FORCOAST

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

PROJECT DELIVERABLE REPORT

Deliverable Number: 3.3
Deliverable Title: DWH request for 2021
Author(s): Daniel Garaboa (MI_ES), Daniel Lowe (MI_ES)
Work Package Number:3
Work Package Title: Service Design

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

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	1st November 2019, 30 months
Project website	https://forcoast.eu/

Deliverable Information	
Work package number	3
Work package title	Service Design
Deliverable number	3.3
Deliverable title	Data Warehouse (DWH) request for 2021
Description	Using the provided template this deliverable states whether ADDITIONAL data is requested for the following year from the DWH. 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
Revision number	1
Revision Date	30/04/2021
Status (Final (F), Draft (D), Revised Draft (RV))	RV
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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Document History			
Revision	Date	Modification	Author
1	28/04/2021	1st Draft	Daniel Garaboa, Daniel Lowe
2	30/04/2021	1st Review	Luis Rodriguez

Approvals				
	Name	Organisation	Date	Signature (initials)
Coordinator	Ghada El Serafy	Deltares		
WP Leaders	Arthur Capet	ULiege		

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

PROPRIETARY RIGHTS STATEMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Executive Summary

This document shows the results obtained from the analysis of the Earth Observation datasets used within the framework of the FORCOAST project by its members, and the planned use of 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, briefly introducing the available datasets in the CSCDA, distinguishing between the CORE and ADDITIONAL datasets, and indicating which datasets are available to the FORCOAST members.

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 Pilots. The results of the survey sent to each of them is presented in Section 3.

Finally, the analysis of the information provided by the Pilots suggests that an increasing number of datasets are expected to be used by the different Pilots to validate their models. However, the Data Warehouse products are L0/L1 processing level datasets, and this needs substantial pre-processing to be translated in terms of oceanic variables relevant for model forcings, validation, or direct use, which makes them less attractive at this stage to be used by the FORCOAST partners.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Table of Contents

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

List of Figures

List of Tables

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Acronyms and abbreviatures

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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

1 Introduction

This deliverable presents the results of a survey conducted among the different Pilots that form the FORCOAST project with the objective of finding out which high-resolution satellite datasets from the Copernicus Data Warehouse website are expected to be used during 2021.

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.

Deliverable 3.2 "Data Warehouse (DWH) use for 2020" presented a short description of both types of datasets and how to access them, intended as a guide to facilitate their access to the partners, summarizing the content of the Data Access Portfolio [1] from CSCDA.

This document focuses on the results obtained from the aforementioned survey, with an emphasis on the differences between the information provided by the Pilots about their dataset usage during 2020 and their plans for 2021, and the role high-resolution datasets might play in their work.

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

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

LR: Low Resolution	res >300m

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

2.1 Data Warehouse datasets

Within the Copernicus Space Component Data Access infrastructure, there are two types of datasets users have access to:

- CORE datasets: Their purpose is to consolidate predefined needs collected from Copernicus services and other activities requesting Earth Observation data. They are characterised by fixed tasking, systematic or not, data processing and delivery timeliness parameters defined in the dataset description. More information can be found in the CORE Datasets section on the ESA-CSCDA website [2].
- ADDITIONAL datasets: These 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. A detailed description of all available datasets and their technical characteristics can be found in Section 6.2 of the Access Data Portfolio document [1] and on the ESA-CSCDA website [3].

2.1.1 Available ADDITIONAL Datasets and assigned quota

In this section, we present all the ADDITIONAL datasets users have access to and the quota assigned to the FORCOAST project.

The following table shows the available ADDITIONAL datasets at the CSCDA Data Warehouse:

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 2: Available ADDITIONAL datasets.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Different user categories are defined by the European Commission to access the available datasets from the Data Warehouse. The European Commission also assigns a quota for each user category and each data type on a yearly basis, expressed in km^2. As an EU research project, the FORCOAST partners were assigned the quotas shown in Table 3.

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 3: 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.

3 Survey: Expected use of ADDITIONAL Datasets in 2021

3.1 Survey introduction

We present a survey that shows the ADDITIONAL datasets expected use among the FORCOAST project partners during 2021, which consisted of an Excel document sent out to each partner to be filled in individually. It contained two sheets:

1. The first sheet asked about datasets currently being used by each partner and the type of use they were given. It also asked about other DWH datasets that are expected to be used during 2021 (Table 4).

