts that SOCIB is undertaking to become part of a globally integrated ocean observing system through the implementation of FAIR principles in its corporate Data Management System.

The SOCIB-ICTS Marine Research Infrastructure approach

The Balearic Islands Coastal Ocean Observing and Forecasting System (SOCIB, http://www.socib.eu), is a Marine Research Infrastructure (ICTS) that provides world-class, quality controlled metocean datasets, in both real time and delayed mode. This is achieved from across its multi-platform, observation and forecasting system, covering coastal to open ocean areas. This multi-platform approach is needed to properly capture oceanographic processes, that take place at different spatial and temporal scales, and that characterise both ocean state and ocean variability. The SOCIB observation system provides physical and biogeochemical variables from different platforms such as the coastal research vessel, a high-frequency (HF) radar system, weather stations, tide gauges, moorings, drifting buoys, ARGO profilers, gliders (autonomous

underwater vehicles) and sea turtle tracking, providing trajectories given by the animals. The forecasting system uses high-resolution numerical models for hydrodynamics (Regional Ocean Model System - ROMS) and waves (Autonomous Wave Forecast System - SAPO).

SOCIB Ocean Observing Increased Value Chain

SOCIB data services and applications have been developed in line with EU-funded initiatives such as the Copernicus Marine Environment Monitoring Service, Jerico-Next, ODIP2, and MedSea Checkpoint (among others), to provide a response to scientific and social needs, by targeting different user profiles such as researchers, technicians, policy and decision makers, educators, students, and society in general. SOCIB also provides applications to (1) allow researchers and technicians to access oceanographic information, (2) provide decision support for oil spill response for open sea and coastal areas, (3) disseminate information about the coastal state for tourists and recreational users, (4) present coastal research in educational programs, and finally, (5) offer easy and fast access to marine information through mobile devices. The following elements represent the fundamentals of the SOCIB Ocean Observing Increased Value Chain - OOIVC (figure 1). The FAIR principles become a relevant framework of reference for a successful implementation and maintenance of the SOCIB OOIVC:

- The utilization of Best Practices to cover the whole SOCIB data life cycle (IODE Ocean Best Practices Repository).
- The development of a Data Management Program and a Data Quality Strategy that contemplates appropriate data formats and interoperable services. This complies with international data and metadata standards, implements world class QA/QC procedures and ensures scientific result traceability through both data and software citation (DOI policy for datasets and software tools).
- The implementation of a Quality Management System (based on the IODE Quality Management Framework) to ensure the successful delivery of oceanographic and meteorological data, products and services.
- The establishment of a structural and diversified Products and Services Strategy, able to reach a broad range of social sectors (including public and private agencies, Defence, tourism, fisheries, etc.).

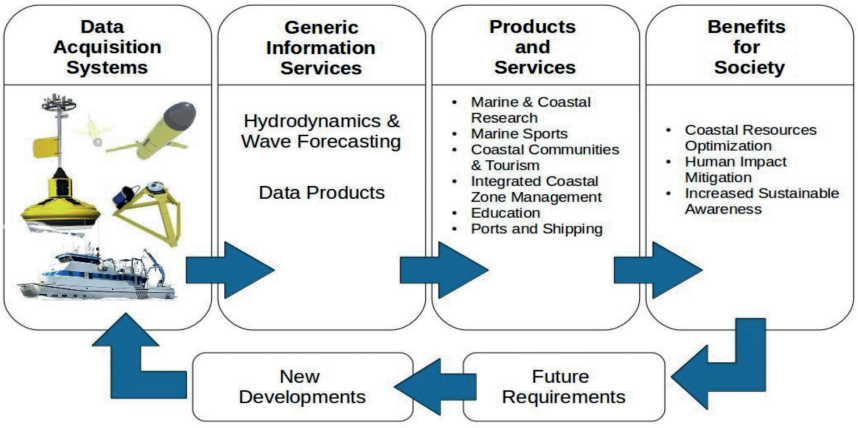


Fig. 1 - SOCIB Ocean Observing Increased Value Chain (adapted from Pinardi N., 2017).

The SOCIB Data Centre plays the main role for both ensuring the implementation of such fundamentals and the safeguarding of the whole data life cycle, leading to the adoption of FAIR data principles. This ranges from data acquisition using SOCIB observational platforms, numerical models and additional information generated by other SOCIB divisions and facilities. This enables further distribution and visualization through the development of specific tools for both discovering and visualising the datasets. It includes the implementation of dedicated web and mobile applications, relying on open source solutions.

Conclusions

SOCIB is a multi-platform distributed and integrated system that provides streams of oceanographic data and modelling services to support operational oceanography in a European and international framework, therefore also contributing to the needs of marine and coastal research in a global change context. The following of FAIR principles in our corporate Data Management Strategy is allowing a complete implementation of international standards and best practices within the SOCIB infrastructure.

References

Wilkinson M.D. *et al*. *The FAIR Guiding Principles for scientific data management and stewardship*. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016).

Pinardi N. 2017. *Moving Ocean Best Practices for Research and Applications to a New Dimension.* In, Evolving and Sustaining Ocean Best Practices Workshop, 15 – 17 November 2017, Intergovernmental Oceanographic Commission, Paris, France: proceedings. AtlantOS/ODIP/OORCN Ocean Best Practices Working Group, 12 slides.

An example of adopting and adapting SeaDataCloud and INSPIRE data models to map EMODnet nutrients data

Elena Partescano, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), epartescano@inogs.it

Alessandra Giorgetti, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), agiorgetti@inogs.it

Alessandro Sarretta, ISMAR Consiglio Nazionale delle Ricerche (Italy), alessandro.sarretta@bo.ismar.cnr.it

Dick Schaap, Marine Information Service 'MARIS' Bv (The Netherlands), dick@maris.nl

In this paper we describe an example of use of INSPIRE Data Specifications to model nutrients data acquired in the Mediterranean Sea.

This work was done to support the obligations defined in the Directive 2008/56/EC (Marine Strategy Framework Directive–MSFD) for the implementation of strategies for achieving or maintaining good environmental status in the marine environment. One of these obligations, described in the article 19(3), prescribes that Member State (MS) shall make data resulting from Article 8 and Article 11 available in agreement with the Directive 2007/2/EC (INSPIRE).

In this context, we contributed, in collaboration with the Technical Group on Marine Data (TG DATA), to realize the recommendations for the publication of datasets under the MSFD Article 19(3). These recommendations include guidelines and propose some examples and best practices (http://cdr.eionet.europa.eu/help/msfd).

The use case builds on the Technical Guidelines related to Environmental Monitoring Facilities (EF) and Oceanographic Geographical Features (OF) themes and to the Observations and Measurements (O&M) data model, and is developed in collaboration with SeaDataCloud (https://www.seadatanet.org) and MEDCIS projects.

The data used for this case study are provided by Croatian Institute of Oceanography and Fisheries (IOF) and are collected and available on the EMODnet Chemistry infrastructure. The metadata format is the SeaDataNet Common Data Index (CDI), while data are provided in Ocean Data View (ODV) format. The CDI provides an ISO19115-ISO19139 based index to individual data sets. The solution proposed in the SeaDataCloud project (we use as guideline the WP8 - Deliverable 8.6 - Part a "Review of data formats, also considering INSPIRE data models (O&M)") was adopted and adapted to deliver data in a INSPIRE-compliant way.

The classes used in this work are (see figure below):

- Environmental Monitoring Facility (EMF);
- Feature of Interest (FoI);
- Procedure (Proc) and
- Observed Property (Obs).

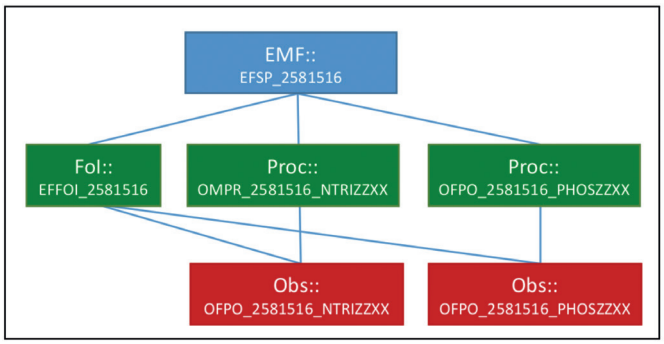


Fig. 1 - Relationship between different INSPIRE classes: in blue stem from the INSPIRE Theme EMF, the objects displayed in red stem from the specialized observations utilized for the INSPIRE Theme OF. The green objects are the area of overlap, utilized by both INSPIRE Themes.

The mapping between SeaDataCloud and INSPIRE elements was done using as source of metadata the CDI and ODV for data (see figure below) http://nodc.ogs.trieste.it/INSPIRE_compliant/INSPIREmatching_MEDCIS.xlsx.

Attribute Association role Constraint	Values/ Enumerations	Multiplicity	Example	Source	Path
Application Schema 'Environmental Monitoring Facilities' (version 4.0)			Application Schema <provide the name of the application schema>		
gml:id	NCName	1	EFSP_2581516	CDI	EFSP_ + [CDI-record id]
inspire Id	Identifier	1			
localId	CharacterString	1	EFSP_2581516	CDI	EFSP_ + [CDI-record id]
namespace	CharacterString	1			
additional Description	CharacterString	0...1		CDI	gmd:MD_Metadata/gmd:identificationInfo/sdn:SDN_DataIdentification/gmd:abstract/gco:CharacterString

Fig. 2 - The mapping between the source data (CDI and ODV format) and the target EF, OF and O&M schemas.

This mapping has been developed using the matching tables for the EF theme, as improved by SeaDataCloud and uploaded in the INSPIRE Thematic Clusters platform: https://themes.jrc.ec.europa.eu/file/view/170503/inspire-ef-matching-table.

A complete version of XML files is downloadable at the following link: http://nodc.ogs.trieste.it/INSPIRE_compliant, where an example of nutrients data acquired in Mediterranean Sea is described by INSPIRE standards. Moreover, the GML file was converted and published on the OGS/NODC Geoserver (Fig. 3). It is accessible by WMS and WFS services (http://nodc.ogs.trieste.it/geoserver/Nodc/ ows?service=WFS&version=1.0.0&request=GetFeature&typeName=Nodc:test_inspire_2581516)

This use case is a first practical example of compliance between INSPIRE themes and SeaDataCloud data and metadata. In the future, it is interest of EMODnet Chemistry to extend the type of data (e.g. contaminants) and contribute to establish together with SeaDataCloud a workflow, developing a centralized tool, to transform data formats in a compliant way.

Fig. 3 - GML available on WMS service (nodc.ogs.trieste.it/geoserver/ Nodc/wms?service=WMS&version=1.1.0&request=GetMap&layers=Nodc:INSPIRE-test&styles=&bbox=15.6768,43.1265,16.7054,44.1716&width=755&height=768&srs=EPSG:4326&format=application/openlayers)

Interoperability of new data type with SeaDataNet Infrastructure: case of flow cytometry data

Soumaya LAHBIB, Mediterranean Institute of Oceanography (France), soumaya.lahbib@mio.osupytheas.fr
Maurice LIBES, Mediterranean Institute of Oceanography (France), maurice.libes@osupytheas.fr
Gwenaëlle Moncoiffé, British Oceanographic Data Center (United Kingdom), gmon@bodc.ac.uk
Gerald Gregori, Mediterranean Institute of Oceanography (France), gerald.gregori@mio.osupytheas.fr
Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), michele.fichaut@ifremer.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
Mathilde Dugenne, Ohio State University (United States of America), dugennem@oregonstate.edu
Michel Denis, Mediterranean Institute of Oceanography (France), michel.denis@mio.osupytheas.fr
Pierre Marrec, Mediterranean Institute of Oceanography (France), pierre.marrec@mio.osupytheas.fr
Melilotus Thyssen, Mediterranean Institute of Oceanography (France), melilotus.thyssen@mio.osupytheas.fr

Background and objectives

Planktonic microbial communities play a major role in the functioning of the global ecosystem. They are good indicators of marine health due to their key role in biogeochemical cycles. Flow cytometry is a powerful technology to investigate them.

Flow Cytometry (FCM) technology measures the optical properties of single (particles) cells in alignment as they cross a light source excitation. It enables to record various fluorescences and light scattering intensities per cell, and determine the abundances of various groups. Typically, groups of pico-, nano- and microphytoplankton, heterotrophic prokaryotes, viruses, heterotrophic nanoflagellates are defined by their inherent optical properties. Some specific instruments are also able to take a picture of the single cell as it flows, giving additional taxonomical identification of cells larger than 20 μm.

This work is part of the SeaDataCloud project (H2020) whose goal is to ingest, validate and provide a long-term storage and easy access to FCM data. Through the SeaDataNet standardized system for marine and ocean data management, FCM datasets will adopt common vocabularies, ISO 19115 metadata standard, Data Transport Formats along with protocols for quality control.

Methods

Before ingesting these data, a special focus was given to the FCM common vocabulary in order to identify the most used metadata and data. First, a brainstorming work was done within small groups of scientists to come up with the most captured FCM parameters (variables). Then, a bibliography work was carried out so as to find the appropriate definitions and descriptions

of these FCM parameters. Finally, a dedicated questionnaire containing 56 questions had for objective to enlarge the impact of the study by reaching out to 180 users located all over the world.

Meanwhile, a new FCM SeaDataNet Ocean Data View (ODV) data transport format had to be created. As both physical and biological, FCM data were not suitable for the existing standardized SeaDataNet transport formats.

The FCM data management method from the instrument acquisition to the SeaDataNet ingestion is described as follows (see Fig. 1): For a project (a cruise for instance), data files acquired by flow cytometry are batch processed, converted and validated through the CytoBase Input Processor (a standalone software built on R programme by Mathilde Dugenne). Then, data integration into CYTOBASE (local database) is processed automatically using Talend (Extract Transform and Load (ETL) tool). Subsequently, CYTOBASE is connected to MIKADO (SeaDataNet tool for metadata production) to generate the Common Data Index (CDI) dataset (aggregation of measurements), the Cruise Summary Report (CSR), the European Directories of Marine Environmental Datasets (EDMED) and Marine Environmental Research Projects (EDMERP). Also, the connexion between MIKADO and CYTOBASE allows the generation of the coupling table which is the association of the CDIs (metadata) and the physical data files. Finally, The CDI and the coupling table are sent to SeaDataNet support team (cdi-support) for validation and ingestion into SeaDataNet infrastructure. The connexion between our data centre and the SeaDataNet Request Status Management service (RSM) is made thanks to the Download Manager (SeaDataNet java component tool) which was installed and has been operating since February 27th, 2018.

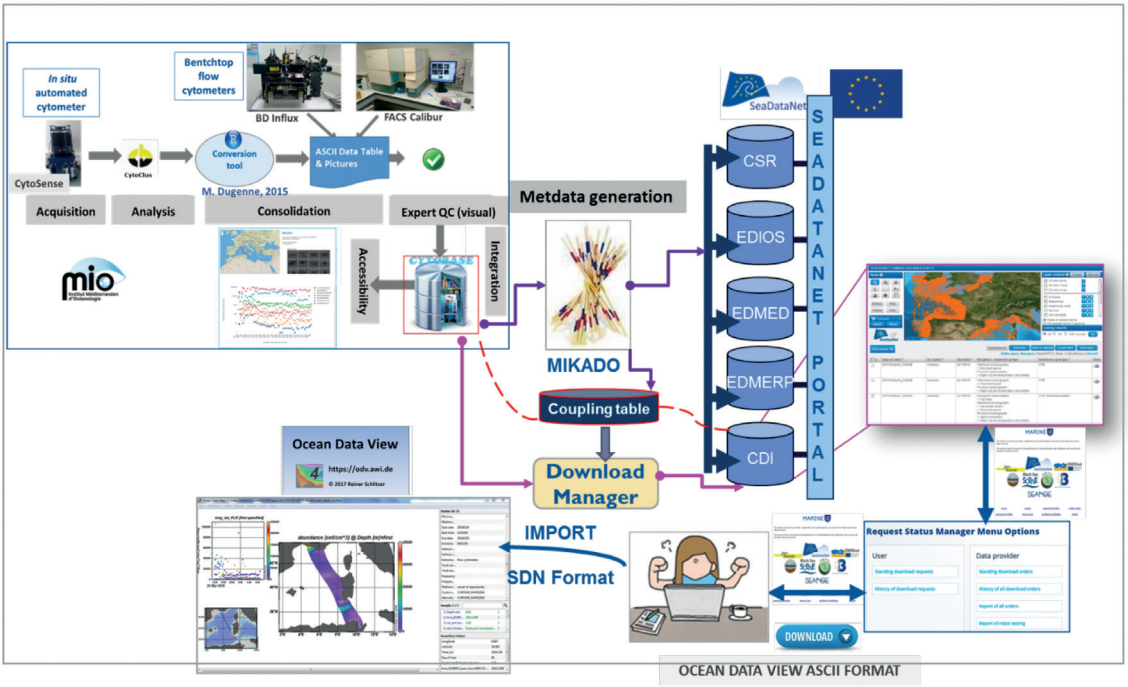


Fig. 1 - Flow Cytometry (FCM) standardized data management workflow of the CNRS-MIO.

Results

Results of this work consist of 44 new FCM common vocabulary terms that were accepted and integrated into the NERC Vocabulary Server (NVS) operated by the British Oceanographic Data Centre (BODC). Results gathered from the questionnaire allowed us to update some existing vocabularies and add new ones. We keep on ingesting new FCM metadata related to the European Directories (EDMED, EDMERP and CSR) and CDIs which are accessible through SeaDataNet infrastructure (http://seadatanet.maris2.nl/v_cdi_v3/search.asp). The CDI service allows the user to search, browse metadata and request access to data sets via a shopping basket system. Downloaded data are specifically formatted into SeadataNet ODV format for Flow Cytometry data. Every ODV file contains a set of FCM data and has a header (description of the column name (subject), column code (object), unit and used instrument) and a data table (fixed and additional fields). Specific documentation on how to use this format and examples of FCM ODV data and FCM CDI metadata files are available on the SeaDataNet webpage (https://www.seadatanet.org/Standards/Data-Transport-Formats)

Finally, FCM technology and mainly the automated sensors are revolutionizing the biological world by acquiring high resolution and real-time data about the first link of the marine food chain. Making these data sustainable, accessible and standardized will be very useful for the marine community as interoperability will greatly facilitate inter-community discussions. There is still a continuous effort to update and/or define common vocabulary, add new metadata and ingest data into SeaDataNet.

EMODnet Physics and river runoff data management

Francisco Campuzano, Maretec, Instituto Superior Técnico, Universidade de Lisboa, (Portugal), campuzanofj.maretec@tecnico.ulisboa.pt
Ana Oliveira, Maretec, Instituto Superior Técnico, Universidade de Lisboa, (Portugal),
Jorge Palma, Maretec, Instituto Superior Técnico, Universidade de Lisboa, (Portugal),
Ramiro Neves, Maretec, Instituto Superior Técnico, Universidade de Lisboa, (Portugal),
Patrick Gorringe, EuroGOOS AISBL (Belgium), patrick.gorringe@smahi.se
Giuseppe Manzella, ETT Solutions Ltd (Italy), giuseppe.manzella@ettsolutions.com
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com

Rivers runoff exert a strong influence in their neighbouring coastal area in several ways, modifying the water stratification, introducing significant fluctuations in circulation patterns and modulating the impact of upwelling events.

In the current context of a global decline of the hydrometric networks, the uncertainties include the river runoff reaching the coast and most of the water properties as temperature, salinity, etc. For this reason, river climatologies are generally imposed in the land boundaries of coastal or regional ocean models, ignoring river variability in flow and other associated properties. Anyhow, the main weakness of river climatologies is its incapacity to include the interannual variability compared to watershed model applications that are in agreement with the main river flow trends.

On the other hand, watershed models tend to overestimate river flows, especially during dry seasons. EMODnet Physics has started integrating and making availble near real time river runoff and in situ river run off trends (monthly and annual means).

Operational observations and watershed modelling forecast for the main rivers and stations near the river discharge area will be increasingly made available to the public and research community through the EMODnet physics webpage (http://www.emodnet-physics.eu/). Currently river flow from some stations in Portugal, Ireland, France, UK, Belgium, Netherlands and Germany have been made available. These data will be completed in terms of watersheds and water properties by catchment models based on the MOHID Land model (http://www.mohid.com/).

Monitoring and forecasting communities are welcoming this new data to improve the current operational thermohaline circulation in coastal areas by a better characterisation of the land-marine boundary conditions. In this presentation the current data flow, format and the future evolution of the service will be presented.

SmartDots: an online international platform for age reading workshops and calibration

Carlos Pinto, International Council for the Exploration of the Sea (Denmark), carlos@ices.dk
Julie Olivia Davies, Denmark Technical University (Denmark), joco@aqua.dtu.dk
Kevin De Coster, Institute for agricultural and fisheries research (Belgium), Kevin.DeCoster@ilvo.vlaanderen.be
Colin Millar, International Council for the Exploration of the Sea (Denmark), colin.millar@ices.dk
Wim Allegaert, Institute for agricultural and fisheries research (Belgium), wim.allegaert@ilvo.vlaanderen.be
Anna Osypchuk, International Council for the Exploration of the Sea (Denmark), Anna.Osypchuk@ices.dk
Rui Catarino, International Council for the Exploration of the Sea (Denmark), rui.catarino@ices.dk
Jane Aanestad Godiksen, Havforskningsinstituttet, (Norway), jane.godiksen@hi.no
Karen Bekaert, Institute for agricultural and fisheries research (Belgium), Karen.Bekaert@ilvo.vlaanderen.be
Neil Holdsworth, International Council for the Exploration of the Sea (Denmark), NeilH@ices.dk
Els Torreele, Institute for agricultural and fisheries research (Belgium), Els.Torreele@ilvo.vlaanderen.be

SmartDots age reading platform (http://smartdots.ices.dk)

SmartDots is a platform which facilitates exchanges and workshops to assess fish age readings. The exchanges have a coordinator who manages the event. Reports are compiled by the system for each event) and delivered to the stock coordinators to facilitate the stock assessment.

The platform is composed by four major components:

a) Web manage interface (smartdots.ices.dk)
b) Web API (web services: http://smartdots.ices.dk/webapi)
c) Software (https://github.com/ices-eg/SmartDots/raw/master/SmartDots/publish/setup.exe)
d) Reporting tool (https://github.com/ices-taf/2018_smartdotsReport)

Web interface

The web interface allows the national coordinators to :

- Manage age readers expertise
- Add age readers to the events
- Set up a new events (Exchange, Workshop, training or a demo)
- Follow the progress of an event

Software

The SmartDots client is a Graphical User Interface which gets and sends annotation data to a Web API. The SmartDots client is an open source (GPL v3.0) Microsoft WPF application.

To participate in an event, end-users authenticate using a token. Once successfully logged on they can make annotations on otolith images. Each annotation contains a line with one or more segments and a number of dots representing the annuli. The estimated age of the fish is determined based on the number of annotated annuli.

Fig. 1 - The view event page that can be accessed by the event coordinator to follow readers annotations in the event.

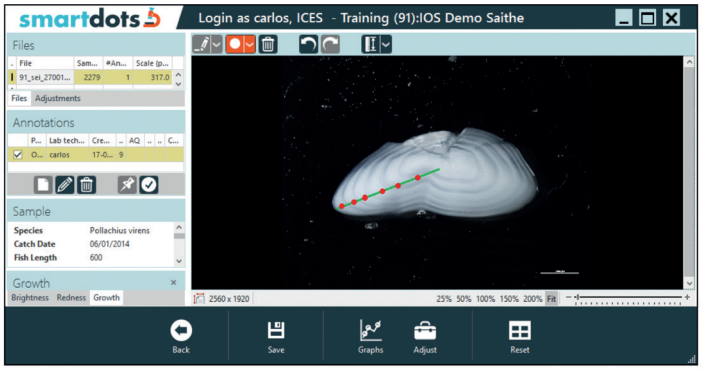


Fig. 2 - SmartDots software was initially developed at ILVO (Institute for Agricultural and Fisheries Research) in Belgium. Since 2017 SmartDots is an open source solution published on GitHub.

Web API

The Web API, the interface between the SmartDots client and the database, contains all business logic. Consequently it decides who has access, how to process and how to store the data.

Reporting tool

The reporting tool is developed in R and runs when the event is closed. The tool reads the annotations of the readers and creates the final report of the event.

Conclusion

We hope the SmartDots age reading platform to facilitate harmonized data analyses and transmission of results within a quality assured process of age estimations across all laboratories delivering data to stock assessment.

The platform is planned to be used for all intercalibrations and validations of age determination. Once the events are published all data (images, annotations and report) becomes publicly available.

The development of SmartDots within ICES is guided by the Working Group on Biological Parameters (WGBIOP).

Automated data ingestion for the Australian Ocean Data Network

Marton Hidas, Integrated Marine Observing System, University of Tasmania (Australia), marty.hidas@utas.edu.au

Roger Proctor, Integrated Marine Observing System, University of Tasmania (Australia), roger.proctor@utas.edu.au

Sebastien Mancini, Integrated Marine Observing System, University of Tasmania (Australia), sebastien.mancini@utas.edu.au

Peter Blain, Integrated Marine Observing System, University of Tasmania (Australia), peter.blain@utas.edu.au

Leigh Gordon, Integrated Marine Observing System, University of Tasmania (Australia), leigh.gordon@utas.edu.au

Angus Scheibner, Integrated Marine Observing System, University of Tasmania (Australia), angus.sheibner@utas.edu.au

Laurent Besnard, Integrated Marine Observing System, University of Tasmania (Australia), Laurent.besnard@utas.edu.au

The Australian Ocean Data Network (AODN) is an interoperable online network of marine and climate data resources. It is a collaboration between six major Australian Commonwealth Agencies, the Integrated Marine Observing System (IMOS), and a growing list of universities, state government offices and other organisations in Australia, New Zealand and the Pacific.

AODN data collections cover a large geographic area (from coast to open ocean, from equator to Antarctica), a wide range of observed parameters (physical, chemical, biological) and are obtained using a variety of platforms and technologies (e.g. ships, autonomous floats, gliders, moorings, satellites, animal tags, coastal radar). The end users include researchers, managers, policy makers, consultants, sailors and fishers. All data products contributed to the AODN are made freely and openly available to the public via the AODN Portal (https://portal.aodn.org.au/).

The development of the AODN Portal and its supporting infrastructure was guided by the diversity of data, the needs of its users, the goals of robustness, scalability, and interoperability with other programs and data sources. Wherever possible, we adopt and extend existing standards and open-source software. In particular, we create OGC-compliant Web Map Services and Web Feature Services using GeoServer, and describe data collections in ISO 19115 metadata records managed with GeoNetwork Opensource software. All software developed by the AODN is also open-source and accessible at https://github.com/aodn.

We focus on recent improvements in the data ingestion "pipeline" software that handle data submitted to the AODN for publication. At the core of this system is a new Python package that manages the common tasks of checking incoming files for compliance with conventions, moving files to public or archive storage, harvesting metadata and data into the database that forms the basis of our web services, and notifying data providers of the end result. Details of these tasks can easily be tailored for each data collection, and additional steps, such as data quality checks, the

creation of derived data products or visualisations, can be added at various stages of the process. Robustness, transparency, and fail-safe operation were key design criteria. Every file received is either published successfully, or saved to an "error" directory, with details of the error logged and email notifications sent, so that the cause can be investigated and rectified. The benefits for end users are faster publication of new data, improved consistency within data collections, compliance (e.g. with the Climate and Forecasting conventions for NetCDF), and interoperability with other collections.

A service for publishing sensors on the web using OGC and W3C standards

Louise Darroch, British Oceanographic Data Centre (United Kingdom), louise.darroch@bodc.ac.uk

Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk

Justin Buck, British Oceanographic Data Centre (United Kingdom), juck@bodc.ac.uk

Thomas Gardner, British Oceanographic Data Centre (United Kingdom), thogar@bodc.ac.uk

Cieran Brazier, British Oceanographic Data Centre (United Kingdom), cbraz@bodc.ac.uk

Gwenaelle Moncoiffé, British Oceanographic Data Centre (United Kingdom), gmon@bodc.ac.uk

Autonomous ocean observation is massively increasing the number of sensors in the ocean. Advances in this data gathering technology mean that we are generating more data than ever before, presenting challenges in data automation, handling geospatial information and linking observations to sensor metadata, the sensor's context that supports scientists on selecting quality sensors and their data. The Open Geospatial Consortium (OGC) has developed the Sensor Web Enablement (SWE) standards to facilitate integration and interoperability of sensor data and metadata. World Wide Web Consortium (W3C) Semantic Web technologies enable machine comprehensibility promoting sophisticated linking and processing of data published on the web. However, there are some practical difficulties when linking sensor content and context using the above-mentioned standards, because of internal hardware bandwidth restrictions and a requirement to constrain data transmission costs. Furthermore, Observations and Measurements (O&M), one of the core standards in the SWE suite, defines a conceptual schema encoding for observations. However, it was primarily designed for handling observations on a data point-by-point basis, requiring administrators to encode data with different dimensions and formats on a case-by-case basis, thereby reducing data integration and automation.

As part of the EU project, SenseOCEAN, the Natural Environment Research Council's (NERC) British Oceanographic Data Centre (BODC) developed the Marine Linked Systems (fig. 1), a SWE and World Wide Web Consortium (W3C) compliant sensor web publication service. Our approach addresses the practical difficulties with hardware bandwidth and transmission costs by uniquely identifying sensor and platform models and instances through URIs that act as Universally Unique Identifiers (UUIDs), which resolve via content negotiation to either OGC's sensor meta-language, sensorML or W3C's Linked Data enriched with Semantic Sensor Network (SSN) ontology. Sensor and platform model URIs and descriptions are created and hosted by the service. Sensor and platform instance URIs are dynamically created prior to and during sensor deployments, by the sensor owner, and are associated with the relevant model URI through an updatable web form, the Sensor Instance Form. Association between platform and sensor URIs

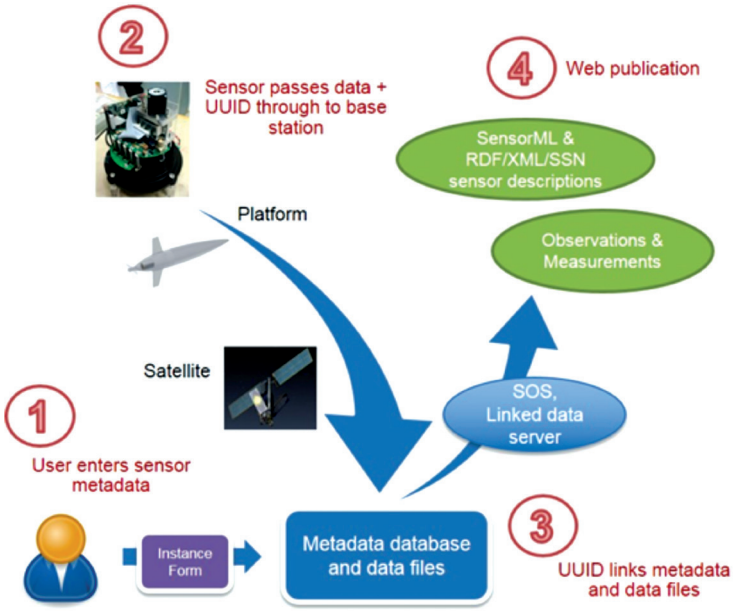


Fig. 1 - The BODC Marine Linked Systems.

is also performed by the end user, representing deployment. When sensors transmit their content, they include their unique URI to refer to their content (Fig. 1).

The service exposes multi-dimensional and formatted sensor observations using OGC's O&M. However, the model used is based on the 'out-of-band' principle, which points to external data resources including Digital Object Identifiers (DOIs) whilst still enabling the response model for the OGC Sensor Observation Service (SOS), in this instance a 52north SOS. Thus, different data types can be added to the system with minimal effort.

In addition, semantic interoperability is enhanced by constraining values using internationally established standardised lists of terms (controlled vocabularies), such as the BODC Parameter Usage Vocabulary (P01), that are published on the NERC Vocabulary Server (NVS). The NVS is a publicly available service for the marine community based on the W3C Simple Knowledge Organization System (SKOS) where each term has a unique URI that is resolvable through a RESTful interface to either HTML or RDF documents through content negotiation.

The use of URIs and the Sensor Instance Form offers both practical and economical benefits to the implementation of SWE and Linked Data standards in near real time systems. Data can be linked to metadata dynamically in-situ while saving on the costs associated to the transmission of long metadata descriptions. The transmission of short URIs also enables the implementation of standards on systems where it is impractical, such as legacy hardware. Exposing sensor observations using O&M based on the 'out-of-band' principle provides a way to easily integrate data into SWE compliant environmental networks and may be particularly cost-effective for sensor systems generating large data volumes where it may not be practical to encode each data point. Using established and web-resolvable controlled vocabularies helps to harmonise sensor data and information across these networks.

The development of this service is supported by the EU projects, SenseOCEAN, AtlantOS, EMSODev, ENVRIplus, BRIDGES and EMODNet.

Interoperable data management and instrument control experiences with the EMSO generic instrument module at OBSEA

Daniel M. Toma, Universitat Politecnica de Catalunya (Spain), daniel.mihai.toma@upc.edu
Enoc Martínez, Universitat Politecnica de Catalunya (Spain), enoc.martinez@upc.edu
Joaquin Del Rio, Universitat Politecnica de Catalunya (Spain), joaquin.del.rio@upc.edu
Óscar Garcia, UTM Consejo Superior de Investigaciones Científicas (Spain), ogarcia@utm.csic.es
Juanjo Dañobeitia, UTM Consejo Superior de Investigaciones Científicas (Spain), juanjo.danobeitia@emso-eu.org

EMSO Generic Instrument Module (EGIM) is designed to consistently and continuously acquire physical parameters of interest to a broad range of marine scientific disciplines covered by the EMSO ERIC (http://emso.eu/projects/). This research infrastructure provides accurate records of marine environmental changes coming from local nodes throughout European Seas. EGIM is able to operate on any EMSO node, mooring line, and seabed station, cabled or standalone, and surface buoy. In fact, a central function of EGIM within the EMSO infrastructure is to have a number of ocean site locations where the same set of core variables are measured homogeneously: using the same hardware, same sensor references, same qualification methods, same calibration methods, same data format and access, and same maintenance procedures.

In this paper, we present the results of the continuous shallow water tests of the EGIM, at the OBSEA site, for almost five months since its deployment and arrangement installation, followed by the acquisition of significant physical parameters in a reasonable quality range (Toma *et al*., 2017(. Also, we present our contribution to the implementation of the EGIM data acquisition system module focusses on the development of a generic software for sensor web enablement. Through this generic software, the EGIM data is directly inserted into a centralised SOS (Sensor Observation Service) server (Bröring *et al*., 2012; 52 North SOS 2.0 implementatio) and into a laboratory monitor system (Zabbix LabMonitor) for recording events and alarms. The generic software for sensor web enablement together with the SOS server is located, in the EMSO Cyberinfrastructure (CI), between the data source (EGIM) and the data management system. The generic software for sensor web enablement has two main functionalities, first to guarantee that the data is recorded properly from the EGIM hardware and second to register and insert the recorded data into a standardize OGC SOS server that works as a gateway for the EMSO data management system, as shown in Figure 1. We have identified several categories of data shared between EGIM and CI. The following define each one:

- Component descriptive data – Description of the platform/instrument configuration including instrument types, serial numbers, position of the deployment, calibration parameters.
- Command data – Commands and associated attributes such as when a command is scheduled to be executed.

- Instrument data – Data produced by the platform instruments, associated time tags, and attributes identifying the specific source instrument.
- Engineering data – Data describing the operational status of the system components.
- Metadata – Data describing the data. Metadata are data describing a resource like an instrument or an information resource.

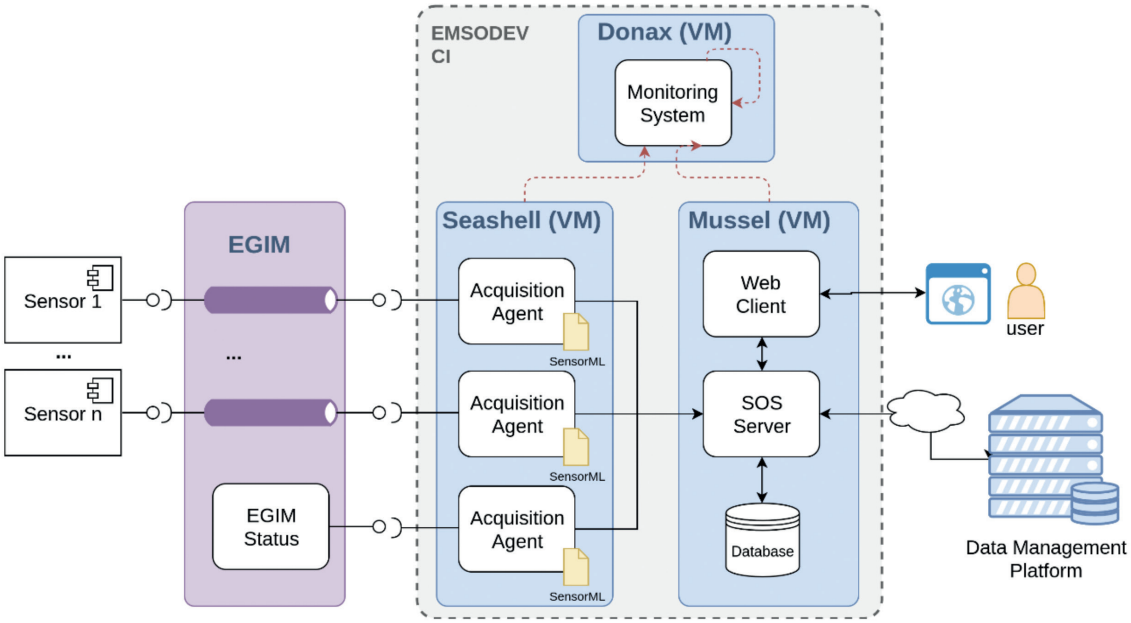


Fig. 1 - Overview of the EGIM Acquisition Components.

To provide the description of all these categories of data we use the SensorML 2.0 standard (Botts and Robin, 2014). SensorML supports the ability to describe the components and encoding of real-time data streams, and to provide a link to the data stream itself (Del Rio *et al*., 2014). Based on the SensorML description of each EGIM component, the generic software for sensor web enablement can automatically connect to a real-time data stream, parse the data stream and generate transaction compliant with Observation & Measurement standard 2.0 which can be directly injected in the OGC SOS server. The acquisition agent (based on SWE Bridge, Martinez *et al*., 2017) reads and decodes this file, encoded in EXI format. With the decoded information it autoconfigures itself, opening a communication port with the EGIM-deployed instrument through an Ethernet connection, with the capability to use both TCP and UDP protocols. Then starts getting information from the instrument in push or pull mode. The data retrieved from the instrument is stored in XML files, following the insertResult format. This format is compliant with the Observation & Measurement standard 2.0 and can be directly injected in the SOS database.

References

Toma D.M., Beranzoli L., Dañobeitia J.J., Del Rio J., Garcia O., Jesus S.M., Kassis D., Lanteri N., Marinaro G., Legrand J., Martinez E., Moreau B., Nogueras, M., Rolin J.-F., 2017. *Data collected during the first deployment of the EMSO Generic Instrument Module (EGIM) at the OBSEA cabled observatory from December, 2016 to April, 2017*. PANGAEA, https://doi.org/10.1594/PANGAEA.883072

Bröring A., Stasch C., Echterhoff J., 2012. "OGC® Sensor Observation Service Interface Standard", http://www.opengis.net/doc/IS/SOS/2.0 , 2012-04-20

52 North SOS 2.0 implementation, http://52north.org/communities/sensorweb/sos/

Botts M. and Robin A., 2014. *"OGC SensorML: Model and XML Encoding Standard,"* Wayland, MA, 01778, USA, 2014.

Del Río J., Toma M.D., O'Reilly T.C., Bröring A, Dana D.R., Bache F., Headley K.L., Manuel-Lazaro A., Edgington D.R., 2014. *"Standards-Based Plug & Work for Instruments in Ocean Observing Systems,"* Oceanic Engineering, IEEE Journal of , vol.39, no.3, pp.430,443, July 2014 doi: 10.1109/JOE.2013.2273277

Martínez E., Toma D.M., Jirka S., and Del Río J., 2017. *"Middleware for plug and play integration of heterogeneous sensor resources into the sensor web,"* Sensors (Switzerland), vol. 17, no. 12, pp. 1–28.

POSTERS

Methodology for evaluating the exploitation of distributed data providers

Evgenii Viazilov, All-Russian Research Institute for Hydrometeorological Information - WDC (Russia), vjaz@meteo.ru

Nick Mikhailov, All-Russian Research Institute for Hydrometeorological Information - WDC (Russia), nodc@meteo.ru

Denis Melnikov, All-Russian Research Institute for Hydrometeorological Information - WDC (Russia), melnikov@meteo.ru

Introduction

To assess the exploitation of distributed, heterogeneous data integration systems, IT infrastructure monitoring systems are used. In these systems the metrics are used that help to understand that data providers of integration system has reached the required indicator values. When monitoring the IT infrastructure are used threshold values of metrics. For example, the time for automatic data upload to the integrated database (IDB) should not exceed 60 minutes. The relevance of the integrated data (updating the data in accordance with the established a regulations in the metadata) should be at least 90% from the number of information resources (IR) supplied to the IDB; or the reliability of the IT infrastructure must be >98%. These indicators allow it to take the necessary actions. For example, this allows loading IR, when the computational resources are available. To increase the reliability of the IT infrastructure in the event of an incident, an automatic software reboot of one or more software components is implementing.

In addition to operational using of the results of IT infrastructure monitoring, it is necessary to monitor long-term trends in the change of various indicators, to compare the work of data providers. Moreover, here already separate indicators are not enough, it is necessary to conduct an integrated assessment of data providers work and to calculate of rating of data providers.

Proposed indicators of integrated systems exploitation

To calculate the rating of data providers, a methodology are developed for the integrated evaluation of the exploitation of distributed data providers. The purpose of developing such a methodology is to compare the exploitation level of data providers. Such an assessment allows data providers to see bottlenecks in the exploitation of the IT infrastructure.

For assessment of the exploitation of the integrated system IT infrastructure, the following groups of indicators defined:

1) reliability of the IT infrastructure of each data provider (the number of data providers in the system; reliability of the IT infrastructure; the downtime of IT infrastructure);
2) the relevance of IR provided by data providers (total number of IR integrated into the system; number of new IR; number of relevance IR; number of free IR);
3) normative accessibility of IR (percentage opened IR from general number of IR);
4) the level of information service of users (total number of views or IR downloads; average number of IR downloads; the specific weight of IR downloads in the total number of resources; the total number of IR transferred by subscription);
5) ensuring rights for access to IR - the number of assigned roles for access to IR;

6) feedback from users of the system - the number of complaints to the data.

The most important groups of indicators from the system's point of view are the reliability of the IT infrastructure, the relevance of IR and the level of information services. These indicators have the greatest contribution to an integrated assessment of the performance of data providers. In the future, when the reliability of the work and the relevance of IR will reach the planned values, their contribution can be reducing, and the contribution of a group of indicators on information services to users will increased. These contributions can install, when you configure an application that implements this method of assessing IT infrastructure. The rating of data providers is equal to the sum of the scores obtained for the first five groups of indicators. The last indicator "feedback from users" subtracts from the sum of the scores.

Most important indicators of integrated systems exploitation

Let's consider some of the important indicators (the work reliability of the data providers and the relevance of IR) in more detail. The reliability of the data providers is the percentage of time that the data provider software were on-line. The IT infrastructure downtime is equal to the time of the problem on any of the software components of the data provider. The total time of downtime is sum of time bands of idle. The monitoring system are controlling a portal, a geographic information system, an integration server, IDB, etc. For each software component, objects such as a virtual machine, network, and application are monitoring. If the network is unavailable, than the problem of network is fixing, since determination of the work other components is physically impossible.

Indicators of the relevance of IR provided by data providers are their number and relevance. The number of IRs is the quantity described resources in the form of structured files, or database tables, or object files on the every data provider. The actual date and time of updating the IR should be in accordance with the data provider's obligation to supply the data declared in their metadata. The relevance indicator is the specific weight of relevant resources in the total number IR of the data provider. He is calculate as the average ratio of the number of regularly updated resources to the total number of resources for a reporting period. This indicator evaluation based on a daily-automated verification of the composition of IR ready for work.

Integral evaluation of the exploitation of data providers is normalizing by calculating the ratio of the values of each indicator to the average value of this indicator as a whole for the entire integrated system.

Conclusions

The methodology has developed for determining the integral indicator for evaluating the exploitation of data providers in the form of their rating, which makes it possible to identify good and poorly working data providers. For the first time in this work, a set of indicators is proposed to assess the performance of data providers, some of which were previously considered only separately or not at all in other systems (for example, normative accessibility and regulation of rights to access to IR).

The considered methodology use for calculate the rating of data providers of the Russian national system "Unified State Information System on the World Ocean" (http://esimo.ru) and showed the significance of the selected indicators. This methodology can apply in the SeaDataCloud project to evaluate the exploitation of the Download Managers, which installed in more than 100 marine centers of European countries, and others system of data integration.

The article was prepared as part of scientific research on the theme "Development of methods and tools of the Integrated Arctic Observing System": a unique identifier for the project № 14.616.21.0103 Ministry of Science Russian Federation.

Open data, open innovation: data, science and industry across paradigm shifts. Can we learn from medieval scholastics?

Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), pdiviacco@inogs.it

Alessandro Busato, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), busato@inogs.it

Massimiliano Iurcev, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), miurcev@inogs.it

Mihai Burca, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mburca@inogs.it

Rodrigo Carbajales, The Abdus Salam International Centre for Theoretical Physics (Italy), rcarbaja@ictp.it

Introduction

Open innovation is an approach where the development of new products or part of them, by industrial companies, is outsourced to external entities. The rationale for this approach is rooted in several economic trend analysis. A first reason for this is to keep up with the speed with which new technologies emerge. Investing in the full cycle of development of a product can be very risky since can take a lot of time and resources. Only very rarely R&D deparments of a single company have the know-how to follow the development of the the full set of product parts. New ideas and products are therefore easier to be assembled leveraging alliances with other entities sharing therefore the risk and the costs of such development.

Open innovation can potentially become a great opportunity for universities and research centres, in fact, being them at the heart of new ideas and trends they seemingly are the perfect partner in open innovation.

Unfortunately bejond this point some issues emerge. The main difficulty is that Universities and Research centres on one side and industrial companies on the other have different missions and approaches. While the costantly under-funded scientific sector has great expectations in the possibility to have incomes from the private sector, the latter needs to protect the investments done with the former.

Open Data

Open data is a trend where data acquired by scientific institutions is shared within the scientific community and theoreticall with any possible stakeholders. Reasons for this approach span a wide spectrum. On one hand the ideals on which Science is built dictate that any finding and how they are found is shared with the scientific community and the society at large. Scientific reproducibility has always been at the very kernel of the scientific method so that results described in a scientific publication should be available for later re-analysis. Data acquired with public

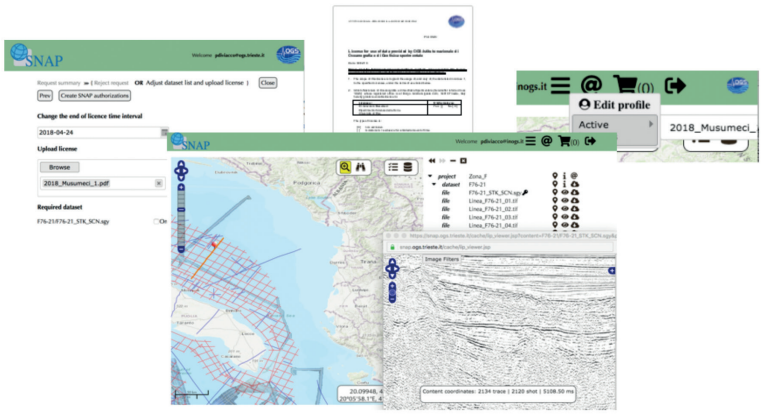


Fig. 1 - Snapshots of previewing and licensing facilities offered by SNAP.

funding, in addition, should be accessible to other scientists or projects to reduce costs sustained by the society, and maximise results and impact of investments. The open data approach triggered a great numer of initiatives, technologies and is changing the way researchers actually work.

The issue

If Open data can be considered a paradigm shift in data accessibility and therefor management, however, there are resistances. The main problem has to deal with the sensibility and economic interest behind certain types of data. In the case of an un-limited and not-controlled open access to some types of data, these can be used by commercial companies to add value to their products without any investment. This is, at least, unfair, since unbalance the relationship between the public and the private sector. In fact, from these data, some commercial entities can acquire competitive advantages, save money not investing in new acquisitions, exploit tax payers money moving revenues outside the fiscal area of the data originator, while often imposing, at the same time, protection on their data. In addition, the dangers of the exploitation of sensitive data such as environmental or some socio-economic data are evident and have been clerarly higlighted by some recent cases of scandals in social media data usage frauds.

Standing on the shoulders of giants

Alcuin of York (735-804 AD), presented us with the puzzle to take across a river a wolf, a goat and a bunch of cabbages (illaesos ultra fluvium transferre lupum et capram et fasciculum cauli). The solution he proposes is creative yet rigorous. Can we imagine something similar to take data, researchers and people from the industry, across the paradigm shifts of Open data and Open Innovation.

First of all it must be realized that open data does not mean unsupervised total possibility to download everything. While in some sectors this is needed, such as in the case of some governamental agencies, in the case of Research Institutions and Universities some filtering can be applied. This can avoid any misunderstanding on the FAIR (Findable, Accessible, Interoperabile, Reusable) principles. Filtering can take place where data (persistent) identifiers redirect users to the data, intercepting users in landing pages where descriptions and previews of

the data are made available (maybe with further copyright protection) and from where data could be made accessible on a user by user approach, depending on a licensing mechanism that has to be negotiated between the data holder and the user. Several license type can be used depending on the data holder policy, such as for example leaving full access to research institutions, or scientific paper reviewers, while allowing access to industrial companies only after the subscription of a full license of use contract, which could introduce also a fair income for the data holder.

SNAP

In the perspective mentioned above, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale – OGS, tailored its web based data management tool named SNAP, to balance Open access and commercial involvement in the domain of Geophysics. This domain is currently living a particular economic trend. The low price of Oil and the recent strict environmental regulations result in the tendency to avoid new surveys. Oil and gas companies then need data they try to obtain from the world of research institution and Universities. SNAP aims at bridging the needs and interests of the commercial word with that of research, supporting at the same time Open Innovation and Open Data.

Integrated data management of several regional projects

Asen Stefanov, Institute of Oceanology Bulgarian Academy of Sciences (Bulgaria), a.stefanov@io-bas.bg
Veselka Marinova, Institute of Oceanology Bulgarian Academy of Sciences (Bulgaria), marinova@io-bas.bg
Atanas Palazov, Institute of Oceanology Bulgarian Academy of Sciences (Bulgaria), palazov@io-bas.bg

The projects

In recent years, Bulgarian Oceanographic Institue(IO-BAS) participated in a series of projects funded by program BG02 "Integrated management of marine and inland waters" financed by the Financial mechanism of the European economic Area (EEA FM) 2019-2014. The aims of these projects were: deploying moored real time monitoring stations; installation of ferry boxes on four ships; collecting samples for priority substances and specific pollutants from water, sediments and biota; creation of tools for assessment of the marine environment through the application of new technologies and best practices to tackle the lack of marine data in two main areas of interest: detection and classification of marine litter in coastal areas, regular monitoring of marine eutrophication of surface waters; filling the gaps of data in preparation of the next assessment of the state of the marine environment, targets and indicators; developing proposals for improved monitoring programs. Data management for the following projects had to be provided - "Improved Maritime Waters Monitoring (IMAMO)", "MARine Litter, Eutrophication and Noise assessment tools (MARLEN)", "Development of the Marine Environment and the IMProvement of monitoring programs developed under MSFD (ISMEIMP)".

The data management

The data management (fig.1) of these projects was assigned to the Bulgarian oceanographic data center (BgODC). The BgOdc had to: handle the real time data coming from AAnderaa real time collector; manage and store data reliably within MS SQL spatial data base; improve and maintenance of BgOdc websites; carry out the real time data QC/QA procedures; registrate maps and data of various formats and sources in GeoServer (most popular open source geospatial server) with combination with GeoNetwork catalog; collate oceanographic data, archive and store them and maximize their utilisation; ensure the availability of high quality oceanographic data for a wide group of users using side kiosks; promote data exchange on national and international level; implement pathways to forward submitted data to the appropriate data repository in the European infrastructures that are driving the EMODnet thematic portal; publicate data on the information kiosks through WEB services to ensure security. All of these changes have greatly increased the performance of the BgODC data center operations in collecting and processing oceanographic data. All SeaDataNet technology and software tools have been used.

The presentation will demonstrate all activities during the management flow.

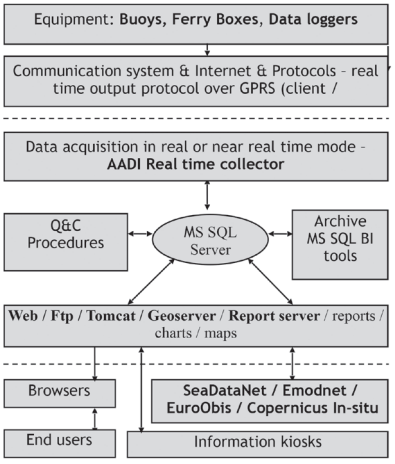


Fig. 1 - Management flow.

Building a bridge between the SeaDataNet data and INSPIRE data models

Raymond Cramer, British Oceanographic Data Centre (United Kingdom), rnc@bodc.ac.uk
Lena Hallin-Pihlatie, Finnish Environment Institute (Finland), Lena.Hallin@ymparisto.fi
Kathi Schleidt, DataCove e.U (Austria), Kathi@DataCove.eu
Seppo Kaitala, Finnish Environment Institute (Finland), Seppo.Kaitala@ymparisto.fi
Dick Schaap, Marine Information Service 'MARIS' Bv (The Netherlands), dick@maris.nl

SeaDataCloud is a four year Horizon 2020 project, which started in November 2016. The project continues the 20 year long endeavor of the SeaDataNet network to develop a Marine Spatial Data Infrastructure for data providers and users in the marine domain. The present SeaDataCloud project strives, amongst other things, to improve the INSPIRE compliance of the data collected and maintained by SeaDataNet.

In 2017 the first steps towards INSPIRE compliance were made; this work started with a review of the INSPIRE data models of relevance to the SeaDataNet community, the INSPIRE Themes Oceanographic Features (OF), Environmental Monitoring Facilities (EF) as well as the Observational Model provided within the INSPIRE Generic Conceptual Framework (GCM). Once the relevant INSPIRE Themes and classes were identified, an alignment of concepts required by the INSPIRE regulations, as well as those stemming from SeaDataNet, was undertaken.

The alignment is documented in a spreadsheet describing the correspondences between the SeaDataNet data/metadata formats and relevant INSPIRE application schemas; this mapping has

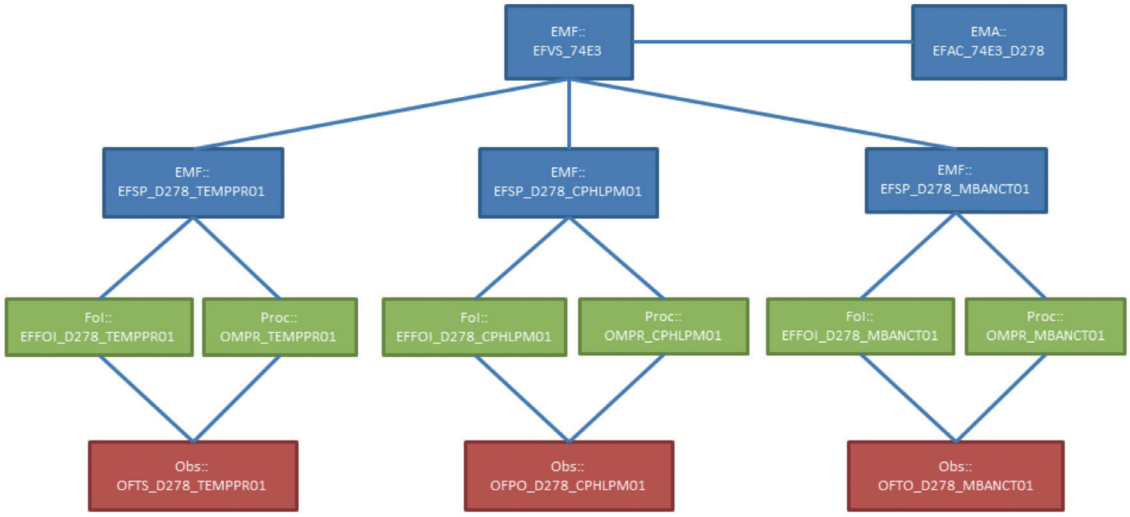


Fig. 1 - The relationships between the individual GML objects. The objects in blue stem from the INSPIRE Theme EF, the objects displayed in red stem from the specialized observations utilized for the INSPIRE Theme OF; the green objects are the area of overlap, utilized by both INSPIRE Themes.

been uploaded on the INSPIRE Thematic Clusters platform: https://themes.jrc.ec.europa.eu/file/view/170503/inspire-ef-matching-table. The fully populated examples are also available: https://themes.jrc.ec.europa.eu/file/view/169534/seadatanet-matching-table.

Type : Environmental Monitoring Facility - Vessel

Attribute Association role Constraint	Values/Enumerations	Multiplicity	Voidable/ Non-voidable	Example	Source	Path	Comment
				Application Schema <provide the name of the application schema>			
gml:id	NCName	1		EFVS_74E3	BODC Vocab C17	EFVS_ + [ConceptID]	
inspire Id	Identifier	1					
localId	CharacterString	1		EFVS_74E3	BODC Vocab C17	EFVS_ + [ConceptID]	
namespace	CharacterString	1					
additionalDescription				NERC research vessel. Rebuilt (including hull lengthening by 10m) in 1991-1992. Decommissioned in December 2012 for disposal in March 2013. She will be replaced by a new vessel of the same name.	BODC Vocab C17	Definition.notes	
name	CharacterString	0...*	v	Discovery	BODC Vocab C17	Preferred label	
media Monitored	MediaValue	1...*		water	constant for SDN		Assuming usually water, but could also be sediment or biota
geometry	GM_Object	0...1					Geometry is optional in order to support mobile facilities (i.e. ships, trucks, drones, satellites); if no explicit geometry is provided, representative Point must be provided
gml:id	NCName	1		EFVS_PT_74E3	BODC Vocab C17	EFVS_PT_ + [ConceptID]	Suggestion: base gmlid with GM_ prefix
srsDimension	positiveInteger	1		2			
srsName	anyURI	1		urn:ogc:def:crs:EPSG:4326			

Fig. 2 - The EF spreadsheet describing the correspondence between SeaDataNet data and INSPIRE schemas.

Once the alignment process was complete, sample XML/GML files were created to illustrate how SeaDataNet data could be provided in an INSPIRE compliant manner.

	CDI	ODV	GML Encoding	Matching Tab
Platform			SDN_EF_Vessel_EFVS_74E3	EF Vessel
Cruise			SDN_CruiseActivity_EFAC_74E3_D278	Activity
Time Series Data			**Temperature values, half daily**	
• **Sampling Point**	2075842	b0686762	SDN_EF_SamplingPoint_EFSP_D278_TEMPPR01	EF SamplingPoint
• **FeatureOfInteres**	2075842	b0686762	SDN_FOI_EFFOI_D278_TEMPPR01	FoI
• **Process**	2075842	b0686762	SDN_Process_OMPR_TEMPPR01	Process
• **Observation**	2075842	b0686762	SDN_PointTimeSeriesObservation_OFTS_D278_TEMPPR01	TimeSeriesObservation
Profile Data			**Chlorophyl at pressure/depths**	
• **Sampling Point**	1597207	b1061981	SDN_EF_SamplingPoint_EFSP_D278_CPHLPM01	EF SamplingPoint
• **FeatureOfInteres**	1597207	b1061981	SDN_FOI_EFFOI_D278_CPHLPM01	FoI
• **Process**	1597207	b1061981	SDN_Process_OMPR_CPHLPM01	Process
• **Observation**	1597207	b1061981	SDN_ProfileObservation_OFPO_D278_CPHLPM01	ProfileObservation
Trajectory Data			**Sea-floor depth {bathymetric depth}**	
• **Sampling Point**	2034903	b1051624	SDN_EF_SamplingPoint_EFSP_D278_MBANCT01	EF SamplingPoint
• **FeatureOfInteres**	2034903	b1051624	SDN_FOI_EFFOI_D278_MBANCT01	FoI
• **Process**	2034903	b1051624	SDN_Process_OMPR_MBANCT01	Process
• **Observation**	2034903	b1051624	SDN_ProfileObservation_OFTO_D278_MBANCT01	TrajectoryObservation

Fig. 3 - An overview of all GML examples created, together with references to their source data.

In our presentation we will present the work done, issues encountered and the recommendations for future topics for the INSPIRE and SeaDataNet communities. The results of the SeaDataCloud project have already been re-used in the TG-DATA group, outlining the guidelines for the usage of INSPIRE in the Marine Strategy Framework Directive reporting.

Scientific results traceability: software citation using GitHub & Zenodo

Charles Troupin, University of Liège (Belgium), ctroupin@uliege.be
Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Juan Gabriel Fernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jfernandez@socib.es
Miquel Àngel Rújula, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mrujula@socib.es

Background

Assigning unique identifiers to various types of objects is a common practice in the research field: data sets, scientists, papers, products... Nevertheless, the connection between the data and the final products or results is not always properly identified and documented, thus putting a damper on the reproducibility and traceability. This connection can be for example the execution of a numerical model or a software tool with a given set of parameters.

Having the code of a software tool in a version control system (VCS), with periodic releases with a unique and persistent identifier for each release contributes to the reproducibility.

