frontiers | Research Topics

Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis

Edited by
Sophie G. Pitois, Elaine Fileman, Mark C. Benfield, Peter H. Wiebe and Fabien Lombard

Published in
Frontiers in Marine Science

ISSN 1664-8714
ISBN 978-2-83251-642-3
DOI 10.3389/978-2-83251-642-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis

Topic editors

Sophie G. Pitois — Centre for Environment, Fisheries and Aquaculture Science (CEFAS), United Kingdom

Elaine Fileman — Plymouth Marine Laboratory, United Kingdom

Mark C. Benfield — Louisiana State University, United States

Peter H. Wiebe — Woods Hole Oceanographic Institution, United States

Fabien Lombard — Sorbonne Universités, France

Citation

Pitois, S. G., Fileman, E., Benfield, M. C., Wiebe, P. H., Lombard, F., eds. (2023). *Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-642-3

Table of contents

frontiers | Frontiers in Marine Science

TYPE Editorial
PUBLISHED 26 January 2023
DOI 10.3389/fmars.2023.1100998

OPEN ACCESS

EDITED AND REVIEWED BY
Won Sang Lee,
Korea Polar Research Institute, Republic of Korea

*CORRESPONDENCE
Sophie G. Pitois
sophie.pitois@cefas.gov.uk

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 17 November 2022
ACCEPTED 16 January 2023
PUBLISHED 26 January 2023

CITATION
Pitois SG, Fileman ES, Benfield MC, Wiebe PH and Lombard F (2023) Editorial: Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis.
Front. Mar. Sci. 10:1100998.
doi: 10.3389/fmars.2023.1100998

Editorial: Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis

Sophie G. Pitois[1]*, Elaine S. Fileman[2], Mark C. Benfield[3], Peter H. Wiebe[4] and Fabien Lombard[5,6]

[1]Centre for Environment, Fisheries and Aquaculture Science (CEFAS), Lowestoft, United Kingdom, [2]Plymouth Marine Laboratory, Plymouth, United Kingdom, [3]Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, LA, United States, [4]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA, United States, [5]Laboratoire d'Océanographie de Villefranche, UMR 7093, Sorbonne Université, CNRS, Villefranche-sur-Mer, France, [6]Institut Universitaire de France (IUF), Paris, France

KEYWORDS

imaging, molecular tools, global monitoring, AI based taxonomy, computer vision, data pipeline, real-time monitoring, plankton observatory network

Editorial on the Research Topic

Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis

The traditional collection of plankton samples, often using nets followed by visual sorting and taxonomic analysis is a labour intensive, time-consuming, and ultimately expensive process. The increasing demand for pelagic data combined with ever reducing budgets for monitoring and the general problem of the taxonomic impediment have driven the development of new tools and techniques for the sampling and analysis of this key ecosystem component.

Technological advances have allowed monitoring of the pelagic environment to evolve towards a more integrated approach to data collection at a finer scale. Nowadays, there is a plethora of methods including for example molecular, optical, remote sensing, and automated techniques that help further our understanding of biodiversity and species interactions within the pelagic ecosystem. Yet, there is still a need for morphological identification of species to enable the verification of the new sensor modalities.

Ongoing and new challenges need to be addressed, for example those associated with taxonomic resolution from image analysis methods; or where the rate of collecting data exceeds that of processing it, a situation arising from collection of huge amounts of fine-resolution data. Furthermore, there is a large variety of data formats including for example, images, acoustics, taxonomy, remote sensing etc … collected at different spatial and temporal scales. There is therefore a need to: (1) improve on how these emerging technologies,

collecting multi-modal data, work together; (2) develop new methods embedding technological advances in data collection and data analytics, for a fully integrated understanding of pelagic processes; and (3) utilise open access databases that can handle diverse data sets.

Articles presented in this topic have a strong focus on:

(1) imaging instruments both for collected samples (e.g., PlanktoScope (Mériguet et al.; Pollina et al.), FlowCam (Mériguet et al.), and *in-situ* deployment (e.g., Underwater Vision Profiler (Barth and Stone; Drago et al.), Plankton Imager (PI) (Giering et al.), *In Situ* Ichthyoplankton Imaging System (ISIIS) (Panaïotis et al.), Scripps Plankton Camera (SPC) (Le et al.), Imaging FlowCytobot (Kraft et al.); Video Plankton Recorder (VPR) (Plonus et al.) and the use of machine learning tools to automatically classify the collected images (e.g., Weldrick).

(2) molecular tools (DNA barcoding (de Vargas et al.), COI metabarcoding (Bucklin et al.) to explore the taxonomic and ecological diversity.

These tools have their own characteristics, with molecular based techniques focussing on taxonomic resolution and imaging instruments on quantifying plankton abundance, biomass, and morphology, but at a lower taxonomic resolution. Both have played an increasing role in the collection of plankton information, and the articles in this collection suggest that this will continue to be the case as technologies are refined and become more affordable. As affordability and ease of use in technologies increases, along with the advances and reliability of machine learning algorithms, it will become possible to move towards consistent and long-term measurements of plankton abundance/biomass and diversity.

Globally, plankton time-series have provided essential information about how planktonic assemblages and constituent taxa respond to climate change. However, more data are needed to inform these powerful tools. Ships of opportunity and citizen science have an increasing role to play, by deploying easy to use instruments and offering an opportunity to collect cost-effective plankton data on a global scale (Pollina et al.; de Vargas et al.).

Imaging systems and traditional nets collect data on scales and from within volumes that are very different. Consequently, comparisons of data collected from imaging systems and traditional methods need further attention if we are to inter-calibrate their measurements (Le et al.). Further improvements to the performance of automated image analysis for taxonomic identification are also needed. This will help to reconcile findings with traditional net-based methods (Barth and Stone), their complementary use, and build trust in Artificial Intelligence (AI) based taxonomy (Giering et al.).

All the methods using novel technologies described in this collection of papers allow for plankton information to be collected at a much higher rate than previously possible using traditional deployment of nets followed by microscopic analysis. The increasing use of digital imaging and molecular tools will enable plankton abundances and biomass distributions, as well as community composition and taxonomic diversity to be mapped at global scales in a much shorter time frame (Drago et al.). Used in combination, they offer new opportunities to monitor and study plankton ecosystems at levels of detail never possible before.

A key element of an ever-increasing amount of data is the development of automated and efficient data processing techniques for seamless data flow. Ship time remains an expensive commodity and it is important to be able to capitalise on information collected at sea in real-time. Analysing continuous streams of high-frequency data calls for development and deployment of novel computer vision and machine learning systems (Panaïotis et al.; Le et al.), while cloud platform and high-performance computing are needed for processing the huge datasets collected. Together they can open new horizons for testing core hypotheses on plankton communities in aquatic ecosystems. Seamless data pipeline and emerging data analytics based on AI further offer the opportunity to move towards real time monitoring of plankton ecosystems (Kraft et al.).

Cost-effective global monitoring will enable us to obtain a complete picture of plankton composition, biogeography and biogeochemistry, opening the way to a plankton observatory network on a planetary scale (Mériguet et al.; Pollina et al.; Drago et al.; de Vargas et al.).

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

FL recieved funding from the European Union's Horizon 2020 research and innovation programme "Atlantic Ecosystems Assessment, Forecasting and Sustainability" (AtlantECO) under grant agreement No 862923 and the JERICO-S3 project under grant agreement No 871153. EF was funded through the UK Natural Environment Research Council's National Capability Long-term Single Centre Science Programme, Climate Linked Atlantic Sector Science, grant number NE/R015953/1, a contribution to Theme 1.3–Biological Dynamics.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

frontiers in Marine Science

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fmars.2021.754375

Automatic Segregation of Pelagic Habitats

Rene-Marcel Plonus[1], Stefanie Vogl[2] and Jens Floeter[1]*

[1] Institute of Marine Ecosystem and Fishery Science, University of Hamburg, Hamburg, Germany, [2] Department of Informatics and Mathematics, Hochschule für Angewandte Wissenschaften München, Munich, Germany

It remains difficult to segregate pelagic habitats since structuring processes are dynamic on a wide range of scales and clear boundaries in the open ocean are non-existent. However, to improve our knowledge about existing ecological niches and the processes shaping the enormous diversity of marine plankton, we need a better understanding of the driving forces behind plankton patchiness. Here we describe a new machine-learning method to detect and quantify pelagic habitats based on hydrographic measurements. An Autoencoder learns two-dimensional, meaningful representations of higher-dimensional micro-habitats, which are characterized by a variety of biotic and abiotic measurements from a high-speed ROTV. Subsequently, we apply a density-based clustering algorithm to group similar micro-habitats into associated pelagic macro-habitats in the German Bight of the North Sea. Three distinct macro-habitats, a "surface mixed layer," a "bottom layer," and an exceptionally "productive layer" are consistently identified, each with its distinct plankton community. We provide evidence that the model detects relevant features like the doming of the thermocline within an Offshore Wind Farm or the presence of a tidal mixing front.

Keywords: machine learning, North Sea, submesoscale, pelagic habitats, plankton patchiness

OPEN ACCESS

Edited by:
Tracey T. Sutton, Nova Southeastern University, United States

Reviewed by:
Adolf Konrad Stips, European Commission, Italy
Marco Uttieri, Anton Dohrn Zoological Station, Italy

****Correspondence:***
Rene-Marcel Plonus
rene-marcel.plonus@uni-hamburg.de; r.plonus@web.de

Specialty section:
This article was submitted to Marine Ecosystem Ecology, a section of the journal Frontiers in Marine Science

Received: *06 August 2021*
Accepted: *15 September 2021*
Published: *04 October 2021*

Citation:
Plonus R-M, Vogl S and Floeter J (2021) Automatic Segregation of Pelagic Habitats. Front. Mar. Sci. 8:754375. doi: 10.3389/fmars.2021.754375

INTRODUCTION

Submesoscale features like eddies, fronts or filaments structure the pelagic realm at spatial scales of order (1–10 km) (Lévy et al., 2012; Shulman et al., 2015; Buckingham et al., 2016) and temporal scales that range from several hours to a few days (Baschek and Maarten Molemaker, 2010; Thompson et al., 2016). Associated processes determine nutrient fluxes (Omand et al., 2015; Thompson et al., 2016) as well as plankton patchiness (Levy and Martin, 2013; Shulman et al., 2015; Lévy et al., 2018) and thereby even shape the seascape for top predators like sea birds (Bertrand et al., 2014).

Recent advances in marine remote sensing technology (Wedding et al., 2011) enabled scientists to separate benthic structures into mosaic-like patterns of different habitat classes (Hinchey et al., 2008; Pittman et al., 2011) following the role model of terrestrial ecosystems. However, what is well known and trivial in landscape ecology can be quite challenging in seascape ecology. While it remains difficult to segregate pelagic habitats, which exhibit no clear boundaries (Hinchey et al., 2008; Pittman et al., 2011; Wedding et al., 2011) and can be quite dynamic on a wide range of scales, benthic habitat maps can give an impression of physically distinct areas that consistently occur together with particular species communities (Harris and Baker, 2012). Some effort has been undertaken to characterize fish habitats (e.g., Bellido et al., 2008; Giannoulaki et al., 2011; Tugores et al., 2011; Laman et al., 2017; Amorim et al., 2018; Friedland et al., 2020; Funk et al., 2020),

but fewer studies focused on zooplankton (e.g., Labat et al., 2009; Alvarez-Berastegui et al., 2014; Espinasse et al., 2014). Thus, mechanisms contributing to the enormous diversity of plankton, a fundamental component of pelagic food webs, are still not fully understood (Sano et al., 2013; North et al., 2016). Understanding the processes shaping plankton communities is essential to improve our knowledge of existing ecological niches (Houliez et al., 2021). Despite the growing awareness of the importance of spatial structure for ecology and management (Pittman et al., 2011; Wedding et al., 2011), there is still a lack of concepts and techniques applicable to characterize the spatial structure of the seascape in pelagic environments (Alvarez-Berastegui et al., 2014). Mainly, because traditional oceanographic methods are inadequate for observing the submesoscale (Baschek and Maarten Molemaker, 2010) due to insufficient resolution and range (Marmorino et al., 2018). Recent advances in instrumentation partially closed this gap, but there still is a need for novel analysis methods to take advantage of the existing data (North et al., 2016). Some machine learning techniques are specifically designed to identify and characterize features in a "sea of data," which makes it very promising to apply them also in this challenging field of research.

Autoencoders (AE) are a common tool in the machine learning community which consist of an encoding and a decoding part (Hinton, 2006). Initially devised to reduce (Encoder) and recover (Decoder) the dimensionality of their inputs (Hinton, 2006), they have been soon applied to a wide range of tasks like denoising (e.g., Vincent et al., 2010) or anomaly detection (e.g., Zhao et al., 2017; Chen et al., 2018).

Autoencoders do not classify or detect specific elements or objects in their inputs, but learn meaningful low dimensional representations, i.e., relevant high-level abstractions, of their inputs (Bengio et al., 2006) so that the original data can be reconstructed as similar as possible by the decoder part. The input data don't need any pre-processing, e.g., labeling of subsets, by humans, since the target the network aims to reconstruct is basically the original input. The compressed representations of the encoder can also be used as input for subsequent modeling, e.g., in a Convolutional Neural Network (CNN) application. In that case the unsupervised pre-training of a CNN embedded in an AE can help to capture more intricate dependencies (Erhan et al., 2009) and better initialize the weights of the extended model (Bengio et al., 2006). Thus, the (local) minimum in the loss surface of the AE corresponds to a good transformation of a high dimensional input to a lower dimensional intermediate output (output of the Encoder-part) (Bengio et al., 2006), which would become the input for the classifier in a CNN. In this setting, the final output of the AE, the reconstructions, are secondary. However, a low reconstruction error of the AE ensures that the compressed signal incorporates the important features of the original high dimensional input data.

In this study we take advantage of this specific application of AEs. Instead of substituting the decoder part with a classification or regression network we use the compressed signal of the encoder as input for a subsequent clustering algorithm. We use a fully connected AE to reduce a high dimensional input consisting of a variety of abiotic and biotic oceanographic measurements to a lower dimensional meaningful representation (intermediate output), skip the decoding part after the training is completed and cluster the encoded features to macro-habitats. Similar micro-habitats lead to similar representations and therefore regions with different characteristics are segregated as different macro-habitats. These macro-habitats correspond to distinct pelagic habitats in the southern North Sea, whose plankton communities are compared and analyzed.

MATERIALS AND PROCEDURES

Data Acquisition and Preparation

Physical and biological oceanographic measurements were recorded on a North Sea summer cruise with the RV Heincke (HE429, July 19–24, 2014) with a MacArtney TRIAXUS Remotely Operated Towed Vehicle (ROTV). For a detailed description of the device see Plonus et al. (2021). The ROTV transects were located in the direct vicinity of two Offshore Wind Farms (OWF) BARD Offshore 1 (BARD) and Global Tech I (GTI; **Figure 1**). The map was generated using QGIS v3.18 (QGIS, 2021) with bathymetric metadata and Digital Terrain Model data products from the EMODnet Bathymetry portal (15.7.21)[1]. The ROTV was towed at a speed of 8 knots (4.1 m s^{-1}) with a three-degree lateral offset to lessen any disturbance from the vessels wake. During most transects the ROTV was undulating with a vertical speed of 0.1 m s^{-1} from $\sim$ 4 m below the sea surface to $\sim$ 8 m above the sea floor. The horizontal resolution between two surface peaks was $\sim$ 560 m, while the vertical resolution was $\sim$ 0.3 m. The ROTV measured water temperature, salinity, oxygen, and chlorophyll-a at a frequency of 1 Hz and was equipped with a Video Plankton Recorder (VPR, Seascan Inc., Falmouth, MA, United States) which provided zoo- and phytoplankton densities on the taxonomic family-, and sometimes even genus-level. For a detailed description of the VPR plankton image classification see Floeter et al. (2017). We used a similar summer cruise with the RV Heincke 5 years later (HE534, June 16–21, 2019) as a test data set. For our analyses we selected the following variables: temperature (°C), salinity (PSU), oxygen (μmol l^{-1}), density (kg $\times$ m^{-1}), and chlorophyll-a (RFU). For each of the variables, we calculated the horizontal (grid cell to the left, i.e., $\sim$ 25 m) and vertical (grid cell above, i.e., 1 m) gradient. Furthermore, we had sufficient density data (N l^{-1}) available for the taxa "Appendicularia," "Copepoda," "Dinoflagellates," "Gastropoda," "Jelly," "Marine snow," "Nauplii," "Ophiuroida," "Pilidium," "Pluteus," and "Polychaeta."

Transect diagrams were generated using Ocean Data View (ODV, Schlitzer, 2020) with the embedded spatial interpolation software DIVA (Troupin et al., 2012) and exported as grids with a resolution of $\sim$ 25 m length $\times$ 1 m depth. Abiotic measurements as well as density values were normalized and rescaled to range from -1 to 1. This was necessary since deep learning models generally perform better with homogeneous, small values (Bishop, 1995).

To check for multicollinearity between our variables we calculated the variance inflation factor (VIF) in R

[1]http://www.emodnet-bathymetry.eu

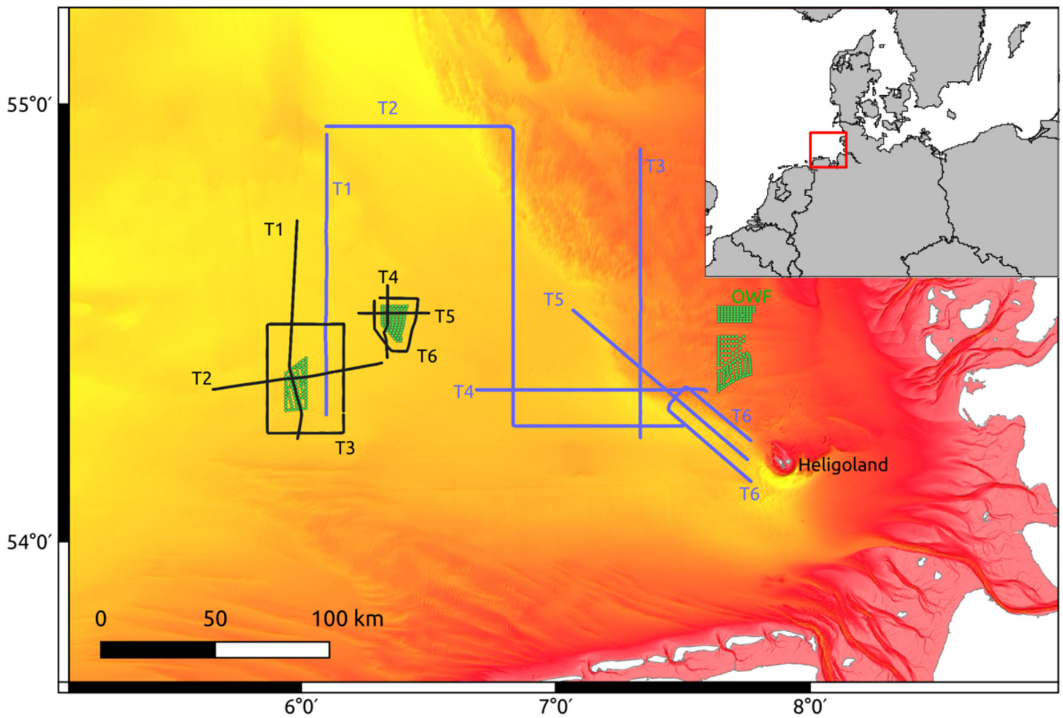


FIGURE 1 | Sampling transects from HE429 (black) and HE534 (blue) in the German Bight of the North Sea. Green dots: Wind turbines. Depth ranges from 10 m (red) to 50 m (yellow). The red box marks the location of the bigger map.

(R Core Team, 2020) using functions provided by Zuur et al. (2009). A threshold of VIF > 3 was applied to identify highly collinear variables and exclude them from further analyses (Zuur et al., 2010). The exported grids for each selected parameter were stacked and transformed into feature-vectors where each grid cell became one vector with four features (1 parameter = 1 feature). In our definition, a pelagic micro-habitat with a spatial extent of ∼25 m × 1 m corresponds to one of those feature-vectors (**Figure 2**).

Based on these feature-vectors the AE was trained to reconstruct the original micro-habitats and thereby learn relevant abstractions that represent important patterns in the pelagic environment. We used a GPU supported TensorFlow backend (Abadi et al., 2015) for Keras (Chollet, 2015) under Python 3.7 (Van Rossum and Drake, 2009) to build and train our AE.

Model Description

The AE consisted of two fully connected layers in the Encoder and Decoder, respectively. The Decoder used the transposed weights of the Encoder in reversed order, e.g., the weights of the first Encoder-Layer were shared with the last Decoder-Layer. The first layer of the Encoder inflated the 4-dimensional feature-vector to a 100-dimensional feature-vector, which was reduced to a 2-dimensional feature-vector by the second layer (4 – 100 – 2). The Decoder did the same in reverse (2 – 100 – 4). The batch size (number of inputs that are processed simultaneously) was set to 38 and the learning rate followed a sawtooth-like scheme, initialized at $5e^{-8}$. Each input feature-vector corresponds to one micro-habitat and includes 1 measurement of each parameter selected for the analyses. The model was trained using the data from HE429 exclusively. Approximately ∼ 13% of the data was separated to validate the training process based on the remaining 87%. Data from HE534 was used as a final test set. As an AE is a gradient-based method, the chosen starting point may be crucial for the final fit of the model (Hinton, 2006), and one way of assessing and reducing the effect of start conditions are multi-start approaches (Subbey, 2018). Therefore, we trained multiple models and selected the one with the smallest final validation RMSE.

Habitat Segregation

By applying the trained Encoder only, we projected all micro-habitats into a *xy*-coordinate system using the 2-dimensional intermediate output. We will refer to the encoded outputs as "Encoded Components" in the following. Micro-habitats with similar characteristics were projected closer to each other than micro-habitats with different characteristics. We used the Euclidean distance to calculate the dissimilarity matrix for the Encoded Components of the micro-habitats, which was clustered by the HDBSCAN algorithm (McInnes et al., 2017). HDBSCAN uses a density-based linkage function, defining clusters by the size of the area in which a certain number of neighbors is found. Micro-habitats in "sufficiently dense" regions were assigned to a macro-habitat. Obviously, the parameters "size of the neighborhood" (Epsilon) and the "critical number of neighbors" (min_samples) are determining the resulting clusters

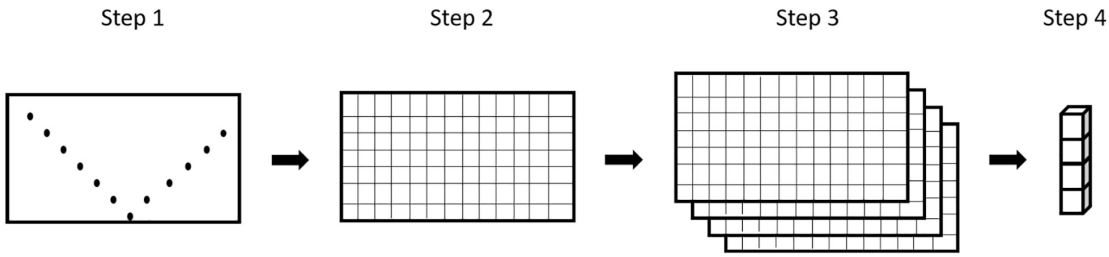


FIGURE 2 | Schematic data processing from measurements to feature-vector. Step 1: original measurements; Step 2: gridded parameter table; Step 3: stacked grids; Step 4: feature-vector for 1 of the 91 grid-cells.

(dense areas) with HDBSCAN. Thus, we checked the resulting macro-habitats for multiple different combinations of these two parameters as well as "min_cluster_size." The parameter "min_cluster_size" is the threshold that separates "sufficiently dense" regions (clusters) from the random background noise. All micro-habitats that were not assigned to a specific macro-habitat by HDBSCAN got the label "−1." Homogeneous regions in the transect produced more dense regions in the 2-dimensional surface that were more likely to trespass the "min_cluster_size" threshold and were separated from other homogeneous water masses by less dense regions. We used the silhouette method (Rousseeuw, 1987) to select the best segregation of micro-habitats. The silhouette score ranges from "−1" to "1" and indicates how well each point fits into the assigned cluster (macro-habitat) and is one of the best performing indices available (Arbelaitz et al., 2013). "−1" is probably wrong labeled, "0" is close to the decision boundary of two clusters and "1" means this specific point is far away from points of other clusters. The silhouette scores were calculated using the scikit-learn module (Pedregosa et al., 2011) for python.

Analyses

We used ODV to add a transect plot of the identified macro-habitats to the original measurements and plankton densities. Isolines of selected parameter measurements were overlayed on the macro-habitat plots to investigate which feature characteristics contributed to the segregation and to assess the associated plankton communities.

We furthermore described the macro-habitat plankton communities by modified Species-Abundance-Plots (SAP). We calculated the relative number of micro-habitats by plankton density and taxonomic group for each cluster. As is common for SAPs we used a log2 scale for density. That way we visualized the shift in specific species densities between the macro-habitats of different segregations of the same ROTV survey transect.

Pelagic submesoscale features often are highly productive areas and aggregate particles (Levy and Martin, 2013; Lévy et al., 2018). Therefore, we calculated Lloyd's mean crowding (Lloyd's MC) and Lloyd's index of patchiness (Lloyd's IP) with the R-function "agg_index" from the "epiphy" package (Gigot, 2018) and compared the results for different segregations of the same transects. Lloyd's index is >1 if species were aggregated, 1 if the distribution is random and <1 indicates an overdispersed distribution compared to a homogeneous distribution. The Index of aggregation proposed by Bez (2000) (Bez's IoA) was calculated in addition to Lloyd's IP.

Data handling was done with R (R Core Team, 2020) and some tidyverse packages (Wickham et al., 2019), namely purrr, tibble, dplyr, ggplot2, and tidyr.

RESULTS

In the initial VIF analysis with the full dataset a couple of parameters had VIF > 3. After removing "density" which had the highest score, no further parameter exceeded this threshold (**Supplementary Table 1**). After a detailed analysis of model sensitivities and reconstruction quality we decided to limit the final parameter selection to (1) vertical temperature difference to the grid cell above, (2) salinity, (3) oxygen, and (4) chlorophyll-a concentration.

Model Training

The root mean squared error (RMSE) after the first epoch ranged roughly between 0.7 and 1.0. Each training epoch took 10–15 s using a graphic card with 768 gpu-cores and we trained each model for 15 epochs until a plateau was reached (**Supplementary Figure 1**). The final training and validation RMSE of our selected model were $RMSE_{Tr} \sim 0.33$ and $RMSE_{Val} \sim 0.35$ (**Figure 3**).

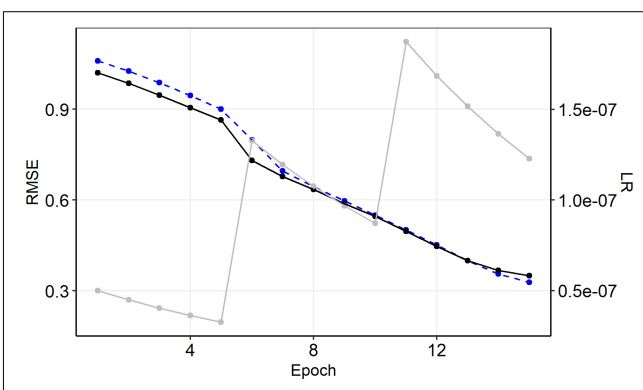


FIGURE 3 | History of model training. Black: validation, blue: training, gray: learning rate (sec. axis).

Clustering

Depending on the HDBSCAN parameter selection, micro-habitats were grouped into 2–20 macro-habitats that ranged in size from <0.1 to 97% of all micro-habitats in a transect. We present exemplary the results for the segregation of T3 into different numbers of macro-habitats. Different parameter combinations could lead to an identical number of segregations. We chose an inverse size-cluster-relationship in the figure since more macro-habitats were usually ecologically less plausible (**Figure 4**). Mostly, "epsilon" had a great impact on the segregation with specific combinations of "min_cluster_size" and "min_samples" but less influence with other tested combinations of those two parameters, indicating that segregations changed discontinuously with slopes and plateaus (**Supplementary Figure 2**).

Even though the label "−1" is used by HDBSCAN to indicate the lack of belonging to a specific cluster, we observed a close relationship between micro-habitats labeled "−1" and exceptionally strong chlorophyll-peaks throughout all transects. Therefore, we decided to treat "−1" as a macro-habitat of its own instead of unclassified micro-habitats. Micro-habitats labeled as "−1" were also frequently located between the BL and the SL.

Projections

While cluster-labeling was not consistent in that cluster "0" always referred to, e.g., the "surface mixed layer," the projections of the "surface mixed layer" micro-habitats were always located in a similar position throughout all projection plots. Thus, while the cluster denotations related to a macro-habitat were not consistent, the position indicated the affiliation to a specific macro-habitat (**Figure 5**).

Silhouette Method

The segregation into three macro-habitats gave the highest average silhouette-scores in most cases: notably high chlorophyll-peaks were merged into one macro-habitat (1) and two further macro-habitats were separated at around 17°C in an upper surface mixed layer (2) and a lower bottom layer (3). There was only one exception from this rule in T1 where in the northern, deeper area the bottom layer (3) was replaced with the layer including the chlorophyll-peaks (1). Another anomaly occurred in T2, where one of the basic macro-habitats was further separated into two "sublayers" so that a total of four macro-habitats were segregated. The highest silhouette-scores ranged from 0.35 (T1) to 0.59 (T5) (**Table 1**).

In the following we will use abbreviations for the three main layers and their sublayers, namely "PL" for the productive layer with the high chlorophyll values, "BL" for the bottom layer and "SL" for the surface mixed layer. A segregation into more than one layer is indicated using numbers, e.g., "SL1"/"SL2" instead of "SL."

Habitat Maps

We present T2 exemplarily for all transects of HE429 (**Figure 6**). Segregating the output of the Encoder into 3 macro-habitats, we got the typical scheme of a SL with temperatures above 17°C, a macro-habitat which was strongly associated with extraordinary high chlorophyll-peaks (PL) and a BL as a third macro-habitat. The average silhouette-score for the entire transect was 0.54 (**Table 1**). However, this clustering did not account, e.g., for the intrusion of marine snow particles into the SL in the eastern half of the transect. When accepting 4 different macro-habitats, the BL and PL macro-habitats were mostly unaffected, while the SL was further separated into 2 different macro-habitats. One corresponded to the area where marine snow particles were predominant while the second macro-habitat corresponded to the area where pluteus larvae were observed in high densities. Segregating characteristics of the two macro-habitats were a salinity difference of 0.2 and a shallowing of the thermocline from 10 m to 5 m water depth. Notably, this change around section distance 18–20 km was located at the entry point of the transect into the Offshore Wind Farm BARD. This segregation increased the average silhouette-score for the entire transect to 0.56 (**Table 1**).

Species Abundance Plots

The segregation into four macro-habitats was further supported by the modified SAPs. The relative amounts of PL and

FIGURE 4 | Number of segregated macro-habitats. The Figure was produced with Epsilon = 0.32. Yellow cross: selected segregation.

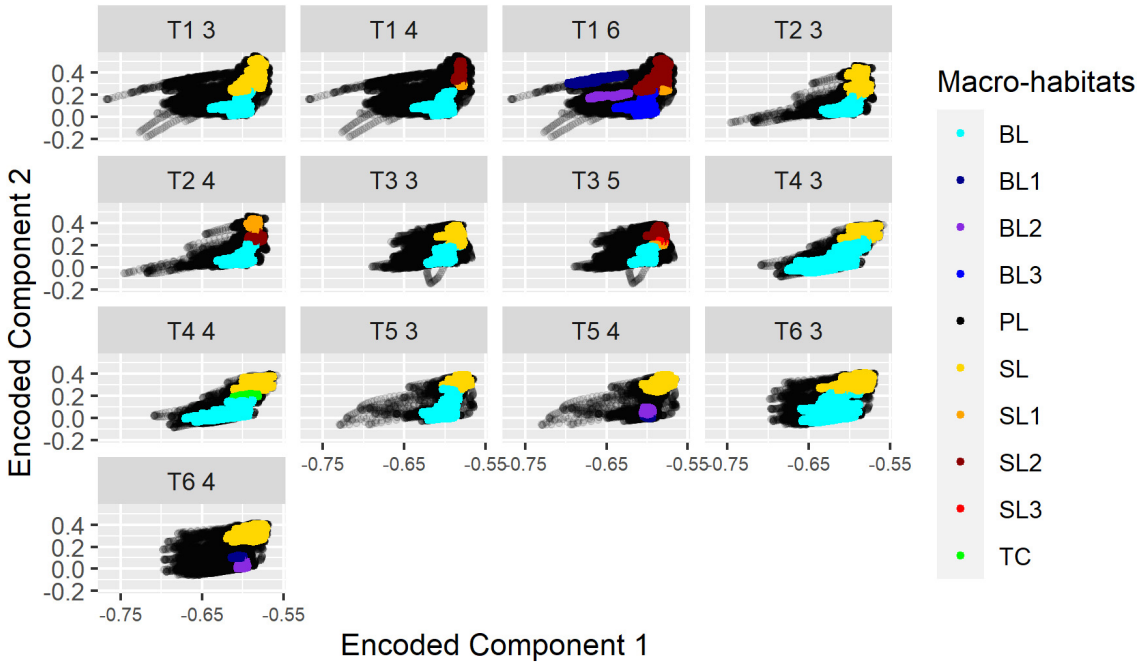


FIGURE 5 | Projections for different segregations of all transects from HE429. The number of segregated macro-habitats is indicated by the last digit in the panel header. BL, Bottom layer (blue); SL, Surface mixed layer (yellow-red); TC, Thermocline Layer; PL, Productive layer (black). Numbers were used to indicate that, e.g., more than 1 "Bottom layer" was segregated (e.g., B1/B2 instead of BL).

BL did not change much between 3 and 4 macro-habitats. However, SL1 included all micro-habitats with copepod densities >8 N l^{-1} and basically all micro-habitats where pluteus larvae occurred. SL2 instead included micro-habitats with copepod densities <8 N l^{-1} and generally less chlorophyll, but most micro-habitats where Appendicularia occurred. Thus, the SL1 and SL2 plankton communities were clearly distinct (**Figure 7**).

TABLE 1 | Silhouette method to select the best clustering.

Transect	Silhouette score	Segregation
HE429 T1	0.35	14583_21220_22015
HE429 T1	0.11	24657_21940_1949_9272
HE429 T1	0.07	12983_596_682_589_21028_21940
HE429 T2	0.54	5221_24912_12213
HE429 T2	0.56	6548_25232_5138_5428
HE429 T3	0.44	18748_21014_56730
HE429 T3	0.34	21660_56167_1371_15859_1435
HE429 T4	0.46	2054_5928_8107
HE429 T4	0.39	3202_5928_605_6354
HE429 T5	0.59	1729_4823_9331
HE429 T5	0.39	4582_5148_990_5163
HE429 T6	0.45	5558_14792_12965
HE429 T6	0.18	14589_12789_1229_4708
HE534 T1	0.40	7488_15344_53469
HE534 T2	0.26	21737_15535_123541
HE534 T2	−0.08	22248_4062_2099_132404
HE534 T2	−0.04	24788_15535_4115_1977_113215_1183
HE534 T3	−0.08	18944_2382_1156_1597_7140_9128

The numbers in "Segregation" give the number of micro-habitats by macro-habitat, e.g., X_Y_Z indicates 3 macro-habitats with X, Y, and Z micro-habitats, respectively.

Lloyd

Lloyd's mean crowding underpinned the SAP results. Patchiness in PL and BL did not change for "marine snow" and "pluteus" but differed clearly between SL1 and SL2, indicating a higher pluteus aggregation in SL1 and a higher aggregation of marine snow in SL2 (**Table 2**).

Test Dataset HE534

The temperature maximum during HE534 was around 15°C, i.e., 2°C lower than the threshold that separated SL and BL in HE429. Consequently, no thermal stratification was detected by the Encoder trained with HE429 measurements. However, this model segregated an oxygen-rich layer that, based on the projections of the Encoder, resembled a similar habitat as the SL in HE429. This oxygen driven stratifications were not consistent over the entire range of a transect and in some areas the macro-habitat with projections similar to BL in HE429 comprised the entire water column, indicating a mixed water column closer toward the coast. Notably, plankton aggregations were commonly located at the border between oxygen-stratified and mixed water columns (**Figure 8**). The highest average silhouette-scores were reached with three segregated macro-habitats. However, the scores were much lower compared to HE429 with a maximum between 0.26 and 0.40.

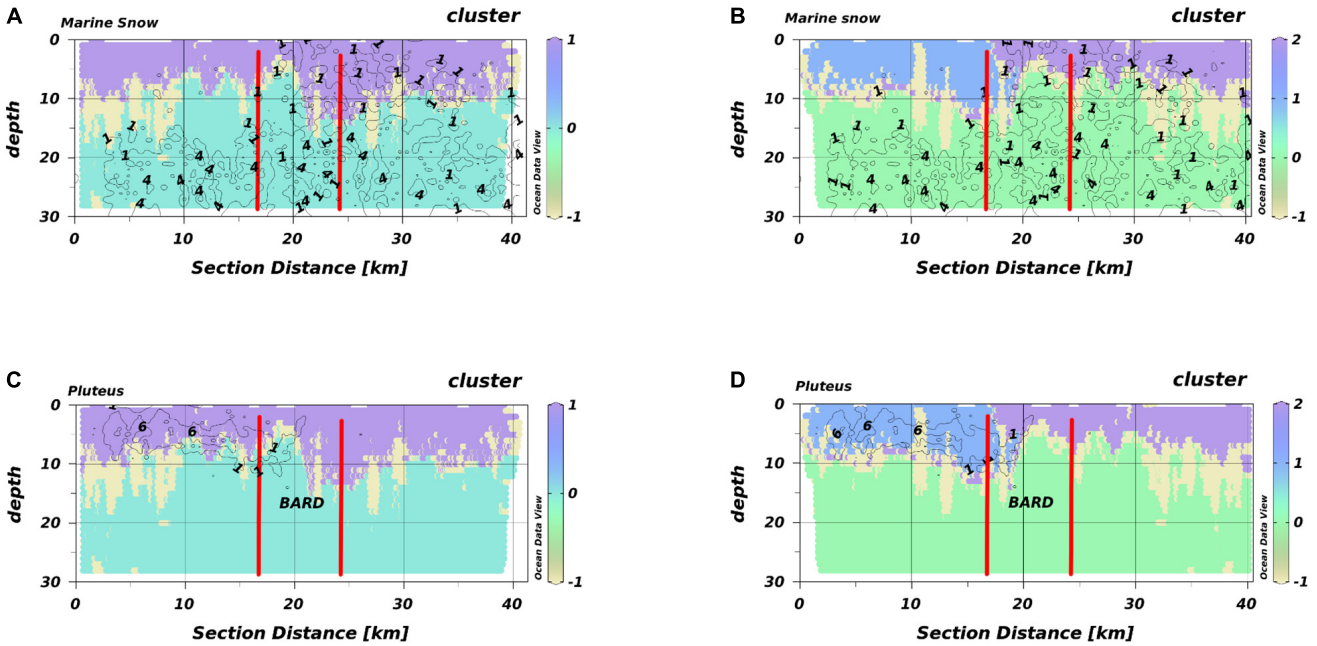


FIGURE 6 | HabitatMap of T2 (HE429) with marine snow density **(A,B)** and pluteus density **(C,D)** as isolines. The OWF BARD is located between the vertical red lines. **(A,C)** 3 segregated macro-habitats, **(B,D)** 4 segregated macro-habitats. Different colors represent different macro-habitats. Different numbers are the respective densities indicated by the isolines.

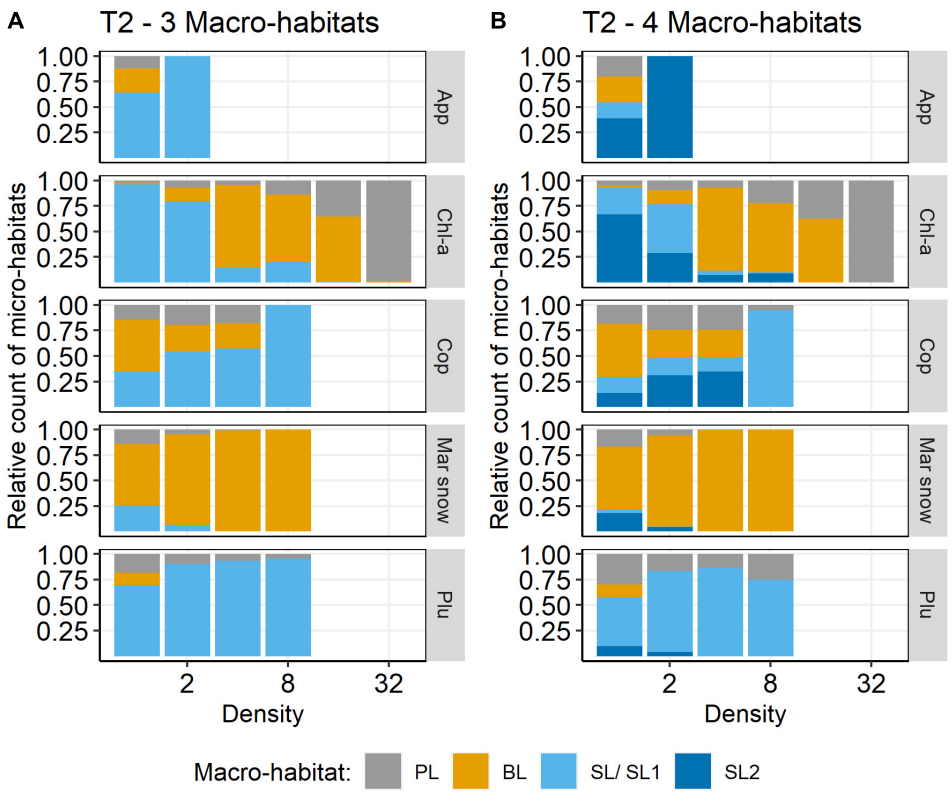


FIGURE 7 | Relative count of density by macro-habitat for selected species on transect T2 (HE429). **(A)** 3 macro-habitats, **(B)** 4 macro-habitats. App, Appendicularia (N l^{-1}); Chl-a, chlorophyll-a (RFU); Cop, Copepoda (N l^{-1}); Mar snow, marine snow (N l^{-1}); Plu, pluteus (N l^{-1}).

TABLE 2 | Lloyd's mean crowding and Bez's Index of aggregation for 3 and 4 segregated macro-habitats for transect T2 (HE429).

Number of macro-habitats	Macro-habitat	Plankton group	Lloyd's MC	Presence	Index of aggregation
3	PL	ms	0.79	0.37	5.61e-04
3	BL	ms	2.35	0.78	6.72e-05
3	SL	ms	0.43	0.25	2.67e-04
4	PL	ms	0.65	0.34	4.50e-04
4	BL	ms	2.35	0.77	6.66e-05
4	SL1	ms	0.23	0.07	1.02e-03
4	SL2	ms	0.43	0.41	4.14e-04
3	PL	plu	2.19	0.14	2.10e-03
3	BL	plu	0.00	0.01	6.75e-04
3	SL	plu	3.56	0.36	3.78e-04
4	PL	plu	3.39	0.20	1.51e-03
4	BL	plu	0.00	0.01	6.58e-04
4	SL1	plu	3.56	0.67	4.63e-04
4	SL2	plu	0.36	0.06	1.57e-03

Presence gives the relative number of micro-habitats with density >0. ms, marine snow; plu, pluteus.

DISCUSSION

Selection of Parameters

When training the model, we got the best results with a limited selection of parameters compared to the entire set of available data. The selected parameters are, however, in accordance with previous findings that physical properties contribute most to differences in habitat utilization by plankton organisms (Schulz et al., 2012; Friedland et al., 2020). In contrast to Alvarez-Berastegui et al. (2014), we did not benefit from the combination of gradients with the original measurements. However, a prior wavelet analysis as in North et al. (2016) could help to identify relevant spatial scales for the derivation of gradients. It is also possible, that the architecture of the model limited the amount of compressed information accessible to the clustering algorithm. In convolutional AEs, the size of the bottleneck (intermediate output) limits the generalization of the model (Manakov et al., 2019). This is also true for the fully connected AE architecture of this model and might limit the potential of including more variables like species densities and environmental gradients.

Reconstruction Loss

The loss for the optimization of an AE is based on the difference between the reconstruction and the original input. However, driving forces behind habitat partitioning vary with study region and season and specific parameters have a higher contribution than others (Schulz et al., 2012; Espinasse et al., 2014; Friedland et al., 2020). Thus, we deemed it more important to accurately reconstruct specific features (parameters) compared to entire vectors (micro-habitats). Accordingly, we calculated the sum of the batchwise RMSE between the specific feature-values (e.g., temperature) of each input and the corresponding feature-values of the reconstructions and not the RMSE of an entire feature-vector (micro-habitat) and its reconstruction. This forced the AE to learn all parameters individually and furthermore made it possible to give specific parameters a higher priority if appropriate.

Aggregation

Lloyd's IP is an area-related quality measure for Lloyd's MC and thus sensitive to zeros. As can be seen in our example (**Table 2**), the "spillover" from a crowded to an empty macro-habitat in the area of the decision boundary leads to misleadingly high Lloyd's IP, and to a lesser degree, misleadingly high Bez's IoA, even though this Index is supposedly insensitive to zeros. In accordance with the recommendation by Bez (2000) we therefore suggest Lloyd's MC as a measure of aggregation within a macro-habitat. Lloyd's IP and Bez's IoA might still be informative if the overall colonialization of the macro-habitat is considered.

Pelagic Habitats

The model segregated three (four) distinct pelagic habitats in HE429: (1) a SL (SL1/SL2) mainly characterized by temperatures >17°C, (2) a BL on the other side of that threshold, and (3) a PL dominated by high chlorophyll concentrations. In contrast to SL and BL, PL was not a true cluster by the definition of HDBSCAN, which indicates a great variability within the micro-habitats belonging to PL. That makes them "special" or at least "different" from common micro-habitats in SL and BL. Micro-habitats of PL were usually located around the 17°C isoline and at the occurrence of exceptionally strong chlorophyll-peaks.

In the North Sea, peaks of primary production following the spring bloom were observed in subsurface layers

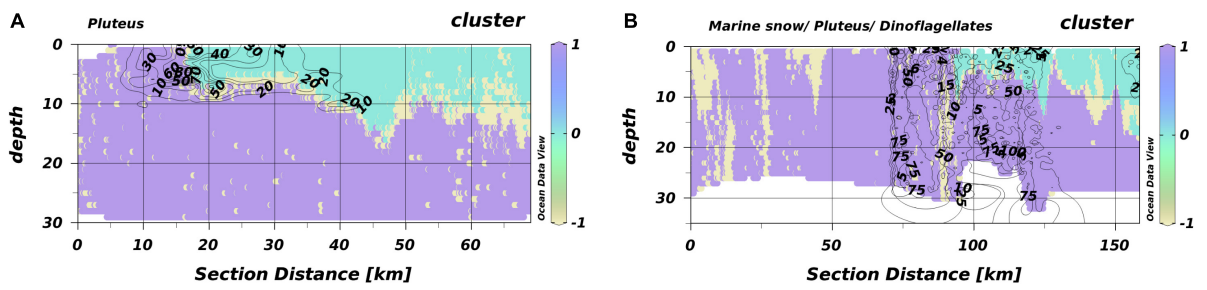


FIGURE 8 | Habitat maps for T1 and T2 from HE534. The isolines give the densities (N l^{-1}) of pluteus for T1. Isolines in the figure of T2 include pluteus, marine snow, and dinoflagellates (all in N l^{-1}). **(A)** T1 and **(B)** T2. Different colors represent different macro-habitats. Different numbers are the respective densities indicated by the isolines.

(Richardson et al., 1998, 2000). The PL most likely resembles such areas of subsurface productivity.

Furthermore, the model detected an upward doming of the thermocline within an OWF, probably caused by enhanced vertical mixing (Segtnan and Christakos, 2015; Floeter et al., 2017; Schultze et al., 2020). The upward doming and the resulting temperature differences are comparable to those observed within cyclonic eddies (Dong and McWilliams, 2007; Marmorino et al., 2018), indicating that OWF's can influence the pelagic realm in the same order of magnitude as natural (sub-) mesoscale processes like eddies. The doming of colder, nutrient-rich water can produce chlorophyll peaks (Munk et al., 1999), indicating the potential for an enhanced primary production in this area.

Cumulative effects of single foundations might lead to a blocking effect around OWF's, similar as observed for islands (Simpson et al., 1982), which has the potential to produce submesoscale eddies (Dong and McWilliams, 2007) in addition to local upwelling fronts (Floeter et al., 2017). Common properties that are used to describe hydrographic eddies and fronts include water velocity, vorticity and the Rossby number (e.g., Lévy et al., 2012; Marmorino et al., 2018), all of which were not available to us, which makes it less likely to detect such features.

The situation during HE534 was fundamentally different from HE429, most likely due to the weather conditions prior to the cruise that dispersed a thermal stratification. However, the projections indicated that similar SL and BL as in HE429 were detected. In case of HE534, segregations occurred along an oxygen isoline ($>235\ \mu mol\ l^{-1}$) instead of temperature ($>17°C$) as during HE429. This is in accordance with findings of Friedland et al. (2020) and references within that the predictive power of variables might change. This variable nature inherently present in pelagic data (Hinchey et al., 2008; Thompson et al., 2016) makes it so challenging to accurately predict pelagic habitats. The temperature isolines in T1 and T2 (**Supplementary Figure 3**) clearly indicate the presence of a tidal mixing front (see Hill et al., 1993). A convergence slick, which is typically associated with such tidal mixing fronts (Hill et al., 1993), would also explain the observed aggregation of plankton particles at the intersection of the two macro-habitats (**Figure 8**).

There exists plenty of evidence that physical properties also structure the marine plankton communities (e.g., Swalethorp et al., 2015; Van Leeuwen et al., 2015; Lindegren et al., 2020). However, the generated habitat maps have only limited explanatory power considering the observed plankton communities. This is not unexpected since physical properties are merely incomplete predictors for the community structure which is most likely further shaped by niche-based processes and interactions (Houliez et al., 2021).

Top predators aggregate in areas with the highest prey-patch densities (not to be confused with the area of highest prey densities!) (Benoit-Bird et al., 2013) and peak abundances of zooplankton and fish larvae are frequently observed in the direct vicinity of frontal convergence zones (Munk et al., 1995, 2002; Höffle et al., 2013; Munk, 2014; Swalethorp et al., 2015). In addition to the horizontal agglomerations, thermo-, and haloclines can produce further vertical structuring (Höffle et al., 2013; Lindegren et al., 2020). Thereby, more pronounced differences lead to a stronger niche separation and less overlap between different species (Lindegren et al., 2020). Changes in nitrate (Scharfe and Wiltshire, 2019) and silicate (Wiltshire et al., 2015) availability produce a temporal succession of different dominant taxa in the tidal advected phytoplankton community. Especially the plankton community is thus shaped by complex spatio-temporal dynamics and local prey patches have the potential to shape the distribution of higher trophic levels (Pope et al., 1994; Burkhard et al., 2011; Benoit-Bird et al., 2013; Defriez et al., 2016), even though this might be of less importance for ecosystem services in a highly diverse and partly functionally redundant plankton communities like that of the North Sea (Atkinson et al., 2015).

COMMENTS AND RECOMMENDATIONS

Future work should aim to include species densities and water current related measurements in order to accurately predict not only physical habitats but also realized ecological niches and hopefully improve our understanding of the complex dynamics shaping the pelagic realm.

Our approach offers beneficial properties to solve this challenge: the AE is a highly non-linear tool to reduce the dimensionality of a nearly unlimited amount of data that can be extended as needed. Additionally, HDBSCAN is a cluster algorithm that makes as few assumptions as possible, i.e., regarding number or shape of clusters. HDBSCAN can also handle outliers on it's own in opposite to, e.g., k-means, and even enables to treat them in our case as an own macro-habitat. While machine learning might not give insight into the underlying mechanistic, it can give a starting point from which to begin future investigations (Friedland et al., 2020).

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

SV and R-MP constructed the Autoencoder used to process the data. R-MP was responsible for the data analyses. JF and R-MP wrote the submitted manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The RV Heincke Research Cruises were supported by grant numbers AWI_HE429_00 and AWI_HE534_00.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2021.754375/full#supplementary-material

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). *TensorFlow: Large-scale Machine Learning on Heterogeneous Distributed Systems*. Available Online at: https://www.tensorflow.org/ (accessed March 30, 2020).

Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M., López-Jurado, J. L., et al. (2014). Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution. *PLoS One* 9:e109338. doi: 10.1371/journal.pone.0109338

Amorim, E., Ramos, S., Elliott, M., and Bordalo, A. A. (2018). 'Dynamic habitat use of an estuarine nursery seascape: ontogenetic shifts in habitat suitability of the european flounder (platichthys flesus)'. *J. Exp. Mar. Biol. Ecol.* 506, 49–60. doi: 10.1016/j.jembe.2018.05.011

Arbelaitz, O., Gurrutxaga, I., Muguerza, J., Pérez, J. M., and Perona, I. (2013). An extensive comparative study of cluster validity indices. *Patt. Recogn.* 46, 243–256. doi: 10.1016/J.PATCOG.2012.07.021

Atkinson, A., Harmer, R. A., Widdicombe, C. E., McEvoy, A. J., Smyth, T. J., Cummings, D. G., et al. (2015). 'Questioning the role of phenology shifts and trophic mismatching in a planktonic food web'. *Progr. Oceanogr.* 137, 498–512. doi: 10.1016/j.pocean.2015.04.023

Baschek, B., and Maarten Molemaker, J. (2010). *Aerial and In Situ Measurements of Submesoscale Eddies, Fronts, and Filaments*. Available Online at: https://ui.adsabs.harvard.edu/abs/2010EGUGA..12.1106B/abstract (accessed April 26, 2021).

Bellido, J. M., Brown, A. M., Valavanis, V. D., Giráldez, A., Pierce, G. J., Iglesias, M., et al. (2008). "Identifying essential fish habitat for small pelagic species in spanish mediterranean waters," in *Essential Fish Habitat Mapping in the Mediterranean*, ed. V. D. Valavanis (Dordrecht: Springer), 171–184.

Bengio, Y., Lamblin, P., Popovici, D., and Larochelle, H. (2006). "Greedy layer-wise training of deep networks," in *Proceedings of the 19th International Conference on Neural Information Processing Systems*, (Cambridge: MIT Press), 153–160.

Benoit-Bird, K. J., Battaile, B. C., Heppell, S. A., Hoover, B., Irons, D., Jones, N., et al. (2013). Prey patch patterns predict habitat use by top marine predators with diverse foraging strategies. *PLoS One* 8:e53348. doi: 10.1371/journal.pone.0053348

Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., et al. (2014). Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds'. *Nat. Commun.* 5:5239. doi: 10.1038/ncomms6239

Bez, N. (2000). On the use of lloyd's index of patchiness. *Fish. Oceanogr.* 9, 372–376. doi: 10.1046/j.1365-2419.2000.00148.x

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press.

Buckingham, C. E., Garabato, A. C. N., Thompson, A. F., Brannigan, L., Lazar, A., Marshall, D. P., et al. (2016). Seasonality of submesoscale flows in the ocean surface boundary layer. *Geophys. Res. Lett.* 43, 2118–2126. doi: 10.1002/2016GL068009

Burkhard, B., Opitz, S., Lenhart, H., Ahrendt, K., Garthe, S., Mendel, B., et al. (2011). Ecosystem based modeling and indication of ecological integrity in the german north sea—case study offshore wind parks. *Ecol. Indic.* 11, 168–174. doi: 10.1016/j.ecolind.2009.07.004

Chen, Z., Yeo, C. K., Lee, B. S., and Lau, C. T. (2018). Autoencoder-based network anomaly detection. *Wireless Telecommun. Symp.* 2018, 1–5. doi: 10.1109/WTS.2018.8363930

Chollet, F. (2015). *Keras. GitHub*. Available Online at: https://github.com/fchollet/keras (accessed March 30, 2020).

Defriez, E. J., Sheppard, L. W., Reid, C., and Reuman, D. C. (2016). Climate change-related regime shifts have altered spatial synchrony of plankton dynamics in the north sea. *Global Change Biol.* 22, 2069–2080. doi: 10.1111/gcb.13229

Dong, C., and McWilliams, J. C. (2007). A numerical study of island wakes in the southern california bight. *Cont. Shelf Res.* 27, 1233–1248. doi: 10.1016/j.csr.2007.01.016

Erhan, D., Manzagol, P.-A., Bengio, Y., Bengio, S., and Vincent. (2009). The difficulty of training deep architectures and the effect of unsupervised pre-training. *Artif. Intell. Stat.* 5, 153–160.

Espinasse, B., Carlotti, F., Zhou, M., and Devenon, J. L. (2014). Defining zooplankton habitats in the gulf of lion (NW mediterranean sea) using size structure and environmental conditions. *Mar. Ecol. Progr. Ser.* 506, 31–46. doi: 10.3354/meps10803

Floeter, J., Beusekom, J. E. E., van, Auch, D., Callies, U., Carpenter, J., et al. (2017). Pelagic effects of offshore wind farm foundations in the stratified north sea. *Progr. Oceanogr.* 156, 154–173. doi: 10.1016/j.pocean.2017.07.003

Friedland, K. D., Bachman, M., Davies, A., Frelat, R., McManus, M. C., Morse, R., et al. (2020). Machine learning highlights the importance of primary and secondary production in determining habitat for marine fish and macroinvertebrates. *Aquat. Conserv.* 13, 1482–1498. doi: 10.1002/aqc.3527

Funk, S., Krumme, U., Temming, A., and Möllmann, C. (2020). 'Gillnet fishers' knowledge reveals seasonality in depth and habitat use of cod (gadus morhua) in the western baltic sea'. *ICES J. Mar. Sci.* 77, 1816–1829. doi: 10.1093/icesjms/fsaa071

Giannoulaki, M., Pyrounaki, M. M., Liorzou, B., Leonori, I., Valavanis, V. D., Tsagarakis, K., et al. (2011). Habitat suitability modelling for sardine juveniles (sardina pilchardus) in the mediterranean sea. *Fish. Oceanogr.* 20, 367–382. doi: 10.1111/j.1365-2419.2011.00590.x

Gigot, C. (2018). *Epiphy: Analysis of Plant Disease Epidemics*. Available Online at: https://CRAN.R-project.org/package=epiphy (accessed May 14, 2020).

Harris, T., and Baker, E. K. (2012). "1 - Why map benthic habitats?" in *Seafloor Geomorphology as Benthic Habitat*, eds T. Harris and E. K. Baker (London: Elsevier), 3–22.

Hill, A. E., James, I., Linden, P., Matthews, J., Prandle, D., Simpson, J., et al. (1993). Dynamics of tidal mixing fronts in the north sea. *Philos. Trans. R. Soc. B Biol. Sci.* 343, 431–446. doi: 10.1098/rsta.1993.0057

Hinchey, E. K., Nicholson, M. C., Zajac, R. N., and Irlandi, E. A. (2008). Preface: marine and coastal applications in landscape ecology. *Landscape Ecol.* 23, 1–5. doi: 10.1007/s10980-007-9141-3

Hinton, G. E. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, 504–507. doi: 10.1126/science.1127647

Höffle, H., Nash, R. D. M., Falkenhaug, T., and Munk, P. (2013). Differences in vertical and horizontal distribution of fish larvae and zooplankton, related to hydrography. *Mar. Biol. Research* 9, 629–644. doi: 10.1080/17451000.2013.765576

Houliez, E., Lefebvre, S., Dessier, A., Huret, M., Marquis, E., Bréret, M., et al. (2021). Spatio-temporal drivers of microphytoplankton community in the bay of biscay: do species ecological niches matter? *Progr. Oceanogr.* 194:102558. doi: 10.1016/j.pocean.2021.102558

Labat, J.-P., Gasparini, S., Mousseau, L., Prieur, L., Boutoute, M., and Mayzaud, P. (2009). Mesoscale distribution of zooplankton biomass in the northeast atlantic ocean determined with an optical plankton counter: relationships with environmental structures. *Deep Sea Res. I Oceanogr. Res. Pap.* 56, 1742–1756. doi: 10.1016/j.dsr.2009.05.013

Laman, E. A., Rooper, C. N., Turner, K., Rooney, S., Cooper, D. W., and Zimmermann, M. (2017). Using species distribution models to describe essential fish habitat in alaska. *Can. J. Fish. Aquat. Sci.* 75, 1230–1255. doi: 10.1139/cjfas-2017-0181

Lévy, M., Ferrari, R., Franks, J. S., Martin, A. P., and Rivière, P. (2012). Bringing physics to life at the submesoscale. *Geophys. Res. Lett.* 39:L14602. doi: 10.1029/2012GL052756

Lévy, M., Franks, J. S., and Smith, K. S. (2018). The role of submesoscale currents in structuring marine ecosystems. *Nat. Commun.* 9:4758. doi: 10.1038/s41467-018-07059-3

Levy, M., and Martin, A. P. (2013). The influence of mesoscale and submesoscale heterogeneity on ocean biogeochemical reactions. *Global Biogeochem. Cyc.* 27, 1139–1150. doi: 10.1002/2012GB004518

Lindegren, M., Thomas, M. K., Jónasdóttir, S. H., Nielsen, T. G., and Munk, P. (2020). Environmental niche separation promotes coexistence among ecologically similar zooplankton species—north sea copepods as a case study. *Limnol. Oceanogr.* 65, 545–556. doi: 10.1002/lno.11322

Manakov, I., Rohm, M., and Tresp, V. (2019). *Walking the Tightrope: An Investigation of the Convolutional Autoencoder Bottleneck*. Available Online at: https://openreview.net/forum?id=ryguP1BFwr [accessed Sep 23, 2020].

Marmorino, G. O., Smith, G. B., North, R. P., and Baschek, B. (2018). Application of airborne infrared remote sensing to the study of ocean submesoscale eddies. *Front. Mech. Eng.* 4:10. doi: 10.3389/fmech.2018.00010

McInnes, L., Healy, J., and Astels, S. (2017). Hdbscan: hierarchical density based clustering. *J. Open Source Softw.* 2:205. doi: 10.21105/joss.00205

Munk, P. (2014). Fish larvae at fronts: horizontal and vertical distributions of gadoid fish larvae across a frontal zone at the norwegian trench. *Deep Sea Res. II Top. Stud. Oceanogr.* 107, 3–14. doi: 10.1016/j.dsr2.2014.01.016

Munk, P., Larsson, P., Danielsen, D., and Moksness, E. (1995). Larval and small juvenile cod gadus morhua concentrated in the highly productive areas of a shelf break front. *Mar. Ecol. Progr. Ser.* 125, 21–30. doi: 10.3354/meps125021

Munk, P., Larsson, P., Danielssen, D., and Moksness, E. (1999). Variability in frontal zone formation and distribution of gadoid fish larvae at the shelf break in the northeastern north sea. *Mar. Ecol. Progr. Ser.* 177, 221–233. doi: 10.3354/meps177221

Munk, P., Wright, J., and Pihl, N. J. (2002). Distribution of the early larval stages of cod, plaice and lesser sandeel across haline fronts in the north sea. *Estuar. Coast. Shelf Sci.* 55, 139–149. doi: 10.1006/ecss.2001.0892

North, R. P., Riethmüller, R., and Baschek, B. (2016). Detecting small–scale horizontal gradients in the upper ocean using wavelet analysis. *Estuar. Coast. Shelf Sci.* 180, 221–229. doi: 10.1016/j.ecss.2016.06.031

Omand, M. M., D'Asaro, E. A., Lee, C. M., Perry, M. J., Briggs, N., Cetinić, I., et al. (2015). Eddy-driven subduction exports particulate organic carbon from the spring bloom. *Science* 348, 222–225. doi: 10.1126/science.1260062

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12, 2825–2830.

Pittman, S. J., Kneib, R. T., and Simenstad, C. A. (2011). Practicing coastal seascape ecology. *Mar. Ecol. Progr. Ser.* 427, 187–190. doi: 10.3354/meps09139

Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S., and Floeter, J. (2021). Automatic plankton image classification—can capsules and filters help cope with data set shift? *Limnol. Oceanogr.* 19, 176–195. doi: 10.1002/lom3.10413

Pope, J. G., Shepherd, J. G., Webb, J., Stebbing, A. R. D., Mangel, M., Beverton, R. J. H., et al. (1994). Successful surf-riding on size spectra: the secret of survival in the sea. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 343, 41–49. doi: 10.1098/rstb.1994.0006

QGIS (2021). *QGIS Geographic Information System*. Switzerland: QGIS Association.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Richardson, K., Nielsen, T., Pedersen, F., Heilmann, J., Løkkegaard, B., and Kaas, H. (1998). Spatial heterogeneity in the structure of the planktonic food web in the north sea. *Mar. Ecol. Progr.Ser.* 168, 197–211. doi: 10.3354/meps168197

Richardson, K., Visser, A. W., and Pedersen, F. B. (2000). Subsurface phytoplankton blooms fuel pelagic production in the north sea. *J. Plankton Res.* 22, 1663–1671. doi: 10.1093/plankt/22.9.1663

Rousseeuw, J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* 20, 53–65. doi: 10.1016/0377-0427(87)90125-7

Sano, M., Maki, K., Nishibe, Y., Nagata, T., and Nishida, S. (2013). Feeding habits of mesopelagic copepods in sagami bay: insights from integrative analysis. *Progr. Oceanogr.* 110, 11–26. doi: 10.1016/j.pocean.2013.01.004

Scharfe, M., and Wiltshire, K. H. (2019). Modeling of intra-annual abundance distributions: constancy and variation in the phenology of marine phytoplankton species over five decades at helgoland roads (north sea). *Ecol. Model.* 404, 46–60. doi: 10.1016/j.ecolmodel.2019.01.001

Schlitzer, R. (2020). *Ocean Data View*. Available Online at: https://odv.awi.de/ [accessed March 30, 2021].

Schultze, L. K. P., Merckelbach, L. M., Horstmann, J., Raasch, S., and Carpenter, J. R. (2020). Increased mixing and turbulence in the wake of offshore wind farm foundations. *J. Geophys. Res. Oceans* 125:e2019JC015858. doi: 10.1029/2019JC015858

Schulz, J., Peck, M. A., Barz, K., Schmidt, J. O., Hansen, F. C., Peters, J., et al. (2012). Spatial and temporal habitat partitioning by zooplankton in the bornholm basin (central baltic sea). *Progr. Oceanogr.* 107, 3–30. doi: 10.1016/j.pocean.2012.07.002

Segtnan, O. H., and Christakos, K. (2015). 'Effect of offshore wind farm design on the vertical motion of the ocean. *Energy Procedia* 80, 213–222. doi: 10.1016/j.egypro.2015.11.424

Shulman, I., Penta, B., Richman, J., Jacobs, G., Anderson, S., and Sakalaukus, P. (2015). Impact of submesoscale processes on dynamics of phytoplankton filaments. *J. Geophys. Res. Oceans* 120, 2050–2062. doi: 10.1002/2014JC010326

Simpson, J. H., Tett, B., Argote-Espinoza, M. L., Edwards, A., Jones, K. J., and Savidge, G. (1982). Mixing and phytoplankton growth around an island in a stratified sea. *Cont. Shelf Res.* 1, 15–31. doi: 10.1016/0278-4343(82)90030-9

Subbey, S. (2018). Parameter estimation in stock assessment modelling: caveats with gradient-based algorithms. *ICES J. Mar. Sci.* 75, 1511–1511. doi: 10.1093/icesjms/fsy060

Swalethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T., and Munk, P. (2015). Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced greenlandic fjord: influence from hydrography and prey availability. *J. Plankton Res.* 37, 102–119. doi: 10.1093/plankt/fbu099

Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, G. M., and Heywood, K. J. (2016). Open-ocean submesoscale motions: a full seasonal cycle of mixed layer instabilities from gliders. *J. Phys. Oceanogr.* 46, 1285–1307. doi: 10.1175/JPO-D-15-0170.1

Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J. M., Brasseur, P., et al. (2012). Generation of analysis and consistent error fields using the data interpolating variational analysis (DIVA). *Ocean Model.* 52-53, 90–101. doi: 10.1016/j.ocemod.2012.05.002

Tugores, M. P., Giannoulaki, M., Iglesias, M., Bonanno, A., Tičina, V., Leonori, I., et al. (2011). Habitat suitability modelling for sardine sardina pilchardus in a highly diverse ecosystem: the mediterranean sea. *Mar. Ecol. Progr. Ser.* 443, 181–205. doi: 10.3354/meps09366

Van Leeuwen, S., Tett, P., Mills, D., and van der Molen, J. (2015). Stratified and nonstratified areas in the north sea: long-term variability and biological and policy implications. *J. Geophys. Res. Oceans* 120, 4670–4686. doi: 10.1002/2014JC010485

Van Rossum, G., and Drake, F. L. (2009). *Python 3 Reference Manual*. Scotts Valley: CreateSpace.

Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y., and Manzagol, P.-A. (2010). Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion. *J. Mach. Learn. Res.* 11, 3371–3408.

Wedding, L. M., Lepczyk, C. A., Pittman, S. J., Friedlander, A. M., and Jorgensen, S. (2011). Quantifying seascape structure: extending terrestrial spatial pattern metrics to the marine realm. *Mar. Ecol. Progr. Ser.* 427, 219–232. doi: 10.3354/meps09119

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., et al. (2019). Welcome to the tidyverse. *J. Open Source Softw.* 4:1686. doi: 10.21105/joss.01686

Wiltshire, K. H., Boersma, M., Carstens, K., Kraberg, A. C., Peters, S., and Scharfe, M. (2015). Control of phytoplankton in a shelf sea: determination of the main drivers based on the helgoland roads time series. *J. Sea Res.* 105, 42–52. doi: 10.1016/j.seares.2015.06.022

Zhao, Y., Deng, B., Shen, C., Liu, Y., Lu, H., and Hua, X.-S. (2017). "Spatio-temporal AutoEncoder for video anomaly detection," in *Proceedings of the 25th ACM International Conference on Multimedia*, (New York: Association for Computing Machinery (MM '17)), 1933–1941.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., and Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer-Verlag.

Zuur, A. F., Ieno, E. N., and Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* 1, 3–14. doi: 10.1111/j.2041-210X.2009.00001.x

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Marine Science

METHODS
published: 12 April 2022
doi: 10.3389/fmars.2022.868420

Automated Plankton Classification With a Dynamic Optimization and Adaptation Cycle

Jan Conradt[1]*, *Gregor Börner*[1], *Ángel López-Urrutia*[2], *Christian Möllmann*[1] *and Marta Moyano*[3]

[1] Institute of Marine Ecosystem and Fisheries Science (IMF), Center for Earth System Research and Sustainability (CEN), Hamburg University, Hamburg, Germany, [2] Centro Oceanográfico de Gijón/Xixón, Instituto Español de Oceanografía, Gijón, Spain, [3] Department of Natural Sciences, Center for Coastal Research, Universitet i Agder, Kristiansand, Norway

With recent advances in Machine Learning techniques based on Deep Neural Networks (DNNs), automated plankton image classification is becoming increasingly popular within the marine ecological sciences. Yet, while the most advanced methods can achieve human-level performance on the classification of everyday images, plankton image data possess properties that frequently require a final manual validation step. On the one hand, this is due to morphological properties manifesting in high intra-class and low inter-class variability, and, on the other hand is due to spatial-temporal changes in the composition and structure of the plankton community. Composition changes enforce a frequent updating of the classifier model *via* training with new user-generated training datasets. Here, we present a Dynamic Optimization Cycle (DOC), a processing pipeline that systematizes and streamlines the model adaptation process *via* an automatic updating of the training dataset based on manual-validation results. We find that frequent adaptation using the DOC pipeline yields strong maintenance of performance with respect to precision, recall and prediction of community composition, compared to more limited adaptation schemes. The DOC is therefore particularly useful when analyzing plankton at novel locations or time periods, where community differences are likely to occur. In order to enable an easy implementation of the DOC pipeline, we provide an end-to-end application with graphical user interface, as well as an initial dataset of training images. The DOC pipeline thus allows for high-throughput plankton classification and quick and systematized model adaptation, thus providing the means for highly-accelerated plankton analysis.

Keywords: machine learning, deep neural networks, plankton community, classification, model adaptation

OPEN ACCESS

Edited by:
Fabien Lombard,
Sorbonne Universités, France

Reviewed by:
Enrique Nogueira,
Spanish Institute of Oceanography (IEO), Spain
Duane Edgington,
Monterey Bay Aquarium Research Institute (MBARI), United States

***Correspondence:**
Jan Conradt
jan.conradt@uni-hamburg.de

Specialty section:
This article was submitted to Marine Ecosystem Ecology, a section of the journal Frontiers in Marine Science

Received: *02 February 2022*
Accepted: *16 March 2022*
Published: *12 April 2022*

Citation:
Conradt J, Börner G, López-Urrutia Á, Möllmann C and Moyano M (2022) Automated Plankton Classification With a Dynamic Optimization and Adaptation Cycle. Front. Mar. Sci. 9:868420. doi: 10.3389/fmars.2022.868420

INTRODUCTION

Plankton is a diverse group of organisms with a key role in marine food-webs and biogeochemical cycles (e.g. Castellani and Edwards, 2017). It is furthermore responsible for about 50% of the global primary production, and they serve as prey for upper trophic levels and as recyclers of organic matter. Changes in their abundance, biogeography or size structure can thus lead to large changes at

the ecosystem level (e.g. Frederiksen et al., 2006; Capuzzo et al., 2017). Climate change in particular can cause major changes in plankton community characteristics. The range of specific research on plankton in the ecological context is wide, covering issues such as the effect of ocean acidification on calcifying organisms (e.g. Stern et al., 2017), migrations of plankton taxa in response to ocean warming (Beaugrand, 2012), or the determination of available food biomass to larval fish at changing hatching times (Asch et al., 2019; Durant et al., 2019). Ultimately, however, many of these address – directly or indirectly – the effects of environmental change on the abundance of commercially exploited marine fish species, which are dependent on plankton either as food for their early life-stages, or as food of their prey. As plankton forms the base of any marine food web, climate effects are propagated to higher trophic levels *via* the response of the plankton community to climate change (Winder and Sommer, 2012; Nagelkerken et al., 2017). Monitoring its composition and abundance is hence of great importance to understanding the effects of climate change on the entire marine ecosystem and services it provides to humanity.

The study of plankton in an environmental context is both quantitative and qualitative in nature. While certain plankton estimates (e.g. phytoplankton biomass) can be inferred from analysis of satellite imagery, most studies require abundance indices of specific taxa that can only be derived from sampling plankton *in situ* and determining its composition. Depending on the research subject, the taxonomic, life-stage and size composition of plankton can e.g. indicate the presence of a community specific to a certain water mass/current (Russell, 1939; Beaugrand et al., 2002), an abundance shift of potentially climate-sensitive species, or the abundance of planktonic food suitable to a particular predator of interest (Dam and Baumann, 2017).

Traditionally, plankton samples have been analyzed by humans with optical devices like microscopes (Wiebe et al., 2017). The accuracy of taxonomic classification was usually high when done by experienced personnel, but it could decrease significantly in complex tasks, such as the differentiation between morphologically similar taxa (Culverhouse et al., 2003). Additionally, sample processing rate is limiting the total number of samples that could be processed using traditional microscopy. The introduction of plankton-image recorders for both *in situ* (e.g. Video Plankton Recorder, VPR, (Davis et al., 1992)) and/or fixed samples (e.g. Flow Cytometer and Microscope [FlowCAM®; Sieracki et al., 1998)], together with the development of image-classification algorithms, has led to great advances in the processing of plankton samples over the last two-to-three decades (e.g. Kraberg et al., 2017; Lombard et al., 2019; Goodwin et al., 2022). Image recording enables the temporally unlimited storage of visual information even for samples that cannot withstand fixing agents for a long time. Furthermore, given that the photographs are stored on disk, all visual information is kept permanently, and is available for discussion, unlike the memories of an expert. However, one of the challenges of these plankton image-recording devices (like VPR or FlowCam) is the large number of images that need to be classified (e.g. > 52 million in Briseño-Avena et al., 2020). So far, classification models are intended to greatly increase classification speed, be it *via* an entire replacement of expert classification with model predictions (Briseño-Avena et al., 2020), or by yielding a rough pre-sorting that alleviates expert validation (Álvarez et al., 2014).

Image classification models were introduced in the late 1980s, first in the form of Neural Networks (NN), which were famously employed for the classification of handwritten digits by the US postal service (LeCun et al., 1989). In the mid-1990s, these were temporally superseded by Support-Vector Machines (SVMs), and for the first time applied for plankton classification in 1998 by Tang et al. (1998). Neural Networks were, at that time, relatively simple in design and could only be applied for simple classification tasks, e.g. discriminating between the clearly-shaped digits. While theory allowed the design of larger NNs for more complex targets like plankton images, constraints in computational power put a temporary constraint on this (e.g. Gu et al., 2018).

SVMs became the tool of choice for plankton classification in the 2000s and early 2010s due to relatively strong performance (e.g. Álvarez et al., 2012). However, they were limited in capability and convenience-of-use by the need for human-defined features for class-discrimination (a limitation not present in NNs). Such "feature-engineering" was required to reduce the enormous amount of information contained in an image (a data point in R^n-dimensional space, n being the number of pixels) to details required to automatically tell classes apart (Scholkopf and Smola, 2002). Many publications of that time concerned the engineering of new features for better class separation, and the problem of the redundancy of devised features (e.g. Tang et al., 1998; Tang et al., 2006; Li et al., 2014). Even then, unique difficulties posed by plankton images became apparent, including the transparent nature of many plankton taxa and morphological similarities between classes.

