LABPLAS

Horizon 2020
European Union Funding
for Research & Innovation

Land-Based Solutions for Plastics in the Sea

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003954

D4.7 - 1st Report on the results for the detection of macroplastics using remote sensing

Due date of deliverable: 31/03/2024

Actual submission date: 26/03/2024

Horizon 2020
European Union Funding
for Research & Innovation

PROJECT INFORMATION

Project number: 101003954
Project acronym: LABPLAS
Project full title: Land-Based Solutions for Plastics in the Sea
Call: H2020-SC5-2018-2019-2020 submitted for H2020-SC5-2020-2 / 03 Sep 2020
Topic: CE-SC5-30-2020 – Plastics in the environment: understanding the sources, transport, distribution and impacts of plastics pollution
Type of action: RIA – Research and Innovation Action
Starting date: June 1st, 2021
Duration: 48 months
List of participants:

No	Participant name	Acronym	Country	Type
1	UNIVERSIDADE DE VIGO	UVI	SPAIN	HES
2	UNIVERSIDADE DA CORUÑA	UDC	SPAIN	HES
3	Bundesanstalt fuer Gewaesserkunde	BfG	GERMANY	RTO
4	LABORATORIO IBERICO INTERNACIONAL DE NANOTECNOLOGIA	INL	PORTUGAL	RTO
5	KATHOLIEKE UNIVERSITEIT LEUVEN	KUL	BELGIUM	HES
6	HELMHOLTZ ZENTRUM FUR OZEANFORSCHUNG KIEL	GEOMAR	GERMANY	RTO
7	NATIONAL OCEANOGRAPHY CENTRE	NOC	UNITED KINGDOM	RTO
8	SORBONNE UNIVERSITE	SU	FRANCE	HES
9	OPEN UNIVERSITEIT NEDERLAND	OUNL	NETHERLANDS	HES
10	LEIBNIZ INSTITUTE FOR BALTIC SEA RESEARCH	IOW	GERMANY	RTO
11	ASSOCIACAO PARA O DESENVOLVIMENTO DO ATLANTIC INTERNATIONAL RESEARCH CENTRE	AC	PORTUGAL	RTO
12	UNIVERSIDADE FEDERAL DO SAO PAULO	UNIFESP	BRAZIL	HES
13	BASF SE	BASF	GERMANY	LE
14	TG ENVIRONMENTAL RESEARCH	ER	UNITED KINGDOM	SME
15	CONTACTICA S.L.	CTA	SPAIN	SME
16	STICHTING EGI	EGI	NETHERLANDS	Non-P
17	STICHTING RADBOUD UNIVERSITEIT	RU	NETHERLANDS	HES
18	UNIVERSIDADE FEDERAL DO PARÁ	UFPA	BRAZIL	HES

DELIVERABLE DETAILS

Document number:	D4.7
Document title:	1st Report on the results for the detection of macroplastics using remote sensing
Dissemination level	PU – Public
Period:	PR2
WP:	WP4
Task:	Task 4.5
Status:	Final – After Review
Author:	ASSOCIAÇÃO PARA O DESENVOLVIMENTO DO ATLANTIC INTERNATIONAL RESEARCH CENTRE
Reviewers:	
Recommended citation format	A. Giusti, E. Castanho, P. Silva, and A. Valente, 2024, 1st Report on the results for the detection of macroplastics using remote sensing, Deliverable 4.7, LABPLAS Grant Agreement No. 101003954 H2020-SC5-2020-2
Executive summary:	The Land-Based Solutions for Plastics in the Sea (LabPlas) project aims to understand the main sources, transport mechanisms, distribution, and impacts of plastic pollution in all environmental compartments by applying technological advances, promoting truly biodegradable novel materials, developing innovative models for the fate, effects and risks of plastics, and presenting results for long term solutions and supporting decision-makers. This report corresponds to Deliverable 4.7 - 1st Report on the results for the detection of macroplastics using remote sensing - resulting from Task 4.5 of the LAPBLAS project. It provides a state-of-the-art insight on current technologies and methodologies for satellite remote sensing of macroplastics with focus on the marine environmental compartment, and presents the work developed regarding collection of satellite data, spectral library construction, and application of Machine Learning models for macroplastics detection. The work was centred on aggregations of macroplastics at the surface of the ocean, in the order of magnitude of meters, that could be detectable by current satellite sensors, in particular the Copernicus Sentinel-2 Mission.

Version	Date	Comments
1	2024-03-26	Final version
2	2024-07-23	Reviewed version with a new sentence to provide information on data repository (last sentence of section 7)

Disclaimer

The views and opinions expressed in this document reflect only the authors' views, and not necessarily those of the European Commission.

TABLE OF CONTENTS

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

Abbreviation / Acronym	Description
AC	Atmospheric Correction
BOA	Bottom-Of-Atmosphere
CNN	Convolutional Neural Networks
CRS	Coordinate Reference System
DNN	Deep Neural Network
EO	Earth Observation
FAI	Floating Algae Index
FDI	Floating Debris Index
FDI	Floating Debris Index
FOV	Field Of View
GEE	Google Earth Engine
kNDVI	kernel Normalized Difference Vegetation Index
LSU	Linear Spectral Unmixing
MARIDA	MARIne Debris Archive
ML	Machine Learning
NDVI	Normalised Vegetation Difference Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infra-Red
PI	Plastic Index
PLP	Plastic Litter Project
RF	Random Forest
ROI	Region Of Interest
SMNP	Small micro and nanoplastics
SWIR	Short-Wave Infra-Red
t-SNE	t-distributed Stochastic Neighbourhood Embedding
TOA	Top-Of-Atmosphere
VHR	Very-High Resolution
VIS	Visible
WSI	Windrows Spectral Index
XGBoost	eXtreme Gradient Boosting

1 INTRODUCTION

1.1 Overview and purpose

Plastics are an integral part of modern life, used in a broad range of applications such as packaging, construction, transport, furniture, medical applications, and machinery. However, this comes at an increasing cost, since a massive accumulation of plastic is currently taking place in the natural environment, with a still poorly understood impact on society, economy and human health.

The Land-Based Solutions for Plastics in the Sea (LabPlas) project addresses the challenge of plastic pollution by applying technological advances regarding sampling, analysis, quantification, up-scalable models, and novel materials to evaluate the interactions of plastics with all compartments (freshwater, marine, terrestrial, atmosphere and aquatic biota) and promote decision making by European agencies. The project also identifies methods for more accurate assessment of the abundance, distribution, and toxicity determination of plastics in the environment, giving the opportunity to new developments of cutting-edge technologies.

The project comprises two outstanding case-studies, where plastic amounts within water, sediment and biota will be determined (Figure 1):

1) The Elbe and Thames river basins feeding into the North Sea (case study site 1 with 18 sampling sites);
2) Northwest Iberian Peninsula with a selected small Atlantic river basin (case study site 2 with 6 sampling sites).

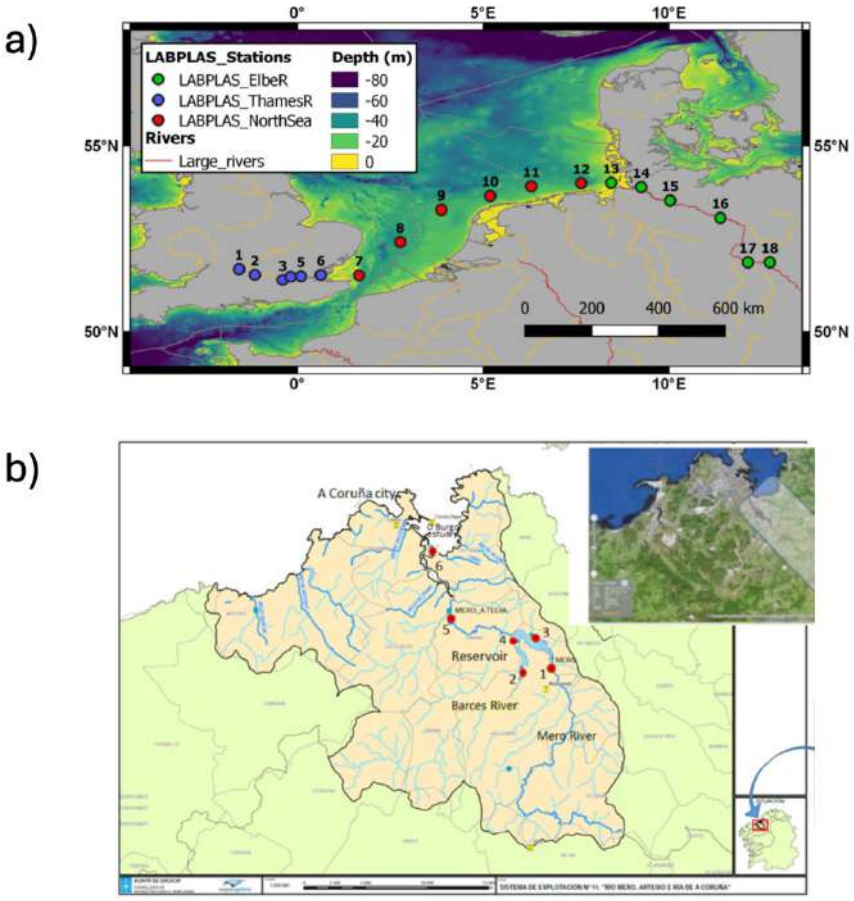

Figure 1 - LabPlas case study areas. a): Thames, North Sea, and Elbe. b): Mero-Barcés river basin.

The WP4 (Smart Hubs) of Labplas is a technological WP that includes advanced analytics proposed to detect small micro and nanoplastics (SMNP), chemical additives, as well as macroplastics. The present document reports the first results of the Task 4.5 "Advanced satellite techniques for determination of macroplastics", where macroplastics detection was carried out using Copernicus satellite data, combined with in-situ and other data sources. Results focus on aggregations of macroplastics at the surface of the ocean, in the order of magnitude of meters, that could be detectable by the spatial resolution of current satellite sensors.

Improving the monitoring for macroplastics is critical for several reasons. For instance, it has been recently estimated that a total of 3200 kilotons of plastic floats in the ocean, with nearly all mass (95%) contained in macroplastics [1]. These larger plastics will fragment into SMNP with a likely increase in bioavailability and risk to organisms and human health as particle size decreases [2]. Furthermore, once fragmentation into SMNP occurs, removal from the ocean becomes nearly impossible. It is thus urgent for developing innovative methods for monitoring macroplastics, due to their relevance in mass budgets and consequences of their fragmentation over time, which could contribute to a better understanding of the sources, transport, distributions and trends of plastic pollution and for providing information to decision-makers.

In present work, the detection of macroplastics by satellites was ensured by selecting the most appropriate Earth Observation (EO) satellite data, namely of the Copernicus Sentinel-2 Mission, from not only the project study areas, but also from locations worldwide with artificially-deployed plastic targets, or known macroplastics accumulations resulting from major pollution events. The satellite imagery was downloaded and passed through several stages of processing to extract the spectral signatures of different features including the macroplastics patches. The extraction of signatures was leveraged by calculating spectral indices that enabled the retrieval of relevant spectral information. The spectral signatures were then compiled into a library that was used as a training dataset for development of several Machine Learning (ML) models for macroplastics detection.

The objective of the deliverable D4.7 is to present the existing literature on detection of macroplastics using satellite remote sensing and the results carried out during the project, with focus on the marine environmental compartment. It is planned that Deliverable D4.8 will complete this document with the results regarding suspected plastic debris detection going into detail about the developed algorithms, more advanced ML models, and the development of an automatic tool encompassing the complete range of developed processing steps from download to final classification of the satellite images.

1.2 Structure

The steps carried out for Task 4.5 and described in this deliverable are:

1) State-of-the-art techniques for macroplastics detection using EO. This is addressed in section 2;
2) Data collection of plastic debris events, artificial plastic targets and the correspondent availability of EO imagery. This regards the work in Subtask 4.5.1 and it is addressed in section 3;
3) Pre-processing methods applied to the satellite data. This concerns results of Subtask 4.5.2 and it is addressed in section 4;
4) Spectral signatures of different materials, including verified agglomerations of macroplastic debris. This outlines efforts of Subtask 4.5.3 and it is addressed in section 5;
5) Machine learning methods for the detection of suspected macroplastics patches on satellite imagery. This reflects the work carried under Subtasks 4.5.4 and 4.5.5 and are addressed in section 6.

2 CURRENT SATELLITE AND MACHINE LEARNING TECHNIQUES FOR MACROPLASTICS DETECTION

The following section presents an overview of the state-of-the-art of satellite-based detection of macroplastics, with a particular focus on available satellite sensors and detection techniques. The presented considerations, results and assumptions from reference literature represented the base for the adopted approach carried out within the scope of Task 4.5.

The use of satellite images for the detection and monitoring of macroplastics in the ocean has been under investigation and development during the last years and different methodologies have been proposed and tested in literature with promising results. Satellites have great advantages with respect to traditional survey methods because they offer regular monitoring and detection capability in hotspots or remote areas, with synoptic coverage over large areas, to guide clean-up and plastic pollution mitigation strategies [3], [4], [5].

Several satellite EO missions are currently orbiting the earth, each having varied capabilities in terms of spectral sensitivity, temporal revisit frequency and spatial resolutions. In the task of floating macroplastics patch detection, suitable satellite sensors ideally produce imagery with a pixel size of at least 10 m x 10 m (High-Resolution), to as fine as 0.3 m x 0.3 m (Very-High Resolution, VHR). However, finer levels of detail determine trade-offs, as Very-High Resolutions (VHR) generally decrease the observable portion of the ocean (swath width) during satellite passages. Therefore, an appropriate balance must be satisfied among detail, coverage, and spectral discretization in the selection of satellite sensors, as well as of the acquisition costs.

Since the initial efforts to detect floating macroplastics patches in the ocean, passive optical sensors have been the most used, in particular though the High-Resolution optical multispectral imagery from Copernicus Sentinel-2 (Table I). VHR commercial satellites have also been tested. WorldView-3 VHR imagery was used for example by [6] for the detection of marine debris in the Sea of Japan, by [7] to study floating plastic accumulation in the North Pacific Garbage Patch or to improve the detection potential of Sentinel-2 through image fusion by [8]. PlanetScope VHR data supported the detection of floating litter in the study carried out by [9] in the Gulf of Honduras or analyse illegal waste detection in riverine environments [10]. Nevertheless, using VHR data comes typically at a cost per image, which is prohibitive for regular monitoring purposes. Other sensors, such as hyperspectral sensors ([11], [12]) while offering more detailed spectral information, have the disadvantage of a lower spatial resolution (30 m, Table I). As for active radar sensors ([13], [14], [15]), while they can detect objects through cloud cover or at night, they the lack the spectral information offered by optical data, which makes plastic detection from space very challenging.

Table I - Satellite sensor and characteristics used in the literature for the detection of macroplastics.

Satellite/Sensor Name	Spatial Resolution (GSD)	Swath width	Spectral Resolution	Revisit Time (days)
Sentinel-2/MSI	10 m (max)	290 km	Multispectral (13 bands)	5
WorldView-3	0.31 cm (max)	13.1 km	Multispectral (29 bands)	1

PlanetScope	3 m (max)	24.6 km	Multispectral (4 bands)	1
Sentinel-1/SAR	5 m (max)	410 km	_	6 (12)
PRISMA	30 m	30 km	Hyperspectral (239 bands)	29

Currently, the Copernicus Sentinel-2 mission with its MultiSpectral Instrument stands our as the most tested and promising source of optical satellite data for macroplastic detection ([10], [15], [16], [17], [18]). This is mainly because it offers the highest spatial resolutions (up to 10 m) with free access. Other freely-available optical sensors (e.g. Landsat) have lower spatial resolutions. Sentinel-2 provides systematic global coverage of the coastal ocean (every 5 days) and data on various spectral bands in visible, near-infrared, and shortwave-infrared (Table II).

Table II - Sentinel-2 spectral bands.

Band	Description	Spatial resolution (m)	Central wavelength (nm)	Bandwidth (nm)
Band 1	Coastal aerosol	60	442.7	21
Band 2	Blue	10	492.4	66
Band 3	Green	10	559.8	36
Band 4	Red	10	664.6	31
Band 5	Vegetation red edge 1	20	704.1	15
Band 6	Vegetation red edge 2	20	740.5	15
Band 7	Vegetation red edge 3	20	782.8	20
Band 8	Near Infrared	10	832.8	106
Band 8A	Narrow NIR	20	864.7	21

Band 9	Water vapour	60	945.1	20
Band 10	Shortwave Infrared - Cirrus	60	1373.5	31
Band 11	SWIR 1	20	1613.7	91
Band 12	SWIR 2	20	2202.4	175

Initial efforts and spectral analysis studies on Sentinel-2 imagery demonstrated the capacity to discriminate pixels containing plastic debris from adjacent water pixels by leveraging the spectral properties of plastic-based materials. Tests conducted in controlled experiments ([13], [19], [20], [21]), by deploying artificial plastic targets of different compositions, sizes and with variable degrees of submersion and biological fouling concluded that the spectral contrast between floating matters and background waters was mostly in the Near Infra-Red (NIR), regardless of the type of floating matter (plastics, wood and others). Basically, this is because floating materials reflects the sunlight in the NIR band, whereas the neighbouring water absorbs light at these wavelengths. The reflectance of floating matter in the Short-Wave Infra-Red (SWIR) bands was several times lower than in the NIR bands. It was also found that the portion of plastic should be at least 30% – 50% of the total area of the pixel to be sufficiently reflective in the NIR band. The presence of biological material on the plastic (biofouling), influenced the spectral behaviour in the Visible (VIS) part of the spectrum. At the same time, submersion and wetness affected the spectral signature, especially in the NIR part of the spectrum, with lower reflectance compared to virgin plastics [22].

Recognizing the ability to distinguish plastic artificial targets from water, plastic patches within satellite imagery have been extracted through various methods, including spectral anomalies [23], spectral unmixing ([19], [21]), the utilization of spectral indexes as well as ML classification algorithms. [24], [25], [26] proposed specific spectral indexes, respectively the Windrows Spectral Index (WSI), the Plastic Index (PI) and the Floating Debris Index (FDI). The latter, based on the floating algae index (FAI) of [27] was employed in combination with the Normalised Vegetation Difference Index (NDVI) to train a Näive Bayes model that differentiated macroplastic plastic patches, from other floating materials (e.g. vegetation, driftwood). [23], [28] on its hand leveraged changes in the Red Edge bands to identify pixels with plastic litter mixed with vegetation.

The vast data produced by satellites and the complex task of discriminating macroplastic from other features in satellite imagery, created opportunities for applications based on ML techniques. ML and deep learning approaches have been recently applied for the classification of spectral and spatial anomalies in Sentinel-2 images. [29] developed two unsupervised (K-means and fuzzy c-means) and two supervised (support vector regression and semi-supervised fuzzy c-means) classification algorithms to identify floating plastics using artificial plastic target data. Using a similar training approach, Support Vector Machines (SVM) and Random Forest (RF) models were developed by [30] based on the spectra of plastic and spectral indexes, such as the kernel Normalized Difference Vegetation Index (kNDVI). [31] overcome the scarcity of ground-validated data by leveraging synthetic data from marine debris, plastic, and wood to train a RF classifier. [32] used eXtreme Gradient Boosting (XGBoost) and their spectral library composed of artificial targets and other ocean features from known events in the literature. A different ML approach was used by [33] who employed a Convolutional

Neural Networks (CNN) U-Net predictor to learn the spatial characteristics of floating objects, rather than using spectral information and distinguishing the different types of objects.