The main reasons for assigning DOI to a software are:

- Reproducibility: any user shall be able to run the same experiment with identical parameters, the same dataset using the same model or software tool should obtain the same results, hence improving their credibility.
- Traceability: all the elements used in the analysis or experiment should be accessible, properly described and uniquely identified.

The present document focuses on real application of software citation in the frame of ocean data management and processing.

Citation in the context of ocean observation

Fig. 1 describes the relationships between the Ocean Observation on one side and the corresponding scientific results on the other side. It is now frequent that the datasets coming from ocean observation are assigned a DOI.

A good and illustrative example is the data citation related to the Argo profilers (see for example http://www.argo.ucsd.edu/Argo_DOIs_AST16.pdf).

The scientific results stemming from the combination of one or several datasets and different analysis methods are frequently published as articles in peer-reviewed journals, and these articles are citable using the corresponding DOI.

The intermediate part (blue items in the diagram) concerns the method or procedure from which results, products or outcomes are obtained, namely:

- the application of a numerical model using a given set of parameters (including boundary and initial conditions, forcing, …) and data or
- the utilisation of an analysis method, implemented as a software tool, on a given dataset.

A proper management of the considered piece of software and its different versions or releases is essential if one wants to guarantee the reproducibility of the results.

Procedure for the attribution of DOI to a software code

In the document *"Emerging standards for ocean data analysis and visualisation"* (submitted to ODIP), different options for software citation are analysed and compared. The following sections present an application using the Zenodo platform. That choice was based on these reasons:

- The code is free and open (https://github.com/zenodo/zenodo);
- It offers DOI versioning: either specific versions of a software or all the versions together can be cited.
- An account on ORCID or on GitHub is sufficient for login.
- There is no need to install anything, everything is stored, managed and backed-up on their servers.

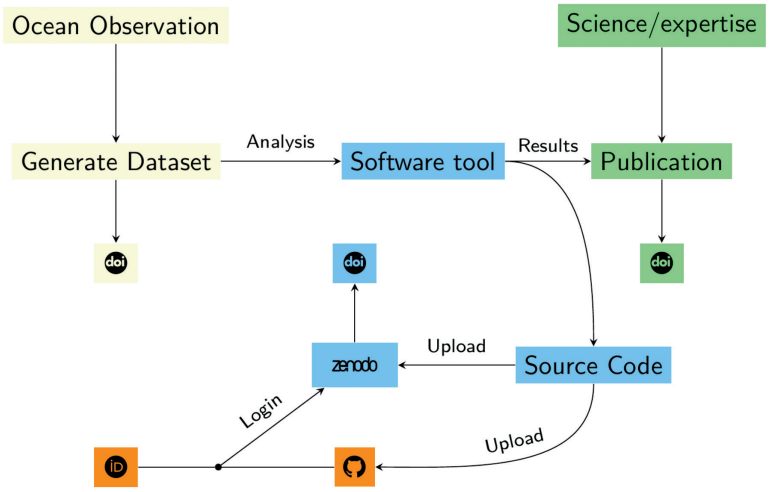


Fig. 1 - Workflow overview.

The DIVA interpolation tool

DIVA (Data-Interpolating Variational Analysis) is one of the reference tools developed within SeaDataNet. It is used to prepare the regional products, which consists of sets of gridded field obtained by spatial interpolation of all the available in situ observations collected by the data centers.

From 2008 on, DIVA code (mainly Fortran and bash) was managed via SVN and periodic releases were performed. In 2016, the code was migrated to GitHub, making it easier for the user to directly access all the versions of the code.

The GitHub repository (https://github.com/gher-ulg/DIVA) was connected to Zenodo, so that every release of the code would result in a new publication in Zenodo, accompanied with the corresponding DOI. With this approach, the DIVA users can know directly cite the version of the

code they employed to prepare the climatologies (either in publications, technical reports or even in the netCDF files themselves).

The SOCIB glider toolbox

The toolbox consists of a set of MATLAB / GNU Octave programs designed to process the raw files provided by different types of gliders. The processing steps include the format conversion, the unit conversion, some corrections on the measurements (e.g. thermal lag) and lead to the creation of netCDF files and various figures (T-S diagrams, sections, …).

The code was managed using GitHub from the beginning (https://github.com/socib/glider_toolbox) and the coupling with Zenodo was performed in 2017.

Conclusions

A platform such as Zenodo constitutes a valuable instrument for scientists and data managers that want to ensure that their software tool is properly cited and identified, with an emphasis on the version of the code employed by the users.

Software citation strives to establish an explicit clearer relationship between:

- the data, more and more published and identified in public infrastructures.
- the result products, often described in peer-reviewed publications.

This methodology allows the description of the procedure that leads to results starting from the data.

MSFD: an opportunity for harmonised data management

Thomas Vandenberghe, Royal Belgian Institute for Natural Sciences (Belgium), tvandenberghe@naturalsciences.be

Ruth Lagring, Royal Belgian Institute for Natural Sciences (Belgium), rlagring@naturalsciences.be

Hong Minh Le, Royal Belgian Institute for Natural Sciences (Belgium), hmle@naturalsciences.be

Serge Scory, Royal Belgian Institute for Natural Sciences (Belgium), sscory@naturalsciences.be

Francis Strobbe, Royal Belgian Institute for Natural Sciences (Belgium), fstrobbe@naturalsciences.be

Karien De Cauwer, Royal Belgian Institute for Natural Sciences (Belgium), kdecauwer@naturalsciences.be

MSFD: INSPIRE used as the reporting standard for metadata and data

The Marine Strategic Framework Directive strives for Good Environmental Status of marine waters by 2020 and requires the Member States to report a wide array of criteria for eleven themes or descriptors. For Belgium, the criteria cover biodiversity, habitats, population health, eutrophication, seafloor morphology, hydrology, contaminants in the environment and in seafood, macrolitter and the introduction of energy (noise). It is the first time that the (meta)data has to be reported according to the INSPIRE requirements. For Belgium, MUMM (Management Unit of the Mathematical Model of the North Sea, OD Nature, RBINS), is coordinating the monitoring activities and collaborates with experts from different scientific institutes to prepare the second assessment of the status of the Belgian marine waters. The monitoring data is managed and disseminated to the EC and the public by the Belgian Marine Data Centre (BMDC). The primary data has been collected by monitoring activities or collated from other sources by several marine specialists.

Harmonised monitoring reporting impossible without transversal approach

The very diverse array of data types (in situ or track, polygon or gridded; many data themes), the INSPIRE requirements and the necessity to maximize the reuse of the collected data have led to the need of a streamlined data flow, that creates new and incorporates existing processes. BMDC's Data and Inventory Tracking System (DITS) (Lagring et al., 2014) codebase was modified to allow the derivation of facets, that can be used to fulfill specific reporting needs and abstract away some of the functionality and metadata fields that are common to a specific reporting theme. Such facets are pluggable in the new website of BMDC. The MSFD facet allows the primary submission of data files and serves three purposes: providing an anchor for the data file(s) during the MSFD reporting by MUMM, providing an entry point for in-situ or track data

to be ingested and data managed in the central oceanographic database (IDOD) of BMDC, and fulfilling the obligation Belgium has with regards to INSPIRE.

The in-situ data falls within the INSPIRE theme 'Oceanographic Geographical Features', which makes use of the Observations and Measurements scheme to describe the data. In IDOD mappings are made with the NERC vocabularies, eg. P02 or P01, that describe the observedProperty in O&M. Surface-based data is represented as shapefiles in a GeoServer system; the shapefiles' attributes are enriched in order to make the INSPIRE transformations as easy as possible. The metadata of a DITS dataset is exposed in ISO 19115:2003 through an API which allows the harvesting by systems like GeoNetwork and the propagation to the Belgian National Spatial Data Infrastructure. Specific data transformations have been written to extract data into the INSPIRE-compliant GML format according to the recommendations by the INSPIRE maintenance and implementation group (MIG) and the TG DATA of MSFD; the transformed data is hosted at the Belgian National Geographic Institute.

Deployment of smart complex system optimizing transmission bandwidth from offshore to open seas

Marcin Wichorowski, Institute of Oceanology Polish Academy of Sciences, wichor@iopan.pl
Krzysztof Rutkowski, Institute of Oceanology Polish Academy of Sciences, kr@iopan.pl

General overview

This paper summarizes experience gained and presents results of design, development and deployment process of communication system of R/V "Oceania", the vessel of Institute of Oceanology, and development of NetBaltic Project: self organizing heterogeneous network in marine areas. The system is composed of components optimising data transfer for low bandwidth connection and foster integration of different communication systems (LTE, LEO and satellite communication respectively) into unique, transparent for users, low cost appliance ensuring access to Internet at offshore and open sea areas, both.

The system has been developed as the response for variety of problems with data transfer occurring when operations on the sea are going on, eg. switching of communication channels on the basis of QoS and cost of the transfer, encryption, security, reduction of network traffic using different techniques, priority of services in multiuser environment and remote management of the system, when IT maintenance is not available onboard.

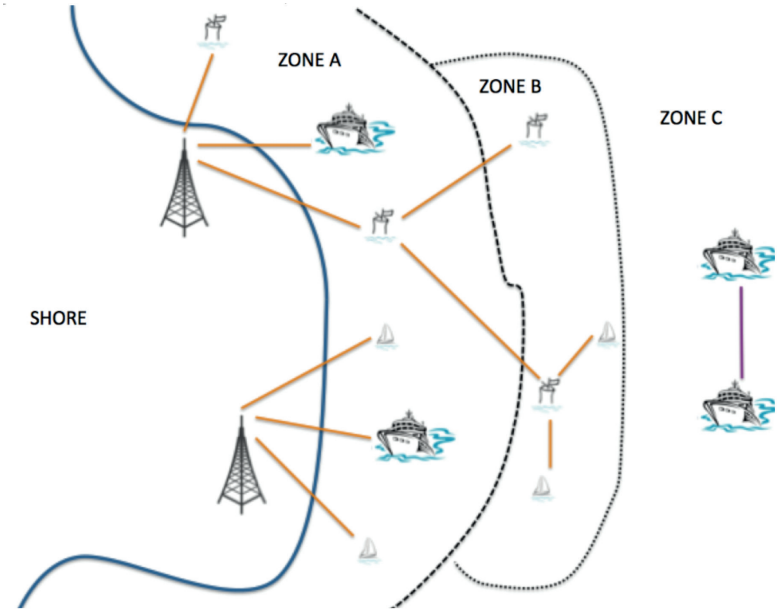


Fig. 1 - Connectivity zones.

The key factor of further development of services (both for e-navigation and scientific research conducted on sea) provided by remote platforms (vessels, buoys, drifters and other autonomous devices) is connectivity. New services demand for expanding range, widening bandwidth of connection, decreasing latency and enabling mechanisms of secure and authorised connections and topology management. In general whole domain could be divided (regarding accessibility properties) into three zones (Fig. 1). Every zone differs from the others in range, bandwidth, topology and available technologies of connectivity.

Zone A covers areas where direct connection with onshore infrastructure is available. In this areas a large variety of different communication channels is available. Remote platforms connecting onshore infrastructure are able to use potentially high bandwidth and low latency technologies. Estimated cost of data transfer is relatively low, considering the costs of the connection terminal (eg. SRC or LTE modem) and the costs of data transfer service charged by telecommunication operators. Onshore infrastructure in most cases can be developed according to the needs of the users.

Zone B extends Zone A range using platforms being in the range of Zone A and providing telecommunication infrastructure for units in range, which are not capable to use onshore infrastructure (are outside Zone A range). Depending on radio wave propagation properties (caused by weather conditions, but construction of antenna and terminal as well) these platforms can use relatively high bandwidth and still relay on low latency transmission of data. Communication depends on availability of proxy platforms in Zone A.

Zone C platforms are out of range of land based communication channels. In this zone platforms can communicate between themselves using short-range radio frequencies or establish connection on the basis of satellite communication. We can assume that platforms operating in this zone can not assure high bandwidth of data transfer and communication has high latency, either has to be buffered. These properties limit the possible applications and disable the services requiring connection protocols for data transfer or services relaying on quality of latency (eg. command and control of measurements and research). Services dependent on high bandwidth can still be provided by satellite telecommunication channels. Accessibility of this communication depends on geographical region (if geostationary satellites are used) and usually have higher requirements for power supply.

References

IMO MSC 81/23/10 Development of an E-Navigation strategy.

Report from 12th session - COMSAR (Sub-Committee on Radiocommunications, Search and Rescue), IMO, Londyn 2008.

Bronk K., Lipka A., Niski R., Waraksa M., Wojnicz P., Zurek J., *Analiza pomiarów jakości i dostępności sieci komórkowych na morzu* (Przegląd Telekomunikacyjny i Wiadomości Telekomunikacyjne) SIGMA NOT 2012 nr 4 s.412-415 CD-ROM

Garroppo R.G., Giordano S., Iacono D., Cignoni A., Falzarano M., *"WiMAX testbed for interconnection of mobile navy units in operational scenarios,"* Military Communications Conference, 2008. MILCOM 2008. IEEE , vol., no., pp.1-7, 16-19 Nov. 2008

Garroppo R.G., Giordano S. *et al.*, *"Experimental analysis of a WiMAX-satellite network for emergency management in sea areas,"* World of Wireless, Mobile and Multimedia Networks & Workshops, 2009. WoWMoM 2009. IEEE International Symposium on a , vol., no., pp.1-6, 15-19 June 2009

Kdouh H., Brousseau C., Zaharia *et al.*, *"Measurements and path loss models for shipboard environments at 2.4 GHz,"* Microwave Conference (EuMC), 2011 pp.408-411, 10-13 Oct. 2011

Kun Yang; Roste T., Bekkadal F., Husby K., Trandem, O. *"Long-Distance Propagation Measurements of Mobile Radio Channel over Sea at 2 GHz,"* Vehicular Technology Conference (VTC Fall), 2011 IEEE, vol., no., pp.1-5, 5-8 Sept. 2011

Le Roux Y.-M., Menard J., Toquin C., Jolivet J.-P., Nicolas F. *"Experimental measurements of propagation characteristics for maritime radio links,"* Intelligent Transport Systems Telecommunications,(ITST), 2009 9th International Conference, vol., no., pp.364-369, 20-22 Oct. 2009

EfficienSea Project http://www.efficiensea.org/

MarCom Project https://sites.google.com/site/marcomcommunity/

Bekkadal F. *"Emerging maritime communications technologies,"* Intelligent Transport Systems Telecommunications,(ITST),2009 9th International Conference, vol., no., pp.358-363, 20-22 Oct. 2009

Bronk K., Lipka A., Niski R. Raport: *"Measurement campaign at the Baltic Sea",* EfficenSea Efficient, Safe and Sustainable Traffic at Sea (EfficienSea), 2012

NetBaltic Project http://netbaltic.pl

The SeaDataCloud monitoring platform

Angelos Iykiardopoulos, Hellenic Centre for Marine Research (Greece), angelo@hcmr.gr
Themis Zamani, Greek Research and Technology Network S.A. (Greece), themis@admin.grnet.gr
Kostas Koumantaros, Greek Research and Technology Network S.A. (Greece), kkoum@admin.grnet.gr
Kostas Kagkelidis, Greek Research and Technology Network S.A. (Greece), kkoum@admin.grnet.gr

SeaDataCloud (SDC) is a standardized system for managing the large and diverse data sets collected by the oceanographic fleets and the automatic observation systems. The SeaDataCloud platform currently includes national oceanographic data centres of 35 countries, active in data collection. The whole platform operates a unique virtual data management system providing integrated data sets of standardized quality on-line. SDC operates a number of distributed services that range from dataset acquisition, management, replication and delivery to the end users. Through the SeaDataCloud (SDC) project the infrastructure is extended in order to efficiently store, replicate and deliver the required datasets by utilizing the EUDAT CDI. As a research infrastructure, SeaDataNet contributes to build research excellence in Europe.

In order to fulfil that it needs to be constantly monitored for the availability and rehabilitee of the whole system as well as for each service separately. In the framework of SeaDataNet a monitoring system was developed based on Nagios as monitoring engine. On top of this a monitoring portal developed form HCMR in order to perform all necessary operations to perform the appropriate availability calculations and provide reporting as well as access to service administrators to the monitoring platform.

In the framework of SDC the whole platform is going to be upgraded adopting the ARGO monitoring system as monitoring platform.

The ARGO Monitoring service is comprised of the following building blocks - components. All these components are deployed and configured to monitor SeaDataCloud:

- The messaging service which is the layover layer between the ARGO monitoring engine and the other components. ARGO Messaging Service is a real-time messaging service that allows you to send and receive messages between independent applications.
- The monitoring engine that executes the service checks against the infrastructure and delivers the metric data (probe check results) to a Messaging network (messaging).
- The POEM service which is used in order to define checks (probes) and associate them to service types.
- The connectors is a collection of libraries that periodically (usually once per day) connect to sources of truth (the monitoring service can connect to multiple external Configuration Management Databases and Service Catalogues) and deliver the information to the analytics engine in a predefined format. The various data sources established in ARGO infrastructure are mainly data from a configuration database (ex: topology, downtimes), but there's also support for fetching alternative topology via various VO feeds, weights information for various sources and POEM metric profiles.

- The analytics engine collects the metric data from the Messaging Network. It is responsible for computing the availability and reliability (A/R) of services using the metric results that are collected from monitoring engine(s), and information from the various sources of truth (ex. topology, downtimes). Results (status and A/R) are passed onto a fast, reliable and distributed data store. It supports stream processing in real time. Monitoring results flow through the messaging service, to the streaming layer (in parallel to the HDFS). The streaming layer is used in order to push raw metric results to the metric result store and to compute status results and push them to the status store in real-time.
- The Web API provides the Serving Layer of ARGO that delivers all computed status and A/R results via a programmatic interface. It is comprised of a high performance and scalable data store and a multi-tenant REST HTTP API.
- The Web UI which is the main interface for the users. It is used in order to represent the status and A/R results graphically and gives the ability to any given user to drill down from the availability of a given resource down to the actual metric results that were recorded and contributed to the computed figures.
- If there is a problem with a service an alert notification is sent via the notification engine. Based on the real-time streaming layer, the alerting functionality is introduced to the ARGO monitoring engine that goes beyond infrastructure monitoring. This component analyses the monitoring results and sends notification based on a set of predefined rules.

A fine scale spatial infrastructure for implementing networks of Marine Protected Areas: the AMAre Geoportal

Valentina Grande, ISMAR Consiglio Nazionale delle Ricerche (Italy), valentina.grande@bo.ismar.cnr.it
Federica Foglini, ISMAR Consiglio Nazionale delle Ricerche (Italy), federica.foglini@bo.ismar.cnr.it
Francesco De Leo, University of Salento (Italy), francescodeleo64@yahoo.it
Simonetta Fraschetti, University of Salento (Italy), simona.fraschetti@unisalento.it

Marine Protected Areas (MPAs) are vital for the conservation of the Mediterranean Sea, since they protect biodiversity and regulate human activities. However, many species and habitats in MPAs are still exposed to a variety of stressors. In most MPAs human activities are not spatially managed with the consequence that they can determine cumulative effects that are, very often, not quantitatively assessed.

AMAre (Actions for Marine Protected Areas) is an Interreg MED Programme project with the aim to develop shared methodologies ad geospatial tools for multiple stressors assessment, coordinated environmental monitoring, multi-criteria analyses and stakeholders' engagements. Transnational cooperation and regulations, development of coordinated best practices, data access to share information and concrete stakeholder and users' involvement are the expected results. The final aim is to scale up strategies and recommendations at transnational level adopting an ecosystem-based approach considering the goals of the Marine Strategy Framework Directive (MSFD) across MPAs.

To store, manage and share data for the monitoring and management of MPAs is foundamental to design a spatial database. Our aim was to implement a geodatabase, a common infrastructure across the MPAs included in the project, starting from the INSPIRE Directive Data Specification and framed on the AMAre needs. The geodatabase represents a critical tool also for networking the protected areas, homogenizing and standardizing data and indicators useful for monitoring changes in biodiversity and the services it provides in replicated areas. The AMAre Geoportal publishes the geodabases in an unique web application. The data are stored in a ORACLE database and published thought ArcGIS server and the Content Management System MOKA. The application is an HTML5 app and provides several functions, from visualisation to advanced quering, printing and downloading. The portal organizes and integrates the principal themes available for MPAs, from biodiversity and monitoring to environmental variables series, hydrology, hydrography, geology and administration issues.

The final aim is:

- create a common space for improving the knowledge within each MPA;
- sharing data in a common language among MPAs;
- develop a user friendly web platform with GIS functions accessible by different stakeholders;
- provide a useful tool supporting MPA managers and decision makers.

Fig. 1 - Overview of the AMAre Geoportal (https://amare.interreg-med.eu/toolbox/geoportal).

The AMAre Geoportal informs about data availability in each MPAs in the Mediterranean Basin and it is proposed as a common infrastructure for formalise and sharing information across MPAs.

Creation of Roshydromet unified parameter vocabulary services using Semantic Web technologies to ensure data integration and interoperability of information systems

Anna Maslennikova, All-Russia Research Institute of Hydrometeorological Information - WDC (Russia), amaslennikova@meteo.ru

Sergey Belov, All-Russia Research Institute of Hydrometeorological Information - WDC (Russia), belov@meteo.ru

Data related to marine environment and marine activities is contained in various sources and can be presented in different ways. To classify and store this kind of data Roshydromet is using unified parameter vocabulary for more than 10 years, which is already used in many existing Roshydromet projects. Unified parameter vocabulary is used in thematic systems for oceanography, geophysics, meteorology. For this reason, there is an emerging task of ensuring the integration of unified parameter vocabulary and replicating it to other systems using modern web technologies, also ensuring compatibility with international parameter vocabularies for its application in foreign projects of Roshydromet.

To ensure a good integration level of the vocabulary information, it is necessary to think over the general structure of the thematic domain, the relationship between the elements of the domain. The use of metadata to describe web resources is a key element in the implementation of search and reference systems on the Web. As a solution, the use of XML can be considered to create a standardized metadata syntax for the purpose of adequate cataloging and subsequent search and presentation of environmental data. But more promising is the use of the Semantic Web to improve the search technology for metadata by using semantic references and the correspondence between the concepts of the subject domain, including in the presence of its several formalized descriptions.

Semantic Web will allow to take into account the semantics of the domain of the parameter vocabulary, organize a general structure for describing the data, create a model of relationships between individual parameters, parameter groups, phenomena and processes to assess their completeness and quality. Using Semantic Web technologies like RDF and SPARQL will help to create services for convenient and quick search and replication of vocabulary information in various systems and international vocabularies based on direct “machine to machine” interaction.

With the help of such Semantic Web technology as OWL, the ontology of the unified parameter vocabulary has been defined, i.e. formalization of the subject domain and its division into classes of objects, division into classes of objects, connections and rules adopted in this field were performed. As a result of overlaying vocabulary information and metadata, the ontology was created with a complete semantic model containing data in the RDF format as provided by the Semantic Web concept. The OWL language has rich and strict semantics that makes possible

to produce explicit modeling and description of the subject area. This capability gives OWL an advantage over SKOS, whose presentation has a weaker semantics that is used for simple data retrieval and navigation tasks.

Using SPARQL (the language of requests and protocol to RDF data) and using Java technologies a REST-service was developed. Similar approaches are implemented in international vocabulary services such as BODC NVS (UK) and SISSVoc (Australia). In present implementation of the vocabulary data model, the main oceanographic parameters of the Roshydromet vocabulary (temperature, oxygen, salinity, etc.) were mapped into the BODC parameter vocabulary using URN identifiers created on the BODC NVS side. Similarly, it is planned to map the main oceanographic parameters of the Roshydromet vocabulary to the identification scheme of the Australian SISSVoc.

The use of Semantic Web technologies and Web service technologies allows expanding, replicating and integrating the existing Roshydromet parameter vocabulary without the risk of changes affecting existing systems. Currently, the services of a unified parameter vocabulary implemented via Semantic Web technology are being tested in RIHMI-WDC. Web user interfaces has been developed for searching parameters (by text, by category, etc.) on the basis of REST-service, allowing to demonstrate the operation of REST services in live mode. Future expansion of the Semantic Web technologies in Roshydromet will be led by RIHMI-WDC in the number of national systems and projects.

An open data network supporting marine planning, science, and policy

Jens Rasmussen, Marine Scotland (United Kingdom), jens.rasmussen@gov.scot
Martyn Cox, Marine Scotland (United Kingdom), martyn.cox@gov.scot
Liam Mason, Marine Scotland (United Kingdom), liam.mason@gov.scot
Drew Milne, Marine Scotland (United Kingdom), drew.milne@gov.scot
David Tulett, Marine Scotland (United Kingdom), david.tulett@gov.scot

Introduction

Marine Scotland, as part of the Scottish Government, is the marine management organisation for Scotland with the purpose of managing Scotland's seas for prosperity and environmental sustainability. This specifically means the responsibility for integrated management of Scotland's Seas, which require clear and transparent policy development and planning that are underpinned by sound evidence. Marine Scotland also conduct scientific monitoring and research to develop parts of the evidence base.

With the development in national and international data sharing, technology, and government strategies to increase transparency, Marine Scotland has developed capabilities to share evidence, and information relevant for policy and planning, not only internally, but with the public as well. Through the development of an Open Data Network of applications, Marine Scotland now provides geospatial services, contextual information, and direct access to reports and open data.

The Open Data network comprises 3 applications that work together with integrated service exchanges of data :

- Marine Scotland Information: A search engine friendly portal, that provides contextual information, lay summaries, and linkages to the Map and Data portals.

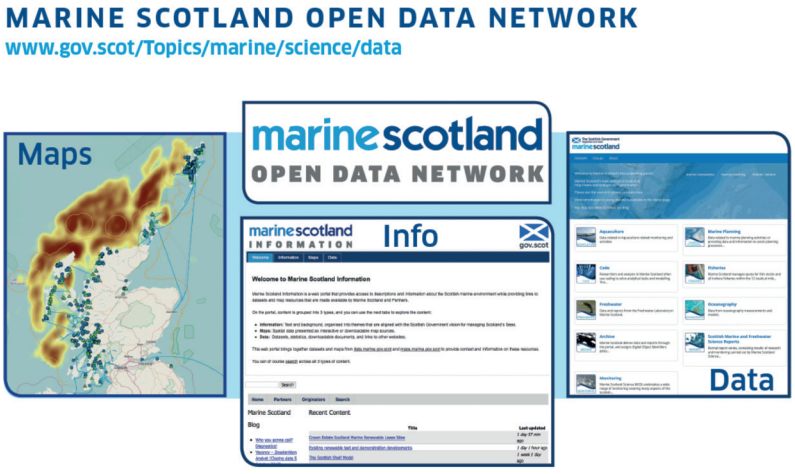


Fig. 1 - The components of the Marine Scotland Open Data Network : Maps, Information, and Data portals.

- Marine Scotland Maps : Provides a web based GIS interface for accessing more than 900 spatial data layers relevant for marine spatial planning on national and regional scales.
- Marine Scotland Data : Provides access to open datasets and reports, minted with persistent identifiers (DOI), and API services for data extraction and exploration.

Open source applications working together

Each of Marine Scotland's Open Data Network components are developed on the basis of open source software. Marine Scotland Maps is an open-source stack using PostGIS, GeoServer and a customised OpenLayers interface, allowing the ingestion and provision of web map services and download services. Marine Scoltand Information is built in Drupal, utilisng ingestion of service feeds from the Map and Data portals to improve linkage between contextual information and real resources. In addition, all content is accessible through machine readable services. The Marine Data portal is also built in Drupal, using a DKAN distribution to standardise metadata reporting to the DCAT standard, while providing data download, exploration and API acess to more than 500 resources in over 180 datasets.

The real strength in the open data network is realised through exchange of data using services, allowing the information portal to provide information pages directly accessed from the geospatial Maps portal, while also listing published open data sets that are relevant evidence to the topic, for e.g. time series.

The Open Data Network has been increasing considerably in popularity, basically trebling traffic to the portals, since all 3 applications started to work together in 2016.

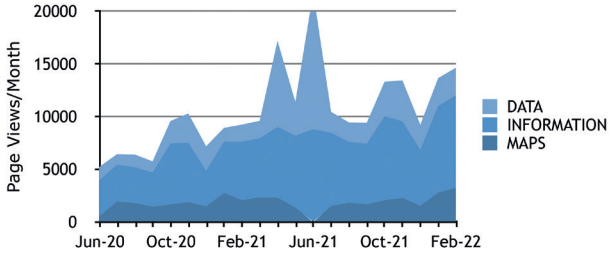


Fig. 2 - Web-traffic growth of the Marine Scotland Open Data Network.

One Network – Multiple Services

Apart from the discrete delivery roles, and the overall provision of data and information about the management of Scotland's Seas, the Open Data Network also fulfills other Marine Scotland obligations. The Data portal further acts as a landing point for Marine Scotland's role as a fisheries data archive centre in UK's Marine Environmental Data and Information Network (MEDIN), and the Maps portal links up to the Scottish public sectorSpatial Data Infrastructure (SSDI) to meet INSPIRE obligations.

In this talk, we demonstrate the capabilities of each component as well as discuss the collaborative work across organisational boundaries that has resulted in a better overall product for all parts of the organisation. We also discuss the design considerations of tackling current challenges while considering future demands and requirements for technical developments in marine information and data management.

Making sea level data FAIR

Andy Matthews, Permanent Service for Mean Sea Level (United Kingdom), antt@noc.ac.uk
Elizabeth Bradshaw, British Oceanographic Data Centre (United Kingdom), elizb@bodc.ac.uk
Mark Hebden, British Oceanographic Data Centre (United Kingdom), mahe@bodc.ac.uk

Sea level records are some of the longest ocean observations available, with the earliest continuous time series beginning in the 18th Century. The length of data available makes creating one complete findable, accessible, interoperable and reusable (FAIR) record a challenge.

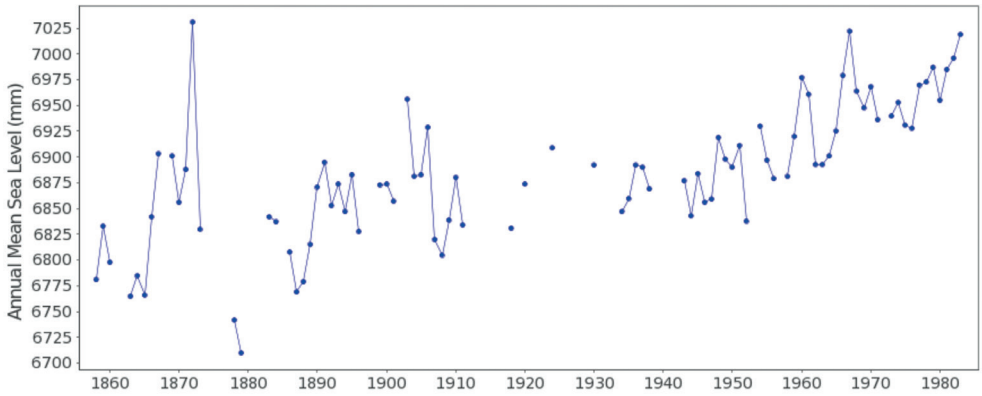


Fig. 1 - Annual mean sea level data from Liverpool, UK.

Are the data findable?

We need to make the data findable and this can be done through creating standalone discovery metadata records, such as European Directory of Marine Environmental Data (EDMED) or NASA's Global Change Master Directory entries. We describe how in the future we will add the discovery metadata to the actual data such as in a netCDF file Attribute Convention for Data Discovery (ACDD).

We can also make sea level data more accessible by assigning persistent and unique identifiers such as Digital Object Identifiers (DOIs). The Global Extreme Sea Level Analysis (GESLA) dataset has been assigned a DOI, but we look at how we will tackle the issue of assigning DOIs to datasets that are growing such as the Permanent Service for Mean Sea Level (PSMSL) and Global Sea Level Observing System (GLOSS) datasets.

Sea level discovery metadata should make use of controlled vocabularies/ontologies/taxonomies and we describe those in common use and areas for development.

How do we make the data accessible?

Global sea level data are relatively accessible compared to other Essential Ocean Variables as they are deposited in the long established GLOSS international data centres e.g. this year

PSMSL celebrates its 85th anniversary. Data in all the GLOSS data centres are freely available and organisations comply with the obligations for GLOSS members (Chapter 8: Obligations of GLOSS member states – http://www.gloss-sealevel.org/publications/documents/GLOSS_Implementation_Plan_2012.pdf). We are however looking to improve data flow between different time streams as some near real time data aren't making it into the PSMSL data bank. One of the possible reasons for this is the cost associated with quality controlling the data.

How can we make data interoperable?

Currently the GLOSS data centres each deliver data in their own format, but we are looking to use a common standard format such as CF netCDF and/or common data models to deliver data in one format to users.

We are also looking at increasing the granularity of our usage metadata. We are developing systems that will use Sensor Web Enablement (SWE) standards to help fully describe how we transform an observable property (such as the length of a piece of wire, the return time of a radar pulse, or the electric charge generated by a crystal under pressure) into a sea level measurement. This will improve the description of a time series where the sensor and platform changes many times.

We also need to ensure that we are using standard vocabularies for simple properties, such as time (ISO 8601) and country names (ISO 3166-1). For long time series, sometimes it turns out that these simple things actually aren't that simple. Is the country code that of the country where the sensor is located or where the organisation processing the data is based? There may be sites where the operating country has changed such as Hong Kong (from Britain to China).

How do we make sea level data reusable?

By storing sea level data in one of the global sea level data centres, we ensure that the data remain useable for the foreseeable future. By keeping comprehensive usage and lineage metadata alongside a dataset we will increase the reuse of the data, but also ensure that proper credit for the creation and preservation of a dataset is given.

Letting a user know what the quality of the data are and the level to which they have been screened will give confidence in the reuse of the data. Unique identifiers for data sets will help in the transparency and replicability of studies.

Operational oceanographic products supporting the e-Navigation in the Baltic Sea

Tarmo Kõuts, Tallinn University of Technology (Estonia), tarmo.kouts@ttu.ee
Leo Käärmann, Estonian Maritime Agency (Estonia), leo.kaarmann@vta.ee

Operational measurement data at sea areas gets more importance as navigation aids is becoming a natural component in e-Navigation systems. Such measurements started in operational mode mainly in coastal stations measuring parameters such as sea level and wind properties and limited number of offshore stations with wave and currents measurements. As in past, data transfer methods were resource demanding, real time oceanographic data were not very well spread among navigational community. Possibilities of nowadays wireless networks and satellite communication form excellent basis for real time data exchange. There is no common standard how operational data are presented and for example in case of the Baltic Sea it's quite hard to find one and universal data portal for mariner sailing in waters of several countries. We develop such an open web based systems, in the frame of the EfficientSea project – integrating, processing and visualizing real time oceanographic data from variety of measurement systems in Estonian waters as a dedicated web-service for e-Navigation.

Wave height in particular sea area is the most critical parameter to be considered for safe navigation. However, it exists comparably limited number of offshore on-line wave gauges. Therefore a pilot study was made in 2010, as part of the EfficienSea project, using navigation buoys as wave measurement instruments. Measurement idea is based on feature, that navigation buoys are in continuous monitoring regime, in same time equipped with motion sensors, reporting about ships collating the buoy. Navigation buoys are located in close vicinity of fairways, this is where actually the wave information is most needed. First aim of our pilot study was to develop automated wave calculation algorithm using acceleration of the buoy (motion) as input and wave height as output. Comparative wave measurements with pressure gauges in vicinity of some sample navigation buoys were made in order to develop and tune wave calculation algorithm. Results of data comparison allowed to conclude that the method is promising and that the automatic wave calculation software prototype WHAPAS (Wave Height and Period Software) gives wave height information for particulate sea area. The software is running on a shore side server, getting input motion data from navigation buoys. Wave data are intended to be broadcasted directly to mariner well AIS channels, together with other coastal station data. In order to guarantee sustainability of the developed real time data systems, low running costs are essential. Existing communication channels and protocols are used in case of navigation buoys, which ensure low power consumption, optimised data transfer etc. Wave height hind cast network based on navigational buoys, consisting of over twenty buoys, showed good stability during the run, also user feedback was positive, which draw perspectives for further development of such systems.

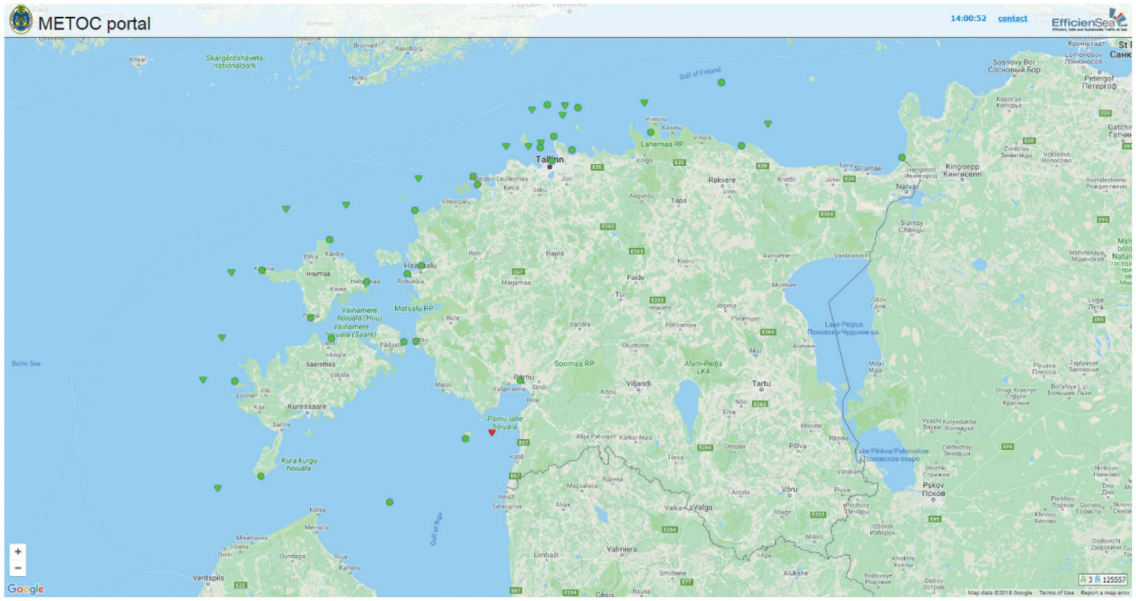


Fig. 1 - Screen view of METOC portal http://on-line.msi.ttu.ee/metoc/.

Operational oceanographic data systems are widely used for decision support by users, like mariners, search and rescue, etc. Most frequent user needs ask for straight forward, easy, and one-stop-shop systems giving right parcel of information in area of interest in right time, when it's actually needed. Recent developments of e-Navigation tools in the Baltic Sea area and particularly in Estonian coastal sea give some hints on how such user oriented navigation support systems could look, pilot solutions are designed and launched. Within the framework of the Baltic Sea Region project EfficienSea, a pilot study was made to harvest available hydro-meteorological on-line information and data and utilise these for navigation aid via AIS (Automatic Identification System) system networks. The data of the closest stations to the ship momentary position meteo- or oceanographic are transmitted to ships AIS receiver automatically as attachment to AIS M8 message. Data package is formed in AIS shore servers and currently it handles successfully with 12 station data on Estonian coast since Nov 2010 as a pilot version of the system, major upgrade made in 2018. METOC portal http://on-line.msi.ttu.ee/metoc/ contain over 50 real time data sources run by different organisations in Estonian coastal sea, more than 1200 active users of the system recorded so far, with over 120 000 clicks over three years.

Operational *in situ* oil spill detection in the Baltic Sea, using FerryBox system equipped with oil sensor

Tarmo Kõuts, Tallinn University of Technology (Estonia), tarmo.kouts@ttu.ee
Siim Pärt, Tallinn University of Technology (Estonia), siim.part@ttu.ee
Kaimo Vahter, Tallinn University of Technology (Estonia), kaimo.vahter@ttu.ee

There is a great potential for gathering scientific data using "ships of opportunity" (SOOPs), especially ferries and cargo ships cruising the same route on a regular basis. Some of the many advantages using SOOPs are: no ship costs, no energy restrictions, regular maintenance, transects are sampled repeatedly and biofouling of the installed system and sensors can be better controlled.

So-called FerryBox systems for automated measurements and water sampling by utilising ships of opportunity have reached reliability status. The installed systems can integrate data from water quality and other sensors with GPS information into a data stream that is automatically transferred from ship to shore. In general, all FerryBox systems employ a similar design - the system consists of a water inlet from where the water is pumped into the measuring circuit containing multiple sensors and gathered data is transmitted to shore via GSM/GPRS connection or satellite communication.

The Baltic Sea, with its high maritime traffic has increased probability for oil pollution occurrence. Spatial distribution of detected oil spills show that they are most probably noted on major ship routes, which lead to the idea to detect and monitor oil in water with FerryBox system on board a ferry between Tallinn and Stockholm.

The FerryBox system developed by Marine Systems Institute at Tallinn University of Technology is used on board of the M/S BALTIC QUEEN (Tallink Group). For detecting and monitoring oil in surface layer of the sea an UviLux (Chelsey Instruments Ltd) and enviroFlu-HC (TriOS) UV-fluorometers were used.

In parallel, basic seawater properties are recorded by the same system in real time – temperature, salinity, turbidity, O2 and pCO2 concentration. On M/S BALTIC QUEEN temperature, conductivity, turbidity, PAH and pCO2 are measured. For conductivity and temperature, High-Precision Pressure Level Transmitter Series 36XiW by KELLER (Switzerland) is used. Turbidity is measured with Seapoint Turbidity Meter by Seapoint Sensors, inc. (USA) and a The OceanPack™ pCO2 analyser by SubCtech (Germany).

Water for analysing is taken from the ferry's sea chest. Parameters are measured in one minute intervals, giving a 100-150m spatial resolution along the fairway. Gathered data is transferred on shore in real time. Such system enables automated asset for detection and monitoring of oil spills on fairways, where occurrence of oil spills is highest.

In general, UV (Ultra-violet) fluorescence is considered to be highly sensitive, reasonably selective, simple, rapid and straightforward method to determine oil-based aromatic compounds in

Fig. 1 - FerryBox system installed on board MS Baltic Queen http://on-line.msi.ttu.ee/GRACEferry.

seawater, even in low concentrations UV fluorescence oil detection method is used both in lab and also on field, as operable UV fluorometers are nowadays compact, with low power consumption and with high sensitivity - up to 0.001 μg/l.

UviLux UV-fluorometer measures oil compounds polycyclic aromatic hydrocarbons (PAH) concentrations (in terms of Carbazole). Sensitivity of the sensor is 0,005 µg/l or 5ppt (Carbazole), calibrated range 0,005 – 2000 µg/l, excitation light 255nm and emission light 360nm. enviroFlu-HC measures oil compounds polycyclic aromatic hydrocarbons (PAH) concentrations (in terms of Phenantrene). Sensitivity of the sensor is 0,5ppt (Phenantrene), calibrated range 0-5000 ppb, excitation light 254nm and emission light 360nm

Data from the sensors in FerryBox is collected by data logger and transferred in real time into on shore FTP server of the Marine Systems Institute, using GSM/GPRS protocol with one minute interval. GPS data and time stamps are added to the FerryBox measurement data.

Special web-based user interface is built to visualise data on-line http://on-line.msi.ttu.ee/GRACEferry where FerryBox data, ship's track, current position and gathered data can be seen both in real-time and in historical views. The web based user interface is equipped with different options: user can make selection of parameters, time periods, construct map view and 2D graphs of single and multiple parameters. Data is available in tabulated form and can be assessed regarding the quality. Individual parameters can be viewed in colour-coded view along ship's track.

All together 60 ship voyages were analysed (19.02 – 19.04.2018). PAH concentrations varied between 1-2,6 μg/L (Carbazole) and 12,4-25,5 μg/L (Phenantrene), having remarkable and quite stable variability patterns, as is with other measured parameters. Measured PAH concentrations are not absolute values, rather relative, but variability patterns can be still estimated.

Sudden concentration rises which would directly indicate oil spills, have not been detected during the observation period, all PAH concentrations stayed below those defining an oil spill. Regular maintenance of the FerryBox system was required during the measurement period, as the optical sensors were sensitive to fouling which has an impact on the data quality.

We have summarised the operational experience gathered from tests of the FerryBox systems equipped with the UV-fluorometers and showed good potential as an oil-spill detection and monitoring tool. Repeated tracks of SOOPs allow to obtain statistics of oil compounds in water in different sea areas. Especially important is the monitoring of small spills, which stay undetected with conventional remote sensing methods, but are most numerous and detectable only with in situ measurements. One drawback of the FerryBox-based oil detection system is the quite high sensitivity to biofouling, which can be handled by automated cleaning system.

We tested technology of automatic oil spill detection on board SOOP's along fairways, where probability of oil spill occurrence is highest. Oil detection is based on search of relative anomalies and not on absolute values of oil compound concentrations, which needs further research in order to compensate for the influence of natural substances like clay and chlorophyll from phytoplankton.

The IEO/NODC relational database for metadata management: improving the operational efficiency and data accessibility

Irene Chamarro León, Instituto Español de Oceanografía (Spain), irene.chamarro@ieo.es
Elena Tel, Instituto Español de Oceanografía (Spain), elena.tel@ieo.es

NODC-IEO databases context

Since 1964, the Spanish Institute of Oceanography (IEO) has headed the National Oceanographic Data Center (CEDO/NODC) is responsible for the collection, storage and distribution of marine data. Recently, the NODC has implemented an infrastructure that reduces the time required to incorporate marine data and information from its acquisition to its the permanent databases. The development and implementation of a relational database (RDB) for metadata management allows to establish a logical structure which determines a set of relationships between the data and the associated metadata.

A relational database (RDB) for metadata management

The implementation of an RDB in the NODC-IEO databases presents important advantages from the point of view of metadata management. Problems of redundancy and data inconsistency are avoided favoring the standardization. In addition, allows a quicker access to the stored information and the regular updates.

The RDB has been implemented using PostGreSQL/PostGis (https://www.postgresql.org), a database management system developed as open source.The RDB is designed to contain the necessary information to generate the Common Data Index (CDI) files used by the Pan-European infrastructure for ocean and marine data management (https://www.seadatanet.org/).The database has been also designed to store information from the Cruise Summary Reports (CSR), Marine Organizations (EDMO) and Observing Systems (EDIOS) by inventary tables, as well as common vocabularies used for all the instrumentation, and meaasured variables.

Linking the RDB with the MIKADO software tool (https://www.seadatanet.org/Software/MIKADO) allows retrieving from the RDB the necessary information to generate the CDIfiles that supports the SeaDataNet system. Moreover, it supports Europeaninitiativeslike SeaDataCloud project (1) (SDC), EMODNET (set 4: Chemistry (2), set:bathymetry (3)) and EMODNET-Data-Ingestion (4), as well as national requeriments.

Conclusions

As a national oceanographic data center, a continuous increase in the volume of information stored in its databases is expected. Regular updates of the RDB will be the key to optimizing

storage efficiency, administration and access to data, achieving the national and international commitments.

References

(1) Project SeaDataCloud (SDC). EU Contract N°730960.

(2) European Marine Observation and Data Network. CONTRACT N° EASME/EMFF/2016/1.3.1.2 - Lot 4/SI2.749773-

(3) European Marine Observation and Data Network. EMODnet Bathymetry Consortium (2016)

(4) European Marine Observation and Data Network Ingestion and safe-keeping of marine data. N° EASME/EMFF/2015/1.3.1.3/SI2.727770

Semantic web application in the context of marine remote sensing data catalogues

Dominique Briand, Institut Français de Recherche pour l'Exploitation de la Mer (France), dominique.briand@ifremer.fr

Context

The constant improvement of technologies, their central role and systematic use in marine scientific works lead to data volume increase and new data management issues. While the data volume augmentation is balanced by data processing and storing capacities augmentation of infrastructures, the context of "open data", the diversity of users and data leads to new specific questions addressed by, among others, people working on centralized services such as Data Information and Access Service (DIAS) to make it easier to discover, search, access, upload and process the data sets online.

In a world of interconnected information systems interoperability is a key word, it allows distant and independent infrastructures to communicate without any effort. Beyond technical aspects for exchanging data (http protocols, API, web services, etc.), international standards are also used to rule the way data are described or, in other words, what the metadata contains and how to serialize it (e.g.: xml specific tags). For geographic information, and specifically imagery, the ISO 19115-2 standard is commonly used. In Europe, the INSPIRE directive defines obligatory fields and controlled vocabulary (theme) that users must adopt.

However, it occurs that describing the data sets using those standards can be tricky: sometimes the fields or theme doesn't fit the data, and sometimes a data fits in various theme depending on the user's point of view. Still, in order to facilitate data discovery based on the metadata records, the use of a list of keywords is also possible and turns out to be essential.

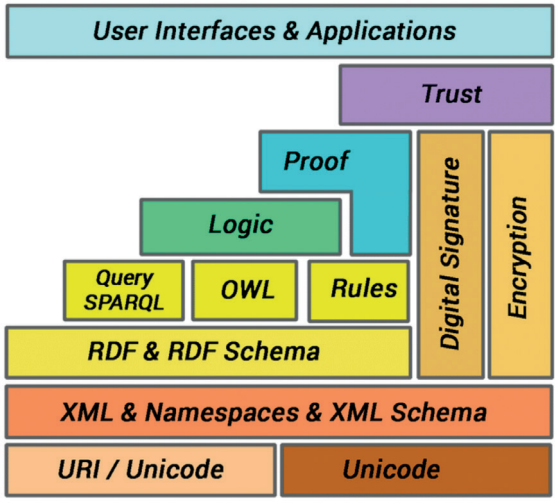


Fig. 1 - The semantic web technology stack.

Controlled vocabularies and knowledge modelling tools from the semantic web (fig. 1), like thesaurus and ontologies, are a complementary approach to metadata standards in order to express better the large spectrum of final data applications and the different user's points of view. In addition to enjoy more flexibility for describing the data sets, the use of technologies from the semantic web (also known as web 3.0 or web of linked data) provides new data searching opportunities. Indeed, the semantic web technologies enable software to better understand the content of the data and even make explicit links between them that were implicit.

Teams involved is this work

The SISMER (Scientific Information System for the Sea) is the service in IFREMER (French Research Institute for Exploitation of the Sea) in charge of the administration of numerous and heterogeneous databases and information systems. Among them, remote sensing data.

The CERSAT (Centre ERS d'Archivage et de Traitement) is the IFREMER multi-mission data centre for archiving, processing and validating data from space borne sensors (such as altimeters, scatterometers, radiometers, SAR,...) with a focus on sea surface parameters and air-sea interactions.

Application in the context of a multi-mission satellite centre

The principal challenge is to provide an efficient search tool that answers the needs of a large spectrum of users (that have different areas of interest). In other words, how to give each dataset the same chance to be retrieved? The cataloguing step is a key one as the search tool is based on the catalogue.

First a metadata application profile is defined, based on the ISO 19115-3 standard and the INSPIRE European recommendation.

Different initiatives for defining and using controlled vocabulary in remote sensing applications are used in order to define technical thesaurus. The Committee on Earth Observation Satellites (CEOS) database is the main one, it allows for every dataset to be described.

The entry point for the search tool is only the measured parameter. Even if the measured parameter is not recorded in the metadata application profile, the sensor is. An ontology is built in order to model the link between the different sensors and the different parameters.

An elastic search index is built, it combines information from the metadata application profile and the ontology. It allows for a quick response and also for users to dynamically display the more accurate facets in order to refine the results.

Perspectives

This work was done with the perspective of in-situ datasets integration, the same methodology will be deployed for cataloguing them in order to have a unique access for remote sensing and in-situ data.

Sensor metadata for automated integration of sensor resources into Research Data Infrastructure

Enoc Martínez, Universitat Politecnica de Catalunya (Spain), enoc.martinez@upc.edu
Daniel M. Toma, Universitat Politecnica de Catalunya (Spain), daniel.mihai.toma@upc.edu
Joaquin Del Rio, Universitat Politecnica de Catalunya (Spain), joaquin.del.rio@upc.edu

The study of global phenomena in the marine environment requires the combination of scientific data coming from various sources. Data is usually acquired by different institutions focused on different fields using different formats and procedures. This disparity difficult the data processing and analysis, resulting in information silos.

Fortunately, in the past years significant steps have been taken towards the standardization and interoperability of scientific data across different domains, such as the Open Geospatial Consortium's Sensor Web Enablement Framework [1]. This framework provides a suite of protocols and standards that permit the data and metadata encoding, archival and retrieval in a standardized and interoperable manner.

However, the ocean observing community uses a vast collection of sensor systems (in short: sensors) from different manufacturers, usually with proprietary communication protocols. Moreover sensors are commonly deployed in observation platforms, such as underwater observatories, buoys, underwater gliders, autonomous surface vehicles, profilers, etc. Each platform has its own architecture, with different power and communications constraints. Thus the re-usability of sensor drivers across different platforms is greatly reduced by the diverse nature of observation platforms. Creating an ad-hoc driver for each sensor-platform combination is a time-consuming task that requires in-depth knowledge of both sensor's protocol and the observation platform's architecture [2].

Moreover once a sensor data is acquired in an observation platform it is required to register or publish the acquired data into existing Research Data Infrastructures (RDIs) in order to make it available to the scientific community.

Nonetheless, it is possible to ease the sensor operation (integration a sensor into an observation platform) and the integration of sensor data into existing RDIs by using metadata structured and coherent format, i.e. Sensor Model Language (SensorML) [3]. Using this standard it is possible to generate machine-readable metadata files that comprehend extensive information about a sensor.

In general terms, sensor metadata can be classified in 4 main sections according to its purpose:

- Sensor Description: identifiers, capabilities, characteristics, etc.
- Sensor Deployment: position of deployment, platform where it is deployed, time of deployment, etc.
- Sensor History: calibrations, past deployments, maintenance operations, etc. Improves the sensor traceability.

- Sensor Interface: communication interface, data streams encoding, modes of operation, etc. Provides the information for sensor drivers to automatically detect, interface and operate the sensor.

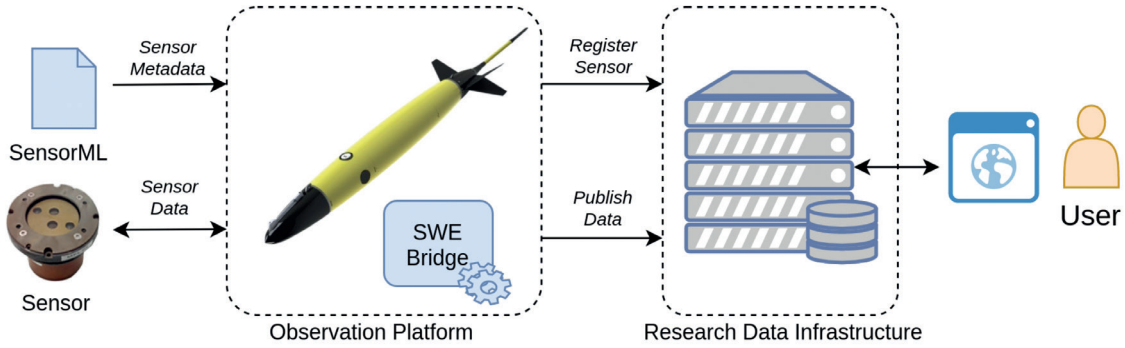


Fig. 1 - Automated Integration of a sensor based on SensorML metadata file into Research Data Infrastructure.

A SensorML metadata file describing a particular sensor deployment (or any other metadata format) can be manually linked to the sensor by the operator, or can be embedded within the sensor itself using the OGC PUCK Protocol [4].

Once all the sensor metadata has been stored into a structured file, it is possible to develop software that interpret this metadata file to automate the integration and operation processes, i.e. the SWE Bridge universal driver [5]. In figure 1, the automatic integration of a sensor into a RDI based on SensorML and the SWE Bridge software is depicted.

The SWE Bridge is a software component that decodes a SensorML file, takes the sensor interface metadata and automatically operates sensor without any a-priori knowledge. Furthermore, the SWE Bridge also takes the deployment, description metadata and automatically generates the necessary registration transactions for Sensor Observation Services [6].

References

[1] Bröring A. *et al.*, *"New generation sensor web enablement,"* Sensors, vol. 11, no. 3, pp. 2652–2699, 2011.

[2] Del Rio J. *et al.*, *"Standards-based plug & work for instruments in ocean observing systems,"* IEEE J. Ocean. Eng., vol. 39, no. 3, pp. 430–443, 2014.

[3] Botts M. and Robin A., *"OGC SensorML: Model and XML Encoding Standard,"* Wayland, MA, 01778, USA, 2014.

[4] "OGC®PUCK Protocol Standard Version 1.4," no. OGC 09-127r2. Wayland, MA, 01778, USA, 2012.

[5] Martínez E., Toma D.M., Jirka S., and Del Río J., *"Middleware for plug and play integration of heterogeneous sensor resources into the sensor web,"* Sensors (Switzerland), vol. 17, no. 12, pp. 1–28, 2017.

[6] Bröring A., Stasch C., and Echterhoff J., *"OGC Sensor Observation Service,"* Open Geospatial Consortium, Wayland, MA, 01778, USA, 2012.

SESSION 3

Marine environmental infrastructures for observation data
(data management and access)

- Operational oceanography metadata/data systems
- Physical and bio-chemical metadata/data systems
- Geophysical and geological metadata/data systems
- Fisheries and biological metadata/data systems
- Cross infrastructure activities

ORAL PRESENTATIONS

IIOE-2 data and information management

Cynthia L. Chandler, Woods Hole Oceanographic Institution (United States of America), cchandler@whoi.edu
Tobias W. Spears, DFO - BIO (Canada), Tobias.Spears@dfo-mpo.gc.ca
Pauline Simpson, Central Caribbean Marine Institute (Cayman Islands)
Peter J. Pissiersens, UNESCO-IOC International Oceanographic Data and Information Exchange Secretariat (Belgium)
Nelly F. Riama, Indonesian Agency for Meteorology, Climatology and Geophysics (Indonesia)
Ntahondi Mcheche Nyandwi, Institute of Marine Sciences, University Dar es Salaam (Tanzania)
Harrison O. Ong'Anda, Kenya Marine and Fisheries Research Institute (Kenya)
Roger Proctor, Integrated Marine Observing System/UTAS (Australia)
E. Pattabhi Rama Rao, Indian National Centre for Ocean Information Services (India)

International Indian Ocean Expedition 2 (IIOE-2)

The second International Indian Ocean Expedition (IIOE-2) is a multinational research program endorsed by the Intergovernmental Oceanographic Commission (IOC) of UNESCO in 2015. The research focus is the Indian Ocean basin with broad, trans-disciplinary science themes designed to encourage wide participation. The current science implementation plan (1) calls for the full range of ocean research platforms including research vessels, vessels of opportunity, mooring arrays, floats, gliders, satellite imagery, lab-based experiments and modelling efforts. Securing funding for such a complex, coordinated, multinational research program is difficult. The plan is to coordinate the research via joint program offices hosted at IOC Perth and the Indian National Centre for Ocean Information Services (INCOIS) in Hyderabad, India, complemented by several national efforts. The research will be carried out by investigator teams that must arrange for their own research grants, funded within their respective nations. While the IIOE-2 researchers recognize that management of data and information is critical to the success of the research objectives, funding a program-specific data and information management effort for IIOE-2 is extremely difficult.

IIOE-2 Data and Information Management Needs

Multinational, transdisciplinary research requires that program participants share and provide access to data and results. To facilitate access to and interpretation of the data, those data must be accompanied by rich metadata. The success of the IIOE-2 science plan will require data to be supported by complete metadata records that describe who, what, where, when, why and how the data were collected, processed and analyzed. Descriptions of sampling and analytical protocols as well as quality assurance and control procedures are essential to accurate interpretation of the data especially when the research themes are transdisciplinary. The majority of IIOE-2 data will be generated by in situ field studies and it is especially important that such environmental data be preserved in archive data centers. Generating standards-compliant metadata and use of open,

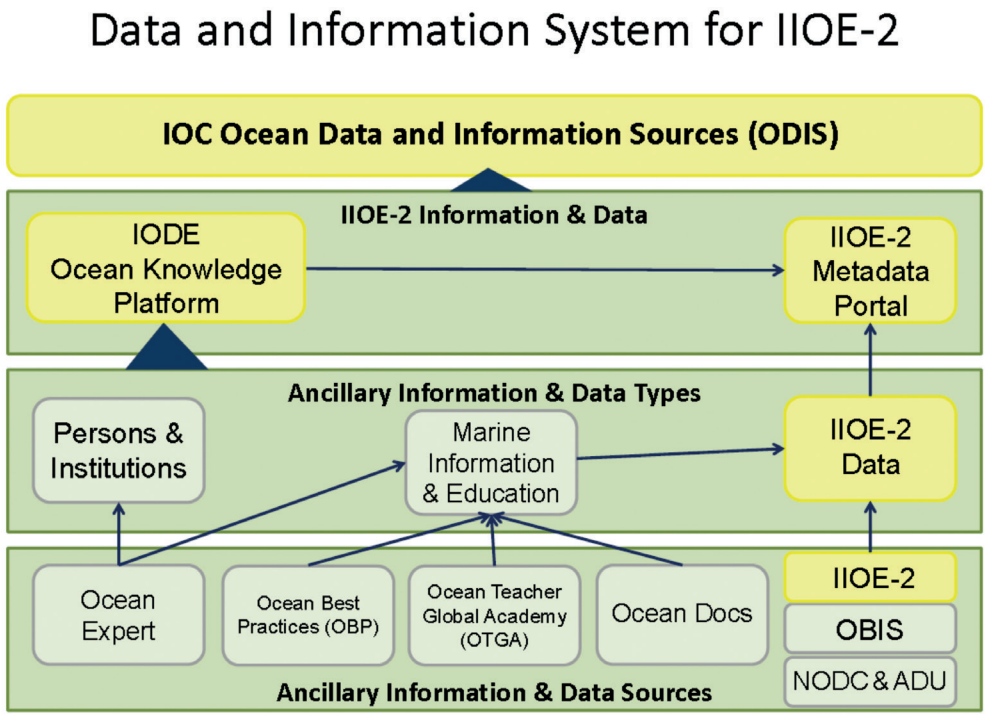


Fig. 1 - Proposed architecture for IIOE-2 data and information management that leverages many existing IODE projects, capabilities and services (green) with new components shown in yellow.

non-proprietary data formats where possible, will help facilitate data and information exchange between centers. Good data management and sharing of best practices have always been integral to successful scientific research, but advances in technology and increased expectations for data sharing have resulted in even greater requirements for management of research data.

