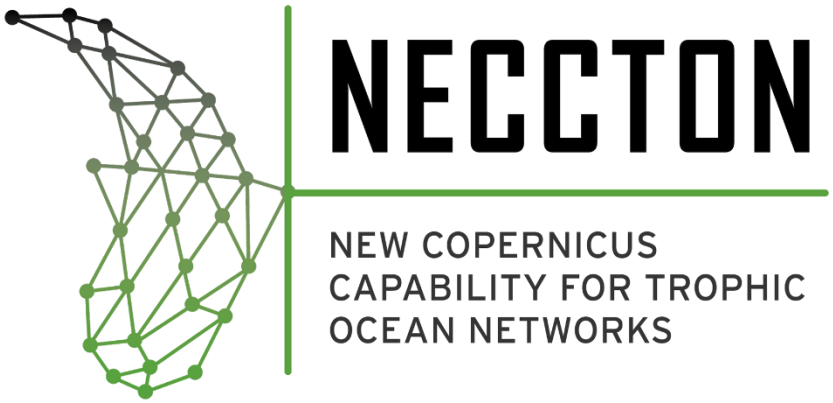
Project	NECCTON No 101081273	**Deliverable**	D1.1
Dissemination	Public	**Nature**	Data Management Plan (DMP)
Date	2023-06-30	**Version**	**1.0**



Action Number: RIA Grant Number 101081273

Action Acronym: NECCTON

Action title: New Copernicus capability for trophic ocean networks

Deliverable 1.1 "Initial data management plan (DMP)"

Date: 30/06/2023

DMP version: V1.0

Deliverable Contributors:	Name	Organisation	Role / Title
Deliverable Leader	Stefano Ciavatta	MOI	Project coordinator
Contributing Author(s)	Aurore Biardeau	MOI	WP2 contributor
	Valentina Giunta	MOI	WP2 contributor
	Corinne Derval	MOI	WP2 leader
	Laia Romero	Lobelia	WP2 contributor
	Federica Braga	CNR	WP4 contributor
Reviewer(s)	Julien Brajard	NERSC	WP4 leader
	Frank Muller-Karger	University of Florida	Advisory Board member
Final review and approval	Stefano Ciavatta	MOI	Project Coordinator

Project	NECCTON No 101081273	**Deliverable**	D1.1
Dissemination	Public	**Nature**	Data Management Plan (DMP)
Date	2023-06-30	**Version**	**1.0**

Document History:

Release	Date	Reason for Change	Status	Distribution
0.0	30/04/2023	Table of content	Revised	Internal
0.1	16/06/2023	Draft document	Revised	Reviewers
1.0	30/06/2023	Completion	Released	EC

Preface

The Horizon Europe Model Grant Agreement requires that a data management plan ('DMP') is established and regularly updated. This document is the Initial DMP of the NECCTON project, namely Deliverable 1.1. A Final DMP will be delivered at month 48. This DMP addresses the requirements for research data management of Horizon Europe as described in article 17 and analysed in the Annotated Grant Agreement, article 17. This document used the EU template for DMPs: "Grants: Data Management Template (HE):V1.0 – 05.05.2021".

To cite this document: Ciavatta S. et al (2023). D1.1 "Initial data management plan (DMP)". Report of the Horizon Europe NECCTON project (grant 101081273). DOI 10.5281/zenodo.8099396

Table of Contents

1 Data Summary 3
2 FAIR data 10
2.1 Making data findable, including provisions for metadata 10
2.2 Making data accessible. 11
2.3 Making data interoperable 12
2.4 Increase data re-use 12
3 Other research outputs 14
4 Allocation of resources 16
5 Data security 16
6 Ethics 16
7 Other issues 16
8 References 16
Appendix 17

1 Data Summary

The objective of NECCTON is to enable the European operational ocean service (namely the Copernicus Marine Service, CMEMS) to predict and protect biodiversity and support the sustainable management of marine ecosystems. New **models and simulation products** for fisheries, pollution and benthic habitats will enable CMEMS to better inform ocean policymakers, managers and publics.

NECCTON will enable CMEMS to deliver **twenty-five** new or improved operational products, in the form of model outputs (see list in Table 1). To achieve Technological Readiness Level (TRL) 5 and 6 before the 2027, as required by the call HORIZON-CL4-2022-SPACE-01 for this project, only products at current TRL 4 (at least) were considered for upgrade in NECCTON. Crucially, these products were selected because of their significance to the specific needs and outcome requested by the call:

- Sixteen NECCTON products are aligned with the UN Sustainable Development Goals (SDG) 2 "zero hunger", 12 "sustainable consumption and production", 13 “climate change” and 14 “life below water” (see column “**SDG**” in Table 1). The SDGs adopted by the UN in 2015 are a collection of 17 global goals designed to be a "blueprint to achieve a better and more sustainable future for all". The UN SDGs are central to European policy, and are promoted in key EU projects, sectoral policies and initiatives.
- Fifteen NECCTON products are related to the ‘Essential Ocean Variables’ (EOVs) and “Essential Biodiversity Variables’ (EBVs) (both listed in the column “**EV**” in Table 1.1). EOVs are physical, chemical and biological variables selected by the Global Climate Observing System (GCOS, gcos.wmo.int) and Global Ocean Observing System (GOOS, www.goosocean.org) to monitor the earth’s climate and the oceans, because of their relevance, feasibility and cost effectiveness. EBVs were introduced by the Group on Earth Observations Biodiversity Observation Network (GEO BON) to monitor the ocean biodiversity. In essence, EBVs are time series of EOVs at one location or as gridded/mapped metrics relevant to the different dimensions of biodiversity (genes, species, ecosystems).
- Fourteen NECCTON products are particularly significant to Copernicus users (column “**Use**” in Table 1), based on the EC Commission Staff Working Document "Expression of user needs for the Copernicus programme". The policy context and organizations targeted in the above document are relevant to NECCTON and include UN SDGs, EU Common Fishery Policy, Marine Strategy Framework Directive, Convention on Biological Diversity, Biodiversity Strategy to 2020, Regional Fisheries Management Organisations (RFMOs), Biodiversity Beyond National Jurisdiction Treaty.

