



DATA MANAGEMENT PLAN



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 776348.

Project no. 776348
Project acronym: CoastObs
Project title: Commercial service platform for user-relevant coastal water monitoring services based on Earth Observation
Instrument: H2020-EO-2017
Start date of project: 01.11.2017
Duration: 36 months
Deliverable title: D2.3: Data Management Plan
Due date of deliverable: Month 6
Organisation name of lead contractor for this deliverable: Water Insight (1)

Author list:

Name	Organisation
Kathrin Poser	Water Insight

Dissemination level		
PU	Public	
CO	Confidential, restricted under conditions set out in Model Grant Agreement	x
CI	Classified, information as referred to in Commission Decision 2001/844/EC	

History			
Version	Date	Reason	Revised by
0.1	06/04/2018	Outline	Kathrin Poser, WI
0.2	24/04/2018	First full draft	Kathrin Poser, WI
1.0	30/04/2018	Contributions and review by partners	Jesus Torres, UVIGO Laurent Barillé, UN Evangelos Spyarakos, USTIR
1.1	19/11/2019	BL logo replacement	Marnix Laanen, WI

CoastObs Project

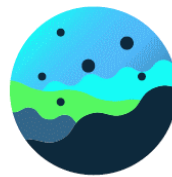
CoastObs is an EU H2020 funded project that aims at using satellite remote sensing to monitor coastal water environments and to develop a user-relevant platform that can offer validated products to users including monitoring of seagrass and macroalgae, phytoplankton size classes, primary production, and harmful algae as well as higher level products such as indicators and integration with predictive models.



phytoplankton



seagrass



harmful algal blooms



primary production

To fulfil this mission, we are in dialogue with users from various sectors including dredging companies, aquaculture businesses, national monitoring institutes, among others, in order to create tailored products at highly reduced costs per user that stick to their requirements.

With the synergistic use of Sentinel-3 and Sentinel-2, CoastObs aims at contributing to the sustainability of the Copernicus program and assisting in implementing and further fine-tuning of European Water Quality related directive.

Partnership



Water Insight BV. (WI)



UNIVERSITY OF
STIRLING

The University of Stirling (USTIR)



Consiglio Nazionale
delle Ricerche

Consiglio Nazionale Delle Ricerche (CNR)



UNIVERSITÉ DE NANTES

Universite de Nantes (UN)



UNIVERSITY
OF APPLIED SCIENCES

HZ University of Applied Sciences (HZ)



UNIVERSIDADE
DE VIGO

Universidad de Vigo (UVIGO)



BIO-LITTORAL
Expertises et Conseils en Environnements Marin et Estuarien

Bio-Littoral (BL)



GEONARDO
STATE-OF-THE-ART AND BEYOND

Geonardo Environmental Technologies Ltd. (GEO)

TABLE OF CONTENTS

Summary	8
1 Data summary	9
1.1 Purpose and utility of the data.....	9
1.2 Types of data	10
1.3 FAIR data	11
2 Earth observation data.....	13
2.1 Data origin	13
2.2 Data ownership, sharing and access	14
2.3 Data formats and metadata	14
3 In situ data.....	16
3.1 Data origin	16
3.2 Data ownership, sharing and access	17
3.3 Data formats and metadata	18
4 Data collection overview.....	19
5 Other data	22
5.1 Data collected from external sources	22
5.2 Personal data.....	22
5.3 Research outputs.....	22

ABBREVIATIONS

List of abbreviations	
Abbreviation	Explanation
CDOM	Coloured dissolved organic matter
Chl-a	Chlorophyll-a
DMP	Data management plan
EO	Earth observation
FAIR	Findable, accessible, interoperable and re-usable
FTP	File Transfer Protocol
GEO	Group on Earth Observation
GEOSS	Global Earth Observation System of Systems (GEOSS)
GHR SST	Group for High Resolution Sea Surface Temperature
HAB	Harmful algae bloom
INSPIRE	Infrastructure for Spatial Information in the European Community
Kd	Vertical diffuse attenuation coefficient
L8	Landsat 8
LIMNADES	Lake Bio-optical Measurements and Matchup Data for Remote Sensing
MB	Management Board
MODIS	Moderate Resolution Imaging Spectrometer
PC	Phycocyanin
PP	Primary production
PSC	Phytoplankton size classes
S2	Sentinel 2
S3	Sentinel 3
SD	Secchi Depth