Leveraging IODE Capabilities and Services

The International Oceanographic Data and Information Exchange (IODE) of the IOC of UNESCO, established in 1961, has grown into a networked community of National Oceanographic Data Centres (NODCs) and Associate Data Units (ADUs) and runs, sponsors or coordinates several programs that can provide effective solutions for IIOE-2 data and information management challenges. In February 2017, the IIOE-2 Scientific Steering Committee decided to leverage existing components of the IODE network to provide comprehensive data and information management for their program (Fig. 1).

IIOE-2 Data and Information Management System

An IIOE-2 metadata portal has been established at INCOIS in Hyderabad, India (http://www.iioe-2.incois.gov.in/IIOE-2/data.jsp). At the moment, all nations proposing IIOE-2 research are IOC Member States and have a NODC, ADU or OBIS node that is part of the IODE network of data repositories. The IIOE-2 Data Policy recommends that data from IIOE-2 endorsed research projects be submitted to the appropriate NODC or ADU, with project and deployment metadata also submitted to the IIOE-2 Data Office at INCOIS. When ready, data packages including

metadata should also be submitted to the appropriate NODC or ADU and eventually to INCOIS. Related information is already being managed in existing IODE systems including: names of marine professionals and their affiliations in OceanExpert (https://www.oceanexpert.net/); an open access repository of digital format marine science publications including preprints, published articles, technical reports, and working papers in OceanDocs (https://www.oceandocs.org/); a secure, permanent repository of published, recommended, community practices including the IOC Manuals & Guides and Ocean Data Standards (https://www.oceanbestpractices.net/); and the extensive training materials available from Ocean Teacher Global Academy (https://classroom.oceanteacher.org/) as well as the facilities of the OTGA Regional Training Centers, four of which are in the Indian Ocean region. The IIOE-2 data and information will then be included in the nascent IOC Ocean Data and Information Sources platform that is being developed to capture results from all IOC endorsed research programs.

References

(1) Hood R.R., *et al.*, 2015. *"Science Plan of the Second International Indian Ocean Expedition (IIOE-2): a Basin-Wide Research Program"*. Scientific Committee on Oceanic Research, Newark, Delaware, USA. http://hdl.handle.net/1834/9675

Enhancing the integration of in-situ Atlantic Ocean observation data and services to users

Valérie Harscoat, Institut Français de Recherche pour l'Exploitation de la Mer (France), Valerie.Harscoat@ifremer.fr

Sylvie Pouliquen, Institut Français de Recherche pour l'Exploitation de la Mer (France), Sylvie.Pouliquen@ifremer.fr

AtlantOS Work Package VII partners, Institut Français de Recherche pour l'Exploitation de la Mer (France), atlantos_wp7_coordination@ifremer.fr

The targeted European data system within AtlantOS project is not a new system but integrates existing data systems, these being enhanced to ingest and deliver more in-situ observation data on Atlantic Ocean and to better serve the users, in a harmonized way across the systems. The existing data systems are diverse in-situ observing network systems operating in the Atlantic Ocean and existing European and international data infrastructures and portals, termed integrators (e.g. Copernicus INS TAC, SeaDataNet NODCs, EMODnet, ICES, EurOBIS, GEOSS).

The actors of such integrated system are overall mature systems with long-term experience and established procedures for data collection and management often agreed at international level. Consequently, trying to implement a sovereign and rigid set of rules for all the actors to comply with, would be highly challenging and not in the best interest for an efficient Atlantic Ocean observing system able to deliver appropriate products to users. Therefore, the AtlantOS community recommended to

- Rely on existing European and international standards and protocols, first focussing on metadata by implementing a minimum set of mandatory information and using agreed vocabularies at all level of the processing chain (Common Vocabulary for parameters, Common Unique ID for Platform and code for Institution, etc.) that allows easier traceability of the observations along the processing and distribution dataflow.
- Encourage open and free data policy.
- Focus on data quality by implementing a set of common near Real Time QC procedures for 7 Essential Variables (Temperature, Salinity, Current,

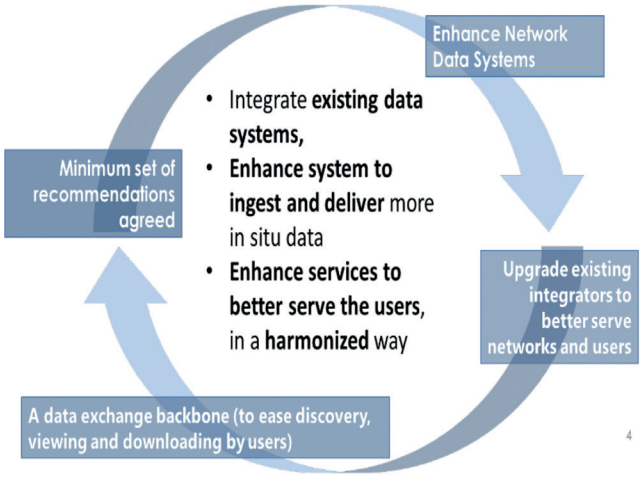


Fig. 1 - Integrated data system of AtlantOS.

Sea-Level, Oxygen, Chlorophyll, Nitrate and Carbon) acquired in near real time (distributed within a few hours to several days).

- Enhance access to network data by setting up a unique entry point to discover and download existing data, either by integrating the data in existing international network Global Data Centres or by setting up new ones.
- Connect to existing integrators.
- Enhance monitoring facilities offered by JCOMMOPS.
- Document the existing services through a catalogue that describes the existing network and integrator services and allows connection to the Global Earth Observation System of Systems (GEOSS).

To move toward the integrated data system, the roadmap setup for networks is to implement the recommendations for harmonization across networks and facilitate access to their data. Improvements are underway in most networks although some actions will probably last beyond the AtlantOS project. The major advances are : (1) setting up two GDACs (Global Data Assembly Centres) for Drifters, one in Europe and one in Canada, as central points of data access in Near Real Time and for the best Delayed Mode version; (2) UK and Norwegian data integration underway in the GDAC for Gliders, to improve access to ADCP data for GO-SHIP (with the long-term goal to set up a GDAC); (3) setting up nodes underway in the SeaDataNet infrastructure (SOCAT carbon data from VOS/SOOP and GO-SHIP networks, physical data from CPR network, planned direct GDAC data flow connection to SeaDataNet for Argo, Gliders, Drifters and OceanSITES networks); (4) integration of data from European Tracking Network in EMODnet Biology.

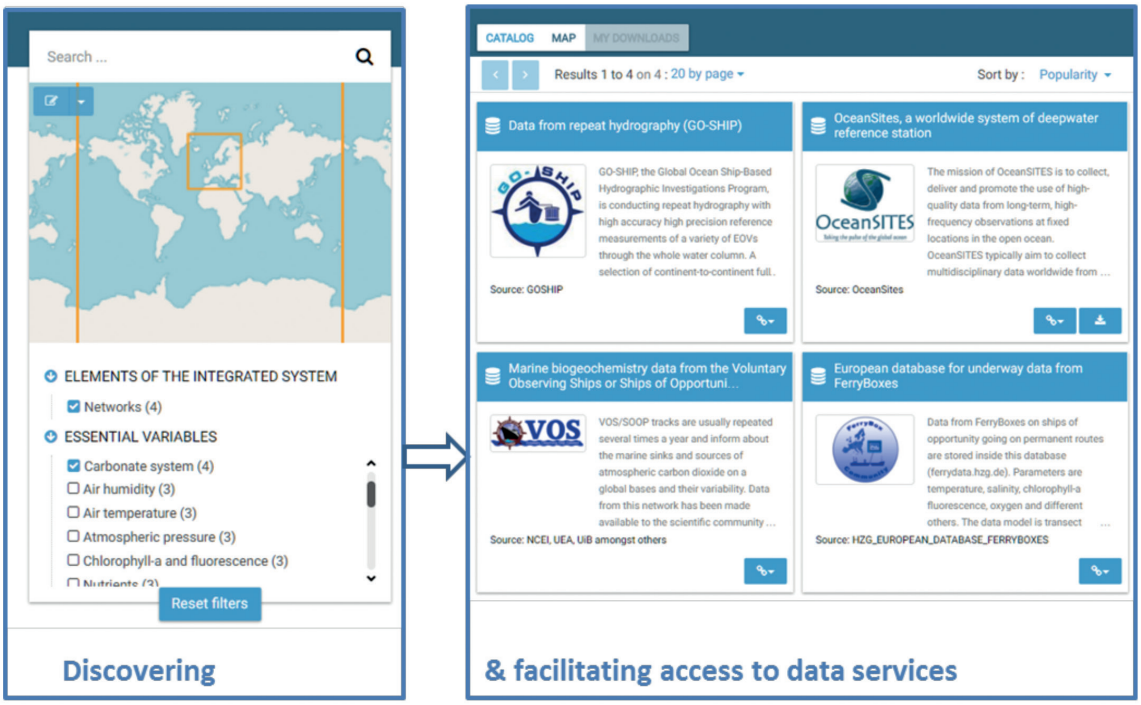


Fig. 2 - AtlantOS catalogue.

For integrators, the roadmap is to enhance their services on network side (ingestion tools to integrate more data, cross network assessments and feedback to networks) and enhance the services on the side of data providers and users (viewing, downloading and traceability/monitoring). Improvements are underway in all integrators where more network data are integrated, and enhanced services are developed. In this framework, the HF Radars have been already integrated in EMODnet Physics and they are going to be integrated in Copernicus INS TAC. The Copernicus INS TAC scope will also include CO2 even after AtlantOS.

Also common services designed within AtlantOS project are made available to users. On the one hand, a catalogue (https://www.atlantos-h2020.eu/project-information/atlantos-catalogue) provides the users with a discovery service of the networks, integrators and products related to Atlantic Ocean, and it facilitates the access to services (viewing, downloading and monitoring) of the existing data systems.

The sustainability at European level of such catalogue beyond AtlantOS will be achieved through GEOSS and EuroGOOS. Catalog content, services, data and products available in the integrated system are progressing as existing systems move forward in achieving the goals of data harmonization and integration.

On the other hand, monitoring services designed under the EuroGOOS and JCOMMOPS umbrellas are made available (monitoring dashboard http://www.emodnet-physics.eu/atlantos/dashboard , browser of monthly status maps of all operational JCOMMOPS platforms on the Atlantic Ocean http://www.jcommops.org/board/wa/Archives?t=JCOMMOPS&availableMaps=jcommops-atlantic&displayedMap=jcommops-atlantic), and also traceability services that aim to give visibility on data usage to data providers.

JCOMMOPS integrated monitoring system

Mathieu Belbéoch, JCOMM in situ Observing Programmes Support Centre (France), mbelbeoch@jcommops.org
Long Jiang, JCOMM in situ Observing Programmes Support Centre (Switzerland), ljiang@jcommops.org
Martin Kramp, JCOMM in situ Observing Programmes Support Centre (France), mkramp@jcommops.org
Anthonin Lizé, JCOMM in situ Observing Programmes Support Centre (France), alize@jcommops.org
Emanuela Rusciano, JCOMM in situ Observing Programmes Support Centre (France), erusciano@jcommops.org

The Joint WMO-IOC Technical Commission for Oceanography and Marine Meteorology *in situ* Observations Programme Support Centre

JCOMMOPS operates within the JCOMM Observations Programme Area and its Observations Coordination Group (OCG), and monitors the status and performance of the following global ocean observing networks: Argo profiling floats, DBCP surface drifters, tropical and coastal moored buoys, tsunameter buoys, ice buoys, SOT (observations from volunteer ships), GO-SHIP hydrographic lines, OceanSITES multidisciplinary reference stations, GLOSS sea level tide gauges, and Ocean Gliders.

JCOMMOPS assists in the implementation and deployment of the observing networks, and in

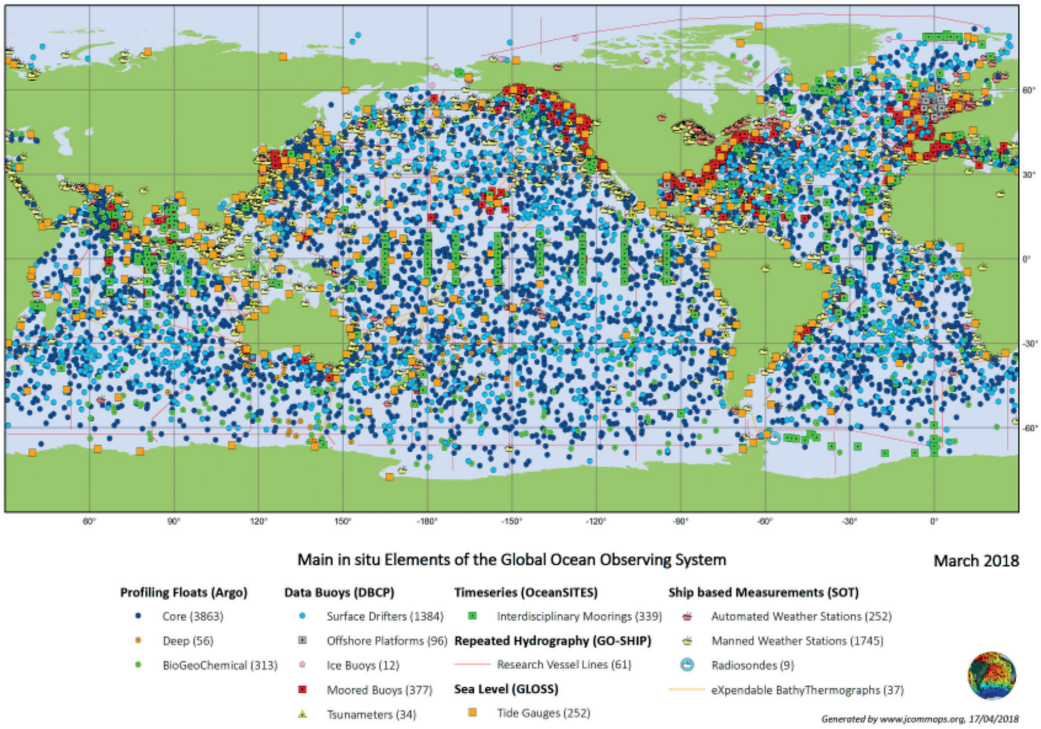


Fig. 1 - Main in situ elements of the Global Ocean Observing System.

data and metadata exchange. It maintains and develops a consistent set of tools needed to monitor the status of the observing networks.

Overall JCOMMOPS is a centralized information and technical support facility, required for coordinating and harmonizing practices of global programmes on a day to day basis, where operators officially register platforms and cruises' metadata to.

One of the goals of the JCOMM OCG and JCOMMOPS is to provide an accurate and homogeneous status of individual and heterogeneous observing networks. JCOMMOPS then faces the challenge of gathering, adapting and analyzing input metadata sets from various sources while being able to deliver unified and accurate output products using core and shared pillars, such as common vocabularies.

The Centre has the responsibility to provide all marine platforms with unique identifiers and to deliver to WMO all metadata on marine instruments through a WIGOS compliant format.

A new metadata and monitoring system

Therefore, JCOMMOPS is releasing a new version of its information system, facilitating access to the observing networks' metadata and enabling users to interact with the whole system through a brand new web experience and set of monitoring tools.

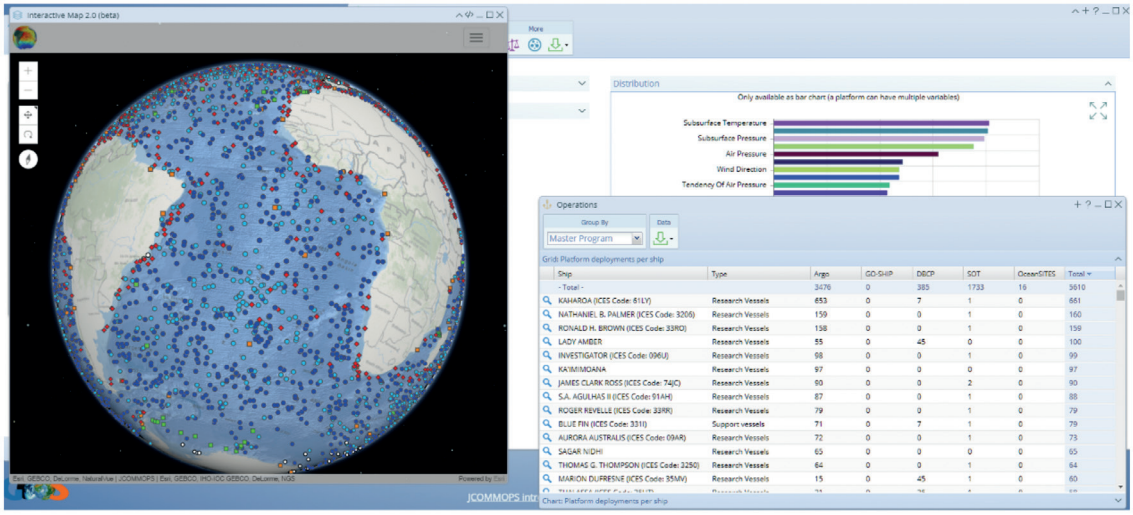


Fig. 2 - Web monitoring tools.

The interactive web-based monitoring application was designed to adapt to the wide spectrum of required perspectives and user needs. This web application is set in motion thanks to a background architecture focusing and relying on interoperable and open data access.

JCOMMOPS will provide a detailed review of its information system, including demonstration use cases of the web interface and available services, as well as future milestones and plans of continuous improvements.

While the metadata and monitoring golden standard is reached for a number of networks, further work is required within the implementing panels and data teams in order to enable a truely integrated Global Ocean Observing System.

An open strategy to connect European gliders: data flow and beyond

EuroGOOS glider data management Task Team:
Victor Turpin, Centre National de la Recherche Scientifique (France), victor.turpin@locean-ipsl.upmc.fr
Justin Buck, British Oceanographic Data Center (United Kingdom), juck@bodc.ac.uk
Thierry Carval, Institut Français de Recherche pour l'Exploitation de la Mer (France), Thierry.carval@ifremer.fr
Miguel Charcos Llorens, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mcharcos@socib.es
Patrick Gorring, EuroGOOS AISBL (Belgium), Patrick.Gorringe@smhi.se
Terry Hannant, Universitetet i Bergen (Norway), terry.hannant@imr.no
Daniel Hayes, OceanGliders (Cyprus), dhayes@ucy.ac.cy
Mark Hebden, British Oceanographic Data Center (United Kingdom), mahe@bodc.ac.uk
Emma Heslop, UNESCO-IOC Global Ocean Observing System (France), e.heslop@unesco.org
Giuseppe Manzella, ETT Solutions Ltd (Italy), giuseppe.manzella@ettsolutions.com
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com

Since completion of the EGO COST Action and the GROOM FP7 project three years ago, many improvements have been made to real time glider data management (new format, new tools, better management of the metadata, new platforms) but the community still struggles to reach a full European harmonization of glider data management with a complete contribution to research, ocean monitoring and operational services.

Moreover, delayed mode data management is becoming a priority for the glider community. Many operators and PIs are making substantial efforts in the qualification of their data sets. Several existing data management tools are freely available through toolboxes and scientific publications. Unfortunately, there is not yet a clear consensus on how these datasets should be shared in delayed mode nor a common strategy to handle these questions at the community level.

The EuroGOOS glider data management Task Team is establishing an open strategy, based on the EGO format and aligned with the objectives of OceanGliders, an associated program of the GOOS, to guide and encourage the glider teams at Operator, Principle Investigator and DAC levels who are willing to contribute to the global effort of ocean observing.

The data management strategy of the EuroGOOS Glider Task Team includes international coordination and harmonisation of methodologies and data. This strategy will allow a better contribution to the Global Ocean Observing Systems and the EuroGOOS Regional Ocean Observing Systems (i.e. Arctic-ROOS, MONGOOS, NOOS, IBI-ROOS, BOOS, Black Sea). Furthermore, it will allow the Task Team to form a realistic implementation plan for the smooth integration of glider data into the GOOS.

This presentation will certainly address the technical and strategic issues related to the EGO (European and abroad) format and real time/delayed mode - physical/biogeochemical glider data

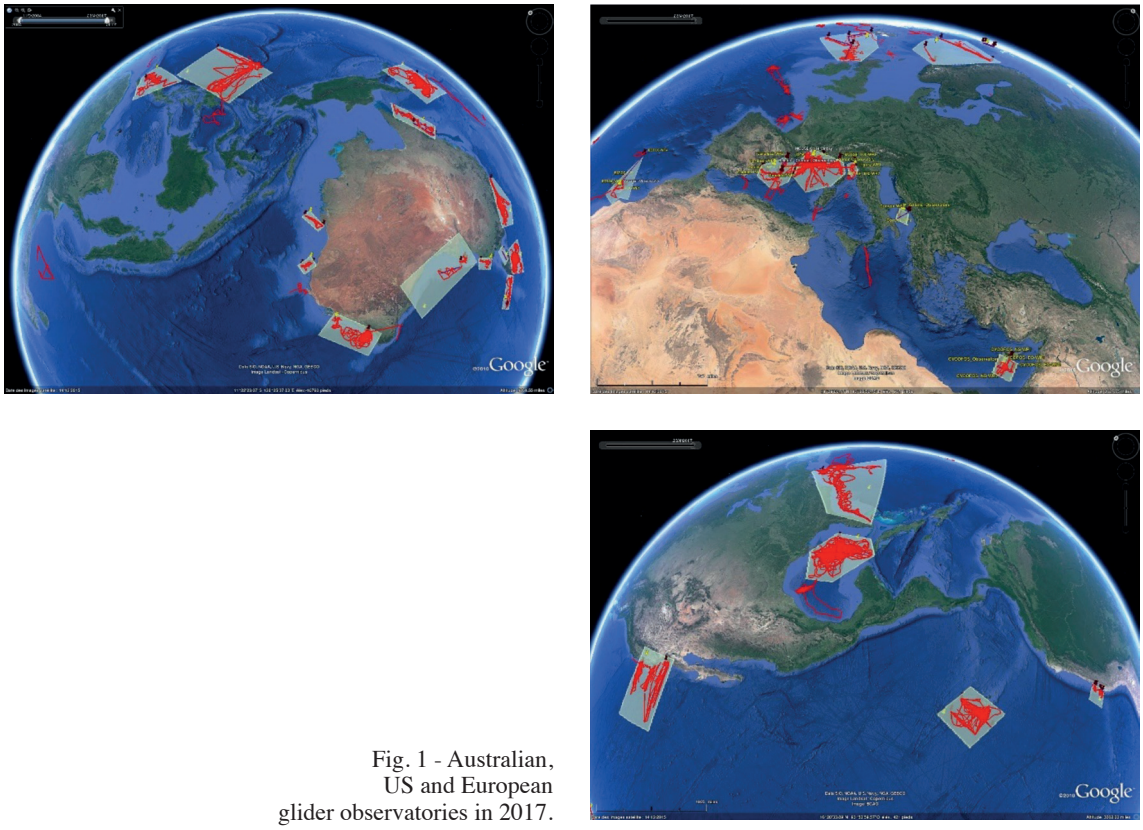


Fig. 1 - Australian, US and European glider observatories in 2017.

management and will address compatibility differences with IOOS (US) and IMOS (Australian) glider data formats in a global context. The presentation will also review the procedures and the different tools available to contribute to the international efforts of ocean data flow and data sharing and the role of DAC and G-DAC in RT and DM data flow. The vision is an international collaborative system able to provide open access to data and information as well as guidance on glider data management, with the goal to enable reproducible research, and to advance research and innovation. This effort will be shared with other initiatives such in IOC-IODE, GOOS/EuroGOOS, DGMARE - EMODnet Physics, Marine Copernicus.

Towards a Regional Database and Estimation System for fisheries

David Currie, Marine Institute (Ireland) David.Currie@Marine.ie
Laurent Dubroca, Institut Français de Recherche pour l'Exploitation de la Mer (France), laurent.dubroca@ifremer.fr
Edvin Fuglebakk, Institute of Marine Research (Norway), edvin.fuglebakk@hi.no
Kirsten Birch Håkansson, National Institute of Aquatic Resources (Denmark), kih@aqua.dtu.dk
Henrik Kjems-Nielsen, International Council for the Exploration of the Sea (Denmark), henrikkn@ices.dk
Twan Leijzer, Wageningen Marine Research (The Netherlands), twan.leijzer@wur.nl
Nuno Prista, Swedish University of Agricultural Sciences (Sweden), nuno.prista@slu.se

From the RDB to RDBES

There is currently a Regional Database (RDB) which is used to store detailed commercial fisheries sampling data. This is a regionally coordinated database platform and covers fisheries in the North Atlantic Ocean, the North Sea and the Baltic Sea. It addresses fishery management needs related to the European Union Common Fisheries Policy. However, it has been recognised for many years that there is a need to have a new version of the RDB which would also store details about how the sampling was performed and enable statistical estimations to be made. This new version is known as the Regional Database and Estimation System, RDBES.

The aims of the RDBES are:

1) To ensure that data can be made available for the coordination of regional fisheries data sampling plans, in particular for the EU DC-MAP Regional Coordination Groups (RCGs),
2) To provide a regional estimation system such that statistical estimates of quantities of interest can be produced from sample data in order to deliver data for ICES stock assessments and advice,
3) To increase the data quality, documentation of data and ensuring of approved estimation methods are used,
4) To serve and facilitate the production of fisheries management advice and status reports,
5) To increase the awareness of fisheries data collected by the users of the RDBES and the overall usage of these data.

It is expected that the new RDBES will replace both the current ICES RDB and InterCatch systems and provide a single platform for countries to produce statistical estimates of quantities of interest (such as discards and age distributions) which will then be used as inputs for fisheries stock assessment working groups.

The RDBES is an evolution of the work already done in defining and using the current RDB and the COST data models and functions. The current RDB data model provides a common data structure to describe commercial sampling data at a disaggregated level, and commercial landings and effort data at an aggregated level. The significant difference in the new RDBES data model is that it provides a common structure to describe both the disaggregated sampling data and, most importantly, how it was sampled.

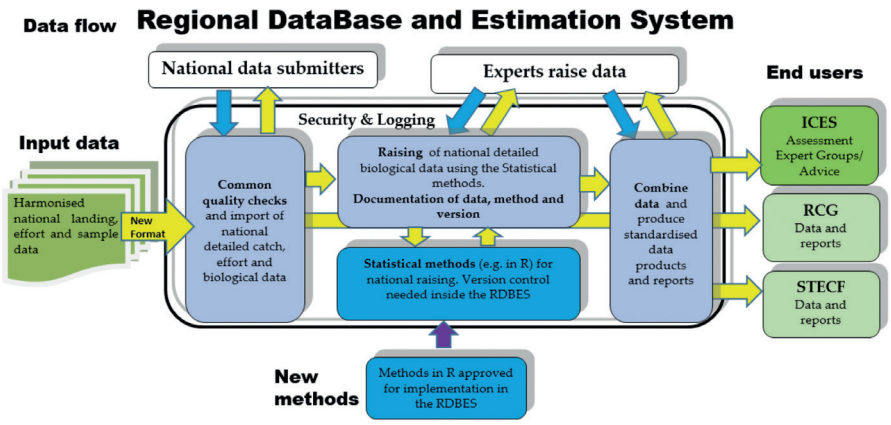


Fig. 1 - Diagram of the data flow and processes including the estimation using R.

The RDBES data model allows a variety of different estimation techniques to be used including the COST functions that make use of Age-Length keys, ratio estimators, and unbiased design based estimations.

The RDBES data model should be seen as part of the movements towards:

- Statistically Sound Sampling Schemes (4S),
- Greater regional coordination,
- Transparent Assessment Framework (TAF),
- Improved estimates to ICES stock assessments and advice.

Whilst the RDBES data model is designed to hold 4S data it will also be able to store data that is not sampled in a statistical manner – this is important so that historical data can be uploaded and stored. However, in doing so, the new RDBES data model flags those data thus allowing their appropriate interpretation during estimation.

It is important that the RDBES uses only approved estimation methods and it is transparent regarding the processing and estimation of data. This is in the interest of the countries, ICES and the European Commission. The countries will benefit from having one system where they can do the estimations. It is known that it is time consuming to do the complex estimations and there is risk of introducing errors, when estimating data for the ICES data calls for stock assessment and advice. The RDBES will help ensure higher data quality and that the users ensure the data is ready for estimation and select the right statistical estimation method. The countries will also benefit from the common repository of all the countries' developed estimation methods, this should also reduce the work load of reinventing estimation methods in the countries in parallel.

The hosting and maintenance of the existing RDB is funded through an Administrative Agreement between the European Commission (EC) and ICES. ICES have provided funding for the development of the RDBES.

References

Regional Database: http://www.ices.dk/marine-data/data-portals/Pages/RDB-FishFrame.aspx

ICES. 2018. *Report of the Workshop on new data model for the Regional Database* (WKRDB-MODEL), 15 – 18 January 2018, ICES HQ, Copenhagen, Denmark. ICES CM 2018/ACOM:41, 44 pp.

MARBEC-Obs: towards a virtual observatory of marine and coastal ecosystems, mainly in Mediterranean and tropical areas

Régis Hocdé, IRD MARine Biodiversity, Exploitation and Conservation (France), regis.hocde@ird.fr
Valérie Derolez, Institut Français de Recherche pour l'Exploitation de la Mer (France), Valerie.Derolez@ifremer.fr
Angélique Jadaud, Institut Français de Recherche pour l'Exploitation de la Mer (France), Angelique.Jadaud@ifremer.fr

The observatories department of MARBEC

MARine Biodiversity, Exploitation and Conservation (MARBEC), is the largest French research unit in marine biodiversity. MARBEC is a joint research unit between IRD, IFREMER, CNRS and University of Montpellier, and it has many international partnerships. Its objectives are the study of marine biodiversity in lagoon, coastal and offshore ecosystems, at different integration levels, mainly in Mediterranean and tropical areas: Indian Ocean, South, West and South-West Pacific, in Asia, Africa and South America. 45 people from MARBEC are involved in the 20 observatories or observation networks led by or involving MARBEC.

www.umr-marbec.fr/en/zones/observatoires,606.html?lang=en.

Fisheries and biological observation services

In ecology, the observation-research interaction stands first in a chronological framework in which the observation precedes any research process, since data production phases coming from the observation of the environment and the measures of conditions 'change, is a prerequisite for processing and analysis.

MARBEC-Obs is in charge of various durable observation missions, contributing to research and expertise on the state of coastal and marine ecosystems, marine biodiversity and the impact of human activities exploiting this biodiversity.

We provide ocean and costal variables from acquisition by the observation networks, i.e. :

- MEDITS (MEDIterranean Trawl Survey), DCMAP-EU: Halieutic, Hydrology, Bony fish, elasmobranchs, cephalopods, crustaceans, contributes also to Marine Strategy Framework Directive, www.sibm.it/SITO%20MEDITS/principaleprogramme.htm
- Observatory of Exploited Tropical Pelagic Ecosystems Obs7: Biological Information Collection, Fisheries Informations, Tuna Fisheries, www.ob7.ird.fr
- ReefTemps (network of temperature, pressure and salinity sensors in the coastal area of the South, West and South-West Pacific): Temperature, Conductivity, Salinity, Pressure, Waves, pH, Acidity, Fluorescence, www.observatoire-gops.org/fr/reeftemps1
- DCE-LAG (EU Water Framework Directive - French Mediterranean Lagoons): monitoring of Mediterranean lagoons, ecological and chemical status (phytoplankton, nutrients,

macrophytes, chemical contaminants), www.ifremer.fr/surval2/consultation.jsp?produit=resultats_par_parametre&carte=Resultats_par_parametre&progCds=RSLHYD#

Observation data management and access

MARBEC-Obs provide access to data and products. Several web data portals - currently for every observation network or part of an ocean - provide a combined array of services and functionalities (metadata exchange, visualisation of data, data acces, transformation of data…). Different types of interoperable services are offered, each tailored to a specific scientific user community (including the use of SOS (Sensor Observation Service), NetCDF services and others). It manages a national fisheries information system (Obs7).

An additional quality control is made. All the steps, with the corrected values, are stored in the information systems. Several observation networks started a national or european labeling process, guaranteeing the quality of the measurements and the traceability of observations and analyzes.

The observation data also feeds the regional and international data centers and web data portals: SeaDataNet, Seanoe/ ODATIS data portal, IMOS, GBIF, EMODnet in a soon future for macrolitters in Mediterranean (MEDITS).

Towards a virtual observatory of marine and coastal ecosystems

The services and functionalities offered by the information systems combine modeling, statistical analysis, data management and data visualisation. We also target to share and to interconnect some datasets, to optimize the exploitation of already acquired data by facilitating access, to allow comparisons between data of different origin and nature (observation data versus model results).

The results, including indicators and innovative products, are designed for Institutional stakeholders, Government Services, EU Water and Marine Strategy Framework Directives and scientific communities. MARBEC-Obs also has developped a Tuna Fisheries Expertise and a wider expertise on marine and coastal ecosystems mainly in Mediterranean and tropical areas.

Future plans. MARBEC-Obs has now different components: from sensors, to data dissemination and interoperability, to data processing combining modeling. We can evolve towards a virtual observatory of marine and coastal ecosystems, mainly in Mediterranean and tropical areas.

DASSH, the archive for marine species and habitats data - international standards for national data management excellence

Anna Luff, The Marine Biological Association (United Kingdom), annluf@mba.ac.uk
Dan Lear, The Marine Biological Association (United Kingdom), dble@mba.ac.uk
Kevin Paxman, The Marine Biological Association (United Kingdom), kevpax@mba.ac.uk
Matthew Arnold, The Marine Biological Association (United Kingdom), matarn@mba.ac.uk

The Archive for Marine Species and Habitats Data (DASSH) is the UK Data Archive Centre for marine biodiversity data, core funded by the UK Department for the Environment, Fisheries and Rural Affairs (DEFRA) and the Scottish Government and hosted by the Marine Biological Association of the United Kingdom. DASSH works in collaboration with, and is accredited by, the Marine and Environmental Data and Information Network (MEDIN), a network of UK organisations committed to sharing and improving access to marine data and information.

On an international scale DASSH is also the UK node of the Ocean Biogeographic Information System (OBIS), and an Associated Data Unit of the International Oceanographic Data and Information Exchange (IODE), giving the Data Archive Centre global recognition.

DASSH is a national and international hub for accessing, safeguarding, managing and disseminating historical, current and future datasets, DASSH also acts as a digital repository for biodiversity images and video.

DASSH provides online, open-source based tools for the description, discovery and download of biodiversity data, including the DASSH data portal, and Global Biodiversity Information Facility Integrated Publishing Toolkit (IPT). Through these tools, data can be downloaded in standard formats for integration into desktop and web GIS and statistical packages including R-based applications and PRIMER.

DASSH also operates a national data and metadata helpdesk and runs workshops to build data management capacity and capability in the UK marine sector.

Within DASSH and the other MEDIN Data Archive Centres, metadata is authored in accordance with the UK GEMINI discovery metadata standard, derived from the INSPIRE regulations, and the international ISO 19115 Geographic Information- metadata standards. The MEDIN community have worked to develop a marine implementation of the INSPIRE discovery metadata, using keywords, terms and definitions from the Vocabulary Service of the British Oceanographic Data Centre, as promoted through SeaDataNet.

As part of implementing the INSPIRE regulations, MEDIN publish data guidelines tailored to suit different methodologies and designed to implement the recording of standardised metadata alongside data, to increase the interoperability, and re-use of marine datasets.

DASSH are committed to applying the FAIR (Findable, Accessible, Interoperable and Reusable, Wilkinson et al, 2016) guiding principles in making marine biological data and associated metadata

open and freely available worldwide, to inform decision making, research, education and other public-benefit purposes, whilst contributing to the natural resource of environmental data. Data are made accessible through a range of data systems and aggregators, including the DASSH online portal, the UK National Biodiversity Network Atlas, EMODnet Biology, EurOBIS and GBIF. Metadata are also made available through the MEDIN portal, containing marine datasets from over 400 UK organisations.

Recent developments within DASSH have focussed on the automation of the archiving and publication of standardised datasets. Using the open-source GeoServer application as the engine for delivery, scripted workflows have been developed to address bottlenecks in data publication. Specifically we have developed an automated ingestion process combining datasets into a shared PostGreSQL database. The harmonisation of data systems drives a dynamic, web front-end for data discovery using geographic and taxonomic based searching.

DASSH have also developed a validation tool for MEDIN data which has vastly increased the accuracy of our QA procedure and decreased the human time required to check incoming datasets. Future developments include the creation of an online version of this tool, allowing users to validate their data before submission.

We have also been able to make progress in conversion between standards, for example from the Marine Recorder marine survey database to MEDIN Data Guideline and from MEDIN to Darwin Core. The latter is a vital component of our role as an OBIS node, allowing the publishing of datasets via IPT.

The presentation will focus on how the application of automated Quality Assurance and ingestion tools can streamline data flow at the national and global scale, and how the UK marine biological data infrastructure is contributing to the development of a global marine biodiversity observation network.

Unlocking European marine biodiversity under EMODnet Biology data using the FAIR principles

Simon Claus, Vlaams Instituut voor de Zee (Belgium), simon.claus@vliz.be
Christos Arvanitidis, Hellenic Centre for Marine Research (Greece), arvanitidis@her.hcmr.gr
Nicoals Bailly, Hellenic Centre for Marine Research (Greece), nbailly@hcmr.gr
Klaas Deneudt, Vlaams Instituut voor de Zee (Belgium), Klaas.deneudt@vliz.be
Daphnis De Pooter, Vlaams Instituut voor de Zee (Belgium), daphnis.depooter@vliz.be
Peter Herman, Deltares (The Netherlands), Peter.Herman@deltares.nl
Dan Lear, Marine Biological Association (United Kingdom), dble@MBA.ac.uk
Paula Oset, Vlaams Instituut voor de Zee (Belgium), paula.oset.garcia@vliz.be
Leen Vandepitte, Vlaams Instituut voor de Zee (Belgium), leen.vandeitte@vliz.be
EMODnet Biology Partnership (http://www.emodnet-biology.eu/partners)

Marine biodiversity data are essential to measure and study the ecosystem health of maritime basins. These data are often collected with limited spatial and temporal scope and are scattered over different organizations in small datasets for a specific species group or habitat. Therefore there is a continuous need to assemble these individual datasets, and process them into interoperable biological data products. The European Marine Observation and Data Network (EMODnet), supported by the EU's integrated maritime policy aims therefore to provide a single access point to European marine biodiversity data and products by assembling individual datasets from various sources and process them into interoperable data products. EMODnet Biology is a long term marine data initiative and is structuring its activities around four main pillars:

Making Marine Biological Data Findable

Detailed inventories of the recent and historical European biological datasets are created and published online via the EMODnet biology data catalogue. This catalogue is based on IMIS, an ISO-19115 compliant metadata system and contains over 1,200 dataset descriptions with information on the what, who, when, where and why a dataset was collected. These datasets include many large monitoring data collections from different European regional seas representing observations of marine species of phytoplankton, zooplankton, macro-algae, angiosperms, benthos, birds, mammals, reptiles or fish. For several of them a Digital Object Identifier (DOI) is equally available.

Making Marine Biological Data Accessible

Today, already 875 EMODnet Biology datasets, representing over 23 milion occurrence records are freely accessible through the EMODnet Biology data download portal. The portal includes 1. a toolbox allowing federated selection queries over different datasets, 2. a description on how the

Fig. 1 - EMODnet-Biology data portal at www.emodnet-biology.eu.

API can be accessed (as a WFS service) if a user wants to access machine readable marine biology data and 3. a link to the IPT resources, being the raw data available in Darwin Core Archive. The data are integrated into the European Ocean Biogeographic Informaiton System (EurOBIS) datasystem which contributes to OBIS, an global strategic alliance of people and organizations sharing a vision to make marine biogeographic data, from all over the world, freely available over the World Wide Web. There are now three subsections under data download.

Making Marine Biological Data Interoperable

All marine biological data is transformed to a common data structure and passes taxonomic (using the WORMS Vocabulary) and geographic (using the Marineregions Vocabulary) quality control procedures. The common data structure and standards used within EMODnet Biology consists of a DarwinCore (Dwc) Event Core in combination with a DwC Occurrence Extension and an enhancement to the DwC MeasurementOrFact Extension. This new structure enables the linkage of measurements or facts - quantitative and qualitative properties - to both sampling events and species occurrences, and includes additional fields for property standardization (using the NERC Vocabulary). The standard also allows to organize, aggregate, and link ocean observation events using “event hierarchy”.

Making Marine Biological Data Reusable

Through liaison with key stakeholders including governments, regulatory authorities, academia, NGOs and industry, it is clear that the real value lies in the the development of scalable, information-rich data products, based on high-quality underlying data. The products developed within EMODnet Biology are now being structured around the Essential Ocean Variables for biodiversity and will be provided through an Atlas of Marine Life. They inlcude 1. Trait based analysis, 2. Gridded species abundance maps, 3. Time series analysis and 4. Species distribution models.

Facilitating the publication of real-time marine observation data: the SeaDataCloud SWE Ingestion Service

Simon Jirka, 52°North GmbH (Germany), jirka@52north.org
Christian Autermann, 52°North GmbH (Germany), c.autermann@52north.org

Near real-time in-situ sensor measurements are of high importance for answering a broad range of scientific questions in the marine domain. At the same time, collected measurements are not only relevant for one specific question but they may be of high value in other research contexts, as well. Thus, there is a significant movement towards the development of research data infrastructures which shall facilitate the sharing of such scientific data sets.

Besides the availability of research data infrastructures, it is also important to consider interoperability aspects. This comprises especially the application of common standards for data access interfaces, data models, and encodings to ensure that data sets from different sources can be easily integrated. An important framework of standards addressing exactly this aspect is the Sensor Web Enablement (SWE) familiy of standards of the Open Geospatial Consortium (OGC).

In order to increase the availability of near-real time in-situ observation data for researchers, the SeaDataCloud project aims, among other aspects, at facilitating the data publication process. For this purpose, the SeaDataCloud SWE Ingestion Service is currently being developed as a tool to support the publication of near real-time in-situ observations in interoperable research data infrastructures. This component is intended to support sensor operators, researchers and data owners during the publication of collected marine observation data by offering a configurable and re-usable publication workflow.

Fig. 1 illustrates the proposed ingestion workflow: Central component is an OGC Sensor Observation Service (SOS) instance which is capable of storing observation data and metadata

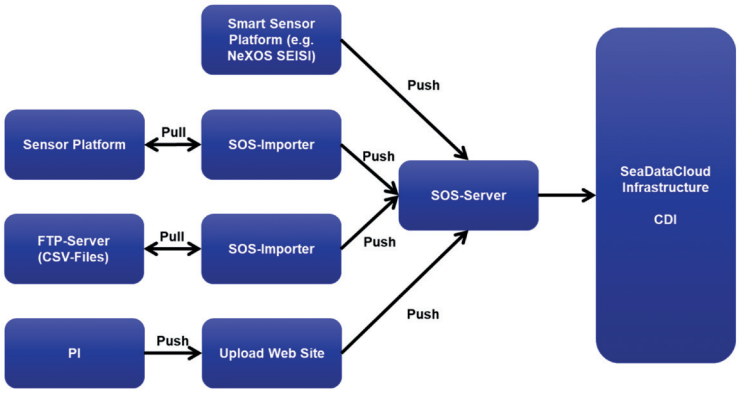


Fig. 1 - Overview of the Proposed Data Ingestion Workflow.

from different sources. Through dedicated adapters and importer components different sources can be linked to the central SOS server:

- Smart sensors (for example sensors supporting the Smart Electronic Interface for Sensor Interoperability (SEISI) developed by the NeXOS project): Such sensors are able to push their metadata and data autonomously to an SOS server.
- Sensor platforms with a well-described data access interface: Such sensor platforms can be regularly queried by an importer for new available data. In this case a formal description of the sensor platform interface is necessary (based on the OGC Sensor Model Language (SensorML) standard) so that an importer can be automatically configured to the commands and data formats of the sensor platform.
- Data archives such as files or servers containing CSV files: Automatic harvesting of data by an importer. In this case a description of the data structure is necessary so that the importer can be configured to the specific format in which the data is provided.
- Uploading (e.g. by researchers) so that the data is automatically processed. Also in this case a description of the data structure is necessary so that the importer can be configured to the specific format in which the data is provided.
- Internet of Things data sources (e.g. MQTT brokers)

The harvested metadata and data can subsequently be coupled with the SeaDataCloud infrastructure. An important element of this data flow is the provision of the collected information through the SeaDataCloud Common Data Index (CDI) so that researchers are able to discover and access the available data sets. In this case, the SOS server will act as a feeder for the CDI through interfaces offered by the CDI. Furthermore, the publication of the data via SOS servers allows the visualisation of the previously ingested data in tools such as the Helgoland Sensor Web Viewer (Fig. 2).

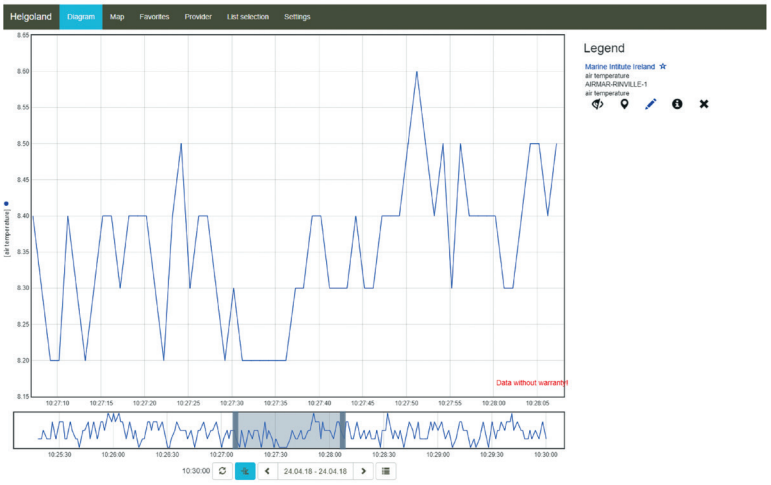


Fig. 2 - Published Observation Data Visualised in the Helgoland Sensor Web Viewer.

In summary, the presented SeaDataCloud SWE Ingestion Service is a powerful tool to publish marine observation data from different sources on interoperable research data infrastructures. This is complemented by the SOS interface for facilitating data distribution and access, and integration into the SeaDataCloud CDI as well as a broad range of data analysis and viewing tools such as the Helgoland Sensor Web viewer.

Pro-active diving. Diving with a purpose

Alexander Smirnov, Arctic Portal (Iceland), alexander@arcticportal.org
Halldor Johannsson, Arctic Portal (Iceland), halldor@arcticportal.org
Erlendur Bogason, Strytan Dive Center (Iceland), erlendur@eyjar.is
Marlies Wolters, Dive O'clock (The Netherlands), marlies@proactivediving.com
Dominik Wyszogrodzki, The Scandinavian Dive Site, (Poland/UK), dom@proactivediving.com
Sune Dadda Karma Harms, The Scandinavian Dive Site, (Denmark), sune@proactivediving.com

Overview

Pro-Active Diving project and Sea Mojo Tool will link people from around the world to become part of the Pro-Active Diving community. This community will combine the diving public's love of the ocean, curiosity about nature, and desire to contribute to marine knowledge. Through a portable underwater measuring device with sensors (designed to be worn by divers) and software package which together form the Sea Mojo Tool, Pro-Active members will monitor marine life as well as physical water conditions, and this data will be merged into an open access database making it accessible to the research community, interested members, and the public in general. Pro-Active Diving and the Sea Mojo Tool are designed to link the sport diving community, dive operators, dive certification organizations, tourism related businesses, educational institutions, and the ocean research community. The program will turn divers into citizen scientists with the training and tools to contribute valuable data to marine scientists and oceanographers who are working to understand how climate change and other factors are affecting the world's oceans.

Project objectives

This project will focus on continuing the development of a device (the Sea Mojo) for collecting physical data from the water while SCUBA, snorkel and freediving enthusiasts are diving recreationally (with the Sea Mojo device), along with a mobile application which will sync the data from the Sea Mojo to their mobile device (phone or tablet), and upload it to an online database (Fig. 1). The project will also continue the development of the database which will be made available to ocean scientists and any other interested parties.

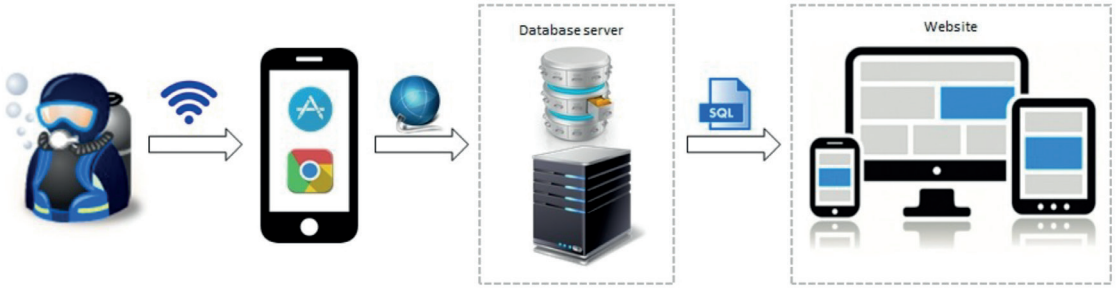


Fig. 1 - Pro-Active Diving technological scheme.

Current status of knoledge

This project will combine three established concepts that have not been integrated before: SCUBA and freediving ecotourism, Citizen science and Portable (wearable) underwater measuring device with sensors (Sea Mojo) with Integrated Upload to an Online Accessible Database.

Despite the potential of SCUBA dive computer temperature data to add to existing temperature data sources, very few attempts have yet been made to compile and assess the quality of this data. Conducted research showed that the absolute temperature differences between diving computers and CTDs at the start of the dive in some cases exceeded 2°C. Such accuracy is unacceptable for scientists to assess regional changes in physical oceanography, or to monitor climate change, assess the response of organisms to environmental change, or validate remote sensed data from satellites in coastal regions.

However, scientific sensor technology has also advanced and relatively small, accurate and affordable sensors can record data such as water temperature, salinity, turbidity (the amount of light reaching a certain depth), and other parameters although less affordable for the intentions of Sea Mojo.

These small sensors can be combined with enough memory to record a large amount of data, and a transmitter (WiFi or Bluetooth), in a package small enough to be easily wearable. Fig. 1 demonstrates a principal scheme of data transmission between the device and a web interface suitable for PCs, tablets and smartphones. The device activates automatically as soon as it gets into water recording current geographical position (GPS), date and time. Such information is crucial for oceanographic station identification. During the dive the device continuously records temperature, salinity and depth with a predefined frequency. After ascending the device records the same metadata parameters at the exit location. Due to a lack of internet connection at the most dive spots the device stores oceanographic data until it connects to a smartphone with the app. When such connectivity is established, the app downloads all the row data from the device and automatically transmits it to the database server. The database server receives the row data, performs required conversions, applies quality and duplicate control procedures and uploads ‘clean’ data into the database which can then make the information accessible to anyone in the world. PostgreSQL open source object-relational database system will be used as DMS (Data Management System). It should be noted, that every dive will give the community at least two oceanographic stations. According to the definition, an oceanographic station is a geographic location at which oceanographic observations are taken. Hence the first station will be taken during descent and the other one will be taken while the driver is ascending.

The Sea Mojo Tool is designed to be expandable, in other words, new sensors to measure different physical parameters can be added or traded out of the device. This means that some sensor elements will be needed for specific dive locations and can be purchased by the user. In addition, as new technology is developed, sensors will be refined and made better, also, sensors able to measure new parameters that have not been available for the Sea Mojo Tool will be released.

The combination of the above ideas is what make Pro-Active Diving and the Sea Mojo tool unique.

Assessment of existing services and new services provided by the Copernicus Marine In Situ Thematic Assembly Centre (INSTAC)

Petit de la Villéon Loïc, Institut Français de Recherche pour l'Exploitation de la Mer (France), Loic.Petit.De.La.Villeon@ifremer.fr

Pouliquen Sylvie, Institut Français de Recherche pour l'Exploitation de la Mer (France), Sylvie.Pouliquen@ifremer.fr

And the Copernicus Marine Environment Monitoring Service In Situ Thematic Centre partners

Context

The Copernicus Marine Environment Monitoring Service (CMEMS) is one of the six operational services of the European Copernicus programme. It is designed to respond to issues emerging in the environmental, business and scientific sectors. The role of the CMEMS Thematic Assembly Centres (TACs) is to collect, process and quality control upstream satellite and in-situ data required both to constrain and validate modelling and data assimilation systems and to directly serve downstream applications and services.

Within CMEMS, the In Situ Thematic Assembly Centre (INSTAC) ensures that a steady supply of in situ ocean measurements is made available to the other service components and to other external intermediate and final users

The In Situ TAC provides vertical profiles and time series data coming from different types of instrument (floats, drifters, moorings, gliders, vessels…) and different physical parameters (temperature, salinity, currents, sea level, …). In the frame of the first phase of CMEMS (May 2015-April 2018), the In Situ TAC has consolidated the Near Real Time products and added to the catalogue reprocessed historical products.

Existing services

During phase 1 (2015-2018) the project set up the operational service and provided data related to the following areas:

- In Near real time
  - Temperature and salinity at both regional and global levels
  - Wave at both regional and global levels
- In delayed mode (reprocessed products)
  - A merged product (1950-2017) between the V1 CMEMS product and ENACT4 product managed by MetOFFICE.
  - A surface current product, designed for reanalysis, that integrates the best available version of in situ data for Ocean surface currents for the period 1990-2017.
  - A wave product that integrates quality-controlled wave data, both near-real time and historical, collected from more than 400 platforms around the globe.
  - A first version of an historical BGC (O2 and Chla) product

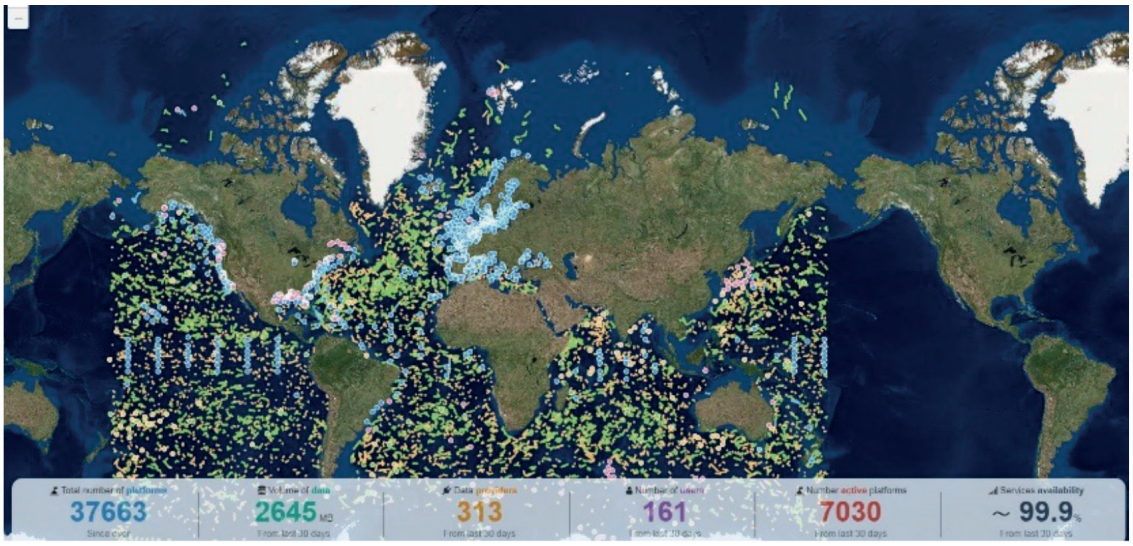


Fig. 1 - Dashboard of the INSTAC service http://www.marineinsitu.eu 30 days of data coverage.

The INSTAC is a distributed centre with a global component closely link the JCOMM networks and 6 regional components developed in partnership with the EuroGOOS ROOSes (Regional Operational Oceanographic Systems).

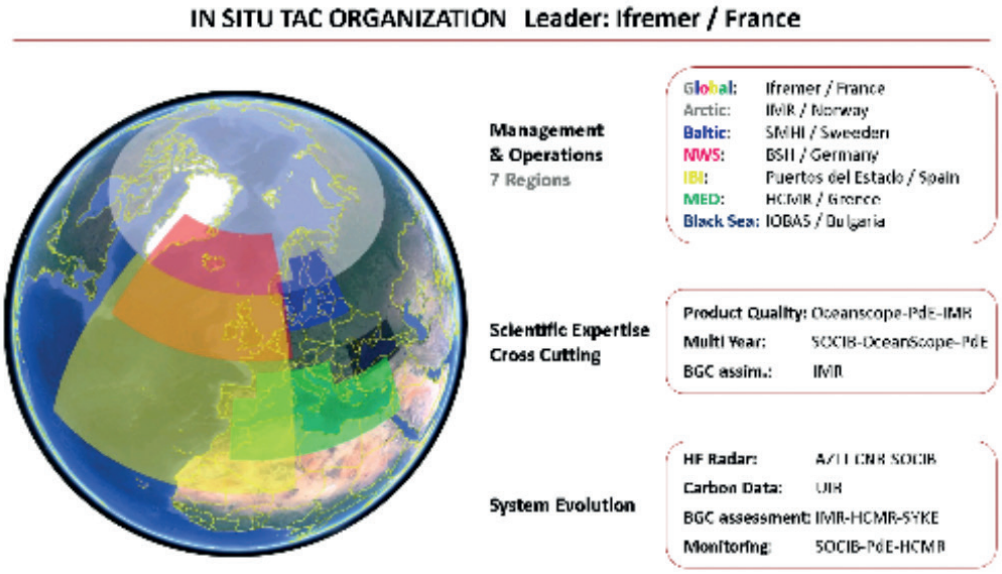


Fig. 2 - INSTAC organisation.

Moreover, the CMEMS INSTAC is one key element of the in situ data management in Europe closely linked to the major European data integration initiatives such as EMODnet Physics, EMODnet Chemistry and SeaDataNet

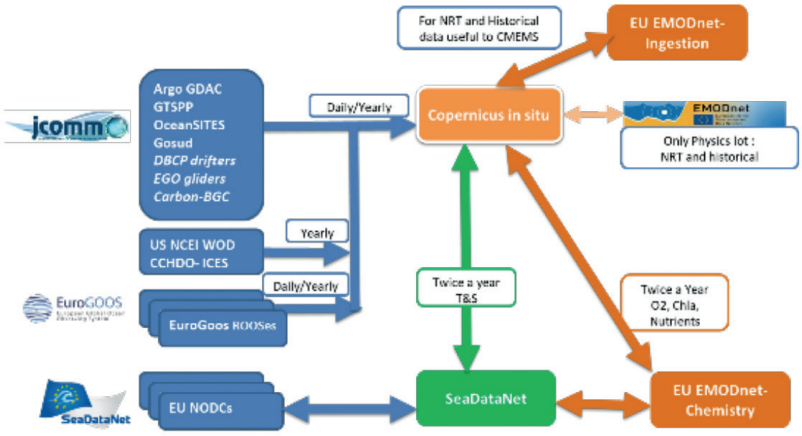


Fig. 3 - CMEMS INSTAC Interfaces with other data systems.

New services planned for the CMEMS INSTAC phase 2 (2018-2021)

In the second phase on CMEMS INSTAC (2018-2021) will extend its activities to the distribution of HF-Radar, the integration of carbon products in link with ICOS and to the extension in time and space coverage of the existing products.

Integrated services such as CMEMS INSTAC products, facilitate and extend the use of existing in -situ observation by a wider community and also help highlighting the existing gaps in the observing systems within an integrated multiplatform design.

The proposed presentation will assess 3 years of IN SITU activity and will detail the ongoing work planned for phase 2

POSTERS

Integrating diverse data into the coastal observation system COSYNA and EMODnet

Gisbert Breitbach, Helmholtz-Zentrum Geesthacht (Germany), gisbert.breitbach@hzg.de
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com

Modern marine observation systems consist of very diverse observational data e.g. data from remote sensing platforms like HF-radar, data from moving platforms like FerryBoxes on ships or data from stationary platforms like underwater nodes. In addition results from model calculations are also part of the system to be able to fill the gaps in observations and get a more complete view of the system.

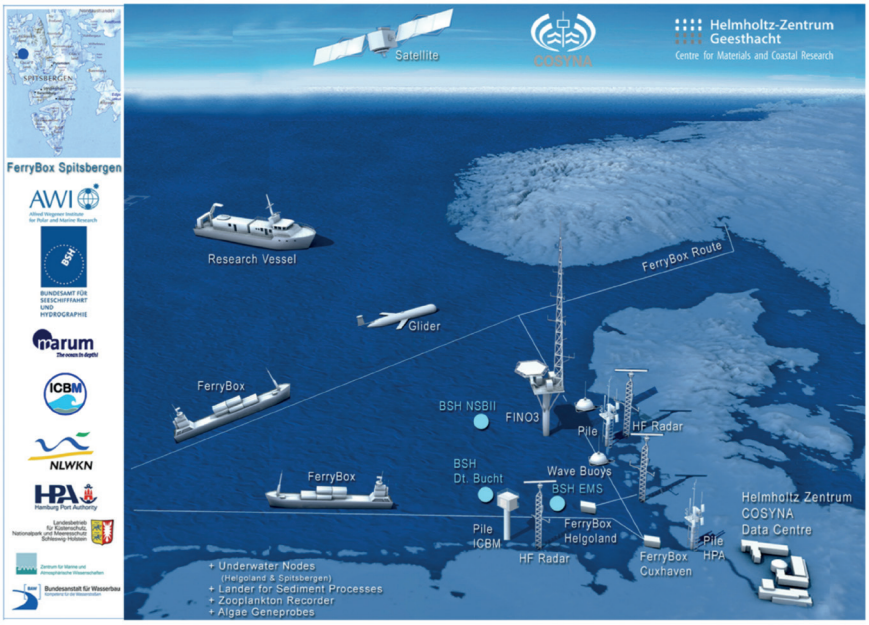


Fig. 1 - Data from diverse sources are integrated into COSYNA.

