Ecological Informatics 91 (2025) 103409

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

A digital-twin strategy using robots for marine ecosystem monitoring

Jacopo Aguzzi [a,b,*], Elias Chatzidouros [c], Damianos Chatzievangelou [a], Morane Clavel-Henry [a,d], Sascha Flögel [d], Nixon Bahamon [a], Michael Tangerlini [k,b], Laurenz Thomsen [e], Giacomo Picardi [a], Joan Navarro [a], Ivan Masmitja [a,f], Nathan J. Robinson [a,g], Tim Nattkemper [h], Sergio Stefanni [b], José Quintana [i], Ricard Campos [i], Rafael García [i,j], Emanuela Fanelli [k], Marco Francescangeli [m], Luca Mirimin [l], Roberto Danovaro [k], Daniel Mihai Toma [f], Joaquín Del Rio-Fernandez [f], Enoc Martinez [f], Pol Baños [f], Oriol Prat [f], David Sarria [f], Matias Carandell [f], Jonathan White [m], Thomas Parissis [n], Stavroula Panagiotidou [n], Juliana Quevedo [a], Silvia Gallegati [k], Jordi Grinyó [a], Erik Simon-Lledó [o], Joan B. Company [a], Jennifer Doyle [m]

[a] Institut de Ciències del Mar (ICM), CSIC, Paseo Marítimo de la Barceloneta 37-49, 08003 Barcelona, Spain
[b] Stazione Zoologica Anton Dohrn, Via Francesco Caracciolo, 80121 Naples, Italy
[c] Engitec Systems International Limited, 143 Spyrou Kyprianou Avenue, 3083 Limassol, Cyprus
[d] GEOMAR Helmholtz Center for Ocean Research Kiel, Wischhofstraße 1-3, 24148 Kiel, Germany
[e] University of Gothenburg, Universitetsplatsen 1, 40530 Gothenburg, Sweden
[f] SARTI Research Group, Electronic Department, Universitat Politecnica de Catalunya (UPC), Rambla Exposición 24, Vilanova i la Geltrú, 08800, Spain
[g] Fundación Oceanogràfic de la Comunitat Valenciana, Carrer d'Eduardo Primo Yúfera 1, 46013 Valencia, Spain
[h] Biodata Mining Group, Faculty of Technology, Bielefeld University, Universitätsstraße 25, 33615 Bielefeld, Germany
[i] Coronis Computing SL, Carrer Pic de Peguera, 11, 17003 Girona, Spain
[j] Computer Vision and Robotics Institute, University of Girona, Carrer de la Universitat de Girona, 6, 17003 Girona, Spain
[k] Department of Life and Environmental Sciences, Polytechnic University of Marche and National Biodiversity Future Center, Via Brecce Bianche, 60131 Ancona, Italy
[l] Department of Natural Resources and the Environment, School of Science and Computing, Atlantic Technological University, Dublin Road, H91 T8NW Galway, Ireland
[m] Marine Institute, Fisheries and Ecosystems Advisory, Rinville, Oranmore, Co. Galway H91 R673, Ireland
[n] STRATAGEM Energy Ltd, Agiou Andreou 365, 3035 Limassol, Cyprus
[o] Ocean BioGeosciences, National Oceanography Center, European Way, SO14 3ZH Southampton, United Kingdom

ARTICLE INFO

Keywords:
Marine monitoring
Robotic platforms
Data integration
Ecological indicators
Machine learning
Spatial modelling

ABSTRACT

Effective marine conservation and management require ecological monitoring in the form of intensive real-time data collection over large spatial scales. The combined use of fixed platforms (e.g., cabled observatories) and research vessels with platforms of different levels of teleoperated autonomy (e.g., remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs) can contribute to the acquisition of large multiparametric biological and environmental data. If those data are spatially combined, sufficient spatial coverage can be achieved for ecological monitoring. A digital twin of the ocean (DTO) approach can then be used as a virtual representation of that monitored space, enabling multiparametric analyses of environmental patterns and processes affecting biodiversity and species distributions, as well as socioeconomic activities. Here, we propose a general architecture for a DTO centred on real-time data collection from local networks on fixed and mobile platforms, such as the physical twin observers (PTO), which is synergistically merged with platforms operating at large geographic scales. We describe a roadmap to achieve this DTO via 4 key steps: (1) acquisition of in situ data with a robotic network of platforms; (2) the application of AI in image processing for extracting biological data; (3) big data management with data bubbles; and (4) development of the resulting DTO framework for providing ecosystem monitoring via the computation of ecological indicators and socioecological modelling.

* Corresponding author at: Paseo Marítimo de la Barceloneta 37-49, 08003 Barcelona, Spain.
E-mail address: jaguzzi@icm.csic.es (J. Aguzzi).

https://doi.org/10.1016/j.ecoinf.2025.103409
Received 20 August 2024; Received in revised form 30 July 2025; Accepted 26 August 2025
Available online 28 August 2025

1. Introduction

On February 21, 2023, the European Commission adopted the Marine Action Plan, which aimed to support sustainable and resilient fisheries within the European Union (EU) by protecting and restoring marine ecosystems. This plan aligns with the EU's 2030 Biodiversity Strategy, in which 30 % of all EU marine habitats should be legally protected by 2030 (European Comission (EU), 2021). To achieve this goal, there is an urgent need to develop technologies and protocols for data acquisition and processing for ecological monitoring over broader spatial and temporal scales than currently possible (Danovaro et al., 2017).

Digital twins (DTs) are virtual representations of physical objects (or systems) based on a complex array of diverse modelling approaches used to simulate their status and behaviour (De Koning et al., 2023; Segovia and Garcia-Alfaro, 2022; Tzilivakis, 2022). DTs are continually updated via real-time data and machine learning approaches to improve model outputs, and "what-if" scenarios can be explored by changing model parameters (Nativi et al., 2021). A digital twin of the ocean (DTO) is therefore a virtual representation of the marine biosphere built with a range of oceanographic, meteorological, biological, and socioeconomic data, enabling multiparametric analyses of environmental patterns and processes, such as ecosystem responses to natural phenomena and anthropogenic impacts (Barbie et al., 2022; Boschert and Rosen, 2016; Brönner et al., 2023; De Koning et al., 2023; Grossmann et al., 2022; Pillai et al., 2022; Schneider et al., 2023; Tzachor et al., 2023; Yu et al., 2024).

The data-reliant nature of DTOs makes them suitable for marine systems for which large amounts of information are available (Murawski et al., 2010). Effective DTOs are therefore dependent on data from various sources, ranging from oceanographic campaigns on vessels, which support classic (e.g., trawl nets, towed sledges, and coring) and novel data collection platforms (e.g., remotely operated vehicles-ROVs and autonomous underwater vehicles-AUVs), to large permanent scientific monitoring (e.g., cable observatories) and industrial infrastructures. Combined, these data-generation platforms can be considered part of a broader physical twin observer (PTO). While the PT corresponds to the real-world system (i.e., the ocean itself), within the framework of adaptive ecological monitoring, the embedded sensor platforms are also integral components of the physical layer. These platforms do not merely collect data but actively interact with the environment, and their operations must be modelled and simulated within the DTO to realistically reflect monitoring dynamics. The function of these platforms can be adapted on the basis of iterative feedback from the DTO modelling process to highlight key areas in need of greater spatial or temporally intensive sampling. Permanently deployed marine platforms, which involve multiparametric biological and environmental data collection, are growing worldwide (Danovaro et al., 2017) and represent the core PTO monitoring units of future DTOs. The European Multidisciplinary Seafloor and Water-column Observatories (EMSO), the Ocean Network Canada (ONC) and the Cubic Kilometer Neutrino Telescope Network (KM3NeT) are ensembles of cabled observatories and moored platforms that perform coordinated, temporally intensive, 4D data acquisition in the benthopelagic realm (Dañobeitia et al., 2023; Moran et al., 2022). These networks also host docked mobile platforms such as crawlers, which collect biological and environmental information at large spatial scales around fixed nodes (Aguzzi et al., 2019; Rountree et al., 2020). Their ecological monitoring functionalities are currently approaching the full virtualization of their deployment areas, a process based on web visualization via the standardization of sensors, data collection procedures and processing methods (Howe et al., 2010; Lantéri et al., 2022).

An established DTO framework for general reference in Europe is under construction and extends beyond the combination of different big data sources for modelling (*European Union,* n.d.). Such a DTO framework can be efficiently implemented in areas where permanent scientific infrastructures, such as PTOs, support decadal ecological monitoring; e.g., networks of benthic cabled observatories. Here, we propose a general architecture for a DTO dedicated to the ecological monitoring via the spatiotemporal integrated collection of biological and environmental multiparametric data. That DTO architecture encompasses data collection by local permanent and nearby transiently operating PTOs with the capacity to collect biological and environmental multiparametric data over a wide range of spatial and temporal scales. Even though these PTOs are not currently modelled to be fully-represented elements of the actual DTO yet, it is envisioned that their structure and functionality will be fully included in the future in the framework of adaptive monitoring needs. We describe a roadmap of how the ecological data collection could be achieved via 4 key steps: (1) acquisition of in situ data with a robotic network of platforms; (2) the application of AI in image processing for extracting biological data; (3) the introduction of a DTO approach for big data management with data bubbles; and (4) development of a DTO framework for providing major services in ecosystem monitoring via ecological indicators and socioecological modelling.

2. Acquisition of in situ data with a robotic network of platforms

Marine robotics refers to the application of autonomous systems in data collection and in situ processes at different spatiotemporal scales (Aguzzi et al., 2024). Ecological monitoring is becoming increasingly independent of the presence of humans in the field with the extension of remote-control technology (Chatzievangelou et al., 2022). In the framework of the marine Internet of Things (Glaviano et al., 2022), the operability of fixed and mobile platforms can be coordinated to establish cooperative networks in strategic marine areas, resulting in the creation of in situ marine laboratories (Aguzzi et al., 2019; Rountree et al., 2020). Consequently, data collection platforms such as PTOs are being integrated into the DTO framework because the functioning of those platforms can be modelled and hence adaptively adjusted on the basis of ongoing monitoring results (Danovaro et al., 2017).

2.1. The use of platforms operating at different spatiotemporal scales

Two key types of PTOs (Fig. 1) serve as the data generators for DTOs: fixed-local PTOs (hereafter defined as F-PTOs) and geographically expanded PTOs (hereafter defined as G-PTOs). F-PTOs include stationary platforms such as cabled observatories and landers with docked locally mobile crawlers or AUVs, as well as anchored or moored water column platforms and surface buoys (Aguzzi et al., 2019, 2020; Aguzzi et al., 2024; Rountree et al., 2020). Additionally, G-PTOs represent mobile long-range technologies, such as vessels and corresponding platforms (e.g., ROVs, towed sledges, AUVs, and pelagic cameras), along with other independent, drifting buoys (e.g., ARGO floats) (Dominguez-Carrió et al., 2021; Jayne et al., 2017; Lambertini et al., 2022; Robinson et al., 2021; Yu et al., 2024). This category also encompasses technologies launched from shore (e.g., wave gliders, autonomous surface vessels (ASVs), drones (Phillips et al., 2019; Zhang et al., 2019) and use of satellites (Mohseni et al., 2022).

The need for building DTOs that can integrate density, biomass and derived biodiversity data at local (F-PTOs) and geographically expanded (G-PTOs) scales is motivated by the high variability in species distributions within the three-dimensional marine environment. Traditional sampling methods such as trawling, beach seines, or visual census, often produce biased results due to the activity rhythms of pelagic and benthic species (Aguzzi and Company, 2010). When measurement repetitions are not scheduled to account for the varying temporal scales of species movements, the observed species assemblage shows high variability as different species rhythmically enter and leave the sampling windows Repeating sampling at hourly scale, and prosecuting that sampling continuously in order to cover the seasonal cycle at a single location is necessary to detect population movements in and out of targeted areas (Aguzzi et al., 2015), which helps explain results on a larger geographic

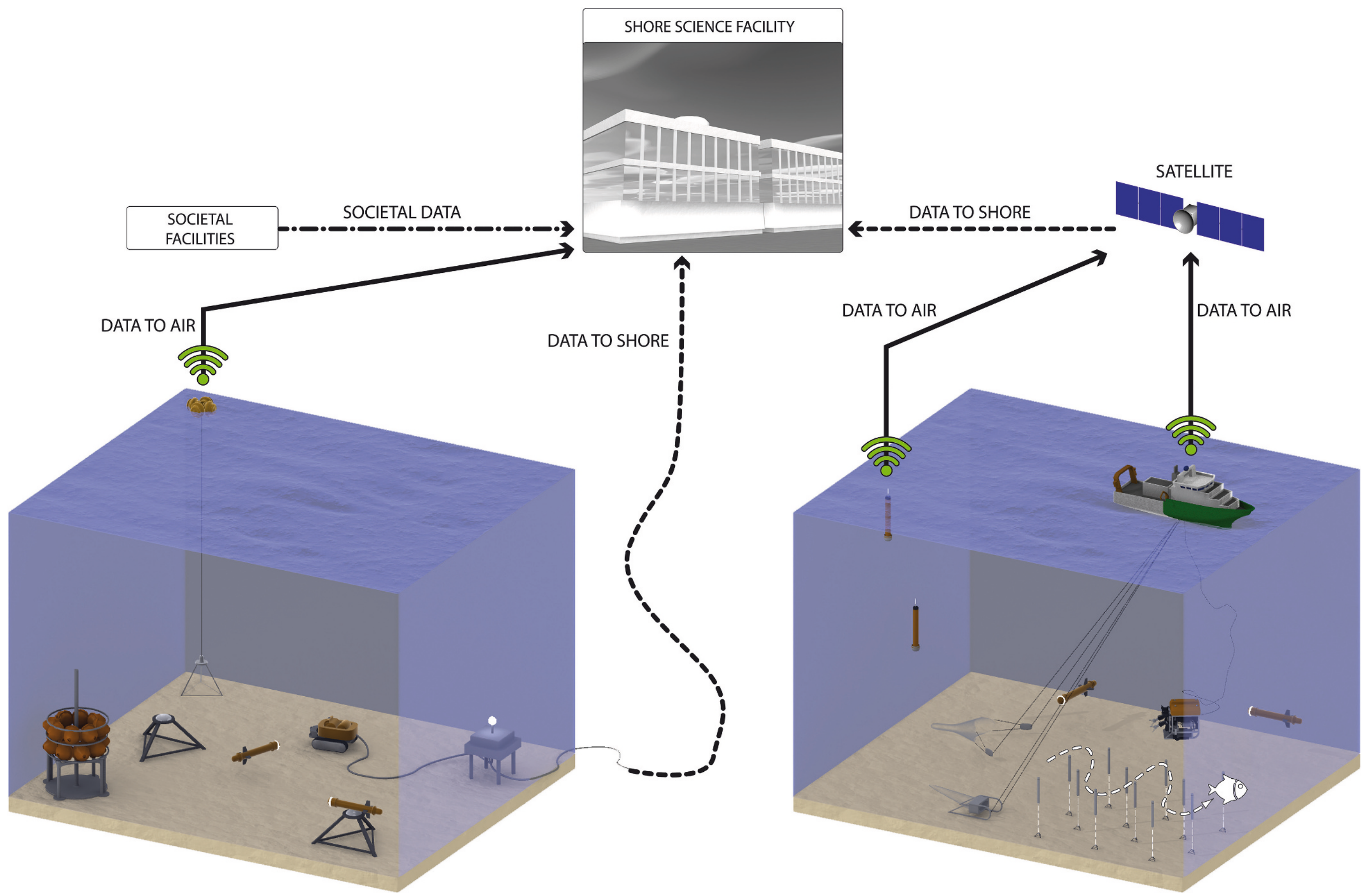

Fig. 1. Schematic illustration of the physical infrastructure supporting the Digital Twin of the Ocean (DTO). The figure shows the spatial layout of fixed and mobile platforms (e.g., moored sensors, seabed observatories, research vessels) and the pathways through which data are transmitted to a central shore-based science facility. This facility functions as a data hub, integrating real-time environmental and biological observations with complementary data streams (e.g., from societal stakeholders), thus enabling centralized control, storage, and further processing toward DTO development.

scale. Integrating local observation nodes within broader geographic networks can enhance the accuracy of video-observations by enabling comprehensive comparisons, thus facilitating species abundance estimations similarly to other mobile sampling methods over larger seabed areas.

Local F-PTO data can be compared with similar data at larger geographic scales provided by G-PTOs, hence addressing the benefits and weaknesses of both monitoring strategies (see Fig. 1); notably, it is important to both build and keep developing DTOs. For example, data on species density, biomass, and overall biodiversity obtained via F-PTO monitoring should be compared with more spatially broad data collected by G-PTOs and nearby fishery operations for ecological representativeness. These broader datasets include conventional ecological monitoring methods such as echo sounder surveys to detect fish schools (e.g., Simmonds and MacLennan, 2005), underwater visual census techniques for assessing species presence and abundance in shallow habitats (e.g., Colton and Swearer, 2010), and fishing surveys which remain essential tools for estimating fish stock distributions and community structure across large areas (e.g., Rufener et al., 2021). Integrating these classical approaches with robotic-based monitoring platforms enhances the capacity of DTO frameworks to capture both fine-scale and regional ecological dynamics.

2.2. Technical limitations and solution affecting the use of F-PTOs in ecological monitoring

A list of technical specifications for platforms and their sensors related to local robotic networks for ecological monitoring (see Fig. 1) is reported in Table 1. In that Table, we also reported the monitored policy variables with the categories for the Essential Biodiversity Variables (EBVs), Essential Oceanographic Variables (EOVs; Muller-Karger et al., 2018) and Marine Strategy Framework Directive Good Environmental Status Descriptors (MSFD GES; Directive 2008/56/EC). The deployment capabilities of these platforms range from AUVs and deep-sea landers to shallow-water cable observatories, each developed to meet specific monitoring requirements. The specifications in Table 1 therefore outline the technological and operational requirements of these platforms, including their sensor configurations, power consumption, data acquisition modes (continuous vs. time-lapsed), onboard processing capabilities, and data transmission methods according to our late experience (see Table 1 caption references). Furthermore, the Table provides insights into deployment and maintenance costs, which are critical for long-term ecological observation networks. The deployment cost column considers not only the price of the asset (platform) itself but also the expenses associated with its transportation, personnel involvement, and operational setup within the observation area. By integrating various platforms, including cabled and non-cabled systems, data collection is extended across a wide range of spatial and temporal scales, enhancing the robustness of DTO initiatives. Continuous data acquisition refers to real-time, uninterrupted measurements, which are essential for capturing rapid environmental changes and dynamic biological activities. In contrast, time-lapsed data acquisition, typically recorded every 30 min, provides a more energy-efficient approach that allows for periodic sampling while still ensuring temporal resolution sufficient for long-term monitoring trends.