SST	Sea Surface Temperature
T	Temperature
TSM	Total suspended matter
TUR	Turbidity
VHR	Very high resolution
WFD	Water Framework Directive
WQ	Water quality

Summary

Objective:

The objective of this data management plan (DMP) is to detail the plan for managing the data generated and collected within the CoastObs project. The DMP describes the data management life cycle for all datasets collected, processed and/or generated by the project. It covers

- what data will be collected, processed or generated
- how the data will be handled during and after the project
- who is considered as an owner of a data set and who it is shared with
- how the sharing of the data within and outside the project is organised
- what formats, meta data and standards the data will adhere to
- how data will be curated and preserved

Rationale:

Sound data management is pivotal for fully realising the benefits of CoastObs as well as for allowing the wider research and user communities to assess and further the achievements of the project. Within the project, well curated data will stimulate and ensure smooth collaboration between the project partners. In interaction with the project user group, sound data management will allow the users to easily evaluation and put to use the data received from the project. For dissemination and exploitation, open access to data generated in the project will help to underpin the credibility and stimulate the uptake of the CoastObs results.

Scope:

The data management plan is expected to evolve during the project. This deliverable presents the status and planning at month 6 of the three year project. Updates to the provisions in this document will be recorded in the periodic reports.

1 Data summary

1.1 Purpose and utility of the data

The quality and ecological status of European and global coastal waters is an area of real concern because of the multiple and conflicting pressures from anthropogenic perturbation and environmental change. The multiple and often conflicting uses of the coastal zone including aquaculture, fisheries, wind, tidal energy, cooling waters, shipping and tourism affect sustainably management of these environments. European directives such as the Marine Strategy Framework Directive (MSFD), the Water Framework Directive (WFD), Bathing Water Directive (BWD), Natura 2000, the Maritime Spatial Planning Directive (MSPD) and Integrated Coastal Management (ICM) are put into place to ensure sustainable management of the marine and coastal resources. These directives ensure that the collective pressure of all activities is kept within levels compatible with the achievement of good environmental status and that the capacity of marine ecosystems to respond to human-induced changes is not compromised, while enabling the sustainable use of marine goods and services by present and future generations. Traditional monitoring, based on in situ sampling, is widely recognised as being inadequate in characterising the ecological status of the these highly dynamic coastal environments in both space and time resulting in significant uncertainties. To overcome these substantial shortcomings, satellite remote sensing can provide synoptic coverage and quantify the spatial and temporal variability in the physical and biogeochemical properties of these vulnerable environments.

The purpose of the data collected and generated in CoastObs is to demonstrate the added value of Earth observation (EO) data for water quality monitoring. In four case study areas, we use standard and novel algorithms to derive relevant water quality parameters from satellite images and convert these to information products that match the users' information needs. Different kinds of users are collaborating with the CoastObs consortium on the demonstration of the utility of the proposed products. The different user groups have different aims and motivations for their interest in the CoastObs products, including, but not limited to:

- System understanding: Many users including public authorities, aquaculture operators or researchers are very interested in a more thorough understanding of the spatio-temporal dynamics and long-term trends of the water bodies they are working with.
- Monitoring: Water management authorities have an obligation to monitor the water quality of the water bodies for which they are responsible. In highly dynamic coastal areas, traditional sampling and laboratory methods are often increasingly recognized as insufficient to adequately capture the ecological status, therefore water managers are looking for new and efficient methods for their monitoring.

- Assess impact of measures: Water management authorities that are responsible for implementing measures to improve or ensure water quality are looking for information to help with evaluating how successful these measures are. Earth observation data allows to go back in time to assess the magnitude and spatio-temporal patterns of ecosystem changes.
- Assess impact of operations on ecosystem / other users: Operators in coastal engineering (such as dredging, harbours) are responsible for monitoring the impact of their operations on the environment to ensure that their impact stays within an acceptable limit. Earth observation can help to understand the scale and intensity of these impacts. This information is of interest to several parties: the engineering companies themselves, the regulators who are overseeing the operations and possible other users of the coastal zones that are feel the impact (for example, aquaculture operators).
- Spatial suitability: Based on long-term records of spatial distribution of certain WQ parameters (such as the MERIS sensor) areas can be identified that are particularly suitable for a certain type of use (e.g. mussel aquaculture) as basis of a management tool to optimize mussel production.
- Near-real time warning: While many of applications of CoastObs services are not time critical, there are also some users that have an interest in near-real time services to warn them immediately if unfavourable water quality conditions are observed, for example aquaculture producers are very much interested in the occurrence of harmful algae blooms that can threaten their production or dredging companies may want to know when the disturbance caused by their operations exceed permitted limits.