The recent publication of the open-source MARIDA (MARIne Debris Archive) spectral library dataset based on Sentinel-2 [34] created more solid opportunities for the application of ML techniques. Particularly suiting the complex task of discriminating plastic debris patches from other features in Sentinel-2 imagery, MARIDA library provided annotated spectral signatures of plastic debris patches from known pollution events, as well as of signatures from other features, such as floating macroalgae, foam, ships and clouds. [34] presented baseline approaches based on RF and a CNN U-Net architecture. [35] used MARIDA to develop a deep neural network (DNN) and uncover the spectral signatures of marine debris. [36] tested MARIDA with another approach named multi-feature pyramid network (MFPN). [37] were able to recently create semi-automated monitoring density maps of potentially plastic-polluted areas using a data pipeline and adaptation of a U-Net model trained with the MARIDA library.

While showing potential, the identification of floating macroplastics patches using satellite imagery, specifically Sentinel-2 images, is still subject to various physical (hazy atmosphere, cloud interference, sunglint, water turbidity, wind-induced sea roughness with whitecaps) and technical (radiometric, spatial, and temporal resolutions of satellite platforms and availability of ground-truth data) constraints ([4], [5]). The technical specifications firstly restrict the detection capability to macroplastics aggregations in the order of meters. For instance, it can identify filament-like structures , known as windrows ([24], [38]), or fronts where macroplastics converge and aggregate into patches ([9], [23], [26]). Secondly, Sentinel-2 has a 5-day revisit time and requires cloud-free conditions, which is suboptimal for collecting continuous data [5].

In addition to the inherent challenges associated with the sensors, the pre-processing methods also impact the ultimate detection results. Atmospheric disturbances such as clouds, aerosols, and water vapour, as well as sea-surface interferences like sunglint, are mitigated through atmospheric correction and pre-processing algorithms. Nevertheless, executing these procedures can be complex and computationally demanding, potentially introducing uncertainties regarding the spectral response of floating objects [5]. As an example, atmospheric correction of Top-Of-Atmosphere (TOA) images is desired to remove the effect of the atmosphere. ACOLITE [39] and Sen2Cor [40] are the common atmospheric correction tools for Sentinel-2 data. However, significant disparities have been observed when applying these tools to the same pixel [4]. Uncertainty can also occur when correcting for sunglint using NIR and SWIR bands, since these wavelengths are also employed for distinguishing floating plastic accumulations from water.

The presence of vessels and vessel wakes, clouds and cloud shadows, whitecaps, wave-breaking and land (e.g. intertidal rocks or islets) affect the accuracy performance as they also reflect in the NIR and could be falsely detected as a macroplastic patch. Therefore, a pre-screening and masking process could potentially identify and exclude pixels with such features that lead to a false detection. In addition, floating macroplastics are not the only floating aggregations found at sea. Pollen, scum-forming phytoplankton blooms, sea snot, foam, can also have similar spectral signatures as macroplastic patches ([5], [21], [34], [41], [42]). Furthermore, the mixing of a naturally occurring material like foam and floating vegetation with plastic debris, may resample the signature of the macroplastic patch. The feasibility of detecting floating plastic from space was also examined more in deep by other studies ([41], [43], [44]) that highlighted spectral distortions and misalignment occurring when resampling all the Sentinel-2 bands to the same resolution that determine a misleading spectral peak in the NIR band. As stressed by [5] further challenges hindering the development of

effective satellite detection methods, include the scarcity of ground-truth datasets resulting in unbalanced class distributions and few annotations made with a high degree of confidence.

The current technology is still in its early stages, and at this time, there has not been a satellite mission particularly designed or dedicated to this specific task. Sentinel-2 represents the state-of-the-art in the detection of suspected floating macroplastics patches, even though with the above-mentioned limitations. The use of VHR satellite imagery (e.g., WorldView-3, PlanetScope) has been advanced, as it offers improved spatial capabilities compared to Sentinel-2. However, these latter have too high associated costs to be used for monitoring extensive areas and regularly over time ([4], [5]). As a result of these challenges, there has been a growing recognition for a dedicated satellite constellation tailored for this purpose, which would optimize spectral bands, spatial resolution, and revisit frequencies for better monitoring of marine debris and floating plastics [45]. Feasibility studies, requirements and specifications for future satellite prototypes have been studied and defined in some initiatives such as OPTIMAL[1], RESMALI[2], and MARLISE[3].

3 DATA COLLECTION

The following section outlines the initial efforts regarding the gathering of data from various sources and the correspondent availability and suitability of satellite imagery, particularly Sentinel-2 imagery. It including compiling data for plastic debris events and other features based on previous related works (section 3.1.1), for artificially deployed targets of plastics (section 3.1.2), as well as for events not previously addressed in the literature of plastic satellite detection (section 3.1.3). Additionally, data from field campaigns conducted under the LabPlas project within the Field Sampling Work Package (WP2) are also presented (section 3.1.4). Following the processing of this data (section 4), the results of this exercise served as background for the analysis and extraction of a spectral signature of floating macroplastic patches and of other materials (section 5) and test the ML methodologies (section 6).

3.1 Collection of data for plastics debris and other events

3.1.1 Reported Cases in Literature

Initial efforts to gather satellite imagery (Figure 2) with occurrences of macroplastic patches and other features have been achieved based on information contained in published literature on satellite detection of macroplastic debris patches (Annex 9.1). [26] selected different locations on coastal waters, often associated to river plumes and fronts. From this work we acquired images for relevant events containing suspected and verified plastic patches, and other floating material, in imagery scenes from Ghana, Scotland, Tonga, and South Africa. A selection of images depicting marine debris were also collected by acquiring scenes in the Gulf of Honduras, a recognized area affected by plastic pollution, as documented by ground-truth data in [9], [34] and in the Haiti where also ground-truth events were available. Additional locations for reported litter windrows were collected in the coastal area of the south-east of the Bay of Biscay based on the study of [38]. Examples of images showing the presence of debris discharged to the sea after heavy rain events are taken from [23], [28], [30]. A selection of images in several coastal areas from [33] showing the presence of undefined floating objects is

[1] https://nebula.esa.int/content/optimal-optical-methods-marine-litter-detection
[2] https://nebula.esa.int/content/resmali-remote-sensing-marine-litter
[3] https://activities.esa.int/sites/default/files/2024-03/Session1_Livens_13.pdf

also taken into consideration to test remote sensing methodologies in our study. We also gathered an event of floating pumice [32] and an outbreak of sea snot [41].

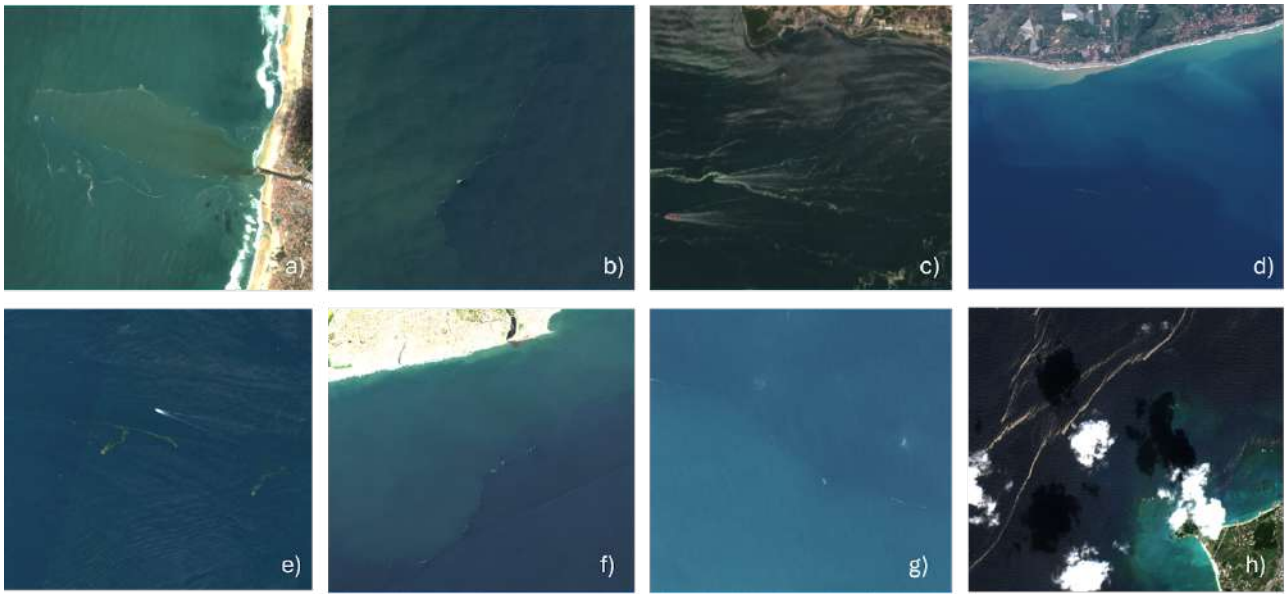

Figure 2 - Examples of Sentinel-2 satellite scenes of reported cases in the literature. a) Bay of Biscay, France (19-04-2018). b) Gulf of Honduras (18-09-2020). c) Marmara Sea, Turkey (27-05-2021). d) Calabria, Italy (22-10-2018). e) North Adriatic, Italy (16-08-2020). f) Accra, Ghana (31-10-2018). g) Da Nang, Vietnam (05-10-2018). h) Okinawa, Japan (26-10-2021).

3.1.2 Relevant image data of in-situ controlled experiment in literature

The Sentinel-2 data collection also included images (Annex 9.2) from controlled experiments (i.e. artificial targets) conducted by [20] in Cyprus. In these experiments, a 3x10 m target made of PET bottles (Figure 3) was deployed. Additionally, a series of Plastic Litter Projects (PLP), as detailed by [13], [21], [46] involved the construction and installation of various artificial targets with different compositions of plastic and organic material on the beaches of Mytilene, Greece (Figure 4). Given the scarcity of real-case scenarios depicting verified floating plastic debris, these images and standardized datasets were utilized for understanding the spectral behaviour of plastics and for testing and calibrating remote sensing algorithms and validating their outcomes.

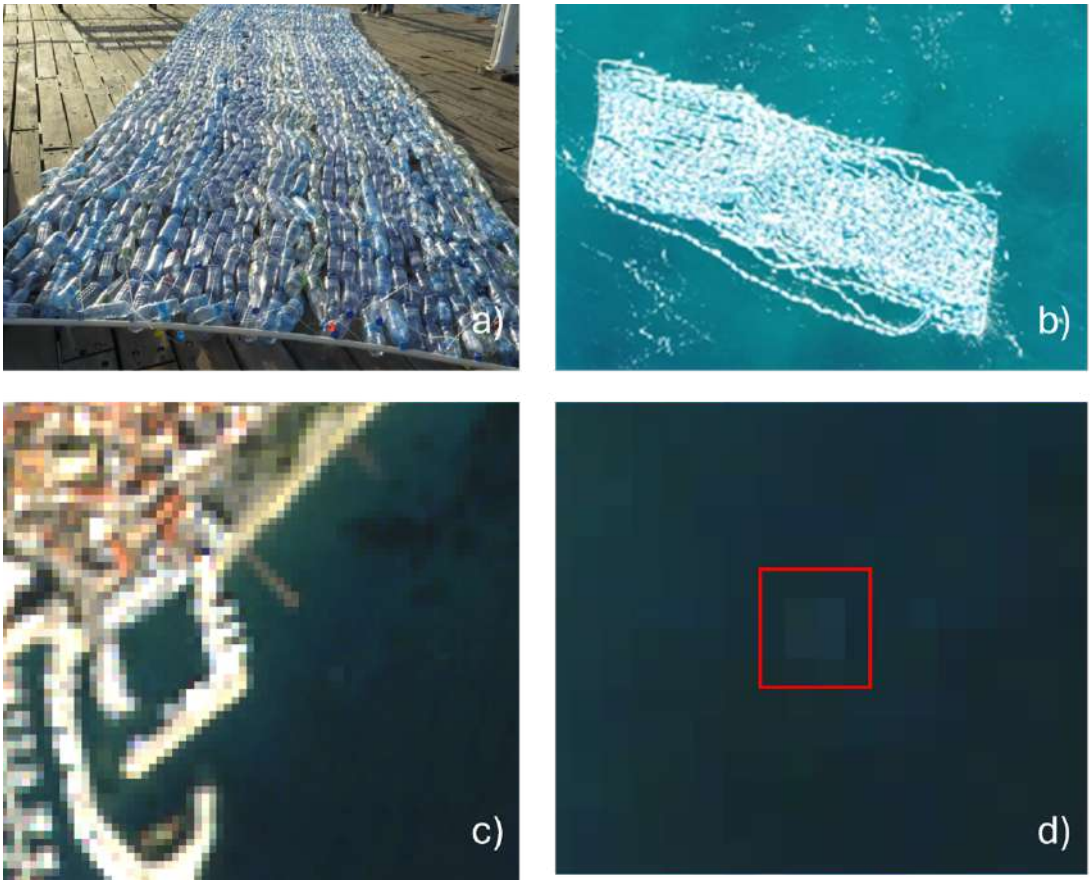

Figure 3 - a,b) 3x10 m target made of PET bottles. c,b) Sentinel-2 Level-1C image true color composite (15-12-2018) showing the pixels covered by the target (red square).

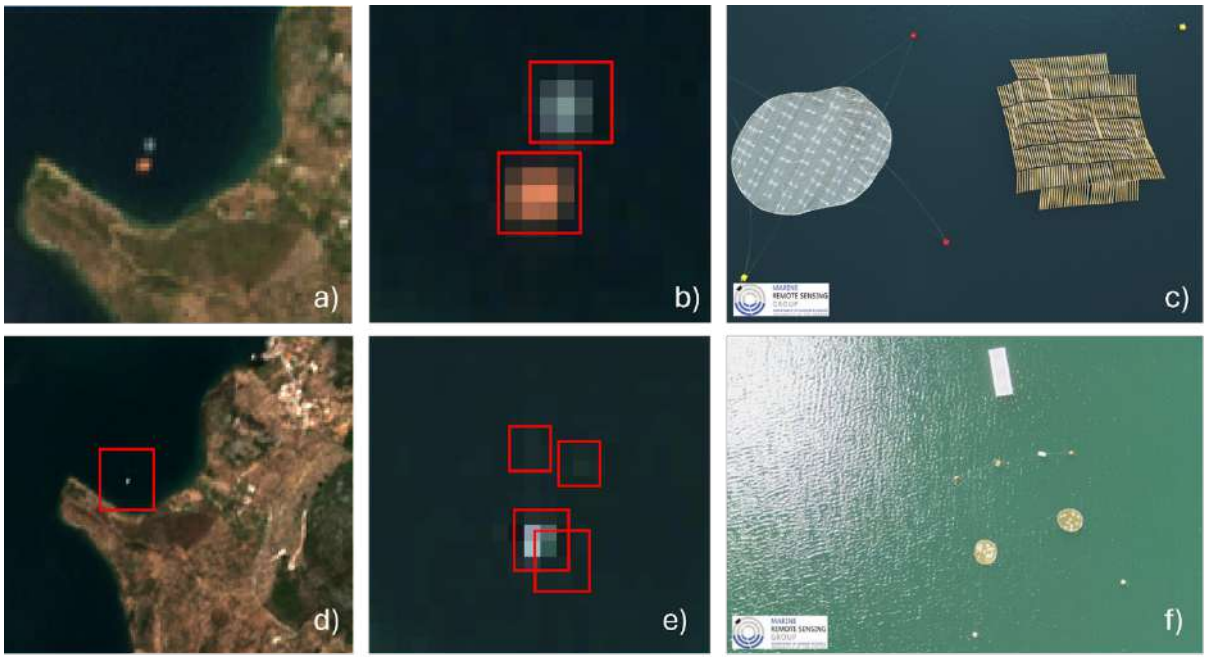

Figure 4 - Sentinel-2 Level-1C images true colour composite on 26-06-2021 (a,b) and 19-09-2022 (d,e) showing the pixels covered by the target (red square). c) 28 m diameter HDPE and 28 m wood/natural debris targets. f) 7 m diameter HDPE target + 5x15 m plastic PVC target.

3.1.3 Additional list of events

In addition to the existing collections of images featuring plastic debris that have been previously analysed in studies focused on detecting plastic debris using Sentinel-2 imagery, we have compiled further relevant events exemplified in Figure 5. This was achieved by consulting scientific reports, Ocean Clean-Up efforts, reviewing news articles and social media posts. Moreover, we relied on input from expert analysts researching possible known cases, in familiar locations, known polluted areas and in the aftermath of natural disasters. The detailed list, including dates, Sentinel-2 products, and reference sources, is provided in Annex 9.3.

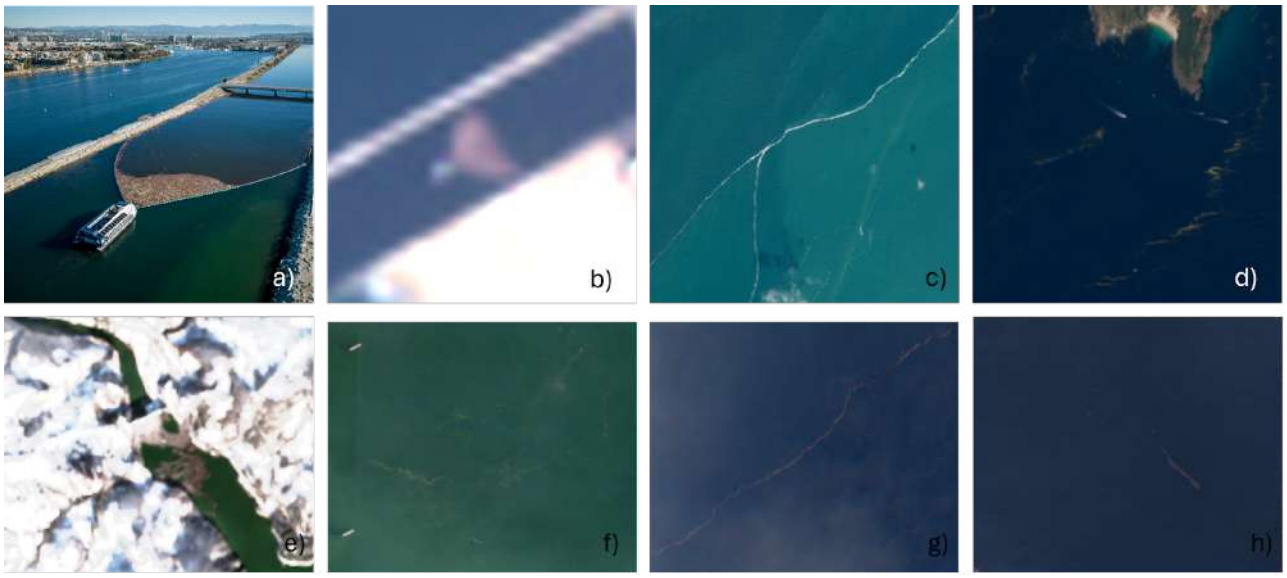

Figure 5 - a,b) Ocean Clean-Up Interceptor Los Angeles (November 2022) and Sentinel-2 Level-1C images true color composite on 09-22-2022 showing the collected debris. c) microalgae bloom (phaeocystis) in Zeebrugge, Belgium (05-01-2016). d) microalgae bloom (noctiluca) in Vigo, Spain (04-09-2021). e) mixed plastic debris at a dam in Potpec, Serbia (16-01-2021). f) suspected marine debris and vegetation in Manila, Phillipines (25-08-2022), g) suspected marine debris in Durban, South Africa (13-04-2022), h) suspected marine debris in Acapulco, Mexico (28-10-2023).