The 25 NECCTON products will cover all the **seven marine areas** of the CMEMS Monitoring and Forecasting Centers (MFCs): 1. the Global Ocean (GLO), 2. Arctic Ocean (ARC), 3. Baltic Sea (BAL), 4. North-West European Shelf (NWS), 5. Iberia-Biscay-Ireland (IBI) area, 6. Mediterranean Sea (MED) and 7. Black Sea (BS) (see Table 1 and the figure). However, some products will not be delivered in all the areas (see the empty cells in the columns from ARC to GLO in Table 1), because of the current low TRL of the product in the area and/or constraints in the resources available in NECCTON for their development. For the same reasons, NECCTON will develop products at different TRL in the different areas (see TRL numbers from ARC to GLO in Table 1).

Some products will be delivered in the same marine area by using **different models** (multi-model ensembles), with the objective of comparing methods and evaluating the uncertainty of the products (e.g., four models are deployed in the global ocean (GLO) to deliver the “Small pelagic” product: Table 1). Overall, we expect to deliver **84 model simulation datasets** and **2 observational datasets** for the 25 products produced for different regions and models (see Table A1 in Appendix).

The full description and specifications of the products will be provided in the Deliverables 5.1, 6.1, 7.1 and 8.1. These will take account of the needs of stakeholders and potential user consulted by NECCTON (see Deliverable D2.1), most notably the “Co-design future product” workshop hold on 28-29 June 2023.

Table 1 List of the products delivered by NECCTON in the domains of the seven CMEMS MFCs. In x^i-r, x ticks the region where the product is produced, i indicates the number of models in the ensemble applied to deliver the product and r indicates the Technological Readiness Level (TRL) of the products. TRL=5 indicates that the product has been validated in NECCTON, TRL=6 the product has been demonstrated in the case studies of NECCTON. The products are developed by the work-packages (WP) of the project. The products support the UN Sustainable Development Goals 2, 12, 13 and 14 (SDG), the Essential Ocean and Essential Biodiversity Variables (EV) and the Copernicus user needs (Use). The datasets used for fusion or validation of the model-output products have been grouped in broad categories for the sake of simplicity. Some categories are described and referenced in the text. Other categories are: i) large scale publicly available datasets, which collect data from a broad range of methods and programs: e.g. Coastal and Oceanic Plankton Ecology, Production, and Observation Database (COPEPOD, www.st.nmfs.noaa.gov/copepod/) British Oceanographic Data Centre (BODC; www.bodc.ac.uk/); International Council for the Exploration of the Sea (ICES, www.ices.dk), International Bottom Trawl Survey in the Mediterranean (MEDITS, www.sibm.it); SeaAroundUS (www.seaaroundus.org); Online portal for marine litter (www.litterbase.awi.de); GEOTRACES (ww.geotraces.org). A reference is provided for local- scale dataset that have been published.

ID	Product	WP	SDG	EV	Use	ARC	NWS	IBI	BAL	MED	BS	GLO	Broad categories of observational dataset
1	Mesozooplankton biomass	5	14	x		x-6	x-6	x-6	x-6			x-5	Acoustic; large-scale dataset (e.g., COPEPOD)
2	Micronekton biomass	5, 7	14	x				x-6				x^2-6	Acoustics; scientific trawling data
3	Suspended particulate matter SPM	5, 6		x	x		x-5			x-5	x-5		PRISMA; Copernicus OLCI
4	POM	5		x		x-6	x-6		x-6	x-5			Copernicus OLCI; BGC-ARGO
5	DOM	5		x		x-6	x-6		x-6				Large-scale dataset; local-scale dataset (UK LOCATE project)
6	Reflectance	4, 5		x			x-5		x-5	x-5	x-5		PRISMA; Copernicus OLCI
7	Oxygen near bottom	6, 5		x			x-6		x-6	x-6	x-6		Large-scale dataset (e.g., ICES, IMARNET); BGC-Argo floats
8	pH near bottom	6, 5	13,14				x-6		x-6	x-5	x-5		Large-scale dataset (e.g., ICES, IMARNET); BGC-Argo floats
9	Light at bottom	6, 5					x-6		x-6	x-6	x-6		Large-scale dataset (e.g., BODC); BGC-Argo floats
10	Carbon flux to bottom	6, 5		x			x-5		x-5	x-5	x-5		Large-scale dataset (e.g., BODC); BGC-Argo floats
11	Carbon in sediment	6, 5					x-5		x-5		x-5		Large-scale dataset (e.g., BODC); local-scale dataset
12	Macrozoobenthos	6	14	x	x		x-5			x-5	x-5		Large-scale dataset (e.g., BODC, EMODNET, ICES, MEDITS)
13	Benthic flora	6		x	x					x-5	x-5		Historical surveys
14	Small-pelagic biomass	7	2,14	x	x		x^6 -6	x^3-6	x-6	x^3-6	x-5	x^4-6	Large-scale dataset (e.g., SeaAroundUS, ICES, MEDITS)
15	Apex predator (e.g., Tuna)	7	12,14	x	x		x^3 -6	x^3-5		x^4-6		x^4-6	Large-scale datasets (e.g., RFMOs, SeaAroundUS, ICES, FAO)
16	Marine mammals	7	14	x	x							x-6	Regional to local-scale datasets (Pleslić et al., 2021)
17	Unspecified fish biomass	7	2,14	x	x	x-6	x^6 -6		x-6				Acoustic; large-scale datasets (e.g., MEDITS, SeaAroundUS)
18	HTL Habitat	7	14	x	x			x^3 -6		x-6	x-5		large-scale datasets (e.g., ICES, RFMOs, FAO)
19	Plastic	8	14		x	x-6	x-6			x-6			large-scale datasets (e.g., Litterbase, EMODnet, UNEP/MED)