Computational power increased strongly in parallel to SVMs reaching their peak of popularity, and NNs eventually regained strong popularity (e.g. Chollet, 2017). In 2012, Krizhevsky et al. won the ImageNet contest with a so –called Deep Convolutional Neural Net (CNN), beating the peak performance achieved in the years prior by a before-unachieved margin. The advances in classification accuracy led to massive investments into the design and application of Deep Neural Nets (the "parent class" of CNNs) in research and economy (Chollet, 2017).

Plankton classification eventually followed suite in this general trend (e.g. Orenstein et al., 2015; Al-Barazanchi et al., 2018), due to the capability of "deep" CNNs to devise and select features themselves; a process colloquially termed "Artificial Intelligence" (AI). CNNs are essentially a complex extension of multinomial regression, whereby the model input, the image, is an array of pixel values, and the output a quasi-"one-hot"-encoded class vector. The vector dimension with maximum value is taken as the predicted class index. Different from simple regression, several "layers" of neurons" – essentially arrays or vectors, lie in-between the model input and output.

These contain abstracted information from the image, with parameters between any element of two adjacent arrays or vectors determining the flow of information (i.e., the filtering-out of information) from lower- to higher-order input representation (LeCun et al., 2010). During model fitting, the backpropagation algorithm transmits classification loss to each parameter using differential calculus, allowing for gradient-based optimization of the complex NN (Rumelhart et al., 1986). Backpropagation essentially allows the model to "learn" to filter information "wisely" by optimizing its parameter values over multiple iterations of fitting (e.g. Goodfellow et al., 2016).

Today, CNN classification models can reach accuracies of well over 95% (e.g. Al-Barazanchi et al., 2018), making automatic plankton classification appearing like a "solved task" at first sight. However, these accuracy values are usually derived from performance on test data originating from the same statistical population as the training data. Thus, these outcomes are only "snapshots" of the range of performances that will occur when a static model is applied to plankton samples that lie outside the "population", where the training data originate from. More precisely, the plankton community tends to vary strongly in time and space, and this variability is precisely what most plankton researchers are interested in. As new taxa appear in a specific location or as formerly less-frequently encountered taxa increase in abundance, a classification model trained on a plankton community, or a pool of communities, from different geographic or temporal origin will likely perform poorly on the respective new samples (dataset shift; Moreno-Torres et al., 2012). González et al. (2016) noted the variability in model performance on samples of different origins and recommended to focus the development of applications robust to various distances between training set and field samples. Also, the non-homogeneous distribution of plankton taxa in the field means that training datasets are often strongly non-homogeneous in distribution of images over classes, as well. This poses a constraint to the successful training of a CNN, since the resulting model will perform well on the dominating classes, but poorly on lower-abundant ones. Note that this is not necessarily reflected in the general accuracy metric, which only accounts for the total number of correctly classified images pooled over all classes.

One further difficulty in automated plankton classification lies in the sometimes high inter-class similarity (e.g. bivalves and some dinoflagellate taxa) (**Figure 1A**) and high intra-class variability in appearance (which is founded in the existence of sub-taxa, different life-stages or different appearances resulting from different imaging angles) (**Figure 1B**) of plankton organisms. Thus, if the intra-class variability is not homogeneously reflected in the training set, the ability of the CNN to discriminate between classes may be limited to only a fraction of the existing sub-classes.

In summary, the current constraints on successful training and application of models for automatic plankton classification are the often limited quality of training sets, and the high spatio-temporal dynamics of the plankton community. Under these

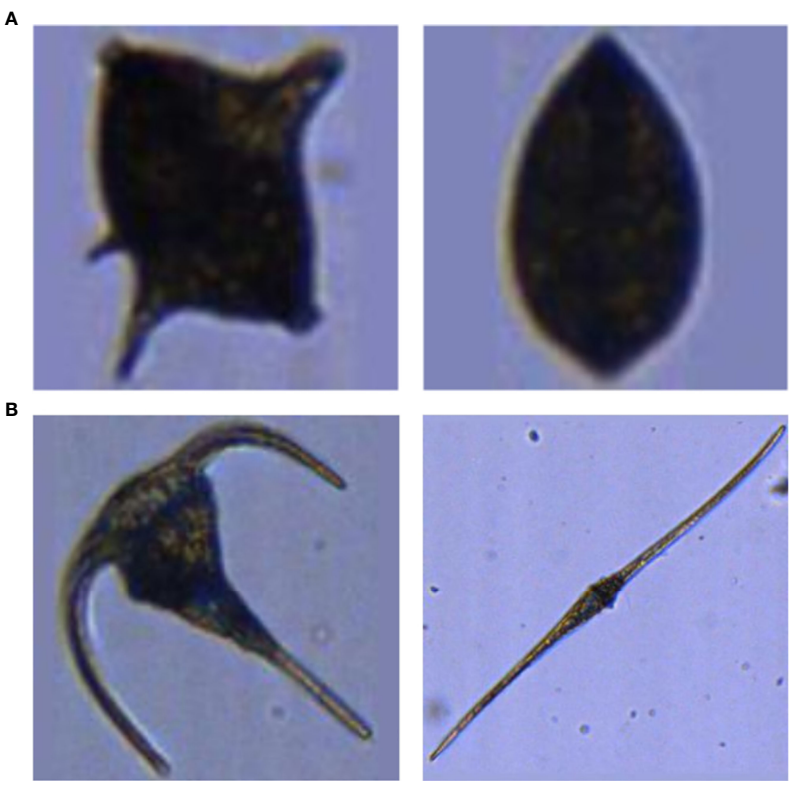


FIGURE 1 | Examples of strong inter-class similarity **(A)** and high intra-class dissimilarity **(B)**. **(A)** A dinoflagellate of the genus *Protoperidinium* spp. (left), and a juvenile bivalve (right). **(B)** Two dinoflagellates: *Ceratium fusus* (left) and *Ceratium tripos* (right).

circumstances, manual validation and correction of the model results is recommended (Gorsky et al., 2010), as is the adaptation of the model to avoid a decrease in classification performance. The latter usually requires the availability of machine-learning expertise, a commodity often lacking in the marine sciences (Malde et al., 2020). Research and development should thus be focused on reducing the time required for the validation task and on improving operability of classifier models by non-AI-experts.

Here, we follow González et al.'s (2016) suggestion and propose a pipeline for alleviating the task of model adaptation to a changing plankton community, and thus for reducing the time for manual validation: A "dynamic optimization cycle" (DOC) for iterative use accessible by non-AI-experts. By making applied use of a trained model on field samples, correcting the classification and evaluating model performance through expert knowledge, and updating the model training set and the model itself (through training on the updated image set), the classifier model adapts to spatial and/or temporal changes in the plankton community. It thus maintains high classification performance, ensuring that validation workload remains relatively constant. The systematization of this procedure, and the implementation of the DOC as an end-to-end application with graphical user interface, removes the requirement for expertise in designing and coding CNNs. The DOC was designed for the classification of FlowCam images and the workflow related to studies using the FlowCam, but is likely applicable for other types of plankton images and different types of workflow, as well.

MATERIALS & EQUIPMENT

Hardware and Software Requirements

Training of NNs was performed with a Nvidia® (Santa Clara/California/US) Quadro P2000 GPU with 4 GiB RAM (driver version 410.79) on a Dell® (Round Rock/Texas/US) Precision 5530 notebook with 32 GiB RAM. CUDA® (Nvidia, Santa Clara/California/US) version 10.0.130 was used for enabling the GPU to be used for general purpose processing. Programming was performed in Python 3.6.8 (van Rossum, 1995) using the Spyder Integrated Developer Environment (Raybaut, 2017) with Ipython version 7.2.0 (Perez and Granger, 2007). Packages used for analyzing classification outputs included NumPy (Oliphant, 2006), pandas (McKinney, 2010) and Dplython (Riederer, 2016). Packages used for image pre-processing included Matplotlib (Hunter, 2007), PIL (Lundh and Ellis, 2019) and Scipy (Oliphant, 2007). Tensorflow 1.12.0 (Abadi et al., 2015) and Keras 2.2.4 (Chollet, 2015) (with Tensorflow backend) Advanced Programming Interfaces were used for building, training and application of the classifier models.

METHODS

Model Design and Training

A convolutional neural net (CNN) was built based on the publicly available "VGG16" network architecture (Simonyan and Zisserman, 2015). This architecture consists of 13 convolutional layers, i.e. 13 intermediate data representations in the form of a stack of matrices that account for positional relationships between pixels of the input image. These layers are arranged in five "blocks" of two-to-three layers each, which are connected *via* non-parameterized information-pooling layers. The sixth block consisting of so-called "dense" layers was removed – as is usually done when applying a pre-defined architecture – and replaced with custom layers: one convolutional layer and two dense layers. The design of this custom "block" of layers - i.e. the number and type of layers, and the number of neurons (i.e. representation dimensions) of each – was the result of a try-and-error approach for achieving satisfying classification performance on training and validation images (Conradt, 2020). Details on the custom layers can be obtained from tab. SI V/2.

Model parameters were initialized with the values provided together with the VGG16 architecture trained on ImageNet data (Deng et al., 2009) for the respective part of the model, and with values drawn randomly from a Glorot uniform distribution (Glorot and Bengio, 2010) for the custom layers, as per default in the Keras software. Model training (i.e. fitting) was started with the custom layers and the final block of convolutional layers of the VGG16 "base" set to trainable. Training was performed by feeding all training images in a sequence of batches of 20 randomly chosen images to the model. All other hyper-parameter settings (e.g. optimizer and learning rate for gradient-based fitting) can be obtained from Tab. SI IV/1. The choice of hyper-parameter settings was based on a series of trial runs for different hyper-parameter set-ups (Conradt, 2020).

The entire set of training images was fed eight times (so-called "epochs") to the model, with an increasing number of the layers of the VGG16 base being set to trainable ("unfrozen") each epoch (Tab. SI V/1). "Unfreezing" is a common procedure applied to ensure that learned features are gradually adapted towards our plankton dataset (VGG16 was originally trained on the ImageNet set of everyday-object images). The chosen number of epochs and the "unfreezing" schedule resulted from optimization through trial-and-error experimentation, as well (Conradt, 2020).They resulted in a steady increase of validation accuracy from approx. 88% to approx. 94% (**Figure SI VIII/1 B**) and a decrease of validation loss from approx. 0.34 to approx. 0.29 when trained on the baseline training set, though validation loss did increase slightly from a minimum value of approx. 0.26 at the third epoch (**Figure SI VIII/1 A**). Validation accuracy was surpassed by training accuracy by the second epoch, which is usually a sign of an onset of over-fitting (e.g. Chollet, 2017); however, the fact that validation accuracy also still increased over the eight epochs was taken as a sign of a robust training schedule.

We did not utilize data augmentation, a technique in which artificial transformations are randomly applied to the training data to reduce model over-fitting and thus improve its generalizability (e.g. Chollet, 2017). While the approach is frequently applied in various image-classification tasks (e.g. Luo et al., 2018; Plonus et al., 2021), previous work had shown that data augmentation did not markedly improve the classification when applied to a partly identical data set of

FlowCam images (Conradt, 2020). This observation has also been made in another instance on an independent plankton data set (Lumini and Nanni, 2019).

While both the set-up of the CNN and the training scheme may not represent an optimal configuration (for example, over-fitting occurred in our experiments), we found the configurations to yield consistently robust results that were sufficient to support routine plankton analysis work. Given the relatively high validation accuracy, our goal was not to further optimize model design or –training, but instead to maintain this satisfactory performance level over changes in the composition of plankton samples.

Image Characteristics

Input image size was set to 120 x 120 x 3 pixels. A size of 256 x 256 x 3 pixels is more commonly used for plankton images (e.g. Orenstein and Beijbom, 2017; Al-Barazanchi et al., 2018; Cui et al., 2018), however preparatory work for the present study had shown that the chosen image size yielded better performance than a larger size, and leads to a faster processing due to the lower data dimensionality (Conradt, 2020). The use of a common square image shape leads to an altered visual appearance of plankton organisms if the original image had a height-length ratio very different from 1. This would increase intra-class variability, an undesirable trait as described above. Therefore, within the DOC pipeline, images are pre-processed *via* padding, i.e. by adding pixels in background color (the mode pixel value of the outermost pixel row for each color layer) to the sides or top and bottom to achieve square format, a common procedure in plankton-image classification (see e.g. Plonus et al., 2021).

Characteristics of the Baseline Training Set

The baseline image dataset, which is updated as part of the adaptive procedures of the DOC pipeline, consists of 27900 RGB FlowCam images of plankton samples gathered from various North Sea surveys over several years. Images in the dataset were sorted into 15 classes, including 13 taxonomic groups as well as a detritus class and a "clumps" class that contains aggregates of plankton organisms and/or detritus. The distribution of images over classes was designed to reflect general, though not empirically determined, patterns of natural relative abundance. However, the very abundant detritus class was reduced in relative proportion in order to avoid the learning of a quasi-binary classification scheme (detritus/non-detritus) by the classifier model. A random 80% of images of each class were used as training images for the baseline model, while 10% each were reserved for validation and testing purposes (see above). The characteristics of the baseline data set are given in Tab. SI VI/1.

Classification Thresholds

Within the DOC pipeline, the model classification is compared with expert validation. For each class, the relative amount of correct predictions is calculated and used as a threshold value against which the maximum probability value of the CNN output vector (the index of which is the class prediction) is compared. Probability values above the threshold lead to acceptance of the classification, as the model classification is deemed "certain". Probability values below the threshold lead to rejection of the model classification, the image is then assigned to an "uncertain-classifications" category. Initially, thresholds were set to 60% for all classes, as the difference between the properties of the baseline training set (on which the baseline model was trained) and the properties of the first station to be classified was deemed to be larger than that between subsequent modified training sets and stations.

This procedure was intended to speed up manual validation by implementing a sort out of images based on probability of miss-classification, which can then be checked more easily than if they were not separated from images with high probability of correct classification.

DOC Pipeline Procedures

The following describes the working steps for applying the DOC onto any given set of plankton samples (see also **Figure 2**). A more thorough user instruction with technical notes of importance is provided in the appendix (SI 1).

1. Classification (**Figure 2A**): The DOC pipeline is typically started by applying the provided classifier model directly on the classification of plankton samples, thus allowing for potentially large initial classification error. However, it is also possible to directly train a custom classifier model if the user has already generated a training set from manually labeled images, and perform the classification with this custom model (for details see SI 1).
2. Validation (**Figure 2B**): Following the classification of two to three plankton samples, the model classification is validated by a plankton expert (by moving images between class folders into which the images were copied by the DOC application). The number of samples required before continuing with the adaptation steps is likely case-specific and might require some initial trial-and-error experimentation. In our case studies, we classified two samples at a time. The validated classification is used as the final classification for further ecological studies. Model classification and expert validation are automatically compared and the correct-classification rate determined for each class.
3. Training-set update and threshold reduction (**Figure 2C**): After expert validation, the original model training set is stocked up with images that were miss-classified by the model. To this end, first the complement of the correct-classification rates is normalized *via* division by the maximum miss-classification rate over all classes (eq. 1, top). These values are then multiplied by the number of miss-classified images of each class to determine the number of images to be added to the training set (eq. 1, bottom). Not selecting all miss-classified images reduces the over-proportionality of naturally-abundant, but well-classified classes, e.g. detritus, in the image supplement, putting more emphasis on poorly-classified classes. The images added are selected randomly.

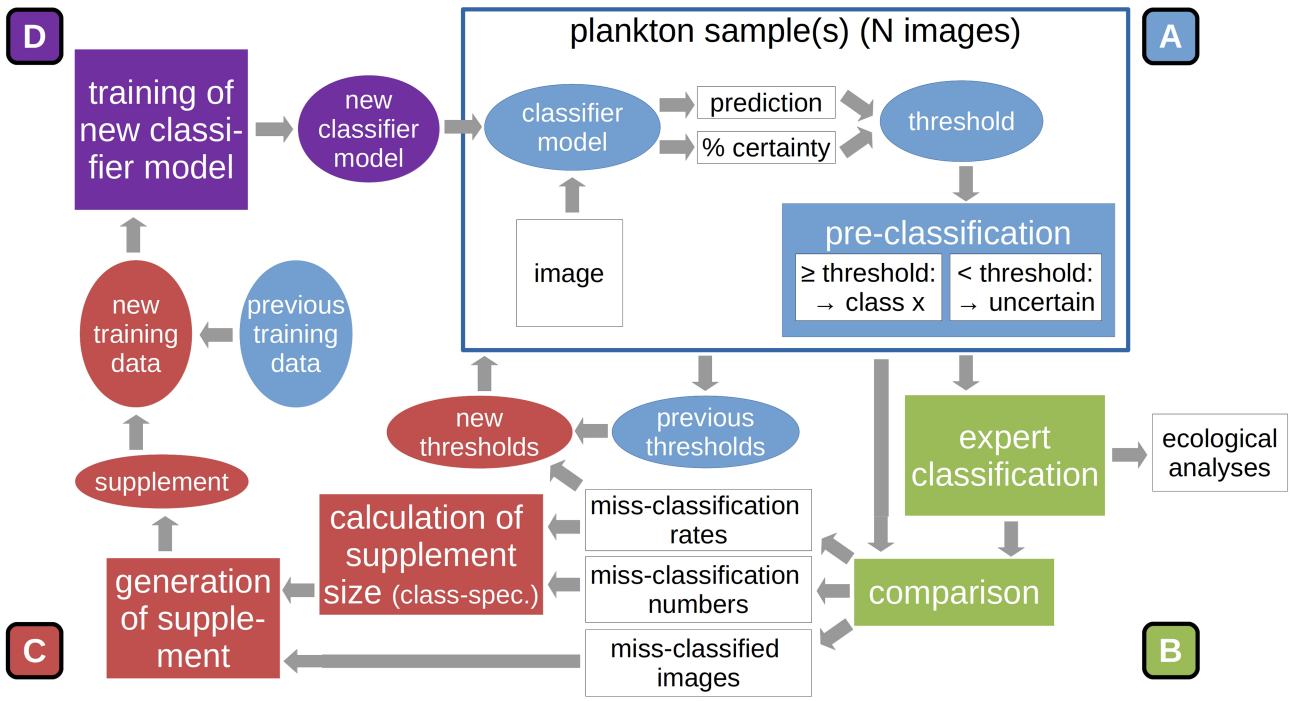


FIGURE 2 | Sequence of main procedures in the DOC pipeline: After the automatic classification **(A)** and expert validation of a set of plankton samples **(B)**, the original model training set is stocked up with a selection of miss-classified images, based on class-specific miss-classification rate. This constrained update reduces the dominance of naturally-abundant classes in the add-on set. Also, classification thresholds used in automatic pre-classification are reduced based on miss-classification rates **(C)**. A new classifier model is trained on this updated training set. The new model is used to classify the next set of plankton samples **(D)**. Further details are given in the text.

$$p_i = \frac{1-\frac{C_i}{N_i}}{\max(1-\frac{C_j}{N_j}\ for\ j \in \{class\ 1 \ldots class\ n\})}$$

$$A = F_i p_i$$

Eq. 1: Calculation of the proportion of miss-classified images to be added to the updated training dataset (top) and calculation of the number of images to be added to the training set (bottom). i = index for classes, p = proportion, C = number of correctly classified images in a given class, N = number of images assigned by expert to that class, A = number of images to be added to the training set, F = number of miss-classified images

The class-specific training-set update is the *first* part of the adaptation procedure. A marked increase in the abundance of a class that was underrepresented in the previous training set will lead to that class being better represented in the adapted version. As a *second* adaptation step, the previous threshold values for automatic culling of likely miss-classified images (see *Classification Thresholds*) are multiplied with the correct-classification rates. This reduces the threshold percentage above which a classification will be deemed correct for classes that receive an increase in training images in the first adaptation step. It is assumed that large threshold values reduce classification performance by the assignment of many in fact correctly-classified images to the "uncertain-classifications" category. By decreasing the classification threshold, the number of images correctly assigned to the predicted classes can theoretically be increased, leading to higher correct-classification rates.

4. Model training (**Figure 2D**): The model is then trained on the updated training set according to the training schedule described above. It should be noted that a completely new model instance is generated and trained. This is done to avoid an over-adaptation of the model on the training data, since re-training would mean training the existing model for an *additional* set of epochs on a still partly identical training set (no original training images are dropped during training-set updates).

After training is completed, the new model can be applied on the next batch of plankton samples, and the adaptation cycle continues anew. The DOC was devised on the notion that plankton communities change on a spatial and/or temporal gradient. It therefore makes sense to process the plankton samples in the same order as they were taken by the research vessel (or along hydrographic gradients).

Further notes on the DOC procedures can be found in SI VII.

User Application

A user application with graphical user interface was designed to aid in the implementation of the DOC pipeline. For practical purposes, it is intended that the DOC pipeline be implemented by a broad user group not necessarily familiar in the use of programming languages and/or Machine-Learning techniques. The DOC application was therefore designed to enable an end-to-end implementation of all pipeline steps described above. It consists of a series of executable, partially nested, Python scripts, one executable Bash (GNU Project, 2007) script that accesses the Python scripts and a comprehensive instruction guide describing the implementation of all DOC-pipeline steps in the application context (SI 1). None of the scripts is protected, which allows users familiar with the Python programming language to edit and change scripts in order to make custom changes to the pipeline processes, if desired.

The DOC application was written in the Python programming language, making extensive use of the *TkInter* package for graphical-user-interface design (Lundh, 2019) and of the *os* package for file-system access. One script utilized to start the application was written in the Bash command language.

The DOC application was designed for use on Linux (The Linux Foundation, San Francisco/CA) operating systems (tested on Ubuntu 18 and Linux Mint 19). It requires hardware and drivers enabling the training and application of Deep Neural Networks for image classification. For the application development and for conducting the case studies, a Nvidia® Quadro P2000 graphics-processing unit (GPU) was utilized. Further system details are given in SI II. The DOC application requires the installation of Python 3 (was tested under Python 3.6) *via* the Anaconda (Anaconda Software Distribution, 2020) distribution, and the creation of a dedicated Python environment containing i.a. the packages Tensorflow (Abadi et al., 2015) and Keras (Chollet, 2015). Full details on the environment setup are given in SI III.

The DOC application is started *via* the Bash script, whereupon each of the DOC processes can be started. The single processes can be executed in the order described above and suggested in the instruction manual, but can also be executed singularly, e.g. when only image classification, but not the implementation of the full DOC pipeline is desired.

The DOC user application, including the baseline training set, is available on zenodo.org (doi: 10.5281/zenodo.6303679).

Case Studies – North Sea Surveys

The DOC pipeline was applied to samples taken on two plankton surveys in order to test the performance of the approach.

The surveys were conducted in autumn and winter 2019 in the Western North Sea. The first survey, undertaken in September 2019, started offshore the East Coast of Scotland at approx. 57.5°N/0°E, and moved gradually closer to the British coast in a zig-zag trajectory between approx. 56.2°N and 57.5°N (**Figure 3A**). Samples were taken at these two latitudes and at approx. 57.9°N. The second survey was conducted in December 2019 in the English Channel, starting at the eastern entrance of the Channel at approx. 51.6°N/2°E, continuing south-westwards until approx. 50.25°N/-1°E, and changing direction north-eastwards, for a route parallel to but closer to the French coast than the initial trajectory (**Figure 3B**). Plankton samples were taken with a PUP net (mesh size: 55 µm) attached to a GULF VII sampler (HYDRO-BIOS Apparatebau GmbH), which was towed in double-oblique hauls.

Plankton samples were stored in 4-%-formaldehyde-seawater solution. Once in the laboratory, samples were processed using a FlowCam, following the FlowCam® Manual V 3.0 (Fluid Imaging Technologies, 2011). The FlowCam flow chamber had a depth of 300 µm, which was also the maximum size of plankton particles processed by the apparatus (the minimum particle size was determined by the PUP net mesh size of 55 µm). Flow rate was set to 1.7 mL min^{-1}, in order to achieve high image quality at

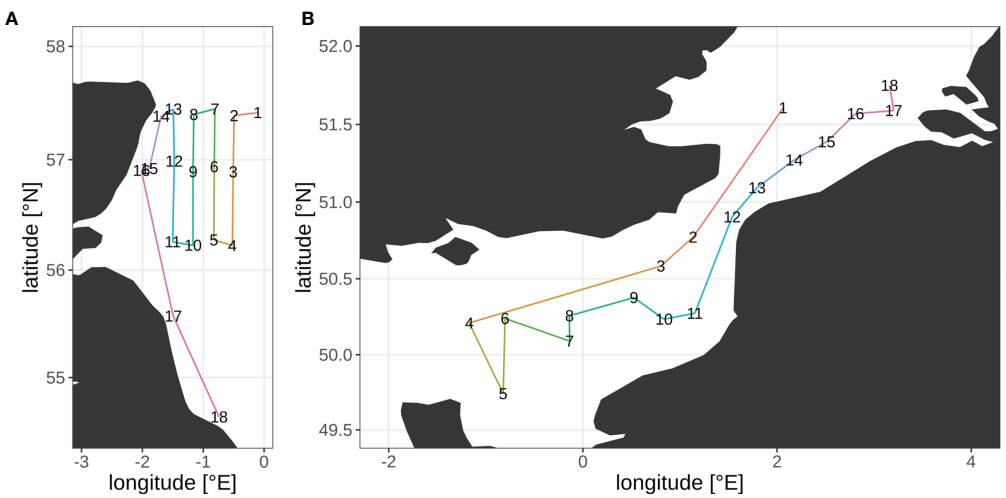


FIGURE 3 | Survey transects and location of the sampling stations from the September **(A)** and December **(B)** surveys.

an acceptable processing speed. Using the AutoImage mode of the FlowCam's Visual Spreadsheet software, images were saved for later processing.

For both surveys, the DOC pipeline was implemented for the classification of 18 samples, with the samples being processed in the sequence they were taken at sea (one sample was taken at each station). The processing sequence equals a spatial and temporal trajectory through plankton habitat. The adaptation procedure was implemented every second station, pooling the images for both stations in order to calculate the misclassification rate and to supply the information for the update of the training set. Classification performance was then calculated for each pair of stations (see below), which in the end yielded a performance trajectory over the survey samples and adaptation steps. Each mark on the trajectory thus constituted the performance of one specific model (trained on one specific version of the training set) applied to one specific set of images. In the Machine-Learning context, this information yielded the *test performance* of the models at the different adaptation steps, i.e. and indicator of their performance on non-training images under constant field conditions (e.g. Chollet, 2017).

In order to assess the importance of the continuous adaptation, a set of reference runs was performed: After each adaptation step, the current model was saved, and all subsequent samples were classified with this model (previous samples were not classified, as images contained in these were introduced into the training set during previous adaptation cycles). This way, we generated a set of reference classification trajectories in which adaptation is stopped after various numbers of samples processed (and thus on different points of the survey trajectory). This set was used to assess the value of continuous adaptation of the training set and the training of new models thereon: By comparing the performance of an adapted model to a non-adapted or less-adapted model at a specific mark on the classification trajectory, the value of adaptation could be determined for a specific sample or point on the survey trajectory. Integrated over all samples, this allowed evaluating the performance of DOC-based adaptation over the survey-/adaptation trajectory, with respect to overall advantage and potential temporal dynamics in the magnitude of adaptation advantage.

With eight adaptation steps, nine different classification trajectories resulted in total: The fully-adaptive pathway (with one adaptation cycle and the usage of a new model every second station), and eight pathways in which adaptation was stopped at a specific station (**Figure 4**).

We implemented the adaptation pathway twice for each survey to account for random effects in the adaptation procedure, generating two replicates each. These primarily include the parameter initialization before training of every model (i.e., at every adaptation step) except the base model (which was always identical) and the selection of miss-classified images for the updating of the training dataset.

We calculated recall and precision to analyze classification performance on overall- and class level, as well as cross-entropy to assess the ability to predict the plankton-community composition (see **Box 1** for details). We compared cross-entropy with class-specific differences between true and predicted relative abundance to analyze the driving factors behind changes in cross-entropy. Means and standard deviations weighted by class abundance of recall and precision were calculated for each pair of stations and each adaptation trajectory. Recall and precision values for "detritus", "clumps" and "uncertain predictions" classes were not included in the calculations of averages in order to focus on the living components of the plankton (which are the target of plankton research). More specifically, miss-classification of detritus is of little concern in research focusing on living biomass, and clumps are miss-classifications per se, since a researcher would need to analyze clumps compositions manually nevertheless. The three classes were excluded from calculation of average precision, as the direct aim of achieving high precision is to reduce the effort of removing miss-classified images from a given class folder. Since detritus, clumps and uncertain classifications are not

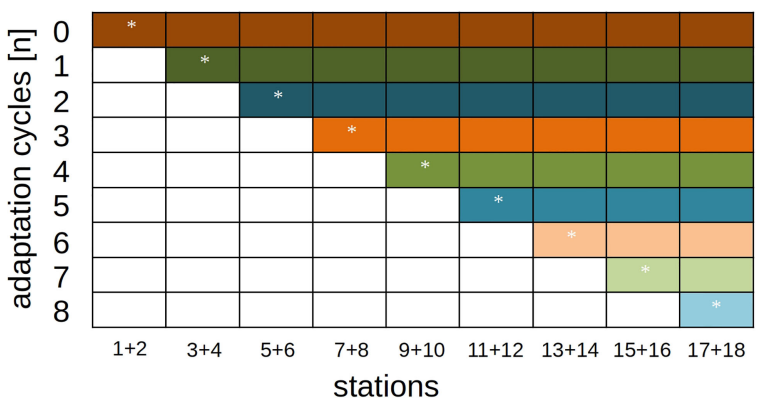


FIGURE 4 | Model-adaptation/station-classification schedule for performance analyses. The diagonal row (marked with stars) represents the fully-adaptive implementation of the DOC pipeline, where an adaptation is implemented every second station. All colored rows show reference runs where samples are classified with an existing model and without further adaptation.

BOX 1 | METRICS FOR THE ANALYSIS OF CLASSIFICATION PERFORMANCE

Recall: Recall is the class-specific ratio of correctly-classified images (true positive classifications) to the total number of images (true positive plus false negative classifications), where the total number is defined by the expert classification (eq. B1, top). This metric indicates the expert effort required to find miss-classified images in all other class folders.

Precision: Precision is the class-specific ratio of correctly-classified images (true positive classifications) to the sum of correctly-classified images (true positive classifications) and miss-classified images (false positive classifications), where the total number is defined by the expert classification (eq. B1, bottom). This metric indicates the expert effort required to find all images that were mistakenly assigned to a specific class folder

$$recall = \frac{n(true\ positive)}{n(true\ positive) + n(false\ negative)}$$

$$precision = \frac{n(true\ positive)}{n(true\ positive) + n(false\ positive)}$$

Eq. B1: Definitions of recall and precision (class-specific metrics)

Categorical cross entropy: Categorical cross entropy (hereafter referred to simply as "cross-entropy") measures the loss between a true and a predicted distribution (eq. B2). This metric is calculated from the true (derived from expert classification) and the predicted (derived from model classification) relative class abundances. Cross-entropy measures the goodness of predicting the quantitative plankton-community composition. In the present study, for classes with a predicted relative number of zero, this value was set to one divided by the total number of images in a given sample (the cross-entropy is not defined for data including zero-values; hence, one correct classification is introduced, which we assume to be a plausible stochastic error given numbers of images per sample of usually more than ten-thousand).

$$\gamma_i = -\sum_{i=1}^{N_c} a_i log\hat{a}_i$$

Eq. B2: Categorical cross entropy (γ). a = true relative abundance, â = predicted relative abundance, N_c = number of classes

Cross-entropy represents information loss between true and predicted distributions, which makes it difficult to interpret single values. Therefore, the metric is used exclusively for comparative purposes (e.g. for comparing different models) in the present study.

directly of interest in plankton research, the desire to achieve "clean" folders for these classes is comparatively low. These classes were also excluded from calculating cross entropy due to them not representing biological taxa.

Analyses and visualization were performed in R version 3.6.3 (R Core Team, 2020), partially using the packages "tidyverse" (Wickham et al., 2019), "viridis" (Garnier, 2018) and "radiant.data" (Nijs, 2020).

RESULTS

Overall performance in the fully-adaptive mode of the DOC was relatively high with regard to recall, with weighted means ranging between approx. 82 and 92% over all survey-station pairs. Precision was lower, with weighted means ranging between approx. 50-75% for the September survey, and approx. 60-80% (with one very low value of 30% at start) for December. Performance was sufficiently large to enable successful usage of the DOC application in the context of experimental research work, which benefitted from the time-savings through semi-automatic classification and model adaptation (Börner, unpubl. data).

Altogether, a fully-adaptive implementation (adaptation cycle implemented every second station) of the DOC frequently achieved comparatively high or top level mean performance in recall and precision metrics, though absolute and comparative performance varied between both survey month, and, more strongly, between classes (for details see below). Performance gains were often largest in the first one to two adaptation cycles, i.e. after the first adaptation of the baseline training set.

Recall

Overall, there were no clear trends in mean recall development over stations for the larger part of the classification trajectory, neither in the fully-adaptive nor in the less-adaptive implementations (**Figure 5**): In the September trajectory, mean recall for the fully-adaptive mode decreased from approx. 90% by approx.10% after the third station pair (stations 5 and 6), and increased again somewhat after stations 11 and 12 in both replicates (**Figures 5A, B**). Mean recall at stations 17/18 was approx. 91%. In the December trajectory, mean recall for the same mode increased strongly between stations 3/4 and stations 5/6, from approx. 20% to slightly over 90% in both replicates (**Figures 5C, D**). Recall remained at a relatively high, though slightly decreasing level, having a final value of approx. 85% at stations 17/18.

Relative performance to less adaptive DOC implementations differed initially strongly between the two surveys, but became more similar thereafter. While in the September samples no large performance difference was visible between the adapted and the baseline model at stations 2/3 (**Figures 5A, B**), recall for the more adaptive model strongly outperformed that of the less adaptive one in the December samples, as a value of over 90% was achieved with the former, while no marked performance difference to the first station (approx. 20% mean recall) was detected in the latter (**Figures 5C, D**). With the exception of the baseline model used for the December samples, which remained at low-level performance of approx. 40% mean over the trajectory, recall of the fully-adaptive mode was not markedly superior or even somewhat inferior (in the December samples) to that of less adaptive approaches, depending on the replicate. Performance of all adaptive modes converged to a relatively similar value (approx. 91%) in the final September sample (see

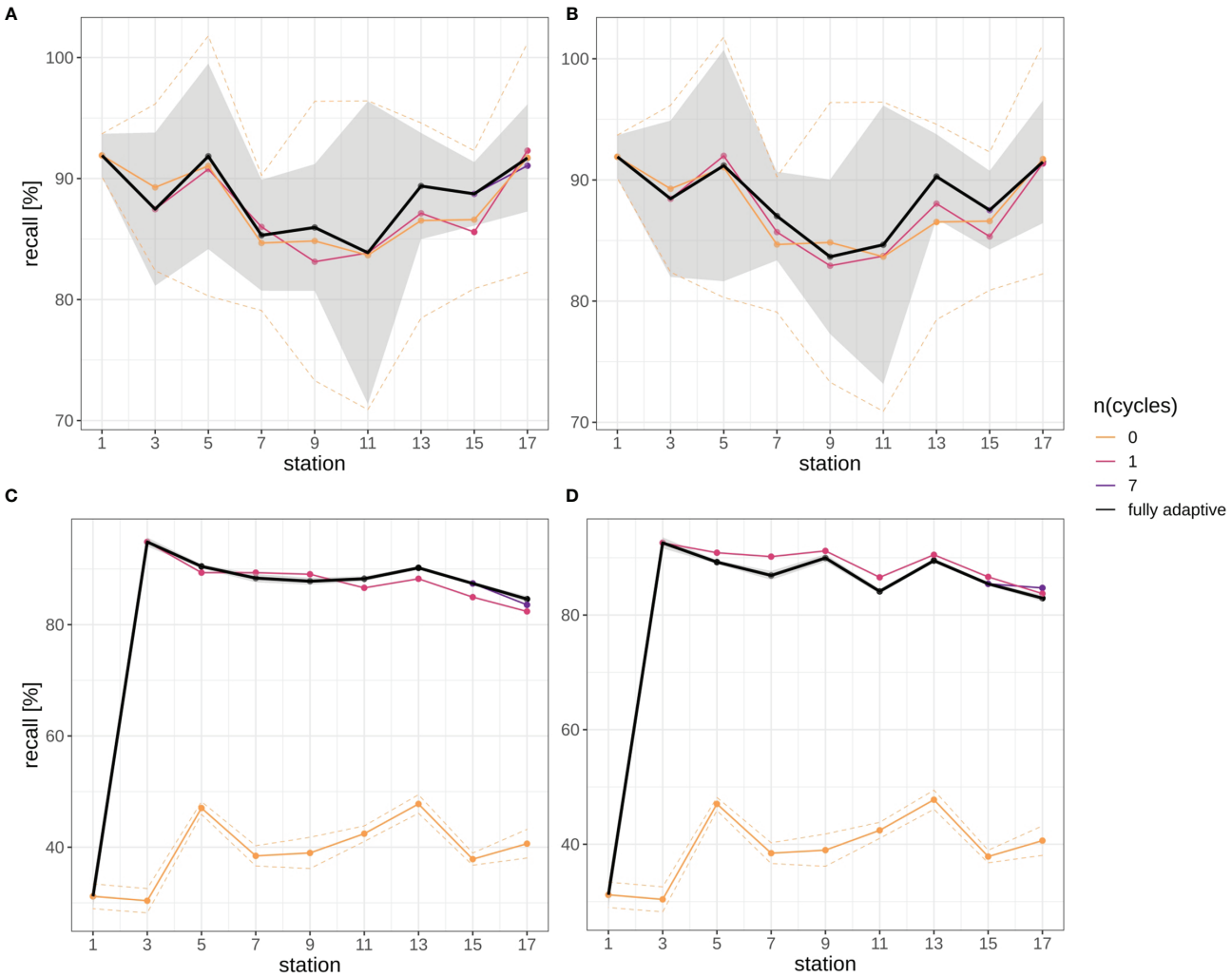


FIGURE 5 | Recall trajectories for different modes of adaptation using the DOC. Solid black line represents weighted mean of the fully-adaptive implementation, grey area denotes the corresponding weighted standard deviation. Colored solid and dashed lines represent weighted mean and weighted standard deviation of less-adaptive implementations (denoted by the number of adaptation cycles). **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*). Trajectories for all nine adaptation modes are shown in **Figure SI XI/1**. Note that weighted standard deviation for the fully-adaptive implementation in the December survey is very small compared to that in the September survey.

also **Figure SI XI/1**). Convergence was not present in the December samples.

Recall trajectories differed strongly between classes, and showed stronger fluctuations between station pairs than the weighted mean trajectory over all classes, with values of zero and 100% being reached occasionally (**Figures 6**, **SI XI/2**). Trajectories for the fully-adaptive implementation of the DOC were relatively similar between replicates, though (compare **Figure 6A vs B**, and **Figure 6C vs D**). For many classes, a recall of markedly over 90% was achieved at least occasionally in fully adaptive mode, although the identity of these classes differed between September (**Figures 6A, B**) and December surveys (**Figures 6C, D**). Classes for which a relatively high recall was frequently achieved (though not necessarily consistently over all stations) included *Ceratium* spp., *Protoperidinium* spp. (September survey only), copepods, detritus and diatoms. All other classes showed relatively high performance at least once in the recall trajectory; thus it is not possible to name classes for which recall was particularly poor. The comparative performance of the fully-adaptive implementation of the DOC varied strongly between classes, as well. Furthermore, performance also varied between surveys, and to a smaller extent between replicates. For some classes, such as bivalves (September), detritus (both surveys), diatoms (both surveys), dinoflagellates (September), foraminiferans (September), unknown taxa A, B and C (only present in September), as well as copepods (December), the fully-adaptive implementation yielded near- or top-level performance over the larger part of the stations trajectory. For other classes,

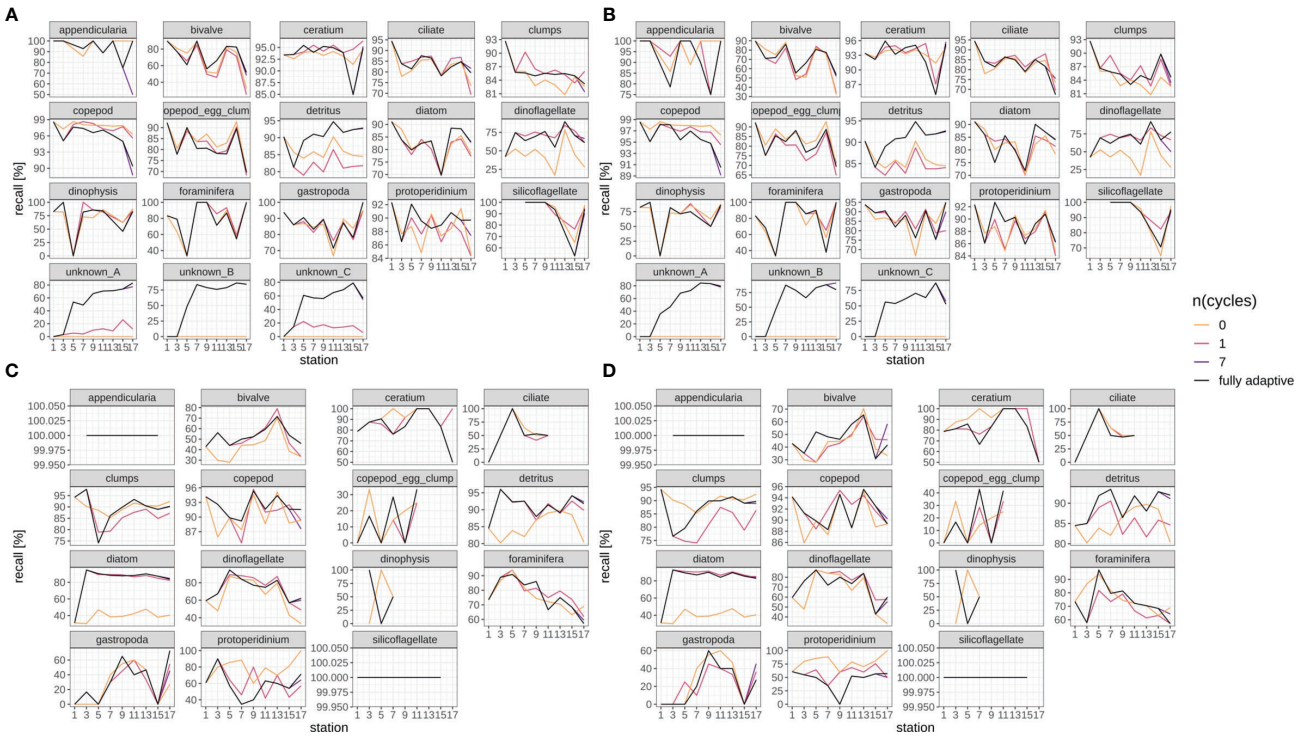


FIGURE 6 | Class-specific recall trajectories. Black line represents fully adaptive DOC implementation (training-set update every second station); colored lines represent less-adaptive implementations (denoted by the number of adaptation cycles). **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*). Trajectories for all nine adaptation modes are shown in **Figure SI XI/2**.

including copepods (September) and *Dinophysis* spp. (September), comparative performance was relatively constantly poor. It should be noted that performance differences between different modes of adaptation were of various magnitudes between classes. In most classes, the recall trajectory of the fully-adaptive implementation followed the general trend shown by all modes of adaptation.

Precision

In general, mean precision increased in both survey trajectories slightly, in all but the two least adaptive implementations of the DOC after a more variable initial phase (first two station pairs) (**Figure 7**). Mean precision increased from approx. 60% at stations 5/6 to approx. 75% at stations 15/16 in the September survey in both replicates (**Figures 7A, B**), and from approx. 65% to approx. 80% in the December survey in both replicates (**Figures 7C, D**). Mean precision then decreased again from stations 15/16 to station 17/18, from the mentioned values to approx. 63% in the September survey, and to approx. 70% in the December survey. Altogether, the trajectory of mean precision was smoother for the December survey, i.e. there was little fluctuation between adjacent station pairs.

Different from the recall trajectories, mean precision of the fully-adaptive mode of the DOC was frequently at top level compared to less-adaptive modes, in both the September and the December survey (for almost every station in the latter; **Figures 7C, D**) (see also **Figure SI XI/3**). The zero-adaptive implementation (use of the baseline model for all classifications) showed markedly lower performance than all other implementations over the full trajectory in the December samples, while lowest performance was achieved by the one-time-adapted model in the September samples. In the latter case, the performance difference was not as pronounced as in the September samples, though. While mean precision for the weakest-performing mode was relatively constant to slightly decreasing in the September survey (approx. 55% at stations 5/6 to approx. 50% at stations 17/18), it did temporarily increase from stations 7/8 to a peak at stations 13/14 (from approx. 20% to approx. 75% to approx. 25% at stations 17/18) in the December survey.

Precision trajectories differed strongly between classes and surveys (**Figures 8**, **SI XI/4**), but were mostly consistent between replicates (compare **Figures 8A vs B** and **Figures 8C vs D**), both with regard to the fully-adaptive implementation of the DOC and to its comparison with less-adaptive implementations. For most classes, precision varied strongly between adjacent stations, and did not bear a clearly increasing or decreasing trend. For many classes in the September survey (**Figures 8A, B**), the fully-

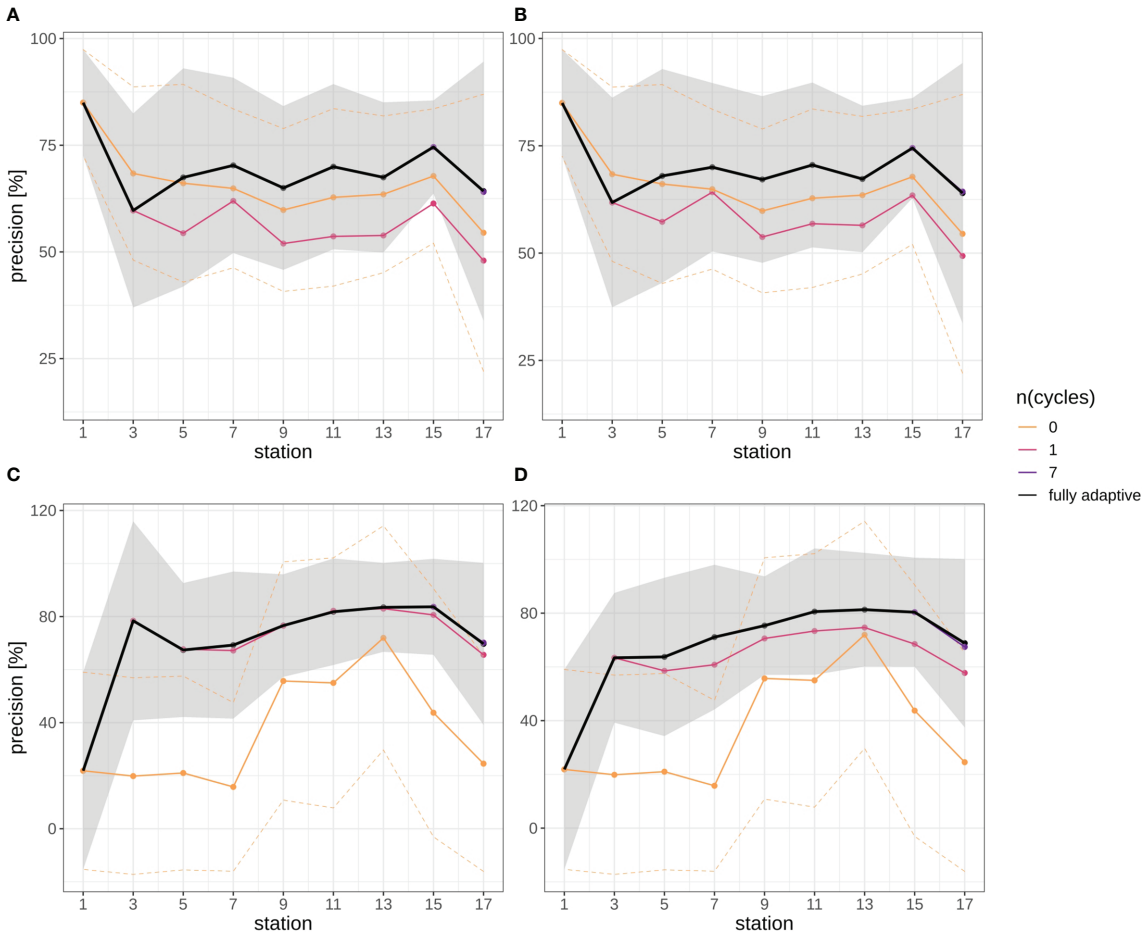


FIGURE 7 | Precision trajectories for different modes of adaptation using the DOC. Solid black line represents weighted mean of the fully-adaptive implementation, grey area denotes the corresponding weighted standard deviation. Colored solid and dashed lines represent weighted mean and weighted standard deviation of less-adaptive implementations (denoted by the number of adaptation cycles). **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*). Trajectories for all nine adaptation modes are shown in **Figure SI XI/3**.

adaptive implementation achieved near- or top-level performance over the larger part of samples; exceptions include the "clumps" class, copepod egg clumps, detritus, dinoflagellates and the two unknown taxa "A" and "B". However, unlike in the case of class-specific recall, a comparatively poor or very poor performance was observed for none of these exceptions. In the December survey (**Figures 8C, D**), the fully-adaptive implementation achieved average performance for the larger number of classes. Exceptions with near- or top-level performance over the larger part of the trajectory include bivalves, *Dinophysis* spp., foraminiferans and *Protoperidinium* spp.; for few additional classes, top-level performance was achieved in only one of the two replicates. Very poor performance was also noted for a few classes (appendicularians, copepod egg clumps, gastropods), but again only in one of the two replicates. As with class-specific recall, performance differences between differently-adaptive modes were of different magnitudes for different classes, and the precision trajectories of the fully-adaptive mode in general followed the trend of all other modes of adaptation.

Cross-Entropy

Cross-entropy in general decreased over the stations trajectory, representing an increasing similarity between true (as defined by classification expert) and predicted distributions of relative abundances of plankton classes (**Figures 9**, **SI XI/5**). By the end of the trajectory (stations 17/18), cross-entropy of the fully-adaptive implementation was decreased to approx. 90% and 40% of its value at the start of the trajectory for the September and December surveys, respectively. The cross-entropy trajectories were markedly smoother for the December survey (**Figures 9C, D**) than that for the September survey (**Figures 9A, B**), which featured an oscillatory pattern from stations five/six onwards. In the September survey, the deviation between true and predicted distributions was driven by a variety of classes, including the

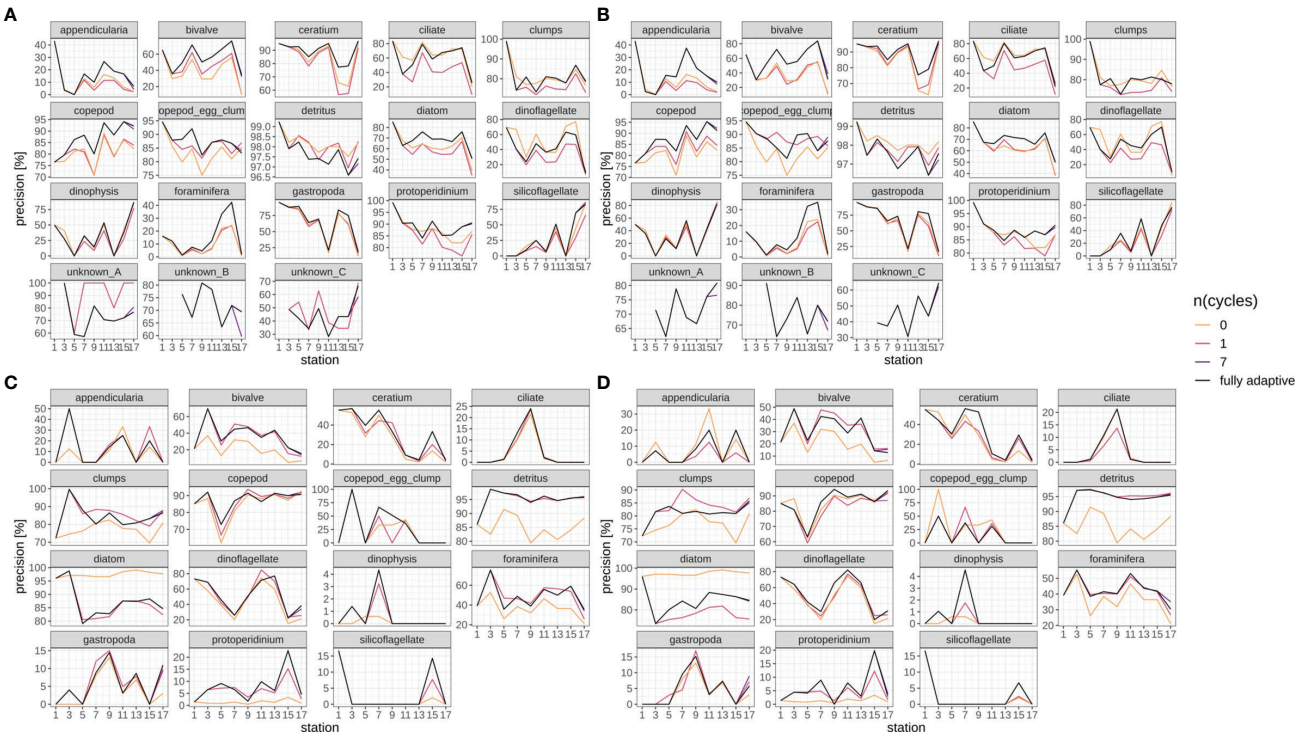


FIGURE 8 | Class-specific precision trajectories. Black line represents fully adaptive DOC implementation (training-set update every second station); colored lines represent less-adaptive implementations (denoted by the number of adaptation cycles). **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*). Trajectories for all nine adaptation modes are shown in **Figure SI XI/4**.

constantly strongly abundant diatoms and *Protoperidinium* spp. classes, as well as the occasionally strongly abundant *Ceratium* spp. class and the little-abundant unknown taxa "B" and "C" (**Figures 10A, B**). The cross-entropy decrease was primarily driven by lowered differences between predicted and true relative abundances of the diatoms class and of the two unknown taxa. Differences were not lowered by a large amount; however, the magnitude of absolute differences was not large (<< 10% at maximum). In the December survey, the deviation was almost exclusively driven by the strongly-abundant diatoms class and the little-abundant *Protoperidinium* spp. class (**Figures 10C, D**). Cross-entropy decrease was notably driven by a decrease in the difference between predicted and true relative abundance for both classes. Differences decreased by a large magnitude, from more than 50% absolute to markedly less than 20%. Cross-entropy trajectories and deviations between true and predicted abundances were very similar between replicates (compare **Figure 9/10A** vs B and **Figure 9/10 C vs D**).

Cross-entropy was lowest over all stations compared to all other adaptation modes, in the fully-adaptive implementation of the DOC (see also **Figure SI XI/5**). It was markedly higher in the two least-adaptive implementations in the September survey, and in the none-adaptive implementation in the December survey, compared to all other implementations. Relative cross-entropy dynamics over time were similar among all adaptation modes.

DISCUSSION

Our results show that adapting a classifier model to changes in the plankton community is vital for ensuring continuously high classification performance. As the comparison between the fully-adaptive and less-adaptive performance trajectories demonstrates, the standardized procedure implemented in the DOC pipeline generates suitable adaptation steps *via* training-set stock-up and reduction of classification thresholds, making the DOC an appropriate tool for implementing model adaptation

Our results confirm that continuous adaptation *via* the DOC pipeline clearly improves classification performance compared to more limited or no adaptation. The fact that performance of the classifier model improved over adaptation steps – primarily in comparison to less-adaptive scenarios, but to some extent also over survey stations, with regard to precision and cross-entropy – shows that the DOC is indeed able to cope with and actively learn from a difficult classification task. However, it is worth noting that improvement was not existing or continuous for all metrics and taxa, with e.g. mean recall not showing clear signs of improvement

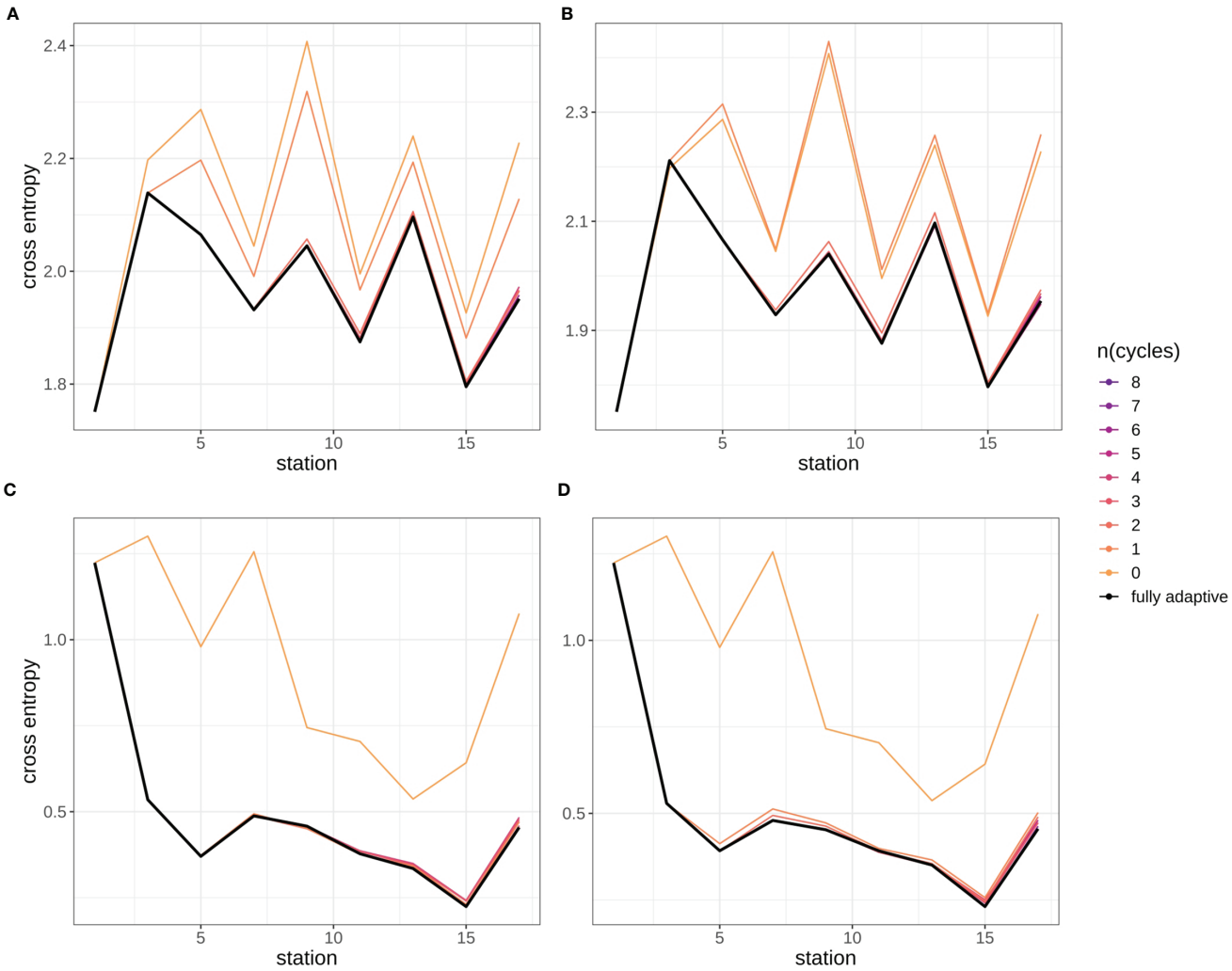


FIGURE 9 | Cross-entropy trajectories for different modes of adaptation using the DOC. Solid black line represents weighted mean of the fully-adaptive implementation, grey area denotes the corresponding weighted standard deviation. Colored solid and dashed lines represent weighted mean and weighted standard deviation of less-adaptive implementations (denoted by the number of adaptation cycles). **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*). Trajectories for all nine adaptation modes are shown in **Figure SI XI/5**.

over stations. Given that neural networks generally require large amounts of data for training (Goodfellow et al., 2016), a larger initial training set and processing of larger samples might have yielded a clearer, more universal performance improvement. Still, in the context of field research, where image data from a new region and/or time period may initially be sparse, the DOC pipeline makes effective use of the incoming data such that best possible performance is frequently achieved.

With regard to precision and cross-entropy metrics, the highest possible performance is achieved for almost every sample by the fully-adaptive implementation of the DOC, while recall performance is often at very high comparative levels. The same is true for a number of single taxa that are of strong importance in the study of the ecological function of marine plankton, e.g. in the determination of planktonic biomass available as food to commercially-harvested fish (e.g. Peck et al., 2012). Thus, fully continuous adaptation yields the best performance possible per sample when integrating over all three performance metrics.

It should be noted that the DOC was not designed with the intention of advancing classification performance in terms of improving accuracy on artificially created validation datasets. Rather, the aim was to design a procedure that achieves acceptably good performance for applied research work that focusses on abundant and broad taxonomic plankton groups, and in particular maintains that level of performance even as the

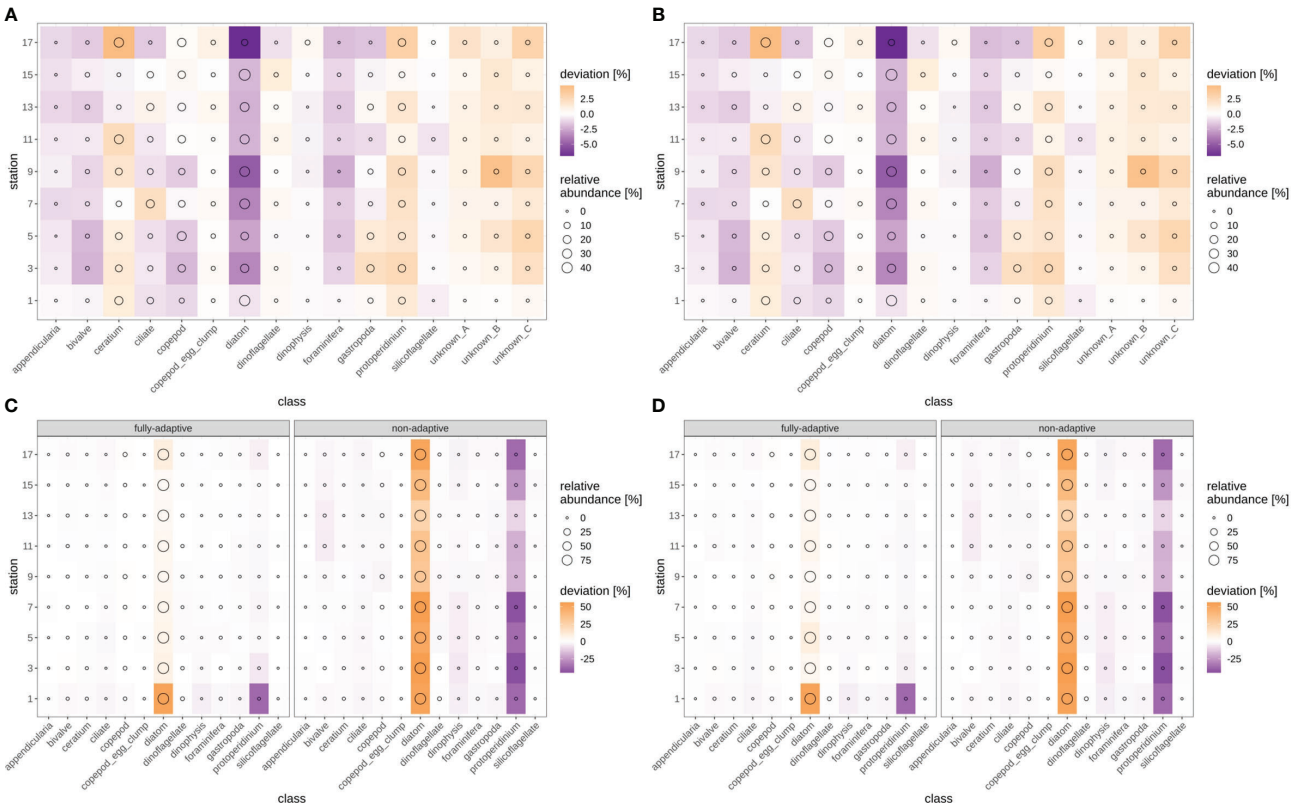


FIGURE 10 | Deviation between true and predicted (via non-validated automatic classification) relative class abundances, for the fully-adaptive DOC implementation and for classification of all samples with the baseline model (no adaptation). Circle size indicates true relative abundance. **(A)**: September survey, first replicate, **(B)** September survey, second replicate; **(C)**: December survey, first replicate, **(D)**: December survey, second replicate. Replicates differ in the random selection of images for training-set updates and random parameter initializations of models before training (see *Case Studies – North Sea Surveys*).

classifier model is confronted with changes in the plankton community. Still, with weighted mean recall ranging from 80 to over 90%, the classification performance of our model is comparable to the current state of the art, which ranges approximately between 80 and 95% (Dai et al., 2016; Luo et al., 2018; Briseno-Avena et al., 2020). Although some studies have reported very high accuracies of over 95% (Al-Barazanchi et al., 2018; Cui et al., 2018), this performance metric appears to depend strongly on the diversity of samples and on the classes chosen to report accuracy on (Luo et al., 2018; Briseno-Avena et al., 2020), which makes model comparisons difficult. Compared to recall, precision of our approach is somewhat low at 60 to 80%, but still similar to the 84% reported by Luo et al. (2018).

Given that speed and easiness of adaptation was also deemed critical for applied usage of the model, the DOC omits a thorough sample-specific model optimization (by means of re-designing the architecture of the Deep Neural Network or changing the training scheme), which might have yielded stronger performance. However, trading in performance optimization for performance reliability and easiness of adaptation did not affect the usefulness of the procedure in the particular research application it was designed for (Börner, unpubl. data) and in routine classification work.

Performance trajectories varied strongly between the two surveys, but to a lesser extent between replicates, both with regard to weighted-mean and to class-specific performance in most classes. This demonstrates that the DOC is affected by natural variability in the plankton community rather than by technical random factors (e.g. the sampling of additional training images during the adaptation procedure). In particular, performance appears to be affected by the complexity of the plankton community, as expressed *via* the degree of homogeneity of relative abundances of the plankton taxa: In the September survey, taxa that made up a very minor part of the total number of plankton organisms of the December samples (e.g. *Ceratium* spp.) were comparatively increased in relative abundance, yielding a more heterogeneous plankton community. Furthermore, the increase varied between survey stations, creating an additional spatial level of heterogeneity. Consequently, the capacity to correctly predict the distribution pattern over classes, as measured by cross-entropy, became lower, as did the capacity to improve that performance by applying the DOC over several stations. As a result, mean precision was also lower for the

September samples, as the increased abundance of non-major classes (for which fewer training images were available) likely led to more miss-classifications that reduced the purity of the model-generated class folders. Given that precision for the September samples increased slightly over stations, and markedly over the number of adaptation steps employed, it becomes visible that the DOC still led to adaptation even in this more difficult classification situation.

The fact that high recall was achieved for the diatom, copepod and some dinoflagellate classes, and that poor precision only occurred in some rather minor classes, makes the DOC useful for research questions addressing abundant plankton taxa. These can include analyses on the amount of potential plankton food available to larval fish, which combine classification with length measurements on the plankton items to calculate taxon-specific biomass estimates (e.g. Menden-Deuer and Lessard, 2000; Kiørboe, 2013). A high classification success on abundant classes thus enables a rapid estimation of the larger part of planktonic biomass, while low classification success on more rare classes does not influence biomass estimation particularly strongly. The distribution of classification performance over classes thus also shows that the DOC is particularly useful for broad quantitative analyses on the plankton community. It is not particularly well suited for qualitative surveys e.g. intended to assess the biodiversity of a certain marine area, which naturally require a classification with higher taxonomic resolution. Still, the DOC can in theory also facilitate expert-based high-level classification, as a performance improvement on a broad taxonomic scale will help the expert to better focus on the finer-scale classification of the taxon of interest. However, this would require the usage of different imaging devices, since FlowCam image resolution only allows for broad taxonomic classification even by experts (sensu Álvarez et al., 2014).

It should be pointed out that the viability of our DOC over longer series of survey samples might not necessarily follow the trends observed on the classification trajectories presented here. While the fact that performance improvements were observed in both the September and December transects indicates stability of the DOC pipeline under various ecological conditions, it remains to be seen how its performance behaves beyond the 18 stations per survey covered here. It is possible that at some point, a manual re-design of the training set might be necessary due to very drastic changes in the plankton community (note that the DOC approach does not discard training images during adaptation, leading to an increase in complexity of the training dataset over samples). Also, the continued decreasing of classification thresholds might at some point prove detrimental to classification precision due to many wrong classifications appearing in class folders instead of the "uncertain-classifications" folder. Some indications of deteriorating performance in the final survey samples (precision in September samples, recall in December samples) were observed in our case study, which might be an indication of the effects mentioned. For applied usage, we suggest to monitor the performance trajectory of the DOC in order to determine whether manual adjustments are advisable. Additionally, depending on the performance level found acceptable and the perceived chance of strong community changes, it may not be necessary to implement the DOC adaptation scheme after each processed sampled. It is up to the user to decide on a good trade-off between the performance improvement achieved through model adaptation and the time saved by not implementing the DOC adaptation steps.

The DOC pipeline proposed by us is not the first attempt at continually maintaining or improving model performance as new plankton samples are classified and validated in applied use: Gorsky et al. (2010) initially made use of a plankton training set not specifically built for their study, and obtained improved classification results once adding validated images from their samples and training a model on this. They continued this procedure until further improvements became marginal. Li et al. (2022) systematized a scheme of human-model interaction, where validated images are added to the training set during applied usage of the classifier. However, neither study has explicitly quantified performance decay nor the effect of training-set updates over a spatial trajectory as presented here. Also, both used expert validation to grow the training set in a rather non-systematized manner, and classification thresholds (to accept or discard a model classification as "uncertain") were not adapted. While a non-systematized growing of the training set achieved marked performance improvements in both studies, our work shows that careful systematized training-set updates and adaptation of classification thresholds initially improve and then maintain classification performance without the need for continuously adding all validated images, which would lead to increased training durations.

Our DOC application joins a growing number of pipelines and applications designed to facilitate the embedding of machine-learning models into the workflow of plankton classification. These include the Prince William Sound Plankton Camera (Campbell et al., 2020), the Scripps Plankton Camera system (Orenstein et al., 2020) and the MorphoCluster clustering workflow (Schröder et al., 2020). All of these applications incorporate a step of manual validation in the workflow; however, none of them incorporate a dedicated standardized scheme for dynamic adaptation, as proposed by our study. The MorphoCluster is an exception to the super-vised classification schemes presented in most other applications, since it makes use of an unsupervised clustering algorithm that groups the plankton images in a data-driven manner. It therefore appears not to require a dedicated dynamic adaptation; however, the interpretation of the resulting clusters may be less straight-forward than the expert check of a machine classification. While the MorphoCluster appears particularly useful for *in-situ* monitoring studies that focus on fine-resolution taxon recognition, we assume that our DOC may be of more convenient use in quantitative studies that primarily address a fixed set of broad taxonomic groups.

Compared to other applications that often present an end-to-end system from field sampling to classification, and related hardware, our DOC covers a relatively small part of the overall workflow. Future extensions of our application would primarily address a more direct coupling to size measurements on the

plankton images (used, together with a class-specific conversion factor, to calculate the biomass of every plankton item (e.g. Menden-Deuer and Lessard, 2000; Kiørboe, 2013), as well as to the preceding photography in the FlowCam. Further extensions could include the incorporation of automatic performance monitoring in order to give advice to the user of when a manual re-design of the training set or a manual adaptation of classification thresholds might be necessary.

CONCLUSIONS

Our DOC proves to be a capable tool for adapting a classifier model on a plankton community changing over the spatial and temporal dimension. Our method continually delivers high or highest performance compared to non- or less-adaptive approaches, especially for abundant classes, though is subject to sample-specific variability in the difficulty of classification. Combined with the streamlining of the adaptation process and the availability of an easy-to-operate user interface, the DOC serves as an aide for quantitative plankton analysis on a broad taxonomic level that performs reliably under changing community patterns.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

JC wrote the manuscript. JC, GB, and MM conceived the study. GB and JC did the classification experiments and analyses. JC did the programming of the DOC application. AL-U, MM, and CM provided input and revisions to the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

JC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820989 (project COMFORT, Our common future ocean in the Earth system – quantifying coupled cycles of carbon, oxygen, and nutrients for determining and achieving safe operating spaces with respect to tipping points). GB was funded by the German Research Foundation (DFG) under project THRESHOLDS (Disentangling the effects of climate-driven processes on North Sea herring recruitment through physiological thresholds, MO 2873-3-1)

ACKNOWLEDGMENTS

The authors wish to like Jens Floeter and Rene Plonus for helpful comments, as well as André Harmer for initial guidance in the programming with Python and Rachel Harmer for help with manual annotation of the baseline training set. Parts of the intellectual content of this manuscript were included, in very preliminary form, in the master's thesis (Conradt, 2020).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.868420/full#supplementary-material

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., et al. (2015). TensorFlow: Large-Scale Machine Learning. Available at: www.tensorflow.org (Accessed on 29th March, 2022).

Al-Barazanchi, H., Verma, A., and Wang, S. X. (2018). Intelligent Plankton Image Classification With Deep Learning. *Int. J. Comput. Vis. Robot.* 8, 561–571. doi: 10.1504/IJCVR.2018.095584

Álvarez, E., López-Urrutia, Á., and Nogueira, E. (2012). Improvement of Plankton Biovolume Estimates Derived From Image-Based Automatic Sampling Devices: Application to FlowCam. *J. Plankton Res.* 34, 454–469. doi: 10.1093/plankt/fbs017

Álvarez, E., Moyano, M., López-Urrutia, Á., Nogueira, E., and Scharek, R. (2014). Routine Determination of Plankton Community Composition and Size Structure: A Comparison Between FlowCAM and Light Microscopy. *J. Plankton Res.* 36, 170–184. doi: 10.1093/plankt/fbt069

Anaconda Software Distribution (2020) *Anaconda Documentation* (Anaconda Inc). Available at: https://docs.anaconda.com/ (Accessed 5th July, 2020).

Asch, R. G., Stock, C. A., and Sarmiento, J. L. (2019). Climate Change Impacts on Mismatches Between Phytoplankton Blooms and Fish Spawning Phenology. *Glob. Change Biol.* 25, 2544–2559. doi: 10.1111/gcb.14650

Beaugrand, G. (2012). Unanticipated Biological Changes and Global Warming. *Mar. Ecol. Prog. Ser.* 445, 293–301. doi: 10.3354/meps09493

Beaugrand, G., Ibañez, F., Lindley, A., and Reid, P. C. (2002). Diversity of Calanoid Copepods in the North Atlantic and Adjacent Seas: Species Associations and Biogeography. *Mar. Ecol. Prog. Ser.* 232, 179–195. doi: 10.3354/meps232179

Briseño-Avena, C., Schmid, M. S., Swieca, K., Sponaugle, S., Brodeur, R. D., and Cowen, R. ,. K. (2020). Three-Dimensional Cross-Shelf Zooplankton Distributions Off the Central Oregon Coast During Anomalous Oceanographic Conditions. *Prog. Oceanogr.* 188, 102436. doi: 10.1016/j.pocean.2020.102436

Campbell, R., Roberts, P., and Jaffe, J. (2020). The Prince William Sound Plankton Camera: A Profiling in Situ Observatory of Plankton and Particulates. *ICES J. Mar. Sci.* 77, 1440–1455. doi: 10.1093/icesjms/fsaa029

Capuzzo, E., Lynam, C. P., Barry, J., Stephens, D., Forster, R. M., Greenwood, N., et al. (2017). A Decline in Primary Production in the North Sea Over 25 Years, Associated With Reductions in Zooplankton Abundance and Fish Stock Recruitment. *Glob. Change Biol.* 24, e352–e364. doi: 10.1111/gcb.13916

Castellani, C., and Edwards, M. (2017). *Marine Plankton* (Oxford, UK: Oxford University Press).

Chollet, F. (2015) *Keras.* Available at: https://www.keras.io (Accessed 5th July, 2020).

Chollet, F. (2017). *Deep Leaning With Python* (New York, NY: Manning Publications Company).

Conradt, J. (2020). *Automated Plankton Image Classification With a Capsule Neural Network. [Master's Thesis]* (Hamburg (DE: Universität Hamburg).

Cui, J., Wei, B., Wang, C., Yu, Z., Zheng, H., Zheng, B., et al. (2018). "Texture and Shape Information Fusion of Convolutional Neural Network for Plankton Image Classification," in *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)* (Washington, DC: IEEE Computer Society) 1–5. doi: 10.1109/OCEANSKOBE.2018.8559156

Culverhouse, P. F., Williams, R., Reguera, B., Herry, V., and González-Gil, S. (2003). Do Experts Make Mistakes? A Comparison of Human and Machine Identification of Dinoflagellates. *Mar. Ecol. Prog. Ser.* 247, 17–25. doi: 10.3354/meps247017

Dai, J., Yu, Z., Zheng, H., Zheng, B., and Wang, N. (2016). ""A Hybrid Convolutional Neural Network for Plankton Classification"," in *Computer Vision – ACCV 2016 Workshops*. Eds. C. S. Chen, J. Lu and K. K. Ma (Cham, CH: Springer), 102–114. doi: 10.1007/978-3-319-54526-4_8

Dam, H. G., and Baumann, H. (2017). ""Climate Change, Zooplankton and Fisheries"," in *Climate Change Impacs on Fisheries and Aquaculture*. Eds. B. F. Phillips and M. Pérez-Ramírez (Hoboken, NJ: Wiley-Blackwell), 851–874.