3.1.4 LabPlas field campaign

To enhance the dataset, efforts were made to obtain ground-truth data on macroplastics through collaborative efforts with Task 2.2 of WP2, and by gathering in-situ data during field campaigns conducted in the North Sea, Thames-Elbe rivers, and the Mero-Barcés river basin. To facilitate this process, a protocol document (Annex 9.4) was provided for reporting any observed sighting event.

Throughout the field campaigns in the North Sea (winter and summer), the AIR Centre also conducted satellite screening and analysis at selected locations (Figure 1-a) by acquiring Sentinel-2 images to detect floating ocean debris, along with VHR GEOSAT images (Annex 9.5) simultaneously and at the same location during in-situ data collection. The goal was to correlate in-situ results with satellite-based observations, test and refine detection algorithms, and validate the outputs of satellite-based detection. Additionally, conducting preliminary satellite screening and acquisition before the ship's passage through the established route provided an opportunity to guide field operations for route adjustment to target relevant ocean surface features (e.g., marine litter windrows or slicks).

During the winter campaign (4 February 2023 to 12 February 2023), no significant floating aggregations or detectable macroplastics were observed or reported during the fieldwork. The Sentinel-2 satellite screening revealed some suspected structures (Figure 6) but were not within the track of the cruise (Annex 9.3). In

addition, three VHR GEOSAT-2 images (Figure 7) were acquired for this campaign with the intention of including the vessel in the satellite FOV, but unfortunately the cloud conditions were unfavourable during the cruise.

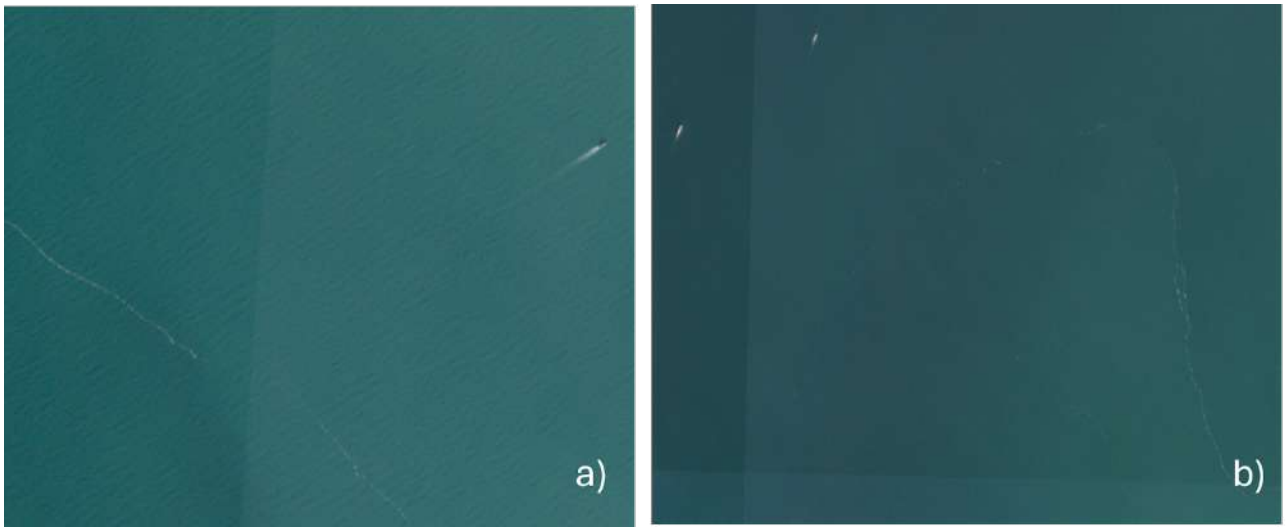

Figure 6 - Sentinel-2 acquired on 06-02-2023 (a) and on 07-02-2023 (b) showing floating structures.

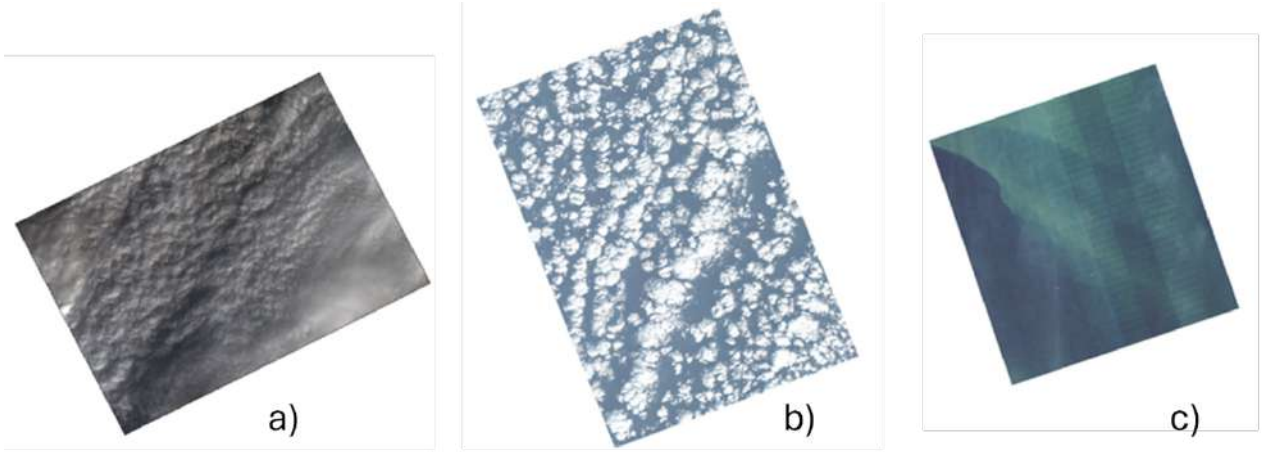

Figure 7 - GEOSAT-2 acquired images on 05-02-2023 (a), 07-02-2023 (b), 08-02-2023 (c).

During the summer field campaign in the North Sea (from 26 June 2023 to 1 July 2023) no major spotting of macroplastics occurred, however a major phytoplankton bloom of *Noctiluca scintillans* was observed both from satellite and on field near station 1 (Figure 8-a, b). Sentinel-2 (Figure 8-d) and GEOSAT-2 (Figure 8-c) data were acquired for this event (Annex 9.3 and Annex 9.5). The extraordinary event provides key information to better discriminate between phytoplankton blooms and other features in satellite images, such as plastic debris accumulations.

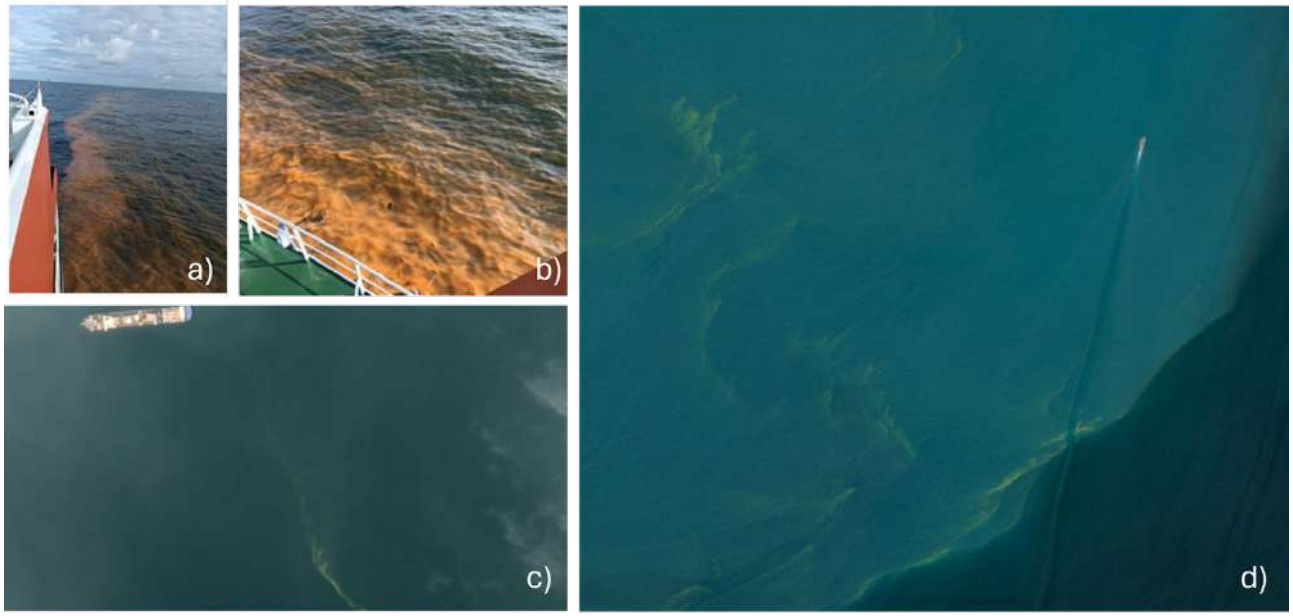

Figure 8 - Photographs taken by LabPlas team showing the Noctiluca scintillans bloom (a,b), GEOSAT-2 acquired images on 30-06-2023 (c), Sentinel-2 acquired images on 24-06-2023 (d).

3.2 EO data sources

The acquisition of satellite data during the LabPlas project was done manually through online platforms with graphical user interface and in an automated way through Application Programming Interfaces (API). Table III presents the satellite data products used during the project and the data providers.

Table III - Satellite products and data providers.

Satellite Product	Data provider
Sentinel-2 Level-1C and Level-2A	Google Earth Engine[4] SentinelHub EO Browser[5] Google Cloud[6] through FeLS[7] Copernicus Data Space Ecosystem[8] trough CDSETool[9]
GEOSAT-2 Level-1C PS3	GEOSAT[10]
WorldView-3 pansharpened RGB	MAXAR Open Data Program[11]

[4] https://earthengine.google.com/
[5] https://www.sentinel-hub.com/explore/eobrowser/
[6] https://cloud.google.com/storage/docs/public-datasets/sentinel-2
[7] https://github.com/vascobnunes/fetchLandsatSentinelFromGoogleCloud
[8] https://dataspace.copernicus.eu/)
[9] https://github.com/SDFIdk/CDSETool
[10] https://geosat.space/
[11] https://www.maxar.com/open-data

Initially, images from Copernicus Sentinel-2 Level-1C and Level-2A with low cloud cover percentage were manually identified for the several sample sites of the two LabPlas study cases. Since the Elbe and Thames river basins feeding into the North sea, and the NW Iberian Peninsula with a selected small Atlantic river basin constitute very extensive regions of interest (ROI), and considering the year 2021 as the analysis period, a script was created in Google Earth Engine (GEE) to facilitate the acquisition process. This script was later replaced by SentinelHub EO Browser platform for rapid verification of events identified in literature, media, and news.

To have more development freedom and save the data locally for future processing, we subsequently created a Python script that automates the acquisition of Sentinel-2 Level-1C products from two main data providers based on the selection of a ROI, sensing period, product name filters and credentials. From FeLS it is possible to access Google Cloud Sentinel data, while CDSETool allows to retrieve data from the new Copernicus Data Space Ecosystem (CDSE) after the ceasing of the Copernicus Open Access Hub (COAH) operations at the end of October 2023. This script outputs a list of available identifiers that are used to download the Sentinel-2 products automatically.

Regarding the VHR satellites, the GEOSAT-2 Level-1C PS3 data was obtained through a partnership based on credits, in which the AIR Centre made a tasking request, and the download was done through a Secure File Transfer Protocol (SFTP) account provided by GEOSAT. In the case of WorldView-3 pansharpened RGB, we used the MAXAR Open Data Program catalogue, which provides free data before and after disaster events. These VHR data were used to correlate in-situ results with satellite-based observations, test and refine detection algorithms, and validate the outputs of satellite-based detection of Sentinel-2 imagery.

4 Processing of EO Data

After identifying and downloading the satellite images, several pre-processing methods were developed related to cropping, atmospheric correction (AC), application of gain and offset values, masking, and reprojection of bands to the same coordinate reference system (CRS) and spatial resolution. After the pre-processing, spectral indices were calculated and information extracted to build a spectral library (section 5) and apply during the ML classification step (section 6).

Initially, Sentinel-2 Level-2A data was downloaded since it was already provided atmospherically corrected, however the operational algorithm for the correction is Sen2Cor[12] which provides Bottom-Of-Atmosphere (BOA) values optimized for terrestrial applications. Therefore, we changed our working focus to Sentinel-2 Level-1C data to explore algorithms dedicated to atmospheric correction in the ocean, namely ACOLITE[13] and Case 2 Regional CoastColour (C2RCC[14]). In order to be consistent with the recently published open-source Sentinel-2 spectral library focused on plastics (MARIDA) that was created using the ACOLITE processor to obtain Rayleigh-corrected reflectances, we decided to follow the same line to expand the library and not initiate a new one, thus adopted the ACOLITE processor.

ACOLITE is a generic processor for atmospheric correction and processing for coastal and inland water applications. It currently supports many sensors, including Sentinel-2 (A/B), and performs the AC by default

[12] Sen2Cor: https://step.esa.int/main/snap-supported-plugins/sen2cor/
[13] ACOLITE: https://odnature.naturalsciences.be/remsem/software-and-data/acolite
[14] C2RCC: https://c2rcc.org/documentation/

using the dark spectrum fitting approach but can be configured to use the exponential extrapolation. Features include the cropping of satellite images to rectangular regions of interest (defined by bounding coordinates), calculation of reflectance derived parameters, and results exported as GeoTIFF. The processor also takes the original UTM/WGS84 projection of Sentinel-2 tiles into consideration and resamples all spectral bands to 10 meters resolution while applying the necessary gain and offset values provided in the metadata following the processing baseline. The source code is available at GitHub[15], and an example between TOA and Rayleigh-corrected obtained with ACOLITE is shown in Figure 9.

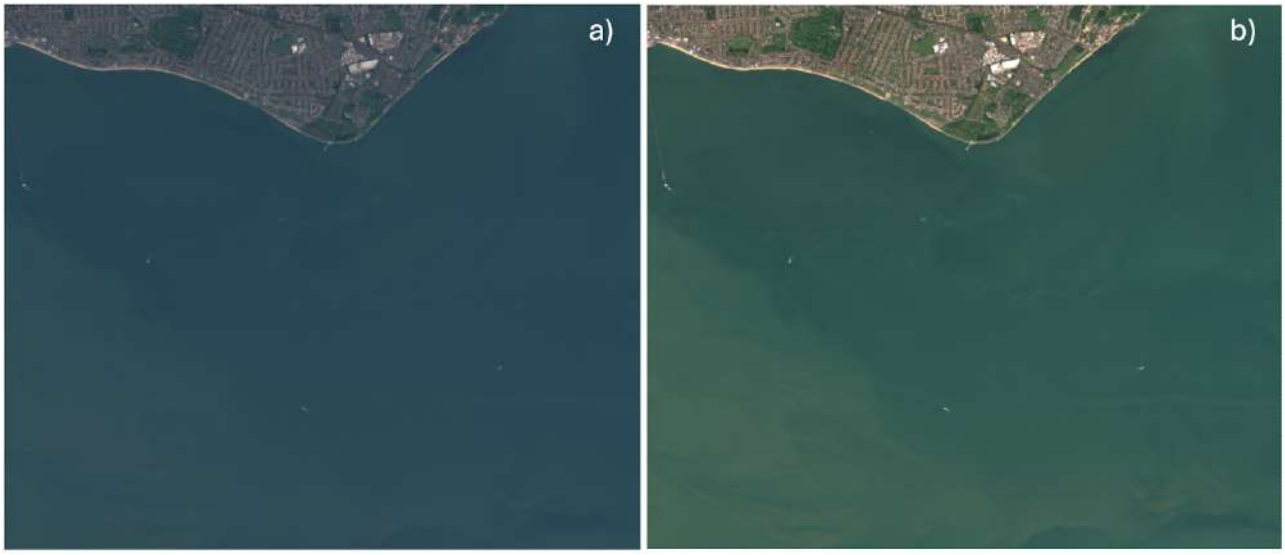

Figure 9 - Atmospheric correction applied to a LabPlas case study 1, site 6, scene in 22-07-2021. a) Sentinel-2 L1C TOA. b) Sentinel-2 ACOLITE Rayleigh-corrected.

Two main processing steps were developed after AC, the creation of a land mask and of a cloud mask. To create the land mask, we download the ESA Worldcover 2021[16] tiles for the given ROI. The product provides a global land cover at 10 meters resolution with eleven generic classes for 2021 based on Sentinel-1 and Sentinel-2 data (Figure 10), and it is delivered in a regular latitude/longitude grid (EPSG:4326) with the ellipsoid WGS 1984. The tiles need to be reprojected to the same CRS as the Sentinel-2 L1C products, merged, cropped, and calculated into a binary mask of only two classes (water and land).

