

FORCOAST



Earth Observation Services For Wild Fisheries, Oystergrounds
Restoration And Bivalve Mariculture Along European Coasts

PROJECT DELIVERABLE REPORT

Deliverable Number: 3.2

Deliverable Title: DWH use for 2020

Author(s): Daniel Garaboa (MI_ES), Daniel
Lowe (MI_ES)

Work Package Number: 3

Work Package Title: Service Design



FORCOAST Project Information	
Project full title	Earth Observation Services For Wild Fisheries, Oystergrounds Restoration And Bivalve Mariculture Along European Coasts
Project acronym	FORCOAST
Grant agreement number	870465
Project coordinator	Ghada El Serafy, Deltares
Project start date and duration	1 st November 2019, 30 months
Project website	https://forcoast.eu/

Deliverable Information	
Work package number	3
Work package title	Service Design
Deliverable number	3.2
Deliverable title	Data Warehouse (DWH) use for 2020
Description	This deliverable states whether ADDITIONAL datasets were used from the Copernicus Space Component Data Access (CSCDA) services Data Warehouse and how they were used. These ADDITIONAL Datasets provide for example higher resolution products which are necessary for aquaculture operation (task 3.1).
Lead beneficiary	MI_ES
Lead Author(s)	Daniel Garaboa, Daniel Lowe
Contributor(s)	Carlos Romay, Luis Rodriguez Galvez, Arthur Capet, Ghada El Serafy
Revision number	4
Revision Date	29/04/2021
Status (Final (F), Draft (D), Revised Draft (RV))	F
Dissemination level (Public (PU), Restricted to other program participants (PP), Restricted to a group specified by the consortium (RE), Confidential for consortium members only (CO))	CO

Document History

Revision	Date	Modification	Author
1	19/04/2021	1 st Draft	Daniel Garaboa, Daniel Lowe
2	23/04/2021	1 st Revision	Luis Rodriguez Galvez, Ghada El Serafy
3	27/04/2021	2 nd Revision	Arthur Capet
4	29/04/2021	Final Revision	Luis Rodriguez Galvez

Approvals				
	Name	Organisation	Date	Signature (initials)
Coordinator	Ghada El Serafy	Deltares	29/04/2021	GES
WP Leaders	Arthur Capet	Liège University	27/04/2021	AC

PROPRIETARY RIGHTS STATEMENT

This document contains information, which is proprietary to the FORCOAST consortium. Neither this document, or the information contained within may be duplicated, used or communicated except with the prior written permission of the FORCOAST coordinator.



Executive Summary

This document is the result of the analysis of the usage to date within the framework of the FORCOAST project of the ADDITIONAL datasets from the Data Warehouse that are part of the Copernicus Space Component Data Access (CSCDA) services.

In the first section, the purpose of this deliverable is described, introducing the available datasets in the CSCDA and their main technical characteristics based on the *Copernicus Space Component Data Access Portfolio: Data Warehouse 2014 – 2022* document [1]. Also, a brief summary about the data availability and their access is introduced.

In the second section, we detail the survey set as methodology to extract the information on the usage of the different datasets by the partners in their respective partners. The results of the survey provided by the partners are presented in Section 3.

Finally, the results presented in this deliverable show a lack of usage of the ADDITIONAL datasets by the partners mainly due to a lack of accessibility-ease of the data warehouse and the lack of expertise of raw data processing to L3 products at pilot level, the consortium as a whole is setting up dedicated additional tasks to tackle this. The results and bottlenecks including recommendations would be described in deliverable D3.4 – DWH use for 2021.

Table of Contents

Executive Summary	5
Table of Contents	6
1 Introduction	8
2. The Copernicus Space Component	8
2.1 The CSCDA Access infrastructure	9
2.1.1 Data Warehouse: The Core Datasets	9
2.1.2 Data Warehouse: The ADDITIONAL Datasets	9
2.1.3 Available Datasets and access request	10
3 Survey: ADDITIONAL Dataset usage in 2020	12
3.1 Survey introduction	12
3.2 Results of the survey	13
3.2.1 Excel survey	13
3.2.2 ADDITIONAL datasets use	16
4 Conclusions and proposed actions	16
4 References	17