20	persistent organic pollutants (POP)	8	14		x	x-6				x-5			local-scale dataset (Li et al., 2017; Berrojalbiz et al., 2011) large-scale dataset (e.g., Emodnet Chemistry)
21	Mercury (Hg)	8	14		x					x-6			local-scale dataset (Cossa et al., 2009, Heimbürger et al., 2010) large-scale dataset (Emodnet Chemistry; GEOTRACES, FixO3)
22	Oil	8	14		x					x-5			SAR and optical imagery
23	Fisheries pressure	8	14		x						x-6		Fishing effort and catch (FAO, ICES, SeaAroundUs, RFMOs)
24	Climate change stressor index	8, 5	13,14				x-6			x-6	x-6		datasets for the single variables composing the index: oxygen, temperature, pH, primary production
25	Multi-stressor index	8	14		x		x-6			x-6	x-5		datasets for the single variables composing the index: 19-24

Although the products and datasets delivered by NECCTON are mainly model outputs coherently with the requirements of the call, the project will also process and use **observational products and datasets** (see last column of Table 1). These observational datasets are already generated and curated by other entities and they will be integrated in the models with the aim of improving and assessing the accuracy of our model products and datasets. Only the observational product of **hyperspectral reflectance** and water quality for selected open ocean sites will be developed and assessed in NECCTON as a new product, responding to the requirements of the call. To deliver this product, NECCTON will assess and develop new processing chains for the PRISMA/ASI space observations. The methods, procedures and quality assessment of the dataset will be presented in the public report in D4.2 "Assessment of L2 PRISMA production chain, outputs and transferability".

The other observational datasets used in the project are not new (see last column of Table 1). They have been collected by previous or ongoing international monitoring infrastructure and implementation programmes: e.g., Copernicus Sentinel 3 OLCI (multi-spectral reflectance and ocean colour products will be used for validation and assimilation in NECCTON), biogeochemical-Argo programme (optical, nutrient and carbonate data will be used for validation of a range of pelagic products, as well as for near-bottom products). These datasets will be acquired from public, on-line data sets (e.g., EMODNET, BODC, and ICES) or requested from the EU Data Collection Framework (e.g. fish biomass surveys of EU member states).

The largest part of the products in Table 1 are new, meaning they are not included in the current catalogue of the Copernicus Marine Service (see 3rd column in Table A1 of the Appendix). This is coherent with the objective of NECCTON to provide CMEMS with innovative capabilities to deliver new products for fisheries, pollution and benthic habitat to an enlarged group of their stakeholders. Some products are already available in the catalogue of some of the CMEMS MFCs (e.g., mesozooplankton, oxygen and pH). However, NECCTON aims at improving the quality of the existing products by developing improved or new simulation models for those same products.

The models used or developed in NECCTON to run the simulations producing the model output datasets are specified in the fourth column of Table A1 of the Appendix. For the pelagic products (blue shadow) they include all the current operational biogeochemical models of the CMEMS MFCs, which are all process-based models. These might be advanced with revised or new modules developed in NECCTON, or complemented by new models (e.g., machine learning models for mesozooplankton). The benthic, higher-trophic-level (HTL) and stressor products are delivered by well-established models adapted by NECCTON to the CMEMS operational framework. The thorough description of the models will be provided in dedicated deliverable of the project.

All the **products** will be delivered using the NetCDF Climate and Forecast (CF) Metadata Conventions. They will all be simulation outputs available on the grid of the model (i.e., L4 products), with the exception of the reflectance observational product, which will be delivered as datasets of geophysical variables at the same resolution and location of the original observations (i.e., as L2 georeferenced product). The procedures to provide the metadata of the products (e.g., names, units) will follow the ones adopted by the Copernicus Marine Service (CMEMS AD9, AD11 and AD19 in the Reference section of this document). This is coherent with the objective specified in the call and the aim of integrating the NECCTON developments in the operational Copernicus Marine Service in the future. The roadmap for such integration is the objective of the Key Deliverable D2.3 of NECCTON.

If a need was identified for a standardized data or metadata format for a product for which no format has been established in the literature, NECCTON will organize dedicated meetings to explore formats in

consultation with the advisory board members and external experts (and tacking account of approaches and standards used for similar data and projects (see Section 2.3).

Key specifications of the model output **datasets** are listed in Table A1 of the Appendix. This includes the expected data volume, marine area covered, spatial resolution and time window of the data.

Product stakeholders were engaged since the preparation of the project proposal to guarantee the utility of the data delivered by NECCTON. In particular, the preliminary product list and their test in thirteen case studies was discussed with stakeholders in the biodiversity conservation, fish management and marine policy sectors, listed in Table 2.

All the work-packages from WP5 to WP9, in collaboration with WP1 and WP2, will continue engaging key stakeholders and experts to refine the product list, define technical specification and identify datasets for model development, validation and dissemination of the products delivered by NECCTON, in consultation with the Stakeholder Committee and the Advisory Board of the project. This is the activity of the initial Tasks 1 “Definition of the new products and services" of each WP, which will all deliver a public report “Technical specifications" at month 10 of the project (October 2023).

Engagement and discussion of the needs and recommendations of stakeholders of the new data products delivered by NECCTON will be collected in dedicated stakeholder workshops (the first to be hold on-line on the 28-29 June 2023), questionaries and surveys planned during the development of the project (see Section 2).

The plan for the continuous engagement of the current and new stakeholders of the NECCTON products, as well as for the dissemination, communication of the NECCTON products, are detailed in the project Deliverable 2.1.

The final recommendations to the stakeholders for the useful and appropriate exploitation of the products will be provided in Deliverable 9.3 “Recommendations on the use of the NECCTON products for fisheries management, biodiversity monitoring and conservation based on the outcomes of the 13 case studies”.

Regarding the **accessibility of the final products**, all the model-based products delivered by NECCTON will be freely accessible by internal and external project users. Differently, the access to PRISMA mission data is currently subject to a License to Use signed by some NECCTON partners with the Italian Space Agency, which required a registration procedure through the PRISMA portal. This License defines the terms and conditions for the access and usage of PRISMA original and derived products and the restrictions to which their provision and use are subject to. In details, PRISMA original products (Level 1, 2B, 2C, 2D PRISMA System Product in HDF5 format) shall not be disclosed, reproduced, transferred or made in any way available to third parties. The use of PRISMA derived products (Products generated by the users from an ORIGINAL Product after applying different processing and/or addition on methods and techniques, i.e., water reflectance derived from the atmospheric correction of L1 PRISMA data) is allowed, only for non-commercial purposes, and shall be freely disclosed, reproduced, transferred or made in any way available to third parties.