[15] https://github.com/acolite/acolite
[16] https://worldcover2021.esa.int/

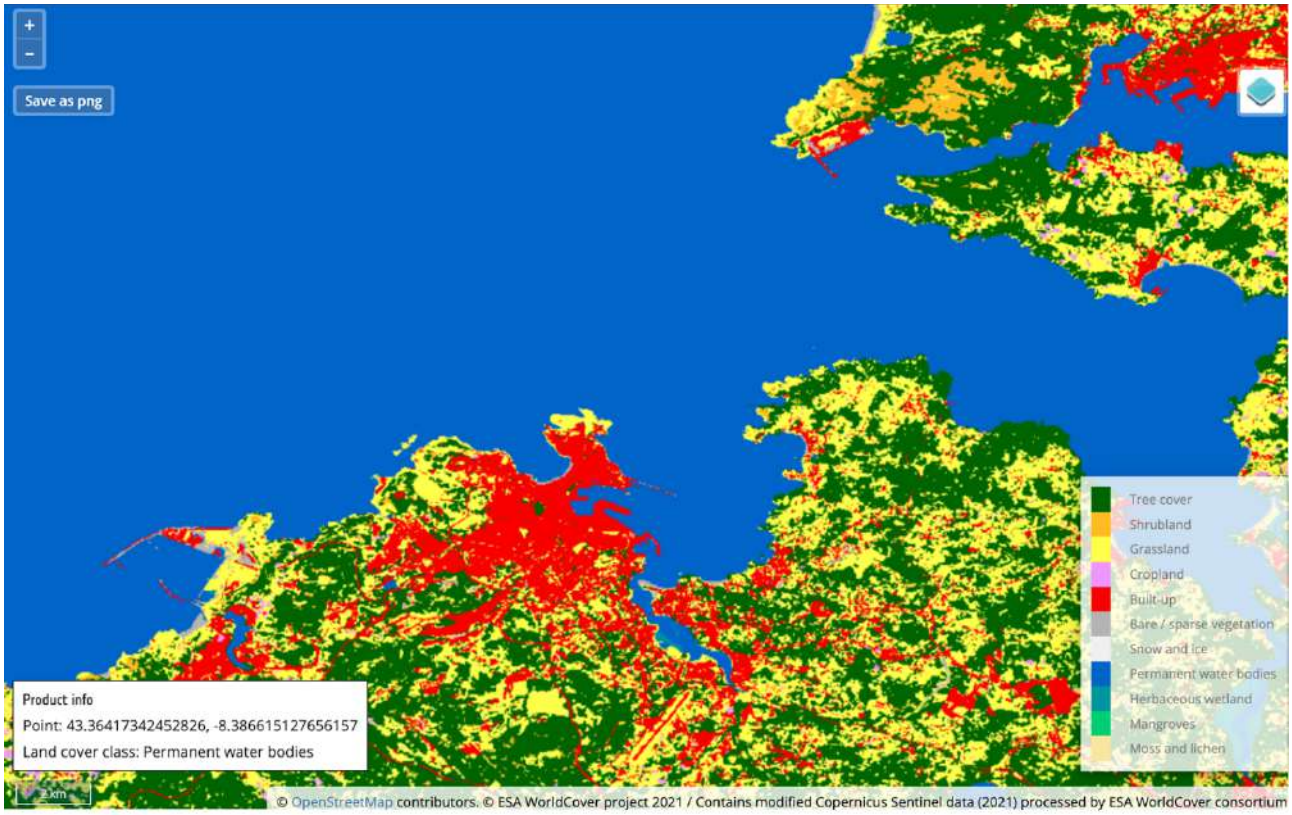

Figure 10 - ESA Worldcover 2021 visualization near LabPlas case study 2, site 6.

For the computing of a cloud mask it was chosen the Sentinel Hub's cloud detector for Sentinel-2 imagery algorithm, also known as s2cloudless[17]. The s2cloudless cloud masking pre-processing was configured to utilize 10 out of the 13 TOA spectral bands (B01, B02, B04, B05, B08, B8A, B09, B10, B11, B12) as input, resampled at a 10 m resolution. The algorithm produces a cloud probability map ranging from 0 to 1, which is then transformed into a cloud mask through thresholding. Additional morphological operations are applied during this transformation, including convolution of the probability map (averaging) and dilation of the binary cloud mask using a disk. To examine the sensitivity of these parameters beyond their default values, we conducted several tests (Figure 11) on various images containing clouds with diverse shapes and characteristics. We tested the sensitivity of the parameters (Figure 11-a,b,c) to ensure optimal masking of very small (Figure 11-d) and thin clouds (Figure 11-e), and cirrus clouds (Figure 11-f).

[17] https://github.com/sentinel-hub/sentinel2-cloud-detector

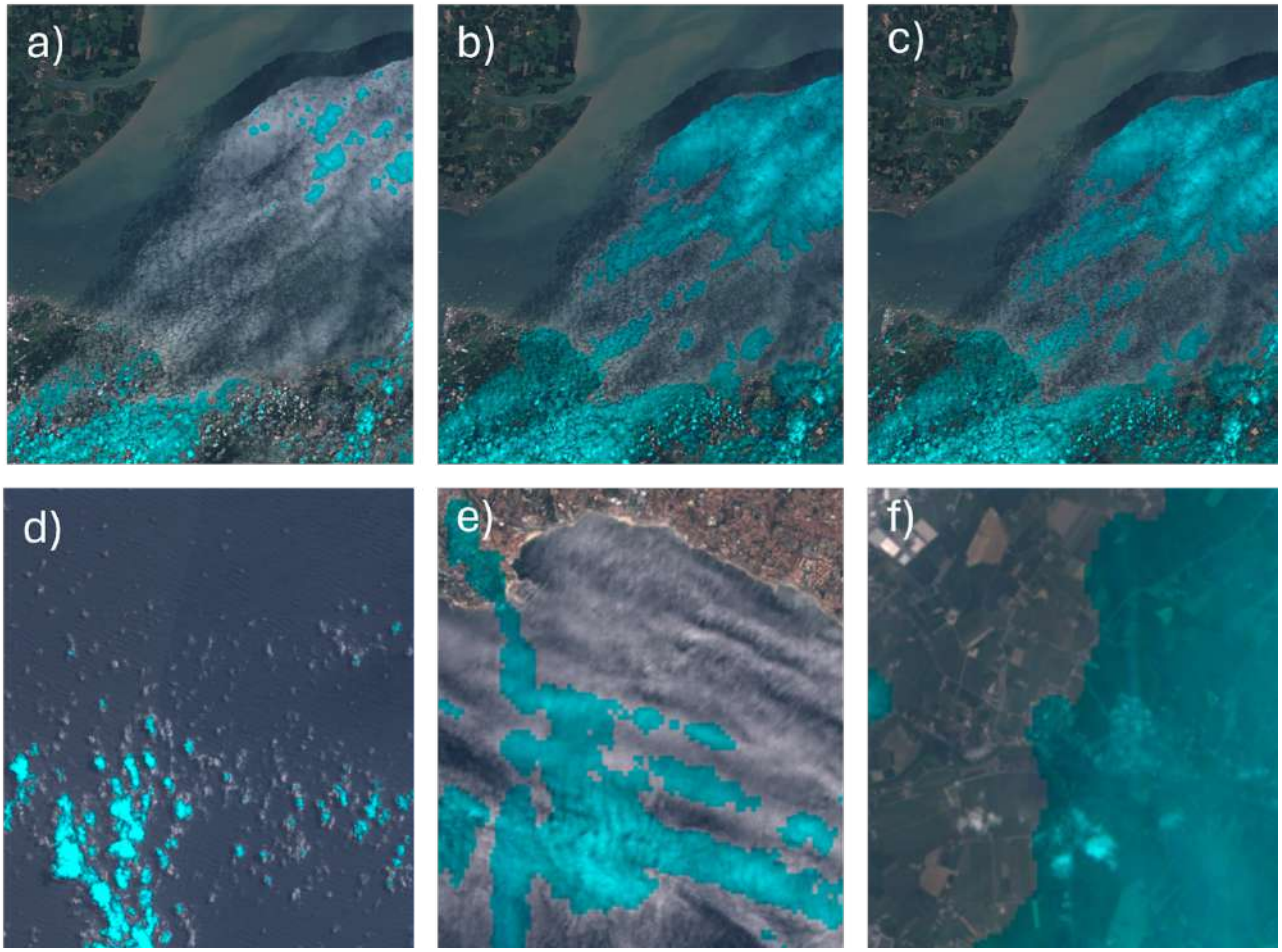

Figure 11 - Testing the performance and sensitivity of the cloud masking processor to different parameter settings. Threshold a) equal to 0.4, b) equal to 0.6, c) equal to 0.8 and different types of clouds (d,e,f). Light blue indicates the detected clouds.

Application of the land and cloud masks to the satellite imagery allow us to focus on the ocean area and reduce the computational load during the classification process. An example of both masks is shown in Figure 12.

Figure 12 - Masks applied to a LabPlas case study 1 scene in 22-07-2021. Land/water (land in brown and water in transparent) and cloud (clouds in purple) masks are overlapped to a RGB Sentinel-2 image.

Concerning the processing of VHR satellite imagery, no additional processing was required, as all acquired images were already processed to Level-2, including atmospheric correction and orthorectification, with the bands pansharpened to achieve the highest resolution.

5 Spectral Signature Analysis

The spectral signature analysis consisted of examining the variation of the reflectance wavelength for different floating materials and features; the unique signatures can change slightly under various conditions, including atmospheric, thus it is important to collect a considerable number of signatures of the same material, such that most of the variations are considered. Through this collection a library of spectral signatures was constructed for subsequent use in the ML detection and classification applications.

We started by exploring the literature to find existing Copernicus Sentinel-2 spectral libraries that could be later expanded by us with new materials. Although the focus was on collecting spectral signatures from patches containing plastic, we also considered signatures from other features, which allow us to obtain a more detailed classification and apply this research to other case studies and regions.

We performed several tests on Sentinel-2 products using manual spectral extraction techniques and evaluated their feasibility for our study case. We calculated plastic abundance using Linear Spectral Unmixing (LSU), which allowed us to obtain an estimate of the percentage of the pixel covered by plastic, thus providing a preliminary level of confidence during extraction. We also applied spectral indices to the satellite images, such as Floating Debris Index (FDI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), among others, which facilitated the distinction of floating plastic debris, floating vegetation, and the background water.

LSU is used to determine the relative abundance of materials that are depicted in multispectral imagery based on the spectral characteristics. The reflectance at each pixel of the image is assumed to be a linear combination of the reflectance of each material (or endmember) present within the pixel. The equation for linear spectral unmixing of a pixel containing plastic and water is the following [21]:

$$R_{mix}(\lambda) = f_p * R_p(\lambda) + \left(1 - f_p\right) * R_w(\lambda) \quad (1)$$

Where R_{mix} is the mixed (plastic and water) pixel reflectance for wavelength λ, R_p is the reflectance for the plastic target endmember for wavelength λ, R_w is the reflectance for the water target endmember for wavelength λ, and f_p is the plastic target endmember abundance fraction.

The LSU and endmembers selection were applied to the artificial plastic targets of the PLP2018 samples [13] (Figure 13) and other PLPs[18]. This technique showed relevant results to the detection and estimation of the abundance of plastics, however real-life scenarios often exhibit non-linear behaviour, and the selection of pure endmembers and the spectral variability of the signatures for the same material, can pose serious limitations.

[18] http://plp.aegean.gr/

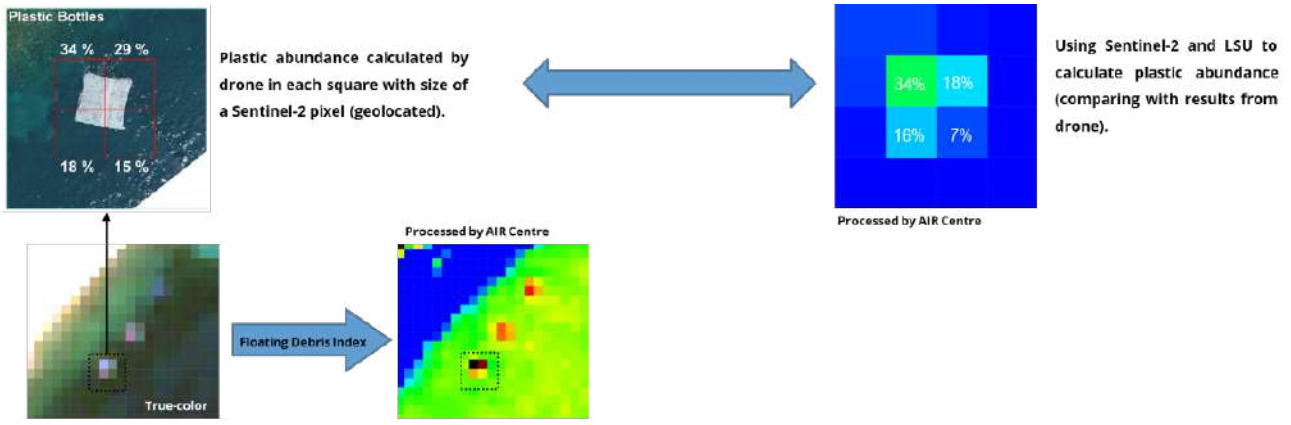

Figure 13 - Application of the LSU method and FDI index to a PLP2018 sample.

Spectral indices are mathematical expressions derived from spectral bands captured by remote sensing sensors. To help on the extraction of signatures we applied the following equations:

$$FDI = R_{NIR} - \left[R_{RE2} + \left(R_{SWIR1} - R_{RE2}\right) * \left(\left(\lambda_{NIR} - \lambda_{RE2}\right)/\left(\lambda_{SWIR1} - \lambda_{RE2}\right)\right) * 10\right] \quad (2)$$

$$NDVI = \left(R_{NIR} - R_{RED}\right)/\left(R_{NIR} + R_{RED}\right) \quad (3)$$

$$NDWI = \left(R_{GREEN} - R_{NIR}\right)/\left(R_{GREEN} + R_{NIR}\right) \quad (4)$$

Where R_{NIR}, R_{RE2}, R_{SWIR1}, R_{RED} and R_{GREEN} are the reflectance values for the bands near-infrared, red-edge 2, shortwave-infrared 1, red and green bands respectively, and λ is the central wavelength for those bands.

FDI was applied in PLP2018 as shown in Figure 13 being able to highlight the three deployed plastic targets. NDVI and FDI were calculated using a script developed in GEE for a Sentinel-2 image of the LabPlas case study 1, site 6, in 22-07-2021 (Figure 14) highlighting a suspected ocean feature of floating vegetation. NDVI, NDWI and FDI were also applied to other Sentinel-2 scenes of LabPlas case study sites after ACOLITE atmospheric correction (Figure 15).

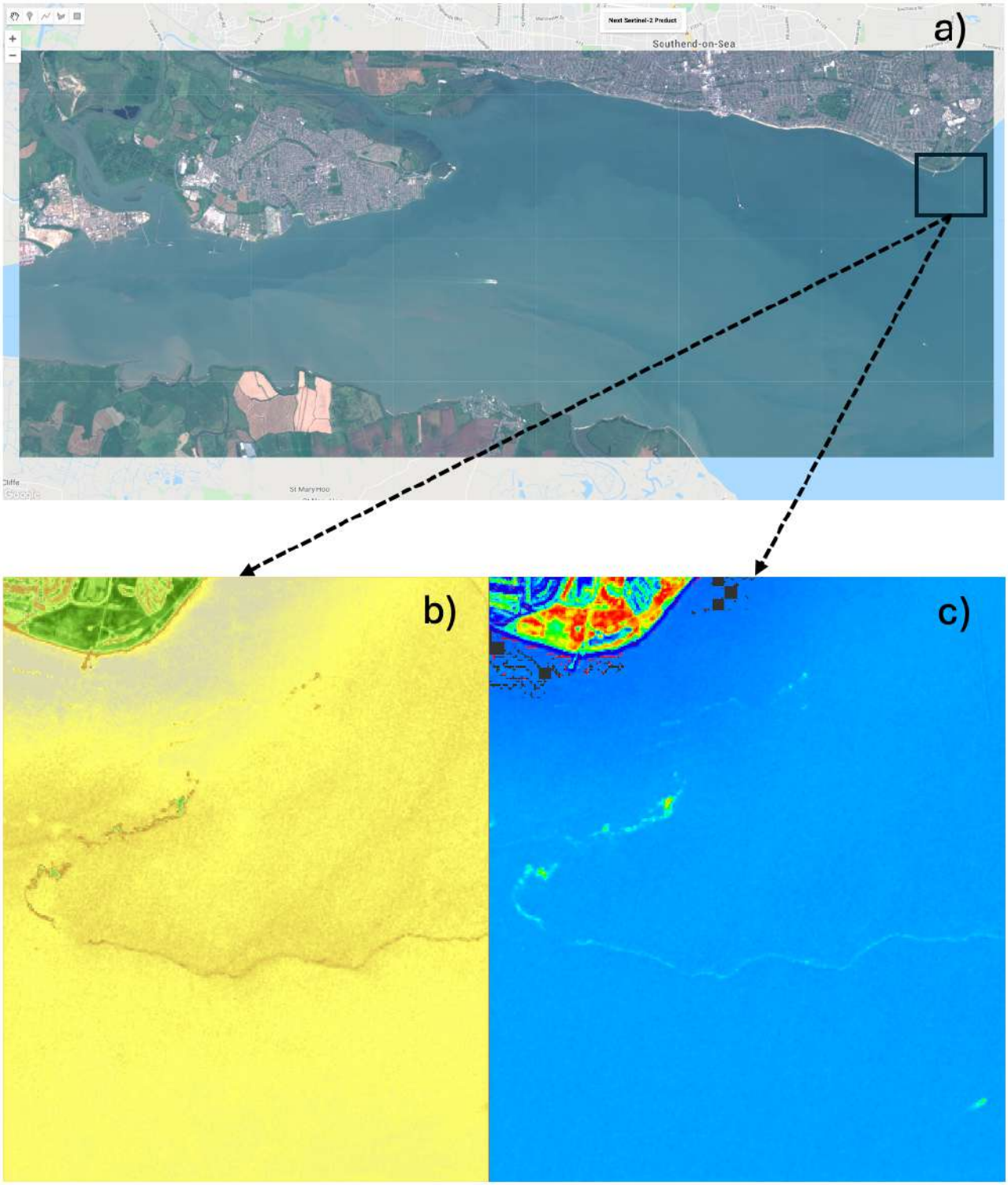

Figure 14 - GEE script for Sentinel-2 data acquisition and spectral indices calculation showcasing (a) true-color, (b) NDVI and (c) FDI images for LabPlas case study 1, site 6, in 22-07-2021.

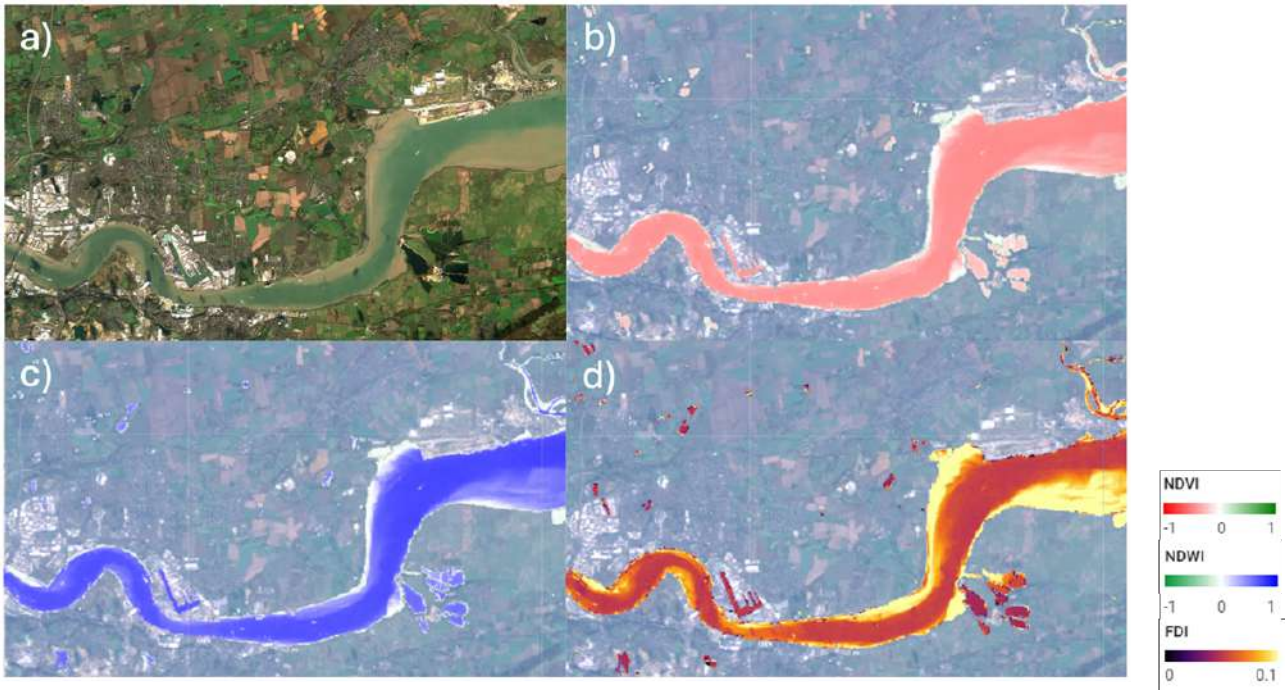

Figure 15 - Application of spectral indices after ACOLITE atmospheric correction to a Sentinel-2 scene of Thames river in 18-11-2018. (a) True-color, (b) NDVI, (c) NDWI and (d) FDI images.