Scientists and technologists are collaboratively developing a

Table 1
Technical details on operational specifications and costs for the establishment of ecological monitoring networks as F-PTOs. Commercially available platforms models can be very variable and specifications have been tailored on the OBSEA cabled observatory (www.obsea.es) and its local network (Del-Rio et al., 2020; Falahzadeh et al., 2023; Masmitja et al., 2024) as well as at the Ocean Network Canada (ONC; Purser et al., 2013). MANSIO-VIATOR specifications are presented in Flögel (2018), while for SLM are from Juva et al. (2020, 2021) and Büscher et al. (2024). Codes fand full policy variable names for EBVs and EOVs are reported in Appendix 1. For EBVs definitions are from https://geobon.org/ebvs/what-are-ebvs/; EOVs definitions are from https://goosocean.org/what-we-do/framework/essential-ocean-variables/; MSFD GES Descriptors definitions are from https://environment.ec.europa.eu/topics/marine-environment/descriptors-under-marine-strategy-framework-directive_en.

Platform	Installed Sensors	Measured Variable	EBVs; EOVs; MSFD GES	Type of Battery	Fuel Cell connectability (W/h)	Energy for Data Collection (Continuous; W/h)	Energy for Data Collection (Time-Lapsed, 30 min; W/h)	Energy for displacement (W/h)	Onboard Processing	Time Requirements of Processing (seconds)	Acoustic Modem Data Transmission	Cost of Deployment (€)	Cost of Maintenance (month; €)
Cabled Observatory	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–18; BE2–4,,6,12; MSFD1–4	–	–	7.5	0.02	NONE (Fixed)	Animal Classification and Counting	0.1	Not required (unless for data exchange with non-wired, satellite platforms)	400,000	12,000
	CTD	Temperature, Salinity, Depth	P6,10,12			0.4	0.000024		None	0.01			
	AWAC	Current Speed and Direction and Wave height, direction and period	P1,4,7–8; MSFD7			9	0.3		None	600			
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			0.24	0.000001		None	0.1			
	O_2	Oxygen concentration	BGC1			0.22	0.000001		None	0.1			
	Ph	Acidity	BGC3			1.02	0.00005		None	0.05			
	PAM	Species Sound and Maritime Noise	EBV5–6,9-10,12–13, 17–18; BE3–4,6; CD3; MSFD1–4,11			2.5	0.08		Animal Identification and Counting	60			
Lander	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–18; BE2–3,12	Li-ion	Possible	0.045	0.045	NONE (Fixed)	Animal Classification and Counting	0.5	Not required (unless for data exchange with non-wired, seabed platforms)	45,000	800
	CTD	Temperature, Salinity, Depth	P6,10,12			0.001	0.000034		None	0.01			
	Fluorometer	Chlorophyll-a, PAR	BGC5; BE1; MSFD5			0.24	0.008		None	0.1			
	PAM	Species Sound and Maritime Noise	EBV5–6,9-10,12–13, 17–18; BE3–4,10; CD3; MSFD1–4,11			1.2	0.04		None	60			
Satellite Lander Module (SLM)	CTD	Temperature, Salinity, Depth	P6,10,12	Primary cells (Li-SOCl2)	Possible	0.2565	0.01	324	None	0,01	Not required (unless for data exchange with non-wired, seabed platforms)	120,000	900
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			0.6	0.02		None	0,1			
	ADCP	Current Speed and Direction	P8			3.78	0.126		None	600			
	O_2	Oxygen	BGC1			0.21744	0.01		None	0,5			

(continued on next page)

Table 1 (*continued*)

Platform	Installed Sensors	Measured Variable	EBVs; EOVs; MSFD GES	Type of Battery	Fuel Cell connectability (W/h)	Energy for Data Collection (Continuous; W/h)	Energy for Data Collection (Time-Lapsed, 30 min; W/h)	Energy for displacement (W/h)	Onboard Processing	Time Requirements of Processing (seconds)	Acoustic Modem Data Transmission	Cost of Deployment (€)	Cost of Maintenance (month; €)
Fuel Cell Lander	CTD	Temperature, Salinity, Depth	P6,10,12	H2O2, LiPo	150,000	0.2565	0.01	NONE (Fixed	None	None	Required for docking of mobile platforms	220,000	3000
Coastal Cabled Crawler	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–18; BE2–4,6,12; CD2; MSFD1–4,6,10	–	Possible	7.5	0.02	57.6	Animal Classification and Counting	0.1	Not required since tethered to the observatory of when in stand-alone fashion, it operates with WiFi surface buoy (acoustic modem can anyway be required for data exchange with other lander platforms)	100,000	400
	CTD	Temperature, Salinity, Depth	P6,10,12			0.4	0.000024		None	0.01			
	ADCP	Current Speed and Direction	P8			9	0.3		None	600			
	Fluorometer	Chlorophyll-a, PAR	BGC5; BE1; CD1; MSFD5			0.24	0.000001		None	0.1			
Deep-Sea untethered Crawler *Rossia*	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–18; BE2–4,6,12; CD2; MSFD1–4,6,10	LiPo	Possible	1	0.016	150	Animal Classification and Counting	0.2	Not required since tethered to the observatory of when in stand-alone fashion, it operates with Junction box within a cabled observatory	150,000	2000
	CTD	Temperature, Salinity, Depth	P6,10,12			0.5	0.00013		None	0.01			
	ADCP	Current Speed and Direction	P8			2	0.35		None	600			
	CH_4	Methane concentration	BGC8; BE11			6.25	0,1		None	0.1			
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			2.9	0.048		None	0.1			
	Particle Camera	Turbidity (particle size >100 µm)	BGC5			1.04	0.02		Particle Counting and Sizing	0.1			
Mansio-Viator (deep-sea vessel-deployed crawler)	CTD	Temperature, Salinity, Depth	P6,10,12	LiPo	Possible	0.2565	0.01	750 (Viator 8000 W/h LiPo, Mansio 4000 W/h LiPo, with inductive transfer)	None	0,01	Required, if not it can run fully autonomous (also, acoustic data transfer possible)	300,000	2500
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5	LiPo		0.6	0.02		None	0,1			
	ADCP	Current Speed and Direction	P8	LiPo		3.78	0.126		None	600			
	O_2	Oxygen concentration	BGC1	LiPo		0.21744	0.01		None	0,5			
	Camera/Line laser	Seabed morphology	EBV19–20; BE7–9; CD2; MSFD6,10	LiPo		6	0.5		Animal Classification and Counting	300			
	Navigational data	IMU/acoustic navigation	–	LiPo		3	0.5		None	180			
Moored Buoy	CTD	Temperature, Salinity, Depth	P6,10,12	LiPo	Possible	0.4	0.000024	NONE (Fixed)	None	0.01	Not required (unless for data exchange with	60,000	600

(*continued on next page*)

Table 1 (*continued*)

Platform	Installed Sensors	Measured Variable	EBVs; EOVs; MSFD GES	Type of Battery	Fuel Cell connectability (W/h)	Energy for Data Collection (Continuous; W/h)	Energy for Data Collection (Time-Lapsed, 30 min; W/h)	Energy for displacement (W/h)	Onboard Processing	Time Requirements of Processing (seconds)	Acoustic Modem Data Transmission	Cost of Deployment (€)	Cost of Maintenance (month; €)
	ADCP	Current Speed and Direction	P8			9	0.3		None	600	non-wired, seabed platforms)		
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			0.24	0.000001		None	0.1			
	PAM	Species Sound and Maritime Noise	EBV5–6,9-10,12–13, 17–18; BE3–4,6; CD3; MSFD1–4,11			1.2	0.04		None	60			
AUV	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–20; BE2–4,6-9,12; CD2; MSFD1–4,6,10	Li	Possible	7.5	0.02	200	Animal Classification and Counting	0.1	Not required (unless for data exchange with non-wired, seabed platforms)	200,000	1100
	Optoacoustic Mosaics & Mapping	Seabed Morphology	EBV5–6,17–20; BE3–4,7-9,12; CD2; MSFD1–4,6,10			20	3.2		3D Habitat Rendering plus Lebensspuren	0.04			
Gliders	CTD	Temperature, Salinity, Depth	P6,10,12	Alkaline, Li	Impossible	0.4	0.000024	3.1	None	0.01		250,000	1250
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			0.24	0.000001		None	0.1			
ASV	HD Camera	Animal counts	EBV5–7,9-10,12–15, 17–20; BE2–4,6-9,12; MSFD1–4,6,10	Li	Impossible	7.5	0.02	200	Animal Classification and Counting	0.1		400,000	1250
	CTD	Temperature, Salinity, Depth	P6,10,12			0.4	0.000024		None	0.01			
	Fluorescence / turbidity meter	Chlorophyll-a, Turbidity	BGC5; BE1; MSFD5			0.24	0.000001		None	0.1			

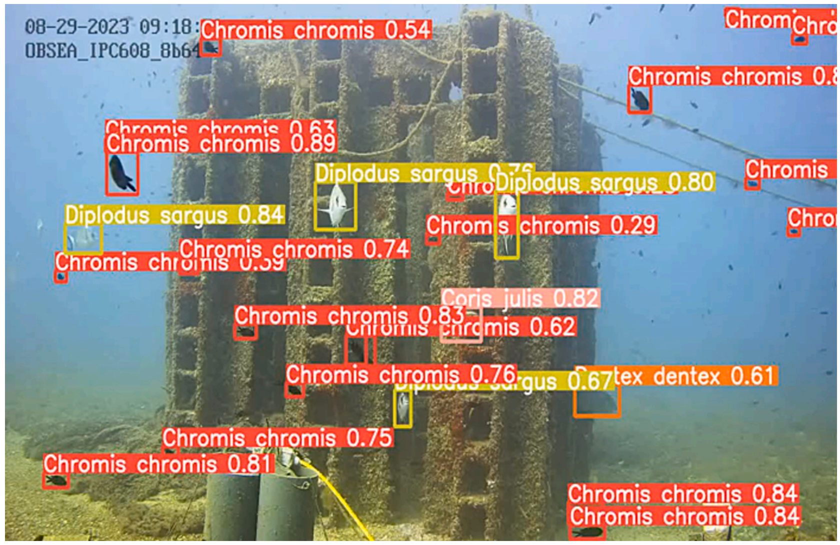

Fig. 2. Automated classification at OBSEA. An example of real-time, automated classification of coastal fishes by the OBSEA network as an example of an F-PTO, which provides biological time series of data to an observatory cyber management infrastructure (Del-Rio et al., 2020; Martínez et al., 2023). Links to the datasets and AI models are available at the Data Availability section.

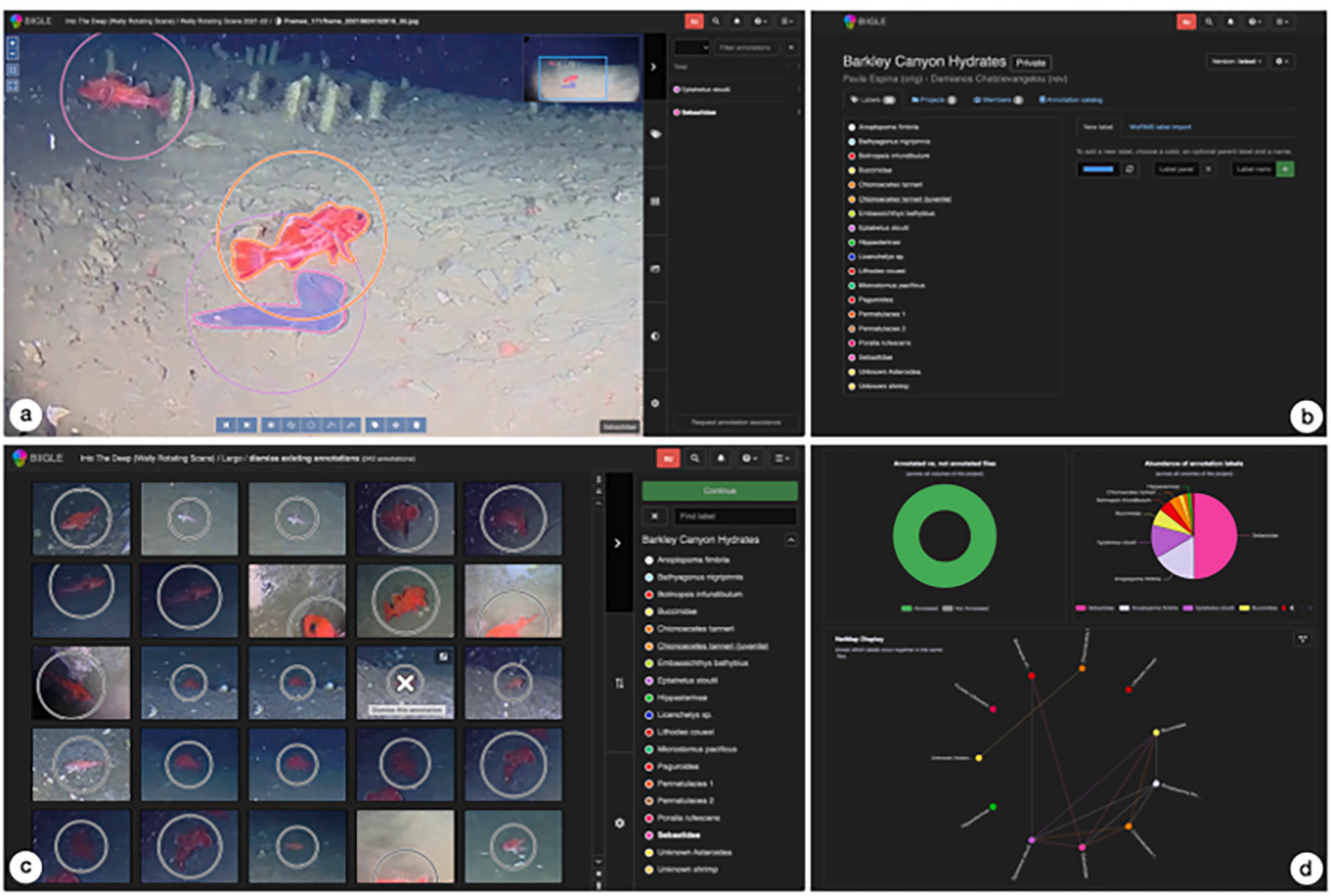

Fig. 3. BIIGLE for image and video annotation. Marine image and video data must be annotated, i.e., the localization and naming of objects of interest. Using modern state-of-the-art online annotation tools such as BIIGLE, image and video data can be browsed, managed, shared, and annotated (e.g., with circles, as shown in (a) with customized labels. Then, (b) quality control (c) or computational segmentation can be performed via incorporated deep learning algorithms such as the Segment Anything Model (SAM) (Kirillov et al., 2023). This is shown for one hagfish (*Eptatretus stoutii*) and one rockfish (*Sebastes* spp.) in the middle of (a) with contour descriptions on the basis of BIIGLE's SAM result. The BIIGLE information visualization window (d) gives a compressed overview of the annotation results obtained by a group of experts or AI tools. The images in (a) and (c) were collected by the crawler "Wally" (Chatzievangelou et al., 2022) at Ocean Networks Canada's (ONC; www.oceannetworks.ca) Barkley Canyon methane hydrate site (870 m depth).

