

FOCCUS

Forecasting and observing the open-to-coastal ocean
for Copernicus users

D1.1 Data Management

DMP- Data Management Plan

27 June 2024 / Version 1.0



Funded by
the European Union

Funded by the European Union (Grant Agreement No. 101133911). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

About this document

Title	D1.1 Data Management
Work Package	WP1 Management Task 1.3 Data Management architecture and IT solutions for the projects tasks Subtask 1.3.1 Data management plan
Lead Partner	Helmholtz-Zentrum Hereon GMBH
Lead Author (Org)	Antonio Novellino, Kelli Johnson (Hereon)
Contributing Author(s)	<ol style="list-style-type: none"> 1. Helmholtz-Zentrum Hereon GMBH (Hereon): 2. Mercator Ocean (MOI): 3. Fondazione Centro Euro-Mediterraneo sui Cambiamenti Climatici (CMCC): 4. Service hydrographique et océanographique de la Marine (Shom): 5. Collecte Localisation Satellites (CLS): 6. Consorcio Para el Diseño, Construcción, Equipamiento y Explotación del Sistema de Observación Costero de las Illes Balears (SOCIB): 7. Consiglio Nazionale delle Ricerche (CNR): 8. Stichting Deltares (DELTARES): Lorinc Meszaros 9. Stiftelsen Nansen Senter for Miljø og Fjernmåling (NERSC): 10. Meteorologisk Institutt (MET.NO): 11. Sveriges Meteorologiska och Hydrologiska Institut (SMHI): 12. +Atlantic Associação Para um Laboratório Colaborativo do Atlântico (+ATLANTIC): Antonio Novellino (ETT) 13. Institut Royal des Sciences Naturelles de Belgique (RBINS): 14. Directia Hidrografica Maritima (MHD): 15. EuroGOOS: 16. Brockmann Consult GMBH (BC): 17. Danmarks Meteorologiske Institut (DMI): 18. Institut Français de Recherche Pour l'Exploitation de la Mer (Ifremer): 19. Seascope Belgium (SSBE):
Reviewers	Joanna Staneva (Hereon)
Type of Document	Report
Due Date	30.06.2024, M6
Submission Date	27.06.2024
Version	1.0

Dissemination Level

- | | |
|-------------------------------------|--|
| <input checked="" type="checkbox"/> | PU: Public |
| <input type="checkbox"/> | PP: Restricted to other programme participants (including the Commission) |
| <input type="checkbox"/> | RE: Restricted to a group specified by the consortium (including the Commission) |
| <input type="checkbox"/> | CO: Confidential, only for members of the consortium (including the Commission) |

FOCCUS: Forecasting and observing the open-to-coastal ocean for Copernicus users is a Research and Innovation action (RIA) funded by the Horizon Europe Work programme topics addressed: HORIZON-CL4-2023-SPACE-01: Strategic autonomy in developing, deploying and using global space based infrastructures, services, applications and data 2023. Start date: 01 January 2024. End date: 31 December 2026.



Funded by
the European Union

Funded by the European Union (Grant Agreement **No. 101133911**). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them.

Table of Contents

1. Executive Summary	6
2. Scope of the document	7
3. Data Summary	8
3.1. Purpose of data generation/reuse.....	9
3.2. FOCCUS in situ data, EO observations, and model data.....	10
3.3. Data types and data formats.....	15
3.4. Data Quality.....	16
3.4.1. Data quality principles.....	16
3.4.2. Data traceability.....	17
3.5. Data collected from stakeholders.....	17
3.6. Data utility outside the project.....	18
4. FAIR Data	19
4.1.1. Data licence statement.....	21
4.2. Other actions to make FOCCUS data/products FAIR.....	21
5. Impact	21
5.1. Contribution to the project objectives.....	21
5.2. Contribution to the expected outcomes.....	22
5.3. Policy relevance.....	22
5.4. Communication, dissemination, and publications.....	22
6. Other research outputs	22
7. Allocation of resources	24
7.1. Costs before making data or other research outputs FAIR.....	24
7.2. Responsible for data management.....	24
7.3. Long term preservation.....	24
8. Data Security	24
9. Ethics and GDPR	25
10. Updates of the DMP	25
11. References	26

1. Executive Summary

FOCCUS provides products to both internal and external users. The data workflow is based on a chain of individual operational components, each contributing either to data production or user interaction functions. The infrastructure implements open and free of charge data access, it adopts international standards, follows European regulations such as INSPIRE, it is accessible under the FOCCUS web domain (<https://foccus-project.eu/>) and offers up-to-date means of data and data products dissemination and interaction tools.

Nevertheless, the developed system cannot be considered finally closed, and for the entire duration of the project, the following evolution principles have to be considered and applied:

- The service shall be state-of-the-art in its scope, and it shall be user-driven.

This means that all service capabilities are defined, allocated and implemented with sound and state-of-the-art technical and scientific justifications and address FOCCUS internal and external user needs. If new needs pop up, it is important to try and match (whenever possible) these new needs.

Providing international level state-of-the-art tools means a constant monitoring and application of most appropriate methodologies at a given time to meet requirements for service quality in data processing, validation, etc.

- Common standards are adopted and applied throughout the different elements of the system.

This means that standards are identified, adopted and applied. If, during the lifetime of the project, rising ocean best practices (OBP) become standards, the project has to try and include (as much as possible) these new OBP.

This also means that the system evolves constantly in response to service requirements, scientific and technical requirements, marine environment products requirements, and stakeholders' requirements.