The overall aim of the CoastObs products can be summarised as providing users from different user communities with fit-for-purpose data that help them to gain a deeper insight into the spatio-temporal dynamics of water quality in coastal areas to help them make better informed decisions.

1.2 Types of data

The two main types of data collected and produced within CoastObs are Earth observation (EO) and in situ data. Water quality information derived from EO data is the main product of CoastObs. The consortium acquires raw satellite data, mainly via the Copernicus Sentinel Data Hub, and processes them into water quality data layers using dedicated algorithms. These data layers are either directly delivered to the users or further processed into higher-level products. Higher level products can take many different forms: they can also be spatial data products, derived maps, time series, statistics, warning messages, or reports.

In situ data are collected by the consortium as reference and validation data in a number of field campaigns in the different study areas. The data collection includes measurements in the field and the laboratory of the target WQ parameters, as well as in situ spectral measurements that allow linking the measured WQ parameters to spectral properties that are also observed by the satellites. These data are used in the development and validation of the CoastObs products.

Details about the data including origin, handling within the project, ownership and access, formats, required metadata, findability, accessibility and security will be discussed in chapters 2 (EO data) and 3 (in situ data).

1.3 FAIR data

In accordance with the H2020 requirements, all CoastObs research data should be 'FAIR', that is findable, accessible, interoperable and re-usable. In the following, a short overview is given of how CoastObs envisions to reach this goal, more detail can be found in chapters 2 (EO data) and 3 (in situ data) for the respective data types. It should be noted that not all details (e.g. how to assign unique identifiers to data) have been worked out at this point, but that data management is considered as an ongoing and evolving activity during the project.

To make CoastObs data findable, accessible and re-usable, our data, or at least the metadata, will be submitted to public repositories. Specific repositories will be discussed in chapters 2 and 3, but as a fallback the general-purpose ZENODO repository (www.zenodo.org) is considered. ZENODO, hosted by CERN, stores and hosts all kinds of research data and makes it findable and accessible via OpenAIRE (www.openaire.eu). All types of data (EO, in situ) will be submitted to the repositories, but only those datasets that can potentially be re-used by new projects or others interested parties. Raw data and intermediate products are not considered to be submitted, only final products, and in some cases examples or compilations (e.g. an average annual Chl-a map instead of daily processed EO imagery).

To increase the findability of specific data to certain communities, where suitable, permalinks from ZENODO will also be provided to other databases or the data will be submitted to other database. These options are discussed per data type in the following sections.

Interoperability is supported by using standard meta data and file formats. The file formats are discussed per data type in the following sections.

CoastObs aims at making all potentially re-usable data available via public repositories at the latest by the end of the project. If partners are at the time in the process of generating a publication and therefore wish data to remain confidential until the publication is accepted/published, this can be decided by the management board. In this case, the lead partners is responsible to ensure submission of the data at the time of publication. At the end

of each reporting period, the management board will review the data generated in that period, ensure that data ownership and access is agreed on and then decide on the timing and procedure for submitting the data.

2 Earth observation data

In CoastObs we retrieve raw satellite data from several data providers and process these to water quality parameters. These processed images of parameters such as chlorophyll-a, phytoplankton size classes or the abundance of sea grass are spatial data products (raster data sets), the first type of product delivered to users by CoastObs. The second data type are higher-level products, which are produced from the individual water quality data layers by aggregation, point extraction, classification, combination or integration into predictive models. These higher-level products come in different formats, not all of which are defined at this time as they very much depend on the users' wishes and requirements. Table 1 gives an overview of how the different levels of satellite data will be managed in CoastObs.

