

---

## iAtlantic Deliverable 2.4

### New imaging and analysis approaches for marine species detection and classification

---

|                     |  |
|---------------------|--|
| Project acronym:    | iAtlantic  |
| Grant Agreement:    | 818123   |
| Deliverable number: | D2.4   |
| Deliverable title:  | New imaging and analysis approaches for marine species detection and classification                                      |
| Work Package:       | 2  |
| Date of completion: | 27.03.2023   |
| Author:             | Timm Schoening, Carlos Dominguez Carrio, Telmo Morato, Touria Baijouk, Maxime Ferrera, Tristan Petit & Aurélien Arnaubec |



*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 818123 (iAtlantic). This output reflects only the author's view and the European Union cannot be held responsible for any use that may be made of the information contained therein.*

## Contents

|  |           |
|--|-----------|
| List of Acronyms .....   | 3         |
| List of Tables.....  | 4         |
| List of Figures.....   | 4         |
| <b>1. Executive Summary.....</b>   | <b>8</b>  |
| <b>2. Introduction.....</b>  | <b>10</b> |
| 2.1 Overview .....   | 10        |
| 2.2 Motivation .....   | 10        |
| <b>3. Low-cost camera systems.....</b>   | <b>11</b> |
| 3.1 Introduction .....   | 11        |
| 3.2 The 'Azor drift-cam' .....   | 12        |
| 3.3 Detailed description of the system.....                                    | 13        |
| 3.4 Use at sea.....  | 25        |
| 3.5 Performance assessment.....  | 29        |
| 3.6 Geological and biological data obtained .....                              | 34        |
| 3.7 Limitations of the Azor drift-cam at the current state of development..... | 38        |
| 3.8 Technology readiness level (TRL).....                                      | 39        |
| 3.9 Conclusions .....  | 39        |
| <b>4. Underwater hyperspectral imaging.....</b>                                | <b>40</b> |
| 4.1 Introduction.....  | 40        |
| 4.2 Overview of underwater hyperspectral imaging technology.....               | 40        |
| 4.3 iAtlantic Experimentation .....  | 45        |
| 4.4 Data processing .....  | 51        |
| 4.4 Benthic feature extraction use cases .....                                 | 57        |
| 4.5 Recommendations towards a UHI data acquisition protocol .....              | 62        |
| 4.6 Conclusion.....  | 64        |
| <b>5. FAIR marine image metadata .....</b>                                     | <b>67</b> |
| 5.1 What is FAIR data? .....   | 67        |
| 5.2 How does FAIR data lead to better science? .....                           | 67        |
| 5.3 What are FAIR marine images? .....   | 67        |
| 5.4 Image FAIR Digital Objects (iFDOs) .....                                   | 68        |
| 5.5 iFDO description .....   | 68        |
| <b>6. Machine-learning for automated image analysis.....</b>                   | <b>73</b> |
| 6.1 Overview .....   | 73        |
| 6.2 Automated underwater image analysis .....                                  | 73        |
| 6.3 Requirements .....   | 73        |
| 6.4 Deep Learning.....   | 74        |
| 6.5 Annotations .....  | 75        |
| 6.6 AI developments focussed on marine images .....                            | 76        |
| 6.7 Recommendations towards a FAIR machine learning environment .....          | 82        |
| <b>7. Future Perspectives.....</b>   | <b>84</b> |
| <b>8. References.....</b>  | <b>86</b> |
| <b>9. Document Information.....</b>  | <b>90</b> |

## List of Acronyms

|               |  |
|---------------|--|
| <b>AI</b>     | Artificial Intelligence                                      |
| <b>AUV</b>    | Autonomous Underwater Vehicles                               |
| <b>CMOS</b>   | Complementary Metal Oxide Semiconductor                      |
| <b>CNNs</b>   | Convolutional Neural Networks                                |
| <b>CTD</b>    | Conductivity, Temperature, Depth                             |
| <b>CWC</b>    | Cold-Water Corals  |
| <b>DN</b>     | Digital Numbers  |
| <b>DOI</b>    | Digital Object Identifier                                    |
| <b>DVL</b>    | Doppler Velocity Log   |
| <b>EEZ</b>    | Exclusive Economic Zone                                      |
| <b>EOSC</b>   | European Open Science Cloud                                  |
| <b>FAIR</b>   | Findability, Accessibility, Interoperability and Reusability |
| <b>FOV</b>    | Field of View  |
| <b>GANs</b>   | Generative Adversarial Networks                              |
| <b>GPU</b>    | Graphics Processing Unit                                     |
| <b>HOV</b>    | Human-Occupied Vehicle                                       |
| <b>iFDO</b>   | Image FAIR Digital Object                                    |
| <b>IMU</b>    | Inertial Measurement Unit                                    |
| <b>INS</b>    | Inertial Navigation System                                   |
| <b>LBL</b>    | Long Base Line   |
| <b>MIR</b>    | Mid-Infrared   |
| <b>ML</b>     | Machine Learning   |
| <b>NIR</b>    | Near-Infrared  |
| <b>NTNU</b>   | Norwegian University of Science and Technology               |
| <b>OUT</b>    | Operational Taxonomic Unit                                   |
| <b>PIDs</b>   | Persistent Identifiers                                       |
| <b>RDA</b>    | Research Data Alliance                                       |
| <b>RF</b>     | Random Forest  |
| <b>RGB</b>    | Red, Green and Blue  |
| <b>ROI</b>    | Region of Interest   |
| <b>ROV</b>    | Remotely Operated Vehicle                                    |
| <b>RX</b>     | Video Receiver   |
| <b>SAM</b>    | Spectral Angle Mapper  |
| <b>SHiPCC</b> | Sea-Going High-Performance Compute Cluster                   |
| <b>SLAM</b>   | Simultaneous Localisation and Mapping                        |
| <b>SVM</b>    | Support Vector Machine                                       |
| <b>TRL</b>    | Technology Readiness Level                                   |
| <b>TX</b>     | Video Transmitter  |
| <b>UHI</b>    | Underwater Hyperspectral Imaging                             |
| <b>USBL</b>   | Ultra-Short Baseline   |
| <b>UIDs</b>   | Universally Unique Identifiers                               |

## List of Tables

|  |    |
|--|----|
| <b>Table 1.</b> List of components that make up the Azor drift-cam and at least one potential manufacturer/supplier for each item. ....  | 22 |
| <b>Table 2.</b> Summary table with information on the dives performed with the Azor drift-cam inside the Azores EEZ during the MapGES cruises performed between 2019 and 2022 .....  | 31 |
| <b>Table 3.</b> Number of dives performed in each sampling area along the Mid-Atlantic Ridge during the MapGES 2019 cruise, with information on the depth range, average length, average dive time and total footage recorded in each seamount. Area (km <sup>2</sup> ) refers to the extent in km <sup>2</sup> of the geomorphological feature above 1,000 m depth inside the sampling area. ....   | 35 |
| <b>Table 4.</b> The behaviour of the Azor drift-cam over the seabed given for each sampling area along the Mid-Atlantic Ridge during the MapGES 2019 cruise. Values are provided as percentage of time considering only the bottom time, which spans from the moment the system reaches the seabed until it starts the ascent back to the vessel. ....   | 36 |
| <b>Table 5.</b> Composition of the main substrate types in the 15 sampling areas explored along the Mid-Atlantic Ridge during the MapGES cruise 2019. Values are provided as percentage of time considering only the bottom time, which spans from the moment the system reaches the seabed until it starts the ascent back to the vessel. ....  | 37 |
| <b>Table 6.</b> Number of morphospecies per phylum identified in each of the sampling areas surveyed along the Mid-Atlantic Ridge during the MapGES 2019 cruise. Morphospecies are classified in the most representative groups of the phyla Cnidaria and Porifera. ....   | 38 |
| <b>Table 7.</b> Current hyperspectral technologies concepts, summarised from Stuart et al. <sup>18</sup> . The main conclusion is that push broom spectral imaging is commonly used, thanks to its suitability for light-weight UAV image acquisition <sup>19</sup> . Snapshot systems can record the spectral image data cube from a unique shot, which is suitable for moving targets, however, these devices are currently limited due to their larger size <sup>18</sup> . . | 41 |
| <b>Table 8.</b> Integration test survey summary .....  | 46 |
| <b>Table 9.</b> Confusion matrix for assessing the accuracy of the Random Forest algorithm applied to classify pixels along a UHI data transect acquired in Lampaul canyon.....  | 62 |
| <b>Table 10.</b> Comparison of underwater hyperspectral imaging system concepts .....  | 65 |
| <b>Table 11.</b> Overview of FAIR recommendations by the RDA <sup>45</sup> and their implementation for marine images. Recommendations are ranked by the RDA as ‘Essential’ (*), ‘Important’ (+) and ‘Useful’.....   | 72 |

## List of Figures

|   |    |
|---|----|
| <b>Figure 1.</b> Image of the Azor drift-cam on the deck of the vessel N/I <i>Arquipélago</i> ready to be deployed to obtain images of the deep sea. ....   | 13 |
| <b>Figure 2.</b> The Azor drift-cam system. Lateral and front view of the main frame with all electronic components mounted on the stainless-steel structure.....   | 14 |
| <b>Figure 3.</b> (a) Shape and dimensions of the stainless-steel frame of the Azor drift-cam, displaying the attachment and brackets used to secure and adjust all electronic components (e.g., camera housings, lights, external batteries, lasers, and sensors). All attachments and brackets can be adjusted along the frame to reach the best configuration. The housings used to hold the batteries and the live-view components are directly attached to the central bar using stainless steel duct clamps (see Figure 2). (b) Shape and measurements of the stabilising wing. (c) Example of a basalt boulder used as ballast. It is recommended to use a low breaking tension line between the weight and the structure (white arrow) to act as a fuse in case of entanglement..... | 15 |
| <b>Figure 4.</b> Electronic components used in the Azor drift-cam. All components are mounted on the metallic structure with the exception of the reader for the depth/temperature sensor, which is kept on board to download the collected data after the dives.....   | 16 |
| <b>Figure 5.</b> Live-view system connections. Layout of the connections placed on the main structure that allow the video signal to be sent from the live-view action camera to the umbilical and the powering of the camera using an external battery source.....   | 17 |
| <b>Figure 6.</b> The umbilical. (a) Image of the polypropylene rope (green) and electrical cable (black) used to control the Azor drift-cam and get the video feed to the surface when underwater. (b) Attachment of the rope and the umbilical every meter or so using duct tape. (c) One of the deep-sea floats employed to keep the drift-cam in an upright position at all times. A thin cord is used to keep the floats in a fixed position  |    |

- on the rope. .... 18
- Figure 7.** Step-by-step illustration of the process for joining the end part of the electric cable of the umbilical to the SubConn connector or for re-joining a broken electric cable. (a) The inner cables on both ends are peeled and cleaned. (b) Using a soldering iron and solder wire, cables are welded together. (c-g) Through a process of iteratively applying several layers of heat-shrink sleeves and vulcanising tape, the joint cables are provided with a resistant coating that prevents water from leaking, even at 1,000 m depth. (h) Final result, with the area welded presenting a larger diameter resulting from the application of multiple protection layers. .... 19
- Figure 9.** PC mounted on board of the N/I *Arquipélago* used to centralise all the information collected with the Azor drift-cam, with the software ArcGIS used to monitor the position of the vessel using the best available bathymetry. .... 21
- Figure 10.** (a) Two members of the ADSR group setting up the cameras and batteries of the Azor drift-cam prior to its deployment. (b-c) Deployment of the Azor drift-cam from the side of the vessel using a crane to lift the system. .... 26
- Figure 11.** Simplicity in the use of the Azor drift-cam. (a) An observer providing indications to the winch operators based on the images provided by the live-view system. (b) Winch operators controlling the amount of cable required to keep a safe distance between the drift-cam and the seabed. The excess cable is stored in a large bucket below the winch. .... 26
- Figure 12.** Screen captures obtained from the video footage recorded with the action camera mounted on the Azor drift-cam during the MapGES 2019 cruise, illustrating some of the different seafloor substrates that were surveyed. The small red stripes in the mid/low part of the image correspond to the reflection of the parallel lasers over the seabed, used for image scaling. The distance between the two marks on the left corresponds to 10 cm. Scale bars = 50 cm. (a) Sand ripples in Monte Alto seamount, 460 m depth. (b) Gravels and boulders in Cavala seamount, 650 m depth. (c) Flat substrate in SE of Pico island, 400 m depth. (d) Sloping substrate in Picoto seamount, 500 m depth. (e) Mixture of sand, gravels and scattered rocks in Gigante seamount, 470 m depth. (f) Large rocky outcrops in Voador seamount, 390 m. (g) Large boulders on a very steep terrain in A3 seamount, 560 m depth. (h) Vertical wall in A3 seamount, 500 m depth. .... 28
- Figure 13.** Differences in the field of view (FOV) of the main camera depending on the distance between the system and the seabed. .... 29
- Figure 14.** Location of the 535 dives performed with the Azor drift-cam inside the Azores EEZ during the MapGES surveys of the past 4 years (2019 to 2022). The dives are colour coded according to the year of the survey. Green: 2019; Yellow: 2020; Red: 2021; Pink: 2022. The map also includes the dives from 2018 (in orange) when the system was still a prototype in early development and was used as a drop-down camera. The blue and black dots correspond to the dives performed with ROVs and towed camera system before the development of the Azor drift-cam and for which the ADSR group has access to the images. .... 31
- Figure 15.** Depth range of the dives performed with the Azor drift-cam inside the Azores EEZ between 2019 and 2022 part of the MapGES cruises. The bar plot on the right side indicates the amount of time spent by the Azor drift-cam (in %) surveying each depth range at 50-m intervals. .... 31
- Figure 16.** Location of the sampling areas explored with the Azor drift-cam along the Mid-Atlantic Ridge during the MapGES cruise 2019 on board of N/I *Arquipélago*. .... 34
- Figure 17.** Examples of different substrate types observed in the images recorded with the Azor drift-cam. (a) Sand. (b) Gravels. (c) Coral rubble. (d) Flat rock. (e) Boulders. (f). Outcropping rock. .... 36
- Figure 18.** Examples of the different megabenthic species observed in the images recorded with the Azor drift-cam in the Mid-Atlantic Ridge. (a) Large colonies of *Callogorgia verticillata*. (b) Aggregation of the yellow sea fan *Dentomuricea aff. meteor*. (c) The whip coral *Viminella flagellum*. (d) A large colony of the red morphology of the species *Paragorgia johnsoni*. (e) The giant sponge *Characella pachastrelloides*. (f) An aggregation of the hexactinellid sponge of the genus *Asconema*. .... 38
- Figure 19.** Schematic diagram of a UHI system<sup>21</sup> (A), an example of ROV used for the UHI-based survey in Trondheimsfjord<sup>22</sup> (B) and the Ecotone UHI (first circular opening indicated by the white circle) mounted on AUV at the right of the lamps (four smaller squares inside the white rectangle)<sup>23</sup> (C). ... 42
- Figure 20.** Annual number of articles published in the field of underwater spectral imaging<sup>21</sup>. .... 43
- Figure 21.** **A1 and A2** respectively show a UHI RGB false-colour image and its corresponding classification of raw data (digital counts) denoting *Hippasteria phrygiana* seastar (Black), brown algae dominated by *Laminaria digitata* kelp (blue), denotes sand/soft bottom (red) and concrete pipeline and holdfast (green), **B1 and B2** respectively show a UHI false-colour image and its supervised classification into

Sand (yellow), leafy red algae (green), red calcareous algae (red), old and dark brown tissue of kelp *L. digitata* (blue) and corresponding new tissue of *L. digitata* (winter growth from meristem). **C1 and C2** respectively show a UHI photomosaic in false-colours (RGB) and its corresponding classification based of seafloor habitat dominated by red calcareous algae (red area), anemones dominated by *Urticina* (yellow) and sea urchins *Strongylocentrotus* spp. (blue) during a Polar night campaign at 15 m depth (from Johnsen et al., 2016). **D1 and D2** respectively show a high-resolution UHI data item and its corresponding quantitative mapping of Chl a by applying the LAUC index regression model on a per-pixel basis to the pseudo-transmittance image, **E1 and E2** respectively show UHI geo-corrected, pseudo-reflectance data in pseudo-RGB and its SVM classification image, **F1 and F2** respectively show an RGB UHI image of a coralligenous site and the corresponding SAM classification (from Foglini et al. 2019) and **(G)** Supervised SAM classification images of a wreck transect. From top to bottom: glass bottle, rust, and ceramics (from Ødegård et al., 2018)..... 44

**Figure 22.** A schematic diagram showing the geometry of UHI acquisition deployed on ROV<sup>24</sup>. ..... 45

**Figure 23.** Underwater Hyperspectral imaging acquired system (A) and a schematic diagram showing the principle of data acquisition (B). ..... 46

**Figure 24.** The *Vortex* IFREMER HROV with Underwater Hyperspectral imager (UHI) mounted in the rear end (Wright) carried out from the TSM *Penzer* vessel (Top left) by IFREMER and Ecotone team with the help of the crew and the PNMI divers (bottom left). ..... 47

**Figure 25.** Left: study areas in the Brest harbour, right: hyperspectral images turned into RGB false-colour images of starfish (top) and brittle stars (bottom) at red algae covered seafloor. .... 48

**Figure 26.** Stitched and georeferenced UHI image transect (A), example of georeferencing issue (B) with UHI navigation in green, Corrected UHI navigation in pink and 3D point cloud from 3D reconstruction in white (C). ..... 48

**Figure 27.** Snapshot of the Ecotone RGB camera (left) and the high-resolution RGB imagery acquired by the *Vortex* HROV still camera (right) for 3D reconstruction for optical navigation. .... 49

**Figure 28.** The *Ariane* HROV with Underwater Hyperspectral imager (UHI) mounted in the front (B) carried out from the *Antea* vessel (B) by IFREMER, Ecotone and Genavir team (C). ..... 50

**Figure 29.** Overview map of the deployment sites during the survey campaign and one example high-resolution image depicting the kind of habitat that was imaged both by the RGB camera and the UHI sensor. ... 51

**Figure 30.** Stereo setup of the two cameras on *Ariane* HROV from IFREMER. The upper pressure housing contains the RGB camera. The bottom pressure housing contains the push-broom HSI camera from Ecotone. The Z axis of the camera coordinate system is in the view direction, the X-axis points to the right and the Y-axis is pointing downwards. The cameras are mounted vertically, thus are having the y-axis as known the baseline between them. .... 51

**Figure 31.** Screenshot of the Immersion software package provided by Ecotone for the operation of their UHI camera. .... 51

**Figure 32.** Dark (left) and gain (right) calibration matrices of the Ecotone UHI Camera operated at IFREMER. The dark matrix is expressed in unit of DN while the gain matrix is given in  $\mu\text{Wcm}^{-2}\text{nm}^{-1}\text{sr}^{-1}$  ..... 52

**Figure 33.** Comparison between raw and calibrated UHI data acquired during the Chereef 2022 cruise (2022-08-06 at 20:51) with, from left to right: RGB composition of the Raw image in DN, RGB composition of the calibrated image in radiance after immersion correction, spectral comparison of one hyperspectral pixel in DN and the radiance curves before (dashed) and after (solid) immersion correction. .... 53

**Figure 34.** Temporal statistics (5% quantile, median and mean) of cross-track pixel values for 200 successive images (197,400 temporal acquisitions) for the Katchof (left) and Chereef (right) cruises with data in DN (top) and radiance (bottom). The red curves are computed from the manufacturer-provided calibration spectrum data. .... 54

**Figure 35.** Downwelling irradiance at sensor level estimated from a laboratory measurement of the ROV lights. .... 55

**Figure 36.** RGB image of sensor (left) and bottom (right) reflectance of UHI data of the Chereef 2022 cruise. ... 56

**Figure 37.** Water body thickness (left, with values ranging between 1.5-2.8 m from blue to red) and bottom albedo at 550 nm (right, with values ranging between 0.02-0.1) estimated through sensor reflectance inversion. .... 56

**Figure 38.** Pipeline of the method for computing hyperspectral 3D reconstructions, from (Ferrera et al., 2021). .... 57

**Figure 39.** Standardised spectral signatures of the main species and substrate types of the Catchoff Mediterranean site (A) and scatterplot of the first two principal components of a PCA (B). .... 58

**Figure 40.** Zoom-in to a selected part of a UHI recording showing noisy spectral bands along and across-track. 59

- Figure 41.** Spectral signatures recorded in the laboratory from samples collected during the 2021 Chereef cruise to Lampaul Canyon (France). ..... 59
- Figure 42.** UHI RGB visualisation (A), RGB image from video recorded simultaneously (B), and regions of interest (ROIs) created for training and validating UHI image classifications. .... 60
- Figure 43.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied on area 1 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the classifiers tested. .... 61
- Figure 44.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied on area 2 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the classifiers tested. The area bounded by the black dotted rectangle within the RF classification shows the detection of an individual of category ‘Narella’, the only place where this species is present in this area. .... 61
- Figure 45.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied to area 3 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the evaluated classifiers. .... 62
- Figure 46.** Setup of image FAIR Digital Objects. Key information and image data is stored in a dedicated infrastructure (yellow squares). iFDOs only contain persistent identifiers to those external information resources. Additionally, specific metadata for marine imaging use-cases is stored inside the iFDO files. iFDOs consist of three sections: (1) the required core part which includes the persistent identifiers as well as licensing information; (2) the recommended capture part that addresses the technical heterogeneity of image acquisition; and the (3) the optional content part that captures semantic information from within the images to address the heterogeneous nature of image data. Together, these three sections constitute one iFDO file. This file contains header information on the entire image data set as well as detailed information on each image item within a defined set of images. .... 69
- Figure 47.** Creation and progression of an iFDO (green circles) and its derived versions. The left part shows how an iFDO uses persistent identifiers to reference itself within the FAIR infrastructure. The middle part shows how iFDO files can be discovered, shared, advertised, and validated. The right part shows how implementing iFDO-compliant APIs to marine science tools facilitates reuse of image data for arbitrary purposes. .... 70
- Figure 48.** Creating an iFDO. Marine image acquisition is guided by OceanBestPractices<sup>44</sup> and creates raw image data and raw position data (for in-situ imaging). Multiple processing steps (blue boxes) create derived data products (green boxes) that are ultimately merged by an iFDO factory process to one iFDO file (green circle). .... 71
- Figure 49.** AI-generated images from the input phrases: ‘a deep-sea seafloor image’ (left), ‘manganese nodule field in the deep sea’ (middle) and ‘a top down looking deep-sea seafloor image with a holothurian’ (right). All images were created by OpenAI’s DALL-E. .... 75
- Figure 50.** The effect of image resolution (given in megapixels per square meter images) on the horizontal axis on various aspects of underwater image analysis. Key metrics reported in scientific papers such as Fauna Density (a) and observed Fauna Size (b) are correlated with resolution. Effort in imaging the seafloor increases with resolution, for the effort in inspecting the images for annotation this increase is disproportionately large. Robustness of annotations (e.g., measured in observer agreement (g), Cohens kappa (i) or uncertainty in annotations (c) also increase with resolution). .... 78
- Figure 51.** Sketch of the proposed FAIR machine learning framework for marine image analysis. Green entities represent data items, e.g., images, annotations or operation procedures. Blue squares represent examples of existing infrastructures that can facilitate operationalising parts of the framework. Red items denote standardisation formalism necessary to advance the TRL of machine learning for marine image analysis beyond demonstrator cases. .... 83

## 1. Executive Summary

Photos and videos taken underwater enable us to get a visual impression of ocean processes which would otherwise remain hidden from our eyes. Marine imaging is used to explore and monitor the marine environment and creates comprehensible material for captivating ocean narratives. Imaging can address questions about the ocean across many spatial and temporal scales from in-situ microscopy to satellite remote sensing, and from slow-motion capture to long-term observatories.

Marine imaging, as a method, covers a range of technical aspects, which makes it a versatile technology employed in many research, industry and public applications. This includes aspects such as sensors, cameras, platforms, illumination, human interaction such as annotation, automated information extraction such as machine learning, or Findable, Accessible, Interoperable and Reusable (FAIR) image publication. And this range is expanding as general technology advancements such as hyperspectral imaging or deep learning are making their way into ocean applications. The wider availability of underwater platforms equipped with camera systems creates ever-increasing volumes of big data. The availability of commercial low-cost camera systems for deep-sea exploration is not only increasing survey coverage, but is proving critical to expand this capacity to areas of the world previously unable to survey deep-sea ecosystems.

During the *iAtlantic* project, selected aspects across the marine imaging data workflow were conceptualised, implemented, deployed and operated. This document provides a summary of the developments that occurred, partly during the global Covid-19 pandemic. It includes details on imaging technology, image data, image metadata, and image processing methods. Four main aspects are described: a) underwater hyperspectral imaging that was trailed and successfully demonstrated by IFREMER; b) a low-cost camera system that was designed and built by IMAR and the successful deployment of this system in Portuguese waters; c) making marine image data FAIR for robust science, and d) efficient machine learning applications, both led by GEOMAR.

Imaging research described in this technical report contributes to *iAtlantic* objectives in several ways. The efforts in making marine image data FAIR and developing a FAIR machine learning infrastructure support to **align and standardise ocean observing** across geographical regions, marine science domains, research institutes and marine sectors. Similarly, the publication of the low-cost camera system and its widespread adoption can lead to the creation of standardised data sets across the Atlantic Ocean. This camera system has been successfully deployed to **map deep and open-ocean ecosystems at local and regional scale** in the Azorean Exclusive Economic Zone (EEZ). An adoption of the open hardware concept by others can similarly support creating ecosystem data at a global scale. These data have already been used in other aspects of *iAtlantic* research to **assess the stability, vulnerability, and tipping points** of ecosystems, e.g., in habitat mapping applications.

However, the most prominent contribution of the imaging work is towards the objective to **build and enhance human and technological capacities** which is represented by the innovative approaches to link underwater hyperspectral imaging with 3D reconstructions, the low-cost camera hardware development, operationalising seagoing high-performance computers and the standardisation efforts and software development for the FAIR marine image and the FAIR machine learning environments. These technologies have been presented and advertised and are potential elements for commercialisation, **supporting a sustainable blue economy**.

The technology readiness levels (TRLs) of the three proposed aspects of marine imaging (low-cost cameras, hyper-spectral imaging, machine learning) were advanced as expected since the start of the



project. The newly developed concept of FAIR marine images added a fourth marine imaging aspect for which the state-of-the-art before *iAtlantic* was TRL 2 which could be advanced through the project efforts to TRL 7. The Azor drift-cam has been advanced through the project from TRL 6 to TRL 7. For the machine learning, selected methods were implemented operationally at TRL 8 (from 7). With the additionally developed FAIR machine learning infrastructure, a generalisation of efficient machine learning towards TRL 9 has been proposed and needs operationalising through future projects. The technology for hyper-spectral imaging was already at TRL 8 through employing a proven off-the-shelf system. Integration into operational, mission-proven equipment at IFREMER is currently progressing towards TRL 9.

The developments described in detail below have led to ideas for future innovation and requirements for their implementation that are described in the outlook section at the end of this document.

## 2. Introduction

### Synopsis:

*This is Deliverable 2.4 of the iAtlantic project - a technical report on the development of new underwater imaging, image management, and image analysis techniques. It includes details on imaging technology, image data, image metadata, and image processing algorithms. Four main topics are described: underwater hyperspectral imaging, low-cost stereo-camera systems and their use, machine learning applications to extract information from images and FAIR image metadata.*

### 2.1 Overview

The term imaging sounds straightforward— yet it disguises a great variety of aspects such as scopes, technology, big data, algorithms, impressions, and visual narratives. All these imaging aspects are linked to each other, though specific applications of underwater imaging target a fraction of these. In this technical report, the focus lies on *four aspects of imaging* technology development that were conducted within the *iAtlantic* project: i) machine-learning approaches for automated image analysis; ii) hyperspectral imaging; iii) the development of low-cost cameras and their deployment for marine science applications; and iv) the development of guidelines and tools for FAIR marine image data. These technology developments were *aligned with the overarching goals of iAtlantic*: to standardise ocean observing methods; to map ecosystems at local to basin scale; to assess ecosystem vulnerability and to build and enhance technological capacities for cost-effective imaging.

### Highlights

Key outcomes of the imaging work conducted within the framework of the *iAtlantic* project that are described within this technical report are:

1. The operationally proven *Azor drift-cam* low-cost camera system.
2. An operational metadata schema enabling *FAIR marine image data*.
3. A status review of *underwater hyperspectral imaging*.
4. A concept of a *FAIR machine learning environment*.

### 2.2 Motivation

Two key aspects of imaging are competing in a never-ending trade-off: the capacity to gain more information with images and the associated costs to do so. Gaining more information is critical for applications such as autonomous vehicles, long-term time-series image acquisition, 4K videos, crowdsourcing, spectral imaging, etc. Each of these methods is associated with an increased cost: a high financial burden for deep-diving robots or hyperspectral cameras, a high personnel burden for analysing big data sets of state-of-the-art image resolution, a high curation effort to standardise heterogeneous data sources such as crowd-sourced material.

Within *iAtlantic*, four imaging technologies were advanced to mitigate this trade-off. Hyperspectral imaging was applied to gain more spectral information, e.g., to assess the health status of ecosystems, yet with the scope of high-throughput analysis. Low-cost cameras were developed and successfully deployed to mitigate the ever-increasing cost of robotic camera platforms and to enable deep seafloor imaging with near-shore shipping capacity available from the blue economy. Machine learning (ML) approaches for automated image analysis focussed on the development of standards for FAIR and open publication of image data and annotations for training ML algorithms to reduce the cost of developing tailored ML systems for new data sets.

### 3. Low-cost camera systems

Contributions from: Carlos Dominguez-Carrió, Telmo Morato and the co-authors of the peer-reviewed paper that contributed to this section<sup>1</sup>.

#### 3.1 Introduction

Large efforts have recently been placed in improving our understanding of the diversity and functioning of deep-sea ecosystems, the ocean's deepest layers. However, and due to its huge extension and difficult access, vast areas remain largely unexplored<sup>2</sup>. The development of marine technology provided us with a series of underwater imaging tools that can gather *high-quality images of the seabed* thousands of metres below the surface, such as Human-Occupied Vehicles (HOVs), Remotely Operated Vehicles (ROVs), Autonomous Underwater Vehicles (AUVs) and towed camera systems<sup>3-6</sup>. These platforms, especially ROVs, currently represent the *most common approach to the study of deep-sea benthic communities*<sup>7</sup>, replacing classical and more intrusive techniques (e.g., epibenthic sledges, beam trawls and grabs). Although underwater camera systems have become essential tools for the study of deep-sea benthic habitats, deep-sea research using cutting-edge technology remains inaccessible to many research teams, notably in the Global South, largely because of the high costs associated with large oceanographic vessels and specialised crews<sup>8-10</sup>.

The current commitments to protect marine biodiversity (e.g., UN 2030 Agenda for Sustainable Development and EU 2030 Biodiversity strategy) have highlighted the national and international *responsibilities to collect large-scale scientific data* on the distribution and conservation status of deep-sea benthic species and habitats. Thus, the information obtained by video tools is of paramount importance to identify priority areas for management and conservation, essential to achieve national and/or international conservation targets (e.g., Convention on Biological Diversity Aichi Biodiversity Target 11 or UN Sustainable Development Goal 14). Therefore, there is a *need for sampling tools* that can generate new deep-sea scientific data at a reasonable cost, making deep-sea exploration easier, simpler, cheaper, and accessible to many. In recent years, several low-cost deep-sea camera systems<sup>11,12</sup>, sensors<sup>13</sup> and video landers<sup>14</sup> have been prototyped. However, few technological developments allow deep-sea exploration with low cost, readily accessible off-the-shelf products.

The Azores region, located in the middle of the Atlantic Ocean at the crossroads of the American, African, and Eurasian tectonic plates, is very complex in terms of deep-sea geomorphology. Its EEZ spans for more than 1,300 kilometres, with over 130 seamounts with their summits shallower than 1,000 m depth. For years, local scientists and policy makers had limited access to underwater imaging devices and deep-sea exploration mostly relied on international research vessels equipped with commercial ROVs and towed camera systems visiting the archipelago. Many these scientific cruises focused on the study of hydrothermal vents, common features along the Mid-Atlantic Ridge, and most of our knowledge on the biological diversity of the deep sea was based on the information collected by the few research cruises that targeted benthic habitats and from the study of organisms accidentally collected in fishing by-catch. Until a few years ago, the deep-sea area prospected in the Azores using visual methods was still relatively small, and comprehensive information regarding the diversity and composition of its benthic communities was only available for a limited number of seamounts. For this reason, and to speed up deep-sea exploration in the region and better inform policy makers, researchers at IMAR worked on the development of a prototype of an underwater imaging system to collect video images of the deep seabed at a reasonable cost, which led to the design of the 'Azor drift-cam'.

### 3.2 The 'Azor drift-cam'

The 'Azor drift-cam' is an affordable and easy-to-use underwater video system designed for a rapid appraisal of benthic habitats down to 1,000 m (Figure 1). It is a modular, light, and easy-to-assemble video platform for the recording of underwater images of the deep seabed, with all its components currently available for purchase in the retail market (i.e., off-the-shelf). The system was designed to reflect the reality of the Azores, aiming to be cost-effective, cover large areas in short periods of time, perform well over rough seafloors, be operational from small vessels and have high chances of escaping lost long-lines, the most common fishing gear in the region.

The development of the Azor drift-cam was driven by a set of guiding principles (GP) that determined the characteristics of the tool and the choices adopted, namely:

- Effective, i.e., suitable for a rapid appraisal of deep-sea benthic habitats.
- Affordable, i.e., total price to be kept low (10–15k €), including the necessary spare parts.
- Simple, i.e., components should be off-the-shelf, light, and easy to assemble.
- User-friendly, i.e., maintenance and operation without highly specialised staff.
- Resilient, i.e., must be operational in most deep-sea areas, including complex topographies and heavily fished grounds.
- Operable, i.e., deployable from small platforms (including local fishing vessels) to keep its operation costs low.
- Portable, i.e., easily moved between sampling areas or countries.

All electronic devices mounted on the system are battery operated, so no power is fed through the umbilical, further reducing the risks associated with its use. The umbilical is an off-the-shelf electrical cable, simple to manipulate and easy to repair in case of breakages occurring. The metallic frame creates an oval shape to protect its components, reducing the likelihood of entanglement in lost fishing gears or complex terrains. The system is relatively small and light (easily lifted by two people) to allow operations from small vessels. Since all pieces are not inter-related but function independently, the system is modular and thus adjustable and easy to reconfigure. More components, new versions or equipment improvements can be easily added without changing the original configuration. Also, the malfunctioning of one component does not affect the functioning of the remaining parts, allowing for a quick identification of the potential problem. The replacement of any damaged component is very straightforward, further reducing the time needed for repair. Most malfunctions can be repaired on board without needing specific tools and complex testing (e.g., cable damage and connections not working), which can be fixed while at sea.

The present-day cost of the Azor drift-cam falls within the 10–15k € range set in guiding principle Affordability. This value is two orders of magnitude lower than a standard working-class ROV capable of reaching 1,000 m (600 k to 3.5 million €<sup>8</sup>) but falls within the price range of low-cost shallow-water ROVs<sup>15</sup> and towed video systems<sup>16</sup>, and also other low-cost deep remote baited camera systems<sup>11</sup>. Keeping the Azor drift-cam affordable meant that some useful add-ons were left out (e.g., fibre-optic cable; ultra-short baseline (USBL) positioning system; conductivity, temperature, and depth (CTD) profiler; extra lighting), but these decisions did not compromise its general capabilities and goals.



**Figure 1.** Image of the Azor drift-cam on the deck of the vessel N/I *Arquipélago* ready to be deployed to obtain images of the deep sea.

The development of the Azor drift-cam was published as an open access article in the journal *Methods in Ecology and Evolution* in 2021<sup>1</sup>, with the objective of sharing the knowledge gathered and provide the tools to make deep-sea exploration more accessible to research groups currently excluded from this field due to monetary and/or offshore research vessel restrictions. A promotional video was also produced to explain in a simpler way the rationale behind the development of the Azor drift-cam, how it is constructed and show images of how it is operated at sea and the data that can be collected<sup>1</sup>. The system has also been promoted by iAtlantic and multiple meetings with a dedicated Capacity Building Training workshop planned 5-8 June 2023.

### 3.3 Detailed description of the system

The Azor drift-cam is composed of 3 parts: the main body, the umbilical, and the on-vessel components. The main body is made up of a stainless-steel structure where all electronic components are attached to, including the cameras, the lights, the sensors, the lasers and the deep-sea housings containing the live video system and the external batteries (Figure 2). A list with all items necessary to build a replica of the Azor drift-cam, together with potential manufacturers and indicative prices is provided in Table 1.

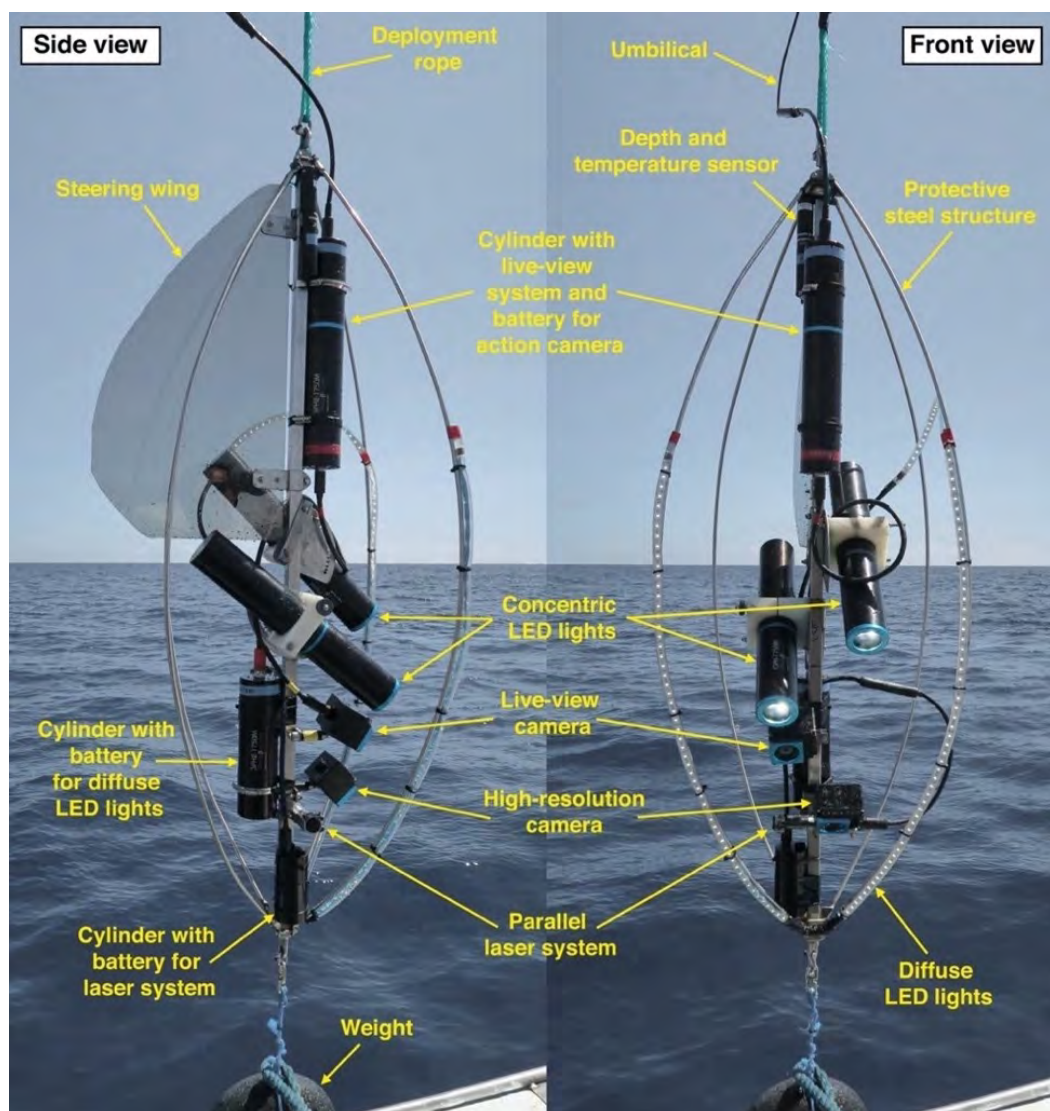
#### External stainless-steel structure

The metallic structure of the Azor drift-cam is made of a 1.5-metre-long square stainless-steel bar (sides of 2.5x2.5 cm), with 4 curved 1.65-metre-long stainless-steel rods (8 mm in diameter) held together on both ends (Figure 3a). These curved bars are designed to protect all the electronic

<sup>1</sup> [https://www.youtube.com/watch?v=agLckO\\_rnPM](https://www.youtube.com/watch?v=agLckO_rnPM)

components and cable connections, from impact and entanglement. All system components are attached to the central bar through custom-made attachments:

- Camera holders. Action camera housings are secured to the main structure using two small custom-made stainless-steel pieces, attached together using two screws. This configuration allows for an easy adjustment along the structure. The tilting of the camera housings can also be easily altered if necessary.
- Light holders. Composed of two custom-made symmetrical Polyvinyl Chloride (PVC) pieces attached with screws. Both pieces have a semi-circular inner section to fit the cylindrical housing of the lights, which is secured simply by tightening the screws. These holders are then attached to a flat metallic piece that allows light angle adjustments.
- Parallel lasers holder. This is a custom-made metal bracket with two adjustable metal rings on both ends to secure the cylindrical laser system.
- Wing holder. Two small metallic flat pieces used to connect and secure the stabilising wing to the upper part of the central bar.