- Quality of the methodology of workflow has to follow and apply well-established engineering methodologies derived from industry (es. ISO9001:2015, ISO15288, etc.) and drive the definition of an adapted approach.

This means that data management also adopts good practices and methodologies to ensure system monitoring, fixing and reporting.

Within this framework, FOCCUS poses itself as up-to-date in terms of data infrastructure, adopts the latest interoperability best practices, and enables data products FAIRness to maximise the FOCCUS data value chain from production to its consumption.

2. Scope of the document

This document presents the preliminary version of the Forecasting and observing the open-to-coastal ocean for Copernicus users (FOCCUS) Data Management Plan (DMP) and it is based on the Horizon Europe Data Management Plan (DMP) template. FOCCUS moves beyond open access in implementing open science practices while following Findability, Accessibility, Interoperability and Reusability (FAIR) principles. It applies to research data collected during the project as well as to publications.

All observational data and derived products will be made freely available via EU public data repositories and a data management strategy will be implemented (WP1) for providing immediate and free access to collected data. Besides adopting open data management software tools and supporting INSPIRE themes (e.g. European harmonised and controlled vocabularies), FOCCUS will contribute to the development of needed extensions and will share these outcomes via open repositories (e.g. Zenodo, GitHub, StackOverflow, Open Science Framework (OSF)) and repositories of partner institutes. Project developed data management software tools will be made freely available under creative commons/open data commons licensing.

The document recalls the scope of the FOCCUS project and describes the FOCCUS framework for the research data management according to the requirements of article 17 and analysed in the Annotated Grant Agreement, article 17. These include: description of data (data summary), approach to FAIRness, research outputs, allocation of resources, security and ethics.

Document Disclaimers:

This document is based on the Horizon Europe Data Management Plan Template [V1.0 May 2021] and adapted to provide the FOCCUS partners and FOCCUS stakeholders a living guide for the project Data Management Plan. The Community is not responsible for any use that might be made of the content of this publication.

3. Data Summary

The coastal zones are the most heavily used and impacted areas of the global ocean, as a consequence of large concentration of the human population inhabiting coastal areas, high number of human activities, and diverse array of resulting human pressures and anthropogenic stressors (e.g., pollution, eutrophication, poor water quality, degradation of coastal ecosystems and habitats).

Copernicus Marine Service is already providing key data and information on the coastal zone, however the use of Earth Observation (EO) data for coastal monitoring and forecasting is not explored enough, thereby the objective of FOCCUS is to address and enhance the coastal extension of the Copernicus Marine Environment Monitoring Service (CMEMS) to better serve coastal users and Member States (MS), and to develop advanced and seamless coastal monitoring and forecasting systems. This is essential and timely so that the existing CMEMS can evolve as a key coastal data and information service for Europe, working in collaboration with the European Marine Observation and Data Network (EMODnet) to ensure optimised and interoperable in situ data provision, to provide the high quality, trusted, marine knowledge needed for evidence-based management and protection of the coastal zones.

In this framework, timely, free and unrestricted exchange of marine and coastal observational data is essential for empowering decision makers with evidence-based information for sustainable management of the very dynamic and complex coastal environment.

Moreover, considering that coastal zone monitoring and processes are coupled with clear policy drivers that require MS to monitor and assess, manage and protect their national waters in order to meet important EU policies (e.g., MSFD, WFD, MSPD, CFP) and wider policy initiatives (e.g., the European Green Deal) and the Horizon Europe Mission: Restore our Ocean and Waters, an open and free data policy and support for data interoperability are mandatory actions.

Notably, interoperability of data systems has become a priority with the development of FAIR principles (2014) (<https://www.force11.org/group/fairgroup/fairprinciples>), a set of guiding principles for scientific data management and stewardship (Wilkinson et al., 2016) was developed to make data Findable, Accessible, Interoperable, and Re-usable.

The availability and accessibility of coastal data is not always easy and often the sources are lacking compliance to international standards, such as using different formats, wide diversity of datasets, and disparate data management structures, among others (Tanhua et al., 2019 described this issue for oceanographic data but it applies well to coastal zones too).

With this background in mind, this document describes the FOCCUS Data Management Plan (DMP). More specifically it describes how the research data collected during the project will be managed to move beyond the FAIR concepts and how users would be able to have immediate and free access to collected data.

3.1. Purpose of data generation/reuse

The objective of FOCCUS is to address and enhance the coastal extension of the Copernicus Marine Environment Monitoring Service (CMEMS) to better serve coastal users and Member States (MS), and to develop advanced and seamless coastal monitoring and forecasting systems. FOCCUS uses a combination of remote sensing and in situ data, and numerical models, each improved using AI methods.