Table 2 *List of the case studies exploiting the NECCTON tools and products and co-developed with stakeholders in the fisheries management and biodiversity conservation sectors. The Monitoring Forecasting Centres (MFC) column lists the seven CMEMS domains where the case studies are developed.*

ID	MFC	Main focus	Title	Co-design stakeholder
C1	MED	Biodiversity conservation	Monitoring a Natura 2000 marine protected area in the Adriatic Sea	Blue World Institute of Marine Research and Conservation
C2	MED	Fishery management	Ensemble of species distribution models to determine hot spots of aggregation supporting management plans and MPA definition in the Mediterranean Sea	FAO-GFCM
C3	MED	Fishery management	Mapping sources of pollutants transported towards Mediterranean aquaculture farm areas	FAO
C4	BS	Biodiversity conservation	Mapping the impact of fish trawling on functional biodiversity of the Black Sea	ICES FBIT
C5	NWS	Biodiversity conservation	Climate-smart MPA planning in the North West European Shelf	UNEP
C6	NWS	Fishery management	Ensemble model evaluation and projection of small pelagics in the Bay of Biscay	ICES WGEAWESS
C7	IBI	Biodiversity conservation	Monitoring plankton diversity and dynamics in the Parc Naturel Marin d'Iroise (PNMI)	Office Français de la Biodiversité (OFB)
C8	ARC	Biodiversity conservation	Monitoring pollution of marine protected areas surrounding Svalbard	IUCN
C9	ARC	Fishery management	Modelling and projecting fisheries potential in the Arctic Atlantic Ocean	ICES-WGINOR and ICES-WGIBAR
C10	BAL	Biodiversity conservation	Marine protected area monitoring and assessment in the Baltic Sea	HELCOM
C11	GLO	Fishery management	Impact of High Seas MPA on tuna stocks and fisheries: the example of the Phoenix Islands Protected Area (PIPA)	Pacific community (SPC), ISSF, IC, FAO

C12	GLO	Biodiversity conservation	Monitoring marine mammals in the Azores region	Azores Regional Directorate of Maritime Affairs
C13	GLO	Fishery management	Forecasting climate change impact on potential catches of open ocean large pelagic fish	Fish-MIP

2 FAIR data

The production and dissemination of the NECCTON model data products, code and software will follow the Findable, Accessible, Interoperable and Reusable (FAIR) principle. The products will be delivered as model output simulation datasets. These are generated by model simulations that cover specific ocean regions and time periods (see Table A1 in Appendix). Crucially we expect that, after the end of the project, the NECCTON products will continue to be produced routinely, FAIR and freely available in the CMEMS data-portal, once the Service will have taken-up the project developments. NECCTON is fostering such take-up through delivering the "roadmap to uptake of NECCTON by CMEMS", i.e., the Key Deliverable D2.3.

In order to make the output FAIR as well as transferable to our main stakeholder CMEMS, we are adopting the same standards, formats, quality and assessment methods, access and dissemination channels, and criteria for use defined by CMEMS for their operational, public products. Some key elements of these CMEMS procedures are reported in the following sub-sections. We refer the reader to the CMEMS AD9, AD11 and AD19 reports for a thorough descriptions of the procedures (see the reference Section 8).

2.1 Making data findable, including provisions for metadata

Following the CMEMS procedures and the Climate and Forecast (CF) metadata convention, the NECCTON data products will be made findable by a unique reference of the product: the product nomenclature rules are specified in the document CMEMS AD9 “CMEMS Product Nomenclature Rules”. Once created, this product reference is persistent over time. As mentioned above, If a need was identified for a standardized data or metadata format for a product for which no format has been established in the literature, NECCTON will organize dedicated meetings to explore formats in consultation with the advisory board members and external experts (e.g., Frank Muller-Karger, AB member and co-chair of the Marine Biodiversity Observation Network (MBON) and member of the GOOS Bio-Eco Panel, Ocean Best Practices System Steering, IMBeR Science Steering Committee) and tacking account of approaches and standards used for similar data (e.g. Core Darwin Standard and Ecological Metadata Language) and projects (e.g., FishMIP, ACTNOW, Marco-Bolo, and EcoScope).

The metadata information provided by the NECCTON partners for the new or updated products will be the same requested for the integration of products in the Copernicus Marine catalogue, i.e.:

- Product’s characteristics: product unique reference, category, title, description, DOI, geographical coverage (longitude and latitude in deg), source (satellite, in situ, model), type (mutli-year, near-real-time, forecast), data assimilation, processing level (L3, L4 etc), documentation references;
- Dataset’s characteristics: dataset unique reference, list of variables, horizontal resolution longitudinal and latitudinal (km or deg), vertical coverage (number of layers), temporal coverage (DD/MM/YYYY), temporal resolution (daily/weekly/monthly/yearly + mean/instantaneous), server addresses, feature type (grid, profile, etc), coordinate reference system (equirectangular, polar etc), update frequency (daily, weekly etc), delivery time (e.g. twice per day at 0:00 and 12:00), WMS title, format (netCDF-3 or 4), PQD metric type
- Product’s contact points: for each role: entity, responsible and email address

- Product's upstream dependencies: upstream product reference, usage and provider

NECCTON will follow the guideline for metadata requirements detailed in CMEMS [AD19] "CMEMS Product Information Guideline". The CF convention will be used, notably for the variable's standard_names.

Title, description, geographical and temporal coverage, list of variables will be set as keywords to make the data findable by users that are interested in specific application of the data.

The metadata are collected in a table that will be shared among data producers and managers. This file is the reference used to upload metadata on the catalogue and will be updated as needed. Each modification of this file is tracked and versioned.

2.2 Making data accessible.

The NECCTON data products will be made freely accessible, visualizable and downloadable from the NECCTON datacube viewer (project Deliverable 2.2), where NECCTON will release data-files as soon as the outputs of the model simulations run in WP5-9 have been validated and released as Milestones (see Table 3).

