



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687289



## Coastal Waters Research Synergy Framework

### Data Management Plan

Document Code: CORESYF-DME-PMD-PMP02-E-R

**Date of delivery:** ABM/UBM-1 (T0+6)  
**Deliverable identifier:** D2.13  
**Version of document:** 1.0 – last updated 23/06/2016]  
**Dissemination level for document:** PU

Table of Signatures			
	Name	Function	Signature
<b>Prepared by</b>	Miguel Terra-Homem		
<b>Reviewed by</b>	Hervé Caumont	WP2 Leader	<i>24/06/2016</i> 
<b>Approved by</b>	Miguel Terra-Homem	Executive Board Chair	
Signatures and approvals appear on original			

<b>Project start date:</b>	01/01/2016
<b>Project duration:</b>	36 months





Revision Records			
Version	Date	Changes	Authors
1.0	23/06/2016	First issue of document	Miguel Terra-Homem Hervé Caumont



## Table of Contents

- 1. Introduction ..... 5
  - 1.1. Purpose and Scope ..... 5
  - 1.2. Document Structure ..... 5
- 2. Data Sets Description..... 6
  - 2.1. SAR\_BATHYMETRY\_DEM ..... 6
  - 2.2. OPT\_BATHYMETRY\_DEM..... 6
  - 2.3. WAT\_BENTHIC\_CLASS ..... 6
  - 2.4. WAT\_QUALITY\_PROP ..... 6
  - 2.5. VESSEL\_DETECTION\_TESTS..... 7
  - 2.6. OIL\_SPILL\_DETECTION\_TESTS..... 7
  - 2.7. WAT\_BOUNDARY\_MAPS ..... 7
  - 2.8. COAST\_ALTIMETRY\_TRACKS..... 8
- 3. Standards and Metadata ..... 8
- 4. Data Sharing..... 9
- 5. Archiving and preservation..... 9
- 6. Reference Documents ..... 10

## List of Tables

No table of figures entries found.

## List of Figures

No table of figures entries found.





Acronyms and Abbreviations	
<b>AIS</b>	Automatic Identification System
<b>API</b>	Application Programming Interface
<b>Co-ReSyF</b>	Coastal Waters Research Synergy Framework
<b>DMP</b>	Data Management Plan
<b>EO</b>	Earth Observation
<b>HR</b>	High Resolution
<b>IOP</b>	Inherent Optical Properties
<b>SAR</b>	Synthetic Aperture Radar
<b>SenSyF</b>	Sentinel Synergy Framework
<b>SLA</b>	Sea Level Anomaly
<b>SSH</b>	Sea Surface Height
<b>SST</b>	Sea Surface Temperature
<b>SWH</b>	Significant Wave Height



## 1. Introduction

The Co-ReSyF project will deploy a dedicated data access and processing infrastructure and user portal, with automated tools, methods and standards to support research applications using Earth Observation (EO) data for monitoring of Coastal Waters, leveraging system components deployed as part of the SenSyF project ([www.sensyf.eu](http://www.sensyf.eu)). The main objective is to facilitate the access to Earth Observation data and processing tools for the Coastal research community, aiming at the provision of new Coastal Water services based on EO data.

Through Co-ReSyF's collaborative front end, even unexperienced researchers in EO will be able to upload their applications to the Cloud Platform, in order to compose and configure processing chains, for easy deployment and exploitation on a cloud computing infrastructure. Users will be able to accelerate their development of high-performing applications taking full advantage of the scalability of resources available from Terradue Cloud Platform's Application integration service. The system's facilities and tools, optimized for distributed processing, include EO data access catalogues, discovery and retrieval tools, as well as a number of pre-processing tools and toolboxes for manipulating EO data. Advanced users will also be able to go further and take full control of the processing chains and algorithms by having access to dedicated cloud back-end services, and to further optimize their applications for fast deployment addressing big data access and processing.

The Co-ReSyF capabilities will be supported and initially demonstrated by a series of early adopters who will develop new research applications for the coastal domain, guide the definition of requirements and serve as system beta testers. Following this, a competitive call will be issued within the project to further demonstrate and promote the usage of the Co-ReSyF release. These pioneering researchers will be given access not only to the Cloud Platform itself, but also to extensive training material on the system and on Coastal Waters research themes, as well as to the project's events, including the Summer School and Final Workshop.

### 1.1. Purpose and Scope

The purpose of the Data Management Plan (DMP) is to provide an analysis of the main elements of the data management policy that will be used by the Co-ReSyF project with regard to all the datasets that will be generated by the project. The DMP is not a fixed document, but evolves during the lifespan of the project.