The COSYNA approach is metadata and web-service based to achieve an integrated view on the data.

- All parameter names are mapped to the corresponding CF standard names to enable a comparison between data from different platforms.
- Data visualisation and download is provided via URL based web-services.
- Data-metadata includes dynamic URLs of web-services.
- A single metadata record exists for time-series at each fixed location. All other metadata are created automatically.
- For a simple access Web Feature Services are constructed from the metadata.

Integrating COSYNA data into EMODnet

Within the EMODnet Physics portal (www.emodnet-physics.eu), one of the European Marine Observation and Data network thematic portals, a combined array of services and functionalities is developed such as capabilities for viewing and downloading, reporting etc. and is currently providing easy access to data and products for the following quantities: wave height and period; temperature and salinity of the water column; wind speed and direction; horizontal velocity of the water column; light attenuation; sea ice coverage and sea level trends (relative and absolute). EMODnet Physics is continuously increasing the number and type of platforms in the system by linking to and providing high quality data from a growing network of providers.

The acquisition of physical parameters is largely an automated process that allows the dissemination of near real time information. The infrastructure for storing and distributing this data is based on the Copernicus Marine Environment Monitoring Service In-Situ Thematic Assembly Center. Both CMEMS and EMODnet are acting as integrators: they do not actually measure themselves, they rather rely on regional systems such as COSYNA for providing the backbone infrastructure for their success.

Moreover where regional systems are exploiting the latest interoperability services (SOS, WFS, WCS etc.) such as COSYNA, the real time integration is faster, more robust and long term sustainable and any new dataset or parameter that the COSYNA system is able to record, is immediately supplied to the European ocean data information infrastructures. The acquisition of physical parameters in COSYNA is an automated process that allows the dissemination of near real time information. These data are then included in the regional product distributed by the CMEMS INSTAC and hence the EMODnet Physics.

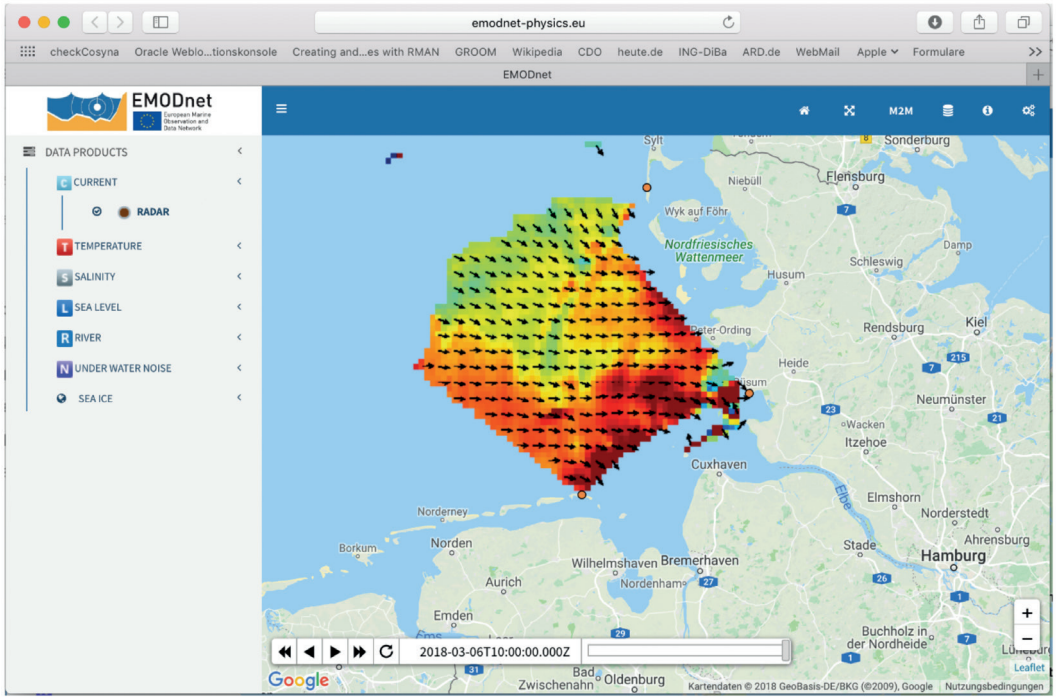


Fig. 2 - Example for the integration of COSYNA HF radar data into EMODnet physics.

SeaDataCloud: Tunisian marine data management

Sana Ben Ismail, Institut National des Sciences et Technologies de la Mer (Tunisia), sana.benismail@instm.rnrt.tn

Nouha Barraj, Institut National des Sciences et Technologies de la Mer (Tunisia), nouha.barraj@instm.rnrt.tn

Mohamed Anis Ben Ismail, Institut National des Sciences et Technologies de la Mer (Tunisia), medanisbenismail@yahoo.fr

Cherif Sammari, Institut National des Sciences et Technologies de la Mer (Tunisia), cherif.sammari@instm.rnrt.tn

Access to oceanographic data is of vital importance for marine research. Data acquired from various observational activities and techniques have problems of heterogeneous data sources, accessibility and standardization. SeaDataNet, SeadataNet II and now SeaDataCloud projects have built a leading infrastructure for marine and ocean data management. It is a Pan-European infrastructure for managing, indexing and providing access to marine data and metadata in compliance with the INSPIRE directive (ISO 19139 standards). National Institute of Marine Sciences and Technologies INSTM and especially its Marine Environment Laboratory (LMM) is being involved in this project since 2011 and is regularly providing new data and metadata in the infrastructure. The majority of Tunisian marine data managed in the framework of SeaDataNet concerns physical oceanography especially Conductivity, Temperature and Deph (CTD) and bottle measurements.

Recently INSTM started to measure high frequency data using a FerryBox system installed on board of the Tunisian ship of opportunity Carthage (CTN, Compagnie Tunisiennes de Navigation). A Tunisian workflow was set up from data acquisition to distribution. The main steps are: error checks, MedAtlas or ODV format file creations, quality control and metadata generation for indexation in Seadatanet catalogues. Finally, Tunisian ocean and marine data of LMM are standardized (ISO19139) and interoperable enabling to improve national and international research.

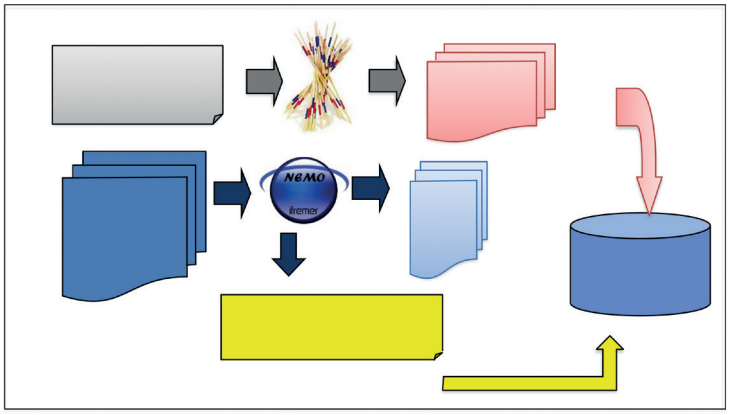


Fig. 1 - INSTM Methodology for setting up CDI, ODV and coupling table.

Oceanographic data in the EU legislation: a temporal evolution analysis

Enrique Wulff, Consejo Superior de Investigaciones Científicas (Spain), enrique.wulff@icman.csic.es
Ernest Abadal, Universitat de Barcelona (Spain), abadal@ub.es

Introduction

The management of oceanographic data (fisheries, marine, maritime, water and spatial data) is an increasingly relevant question for European environmental policy and management. To get familiar with the services and tools for data management (access, quality, reutilization), they must be based on clear legal grounds in terms of their access, use and practical support.

The Community Institutions (especially the European Commission) introduced over the past 40 years, a set of legal texts on the oceanographic data concerned with issues related to:

- the guarantees that these data can be used and re-used as widely as possible;
- the best way to interconnect the new and existing oceanographic data infrastructures;
- the coordination of already existing aids in favour of the data infrastructures.

But EU searching to legally commit governments to grant access to ocean data and allow their re-use, does not automatically apply to the large pools of data held by research institutions for the benefit of National Oceanographic Data Centers (NODC).

Objectives

This proposal attempts to analyse the existing information at the European legislation level on oceanographic data. That means that this will focus on examining the temporal distribution of the legal texts, the type of document, the organization that issued them, the type of data which is referred (fisheries, marine, maritime, water and spatial data) and the actions of data management that correspond to their retrieval and compilation, process, custody, distribution and re-use.

Methodology

A database has been compiled with all the valid and binding EU legislation referred to the European Seas and Oceans, available from the EU official gateway to European legislation, EUR-Lex, since 1976 to 2017. This database contains a total of 216 records. A content analysis was performed associated with each of the legal texts, by using a set of seven indicators: type of document, the organization that has approved them, content description, date and place of adoption, geographical scope, type of data, actions applied to these research data.

Results

The analysis unit identifies 20 different types of documents, pointing out that regulations (31%), communications (25%) and directives (14%) are the main sources for the study of the European legislation on oceanographic data; also important were the decisions (8%), the

international treaties (5%) and the laws (4%). Other legal documents (13%) complement the types of legislation. Concerning the body that promotes and approves the applicable legislation, the European Commission should be mentioned with the greater number of legal provisions issued (40%); above the European Parliament (23%) and the European Council (20%).

Oceanographic data management legislation in Europe, begins with the Barcelona Convention for the Protection of the Marine Environment and the Coastal Region of the Mediterranean, originally signed in 1976 by the European Economic Community. Inside the Convention a main activity is data compilation. The last document analyzed is a Report from the Commission to the European Parliament in October 2017. The temporal distribution within these limits shows two main peaks: 2006, when the EU outlined its Integrated Maritime Policy (IMP) and its European Marine Observation and Data Network (EMODnet) and 2008 when the Marine Strategy Framework Directive (MSFD) was promulgated.

According to their geographic framework of application, all the analyzed documents have the overall geographic aim to cover the coasts, seas and oceans of the 23 EU coastal Member States. In relationship to the typology of data managed, reference is made to fisheries (35%), maritime (13%), marine (17%), water (4%), spatial (7%) and other (24%) kind of data. Concerning the actions that can be taken to manage the data and which explicitly appear in the legal texts, data retrieval and compilation rose 64%, process 11% and, to a lesser extent, data custody 6%, distribution 4% and other tasks 15%.

ODIN2 - user-friendly, web-based access to more than 70 million oceanographic readings

Steffen Bock, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), steffen.bock@io-warnemuende.de

Susanne Feistel, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), susanne.feistel@io-warnemuende.de

The Oceanographic Database of IOW (IOWDB) had originally been designed for particular internal requirements of the Leibniz Institute for Baltic Sea Research (IOW). IOWDB has always been aimed at the management of historical and recent oceanographic measurements. Most recently, the research tool ODIN2 was published to provide those research data to the public in a user-friendly way.

Fig. 1 - QR code for https://odin2.io-warnemuende.de

The data stock possessed by IOW was declared as highly valuable by independent commissions. Research cruises have been conducted since 1949, and their data systematically collected. The post-processing and analysis of the resulting observational data was carried out by varying methods and with changing quality. A substantial fraction of legacy data was successively transferred to modern storage media, their quality was controlled and, if possible and necessary, improved by detailed individual scientific inspection.

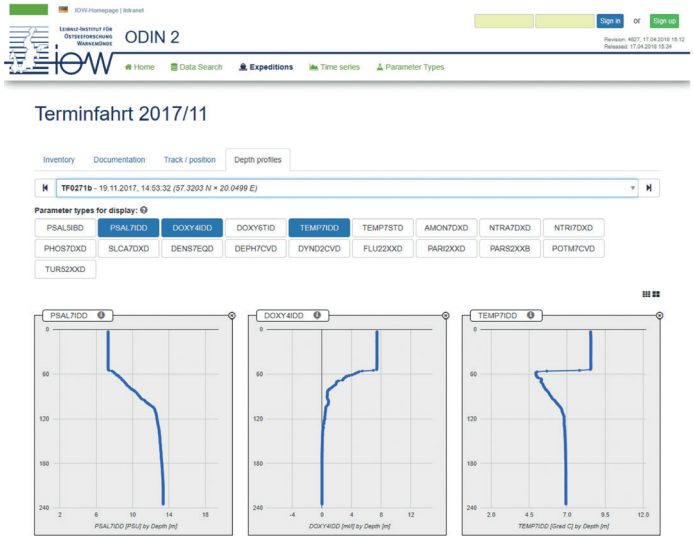


Fig. 2 - Depth profiles available in ODIN2: A single CTD cast at the Gotland Deep in the Baltic Sea, taken November 19, 2017, from surface (top) to 240 m depth (bottom). Profiles from left to right: practical salinity ranging between 7 and 14 PSU, dissolved oxygen concentration (8 to 0 ml/l) and temperature (8 to 4 °C). The stratification of the Baltic Sea is instantly visible.

The content of our database includes oceanographic readings and metadata (mainly Baltic Sea) from 1877 to 2018 obtained during 932 research campaigns of the IOW (the former Institut für Meereskunde, IfM) and cooperating institutions. As of April 2018, the IOWDB contains more than 70 million measured samples representing georeferenced point data from the water column, primarily from CTD profiles, hydrochemical and biological sampling, current-meter time series, trace metal sampling and long-term monitoring. Phyto- and zooplankton data are available for 1988 to 2017.

To access this data treasure the IOW released ODIN2 for the public (https://odin2.io-warnemuende.de). It is a web-based search application to access all of the data stored in the IOWDB. Instantly accessible are, for example, either ship tracks or station lists for each campaign as well as on-the-fly depth profiles of each CTD cast taken (Fig. 2). For the long-term monitoring stations in the Baltic Sea, ODIN2 offers automatic visualisation of up to 50-year time series for regularly sampled parameters at standard depths (Fig. 3).

To explore the research data in IOWDB, ODIN2's individual search engine is the most powerful and fastest tool. It offers complex search and filter options for geographic areas, time spans, seasons, depth range, parameters, cruises or standard positions. After executing a search, the results are integrated in user-friendly digital export formats such as plain text, xls or netCDF.

ODIN2, and along with it IOW's long-term monitoring data, have been made publicly available worldwide since April, 2018. For convenience, ODIN2 offers the creation of a URL as a permanent link to repeat or share any previously defined search.

All research data available via ODIN2 are licenced under CC BY 4.0.

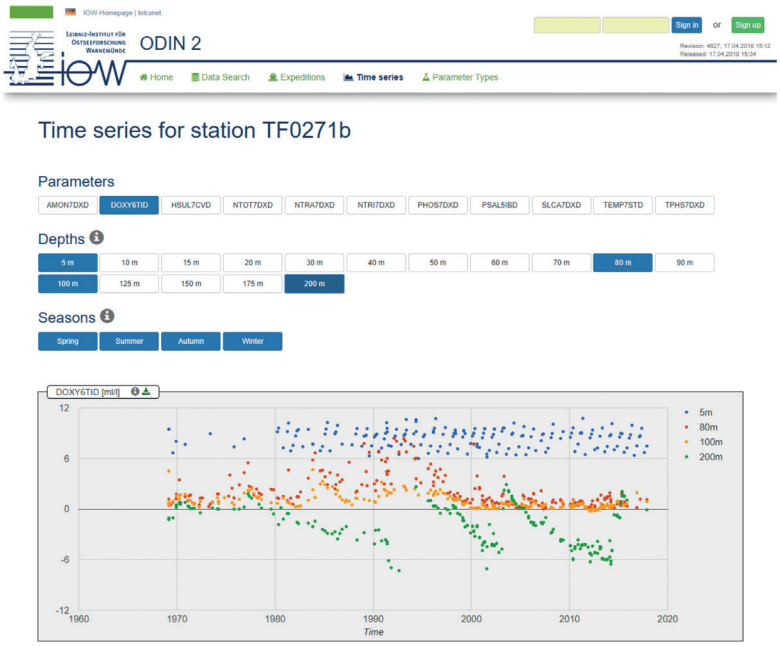


Fig. 3 - Visualisation of long-term measurements of dissolved oxygen and hydrogen sulphide at the Gotland Deep in the central Baltic Sea. The time series starts in 1969 and continues to 2017. Displayed are data from the surface (5 m, blue dots), from the halocline depth (80 m, red dots; 100 m, yellow dots) and from the near-bottom depth (200 m, green dots). Clearly discernible are the Major Baltic Inflow events in 1993, 2003 and 2014.

SOOSmap brings circumpolar Southern Ocean data to a computer near you

Pip Bricher, Southern Ocean Observing System (Australia), data@soos.au
Patrick Gorringe, Sveriges Meteorologiska och Hydrologiska Institut (Sweden), patrick.gorringe@smhi.se
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com
Marco Alba, ETT Solutions Ltd (Italy), marco.alba@ettsolutions.com
Jie Zhang, Polar Research Institute of China (China), zhangjie@pric.org.cn
Roger Proctor, IMOS Australian Ocean Data Network (Australia), Roger.Proctor@utas.edu.au

Data discovery and accessibility are constant challenges for scientists, especially those working in inherently international disciplines, such as oceanography. The Southern Ocean Observing System (SOOS) and EMODnet Physics groups are collaborating to remove some of these challenges by developing SOOSmap. SOOSmap builds on the data aggregation and sharing infrastructure of EMODnet to bring circumpolar datasets into a single web-based discovery portal. Through SOOSmap, users can discover, plot, explore, and download datasets of relevance to biologists, ecologists, ice scientists, and physical oceanographers. They can also use it to identify key spatial and temporal gaps in the observing infrastructure of the Southern Ocean. Until now, EMODnet's focus has been on European waters, although it houses several global datasets. The collaboration with SOOS brings Southern Ocean-specific datasets into EMODnet's data-handling infrastructure. For SOOS, accessing the coding skills and data aggregating infrastructure of EMODnet allows it to develop the data-sharing tools it needs without duplicating existing infrastructure and without placing undue burden on its member organisations. In this presentation, we will share our lessons from this new collaboration, demonstrate the progress so far in sharing circumpolar datasets, and describe our future development plans.

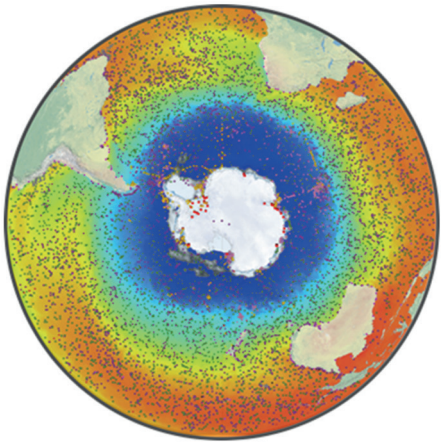


Fig. 1 - The SOOSmap interface, showing all observation points overlaid on interpolated sea surface temperature and sea-ice concentration layers.

Geoinformation system of the Russian Black Sea coastal zone

Elena Zhuk, FSBSI «Marine Hydrophysical Institute of RAS», (Russia), alenixx@gmail.com
Eugeny Godin, FSBSI «Marine Hydrophysical Institute of RAS», (Russia), godin_ea@mhi-ras.ru
Andrey Ingerov, FSBSI «Marine Hydrophysical Institute of RAS», (Russia), ingerov@rambler.ru
Elena Isaeva, FSBSI «Marine Hydrophysical Institute of RAS», (Russia), isaeva-ea@mail.ru

Introduction

Successful research and analysis of the current state of seas and oceans coastal areas require reliable information support. The most convenient instrument to solve this problem is GIS. At present, a GIS of the Russian Black Sea coastal zone is being developed in FSBSI "Marine Hydrophysical Institute of RAS" (MHI). While designing the GIS, particular attention was paid both to the software and to the databases used.

Coastal zone data access software

The Black Sea GIS software developed in MHI [1,2] was taken as a basis to provide data access via Internet. It was developed using the client server architecture, MapServer was used as a map service, and MySQL was applied as a database management system. This software enables selecting and visualizing tabular data kept in relational databases, as well as textual and graphical information. The Black Sea GIS module structure gives an opportunity of further extending the specialized database.

Databases

The specialized database for information support of the coastal research consists of two main blocks. One of them includes data of the oceanographic research in the costal area; the other one contains data obtained during the coastal research and remote observation for the coastal zone (Fig. 1).

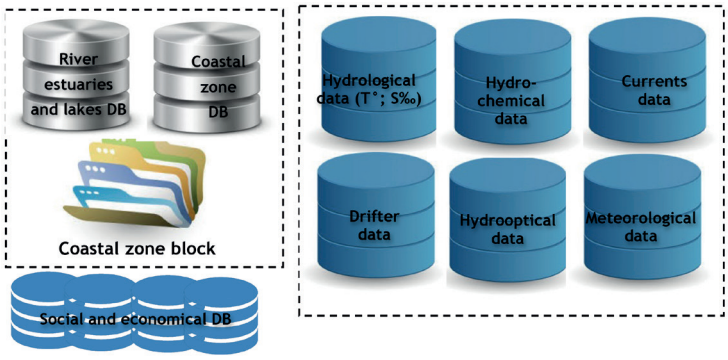


Fig. 1 - The structure of the specialized database for information support of coastal research.

The oceanographic block is constructed using the Black Sea specialized database of Oceanographic data bank of Marine Hydrophysical Institute RAS [3]. It includes oceanographic and meteorological data. The Specialized database includes observations along the Crimean and Caucasian coasts of the Black sea. The database contains over 55,000 oceanographic and over 10,500 hydrochemical stations.

The Coastal zone block consists of three databases:

- the Database for the coastal zone study (the DB includes granulometric and mineral composition of sediments; morphological characteristics of the submerge slope, etc.);
- the Database for the river estuaries and coastal lakes;
- the Database of aerial and satellite images (the DB includes materials of aerial photography, satellite images, and photos and video from an intelligent camera installed on a drone [4]).

Data access examples

The developed software provides data access and visualisation of different kinds of coastal data which are storaged in several databases. The results of queries to the Database for the coastal zone study (photo of landscape for the selected profile) and the Database of aerial and satellite images (photo of Bakalskaya spit from the drone) are shown as an example in Fig. 2.

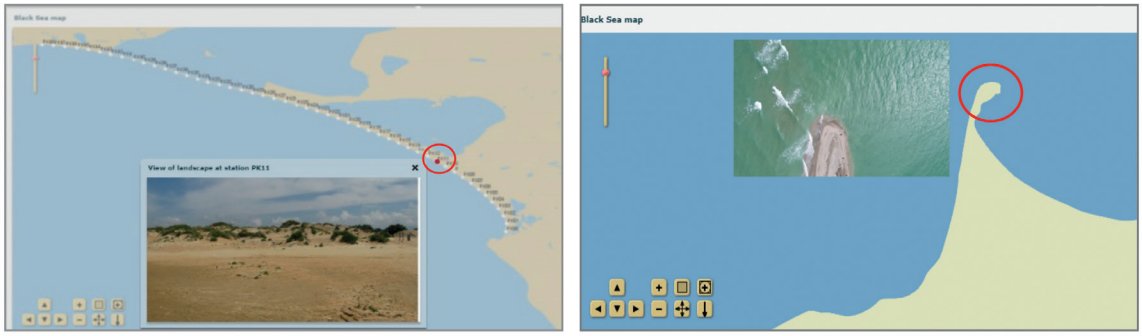


Fig. 2 - a - Photo of landscape for the profile; b - Photo of Bakalskaya spit from drone.

Conclusions

The current version provides a convenient access and presentation of the data obtained while studying the Russian Black Sea coastal zone. At present, the activities aimed at the further GIS development, related both to improving the software and connecting new databases, are ongoing.

Acknowledgements. The GIS is developed within the framework of the state task on theme No. 0827-2018-0004. The adaptation of the database for filling with data about Black sea coasts was financially supported by the Russian Science Foundation (RSF grant 14-17-00547).

References

[1] Zhuk E., Khaliulin A., Zodiatis G., Nikolaidis A., Isaeva E., 2016, *"Black Sea GIS developed in MHI,"* Proc. SPIE 9688 (96881C) doi: 10.1117/12.2241631.

[2] Krylenko M., Zhuk E., Khaliulin A. *Using of GIS technology for access to coastal data // The XIII International MEDCOAST Congress on Coastal and Marine Sciences, Engineering, Management and Conservation.* Mugla: MEDCOAST Foundation, 2017. V.2. P. 757-764.

[3] Godin E., Ingerov A., Khaliulin A., Kosyan R. *Information Support of Coastal Research // Proceedings of the Thirteenth International MEDCOAST Congress on Coastal and Marine Sciences, Engineering, Management and Conservation* (Editor E. Ozhan), MEDCOAST 17, 31 Oct - 04 Nov 2017, Mellieha, Malta, MEDCOAST, Mediterranean Coastal Foundation, Dalyan, Mugla, Turkey, 2017, vol. 2., P. 129-135.

[4] Goryachkin Y., Godin E., Kosyan R., Krylenko M., Kharitonova L. *Monitoring of the West Crimean coast by drone // Innovations in Geology, Geophysics and Geography-2017.* Conference materials of the 2nd International Youth Scientific and Practice Conference, 07-09 July, 2017. - Moscow. Pero, 2017. P.44.

The data life cycle management at SOCIB: responding to science and societal needs

Xisco Notario, Sistema d'observació i predicció costaner de les Illes Balears (Spain), xnotario@socib.es
Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Inmaculada Ruíz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), iruiz@socib.es
Miguel Charcos, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mcharcos@socib.es
Miquel Àngel Rújula, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mrujula@socib.es
Paz Rotllán, Sistema d'observació i predicció costaner de les Illes Balears (Spain), protllan@socib.es
Sonia Gómara, Sistema d'observació i predicció costaner de les Illes Balears (Spain), sgomara@socib.es
Miquel Gomila, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mgomila@socib.es
Charles Troupin, University of Liège (Belgium), ctroupin@ulg.ac.be
Juan Gabriel Fernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jfernandez@socib.es
Joaquín Tintoré, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jtintore@socib.es

The SOCIB-ICTS Marine Research Infrastructure approach

The Balearic Islands Coastal Ocean Observing and Forecasting System (SOCIB, http://www.socib.eu), is a Marine Research Infrastructure (ICTS) that provides world-class, quality controlled metocean datasets, in both real time and delayed mode. This is achieved from across its multi-platform, observation and forecasting system, covering coastal to open ocean areas. This multi-platform approach is needed to properly capture oceanographic processes, that take place at different spatial and temporal scales, and that characterise both ocean state and ocean variability.

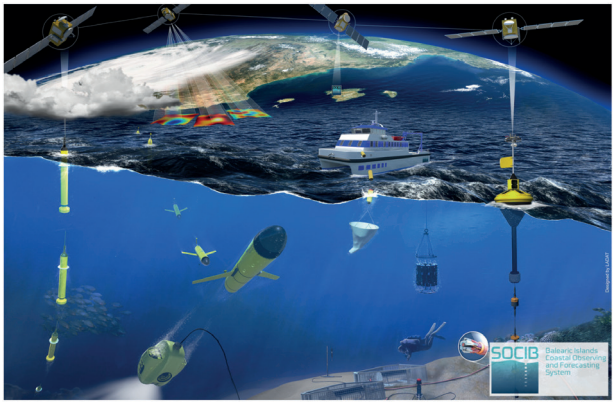


Fig. 1 - SOCIB multi-platform observing & forecasting system.

The SOCIB observation system provides physical and biogeochemical variables from different platforms such as the coastal research vessel, a high-frequency (HF) radar system, weather stations, tide gauges, moorings, drifting buoys, ARGO profilers, gliders (autonomous underwater vehicles) and sea turtle tracking, providing trajectories given by the animals. The forecasting system uses high-resolution numerical models for hydrodynamics (ROMS) and waves (SAPO).

The role of the Data Centre

The DC is responsible for the different stages of data management and covers the whole data life cycle, ranging from data acquisition using SOCIB observational platforms, numerical models or information generated by other divisions, to distribution and visualization through the development of specific tools for visualising the data sets, including both dedicated web and mobile applications. The implemented system relies on open source solutions.

In order to cope with a wide range of platforms, automatic management and processing are necessary. Here we present some of the applications developed to perform the oceanographic data management of the different platforms and a specific example developed for gliders.

- Instrumentation: a database that contains the inventory of materials, the activities performed with them and the processing applied on the collected datasets.
- Processing: an application designed to extract metadata of the deployed equipment from Instrumentation and to perform the data ingestion, processing, quality control and standardization.
- Glider toolbox (https://github.com/socib/glider_toolbox): a complete set of MATLAB/Octave scripts that automates glider data processing function, including thermal lag correction, quality control and graphical outputs.

Applications

Based on the available data and using a set of web services, several applications were build:

- SEABOARD (http://seaboard.socib.es), a dashboard combining different sources of information in real time for different types of users.
- Smartphone apps to access data, platform trajectories and forecasts in real-time.
- "Medclic: the Mediterranean in one click" (http://www.medclic.es/en/), a web dedicated to the Mediterranean Sea monitoring, with scientific and an outreach components.

Other applications are currently being developed as an adaptation to different sectors within the new SOCIB Products and Services 2017 strategy (beach lifeguard and Bluefin tuna apps).

Conclusions

SOCIB organizational and conceptual structure as a facility of facilities including the Data Centre and its developed components is a good example of Marine Information System within the framework of new Ocean Observatories and/or Marine Research Infrastructures, a system of systems that through FAIR principles, generates added value to both cover the scientific community demands and respond to the general societal needs.

Data processing, storing and publishing for the autonomous moored profiler Aqualog

Vladimir Solovyev, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia), sol@ocean.ru

Tamara Shiganova, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia), shiganov@ocean.ru

Alexander Ostrovskii, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia), osasha@ocean.ru

Andrey Zatsepin, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia)

Dmitry Shvoev, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia)

Andrey Tsybulsky, P.P. Shirshov Institute of Oceanology of Russian Academy of Sciences (Russia)

Aqualog is an observational platform that moves down and up along a mooring line, which is taut vertically between a subsurface flotation and an anchor (Fig. 1).

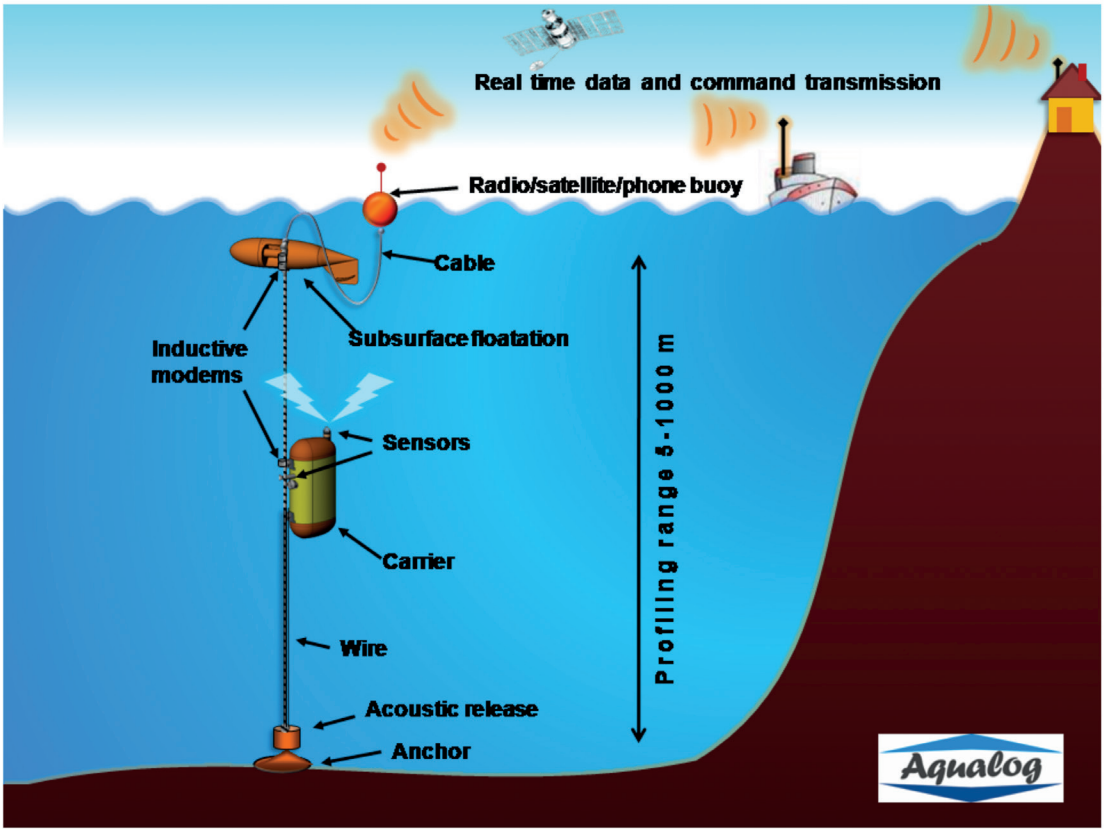


Fig. 1 - The autonomous moored Aqualog profiler observation platform.

The profile was developed in the Shirshov Institute of Oceanology, RAS.

The Aqualog comprises a carrier with a load of oceanographic instruments. The sensors include but are not limited to Teledyne RDI Citadel CTD-ES probe, Nortek Aquadopp-3D current meter, and AANDERAA fast Oxygen Optode4330F. The acoustic Doppler current meter is a horizontal beam single-frequency 2-MHz instrument remotely sensing the water volume in the range of about 0.35–1.85 m from the profiler with a sampling frequency of 23 Hz. When the carrier is moving with the speed of 0.25 m/s, the vertical profiles are measured with a vertical resolution of 0.1 m for pressure, conductivity, and temperature; 0.2 m for acoustic backscatter signal and horizontal current speed; and 1 m for dissolved oxygen.

Besides the cost optimization, the tethered multisensory profiler has other advantages. Unlike conventional mooring where the equipment is placed on fixed depths, Aqualog conducts continuous measurements of vertical profiles applicable for assessing both integral and differential characteristics of the ocean fine structure. By combining pressure, conductivity, temperature, and horizontal current velocity data, it is possible to evaluate vertical mixing. The joint analysis of dissolved oxygen data and the strength of the acoustic backscatter signal give a better understanding of the variability of the marine ecosystem vertical structure at multiple time scales.

There are slots in the profiler reserved to expand the set of sensors by installing additional probes such as the AQUAscat (by Aquatec Group Ltd.) for observations of the particulate matter and biota at multiple acoustic frequencies.

Shirshov Institute of Oceanology has three permanent working the Aqualog profilers. One Aqualog works in the Black Sea since 2010 and two profiler in the Baltic Sea since 2012.

The data obtained by Aqualog profilers are described using the SensorML standard was transferred to the data center of the Shirshov Institute of Oceanology. In the data center, quality control was carried out to verify the following: all the data was in the regional parameter range, ascending and descending profile data variance remained within the predefined limits, pressure and density inversions were eliminated, spikes were tested, proper visual QC was obtained, and etc. After QC check the data store in the oceanographic data warehouse OceanDB.

The oceanographic data warehouse OceanDB has been developed in P.P. Shirshov Institute of Oceanology Russian Academy of Sciences.

OceanDB provides the structured storage of diverse oceanographic data on the following scientific sections: physics of ocean, ocean chemistry, marine biology and ecology and geology. CDI indexes and records for Download manager coupling table are automatically processed for data publish in EMODnet and SeaDataCloud system.

Complying with data interoperability standards in MyCoast project

Pedro Montero, Instituto Tecnolóxico para o Control do Medio Mariño de Galicia (Spain), pmontero@intecmar.gal

Julien Mader, Anna Rubio, Jose Luis Asensio, AZTI (Spain)

Manuel Ruiz-Villarreal, Elena Tel, Alicia Lavín, Carmen Rodríguez, Instituto Español de Oceanografía (Spain)

Garbiñe Ayensa, Begoña Vila, Silvia Allen-Perkins, Instituto Tecnolóxico para o Control do Medio Mariño de Galicia (Spain)

Tomazs Dabrowsky, Marine Institute (Ireland)

Vicente Pérez-Muñuzuri, Universidade de Santiago de Compostela (Spain)

Olivier Cann, Service Hydrographique et Océanographique de la Marine (France)

Marta de Alfonso, Marcos García-Sotillo, Enrique Álvarez, Puersto del Estado (Spain)

Guillaume Charria, Institut Français de Recherche pour l'Exploitation de la Mer (France)

Liam Fernand, Luz García, Centre for Environment, Fisheries and Aquaculture Science (United Kingdom)

Ricardo Torres, Olly Clements, Pierre Cazevane, Plymouth Marine Laboratory (United Kingdom)

Alejandro Gallego, Jenny Hindson, Jens Ramussen, Bill Turrell, Marine Scotland Science (United Kingdom)

Juan Taboada, María Luz Macho, MeteoGalicia (Spain)

Carlos S. Fernandes, Margarida Melo, Instituto Hidrografico (Portugal)

Maria Fernandes, Pedro Agostinho, Qualitas Remos (Portugal)

MyCOAST

In the last decade, major efforts have been devoted to develop large-scale networks and e-infrastructures aiming to share and distribute oceanographic data in the European area using state-of-the-art technologies and standard formats. At the same time several regional coastal ocean observatories have been implemented along the North Atlantic Coast with the aim of complementing existing networks and to cover near-shore areas with higher spatial resolution observations and forecasts, e.g. RAIA, LOREA, the Western Channel Observatory, SmartBay, HOSEA and the upcoming OCASO. The collection and dissemination of data gathering to satisfy local needs was the main objective in the implementation of these observatories and very weak interactions between these different networks and specific e-platforms have taken place. A rationalization and capitalization of the initial efforts through networking and upgrading activities is needed so that a seamless access to the existing oceanographic data, products and services is provided in a homogeneous and standardized way along the Atlantic Coast. The EU Project MyCOAST will integrate these observatories across the entire Atlantic Area and will serve as a link between the large-scale initiatives and the end users. This transnational consistent approach will generate transferable products and services and will provide a vital support to a wide array of

applications, e.g. it is a key requirement in the implementation of MSFD.

MyCoast aims to build a coordinated Atlantic Coastal Operational Observatory in the Atlantic area joining capabilities from all the 5 countries and from existing cross-border cooperation activities, all targeted towards the improvement of coastal monitoring and forecasting tools to support threat and emergency response. The technical networking and specific synergies will strengthen the use and the dissemination of downstream applications of the Copernicus Marine and Environmental Monitoring Service (CMEMS) in order to address the common challenge of resilience of the coastal to risk. The proposed data management tools will promote information sharing and interoperability between coastal observatories and the common European information sharing systems.

Complying wiht data interoperability standards

The effective implementation of a transnational coastal system depends on the capability of multiple systems to work with the same sets of data and metadata in an interoperable fashion. Metadata records associated with one resource should be accessed, accurately interpreted and subsequently used by a system.

MyCOAST aims to homogenize and develop the interconnection of the platforms of coastal data with the existing networks (CMEMS, SeaDataNet); adopting standard conventions (OGC, ISO); fulfilling the INSPIRE Directive; increasing the impact of the transferable.

In order to achieve this goal, the sequence of the tasks will be:

- Explore the state of the art and the level of homogenization among coastal observatories and with European Systems (CMEMS, SeaDataNet). MyCOAST will focus on exploring the state of the art of the standardization of data sets in the different coastal observatories to know the level of harmonization thereof: used standards, ingestion and alignment to international networks, as CMEMS, level of maturity of standards and services. The aim is to find the starting point for choosing which datasets from the coastal observatories can be homogenized in the next actions, as well as which standards and protocols are most suitable.
- Promote a low-level standardization of chosen datasets for all providers gathering the end users necessities: Adoption of a low level of interoperability between different coastal observatories, which is achieved when a tool is able to address datasets using standards but there are discrepancies about internal structures of datasets, files and folders. It will be necessary totake inventory of the data sets to be adapted, and they will be transformed adopting international standards as THREDDS, NetCDF and CF conventions for models or OGC-INSPIRE standards to visualize, download or catalogue the data set.
- Promote a high-level standardization of several datasets and adopt methods to use them : This task will focus on trying to obtain similar structures of datasets of similar measurements or platforms regardless of the coastal observatory of origin. To achieve this, it will be necessary to agree on which variables have to be mandatory, which vocabularies have to be used, etc. In this way, any final tool that uses the data of one of the observatories can use the data of any other, since the interoperability between observatories will be guaranteed.
- Share documents, software and tools in order to use and adopt the defined standards In order to help coastal observatories to make them interoperable between them, a series of guides, manuals and scripts will be developed in order to facilitate the adaptation of their data to ones with a common structure as well as software libraries to be use by the developed tools.

Connecting French coastal ecosystem monitoring databases to SeaDataNet infrastructure: SOMLIT and RESOMAR Networks

Yolanda Del Amo, EPOC University of Bordeaux / CNRS, yolanda.del-amo@u-bordeaux.fr
Nicolas Savoye, EPOC University of Bordeaux / CNRS, nicolas.savoye@u-bordeaux.fr
Serge Heussner, CEFREM University of Perpignan / CNRS, heussner@univ-perp.fr
Nathalie Simon, SBR Sorbonne University / CNRS, nathalie.simon@sb-roscoff.fr
Nicolas Lavesque, EPOC University of Bordeaux / CNRS, nicolas.lavesque@u-bordeaux.fr
Antoine Grémare, EPOC University of Bordeaux / CNRS, antoine.gremare@u-bordeaux.fr
Soumaya Lahbib, MIO Aix-Marseille University & OASU University of Bordeaux / CNRS,
Fabrice Mendès, OASU University of Bordeaux / CNRS, fabrice.mendes@u-bordeaux.fr
Franck Delalee, OASU University of Bordeaux / CNRS, franck.delalee@u-bordeaux.fr
Mark Hoebeke, SBR Sorbonne University / CNRS, mark.hoebeke@sb-roscoff.fr
Fabienne Rigaut Jalabert, SBR Sorbonne University / CNRS, jalabert@sb-roscoff.fr
Pascal Claquin, CREC BOREA University of Caen Normandie / CNRS, pascal.claquin@unicaen.fr
Maurice Libes, MIO Aix-Marseille University / CNRS, maurice.libes@osupytheas.fr
Gérald Grégori, MIO Aix-Marseille University / CNRS, Gerald.gregori@mio.osupytheas.fr
Mélilotus Thyssen, MIO Aix-Marseille University / CNRS, melilotus.thyssen@mio.osupytheas.fr

The SOMLIT is the French Coastal Monitoring Network (http://somlit.epoc.u-bordeaux1.fr/fr/) of the CNRS and marine universities. The network aims at 1) studying multi-decadal changes of coastal systems and identifying climate and anthropic forcings and 2) disseminate data sets for scientific, educational and policy purposes. It gathers eleven teams of Marine Laboratories that carry out time series of sub-monthly resolution in order to understand seasonal variability and long-term changes in coastal systems. Twelve ecosystems distributed over the three main maritime facades of France are monitored since the late 90's. Surface water is sampled for Essential Ocean Variables, among which temperature, salinity, pH, dissolved oxygen, nutrients, SPM, chlorophyll a, particulate organic carbon and nitrogen (POC and PN, respectively), stable isotopes of POC and PN, and pico- and nano-plankton diversity and characteristics. The SOMLIT Network shares best practices among the involved teams and has developed quality check procedures and tools in order to ensure the homogeneity of the practices and the comparability of the results.

Linked to the SOMLIT, the RESOMAR (http://resomar.cnrs.fr/) is the French National Network of Marine Stations and Observatories. It is leaded by the CNRS in collaboration with other national institutions like IFREMER and the National Museum for Natural History. It aims at coordinating actions of national interest, promoting structuring research projects (on physics, biogeochemistry, biology and geology) and being a French interlocutor at the international level. The network was also given the task of establishing and operating two biological databases for benthic and pelagic components of coastal ecosystems.

Fig. 1 - SOMLIT network and RESOMAR network.

At the national level, french observation networks are structured within the French Research Infrastructure for Coastal Ocean and Seashore ILICO, and they strive for European and international connection in order to better understand multi-decadal changes of coastal ocean systems. Within the European H2020 program SeaDataCloud, these coastal observing databases will be in a near future connected to the SeaDataNet portal. For the moment, SOMLIT Network uses a specific MySQL database with a PHP website. In order to achieve the interoperability between systems, metadata had to be added into the database schema. Then, a full mapping between parameters was required in order to allow the SeaDataCloud connection.

The IGME Marine geo-information system: integrating international standards towards INSPIRE-compliancy

Irene Zananiri, Institute of Geology and Mineral Exploration (Greece), izanan@igme.gr
Andreas Valaouris, Institute of Geology and Mineral Exploration (Greece), andywalas@gmail.com
Ioannis Vakalas, Institute of Geology and Mineral Exploration (Greece), vakalasjohn@gmail.com
Vaggelis Zimianitis, Institute of Geology and Mineral Exploration (Greece), zimianitis@igme.gr
Eleftheria Drosopoulou, Institute of Geology and Mineral Exploration (Greece), eleftheria.drosopoulou@gmail.com

During more than 30 years of activity and participation in numerous research projects, the Institute of Geology and Mineral Exploration of Greece has carried out extensive work in marine geology. As a result a vast amount of data has been collected, including marine sediment samples and cores, shallow and medium penetration seismic profiling, bathymetry and side scan sonar data.

Up to a few years ago, management of available information was carried out by means of hard copy maps and reports, archive of seismic paper rolls and data stored in spreadsheets. In order for those data to be readily available and used by local authorities and the international scientific community, a need for standardization and harmonization was imminent. Towards this scope, our involvement in EU projects (e.g. SEISCANEX, Geo-Seas, EMODnet Geology) provided standards for the development of an integrated digital data management scheme in GIS invironment.

The IGME Marine Geology database (IMGdb), used for storage, management, analysis and cartographic representation of the Institute's marine data, was designed as a normalized schema and implemented in an ESRI ArcGIS file geodatabase, using layers and related tables, annotations, raster datasets and relations (topology). It consists of a "back end" that contains all the raw analytical data and supporting metadata, and a ''front end'' comprising data interpretations and syntheses (t.i. maps).

Data are organized in several feature datasets, listed below, as well as raster datasets (e.g. DEM):

- Bathymetry (point data and polylines)
- Bibliography (literature data were scanned, georeferenced and digitized)
- CrossSections (2D interpretations, e.g. Holocene thickness)
- GeophysicalData (data from interpreted seismic profiles)
- Geochemistry (point data from laboratory analyses and interpolation polygon data)
- Geomorphology (point, line, polygon features)
- Sedimentology (point data from laboratory analyses and interpolation polygon data)
- LandGeology (onshore lithology, age and tectonic lines)
- Mapframe (the IGME Marine Geology 1:200K and frames of other datasets)
- Minerals (point data from laboratory analyses and polygon data from interpretations)

- Palaeogeography (derived palaeogeography information)
- Profiles (metadata information)
- ResidentData (administrative information, shoreline)
- SamplingCoring (metadata information)
- TectonicData (derived tectonic interpretations)
- Topography (on-shore features)

The spatial distribution of available information extends over the Hellenic EEZ, also including on-shore water bodies (t.i. lakes). The EPSG:4326 projection system was chosen, to allow interoperability with international datasets. The database design, vocabulary terms and portrayal rules were determined by use of existing common European protocols and standards (e.g. ISO 19115, INSPIRE); common standards were applied for on-shore datasets to allow integrated studies. Auxiliary metadata are INSPIRE-compliant and in accordance with requirements from Geo-Seas and EMODnet e-infrastructures.

The IMGdb is also equipped with several Toolboxes, desighed to serve the specific needs of data analysis and processing: e.g. Folk.tbx (grain size data manipulation using Folk classification scheme), Roundness.tbx (GIS-based evaluation of grain roundness), IndiKrig.tbx (geostatistical mapping using indicator krigging).

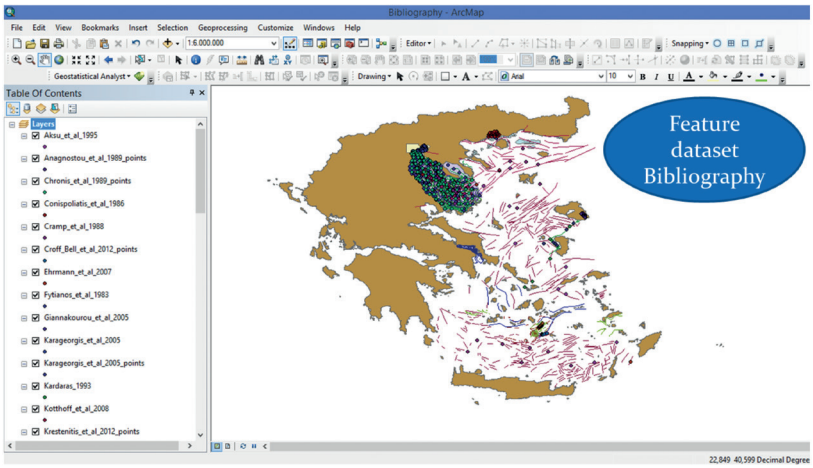


Fig. 1 - An example of the bibliography feature dataset.

Benefits of this geo-information system: all data are stored in a common format, readily available to be used in various studies (selection by area), allowing for combined interpretations and accurate digital mapping. Users can easily discover, access and reuse collected data; auxiliary metadata provide the required information to evaluate data quality and appropriateness.

The Marine Geology database of IGME is a fundamental tool that provides guidelines and standards for future projects. The integrated data management supports Greece towards implementation of the Marine Strategy Framework Directive and the INSPIRE directive, promoting the establishment of Maritime Spatial Planning.

Future steps in the IGME Marine geo-information system comprise full INSPIRE-compliancy and provision of WMS, WFS services. Moreover multiscale GIS layers from various maps and case studies will be visualized and queried interactively through WebGIS.

Building strong foundations towards the pan-European high frequency radar network

Lorenzo Corgnati, ISMAR Consiglio Nazionale delle Ricerche (Italy), lorenzo.corgnati@sp.ismar.cnr.it
Carlo Mantovani, ISMAR Consiglio Nazionale delle Ricerche (Italy), carlo.mantovani@cnr.it
Anna Rubio, Marine Research (Spain), arubio@azti.es
Jose Luis Asensio Igoa, Marine Research, Pasaia (Spain), jlasensio@azti.es
Emma Reyes, Sistema d'observació i predicció costaner de les Illes Balears (Spain), ereyes@socib.es
Antonio Novellino, ETT Solutions Ltd (Italy), antonio.novellino@ettsolutions.com
Patrick Gorringe, EuroGOOS AISBL (Belgium), patrick.gorringe@smhi.se
Annalisa Griffa, ISMAR Consiglio Nazionale delle Ricerche (Italy), annalisa.griffa@sp.ismar.cnr.it
Julien Mader, Marine Research (Spain), jmader@azti.es

High Frequency Radar (HFR) is a land-based remote sensing technology offering a unique insight to coastal ocean variability, as it maps coastal ocean surface currents over wide areas (reaching distances from the coast of over 200 km) with high spatial (a few kms or higher) and temporal resolution (hourly or higher). HFR products are directly used in different sectors, such as Search and Rescue, renewable energy, fishery management and monitoring of pollutants and biological quantities, offering an unprecedented potential for the integrated management of coastal zones. Moreover, in the next years it is expected that HFR surface current data will be systematically ingested in Data Assimilation (DA) processes necessary for predictive model adjustment. It is then crucial to promote and distribute high quality HFR data for scientific, operational and societal applications.

HFR technology is rapidly expanding in Europe, where the number of systems is growing at a rate of 6 new sites per year, with over 58 systems currently deployed and operational and a number in the planning stage. Since the European HFR systems are playing an increasing role in the overall operational oceanography marine services, many initiatives are now active in Europe aiming at building an operational pan-European HFR network based on a coordinated data management. In 2014, EuroGOOS launched the HFR Task Team to achieve the harmonization of system requirements and design, data quality and standardization of HFR data access and tools. In 2015, a pilot action coordinated by EMODnet Physics, begun to develop a strategy for assembling HFR metadata and data products within Europe in a uniform way to make them easily accessible, and more interoperable. The SeaDataCloud (SDC) project, launched in 2016, is contributing to the integration and long-term preservation of historical time series from HFR into the SeaDataNet infrastructure by defining standard interoperable data and Common Data Index (CDI) derived metadata formats and Quality Control (QC) standard procedures for historical data. Recently, the Copernicus Marine Environment Monitoring Service (CMEMS) Service Evolution Call supported the INCREASE project, which set the bases for the integration of existing European HFR operational systems into the CMEMS-INSTAC (In Situ Thematic Assembly Center). In parallel, EU project JERICO-NEXT is working to provide procedures and methodologies to

enable HFR data to comply with the international standards regarding their quality and metadata, within the overall goal of integrating the European coastal observatories.

The results of these integrated efforts are significant and promising. The European common data and metadata model for real-time surface current HFR data has been defined and implemented, compliant with Climate and Forecast Metadata Convention version 1.6 (CF-1.6), OceanSITES convention, CMEMS-INSTAC requirements and INSPIRE directive. Furthermore, the list of the QC tests to be applied to HFR data has been defined according to the DATAMEQ working recommendations on real-time QC and building on the Quality Assurance/Quality Control of Real-Time Oceanographic Data (QARTOD) manual produced by the US Integrated Ocean Observing System (IOOS). Thanks to these achievements, the inclusion of HFR data into CMEMS-INSTAC and into SDC Data Access was decided to ensure the improved management of several related key issues as Marine Safety, Marine Resources, Coastal & Marine Environment, Weather, Climate & Seasonal Forecast.

CMEMS-INSTAC and SDC operate through a decentralized architecture based on National Oceanographic Data Centres (NODC), Production Units (PUs) organized by region for the global ocean and the six European seas and a Global Distribution Unit (DU). In particular, CMEMS-INSTAC implements the functions of data acquisition, QC, validation/assessment and distribution. The core of CMEMS-INSTAC and SDC is to guarantee that for the users the quality of the product delivered is equivalent wherever the data are processed.

HFR data are in situ gridded data in time (big data), therefore the standard in situ data management infrastructures have to be organized and adapted to allow INSTAC PUs, other CMEMS Thematic Centres (TAC) and Marine Forecasting Centers (MFC) to efficiently manage this type of data. The establishment of the HFR data stream has to be organized in the coordinated framework formed by the existing main European infrastructures and actors. Given the importance of the data type and the diversity with the already available data streams and quality check procedures, the implementation of the HFR data stream has to come together with the development of a centralized European competence centre.

Thus, the development of a centralized HFR node goes towards three additional main steps:

i) to set up a data centre dedicated to link all the available data providers and collect and process HFR data;
ii) to develop and upgrade the software tools for the harmonization of data and metadata of HFR data coming from different sources;
iii) to apply data processing, both in real time and delayed mode, and create catalogues of HFR data compliant with the requirements of CMEMS-INSTAC and SDC.

Its implementation should be based on a hierarchical infrastructure to facilitate management and integration of any potential data provider according to a simple and very effective rule: if the data provider can set up the data flow according the defined standards, the HFR central node only has to link and include the new catalogue and data stream. If the data centre cannot setup the data flow (because of lack of experience, technical capacity, etc.), the HFR node will work on harvesting the data from the provider, harmonize and format these data and make them available.

The integration and assessment of the HFR data in a centralized data system will allow a second harmonized level of quality check assessment, interoperable data products and a more efficient implementation of tools for downstream services. For all these reasons the establishment of a centralized HFR node should be the cornerstone of the operational European HFR network.

EMODnet PP: Portugal presence

Sara Almeida, Instituto Hidrográfico (Portugal), sara.almeida@hidrografico.pt
Eric Magalhães, Instituto Hidrográfico (Portugal)
José Aguiar, Instituto Hidrográfico (Portugal)

"The overall objective of EMODnet Physics is to provide access to archived and near real-time data on the physical conditions of European seas and oceans based on observation of the sea, such us wave height and period; temperature of the water column; wind speed and direction; …sea-level; data from rivers'' from the portal, www.emodnet.eu/physics.

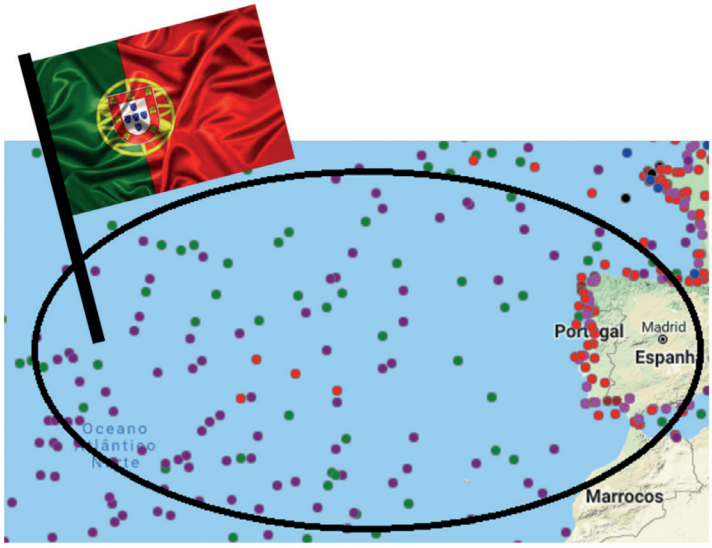


Fig. 1 - EMODnet PP - data from Portuguese waters.

Due to the national responsibility of managing the existing observation system network, Hidrografico was present at EMODnet from the launch of the pilot portal for Physical Parameters in 2011. Since then, another three Portuguese entities began contributing to the this general purpose: two Centers from Universities - Azores and Lisbon - and the national environmental agency. This provided access to a wider range of in-situ measurements.

Hidrografico manages a network with a variety of fixed stations with several oceanographic and meteorological monitoring capabilities. The mutiparametric buoys component of the network comprises four platforms installed in deep waters near Leixões, Nazaré (two units) and Faro. They collect data on wave dynamics, sea surface temperature, wind speed and direction, air temperature and atmospheric pressure. Data is collected up to a depth of 2000 m.

Near real time data from five tide gauges, distributed along the continental coast, and from three HF Radar stations measuring of surface currents are made available through the IBIRoos' Assembly Center (Puertos del Estado) and, subsequently, to the EMODnet portal.

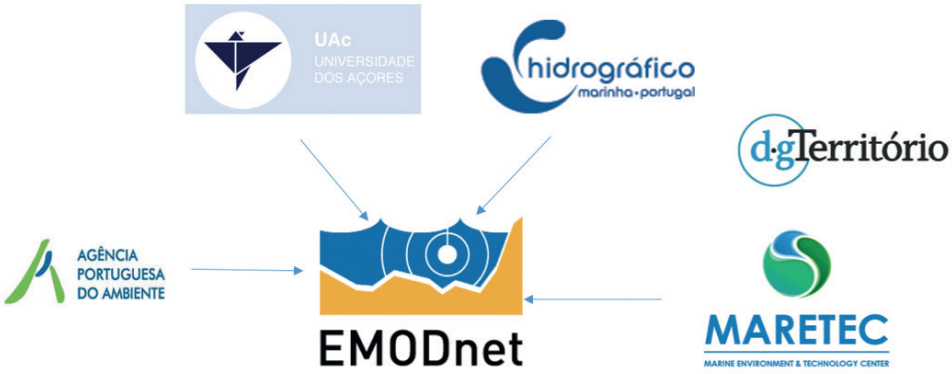


Fig. 2 - Portuguese data providers.

The Center for Climate, Meteorology and Global Change Studies from the University of Azores, contributes with near real time data from waverider buoys with wave parameters and sea surface temperature. They are spread geographically in the North Atlantic Ocean over six islands.

Maretec (Research Centre of IST for Marine, Environment and Technology) from University of Lisbon, assembles and provides data from the river stations, regarding the water flow.

The pertinent point is that data made available to the EMODnet portal follows standard automatic quality control procedures so as to provide marine data that is harmonised consistent. This allows for the interoperability from a variety of sources, which proved to be an added value for the users as a single point of entry allows for a multitude of uses.

Near future

Before the end of 2018, Hidrografico intends to aggregate two new providers, both from Madeira Island. The APRAM (Administração dos Portos da Região Autónoma da Madeira) with wave parameters and sea surface temperature from Madeira buoys and the Oceanic Observatory of Madeira with HF Radar data from two platforms, Calheta and Camara de Lobos.

And from it's own resources, Hidrografico intends to include the HF radar stations, covering the Lisbon area, in the system.

References

http://www.emodnet-physics.eu/portal/
http://www.uac.pt/
http://www.maretec.org
http://www.ibi-roos.eu

EMODnet Geology - discover Europe's seabed geology

Henry Vallius, Geological Survey of Finland (Finland), henry.vallius@gtk.fi
Irene Zananiri, Institute of Geology and Mineral Exploration (Greece), izanan@igme.gr
and **The EMODnet Geology Team**

The European Marine Observation and Data Network (http://www.emodnet.eu) is financed by the European Union, currently under Regulation (EU) No 508/2014 of the European Parliament and of the Council of 15 May 2014 on the European Maritime and Fisheries Fund. It consists of more than 160 organisations assembling marine data, products and metadata to make these fragmented resources more accessible to public and private users, relying on quality-assured, standardised and harmonised marine data which are interoperable and free of restrictions on use. EMODnet is currently in its third development phase with the target to be fully deployed by 2020 and it's main scope, as part of the Integrated Maritime Policy Action Plan, is to support "Marine Knowledge 2020".