Davis, C. S., Gallager, S. M., Berman, S. M., Haury, L. R., and Strickler, J. R. (1992). The Video Plankton Recorder (VPR): Design and Initial Results. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* 36, 67–81.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). "Imagenet: A Large-Scale Hierarchical Image Database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009)* (Washington, DC: IEEE Computer Society), 248–255. doi: 10.1109/CVPR.2009.5206848

Durant, J. M., Molinero, J.-C., Ottersen, G., Reygondeau, G., Stige, L. C., and Langangen, Ø. (2019). Contrasting Effects of Rising Temperatures on Trophic Interactions in Marine Ecosystems. *Sci. Rep.* 9, 15213. doi: 10.1038/s41598-019-51607-w

Fluid Imaging Technologies (2011). FlowCam® Manual Version 3.0. Available at: http://www.ihb.cas.cn/fxcszx/fxcs_xgxz/201203/P020120329576952031804.pdf. (Accessed on 28th March, 2022).

Frederiksen, M., Edwards, M., Richardson, A. J., Halliday, N. C., and Wanless, S. (2006). From Plankton to Top Predators: Bottom-Up Control of a Marine Food Web Across Four Trophic Levels. *J. @ Anim. Ecol.* 75, 1259–1268. doi: 10.1111/j.1365-2656.2006.01148.x

Garnier, S. (2018) *Viridis: Default Color Maps From 'Matplotlib'*. Available at: https://CRAN.R-project.org/package=viridis.

Glorot, X., and Bengio, J. (2010). Understanding the Difficulty of Training Deep Forward Neural Networks. *J. Mach. Learn. Res. – Proceedings Track*, 9. 249–256.

GNU Project (2007). *Free Software Foundation* (Bash [Unix shell program]). Available at: https://www.gnu.org/gnu/gnu.html. (Accessed on 28th March, 2022).

González, P., Álvarez, E., Díez, J., López-Urrutia, Á., and del Coz, J. J. (2016). Validation Methods for Plankton Image Classification Systems. *Limnol. Oceanogr. Methods* 15, 221–237. doi: 10.1002/lom3.10151

Goodfellow, I. J., Bengio, Y., and Courville, A. (2016). *Deep Learning* (Cambridge, MS: The MIT Press).

Goodwin, M., Halvorsen, K. T., Jiao, L., Knausgård, K. M., Martin, A. H., Moyano, M., et al. (2022). Unlocking the Potential of Deep Learning for Marine Ecology: A Review Exemplified Through Seven Established and Emerging Applications. *ICES J. Mar. Sci.* 79, 319–336. doi: 10.1093/icesjms/fsab255

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital Zooplankton Image Analysis Using the ZooScan Integrated System. *J. Plankton Res.* 32, 285–303. doi: 10.1093/plankt/fbp124

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., et al. (2018). Recent Advances in Convolutional Neural Networks. *Pattern Recognit.* 77, 354–377. doi: 10.1016/j.patcog.2017.10.013

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55

Kiørboe, T. (2013). Zooplankton Body Composition. *Limnol. Oceanogr.* 58, 1843–1850. doi: 10.4319/lo.2013.58.5.1843

Kraberg, A., Metfies, K., and Stern, R. (2017). ""Sampling, Preservation and Counting of Samples I: Phytoplankton"," in *Marine Plankton*. Eds. C. Castellani and M. Edwards (Oxford, UK: Oxford University Press), 91–103.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet Classification With Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* 1, 1097–1105. doi: 10.1145/3065386

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., et al. (1989). Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Comput.* 1, 541–551. doi: 10.1162/neco.1989.1.4.541

LeCun, Y., Kavukcuoglu, K., and Farabet, C. (2010). "Convolutional Networks and Applications in Vision," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems* (Washington, DC: IEEE Computer Society), 253–256. doi: 10.1109/ISCAS.2010.5537907

Li, J., Chen, T., Yang, Z., Chen, L., Liu, P., Zhang, Y., et al. (2022). Development of a Buoy-Borne Underwater Imaging System for *in Situ* Mesoplankton Monitoring of Coastal Waters. *IEEE J. Ocean. Eng.* 47, 88–110. doi: 10.1109/JOE.2021.3106122

Li, Z., Zhao, F., Liu, J., and Qiao, Y. (2014). Pairwise Nonparametric Discriminant Analysis for Binary Plankton Image Recognition. *IEEE J. Ocean. Eng.* 39, 695–701. doi: 10.1109/JOE.2013.2280035

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Lumini, A., and Nanni, L. (2019). Deep Learning and Transfer Learning Features for Plankton Classification. *Ecol. Inform.* 51, 33–43. doi: 10.1016/j.ecoinf.2019.02.007

Lundh, F. (2019) *An Introduction to TkInter*. Available at: http://www.pythonware.com/library/tkinter/introduction/index.htm (Accessed 5th July, 2020).

Lundh, F., and Ellis, M. (2019) *Python Imaging Library (PIL)*. Available at: http://www.pythonware.com/products/pil/ (Accessed 5th July, 2020).

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., et al. (2018). Automated Plankton Image Analysis Using Convolutional Neural Networks. *Limnol. Oceanogr. Methods* 16, 814–827. doi: 10.1002/lom3.10285

Malde, K., Handegard, N. O., Eikvil, L., and Salberg, A.-B. (2020). Machine Intelligence and the Data-Driven Future of Marine Science. *ICES J. Mar. Sci.* 77, 1274–1285. doi: 10.1093/icesjms/fsz057

McKinney, W. (2010). "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference (SciPy), Austin / TX, USA*, eds. van der Walt, S., and Millman, J. (London, UK: lulu.com), 51–56. doi: 10.25080/Majora-92bf1922-00a

Menden-Deuer, S., and Lessard, E. J. (2000). Carbon to Volume Relationships for Dinoflagellates, Diatoms, and Other Protist Plankton. *Limnol. Oceanogr.* 45, 569–579. doi: 10.4319/lo.2000.45.3.0569

Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrerra, F. (2012). A Unifying View on Dataset Shift in Classification. *Pattern Recogn.* 45, 521–530. doi: 10.1016/j.patcog.2011.06.019

Nagelkerken, I., Goldenberg, S. U., Ferreira, C. M., Russell, B. D., and Connell, S. D. (2017). Species Interactions Drive Fish Biodiversity Loss in a High-CO_2 World. *Curr. Biol.* 27, 2177–2184. doi: 10.1016/j.cub.2017.06.023

Nijs, V. (2020) *Radiant.Data: Data Menu for Radiant: Business Analytics Using R and Shiny*. Available at: https://CRAN.R-project.org/package=radiant.data.

Oliphant, T. E. (2006). *A Guide to NumPy* (US: Trelgol Publishing).

Oliphant, T. E. (2007). Python for Scientific Computing. *Comput. Sci. Eng.* 9, 10–20. doi: 10.1109/MCSE.2007.58

Orenstein, E. C., and Beijbom, O. (2017). "Tranfer Learning and Deep Feature Extraction for Planktonic Image Data Sets," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. (Washignton, DC: IEEE Computer Society). 1082–1088. doi: 10.1109/WACV.2017.125

Orenstein, E. C., Beijbom, O., Peacock, E. E., and Sosik, H. (2015) *WHOI-Plankton – A Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification* (arXiv). Available at: https://arXiv:1510.00745 (Accessed 5th July, 2020).

Orenstein, E. C., Ratelle, D., Brieseño-Avena, C., Carter, M. L., Franks, P. J. S., Jaffe, J. S., et al. (2020). The Scripps Plankton Camera System: A Framework and Platform for *in Situ* Microscopy. *Limnol. Oceanogr. Methods* 18, 681–695. doi: 10.1002/lom3.10394

Peck, M. A., Huebert, K. B., and Llopiz, J. K. (2012). Intrinsic and Extrinsic Factors Driving Match-Mismatch Dynamics During the Early Life History of Marine Fishes. *Adv. Ecol. Res.* 47, 177–302. doi: 10.1016/B978-0-12-398315-2.00003-X

Perez, F., and Granger, B. E. (2007). IPython: A System for Interactive Scientific Computing. *Comput. Sci. Eng.* 9, 21–29. doi: 10.1109/MCSE.2007.53

Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S., and Floeter, J. (2021). Automatic Plankton Image Classification – Can Capsules and Filters Help Cope With Data Set Shift? *Limnol. Oceanogr. Methods* 19, 176–195. doi: 10.1002/lom3.10413

Raybaut, P. (2017) *Spyder Documentation – Release 3*. Available at: http://pythonhosted.org.

R Core Team (2020). *R: An Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing). Available at: https://www.R-project.org/.

Riederer, C. (2016) *Welcome to Dplython's Documentation*. Available at: https://pythonhosted.org/dplython/ (Accessed 5th July, 2020).

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning Representations by Back-Propagating Errors. *Nature* 323, 533–536. doi: 10.1038/323533a0

Russell, F. S. (1939). Hydrographical and Biological Conditions in the North Sea as Indicated by Plankton Organisms. *ICES J. Mar. Sci.* 14, 171–192. doi: 10.1093/icesjms/14.2.171

Scholkopf, B., and Smola, A. (2002). *Learning With Kernels: Support Vector Machines, Regularization, Optimization and Beyond* (Cambridge, MS: The MIT Press).

Schröder, S.-M., Kiko, R., and Koch, R. (2020). MorphoCluster: Efficient Annotation of Plankton Images by Clustering. *Sensors* 20, 3060. doi: 10.3390/s20113060

Sieracki, C. K., Sieracki, M. E., and Yentsch, C. S. (1998). An Imaging-in-Flow System for Automated Analysis of Microplankton. *Mar. Ecol. Prog. Ser.* 168, 285–296. doi: 10.3354/meps168285

Simonyan, K., and Zisserman, A. (2015) *Very Deep Convolutional Networks for Large-Scale Image Recognition* (arXiv). Available at: https://arxiv.org/abs/1409.1556 (Accessed 5th July, 2020).

Stern, R., Taylor, C., and Sadri, S. (2017). ""Protozooplankton: Foraminifera"," in *Marine Plankton*. Eds. C. Castellani and M. Edwards (Oxford, UK: Oxford University Press), 194–197.

Tang, X., Lin, F., Samson, S., and Remsen, A. (2006). Binary Plankton Image Classification. *IEEE J. Ocean. Eng.* 31, 728–735. doi: 10.1109/JOE.2004.836995

Tang, X., Stewart, W. K., Huang, H., Gallager, S. M., Davis, C. S., Vincent, L., et al. (1998). Automatic Plankton Image Recognition. *Artif. Intell. Rev.* 12, 177–199. doi: 10.1023/A:1006517211724

van Rossum, G. (1995). *Python Tutorial* (Amsterdam, NL: Centrum voor Wiskunde en Informatica (CWI). Technical Report CS-R9526.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., Francois, R., et al. (2019). Welcome to the Tidyverse. *J. Open Source Software* 4, 1686. doi: 10.21105/joss.01686

Wiebe, P. H., Bucklin, A., and Benfield, M. (2017). ""Sampling, Preservation and Counting of Samples II: Zooplankton"," in *Marine Plankton*. Eds. C. Castellani and M. Edwards (Oxford, UK: Oxford University Press), 104–135.

Winder, M., and Sommer, U. (2012). Phytoplankton Response to a Changing Climate. *Hydrobiologia* 698, 5–16. doi: 10.1007/s10750-012-1149-2

Author Disclaimer: The work reflects only the authors' view; the European Commission and their executive agency are not responsible for any use that may be made of the information the work contains.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

ORIGINAL RESEARCH
published: 22 April 2022
doi: 10.3389/fmars.2022.867893

COI Metabarcoding of Zooplankton Species Diversity for Time-Series Monitoring of the NW Atlantic Continental Shelf

Ann Bucklin[1], Paola G. Batta-Lona[1], Jennifer M. Questel[2], Peter H. Wiebe[3], David E. Richardson[4], Nancy J. Copley[3] and Todd D. O'Brien[5]*

[1] Department of Marine Sciences, University of Connecticut, Groton, CT, United States, [2] College of Fisheries and Ocean Sciences, University of Alaska, Fairbanks, AK, United States, [3] Department of Biology, Woods Hole Oceanographic Institution, Woods Hole, MA, United States, [4] National Oceanic and Atmospheric Association (NOAA) Northeast Fisheries Science Center, Narragansett, RI, United States, [5] National Oceanic and Atmospheric Association (NOAA) Fisheries, Office of Science and Technology Science Center, Silver Spring, MD, United States

OPEN ACCESS

Edited by:
Xosé Anxelu G. Morán, Spanish Institute of Oceanography, Spain

Reviewed by:
John Kenneth Pearman, Cawthron Institute, New Zealand
Mie Hylstofte Sichlau Winding, Greenland Climate Research Centre, Greenland

***Correspondence:**
Ann Bucklin
ann.bucklin@uconn.edu

Specialty section:
This article was submitted to Marine Ecosystem Ecology, a section of the journal Frontiers in Marine Science

Received: *01 February 2022*
Accepted: *25 March 2022*
Published: *22 April 2022*

Citation:
Bucklin A, Batta-Lona PG, Questel JM, Wiebe PH, Richardson DE, Copley NJ and O'Brien TD (2022) COI Metabarcoding of Zooplankton Species Diversity for Time-Series Monitoring of the NW Atlantic Continental Shelf. Front. Mar. Sci. 9:867893. doi: 10.3389/fmars.2022.867893

Marine zooplankton are rapid-responders and useful indicators of environmental variability and climate change impacts on pelagic ecosystems on time scales ranging from seasons to years to decades. The systematic complexity and taxonomic diversity of the zooplankton assemblage has presented significant challenges for routine morphological (microscopic) identification of species in samples collected during ecosystem monitoring and fisheries management surveys. Metabarcoding using the mitochondrial Cytochrome Oxidase I (COI) gene region has shown promise for detecting and identifying species of some – but not all – taxonomic groups in samples of marine zooplankton. This study examined species diversity of zooplankton on the Northwest Atlantic Continental Shelf using 27 samples collected in 2002-2012 from the Gulf of Maine, Georges Bank, and Mid-Atlantic Bight during Ecosystem Monitoring (EcoMon) Surveys by the NOAA NMFS Northeast Fisheries Science Center. COI metabarcodes were identified using the MetaZooGene Barcode Atlas and Database (https://metazoogene.org/MZGdb) specific to the North Atlantic Ocean. A total of 181 species across 23 taxonomic groups were detected, including a number of sibling and cryptic species that were not discriminated by morphological taxonomic analysis of EcoMon samples. In all, 67 species of 15 taxonomic groups had ≥ 50 COI sequences; 23 species had >1,000 COI sequences. Comparative analysis of molecular and morphological data showed significant correlations between COI sequence numbers and microscopic counts for 5 of 6 taxonomic groups and for 5 of 7 species with >1,000 COI sequences for which both types of data were available. Multivariate statistical analysis showed clustering of samples within each region based on both COI sequence numbers and EcoMon counts, although differences among the three regions were not statistically significant. The results demonstrate the power and potential of COI metabarcoding for identification of species of metazoan zooplankton in the context of ecosystem monitoring.

Keywords: zooplankton, metabarcoding, cytochrome oxidase I, species diversity, ecosystem monitoring, Northwest Atlantic continental shelf

INTRODUCTION

Metabarcoding of Zooplankton Diversity

Patterns of species diversity of the many taxonomic groups of marine zooplankton are key characteristics of ocean ecosystems, determining their function, sustainability, and responses to environmental variation and anthropogenic impacts, including climate change (Sherman et al., 2002; Friedland et al., 2020). Pelagic ecosystems of the NW Atlantic Ocean have been monitored and studied over many decades, providing an invaluable time-series record of zooplankton diversity and abundance (Wiebe et al., 2012; O'Brien et al., 2013).

DNA metabarcoding of zooplankton samples has been used in association with ecosystem monitoring and management in recent years (Mohrbeck et al., 2015; Deagle et al., 2017; Djurhuus et al., 2018; Blanco-Bercial, 2020; Matthews et al., 2021). These efforts have used different molecular protocols and bioinformatics pipelines (Bucklin et al., 2021a). A subset of the studies have focused on discrimination and identification of species diversity across the zooplankton assemblage based on metabarcoding analysis using portions of the mitochondrial cytochrome oxidase I (COI) gene (reviewed by Bucklin et al., 2021b), which has been widely used to identify species of marine organisms (Bucklin et al., 2011).

The potential of COI metabarcoding has been examined from many perspectives, including accuracy of species-level resolution of biodiversity (Brown et al., 2015; Leray and Knowlton, 2017; Schroeder et al., 2021), availability of reference sequence databases resulting from DNA barcoding efforts (Andújar et al., 2018; Porter and Hajibabaei, 2018; Bucklin et al., 2021b; Singh et al., 2021), and prospects for quantitative analysis related to species abundance and/or biomass (Lamb et al., 2018; Matthews et al., 2021). Challenges and disadvantages of COI as a metabarcode include lack of universal primers and missing groups due to primer-mismatch (Deagle et al., 2014; Clarke et al., 2017; Hajibabaei et al., 2019). A number of studies have evaluated the results using COI metabarcodes based on parallel analysis using multiple metabarcode gene regions, including hypervariable regions of 18S rRNA (Djurhuus et al., 2018; Steffani et al., 2018; Giebner et al., 2020; Pitz et al., 2020; Brandão et al., 2021; Pappalardo et al., 2021; Questel et al., 2021; Zhao et al., 2021).

An important consideration for species identification based on COI metabarcoding is the availability of COI barcode sequences for identified specimens (Andújar et al., 2018; Porter and Hajibabaei, 2018; Steinke et al., 2021). The goal of taxonomically-complete COI reference databases for marine zooplankton, with search capacities for targeted taxonomic groups and geographic regions (Bucklin et al., 2021b), is a priority for an international collaborative effort sponsored by the Scientific Committee for Oceanic Research (SCOR), MetaZooGene (WG157; see https://metazoogene.org/).

There is broad interest and considerable enthusiasm for potential applications of metabarcoding for ecosystem assessment and fisheries management (Bourlat et al., 2013; Ji et al., 2013; Kelly et al., 2014; Kelly, 2016; Goodwin et al., 2017; Aylagas et al., 2018). A particular focus is the importance of rapid, accurate, and reliable species-level characterization of time-space variability of the taxonomically-complex pelagic assemblage (Andújar et al., 2018).

Monitoring the NW Atlantic Continental Shelf

Time-series monitoring of the NW Atlantic continental shelf was established by the Northeast Fisheries Science Center (NEFSC) of the National Marine Fisheries Service (NMFS) in 1977 and continues into the present (Cox and Wiebe, 1979; Hare and Kane, 2012). Surveys are scheduled to sample six times each year throughout four regions (**Figure 1**), with collection of environmental (hydrographic) data and samples at a stratified-random selection of standard station locations (see https://www.fisheries.noaa.gov/new-england-mid-atlantic/ecosystems/monitoring-ecosystem-northeast). Morphological microscopic examination of zooplankton samples from NEFSC surveys has allowed analysis and interpretation of temporal and spatial patterns of variability across the region (Kane, 2007; Kane, 2011; O'Brien et al., 2013). The time-series records have provided clear evidence that the region is experiencing rapid climate change (O'Brien et al., 2013; Friedland et al., 2020). Regime shifts (i.e., persistent changes in the structure and function of ecosystems) have been documented during the 1990s and 2000s in pelagic community structure (Pershing et al., 2005; Walsh et al., 2015; Morse et al., 2017), including zooplankton diversity (Head and Sameoto, 2007; Record et al., 2010; Johnson et al., 2011; Bi et al., 2014). NEFSC time-series records revealed a marked increase in zooplankton displacement volume on Georges Bank (GB) and the Gulf of Maine (GoM) in ~1990 (O'Brien et al., 2013; **Figure 2**). Another regime shift was evident in ~2000, when the earlier changes were partially reversed. In the Gulf of Maine (GoM), zooplankton diversity increased markedly during the early 1990s and decreased rapidly about 2000 (Record et al., 2010; Johnson et al., 2011).

Monitoring efforts in North Atlantic pelagic ecosystems (Wiebe et al., 2012; O'Brien et al., 2013) have provided an essential foundation for recognizing and understanding regime shifts (Borja, 2014; Conversi et al., 2015; Morse et al., 2017; Stern et al., 2018). Regular, standardized, and sustained analysis of zooplankton diversity and abundance based on morphological microscopic examination of zooplankton samples has also provided opportunities for evaluating the accuracy, reliability, and power of rapidly-developing approaches to molecular analysis of zooplankton biodiversity, including DNA metabarcoding (Bucklin et al., 2016; Bucklin et al., 2019).

Integrative Morphological – Molecular Analysis of Zooplankton Diversity

The EcoMon zooplankton database includes records from morphological (microscopic) taxonomic counts reported as numbers per $10m^2$ for zooplankton taxa, which are identified to species when possible (Kane, 2007; Kane, 2011). A total of 186 zooplankton species of 14 taxonomic groups of metazoan zooplankton have been detected in net samples collected during surveys of the NW Atlantic continental shelf since 1977; of these, 43 species of 9 taxonomic groups have been

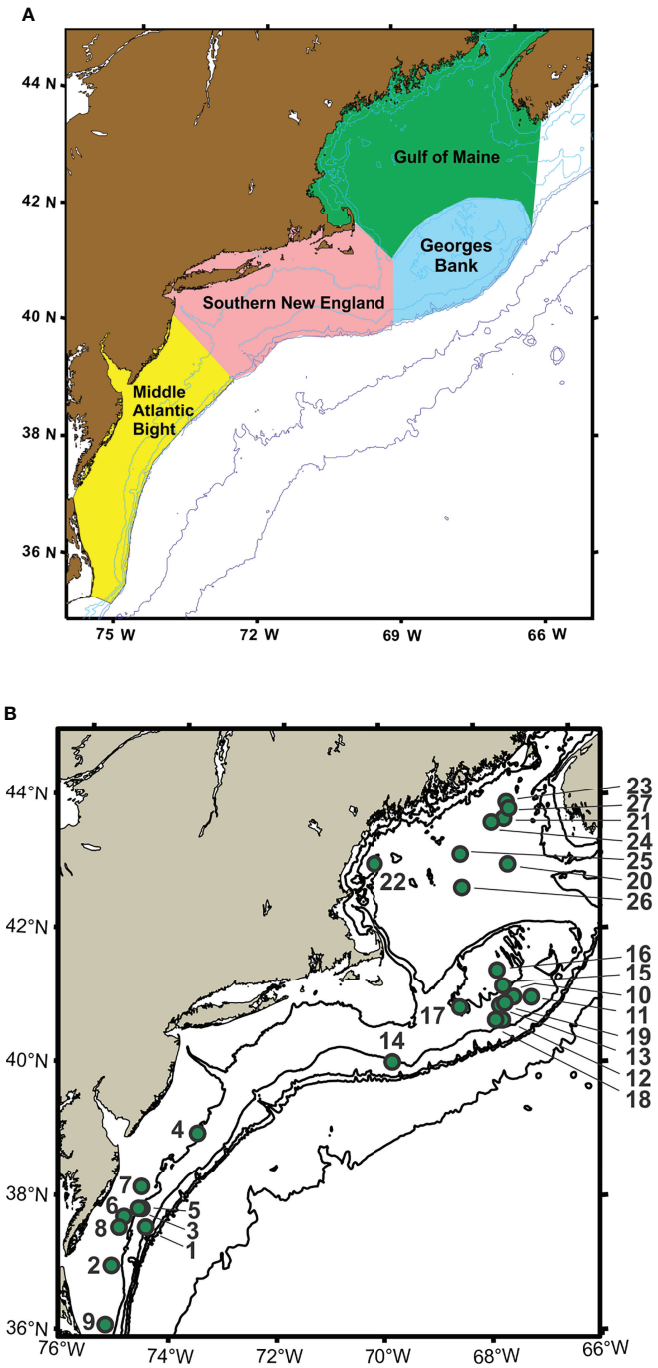


FIGURE 1 | Maps showing regions **(A)** and sample collection locations **(B)** for Northeast Fisheries Science Center (NEFSC) Ecosystem Monitoring Surveys (EcoMon). Stations are identified by number (see **Table 1**). Modified from Bucklin et al. (2019).

recorded at >1% frequency of occurrence among all samples (David Richardson, pers. comm.). An additional 237 taxa, including groups of congeneric species, genera, and higher taxonomic groups, are are listed in the EcoMon database, https://www.ncei.noaa.gov/archive/accession/0187513; accessed May 23, 2021 (NMFS/NEFSC, 2019).

Molecular analysis of EcoMon samples has been carried out since 2000, first in partnership with the international project, ZooGene (http://www.zoogene.org/), and from 2004 to 2010 with the Census of Marine Zooplankton (CMarZ; http://www.cmarz.org/), an ocean realm field project of the Census of Marine Life (CoML; Bucklin et al., 2010). The partnership, with

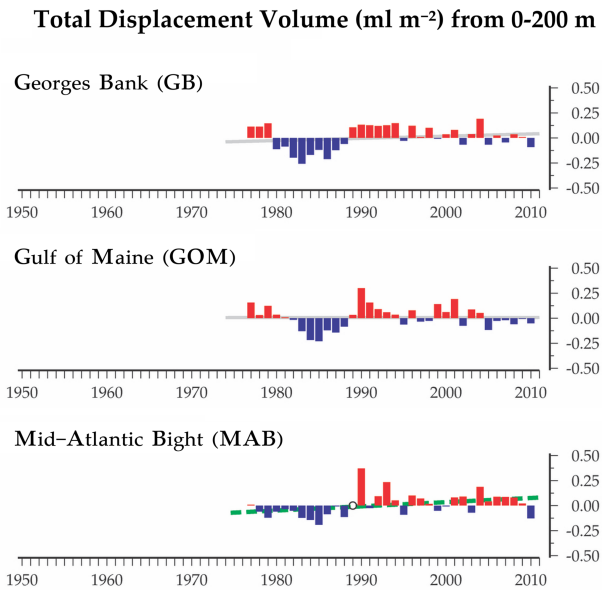


FIGURE 2 | Time-series data for total displacement volume of zooplankton from EcoMon regions: Mid-Atlantic Bight (MAB), Georges Bank (GB), and Gulf of Maine (GoM). Graphs show anomalies based on average values from 1977 to 2010. Positive anomalies are red; negative anomalies are blue. The lines represent the linear regression of the annual anomalies versus year; colors of the lines indicate the statistical significance of the relationship: dashed green ($p < 0.05$); grey (non-significant). Figure modified from O'Brien et al. (2013).

collection and preservation of samples for molecular analysis, continues today in association with another international program, MetaZooGene (https://metazoogene.org/) Working Group 157 of the Scientific Committee for Oceanic Research (SCOR).

Bucklin et al. (2019) analyzed 27 EcoMon samples collected from 2002 – 2012 using DNA metabarcoding of the V9 hypervariable region of 18S rRNA, with sequences classified into 28 taxonomic groups of zooplankton. The conserved nature of the 18S rRNA gene allows detection of taxa across the spectrum of marine zooplankton, but does not accurately resolve or identify species (Blanco-Bercial, 2020; Govindarajan et al., 2021; Questel et al., 2021). Bucklin et al. (2019) reported significant positive correlations between V9 18S rRNA sequence numbers and microscopic counts for 7 taxonomic groups for which both types of data were available, with significant regressions for three groups: Calanoida, Gastropoda, and Chaetognatha. These results provided promising evidence that DNA metabarcoding using V9 18S rRNA can provide accurate classification and relative quantification for targeted zooplankton groups, which are important goals for applications for ecosystem monitoring (Lamb et al., 2018; Matthews et al., 2021).

MATERIALS AND METHODS

Collection and Selection of Samples for Analysis

Zooplankton samples for this study were collected by the NOAA Northeast Fisheries Science Center (NEFSC) Oceans and Climate Branch during surveys by the Ecosystem Monitoring Program (EcoMon) of the NW Atlantic continental shelf (Kane, 2007; Kane, 2011; Hare and Kane, 2012; Bucklin et al., 2019). Surveys are designed to sample four regions of the shelf ecosystem (**Figure 1**). Samples for morphological taxonomic analysis were collected following a standard protocol (Richardson et al., 2010), with both day and night sampling using a 61-cm bongo net fitted with a 333-µm mesh net; oblique tows were a minimum of 5-min in duration and sampled from the surface to within 5 m of the seabed or to a maximum depth of 200 m. A mechanical flowmeter was fitted in the mouth of each net to record the volume sampled. Samples were preserved in 5% formalin and archived at the NEFSC.

Zooplankton samples for genetic analysis were collected during EcoMon survey cruises at 5 randomly-selected locations in each region. Sampling was done using a 20-cm bongo net with 165-µm mesh nets, which was attached to the same cable and deployed with the 61-cm bongo nets. Differences in the opening diameter and mesh size of the nets used for collection of samples for genetic analysis may have resulted in differences between the sets of samples, but the methods were unchanged across all years, regions, and stations, and the resulting time-series patterns of variability were evaluated with this caveat. Samples were preserved immediately in 95% undenatured ethanol, which was changed 24 hr after collection. Samples were transported to and archived at the University of Connecticut, with long-term storage in walk-in freezers (-20°C)

Morphological Taxonomic Analysis

Sample sorting and identification was done at the Morski Instytut Rybacki Plankton Sorting and Identification Center

(Szczecin, Poland). Zooplankton samples were split to an aliquot containing approximately 500 specimens; individuals were sorted, counted, and identified to the lowest possible taxon (Kane, 2007; Kane, 2011). Data recorded include abundance measured by area (conc/10m^2) and volume (conc/100m^3) for selected taxononomic groups and species of zooplankton, including fish larvae. The morphological species count data were downloaded from: https://www.ncei.noaa.gov/archive/accession/0187513, accessed May 23, 2021 (NMFS/NEFSC, 2019).

Metabarcoding Analysis

A total of 27 samples was selected for metabarcoding analysis, including one sample collected in each of three regions, Georges Bank (GB), Gulf of Maine (GoM), and Mid-Atlantic Bight (MAB), during May/June of 2002 – 2012 (**Figure 1**; **Table 1**). The samples are the same ones analyzed by metabarcoding using the V9 hypervariable region of 18S rRNA by Bucklin et al. (2019). There were a number of sampling gaps due to cancelled cruises, bad weather, and other causes: no samples were analyzed for MAB in 2003 or GoM in 2006; no samples were analysed for 2008; and only a single GB sample was analysed for 2012; COI metabarcoding data are missing for DE1105-25; EcoMon count data are missing for DE0305-38 and DE1105-127 (**Table 1**). The collection site of one sample (AL0605-53 #13) is correctly shown on GB, although was listed within the EcoMon region of Southern New England (SNE).

Extraction and Quantification of Genomic DNA

Samples were quantitatively sub-divided using a box splitter (Motoda, 1959) to reduce zooplankton volume to ~25 mL. The sample was then washed with distilled water; inserted into a 50 mL Falcon tube above 35 μm Nitex mesh, which served to suspend the material and dry the pellet; and centrifuged at 3500 g for 4 min. The pellet was moved to a new 50 mL Falcon tube, and SDS buffer (Tris-HCl, 10 mM; EDTA, pH 8.0, 100mM; NaCl, 200mM; SDS 1%) 3 mL or equal to pellet volume, whichever was smaller) was added. The sample was homogenized using a hand-held homogenizer (D1000, Thomas Scientific) with saw tooth blade for 4 min at level 5. Proteinase K (MP Biomedicals) was added (0.2 mg/mL of sample) and tubes were incubated overnight in a water bath at 55-56°C. After centrifugation (3500 g for 15 min), 400 uL of the supernatant was transferred to individual sterile 2 mL Eppendorf tubes for storage as necessary at -20 or -80°C. Total genomic DNA was extracted using the E.Z.N.A Mollusc DNA kit (Omega Bio-tek) following manufacturer instructions. All samples yielded DNA of sufficient quantify and quality for metabarcoding analysis. Total genomic DNA was quantified on a Thermo-Fisher NanoDrop 2000 and normalized to a final concentration of 5 ng/μl.

PCR Amplification, Library Preparation, and Sequencing

Purified DNA was used to amplify a 313 base-pair (bp) region of mitochondrial cytochrome oxidase I (COI) using the primers: mlCOIintF and jgHCO2198 (Geller et al., 2013; Leray et al., 2013). Forward and reverse primers were altered for multiplexed sequencing by adding 5' adapters (Illumina, Inc., San Diego, CA). The PCR reaction used 20ng of DNA, with Platinum Taq reagents, 4μL buffer, 2.4μL MgCl, 0.8 μL dNTPs, 0.2μL HiFi Taq Polymerase, and 0.8μL of each primer (10μM), with the following protocol: one denaturation cycle at 94°C for 60 sec; 38 cycles of 94°C for 30 sec, 46°C for 30 sec, 72°C for 90 sec; a single extension cycle of 72°C for 5 min; and an infinite hold at 4°C. COI amplicons were checked for successful amplification by running in a 2% agarose gel with a 50 bp marker.

Library preparation entailed adding index primers in a second PCR amplification of the purified amplicons using a master mix composed of (per sample): 5.0 μl purified PCR product; 5 μl Nextera XT Index 1 Primer; 5 μl Nextera XT Index 2 Primer; 25 μl 2x KAPA HiFi HotStart ReadyMix; 10 μl PCR-grade water; for a total volume of 50 μl. The PCR protocol was: 95°C for 3 min; 8 cycles of: 95°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec; and 1 cycle of 72°C for 5 min. The indexed PCR product was purified using AMPure XP beads, with a final elution volume of 25 μL. Successful library attachment was verified using an TapeStation 4200 D1000 High Sensitivity assay (Agilent Technologies). Libraries were quantified using a Qubit 3.0 fluorometer, normalized according to amplicon size, pooled, and denatured with 0.2 N NaOH. Bi-directional sequencing was carried out at the University of Connecticut Center for Genomic Innovation (CGI; https://cgi.uconn.edu/) using an Illumina MiSeq sequencer using the MiSeq Reagent Nano Kit Ver.2 (500 cycles; 1 million clusters) spiked with a minimum of 20% PhiX (Illumina, Inc.). All samples were analyzed in a single MiSeq run, including negative controls and replicate samples, to allow full intercomparison of samples (e.g., sequencing depth).

Sequence Quality Assessment and Bioinformatics

Demultiplexed reads for the COI region were processed using a custom script for the Mothur pipeline (Ver. 1.44.3; Schloss et al., 2009) and run on the Xanadu computing cluster of the UConn Computational Biology Core (CBC; https://bioinformatics.uconn.edu/). Contiguous sequences (contigs) were assembled from forward and reverse Illumina MiSeq reads and trimmed to the overlapping section. Sequences were trimmed to a uniform length by removing the beginning and terminal ends of sequences that extended beyond the targeted COI gene region. Sequences containing ambiguous bases, quality Phred scores < 30, and with lengths shorter than 150 bp were removed from analysis. Concerns that PCR error may contribute to errors in biodiversity assessment (Kelly et al., 2019) were addressed by using the UNOISE method (Edgar and Flyvbjerg, 2015) within Mothur (Ver. 1.44.3) to de-noise aligned sequences. Sequences were screened for chimeras using the VSEARCH command (Rognes et al., 2016); sequences with chimeras were removed from analysis.

Among the 27 samples used for COI metabarcoding, 12 were selected at random for sequencing in the same MiSeq run, using a second aliquot from the PCR product for the target COI metabarcode region. The results from these samples were

TABLE 1 | Collection locations and dates for EcoMon Survey samples analyzed.

Sequential Stn No.	Region	Cruise	Station	Latitude (N)	Longitude (W)	Collection Date	Time (UTC)	Temp Surface (°C)	Salinity Surface (ppt)	Temp max depth (°C)	Salinity max depth (ppt)	Water column depth (m)	Day / Night
1	MAB	AL0206	11	37.582	74.498	24-MAY-2002	03:20	15.14	33.79	11.19	34.37	69	N
2	MAB	AL0405	14	36.987	75.107	26-MAY-2004	02:00	21.32	29.41	6.69	33.37	39	N
3	MAB	AL0505	24	37.857	74.582	26-MAY-2005	10:22	10.21	31.88	8.99	31.97	52	D
4	MAB	AL0605	14	39.017	73.572	02-JUN-2006	10:51	15.59	31.92	13.34	34.60	51	D
5	MAB	DE0706	15	37.857	74.645	24-MAY-2007	01:00	13.90	33.30	9.90	33.46	44	N
6	MAB	DE0905	35	37.727	74.913	31-MAY-2009	09:06	18.30	32.38	9.80	33.08	32	N
7	MAB	DE1004	25	38.188	74.608	26-MAY-2010	09:48	15.39	31.24	7.16	32.35	43	D
8	MAB	DE1105	25	37.563	74.997	05-JUN-2011	07:23	23.14	27.33	9.00	32.65	31	N
9	MAB	HB1202	33	36.100	75.170	3-Jun-2012	14:40	20.00	32.50	14.60	33.20	38	D
10	GB	AL0206	75	41.023	67.373	31-MAY-2002	08:33	12.02	32.64	7.76	32.68	68	N
11	GB	DE0305	15	41.008	67.023	26-MAY-2003	03:18	7.65	32.71	9.01	33.45	69	N
12	GB	AL0405	53	40.682	67.612	03-JUN-2004	01:38	8.63	32.89	5.95	33.27	150	N
13	GB	AL0505	70	40.900	67.650	02-JUN-2005	10:22	9.39	32.01	6.21	32.27	205	D
14	GB	AL0605	53	40.100	69.790	06-JUN-2006	11:50	15.57	32.49	13.65	35.68	106	D
15	GB	DE0706	76	41.193	67.575	30-MAY-2007	10:04	10.28	33.13	9.61	33.13	41	D
16	GB	DE0905	73	41.437	67.678	05-JUN-2009	02:47	10.20	32.68	10.20	32.76	36	N
17	GB	DE1004	73	40.895	68.443	03-JUN-2010	02:50	10.71	32.45	10.71	32.45	54	N
18	GB	DE1105	83	40.688	67.745	10-JUN-2011	08:22	13.97	32.21	7.65	32.77	78	N
19	GB	HB1202	71	40.933	67.550	11-JUN-2012	11:10	12.00	32.90	9.20	33.00	73	D
20	GoM	AL0206	108	43.015	67.382	04-JUN-2002	01:00	9.47	32.33	8.66	34.69	224	N
21	GoM	DE0305	38	43.700	67.425	29-MAY-2003	02:00	7.07	32.50	7.36	34.09	210	N
22	GoM	AL0405	117	43.067	70.110	07-JUN-2004	12:42	10.90	31.52	3.11	32.76	133	D
23	GoM	AL0505	111	43.945	67.360	06-JUN-2005	12:34	8.54	31.81	6.90	33.95	218	D
24	GoM	DE0706	112	43.650	67.683	03-JUN-2007	10:43	9.71	32.36	7.04	34.20	232	D
25	GoM	DE0905	112	43.187	68.343	09-JUN-2009	08:38	10.50	31.84	6.60	33.99	192	N
26	GoM	DE1004	129	42.688	68.330	08-JUN-2010	05:51	12.98	31.83	9.12	34.57	204	N
27	GoM	DE1105	127	43.848	67.315	14-JUN-2011	05:44	8.46	31.79	8.29	34.21	197	N

Region names are abbreviated: Mid-Atlantic Bight (MAB), Georges Bank (GB), Gulf of Maine (GoM).
EcoMon station numbers indicated for each cruise are based on sampling carried out at a subset of stations, which are assigned sequential station numbers.

treated as technical replicates and examined using the Wilcoxon rank sum test carried out in MatLab (Ver. 2020B), which indicated no statistical differences among the replicates. In all cases, data from only one of each pair technical replicates were used for definitive statistical analysis.

Taxonomic Assignment to Zooplankton Groups and Species

Taxonomic identification of COI metabarcode sequences was determined using a naïve Bayesian classifier algorithm in Mothur (Ver. 1.44.3). Taxonomic assignments for species-level

identifications used bootstrap values ≥ 97% after 100 iterations. Before performing zooplankton community analyses, sequences with abundances < 2 (i.e., global singletons) across the entire dataset were removed. Taxonomic classification and species identification were based upon the MetaZooGene Barcode Atlas and Database (MZGdb; https://metazoogene.org/MZGdb), which includes publicly available COI barcode sequences downloaded from GenBank and BOLD (Bucklin et al., 2021b). The results reported in this study used the North Atlantic Atlas and Database (https://metazoogene.org/MZGdb-NATL), (which includes barcodes for all holo- or mesozooplankton species reported from the region. To ensure completeness of the regional database, DNA barcodes are included in the North Atlantic MZGdb, even if the specimen used for DNA sequencing was obtained from a different ocean region.

The MZGdb allows targeted searches by taxonomic groups and geographic regions of interest and provides the capacity to map and visualize the geographical distribution of species observations and collection locations of specimens used for DNA sequencing on global to regional scales. The MZGdb is a collaboratively-developed product of MetaZoogene (SCOR WG157; https://scor-int.org/group/157/) and the Coastal & Oceanic Plankton Ecology, Production, & Observation Database (COPEPOD, https://www.st.nmfs.noaa.gov/copepod/). The MZGdb builds upon a taxonomically-arranged database of zooplankton observations, biometric traits, photographs, and DNA barcode data in COPEPEDIA (https://copepedia.org/), which stores and compiles information at multiple taxonomic levels, including species and taxonomic groups.

Analysis of COI Sequence Numbers and EcoMon Counts

COI sequence numbers are reported for species of the same 17 taxonomic categories used to classify metabarcoding results for V9 18S rRNA by (Bucklin et al. 2019) (**Supplementary Table S1**). Morphological (microscopic) abundance counts per $10m^2$ were recorded for these same taxonomic groups (**Supplementary Table S2**). Results were further analyzed for six taxonomic groups (Calanoida, Cyclopoida, Eucarida, Chaetognatha, Hydrozoa, and Gastropoda) for which both metabarcoding results and microscopic taxonomic counts were available for most samples; groups with zeros or missing data for many of the samples were not included in the analyses. All data were transformed (Log10+1) prior to analysis.

Numbers of COI sequences and abundance counts from morphological taxonomic analysis reported in the EcoMon database were statistically compared for the six groups using functional regression analysis (Ricker, 1973). Multivariate statistical analyses of COI sequence numbers for the six zooplankton groups were carried out to examine patterns of variation between regions and years using MatLab (Ver. 2020B). One distance measure used was Bray-Curtis dissimilarity coefficient (Bray and Curtis, 1957; McCune et al., 2002), with results displayed by cluster diagrams. Differentiation among the 3 regions was evaluated by Non-Metric Multidimensional Scaling (NMDS) and by Nonparametric (Permutation-based) MANOVA using the FATHOM Toolbox for MatLab (Jones, 2017; https://www.usf.edu/marine-science/research/matlab-resources/index.aspx/). The Shannon (H) and Simpson (D) Diversity Indices (Pielou, 1977) were calculated using COI sequence numbers for the 6 taxonomic groups for each sample. Regression analysis of values for the two indices showed no difference in all cases and results are reported for the Shannon Index (H).

COI metabarcodes identified a total of 181 species across 23 taxonomic groups of metazoan zooplankton listed in the summary file (Wang et al., 2007) generated by Mothur (Ver. 1.44.3; Schloss et al., 2009). Many of these species showed very low sequence numbers including many zeros (**Supplementary Table S3**), which prevented statistical analysis. Multivariate statistical analysis was carried out for 23 species with total COI sequence numbers >1,000 across all samples (**Table 2**). All data were transformed (Log10+1) for analysis. Patterns of variation in COI sequence numbers for the 23 species were statistically evaluated between regions and years in MatLab (Ver. 2020B), using the same tests as for the group comparisons, including NMDS and Nonparametric (Permutation-based) MANOVA (Jones, 2017; https://www.usf.edu/marine-science/research/matlab-resources/index.aspx/), Bray-Curtis dissimilarity coefficient (Bray and Curtis, 1957; McCune et al., 2002), and Shannon Diversity Index (H; Pielou, 1977).

Morphological counts are available in the EcoMon database for 7 of the 23 species with >1,000 COI sequences; the remaining 16 species were either grouped at a higher taxonomic level or were not listed among species to be identified (NMFS/NEFSC, 2019; https://www.ncei.noaa.gov/archive/accession/0187513). Numbers of COI sequences and miscroscopic counts were statistically compared for these species using functional regression analysis (Ricker, 1973).

RESULTS

Comparative Molecular - Morphological Analysis of Zooplankton Groups

Metabarcoding using a portion of the COI barcode region was carried out for 27 samples collected from three EcoMon Survey regions (GB, GoM, MAB) during 2002-2012 (**Table 1**). The taxonomic groups selected were used in a previous metabarcoding study by Bucklin et al. (2019) that analyzed the V9 hypervariable region of 18S rRNA for the same EcoMon samples. Metabarcoding yielded a total of 4,992,468 COI sequences and 1,404,242 Amplified Sequence Variants (ASVs). Considering all samples together, numbers of COI sequences and ASVs for species with >50 sequences (**Supplementary Tables S3**, **S4**) were highly significantly correlated across all taxonomic groups ($r = 0.978$, $p = 6.102\ e^{-11}$). The definitive analysis used COI sequence numbers.

Statistical analysis focused on 6 taxonomic groups (Calanoida, Cyclopoida, Eucarida, Gastropoda, Hydrozoa, Chaetognatha) for which sufficient numbers of observations (non-zero) were

TABLE 2 | Abundant species with >1000 COI sequences totaled across 27 samples from EcoMon Surveys 2002-2012.

	Taxonomic Group & Species	COI Seqs
	Annelida	
	Paramphinome jeffreysii	8,481
	Polygordius jouinae	1,008
	Calanoida	
*	*Calanus finmarchicus*	1,153,889
	Calanus hyperboreus	313,176
*	*Nannocalanus minor*	2,177
*	*Centropages hamatus*	361,230
	Clausocalanus pergens	1,015
	Microcalanus pusillus	5,483
*	*Pseudocalanus moultoni*	21,437
*	*Pseudocalanus newmani*	90,847
*	*Temora longicornis*	9,735
	Eucarida	
*	*Meganyctiphanes norvegica*	249,717
*	*Thysanoessa longicaudata*	1,465
	Peracarida	
	Gammarus annulatus	1,077
*	*Evadne nordmanni*	18,206
	Teleostei	
	Hygophum hygomii	1,076
	Hydrozoa	
	Obelia geniculata	35,278
	Melicertum octocostatum	1,370
	Siphonophorae	
	Nanomia cara	208,667
	Gastropoda	
*	*Limacina retroversa*	2,614
*	*Clione limacina*	1,628

Data for these species were used for multivariate statistical analysis of patterns of zooplankton diversity among years and EcoMon regions. Asterisks (*) indicate species that are counted by morphological (microscopic) analysis in EcoMon Survey samples.

available for both molecular (metabarcoding) and morphological (counts) data. These groups were the same ones analyzed for V9 18S rRNA metabarcodes by Bucklin et al. (2019) for these same samples; a seventh group (Peracarida) could not be analyzed due to many zero values for COI sequence numbers (**Supplementary Table S1**). Functional regression analysis (Ricker, 1973) of COI sequence numbers versus morphological counts was statistically significant for five of the six taxonomic groups; the exception was the Hydrozoa (**Figure 3**).

Interannual and regional patterns of diversity of the 6 taxonomic groups based on COI sequence numbers and morphological counts analyzed by NMDS (Jones, 2017) revealed similar patterns, with considerable overlap among the 3 EcoMon regions, but some evidence of distinctive samples in some regions and years for both molecular and morphological analysis (**Figure 4**). Based on Nonparametric (Permutation-based) MANOVA (Jones, 2017) analysis, the 6 groups showed different patterns of variation among regions and years ($p < 0.001$), but samples did not differ statistically significant among regions ($p = 0.772$). Cluster diagrams based on the Bray-Curtis dissimilarity coefficient (Bray and Curtis, 1957; McCune et al., 2002) based on both COI sequence numbers and morphological counts showed clear differentiation of MAB samples, with some overlap between GB and GoM (**Figure 5**). There are two groups of GB stations based on both COI sequence numbers and EcoMon counts (**Figure 5**), although the two GB clusters grouped together for EcoMon counts (**Figure 5B**), while GB #10-14 grouped with MAB and GB #15-19 grouped with GoM for COI sequence numbers (**Figure 5A**).

The Shannon Diversity Index (H; Pielou, 1977) showed lower levels for GoM during 2002-2005, based on both COI sequence numbers and morphological counts, with more variation among the 3 regions based on COI sequence numbers for 2007, 2009, 2010 and 2011 (**Figure 6**). The Simpson Index was also calculated from the same data, with results that were statistically nearly identical to the Shannon Index based on regression analysis (COI sequences: $r = 0.965$, $p = 2.138\ e^{-15}$; EcoMon counts: $r = 0.982$, $p = 2.636\ e\text{-}^{17}$).

COI Metabarcoding of Species Diversity

COI metabarcodes identified 181 species across 23 groups at varying taxonomic levels; 67 species of 15 groups had > 50 COI sequences; 23 species had >1,000 COI sequences (**Supplementary Table S3**). Classification and identification of species based on COI sequences used the North Atlantic regional MetaZooGene Atlas and Database (MZGdb), which includes 75,976 barcodes for 12,985 zooplankton species reported to occur in the region (https://metazoogene.org/mzgdb-natl, accessed March 13, 2022). Selection of species for the MZGdb is based upon the COPEPOD database (https://www.st.nmfs.noaa.gov/copepod/) and is designed to ensure an accurate, reliable, and taxonomically-complete reference sequence database with appropriate geographic coverage (Bucklin et al., 2021b).

Multivariate statistical analysis of metabarcoding results focused on 23 species with COI sequence numbers >1000 across all samples (**Table 2**). Two-dimensional NMDS results for the 23 species showed grouping of samples for each EcoMon region, with a distinct cluster of GoM samples and overlap between MAB and GB samples, with the notable exception of the 2010 MAB sample (**Figure 7**), which was dominated by the siphonophore, *Nanomia cara* (**Supplementary Table S3**). The 23 species showed some variation among regions and years ($p = 0.038$) based on Nonparametric (Permutation-based) MANOVA (Jones, 2017) analysis, but did not differ statistically significant among regions ($p = 0.359$). The Bray-Curtis similarity index cluster diagram based on COI sequence numbers for 23 species also showed clear regional patterns, except for the 2002 GB sample, which clustered with MAB (**Figure 8**).

The pattern of variation of the Shannon Diversity Index (H) based on COI sequence numbers for 23 species showed similar patterns of year-to-year variation in each of the 3 EcoMon regions, with marked differences between 2002-2005 versus 2007-2011 (**Figure 9**), similar to the interannual variation of H based on COI sequence numbers for the 6 groups (**Figure 6**).

Functional regression analysis of COI sequence numbers versus morphological counts from the EcoMon database were done for 7 species for which both types of data were available for most stations. Of these, 5 species showed significant regression relationships: *Calanus finmarchicus, Centropages hamatus, C. typicus, Pseudocalanus* spp., and *Temora longicornis* (**Figure 10**); regressions were not significant for two species (not shown):

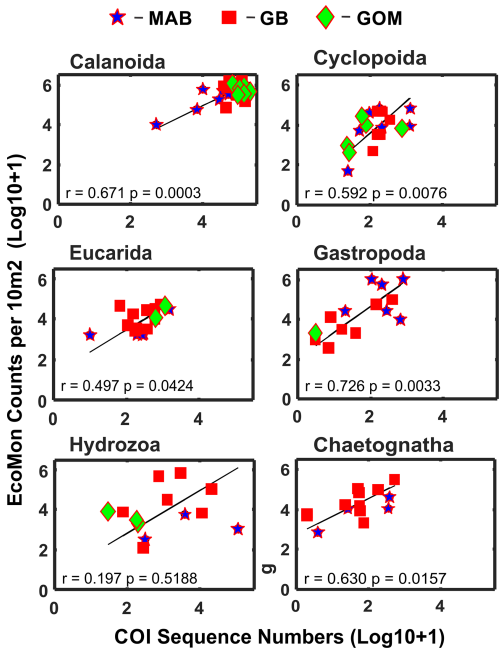


FIGURE 3 | Functional regression analysis of COI sequence numbers versus morphological microscopic counts per $10m^2$ for selected taxonomic groups of zooplankton. Symbols indicate regions for sample collections: Mid-Atlantic Bight (MAB), Georges Bank (GB), Gulf of Maine (GoM). Numbers are Log10+1 conversions; regression equation coefficient (r) and statistical significance (p) are indicated for each group.

Acartia longiremis (r = 0.607, p = 0.277) and *Nannocalanus minor* (r = 0.439, p = 0.711). In all cases, the species showing significant regressions had higher numbers of COI sequences, higher counts, and had data of both types for more stations; the species with insignificant regressions had more missing observations and recorded zeroes.

DISCUSSION

The NW Atlantic Continental Shelf

The NW Atlantic continental shelf was designated as a Large Marine Ecosystem (LME) based on the importance of the region for commercial harvesting and the need for conservation measures (Sherman et al., 2002). Despite many challenges over recent decades, including rapid warming from climate change (Friedland et al., 2020), the region has remained an important and productive region for commercial harvesting of numerous species. The importance of the pelagic community, and the zooplankton assemblage in particular, in ecosystem function and services has been acknowledged and examined for many decades (Sherman and Duda, 1999; Walsh et al., 2015; Friedland et al., 2019). Marked differences have been observed among the regions of the NW Atlantic continental shelf ecosystem (**Figure 1**) in temporal patterns of variation in ecosystem dynamics, including zooplankton diversity and biomass (O'Brien et al., 2013; **Figure 2**).

The importance of biodiversity in the functioning of marine ecosystems is well established (Gamfeldt et al., 2015). A number of studies have examined the more specific question of the role of species diversity of zooplankton in sustaining ecosystem services, including commercial fisheries (Byron and Link, 2010; Bi et al., 2014; Morse et al., 2017). Analysis of the taxonomic composition, diversity, abundance, and biomass of the zooplankton assemblage can serve as an early indicator of climate impacts and regime shifts in the region (Johnson et al., 2011; Borja, 2014; Stern et al., 2018).

A number of previous studies have established the importance of identifying and discriminating zooplankton species, even closely-related and morphologically-cryptic species, to allow understanding of ecosystem function and prediction of impacts of environmental variation and climate change (Johnson et al., 2011; Hare and Kane, 2012; O'Brien et al., 2013), and also to guide fisheries assessment and management (Kelly, 2016; Goodwin et al., 2017; Aylagas et al., 2018). The increasing evidence that COI metabarcoding can provide accurate and reliable species-level identification across the zooplankton assemblage is especially relevant and important for these applications (Andújar et al., 2018).

Integrative Molecular (Metabarcoding) and Morphological (Microscopic) Analysis

This study reports the results of comparative molecular (COI metabarcode sequence numbers) and mophological (EcoMon database records for microscopic counts) analysis of six taxonomic groups of marine zooplankton for which both types of data are available (**Supplementary Tables S1**, **S2**). Five of the 6 groups showed significant correlations between COI sequence

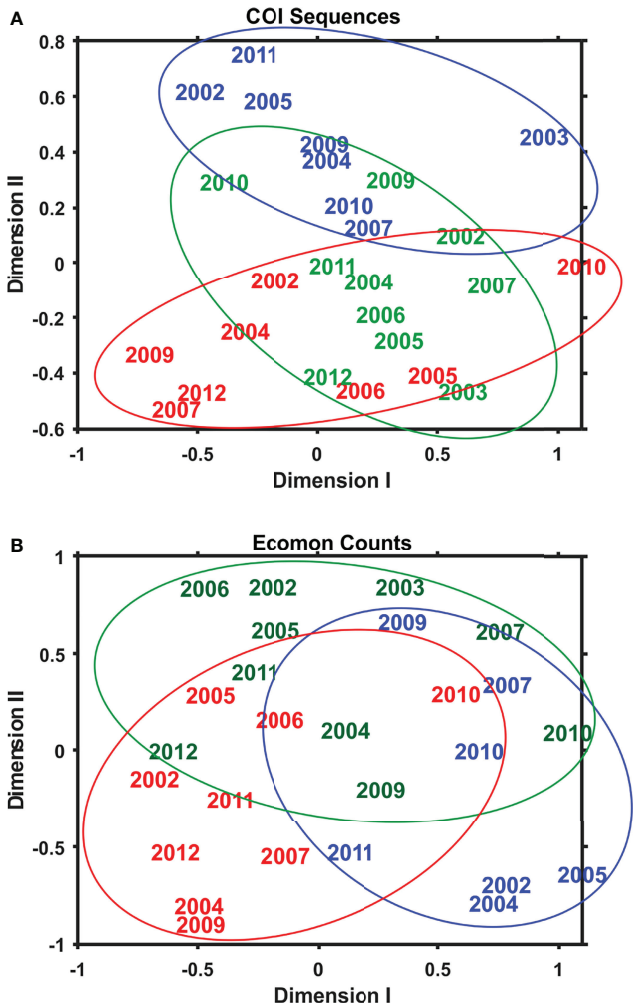


FIGURE 4 | Two-dimensional Nonmetric Multidimensional Scaling (NMDS) analysis of regional variation based on **(A)** COI sequence numbers and **(B)** EcoMon morphological counts for 6 taxonomic groups: Calanoida, Cyclopoida, Eucarida, Gastropoda, Hydrozoa, Chaetognatha. The plot indicates the year of sample collection; colors indicate regions: MAB (red), GB (green), GoM (blue).

numbers and morphological counts: Calanoida, Cyclopoida, Eucarida, Gastropoda, and Chaetognatha, but not Hydrozoa (**Figure 3**). These results provide further evidence of accurate quantitative measurements for some – but not all – taxonomic groups of zooplankton in some – but likely not all – circumstances.

The Shannon Diversity Index (H) based on COI metabarcoding revealed interannual variation for each region, including changes between 2002-2005 and 2007-2011 in the GoM for H index values for 6 taxonomic groups (**Figure 6**) and 23 species (**Figure 9**). The changes in some cases were consistent with variation of H index values based on EcoMon morphological counts (**Figure 6**) and with time-series records of total zooplankton displacement volume in the GoM (O'Brien et al., 2013; **Figure 2**). This finding provides further support for the potential value of COI metabarcoding for revealing and analyzing time-series variation of the zooplankton assemblage and monitoring of ocean ecosystems.

Another positive result, in terms of potential applications of metabarcoding for monitoring of pelagic ecosystems, is the power of COI metabarcoding for detection of species across a number of diverse taxonomic groups of the marine zooplankton assemblage. In this study, a total of 181 species across 23 taxonomic groups were identified based on comparison with the MetaZooGene Database (https://metazoogene.org/MZGdb; **Supplementary Table S3**). These numbers are similar to the numbers of species in EcoMon Survey records, which list 186 zooplankton species across 14 taxonomic groups of metazoan zooplankton detected in NEFSC records since 1977 (NMFS/NEFSC, 2019). However, there is marked lack of overlap in the species detected: only 53 species were found in common between the lists for COI sequences and EcoMon counts. A total of 24 species were detected by both metabarcoding and morphology, considering only species with more frequent observations (including 67 species of 15 groups

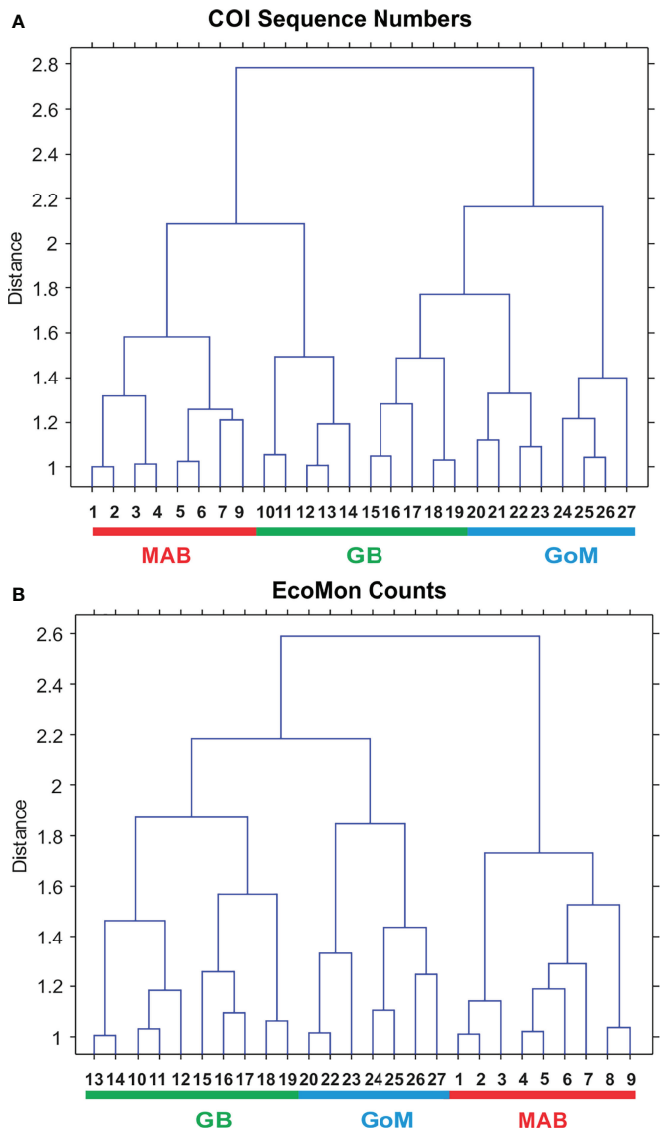


FIGURE 5 | Bray Curtis similarity cluster diagram showing analysis of regional variation based on **(A)** COI sequence numbers and **(B)** EcoMon microscopic counts for 6 taxonomic groups. Numbers are EcoMon sample numbers; see **Table 1** for collection information. The three regions are indicated as Mid-Atlantic Bight (MAB), Georges Bank (GB), and Gulf of Maine (GoM).

with > 50 COI sequences and 43 species of 9 groups with >1% frequency of occurrence in EcoMon Survey samples since 1977).

Additional analysis and intercomparison of results from metabarcoding and microscopic counts of the same samples are needed to evaluate and understand both the similarities in total numbers and limited overlap in species identified. One concern is the significant challenges of accurate identification of congeneric and closely-related species of zooplankton based on morphological characters. The discrimination and identification of morphologically cryptic species is a particular power of COI barcoding and metabarcoding (Bucklin et al., 2016; Leray and Knowlton, 2017). In this study, COI metabarcodes detected multiple species of several taxonomically-challenging copepod genera for which species are frequently over-looked or ignored – and rarely counted – in morphological taxonomic analysis, including EcoMon Survey data (NMFS/NEFSC, 2019). Four species of *Calanus* were detected: *C. finmarchicus* and *C. hyperboreus* predominated (**Table 2**); a few sequences were identified to *C. helgolandicus* in the 2012 GB sample and to *C. glacialis* in the 2002 GB and 2005 GoM samples. Seven species of *Clausocalanus* were detected: *C. pergens* was most abundant, but *C. furcatus, C. jobei, C. lividus, C. mastigophorus, C. parapergens*, and *C. paululus* were detected with small numbers of sequences in several stations; also noteworthy was the detection of four species of *Pseudocalanus*, including *P. acuspes* and *P. minutus*, as well as the cryptic species, *P. moultoni* and *P. newmani* (**Supplementary Table**

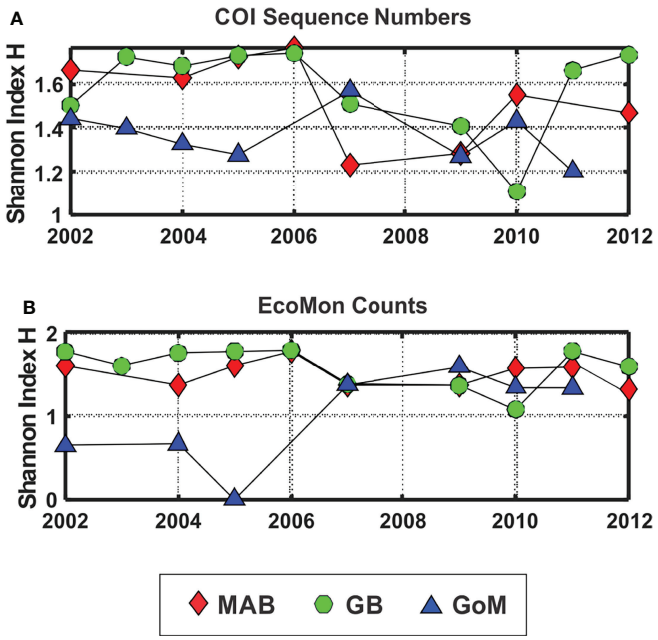


FIGURE 6 | Shannon Diversity Index (H; Pielou, 1977) based on **(A)** COI sequence numbers and **(B)** EcoMon morphological counts for 6 taxonomic groups. Legend indicates symbols for the three regions: Mid-Atlantic Bight (MAB), Georges Bank (GB), Gulf of Maine (GoM).

S3). COI sequence totals across all four *Pseudocalanus* species were significantly correlated to morphological counts for *Pseudocalanus* spp. in EcoMon records (**Figure 10**). Congeneric, sibling, and cryptic species of these copepod genera can be difficult to discriminate morphologically (Hill et al., 2001; Bucklin et al., 2003; Bucklin and Frost, 2009; Crouch et al., 2020), yet species-specific patterns of distribution and abundance are important indicators of seasonal-to-decadal patterns of environmental variation, climate change, and regime shifts (Johnson et al., 2011; Greene et al., 2013; Conversi et al., 2015; Morse et al., 2017).

Further evidence of the usefulness of COI metabarcoding for monitoring of ocean ecosystems is the finding of significant correlations between COI sequence numbers and EcoMon microscopic counts for abundant species for which both types

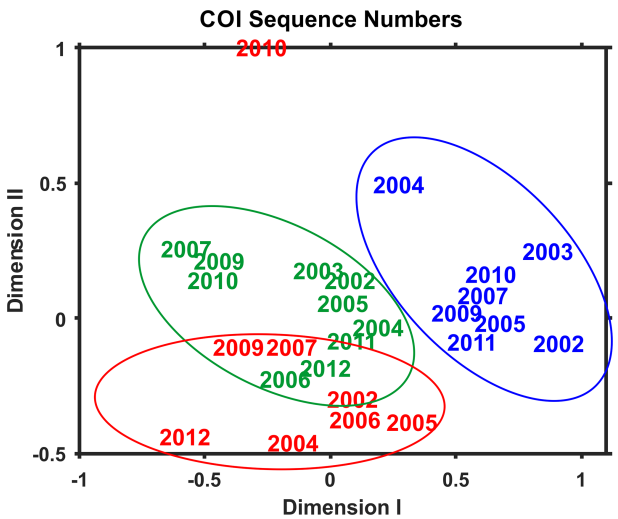


FIGURE 7 | Two-dimensional Nonmetric Multidimensional Scaling (NMDS) analysis of regional variation based on COI sequence numbers for 23 species with total sequence numbers >1000. See **Table 2** for list of species. The plot indicates the year of sample collection; colors indicate regions: MAB (red), GB (green), GoM (blue). Note that the MAB 2010 sample is not included within the circle defining the MAB region.

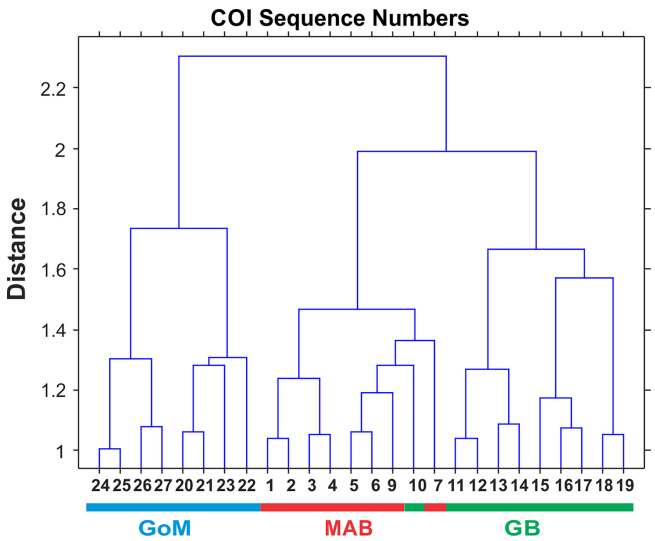


FIGURE 8 | Bray-Curtis similarity index cluster diagram for zooplankton samples collected throughout EcoMon regions based COI sequence numbers for 23 identified species with totals >1000 COI sequence numbers (see **Table 2**). Regions are indicated as Mid-Atlantic Bight (MAB), Georges Bank (GB), Gulf of Maine (GoM).

of data are available (**Figure 10**). The accuracy and reliability of metabarcoding for (semi)quantitative analysis, including abundance or biomass, of zooplankton has been evaluated in previous studies (Elbrecht and Leese, 2015; Thomas et al., 2016; Bucklin et al., 2019). A consistent finding has been that quantitative estimates are more accurate for highly abundant taxonomic groups or species (e.g., Matthews et al., 2021).

An acknowledged limitation of COI metabarcoding of zooplankton diversity is the uneven detection of species across the broad span of taxonomic groups in the pelagic assemblage (Deagle et al., 2014; Clarke et al., 2017; Hajibabaei et al., 2019). Various solutions have been proposed, including using multiple COI sub-regions, with specially-designed primers for target groups (Leray et al., 2013; Corell and Rodríguez-Ezpeleta, 2014; Elbrecht and Leese, 2017; Elbrecht et al., 2019) and integrative multi-region sequence analysis and bioinformatics (Antich et al., 2021; Creedy et al., 2021). Most importantly, classification and identification of species based on COI metabarcodes requires a taxonomically-complete and geographically-appropriate reference sequence database (Leray and Knowlton, 2017; Singh et al., 2021). Continued effort is needed to allow and ensure progress toward inclusion of COI barcode sequences for all zooplankton species, including sibling and cryptic species, recorded from regions throughout the global ocean (Bucklin et al., 2021b).

Comparative Assessment of Metabarcoding Using COI versus V9 18S rRNA

Marine zooplankton diversity and distribution have been examined using metabarcoding based on a number of different gene regions

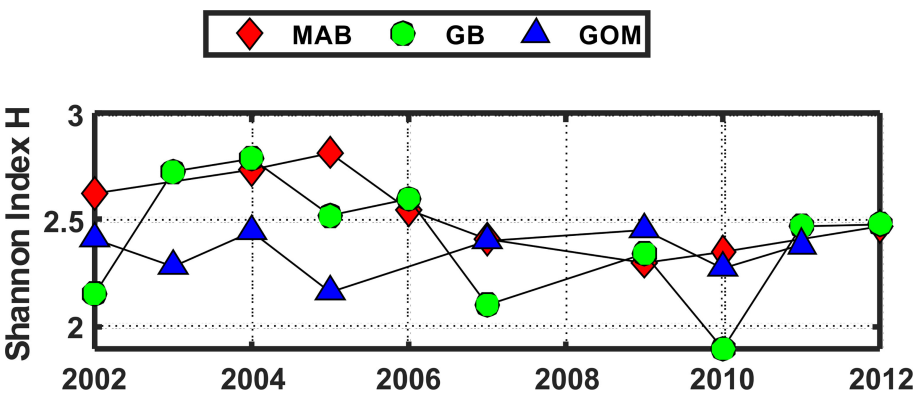


FIGURE 9 | Shannon Diversity Index (H) for the three regions based on COI sequence numbers for 23 abundant species with total sequence numbers >1000. See **Table 2** for list of species. Values of the Shannon Index (H) shown here were significantly correlated with Simpson Index values ($r = 0.953$, $p = 6.657\ e^{-14}$). Regions are: Mid-Atlantic Bight (MAB), Georges Bank (GB), Gulf of Maine (GoM).

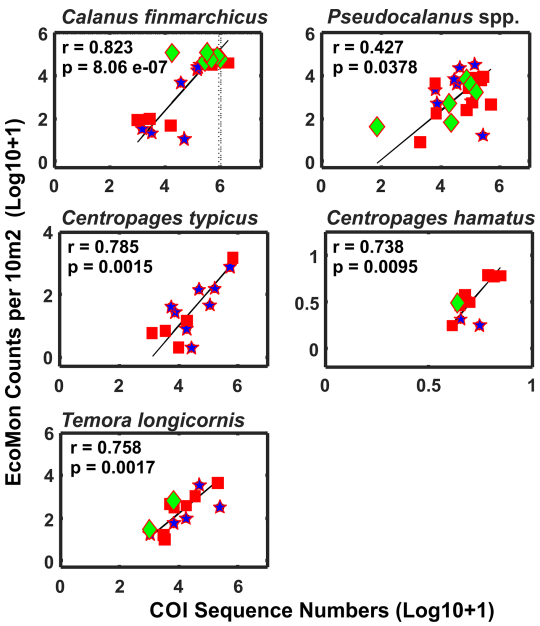


FIGURE 10 | Functional regression analysis of COI sequence numbers and EcoMon microscopic counts per 10 m^2 for abundant copepods that were detected by COI metabarcoding and also counted by EcoMon microscopic analysis (**Supplementary Tables S1**, **S2**). COI sequences for all *Pseudocalanus* species detected were added for comparison with *Pseudocalanus* spp. counts from the EcoMon database. Species were selected based on availability of molecular and morphological data for sufficient numbers of samples to allow regression analysis. Numbers are Log 10 + 1 conversions.

(Bucklin et al., 2016). A previous study (Bucklin et al., 2019) analyzed the V9 hypervariable region of 18S r RNA for the same set of samples from the NEFSC EcoMon Surveys, providing an opportunity to compare and contrast the results and conclusions based on the two marker gene regions. The 18S r RNA "tree of life" gene occurs in all living organisms on Earth; the gene sequence is both universal and highly conserved, ensuring detection across all major groups of metazoan zooplankton (Amaral-Zettler et al., 2009).

A total of 21 taxonomic groups of zooplankton, ranging from phylum to order, were detected and classified based on V9 metabarcoding of EcoMon samples (Bucklin et al., 2019; **Supplementary Table S1**). Statistical comparison of sequence numbers and abundance counts for these same groups revealed similar patterns of temporal (among years) and spatial (among regions) variation based on 27 samples collected during EcoMon Surveys from 2002-2012. Functional regression analysis for 7 taxonomic groups revealed positive correlations between V9 sequence numbers and abundance counts, with significant correlations ($p < 0.05$) for 3 groups: Calanoida, Chaetognatha, and Gastropoda. Comparison between the results from V9 and COI metabarcoding analysis of the same set of EcoMon samples demonstrates the power and accuracy of species-level identifications of marine zooplankton.

CONCLUSIONS

Zooplankton are key components of ocean ecosystems that provide early indicators of the impacts of seasonal-to-decadal patterns of environmental variation, including climate-driven regime shifts. Ecosystem monitoring and fisheries assessment programs provide invaluable time-series records of biodiversity of the zooplankton assemblage, based primarily on morphological taxonomic examination of plankton net samples. Analysis of these samples by DNA metabarcoding using a short region of the cytochrome oxidase I (COI) barcode gene allowed rapid and cost-effective characterization of biodiversity, including discrimination and identification of the 100s of species across numerous taxonomic groups that comprises the zooplankton assemblage. This study used COI metabarcoding of zooplankton diversity in samples collected from three regions (Mid-Atlantic Bight, Georges Bank, and Gulf of Maine) of the NW Atlantic continental shelf during 2002 – 2012 by the NOAA NMFS NEFSC Ecosystem Monitoring Program. Results show significant correlation between metabarcoding (COI sequence numbers) and morphological (microscopic counts) for abundant taxonomic groups and species, confirming the power of metabarcoding for (semi) quantitative measurements. COI metabarcoding also identified a number of sibling and cryptic species of copepods that were not discriminated in morphological counts. Limitations of the approach included failure to detect species within all taxonomic groups of zooplankton. Future development of COI metabarcoding for ecosystem monitoring of zooplankton diversity will require continued improvements in molecular protocols (e.g., COI primer design), completion of COI reference databases for species identification, and training of morphological taxonomic experts for marine zooplankton.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories of the NCBI GenBank SRA and are publicly available: BioProject PRJNA803165, Accession Numbers SAMN25654816–SAMN25654853.

AUTHOR CONTRIBUTIONS

Designed research, AB, PB-L, JQ, and PW. Performed research, PB-L, JQ, NC, and TO'B. Analyzed data, PB-L, JQ, PW, and DR. Wrote the paper, AB, DR, and PW. All authors agree to be accountable for the content of the work. All authors contributed to the article and approved the submitted version.

FUNDING

This publication resulted in part from support provided by the Scientific Committee on Oceanic Research (SCOR). Funds were also contributed by the U.S. National Science Foundation (Grant OCE-1840868) and by national SCOR committees.

ACKNOWLEDGMENTS

This study examined samples and data provided by the NOAA Northeast Fisheries Science Center, which oversees the Ecosystem Monitoring (EcoMon) Program of NW Atlantic continental shelf. We gratefully acknowledge the efforts of NEFSC managers, staff, and shipboard scientists and crew, with our special appreciation to Jon Hare and Jerry Prezioso (NOAA NEFSC, Narragansett, RI, USA). We thank Leocadio Blanco-Bercial (Bermuda Institute for Ocean Science, Bermuda) for providing advice and comments on the manuscript. Expert technical and bioinformatics advice and assistance was provided by professional staff at the Institute for Systems Genomics, University of Connecticut (Storrs, CT, USA), including Bo Reese and Lu Li at the UConn Center for Genome Innovation (CGI) and Vijender Singh and Jill L. Wegrzyn at the UConn Computational Biology Core (CBC).
This publication is a contribution from the Scientific Committee on Oceanic Research (SCOR) Working Group 157, MetaZooGene: Toward a new global view of marine zooplankton biodiversity based on DNA metabarcoding and reference DNA sequence databases (https://metazoogene.org/ and https://scor-int.org/group/157/). The authors acknowledge contributions from the Working Group on Integrated Morphological and Molecular Taxonomy (WGIMT) and the Working Group for Zooplankton Ecology (WGZE) of the International Council for the Exploration of the Sea (ICES).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.867893/full#supplementary-material

Supplementary Table 1 | COI sequence numbers for EcoMon samples for 17 taxonomic groups of marine zooplankton.