We identified an existing state-of-the-art spectral library created specifically for Copernicus Sentinel-2 called MARIDA [34]. The dataset was developed through a meticulous process of marine debris events compilation covering a 2015 to 2021 period around several worldwide locations. This information was gathered based on reports from citizen scientists, the media, and scientific literature. The events of verified plastic debris are mostly from coastal regions affected by high plastic pollution, distributed over eleven countries (Honduras, Guatemala, Haiti, Santo Domingo, Vietnam, South Africa, Scotland, Indonesia, Philippines, South Korea, and China). Weather data, spectral indices calculation and image interpretation were used to select the Sentinel-2 images that contained marine debris and other relevant ocean features. The TOA reflectances of Level-1C were converted to Rayleigh-corrected reflectance using the ACOLITE atmospheric correction algorithm.

The selection and categorization of ocean features in MARIDA were made with polygons overlapping the Sentinel-2 images. This procedure resulted in a vector dataset of the digitized polygons, in shapefile format. The dataset was converted into a raster structure, which was finally cropped into non-overlapping 256x256 pixel-sized patches. Each patch includes a raster stack with 11 bands of Rayleigh-corrected reflectances, a mask with 15 thematic classes and a mask with three values of confidence. The raster stack and the class mask for each patch were prepared to be used with the U-Net model. Furthermore, the data from all patches were organized into a tabular structure to be used with RF.

Figure 16 illustrates the percentage of pixels for each class in the MARIDA library, for the training, validation, and testing. The dataset contains an imbalance in terms of the representativeness of each class. The majority of pixels were classified as Sediment-Laden Water, while Natural Organic Material represents a small percentage of the total pixels classified. Marine debris, defined as “floating plastics or other polymers, mixed with anthropogenic debris”, are also underrepresented. Each class bar represents the percentage of pixels

allocated to training (blue), validation (yellow) and test (green). The pink bars represent the amount of total pixels allocated for the training, validation and testing sub-datasets.

Class	Description	Pixels Percentage [%]
Marine Debris	Floating plastics or other polymers, mixed anthropogenic debris.	0,5%
		0,5%
MD		0,2%
Dense Sargassum	Dense floating *Sargassum* macroalgae.	0,2%
		0,5%
DenS		0,4%
Sparse Sargassum	Sparse floating Sargassum macroalgae.	0,3%
		0,2%
SpS		0,5%
Natural Organic Material	Vegetation and Wood.	0,2%
		0,0%
NatM		0,0%
Ship	Sailing and Anchored Vessels.	0,8%
		0,6%
Ship		0,6%
Clouds	Clouds including thin Clouds.	15,2%
		9,0%
Cloud		16,9%
Marine Water Super Class	Clear water, water near floating materials, waves (including waves from sailing vessels), wakes and cloud shadows.	23,5%
		11,4%
MWaterSC		15,7%
Sediment-Laden Water	High-Sediment river discharges with brown colour.	35,9%
		58,9%
SLWater		47,7%
Foam	Foam recorded at river fronts or coastal wave breaking area.	0,1%
		0,2%
Foam		0,2%
Turbid Water	Turbid waters close to coastal areas.	20,2%
		18,1%
Twater		16,5%
Shallow Water	Coastal waters, including coral reefs and submerged vegetation.	3,2%
		0,5%
Swater		1,3%
Pixels Number	Train	429412
	Validate	213102
	Test	194863
	Total	**837377**

Figure 16 - MARIDA spectral library summary. Classes full and short names, their description, and pixel distribution for train, validation, and test.

To the MARIDA dataset, we added spectral signatures extracted from various literature sources, but also through the analysis of satellite images and natural events reported in the media, totalizing 115010 extractions. All extractions are compiled into a single dataset, being validated with in-situ observations, and normalized into to the MARIDA format standardizations for consistency and replication. The added extractions included aquaculture structures from Portugal and China, plastic, and wood targets from the PLPs and natural events, reported phytoplankton blooms, natural slicks, and clouds, foam and different water types registered in the Azores, with some examples given in Figure 17.

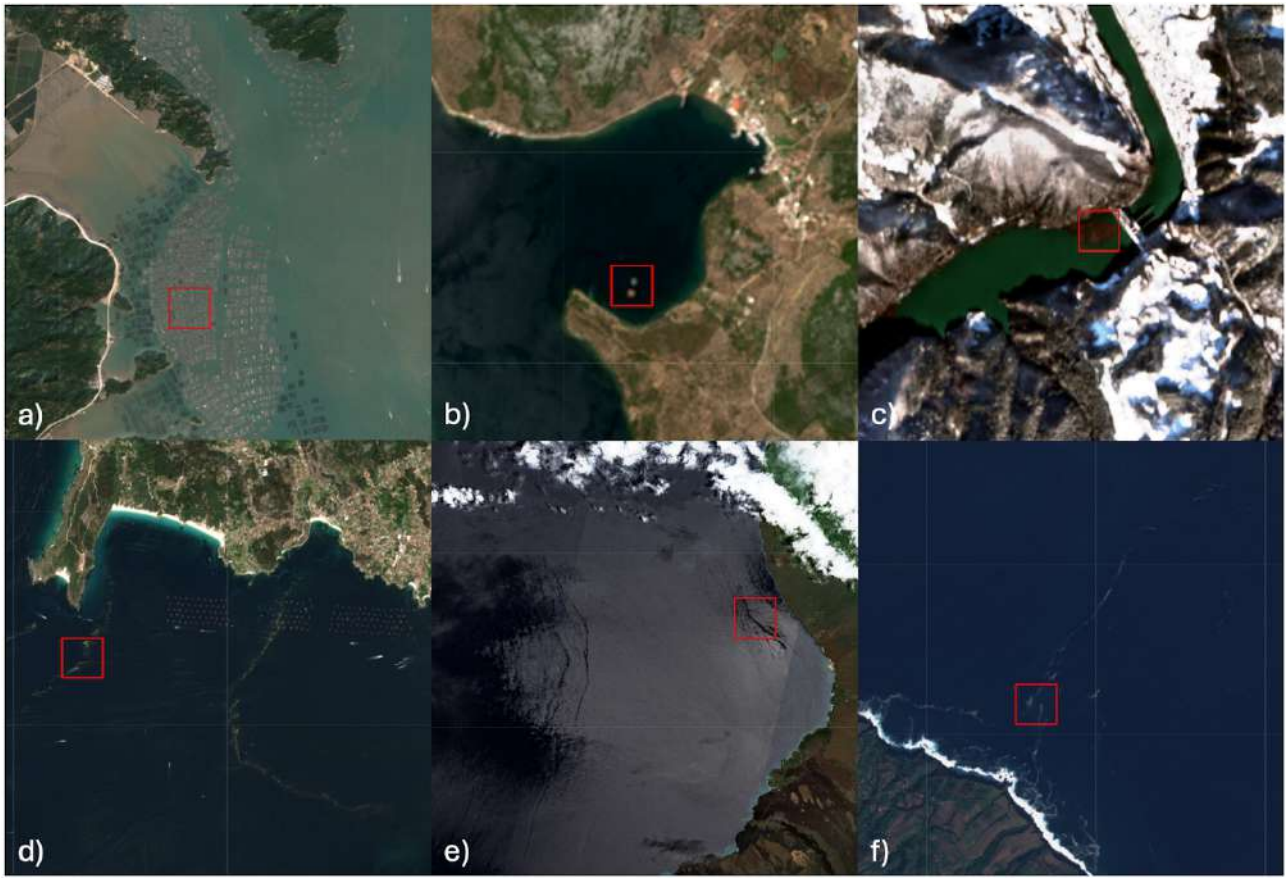

Figure 17 - Examples of Sentinel-2 imagery used on extraction of signatures. a) Aquaculture cages in China. b) Targets of PLP2021. c) Debris in a dam near Visegrad. d) Noctiluca in Vigo. e) Natural slicks in Hawaii. f) Foam filaments in Azores.

The mean Rayleigh Corrected spectra of ten signatures from the final spectral signature are shown in Figure 18-a. Only 10 classes are shown to facilitate the visualization. On Figure 18-b we focus on the related plastic debris signatures of the targets and natural events and compare with the marine water class.

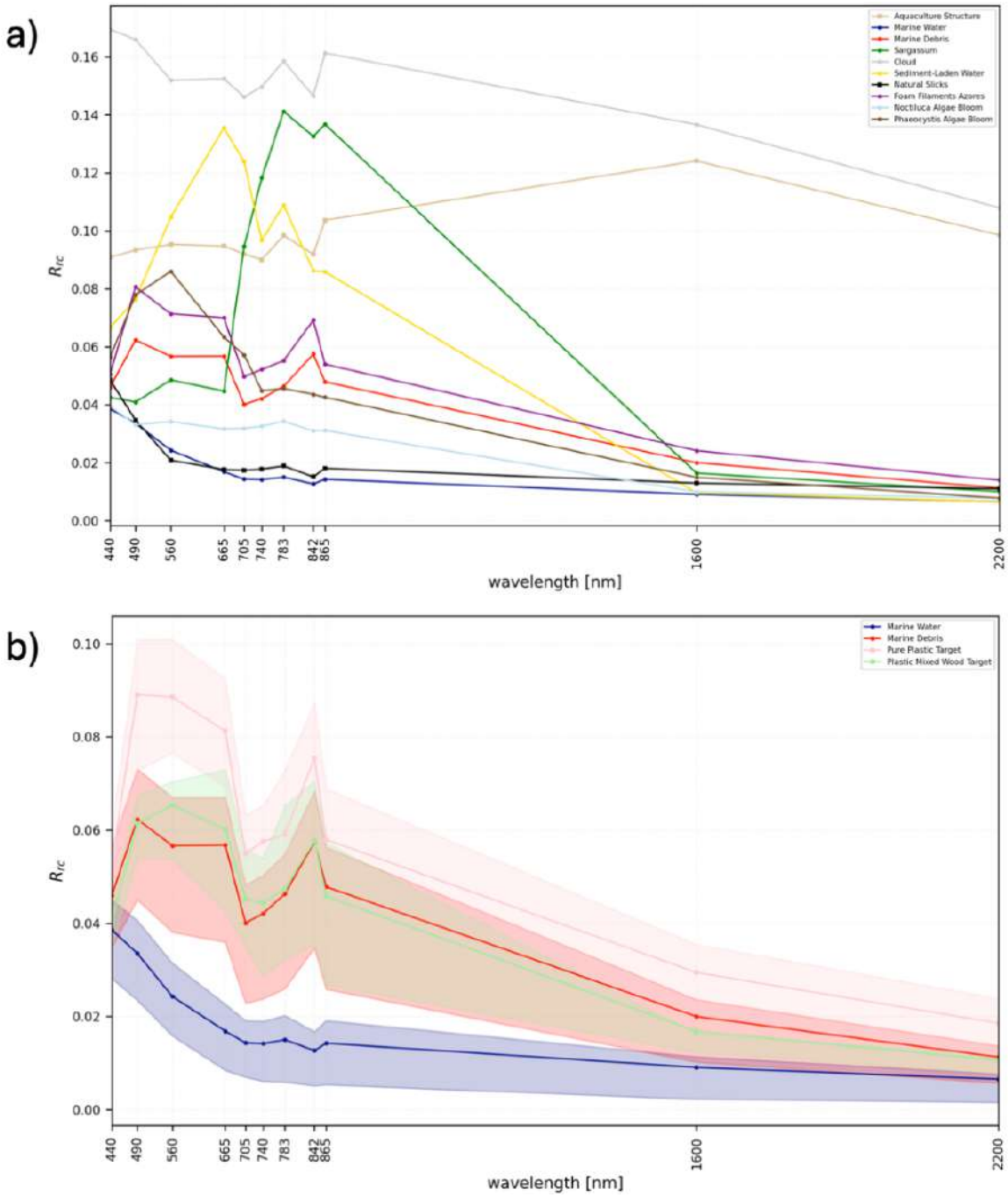

Figure 18 - Spectral signatures from the compiled dataset. a) Ten mean spectral signatures. b) Spectral comparison between Marine Debris from MARIDA and Plastic related targets from PLP using 25 and 75 percentiles together with mean.

We also applied a two principal component t-distributed Stochastic Neighborhood Embedding (t-SNE) analysis using Spectral Angle Mapping (SAM) as a distance metric to the spectral library to find possible clusters and facilitate the class reduction during the classification process. The analysis was done with a maximum of 100 pixels for each class of the dataset for computational demands. Figure 19 presents evident clusters regarding sediment-laden water (SLWater), floating vegetation such as Sargassum (DenS and SpS), aquaculture structures (AquSt, with difference between the ones in China and Portugal), among others. However, Marine

debris (MD) dispersion in the t-SNE analysis, shows that it could be confused with other materials, namely foams, such as the one from the filaments in the Azores (FilFoAz), due to spectral signature similarity, as also patent in Figure 18-a.

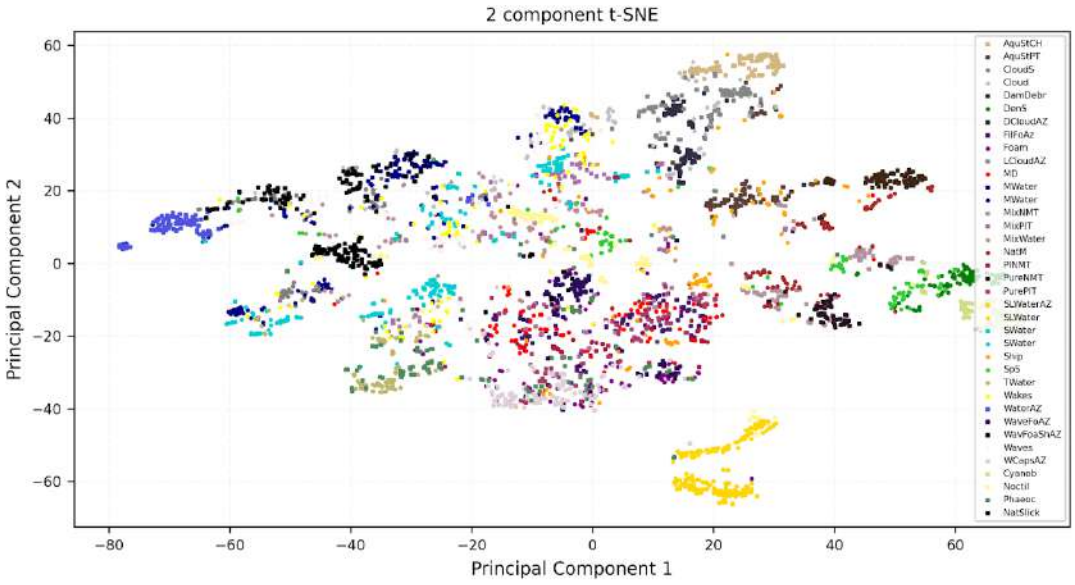

Figure 19 - t-SNE analysis of 100 pixels of each class from the compiled dataset. Squares correspond to MARIDA extractions and circles are AIR Centre.

A look into four main classes of the MARIDA dataset using the same t-SNE methodology is done in Figure 20, where is evident the separation between Sargassum and Marine Debris, but at the same time a spectral similarity between Natural Organic Material and Marine Debris, thus indicating overlapping optical signatures between these two classes.

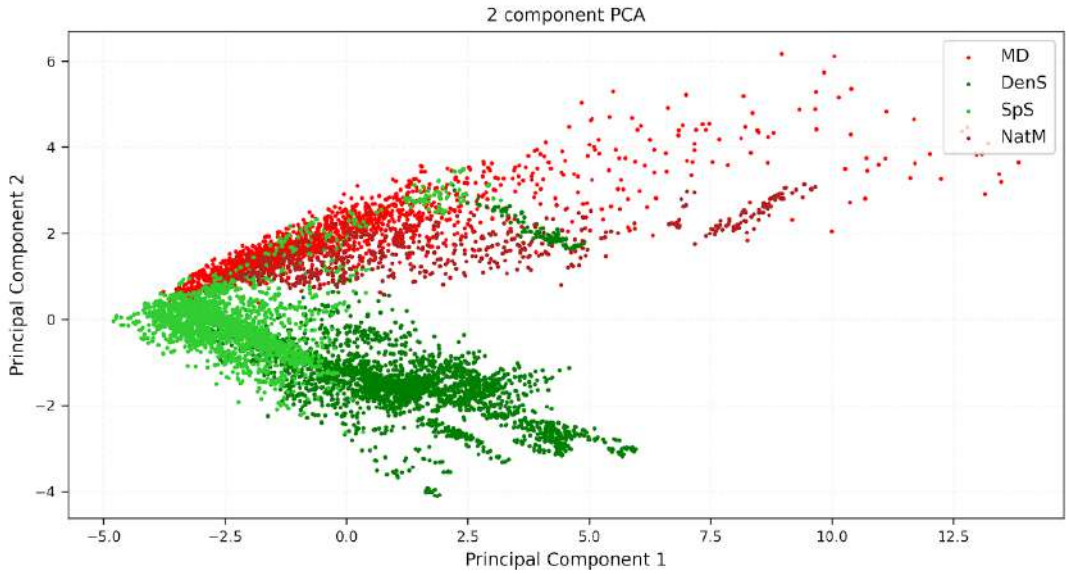

Figure 20 - t-SNE analysis of four main classes in MARIDA.

By using whisker plots with Sentinel-2 band 8 (B08) NIR reflectance and thresholds (Figure 21-a), we were able to identify some outliers in the MARIDA dataset. For example, although the pixel in Figure 21-b was labelled as marine water it presents a high value of B08 (higher than 0.4) indicating that the manual labelling might be incorrect, in fact a separated visual analysis revealed that the pixel corresponds to cloud. For an advanced quality control, these outliers must be removed from the final dataset, although they do not have a major impact during the ML process as they correspond to a very small percentage of the entire dataset. In the next section some of the classes of the compiled spectral library are used in the Machine Learning process, though a deeper focus on ML methods will be provided in deliverable D4.8.