More specifically FOCCUS sets out three pillars to pursue this innovative vision:

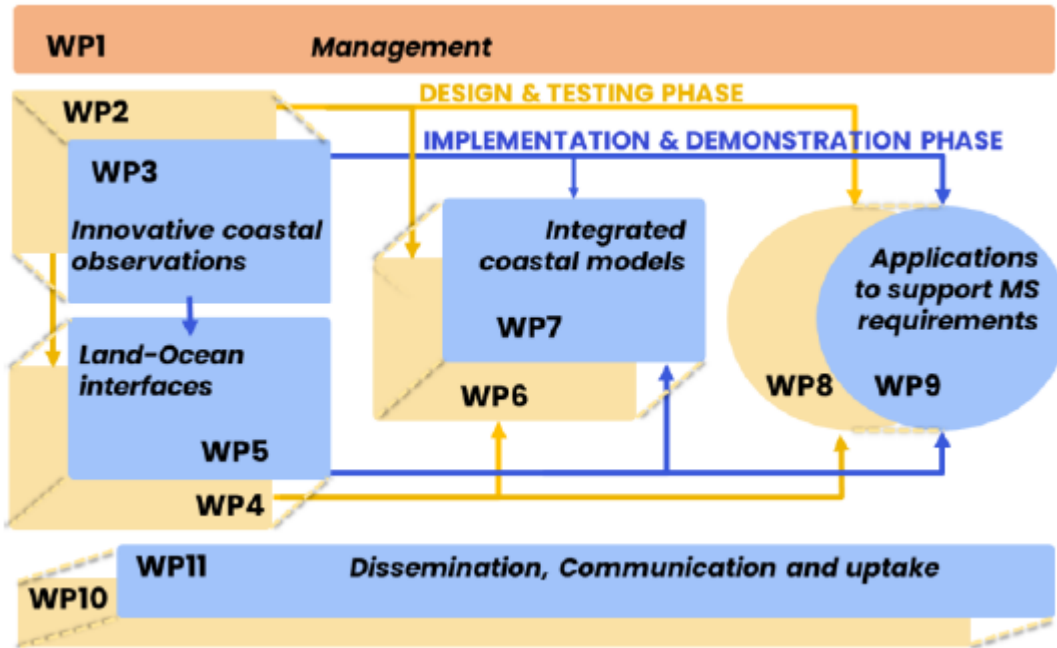
- to support the exploitation of next generation coastal INS and R/S data and link to hydrology
- to develop the first end-to-end pan-European operational coastal models
- to set the foundation of fully integrated Copernicus and MS coastal systems (MSCS)

These pillars strongly deal with digital data, more specifically they set some “data” goals:

- Physical (PHYS) and biogeochemistry (BGC) essential ocean variables (EOVs) in coastal areas are yet poorly represented in CMEMS
- Real-time or forecast river discharges are not sufficiently used in the model forcing
- Estuarine Box Model (EBMs) are not fully exploited by CMEMS and MSCS to resolve estuary dynamics and to connect to hydrological models
- MSCS are not yet making optimal use of CMEMS data (e.g., satellite data, models), or sufficiently include new modelling technologies

More specifically FOCCUS is going to reuse available data (Copernicus Marine Service and EMODnet), will facilitate the discovery of new in situ data to be linked in and included in both CMEMS and EMODnet in situ (coastal) data collections, will improve the interoperability and the methodology to process these data, will develop/improve models to serve MSCSs users. FOCCUS is not deploying new sensors or platforms and it is not generating new in situ data. FOCUS is developing new models that are going to be included in the CMEMS catalogue and new virtual in situ data (from the EBMs) that will be openly available as soon as possible, as soon as they have been scientifically validated,

The figure shows the FOCCUS structure and addressed goals:



Each WP is organised in tasks and each task is either producing, assessing or developing a data/product. In particular, WP2 to WP9 are the data and product generators.

3.2. FOCCUS in situ data, EO observations, and model data

As anticipated in the previous section, FOCCUS is organised in WPs and tasks, each (WP2-WP9) task is consuming or producing data products being in one of these three main categories:

1. in situ (INS) observational data
2. satellite derived Earth Observational (EO) products, both new data production and consumption of existing data
3. numerical models (NMO)

The following table summarises this planning:

	INS	EO (producer)	EO (consumer)	NMO
WP2, WP3	x	x	x	
WP4, WP5	x		x	x
WP6, WP7	x		x	x
WP8, WP9	x		x	x

Table 1: Data use details, per WP

In WP2 “New insight of high-resolution coastal observations” and WP3 “Validation and Implementation of high-resolution coastal observations”, FOCCUS aims to improve pan-European innovative coastal observations in response to society and EU policy needs in the context of climate change. The improvement of the pan-European innovative coastal observations will then be assessed. Details are given in the table below:

WP	2 and 3
Partner	HEREON, CLS, SOCIB, CNR, NERSC, MET.NO, +ATLANTIC, MHD, BC, DMI, Ifremer, SSBE
Contact person and email	Emma Reyes (ereyes@socib.es) and Antonio Bonaduce (antonio.bonaduce@nersc.no)
Dataset name / brief description	<p>In WP2/3 novel remote-sensing data products will be designed and validated to:</p> <ul style="list-style-type: none"> - characterise: essential climate variables (such as sea level, sea surface temperature, sea ice change) -enhance weather and marine services (e.g. surface ocean currents) - monitor ocean health and ecosystem (biogeochemical anomalies and trends, seagrass and macroalgae) and - monitor coastal changes (shoreline positions). Data-fusion techniques will be employed to merge remote-sensing and in-situ data, providing improved estimates of coastal ocean circulation, harmful algal blooms (HABs), and inland-marine water connectivity. <p>Additionally, new functionalities, algorithms and AI techniques will be developed and tested to improve the quality of coastal observations (e.g., beach morphodynamics) and the traceability of coastal in-situ and river data.</p> <p>To be detailed in the deliverable 2.1 (section 3)</p>
Preexisting or generated in FOCCUS?	Generated, developed and/or improved in FOCCUS
If already existing, DOI or link	NA
Purpose, target user	To be detailed in the deliverable 2.1 (section 4)
in situ, remote sensing, model?	In-situ observations, remote sensing, ML model
Data source (instrument, sensor, model short name, samples: sampler, analysis instrument)	Sentinel-6A HR mode, SWOT, Sentinel 3, Sentinel 3, SAR, Sentinel-1A/B SAR, Landsat, multi-platform observations