Crucially, the NECCTON datacube is compatible with the current MyOcean Viewer used by the Copernicus Marine Service, which has about 50000 users. The use of the datacube will facilitate the expected transfer of the NECCTON model and products to CMEMS for their dissemination after the completion of the project.

Table 3: Model data products that will be made accessible to project partners and stakeholders during the months of development of the project.

Milestone Name	Means of dissemination	Month
MS6.4 Validated hindcast of all the six pelagic products and seven benthic products shared in WP9.1 and KPIs	Files in the datacube; skill and Kee Performance Indicators (KPI) metrics tables in repository	35
MS8.3 Validated hindcasts of 7 pollution and stressor products and hindcast and ensemble simulations of 5 HTL products shared in WP9.1 and KPIs	Files in the datacube; skill and KPI metrics tables in repository	36
MS9.1 The global and regional LTL climate projections are shared	Files in datacube	36
MS9.2 The global and regional HTL climate projections are shared	Files in datacube	43

The datacube is based on innovative cloud-based technologies and will use a serverless architecture that allows to connect directly to files and not to a server. This viewer will guarantee high-availability, visual analysis, flexible data dissemination, and will enable the user to explore the data across all dimensions: longitude, latitude, depth, time. This datacube exploratory viewer will promote free and open data dissemination and allow the sub-setting and analysis of diverse datasets independently of existing HTTP interfaces such as the Web Map Service (WMS) protocol.

The datacube developed by LOBELIA will act as trusted repository of the NECCTON data products, with capacity to resolve the identifier of the digital objects adopted in NECCTON. Native NetCDF datasets as well as any other data format provided will be stored in object storage in a public cloud repository. Currently, two European cloud providers have been contacted and have committed to competitive unit prices. Furthermore, since the system will remain agnostic in terms of the cloud providers, the most appropriate arrangement will be put in place in due time considering the sustainability of the service. LOBELIA, who committed to the arrangements described in the grant agreement of NECCTON as project partner, has also developed and is managing the MyOcean Viewer currently used by CMEMS.

Data in the datacube will be made openly available to the public. Users will be invited to register into a user database which will follow GDPR criteria. Data visualization and downloading will be accessible through free and standardized access protocols consistent with the procedures adopted in the CMEMS.

Metadata will be made openly available and licensed under a public domain (Creative Commons) agreed by the partners.

The objective of the project is to make data remain available to the users. First, through a dedicated domain and ideally through the CMEMS catalogue afterwards.

The datasets will be viewable through the NECCTON viewer based on Background Intellectual Property Rights and accessible directly in object storage, in principle without the need of specific access software. All software developed under this agreement will be included as source code and the documentation for its use will be released along with the public Deliverable D2.2." Tools of dissemination and visualization of the NECCTON products".

Different conditions hold for the accessibility of the original PRISMA data and the PRISMA-derived products delivered by NECCTON. Such conditions are stated in the License to Use of the PRISMA data, as mentioned in Section 1. This implies that the availability of original PRISMA products is allowed only for internal use, but the provision of Reflectance data, generated by NECCTON, are freely available for testing experimental/demonstrative activities and assimilation methods for the scope of the project.

2.3 Making data interoperable

Following the CMEMS procedures (CMEMS AD9, AD11, AD19), datafiles will be made interoperable by producing data and metadata following consolidate standards and guidelines: network Common Data Form (NetCDF) gridded data files with Climate and Forecast (CF) metadata conventions, following the best practices of the Open Geospatial Consortium (OGC). As mentioned above, if a need was identified for a standardized data or metadata format for a product for which no format has been established in the literature, NECCTON will organize dedicated meetings to explore formats in consultation with the advisory board members and external experts and tacking account of approaches and standards used for similar data and projects.

2.4 Increase data re-use

To boost the re-use of the NECCTON model data products by stakeholders, the validated datasets will be made freely available to the public through the datacube, with open license CC BY (see Section 2.2 and Table 3).

All the information on the 25 NECCTON products, as well as on the models and tools for their simulation, will be reported in the project report deliverables, which will be all public, disseminated in preprint servers (e.g., Zenodo, with DOI) with links highlighted in the datacube sheets and in the project website.

After the completion of the project, we expect that CMEMS will take-up the NECCTON model developments in their future operational systems. This would ensure that the NECCTON product will continue to be produced operationally and maintained easily and freely accessible by users via the CMEMS on-line catalogue and data viewer.

In addition, we will engage stakeholders and potential end-users (e.g., MPA managers) and civil society organizations (e.g., Conservation International) in co-design the products and thirteen case studies. To maximize the likelihood of external users' exploitation of the NECCTON products, we have planned six events dedicated to stakeholder engagement and training (see Table 4), the first of which took place on-line on the 28-29 June 2023, with more than 100 attendees.

The models and procedures used or developed in NECCTON for the production and validation of the simulation products will be published journals that are open-access or support self-archiving (also called green open-access), including open-peer-review journals, to guarantee their highest standard and largest dissemination.

Table 4 Specific events planned to align NECCTON to the stakeholder needs and foster the use of the NECCTON products.

NECCTON Events	**Month**
Workshop Stakeholders 1 – Needs and co-design of future products	M6
Survey CMEMS users – Refine training content	M32
Training-workshop 1 – Using NECCTON products (biodiversity conservation)	M38
Training-workshop 2 – Using NECCTON products (fisheries management)	M43
Survey NECCTON Advisory Board and of independent stakeholders and training attendees	M44
Final Workshop Stakeholders 2 – Demonstration: Exploration and validation of new products	M48

3 Other research outputs

NECCTON research outputs will include also software and model code (see the project Deliverables and Milestones in Table 5 and Table 6, respectively), besides the model products described in the above Sections 1 and 2.

NECCTON will deliver innovative algorithms based on Machine Learning (ML) and Data Assimilation (DA). These algorithms will be distributed within and outside the project using dedicated open repositories (e.g., GitHub). Each algorithm will be associated with a Jupyter notebook freely available to be run by users. This will ensure an efficient distribution of the algorithms and will deliver a toolbox of algorithms to be used by a large community.