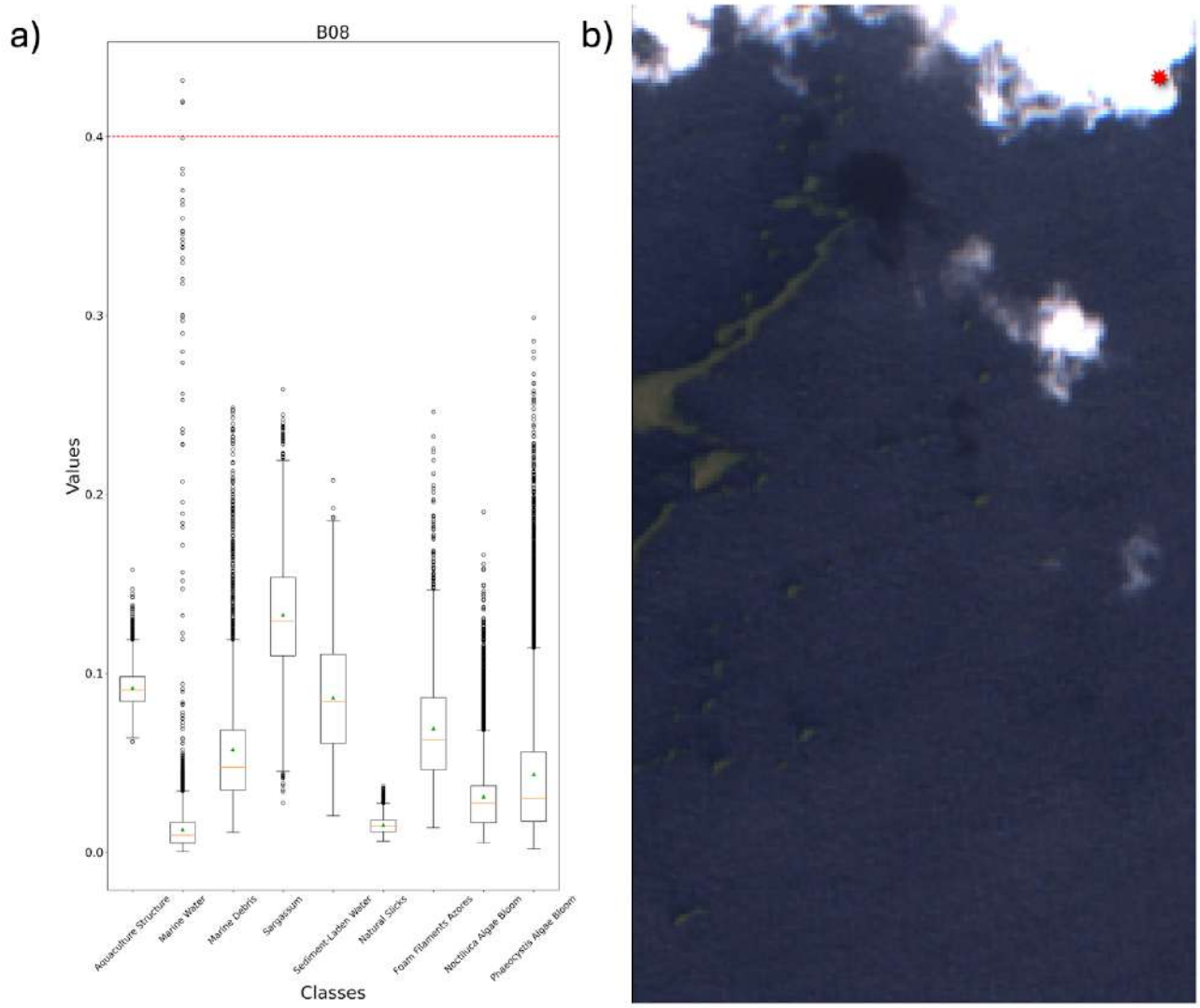

Figure 21 - a) Whisker plot for some of the dataset classes with B08 threshold of 0.4. b) Thresholds allow to find possible outliers, such as cloud and marine water (red star), on MARIDA's tile 16QED, image 1 of date 08-03-2018.

6 PLASTIC DETECTION WITH ML

Machine learning (ML) is a field of study in artificial intelligence dedicated to algorithms that can learn from data, identify patterns and perform decisions with reduced human intervention. Our research focused on three types of ML models, Random Forests (RF), eXtreme Gradient Boosting (XGBoost) and Convolutional Neural Networks (specifically, the U-Net).

RF supervised models have been widely used in various fields and applications due to their effectiveness in handling complex data and producing accurate predictions. They are a prime example of ensemble modelling, which involves generating multiple estimators during the modelling process and then combining their outputs to produce the final prediction [47]. RF is also a meta estimator rooted in decision trees, this means that the algorithm trains numerous decision trees on random subsets of both sample observations and predictive variables using techniques like bagging (bootstrap aggregation). The Decision Tree is a non-parametric

supervised ML algorithm that makes splits based on all input features, leading to subsequent splits guided by the homogeneity of resulting sub-nodes. Each decision tree is essentially a "weak learner" that might not perform exceptionally on its own, but their combination significantly improves overall predictive accuracy and helps prevent overfitting [48]. The final class prediction for each data point (in image analysis, a pixel) is determined by a majority vote among the predictions from all the individual trees [49].

XGBoost is a computationally efficient and highly effective implementation of the gradient boosting algorithm. Gradient Boosting stands as a tree-based ensemble ML algorithm that is trained using the boosting technique. Boosting is used to improve the predictive performance of models by sequentially training weak learners and giving more emphasis to misclassified data points in subsequent iterations. By iteratively correcting the errors made by previous models, boosting reaches a more accurate model [50].

XGBoost models are fit using a loss function and gradient descent optimization technique. Gradient descent is a core optimization technique used in ML that helps the model to learn and improve its performance over time by finding the optimal set of parameters that results in the lowest possible value of the loss function. This process iteratively computes the next point in the parameter space using the negative gradient's guidance at the current position. To prevent the models from overfitting, XGBoost also computes the second-order gradients of the loss function and can use both L1 and L2 regularizations to penalize the values of the weights [32].

The U-Net is a convolutional neural network whose architecture was initially proposed by [51] for biomedical image segmentation. The model receives input as three-dimensional arrays where two are spatial dimensions and the third is the number of bands (channels) in the image. The output of the U-Net is also a three-dimensional array with the same two spatial dimensions. The additional dimension corresponds to the number of class-labels used. The output represents the probability that a given pixel belongs to a certain class.

The purpose of the U-Net is to assign a class label (with respective probability) to every pixel of the input image. The U-Net architecture is similar to a convolutional auto-encoder. For auto-encoders, the spatial dimensions of the image are successively reduced via the application of convolutional layers (typically increasing the number of channels) followed by pooling layers (encoder part). From the compressed image representation, an image with the full spatial size is obtained successively upscaling the spatial dimensions via the application of convolution layers and interpolation layers (decoder part). The interpolation layers can also be replaced by transposed convolution operations.

To check the functioning and efficiency of the spectral library in validation targets and LabPlas regions of interest, we started with a preliminary exercise of training a RF model with only six classes of interest from MARIDA dataset (Marine Debris, Sargassum, Natural Organic Material, Marine Water Super Class, Sediment-Laden Water and Foam) and the fifteen most relevant ML features (all available bands, and NDVI, FDI, NDWI and NDMI). The Sargassum class includes Dense and Sparse Sargassum, the Marine Water Super Class includes Turbid, Shallow, Mixed and Marine Waters, in addition to Waves and Wakes. All other MARIDA classes were ignored in this preliminary analysis. The training of the XGBoost model, application of buffers and thresholds, and the final U-Net model will be described in deliverable D4.8.

The RF model was trained with the classifier ensemble.RandomForestClassifier from Scikit-learn's[19] (v1.1.0) open source ML library for Python (v3.9). Different versions of the Scikit-learn library may give slightly different classification results for some pixels. The split of the MARIDA dataset was done only for train and test with

[19] https://scikit-learn.org

percentages of 75% and 25%, respectively. Training parameters included 170 n_estimaters, 2 min_samples_split, 1 min_samples_leaf, max_depth of 25, class_weight as balanced_subsampled, bootstrap as True and random_state equal to 42, the other hyperparameters were kept at default.

The metrics for evaluating the performance of the model are characterized by the following equations:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \quad (7)$$

Where TP is true positive, TN is true negative, FP is false positive, and FN is false negative. The Precision measures how accurate the prediction is, this is, the percentage of correct predictions. The Recall measures how good the model finds all the positives. Finally, the F1-Score combines the precision and recall metrics into a single metric giving equal weight to both.

Table IV presents the metrics for each of the classes of the trained RF model. The RF model presents very good metrics (close to 1), indicating a successful detection of nearly all MD pixels for the test dataset. Figure 22-a shows the confusion matrix for the model with only a few false detections, especially between Natural Organic Material and Marine Water Super Class, and between Foam and Marine Water Super Class. Figure 22-b shows the ML features importance, highlighting band B03, and spectral indices NDVI and NDWI. With the increase in the number of used classes in future models, it is expected a reduction in these metrics. Hence, when applied the present model to images, it may present false detections mainly related to classes that were not included in this preliminary model.

Table IV - Metrics of the preliminary RF model.

Class	**Precision**	**Recall**	**F1-Score**
Marine Debris (MD)	0.982	0.977	0.980
Sargassum (Sarg)	0.997	0.985	0.991
Natural Organic Matter (NatM)	0.992	0.933	0.962
Marine Water Super Class (MWaterSC)	0.999	1.000	1.000
Sediment-Laden Water (SLWater)	1.000	1.000	1.000
Foam (Foam)	0.947	0.909	0.928
Macro Average	**0.986**	**0.967**	**0.977**

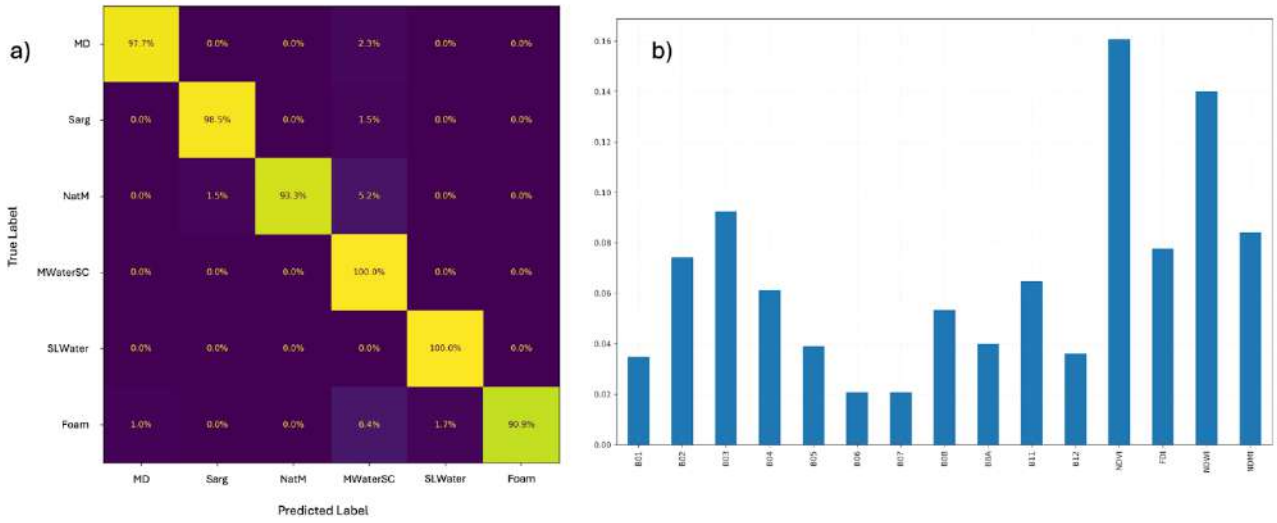

Figure 22 - Additional information about the preliminary RF model. a) Confusion matrix. b) Feature importance.

We give two examples of application of the trained RF. In Figure 23, the model is applied to PLP2021 containing two main artificial targets (one made of plastic and the other made of wood), thus acting as in-situ validation. In Figure 24, the model is used on one of the LabPlas sampling sites (6 - Thames estuary) from case study site 1, thus acting as a real case scenario applied to the project.

We see that in the case of PLP2021, the plastic target is successfully detected by the model, while the wooden target is confused with SLWater and Sargassum, with only two pixels detected as NatMat. A possible reason for this confusion is the similarity in the composition of both floating vegetation and wood, as well as the brownish colour of the water with high sediment amount and wood. Still in PLP2021, it is possible to observe the detection of two small targets and some structures near the coast.

In the case of the Thames estuary, the model detects a patch of Sargassum macroalgae which is possibly floating vegetation from the river. Some pixels seemed be detected near the coast as false detections of MD that correspond to human made structures that were not removed by the land mask. Pixels classified as MD close to the floating vegetation patch can however be real macroplastic patches, likely mixed with other material. These pixels classified as MD by the model are here considered as suspected marine plastic patches.

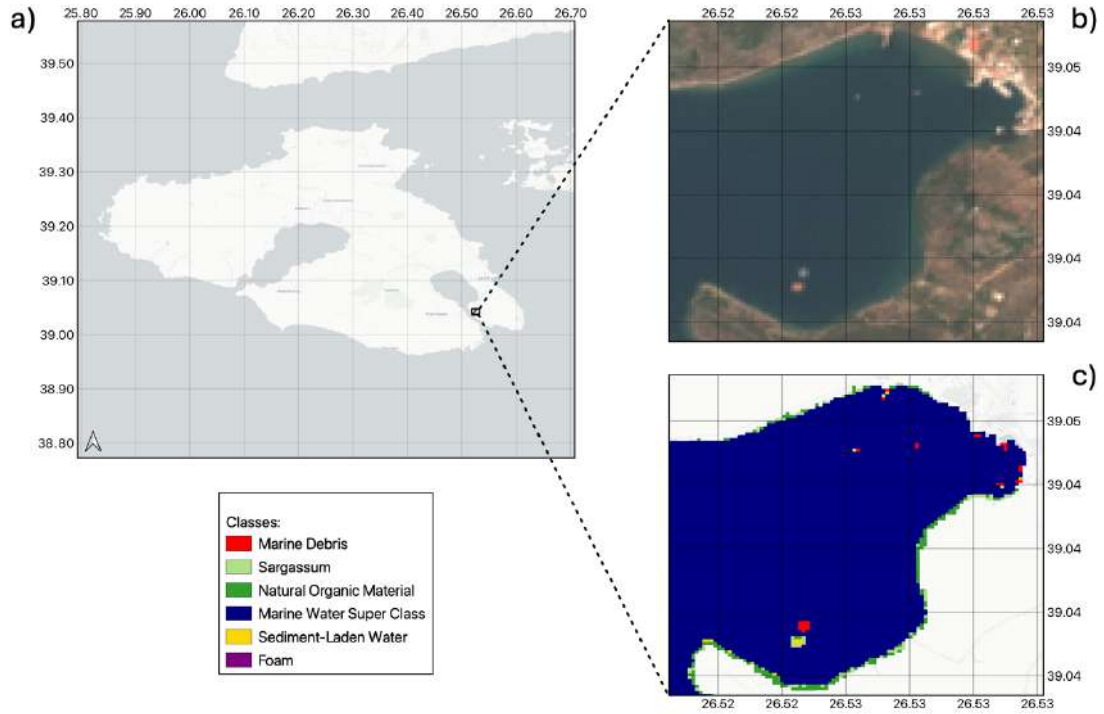

Figure 23 - Application of the preliminary RF model on the PLP2021 use case, 26-06-2021. a) Skala Loutron site in the Gulf of Gera, Island of Lesvos (Greece). b) RGB Sentinel-2 image after ACOLITE. c) RF model classification.

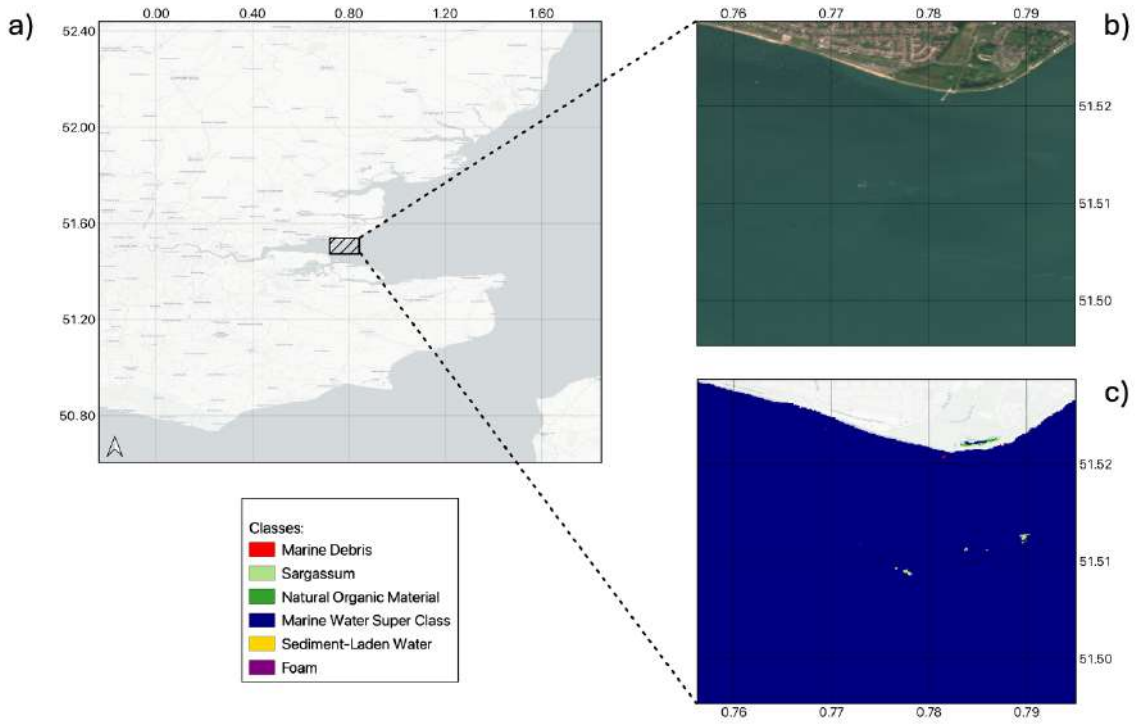

Figure 24 - Application of the preliminary RF model close to sampling site 6 of case study site 1, 22-07-2021. a) River Thames estuary, United Kingdom. b) RGB Sentinel-2 image after ACOLITE. c) RF model classification.

7 CONCLUSIONS

This report presented the preliminary developments and description of the research work conducted in the LabPlas project concerning advanced satellite techniques for determination of macroplastics.

The preliminary results of this study showed that the detection of macroplastic via satellite is feasible. Nevertheless, although there are a considerable number of recent studies on the use of EO satellite remote sensing combined with the latest advances in the area of Artificial Intelligence, there are still many technical limitations in the spatial, spectral, and temporal resolution of current EO technologies, freely-available satellite data, as well as challenges due to the complexity of the task itself due to overlapping optical signatures between different materials.