Level of processing and description	To be detailed in the deliverable 2.1 (section 4)
Variables and units	To be detailed in the deliverable 2.1 (section 4)
Expected accuracy of data (description for each variable)	To be detailed in the deliverable 3.1
Link to expected accuracy of data (es. datasheet of the sensor)	To be detailed in the deliverable 3.1
Spatial extent and resolution	To be detailed in the deliverable 2.1 (section 3)
Time span and resolution	To be detailed in the deliverable 2.1 (section 3)
Metadata provided	To be detailed in the deliverable 2.1 (section 4)
File format	To be detailed in the deliverable 2.1 (sections 2 and 3)
applied vocabularies / conventions	To be detailed in the deliverable 2.1 (section 4)
link to product	To be detailed in the deliverable 2.1 (sections 1 and 2)

In WP4 “Design of improved model interfaces in the Land-Ocean Continuum” and WP5 “Building, validation and demonstration of improved model interfaces in the Land-Ocean Continuum”,

FOCCUS aims at improving hydrological modelling (E-HYPE and LISFLOOD) , estuarine models (i.e. upgrades of physics and ML based Estuary Box Model, learning dataset by unstructured grid to support EBM training, EO satellite to support EBM testing) and land-sea boundary conditions targeting improvements of estuarine dynamics and representation of river discharge at the riverine outlets. Hydrological models will be improved through integration of nutrient and sediment transport, which sources are estimated from European databases and calibrated to instream water quality observations. Other foreseen upgrades entail the merging of physical and ML/AI-based approaches for estuarine dynamics and better coupling techniques at river mouths. In WP5 the advancements of the hydrological and estuarine models will be implemented and validated in regional and coastal models across the EU Copernicus Marine Service domain, using EMODnet near-real time data and EO satellite data in both estuarine and plume areas. Details are given in the table below:

WP	4 and 5
Partner	SMHI, CMCC, DMI, +ATLANTIC, DELTARES, HEREON, CNR, MHD, MET.NO,
Contact person and email	Alena Bartosova (alena.bartosova@smhi.se) and Giorgia Verri (giorgia.verri@cmcc.it)
Dataset name / brief description	Base time series with freshwater inputs to coasts (volume flux, physical tracers as temperature and salinity , nutrients, sediment) for current and improved hydrological models and by adding an

	estuarine box model downstream of the hydrology ones to solve the estuarine water exchange at the net releases at river mouths
Preexisting or generated in FOCCUS?	Preexisting, improved and/or new products within FOCCUS
If already existing, DOI or link	
Purpose, target user	Inputs for regional to coastal models
in situ, remote sensing, model?	model
Data source (instrument, sensor, model short name, samples: sampler, analysis instrument)	model (E-HYPE, LISFLOOD) 1D ML-based Estuary Box Model 3D physically based finite element model (SHYFEM)
Level of processing and description	
Variables and units	Volume flux (m ³ /s), temperature (celsius degree), salinity (psu), nutrients and sediment (weight / volume, alt. weight / time step)
Expected accuracy of data (description for each variable)	to be described in D4.2
Link to expected accuracy of data (es. datasheet of the sensor)	N/A
Spatial extent and resolution	to be described in D4.2
Time span and resolution	to be described in D4.2
Metadata provided	to be described in D4.2
File format	text file, netCDF
applied vocabularies / conventions	
link to product	

In WP6 “Implementation of new interfaces and methodologies in coastal systems” and WP7 “Coastal Modelling Systems - validation and demonstration”, FOCCUS targets seamless modelling capabilities of the pan-European coasts, from land to open ocean, through the integration of existing Copernicus products into the different Member State Coastal modelling System (MSCS). These data include both numerical and observational data, the former being principally used to force MSCS while the latter being principally used for data assimilation or for MSCS validation. Improvements of the MSCS themselves are also targeted, including ML/AI-assisted methods or ensemble approaches, and the added value of these new types of data for decision support and other stakeholders needs will be assessed. Details are given in the table below:

WP	6 and 7
Partner	MET.NO, DELTARES, +ATLANTIC, SHOM, RBINS, HEREON, SMHI, CMCC, DMI, MOi, MHD, CNR, NERSC, SOCIB
Contact person and email	Kai Håkon Christensen (kaihc@met.no) and Quentin Jamet (quentin.jamet@shom.fr)
Dataset name / brief description	Data delivery from several coastal systems, as detailed in inventory. Main outputs are physical variables such as salinity, temperature, mean surface heights, ocean currents. For this plan, we restrict ourselves to variables that are specifically addressed in FOCCUS tasks.
Preexisting or generated in FOCCUS?	Preexisting.
If already existing, DOI or link	To be detailed in D9.2.
Purpose, target user	General public and MSCS downstream decision support systems.
in situ, remote sensing, model?	Model.
Data source (instrument, sensor, model short name, samples: sampler, analysis instrument)	See MSCS inventory.
Level of processing and description	To be detailed in D9.2.
Variables and units	To be detailed in D9.2.
Expected accuracy of data (description for each variable)	To be detailed in D9.2.
Link to expected accuracy of data (es. datasheet of the sensor)	N/A
Spatial extent and resolution	To be detailed in D9.2.
Time span and resolution	To be detailed in D9.2.
Metadata provided	Yes
File format	To be detailed in D9.2.
applied vocabularies / conventions	CF compliant
link to product	To be detailed in D9.2.