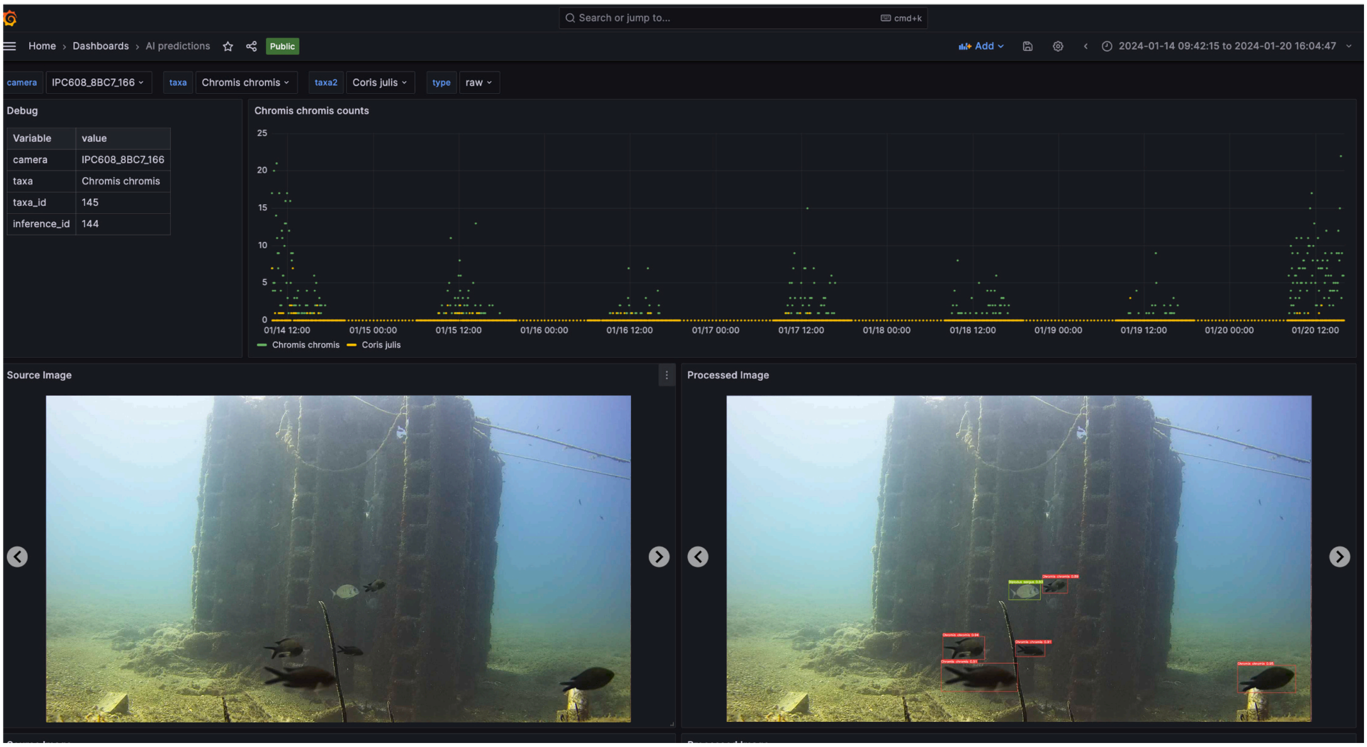

Fig. 4. Grafana visualization tool. Consecutive AI-processed images of the OBSEA artificial reef (see Fig. 2) are visualized to evaluate the performance of automation in species classification and individual counting. Moreover, time series of counts can be generated to obtain species count trends. In the example, two common coastal fish species (*Chromis chromis* in green and *Coris julis* in yellow) are automatically identified, and their time series are depicted over consecutive days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

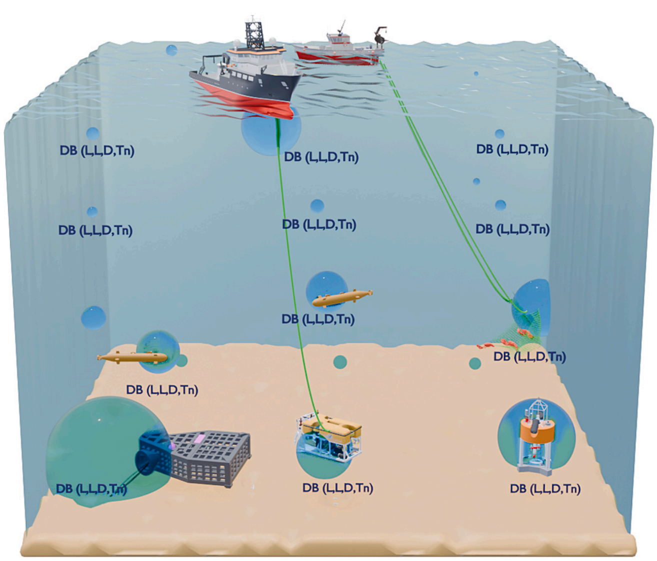

Fig. 5. Data bubbles. DTO input data are represented by a set of data bubbles (DBs) associated with any PTO component in a spatial monitoring network that includes fixed and mobile platforms with different levels of vessel tele-operations and nearby commercial operations (such as trawling fishing boats catching shrimps in this example). Platforms and tools for data collection cover a specific underwater volume during any specific time interval, so all biological and environmental data are labelled with at least 4 metadata tags: latitude and longitude (*L*,*L*), depth (D) and time (Tn). Therefore, storage in a common database requires the same typology to be applied to any other type of archived or contemporary data in nearby areas (e.g., trawler data for demersal crustacean resources in this case).

spatially adaptive, non-invasive modular platform made up of vessel-independent, wirelessly connected benthic stations and AUVs to monitor and map marine ecosystems over extended periods, autonomously. One challenge is the mission duration, which limits the use of landers and AUVs for ecological monitoring in remote areas without cables to shore sources. Landers operativity can be increased with Fuel Cells (Aguzzi et al., 2020) and that solution can also be extended to AUVs by developing docking stations. These stations provide protection, allow battery charging, and enable data transmission without adding mechanical parts to the AUV. Although various types of docking stations have been proposed in literature, funnel-based ones are the most popular for torpedo-shaped AUVs (Palomeras et al., 2018). Underwater wireless recharging techniques for AUVs have been explored in recent years, which offer a safe and reliable method for power transfer between a charging station and a vehicle (Teeneti et al., 2019).

In a cooperative network of fixed observatories with docked AUVs, wireless communication is crucial. Acoustic systems have achieved significant success underwater due to their ability to communicate over many kilometres (Song et al., 2019). However, their performance is influenced by the physical properties of the water environment, which limit bandwidth, cause high latency, produce high transmission losses, lead to time-varying multipath propagation, and create Doppler spread (Stojanovic and Preisig, 2009). These limitations prevent AUVs from transmitting real-time imaging products via acoustic communication. Therefore, a complementary technology is necessary to achieve broadband underwater communications. Visible-Light Communication (VLC) technology has the potential to address this issue. Nevertheless, VLC systems are currently limited to short-range use, and only a few commercial systems are available (e.g., BlueComm, www.sonardyne.com).

A complementary challenge in underwater communications is achieving remote access to underwater platforms "from the office" (see next section). Moored buoys can be used to monitor oceanic and near-

Fig. 6. Norway lobster photomosaic. Segment of a photomosaic of a two-sided video transect during the Norway lobster (*Nephrops norvegicus*) 2018 stock assessment surveys in the Kattegat area (~100 m depth; 57.86706–57.86694°N; 9.79335–9.79373°E) by the ICES Working Group on *Nephrops* (WGNEPS). The date, hour, and frames stitched to create each part of the mosaic are provided in the upper part of the figure (right-hand side in the full vertical mosaic). The two green lines (70 cm in distance) depict the useful area for data extraction. This approach allows for the geospatial and temporal referencing of biological data from video surveys (e.g., number of individuals and their burrow entrances) with all other data within a DTO database. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

surface atmospheric conditions and facilitate data transmission (e.g., Bahamon et al., 2011). However, these moored systems face limitations such as deployment complexity, maintenance requirements, and compliance with standards and governmental regulations. To obtain reliable and real-time data for tracking and responding to rapidly evolving environments, the use of Unmanned Surface Vehicles (USVs) is becoming increasingly popular (Zhang et al., 2021). USVs can be rapidly deployed and operate at sea for several weeks at very low operational costs, offering a clear advantage over using crewed ships for monitoring deployed instruments. Additionally, innovative pop-up buoys provide complementary information. These buoys can be connected to landers and released to the surface without any surface infrastructure, sending information via satellite/GSM communications (Carandell et al., 2023). These buoys can last more than three months, acting as Lagrangian platforms for oceanographic purposes and transmitting their positions for recovery. This is especially useful in areas where, due to weather or security reasons, USVs are not a feasible option.

2.3. The control Centre for platform navigation and data collection

A DTO framework for ecological monitoring encompasses not only on simulations of ecosystem dynamics but also on the PTO, which provides input data. A remote-control centre is valuable for the adaptive modification of platform operations on the basis of simulations established considering energy autonomy constraints in relation to the quality of the acquired data. A control centre can streamline in situ data transmission and mission planning and coordinate data collection remotely. Mission changes can be made on the basis of real-time data with a mission planning tool, such as the planning domain definition language (PDDL) (Kootbally et al., 2015). Within such a framework, the mission plan for each robotic platform can be generated on the basis of the mission tasks and goals specified by the user and monitored by an operator through a graphical interface. After checking the generated plan, the operator can confirm it or alter it. Operators can record mission execution information before deployment and assess the status of each step as “not completed”, “in progress” or “completed”. Another functionality of the control centre is the simulation of F-PTO platform operations. The control architectures of the newest robotic platforms and prototypes are based on the Robotic Operating System (ROS) (Macenski et al., 2022; Quigley, 2009), an open-source software tool for control, simulation, and planning. The mission goals for autonomous vehicles are routine data collection, surveying specific areas to confirm DTO predictions, and implementing reactive navigation schemes on the basis of environmental variables. It is crucial to update planning algorithms’ knowledgebases with DTO predictions. Finally, deploying a ROS-based architecture with nonrobotic F-PTO nodes (e.g., buoys, cabled observatories and landers) may improve intraplatform communication and knowledgebase updates and enhance assessments of DTO planning outcomes.

3. AI in image processing for extracting biological data

A critical portion of the ecological data input into DTOs currently comes from imaging applications. Image-based methods for animal detection, classification, and tracking (see e.g., Durden et al., 2016) are relevant for the automated generation of biological data by both F-PTO fixed infrastructures (e.g., cabled observatories and landers; Fig. 2) and mobile platforms that deliver seabed mapping products (e.g., photomosaics; see section 4). Deep learning (DL) and convolutional neural networks (CNNs) have already proven useful in many different ecological scenarios for partially or fully automatically classifying specific marine animal species from large images and video sources (Cline et al., 2009; Cuvelier et al., 2024; Garcia et al., 2020; Möller and Nattkemper, 2021; Purser et al., 2009; Salman et al., 2016; Schoening et al., 2022; Villon et al., 2018, 2020; Xu and Matzner, 2018; Zurowietz et al., 2018).

Relevant consideration in image transference and processing can be described in the case of the OBSEA network of ecological monitoring (see Table 1). Images are collected by the observatory and transferred for real-time processing through AI and machine learning approaches in shore centralized facilities (see Table 1). The computational detection and classification of species (e.g., fishes) requires an average time 0.1 s/image for a 10 megapixels image (i.e., full HD) or video frame, considering a standard neural architecture as YOLO. An example of this network for OBSEA can be seen in Baños Castelló et al. (2025), where the CNN was trained with images from 23 different classes. In case of larger images (e.g., 20 megapixel), the processing would occur in tiles of 800 × 800 pixel and processing time would be N x 0.1 s. If other more sophisticated architectures are used, than the processing time can increase to 1–10s. For all other non-cabled platforms (i.e., mobile of fixed; see Fig. 1), edge-computing functionalities in onboard image treatment would be required and the processing time and energy requirements will depend upon resolution, tasks (animals’ detection and classification or mapping), and available processing hardware (GPU power and number of units available; Ortenzi et al., 2024).Automated biological classification models can be trained with archived imagery data and can be periodically retrained or adapted to site-specific knowledge as new PTO data are input into DTOs (e.g., Ottaviani et al., 2022). In essence, imagery characterizing target communities or populations is needed (Durden et al., 2024), and it can be sourced from available repositories (e.g., Zenodo and datasets presented in Garcia et al., 2020), along with associated 'real-world conditions' recorded during image acquisition or metadata (depth, environmental light levels, changing seabed background and turbidity, among others). Marine imaging data typologies, collection methods, and processing techniques can be highly diverse (e.g., Durden et al., 2016), and the application of each of these methods for automated classification can vary from small-scale approaches (e.g., for model training) to extensive manual tagging efforts (Marini et al., 2018; Ottaviani et al., 2022). Online initiatives are a step toward overcoming data processing issues since they can facilitate interoperability across institutions and quality control and data sharing methods, increasing the image analysis capability and use efficiency of CNNs.

Server-stored or locally stored videos and images can be readily annotated with available online software platforms such as BIIGLE (equipped with computer vision tools such as DELPHI and MAIA) (Fig. 3), FathomNet or RoboFlow (Katija et al., 2022; Langenkämper et al., 2017; Schoening et al., 2015; Zurowietz et al., 2018; Zurowietz and Nattkemper, 2020). In biological analyses, the term "annotation" refers to (1) detection, the presence of an animal in a given location (a point, line or circle in a georeferenced image or video frame), and (2) classification, the assignment of an animal to a semantic or taxonomic category (e.g., species) selected from a standardized image-based catalogue (a species list; see e.g., Simon-Lledó et al., 2023). Since the level of taxonomic precision that can be achieved using imagery can vary (e.g., morphotypes, from species to family or order level) depending on the collection methods (e.g., the platform characteristics, camera settings, or resolution) and across animal groups (e.g., requiring physical examination for species determination), the use of an open taxonomic nomenclature for image-based identification (aligned to the World Register of Marine Species; e.g., Horton et al., 2021) can improve the reliability of biological datasets and their usability and comparability across DTOs. In both detection and classification, AI has played a vital and growing role, as the volume of recorded data has, in most instances, already exceeded the volume at which it can be effectively processed by human domain experts (marine biologists or engineers/inspectors).

Grafana is an interactive web application for data visualization that can be used to plot time series of detected species counts in real time via image processing on the basis of a trained machine learning algorithm for animal classification and counting (Fig. 4). Once the camera source and a particular species have been selected, consecutive AI-processed images can be visualized along with time series of counts. Time series visualization can be performed for multiple species at once, providing a clear representation of dynamics.

4. Big data management with data bubbles

Effective data management is a critical component in the development of DTOs, particularly when handling high-frequency, multi-modal data streams originating from heterogeneous sensor platforms, such as fixed-local (F-PTOs) and geographically expanded physical twins observers (G-PTOs). These platforms generate large volumes of structured and unstructured data in various native formats (e.g., ASCII, BIN, NMEA, JPEG, WAV). To ensure interoperability, long-term usability, and seamless integration into modelling and visualization systems, these data streams are harmonized and standardized across institutions and platforms.

Structured datasets are converted into the Network Common Data Form (NetCDF) format (Rew and Davis, 1990), following the widely adopted Climate and Forecast (CF) metadata conventions (Eaton et al., 2023). These standards enforce consistency in variable naming, units, coordinate reference systems, and temporal structures. Depending on the characteristics of the data, the appropriate CF feature type is assigned:

- Point – Instantaneous, discrete observations (e.g., single sensor readings)
- TimeSeries – Repeated measurements at a fixed location (e.g., moored sensors)
- Profile – Vertical sampling at a single time point (e.g., CTD casts)
- Trajectory – Time-stamped data along a horizontal path (e.g., drifters)
- TimeSeriesProfile – Repeated vertical profiles at a fixed site (e.g., moored profilers)
- TrajectoryProfile – Vertical profiles along a moving platform path (e.g., gliders)
- Grid – Spatially continuous data gridded over a region (e.g., remote sensing or model outputs)

Each NetCDF file is enriched with metadata describing deployment context, sensor specifications, calibration parameters, quality control flags, and standardized attributes, conforming to INSPIRE and SeaDataNet guidelines. This enables semantic interoperability and enhances machine readability. Standardized datasets are made accessible through an ERDDAP server Simons, 2017), which supports RESTful and OPeNDAP API access, facilitating remote data retrieval, real-time integration, and downstream applications in forecasting or DTO visualization frameworks.

4.1. The application of data bubbles to merge multiparametric data

A DTO should be able to acquire, store and process multiparametric biological and environmental data in a spatiotemporally collated fashion. The diversity of sensor types and data formats calls for formal frameworks that allow datasets to be interoperable and machine-readable. Semantic interoperability among heterogeneous sensing systems is essential for the effective integration and interpretation of ecological big data. In this context, dedicated ontologies (e.g., BiodivOnto; Abdelmageed et al., 2021), can play a central role by providing shared vocabularies and defining relationships among biological and environmental variables collected in different areas at different time. This creates the need for a spatiotemporal sampling unit (referred to as a data bubble), which can be treated as a semi-independent DTO element used to merge data (i.e., standardize and homogenize data). The term "data bubbles" is not standard in oceanographic literature, which more commonly refers to "data cubes" for a posteriori organizing large-scale spatiotemporal datasets (e.g., Montero et al., 2024). While data cubes are well suited for post-hoc integration and retrospective analysis across broad spatial domains, we introduce "data bubbles" as a novel and complementary concept tailored to in-situ ecological monitoring. Data bubbles represent localized, multi-parametric datasets collected dynamically by monitoring platforms within a known radius their immediate environment, effectively capturing information in a 360-degree horizontal and vertical sphere. Each bubble encapsulates high-resolution physical, chemical, biological, and behavioural data, rooted in the real-time context of ecological processes.

A data bubble (Fig. 5) is a geospatially defined location (including depth), within which all relevant biological and environmental data are collated and stored with appropriate time stamps. The operational metadata labels associated with each type of biological and environmental information can be obtained via the precise geo-referencing of platform positions with active acoustic communication tools (reviewed by Aguzzi et al., 2024).

While fixed platforms such as cabled observatories and landers provide synchronous biological and environmental data (e.g., counts of individuals of different species) within the same imaged seascape of a few cubic metres (e.g., HD and multibeam acoustic imaging can reach a distance from 2 to 3 to 12–15 m of field projection) at minute time

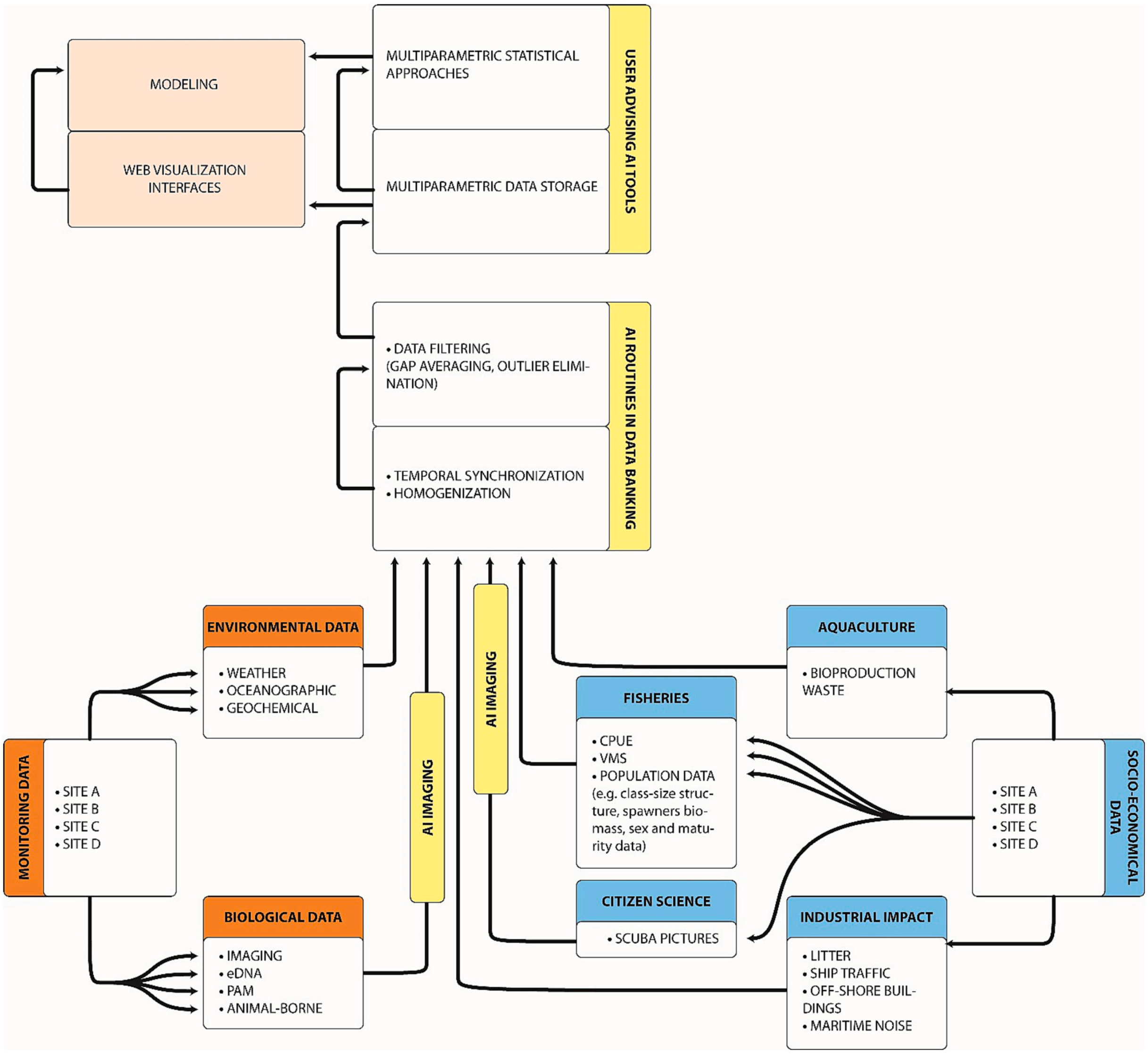

Fig. 7. DTO flow chart for relevant information processing. Ecological information flow processing (bottom-up) from the F-PTO (on the left) and G-PTO (on the right, with different types of societal data) at the core of the DTO structure. Data bank activities associated with pre-processing, storage, processing, visualization, and analysis are based on AI functionalities. CPUE is the catch per unit effort, and VMS is the vessel monitoring system.

intervals, mobile platforms can generate image-based data for ecosystems ranging from a few hundred metres to tens of kilometres (Aguzzi et al., 2019). Thus, biological data can be extracted by mapping the occurrence and detailed positions of megafaunal samples per unit of video-captured surface in created photomosaics to derive, for example, the spatial range or standing stock of an animal (Fig. 6). Since mobile platforms provide temporally continuous imaging products, the obtained biological information can be processed into standardized time–lapse intervals for comparison with data from other platforms (e.g., by subdividing transects into subsections of video-captured seabed surfaces and by estimating species densities at certain intervals (e.g., minutes) (Chatzievangelou et al., 2020).