Standards, procedures and best practices followed and recommended for the model and algorithm development and application are the specific object of two tasks (WP3.1 and WP4.1) and two public reports:

- D3.1 Best practices for implementing new biogeochemical process modules in CMEMS/FABM
- D4.1 Github organisation and guide of good practice

NECCTON will pursue the recommended practices of open science and FAIR principles also for these other research outputs, by granting open access to the code software, models, algorithms, workflows, protocols, simulations to users. Model components will be built on the open-source, publicly available framework FABM (Framework for Aquatic Biogeochemical Modelling: https://github.com/fabm-model/fabm) in the work-package (WP) 3 of NECCTON. Both the CMEMS Lower-Trophic-Level models ported in FABM (WP3) and newly developed building blocks from other WPs 5-8 (e.g., spectral light, Suspended Particulate Matter, zoobenthos, plastic) will be made available in public code repositories (e.g., GitHub) with documentation allowing reproducible results, enabling their use in CMEMS modelling systems and beyond. The documented code will be released in repositories in 8 Milestones and in the Key Deliverable 3.3 of the inter-linked modelling framework (see Tables 5 and 6).

To promote the use of the models and algorithms by stakeholders, the code stored in open repositories will be documented with clear installation instructions (e.g., GitHub). The NECCTON software will use open-source licences used for the background intellectual property rights.

The final recommendations to CMEMS for the useful and appropriate exploitation of the models will be provided in: Deliverable 9.1 “Synthesis of the performance of the integrated modelling systems and recommendations on all 25 products for use by CMEMS”.

Table 5: Model developments and assessments that will be released as public deliverables to project partners and stakeholders during the months of development of the project.

Deliverable	Title	Type	Dissemination
D3.2	Offline simulation tool	Code	Public
D3.3	Inter-linked modelling framework	Code	Public
D4.3	Software repository of ML and DA algorithms of T4.2.1-T4.4.3	Code	Public
D4.4	Report on ML and DA developments for data-model integration in T4.2.1-T4.4.3	Report	Public
D5.2	Report on all the integrated pelagic model developments in T5.2.1-4	Report	Public
D6.2	Report on the integrated model developments for SPM and habitat suitability from T6.2.1-T6.2.3	Report	Public
D6.3	Report on the model developments for benthic processes in T6.2.4	Report	Public
D7.2	Report on all 16 model coupling developments in T7.2	Report	Public
D8.2	Report on integrated modelling of pollutants and single pressures in T8.2-T8.3	Report	Public
D8.3	Report on developments of tools for mapping integrated pressures, exposure, and risk in T8.4	Report	Public

Table 6: Model code that will be released as milestones to project partners and stakeholders during the months of development of the project.

Milestone Name	**Means of dissemination**	**Month**
MS3.1 First version of offline simulator for 2D simulation (benthic fauna) with CMEMS output	Documented code in Github	13
MS3.2 New CMEMS biogeochemical models implemented in the unified framework	Documented code in Github	24
MS6.2 Release of integrated model code for SPM and habitat suitability from T6.2.1- T6.2.3	Documented Code in Github	30
MS6.3 Release of model code for benthic processes from T6.2.4	Documented Code in Github	31
MS7.1 Coupling of all the 16 HTL models in T7.2 completed	Code in repositories	29
MS8.1 Release of model code for all pelagic processes pollutants and single pressures	Documented Code in Github	24
MS8.2 Release of tools for mapping integrated pressures, exposure, and risk from T8.4	Documented Code in Github	29

4 Allocation of resources

LOBELIA will be responsible for the data management in NECCTON from the time the output datasets are released.

Adequate budget for making the models and model output product FAIR have been allocated to all the models and data producers, as well as to the developer and responsible of the datacube and software for data storage, dissemination and visualization. These costs were specified in the Grant Agreement Annex 1 Description of the Action (Part B), Table 3.1h for LOBELIA, and in Annex 2 Estimated budget for the action for all the other partners.

Funding for 35 open-access publications has been allocated to boost the dissemination of the products and impact of the project.

Funding for training user of the NECCTON products have been allocated for two training workshops and four co-design events (see Table 4).

5 Data security

NECCTON datasets will be safely stored in trusted repositories for long term preservation and curation (see Section 2.2). They will not include sensitive data. If user information needs to be registered, best practices in terms of data security at rest and data transmission will be enforced. Data recovery will be available from a secure object storage with a backup copy.

6 Ethics

No ethical or legal issues are foreseen in relation to the sharing of the NECCTON products planned in this document, since it follows the principles and practices agreed by the partners in both the Grant and Consortium Agreements.

Ethical aspects are mainly relevant for data of a personal nature. These data will be treated according to the ethics procedures laid out in the Ethics section of the Grant Agreement (Section 4).

When consent is collected through questionnaires or any form of collection of personal data (e.g., interviews, surveys), the data generated will not be re-used directly and the processes will be treated / anonymized / secured to be compliant with the General Data Protection Regulation (GDPR) No 2016/679 obligations.

7 Other issues

This Data Management Plan embraces the FAIR procedures and best practises of the Copernicus Marine Service, coherently with the call requirements and project objective. No other procedures for the management of the NECCTON products will be used if not planned, or at least coherent with the Open Research and FAIR principles of this Data Management Plan.

8 References

CMEMS AD9, CMEMS Product Nomenclature Rules", Mercator Ocean International

CMEMS AD11, Spreadsheet of variables definitions, V1.0 November 2022, Mercator Ocean International

CMEMS AD19, Product information guideline, V2.2 April 2023, Mercator Ocean International

Appendix

Table A1 Specification of the model products and simulation output datasets delivered by NECCTON. List of the 25 products of NECCTON delivered as 84 simulations datasets produced by different models for different regions and periods (L4 mapped) and as 2 derived observational datasets (namely the reflectance L2 georeferenced products 6.1. and 6.2). The table uses the same colours used in Table 1 of this document for the products.