In WP8 “Design of Coastal Applications” and WP9 “Building, validation and demonstration of Coastal Applications”, FOCCUS aims to evaluate the improvements made with respect to coastal observations, land-ocean coupling (hydrology), and modelling interfaces for coastal

applications along the identified key Environmental and Societal Challenges. WP8 and WP9 will therefore incorporate the improved data products delivered in WP2-WP7. The coastal applications are local configurations of the member state coastal systems that require initial conditions, forcing fields, boundary conditions, river inputs but also independent observations for calibration and validation (in situ and satellite). These inputs originate from various data providers. The coastal applications produce numerical model outputs in the form of time series, two- or three-dimensional environmental fields or aggregate statistics such as fluxes. The model outputs are either distributed in their raw netCDF data formats or post-processed into information services, such as figures, tables, and textual summary. Details are given in the table below:

WP	8 and 9
Partner	CMCC, DELTARES, RBINS, CNR, HEREON, MHD, BC, MET.NO, NERSC, SHOM, MOi, DMI, +ATLANTIC, SOCIB
Contact person and email	Ivan Federico (ivan.federico@cmcc.it) and Lorinc Meszaros (lorinc.meszaros@deltares.nl)
Dataset name / brief description	Coastal Applications: The coastal applications are organised in three groups addressing environmental and societal challenges. The first group of applications aims at supporting better management and protection of the coastal area against pollution and coastal erosion; The second group of applications aims at supporting multi-use of coastal and offshore operations with early warning systems for risks associated to for e.g. HABs, pathogens in aquaculture and cascade impacts from offshore installations; The third group of applications aims at supporting resilience to climate change through for e.g. NBS (seagrass restoration) and prevent natural and anthropogenic hazards and extreme events (e.g. MHWs and storm surges).
Preexisting or generated in FOCCUS?	Preexisting, improved and/or new products within FOCCUS: New products based on pre-existing models (WP6-7), new observation data (WP2-3), improved hydrological data (WP4-5).
If already existing, DOI or link	To be detailed in the deliverable 9.2
Purpose, target user	Coastal applications seek to better address the selected Environmental and Societal Challenges: ESC 1 (“Better Manage and Protect the Coastal Area.”), ESC 2 (“Enhance the Blue Economy.”) and ESC 3 (“Natural and Anthropogenic Hazards and Building Resilience to Climate Change”). Target users: MSCS operators with reference users

	community linked to climate, weather and marine services, ocean health and ocean changes.
in situ, remote sensing, model?	All these features are combined in the coastal applications.
Data source (instrument, sensor, model short name, samples: sampler, analysis instrument)	Data sources for FOCCUS coastal applications combine data sources used for WP6/7 (see MSCS inventory), WP4/5 (Land Ocean models; E-HYPE, LISFLOOD, 1D ML-based Estuary Box Model, 3D physically based finite element model, SHYFEM), WP2/3 (observation data; Sentinel-6A HR mode, SWOT, Sentinel 3, Sentinel 3, SAR, Sentinel-1A/B SAR, Landsat, multi-platform observations).
Level of processing and description	To be detailed in deliverable 9.2
Variables and units	To be detailed in deliverable 9.2
Expected accuracy of data (description for each variable)	To be detailed in deliverable 9.2
Link to expected accuracy of data (es. datasheet of the sensor)	N/A
Spatial extent and resolution	To be detailed in deliverable 9.2
Time span and resolution	To be detailed in deliverable 9.2
Metadata provided	To be detailed in deliverable 9.2
File format	To be detailed in deliverable 9.2
applied vocabularies / conventions	CF compliant
link to product	To be detailed in deliverable 9.2

3.3. Data in Coastal Systems Inventory

As part of the publicly available Milestone MS6.1 Interfaces Roadmap, existing uses of data in MS Coastal Systems are inventoried and roadmaps of each MS Coastal System have been created for their future development.

- The inventory can be found in the shared project folder here: https://docs.google.com/spreadsheets/d/1JxFKYBhboVHDFwAgmMikijJ_9N2UFUu6xNmFOAGB21M/edit?usp=sharing
- The roadmaps can be found in the shared project folder here: https://drive.google.com/drive/folders/1tlmc7NZxheF69Is86elyiv3eLRsgggPH?usp=drive_link

3.4. Data types and data formats

As anticipated, FOCCUS is developing new tools to enhance the coastal extension of Copernicus Marine Service (CMEMS) to better serve coastal users and Member States (MS), thereby FOCCUS will adopt the CMEMS data types and formats. The vast majority of Copernicus Marine products are produced in NetCDF format (.nc). With the entry into service of the new CMEMS Marine Data Store, the infrastructure also supports cloud based formats such as Zarr. Other important file formats for the FOCCUS scope are GeoTIFF and Shapefile data.

The NetCDF (Network Common Data Form) format is a self-describing, platform-independent binary data format commonly used for storing and sharing scientific data, particularly in fields such as climate science, meteorology, oceanography, and geophysics. NetCDF files are designed to be flexible, efficient, and portable, making them suitable for handling large volumes of multidimensional data. As described above, all Copernicus Marine products are delivered in NetCDF-4 format by default. Notably, NetCDF is also the recommended data format for the in situ observational data by both CMEMS INSTAC and EMODnet.