The georeferenced and time-stamped biological and environmental data collected by the F-PTO and G-PTO platforms can be synchronized with other historical and up-to-date oceanographic and biogeochemical data from both public and private sources (see section 4.2). In this process, datasets provided by societal actors should have the same metadata labels and nomenclature typology as the PTO data, including a range of essential data collection parameters, ranging from the time and exact positioning of an observation to the sampling platform type and the measuring sensor/tool used (as described in Fig. 5).

4.2. The architecture for data storage and management

A DTO requires an infrastructure for data storage and management, as well as computational tools, to automatically compose data workflows on the basis of ecological information. That architecture should be built with a bottom-up approach. The base layer should be able to process a very diverse range of biological and environmental data from multiple sources, such as in situ PTOs (with their platforms and sensors; see Fig. 1), as well as any other locally collected data. Then, workflows capable of automated data processing should be embedded for the computation of ecological metrics, such as indicators (Aguzzi et al., 2019, 2020), yielding new scientific knowledge for a diversified class of end-users, spanning from scientists to citizens and stakeholders.

The workflow for data treatment, which influences the final DTO architecture, is presented in Fig. 7. A central repository should be created to store all biological (including image-based), oceanographic and geochemical data with associated metadata (see the data bubble labelling requirements in Fig. 5). The data bubbles associated with PTO platforms need to be temporarily stored with other previously and continuously obtained similar information. Example of such information are the catch per unit effort of nearby commercial fishing boats,

Table 2
Overview of data characteristics and standardization strategies for representative sensing platforms integrated into the DTO framework. File formats, data structures, sampling rates, daily volumes, and harmonization methods (e.g., NetCDF feature types and metadata standards) are based on existing deployments, including OBSEA (www.obsea.es), Ocean Network Canada (Purser et al., 2013), and Sentinel/Planet satellite data services. Posidonia mapping formats and rates reflect Sentinel-2 (Copernicus) and PlanetScope (Planet Labs) usage.

Sensor type	Measured variables	File format (s)	Data structure	Sampling rate	Volume/ day	Storage requirement (GB/day)	Standardization format
HD Camera (stills)	Images (Biodiversity, Richness)	JPEG	Unstructured	1 image/min (1440/day)	~720 MB	0.72 GB	NetCDF (metadata + AI detections)
HD Camera (videos)	Short videos (faunal behaviour)	MP4 (H.264)	Unstructured	1 video/min (10s, 1440/day)	~14,400 MB	14.4 GB	NetCDF (metadata + AI detections)
CTD	Temperature, Salinity, Depth	ASCII, BIN	Structured	1 Hz	~2 MB	0.002 GB	NetCDF (Point, TimeSeries)
ADCP	Currents, Wave height/ direction	BIN	Structured	1 Hz	~5 MB	0.005 GB	NetCDF (TimeSeries)
FTU (Chlorophyll/ Turbidity)	Fluorescence, Turbidity	ASCII, CSV	Structured	Every 5 min	~1–2 MB	0.001–0.002 GB	NetCDF (TimeSeries)
PAM Hydrophone	Marine noise, species calls	FLAC	Unstructured	Continuous @96 kHz, 16-bit	Raw: ~15.5 GB	~8 GB	NetCDF (metadata + FLAC)
Oxygen Sensor (O_2)	Dissolved oxygen	ASCII, CSV	Structured	1 reading/5 min	~1 MB	0.001 GB	NetCDF (TimeSeries)
Particle Camera	Sediment & particle sizes	JPEG, CSV	Mixed	~200 images/day	~2 GB	2 GB	NetCDF (metadata + imagery)
Seafloor Crawler	HD Video, CTD, Mapping	MP4, BIN, ASCII	Mixed	Continuous mission (8 h/day)	~5–10 GB	5–10 GB	NetCDF (TrajectoryProfile)
Gliders / AUVs	CTD, Video, Currents, Mapping	BIN, ASCII, MP4	Mixed	Mission-dependent	0.5–2 GB	0.5–2 GB	NetCDF (TrajectoryProfile)
Satellite	Habitat cover (e.g., Posidonia), RGB/NIR	GeoTIFF (.tif)	Gridded Raster	5–10 scenes/day (3–5 m res.)	~0.5–2 GB	0.5–2 GB	NetCDF (Grid)
Satellite Remote Sensing	SST, Altimetry, Ocean colour (global)	NetCDF, HDF5	Gridded	Daily pass (~4–8/ day)	~1–10 GB	1–10 GB	NetCDF (Grid)

information for commercial species size classes, sex and maturity, as obtained via government- or EU-funded monitoring programs (e.g., the Data Collection Framework), or the results of citizen science (e.g., SCUBA diving images and information from recreational fishing activities).

Acquiring a high volume of data from various F-PTO and G-PTO components, at times with latency, and harmonizing them over time are notable challenges. To illustrate the diversity of data sources, formats, and standardization strategies across the monitoring infrastructure, Table 2 summarizes representative sensor types and platforms integrated into the DTO framework. The table includes key information such as measured variables, native file formats, underlying data structures, sampling rates, and typical daily data volumes. Additionally, it outlines the harmonization formats applied during data processing—such as NetCDF feature types and CF-compliant metadata standards—used to ensure semantic interoperability and long-term usability. This overview highlights the heterogeneity of the data ecosystem and the necessity of a robust standardization pipeline to support integration into downstream analysis, modelling, and visualization systems.

These data-specific standardization strategies constitute the foundational layer of our DTO architecture, enabling scalable and harmonized data ingestion pipelines that support downstream AI-driven analysis as well as spatiotemporal modelling workflows. In our framework, AI routines embedded in a central DTO data infrastructure are not only used for basic tasks such as data cleaning, temporal alignment, and outlier removal, but are also central to more advanced processes like predictive modelling and inferential gap filling. We particularly support the view that probabilistic approaches are essential when direct observations are incomplete or sporadic (Price et al., 2025). For example, spatiotemporal models informed by time series of species detections and associated environmental parameters allow for the forecasting of likely presence or abundance in unsampled intervals or locations. These predictions can be strengthened by leveraging data from ecologically analogous areas, using transfer learning or environmental similarity metrics. This is not merely a workaround for data limitations, but a functional feature of the DTO, which aims to produce actionable ecological indicators even in data-sparse contexts.

The application of FAIR (Findability, Accessibility, Interoperability, and Reusability) data principles can facilitate data retrieval by the scientific community (Wilkinson et al., 2016). To ensure sustainable multimodal data collection, integration, and interpretation, metadata standards such as iFDO (Schoening et al., 2022), DublinCore (Weibel and Koch, 2000), Audubon (Morris et al., 2013), SmartarID (Howell et al., 2019), and PDS4 (Hughes et al., 2014) have been created and applied.

Finally, the DTO architecture should be presented to the public via web visualization interfaces as interactive windows to visualize all types of biological and environmental information. Synthetic graphic outputs from different multivariate statistics and time series analysis approaches can be selected and displayed for different combinations of data and spatiotemporal frameworks (Aguzzi et al., 2020).

The storage, open-access querying, and downloading of data from PTO platforms should be managed by an associated Application Programming Interface (API) (Martínez et al., 2023) that is connected to other portals and applications. For organizations that carry out specific experiments or target surveys and host their own data, metadata harvesting mechanisms (using standard APIs) can be established, facilitating the discovery of datasets stored at distributed nodes via a search of the central catalogue. This interconnection capability can be upgraded by employing the SensorThings API to provide an Open Geospatial Consortium (OGC) standard (Liang et al., 2021), resulting in an open and unified framework to interconnect different platforms and sensors, data, and applications via the SensorThings API.

4.3. Graphical user interfaces

A DTO offers worldwide end-users the opportunity to investigate ecological processes through Graphical User Interfaces (GUIs) as central parts of web visualization interfaces (Fig. 8). These interfaces should support the visualization of complex ecological (biological and environmental) information within a spatially simulated representation of the monitored marine environment, where the trends of changes in different biological and environmental variables can be assessed. The graphic outputs used should highlight significant trends for ecological components in a given space, supported by the relevant AI tools for data treatment, sensor integration, and data banking (Chen et al., 2023).

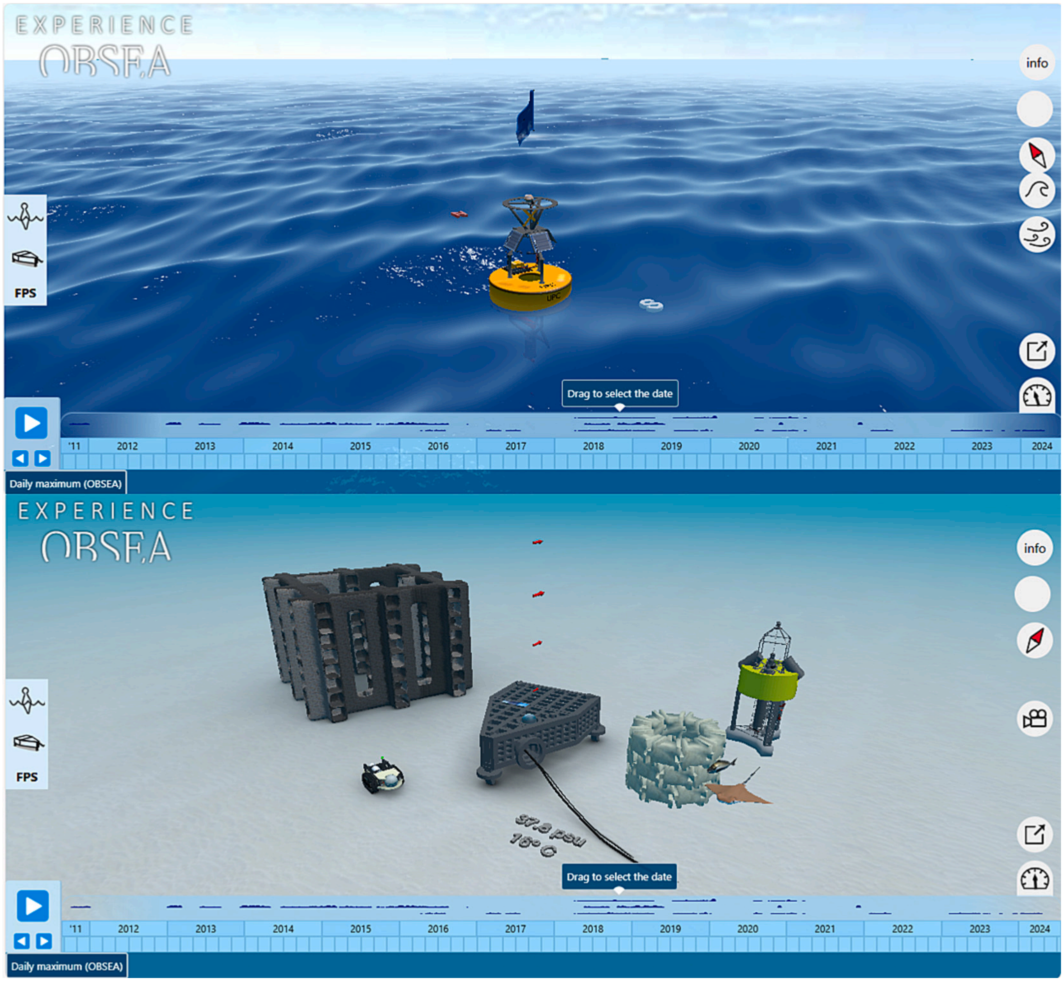

Fig. 8. OBSEA visualization interface as GUI. Web visualization interface for the OBSEA F-PTO, incorporating a surface meteorological buoy, a seabed-cabled multiparametric video observatory, a docked video crawler in front of an artificial reef, namely, a slag reef of recycled metal compounds, and a satellite lander (link to the GUI https://cgi-dto.github.io/OBSEA/ and link to the source code of the GUI https://github.com/BlueNetCat/OBSEA). This 3D GUI allows the spatial representation of the dynamics of ecological variables on the basis of the selection of time windows for different time series of data in the bottom panel (i.e., the columns with red arrows indicate the current direction components). Descriptions of the crawler and OBSEA platforms were presented by Del-Rio et al. (2020) and Falahzadeh et al. (2023). While the current interface primarily displays physical oceanographic parameters such as sea temperature, salinity, wave height, wave direction, and air temperature, it forms a foundational layer of the broader DTO framework. The integration of ecological variables such as species presence, abundance, and activity, is currently under development, supported by the implementation of AI-driven image and acoustic analysis pipelines. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

GUIs can be developed to visualize trends in environmental parameters by selecting various data time series collected over different temporal windows (Llorach-Tó et al., 2023) (see Fig. 8). For example, the occupancy of each spatial subunit by individuals of different species can be monitored and modelled. Although this approach is relatively straightforward in 2D space (e.g., the heatmap approach of Chatzievangelou et al., 2020 and Doya et al., 2016), 3D benthopelagic habitats can make this process challenging. Multiple imaging tools operating simultaneously and focusing on the same areas can provide information to help assess not only the presence but also the trajectory of megafauna.

5. DTO framework for ecosystem monitoring via ecological indicators and socioecological modelling

5.1. Computation of ecological indicators

Ecological indicators are at the centre of current management and conservation policy frameworks (Danovaro et al., 2020; Miloslavich et al., 2018) as per the information presented in Table 1. A description of how to implement a DTO based on a combination of different F-PTO and G-PTO platforms (and their sensors) is presented in Table 3 in relation to three different categories of ecological monitoring indicators and their objectives.

5.1.1. Biodiversity in data bubbles

The measurement of marine biodiversity is strategically important for conservation and management policies (Duarte et al., 2020; Mokany et al., 2020) and is thus a key aspect of DTO functionality. This functionality is based on several key operational assumptions for both assessing the present status of DTOs and modelling changes at different spatiotemporal scales.

First, F-PTO and G-PTO platforms can support data collection over different temporal scales, allowing the precise characterization of the ecological niches of species. The abiotic niches of individuals are represented by “imaged spaces” (i.e., data bubbles; see Fig. 5), with habitat features that are synchronously monitored by oceanographic and geochemical sensors installed near platform cameras (Aguzzi et al., 2020). Increasing the number of sensors utilized by platforms in a given area can enhance the precision of niche delineation. This precision is fundamental to analyses and modelling (Fig. 9): locally acquired

Table 3
Ecological indicators. The different ecological indicators calculated with data collected by F-PTOs (cabled observatories, satellite landers, crawlers, moored buoys, and docked AUVs for surveying, as described in Fig. 1, plus sensor payloads), by G-PTOs (ROVs, towed sledges and drop cams), by shore-deployed ASVs with coupled AUVs and drones, are combined with various types of data collected by the public to facilitate DTO monitoring and enhance the modelling capacity. PAM is passive acoustic monitoring; ADCP is an acoustic Doppler current profiler; CTD is conductivity, temperature and depth; and eDNA is environmental DNA.

Ecological indicator	Description	Monitoring objective	PTO components	Ecological monitoring strategy	How to implement the DTO
Temporal biodiversity	Variations in abundance, richness and evenness of mobile megafauna over 24-h and seasonal scales	Understand the role of community turnover based on activity rhythms and their influences on species abundance and community composition and	F-PTOs	Day-night time-lapse optoacoustic imaging (species counts fluctuations and co-presence as proxy of food web interactions), plus eDNA and PAM sampling to enhance richness determination	Build multiparametric data repositories where species abundance and biomass data from PTOs in nearby areas are temporally collated with equivalent institutional and governmental actions (e.g., research cruise programs), entries from logbooks of local fishery associations and information from SCUBA diving centres (imaging material and citizen science observations)
Spatial biodiversity	Variations in abundance, richness and evenness of mobile megafauna across spatial intervals of ecological heterogeneity (α, β, and γ biodiversity)	Understanding the effects of geomorphology and oceanography on the species distribution, incrementing monitoring and modelling the predictive capacity of	G-PTOs	Kilometer-scale video-transecting and photo mosaicking plus eDNA sampling to enhance richness and biomass determination on the basis of local F-PTOs	
Connectivity	Flux of gametes, larvae, and adults at different levels of development that contribute to the local demography and genetic diversity	Estimate the contribution from the transformation of individuals at different levels of development to the local demography and genetic diversity		Larval, DNA, and eDNA sampling across depth strata plus data collection with animal-borne data loggers including physical (e.g., plastic) and acoustic tagging	Identify geographic corridors for the dispersal of individuals across variable seascapes using global animal tagging/tracking data (e.g., the acoustics solutions obtained by the Ocean Tracking Network)

knowledge of species presence and abundance can be inferred for any other area for which only some in situ data on seascape conditions are available (e.g., discrete or gridded data stored in EMODNet). This implies the capacity to relate the status of biological data with the status of habitat variables and identify conditioning drivers (e.g., via multivariate statistics) to be used for further inference.