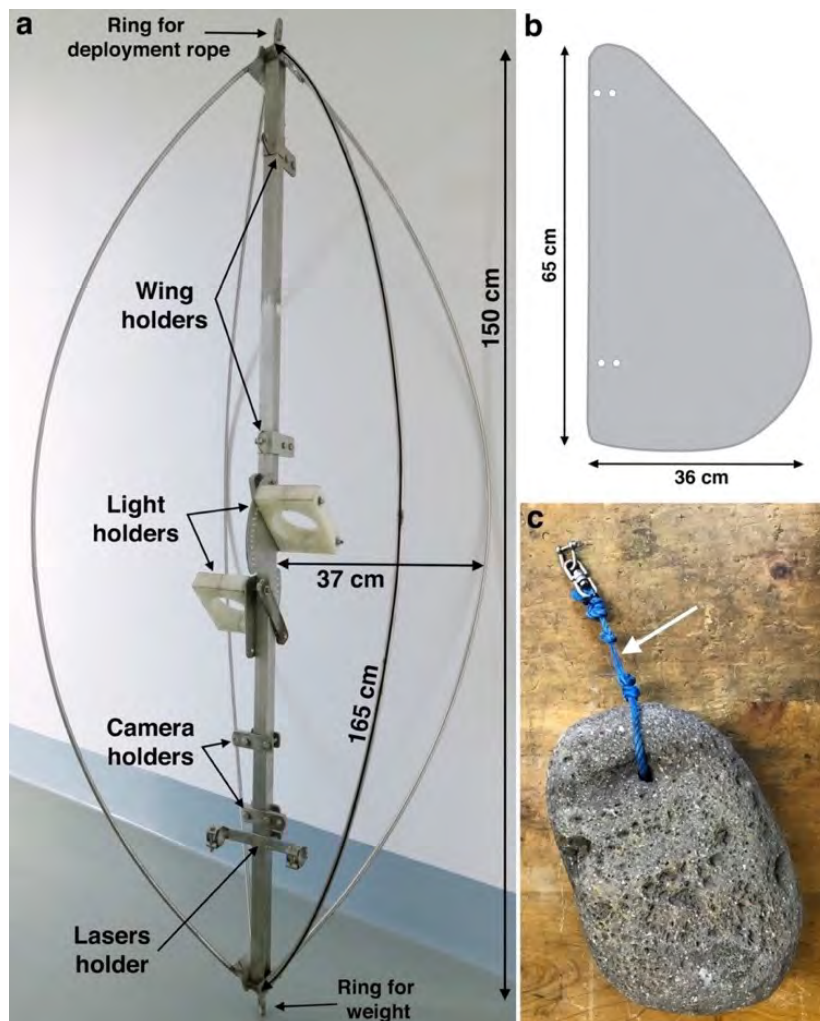


**Figure 2.** The Azor drift-cam system. Lateral and front view of the main frame with all electronic components mounted on the stainless-steel structure.

To maintain a forward-facing orientation when drifting over the seabed, the system uses a stabilising wing, strong enough so it does not bend through time. The best solution found so far is D-shaped methacrylate wing, 65 cm high and 36 cm wide (Figure 3b). It is advised to keep the wing contained

within the protective frame as much as possible to reduce potential entanglement with abandoned fishing lines.

The central bar should come with rings on both ends, which are used to attach the deployment rope (at the top) and the weight (at the bottom). The 15–20 kg weight helps maintaining the umbilical in a vertical position with respect of the boat as much as possible. It is recommended that the weight used is of a tear or spherical shape (Figure 3c) in order to: (1) reduce its spin and drag during deployment and recovery, and (2) increase the likelihood of disentanglement from a fishing line/rope or seafloor structure in the case of getting caught. A thin piece of rope with a low breaking tension should be used as fuse between the weight and the structure (arrow in Figure 3c). This thinned rope should allow for an easy release of the weight in case of severe entanglement, facilitating the recovery of the main body.



**Figure 3.** (a) Shape and dimensions of the stainless-steel frame of the Azor drift-cam, displaying the attachment and brackets used to secure and adjust all electronic components (e.g., camera housings, lights, external batteries, lasers, and sensors). All attachments and brackets can be adjusted along the frame to reach the best configuration. The housings used to hold the batteries and the live-view components are directly attached to the central bar using stainless steel duct clamps (see Figure 2). (b) Shape and measurements of the stabilising wing. (c) Example of a basalt boulder used as ballast. It is recommended to use a low breaking tension line between the weight and the structure (white arrow) to act as a fuse in case of entanglement.

## Electronic components

The Azor drift-cam is equipped with two action cameras contained on off-the-shelf deep-sea housings (GroupBinc, USA) rated to 2,500 m depth (Figure 4). The camera housings are mounted on a similar position on the lower part of the central bar, oriented in the same direction but with slightly different angles (see side view in Figure 2). The high-definition action camera used to provide the live-view signal is placed on a higher position to provide a wider view of the habitat explored and assist navigation. The 4k-resolution action camera is placed below the live-view camera and slightly more tilted towards the seabed, to increase the quality of the images recorded for data processing (i.e., taxonomic identification of the species, annotation of substrate and fauna for quantitative studies and size measurements). The use of the mode 'Linear Field of View' is recommended, which eliminates the barrel distortion of the action camera (fish-eye effect).

The live-view camera is connected to the umbilical to send an analogue video feed to the surface and guide piloting the system (that is, avoiding obstacles or keeping a safe distance with the seafloor). The connections required to send the live-view video signal from the camera to the umbilical are shown in Figure 5. A 90-degree mini-USB plug (GroupB, USA) is connected to the camera for video output and energy input to power the camera (Figure 5). The camera is powered using a rechargeable LiPo battery (e.g., 7.4 V, 5300 mAh) placed inside the main housing, which is attached to the upper side of the central bar (Figure 2). Since the battery generates a higher voltage than that required by the action camera, a voltage regulator is required (Figure 5; see specifications on the camera used). The video signal is sent from the camera housing to the main cylinder (containing the video transmitter, TX) using a 4-pin female-female SubConn connector, which is also used to send the battery power from the cylinder to the camera (Figure 5). The video transmitter (TX) allows the video feed to be transmitted to the surface through a simple dual copper rubber coated wire cable. At surface, a video receiver (RX) and an analogue-to-digital converter (RCA to HDMI) are used to display the images on a PC monitor or a TV screen.



**Figure 4.** Electronic components used in the Azor drift-cam. All components are mounted on the metallic structure with the exception of the reader for the depth/temperature sensor, which is kept on board to download the collected data after the dives.

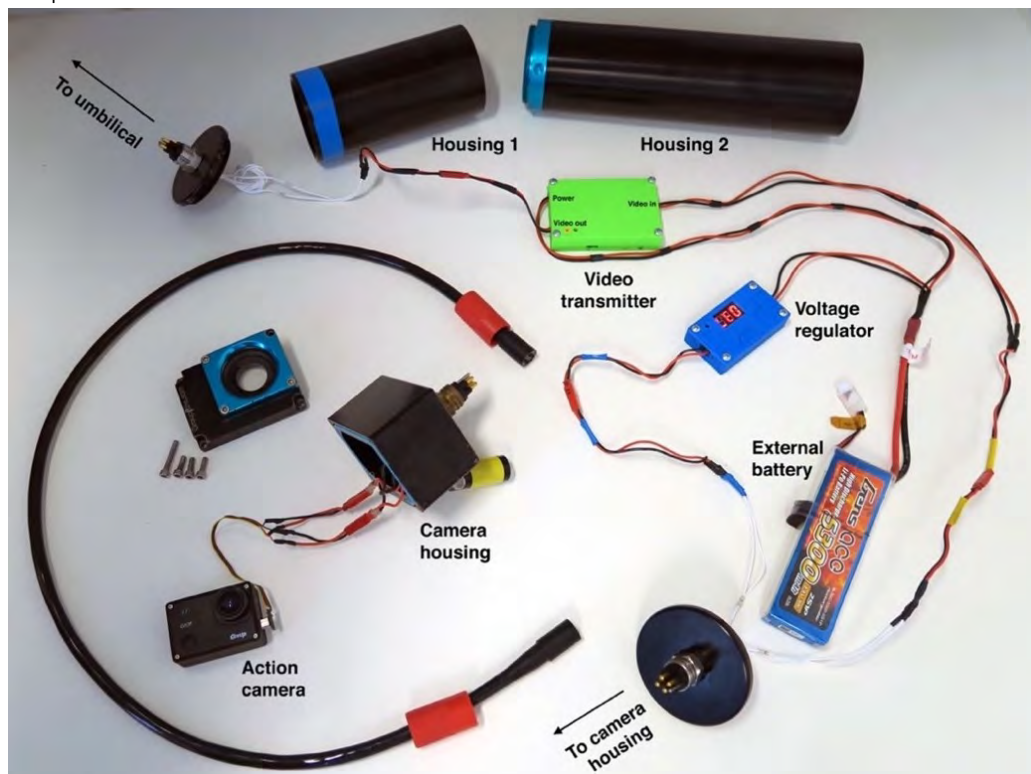
The 4k action camera works independently from the live-view feed, and the video images can only be viewed upon recovery. Images are stored on an internal micro-SD memory card (e.g., 128 Gb or 256 Gb),



which is generally sufficient to record 2.7k or 4k images during a 2-hour dive. This action camera is powered by its own internal battery, which should provide enough energy for these 2 hours. There are currently no off-the-shelf connectors available in the market that allow the latest action camera models to be fed using an external battery.

### Lighting system

The Azor drift-cam is equipped with two different lighting systems designed to provide a powerful yet concentric beam of light and also a diffuse flood light (Figure 2). Both systems provide enough light for navigation purposes while at the same time generate a clean ambience light to obtain sharp footage, necessary for species identification and taxa annotation. The two powerful LED lights (GroupB, USA; Figure 4) generate 3,500 lumens of concentric neutral white light each and are powered by a 11.1 V external LiPo battery (rechargeable). To obtain better quality images, the two flashlights should be attached to the central bar at different angles (see side view in Figure 2), aiming to illuminate (1) the seabed being recorded by the 4k-resolution action camera, immediately in front of the system, and also (2) the seabed being recorded by the live-view camera, some meters away from the system, to facilitate navigation. The diffuse flood light is produced with an LED strip (a flexible circuit board with LEDs) housed inside a transparent polypropylene tube filled with liquid paraffin (Figure 4). The LED strip, measuring approximately 2 m, is attached to the two stainless-steel curved rods on the front side of the structure using plastic cable ties (Figure 2). This LED strip is powered with a rechargeable external LiPo battery (e.g., 7.4 V, 5,300 mAh) placed on an external cylinder, which provides enough power for up to 5–6 hours of bottom time.



**Figure 5.** Live-view system connections. Layout of the connections placed on the main structure that allow the video signal to be sent from the live-view action camera to the umbilical and the powering of the camera using an external battery source.

### Data logger and lasers

To have a size reference on the images recorded, a compact laser system (Outland Technology, USA) is used, which provides a 10 cm-wide laser reflection over the seabed (Figure 4). The parallel lasers

are powered with dedicated external 9V batteries, housed on a cylinder and connected with a 2-pin Subcon connector (Figure 4). The batteries should always be disconnected once the system is on board, and based on our experience, replaced every 10–15 dives (i.e., every 20 to 30 hours). The lasers are attached below the lower camera (Figure 2) and oriented in a way that their reflection over the seabed is contained in the view of both cameras. Finally, a small logger with depth and temperature sensors tested for 1,000 m depth (Aquatec Group, UK) is attached to the structure (Figure 4) to record data at short time intervals (e.g., 1 second). It should be noted that no underwater positioning system (e.g., ultra-short baseline, USBL) is used since it would violate the guiding principle of affordability due to its high commercial prices. Alternatively, a regular GPS device is used to track the vessel's position. To generate more accurate data on the GPS position of the drift-cam system, depth values registered with the depth/temperature logger can be coupled with the GPS position of the vessel and the bathymetric chart to determine the lag existing between the system and the vessel's registered position.

### Umbilical

The distance of the Azor drift-cam with respect to the seafloor is controlled using a polypropylene rope (thickness of 10 mm; Figure 6a). Ideally, the rope selected should have some elasticity to absorb part of the oscillating movement produced by the vessel. Stretchy polypropylene ropes significantly reduce the up-and-down movement of the drift-cam over the seabed, which ultimately translates into a safer navigation and better image quality for annotation purposes. The live-view images are sent from the video transmitter to the vessel via a dual copper rubber coated wire cable (F2w 2x0.75mm) (Figure 6a). When designing the Azor drift-cam, one of the main commitments was that it could be operated over complex terrains, including heavily fished grounds. These areas have a high risk of entanglement due to the abundance of lost fishing gear and anchor lines. In the case of entanglement, some stress will be applied to the umbilical, especially during release attempts, which can damage the umbilical transmission cable. This is why this affordable cable was selected, in line with the guiding principles Affordability, Simplicity and User-friendly. The cable used has to be easy to find on the retail market, relatively light to be moved around, and at an affordable cost. At the same time, its repair has to be simple, not requiring specialised technicians on board or special tools. For these reasons, a dual core electric cable was chosen, which is believed to be the most suited solution given the guiding principles for development, instead of high-end communication cables (including fibre optic), which are more expensive and very difficult to repair at sea.

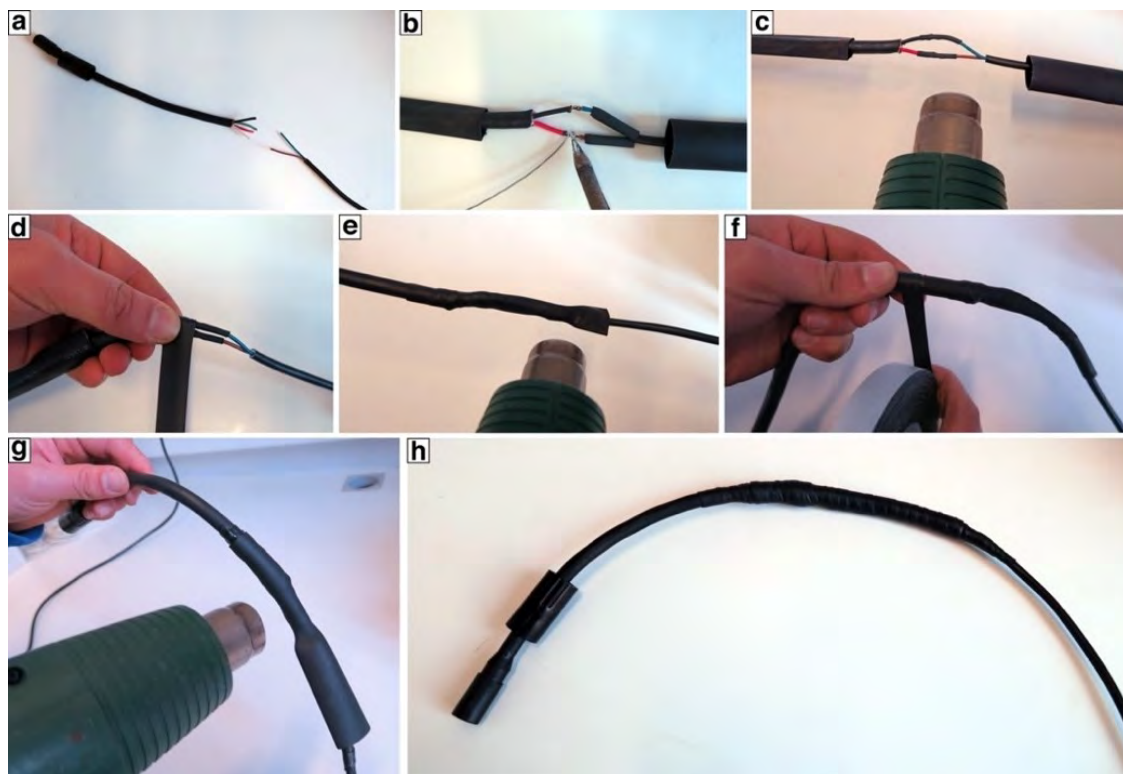


**Figure 6.** The umbilical. (a) Image of the polypropylene rope (green) and electrical cable (black) used to control the Azor drift-cam and get the video feed to the surface when underwater. (b) Attachment of the rope and the umbilical every meter or so using duct tape. (c) One of the deep-sea floats employed to always keep the drift-cam in an upright position. A thin cord is used to keep the floats in a fixed position on the rope.

In order to operate the rope and umbilical as a single unit, both cables are kept together using duct tape, placed every meter or so (Figure 6b). When attaching the electrical cable to the polypropylene rope, it is recommended to have the polypropylene rope as stretched as possible, since it will prevent

its elasticity to umbilical stretching that damages the electrical cable once underwater. Further, when attaching the cable to the rope with tape, it is important to leave some slack on the electrical cable, in a way that all the force generated by the drift-cam is only supported by the polypropylene rope. To keep the main body of the drift-cam in an upright position when submerged, 3–5 cylindrical deep-sea floats are placed on the rope/umbilical (**Figure 6c**), generally 2–3 m above the metallic structure. It is recommended to use holed floats because they have no sharp edges or loops, which is essential to prevent entanglement.

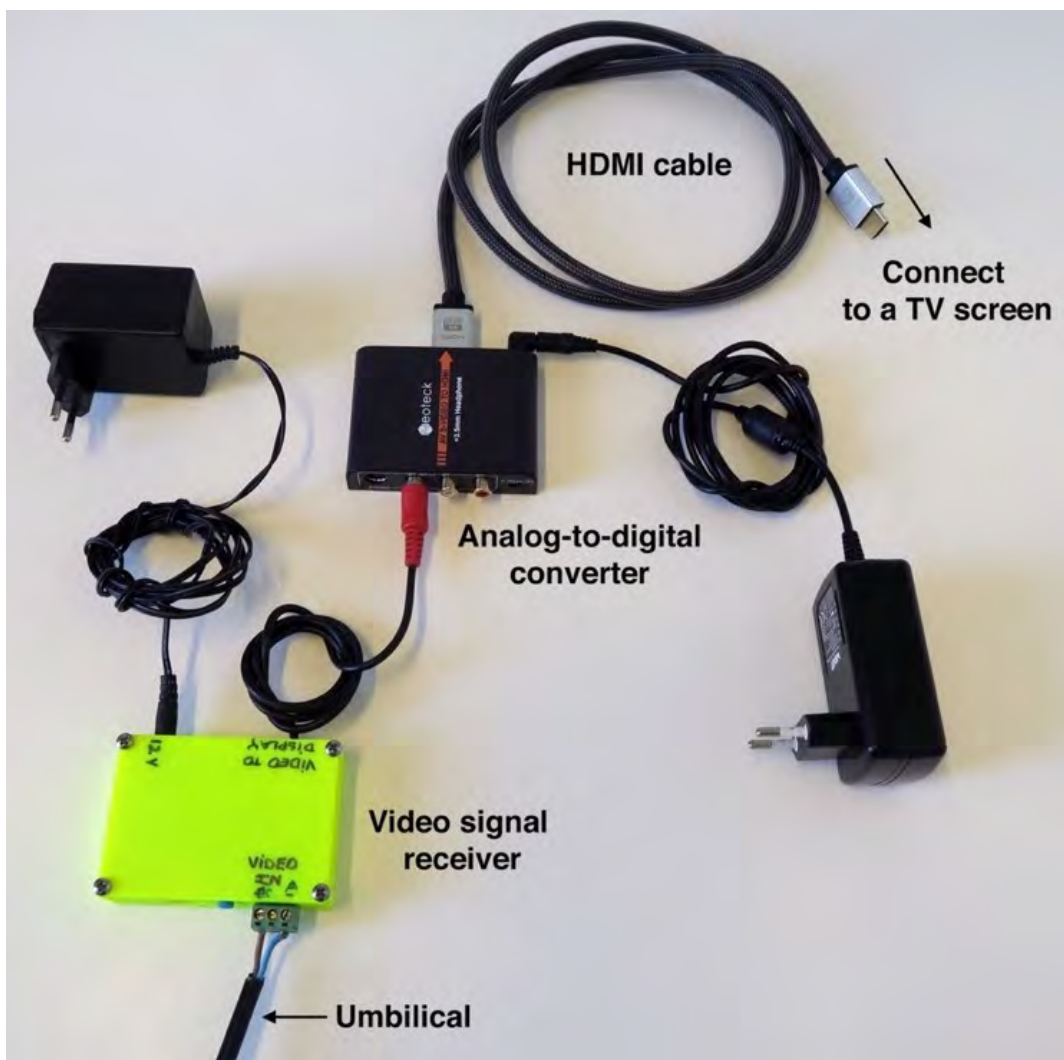
The end of the electric cable is connected to the 4-pin bulkhead placed on the main cylinder with a SubConn connector (MacArtney, Denmark). This connection is necessary to feed the live-view signal from the TX transmitter placed inside the cylinder to the copper cable of the umbilical. This is achieved by simply soldering together the cables of the SubConn connector to those of the umbilical and securing the connection with heat-shrink sleeves and vulcanising tape (process shown in **Figure 7**). After some attempts with different components and resins, it was concluded that this method is the fastest yet more secure way to join cables for the Azor drift-cam, providing excellent results at depths of up to 1,000 m. At present, none of the joints connected using this procedure have failed (loss of signal due to tension or infiltration of water). In the case of a breakage on any other part of the electric cable due excessive tension (which likely occurs during entanglements), loose ends of the electric cable can be re-joined following the same steps shown in **Figure 7**.



**Figure 7.** Step-by-step illustration of the process for joining the end part of the electric cable of the umbilical to the SubConn connector or for re-joining a broken electric cable. (a) The inner cables on both ends are peeled and cleaned. (b) Using a soldering iron and solder wire, cables are welded together. (c-g) Through a process of iteratively applying several layers of heat-shrink sleeves and vulcanising tape, the joint cables are provided with a resistant coating that prevents water from leaking, even at 1,000 m depth. (h) Final result, with the area welded presenting a larger diameter resulting from the application of multiple protection layers.

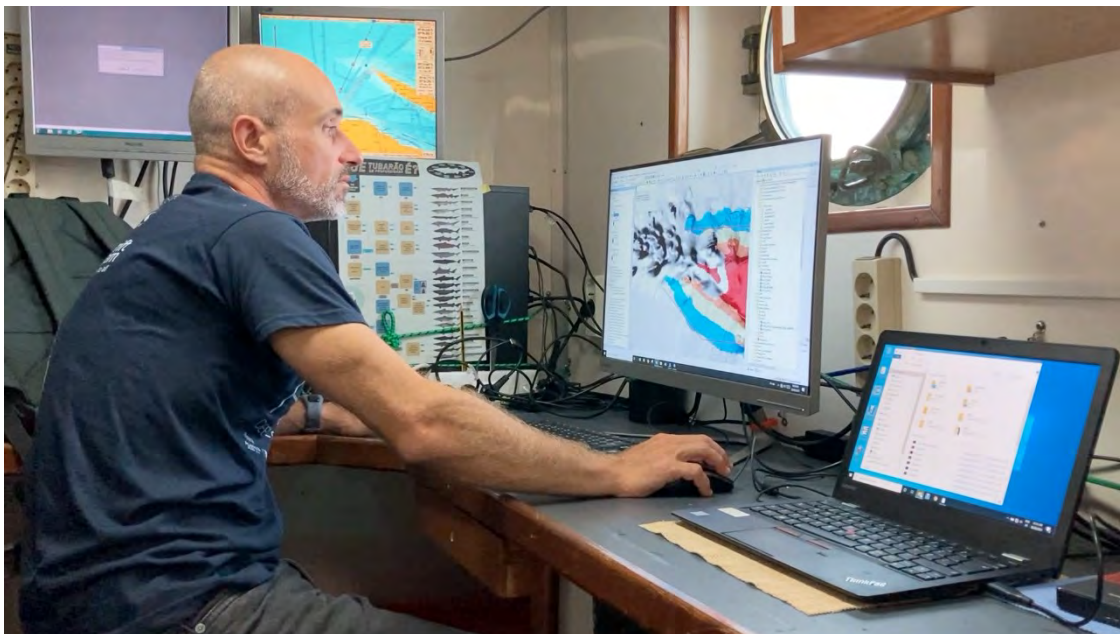
### On-vessel components

Once the video signal reaches the surface, it needs to be converted to a digital feed for screen display. This is achieved by using RX receiver and an analogue-to-digital converter (RCA to HDMI; [Figure 8](#)), which transfers the image signal to a PC monitor or TV screen via an HDMI cable. The images displayed on board are shown at a lower quality than that produced by the camera due to signal degradation along the cable, with image quality inversely related to umbilical length. In any case, the signal that reaches the surface after going through a 1,200 m long electrical cable, although in black and white and with less sharpness than the original feed produced by the action camera, is of sufficient quality for a safe underwater navigation. The images displayed on the TV monitor do not need to be recorded on board but are retrieved after the dive from the micro-SD card placed inside the live-view camera.



**Figure 8.** On-vessel components used to convert the live-view signal sent from the camera through the umbilical to a digital signal displayed on a TV screen. Both converters can be fed using 12V batteries if there is no 220V available on board.

Since all the electronic equipment mounted in the Azor drift-cam is battery fed, battery chargers are necessary to keep charging the batteries after they have been used, especially if spending more than one day at sea. If the vessel is of a small size and has no 220V available, sufficient batteries can be brought to keep replacing the ones being used. For some equipment (such as the LED lights), one battery fully charged should be enough for at least two dives. For other, less energy demanding equipment, such as the cameras and the lasers, one fully charged battery of 11.1V should last for a full day. It is advisable to have the batteries in protective bags when not being used or while charging to avoid any potential risks due to them catching on fire because of overheating or a short-circuit. To assist with the underwater dives, it is advisable to bring a computer with a Geographic Information System (GIS) software installed and connected to an external GPS antenna, which feeds the vessel's position into the computer (Figure 9). This allows the position of the vessel, and hence the drift over the seabed, to be monitored continuously. It is also necessary to bring external hard drives to copy the data recorded by the cameras, with compatible card readers to connect them to the computer.



**Figure 9.** PC mounted on board of the N/I *Arquipélago* used to centralise all the information collected with the Azor drift-cam, with the software ArcGIS used to monitor the position of the vessel using the best available bathymetry.

**Table 1.** List of components that make up the Azor drift-cam and at least one potential manufacturer/supplier for each item.

| <b>Main body</b>                                |           |   |   |
|---|-----------|---|---|
| <b>Item</b>                                     | <b>n°</b> | <b>Description</b>  | <b>Potential supplier</b>   |
| Camera housing 1 (GoPro 5/6/7)                  | 1         | Deep-sea housing used to protect the action camera that records 4K images of the seabed ( <b>Figure 4</b> )   | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Camera housing 2 (GitUp Git2P)                  | 1         | Deep-sea housing used to protect the action camera that provides live-view images of the seabed ( <b>Figure 4</b> )   | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Mountings for camera housings                   | 2         | Metallic pieces screwed to the camera housings for support ( <b>Figure 4</b> )  | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Mountings to attach camera housing to structure | 2         | Metallic cylindrical pieces to attach the camera housings to the frame ( <b>Figure 4</b> )  | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Housings for batteries and live-view elements   | 3         | Pressure-resistant cylinders that hold the live-view components, the external batteries and the voltage regulator for the live-view camera (x2; <b>Figure 5</b> ) and the battery for the LED diffuse lights (x1; <b>Figure 5</b> ) | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Adjustable Voltage regulator                    | 1         | Voltage regulator used to provide the correct voltage from the external battery to the live-view camera ( <b>Figure 5</b> )   | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| USB connectors for GitUp camera                 | 1         | Mini-USB connector to get the video feed from the camera ( <b>Figure 5</b> )  | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| Video RX/TX kit                                 | 1         | TX video transmitter to send the video feed through the umbilical to surface ( <b>Figure 5</b> ) and a RX receiver located on board ( <b>Figure 7</b> )   | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| LED lights, with batteries                      | 2         | Battery-powered lights used to illuminate the seabed to obtain good image quality and to allow a safe navigation ( <b>Figure 4</b> )  | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |
| LED strip (plus tube and resin)                 | 1         | Strip of LED lights that provide the diffuse lighting ( <b>Figure 4</b> )   | Local supplier  |
| GoPro 5/6/7 action camera                       | 1         | Action camera to obtain high-quality images of the seabed ( <b>Figure 4</b> )   | GoPro, USA<br><a href="http://www.gopro.com">www.gopro.com</a>  |
| Extra GoPro batteries & charger                 | 1         | Spare batteries used for a quick replacement between dives  | GoPro, USA <a href="http://www.gopro.com">www.gopro.com</a>   |
| GitUp Git2P action camera                       | 1         | Action camera to obtain live-view images of the seabed ( <b>Figure 4</b> )  | GitUp, China <a href="http://www.gitup.com">www.gitup.com</a>   |
| 4-pin bulkhead connector                        | 4         | 4-pin connectors used on the live-view camera housing and the cylinders that hold the batteries/live-view system ( <b>Figure 5</b> )  | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a> ; MacArtney, Denmark <a href="http://www.macartney.com">www.macartney.com</a> |
| 4-pin female-female connector                   | 2         | Cables to connect (1) the camera housing to the live-view cylinder and (2) the LED strip to the cylinder with the battery <b>Figure 4</b> – <b>Figure 5</b> )   | Group B inc., USA<br><a href="http://www.groupbinc.com">www.groupbinc.com</a>   |

Table 1: Continued

| <b>Main body</b>                     |           |   |   |
|--------------------------------------|-----------|---|---|
| <b>Item</b>                          | <b>n°</b> | <b>Description</b>  | <b>Potential supplier</b>   |
| Parallel lasers                      | 1         | Used for image scaling. Should be of small size and powered using external batteries ( <b>Figure 4</b> )  | Outland technology, USA<br><a href="http://www.outlandtech.com">www.outlandtech.com</a> |
| Depth/pressure logger                | 1         | Should have its own internal batteries ( <b>Figure 4</b> ).   | Aquatech Group, UK<br><a href="http://www.aquatechgroup.com">www.aquatechgroup.com</a>  |
| 11.1V battery                        | 4         | For live-view camera and to power the strip of LED lights. Should fit into the deep-sea housing ( <b>Figure 4, Figure 5</b> )   | Gens ace, Germany<br><a href="http://www.gensace.de">www.gensace.de</a>                 |
| Stabilising wing                     | 1         | Made of methacrylate to avoid bending through time ( <b>Figure 3</b> )  | Local supplier  |
| Steel structure                      | 1         | Should be made of stainless steel ( <b>Figure 3</b> ). It contains several pieces to attach all the electronic components to the central bar.   | Local supplier  |
| Weight                               | 1         | Round shape. 15–20 kg (based on boat and sampling depth; <b>Figure 3</b> )  | n/a   |
| <b>Umbilical</b>                     |           |   |   |
| <b>Item</b>                          | <b>n°</b> | <b>Description</b>  | <b>Potential supplier</b>   |
| Polypropylene rope, Ø 10 mm, 1,200 m | 1         | Three- or four-strand rope ( <b>Figure 6</b> ). Should have the least number of breaks as possible. If it does not come in a single piece, ropes should be joint using short splices to allow a better pass through the winch | Local supplier  |
| Dual core electrical cable, 1,200 m  | 1         | F2w 2x0.75 mm ( <b>Figure 6</b> ). Should come as one single cable to reduce signal loss due to cable joints  | Local supplier  |
| Deep-sea floats                      | 3–4       | Should have a central hole large enough to fit rope and cable. Maximum working depth of at least 1,000 m ( <b>Figure 6</b> )  | Engel Netze, Germany<br><a href="http://www.engelnetze.com">www.engelnetze.com</a>      |
| Swivel                               | 1         | Stainless-steel swivel large enough to fit the 10 mm rope ( <b>Figure 3</b> )   | Local supplier  |
| Large bucket for umbilical storage   | 1         | Large enough to fit the whole length of the umbilical (rope + cable). Should be holed at the bottom to avoid water accumulation   | Local supplier  |
| <b>Live-view components</b>          |           |   |   |
| <b>Item</b>                          | <b>n°</b> | <b>Description</b>  | <b>Potential supplier</b>   |
| RCA to HDMI adapter                  | 1         | Converts video signal from analogic to digital ( <b>Figure 8</b> )  | NeoTeck, <a href="http://www.neoteck.cn">www.neoteck.cn</a>                             |
| HDMI to HDMI cable                   | 1         | Standard HDMI cable ( <b>Figure 8</b> )   | Local supplier  |
| TV screen                            | 1         | Any HD PC monitor/TV screen with an HDMI connection   | Local supplier  |

Table 1: Continued

| <b>Consumables</b>                              |           |   |  |
|---|-----------|---|--|
| <b>Item</b>                                     | <b>n°</b> |   | <b>Potential supplier</b>  |
| 2 Tb hard drive (can store up to 50 dives each) | 2         | External hard drive to store images and data produced. 2 Tb should be enough for 40–50 dives. A back-up disk is highly recommended  | Toshiba, WD, Seagate, Lacie  |
| GPS   | 1         | A portable GPS to record position data at 1-sec intervals. Spare rechargeable batteries are highly recommended  | Garmin, USA <a href="http://www.garmin.com">www.garmin.com</a>                     |
| microSD card                                    | 2         | Memory cards used for action cameras. 64 Gb should be enough for live-view camera and 128Gb for 4K camera. Spares are recommended to speed up intervals between dives                               | SanDisk, USA<br><a href="http://www.westerndigital.com">www.westerndigital.com</a> |
| Batteries for parallel lasers, 9V               | 2         | Any 9V batteries. Spares are highly recommended   | Local supplier   |
| Tool box  | 1         | A box with the basic tools to be taken on board   | Local supplier   |
| Other consumables                               | n         | A set of general consumables that could be required on board (cable ties, electrical tape, duct tape, screws of different sizes, o-rings, heat-shrink sleeves, vulcanising tape, solder wire, etc.) | Local supplier   |



### 3.4 Use at sea

#### Equipment preparation

The simplicity of the Azor drift-cam makes equipment preparation relatively easy and fast prior each deployment, with only a few electronic devices mounted on the structure that need to be prepared: data logger, cameras, and external batteries. The data logger used in the Azor drift-cam is a simple depth/pressure sensor, which has to be activated before the dives in order to start collecting data. It is important that the logger is synchronised with the time of the main PC and set to collect depth/pressure values every second, so data can then be easily merged with the positioning data. The Azor drift-cam uses two cameras, an HD action camera connected to live view (Git2 from GitUp) and a 4k action camera used to collect higher-resolution images (GoPro Hero 5, 6 or 7 models). Both cameras should have their internal microSD cards emptied (to maximise the amount of footage recorded) and formatted according to the model of the camera before the dive. It is important that the time on the cameras is set with the main PC to have all data collected synchronised to the same second. The settings of both cameras should be configured to maximise the quality of the images collected:

- GitUp Git2 camera: HD resolution (1920x1080 pixels), fisheye for a wider field of view that helps during navigation, auto-start on, date/time displayed on the image.
- GoPro Hero 5, 6 or 7 camera: 4k or 2.7k resolution, 25fps to increase the amount of light going in the sensor, linear field of view to avoid image distortion.

Just a few minutes before the vessel is in position to start the deployment, the logger can be mounted on the steel structure and the cameras placed inside the camera housings (Figure 10a). The three external batteries can then be placed inside the cylinders to power feed the HD camera, the live-view system, the LED lights, the LED strip and the lasers. Once all items have been connected, the live-view images from the HD camera should start being displayed on the main screen used to assist during underwater navigation.

#### Deployment, underwater image collection and recovery

Due to its low overall size and weight, the Azor drift-cam can be deployed by hand from the side of a small vessel without the need of a crane, although its use is recommended (Figure 10b-c). Once the system is released from the crane, it freely descends through the water column towards the seabed at speeds of 1.4–1.5 m/s (that is, it reaches 500 m depth in about 5–6 mins). The images provided by live-view system indicate when the system has reached the seabed but placing marks on the umbilical (in the form of colour bands every 50 or 100 m) are recommended to provide a rough estimate of how far the system is from the seabed at every moment. Once the system reaches the bottom, some cable must be recovered until the weight of the structure is lifted from the seabed. At this moment, the system is left to cruise over the seabed following the vessel's drift. The distance between the weight and the seafloor is controlled via a hydraulic winch, giving, or recovering umbilical as required. Indications to winch operators based on the live feed provided by the navigation camera are given by an observer watching the images on a screen placed on the vessel (Figure 11b). The live feed, although in black and white and with less sharpness than the original feed produced by the action camera, is of sufficient quality for a safe underwater navigation. A GPS device is used to record the vessel's position at short time intervals (1–2 s), and data collected by the temperature/depth sensor can be later used to adjust the position of the system over the seabed.

At the end of the dive, generally after 1 hour of cruising over the seabed, the system is recovered back to the vessel using the hydraulic winch, letting the umbilical accumulate inside a large bucket (Figure 11b). Recovery time depends on winch capacity and speed, but a limit of 0.7–0.8 m/s is recommended to avoid damaging the electrical cable during the operation. At this pace, the system can reach the surface from 500 m depth in about 12–14 min. Once on deck, the time required to prepare the equipment for

the following deployment is short, since only SD cards and batteries for lights/cameras must be replaced. It is recommended to have at least one set of spare batteries for each electronic device, replaced every 3–4 deployments.



Figure 10. (a) Two members of the Azores deep-sea research group setting up the cameras and batteries of the Azor drift-cam prior to its deployment. (b–c) Deployment of the Azor drift-cam from the side of the vessel using a crane to lift the system.

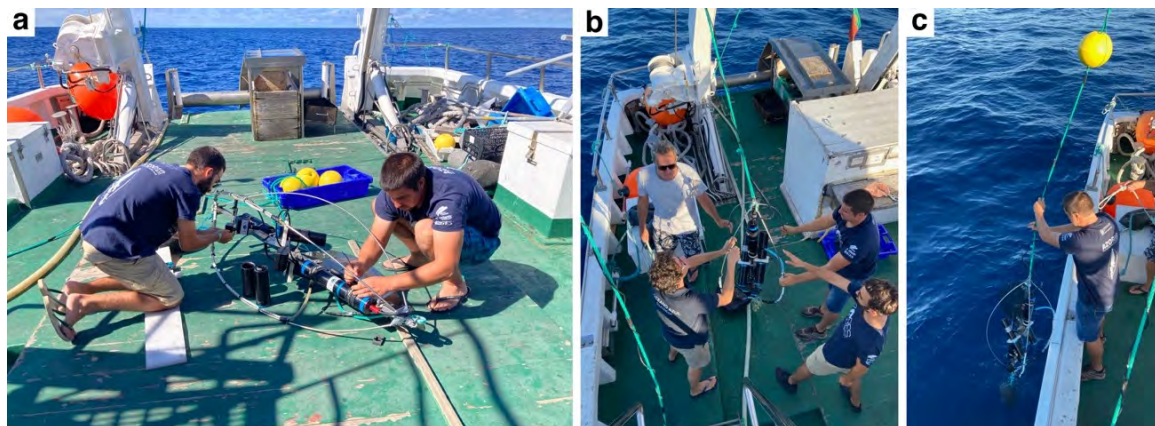


Figure 11. Simplicity in the use of the Azor drift-cam. (a) An observer providing indications to the winch operators based on the images provided by the live-view system. (b) Winch operators controlling the amount of cable required to keep a safe distance between the drift-cam and the seabed. The excess cable is stored in a large bucket below the winch.