The EMODnet Geology thematic lot, initiated as a pilot project in 2009 (ur-EMODnet), is now running it's third phase (2017-2019), coordinated by GTK, with the participation of 34 partners and 5 subcontractors from 31 countries. It has succeded in providing full coverage of all European regional seas: Adriatic Sea, Aegean and Levantine Sea, Baltic Sea, Barents Sea, Bay of Biscay and Iberian Coast, Black Sea, Celtic Sea, Faeroe Islands EEZ, Greater North Sea, Iceland Sea, Ionian and Central Mediterranean Sea, Macaronesia, Norwegian Sea, Western Mediterranean Sea, White Sea and the Wider Atlantic, while the target of the current phase is to consolidate the existing data products with higher resolution (scale 1:100.000 or finer) and more contents.

The geology data compiled in the frame of the project, along with deliverables of the preparatory phase and phase II, will during this phase of EMODnet Geology be available through the portal http://www.emodnet-geology.eu, and include:

- Sea-bed substrate (sediment layer at the sea floor & sediment accumulation rate);
- Sea-floor geology - lithology (bedrock geology beneath the surficial sediment: Quaternary deposits and pre-Quaternary);
- Geomporphological features of the sea-floor;
- Coastal behavior (migration, resilience and vulnerability);
- Geological events and probabilities (e.g. submarine landslides, volcanic centres, earthquakes);
- Mineral occurences (e.g. oil and gas, aggregates, metallic minerals);
- Submerged landscapes.

The information included in the project is principally that held by the partners, with the addition of connections to other data providers using Web Map Services (WMS), as for example by linking to the European-Mediterranean Seismological Centre (EMSC) for earthquake activity data.

A characteristic of the EMODnet Geology project is that the main focus is on harmonised

interpreted map information rather than the underlying data that have been used to create the interpreted geological outputs. However, the web delivery mechanism, using open source standards to ensure long-term sustainability, also aims at providing access to data catalogues of information held by each partner. Through the EMODnet Geology portal a range of services and functionalities for viewing and downloading geology data and products is available:

- A product catalogue and access service (WMS) with an associated description for each dataset.
- Composite Products Discovery and Access Service: WMS, WFS, CSW. Each result links to a full metadata record. Data products can also be found using the EGDI metadata database.
- Map viewer with layer selection and configuration menu.

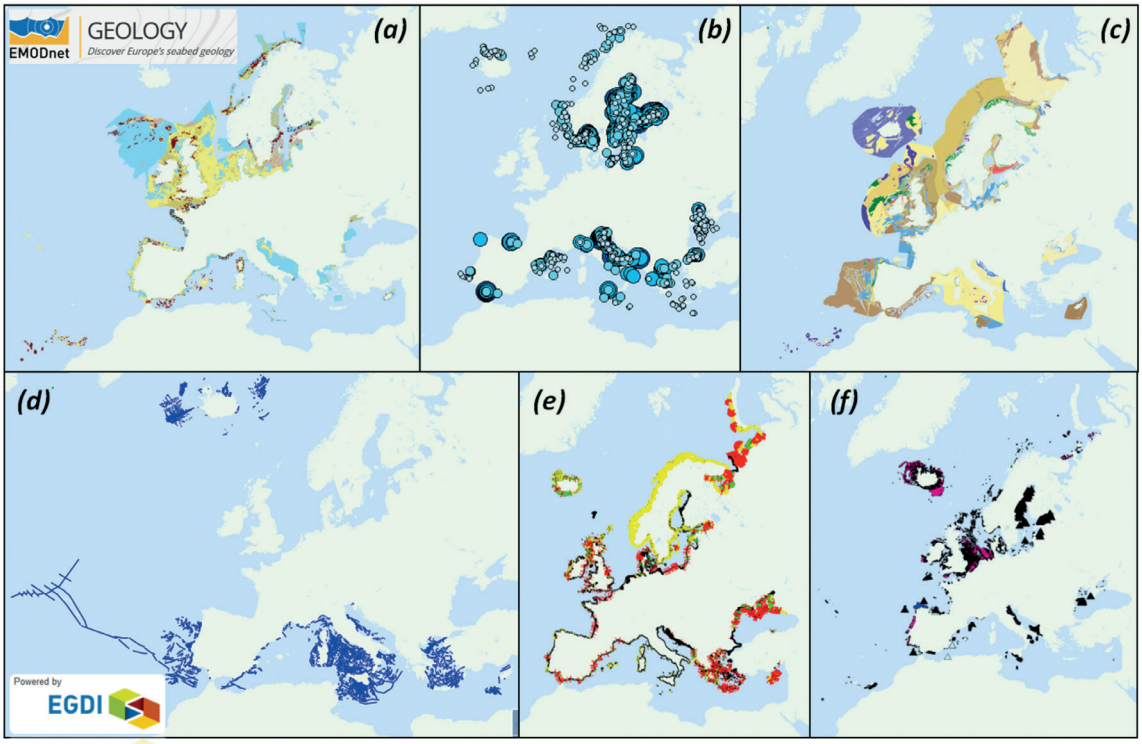


Fig. 1 - Representative datasets available through the EMODnet Geology portal (a) Seabed substrate map 1:250k, (b) Sediment Accumulation Rates, (c) Sea-floor (bedrock) lithology, (d) Tectonic lines, (e) Coastal migration, (f) Mineral occurrences (aggregates, evaporites, hydrocarbons).

EMODnet Geology is an ongoing dynamic project, where - apart from creating new products - existing datasets are continously updated and completed with new data. Thus to ensure sustainability these datasets are also available through the European Geological Data Infrastructure (EGDI). Moreover, the EMODnet Data Ingestion initiative has been developed, reaching out to potential data providers from the public and private sector that are not yet connected, who can easily release their data for safekeeping and subsequent distribution through EMODnet.

Arctic Research Icebreaker Consortium (ARICE) and its oceanographic data management

Alexander Smirnov, Arctic Portal (Iceland), alexander@arcticportal.org
Veronica Willmott, Alfred Wegener Institute (Germany), Veronica.Willmott@awi.de
Halldor Johannsson, Arctic Portal (Iceland), halldor@arcticportal.org

About the ARICE project

The recent changes of the Arctic and the resulting increased economic activity in the region have triggered a societal demand for accurate sea-ice and weather predictions, information on the status of the Arctic Ocean and its marine life, and complex predictions of future scenarios. To address these issues of particular environmental and societal concern and to develop policy recommendations for a sustainable usage of the Arctic Ocean and its resources, the international Arctic science community must have access to world-class research icebreakers to access the ice-covered Arctic Ocean. The lack of availability of icebreakers in Europe and beyond and a not optimally coordinated polar research fleet and polar programmes impedes Europe's capacity to investigate this region. There is thus an urgent need for providing European researchers with better research icebreaker capacities for the Arctic. ARICE aims at reaching this goal with the existing polar fleet by:

1) Networking:

ARICE will develop strategies to ensure the optimal use of the existing polar research vessels at a European and international level, working towards an International Arctic Research Icebreaker Consortium which shares and jointly funds operational ship time on the available research icebreakers.

2) Transnational access:

ARICE will provide transnational access to a set of six key European and international research icebreakers for European scientists, based on scientific excellence of submitted proposals:

- PRV Polarstern, Germany;
- IB Oden, Sweden;
- RV Kronprins Haakon, Norway (under construction, to be operative in 2017);
- RRS Sir David Attenborough, United Kingdom (under construction, to be operative in 2018);
- CCGS Amundsen, Canada;
- RV Sikuliaq, United States of America.

3) Joint research activities:

ARICE will improve the research icebreakers' services by partnering with maritime industry on a "ships and platforms of opportunity" programme and by exploring into new key technologies

that could lead to an improvement of ship-based and autonomous measurements in the Arctic Ocean. ARICE will also implement virtual and remote access of data via an innovative 3D Virtual Icebreaker.

Data management in ARICE

ARICE will generate a variety of data sets which can be sub-divide into three categories :

1) Oceanographic observations : bathymetry, water temperature, salinity, currents, sea state (wave height), ocean surface heat flux, ocean colour, pH, nutrients, carbon (inorganic or organic), oxygen or nitrous oxide (N2O), primary production / chlorophyll, fish abundance, marine mammals and polar bears ;
2) Atmospheric parameters : air temperature, wind speed and direction, air pressure, water vapor or humidity, precipitation, cloud fraction, cloud base height, CO2, methane or other greenhouse gases, ozone or aerosols, radiation budget ;
3) Sea ice / snow parameters : sea ice thickness, sea ice concentration, sea ice extent/sea ice edge, sea ice drift, ice load (on the hull), floe size, ice ridge density, compression in the ice pack, snow/ice surface topography, snow thickness, albedo and melt ponds (fraction or appearance);

ARICE aims at its research data to be findable, accessible, interoperable and reusable (FAIR).

Data management in ARICE will be coordinated with relevant Arctic and oceanographic data management systems and adopts the Arctic Data Committee standards. The Arctic Data Committee can be visited at https://arcticdc.org/.

All data generated under ARICE funding will be accessible to the user group who collected the data. The Principal Investigator of an ARICE cruise must submit the data generated together with sufficient metadata to the respective IODE National Oceanographic Data Centre (NODCs) directly after the cruise. The NODC will then make sure that the data sets are quality controlled, archived in the NODCs and linked to the metadata of the respective cruise. All Datasets will also be integrated into the ARICE Data Management System, the ARICE project's database and disseminated in an interoperable open format through the 3D Icebreaker outreach tool. Access to the datasets, apart from metadata, can be is restricted to the scientific party and its designated partners for a period of 1 to 3 years after the cruise takes place (depending on specific RV procedures). Requests of external users for data access during this time will be forwarded to the data originators for their decision.

ARICE is a part of the H2020 Open Research Data Pilot which aims at improving and maximising the access to and re-use of research data generated projects.

A 9-year monitoring of environmental changes over the continental shelf in the Catalan Sea from multiparametric measurements

Nixon Bahamon, Consejo Superior de Investigaciones Científicas (Spain), bahamon@ceab.csic.es
Miguel-Angel Ahumada, Universidad del Mar (México), ahumada@angel.umar.mx
Raffaele Bernardello, Barcelona Supercomputing Centre (Spain), raffaele.bernardello@bsc.es
Charlotte Reuschel, Fresenius University of Applied Sciences (Germany), reuschel.charlotte@stud.hs-fresenius.de
Joan Baptista Company, ICM Consejo Superior de Investigaciones Científicas (Spain), batista@icm.csic.es
Jacopo Aguzzi, ICM Consejo Superior de Investigaciones Científicas (Spain), jaguzzi@icm.csic.es
Antonio Cruzado, Oceans.cat (Spain), acruzado@oceans.cat

Introduction

Long-term global warming trends reported in the last decades by NOAA-NASA may be accelerating in recent years (Peñuelas *et al*., 2017). Extreme global average temperatures reported since 2014 are consistent with the largest annual atmospheric CO2 what may be a signal of slowdown of carbon sequestration, in oceans and continents. The Mediterranean Sea warming trend is in line with global warming (Schroeder *et al*., 2016).

Coastal marine ecosystems in the Mediterranean are particularly exposed to the effect of warming, as they hold relatively high biodiversity at shallow water layers. Longer and warmer summer periods are related to massive mortalities of benthic (deep-sea) organisms (Bensoussan *et al*., 2010), because the specific temperature ranges of some species may be overtaken.

Rationale

Though long-term coastal monitoring systems allow better approaching the effect of water heating on marine biota, those monitoring systems are very scarce. The Operational Observatory of the Catalan Sea (OOCS) was one of those expectations (Bahamon *et al*., 2011), operating from March 2009 to March 2018. The OOCS collected in-situ data from fixed coastal observation stations over the continental shelf in the Catalan Sea, fuelling time series data of physical and biogeochemical characteristics of the marine environment from hourly to fortnightly time scales.

Data were collected from instrumentation in a meteorological and oceanographic buoy (Fig. 1) and from instrumentation on-board a research boat.

Fig. 1 - Oceanographic buoy moored at the observation station in the Blanes canyon head.

Products

On-board measurements for the period between 2014 to 2017 indicated that yearly temperature of the water column over the continental shelf was 0.4°C higher than temperature in the period between 2010 to 2013 (Fig. 2).

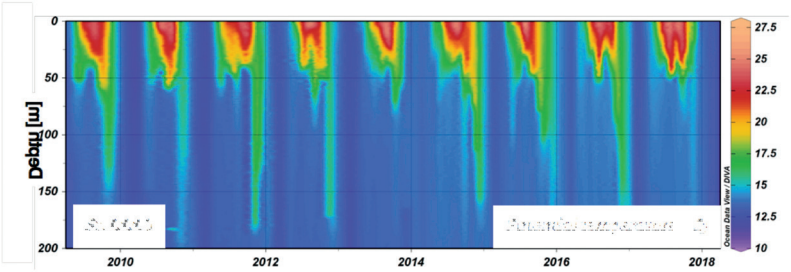


Fig. 2 - Time series of water column temperature over the continental shelf in the Blanes canyon head.

Simultaneously collected in-situ measurements from the buoy instrumentation (e.g. atmospheric pressure, relative humidity, wind speed and air and sea temperatures) along with reanalysis products from NOAA (sea surface albedo and cloud cover), allowed monitoring daily net air-sea heat fluxes (Bensoussan *et al*., 2010).

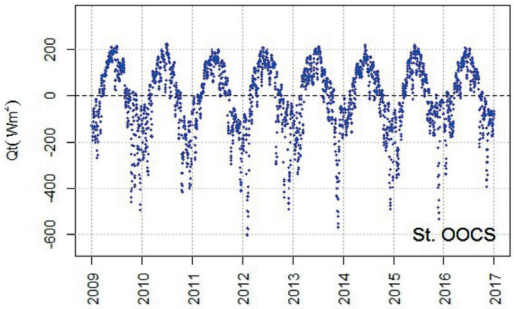


Fig. 3 - Estimated net sea-surface heat fluxes at the observation station in the Blanes canyon head. Negative values indicate heat losses from seawater, whereas positive values indicate heat gains by seawater.

Since 2013, turning points from negative to positive heat flux started earlier, whereas the turning points from positive to negative starter later. This made the summer periods since 2013 were about 13 days longer than summers before 2013. Persistence of longer and warmer summers are expected to amplified the negative effects on the dynamics of the pelagic ecosystem (Bensoussan *et al*., 2010; Williamson *et al*., 2016).

References

Peñuelas J., Ciais P., Canadell J. G., Janssens I. A., Fernández-Martínez M., Carnicer J., Obersteiner M., Piao S., Vautard R. & Sardans J., 2017. *Shifting from a fertilization-dominated to a warming-dominated period.* Nature Ecology & Evolution, 1, 1438–1445.

Schroeder K., Chiggiato J., Bryden H. L., Borghini M. & Ben Ismail S., 2016. *Abrupt climate shift in the Western Mediterranean Sea.* Scientific Reports, 6: 23009. doi:10.1038/srep23009.

Bensoussana N., Romano J.-C., Harmelin J.-G. & Garrabou J., 2010. *High resolution characterization of northwest Mediterranean coastal waters thermal regimes: To better understand responses of benthic communities to climate change.* Estuarine, Coastal and Shelf Science, 87 (3), 431-441.

Bahamon N., Aguzzi J., Bernardello R., Ahumada-Sempoal M.-A., Puigdefabregas J., Cateura J., Muñoz E., Velásquez Z. and Cruzado A., 2011. *The New Pelagic Operational Observatory of the Catalan Sea (OOCS) for the Multisensor Coordinated Measurement of Atmospheric and Oceanographic Conditions,* Sensor, 11(12), 11251-11272. https://doi.org/10.3390/s111211251

Williamson P., Smythe-Wright D. and Burkill P., Eds., 2016. *Future of the Ocean and its Seas: a non-governmental scientific perspective on seven marine research issues of G7 interest.* ICSU-IAPSO-IUGG-SCOR, Paris.

How a virtual appliance facilitates the data management, the SeaDataCloud project experience

Leda Pecci, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (Italy), leda.pecci@enea.it

Marcello Galli, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile (Italy), marcello.galli@enea.it

The SeaDataCloud project has in its main objectives the innovation of the SeaDataNet infrastructure, reached by means of the adoption of cloud computing as well as the High Performance Computing (HPC). In the framework of the project's activity there is also a challenge to simplify the management of the network nodes, especially for new data centres, of the semi-distributed system architecture, that is the base of the infrastructure. The SeaDataNet infrastructure allows to access to a very large number of ocean and marine data sets stored by more than 110 data centres deployed in a territory that extends beyond Europe. The central portal plays the role of providing data visualisation, discovery and access services, the latter service is based on the CDI (Common Data Index) system which includes a key component that allows data centres to manage data sets requests. The component, called Download Manager (DM), is a Java application that communicates with the central portal and provides data to SeaDataNet users.

A solution that helps the system managers of data centres from installation, configuration and maintenance of the SeaDataNet node has been carried out. Data providers that intend to become a node of the infrastructure have to install, configure and maintain a server that contains the D. M. application, that manages the delivery of the data files ordered, and the software required to run the application. Virtualization technology can optimise the connection of data centres to the infrastructure and delivers an alternative over traditional hardware servers. We make available a pre-built virtual appliance, containing the D. M. component, that is a special purpose virtual machine, based on a linux operating system and contains only the necessary software applications (e.g. web and application servers) with an operating system that is enough for it to run optimally and, thanks to which the system is less vulnerable to security breaches. This virtual appliance allows delivering and deployment of the D. M. application, which requires not easy configuration after being installed, in an easy way.

The SeaDataNet virtual appliance includes all necessary data management tools, it is easily deployable into a compatible virtualisation environment, ready to use with minimal setup, which eases the process of maintenance and management of the distributed nodes. We have also prepared a bash procedure for a semi-automatic installation of the appliance software and some testing.

The virtual appliance is realised using a TurnKey Linux 14 (TKL), based on Debian 8 (Jessie) a stable version of the operating system, suitable for the production environment, well known for its robustness and reliability. The operating system updates and software patches contained in the

SeaDataNet virtual appliance are automatically downloaded and installed. The virtual appliance contains only open source software.

Virtual appliances eliminate the need for physical hardware because they run on virtual platform solutions. The advantages of adopting this solution are different, data centres save money on hardware, software, the availability and reliability of the global infrastructure increases, because the disaster recovery of the nodes is simplified, the operating system and the data can be easily recovered by duplicating images of the virtual appliance and migrating them in another host system. Furthermore virtualization provides the ability to adapt the hardware to real needs, increasing the amount of RAM, the size of hard disks, etc., via software.

On a single physical platform multiple virtual machines can run, this allows to have lower energy consumption, that means lower costs and benefits for the environment.

The appliance is usually managed via a ssh shell, but can also be managed by a web interface, which uses the tools shellinabox and webmin. Currently, the time required to install and configure the download manager is reduced and the maintenance of the system is easier.

Sea2Data – from acquisition to advice

Sjur Ringheim Lid, Institute of Marine Research (Norway), sjur.ringheim.lid@hi.no
Åge Fotland, Institute of Marine Research (Norway), aage.fotland@hi.no
Harald Gjøsæter, Institute of Marine Research (Norway), harald@hi.no
Espen Johnsen, Institute of Marine Research (Norway),espen.johnsen@hi.no
Jens-Otto Krakstad, Institute of Marine Research (Norway), jens.otto.krakstad@hi.no
Helge Sagen, Institute of Marine Research (Norway), helge@hi.no
Edvin Fuglebakk, Institute of Marine Research (Norway), edvin.fuglebakk@hi.no
Nils Olav Handegard, Institute of Marine Research (Norway), nilsolav@hi.no

Institute of Marine Research in Norway (IMR) provides advice to decision makers on questions related to anthropogenic impact on the marine environment and how to harvest marine living resources. It is required that our advisory processes are transparent, has efficient access to data and uses standardized and quality-assured data-processing tools.

IMR has over the last years been building an e-infrastructure addressing these requirements using fisheries management as a use case. This infrastructure includes data acquisition, data transfer and management, data storage and data processing. The design is modular and the various modules communicates through application programming interfaces (APIs) that allow us to upgrade, maintain and replace the modules independent of each other. The current modules on data acquisition include data registration software and data storage on our research vessels, data transfer to shore based stations, a shore based data centre module that allows versioning of the data, a light weight data browser for efficient data access, an R-package to access and process data, and a standalone software to estimate indices of fish abundance. Historical data is converted and incorporated in the infrastructure to allow estimation of historical survey time series. Without a close collaboration between developers and users through an agile approach, the project would not have met its objectives.

Is single-window data search a myth or can it be made real? Developing federated search for polar oceanographic and terrestrial data

Taco de Bruin, Royal Netherlands Institute for Sea Research (The Netherlands), Taco.de.Bruin@nioz.nl
Roger Proctor, Integrated Marine Observing System, University of Tasmania (Australia), Roger.Proctor@utas.edu.au
Alexander Smirnov, Arctic Portal (Iceland), Alexander@arcticportal.org
Anton Van de Putte, biodiversity.aq (Belgium), antonarctica@gmail.com
William Manley, Institute of Arctic and Antarctic Research (United States of America), William.Manley@colorado.edu
Brendan Billingsley, Billingsley Custom Software (United States of America), billingb@gmail.com
Halldor Johansson, Arctic Portal (Iceland), Halldor@arcticportal.org
Marten Tacoma, Royal Netherlands Institute for Sea Research (The Netherlands), Marten.Tacoma@nioz.nl
Peter Pulsifer, National Snow and Ice Data Center (United States of America), Peter.Pulsifer@colorado.edu
Stein Tronstad, Norwegian Polar Institute (Norway), Stein.Tronstad@npolar.no
Thomas Vandenberghe, Belgian Marine Data Centre (Belgium), Tvandenberghe@naturalsciences.be
Pip Bricher, Southern Ocean Observing System (Australia), data@soos.aq

Single-window metadata search across multiple data repositories has been a constant desire of science communities in high latitudes environments. Near the poles, observations are sparse and logistical challenges severe. These factors make existing observation data particularly valuable to researchers. However, due to the international and interdisciplinary nature of these science communities, the data are fragmented across a bewildering array of data centres and catalogues. For example, the Southern Ocean Observing System (SOOS) has identified at least 50 catalogues that hold data of interest to the Southern Ocean science community. Federated search is a promising set of technologies that may allow scientists to achieve their dream of single-window search.

SOOS, the Standing Committee on Antarctic Data Management (SCADM) and the Arctic Data Committee (ADC) have formed POLDER, a working group to develop federated search for polar regions. POLDER is investigating and working with the available tools and initiatives in this field to ensure that the polar data communities are working together with our colleagues in the rest of the world to avoid reinventing any wheels.

In this presentation, we introduce POLDER and report on its progress to solving the technical and human challenges that have so far prevented the scientific community of meeting its goal of finding "all" the relevant data in a single-window search. We invite the marine data community to work with us to solve this common problem.

SESSION 4

Data products, information and knowledge

- Quality Check procedures and data consistency assessment for historical data collections for climatic studies
- Machine learning and its application to data collection and data quality check
- Regional and global data collections and marine data aggregations
- Validation of data products and uncertainty estimation
- Distribution and publication of data products
- Data products for multiple stakeholders/end users including the Blue Economy/Blue Growth
- Data integration and evaluation of data portal solutions
- Transfer of technologies and knowledge
- Systems and products for outreach and dissemination, including Ocean Literacy

ORAL PRESENTATIONS

A new world ocean temperature profile product (v0.1): the International Quality controlled Ocean Database (IQuOD)

The IQuOD Project Team (including SCOR working group 148 and IOC/IODE SG-IQuOD),
IQuOD.org, catia.domingues@utas.edu.au

Reliable long-term ocean subsurface temperature measurements are critical for understanding changes in the Earth's energy imbalance, ocean temperature, sea level, and also separating natural variability from anthropogenic factors.

The International Quality Controlled Ocean Database (IQuOD, www.iquod.org), with support from programmes, such as CLIVAR, SCOR and IODE, is the first internationally-coordinated community effort to enhance the content, utility, and quality of the global historical profile database of subsurface temperature observations, and with the potential to expand to other oceanographic variables (salinity, oxygen, etc.).

Progress to date includes inclusion of uncertainty estimates to temperature profiles on each depth (pressure) and intelligent metadata for unknown probe types (identified through a deterministic approach using country, year, depth, etc, Palmer *et al*., 2018) for a major instrument of the historical record. This initial progress has been recently released as part of an interim product, IQuOD v0.1, which is now part of the existing NOAA's World Ocean Database (WOD)

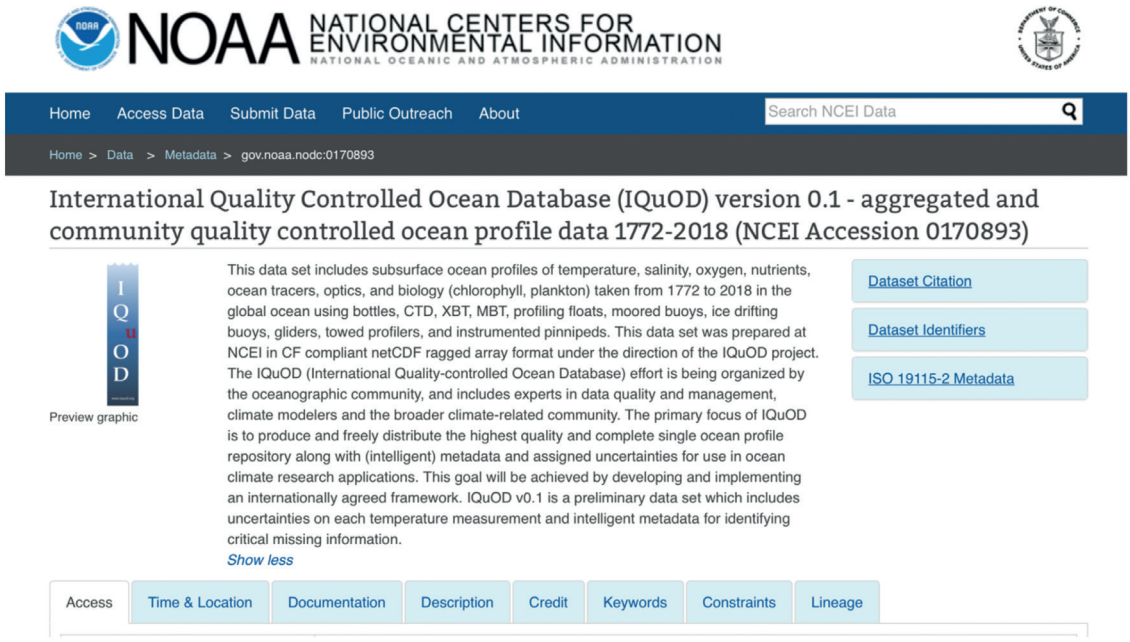


Fig. 1 - NOAA'S NCEI landing page for IQuOD v0.1.

with IQuOD quality flags. Users can publicly download this WOD/IQuODv0.1 through the WODselect system in a Climate-Forecast (CF) compliant netCDF ragged array format. This interim product will also be distributed to the public via other data centers in Australia, France, Germany, Japan and the UK.

Additional efforts are underway to establish an internationally-agreed optimal set of automatic quality control procedures (https://github.com/IQuOD/AutoQC) and to incorporate machine learning into expert quality control. These ongoing efforts will be published as part of future IQuOD interim product releases.

To download and view the documentation for IQuOD v0.1 profile dataset, please visit:

https://data.nodc.noaa.gov/cgi-bin/iso?id=gov.noaa.nodc:0170893

CLIVAR Global synthesis and observations panel: http://www.clivar.org/clivar-panels/gsop

SCOR: http://www.scor-int.org/SCOR_WGs_WG148.htm

IODE: https://www.iode.org/index.php?option=com_content&view=article&id=461&Itemid=100199

References

Palmer M., T. Boyer, R. Cowley, S. Kizu, F. Reseghetti, T. Suzuki, and A. Thresher, 2018: *An algorithm for classifying unknown expendable bathythermograph (XBT) instruments based on existing meta data.* J. Atmos. Oceanic Technol. doi:10.1175/JTECH-D-17-0129.1, https://journals.ametsoc.org/doi/abs/10.1175/JTECH-D-17-0129.1

OceanBestPractices System: a global resource to facilitate harmonizing practices in ocean observation, data and information

Peter Pissierssens, UNESCO-IOC International Oceanographic Data and Information Exchange (Belgium), p.pissierssens@unesco.org

Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es

Johannes Karstensen, Helmholtz Centre for ocean Research Kiel (Germany), jkarstensen@geomar.de

Jay Pearlman, Institute of Electrical and Electronics Engineers (United States of America), jay.pearlman@ieee.org

Pauline Simpson, UNESCO-IOC/IODE - Central Caribbean Marine Institute (Cayman Islands), psimpson@reefresearch.org

Mark Bushnell, NOAA Integrated Ocean Observing System (United States of America), mark.bushnell@noaa.gov

Pier Luigi Buttigieg, Alfred Wegener Institute (Germany), pbuttigi@mpi-bremen.de

Juliet Hermes, South African Environmental Observation Network (South Africa), juliet@saeon.ac.za

Emma Heslop, UNESCO-IOC Global Ocean Observing System (France), e.heslop@unesco.org

Frank Muller-Karger, University South Florida, (United States of America), carib@usf.edu

Francoise Pearlman, Institute of Electrical and Electronics Engineers (United States of America), francoise.pearlman@ieee.org

Introduction

There is an ever-present need for the identification and dissemination of best practices in the multidisciplinary field of ocean observation and data management. However, the complexity of these domains and the diversity of its stakeholders make discovering relevant best practices (BP) a challenge. Addressing this challenge requires not only the creation of a repository enabling efficient discovery and access of documented best practices, but also expanding means of community engagement, including peer review and training. The AtlantOS Project (through its Best Practices Working Group), the ODIP Project, Frontiers in Marine Science and the IODE of the IOC of UNESCO and others are collaborating on the development of an enhanced Ocean Best Practices System, based upon expanding the already existing UNESCO-IOC/IODE OceanBestPractices Repository www.OceanBestPractices.net and using new means to foster OBP use by a broader ocean community. This paper presents the progress of the collaborative efforts in developing the System.

The Ocean Best Practices System

As noted above, an Ocean Best Practices System (the "OBP-S") has been designed expressly to address the challenges of multidisciplinary research needed to answer the global challenges such as climate change and others (Pearlman *et al*., 2017). This solution (see Fig. 1) covers the

entire range of ocean observations including data management and user support and draws on the developing fields of natural language processing and ocean vocabularies (see companion paper on the OBP repository). But more than a technology implementation is needed for effective community engagement. Thus, the OBP-S will provide mechanisms for community dialogues and to facilitate publishing BPs. One aspect of this effort is the recently created Research Topic in Frontiers of Marine Science (https://www.frontiersin.org/research-topics/7173/best-practices-in-ocean-observing), which will become a place of commentary and dialog. Peer review of best practice documents is a means of promoting community adoption and providing increased

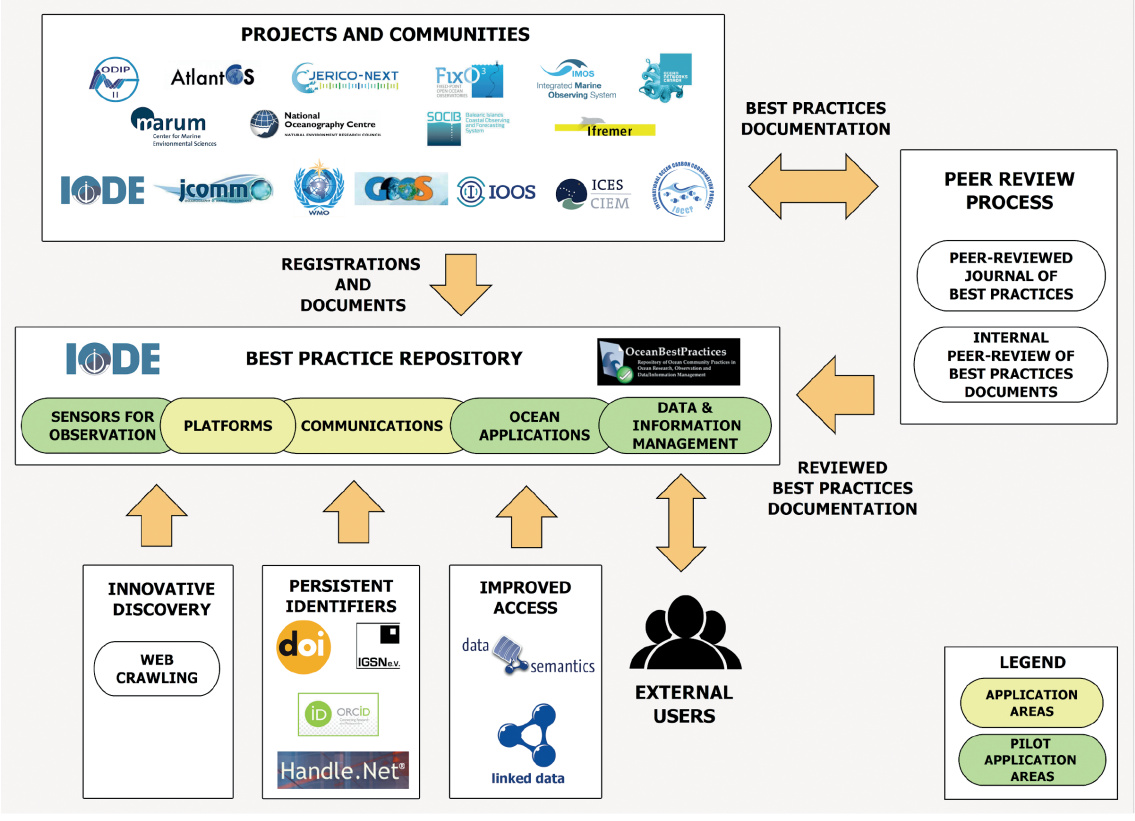


Fig. 1 - OBP-R end-to-end process model.

visibility of methodologies. It also has significant benefits for those in universities and elsewhere that use number and quality of publications as a metric for advancement. Working together with Frontiers, the Research Topic offers this as the media to describe and disseminate robust and high-quality methodologies and interoperability, linked and referenced to the OBP repository documents as appropriate.

As the System depends on the quality of documents provided by its stakeholders, the project will pay substantial attention to community building. Thus, OBP-S includes (1) open access from central repository, (2) multiple document locations are possible, (3) the BP developer retains control of the content, and (4) journal articles that are methods papers describing implementation

and use of corresponding entries of full document in OBP-Repository. These are important, but they do not stand alone for community building. Training and capacity building are also an essential elements for BP adoption. For OBP-S, this will be done working closely with established organizations such as IODE Ocean Teacher, POGO and the SCOR Committee on Capacity Building as well as other activities such as the summer schools run by IMBER, SOLAS, GEOTRACES, etc. One of the main established platforms for providing training is the Ocean Teacher Global Academy (OTGA). OTGA aims at building equitable capacity related to ocean data and information management (as well as other IOC related disciplines) in all IOC Member States. This will ensure training and access to best practices from a global perspective and also get feedback as to the appropriateness of particular documented methodologies in the OBP-R. To further this direction, access to web based tools including information via mobile devices will be available. An advantage of this approach is that people have these as tools when they get back from training and can continue using the information and experience of the training.

References

Pearlman J., Buttigieg P.L., Simpson P., Muñoz Mas C., Heslop E., and Hermes J. 2017. *Accessing Existing and Emerging Best Practices for Ocean Observation - a new approach for end-to-end management of Best Practices.* IEEE/MTS OCEANS'17 Proceedings – Anchorage, 2017, pp 1-7.

Generating ocean climatologies from in situ observations

Alexander Barth, Université de Liège (Belgium), a.barth@uliege.be
Charles Troupin, Université de Liège (Belgium), c.troupin@uliege.be
Sylvain Watelet, Université de Liège (Belgium), swatelet@uliege.be
Aida Alvera-Azcárate, Université de Liège (Belgium), a.alvera@uliege.be
Jean-Marie Beckers, Université de Liège (Belgium), jm.beckers@uliege.be

Methodology

DIVAnd (Data-Interpolating Variational Analysis - n-dimensional) is an analysis tool that allows one to generate gridded data products from in situ observations. DIVAnd extends the 2D capabilities of the DIVA tool and allows the interpolation of observations on curvilinear orthogonal grids in an arbitrary high dimensional space by minimizing a cost function. This cost function penalizes the deviation from the observations, the deviation from a first guess and abruptly varying fields based on a given correlation length (potentially varying in space and time). Physical constraints can be added to this cost function such as an advection constraint, diffusion or source terms. One major advantage of the method is that it naturally decouples basins that are not connected and where water masses often have very different properties.

While the classical DIVA was written in Fortran and called via shell scripts, DIVAnd was rewritten from scratch using generalized mathematical formulation, in the programming language Julia which allows a high-level programming style but it compiled to machine code for better performance. The quality of the analysis depends strongly on the two main analysis parameters:

1. the correlation length, which represents over which space-time distance two location are significantly correlated) and
2. the observational error variance, i.e. how accurate and representative the observations are. Several tools have been written to estimate the correlation and observational error variance to guide the user to choose suitable values for these parameters.

Observational data

DIVAnd can load observations in NetCDF or ODV spreadsheet formats. There are also functions implemented to directly query online the following databases: World Ocean Database, Copernicus Marine Environment Monitoring Service (CMEMS) and EMODnet Physics. The bathymetries, needed to delimit the domains of interpolation, based on GEBCO and EMODnet Bathymetry have been adapted to work with DIVAnd.

Additional tools around DIVAnd have been developed for automatic outlier detection and duplicates check. The outlier detection is based on a DIVAnd analysis to evaluate how consistent a data value is with other data values in its surroundings. The duplicated check implements an n-dimension Quadtree to quickly searches for observations nearby a given location.

Applications

An outlook on new developpements of DIVAnd will also be given, including the combination of DIVAnd with non-linear analysis techniques from machine learning (neural networks). New practical application in the context of SeaDataCloud for surface currents high-frequency radar data, EMODnet Physics data and EMODnet Biology data with initial test of DIVAnd coupled with a neural networks will be given.

Using Jupyter Notebooks as a data scientist "Work bench" for quality assurance of data processing flows and quality control of data series

Rob Thomas, Marine Institute (Ireland), rob.thomas@marine.ie
Sarah Flynn, Marine Institute (Ireland), sarah.flynn@marine.ie
Will Meaney, Marine Institute (Ireland), will.meaney@marine.ie
Siobhan Moran, Marine Institute (Ireland), siobhan.moran@marine.ie
Ramona Carr, Marine Institute (Ireland), ramona.carr@marine.ie
Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie

Abstract

Marine data are increasingly being utilised for purposes or used by organisations beyond those intended at collection. In determining appropriate use, quality assurance is given greater importance to provide visibility of the processes used in generation, QC, curation and serving of data. This presentation describes the process of creating and implementing a Data Quality Framework in a cross-discipline organisation, showcases examples and benefits of using Jupyter notebooks as a data scientist workbench, also demonstrating the implementation of open source tools to provide an improved experience for scientists involved with data QC.

Background

The Marine Institute (MI) has published its Strategy for 2018-2021 and the data it produces and manages are clearly identified as a "Strategic Enabler" for the Strategic Focus Areas. Supporting the role of data as a Strategic Enabler in the MI Strategy, the MI Data Strategy is being implemented with the vision that "Irish marine data will underpin the development of Ireland's marine sectors and the sustainable development of Ireland's marine resource". Quality is one component in achieving this vision.

In order to meet the Quality element of the MI Data Strategy, a Working Group has been tasked with drafting and then implementing a Quality Management Framework for the environmental data collected by the MI. The IODE has recently introduced training to support data centres producing and implementing a Quality Framework. As a member of the IODE NODC network official recognition of the MI for its data management practices can be achieved by achieving accredited NODC status with the IODE. This then confers accreditation by the World Data System (WDS) through IODE's membership of the WDS.

The MI has chosen to follow the IODE Quality Framework guidelines, with the structure following the ISO9001:2015 standard, and is documenting its processes through process flows to describe the high-level "what" a processes does and documenting standard operating procedures (SOPs) to describe "how" a process should be carried out.

For the MI Oceanographic Sciences section Jupyter Notebooks have become a key tool in the documentation of scripted workflows and are being used to implement a "Data Scientist

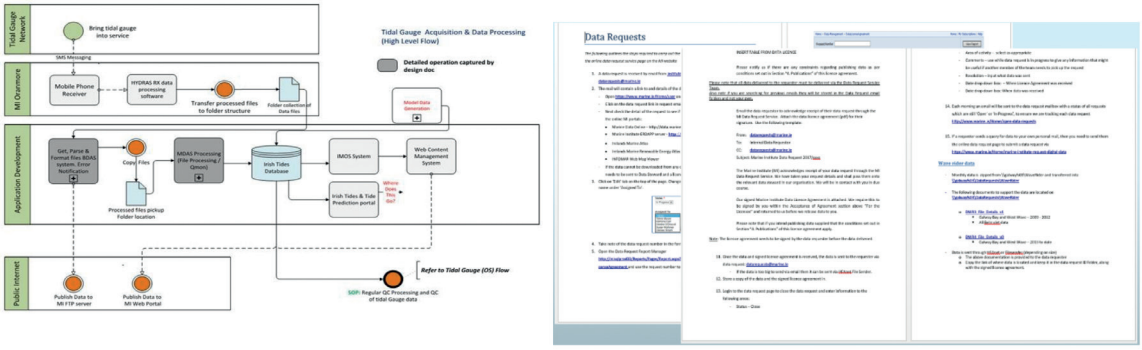


Fig. 1 - Examples of a process flow documented in the Business Process Management Notation and documented SOP.

Workbench" for the different data types handled by the section. The notebooks provide an interactive tool that allow users with limited scripting experience to work through a defined process or analysis with a consistent approach, while capturing the expert judgement decisions that need to be made on a case by case basis. These notebooks allow reproducibility with the flexibility to provide a mechanism to document different runs of a process (e.g. underway data processing documentation on a cruise by cruise basis). In this way they provide a great option when a process requires inputs that prevent the end-to-end process being automated.

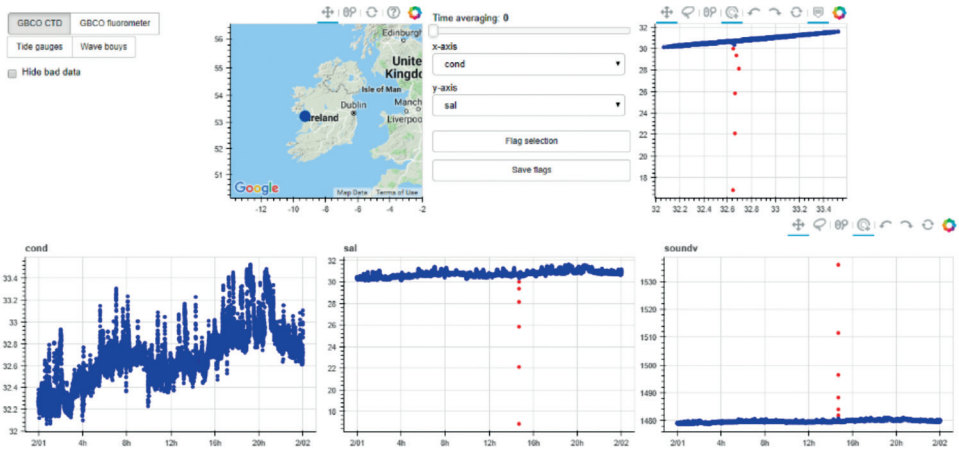


Fig. 2 - Example of data screening GUI implemented using interactive visualisation toolbox.

By utilising open-source toolboxes available for use with Jupyter notebooks, data visualisation can be incorporated as a seamless part of the workflow and using interactive plotting toolboxes (e.g. Bokeh for Python) enables screening of data streams through a GUI. One advantage of such tools is that dynamic manipulation of high resolution data streams allows a scientist to see the impact of screening/flagging choices on binning of data by downstream users.

In providing clear processes, procedures and tools in support of a Quality Framework data this supports MI staff in meeting the Quality element of the MI Data Strategy and recognises the value of MI data as a strategic enabler in its future strategy. IODE accreditation also provides an external acknowledgement of the quality assurance for data collected and curated by the Marine Institute.

SeaDataCloud temperature and salinity data collections

Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it
Christine Coatanoan, Institut Français de Recherche pour l'Exploitation de la Mer (France), christine.coatanoan@ifremer.fr
Volodymyr Myroshnychenko, Middle East Technical University (Turkey), volodymyr@ims.metu.edu.tr
Örjan Bäck, Sveriges Meteorologiska och Hydrologiska Institut (Sweden), orjan.back@smhi.se
Helge Sagen, Institute of Marine Research (Norway), helge@hi.no
Serge Scory, Royal Belgian Institute of Natural Sciences (Belgium), sscory@naturalsciences.be
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de
Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), michele.fichaut@ifremer.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl

Two versions of temperature and salinity historical data collections for each European marginal sea (Arctic Sea, Baltic Sea, Black Sea, North Sea, North Atlantic Ocean, and Mediterranean Sea) have been published within the framework of SeaDataNet2 Project. They represent a snapshot of the SeaDataNet database content at two different times: V1.1 (Jan 2014, Simoncelli *et al*., 2014) and V2 (Mar 2015, Simoncelli *et al*., 2015 and 2016). A Quality Control Strategy (QCS) was developed in SeaDataNet2 and continuously refined in order to improve the quality of the data and create the best data products. The QCS iterative approach facilitates the upgrade of the data and it allows a versioning of data products.

A newer version of temperature and salinity historical data collections has been released within SeaDataCloud Project in June 2018. The objective of this presentation is to briefly overview the existing SeaDataNet products and to present the first release of SeaDataCloud temperature and salinity historical data collections (SDC_DATA_TS_V1), spanning the time period 1900-2017, their characteristics in terms of space-time data distribution and their usability. A particular focus will be dedicated to the Mediterranean Sea collection.

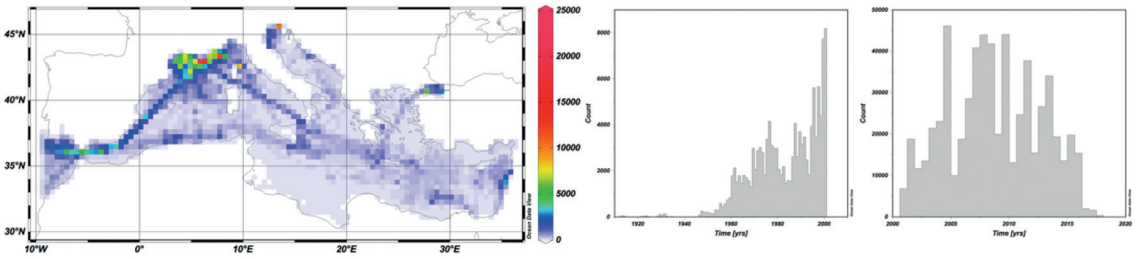


Fig. 1 - Results of the QC analysis for the Mediterranean Sea data set (SDC_MED_DATA_TS_V1): (left) data density map; (middle) temporal data distribution 1900-2000; (right) temporal data distribution 2001-2017.

Temperature and Salinity data sets were analyzed at regional level to assess and report on their quality. A common basic QC analysis was performed using ODV software (5.0.0) and following common QC guidelines. Product Information Documents (PIDocs) contain all specifications about the general products' characteristics (space-time coverage, resolution, format, usability) and quality (validation methodology and results). Fig. 1 shows an example in the Mediterranean Sea of data density map and time distribution histogram produced per each European basin. Fig. 2 is an example of the scatter diagrams produced per each region and contained in the PIDocs.

Statistics about the SeaDataNet infrastructure population in terms of temperature and salinity data per sea basin show a progressive increase of available data. Data quality also improved thanks to the introduction of additional checks by regional experts, exploiting the complete metadata description. The statistics about the quality flags after the quality assessment presents very high percentages of good (QF=1) or probably good data (QF=2): ~99% for the Mediterranean Sea; 98-99% for the Black Sea; ~99% Arctic Sea; ~99% Baltic Sea; 98-99% for the North Sea and 96(S)-99% for the North Atlantic Ocean. In fact, the analysis could be performed by instrument type to verify the data set completeness and consistency, and per data originator to identify systematic data anomalies. The derived metadata statistics per sea basin allow monitoring the European data

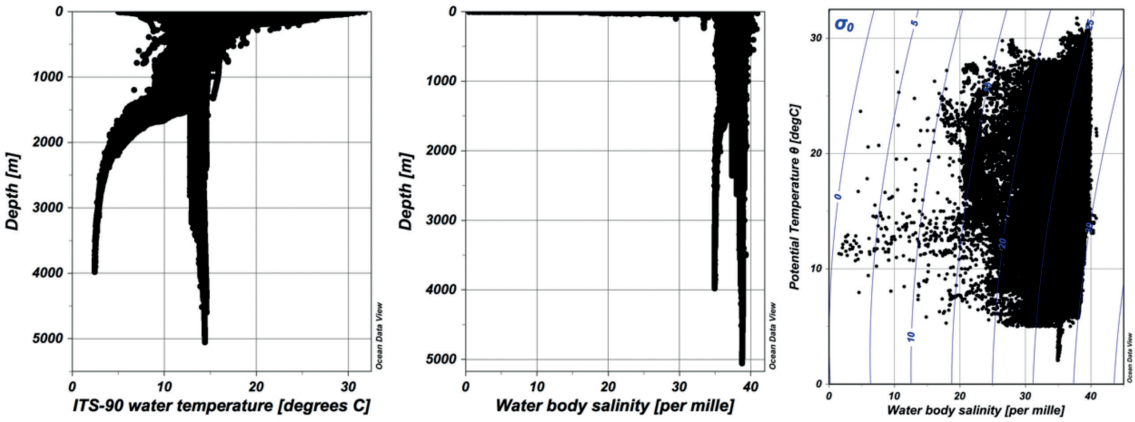


Fig. 2 - Temperature (left), salinity (middle) and TS (right) scatter plots of good (Quality Flag 1) and probably good (Quality Flag 2) data after Quality Check analysis for the Mediterranean Sea data set.

sharing landscape per sea basin and the advent of new sensors, which require particular efforts in data management and quality assessment.

Conclusions and Developments

All SeaDataCloud products are available as ODV collections through a web catalog (https://www.seadatanet.org/Products) together with their associated Digital Object Identifier (DOI) and Product Information Document (PIDoc) containing the specifications about product's generation, quality assessment and technical details to facilitate users' uptake.

The progressive automation of the QCS in the SeaDataCloud Virtual Research Environment will speed up the basic quality check process of the data and further improve the quality of the

SeaDataNet infrastructure content and the derived products, which could be delivered with a regular time schedule.

References

Simoncelli S., Tonani M., Grandi A., Coatanoan C., Myroshnychenko V.r, Sagen H., Bäck Ö., Scory S., Schlitzer R., Fichaut M. (2014). *First Release of the SeaDataNet Aggregated Data Sets Products.* WP10 Second Year Report - DELIVERABLE D10.2. http://doi.org/10.13155/49827

Simoncelli S., Coatanoan C., Myroshnychenko V., Sagen H., Bäck Ö., Scory S., Grandi A., Schlitzer R., Fichaut M. (2015). *Second release of the SeaDataNet aggregated data sets products.* WP10 Fourth Year Report - DELIVERABLE D10.4. http://doi.org/10.13155/50382

Simoncelli S., Coatanoan C., Back O., Sagen H., Scory S., Myroshnychenko V., Schaap D., Schlitzer R., Iona S., Fichaut M. (2016). *The SeaDataNet data products: regional temperature and salinity historical data collections.* EGU 2016 - European Geosciences Union General Assembly 2016. 17-22 April 2016, Austria.

HELCOM map and data service - system for making available assessment data products and underlying data for various use cases

Joni Kaitaranta, Helsinki Commission Secretariat, joni.kaitaranta@helcom.fi
Andžej Miloš, Helsinki Commission Secretariat, andzej.milos@helcom.fi

Introduction

HELCOM (Baltic Marine Environment Protection Commission - Helsinki Commission) is the governing body of the Convention on the Protection of the Marine Environment of the Baltic Sea Area, known as the Helsinki Convention. The Contracting Parties are Denmark, Estonia, the European Union, Finland, Germany, Latvia, Lithuania, Poland, Russia and Sweden. HELCOM works to protect the Baltic Sea from all sources of pollutions through intergovernmental cooperation. HELCOM is a regional sea convention (RSC) for the Baltic Sea and policy maker for the Baltic Sea area by developing common environmental objectives and actions.

One of HELCOM's task is to act as an environmental focal point providing information about the state of and trends in the marine environment, the efficiency of measures to protect it and common initiatives and positions, which can form the basis for decision-making in other international fora. This task requires dedicated tools for providing information and data used in assessment carried out by HELCOM. HELCOM Map and Data service (HELCOM MADS) has been desigined to fulfill that requirement.

This presentation aims at providing information on:

- Requirements and use cases for making available and distributing reginal data products and underlying data
- Technical solutions used for HELCOM MADS, developed with support by EU co-funded project TAPAS.

Requirements and use cases

Fundamental requirement for tools used in making available data stems from HELCOM Monitoring and Assessment Strategy, which contains attachment of Data and Information Strategy. The strategy (which is currently undergoing update) sets out following giding principle: The HELCOM data and information activities should facilitate access of the general public to environmental information.

This includes requirements for various thematic datasets ranging from in-situ monitoring of eutrophication, biodiversity and hazardous substances related data to more aggregated data products on status of marine environment, nutrient loading and trends and overview of maritime traffic related activities.

HELCOM Contracting Parties that are also EU Member State have obligation stemming from Marine Strategy Frame Work Directive (MSFD). The Directive requires coherence and

coordination within regions and makes possible reference to assessements/data that is done by Regional Sea Conventions. This sets requirements in relation to data products being made available on regional level and according to INSPIRE Directive.

Maritime Spatial Planning (MSP) requires also regional coordination and coherence of plans across the region and between neighboring countries. According to the MSP Directive, plans should be harmonized across the region and this requires regional level work and agreement on data harmonization. Spatial data should be also spatially very accurate and up-to-date to be useful in the planning process.

According to HELCOM data and information strategy, the target group for information and data products should be also general public / decision makers. This sets a requirement on complexity of information and terms used for displaying results. This purpose sets a requirement to be able to pinpoint to a specific aggregated and simplified map product in the service.

Technical solutions

HELCOM MADS is based on two user interfaces, which are seamlessly intelinked:

- Map and Data service map viewer part (ESRI ArcGIS Server and ArcGIS API for Javascript) http://maps.helcom.fi/website/mapservice/index.html
  - Viewing a dataset and feaure attributes on map viewer;
  - Searching dataset in table of contents;
  - Viewing external datasets (WMS);
  - Data service addresses (ArcGIS Rest / OGC WMS);
  - Accessing datasets and features by URL parameters;
- Metadata catalogue part (Geonetwork); http://metadata.helcom.fi/geonetwork/srv/eng/catalog.search#/home;
  - Viewing, downloading and searching metadata records (INSPIRE compliant).

Each dataset has a unique ID, that is used to link a record in the metadata catalogue with a dataset in the map viewer. Linkage is done by including ID in the map viewer and metadata catalogue URL. At anytime user can swich between two systems: to view dataset in the map viewer or read metadata and download dataset in the metadata catalogue (Fig. 1).

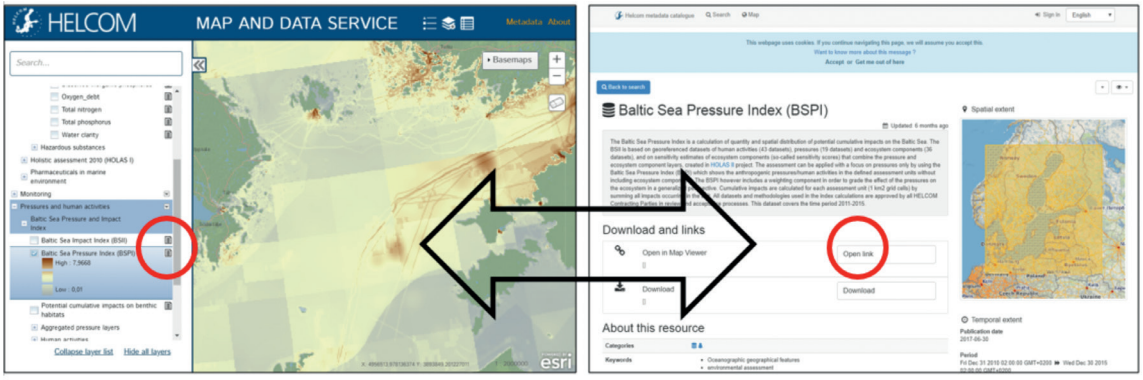


Fig. 1 - Visualization of a dataset in map viewer part and metadata record in metadata catalogue.

Map and Data service viewer is based on datasets stored in file Geodatabases and published in ArcGIS Server. and Uuser interface is developed by HELCOM Secretariat using ArcGIS API for Javascript. The main features of map viewer are: searchable table of contents, linkage between to metadata catalogue, accessing datasets and features via URL, feature identification, attribute table widget.

HELCOM Metadata catalogue is application of Geonetwork and utilizing INSPIRE compatible metadata template with GEMET thesauri.

Source code of HELCOM MADS can be found at GitHub under GNU General Public License v3.0: https://github.com/helcomsecretariat/MADS

Error characterization of sea surface salinity products using triple collocation analysis

Nina Hoareau, ICM Consejo Superior de Investigaciones Científicas (Spain), nhoareau@icm.csic.es
Marcos Portabella, ICM Consejo Superior de Investigaciones Científicas (Spain), portabella@icm.csic.es
Wenming Lin, Nanjing University of Information Science and Technology (China), wenminglin@nuist.edu.cn
Joaquim Ballabrera-Poy, ICM Consejo Superior de Investigaciones Científicas (Spain), joaquim@icm.csic.es
Antonio Turiel, ICM Consejo Superior de Investigaciones Científicas (Spain), turiel@icm.csic.es

Abstract

The monitoring of the global distribution of sea surface salinity (SSS) is vital to understand the ocean's role in the Earth's climate. Until the advent of the spaceborne L-band radiometers, SSS observations were mainly acquired by *in-situ* sensors (moored buoys, drifters, and thermosalinographs). As a result, knowledge of the spatial and temporal variability of salinity has been scarce due to the lack of a comprehensive set of salinity observations. While *in-situ* data (e.g., Argo floats) were being used in a growing number of studies, numerical models were also widely used as a complementary source of such information. The spatiotemporal resolution achieved by satellite salinity measurements has no equivalent among the other existing salinity observations systems. Since the launch of the Soil Moisture and Ocean Salinity (SMOS) mission (2009) and then the Aquarius mission (2001), more than seven years of satellite-derived SSS data, with a spatial and temporal resolution adequate for climate and ocean general circulation studies, have become available. The L-band radiometers onboard SMOS and Aquarius have proven to be challenging and various spatial and temporal averaging and data fusion techniques have been implemented to better recover structured and meaningful geophysical information from remote sensing SSS retrievals. A comprehensive validation is therefore essential to characterize the information provided by the different salinity products.

Table 1 - Representativenesse error (r^2) for the following triplets: TAO-AV4-WOA09 (TA09) and TAO-SOA-WOA09 (TS09) denoted as T-09; TAO-AV4-WOA13 (TA13) and TAO-SMOS-WOA13 (TS13) denoted as T-13; and TAO-GLORYS2V3-AV4 (TGA) and TAO-GLORYS2V3-SOA (TGS) denoted as TG-; and TAO-AV4-SOA.

	TA09 (TS09)	TA13 (TS13)	TGA (TGS)	TAS
Representativeness error (r^2)	0.034 (0.020)	0.027 (0.011)	0.009 (0.023)	0.015

The triple collocation (TC) technique allows the simultaneous calibration of three independent, collocated data sources, while providing an estimate of their accuracy. Here, the TC is adapted to validate different salinity data products along the tropical band. The representativeness error

(i.e., the true variance resolved by the relatively high-resolution systems but not by the relatively low-resolution system) is accounted for in the validation process. A method based on the intercalibration capabilities of TC is used to estimate the representativeness error for each triplet. Such error (see Table 1) is found to impact between 15% and 50% the total error estimation of the different products (see Tables 2 and 3). The method also sorts the different products in terms of their resolving spatiotemporal scales. This calibration study is limited to the year 2013, a year when all the products are available.

The six salinity products used in this study are: the in situ data from the Global Tropical Moored Buoy Array (TAO) available at https://www.pmel.noaa.gov/; the GLORYS2V3 ocean reanalysis output (from NEMO version 3.1) provided by Copernicus; the climatology maps provided by the World Ocean Atlas 2009 (WOA09) and 2013 (WOA13); and the satellite-derived SSS maps from Aquarius Level 3 version 4 (AV4) available at ftp://podaac-ftp.jpl.nasa. gov/allData/aquarius/L3/mapped/ and the Soil Moisture and Ocean Salinity (SMOS) objectively analyzed (SOA). The SOA product correspond to the Objective Analysis product distributed by the Barcelona Expert Center (http://bec.icm.csic.es/).

Table 2 - Estimated Standard deviation errors of the different salinity measurements at AV4 scales.

Errors at AV4 scale						
SSS measurement	TAO	GLORYS2V3	AV4	SOA	WOA13	WOA09
Standard deviation error	0.178	0.175	0.167	0.244	0.290	0.314

Table 3 - Estimated Standard deviation errors of the different salinity measurements at SOA scales.

Errors at SOA scale						
SSS measurement	TAO	GLORYS2V3	AV4	SOA	WOA13	WOA09
Standard deviation error	0.218	0.208	0.211	0.204	0.257	0.293

This validation approach aims to assess the quality of the different salinity products at the satellite-resolved spatiotemporal scales, i.e., those of AV4 (see Table 2) and SOA (see Table 3). The results show that the resolved scales of SOA are coarser than those of AV4. At the AV4 resolved scales, the Aquarius product (AV4) has an error of 0.17, and outperforms TAO, GLORYS2V3, and the SOA maps. However, at SOA resolved scales, the SMOS product has an error of 0.20, slightly lower than that of GLORYS2V3, AV4 and TAO. The WOA products show the highest errors. High order calibration may lead to a more accurate assessment of the quality of the climatological products.