### 1.2. Document Structure

The structure of the document is as follows:

- Chapter 2 : Description of the datasets to be produced
- Chapter 3: Description of the standards and metadata used
- Chapter 4: Description of the method for sharing the data
- Chapter 5: Solutions for the archiving and preservation



## 2. Data Sets Description

### 2.1. SAR\_BATHYMETRY\_DEM

Digital Elevation model of the sea bed surface derived from a collection of SAR images of the area of interest. The derived bathymetry data will be applicable to the region of the sea bed from the coastline ranging from depths lower than 200m and greater than 10m.

The data may allow monitoring the coastal bathymetry evolution from multiple images (at multiple times). Also in rapidly morphologically changing conditions, such as during coastal storms or tsunamis, if the SAR-image conditions are valid, then one can obtain pre and post-disaster bathymetries, which are extremely useful for disaster/risk management and coastal management in general. The bathymetry evolution would also be of extreme importance for data assimilation studies with the numerical morphodynamical or storm-surge models. Other uses of this product are: coastal engineering studies, coastal morphological evolution studies, coastal wave and current numerical modelling, etc...

### 2.2. OPT\_BATHYMETRY\_DEM

Digital Elevation model of the sea bed surface derived from a collection of HR optical images of the area of interest. The derived bathymetry data will be applicable to the region of the sea bed from the coastline ranging from depths from 0m to 10m (shallow waters).

This data will be complementary to the SAR\_BATHYMETRY\_DEM, covering the shallow waters region that cannot be covered by the SAR technique. The usage of the data will be the same as for the SAR\_BATHYMETRY\_DEM (refer to Section 2.1).

### 2.3. WAT\_BENTHIC\_CLASS

Data containing the classification of the sea floor by its class of sea bottom albedo type. The derived data will be applicable to the region of the sea bed from the coastline ranging from depths from 0m to 10m (shallow waters).

For shallow ocean waters, knowledge of the optical bottom albedo is necessary to model the underwater and above-water light field, to enhance underwater object detection or imaging, and to correct for bottom effects in the optical remote sensing of water depth or inherent optical properties (IOP's). Measurements of the albedo can also help one identify the bottom-sediment composition, determine the distribution of benthic algal or coral communities, and detect objects embedded in the sea floor.

### 2.4. WAT\_QUALITY\_PROP

Data containing the water quality properties (Chlorophyll-a concentration, particulates backscattering coefficient, and absorption of Coloured dissolved organic materials), for a region



of interest. The derived data will be applicable to the region of the sea bed from the coastline ranging from depths from 0m to 10m (shallow waters).

Knowledge of the water quality properties can be used for environmental analysis of the quality of the coastal waters and their evolution with the growing of the coastal city areas. Knowledge of the Chlorophyll-a concentration (biomass) is one of the most useful measurements in limnology and oceanography. The biomass can be used for studying phytoplankton community structure, the size frequency distribution of the algal cells and seasonal shifts within the plankton community. Phytoplankton abundance is related to natural cycles in nutrient availability and to the input of phosphate and nitrate. Excess phosphate and nitrate can come from groundwater or water treatment plants and sewer overflow (nitrate and phosphate are not removed in most sewage treatment plants). Excess nutrients can cause blooms of phytoplankton, which can contribute to bottom water anoxia under stratified conditions.

## 2.5. VESSEL\_DETECTION\_TESTS

Data containing the position of the detected vessels for a time interval and region of interest, and the real position of vessels from historical AIS databases for the same time interval and region of interest.

The data may be used for researchers to analyse the performance of their detection algorithms with respect to the number of false alarms and true positives for the ships detection. Ship detection plays an important role in monitoring illegal ships activities and controlling the country's borders. It can also be used for statistics for marine traffic in order to identify the areas with intensive ship routes.

## 2.6. OIL\_SPILL\_DETECTION\_TESTS

Data containing the position of the detected oil spills for a time interval and region of interest, and the real position of the oil spills from historical databases for the same time interval and region of interest.

The data may be used for researchers to analyse the performance of their detection algorithms with respect to the number of false alarms and true positives for the oil spills detection. Oil spill detection is used to monitor the illegal dumping of oil from vessels and maintaining the environmental integrity of the coastline. The identification of oil spills may also play an important role in case of environmental disasters in order to assist in the clean-up procedures of the environmental agencies.

## 2.7. WAT\_BOUNDARY\_MAPS

The data is composed of two ocean boundary map for the period and region of interest in question. One map is delineating the boundaries between different pixels exhibiting different seasonalities, where prime zones for water mass mixing can be found. And another map is providing a randomised SST/Chl boundary output, to determine whether the patterns evident in the true boundary data, are indeed patterns worth basing scientific research on.



Water mass convergence zones are critically important to maintaining our ocean's fisheries and ecological systems, frequently representing areas where colder, nutrient rich water mixes with warmer waters. This interaction fuels increased plant growth at the base of the ocean food chain, feeding the system that produces the overwhelming majority of the fish we eat. As ocean surface phytoplankton grows, sickens, and dies, their photosynthetic activity fluctuates accordingly. Satellite-derived measurements of chlorophyll activity provide a measurable estimate of this photosynthetic activity. Furthermore, with near-daily repeat times over fixed points, some optical datasets currently extend over a continuous 15 years.