Image data produced

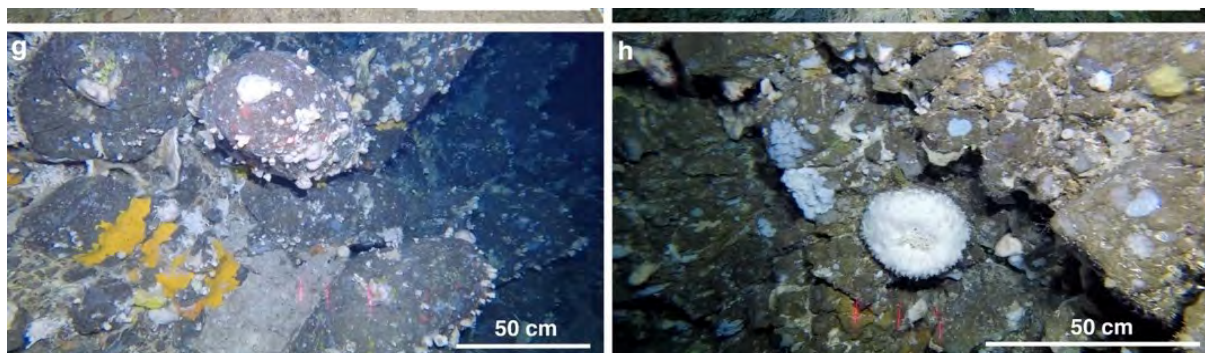
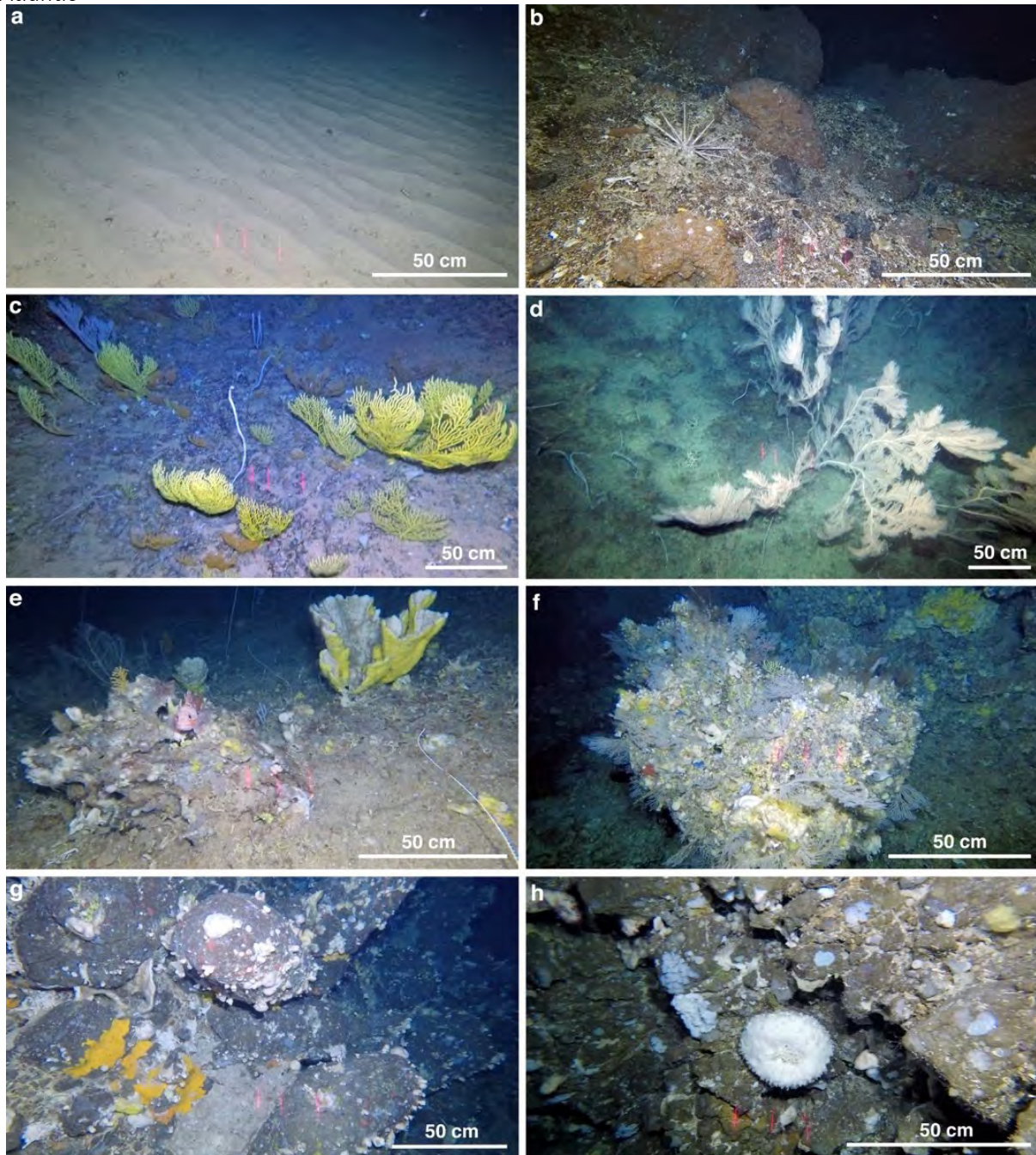


Figure 12. More images can be viewed in the YouTube Channel of the Azores Deep-Sea Research group, which will be updated regularly to add new images collected in seamounts of the Azores<sup>2</sup>. The presence of parallel lasers on-screen allows for species counts to be transformed into density measures. This can

<sup>2</sup> <https://www.youtube.com/channel/UCrUCCK9866Ym8vq7ZwwZoQ>

be easily achieved by generating an estimate of the swept surface using the computed average field of view and the distance travelled over the seabed provided by the GPS track.



**Figure 12.** Screen captures obtained from the video footage recorded with the action camera mounted on the Azor drift-cam during the MapGES 2019 cruise, illustrating some of the different seafloor substrates that were surveyed. The small red stripes in the mid/low part of the image correspond to the reflection of the parallel lasers over the seabed, used for image scaling. The distance between the two marks on the left corresponds to 10 cm. Scale bars = 50 cm. (a) Sand ripples in Monte Alto seamount, 460 m depth. (b) Gravels and boulders in Cavala seamount, 650 m depth. (c) Flat substrate in SE of Pico Island, 400 m depth. (d) Sloping substrate in Picoto seamount, 500 m depth. (e) Mixture of sand, gravels and scattered rocks in Gigante seamount, 470 m depth. (f) Large rocky outcrops in Voador seamount, 390 m. (g) Large boulders on a very steep terrain in A3 seamount, 560 m depth. (h) Vertical wall in A3 seamount, 500 m depth.

The oscillating movement of the vessel over the sea is generally transmitted throughout the umbilical to the Azor drift-cam, with the magnitude of the up-and-down movement observed in the images largely depend on wave height and position of the vessel against the waves. If the type of rope selected has some flexibility, some of the oscillation produced by the waves can be absorbed, improving image quality, and facilitating navigation. Due to the oscillating movement and differences in depth along the dive, the field of view (FOV) of the video image experiences certain variability. For any given dive, the FOV is generally smaller when the system is moving up the slope and will be much larger when going through flat terrains or downslope. Based on the analysis of 475 images randomly obtained from the footage recorded by the Azor drift-cam in 2019 with the 4k camera, the FOV of the images obtained ranged between 1.5 and over 6 meters, displaying an average value of 3.5 m (

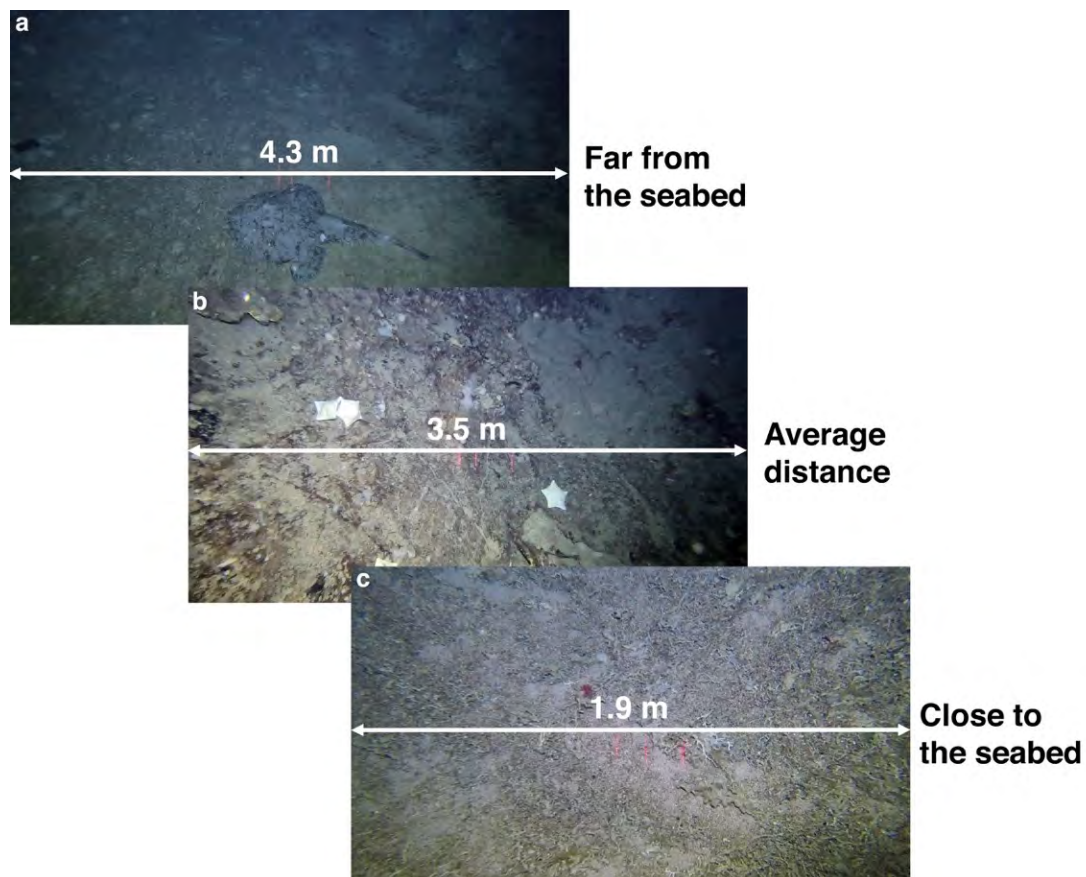


Figure 13).

**Figure 13.** Differences in the field of view (FOV) of the main camera depending on the distance between the system and the seabed.

### 3.5 Performance assessment

The performance of the Azor drift-cam was assessed at two different levels. First, the metadata obtained from the different surveys performed at sea during the past 4 years was analysed to determine whether the design of the Azor drift-cam is suited for the exploration of deep-sea benthic habitats, indicating its strengths and limitations. Second, the imagery data obtained in the MapGES 2019 cruise along the Mid-Atlantic Ridge was analysed to determine whether it is of sufficient quality to successfully characterise deep-sea benthic habitats and their associated biological diversity.

#### Four years of surveys in the Azores

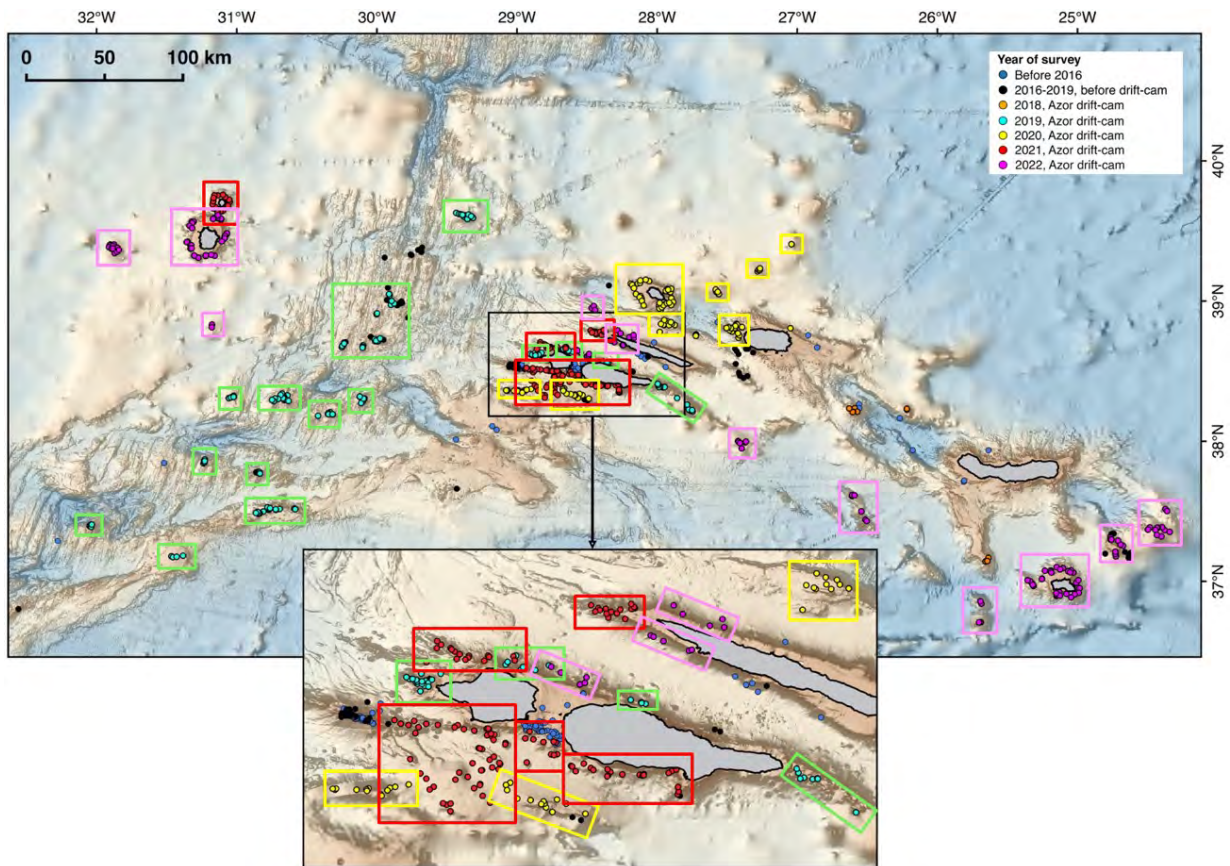
The capacity of the Azor drift-cam to collect underwater images of the seabed was assessed from the information collected during the MapGES cruises carried out in the Azores during the past 4 summers

(2019 to 2022). The surveys were performed on board of N/1 *Arquipélago*, a 25 m-long research vessel from the Government of the Azores, as well as on board of 4 different local fishing vessels (lengths between 7 and 13 m). In general, smoother seafloors were easier to survey than rough terrains, especially if found on positive slopes. The type of substrate encountered did not seem to exert a strong influence on the quality of the images obtained, with sandy bottoms similarly surveyed to mixed substrates with rocks and boulders or areas dominated by large rocky outcrops. Very steep slopes and large vertical walls, however, were the exception since the Azor drift-cam had to be lifted at high speeds with the hydraulic winch to avoid entanglement, which significantly limits image quality.

### Effectiveness

In the last 4 years (2019–2022), 535 successful dives have been performed with the Azor drift-cam in the Azores region, from a total of 557 dives attempted (Figure 14). The approximately 20 dives that had to be aborted before the system reached the seabed responded to several reasons, including the live-view system giving problems during the descent, the direction of the drift changing towards deeper areas or the surface current being too strong to keep the Azor drift-cam in a vertical position, hence limiting a safe navigation. Considering that the surveys with the Azor drift-cam were only done during daily hours (no night shifts), the 500+ successful dives performed in the Azores in 100 days of work at sea represented an average of  $5.04 \pm 1.54$  dives per day. As shown on Figure 14, these 500+ dives cover a very wide latitudinal (37.15 to 39.65 °N) and longitudinal gradient (29.30 to 32 °W), with varying sampling efforts in the different areas surveyed, which was based on seamount size and ecological relevance. So far, 15 seamount-like structures have currently been visually evaluated with the Azor drift-cam along the Mid-Atlantic Ridge, as well as other 17 seamounts close to the islands of the western (e.g., Cachalote, Diogo de Teive), central (e.g., Baixo de São Mateus, Condor de Fora) and eastern groups (e.g., Margrette, Formigas). 303 dives have also been performed on the slopes of the islands of the western (48 dives), central (225) and eastern groups (30), except for the island of São Miguel, yet to be surveyed.

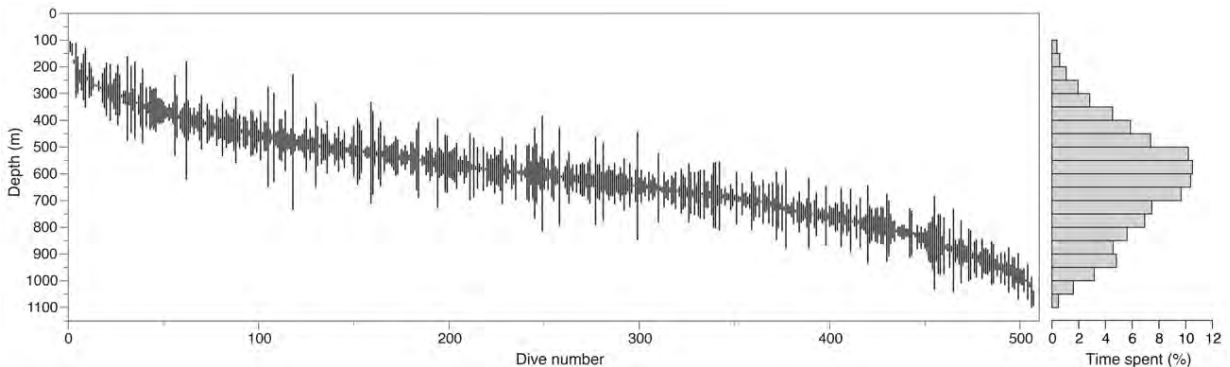
More than 295 km of seafloor between 50 and 1,080 m depth have been surveyed in the past 4 years, adding up to approximately 450 hours of bottom time (Table 2). The time travelled over the seabed averaged 50.61 min per dive, totalling  $4.55 \pm 1.42$  hours of bottom time recorded on average each day at sea. This represents an average distance over the seabed of 562 m per dive, which adds up to an average of roughly 3 km of seafloor explored per day. The average deployment speed was  $1.37 \pm 0.41$  m/s, which represents descending 500 m along the water column in approximately 6 minutes. The recovery speed was much lower, averaging  $0.70 \pm 0.38$  m/s (that is, 500 m of cable recovered in 12 minutes). The average time at surface between dives was generally kept low, at around half an hour. The data collected for the different surveys (Table 2) show that the number of dives per day remained relatively stable through the years (around 5.5), with the average dive time increasing from 45 minutes in 2019 to almost 1 hour in 2022. The time spent at surface preparing the equipment for the following deployment showed a significant decrease through the years, from 41 minutes in 2019 to just over 23 minutes in 2022. The maximum depth reached by the system also increased every year, mainly due to improvements in cable configuration and the confidence in underwater navigation. In 2019, the Azor drift-cam was deployed to a maximum depth of 850 m, and that depth increased to reach almost 1,100 m in 2022. Figure 15 shows the depth range of each of the 535 underwater dives performed with the Azor drift-cam in the past 4 years, ordered by their average depth. Although the dives were planned to explore all depth ranges within each geomorphological structure to obtain a representative image of the different habitats present, a large proportion of the time spent underwater was allocated to depths between 400 and 700 m, with the shallower and deeper strata proportionately less explored.



**Figure 14.** Location of the 535 dives performed with the Azor drift-cam inside the Azores EEZ during the MapGES surveys of the past 4 years (2019 to 2022). The dives are colour coded according to the year of the survey. Green: 2019; Yellow: 2020; Red: 2021; Pink: 2022. The map also includes the dives from 2018 (in orange) when the system was still a prototype in early development and was used as a drop-down camera. The blue and black dots correspond to the dives performed with ROVs and towed camera system before the development of the Azor drift-cam and for which the ADSR group has access to the images.

**Table 2.** Summary table with information on the dives performed with the Azor drift-cam inside the Azores EEZ during the MapGES cruises performed between 2019 and 2022.

| Year         | n° of dives | Av. n° dives/day | Av. bottom time (min) | Total bottom time (h) | Avg. dive length (m) | Total dive length (km) | Av. surface time (min) | Depth range(m) |
|--------------|-------------|------------------|-----------------------|-----------------------|----------------------|------------------------|------------------------|----------------|
| 2019         | 155         | 5.74±1.85        | 44.96                 | 116.18                | 565.59               | 87.22                  | 41.10                  | 108–848        |
| 2020         | 100         | 5.00±1.83        | 44.36                 | 73.90                 | 541.65               | 52.54                  | 32.08                  | 86–915         |
| 2021         | 146         | 5.21±1.47        | 54.30                 | 133.03                | 575.24               | 83.29                  | 27.16                  | 48–1022        |
| 2022         | 134         | 5.58±1.02        | 57.83                 | 128.18                | 559.19               | 72.13                  | 23.66                  | 192–1086       |
| <b>Total</b> | <b>535</b>  | <b>5.40±1.58</b> | <b>50.61</b>          | <b>451.31</b>         | <b>562.26</b>        | <b>295.18</b>          | <b>31.15</b>           | <b>48–1086</b> |



**Figure 15.** Depth range of the dives performed with the Azor drift-cam inside the Azores EEZ between 2019 and 2022 part of the MapGES cruises. The bar plot on the right side indicates the amount of time spent by the Azor drift-cam (in %) surveying each depth range at 50-m intervals.

**Affordability**

The cost to construct an Azor drift-cam, considering all the necessary spare parts, falls within the predefined 10–15k € price range that was set when starting to develop the first prototype. This value is two orders of magnitude lower than that of standard working-class ROVs capable of reaching 1,000 m depth (600 k to 3.5 million €<sup>8</sup>) but falls within the price range of low-cost shallow-water ROVs<sup>15</sup> and towed video systems<sup>16</sup>, and also other low-cost deep remote baited camera systems. Keeping the Azor drift-cam affordable implied that some useful add-ons were left out, always without compromising its general capabilities and goals. Some of the add-ons that could be very useful to improve the functionality of the system, but that would significantly increase the final price, would be (a) the use of fibre optic in the umbilical for a better image quality received at surface, (b) the use of an ultra-short baseline acoustic positioning system (USBL) to generate accurate positioning data when the system is underwater, (c) the use of a high-end CTD profiler to better characterise the environmental conditions at the seabed, and (d) the use of extra LED lights for better image quality, among others.

The Azor drift-cam can be deployed from small local vessels, including fishing boats equipped with a hydraulic winch. Not having to rely on large oceanographic vessels to carry out deep-sea exploration significantly reduces the operational costs associated to its use. Another advantage of employing a relatively low-cost imaging tool is that greater risks can be taken when exploring geomorphologically complex areas or heavily fished grounds, in opposition to other, more expensive imaging systems that generally take great precautions when moving underwater to avoid potential entanglements.

**Simplicity**

The Azor drift-cam resulted in a modular, simple, light, and easy-to-assemble video platform composed of 'off-the-shelf' components, all of which can be easily acquired locally or from online distributors. Since all pieces are not inter-related but function independently, the system has a series of advantages with respect to more complex devices. First, the Azor drift-cam is fully adjustable and easy to reconfigure. This means that it can be easily adapted to the specific needs of each survey, with more components being added if necessary. Also, the electronic equipment can be easily replaced for newer versions without changing the original configuration. Second, the malfunctioning of an electronic component does not affect the functioning of the remaining parts. This in turn allows for a quick identification of the potential problem by isolating each of the parts and testing their operability. Third, the replacement of any damaged component is very straightforward, further reducing the time needed for repair. And finally, most types of malfunctioning can be repaired on board of the vessel without the need of specific tools and complex testing. Most incidents that generally occur at sea (e.g., cable damage, connections not working) can generally be fixed while at sea without the need of coming back to shore, further reducing the economic costs associated to its use.

**User-friendly**

The Azor drift-cam has always been operated by researchers and technicians of the ADSR group without the assistance of highly qualified engineers. Since the design and development of the system was led by natural scientists rather than electronic engineers, all the decisions taken along the way searched for the simplest solution, reaching a prototype that can be mounted, operated, and repaired by any member of a research team with basic engineering expertise. Learning how to use the system is generally a fast process, and only requires the familiarisation with the different electronic components and the most common typed of incidences that can occur during the work at sea (e.g., cable breakages, bad connections, misconfigurations of the cameras). This approach makes the Azor drift-cam ideally suited for scientists without access to specialist engineering support.

**Resiliency**



The main risk when operating deep-sea video platforms attached to the support vessel is the entanglement of the umbilical on large rocky outcrops, vertical walls or abandoned fishing gears, which could damage some of the equipment, or even lead to the loss of the device. Longlines are the most common type of abandoned fishing gear encountered in the Azores and were observed in a large percentage of the dives performed with the Azor drift-cam during the MapGES cruises, especially in shallow seamounts. Longlines lying flat over the seabed could be easily avoided by pulling the frame upwards through the water column, with visual navigation proving sufficient to detect such hazards. Longlines found suspended several meters over the ground are not detectable in the video feed and are almost impossible to avoid when cruising close to the seafloor. In these situations, the Azor drift-cam might temporarily get caught, in a situation like that found when encountering large vertical walls, since the recovery speed of the winch not always sufficiently fast to overcome the hazard. The oval shape of the metal frame, in conjunction with rapid recovery/release of the cable and manoeuvring the vessel away from the hazard, has proven the best practice to free the system from entanglements. During the MapGES cruises of the past 4 years, the system was successfully released from more than 15 entanglements, most of which were longlines. In those cases, no loss of equipment was reported, with the only damage related to breakages in the umbilical (electrical cable). It should be noted, however, that three complete systems were lost due to entanglements in fishing lines or vertical walls during the MapGES cruises, one per year between 2019 and 2021.

### **Operability**

The simplicity in the use of the Azor drift-cam represents one of its main advantages with respect to commercial ROVs or towed camera systems. The Azor drift-cam can be released manually from the side of the vessel and requires only a hydraulic winch to be operated. It can be used in weather conditions similar or worse to those required for ROV operations, incrementing the number of potential working days. Its fast deployment and recovery allow for a large number of dives to be performed per day, which can help increase the spatial coverage at a reduced cost.

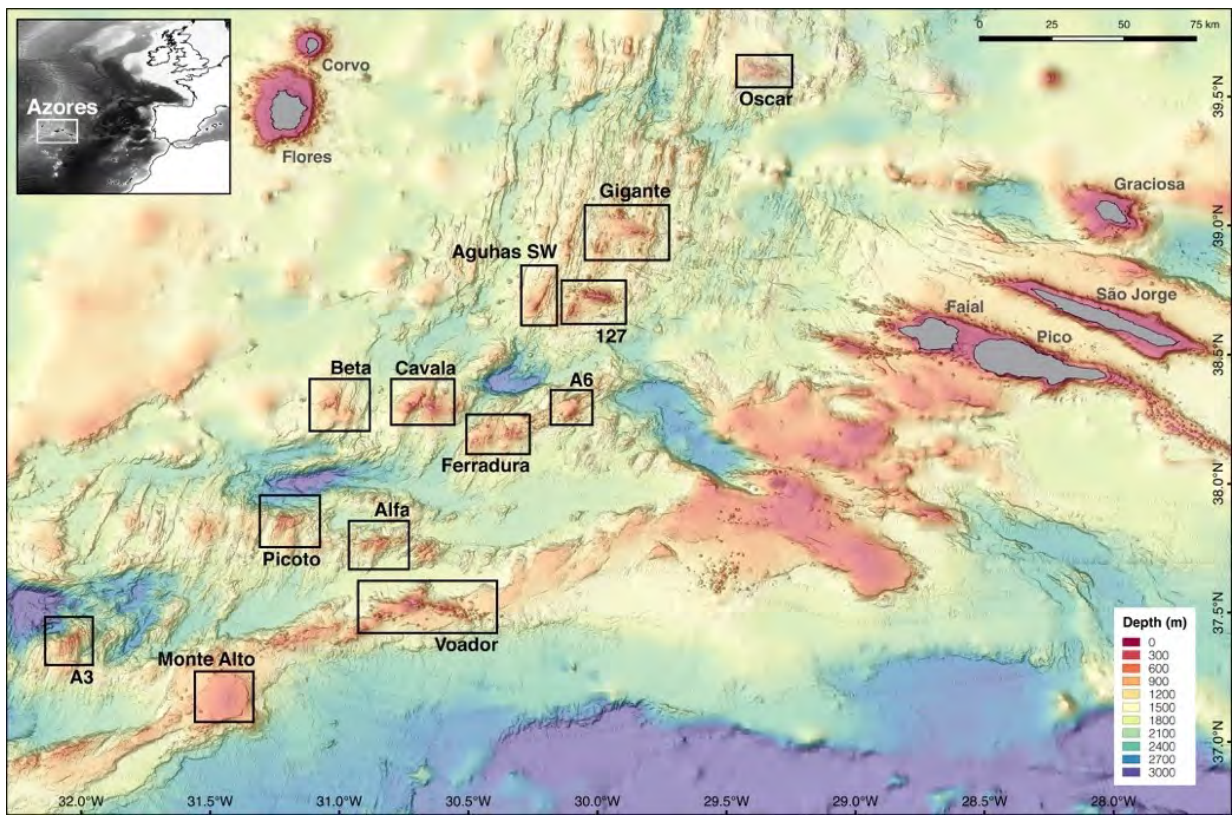
During the past 4 years, the Azor drift-cam has been successfully operated from a 25 m-long research vessel (N/I *Arquipélago*) and from four local fishing vessels (7 to 12 m in length) based in the islands of Faial, Graciosa and Corvo. Although the largest part of the dives (463, 86%) were performed on board of N/I *Arquipélago*, up to 72 (14%) have now been executed from small fishing vessels. The largest number of dives on board of a fishing vessel was achieved in the island of Graciosa (MapGES cruise 2020), with 36 deployments in 8 days of work at sea. During this survey, over 25 hours of bottom time were recorded at depths between 298 and 771 m, exploring over 17 km of seabed from the slope and small seamounts around the island.

### **Portability**

Most parts of the Azor drift-cam are easily packed and transported. The frame, the housings, the action cameras, the lights, the lasers and the sensors can be shipped to any part of the world either by land, air or sea. The International Air Transport Association (IATA) Dangerous Goods Regulations, however, restricts the shipment of certain types of lithium cells and batteries, which may have to be shipped in advance by sea or land. The umbilical, composed of polypropylene rope and electric cable, weighs over 100 kg, making shipment by air expensive. These materials are found everywhere and therefore a new umbilical can be built upon arrival. During the surveys made in the Azores, the whole system has been moved twice, first using a regular passenger ferry from the island of Faial to Graciosa (2020), and then using a small cargo ship (2021) to be moved to the island of Corvo. In both surveys, the system was successfully operated from local fishing vessels to collect new images of the seabed down to depths below 950 m.

### 3.6 Geological and biological data obtained

Several images recorded with the Azor drift-cam were evaluated in further detail in order to assess the potential use of the footage obtained with this system to characterise deep-sea benthic habitats and their associated biological communities. Information was extracted from the images at 3 different levels: the behaviour of the Azor drift-cam when cruising over the seabed, the composition of the dominant substrate and the number of benthic species identified. For this exercise, 112 dives performed in 15 seamounts and ridges along the Mid-Atlantic Ridge performed during the MapGES cruise in 2019 were evaluated (Figure 16;



**Table 3).** The amount of underwater video footage analysed added up to 84 hours of bottom time. Sampling effort was unevenly distributed in the different areas. For instance, the two seamounts with the highest sampling effort were Cavala (15 dives) and Voador (16), both with more than 11 hours of video footage recorded. On the other hand, the seamounts least explored had less than 3 hours of bottom time each, and corresponded to Gigante Agulhas NE (2 dives), Gigante Agulhas S (3 dives) and Monte Alto (4 dives). The depth range explored was also variable across sites, although generally found between 400 and 700 m, but reaching depths below 800 m during the deepest dives.

**Figure 16.** Location of the sampling areas explored with the Azor drift-cam along the Mid-Atlantic Ridge during the MapGES cruise 2019 on board of *N/I Arquipélago*.

**Table 3.** Number of dives performed in each sampling area along the Mid-Atlantic Ridge during the MapGES 2019 cruise, with information on the depth range, average length, average dive time and total footage recorded in each seamount. Area (km<sup>2</sup>) refers to the extent in km<sup>2</sup> of the geomorphological feature above 1,000 m depth inside the sampling area.

| Seamount     | Surface            | Dives      | Depth (m)      | Av. dive   | Av. bottom      | Total bottom    |
|--------------|--------------------|------------|----------------|------------|-----------------|-----------------|
|              | (km <sup>2</sup> ) |            | min–max        | length (m) | time (hh:mm:ss) | time (h)        |
| 127 Smt      | 75                 | 7          | 235–563        | 525        | 0:41:30         | 4:50:28         |
| A3           | 104                | 6          | 504–687        | 624        | 0:40:22         | 4:02:15         |
| A6           | 152                | 7          | 484–848        | 519        | 0:38:45         | 4:31:17         |
| Alfa         | 107                | 6          | 509–672        | 398        | 0:41:38         | 4:09:50         |
| Beta         | 181                | 5          | 562–745        | 813        | 1:03:09         | 5:15:44         |
| Cavala       | 274                | 15         | 367–779        | 565        | 0:47:05         | 11:46:17        |
| Ferradura    | 191                | 7          | 521–816        | 692        | 0:57:55         | 6:45:25         |
| Gigante      | 324                | 7          | 187–797        | 608        | 0:34:20         | 4:00:19         |
| Gigante NE   | 82                 | 2          | 588–700        | 431        | 0:58:49         | 1:57:38         |
| Gigante S    | 63                 | 3          | 502–684        | 481        | 0:47:43         | 2:23:08         |
| Gigante SW   | 131                | 7          | 460–640        | 620        | 0:47:07         | 5:29:51         |
| Monte Alto   | 767                | 4          | 414–584        | 640        | 0:38:11         | 2:32:42         |
| Oscar        | 192                | 15         | 562–724        | 438        | 0:38:38         | 9:39:35         |
| Picoto       | 90                 | 5          | 528–731        | 781        | 1:01:59         | 5:09:56         |
| Voador       | 1087               | 16         | 278–692        | 548        | 0:43:23         | 11:34:15        |
| <b>Total</b> | <b>3820</b>        | <b>112</b> | <b>187–848</b> | <b>567</b> | <b>0:45:05</b>  | <b>84:08:40</b> |

The first step to evaluate the utility of the images recorded to provide an accurate characterisation of the seabed corresponds to the analysis of the behaviour of the Azor drift-cam while surveying the seabed. This analysis identifies the sections of the video footage that can be used for annotation purposes, and were defined as follows:

- Stopped over the seabed, when the Azor drift-cam is kept stationary in the same position. This generally occurs at the beginning of the dive or when the system gets caught on a fishing line.
- Linear transect, when the camera of the Azor drift-cam moves on a forward-facing position over the seabed.
- Moving sideways, when the Azor drift-cam has its camera turned towards its side (generally due to the effect of bottom currents), but the system keeps moving forward.
- Erratic movement, when the Azor drift-cam does not follow a linear trajectory and the camera keeps moving erratically from one side to the other.
- Off bottom, when the Azor drift-cam cruises too far from the seabed for any information to be collected from the images.

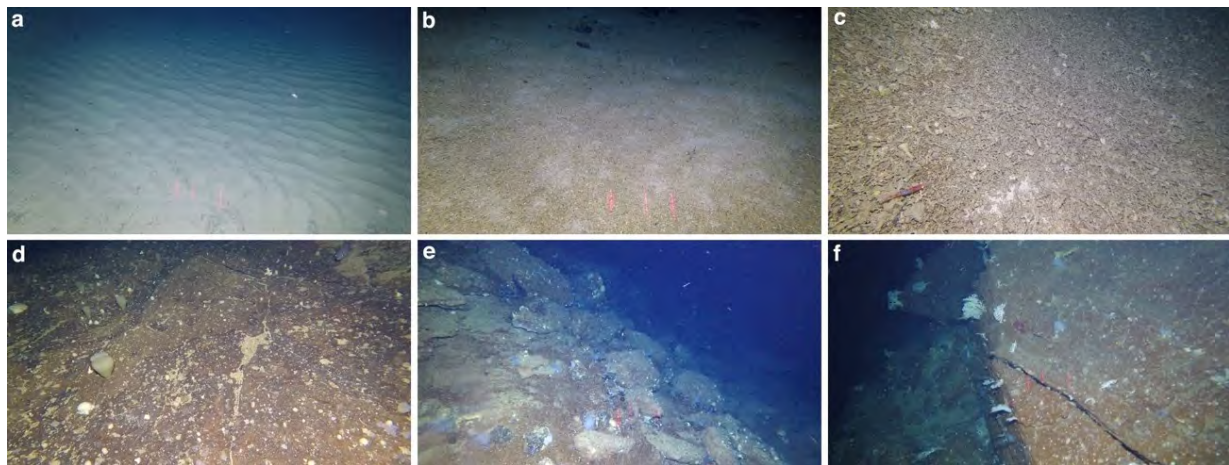
The behaviour of the Azor drift-cam in the images recorded during the MapGES 2019 cruise is provided in

**Table 4.** The percentage of footage considered 'off-bottom' remained below 10% in almost all dives, with an average of 5%. This result indicates that a large percentage of the images collected by the Azor drift-cam are useful for annotation, and that the system is suited to survey benthic habitats of the deep sea. In fact, most of the footage collected corresponded to 'linear transect', in some dives even reaching 90% of the time spent at bottom. This footage can be considered of the best quality for image annotation, and would allow for species occurrences to be converted into density values. The second category with the highest percentage (20%) was 'erratic movement', which consists in those situations where a clear path over the seabed cannot be determined. In those cases, the identification of benthic species to characterise the diversity of the area is still useful, although the data extracted have less potential to be converted into quantitative measures.

**Table 4.** The behaviour of the Azor drift-cam over the seabed given for each sampling area along the Mid-Atlantic Ridge during the MapGES 2019 cruise. Values are provided as percentage of time considering only the bottom time, which spans from the moment the system reaches the seabed until it starts the ascent back to the vessel.

| Seamount       | Behaviour over the seabed (%) |             |             |            |            |
|----------------|-------------------------------|-------------|-------------|------------|------------|
|                | Stopped                       | Erratic     | Linear      | Sideways   | Off bottom |
| 127 Smt        | 15.1                          | 22.3        | 61.0        | 0.0        | 1.7        |
| A3             | 1.3                           | 26.1        | 63.2        | 0.0        | 9.4        |
| A6             | 9.7                           | 21.0        | 47.7        | 3.3        | 18.3       |
| Alfa           | 6.5                           | 31.3        | 53.2        | 0.0        | 9.1        |
| Beta           | 2.4                           | 23.4        | 69.6        | 0.0        | 4.6        |
| Cavala         | 12.0                          | 25.4        | 57.2        | 0.0        | 5.5        |
| Ferradura      | 3.5                           | 18.6        | 72.2        | 0.0        | 5.8        |
| Gigante        | 1.7                           | 21.4        | 75.2        | 0.0        | 1.7        |
| Gigante NE     | 0.8                           | 5.6         | 90.0        | 0.0        | 3.6        |
| Gigante S      | 3.3                           | 18.3        | 71.8        | 0.0        | 6.6        |
| Gigante SW     | 1.1                           | 14.1        | 79.0        | 0.0        | 5.8        |
| Monte Alto     | 2.3                           | 12.9        | 79.8        | 0.0        | 5.0        |
| Oscar          | 1.5                           | 11.1        | 84.6        | 0.4        | 2.3        |
| Picoto         | 8.2                           | 4.7         | 83.1        | 0.0        | 4.1        |
| Voador         | 1.4                           | 25.2        | 69.8        | 0.0        | 3.6        |
| <b>Average</b> | <b>5.1</b>                    | <b>19.9</b> | <b>69.3</b> | <b>0.2</b> | <b>5.4</b> |

The type of substrate was visually evaluated along the full length of the underwater dives, indicating those sections where each substrate category was dominant. The selected substrate types were based on bottom composition and sediment grain size and correspond to categories recurrently used in ecological studies of the deep sea: Mud, Sand, Gravel, Cobbles and pebbles, Boulders, Flat rock, Outcropping rock, Vertical walls, Coral rubble and Coral framework. Some examples of the main categories are provided in [Figure 17](#). Considering that substrate type can vary at very small scales, and combinations of different substrate types are common in the deep sea, the annotation of the images focused only on identifying the primary substrate, understood as the category with the highest percentage cover at any given time.



**Figure 17.** Examples of different substrate types observed in the images recorded with the Azor drift-cam. (a) Sand. (b) Gravels. (c) Coral rubble. (d) Flat rock. (e) Boulders. (f). Outcropping rock.