Supplementary Table 2 | EcoMon morphological (microscopic) counts for 17 taxonomic groups of marine zooplankton for samples collected at the same stations for which samples were analyzed by metabarcoding.

Supplementary Table 3 | COI sequence numbers for species detected in any sample analyzed by metabarcoding (see for collection information). Species are listed by taxonomic group; samples are organized by region. Asterisks (*) indicate species that were detected by morphological (microscopic) analysis of the same set of EcoMon Survey samples (except DE0305-38 and DE1105-127, for which counts are not available).

Supplementary Table 4 | COI Amplicon Sequence Variants (ASVs) for species detected by COI metabarcoding for which >50 sequences were observed for the samples analyzed. Species are listed by taxonomic group; samples are organized by region. Asterisks (*) indicate species that were detected by morphological (microscopic) analysis of the same set of EcoMon Survey samples (except DE0305-38 and DE1105-127, for which counts are not available).

REFERENCES

Amaral-Zettler, L. A., McCliment, P., Ducklow, H. W., and Huss, S. M. (2009). A Method for Studying Protistan Diversity Using Massively Parallel Sequencing of V9 Hypervariable Regions of Small-Subunit Ribosomal RNA Genes *PLoS One* 4, e6372. doi: 10.1371/journal.pone.0006372

Andujar, C., Arribas, P., Yu, D. W., Vogler, A. P., and Emerson, B. C. (2018). Why the COI Barcode Should be the Community DNA Metabarcode for the Metazoa? *Molec. Ecol.* 27, 3968–3975. doi: 10.1111/mec.14844

Antich, A., Palacin, C., Wangensteen, O. S., and Turon, X. (2021). To Denoise or to Cluster, That is Not the Question: Optimizing Pipelines for COI Metabarcoding and Metaphylogeography. *BMC Bioinf.* 22, 177. doi: 10.1186/s12859-021-04115-6

Aylagas, E., Borja, Á., Muxika, I., and Rodríguez-Ezpeleta, N. (2018). Adapting Metabarcoding-Based Benthic Biomonitoring Into Routine Marine Ecological Status Assessment Networks. *Ecol. Indic.* 95, 194–202. doi: 10.1016/j.ecolind.2018.07.044

Bi, H., Ji, R., Liu, H., Jo, Y. H., and Hare, J. A. (2014). Decadal Changes in Zooplankton of the Northeast U.S. Continental Shelf. *PloS One* 9, e87720. doi: 10.1371/journal.pone.0087720

Blanco-Bercial, L. (2020). Metabarcoding Analyses and Seasonality of the Zooplankton Community at BATS. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00173

Borja, A. (2014). Grand Challenges in Marine Ecosystems Ecology. *Front. Mar. Sci.* 1, 1. doi: 10.3389/fmars.2014.00001

Bourlat, S. J., Borja, A., Gilbert, J., Taylor, M. I., Davies, N., Weisberg, S. B., et al. (2013). Genomics in Marine Monitoring: New Opportunities for Assessing Marine Health Status. *Mar. Pollution Bull.* 74, 19–31. doi: 10.1016/j.marpolbul.2013.05.042

Brandão, M. C., Comtet, T., Pouline, P., Cailliau, C., Blanchet-Aurigny, A., Sourisseau, M., et al. (2021). Oceanographic Structure and Temporal Variations Contribute to High Heterogeneity in Mesozooplankton Over Small Spatial Scales. *ICES J. Mar. Sci.* 78, 3288–3302. doi: 10.1093/icesjms/fsab127

Bray, J. R., and Curtis, J. T. (1957). An Ordination of Upland Forest Communities of Southern Wisconsin. *Ecol. Monogr.* 27, 325–349. doi: 10.2307/1942268

Brown, E. A., Chain, F. J. J., Crease, T. J., MacIsaac, H. J., and Cristescu, M. E. (2015). Divergence Thresholds and Divergent Biodiversity Estimates: Can Metabarcoding Reliably Describe Zooplankton Communities? *Ecol. Evol.* 5, 2234–2251. doi: 10.1002/ece3.1485

Bucklin, A., and Frost, B. W. (2009). Morphological and Molecular Phylogenetic Analysis of Evolutionary Lineages Within *Clausocalanus* (Crustacea, Copepoda, Calanoida). *J. Crustacean Biol.* 29, 111–120. doi: 10.1651/07-2879.1

Bucklin, A., Frost, B. W., Bradford Grieve, J., Allen, L. D., and Copley, N. J. (2003). Molecular Systematic Assessment of Thirty-Four Calanoid Copepod Species of the Calanidae and Clausocalanidae Using DNA Sequences of mtCOI and Nuclear 18S rRNA. *Mar. Biol.* 142, 333–343. doi: 10.1007/s00227-002-0943-1

Bucklin, A., Lindeque, P. K., Rodriguez-Ezpeleta, N., Albaina, A., and Lehtiniemi, M. (2016). Metabarcoding of Marine Zooplankton: Progress, Prospects and Pitfalls. *J. Plankton Res.* 38, 393–400. doi: 10.1093/plankt/fbw023

Bucklin, A., Nishida, S., Schnack-Schiel, S., Wiebe, P. H., Lindsay, D., Machida, R. J., et al. (2010). "A Census of Zooplankton of the Global Ocean," in *Marine Life: Diversity, Distribution, and Abundance*. Ed. A. McIntyre (Oxford, UK: Wiley-Blackwell), 247–265.

Bucklin, A., Peijnenburg, K. T. C. A., Kosobokova, K. N., and Machida, R. J. (2021a). New Insights Into Biodiversity, Biogeography, Ecology, and Evolution of Marine Zooplankton Based on Molecular Approaches. *ICES J. Mar. Sci.* 78, 3281–3287. doi: 10.1093/icesjms/fsab198

Bucklin, A., Peijnenburg, K. T. C. A., Kosobokova, K. N., O'Brien, T. D., Blanco-Bercial, L., Cornils, A., et al. (2021b). Toward a Global Reference Database of COI Barcodes for Marine Zooplankton. *Mar. Biol.* 168, 78. doi: 10.1007/s00227-021-03887-y

Bucklin, A., Steinke, D., and Blanco-Bercial, L. (2011). DNA Barcoding of Marine Metazoa. *Ann. Rev. Mar. Sci.* 3, 471–508. doi: 10.1146/annurev-marine-120308-080950

Bucklin, A., Yeh, H. D., Questel, J. M., Richardson, D. E., Reese, B., Copley, N. J., et al. (2019). Time-Series Metabarcoding Analysis of Zooplankton Diversity of the NW Atlantic Continental Shelf. *ICES J. Mar. Sci.* 76, 1162–1176. doi: 10.1093/icesjms/fsz021

Byron, C. J., and Link, J. S. (2010). Stability in the Feeding Ecology of Four Demersal Fish Predators in the US Northeast Shelf Large Marine Ecosystem. *Mar. Ecol. Progr. Ser.* 406, 239–250. doi: 10.3354/meps08570

Clarke, L. J., Beard, J. M., Swadling, K. M., and Deagle, B. E. (2017). Effect of Marker Choice and Thermal Cycling Protocol on Zooplankton DNA Metabarcoding Studies. *Ecol. Evol.* 7, 873–883. doi: 10.1002/ece3.2667

Conversi, A., Dakos, V., Gardmark, A., Ling, S., Folke, C., Mumby, P. J., et al. (2015). A Holistic View of Marine Regime Shifts. *Phil. Trans. R. Soc B.* 370, 20130279. doi: 10.1098/rstb.2013.0279

Corell, J., and Rodríguez-Ezpeleta, N. (2014). Tuning of Protocols and Marker Selection to Evaluate the Diversity of Zooplankton Using Metabarcoding. *Rev. Invest. Mar. AZTI-Tec.* 21, 19–39.

Cox, J., and Wiebe, P. H. (1979). Origins of Oceanic Plankton in the Middle Atlantic Bight. *Est. Coast. Mar. Sci.* 9, 509–527. doi: 10.1016/0302-3524(79)90076-8

Creedy, T., Andujar, C., Meramveliotakis, E., Noguerales, V., Overcast, I., Papadopoulou, A., et al. (2021). Coming of Age for COI Metabarcoding of Whole Organism Community DNA: Towards Bioinformatic Harmonisation. *Authorea. Molec. Ecol. Res.* 22, 847–861. doi: 10.22541/au.162141276.61766048/v2

Crouch, K. E., Blanco-Bercial, L., Richardson, D. E., Copley, N. J., Wiebe, P. H., and Bucklin, A. (2020). Species-Specific Patterns of Distribution and Abundance of the Cryptic Copepods *Pseudocalanus Moultoni* and P. Newmani on Georges Bank (NW Atlantic Ocean) During Spring 1995–2012. *Cont. Shelf Res.* 208, 104242. doi: 10.1016/j.csr.2020.104242

Deagle, B. E., Clarke, L. J., Kitchener, J. A., Polanowski, A. M., and Davidson, A. T. (2017). Genetic Monitoring of Open Ocean Biodiversity: An Evaluation of DNA Metabarcoding for Processing Continuous Plankton Recorder Samples. *Molec. Ecol. Res.* 2017, 1–16. doi: 10.1111/1755-0998.12740

Deagle, B. E., Jarman, S. N., Coissac, E., Pompanon, F., and Taberlet, P. (2014). DNA Metabarcoding and the Cytochrome C Oxidase Subunit I Marker: Not a Perfect Match. *Biol. Lett.* 10, 20140562. doi: 10.1098/rsbl.2014.0562

Djurhuus, A., Pitz, K., Sawaya, N. A., Rojas-Márquez, J., Michaud, B., Montes, E., et al. (2018). Evaluation of Marine Zooplankton Community Structure Through Environmental DNA Metabarcoding. *Limnol. Oceanogr. Meth.* 16, 209–221. doi: 10.1002/lom3.10237

Edgar, R. C., and Flyvbjerg, H. (2015). Error Filtering, Pair Assembly and Error Correction for Next-Generation Sequencing Reads. *Bioinf* 31, 3476–3482. doi: 10.1093/bioinformatics/btv401

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., et al. (2019). Validation of COI Metabarcoding Primers for Terrestrial Arthropods. *PeerJ* 7, e7745. doi: 10.7717/peerj.7745

Elbrecht, V., and Leese, F. (2015). Can DNA-Based Ecosystem Assessments Quantify Species Abundance? Testing Primer Bias and Biomass—Sequence Relationships With an Innovative Metabarcoding Protocol. *PloS One* 10, e0130324. doi: 10.1371/journal.pone.0130324

Elbrecht, V., and Leese, F. (2017). Validation and Development of Freshwater Invertebrate Metabarcoding COI Primers for Environmental Impact Assessment. *Front. Envi. Sci.* 5, 11. doi: 10.7287/peerj.preprints.2044v4

Friedland, K. D., McManus, M. C., Morse, R. E., and Link, J. S. (2019). Event Scale and Persistent Drivers of Fish and Macroinvertebrate Distributions on the Northeast US Shelf. *ICES J. Mar. Sci.* 76, 1316–1334. doi: 10.1093/icesjms/fsy167

Friedland, K. D., Morse, R. E., Shackell, N., Tam, J. C., Morano, J. L., Moisan, J. R., et al. (2020). Changing Physical Conditions and Lower and Upper Trophic Level Responses on the US Northeast Shelf. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.567445

Gamfeldt, L., Lefcheck, J. S., Byrnes, J. E. K., Cardinale, B. J., Duffy, E., and Griffin, J. N. (2015). Marine Biodiversity and Ecosystem Functioning: What's Known and What's Next? *Oikos* 124, 252–265. doi: 10.1111/oik.01549

Geller, J. B., Meyer, C. P., Parker, M., and Hawk, H. (2013). Redesign of PCR Primers for Mitochondrial Cytochrome C. Oxidase Subunit I for Marine Invertebrates and Application in All-Taxa Biotic Surveys. *Molec. Ecol. Res.* 13, 851–861. doi: 10.1111/1755-0998.12138

Giebner, H., Langen, K., Bourlat, S. J., Kukowka, S., Mayer, C., Astrin, J. J., et al. (2020). Comparing Diversity Levels in Environmental Samples: DNA Sequence Capture and Metabarcoding Approaches Using 18S and COI Genes. *Molec. Ecol. Res.* 20, 1333–1345. doi: 10.1111/1755-0998.13201

Goodwin, K. D., Thompson, L. R., Duarte, B., Kahlke, T., Thompson, A. R., Marques, J. C., et al. (2017). DNA Sequencing as a Tool to Monitor Marine Ecological Status. *Front. Mar. Sci.* 4, 107. doi: 10.3389/fmars.2017.00107

Govindarajan, A. F., Francolini, R. D., Jech, J. M., Lavery, A. C., Llopiz, J. K., Wiebe, P. H., et al. (2021). Exploring the Use of Environmental DNA (eDNA) to Detect Animal Taxa in the Mesopelagic Zone. *Front. Ecol. Evol.* 9. doi: 10.3389/fevo.2021.574877

Greene, C. H., Meyer-Gutbrod, E., Monger, B. C., McGarry, L. P., Pershing, A. J., Belkin, I. M., et al. (2013). Remote Climate Forcing of Decadal-Scale Regime Shifts in Northwest Atlantic Shelf Ecosystems. *Limnol. Oceanogr.* 53, 803–816. doi: 10.4319/lo.2013.58.3.0803

Hajibabaei, M., Porter, T. M., Wright, M., and Rudar, J. (2019). COI Metabarcoding Primer Choice Affects Richness and Recovery of Indicator Taxa in Freshwater Systems. *PloS One* 14, e0220953. doi: 10.1371/journal.pone.0220953

Hare, J. A., and Kane, J. (2012). "Zooplankton of the Gulf of Maine—A Changing Perspective," in *Advancing an Ecosystem Approach in the Gulf of Maine*. Eds. R. Stephenson, J. Annala, J. Runge and M. Hall-Arber, 115–137. American Fisheries Society, Bethesda, MD, USA.

Head, E. J. H., and Sameoto, D. D. (2007). Inter-Decadal Variability in Zooplankton and Phytoplankton Abundance on the Newfoundland and Scotian Shelves. *Deep-Sea Res. II* 54, 2686–2701. doi: 10.1016/j.dsr2.2007.08.003

Hill, R. S., Allen, L. D., and Bucklin, A. (2001). Multiplexed Species-Specific PCR Protocol to Discriminate Four N. Atlantic *Calanus* Species, With a mtCOI Gene Tree for Ten *Calanus* Species. *Mar. Biol.* 139, 279–287. doi: 10.1007/s002270100548

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura,, et al. (2013). Reliable, Verifiable and Efficient Monitoring of Biodiversity *via* Metabarcoding. *Ecol. Lett.* 16, 1245–1257. doi: 10.1111/ele.12162

Johnson, C. L., Runge, J. A., Curtis, K. A., Durbin, E. G., Hare, J. A., Incze, L. S., et al. (2011). Biodiversity and Ecosystem Function in the Gulf of Maine: Pattern and Role of Zooplankton and Pelagic Nekton. *PloS One* 6, e16491. doi: 10.1371/journal.pone.0016491

Jones, D. L. (2017). *Fathom Toolbox for MATLAB: Software for Multivariate Ecological and Oceanographic Data Analysis* (St. Petersburg, FL, USA: College of Marine Science, University of South Florida). Available at: https://www.usf.edu/marine-science/research/matlab-resources/index.aspx/.

Kane, J. (2007). Zooplankton Abundance Trends on Georges Ban-2004. *ICES J. Mar. Sci.* 64, 909–919. doi: 10.1093/icesjms/fsm066

Kane, J. (2011). Inter-Decadal Variability of Zooplankton Abundance in the Middle Atlantic Bight. *J. NW Atl. Fish. Sci.* 43, 81–92. doi: 10.2960/J.v43.m674

Kelly, R. P. (2016). Making Environmental DNA Count. *Molec. Ecol. Res.* 16, 10–12. doi: 10.1111/1755-0998.12455

Kelly, T. B., Davison, P. C., Goericke, R., Landry, M. R., Ohman, M. D., and Stukel, M. R. (2019). The Importance of Mesozooplankton Diel Vertical Migration for Sustaining a Mesopelagic Food Web. *Front. Mar. Sci.* 6, 1–18. doi: 10.3389/fmars.2019.00508

Kelly, R. P., Port, J. A., Yamahara, K. M., Martone, R. G., Lowell, N., Thomsen, P. F., et al. (2014). Harnessing DNA to Improve Environmental Management. *Science* 344, 1455–1456. doi: 10.1126/science.1251156

Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., and Taylor, M. I. (2018). How Quantitative is Metabarcoding: A Meta-Analytical Approach. *Molec. Ecol.* 28, 420–430. doi: 10.1111/mec.14920

Leray, M., and Knowlton, N. (2017). Random Sampling Causes the Low Reproducibility of Rare Eukaryotic OTUs in Illumina COI Metabarcoding. *PeerJ* 5, e3006. doi: 10.7717/peerj.3006

Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., et al. (2013). A New Versatile Primer Set Targeting a Short Fragment of the Mitochondrial COI Region for Metabarcoding Metazoan Diversity: Application for Characterizing Coral Reef Fish Gut Contents. *Front. Zool.* 10, 1. doi: 10.1186/1742-9994-10-34

Matthews, S. A., Goetze, E., and Ohman, M. D. (2021). Recommendations for Interpreting Zooplankton Metabarcoding and Integrating Molecular Methods With Morphological Analyses. *ICES J. Mar. Sci.* 78, 3387–3396. doi: 10.1093/icesjms/fsab107

McCune, G., Grace, J. B., and Urban, D. L. (2002). *Analysis of Ecological Communities* (Oregon: Gleneden Beach: MjM Sofware Design).

Mohrbeck, I., Raupach, M. J., Martínez Arbizu, P., Knebelsberger, T., and Laakmann, S. (2015). High Throughput Sequencing - The Key to Rapid Biodiversity Assessment of Marine Metazoa? *PloS One* 10, e0140342. doi: 10.1371/journal.pone.0140342

Morse, R. E., Friedland, K. D., Tommasi, D., Stock, C., and Nye, J. (2017). Distinct Zooplankton Regime Shift Patterns Across Ecoregions of the U.S. Northeast Continental Shelf Large Marine Ecosystem. *J. Mar. Syst.* 165, 77–91. doi: 10.1016/j.jmarsys.2016.09.011

Northeast Fisheries Science Center, NMFS/NEFSC (2019) *Zooplankton and Ichthyoplankton Abundance and Distribution in the North Atlantic Collected by the Ecosystem Monitoring (EcoMon) Project From 1977-02-13 to 2017-11-16 (NCEI Accession 0187513)*. Available at: https://www.ncei.noaa.gov/archive/accession/0187513 (Accessed May 17, 2021).

O'Brien, T. D., Wiebe, P. H., and Falkenhaug, T. (2013). "ICES Zooplankton Status Report 2010/2011," in *ICES Coop. Res. Reports*, International Council for the Exploration of the Sea. Copenhagen, Denmark vol. 318, 1–208.

Pappalardo, P., Collins, A. G., Pagenkopp Lohan, K. M., Hanson, K. M., Truskey, S. B., Jaeckle, W., et al. (2021). The Role of Taxonomic Expertise in Interpretation of Metabarcoding Studies. *ICES J. Mar. Sci.* 78, 3397–3410. doi: 10.1093/icesjms/fsab082

Pershing, A. J., Greene, C. H., Jossi, J. W., O'Brien, L., Brodziak, J. K. T., and Bailey, B. A. (2005). Interdecadal Variability in the Gulf of Maine Zooplankton Community, With Potential Impacts on Fish Recruitment. *ICES J. Mar. Sci.* 62, 1511–1523. doi: 10.1016/j.icesjms.2005.04.025

Pielou, E. C. (1977). *Mathematical Ecology* (New York: Wiley).

Pitz, K. J., Guo, J., Johnson, S. B., Campbell, T. L., Zhang, H., Vrijenhoek, R. C., et al. (2020). Zooplankton Biogeographic Boundaries in the California Current System as Determined From Metabarcoding. *PloS One* 15, e0235159. doi: 10.1371/journal.pone.0235159

Porter, T. M., and Hajibabaei, M. (2018). Over 2.5 Million COI Sequences in GenBank and Growing. *PloS One* 13 (9), e0200177. doi: 10.1371/journal.pone.0200177

Questel, J. M., Hopcroft, R. R., DeHart, H. M., Kosobokova, K., Smoot, C., and Bucklin, A. (2021). Metabarcoding Analysis of Zooplankton Biodiversity of the Pacific-Arctic Chukchi Borderlands Region. *Mar. Biodiv.* 51, 4. doi: 10.1007/s12526-020-01136-x

Record, N. R., Pershing, A. J., and Jossi, J. W. (2010). Biodiversity as a Dynamic Variable in the Gulf of Maine Continuous Plankton Recorder Transect. *J. Plankton Res.* 32, 1675–1684. doi: 10.1093/plankt/fbq050

Richardson, D. E., Hare, J. A., Overholtz, W. J., and Johnson, D. L. (2010). Development of Long-Term Larval Indices for Atlantic Herring (*Clupea Harengus*) on the Northeast US Continental Shelf. *ICES J. Mar. Sci.* 67, 617–627. doi: 10.1093/icesjms/fsp276

Ricker, W. E. (1973). Linear Regressions in Fishery Research. *J. Fish. Res. Bd. Canada* 30, 409–434. doi: 10.1139/f73-072

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: A Versatile Open Source Tool for Metagenomics. *PeerJ* 4, e2584. doi: 10.7717/peerj.2584

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing Mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Envi. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-09

Schroeder, A., Pallavicini, A., Edomi, P., Pansera, M., and Camatti, E. (2021). Suitability of a Dual COI Marker for Marine Zooplankton DNA Metabarcoding. *Mar. Envi. Res.* 170, 105444. doi: 10.1016/j.marenvres.2021.105444

Sherman, K., and Duda, A. M. (1999). Large Marine Ecosystems: An Emerging Paradigm for Fishery Sustainability. *Fisheries* 24, 15–26. doi: 10.1577/1548-8446(1999)024<0015:LME>2.0.CO;2

Sherman, K., Kane, J., Murawski, S., Overholtz, W., and Solow, A. (2002). "The U.S. Northeast Shelf Large Marine Ecosystem: Zooplankton Trends in Fish Biomass Recovery," in *Large Marine Ecosystems*. Eds. K. Sherman and H. R. Skjoldal (Elsevier Publishers, Cambridge, MAssachusetts, USA), 195–215.

Singh, S. P., Groeneveld, J. C., Huggett, J., Naidoo, D., Cedras, R., and Willows-Munro, S. (2021). Metabarcoding of Marine Zooplankton in South Africa, African J. *Mar. Sci.* 43 (2), 147–159. doi: 10.2989/1814232X.2021.1919759

Stefanni, S., Stanković, D., Borme, D., de Olazabal, A., Juretić, A., Pallavicini, A., et al. (2018). Multi-Marker Metabarcoding Approach to Study Mesozooplankton at Basin Scale. *Sci. Rep.* 8, 12085. doi: 10.1038/s41598-018-30157-7

Steinke, D., deWaard, S. L., Sones, J. E., Ivanova, N. V., Prosser, S. W. J., Perez, K., et al. (2021). Message in a Bottle – Metabarcoding Enables Biodiversity Comparisons Across Ecoregions. *bioRxiv* 5, e59201. doi: 10.1101/2021.07.05.451165

Stern, R., Kraberg, A., Bresnan, E., Kooistra, W. H. C. F., Lovejoy, C., Montresor, M., et al. (2018). Molecular Analyses of Protists in Long-Term Observation Programmes - Current Status and Future Perspectives. *J. Plankton Res.* 40, 519–536. doi: 10.1093/plankt/fby035

Thomas, A. C., Deagle, B. E., Eveson, J. P., Harsch, C. H., and Trites, A. W. (2016). Quantitative DNA Metabarcoding: Improved Estimates of Species Proportional Biomass Using Correction Factors Derived From Control Material. *Molec. Ecol. Res.* 16, 714–726. doi: 10.1111/1755-0998.12490

Walsh, H. J., Richardson, D. E., Marancik, K. E., and Hare, J. A. (2015). Long-Term Changes in the Distributions of Larval and Adult Fish in the Northeast U.S. *Shelf Ecosystem. PloS One* 10, e0137382. doi: 10.1371/journal.pone.0137382

Wang, Q., Garrity, G. M., Tiedje, J. M., and Cole, J. R. (2007). Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences Into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* 73, 5261–5267. doi: 10.1128/AEM.00062-07

Wiebe, P. H., Rudels, B., Cadrin, S. X., Drinkwater, K. F., and Lavin, A. (2012). Introduction to Variability of the North Atlantic and its Marine Ecosystem–2009; Proceedings of an ICES/NAFO Symposium Held in Santander, Spain (10–12 May 2011). *ICES J. Mar. Sci.* 69, 697–702. doi: 10.1093/icesjms/fss090

Zhao, L., Zhang, X., Xu, M., Mao, Y., and Huang, Y. (2021). DNA Metabarcoding of Zooplankton Communities: Species Diversity and Seasonal Variation Revealed by 18S rRNA and COI. *PeerJ* 9, e11057. doi: 10.7717/peerj.11057

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

METHODS
published: 29 April 2022
doi: 10.3389/fmars.2022.869252

A Promising Approach to Quantifying Pteropod Eggs Using Image Analysis and Machine Learning

*Christine K. Weldrick**

Australian Antarctic Program Partnership, Institute for Marine & Antarctic Studies, University of Tasmania, Hobart, TAS, Australia

A newly developed protocol to semi-automate egg counting in Southern Ocean shelled (thecosome) pteropods using image analysis software and machine learning algorithms was developed and tested for accuracy. Preserved thecosome pteropod (*Limacina helicina antarctica*) egg masses collected from two austral summer research voyages in East Antarctica were digitally photographed to develop a streamlined approach to enumerate eggs within egg masses using Fiji/ImageJ and the associated machine learning plugin known as Trainable Weka Segmentation. Results from this semi-automated approach were then used to compare with manual egg counts from eggs dissected from egg masses under stereomicroscope. A statistically significant correlation was observed between manual and semi-automated approaches (R^2 = 0.92, $p < 0.05$). There was no significant difference between manual and automated protocols when egg counts were divided by the egg mass areas (mm^2) ($t(29.6) = 1.98$, $p = 0.06$). However, the average time to conduct semi-automated counts (M = 7.4, SD = 1.2) was significantly less than that for the manual enumeration technique (M = 35.9, SD = 5.7; $t(30) = 2.042$, $p < 0.05$). This new approach is promising and, unlike manual enumeration, could allow specimens to remain intact for use in live culturing experiments. Despite some limitations that are discussed, this user-friendly and simplistic protocol can provide the basis for further future development, including the addition of macro scripts to improve reproducibility and through the association with other imaging platforms to enhance interoperability. Furthermore, egg counting using this technique may lead to a relatively unexplored monitoring tool to better understand the responses of a species highly sensitive to multiple stressors connected to climate change.

Keywords: egg counting, egg masses, image analysis, machine learning, pteropods, Southern Ocean, thecosomes, zooplankton

OPEN ACCESS

Edited by:
Sophie G. Pitois, Centre for Environment, Fisheries and Aquaculture Science (CEFAS), United Kingdom

Reviewed by:
Clara Manno, British Antarctic Survey (BAS), United Kingdom
David Legland, Institut National de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), France

***Correspondence:**
Christine K. Weldrick
christine.weldrick@utas.edu.au

Specialty section:
This article was submitted to Marine Ecosystem Ecology, a section of the journal Frontiers in Marine Science

Received: *04 February 2022*
Accepted: *04 April 2022*
Published: *29 April 2022*

Citation:
Weldrick CK (2022) A Promising Approach to Quantifying Pteropod Eggs Using Image Analysis and Machine Learning. Front. Mar. Sci. 9:869252. doi: 10.3389/fmars.2022.869252

INTRODUCTION

It is now widely recognized that a multitude of concurrent biological, chemical and physical stressors caused by human activities are posing significant threats to global marine ecosystems and their components (IPCC, 2022). In polar regions, research has shown that changes to the development and reproduction of many marine organisms, including zooplankton, are

particularly vulnerable to warming and ocean acidification (Johnston et al., 2022). Some zooplanktonic groups, such as gastropod molluscs known as thecosome (shelled) pteropods, are regarded as early responders to climate change (Bednaršek et al., 2016), as they produce fragile, aragonite shells that are highly susceptible to dissolution linked to high CO_2 partial pressures (pCO_2) due to increasing ocean acidification (Riebesell et al., 2000; Orr et al., 2005; Kroeker et al., 2013). Whilst recent studies have shown a relatively higher capacity to withstand such effects than previously assumed (Peck et al., 2018), it is the early developmental stages of thecosome pteropods that are at greatest risk to changing ocean chemistry (Gardner et al., 2018). These risks will undoubtedly present wider consequences throughout marine ecosystems as, like many zooplankton taxa, thecosome pteropods provide a key energetic link between basal and higher trophic levels as well as an important contributor to the global export of carbon and carbonate to the deep sea through the fluxing of fast-sinking fecal pellets and shells post-mortem (Manno et al., 2010; Manno et al., 2018).

Shell dissolution in thecosome pteropods has often been studied using the common species *Limacina helicina* from the Northern Hemisphere (Comeau et al., 2010; Lischka et al., 2011; Comeau et al., 2012b; Bednaršek et al., 2014), and its Southern Ocean congener species, *Limacina helicina antarctica* (Manno et al., 2007; Seibel et al., 2012; Johnson and Hofmann, 2017; Gardner et al., 2018). Whilst one study by Bednaršek et al. (2012) revealed *in situ* shell dissolution of juvenile *L. h. antarctica* from the Scotia Sea, situated in the Atlantic sector of the Southern Ocean, much of the effects of climate change on early life development of thecosome pteropods have been observed through laboratory-based manipulation experiments. Examined under predicted levels of ocean acidification and warming, incubated thecosome pteropods have shown a range of adverse responses, including degradation, reduction in and/or lack of shell development (Lischka et al., 2011; Comeau et al., 2012a; Gardner et al., 2018), increased larval mortality (Lischka et al., 2011; Thabet et al., 2015; Gardner et al., 2018), and a decrease in the proportion of eggs developing to advanced embryogenetic stages (Manno et al., 2016). These responses are bound to have wider ecological and long-term ramifications related to population stability and recruitment.

Thecosome pteropods are holoplanktonic with unique life history strategies. Most species begin life as males until they reach a particular size (e.g., shell diameter of ~4 mm for *L. helicina*) then subsequently develop female organs and mature into females whilst their male organs are resorbed (Lalli and Wells, 1978; Lalli and Gilmer, 1989), which characterizes them as protandrous hermaphrodites. Females spawn tens of thousands of transparent eggs during their lifespan which are embedded into ribbons within gel matrix egg masses. Embryogenetic development occurs within these clutches, and hatching generally occurs at the trochophore larval stage (Lalli and Gilmer, 1989; Thabet et al., 2015; Wakabayashi, 2017). Optimal clutch size theory posits that mature females will spawn variable numbers of eggs to maximize the offspring fitness as it relates to resource availability, intraspecific competition, and mortality (Godfray et al., 1991). Different forms of parental care exist in marine gastropods, but for many species, females control the number of eggs contained within egg masses in an effort to manage their fecundity under changing conditions (Spight and Emlen, 1976; Perron, 1981).

Challenges related to estimating fecundity in thecosome pteropods can be attributed to the high number of microscopic eggs embedded within each egg mass. Manually counting them can be time consuming and using abundance of mature-aged adults is a relatively inaccurate alternative given the range of egg masses released by each pteropod adult. Manually counting thecosome eggs has previously involved dissecting the egg ribbons from the egg mass, which may introduce stress, particularly if eggs are being placed in live culture for subsequent observational studies (Manno et al., 2016). One study by Lalli and Wells (1978) used a conversion factor of 35 eggs mm^{-2} for *L. helicina* egg masses collected from Eastern Canada which is derived from estimating the number of eggs per area of egg mass measured, however this average value was based on complete measurements taken from only five egg masses. These challenges may be minimized with the use of image analyses platforms.

Autonomous image analysis techniques have previously been tested in plankton research involving the counting and measuring of round objects in aqueous solution, including the use of images and on-board, large-volume samples (Gorsky et al., 1989; Colas et al., 2018). Several studies have employed software platforms to automatically enumerate microscopic eggs of invertebrates from images with high degrees of success (Collin, 2010; Rosati et al., 2015; da Silva Júnior et al., 2018). The purpose of this study is to develop and validate a workflow that uses a combination of image segmentation and a supervised machine learning algorithmic approach to perform semi-automatic detection of thecosome pteropod eggs embedded within egg masses. This study aims to efficiently and accurately enumerate thecosome pteropod eggs embedded in their egg masses using the workflow developed in this study, and statistically compare this method to manual egg enumeration, which involves dissection under stereomicroscope. Reliably predicting the number of eggs within thecosome pteropod egg masses through non-destructive data imaging techniques can be beneficial to monitoring the health of marine ecosystems particularly prone to rapid chemical change.

MATERIALS AND WORKFLOW CONSTRUCTION

Study Area and Sampling

Plankton sampling was conducted along the East Atlantic region of the Southern Ocean during two separate research voyages (**Figure 1**). The first was aboard the RV *Aurora Australis* as part of the Kerguelen Axis (K-Axis) program (January-February 2016) within the southern extent of the Kerguelen Plateau. Sampling for K-Axis spanned a region from 62.7°E to 93.5°E,

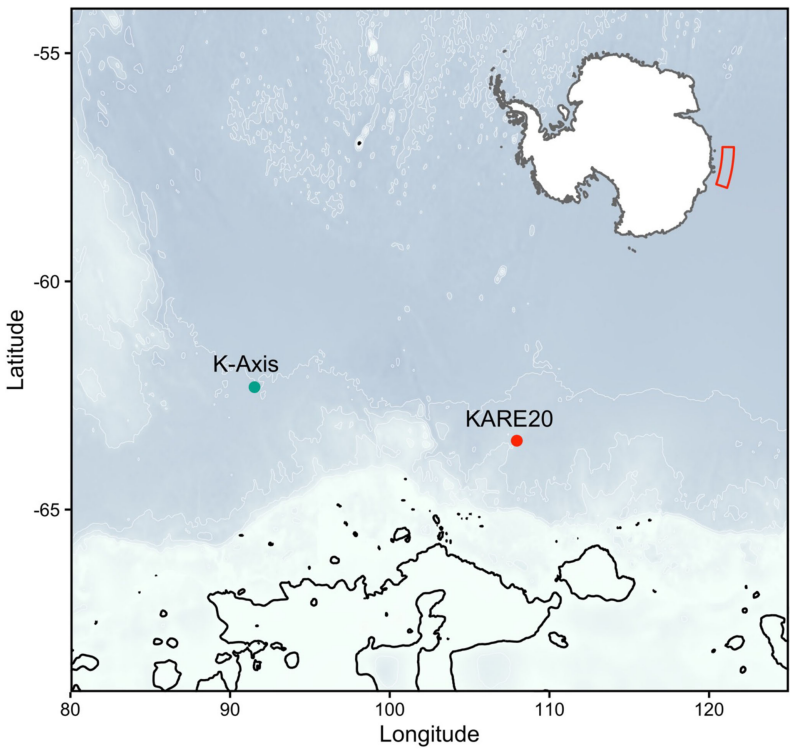


FIGURE 1 | Map of two sampling stations representing two separate research voyages aboard the TRV *Umitaka-maru* (KARE20) and the RV *Aurora Australis* (KAXIS), surveyed during the 2017 and 2016 austral summer, respectively.

and 57.6°S to 65.2°S. The second was aboard the TRV *Umitaka-maru* as part of the 20th Kaiyodai Antarctic Research Expedition (KARE20) program (January 2017) which covered a repeat transect southward along the 110°E longitudinal line.

Mesozooplankton samples from K-Axis were obtained using a Rectangular Midwater Trawl (RMT 1 + 8) net with a mouth area of 8 m^2 and a mesh size of 4.5 mm that tapered to a mesh size of 1.5 mm in the last 1.8 m of net [see Hosie et al. (2000) for more details]. Undamaged specimens collected with the RMT1 net, with a mesh size of 315 µm and a mouth area of 1 m^2, were measured for this study. Samples from KARE20 were obtained using an Ocean Research Institute (ORI) net with a mouth diameter of 160 cm and a mesh size of 500 µm [see Sakurai et al. (2018); Sakurai et al. (2020) for more details]. Both zooplankton collection methods sampled from a maximum depth of 200 m. All samples were preserved in 5% buffered formaldehyde and seawater solution and transported back to the Institute for Marine and Antarctic Studies in Hobart, Tasmania.

Pteropod egg masses (**Figure 2**) selected for this study were obtained from two sampling sites determined to have the highest number of intact egg masses, one from each voyage. The sampling site selected from K-Axis was located at -62.318°S and 91.531°E, and the site selected from KARE20 was located 453 nm away at -63.491°S and 107.958°E (**Figure 1**).

Manual Counting, Image Capturing, Pre-Processing, Calibration and Threshold Setting

A workflow for image pre-processing, segmenting images and enumerating eggs within thecosome pteropod egg masses is shown in **Figure 3**. Separated pteropod egg masses (n = 20) were rinsed in filtered seawater and transferred to glass petri dishes in preparation for imaging. Sharpened metal needles were used to gently remove any debris that may affect the segmentation process. Photographs of egg masses were taken with a Canon EOS Mark II 5D camera mounted on a Leica M165 C stereoscopic microscope and using EOS Utility software (Canon USA), while taking note of magnification. For converting measurements from pixels to µm, photographs were also taken of a micrometer slide at the same magnification used for the egg mass images. A selection of images (n = 16) was chosen to include all variations of typically encountered characteristics (e.g., eggs, matrix, phytoplankton cells), and the egg ribbons from the egg masses featured in these images were then carefully dissected under the microscope using a sharpened needle. The eggs from each ribbon were then enumerated to ground truth counts estimated from the automated technique image analysis.

Each digital image was opened in the Fiji/ImageJ software (RRID : SCR_002285) v. 2.3.1 (Schindelin et al., 2012) and a Wacom Intuos drawing tablet and pen (CTL-6100WL) was

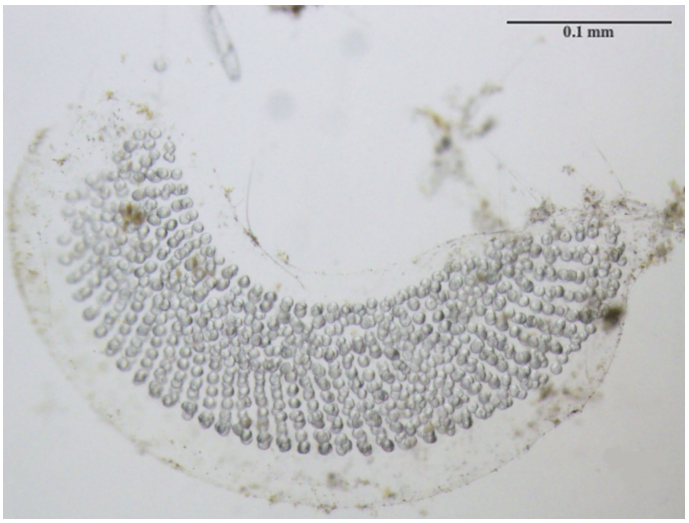


FIGURE 2 | Newly hatched egg mass consisting of hundreds of oval-shaped eggs arranged in a ribbon and embedded within an outer gelatinous matrix.

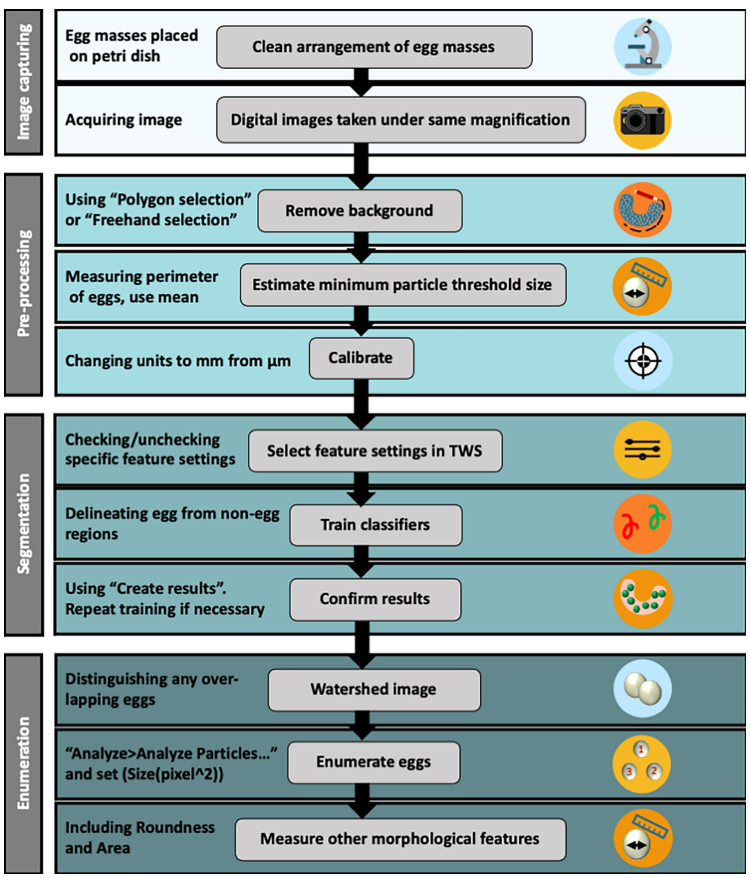


FIGURE 3 | Workflow procedures detailing steps beginning at image capturing, followed by data image pre-processing, segmentation and egg enumeration. See **Supplementary Figure 1** for a more detailed version of the workflow procedure.

used for accurate digital drawing on images. To exclude material around the egg masses, the "Polygon selection" or "Freehand selection" tools from the toolbar were used to draw an overlap around the perimeter of the egg masses (**Figure 4A**). The "Clear outside" function, located under the Edit drop-down menu, was selected and each image was then saved. To calibrate measurements for each egg mass image, the micrometer slide images were opened first, and the "Straight line" tool from the toolbar was superimposed over the micrometer slide ruler using the "Analyze>Set Scale…" function. A value in pixels was then linked to a known distance value of 1 mm from the line drawn over the slide ruler in the image (spatial calibration value = 1224.06 pixels/mm at 3.2x magnification). A threshold cell size needed to be set due to both the large concentration of non-egg material (e.g., phytoplankton cells) and egg overlapping within the egg masses. To estimate this, the areas (μm^2) of a random subset of 20 eggs from four egg mass images were individually measured using the "Polygon selections" tool to draw around the perimeter of each egg and determined to be an average of 10 μm^2 from the values in the results table using the "Analyze>Measure" function. Egg mass lengths and areas were also determined using Fiji/ImageJ, using the "Straight line" and "Freehand selection" tools, respectively.

Segmentation and Egg Enumeration

For Fiji/ImageJ to perform semi-automated enumeration, the eggs need to be differentiated from the backgrounds within each data image by a process known as segmentation. Built into Fiji/ImageJ is the Trainable Weka Segmentation (TWS; RRID: SCR_001214) plugin (http://imagej.net/Trainable_Weka_Segmentation) which is a tool that leans on machine learning and user-directed guidance to partition digital images into multiple segments, or classifiers, and subsequently perform automatic quantitative segmentation (Arganda-Carreras et al., 2017). Once both the image and TWS (version 3.3.2 was used in this study) are opened, classes were defined and renamed as "eggs" and "not eggs" (**Figure 4B**; **Supplementary Figure 1**). Training feature settings will generally depend on the quality of the image. For images taken for this study, Gaussian blur, Hessian, Membrane projections, Sobel filter, Difference of gaussians, Variance, and Structure were selected; Membrane thickness, Membrane patch size and Minimum sigma were kept at default settings, and Maximum sigma was changed to 32.0. The "Freehand" tool from the toolbar was used to mark the regions of each image under each class. For accuracy, a minimum of 10 marks per class was defined, before classifier training began. Training was repeated depending on the quality of the image. The image results from this segmentation workflow were saved and used for the final steps for counting eggs, detailed in the next section.

The binary result images generated from the TWS plugin were opened in Fiji/ImageJ (**Figure 4C**). Many eggs appeared fused due to the overlapping in regions and the watershed operator was applied to correct for this. The minimum size threshold was set to the previously determined 10 μm^2 under Analyze>Analyze Particles … and egg counts were determined from the display results window (**Figure 4D**). A stopwatch was used to time both manual and semi-automated counting techniques.

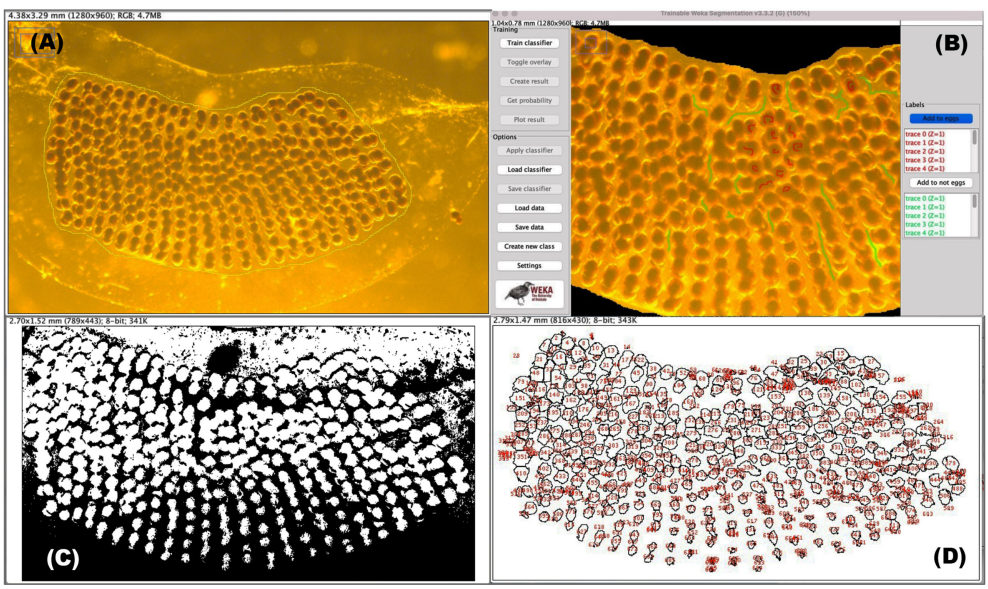


FIGURE 4 | Workflow steps leading to egg enumeration using supervised machine learning. **(A)** Selection of the area surrounding all eggs by using the "freehand" tool. **(B)** Supervised training in the TWS graphical user interface corresponding to eggs (default red) and non-egg (default green) regions. **(C)** Segmentation is first achieved by using a binary image (which is generated by the TWS plugin) prior to employing the watershed function. **(D)** Selecting "Analyse > Analyse Particles…" gives a results table (not shown) and this image window showing estimations of egg each numbered in red.

Statistical Analyses

All statistical analyses were performed using RStudio (RRID : SCR_000432) version 2021.09.0 (Team, R.C. 2014). A linear regression was performed from a correlation calculated between manual and semi-automated counts; the latter performed through Fiji/ImageJ. Two-sample *t*-tests were used to statistically compare manual with automated estimations of egg per mass area (mm^2) and duration of technique (minutes).

WORKFLOW ASSESSMENT RESULTS

Comparing Egg Enumeration Methods

Lengths and areas of pteropod egg masses measured digitally in this study varied between 2.9 mm and 12.2 mm and 4.1 mm^2 and 21.4 mm^2, respectively (**Table 1**). The averages calculated for the number of eggs per egg mass lengths were 48.4 eggs mm^{-2} (± 9.0 SD) for the manual enumeration technique and 41.7 eggs mm^{-2} (± 10.1 SD) for the semi-automated counting technique with Fiji/ImageJ.

The comparison of egg counting techniques showed a statistically significant correlation between manual and semi-automated egg counts ($R^2 = 0.92$, $p < 0.05$; **Figure 5**). In all but two of the totals, the manual counting technique produced higher egg counts as compared to those conducted digitally. When egg counts were divided by the egg mass areas (mm^2), these values were compared between the manual and semi-automated methods and there was no significant difference [$t(29.6) = 1.98$, $p = 0.06$; **Figure 5** inset plot]. The semi-automated counting technique averaged 7.4 minutes (± 1.2 SD) in duration, which took statistically significantly less time than manual egg counting, which averaged 35.9 minutes (± 5.7 SD) to complete ($t(30) = 2.042$, $p < 0.05$).

Limitations and Suggested Improvements

Before discussing the ecological implications associated with the egg counting outputs originating from the Fiji/ImageJ platform, it is critical to determine the reliability of these results. Many of the images depicted egg mass samples surrounded by non-egg particles, such as phytoplankton cells, that would likely also be counted by the software platform. The inability of the platform to distinguish eggs from other materials, identify egg and egg mass abnormalities or differentiate between eggs within close proximity are all limitations of this technique. Precision is enhanced through pre-analysis image preparation, involving setting size threshold limits, and drawing regions of interest (ROIs) encompassing high concentrations of intended materials. Despite these limitations, the statistically significant similarity obtained between manual and automated techniques validate the latter as a suitable solution for developing future studies that estimate fecundity.

Suggested improvements to these limitations should focus on the pre-processing of the egg mass samples prior to imaging. This could be through staining the sample with an agent that would enable the saturation value of eggs to be detected easily from non-egg materials, a step that is often used in medical imaging and histological studies. This was demonstrated by Malhan et al. (2018) who used various stains to distinguish, by color, various elements of connective tissue, including mineralized bone, cartilage, elastic fibers and muscles. Future studies are encouraged require the identification of constituents within and typically adjacent (e.g., phytoplankton, marine snow) to thecosome eggs and egg masses to select the appropriate stain used to separate these constituents by color or other identifier in preparation for pixel-based segmentation. Pre-process staining may eliminate the early workflow steps that focus on manually selecting ROIs and size threshold limits from images as well as decrease machine learning supervision while improving the overall speed and reproducibility of this methodology. Staining, though an extra step in the pre-processing stage, may effectively reduce the inclusion of background noise while also enabling fully automated batch processing of multiple images through scripts.

Other potential improvements to semi-automated egg enumeration involve advanced machine learning strategies, such

TABLE 1 | Length (mm), area (mm2) and count data of randomly selected pteropod egg masses from manual and semi-automatic techniques.

Sample image ID	Egg mass length (mm)	Egg mass area (mm^2)	Manual egg count	ImageJ egg count*	Manual count per area (eggs mm^{-2})	ImageJ count per area (eggs mm^{-2})
1	12.2	21.4	781 (43)	764 (8)	36.4	35.6
2	7.5	11.4	721 (44)	633 (9)	63.4	55.7
3	9.1	12.0	548 (39)	695 (7)	45.6	57.8
4	7.7	10.3	565 (41)	432 (8)	54.8	41.9
5	9.1	15.9	575 (42)	460 (9)	36.2	29.0
6	9.4	13.7	667 (44)	490 (7)	48.6	35.7
7	7.5	13.6	569 (37)	416 (9)	44.7	39.5
8	7.0	10.6	522 (32)	497 (7)	49.4	47.1
9	7.0	8.2	480 (30)	401 (7)	58.3	48.7
10	5.4	7.6	292 (29)	222 (9)	38.2	29.1
11	7.3	8.3	448 (32)	419 (7)	53.8	50.3
12	4.3	5.5	350 (31)	245 (5)	63.2	44.3
13	2.9	4.1	161 (27)	90 (8)	39.0	21.8
14	6.5	11.1	483 (35)	386 (7)	43.4	34.7
15	7.8	9.1	398 (32)	446 (6)	43.8	49.1
16	7.1	9.8	542 (36)	458 (6)	55.5	46.9

**Duration of technique (in brackets) starts from pre-processing of images to enumeration, but does not include steps necessary to conduct once, including calibration and scale setting. Approximate duration, in minutes, is recorded in brackets for each egg count and technique.*

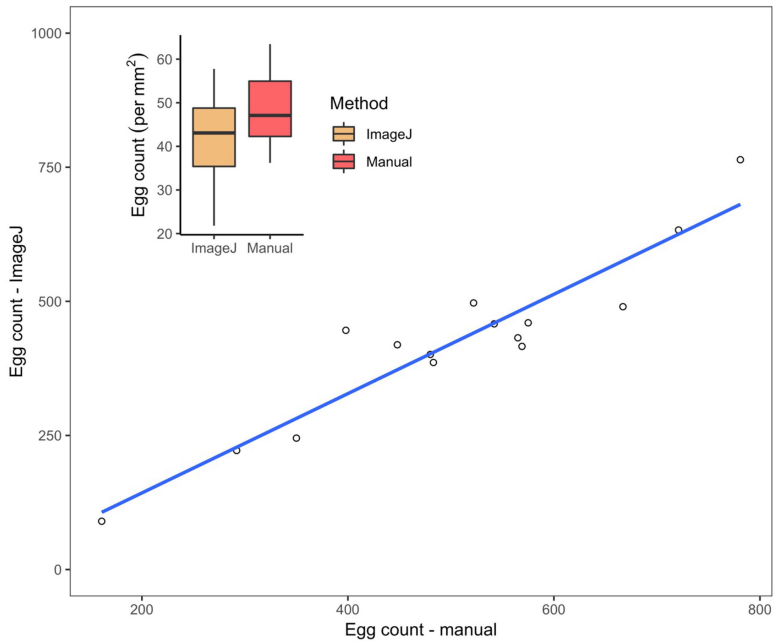


FIGURE 5 | Estimated counts of eggs within *L. h. antarctica* egg masses (*n*=16). Linear regression from the correlation calculated between manual and semi-automated counts obtained through Fiji/ImageJ is y = 0.9787x – 55.9883, where y is the predicted number of eggs estimated through automation, and x is manual count variable; R^2 = 0.9217, p <0.05. Dashed line is 1:1 reference. Inset plot: Results of *L. h. antarctica* egg counts mm^{-2} conducted by automation and manually. Median values of egg counts per egg mass area are depicted by horizontal lines within the 50% interquartiles (boxes). Upper and lower vertical lines, or "whiskers" refer to maximum and minimum dependent values, respectively. No significant difference was observed between methods, p > 0.05.

as deep learning-based methods with a particular focus on the segmentation of nucleus-like shapes and overlapping objects of interest. Deep learning methods have become increasingly popular in recent years, with many applications used in medical research (Hesamian et al., 2019). Emerging deep learning image analysis tools have been developed to address these issues, including the Fiji plugin StarDist, for cell and nuclei detection from images that detects star-convex shaped objects (Weigert et al., 2020), and Cellpose, which is a segmentation algorithm designed to efficiently segment cells stained for a variety of markers (Stringer et al., 2021). Future studies would benefit from testing deep learning algorithms in zooplankton research.

Anticipated Results Using Semi-Automated Egg Enumeration

Measuring the efficiency of a newly constructed, semi-automated egg enumeration technique is difficult when the eggs are microscopic in size, numerous, and the gel matrices in which they are embedded have other particles present, which creates noise, and consequently, potential for error. Therefore, using data retrieved from manual egg enumeration to compare techniques can enable an appropriate assessment of the effectiveness of the semi-automated technique described here. Pteropod egg counts determined accurately and efficiently can then be used to model drivers of both spatial and temporal patterns of early life development and fecundity throughout a rapidly changing Southern Ocean. Egg count data may thus lead to fruitful gains in assembling monitoring programs used to forecast how spatial, ecological and environmental cues affect variability in egg production of a sentinel species in response to increased ocean acidification, deoxygenation and temperatures (Bednaršek et al., 2016; Manno et al., 2017).

DISCUSSION

Shelled pteropods perform essential ecological roles in polar regions and serve as sentinels of climate change, though much work is yet to be done to better understand how these changes affect their early life development. This study constructed a framework to perform automatic, albeit supervised, enumeration of microscopic eggs from thecosome pteropod egg masses using image analysis and machine learning algorithms. Prior to this study, egg counting from thecosome pteropod egg masses had been performed manually under stereomicroscope either through counting eggs along ribbons dissected from each clutch or through estimating the number of eggs over a known area (mm^2) then extrapolating this value over the entire length of the egg mass. While the former is more accurate, this method is far more invasive and destructive, whereas the latter method does not account for high variability in egg density present along the length of the egg masses. The purpose of this study was to determine if a digital protocol for egg counting could be as accurate and efficient as the more invasive manual counting

method and results here reveal no statistically significant difference between methods. There was a strong correlation found between semi-automatic and manual counting methods.

There are very few published data focused on thecosome pteropod egg numbers and morphology, and how these attributes change over time and under future predicted climate scenarios, though spatio-temporal studies on manually enumerated eggs have been conducted in other marine gastropod species for decades (Berry, 1987; Mandal et al., 2010). However, very few studies have examined gastropod egg number variation along a gradient of environmental factors (Przeslawski, 2014). An accurate egg enumeration workflow has the potential to answer questions pertaining to early life responses of shelled pteropods to climate change, and the application of machine learning within these studies allows for the automation and simplified analyses of large-sized datasets. While only a few studies have counted thecosome eggs for various research purposes, at the time of this study, no other studies have closely analyzed different enumeration techniques for pteropod eggs nor developed an image processing technique incorporating supervised or unsupervised machine learning algorithms. This is the first study to develop and propose a framework to analyze thecosome pteropod eggs digitally using open-source image analysis software and machine learning algorithms.

Digital egg enumeration has advantages over manual counting. Namely, it does not impose damage and potential stress to the individual eggs, thus allowing the eggs and egg masses to be maintained in live cultures for further ontogenetic studies. Images can be captured while the live egg masses are placed in petri dishes or well slides under stereomicroscope, and subsequently available to use for ontogenetic experiments. This can facilitate more research into understanding uncertainties related to early life development of species sensitive to ocean acidification and ocean sea surface temperature change.

There is capacity for improving detection accuracy and speed of operation on the workflow presented here. Firstly, the discovery of a stain that would easily differentiate eggs from non-egg materials would be a fruitful next step. This would require a deeper understanding of the constituents that make up organic and inorganic materials within and adjacent to the egg masses and result in more accurate segmentation by the TWS plugin. Secondly, there are additional Fiji/ImageJ-based plugins and tools that have shown promising results in pre- and postprocessing cell enumeration, including cell staining followed by the in-built Fiji/ImageJ Color Deconvolution plugin for color segmentation (Ruifrok and Johnston, 2001). This research included a single observer, an assessment of variability in results between multiple observers would be recommended to test user-induced bias and standardize steps, beginning with settings (e.g., magnification) and equipment (e.g., microscope and camera make and model) related to the acquisition of digital images. And finally, through the development of new macros with customizable parameters (based on egg roundness, area, diameter, etc.) that would enable batch processing of multiple image files, rendering the process more automatic. The workflow described here can serve as a baseline for future development with new functionality.

In conclusion, the semi-automatic machine learning approach to analyzing pteropod egg mass images developed here is a promising user-friendly, non-destructive, and highly practical methodology for enumerating eggs within their gel matrices. This study outlines a simple, stepwise workflow necessary to accomplish accurate pixel-based segmentation of pteropod egg mass images using the image analysis software, Fiji/ImageJ, and the in-build TWS plugin. The effectiveness of this workflow was shown through a comparative analysis with manual counting requiring dissection of egg ribbons embedded within the egg mass gel matrix under a stereomicroscope that revealed high correlation.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

The author confirms all responsibility for study conception and design, sample and data collection, analyses and interpretation of results, and manuscript preparation.

FUNDING

This study was financially supported through a grant provided by the Holsworth Wildlife Research Endowment (grant #109804) of the Ecological Society of Australia, as well as funding from the Australian Government as part of the Antarctic Science Collaboration Initiative program (grant # ASCI000002). The Australian Antarctic Program Partnership is led by the University of Tasmania, and includes the Australian Antarctic Division, CSIRO Oceans and Atmosphere, Geoscience Australia, the Bureau of Meteorology, the Tasmanian State Government and Australia's Integrated Marine Observing System.

ACKNOWLEDGMENTS

This research was made possible through the assistance of those involved in plankton sampling, including the master, crew, scientists and technical support teams aboard both the RV *Aurora Australis* and TRV *Umitaka-maru*. I am grateful to the anonymous reviewers who provided constructive feedback that significantly improved the quality of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.869252/full#supplementary-material

REFERENCES

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K. W., Schindelin, J., Cardona, A., et al. (2017). Trainable Weka Segmentation: A Machine Learning Tool for Microscopy Pixel Classification. *Bioinformatics* 33, 2424–2426. doi: 10.1093/bioinformatics/btx180

Bednaršek, N., Feely, R. A., Reum, J. C. P., Peterson, B., Menkel, J., Alin, S. R., et al. (2014). Limacina Helicina Shell Dissolution as an Indicator of Declining Habitat Suitability Owing to Ocean Acidification in the California Current Ecosystem. *Proc. R. Soc B Biol. Sci.* 281, 20140123. doi: 10.1098/rspb.2014.0123

Bednaršek, N., Harvey, C. J., Kaplan, I. C., Feely, R. A., and Možina, J. (2016). Pteropods on the Edge: Cumulative Effects of Ocean Acidification, Warming, and Deoxygenation. *Prog. Oceanogr.* 145, 1–24. doi: 10.1016/j.pocean.2016.04.002

Bednaršek, N., Tarling, G. A., Bakker, D. C., Fielding, S., Cohen, A., Kuzirian, A., et al. (2012). Description and Quantification of Pteropod Shell Dissolution: A Sensitive Bioindicator of Ocean Acidification. *Glob. Change Biol.* 18, 2378–2388. doi: 10.1111/j.1365-2486.2012.02668.x

Berry, A. J. (1987). Reproductive Cycles, Egg Production and Recruitment in the Indo-Pacific Intertidal Gastropod Umbonium Vestiarium (L.). *Estuar. Coast. Shelf Sci.* 24, 711–723. doi: 10.1016/0272-7714(87)90109-0

Colas, F., Tardivel, M., Perchoc, J., Lunven, M., Forest, B., Guyader, G., et al. (2018). The ZooCAM, a New in-Flow Imaging System for Fast Onboard Counting, Sizing and Classification of Fish Eggs and Metazooplankton. *Prog. Oceanogr.* 166, 54–65. doi: 10.1016/j.pocean.2017.10.014

Collin, R. (2010). Repeatability of Egg Size in Two Marine Gastropods: Brood Order and Female Size do Not Contribute to Intraspecific Variation. *Mar. Ecol. Prog. Ser.* 410, 89–96. doi: 10.3354/meps08638

Comeau, S., Alliouane, S., and Gattuso, J. P. (2012a). Effects of Ocean Acidification on Overwintering Juvenile Arctic Pteropods *Limacina Helicina*. *Mar. Ecol. Prog. Ser.* 456, 279–284. doi: 10.3354/meps09696

Comeau, S., Gattuso, J.-P., Nisumaa, a.-M., and Orr, J. (2012b). Impact of Aragonite Saturation State Changes on Migratory Pteropods. *Proc. R. Soc B Biol. Sci.* 279, 732–738. doi: 10.1098/rspb.2011.0910

Comeau, S., Jeffree, R., Teyssie, J. L., and Gattuso, J. P. (2010). Response of the Arctic Pteropod *Limacina Helicina* to Projected Future Environmental Conditions. *PloS One* 5, 1–7. doi: 10.1371/journal.pone.0011362

da Silva Júnior, F. A., Monteiro da Silva, C., de Almeida, F. B., and Rodrigues-Silva, R. (2018). Digital Image Analysis to Estimate the Minimum Number of *Eurytrema Coelomaticum* Eggs in the Uterus of Adult Specimens. *Helminthologia* 55, 204–212. doi: 10.2478/helm-2018-0014

Gardner, J., Manno, C., Bakker, D. C. E., Peck, V. L., and Tarling, G. A. (2018). Southern Ocean Pteropods at Risk From Ocean Warming and Acidification. *Mar. Biol.* 165, 8. doi: 10.1007/s00227-017-3261-3

Godfray, H. C. J., Partridge, L., and Harvey, P. H. (1991). Clutch Size. *Annu. Rev. Ecol. Syst.* 22, 409–429. doi: 10.1146/annurev.es.22.110191.002205

Gorsky, G., Guilbert, P., and Valenta, E. (1989). The Autonomous Image Analyzer - Enumeration, Measurement and Identification of Marine Phytoplankton. *Mar. Ecol. Prog. Ser.* 58, 133–142. doi: 10.3354/meps058133

Hesamian, M. H., Jia, W., He, X., and Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *J. Digit. Imaging* 32, 582–596. doi: 10.1007/s10278-019-00227-x

Hosie, G. W., Schultz, M. B., Kitchener, J. A., Cochran, T. G., and Richards, K. (2000). Macrozooplankton Community Structure Off East Antarctica (80-150° E) During the Austral Summer of 1995/1996. *Deep Sea Res. Part II* 47, 2437–2463. doi: 10.1016/S0967-0645(00)00031-X

Hosie, G. W., Schultz, M. B., Kitchener, J. A., Cochran, T. G., and Richards, K. (2000). Macrozooplankton Community Structure Off East Antarctica (80-150° E) During the Austral Summer of 1995/1996. *Deep Sea Res. Part II* 47, 2437–2463. doi: 10.1016/S0967-0645(00)00031-X

IPCC. (2022) *Climate Change 2022: Impacts, Adaptation, and Vulnerability*. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Eds. H. -O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, et al. (Cambridge University Press). In Press.

Johnston, N. M., Murphy, E. J., Atkinson, A. A., Constable, A. J., Cotté, S., Cox, M., et al. (2022). Status, Change and Futures of Zooplankton in the Southern Ocean. *Front. Mar. Sci* 9, 624692. doi: 10.3389/fevo.2021.624692

Kroeker, K. J., Kordas, R. L., Crim, R., Hendriks, I. E., Ramajo, L., Singh, G. S., et al. (2013). Impacts of Ocean Acidification on Marine Organisms: Quantifying Sensitivities and Interaction With Warming. *Glob. Change Biol.* 19, 1884–1896. doi: 10.1111/gcb.12179

Lalli, C. M., and Gilmer, R. W. (1989). *Pelagic Snails: The Biology of Holoplanktonic Gastropod Mollusks* (Stanford, CA: Stanford University Press).

Lalli, M., and Wells, F. E. (1978). Reproduction in the Genus *Limacina* (Opisthobranchia: Thecosomata). *J. Zool. London* 186, 95–108. doi: 10.1111/j.1469-7998.1978.tb03359.x

Lischka, S., Budenbender, J., Boxhammer, T., and Riebesell, U. (2011). Impact of Ocean Acidification and Elevated Temperatures on Early Juveniles of the Polar Shelled Pteropod *Limacina Helicina*: Mortality, Shell Degradation, and Shell Growth. *Biogeosciences* 8, 919–932. doi: 10.5194/bg-8-919-2011

Malhan, D., Muelke, M., Rosch, S., Schaefer, A. B., Merboth, F., Weisweiler, D., et al. (2018). An Optimized Approach to Perform Bone Histomorphometry. *Front. Endocrinol. (Lausanne)* 9. doi: 10.3389/fendo.2018.00666

Mandal, S., Tamaki, A., Ohashi, S., Takeuchi, S., Agata, Y., Takahara, Y., et al. (2010). How Newly Recruited Cohorts are Formed in the Trochid Gastropod Population (Umbonium Moniliferum) on an Intertidal Sandflat in Western Kyushu, Japan. *J. Exp. Mar. Bio. Ecol.* 389, 18–37. doi: 10.1016/j.jembe.2010.04.001

Manno, C., Bednaršek, N., Tarling, G. A., Peck, V. L., Comeau, S., Adhikari, D., et al. (2017). Shelled Pteropods in Peril: Assessing Vulnerability in a High CO_2 Ocean. *Earth-Sci. Rev* 169, 132–145. doi: 10.1016/j.earscirev.2017.04.005

Manno, C., Giglio, F., Stowasser, G., Fielding, S., Enderlein, P., and Tarling, G. A. (2018). Threatened Species Drive the Strength of the Carbonate Pump in the Northern Scotia Sea. *Nat. Commun.* 9, 4592. doi: 10.1038/s41467-018-07088-y

Manno, C., Peck, V. L., and Tarling, G. A. (2016). Pteropod Eggs Released at High Pco_2 Lack Resilience to Ocean Acidification. *Nat. Publ. Gr.* 6, 25752. doi: 10.1038/srep25752

Manno, C., Sandrini, S., Tositti, L., and Accornero, A. (2007). First Stages of Degradation of *Limacina Helicina* Shells Observed Above the Aragonite Chemical Lysocline in Terra Nova Bay (Antarctica). *J. Mar. Syst.* 68, 91–102. doi: 10.1016/j.jmarsys.2006.11.002

Manno, C., Tirelli, V., Accornero, A., and Fonda Umani, S. (2010). Importance of the Contribution of Limacina Helicina Faecal Pellets to the Carbon Pump in Terra Nova Bay (Antarctica). *J. Plankton Res.* 32, 145–152. doi: 10.1093/plankt/fbp108

Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., et al. (2005). Anthropogenic Ocean Acidification Over the Twenty-First Century and its Impact on Calcifying Organisms. *Nature* 437, 681–686. doi: 10.1038/nature04095

Peck, V. L., Oakes, R. L., Harper, E. M., Manno, C., and Tarling, G. A. (2018). Pteropods Counter Mechanical Damage and Dissolution Through Extensive Shell Repair. *Nat. Commun.* 9, 264. doi: 10.1038/s41467-017-02692-w

Perron, F. E. (1981). The Partitioning of Reproductive Energy Between Ova and Protective Capsules in Marine Gastropods of the Genus Conus. *Am. Nat.* 118, 110–118. doi: 10.1086/283805

Przeslawski, R. (2014). A Review of the Effects of Environmental Stress on Embryonic Development Within Intertidal Gastropod Egg Masses. *Molluscan Res.* 24, 43–63. doi: 10.1071/MR04001

Riebesell, U., Zondervan, I., Rost, B., Tortell, P. D., Zeebe, R. E., and Morel, F. M. (2000). Reduced Calcification of Marine Plankton in Response to Increased Atmospheric Co2. *Nature* 407, 364–367. doi: 10.1038/35030078

Rosati, J. Y., Pacheco, V. A., Vankosky, M. A., and Vanlaerhoven, S. L. (2015). Estimating the Number of Eggs in Blow Fly (Diptera: Calliphoridae) Egg Masses Using Photographic Analysis. *J. Med. Entomol.* 52, 658–662. doi: 10.1093/jme/tjv053

Ruifrok, A., and Johnston, D. (2001). Quantification of Histochemical Staining by Color Deconvolution. *Anal. Quantif. Cytol. Histol. Histol.* 23, 291–299.

Sakurai, H., Moteki, M., Mizobata, K., Shimada, K., Matsuno, K., Okubo, Y., et al. (2018). Plankton Sampling by the Training Vessel Umitaka-Maru in the Indian Sector of the Indian Sector of the Southern Ocean in the Austral Summer of 2017. *JARE Data Rep.* 363, 1–13. doi: 10.15094/00014871

Sakurai, H., Moteki, M., Shimada, K., Tachibana, A., Yasuda, M., Tanaka, A., et al. (2020). Japanese Antarctic Research Expedition (JARE) Serial Data. Plankton Sampling by the Training Vessel Umitaka-Maru in the Indian Sector of the Southern Ocean in the Austral Summer of 2020. 366, 1–12. doi: 10.15094/00015897

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: An Open-Source Platform for Biological-Image Analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Seibel, B. A., Maas, A. E., and Dierssen, H. M. (2012). Energetic Plasticity Underlies a Variable Response to Ocean Acidification in the Pteropod, Limacina Helicina Antarctica. *PloS One* 7, e30464. doi: 10.1371/journal.pone.0030464

Spight, T. O. M. M., and Emlen, J. (1976). Clutch Sizes of Two Marine Snails With a Changing Food Supply. *Ecology* 57, 1162–1178. doi: 10.2307/1935042

Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: A Generalist Algorithm for Cellular Segmentation. *Nat. Methods* 18, 100–106. doi: 10.1038/s41592-020-01018-x

Team, R. C. (2014) *R: A Language and Environment for Statistical Computing*. Available at: http://www.r-project.org/.

Thabet, A. A., Maas, A. E., Lawson, G. L., and Tarrant, A. M. (2015). Life Cycle and Early Development of the Thecosomatous Pteropod *Limacina Retroversa* in the Gulf of Maine, Including the Effect of Elevated CO_2 Levels. *Mar. Biol.* 162, 2235–2249. doi: 10.1007/s00227-015-2754-1

Wakabayashi, K. (2017). Embryonic Development of the Sea Butterfly *Desmopterus Papilio* (Gastropoda: Thecosomata). *Invertebr. Reprod. Dev.* 61, 142–146. doi: 10.1080/07924259.2017.1311952

Weigert, M., Schmidt, U., Haase, R., Sugawara, K., and Myers, G. (2020). Star-Convex Polyhedra for 3D Object Detection and Segmentation in Microscopy. *Proc. - 2020 IEEE Winter Conf. Appl. Comput. Vision WACV 2020*, 3666–3673. doi: 10.1109/WACV45572.2020.9093435

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

ORIGINAL RESEARCH
published: 06 June 2022
doi: 10.3389/fmars.2022.870005

Content-Aware Segmentation of Objects Spanning a Large Size Range: Application to Plankton Images

Thelma Panaïotis[1], Louis Caray–Counil[1], Ben Woodward[2], Moritz S. Schmid[3], Dominic Daprano[4], Sheng Tse Tsai[4], Christopher M. Sullivan[4], Robert K. Cowen[3] and Jean-Olivier Irisson[1]*

[1] Laboratoire d'Océanographie de Villefranche, Sorbonne Université, Villefranche-sur-Mer, France, [2] CVision AI, Medford, MA, United States, [3] Hatfield Marine Science Center, Oregon State University, Newport, OR, United States, [4] Center for Quantitative and Life Science, Oregon State University, Corvallis, OR, United States

OPEN ACCESS

Edited by:
Mark C. Benfield, Louisiana State University, United States

Reviewed by:
Damianos Chatzievangelou, Self-employed, Greece
Francesco Pomati, Swiss Federal Institute of Aquatic Science and Technology, Switzerland

***Correspondence:**
Thelma Panaïotis
thelma.panaiotis@imev-mer.fr

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *05 February 2022*
Accepted: *04 May 2022*
Published: *06 June 2022*

Citation:
Panaïotis T, Caray–Counil L, Woodward B, Schmid MS, Daprano D, Tsai ST, Sullivan CM, Cowen RK and Irisson J-O (2022) Content-Aware Segmentation of Objects Spanning a Large Size Range: Application to Plankton Images. Front. Mar. Sci. 9:870005. doi: 10.3389/fmars.2022.870005

As the basis of oceanic food webs and a key component of the biological carbon pump, planktonic organisms play major roles in the oceans. Their study benefited from the development of *in situ* imaging instruments, which provide higher spatio-temporal resolution than previous tools. But these instruments collect huge quantities of images, the vast majority of which are of marine snow particles or imaging artifacts. Among them, the *In Situ* Ichthyoplankton Imaging System (ISIIS) samples the largest water volumes (> 100 L s^{-1}) and thus produces particularly large datasets. To extract manageable amounts of ecological information from *in situ* images, we propose to focus on planktonic organisms early in the data processing pipeline: at the segmentation stage. We compared three segmentation methods, particularly for smaller targets, in which plankton represents less than 1% of the objects: (i) a traditional thresholding over the background, (ii) an object detector based on maximally stable extremal regions (MSER), and (iii) a content-aware object detector, based on a Convolutional Neural Network (CNN). These methods were assessed on a subset of ISIIS data collected in the Mediterranean Sea, from which a ground truth dataset of > 3,000 manually delineated organisms is extracted. The naive thresholding method captured 97.3% of those but produced ~340,000 segments, 99.1% of which were therefore not plankton (i.e. recall = 97.3%, precision = 0.9%). Combining thresholding with a CNN missed a few more planktonic organisms (recall = 91.8%) but the number of segments decreased 18-fold (precision increased to 16.3%). The MSER detector produced four times fewer segments than thresholding (precision = 3.5%), missed more organisms (recall = 85.4%), but was considerably faster. Because naive thresholding produces ~525,000 objects from 1 minute of ISIIS deployment, the more advanced segmentation methods significantly improve ISIIS data handling and ease the subsequent taxonomic classification of segmented objects.

The cost in terms of recall is limited, particularly for the CNN object detector. These approaches are now standard in computer vision and could be applicable to other plankton imaging devices, the majority of which pose a data management problem.

Keywords: plankton images, ISIIS, image processing, image segmentation, object detection, convolutional neural network, computer vision

1 INTRODUCTION

1.1. Plankton Imaging Enables Fine Scale Studies

Planktonic organisms play crucial roles in the ocean: photosynthetic phytoplankton is responsible for about half of the primary production of the biosphere (Field et al., 1998) and is the basis of oceanic food webs (Falkowski, 2012); zooplankton acts as a trophic link between phytoplankton and higher trophic levels (Ware and Thomson, 2005; Frederiksen et al., 2006) and is a key component of the biological carbon pump, sequestering organic carbon at depth (Longhurst and Glen Harrison, 1989). Plankton comprises organisms from very diverse taxonomic groups (de Vargas et al., 2015) that span from micrometer scale picoplankton to meter-long Cnidarians (Lombard et al., 2019). Given this very wide size range, plankton sampling instruments cannot tackle all organisms at once and typically target a reduced size range instead (Lombard et al., 2019).

The power law underlying plankton or marine snow particle size spectra means that concentration drastically increases when size decreases: the relationship is linear in log-log form (Sheldon and Parsons, 1967; Sheldon et al., 1972; Stemmann and Boss, 2012; Lombard et al., 2019). The larger organisms, which each contribute significantly to biomass, are rare but easy to detect. Yet, it is critical to also focus on the smaller objects, to avoid artificially cutting the effective size range of any instrument, thus potentially discarding the most numerous objects in the sample (Lombard et al., 2019). Moreover, as marine snow particles cannot grow past a few centimeters because of disaggregation (Alldredge and Silver, 1988; Alldredge et al., 1990), the ratio of particles to plankton also decreases with increasing size. Therefore, while targeting small planktonic organisms is desirable, it comes with the difficulty of separating them from the largely dominant particles within the same size range.