Second, F- and G-PTO data can be used to interpret the temporal trends for species in historical datasets at large geographic ranges (with different types of archived data). For motile species, for example, time series of species counts, with peaks and troughs at 24-h and seasonal scales, as recorded by cabled observatories and landers, can be compared to concomitant changes in species densities over larger surrounding areas derived via ROV and AUV direct inspection and even from the reports of commercial fishing operations (e.g., the established case of the Norway lobster as seen in Aguzzi et al., 2022).

On the basis of both assumptions, the fusion of real-time monitoring data with those historically archived for the same PTO zones would provide a homogenized and detailed dataset for the spatiotemporal modelling of biodiversity on the basis of the detection of cause–effect relationships for species and community responses (abundance fluctuations) under changing environmental conditions. From a perspective encompassing species assemblages and biodiversity, the Joint Species Distribution Model (JSDM) (Franklin, 2023) is a valuable tool for interpreting geographic patterns of biodiversity by correlating known species occurrence or abundance records with environmental conditions (Warton et al., 2015; Wilkinson et al., 2021). The modelling strategy, along with multivariate statistical approaches, can be used to identify functional relationships among biological and environmental parameters, hence identifying key drivers of the observed conditions and supporting forecasting methods for diverse scenarios. The JSDM supports the upscaling of environmental and occurrence data after the adequate treatment of boundary predictors (Meynard et al., 2023; Wilkinson et al., 2021) beyond the data bubbles associated with a DTO.

5.1.2. The DTO for marine functional connectivity

Marine functional connectivity refers to the flux of individuals (and their genes) at any level of development (i.e., from gametes to adult stages) across three-dimensional spatial scales (Darnaude et al., 2024). For highly motile organisms, measuring connectivity at the DTO scale is challenging, as it requires tracking specific individuals via electronic devices (Espinoza et al., 2015), intrinsic biomarkers or photo-identification processes (Ferreira et al., 2021). For low-motility organisms, oceanographic factors play pivotal roles in dispersing individuals in weak pelagic early-life stages away from their natal origins. Assessing the connectivity of these species can be achieved through an ecological modelling approach on the basis of an individual-based model (IBM). In IBMs, oceanographic data are used to estimate the potential dispersion pathways of organisms, integrating ecological parameters such as spawning times, larval durations, and vertical displacement (Clavel-Henry et al., 2020; Fobert et al., 2019; Matos et al., 2024; Sciascia et al., 2022).

DTOs should encompass in situ hydrodynamics and gridded oceanographic information from external datasets (e.g., COPERNICUS), thus providing minimum inputs for running the applications associated with an IBM. These applications can be explored to obtain connectivity insights regarding DTO through the utilization of metrics and indicators derived from an approach adapted within the IBM context and considering the IBM objectives. For example, these approaches involve simple dispersal metrics (e.g., the number of particles arriving at a DTO and the dispersed distance from a DTO), metrics related to connectivity between sites (e.g., links between deployed PTO sites) via a graph theory approach (Treml et al., 2008), and indicators of temporal variability regarding connectivity between sites (Clavel-Henry et al., 2024).

The real-time and forecasted hydrodynamic products provided through a DTO can be used to explore development in the computing of operational connectivity; e.g., automating the estimates of particle

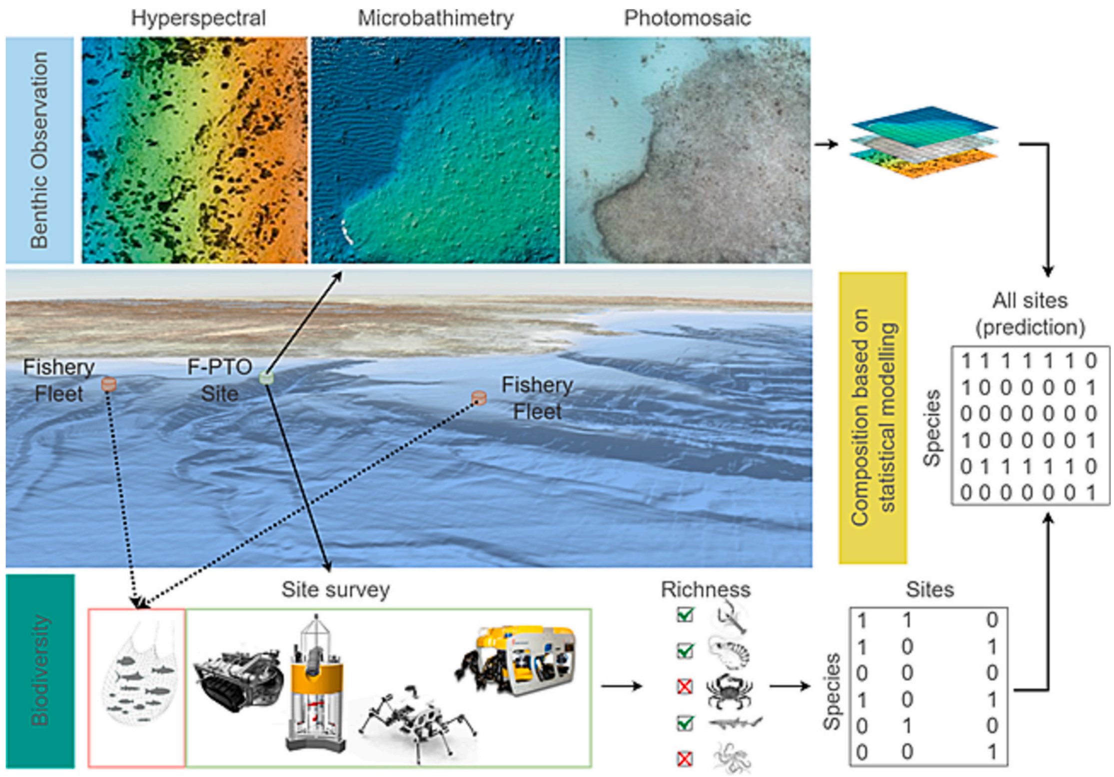

Fig. 9. Site-specific ecological knowledge obtained via multiparametric monitoring, which is spatially scaled for predictions of species distributions and abundances. The presence and abundance of benthic species can be tracked over time at any monitored and mapped site (by F-PTOs, via the combination of different imaging approaches). The results can be compared with similar data from nearby areas (e.g., fishery-dependent trawl data, with commercial vessels being a part of the G-PTO framework). Predictions of species presence and abundance can then be formulated for any other non-surveyed area for which only seascape data are available.

transport on the basis of the most recent hydrodynamic data if the simulated period of the trajectories can be assessed. This approach, rooted in numerical modelling, can also be expanded to other key processes for conservation and management, such as the dispersion of contaminants (Keramea et al., 2022) and sediment clouds due to trawling and mining (Weaver et al., 2022). Recent studies have substantiated the efficacy of such approaches (Andruszkiewicz et al., 2019; Córdova and Flores, 2022; Payo-Payo et al., 2017), highlighting their importance in advancing environmental monitoring and response capabilities. Moreover, adopting an operational approach would allow DTO managers to swiftly respond to authorities during emergencies such as human rescues, ship drifts, and oil spills within the scope of a given (Keramea et al., 2022; Pereiro et al., 2021).

Ecological modelling on the basis of F-PTO and G-PTO data allows the visualization and prediction of species distributions and dispersal in the ocean. Modelling methods can be added to DTO frameworks, assuming that ecological models such as IBMs or the JSDM (see section 5.1.1) have been established (Fig. 10). Moreover, ecological predictions

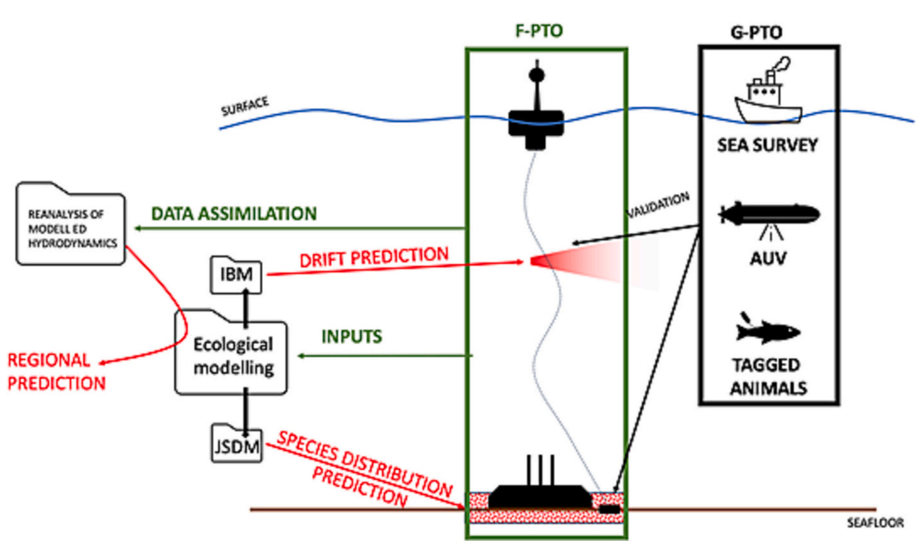

Fig. 10. Ecological modelling flow chart. The ecological modelling approach is based on an individual-based model (IBM, for larval transport) and, to a lesser extent, a JSDM. F-PTO and G-PTO data are used as follows: data assimilation for modelling local marine ecosystem features, assessing the deployment of mobile observatory tools, and validating model outputs.

can be produced at a local scale (i.e., at the site of real-time measurements) or at a regional scale (i.e., when PTO data are assimilated into oceanographic models) and can be directly validated with the data collected at observatories (see Fig. 1). The overall output can be informative for the initiation of seawater sampling activities, whether autonomously or mechanically implemented, and for enhancing model reliability (see Table 2).

5.2. Socioecological modelling

A holistic DTO requires the implementation of socioecological approaches and models to capture the complexity of the interactions between nature and people. The social dimension of socioecological systems includes a diversity of actors, institutions, cultures, and economies, whereas the ecological dimension includes the relevant species and the ecosystems they inhabit (IPBES, 2024).

Socioecological systems typically consist of three inextricable dimensions: (1) the environment, (2) society, and (3) the economy. Each dimension interacts in many ways over time based on a hierarchical dimensional scale, now widely accepted in the holistic approach of ecological economics (Daly and Farley, 2011). This means that in the socioecological approach, the interactions between human behaviour and the environment are considered part of a complex socioecological system.

The goal of introducing a socioecological approach and corresponding modelling into a DTO is to identify the socioeconomic components of marine and coastal systems in relation to different ecological and environmental components (Davies et al., 2016). A socioecological approach requires the development of a conceptual model in which the different components can interact, potential interactions (positive, neutral, or negative) can be identified and the weights of these components can be determined. The interactions can produce emergent properties, and the aggregate result may differ from what it would be if each agent were isolated from others.

One of the tools used to develop socioecological models is the Agent-Based Model (ABM), which is a computational model populated by many heterogeneous agents independently interacting with each other (a bottom-up approach, without a coordinator), without feedback or externalities (Caiani et al., 2016). The strength of an ABM is its ability to assess emergent patterns that result from the dynamic behaviour, adaptation, and learning of each individual independently. As such, questions about how a system would react to a certain stimulation are explored, as are the corresponding outcomes, and new hypotheses regarding the functioning of a system for which not all data are available can be investigated (Heckbert et al., 2010). ABMs have been widely applied in economics (Gallegati et al., 2017; Hamill and Gilbert, 2015) but are also gaining popularity in ecological assessments, land-use decision-making (Matthews et al., 2007), fishery management (Lindkvist et al., 2020; Moran et al., 2021), pollutant emission assessments (Ghazi et al., 2014; Newth and Gunasekera, 2012), and natural resource management (Loomis et al., 2009), among other fields.

Integrating the socioecological approach with a DTO can significantly enhance the understanding of the marine environment and its impacts on society, and vice versa, in various ways: (1) models can be used to assimilate the data obtained during fishing activities and fishing catch observations with microdata collected from the fishery community to consequently inform policies and strategies related to marine conservation, sustainable use of fishing resources, ecosystem restoration, the rebuilding of depleted stocks, and planning and management strategies at multiple spatial scales; (2) multivariate statistical analyses can be performed to answer questions about the efficiency and benefits of integrated coastal and marine management actions related to the mitigation of anthropogenic impacts on ecosystems (caused by overfishing, pollution, or habitat destruction) or, conversely, the impacts of changing ecosystems on human life; and (3) fluctuations in fish populations can be predicted, and direct impacts on the fishing industry and coastal communities can be determined.

The integration of socioeconomic data, gathered at the microscale via ABMs, into a DTO can be useful for the development of recommendations and guidelines in the context of various objectives:

- Mapping human activities and ocean use patterns: Socioeconomic data can be used to map human activities in the ocean. Shipping, fisheries, marine tourism, resource extraction, and coastal infrastructure are some of the sectors that can be investigated. Data may include information regarding fleet sizes, the geographic distribution of activities, and changes over time. The environmental data modelled by a DTO can be combined with socioeconomic data to understand how changes in ocean conditions affect local and regional economies.
- Increasing the forecasting ability of a DTO through the use of regional oceanographic, environmental, and biological/fishery data can be accomplished via the mapping of areas with F-PTOs and analyses of spatial and temporal biological community structures in addition to environmental characteristics. This process relies on available databases with information regarding marine environmental conditions from remote (e.g., satellites) and in situ (e.g., moored buoys and ship-borne sensors) observations.

The resulting data can be used within a DTO framework to construct

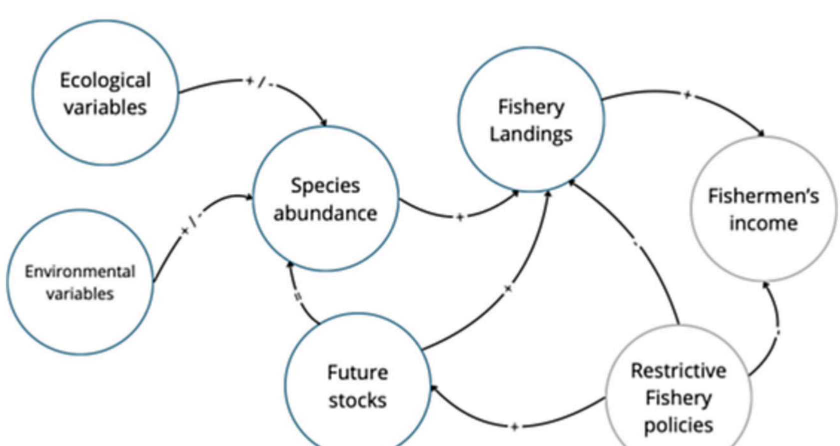

Fig. 11. Causal loop diagram of fishery socio-ecological dynamics and interactions. Arrows indicate causal relationships between variables, with their respective effects (positive, negative, or neutral).

an operating model that projects the future dynamics of marine ecosystems to assess the possible impacts on society under different management scenarios; e.g., a reduction in fish catches could result in reduced disposable income for households relying on the fishery economy.

While the integration of socioecological models into DTOs offers significant potential benefits, the actual implementation process faces substantial challenges. These include limited access to comprehensive socioeconomic datasets across different geographical regions, heterogeneity of information across different fisheries categories, and discrepancies between monitoring observations and fisheries data (Addison et al., 2017; Leenhardt et al., 2015; Saunders et al., 2015). These data integration issues can affect consistent estimation of socioecological impacts. However, once socioecological models are established and validated with appropriate data inputs, their technical integration into DTOs should follow similar protocols as other types of models (species population, distribution, connectivity and hydrodynamic models), albeit with necessary adjustments for human dimension components. Fig. 11 illustrates a causal loop diagram depicting the complex relationships between ecological variables, fishery activities, and socioeconomic impacts.

Implementation will likely follow an iterative process (Gallegati et al., 2024), beginning with data collection and integration from demonstration sites. These sites will serve as test cases for developing methodologies to bridge the gap between environmental and socioeconomic domains, with continuous refinement based on stakeholder feedback and evolving technical capabilities. The development of socioecological DTOs therefore represents not only a technical data integration challenge, but a complex socio-technical effort that requires careful consideration of human behaviour, policy and governance structures along with technical implementation aspects that will be addressed by systems engineers and data scientists.

6. Conclusion

This study outlines a comprehensive and forward-looking strategy for marine ecosystem monitoring based on the development of a Digital Twin of the Ocean (DTO). Current robotic platforms, such as those we describe, can generate high-resolution data, however, challenges remain in achieving adequate spatial coverage, temporal continuity, and ecological representativeness. One of the key messages of our manuscript is that technological capability alone is not sufficient. It must be coupled with adaptive monitoring strategies, data integration protocols, and modelling approaches to effectively address the complexity of marine ecosystems. We therefore do not claim that existing systems are universally sufficient, but rather that they represent a foundational step toward more robust and scalable monitoring frameworks. A concise Strengths, Weaknesses, Opportunities, Threats (SWOT) analysis can be proposed as follows. Strengths of the elaborated DTO concept rely on its highly innovative and integrative approach for data collection grounded in robotic platforms and AI-driven image analysis. Its strength lies in the interdisciplinary and modular design, including fixed and mobile robotic monitoring assets operating at local and larger geographic scales, and producing data that can be homogenized based on data bubbles, where biological data on species presence and abundance can be constantly updated (as core functionality for monitoring) via AI pipelines for image processing. The detailed technological roadmap and the potential for real-time, spatially explicit ecological monitoring position this approach as a future-proof solution for marine conservation and policy implementation, with in situ collected data usable for species distribution and environmental modelling. Despite its conceptual robustness, the proposed framework faces weaknesses as practical limitations regarding the heterogeneity and interoperability of data from diverse platforms, particularly in areas with low infrastructure or inconsistent monitoring protocols. Additionally, high initial costs for deployment and maintenance (especially for cabled observatories and deep-sea platforms) might hinder widespread adoption, especially in developing regions. Notwithstanding, there are opportunities because the proposed technological development in platforms deployment, data collection and treatment meet the ongoing EU strategies for resource management and ecosystems conservation (e.g., EU Biodiversity Strategy 2030). The framework offers an opportunity to standardize marine monitoring, improve ecological forecasting, and foster socioecological integration via agent-based modelling. Finally, hidden threats are associated to the DTO implementation itself that relies on stable funding to continue producing real-world data and updating its datasets through the PTO, thus avoiding cumulative overreliance on AI-generated information. In addition, maintaining of it, functionality also relies on technological readiness, and data-sharing agreements. Moreover, the dependence on advanced AI and high-volume data processing may limit resilience in case of system failures. The potential ecological disturbance by increased deployment of mobile units in sensitive habitats could also raise environmental concerns if not managed properly.