The composition of the dominant substrate in the different areas surveyed during the MapGES 2019 cruise, as identified from the images recorded with the Azor drift-cam, is provided in [Table 5](#). ‘Outcropping rock’ was the dominant category in all areas but Monte Alto, in some cases reaching values of around 80%. This result could be expected due to the complex geomorphology of the seamounts and ridges that make up the Mid-Atlantic Ridge, and also due to the sampling strategy followed by the ADSR team during the surveys, in which areas of positive slopes are targeted in order to perform dives moving up the slope towards the summit. Interestingly, four seamounts had more than 20% of the area surveyed characterised by large deposits of coral rubble, which likely derive from the breaking up of once alive coral framework that thrived in shallower areas of the seamount. The percentage of unconsolidated sediments (sand, gravels) showed high variability across the different geomorphological feature, being more common in flat areas of the seamounts, such as the summit.

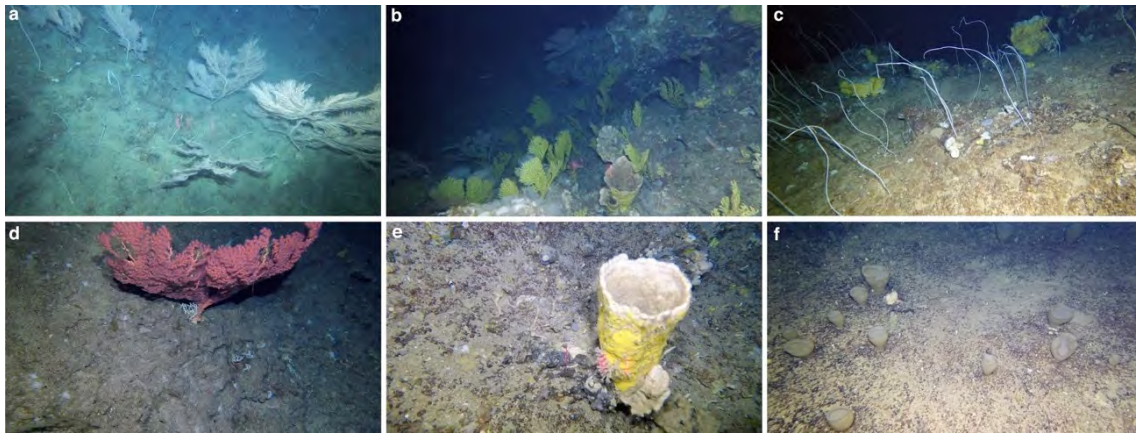
**Table 5.** Composition of the main substrate types in the 15 sampling areas explored along the Mid-Atlantic Ridge during the MapGES cruise 2019. Values are provided as percentage of time considering only the bottom time, which spans from the moment the system reaches the seabed until it starts the ascent back to the vessel.

| Seamount     | Sand       | Gravels     | Coral rubble | Pebbles & Cobbles | Boulders   | Outcropping rock | Flat rock  | Vertical rock |
|--------------|------------|-------------|--------------|-------------------|------------|------------------|------------|---------------|
| 127 Smt      | 11.6       | 33.7        | 0.0          | 0.5               | 0.0        | 54.1             | 0.0        | 0.0           |
| A3           | 8.2        | 4.7         | 15.0         | 0.2               | 0.0        | 63.6             | 0.0        | 8.4           |
| A6           | 9.3        | 13.0        | 35.3         | 0.0               | 0.0        | 42.4             | 0.0        | 0.0           |
| Alfa         | 5.5        | 10.4        | 20.8         | 0.0               | 0.0        | 63.3             | 0.0        | 0.0           |
| Beta         | 0.0        | 14.5        | 17.4         | 0.0               | 0.0        | 68.2             | 0.0        | 0.0           |
| Cavala       | 4.7        | 10.3        | 5.0          | 1.0               | 0.0        | 73.9             | 0.0        | 5.2           |
| Ferradura    | 0.9        | 23.9        | 26.1         | 0.0               | 0.0        | 49.1             | 0.0        | 0.0           |
| Gigante      | 2.7        | 25.7        | 2.7          | 2.4               | 1.8        | 64.6             | 0.0        | 0.0           |
| Gigante NE   | 0.0        | 0.5         | 18.6         | 0.0               | 0.0        | 79.4             | 0.0        | 1.6           |
| Gigante S    | 15.6       | 6.3         | 12.4         | 0.5               | 0.0        | 65.2             | 0.0        | 0.0           |
| Gigante SW   | 13.6       | 8.5         | 25.9         | 0.0               | 0.0        | 49.2             | 0.0        | 2.8           |
| Monte Alto   | 70.9       | 1.3         | 0.0          | 0.0               | 0.0        | 27.8             | 0.0        | 0.0           |
| Oscar        | 3.5        | 11.6        | 18.1         | 0.2               | 0.3        | 62.3             | 0.4        | 3.6           |
| Picoto       | 17.4       | 30.0        | 3.4          | 0.0               | 0.0        | 42.6             | 5.2        | 1.4           |
| Voador       | 4.9        | 5.7         | 6.1          | 0.0               | 0.0        | 83.1             | 0.0        | 0.1           |
| <b>Total</b> | <b>8.3</b> | <b>13.6</b> | <b>13.2</b>  | <b>0.3</b>        | <b>0.1</b> | <b>62.3</b>      | <b>0.4</b> | <b>1.9</b>    |

The assessment of the biological diversity that characterises each of the areas explored was accomplished through the annotation of the benthic megafauna observed in the images. Benthic megafauna is defined as those sessile (or low mobility) invertebrate species of a size greater than several centimetres that can be identified from underwater images. Each (morpho)species observed was classified and given an Operational Taxonomic Unit (OTU) number. Each OTU was registered for each dive, providing a rapid assessment of the observed diversity along large geographic areas. This rapid analysis generates a database of benthic megafauna occurrences at the scale of approx. 1,000 m that allows species distribution maps to be rapidly constructed, with data to be used in other large-scale studies, as for example, the development of predictive models of species distributions.

A total of 171 morphospecies were identified from the images recorded by the Azor drift-cam, most of which belonging to the phylum Cnidaria (74) and Porifera (65). The most representative group of Cnidaria was Octocorallia with 36 morphospecies, followed by Scleractinia (13) and Antipatharia (10). The remaining phyla were proportionately less represented, with only 13 morphospecies belonging to the phylum Arthropoda, and 10 to Echinodermata. Images with some of the most conspicuous species of Cnidaria and Porifera observed in the images are provided in

**Figure 18**, and include the octocorals *Callogorgia verticillata*, *Dentomuricea* aff. *meteor*, *Viminella flagellum* and *Paragorgia johnsoni*, and the sponges *Characella pachastrelloides* and *Asconema* sp. The number of morphospecies identified in each of the sampling areas surveyed along the Mid-Atlantic Ridge was highly variable (**Table 6**), ranging from 113 in Cavala seamount to just 39 in Monte Alto. Octocorals were always the most diverse group of Cnidaria, with areas hosting up to 22 different morphospecies (Beta and Cavala). The number of black corals as higher in Gigante seamount, with a total of 7 morphospecies identified. Cavala was not only the most diverse seamount; it also hosted the largest number of Scleractinia and Hexactinellida morphospecies. Although differences in species composition and diversity across seamounts can be partly attributed to a series of methodological constraints, such as sampling effort and depth range explored, part of that variability should be attributed to differences in the environmental conditions found in each seamount. A more profound analysis of these results will provide better clues to understand differences in the composition of the benthic fauna in seamounts along the Mid-Atlantic Ridge.



**Figure 18.** Examples of the different megabenthic species observed in the images recorded with the Azor drift-cam in the Mid-Atlantic Ridge. (a) Large colonies of *Callogorgia verticillata*. (b) Aggregation of the yellow sea fan *Dentomuricea aff. meteor*. (c) The whip coral *Viminella flagellum*. (d) A large colony of the red morphology of the species *Paragorgia johnsoni*. (e) The giant sponge *Characella pachastrelloides*. (f) An aggregation of the hexactinellid sponge of the genus *Asconema*.

**Table 6.** Number of morphospecies per phylum identified in each of the sampling areas surveyed along the Mid-Atlantic Ridge during the MapGES 2019 cruise. Morphospecies are classified in the most representative groups of the phyla Cnidaria and Porifera.

|                 | 127 Smt   | A3        | A6        | Alfa      | Beta      | Cavala     | Ferradura | Gigante   | Gigante NE | Gigante S | Gigante SW | Monte Alto | Oscar     | Picoto    | Voador     |
|-----------------|-----------|-----------|-----------|-----------|-----------|------------|-----------|-----------|------------|-----------|------------|------------|-----------|-----------|------------|
| <b>Cnidaria</b> |           |           |           |           |           |            |           |           |            |           |            |            |           |           |            |
| Actinaria       |           |           | 1         |           |           |            | 1         | 1         | 1          |           | 1          | 1          | 1         |           | 1          |
| Antipatharia    | 4         | 4         | 2         | 2         | 4         | 4          | 4         | 7         | 2          | 2         | 5          |            | 3         | 3         | 4          |
| Hydrozoa        | 3         |           | 2         | 1         | 1         | 2          | 1         | 1         |            |           | 2          |            | 3         | 1         | 4          |
| Octocorallia    | 14        | 10        | 19        | 11        | 22        | 22         | 19        | 18        | 9          | 7         | 14         | 6          | 18        | 14        | 20         |
| Scleractinia    | 5         | 4         | 2         | 2         | 3         | 8          | 3         | 5         | 3          | 3         | 3          | 1          | 5         | 1         | 5          |
| Stylasteridae   |           |           | 2         | 1         |           | 2          | 2         |           |            |           |            |            |           |           | 3          |
| <b>Porifera</b> |           |           |           |           |           |            |           |           |            |           |            |            |           |           |            |
| Desmospongia    | 15        | 12        | 16        | 11        | 11        | 21         | 15        | 18        | 8          | 12        | 17         | 15         | 12        | 11        | 19         |
| Hexactinellida  | 1         | 3         | 7         | 4         | 5         | 7          | 5         | 5         | 5          | 1         | 1          |            | 4         | 2         | 4          |
| Indet.          | 24        | 19        | 19        | 17        | 19        | 26         | 19        | 26        | 14         | 10        | 22         | 13         | 19        | 19        | 26         |
| Bryozoa         | 2         |           | 1         |           | 1         | 2          | 1         | 1         |            |           | 1          |            |           |           | 2          |
| Mollusca        | 1         | 2         |           | 1         | 1         | 3          | 1         | 1         |            | 1         | 3          |            | 2         | 2         | 3          |
| Arthropoda      | 6         | 6         | 8         | 5         | 6         | 10         | 8         | 6         | 4          | 3         | 6          | 1          | 7         | 8         | 6          |
| Echinodermata   | 4         | 2         | 5         | 2         | 3         | 5          | 4         | 7         | 4          | 4         | 3          | 2          | 7         | 4         | 3          |
| Brachiopoda     |           |           | 1         |           |           | 1          |           | 1         |            |           |            |            |           |           |            |
| <b>Total</b>    | <b>79</b> | <b>62</b> | <b>85</b> | <b>57</b> | <b>76</b> | <b>113</b> | <b>83</b> | <b>97</b> | <b>50</b>  | <b>43</b> | <b>78</b>  | <b>39</b>  | <b>81</b> | <b>65</b> | <b>100</b> |

### 3.7 Limitations of the Azor drift-cam at the current state of development

Although the Azor drift-cam has shown the capacity to offer a series of advantages with respect to other deep-sea imaging platforms, especially regarding its cost, simplicity and operability, there are still some drawbacks that should not be overlooked. One of the major limitations relates to the impossibility of collecting biological samples necessary for a correct taxonomic identification of the species observed, in opposition to the ability of ROVs and HUVs to collect specimens using their hydraulic arms. Furthermore, the inability of the Azor drift-cam to generate close ups of organisms of interest, or even targeted filming, can also be an important limitation when surveying areas not yet explored, for which limited information is available or where the composition of the benthic fauna has not been fully characterised. Another important limitation relates to the maximum operational depth that can be achieved by the Azor drift-cam, currently constrained to 1,000 m depth. The low-cost standard electrical cable used as umbilical to send the video feed to the surface produces a loss of video signal that increases with cable length. The use of ~1,200 m of electrical cable implies that the image received on the vessel has lost its original colour (black and white image) and some of its sharpness, although it is still sufficient for navigation purposes. Further increases in cable length would continue reducing image quality to a point where maintaining a safe navigation over the seabed would become extremely difficult, hence putting the system at risk and being prone to potential entanglements with rocks or overhangs.



In terms of operability, since the Azor drift-cam primarily follows the direction taken by the vessel based on its drift, targeting specific areas, or following a predefined path over the seabed becomes very difficult, unless weather conditions allow. Targeting areas of interest can be partially achieved by evaluating the vessel's drift and the bathymetric chart prior to each dive, and adjusting the position of the vessel once the system is deployed. Another aspect that can limit the performance of the Azor drift-cam relates to cross-side currents over the seafloor. These water flows can turn the system sideways, leaving the camera at an angle from the path being followed. This situation not only limits image quality, but it can also constrain video-based navigation and increases entanglement risk. Surveying areas with a negative slope can also be a complex task since light is absorbed through the water column, limiting video quality for fauna annotation. The static position of the camera angle implies that it has to be fixed before the dive, and no adjustments can be made during the transect over the seabed to compensate for differences on the slope of the seabed.

Finally, one of the main limitations of the Azor drift-cam in its current state of development relates to the lack of a real-time GPS positioning of the system underwater. Commercial acoustic positioning systems (such as those provided by the Ultra-Short Baseline method; USBL) are generally very expensive and would break the guiding principle on Affordability. The development of marine technology will eventually lead to affordable USBL positioning systems to be available in the market, allowing for a more accurate position of the Azor drift-cam over the seabed, which will further increase the accuracy of the data produced.

### 3.8 Technology readiness level (TRL)

Based on the information collected in the different surveys carried out at sea in the past years, the Azor drift-cam is now at Technology Readiness Level 7 (TRL 7), with the system fully verified in an operational environment.

### 3.9 Conclusions

There is a diverse range of state-of-the-art underwater technology for different deep-sea applications. However, simple yet affordable systems that take advantage of the power of small commercial action cameras remain unavailable. The Azor drift-cam, built with off-the-shelf components, has proven cost-effective, easy to operate and reliable to explore deep-sea habitats to 1,000 m depth. This system aims to provide a simple yet versatile tool to facilitate the access to deep-sea exploration, not intending to become a substitute for other, more sophisticated, video and photography platforms (ROVs, AUVs or manned submersibles). In fact, both systems could complement each other, with the Azor drift-cam providing rapid assessments of the seabed over large areas, allowing relevant locations to be identified and subsequently targeted for a more detailed examination. This could further optimise and reduce the costs of ROV operations. Hence, the Azor drift-cam could play an important role in the Deep-Ocean Observing Strategy<sup>17</sup> and the measurement of Essential Ocean Variables for deep-sea monitoring and conservation strategies<sup>18</sup>. This technological development can contribute to open the doors to many researchers/countries currently excluded from this research field due to monetary and technological restrictions.

### Acknowledgements

The authors would like to especially acknowledge all those involved during the development of the Azor drift-cam, both in terms of technical assistance to prepare the different prototypes and their participation during the first trials at sea and the subsequent deep-sea surveys (MapGES 2018, 2019, 2020, 2021), namely Sérgio Gomes, Gerald H. Taranto, Victor Rosa, Gonçalo Graça, Marco Dutra, Manuela Ramos, Guilherme Gonçalves, Inês Carneiro, Laurence Fauconnet, Joana Brito, Jordi Blasco-Ferre, Cristina Gutiérrez-Zárate, Luís Rodrigues and Marina Carreiro-Siva. We are also grateful for all the help provided by the crews of the research vessel 'N/I *Arquipélago*' during all those years and by the fishing vessels 'Tatiana', 'Garantia', 'Galinha' and 'Gotimar'. Finally, we would like to thank Group B Distribution Inc. for their technical assistance whenever needed.

## 4. Underwater hyperspectral imaging

Contributions from: Touria Baijouk, Maxime Ferrera, Tristan Petit & Aurélien Arnaubec

### 4.1 Introduction

While human eyes and conventional cameras sense colours through three broad red, green and blue (RGB) spectral bands in the visible range of light, hyperspectral imagers separate light into numerous narrow bands in the visible, near-infrared (NIR) or even mid-infrared (MIR) spectral regions<sup>19</sup>. This allows additional details and features to be captured, thereby potentially improving scientific characterisation of objects.

Based on passive sensors installed either on aerial platforms or satellites, hyperspectral imaging has been successfully used to detect, characterise, and quantify a wide range of objects around the world. These sensors use the sun as the light source which limits their use to above-water applications with limited feasibility for shallow water use cases. In deeper waters, where sunlight is negligible or absent, specific underwater hyperspectral imaging (UHI) equipment is needed where the underwater scene is illuminated by artificial light.

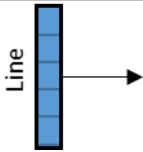
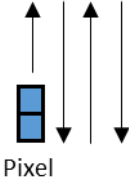
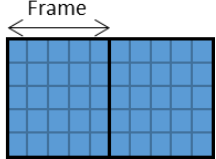
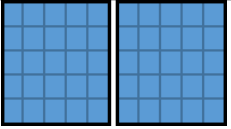
As part of the Life MarHa project and the imaging technologies showcases being developed in the *iAtlantic* project, UHI equipment was tested for different scenarios. The main objective was to evaluate the potential of this technology for deep-sea mapping and monitoring with a focus on assessing the maturity and reliability of the system for benthic habitat identification and to assess the ecological status of habitats.

### 4.2 Overview of underwater hyperspectral imaging technology

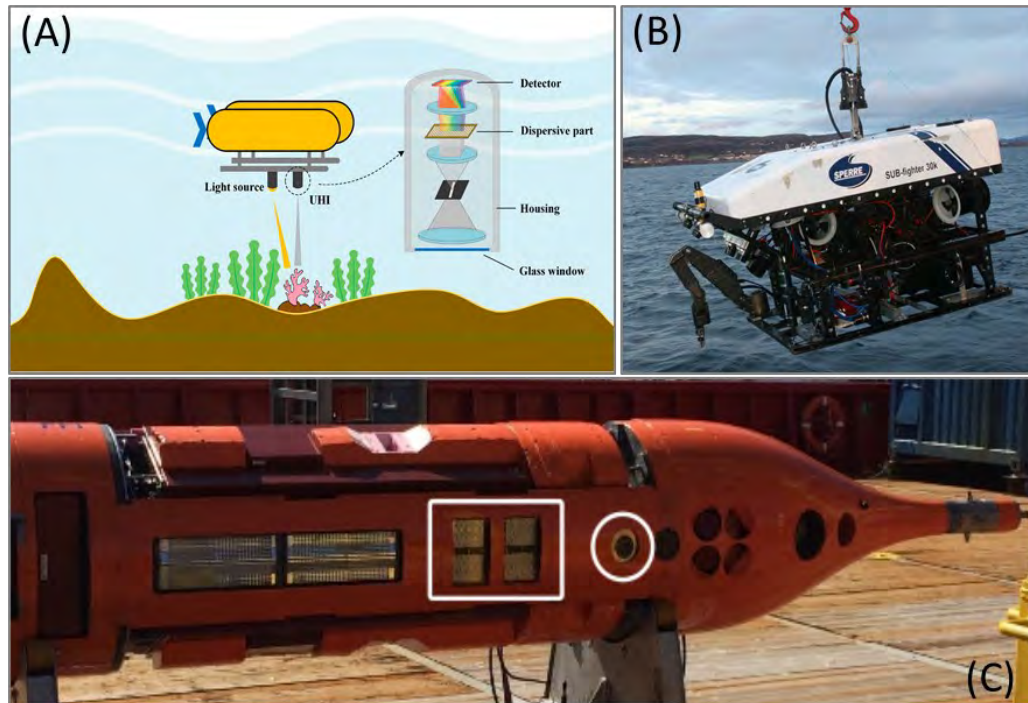
#### Sensors

Represented as a 3-Dimensional data cube, hyperspectral images are characterised by both their two spatial dimensions and one spectral dimension. Sensors are thus characterised by the arrangement and/or the number of spectral bands involved in the instrumental architecture, as well as the employed image capturing method. Stuart et al. provides a review of current hyperspectral technologies and their integration into the environmental monitoring field<sup>19</sup>. The concept and main characteristics are summarised in [Table 7](#).

**Table 7.** Current hyperspectral technologies concepts, summarised from Stuart et al.<sup>19</sup>. The main conclusion is that push broom spectral imaging is commonly used, thanks to its suitability for light-weight unmanned aerial vehicle (UAV) image acquisition<sup>20</sup>. Snapshot systems can record the spectral image data cube from a unique shot, which is suitable for moving targets, however, these devices are currently limited due to their larger size<sup>19</sup>.

|               | Push broom   | Whisk broom  | Snapshot   |   |
|---------------|--|--|--|---|
|               |  |  | Framing  | Windowing   |
| Approach      |   |   |    |    |
|               | line-by-line image acquisition process of spectral information per exposure is recorded  | image a single pixel at a time, using a rotating mirror to sweep a scan line perpendicular to the direction of the sensor platform's movement  | Capture scenes through 2-Dimensional images with additional optics that focus on either an individual wavelength or wavelength bands   | Employ a 2-Dimensional Field of View (FOV) that moves across a scene in a continuous fashion  |
| Advantages    | Significantly greater spatial and spectral resolution  | Demonstrated the pioneering of low-cost whiskbroom image collection suitable for UAV deployment.   | The sensor design significantly simpler than those of push broom and whiskbroom sensors  |   |
| Disadvantages | <ul style="list-style-type: none"> <li>- difficulties in post-processing</li> <li>- confining potential data misalignments to between lines</li> <li>- needs the support of inertial navigation and other auxiliary systems</li> </ul> | <ul style="list-style-type: none"> <li>- difficulties in post-processing</li> <li>- provide inherently slower frame rates than Push Broom units, resulting in lengthier data acquisition periods where all other things are equal.</li> <li>- the rotation of the optics can result in spatial distortions in the image outputs</li> </ul> | <ul style="list-style-type: none"> <li>- the use of spectral filtering substantially reduces the intensity of light captured at the sensor, limiting signal to noise.</li> <li>- Limited spectral resolution due to the filtering of spectral bands</li> </ul> | <ul style="list-style-type: none"> <li>- acquire a distinct exposure each time the FOV moves forward, with no integration between exposures.</li> <li>- Limited spectral resolution due to the filtering of spectral bands</li> </ul> |

Three main components are necessary for underwater hyperspectral imaging (UHI) deployment: a light source, the hyperspectral imager sensor and a platform for its deployment<sup>21</sup>. Thanks to mobile underwater robots, not only can snapshots of stationary objects be taken but there is a potential to map large areas of interest with hyperspectral sensors (Figure 19).



**Figure 19.** Schematic diagram of a UHI system<sup>22</sup> (A), an example of ROV used for the UHI-based survey in Trondheimsfjord<sup>23</sup> (B) and the Ecotone UHI (first circular opening indicated by the white circle) mounted on AUV at the right of the lamps (four smaller squares inside the white rectangle)<sup>24</sup> (C).

As described by Johnsen et al., two types of marine robotic vehicles are mainly used to deploy UHI sensors<sup>25</sup>:

- Remotely operated vehicles (ROVs) are tethered from a ship during the entire deployment to provide both power and communication to the vehicle. The main advantage of using a ROV as UHI platform is the on-line control of the instrument and collection of the data stream. Thus, dysfunctions are identified during data acquisition and controls can be sent to the platform to adjust the settings of the operated instrument. The use of ROVs has, however, some limitations. The spatial coverage is limited by the tether length (generally 500–1,000 m long). On top of that, their external thruster motors may resuspend sediments and reduce image quality if the UHI is operated close to soft sediments.
- Autonomous underwater vehicles (AUV) can either be a glider (not ideal platforms for UHI as they have very low power capacity and limited sampling capabilities near the seafloor) or, in contrast, propeller-driven systems that offer higher payloads and more sampling capabilities. A large AUV may be the best platform for UHI thanks to their high degree of speed/direction, altitude and pitch, roll, and yaw control. In addition, the ability of hover AUVs to map vertical walls and rough bottoms with a complex structure, can make seafloor mapping quite efficient.

### Navigation requirements

As pointed out by Johnsen et al.<sup>25</sup>, the targeted accuracy for the acquired UHI images determines the navigation equipment and survey procedures. Navigation data for the underwater vehicle on which the UHI is mounted are thus required for georeferencing the image cube. There are two main tools to perform underwater navigation for UHI:

- Acoustic underwater navigation: a concept that includes Ultra-Short Base Line (USBL) and Long Base Line (LBL) systems, the most common methods used for diving robots to determine a geographical position. For highly accurate USBL navigation, a ship, a transceiver, and a transponder mounted on the robot are required, LBL positioning system relies on an array of transponders placed on the seabed and a transceiver mounted on the moving platform.
- Dead reckoning: mainly used for AUVs, this method estimates locations from the initial position, heading, time, and velocity. Mathematically, speed and acceleration are integrated over time to establish position estimates. Common sensors used for dead reckoning are gyro, Doppler Velocity Log (DVL), Inertial Measurement Unit (IMU), and Inertial Navigation System (INS).

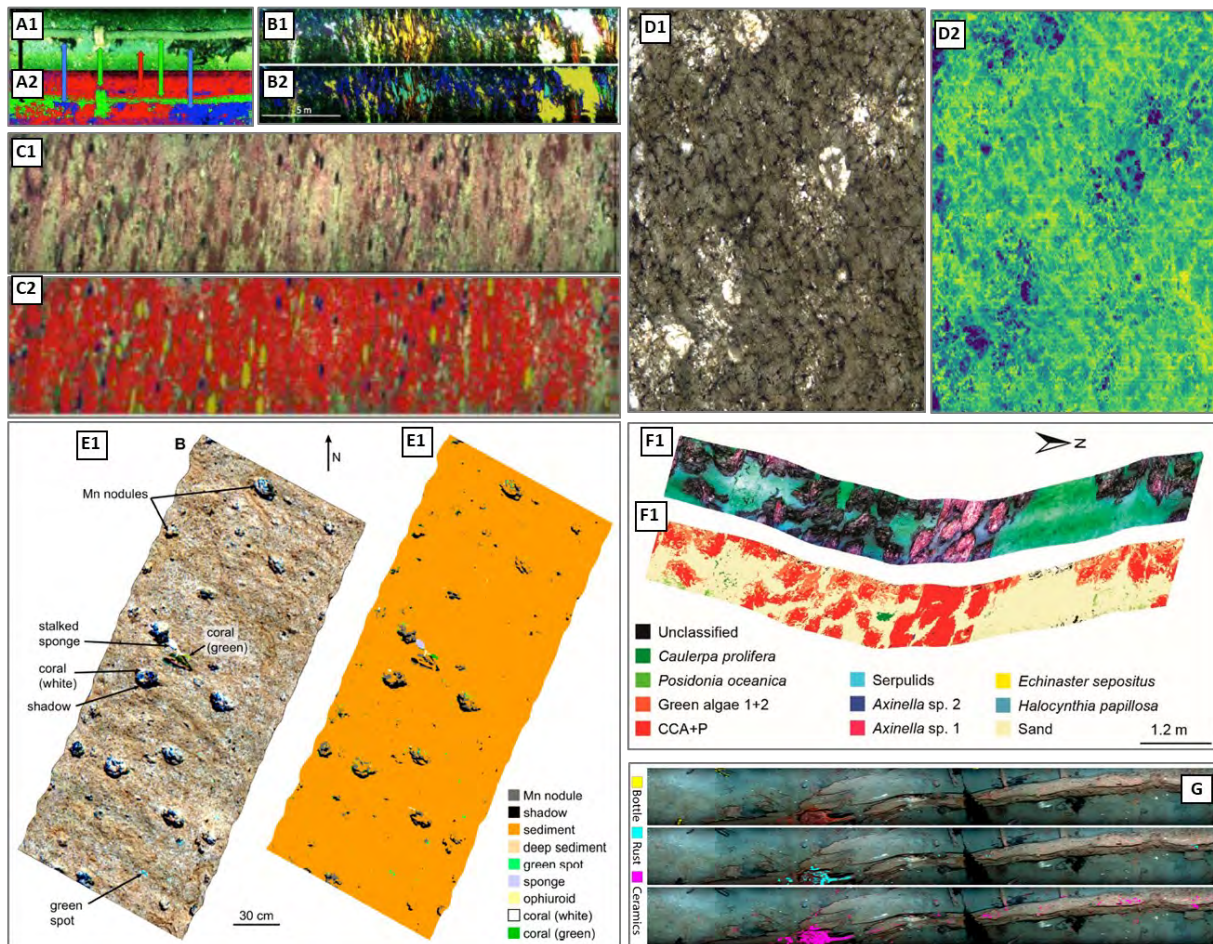
### Applications

As highlighted in a recent review<sup>22</sup>, the number of articles in the field of UHI technology has increased in recent years (**Figure 20**). UHI seems to offer a new perspective for a variety of environmental monitoring researches and operations, with its focus on non-destructive data acquisition, often a key requirement for data acquisition in protected areas<sup>19</sup>.



**Figure 20.** Annual number of articles published in the field of underwater spectral imaging<sup>22</sup>.

The application of HS imagery, from which physical and ecological characteristics can be extracted, cover several fields (**Figure 21**). Johnsen et al. have applied supervised classifications to UHI data in the visible range with up to millimetre spatial resolution and 1 nm spectral resolution<sup>26</sup>. They were able to automatically identify and map several sea bottom types, including kelp forest, sponges, and deep-water coral reefs as well as man-made structures such as pipelines. More recently, vertical and horizontal distribution of sea-ice algae were estimated at fine scale using an *in situ* under-ice HS technology thanks to correlation of spectral indices with fluorometrically derived Chl a values<sup>27</sup>. Foglini et al. have experimented for the first time with a UHI systems to characterise cold-water coral habitat and coralligenous habitat in the Southern Adriatic Sea (Mediterranean Sea)<sup>28</sup>. Even if this method provided inaccurate discrimination of some seabed types (probably due to spectral similarities or artifacts such as shadows and distortions), SAM (Spectral Angle Mapper) supervised classification was efficient for recognising colonial cnidarians and sponges. Coraline crust algae and associated organisms were also correctly classified despite habitat heterogeneity.



**Figure 21.** A1 and A2 respectively show a UHI RGB false-colour image and its corresponding classification of raw data (digital counts) denoting *Hippasteria phrygiana* seastar (Black), brown algae dominated by *Laminaria digitata* kelp (blue), denotes sand/soft bottom (red) and concrete pipeline and holdfast (green), B1 and B2 respectively show a UHI false-colour image and its supervised classification into sand (yellow), leafy red algae (green), red calcareous algae (red), old and dark brown tissue of kelp *L. digitata* (blue) and corresponding new tissue of *L. digitata* (winter growth from meristem). C1 and C2 respectively show a UHI photomosaic in false-colours (RGB) and its corresponding classification based of seafloor habitat dominated by red calcareous algae (red area), anemones dominated by *Urticina* (yellow) and sea urchins *Strongylocentrotus* spp. (blue) during a Polar night campaign at 15 m depth (from Johnsen et al., 2016). D1 and D2 respectively show a high-resolution UHI data item and its corresponding quantitative mapping of Chl a by applying the LAUC index regression model on a per-pixel basis to the pseudo-transmittance image, E1 and E2 respectively show UHI geo-corrected, pseudo-reflectance data in pseudo-RGB and its SVM classification image, F1 and F2 respectively show an RGB UHI image of a coralligenous site and the corresponding SAM classification (from Foglini et al. 2019) and (G) Supervised SAM classification images of a wreck transect. From top to bottom: glass bottle, rust, and ceramics (from Ødegård et al., 2018).

Based only on reflectance spectra, a machine learning models have been used to detect changes in the health status of both orange and white morphotypes of the cold-water coral *Lophelia*, achieved by exposing 60 coral samples to the toxic compound 2-methylnaphthalene (Letnes et al., 2019). Corals were successfully classified according to lethal concentration levels of 5% and 25% mortality, thus opening prospects for environmental impact assessment over larger areas. Likewise, as mentioned by (Stuart et al., 2019), underwater hyperspectral technology may also be suitable for extreme environment monitoring by enabling the acquisition of high spatial resolution where ground-based field surveys would be difficult to operate. The addition of field hyperspectral sensing to glacial settings can provide a significant improvement of knowledge related to identification of supraglacial debris composition (Di Mauro et al., 2017).

## Main challenges

Currently, push broom and whisk broom cameras need to be deployed from a mobile platform to complete the image stitching required for contiguous mapping. Hence, sensor stability is low and subject to limitations of georectification<sup>19</sup>. As highlighted in Ferrera et al., most studies using UHI cameras have relied on mosaicking techniques to exploit and analyse the gathered hyperspectral data<sup>29</sup>. Yet, the creation of such mosaics is performed by interpolating the UHI camera trajectory using drifting dead-reckoning navigation systems, leading to approximate trajectory estimation. Furthermore, mosaicking techniques also use a flat surface assumption on the imaged scene, leading to distorted results when this assumption is no longer correct. With higher control of speed/direction, altitude and pitch, roll, and yaw (Figure 22), the UHI images must be previously geo-processed to provide accurate ortho-images and to recover an exploitable material for scientific analysis, more specifically for non-flat areas.

Alongside the improvement of geometric processing, radiometric aspects are also challenging due to scene dependence of most of the underwater imaging techniques. Due to the attenuation effect caused by water, the spectral intensity of the underwater target is indeed different at different distances<sup>21</sup>. In addition, the scattering phenomenon caused by particles suspended in highly turbid water strongly affects the recorded hyperspectral signal. The recovery of reflectance data through careful measurement, calibration, and mitigation of light in the underwater environment, as indicated by Bongiorno et al., are required to ensure accurate information extraction, especially for spatio-temporal studies of process dynamics<sup>30</sup>.

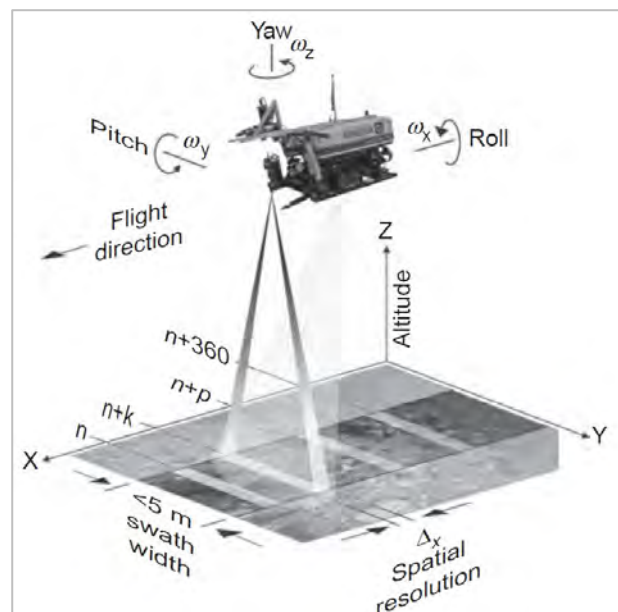


Figure 22. A schematic diagram showing the geometry of UHI acquisition deployed on ROV24.

## 4.3 iAtlantic Experimentation

### Hyperspectral camera

Johnsen, from the Norwegian University of Science and Technology (NTNU), invented the first underwater hyperspectral imagers before developing a push-broom system based on grating diffraction that was the only commercial one available at the time<sup>31</sup>. Later, NTNU researchers together with the Ecotone company developed several push-broom UHI systems with large field of view ( $\sim 60^\circ$ ) and diving capacity down to 6,000 m.

As part of the European MarHa

project, coordinated by the French Agency for Biodiversity, IFREMER was able to purchase a hyperspectral underwater camera (UHI) to extend the use of this technology to shallow and deep marine

habitats. The UHI push-broom sensor was chosen, as, unlike snapshot hyperspectral cameras which are quite limited in terms of spatial and spectral resolutions (typically  $\leq 500 \times 500$  px and  $\leq 25$  spectral bands), push-broom cameras provide wide spatial and spectral resolutions (typically  $\geq 960$  px and  $> 100$  spectral bands)<sup>22</sup>.

The UHI is a scanner that continuously records intensities of reflected light in the 378–800 nm spectral range (Figure 23). The housing is made of titanium alloy. This push-broom scanner with beam widths of 60° (transverse) and 0.4° (longitudinal) can be mounted looking vertically downwards to record lines of 1,900 pixels perpendicular to the track direction. Intensities of reflected light are measured for spectral bands between 378 and 800 nm with a spectral resolution of 4 nm.

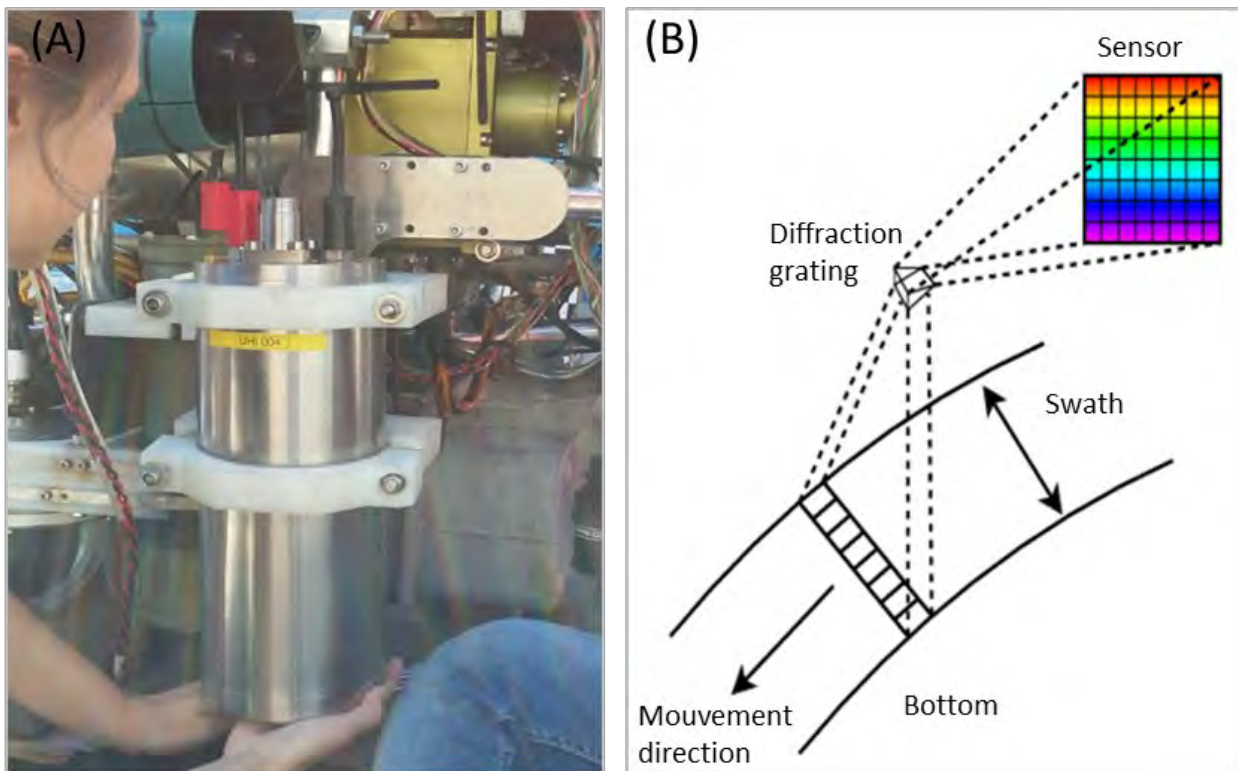


Figure 23. Underwater Hyperspectral imaging acquired system (A) and a schematic diagram showing the principle of data acquisition (B).

### UHI integration

Integration of the UHI into IFREMER’s ROVs was tested during two surveys. The first one was performed in the framework of the MarHa project in the Mer d’Iroise coastal area (Brest, France) by IFREMER and Ecotone. The integration of this camera was also performed on the HROV *Ariane*. This latest test was carried out as part of the European EU MarineRobots project in Mediterranean (Table 8).