While large scale plankton distribution patterns are resolved to a certain extent (Rutherford et al., 1999; Rombouts et al., 2009; Tittensor et al., 2010; Ibarbalz et al., 2019; Brandão et al., 2021), much remains to be discovered regarding fine scale distribution, in particular for zooplankton. For phytoplankton, submesoscale dynamics are known to influence their distribution and concentration: vertical currents may affect nutrient and cell distribution relative to the euphotic zone, thus affecting growth rate, horizontal currents can stir patches into filaments. These changes are expected to propagate to higher trophic levels (zooplankton, fish, etc.) (Lévy et al., 2018). Indeed, the trophic and reproductive interactions of zooplankton occur at the scale of organisms (µm to cm). Therefore, a local concentration of phytoplankton, in a thin layer for example, has more immediate consequences on the survival and development of zooplanktonic grazers than the average chlorophyll *a* concentration in the region. Thus, studying zooplankton distribution at fine scales, in relation with submesoscale dynamics, becomes relevant to understand the processes driving its distribution at regional scale.

Our lack of knowledge regarding the fine scale distribution of plankton partly stems from the difficulty to adequately sample it at such a small scale. Traditional plankton collection methods such as pumps, nets, and bottles typically integrate organisms over some vertical and/or horizontal distance and make it difficult to associate organism concentrations with their immediate environmental context (Remsen et al, 2004; Benfield et al., 2007; Lombard et al., 2019). Moreover, most damage fragile organisms and fail to sample some of them properly (Remsen et al., 2004).

As an alternative, *in situ* pelagic imaging instruments such as the Imaging FlowCytoBot (IFCB) (Olson and Sosik, 2007), the *In Situ* Ichthyoplankton Imaging System (ISIIS) (Cowen and Guigand, 2008), the Underwater Vision Profiler (UVP) (Picheral et al., 2010), and the Scripps Plankton Camera (SPC) (Orenstein et al., 2020) (see Lombard et al. (2019) for a detailed list) allow studying plankton distribution at all scales: from the fine ones they resolve in each sample to long time scales and global spatial coverage through the accumulation of individual samples (Stemmann et al., 2008; Forest et al., 2012; Robinson et al., 2021; Irisson et al., 2022). As a non-destructive sampling approach, these instruments allow investigating fragile planktonic objects, such as Rhizaria (Dennett et al., 2002; Biard et al., 2016; Biard and Ohman, 2020), Cnidaria and Ctenophora (Luo et al., 2014), or marine snow aggregates (Guidi et al., 2008; Guidi et al., 2015). Still, *in situ* imaging systems typically sample smaller volumes than plankton nets (Lombard et al., 2019), limiting their quantitative application to abundant taxa. To quantify rarer planktonic groups, sampling effort has to be increased to improve the chances of detection. For example, the ISIIS was initially developed with a very high sampling volume to study the very sparsely distributed fish larvae. Because of this, all *in situ* imaging instruments collect vast amounts of data, although the acquisition rate varies from one instrument to the next. ISIIS, for instance, collects up to 11 million objects per hour of sampling, while IFCB collects images at a rate of ~10,000 per hour (Sosik and Olson, 2007). Thus all these systems need efficient and automated data processing approaches, albeit with different stringency.

In addition, high resolution sampling is required to tackle questions that used to be out of reach, such as fine-scale plankton

distribution in relation with environmental conditions (McClatchie et al., 2012; Greer et al., 2015; Briseño-Avena et al., 2020), plankton patch structure (Robinson et al., 2021), interactions between zooplankton and phytoplankton fine layers (Greer et al., 2013; Greer et al., 2020a; Schmid and Fortiers, 2019) or co-occurrences revealing biological interactions such as predation (Greer et al., 2014; Schmid et al., 2020; Swieca et al., 2020; Greer et al., 2021).

1.2. Objects Need to be Extracted Automatically From Pelagic Images

The first data processing step is separating relevant organisms and particles from the background in raw images, i.e. image segmentation. Various segmentation methods have been applied for images collected by commonly used *in situ* imaging devices: the UVP relies on a fixed gray level threshold (Picheral et al., 2010), the IFCB uses an algorithm based on edge detection (Olson and Sosik, 2007), the SPC (Orenstein et al., 2020) runs a canny edge detector to initialize the segmentation of its dark-field microscopy images. To segment images generated by the Zooglider, a glider equipped with a shadowgraph, Ohman et al. (2019) also applied a canny edge detector. Finally, to segment shadowgrams from the ISIIS, Tsechpenakis et al. (2007) and Iyer, (2012) used statistical modeling of the background of the image and identified anomalies over this background as objects of interest.

The ISIIS is deployed in an undulating manner, between the surface and a given depth (Cowen and Guigand, 2008). It targets organisms in the range 250 µm - 10 cm. Together with grayscale images, it continually records environmental variables (temperature, salinity, fluorescence, dissolved oxygen and irradiance). The use of shadowgraphy combined with a specific lens and lighting system provide a large depth of field and allow a high sampling rate (28 kHz line scan camera). Therefore, the ISIIS is capable of sampling volumes of waters larger than all other *in situ* imaging instruments [> 100 L s^{-1}; Lombard et al. (2019)]. This optical design also ensures that the organism's size is not affected by its position within the depth of field. Shadowgraphs are also able to detect heterogeneities in the medium that is traversed by the light, which makes them excellent to image transparent organisms such as plankton, gelatinous organisms in particular. But it also makes them sensitive to other sources of heterogeneity, such as suspended particles or water density changes. ISIIS may thus generate noisy images when deployed in turbid waters (Luo et al., 2018; Greer et al., 2018) or across strong density gradients (**Figures 1D–F**) (Faillettaz et al., 2016). Furthermore, the use of a line scan camera means that marks or dust on the lens cause continuous streaks in the generated images (the line continuously scans the same speckle; **Figures 1A, D**). Those can be partially removed by applying a flat-fielding procedure, whereby the average gray value computed per row over a few thousand scanned lines is subtracted from the incoming new values (**Figures 1B, E**) (Faillettaz et al., 2016; Luo et al., 2018; Greer et al., 2018).

The very characteristics that give the ISIIS its qualities as a plankton imager (large sampling volume, high speed, ability to detect transparent objects) also mean that it creates a huge amount of images, the background of which is often non-uniform. This makes segmentation of planktonic objects from raw images far from trivial. To perform this segmentation, the processing pipeline was initially based on anomalies from a gaussian mixture model of the background gray levels without flat-fielding (Tsechpenakis et al., 2007) and later on k-harmonic means clustering on flat-fielded images (Iyer, 2012). This latter method was used in several studies (Luo et al., 2018; Greer et al., 2018; Schmid et al., 2020) and the full pipeline was open sourced in order to make plankton imaging more accessible and lower entry barriers (Schmid et al., 2021). Other studies relied on flat-fielding followed by segmentation above a fixed gray level (Faillettaz et al., 2016; Greer et al., 2020a; Greer et al., 2020b). However, most of these studies focused on the larger end of size range targeted by the ISIIS, by considering only objects above a given size threshold (**Table 1**), often because those were desirable targets, not noise. Similarly, for their canny edge detector applied to ZooGlider images, Ohman et al. (2019) considered objects larger than 100 pixels (Equivalent Spherical Diameter, or ESD of 0.45 mm). However, the algorithm failed when too many particles were present and had to fall back to a less sensitive (i.e. higher) gray threshold. As shown above, both planktonic organisms and particles are much more abundant towards the smaller end of the spectrum, meaning that such methods had to ignore a non-negligible part of planktonic organisms and marine snow in order to discard the background noise.

1.3. Marine Snow and Imaging Artifacts Dominate *In Situ* Images and Complicate Plankton Detection

Marine snow particles are much more abundant than plankton in the ocean (Lombard et al., 2019), which means that the vast majority (often > 85%) of images captured by *in situ* plankton imaging instruments are actually of various marine snow items (fecal pellets, large aggregates, small organism pieces, etc.; (Stemmann et al., 2000; Picheral et al., 2010; Stemmann and Boss, 2012)). Therefore, for plankton ecology studies, the bottleneck has often become the processing and filtering of collected images (Irisson et al., 2022). To reduce the proportion of detrital particles and focus on photosynthetic plankton, the IFCB and the FlowCam can use fluorescence image triggering, hence imaging only items that contain chlorophyll (Sieracki et al., 1998; Sosik and Olson, 2007). This is not possible over the large volumes and for the non-photosynthetic organisms that ISIIS or other zooplankton imagers target. Furthermore, density anomalies lead to the characteristically noisy shadowgrams presented above (**Figures 1D–F**), from which numerous artifactual "particles" are detected by the usual image processing pipelines. Those artifacts or noise, together with marine snow, can constitute 99% of the objects detected. Such an extreme class imbalance makes the automatic classification of these objects through machine learning a very arduous task (Lee et al., 2016).

Even for a trained human operator, the differentiation of some planktonic classes from the proteiform marine snow aggregates and noise, as well as distinction between marine

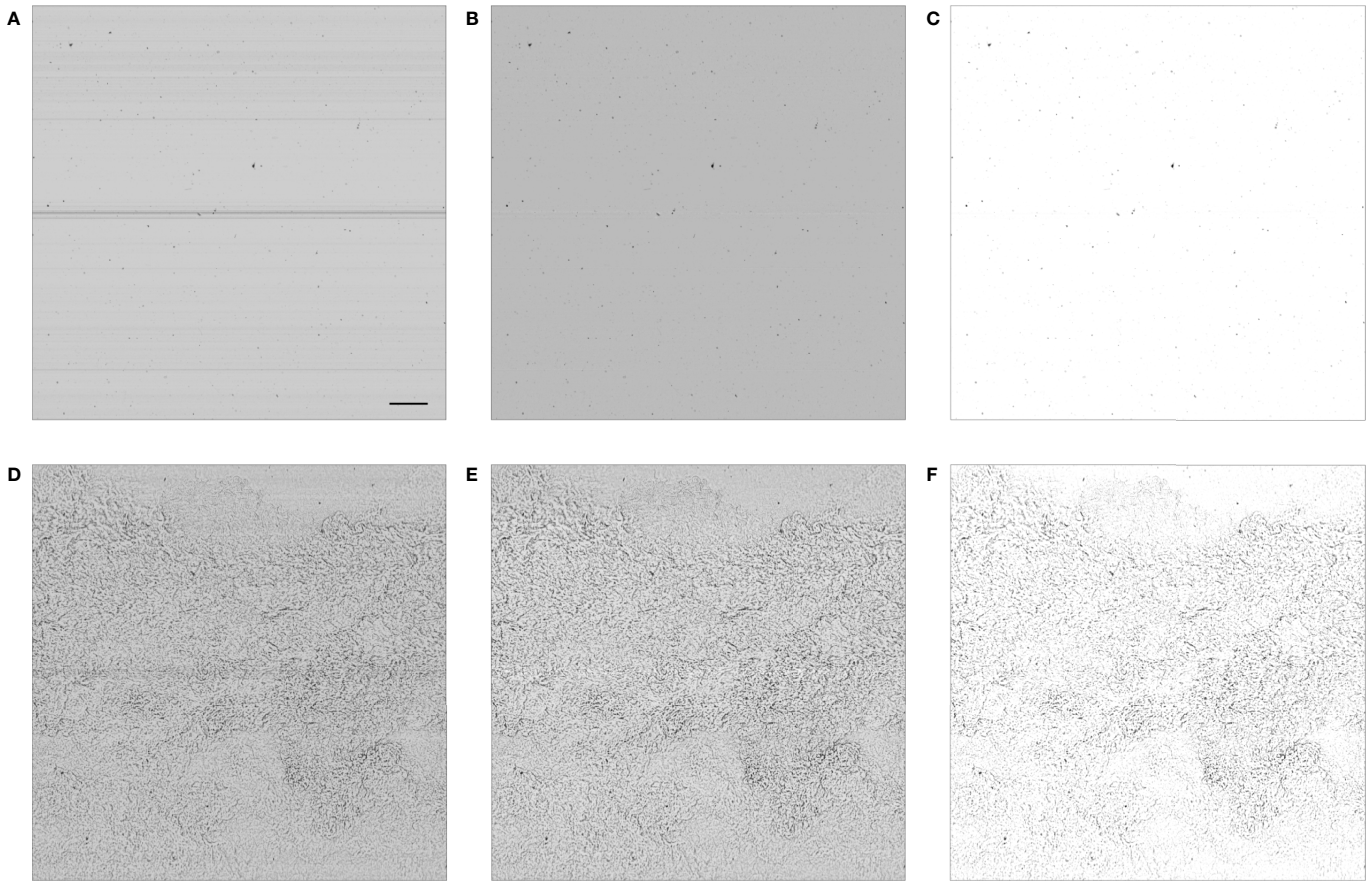


FIGURE 1 | ISIIS frames in clean waters **(A–C)** and across a density change **(D–F)**. The signature of this density change is similar to what a shadowgraph would image in air, above a burning candle. The panels are: **(A, D)** raw output; **(B, E)** after flat-fielding; **(C, F)** after contrasting. The camera scans vertically and the image is acquired from the right edge, as ISIIS moves through the water. In panel **(A)**, the scale bar represents 1 cm and is applicable to other panels.

TABLE 1 | Threshold in object area (number of pixels considered as part of the object) in studies exploiting ISIIS data.

Reference	Area threshold (px)	ESD (mm)
Schmid et al. (2020)	7	0.2
Luo et al. (2018)	50	0.53
Faillettaz et al. (2016)	250	0.92
Greet er al. (2020b)	400	0.95
Greer et al. (2020a)	900	1.4
Greer et al. (2021)	2000	3.0
Greer et al. (2018)	5000	5.4

[a]*The conversion factor from area (px) to Equivalent Spherical Diameter (ESD, mm) depends on the ISIIS configuration.*

snow and noise themselves, can be very challenging. Towards the smaller end of the size spectrum it becomes virtually impossible. Indeed, once these small objects are segmented out, the low pixel count combined with the lack of information regarding their context in the image makes their identification very difficult, for humans and computers alike (Parikh et al., 2012). Hence, one solution could be to focus solely on planktonic organisms from the segmentation step already and try to avoid segmenting non-planktonic objects, thanks to their broader context in the image, still accessible at this step. This should result in a much more manageable amount of data to classify and a lesser class imbalance. This approach requires the development of specific and "intelligent" segmentation methods that target specific objects only. The purpose of this work was (i) to develop such "intelligent" segmentation approaches and (ii) to compare them with classic methods to test whether they significantly improve the data processing pipeline. With this in mind, we benchmarked three segmentation methods against a ground-truth human segmentation using a dataset collected by the ISIIS in the North-Western Mediterranean Sea.

2 MATERIALS AND METHODS

2.1. Image Segmentation Methods

2.1.1. Threshold-Based Segmentation

The simplest segmentation method is to threshold pixels below a given gray level: adjoining pixels darker than the threshold are considered as segments. This threshold can be a value fixed *a priori* or dynamically computed from the properties of each image. For example, the classic method of Otsu (1979) is to examine the histogram of intensity levels and define the threshold so that it separates pixels into two relatively homogeneous intensity classes. Here either a fixed threshold was set or the threshold was defined based on a quantile of the histogram of gray levels. This quantile-based approach resulted in a darker segmentation threshold on noisy images, such as those captured around the strong density gradient induced by the thermocline (**Figures 1D–F**), which were richer in dark pixels. It was well adapted to limit the number of artifact segments generated from these images. Moreover, the first quartile is barely affected by the presence of relatively large dark objects such as jellyfish tentacles, making the segmentation threshold robust to these natural occurrences. After thresholding, segments defined by connected components were dilated by 3 pixels and eroded by 2 pixels to fill potential holes in transparent organisms and reconnect thin appendages to the organisms bodies. Finally, only segments larger than 50 pixels (400 μm in ESD) were retained, because it was the minimum size at which taxonomists could recognise organisms.

2.1.2. Threshold-MSER (T-MSER) Segmentation

This approach uses a signal-to-noise ratio (SNR) cutoff, calculated on images after flat-fielding, to determine whether the frame should be segmented using a Maximally Stable Extremal Region approach (MSER, Matas et al. (2004)), or if areas of high noise should first be filtered out using a naive thresholding approach before applying MSER. MSER was successfully applied to the segmentation of ZOOVIS imagery (Bi et al., 2015; Cheng et al., 2019). SNR can be used to determine the relative noise level in an image and was computed as

$$\mathrm{SNR} = 20 \times \log\left(\frac{S}{N}\right) \quad (1)$$

where S is the signal, defined as the mean of the input data, and N is the noise, computed as the standard deviation around that mean. Here, flat-fielded frames with low SNR (i.e. high noise) were binarized using a fixed thresholding in order to extract continuous regions of interest with darker pixel values. The regions identified in this way were then extracted using a mask and subsequently re-segmented using the MSER approach. MSER detects stable connected regions in images, which are areas that stay nearly unchanged over a wide range of grayscale thresholds. MSER can be tuned to allow for varying degrees of stable region area and the range of pixel gray values tested in the dynamic thresholding. High SNR frames are directly segmented using the MSER approach (**Figure 2** skip from step B to step D). Going from a pure MSER approach to the threshold+MSER (T-MSER) on low SNR (< 50) frames increased the recall on the test data from 65% to 85%, while also substantially increasing precision. This SNR and MSER method is written in C++17. The OpenCV and OpenMP Python packages were used for general computer vision and parallel processing for high processing efficiency, respectively.

2.1.3. Threshold-CNN (T-CNN) Segmentation

Another solution is to use Convolutional Neural Networks to either detect (i.e. define bounding boxes around) or segment (i.e. define a pixel mask of) objects of interest. Such approaches open the possibility to focus the detection on some types of objects (here, plankton) and ignore others (here, marine snow and artifacts); this is also called content-aware object detection or segmentation. However, CNNs tend to underperform at detecting objects across a large size range, especially for objects starting from a few dozen pixels (Cai et al., 2016). They work best when the target objects are of the same size as the receptive field of the model (Eggert et al., 2016). Thus, the development of detectors implementing receptive fields of various sizes constituted a major improvement, as they allowed detecting objects across a larger size range (Cai et al., 2016). In particular, we chose the Detectron2 library (Wu et al., 2019) developed by Facebook AI Research, which provides state-of-the-art object detection and segmentation algorithms, as well as pre-trained models for such tasks. Detectron2 includes a feature pyramid

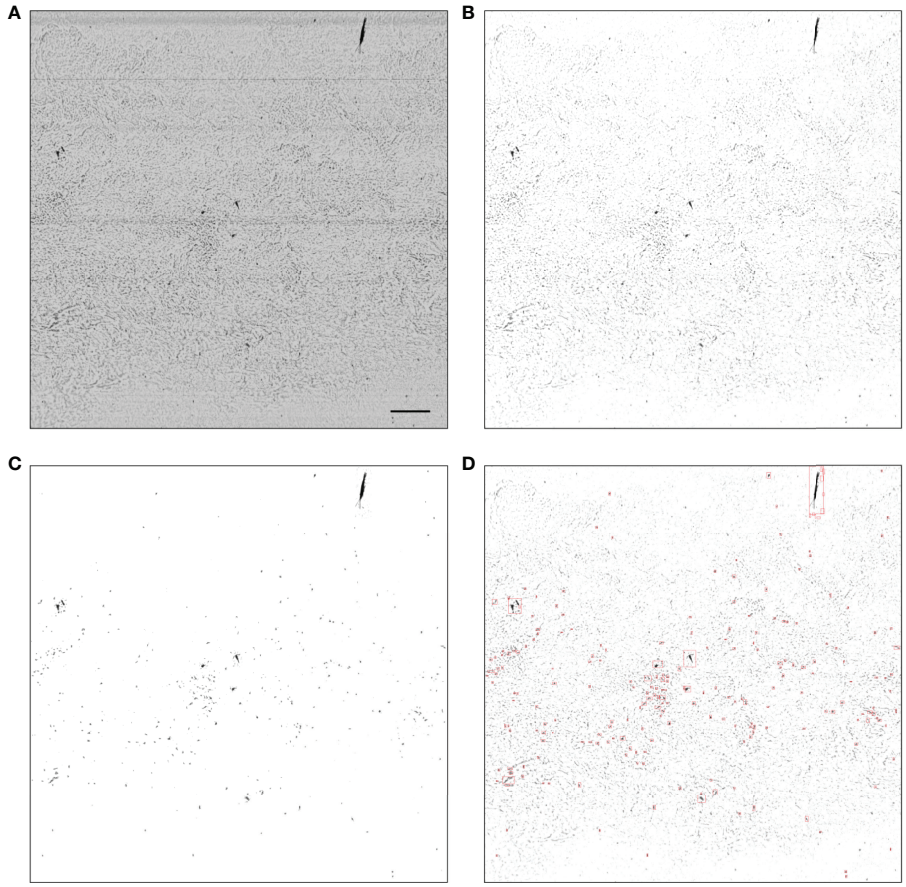


FIGURE 2 | Example MSER segmentation of a noisy raw frame (with low SNR). **(A)** Raw output; **(B)** after flat-fielding; (**C)** regions of interest created through naive thresholding; **(D)** regions of interest and their bounding boxes created by applying MSER to **(C)**. In a low SNR frame such as the one above the processing steps are **(A–D)**, while in a high SNR frame the processing steps are **(A, B, D)**. In panel **(A)**, the scale bar represents 1 cm and is applicable to other panels.

network (Lin et al., 2017) backbone that extracts feature maps across multiple scales to enable the detection of objects of various sizes, which was critical in our application to plankton images. Yet, this was not enough to cover the very large size range of organisms imaged by the ISIIS (from 50 to hundreds of thousands of pixels in area).

As explained above, marine snow particles and density-induced imaging artifacts are especially dominant compared to plankton in the smaller size classes. Therefore, our CNN pipeline was set up to segment the smaller objects, from 50 to 400 pixels in area, where the ability to specifically segment plankton makes the most difference. Above 400 pixels, the quantile-based threshold approach, with dilation and erosion, was used because it was simple and did not generate too many non-plankton segments.

In Detectron2, we used Mask R-CNN (He et al., 2017), which allows simultaneous bounding box detection and instance segmentation. The model was initialized with weights trained on the COCO reference dataset[1] but, for it to detect planktonic organisms on ISIIS images, it has to be fine-tuned on a dataset of ground truth bounding boxes and masks of such organisms. This dataset was generated by manually delineating all recognizable planktonic organisms in a set of ISIIS images, using a digital pen on a tablet computer. This produced 23,197 ground truth masks, from which bounding boxes were computed. Among those, 10,878 object were in the 50-400 pixels area range and usable. A 524×524 pixels crop was generated around every ground truth object (pushing the crop back inside the image when it crossed the edges). The choice of this particular size is a tradeoff between the maximum size of planktonic organisms that can be detected and the memory available on the graphics card. Moreover, it is in the line with common input sizes for segmentation models and was convenient to generate a tiling on ISIIS images. Several objects could be present in a crop. The crops were then split into 70% for training, 15% for validation, and 15% for testing. This split was stratified by the average gray level of the crop to ensure that both noisy (darker) and clean (lighter) images were present in each split, so that the model was presented with all kinds of images during training. Indeed, a model trained on clean images only would have performed poorly on noisy ones.

[1] https://github.com/facebookresearch/detectron2/blob/main/configs/COCO-InstanceSegmentation/mask_rcnn_R_50_FPN_3x.yaml/

Detectron2 can perform multiclass object detection or segmentation, meaning that objects are both detected/segmented and classified in a single step. However, it requires sufficient examples in each class for training. This condition could not be satisfied here, given how time-consuming it was to obtain pixel-level masks for every object and because plankton samples are usually dominated by a few abundant taxa while most others are very rare (Ser-Giacomi et al., 2018). Since the focus of this study is on segmentation, we decided to perform one-class object detection/segmentation, thus training the model to recognize planktonic organisms of any taxon. This implies that classification needs to be done after segmentation. Once an object is detected, this sequential, rather than concurrent, approach does not affect the result of the classification, since the same information is available to the subsequent classifier as to the concurrent one. Furthermore, focusing on segmentation only is also more comparable with the two other methods described above.

The model was trained for 30,000 iterations, and evaluation was run on the validation set every 1,000 iterations to ensure that the validation loss reached a plateau. The learning rate was set to 0.0005 initially and decreased 10 fold after 10,000 and 20,000 iterations. To increase the generality of the detector, data augmentation was used in the form of random resizing of the 524 pixels crops (to 640, 672, 704, 736, 768 or 800 pixels) and random horizontal flipping. The test set was used to assess theoretical performance after training and guide the choice of model settings; the actual performance was assessed on a separate, real-world dataset (presented below).

To apply the trained model to new images, a tiling of 524×524 pixels crops (the size used during model training) was generated over each input image, resulting in an overlap of 143 pixels vertically and 135 pixels horizontally. The overlap ensured that detectable objects spread over two crops were not missed. Crops were upscaled to 900×900 pixels to improve detection of small objects (Eggert et al., 2016). For each crop, the model predicted the bounding boxes of objects and their masks. We only considered the boxes, resolved overlaps in detections caused by overlapping crops, and submitted each box to exactly the same quantile-based thresholding as what was used above 400 pixels. This was preferred over using Detectron's mask proposals because their outline was not as detailed or replicable as the threshold-based ones. Furthermore, it also ensured that morphometric measurements performed on the masks (area in particular) were exactly comparable between the objects that went through the CNN and those above 400 pixels that were defined by simple thresholding. For each bounding box proposal, the model computes a confidence score. We retained all boxes with a score over 0.1, which is a quite low confidence threshold designed to increase the chance of detecting all objects of interest (i.e. favor recall) at the cost of some false positive detections (i.e. lower precision). Those false positives (i.e. segmented objects that are not plankton) will have the opportunity to be eliminated later, when segments are classified taxonomically.

The CNN was coded in Python with PyTorch, the original implementation library for Detectron2. Training was conducted on an Nvidia Quadro RTX 8000 GPU and the code is available at https://github.com/ThelmaPana/Detectron2_plankton_training. The combined CNN and threshold segmentation pipeline is implemented in https://github.com/jiho/apeep and this was run in several Linux-based environments, using various Nvidia GPUs.

2.2. Application to ISIIS Data from VISUFRONT Campaign

We evaluated these segmentation methods on ISIIS data from the VISUFRONT campaign, which sampled the Ligurian current front (North Western Mediterranean Sea), in the 0-100 m depth range, during summer 2013. Towed at a speed of 2 m s^{-1} (4kts) and set for a 28 kHz scanning rate, the ISIIS sampled 108 L per second. The 2048 pixels high continuous image strip created by the line scan camera moving in the water was cut in 2048×2048 pixels frames for storage. The ISIIS captured marked volutes caused by water density variations (**Figures 1D–F**), mostly driven by temperature changes around the thermocline, previously described by Faillettaz et al. (2016).

The continuous image strip was reassembled from the stored 2048×2048 pixels frames. Each line of pixels was flat-fielded by subtracting the row-wise average over a 8000 pixels moving window, hence removing streaks (**Figures 1A, B, D, E**). The cleaned image was cut into 10,240 pixels long images (5 frames, instead of 1) to reduce the probability of cutting objects across images while keeping the memory footprint of each image manageable. Finally, the image was contrasted by stretching the intensity range between percentiles 0 and 40 (**Figures 1B, C, E, F**). These values were chosen by iteration, through discussions with the taxonomist in charge of delineating planktonic organisms from raw images, as to achieve the highest distinguishability for those.

A ground truth dataset was generated by manually delineating all planktonic organisms (using a digital pen and tablet) in 106 10,240×2048 pixels images, regularly spread across a full transect, hence representative of different environments. This resulted in 3,356 objects that were later taxonomically sorted into 24 taxa (**Figure 3**), in the Ecotaxa web application (Picheral et al., 2017). This dataset was completely independent from the one that was used to train, validate and test the Detectron2 model. Some images were checked by two independent operators to check their consistency; when this was done, no differences were found.

Segments from each of the three automated methods were matched with ground truth segments of the same image. A bounding box intersection over union (IoU) score higher than 10% was considered as a match between segments. This threshold was set after manually inspecting a set of potential matches with various IoU values and was found to be the best value to discriminate between true and false matches. In case a ground truth segment matched multiple automatic segments, only one match was retained, to avoid inflating artificially the number of matches from the automated pipelines. In case an automatic segment matched multiple ground truth segments, the match was not counted either because it corresponded to a large segment that encompassed several organisms likely belonging to

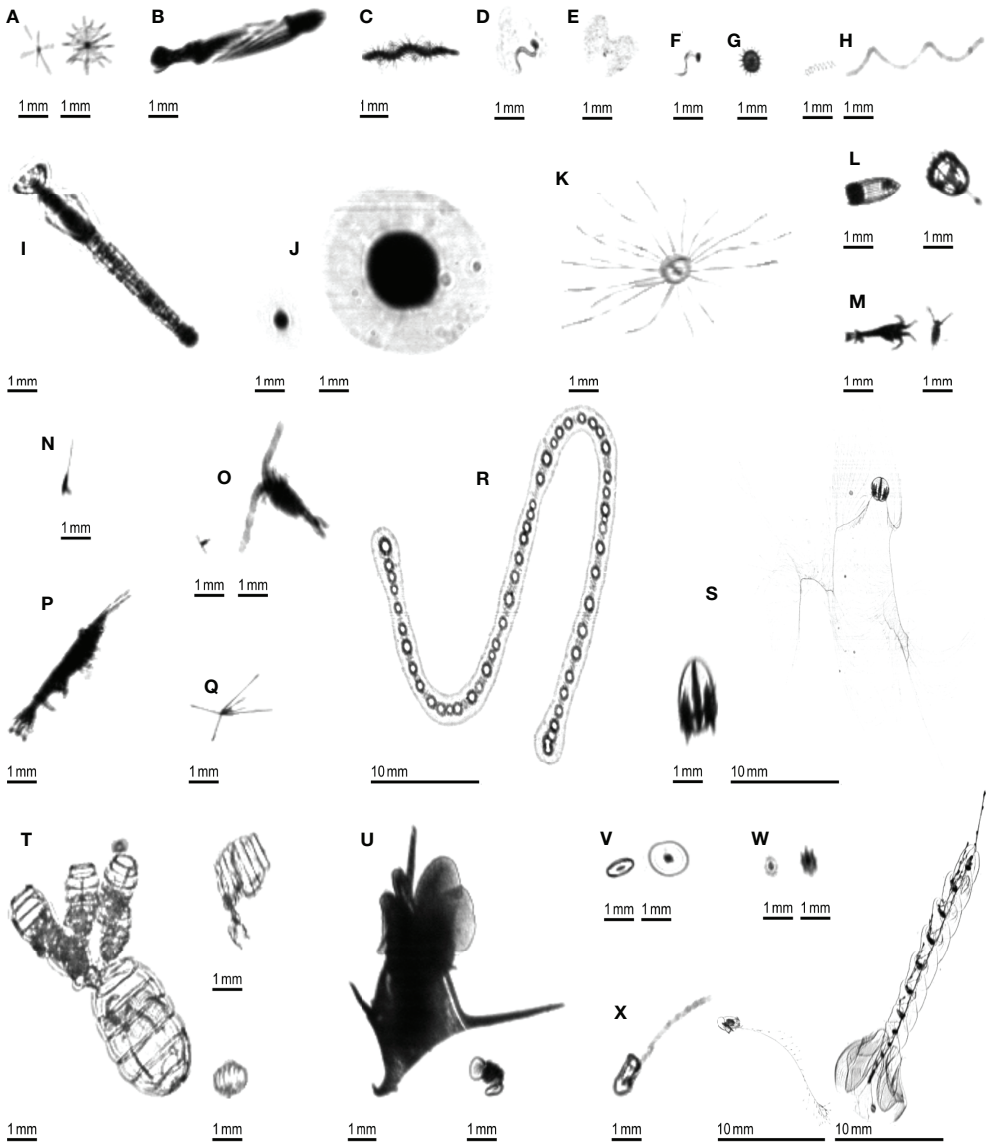


FIGURE 3 | Examples of planktonic organisms imaged by the ISIIS. **(A)** Acantharea; **(B)** Actinopterygii; **(C)** Annelida; **(D)** Appendicularia; **(E)** Appendicularia (house only); **(F)** Appendicularia (body only); **(G)** Aulacanthidae; **(H)** Bacillariophyceae; **(I)** Chaetognatha; **(J)** solitary Collodaria; **(K)** Hydrozoa; **(L)** Cnidaria (other than Hydrozoa); **(M)** Crustacea (other than Harpacticoida, Copepoda and Eumalacostraca); **(N)** Harpacticoida; **(O)** Copepoda (other than Harpacticoida); **(P)** Eumalacostraca; **(Q)** Echinodermata (pluteus larva); **(R)** colonial Collodaria; **(S)** Ctenophora; **(T)** Doliolida; **(U)** Mollusca; **(V)** Pyrocystis; **(W)** Rhizaria (other than Acantharea; Aulacanthidae and Collodaria); **(X)** Siphonophorae.

different taxa, which would make it unexploitable ecologically. Both choices made the match metrics conservative.

From these matches, global precision and recall were computed to summarize performance. Precision was computed as the proportion of automatic segments that matched ground truth segments. A 100% precision means that the algorithm only extracted ground truth segments. Recall was computed as the proportion of ground truth segments detected by the automated segmentation algorithm. A 100% recall means the algorithm did segment every manually delineated organism. Precision and recall scores were also computed per size class, where size was defined as the length of the diagonal of the bounding box; size classes were defined as intervals of 10 pixels, from 10 to 100 pixels, plus a class > 100 pixels. These size classes do not aim at reflecting any ecological groups but were designed to split segments into roughly balanced classes. Recall was also computed for each taxonomic group defined in the ground truth segments. Precision does not make sense for taxonomic groups since it would only reflect the performance of the classification, not of the segmentation. The particle matching and metric computation code is available at https://github.com/ThelmaPana/segmentation_benchmark.

3 RESULTS

3.1. Number and Size Distribution of Segments

On the 106 images of the segmentation benchmark dataset, 3,356 organisms were manually segmented, whereas the automated pipelines generated many more segments, especially the threshold-based one (**Table 2**).

The normalized abundance size spectra (NASS) (**Figure 4**) display the expected linear decrease of abundance with size in log-log scale. For the ground truth segments, the curve dips below this linear relationship for objects of 25 pixels in diagonal and smaller (dotted vertical line on **Figure 4**). Since this dataset specifically targeted recognisable planktonic organisms, this dip highlights that not all organisms below this size could be detected by a human taxonomist upon detailed examination of the images (Lombard et al., 2019). The discontinuity is towards smaller diagonal sizes in the automated pipelines, but likely because many of the small segments are of non-plankton objects.

All automated pipelines have NASS curves above the ground truth, which highlights the fact that they segmented non-plankton objects. This was true over the entire size range but was particularly pronounced for the smaller size classes. Above 10 mm/200 pixels in diagonal, the T-MSER pipeline produced a number of segments comparable to the ground truth, which is satisfying, although it does not guarantee that those are of the same objects (it might have missed some plankton and segmented marine snow/artifacts in the same size range; see precision and recall performances for the largest size class in **Figure 5** below). From the maximal size down to ~70 pixels in diagonal, the T and T-CNN pipelines produced the same segments. This coincides with the critical size of 400 pixels in area at which the segmentation method switched from threshold-based to content-aware. Indeed, the conversion from area to bounding box diagonal is not linear because it depends on the shape of the objects. For an object of 400 pixels in area, the bounding box diagonal is between 30 and 70 pixels. This shows that the T-CNN pipeline was effective in reducing the number of segments compared to naive thresholding, because the NASS diverges below that size.

A linear regression performed on the linear portion of the NASS (diagonal values between 30 and 500 pixels) followed by an analysis

TABLE 2 | Number of segments generated by each pipeline on the 106 benchmark images and estimation of the amount of segments they would produce on one minute of ISIIS data.

Segmentation pipeline	Number of segments on benchmark images	Average number of segments per minute of ISIIS deployment
Ground truth	3,356	~5,000
Threshold	339,907	~525,000
Threshold-MSER	82,731	~130,000
Threshold-CNN	19,048	~30,000

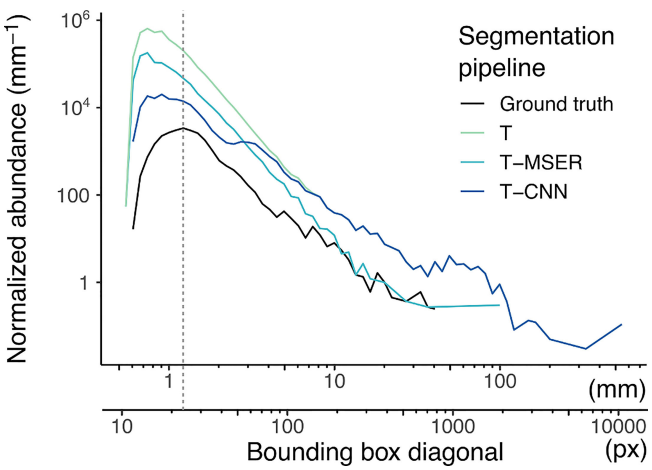


FIGURE 4 | Normalized abundance size spectra (NASS) of all segments generated by the benchmarked pipelines and ground truth segmentation. To compute the NASS, segments were grouped into size classes on a log2 scale, each class size being two times wider than the previous one. Normalized abundance was computed by dividing the number of segments in each class by the size class width, resulting in an adimensional quantity (number of segments) divided by a length (mm here). The double x-axis is the length of the diagonal bounding box displayed both in pixels and after conversion in mm. The dotted vertical line highlights the slope discontinuity in the size spectrum of ground truth segments. Note that both axes use log10 scaling. T, threshold-based; T-MSER, threshold-MSER; T-CNN, threshold-CNN.

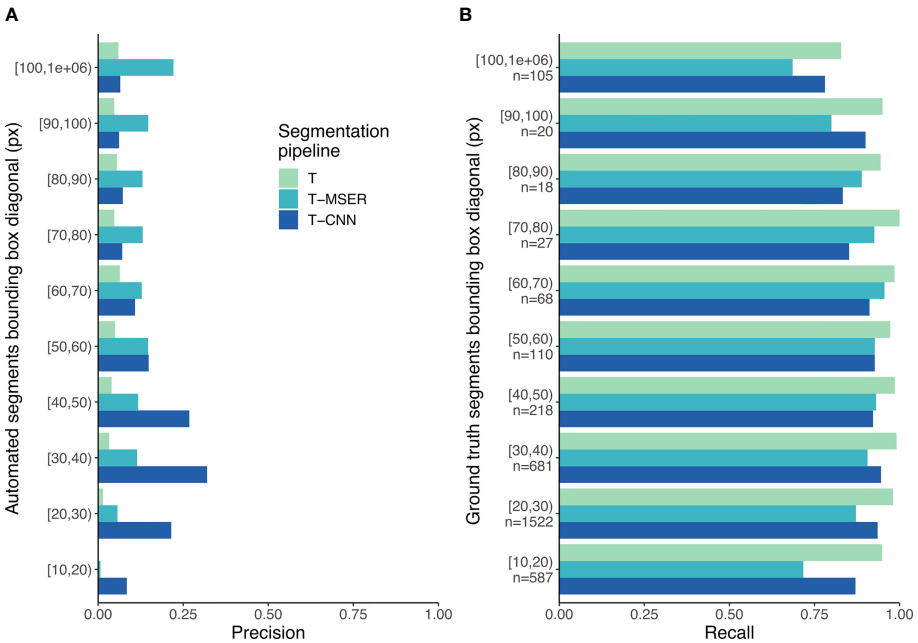


FIGURE 5 | Precision **(A)** and recall **(B)** scores per size class. In **(B)**, n indicates the number of segments per size class for the ground truth dataset. T, threshold-based; T-MSER, threshold-MSER; T-CNN, threshold-CNN.

of covariance demonstrated significant difference in slopes between the segmentation methods: $F(3,105) = 133.07$; $p < 0.001$ (**Table S1**). *Post hoc* analysis showed a significant difference between all segmentation methods ($p < 0.001$ for all pairs) (**Table S2**).

3.2. Global Performance Statistics

Overall, the three pipelines demonstrate good recall: when looking at the total number of segments, they all captured over 85% of the ground truth organisms. The T-CNN pipeline largely outperformed both the threshold-based and T-MSER pipelines in terms of precision (**Table 3**). In other words, although it segmented almost all planktonic objects, the threshold-based pipeline generated mostly non-plankton segments (~99%), composed of both marine snow and density volutes artifacts. The T-CNN pipeline also produced non-planktonic segments but they "only" represented 84% of segments, while still segmenting a good proportion of planktonic objects. The T-MSER performed somewhere in between those two extremes.

3.3. Performances Per Size Class

Because the behavior of the pipelines seems to vary with size (**Figure 4**), it seems relevant to break down the matching statistics per size class. With the threshold-based pipeline, precision decreased with size: smaller segments included a lower proportion of planktonic organisms than larger ones (**Figure 5A**). The T-CNN pipeline had better precision than the others for small segments while T-MSER had a better precision for larger segments. In terms of recall, the threshold-based pipeline always performed better than the others, regardless of size class (**Figure 5B**). The T-MSER pipeline performed as well as the T-CNN pipeline on middle size classes, but achieved a lower recall for both very small and very large segments.

TABLE 3 | Precision and recall values of the automated pipelines evaluated against the 3,356 ground truth organisms.

Pipeline	Precision	Recall
Threshold	0.9%	97.3%
Threshold-MSER	3.5%	85.4%
Threshold-CNN	16.3%	91.9%

3.4. Performances Per Taxonomic Group

In the ground truth dataset, half of the 24 detected taxa were represented by fewer than 18 individuals (median is 18.5), hence inducing little resolution and large variance in the performance statistics of segmentation pipelines. Among the other half of the taxa, the recall of the T-CNN pipeline was lower than that of the threshold pipeline by more than 10% for only two taxa (Bacillaryophycea and Doliolida) and for only four in the case of the T-MSER pipeline (Bacillariophyceae, Ctenophora, Acantharea, and other Rhizaria; **Figure 6**). The lowest recall values were reached for Bacillariophyceae and Ctenophora, for all pipelines. In concordance with the consistent recall performance across size classes, taxa-wise recall performance of

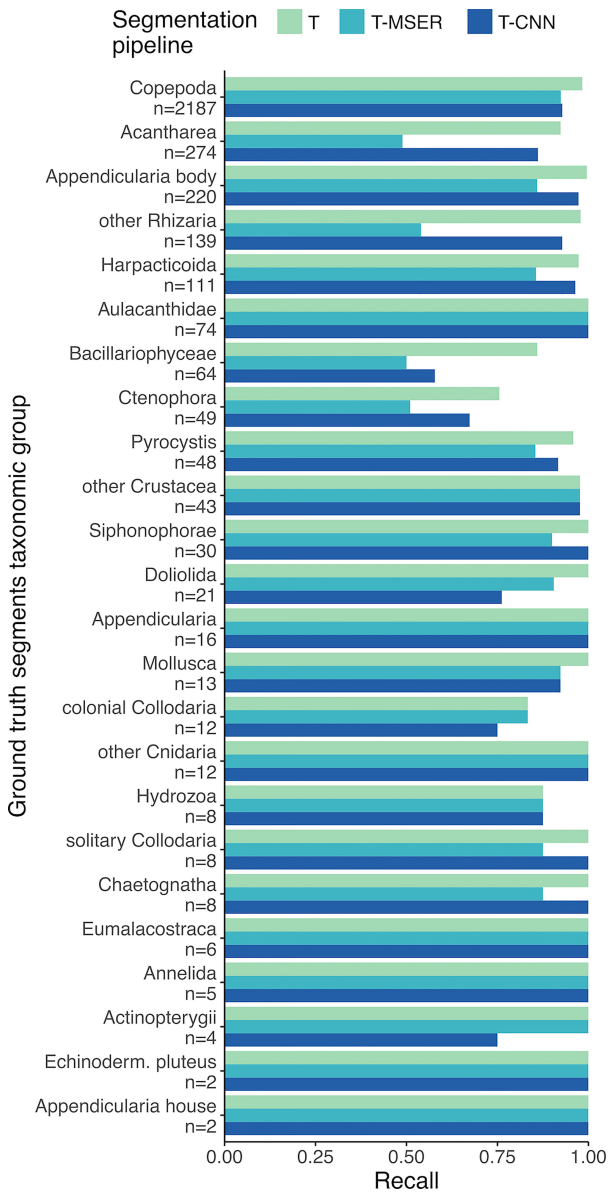


FIGURE 6 | Recall scores per taxon. n is the number of individuals from each taxon in the 106 benchmark images and taxa are sorted in decreasing order of abundance. T, threshold-based; T-MSER, threshold-MSER; T-CNN, threshold-CNN.

the T-CNN pipeline do not seem linked to organism size: small organisms (e.g. Acantharea, Pyrocystis) were accurately detected.

4 DISCUSSION

4.1. Summary of Results

The threshold-based pipeline performed an exhaustive segmentation: planktonic organisms were almost all properly detected, yet they were drowned in the overwhelming majority of non-planktonic objects (**Table 2**). The T-CNN pipeline reduced this problem, significantly increasing precision (**Table 3** and **Figure 5A**) while still achieving a very good detection of plankton across the entire size range targeted by ISIIS. The T-MSER pipeline also reduced the segmentation of non-planktonic objects, especially at the top-end of the size range, but detected fewer planktonic organisms than the other pipelines (**Figure 5B**). Despite the large decrease in number of segmented objects, for most taxa, the MSER or CNN pipelines reduced recall by less than 10% (**Figure 6**). One explanation for these differences is that naive thresholding captured a lot of noise (i.e. density volutes) and, additionally, broke it into many small segments. The use of either MSER or a CNN allowed ignoring these noise segments and/or not breaking them apart, hence producing much fewer non-planktonic segments. The decrease in abundance below the expected slope at the smaller end of the size spectrum of ground truth segments (**Figure 4**) suggests that identification of planktonic organisms becomes non-exhaustive below 25 pixels in bounding box diagonal. Below this size, which amounts to 600 µm in ESD on average, some organisms can still be detected. This means that relative concentrations between locations/times can likely be exploited within a taxon but that further filtering and corrections are needed to reach absolute concentrations.

The statistical difference between NASS slopes (**Figure 4**) indicates that they segment different kinds and amounts of non-planktonic objects, compared to the all-plankton ground truth. This implies that the output of different segmentation approaches should not be directly compared in terms of size distribution. Segmentation methods were already shown to have an impact on the definition of particle size and shape, which propagates to subsequent analyses such as particle flux estimates (Giering et al., 2020). This slope discrepancy as well as the vastly larger intercept of the NASS of automated pipelines compared to the ground truth means that the computation of an appropriate plankton size spectrum requires a classification step that would exclude non-planktonic objects.

4.2. Targeted Organisms

Some taxa were systematically less often detected than others. Some of the not detected Bacillariophyceae were large, blurry, and too translucent (**Figure 3H**) to be caught by the threshold-based branch of the T-CNN pipeline or by the T-MSER method. The other, smaller, ones that were missed by the content-aware branch of T-CNN were not detected because they were quite different from the ones used during training (blurrier). Integrating more representative examples of Bacillariophyceae for CNN training could have improved performance on this taxon. Similarly, doliolids (**Figure 3T**), that were often large, should have been segmented by the threshold-based branch of T-CNN as well as by T-MSER. The ones missed, mostly by T-CNN, were also blurry and too translucent for intensity-based thresholding with a single threshold. Ctenophores (likely of the Mertensiidae family, **Figure 3S**) displayed thin, translucent tentacles that were often missed by threshold-based methods. Therefore, only the body was segmented, which resulted in a bounding box IoU value < 0.1, too low to be considered a match with the ground truth segment that included the tentacles. Still, a

later CNN classifier should be able to correctly identify even such portions of organisms, as CNNs were shown to mostly rely on local shape and texture features instead of on the global shape (Baker et al., 2018; Baker et al., 2020). Finally, the T-MSER pipeline resulted in a lower recall for Acantharea and other Rhizaria (**Figures 3A, W**). This seems to stem from a too aggressive thresholding step in low SNR high noise frames, the pre-processing step before MSER is applied. Further fine-tuning would likely allow it to retain more or all Acantharea and other Rhizaria images.

In the present study, we aimed at performing an exhaustive detection of every planktonic organism across the size range targeted by the ISIIS. However, in general, the segmentation algorithm should be chosen according to the target organisms. For example, to focus on organisms towards the larger end of the ISIIS size range (e.g. > 10 mm), where particles — mostly marine snow aggregates — are much less abundant, a simple gray-level threshold seems sufficient.

4.3. Processing Time and Cost

The quantile-based thresholding pipeline ran on a single CPU core at a rate of 30 minutes of processing for 1 minute of ISIIS data (0.03x), on an Intel Xeon E5-2643 v3 (3.40 GHz). Its memory requirements were limited so it was easy to run simultaneous processing of multiple batches of data on a multi-core/multi-processor machine, but the treatment of ISIIS data as a continuous stream for flat-fielding prevented automatic multithreading. The T-CNN pipeline required a GPU with sufficient memory (48 GB, on aNvidia Quadro RTX 8000 in our case) to efficiently train the CNN portion and to fit ISIIS images in at evaluation time. It processed data at the same rate as the threshold-based pipeline (30 min processing for 1 min of data, or 0.03x). The T-MSER pipeline was optimized for speed and utilized the 8 cores of an AMD Ryzen 3700, processing one minute of ISIIS data in 50 seconds (1.2x), or 6 min 40 s of processing for 1 min of ISIIS data (0.15x) when considering running on one core.

The MSER implementation followed Matas et al. (2004) closely. The optimization of the T-MSER approach stems from adding the SNR switch, which leads to the pre-processing of high-noise images with naive thresholding, while going straight to the MSER-based detection in low noise images. Adding these changes increased segmentation recall from 65% to 85%. Further optimization included making the code multi-thread ready for deployment on High Performance Computing infrastructures. Using the specialized CPUs of these infrastructures, such as the AMD EPYC 7742 (64 cores, 128 threads) performance could improve well above 1.2x. At current data collection rates of 75-100 h of ISIIS data per scientific cruise, a real time or faster than real time segmentation approach constitutes a substantial benefit.

At first glance, the T-CNN pipeline seems expensive in terms of set up and architecture: it requires a GPU with sufficient memory to operate, implies the use of relatively new deep learning coding frameworks and the preparation of a training set with manual delineation of thousands of planktonic organisms. But these costs are offset by the time gained not processing a multitude of particles in each image, resulting in a processing rate comparable to that of the pure threshold-based pipeline, as stated above. Furthermore, the fact that T-CNN produced 20 times fewer segments will also considerably reduce the classification time (often CNN based too). Finally, since recall barely decreased, the objects ignored were mostly the dominant non-plankton objects, as per design; this will diminish the imbalance among classes that classifiers are sensitive too, further improving the classification step. Moreover, both the Detectron2 library and the baseline model on which the T-CNN pipeline relies are easily downloadable and well documented[2]. With GPU resources becoming increasingly available for scientific research and the associated frameworks becoming easier to use, such tools are poised to become more powerful and accessible.

4.4. Detection of Small Objects by CNN Models

The detection of objects measuring just a few pixels is still a research problem in its own right in computer sciences (Eggert et al., 2017), coined very low resolution recognition problems (Wang et al., 2016). They are characterized by targets smaller than 16×16 pixels, which can be challenging even for the perceptual abilities of human experts. They target applications for company logo detection (Eggert et al., 2016; Eggert et al., 2017), face recognition from video surveillance, or text recognition (Wang et al., 2016). The receptive fields of common object detection architectures match the target object size and range from 50×50 to 450×450 pixels which is much larger than the small objects targeted in low resolution studies (Eggert et al., 2017). Here, the smallest organisms targeted had an area of 50 pixels, which corresponded to a bounding box diagonal of 12 pixels, or an 8x8 pixels square. Thus the exhaustive detection of plankton organisms in ISIIS images, including the smaller ones, clearly falls in the domain of very low resolution recognition. A common solution is image upscaling, as highlighted by Eggert et al. (2016), which we implemented in the present work. The 524×524 pixels crops were upscaled to 900×900 pixels before evaluation in the Detectron2 model. The 900 pixels size is a compromise between detection accuracy, usage of the GPU memory, and processing time. Other approaches for multi-scale object detection are described by (Cai et al., 2016) and include magnification of regions susceptible to contain small objects (Eggert et al., 2016) or the integration of contextual information outside of regions of interest (Bell et al., 2016).

No automated segmentation method is perfect; depending on their settings, they either avoid objects other than their targets but miss some objects of interest (high precision, low recall) or detect most objects of interest but also many others (high recall, low precision). If the segmentation or object detection task is followed by a classification step, which is always the case for plankton imaging, we advocate in favor of recall over precision during segmentation, provided that the amount of data remains manageable. Hence, a maximum number of planktonic objects have the opportunity to be classified. The precision can be improved after classification, by filtering out low confidence, usually error prone, predictions

based on the score given by the classifier (Faillettaz et al., 2016; Luo et al., 2018).

To extract planktonic organisms of various taxa from ISIIS images, full instance segmentation would have been the most elegant approach, outputting classified mask instances in a single step (Dai et al., 2016). Several obstacles still lay ahead for this approach to be applicable. First, training an instance segmentation model to recognize each taxonomic group would require hundreds to thousands of ground truth (i.e. human-produced) masks of all taxa. Given the long tailed distribution of taxa concentrations in the planktonic world, with many rare taxa, in particular the largest ones, this would require a considerable amount of searching and labeling effort. Indeed, assembling enough examples to train classifications models is already challenging (Irisson et al., 2022) and manual delineation of each organism is much more time consuming than manual classification. A second obstacle is the size range of organisms imaged by ISIIS. Although Detectron2 does produce multi-scale feature maps through a Feature Pyramid Network in order to apply receptive fields of multiple size, the ratio between the largest and the smallest feature maps is only 16. Here, the ratio between the smallest and largest bounding box diagonals of manually segmented organisms is 65 and can reach > 180 in more exhaustive ISIIS datasets. To tackle this span, one could theoretically set up an ensemble of detectors, fed with crops of different sizes, each one targeting a restricted size range. Yet, this would be a particularly computationally demanding and complex set up, for a gain yet to be determined since, for larger sizes, the proportion of non-plankton objects, and therefore the advantage of a CNN-based segmentation, diminishes. Finally, masks generated by instance segmentation models currently lack both precision (their outline is smoothed, not matching the fine appendages of plankton) and reproducibility (because of the randomness included during training to avoid overfitting, two models trained on the same data will output different masks). These drawbacks are particularly critical for plankton application, where the size of the organisms, computed from their masks, is often of interest.

5 CONCLUSION AND PERSPECTIVES

We developed combined segmentation pipelines able to detect planktonic organisms spanning a broad size range. The fact that all methods comprised a deterministic, threshold-based segmentation ensured that particle shapes and measurement were consistent over the whole size range. Still, the segmentation method affected the shape of the size spectrum and additional processing steps (including classification) are needed to extract the correct size structure of living organisms. The MSER method limited over-segmentation of background noise objects and extracted more consistent segments, at a very high processing rate. This speed opens the possibility for near-real time processing, which is particularly relevant for adaptive sampling during a cruise or an early warning system in a time series context. Although at the lower limit of the detection capabilities of CNNs, our content-aware approach was able to detect planktonic organisms among an overwhelming number of marine snow and noise images, exhibiting the best recall of the three methods. Therefore, the ideal segmentation approach depends on the study objectives and operational constraints.

These approaches seem relevant for imaging studies focused on living planktonic organisms, since they reduce the number of objects from non-plankton classes that are extracted. In turn, this dampens the imbalance towards these classes, laying the foundations for easier, faster, and more accurate subsequent object classification by (i) reducing the amount of work needed to generate a training set with similar class distribution, which is essential to avoid the caveat of dataset shift (Moreno-Torres et al., 2012); (ii) decreasing the computation time because there are fewer objects; and (iii) limiting the contamination of the rare planktonic classes by the dominant, non-plankton, ones.

Although CNN-based object detection may seem overwhelming at first, both in terms of set up and processing time, it actually is fast enough and within the reach of marine ecologists, particularly now that artificial intelligence frameworks and GPU computing are being made more accessible. This work constitutes a step towards the "intelligent" segmentation of ecological images, even at low resolution, which could find even wider applications such as the automated separation of objects overlapping onto each other on an image for more accurate species counts, the detection and classification in a single step for more automated surveys, or the extraction of individual-level traits to track e.g., reproductive organs development, for a richer exploitation of ecological images (Orenstein et al., 2021). Such tasks are in no way limited to plankton images and are common in data collected by trawl cameras, benthic observations or surveying cameras, vessel monitoring cameras, etc.

In this era of data-driven oceanography, the volume of data collected is increasing sharply, thanks to technological advances such as high frequency imagery, autonomous instruments (e.g. floats, gliders), satellite-based methods as well as environmental -omics approaches permitted by high throughput sequencing. In this context of abundant data, the development of automated and efficient data processing techniques becomes a key element in drawing a holistic understanding of oceanic ecosystems; it is needed to provide an extensive description of biodiversity, including species distributions as well as estimates of biomass and abundance.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

J-OI and TP conceptualized the study. LC-C and TP generated and taxonomically sorted ground truth plankton segments. BW and TP developed the CNN segmentation method. J-OI and TP developed the threshold-based and the T-CNN processing

pipelines. MS, DD, ST, CS, and RC set up and ran the T-MSER method. TP prepared the original draft. All co-authors proofread the manuscript prior to submission. All authors read and approved the final manuscript.

FUNDING

This study is part of project "World Wide Web of Plankton Image Curation", funded by the Belmont Forum through the Agence Nationale de la Recherche ANR-18-BELM-0003-01 and the National Science Foundation (NSF) #ICER1927710. Funding also came from NSF #OCE2125407. The Extreme Science and Engineering Discovery Environment (XSEDE) provided computing resources to the US team through grant #TG-OCE170012. Data acquisition during the VISUFRONT cruise was funded by the Partner University Fund and supported by the French Oceanographic Fleet through ship time. TP's doctoral fellowship was granted by the French Ministry of Higher Education, Research and Innovation (#3500/2019).

ACKNOWLEDGMENTS

The authors would like to thank the officers and crew of the R/V Tethys 2 who made the VISUFRONT campaign a success, as well as the additional scientists who took part in the cruise: R Faillettaz, C M Guigand, F Lombard, M Lilley, and J Luo.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.870005/full#supplementary-material

REFERENCES

Alldredge, A. L., Granata, T. C., Gotschalk, C. C., and Dickey, T. D. (1990). The Physical Strength of Marine Snow and Its Implications for Particle Disaggregation in the Ocean. *Limnol. Oceanog.* 35, 1415–1428. doi: 10.4319/lo.1990.35.7.1415

Alldredge, A. L., and Silver, M. W. (1988). Characteristics, Dynamics and Significance of Marine Snow. *Prog. Oceanog.* 20, 41–82. doi: 10.1016/0079-6611(88)90053-5

Baker, N., Lu, H., Erlikhman, G., and Kellman, P. J. (2018). Deep Convolutional Networks do Not Classify Based on Global Object Shape. *PloS Comput. Biol.* 14, e1006613. doi: 10.1371/journal.pcbi.1006613

Baker, N., Lu, H., Erlikhman, G., and Kellman, P. J. (2020). Local Features and Global Shape Information in Object Classification by Deep Convolutional Neural Networks. *Vision Res.* 172, 46–61. doi: 10.1016/j.visres.2020.04.003

Bell, S., Zitnick, C. L., Bala, K., and Girshick, R. (2016). "Inside-Outside Net: Detecting Objects in Context With Skip Pooling and Recurrent Neural Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2874–2883.

Benfield, M., Grosjean, P., Culverhouse, P., Irigolen, X., Sieracki, M., Lopez-Urrutia, A., et al. (2007). RAPID: Research on Automated Plankton Identification. *Oceanography* 20, 172–187. doi: 10.5670/oceanog.2007.63

Biard, T., and Ohman, M. D. (2020). Vertical Niche Definition of Test-Bearing Protists (Rhizaria) Into the Twilight Zone Revealed by *in Situ* Imaging. *Limnol. Oceanog.* 65, 2583–2602. doi: 10.1002/lno.11472

Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., et al. (2016). *In Situ* Imaging Reveals the Biomass of Giant Protists in the Global Ocean. *Nature* 532, 504–507. doi: 10.1038/nature17652

Bi, H., Guo, Z., Benfield, M. C., Fan, C., Ford, M., Shahrestani, S., et al. (2015).). A Semi-Automated Image Analysis Procedure for *In Situ* Plankton Imaging Systems. *PloS One* 10, e0127121. doi: 10.1371/journal.pone.0127121

Brandão, M. C., Benedetti, F., Martini, S., Soviadan, Y. D., Irisson, J.-O., Romagnan, J.-B., et al. (2021). Macroscale Patterns of Oceanic Zooplankton Composition and Size Structure. *Sci. Rep.* 11, 15714. doi: 10.1038/s41598-021-94615-5

Briseño-Avena, C., Schmid, M. S., Swieca, K., Sponaugle, S., Brodeur, R. D., and Cowen, R. K. (2020). Three-Dimensional Cross-Shelf Zooplankton Distributions Off the Central Oregon Coast During Anomalous Oceanographic Conditions. *Prog. Oceanog.* 188, 102436. doi: 10.1016/j.pocean.2020.102436

Cai, Z., Fan, Q., Feris, R. S., and Vasconcelos, N. (2016). "A Unified Multi-Scale Deep Convolutional Neural Network for Fast Object Detectionv," in *European Conference on Computer Vision (ECCV)*, 1607.07155.

Cheng, K., Cheng, X., Wang, Y., Bi, H., and Benfield, M. C. (2019). Enhanced Convolutional Neural Network for Plankton Identification and Enumeration. *PloS One* 14, e0219570. doi: 10.1371/journal.pone.0219570

Cowen, R. K., and Guigand, C. M. (2008). *In Situ* Ichthyoplankton Imaging System (ISIIS): System Design and Preliminary Results. *Limnol. Oceanog.: Methods* 6, 126–132. doi: 10.4319/lom.2008.6.126

Dai, J., He, K., and Sun, J. (2016). "Instance-Aware Semantic Segmentation *via* Multi-Task Network Cascades," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3150–3158.

Dennett, M. R., Caron, D. A., Michaels, A. F., Gallager, S. M., and Davis, C. S. (2002). Video Plankton Recorder Reveals High Abundances of Colonial Radiolaria in Surface Waters of the Central North Pacific. *J. Plank. Res.* 24, 797–805. doi: 10.1093/plankt/24.8.797

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015). Eukaryotic Plankton Diversity in the Sunlit Ocean. *Science* 348, 1261605. doi: 10.1126/SCIENCE.1261605

Eggert, C., Brehm, S., Winschel, A., Zecha, D., and Lienhart, R. (2017). "A Closer Look: Small Object Detection in Faster R-CNN," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, 421–426. doi: 10.1109/ICME.2017.8019550

Eggert, C., Winschel, A., Zecha, D., and Lienhart, R. (2016). "Saliency-Guided Selective Magnification for Company Logo Detection," in *23rd International Conference on Pattern Recognition (ICPR)*, 651–656. doi: 10.1109/ICPR.2016.7899708

Faillettaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K., and Irisson, J.-O. (2016). Imperfect Automatic Image Classification Successfully Describes Plankton Distribution Patterns. *Methods Oceanog.* 15–16, 60–77. doi: 10.1016/J.MIO.2016.04.003

Falkowski, P. (2012). Ocean Science: The Power of Plankton. *Nature* 483, S17–S20. doi: 10.1038/483S17a

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* 281, 237–240. doi: 10.1126/science.281.5374.237

Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., et al. (2012). Size Distribution of Particles and Zooplankton Across the Shelf-Basin System in Southeast Beaufort Sea: Combined Results From an Underwater Vision Profiler and Vertical Net Tows. *Biogeosciences* 9, 1301–1320. doi: 10.5194/bg-9-1301-2012

Frederiksen, M., Edwards, M., Richardson, A. J., Halliday, N. C., and Wanless, S. (2006). From Plankton to Top Predators: Bottom-Up Control of a Marine Food Web Across Four Trophic Levels. *J. Anim. Ecol.* 75, 1259–1268. doi: 10.1111/j.1365-2656.2006.01148.x

Giering, S. L. C., Hosking, B., Briggs, N., and Iversen, M. H. (2020). The Interpretation of Particle Size, Shape, and Carbon Flux of Marine Particle Images Is Strongly Affected by the Choice of Particle Detection Algorithm. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00564

Greer, A. T., Boyette, A. D., Cruz, V. J., Cambazoglu, M. K., Dzwonkowski, B., Chiaverano, L. M., et al. (2020a). Contrasting Fine-Scale Distributional Patterns of Zooplankton Driven by the Formation of a Diatom-Dominated Thin Layer. *Limnol. Oceanog.* 65, 2236–2258. doi: 10.1002/lno.11450

Greer, A. T., Chiaverano, L. M., Luo, J. Y., Cowen, R. K., and Graham, W. M. (2018). Ecology and Behaviour of Holoplanktonic Scyphomedusae and Their Interactions With Larval and Juvenile Fishes in the Northern Gulf of Mexico. *ICES. J. Mar. Sci.* 75, 751–763. doi: 10.1093/icesjms/fsx168

Greer, A. T., Chiaverano, L. M., Treible, L. M., Briseño-Avena, C., and Hernandez, F. J. (2021). From Spatial Pattern to Ecological Process Through Imaging Zooplankton Interactions. *ICES. J. Mar. Sci* 78(8):2664–74. doi: 10.1093/icesjms/fsab149

Greer, A. T., Cowen, R. K., Guigand, C. M., and Hare, J. A. (2015). Fine-Scale Planktonic Habitat Partitioning at a Shelf-Slope Front Revealed by a High-Resolution Imaging System. *J. Mar. Syst.* 142, 111–125. doi: 10.1016/j.jmarsys.2014.10.008

Greer, A. T., Cowen, R. K., Guigand, C. M., Hare, J. A., and Tang, D. (2014). The Role of Internal Waves in Larval Fish Interactions With Potential Predators and Prey. *Prog. Oceanog.* 127, 47–61. doi: 10.1016/j.pocean.2014.05.010

Greer, A. T., Cowen, R. K., Guigand, C. M., McManus, M. A., Sevadjian, J. C., and Timmerman, A. H. V. (2013). Relationships Between Phytoplankton Thin Layers and the Fine-Scale Vertical Distributions of Two Trophic Levels of Zooplankton. *J. Plank. Res.* 35, 939–956. doi: 10.1093/plankt/fbt056

Greer, A. T., Lehrter, J. C., Binder, B. M., Nayak, A. R., Barua, R., Rice, A. E., et al. (2020b). High-Resolution Sampling of a Broad Marine Life Size Spectrum Reveals Differing Size- and Composition-Based Associations With Physical Oceanographic Structure. *Front. Mar. Sci.* 7 (8), 2664–2674. doi: 10.3389/fmars.2020.542701

Guidi, L., Jackson, G. A., Stemmann, L., Miquel, J. C., Picheral, M., and Gorsky, G. (2008). Relationship Between Particle Size Distribution and Flux in the Mesopelagic Zone. *Deep. Sea. Res. Part I.: Oceanog. Res. Pap.* 55, 1364–1374. doi: 10.1016/J.DSR.2008.05.014

Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., and Henson, S. A. (2015). A New Look at Ocean Carbon Remineralization for Estimating Deepwater Sequestration. *Global Biogeochem. Cycle.* 29, 1044–1059. doi: 10.1002/2014GB005063

He, K., Gkioxari, G., Dollar, P., Girshick, R.Mask R-CNN (2017). *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2961–2969.

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., et al. (2019). Global Trends in Marine Plankton Diversity Across Kingdoms of Life. *Cell* 179, 1084–1097.e21. doi: 10.1016/j.cell.2019.10.008

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. (2022). Machine Learning for the Study of Plankton and Marine Snow From Images. *Ann. Rev. Mar. Sci.* 14, 277–301. doi: 10.1146/annurev-marine-041921-013023

Iyer, N. (2012). *Machine Vision Assisted* in Situ *Ichthyoplankton Imaging System* (Purdue University).

Lee, H., Park, M., and Kim, J. (2016). "Plankton Classification on Imbalanced Large Scale Database *via* Convolutional Neural Networks With Transfer Learning," in *IEEE International Conference on Image Processing (ICIP)*, 3713–3717. doi: 10.1109/ICIP.2016.7533053

Lévy, M., Franks, P. J. S., and Smith, K. S. (2018). The Role of Submesoscale Currents in Structuring Marine Ecosystems. *Nat. Commun.* 9, 4758. doi: 10.1038/s41467-018-07059-3

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017). "Feature Pyramid Networks for Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1612.03144.

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6, 196. doi: 10.3389/fmars.2019.00196

Longhurst, A. R., and Glen Harrison, W. (1989). The Biological Pump: Profiles of Plankton Production and Consumption in the Upper Ocean. *Prog. Oceanog.* 22, 47–123. doi: 10.1016/0079-6611(89)90010-4

Luo, J. Y., Grassian, B., Tang, D., Irisson, J.-O., Greer, A. T., Guigand, C. M., et al. (2014). Environmental Drivers of the Fine-Scale Distribution of a Gelatinous Zooplankton Community Across a Mesoscale Front. *Mar. Ecol. Prog. Ser.* 510, 129–149. doi: 10.3354/meps10908

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., et al. (2018). Automated Plankton Image Analysis Using Convolutional Neural Networks. *Limnol. Oceanog.: Methods* 16, 814–827. doi: 10.1002/lom3.10285

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust Wide-Baseline Stereo From Maximally Stable Extremal Regions. *Imag. Vision Comput.* 22, 761–767. doi: 10.1016/j.imavis.2004.02.006

McClatchie, S., Cowen, R., Nieto, K., Greer, A., Luo, J. Y., Guigand, C., et al. (2012). Resolution of Fine Biological Structure Including Small Narcomedusae Across a Front in the Southern California Bight. *J. Geophys. Res.: Ocean.* 117, C04020. doi: 10.1029/2011JC007565

Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrera, F. (2012). A Unifying View on Dataset Shift in Classification. *Pattern Recognit.* 45, 521–530. doi: 10.1016/j.patcog.2011.06.019

Ohman, M. D., Davis, R. E., Sherman, J. T., Grindley, K. R., Whitmore, B. M., Nickels, C. F., et al. (2019). Zooglider: An Autonomous Vehicle for Optical and Acoustic Sensing of Zooplankton. *Limnol. Oceanog.: Methods* 17, 69–86. doi: 10.1002/lom3.10301

Olson, R. J., and Sosik, H. M. (2007). A Submersible Imaging-in-Flow Instrument to Analyze Nano-and Microplankton: Imaging FlowCytobot. *Limnol. Oceanog.: Methods* 5, 195–203. doi: 10.4319/lom.2007.5.195

Orenstein, E. C., Ayata, S.-D., Maps, F., Biard, T., Becker, É.C., Benedetti, F., et al. (2021). Machine Learning Techniques to Characterize Functional Traits of Plankton From Image Data. hal–03482282

Orenstein, E. C., Ratelle, D., Briseño-Avena, C., Carter, M. L., Franks, P. J. S., Jaffe, J. S., et al. (2020). The Scripps Plankton Camera System: A Framework and Platform for *in Situ* Microscopy. *Limnol. Oceanog.: Methods* 18, 681–695. doi: 10.1002/lom3.10394

Otsu, N. (1979). A Threshold Selection Method From Gray-Level Histograms. *IEEE Trans. Sys. Man. Cybernet.* 9, 62–66.

Parikh, D., Zitnick, C. L., and Chen, T. (2012). Exploring Tiny Images: The Roles of Appearance and Contextual Information for Machine and Human Object Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 34, 1978–1991. doi: 10.1109/TPAMI.2011.276

Picheral, M., Colin, S., and Irisson, J.-O. (2017). EcoTaxa, a Tool for the Taxonomic Classification of Images.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The Underwater Vision Profiler 5: An Advanced Instrument for High Spatial Resolution Studies of Particle Size Spectra and Zooplankton. *Limnol. Oceanog.: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462

Remsen, A., Hopkins, T. L., and Samson, S. (2004). What You See is Not What You Catch: A Comparison of Concurrently Collected Net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder Data From the Northeast Gulf of Mexico. *Deep. Sea. Res. Part I.: Oceanog. Res. Pap.* 51, 129–151. doi: 10.1016/J.DSR.2003.09.008

Robinson, K. L., Sponaugle, S., Luo, J. Y., Gleiber, M. R., and Cowen, R. K. (2021). Big or Small, Patchy All: Resolution of Marine Plankton Patch Structure at Micro- to Submesoscales for 36 Taxa. *Sci. Adv.* 7, eabk2904. doi: 10.1126/sciadv.abk2904

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., and Legendre, L. (2009). Global Latitudinal Variations in Marine Copepod Diversity and Environmental Factors. *Proc. R. Soc. B.: Biol. Sci.* 276, 3053–3062. doi: 10.1098/rspb.2009.0742

Rutherford, S., D'Hondt, S., and Prell, W. (1999). Environmental Controls on the Geographic Distribution of Zooplankton Diversity. *Nature* 400, 749–753. doi: 10.1038/23449

Schmid, M. S., Cowen, R. K., Robinson, K., Luo, J. Y., Briseño-Avena, C., and Sponaugle, S. (2020). Prey and Predator Overlap at the Edge of a Mesoscale Eddy: Fine-Scale, *in-Situ* Distributions to Inform Our Understanding of Oceanographic Processes. *Sci. Rep.* 10, 1–16. doi: 10.1038/s41598-020-57879-x schmid2020Prey

Schmid, M. S., Daprano, D., Jacobson, K. M., Sullivan, C., Briseño-Avena, C., Luo, J. Y., et al. (2021). A Convolutional Neural Network Based High-Throughput Image Classification Pipeline - Code and Documentation to Process Plankton Underwater Imagery Using Local HPC Infrastructure and NSF's XSEDE. *Zenodo*. doi: 10.5281/zenodo.4641158

Schmid, M. S., and Fortier, L. (2019). The Intriguing Co-Distribution of the Copepods Calanus Hyperboreus and Calanus Glacialis in the Subsurface Chlorophyll Maximum of Arctic Seas. *Element.: Sci. Anthropocene.* 7, 50. doi: 10.1525/elementa.388

Ser-Giacomi, E., Zinger, L., Malviya, S., De Vargas, C., Karsenti, E., Bowler, C., et al. (2018). Ubiquitous Abundance Distribution of non-Dominant Plankton Across the Global Ocean. *Nat. Ecol. Evol.* 2, 1243–1249. doi: 10.1038/s41559-018-0587-2

Sheldon, R. W., and Parsons, T. R. (1967). A Continuous Size Spectrum for Particulate Matter in the Sea. *J. Fish. Res. Board. Canada.* 24, 909–915. doi: 10.1139/f67-081

Sheldon, R. W., Prakash, A., and Sutcliffe, W. H. (1972). The Size Distribution of Particles in the Ocean. *Limnol. Oceanog.* 17, 327–340. doi: 10.4319/lo.1972.17.3.0327

Sieracki, C. K., Sieracki, M. E., and Yentsch, C. S. (1998). An Imaging-in-Flow System for Automated Analysis of Marine Microplankton. *Mar. Ecol. Prog. Ser.* 168, 285–296. doi: 10.3354/meps168285

Sosik, H. M., and Olson, R. J. (2007). Automated Taxonomic Classification of Phytoplankton Sampled With Imaging-in-Flow Cytometry. *Limnol. Oceanog.: Methods* 5, 204–216. doi: 10.4319/lom.2007.5.204

Stemmann, L., and Boss, E. (2012). Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. *Annu. Rev. Mar. Sci.* 4, 263–290.

Stemmann, L., Hosia, A., Youngbluth, M. J., Søiland, H., Picheral, M., and Gorsky, G. (2008). Vertical Distribution (0–1000 M) of Macrozooplankton, Estimated Using the Underwater Video Profiler, in Different Hydrographic Regimes Along the Northern Portion of the Mid-Atlantic Ridge. *Deep. Sea. Res. Part II.: Top. Stud. Oceanog.* 55, 94–105. doi: 10.1016/J.DSR2.2007.09.019

Stemmann, L., Picheral, M., and Gorsky, G. (2000). Diel Variation in the Vertical Distribution of Particulate Matter (>0.15mm) in the NW Mediterranean Sea Investigated With the Underwater Video Profiler. *Deep. Sea. Res. Part I.: Oceanog. Res. Pap.* 47, 505–531. doi: 10.1016/S0967-0637(99)00100-4

Swieca, K., Sponaugle, S., Briseño-Avena, C., Schmid, M. S., Brodeur, R. D., and Cowen, R. K. (2020). Changing With the Tides: Fine-Scale Larval Fish Prey Availability and Predation Pressure Near a Tidally Modulated River Plume. *Mar. Ecol. Prog. Ser.* 650, 217–238. doi: 10.3354/meps13367

Tittensor, D. P., Mora, C., Jetz, W., Lotze, H. K., Ricard, D., Berghe, E. V., et al. (2010). Global Patterns and Predictors of Marine Biodiversity Across Taxa. *Nature* 466, 1098–1101. doi: 10.1038/nature09329

Tsechpenakis, G., Guigand, C., and Cowen, R. K. (2007). Image Analysis Techniques to Accompany a New *In Situ* Ichthyoplankton Imaging System. *OCEANS 2007 - Europe*, 1–6. doi: 10.1109/OCEANSE.2007.4302271

Wang, Z., Chang, S., Yang, Y., Liu, D., and Huang, T. S. (2016). "Studying Very Low Resolution Recognition Using Deep Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 4792–4800.

Ware, D. M., and Thomson, R. E. (2005). Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific. *Science* 308, 1280–1284. doi: 10.1126/SCIENCE.1109049

Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., and Girshick, R. (2019). Detectron2.