CRediT authorship contribution statement

Jacopo Aguzzi: Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Elias Chatzidouros:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization. **Damianos Chatzievangelou:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Morane Clavel-Henry:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Sascha Flögel:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization. **Nixon Bahamon:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Michael Tangerlini:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Laurenz Thomsen:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization. **Giacomo Picardi:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Joan Navarro:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Ivan Masmitja:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Nathan J. Robinson:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Tim Nattkemper:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Sergio Stefanni:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **José Quintana:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Ricard Campos:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Rafael García:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization. **Emanuela Fanelli:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Marco Francescangeli:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Luca Mirimin:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Roberto Danovaro:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization. **Daniel Mihai Toma:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Joaquín Del Rio-Fernandez:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Conceptualization. **Enoc Martinez:** Data curation. **Pol Baños:** Data curation. **Oriol Prat:** Data curation. **David Sarria:** Data curation. **Matias Carandell:** Data curation.

Jonathan White: Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Thomas Parissis:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization. **Stavroula Panagiotidou:** Writing – review & editing, Methodology, Conceptualization. **Juliana Quevedo:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Silvia Gallegati:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Jordi Grinyó:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Erik Simon-Lledó:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Joan B. Company:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Jennifer Doyle:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Funding acquisition, Conceptualization.

Funding

This work was supported by the European Union through the projects DIGI4ECO [grant number 101112883 - GAP-101112883], REDRESS [grant number 101135492-GAP-101135492], MERLIN [grant number GAP-01189796], SUN-BIO [grant number GAP-101157493] and ENDURUNS [grant number 824348]; Ministerio de Ciencia e Innovación through the project PLOME [grant number PLEC2021-007525]. Authors want to acknowledge Bluenetcat network (https://www.bluenetcat.eu/) and GL for the funding and development of the first version of the OBSEA DTO shown in Fig. 8. IM, NJR, and ESL were funded by the Spanish National Program Ramón y Cajal [grant numbers RYC2022-038056-I, RYC2021-034381-I, and RYC2023-043275-I], financed by MCIN/AEI/10.13039/501100011033 and the European Union Next Generation EU/PRTR. DC was funded by a Juan de la Cierva Formación Postdoctoral Fellowship financed by MCIN/AEI/10.13039/501100011033 and European Union Next Generation EU/PRTR funds [grant number FJC2021-047734-I], and received aditional funding from the European Union under the Horizon Europe MSCA grant 101104596. GP received financial support through the Horizon Europe MSCA grant 101061354. Views and opinions expressed, are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

This work used the EGI infrastructure with the dedicated support of EGI-IFCA-STACK.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Jacopo Aguzzi reports financial support was provided by European Union. Jacopo Aguzzi reports financial support was provided by Spain Ministry of Science and Innovation. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research has been carried out within the framework of the project "Severo Ochoa Center of Excellence" financially supported by ICM-CSIC (CEX2019-000928-S) and the Research Unit *Tecnoterra* (ICM-CSIC/UPC) of the Spanish Government. The photomosaic (ref. S1C1210_1-002, in the folder S1C1210_1) of Fig. 6 is a courtesy of Dr. K. Wieland (DTU Aqua, Denmark).

Appendix A. Codes and names for the different EBVs, EOVs and MSFD GES Descriptors reported in Table 1

N	EBV	P (EOV - Physics)	BGC (EOV –Biogeochemistry)	BE (EOV - Biology and Ecosystems)	CD (EOV –Cross-disciplinary)	MSFD GES descriptors
1	Genetic diversity (richness and heterozygosity)	Sea state	Oxygen	Phytoplankton biomass and diversity	Ocean colour	Marine biodiversity
2	Genetic differentiation (number of genetic units and genetic distance)	Ocean surface stress	Nutrients	Zooplankton biomass and diversity	Marine debris (*pilot)	Non-indigenous species
3	Effective population size	Sea ice	Inorganic carbon	Fish abundance and distribution	Ocean sound	Commercial fish and shellfish
4	Inbreeding	Sea surface height	Transient tracers	Sea turtles abundance and distribution		Food webs
5	Species distributions	Sea surface temperature	Particulate matter	Seabirds abundance and distribution		Eutrophication
6	Species abundances	Subsurface temperature	Nitrous oxide	Marine mammal abundance and distribution		Seabed integrity
7	Morphology	Surface currents	Stable carbon isotopes	Coral cover and composition		Hydrographical conditions
8	Physiology	Subsurface currents	Dissolved organic carbon	Seagrass cover and composition		Contaminants
9	Phenology	Sea surface salinity		Macroalgal canopy cover and composition		Contaminants in seafood
10	Movement	Subsurface salinity		Mangrove cover and composition		Marine litter
11	Reproduction	Ocean surface heat flux		Microbe biomass and diversity (*pilot)		Energy, including underwater noise
12	Community abundance	Ocean bottom pressure		Benthic invertebrate abundance and distribution (*pilot)		
13	Taxonomic/phylogenetic diversity	Turbulent diapycnal fluxes (*pilot)				
14	Trait diversity					
15	Interaction diversity					
16	Primary productivity					
17	Ecosystem phenology					

(*continued on next page*)

(*continued*)

N	EBV	P (EOV - Physics)	BGC (EOV –Biogeochemistry)	BE (EOV - Biology and Ecosystems)	CD (EOV –Cross-disciplinary)	MSFD GES descriptors
18	Ecosystem disturbances					
19	Live cover fraction					
20	Ecosystem distribution					
21	Ecosystem Vertical Profile					

Data availability

Raw and processed data and metadata presented here is public, online and free following FAIR principles.-

Physical data from Obsea sensors is available at○

https://data.obsea.es/erddap/tabledap/index.html?page=1&itemsPerPage=1000-

Raw images, labelled images and AI models are available here○

Baños Castelló, P., Prat Bayarri, O., Martínez Padró, E., Francescangeli, M., Aguzzi, J., & del Rio, J. (2025). Labelled Images at OBSEA for Object Detection Algorithms [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14888328○

Martínez Padró, E., Prat i Bayarri, O., Baños Castelló, P., Francescangeli, M., Toma, D. M., Nogueras-Cervera, M., Artero-Delgado, C., Carandell Widmer, M., Cadena muñoz, F. J., Bghiel Bensalah, I., Batet Xaus, G., Aguzzi, J., & del Rio, J. (2025). AI-based fish detections at OBSEA Underwater Observatory [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14916451○

Baños Castelló, P., Prat I Bayarri, O., Martínez Padró, E., Francescangeli, M., & del Rio, J. (2025). Underwater images from OBSEA fish detection training dataset (YOLO) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14888440○

Pol, B., Oriol, P., Martínez Padró, E., & del Rio, J. (2025). OBSEA fish detector AI model (YOLO). Zenodo. https://doi.org/10.5281/zenodo.14910365-

Abyssal NE Pacific Seafloor Megafauna Dataset is available at:○

Erik Simon-Lledó, Amon, D. J., Bribiesca-Contreras, G., Daphne Cuvelier, Jennifer M. Durden, Sofia P. Ramalho, Katja Uhlenkott, Pedro Martinez Arbizu, Noëlie Benoist, Jonathan Copley, Thomas G. Dahlgren, Adrian G. Glover, Bethany Fleming, Tammy Horton, Se-Jong Ju, Alejandra Mejia-Saenz, Kirsty McQuaid, Ellen Pape, Chailinn Park, … Daniel O. B. Jones. (2023). Abyssal NE Pacific Seafloor Megafauna Dataset (Versión 1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.7982462-

Source code of the GUI visualization tool is available at○

https://github.com/BlueNetCat/OBSEA

References

Abdelmageed, N., Algergawy, A., Samuel, S., König-Ries, B., 2021. BiodivOnto: towards a core ontology for biodiversity. In: European Semantic Web Conference. Springer International Publishing, Cham, pp. 3–8.

Addison, P.F.E., Collins, D.J., Trebilco, R., Howe, S., Bax, N., Hedge, P., Jones, G., Miloslavich, P., Roelfsema, C., Sams, M., Stuart-Smith, R.D., Scanes, P., von Baumgarten, P., McQuatters-Gollop, A., 2017. A new wave of marine evidence-based management: emerging challenges and solutions to transform monitoring, evaluating, and reporting. ICES J. Mar. Sci. 75 (3), 941–952. https://doi.org/10.1093/icesjms/fsx216. Aguzzi, J., Albiez, J., Flögel, S., Rune, Godø, O., Grimsbø, E., Marini, S., Pfannkuche, O., Rodriguez, E., Thomsen, L., Torkelsen, T., Valencia, J., López-Vázquez, V., Wehde, H., Zhang, G., 2020a. A flexible autonomous robotic observatory infrastructure for bentho-pelagic monitoring. Sensors-Basel, 20(6), 1614. 10.3390/s20061614.

Aguzzi, J., Company, J.B., 2010. Chronobiology of deep-water decapod crustaceans on continental margins. Adv. Mar. Biol. Ann. Rev. 58, 155–225. https://doi.org/10.1016/B978-0-12-381015-1.00003-4.

Aguzzi, J., Doya, C., Tecchio, S., De Leo, F., Azzurro, E., Costa, C., Sbragaglia, V., Del Río, J., Navarro, J., Ruhl, H.A., Company, J.B., Favali, P., Purser, A., Thomsen, L., Catalán, I.A., 2015. Coastal observatories for monitoring of fish behaviour and their responses to environmental changes. Rev. Fish Biol. Fisher. 25, 463–483. https://doi.org/10.1007/s11160-015-9387-9.

Aguzzi, J., Chatzievangelou, D., Marini, S., Fanelli, E., Danovaro, R., Flögel, S., Lebris, N., Juanes, F., De Leo, F.C., Del Rio, J., Thomsen, L., Costa, C., Riccobene, G., Tamburini, C., Lefevre, D., Gojak, C., Poulain, P.-M., Favali, P., Griffa, A., Purser, A., Cline, D., Edgington, D., Navarro, J., Stefanni, S., D'Hondt, S., Priede, I.G., Rountree, R., Company, J.B., 2019. New high-tech flexible networks for the monitoring of Deep-Sea ecosystems. Environ. Sci. Technol. 53, 6616–6631. https://doi.org/10.1021/acs.est.9b00409.

Aguzzi, J., Chatzievangelou, D., Francescangeli, M., Marini, S., Bonofiglio, F., Del Rio, J., Danovaro, R., 2020. The hierarchic treatment of marine ecological information from spatial networks of benthic platforms. Sensors 20, 1751. https://doi.org/10.3390/s20061751.

Aguzzi, J., Chatzievangelou, D., Robinson, N.J., Bahamon, N., Berry, A., Carreras, M., Company, J.B., Costa, C., Del Rio Fernandez, J., Falahzadeh, A., Fifas, S., Flögel, S., Grinyó, J., Jónasson, J.P., Jonsson, P., Lordan, C., Lundy, M., Marini, S., Martinelli, M., Masmitja, I., Mirimin, L., Naseer, A., Navarro, J., Palomeras, N., Picardi, G., Silva, C., Stefanni, S., Vigo, M., Vila, Y., Weetman, A., Doyle, J., 2022. Advancing fishery-independent stock assessments for the Norway lobster (Nephrops norvegicus) with new monitoring technologies. Front. Mar. Sci. 9, 969071. https://doi.org/10.3389/fmars.2022.969071.

Aguzzi, J., Thomsen, L., Flögel, S., Robinson, N.J., Picardi, G., Chatzievangelou, D., Bahamon, N., Stefanni, S., Grinyó, J., Fanelli, E., Corinaldesi, C., Del Rio Fernandez, J., Calisti, M., Mienis, F., Chatzidouros, E., Costa, C., Violino, S., Tangherlini, M., Danovaro, R., 2024. New Technologies for Monitoring and Upscaling Marine Ecosystem Restoration in Deep-Sea environments. Engineering S2095809924000286. https://doi.org/10.1016/j.eng.2023.10.012.

Andruszkiewicz, E.A., Koseff, J.R., Fringer, O.B., Ouellette, N.T., Lowe, A.B., Edwards, C. A., Boehm, A.B., 2019. Modeling environmental DNA transport in the Coastal Ocean using Lagrangian particle tracking. Front. Mar. Sci. 6, 477. https://doi.org/10.3389/fmars.2019.00477.

Bahamon, N., Aguzzi, J., Bernardello, R., Ahumada, M.-A., Puigdefabregas, J., Muñoz, E., Velásquez, Z., Cruzado, A., 2011. The new pelagic operational Observatory of the Catalan Sea (OOCS) for the multisensor coordinated measurement of atmospheric and oceanographic conditions. Sensors-Basel 11, 11251–11272. https://doi.org/10.3390/s111211251.

Baños Castelló, P., Prat I Bayarri, O., Martínez Padró, E., Francescangeli, M., del Rio, J., 2025. Underwater Images From OBSEA Fish Detection Training Dataset (YOLO) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.14888440.

Barbie, A., Pech, N., Hasselbring, W., Flogel, S., Wenzhofer, F., Walter, M., Shchekinova, E., Busse, M., Turk, M., Hofbauer, M., Sommer, S., 2022. Developing an underwater network of ocean observation systems with digital twin prototypes—a field report from the Baltic Sea. IEEE Internet Comput. 26, 33–42. https://doi.org/10.1109/MIC.2021.3065245.

Boschert, S., Rosen, R., 2016. Digital twin—The simulation aspect. In: Hehenberger, P., Bradley, D. (Eds.), Mechatronic Futures. Springer International Publishing, Cham, pp. 59–74. https://doi.org/10.1007/978-3-319-32156-1_5.

Brönner, U., Sonnewald, M., Visbeck, M., 2023. Digital twins of the ocean can foster a sustainable blue economy in a protected marine environment. Int. Hydrogr. Rev. 29, 26–40. https://doi.org/10.58440/ihr-29-a04.

Büscher, J.V., Juva, K., Flögel, S., Wisshak, M., Rüggeberg, A., Riebesell, U., Form, A.U., 2024. Water mass characteristics and hydrodynamics at an inshore versus an offshore mid-Norwegian cold-water coral reef habitat. Front. Mar. Sci. 11. https://doi.org/10.3389/fmars.2024.1363542.

Caiani, A., Russo, A., Palestrini, A., Gallegati, M. (Eds.), 2016. Economics with Heterogeneous Interacting Agents: A Practical Guide to Agent-Based Modeling, New Economic Windows. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-44058-3.

Carandell, M., Hernández, S., Toma, D.M., Martínez, E., Nogueras, M., Aguzzi, J., del Río, J., 2023, June. Expanding the underwater communication capabilities of seafloor ecosystem monitoring stand-alone platforms using pop-up buoys. In: OCEANS 2023-Limerick. IEEE, pp. 1–7. https://doi.org/10.1109/OCEANSLimerick52467.2023.10244428.

Chatzievangelou, D., Aguzzi, J., Ogston, A., Suárez, A., Thomsen, L., 2020. Visual monitoring of key deep-sea megafauna with an internet operated crawler as a tool for ecological status assessment. Prog. Oceanogr. 184, 102321. https://doi.org/10.1016/j.pocean.2020.102321.

Chatzievangelou, D., Thomsen, L., Doya, C., Purser, A., Aguzzi, J., 2022. Transects in the deep: opportunities with tele-operated resident seafloor robots. Front. Mar. Sci. 9, 833617. https://doi.org/10.3389/fmars.2022.833617.

Chen, G., Yang, J., Huang, B., Ma, C., Tian, F., Ge, L., Xia, L., Li, J., 2023. Toward digital twin of the ocean: from digitalization to cloning. Intell. Mar. Technol. Syst. 1, 3. https://doi.org/10.1007/s44295-023-00003-2.