The research carried out within the framework of LabPlas has contributed to expanding the Sentinel-2 spectral library with new classes, as well as more signatures for the existing classes, however it is still a challenge to find events with in-situ observations that can be used as validation for the manual extraction of various features from the ocean surface.

More sophisticated ML models will be presented in deliverable D4.8, in which they were trained with more classes and for a longer period using higher computational power to obtain the best hyperparameters. Furthermore, it is envisaged to merge all methodological steps, from downloading to classification, in one unique tool, as well as applied in the LabPlas regions of interest.

A tool called POS2IDON is already being tested in the framework of LabPlas. The programming code has been made freely available in GitHub (https://github.com/AIRCentre/POS2IDON). As the tests and developments are concluded in the upcoming update of this deliverable (i.e., in D4.8) the source code will be provided in a suitable data repository (e.g. ZENODO) providing a Digital Object Identifier (DOI).

8 BIBLIOGRAPHY

[1] M. L. A. Kaandorp, D. Lobelle, C. Kehl, H. A. Dijkstra, and E. van Sebille, “Global mass of buoyant marine plastics dominated by large long-lived debris,” *Nat Geosci*, vol. 16, no. 8, pp. 689–694, Aug. 2023, doi: 10.1038/s41561-023-01216-0.

[2] G. G. N. Thushari and J. D. M. Senevirathna, “Plastic pollution in the marine environment,” *Heliyon*, vol. 6, no. 8, p. e04709, Aug. 2020, doi: 10.1016/j.heliyon.2020.e04709.

[3] V. Martínez-Vicente *et al.*, “Measuring Marine Plastic Debris from Space: Initial Assessment of Observation Requirements,” *Remote Sens (Basel)*, vol. 11, no. 20, p. 2443, Oct. 2019, doi: 10.3390/rs11202443.

[4] K. Topouzelis, D. Papageorgiou, G. Suaria, and S. Aliani, “Floating marine litter detection algorithms and techniques using optical remote sensing data: A review,” *Mar Pollut Bull*, vol. 170, p. 112675, Sep. 2021, doi: 10.1016/j.marpolbul.2021.112675.

[5] O. Karakuş, “On advances, challenges and potentials of remote sensing image analysis in marine debris and suspected plastics monitoring,” *Frontiers in Remote Sensing*, vol. 4, Nov. 2023, doi: 10.3389/frsen.2023.1302384.

[6] T. Aoyama, "Extraction of marine debris in the Sea of Japan using high-spatial-resolution satellite images," R. J. Frouin, S. C. Shenoi, and K. H. Rao, Eds., May 2016, p. 987817. doi: 10.1117/12.2220370.

[7] Y.-J. Park, S. P. Garaba, and B. Sainte-Rose, "Detecting the Great Pacific Garbage Patch floating plastic litter using WorldView-3 satellite imagery," *Opt Express*, vol. 29, no. 22, p. 35288, Oct. 2021, doi: 10.1364/OE.440380.

[8] M. Kremezi *et al.*, "Increasing the Sentinel-2 potential for marine plastic litter monitoring through image fusion techniques," *Mar Pollut Bull*, vol. 182, p. 113974, Sep. 2022, doi: 10.1016/j.marpolbul.2022.113974.

[9] A. Kikaki, K. Karantzalos, C. A. Power, and D. E. Raitsos, "Remotely Sensing the Source and Transport of Marine Plastic Debris in Bay Islands of Honduras (Caribbean Sea)," *Remote Sens (Basel)*, vol. 12, no. 11, p. 1727, May 2020, doi: 10.3390/rs12111727.

[10] D. Magyar, M. Cserép, Z. Vincellér, and A. D. Molnár, "Waste Detection and Change Analysis based on Multispectral Satellite Imagery," Mar. 2023.

[11] M. Kremezi *et al.*, "Pansharpening PRISMA Data for Marine Plastic Litter Detection Using Plastic Indexes," *IEEE Access*, vol. 9, pp. 61955–61971, 2021, doi: 10.1109/ACCESS.2021.3073903.

[12] N. Taggio *et al.*, "A Combination of Machine Learning Algorithms for Marine Plastic Litter Detection Exploiting Hyperspectral PRISMA Data," *Remote Sens (Basel)*, vol. 14, no. 15, p. 3606, Jul. 2022, doi: 10.3390/rs14153606.

[13] K. Topouzelis, A. Papakonstantinou, and S. P. Garaba, "Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018)," *International Journal of Applied Earth Observation and Geoinformation*, vol. 79, pp. 175–183, Jul. 2019, doi: 10.1016/j.jag.2019.03.011.

[14] S. Savastano, I. Cester, M. Perpinya, and L. Romero, "A First Approach to the Automatic Detection of Marine Litter in SAR Images Using Artificial Intelligence," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, Jul. 2021, pp. 8704–8707. doi: 10.1109/IGARSS47720.2021.9737038.

[15] S. Lavender, "Detection of Waste Plastics in the Environment: Application of Copernicus Earth Observation Data," *Remote Sens (Basel)*, vol. 14, no. 19, p. 4772, Sep. 2022, doi: 10.3390/rs14194772.

[16] À. Solé Gómez, L. Scandolo, and E. Eisemann, "A learning approach for river debris detection," *International Journal of Applied Earth Observation and Geoinformation*, vol. 107, p. 102682, Mar. 2022, doi: 10.1016/j.jag.2022.102682.

[17] A. Mohsen, T. Kiss, and F. Kovács, "Machine learning-based detection and mapping of riverine litter utilizing Sentinel-2 imagery," *Environmental Science and Pollution Research*, vol. 30, no. 25, pp. 67742–67757, Apr. 2023, doi: 10.1007/s11356-023-27068-0.

[18] A. D. Sakti *et al.*, "Identification of illegally dumped plastic waste in a highly polluted river in Indonesia using Sentinel-2 satellite imagery," *Sci Rep*, vol. 13, no. 1, p. 5039, Mar. 2023, doi: 10.1038/s41598-023-32087-5.

[19] K. Topouzelis, D. Papageorgiou, A. Karagaitanakis, A. Papakonstantinou, and M. Arias Ballesteros, "Remote Sensing of Sea Surface Artificial Floating Plastic Targets with Sentinel-2 and Unmanned Aerial

Systems (Plastic Litter Project 2019)," *Remote Sens (Basel)*, vol. 12, no. 12, p. 2013, Jun. 2020, doi: 10.3390/rs12122013.

[20] K. Themistocleous, C. Papoutsa, S. Michaelides, and D. Hadjimitsis, "Investigating Detection of Floating Plastic Litter from Space Using Sentinel-2 Imagery," *Remote Sens (Basel)*, vol. 12, no. 16, p. 2648, Aug. 2020, doi: 10.3390/rs12162648.

[21] D. Papageorgiou, K. Topouzelis, G. Suaria, S. Aliani, and P. Corradi, "Sentinel-2 Detection of Floating Marine Litter Targets with Partial Spectral Unmixing and Spectral Comparison with Other Floating Materials (Plastic Litter Project 2021)," *Remote Sens (Basel)*, vol. 14, no. 23, p. 5997, Nov. 2022, doi: 10.3390/rs14235997.

[22] S. P. Garaba, M. Arias, P. Corradi, T. Harmel, R. de Vries, and L. Lebreton, "Concentration, anisotropic and apparent colour effects on optical reflectance properties of virgin and ocean-harvested plastics," *J Hazard Mater*, vol. 406, p. 124290, Mar. 2021, doi: 10.1016/j.jhazmat.2020.124290.

[23] A. C. Ciappa, "Marine Litter Detection by Sentinel-2: A Case Study in North Adriatic (Summer 2020)," *Remote Sens (Basel)*, vol. 14, no. 10, p. 2409, May 2022, doi: 10.3390/rs14102409.

[24] M. Arias *et al.*, "Advances on Remote Sensing of Windrows as Proxies for Marine Litter Based on Sentinel-2/MSI Datasets," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, IEEE, Jul. 2021, pp. 1126–1129. doi: 10.1109/IGARSS47720.2021.9555139.

[25] K. Themistocleous, C. Papoutsa, S. Michaelides, and D. Hadjimitsis, "Investigating Detection of Floating Plastic Litter from Space Using Sentinel-2 Imagery," *Remote Sens (Basel)*, vol. 12, no. 16, p. 2648, Aug. 2020, doi: 10.3390/rs12162648.

[26] L. Biermann, D. Clewley, V. Martinez-Vicente, and K. Topouzelis, "Finding Plastic Patches in Coastal Waters using Optical Satellite Data," *Sci Rep*, vol. 10, no. 1, p. 5364, Dec. 2020, doi: 10.1038/s41598-020-62298-z.

[27] C. Hu, "A novel ocean color index to detect floating algae in the global oceans," *Remote Sens Environ*, vol. 113, no. 10, pp. 2118–2129, Oct. 2009, doi: 10.1016/j.rse.2009.05.012.

[28] A. C. Ciappa, "Marine plastic litter detection offshore Hawai'i by Sentinel-2," *Mar Pollut Bull*, vol. 168, p. 112457, Jul. 2021, doi: 10.1016/j.marpolbul.2021.112457.

[29] B. Basu, S. Sannigrahi, A. Sarkar Basu, and F. Pilla, "Development of Novel Classification Algorithms for Detection of Floating Plastic Debris in Coastal Waterbodies Using Multispectral Sentinel-2 Remote Sensing Imagery," *Remote Sens (Basel)*, vol. 13, no. 8, p. 1598, Apr. 2021, doi: 10.3390/rs13081598.

[30] S. Sannigrahi, B. Basu, A. S. Basu, and F. Pilla, "Development of automated marine floating plastic detection system using Sentinel-2 imagery and machine learning models," *Mar Pollut Bull*, vol. 178, p. 113527, May 2022, doi: 10.1016/j.marpolbul.2022.113527.

[31] M. Nagy, L. Istrate, M. Simtinică, S. Travadel, and P. Blanc, "Automatic Detection of Marine Litter: A General Framework to Leverage Synthetic Data," *Remote Sens (Basel)*, vol. 14, no. 23, p. 6102, Dec. 2022, doi: 10.3390/rs14236102.

[32] M. M. Duarte and L. Azevedo, "Automatic Detection and Identification of Floating Marine Debris Using Multispectral Satellite Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023, doi: 10.1109/TGRS.2023.3283607.

[33] J. Mifdal, N. Longépé, and M. Rußwurm, "TOWARDS DETECTING FLOATING OBJECTS ON A GLOBAL SCALE WITH LEARNED SPATIAL FEATURES USING SENTINEL 2," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. V-3–2021, pp. 285–293, Jun. 2021, doi: 10.5194/isprs-annals-V-3-2021-285-2021.

[34] K. Kikaki, I. Kakogeorgiou, P. Mikeli, D. E. Raitsos, and K. Karantzalos, "MARIDA: A benchmark for Marine Debris detection from Sentinel-2 remote sensing data," *PLoS One*, vol. 17, no. 1, p. e0262247, Jan. 2022, doi: 10.1371/journal.pone.0262247.

[35] M. Olyaei, A. Ebtehaj, and J. Hong, "Optical Detection of Marine Debris Using Deep Knockoff," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–12, 2022, doi: 10.1109/TGRS.2022.3228638.

[36] S. Gupta, Y. Khurana, J. Atrey, S. Gupta, and P. Krishnamoorthy, "Marine debris detection using a multi-feature pyramid network," *Remote Sensing Letters*, pp. 231–241, Apr. 2023, doi: 10.1080/2150704X.2023.2183480.

[37] H. Booth, W. Ma, and O. Karakuş, "High-precision density mapping of marine debris and floating plastics via satellite imagery," *Sci Rep*, vol. 13, no. 1, p. 6822, Apr. 2023, doi: 10.1038/s41598-023-33612-2.

[38] I. Ruiz *et al.*, "Litter Windrows in the South-East Coast of the Bay of Biscay: An Ocean Process Enabling Effective Active Fishing for Litter," *Front Mar Sci*, vol. 7, May 2020, doi: 10.3389/fmars.2020.00308.

[39] Q. Vanhellemont and K. Ruddick, "Atmospheric correction of metre-scale optical satellite data for inland and coastal water applications," *Remote Sens Environ*, vol. 216, pp. 586–597, Oct. 2018, doi: 10.1016/j.rse.2018.07.015.

[40] M. Main-Knorn, B. Pflug, J. Louis, V. Debaecker, U. Müller-Wilm, and F. Gascon, "Sen2Cor for Sentinel-2," in *Image and Signal Processing for Remote Sensing XXIII*, L. Bruzzone, F. Bovolo, and J. A. Benediktsson, Eds., SPIE, Oct. 2017, p. 3. doi: 10.1117/12.2278218.

[41] C. Hu, L. Qi, Y. Xie, S. Zhang, and B. B. Barnes, "Spectral characteristics of sea snot reflectance observed from satellites: Implications for remote sensing of marine debris," *Remote Sens Environ*, vol. 269, p. 112842, Feb. 2022, doi: 10.1016/j.rse.2021.112842.

[42] P. Mikeli, K. Kikaki, I. Kakogeorgiou, and K. Karantzalos, "HOW CHALLENGING IS THE DISCRIMINATION OF FLOATING MATERIALS ON THE SEA SURFACE USING HIGH RESOLUTION MULTISPECTRAL SATELLITE DATA?," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLIII-B3-2022, pp. 151–157, May 2022, doi: 10.5194/isprs-archives-XLIII-B3-2022-151-2022.

[43] C. Hu, "Remote detection of marine debris using satellite observations in the visible and near infrared spectral range: Challenges and potentials," *Remote Sens Environ*, vol. 259, p. 112414, Jun. 2021, doi: 10.1016/j.rse.2021.112414.

[44] C. Hu, "Remote detection of marine debris using Sentinel-2 imagery: A cautious note on spectral interpretations," *Mar Pollut Bull*, vol. 183, p. 114082, Oct. 2022, doi: 10.1016/j.marpolbul.2022.114082.

[45] V. Martinez-Vicente, "The need for a dedicated marine plastic litter satellite mission," *Nat Rev Earth Environ*, vol. 3, no. 11, pp. 728–729, Nov. 2022, doi: 10.1038/s43017-022-00360-2.

[46] K. Topouzelis, D. Papageorgiou, A. Karagaitanakis, A. Papakonstantinou, and M. Arias Ballesteros, "Remote Sensing of Sea Surface Artificial Floating Plastic Targets with Sentinel-2 and Unmanned Aerial Systems (Plastic Litter Project 2019)," *Remote Sens (Basel)*, vol. 12, no. 12, p. 2013, Jun. 2020, doi: 10.3390/rs12122013.

[47] L. Breiman, "Random Forests," *Mach Learn*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[48] A. Mohsen, T. Kiss, and F. Kovács, "Machine learning-based detection and mapping of riverine litter utilizing Sentinel-2 imagery," *Environmental Science and Pollution Research*, vol. 30, no. 25, pp. 67742–67757, Apr. 2023, doi: 10.1007/s11356-023-27068-0.

[49] J. V. Solórzano, J. F. Mas, Y. Gao, and J. A. Gallardo-Cruz, "Land Use Land Cover Classification with U-Net: Advantages of Combining Sentinel-1 and Sentinel-2 Imagery," *Remote Sens (Basel)*, vol. 13, no. 18, p. 3600, Sep. 2021, doi: 10.3390/rs13183600.

[50] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Mar. 2016, doi: 10.1145/2939672.2939785.

[51] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

9 ANNEXES

9.1 List of collected Sentinel-2 products from reported cases in literature.

Date	Location	Lat	Long	Floating Material	Sentinel-2 product	Reference
2018-10-31	Accra, Ghana	5.52	-0.1	macroalgae, sea foam, suspected plastic	S2B_MSIL1C_20181031T101139_N0206_R022_T30NZM_20181031T135633	Biermann et al. (2020)
2018-11-01	Accra, Ghana	5.52	-0.1	macroalgae, sea foam, suspected plastic	S2B_MSIL1C_20181031T101139_N0206_R022_T30NYM_20181031T135633	Biermann et al. (2020)
2019-08-16	Late Island, Tonga	-18.97	-175.58	pumice rock	S2A_MSIL1C_20190816T215921_N0208_R086_T01KFV_20190816T232135	Biermann et al. (2020)
2018-04-20	Scotland,UK	56.18	-2.48	macroalgae, sea foam, suspected plastic	S2A_MSIL1C_20180420T112121_N0206_R037_T30VWH_20180420T114417	Biermann et al. (2020)
2019-04-24	Durban, South Africa	-29.68	31.19	verified plastic	S2A_MSIL1C_20190424T101031_N0207_R022_T30NZM_20190424T170811	Biermann et al. (2020)
2015-11-29	Gulf of Honduras, Caribbean Sea	16.06	86.39	marine debris	S2A_MSIL1C_20151129T161622_N0204_R140_T16PEC_20151129T162644	Kikaki et al. (2020)
2016-09-04	Gulf of Honduras, Caribbean Sea	14.98	89.54	marine debris	S2A_MSIL1C_20160904T161342_N0204_R140_T16PCC_20160904T162740	Kikaki et al. (2020)
2020-09-18	Gulf of Honduras, Caribbean Sea	15.83	88.02	marine debris	S2B_MSIL1C_20200918T160839_N0209_R140_T16PCC_20200918T19453	Kikaki et al. (2022)
2020-09-18	Gulf of Honduras, Caribbean Sea	16.14	87.62	marine debris	S2B_MSIL1C_20200918T160839_N0500_R140_T16PDC_20230416T192938	Kikaki et al. (2022)
2020-09-23	Gulf of Honduras, Caribbean Sea	16.04	86.43	marine debris	S2A_MSIL1C_20200923T161011_N0500_R140_T16PEC_20230312T125328	Kikaki et al. (2022)
2020-09-23	Gulf of Honduras, Caribbean Sea	16.04	86.43	marine debris	S2A_MSIL1C_20200923T161011_N0500_R140_T16QED_20230312T125328	Kikaki et al. (2022)
2020-03-14	Port au Prince, Haiti	18.44	-72.75	marine debris	S2B_MSIL1C_20200314T152639_N0500_R025_T18QYF_20230412T160106	Kikaki et al. (2022)
2018-04-19	Bay of Biscay, France	43.49	-1.61	marine debris	S2B_MSIL1C_20180419T105619_N0206_R094_T30TXP_20180419T121006	Ruiz et al. (2020)
2018-04-19	Bay of Biscay, France	43.477	-2.75	marine debris	S2B_MSIL1C_20180419T105619_N0206_R094_T30TWP_20180419T121006	Ruiz et al. (2020)
2018-10-22	Calabria, Italy	37.93	15.7	marine debris	S2B_MSIL1C_20181022T094029_N0206_R036_T33SWB_20181022T133717	Sannigrahi et al. (2022)
2018-10-13	Beirut, Lebanon	33.92	35.59	marine debris	S2A_MSIL1C_20181013T081901_N0206_R121_T36SYC_20181013T094237	Sannigrahi et al. (2022)
2020-10-14	Hawai,USA	19.71	-154.85	marine debris and organic material	S2A_MSIL1C_20201014T205941_N0500_R014_T05QKC_20230325T021542	Ciappa (2021)
2020-08-16	North Adriatic, Italy	45.66	13.48	marine debris and organic material	S2A_MSIL1C_20200816T101031_N0209_R022_T33TUL_20200816T122432	Ciappa (2022)
2021-08-11	North Adriatic, Italy	44.87	12.9	marine debris and organic material	S2A_MSIL1C_20210811T101031_N0301_R022_T32TQR_20210811T143913	Ciappa (2022)