Dataset Variables	Type of use during 2021		
	Direct Use	Model Forcing	Model Validation
	New data from DWH that you expect to use in 2021		
	Direct Use	Model Forcing	Model Validation

Table 4: First Excel sheet extracted from the document shared with the partners.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

2. The second sheet asked explicitly about the ADDITIONAL datasets that each partner is expected to use in 2021, including the region coordinates and the satellite mission covering the selected region (Table 5).

Dataset	Region of operation and M	Comments about expected use in 2021
Ex: Archive_standard_Optical_VHR1	AOI = - (lat,lon) = - Missions = -	
Archive_standard_SAR_VHR2	AOI = - (lat,lon) = - Missions = -	
New acquisition_standard_Optical_VHR1	AOI = - (lat,lon) = - Missions = -	
New acquisition_standard_SAR_VHR2	AOI = - (lat,lon) = - Missions = -	

Table 5: Second Excel sheet extracted from the document shared with the partners.

The use categories included in the first sheet 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.

3.2 Results of the survey

3.2.1 Excel survey

To analyse the results, we will first focus on the ADDITIONAL datasets. The main conclusion is that only a quarter of the Pilots report that they expect to use these datasets in 2021, while the remaining three-quarters do not expect to use them at all. To understand these results, we will first analyse the information filled in by each Pilot in the both sheets of the Excel document, which contains data about the datasets currently used by them. Figure 1 shows this information grouped by dataset name.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

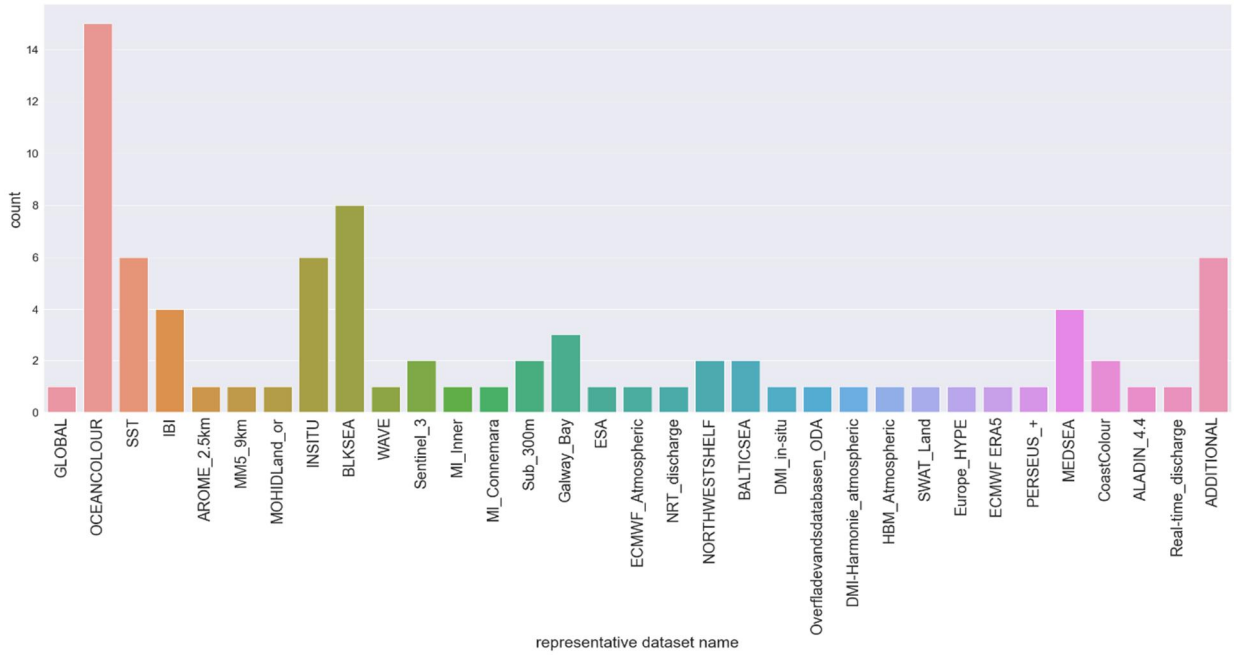

Figure 1: Datasets grouped by their representative names and usage.

To understand the reasons for the low expectations of usage, the survey results were analysed comparing the results of usage for 2020 (Deliverable 3.2) within the prevision of usage for 2021.