Figures and Tables

Figure 1: Work flow to request standard ADDITIONAL datasets.	12
Figure 2: Datasets grouped by their representative names and usage.....	13
Figure 3: Level data processing and percentage of use.....	14
Figure 4: Type of dataset use and percentage over total number of datasets.....	15
Figure 5: Type of use of datasets grouped by their processing level.	16
Table 1: Copernicus Space Segment mission classification by sensor type and resolution.	9
Table 2: Copernicus Space Segment mission classification for ADDITIONAL datasets based on service type.....	10
Table 3: Available ADDITIONAL datasets.....	10
Table 4: Main authorities and their accessibility to the ADDITIONAL datasets (D=Download, V=View).	11
Table 5: Quotas assigned for ADDITIONAL datasets to the FORCOAST project.	11
Table 6: Excel sheet shared with the partners.....	12

Acronyms and abbreviations

VHR: Very High Resolution

ESA: European Spatial Agency

EO: Earth Observation

HR: High Resolution

MR: Medium Resolution

LR: Low Resolution

SAR: Synthetic Aperture Radar

EDAP: Earthnet Data Assessment Pilot

CSCDA: Copernicus Space Component Data Access

CCME: Copernicus Contributing Mission Entity

CSC: Copernicus Space Component

1 Introduction

This deliverable presents a survey to assess the usage of high-resolution satellite data from the Data Warehouse Copernicus Space Component Data Access by each FORCOAST partner during the year 2020. The use of high-resolution satellite data can provide an opportunity to complement or improve the data available from other datasets, as a direct source to services provision, boundary conditions or validation for FORCOAST pilot models.

Inside the Copernicus Data Warehouse, the Copernicus Space Component Data Access (CSCDA) provides access to satellite observations, separating them into two main categories: CORE Datasets and ADDITIONAL Datasets. The CORE datasets are characterised by a predefined large coverage, fixed tasking, data processing and delivery in timeliness parameters, while the ADDITIONAL datasets constitute the on-demand part of the data, providing more specific and flexible needs not covered by the CORE datasets.

In this document, we present a short description of both types of datasets and how to access them. We seek to create a guide to introduce them and facilitate their access to the partners, summarizing the content of the Data Access Portfolio [1] from CSCDA.

Then, we present a short survey to evaluate their use by the pilots during the year 2020. The main goal is to have feedback provided by the different pilots to identify which and how the datasets have been used.

This document will serve to estimate their uptake and identify potential gaps in used/not used data or even lack of knowledge about them, providing the guidelines to address the potential issues detected on the scope of the project and for future projects that need to deploy a strategy to introduce the ADDITIONAL datasets.

2. The Copernicus Space Component

The Copernicus Space Segment (one of the two elements of the Copernicus Space Component (CSC)) represents the geo-spatial resources for the Copernicus programme; it relies on a constellation of dedicated missions - the Sentinels - as well as on a set of EO space missions – the Copernicus Contributing Missions – contributing according to agreed operational scenarios and service levels. ESA is entrusted by the EU to guarantee a harmonised data provision to the Copernicus Space Segment. All missions of the Space Segment are classified primarily by sensor type (SAR or Optical) but also by resolution classes defined as follows:

Sensor Type	
SAR	
Optical	
Resolution	
VHR1: Very High-Resolution 1	<=1m
VHR2: Very High-Resolution 2	1m <res<=4m
HR1: High Resolution 1	4m <res<=10m
HR2: High Resolution 2	10m <re<=30m

MR1: Medium Resolution 1	30m <res<=100m
MR2: Medium Resolution 2	100m <res<=300m
LR: Low Resolution	res >300m

Table 1: Copernicus Space Segment mission classification by sensor type and resolution.

2.1 The CSCDA Access infrastructure

The Copernicus Space Component Data Access infrastructure, developed by ESA within the framework of the implementation of the Data Warehouse Phase 2 operations, relies on the ground segments of various data providers plus a set of coordinating functions in charge of ensuring internal coordination and control guaranteeing the users with a harmonized access to the data portfolio. The Copernicus Contributing Missions Access Support Functions and Platform (PRISM) is the core of the operational system. It is in charge of coordinating the data flow between the data providers and the Copernicus users, through the implementation of predefined datasets and the processing of asynchronous data requests based on mission capacity analysis as well as on the established agreements with CCMEs. PRISM ensures a harmonised user interface for the Copernicus users. It also supports the production and dissemination of multi-mission collections and provides a dedicated storage for CSC long-term data preservation.