Clavel-Henry, M., Solé, J., Kristiansen, T., Bahamon, N., Rotllant, G., Company, J.B., 2020. Modeled buoyancy of eggs and larvae of the deep-sea shrimp Aristeus antennatus (Crustacea: Decapoda) in the northwestern Mediterranean Sea. PloS One 15, e0223396. https://doi.org/10.1371/journal.pone.0223396.
Clavel-Henry, M., Bahamon, N., Aguzzi, J., Navarro, J., López, M., Company, J.B., 2024. Indicators to Assess Interannual Variability in Marine Connectivity Processes: A Semi-theoretical Approach. https://doi.org/10.1101/2024.01.12.575479.
Cline, D.E., Edgington, D.R., Smith, K.L., Vardaro, M.F., Kuhnz, L., Ellena, J.A., 2009. An automated event detection and classification system for abyssal time-series images of station M. In: NE Pacific, in: OCEANS 2009. Presented at the OCEANS 2009. IEEE, Biloxi, MS, pp. 1–4. https://doi.org/10.23919/OCEANS.2009.5422292.
Colton, M.A., Swearer, S.E., 2010. A comparison of two survey methods: differences between underwater visual census and baited remote underwater video. Mar. Ecol. Prog. Ser. 400, 19–36.
Córdova, P., Flores, R.P., 2022. Hydrodynamic and particle drift modeling as a support system for maritime search and rescue (SAR) emergencies: application to the C-212 aircraft accident on 2 September, 2011, in the Juan Fernández Archipelago, Chile. J. Mar. Sci. Eng. 10, 1649. https://doi.org/10.3390/jmse10111649.
Cuvelier, D., Zurowietz, M., Nattkemper, T.W., 2024. Deep learning–assisted biodiversity assessment in deep-sea benthic megafauna communities: a case study in the context of polymetallic nodule mining. Front. Mar. Sci. 11, 1366078. https://doi.org/10.3389/fmars.2024.1366078.
Daly, H.E., Farley, J., 2011. Ecological economics: principles and applications, 2nd ed. Island press, Washington DC.
Dañobeitia, J.J., Pouliquen, S., Pade, N., Arvanitidis, C., Sanders, R., Stanica, A., Gourcuff, C., Petihakis, G., Tegas, V., Favali, P., 2023. The role of the marine research infrastructures in the European marine observation landscape: present and future perspectives. Front. Mar. Sci. 10, 1047251. https://doi.org/10.3389/fmars.2023.1047251.
Danovaro, R., Aguzzi, J., Fanelli, E., Billett, D., Gjerde, K., Jamieson, A., Ramirez-Llodra, E., Smith, C.R., Snelgrove, P.V.R., Thomsen, L., Dover, C.L.V., 2017. An ecosystem-based deep-ocean strategy. Science 355, 452–454. https://doi.org/10.1126/science.aah7178.
Danovaro, R., Fanelli, E., Aguzzi, J., Billett, D., Carugati, L., Corinaldesi, C., Dell'Anno, A., Gjerde, K., Jamieson, A.J., Kark, S., McClain, C., Levin, L., Levin, N., Ramirez-Llodra, E., Ruhl, H., Smith, C.R., Snelgrove, P.V.R., Thomsen, L., Van Dover, C.L., Yasuhara, M., 2020. Ecological variables for developing a global deep-ocean monitoring and conservation strategy. Nat. Ecol. Evol. 4, 181–192. https://doi.org/10.1038/s41559-019-1091-z.
Darnaude, A., Tanner, S., Hunter, E., Costantini, F., 2024. Advancing research in marine functional connectivity for improved policy and management. Mar. Ecol. Prog. Ser. 731, 1–8. https://doi.org/10.3354/meps14550.
Davies, N., Field, D., Gavaghan, D., Holbrook, S.J., Planes, S., Troyer, M., Bonsall, M., Claudet, J., Roderick, G., Schmitt, R.J., Zettler, L.A., Berteaux, V., Bossin, H.C., Cabasse, C., Collin, A., Deck, J., Dell, T., Dunne, J., Gates, R., Harfoot, M., Hench, J. L., Hopuare, M., Kirch, P., Kotoulas, G., Kosenkov, A., Kusenko, A., Leichter, J.J., Lenihan, H., Magoulas, A., Martinez, N., Meyer, C., Stoll, B., Swalla, B., Tartakovsky, D.M., Murphy, H.T., Turyshev, S., Valdvinos, F., Williams, R., Wood, S., 2016. Simulating social-ecological systems: the island digital ecosystem avatars (IDEA) consortium. GigaScience 5, 14. https://doi.org/10.1186/s13742-016-0118-5.
De Koning, K., Broekhuijsen, J., Kühn, I., Ovaskainen, O., Taubert, F., Endresen, D., Schigel, D., Grimm, V., 2023. Digital twins: dynamic model-data fusion for ecology. Trends Ecol. Evol. 38, 916–926. https://doi.org/10.1016/j.tree.2023.04.010.
Del-Rio, J., Nogueras, M., Toma, D.M., Martinez, E., Artero-Delgado, C., Bghiel, I., Martinez, M., Cadena, J., Garcia-Benadi, A., Sarria, D., Aguzzi, J., Masmitja, I., Carandell, M., Olive, J., Gomariz, S., Santamaria, P., Manuel Lazaro, A., 2020. Obsea: A Decadal Balance for a Cabled Observatory Deployment, 8. IEEE Access, pp. 33163–33177. https://doi.org/10.1109/ACCESS.2020.2973771.
Dominguez-Carrió, C., Fontes, J., Morato, T., 2021. A cost-effective video system for a rapid appraisal of deep-sea benthic habitats: the Azor drift-cam. Methods Ecol. Evol. 12, 1379–1388. https://doi.org/10.1111/2041-210X.13617.
Doya, C., Aguzzi, J., Chatzievangelou, D., Costa, C., Company, J.B., Tunnicliffe, V., 2016. The seasonal use of small-scale space by benthic species in a transiently hypoxic area. J. Mar. Syst. 154, 280–290. https://doi.org/10.1016/j.jmarsys.2015.09.005.
Duarte, C.M., Agusti, S., Barbier, E., Britten, G.L., Castilla, J.C., Gattuso, J.-P., Fulweiler, R.W., Hughes, T.P., Knowlton, N., Lovelock, C.E., Lotze, H.K., Predragovic, M., Poloczanska, E., Roberts, C., Worm, B., 2020. Rebuilding marine life. Nature 580, 39–51. https://doi.org/10.1038/s41586-020-2146-7.
Durden, J., Schoening, T., Althaus, F., Friedman, A., Garcia, R., Glover, A., Greinert, J., Stout, N., Jones, D., Jordt, A., Kaeli, J., Koser, K., Kuhnz, L., Lindsay, D., Morris, K., Nattkemper, T., Osterloff, J., Ruhl, H., Singh, H., Bett, B., 2016. Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. Oceanogr. Mar. Biol. 1–72.
Durden, J.M., Schoening, T., Curtis, E.J., Downie, A., Gates, A.R., Jones, D.O.B., Kokkinaki, A., Simon-Lledó, E., Wright, D., Bett, B.J., 2024. Defining the target population to make marine image-based biological data FAIR. Ecol. Inform. 80, 102526. https://doi.org/10.1016/j.ecoinf.2024.102526.
Eaton, B., Gregory, J., Drach, B., Taylor, K., Hankin, S., Caron, J., Signell, R., 2023. NetCDF Climate and Forecast (CF) Metadata Conventions Version 1.11. https://cfconventions.org.
Espinoza, M., Lédée, E.J.I., Simpfendorfer, C.A., Tobin, A.J., Heupel, M.R., 2015. Contrasting movements and connectivity of reef-associated sharks using acoustic telemetry: implications for management. Ecol. Appl. 25, 2101–2118. https://doi.org/10.1890/14-2293.1.
European Comission (EU), 2021. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM (2021) 609 Final.
European Union, n.d. EDITOInfra European Digital Twin Ocean.
Falahzadeh, A., Toma, D.M., Francescangeli, M., Chatzievangelou, D., Nogueras, M., Martínez, E., Carandell, M., Tangerlini, M., Thomsen, L., Picardi, G., Le Bris, M., Dominguez, L., Aguzzi, J., Del Río, J., 2023. A new coastal crawler prototype to expand the ecological monitoring radius of OBSEA cabled observatory. J Mar Sci Eng 11, 857. https://doi.org/10.3390/jmse11040857.
Ferreira, R., Dinis, A., Badenas, A., Sambolino, A., Marrero-Pérez, J., Crespo, A., Alves, F., 2021. Bryde's whales in the north-East Atlantic: new insights on site fidelity and connectivity between oceanic archipelagos. Aquat. Conserv. Mar. Freshw. Ecosyst. 31, 2938–2950. https://doi.org/10.1002/aqc.3665.
Flögel, S., 2018. A new deep-sea crawler system - MANSIO-VIATOR. In: 2018 OCEANS - MTS/IEEE Kobe Tecehno-Oceans (OTO), pp. 1–10. https://doi.org/10.1109/OCEANSKOBE.2018.8559368.
Fobert, E.K., Treml, E.A., Swearer, S.E., 2019. Dispersal and population connectivity are phenotype dependent in a marine metapopulation. Proc. R. Soc. B Biol. Sci. 286, 20191104. https://doi.org/10.1098/rspb.2019.1104.
Franklin, J., 2023. Species distribution modelling supports the study of past, present and future biogeographies. J. Biogeogr. 50, 1533–1545. https://doi.org/10.1111/jbi.14617.
Gallegati, M., Palestrini, A., Russo, A., 2017. Introduction to Agent-Based Economics. Academic Press. https://doi.org/10.1016/C2015-0-00736-5.
Gallegati, G., Gallegati, M., Landini, S., 2024. Agent-Based Modelling: A Tool for Complexity. Cambridge University Press.
Garcia, R., Prados, R., Quintana, J., Tempelaar, A., Gracias, N., Rosen, S., Vågstøl, H., Løvall, K., 2020. Automatic segmentation of fish using deep learning with application to fish size measurement. ICES J. Mar. Sci. 77, 1354–1366. https://doi.org/10.1093/icesjms/fsz186.
Ghazi, S., Khadir, T., Dugdale, J., 2014. Multi-agent based simulation of environmental pollution issues: A review. in: Corchado, J.M., Bajo, J., Kozlak, J., Pawlewski, P., Molina, J.M., Gaudou, B., Julian, V., Unland, R., Lopes, F., Hallenborg, K.,. In: García Teodoro, P. (Ed.), Highlights of Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection, Communications in Computer and Information Science. Springer International Publishing, Cham, pp. 13–21. https://doi.org/10.1007/978-3-319-07767-3_2.
Glaviano, F., Esposito, R., Cosmo, A.D., Esposito, F., Gerevini, L., Ria, A., Molinara, M., Bruschi, P., Costantini, M., Zupo, V., 2022. Management and sustainable exploitation of marine environments through smart monitoring and automation. J Mar Sci Eng 10, 297. https://doi.org/10.3390/jmse10020297.
Grossmann, V., Nakath, D., Urlaub, M., Oppelt, N., Koch, R., Köser, K., 2022. Digital twinning in the ocean – challenges in multimodal sensing and multiscale fusion based on faithful visual models. In: ISPRS Ann. Photogramm. Remote Sens. Spat. Inform. Sci. V-4–2022, pp. 345–352. https://doi.org/10.5194/isprs-annals-V-4-2022-345-2022.
Hamill, L., Gilbert, N., 2015. Agent-based Modelling in Economics, 1st ed. Wiley. https://doi.org/10.1002/9781118945520.
Heckbert, S., Baynes, T., Reeson, A., 2010. Agent-based modeling in ecological economics. Ann. N. Y. Acad. Sci. 1185, 39–53. https://doi.org/10.1111/j.1749-6632.2009.05286.x.
Horton, T., Marsh, L., Bett, B.J., Gates, A.R., Jones, D.O.B., Benoist, N.M.A., Pfeifer, S., Simon-Lledó, E., Durden, J.M., Vandepitte, L., Appeltans, W., 2021. Recommendations for the standardisation of open taxonomic nomenclature for image-based identifications. Front. Mar. Sci. 8, 620702. https://doi.org/10.3389/fmars.2021.620702.
Howe, B.M., Chao, Y., Arabshahi, P., Roy, S., McGinnis, T., Gray, A., 2010. A smart sensor web for ocean observation: fixed and mobile platforms, integrated acoustics, satellites and predictive modeling. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 3, 507–521. https://doi.org/10.1109/JSTARS.2010.2052022.
Howell, K.L., Davies, J.S., Allcock, A.L., Braga-Henriques, A., Buhl-Mortensen, P., Carreiro-Silva, M., Dominguez-Carrió, C., Durden, J.M., Foster, N.L., Game, C.A., Hitchin, B., Horton, T., Hosking, B., Jones, D.O.B., Mah, C., Laguionie Marchais, C., Menot, L., Morato, T., Pearman, T.R.R., Piechaud, N., Ross, R.E., Ruhl, H.A., Saeedi, H., Stefanoudis, P.V., Taranto, G.H., Thompson, M.B., Taylor, J.R., Tyler, P., Vad, J., Victorero, L., Vieira, R.P., Woodall, L.C., Xavier, J.R., Wagner, D., 2019. 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. https://doi.org/10.1371/journal.pone.0218904.
Hughes, J.S., Crichton, D., Hardman, S., Law, E., Joyner, R., Ramirez, P., 2014. PDS4: A model-driven planetary science data architecture for long-term preservation. In: In: 2014 IEEE 30th International Conference on Data Engineering Workshops. Presented at the 2014 IEEE 30th International Conference on Data Engineering Workshops (ICDEW). IEEE, Chicago, IL, USA, pp. 134–141. https://doi.org/10.1109/ICDEW.2014.6818317.
IPBES, 2024. socio-ecological system (IPBES).
Jayne, S., Roemmich, D., Zilberman, N., Riser, S., Johnson, K., Johnson, G., Piotrowicz, S., 2017. The argo program: present and future. Oceanography 30, 18–28. https://doi.org/10.5670/oceanog.2017.213.
Juva, K., Flögel, S., Karstensen, J., Linke, P., Dullo, W.-C., 2020. Tidal dynamics control on cold-water coral growth: a high-resolution multivariable study on eastern Atlantic cold-water coral sites. Front. Mar. Sci. 7, 132. https://doi.org/10.3389/fmars.2020.00132.
Juva, K., Kutti, T., Chierici, M., Dullo, W.-C., Flögel, S., 2021. Cold-water coral reefs in the Langenuen Fjord, southwestern Norway – a window into future environmental change. Oceans 2, 583–610. https://doi.org/10.3390/oceans2030033.