Table 8. Integration test survey summary

|             | Brest (Atlantic)   | Cassidaigne (Mediterranean)      |
|-------------|--------------------|----------------------------------|
| Framework   | Life MarHa project | European EU MarineRobots project |
| Depth       | 10 to 30 m         | 200 to 500 m                     |
| Survey date | 2018-09-17-21      | 01/08/2019 to 08                 |
| Vessel      | TSM <i>Penzer</i>  | <i>Antea</i>                     |
| ROV         | HROV <i>Vortex</i> | <i>Ariane</i>                    |

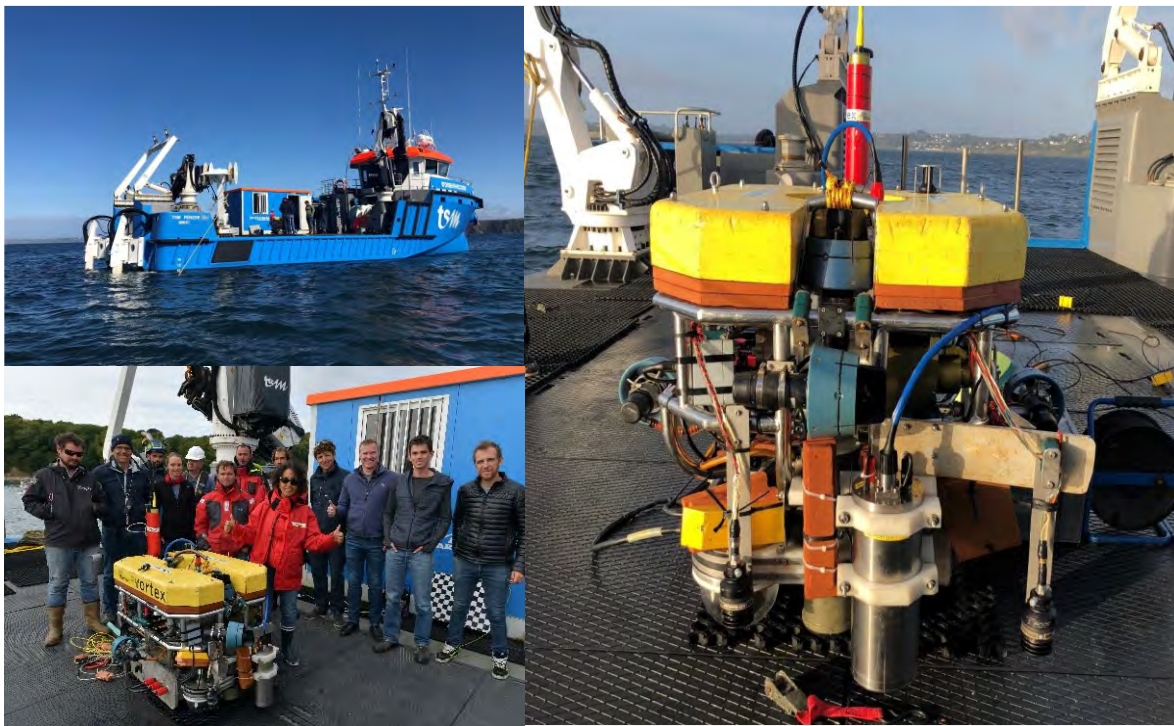


|              | Brest (Atlantic)   | Cassidaigne (Mediterranean)  |
|--------------|--|--|
| UHI          | 4-3  | 4-3  |
| Illumination | 2xLED<br>DEEPSEA Sealite LSL-2000, 8,000 lumen (5,000 K ~ 6,500 K) | 2xLED<br>DEEPSEA Sealite LSL-2000, 8,000 lumen (5,000 K ~ 6,500 K) |
| Challenges   | First integration of the UHI to IFREMER HROV                       | Improving UHI navigation estimate and data quality                 |

### Survey 1: Brest

The main objective was the integration of the Underwater hyperspectral imaging (UHI) to one of IFREMER's ROVs, the *Vortex Hybrid* ROV. For this first test, the demonstration area was related to selected coastal seabed habitat areas in the Iroise Natural Marine Park (Figure 25).

The survey was carried out by IFREMER and Ecotone, onboard the vessel TSM *Penzer*, using UHI 4-3 mounted on HROV *Vortex*, on September 17<sup>th</sup>–21<sup>st</sup>2018. The data collection was carried out from the vessel TSM *Penzer*. The HROV *Vortex* operated by IFREMER was used as platform carrier (Figure 24).



**Figure 24.** The *Vortex* IFREMER HROV with Underwater Hyperspectral imager (UHI) mounted in the rear end (Wright) carried out from the TSM *Penzer* vessel (Top left) by IFREMER and Ecotone team with the help of the crew and the PNMI divers (bottom left).

The HROV had an umbilical to a surface buoy, and wi-fi signal from the buoy to control room onboard vessel. The UHI computer was on ROV side, and the UHI was remotely accessed from control room. Navigation data were recorded. Flight altitude was mainly 1 m throughout the survey.

The integration was done before the cruise, while the vessel was docked in the harbour of Brest. For seafloor illumination, two LEDs were used (DEEPSEA Sealite LSL-2000, 8,000 lumen, 5,000-6,500 K). Calibration tests were carried out at several locations, either on deck or under water, aided by divers. Sites with different habitats (maerl, dead and live *Crepidula* gastropod mollusc bed, red algae, and sea grass) in the proximity of Brest were visited. At each site, a set of transect lines were carried out to determine adequate parametrisation. In addition, at the last part of the survey, the UHI was run in parallel

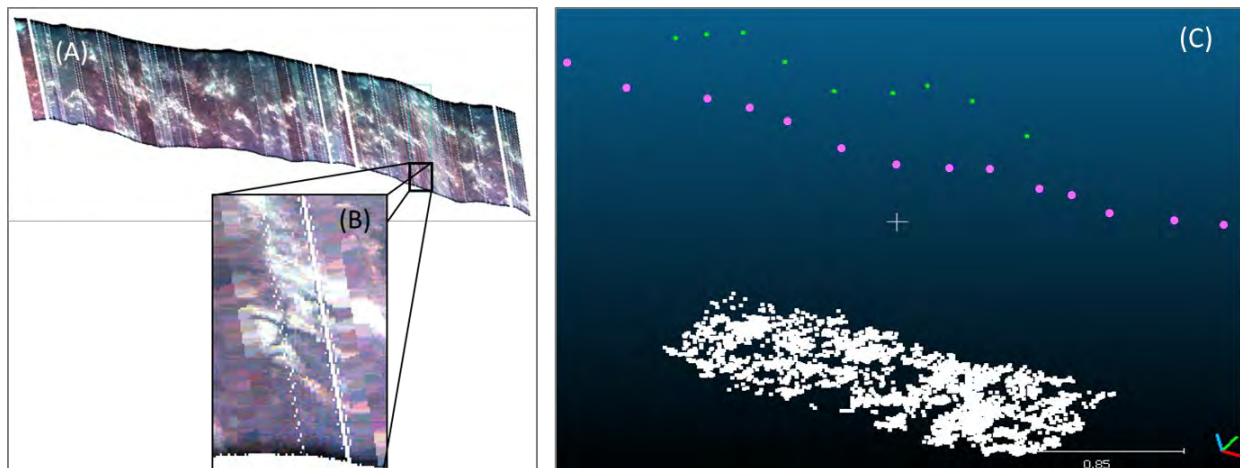
with one of IFREMER’s high resolution underwater cameras. In this case, only 1 LED lamp was allocated to the UHI.



**Figure 25.** Left: study areas in the Brest harbour, right: hyperspectral images turned into RGB false-colour images of starfish (top) and brittle stars (bottom) at red algae covered seafloor.

At the end of the survey, the integration of UHI into the HROV worked well, enabling acquisition of hyperspectral images. However, some issues arose when georeferencing the images. As the UHI sensor is a line sensor and the image data are created by moving the platform, imaging accuracy depends on the accuracy of the georeferencing data.

By using the proprietary Ecotone software, a data processed pipeline assuming data collected by flying drones or planes were employed without including relative motion or attitude information. While the UHI data are recorded at a spatial resolution of millimetres per pixel, navigation data were only available in the order of tens of centimetres. It could be seen that this navigation accuracy is not sufficient to properly align UHI data for robust interpretation (**Figure 26**).



**Figure 26.** Stitched and georeferenced UHI image transect (A), example of georeferencing issue (B) with UHI navigation in green, Corrected UHI navigation in pink and 3D point cloud from 3D reconstruction in white (C).

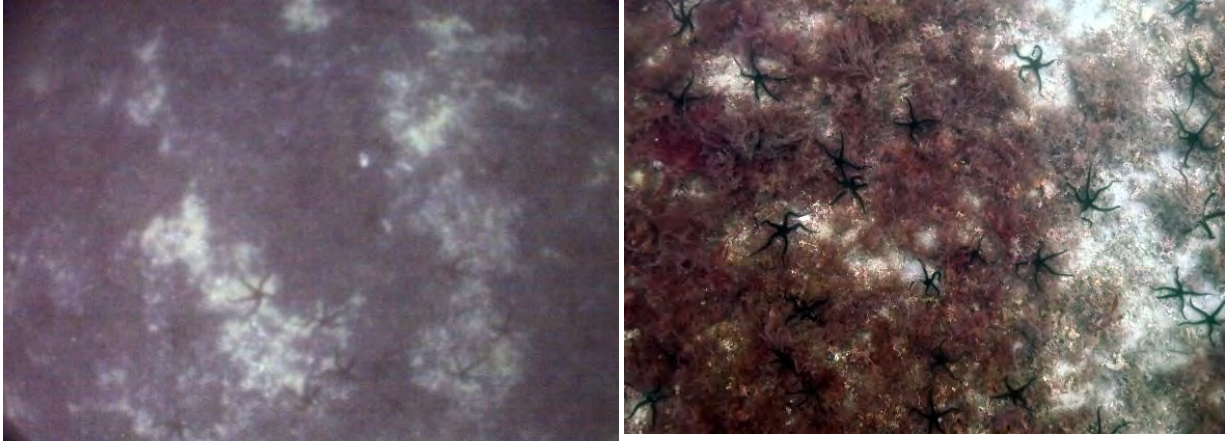
To overcome the identified georeferencing issue, another optical sensor may be used to improve navigation accuracy. The idea is to use a traditional high-resolution RGB camera to create a 3D model of the illuminated scene that is seen by both the HD camera and the UHI sensor.

By deploying a video camera, a sufficient image overlapping and corresponding high georeferencing accuracy can be created by using e.g., structure-from-motion techniques. From the reconstructed 3D scene, a very accurate navigation at millimetre resolution can be back calculated to then project

the hyperspectral transect data as a drape onto the 3D scene with comparable navigation and resolution orders of magnitude.

### Survey 2: Cassidaigne

This survey was an opportunity to test the Hyperspectral camera during deployments of the *Ariane* HROV. From the lessons learnt during the previous test survey, it was decided to couple the hyperspectral camera with an HD video camera looking at the exact same scene to enable optical navigation and 3D reconstruction. Hence, hyperspectral data can be projected on the 3D model to obtain 3D hyperspectral imaging data. To generate optical navigation data the *Vortex* camera of the *Ariane* HROV was used.

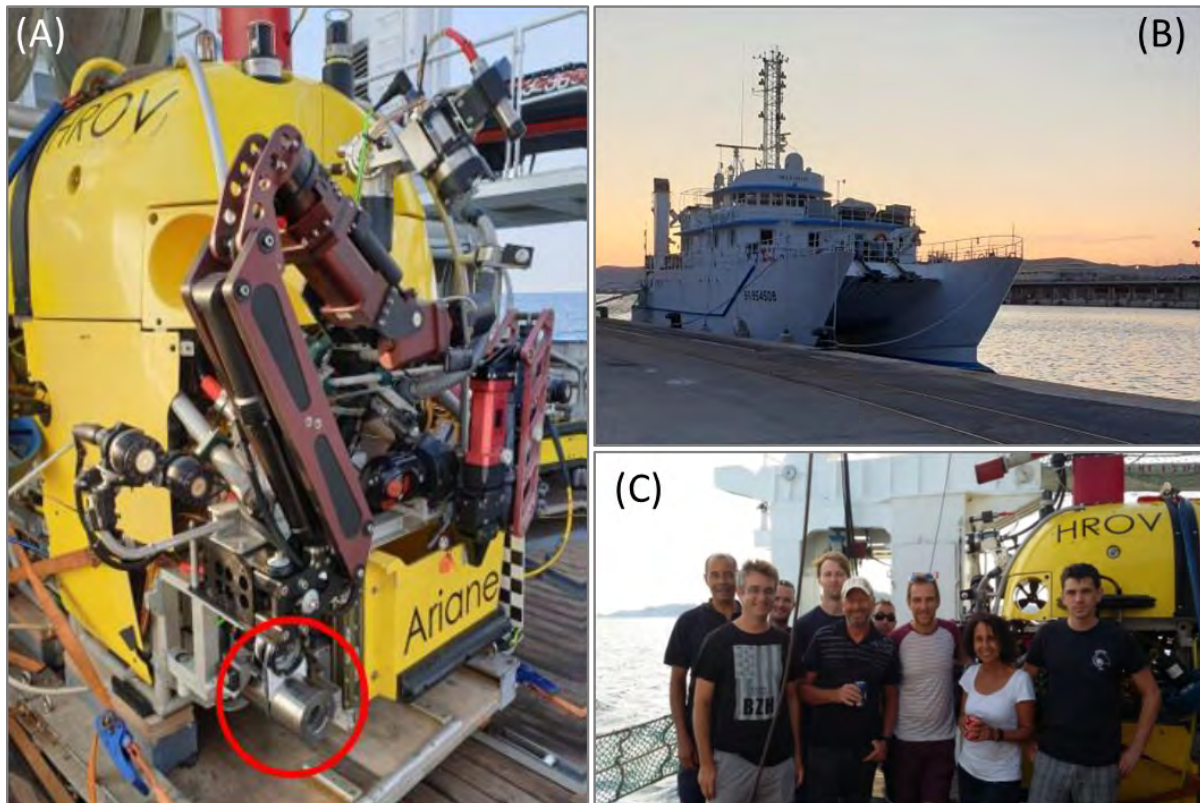


**Figure 27.** Snapshot of the Ecotone RGB camera (left) and the high-resolution RGB imagery acquired by the *Vortex* HROV still camera (right) for 3D reconstruction for optical navigation.

Ecotone participated in the survey organised by an IFREMER team from Toulon, France (August 3<sup>rd</sup> – 6<sup>th</sup> 2019). The survey was part of the project ‘Hyperspectral Imaging – Scientific payload integration and evaluation on hybrid ROV’ within the program EU Marine Robots (EUMIR) program. The first part of the survey was used for technical integration of the hyperspectral imager into the ROV system (**Figure 28**). Data collection was carried out from the vessel *Antea* with scientists from IFREMER and one operator from Ecotone on board.

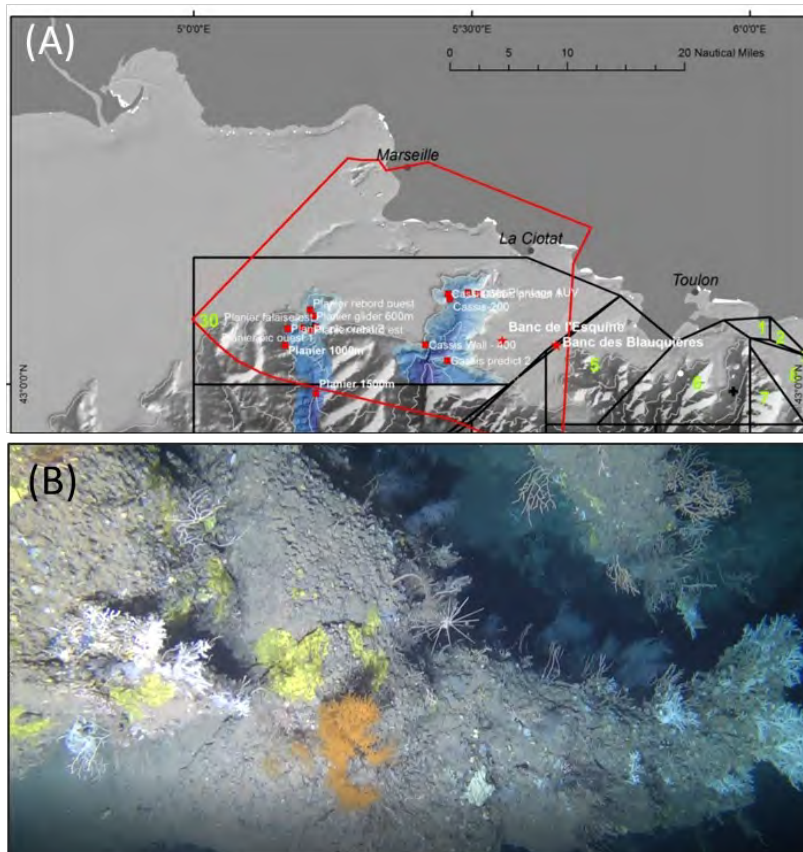
The test survey was carried out near Toulon between Marseille and Saint Tropez within the Zonex 30 area (**Figure 29**). This area has been chosen for these tests because it is large enough to allow the planned dives to be adapted according to the authorisations, the weather and technical goals.

The first part of the survey was dedicated to technological tests of the UHI camera prototype provided by Ecotone. After two days of tuning, it was integrated to be operated in parallel to *Ariane's* *Vortex* science camera, thus allowing for 3D reconstruction. Two full dives were made with this configuration on the head of the Cassidaigne canyon between 200 to 500 m in cold water coral habitats (**Figure 29**).



**Figure 28.** The Ariane HROV with Underwater Hyperspectral imager (UHI) mounted in the front (B) carried out from the Antea vessel (B) by IFREMER, Ecotone and Genavir team (C).

These tests made it possible to produce more data sets with different camera settings to evaluate the impact on the quality of the data. Data were acquired at altitude ranges from 1 to 4 m, exposure times of 10 to 40 milliseconds and at varying survey speed, to define the protocols for future dives. The tests were carried out both in vertical and horizontal orientations in relatively flat areas and on 3D structures.



**Figure 29.** Overview map of the deployment sites during the survey campaign and one example high-resolution image depicting the kind of habitat that was imaged both by the RGB camera and the UHI sensor.

### 3030314.4 Data processing

#### Radiometric calibration

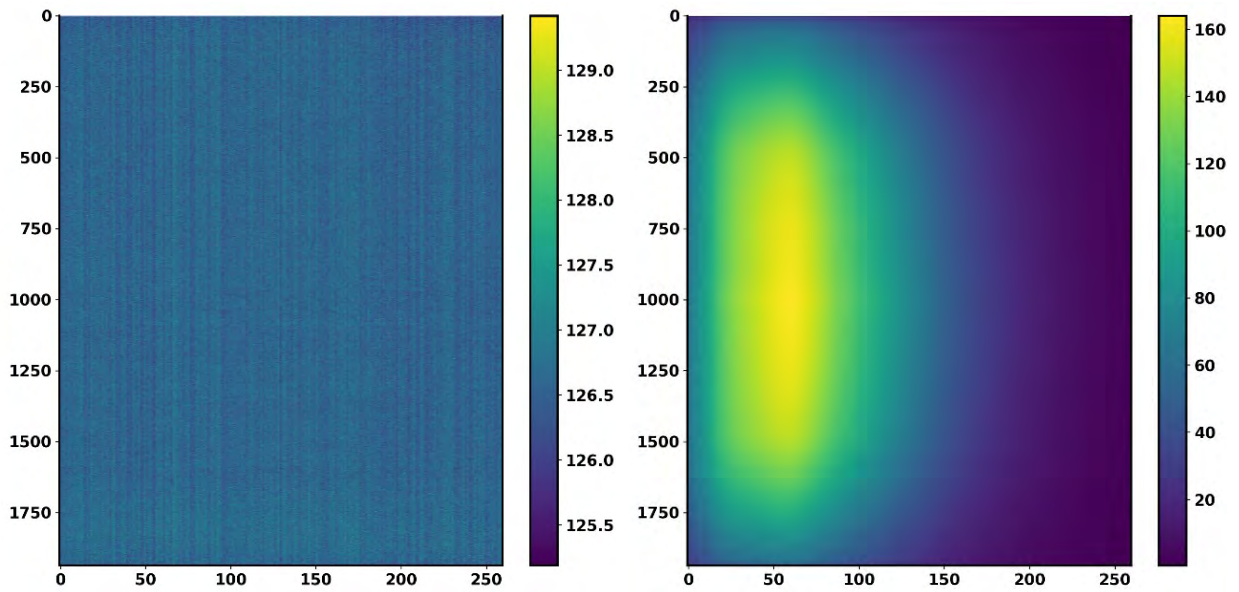
UHI signatures are distorted due to absorption and scattering of light by the water present between the camera and the sea floor. The thickness of this water body, i.e., the bottom depth relative to the camera for each pixel, can be estimated from the radiative transfer-based method of Lee et al., developed for airborne HI, which implies prior calibration and conversion into reflectance of the UHI data<sup>32</sup>.

The raw data accessible from the Ecotone UHI push broom camera are given in Digital Numbers (DN) recorded for each cell of the 2D complementary metal oxide semiconductor (CMOS) sensor as a function of time, leading to the so-called 3D hypercube. However, raw data recorded in units of DN cannot be used by any subsequent physics-based algorithm and first needs to be converted into a physical quantity: the spectral radiance. Omitting any artefacts linked to sensor imperfect design (e.g., vignetting, smile and keystone effects), derives the in-air spectral radiance  $L_{air}$  (in  $Wm^{-2}sr^{-1}nm^{-1}$ ) from a simple empirical relationship. It can be computed for each spectral band  $\lambda_i$  and spatial position  $(x_j, y_k)$  using the following equation:

$$L_{air}(\lambda_i, x_j, y_k) = \frac{(DN(i, j, k) - dark(i, j))}{t_{int} \cdot gain(i, j)} \quad (1)$$

with  $(i, j)$  the row and column index of the CMOS matrix,  $k$  the temporal index of the current frame,  $t_{int}$  the integration time (in seconds),  $dark$  a matrix provided by the manufacturer representing the average of the dark current noise, and  $gain$  a matrix provided by the manufacturer taking into account the CMOS quantum efficiency, sensor electronic characteristics as well as the loss of photons within the camera. The specific calibration matrices of the Ecotone UHI are given in Figure 32, with rows corresponding to the spatial dimension and columns corresponding to the spectral dimension. The dark noise matrix has

very low variability (125–129 DN), while the gain matrix has values ranging between 20–160  $\mu\text{Wcm}^{-2}\text{nm}^{-1}\text{sr}^{-1}$ . The spectral variability of this matrix is mainly due to variations in the quantum efficiency of the CMOS sensor cells.

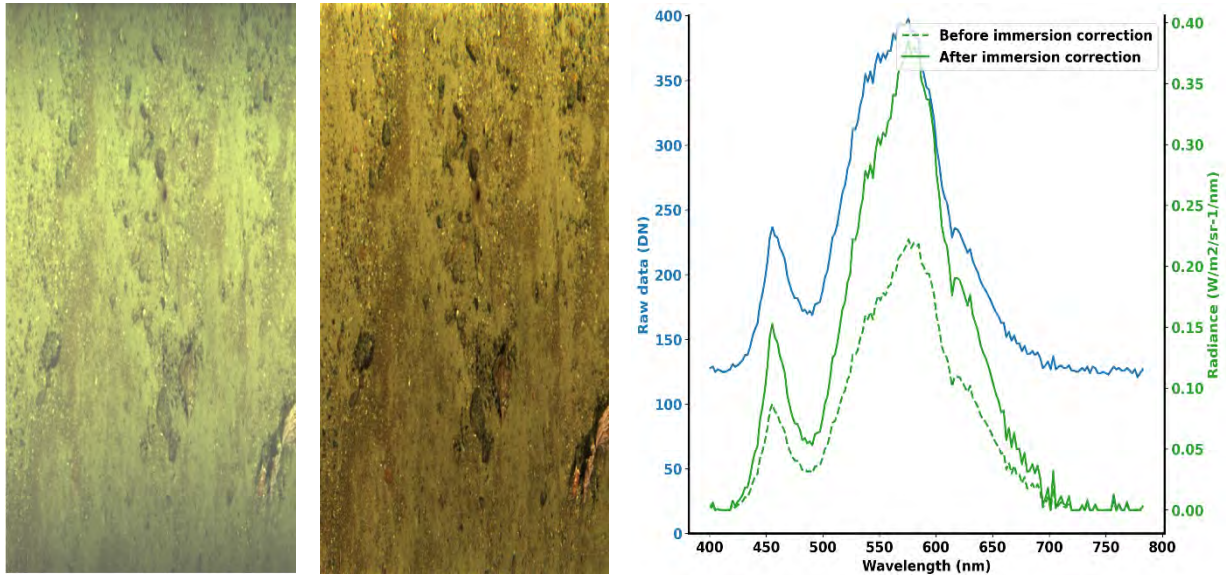


**Figure 32.** Dark (left) and gain (right) calibration matrices of the Ecotone UHI Camera operated at IFREMER. The dark matrix is expressed in unit of DN while the gain matrix is given in  $\mu\text{Wcm}^{-2}\text{nm}^{-1}\text{sr}^{-1}$ .

When the camera is immersed into a medium with refractive index different from the one of air, a second step must be performed, as the in-air calibration given in Equation 1 is no longer completely valid due to changes in the field-of-view, transmittance, and reflection at the medium-glass interface. For obtaining a valid in-water radiance  $L_{\text{water}}$ , it is thus necessary to account for these phenomena through a multiplicative factor  $I_f$  called the immersion factor, to derive  $L_{\text{water}}$  as follows:

$$L_{\text{water}}(\lambda_i, x_j, y_k) = I_f(\lambda_i) \cdot L_{\text{air}}(\lambda_i, x_j, y_k) \quad (2)$$

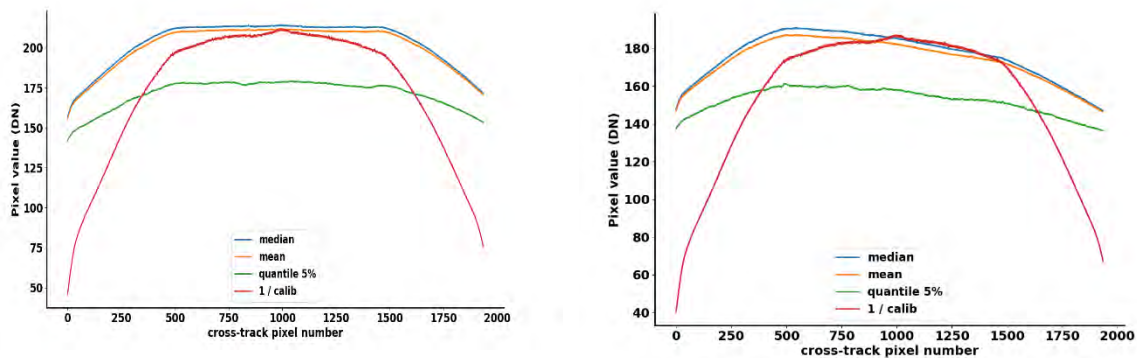
The state-of-the-art theoretical method given in Equation 9 of (Zibordi, 2006) can be used for estimating  $I_f$ . A visual comparison of a raw image in units of DN and a calibrated image in  $\mu\text{Wcm}^{-2}\text{nm}^{-1}\text{sr}^{-1}$  is presented on the left and middle panels of **Figure 33**. The crucial role of the immersion factor correction can be seen from the right panel of **Figure 33**, where the corrected spectrum attains values about 1.8 times higher than the uncorrected one.

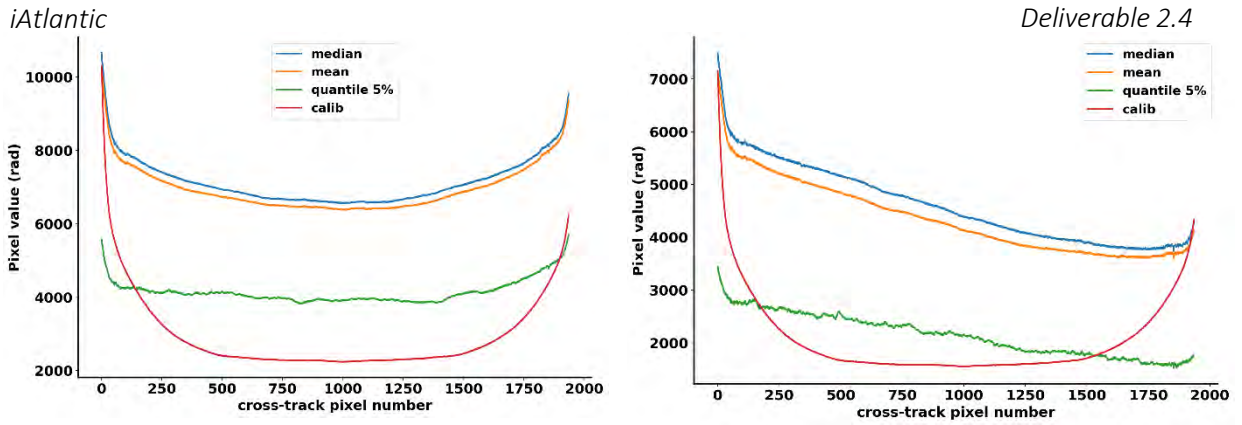


**Figure 33.** Comparison between raw and calibrated UHI data acquired during the Chereef 2022 cruise (2022-08-06 at 20:51) with, from left to right: RGB composition of the Raw image in DN, RGB composition of the calibrated image in radiance after immersion correction, spectral comparison of one hyperspectral pixel in DN and the radiance curves before (dashed) and after (solid) immersion correction.

**Vignetting correction assessment**

**Figure 34** shows temporal statistics derived for each spatial pixel (i.e., the across-track dimension of UHI imaging) of a given spectral band on two datasets corresponding to two distinct cruises: Katchof (left) and Chereef (right). The top panels correspond to data in units of DN while the bottom ones correspond to data in radiance. The 5% quantile (green), median (blue) and mean (orange) curves were computed in each case considering the same 100 successive temporal images, corresponding to about 100,000 temporal samples per spatial pixel. The manufacturer calibration data were added to these plots as the red curve for comparison. An expected decrease in brightness toward the edges of the camera can be seen in the data. Under homogeneous lighting conditions, it would have been expected that the green, orange and blue curves become flat once the data have been calibrated in radiance. However, this is not the case for the statistical curves in radiance (**Figure 34**, bottom) as they present opposite patterns as compared to the one in DN. In addition, it can be seen that a strong correlation exists between these patterns and the increase in the associated calibration (red) curve, especially for the Katchof cruise. This tends to show that the manufacturer calibration data are overcorrecting the vignetting phenomenon occurring within the camera. In addition, regarding the Chereef cruise, it can be observed that an overall decrease exists of the statistical curves from the first pixel (index 0) to the last pixel (index 1920) across the track. It can thus be assumed that a decrease in the light source intensity along this direction exists.





**Figure 34.** Temporal statistics (5% quantile, median and mean) of cross-track pixel values for 200 successive images (197,400 temporal acquisitions) for the Katchof (left) and Chereef (right) cruises with data in DN (top) and radiance (bottom). The red curves are computed from the manufacturer-provided calibration spectrum data.

### Reflectance computation

The upwelling radiance  $L_{up}$  measured by the UHI camera not only reflects the optical properties of the natural environment but is also closely linked to the spectral characteristics of the artificial light source. Converting  $L_{up}$  into sensor reflectance  $R_{sensor}$  enables normalising the data with respect to the downwelling irradiance  $E_{down}$  light source. The reflectance is obtained as follows:

$$R_{sensor}(\lambda_i) = \frac{\pi \cdot L_{up}(\lambda_i)}{E_{down}(\lambda_i)} \quad (3)$$

Compared to passive remote sensing when the sun is used as a point-source, estimating  $E_{down}$  in the context of underwater imagery is challenging due to the use of several artificial lamps. Here  $E_{down}$  is estimated from the radiance spectrum of one of these light sources, as measured by the UHI in laboratory. In order to take the water attenuation into account, which is occurring between the light source and the sensor position,  $E_{down}$  can be derived from the light source radiance  $L_{lab}$  as measured in a controlled environment using the following equation:

$$E_{down}(\lambda) = \pi \cdot L_{down}(\lambda) \cdot e^{-k(\lambda) \cdot h} \quad (4)$$

with  $k$  the downwelling diffuse attenuation coefficient of water and  $h$  the height difference between the light source and the camera positions. It can be seen in **Error! Reference source not found.** that the resulting value  $E_{down}$  has two emission maxima, one located around 450 nm and a second one located around 600 nm.



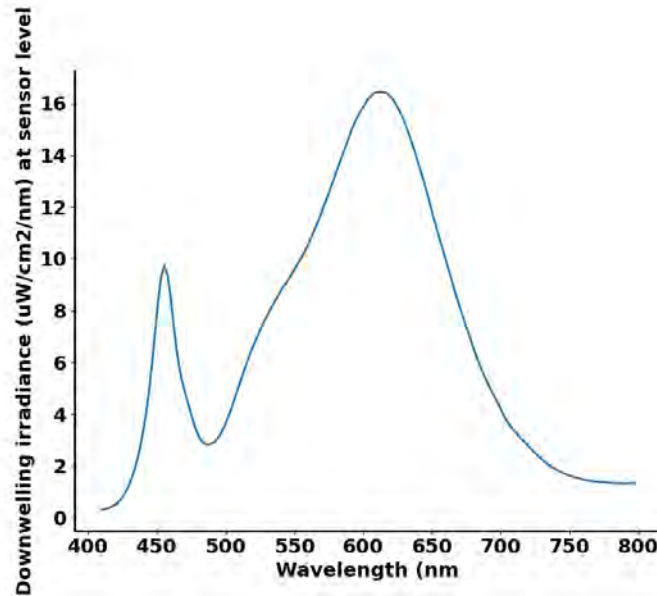


Figure 35. Downwelling irradiance at sensor level estimated from a laboratory measurement of the ROV lights.

### Water body thickness estimation

The reflectance at the sensor level  $R_{sensor}$  is not only a function of seabed albedo but also of the optical properties of the water body, mainly depending on its thickness but also on its optically active constituents and their concentrations. In the current context of this benthic study the bottom reflectance  $R_{bottom}$  was estimated from  $R_{sensor}$ . Following the widely used approach of Lee et al.<sup>32</sup>, this implies modelling  $R_{sensor}$  as a function of  $R_{bottom}$ , the water body thickness  $z$ , the water diffuse attenuation coefficient  $k$  as well as the photon distribution functions  $D_{uc}$  and  $D_{ub}$ , and the water reflectance  $R_{water}$ . This is achieved through the following Beer-Lambert-based relationship:

$$R_{sensor}(\lambda) = R_{water}(\lambda) \left( 1 - e^{-k(\lambda) \left( \frac{1}{\cos \theta_s} + \frac{D_{uc}(\lambda)}{\cos \theta_v} \right) z} \right) + R_{bottom}(\lambda) e^{-k(\lambda) \left( \frac{1}{\cos \theta_s} + \frac{D_{ub}(\lambda)}{\cos \theta_v} \right) z} \quad (5)$$

with  $\theta_s$  and  $\theta_v$  the zenithal angles of the source and the camera, respectively. The second step of the approach of Lee et al. is to co-estimate all the unknown parameters together by iterative minimisation of the distance between the modelled and measured sensor reflectance. However, this method was designed for aerial remote sensing of coastal waters, with highly varying optical properties due to variations in phytoplankton and terrestrial inputs concentrations. These concentrations may be set to constant values as only minor spatial variations are expected in deeper areas and this greatly simplifies the numerical problem to be solved, partly avoiding overfitting issues. Then the  $R_{bottom}$  estimation can be conducted in two successive steps: (i) estimating the water layer thickness  $z$  and the bottom reflectance amplitude with a neutral (sand) bottom reflectance spectral shape fixed within the model, and (ii) computing  $R_{bottom}$  directly from Equation 4 using previously estimated values for the other parameters of this equation.

Figure 36 shows an RGB composition of the sensor reflectance. The greenish aspect of this image is due to the spectral shapes of absorption and scattering occurring within the water layer present between the camera and the sea floor. The bottom reflectance is presented on the right panel of Figure 37. It can be noticed that the greenish aspect is no longer present which proves the effectiveness of the radiative transfer-based inversion. Two products derived by inversion are shown on Figure 37. The water layer

thickness (Figure 37, left) shows values between 1.5–2.8m, which is consistent with altitude data recorded by the robot vehicle. The high correlation observed between decrease in this thickness and the presence of objects on the sea floor provides additional support to the effectiveness of the inversion procedure. The seabed albedo at 550 nm shown on the right panel of Figure 37 has values ranging between 0.02–0.1. Its values gradually increase from the bottom to the top of the image, which may be linked to potential concomitant increase in the scene illumination by the artificial light source.

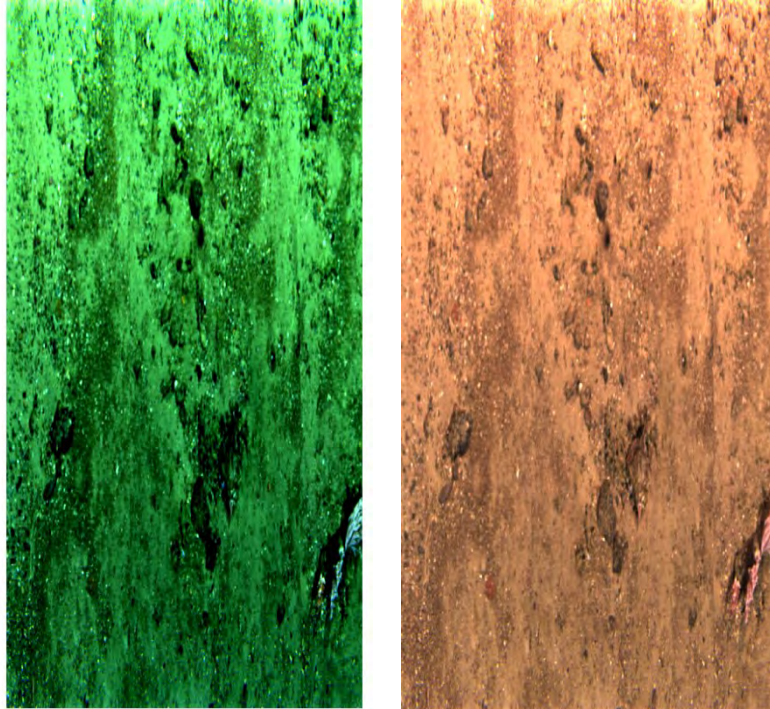


Figure 36. RGB image of sensor (left) and bottom (right) reflectance of UHI data of the Chereef 2022 cruise.

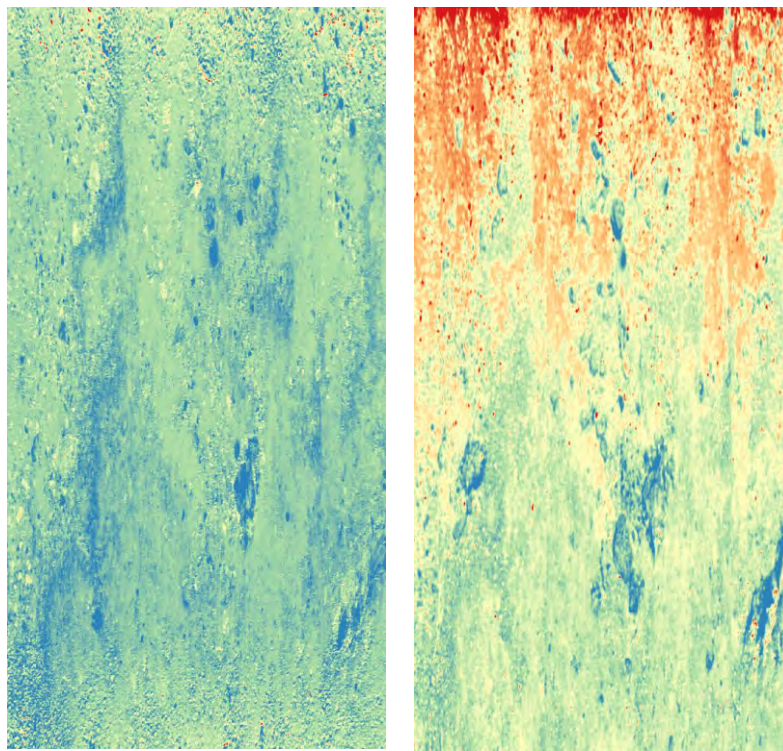
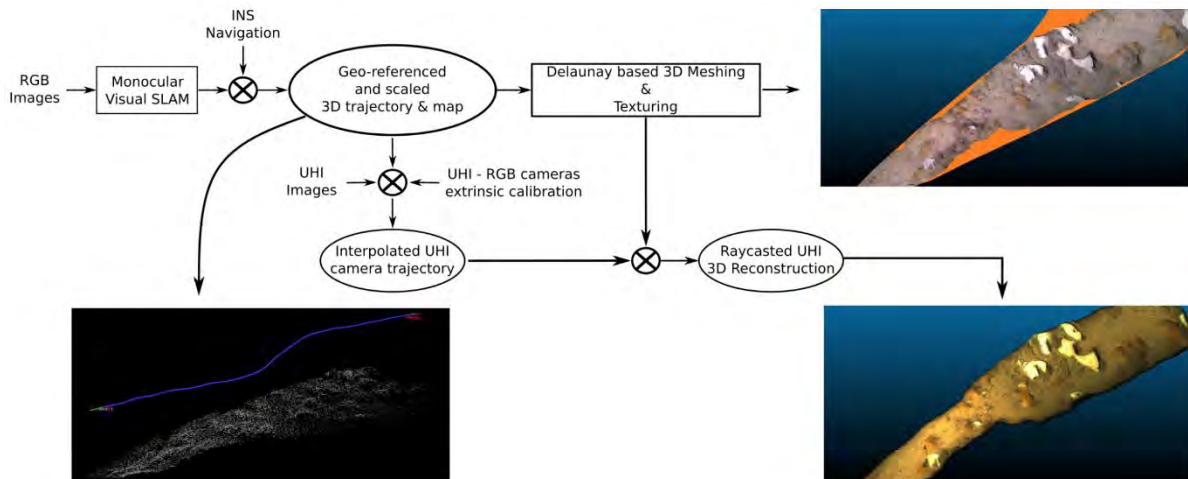


Figure 37. Water body thickness (left, with values ranging between 1.5–2.8 m from blue to red) and bottom albedo at 550 nm (right, with values ranging between 0.02–0.1) estimated through sensor reflectance inversion.