2.1.1 Data Warehouse: The Core Datasets

The purpose of the CORE datasets is to consolidate predefined needs collected from Copernicus services and other activities requesting Earth Observation data, whether financed by the Union or related to policies of the Union, such as Union financed research projects or the activities of Union agencies (EEA, EMSA, SatCen, etc.). This approach allows a more robust and cost-effective access mechanism, which will be offered to a broad range of users and activities. CORE datasets are characterised by fixed tasking, systematic or not, data processing and delivery timeliness parameters defined in the dataset description.

A detailed description of the available datasets and their technical characteristics can be found in Section 5.1 of the Data Portofilo document [1] or directly through the ESA-CSCDA website [core-datasets](#) [3] section.

2.1.2 Data Warehouse: The ADDITIONAL Datasets

The ADDITIONAL Datasets are the on-demand part of the data portfolio, assumed flexible enough to accommodate specific requirements which are not covered by the CORE datasets and not foreseeable in advance. They represent a large volume of EO data classified by resolution type to cover a range of specifications by Copernicus users, e.g., geographical area, processing levels, lead times and delivery mechanisms to be selected at ordering time among the proposed ones for the specific dataset.

They are typically related to Emergency Management Services and Security Services activities, which require fresh data over critical areas. However, on-demand ADDITIONAL data are available also for non-rush applications, for which requirement details are only known at short notice (e.g. areas of interest, time window of interest, data type, etc.).

The ADDITIONAL datasets are needed to complement the CORE datasets, as many data characteristics are not known in advance (e.g., satellite tasking for rapid mapping or security applications, selected

areas for biodiversity monitoring). As for the CORE datasets, these ADDITIONAL datasets are expected to fulfil the needs collected from Copernicus services and other activities requesting Earth Observation data, whether financed by the Union or related to policies of the Union, like Union financed research projects and the activities of Union agencies (EEA, EMSA, SatGen, etc.).

In addition to the classification introduced by the sensor type and resolution for CORE Datasets (see Table 1), the ADDITIONAL datasets are classified including the Service Type:

Service Type	
Rush	Archive / New acquisition
Standard	Archive / New acquisition

Table 2: Copernicus Space Segment mission classification for ADDITIONAL datasets based on service type.

2.1.3 Available Datasets and access request

In this section, we present all the ADDITIONAL datasets users have access to. A detailed description of all available datasets and their technical characteristics can be found in Section 6.2 of the Access Data Portofilo document [1] and also on the ESA-CSCDA website: [ADDITIONAL-datasets](#)

Dataset ID	Dataset Title
DWH_MG2_CORE_01	2 Optical HR Pan EU coverages (Image2012)
DWH_MG2_CORE_02	2 Optical HR Pan EU coverages each (Image2006 and Image2009)
DWH_MG2b_CORE_03	Optical VHR2 coverage EU 2011-2013 and Riparian zones (gap filling from DWH phase 2)
DWH_MG2-3_CORE_08	Monthly composites EU 2011-2012
DWH_MG2_CORE_09	Sub-Saharan Optical coverage HR2 2011-2013
DWH_MG1_CORE_11	Sea Ice monitoring 2011-2014

Table 3: Available ADDITIONAL datasets.

Different user categories are defined by the European Commission to access the available datasets from the Data Warehouse. In Table 4 the availability of the ADDITIONAL datasets to each of these categories is shown.

Datasets	Copernicus Services	Union Institution	EU Research Projects space	EU Research Projects non-space	Public Authority	International Organisations and NGOs	Public
All ADDITIONAL datasets 2014-2022	D	D	D	D	D		
All ADDITIONAL datasets of DWH phase 1	D	D	D				

Table 4: Main authorities and their accessibility to the ADDITIONAL datasets (D=Download, V=View).

The European Commission also assigns a quota for each user category and each data type on a yearly basis, expressed in km². As an EU research project, the FORCOAST partners were assigned the quotas shown in Table 5.

Dataset title	Dataset ID	Quota (km ²)
Archive_standard_Optical_VHR1	D2_MG2b_FORC_011a	5,690
New acquisition_standard_Optical_VHR1	D2_MG2b_FORC_012a	4,800
Archive_standard_SAR_VHR2	D2_MG1_FORC_015b	4,800

Table 5: Quotas assigned for ADDITIONAL datasets to the FORCOAST project.