ID	Product	Partner	New?	Method	Min size (Gb)	Max size (Gb)	Spatial resolution	Area	Period
1.1	Mesozooplankton biomass	NERSC	Yes	ECOSMO	10	200	12 km	Arctic	1991-2020
1.2	Mesozooplankton biomass	UKMO	Yes	ERSEM	30	260	7 km	North West Shelf	1991-2020
1.3	Mesozooplankton biomass	BSH	no	ERGOM	45	225	2 km	Baltic	2-5 years
2.1	Micronekton biomass GLO	CLS	Yes	SEAPODYM	2	15	1/4°	Global	1998-2019
2.2	Micronekton IBI	CLS	Yes	SEAPODYM	50	200	1/36°	IBI	1998-2019
3.1	Suspended particulate matter SPM	UKMO	Yes	ERSEM (+SPM)	30	1000	7 km	North West Shelf	1991-2020
3.2	Suspended particulate matter SPM	UoL	Yes	BAMHBI	13	13	2,5km and 59 levels	Black Sea.	2 years sample
4.1	Particulate Organic Matter (POM)	NERSC	Yes	ECOSMO	10	600	12 km	Arctic	1991-2020
4.2	POM	OGS	yes	BFM	200	600	4.5km	Mediterranean Sea	1999-2020
4.3	POM	UKMO	Yes	ERSEM	30	1000	7 km	North West Shelf	1991-2020
4.4	POM	BSH	Yes	ERGOM	45	225	2 km	Baltic	2-5 years
5.1	Dissolved Organic Matter (DOM)	NERSC	Yes	ECOSMO	30	600	12 km	Arctic	1991-2020
5.2	DOM	UKMO	Yes	ERSEM	30	1000	7 km	North West Shelf	1991-2020
5.3	DOM	BSH	Yes	ERGOM	45	225	2 km	Baltic	2-5 years
6.1	Reflectance (satellite derived product)	CNR	Yes	ASI standard and ACOLITE	300	450	30 m	North West Shelf	2019-2023
6.2	Reflectance (satellite derived product)	CNR	Yes	ASI standard and ACOLITE	300	450	30 m	Black Sea	2019-2023

6.3	Reflectance (satellite derived product)	CNR	Yes	ASI standard and ACOLITE	300	450	30 m	Mediterranean Sea	2019-2023
6.4	Reflectance	OGS	yes	BFM	10	30	4.5km	Mediterranean Sea	1999-2000
6.5	Reflectance	BSH	yes	ERGOM	2	4	2km	Baltic	2-5 years
6.6	Reflectance	UoL	yes	BAM + RADTRANS	2	4	2km	Black Sea	2019-2023
ID	**Product**	**Partner**	**New?**	**Method**	**Min size (Gb)**	**Max size (Gb)**	**Spatial resolution**	**Area**	**Period**
7.1	Oxygen near bottom	OGS	no	BFM	10	10	4.5km	Mediterranean Sea	1999-2020
7.2	Oxygen near bottom	UoL	No	BAMHBI	0.2	0.2	2,5km	Black Sea.	2 years of sampled products validated with in-situ data
7.3	Oxygen near bottom	UKMO	No	ERSEM	2	15	7 km	North West Shelf	1991-2020
7.4	Oxygen near bottom	BSH	No	ERGOM	2	4	2km	Baltic	2-5 years
8.1	pH near bottom	OGS	no	BFM	10	10	4.5km	Mediterranean Sea	1999-2020
8.2	pH near bottom	UoL	Yes	BAMHBI	0.2	0.2	2,5km	Black Sea.	2 years of sampled products validated with in-situ data
8.3	pH near bottom	UKMO	No	ERSEM	2	15	7 km	North West Shelf	1991-2020
8.4	pH near bottom	BSH	No	ERGOM	2	4	2km	Baltic	2-5 years
9.1	Light at bottom	OGS	yes	BFM	10	30	4.5km	Mediterranean Sea	1999-2020
9.2	Light at bottom	UoL	Yes	BAMHBI +RADTRANS	0.2	0.2	2,5km	Black Sea.	2 years of sampled products validated with in-situ data
9.3	Light at bottom	UKMO	Yes	ERSEM	2	15	7 km	North West Shelf	1991-2020
9.4	Light at bottom	BSH	Yes	ERGOM	2	4	2km	Baltic	2-5 yaers

10.1	Carbon flux to bottom	OGS	Yes	BFM	10	10	4.5km	Mediterranean Sea	1999-2020
10.2	Carbon flux to bottom	UoL	Yes	BAMHBI	0.2	0.2	2,5km	Black Sea.	2 years of sampled products validated with in-situ data
10.3	Carbon flux to bottom	UKMO	Yes	ERSEM	2	100	7 km	North West Shelf	1991-2020
10.4	Carbon flux to bottom	BSH	Yes	ERGOM	2	4	2km	Baltic	2-5 years
ID	**Product**	**Partner**	**New?**	**Method**	**Min size (Gb)**	**Max size (Gb)**	**Spatial resolution**	**Area**	**Period**
11.2	Carbon in sediment	BS(UoL)	Yes	BAMHBI	0.0025	0.0025	few kms.	Black Sea.	Climatological averaged (e.g. seasonal averaged)
11.3	Carbon in sediment	UKMO	Yes	ERSEM	2	50	7 km	North West Shelf	1991-2020
11.4	Carbon in sediment	BSH	Yes	ERGOM	2	4	2km	Baltic	2-5 years
12.1	Macrozoobenthos	MED(OGS),	YEs	HSM	0.5	0.5	4.5km	Mediterranean Sea	Climatological averaged, 2000-2021
12.2	Macrozoobenthos	UKMO, PML	Yes	ERSEM	2	50	7 km	North West Shelf	1991-2020
12.3	Macrozoobenthos	NIOZ, UoL	Yes	DEB model	0.01	0.01	2,5km	Black Sea.	2 years sample, averaged
13.1	Benthic flora	MED (OGS)	Yes	HSM	0.5	0.5	4.5km	Mediterranean Sea	Climatological averaged, 2000-2021