Table 1 – Earth observation data management overview

	Source	Owner	Storage	Internal sharing	External sharing
Raw data	EO data providers (mainly Copernicus, NASA)	EO data providers	During product development phase : WI or partner working on product development, operational phase : WI	Only when necessary	No
Processed images	Own processing	Shared ownership between partners who developed, processed and validated the product, details decided by MB		Yes, via access to data cube software	Yes, selected case studies via data portal
Higher level products	Own processing	Shared ownership between partners who developed, processed and validated the product, details decided by MB		Yes, how depends on type of product	Yes, how depends on type of product

2.1 Data origin

EO data is all the data that originates from satellite imagery of various providers. The most important provider of raw satellite imagery in CoastObs is the Copernicus programme via the Sentinel Data Hub (from the European Space Agency for Sentinel-2, European Organisation for

the Exploitation of Meteorological Satellites for Sentinel-3). Other satellite imagery will be obtained via the Copernicus contributing missions Data Ware House from other (commercial) providers and the from the National Aeronautics and Space Agency (Landsat-8 and sea surface temperature data).

2.2 Data ownership, sharing and access

The raw satellite data will be retrieved from the respective data provider either by WI as the operational service provider, or during the product development phase, by the partners working on the respective products. The data owner is the respective data provider. Data are shared within the consortium only where necessary, in principle each partner can download them individually. Raw satellite data are not shared outside the consortium (for the publicly available data, it is considered as unnecessary, whereas data from commercial providers may come with restrictions for distribution of the raw data).

All processed and higher-level EO-based data generated in CoastObs are considered as CoastObs products. Partners that have contributed to each product (by method development, processing or validation) will be considered as shared owners of the products. A management board meeting at the end of each reporting period will be used to agree on the owners of each of the products developed in this period. All partners have access to all EO products, this access will be realised based on the technical solution for data cube for the processed images. For the higher level products, it depends on the type and size of data. They will be shared either via the project Dropbox (e.g. for time series or reports) or via FTP, the data portal or also via the data cube (for spatial data products).

The EO products will be made available to users via the agreed channels (FTP, portal/map service/data service), details can be found in the user requirements document (D2.1).

Selected EO data from the case studies („success stories”) will be submitted to ZENODO-OpenAIRE. Permalinks will be provided to the GEOSS portal (www.earthobservations.org/geoss.php) of the Group on Earth Observation (GEO). The consortium will also explore the possibility to make (meta)data available via the European Marine Observation and Data Network (EMODnet).

2.3 Data formats and metadata

The EO data products in CoastObs are in raster format. The standard output formats that will be produced are:

- GeoTiff or NetCDF (data file)

For higher-level products the format depends on the type of product, the following formats will be considered:

- GeoTiff or NetCDF (data file) for spatial raster data
- .shp (data file) for vector data
- .png, .jpg or .pdf (figure) for maps and plots
- .mpg for animations
- .csv or .xlsx for point and time series data

All data that is made available by CoastObs will be stored including a basic set of the meta data, with additional meta data depending on the user requirements. The minimum required metadata are:

- Parameter
- Unit
- Date, time (including time zone)
- Location information, spatial reference system
- Lineage (how was the data produced)
- Ownership, contact person or email

EO data that will be made available outside the consortium will be accompanied by meta data according to the INSPIRE guidelines and/or ISO 19115:2003- Geographic information -Metadata standard.

3 In situ data

In situ data are all the data collected in the field. Within CoastObs a number of field campaigns will be performed by the partners to collect reference data for the development and validation of the EO products. In this campaigns the partners will be collecting data on the target water quality parameters of CoastOb, as well as in situ spectral data to allow linking the WQ parameters to the spectral signal that can also be observed by the satellites. In addition, ancillary data are collected that can help in explaining observed properties (such as wave and weather data, photographs). Similar in situ data are sometimes also provided by the users or collected from external data sources.

While the primary goal of in situ data collection as stated in the project proposal is the development and validation of the products within the project, it is now also being discussed to explore the development of an integrated service that provides the user with both EO data and matching in situ validation data to ensure continuous validation and quality control.

Table 2 gives an overview of how the different in situ datasets be managed in CoastObs.