This quota is available for standard ADDITIONAL datasets. They are non-rush datasets populated by on-demand standard user requests for new or archived data.

Eligible FORCOAST partners with quota assigned for the required dataset shall submit a data request via the "Standard Data Request" interface after login within the CSCDA portal (<https://spacedata.copernicus.eu/>).

The detailed instructions to access the ADDITIONAL datasets, given in the Data Access Portfolio document [1], are the following:

1. Data request submission by the user: All order options shall be specified (AOI, acquisition window or archive time period, product type, required bands, processing level, polarization, delivery timeliness). The user cannot select a specific Contributing Mission but only a data type via the selection of a dataset; during order creation, the SCI (Services Coordinated Interface) will select the most suitable mission from those contributing to the dataset according to predefined criteria (minimum order size vs AOI, data cost, etc.).
2. Data request validation and order submission by the SCI team to the selected data provider: Order parameters and quota availability are checked and the order is sent to the most suitable Copernicus Contributing Mission Entity (CCME).
3. Order implementation by selected data provider: The CCME implements tasking (if needed) and data processing according to the parameters specified in the order.
4. Data delivery: CCME circulates the resulting product(s) to PRISM archive, according to delivery timeliness specified by the user.

At this point in time, the data are available for access by the originator of the request and for re-use by other eligible users.

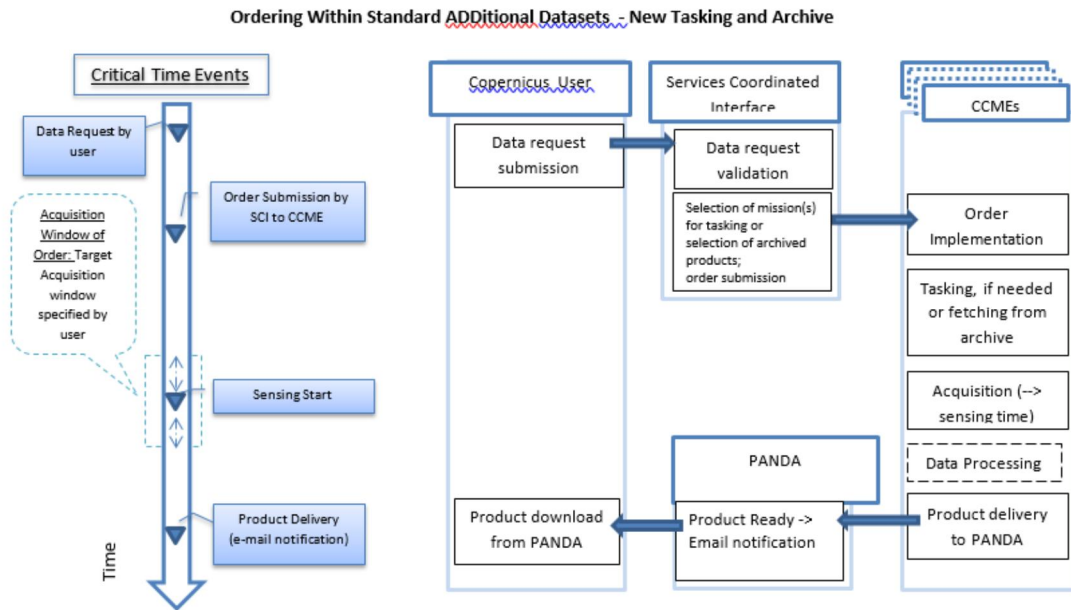


Figure 1: Work flow to request standard ADDITIONAL datasets.

3 Survey: ADDITIONAL Dataset usage in 2020

3.1 Survey introduction

We present a survey that shows the ADDITIONAL datasets use among the FORCOAST project partners. The following steps were taken:

- 1) We provided the Excel survey template below (Table 6) to be filled by each partner individually.

Dataset Variables	Type of use during 2020		
	Direct Use	Model Forcing	Model Validation
Other data from DWH used in 2020			

Table 6: Excel sheet shared with the partners.

The use categories are based on the following aspects.