## 2.8. COAST\_ALTIMETRY\_TRACKS

Data containing the SSH, SLA, SWH and Wind Speed data, and its respective geographical coordinates and time stamp, derived from the ALES retracker algorithm for the region of interest and the selected time period with a sampling frequency of 20Hz. In addition an extra dataset with the range and the applied corrections in order to derive the main parameters will also be part of the data.

Sea level rise is an important indicator of climate change and one of its greatest impacts on society. Due to sea level rise many regions of the world's coasts will be at much increased risk of flooding during the course of the 21st century and hundreds of millions of people currently living just above may have to be relocated, and coastal infrastructure to be moved or rebuilt. The rise rate shows huge regional variations so it is essential for coastal planners to have regional observations of sea level rise rate, which in combination with regional models will lead to regional projections. Tide gauges measure sea level variation but are affected by local vertical land movement, for instance due to subsidence. Altimetry, and coastal altimetry in particular, provide complementary measurements that are not affected by vertical land variation. In essence the coastal planners need the integration of both types of measurement.

## 3. Standards and Metadata

The Co-ReSyF catalogue uses the OGC® OpenSearch Geo and Time extensions standard to expose data + metadata. The baseline of the standard is the OpenSearch v1.1 specification (A9.com, Inc, n.d.). The Geo and Time extensions provide with the main queryables to manage geospatial data products, and the specification is standardized at the Open Geospatial Consortium (OGC, n.d.).

According to the selected standard baseline and extensions, the data catalogue queryable elements are as follows:

- count={count?}
- startPage={startPage?}
- startIndex={startIndex?}
- sort={sru:sortKeys?}
- q={searchTerms?}



- start={time:start?}
- stop={time:end?}
- timerelement={time:relation?}
- name={geo:name?}
- uid={geo:uid?}
- bbox={geo:box?}
- geometry={geo:geometry?}
- submitted={dct:dateSubmitted?}

## 4. Data Sharing

The Cloud Platform will be the privileged storage and cataloguing resource for managing the information layers produced by the project and described above. The Data Agency hosted by Terradue is the dedicated Cloud Platform service in charge of this type of operations. The catalogue entries of the Data Agency are open and web accessible.

The data products referenced by a catalogue entry might require user authentication to allow download from a storage repository, depending on the policies (e.g. embargo period before full public release) applied by the producer organization. User registration procedures shall be described on the Co-ReSyF portal and be simple to follow by registrants (based on a single sign-on principle).

To query the Data Agency catalogue from within application scripts, users can take advantage of the opensearch-client tool (Terradue, n.d.). Here are the available keywords to receive direct results:

- wkt -> retrieve the product geometry in Well Known Text format
- enclosure -> the URL to download the product
- startdate -> Product Start Datetime
- enddate -> Product End Datetime
- identifier -> Product ID

## 5. Archiving and preservation

While Terradue Cloud Platform provides the Data management and operations backbone from where data products can be directly accessed, the Cloud Platform supports distributed computing protocols allowing managing different storage locations for a same dataset. Long term data preservation of the produced datasets is foreseen to follow two main tracks:



1. Management of product copies onto the European Science Cloud, with as of today, physical resources being provided by the organization EGI (European Grid Infrastructure).
2. Management of products copies onto partners own storages offering an access point to the Platform to perform the data staging operations.

Option 1 is the default option for long-term preservation of the product copies and will be assured as long as the Co-ReSyF platform is operational with no additional cost to the partners for storing the data.

Option 2 is an alternative option in case the generator of the dataset prefers to keep control of the data. This option is not foreseen to be used for the datasets described above and in case it is used it is the responsibility of the owner of the repository to arrange for financial support for ensuring the preservation of the data.

## 6. Reference Documents

- Co-ReSyF. (2016). *GRANT AGREEMENT-687289*. European Commission, Research Executive Agency.
- Terradue (n.d.). opensearch-client. Terradue Developer Cloud Sandbox. Retrieved June 22<sup>nd</sup>, 2016, from [http://docs.terradue.com/developer-sandbox/reference/man/bash\\_commands\\_functions/catalogue/opensearch-client.html?highlight=opensearch](http://docs.terradue.com/developer-sandbox/reference/man/bash_commands_functions/catalogue/opensearch-client.html?highlight=opensearch)
- A9.com, Inc (n.d.). OpenSearch Specifications 1.1. OpenSearch.org. Retrieved June 22<sup>nd</sup>, 2016, from <http://www.opensearch.org/Specifications/OpenSearch>
- OGC (n.d.). OpenSearch Geo and Time Extensions. Opengeospatial.org. Retrieved June 22<sup>nd</sup>, 2016, from <http://www.opengeospatial.org/standards/opensearchgeo>



END OF DOCUMENT