## Geometric correction

The push-broom technique used for the experimented UHI relies on the navigational data to provide an image data cube from recorded spectral lines. In the framework of the *iAtlantic* project, IFEMER designed and implemented a method for creating accurate underwater hyperspectral 3D reconstructions by fusing the measurements of a UHI camera, an RGB camera and an Inertial Navigation System (INS) embedded on an ROV<sup>29</sup>.

The UHI sensor was coupled with a HD video camera and mounted on a tiltable mechanical support. This enabled acquiring data on any type of terrain, from flat to very steep areas. Furthermore, thanks to the accurate CAD model of the HROV, we obtained an approximate 3D transformation (extrinsic calibration) between the HD and UHI cameras. This method is based on the use the RGB camera to accurately estimate the trajectory followed by the ROV at the video's frame-rate in conjunction with the INS predictions for scaling purposes. This allows us to interpolate the trajectory followed by the UHI camera. In addition, we produced a dense 3D reconstruction using the acquired RGB images that we then linked to the UHI images thanks to an approximately known transformation between both cameras. Doing so, we were able to map the 3D mesh with hyperspectral textures, thus producing accurate and reliable 3D hyperspectral reconstructions that can be used to produce non-distorted ortho-images for further scientific analysis. The proposed pipeline is illustrated in [Figure 38](#).



**Figure 38.** Pipeline of the method for computing hyperspectral 3D reconstructions, from (Ferrera et al., 2021).

We used the RGB camera mounted on the ROV to estimate the trajectory followed by the UHI. More precisely, we employed a monocular visual Simultaneous Localisation and Mapping (SLAM) algorithm to get the estimate of the trajectory followed by the RGB camera. In addition, we fused the visual SLAM estimates with the prediction from the INS embedded on the ROV to recover scaled and geo-referenced estimates. Using the scaled SLAM results, we then computed a dense 3D mesh of the surveyed environment. Finally, we used the approximately known extrinsic calibration between the RGB camera and the UHI to get an estimate of the UHI trajectory and then raycast the hyperspectral images on the dense 3D model. For more method description, see Ferrera et al. and Kumar et al.<sup>29,33</sup>.

## 4.4 Benthic feature extraction use cases

### Seabed benthic habitat mapping (Catchoff Mediterranean site)

The data for this use case were acquired at the Catchoff site in the Mediterranean Sea during a UHI integration test survey with HROV *Ariane*. This site is recognised to host a habitat of gorgonians. Indeed, coralligenous reefs, owing to their extent, species diversity, productivity, and carbonate production, are recognised to have a paramount role in the Mediterranean coastal system.

First, we extracted from the image recorded in DN, spectra on pixels corresponding to species or types of background recognised using video images acquired simultaneously with the hyperspectral images. To determine the discrimination potential of UHI, spectra were standardised by a homogeneous sand spectrum before applying a multivariate analysis based on PCA. Only the spectral responses in the visible were retained knowing the level of noise and absorption power in the PIR spectral range (Figure 39).

Based only on the two first principal components, resulting PCA plots shows a great potential for discrimination except for *Bonellia* echiuran worm species that display a spectral characteristic overlapping that of hard substrate (Figure 39).

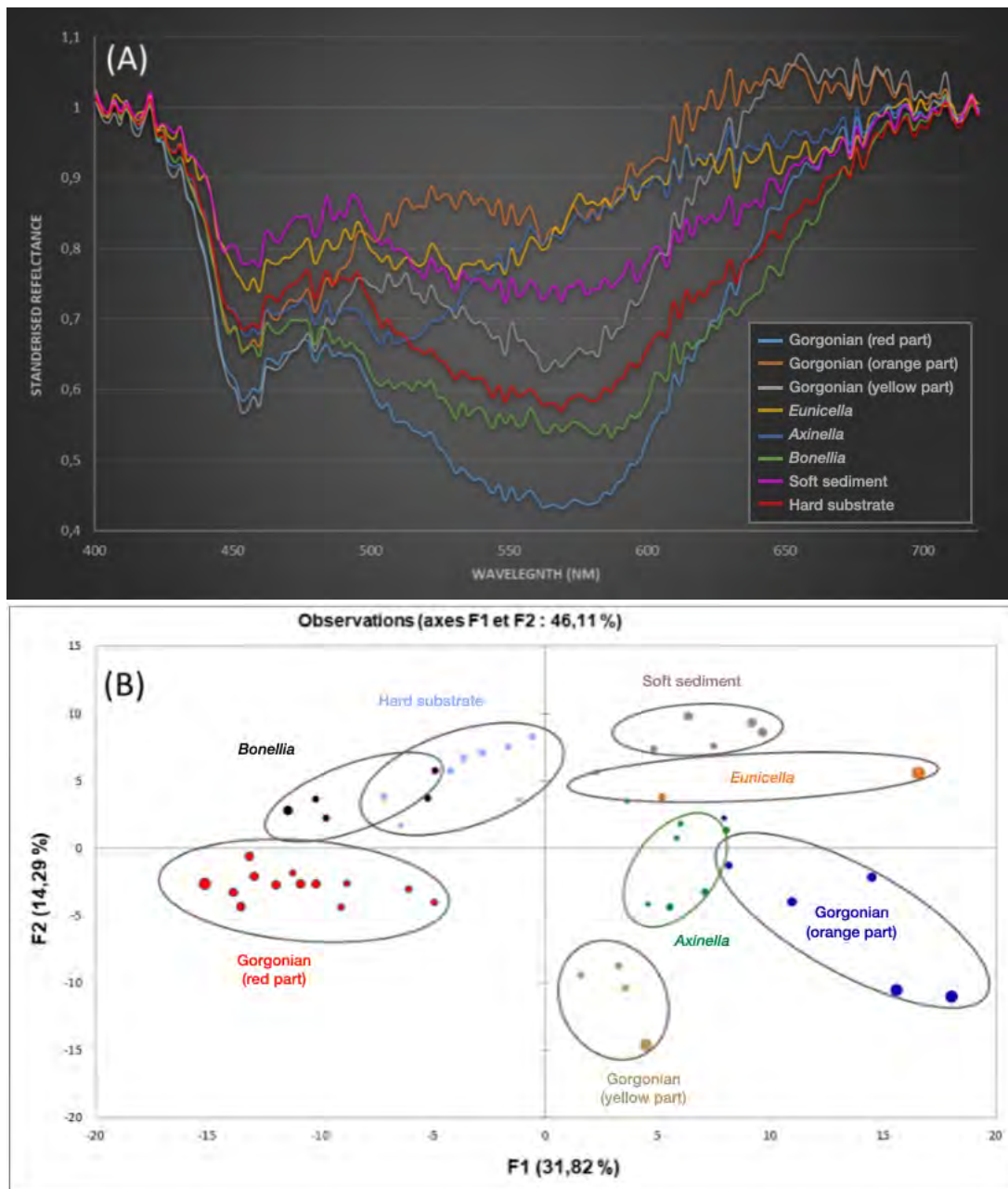


Figure 39. Standardised spectral signatures of the main species and substrate types of the Catchoff Mediterranean site (A) and scatterplot of the first two principal components of a PCA (B).

The first analysis of the suitability of underwater hyperspectral imagery for the automatic classification of habitats is still ongoing. It became obvious that pre-processing is essential to (i) remove noise from the

recorded signal (Figure 40) and (ii) to perform geometric and radiometric calibration before geospatial processing.



Figure 40. Zoom-in to a selected part of a UHI recording showing noisy spectral bands along and across-track.

### Mapping the ecological status of cold-water corals (Lampaul Atlantic site)

Regarding the exploitation of the acquired UHI data to evaluate the potential for detection and discrimination of bottom types, we used the Chereef campaign in the summer of 2021 to collect samples of coral species and substrate by the ROV *Ariane* that is equipped with a manipulator and sampling arm. It was then possible to carry out measurements of spectral signatures in the onboard laboratory by the UHI. Figure 41 shows the raw signal for live and dead coral compared to a sandy substrate. It can be seen that these signatures have spectral amplitudes and reflectance maxima at different wavelengths. This is promising for the potential for automatic discrimination of live and dead coral from substrate from UHI images.

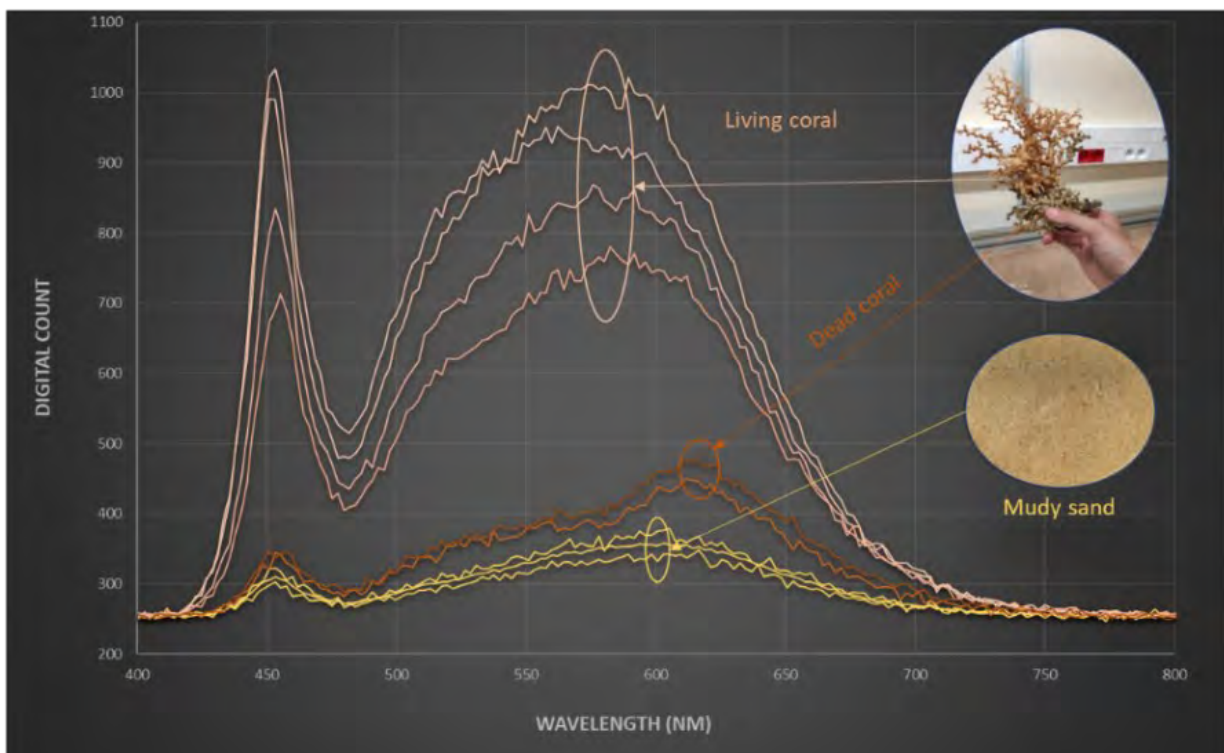


Figure 41. Spectral signatures recorded in the laboratory from samples collected during the 2021 Chereef cruise to Lampaul Canyon (France).

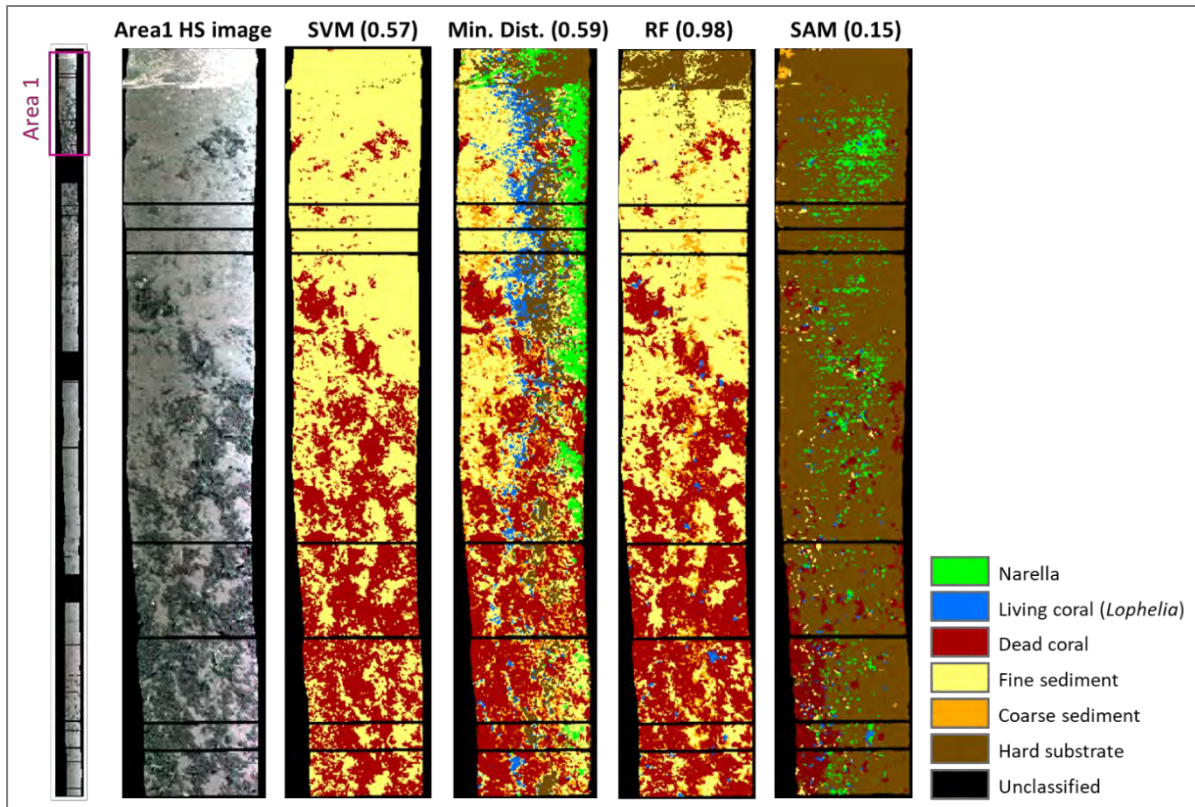


**Figure 42.** UHI RGB visualisation (A), RGB image from video recorded simultaneously (B), and regions of interest (ROIs) created for training and validating UHI image classifications.

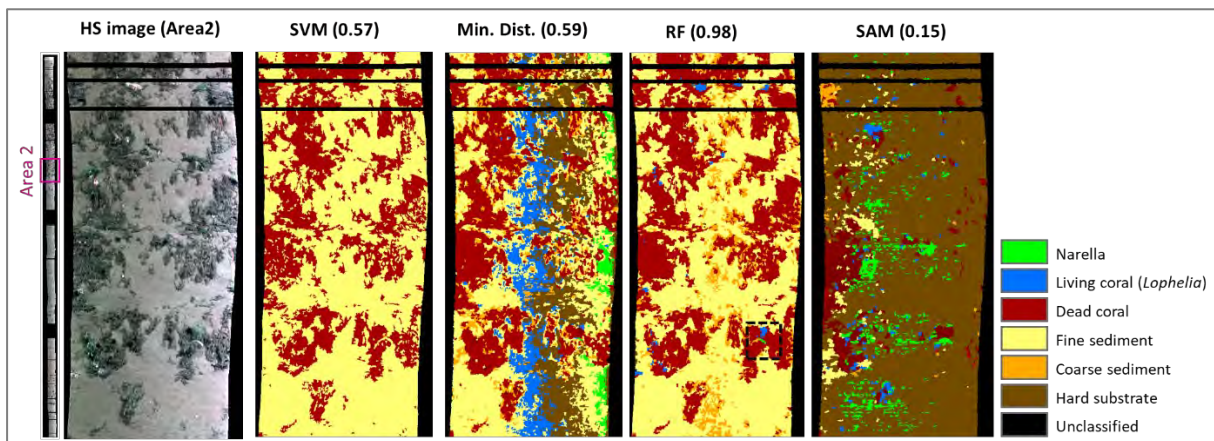
For this use case, we have chosen to process a long profile, which best represents the diversity of habitats found at the site. **Figure 42** shows a snapshot of the UHI data, compared to the RGB footage from the same scene as well as manually annotated patches of the UHI data, used for training ML methods. **Figure 43** shows a comparison of the classification results for the different algorithms tested for the selected subarea 1 of this profile. The black lines that appear on the hyperspectral images correspond to interrupted recordings due to a malfunction of the UHI.

In general, as well as for the entire profile, the spectral angle mapper (SAM) with an overall accuracy of 15% presents the worst performance. Most habitats are poorly detected and erroneously assigned to the category 'Hard substrate' by this classifier. Slightly better performance was achieved by the Support Vector Machine (SVM) and Minimum Distance (Min.Distance) methods, which display accuracy values of 57% and 59% respectively. These two algorithms, if they detect the 'Dead coral' well, show inferior performance for the other habitat categories. SVMs are not able to discriminate 'Living coral' from other categories. With an overall accuracy of 98%, the Random Forest (RF) method is the algorithm that provides the best result of discriminating seabed types and cold-water coral (CWC). These observations are supported by the results of area 2 (**Figure 44**). While Min.Distance and SAM overestimate the presence of the octocoral *Narella* and SVMs does not detect this category at all, the RF method correctly identified the only individual present in this area.

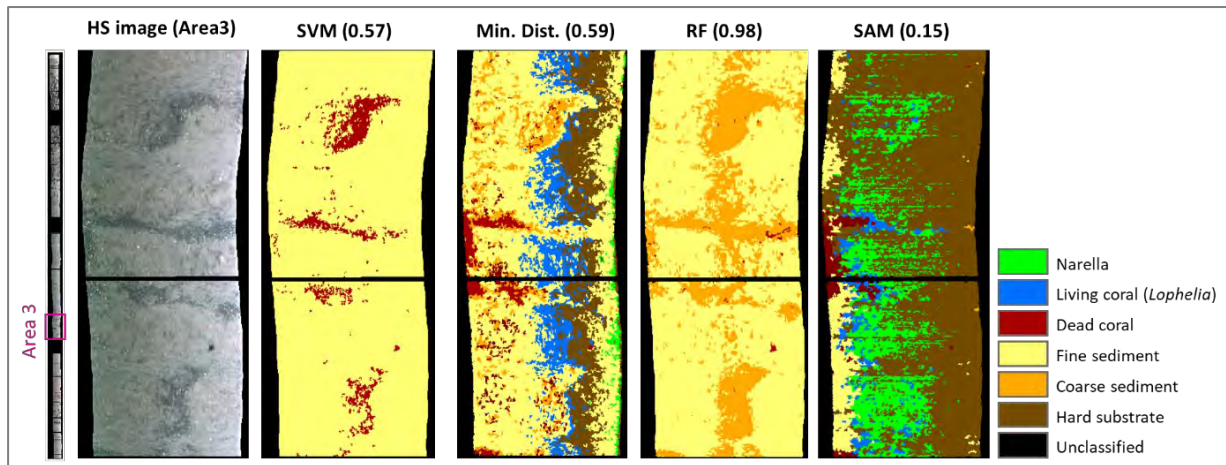
In area 3, only soft sediments are present while other habitat categories are incorrectly detected by different algorithms, in particular the Min.Distance and SAM methods. Once more, this result shows the best performance by the RF in this area for which this algorithm also shows few confusions between the 'Coarse sediment' and the 'Dead coral' categories.



**Figure 43.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied on area 1 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the classifiers tested.



**Figure 44.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied on area 2 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the classifiers tested. The area bounded by the black dotted rectangle within the RF classification shows the detection of an individual of category ‘Narella’, the only place where this species is present in this area.



**Figure 45.** UHI RGB visualisation and comparison between supervised classifications obtained from Support Vector Machine (SVM), Minimum Distance (Min.Dist.), Random Forest (RF) and Spectral Angle Mapper (SAM) algorithms applied to area 3 of the entire track, delimited by the pink rectangle in the left-most part. The values in brackets correspond to the overall accuracies of the evaluated classifiers.

Table 9 shows the confusion matrix for the RF algorithm as applied to the entire track. The confusions are mainly seen between the categories ‘Hard substrate’ and ‘Fine sediment’ as well as between ‘Dead coral’ and ‘Coarse sediment’ with misclassification errors of the order of 11% and 3% respectively. It is the class of ‘Living coral’ which shows confusion with a greater number of habitat categories, but the misclassifications remain between 0.2% and 1.6%.

**Table 9.** Confusion matrix for assessing the accuracy of the Random Forest algorithm applied to classify pixels along a UHI data transect acquired in Lampaul canyon.

| Overall accuracy (98%) |                 | Ground truth    |                |             |               |              |          | Total |
|------------------------|-----------------|-----------------|----------------|-------------|---------------|--------------|----------|-------|
|                        |                 | Coarse sediment | Hard substrate | Dead coral  | Fine sediment | Living coral | Narella  |       |
| Predicted (RF)         | Unclassified    | 0               | 0              | 0           | 0             | 3            | 0        | 3     |
|                        | Coarse sediment | 17488 (99 %)    | 0              | 147         | 0             | 25           | 0        | 17660 |
|                        | Hard substrate  | 0               | 6234 (89 %)    | 0           | 25            | 18           | 0        | 6277  |
|                        | Dead coral      | 92              | 0              | 5405 (97 %) | 0             | 21           | 0        | 5518  |
|                        | Fine sediment   | 0               | 768            | 1           | 21578 (99%)   | 21           | 0        | 29498 |
|                        | Live coral      | 0               | 0              | 0           | 0             | 1201 (93 %)  | 0        | 1201  |
|                        | Narella         | 0               | 0              | 0           | 0             | 0            | 34 (1 %) | 34    |
|                        | Total           | 17580           | 7002           | 5553        | 28733         | 1289         | 34       | 60191 |

From the Random Forest classification, spatial statistics were computed for the entire track, which covers a total area of approximately 118 m<sup>2</sup>. Within this track, CWCs represent only 16.9% of the total surveyed profile. The proportion of ‘Living coral’ is comparatively low, representing only about 2% of the coral fragments in this area. The present results are only the initial steps of the data planned interpretation for this UHI data set. Postponement of surveys due to the global COVID-19 pandemic led to substantial delays in data acquisition, but this work will be continued by the IFREMER team to enrich the acquired data with other metrics such as geomorphology, landscape ecology and functional traits.

#### 4.5 Recommendations towards a UHI data acquisition protocol

Based on our experimentation from all carried out surveys, some best practices are summarised below for ensuring optimal quality of hyperspectral image data using an ROV.

##### Survey preparation

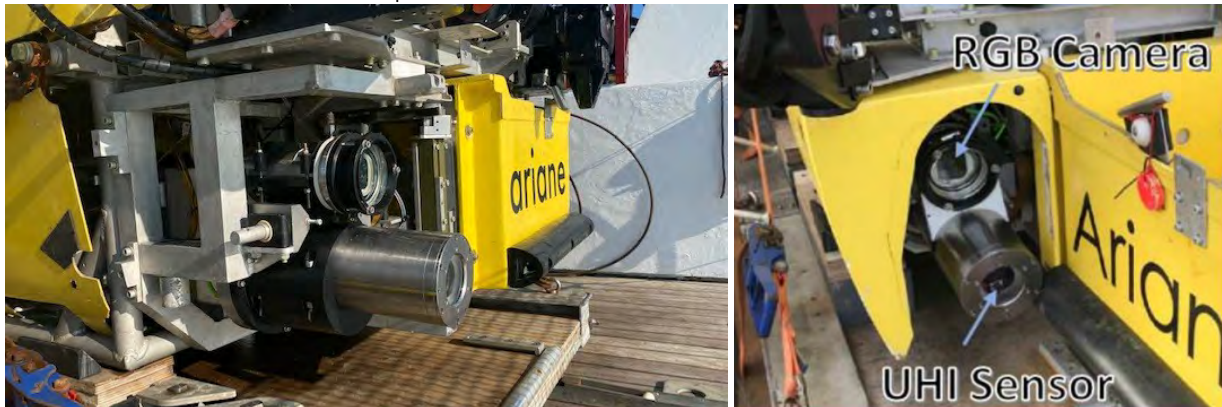
- Ensure that all research permits, safety plans and approvals have been created and obtained (General research preparation requirement, not specific to UHI).



- Define the required transect design: the decision to which transect design is most appropriate is driven by the question being addressed, as well as the environment, available time, and logistics of AUV deployment.
- Ensure that all cameras to be used are correctly calibrated.
- Ensure that required processing software is installed on onboard computers and that users have previously been trained to their use.
- Ensure that sufficient data storage capacity is available onboard for on the planned acquisition.

### Deployment preparation

- Ensure that the UHI camera and HD camera are correctly mounted and aligned (30).
- Ensure that appropriate illumination is mounted in a way that the scene imaged by the hyper-spectral imager (HSI), and other cameras is uniformly illuminated.
- Ensure that enough free data storage capacity is available on the HSI storage in case of offline data recording.
- Ensure that the power supply for the camera systems and illumination system is operational.
- Ensure that all communication channels and the underwater navigation systems are operational.
- Potentially ensure that additionally required equipment, e.g., for underwater image calibration, is available to the robotic platform.



**Figure 46.** Stereo setup of the two cameras on *Ariane* HROV from IFREMER. The upper pressure housing contains the RGB camera. The bottom pressure housing contains the push-broom HSI camera from Ecotone. The Z axis of the camera coordinate system is in the view direction, the X-axis points to the right and the Y-axis is pointing downwards. The cameras are mounted vertically, thus are having the y-axis as known the baseline between them.

### UHI deployment

- Monitor the ROV descent to the seabed and towards the start of transect location by using appropriate software.
- Ensure accurate recording of image metadata to enable creating FAIR UHI data, e.g., based on the image FAIR Digital Object (iFDO) format described in the next chapter.

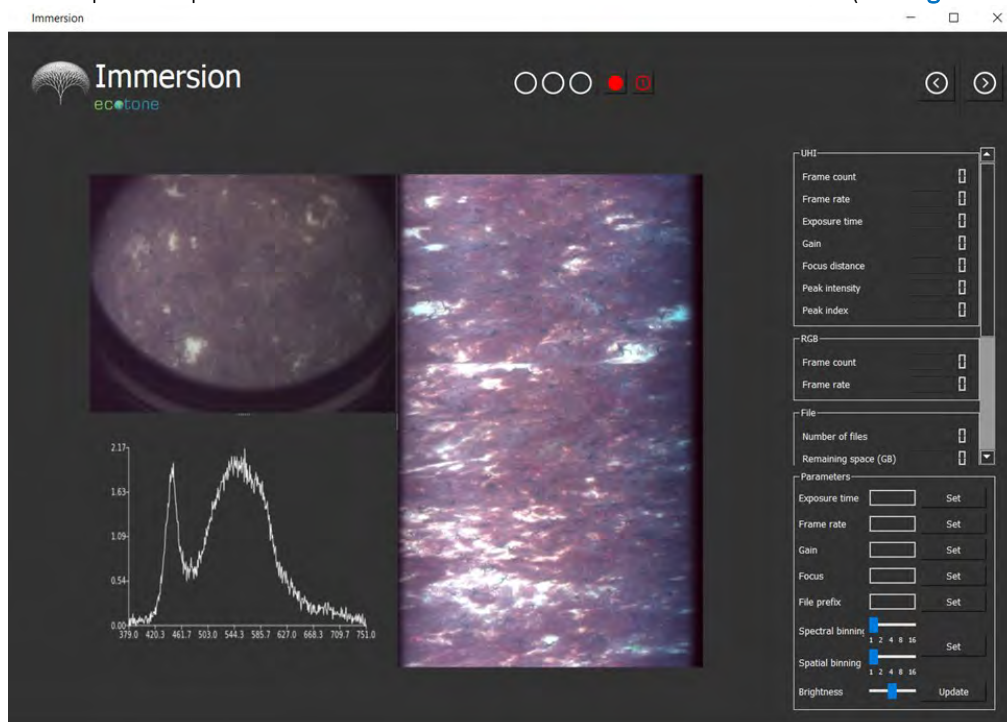
Minimum metadata recorded for each transect should contain:

- Survey name/identifier (iFDO field: image-project)
- Station name/ID (iFDO field: image-event)
- Platform (iFDO field: image-platform)
- Geographical coordinate (iFDO fields: image-latitude, image-longitude, image-coordinate-reference-system)

It is recommended to use the WGS1984 system at a minimum of 7 decimal places to reach meter accuracy.

- Altitude (iFDO field: image-altitude)
- UTC date and time (iFDO field: image-datetime)
- ROV orientation (iFDO fields: image-camera-yaw-degrees, image-camera-pitch-degrees, image-camera-roll-degrees or image-camera-pose)
- ROV speed
- Navigation data quality such as type of navigation system used and its associated errors (iFDO fields: image-coordinate-uncertainty-meters, image-navigation)

- UHI acquisition parameter settings such as integration time, frame rate, gain, ... (iFDO field: image-acquisition-settings)
- Comments on the deployment (iFDO fields: image-abstract, image-objective, image-target-environment, image-target-timescale, image-spatial-constraints, image-temporal-constraints)
- Acquisition of radiometric calibration data once for each new acquisition site by deploying a calibration grid marker that is imaged by the UHI system in-situ. For radiometric calibration, it is crucial to measure the spectral signature of the illumination system being used as well.
- For data acquisition with the Ecotone system in a specific targeted area, the following steps are recommended:
  1. Positioning of the ROV at the planned distance from the bottom depending on the targeted spatial resolution.
  2. Recording the required acquisition metadata.
  3. Start the recording for the HD camera.
  4. Start the recording for the UHI sensor from the Immersion software interface.
  5. Turn the lights off, wait about 2 seconds, then turn the lights back on to synchronise the two cameras HD camera for 3D reconstruction and the UHI camera signals.
  6. As you acquire image data, continuously fill in a dive protocol to note modifications of acquisition parameters as well comments on the recorded habitats (see [Figure 31](#)).



**Figure 47.** Screenshot of the Immersion software package provided by Ecotone for the operation of their UHI camera.

### Post-deployment

- After the last deployment of the day, wash down the UHI with fresh water.
- Download the recorded raw image data from the cameras.
- Collect the associated navigation data.
- Ensure accurate recording of metadata and finalise the deployment protocol.

## 4.6 Conclusion

### Lessons learned from UHI experimentation in the context of *iAtlantic*

The implementation of UHI technology has been deployed for the first time in France, thanks to the synergy between two European projects: MarHa and iAtlantic. The sensor used is the UHI (Underwater Hyperspectral Imager) marketed by ECOTONE. It was acquired as part of the MarHa project.

Several test surveys were necessary to be able to integrate the UHI into different ROVs, finally made possible thanks to the technical support of the ECOTONE teams. The surveys were carried out in the Atlantic Ocean and coastal areas of the Mediterranean Sea.

The main difficulties we identified are related to the geometric and radiometric calibration of the acquired images. Targeted technology developments have thus been carried out by the IFREMER teams to improve the acquired UHI data for scientific exploitation. To date, it was not possible to process all the acquired data or test all the planned processing methods. First results presented above already show that the hyperspectral imaging technology has the potential to automatically extract relevant information on deep-sea environments, aiming at spatio-temporal ecosystem monitoring.

### Choice of UHI gear

The first step when entering the field of UHI is to choose an appropriate sensor for the intended use case<sup>22</sup>. Two types of sensors received particular attention during this study (Table 10).

**Table 10.** Comparison of underwater hyperspectral imaging system concepts.

| UHI sensor concepts | Spectral Resolution | Imaging Speed | Geometric Correction | Usability  |
|---------------------|---------------------|---------------|----------------------|--|
| Push-broom          | High                | High          | Difficult            | Commonly used, very suitable for seabed mapping in relatively flat areas |
| Snapshot            | Low                 | High          | Easy                 | Suitable for moving targets, may become a new trend in the future        |

Although the push-broom and snapshot type sensors both have high acquisition costs to date, they nevertheless have technical characteristics that distinguish them, which can guide the choice in terms of spectral resolution and acquisition speed requirements as well as according to the difficulty of pre-processing, especially for geometric corrections.

While the commonly used push-broom sensors, with its high spectral resolution, is suitable for seabed mapping in relatively flat areas, the low spectral resolution snapshot architecture has the ability to acquire 2D images by a single exposure. This method might thus become the state-of-the-art in the future<sup>22</sup>.

### UHI Deployment

Our experience with the UHI test deployments meets the recommendations of the literature and more specifically those summarised by Fogliani et al. on the mapping of benthic habitats<sup>28</sup>. In addition to the protocol described above, three aspects should be considered, in particular, for a better quality of underwater hyperspectral data acquisition:

- Platform behaviour: to ensure a rigorous UHI survey, an efficient positioning system and accurate and dense navigation data for the camera platform are necessary. High stability, constant speed and altitude above the seafloor are beneficial for creating robust UHI data.
- Light source: currently, there is no standardised light source designed for the UHI<sup>22</sup>. Thus, an appropriate lamp system needs to be rigorously adapted to uniformly illuminate the surveyed area to avoid spatial changes in light and presence of shade. Any deployed light source should have a sufficient intensity to illuminate targets depending on water depth, water body thickness and sunlight available. The spectrum of the light source needs to be measured in a controlled environment for qualitative and quantitative analysis. This is particularly important when planning radiometric

calibration.

- Ancillary sensors and equipment: these can include (i) an HD RGB camera mounted with a similar view direction like the UHI camera to record concomitantly the seafloor for either target identification or to support geometric correction; and (ii) sensors to measure the optical properties of the water body through which the UHI is capturing images. (e.g., Chlorophyll a, 'coloured dissolved organic matter' (CDOM), or the total suspended matter (TSM) amount); and (iii) a radiometric calibration target for in-situ reflectance quantification.

### **The importance of data pre-processing**

Achieving the goal of monitoring seafloor habitats spatio-temporally with UHI sensors requires both accurate and standardised data to be able to identify gradual changes as well as sudden changes without bias. It is therefore important to consider two types of UHI data pre-processing before conducting any extraction of semantic information from the images.

The first pre-processing concerns radiometric calibration: UHI signatures are distorted due to absorption and scattering of light by the water layer present between the camera and the sea floor. This implies that a prior calibration for the conversion of raw UHI data into reflectance measurements is needed. So far, benthic UHI studies have either ignored this distortion or have used sophisticated robotic systems for deploying an optical reference plate during the UHI acquisition<sup>34</sup>. However, this solution is difficult to operate in practice and does not consider variations in the optical properties of the water layer due to varying bottom distance (e.g., in the presence of corals or rocks). The radiative transfer-based method of Lee et al.<sup>32</sup>, developed for airborne hyperspectral imaging, is one of the possibilities for estimating bottom reflectance. The radiometric corrections require information recorded by the UHI sensor alongside the raw data. It is recommended to apply the radiometric corrections before applying geometric corrections.

The second pre-processing is relative to the geometrical correction: the comparison over time or the extraction of metrics on the surface require precise geographical positioning. These can be achieved by a precise navigation system, but the present study has shown that 3D reconstruction based on structure-from-motion methods can significantly improve the geometric correction qualities of UHI data.

### **Acknowledgement**

We are grateful to the ECOTONE company team, more particularly Ivar Erdal, Ingrid Myrnes Hansen, Øystein Bergsagel and Jacob Støren, for the valuable technical support they provided us throughout this UHI integration and experimentation project. We sincerely thank all the heads of the different cruises namely Lorenzo Brignone (MarHa 2019), Ewen Raugel (ESSHROV 2019), Laurent Artzner (ESSHROV 2020-21), Lénaïck Menot and Julie Tourolle (CheReef 2021, 2022) as well as the crews of the different research vessels that allowed us to integrate, to test and to deploy this technology. We would also like to thank Marie Claire Fabri and Stéphane Sartoretto for all their help in identifying coralligenous habitats at the Catchoff site.

## 5. FAIR marine image metadata

Contributions from: Timm Schoening and the co-authors of the peer-reviewed paper that contributed to this section<sup>35</sup>

### 5.1 What is FAIR data?

The vision behind *FAIR data* is to create a world where data are treated as a *valuable and shared resource* that is accessible to all, and where data are used in a way that is transparent and accountable<sup>36</sup>. FAIR data seeks to ensure that data are collected, stored, and used in a way that is *transparent* to all stakeholders, including individuals, communities, organisations, and society as a whole<sup>37</sup>.

At its core, the idea of FAIR data is based on four guiding principles:

- 1) *Findable*: Data should be easy to find and discover by those who need it. This means that data should be well-documented, with clear and consistent metadata that allows others to understand what the data are, how it was collected, and how it can be used. Additionally, data should be assigned a unique and persistent identifier, such as a Digital Object Identifier (DOI), that enables others to easily locate and reference the data.
- 2) *Accessible*: Data should be available to all who have a legitimate interest in it. This means that data should be stored in an open format that is easily accessible, and that barriers to access should be minimised.
- 3) *Interoperable*: Data should be able to be used together with other data, regardless of where it came from or how it was collected. This means that data should be standardised, with consistent formats and structures that allow different data sets to be combined and analysed together.
- 4) *Reusable*: Data should be able to be used and reused by others. This means that data should be free to use, with clear licensing and copyright agreements, and that data should be preserved and curated to ensure that it remains useful over time.

By adhering to these principles, the FAIR data idea seeks to create a world where data are used for the greater good, rather than for the benefit of a few<sup>38</sup>. At their core, the FAIR data principles aim to enable answering the scientific questions of the future.

### 5.2 How does FAIR data lead to better science?

By providing open and accessible data, FAIR data practices enable researchers to *reproduce experiments and studies*<sup>39</sup>. This promotes transparency and accountability, allowing others to validate research findings and ensuring that scientific results are reliable and trustworthy. FAIR data practices *facilitate collaboration* among researchers, as data can be easily shared and integrated across different research groups and disciplines<sup>40</sup>. This enables researchers to pool resources and expertise, leading to *more comprehensive and impactful research* outcomes<sup>37</sup>.

FAIR data practices can *spur innovation* by enabling new types of analyses and discoveries. Access to large, diverse, and high-quality datasets can help researchers identify new patterns and connections, leading to new insights and discoveries. FAIR data practices can *increase efficiency* in the scientific research process<sup>40</sup>. Researchers can build on existing data and results, rather than starting from scratch, saving time and resources. FAIR data practices *promote accountability* by requiring researchers to clearly document and describe their methods and data<sup>37</sup>. This ensures that research findings are based on sound methodology and data that can be verified and validated.

### 5.3 What are FAIR marine images?

How to achieve FAIRness for different data types or data use communities is an ongoing research topic<sup>38,39,41</sup>. Organisations such as the European Open Science Cloud (EOSC) or the Research Data Alliance (RDA) publish recommendations on FAIRness, some of which are directly applicable to marine image data. These include human- and machine-readable standardised vocabularies or the use of persistent identifiers (PIDs). However, the massive data set sizes in marine imaging create additional requirements

for FAIRness such as the detachment of the image data from associated data (called metadata) to maintain efficient handling of data and facilitate accessibility and reuse. Finally, the inherent lack of semantic structure in image data, i.e., that information in images is encoded in millions of pixels that only together create a picture by means of interpretation by a brain or computer, requires standards for tools that create and use information derived from images such as annotations<sup>7</sup>.