- Model validation: The datasets are used to validate the results produced by the partner.
 - Model forcing: The datasets are used as boundary or initial conditions for the models developed by each partner.
 - Direct use: The datasets are used directly to provide a service without any processing affecting the dataset itself. For example, visualization through a platform.
 - Other: The use of the dataset does not fit in the previous categories. *Other* is considered when the partner does not specify the type of use assigned to the dataset. For example: operational download to use it in the future, either for an already-defined purpose or not.
- 2) Each partner was asked to answer the following questions:
1. What was the use of datasets in 2020?
 2. If none were used, explain the reasons why this was the case.

3.2 Results of the survey

3.2.1 Excel survey

Figure 2 shows the different datasets used by the different FORCOAST partners grouped by the main keys in each dataset name.

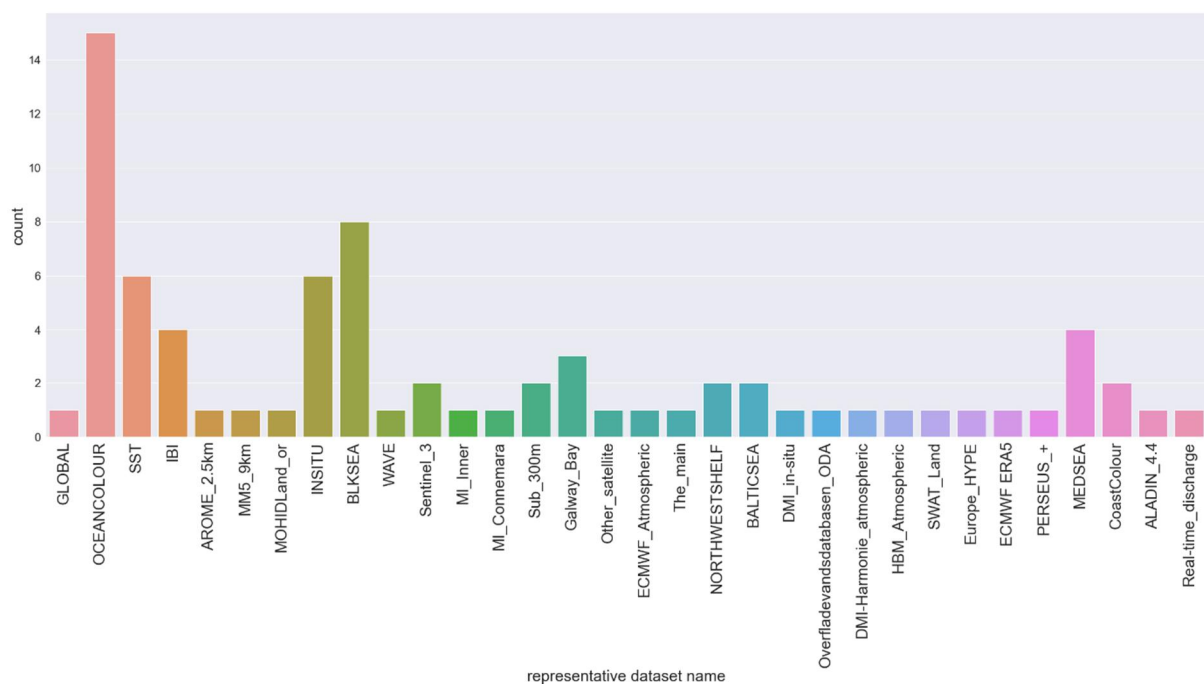


Figure 2: Datasets grouped by their representative names and usage.

The main conclusion reached after analysing the survey results is that none of the FORCOAST partners uses any datasets from the ADDITIONAL category in the CSCDA portfolio.

The datasets used by the different partners can be classified into three types:

- Satellite observation

- In situ observation
- Models

Processing levels go from level 0 (L0) raw data all the way up to level 4 (L4) modelled data received from multiple measurements, and are an indicator of how much processing has been applied to the raw data registered by, for example, an EO satellite.

To get deeper into the reasons for the lack of use of the ADDITIONAL datasets, the survey results were further analysed. First, a classification attending to the *processing level* of the datasets was considered (Figure 3).