References

N. Hoareau, M. Portabella, W. Lin, J. Ballabreray-Poy and A. Turiel, *"Error Characterization of Sea Surface Salinity Products Using Triple Collocation Analysis"*, IEEE Trans. Geoscien. Remot. Sens., in press, 2018, doi:10.1109/TGRS.2018.2810442

IBISAR: skill assessment service for real-time ranking of met-ocean data products in the IBI area for emergency and SAR operators

Emma Reyes, Sistema d'observació i predicció costaner de les Illes Balears (Spain), ereyes@socib.es
Baptiste Mourre, Sistema d'observació i predicció costaner de les Illes Balears (Spain), bmourre@socib.es
Paz Rotllán, Sistema d'observació i predicció costaner de les Illes Balears (Spain), protllan@socib.es
Ismael Hernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), ihernandez@socib.es
Eric Comerma, RPS Ocean Science (United States of America), eric.comerma@rpsgroup.com
Tayebeh Tajalli Bakhsh, RPS Ocean Science (United States of America), Tayebeh.TajalliBakhsh@rpsgroup.com
Anna Rubio, AZTI Marine Research (Spain), arubio@azti.es
Julien Mader, AZTI Marine Research (Spain), jmader@azti.es
Luis Ferrer, AZTI Marine Research (Spain), lferrer@azti.es
Christian De Lera Fernandez, Salvamento Maritímo (Spain), christiandlf@centrojovellanos.es
Enrique Álvarez-Fanjul, Puertos del Estado (Spain), enrique@puertos.es
Alejandro Orfila, Instituto Mediterráneo de Estudios Avanzados (Spain), aorfila@imedea.uib-csic.es
Joaquín Tintoré, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jtintore@socib.es

To support marine response and crisis management, a set of accurate near real-time (NRT) operational data and model forecasts is crucial. Search & Rescue (SAR) and environmental risk modelling applications are mostly based on the Lagrangian discrete particle algorithms. They rely on appropriate currents and wind forecasts to accurately predict oil's future drift or trajectories of drifting targets to determine an optimal search region. With multiple ocean forecasting datasets now available, the growing strategy is to ensure the availability of NRT datasets in a specific region, and to use all of them to develop a consensus of opinions from the various outcomes on what might occur.

If different forecast datasets result in disparate trajectories, there are multiple practical outcomes, and a low level of confidence in each of the predictions. In such context, the key overarching concern of many targeted users is the impact of inaccurate data on decision-making, since it can result in erroneous estimation of the necessary resources and even in some extreme cases in lost lives. Moreover, the increasing complexity of high-resolution ocean models may require Skill Assessment (SA) methods that quantify the model performance, by providing easily interpretable and understandable metrics.

The other part of the answer lies in providing a confidence indicator of the forecast in a systematic and long-term routine manner, through user-friendly tools. The IBISAR SA service, built on an existing operational service, consists of an automated process that first simulates the trajectory of a particle for the available met-ocean datasets. Then, the process evaluates

the model performance, by comparing pairs of observed and predicted trajectories to provide a dimensionless skill score metric. Surface current data from multi-platform observing systems (e.g. drifting buoys, satellite-derived observations, High-Frequency Radar -HFR- and moorings) are used to help estimate errors in NRT in the forecast outcomes, using Eulerian or Lagrangian approaches.

IBISAR is a restricted service available then only for registered users. A link to the IBISAR private service will be available on the dedicated webpage (draft as shown in Fig. 1). Subscribers (targeted users) will be able to easily access the met-ocean data for downloading, visualizing and comparing different data products, assessing the quality (SA) of different model predictions and HFR performance in comparison with real-time or historical measurements. The service focuses on the IBI (Iberian-Biscay-Ireland) region, supporting a myriad of socio-economically important

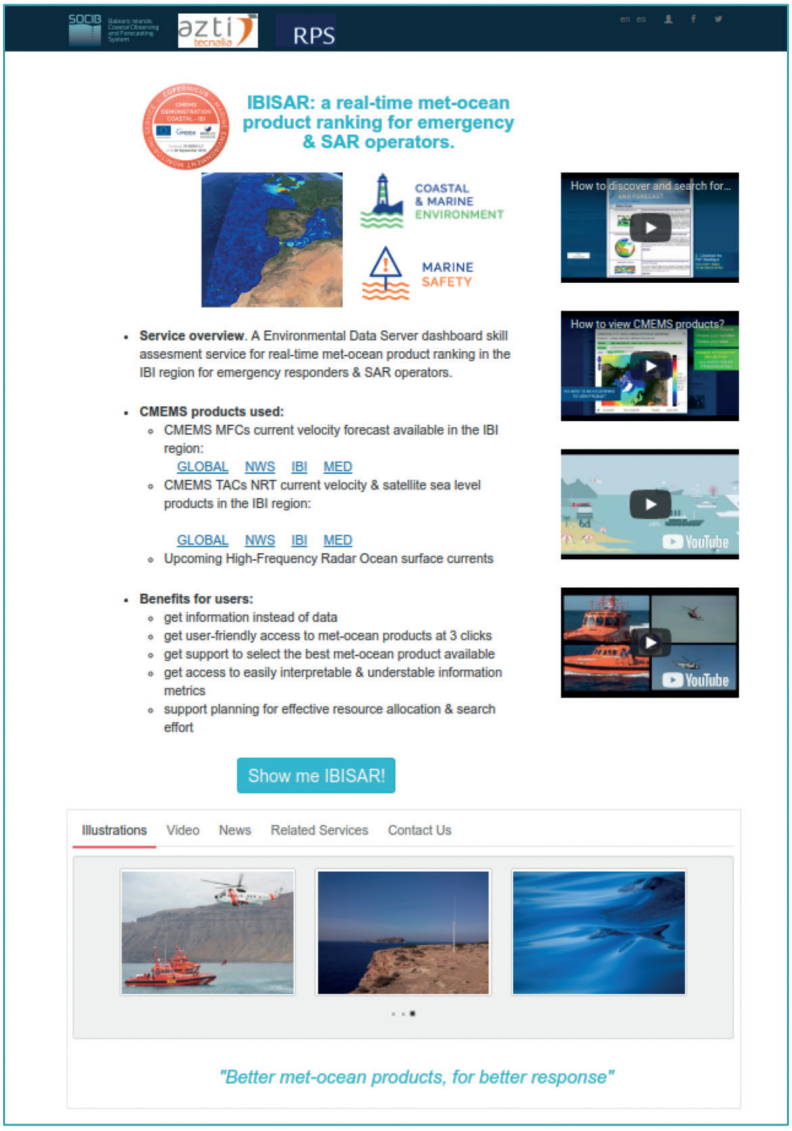


Fig. 1 - Draft of the dedicated webpage for the demonstration of the IBISAR downstream service.

activities including fisheries, oil and gas shipping, commercial ship traffic, coastal and marine environment management and protection, marine safety and energy production.

The improvement, validation and promotion of IBISAR service are being carried out by a Public-Private Partnership (PPP) between an advanced Marine Research Infrastructure and data provider (SOCIB, public sector), a Technological Centre with expertise on marine and food research (AZTI, private non-profit organization) and a Downstream Service Provider (RPS Ocean Science, private sector), under the umbrella of CMEMS (Copernicus Marine Environment Monitoring Service) User Uptake programme. Furthermore, the Spanish SAR Agency (one of the prime targeted users of this service) has been engaged with support of RPS Ocean Science. RPS has been providing added-value products and tailored services to this specific sector since more than a decade.

Knowing and understanding targeted users (e.g. SAR operators, emergency responders and other maritime safety, coastal and marine environment actors) needs, will enable IBISAR SA to serve as a decision-support tool to help guide users to choose the best available met-ocean product as input for their SAR and environmental risk modelling applications. Moreover, IBISAR service will help them to:

- Get information instead of data through a user-friendly, quick, and intuitive process.
- Get access to a dynamic and continually updated inventory of datasets.
- Get access to models assessments by means of easily interpretable metrics.
- Better support emergency management preparations at sea.
- Support immediate response and more secure search operations.
- Support planning for the most optimal search area.

Coastal data portals to support marine science and management - the coastMap approach

Linda Baldewein, Helmholtz-Zentrum Geesthacht (Germany), linda.baldewein@hzg.de
Ulrike Kleeberg, Helmholtz-Zentrum Geesthacht (Germany), ulrike.kleeberg@hzg.de
Marcus Lange, Helmholtz-Zentrum Geesthacht (Germany), marcus.lange@hzg.de
Dietmar Sauer, Helmholtz-Zentrum Geesthacht (Germany), dietmar.sauer@hzg.de

Managing ecosystems in the marine environment requires cooperation between the spheres of policy and science, and between the differing perspectives from the various scientific disciplines. Modern IT solutions allow for the combination of in situ measurements, laboratory-analyzed samples, interpolated point data and model outcomes. To support integrated advice for management on a regional sea basin scale, this union is essential.

The coastMap portal (www.coastmap.org) addresses a large variety of topics with a particular focus on their effect on the North Sea. Within the data portal, the combination of analyses and model data relevant to the conditions of the seafloor, the overlying water column and the atmosphere is used to explore issues, such as human impacts on coastal zones from shipping emissions, as well as habitat changes from wind farm constructions and marine pollutants. In 'Spotlights', these important research topics are clearly explained for the benefit of the interested public and policy makers using edited science writings that are supported by interactive maps, videos and links to further information.

Data products freely available to the public on the portal include campaign data, preprocessed maps and large-scale oceanographic and atmospheric models. coastMap offers applications to visualise and download field and laboratory work and to connect the information with interactive maps. Using filter functions, the user can search a relational campaign database for both general and specific topics. The broad data pertaining to a marine field of interest, for example, or the specific data pertaining to a selected ship campaign or of one of 1000 measured parameters can all be explored. The workflow for scientists who want to submit data has been simplified as to encourage high submission rates and quick response times. For example, during ship campaigns, researchers can run an application that collects the metadata for their samples on waterproof tablets.

Oftentimes, model output for the coastal regions is in the range of terabytes and is stored in a way that only experts can retrieve and further use the results. The Model Analysis Tool differs from this and uses a "Big Data" approach, giving non-modelling experts access to detailed and high-resolution oceanographic model data. An interface allows statistically examining and on the fly downloading subsets of model-derived data. Thus, the tool bridges the gap between modelling scientists and experts of other disciplines, who are required to use model results, but who, in general, lack the knowledge of how best to approach them.

Data stored and refined in coastMap may provide the evidence based knowledge for the creation of new models and may help scientists and laymen to understand ecosystem connectivity. Such tools, on a regional scale, support policymakers in their decisions toward fashioning a sustainable development of the coastal resources, and allows scientists access to existing regional data to develop and test new theories.

Marine base maps in coastal Norway: a case for developing sustainable blue growth in coastal communities

Minika Ekanem, Norwegian Hydrographic Service (Norway), minika.ekanem@gmail.com

Introduction

Today, there is no coordinated collection and publishing of marine base maps. The presentation demonstrates the stakeholder needs for marine base maps (the data products), the organisation of a pilot project and the potential socio economic benefits to be acheived as a basis for better coastal zone planning and management and sustainable blue growth in Norway.

Fig. 1 - Illustration of terrain modell of Søre Sunnmøre, Norway.

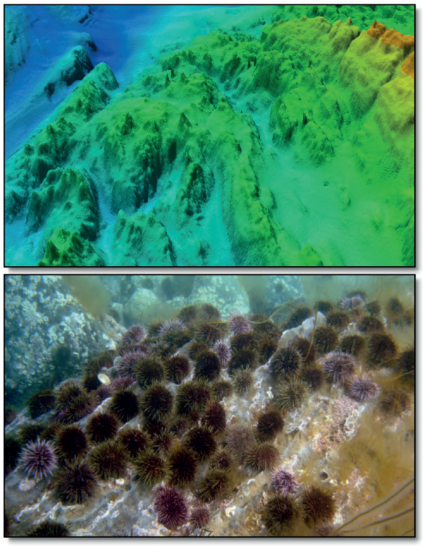


Fig. 2 - Marine Base Maps examples.

Marine Base Map Project

With today's survey processes, the ambition for complete coverage with modern data along the Norwegian coastline will take many years to achieve, hence the Marine Basemaps pilot project. The primary goals of the project are to collect knowledge of the seabed along the coast, produce standardized products and services and improve the infrastructure for free access to all stakeholders via the national geodata portal. The project will map the physical, geological, biological and chemical properties of the seabed and subsea environment starting from the coastline and up to one nautical mile beyond the baseline.

Pilot Project

3 pilot areas in different geographical areas are selected each with unique challenges and multiple stakeholder needs. New technology and methods for collecting data such as unmanned underwater vehicles will be tested to assess its potential capacity, profitability and suitability for use on a larger scale in a nationwide programme. It is intended that the Marine Base Maps project will become a national programme after the pilot phase. The presentation will highlight the current challenges in the pilot areas and the potential benefits to be obtained from the project.

Inter-Governmental

The pilot project will build better synergy between the 3 governmental project cooperation partners to create a seamless process from start to finish, the effect of which will be effectiveness and cost savings in each organisation. The presentation will highlight the plans for building synergy.

Fig. 3 - Testing new technology at Runde Environmental centre, Norway (June 2017).

Socio Economic Analysis of Marine Base Maps

A simplified socio-economic analysis of the proposed pilot areas concludes that marine base maps can lead to a productivity and revenue growth along these coastal communities and can potentially give a net present value of NOK 156million. The presentation will compare this against reports from other blue growth investments projects in Norway.

An update on Ireland's integrated Digital Ocean

Eoin O'Grady, Marine Institute (Ireland), eoin.ogrady@marine.ie
Adam Leadbetter, Marine Institute (Ireland), adamleadbetter@marine.ie

Introduction

A shared understanding of data through coherent acquisition, integration, analysis and delivery is a key component of aligning Information Systems and Information Technology, and is an essential enabler in adding value to data (Nagle and Sammon, 2017). Typically, the presentation and organisation of data is said to give rise to information and the integration of and conversations around information to wisdom (Rowley, 2016). In order to present marine data and integrate marine information from around Ireland's seas Ireland's Integrated Digital Ocean was introduced by Leadbetter, O'Grady and Burke (2016). Organisations contributing data and information to the Digital Ocean platform and portal include:

- Commissioners of Irish Lights
- Dublin Bay Biosphere
- Dublin City University
- Electricity Supply Board
- Environmental Protection Agency
- Failté Ireland
- INFOMAR seabed mapping programme
- Irish Underwater Council
- National Parks and Wildlife Service
- SmartBay Ireland
- Sustainable Energy Authority of Ireland
- University College Cork

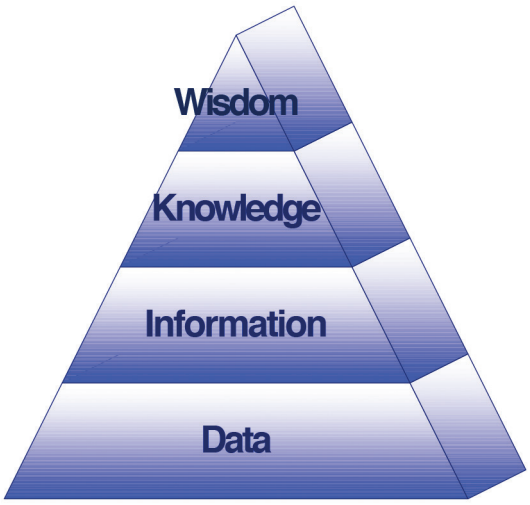


Fig. 1 - Data Information Knowledge-Wisdom pyramid (after Rowley, 2016).

Digital Ocean Platform Developments

Since the introduction of the Digital Ocean platform, a number of technical developments have been undertaken in order to increase the breadth of data in the portal. These have included the introduction of tile services, over Web Map Services, for INFOMAR sea bed mapping data to the Digital Ocean as the WMS layers proved to be non-performant for end-users. Modelled data from the Marine Institute's operational forecasts have been incorporated into the platform through extension of the Leaflet.js mapping toolbox's Time Dimension plugin[1], originally developed by the

[1] https://github.com/socib/Leaflet.TimeDimension.

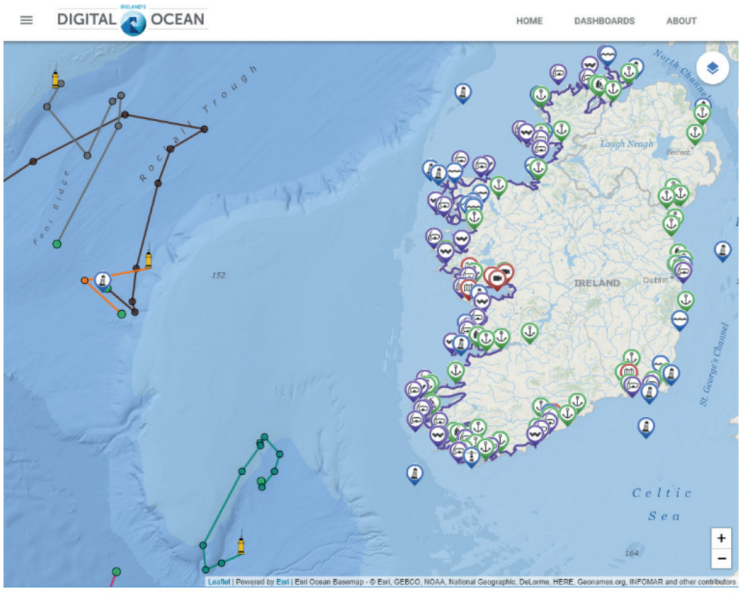


Fig. 2 - Ireland's Integrated Digital Ocean portal (http://www.digitalocean.ie) showing locations of live data feeds and static information points.

Balearic Islands Coastal Observing and Forecasting System. This extension has been contributed back to the community, and allows for fade between frames of a time based Web Map Service, in the case of the model data delivered from an instance of NOAA's Erddap data server. Finally, a recent overhaul of the portal's user interface has allowed a switch to a Material Design based look-and-feel (Fig. 2) which allows users to more intuitevely navigate the data supplied through the platform.

Beyond the web portal, technical developments have been focussed in three areas. First has been the definition of a software architecture which both acknowledges the Internet of Things research conducted in the Marine Institute around the Galway Bay Cable Observatory and offers a production ready, resilient framework. The architecture is also deployable at data nodes to feed into the Digital Ocean platform. Second has been the adaptiation of the GraphQL system[2], developed by Facebook as a query language for Application Programming Interfaces, to allow the querying of data published through the Digital Ocean platform, specifically from Erddap endpoints. Third, the API for the Digital Ocean platform has been used to provide data access to Mixed Reality applications on the Microsoft HoloLens platform (Fig. 3).

Fig. 3 - A 3 D model of a shipwreck, surveyed by the INFOMAR seabed mapping program, made available to a Microsoft HoloLens device from the Digital Ocean.

[2] https://graphql.org/.

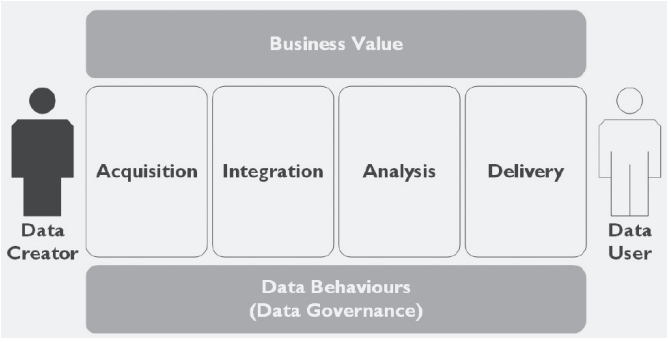


Fig. 4 - The Data Value Map (after Nagle and Sammon, 2017).

Conclusions

The development of the Digital Ocean platform has allowed the creation of value from marine data and information through integration of these assets in a single portal, and delivery of these data to end users; thus reducing the gap between data creators and data users (Fig. 4). The Digital Ocean platform has also allowed the use of emerging technological paradigms to create new visualisation techniques in marine informatics. Over the coming years, the sustainability and growth of the network providing data to the platform will be the key challenges.

References

Leadbetter, A., O'Grady, E., and Burke, N. (2016). *Ireland's Integrated Digital Ocean*. Bollettino di Geofisica teorica ed applicata 57 (supplement), 224-226.

Nagle, T. & Sammon, D. (2017). *"The Data Value Map: a framework for developing shared understanding on data initiatives"*. In Proceedings of the 25th European Conference on Information Systems (ECIS), Guimarães, Portugal, June 5-10.

Rowley, Jennifer (2007). *The wisdom hierarchy: representations of the DIKW hierarchy.* Journal of Information and Communication Science. 33 (2): 163–180. doi:10.1177/0165551506070706.

MARMOD database: Marmara Sea marine data

Devrim Tezcan, Middle East Technical University (Turkey), devrim@ims.metu.edu.tr
Volodymyr Myroshnychenko, Middle East Technical University (Turkey), volodymyr@ims.metu.edu.tr
Serap Kantarlı, Ministry of Environment and Urbanization (Turkey), serap.kantarli@csb.gov.tr
Hacer Selamoğlu Çağlayan, Ministry of Environment and Urbanization (Turkey), hacer.caglayan@csb.gov.tr
Hüsne Altıok, Istanbul University (Turkey), altiokh@istanbul.edu.tr
Barış Salihoğlu, Middle East Technical University (Turkey), baris@ims.metu.edu.tr

Marmara Sea

The Çanakkale Strait (Dardanelles), the Marmara Sea and the İstanbul Strait (Bosphorus) are the components of the Turkish Strait System (TSS). TSS is the only path for the water exchange between the Mediterranean and the Black Sea. Because of the density difference there is a two layer system in the TSS: the low salinity Black Sea waters flow on top, while the high salinity Mediterranean waters flow at bottom.

MARMOD Project: "An integrated modelling system for the Marmara Sea"

Ministry of Environment and Urbanization of Turkey, General Directorate of Environmental Impact Assessment Permits and Control, Laboratory, Measurement and Monitoring Department is officially responsible to coordinate the marine pollution monitoring projects in Turkish seas including the Marmara Sea.

In 2017 the Ministry of Environment and Urbanization has initiated the MARMOD project in order to perform assessment of the environmental and biogeochemical properties of the Marmara Sea. The project is coordinated by the Ministry of Environment and Urbanization and Institute of Marine Sciences of Middle East Technical University (IMS-METU). The main objective of the project is to apply a coupled hydrodynamic biogeochemical model to identify and predict effects of the Black Sea, urbanisation and industry on the environmental health of Marmara Sea ecosystem.

MARMOD Database

Setting up the hydrodynamic biogeochemical model requires comprehensive environmental data for model initialization, verification and parameterisation. For these purposes the MARMOD project compiled integrated database that includes data from all available sources. The main data supplier to the database is IMS-METU that has been doing investigation in the Marmara Sea and the straits since 1980s. Tens of cruises were performed, the physical, chemical and biological data were collected from about 7000 stations. The IMS-METU data were complemented with the data of environmental monitoring that is been implemented in the sea by the Ministry since 2000 with involvement of several subcontractors. The database contains data from more than 9000 stations

for period 1985-2017 on the following parameters: temperature, salinity, dissolved oxygen, fluorescense, turbidity, PAR, nutrients (NO2, NO3+NO2, NH4, PO4, Si, TN, TP), pH, Chl-a, and Secchi Disk Depth. The quality control of the data has been performed with the help of ODV (R. Schlitzer, Ocean Data View, 2017, available at http://odv.awi.de) following the SEADATANET procedures.

The station map is presented in Fig. 1.

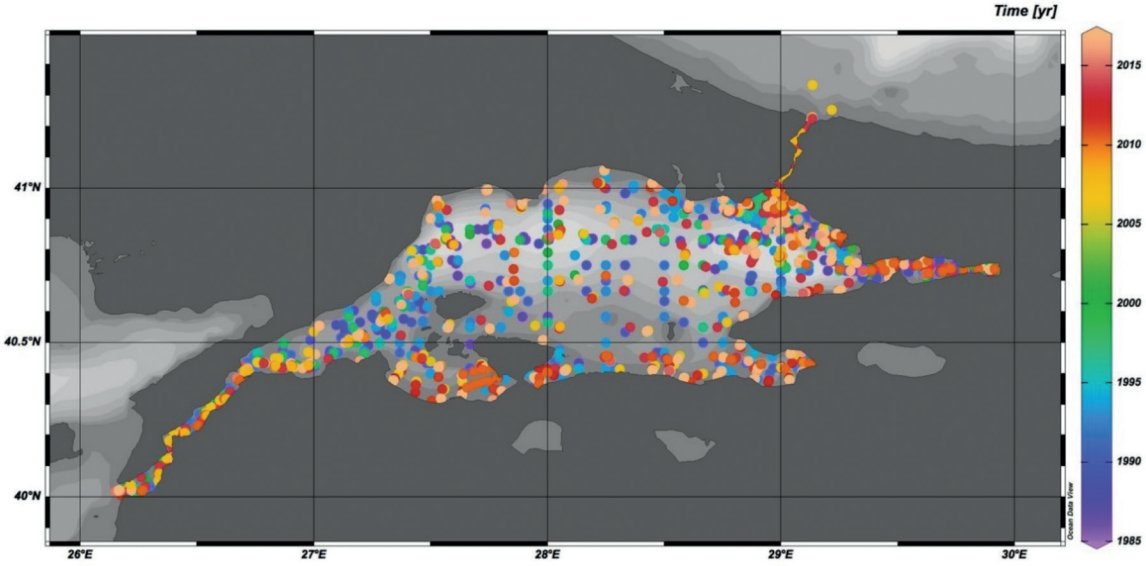


Fig. 1 - Station location map of MARMOD Database.

The most intensive observations were performed in the area of İstanbul strait, while other parts of the sea are less covered with data, moreover, the data are irregularly scattered over time. Although the database includes a huge amount of marine data, the more dense dataset on a regular grid is needed for biogeochemical models. For this reason four seasonal cruises are planned to be performed in the second phase of the MARMOD project (2018-2019). The data from the cruises will be included in the MARMOD Database as well as the coastal data that are collected periodically at the discharge point of the cities around the Marmara Sea.

Enlarging the EMODnet Chemistry focus with the EU marine litter data challenge

Matteo Vinci, Istituto Nazionale di Oceanografia e Geofisica Sperimentale (Italy), mvinci@inogs.it
Alessandra Giorgetti, Istituto Nazionale di Oceanografia e Geofisica Sperimentale (Italy), agiorgetti@inogs.it
Maria Eugenia Molina Jack, Istituto Nazionale di Oceanografia e Geofisica Sperimentale (Italy), mmolina@inogs.it
Alberto Brosich, Istituto Nazionale di Oceanografia e Geofisica Sperimentale (Italy), abrosich@inogs.it
Maria del Mar Chaves Montero, Istituto Nazionale di Oceanografia e Geofisica Sperimentale (Italy), mchavesmontero@inogs.it
Anna Maria Addamo, Joint Research Centre-European Commission (EU), anna.addamo@ec.europa.eu
Georg Hanke, Joint Research Centre-European Commission (EU), georg.hanke@ec.europa.eu
Francois Galgani, Institut Français de Recherche pour l'Exploitation de la Mer (France), Francois.Galgani@ifremer.fr

EMODnet

The European Marine Observation and Data Network as a long-term initiative of the European Commission is continuing its path to become a reliable pillar of the European Maritime policy for the development of the Blue-Green Economy.

EMODnet Chemistry

EMODnet Chemistry represents one of the eight thematic data portals, each one related to a different area of expertise. Started in 2009 with a pilot phase, it was born to ensure that European marine data would become easily accessible, interoperable, and without restrictions on use. From the outset, and also thanks to the positive evaluation of the pilot phase, it has been increasingly enlarged to cover a wider area and more environmental parameters. The main activities are collection, standardization, aggregation and dissemination of data and products related to nutrients and contaminants in three matrices: water, sediment and biota. The scope of EMODnet Chemistry has been extended recently to marine litter. Marine litter is a global concern that represents a thread for all the oceans and seas, it is recognized in the Marine Strategy Framework Directive (MSFD) as one of the descriptors (Descriptor 10) essential for achieving Good Environmental Status of EU marine waters and protecting the marine resources. Marine plastics and its reduction is the core of a political action through the new EU Plastics Strategy.

MSFD Descriptor 10 Marine Litter

MSFD Descriptor 10 considers several marine litter compartments, though the EMODnet Chemistry focus has been limited for the moment to beach litter, seafloor litter (i.e. for now on litter collected by bottom trawl fishing surveys), and micro plastic in the sea surface layer.

EU marine litter data management presents a complex and heterogeneous situation in which different levels of protocols and standards have been developed (e.g. regional, national, etc.). The plan of the consortium is either to adopt consolidated data formats when available, to adapt them when it is possible or to develop new solutions when needed.

The initial ambition was to develop a single database able to handle all the EU information from the three marine litter compartments. After the analysis of the available information, three different approaches have been proposed in order to develop a tailor-made standard procedure on a European scale, each based on the best available reference documents for each litter compartment:

- the beach litter approach is based on the OSPAR (Convention for the Protection of the Marine Environment of the North-East Atlantic) experience. The proposed data format provides the chance to describe in a detailed way the surveyed beaches, the surveys, the eventual animals harmed by litter entanglement or ingestion and the items found in the beaches. The items have been described following any of the three main litter reference lists available in EU: OSPAR, MSFD or UNEP/MAP (Barcelona Convention for the Protection of the Marine Environment and the Coastal Region of the Mediterranean);
- the bottom trawl litter approach follows the ICES DATRAS (International Council for the Exploration of the Sea Database of Trawl Surveys) experience. The data format provides an essential description of the bottom trawling survey and the collected litter items. The litter have been described using the three main litter reference lists available in EU: ICES, MSFD or MEDITS (International Bottom Trawl Survey in the Mediterranean);
- the floating marine micro litter approach adopts SeaDataNet formats that have been fitted after the comparison with the available information in Europe. CDI ISO XML format with little modifications is used to host micro litter metadata. Bio-ODV format has been adapted to include the information related to the surveys and the items collected in the water surface. For this scope a set of new standard oceanographic vocabularies has been minted.

The biggest challenge faced during the first phase has been dealing with the heterogeneity related to differences in formats, protocols and data quality. A huge effort has been made to handle a varied set of data in order to guarantee an efficient data management of the information. As a result, a series of basic products (including aggregated datasets and maps) has been released through the web site and the product visualization service. A more dynamic and tailored set of products is currently under discussion and development.

The relevance and actuality of the marine litter topic has favored strong interactions between the consortium and stakeholders or institutions dealing with marine litter data management (i.e.TG-ML MSFD Technical Group on Marine Litter, JRC European Commission Joint Research Centre, Regional Sea Conventions, Member States, EEA European Environmental Agency, etc.). These collaborations pursue the common goal of defining the best possible data management strategies and identifying the most valuable products. One outcome from the assessment of the available EU marine litter data is the contribution to the ongoing revision of the MSFD TGML Master List of Litter Item Categories Items with the aim of improving the data quality in future surveys. In addition, the first aggregated maps of marine seafloor litter have been used to support a preliminary assessment of single use plastics and fishing gear impact. These types of interactions demonstrate the necessity and utility of a consistent, harmonized marine litter dataset, and encourage the continuation of the work.

How to encourage new marine data users and providers through communication and outreach? The examples of EMODnet and the European Atlas of the Seas

Andrée-Anne Marsan, European Marine Observation and Data network (Belgium), andree-anne.marsan@emodnet.eu

Jan-Bart Calewaert, European Marine Observation and Data network (Belgium), janbart.calewaert@emodnet.eu

Oonagh McMeel, European Marine Observation and Data network (Belgium), oonagh.mcmeel@emodnet.eu

At the onset of its third development phase (2017-2020), the European Marine Observation and Data Network (EMODnet) had reached a maturity level where it could provide real value for industry and the public sector, including Member States.

In line with this, the focus of its communication strategy is therefore to increase communication and dissemination efforts to encourage new users to exploit this immense resource of marine data, products and services. EMODnet sits in an increasingly complex European marine observation data and information sharing landscape. This complexity is confusing, off-putting and limits both the visibility and the usability of EMODnet.

Through four examples, this abstract will illustrate how EMODnet has overcome this limit in clearly defining and communicating what its unique role is, what services and resources it provides as well as its economic and societal relevance. This abstracts will also cover the promotion of the European Atlas of the seas, a tool developed under the coordination of EMODnet since October 2017.

An updated visual identity

The EMODnet Website has undergone minor graphical and structural improvements since its inception in 2013. A major visual and structural improvement has been carried out in the course of 2017 to provide visitors with a visually attractive, state-of-the-art, and user-friendly interface.

This revamp of the Central Portal also provided the basis for a coherent EMODnet visual identity permeating through all sub-portals and delivering a user-oriented focus in line with the mature and operational stage EMODnet had reached. The new visual guideline allowed EMODnet to become a more recognizable brand for users worldwide.

In addition, a general online survey is implemented on each portal once a year to gather user's recommendations for improvements. Moreover, the friendliness of the portals is assessed each year by a professional communication firm.

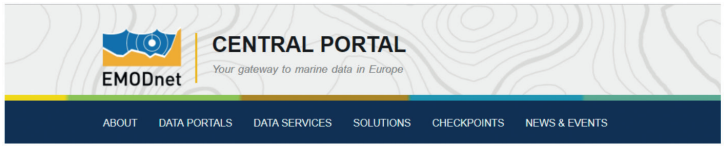


Fig. 1 - New visual identity for the EMODnet Central Portal.

The EMODnet for business campaign

The number of businesses downloading EMODnet datasets is still limited and many of them continue to pay other companies to gather data that are freely and easily accessible on EMODnet. Get more businesses involved as marine data users and data providers is therefore a major priority for EMODnet. This pushes the intermediate service providers to provide more advanced added-value products and services rather than selling what is available publicly.

In order to better meet the needs of industry and show the numerous benefits of using EMODNet, a Marine Knowledge Expert Group and an EMODnet for business campaign have been set up in 2018.

These were opportunities to highlight concrete uses of EMODnet and success stories from industry through different channels (social media, presentation to associations, etc.). A new Associated partners brochure, an EMODnet for business leaflet and a data products catalogue were created at the same occasion to get more users on board and show how they can get value from ocean data.

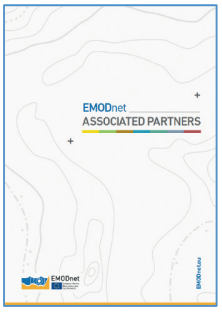


Fig. 2 - EMODnet Associated Partners Brochure.

The EMODnet Open Sea Lab hackathon

Huge efforts are being undertaken to develop and improve EMODnet and this is only worthwhile if the data and services are actually used by the various user communities. The Open Sea Lab hackathon took place in Antwerp in November 2017 to provide concrete, innovative and day-to-day usages of EMODnet portals data.

Fig. 3 - Visual identity of the EMODnet Open Sea Lab Hackathon.

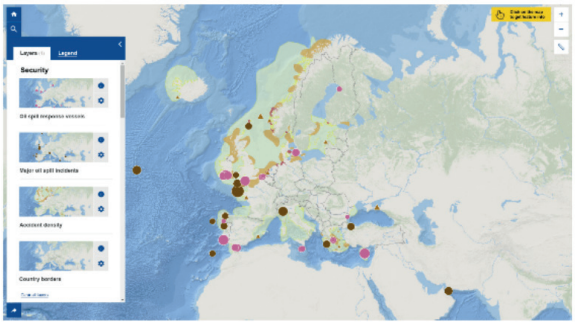


Fig. 4 - Visual identity of the EMODnet Open Sea Lab Hackathon.

This 3-day hackathon was not only an opportunity to reach new types of users, but it appears to be a powerful communication tool and an operational stress-test for EMODnet services.

Partnerships with schools and institutions

2018 is an important year for the European Atlas of the Seas as a new version of this gateway of interactive maps will be launched in June. With the ambition to become the central access point to marine information and maps for citizens, partnerships with schools and aquariums have been set up to better meet their needs. A pilot phase will be held in Automn 2018 with few institutions and wider partnerships in all EU Member states will be put in place in 2019.

EMODnet High Resolution Seabed Mapping - further developing a high resolution digital bathymetry for European seas

Thierry Schmitt, Service Hydrographique et Océanographique de la Marine (France), thierry.schmitt@shom.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl
George Spoelstra, GGS Geo Consultancy (The Netherlands), george@ggsgc.eu

Access to marine data is a key issue for the EU Marine Strategy Framework Directive and the EU Marine Knowledge 2020 agenda and includes the European Marine Observation and Data Network (EMODnet) initiative. EMODnet aims at assembling European marine data, data products and metadata from diverse sources in a uniform way.

The EMODnet Bathymetry project has developed Digital Terrain Models (DTM) for the European seas. These have been produced from survey and aggregated data sets that are indexed with metadata by adopting the SeaDataNet Catalogue services. SeaDataNet is a network of major oceanographic data centres around the European seas that manage, operate and further develop a pan-European infrastructure for marine and ocean data management. The latest EMODnet Bathymetry DTM release has a grid resolution of 1/8 arcminute (ca 250 * 250 meters) and covers all European sea regions. Use has been made of circa 7800 gathered survey datasets and composite DTMs. Catalogues and the EMODnet DTM are published at the dedicated EMODnet Bathymetry portal including a versatile DTM viewing and downloading service.

End December 2016 the Bathymetry project has been succeeded by EMODnet High Resolution Seabed Mapping (HRSM). This includes continuing gathering of bathymetric in-situ data sets with extra efforts for near coastal waters and coastal zones. So far a major increase has been

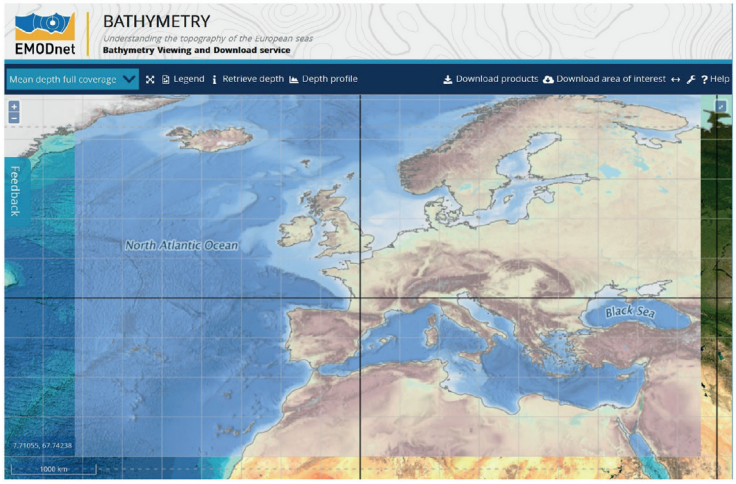


Fig. 1 - Current EMODnet Digital Terrain Model for all European seas.

achieved of the total number of survey data sets included in the CDI service from 14791 to 27078 records and Composite DTM entries from 78 to 115. The latter also includes 18 satellite derived Composite DTMs generated from Landsat images for the Mediterranean coastal zones of Spain, Greece and Libya. The total number of data providers has increased from 28 to 39. The Satellite Derived Bathymetry data are included to fill gaps in coverage of the coastal zones. The extra input data is used for producing the new EMODnet DTM with a common resolution of 1/16 arc minutes (circa 125 * 125 meters). This product will be released by mid-2018. The Bathymetry Viewing and Download service is also being upgraded to provide a multi-resolution map and including 3D viewing capability. The viewer will be based on Cesium. Current open source implementations of Cesium all make use of a height map (raster based) tile structure. Although workable, the performance on an average computer is not optimal and in areas below sea level, artefacts (tile joints) may be visible. Therefore, to overcome these issues it was decided to develop a data structure based on a triangulated irregular network (TIN). Having a TIN instead of a regular grid enables faster representation of the complexity of the map (i.e. the number and size of triangles). As there was no open-source tool available for creating tiles in quantized mesh format out of a raster format such as used in EMODnet, this software is developed by the project. This opens the door to trying different approaches for simplifying the terrain. The software will be ready in time for delivery of the new EMODnet HRSM DTM.

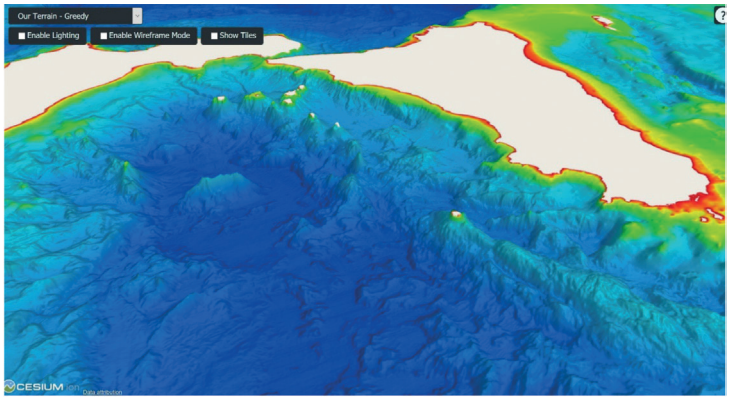


Fig. 2 - Beta 3D visualisation of the Tyrrhenian Sea DTM.

Also good progress is made with determining best-estimates of the European coastline for a range of tidal levels. A methodology has been developed for automatic extraction of a coastline from high-resolution bathymetry and from optical satellite images (typically Sentinel-2 and Landsat-8). This includes retrieving a large number of satellite images. By linking the satellite images to information about the sea-level at the time and place that the image was taken, one can derive coastlines. Coastline contours will be computed for various levels, such as LAT (Lowest Astronomical Tide), MSL (Mean-Sea-Level), and MHW (Mean-High-Water). The sea-level data are derived from the Global Tide Surge Model (GTSM). This model will also facilitate to provide the next version of the EMODnet DTM both relative to LAT and MSL.

The presentation will highlight key details of the EMODnet Bathymetry results and the way how challenges of the new HRSM project have been tackled.

EyeonWater: advancing in adopting citizen science for water quality monitoring

Peter Thijsse, Mariene Informatie Service 'MARIS' Bv (The Netherlands), peter@maris.nl
Hans van der Woerd, Vrije Universiteit Amsterdam (The Netherlands), h.j.vander.woerd@vu.nl

Short background

People have always been interested in observing their surroundings. Whereas costly observation satellites and in-situ measuring stations are set up to monitor vast areas of ocean coastlines, this can now be complemented more and more by observations carried out by citizens. The data will be used by scientists (oceanographers, limnologists) and water authorities for statistical and long-term water quality analysis in conjunction with e.g. climate research. For an optimised result and involvement of the citizen a feedback loop with results will need to be created.

The concept of EyeOnWater

Water quality of natural waters can be estimated via their colour and transparency, since these optical properties are affected by the substances that are either suspended or dissolved in it. The EOW Colour App facilitates the monitoring of colour changes of natural waters around you. The App was developed within the EU FP7 - Citclops consortium to estimate the colour of water bodies by citizens. The App is based on an old oceanographic colour standard, the Forel-Ule (FU) scale. The colour of the waterbody is compared with 21 on-screen colour bars, additionally images are also processing automatically on the server.

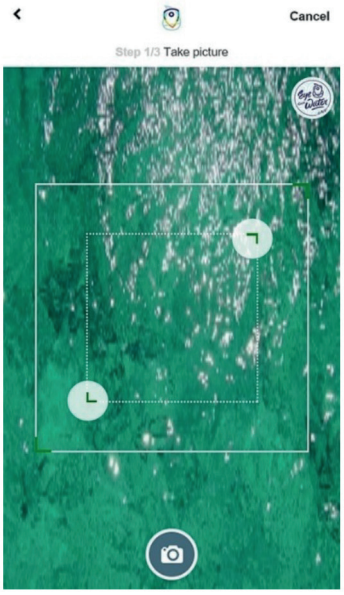


Fig. 1 - The EyeOnWater Colour App.

The app can be found in the IOS and Android stores:

https://itunes.apple.com/us/app/eyeonwater-colour/id1021542366?mt=8

https://play.google.com/store/apps/details?id=nl.maris.citclops.crosswalk

The observations from the app are sent to the server. The data concerns the image taken, the FU-index as marked by the user and a set of metadata: location, date/time, device type, angle, azimuth angle, installation ID etc. All data and metadata is stored in a database immediately at import. Incoming data is validated by running an algorithm on the image extracting RGB values and thereby processing an automatic FU index value. This can be compared to the user value and sets a quality flag.

The central EyeOnWater website www.eyeonwater.org has three main functions:

- The website provides app users a personal experience and shows all observations to all interested.

- Validation by users: Users can check other observations (e.g. around their own) and flag them if they are of insufficient quality.
- Data access services:
  - OGC compliant WMS and WFS service to serve out data to other systems
  - A search and download service for manual access to the data and metadata.

Current developments and possible extensions and use cases

The IMDIS presentation will report on the concept and especially on the following interesting projects and lessons learnt.

EyeOnWater Australia: CSIRO in Australia has started the EyeOnwater Australia project. The original EOW app and services are updated and extended. The data from citizen scientists will be used to improve interpretation of satellite data. Students are using the Eye on Water app to assess the water colour using the Forel-Ule scale. Students also use a chemical backpack to make physical and chemical measurements in their environments as part of a year 9 & 10 biodiversity and ecology unit, and the backpack will also be handed out to ranger in Northern Australia. All data is collected via the app. Via the kit the following parameters are measured: Temperature, pH, Conductivity, and later via lab processing a.o. ammonia, CO2, Nitrate, etc. The Australian version of the app is freely available and records all measurements in a national database.

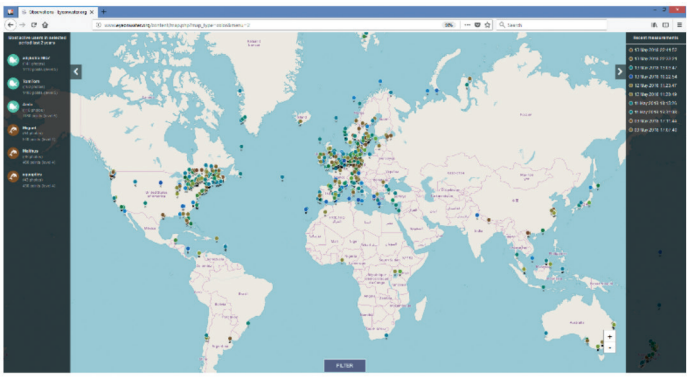


Fig. 2 - Map of colour observations.

In the Netherlands there are two varieties/spin-offs of the EOW App released:

- Monitoring biodiversity in Dutch polder canals via a set of indicator flora. Dutch farmers are involved in a project to fertilise their lands, and maintain their canals in a different way to increase biodiversity. The app supports monitoring progress.
- Monitoring waterplant nuisance for sailors. Large Dutch waters tend to have large quantities of waterplants that hinder sailors. Government maintains the waterways but dynamics of the plants are high, the sailors can now send alerts of dense areas that will trigger maintenance actions from government.

There is a high potential for near-future extensions. Internationally there is large interest for the EOW App + backpack option. But there is also a request for local applications aimed at education and water quality awareness raising by Dutch Water authorities, while cooperating closely with nature voluteers, recreational fishermen, etc. Integration and use of water quality citizen science data seems to be only just starting, but the use cases are almost endless.

The key to success is making the connection to both the volunteer (citizen, student) and to the user of the data. In the ideal case there should be a win-win for both: Data for the research organisation, improved exposure, data and water awareness for the water authority, knowledge, education or problem fixing for the citizen.

SOCIB and Balearic islands lifeguards: an example of data products and services that provide direct social benefits

Paz Rotllán, Sistema d'observació i predicció costaner de les Illes Balears (Spain), protllan@socib.es
Biel Frontera, Sistema d'observació i predicció costaner de les Illes Balears (Spain), bfrontera@socib.es
Emma Heslop, IOC-UNESCO (France), e.heslop@unesco.org
Baptiste Mourre, Sistema d'observació i predicció costaner de les Illes Balears (Spain), bmourre@socib.es
Melanie Juza, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mjuza@socib.es
Amaya ÁLvarez, Instituto Mediterráneo de Estudios Avanzados (Spain), amaya@imedea.uib-csic.es
Lluís Pujol, Universitat de les Illes Balears (Spain), lgomez-pujol@uib.cat
Miquel Àngel Rújula, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mrujula@socib.es
Inmaculada Ruíz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), iruiz@socib.es
Xisco Notario, Sistema d'observació i predicció costaner de les Illes Balears (Spain), xnotario@socib.es
Cristian Muñoz, Sistema d'observació i predicció costaner de les Illes Balears (Spain), cmunoz@socib.es
Miguel Charcos, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mcharcos@socib.es
Emma Reyes, Sistema d'observació i predicció costaner de les Illes Balears (Spain), ereyes@socib.es
Miquel Gomila, Sistema d'observació i predicció costaner de les Illes Balears (Spain), mgomila@socib.es
Sonia Gómara, Sistema d'observació i predicció costaner de les Illes Balears (Spain), sgomara@socib.es
Juan Gabriel Fernández, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jfernandez@socib.es
Joaquín Tintoré, Sistema d'observació i predicció costaner de les Illes Balears (Spain), jtintore@socib.es

Introduction

The Balearic Islands Coastal Ocean Observing and Forecasting System (SOCIB, http://www.socib.eu), is a Marine Research Infrastructure (ICTS) that provides world-class, quality controlled metocean datasets, in both real time and delayed mode. This is achieved from across its multi-platform, observation and forecasting system, covering coastal to open ocean areas. This multi-platform approach is needed to properly capture oceanographic processes, that take place at different spatial and temporal scales, and that characterise both ocean state and ocean variability. The SOCIB observation system provides physical and biogeochemical variables from different platforms such as the coastal research vessel, a high-frequency (HF) radar system, weather stations, tide gauges, moorings, drifting buoys, ARGO profilers, gliders (autonomous underwater vehicles) and sea turtle tracking, providing trajectories given by the animals. SOCIB provides 72-hour predictions of ocean temperature, currents and waves through the WMOP (hydrodynamics)

and SAPO (waves) forecasting systems. Moreover, a local implementation of the WRF model generates 48-hour predictions of atmospheric variables.

SOCIB Products and Services Strategy

After three years of full operational capacity and despite of developing a broad range of quality-controlled, freely available datasets as well as a range platform-focused apps for visualising and experiencing an increase in our web site visitors since 2015, there was no clear evidence of widespread use of the data beyond regional scientists.

In order to increase the use and utility of the available data for key societal stakeholders, in 2016 SOCIB developed a Products and Services strategy. An analysis was undertaken (web-based research, publications, meetings with key leaders in the area, participation in workshops, SOCIB facilities inputs, meetings with existing users…) and 10 sectors (groups of users) with common ocean data interests, that are regionally and globally important, and for which SOCIB has data/ intellectual property of value, were identified:

- Marine and coastal research
- Marine sports
- Beach/ coastal communities and tourism
- Coastal protection, planning and governance
- Sustainable marine ecosystems
- Integrated coastal zone and ocean management
- Ports and shipping
- Marine safety
- Education and kids
- Cross-sector products

A total of 35 products per sector are currently available in SOCIB products and services catalogue.

SOCIB Beach Lifeguards Product

SOCIB Beach Lifeguards Product was specifically designed to cover the requirements from the Balearic Islands General Directorate for Emergencies (Dirección General de Emergencias e Interior - Govern Illes Balears) in regards of beach safety, providing benefits to the operators in terms of:

- Ease daily reports fill-in (before: information in different portals).
- Hazard mitigation (close/open beaches regarding forecast).

Two tools were developed within this product, (1) SOCIB Beach Lifeguards Seaboard and (2) SOCIB Beach Lifeguards Mobile App.

SOCIB Beach Lifeguards Seaboard

SOCIB Seaboards are single 'dashboard' visualizations of the real time and forecast ocean data currently provided by SOCIB, from different coastal and ocean monitoring locations around the Balearic Islands (http://seaboard.socib.es/).

SOCIB Beach Lifeguards Seaboard was born as an in-browser application within the family of Seaboards. This product is oriented to the beach/coastal communities and tourism sector. It supplies quality controlled information about both weather and sea state forecast for the Balearic

Islands beach catalogue. In this Seaboard the user is able to browse and access past (calendar selection), present (default view) and future (48 hours ahead) critical meteo (wind speed, wind direction, rain, pressure etc.) and oceanographic (height, direction and period of waves etc) information for up to 353 beaches through a select menu and input search.

SOCIB Beach Lifeguards Mobile App

Beach Lifeguards also needed to be able to set an alarm system over the wide range of variables available for each beach, thus a mobile app was built in collaboration with *Apploading Ltd* in order to cover such requirement. The mobile app has the following additional features:

- Browse and access through up to 353 beaches using a mobile-friendly interactive map.
- Create favourites (shortcuts for a quick access to certain beaches)
- Create alarms over a set of variables for any beach
- Receive notifications and check them whenever needed.

Conclusions

SOCIB, the Balearic Islands Coastal Observing and Forecasting System, is an initiative to supply real time coastal and ocean data products and services. It provides responses to both international scientific priorities and pressing societal needs, such as beach erosion, sea level rise, and climate variability. Through this we contribute to the development and implementation of new technologies, enhancing our capacity to sustainably manage our interaction with the coastal and open ocean.

POSTERS

Tropical research and scientific data management - why one doesn't work without the other

Nicolas Dittert, Leibniz Centre for Tropical Marine Research (Germany),
nicolas.dittert@zmt-bremen.de

On June 29 2016 the first ever International Day of the Tropics was celebrated. This date marks a different perspective on the world and highlights the increasing importance of this region and its global implications as the impact of climate change and environmental degradation, poor health and educational outcomes, extreme poverty, and political and economic instability.

Human activity is causing widespread changes to the oceans' physical, chemical and biological systems. Habitats, lower-trophic-level productivity, benthic communities, fish, seabirds or marine mammals populations have been severely altered.

The international Census of Marine Life programme provided significant new data and analyses of patterns that continue to emerge today. The diversity and productivity of coral reefs and associated ecosystems (mangroves, seagrasses and pelagic habitats), for example, are among the highest globally, providing essential ecosystem services to tropical countries. Decision-makers are required to pursue policies that cover many different, yet plausible, estimates of the likelihood of alternative, future climate development. These climate-change scenarios are frequently derived from climate models, whose correctness can be measured only by access to raw data. However, many reports highlight the lack of information we have in almost every part of the world for making science-based decisions. An integrated assessment by definition needs to include environmental, social and economic information relevant to human activities, and all the components of relevant ecosystems, with input and information from a variety of geographic locations.

Explicit data management plans were developed in the late 1950s for the newly invented International Council of Scientific Unions' World Data Center (ICSU WDC) system. Since then, the compilation and exchange of scientific data had been transformed by immense technological advances. The replacement of analogue devices with digital computers, the invention of relational data base management systems, etc., have made it possible to operate data at any place and any time.

There is a seemingly inexhaustible number of scientific databases worldwide. Nevertheless, the number of those is conceivably low, who (1) support a highly diverse data model (in order to cover many scientific domains), (2) achieve maximum data consistency, (3) generate maximum data quality through the stringent application of internationally valid metadata standards, (4) commit themselves to open access, (5) offer access free of charge, and (6) guarantee long-term availability of data through independent financing.

Global Change science, requires a good availability of enormous amounts of scientific data. Yet, there are no international regulations requiring scientists to store results as raw data and

Fig. 1 - Tropical research requires proper scientific data management.

accompanying meta-information in publicly accessible archives – as for example World Data Centres. But many principal investigators or authors still refuse to archive data individually in appropriate databases, still refuse to make their published data publicly available as raw data, and still ignore the potential benefit from data networking. As a consequence vast amounts and inestimable values of scientific data are irretrievably lost every day – economically, socially, and scientifically. As a remedy we must address issues pertaining to methodology, quality, technology, data access, legislation, and privacy of data management in order to contribute to a healthy, self-determined, secure, educated life with a personal perspective for people in tropical regions.

In order to assist the Mauritanian Institute for Oceanographic Research and Fisheries (IMROP) towards the establishment of a regional (geo)graphical information systems on fishery ecology, biogeochemistry and sedimentology, one workshop at IMROP in Nouadhibou and one at the Leibniz Centre for Tropical Research (ZMT) in Bremen created the comprehensive picture of the highly heterogeneous nature of scientific data. As result a mutual back-up solution according to the so-called FAIR (findable, accessible, interoperable, and re-usable) guiding principles of scientific data management and stewardship was developed, and quantitative as well as qualitative, natural and social science data, raw data such as processed information, geo-coded and spatially unrelated data, point information such as spatial data sets or time series, historical as well as newly collected data, etc. are being digitized and safeguarded. At the same time, joint interdisciplinary studies are initiated together with IMROP in order (1) to gain a better understanding of subtle or previously unknown patterns of the causality of complex interrelationships between natural and by men influenced systems and their effects on the tropics, and (2) promote the so-called cross-referencing of scientific data and publication to its author(s) through a technical www link between data and publication.

Extremes of extratropical storms over North Atlantic based on cyclone indicators in ESIMO

Natalia Viazilova, All-Russian Research Institute of Hydrometeorological Information - WDC (Russia)
nav@meteo.ru

The goal of this work is to show the changes in extremes of extratropical storms using of cyclone activity indicators. The list of cyclone parameters includes number of cyclone tracks, cyclone frequency, the cyclone activity, maximum storm wind and maximum precipitation near cyclone centres that are available on the Unified State System of Information on the Global Ocean (ESIMO) portal [http://portal.esimo.ru]. The cyclone parameters were calculated based on automated cyclone identification and tracking algorithm using the 6-hourly SLP and surface wind from the NCEP/NCAR DOE reanalyses. The cyclone frequency and index of cyclone activity are calculated as number of the cyclones centres and sum of pressure anomaly in cyclones centres during the month in every grid point.

Focus of this study is intensity of extratropical storms, that here is the maximum storm wind in cyclones. The maximum wind strength was defined from a 5° spherical radius near cyclone centre. The data with cyclone track, cyclones centres and maximum wind coordinates are presented on ESIMO portal for last calendar month (Fig.1). Storm activity indicators are calculated, as number cyclones with different maximum wind intensity, using scale Beaufort, and are presented for North Atlantic and Baltic Sea for every month and winter and summer season for period from 1999/12 to present (Fig.2).

Fig. 1 - Maximum wind speed near cyclones centres for last month.

a) b)

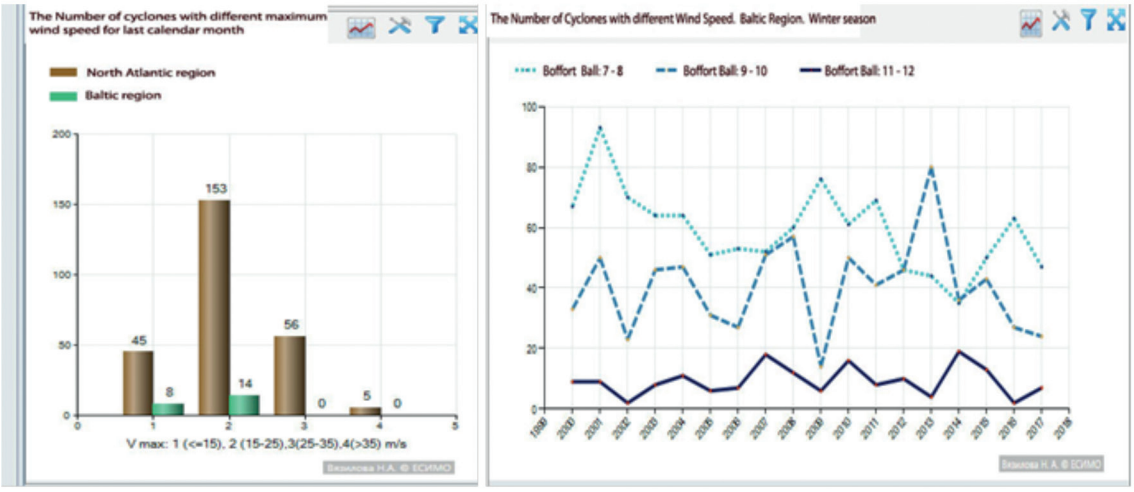


Fig. 2 - The number of cyclones with different maximum wind speed: a) for North Atlantic and Baltic Sea region during last calendar month, b) Baltic Sea region for period from 1999 to present. Winter season.

Extreme storms are usually defined as cyclones with extreme wind, wind strength 24 m/sec and more, using scale Beaufort. Analyses of distribution for cyclones with different maximum wind intensity (weak strong, moderately strong, extreme wind) shows, that in high-latitude North Atlantic most part of storms with extreme wind are the cyclones with extreme depth of pressure, with an MSLP of 970 hPa or less. On other side, in low- latitude North Atlantic extreme wind strength often is observed in cyclones with moderate pressure depth. It is important to note, that as in north, as in south region of North Atlantic, the part of extreme storms, the storms with wind strength from 24 m/sec and more, increases during past years.

Black Sea water quality database within the EMBLAS project

Yevhen Ivchenko, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), yevhen.ivchenko@outlook.com

Artem Kruhlov, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), kruhlov.artem@gmail.com

Oleksandr Leposhkin, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), cephei7@gmail.com

Oleksandr Neprokin, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), o.neprokin@gmail.com

Olena Miasnikova, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), olena.miasnikova@gmail.com

Maksym Motylov, Ukrainian Scientific Centre of Ecology of the Sea (Ukraine), maxmotylev@gmail.com

The EMBLAS project

The overall objective of the project is to improve the protection of the Black Sea environment. The project is addressing the overall need for support in protection and restoring the environmental quality and sustainability of the Black Sea.

The specific objectives: are as follows:

- Improve availability and quality of data on the chemical and biological status of the Black Sea, in line with expected MSFD and Black Sea Strategic Action Plan needs;
- Improve partner countries' ability to perform marine environmental monitoring along MSFD principles, taking into account Black Sea Diagnostic Report.