The Zarr format is a Python library and file format designed for efficient storage and manipulation of multi-dimensional array data. Zarr is particularly well-suited for handling large-scale scientific datasets, especially in the context of parallel and distributed computing. This is a key data format to facilitate uninhibited flow of data to the EU Digital Twin Ocean initiative (EU DTO) - EDITO data lake (<https://www.edito.eu/>).

The GeoTIFF format, short for Georeferenced Tagged Image File Format, is a type of raster image file that contains spatially referenced information, making it suitable for use in Geographic Information System (GIS) applications and remote sensing analysis.

The Shapefile format is a popular geospatial vector data format developed by Esri (Environmental Systems Research Institute) for storing and exchanging geographic data in a simple, non-topological format.

3.5. Data Quality

The platform data.europa.eu (i.e. the official portal for European Union (EU) data) has recently published the data quality guidelines for delivering high-quality open data. These recommendations are directed at data providers to support them in preparing their data, developing their data strategy and ensuring data quality.

FOCCUS is adopting these recommendations and in particular is applying measures for making data FAIR, using, as described in this document, well known and standardised data models and formats, using common vocabularies and standards (ISO and community defined dictionaries), adopting easy data access and data download tools.

Moreover, all data collected and produced in FOCCUS will refer/document the applied (scientific) methodology (citation of references). The following data quality principles will be applied:

3.5.1. Data quality principles

The principles aim to encourage all involved in research to maintain, and to respond adequately to possible threats or violations of research integrity. FOCCUS researchers should be guided by these principles at each stage of the project work and in particular they have to ensure:

3.5.1.1. Academic Excellence

Researchers have to apply sound methodology, suitable methods, standard procedures and documented protocols where appropriate, to ensure the highest quality of work, its dissemination and replicability.

3.5.1.2. Honesty

Researchers must acknowledge the direct and indirect contributions of colleagues, collaborators and others. The basic principle of any scientific activity is the need for researchers to be honest in respect of their own actions and in their responses to the actions of other scientists.

3.5.1.3. Accountability

Researchers should recognise their responsibility to the general public, and should take all reasonable measures to ensure that their research complies with any agreement, related policies and professional bodies' guidance; and allows for proper governance and transparency.

3.5.1.4. Care and Respect

Researchers should avoid any unreasonable risk or harm to research subjects and researchers themselves.

3.5.2. Data traceability

In line with the above described principles, FOCCUS is adopting all necessary measures to ensure data traceability. In all aspects of research, substantive contributions of formal collaborators and all others who directly assist or indirectly support the research must be properly acknowledged. This applies to any circumstances in which statements about the research are made, including provision of information about the nature and process of the research, and in publishing the outcome. Failure to acknowledge the contributions of others is regarded as unethical conduct.

This applies to data and data products that are going to be (re)used in the project and is achieved by documenting the sources and the methods (references, DOI, or any other identifier whenever available/applicable). It also applies data products that are collected and produced by FOCCUS by properly documenting and by assigning a DOI to each of them.

3.6. Data collected from stakeholders

In FOCCUS, responses from different stakeholder surveys will be used and additional surveys could be developed to co-design FOCCUS R&I and to gather needs from stakeholders (WP10-WP11).

These data will be handled following the General Data Protection Regulation (GDPR).

3.7. Data utility outside the project

FOCCUS (Forecasting and Observing the Open-to-Coastal Ocean for Copernicus Users) is part of the 2023 Horizon Europe calls for Space supervised by HaDEA, the Health and Digital Executive Agency, and aims at enhancing the Copernicus Marine Service's existing capabilities and developing innovative coastal monitoring and forecasting products. This includes changes in the use of new space and in situ coastal observations, innovations in data fusion, data processing and visualisation, together with seamless numerical forecasting from regional models of the Copernicus Marine Service covering the EU regional seas, to MS coastal forecasting systems

Coastal zones have large social, economic and environmental values. They are more densely populated than the hinterland and concentrate large economic assets, critical infrastructures and human activities such as tourism, fisheries, navigation. Furthermore, the coastal ocean is home to a wealth of living marine resources and very productive ecosystems. Yet, coastal zones are exposed to various natural and anthropogenic hazards, such as pollution, eutrophication, and habitat degradation. Natural hazards can be exacerbated by climate change and induced ocean warming, acidification, deoxygenation, sea level rise, increases in harmful algae blooms and invasive species.

While Europe has a robust capability for ocean monitoring, particularly through the Copernicus Marine Service, the use of Earth observation (EO) data for coastal monitoring is not fully explored. The FOCCUS project aims to enhance the Copernicus Marine Service's coastal dimension by fostering collaboration between European Research and Innovation entities (R&I), public authorities (including coastal services) and industry stakeholders. This involves producing coastal information through the creation of new computational models utilising data from both MSCS and the Copernicus Marine Service for improved monitoring and forecasting of coastal zones.

The initiative aligns with EU policies, including the European Green Deal, and contributes to global efforts such as the United Nations Decade of Ocean Science for Sustainable Development.

4. FAIR Data

Over the last three decades great progress was made at European level in terms of advancing these standards and facilitating machine-to-machine data processing, and programs and projects such as Copernicus Marine Service, EMODnet, SeaDataNet, H2020 EuroSea etc are continuously updating these recommendations (see Tanhua et al. 2019 for an exhaustive review). These recommendations are made according to the FAIR principles broken down into the 15 characteristics laid down by the FORCE11 collective and clearly described in Quimbert et al. (2022).