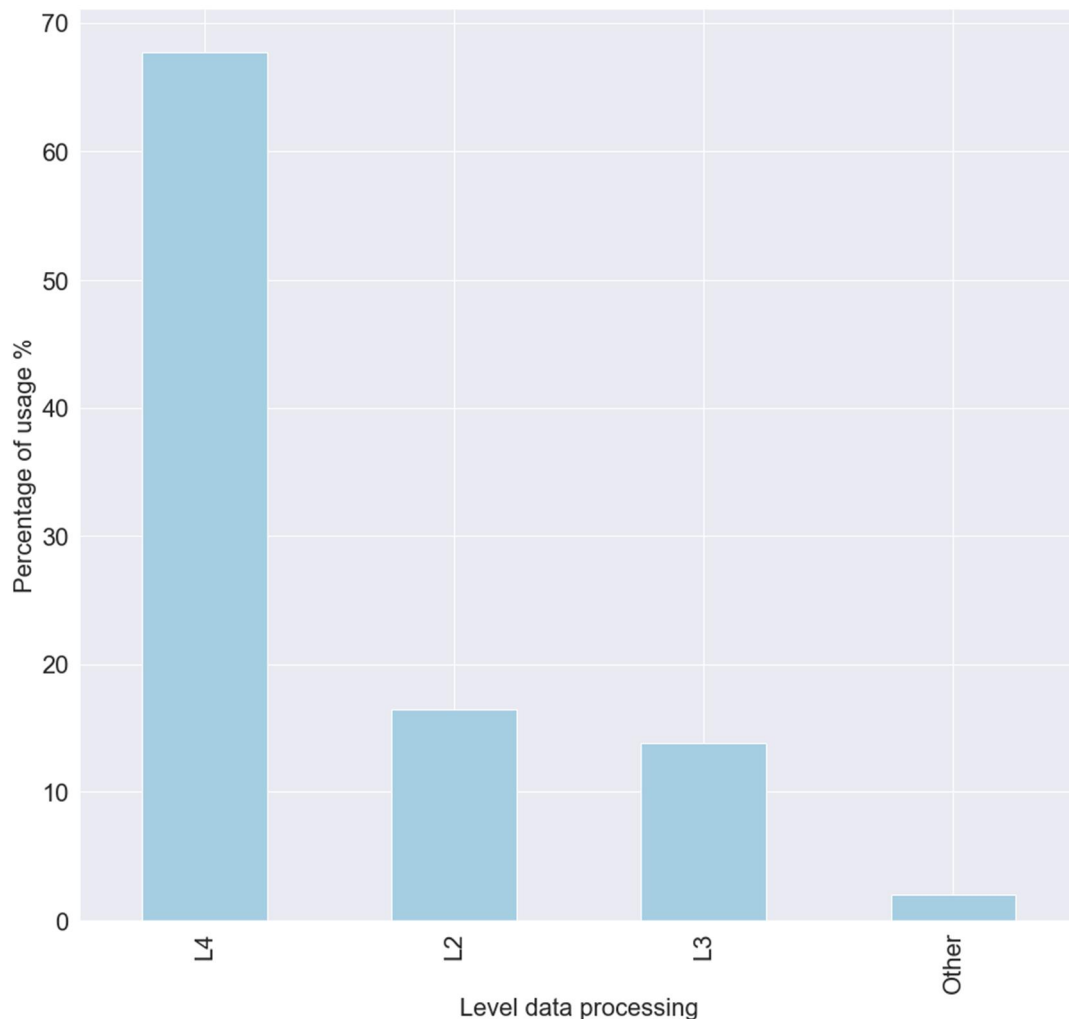


Figure 3: Level data processing and percentage of use.

At first glance, it can be seen that neither L0 nor L1 processing level data were used by any of the FORCOAST partners: all the satellite observation data (around 31% of all data) used by them correspond to L2, L3 or L4 processing level datasets. It must be noted that the ADDITIONAL datasets correspond to L0 or L1 processing level data.

Around half of the datasets used contain L4 processing level data (satellite observation or models). Most of the datasets used are derived from model outputs where assimilation is used. The *other*

category corresponds to datasets a processing level cannot be applied to, such as model outputs without an assimilation process, or those where we cannot find the source of the data to identify their level. L3 datasets make up around 13% of all the datasets, and L2 datasets, 15%. Most of the data from in situ stations are included in processing level L2 (40% of the L2 datasets used are made up of data collected from in-situ stations).