The overall aim of the "web-based Black Sea Water Quality Database" (BS WQDB) project

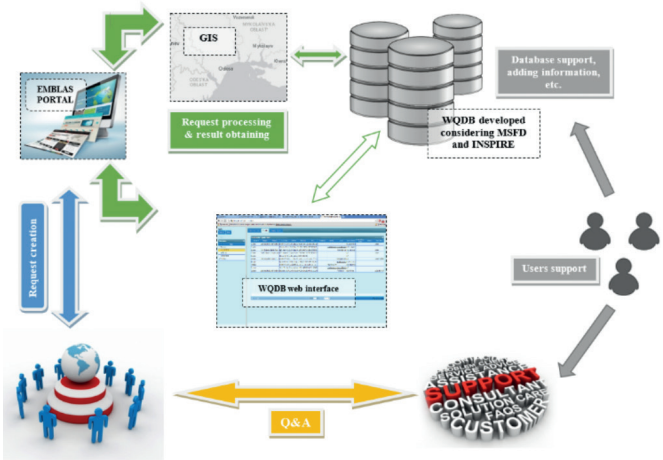


Fig. 1 - WQDB data flow.

activity is to further develop the database, put it on-line and promote its sustainable update, maintenance and use by stakeholders. It will allow storage of the monitoring data already collected and newly collected during the project, with focus on upgrade of the Black Sea Information System (BSIS). The data are being collected through the Data Collection Templates that are customized to the Black Sea JOSS and NPMS cruises.

The database

The large volumes of data were collected During the lifespan of the Project from the different sources including scientific marine monitoring cruises, observations aboard of the ships, etc. The information received on the following features: biological parameters (Phytoplankton, Zooplankton, Zoobenthos), chemical parameters (Pollution in water, bottom sediments, biota), indicators of eutrophication, marine litter, marine mammals and birds observations.

Database developed using the Microsoft SQL Server software with respect to the Marine Strategy Framework Directive requirements. Due to the use of the modern approach and methods during the development it became possible to provide different interactive map integration (interactive map services publication by means of ESRI ArcGIS) and data exchange services integration (data sets automation for the EMODNET and SeaDataCloud marine data exchange Projects from the BS WQDB).

The processed data of the surveys are inserted into the specially developed templates by the researchers and later uploaded into the WQDB using parser software also developed within the project.

User friendly web interface developed for the convenient search and query from the WQDB.

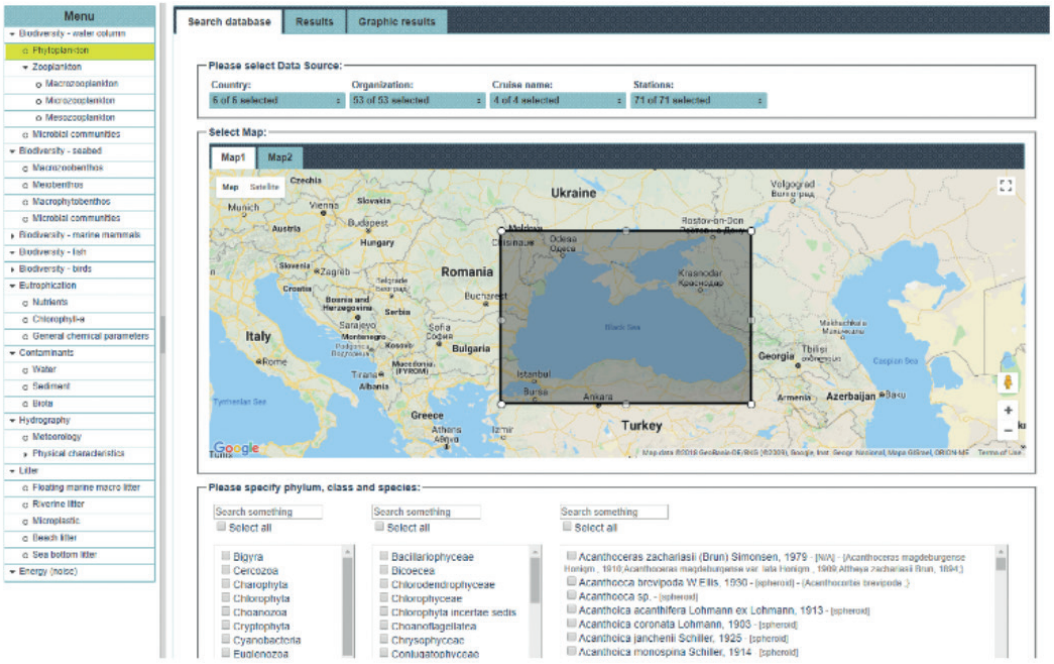
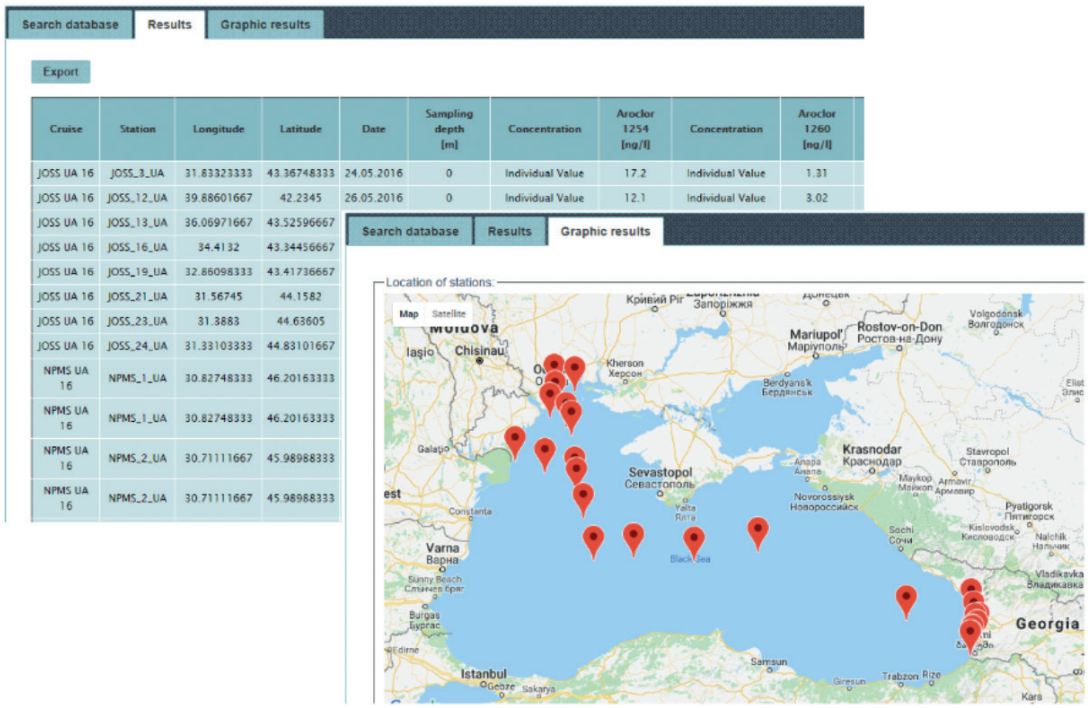


Fig. 2 - WQDB web interface (data query).



Cruise	Station	Longitude	Latitude	Date	Sampling depth [m]	Concentration	Aroclor 1254 [ng/l]	Concentration	Aroclor 1260 [ng/l]
JOSS UA 16	JOSS_3_UA	31.83323333	43.36748333	24.05.2016	0	Individual Value	17.2	Individual Value	1.31
JOSS UA 16	JOSS_12_UA	39.88601667	42.2345	26.05.2016	0	Individual Value	12.1	Individual Value	3.02
JOSS UA 16	JOSS_13_UA	36.06971667	43.52596667						
JOSS UA 16	JOSS_16_UA	34.4132	43.34456667						
JOSS UA 16	JOSS_19_UA	32.86098333	43.41736667						
JOSS UA 16	JOSS_21_UA	31.56745	44.1582						
JOSS UA 16	JOSS_23_UA	31.3883	44.63605						
JOSS UA 16	JOSS_24_UA	31.33103333	44.83101667						
NPMS UA 16	NPMS_1_UA	30.82748333	46.20163333						
NPMS UA 16	NPMS_1_UA	30.82748333	46.20163333						
NPMS UA 16	NPMS_2_UA	30.71111667	45.98988333						
NPMS UA 16	NPMS_2_UA	30.71111667	45.98988333						

Fig. 3 - WQDB web interface (query result and stations map).

The web interface

Various development tools like HTML, CSS, PHP, Ajax, JQuery, Google Maps API were selected to build the web interface. Interface contains main menu (allows to display filters set for the required features) and work space divided in to 3 functional areas – filters set, results display and monitoring station display.

Web interface key features:

- Specific criteria data query filters;
- Results table view;
- Different formats of the data export (Excel, CSV, etc.);
- Query results for monitoring stations display on map (Query results for monitoring stations displaying on map);
- Displaying main statistical parameters.

MarineRegions.org - a world reference for marine placenames and Maritime Boundaries

Lennert Schepers, Vlaams Instituut voor de Zee (Belgium)
Paula Oset Garcia, Vlaams Instituut voor de Zee (Belgium)
Britt Lonneville, Vlaams Instituut voor de Zee (Belgium)
Simon Claus, Bart Vanhoorne, Vlaams Instituut voor de Zee (Belgium)
Francisco Hernandez, Vlaams Instituut voor de Zee (Belgium)
Jan Mees, Vlaams Instituut voor de Zee (Belgium)
info@marineregions.org

Marine Regions is a global geographic database containing georeferenced geoobjects on (i) marine placenames and (ii) maritime boundaries. The database containing both types of information is accessible through MarineRegions.org and by various web service applications.

The marine placenames gazetteer started with geographic, mainly marine names such as seas, sandbanks, seamounts, ridges, bays or even standard sampling stations used in marine research. The geographic cover initially focused on the Belgian Continental Shelf and the Scheldt Estuary, but gradually more regional and global geographic information were added. The gazetteer now includes almost 50,000 placenames, including national, regional as well as global marine gazetteers such as the GEBCO and ACUF gazetteers.

The basis for the Maritime Boundaries dataset is the United Nations Convention on the Law of the Sea (UNCLOS), which was signed in 1982 and came into force in 1994. This convention has defined a series of maritime zones (internal waters, archipelagic waters, territorial sea, contiguous zone, exclusive economic zone and continental shelf) and establishes the degree of rights and obligations of a country in each of those areas. The Exclusive Economic Zone (EEZ) is a country's basic geo-unit for the management of marine natural resources. Despite the strategic significance of EEZs, a standard georeferenced product with maritime boundaries was not available at the global level (Claus *et al*., 2014), until it was developed and made available by the Flanders Marine Institute (VLIZ) in 2006 (Deckers and Vanden Berghe, 2006). Importantly, the Maritime Boundaries database is regularly updated with the latest international treaties. The latest version,

Marineregions.org
a standard for georeferenced marine names

Fig. 1 - Marine Regions Logo.

Version 10, was launched in February 2018 by the portal website MarineRegions.org and has already been downloaded 2173 times (as of April 2018).

The Maritime Boundaries is the most popular product available at Marineregions.org. In all its different versions, the product has been downloaded more than 50,000 times, representing 65% of the total downloads. These boundaries and other derived products are used in many projects related to bio-geographic research and conservation, such as the World Register of Marine Species (WoRMS), The Sea Around Us (Pauly and Zeller, 2015), Global Fishing Watch or The Ocean Health Index (Halpern, 2012).

The MarineRegions products are accessible by the website MarineRegions.org, which offers a map viewer, a search page and a download page. Marine Regions also provides several web services which allow the user to have direct access to the geographic data, maps and metadata from a GIS desktop or other applications.

References

Claus, S.; De Hauwere, N.; Vanhoorne, B.; Deckers, P.; Souza Dias, F.; Hernandez, F.; Mees, J. (2014). *Marine Regions: Towards a global standard for georeferenced marine names and boundaries.* Mar. Geod. 37(2): 99-125.

Deckers, P., and E. Vanden Berghe, 2006. *The VLIZ maritime boundaries geodatabase as a bio-geographical tool.* Book of Abstracts ICES 2006 Annual Science Conference, Maastricht, The Netherlands

Pauly D. and Zeller D. (Editors), 2015. *Sea Around Us Concepts, Design and Data* (seaaroundus.org)

Halpern, B. S. *et al. An index to assess the health and benefits of the global ocean.* Nature Nature 488, 615-620 (30 August 2012).

Black Sea monitoring to evaluate the ecological state of the sea water in the Georgian coastal area

Robert Diakonidze, Tsotne Mirtskhulava Water Management Institute of Georgian Technical University (Georgia), robertdia@mail.ru

Kakhaber Bilashvili, Institute of Oceanography and Hydrology of Iv. Javakhishvili Tbilisi State University (Georgia), wocean@consultant.com

Vazha Trapaidze, Institute of Oceanography and Hydrology of Iv. Javakhishvili Tbilisi State University (Georgia), vazha.trapaidze@gmail.com

Irine Baramidze, Batumi State Maritime Academy (Georgia), irine.baramidze@gmail.com

Tamriko Supatashvili, Tsotne Mirtskhulava Water Management Institute of Georgian Technical University (Georgia), tamunasupatashvili@gmail.com

Bela Diakonidze, Tsotne Mirtskhulava Water Management Institute of Georgian Technical University (Georgia) bella_diakonidze@mail.ru

Nino Machitadze, Scientific Research Firm GAMMA (Georgia), ninomachitadze@gamma.ge

Nino Gelashvili, Scientific Research Firm GAMMA (Georgia), ninogelashvili@gamma.ge

Vakhtang Gvakharia, Scientific Research Firm GAMMA (Georgia), vakhtangvakharia@gamma.ge

The Black Sea coastal area in Georgia is the most vulnerable eco-system to the modern environmental change. It is under both, natural and intense anthropogenic impact. Water quality is the severest ecological problem of the Black Sea. On the background of anthropogenic loads, a risk of an increasing concentration of the substances in the surface waters able to change the natural background increases further. The requirements of the Association Agreement between the European Union and Georgia envisage permanent improvement of the qualitative and quantitative indicators of the sea data by using an eco-system approach, development of sea research and sea environment monitoring. The principal guiding document for the sea ecological state is Water Framework Directive (WFD) together with the Marine Strategy Framework Directive (MSFD), which came into force on July 15, 2008. Within the scope of the EC-funded scientific projects, including ongoing projects SeaDataCloud, EMODnet, EMODnet Data Ingestion, EMBLAS II, it became possible to obtain oceanographic data, provide e-cataloguing of meta-data and to publish them in the relevant European bases after due formatting.

In order to estimate the sea water, the territorial waters of the popular Georgian resorts of the Black Sea (Batumi, Kobuleti, Ureki, Grigoleti, Maltakva and Anaklia) were selected. The water samples were taken near the shore and at about 150 m from the shore, at the depth of 30 cm. The quality of water in the rivers flowing into the Black Sea on the territory of Georgia (Chorokhi, Kintrishi, Chakvistskali, Natanebi, Supsa, Rioni, Khobi and Enguri) was also studied. Total of 30 indicators were estimated. The water quality was estimated by using the European and local (Georgian) regulatory guidelines.

Results of the study

As the study results suggest, the water in the Black Sea, in the territorial waters of Georgia are mostly polluted with heavy metals: zinc, copper, nickel, arsenic, lead and phenols, with their values exceeding maximum admissible concentrations (MAC) by 0,1-0,5 times on average and with their maximum values exceeding MAC by 1,5 to 2,5 times. The values of the pollution, according to the above-mentioned estimate, are not harmful for human health. On the other turn, the effluent waters in the sea are badly polluted with various substances, including microbiological substances.

As for the presence of hydrogen sulphide in the sea (values measured 3 miles from the city of Poti), the water with H2S begins at 180 m, where the depth of the sea is over 500 m and this value coincides with a commonly recognized value and its limit in the territorial waters of Georgia.

Conclusions

Sampling and analysis of the prioritized substances determining the chemical state of the sea water are quite labour-consuming. Despite the fact that there is much done as a result of the field and theoretical studies, the issue cannot be considered as solved to a desirable level. Based on the available and obtained data, the work to visualize and analyse the data and information will be continued to provide the final product using contemporary approaches and software.

Geophysical data valorisation and dissemination in the framework of the EMODnet Ingestion project. The OGS Experience

Paolo Diviacco, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), pdiviacco@inogs.it

Mihai Burca, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), mburca@inogs.it

Giuseppe Brancatelli, Istituto Nazionale di Oceanografia e di Geofisica Sperimentale (Italy), gbrancatelli@inogs.it

Introduction

The EMODnet Ingestion project aims at sharing data from public and private sectors that otherwise would remain inaccessible. This could be due to many reasons, such as the case when originators are not "connected" with the main international data sharing initiatives, or do not have the know-how to disseminate them, or simply because they do not have the resources to do it.

Geophysical data

The case of Geophysical data is particularly interesting because this domain is living a complex moment due to restrictions in its activities. From a scientific point of view, funding opportunities have been reduced rather severely, and at the same time, from the commercial point of view, exploration for Oil and Gas has also been reduced due to the current difficult economic trend and the introduction in some areas of strict environmental regulation. All this resulted in a reduction of the activity of acquisition of new data and in the increase in the demand for already available data. In addition, data acquired in the '70s and '80s have great possibilities that often are not available in modern data. For instance, at that time, the lack of environmental awareness allowed the use of large seismic sources that nowadays are not permitted anymore, while often the geopolitical setting of that time allowed regional surveys that currently would not be possible. On the other had that data have been acquired with outdated techniques and are available on media and in formats that introduce the need to invest a lot of resources to recover them.

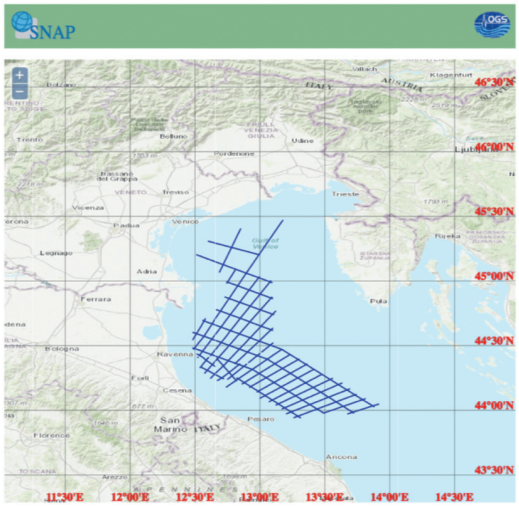


Fig. 1 - Map of the Zona A seismic database.

The Zona A

In 1964 the Italian National Agency for Hydrocarbon exploitation (AGIP) performed a

large survey of the Italian offshore areas. So far, this large dataset was available only as paper prints that of course cannot be used with modern software and methodologies. OGS, though an agreement with the Italian Ministry of Economic Development (MISE) recovered and vectorialized that dataset to produce data that could be used within the current practices.

EMODnet Ingestion

Vectorialization has been done for the actual seismic sections but of course on

Data Processing

Further to vectorialization, since proper digital data have been made available, the dataset underwent a phase of reprocessing. A particular processing sequence has been adopted in order to obtain a final stack section from a 100% fold coverage original data. This sequence consists of geometry assignment, bias removal, shots interpolation, random noise attenuation, dip filtering, predictive deconvolution, stack and time variant filtering. The final reprocessed section shows the improvements in the shallow area and in the continuity of signals (Fig. 2) where the strong sea floor reverberation have been attenuated.

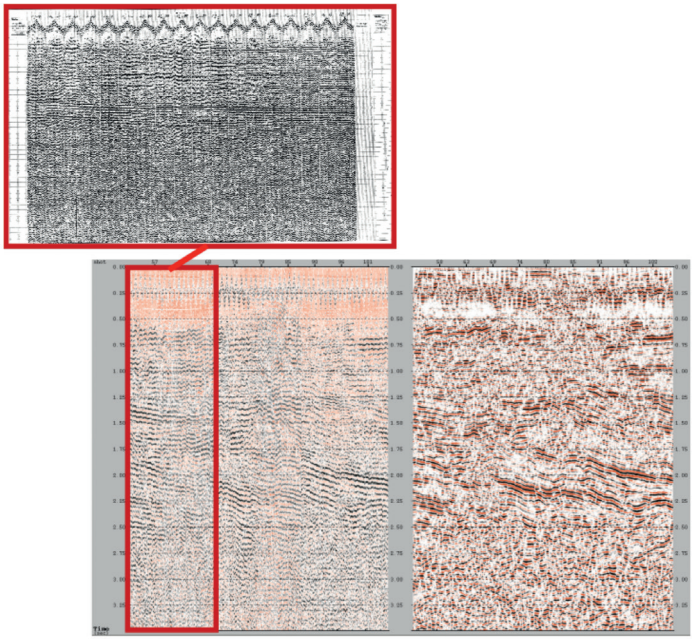


Fig. 2 - Original paper section (upper left), Vectorialized data that shows also that several paper sections are integrated to get a complete seismic section (lower left) Reprocessed data that shows the improvements in the shallow area and in the continuity of signals (lower right).

EMODnet Ingestion

Not only data have been restored but navigation has been extracted from paper positioning maps and after this data have been georeferenced. Metadata extracted and conformed to the EMODnet and SeaDataNet/GeoSeas standards and made available through the SNAP portal

Conclusions

We have demonstrated the importance of recovering, valorisation and dissemination of vintage seismic data. The work is complex, and several steps have to be done manually but yet the results that can be obtained can be very relevant both for the scientific community and for the industry.

Processing 50 years of oxygen and hydrogen-sulphide observations in the Baltic Sea

Susanne Feistel, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), susanne.feistel@io-warnemuende.de

Michael Naumann, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), michael.naumann@io-warnemuende.de

Günther Nausch, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), guenther.nausch@io-warnemuende.de

Anne Hiller, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), anne.hiller@io-warnemuende.de

Philipp Paysen, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), philipp.paysen@io-warnemuende.de

Martin Hansson, Swedish Meteorological and Hydrological Institute (Sweden), martin.hansson@smhi.se

Lars Andersson, Swedish Meteorological and Hydrological Institute (Sweden), lars.andersson@smhi.se

Lena Viktorsson, Swedish Meteorological and Hydrological Institute (Sweden), lena.viktorsson@smhi.se

Elzbieta Lysiak-Pastuszak, Institute of Meteorology and Water Management (Poland), elzbieta.lysiak-pastuszak@imgw.pl

Rainer Feistel, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), rainer.feistel@io-warnemuende.de

H.E. Markus Meier, Leibniz Institute for Baltic Sea Research Warnemünde (Germany), markus.meier@io-warnemuende.de

The Baltic Sea is a complex estuary characterized by a strongly fluctuating, fragile balance between high freshwater runoff and saline water inflows, a persisting vertical stratification, and a bottom topography of several connected basins. By the sensitivity of this system, climatological fluctuations appear amplified on the decadal time scale. Changes that may be insignificant in the open ocean typically constitute significant indicators in the Baltic Sea. Salt and nutrients remain present in the estuary for 20 and more years before being flushed to the Atlantic along with the brackish water export. This long residence time attenuates short-time fluctuations in environmental conditions, but highlights systematic, even small long-term anomalies.

Lateral property distribution maps published in 2016 by Feistel et al. (https://www.io-warnemuende.de/msr-2016-0100.html) allow an evaluation of occasional inflow events, of the progress of oxygen-consuming processes and of the development of hydrogen sulphide distributions over longer periods of time. The oceanographic database IOWDB (https://odin2.io-warnemuende.de) serves as a central primary data source and contains harmonized, quality-controlled oxygen and hydrogen-sulphide data from regular seasonal monitoring cruises that have visited the western and central Baltic Sea since 1969. Furthermore, our research task combines

IOW long-term data with those from cooperating partner institutions in Sweden and Poland. The result presents the most comprehensive dataset currently available and can be used to illustrate and analyze the shifting levels of oxygen deficiency.

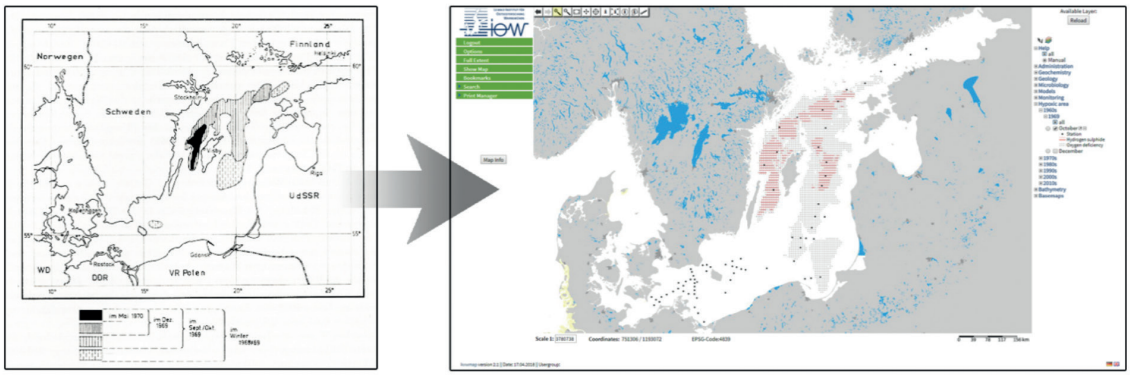


Fig. 1 - Change over time: The visualisation of the Baltic Sea with hypoxic and euxinic areas in 1969. On the left a hand-drawn map from 1971, on the right as a shape file and web map service from 2018.

The applied database- and software-based methods represent transparently and reproducibly the distribution of hypoxic and euxinic waters in the near-bottom layer of the Baltic Sea (Fig. 1). The datasets of IOWTOPO (https://www.io-warnemuende.de/topography-of-the-baltic-sea.html) and RANGS (https://www.io-warnemuende.de/rangs-en.html) form the topological framework for all created maps. As a new feature, we now present all extrapolated map data as so-called shape files for GeoInformationSystems and offer them as open-access web service for public use (http://www.io-warnemuende.de/baltic-sea-atlas). The maps permit detailed statistical analyses covering the central basins in all seasons, spanning several decades as outlined in Naumann et al. 2017.

For example, Fig. 2a shows long-term change of hypoxic and euxinic areas in the Baltic basins, where maximum extents were reached in the early 1970's as well as from 2000 on. Fig. 2b shows the change over time of the mean redoxcline depth which is more or less constant at about -110 m and is temporarily lowered during inflow situations. Under recent meteorological and hydrographic conditions, a further expansion seems to be restricted by the pycnocline depth.

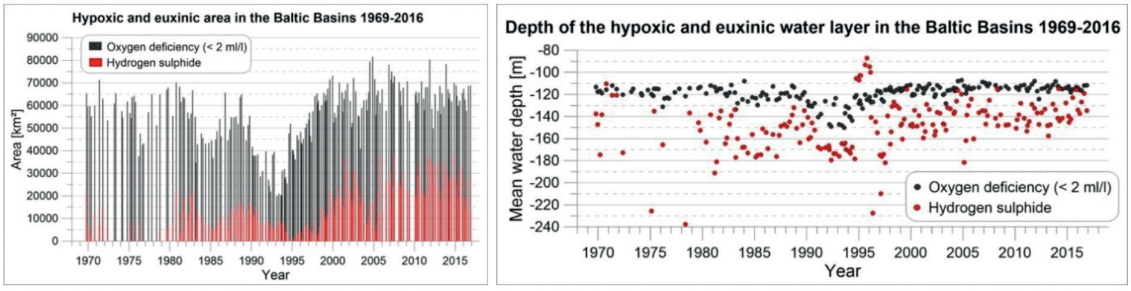


Fig. 2 - Analytical results of available datasets: a) Spatial analysis of hypoxic to euxinic conditions in the Baltic Sea since 1969; b) Mean water depth showing hypoxia and euxinia, considering all deep basins (after Naumann *et al.*, in prep).

Integration of underwater noise measurements into EMODnet Physics

Joaquin Del Rio, Universitat Politecnica de Catalunya (Spain), joaquin.del.rio@upc.edu
Michel Andre, Universitat Politècnica de Catalunya (Spain), michel.andre@upc.edu
Thomas Folegot, Quite Oceans (France), thomas.folegot@quiet-oceans.com
Mike van der Schaar, Universitat Politecnica de Catalunya (Spain), mike.vanderschaar@upc.edu
Patrick Gorringe, EMODnet Physics (Sweden), patrick.gorringe@smhi.se
Antonio Novellino, EMODnet Physics (Italy), antonio.novellino@ettsolutions.com

Underwater noise has been significantly raising in the past decades due to an increment of human-related activities in the oceans such as shipping, industrial activities, seismic explorations, etc. These activities may have adverse effects on fish and mammals, such as communications masking and modifying predator–prey interactions.

In order to assess and limit the impact of these, the European Commission approved the Marine Strategy Framework Directive (MSFD) which aims to achieve a good environmental status in European waters. Within this directive different environmental challenges are addressed, including the long-term monitoring of underwater noise throughout European waters.

EMODnet Physics, one of the European Marine Observation and Data network thematic portals, which is is currently providing easy access to data and products of: wave height and period; temperature and salinity of the water column; wind speed and direction; horizontal velocity of the water column; light attenuation; sea ice coverage and sea level trends. EMODnet Physics is continuously increasing the number and type of platforms in the system by unlocking and providing high quality data from a growing network. EMODnet Physics has recently started working on water noise with the aim of making available more operational data (in terms of parameters and format that are close to MSFD I.11 requirements), offer a single European entry point to impulsive noise registries (MSFD I.11.1) and work on (regional) sound maps are three key identified activities for Physics. Furthermore the very first operational under water noise data (i.e. Sound Pressure Level – SPL), and HELCOM and OSPAR impulsive sounds registry were connected and are now available on the portal. Exploiting the LIDO (Listen to Deep Ocean) knowledge and the BIAS project (https://biasproject.wordpress.com/) experience EMODnet Physics will develop and make available monthly sound maps.

In this presentation, we give an overview of how EMODnet Physics is organized, with a particular focus on this new data flow and its perspectives.

Learning from 2011 Japan's tsunami warning system's short comings

Masa Hayashi, Live Warning Inc. (United States of America), masa@livewarning.org
Art Honda, Live Warning Inc. (United States of America), art@livearning.org

2011 Japan's Earthquake/Tsunami killed an estimated 19,000+ people, and Japan had the world's most advanced tsunami warning systems. However, the stunning fact is that, even though the majority of the 19,0000 victims KNEW and HEARD the tsunami warnings and had time to evacuate prior to the tsunami arrival, they were still victimized. The current tsunami warning systems (DARTs) including Japan's latest deep ocean cable pressure sensor warning system are based on how fast and how accurately the public can be warned. However, the end public interface is still based on text, radio, TV, and the local communication systems (often knocked out), and in addition, they do not convince some of the public to react. Learning from the Japan Tsunami catastrophe, the current tsunami warning systems throughout the world are missing the most important aspect, a way of CONVINCING the public to react to these tsunami warnings. Over time, the human nature of underestimation, tsunami ignorance, turnover of generations, chronic false tsunami warnings, inaccurate tsunami info. Etc. have resulted in staggering apathy.

The 2011 Japan earthquake occurred at 75km off the Northeast coast at 2:46PM, and the 1st tsunami wave arrived at Rikuzen Takada city (killed 1,939) at 3:16PM, 30 minutes later. Then, the 1st wave arrived at Ofunato city (killed 446) at 3:20PM, Otsuchi city (killed 1,378) at 3:22PM, Ishinomaki city (killed 3,890) at 3:43PM, Higashi Matsuyama city (killed 1,138) at 3:45PM, Sendai city (killed 919) at 4:04PM, Soma city (killed 1,122) at 4:10PM, and so on (Fig. 1).

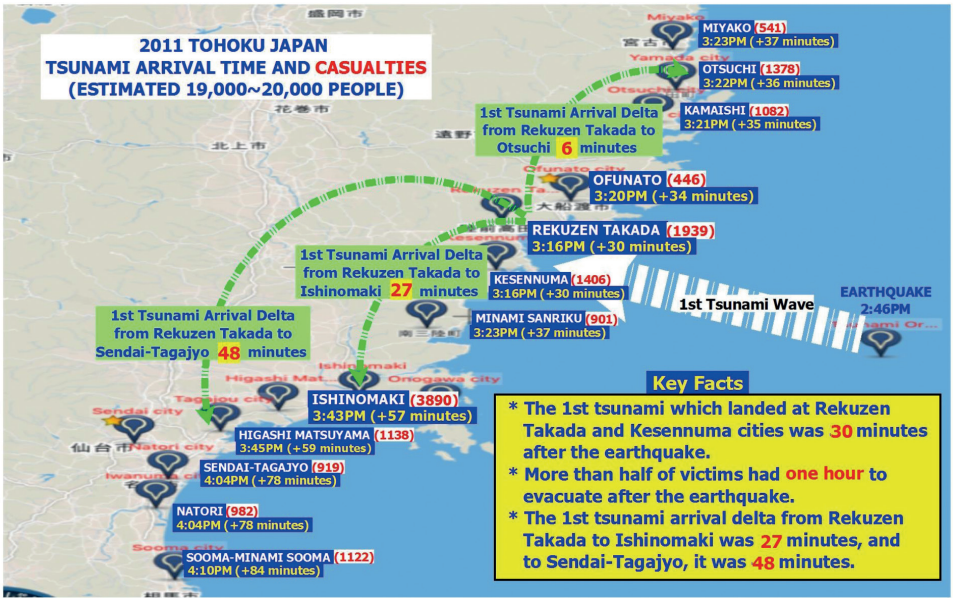


Fig. 1 - Location of the 2011 Japan's Earthquake/ Tsunami.

Based on each city's 1st tsunami wave arrival time and casualties, the majority of the 19,000 victims had more than one hour to evacuate. More importantly, when the 1st tsunami landed in the Rekuzen Takada city, a greater number of victims had more than 30 minutes to evacuate. These time differences could have been utilized to save the majority of potential victims, if the potential victims had been convinced to take evacuation action immediately.

The tsunami warning's biggest hurdle is the infrequency of a huge tsunami event occurring at the same location, and the next huge tsunami event will occur some place in the world, not if but WHEN. The last huge earthquake/tsunami catastrophe in Europe occurred on November 1st, 1755, and an estimated 40,000 people were victimized in Lisbon alone.

The solution to save the majority of the people, who do not know anything about a tsunami danger, or do not have any urgency to evacuate, is to show a LIVE tsunami picture/image, video, with a warning sound, and tsunami warning messages to the individual public's smart phones, tablets, and also the other instant visual media.

The Live Visual Tsunami Warning System's done stations/communication hubs would be placed in the key coast cities, and would be linked to cell communication. When an earthquake of a greater than set magnitude is detected (Ex: >4M), all the drones would be deployed at these key stations/hubs autonomously and simultaneously to their cities' predetermined sea locations to search for a 1st tsunami wave. If the 1st tsunami wave is sighted by a city, then in real time, the 1st wave sighted video with text and auditory warning messages, containing the name of the 1st wave sighted city, the wave expected arrival time, the expected wave height, and the city's evacuation information, would be transmitted to cells, tablets, and TV stations, ahead of the 1st tsunami arrival, to warn the public in this city, as well as all others in the region and in the nation. Furthermore, when the 1st tsunami wave destroys the initial city, the drones, and the drone based communication network would continue to transmit the initial city's real time DESTRUCTION video and warning messages to the initial city as well as the other cities in the region to provide further urgency for the public to evacuate immediately (Fig. 2).

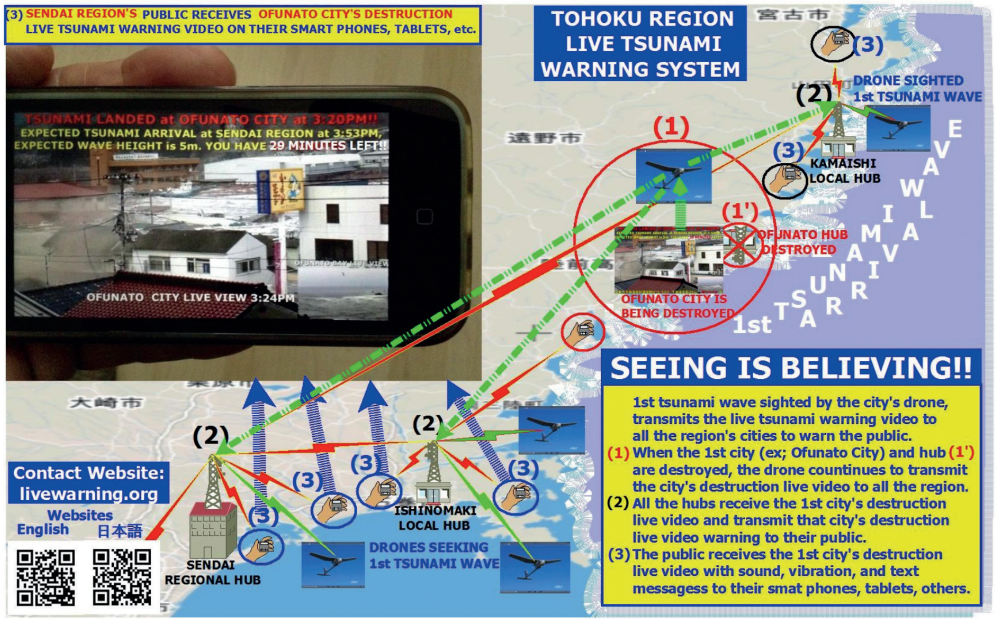


Fig. 2 - The Live Visual Tsunami Warning System's.

From oceanographic data integration systems to knowledge management - issues and approaches

Nurhan Hasanov, Institute of oceanology - BAS (Bulgaria), nurhi@abv.bg
Vasil Donev, Institute of oceanology - BAS (Bulgaria), donev@sirma.bg
Atanas Palazov, Institute of oceanology - BAS (Bulgaria), palazov@io-bas.bg

Oceanographic raw/primary data increases exponentially last decades and this phenomenon strongly desired by the scientists in the past now threatens to become a serious issue while interpreting or simply searching for a particular type of information. Such newly emerged challenge has been seriously considered both institutionally and professionally and is now addressed with some preventive measures in the field of the information technologies sciences which already provide some kind of order in respect of data storage and integration. However, with such vast amount of information now easily available the question of its correct analyzing, rationalization and finally getting certain new facts out of it pose serious task for all participants in the process – scientists, government officials, business bodies, analysts. The bottom line of all these developments is that the focus moves from gathering of primary information to producing new knowledge out of all available information sources. This in turn changes the approaches for managing the information flows – from simple arrangement and search for certain data in the databases to obtaining some kind of existing or new knowledge. The issue is not unknown in other fields of human activities, for instance the business management. For the oceanology science however, the matter is further complicated by the extensive kinds of data being otherwise reliably stored and we may assume nowadays relatively easily available. The transition from data management to knowledge management obviously requires implementation of new instruments for information processing or at least some kind of an upgrade of contemporary data management tools.

Based on the above considerations the article deals mainly with the issues that go in line with the data integration as basis for further knowledge creation or access. Following this path, an attempt is made to tie up the conclusions and the recommendations concerning the knowledge management process to the needs of broader span of end users rather than limiting the results for the oceanographers' community only. Such open approach is determined first of all by the constant deepening of the interrelations between the science and the economy deemed to be an economy of knowledge as often called nowadays. An analysis of the major drawbacks of some of the most used systems for oceanographic data integration is presented and the results of this analysis outline the main directions which should be followed by the institutions in search of extra value from the yet accumulated oceanographic information. Given the complex nature of the oceanographic information, the still existing variety of formats, storages and forms of data, the broad span of interests for this information as well as the new technological instruments for

processing, a conclusion can be drawn up that further improvement of the usage of the existing oceanographic information could be reached only through the implementation of semantic technologies in the search process for the information. The idea is exemplified through several approaches adapted to the needs of the institution. On the basis of the proposed solution the presentation also examines the opportunities for the creation of an expert system for niche studies which draws resources from the already shown model of semantic integration system.

The fundamental idea that lies behind the approach described in the presentation is the understanding that both science and business are mutually dependent in respect of the progress they seek for and only through creations of technical instruments which open the information space of the science to wider community of users – business, administration, social and nongovernmental entities - we can reach maximum efficiency in transformation of data into knowledge. To prove this the presentation demonstrates both technical and methodological solutions for data and knowledge management. A brief description of the structure of the combined data management system is given with elements of analysis for the most critical parts of it. Finally, the presentation opens a discussion on some further fields of development in the context of the latest achievements of the information technologies and how these in our opinion may affect the functionality and the access features of the system.

The EMODnet Seabed Habitats mapping portal

Eleonora Manca, Joint Nature Conservation Committee (United Kingdom),
Eleonora.Manca@jncc.gov.uk

Introducing the EMODnet Seabed Habitats mapping portal

The EMODnet Seabed Habitats mapping portal (www.emodnet-seabedhabitats.eu) provides a permanent single access point to freely view and download marine habitat data and maps from across Europe. EMODnet (European Marine Observation and Data Network) is a data initiative funded by the European Commission, aimed at assembling fragmented resources and providing easy access to quality-assured and harmonised marine data, through its online portals. There are seven other portals within the EMODnet initiative covering all marine data types from bathymetry to human activities.

The presentation will provide an overview of the main contents and services provided by the portal, and will include user cases and planned activities by the EMODnet Seabed Habitats consortium.

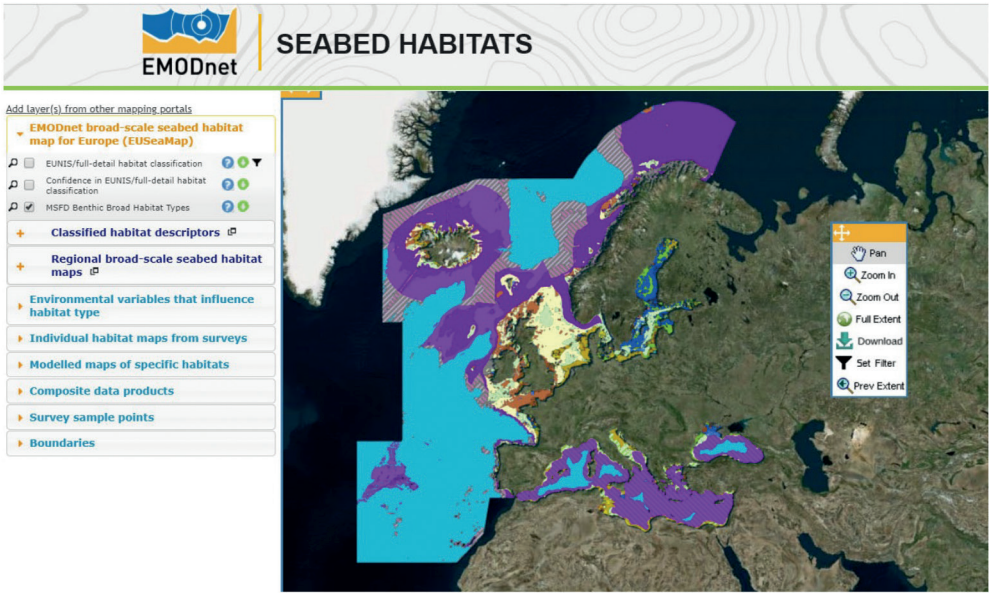


Fig. 1 - A screenshot of the EMODnet Seabed Habitats mapping portal displaying EUSeaMap, a broad-scale seabed habitats map covering all European seas, available for public download.

Contents of the EMODnet Seabed Habitats portal include:

- The EMODnet broad-scale seabed habitat map for Europe (called EUSeaMap) – A predictive habitat map derived from physical data, covering over 23 million km2 of European seas, including North East Atlantic waters from Norway and Iceland to the Azores and Canary Islands.

- Marine physical data layers – Datasets such as kinetic energy due to waves and currents at the seabed, bathymetry, substrate type and modelled biological zone, that can be used as predictor variables for seabed habitats, as used in making EUSeaMap.
- Detailed habitats maps from survey – A collection of habitat maps at various scales from marine surveys across Europe.
- OSPAR threatened and/or declining habitats – The portal is the official repository for data on OSPAR threatened and/or declining habitats in the North East Atlantic.
- Modelled maps of specific habitats – such as probablility distributions of kelp beds along the cost of Norway.
- Confidence maps – all habitat maps also have associated confidence maps.

Accessing European habitat datasets

The portal is fully interactive: users can view data as map layers, build a custom map, query the layers or filter per habitat. Information can also be downloaded in GIS format to be easily overlaid with other datasets. Using the Web Map Service, all layers can be accessed by personal desktop GIS applications or by other web mapping portals. The Portal uses the ICES metadata catalogue for storing metadata related to all the habitat maps and models.

Contribute your data

Our active data collation is ongoing and we welcome contributions from any source (https://emodnet-ingestion.eu/). Over the coming years we will continue gathering habitat maps from surveys, collating outputs of habitat distribution models, and updating EUSeaMap. Importantly we are working to expand the collation of habitat point data and to implement, for the first time, a newly updated international standard data schema which will allow the submission of habitat/biotope point data to a recognised international database (Ocean Biogeographic Information System).

SeaDataCloud quality control of data collections

Christine Coatanoan, Institut Français de Recherche pour l'Exploitation de la Mer (France), christine.coatanoan@ifremer.fr
Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it
Volodymyr Myroshnychenko, Middle East Technical University (Turkey), volodymyr@ims.metu.edu.tr
Örjan Bäck, Sveriges Meteorologiska och Hydrologiska Institut (Sweden), orjan.back@smhi.se
Helge Sagen, Institute of Marine Research (Norway), helge@hi.no
Serge Scory, Royal Belgian Institute of Natural Sciences (Belgium), sscory@naturalsciences.be
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de
Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), michele.fichaut@ifremer.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl

During the SeaDataNet II (SDN) EU-project, the Quality Control Strategy (QCS) has been implemented and continuously reviewed aiming at improving the quality of the global dataset and creating the best products. This QCS has also been used for the first aggregated dataset provided in SeaDataCloud (SDC). New regional temperature and salinity data collections covering the time period 1900-2017 have been released within the SeaDataCloud (SDC) project in 2018. A general description of these datasets, their data quality assessment procedure and results are presented.

The specific procedure implemented during SDN II allows assuring and certifying the best quality for the datasets (Fig. 1). After the data harvesting from the central CDI catalogue, QC has been performed at regional levels in a coordinated way, using the ODV software (5.0.0) as common and basic QC analysis tool. In SDC the additional checks have been performed per basin to consider the specific water masses characteristics, per instrument type to investigate data completeness and consistency, per data provider to better identify data anomalies.

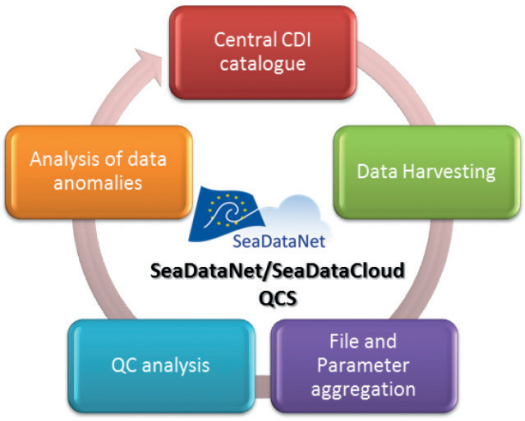


Fig. 1 - Quality Check Strategy implemented during SDN.SDC project.

This QCS allowed to highlight doubtful data and to organize the data anomalies in lists that have been sent to each concerned data originator together with guidelines explaining the expected corrections. The National Oceanographic Data Centers (NODC) have been asked, on the base of those lists, to check and eventually correct the original data and resubmit them in the SDC dataflow. The iterative procedure has been designed to facilitate the update and improvement of SDC database content.

A detailed description of each regional

dataset (Fig. 2) is contained in a Product Information Document (PIDoc): the general products' characteristics (space-time coverage, resolution, format), its quality (validation methodology results) together with experts' recommendations for its usability. ODV qualified dataset collections and PIDocs are available at https://www.seadatanet.org/Products.

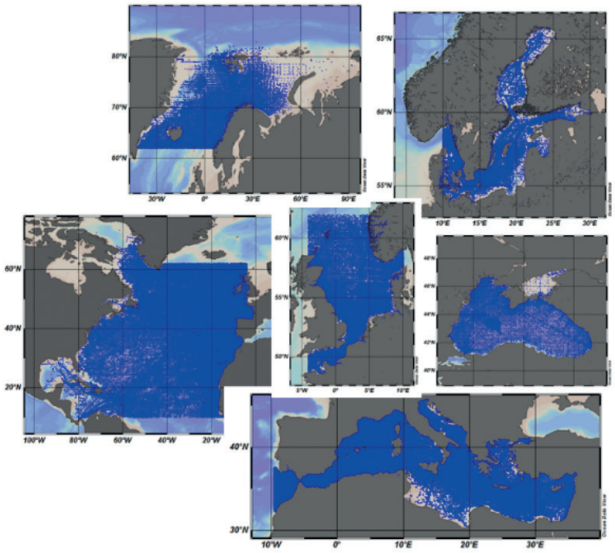


Fig. 2 - Regional data distribution maps of SDC temperature and salinity data collections.

Within SeaDataCloud, the implementation of a cloud environment (Virtual research environment, VRE in Figure 3) aims to optimize and automate the QCS at the central level assuring a continuous monitoring of the database content and its quality. The VRE gives the possibility of generating database snapshots on a regular basis, it facilitates data products versioning and it allows to combine data with subsets from external sources.

The VRE will offer to the users the opportunity to access SDC data and services in the cloud thus providing the possibility of generating their own temperature and salinity data products as well as products for other parameters.

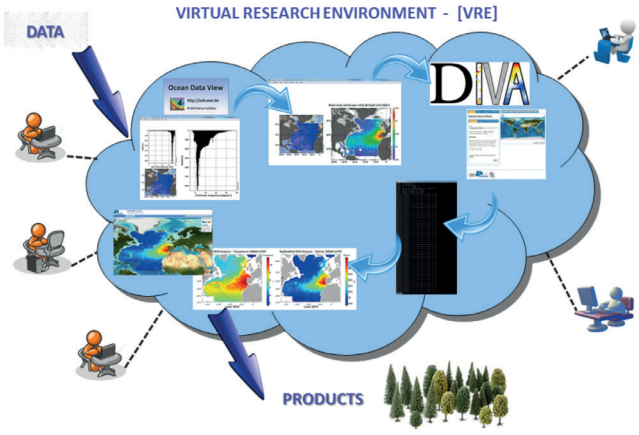


Fig. 3 - Virtual Research Environment.

Real-time temperature and salinity quality control based on minimum/maximum estimates from the known local variability

Jérôme Gourrion, OceanScope (France), jerome.gourrion@ocean-scope.com
Tanguy Szekely, OceanScope (France), tanguy.szekely@ocean-scope.com
Christine Coatanoan, Institut Français de Recherche pour l'Exploitation de la Mer (France), christine.coatanoan@ifremer.fr

In the past years, delayed-time quality control (QC) procedures of temperature and salinity measurements at the French CORIOLIS facility have improved significantly thanks to the implementation and use of new methods with reduced statistical assumptions. In the context of the Copernicus Marine Environment Monitoring Service, this success led us to introduce the same concept into real-time processing.

With such an approach, observations are compared to the known local variability through validity intervals built from historical estimates of minimum and maximum values of the parameter of interest. No a priori assumption on the local parameter distribution shape is required, and natural skewness and kurtosis can be accounted for during the detection process.

In a delayed-time context, such a QC procedure is used to raise alarms that an operator will then visualize and potentially confirm. Thus, an improved automatic detection procedure essentially allows saving operator time through reduction of the number of false alarms.

In a real-time or near-real-time context, the available operator time is much reduced or even null. In order to implement such a procedure in an operational chain, it is fundamental to have a fine control of the total number of alarms. The method shall be adjusted to raise a manageable number of alarms, allowing small anomalies to pass through the filter while ensuring that the largest ones are systematically caught, being more likely associated to gross observation errors.

Consequently, such a QC procedure needs to be adapted and tuned before implementation in real-time processing. This presentation will focus on the design and the validation results conducted at CORIOLIS in 2018.

Nippon Foundation-GEBCO Seabed 2030 project - aiming to map the global ocean floor by 2030

Pauline Weatherall, British Oceanographic Data Centre (BODC) (UK), paw@bodc.ac.uk
Satinder Bindra, Nippon Foundation-GEBCO Seabed 2030 Project Director, satinder.bindra@seabed2030.org
Boris Dorschel, Alfred Wegener Institute (Germany), boris.dorschel@awi.de
Vicki Ferrini, Lamont-Doherty Earth Observatory, Columbia University (USA), ferrini@ldeo.columbia.edu
Martin Jakobsson, Stockholm University (Sweden), martin.jakobsson@geo.su.se
Geoffroy Lamarche, National Institute of Water and Atmospheric Research and School of Environment, University of Auckland, Auckland (New Zealand), geoffroy.lamarche@niwa.co.nz
Larry Mayer, Center for Coastal and Ocean Mapping, University of New Hampshire (USA), larry@ccom.unh.edu
Helen Snaith, British Oceanographic Data Center (United Kingdom), h.snaith@bodc.ac.uk

Despite many of years of effort, less than 20 per cent of the global seafloor has been mapped using modern echo-sounding techniques. The percentage depends on the target mapping resolution. This greatly diminishes our ability to explore and understand critical ocean and seafloor processes.

To help address these issues, the General Bathymetric Chart of the Oceans (GEBCO) and the Nippon Foundation of Japan have collaborated to create the Nippon Foundation-GEBCO Seabed 2030 Project, an international effort with the aim to bring together all available bathymetry data to produce the definitive map of the world ocean floor by 2030 and make it available to all.

The project was launched at the United Nations (UN) Ocean Conference in June 2017 and is aligned with the UN's Sustainable Development Goal #14 to conserve and sustainably use the oceans, seas and marine resources.

To achieve its goal, the Seabed 2030 Project will work and collaborate with international organizations, mapping initiatives, the scientific community and maritime industries to bring together existing data and help to prioritize new survey operations through highlighting gaps in existing data coverage.

Seabed 2030 Structure

The project sits within the existing International Hydrographic Organization (IHO)-Intergovernmental Oceanographic Commission (IOC) of UNESCO GEBCO framework and greatly benefits from the input and work of GEBCO's committees and working groups.

Seabed 2030 has four Regional Centers. Each Center has a dedicated team of experts who are responsible for championing mapping activities; assembling and compiling bathymetric information and collaborating with existing mapping initiatives within their region. A Global Center is responsible for assembling and delivering centralized GEBCO products. The work is led by a Project Director.

The Regional and Global Centers will not act as a permanent archive or distributor of the data they are assembling, i.e. they will not be duplicating the responsibilities of existing international and national data centres or regional data compilation efforts but will work with these organisations to contribute and access data.

Seabed 2030 will maintain a strong working relationship with the IHO Data Center for Digital Bathymetry, hosted by the US National Centers for Environmental Information, USA, which is the recognized IHO repository for all deep ocean bathymetric data (deeper than 100m) and recommend it as a data archive to potential data contributors.

Seabed 2030 Centers:

- Southern Ocean - hosted at the Alfred Wegener Institute (AWI), Germany
- South and West Pacific Ocean - hosted at the National Institute of Water and Atmospheric Research (NIWA), New Zealand
- Atlantic and Indian Ocean - hosted at the Lamont-Doherty Earth Observatory (LDEO), Columbia University, USA,
- Arctic and North Pacific Ocean - hosted at Stockholm University (SU), Sweden and the Center for Coastal and Ocean Mapping (CCOM) at the University of New Hampshire (UNH), USA
- Global Data Center - hosted at the British Oceanographic Data Centre (BODC), NOC, UK

As well as assembling and processing existing bathymetric data sets, the Regional Centers will also work to identify gaps in our current data coverage and look at new data collection opportunities.

At what resolution will Seabed 2030 map the ocean floor?

Seabed 2030 will map the ocean floor at the best possible resolution within practical limits. However, gathering bathymetric data gets more difficult as the ocean gets deeper. Taking this into account, Seabed 2030 we will establish a depth-variable resolution goal and consider the seafloor "mapped" if at least one sounding falls in a grid cell of the size described in Table 1.

Table 1 - Feasible resolution based on state-of-the-art 2°×2° deep-water multibeam installed on surface vessel, calculated at 60° from nadir.

Depth Range	Grid-Cell Size	% of World Ocean
0-1500 m	100 × 100 m	13.7
1500-3000 m	200 × 200 m	11
3000-5750 m	400 × 400 m	72.6
5750-11,000 m	800 × 800 m	2.7

Marine data incorporated into experimental hardware and software system for climate change monitoring and forecast in Svalbard and the western Arctic zone of the Russian Federation

Anastasiya Vyazilova, Arctic and Antarctic Research Institute (Russia), vae@aari.ru
Alekseev Genrikh V., Arctic and Antarctic Research Institute (Russia), alexgv@aari.ru
Alexander Smirnov, Arctic and Antarctic Research Institute (Russia), Arctic Portal (Iceland), avsmir@aari.ru

Experimental hardware and software system for climate change monitoring and forecast in Svalbard and the Western Arctic zone of the Russian Federation was being developed in 2014–2016 at Arctic and Antarctic Research Institute. The system was developed under a project of Ministry of Education and Science of the Russian Federation titled: "Development of new methods and hardware for hydrometeorological and geophysical monitoring in Svalbard and the Western Arctic zone of the Russian Federation".

The core of the system is a specialized database (DB) for assessment and forecast of climate changes and extreme weather phenomena. Existing hydrometeorological data resources were used for DB developing such as oceanographical, meteorological, sea ice data, historical datasets, elements of Svalbard cryosphere, maps and plots of extreme weather phenomena occurence, extremum values, etc. This data is categorized into the following groups: Oceanography, Meteorology, Sea ice and Cryosphere.

Marine data

The Oceanography section includes data sets of temperatures and salinities in the region of the Barents and the Kara Seas and the neighboring part of the Norwegian and the Greenland Seas (63°–83° N, 0°–100° E), monthly arrays of gridded sea surface temperatures derived from both HadISST and ERA Interim reanalyses, monthly sea surface temperatures for the period 1948 - 2071 based on model results obtained from Institute of Numerical Mathematics Ocean, extreme current speed frequencies in the Svalbard region and extreme sea level frequencies at the shores of the Barents and the Kara Seas.

Gridded arrays of water temperatures and salinities at standard level depths were calculated based on objective analysis of oceanographic data for the area 63 - 83° N, 0 - 100° W (299077 stations univenly distributed in space and time) (Fig. 1).

Yearly distribution of oceanographic stations (Fig. 2) shows that the majority of measurements were made in 1950–1990. All collected data passed through several stages of quality control. Spatial interpolation was performed by using DIVA (Data-Interpolating Variational Analysis) resulting in creation of a set of gridded fields of temperature and salinity (0.25° X 0.25°, 29 standard levels). The following fields were calculated and included into the system: 30-year

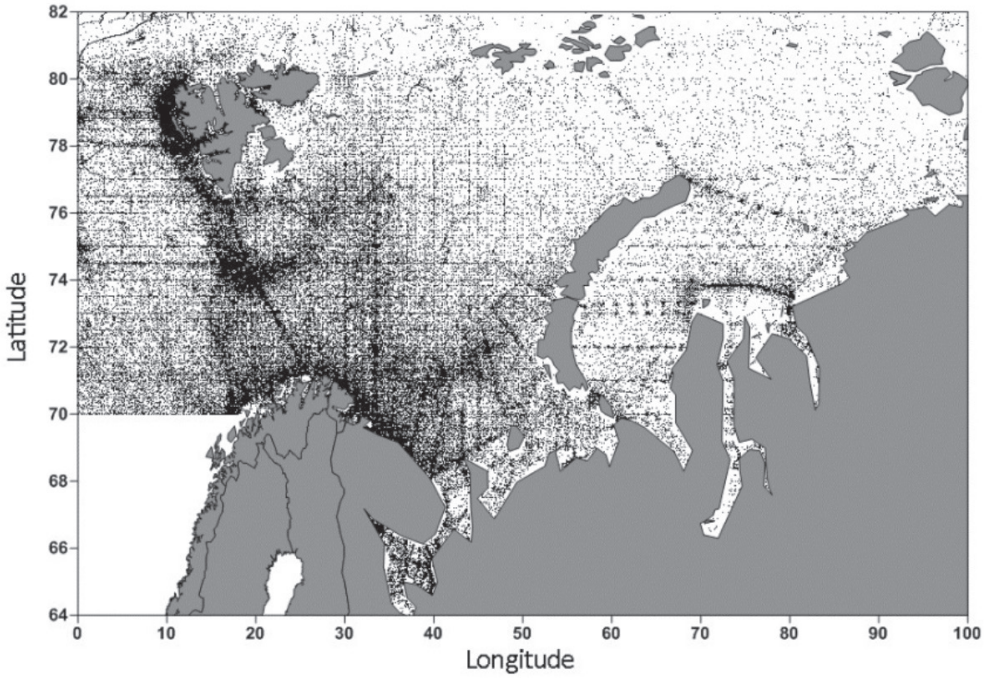


Fig. 1 - Spatial distribution of oceanographic stations.

summer period average fields (1900-1930, 1931-1960, 1961-1990, 1991-2014), 10-year average fields (1900-1910, 1911-1920 etc.), average for 1900–2014 and monthly fields for 1900-2014.

Such parameters as adverse wind and wave occurrence, spray freezing of ships, maximum surface current speeds, sea ice conditions in Barents and Kara Seas were also a part of the Oceanography section. The experimental hardware and software system for climate change monitoring and forecast includes archives of meteorological data (measurements, reanalysis, forecast), sea ice data, ice cap and snow cover data at Svalbard.

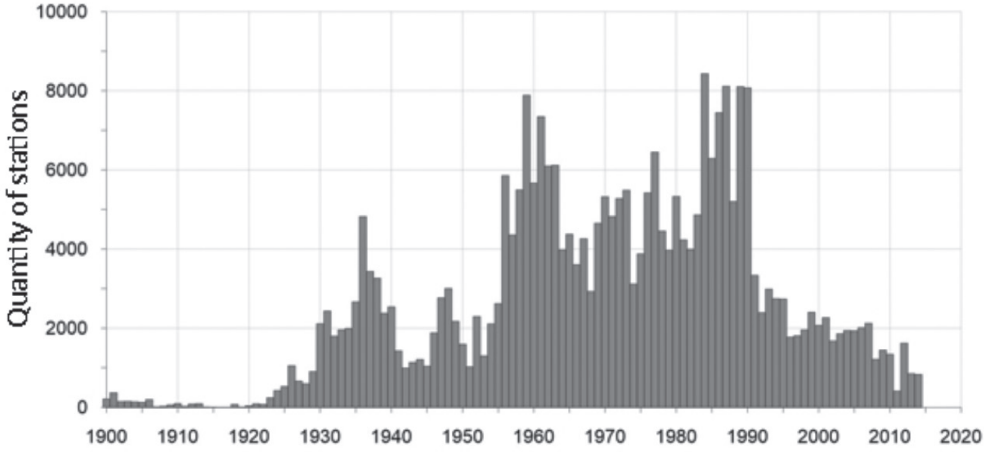


Fig. 2 - Yearly amount of oceanographic stations.