Conflict of Interest: Author Ben Woodward was employed by company CVision AI.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

ORIGINAL RESEARCH
published: 10 June 2022
doi: 10.3389/fmars.2022.869088

Benchmarking and Automating the Image Recognition Capability of an *In Situ* Plankton Imaging System

*Kevin T. Le[1], Zhouyuan Yuan[1], Areeb Syed[1], Devin Ratelle[2], Eric C. Orenstein[2,3], Melissa L. Carter[2], Sarah Strang[2], Kasia M. Kenitz[2], Pedro Morgado[1], Peter J. S. Franks[2], Nuno Vasconcelos[1] and Jules S. Jaffe[2]**

[1] Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA, United States, [2] Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, United States, [3] Bioinspiration Laboratory, Monterey Bay Aquarium Research Institute, Moss Landing, CA, United States

OPEN ACCESS

Edited by:
Mark C. Benfield, Louisiana State University, United States

Reviewed by:
Aletta T. Yñiguez, University of the Philippines Diliman, Philippines
Hongsheng Bi, University of Maryland Center for Environmental Science (UMCES), United States
Antonio Carlos Domínguez-Brito, University of Las Palmas de Gran Canaria, Spain

***Correspondence:**
Jules S. Jaffe
jjaffe@ucsd.edu

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *03 February 2022*
Accepted: *08 April 2022*
Published: *10 June 2022*

Citation:
Le KT, Yuan Z, Syed A, Ratelle D, Orenstein EC, Carter ML, Strang S, Kenitz KM, Morgado P, Franks PJS, Vasconcelos N and Jaffe JS (2022) Benchmarking and Automating the Image Recognition Capability of an In Situ Plankton Imaging System. Front. Mar. Sci. 9:869088. doi: 10.3389/fmars.2022.869088

To understand ocean health, it is crucial to monitor photosynthetic marine plankton – the microorganisms that form the base of the marine food web and are responsible for the uptake of atmospheric carbon. With the recent development of *in situ* microscopes that can acquire vast numbers of images of these organisms, the use of deep learning methods to taxonomically identify them has come to the forefront. Given this, two questions arise: 1) How well do deep learning methods such as Convolutional Neural Networks (CNNs) identify these marine organisms using data from *in situ* microscopes? 2) How well do CNN-derived estimates of abundance agree with established net and bottle-based sampling? Here, using images collected by the *in situ* Scripps Plankton Camera (SPC) system, we trained a CNN to recognize 9 species of phytoplankton, some of which are associated with Harmful Algal Blooms (HABs). The CNNs evaluated on 26 independent natural samples collected at Scripps Pier achieved an averaged accuracy of 92%, with 7 of 10 target categories above 85%. To compare abundance estimates, we fit a linear model between the number of organisms of each species counted in a known volume in the lab, with the number of organisms collected by the *in situ* microscope sampling at the same time. The linear fit between lab and *in situ* counts of several of the most abundant key HAB species suggests that, in the case of dinoflagellates, there is good correspondence between the two methods. As one advantage of our method, given the excellent correlation between lab counts and *in situ* microscope counts for key species, the methodology proposed here provides a way to estimate an equivalent volume in which the employed microscope can identify in-focus organisms and obtain statistically robust estimates of abundance.

Keywords: underwater imaging, microscopy, harmful algal blooms, convolutional neural network, deep learning, automated image analysis, underwater microscopy

1 INTRODUCTION

Small plankton are an extremely diverse group of single-celled underwater organisms with profound effects on ocean health (Field et al., 1998): they form the foundation of the food web, contribute to the early developmental stages of commercially harvestable species, and their abundance and composition are tightly related to hydro-climatic change (Lombard et al., 2019). Planktonic organisms can also adversely affect the marine ecosystem by forming dense blooms, known as Harmful Algal Blooms (HABs), that can sicken or kill both marine organisms and humans *via* a variety of mechanisms. The appearance and composition of these HAB taxa is a topic of intense research since they have deleterious effects on human health, negatively affect fish stocks, and are linked to eutrophication that is likely to increase in the coming years (Sinha et al., 2017). These biological impacts have serious economic ramifications and there is urgent interest in developing inexpensive, automated ways to detect HABs and quantify their abundance (Lefebvre et al., 1999; Scholin et al., 2000; Kim et al., 2009; Smith et al., 2018). The main goal of this study is to examine the potential for *in situ* imaging microscopy, supported by automated deep learning algorithms, for providing reliable estimates of a variety of plankton including HAB species.

Most HAB monitoring programs use traditional plankton sampling techniques, such as net tows and bottle sampling (Castellani, 2010) to estimate *in situ* abundance. These approaches require physically collecting the samples, chemically preserving the organisms, and manually enumerating species with a lab microscope. This laborious process is severely limited by a number of factors: net tows can damage delicate organisms during collection (Hamner et al., 1975; Omori and Hamner, 1982); certain organisms may dissolve in the preservation solution without proper treatment (Beers and Stewart, 1970); and critically, physical collection and subsequent analysis of the samples is expensive in terms of cost and human labor, resulting in less frequent sampling than is desirable.

Due to these factors, there is increasing interest in the use of imaging systems to monitor plankton populations. These systems have the capability to quantify organisms at very local spatial and fine temporal resolution, therefore providing a more scalable solution for long-term analysis (Olson and Sosik, 2007; Iyer, 2012; Cowen et al., 2013; Culverhouse et al., 2014; Lombard et al., 2019). Currently, underwater microscopes either continuously take images of plankton as they freely flow through the camera's view (Picheral et al., 2010; Orenstein et al., 2020a; Picheral et al., 2021) or are sampled discretely *via* microfluidic systems (Olson and Sosik, 2007). These systems do not require manual collection or concentration of water, chemical treatment of samples, or the use of counting chambers. An additional benefit of *in situ* imaging is that the digital archives can be easily preserved for future re-analyses and wide scale dissemination. However, the major bottleneck for using *in situ* imaging instruments for monitoring is the sheer volume of data they collect. To speed up analysis, scientists have begun using automated classification methods, such as Support Vector Machines and Convolutional Neural Networks (CNNs) that are capable of processing these large imaging libraries (Sosik and Olson, 2007; LeCun et al., 2015; Orenstein and Beijbom, 2017; Luo et al., 2018; Ellen et al., 2019). The results indicate that CNNs can successfully identify a variety of marine organisms such as zooplankton, phytoplankton, coral, and fish (Orenstein et al., 2015; Salman et al., 2016; González et al., 2019). A recent review highlights the use of these methods, specifically, for plankton (Irisson et al., 2022).

Although the utilization of automated imaging and recognition systems for estimating plankton abundance promises to expand *in situ* observational capacity, the methodology has yet to be widely adopted for both scientific studies and monitoring programs. Several recent studies have been dedicated to comparing submerged instruments against traditional lab counting methods, but an important difference in those vs our study is that their image data was manually – not automatically – classified. Whitmore et al. (2019) explicitly compared the Zooglider's abundance estimates against MOCNESS net tows and acoustic data. Likewise, Sosik and Olson (2007) compared manual counts from the IFCB images to manual bench top counts.

Conversely, other related studies focused on validating the automated estimation of plankton abundance but did not seek to compare the results to traditional methods. Wang et al. (2017) suggested that an automated classifier's performance can be improved by attempting to match the training set class distribution to the eventual target population. González et al. (2019) proposed a number of automated quantification algorithms to improve plankton abundance estimates. Orenstein et al. (2020b) proposed similar methods to reduce human annotators' validation labor while reliably reproducing plankton distributions. However, the comparison of automated workflows that employ imaging paired with trained CNN classifiers with plankton population estimates that use the more traditional lab counting methods remains an interesting research question that has not been addressed.

Here, we quantify the relationship between plankton population estimates derived from an *in situ* imaging system, the Scripps Plankton Camera (SPC), with those obtained from concurrent bottle-based samples manually enumerated by a trained taxonomist. The SPC system, located at the Scripps Pier, consists of two underwater microscopes that image undisturbed volumes of water that can freely flow between a light source and a camera system. It has been operating nearly continuously for 6 years, resulting in the collection of more than a billion images of ROIs that includes plankton, detritus, sand, as well a host of other suspended microscopic inhabitants. Using data from the SPC microscopes, CNNs have been trained to sort the resulting data and speed up ecological analyses (Orenstein and Beijbom, 2017; Kenitz et al., 2020; Orenstein et al., 2020a; Orenstein et al., 2020b). The Scripps Pier is also a sampling location for the on-going Southern California Coastal Ocean Observing System (SCCOOS) HABMAP monitoring program (Kim et al., 2020) that has been enumerating HAB taxa from weekly water samples since 2008. The methodology employs hand-acquired water samples and a modern variant of the

Utermöhl method to count a variety of plankton and estimate the abundance of HAB formers (Utermöhl, 1931; Utermöhl, 1958; Karlson et al., 2010). Here, we reference those lab-based abundance estimates as the most widely accepted and traditional method that provides a baseline for comparing our automated methods that are based on automatically classified SPC data. If successful, the automated analysis workflow would provide an efficient, continuous monitoring system to detect and monitor phytoplankton and provide real-time, detailed, and reliable HAB warnings. The detection performance of both the imaging system itself and automated classification is evaluated in this study.

In this study, we compare the automated workflow for plankton count estimates obtained *via* CNN classification of the SPC images (SPC+CNN-Pier) to those derived by a plankton taxonomist counting hand-acquired, preserved samples under a microscope in the lab (Lab-micro). As a bridge between the two methods, a subsample of the hand-acquired bottle sample was imaged by a benchtop version of the SPC (SPC-Lab) and classified with an identically trained CNN (SPC+CNN-Lab). The complimentary analyses of images collected by SPC-Pier with the (Lab-micro) images allowed us to quantify the "effective" imaging volume of the SPC Lab and Pier systems. The complication arises as they employ a dark field method of illumination (Orenstein et al., 2020a) that we have found to produce optimal contrast to aid in identification. This leads to some ambiguities in the sampling volume. Another factor is that the orientation dependence of plankton may provide views that are hard to assign to a specific organism.

2 MATERIALS AND METHODS

Data for this study were obtained from three methods: (i) lab-based manual enumeration of collected water samples (Lab-micro), (ii) lab-based imagery of collected water samples (SPC-Lab), and (iii) imagery of plankton communities *in situ* (SPC-Pier). Water samples for lab-based analyses were collected from the Ellen Browning Scripps Memorial Pier in La Jolla, CA (32°52.02′N, 117°15.300′W) twice a week in the morning from May through October 2019. Five 2-liter bucket samples (total 10L) were collected from the surface at a depth of approximately 0.5 m. 2 L were then allocated for enumeration using traditional microscopy with the remaining 8 L imaged by the benchtop version of the SPC.

2.1 Traditional Microscopy Analysis: Lab-Micro

Plankton were enumerated using the Utermöhl method for quantitative phytoplankton analysis *via* the routine monitoring program carried out by SCCOOS, referred to as "Lab-micro" throughout this paper. Seawater was concentrated in sedimentation chambers after being fixed in a 4% formaldehyde solution prior to manual counting. Once the sample settles, the upper chamber is removed and replaced with a glass cover slip that is placed under an inverted microscope. Cells are then classified to the lowest possible taxonomic level at 200x magnification and counted by a human expert (Utermöhl, 1931; Utermöhl, 1958; Karlson et al., 2010). SCCOOS technicians typically examine the organisms from settling 10 or 50 mL of seawater. However, the sample volume enumerated here, ranged from 1.25 mL to 2.68 mL based on the abundance of phytoplankton. To account for the variation in settling volumes, we normalized the counts as the fraction of organisms that would have been observed if the volume was 1.76 mL volume. Although SCCOOS monitors a variety of species, here, we focus on the following 9 taxa: *Akashiwo sanguinea, Ceratium falcatiforme and fusus, Ceratium furca, Chattonella* spp., *Cochlodinium* spp., *Gyrodinium* spp., *Lingulodinium polyedra, Prorocentrum micans*, and *Pseudo-nitzschia* spp. as reported in absolute counts from the observed sample volume that was rescaled, if needed, to 1.76 ml. The input data from the Lab-micro system was therefore the number of counts of the organisms in the equivalent volume as a function of identified taxa and the date of collection.

2.2 Automated Imaging Systems: SPC-Pier and SPC-Lab

The SPC system is a set of two *in situ* underwater microscopes (Orenstein et al., 2020a). An onboard embedded computer identifies and segments out suspected plankton as Regions of Interest (ROIs). Here, two versions of the SPC-SPCP2 were used: (i) the SPC-Pier system, installed *in situ* at the Scripps Pier; and (ii) the SPC-Lab system – a lab-based version for benchtop imaging. The microscope uses a 5x objective to image a 2.5 mm x 2.5 mm field of view using dark field illumination that yields 40% contrast transmittance at 5.0 μm resolution with an image plane pixel size of 0.74 μm. Using both systems, ROIs (Regions of Interest) were selected that ranged between 40 μm and 120 μm in the maximal size dimension of the organism.

The SPC-Pier system was moored at a tidally dependent average depth of 3 meters (**Figure 1A**) and collected images at a rate of 8 frames per second throughout the study period, with a brief pause in September due to heavy biofouling. To enumerate "counts" an arbitrary temporal window of +/- 1000 seconds, yielding 16,000 images, was chosen for evaluation that was centered around the exact time of the hand-acquired sample.

The SPC-Lab is a reconfigured benchtop version of the SPC-Pier. To support the imaging of hand drawn samples, it was augmented with a gravity flow water system so that each 8L water sample passed through a clear acrylic chamber positioned in the field of view of the system (**Figure 1B**). The sample was put through the system at a constant flow rate by routinely replenishing the elevated water bucket with more seawater to maintain a minimum of 2 L of fluid. The flow system was flushed with filtered seawater between samples to prevent cross-contamination.

2.3 Species Selection and Manual Classification

To form a data set for comparing the observed image counts from the two SPC systems with those of the Lab-micro, a team of 3 taxonomists sorted all images collected by both SPCs into 10 classification categories, or classes: 9 taxonomy-based categories that captured each of the target organisms (**Figure 2**) in the 30 μm and 60 μm size range, and a category 'other' that included images of remaining organisms and particles imaged by the system. The 'other' class is necessary to give both the taxonomists

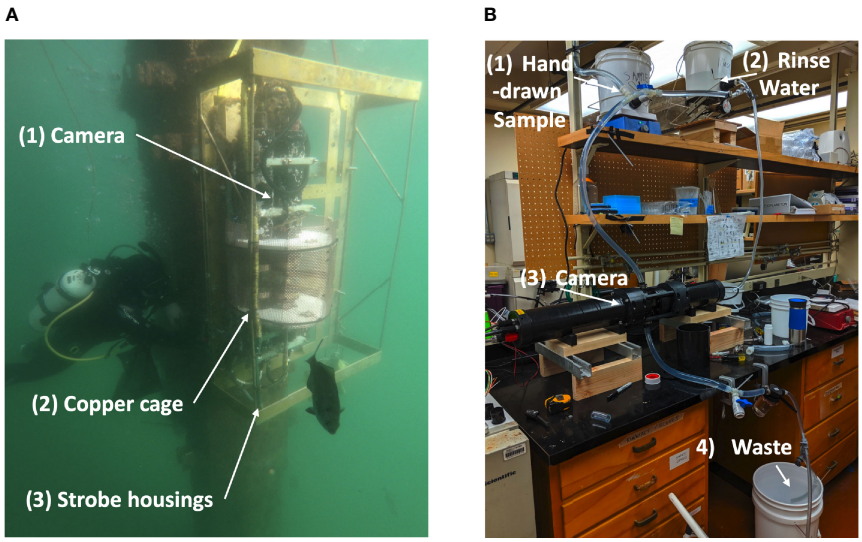


FIGURE 1 | The Imaging Systems. **(A)** SPC-Pier, SPC-MICRO Underwater Camera. **(B)** SPC-Lab, Benched laboratory configuration of SPC-MICRO.

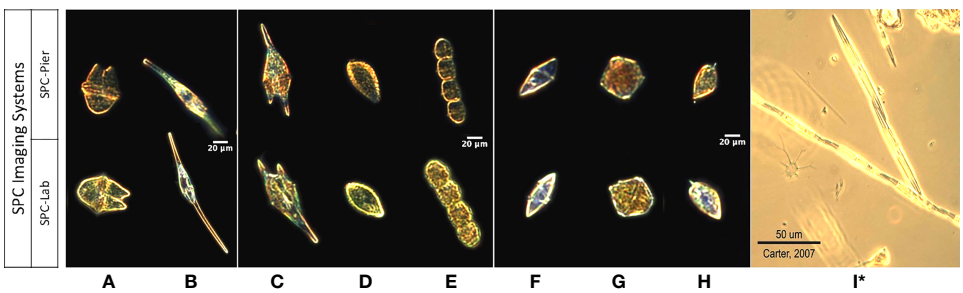


FIGURE 2 | Images of 9 taxa from the SPC-Pier, SPC-Lab systems. **(A–I)** *Akashiwo sanguinea*, *Ceratium furca*, *Chattonella* spp., *Cochlodinium* spp., *Gyrodinium* spp., *Lingulodinium polyedra*, *Prorocentrum micans*, and *Pseudo-nitzschia* spp. **(I)*** Is a lab microscopy photo of Pseudo-nitzschia sp. as the SPC imaging systems produced unsuitable images.

and the automated classifiers a place to put ambiguous objects and avoid high false-positive rates (Dhamija et al., 2018).

2.4 Sample Selection

Over the course of 5 months, 43 independent plankton samples were acquired *via* Lab-micro, SPC-Lab, and SPC-Pier. After a preliminary data analysis, a subset of 26 days were deemed suitable for analysis as there were complementary Lab-micro samples with suitable abundances. We note that these abundances were suitable if, at least, tens of organisms were sampled on a fraction of the days. Using the data from the 26 days, a data set consisting of the measurement of plankton "counts" using 5 methods (**Figure 3**) was assembled: (i) traditional microscopy counts provided by SCCOOS (Lab-micro), (ii) manual classification of (SPC-Lab), (iii) automated classification, using a CNN, of images collected by the SPC-lab system (SPC +CNN-Lab), and, similarly for images collected by the *in situ* SPC system, (iv) manual (SPC-Pier) and (v) automated (SPC+CNN-Pier) classification of images collected by the SPC-Pier.

2.5 Automated Imaging Classification Using Convolutional Neural Networks

To test the accuracy of automated image classification, we trained a collection of convolutional neural networks (SPC+CNN-Lab and SPC+CNN-Pier) and tested them on SPC-Lab and SPC-Pier images. The details of the implementation of the convolutional neural network methods are described below.

In training the Neural Networks we use the Residual Neural Network (He et al., 2015) architecture with 18 layers (ResNet-18). The relatively shallow network design is quick to train and less likely to overfit to the relatively small training sets we collected (Tetko et al., 1995). Network training followed standard practices in the machine learning literature, namely using stages of training, cross-validation, and testing (**Table 1**).

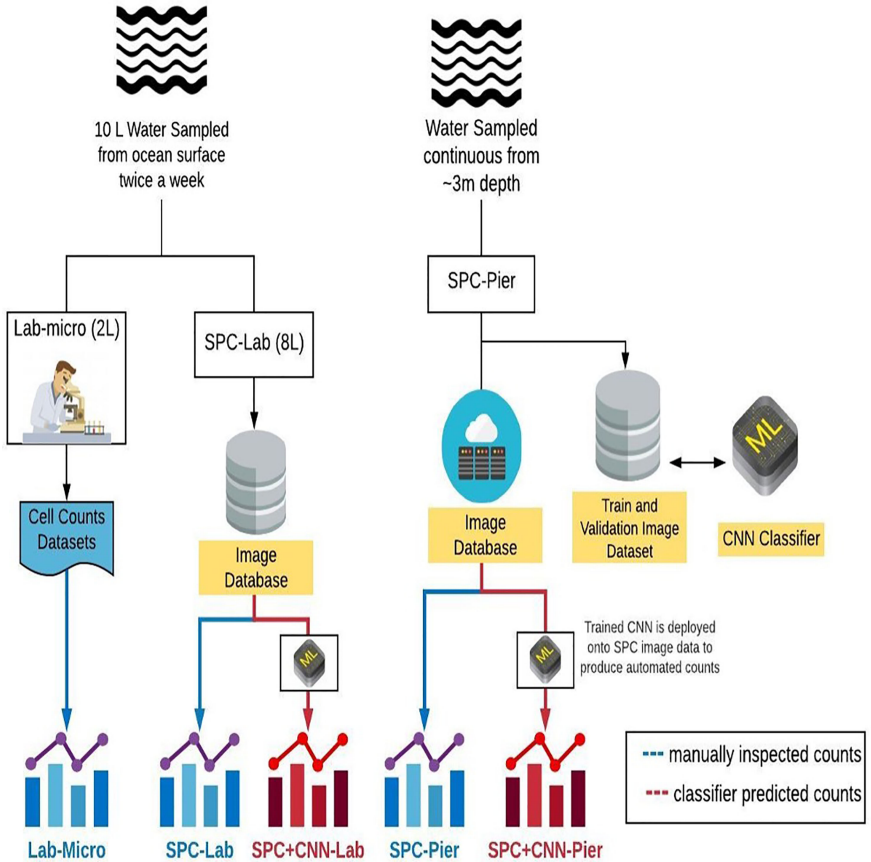


FIGURE 3 | A diagram that compares the sampling methods and the 5 resultant data sets: Lab-micro, SPC-Lab, SPC+CNN-Lab, SPC-Pier and SPC+CNN-Pier.

TABLE 1 | Overview of training, validation, and test datasets to train the SPC+CNN.

Dataset	Fine-tuning Stage	Data	# Classes	# Images
Phytoplankton-Train	1	Train	30	29,196
Phytoplankton-Val	1	Validation	30	19,773
SPC-Pier (n=25 dates)	2	Train	10	avg ~19,000
SPC-Pier (n=1 date)	2	Test	10	778
SPC-Lab (n=1 date)	2	Test	10	745

In all experiments, images were subject to random affine transformations – rotations and translations. This type of data augmentation enables the creation of additional training examples. Prior to the random affine transformations, images are padded into a square image and resized into 224 x 224 pixels. All networks were trained with the cross-entropy loss for 50 epochs. However, throughout each phase of the training procedure, the loss was weighted inversely proportionally to the class distribution of the corresponding training dataset, to mitigate potential class imbalance problems (Wang et al., 2017). Note this also includes recomputing the weight of the loss of each during cross-validation. Model weights that achieved the lowest loss on the validation set during training the 50 epochs were utilized.

The SPC+CNN for both lab and pier was trained in two fine tuning stages: (i) We fine-tuned a ResNet-18 pre-trained on the ImageNet database (Deng et al., 2009) with SPC phytoplankton images. (ii) The resulting network was again fine-tuned on just the ten classes of interest using images collected by either SPC-Pier or SPC-Lab. Fine tuning repurposes the parameters of a network trained for a particular task to a different target. The procedure reduces training time and improves accuracy when training with small datasets (Yosinski et al., 2014). Double fine tuning further adapts each network to subtle differences between the SPC-Pier and SPC-Lab data after learning more general representations of plankton (Orenstein and Beijbom, 2017).

The first fine-tuning step uses a labeled phytoplankton training set from the SPC-Pier system that comprised of 37,147 images

spanning 51 classes (Kenitz et al., 2022). This dataset was produced by 15 expert taxonomists and 5 non-taxonomists from the US West Coast during a two-day workshop whose main goal was to collect images for training the CNNs. This workshop dataset came from an earlier portion of the SPC-Pier time series and has no temporal overlap with the images acquired in our experiment. Experts sorted the annotated images into 44 taxonomic classes and 7 noise categories, which included the 9 species of interest. The workshop dataset was adjusted by combining categories of the same species tagged with semantic descriptors such as the number of cells (e.g., Ceratium furca pair vs. single) and eliminating categories with fewer than 300 images. This resulted in a total of 30 classes, 24 identifiable species and 6 noise categories. 80% of the 36,496 images were then randomly chosen for training (Phytoplankton-Train) and the remaining 20% used for validation (Phytoplankton-Val) (**Table 1**).

The second fine-tuning step had two objectives: (i) force the network to recognize only the 9 species of interest and the background class 'other' of our study; and (ii) account for dataset shift, the well-known property of classifiers to be sensitive to changes in the input data, both the appearance of the images and the relative distribution of the classes between training and testing (Moreno-Torres et al., 2012; He et al., 2015; González et al., 2019; Orenstein et al., 2020b). In this step, the classifier is fine-tuned to the collected SPC-Pier dataset, which was partitioned in a leave one-out cross-validation manner for training and testing. Specifically, the model is trained on data from all dates from the SPC-Pier except for one, which is used as a held-out test set (**Table 1**). The same procedure is repeated several times with each sampled date being used as a held-out set once, and performance metrics are averaged across all 26 days. The training sets for each cross-validation iteration contain approximately 39,000 images, and test sets respectively holding out 745 and 778 for the SPC-Pier and SPC-Lab.

In implementing the first stage, the base ResNet-18 model pretrained on ImageNet was fine-tuned for 50 epochs on the 30-class phytoplankton taxonomy workshop dataset. Model weights that achieved the lowest loss on the validation set during the 50 epochs were utilized. In this stage, the model achieved an accuracy of 95.5% on the Phytoplankton-Train set and accuracy of 95.2% on the Phytoplankton-Val set. The second stage was initialized with the model weights learned in the first stage, where the final layer was replaced with a layer of 10 outputs (9 categories of interest plus Other). Fine-tuning to the leave-one-out cross-validation training datasets was performed for an additional 50 epochs with model weight selection corresponding to the lowest training loss. This resulted in a collection of 26 trained models, where each model is tested on an independent date from the SPC-Pier and SPC-Lab dataset.

All models were trained with an initial learning rate of 0.001 and a batch size of 16 using the Adam optimizer (Kingma and Ba, 2014). Models were trained on an NVIDIA Titan Xp GPU. Python code used to train and evaluate the models is available at https://github.com/hab-spc/hab-ml.

There are several examples of dataset shift between our training sets, notably the slight variations in illumination between images captured by the SPC-Pier and SPC-Lab systems (**Figure 2**). The restriction of fine-tuning to only the SPC-Pier image dataset is specifically designed to examine the potential effects of dataset shift when the classifier is deployed on a new target domain, in our case the SPC-Lab. Training on SPC-Pier and testing on SPC-Lab data is a proxy for the more general transfer of a classifier trained on an *in-situ* imaging system to an *in vitro* imaging system.

2.6 Analyses of the Three Sampling Methods

To compare the three sampling methods, we used the total number (counts) of organism identification for each of the 10 categories collected on each of the 26 independent days. Given the species-specific counts, we performed 1) an assessment of the classifier performance and 2) a comparison between the Lab-micro counts and SPC+CNN counts. A comparison between the Lab-micro counts and manually enumerated SPC counts is also included to establish a baseline unaffected by CNN classifier errors. Although relative abundance is a widely used measure of plankton distributions, we use the number of counts of each species as a function of date, for two reasons: (1) Comparisons of relative abundance are sensitive to numerical instability caused by frequent counts of 0 or 1. (2) The effective interrogation volume of the SPC systems varies from species to species, due to both the focus dependent darkfield imaging as well as the effects of random orientations. As such, an important aspect of our work is the estimation of an "effective sampling volume" for each species as elaborated below.

2.6.1 Volume Computation Analyses

In all experiments, the volume $V_{\text{Lab-micro}}$ of water used by lab procedure was standardized to 1.76 mL, while the counts of the SPC-Pier system were integrated over 2000 seconds of images taken at 8 Hz, resulting in 16,000 images. Under the assumption that the concentration of species counted by each method is the same,

$$\frac{C_{\text{SPC+CNN}}}{V_{\text{SPC+CNN}}} = \frac{C_{\text{Lab-micro}}}{V_{\text{Lab-Micro}}}, \qquad \text{Equation 1}$$

where $V_{\text{SPC-CNN}}$ is the effective volume imaged by the SPC system and C denotes counts. Now, defining the ratio α between the two volumes as

$$V_{\text{SPC+CNN}} = \alpha V_{\text{Lab-micro}} \qquad \text{Equation 2}$$

leads to the linear relationship

$$C_{\text{SPC+CNN}} = \alpha C_{\text{Lab-micro}}, \qquad \text{Equation 3}$$

between SPC+CNN and Lab-micro raw counts. This was the model used to relate the SPC+CNN counts of both the Pier and lab implementations to Lab-micro counts in our study. The scaling factor α was estimated by computing a linear regression between each pair of counting methods.

2.6.2 Counting Analyses

Counts are compared across the 3 sampling methods, for both manually enumerated SPC and automated SPC+CNN counts. A separate model is fit for each of the 9 species using the linear regression model of (3) across all 3 pairs: SPC+CNN-Lab vs. Lab-

micro; SPC+CNN-Pier vs. Lab-micro; and SPC+CNN-Pier vs. SPC+CNN-Lab. The model of (3) was fit with a linear-linear least squares estimator, assuming zero intercept. However, for display purposes, the data was transformed to a log scale. Given the computed linear regression between each pair of counting methods, quantitative comparisons are obtained by computing the Pearson correlation coefficient, a measure of linear correlation between two variables, and the percentage $R^2$of the variance explained by the model relative to the total variance. In conjunction with the factor α, these measurements express how related the counting methods are.

2.6.3 Classification Analyses

Our collection of double fine-tuned classifiers is applied to the 26 daily test sets from which 21,211 images were extracted from the SPC-Lab and 20,148 images from the SPC-Pier that were then manually classified into the 10 categories. CNN Performance results are then averaged across the test sets. Classification performance is assessed by 1) accuracy (ACC), the fraction of correct predictions, 2) mean class accuracy (MCA), the average correct predictions over each individual class, and 3) the F1 score, a commonly used metric for scoring class-imbalanced problems. Together, these metrics capture both model generalization ability and bias towards highly populated classes – ACC characterizes the overall classifier performance while MCA and F1 scores assess how well the system does on a per-class basis. Significant differences between the three metrics indicate that a method favors common classes while underperforming for rare ones.

3 RESULTS

Given the 26 independent samples, the datasets were largely dominated by the 'other' category (83% of the SPC-Pier total and 92% of the SPC-Lab total). The resulting manual counts are denoted as SPC counts. CNN-produced counts on the same dataset are denoted SPC+CNN counts. Lab-micro counts were produced by a biologist, using traditional microscopy.

3.1 Analysis of the Neural Net Results

In general, Lab-micro collected more total counts of the 9 target species, over the set of images, than the SPC systems (**Figure 4A**). Averaged over all 26 independent samples, Lab-micro count data was predominantly composed of 3 common species: *Pseudo-nitzschia* spp., *Lingulodinium polyedra*, and *Prorocentrum micans* (**Figure 4B**). The latter two also dominated SPC-Lab and SPC-Pier counts. However, in the case of the SPCs, the *Pseudo-nitzschia* spp. counts were notably fewer. Although there is some uncertainty in the inability of the SPCs to reliably detect the *Pseudo-nitzschia* spp., we suspect that it is likely because the thickness of this pennate diatom is close to the resolution limit of the system as well as the fact that a needle like structure, when subject to a uniformly random 2-dimensional view will be difficult to see in many of its orientations. The remaining taxa of interest, namely *Akashiwo sanguinea*, *Ceratium falcatiforme or C. fusus*, *Ceratium furca*, *Chattonella* spp., *Cochlodinium* spp., and *Gyrodinium* spp., were more often observed by the SPCs than the Lab-micro suggesting that the methodology has some taxonomic dependence.

Inspection of the confusion matrices for the CNN performance of SPC-Pier versus SPC-Lab (**Figure 5A**) confirmed that the CNN performed significantly better on the SPC-Pier than on the SPC-Lab data, as expected from the MCA and F1 score difference shown in **Table 2**. For half of the tested species (*Akashiwo sanguinea*, *Ceratium furca*, *Cochlodinium* spp., *Lingulodinium polyedra*, *Prorocentrum micans*) the accuracy dropped more than 10% from SPC -Pier to SPC-Lab, especially *Lingulodinium polyedra* (**Figure 5B**). This is a

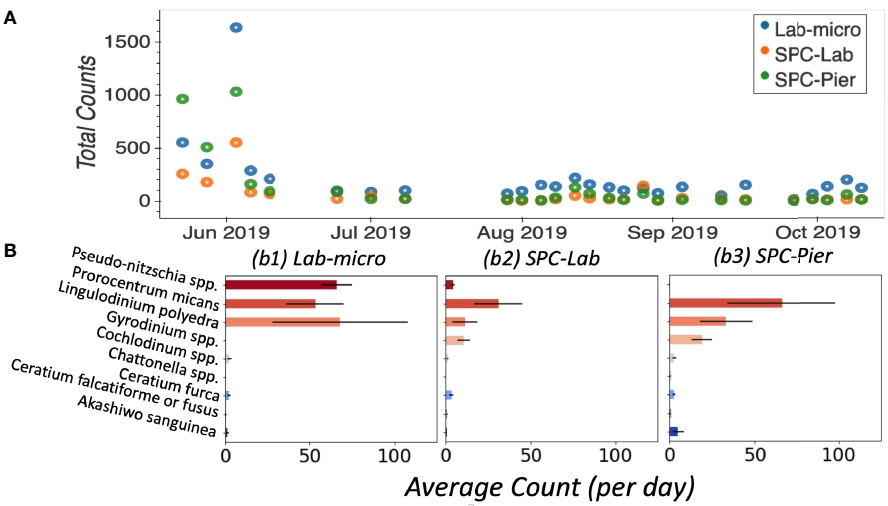


FIGURE 4 | Enumerated plankton taxa. **(A)** Time series of total counts as obtained by traditional methods (LAB-micro) and manual image classification of lab samples (SPC-Lab) and *in situ* (SPC-Pier). **(B)** Average count per day per species collected by each method.

manifestation of the domain shift between the SPC-Pier and SPC-Lab imaging methods: the lab flow-through system appeared to result in more out-of-focus images that rendered the species differentiation more difficult. This was not unexpected, given that the model is only trained on SPC-Pier data, however, it does illustrate that deployment of the same imaging system can vary, likely due to difference in lighting and any orientation effects that are due to flow.

3.2 Classification Performance and Comparison of the Lab Micro vs SPC+CNN

Results indicate that the CNN achieved averaged test accuracies of 92% on both the SPC-Lab and SPC-Pier data (**Table 2**). The averaged ACC, MCA, and F1 Score performance was measured for a CNN tested on independent samples from the 26 SPC-Pier and SPC-Lab image datasets. The MCAs were lower (68 and 74%) suggesting an unbalanced performance across classes. This discrepancy between the metrics is originated by class population imbalance, due to the fact that some species were observed relatively rarely under the SPC setting (e.g. *Ceratium falcatiforme* or *C. fusus*, *Chattonella* spp., and *Pseudo-nitzschia* spp.). This results in less training data to effectively learn the species' morphology. The F1 scores were the lowest of the three (.47 and.64), due to the CNNs' frequent overestimation of the count of HAB species, which is penalized in the F1 score for poor precision. These results show that the CNN performs with high accuracy for the classes that are relatively abundant in the training data. Class imbalance in the training dataset can have a large effect on the learned model and is a well-established feature of training CNNs on natural populations.

The Pearson correlation analysis on the intermediary comparison of the Lab-micro and manually enumerated SPC counts (**Figure 6**) reveals high-to-very high correlations between the sampling methods on 4 out of the 9 species – *Akashiwo sanguinea*, *Cochlodinium* spp., *Lingulodinium polyedra*, and *Prorocentrum micans* – representing a mix of abundant and rare organisms (**Figure 6A**). The comparison for *Ceratium furca* revealed moderate correlations between both SPC methods and the Lab-micro (0.58 and 0.70). The other 4 species, *Ceratium*

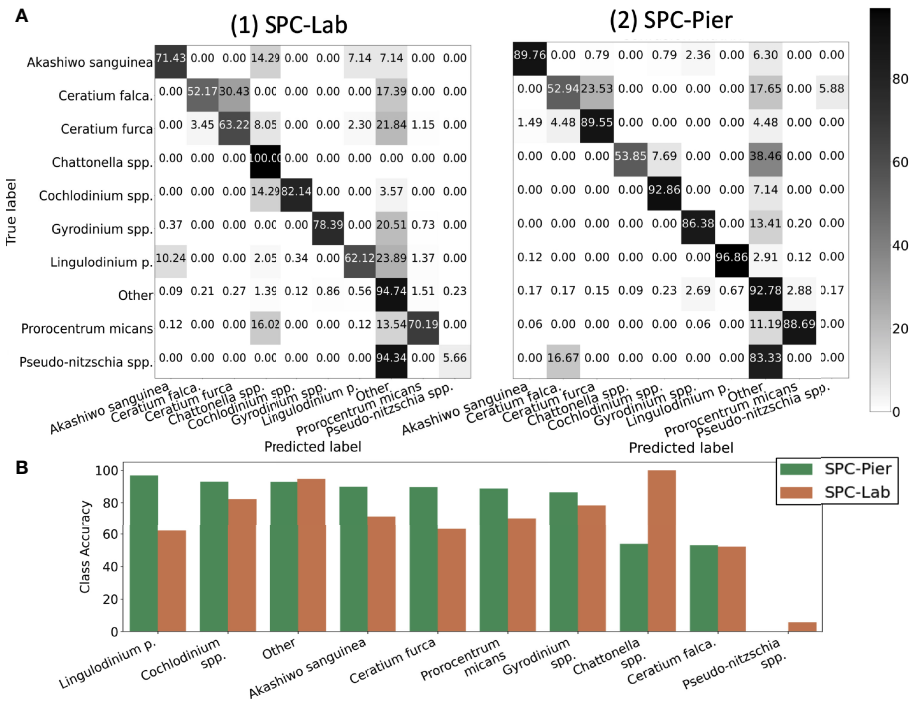


FIGURE 5 | Quantification of the classification accuracy for SPC test sets. **(A)** Confusion Matrix. **(B)** Diagonal class accuracies of confusion matrix sorted in a descending fashion from left to right.

TABLE 2 | Average classification results of a double fined-tuned model tested on independent held out samples collected by the SPC-Pier and SPC-Lab.

Dataset	ACC	MCA	F1 Score
SPC-Lab	0.92	0.68	0.47
SPC-Pier	0.92	0.74	0.64

Evaluation metrics used are accuracy (ACC), mean class accuracy (MCA), F1 Score.

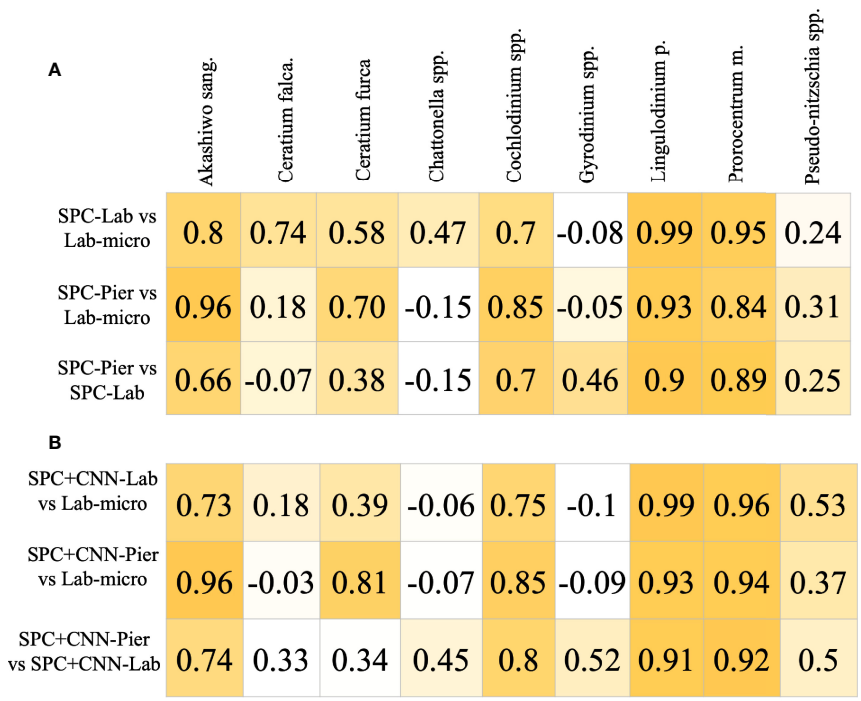


FIGURE 6 | Pearson Corrélation Coefficient Matrices. Each row compares two of the resultant data and/or CNN estimation of taxonomic presence. Each column is a corresponding species. Coefficient values are color coded with respect to the species correlation value of the compared setting, in an ascending fashion. **(A)** Correlation of Lab-micro vs. manually enumerated SPC counts. **(B)** Correlation of Lab-micro vs. SPC+CNN counts.

falcatiforme or *C. fusus*, *Chattonella* spp., *Gyrodinium* spp., and *Pseudo-nitzschia* spp., demonstrated a pattern of low correlation scores scoring correlation scores for two out of the three pairs.

In general, the SPC+CNN *vs*. Lab-micro correlations produced similar results to the baseline correlation values between the manually enumerated SPC vs. Lab-micro counts (**Figure 6B**). The same 4 species that previously produced high-to-very high correlations were consistent when using SPC+CNN counts, with correlation value differences up to 10%. The correlation differences were due to the previously mentioned unbalanced performances across the classes from the SPC+CNN, that arises from using imbalanced training data. In the case of the SPC+CNN-Lab vs. Lab-micro, we observed that many correlation scores dropped, which can be attributed to the domain-shift problem.

Figures 7, **8** display the linear fit between the enumerated counts for each of the sampling methods across the various taxa as computed by the regression model across the 3 possible data sources (Lab-micro, SPC+CNN-Lab and SPC+CNN-Pier). As can be seen, the proportionality approximation conveys that the SPC +CNN-Pier's sampling of an aggregate volume over the 2000 seconds recorded nearly twice the number of images of the SPC +CNN-Lab. In addition, a majority of the five species showed non-existing-to-poor linear relationships between the Lab-micro and SPC+CNN counts. The linear fit for the *Pseudo-nitzschia* spp. showed little ability to model the relationship between the SPC +CNN and Lab-micro, as the SPCs detected the species poorly. *Gyrodinium* spp. were mostly absent from the Lab-micro, preventing a comparison *via* linear regression between the sampling methods. Species that had previously demonstrated low classification performance resulted in poorer relationships when computing the linear regression for the CNN-based pairs of counting methods. Compared to the manually enumerated-based linear regressions, *Ceratium falcatiforme* or *C. fusus*, and *Chattonella* spp. showed small R^2 values and fit to the slopes across all 3 possible pairs, suggesting that, possibly, poor classification performance negatively impacted the linearly modeled relationships. *Ceratium furca* also showed some fluctuations when comparing automated vs manual regressions, but generally showed only a lack of a linear relationship between the two data generation methods (**Figure 7**). **Figure 8** shows the other 4 species where, we note, *Akashiwo sanguinea* and *Cochlodinium* spp. demonstrated a poor fit to the linear correlation while the *L. polyedra* and the *P. micans* were quite good with R^2 scores of (0.97, 0.89). We note that these two species had the highest number of counts across all three sampling methods and, conjecturally, the highest concentrations.

As shown in **Figure 8**, in a manner like the results of the Pearson correlation analysis for the pair of SPC+CNN-Pier vs. Lab-micro, we found high R^2 values for two of the less-abundant species (*Akashiwo sanguinea*, *Cochlodinium* spp.), and two of the more-abundant species (*Lingulodinium polyedra*, *Prorocentrum micans*). We also observed that the sizes of the prediction and confidence bands were related to the frequency of occurrence of

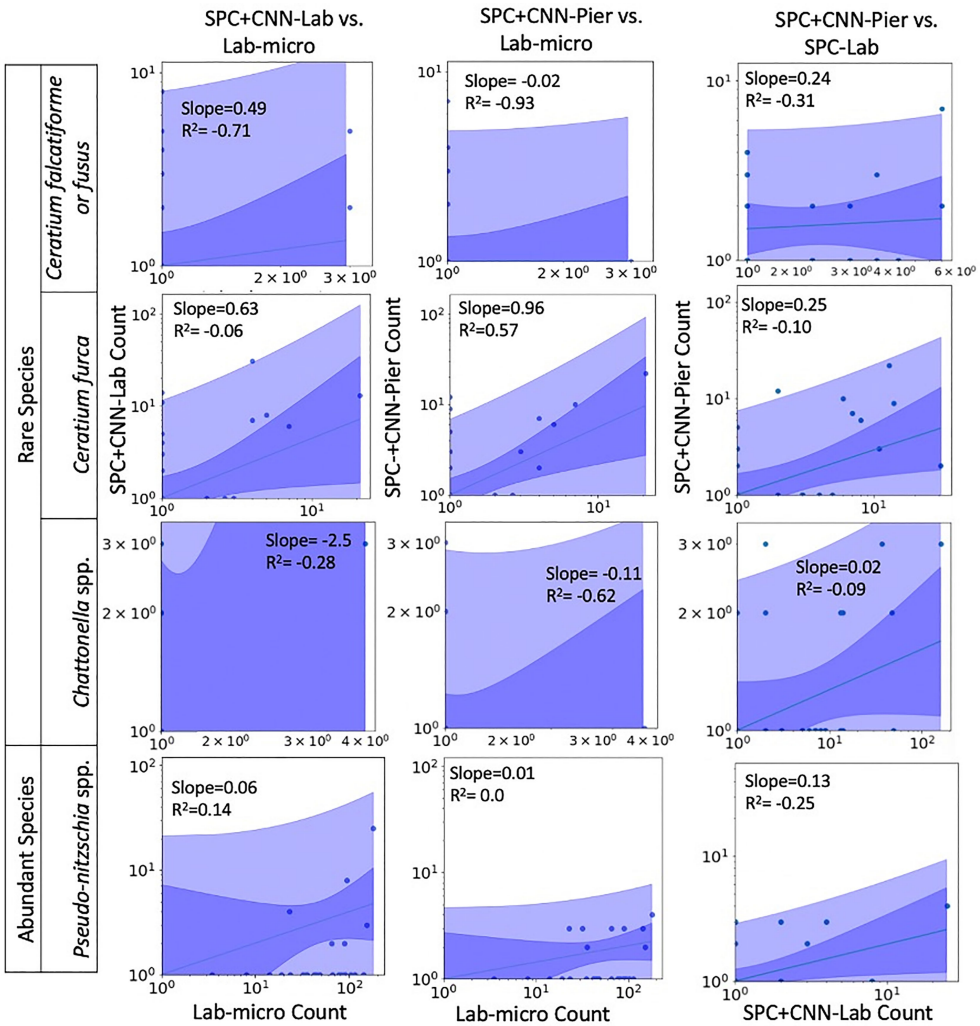


FIGURE 7 | Relationships between counts of Lab-micro and SPC+CNN methods (less abundant species). Columns highlight pairs of counting methods, rows demarcate species. The solid line indicates a linear regression model that is coupled with multiple shaded areas indicating the 95% prediction (dark shade) and confidence interval (light shade). The slope and R^2 of the model fit are indicated.

the species. The two abundant species showed much narrower prediction and confidence bands, in contrast to the two rare species, which exhibited wider bands. Discrepancy of the size of the bands could be due to the low cell counts of the relatively rare species.

3.3 Volume Computation

An important feature of this work is the computation of the "effective sampling volume" for the SPC+CNN results. This then permits the estimate of abundance. Considering the most abundant and highly correlated species (*Lingulodinium polyedra* and *Prorocentrum micans*) equation (3) can be used to compute this volume using the slope of the fit as shown in **Figure 7**. Given that this slope is (0.39, 2.02) for (*L. polyedra*, *P. micans*) and that the reported Lab-micro samples a 1.76 mL volume, our cumulative sampling volume for 2000 seconds of images at 8 Hz is (0.69, 3.56) mL. Then, the "effective sampling volume" per image is estimated as (0.043, 0.22) µL after dividing by the 16000 frames. We note that the R^2 values for the other 4 categories were too low to be considered and are therefore not reported.

3.3 Continuous Observation Data

One major advantage of *in situ* microscopes like the SPC-Pier system is that they can observe plankton continuously in time. This permits post processing with a variable integration time to compute species dependent total counts. In this study, we used a 2000 second integration window that provided 16,000 image samples (at 8 Hz) that occurred over the period from the end of May until October 2019 (**Figure 9**). Here, the continuous grey line indicates counts of the 4 species that were most confidently estimated from the SPC+CNN-Pier during both the lab sampling

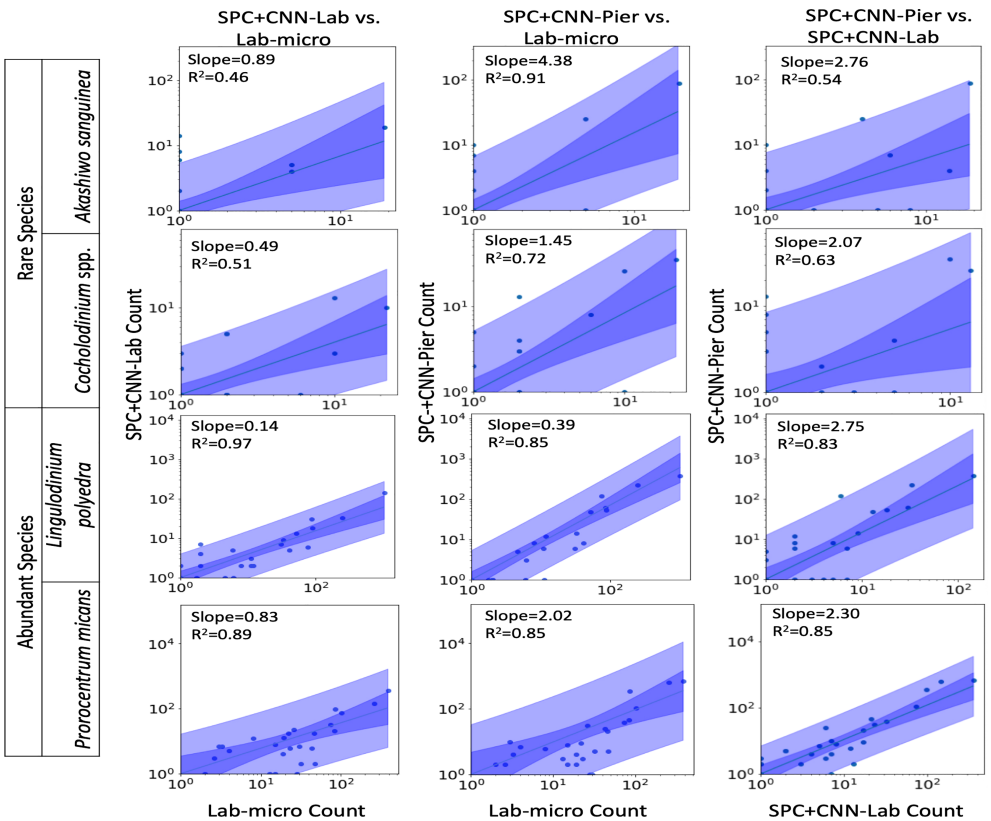


FIGURE 8 | Relationships between counts of Lab-micro and SPC+CNN methods. Columns highlight pairs of counting methods, rows demarcate species. The solid line indicates a linear regression model that is coupled with multiple shaded areas indicating the 95% prediction (dark shade) and confidence interval (light shade). The slope and R^2 of the model fit are indicated. Note that data is displayed logarithmically but was fit linearly.

occurrences as well as other times where there were no manually collected samples. We note that there was an increase in the *Akashiwo sanguinea* and *Cochlodinium* spp. during and in-between the lab samples as well as the absence of increased abundance for the *Lingulodinium poleydra* as well as the *Prorocentrum micans* that were not observed by the Lab-micro sampling as it was less frequent. The results highlight the advantages of continuous sampling that is facilitated by an *in situ* instrument. Moreover, the agreement when both the Lab-micro and the SPC+CNN-Pier data were available provides support to interpret the SPC+CNN-Pier system as valid, with, naturally, some error bound.

4 DISCUSSION

In recounting the goals of the work reported here, we first sought to explore the ability of CNNs to correctly classify the images that were recorded from the SPC systems. Although lab-based identification of the phytoplankton species is well established, the correspondence between the traditional methods and our dark-field microscopes had not been established. In examining the potential differences between the two methods, Lab-micro vs. SPC+CNN, there are several factors to consider: the samples observed by the SPC microscopes experience range-dependent defocus that is a necessary consequence of the dark-field illumination. In addition, since the SPC microscopes image organisms that are freely drifting in the field of view of the system, a natural assumption is that their orientation, relative to the viewpoint of the camera, is uniformly distributed. In contrast to larger zooplankton, such as copepods, our organisms of interest have fewer morphological differences that are also confounded by the aspect-dependent views acquired. This makes the identification more difficult for automated systems as well as taxonomists viewing the resultant SPC images.

In considering the success of the CNNs to classify the species present in the images, we found that the imbalanced nature of datasets significantly influenced the performance of the system. Class imbalance is a well-studied problem that exists in many real-world ocean ecosystem datasets (e.g. WHOI-plankton: (Orenstein et al., 2015), EILAT and RAMAS coral dataset (Shihavuddin, 2017) in which rare species have far fewer images than abundant species). To combat this problem, we applied transfer learning from a less-imbalanced and filtered dataset to a more-imbalanced and unfiltered one. We also applied cost-sensitive learning, one of the techniques

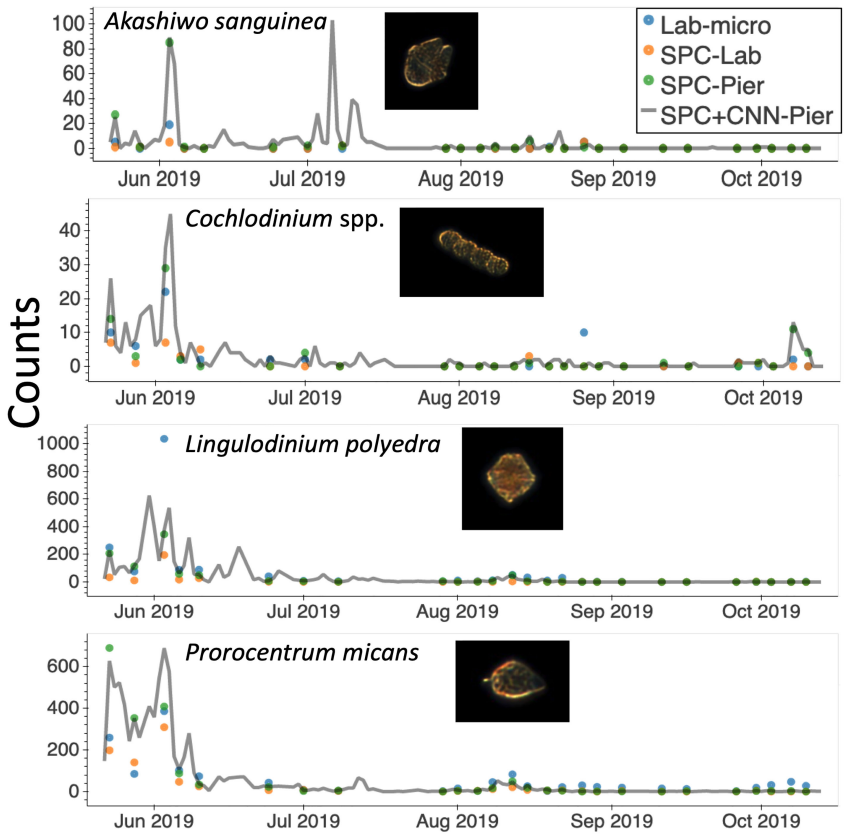


FIGURE 9 | A time series of species presence *via* "counts" or number of observations by the SPC-pier and the SCCOOS monitoring program during 2019. Automated image classification was used to produce counts on continuous periods. Most of which were not sampled by the SCCOOS program. Plots are shown for only the highly correlated abundant and rare species.

commonly used to improve the performance of class imbalanced classification (Wang et al., 2017). However, the comparatively low performance on rare classes suggests a limited capability of our techniques to mitigate the class imbalance problem. For model improvement, it would be worth experimenting with other methods, such as an ensemble of CNN models (Lumini and Nanni, 2019) or applying transfer learning by pre-training with class-normalized data (Lee et al., 2016). Class imbalance can also be addressed by collecting data over an extended period, especially days with significant presence of the organisms from the classes under-populated in our training set. This is left for future studies.

Compared to the class imbalance problem of the SPC+CNN, domain shift is less discussed in deep learning applications in the ecological literature. However, our results suggest that this problem deserves critical consideration when deep learning systems are to be deployed in an environment different from that used for training. Many zooplankton detection systems, such as ZooplanktoNet (Dai et al., 2016) and Zooglider (Whitmore et al., 2019), did not explicitly address and investigate their deep learning models' capability to transfer across domains. When trained purely on SPC-Pier image data, our model was not able to replicate its high performance to the SPC-Lab data, showing noticeably lower-class accuracies (for example for *Prorocentrum micans* or *Lingulodinium polyedra*) relative to the SPC-Pier. In future research, experimenting with other domain adaptation techniques, such as similarity learning (Pinheiro, 2018), or image-to-image translation (Murez et al., 2018), can help further improve our model. Solving the domain shift problem is essential to ensuring the reliability of deep learning automated systems in different environments.

Considering the nine species, or classification categories, investigated here, the significant correlation between the Lab-micro counts and the SPC+CNN-Pier data for *Prorocentrum micans* and *Lingulodinium polyedra* indicates that, under the environmental and lab identification procedures developed here, the *in situ* system counts can be transformed into estimates of concentration that are consistent with traditional microscopy observations. These correlations were also consistent when using the manually enumerated SPC counts instead of the SPC+CNN. The use of the multiplicative scaling factor α in our volume computation analysis mitigates these effects.

Both the SPC+CNN methods and Lab-micro show gaps in their ability to detect certain species. Firstly, Lab-micro only detected *Gyrodinium* spp. on one day, while both SPC methods

detected it on more than 20 of the 26 days. This is, presumably, due to the formaldehyde treatment that leads to dissolution and subsequent misidentification of "naked" species like *Gyrodinium* spp. (Costas et al., 1995). On the other hand, the SPC methods have problems detecting *Pseudo-nitzschia* spp. Whether this is due to the inefficiency of the darkfield imaging technique or, rather, effects related to their chain-like structure when viewed in 3D is unknown. We do note, however, that there may be some advantages to observing settled samples.

The other major goal of this research was to estimate the "effective sampling volumes" so that abundance could be estimated from the SPC+CNN-Pier data. Here, we note that, as reported on the web site, (spc.ucsd.edu) the SPC2 camera used here has a "high-resolution image volume" of 0.1 μL and a "Blob detection volume" of 10 μL. The sample volumes reported in **Table 3** of 0.043 μL and 0.22 μL for *Lingulodinium polyedra* and *Prorocentrum micans*, respectively, for the SPC+CNN-Pier are not inconsistent, likely due to the system's single view angle resulting in ambiguities that prevent the unique identification of the species. We also note that in comparing the SPC+CNN-Lab values vs Lab-micro, the proportionalities indicate that the lab system detected approximately half of those detected by the SPC +CNN-Pier. The discrepancy may be because the SPC-Lab samples were taken from the near-surface of the ocean (~ 0.5 m), whereas the SPC-Pier samples from a tidally dependent depth of 3 m. The differences may also arise from orientation-dependent effects that result from the water flowing past the SPC-Lab, or differences in the two optical systems, such as illumination intensity. Less-abundant species (e.g., *Akashiwo sanguinea* and *Cochlodinium* spp.) had reasonable fits between the SPC-Pier and the Lab-micro, with the SPC-Pier having a larger slope and hence, a larger estimated sampling volume. However, the uncertainty of these values is higher due to the small number of samples.

A distinguishing feature of this analysis is that the "effective sampling volumes" as computed *via* comparison with the Lab-micro calibrations are different for each species (e.g., *Lingulodinium polyedra* and *Prorocentrum micans*). These differences in estimated sampling volumes were not entirely unanticipated, as our dark-field illumination setup acquires an orientation-dependent image of these organisms, causing CNNs and expert taxonomists to be less capable of determining the exact identity of each species. Consequently, our linear fit for each of the species has a different slope, leading to different effective sampling volumes that are species dependent.

An important aspect of *in situ* sampling is that it is capable of detecting organisms on a 24/7 basis: the *in situ* microscope can provide continuous, real-time sampling during periods when there was no manual data collection (**Figure 9**). The period from May to October 2019 provided roughly 128k images *via* the automated sampling. The values obtained for *Lingulodinium polyedra* and *Prorocentrum micans* study showed realistic abundance increases and decreases of both, that occurred before and after a detected bloom. Rarer taxa, such as *Akashiwo sanguinea* and *Cochlodinium* spp., showed similar trends, but increases in abundance recorded by the imaging system were missed by the manual sample collection. Although a more detailed analysis would be needed to estimate the confidence in these observations, it seems that these transient changes in abundance were simply undetected because of the less frequent sampling by the Lab-micro. This, in turn, highlights the need for real-time continuous monitoring with less human effort. Furthermore, the low counts generated by the SPC systems between July and October indicated that there were no significant blooms during that time. Given the continuous nature of the SPC data stream, a set of algorithms could be implemented to deploy adaptive sampling that would improve the dynamic range of lab quantification.

One advantage of systems like SPC+CNN that produce real-time data is their potential for use as an early detection system. Data-driven insights would then inform decision making in monitoring programs, such as SCCOOS, for which shore station leaders have limited information on the daily abundance level of the HAB species. For example, previous studies show that it can be advantageous to know the initial and final periods of a bloom (Stroming et al., 2020). Stroming et al. (2020) showed the socioeconomic benefit of early HAB detection and estimated a saving of $370,000 following the early warning of a 2017 cyanoHAB event in Utah Lake. Given the statistically robust signals found in the present study for estimating HAB abundances, the recommended next steps would be to explore the use of the SPC for supporting decision-making in such settings.

5 COMMENTS AND RECOMMENDATIONS

The SPC+CNN workflow has shown its capability to provide real-time, high accuracy detection of certain HABs species, such as *Akashiwo sanguinea Cochlodinium* spp., *Lingulodiniumpolyedra* and *Prorocentrum micans*. Although its performance is species-dependent, it has shown a high correlation with the Lab-micro counts in certain cases. Moreover, this automated workflow can detect rare species more frequently than the manual method. It also minimizes manual labor and can provide continuous sampling at a high spatial and temporal resolution. All of these benefits make the

TABLE 3 | Calibrated SPC+CNN-Pier Sampling Volume Per Image.

Species	Proportionality (α)	Volume In-Focus (mL)
Lingulodinium polyedra	0.39	4.29×10^{-5}
Prorocentrum micans	2.02	2.22×10^{-4}

This calibration assumes a sampled Lab-micro volume of 1.76 mL and 2000 seconds of images at 8 Hz that were collected and classified for each species by the SPC+CNN-Pier.

SPC+CNN a potentially important tool with the capability to advance the study of imaging, recognition, and monitoring of HAB-related phytoplankton. The results suggest that image-based monitoring systems, supported by high-throughput automated classifiers, can be a reliable alternative to time-consuming manual sampling campaigns. Moreover, our experimental techniques and analyses provide a framework for future intercalibration studies of innovative new plankton sampling modalities.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

KTL conceived the ideas and designed the methodology along with JSJ; ECO, MLC, KMK and SS collected the data; DR built the lab imaging instrument; ZY and PM trained the classifiers; KL, ZY, and AS analyzed the data; KL, JJ, NV, ECO, and PJSF led the writing of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors are grateful to the National Science Foundation for funding this research under Grant #1546351 and Grant # 1546305 that was awarded from the Div of Information and Intelligent Systems.

REFERENCES

Beers, J. R., and Stewart, G. L. (1970). The Preservation of Acantharians in Fixed Plankton Samples1. *Limnol Oceanog.* 15 (5), 825–827. doi: 10.4319/lo.1970.15.5.0825

Castellani, C. (2010). Plankton: A Guide to Their Ecology and Monitoring for Water Quality. *J. Plankt. Res.* 32 (2), 261–262. doi: 10.1093/plankt/fbp102

Costas, E., Zardoya, R., Bautista, J., Garrido, A., Rojo, C., and López-Rodas, V. (1995). Morphospecies Vs. Genospecies in Toxic Marine Dinoflagellates: An Analysis of Gymnodznium Catenatum/Gyrodinium Impudicum and Alexandrium Minutum/a. Lusitanicum Using Antibodies, Lectins, and Gene Sequences1. *J. Phycol.* 31 (5), 801–807. doi: 10.1111/j.0022-3646.1995.00801.x

Cowen, R., Greer, A., Guigand, C., Hare, J., Richardson, D., and Walsh, H. (2013). Evaluation of the *In Situ* Ichthyoplankton Imaging System (Isiis): Comparison With the Traditional (Bongo Net) Sampler. *Fish. Bull.* 111, 1–12. doi: 10.7755/FB.111.1.1

Culverhouse, P. F., Macleod, N., Williams, R., Benfield, M. C., Lopes, R. M., and Picheral, M. (2014). An Empirical Assessment of the Consistency of Taxonomic Identifications. *Mar. Biol. Res.* 10 (1), 73–84. doi: 10.1080/17451000.2013.810762

Dai, J., Ruchen Wang, H. Z., Guangrong, J., and Xiaoyan, Q. (2016). "Zooplanktonet: Deep Convolutional Network for Zooplankton Classification." in *Oceans 2016* (Shanghai: IEEE), 1–6.

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). "Imagenet: A Large-Scale Hierarchical Image Database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255.

Dhamija, A. R., Günther, M., and Boult, T. E. (2018). Reducing Network Agnostophobia. *Adv. Neutr. Inform. Proc. Syst.* 31, 1811.04110. doi: 10.48550/arXiv.1811.04110

Ellen, J. S., Graff, C. A., and Ohman, M. D. (2019). Improving Plankton Image Classification Using Context Metadata. *Limnol. Oceanog.: Methods* 17 (8), 439–461. doi: 10.1002/lom3.10324

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* 281 (5374), 237. doi: 10.1126/science.281.5374.237

González, P., Castaño, A., Peacock, E. E., Díez, J., Del Coz, J. J., and Sosik, H. M. (2019). Automatic Plankton Quantification Using Deep Features. *J. Plankt. Res.* 41 (4), 449–463. doi: 10.1093/plankt/fbz023

Hamner, W. M., Madin, L. P., Alldredge, A. L., Gilmer, R. W., and Hamner, P. P. (1975). Underwater Observations of Gelatinous Zooplankton: Sampling Problems, Feeding Biology, and Behavior1. *Limnol. Oceanog.* 20 (6), 907–917. doi: 10.4319/lo.1975.20.6.0907

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep Residual Learning for Image Recognition. *ArXiv*, 1512.03385. doi: 10.48550/arXiv.1512.03385

Irisson, J. O., Ayata, S. D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. (2022). Machine Learning for the Study of Plankton and Marine Snow From Images. *Annu. Rev. Mar. Sci.* 3, 14. doi: 10.48550/arXiv.1512.03385

Iyer, N. (2012). Machine Vision Assisted *In Situ* Ichthyoplankton Imaging System.

Karlson, B., Cusack, C., and Bresnan, E. (2010). *Microscopic and Molecular Methods for Quantitative Phytoplankton Analysis* (UNESCO). Available at: https://repository.oceanbestpractices.org/handle/11329/303.

Kenitz, K. M., Orenstein, E. C., Roberts, P. L., Franks, P. J., Jaffe, J. S., Carter, M. L., et al. (2020). Environmental Drivers of Population Variability in Colony-Forming Marine Diatoms. *Limnol. Oceanog.* 65 (10), 2515–2528. doi: 10.1002/lno.11468

Kenitz, K. M., Anderson, C. R., Carter, M. L., Eggleston, E., Seech, K., Shipe, R., et al. (2022). Training Image Data For: Environmental and Ecological Drivers of Harmful Algal Blooms in the Southern California Bight. *UC San Diego Library Digital Collections*. doi: 10.6075/J00865GT

Kim, H.-J., Miller, A. J., McGowan, J., and Carter, M. L. (2009). Coastal Phytoplankton Blooms in the Southern California Bight. *Prog. Oceanog.* 82 (2), 137–147. doi: 10.48550/arXiv.1512.03385

Kingma, D. P., and Ba, J. (2014). Adam: A Method for Stochastic Optimization. *ArXiv. Prepr. ArXiv.*, 1412.6980. doi: 10.48550/arXiv.1512.03385

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature* 521 (7553), 436–444. doi: 10.1038/nature14539

Lee, H., Park, M., and Kim, J. (2016). "Plankton Classification on Imbalanced Large Scale Database *Via* Convolutional Neural Networks With Transfer Learning," in *2016 IEEE International Conference on Image Processing (ICIP)*. 3713–3717. doi: 10.1109/ICIP.2016.7533053

Lefebvre, K. A., Powell, C. L., Busman, M., Doucette, G. J., Moeller, P. D., Silver, J. B., et al. (1999). Detection of Domoic Acid in Northern Anchovies and California Sea Lions Associated With an Unusual Mortality Event. *Nat. Tox.* 7 (3), 85–92. doi: 10.1002/(sici)1522-7189(199905/06)7:3<85::aid-nt39>3.0.co;2-q

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Lumini, A., and Nanni, L. (2019). Deep Learning and Transfer Learning Features for Plankton Classification. *Ecol. Inf.* 51, 33–43. doi: 10.1016/j.ecoinf.2019.02.007

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., et al. (2018). Automated Plankton Image Analysis Using Convolutional Neural Networks. *Limnol. Oceanog.: Methods* 16 (12), 814–827. doi: 10.1002/lom3.10285

Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrera, F. (2012). A Unifying View on Dataset Shift in Classification. *Pattern Recog.* 45 (1), 521–530. doi: 10.1016/j.patcog.2011.06.019

Murez, Z., Kolouri, S., Kriegman, D., Ramamoorthi, R., and Kim, K. (2018). "Image to Image Translation for Domain Adaptation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 4500–4509.

Olson, R. J., and Sosik, H. M. (2007). A Submersible Imaging-in-Flow Instrument to Analyze Nano-and Microplankton: Imaging Flowcytobot: *In Situ* Imaging of

Nano- and Microplankton. *Limnol. Oceanog.: Methods* 5 (6), 195–203. doi: 10.4319/lom.2007.5.195

Omori, M., and Hamner, W. M. (1982). Patchy Distribution of Zooplankton: Behavior, Population Assessment and Sampling Problems. *Mar. Biol.* 72 (2), 193–200. doi: 10.1007/BF00396920

Orenstein, E. C., and Beijbom, O. (2017). "Transfer Learning and Deep Feature Extraction for Planktonic Image Data Sets," in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 1082–1088. doi: 10.1109/WACV.2017.125

Orenstein, E. C., Beijbom, O., Peacock, E. E., and Sosik, H. M. (2015). Whoi-Plankton-a Large Scale Fine Grained Visual Recognition Benchmark Dataset for Plankton Classification. *ArXiv. Prepr. ArXiv.*, 1510.00745.

Orenstein, E. C., Kenitz, K. M., Roberts, P. L. D., Franks, P. J. S., Jaffe, J. S., and Barton, A. D. (2020b). Semi- and Fully Supervised Quantification Techniques to Improve Population Estimates From Machine Classifiers. *Limnol. Oceanog.: Methods* 18 (12), 739–53. doi: 10.1002/lom3.10399

Orenstein, E. C., Ratelle, D., Briseño-Avena, C., Carter, M. L., Franks, P. J. S., Jaffe, J. S., et al. (2020a). The Scripps Plankton Camera System: A Framework and Platform for *in Situ* Microscopy. *Limnol. Oceanog.: Methods* 18 (11), 681–695. doi: 10.1002/lom3.10394

Picheral, M., Catalano, C., Brousseau, D., Claustre, H., Coppola, L., Leymarie, E., et al. (2021). The Underwater Vision Profiler 6: An Imaging Sensor of Particle Size Spectra and Plankton, for Autonomous and Cabled Platforms. *Limnol. Oceanog.: Methods* 20 (2), 115–29. doi: 10.1002/lom3.10475.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The Underwater Vision Profiler 5: An Advanced Instrument for High Spatial Resolution Studies of Particle Size Spectra and Zooplankton. *Limnol. Oceanog.: Methods* 8 (9), 462–473. doi: 10.4319/lom.2010.8.462

Pinheiro, P. O. (2018). "Unsupervised Domain Adaptation With Similarity Learning," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp 8004–8013.

Salman, A., Jalal, A., Shafait, F., Mian, A., Shortis, M., Seager, J., et al. (2016). Fish Species Classification in Unconstrained Underwater Environments Based on Deep Learning. *Limnol. Oceanog.: Methods* 14 (9), 570–585. doi: 10.1002/lom3.10113

Scholin, C. A., Gulland, F., Doucette, G. J., Benson, S., Busman, M., Chavez, F. P., et al. (2000). Mortality of Sea Lions Along the Central California Coast Linked to a Toxic Diatom Bloom. *Nature* 403 (6765), 80–84. doi: 10.1038/47481

Shihavuddin, A. (2017). Coral Reef Dataset, v2. *Mendeley Data*. Available at: https://data.mendeley.com/datasets/86y667257h/2

Sinha, E., Michalak, A. M., and Balaji, V. (2017). Eutrophication Will Increase During the 21st Century as a Result of Precipitation Changes. *Science* 357 (6349), 405–408. doi: 10.1126/science.aan2409

Smith, J., Connell, P., Evans, R. H., Gellene, A. G., Howard, M. D. A., Jones, B. H., et al. (2018). A Decade and a Half of Pseudo-nitzschia Spp. And Domoic Acid Along the Coast of Southern California. *Harmf. Algae.* 79, 87–104. doi: 10.1016/j.hal.2018.07.007

Sosik, H. M., and Olson, R. J. (2007). Automated Taxonomic Classification of Phytoplankton Sampled With Imaging-in-Flow Cytometry. *Limnol. Oceanog.: Methods* 5 (6), 204–216. doi: 10.4319/lom.2007.5.204

Stroming, S., Robertson, M., Mabee, B., Kuwayama, Y., and Schaeffer, B. (2020). Quantifying the Human Health Benefits of Using Satellite Information to Detect Cyanobacterial Harmful Algal Blooms and Manage Recreational Advisories in U.S. Lakes. *GeoHealth* 4 (9), e2020GH000254. doi: 10.1029/2020GH000254

Tetko, I. V., Livingstone, D. J., and Luik, A. I. (1995). Neural Network Studies. 1. Comparison of Overfitting and Overtraining. *J. Chem. Inf. Model.* 35 (5), 826–833. doi: 10.1021/ci00027a006

Utermöhl, H. (1931). Neue Wege in Der Quantitativen Erfassung Des Plankton. (Mit Besonderer Berücksichtigung Des Ultraplanktons). *SIL. Proc.* 5 (2), 567–596. doi: 10.1080/03680770.1931.11898492

Utermöhl, H. (1958). Methods of Collecting Plankton for Various Purposes are Discussed. *SIL. Communicat. 1953-1996* 9 (1), 1–38. doi: 10.1080/05384680.1958.11904091

Wang, Y.-X., Ramanan, D., and Hebert, M. (2017). "Learning to Model the Tail," in *Advances in Neural Information Processing Systems*, vol. 30 . Eds. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan and R. Garnett (Curran Associates, Inc), 7029–7039. Available at: http://papers.nips.cc/paper/7278-learning-to-model-the-tail.pdf.

Whitmore, B. M., Nickels, C. F., and Ohman, M. D. (2019). A Comparison Between Zooglider and Shipboard Net and Acoustic Mesozooplankton Sensing Systems. *J. Plankt. Res.* 41 (4), 521–533. doi: 10.1093/plankt/fbz033

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How Transferable are Features in Deep Neural Networks? *Adv. Neural Inf. Process. Syst.* 27. 1–9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

METHODS
published: 21 June 2022
doi: 10.3389/fmars.2022.898057

Comparison of an *In Situ* Imaging Device and Net-Based Method to Study Mesozooplankton Communities in an Oligotrophic System

Alexander Barth * *and Joshua Stone*

Biological Sciences Department, University of South Carolina, Columbia, SC, United States

In the past several years, the capabilities of optical tools and *in situ* imaging devices have greatly expanded and are now revolutionizing the field of plankton research. These tools have facilitated the discovery of new plankton and enhanced the understanding of populations of fragile and gelatinous zooplankton. Imaging devices are becoming more accessible and regularly deployed on oceanographic studies and monitoring efforts. However, despite the increasing use of these tools, there are few studies which offer direct comparisons between *in situ* imaging devices and traditional-net based methods, especially in open-ocean, oligotrophic systems where plankton are sparser and less intensively sampled. This study compares estimates of mesozooplankton abundance calculated by net-tows and an Underwater Vision Profiler 5 (UVP5HD-DEEP) imaging system. Net tows were conducted with a Multiple Opening and Closing Nets with Environmental Sensing System (MOCNESS) device equipped with 153μm mesh. In total, four tows, each sampling eight distinct depth bins, were conducted aboard two cruises in the Sargasso Sea. Along each cruise, *in situ* images were collected using an Underwater Vison Profiler 5 (UVP5HD-DEEP). Using these methods, we estimated abundance of different mesozooplankton groups (>0.5 mm). Using established biovolume-biomass conversions, we also estimated the dry mass of certain zooplankton taxa. Furthermore, we address two methods for calculating density and biomass concentration from UVP data. Estimates of mesozooplankton abundance and biomass concentration were generally higher from MOCNESS methods than the UVP estimates across all taxa. It was found that there is not a reliable relationship between UVP estimates and MOCNESS estimates when directly comparing similar depth bins. Nonetheless, when integrating density and biomass concentrations throughout the water column, estimates are not significantly different between the methodology. This study addresses several important considerations for using *in situ* imaging tools and how to reconcile findings with traditional net-based methods.