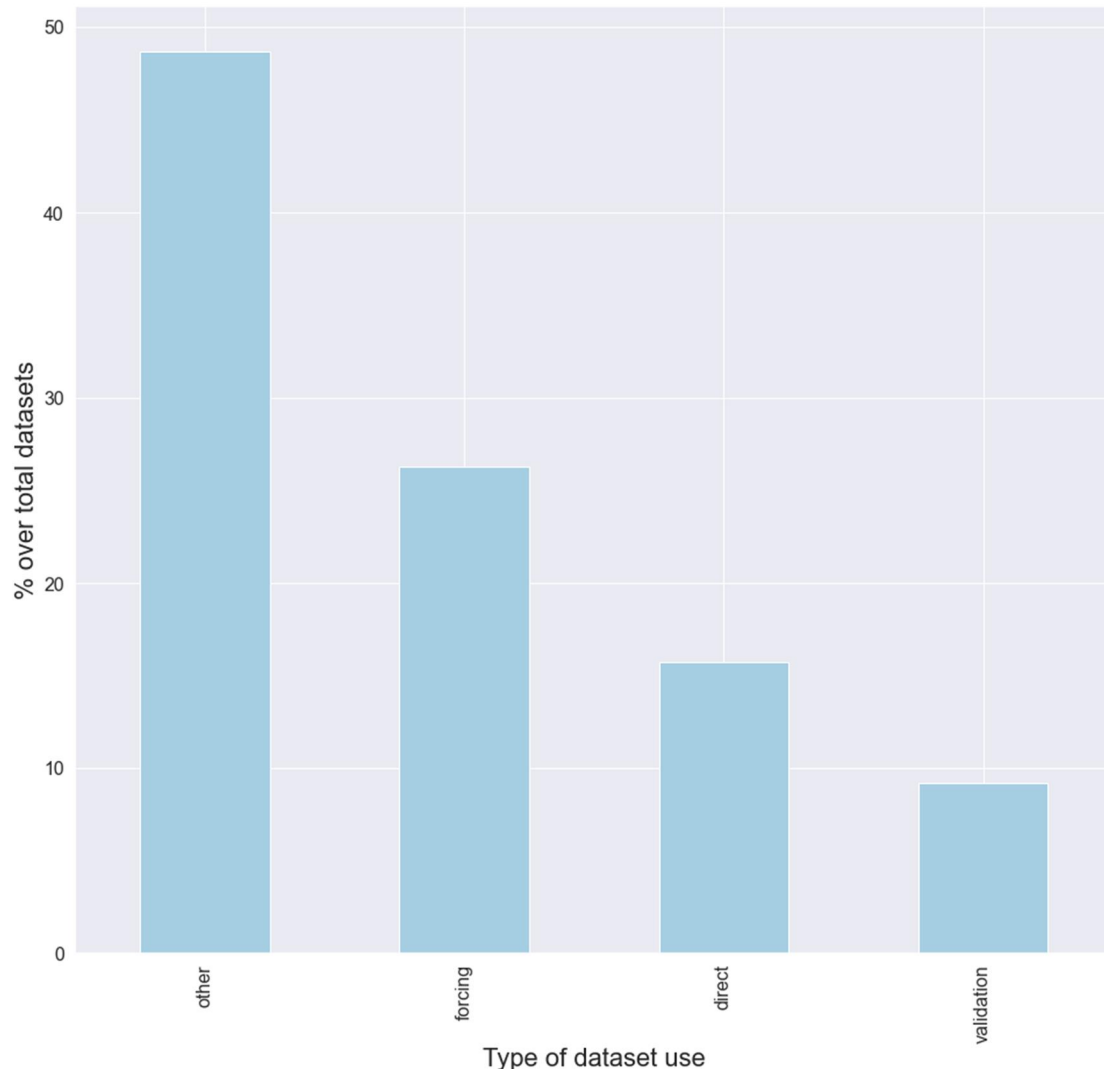


Figure 4: Type of dataset use and percentage over total number of datasets.

When the datasets are classified by their *Type of Use* (Figure 4), nearly half of the datasets are used for purposes different than *forcing*, *validation* or *direct*, 26% percent are used to force the models developed by each partner, and 16% are used as is (*direct use*). Finally, 9% are used for *validation* purposes.

To analyse better the *Type of use*, we group the datasets according to their *processing level* in Figure 5.