13.2	Benthic flora	BS(NIOZ, UoL)	Yes	Macroalgae model	0.01	0.01	2,5km	Black Sea.	Climatological averaged (e.g. seasonal averaged)
14.1	Small-pelagic biomass	HCMR	Yes	ERSEM-SPF	1	10	0.1 degree grid	Mediterranean Sea	2000-2020 monthly
14.2	Small-pelagic biomass	Ifremer	yes	OSMOSE	1	1	10' lat x 15' lon	Bay of Biscay	2000-2020
14.3	Small-pelagic biomass	Ifremer	yes	DEB-IBM	20	50	2,5km	Bay of Biscay	2000-2020
14.4	Small-pelagic biomass	PML-DBEM	yes	SS-DBEM	6	50	0.5 degree grid	NWS	2000-2020 annual
14.5	Small-pelagic biomass	PML-MIZER	yes	ERSEM-MIZER	6	50	7km	NWS	1991-2020
14.6	Small-pelagic biomass	DTU	Yes	FEISTY	1	1	2 km	NWS	1990-2020
14.7	Small-pelagic biomass	DTU	Yes	FEISTY	1	1	1/12 deg	IBI	1992-2022
14.8	Small-pelagic biomass	DTU	Yes	FEISTY	1	1	1 nm	BAL	2000-2020
14.9	Small-pelagic biomass	DTU	Yes	FEISTY	1	1	1/24 deg	MED	2011-2020
14.1	Small-pelagic biomass	DTU	Yes	FEISTY	1	1	1/12 deg	Global	1992-2022
14.11	Small-pelagic biomass (Mackerel)	CLS	yes	SEAPODYM	6	50	1/4°	North Atlantic	1998-2020
14.12	Small-pelagic biomass	OGS	yes	E-SDM	2	2	4.5 km	Mediterranean Sea	2000-2020
ID	**Product**	**Partner**	**New?**	**Method**	**Min size (Gb)**	**Max size (Gb)**	**Spatial resolution**	**Area**	**Period**
14.13	Small-pelagic biomass	MOi	Yes	SEAPODYM	1	5	1/4; 1/12; 1/36?	Bay of Biscay	2000-2020
14.14	Small-pelagic biomass	UoL	yes	DEB/FESTY	2	2	2,5km	Black Sea.	2 years sample
14.15	Small-pelagic biomass	EII	Yes	EcoOcean	1	1	1 degree grid	Global	1950-2020
15.1	Apex predator (e.g., Tuna)	PML-DBEM	yes	SS-DBEM	6	50	0.5 degree	Global	2000-2020 annual

							grid		
15.2	Apex predator (e.g., Tuna)	DTU	Yes	FEISTY	1	1	2 km	NWS	1990-2020
15.3	Apex predator (e.g., Tuna)	DTU	Yes	FEISTY	1	1	1/12 deg	IBI	1992-2022
15.4	Apex predator (e.g., Tuna)	DTU	Yes	FEISTY	1	1	1 nm	BAL	2000-2020
15.5	Apex predator (e.g., Tuna)	DTU	Yes	FEISTY	1	1	1/24 deg	MED	2011-2020
15.6	Apex predator (e.g., Tuna)	DTU	Yes	FEISTY	1	1	1/12 deg	Global	1992-2022
15.7	Apex predator (e.g., Tuna)	MOi	yes	SEAPODYM	5	10	1 degree grid	Global	1979-2050
15.9	Apex predator (e.g., Tuna)	EII	Yes	EcoOcean	1	1	1 degree grid	Global	1950-2020
16.1	Marine mammals	IMAR	Yes	GAM	1	1	1/4°	Azores	2010-2018
17.2	Unspecified fish biomass	Ifremer	Yes	OSMOSE	1	1	10' lat x 15' lon	Bay of Biscay	2000-2020
17.3	Unspecified fish biomass	PML-DBEM	yes	SS-DBEM	6	50	0.5 degree grid	NWS	2000-2020 annual
17.4	Unspecified fish biomass	PML -MIZER	Yes	ERSEM-MIZER	6	50	7km	NWS	1991-2020
17.5	Unspecified fish biomass	Hereon	Yes	E2E	10	200	12 km	ARC	1991-2020
17.6	Unspecified fish biomass	Hereon	Yes	E2E	10	200	2km BAL	BAL	2-5 years
17.7	Unspecified fish biomass	OGS	Yes	E-SDM	2	2	4.5 km	Mediterranean Sea	2000-2020
17.8	Unspecified (demersal) fish biomass	DTU	Yes	FEISTY	1	1	2 km	NWS	1990-2020
17.9	Unspecified (demersal) fish biomass	DTU	Yes	FEISTY	1	1	1/12 deg	IBI	1992-2022
17.10	Unspecified (demersal) fish biomass	DTU	Yes	FEISTY	1	1	1 nm	BAL	2000-2020

ID	Product	Partner	New?	Method	Min size (Gb)	Max size (Gb)	Spatial resolution	Area	Period
17.11	Unspecified (demersal) fish biomass	DTU	Yes	FEISTY	1	1	1/24 deg	MED	2011-2020
17.12	Unspecified (demersal) fish biomass	DTU	Yes	FEISTY	1	1	1/12 deg	Global	1992-2022
18.1	HTL Habitat	OGS	Yes	Ecospace	2	2	4.5 km	Adriatic Sea	2005-2020
19	Plastic	MINDS	Yes	HCMR-plastic model	2	20	0.1 x 0.1 degree (~10Km)	Mediterranean Sea	2010-2020 seasonally averaged and monthly climatology
20	persistent organic pollutants (POP)	Hereon	Yes	CanMETOP	10	200	1 x 1 degree	Arctic	2010-2020
21	Mercury (Hg)	OGS	Yes	BFM-Hg	20	340	1/16° and 70 vertical levels	Mediterranean Sea	2005-2021
22	Oil	CMCC	Yes	MEDSLIK-II	0.2	2.5	CMCC: GEBCO 30''	Mediterranean Sea	2020-2021
23	Fisheries pressure	NIOZ	Yes	combined stressor index	1	10	depends on the data availability - target is 2,5 km	Black Sea.	2 year averages
24	Climate change stressor index	PML	Yes	multiple indexes	50	100	~7km horizontal	NW European shelf	1990-2099
25	Multi-stressor index	PML	Yes	combined stressor index	5	5	~7km horizontal	NW European shelf	1990-2099