2020-11-15	Kolkata, India	20.71	88.43	floating objects	S2A_MSIL1C_20201115T044051_N0209_R033_T45QXD_20201115T055800	Mifdal et al. (2021)
2019-01-01	Lagos, Nigeria	6.31	3.71	floating objects	S2A_MSIL1C_20190101T100411_N0207_R122_T31NEH_20190101T115512	Mifdal et al. (2021)
2020-05-05	Lagos, Nigeria	6.04	2.93	floating objects	S2A_MSIL1C_20200505T100031_N0209_R122_T31NEG_20200505T115737	Mifdal et al. (2021)
2018-11-02	Long Xuyen, Vietnam	9.94	104.88	floating objects	S2A_MSIL1C_20181102T031901_N0206_R118_T48PVS_20181105T173148	Mifdal et al. (2021)
2018-03-14	Mandaluyong, Philippines	14.59	120.83	floating objects	S2B_MSIL1C_20180314T021559_N0206_R003_T51PTS_20180314T040537	Mifdal et al. (2021)
2019-04-25	Panama	8.61	-79.64	floating objects	S2B_MSIL1C_20190425T154559_N0207_R111_T17PPK_20190425T190518	Mifdal et al. (2021)
2018-09-28	Venice, Italy	45.62	13.35	floating objects	S2B_MSIL1C_20180928T100019_N0206_R122_T33TUL_20180928T151952	Mifdal et al. (2021)
2018-04-23	Vung Tau, Vietnam	9.43	106.69	floating objects	S2A_MSIL1C_20180423T030541_N0206_R075_T48PXR_20180423T080516	Mifdal et al. (2021)
2018-10-05	Da Nang, Vietnam	16.15	108.24	floating objects	S2B_MSIL1C_20181005T030549_N0206_R075_T48PZC_20181005T064938	Mifdal et al. (2021)
2021-10-26	Okinawa, Japan	26.72	127.95	pumice rock	S2A_MSIL1C_20211026T020801_N0500_R103_T52RCQ_20230103T132830	Duarte et al. (2023)
2021-05-27	Marmara Sea, Turkey	40.74	29.28	sea snot	S2B_MSIL1C_20210527T085559_N0300_R007_T35TPF_20210527T100606	Hu et al. (2022)

9.2 List of collected Sentinel-2 products from controlled experiments.

Date	Location	Lat	Long	Floating Material	Sentinel-2 product	Reference
2018-12-15	Limassol, Cyprus	34.63	33.011	3x10m target PET, + boats + fish farms	S2A_MSIL2A_20181215T083341_N0211_R021_T36SVD_20181215T094639	Themistocleous et al. (2020)
2018-06-07	Mytilene, Greece	39.1	26.56	10x10m targets (LDPE, PET, Nylon Fishing Net)	S2A_MSIL2A_20180607T085601_N0208_R007_T35SMD_20180607T114919	Topouzelis et al. (2019); PLP2018[20]
2019-04-18	Mytilene, Greece	39.1	26.56	5x20m unique target plastic PET + LDPE	S2B_MSIL2A_20190418T085559_N0211_R007_T35SMD_20190418T113005	Topouzelis et al. (2020); PLP2019[21]
2019-05-03	Mytilene, Greece	39.1	26.56	5x5m four targets , two PET + two LDPE, + 1x10m two targets PET+LDPE	S2A_MSIL2A_20190503T085601_N0211_R007_T35SMD_20190503T113356	Topouzelis et al. (2020); PLP2019[21]
2019-05-18	Mytilene, Greece	39.1	26.56	5x5m four targets , two PET + two LDPE, + 2x5m target PET	S2B_MSIL2A_20190518T085609_N0212_R007_T35SMD_20190518T115810	Topouzelis et al. (2020); PLP2019[21]
2019-05-28	Mytilene, Greece	39.1	26.56	5x10m two targets, one PET + one LDPE, + 2x5m target PET	S2B_MSIL2A_20190528T085609_N0212_R007_T35SMD_20190528T123615	Topouzelis et al. (2020); PLP2019[21]

[20] https://plp.aegean.gr/plastic-litter-project-2018/
[21] https://plp.aegean.gr/plastic-litter-project-2019/

2019-06-07	Mytilene, Greece	39.1	26.56	10x10m single mixed target PET+LDPE, 5x10m NatDeb target	S2B_MSIL2A_20190607T085609_N0212_R007_T35SMD_20190607T120913	Topouzelis et al. (2020); PLP2019[21]
2021-06-21	Mytilene, Greece	39.03	26.52	28m diameter HDPE & 28m NatDeb targets	S2A_MSIL2A_20210621T085601_N0300_R007_T35SMD_20210621T121311	Papageorgiou et al. (2022) ; PLP2021[22]
2021-08-10	Mytilene, Greece	39.03	26.52	28m diameter HDPE(+biofouling) & 28m NatDeb(+biofouling) targets	S2A_MSIL2A_20210810T085601_N0301_R007_T35SMD_20210810T120517	Papageorgiou et al. (2022) ; PLP2021[22]
2021-08-20	Mytilene, Greece	39.03	26.52	28m diameter HDPE(+biofouling+underwater) & 28m NatDeb(+biofouling) targets	S2A_MSIL2A_20210820T085601_N0301_R007_T35SMD_20210820T122214	Papageorgiou et al. (2022) ; PLP2021[22]
2021-09-04	Mytilene, Greece	39.03	26.52	28m diameter mixed target HDPE + NatDeb	S2B_MSIL1C_20210904T085549_N0301_R007_T35SMD_20210904T100110	Papageorgiou et al. (2022) ; PLP2021[22]
2021-10-04	Mytilene, Greece	39.03	26.52	28m diameter mixed target HDPE + NatDeb (submerged biofouled)	S2B_MSIL1C_20211004T085729_N0301_R007_T35SMD_20211004T095402	Papageorgiou et al. (2022) ; PLP2021[22]
2022-06-16	Mytilene, Greece	39.03	26.52	7m diameter HDPE target + 5x5m plastic PVC target	S2A_MSIL2A_20220616T085611_N0400_R007_T35SMD_20220616T152825	PLP2022[23]
2022-08-25	Mytilene, Greece	39.03	26.52	7m diameter HDPE target + 5x5m plastic PVC target	S2A_MSIL1C_20220825T085611_N0400_R007_T35SMD_20220825T141108	PLP2022[23]
2022-09-14	Mytilene, Greece	39.03	26.52	7m diameter HDPE target + 5x15m plastic PVC target	S2A_MSIL1C_20220914T085611_N0400_R007_T35SMD_20220914T105958	PLP2022[23]
2022-10-04	Mytilene, Greece	39.03	26.52	7m diameter HDPE target + 5x15m plastic PVC target	S2A_MSIL1C_20221004T085821_N0400_R007_T35SMD_20221004T110001	PLP2022[23]

9.3 List of collected Sentinel-2 products from additional events.

Date	Location	Lat	Long	Floating Material	Sentinel-2 product	Reference
2021-01-19	Visegrad, Bosnia	43.74	19.26	plastic bottles, rusty barrels, used tires, old furniture and other rubbish	S2B_MSIL1C_20210119T094229_N0209_R036_T34TCP_20210119T105934	https://apnews.com/article/environment-serbia-hydroelectric-power-95866b7e3af63b9608218e89791df5d0
2021-02-23	Visegrad, Bosnia	43.74	19.26	plastic bottles, rusty barrels, used tires, old furniture and other rubbish	S2A_MSIL1C_20210223T094031_N0209_R036_T34TCP_20210223T110016.SAFE	https://apnews.com/article/environment-serbia-hydroelectric-power-95866b7e3af63b9608218e89791df5d0
2021-02-25	Visegrad, Bosnia	43.74	19.26	plastic bottles, rusty barrels, used tires, old furniture and other rubbish	S2B_MSIL1C_20210225T093029_N0209_R136_T34TCP_20210225T105425.SAFE	https://apnews.com/article/environment-serbia-hydroelectric-power-

[22] https://plp.aegean.gr/plastic-litter-project-2021/
[23] https://plp.aegean.gr/plastic-litter-project-2022/

						95866b7e3af63b9608218e89791df5d0
2022-11-09	Los Angeles, USA	33.96	-118.45	wood, plastic, other rubbish	S2B_MSIL1C_20221109T183559_N0400_R027_T11SLT_20221109T202814.SAFE	https://theoceancleanup.com/media-gallery/interceptors-in-operation/#&gid=1&pid=1
2021-01-18	Svoge, Bulgaria	42.94	23.37	wood, plastic, other rubbish, snow	S2A_MSIL2A_20210118T092321_N0500_R093_T34TFN_20230527T115723	https://news.cgtn.com/news/2021-01-15/Heavy-rains-turn-Bulgaria-s-longest-river-into-garbage-river--X4seceiRR6/index.html
2021-01-16	Potpec, Serbia	43.52	19.57	wood, plastic, other rubbish, snow	S2B_MSIL1C_20210116T093239_N0500_R136_T34TCP_20230611T030950	https://www.dailymail.co.uk/news/article-9202723/Workers-start-clearing-plastic-junk-Serbias-lake-trash.html
2017-08-29	La Gomera, Spain	28	-17.24	microalgae bloom	S2A_MSIL1C_20170829T115221_N0205_R123_T28RBS_20170829T115218	Benavides & Arístegui (2020)
2017-08-09	Las Palmas, Canary Islands	27.82	-15.79	microalgae bloom	S2A_MSIL1C_20170809T115221_N0205_R123_T28RDR_20170809T115220	Benavides & Arístegui (2020)
2015-08-07	Baltic Sea	59.05	21.69	microalgae bloom (cyanobacteria)	S2A_MSIL1C_20150807T095006_N0204_R079_T34VEL_20150807T095007	https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Sentinel-2_catches_eye_of_algal_storm
2019-07-20	Baltic Sea	59.05	21.69	microalgae bloom (cyanobacteria)	S2A_MSIL1C_20190720T100031_N0208_R122_T34VEL_20190720T134457	https://www.esa.int/ESA_Multimedia/Images/2019/12/Baltic_blooms
2020-05-15	Zeebrugge, Belgium	51.31	2.63	microalgae bloom (phaeocystis)	S2B_MSIL1C_20200515T104619_N0209_R051_T31UDS_20200515T132523	https://marine.copernicus.eu/news/touring-europes-coasts-belgium-and-foaming-algae
2016-05-01	Zeebrugge, Belgium	51.41	3.23	microalgae bloom (phaeocystis)	S2A_MSIL1C_20160501T105032_N0201_R051_T31UES_20160501T105310	https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-2/Near-shore_phytoplankton_bloom_captured_from_space
2021-04-18	Gabriola Island, Canada	49.17	-123.77	microalgae bloom (noctiluca)	S2A_MSIL1C_20210418T191911_N0500_R099_T10UDV_20230511T040158	https://twitter.com/UVicSpectral/status/1385303463282692100
2021-09-04	Vigo, Spain	42.23	-8.86	microalgae bloom (noctiluca)	S2A_MSIL1C_20210904T113321_N0500_R080_T29TNG_20230117T101616	Detoni et al. (2023)
2017-09-25	Vigo, Spain	42.2	-8.85	microalgae bloom (noctiluca)	S2A_MSIL1C_20170925T113321_N0205_R080_T29TNG_20170925T113818	Detoni et al. (2023)

2022-04-13	Durban, South Africa	-29.65	31.46	suspected marine debris	S2A_MSIL1C_20220413T073621_N0400_R092_T36JUN_20220413T094538.SAFE	https://www.dailymaverick.co.za/article/2022-06-14-huge-volume-of-plastic-flood-debris-trapped-inshore-by-agulhas-current/
2022-04-15	Durban, South Africa	-29-88	31.8	suspected marine debris	S2B_MSIL1C_20220415T072609_N0400_R049_T36JUN_20220415T094807	https://www.dailymaverick.co.za/article/2022-06-14-huge-volume-of-plastic-flood-debris-trapped-inshore-by-agulhas-current/
2022-04-18	Durban, South Africa	-29.95	31	suspected marine debris	S2B_MSIL1C_20220418T073609_N0400_R092_T36JUM_20220418T094125.SAFE	https://www.dailymaverick.co.za/article/2022-06-14-huge-volume-of-plastic-flood-debris-trapped-inshore-by-agulhas-current/
2022-04-20	Durban, South Africa	-29.3	31.65	suspected marine debris	S2A_MSIL1C_20220420T072621_N0400_R049_T36JUN_20220420T112421	https://www.dailymaverick.co.za/article/2022-06-14-huge-volume-of-plastic-flood-debris-trapped-inshore-by-agulhas-current/
2024-01-08	Congo River Delta	-5.5	11.43	suspected marine debris	S2B_MSIL1C_20240108T092259_N0510_R093_T32MQU_20240108T112129	https://www.reuters.com/world/africa/hundreds-dead-congo-river-basin-submerged-by-generational-floods-2024-01-11/
2020-01-06	Luanda, Angola	-9	13.04	suspected marine debris	S2B_MSIL1C_20200106T091249_N0208_R050_T33LTK_20200106T110709	https://www.youtube.com/watch?v=oNJ-Td5IJY8
2023-10-28	Acapulco, Mexico	16.7	-100-26	suspected marine debris	S2B_MSIL1C_20231028T170419_N0509_R069_T14QLD_20231028T204250	https://bnn.network/politics/fact-checker/acapulco-in-peril-the-aftermath-of-hurricane-otis/
2021-11-09	Faial, Açores	38.52	-28.8	suspected floating vegetation	S2B_MSIL1C_20211109T125039_N0301_R095_T26SLH_20211109T131946	N/A
2021-03-24	Faial, Açores	38.61	-28.61	suspected bloom	S2B_MSIL1C_20210324T125039_N0209_R095_T26SLH_20210324T161829	N/A
2022-08-25	Manila, Phillipines	14.58	120.77	suspected marine debris and vegetation	S2A_MSIL1C_20220825T021541_N0400_R003_T51PTS_20220826T161133	Cruz & Shimozono (2021)
2022-06-11	Manila, Phillipines	14.48	120.71	suspected marine debris and vegetation	S2B_MSIL1C_20220611T021539_N0400_R003_T51PTS_20220611T043923	Cruz & Shimozono (2021)
2021-10-19	Manila, Phillipines	14.55	120.71	suspected marine debris and vegetation	S2A_MSIL1C_20211019T021721_N0500_R003_T51PTS_20230110T151745	Cruz & Shimozono (2021)
2022-06-15	Playa del Carmen, Mexico	20.6	-87.06	floating sargassum	S2A_MSIL1C_20220615T160911_N0400_R140_T16QDH_20220615T212938	N/A
2022-05-22	Bremenhaven, Germany	53.83	8.04	sea foam suspected debris	S2B_MSIL1C_20220522T103629_N0400_R008_T32UME_20220522T113002	N/A

2021-07-22	River Thames, England	51.5	0.77	suspected floating vegetation	S2B_MSIL1C_20210722T105619_N0301_R094_T31UCT_20210722T115604	N/A
2021-03-29	Madeira, Portugal	32.66	-17.23	suspected floating vegetation	S2B_MSIL1C_20210329T120219_N0209_R023_T28SBB_20210329T152940	N/A
2021-08-03	North Sea	53.62	3.05	suspected bloom	S2A_MSIL1C_20210803T105031_N0301_R051_T31UEV_20210803T130429	N/A
2021-08-11	North Sea	53.84	3.68	suspected bloom	S2B_MSIL1C_20210811T105619_N0301_R094_T31UEV_20210811T114330	N/A
2021-09-05	North Sea	53.62	3.9	suspected bloom	S2A_MSIL1C_20210905T105621_N0301_R094_T31UEV_20210905T130404	N/A
2016-12-17	Sao Miguel, Azores	37-76	-25.3	sea foam	S2A_MSIL1C_20161217T124032_N0204_R052_T26SPG_20161217T124034.SAFE	N/A
2022-04-30	Terceira, Azores	38.69	-27.38	sea foam	S2A_MSIL1C_20220430T124051_N0400_R052_T26SMH_20220430T142931.SAFE	N/A
2021-04-20	Vancuver, Canada	49.45	-123.51	rafted timber	S2B_MSIL2A_20210420T190909_N0300_R056_T10UDV_20210420T223037	N/A
2019-10-21	Aquaculture Sites, China	26.51	119.71	floating wooden rafts, plastic foam floats, semisubmerged seaweed	S2B_MSIL2A_20191021T023729_N0213_R089_T50RQQ_20191021T064010	N/A
2023-02-06	North Sea	53.94	7.68	not defined floating objects	S2B_MSIL1C_20230206T104129_N0509_R008_T31UGV_20230206T123935	Acquisition during Labplas Field Campaign
2023-02-07	North Sea	53.14	4.71	not defined floating objects	S2A_MSIL1C_20230207T110221_N0509_R094_T31UFU_20230207T130025.SAFE	Acquisition during Labplas Field Campaign
2023-06-24	North Sea	54.02	6.19	microalgae bloom (noctiluca)	S2A_MSIL1C_20230624T104621_N0509_R051_T32ULE_20230624T142921	Acquisition during Labplas Field Campaign
2023-06-24	North Sea	53.65	3.62	microalgae bloom (noctiluca)	S2A_MSIL1C_20230624T104621_N0509_R051_T31UEV_20230624T142921	Acquisition during Labplas Field Campaign

9.4 Datasheet shared with project partner for field data collection.

LABPLAS AIR CENTRE

Floating Debris Observation Datasheet

Partner		
Case-study		
Sampling Site		
Date		
Hour		
Coordinates	N	W
Were you able to identify any floating debris agglomerate?		
Did you take pictures? If yes include in attachment		
Comments		

9.5 List of acquired GEOSAT-2 products.

Date	Product ID
05-02-2023	DE2_PS3_L1C_000000_20230205T083322_20230205T083325_DE2_46780_0522
07-02-2023	DE2_PS3_L1C_000000_20230207T090143_20230207T090148_DE2_46810_DE02
08-02-2023	DE2_PS3_L1C_000000_20230208T091612_20230208T091615_DE2_46825_E60B
27-06-2023	DE2_PS3_L1C_000000_20230627T092938_20230627T092941_DE2_48890_DE27
28-06-2024	DE2_PS3_L1C_000000_20230628T094248_20230628T094251_DE2_48905_AE00
29-06-2025	DE2_PS3_L1C_000000_20230629T095652_20230629T095655_DE2_48920_1AC7
30-06-2026	DE2_PS3_L1C_000000_20230630T083428_20230630T083432_DE2_48934_92F8