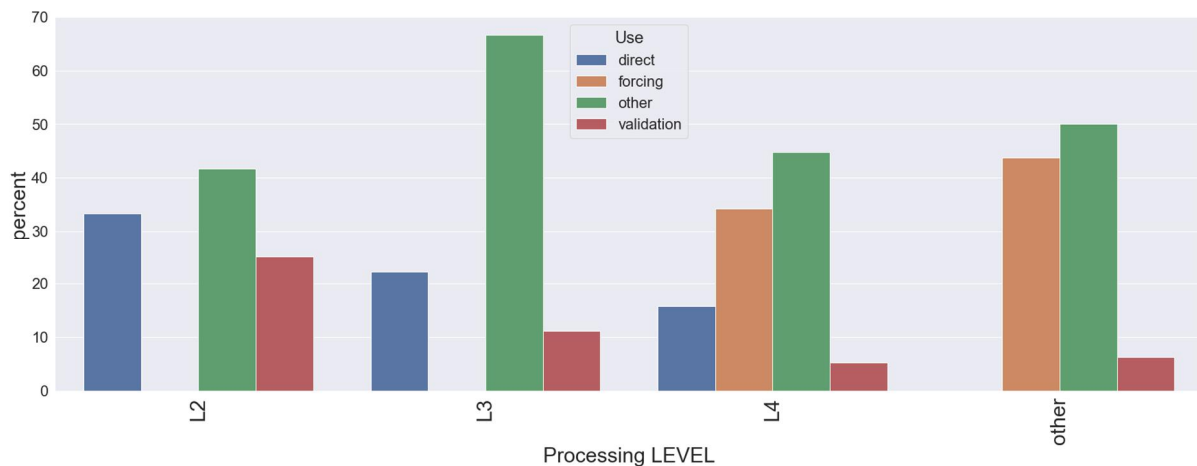


Figure 5: Type of use of datasets grouped by their processing level.

Figure 5 shows that datasets with processing levels lower than L4 are not used to force any models; however, their use increases for *validation* and *direct use*. Creating datasets for *forcing* conditions requires an assimilation process that the partners do not perform by themselves, they instead use already available datasets with higher processing levels.

3.2.2 ADDITIONAL datasets use

The different pilots were asked to provide ADDITIONAL information about why the ADDITIONAL datasets were not used.

The main reasons provided were:

1. They did not know about the benefits/existence of the ADDITIONAL datasets.
2. They do not find them useful for their project pilots.
3. The Pilots do not indicate usefulness of the datasets at this stage.

4 Conclusions and proposed actions

Despite most of the partners having both an established knowledge and widespread use of different Copernicus datasets and products, the lack of use of the ADDITIONAL datasets in 2020 comes down to four main reasons that are explained below and that should be addressed in the scope of the FORCOAST project:

- The lower the processing level, the higher the proportion of these datasets used for *direct use* or *validation*. Taking this into account, and considering the characteristics of the ADDITIONAL datasets, we can presume their main purpose could also be for *direct use* or *validation*. At this early stage of the project, the different pilots might feel more comfortable using standard and well known datasets. Therefore, the introduction of ADDITIONAL datasets could be considered once the pilot actions are established.

- The high percentage of datasets that fall into the *other* category (48%) shows that a large proportion of partners do not use the datasets at their disposal for a specific and predefined purpose. As the access to the ADDITIONAL datasets is limited and the assigned quota is shared among the different members of the FORCOAST project, their use must be analysed in advance and in detail before requesting any data downloads.
- As shown in Figure 4, 26% of the datasets' use fall into the *forcing* category. The ADDITIONAL datasets are not useful for this purpose as was noted in the previous section.

In general terms, one of the main reasons for lack of use is that DWH data are L0/L1, and this needs substantial pre-processing to be translated in terms of oceanic variables relevant for model forcings, validation, or direct use in terms of service.

To address the lack of use of ADDITIONAL datasets and to enhance the use of Copernicus products, an active policy related to the access of these datasets should be followed in the upcoming year. It should inform about the benefits they can bring to the project, their characteristics, how to request them and how to apply the requested data. To that end, a Remote Sensing working group has been created with the following goals:

1. Explain the benefits of using the data these datasets provide.
2. Support the access to these datasets. Coordinate the requests among the different project partners.
3. Support the partners on dataset selection according to their pilot actions and needs.

Pilot-per-Pilot meetings are taking place over the next weeks to address on an individual basis the potential applications of Copernicus (and Remote Sensing) for their services that are being developed.

4 References

1. [Copernicus Space Component Data Access Portfolio](#)
2. [CSC: Reference Card](#)
3. [CSCDA: Core Datasets](#)