Keywords: *in situ*, plankton, oceanography, Sargasso Sea, sampling, ocean optics, copepod

OPEN ACCESS

Edited by:
Elaine Fileman,
Plymouth Marine Laboratory,
United Kingdom

Reviewed by:
Akash R. Sastri,
Institute of Ocean Sciences, Canada
Helena Hauss,
GEOMAR Helmholtz Center for Ocean Research Kiel, Germany

****Correspondence:***
Alexander Barth
AB93@email.sc.edu

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *16 March 2022*
Accepted: *23 May 2022*
Published: *21 June 2022*

Citation:
Barth A and Stone J (2022) Comparison of an In Situ Imaging Device and Net-Based Method to Study Mesozooplankton Communities in an Oligotrophic System. Front. Mar. Sci. 9:898057. doi: 10.3389/fmars.2022.898057

1 INTRODUCTION

To understand ocean ecosystems, it is necessary to understand zooplankton community structure. Zooplankton have a wide range of complex life history strategies, body types, and feeding strategies (Kiørboe et al., 2011). Trophic interactions and behavior of zooplankton can have large impacts on the biological carbon pump, and thus the global carbon cycle (Steinberg and Landry 2017). However, zooplankton communities are extremely dynamic, and populations of different plankton can fluctuate largely over fine temporal and spatial scales. Therefore, studying zooplankton populations and communities can be a great challenge.

Historically, zooplankton have been collected using net-based approaches. Mesh nets allow for the concentration of large volumes of water to sample zooplankton and accurately estimate their abundance. Over the past several decades, advances in net technology includes opening and closing net systems, such as the MOCNESS and MultiNet, which allow for the study of zooplankton communities' vertical structure (Wiebe and Benfield, 2003). However, there are limitations to net-based study of plankton. Specifically, nets can be destructive and do not adequately sample gelatinous or fragile bodied zooplankton. Additionally, even with open-closing net systems, nets do not offer fine enough scales of vertical resolution to study zooplankton which can occur in dense, thin layers (Holliday et al., 2003).

Recently, developments in imaging technology have offered a new way to study zooplankton. *In situ* imaging tools offer a large advantage over nets because they can sample a plankter's exact position in the water column. The frequency of image collection can be fairly close to the frequency of data collection for physical parameters. This information allows for the study of plankton in context with small-scale changes in physical features of the water column (Ohman, 2019) and ecological interactions such as thin layers. Furthermore, *in situ* imaging allows for the characterization of plankton's natural state and traits, while nets can disturb and damage plankton. In the past several years, plankton ecologists are increasingly utilizing a trait-based approach to characterize zooplankton communities (Litchman et al., 2013; Kiørboe et al., 2018). Recently, there has been advances in combining *in situ* imaging and trait-based methodology to study zooplankton (Ohman, 2019; Vilgrain et al., 2021; Orenstein et al., 2021). Studying plankton *in situ* is particularly important for the study of fragile and gelatinous organisms. *In situ* plankton imaging devices have recently shed light on some of the major community roles that previously under-described taxa have in ocean ecosystems (Biard et al., 2016; Christiansen et al., 2018; Hoving et al., 2019; Stukel et al., 2019).

There are a wide range of tools which facilitate the study of zooplankton *in situ.* Examples of this technology include the zooglider (Ohman et al., 2018), ISIIS (Cowen and Guiginad, 2008), LOKI (Schulz et al., 2010), LOPC (Herman et al., 2004), PELAGIOS (Hoving et al., 2019), VPR (Davis et al., 1992), and UVP (Picheral et al., 2010; Picheral et al., 2022) (see Lombard et al., 2019 for a complete review of optical tools). Although these tools accomplish a similar goal, they have vastly different approaches and outcomes. Some devices are independently towed (Cowen and Guiginad 2008; Hoving et al., 2019), while others are designed to be incorporated with oceanographic instrument rosettes (Picheral et al., 2010; Picheral et al., 2022). Illumination and imaging technology also varies between devices. Plankton cameras can include white-light (Hoving et al., 2019), single beam and two-beam red-light (Picheral et al., 2010; Picheral et al., 2022), shadowgraphy (Cowen and Guiginad 2008; Ohman et al., 2018), holography (Nayak et al., 2021), dark-field microscopy (Orenstein et al., 2020), and more. Additionally, these devices range in the quality of image taken, frequency of data collection, and volume sampled in a given profile (Lombard et al., 2019). In this paper, we focus on the Underwater Vision Profiler 5 (UVP5; Picheral et al., 2010). The UVP5 has been commercially available for several years and is a popular tool due to its ability to study both particles and zooplankton. Additionally, the UVP is designed to be able to integrate into CTD-rosette instrument packages and collect data semi-autonomously. This allows for collection of data alongside with standard oceanographic research and no additional wire-time. To date, there have been several studies utilizing UVPs to study particles (Forest et al., 2013; Puig et al., 2013; Martin et al., 2013; Jouandet et al., 2014; Miquel et al., 2015; Waite et al., 2016; Turner et al., 2017; Hoving et al., 2020); zooplankton (Forest et al., 2012; Biard et al., 2016; Hauss et al., 2016; Donoso et al., 2017; Christiansen et al., 2018; Vilgrain et al., 2021), and cyanobacteria (Guidi et al., 2012; Sandel et al., 2015). Studies utilizing the UVP have been conducted in a wide range of environments including the Mediterranean (Donoso et al., 2017; Durrieu de Madron et al., 2017; Severin et al., 2017), Equatorial (Kiko et al., 2017), Atlantic (Thomsen et al., 2019; Christiansen et al., 2018), Pacific (Turner et al., 2017; Stukel et al., 2019), Artic (Miquel et al., 2015; Vilgrain et al., 2021) and Antarctic (Martin et al., 2013). However, there are few studies in oligotrophic regions (Sandel et al., 2015).

One challenge of sampling zooplankton in oligotrophic systems is that zooplankton densities are very low. Thus, large volumes of water are required to adequately study their populations. Some studies with the UVP observed that the volume sampled was too low to adequately describe zooplankton populations (Donoso et al., 2017). However, Forest et al. (2012) found that in a copepod dominated system, the UVP can yield similar density estimate to net-based systems. This suggests the need for regional analyses to assess how effective the UVP measurements are compared to net-based systems. Additionally, recent developments in the UVP have made it available with much higher sampling frequencies, facilitating a larger sampling volume. This can increase the reliability of UVP data collection. However, high sampling frequencies also introduce new challenges like double imaging of individual particles.

Overall, the UVP is an attractive choice for studying zooplankton due to its ability to both study organisms *in situ* as well as the ease of incorporating it into standard sampling programs. However, there is a clear need to assess how UVP

estimates of zooplankton populations compare to net-based systems. In the present study, we offer a comparison of zooplankton abundance and biomass calculations using an *in situ* imaging device (UVP5-HD) and a depth-specific net system (MOCNESS). This study addresses particular challenges working with high-frequency imaging systems and systems with low-organism density. Finally, we describe the reliability of sampling different devices with both devices.

2 MATERIALS AND METHODS

2.1 Sample Collection

Data were collected onboard the R/V Atlantic Explorer during 5-day cruises as part of the Bermuda Atlantic Time-series Study (BATS; Steinberg et al., 2001), which conducts monthly long-term monitoring sampling ~80 km southeast of Bermuda. This study utilizes data collected during the June and July BATS cruises of 2019 (AE1912 and AE1917 respectively).

2.1.1 *In Situ* Imaging of Plankton

An Underwater Vision Profiler (UVP5-HD, sn:209, Hydroptic, Picheral et al., 2010) was attached to the CTD Rosette aboard the R/V Atlantic Explorer. This model of the UVP5-HD has a 4.2-megapixel camera which images a 3.11cm x 18.8 cm x 18.8cm (H x W x L) field of view (1.1L). The pixel size is 92µm. The UVP is designed to measure both particle abundances and collect *in situ*zooplankton imagery. Under mixed acquisition mode, the UVP automatically segments and measures all particles larger than 125µm according to equivalent spherical diameter (ESD). All particles larger than 500µm ESD were recorded and stored as individual images (vignettes). Images are collected at a rate of approximately 15Hz.

The UVP was attached to the CTD rosette on all cruises and configured for automatic acquisition of data for all profiles during each cruise (~18 casts per cruise). On average, a UVP cast sampled 9.31m^3 in the epipelagic (0-250m) and 14.38m^3 in the mesopelagic (250m-1000m) (**Table 1**; **Supplemental Figure 1**; **Supporting Information**). Images are collected during down casts only, then the UVP is programmed to turn off once it has ascended more than 30m. Data are downloaded from the UVP onboard then processed in Zooprocess and trimmed to remove any data collected during the rinse cycle or the first 30m of the upcast. UVP data are then uploaded to the EcoPart web application (Picheral et al., 2017; https://ecopart.obs-vlfr.fr/), which applies a descent filter to account for variation in the data from ship rock or variable CTD descent speeds. The descent filter excludes any images which were taken at a shallower depth than the preceding image. However, UVP images can still overlap if the UVP is descending at a slow enough rate (<0.622ms^{-1}; **Supplemental Figure 1**). The average UVP descent rate in the epipelagic was 0.653ms^{-1} and 1.099ms^{-1} in the mesopelagic. We did find that at the bottom 50m of each profile, the slowing of the CTD rosette could lead to the high potential of re-imaging particles (**Supplemental Figure 1**). To account for this, we removed data collected from the bottom 50m of each profile. The typical UVP cast descended to 1200m, although several descended to approximately 500m. Consequently, the removal of the bottom 50m has minimal impact on the data available for this analysis.

2.1.2 Net-Based Plankton Collection

Plankton were collected using a Multiple Opening and Closing Net with Environmental Sensing System (MOCNESS; Wiebe et al., 1976, Wiebe et al., 1985). The MOCNESS has a 1m^{-2} opening and was equipped with 153µm mesh and was deployed on oblique tows. The MOCNESS was used to sample plankton at discrete depth-bins following an adaptive profiling method to sample ecologically relevant regions of the water column. Specifically, bins were targeted to capture variation around the deep chlorophyll-a maximum (DCM) which was determined prior to each tow using a CTD cast with an attached fluorometer (Chelsea Instruments). It should be noted that tows between the two cruises did have different maximum depths (**Table 2**). During the June cruise (AE1912), one night tow (AE1912m1) was conducted to a maximum depth of 1000m. During the July cruise (AE1917), three tows were conducted. Two day-time tows (AE1915m14 & AE1917m15) with bottom depths to 270m & 260m respectively, and a night-time tow (AE1917m16) to 260m. Due to the difference in maximum depth between the tows, when warranted by analyses, we distinguish the June night tow into an epipelagic section and a mesopelagic section, defined by above or below 250m. Once on-board the plankton samples were split, and half were fixed in buffered 4% formalin to be used in the present study.

2.2 Laboratory Processing of Net Samples

Fixed samples of plankton were transported back to the lab where they were measured using a ZooScan (Hydroptic; Gorsky et al., 2010). Samples collected from the June cruise (AE1912) were scanned at the University of South Carolina at 2400dpi. Samples from the July cruise (AE1917) were scanned at the Bermuda Institute of Oceanography at 4800dpi. To optimize segmentation (extracting vignettes of individual particle images from scanned samples), samples were size fractioned and split so that there were not too many particles in any given scan. Samples from the June cruise were split into two size fractions using a 1500µm sieve. Samples from the July cruise were split into three size fractions; all individual organisms larger than 2mm were removed by hand and imaged, then the samples were split using a 1000µm sieve. For all samples, the larger size fractions were split using a Motoda splitter (Motoda, 1959). All splits from the larger size fractions were scanned. For the smallest size fraction, it was important that there were not too many organisms in any single scan because this can impact the extraction of individuals. Samples from the June cruise were diluted while those from the July cruise were split to reduce concentrations in individual scan. For both approaches, enough scans were done so that there were at least 1500 objects scanned from each net. Scans were then processed using Zooprocess (Gorsky et al., 2010) to extract vignettes of individual objects. The default setting of the ZooScan extracts all objects larger than 300µm, a size much smaller than what is characterized by the UVP. For direct

TABLE 1 | Metadata for UVP casts.

Cast Group	UVP Casts	Deployment Time (UTC - 3)	Latitude	Longitude	Comparable Depth Range [m]	Total Volume Sampled [m^3]
June Night	gf360c11	Jun 08 2019, 23:18	31.66 N	64.17 W	0-250	9.537
					250-500	4.788
	gf360c17	Jun 09 2019, 00:55	32.15 N	64.02 W	0-250	9.139
					25-1000	16.805
	gf360c5	Jun 06 2019, 03:21	31.97 N	64.38 W	0-250	8.888
					250-1000	16.94
	gf360c9	Jun 07 2019, 03:07	31.17 N	64.32 W	0-250	9.217
					25-1000	17.375
June Day-A	bats361_ctd1	Jul 14 2019, 10:10	32.337 N	64.59 W	0-253	10.537
	bats361_ctd12	Jul 15 2019, 18:20	31.17 N	64.32 W	0-270	9.285
	bats361_ctd14	Jul 16 2019, 12:39	31.67 N	64.15 W	0-270	8.845
	bats361_ctd15	Jul 16 2019, 17:16	31.49 N	64.53 W	0-270	10.659
	bats361_ctd19	Jul 17 2019, 19:13	31.74 N	64.22 W	0-270	8.544
	bats361_ctd2	Jul 14 2019, 10:56	32.30 N	64.57 W	0-270	10.413
	bats361_ctd23	Jul 18 2019, 09:35	32.02 N	63.44 W	0-270	10.13
	bats361_ctd24	Jul 18 2019, 13:00	32.33 N	63.65 W	0-270	9.928
	bats361_ctd25	Jul 18 2019, 17:10	32.05 N	64.15 W	0-270	9.55
	bats361_ctd3	Jul 17 2019, 12:17	32.26 N	64.55 W	0-270	7.977
July Day-B	bats361_ctd1	Jul 14 2019, 10:10	32.337 N	64.59 W	0-253	10.537
	bats361_ctd12	Jul 15 2019, 18:20	31.17 N	64.32 W	0-260	9.062
	bats361_ctd14	Jul 16 2019, 12:39	31.67 N	64.15 W	0-260	8.625
	bats361_ctd15	Jul 16 2019, 17:16	31.49 N	64.53 W	0-260	10.418
	bats361_ctd19	Jul 17 2019, 19:13	31.74 N	64.22 W	0-260	8.329
	bats361_ctd2	Jul 14 2019, 10:56	32.30 N	64.57 W	0-260	10.167
	bats361_ctd23	Jul 18 2019, 09:35	32.02 N	63.44 W	0-260	9.892
	bats361_ctd24	Jul 18 2019, 13:00	32.33 N	63.65 W	0-260	9.695
	bats361_ctd25	Jul 18 2019, 17:10	32.05 N	64.15 W	0-260	9.301
	bats361_ctd3	Jul 17 2019, 12:17	32.26 N	64.55 W	0-260	7.74
July Night	bats361_ctd16	Jul 17 2019, 22:09	31.76 N	63.99 W	0-260	11.057
	bats361_ctd17	Jul 17 2019, 00:21	31.67 N	64.17 W	0-260	10.018
	bats361_ctd20	Jul 17 2019, 23:19	31.66 N	64.17 W	0-260	10.055
	bats361_ctd21	Jul 18 2019, 01:04	31.54 N	63.60 W	0-260	10.121
	bats361_ctd26	Jul 18 2019, 23:39	31.66 N	64.02 W	0-260	9.871
	bats361_ctd27	Jul 19 2019, 01:39	32.16 N	64.02 W	0-260	9.489
	bats361_ctd7	Jul 14 2019, 22:32	31.75 N	63.99 W	0-260	10.57
	bats361_ctd8	Jul 15 2019, 01:06	31.84 N	63.80 W	0-260	8.082

Cast group indicates which casts were aggregated to compare to MOCNESS tows.

comparisons to the UVP dataset, the MOCNESS data was trimmed to only include plankton which were equal to or larger than the smallest identified plankton from UVP data. This size cutoff was determined to be 934μm.

2.3 Classification of Images and Morphology

The Ecotaxa web application was used to sort vignettes from both instruments (UVP and Zooscan) using a random forest classifier (Picheral et al., 2017; https://ecotaxa.obs-vlfr.fr/). All predicted identification were subsequently verified or reclassified by the same trained annotator. Generally, images collected by the UVP cannot be reliably classified to the same taxonomic resolution as images collected by the MOCNESS/ZooScan due to lower image resolution. As a result, many taxa were grouped into broad categories for comparison. Most notably, all Copepoda (Class: Hexanauplia, Subclass Copepoda) taxa were grouped to "copepod", Decapods (Class: Malacostraca, Superorder: Eucardia, Order: Decapoda) and Euphausiids (Class: Malacostraca, Superorder: Eucardia, Order: Euphausiacea) were grouped to "shrimp-like crustaceans", and ostracods (Class: Ostracoda) and cladocerans (Class: Brachiopoda, Subclass: Phyllopoda, Superorder: Diplostraca) were grouped to "Ostracod/Cladoceran". Morphologically relevant metrics for each particle (major axis, minor axis, grey level, etc.) are computed in Zooprocess.

2.3.1 Management of ZooScan Vignettes With Multiple Organisms in a Single Vignette

Processing samples with the ZooScan requires manual separation of particles to facilitate the segmentation algorithm in zooprocess. However, it is inevitable that a few individual objects will not be separated during segmentation. Zooprocess allows for the post-processing of unseparated individuals. However, this can result in straight-lines and alter the accuracy of computed morphometrics. Additionally, there is a small portion of organisms which cannot be separated, even in post-processing if they are entangled or overlapping in a scan. To manage this challenge, we manually flagged all vignettes with multiple individuals during Ecotaxa classification. These vignettes were then re-examined, and individuals were counted after Ecotaxa classification. Because the morphometrics

TABLE 2 | MOCNESS metadata for the four tows.

MOCNESS TOW	Location	Times (UTC-3)	Depth Bins [m]	Volume Filtered [m^3]
June Night	**Deployed**:	**Deployed**: Jun 06 2019, 21:21	791.6-995.7	980.4
	31.65N 64.15W	**Retrieved**: Jun 07 2019, 00:59	590.4-791.6	1762.4
	Retrieved:		348.3-590.4	2257.7
	31.6N 64.1W		253.7-348.3	941.1
			150.5-253.7	643.8
			48.0-150.5	656.5
			5.1-48.0	535.7
July Day A	**Deployed:**	**Deployed:** Jul 16 2019, 16:24	219.7-270.4	620.7
	31.65N 64.15W		168.7-219.7	387.4
	Retrieved:	**Retrieved:** Jul 16 2019, 18:07	140.1-168.7	242.8
	31.62N 64.15W		111.1-140.1	448.1
			80.6-111.1	471.7
			50.7-80.6	362.7
			20-50.7	378.7
			0.8-20	233.1
July Day B	**Deployed:**	**Deployed:** Jul 17 2019, 17:12	221.4-260.5	299.6
	31.72N 64.18W	**Retrieved:** Jul 17 2019, 18:45	180.9-221.4	355.7
	Retrieved:		150.6-180.9	236.9
	31.72N 64.22W		119.4-150.6	265.8
			89.0-119.4	198.2
			59.0-89.0	180.1
			30.7-59.0	277.1
			0-30.7	233.5
July Night	**Deployed:**	**Deployed:** Jul 19 2019, 01:21	220.7-259.8	323.5
	31.67N 64.17W	**Retrieved:** Jul 19 2019, 03:07	180.8-220.7	713.5
	Retrieved:		151.0-180.8	271.9
	31.68N 64.17W		120.8-151.0	370.0
			89.1-120.8	731.1
			59.7-89.1	300.9
			30.4-56.7	510.7
			0-30.4	338.1

associated with these vignettes are inaccurate, we assigned each individual a set of randomly selected morphometric values from like taxa. Using this method assumes that the morphology of an organism does not influence the likelihood of it being caught in a multiple vignette, which for our dataset appeared to be a reliable assumption. This only affected 3% of our total MOCNESS data (**Supplemental Information**).

2.3.2 Assessing the Impact of Twice-Imaged Organisms in UVP Images

During validation of UVP vignettes, we noticed that there were cases where the same individual organism was imaged multiple times (**Supplemental Figure 2**). This can occur when the UVP is descending at a rate slower than the rate required to avoid overlap of images. It appears to be a concern primarily with larger, darker organisms. To assess the impact of multiple recordings of individuals, for all casts aboard AE1912, we sorted vignettes which clearly were multiple recordings into a distinct category. Estimates of those specific taxa's density were then estimated for each profile in 20-m bins following two methods. In the first method, multiple-imaged organisms were treated as standard observations and counted, then divided by the total volume sampled in that 20-m bin. In the second method, multiple-imaged organisms had all but one vignette removed, then to account for this removal, the recorded volume in a 20-m bin was reduced to the maximum possible volume for non-overlapping UVP images in a 20-m stack ($0.643 m^3$).

2.4 Data Processing

To compare data collected from UVP casts to MOCNESS tows, UVP casts were categorized as either day or night for each month (**Table 1**). Sunrise and sunset times were calculated using the NOAA ESRL Solar Calculator (https://gml.noaa.gov/grad/solcalc/). To account for any potential diel vertical migration, casts which occurred within an hour before or after sunrise/sunset were marked as twilight and not included. There were no MOCNESS tows conducted near twilight hours.

First the instruments were compared on their scope of sampling to assess which taxa can be compared between the two devices. While the UVP was set to record all particles above 500 μm, larger sizes were required to reliably identify objects as living organisms. The smallest living organism recorded by the UVP was 0.934 mm. Because of this, MOCNESS-collected plankton which were smaller than 0.934 mm were excluded from all analyses. Then the relative contribution of different taxonomic groups were compared between the instruments. Finally, Annelids, Copepods, Chaetognaths, Shrimp-like crustaceans, and Ostracod/Cladocerans were selected for direct comparison.

2.4.1 UVP Cast Binning and Aggregating

UVP casts were binned into distinct depth bins in which concentration could be calculated. For depth specific bin comparison, the UVP bins were set to match the MOCNESS depth bins. However, a benefit of the UVP is the ability to resolve finer-scale patterns in mesozooplankton. Thus, for visualization and depth-integration, UVP bins were independently set. To select the independent bin size, depth integrated abundance was calculated for different taxa in individual UVP casts using trapezoidal integration across a range bin sizes. The smallest bin size which still yielded stable estimates was found to be 20-m (**Supplemental Figure 4**; **Supplemental Information**).

There were two methods utilized to aggregate the several UVP casts which correspond to a single MOCNESS tow. The first method is a pooled-cast approach. In this method, all similar casts are pooled into one representative profile. Then the concentration of observations (either counts or summed biomass) were calculated for each depth bin (Equation 1). This approach is common in UVP studies as it can increase the volume sampled in an individual depth bin.

$$\frac{\sum_{I}^{N} Observation_i}{\sum_{i}^{N} Vol - Sampled_i}$$

Equation 1:

Pooled-cast calculation for a UVP depth bin concentration of i observations (counts or biomass) for all N casts in a depth bin.

The other method was an average-cast approach. This calculated concentration in a depth bin in each individual cast, then took the mean of all similar casts (Equation). This approach allows for the characterization of mean and standard deviation between similar casts.

$$\frac{\sum_{i}^{N} \frac{Observation}{Volume\ \ Sampled_i}}{N}$$

Equation 2.

Average-cast method for a UVP depth bin. The concentration of i observations summed across all N casts then divided by the number of casts.

2.5 Taxa-Specific Comparison of Density

To assess patterns of density throughout the water column, the concentrations of each comparable taxa were plotted using independent UVP bins with both pooled-cast and average-cast methods overlayed on MOCNESS data. Then, to quantify the difference in depth specific density estimates between the two sampling methods, linear regressions were conducted between the estimated concentration of each comparable taxa. For this analysis, the concentration of organisms was calculated in each depth-bin as determined by the MOCNESS tows. Regressions were done between pooled-cast UVP data versus MOCNESS data and average-cast UVP versus MOCNESS. For the average-cast approach, the mean concentration was used.

Then, the depth integrated abundance was calculated for all UVP and MOCNESS profiles. Due to the difference in sampling methodology, the June cruise was split into an epipelagic zone and mesopelagic zone. This provided two split integrations from data collected at the same location. Depth integration used trapezoidal integration with linear approximation between the mid-points of each depth bin (**Supplemental Information**; github.com/thealexbarth/EcotaxaTools). UVP casts used independent bin sizes for this integration. For the pooled-cast method, one integration was done over the whole pooled-cast. For the average-cast methods, integrations were done on each individual UVP cast, then the mean depth integrated abundance was found for similar casts. Paired Wilcox signed rank tests were used to compare depth integrated abundance between the different methods for each taxa.

2.6 Taxa-Specific Comparison of Biomass

For all comparable taxa, the volume of each individual vignette was calculated following assuming an ellipsoidal shape (Equation 3). This required the conversion of pixels to mm, which was a different conversion for each device (**Supplemental Information**).

$$Biovolume = \frac{4}{3}\pi(major\ \ axis)(minor\ \ axis)^2$$

Equation 3.

Biovolume estimation for an individual plankton vignette assuming an ellipsoidal shape.

Then, dry mass was calculated for each individual using biovolume to mass conversions described in Maas et al. (2021). Because the UVP does not facilitate high taxonomic specificity, we assigned all copepods the conversion factor for Calanoida, and all shrimp-like crustacean the conversion factor for Decapoda. Annelids were excluded from this analysis because there was not an available conversion factor. The biomass concentration (mg m^{-3}) was calculated for each depth bin by summing the biomass of all individuals of a given taxa then dividing by the total volume sampled in that depth bin. This was done using both the pooled-cash approach and average-cast approach for the UVP. Again, linear regressions were used to compare the direct calculations of biomass concentration between the MOCNESS and the two UVP approaches. Then the depth integrated biomass was calculated following the same steps as for the abundance. Paired Wilcox signed rank tests were used to compare depth integrated biomass between the different methods for each taxon.

All data were processed using R ver. 4.0 (R Core Team). Data were processed largely using the EcotaxaTools package (github.com/TheAlexBarth/EcotaxaTools). All data and code are available in **Supplemental Information 1**.

3 RESULTS

3.1 Scope of Instruments

Images of the MOCNESS-collected plankton acquired by the ZooScan are generally much higher resolution than the *in situ* images acquired by the UVP (**Figure 1**). Although the UVP does acquire images which are capable of identifying several taxa

in situ, the MOCNESS facilitates identification to a higher taxonomic resolution than the UVP. Notably, copepods sampled by the MOCNESS/ZooScan can be separated into at least the order level, and at times the family level. While some larger Calanoid copepods can easily be identified from UVP images, smaller copepods cannot be reliably identified to higher taxonomic resolution (**Figure 1**). From the MOCNESS tows, it appears that the majority of copepods sampled in this system are Calanoids or Cyclopoids, with a smaller percentage of Harpacticoids. The UVP can detect both decapods and euphausiids, although these cannot be reliably distinguished in most vignettes, so they were grouped to Eumalacostraca (referred to as shrimp-like crustacean). The MOCNESS/ZooScan images can be consistently distinguished as euphausiids or decapods, although for comparison to the UVP, we combined these as shrimp-like crustaceans. Additionally, the MOCNESS is able to sample meroplankton, larval forms, and fish (**Figure 1**). A few fish were sampled by the UVP, although these were often while in motion (**Figure 1**).

Recording multiples of an individual did not have a clear effect on the UVP density estimates. After visually investigating the difference between taxon-specific density estimates for all June UVP casts when including multiple-recorded individuals and excluding them, we found no observable pattern (**Supplemental Figure 3**). This issue was most noticeable in select rhizaria and *Trichodesmium* images and inclusion of multiples would slightly increase density estimates. Alternatively, the exclusion method of multiple images also at times increased density estimates (**Supplemental Figure 3**). Thus, we determined it would be best to include multiples, particularly because the exclusion requires alteration of the volume sampled measurement, which can then decrease the accuracy of concentration estimate for all other taxa.

As expected, the MOCNESS sampled a much larger size range than the UVP. The ZooScan is set to record all individual particles larger than 300µm ESD while the UVP is set to record all individual particles larger than 500µm ESD. However, the images collected by the UVP could only be reliably identified for much larger particles. Thus, the smallest living organism collected by the UVP identified was 934µm. For comparison, all MOCNESS-collected plankton below this size were excluded. This exclusion removes a large portion of the plankton collected by the MOCNESS (**Supplemental Figure 5**). Notably, 91.2% and 96.7% of copepods and ostracods/cladocerans respectively, sampled by the MOCNESS were smaller than 934µm. For other MOCNESS-collected plankton, 30% of annelids were below this size cut-off while only 11% of chaetognaths and 6.98% of Shrimp-like crustaceans. With the size trimmed MOCNESS data, there was a considerable overlap with the UVP in the size distribution of all plankton (**Figure 2**). The median MOCNESS size was 1.87mm. The median UVP size was 1.56mm. For specific taxa, the MOCNESS generally sampled across sizes more evenly than the UVP, which had its size distributions more concentrated (**Figure 2**). There was a large size overlap for copepods although the MOCNESS median (1.51mm) was slightly larger than the UVP median (1.31mm) (**Figure 2C**; **Supplemental Information**). Interestingly, the MOCNESS seemingly sampled larger chaetognaths and shrimp-like crustaceans better than the UVP. The MOCNESS size distribution for those taxa had secondary peaks between 3-3.75mm, where there were very few UVP-imaged plankton at those sizes (**Figures 2B, E**). Alternatively, the UVP size distributions for annelids and ostracods/cladocerans were

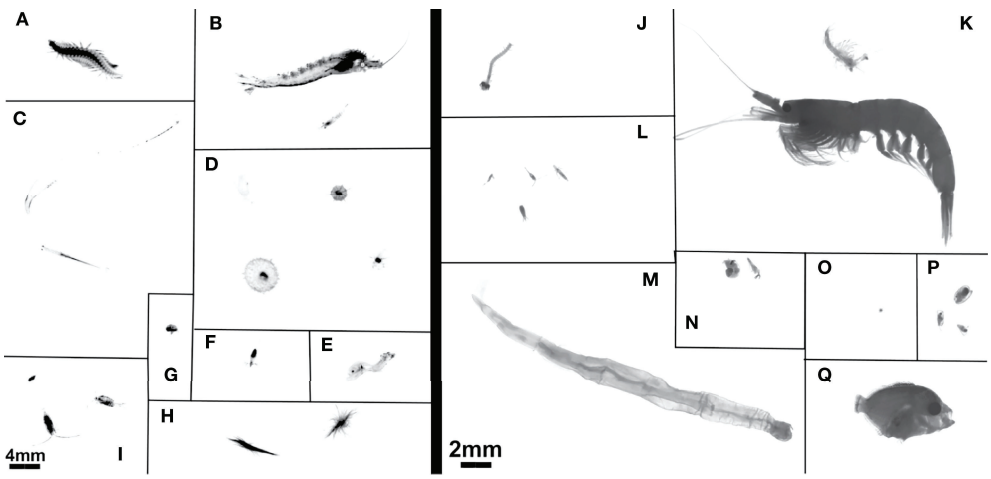


FIGURE 1 | Example images from **(A–I)** the UVP and **(J–Q)** MOCNESS. UVP images have 4mm scale bar in bottom left. MOCNESS images use 2mm scale bar in bottom right. UVP images are **(A)** Annelida, **(B)** Shrimp-like Crustacea, **(C)** Chaetognatha, **(D)** Rhizaria, **(E)** Actinopterygii, **(F)** Mollusca, **(G)** Ostracoda, **(H)** Trichodesmium, **(I)** Copepoda. MOCNESS images are **(J)** Annelida, **(K)** Shrimp-like Crustacea, **(L)** Copepoda, **(M)** Chaetognatha, **(N)** Mollusca, **(O)** Acantharea, **(P)** Ostracoda, **(Q)** Actinopterygii.

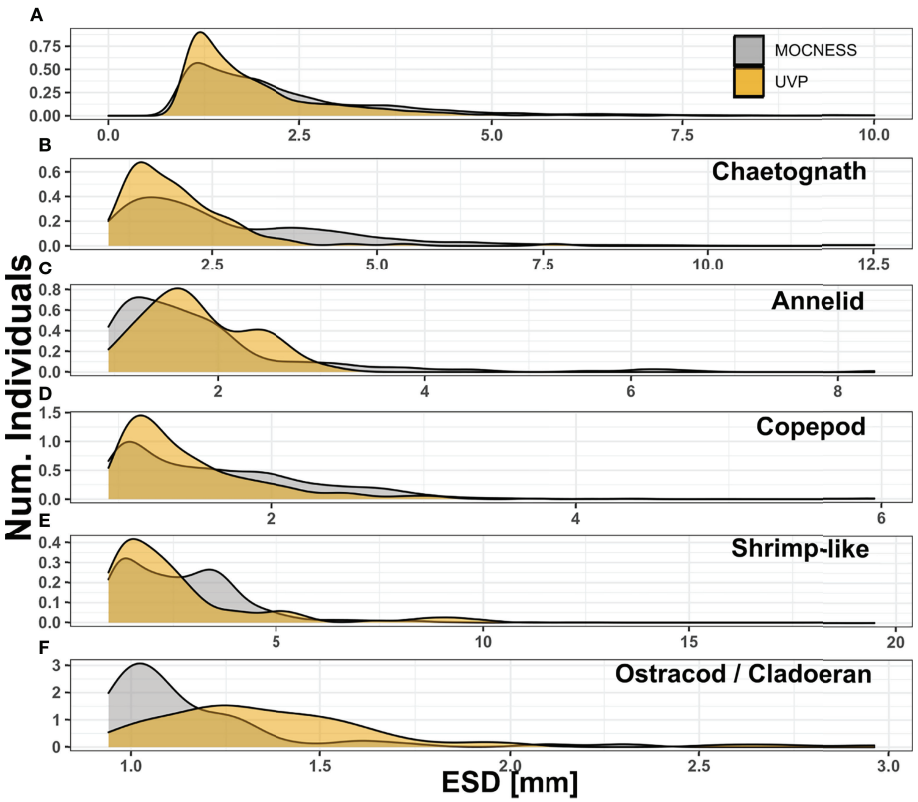


FIGURE 2 | Size distribution compared between MOCNESS-collected plankton (excluding those smaller than 934μm) and UVP-imaged plankton for **(A)** All living organisms, **(B)** chaetognaths, **(C)** annelids, **(D)** copepods, **(E)** shrimp-like crustaceans, **(F)** ostracods/cladocerans. For all living organisms, those larger than 10mm ESD were excluded to allow for visualization. Notes that between each panel y and x axis differ.

shifted upwards and had little overlap with the MOCNESS distributions (**Figures 2C, F**).

Even with the size trimmed data, the MOCNESS tow recorded a much larger diversity of taxa than the UVP (**Figure 3**). Across all MOCNESS tows, copepods were proportionally the most abundant organisms, representing a third of all recorded organisms (**Figure 3**). The proportion of copepods observed by the UVP was slightly smaller, at 27% of all living organisms across all casts (**Figure 3**; **Supplemental Information**). Of all recorded particles from UVP casts, approximately 84.5% were detritus or unidentifiable particles. Among living organisms, Rhizaria and *Trichodesmium* made up

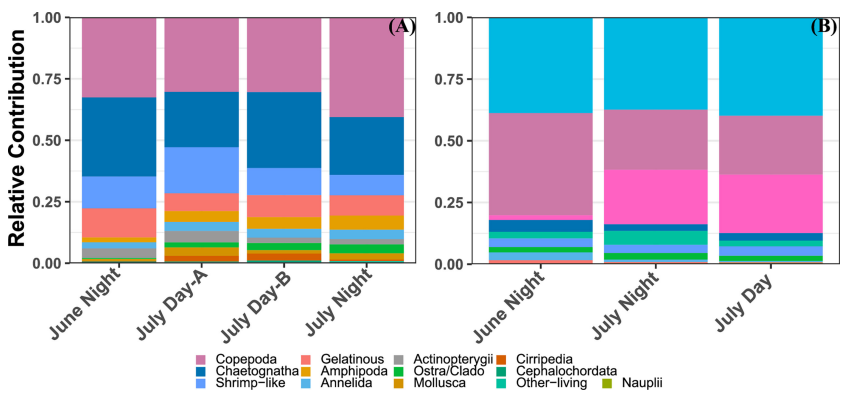


FIGURE 3 | Relative contribution of different zooplankton groups to the total living abundance from each profile from **(A)** the MOCNESS/ZooScan plankton above the 934μm size cut-off and **(B)** the UVP.

39.4% and 17.5% of all UVPs observations respectively. Interestingly, *Trichodesmium* abundance was greatly varied between the two cruises (**Figure 3**). The MOCNESS tows did not record any *Trichodesmium* while it did record a few rhizaria cells. Typically, rhizaria cells sampled by the MOCNESS were small acanthareans or foraminiferans. While both these taxa are sampled by the UVP, *in situ* vignettes collected by the UVP reveal a much larger diversity of Rhizaria, including many large phaeodarians and radiolarians. Mollusca, generally pteropods, were a sizeable portion of MOCNESS sampled organisms however they were not sampled adequately by the UVP (**Figure 3**). Organisms which were sampled by both instruments in sizeable numbers were copepods, shrimp-like crustaceans, chaetognaths, ostracods/cladocerans, and annelids. It should be noted that in UVP casts, shrimp-like crustacean and chaetognath proportions were roughly equivalent (3.2% and 3.4% respectively) (**Figure 3**). However, in MOCNESS samples, the chaetognath proportion was generally over double shrimp-like crustacean (**Figure 3**).

3.2 Comparison of Density Estimates Between Sampling Methods

In general, the UVP had much lower estimates of abundance across all five investigated taxa (**Figure 4** and **Supplemental Figure 6**). However, the average-cast approach revealed that there was large variation between individual casts. Generally, the pooled-cast approach and average cast approach led to the similar estimates. For taxa which had higher concentrations of individuals (copepods, chaetognaths, and shrimp-like crustaceans), the UVP was able to partially capture vertical patterns (**Figures 4A–C**). However, this result was inconsistent, particularly for chaetognaths and shrimp-like crustaceans for which the UVP had much more variation between casts and did not follow MOCNESS patterns.

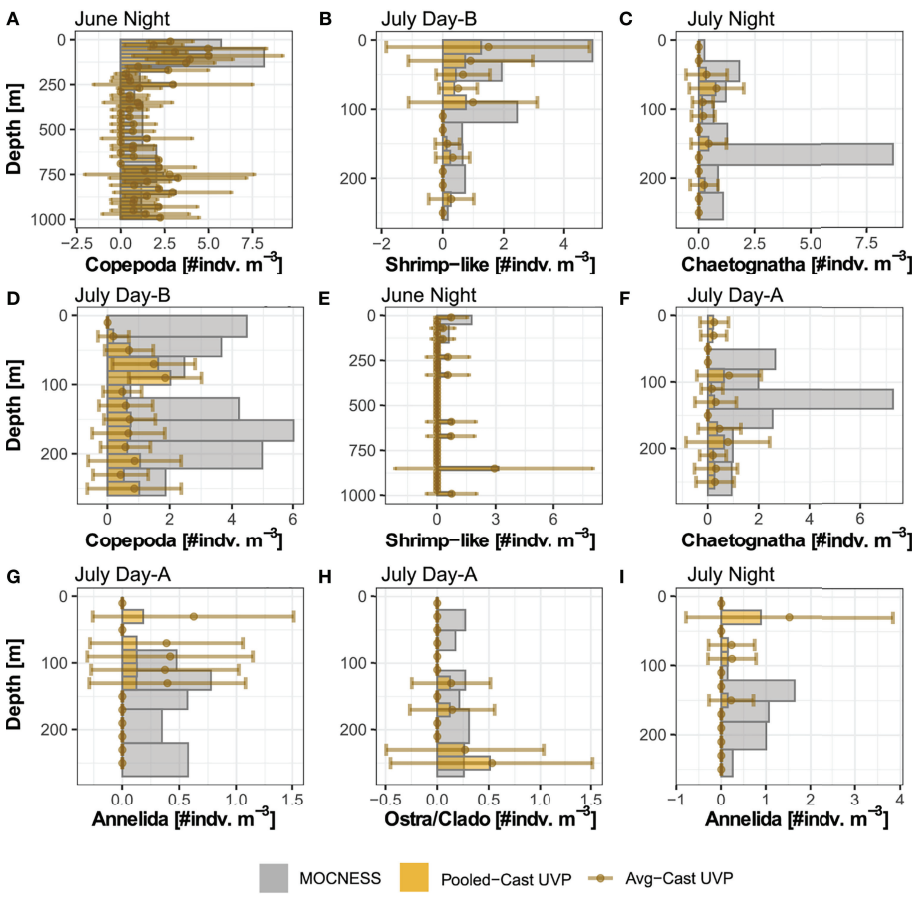


FIGURE 4 | Selected profiles of taxa density estimates by the MOCNESS, pooled-cast UVP, and average-cast UVP calculations. Pooled-cast UVP is shown as bars while average-cast UVP is shown mean points with standard deviation. MOCNESS depth-bins are determined by net deployment while UVP bins are 20m Profiles are selected to show **(A–C)** cases where the vertical pattern of plankton shown by the MOCNESS is captured by the UVP, **(D–F)** cases where the UVP does not emulate the MOCNESS, and **(G–I)** cases of annelids and ostracod/cladocerans. **(A)** Copepod density estimates from June night. **(B)** Shrimp-like density estimates from July Day-B. **(C)** Chaetognath density estimates from July Night. **(D)** Copepod from July Day-B. **(E)** Shrimp-like density estimates from June night. **(F)** Chaetognath density estimates from July Day-A. **(G)** Annelid density estimates from July Day-A. **(H)** Ostracod/Cladoceran density estimates from July Day-A. **(I)** Annelid density estimates from July night. All profiles are available in **Supplemental Figure 6**.

(**Figures 4D–F**). For the taxa with lower concentrations (ostracods/cladocerans and annelids), the UVP did not capture the vertical structure of those populations. Interestingly however, the UVP did detect both annelids and ostracods in regions of the water column where the MOCNESS did not (**Figures 4G–I**). Although variation in the average-cast approach for these taxa was very large.

UVP concentrations calculated in matching depth bins to the MOCNESS were analyzed with linear regressions to quantify if there was a predictable pattern of under/over sampling. For shrimp-like crustaceans, there was a significant relationship between the pooled-cast UVP estimates and MOCNESS estimates (b_1 = 0.073, p-value = 0.01, r^2 = 0.21) (**Figure 5I**) and a significant relationship between average-cast UVP estimates and MOCNESS estimates (b_1 = 0.094, p = 0.007, r^2 = 0.23) (**Figure 5J**). However, this relationship appears to be spurious as there is high heteroskedasticity around the regression line, with only a few, influential observations at higher concentrations. For all other taxa, no significant relationships were found between either the pooled-cast UVP or the average-cast UVP and the MOCNESS (**Supporting Information**). In general, UVP estimates fell below the 1:1 line with MOCNESS estimates (**Figure 5**). However, annelids and ostracods/cladocerans had more observations closer to the 1:1 line. Yet, for both those taxa, there were more observations where only one instrument measured any individuals, and the other did not (**Figure 5**).

Once integrating abundance throughout the water column, the taxon-specific estimates were closer between the different methods. For all taxa, the MOCNESS depth integrated abundance was generally larger than the both the pooled-cast and average-cast UVP depth integrated abundances (**Figure 6**). However, this trend was not a statistically significant difference for any taxa (Paired Wilcox sign rank test, p > 0.05; **Supporting Information**). Interestingly, in the UVP integrated abundance estimates in the mesopelagic were much closer, and in cases, higher than the MOCNESS estimates (**Figure 6**). Additionally, it was found for all taxa that there was no significant differences between the pooled-cast and average-cast UVP approach (Paired Wilcox sign rank test, p>0.05; **Supporting Information**).

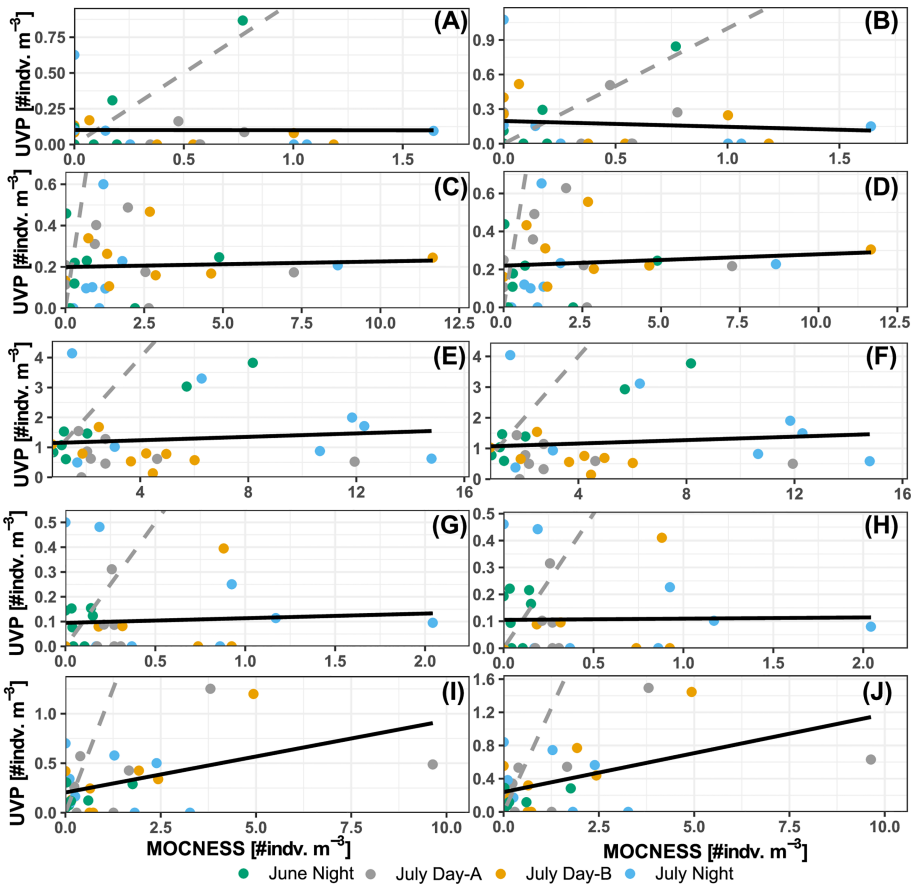


FIGURE 5 | Comparison of MOCNESS density estimates and **(A, C, E, G, I)** pooled-cast UVP density. Comparison of MOCNESS density estimates and **(B, D, F, H, J)** average-cast UVP density estimates. Points show density estimates in matching depth bins for **(A, B)** Annelids, **(C, D)** chaetognaths, **(E, F)** copepods, **(G, H)** shrimp-like crustacean, and **(I, J)** ostracod/cladocerans. Points are colored by corresponding profile. Dotted line represents the 1:1 line between the two estimates. Solid line shows the line of best fit identified through least squares regression.

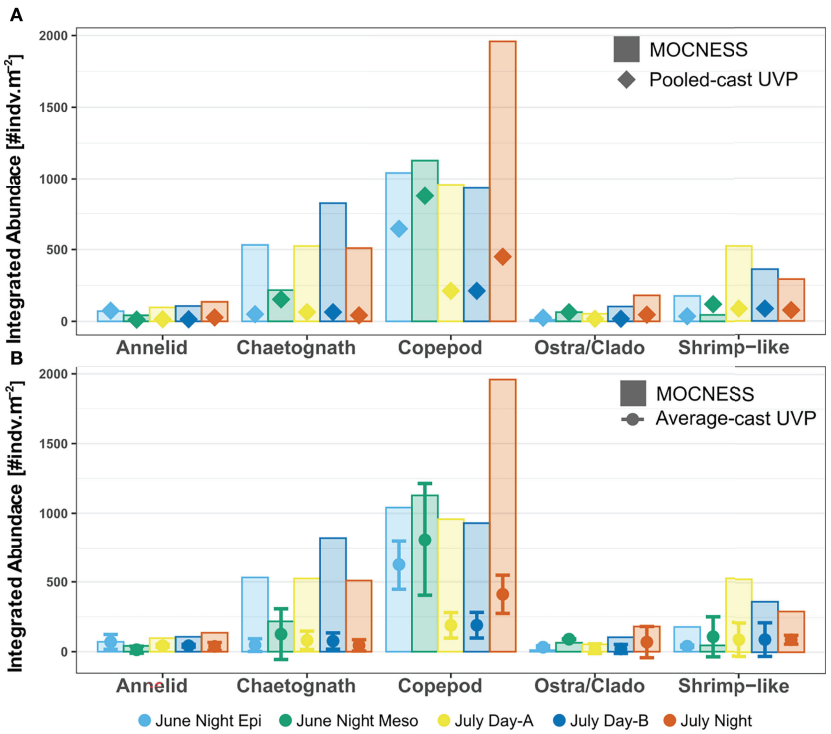


FIGURE 6 | Depth integrated abundance for each specific taxon comparing MOCNESS estimates to **(A)** pooled-cast UVP and **(B)** average-cast UVP calculations. For average-cast calculations, error bars indicate standard deviation between depth integrated abundance of similar UVP profiles. Colors indicate the corresponding integrated profile/tow. Note that the June night tow was integrated as an epipelagic region (0-250) and a mesopelagic region (250-1000). There were no significant differences in integrated abundance for taxon-specific comparisons between the MOCNESS vs pooled-cast UVP; MOCNESS vs average-cast UVP; nor the pooled-cast vs average-cast calculation (Paired Wilcox sign rank test, $p > 0.05$).

3.3 Comparison of Biomass Calculation Between Sampling Methods

Biomass concentration (mg m^{-3}) was then estimated for annelids, chaetognaths, copepods, and shrimp-like crustaceans. Similar to the abundance profiles, the MOCNESS generally had larger biomass concentrations than both UVP methods in comparable areas of the water column (**Figure 7**). The biomass concentration for shrimp-like crustacean and chaetognaths were extremely variable between UVP casts (**Figures 7B, C**). There was no significant relationship of chaetognath biomass concentration estimates between the pooled-cast UVP and MOCNESS nor between the average-cast UVP and MOCNESS (**Figures 8A, B**; **Supporting Information**). However, there were significant relationships between both the UVP methods and the MOCNESS estimates for biomass concentrations of shrimp-like crustaceans, copepods, and ostracod/cladocerans (**Figures 8C–H**; **Supporting Information**). This finding is particularly surprising, given that there was not a meaningful relationship between the abundance estimates between the two devices. It is likely that the regression slopes between the UVP and MOCNESS are not meaningful despite their statistical significance.

Finally, depth integrated biomass concentration estimates calculated by the UVP methods were close, yet lower than those calculated by the MOCNESS (**Figure 9**). This trend was not significantly different for any of the taxa for both the pooled-cast versus MOCNESS nor the average-cast versus MOCNESS (Paired Wilcox signed-rank test, p-value > 0.05, **Supporting Information**). Additionally, there was no significant difference in depth integrated biomass concentration estimates between either of the UVP methods, for all taxa (Paired Wilcox signed-rank test, p-value > 0.05, **Supporting Information**).

4 DISCUSSION

4.1 Scope of Instruments

Generally, the MOCNESS/ZooScan produces higher quality images allowing for superior taxonomic resolution compared to the UVP. Use of the MOCNESS is necessary to sample a large portion of the copepod and ostracod community in the oligotrophic ocean which are not able to be sampled by the UVP due to their small size. Once looking at comparable size ranges however, the UVP and MOCNESS had copepods represent a similar proportion of the total sampled organisms. However, aside from copepods, the relative abundance of taxa varied between the devices. The, MOCNESS's next largest categories of sampled plankton were chaetognaths and shrimp-

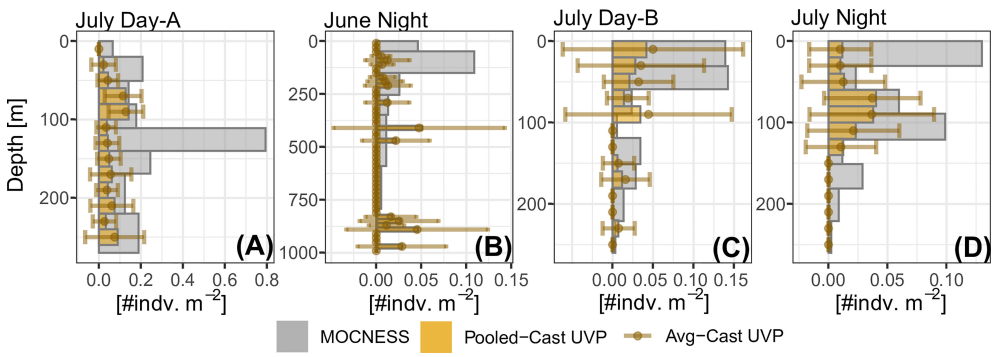


FIGURE 7 | Select profiles of taxa biomass concentration calculated from MOCNESS, pooled-cast UVP approach, and average-cast UVP approach. Profiles are selected to show examples for **(A)** copepods, **(B)** chaetognaths, **(C)** shrimp-like crustaceans, and **(D)** ostracod/cladocerans. All profiles can be seen in **Supplemental Figure 7**.

like crustaceans. The UVP's other common organisms by abundance were Rhizaria and *Trichodesmium*. Fragile taxa such as these are likely destroyed in the MOCNESS and formalin preservation process and thus under sampled by net-based methods. MOCNESS destruction or fragmentation of annelids could also explain why the UVP sampled a seemingly larger size distribution of that taxa. The effectiveness of the UVP for sampling such fragile taxa have been demonstrated in previous studies (Biard et al., 2016; Stukel et al., 2019). Additionally, while in the present study, all rhizaria taxa were

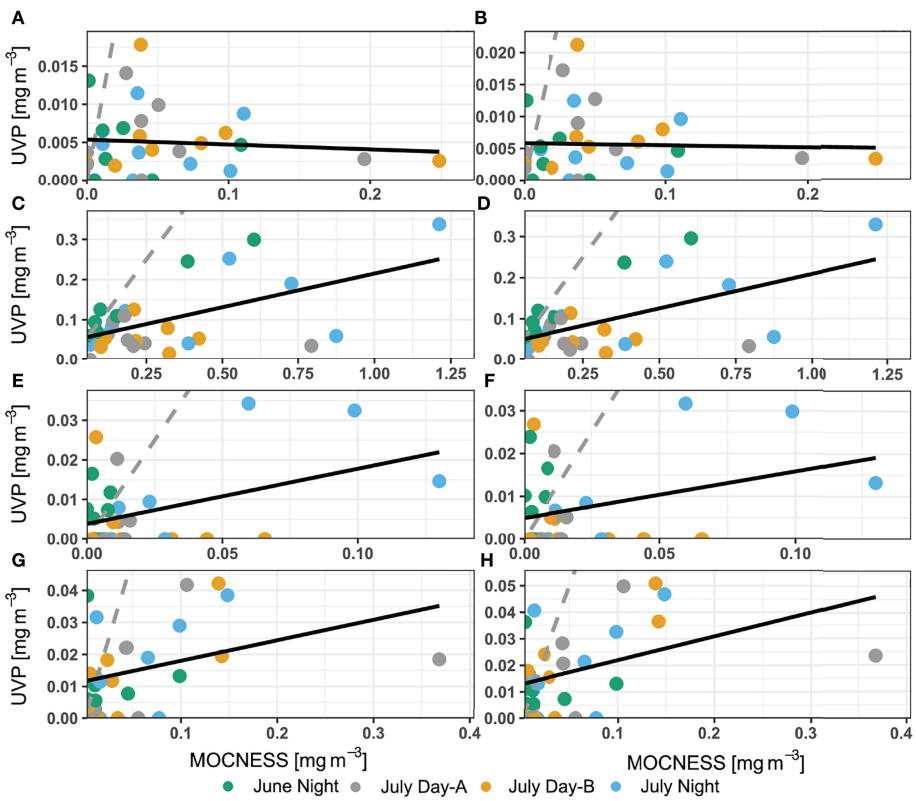


FIGURE 8 | Comparison of biomass concentration estimates between MOCNESS to **(A, C, E, G)** pooled-cast UVP method and between MOCNESS to **(B, D, F, H)** average-cast UVP method. Taxon-specific regressions are for **(A, B)** chaetognaths, **(C, D)** copepods, **(E, F)** ostracods/cladocerans, and **(G, H)** shrimp-like crustaceans. Dotted grey line indicates 1:1 relationship while the solid grey line identifies the line of best fit determined by ordinary least squares regression.

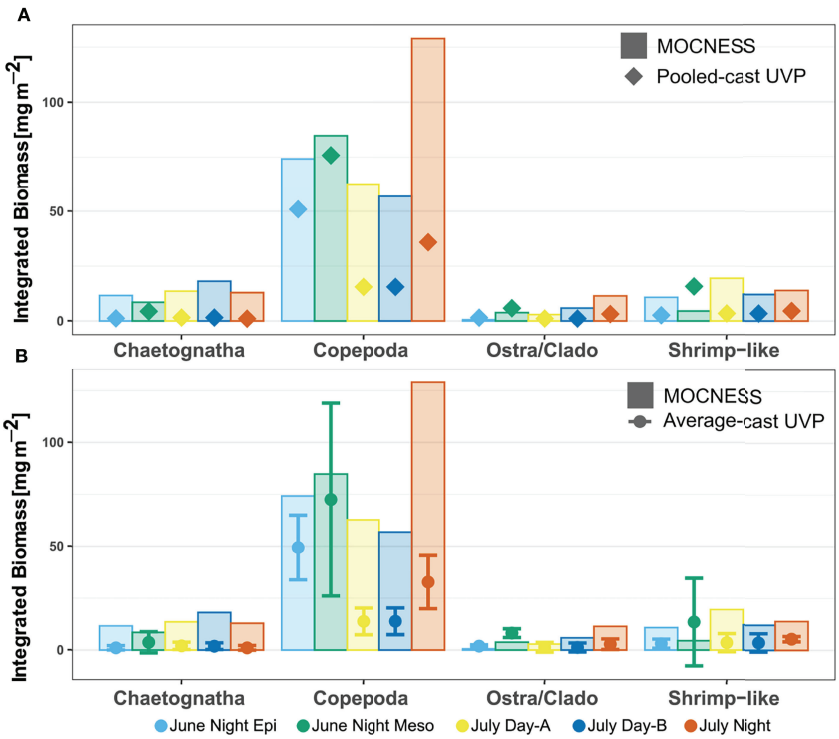


FIGURE 9 | Depth integrated biomass for each specific taxon comparing estimates from MOCNESS to **(A)** pooled-cast UVP calculations and to **(B)** average-cast UVP calculations. Average-cast UVP calculations show standard deviation between similar UVP casts. There were no significant differences in the taxon-specific estimated depth integrated biomass between MOCNESS to pooled-cast UVP, MOCNESS to average-cast UVP, nor pooled-cast to average-cast UVP.

grouped broadly, it is possible to use UVP images to categorize rhizaria into families (**Figure 1**; Biard et al., 2016). Interestingly, in our study we found rhizaria to be a higher proportion of all mesozooplankton by abundance compared to previous studies in the region (Biard et al., 2016), indicating that there is likely large variation in rhizaria abundance across some time scales.

The MOCNESS/ZooScan also sampled a much larger size range than the UVP did. This was an unsurprising finding given that the ZooScan can record particles above 300µm while the UVP is set to only save vignettes of particles larger than 500µm. However, the UVP vignettes at the smaller sizes (500µm – 1000µm) were too coarse to identify as living organisms. Additionally, another consideration for the size estimates of UVP organisms is that plankton can be oriented in any direction during imaging. Thus, if a plankter is positioned orthogonally to the UVP's camera its true size might be underestimated. Additionally small plankton in such orientations might be difficult to identify. As a result, the smallest UVP-imaged particles identified to be a living organism were at least 934µm and for some taxa, they were over a millimeter. Forest et al. (2012) also observed that the UVP did not sample copepods well below a 1mm. In oligotrophic systems like the Sargasso Sea, there are a large portion of the zooplankton community which is not sampled by the UVP because they are smaller than 900µm (**Supplemental Figure 5**). Certainly, there is a portion of the particles recorded by the UVP which are truly living organisms, yet the images of them are too coarse to distinguish as living organisms. It is likely that the upward shift in the UVP ostracod size distribution was caused by the difficulty of distinguishing smaller ostracods and cladocerans from particles. While the particle data do not allow for taxon-specific density or biomass estimation, this information can be used to characterize communities based on particle size spectra (Sprules and Barth, 2016; Lombard et al., 2019).

While the majority of the taxa sampled by the MOCNESS were smaller than those sampled by the UVP, there were a few taxa which had larger size classes that the UVP did not sample either. This was most notable with the chaetognaths. A sizable portion of the chaetognaths measured by the MOCNESS were 3mm to 7.5mm ESD. The UVP hardly imaged any chaetognaths in this size range. It is likely that larger organisms are able to avoid the CTD-rosette. Many of the chaetognaths which were imaged by the UVP were actively in motion (**Figure 1**). Additionally, many fish and shrimp-like crustaceans measured by the UVP were also in motion. Other *in situ* imaging devices have documented krill showing an escape response when encountered with the device (Hoving et al., 2019). However, Hoving et al. (2019) used a white-light system. For imaging devices to minimize zooplankton response, the device must be designed specifically to reduce disturbance (Ohman et al., 2018).

The UVP5 is equipped with red LED lights, intending to reduce escape response by zooplankton. The cause of plankton disturbance in this study is unclear. The UVP was positioned inside a large CTD rosette, thus it could have been the physical turbulence caused by the large frame initiating the escape response and not the light. Chaetognaths in particular have long been known to rely on tactile rather than visual cues to initiate movement (Horridge and Boulton 1967). Avoidance is also a challenge for net-based systems; however, our results indicate that avoidance is also a potential large issue for studying certain taxa with the UVP.

4.2 Density Estimates and Biomass Calculations

Generally, the UVP underestimated density across all compared taxa. For annelids and ostracods/cladocerans, there are likely too few organisms for the UVP to adequately sample. This is clearly observed in the depth profiles for these taxa which show small, infrequent peaks and high variation in the average-cast UVP profiles (**Figure 4** and **Supplemental Figure 5**). The UVP sampling volume, even with pooled casts, is still too low to adequately sample these sparser zooplankton. Increasing the imaged volume is a critical step for *in situ* optical tools (Cowen and Guigand, 2008; Lombard et al., 2019). Chaetognath, shrimp-like crustaceans, and copepods were all well sampled by the UVP, however the density estimates were still much lower than the MOCNESS. Other studies have used a "relative index" for large copepods sampled by the UVP (Donoso et al., 2017), however our results did not support a clear relationship for estimates of any taxa from the UVP to the MOCNESS. A contributing factor to the under sampling of some of these organisms is likely the mobility of these plankton and their avoidance of the CTD rosette as it descends through the water column. Copepods are decently sampled by the UVP, yet because the UVP only sampled large copepods, it is missing a sizable portion of the oligotrophic copepod community. Large copepods are inherently less abundant and thus require larger volumes filtered to adequately study. Our study did find a reliable relationship between UVP and MOCNESS estimates of biomass concentration for three of the four investigated taxa. However, because we know the UVP is estimating a different number of organisms than the MOCNESS, the biomass concentrations are likely faulty. This indicates that existing biovolume to dry mass relationships for net collected plankton may not be reliable for *in situ* imaged organisms.

Depth integration for both abundance and biomass concentration led to UVP estimates more similar to those of MOCNESS estimates. While estimates in matching depth bins were not similar, the similarity of depth integrated estimates can be explained by a few possibilities. First, depth integration effectively increases the volume filtered by combining several depth bins. Secondly, plankton are patchily distributed throughout the water column so populations of plankton may be a few meters deeper or shallower between nearby profiles. Finally, although our study did not find a significant difference in depth integrated estimates between MOCNESS and either UVP method, this could be a result of the low statistical power from the small sample size. There was a notable trend of lower UVP estimates in paired MOCNESS integrations.

4.3 Conclusion and Recommendations for *In Situ* Imaging

It is clear that the UVP under samples many categories of zooplankton compared to a MOCNESS. In more eutrophic systems, or areas where average body sizes are larger, the *in situ* imaging like will be more effective at estimating zooplankton abundance (Forest et al., 2012; Vilgrain et al., 2021). The mobility and escape response of zooplankton also need to be considered when attempting to characterize large zooplankton populations. *In situ* imaging studies should consider both the light and turbulence disruption caused by the sampling device.

This study identifies several methodological considerations for *in situ* imaging studies. Previous UVP studies have pooled similar casts, however this study shows that there is no significant improvement to pool casts rather than average them. We argue that averaging casts provides more information because the variation between casts is clearly represented. While some variation between casts may be due to the small sampling volume, patchiness can also be characterized for more abundant taxa. Selection of depth bin width to study plankton is also an important consideration. While increasing bin width does increase the volume sampled in a depth region, it sacrifices ecologically relevant information about plankton distributions. However, using too small of bin sizes can be misrepresentative. We encourage authors using *in situ* imaging tools to investigate the smallest reliable bin size to use in their systems (**Supplemental Figure 4**). Finally in our system, estimates of density and biomass were not affected by multiple individuals being imaged twice. However, this finding may not hold true in other systems or if the rate of descent for the UVP is decreased.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Materials**. Further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JS contributed to the design of study and deployment of instruments. AB processed lab samples and managed data and data analysis. JS and AB contributed to the writing and revision of the manuscript.

FUNDING

Funding for this data collection was provided by Simons's Foundation International's as part of BIOS-SCOPE to L.B.B. and A.E.M. Ship time was supported by the Bermuda Atlantic

Time Series Study via the OCE grant #1756105 and NSF OCE grant #1756312.

ACKNOWLEDGMENTS

We thank Rod Johnson, the BATS technicians, and BATS program for facilitating data collection and assisting in UVP piloting. The R/V Atlantic Explorer staff and crew were instrumental in data collection. Ryan Rykaczewski helped the set-up of the UVP and with data collection. Leo Blanco-Bercial, Amy Maas, and Hannah Gossner assisted in zooscanning of MOCNESS samples and guidance on this project. The content of this manuscript was greatly improved through peer-reviewers' thoughtful comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.898057/full#supplementary-material

All data and code can be accessed at https://github.com/TheAlexBarth/Oligotrophic_in-situ_net_comparison

REFERENCES

Biard, T., Stemmann, L., Picheral, M., Mayot, N., and Vandromme, P.. (2016). In Situ Imaging Reveals the Biomass of Giant Protists in the Global Ocean. *Nature* 532, 504–507. doi: 10.1038/nature17652

Christiansen, S., Hoving, H.-J., Schütte, F., Hauss, H., Karstensen, J., Kortzinger, A., et al. (2018). Particular Matter Flux Interception in Oceanic Mesoscale Eddies by the Polychaete *Poeobius* Sp. *Limnol. Oceanogr.* 63, 2093–2109. doi: 10.1002/lno.10926

Cowen, R. K., and Guigand, C. M. (2008). *In Situ* Ichthyoplankton Imaging System (ISIIS): System Design and Preliminary Results. *Limnol. Oceangr.: Methods* 6, 126–132. doi: 10.4319/lom.2008.6.126

Davis, C. S., Gallage, S. M., Berman, M. S., Haury, L. R., and Strickler, J. R. (1992). The Video Plankton Recorder (VPR): Design and Initial Results. *Arch. Hydrobiol. Beih. Ergebn. Limnol.* 36, 1651–1653.

Donoso, K., Carlotti, F., Pagano, M., Hunt, B. P. V., Escribano, R., and Berline, L. (2017). Zooplankton Community Response to the Winter 013 Deep Convection Process in the NW Mediterranean Sea. *J. Geophys. Res.: Ocean.* 122, 2319–2338. doi: 10.1002/2016JC012176

Durrieu de Madron, X., Ramondenc, S., Berline, L., Houpret, L., Bosse, A., Martini, S., et al. (2017). Deep Sediment Resuspension and Thick Nepheloid Layer Generation by Open-Ocean Convection. *J. Geophys. Res.: Ocean.* 122, 2291–2318. doi: 10.1002/2016JC012062

Forest, A., Babin, M., Stemmann, L., Picheral, M., Sampie, M., Fortier, L., et al. (2013). Ecosystem Function and Particle Flux Dynamics Across the Mackenzie Shelf (Beaufort Sea, Artic Ocean): An Integrative Analysis of Spatial Variability and Biophysical Forcings. *Biogeosciences* 10 (5), 2833–2866. doi: 10.5194/bg-10-2833-2013

Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., et al. (2012). Size Distribution of Particles and Zooplankton Across the Shelf-Basin System in Southeast Beaufort Sea: Combined Results From an Underwater Vision Profiler and Vertical Net Tows. *Biogeosciences* 9 (4), 1301–1320. doi: 10.5194/bg-9-1301-2012

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital Zooplankton Image Analysis Using the ZooScan Integrated System. *J. Plankton. Res.* 322, 285–303. doi: 10.1093/plankt/fbp124

Guidi, L., Calil, P. H. R., Duhamel, S., Bjorkman, K. M., Doney, S. C., Jackson, G. A., et al. (2012). Does Eddy-Eddy Interaction Control Surface Phytoplankton Distribution and Carbon Export in the North Pacific Subtropical Gyre? *J. Geophys. Res.: Biogeosci.* 117, G02024. doi: 10.1029/2012JG001984

Hauss, H., Christiansen, S., Schütte, F., Kiko, R., Lima, M. E., Rodrigues, E., et al. (2016). Dead Zone or Oasis Iin the Open Ocean? Zooplankton Distribution and Migration in Low-Oxygen Modewater Eddies. *Biogeosciences* 13 (6), 1977–1989. doi: 10.5194/bg-13-1997-2016

Herman, A. W., Beanlands, B., and Phillips, E. F. (2004). The Next Generation of Optical Plankton Counter: The Laser-OPC. *J. Plankton. Res.* 10 (26), 1135–1145. doi: 10.1093/plankt/fbh095

Holliday, D. V., Donaghay, P. L., Greenlay, C. F., McGehee, D. E., McManus, M. M., Sullivan, J. M., et al. (2003). Advances in Defining Fine- and Micro-Scale Pattern in Marine Plankton. *Aquat. Liv. Resour.* 16 (3), 131–136. doi: 10.1016/S0990-7440(03)00023-8

Horridge, G. A., and Boulton, P. S. (1967) Prey Detection by Chaetognatha *via* Vibration Sense. *Proc. Roy. Soc B.* 168, 413–419. doi: 10.1098/rspb.1967.0072

Hoving, H.-J., Christianesn, S., Fabrizius, E., Hauss, H., Kiko, R., Linke, P., et al. (2019). The Pelagic *In Situ* Observation System (PELAGIOS) to Reveal Biodiversity, Behavior, and Ecology of Elusive Oceanic Fauna. *Ocean. Sci.* 15, 1327–1340. doi: 10.5194/os-15-1327-2019

Hoving, H.-J., Neitzel, P., Hauss, H., Christiansen, S., Kiko, R., Robison, B. H., et al. (2020). *In Situ*Observations Show Vertical Community Structure of Pelagic Fauna in the Eastern Tropical North Atlantic Off Cape Verde. *Sci. Rep.* 10, 10.21798. doi: 10.1038/s41598-020-780255-9

Jouandet, M.-P., Jackson, G. A., Carlotti, F., Picheral, M., Stemmann, L., and Blain, S. (2014). Rapid Formation of Large Aggregates During the Spring Bloom of Kerguelen Island: Observations and Model Comparisons. *Biogeosciences* 11, 4393–4406. doi: 10.5194/bg-11-4393-2014

Kiorboe, T. (2011). What Makes Pelagic Copepods So Successful? *J. Plank. Res.* 33 (5): 677–685. doi: doi: 10.1093/plankt/fbq159

Kiørboe, T., Visser, A., and Andersen, K. H. (2018). A Trait-Based Approach to Ocean Ecology. *ICES. J. Mar. Sci.* 75, 1849–1863. doi: 10.1111/j.1365-2427.2009.02298.x

Kiko, R., Biastoch, A., Brandt, P., Cravatte, S., Hauss, H., Hummels, R., et al. (2017). Biological and Physical Influences on Marine Snowfall at the Equator. *Nat. Geosci.* 10 (11), 852–858. doi: 10.1038/NGEO3042

Litchman, E., Ohman, M. D., and Kiørboe, T. (2013). Trait-Based Approaches to Zooplankton Communities. *J. Plankton. Res.* 35 (3), 473–484. doi: 10.1093/plankt/fbt019

Lombard, F., Boss, E., Waite, A.M., Vogt, M., and Uitz, J.. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6: 196. doi: 10.3389/fmars.2019.00196

Maas, A. E., Gossner, H., Smith, M. J., and Blanco-Bercial, L. (2021). Use of Optical Imaging Datasets to Assess Biogeochemical Contributions of the Mesozooplankton. *J. Plank. Res.* 43 (3), 475–491. doi: 10.1093/plankt/fbab037

Martin, P., van der Loeff, M. R., Cassar, N., Vandromme, P., d'Ovidio, F., Stemmann, L., et al. (2013). Iron Fertilization Enhanced Net Community Production But Not Downward Particle Flux During the Southern Ocean Iron Fertilization Experiment LOHAFEX. *Global Biogeochem. Cycles.* 27, 871–881. doi: 10.1002/gbc.20077

Miquel, J.-C., Gasser, B., Martin, J., Marec, C., Babin, M., Fortier, L., et al. (2015). Downward Particle Flux and Carbon Export in the Beaufort Sea, Artic Ocean; the Role of Zooplankton. *Biogeosciences* 12, 5103–5117. doi: 10.5194/bg-12-5103-2015

Motoda, S. (1959). Devices of Simple Plankton Apparatus. *Memoirs of the Faculty of Fisheries Hokkaido University*. 7(1-2): 73–94

Nayak, A. R., Malkiel, D., McFarland, M. N., Twardowski, M. S., and Sullivan, J. M. (2021). A Review of Holography in the Aquatic Sciences: *In Situ* Characterization of Particles, Plankton, and Small Scale Biophysical Interactions. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.572147

Ohman, M. D. (2019). A Sea of Tentacles: Optically Discernible Traits Resolved From Planktonic Organisms *in Situ*. *ICES. J. Mar. Sci.* 76 (7), 1959–1972. doi: 10.1093/icesjms/fsz184

Ohman, M. D., Davis, R. E., Sherman, J. T., Grindley, K. R., Whitmore, B. M., Nickels, C. F., et al. (2018). *Zooglider*: An Autonomous Vehicle for Optical and Acoustic Sensing of Zooplankton. *Limnol. Oceangr.: Methods* 17), 69–86. doi: 10.1002/lom3.10301

Orenstein, E., Ayata, S.-D., Maps, F., Biard, T., Becker, E., Benedetti, F., et al. (2021). Machine Learning Techniques to Characterize Functional Traits of Plankton From Image Data. *HAL Open Sci*. hal-03482282.

Orenstein, E. C., Ratelle, D., Briseño-Avena, C., Carter, M. L., Franks, P. J. S., Jaffe, J. S., et al. (2020). The Scripps Plankton Camera System: A Framework and Platform for *in Situ* Microscopy. *Limnol. Oceanogr.: Methods* 18, 681–695. doi: 10.1002/lom3.10394

Picheral., M., Catalano, C., Brousseau, D., Claustre, H., and Coppola, L.. (2022). The Underwater Vision Profiler 6: An Imaging Sensor of Particle Size Spectra and Plankton, for Autonomous and Cabled Platforms. *Limnol. Oceanogr.: Methods* 20: 114–129. doi: 10.1002/lom3.10475

Picheral, M., Catalano, C., Brousseau, D., Claustre, H., Coppola, L., Leymarie, E., et al. (2021). The Underwater Vision Profiler 6: An Imaging Sensor of Particle Size Spectra and Plankton, for Autonomous and Cabled Platforms. *Limnol. Oceangr.: Methods* 20), 1775–1778. doi: 10.1002/lom3.10475

Picheral, M., Colin, S., and Irrison, J.-O. (2017) *EcoTaxa, a Tool for the Taxonomic Classification of Images*. Available at: http://ecotaxa.obs-vlfr.frhttp://ecopart.obs-vlfr.fr.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Ghizlaine, I., and Gorsky, G. (2010). The Underwater Vision Profiler 5: An Advanced Instrument for High Spatial Resolution Studies of Particle Size Spectra and Zooplankton. *Limnol. Oceanogr.: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462

Puig, P., de Madron, X. D., Salat, J., Schroeder, K., Martin, J., Karageorgis, P., et al. (2013). Thick Bottom Nepheloid Layers in the Western Mediterranean Generated by Deep Dense Shelf Water Cascading. *Prog. Oceanogr.* 111, 1–23. doi: 10.1016/j.pocean.2012.10.003

Sandel, V., Kiko, R., Brandt, P., Dengler, M., Stemmann, L., Vandromme, P., et al. (2015). Nitrogen Fuelling of the Pelagic Food Web of the Tropical Atlantic. *PLoS-One* 10 (6), e0131258. doi: 10.1371/journal.pone.0131258

Schulz, J., Barz, K., Ludtke, A., Zielinski, O., Mengedoht, D., and Hirche, H.-J. (2010). Imaging of Plankton Specimens With the Lightframe on-Sight Keyspecies Investigation (LOKI) System. *J. Euro. Opt. Soc* 5, 100175. doi: 10.2971/jeos.2010.10017s

Severin, T., Kessouri, F., Rembauville, M., Sanchez-Perez, E. D., Oriol, L., Caparros, J., et al. (2017). Open-Ocean Convection Process: A Driver of the Winter Nutrient Supply and the Spring Phytoplankton Distribution in the Northwestern Mediterranean Sea. *J. Geophys. Res.: Ocean.* 122, 4587–4601. doi: 10.1002/2016JC012664

Sprules, W. G., and Barth, L. E. (2016). Surfing the Biomass Size Spectrum: Some Remarks on History, Theory, and Application. *C. J. Fish. Aquati. Sci.* 73, 477–495. doi: 10.1139/cjfas-2015-0115

Steinberg, D. K., Carlson, C. A., Bates, N. R., Johnson, R. J., Michaels, A. F., and Knap, A. H. (2001). Overview of the US JGOFS Bermuda Atlantic Time-Series Study (BATS): A Decade-Scale Look at Ocean Biology and Biogeochemistry. *Deep-sea. Res. II.* 48, 1405–1447. doi: 10.1016/S0967-0645(00)00148-X

Steinberg, D.K., and Landry, M.R. (2017). Zooplankton and the Ocean Carbon Cycle. *Annu. Rev. Mar. Sci.* 9: 413–444. doi: 10.1146/annurev-marine-010814-015924

Stukel, M. R., Ohman, M. D., Kelley, T. B., and Biard, T. (2019). Feeding and Flux-Feeding Zooplankton as Gatekeepers of Particle Flux Into the Mesopelagic Ocean in the Northeast Pacific. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00397

Thomsen, S., Karstensen, J., Kiko, R., Krahmann, G., Dengler, M., and Engel, A. (2019). Remote and Local Drivers of Oxygen and Nitrate Variability in the Shallow Oxygen Minimum Zone Off Mauritania in June 2014. *Biogeosci. Disc.* 16, 979–998. doi: 10.5194/bg-16-979-2019

Turner, J. S., Pretty, J. L., and McDonnell, A. M. P. (2017). Marine Particles in the Gulf of Alaska Shelf System: Spatial Patterns and Size Distributions From *in Situ* Optics. *Continent. Shelf. Res.* 145, 13–20. doi: 10.1016/j.csr.2017.07.002

Vilgrain, L., Maps, F., Picheral, M., Babin, M., Aubrey, C., Irrison, J.-O., et al. (2021). Trait-Based Approach Using *in Situ* Copepod Images Reveals Contrasting Ecological Patterns Across an Artic Ice Melt Zone. *Limnol. Oceangr.* 9999, 1–13. doi: 10.1002/lno.11672

Waite, A. M., Stemmann, L., Guidi, L., Calil, P. H. R., Hogg, A. M. C., Feng, M., et al. (2016). The Wineglass Effect Shapes Particle Export to the Deep Ocean in Mesoscale Eddies. *Geophys. Res. Let.* 43, 9791–9800. doi: 10.1002/2015GL066463

Wiebe, P. H., and Benfield, M. C. (2003). From the Hensen Net Toward Four-Dimensional Biological Oceanography. *Prog. Oceanog.* 56 (1), 7–136. doi: 10.1016/S0079-6611(02)00140-4

Wiebe, P. H., Burt, K. H., Boyd, S. H., and Morton, A. W. (1976). A Multiple Opening/Closing Net and Environmental Sensing System for Sampling Zooplankton. *J. Mar. Res.*, 34 (3), 1822–1824.