First, a classification attending to the *processing level* of the datasets was considered. 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.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

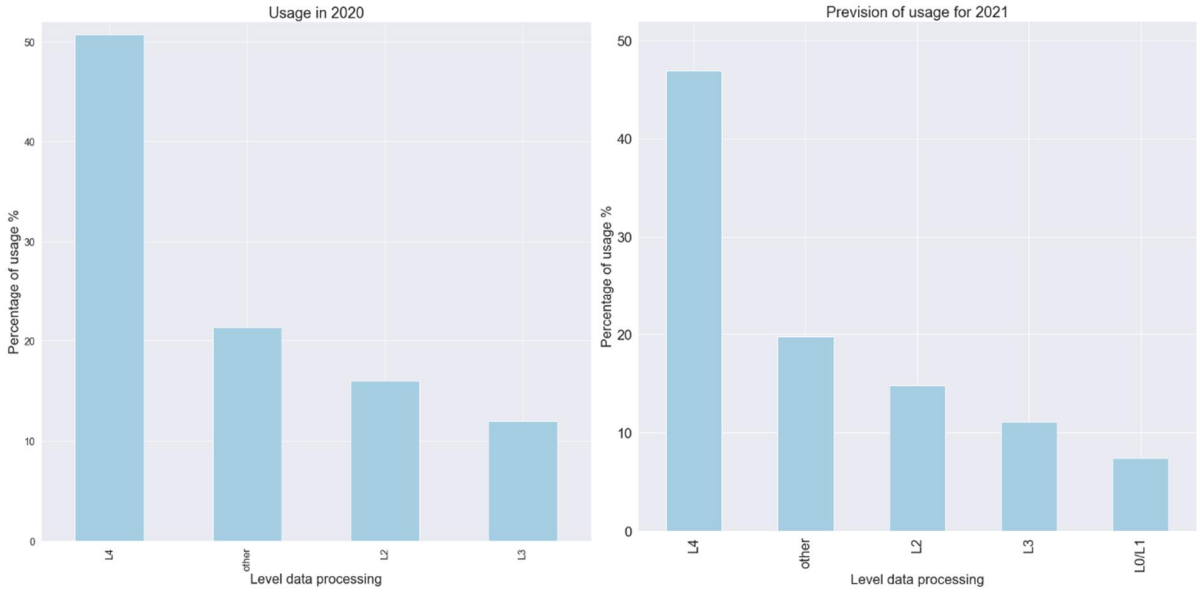

Figure 2 Level data processing and percentage of use for 2020 (use) and 2021 (prevision of use).

At first glance, in Figure 2, there is an increase in the prevision of usage in a 7.4% with datasets with a processing level of L0/L1 with respect to the usage in 2020. This is a direct consequence of the expected usage of ADDITIONAL datasets reported by the quarter of the pilots.