Findable	
F1	(meta)data are assigned a globally unique and eternally persistent identifier
F2	data are described with rich metadata
F3	(meta)data are registered or indexed in a searchable resource
F4	metadata specify the data identifier
Accessible	
A1	(meta)data are retrievable by their identifier using a standardised communications protocol
A1.1	the protocol is open, free, and universally implementable
A1.2	the protocol allows for an authentication and authorization procedure, where necessary
A2	metadata are accessible, even when the data are no longer available
Interoperable	
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2	(meta)data use vocabularies that follow FAIR principles
I3	(meta)data include qualified references to other (meta)data
Reusable	
R1	(meta)data have a plurality of accurate and relevant attributes
R1.1	(meta)data are released with a clear and accessible data usage licence
R1.2	(meta)data are associated with their provenance
R1.3	(meta)data meet domain-relevant community standards

Table 2: FAIR principles proposed by the FORCE11 community

Notably, these recommendations always include three main elements for implementing interoperability: metadata (including data quality), data format, and data services.

More specifically, according to our Grant Agreement (Art. 187), metadata of deposited publications must be open under a Creative Commons Public Domain Dedication (CC 0) or equivalent, in line with the FAIR principles (particularly machine-actionable), and must provide information on at least the following:

- Title
- Date of publication
- Publication venue
- Horizon Europe funding: HORIZON-CL4-2023-SPACE-01
- Grant project name, acronym, and number: FOCCUS, Funded by the European Union, Grant Agreement No. 101133911
- Licensing terms
- Persistent identifiers for the publication
- Authors involved in the action (preferably using ORCID) and, if possible, their organisations (preferably using the PIC and EDMO code)

The usability of data and outcomes strongly depends on the data policy license. There is an increasing push for adopting the Creative Commons framework, particularly the CC-BY license (the only limitation being that credit must be given to the creator).

Embargoes on data and other CC licenses can also be applied where the policy is 'As open as possible, as closed as necessary'. Partners can avoid offering open access in two cases:

- 1) If open access would be against the beneficiary's legitimate interests, including commercial exploitation
- 2) If open access would be contrary to any other constraints, particularly the EU's competitive interests or the beneficiary's obligations under this Agreement

If open access is not provided (to some or all data), this must be justified in the DMP (as annex).

Going narrower into the dataset/product description, according to the ISO-19115, the minimum set of metadata to ensure comprehensive identification of the dataset/product has to include its spatial dimension, temporal dimension, depth, source and quality control for each element.

For the spatial dimension, the WGS84 standard must be used to ensure accurate geographic positioning. For the temporal dimension, the ISO 8601 standard must be used, which provides a clear and consistent format for representing dates and times. The vocabulary used in keywords should be consistent with the GCMD to ensure consistency and facilitate data discovery.

For the parameters and units, the Copernicus Marine Service conventions should be applied.

4.1.1. Data licence statement

FOCCUS is an EU supported research project, hence all the collected data have to be freely available to the community at no cost and limitation. As described, in FOCCUS, all contributions will be Open Access and data will be openly available with licences following the CC BY 4.0 "unrestricted" standard (or equivalent), while metadata will follow the CC 0 licence (see previous section).

Digital data will be made available through the Copernicus Marine Service.

4.2. Other actions to make FOCCUS data/products FAIR

FOCCUS outcomes would be either included into the CMEMS Data Store and listed into the CMEMS catalogue (<https://data.marine.copernicus.eu/products>) or may be presented and described as CMEMS Use Cases (<https://marine.copernicus.eu/services/use-cases>).

Moreover, FOCCUS project partners are encouraged to contribute their models to the DCC repository using the form provided on the OceanPrediction DCC website, ensuring that they are accessible to the wider ocean prediction community.

OceanPrediction DCC website - <https://www.unoceanprediction.org/en/atlas/models>

The OceanPrediction Decade Collaborative Centre (DCC) is emerging as a key initiative within the European Union's strategic actions for ocean science and sustainable development. In line with the United Nations Decade of Ocean Science for Sustainable Development, this European initiative, hosted by Mercator Ocean International, represents a concerted effort to harness collective expertise and resources to achieve a sustainable global ocean by 2030. Incorporating models into the work of the DCC can provide powerful ocean prediction capabilities around the world, widening FOCCUS stakeholders and facilitating sustainable ocean management practices.

5. Impact

5.1. Contribution to the project objectives

The FOCCUS DMP sets the schemes and rules for project data management and sharing. This applies to all project Objectives, but more specifically it is a pillar for the implementation of these initial Objectives:

- Objective 1: Improve pan-European innovative coastal observations in response to society and EU policy needs.

- Objective 2: Improve land-coast-open ocean interfaces.

In the implementation of these Objectives, new coastal data products will be created by the fusion of in situ and remote sensing coastal data and biogeochemical parameters, in addition to hydrological data. This includes an inventory of the current status of coastal data available in different EU marine data services (i.e., CMEMS, EMODnet, JERICO-RI, DANUBIUS-RI) that will be performed allowing the identification of respective gaps. Thereby, the foundation will be laid for further work on MSCS modelling and data fusion, for integration into established CMEMS global/regional systems and production of new applications in Objective 3 and Objective 4.