#### 5.4 Image FAIR Digital Objects (iFDOs)

As a solution, the concept of image FAIR Digital Objects (iFDOs) was developed in the context of the iAtlantic project as well as an infrastructure environment to create and exploit such iFDOs<sup>35,42</sup>. The iFDO format solves the issues of technical heterogeneity and semantic un-structuredness. The infrastructure environment solves the challenge of handling vast image data sets. But only in combination can the iFDOs and the FAIR marine image infrastructure environment that works with iFDOs achieve FAIRness of marine images<sup>35</sup>. iFDOs are designed to be applied to all marine image data: in-situ and ex-situ imagery, photos and videos and datasets consisting of single images to thousands of images.

With the iFDO concept implemented and the required infrastructure environment in place, most recommendations on FAIRness can be addressed.

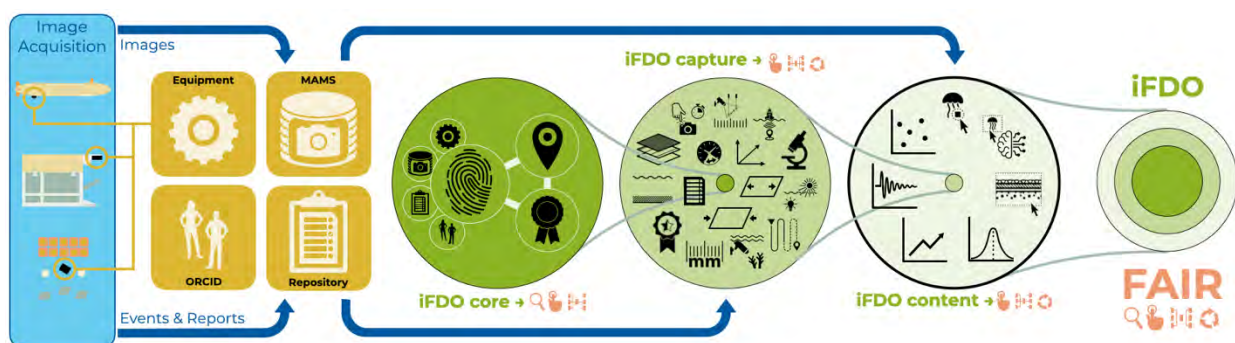
**Table 11** provides an overview on the status for marine images using iFDOs and the infrastructure environment. Of the 41 RDA recommendations, 36 are achieved, 2 are not viable in the image data context and 3 are resolved by the currently developed iFDO update. All essential and important FAIR indicators are being addressed by iFDOs and the marine image infrastructure environment.

## 5.5 iFDO description

An iFDO file consists of two *iFDO parts*: the image-set-header part and the image-set-items part. The header part contains default values for all items. The image-items part contains all values that deviate from the default values for this specific image item. They supersede the default values. All metadata of an image item is provided in a list. For moving images, the first entry of the list contains an object of those metadata fields that are defaults for an entire video. All subsequent list entries correspond to specifications of the metadata for one given timepoint of a video. For still images this list may contain only one entry of its field values or it is rather a single object of field values for this particular photo.

Making data FAIR requires that data are assigned a persistent identifier (PID). Many such PID systems exist, and usually they are based on the handle system (e.g., DOI, ORCID) and alpha-numerical IDs that are globally unique. For iFDOs, it was decided to choose Universally Unique Identifiers (*UUIDs as a base for the PID*) record There are ca.  $5 \times 10^{36}$  possible UUID4s, making it almost impossible that the same UUID is created more than once. This UUID is the alphanumeric identifier that must be assigned to each image and to each image set. The UUID for the image file (a photo or video) must be written into the metadata header of the image file itself to be included into the file hash later (see below). How this step can be done depends on the image file format used. Mostly, software tools like exiftool or ffmpeg are used which is also the case in the MarIQT Python package. The image-set-uuid only needs to be part of the iFDO file and not written to any image file. Using UUIDs for persistent identification is essential to achieve FAIRness by making images uniquely identifiable. Those UUIDs need to be registered in a long-term data repository and data portal to facilitate Findability of images.

iFDOs use *file hashes to monitor the integrity* of images. A hash is a fingerprint of a file, computed from the file's byte content. A file hash can be used to assert that a file is not broken or that a particular file is a specific version. Checking the integrity of a file with hashes requires that the byte content does not change. It is therefore essential, that the UUID is written to the image file's metadata header before the hash for that file is computed!



**Figure 48.** Setup of image FAIR Digital Objects. Key information and image data are stored in a dedicated infrastructure (yellow squares). iFDOs only contain persistent identifiers to those external information resources. Additionally, specific metadata for marine imaging use-cases is stored inside the iFDO files. iFDOs consist of three sections: (1) the required core part which includes the persistent identifiers as well as licensing information; (2) the recommended capture part that addresses the technical heterogeneity of image acquisition; and the (3) the optional content part that captures semantic information from within the images to address the heterogeneous nature of image data. Together, these three sections constitute one iFDO file. This file contains header information on the entire image data set as well as detailed information on each image item within a defined set of images.

The standardised fields in iFDOs are grouped into *three iFDO sections* (see [Figure 46](#)). An entire image set is defined by project-specific metadata (e.g., deployment, station, dive, mission) and requires information

on the ownership and allowed usage of the collection. All these metadata fields are placed in the iFDO core section. Numerical metadata is required for each image to document its acquisition position. The image-set-header part of an iFDO must contain three fields that must not be superseded by values in the image-set-items part. These are image-set-name, image-set-uuid and image-set-handle. All other fields may be provided as default values in the image-set-header, potentially superseded by values in the image-set-items part. These correspond to imaging context (image-context, image-project, image-event, image-platform, image-sensor, image-abstract), the navigation data (image-datetime, image-latitude, image-longitude, image-altitude, image-coordinate-reference-system, image-coordinate-uncertainty-meters), image identification (image-uuid, image-hash-sha256) and the reuse permissions (image-pi, image-creators, image-license, image-copyright).

Information on how image data was captured can be crucial to understand the information subsequently extracted from images. It is thus highly recommended to provide this information as part of the iFDOs capture section. This section is expected to grow with time, as additional (marine) imaging domains make use of the iFDO concept and extend the standard with fields for their specific use case. Existing iFDO capture fields currently fall into four categories. First, iFDO capture fields with restricted values that characterise the image material (e.g., image-acquisition: photo, video, slide; image-deployment: mapping, stationary, survey, ...; image-illumination: sunlight, artificial-light, mixed-light; image-capture-mode: timer, manual, mixed; etc.). Second, numerical information on image capture (e.g., image-area-square-meter, image-meters-above-ground, image-overlap-fraction, etc.). Third, pose and calibration information (e.g., image-camera-housing-viewport, image-camera-calibration-model, etc.). Fourth, information on the scope and limitations of the image acquisition (e.g., image-objective, image-target-environment, image-fauna-attraction, image-temporal-constraints, etc.). Not all iFDO capture fields have to be populated for each data set. But providing a more complete set of field values for a data set is expected to increase its value and reuse potential significantly, in particular for yet unknown uses. Investing into the iFDO capture section will generate the visibility and credit for imaging efforts. By limiting some fields to restricted values, it is possible to classify and filter image data sets in data portals and to rapidly visualise data characteristics. See the iFDO fixed term icons section below for details.

The iFDO content section fields are a mechanism to encode the semantic content of image data. Some fields encode quantitative scalar data extracted from the image material (e.g., image-entropy, image-particle-count, image-average-colour). Other fields encode higher-dimensional feature descriptors (e.g., image-mpeg7-colourstructure).

Most relevant for marine science is the iFDO content section for annotations. These are semantic classifications of groups of pixels in the images assigned by humans or algorithms. By establishing iFDOs as the standard format for image annotations, exchange and reuse is facilitated of this derived data while in parallel also making annotation data FAIR. Annotations require a set of semantic labels in the image-annotation-labels field and a set of identifiers of humans or machines in the image-annotation-creators field. The image annotations are then encoded in the image-annotations field as a list of objects. These objects consist of a set of pixel coordinates, one or many label IDs, one or many annotator IDs and an optional confidence value. The format of the pixel coordinates is flexible such that point annotations, bounding boxes, polygons or whole image annotations can be stored.



By standardising the format and schema by which image data is characterised together with the semantic annotations of objects visible in images allows to implement operational machine learning systems for efficiently and effectively. As the iFDO format now provides a clearly structured schema as an interface between tools in the ML pipeline (e.g., image annotation tools such as BIIGLE<sup>43</sup> and deep learning frameworks such as YOLO<sup>44</sup>) data scientists can now focus on the key aspect of tuning AI parameters and optimising to a specific marine science use case (see Figure 47). The creation of iFDO files requires to include all relevant metadata that are collected during image acquisition and initial pre-processing for quality management. This may include information on equipment (cameras and platforms), mechanisms for sensor time synchronisation, pixel scale computation and more. All this information needs to be bundled by an iFDO factory mechanism such as the one provided by the MarIQ Python library (see Figure 48).

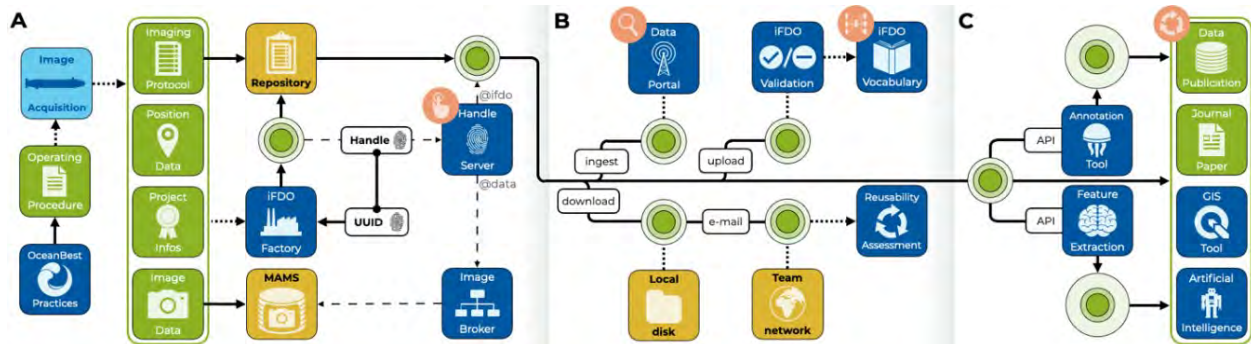


Figure 49. Creation and progression of an iFDO (green circles) and its derived versions. The left part shows how an iFDO uses persistent identifiers to reference itself within the FAIR infrastructure. The middle part shows how iFDO files can be discovered, shared, advertised, and validated. The right part shows how implementing iFDO-compliant APIs to marine science tools facilitates reuse of image data for arbitrary purposes.

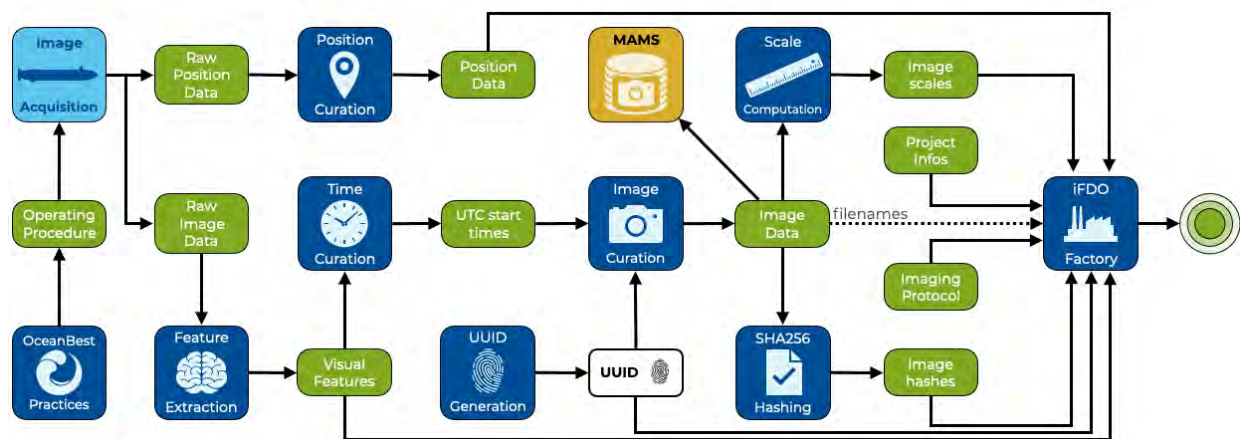


Figure 50. Creating an iFDO. Marine image acquisition is guided by OceanBestPractices<sup>45</sup> and creates raw image data and raw position data (for in-situ imaging). Multiple processing steps (blue boxes) create derived data products (green boxes) that are ultimately merged by an iFDO factory process to one iFDO file (green circle).

**Table 11.** Overview of FAIR recommendations by the RDA<sup>46</sup> and their implementation for marine images. Recommendations are ranked by the RDA as 'Essential' (\*), 'Important' (+) and 'Useful'.

| Recommendation | Description   | Implementation for marine images |
|----------------|---|----------------------------------|
| RDA-F1-01M*    | <i>Metadata are identified by a persistent identifier</i>                   | image-set-handle                 |
| RDA-F1-01D*    | <i>Data is identified by a persistent identifier</i>                        | Image-set-handle                 |
| RDA-F1-02M*    | <i>Metadata are identified by a globally unique identifier</i>              | Image-set-uuid                   |
| RDA-F1-02D*    | <i>Data are identified by a globally unique identifier</i>                  | image-set-uuid                   |
| RDA-F2-01M*    | <i>Rich metadata are provided to allow discovery</i>                        | iFDO format                      |
| RDA-F3-01M*    | <i>Metadata include the identifier for the data</i>                         | Image-set-uuid                   |
| RDA-F4-01M*    | <i>Metadata are offered in a way to be harvested and indexed</i>            | OAI-PMH catalogue                |
| RDA-A1-01M+    | <i>Metadata contain info to enable accessing the data</i>                   | Image-set-handle                 |
| RDA-A1-02M*    | <i>Metadata can be accessed manually</i>                                    | Image-set-handle                 |
| RDA-A1-02D*    | <i>Data can be accessed manually</i>  | Image-set-handle                 |
| RDA-A1-03M*    | <i>Metadata identifier resolves to a metadata record</i>                    | Image-set-handle                 |
| RDA-A1-03D*    | <i>Data identifier resolves to a digital object</i>                         | Image-set-handle                 |
| RDA-A1-04M*    | <i>Metadata are accessed through standardised protocol</i>                  | https                            |
| RDA-A1-04D*    | <i>Data are accessible through standardised protocol</i>                    | https, nfs, ...                  |
| RDA-A1-05D+    | <i>Data can be accessed automatically</i>                                   | Image Broker API                 |
| RDA-A1.1-01M*  | <i>Metadata are accessible through a free protocol</i>                      | https                            |
| RDA-A1.1-01D+  | <i>Data are accessible through a free access protocol</i>                   | https                            |
| RDA-A1.2-01D   | <i>Data are accessible with authentication and authorisation</i>            | https                            |
| RDA-A2-01M*    | <i>Metadata are guaranteed to remain available after data loss</i>          | handle system                    |
| RDA-I1-01M+    | <i>Metadata use knowledge representation in standard format</i>             | iFDO                             |
| RDA-I1-01D+    | <i>Data use knowledge representation in standard format</i>                 | e.g. jpg, png, mov               |
| RDA-I1-02M+    | <i>Metadata use machine-understandable knowledge representation</i>         | iFDO schema                      |
| RDA-I1-02D+    | <i>Data use machine-understandable knowledge representation</i>             | e.g. jpg, png, mov               |
| RDA-I2-01M+    | <i>Metadata use FAIR-compliant vocabularies</i>                             | iFDO schema                      |
| RDA-I2-01D     | <i>Data use FAIR-compliant vocabularies</i>                                 |                                  |
| RDA-I3-01M+    | <i>Metadata include references to other metadata</i>                        | e.g. image-creators              |
| RDA-I3-01D     | <i>Data include references to other data</i>                                | image-uuid                       |
| RDA-I3-02M     | <i>Metadata include references to other data</i>                            | image-related-material           |
| RDA-I3-02D     | <i>Data include qualified references to other data</i>                      | N/A for images                   |
| RDA-I3-04M     | <i>Metadata include qualified references to other data</i>                  | Image-related-material           |
| RDA-I3-03M+    | <i>Metadata include qualified references to other metadata</i>              | e.g. image-creators              |
| RDA-R1-01M*    | <i>Plurality of accurate and relevant attributes are provided for reuse</i> | iFDO fields                      |
| RDA-R1.1-01M*  | <i>Metadata include information about the reuse license</i>                 | Image-license                    |
| RDA-R1.1-02M+  | <i>Metadata refer to a standard reuse license</i>                           | Image-license                    |
| RDA-R1.1-03M+  | <i>Metadata refer to a machine-understandable reuse license</i>             | Image-license                    |
| RDA-R1.2-02M   | <i>Metadata include cross-community provenance info</i>                     | Image-provenance                 |
| RDA-R1.2-01M+  | <i>Metadata include community-specific provenance info</i>                  | Image-provenance                 |
| RDA-R1.3-01M*  | <i>Metadata comply with a community standard</i>                            | iFDO                             |
| RDA-R1.3-01D*  | <i>Data comply with a community standard</i>                                | e.g. jpg, png, mov               |
| RDA-R1.3-03M*  | <i>Metadata use a machine-understandable standard</i>                       | JSON                             |
| RDA-R1.3-02D   | <i>Data use a machine-understandable standard</i>                           | e.g. jpg, png, mov               |

## 6. Machine-learning for automated image analysis

Contributions from: Timm Schoening, Pedro Juan Soto Vega, Marjolaine Matabos, and the co-authors of the peer-reviewed papers that contributed to this section<sup>35,47–52</sup>.

### 6.1 Overview

Automated image analysis refers to the process of *extracting meaningful information from images*. It can involve a wide range of techniques and methods, depending on the type of image and the information being sought. The goal of automated image analysis is to extract meaningful information from images to better understand and interpret the visual data<sup>7</sup>.

Image analysis includes steps such as image pre-processing, feature extraction, sometimes learning, and data interpretation. Image pre-processing typically involves adjusting the image to correct for things like brightness, contrast, and noise – which is essential for many subsequent processing steps such as machine learning<sup>7</sup>. Feature extraction involves identifying and extracting specific information from the image, such as the shapes and sizes of objects, the texture of surfaces, and the distribution of colours. Modern machine learning applications automate this step in the first layers of deep neural networks<sup>53</sup>. By automating this task, human-induced subjectivity is reduced, at the cost of increased computational effort to also tune these parameters<sup>52</sup>. Modern environments provide pre-trained feature descriptors to the community which can be tuned to specific automated image analysis tasks<sup>54</sup>. Finally, data interpretation involves using statistical methods, image annotations, crowd-sourcing, or other techniques to analyse the extracted features and draw conclusions about the image, i.e., to add semantic meaning to the visual image content<sup>43,47</sup>. The order of steps is nowadays blurred or intertwined as often there are multiple processing and interpretation steps involved in an iterative and complex workflow for automated image analysis<sup>55</sup>.

### 6.2 Automated underwater image analysis

Automated underwater image analysis is a subset of automated image analysis that is specifically designed to address the *unique challenges* of analysing images captured in an underwater environment, such as low light levels, murky water, limited visibility and lens distortion<sup>56</sup>.

Some common applications of automated underwater image analysis include monitoring and conserving marine ecosystems, detecting and mapping underwater pollutants, and studying the distribution and behaviour of marine organisms<sup>7,57,58</sup>. However, it is important to note that the application of automated underwater image analysis requires expert knowledge to extract robust and meaningful information for decision support. For example, the *limited visibility and poor lighting* conditions in the underwater environment can make it difficult for computer algorithms to accurately identify and extract objects from images<sup>59</sup>. Additionally, the presence of suspended particles, bubbles, and other underwater debris can also complicate the analysis process<sup>60,61</sup>. Despite these challenges, advances in computer vision and machine learning are helping to overcome these limitations and making automated underwater image analysis an increasingly powerful tool for understanding and conserving the underwater environment.

### 6.3 Requirements

To do machine learning with images, you will typically need several resources and capacities. First, a *dataset of training images* is needed that are relevant to the machine learning task<sup>62,63</sup>. The images should be annotated with labels or other metadata that indicate what is contained in the images<sup>43,64</sup>. For object detection, the dataset should include images with bounding boxes around the objects of interest<sup>51</sup>. *Resolution of the images* and imaged objects is strongly impacting the machine learning method quality<sup>50</sup>. The big breakthroughs discussed in the media such as ImageNet were obtained with objects consisting of millions of pixels<sup>62</sup>. In underwater imaging, though, the objects of interest often consist of as little as a few hundred pixels, constituting a much greater algorithmic challenge.

Before a machine learning model can be trained, it may be necessary to pre-process the images to standardise their format and size, adjust their brightness and contrast, or perform other operations to improve their quality<sup>59</sup>. Normalisation is particularly important when dealing with underwater imagery and a multitude of methods has been proposed to solve this challenge<sup>59</sup>. Future deep learning tools might be able to overcome this challenge as they may be capable to overlook the visual challenges based on a large and diverse enough training set. One project that aims in that direction is FathomNet<sup>44,63</sup>.

To be able to operate machine learning methods it is then required to extract relevant features from the images. There are many traditional feature extraction techniques available, including colour histograms, texture analysis, edge detection, but recently deep learning has turned into the method of choice. Deep learning techniques such as convolutional neural networks (CNNs) have been very successful for image feature extraction as they learn the feature representation from the data without the need for tuning based on human expert knowledge<sup>44</sup>. Learning feature representations with deep learning is now the state-of-the-art in image-based machine learning<sup>47</sup>.

Once features have been extracted from the images, a machine learning algorithm can be trained into a model that can classify, detect, or segment objects in the images. Some traditional machine learning algorithms for image-based tasks include decision trees, support vector machines (SVMs), random forests, but nowadays neural networks are the dominant state-of-the-art for this step as well.

Training machine learning models with images can be computationally intensive, especially when tuning deep learning architectures. This may require a powerful computer with a fast and large Graphics Processing Unit (GPU) to train models efficiently<sup>52</sup>. Alternatively, cloud-based services such as Amazon Web Services (AWS) or Google Cloud Platform (GCP) provide pre-configured environments for training machine learning models.

## 6.4 Deep Learning

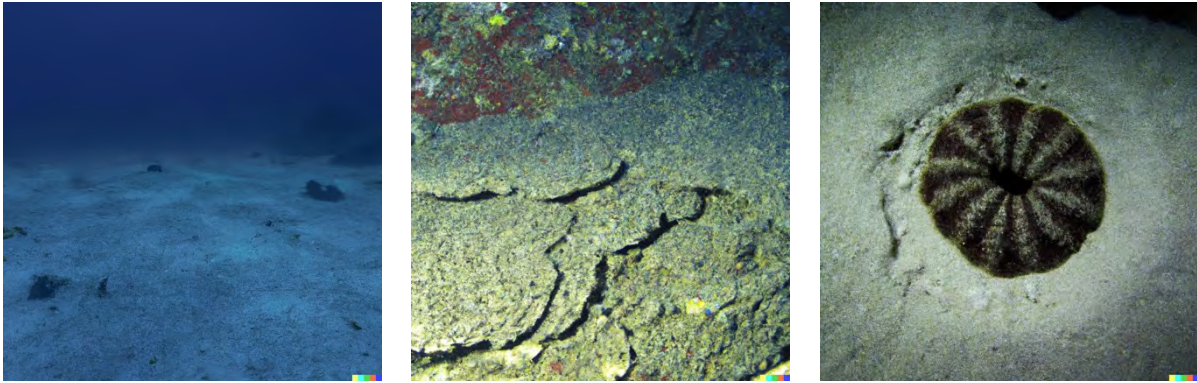
Deep learning, as a subfield of machine learning and the state-of-the-art in automated image analysis, has additional requirements to be implemented and used effectively. Firstly, large amounts of high-quality data are necessary. The data must be representative of the problem being solved and must be labelled accurately so that the algorithm can learn from it. Several hundreds of robust annotations per category are often stated as the bare minimum to allow the complex neural networks to adapt to the given task (as stated in the YOLOv5 documentation).

Choosing the appropriate deep learning architecture, such as a convolutional neural network or a recurrent neural network, is crucial for success. Expertise in deep learning architectures and frameworks is required, including a deep understanding of the underlying mathematics, algorithms, and programming frameworks. This is key to enable training robust models that are not just overfitting to the training data but are capable of generalising across data sets.

By building on successful deep learning frameworks and tools, such as TensorFlow, PyTorch, or Keras, the implementation of machine learning workflows can be done efficiently.

The diversity of methods and applications that fall under the umbrella term deep learning is growing and evolving rapidly. About two years back, e.g., the discussion about creative art works produced through so called Generative Adversarial Networks (GANs) showed that artificial intelligence (AI) is capable of further tasks that were thought to remain human-centred for a while<sup>65</sup>. Now even GANs are not state of the art anymore but tools like the currently hyped DALL-E are trained on diffusion-based methods<sup>66</sup>. These allow to create artificial images from text inputs. Variations of this concept have been developed by various AI research teams, e.g., DALL-E was created by OpenAI which is also behind ChatGPT. Others like Imagen by Google also use backwards diffusion whereas StabilityAI uses stable diffusion for their creative AI. All of these are trained on vast image data bases but are hence also restricted by the type of

training data that are available to them. [Figure 51](#) shows, that deep-sea imagery was clearly not part of the training data. Efforts like FathomNet and standardisation of image data for publication, e.g., using iFDOs can help in the future to develop generative deep learning models that can also comprehend how deep-sea images look like<sup>35,63</sup>.



**Figure 51.** AI-generated images from the input phrases: ‘a deep-sea seafloor image’ (left), ‘manganese nodule field in the deep sea’ (middle) and ‘a top down looking deep-sea seafloor image with a holothurian’ (right). All images were created by OpenAI’s DALL-E.

The stronghold and deep learning with images came from the development of ImageNet by Google and the subsequent rapid evolution of methods<sup>62</sup>. This led to powerful networks for object detection and object classification (i.e., identifying which pixels in an image are interesting and identifying to which semantic / human category these pixels belong). Back bone examples are VGG16, ResNet, DarkNet<sup>67,68</sup>. These were used and extended in different scenarios for examples in two-stage architectures that are comparatively slower yet attain higher accuracy such as R-CNN, Faster R-CNN or SPP-Net as well as one stage architectures that are very fast such as YOLO up to its current version YOLO v8<sup>44,69</sup>. Other examples are RetinaNet for pose estimation and instance segmentation or CenterNet<sup>70</sup>.

In other deep learning tasks, where a complete segmentation (i.e., every pixel is assigned to a known or unknown group rather than extracting an area of interest) of image is targeted a different family of methods has been developed. These include encoder-decoder architectures such as UNet, Seg-Net, LinkNet or TernausNetV2. More information is included by architectures that concatenate the learned features from the backbone networks such as DeepMask, SharpMask or Frnet. Other architectures use Spatial Pyramid Pooling or Dilated Convolutions and the list of specialties and application scenarios evolves and evolves and it is most certainly impossible to keep track of the rapidly evolving field. It is also often impossible to properly reference the developments as research progress is usually published through pre-print or as posts in company blogs and social media. This makes it nearly impossible to compare qualities of AI methods and to calibrate across tuned architectures for a specific use case. More worryingly, the current state of the art requires that an AI system must be tuned for each and every new application in case one of the acquisition parameters changes. This can include a change in illumination, change in camera orientation, change in resolution, etc.

## 6.5 Annotations

Robust annotations are the key ingredient to make AI work. The breakthrough discoveries in the AI community are based on well-annotated data, rather than the network architectures or compute power<sup>62</sup>. The number of annotations needed to train a state-of-the-art object detection network such as YOLOv5, depends on several factors, including the complexity of the task, the quality and diversity of the data, and the size of the network<sup>44</sup>. As a general rule, the more complex the task, the more annotations are needed. For example, training a model to detect objects in cluttered scenes with many overlapping objects will require more annotations than training a model to detect objects in simple, clear scenes.

The quality and diversity of the data also play a role in the number of annotations needed. If the data are high-quality and diverse, with examples of the objects of interest from a variety of viewpoints, scales, and orientations, then fewer annotations may be needed. The size of the network can also impact the number of annotations needed. Larger networks typically require more annotations to achieve good performance, while smaller networks may be able to achieve good performance with fewer annotations. In general, it is difficult to give a specific number of annotations needed to train YOLOv5 or any other object detection model, as this will vary depending on the specific task and data<sup>44</sup>. However, it is common to use thousands or tens of thousands of annotations for object detection tasks, and the number may be even higher for complex or challenging tasks.

To make AI widely applicable to underwater image data hence requires a step-change in sharing and publishing image annotations. The community needs open-sourced annotations that are provided in a standardised format for interoperability between annotation tools, AI training methods and for cross-calibrating tuned models<sup>43</sup>.

## 6.6 AI developments focussed on marine images

### Approaches to standardise marine image analysis

A growing number of initiatives is currently aiming in this direction. The SMarTar-ID scheme can provide a catalogue of operational categories for annotation labels that is based on standardised fauna data bases and metadata formats such as the World Register of Marine Species, OBIS and DarwinCore<sup>71,72</sup>.

The iFDO concept (see above) developed through iAtlantic provides a metadata schema and file format for the image data that needs to be openly shared for AI, as well as for the image annotations that need to be open-sourced<sup>35</sup>. This format is currently being picked up by several interested stakeholders and implemented into community tools such as BIIGLE<sup>43</sup>.

The FathomNet initiative likely provides the unifying component to assemble the federated data sets that are hosted nationally or institutionally into one data base for underwater images<sup>63</sup>. The standardisation efforts such as SmarTarID or iFDOs can provide one efficient mechanism to widen the data contributions to FathomNet.

### How to acquire images

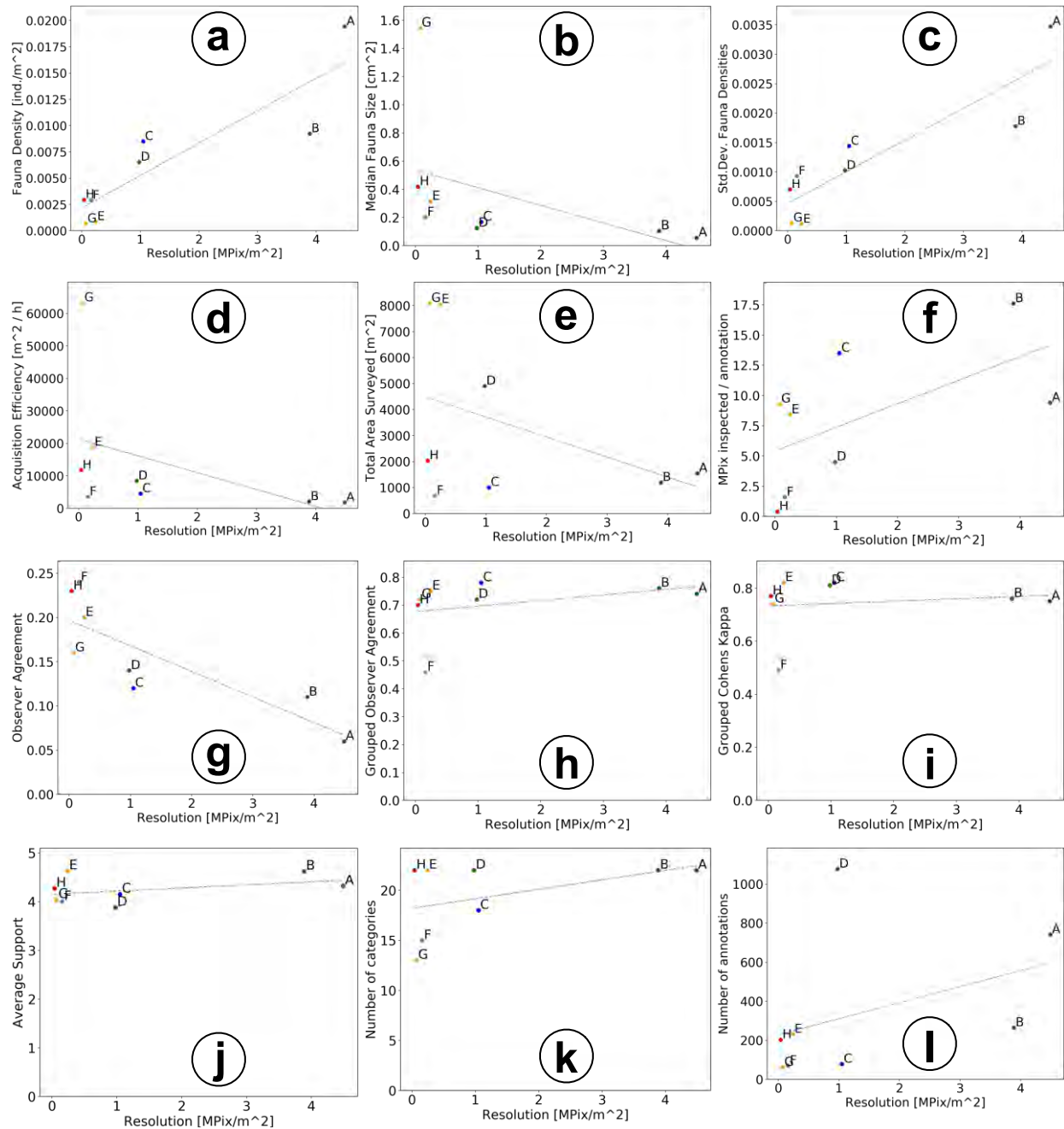
The platforms and sensors used to acquire underwater image data are highly customisable. With cameras, illumination sources and deployment protocols changing rapidly, even during one survey cruise, image data of differing quality is creating that may serve different analysis needs<sup>7</sup>. The extraction of information from heterogeneous imagery requires particular attention. This affects automated approaches using AI, but similar attention is required during manual image annotation for non-automatic analysis or the generation of training data for machine learning applications<sup>50,51</sup>.

The area of seafloor which may be imaged by an optical platform is determined by the lens parameters used in the camera system, distance and orientation to the seafloor, sensitivity of the system to motion and illumination, and a range of other factors<sup>73</sup>. Larger areas of the seafloor can be imaged with wide-angle or 'fisheye' camera systems<sup>56</sup>, though there is an associated vignetting effect rendering the details collected from the extremities of an image less rich than areas of seafloor more directly located below the lens center<sup>74</sup>. The raw images collected by those camera systems can appear quite distorted, and manual labelling of fauna within these images is more difficult towards the edges of each image. Digital post-processing of these distorted images can be reasonably straightforward when the arrangement of optics for an imaging platform is known, and for larger fauna these processed images can be suitable for subsequent analysis<sup>75,76</sup>. However, image processing cannot create 'newly improved' data, and therefore there will always be a loss of information at the image boundaries after lens correction. Image analysis – manual as well as automated – could therefore focus on central parts of the image, and the boundary

area of images could be used to display, for example, navigation metadata. Lenses of a more ‘telephoto’ or narrower angle will allow collection of less distorted images, though these collected images will capture a significantly smaller area of seafloor than may be achieved with wider-angle systems.

The deep sea is a dark environment with no sunlight penetration. It is therefore essential that camera systems are supplemented by artificial illumination. To provide sufficient illumination for video and still-camera systems, abundant power reserves must either be mounted on the platform or delivered via a cable from the support vessel. Positioning of the lights on an imaging platform can be difficult, and optimising the spread of light, i.e., maintaining an equal light balance across the imaged area, can be challenging. Illumination vignetting can be partially addressed prior to analysis by excluding the image edges from analysis <sup>74</sup>. Additionally, when the lights and camera are mounted close to each other, a significant amount of light might be scattered by the water column into the camera, leading to a degraded ‘foggy’ image, which is an issue for small platforms and/or high-altitude photography. Finally, the colour spectrum of the light also needs to be considered, as for instance the returned yellow, orange, and red components of the signal may be too weak to support taxonomic identification, depending on the type of light source. This spectrum change can affect human interpretation, reduce discriminability by automated systems due to a reduced contrast range available and is particularly relevant in UHI applications. Any illumination system needs to be set up to accommodate the target altitude of the camera platform above the seafloor as well as the expected altitude variation.

Image resolution is derived from a combination of the camera optics and the deployment altitude and allows comparing image datasets numerically. The camera optics determine the pixel resolution (usually in the tens of megapixels for state-of-the-art camera systems). The field of view of the camera objective lens and the deployment altitude determine the image footprint, i.e., the area in square meters that is covered by a single image acquisition. These two values can be combined to a measure of megapixels per square meter ( $\text{MP m}^{-2}$ ) – or the numerically identical value given in pixels per square millimetre ( $\text{P mm}^{-2}$ ) – to e.g., analyse annotator performance or fauna density estimates consistently.



**Figure 52.** The effect of image resolution (given in megapixels per square meter images) on the horizontal axis on various aspects of underwater image analysis. Key metrics reported in scientific papers such as Fauna Density (a) and observed Fauna Size (b) are correlated with resolution. Effort in imaging the seafloor increases with resolution, for the effort in inspecting the images for annotation this increase is disproportionately large. Robustness of annotations (e.g., measured in observer agreement (g), Cohens kappa (i) or uncertainty in annotations (c) also increase with resolution). Figure from Schoening et al. 2020<sup>3</sup>

A study conducted within *iAtlantic* explored the effect of image resolution on manual fauna annotation. The study looked at the effect of using various camera systems of different resolution at different altitudes over ground. A group of trained annotators marked megafauna in the deep seafloor images which allowed to assess derived annotation statistics as well as to compare annotators' performances. It showed an obvious increase in observed fauna density with imaging resolution (see Figure 52a). This trend was mirrored in the observation that the median size of the annotated fauna decreases with increasing resolution (see Figure 52b). Together, it was reasoned that the increased resolution allows annotating smaller objects, increasing the total amount of individuals annotated. Nevertheless, it is also

<sup>3</sup> <https://bg.copernicus.org/articles/17/3115/2020/>



obvious that the increased resolution comes with an increase in observer disagreement (see [Figure 52c](#)) shown by an increase in the standard deviation of fauna densities estimated the five annotation experts. This means that trained annotation experts tend to disagree more on image data of higher resolution. [Figure 52\(d–f\)](#) highlights the trade-off between image resolution and effort involved in inspecting the seafloor images. In [Figure 52d](#) the increase in resolution comes with a decrease in acquisition efficiency in terms of the area per hour ( $\text{m}^2 \text{h}^{-1}$ ) that can be imaged. [Figure 52f](#) shows, that although higher densities of fauna are detected for high-resolution datasets, it still requires manually inspecting more megapixels per annotation compared to lower-resolution datasets. The annotation effort for such high-resolution datasets is thus disproportionately large.

From the abovementioned observations result some general guidelines on the collection of good-quality image data.

- a) A well-documented camera system should be used: aperture, sensitivity, lens arrangement and mounting angle.
- b) For a given study location, a comparable survey deployment plan should be used at each time step of analysis: the same sensor payload, instrument platform altitude, deployment speed, seafloor area imaged and sample unit size.
- c) Illumination should be maintained across deployments: intensity, wavelength and mounting angle.
- d) The lowest feasible altitude above seabed using a given platform will always provide higher-resolution data and higher taxonomical resolution in the faunal identification.
- e) Annotations by several observers need to be collected and thoroughly merged to create robust data for interpretation.

Guideline a) was thoroughly supported by the work conducted in iAtlantic on making marine image data FAIR with iFDOs<sup>35</sup>.

Guidelines b) – c) are recommendations that can potentially be overcome by developments in AI methodology that can circumvent aspects of heterogeneous image data. It was aspired to also address these during *iAtlantic* – unfortunately the global pandemic did not allow to acquire the necessary baseline data for this far-reaching goal.

Guideline e) was a necessity for creating robust annotations: for manual scientific interpretation, for assimilating robust data into disruptive AI training data sets as well as for creating training data for AI applications. Despite the cost involved with this guideline, following it can lead to better publications and better automation of the image analysis challenge.

### How to annotate

In any image annotation task, the results should be of a quality to allow including them in a scientific study or in a decision process<sup>50</sup>. In marine image annotation, the concept for a high or low annotation quality is often not clearly defined<sup>51</sup>. This section describes quality measures for image annotation as well as ways to create reference data sets for comparison, also referred to as gold standards and is based on a previously published paper<sup>51</sup>.

In contrast to other scientific imaging domains, such as digital pathology, there are no protocols and reports available that guide users (often referred to as observers) in the non-trivial process of assigning semantic categories to whole images, regions of interest (ROI) in images, or objects of interest (OOI). These protocols are crucial to facilitate image analysis as a robust scientific method, particularly in case of annotations to be used as a basis for training machine learning methods.