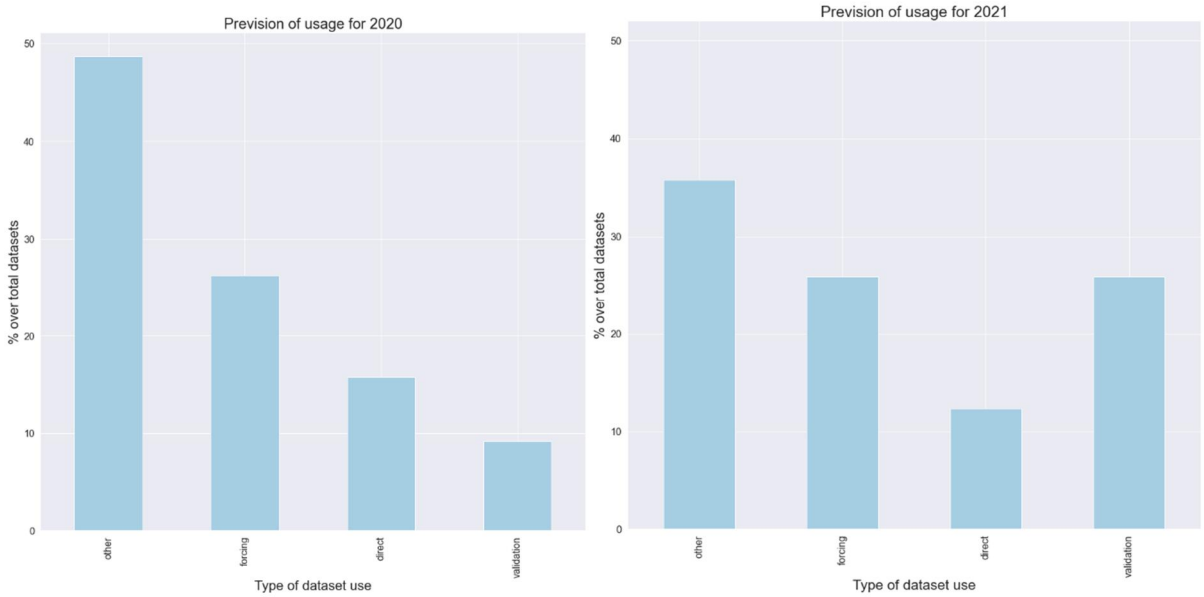

Figure 3 Type of dataset use and percentage over total number of datasets for 2020 (use) and 2021(prevision of use).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Analyzing the *Type of dataset use*, the survey shows an increase of usage of datasets for validation, going from 9% in 2020 to a prevision of 26% in 2021. Also, in datasets categorized as *other*, (the partners do not specify a use in the other categories) the usage reduced from 48% to 36%. This reduction is linked with the increase of usage for *validation*, showing that some datasets, which do not have a clear *type of use* in 2020, will have it in the present year, mainly for validation purposes. It can be presumed that many partners have an expectation of entering in a validation stage (mainly those ones related to operational models) of their results in this stage of the project.

In agreement with this, there is no change of usage in the *forcing* category. The usage of datasets as an input for numerical models is expected to remains the same during 2021.

To understand what datasets are used for validation, we analyze the survey categorizing the results by the *type of use* and the *type of the dataset,*