Table 2 – In situ data management overview

	Source	Owner	Storage	Internal sharing	External sharing
CoastObs in situ data	Own collection	Partners involved in collecting and processing the samples / data	Dropbox or FTP, for spectral measurements : WISPweb	Yes	Yes, via ZENODO or dedicated repository
User in situ data / data from external sources	CoastObs users / various sources (national monitoring networks, other researchers)	Original data owner		If permitted by data owner	No

3.1 Data origin

In situ data is collected and processed by the partners for development and validation. Data are collected in the field, i.e. either from a boat or from the shore (e.g. sea grass in intertidal areas). Data collection includes optical inspection, using in situ probes or spectrometers, or collecting water samples, which are subsequently processed in a lab.

Data provided by users or external data sources will be collected from the data owners for the use within the CoastObs project. Usage rights and obligations will be discussed with each data provider.

3.2 Data ownership, sharing and access

All in situ data generated in CoastObs are considered as CoastObs data. Partners that have contributed to each data set (by collecting or processing field samples) will be considered as shared owners of the data. All in situ data generated in the project will be shared within the consortium, and will be made publicly available.

The WISP data will be stored in the WISPweb database at WI servers. CoastObs data will be included in the 'CoastObs group', which means that all partners can access all WISP-3 data collected within CoastObs. Other in situ data will be shared as comma separated text files via Dropbox or FTP within the consortium.

To make the in situ data publicly findable and accessible, in situ data will be submitted to ZENODO-OpenAIRE and LIMNADES, or one of these two, creating links on the other. LIMNADES is a centralised global database of ground bio-optical measurements from lakes developed through voluntary cooperation across the international scientific community. LIMNADES provides a repository for (1) inherent and apparent optical property data and associated water constituent measurements; and (2) in situ water constituent measurements for satellite validation. LIMNADES was developed by and managed by the UK-Globolakes project and is curated by CoastObs partners USTIR. Although LIMNADES started as a database on lakes, including data from coastal regions is seen favourably by USTIR.

The timing of data publication will be discussed in the management board meeting at the end of each reporting period to allow partners time for producing publications before giving the public access to the data where this is necessary.

3.3 Data formats and metadata

In situ data will be shared within the project as comma-separated text files (.csv). An example of such a file is given in Figure 1

```
CoastObs field campaign region XXX
Month year, partner Y, if applicable user Z
contact person: Name, contactperson@partnerY.eu
Notes: also spectral data and photos collected

Date, time UTC, lat, lon, station, parameter, unit, value, method, notes
Yyyymmdd, hh:mm, deg.mmmmm, deg.mmmmm, Station 1, Chl-a, mg/m3-1, X.XX, HPLC, cloudy
yyymmdd, hh:mm, deg.mmmmm, deg.mmmmm, Station 1, TSM, g/m3-1, X.XX, filterpad, cloudy
yyymmdd, hh:mm, deg.mmmmm, deg.mmmmm, Station 2, Chl-a, mg/m3-1, X.XX, HPLC, S2 matchup
yyymmdd, hh:mm, deg.mmmmm, deg.mmmmm, Station 2, TSM, g/m3-1, X.XX, filterpad, S2 matchup
```

Figure 1 - Example .csv file for in situ data sharing

All data that is made available by CoastObs will be stored including a basic set of the meta data, with additional meta data depending on the user requirements. The minimum required metadata are:

- Parameter
- Unit
- Method (instrument / procedure used, number of repeat measurements, ...)
- Date, time (including time zone)
- Location information, preferably as latitude, longitude
- Ownership, contact person or email

4 Data collection overview

Table 3 gives an overview of all data that is expected to be collected and produced by the consortium for the case study sites. The overview is at this moment still tentative as it reflects the user requirements, which in some cases need some feasibility check and more discussion with the respective users.

Table 3 – Overview of data collected/produced in CoastObs

Area	EO standard and innovative products (from satellite)	Higher level products	In situ data campaigns	User-provided data	Other data
Wadden Sea and Eastern Scheldt	Chl-a (S2, S3), TSM, TUR, Kd (S2, L8, S3), CDOM (S3), SST (L8, GHRSSST), PSC (S3), HAB (S3)	Spatial suitability maps, integration with predictive modelling of mussel growth, coastal erosion and accretion, source identification, time series, statistics, trends	Chl-a, TUR, T, spectral data, PSC	Chl-a, TSM, Kd, SST, HAB, algae toxins	
Western Scheldt, entire Dutch coast (optional)	Chl-a (S2, S3), TSM, TUR, Kd (S2, L8, S3), CDOM (S3), SST (L8, GHRSSST), HAB (S3)	Coastal erosion and accretion, source identification, time series, statistics, trends	-	Chl-a, TSM, Kd, SST, HAB	
Bourgneuf Bay & Loire Estuary	Chl-a (S2, S3, MODIS), TSM, TUR, Kd (S2, L8, S3), CDOM (S3), PC	Phytoplankton bloom, time series, trends,	Spectral data, ground-truthing data	Oyster/mussel density, T, Bivalve	Phytoplankton (REPHY)