To assess annotation quality, a test set of images with a gold standard annotation result is required. The test set must represent the diversity of input signals and structures monitored, which means that each category considered in the study shall be represented with an appropriate number of examples. The gold

standard must be accepted by the scientific community to be an acceptable approximation of the ground truth. To acquire a gold standard for a test set of images three different strategies have been applied in the past: i) basing the gold standard on the annotations of one renowned expert; ii) creating a consensus expert that represents the joined expertise of a group of trained annotators; iii) ground truth data where e.g., sampling allows to robustly identify taxa visible in images.

Once a gold standard has been created, annotations can be evaluated against this gold standard. Many studies report different quality metrics and mention terms like accuracy or precision to denote the quality of an annotation process, sometimes even using these metrics in a wrong way or without a concise definition. In general, two overarching quality metrics need to be considered for annotation. First, an accuracy metric that relates an annotation set to a gold standard, and second, a reproducibility metric that assesses how robust an annotation set is, i.e., how well it can be reproduced, e.g., by the same annotator.

Accuracy thereby describes the ability of an observer to make decisions or classifications that can be considered as correct, condition to an available gold standard representing the state-of-art knowledge. To assess the accuracy of an annotation, a test set of image data must be provided together with a gold standard and the accuracy is usually described by statistics such as Precision, Recall, Positive Predictive Value, etc.

The ability to reproduce results from an experiment is one cornerstone in good scientific practice and must be considered in the development and application of new scientific methods. In the context of marine image annotation this term describes the ability of one observer to either reproduce her/his results or the results of a second observer for the same data set. The reproducibility should be discussed separately from the accuracy since the significance of an accuracy assessment depends on the availability of an accepted gold standard. But even if such a gold standard is missing, the ability of one observer to reproduce her/his results or the results of a second observer must be shown to demonstrate the potential of an imaging-based approach to produce significant results. In fact, if the classification task is very complicated and difficult to perform by the observers due to the image quality, a reproducibility analysis should be carried out before collecting gold standard annotations or should be an integrated part of this procedure.

### **In-situ computation capacity**

Marine imaging easily creates Terabyte-sized datasets for single camera deployments. Image processing both by humans and machines requires substantial compute capacity to extract information from such big data sets. Algorithms employed to automate the analysis commonly rely on large-scale compute infrastructure. So far, such an infrastructure has only been available on-shore. There is a clear trend towards bringing analysis capacity out to sea. For image data specialised GPU (Graphics Processing Unit) compute capacity is required. Within *iAtlantic*, the Sea-going High-Performance Compute Cluster (SHiPCC) units were completed, tested and deployed during months-long expeditions for image analysis<sup>52</sup>. They are mobile, robustly designed to operate with electrically impure ship-based power supplies and based on off-the-shelf computer hardware. Each unit comprises of up to eight compute nodes with graphics processing units for efficient image analysis and an internal storage to manage the big image data sets. Successful deployment of a SHiPCC unit at sea has allowed to extract semantic and quantitative information from a Terabyte-sized image data set within 1.5 h, representing a relative speedup of 97% compared to a single four-core CPU computer. Enabling such compute capability out at sea allows to include image-derived information into the cruise research plan, for example by determining promising sampling locations. The SHiPCC units are envisioned to generally improve the relevance and importance of marine imagery for marine science.

At present, this development is advanced further to bring the compute capacity – and thus intelligence – right into the diving camera system itself. With embedded GPU compute components such as the NVIDIA Jetson boards, image processing capacity can be deployed in-situ for edge computing scenarios. This innovation development was directly inspired by the results achieved during *iAtlantic* and the implementation of the standardised image processing methods for the AI applications. The goal is to bring automated image processing, image quality assurance and automated image information extraction (e.g., object detection and classification) to depths, enabling to adjust deployment missions live as well as to inform robotic platforms with an image-based backseat-driver capability.

### Examples of current machine learning for underwater imaging

GEOMAR developed an automated and fully-integrated seafloor classification workflow in the context of *iAtlantic*, aimed at classifying the seafloor into habitat categories based on an automated analysis of optical underwater images with a minimal number of human annotations<sup>49</sup>. The workflow employs automated laser point detection for scale determination and colour normalisation as well as a semi-automatic generation of a training data set for tuning a seafloor classifier. The global COVID-19 pandemic and technical gear failure of AUV Autosub6000 during the iMirabilis expedition led to the unavailability of a well-annotated image dataset from iAtlantic target areas, acquired by heterogeneous acquisition devices (camera platforms and cameras). The workflow was hence applied to an image dataset from the Pacific Ocean.

The proposed workflow operates on optical RGB still images recorded by any moving platform such as an ROV, AUV or towed camera system. The workflow starts with automatically detecting laser points and using these to infer the photo scale required for colour normalisation. This normalisation is implemented by matching the histogram of all other images to the respective reference image. The images are further rescaled to a defined resolution, and central parts of the image are cropped out to create data with equal spatial footprint in square meters. A labelled training set is then generated semi-automatically, and used to tune a convolutional neural network (CNN) classification model using the Inception V3 architecture<sup>77</sup>. Further, an empirical comparison of four sampling strategies is done to identify the optimal strategy to generate training data for tuning an unsupervised seafloor classification model.

Although the workflow was tuned for the Pacific data set, the method is generally applicable to characterise other seafloor settings, given appropriate training data. This workflow will be modularised, and its components standardised according to the FAIR machine learning infrastructure described below.

IFREMER implemented Deep Learning (DL) methods in the context of *iAtlantic* to assist in recognition procedures to produce local high-resolution maps of vent and non-vent habitats and communities at the Lucky Strike vent field along the mid-Atlantic Ridge. The DL was used due to its ability to learn and extract useful information from the images automatically. The development was carried out with 2D images to identify fauna, sea-bed substratum, and anthropogenic debris. Multiple DL architectures were evaluated in the substratum characterisation task with a focus on Convolutional Neural Networks (CNNs) such as VGG, ResNet and Xception. The model training was conducted in a multi-class and multi-label image classification scheme, generating models for the recognition of the substratum in all characterisation criteria. Following model training, an ensemble of classifiers was formed, optimising the robustness of the entire classification system. Results show the suitability of such techniques for correctly categorising the substratum, a fundamental driver of species distribution. A novel procedure based on uncertainty measures has been introduced in the substratum characterisation to reduce human annotators' effort and increase confidence in the DL-based model predictions. The workflow helps in defining which images need to be audited after the models have predicted them in order to optimise image classification accuracy.

Another innovative contribution to automatically interpret underwater images with AI methods includes the development to characterise and visualise multi-hectare areas of seafloor by means of AUV imaging, automated image correction, and algorithms such as autoencoders that learn from visual and spatial information<sup>78</sup>. This approach has shown robustness and applicability in multiple use case scenarios such as monitoring marine protected areas or detecting anthropogenic structures and biologically active sections around chemosynthetic habitats.

Some AI methods have already found their way into selected community best-practice tools such as BIIGLE which implements methods to rapidly create and validate annotations provided by humans or AI-based methods such as MAIA which supports automated classification<sup>47</sup>. In order to detect regions and objects of interest, other methods are implemented within BIIGLE such as the Unsupervised Knowledge Transfer method that can transfer knowledge learned from one image data set and its annotations to another dataset with comparable, yet different visual characteristics<sup>48</sup>.

## 6.7 Recommendations towards a FAIR machine learning environment

Developments in general machine learning methods for image analysis have been rapid and focussed on the effectiveness of the methods and the efficiency of their execution rather than efficiency of their use or re-use. Applications of those ML methods to specific marine image analysis by tuning to use case data sets has also rapidly increased in number, yet, a generalisation or standardisation of such adaptations is still lacking robustness towards data changes, development sustainability, reusability, and scientific reproducibility, i.e., FAIR-ness of machine learning for marine image analysis is lacking. Most important in reaching robustness and reproducibility is standardisation of components of the machine learning framework and open and FAIR publication of data items used in the data flow across the framework (e.g., image data, annotation data, processing instructions how to turn images and annotations into ML models). While for some components of a FAIR ML environment (see [Figure 53](#)) standardisation exist (e.g., DarwinCore metadata standard for fauna observation metadata required for annotation catalogues) others are still under development and yet to be adopted by the marine imaging community. The iFDO format, for example, is currently not a standard, as its widespread adoption by the marine science community is still needed. Once iFDOs becomes a standard for image metadata exchange, they will provide a standardised mechanism for the annotation data required for training a FAIR machine learning system.

It is yet unclear which technology will become a common solution for developing ML training systems. A potential candidate is Docker that has become a de-facto standard in software development, DevOps, and for the orchestration of even complex IT infrastructure. It represents a powerful mechanism by being maintained internationally, being supported by many software tools and by representing a means to share procedural information in a compact format that allows for efficient sharing, documentation for reproduction, adaptation for re-use, etc.

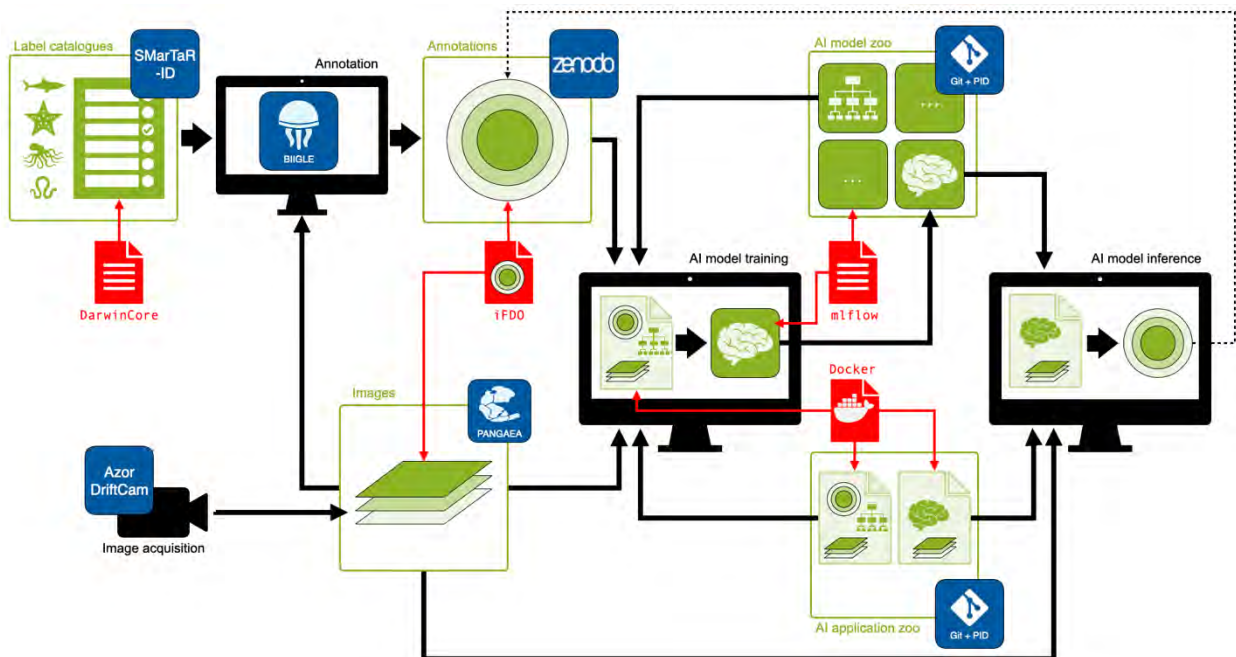
The second generalisation that could support a more widespread use of ML systems is the adoption of a standardised format for trained ML systems. A prominent solution that is yet underused in marine science is mlflow<sup>4</sup>. This platform is being developed with a comprehensive industry backing. It allows users to describe, save and load trained ML systems for distribution and re-use. Building an ML model for a marine imaging use case with this platform would thus give the opportunity to publish a re-usable product alongside a journal publication with its own data interpretation. Given the rapid evolution of foundational machine learning methods (see section Deep Learning) it cannot be expected that all these innovative products follow the mlflow paradigms. It is thus even more important to share adaptations of those foundational models for marine imaging use cases through the Docker mechanism.

---

<sup>4</sup> <https://mlflow.org>

To add operation ability to the standardised open and FAIR data components in the FAIR ML framework requires several software tools. Potential candidates for acquiring images are as diverse as the marine imaging community. One example would be the Azor drift-cam described above<sup>1</sup>. Storing the image data (i.e., photos and videos) requires substantial disk capacity, available through large, centralised data repositories such as Pangaea or Zenodo or through federating FAIR institutional media asset management systems. Creation of annotation data for training requires annotation software such as BIIGLE or Squidle+. The created annotation data needs to be made publicly available through data repositories such as Zenodo as well. Once these requirements are met, FAIR ML applications can be built. These can be tuned and operated on various computer hardware such as desktop computers, mobile GPU clusters (see SHiPCC section above) or on large-scale HPC (High-performance computing) environments in the cloud. Importantly the Docker instructions to build the models have to be published in repositories as well. As they represent code that is subject to evolving change it might be most efficient and effective to keep these instructions in code repositories such as Git (e.g., Gitlab). To make them sustainable and available to future FAIR use it would be required to publish them with a persistent identifier assigned (e.g., a DOI for a static Docker file publication, or a Handle URL pointing towards the current head of a code repository). The same consideration applies to the trained AI models that would likely be published through Git- and DOI-based AI model zoos. A simpler solution would be to publish Docker files and/or trained models in data repositories such as Zenodo as well.

Only if all data items and processing steps are available FAIR and open in a reproducible fashion will the FAIR ML environment become reality. The benefits will lie in easier adoptability to new data sets and use cases; an efficient sharing of results and trained products and a general increase in robustness of methods and the interpretation results needed for researching the oceans and informing important decisions.



**Figure 53.** Sketch of the proposed FAIR machine learning framework for marine image analysis. Green entities represent data items, e.g., images, annotations, or operation procedures. Blue squares represent examples of existing infrastructures that can facilitate operationalising parts of the framework. Red items denote standardisation formalism necessary to advance the TRL of machine learning for marine image analysis beyond demonstrator cases. Currently lacking in the community are published AI models and published operation procedures to reproduce scientific studies.

## 7. Future Perspectives

Marine imaging will remain a key technology in exploring and monitoring the ocean and human activities within. Technological evolutions like the developments described in this report will contribute to the growing importance of imaging in creating ocean narratives and in the widespread application of marine imaging to answer scientific questions and to address scenarios in decision making processes. Developments, such as those described in this report will emerge around the world and contribute to building and enhancing human and technological capacities around marine imaging.

Obvious future advances in imaging can be expected from ongoing technological developments. ChatGPT AI system was asked to predict the advances of marine imaging and it answered with the following main areas — high resolution imaging, 3D underwater imaging, the availability of autonomous vehicles and advances in AI and augmented reality using marine images. In addition to these AI predictions, we can expect technological advances such as low-light acquisition capacity, long-range/long-term camera platforms and integration of components of marine imaging into seamless workflows.

Bridging gaps between solitary technology developments is one key challenge for future research and implementation. Making marine image data FAIR needs cameras, deployments and image workflows that build upon the added value of FAIR-ness. Low-cost camera systems need efficient interpretation workflows to cope with the increasing data amounts created by simplified data acquisition. Hyperspectral imaging systems need 3D reconstruction for robust scene understanding. And bridging the gaps towards non-imaging developments is essential for integrating image data into information transfer processes as well. Predictive habitat mapping, e.g., requires ground-truth observation data, often created from marine images. With more habitat information available from calibrated hyperspectral data and from widespread observation capacity by reduced camera cost, habitat maps can be built with more data and more informative data, ideally leading to increased robustness. With FAIR and thus transparent and provenance-tracked data and data processing, habitat maps will gain more trustworthiness as interpretation and processing steps and decisions in creating ground-truth data become reproducible. Of course, the machine learning for habitat mapping itself should also become open and FAIR, comparable to the proposed FAIR machine learning infrastructure. By replacing image-based models in the infrastructure scheme with habitat mapping models, the same concept becomes applicable to the habitat mapping use case as well.

With image-based observation data being available to the research community, with machine learning models becoming operational in many marine science domains and with standardised and interactive visualisation tools entering the research process, all components to build Digital Twins for image-based workflows are now available. Operational Digital Twins that allow stakeholder engagement and decision-making support through exploring ‘What-if’-scenarios around image data are still missing, though. But demonstrators and larger visions can be expected within the next few years. Once technology developments, such as those described in this report, become inherently compatible with Digital Twin architectures, the uptake of image-derived information into scientific studies or monitoring can be substantially increased in volume and speed. Of course, this requires further investment in researching, developing, and implementing standards of data formats and interfaces between tools across image data workflows. By doing so, marine image data can contribute to Digital Twins in many ways. It will provide very high-resolution seafloor mapping (i.e., at sub-millimetre resolution) also in extreme environments for localised habitat predictions, e.g., around vents as demonstrated within *iAtlantic* WP2. It can enable marine organism tracking by turning traditional camera platforms such as AUVs and ROVs into smart monitoring robots that use camera data and automatically extracted information to implement a ‘backseat-driver’ functionality that keeps objects of interest within the field of view of the camera. This requires a Digital Twin of the vehicle platform that models the robotic system state and uses image-

derived observations to predict the required control commands to precisely manoeuvre the camera platform and illumination source. It further requires implementing edge computing capacity on camera platforms, i.e., bringing GPU computers down to the deep ocean.

The Digital Twin concept also holds promise to open research processes, solving the demands of funders to communicate and transfer knowledge while simultaneously providing novel research capabilities to researchers. This appears to be a win-win-scenario within which researchers themselves become drivers of the cultural change towards digital science.

The Digital Twin concept allows efficiently involving a wider group of stakeholders into information extraction and decision-making processes. Other aspects of marine imaging also require a more efficient and effective incorporation of human expertise, particularly in this time of a growing AI footprint in marine imaging. While it is simple to run a pre-tuned AI system and generate and publish impressive training results it remains yet to be seen whether a leap in generalisation capability can be achieved with marine images and AI. The same is true for habitat mapping. Solving this issue would require the AI technology for image interpretation to be made widely applicable across data sets and ocean basins. Key in that direction are large and robust training data sets that rely on quality-controlled annotations. To efficiently and effectively collect these, it is key that annotation protocols encourage the use of methods for automated assistance and the adherence of annotation best practices wherever possible. There is still room for improvement within the marine image analysis community to adopt modern image annotation tools and procedures.

To further align and standardise ocean observing across geographical regions, marine science domains, research institutes and marine sectors, marine imaging needs to expand on concepts such as the FAIR Digital Objects for images. Standardisation of protocols, interfaces and formats is essential to increase the efficiency and effectiveness of marine image analysis. Further research and implementation efforts are needed to turn these concepts into operable reality for researchers and the public.

Imaging will further contribute to map deep and open-ocean ecosystems at local, regional and global scale. With more cameras available, simplified deployment strategies and citizen science applications entering deeper waters, the footprint of marine imaging will rise across space, time, resolution, and wavelength. A particular focus for future mapping applications is the open pelagic ocean that promises to hold unexplored aspects of oceanic life and processes. With dark imaging and low-noise and low-electromagnetic emission platforms, pioneering research can be expected.

To fully understand oceanic processes across compartments and domains, such new insights are required and can partly be provided by marine imaging. By linking research domains and methodologies, such as in habitat mapping for the 3D pelagic environment, society can advance its capacity to assess the stability, vulnerability and tipping points of ecosystems.

Alongside these research and curiosity-driven developments and the formalisation and standardisation will come a transfer of methods, gear, and capacities for a sustainable blue economy. Digital developments that make marine image data FAIR and open already contribute to international platforms such as EMODnet for data sharing, access, and interoperability. With ongoing efforts to build a digital marketplace within GAIA-X, these FAIR and open contributions of image data – as well as image analysis methods – can also create a financially relevant value chain, for researchers and research institutes, to sustainably support future mapping, observing, monitoring, and automating projects.

## 8. References

1. Dominguez-Carrió, C., Fontes, J. & Morato, T. A cost-effective video system for a rapid appraisal of deep-sea benthic habitats: The Azor drift-cam. *Methods Ecol Evol* **12**, 1379–1388 (2021). <https://doi.org/10.1111/2041-210X.13617>
2. Ramirez-Llodra, E. *et al.* Deep, diverse and definitely different: Unique attributes of the world's largest ecosystem. *Biogeosciences* **7**, 2851–2899 (2010). <https://doi.org/10.5194/bg-7-2851-2010>
3. Bowden, D. A. & Jones, D. O. B. Towed Cameras. in *Biological Sampling in the Deep Sea* 260–284 (John Wiley & Sons, Ltd, 2016). doi:10.1002/9781118332535.ch12. <https://doi.org/10.1002/9781118332535.ch12>
4. Humphris, S. E. & Soule, A. Vehicles for Deep-Sea Exploration. in *Encyclopedia of Ocean Sciences* 21–30 (Elsevier, 2019). doi:10.1016/B978-0-12-409548-9.11213-8. <https://doi.org/10.1016/B978-0-12-409548-9.11213-8>
5. Huvenne, V. A. I. *et al.* ROVs and AUVs. in 93–108 (2018). doi:10.1007/978-3-319-57852-1\_7. [https://doi.org/10.1007/978-3-319-57852-1\\_7](https://doi.org/10.1007/978-3-319-57852-1_7)
6. Kelley, C., Kerby, T., Sarradin, P.-M., Sarrazin, J. & Lindsay, D. J. Submersibles and Remotely Operated Vehicles. in *Biological Sampling in the Deep Sea* 285–305 (John Wiley & Sons, Ltd, 2016). doi:10.1002/9781118332535.ch13. <https://doi.org/10.1002/9781118332535.ch13>
7. Durden, J. M. *et al.* Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. *Oceanography and Marine Biology: An Annual Review* **54**, 1–72 (2016). <https://doi.org/10.1201/9781315368597>
8. Teague, J., Allen, M. J. & Scott, T. B. The potential of low-cost ROV for use in deep-sea mineral, ore prospecting and monitoring. *Ocean Engineering* **147**, 333–339 (2018). <https://doi.org/10.1016/j.oceaneng.2017.10.046>
9. Thorsnes, T., Misund, O. A. & Smelror, M. Seabed mapping in Norwegian waters: programmes, technologies and future advances. *Geological Society, London, Special Publications* **499**, 99–118 (2020). <https://doi.org/10.1144/SP499-2019-69>
10. Roberts, J. M. *et al.* A blueprint for integrating scientific approaches and international communities to assess basin-wide ocean ecosystem status. *Commun Earth Environ* **4**, 12 (2023). <https://doi.org/10.1038/s43247-022-00645-w>
11. Giddens, J., Turchik, A., Goodell, W., Rodriguez, M. & Delaney, D. The National Geographic Society Deep-Sea Camera System: A Low-Cost Remote Video Survey Instrument to Advance Biodiversity Observation in the Deep Ocean. *Front Mar Sci* **7**, (2021). <https://doi.org/10.3389/fmars.2020.601411>
12. Phillips, B. T. *et al.* DEEPi: A miniaturized, robust, and economical camera and computer system for deep-sea exploration. *Deep Sea Research Part I: Oceanographic Research Papers* **153**, 103136 (2019). <https://doi.org/10.1016/j.dsr.2019.103136>
13. Jiang, H., Hu, Y., Yang, H., Wang, Y. & Ye, S. A Highly Sensitive Deep-Sea *In-Situ* Turbidity Sensor With Spectrum Optimization Modulation-Demodulation Method. *IEEE Sens J* **20**, 6441–6449 (2020). <https://doi.org/10.1109/JSEN.2020.2977348>
14. Hannah, R. W. & Blume, M. T. O. Tests of an experimental unbaited video lander as a marine fish survey tool for high-relief deepwater rocky reefs. *J Exp Mar Biol Ecol* **430–431**, 1–9 (2012). <https://doi.org/10.1016/j.jembe.2012.06.021>
15. Lund-Hansen, L. C. *et al.* A low-cost remotely operated vehicle (ROV) with an optical positioning system for under-ice measurements and sampling. *Cold Reg Sci Technol* **151**, 148–155 (2018). <https://doi.org/10.1016/j.coldregions.2018.03.017>
16. Sheehan, E. V. *et al.* An experimental comparison of three towed underwater video systems using species metrics, benthic impact and performance. *Methods Ecol Evol* **7**, 843–852 (2016). <https://doi.org/10.1111/2041-210X.12540>
17. Levin, L. A. *et al.* Global observing needs in the deep ocean. *Front Mar Sci* **6**, 241 (2019). <https://doi.org/10.3389/fmars.2019.00241>
18. Danovaro, R. *et al.* Ecological variables for developing a global deep-ocean monitoring and conservation strategy. *Nat Ecol Evol* **4**, 181–192 (2020). <https://doi.org/10.1038/s41559-019-1091-z>



19. Stuart, McGonigle & Willmott. Hyperspectral Imaging in Environmental Monitoring: A Review of Recent Developments and Technological Advances in Compact Field Deployable Systems. *Sensors* **19**, 3071 (2019). <https://doi.org/10.3390/s19143071>
20. Freitas, S., Silva, H., Almeida, J. & Silva, E. Hyperspectral Imaging for Real-Time Unmanned Aerial Vehicle Maritime Target Detection. *J Intell Robot Syst* **90**, 551–570 (2018). <https://doi.org/10.1007/s10846-017-0689-0>
21. Shen, Y., Zhao, C., Liu, Y., Wang, S. & Huang, F. Underwater Optical Imaging: Key Technologies and Applications Review. *IEEE Access* **9**, 85500–85514 (2021). <https://doi.org/10.1109/ACCESS.2021.3086820>
22. Liu, B. *et al.* Underwater Hyperspectral Imaging Technology and Its Applications for Detecting and Mapping the Seafloor: A Review. *Sensors* **20**, 4962 (2020). <https://doi.org/10.3390/s20174962>
23. Tegdan, J. *et al.* Underwater hyperspectral imaging for environmental mapping and monitoring of seabed habitats. in *OCEANS 2015 - Genova* 1–6 (IEEE, 2015). doi:10.1109/OCEANS-Genova.2015.7271703. <https://doi.org/10.1109/OCEANS-Genova.2015.7271703>
24. Sture, O., Ludvigsen, M., Soreide, F. & Aas, L. M. S. Autonomous underwater vehicles as a platform for underwater hyperspectral imaging. in *OCEANS 2017 - Aberdeen* 1–8 (IEEE, 2017). doi:10.1109/OCEANSE.2017.8084995. <https://doi.org/10.1109/OCEANSE.2017.8084995>
25. Johnsen, G. *et al.* Underwater hyperspectral imagery to create biogeochemical maps of seafloor properties. in *Subsea Optics and Imaging* 508–540e (Elsevier, 2013). doi:10.1533/9780857093523.3.508. <https://doi.org/10.1533/9780857093523.3.508>
26. Johnsen, G., Ludvigsen, M., Sørensen, A. & Sandvik Aas, L. M. The use of underwater hyperspectral imaging deployed on remotely operated vehicles - methods and applications. *IFAC-PapersOnLine* **49**, 476–481 (2016). <https://doi.org/10.1016/j.ifacol.2016.10.451>
27. Cimoli, E. *et al.* Mapping the in situ microspatial distribution of ice algal biomass through hyperspectral imaging of sea-ice cores. *Sci Rep* **10**, 21848 (2020). <https://doi.org/10.1038/s41598-020-79084-6>
28. Foglini, F. *et al.* Application of Hyperspectral Imaging to Underwater Habitat Mapping, Southern Adriatic Sea. *Sensors* **19**, 2261 (2019). <https://doi.org/10.3390/s19102261>
29. Ferrera, M., Arnaubec, A., Istenic, K., Gracias, N. & Bajjouk, T. Hyperspectral 3D Mapping of Underwater Environments. in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* 3696–3705 (IEEE, 2021). doi:10.1109/ICCVW54120.2021.00413. <https://doi.org/10.1109/ICCVW54120.2021.00413>
30. Bongiorno, D. L., Bryson, M., Bridge, T. C. L., Dansereau, D. G. & Williams, S. B. Coregistered Hyperspectral and Stereo Image Seafloor Mapping from an Autonomous Underwater Vehicle. *J Field Robot* **35**, 312–329 (2018). <https://doi.org/10.1002/rob.21713>
31. Johnsen, G. Underwater hyperspectral imaging. Preprint at (2013).
32. Lee, Z., Carder, K. L., Mobley, C. D., Steward, R. G. & Patch, J. S. Hyperspectral remote sensing for shallow waters: 2. Deriving bottom depths and water properties by optimization. *Appl Opt* **38**, 3831–3843 (1999). <https://doi.org/10.1364/AO.38.003831>
33. Kumar, P. *et al.* Combined use of a frame and a linear pushbroom camera for deep-sea 3D hyperspectral mapping. in *OCEANS 2021: San Diego – Porto* 1–9 (IEEE, 2021). <https://doi.org/10.23919/OCEANS44145.2021.9706053>.
34. Montes-Herrera, J. C. *et al.* Underwater Hyperspectral Imaging (UHI): A Review of Systems and Applications for Proximal Seafloor Ecosystem Studies. *Remote Sens (Basel)* **13**, 3451 (2021). <https://doi.org/10.3390/rs13173451>
35. Schoening, T. *et al.* Making marine image data FAIR. *Sci Data* **9**, 414 (2022). <https://doi.org/10.1038/s41597-022-01491-3>
36. Wilkinson, M. D. *et al.* Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 1–9 (2016). <https://doi.org/10.1038/sdata.2016.18>
37. Collins, S. *et al.* Turning FAIR into reality: Final report and action plan from the European Commission expert group on FAIR data. Preprint at (2018). <https://doi.org/10.2777/1524>
38. Genova, F. *et al.* *Recommendations on FAIR metrics for EOOSC*. (Publications Office of the European Union, 2021). <https://doi.org/10.2777/70791>
39. Group, R. D. A. F. D. M. M. W. & others. FAIR Data Maturity Model: specification and guidelines.

- Res. Data Alliance 2019–2020 (2020). <https://doi.org/10.15497/RDA0050>
40. Lamprecht, A.-L. *et al.* Towards FAIR principles for research software. *Data Science* **3**, 37–59 (2019). <https://doi.org/10.1038/s41597-022-01710-x>
  41. Genoveva, F. & Jones, S. Recommendations on FAIR Metrics for EOSC Report from the. (2021).
  42. De Smedt, K., Koureas, D. & Wittenburg, P. FAIR Digital Objects for Science: From Data Pieces to Actionable Knowledge Units. *Publications* **8**, 21 (2020). <https://doi.org/10.3390/publications8020021>
  43. Langenkämper, D., Zurowietz, M., Schoening, T. & Nattkemper, T. W. BIIGLE 2.0 - browsing and annotating large marine image collections. *Front Mar Sci* **4**, 1–10 (2017). <https://doi.org/10.3389/fmars.2017.00083>
  44. Jiang, P., Ergu, D., Liu, F., Cai, Y. & Ma, B. A Review of Yolo Algorithm Developments. *Procedia Comput Sci* **199**, 1066–1073 (2022). <https://doi.org/10.1016/j.procs.2022.01.135>
  45. Schoening, T. iFDO Creation Version 1.0.0. and Supplement, Version 1.0.0. (2021). <http://dx.doi.org/10.5281/zenodo.5681429>
  46. RDA FAIR Data Maturity Model Working Group. FAIR Data Maturity Model: specification and guidelines. *Research Data Alliance 2019–2020* (2020). <https://doi.org/10.15497/RDA00050>
  47. Zurowietz, M., Langenkämper, D., Hosking, B., Ruhl, H. A. & Nattkemper, T. W. MAIA—A machine learning assisted image annotation method for environmental monitoring and exploration. *PLoS One* **13**, e0207498 (2018). <https://doi.org/10.1371/journal.pone.0207498>
  48. Zurowietz, M. & Nattkemper, T. W. Unsupervised Knowledge Transfer for Object Detection in Marine Environmental Monitoring and Exploration. *IEEE Access* **8**, 143558–143568 (2020). <https://doi.org/10.1109/ACCESS.2020.3014441>
  49. Mbani, B., Schoening, T., Gazis, I.-Z., Koch, R. & Greinert, J. Implementation of an automated workflow for image-based seafloor classification with examples from manganese-nodule covered seabed areas in the Central Pacific Ocean. *Sci Rep* **12**, 15338 (2022). <https://doi.org/10.1038/s41598-022-19070-2>
  50. Schoening, T. *et al.* Megafauna community assessment of polymetallic-nodule fields with cameras: platform and methodology comparison. *Biogeosciences* **17**, 3115–3133 (2020). <https://doi.org/10.5194/bg-17-3115-2020>
  51. Schoening, T., Osterloff, J. & Nattkemper, T. W. RecoMIA—Recommendations for Marine Image Annotation: Lessons Learned and Future Directions. *Front Mar Sci* **3**, 59 (2016). <https://doi.org/10.3389/fmars.2016.00059>
  52. Schoening, T. SHiPCC—A Sea-going High-Performance Compute Cluster for Image Analysis. *Front Mar Sci* **6**, (2019). <https://doi.org/10.3389/fmars.2019.00736>
  53. Cichy, R. M. & Kaiser, D. Deep Neural Networks as Scientific Models. *Trends Cogn Sci* **23**, 305–317 (2019). <https://doi.org/10.1016/j.tics.2019.01.009>
  54. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv Neural Inf Process Syst* 1–9 (2012) <https://doi.org/10.1016/j.protcy.2014.09.007>
  55. Schoening, T., Kuhn, T., Bergmann, M. & Nattkemper, T. W. DELPHI-fast and adaptive computational laser point detection and visual footprint quantification for arbitrary underwater image collections. *Front Mar Sci* **2**, (2015). <https://doi.org/10.3389/fmars.2015.00020>
  56. Kwasnitschka, T. *et al.* DeepSurveyCam—A deep ocean optical mapping system. *Sensors (Switzerland)* **16**, (2016). <https://doi.org/10.3390/s16020164>
  57. Durden, J. M. *et al.* Comparison of image annotation data generated by multiple investigators for benthic ecology. *Mar Ecol Prog Ser* **552**, 61–70 (2016). <https://doi.org/10.3354/meps11775>
  58. Simon-Lledó, E. *et al.* Megafaunal variation in the abyssal landscape of the Clarion Clipperton Zone. *Prog Oceanogr* **170**, (2019). <https://doi.org/10.1016/j.pocean.2018.11.003>
  59. Osterlo, J. *et al.* Ranking color correction algorithms using cluster indices. in *Proceedings - 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery, CVAUI 2014* 41–48 (2014). doi:10.1109/CVAUI.2014.13. <https://doi.org/10.1109/CVAUI.2014.13>
  60. Thomanek, K., Zielinski, O., Sahling, H. & Bohrmann, G. Automated gas bubble imaging at sea floor - A new method of in situ gas flux quantification. *Ocean Science* **6**, 549–562 (2010). <https://doi.org/10.5194/os-6-549-2010>

61. Jordt, A., Zelenka, C., von Deimling, J. S., Koch, R. & Köser, K. The Bubble Box: Towards an Automated Visual Sensor for 3D Analysis and Characterization of Marine Gas Release Sites. *Sensors* **15**, 30716–30735 (2015). <https://doi.org/10.3390/s151229825>
62. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun ACM* **60**, 84–90 (2017). <https://doi.org/10.1145/3065386>
63. Katija, K. *et al.* FathomNet: A global image database for enabling artificial intelligence in the ocean. *Sci Rep* **12**, 15914 (2022). <https://doi.org/10.1038/s41598-022-19939-2>
64. Zurowietz, M. & Nattkemper, T. W. Current trends and future directions of large scale image and video annotation: Observations from four years of BIIGLE 2.0. *Front Mar Sci* (2021). <https://doi.org/10.3389/fmars.2021.760036>
65. Creswell, A. *et al.* Generative Adversarial Networks: An Overview. *IEEE Signal Process Mag* **35**, 53–65 (2018). <https://doi.org/10.48550/arXiv.1710.07035>
66. Jascha Sohl-Dickstein, Eric A. Weiss, Niru Maheswaranathan & Surya Ganguli. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. <https://doi.org/10.48550/arXiv.1503.03585>
67. Qassim, H., Verma, A. & Feinzimer, D. Compressed residual-VGG16 CNN model for big data places image recognition. in *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)* 169–175 (IEEE, 2018). doi:10.1109/CCWC.2018.8301729. <https://doi.org/10.1109/CCWC.2018.8301729>
68. Koonce, B. ResNet 50. in *Convolutional Neural Networks with Swift for Tensorflow* 63–72 (Apress, 2021). doi:10.1007/978-1-4842-6168-2\_6. [https://doi.org/10.1007/978-1-4842-6168-2\\_6](https://doi.org/10.1007/978-1-4842-6168-2_6)
69. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mach Intell* **39**, 1137–1149 (2017). <https://doi.org/10.48550/arXiv.1506.01497>
70. Wang, Y., Wang, C., Zhang, H., Dong, Y. & Wei, S. Automatic Ship Detection Based on RetinaNet Using Multi-Resolution Gaofen-3 Imagery. *Remote Sens (Basel)* **11**, 531 (2019). <https://doi.org/10.3390/rs11050531>
71. Howell, K. L. *et al.* A framework for the development of a global standardised marine taxon reference image database (SMarTaR-ID) to support image-based analyses. *PLoS One* **14**, e0218904 (2019). <https://doi.org/10.1371/journal.pone.0218904>
72. Wieczorek, J. *et al.* Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLoS One* **7**, e29715 (2012). <https://doi.org/10.1371/journal.pone.0029715>
73. Jaffe, J. S. Computer modeling and the design of optimal underwater imaging systems. *Oceanic Engineering, IEEE Journal of* **15**, 101–111 (1990). <https://doi.org/10.1109/48.50695>
74. Purser, A., Bergmann, M., Lundqvist, T., Ontrup, J. & Nattkemper, T. W. Use of machine-learning algorithms for the automated detection of cold-water coral habitats: A pilot study. *Mar Ecol Prog Ser* **397**, 241–251 (2009). <http://dx.doi.org/10.3354/meps08154>
75. Schoening, T., Kuhn, T., Jones, D. O. B., Simon-Lledo, E. & Nattkemper, T. W. Fully automated image segmentation for benthic resource assessment of poly-metallic nodules. *Methods in Oceanography* **15–16**, 78–89 (2016). <https://doi.org/10.1016/j.mio.2016.04.002>
76. Schoening, T., Jones, D. O. B. & Greinert, J. Compact-Morphology-based poly-metallic Nodule Delineation. *Sci Rep* **7**, 1–12 (2017). <https://doi.org/10.1038/s41598-017-13335-x>
77. Jahandad, Sam, S. M., Kamardin, K., Amir Sjarif, N. N. & Mohamed, N. Offline Signature Verification using Deep Learning Convolutional Neural Network (CNN) Architectures GoogLeNet Inception-v1 and Inception-v3. *Procedia Comput Sci* **161**, 475–483 (2019). <https://doi.org/10.1016/j.procs.2019.11.147>
78. Yamada, T., Prügel-Bennett, A. & Thornton, B. Learning features from georeferenced seafloor imagery with location guided autoencoders. *J Field Robot* **38**, 52–67 (2021). <https://doi.org/10.1002/rob.21961>

## 9. Document Information

|                        |   |                |           |
|------------------------|---|----------------|-----------|
| <b>EU Project N°</b>   | 818123  | <b>Acronym</b> | iAtlantic |
| <b>Full Title</b>      | Integrated Assessment of Atlantic Marine Ecosystems in Space and Time |                |           |
| <b>Project website</b> | <a href="http://www.iatlantic.euht">www.iatlantic.euht</a>            |                |           |

|                     |           |     |              |   |
|---------------------|-----------|-----|--------------|---|
| <b>Deliverable</b>  | <b>N°</b> | 2.4 | <b>Title</b> | New imaging and analysis approaches for marine species detection and classification |
| <b>Work Package</b> | <b>N°</b> | 2   | <b>Title</b> | Mapping Atlantic Ecosystems   |

|                            |                    |  |               |    |
|----------------------------|--------------------|--|---------------|----|
| <b>Date of delivery</b>    | <b>Contractual</b> | PU   | <b>Actual</b> | PU |
| <b>Dissemination level</b> |                    | PU Public, fully open, e.g. web  |               |    |
|                            |                    | CO Confidential restricted under conditions set out in Model Grant Agreement |               |    |
|                            |                    | CI Classified, information as referred to in Commission Decision 2001/844/EC |               |    |

|                             |  |                |               |                      |
|-----------------------------|--|----------------|---------------|----------------------|
| <b>Authors (Partner(s))</b> | Carlos Dominguez Carrio, Telmo Morato, Touria Baijouk, Maxime Ferrera, Tristan Petit & Aurélien Arnaubec |                |               |                      |
| <b>Responsible Authors</b>  | <b>Names</b>   | Timm Schoening | <b>Emails</b> | tschoening@geomar.de |