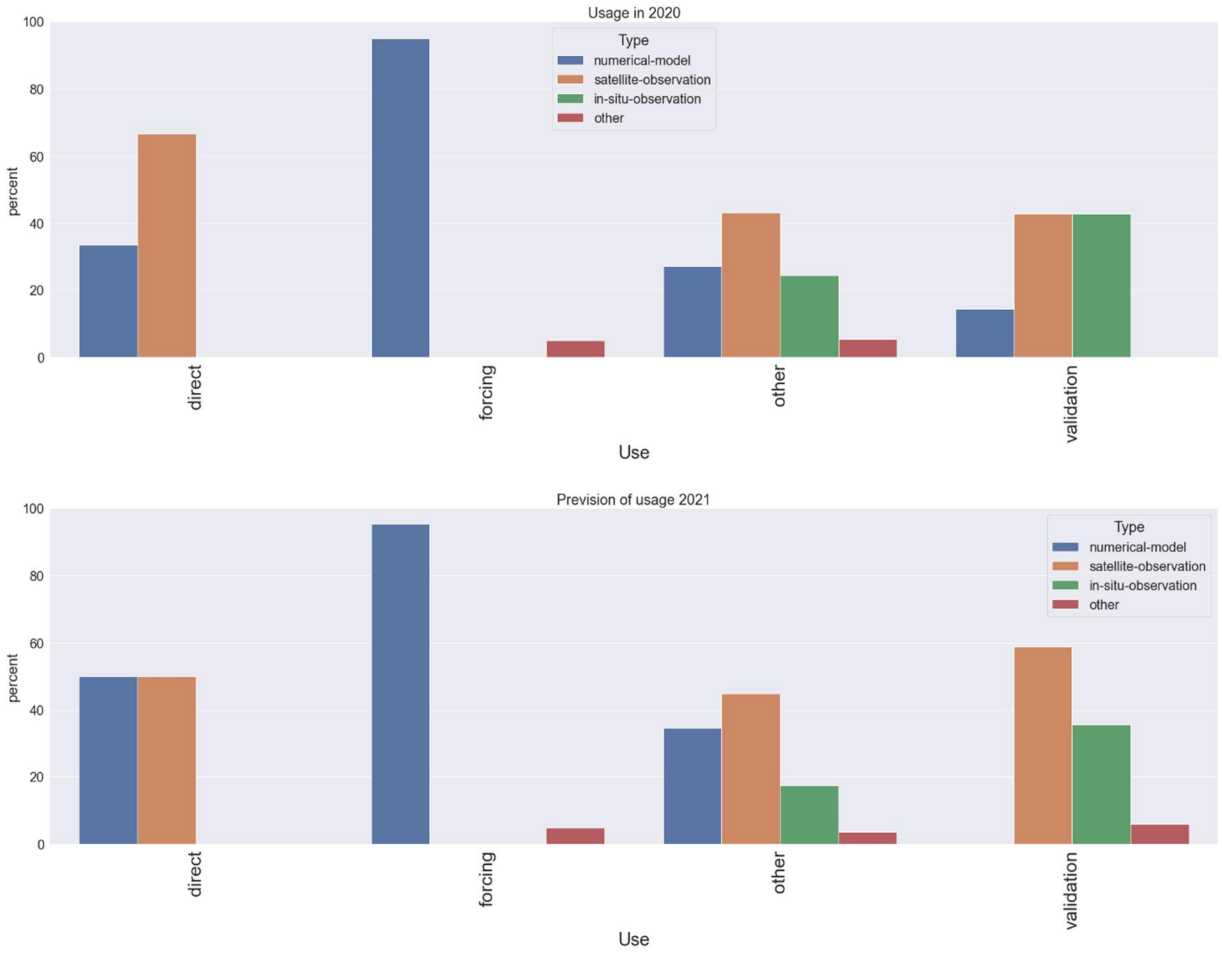

Figure 4 Type of dataset use and percentage over total number of datasets for 2020 (use) and 2021(prevision of use).

Focusing on *validation* category in Figure 4, we can observe that the prevision of use of *satellite-observation* datasets for *validation* increases 18% with respect to 2020, being close to the 60% of the

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

satellite datasets used for *validation*, and hence being a source for validation higher than *in-situ-observation*. However, this increase does not totally correspond with the expected use of ADDITIONAL datasets for 2021.

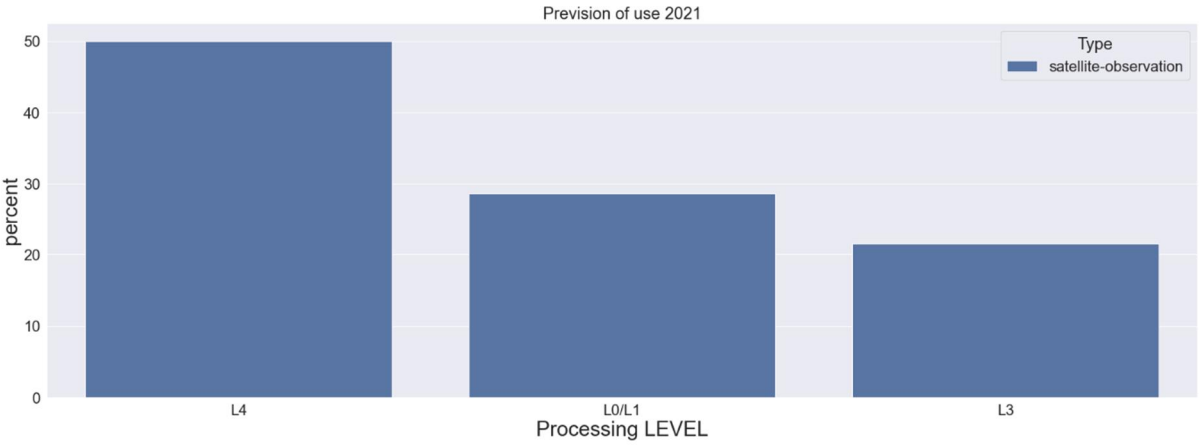

Figure 5 level of processing of satellite data for validation purposes.

In Figure 5, *satellite* datasets with L3 and L4 represent around the 70% of datasets used for *validation*. The results confirm that the partners find more attractive to use datasets with higher level of processing than L0/L1 to validate their results.

4 Conclusions and proposed actions

With respect to the previous year, the previsions of usage for ADDITIONAL datasets has increased from 2020. At least a quarter of the partners report that they are going to use the ADDITIONAL datasets. We should notice that this survey is about the expectations that the partners have about the usage which cannot translate in a real usage.

The main results extracted from the survey are:

- The expectation of usage for ADDITIONAL datasets is a 7% over the total datasets used.

- Based on the prevision of the dataset usage, the partners expect to enter in a validation stage of their results in the year 2021. A prevision of an increase of 26 % with respect to 2020 is expected in the use of datasets for validation. The validation stage expected to use the satellite data in 60% and in-situ observation in 40%.

- The expected increase of use of satellite for validation does not translate into the usage of ADDITIONAL datasets. Non-ADDITIONAL datasets represent the 70% of the total satellite data.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.

Partners prefer to use well know and established datasets with higher processing levels to avoid the pre-processing stages to produce compatible datasets for their needs. We can presume that the pre-processing of raw data is a big effort that is outside of the scope of the project, and it could be considered for some partners once their pilot actions are fully established with well know datasets.

5 References

1. Copernicus Space Component Data Access Portfolio
2. CSCDA: CORE Datasets
3. CSCDA: ADDITIONAL Datasets

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 870465.