As described in the previous section, FOCCUS is setting up a data management infrastructure that is adopting and adapting the recommendations and lessons learnt from major European Marine Data initiatives and is interoperable towards both EMODnet and Copernicus Marine Service. It applies FAIR principles and implements tools to be INSPIRE compliant.

5.2. Contribution to the expected outcomes

The DMP is supporting the FOCCUS data management towards all its EOs.

5.3. Policy relevance

The FOCCUS Data Management Plan is setting (proposing) a better and more clear schema for data delivery and Commission assessment on the Horizon Europe project data FAIRness, more specifically it is embracing and promoting the adoption of CC and in particularly CC-BY whenever possible, and it is proposing to have clear statements on embargoed data, meaning that some data will may go under embargo for enabling scientific processing and fine tuning (and publication), but since now the project is declaring that all the data will be available and linkable. New observational capacity (modelling and new coastal products) is already declared and at the end of the project it can be easily tracked if and how much the project was compliant to this statement.

Being in contact – cooperating with the major European Marine Data Integrators (EMODnet and Copernicus Marine Service) is important to have a further channel to make value of the project data. Some FOCCUS data may be not included in these integrators (because of the scope, type ...) but it is important that the ones that are of interest to those initiatives can easily flow to/interoperate with them. The adoption of open data policy and open science principles are further cornerstones of this process and should be encouraged as a EU policy for research projects.

5.4. Communication, dissemination, and publications

Communication, Dissemination and Publication are reported in the non-public facing document: Deliverable 10.1 Communication, Dissemination, and Exploitation Plan.

6. Other research outputs

FOCCUS is an ambitious project and there will be several outputs going from new data and products to deliverables, to publications, to dissemination materials. These outputs are planned and this approach matches the open science practices. This document is describing the approach on data and data-products data management approach, while plans for communication strategy, deliverable quality, document management, web and social communication are addressed in other deliverables.

The project's data and data-products are digital objects and they will be published together with (FAIR) metadata, be assigned a digital object identifier (DOI) and be available/linkable in the FOCCUS web portal or web data infrastructure.

7. Allocation of resources

7.1. Costs before making data or other research outputs FAIR

Besides having a DM team, each partner has some allocated budget to implement fairness according to the principles and recommendations described in this document.

7.2. Responsible for data management

Data Management is included in a dedicated project work package (WP1), led by Hereon.

Responsible for the data management is Kelli Johnson (HEREON), who is the Scientific Project Manager for FOCCUS at Hereon, with a background in oceanography and data science.

Considering the complexity of the project, the Data Manager is assisted by a team of experts, with strong background in remote sensing data management, in situ data, and numerical modelling. More specifically, she will be supported by:

Joanna Staneva (HEREON), is a Senior Scientist and Head of the Department "Hydrodynamics and Data Assimilation" at Helmholtz Centre Hereon, Germany. She holds a Diploma in Physics and a PhD in Physical Oceanography from the University of Sofia, Bulgaria. Dr. Staneva has expertise in marine research, research management, and interdisciplinary collaboration. Her work includes modeling the marine environment, earth system modeling, ocean predictability, climate dynamics, and coastal and regional oceanography. She actively contributes to scientific advancement by coordinating the Horizon Europe project FOCCUS and participating in international and national projects.



Francisco Campuzano (+ATLANTIC), who has been the leader of the EMODnet physics river node and also of operational numerical models in Portugal. He is member of several EuroGOOS working groups where data and model results standard practices were defined and implemented.

Antonio Novellino (ETT, subcontractor of +ATLANTIC), who has long experience in data management and ocean data management and sharing for both Copernicus Marine Service (he was the CMEMS Dissemination Unit deputy), EMODnet (he is the Coordinator of EMODnet Physics, and leader in EMODnet Ingestion operational data ingestion flow), and several H2020 and HEurope marine data and services projects.

Whenever a FOCCUS data product is hosted, safekept into any other well known/TRUSTed repository, WP1 will list the proper links.

7.3. Long term preservation

For the duration of the project, data is going to be hosted by the FOCCUS data management infrastructure. This infrastructure is based on the key European integrating data ocean infrastructures and programs (CMEMS, EMODnet) and FOCCUS outcomes will be directly deposited in these long-term safe keeping systems. FOCCUS outcomes will be always be visible and identifiable by means of data-products metadata (and citation statement)

8. Data Security

Data security is not a concern for the specific goals of the project: FOCCUS is developing open science and open products that are going to be made available in CC-BY (or equivalent) as soon as possible.

9. Ethics and GDPR

Ethics and GDPR are covered by the FOCCUS Milestone 1.1 (MS1.1): GDPR-compliant project management portal created, including risk register.

10. Updates of the DMP

This DMP will be updated in months 18 and 36.

11. References

Tanhua, T., McCurdy, A., Fischer, A., Appeltans, W., Bax, N., Currie, K., DeYoung, B., Dunn, D., Heslop, E., Glover, L. K., Gunn, J., Hill, K., Ishii, M., Legler, D., Lindstrom, E., Miloslavich, P., Moltmann, T., Nolan, G., Palacz, A., ... Wilkin, J. (2019). What We Have Learned From the Framework for Ocean Observing: Evolution of the Global Ocean Observing System. *Frontiers in Marine Science – Ocean Observation*, 6. <https://doi.org/10.3389/fmars.2019.00471>

Quimbert Erwann, Fichaut Michele, Maudire Gilbert (2022). Guide principes FAIR. Principes FAIR dans le contexte du pôle ODATIS. <https://doi.org/10.13155/87107>