Katija, K., Orenstein, E., Schlining, B., Lundsten, L., Barnard, K., Sainz, G., Boulais, O., Cromwell, M., Butler, E., Woodward, B., Bell, K.L.C., 2022. FathomNet: A global image database for enabling artificial intelligence in the ocean. Sci. Rep. 12, 15914. https://doi.org/10.1038/s41598-022-19939-2.
Keramea, P., Kokkos, N., Gikas, G., Sylaios, G., 2022. Operational modeling of North Aegean oil spills forced by Real-time Met-Ocean forecasts. J Mar Sci Eng 10, 411. https://doi.org/10.3390/jmse10030411.
Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R., 2023. Segment anything. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Paris, pp. 3992–4003.
Kootbally, Z., Schlenoff, C., Lawler, C., Kramer, T., Gupta, S.K., 2015. Towards robust assembly with knowledge representation for the planning domain definition language (PDDL). Robot. Comput. Integr. Manuf. 33, 42–55. https://doi.org/10.1016/j.rcim.2014.08.006.
Lambertini, A., Menghini, M., Cimini, J., Odetti, A., Bruzzone, G., Bibuli, M., De Marchi, L., 2022. Underwater drone architecture for marine digital twin: lessons learned from sushi drop project. Sensors 22 (3), 744. https://doi.org/10.3390/s22030744.
Langenkämper, D., Zurowietz, M., Schoening, T., Nattkemper, T.W., 2017. BIIGLE 2.0 - browsing and annotating large marine image collections. Front. Mar. Sci. 4, 83. https://doi.org/10.3389/fmars.2017.00083.
Lantéri, N., Ruhl, H.A., Gates, A., Martínez, E., Del Rio Fernandez, J., Aguzzi, J., Cannat, M., Delory, E., Embriaco, D., Huber, R., Matabos, M., Petihakis, G., Reilly, K., Rolin, J.-F., Van Der Schaar, M., André, M., Blandin, J., Cianca, A., Francescangeli, M., Garcia, O., Hartman, S., Lagadec, J.-R., Legrand, J., Pagonis, P., Piera, J., Remirez, X., Toma, D.M., Marinaro, G., Moreau, B., Santana, R., Wright, H., Dañobeitia, J.J., Favali, P., 2022. The EMSO generic instrument module (EGIM): standardized and interoperable instrumentation for ocean observation. Front. Mar. Sci. 9, 801033. https://doi.org/10.3389/fmars.2022.801033.
Leenhardt, P., Teneva, L., Kininmonth, S., Darling, E., Cooley, S., Claudet, J., 2015. Challenges, insights and perspectives associated with using social-ecological science for marine conservation. Ocean Coast. Manag. 115, 49–60. https://doi.org/10.1016/j.ocecoaman.2015.04.018.
Liang, S., Khalafbeigi, T., van der Schaaf, H., Miles, B., Schleidt, K., Grellet, S., Beaufils, M., Alzona, M., 2021. OGC SensorThings API Part 1: Sensing Version 1.1.
Lindkvist, E., Wijermans, N., Daw, T.M., Gonzalez-Mon, B., Giron-Nava, A., Johnson, A. F., Van Putten, I., Basurto, X., Schlüter, M., 2020. Navigating complexities: agent-based modeling to support research, governance, and Management in Small-Scale Fisheries. Front. Mar. Sci. 6, 733. https://doi.org/10.3389/fmars.2019.00733.
Llorach-Tó, G., Martínez, E., Fernández, J.D.R., García-Ladona, E., 2023. Experience OBSEA: a web-based 3D virtual environment of a seafloor observatory. In: OCEANS 2023 - Limerick. Presented at the OCEANS 2023 - Limerick, IEEE, Limerick, Ireland, pp. 1–6. https://doi.org/10.1109/OCEANSLimerick52467.2023.10244688.
Loomis, J., Bond, C., Harpman, D., 2009. The potential of agent-based modelling for performing economic analysis of adaptive natural resource management. J. Nat. Resour. Policy Res. 1, 35–48. https://doi.org/10.1080/19390450802509773.
Macenski, S., Foote, T., Gerkey, B., Lalancette, C., Woodall, W., 2022. Robot Operating System 2: Design, Architecture, and Uses in the Wild. https://doi.org/10.48550/ARXIV.2211.07752.
Marini, S., Fanelli, E., Sbragaglia, V., Azzurro, E., Del Rio Fernandez, J., Aguzzi, J., 2018. Tracking fish abundance by underwater image recognition. Sci. Rep. 8, 13748. https://doi.org/10.1038/s41598-018-32089-8.
Martínez, E., Garcia-Benadí, A., Toma, D.M., Carandell, M., Nogueras, M., Del Río, J., 2023. An e-infrastructure for FAIR data management of underwater observatories. In: OCEANS 2023 - Limerick. Presented at the OCEANS 2023 - Limerick, IEEE, Limerick, Ireland, pp. 1–6. https://doi.org/10.1109/OCEANSLimerick52467.2023.10244451.
Masmitja, I., Toma, D., Bahamon, N., Carandell, M., García, del Arco, J.A., Chatzievangelou, D., Real, M., Sánchez-Márquez, A., Batet, G., Bonin-Font, F., Muntaner, C., Hurtós, N., López, J.M., Ripoll, A.R., de Arcas, G., Brinyó, J., Gomaritz, S., Katija, K., Aguzzi, J., 2024. Interconnected Robotic Platforms Inform Deep-Sea Ecological Restoration Trends. Marine Pollution Bulletin, Submitted.
Matos, F.L., Aguzzi, J., Company, J.B., Cunha, M.R., 2024. Gone with the stream: functional connectivity of a cold-water coral at basin scale. Limnol. Oceanogr. 69, 217–231. https://doi.org/10.1002/lno.12444.
Matthews, R.B., Gilbert, N.G., Roach, A., Polhill, J.G., Gotts, N.M., 2007. Agent-based land-use models: a review of applications. Landsc. Ecol. 22, 1447–1459. https://doi.org/10.1007/s10980-007-9135-1.
Meynard, C.N., Piou, C., Kaplan, D.M., 2023. A theoretical framework for upscaling species distribution models. Methods Ecol. Evol. 14, 2888–2899. https://doi.org/10.1111/2041-210X.14207.
Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley, D.M., Chiba, S., Duffy, J.E., Dunn, D.C., Fischer, A., Gunn, J., Kudela, R., Marsac, F., Muller-Karger, F.E., Obura, D., Shin, Y., 2018. Essential Ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Change Biol. 24, 2416–2433. https://doi.org/10.1111/gcb.14108.
Mohseni, F., Saba, F., Mirmazloumi, S.M., Amani, M., Mokhtarzade, M., Jamali, S., Mahdavi, S., 2022. Ocean water quality monitoring using remote sensing techniques: A review. Mar. Environ. Res. 180, 105701. https://doi.org/10.1016/j.marenvres.2022.105701.
Mokany, K., Ferrier, S., Harwood, T.D., Ware, C., Di Marco, M., Grantham, H.S., Venter, O., Hoskins, A.J., Watson, J.E.M., 2020. Reconciling global priorities for conserving biodiversity habitat. Proc. Natl. Acad. Sci. 117, 9906–9911. https://doi.org/10.1073/pnas.1918373117.
Möller, T., Nattkemper, T.W., 2021. ALMI—A generic active learning system for computational object classification in marine observation images. Sensors 21, 1134. https://doi.org/10.3390/s21041134.
Montero, D., Kraemer, G., Anghelea, A., Aybar Camacho, C.L., Brandt, G., Camps-Valls, G., Cremer, F., Flik, I., Gans, F., Habershon, S., Ji, C., Kattenborn, T., Martínez-Ferrer, L., Martinuzzi, F., Reinhardt, M., Söchting, M., Teber, K., Mahecha, M., 2024. Earth system data cubes: avenues for advancing Earth system research. Environ. Data Sci. 3, e27. https://doi.org/10.1017/eds.2024.22.
Moran, J., Fosset, A., Kirman, A., Benzaquen, M., 2021. From ants to fishing vessels: a simple model for herding and exploitation of finite resources. J. Econ. Dyn. Control 129, 104169. https://doi.org/10.1016/j.jedc.2021.104169.
Moran, K., Juniper, S.K., Bligh, S., Loock, D., Kulin, I., Paulson, M., Pirenne, B., 2022. Canada's internet-connected ocean. Front. Mar. Sci. 8, 805134. https://doi.org/10.3389/fmars.2021.805134.
Morris, R.A., Barve, V., Carausu, M., Chavan, V., Cuadra, J., Freeland, C., Hagedorn, G., Leary, P., Mozzherin, D., Olson, A., Riccardi, G., Teage, I., Whitbread, G., 2013. Discovery and publishing of primary biodiversity data associated with multimedia resources: the Audubon Core strategies and approaches. Biodivers. Inform. 8. https://doi.org/10.17161/bi.v8i2.4117.
Muller-Karger, F.E., Miloslavich, P., Bax, N.J., Simmons, S., Costello, M.J., Sousa Pinto, I., Canonico, G., Turner, W., Gill, M., Montes, E., Best, B.D., Pearlman, J., Halpin, P., Dunn, D., Benson, A., Martin, C.S., Weatherdon, L.V., Appeltans, W., Provoost, P., Klein, E., Kelble, C.R., Miller, R.J., Chavez, F.P., Iken, K., Chiba, S., Obura, D., Navarro, L.M., Pereira, H.M., Allain, V., Batten, S., Benedetti-Checchi, L., Duffy, J.E., Kudela, R.M., Rebelo, L.-M., Shin, Y., Geller, G., 2018. Advancing marine biological observations and data requirements of the complementary Essential Ocean variables (EOVs) and essential biodiversity variables (EBVs) frameworks. Front. Mar. Sci. 5, 211. https://doi.org/10.3389/fmars.2018.00211.
Murawski, S.A., Steele, J.H., Taylor, P., Fogarty, M.J., Sissenwine, M.P., Ford, M., Suchman, C., 2010. Why compare marine ecosystems? ICES. J. Mar. Sci. 67, 1–9. https://doi.org/10.1093/icesjms/fsp221.
Nativi, S., Mazzetti, P., Craglia, M., 2021. Digital ecosystems for developing digital twins of the earth: the destination earth case. Remote Sens 13, 2119. https://doi.org/10.3390/rs13112119.
Newth, D., Gunasekera, D., 2012. An integrated agent-based framework for assessing air pollution impacts. J. Environ. Prot. 03, 1135–1146. https://doi.org/10.4236/jep.2012.329132.
Ortenzi, L., Aguzzi, J., Costa, C., Marini, S., D'agostino, D., Thomsen, L., De Leo, F., Chatzievangelou, D., 2024. Automated species classification and counting by deep-sea mobile crawler platforms using YOLO. Eco. Inform. 82, 102788. https://doi.org/10.1016/j.ecoinf.2024.102788.
Ottaviani, E., Francescangeli, M., Gjeci, N., Del Rio Fernandez, J., Aguzzi, J., Marini, S., 2022. Assessing the image concept drift at the OBSEA coastal underwater cabled observatory. Front. Mar. Sci. 9, 840088. https://doi.org/10.3389/fmars.2022.840088.
Palomeras, N., Vallicrosa, G., Mallios, A., Bosch, J., Vidal, E., Hurtos, N., Ridao, P., 2018. AUV homing and docking for remote operations. Ocean Eng. 154, 106–120. https://doi.org/10.1016/j.oceaneng.2018.01.114.
Payo-Payo, M., Jacinto, R.S., Lastras, G., Rabineau, M., Puig, P., Martín, J., Canals, M., Sultan, N., 2017. Numerical modeling of bottom trawling-induced sediment transport and accumulation in La Fonera submarine canyon, northwestern Mediterranean Sea. Mar. Geol. 386, 107–125. https://doi.org/10.1016/j.margeo.2017.02.015.
Pereiro, D., Lyons, K., Orthmani, A., Fuller, R., Nolan, G., Dabrowski, T., 2021. DEVELOPMENT OF AN APPLICATION TO TRACK MISSING PERSONS AT SEA. Presented at the 9th EuroGOOS International Conference, Shom; Ifremer; EuroGOOS AISBL, HAL Open Science, Brest, France, pp. 414–419.
Phillips, B.T., Licht, S., Haiat, K.S., Bonney, J., Allder, J., Chaloux, N., Shomberg, R., Noyes, T.J., 2019. DEEPi: a miniaturized, robust, and economical camera and computer system for deep-sea exploration. Deep Sea Res. 153, 103136. https://doi.org/10.1016/j.dsr.2019.103136.
Pillai, U.P.A., Pinardi, N., Alessandri, J., Federico, I., Causio, S., Unguendoli, S., Valentini, A., Staneva, J., 2022. A digital twin modelling framework for the assessment of seagrass nature based solutions against storm surges. Sci. Total Environ. 847, 157603. https://doi.org/10.1016/j.scitotenv.2022.157603.
Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T.R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, Shakir, Battaglia, Peter, Remi, L., Willson, M., 2025. Probabilistic weather forecasting with machine learning. Nature 637, 84–90. https://doi.org/10.1038/s41586-024-08252-9.
Purser, A., Bergmann, M., Lundälv, T., Ontrup, J., Nattkemper, T., 2009. Use of machine-learning algorithms for the automated detection of cold-water coral habitats: a pilot study. Mar. Ecol. Prog. Ser. 397, 241–251. https://doi.org/10.3354/meps08154.
Purser, A., Thomsen, L., Barnes, C., Best, M., Chapman, R., Hofbauer, M., Menzel, M., Wagner, H., 2013. Temporal and spatial benthic data collection via an internet operated Deep Sea crawler. Methods Oceanogr. 5, 1–18. https://doi.org/10.1016/j.mio.2013.07.001.
Quigley, M., 2009. ROS: an open-source robot operating system. In: IEEE International Conference on Robotics and Automation.
Rew, R., Davis, G., 1990. NetCDF: an interface for scientific data access. In: IEEE Computer Graphics and Applications, 10, pp. 76–82. https://doi.org/10.1109/38.56302.
Robinson, N.J., Johnsen, S., Brooks, A., Frey, L., Judkins, H., Vecchione, M., Widder, E., 2021. Studying the swift, smart, and shy: unobtrusive camera-platforms for observing large deep-sea squid. Deep Sea Res. 172, 103538. https://doi.org/10.1016/j.dsr.2021.103538.

Rountree, R.A., Aguzzi, J., Marini, S., Fanelli, E., De Leo, F.C., Del Rio, J., Juanes, F., 2020. Towards an optimal design for ecosystem-level ocean observatories. In: Allcock, A.L., Bates, A.E., Evans, A.J., Firth, L.B., McQuaid, C.D., Russell, B.D., Smith, I.P., Swearer, S.E., Todd, P.A. (Eds.), Hawkins, S.J. CRC Press, Oceanography and Marine Biology, pp. 79–105. https://doi.org/10.1201/9780429351495-2.
Rufener, M.C., Kristensen, K., Nielsen, J.R., Bastardie, F., 2021. Bridging the gap between commercial fisheries and survey data to model the spatiotemporal dynamics of marine species. Ecol. Appl. 31 (8), e02453.
Salman, A., Jalal, A., Shafait, F., Mian, A., Shortis, M., Seager, J., Harvey, E., 2016. Fish species classification in unconstrained underwater environments based on deep learning: fish classification based on deep learning. Limnol. Oceanogr. Methods 14, 570–585. https://doi.org/10.1002/lom3.10113.
Saunders, J., Beaumont, N., Atkins, J.P., Lannin, A., Lear, D., Ozdemiroglu, E., Potts, T., 2015. A review of marine and coastal ecosystem services data and tools to incorporate this into decision-making. In: Turner, R.K., Schaafsma, M. (Eds.), Coastal Zones Ecosystem Services: From Science to Values and Decision Making. Springer International Publishing, Cham, pp. 145–164.
Schneider, J., Klüner, A., Zielinski, O., 2023. Towards digital twins of the oceans: the potential of machine learning for monitoring the impacts of offshore wind farms on marine environments. Sensors 23, 4581. https://doi.org/10.3390/s23104581.
Schoening, T., Kuhn, T., Bergmann, M., Nattkemper, T.W., 2015. DELPHI - fast and adaptive computational laser point detection and visual footprint quantification for arbitrary underwater image collections. Front. Mar. Sci. 2. https://doi.org/10.3389/fmars.2015.00020.
Schoening, T., Durden, J.M., Faber, C., Felden, J., Heger, K., Hoving, H.-J.T., Kiko, R., Köser, K., Krämmer, C., Kwasnitschka, T., Möller, K.O., Nakath, D., Naß, A., Nattkemper, T.W., Purser, A., Zurowietz, M., 2022. Making marine image data FAIR. Sci Data 9, 414. https://doi.org/10.1038/s41597-022-01491-3.
Sciascia, R., Guizien, K., Magaldi, M.G., 2022. Larval dispersal simulations and connectivity predictions for Mediterranean gorgonian species: sensitivity to flow representation and biological traits. ICES J. Mar. Sci. 79, 2043–2054. https://doi.org/10.1093/icesjms/fsac135.
Segovia, M., Garcia-Alfaro, J., 2022. Design, modeling and implementation of digital twins. Sensors 22, 5396. https://doi.org/10.3390/s22145396.
Simmonds, J., MacLennan, D.N., 2005. Fisheries Acoustics: Theory and Practice. Blackwell Science Ltd.
Simon-Lledó, E., Amon, D.J., Bribiesca-Contreras, G., Cuvelier, D., Durden, J.M., Ramalho, S.P., Uhlenkott, K., Arbizu, P.M., Benoist, Noëlie, Copley, J., Dahlgren, T. G., Glover, A.G., Fleming, B., Horton, T., Se-Jong, Ju, Mejia-Saenz, A., McQuaid, K., Pape, E., Chailinn, Park, Smith, C.R., Jones, D.O.B., 2023. Abyssal NE Pacific Seafloor Megafauna Dataset. https://doi.org/10.5281/ZENODO.7982462.
Simons, R.A., 2017. ERDDAP: A Data Server that Gives you a Simple, Consistent Way to Download Subsets of Scientific Datasets in Common File Formats and Make Graphs and Maps. NOAA/NMFS/SWFSC. Available at: https://coastwatch.pfeg.noaa.gov/erddap/.
Song, A., Stojanovic, M., Chitre, M., 2019. Editorial underwater acoustic communications: where we stand and what is next? IEEE J. Ocean. Eng. 44 (1).
Stojanovic, M., Preisig, J., 2009. Underwater acoustic communication channels: propagation models and statistical characterization. IEEE Commun Mag 47 (1), 84–89. https://doi.org/10.1109/MCOM.2009.4752682.
Teeneti, C.R., Truscott, T.T., Beal, D.N., Pantic, Z., 2019. Review of wireless charging systems for autonomous underwater vehicles. IEEE J. Ocean. Eng. 46 (1), 68–87. https://doi.org/10.1109/JOE.2019.2953015.
Treml, E.A., Halpin, P.N., Urban, D.L., Pratson, L.F., 2008. Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. Landsc. Ecol. 23, 19–36. https://doi.org/10.1007/s10980-007-9138-y.
Tzachor, A., Hendel, O., Richards, C.E., 2023. Digital twins: a stepping stone to achieve ocean sustainability? Npj Ocean Sustain. 2, 16. https://doi.org/10.1038/s44183-023-00023-9.
Tzilivakis, K., 2022. A Tale of Two Oceans: Scientists Are Building Digital Twins of the Ocean.
Villon, S., Mouillot, D., Chaumont, M., Darling, E.S., Subsol, G., Claverie, T., Villéger, S., 2018. A deep learning method for accurate and fast identification of coral reef fishes in underwater images. Ecol. Inform. 48, 238–244. https://doi.org/10.1016/j.ecoinf.2018.09.007.
Villon, S., Mouillot, D., Chaumont, M., Subsol, G., Claverie, T., Villéger, S., 2020. A new method to control error rates in automated species identification with deep learning algorithms. Sci. Rep. 10, 10972. https://doi.org/10.1038/s41598-020-67573-7.
Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C., Hui, F.K.C., 2015. So many variables: joint modeling in community ecology. Trends Ecol. Evol. 30, 766–779. https://doi.org/10.1016/j.tree.2015.09.007.
Weaver, P.P.E., Aguzzi, J., Boschen-Rose, R.E., Colaço, A., De Stigter, H., Gollner, S., Haeckel, M., Hauton, C., Helmons, R., Jones, D.O.B., Lily, H., Mestre, N.C., Mohn, C., Thomsen, L., 2022. Assessing plume impacts caused by polymetallic nodule mining vehicles. Mar. Policy 139, 105011. https://doi.org/10.1016/j.marpol.2022.105011.
Weibel, S.L., Koch, T., 2000. The Dublin Core Metadata Initiative: Mission, Current Activities, and Future Directions. -Lib Mag, 6. https://doi.org/10.1045/december2000-weibel.
Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 'T Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., Van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR guiding principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18.
Wilkinson, D.P., Golding, N., Guillera-Arroita, G., Tingley, R., McCarthy, M.A., 2021. Defining and evaluating predictions of joint species distribution models. Methods Ecol. Evol. 12, 394–404. https://doi.org/10.1111/2041-210X.13518.
Xu, W., Matzner, S., 2018. Underwater fish detection using deep learning for water power applications. In: In: 2018 International Conference on Computational Science and Computational Intelligence (CSCI). Presented at the 2018 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, Las Vegas, NV, USA, pp. 313–318. https://doi.org/10.1109/CSCI46756.2018.00067.
Yu, Z., Du, P., Yi, L., Luo, W., Li, D., Zhao, B., Yuan, L., 2024. Coastal Zone Information Model: A Comprehensive Architecture for Coastal Digital Twin by Integrating Data, Models, and Knowledge. In Press, Fundamental Research. https://doi.org/10.1016/j.fmre.2024.06.003.
Zhang, Y., Ryan, J.P., Kieft, B., Hobson, B.W., McEwen, R.S., Godin, M.A., Harvey, J.B., Barone, B., Bellingham, J.G., Birch, J.M., Scholin, C.A., Chavez, F.P., 2019. Targeted sampling by autonomous underwater vehicles. Front. Mar. Sci. 6, 415. https://doi.org/10.3389/fmars.2019.00415.
Zhang, Y., Ryan, J.P., Hobson, B.W., Kieft, B., Romano, A., Barone, B., Scholin, C.A., 2021. A system of coordinated autonomous robots for Lagrangian studies of microbes in the oceanic deep chlorophyll maximum. Science. Robotics 6 (50), eabb9138. https://doi.org/10.1126/scirobotics.abb913.
Zurowietz, M., Nattkemper, T.W., 2020. Unsupervised Knowledge Transfer for Object Detection in Marine Environmental Monitoring and Exploration, 8. IEEE Access, pp. 143558–143568. https://doi.org/10.1109/ACCESS.2020.3014441.
Zurowietz, M., Langenkämper, D., Hosking, B., Ruhl, H.A., Nattkemper, T.W., 2018. MAIA—A machine learning assisted image annotation method for environmental monitoring and exploration. PloS One 13, e0207498. https://doi.org/10.1371/journal.pone.0207498.

Glossary

ROVs: remotely operated vehicles
AUVs: autonomous underwater vehicles
ASVs: autonomous surface vessels
DTO: digital twin of the ocean
PTO: physical twin observer
F-PTOs: fixed, local physical twin observer
G-PTOs: geographically expanded physical twin observer
EMSO: European Multidisciplinary Seafloor and Water-column Observatories
PDDL: planning domain definition language
ROS: Robotic Operating System
CNNs: convolutional neural networks
DBs: data bubbles
WGNEPS: Working Group on *Nephrops norvegicus*
FAIR: findability, accessibility, interoperability, and reusability
API: application programming interface
OGC: open geospatial consortium
GUIs: graphical user interfaces
eDNA: environmental DNA
JSDM: Joint Species Distribution Model
IBM: individual-based model