Functions and capabilities of the experimental hardware and software system

The hardware part of the system sonsists of a server, network automated database storage and a set of personal computers linked altogether into a network. The software part has the following modules: control module, module of visualization, storage module, module of representation. The control module has a series of capabilities: selection of quantitive values of climatic parameters, climate forcast of sea surface temperature, air temperature, ice concentration and ice extent, formation of extreme weather parameters time series, calculation of surface components vulnerability. The experimental hardware and software system has a user-friendly interface that allows quick access to data and data visualization by module of representation (ESRI ArcGIS, Golden Software Grapher etc.).

The experimental hardware and software system provides: automated generation of maps and plots of climatological parameters, automated calculation of predicted values of main climatological parameters in the Svalbard region, automated obtaining of extreme characteristics time series, automated calculation of ground components vulnerability, control of data visualization in ArcView 10.1.

The output of the experimental hardware and software system is climatic data represented in maps, plots and tables. This information provides an insight into climate change dynamics in the region. The data resources incorporated into the system allow to assess observed and predicted climate changes and extreme weather conditions in the region.

The Scottish Coastal Observatory

Jenny Hindson, Marine Scotland Science (United Kingdom), j.hindson@marlab.ac.uk
Barbara Berx, Marine Scotland Science (United Kingdom), b.berx@marlab.ac.uk
Sarah Hughes, Marine Scotland Science (United Kingdom), sarah.hughes@prospect.org.uk
Pamela Walsham, Marine Scotland Science (United Kingdom), pamela.walsham@gov.scot
Margarita Machairpoulou, Marine Scotland Science (United Kingdom), margarita.machairpoulou@gov.scot
Eileen Bresnan, Marine Scotland Science (United Kingdom), eileen.bresnan@gov.scot
Bill Turrell, Marine Scotland Science (United Kingdom), bill.turrell@gov.scot

Long term time series are extremely valuable in understanding how the marine ecosystem in coastal waters functions and to describe long term environmental change. Scientists in Marine Scotland Science (MSS), along with a group of volunteer samplers, have been monitoring the physics, chemistry and biology at a number of sites in Scotland's coastal waters since 1997. This monitoring programme is called the Scottish Coastal Observatory (SCObs).

A report providing a basic description of the seasonality and variability of the main parameters, including temperature, salinity, nutrients, carbonate chemistry, chlorophyll 'a', phytoplankton, algal toxins and zooplankton, examined between 1997 and 2013, has been published (https://data.marine.gov.scot/dataset/scottish-coastal-observatory-1997-2013-parts-1-3). A dataset of monthly means has also been published (doi: 10.7489/1761-1) and in the summer 2018 the entire SCObs dataset will be published at a higher temporal resolution. These data are being used to fulfil the monitoring requirements of the Marine Strategy Framework Directive, the Water Framework Directive, and are freely available to the public. Analytical quality is ensured by accreditation under the data quality standard ISO 17025 or Joint Code of Practice. While the data quality of each SCObs parameter is assessed separately, they are also assessed as a whole before a quality flag assigned to each data point. The flagging system uses the SeaDataNet QC flag scale, which evaluates the data, but no changes are made to the original data values. The process taken to assign these quality flags, by combined oceanographic, chemistry and plankton expert criteria, will be discussed.

The dataset reveals regional differences in Scotland's coastal environment and ecosystem, with temperature, salinity and plankton community differences observed between sites. A high degree of interannual variability has also been observed in the time series.

The SCObs dataset will contribute to the Interreg Atlantic Area MyCOAST project, which aims to harmonise coastal monitoring programmes across the European Atlantic Area and apply common coastal risk management tools.

Semantic fisheries data integration and analytics

Aileen Kennedy, National University of Ireland Galway (Ireland), a.kennedy28@nuigalway.ie
David Currie, Marine Institute (Ireland), david.currie@marine.ie
Enda Howley, National University of Ireland Galway (Ireland), enda.howley@nuigalway.ie
Jim Duggan, National University of Ireland Galway (Ireland), jim.duggan@nuigalway.ie

Background

The Marine Institute generates and consumes data from a number of different sources. In particular, Fisheries Ecosystem Advisory Services (FEAS) maintains databases containing information on:

- *catch value and volume*, including commercial landings, fishing effort and fleet capacity;
- *biological data* - which includes data on variables such as the number, length, weight, sex and age of fish species in a given location.

Fine-scale spatio-temporal data about vessel position is also stored. The fisheries data typically becomes available in batches rather than being streamed – these batches can range in size from an individual fishing trip, up to all vessel positions in a calendar year. These databases are typically siloed, so querying for information that are stored across many databases can be a significant challenge.

Semantic Integration and Analysis

This project aims to semantically integrate and analyse fisheries and marine data - it will use a pipeline approach to data analytics, whereby heterogeneous marine data sources can be combined within an analytics framework (e.g. R), so that key hypotheses can be explored and evaluated by decision makers.

This system will allow queries to be expressed in terms of the items of interest rather than requiring structural knowledge of the underlying databases. This semantic layer will add extra querying power by allowing inferences to be made that are difficult or impossible to process using traditional queries to the underlying relational model. It will also allow the work to be more transferable since it will be based on meaning, rather than a specific database structure.

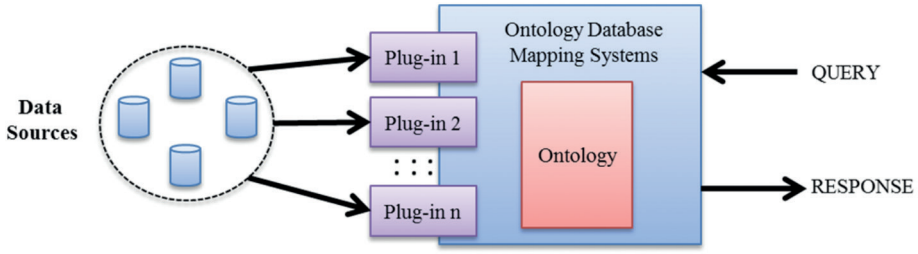


Fig. 1 - A possible mapping between a semantic model and relational databases.

Individual data sources have their own characteristics with respect to noise, accuracy and completeness and data will also be available at different aggregation levels. Techniques for combining data sources must be able to incorporate these features and use them to inform the outputs. The semantic system designed must be extendable, both to facilitate sharing with other researchers and to allow new data sources to be added easily to the system as the project progresses.

The key outputs of the project will be:

1. A novel, open, generic framework for semantic data integration so that heterogeneous marine data sources can be integrated in a meaningful way.
2. An algorithm that will map between semantic queries and different underlying data structures that improves on the performance of the current state-of-the-art whilst retaining expressiveness.
3. A pilot analytics engine to allow the semantic model to be used to answer policy and decision support questions. This engine could use a number of different techniques such as machine learning, simulation, or game theory.

Discussion

This presentation will discuss the progress that has been made on the project and present initial results - these will be targeted at modelling the biological measurements of a fish using a suitable ontology and mapping semantic queries in that domain to the data stored in relational databases.

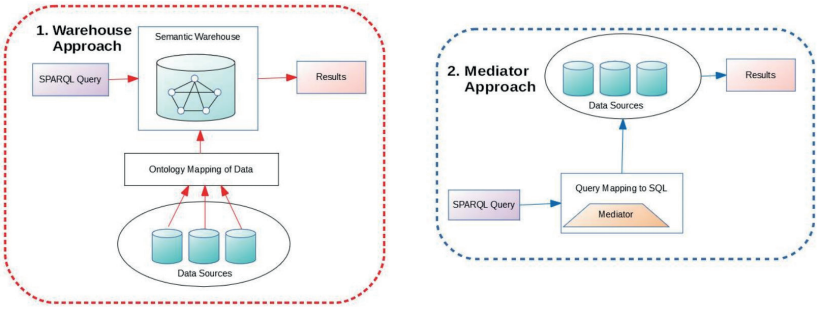


Fig. 2 - Approaches for data integration.

Two different approaches to data integration will be designed and evaluated. In the first approach, the Warehouse Approach, the data is extracted from the relational databases, transformed using the ontology and stored in a semantic warehouse. The semantic query language SPARQL is used to query the semantic warehouse directly to obtain the results. In the second approach, the Mediator Approach, the SPARQL query is mapped, using an ontology, to SQL - the relational databases are then queried using the SQL query. The two approaches are evaluated with respect to defined performance criteria with the aim of providing a semantic structure to support the design of the analytics engine.

References

Currie D., Howley E., and Duggan J. (2016) *A Data Analytics Framework for Ecosystem-Based Fisheries Management.* ICES Annual Science Conference 2016

Tzitzikas Y. *et al.* (2013) *Integrating Heterogeneous and Distributed Information about Marine Species through a Top Level Ontology.* In: Garoufallou E., Greenberg J. (eds) Metadata and Semantics Research. MTSR 2013. Communications in Computer and Information Science, vol 390. Springer, Cham.

Advances in Spanish marine data availability in the framework of EMODNET data ingestion EU-project

Elena Tel, Instituto Español de Oceanografía (Spain), elena.tel@ieo.es
Gemma Ercilla, Instituto de Ciencias del Mar CSIC (Spain), gemma@icm.csic.es
Javier Valencia, Lyra Ingenieria (Spain), javi.valencia.m@gmail.com
Belen Alonso, Instituto de Ciencias del Mar CSIC (Spain), belen@icm.csic.es
Irene Chamarro, Instituto Español de Oceanografia (Spain), irene.chamarro@ieo.es

The Spanish coast extends for 7879 km in which 58% of the Spanish population lives, almost 23 million people. In addition, Spain has more than 1 million km2 of marine waters, more or less twice the land area of the country. These characteristics favor that an important part of its economic activity is linked to the use of the resources that the sea offers, and generate a valuable amount of marine information. However, these data are scattered in private sectors and public administrations.

EMODNET Data Ingestion is an EU-funded iniciative that promotes the incorporation of marine data into databases for publishing as open data and contributing to applications for society. Marine data remains stored and dispersed throughout the European Union countries, and provide a legacy of continuing usefulness and importance. Their secondary use increases the knowledge and understanding of the marine environment. Promotion of this data accessing and sharing has different responses depending on their confidentiality issues or the lack of them. Already ingested, data show the variated information that will be reused in the framework of EMODNET iniciatives, but also in different aproaches that are not previously though.

As an example of unexpected uses, the coastal temperature timeseries of the El Bocal aquaculture facility on the North coast of Spain, are nowadays being crossed with medical data in the search for environmental factors that can affect health. Also, those timeseries can be useful in climatic variability assessments.

The bathymetric and geological data from old research campaigns contribute to improve the knowledge of the seafloor nearsurface, allowing a better approach to future local and regional studies about global climate change, pollution control and assessment, slope stability etc. Aditionally, those data contribute tothe implementation of improved digital terrain models that can be used, for example, to perform useful simulations in tsunami warning systems or other coastal risks.

As expected, all these ingested data are evaluated, metadata to preserve their authorship and favor their reuse. Finally, they are incorporated into European databases in internationally accepted formats and depending on their discipline used for the development of EMODNET products: bathymetries, distribution maps of different ocean variables, climatologies and trends, among others. In consequence the effort to promote EMODNET-DI is valuable, although the response may be uneven.

Knowledge transfer: the key to creating societal and economic benefits from marine observations

Jan-Bart Calwaert, European Marine Observation and Data network (Belgium), janbart.calewaert@emodnet.eu
Oonagh McMeel, Seascape Consultants Ltd (Belgium), oonagh.mcmeel@seascapeconsultants.co.uk
Belén Martín Míguez, European Marine Observation and Data network Secretariat (Belgium) belen.martin-miguez@emodnet.eu

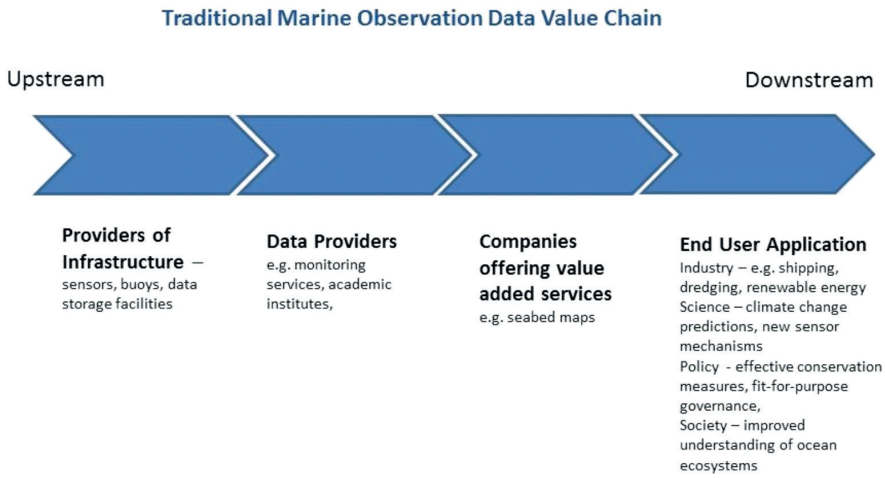


Fig. 1 - The marine observation data value chain.

An effective marine observation and data-sharing system, delivering societal and economic benefits requires the coordination of efforts between multiple sectors in the value chain. These include the scientific community, oceanographic data centres, federated data infrastructures, national and regional agencies and authorities with competency for marine environment and maritime economy, actors from civil society and the private sector.

Key to this change is effective knowledge transfer at all stages of the marine observations to user application value chain, from the development and deployment of innovative new sensing technologies to the application of marine data and products by stakeholders from the public and private sectors and by civil society.

Capturing the Data: Innovative Marine Sensing Technologies

There is a need for a concerted effort to create a supportive environment to stimulate progress along the maritime sensing technology value chain from identification of requirements, research and development up to full market uptake and application by end-users from public sector, research, business and society as whole. This is important because there are significant opportunities for

Europe to develop a dynamic and thriving market for maritime sensing technologies. Advances in sensor technology will allow us to do more and do it more cheaply. This will strengthen our research capacity, increase the confidence in scientific outputs and improve our understanding of marine environmental processes and the impacts of climate change. Robust, low cost, long-life ocean sensing and observing technologies will also improve our capacity to sustainably manage maritime activities. This is crucial for delivering important EU policy objectives such as progress towards Good Environmental Status under the Marine Strategy Framework Directive.

Marine data sharing and utilisation

The mere collection, safeguarding and sharing of marine observation and monitoring data provides huge societal benefits. Data and information on the state and variability of the marine environment is crucial for understanding changes that may result from human activity, including the effects of human-induced climate change and ocean acidification. Long-term time series are particularly valuable to support both scientific research to elucidate the causes, drivers and impacts of environmental change and, in turn, evidence based policy making. Moreover, they are invaluable for establishing the baselines for accurate resource assessment, essential for spawning private initiative in sectors such as marine renewable energy or aquaculture. Marine data also feed into the provision of ocean forecasts and reanalysis such as those delivered in the Copernicus Marine Service (CMEMS). Access to accurate and adequate data underpins the implementation of the Marine Strategy Framework Directive and supports the implementation of the Maritime Spatial Planning Directive.

However coastal and ocean observatories and public data-sharing initiatives face common challenges in their efforts to unlock the full societal and economic potential of the wealth of European marine data and observations at European, national, regional or local level and demonstrating their use and positive contribution to sustainable blue growth.

Changing the status quo: Challenges, solutions and ways forward

This poster will highlight some of the key challenges along the entire value chain, from the development and operational use of novel marine sensing technologies to the application of marine data and products by diverse stakeholders and propose some mitigating solutions to these, drawing from the experience of EMODnet as well as work carried out in the framework of the H2020 AtlantOS and COLUMBUS projects. These include: more effective and targeted stakeholder engagement to move beyond the traditional players in the marine observations landscape; mapping potential users and ensuring that products meet their needs by involving them in product development; understanding the policy landscape and ensuring that product development is fit-for-purpose in this context and ensuring that marketing and brokerage is considered as an important aspect of bringing products to market.

For this potential to be realized, it is important to ensure that the significant investment (both at EU and national level) in sensor technology development, monitoring programmes and data-sharing initiatives is targeted at the right areas, is focused on addressing real-life marine monitoring challenges and opportunities, and that new intellectual property (IP) generated is taken up to drive towards competitive and marketable sensor technologies and state-of-the art data-sharing initiatives to realise Europe's full potential as global leader in marine observations and data sharing.

Benefits of interpreted vector programming and hierarchical data format for statistic ocean model evaluation

Paolo Oliveri, Istituto Nazionale di Geofisica e Vulcanologia (Italy), paolo.oliveri@ingv.it
Simona Simoncelli, Istituto Nazionale di Geofisica e Vulcanologia (Italy), simona.simoncelli@ingv.it
Alessandro Grandi, Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy), alessandro.grandi@cmcc.it
Emanuela Clementi, Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy), emanuela.clementi@cmcc.it

Big Data world is increasing incredibly fast, as High Performance Computing is near to overcome the Exascale and Artificial Intelligence techniques are invading our lives in new and different manners every day.

From the ocean models point of view, this means that now we have the capacity to build and run newer, more resolute and accurate models, using more or less the same time that we used in the past with less resources; however, this causes the production of more complicated and big datasets that anyway have to be analysed and evaluated.

From the observations point of view, this means that we can install, use, maintain and store very huge amounts of devices, with the same resources that we used in the past; unfortunately, the increasing number of recorded devices data and the temporal dilatation of the time series is producing very huge datasets, even if the single device data is pointwise.

Further, ocean recorded data are not completely reliable yet, because marine environment is not the ideal condition where electronic can work long and optimally, causing sampling discontinuities and needing frequent maintenance and sensors' calibration.

In order to make these two different worlds interact we took advantage of:

1) Interpreted language (for ease of use and portability) -> Python 3.x;
2) Vectorized numerical data analysis (for flexibility and speed) -> NumPy;
3) Output data storage (for efficiency and standardization) -> NetCDF-4 (HDF-5 extension);
4) Software maintenance and update (for security and speed) -> Git,

All of the selected software is free of charge and open source.

A Python module has been set up to automatically do:

1) Data quality assessment of a set of sea observations, using original quality controls (if available), spike removal (gross check) and redundant statistic quality check, the last one by computing standardized anomaly and the probability distribution (kernel density estimation), finally returning the best possible hourly and daily time series for the analysts' purposes;
2) Extraction and aggregation in time of ocean model time series at the horizontal nearest

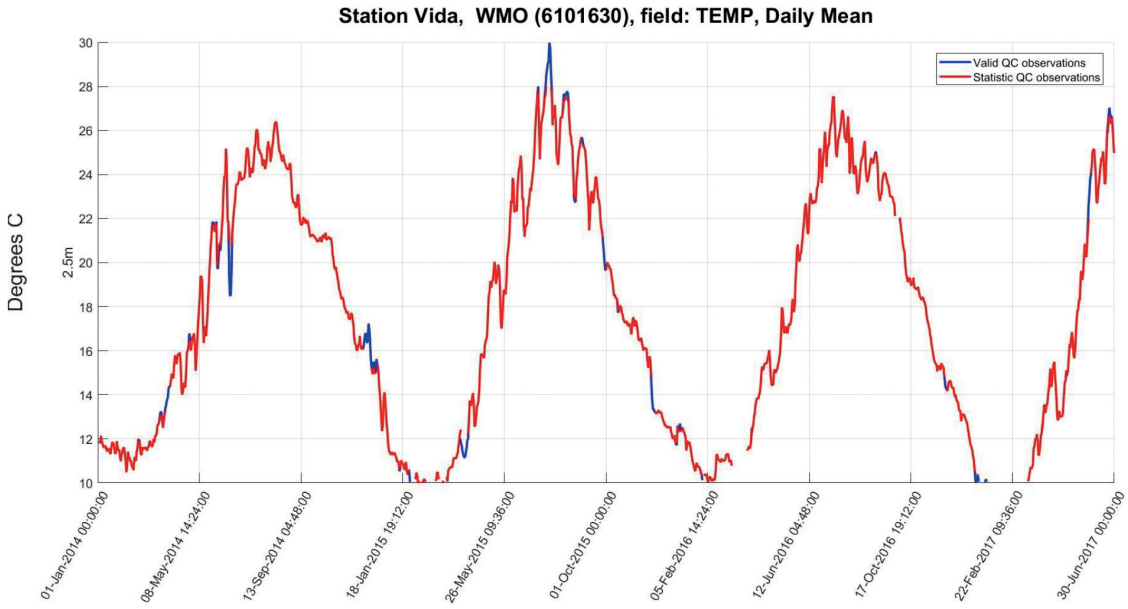


Fig. 1 - Difference of daily mean post processed insitu data of station Vida using original (blue line) versus 3 iterations statistic quality flag (red line) application.

point to the observations, and then interpolating the model depth levels to the observations depth levels;

3) Running evaluation methods in terms of RMSE (root mean square error) and statistical bias, both for each device and for the Mediterranean standard Copernicus regions.

In order to create and set up the system for the first time we have chosen:

1) Water temperature, salinity, sea level, and 3D speed as test variables;
2) Fixed observatories data (moorings) located in the Mediterranean Sea downloaded from the European Copernicus Marine Environment Monitoring Service (CMEMS) in situ TAC (Thematic Assembly Center) Med FTP server (the data is accessible after free registration to Copernicus marine service portal);
3) Two model data from the RITMARE Italian project and from CMEMS.

The results are promising, both from the scientific and the technic point of view:

- An historical evaluation can be processed in acceptable times and it can be easily repeated on any machine equipped by the necessary portable software;
- A real time - production evaluation can be as easily set up as an “one off” one;
- Taking into account of the internal quality data processing of the observations, the evaluation results can provide more reliable and model helpful skill scores.

T-MEDNet observation network and resource platform on climate change effects in Mediterranean coastal ecosystems

Nathaniel Bensoussan, ICM Consejo Superior de Investigaciones Científicas (Spain), nbensoussan@gmail.com
Joaquim Garrabou, ICM Consejo Superior de Investigaciones Científicas (Spain), garrabou@icm.csic.es
T-MEDNet Network*

Context: adaptation pressure on Mediterranean coastal ecosystems

Climate change (CC) is having an increasing influence on the abundance, distribution and wellbeing of living organisms worldwide. In the Mediterranean, Sea Surface Temperature (SST) warming is obvious and ecosystems are already impacted by climate change, raising concerns for the conservation of essential marine coastal habitats. The increasing occurrence of mass mortality events affecting the benthic biota and success of thermophilic biota colonizing the Mediterranean Sea are some clear examples of ongoing consequences of environmental changes. Marine Protected Areas (MPAs) are natural infrastructures for the conservation of marine biodiversity. However, the development of sound management and adaptation plans is often constrained by the lack of information at the appropriate time and space scales. In particular, regarding the strong variability, both spatial and temporal, of hydrological conditions in nearshore and coatal areas.

Enhance monitoring and understanding of climate change effects

T-MEDNet initiative is devoted to develop an observation network on climate change effects in marine coastal ecosystems, firstly by spreading the implementation of standard monitoring protocols on seawater temperature and biological indicators over the long-term. This is achieved through fostering international cooperation between marine scientists and MPA managers working in the coastal zone at Mediterranean scale (Fig. 1).

In T-MEDNet, in situ temperature is continuously sampled at hourly frequency using data loggers deployed at standard depth levels, generally every 5 m down to 40 m depth or more. To date, the network counts 30+ members conducting observations in 40+ sites, since 1999 for the longest series, resulting in 12+ million in situ T samples. This dataset allows for the very first time to draw robust base lines on thermal regimes and track signals of change in coastal areas in different Mediterranean sub-basins. These data are being key to analyze biological responses to warming, considering episodic events, in particular the onset of mass mortality events, but also changes in distribution and phenology.

* T-MEDNet network co-authors: Linares C., Cebrian E., Kipson S., Teixido N., Romano J.C., Ramos-Esplà A., Kersting D.K., Riera i Arago V., Payrot J., Blouet S., Bachet F., Charbonnel E., Zuberer F., Pairaud I., Sartoretto S., Bergere H., Peyrache M., Cappanera V., Dominici J.M., Negre-Santucci N., Santoni M.C., Bursic M., Harmelin J.G., Ruiz J.M., Ledoux J.B., Zabala M., Giardina F., Topcu E., Ozturkb B., Sini M., Bensoussi J., Ghanem R., Ocaña O., Reñones O., Díaz D., Kızılkaya Z. *et al*.

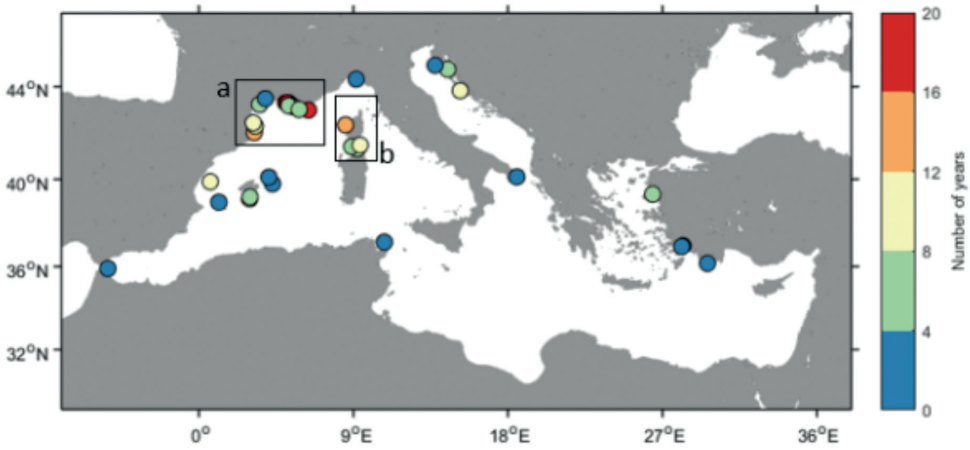


Fig. 1 - Location map on T-MEDNet multiyear temperature time series in the Mediterranean Sea. Colors show the length of the time series (in year) as indicated by the color bar.

T-MEDNet platform on coastal warming and biological impacts

T-MEDNet network and resource platform functionalities (www.t-mednet.org) ensure data quality check and management, data exploration, periodic reporting to members and public information display for stakeholders and sea practitioners. The obtained in situ baselines and trends are being used to complement classical CC approaches based on satellite derived SST. In the framework of the MPA-ADAPT project (https://mpa-adapt.interreg-med.eu/), data and knowledge transfer is being conducted for the development of sound adaptation plans for Marine Protected Areas face to climate change.

Bjerknes Climate Data Centre

Benjamin Pfeil, Bjerknes Climate Data Centre, University of Bergen and RI Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (Norway), benjamin.pfeil@gfi.uib.no

Steve Jones, Bjerknes Climate Data Centre, University of Bergen and RI Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (Norway), steve.jones@uib.no

Camilla Stegen Landa, Bjerknes Climate Data Centre, University of Bergen and RI Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (Norway), Camilla.Landa@uib.no

Jonas Fagnastol Henriksen, Bjerknes Climate Data Centre, University of Bergen and RI Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (Norway), Jonas.Henriksen@uib.no

Rocio Castano Primo, Bjerknes Climate Data Centre, University of Bergen and RI Integrated Carbon Observation System (ICOS) Ocean Thematic Centre (Norway), Rocio.Primo@uib.no

The Bjerknes Climate Data Centre (BCDC) of the Bjerknes Centre for Climate Research was initiated in 2014 and is hosted at the University of Bergen, Norway. BCDC has the following aims:

1. To serve the data obtained, and data syntheses (SOCAT, GLODAP, DATED) assembled in international collaboration with researchers from the Bjerknes Centre for Climate Research cluster. All data from the different disciplines (e.g. geology, oceanography, biology, model output community) are made Findable, Accessible, Interoperable and Re-usable (FAIR) and made publicly available by the BCDC. It will, however, be open for all interested scientists independent of institution.

2. To provide an online access portal to climate data of all kinds. Currently is in contact with the worldwide data access network established by the International Council for Science called World Data System (ICSU WDS, former ICSU World Data Centre System, http://www.icsu-wds.org/) to become the first partner in Norway. BCDC follows the idea of having a common globally interoperable distributed data system within the climate (change) community. BCDC also cooperates with data centres that are not part of the ICSU-WDS, (prominent examples are NMD and NORSTORE), in order to use these for storage, and to provide access to their data holdings.

3. Data management services for high impact research infrastructures. BCDC hosts data management services for the marine part of the European Research ICOS (Integrated Carbon Observation System Ocean Thematic Centre (ICOS OTC), initiates a Global Data Assembly Centre for Marine Biogeochemistry with US colleagues and contributes to Copernicus Marine Environmental Monitoring Services. In addition BCDC has its main expertise in the field of marine biogeochemistry and heavily supports the Surface Ocean CO2 Atlas (SOCAT) and the Global Ocean Data Analysis Project (GLODAP). BCDC also supports internation bodeis liek IOC UNESCO in addressing data management service towards the Sustainable Development Goal 14.3 (minimize and address the impacts of ocean acidification, including through enhanced scientific cooperation at all levels) by the the United Nations

The Bjerknes Climate Data Centre was acknowledged to be one of the leading and most functional data centres in its field by the Research Council of Norway. This poster will highlight the international interconnection, functionality and services provided.

Eutrophication and contaminants Black Sea data management in the framework of EMODnet Chemistry

Luminita Buga, National institute for Marine Research and Development "Grigore Antipa" (Romania), lbuga@web.de

George Sarbu, National institute for Marine Research and Development "Grigore Antipa" (Romania), gsarbu@alpha.rmri.ro

In order to unlock fragmented and hidden European marine data resources, to improve Europe's marine data infrastructure, increase the availability of high quality data and assemble them under a common framework and to make these available to individuals and organisations (public and private), DG-Mare launched in 2009 a new initiative: European Marine Observation and Data Network (EMODNet) as proposed in the EU Green Paper on Future Maritime Policy. In present EMODNet (www.emodnet.eu) provides access to marine data and derived data products from eight thematic portals: bathymetry, geology, sea bed habitats, chemistry, biology, physics, human activity and coastal mapping.

Through a stepwise approach, EMODnet Chemistry (www.emodnet-chemistry.eu) aims to collect, standardize, check the quality of data developing new services to share and visualize information and products at the scale of regions and sub-regions defined by the Marine Strategy Framework Directive.The third phase of EMODnet Chemistry partnership involves 45 institutes from 27 countries and 3 international organisations (ICES, Black Sea Commission, UNEP/MAP) from all European Seas.

The Black Sea is one of the regional seas in EMODNet Chemistry. Black Sea is a unique marine environment representing the largest land-locked basin in the world. Its waters are in a state of almost complete isolation from the world ocean, as a result of the restricted exchange with the Mediterranean Sea through the Turkish Straits System. As a result, the basin is almost completely anoxic, containing oxygen in the upper 150 m depth (13% of the sea volume) and hydrogen sulphide in the deep waters. (Özsoy and Ünlüata, 1997).

Data products on eutrophication (Nutrients, Chlorophyll, Oxygen) and ocean acidification will be presented, highlighting the availability of historic and present time data and problems encountered with the datasets from contributors.

National Institute for Marine Research and Development "Grigore Antipa" - NIMRD, as Regional leader for the Black Sea, receives harvested data collections as "raw' data for further QC and validation and data products.

The MSFD Black Sea area (enlarged with Marmara Sea besides Black Sea and Sea of Azov) data collection contained datasets originated from 24 CDI-partners from 41 data. Out of total number of the Black Sea CDIs, 82% are non-restricted data (SDN License/unrestricted) while 18% are restricted (by negotiation/academic/moratorium).

NIMRD applied established standard quality control to the data which consists in: data format checks, units conversions, broad range control checks to exclude erroneous high values, negatives, handling of LOD, duplication eliminations and comparison of interpolated data with spatially averaged profiles. The parameters aggregation was done automatically by the new enhanced ODV built-in aggregation module. For Data Aggregation and Data Quality Control the common project methodology has been followed.

EMODnet Chemistry developed products suitable to visualize the time evolution of a selected group of measurements and to calculate spatially distributed data products specifically relevant for MSFD descriptor 5 (eutrophication), 8 (chemical pollution), and 9 (contaminants in seafood). Two type of data products are prepared:

- Interpolated maps of specific parameters in time and depth per sea region;
- Graphical time series of specific parameters at point locations

The interpolated maps have been produced with the Variational Inverse Method (VIM) using the software DIVA (Data-Interpolating Variational Analysis). DIVA is an appropriate numerical implementation of VIM suitable for oceanographic data spatial analysis as it is designed to obtain a gridded field from the availability of non-uniformly distributed observations and it gives major benefits above standard interpolations. Seasonal DIVA products (concentration maps) are available for visualization and downloading at OceanBrowser Viewing Service EMODnet Chemistry - Map Server (http://oceanbrowser.net/emodnet/). An example of a DIVA horizontal map for oxygen parameter in Black Sea is given in Fig. 1.

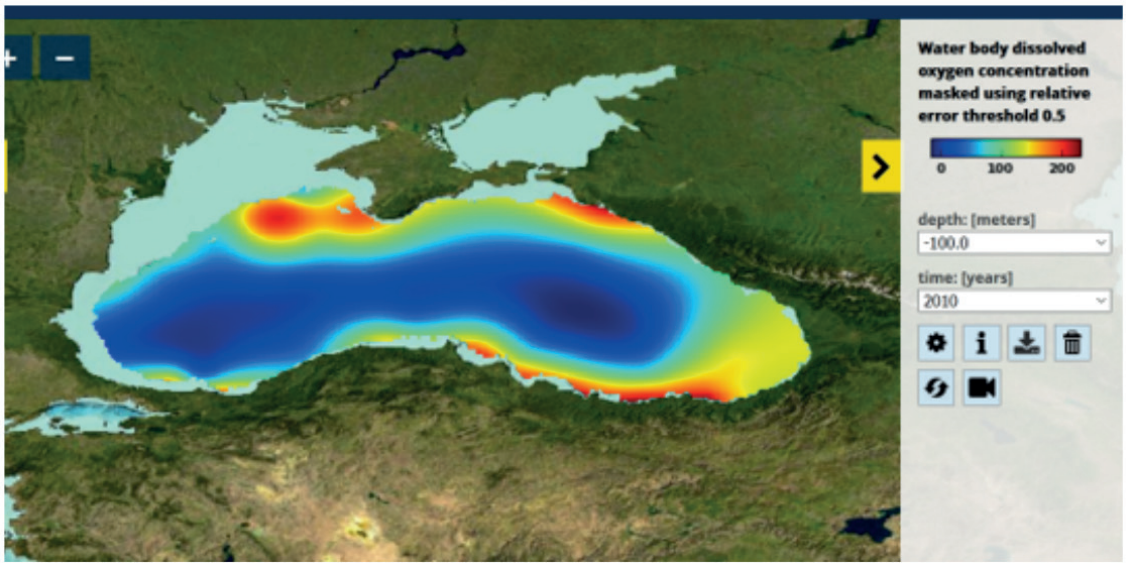


Fig. 1 - DIVA horizontal map for Water body dissolved oxygen concentration in the Black Sea in spring, at 100 m depth, decade 2005-2014, centred on year 2010.

The harmonized and validated data collections were sent to Central Buffer and used for EMODNET Chemistry - Dynamic Plots (Fig. 2)

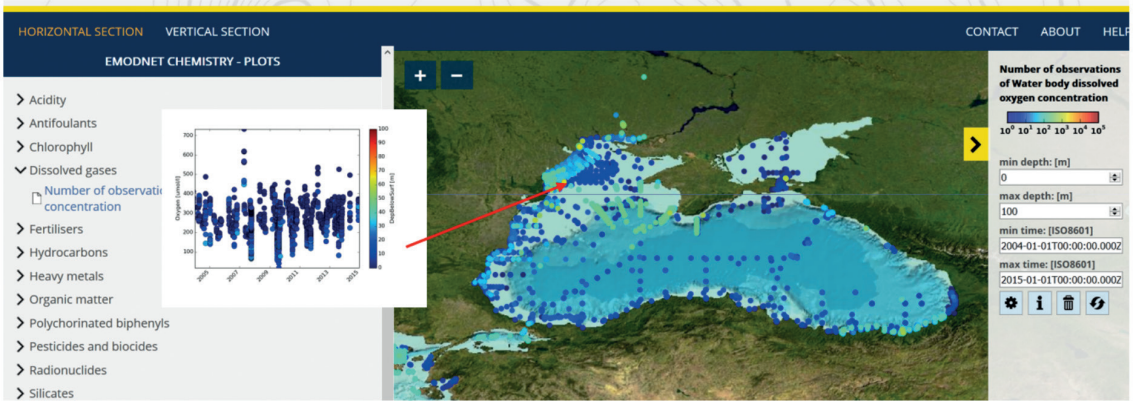


Fig. 2 - The distribution of sampling stations, number of observations and time series for Water body dissolved oxygen concentration in the Black Sea, period: 2004-2015.

Acknowledgements. This work was supported by DG MARE, Call for tenders ASME/EMFF/2016/006 Lot 4 Chemistry. Authors acknowledge the contribution of the EMODnet Chemistry partners involved in the technical infrastructure development, in marine chemical data gathering, harmonization, standardization and quality control, and in data products preparation, as well as all Black Sea data providers

XBT data management and quality control in Japan (III)

Toru Suzuki, Marine Information Research Center (Japan), suzuki@mirc.jha.jp

We reassembled historical expendable bathythermograph (XBT) data in order to improve an ocean subsurface temperature database. Over 4,000 strip charts by XBT chart recorder in the 1980's were collected from Japan Meteorological Agency and Japan Coast Guard, and digitized and stored by image scanner as TIFF format image file (left panel in Fig. 1). The image file was imported by Adobe Illustrator and temperature profile on strip chart was traced and saved as DXF format file. The DXF file was converted as a function of temperature and elaped time from XBT probe launch by affine transformation and depth was calculated by manufacture's fall rate equation. The traced temperature, howerver, was lower than visual reading temperature in existing database of Japan Oceanographic Data Center at deep layer (right panel in Fig. 1). The result of all comparison also indicated the traced temperature was lower than the visual reading temperature at the deeper standard depths (left panel in Fig. 2). We recognized that it is caused by non-linear temperature scale of strip chart. Unfortunately the specification of strip chart was undocumented so that we investigated the temperature scale, i.e., horizontal axis of strip chart by the same way as temperature tracing, and then we determined approximate functions for traced temperature. After correction, more thant 93% of the absolute of difference between traced temperature and visual reading temperature at the standard depths are within 0.1 degree Celsius (right panel in Fig. 2). It is clarify that the traced temperature profile can be replaced with existing database. The depth of traced temperature was finally calculated by new fall rate equation by Hanawa *et al*. (1995) for T-6 and by Kizu *et al*. (2005) for T-5, and those profile have high vertical resolution at 1m intervals as the same as output by recently digital converter. In addition we found the vertically high resolution data by digital converter in 1990's by R/V Ryofu-Maru of Japan Meteorological Agency, so nearly 2,600 profiles were replaced (left panels in Fig. 3) and the number of data remarkable increased in comparison with the existing database (right panels in Fig. 3).

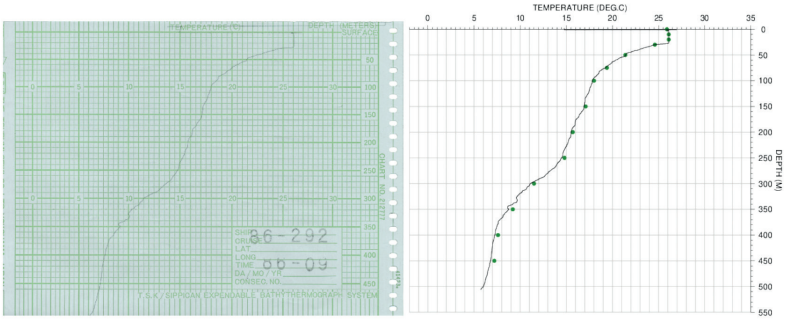


Fig. 1 - Examples of XBT strp chart by R/V Ryofu-Maru of Japan Meteologocal Agency (left panel), and traced temperature profile (black line in right panel) and visual reading temperature at the standard depths (green circle in right panel).

References

HANAWA, K., P.RAUL, RICK BAILY, A.SY, AND M.SZABADOS: *A new depth-time equation for Sippican or TSK T-7, T-6 and T-4 expendable bathythermographs (XBT),* Deep Sea Res., Part I, 42, 1423-1451, doi:10.1016/0967-0637(95)97154-Z, 1995.

KIZU S., H. YORITAKA AND K. HANAWA: *A New Fall Rate Equation for T-5 Expendable Bathythermograph (XBT) by TSK*, J. Oceanogr., 61, No.1, 115-121, 2005.

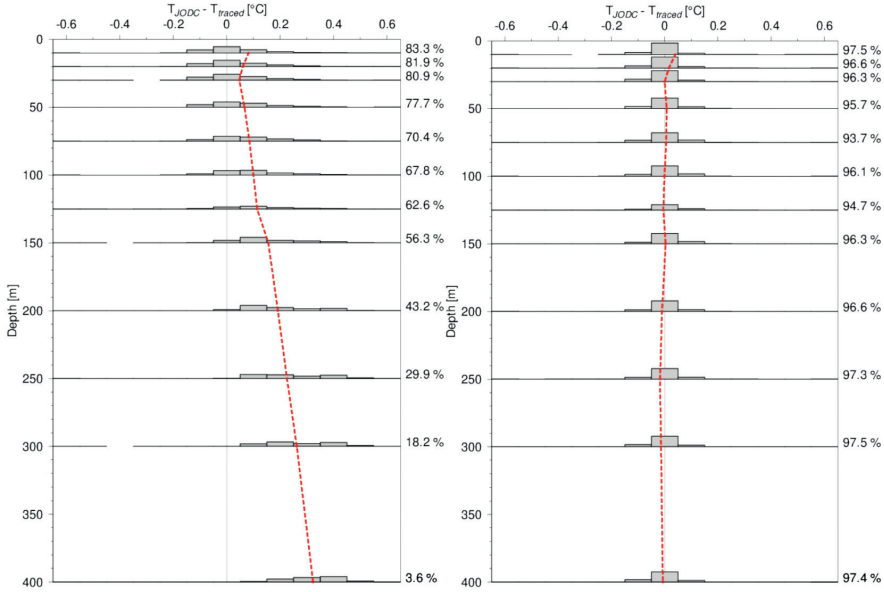


Fig. 2 - Histogram of differences of visual reading temperature and tracing temperature at the standard depths. Left panel show a result by tracing temperature using affine transformation only, and right panel show the after collection by determined approximate functions in this study. Red line shows mean of differences and label at right vertical axis shows a ratio of absolute differences within 0.1 degreee Celcius.

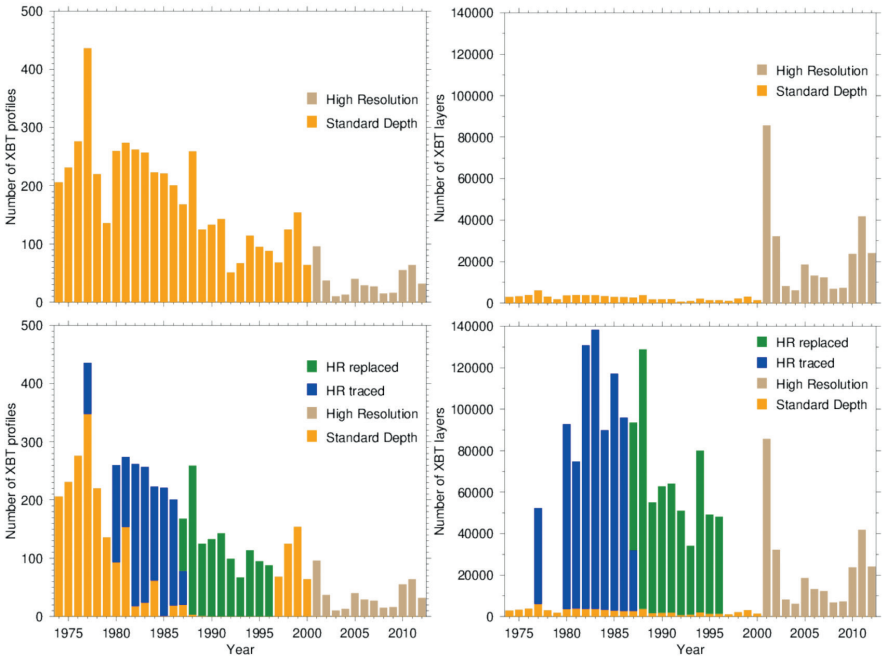


Fig. 3 - Annual changes of the number of XBT profiles (left panels) and XBT layers (right panels) by R/V Ryofu-Maru. Upper panels shows existing database of Japan Oceanographic Data Center and lower panels show the new one in this study. Blue indicates traced data and grean indicates replaced with high resolution data by digital converter which were funded by this study.

Building trust through transparency in the NERC Vocabulary Server (NVS)

Alexandra Kokkinaki, British Oceanographic Data Centre (United Kingdom), alexk@bodc.ac.uk
Gwenaelle Moncoiffe, British Oceanographic Data Centre (United Kingdom), gmon@bodc.ac.uk
Quyen Luong, British Oceanographic Data Centre (United Kingdom), qtl@bodc.ac.uk
Adam Leadbetter, Marine Institute (Ireland), adam.leadbetter@marine.ie
Rob Thomas, Marine Institute (Ireland), rob.thomas@marine.ie
Simon Cox, CSIRO Land and Water (Australia), Simon.Cox@csiro.au

The National Environmental Research Council's (NERC) Vocabulary Server (NVS) has been serving the marine and wider community with controlled vocabularies for over a decade. NVS provides access to standardised lists of terms which are used for data mark-up, facilitating interoperability and discovery in the marine and associated earth science domains. The NVS controlled vocabularies are published as Linked Data on the web using the data model of the Simple Knowledge Organisation System (SKOS). They can also be accessed as web services or through a sparql endpoint. Some of the principles of NVS governance, are the use of deprecation versus deletion as a way to avoid broken links, the access to older versions of a vocabulary on the vocabulary level, the inclusion of external users in content governance discussions and the provision of internal or external concept mappings.

Expanding on the above mentioned principles, work has been undertakento provide greater trust and transparency to NVS users by: adding version control on the concept level, allowing users to access previous versions of a concept; enabling external users to interact with NVS via collaborative platforms with contact details to these platforms and finally provide provenance information related to the creation of mappings.

In this presentation, we will explain how NVS concept version control works and demonstrate how we expanded the current NVS linked data model with World Wide Web Consortium (W3C) compliant ontologies in order to accommodate concept versioning, the provenance of mappings and improve the transparency of the governance model.

SeaDataCloud temperature and salinity data collection for the Black Sea: analysis of data quality problems

Volodymyr Myroshnychenko, Middle East Technical University (Turkey), volodymyr@ims.metu.edu.tr
Reiner Schlitzer, Alfred Wegener Institute (Germany), reiner.schlitzer@awi.de
Michèle Fichaut, Institut Français de Recherche pour l'Exploitation de la Mer (France), michele.fichaut@ifremer.fr
Dick Schaap, Mariene Informatie Service 'MARIS' Bv (The Netherlands), dick@maris.nl

Created for the first time in 2014 within the SeaDataNet II (SDN) EU-project the regional temperature and salinity data collection for Black Sea then underwent 2 updates: one was done in course of the SDN II project while the latest version was released within the SeaDataCloud (SDC) project in 2018. The data collection is managed with ODV software (Schlitzer, R., Ocean Data View, odv.awi.de, 2017) which provides various possibilities for performing quality control (QC), including flagging of wrong and doubtful data, identification of duplicates and data anomalies, etc.

The periodical update intends to supply researches and other end users with the most complete and qualitative data products based on the up to date information retrieved from the expanding SDN infrastructure. Each cycle of the data collection update consists of the following steps:

1. Data harvesting from the central Common Data Index (CDI).
2. File and parameter aggregation into regional collection.
3. Quality check at regional level.
4. Release of final QC-ed collection.
5. Feedback to data providers on found problems and suggested quality flagging.

It is expected that each update cycle will contribute to improvement of the overall quality of data and information in the SDN infrastructure assuming that data originators and distributors will apply the recommended corrections. However the experience earned from the two cycles of collection update suggests that this process is going rather slowly and that some problems continue to persist through update cycles.

Hereby we provide analysis of the problems and factors that affect data quality of the latest SDC temperature and salinity data collection for Black Sea. They can be divided into 2 groups: metadata-related and data-related. The problems from the first group are more serious since they affect whole temperature-salinity profiles and, usually, can't be corrected during QC of the aggregated dataset because require communication and actions from data providers that takes time. As a result, the affected profiles should be eliminated from the data collection.

The major metadata-related problems were identified as follow:

- Duplicates. This is most significant problem in the SDN Black Sea dataset. The duplicates can introduce bias into derived data products, e.g. in climatologies, therefore they should be eliminated from the collection.

- Wrong location (on land).
- Mismatch between CDIs and datasets with respect to parameters, i.e. when CDI record indicates presence of temperature or salinity while the respective dataset (profile) does not contain nor temperature nor salinity. This kind of error can mislead users who are searching data via CDI interface.

The effect of metadata-related problems on the data quality of the SDC temperature and salinity data collection for Black Sea is presented at Fig. 1. More than 10% of profiles are affected, all of them were eliminated from the collection.

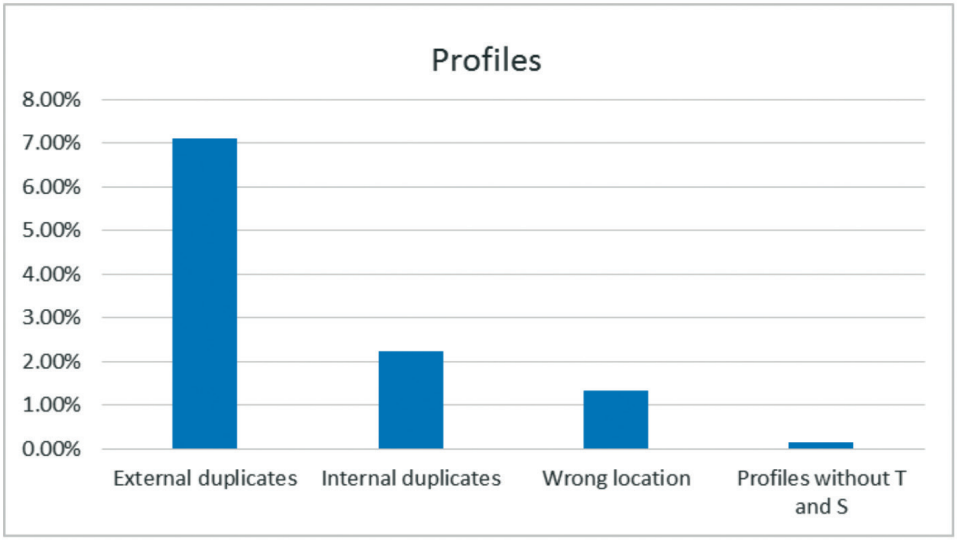


Fig. 1 - Percentage of profiles with metadata-related problems.

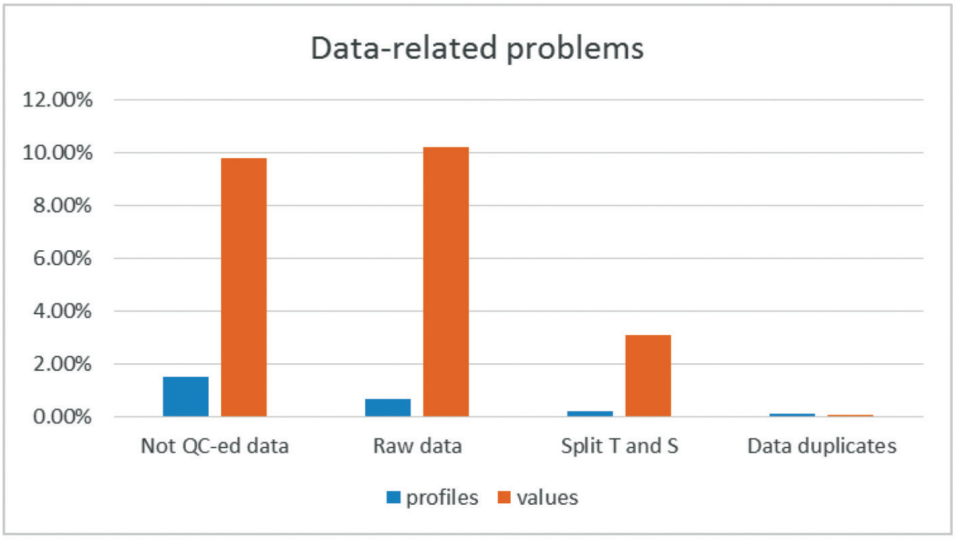


Fig. 2 - Effect of data-related problems.

The data-related problems include:

- Not QC-ed data.
- Raw data. Although the raw data are accepted, they are noisy, the profiles may contain density inversions.
- Profiles with split temperature and salinity, i.e. profiles that consist of pairs of data rows: one row contain temperature, other – salinity. Such pairs should be merged, otherwise it is not possible to calculate derived physical properties.
- Data duplicates within profiles.

In terms of profiles the effect of the data-related problems on data quality in the collection is not large - just ~2.5%, however in terms of data values it is significant - >22% - because the data are coming from the high-resolution CTD profiles (Fig. 2).

The performed analysis allowed to identify most serious problems that affect quality of more than 10% profiles and more than 20% data values in the SDC temperature and salinity data collection for Black Sea. The first problem to be resolved is elimination of duplicates, however it requires close cooperation of involved data providers under coordination of SDN managers and readiness of data providers to withdraw duplicates notwithstanding that it will decrease their scores in SDN. The mandatory QC of data and processing of raw data followed by resubmission of final CTD profiles are the other tasks to be performed with high priority by corresponding data providers. The recommendations on actions to be performed have been elaborated and send to data providers.

Second Black Sea Checkpoint Data Adequacy Report

Vladyslav Lyubartsev, Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy), vladyslav.lyubartsev@cmcc.it

Nadia Pinardi, Centro Euro-Mediterraneo sui Cambiamenti Climatici (Italy), n.pinardi@sincem.unibo.it

Atanas Palazov, Institute of Oceanology Bulgarian Academy of Sciences (Bulgaria), palazov@io-bas.bg

Violeta Slabakova, Institute of Oceanology Bulgarian Academy of Sciences (Bulgaria), v.slabakova@io-bas.bg

Luminita Buga, National Institute for Marine Research and Development "Grigore Antipa" (Romania), lbuga@web.de

Frédérique Blanc, Collecte Localisation Satellites (France), fblanc@cls.fr

Eric Moussat, Institut Français de Recherche pour l'Exploitation de la Mer (France), eric.moussat@ifremer.fr

The aim of the EMODnet Black Sea Checkpoint project is to assess the basin scale monitoring systems on the basis of input data sets for 11 prescribed Challenges, such as: Windfarm siting, Marine Protected Areas, Oil Platform leak, Climate, Coast, Fishery Management, Fishery Impact, Eutrophication, River inputs, Bathymetry and Alien Species. The goal is to assess the "Availability" and "Appropriateness" of the monitoring data sets used to produce the Challenge outputs. The second Black Sea Data Adequacy Report (DAR) concludes the development and implementation of the first assessment of basin monitoring gaps emerging from the generation of 61 Targeted Data Products for 11 Challenges.

The methodology of the DAR follow closely the one developed for the Mediterranean Sea which is based upon ISO and INSPIRE principles and the development of indicators. The indicators are constructed from the Black Sea Checkpoint metadatabase, which contains information on the upstream data used to construct the Challenge products. For each Challenge product, Checkpoint information on "What, Why, Where, When, How" data have been used to develop targeted products is given and statistically analyzed.

The metadatabase contains 503 data set descriptors related to 42 characteristics, i.e. monitoring environmental and human activity information. These descriptors identify potentially usable information for the construction of the Challenge products. Targeted products were constructed from 253 input data sets for the fulfillment of the Challenge products.

The assessment methodology is providing quantitative and qualitative information on "How" the input data sets are made available to Challenges (Availability Indicators) and "What" is the quality of the monitoring data for the Challenge products (Appropriateness Indicators). The assessment methodology has been based on five elements:

- the potential input data sets metadatabase and the availability indicators,
- the Data Product Specification (DPS) and related quality elements,
- the Targeted Data Products (TDP - requested by the call) information and the related quality elements;

- the Ustream Data (UD) used for the products and the related quality elements,
- the calculation of appropriateness indicators from the DPS, UD and TDP quality elements.

Indicator values have been grouped in three color codes (red-inadequate; yellow-partially adequate and green-totally adequate) in order to increase the readability of the results. Results are presented separately for the availability and appropriateness indicators and then they are combined to extract the monitoring gaps. Seventeen monitoring characteristics are found not adequate for the availability indicators. Six are instead found not adequate for appropriateness indicators from the metadatabase analysis. However, it is believed that this evaluation was biased by the fact that the Data Product Specification was not really about what it should have been expected but more what was available. Thus we added the appropriateness scores coming from expert opinion and this raised the inadequate monitoring characteristics to 10 (mostly horizontal and temporal coverage).

In conclusion basin monitoring gaps emerging from this analysis point out to 23 different characteristics that are not monitored adequately in order to construct the 11 Challenge products requested by DG MARE.

In synthesis the Black Sea Checkpoint demonstrated that a quality assessment framework can be defined for the marine environment at basin scales. The framework allows for the first time to assess the monitoring from a customized end-product user point of view. Recommendations for the future development of the service are given in the conclusions.

Acknowledgements. Funding by Service Contract EASME/EMFF/2014/1.3.1.3/LOT4/SI2.709436 Sea-Basin. Checkpoints LOT 4 “Black Sea” - CALL MARE/2014/09 is acknowledged.

Quality control and management of the long-term time series data holdings at NIOZ Royal Netherlands Institute for Sea Research

Taco de Bruin, Royal Netherlands Institute for Sea Research (The Netherlands), Taco.de.Bruin@nioz.nl

Marten Tacoma, Royal Netherlands Institute for Sea Research (The Netherlands), Marten.Tacoma@nioz.nl

Data become more valuable over time. This is especially true for data being used in studies on climate change or the anthropogenic effects on the environment. However, data only become more valuable over time, if the data are managed carefully and professionally. This includes the completeness of the metadata, describing the methods used and circumstances of the measurements; the provenance of the data, describing what has happened with the data since they were collected and the quality control (QC) applied to the data.

NIOZ Royal Netherlands Institute for Sea Research, one of the oldest oceanographic research institutes in the world whose history dates back to 1876, operates all over the world, from the Wadden Sea along the Dutch coast to the remote polar oceans. Its data holdings therefore contain many different data types, ranging from some 6000 CTD-profiles from the deep sea to sediment data from over 3000 annually sampled sites in the Dutch Wadden Sea. Many of these data are part of long-term monitoring series, the crown jewel being a time-series of daily sampled coastal temperatures starting in 1861 (!) and continued to the present day.

A large part of these data are available via the SeaDataNet infrastructure, and thus contribute to several EMODnet projects. Other (biological) data types are accessible through the NIOZ IPT and GBIF.

This presentation provides an overview of the data holdings of NIOZ. Special emphasis is on the QC of extremely long-term time series or monitoring activities, where the duration of these time series may range from several years to over 40 years and, in one case, even to over 150 years!

The integrated information system to support the implementation of the Greek Marine Strategy Framework Directive

Athanasia Iona, Hellenic Centre for Marine Research/Institute of Oceanography/Hellenic National Oceanographic Data Centre (Greece), sissy@hnodc.hcmr.gr

Angelo Lykiardopoulos, Hellenic Centre for Marine Research (Greece), angelo@hcmr.gr

Paraskevi Drakopoulou, Hellenic Centre for Marine Research)/Institute of Oceanography (Greece), vivi@hcmr.gr

Stefanos Kavadas, Hellenic Centre for Marine Research/Institute of Marine Biological Resources & Inland Waters (Greece), stefanos@hcmr.gr

Panagiotis Panagiotidis, Hellenic Centre for Marine Research/Institute of Oceanography (Greece), ppanag@hcmr.gr

An integrated information system was designed to support the input and quality control of the Marine Strategy Framework Directive (MSFD) data, the analysis and indicators estimation and provide network services (WFS, WMS, CSW, etc) to the end users (managers, researchers, technical scientific groups, etc). Data and metadata should be made available through network services supporting the search, display, transformation, and capture of alphanumeric or geospatial data sets. The datasets to be made available within the proposed system should be compatible with the INSPIRE interoperability rules. The architecture of the system is structured on four levels:

1. data ingestion services including the collection and storage of primary data that will be homogenized, normalized and complied according to the INSPIRE Directive. Use will be made of the EMODnet data ingestion services to allow data collectors to upload their data, document them properly and integrate them at the relevant national and European marine data repositories and infrastructures.
2. database management system to support the organization and management of the data collected and processed at the first level,
3. middleware services where the INSPIRE network services are implemented, namely view services, discovery services, transformation services, download services and invoke spatial data services,
4. internet applications to access the network services and the data analysis. Open source technologies are adopted ensuring the scalability of the system without additional cost.

The system will make use of the data services from European infrastructures (such as EMODnet, SeaDataNet, Data Collection Regulation, Data Collection Framework, etc) to integrate additional data sources where it is appropriate and support more efficiently the reporting needs.

Authors Index

G

H

T

U

V

W

Z

Lugliopint - Trieste