	(S3), SST (L8, GHRSSST), HAB (S3)	phenology, WFD indicators, spatial suitability maps, coastal erosion and accretion	(seagrass, bivalve beds, macroalgae)	growth and mortality	network IFREMER)
Glénans Archipelago	Chl-a (S2, S3), TSM, TUR, Kd (S2, S3), CDOM (S3), SST (L8, GHRSSST), HAB (S3), Seagrass/macroalgae (S2, L8), HAB (S3)	Phytoplankton bloom phenology	Spectral data, ground truthing data (vegetation percentage cover)		
Northern Adriatic coast (or subset thereof)	Chl-a (S2, S3, MODIS), TSM, TUR, Kd (S2, L8, S3), SST (L8, GHRSSST), Seagrass/macroalgae (S2, L8), Sublittoral habitats (S2, VHR), HAB (S3), PSC (S3), PP (S3)	Phytoplankton bloom phenology, WFD indicators, spatial suitability maps, source identification, coastal erosion and accretion	Spectral data, ground truthing data, PSC, PP	Chl-a, TUR, SD, T, HAB, Presence and status of seagrass, TSM	
Galician coast	Chl-a (S2, S3, MODIS), TSM, TUR, Kd (S2), CDOM (S3), SST (L8, GHRSSST), Seagrass/macroalgae (S2, L8), HAB (S3), PSC (S3), PP (S3)	Integration with predictive HAB modelling	Spectral data, PP, Chl-a, CDOM, PSC, TSM, environmental variables, Phyto. Taxonomy	Environmental variables, Closure Date and Zone by HABs	Data Bases from INTECMAR (Ambiental variables, Phyto. Taxonomy) Upwelling Index

Mediterranean and Black seas	TSM, TUR, Kd (S2, L8, S3), SST (L8, GHRSSST), Seagrass/macroalgae (S2, L8), HAB (S3)		-		
-------------------------------------	--	--	---	--	--

5 Other data

In addition to the two main data types generated in CoastObs, some other data are also collected. A short description of these and their handling is given below.

5.1 Data collected from external sources

For the product development and validation activities, the partners are also collecting data from various sources, including national monitoring networks or other research groups. Where possible, these are shared within the consortium, but not outside the consortium. For the internal sharing of the data, the same principles (formats, storage location, meta data) should be applied as for the CoastObs data.

5.2 Personal data

For some of the activities carried out during the project (in particular interacting with the users and outreach and exploitation activities) it will be necessary to collect basic personal data (name, contact details, background) of project stakeholders. The amount of personal data collected will be limited to data absolutely essential for the project activities. These data will be protected in compliance with the EU's Data Protection Directive aiming at protecting personal data. Where necessary, also national regulation concerning personal data protection will be implemented.

5.3 Research outputs

Research outputs, such as publications, abstracts, posters and presentations, will – as much as possible and relevant – be made publicly available via open access channels. All peer-reviewed publications will be open access, preferably using the “gold” route of publishing in an open access journal. If, for some reason, that is not possible, open access will be granted via the “green” route of making the publication available via an institutional, disciplinary or general-purpose repository. It is the responsibility of the lead author to ensure that open access is provided to each publication according to the H2020 rules. Other publications such as popular science articles, abstracts, posters and presentations will as much as possible also be made publicly available.

To increase the findability of the research results, publication meta-data (and, where applicable, full texts) will also be made available via OpenAire, Research Gate (CoastObs at Research Gate: <https://www.researchgate.net/project/Commercial-service-platform-for-user-relevant-coastal-water-monitoring-services-based-on-Earth-observation-CoastObs>), and also via institutional websites and the CoastObs website.