Wiebe, P. H., Morton, A. W., Bradley, A. M., Backus, R. H., Craddock, J. E., Barber, V., et al. (1985). New Development in the MOCNESS, an Apparatus for Sampling Zooplankton and Micronekton. *Mar. Biol.* 87, 313–323. doi: 10.1007/BF00397811

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Marine Science

ORIGINAL RESEARCH
published: 05 July 2022
doi: 10.3389/fmars.2022.916025

Basin-Scale Underway Quantitative Survey of Surface Microplankton Using Affordable Collection and Imaging Tools Deployed From *Tara*

Zoé Mériguet[1], Anna Oddone[2], David Le Guen[2], Thibaut Pollina[2], Romain Bazile[2], Clémentine Moulin[2,3,4], Romain Troublé[3,4], Manu Prakash[5], Colomban de Vargas[2,3,6]* and Fabien Lombard[1,2,3]**

[1] Laboratoire d'Océanographie de Villefranche, UMR 7093, Sorbonne Université, CNRS, Villefranche sur mer, France, [2] Plankton Planet Non Governmental Organization (NGO), Station Biologique de Roscoff & Atelier PontonZ Morlaix, Roscoff, France, [3] Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/Tara GOSEE, Paris, France, [4] Tara Ocean Foundation, Paris, France, [5] Department of Bioengineering, Stanford University, California, Stanford, CA, United States, [6] Station Biologique de Roscoff, UMR 7144, Sorbonne Université, CNRS, ECOMAP, Roscoff, France

OPEN ACCESS

Edited by:
Griet Neukermans,
Ghent University, Belgium

Reviewed by:
Bob Brewin,
University of Exeter, United Kingdom
Hiroyuki Yamamoto,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC) Japan
Enrique Nogueira,
Spanish Institute of Oceanography (IEO), Spain

***Correspondence:**
Zoé Mériguet
zoe.meriguet@gmail.com
Colomban de Vargas
vargas@sb-roscoff.fr
Fabien Lombard
lombard@obs-vlfr.fr

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 08 April 2022
Accepted: 02 June 2022
Published: 05 July 2022

Citation:
Mériguet Z, Oddone A, Le Guen D, Pollina T, Bazile R, Moulin C, Troublé R, Prakash M, de Vargas C and Lombard F (2022) Basin-Scale Underway Quantitative Survey of Surface Microplankton Using Affordable Collection and Imaging Tools Deployed From Tara. Front. Mar. Sci. 9:916025. doi: 10.3389/fmars.2022.916025

World ocean plankton quantitative biodiversity data are still severely limited due to the high cost and logistical constraints associated to oceanographic vessels and collection/analytic devices. Here, we report the first use of an affordable and open-source plankton collection and imaging kit designed for citizen biological oceanography, composed of a high-speed surface plankton net, the *Coryphaena*, together with a portable in-flux automated imaging device, the *PlanktoScope*. We deployed this kit in December 2020 along a latitudinal transect across the Atlantic Ocean on board the schooner *Tara*, during the first Leg of her 'Mission Microbiomes'. The citizen-science instruments were benchmarked and compared at sea to state-of-the-art protocols applied in previous *Tara* expeditions, i.e. on-board water pumping and filtration system and the FlowCam to respectively sample and image total micro-plankton. Results show that the *Coryphaena* can collect pristine micro-plankton at speed up to 11 knots, generating quantitative imaging data comparable to those obtained from total, on-board filtered water, and that the *PlanktoScope* and FlowCam provide comparable data. Overall, the new citizen tools provided a complete picture of surface micro-plankton composition, biogeography and biogeochemistry, opening the way toward a global, cooperative, and frugal plankton observatory network at planetary scale.

Keywords: citizen sciences, microplankton, *Tara* Mission microbiomes, *Coryphaena net*, *PlanktoScope*, global ecology

1 INTRODUCTION

The oceans are home to a large diversity of planktonic organisms. The sensitivity of these organisms to their environment makes them exceptional sentinels of environmental changes, such as temperature rise (Beaugrand, 2005), or variation in currentology (Borkman and Smayda, 2009). Due to the non-linear response of plankton to environmental changes, plankton reaction to subtle environmental variations can be amplified, making plankton a potentially better indicator of environmental change than the environmental variables themselves (Taylor et al., 2002). Response of plankton to

environmental changes is also rapid, due to the relatively short life cycle of phytoplankton (order of days) when compared with terrestrial plants (order of years to decades). For these reasons, plankton are often referred to as essential oceanic variables (EOV) and essential climate variables (ECV; Global Ocean Observing System; Global Climate Observing System; Bax et al., 2019). Sub-surface (<5m depth) planktonic communities are particularly sensitive to climate change (Bopp et al., 2013), while also being critical actors of biogeochemical cycles (Falkowski et al., 2008). Indeed, these communities face different environmental constraints than plankton thriving in deeper layers, notably in tropical oceans where water column stratification (thermocline/pycnocline) generates a barrier to nutrients upflow from the deep sea, and will increase in our warming world (IPCC 2022). Consequently, sub-surface plankton are more dependent on land or atmospheric inputs (e.g. aerosols, diazotrophy), and serve as a gateway to various nutrient inputs essential to the structuring of epipelagic planktonic ecosystems. Ocean surface layers are also a place of increased environmental stress for plankton such as waves, winds, and solar radiations. Therefore, the processes controlling the abundance and diversity of surface plankton may be significantly different from those observed for biota living in deeper layers (Ibarbalz et al., 2019).

Monitoring (sub)surface plankton in a global change context would require repeated, systematic, large-scale and high-resolution observations, a task that is hardly achievable with oceanographic vessels, which are too expensive to be used for this purpose (the operational cost of an ocean research vessel reaches typically >US$30,000 per day, excluding the cost of scientists, engineers, and the research itself; Lauro et al., 2014). On the other hand, thousands of sailing boats and larger vessels are permanently crossing the oceans, and could be used to this end. Brewin et al. (2017) demonstrated the potential for increased oceanographic data by exploiting these other vessels. A first example of this approach is the Continuous Plankton Recorder, which has generated a successful network of observations through cargo boats over the last 81 years (Batten et al., 2019). A complementary approach consists in engaging citizen sailors in the collection of planetary plankton, such as the ones engaged in the Indian Ocean (Lauro et al., 2014) or more globally at planetary scale in the 'Plankton Planet' initiative (de Vargas et al., 2020). Citizen science strategies require frugal, affordable, and scientifically-sound instruments, sufficiently agile and robust to be used by non-scientists.

We achieved a proof-of-principle for citizen oceanography in 2015/16, collaborating with 20 citizen sailors who performed plankton biomass sampling at more than 250 sites spanning the planetary oceans. The dried plankton samples were simply mailed by the sailors to a single laboratory, generating the first global-scale, high-quality DNA metabarcoding overview of plankton (>20µm) populations for a fraction of the putative cost associated to similar spatio-temporal sampling realized by a standard oceanographic vessel (de Vargas et al., 2020). The results of this first experiment were very promising but highlighted two main limitations. Firstly, sailors were asked to slow their boats down to less than two knots in order to deploy classical plankton nets without breaking the mesh. This requests uncomfortable sailing operations impacting the cruising speed, and it was identified as the primary limiting factor for denser sampling. Secondly, sailors expressed frustration for not being able to observe plankton while realizing the biomass-concentration protocol. Indeed, plankton imaging, which provides critical and complementary, quantitative and morphological information (Lombard et al., 2019), was not implemented due to prohibitive costs and complexity of existing instruments. To address these issues and promote large scale collection and monitoring of plankton in the 20-200µm range by sailors, we developed two new frugal tools for citizen oceanography: the '*Coryphaena*', a high-speed net to collect plankton at cruising speeds, and the '*PlanktoScope*', a frugal, microfluidic, quantitative imaging microscope (Pollina et al., 2020).

In this study, we tested the efficiency of both the *Coryphaena* and the *PlanktoScope* against established standards. Along a transect from Lorient (France) to Punta Arenas (Chile) carried out by the schooner *Tara* in December 2020, we compared the *Coryphaena* net to the Decknet system (DN; Gorsky et al., 2010), a suspended, on-board net that filters surface seawater collected by a high flow pump, and the *PlanktoScope* to the FlowCam (Sieracki et al., 1998), a standard flow-imaging system used in plankton research. The abundance, taxonomic and morphological diversity data from surface micro-plankton (20-200µm; analysis were performed in the 50-150µm size range) communities were used to assess the efficiency of each combination of instruments, and demonstrate the power of our new frugal tools for global-scale plankton ecology.

2 MATERIALS AND METHODS

2.1 Sampling Methods

During the trans-Atlantic journey of the schooner *Tara* from Lorient (15/12/2020; France) to Punta Arenas (04/02/2021; Chile), 35 sampling stations were carried out daily (**Figure 2**). On board, two nets allowing sub-surface plankton sampling were deployed: the *Coryphaena* high speed net deployable up to 11 knots, and the Decknet (DN), suspended on the boat's deck and coupled to a high-flow pumping system, validated and used during various previous *Tara* campaigns (Pesant et al., 2015; Gorsky et al., 2019). The DN filters the entire amount of water pumped on board by using a water inlet called the 'Dolphin' (Gorsky et al., 2019). The seawater is pre-filtered through a 2mm metal screen and subsequently concentrated through the DN suspended on the deck (**Figure 1C**). The volume of water concentrated in the DN was measured using a flow meter, ranging from 0.5 to 8 m^3 (see **Supplementary Table II**), depending on local plankton density. The newly designed *Coryphaena* (**Figure 1A**), inspired from the Small Plankton Sampler (Glover 1953; Wiebe and Benfield, 2003), aims at collecting plankton >50µm at cruising speed (i.e. 1 to 11 knots; see **Supplementary Table II**). The *Coryphaena* has a mouth opening of 4 cm, a length of 80 cm, and a lead weight of 750 grams. Preliminary tests had shown that the *Coryphaena* is stable underwater at speeds below 11 knots while collecting seemingly pristine plankton. Higher speeds make it lift out of the water. The design of the *Coryphaena* is based on the use of

(*i*) a 3D printed head that provides good aerodynamics while reducing the flow water into the net, in order to preserve both the net and plankton at high speed collection, (*ii*) a 50µm mesh supported externally by a 200µm mesh providing greater strength (**Figure 1B**), and (*iii*) an impermeable outside skirt increasing the filtration through the mesh by Venturi effect. Due to its small dimension, the placement of a flowmeter in the net was not possible. We therefore calculated the volume filtered, as its theoretical maximum in the absence of backflow, using the initial and final deployment coordinates and the net mouth opening. Wherever possible, two samples (on board DN and *in situ Coryphaena*) were acquired at the same station simultaneously. For practical comparison purposes, it was initially decided to use a DN with a 50µm mesh in contrast to previous *Tara* campaigns (Gorsky et al., 2019). However, as shown by results from the first 10 stations, this configuration led to over-efficient filtration damaging fragile organisms by abrasion on the drained silk. DN results from stations 1 to 10 were thus disregarded. A 20µm DN was thus used for the subsequent stations 11 to 35 while only considering organisms >40µm in the imaging results. A complete replacement of the *Coryphaena* net was carried out at station 21 following its destruction by, presumably, a swordfish.

2.2 Image Acquisition

After collection, plankton from both *Coryphaena* and DN samples were filtered through a 200µm mesh to remove larger organisms which may clog the fluidic system of both the FlowCam and the *PlanktoScope*. The *PlanktoScope* (**Figure 1D**) is a cost-effective microscope (<800€ of hardware parts) allowing quantitative imaging of microplankton (in the 20-200µm size range). Full description and prior quality test are available in a companion article (Pollina et al., this issue). Initial tests generated data of a quality comparable to that produced by the FlowCam, an automated commercial microscope taking digital image of microscopic particles flowing through a capillary imaging chamber (Sieracki et al., 1998). The reliability of medium/high throughput imaging instruments for quantitative analysis of marine plankton is evidenced by a growing number of studies in the scientific community using these methods (Irisson et al., 2022). Notably FlowCam data have been compared and validated against microscopy analyses as regard to organismal size (Sieracki et al., 1998; Buskey and Hyatt, 2006; Ide et al., 2007; Álvarez et al., 2014; Le Bourg et al., 2015) and biovolume (Hrycik et al., 2019).

The four configurations, (1) *Coryphaena* - FlowCam, (2) *Coryphaena* - *PlanktoScope*, (3) DN - FlowCam and (4) DN - *PlanktoScope*, were tested in parallel whenever possible (**Figure 2A**). Images generated by the FlowCam were processed using the ZooProcess software (Gorsky et al., 2010), and images generated by the *PlanktoScope* were processed using a custom-made equivalent script in Matlab, a prototype of the segmentation script currently encoded into the *PlanktoScope* computer (see https://github.com/PlanktoScope/PlanktoScope). This allows similar extraction of the segmented objects as vignettes, together with a series of morphometric measurements that are then imported into the EcoTaxa web platform (http://ecotaxa.obs-vlfr.fr) for taxonomic classification. The taxonomic categories predicted by image recognition algorithms were validated or corrected by a trained taxonomist. Overall, 398, 466 vignettes (88, 465 for DN – FlowCam, 66, 243 for *Coryphaena* – FlowCam, 132, 322 for DN – *PlanktoScope*, 111, 436 for *Coryphaena* – *PlanktoScope*) were classified into 179 taxa (list **Supplementary Table I**; 34% of taxonomic categories correspond to the genus level, 23% to the species levels and the 43% to the other levels such as class, order or phylum). Examples of images from the

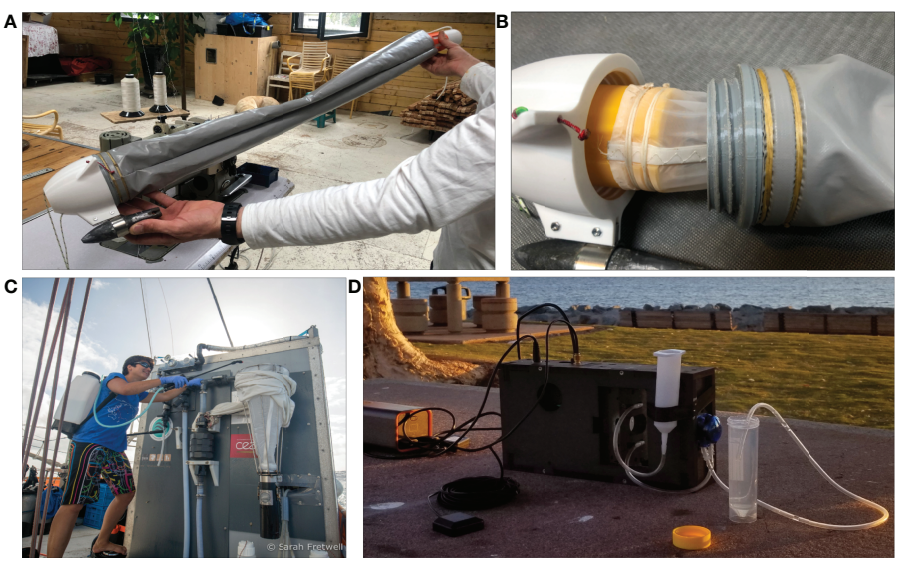


FIGURE 1 | (A) The '*Coryphaena*' high speed net, able to collect plankton >50µm at speed up to 11 knots. **(B)** The 50µm mesh in the *Coryphaena* is supported and protected externally by a 200µm mesh allowing for greater strength, as well as an impermeable skirt (gray) improving the flow of water into the net by Venturi effect. **(C)** The Decknet (DN) pumping and filtration system installed on board *Tara*. **(D)** The *PlanktoScope* allowing quantitative imaging of micro-plankton.

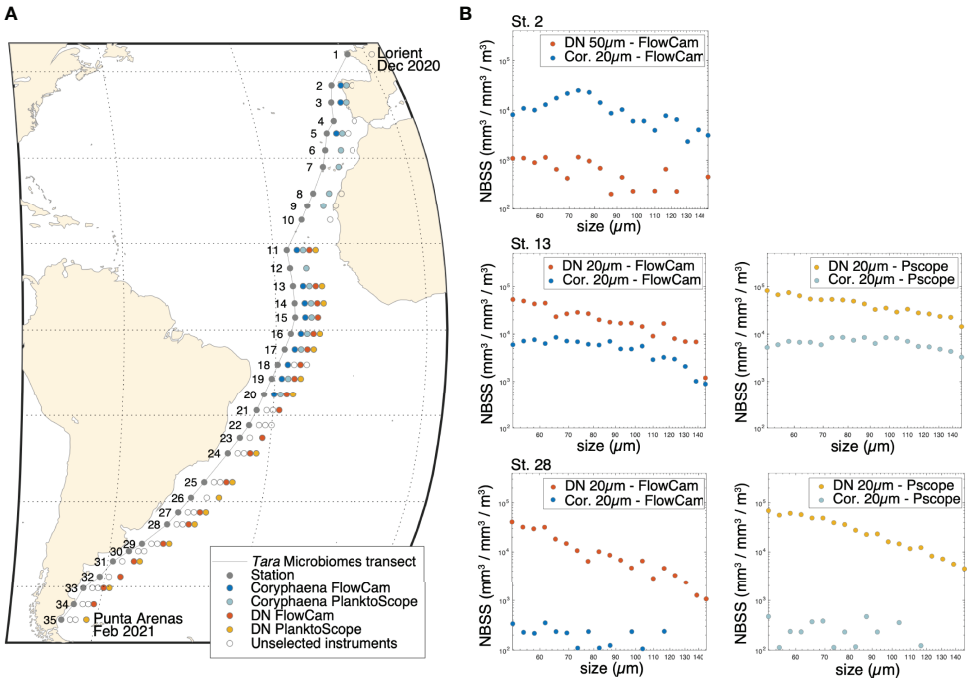


FIGURE 2 | (A) Map of the *Tara* Mission Microbiomes Atlantic transect. The colored dot (see legend) indicates instruments' configuration deployed at each of the 35 daily stations (gray dots). **(B)** Examples of Normalized Biovolume Size Spectra (NBSS) displaying the number of live organisms per size class (in $mm^3\ mm^{-3}\ m^{-3}$) sampled by each net (DN - Decknet; Cor. - *Coryphaena*), at stations 2, 13 and 28. The comparison of the number of live organisms sampled between the two nets is shown for both FlowCam and *PlanktoScope* (Pscope) analyses at stations 13 and 28.

PlanktoScope and FlowCam can be explored and compared in the supplementary material (**Supplementary Figure 2**) as well as on the EcoTaxa web platform (see project link in *Data Availability Statement*).

2.3 Environmental Data

On board *Tara*, surface seawater was continuously pumped through a hull inlet located 1.5m below the waterline and distributed to various instruments such as a ThermoSalinoGraph (TSG; SeaBird Electronics SBE45/SBE38) and a multispectral spectrophotometer (ACS; WETLabs), as performed during the *Tara* Pacific expedition (Gorsky et al., 2019). The ACS measures hyperspectral attenuation and absorption (resolution ~4nm) in the visible and near infrared, allowing notably to derive estimates of chlorophyll-a concentrations. The TSG measures surface temperature and conductivity at a sampling rate of 0.1 Hz. Additional environmental data were extracted from satellite data and/or the copernicus-mercator model (https://marine.copernicus.eu/fr). Satellite data were extracted *via* NASA ocean color (8-day average 4km/pixel) and 12 pixels (50km) around the sampling position and at the date of sampling were averaged to provide a single mean. The environmental data for the mercator model are extracted from marine Copernicus (GLOBAL_ANALYSIS_FORECAST_PHY_001_024-TDS and GLOBAL_ANALYSIS_FORECAST_BIO_001_028-TDS). A single, homogeneous environmental database was created from these multiple sources; missing TSG and ACS data were replaced by satellite data first, then by mercator model data. This database contains: sea surface temperature (SST; °C), salinity (psu), chlorophyll a (Chl; $mg.m^{-3}$), O_2 ($mmg.m^{-3}$), NO_3 ($mmg.m^{-3}$), PO_4 ($mmg.m^{-3}$), Si ($mmg.m^{-3}$), Fe ($mmol.m^{-3}$), particulate inorganic carbon (PIC; $mol.m^{-3}$) and pH, and is available with the associated sources of each environmental value (**Supplementary Table IV**).

2.4 Numerical and Statistical Analysis

For each database, we calculated organismal abundance ($ind.m^{-3}$) and biovolume ($mm^3.m^{-3}$) for each taxa and functional group living *versus* non-living (see **Supplementary Table I**), taking into account the volumes of water filtered by the plankton collecting devices. Major and minor axes of the best ellipsoidal approximation are used to estimate the biovolume ($mm^3.m^{-3}$) of each object following Vandromme (Vandromme et al., 2012). Size is expressed as equivalent spherical diameter (ESD, µm). The individual biovolumes of the organisms are arranged in Normalized Biomass Size Spectra (NBSS) as described by Platt (1978) along an harmonic range of biovolume such as minimal and maximal biovolume of each class are linked such as:

$$Bv_{max} = 2^{0.25} \times Bv_{min}$$

The NBSS is obtained by dividing the total biovolume of each size class by its biovolume interval:

$$Bv_{range} = Bv_{max} - Bv_{min}$$

The NBSS ($mm^3.mm^{-3}.m^{-3}$) is directly proportional to the number of organisms per size class. Biovolume analyses were only performed in the 50-150µm size range due to underestimation of the number of living organisms <50µm induced by undersampling of nets and/or difficulty in taxonomic identification, and misrepresentation of organisms >150µm which were too rare beyond this size (Tranter and Smith, 1968; Pollina et al., 2020). First, we performed a quality control of the instruments to detect putative misfunctioning along the *Tara* transect, using NBSS which can reveal over or under sampling of one net and/or one imaging instrument compared to another. NBSS were also used to establish whether difference of sampling between the two nets affected all size classes similarly. We then used the various observations collected by the different combination of instruments to produce a homogeneous – intercalibrated global overview of plankton at the scale of the Atlantic Ocean. For this, we determined a correction coefficient using NBSS of living organisms in the 50 to 150µm size range. Using the DN-20µm - FlowCam dataset as a reference, we produced a correction coefficient (cross-size classes average correction coefficient) for each station, and further averaged across stations (after checking that no significant effect was visible across stations). After correcting for this sampling efficiency, we further inspected if some residual effect was visible on the species composition. For this a principal component analysis (PCA) was performed on a database that separates the 4 instrumental configurations adjusted with these coefficients. This PCA was performed both using abundance (log+1 transformed) or composition (Hellinger transformed) data. For these analyses, imaging data were clustered both taxonomically (179 taxa identified) and functionally (9 functional groups).

Finally, we used the various correction factors to produce a single cross-calibrated database providing microplankton average abundance and biovolume between the 4 instrumental configurations per station. This synthetic database was used to analyze the global structure of micro-plankton populations at the scale of the Atlantic Ocean. Diversity was calculated with the Shannon index (H) taking into account the 179 taxa identified. Hierarchical clustering analyses (using descriptive complete link method, and Hellinger distance) were performed using the 9 functional groups. Environmental data were integrated into the PCA to assess their impact on taxonomic composition at each station. Spearman correlation tests were performed between different variables (alpha risk set at 0.05%).

A morphological analysis partly based on plankton colors was performed on the vignettes from samples collected with the two nets and imaged with the *PlanktoScope* (the FlowCam model used generates black and white images). As this analysis focuses on the morphological properties of the objects and not their quantity, the difference in sampling between the 2 nets does not induce biases. Only vignettes corresponding to living organisms were considered, while detritus and optical artifacts were discarded. Following previous methodology (Trudnowska et al., 2021; Vilgrain et al., 2021), the data from 15 morphometric measurements were normalized by a non-linear Yeo-Johnson transformation prior to a PCA analysis. Station averages of the morphological values of the PCA axes were then calculated allowing the extraction of morphological metrics at the station scale.

3 RESULTS

3.1 Quality Control and Comparison of the Instruments

3.1.1 Instruments' Quality Control

While *Tara* was cruising southward through the Atlantic Ocean, we used the Normalized Biomass Size Spectra (NBSS, roughly equivalent to organismal abundances per size class) produced by the different plankton collection tools, i.e. the *Coryphaena* and the Decknet, to check and compare their efficiency (**Figure 2**). We first observed a severe under-sampling of the DN-50µm as compared to the *Coryphaena*-20µm from stations 1 to 10. The *Coryphaena* samples were on average 10.21 (± 7.42) more abundant than the DN-50µm samples, regardless of the imaging instrument, **Figure 2A**). Starting from station 11 (**Figure 2A**), we therefore replaced the DN-50µm with a DN-20µm. Between stations 11 and 20, the NBSS from both the DN-20µm and the *Coryphaena* displayed about the same order of magnitudes of abundances (e.g. station 13, **Figure 2B**, see also next section: *Coryphaena and PlanktoScope characterization*). Between station 20 and 21 the initial *Coryphaena* was lost, and the new *Coryphaena* used from station 21 displayed strong under sampling with *Coryphaena*/DN sampling coefficients averaging 0.35 (± 0.76) between stations 21 to 31, regardless of the imaging instrument (e.g. station 28, **Figure 2B**). The *Coryphaena* data were thus not used after station 21. We then compared the results obtained with the *PlanktoScope versus* the FlowCam. Samples imaged with the *PlanktoScope* displayed slightly higher abundances of living organisms than those imaged with the FlowCam (e.g. station 13, **Figure 2B**, and see next section). Data generated from both imaging instruments were used. All values of NBSS spectra per station (station 1 to 35) can be found in **Supplementary Table III**.

3.1.2 Coryphaena and PlanktoScope Characterization

We compared the 4 quality-controlled and filtered databases from the 4 configurations to determine a cross-size classes average correction coefficient between the instruments based on the NBSS biovolumes of living organisms from 50 to 150µm. The correction coefficient between the two nets is equal to 0.35 (standard deviation of 0.34) meaning that the *Coryphaena* under-samples live organisms by about one third compared to the DN. The correction coefficient between the two imaging devices is 1.86 (standard deviation of 1.17), indicating that more live organisms (+86%, ± 17%) were observed in the *PlanktoScope* compared to the FlowCam. The correction factors were applied to the different datasets, and a PCA was used to reveal putative residual effects of the sampling method. The first 3 axes of the PCA were considered and color-coded in RGB to visually inspect coherence between the plankton collection and imaging devices (**Figure 3A**). Per station, the symbols share a similar color (**Figure 3B**) therefore exhibiting similar plankton taxo-functional composition

regardless of instrumental configuration (**Figures 3A, B**). No significant effect of the different modes of sampling/imaging is observed, the variance resulting mainly from geographical difference in taxo-functional composition. Similar PCA analyses based on the 179 identified taxa (dot color corresponds to instrumental configuration; **Figure 3C**) revealed a difference between the two nets, with the *Coryphaena* samples enriched in robust plankton (e.g. *Neoceratium* spp. or *Rhabdonella* spp.) as opposed to fragile ones (e.g. diatoms like *Hemiaulus* or *Eucampia* spp.). However, these differences are only visible in the second axis of the PCA (13% variance explained), suggesting that this bias is essentially concentrated on specific taxa. Even at the scale of 179 taxa, we observed a good agreement between the two imaging instruments (good overlap between *PlanktoScope* and FlowCam points on **Figure 3C**).

3.2 Surface Microplankton Communities in Relation to Environmental Characteristics

By combining the different datasets with the correction factors, we obtained a single homogenized dataset for microplankton along the *Tara* Mission Microbiomes Atlantic transect minimizing biases due the heterogeneous sampling and imaging. Microplankton absolute abundance values vary from a minimum of ca. 2K ind.m^{-3} at station 9 to a 200 times higher maximum of ca. 4Mio ind.m^{-3} at station 23, of which 3.5 Mio ind.m^{-3} (or 3500 cells/L) are diatoms of the genus *Hemiaulus* (**Figure 4B**). The Shannon H indices range from 3.32 (station 19) to 0.54 (station 23) along the *Tara* track (**Figure 4C**) and display a significant inverse correlation (p = 0.0008<0.05; R^2 = -0.59) to absolute abundance. We performed a clustering analysis (descriptive complete link method, Hellinger distance) based on the relative abundances of the 9 plankton taxo-functional groups. Eight clusters of stations emerged based mainly on differences in their diatoms, cyanobacteria, and dinoflagellates composition (**Figure 5A**). These clusters correlate to specific environmental (**Figure 5A**) and biogeographic (**Figures 4**, **5B**) features. The oligotrophic zone (stations 9 to 14) is characterized by microplankton communities dominated by cyanobacteria and associated to high sea surface temperatures (SST) and iron (Fe) concentrations. Conversely, coastal and temperate zones plankton are dominated by diatoms associated with high NO_3 concentrations (stations 2 to 8, 16, 23, 25 to 32 and 35). PO_4-rich areas deprived of iron (stations 15, 17 to 21 and 33) are associated with microplankton communities rich in dinoflagellates.

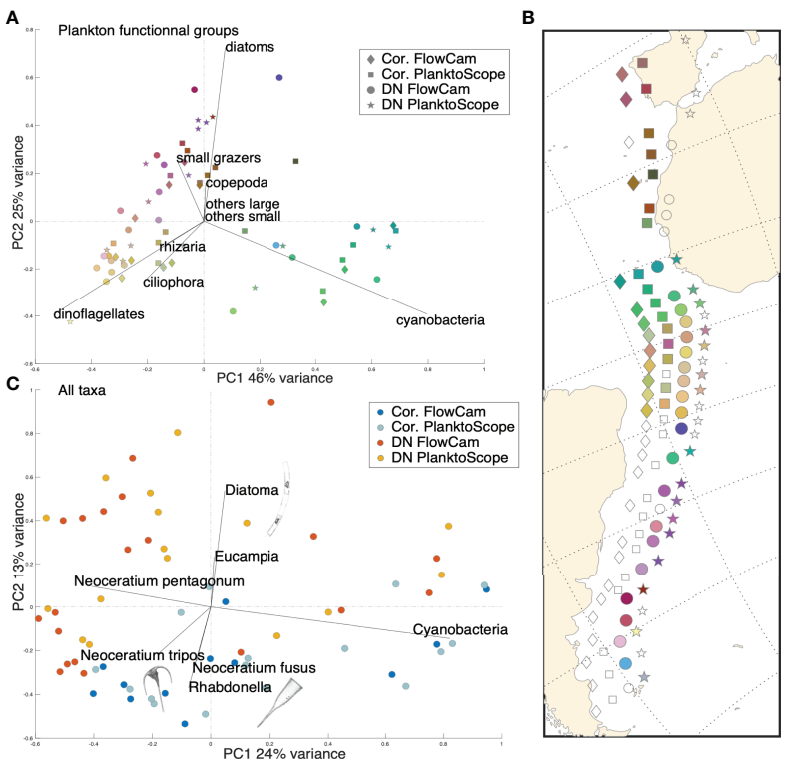


FIGURE 3 | (A) PCA performed on microplankton functional composition (Hellinger transformed data, 9 taxo-functional groups). The 4 instrumental configurations (sampling-imaging devices couples) are represented by the 4 different shapes. Colors are derived from the position of the points on the first 3 axes of the PCA. **(B)** Geographic projection of each point along the *Tara* Atlantic transect (exact position of the stations: diamond shape Cor. FlowCam). The empty shapes represent the instrumental configurations deployed but not selected in our analysis (see **Figure 2** and section: *Instrument's Quality Control*). **(C)** PCA performed on microplankton taxonomic composition (Hellinger transformed data; 179 taxa). The 4 instrumental configurations are represented by the 4 different colors. Only taxa with a contribution greater than 0.23 to the PCA axes are shown for ease of reading. Plankton images taken as examples are from FlowCam; the image close to *'Diatoma'* corresponds to a diatom of the genus Hemiaulus, the image close to *'Rhabdonella'* corresponds to a ciliate of the genus *Rhabdonella* and the image close to *'Neoceratium tripos'* corresponds to a dinoflagellate of the species *Neoceratium tripos*.

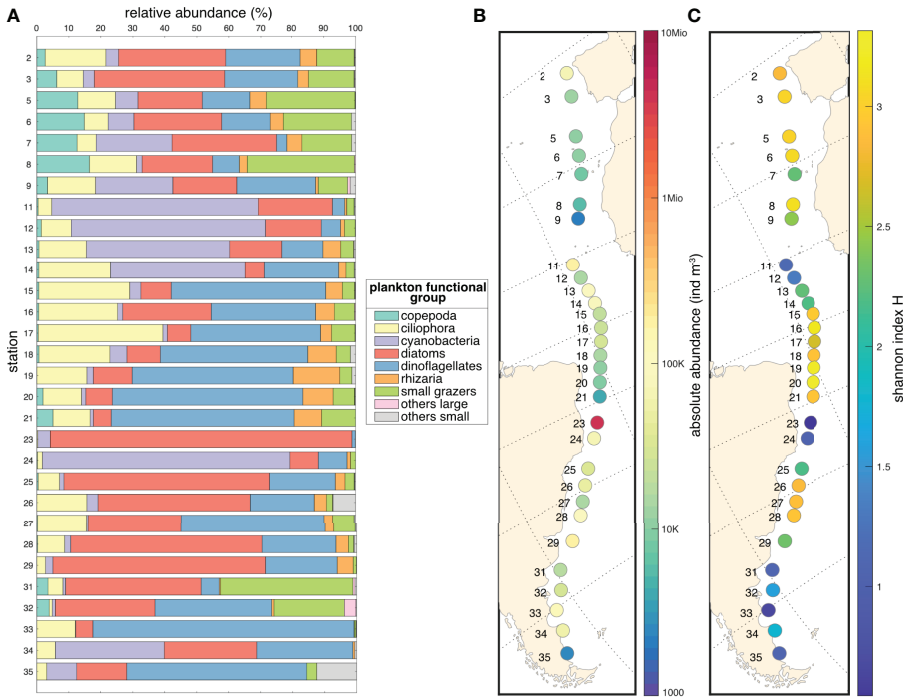


FIGURE 4 | (A) Relative abundance of the nine microplanktonic taxo-functional groups per station. **(B)** Absolute microplankton abundance values in individuals m^{-3} along the *Tara* Mission Microbiomes Atlantic transect. **(C)** Microplankton diversity (Shannon index H) at each station. H-index values are determined from the 179 taxa.

3.3 Morphological Analysis of Surface Microplankton

The PCA analysis performed on 15 morphological variables (**Figure 6A**) defined a typical morphometric space on the first axis (40% of variance explained) with small round organisms on one end, and larger, elongated organisms on the other end (positive values; **Figure 6B**). The second PCA axis (23% of variance explained) corresponds to a color space, with green and red colored organisms for positive values and transparent, lightly blue-colored organisms for negative values. At the

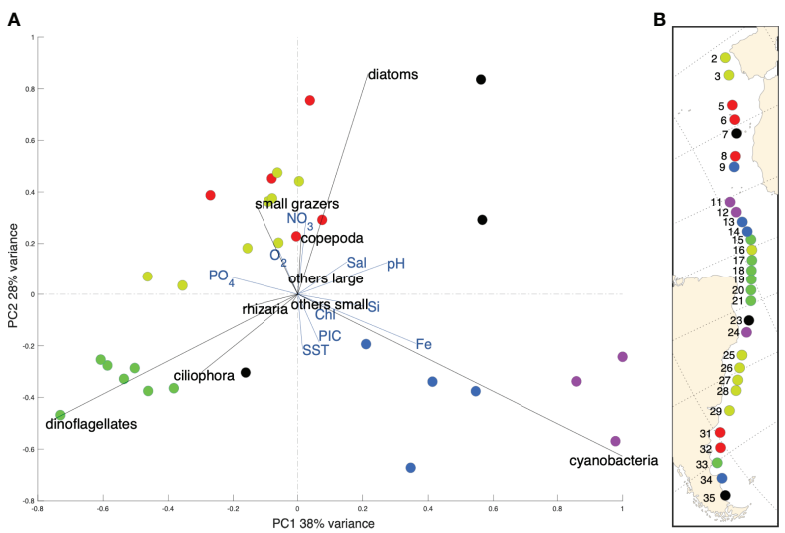


FIGURE 5 | (A) PCA performed on the taxonomic composition at the functional scale (9 taxo-functional groups, Hellinger transformed data) and including key environmental variables. The colors correspond to the 8 characteristic clusters determined via an independent hierarchical analysis based on Euclidean distance. **(B)** The taxo-functional clusters are projected on the Mission Microbiomes Atlantic transect with corresponding station numbers.

station level across the *Tara* Mission Microbiomes Atlantic transect (**Figure 6B**), a trend in the size and shape of organisms (axis 1; **Figure 6B**) is observed. Microplankton communities are dominated by relatively large, elongated organisms at the beginning of the transect (stations 1 to 12), and communities characterized by increasingly small and round organisms south of the equator in the more coastal stations 16 to 35. Stations 11 and 12 displaying very low diversity in the North Atlantic showed clear morphological signals corresponding to communities dominated by *Trichodesmium* cyanobacteria (>60% that are large, elongated and poorly colored cells; see **Figure 4A**).

4 DISCUSSION

4.1 Characterization of our New Citizen Plankton Sampling and Imaging Gears

The concurrent deployment of validated and novel plankton sampling devices allowed quality check of our new frugal tools. Comparison of the two nets (DN and *Coryphaena*) first revealed a significant under-sampling of the DN-50µm, hypothetically explained by a too large mesh size (Heron, 1968) and the resulting abrasion of planktonic organisms flowing onto the dry silk leading to strong degradation. On the other hand, the silk of the DN-20µm stays immersed in water due to the slower filtration process, leading to better plankton preservation and good-quality samples that could be used as standard for further comparison with the *Coryphaena*. Such comparison allowed us to identify significant under-sampling of the second *Coryphaena*, when it was replaced due to the loss of the original net. The new *Coryphaena* net probably had a manufacturing defect such as hole(s) in the collector mesh; future versions will need to integrate solutions to quality-check the material before deployment in the field.

4.1.1 Coryphaena

The *Coryphaena* net was deployed while *Tara* was cruising at speeds between 4 and 11 knots. A reduction of the flow due to filtration resistance through the mesh (Tranter and Smith, 1968) is thus expected, in opposition to the DN where all the water collected is filtered through the system (Gorsky et al., 2019). Consistently, the *Coryphaena* sampled less than the DN-20µm net (correction factor=0.35). Comparatively, replicate water collections using the same type of net display 17% variability on average on plankton biomass, and between 20% and 50% variability between two different nets types (Skjoldal et al., 2013). The variability between the 2 plankton collection gears observed herein (~35%) can therefore be considered as relatively low, and thus validates the sampling efficiency of the *Coryphaena*. Such discrepancies between sampling gears have been shown in many past studies (e.g., Herdman, 1921; Barnes and Marshall, 1951; Anraku, 1956 and Wiebe and Wiebe, 1968), and are typically due to net avoidance, mesh extrusion, escapement, and especially to non-random distribution of plankton (local plankton patchiness; Robinson et al., 2021). Indeed, unlike laboratory experiments

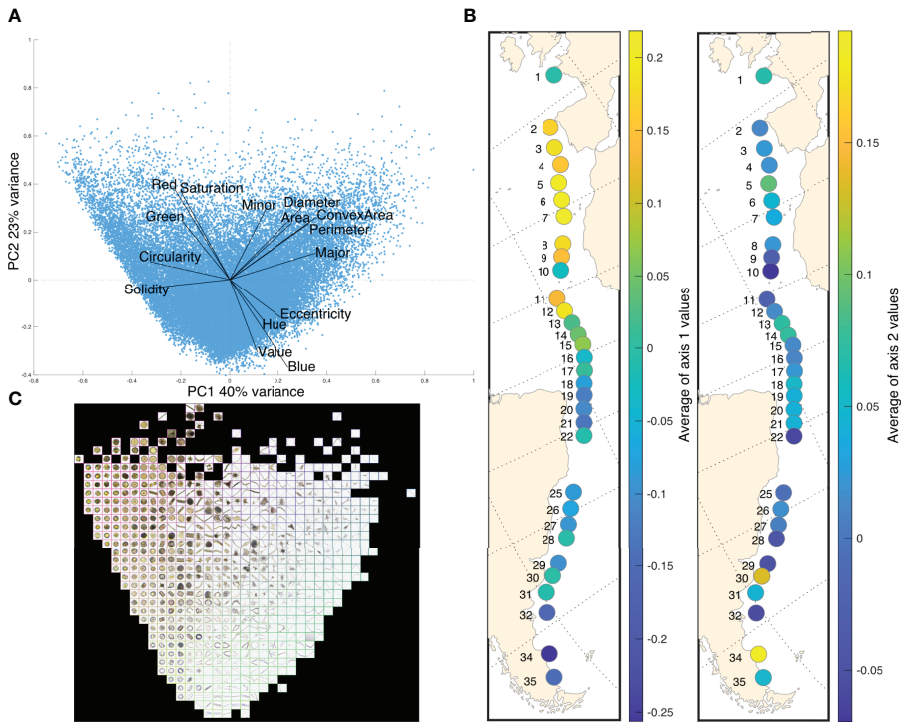


FIGURE 6 | (A) PCA performed on morphological variables characterizing the size, shape and color of microplankton (Euclidean distance). **(B)** Morphological characteristics of microplankton along the *Tara* Mission Microbiomes Atlantic transect, represented by morphometric (left side, PCA axis 1 values, average per station) and color (right side, PCA axis 2 values, average per station) features. **(C)** Representation of the PCA space by characteristic *PlanktoScope* vignettes.

where all variables are isolated and controlled, field trials to validate technologies such as the *Coryphaena* and DN, induce variability dependent on local conditions. Part of the variability observed between instruments could therefore result from plankton heterogeneity in the ocean (Robinson et al., 2021). Notably, although performed in the same area, the *Coryphaena* sampling took on average 25 min (maximum 55 min), while the DN needed between 1-2 hours to filter ca. equivalent volumes of surface sea-waters (0.5 to 4 cubic meters; see **Supplementary Table II**). However, the correction factors between the two nets (0.35) allowed us to adjust their quantitative biases toward a global, surface plankton analysis. The *Coryphaena* adjusted data display minor differences as compared to the DN data, with notable under-sampling of certain taxa. This slight difference is likely due to the relatively high sampling speed that generates increasing pressure across the mesh (Keen, 2013) and damage some organisms. This explains our results showing higher sampling of fragile taxa, such as *Diatoma* and *Eucampia*, by the DN when compared to the *Coryphaena* (**Figure 3C**). This adds up to putative 'mesh selection' effect (Heron, 1968; Vannucci, 1968) related to the elongated shapes of certain fragile plankton (see the FlowCam image of the taxa Diatoma on **Figure 3C**), i.e. these can get stuck in the 50µm-mesh of the *Coryphaena* net and/or be more prone to escape through the mesh and thus not be analyzed by quantitative imaging.

4.1.2 PlanktoScope

The *PlanktoScope* and the FlowCam were previously compared on a single plankton sample collected offshore the Mediterranean marine laboratory of Villefranche/Mer (Pollina et al., this issue) showing a higher abundances of living organisms data collected by the *PlanktoScope* for equivalent volume of water analyzed (correction factor=2.24). Here, we carried out an extensive characterization of the *PlanktoScope* performances over an Atlantic transect on board *Tara*. This comparison reinforces the higher abundances of living organisms data collected by the *PlanktoScope* with respect to the FlowCam (correction factor=1.86). This difference could be explained by the FlowCam operating protocol involving a better homogenization of the sample in the syringe injecting plankton into the system. Indeed, low plankton mixing favors sedimentation at the bottom of the admission syringe of the *PlanktoScope*, putatively driving larger and biased concentrations into the system. Tests confirming such sedimentation bias within the *PlanktoScope* have been performed lately, allowing adjustments of the hardware and protocol to avoid this shortcoming in future *PlanktoScope* deployments.

4.2 Accurate and Underway, Citizen-Tools Based Assessment of Microplankton at Basin-Scale

4.2.1 Microplankton Taxonomic Composition Across the Atlantic Ocean

Overall, our study has allowed consistent description of surface micro-plankton taxonomic composition in direct relation to environmental constraints and biogeography. The correlations we found between taxo-functional groups and environmental features (**Figures 4**, **5**) are consistent with the plankton-environment associations summarized in Margalef's revisited mandala (Glibert, 2016), and highlight the central role of various nutrient limitations in the structure of surface microplankton composition and their abundance as described by Moore et al. (2013). These consistencies thus show the power of our new frugal tools to assess plankton ecology on a global scale.

In our dataset diatoms correlate with high NO_3 concentrations and are found in eutrophic and cold areas (**Figure 5**), which is consistent with the physiological appetence of diatoms to nitrate absorption and storage (Glibert et al., 2016). *Trichodesmium* cyanobacteria negatively correlated with macronutrients (NO_3 and PO_4) and dominated warm oligotrophic zones, which is consistent with their diazotrophy allowing them to fix dissolved N_2. Since ca. 99% of ocean nitrogen is in the form of dissolved N_2 (Gruber and Galloway, 2008), diazotrophic cyanobacteria have a major ecological advantage in oligotrophic areas, however they require 2.5 to 100 times more iron than non-diazotrophic organisms (Zehr, 2011), which explains their positive association with iron in our results (**Figure 5**). The geographic distribution of *Trichodesmium* cyanobacteria in our study (stations 9 to 14; **Figures 4A**, **5**) is otherwise broadly consistent with that observed across 8 Atlantic Meridional Transect (AMT) cruises (Tyrrell, 2003), demonstrating predominance in the region between 0 and ~15°N, with an average filament concentration in the surface layer of 300 ± 101 filaments l^{-1} and a maximum of >600 filaments l^{-1}. The observed correlation between PO_4 and dinoflagellates is also found in Margalef's revisited mandala (Glibert, 2016). However, (bio)chemicals factors such as nutrients limitations are incomplete predictors of plankton community structure (Lima-Mendez et al., 2015). Plankton symbiotic relationships must be considered, especially in the oligotrophic water masses at tropical and subtropical latitudes where mutualistic species interactions are prevalent (Massana, 2015). Of note, our absolute abundance data point to a bloom of the colonial diatom *Hemiaulus hauckii* at station 23 (**Figure 4**), an area where such bloom was previously reported (Carpenter et al., 1999). This diatom bloom occurs in NO_3-poor waters, which is explained by the presence of the endosymbiontic diazotrophic cyanobacterium *Richelia* in *Hemiaulus* cells, providing to the diatom host the nitrogen needed to thrive in these otherwise oligotrophic waters (Villareal, 1992; Carpenter et al., 1999). Images from the *PlanktoScope* allow direct confirmation of this biotic interaction in the sampled populations (**Supplementary Figure 1**). Station 24, characterized by an even stronger nitrate limitation but with higher iron concentrations, was dominated by cyanobacteria (**Figures 4**, **5**). We thus detected a shift from diazotrophic symbiotic diatoms to diazotrophic cyanobacteria, likely due to different levels of nitrogen *versus* iron limitation between two consecutive stations separated by 372 km.

4.2.2 Exploring the Morphometric and Color Spaces of Surface Atlantic Microplankton

The relatively large image dataset collected here (370 175 images) by the *Coryphaena/PlanktoScope* frugal kit allows exploration of the morphological traits of surface-water microplankton across large environmental and geographic scales, independently of the

tedious semi-automated taxonomic annotation of all vignette individually. Our results (**Figure 6**) show only a very weak morphological signal (mean PCA value close to 0; **Figures 6C, D**). This high variability highlights the extreme diversity of plankton morphological characteristics (size, shape, and color) previously described in the literature (e.g. recently, Ibarbalz et al., 2019; Ryabov et al., 2021). Only a few stations with low Shannon diversity (but high dominance of a single taxon, e.g. station 11 an 12; **Figures 5**, **6**) display distinct morphological components that match the morphological traits of the dominant organism. The majority of the variance (first axis 40% of variance; **Figure 6**) is explained by a typical morphological space opposing different shapes and sizes. This morphometric space is echoed in a study by Ryabov et al. (2021) where cell elongation and cell volume together explained up to 92% of the total variance. Indeed, it is known that environmental conditions, such as nutrients, light or temperature, affect the shape and size distributions of plankton (Naselli-Flores et al., 2007; Stanca et al., 2013; Ryabov et al., 2021) confirming that both size and shape are crucial determinants of fitness. Given that our study focused on surface plankton, we would expect a predominance of round shapes while elongated shapes are mostly found in deep waters as they would optimize chloroplast aggregation along the cell surface and increase light harvesting (O'Farrell et al., 2007). However, a predominance of round shapes is not clearly visible in our results, and is highly counteracted by the large presence of *Trichodesmium* filaments. The fact that including color information gathers 26% variance in our dataset (**Figure 6**), further shows that coloration is an important plankton trait (Martini et al., 2021) that previous morphologic studies conducted only on shape and size have deeply ignored because of technological constraints. The onset of a new generation of instruments with color capabilities, like the *PlanktoScope*, will allow us to tackle such unexplored plankton traits.

These morphological methods are very promising for large datasets, and will prove valuable for the work we propose in the context of large-scale citizen science observations. For in-depth analyses of plankton morphological traits, beyond the addition of color information, improvements can still be made, such as analyses on more precise taxonomic groups like in Ryabov et al. (2021) which showed distinct and different diversities within each taxonomic group or a clustering method on PCA coordinates in order to distinguish distinct morphotypes as done by Ibarbalz et al. (2019) on plankton or by Trudnowska et al. (2021) on marine snow.

5 CONCLUSION

This study demonstrates that frugal and affordable tools for biological oceanography can match the quality of validated scientific instruments. The *PlanktoScope*, a simple imaging system, yielded results comparable to that of the Flowcam, a state-of-the-art scientific instrument. The *Coryphaena*, a 3D-printed net allowing collection of micro-plankton at speeds up to 11 knots, recovered plankton communities matching the ones sampled by a validated concentration system. Improvements can certainly be made to these instruments, notably to increase their robustness; however, these represent great perspectives for cooperative plankton studies over unique spatio-temporal scales by citizen sailors. Furthermore, we have also shown how our new frugal tools enabled low-cost collection of consistent plankton data at basin scale allowing taxonomic and morphological assessment and analysis of surface plankton over a 6 months time frame from plankton sampling to statistical analysis of the data. Our results are in agreement with previous observations, showing that the taxonomic and morphological compositions of surface plankton are essentially controlled by different nutrient limitations selecting specific phytoplanktonic functional groups and symbiotic associations. Overall, this shows that long-term collaborative plankton monitoring at planetary scale is not anymore a dream, and such endeavor would provide the 'essential oceanic and climatic variable' (Bax et al., 2019) critically needed to model oceanic ecosystems facing global changes.

DATA AVAIBILITY STATEMENT

The datasets analyzed for this study can be found in the EcoTaxa web platform:
https://ecotaxa.obs-vlfr.fr/prj/3891
https://ecotaxa.obs-vlfr.fr/prj/3892
https://ecotaxa.obs-vlfr.fr/prj/4343
https://ecotaxa.obs-vlfr.fr/prj/4356

AUTHOR CONTRIBUTIONS

ZM: taxonomic annotation, data analysis, wrote the manuscript AO: designed the study, sample acquisition, constructive comments, revised the manuscript. DG: designed the study, sample acquisition, constructive comments. TP: designed the study, data analysis, constructive comments. RB: designed the study, data analysis, constructive comments. CM: designed study, logistical support. RT: designed study, logistical support. MP: designed study, constructive comments. CV: designed the study, constructive comments, revised and edited the manuscript. FL: designed the study, constructive comments, revised and edited the manuscript, supervised the study. All authors contributed to the article and approved the submitted version.

FUNDING

This study used EU Copernicus Marine Service Information and Mercator Ocean products. This project has received funding from the European Union's Horizon 2020 research and innovation programme "Atlantic Ecosystems Assessment, Forecasting and Sustainability" (AtlantECO) under grant agreement No 862923 (ZM, FL, AO and CV), the CNRS MITI OOSP grant (AO), the 'Plankton Arts' grant from the Fondation d'Entreprise Total (TP, MP, CdV), and the COMPLEx team (ZM fellowship).

ACKNOWLEDGMENTS

We would like to thank the 'Plankton Planet' team (https://planktonplanet.org) who has developed the *Coryphaena* and *PlanktoScope* over the last few years, and in particular Niels Haentjens, Guillaume Bourdin, and Emmanuel Boss for the installation of the inline biophysics sensors on board *Tara*. We thank the CNRS, CEA, CMM, Sorbonne Université, Institut Universitaire de France (IUF), as well as the *Tara* Ocean Foundation team and its partners: agnès b., BIC, Capgemini Engineering, Fondation Groupe EDF, Compagnie Nationale du Rhône, L'Oréal, Biotherm, Région Bretagne, Lorient Agglomeration, Billerudkorsnas, Havas Paris, Fondation Rothschild, Office Français de la Biodiversité, AmerisourceBergen, Philgood Foundation, UNESCO-IOC, Etienne Bourgois. This article is contribution number 2 of *Tara* Mission Microbiomes and number 4 of Plankton Planet.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.916025/full#supplementary-material

REFERENCES

Álvarez, E., Moyano, M., López-Urrutia, Á., Nogueira, E. and Scharek, R. (2014). Routine Determination of Plankton Community Composition and Size Structure: A Comparison Between FlowCAM and Light Microscopy. *J. Plank. Res.* 36, 170–184. doi: 10.1093/plankt/fbt069

Anraku, M. (1956). Some Experiments on the Variability of Horizontal Plankton Hauls and on the Horizontal Distribution of Plankton in a Limited Area. *North Sea Research Report of the Department of Aquaculture, Hokkaido University* 7 (1), 1–16.

Barnes, H. T. and Marshall, S. M. (1951). On the Variability of Replicate Plankton Samples and Some Applications of 'Contagious' Series to the Statistical Distribution of Catches Over Restricted Periods. *J. Mar. Biol. Assoc. Unite. Kingdom.* 30 (2), 233–263. doi: 10.1017/S002531540001273X

Batten, S. D., Abu-Alhaija, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., et al. (2019). A Global Plankton Diversity Monitoring Program. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00321

Bax, N. J., Miloslavich, P., Muller-Karger, F. E., Allain, V., Appeltans, W., Batten, S. D., et al. (2019). A Response to Scientific and Societal Needs for Marine Biological Observations. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00395

Beaugrand, G. (2005). Monitoring Pelagic Ecosystems Using Plankton Indicators. *ICES. J. Mar. Sci.* 62, 333–338. doi: 10.1016/j.icesjms.2005.01.002

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple Stressors of Ocean Ecosystems in the 21st Century: Projections With CMIP5 Models. *Biogeosciences* 10, 6225–6245. doi: 10.5194/bg-10-6225-2013

Borkman, D. G. and Smayda, T. J. (2009). Gulf Stream Position and Winter NAO as Drivers of Long-Term Variations in the Bloom Phenology of the Diatom Skeletonema Costatum "Species-Complex" in Narragansett Bay, RI, USA. *J. Plank. Res.* 31, 1407–1425. doi: 10.1093/plankt/fbp072

Brewin, R. J. W., Hyder, K., Andersson, A. J., Billson, O., Bresnahan, P. J., Brewin, T. G., et al. (2017). Expanding Aquatic Observations Through Recreation. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00351

Buskey and Hyatt, 2006: Buskey, E. J., and Hyatt, C. J. (2006). Use of the FlowCAM for semi-automated recognition and enumeration of red tide cells (Karenia brevis) in natural plankton samples. *Harmful Algae* 5, 685–692. doi: 10.1016/j.hal.2006.02.003

Carpenter, E., Montoya, J., Burns, J., Mulholland, M., Subramaniam, A. and Capone, D. (1999). Extensive Bloom of a N2-Fixing Diatom/Cyanobacterial Association in the Tropical Atlantic Ocean. *Mar. Ecol. Prog. Ser.* 185, 273–283. doi: 10.3354/meps185273

de Vargas, C., Pollina, T., Romac, S., Le Bescot, N., Henry, N., Berger, C., et al. (2020). *Plankton Planet* : 'Seatizen' Oceanography to Assess Open Ocean Life at the Planetary Scale. *Ecology. bioRxiv* 08 (31), 263442. doi: 10.1101/2020.08.31.263442

Falkowski, P. G., Fenchel, T. and Delong, E. F. (2008). The Microbial Engines That Drive Earth's Biogeochemical Cycles. *Science* 320, 1034–1039. doi: 10.1126/science.1153213

Glibert, P. M. (2016). Margalef Revisited: A New Phytoplankton Mandala Incorporating Twelve Dimensions, Including Nutritional Physiology. *Harm. Algae.* 55, 25–30. doi: 10.1016/j.hal.2016.01.008

Glibert, P. M., Wilkerson, F. P., Dugdale, R. C., Raven, J. A., Dupont, C. L., Leavitt, P. R., et al. (2016). Pluses and Minuses of Ammonium and Nitrate Uptake and Assimilation by Phytoplankton and Implications for Productivity and Community Composition, With Emphasis on Nitrogen-Enriched Conditions: Pluses and Minuses of NH4+ *and* NO3–. *Limnol. Oceanogr.* 61, 165–197. doi: 10.1002/lno.10203

Glover, 1953: Glover, R. S. (1953). The Hardy plankton indicator and sampler: a description of the various models in use. *Bulletin of Marine. Ecology*, 4, 7–20.

Gorsky, G., Bourdin, G., Lombard, F., Pedrotti, M. L., Audrain, S., Bin, N., et al. (2019). Expanding Tara Oceans Protocols for Underway, Ecosystemic Sampling of the Ocean-Atmosphere Interface During Tara Pacific Expedition, (2016–2018). *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00750

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital Zooplankton Image Analysis Using the ZooScan Integrated System. *J. Plank. Res.* 32, 285–303. doi: 10.1093/plankt/fbp124

Gruber, N. and Galloway, J. N. (2008). An Earth-System Perspective of the Global Nitrogen Cycle. *Nature* 451, 293–296. doi: 10.1038/nature06592

Herdman, W. A. (1921). Variation in Successive Vertical Plankton Hauls at Port Erin. *In. Proc. Trans. Liverpool. Biol. Soc.* 35, 161–174. doi:10.1017/S002531540001273X

Heron, A. C. (1968). "Plankton gauze" in *Zooplankton Sampling*, ed. The United Nations Educational, Scientific and Cultural Organization (UNESCO), ISBN 92-3101194-4, 19-27.

Hrycik, A. R., Shambaugh, A. and Stockwell, J. D. (2019). Comparison of FlowCAM and Microscope Biovolume Measurements for a Diverse Freshwater Phytoplankton Community. *J. Plank. Res.* 41, 849–864. doi: 10.1093/plankt/fbz056

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., et al. (2019). Global Trends in Marine Plankton Diversity Across Kingdoms of Life. *Cell* 179, 1084–1097.e21. doi: 10.1016/j.cell.2019.10.008

Ide, K., Takahashi, K., Kuwata, A., Nakamachi, M. and Saito, H. (2007). A Rapid Analysis of Copepod Feeding Using FlowCAM. *J. Plank. Res.* 30, 275–281. doi: 10.1093/plankt/fbm108

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L. and Stemmann, L. (2022). Machine Learning for the Study of Plankton and Marine Snow From Images. *Annu. Rev. Mar. Sci.* 14, 277–301. doi: 10.1146/annurev-marine-041921-013023

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press: In Press.

Keen, 2013: Keen, E. (2013). A practical designer's guide to mesozooplankton nets. Environmental Science.

Lauro, F. M., Senstius, S. J., Cullen, J., Neches, R., Jensen, R. M., Brown, M. V., et al. (2014). The Common Oceanographer: Crowdsourcing the Collection of Oceanographic Data. *PloS Biol.* 12, e1001947. doi: 10.1371/journal.pbio.1001947

Le Bourg, B., Cornet-Barthaux, V., Pagano, M. and Blanchot, J. (2015). FlowCAM as a Tool for Studying Small (80–1000 μm) Metazooplankton Communities. *J. Plank. Res.* 37, 666–670. doi: 10.1093/plankt/fbv025

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., et al. (2015). Determinants of Community Structure in the Global Plankton Interactome. *Science* 348, 1262073. doi: 10.1126/science.1262073

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Martini, S., Larras, F., Boyé, A., Faure, E., Aberle, N., Archambault, P., et al. (2021). Functional Trait-Based Approaches as a Common Framework for Aquatic Ecologists. *Limnol. Oceanogr.* 66, 965–994. doi: 10.1002/lno.11655

Massana, R. (2015). "Protistan Diversity in Environmental Molecular Surveys," in *Marine Protists*. Eds. Ohtsuka, S., Suzaki, T., Horiguchi, T., Suzuki, N. and Not, F. (Tokyo: Springer Japan), 3–21. doi: 10.1007/978-4-431-55130-0_1

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., Boyd, P. W., et al. (2013). Processes and Patterns of Oceanic Nutrient Limitation. *Nat. Geosci.* 6, 701–710. doi: 10.1038/ngeo1765

Naselli-Flores, L., Padisák, J. and Albay, M. (2007). Shape and Size in Phytoplankton Ecology: Do They Matter? *Hydrobiologia* 578, 157–161. doi: 10.1007/s10750-006-2815-z

O'Farrell, I., de Tezanos Pinto, P. and Izaguirre, I. (2007). Phytoplankton Morphological Response to the Underwater Light Conditions in a Vegetated Wetland. *Hydrobiologia* 578, 65–77. doi: 10.1007/s10750-006-0434-3

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015). Open Science Resources for the Discovery and Analysis of Tara Oceans Data. *Sci. Data* 2, 150023. doi: 10.1038/sdata.2015.23

Platt, T. (1978). "Spectral Analysis of Spatial Structure in Phytoplankton Populations," in *Spatial Pattern in Plankton Communities*. Ed. Steele, J. H. (Boston: MA: Springer US), 73–84. doi: 10.1007/978-1-4899-2195-6_4

Pollina, T., Larson, A. G., Lombard, F., Li, H., Colin, S., de Vargas, C., et al. (2020). PlanktonScope: Affordable Modular Imaging Platform for Citizen Oceanography. *Bioengineering*. 4 (23), 056978. doi: 10.1101/2020.04.23.056978

Robinson, K. L., Sponaugle, S., Luo, J. Y., Gleiber, M. R. and Cowen, R. K. (2021). Big or Small, Patchy All: Resolution of Marine Plankton Patch Structure at Micro- to Submesoscales for 36 Taxa. *Sci. Adv.* 7, eabk2904. doi: 10.1126/sciadv.abk2904

Ryabov, A., Kerimoglu, O., Litchman, E., Olenina, I., Roselli, L., Basset, A., et al. (2021). Shape Matters: The Relationship Between Cell Geometry and Diversity in Phytoplankton. *Ecol. Lett.* 24, 847–861. doi: 10.1111/ele.13680

Sieracki, C., Sieracki, M. and Yentsch, C. (1998). An Imaging-in-Flow System for Automated Analysis of Marine Microplankton. *Mar. Ecol. Prog. Ser.* 168, 285–296. doi: 10.3354/meps168285

Skjoldal, H. R., Wiebe, P. H., Postel, L., Knutsen, T., Kaartvedt, S. and Sameoto, D. D. (2013). Intercomparison of Zooplankton (Net) Sampling Systems: Results From the ICES/GLOBEC Sea-Going Workshop. *Prog. Oceanog.* 108, 1–42. doi: 10.1016/j.pocean.2012.10.006

Stanca, E., Cellamare, M. and Basset, A. (2013). Geometric Shape as a Trait to Study Phytoplankton Distributions in Aquatic Ecosystems. *Hydrobiologia* 701, 99–116. doi: 10.1007/s10750-012-1262-2

Taylor, A. H., Allen, J. I. and Clark, P. A. (2002). Extraction of a Weak Climatic Signal by an Ecosystem. *Nature* 416, 629–632. doi: 10.1038/416629a

Tranter and Smith, 1968: Tranter, D.J., and Smith, P.E. (1968). "*Filtration performance" in Zooplankton Sampling*, ed. The United Nations Educational, Scientific and Cultural Organization (UNESCO), ISBN 92-3101194-4, 27-57

Trudnowska, E., Lacour, L., Ardyna, M., Rogge, A., Irisson, J. O., Waite, A. M., et al. (2021). Marine Snow Morphology Illuminates the Evolution of Phytoplankton Blooms and Determines Their Subsequent Vertical Export. *Nat. Commun.* 12, 2816. doi: 10.1038/s41467-021-22994-4

Tyrrell, T. (2003). Large-Scale Latitudinal Distribution of Trichodesmium Spp. In the Atlantic Ocean. *J. Plank. Res.* 25, 405–416. doi: 10.1093/plankt/25.4.405

Vandromme, P., Stemmann, L., Garcìa-Comas, C., Berline, L., Sun, X. and Gorsky, G. (2012). Assessing Biases in Computing Size Spectra of Automatically Classified Zooplankton From Imaging Systems: A Case Study With the ZooScan Integrated System. *Methods Oceanog.* 1–2, 3–21. doi: 10.1016/j.mio.2012.06.001

Vannucci, M. (1968). "Loss of organisms through the meshes" in *Zooplankton Sampling, ed. The United Nations Educational, Scientific and Cultural Organization (UNESCO), ISBN* 92-3101194-4, 19-27

Vilgrain, L., Maps, F., Picheral, M., Babin, M., Aubry, C., Irisson, J., et al. (2021). Trait-Based Approach Using *in Situ* Copepod Images Reveals Contrasting Ecological Patterns Across an Arctic Ice Melt Zone. *Limnol. Oceanogr.* 66, 1155–1167. doi: 10.1002/lno.11672

Villareal, T. A. (1992). "Marine Nitrogen-Fixing Diatom-Cyanobacteria Symbioses," in *Marine Pelagic Cyanobacteria: Trichodesmium and Other Diazotrophs*. Eds. Carpenter, E. J., Capone, D. G. and Rueter, J. G. (Dordrecht: Springer Netherlands), 163–175. doi: 10.1007/978-94-015-7977-3_10

Wiebe, P. H. and Benfield, M. C. (2003). From the Hensen Net Toward Four-Dimensional Biological Oceanography. *Prog. Oceanog.* 56, 7–136. doi: 10.1016/S0079-6611(02)00140-4. doi: 10.4319/lo.1968.13.2.0315

Wiebe, P. H. and Wiebe, P. H. (1968). Plankton Patchiness: Effects on Repeated Net Tows 1. *Limnol. Oceanog.* 13 (2), 315–321. doi: 10.4319/lo.1968.13.2.0315

Zehr, J. P. (2011). Nitrogen Fixation by Marine Cyanobacteria. *Trends Microbiol.* 19, 162–173. doi: 10.1016/j.tim.2010.12.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

METHODS
published: 22 July 2022
doi: 10.3389/fmars.2022.949428

PlanktoScope: Affordable Modular Quantitative Imaging Platform for Citizen Oceanography

Thibaut Pollina[1,2†], *Adam G. Larson*[1,2†], *Fabien Lombard*[2,3,4,5], *Hongquan Li*[1], *David Le Guen*[2], *Sébastien Colin*[2,6], *Colomban de Vargas*[2,3,7*] *and Manu Prakash*[1,2*]

[1]Bioengineering, Stanford University, Stanford, CA, United States, [2]Plankton Planet Non-Governmental Organization (NGO), Station Biologique de Roscoff, Roscoff, France, [3]Research Federation for the Study of Global Ocean Systems Ecology and Evolution, FR2022/Tara Global Ocean Systems Ecology & Evolution (GOSEE), Paris, France, [4]Sorbonne Université, Centre National de la Recherche Scientifique (CNRS), Laboratoire d'Océanographie de Villefranche, Villefranche-sur-mer, France, [5]Biology Institut Universitaire de France (IUF), Paris, France, [6]Max Planck Institute for Biology Tübingen, BioOptics Facility, Max-Planck-Ring 5, Tübingen, Germany, [7]CNRS, Sorbonne Université, Station Biologique de Roscoff, UMR7144, ECOMAP - Ecology of Marine Plankton, Roscoff, France

OPEN ACCESS

Edited by:
Martin Edwards,
Plymouth Marine Laboratory,
United Kingdom

Reviewed by:
David McKee,
University of Strathclyde,
United Kingdom
Rohit Ghai,
Academy of Sciences of the Czech Republic (ASCR), Czechia

***Correspondence:**
Manu Prakash
manup@stanford.edu
Colomban de Vargas
vargas@sb-roscoff.fr

†These authors share first authorship

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 20 May 2022
Accepted: 23 June 2022
Published: 22 July 2022

Citation:
Pollina T, Larson AG, Lombard F, Li H, Le Guen D, Colin S, de Vargas C and Prakash M (2022) PlanktoScope: Affordable Modular Quantitative Imaging Platform for Citizen Oceanography. Front. Mar. Sci. 9:949428. doi: 10.3389/fmars.2022.949428

The oceans represent 97% of all water on Earth and contain microscopic, drifting life, plankton, which drives global biogeochemical cycles. A major hurdle in assessing marine plankton is the planetary scale of the oceans and the logistical and economic constraints associated with their sampling. This difficulty is reflected in the limited amount of scientifically equipped fleets and affordable equipment. Here we present a modular hardware/software open-source strategy for building a versatile, re-configurable imaging platform - the PlanktoScope - that can be adapted to a number of applications in aquatic biology and ecology. We demonstrate high-throughput quantitative imaging of laboratory and field plankton samples while enabling rapid device reconfiguration to match the evolving needs of the sampler. The presented versions of PlanktoScope are capable of autonomously imaging 1.7 ml per minute with a 2.8 µm/px resolution and can be controlled from any WiFi-enabled device. The PlanktoScope's small size, ease of use, and low cost - under $1000 in parts - enable its deployment for customizable monitoring of laboratory cultures or natural micro-plankton communities. This also paves the way toward consistent and long-term measurement of plankton diversity by an international fleet of citizen vessels at the planetary scale.

Keywords: PlanktoScope, microplankton, frugal microscopy, quantitative imaging, open source modularity

1 INTRODUCTION

Life drifting in water - plankton - forms the foundation of ecological networks and biodiversity in aquatic ecosystems (Fenchel, 1988). It is a major driver of global geochemical processes, by generating nearly half of the planet's oxygen (Field, 1998) and maintaining a flux of photosynthetically fixed carbon to deeper layers of the ocean and its floor (Field, 1998, Henson et al., 2012). However, we still know little about the ecological and evolutionary dynamics of planktonic communities or the extent of the anthropogenic impact on these communities. Unlocked by the revolution in environmental DNA sequencing, our knowledge about plankton diversity has dramatically improved over the last two decades, notably through global-scale expeditions led by biologists, including the

Global Ocean Sampling (Venter et al., 2004), *Tara* Oceans (Karsenti et al., 2011; Duarte, 2015). In particular, *Tara* Oceans (2009 - 2013) has applied a standardized, eco-systems biology strategy to explore plankton diversity from genes to communities, from viruses to animals, and across coarse but planetary spatial and seasonal scales (Sunagawa et al., 2020). The combination of global ocean DNA metabarcoding, metagenomic, and metatranscriptomic datasets [e.g. (de Vargas et al., 2015) (Sunagawa et al., 2015) Carradec et al. (2018)] has unveiled the basic structures of open ocean plankton taxonomic diversity and generated hypotheses about its interactions (Chaffron et al., 2021), biogeography (Ruuskanen et al., 2021), and roles in critical ocean processes such as the carbon pump Guidi et al., 2016.

However, understanding the eco-evolutionary dynamics of plankton will require far more information across the four dimensions of the world ocean. In addition, if the molecular 'omics' data bring a wealth of taxonomic and metabolic knowledge, they convey relatively poor information about the phenotypes, abundances, interactions, and behaviors at the organismal level, which are driving a large extent of plankton ecology and function (Martini et al., 2021). Today, it is critical to complement the ocean 'omics' layer of information with quantitative imaging data as it is classically performed in cell biology, and this should be done across relevant Spatio-temporal scales of the ocean system, from micro- to meso-, to planetary scales. Quantitative imaging methods allow monitoring of both the quantity and morphological diversity of plankton communities between a few µm to and a few mm in size Lombard et al., 2019, together with measures of the many environmental or anthropogenic factors Kautsky et al., 2016 shaping them. The few existing high-throughput, automated imaging instruments, such as the FlowCam (Sieracki et al., 1998) or the IFCB (Sosik and Olson, 2007), are expensive, bulky, and not suitable for large-scale community deployment. In the 'Plankton Planet' initiative (de Vargas et al., 2022), we propose to harness the creativity of researchers, mariners, and makers, to co-develop a suite of user-friendly and cost-effective tools for a cooperative, global, and long-term measure of microbial aquatic life. Frugal yet scientifically sound tools shared with a large community become an effective way to tackle the problem of the cost associated with classical oceanographic instruments and vessels. For example, the Foldscope (Cybulski et al., 2014), with over two million copies distributed in 164 countries around the world, has enabled a community of citizen microscopists to share their data and discoveries at a planetary scale (http://microcosmos.foldscope.com/). Plankton ecology would greatly benefit from a low-cost portable quantitative microscope that can be used directly at sea or on the shore by the vast community of mariners enjoying and/or living from the ocean.

Here, we used modularity - a natural way to make complexity manageable and accommodate uncertainty in the evolution of design (Efatmaneshnik and Ryan, 2016) - to construct the PlanktoScope, a miniaturized modular open-source imaging platform for quantitative imaging of micro-plankton that matches the quality of much larger and more expensive commercial instruments, for costs that are affordable for personal assembly and use. Even though we develop the canonical versions of the PlanktoScope for a global homogenous measure of plankton life, every module encapsulates a simple function allowing scientists and makers to adopt the platform for their needs. This strategy enables the device to be easily upgraded instead of replaced as a whole, providing a way to take on unforeseen future applications. We demonstrate the efficiency of the PlanktoScope in obtaining high-throughput imaging from both laboratory and field samples while enabling rapid reconfiguration to match the evolving needs of aquatic ecology. Since sharing PlanktoScope with community researchers, we have recorded more than 30+ replications of the instrument worldwide - demonstrating the replicability and scale-up of our approach driven by an organic community built on the collaboration of professional and amateur scientists.

2 MATERIALS AND EQUIPMENT

2.1 Designs of Two PlanktoScope Prototypes

To design the modular version of the PlanktoScope (v.1) made of six units that can be stacked on top of each other (**Figures 1A–D**), we used Autodesk Fusion 360 (v2.0.5688) to create a parametric design optimizing the physical interface common to all modules. Different parameters define the interface's areas, such as the electronic connection area, the magnetic linkage, and the optical path. The thickness of the material and the outer dimensions of all the electronics used inside the instrument were critical to characterizing the interface. The shareable online 3D environment provided by Fusion 360 contains the main 3D model, together with other models that form the electronic and optical parts. Most of these models have been generated by measuring existing objects but some have been downloaded from the online GrabCad library (https://grabcad.com/). Once the different iterations of the 3D model were ready to be machined, the sketches were extracted as DXF files from Fusion 360 and nested in Adobe Illustrator CC (version 22.1) to fit the dimensions of the sheets of used material. The parts were then machined on a 3 mm thick acrylic sheet by a laser cutter machine (RS-1610L) with an optimal resolution of 25 µm, at the UBO Open Factory in Brest, France. These laser cutter instruments are common at universities as well as a growing worldwide network of maker/fabrication spaces. Such spaces often provide user access to machines after proper training, though work can often be commissioned for a few hundred dollars. All that is required is sharing of the file found on the PlanktoScope website. The assembly of v.1 was performed manually and took c.a. 8 hours. On the other hand, the monolithic version of the PlanktoScope (v.2) (**Figure 1E**) has been designed for fluidic-based, quantitative observations, and thus employs a much simpler assembly process. Its form factor and robustness allow it to be carried in a backpack for field trips without risking damage. Modularity remains in the objective lens that can be swapped magnetically as well as the Ibidi Luer Slide holder, while other components such as focus stages and electronics remain fixed. The PlanktoScope v.2 can be assembled in less than 4 hours.

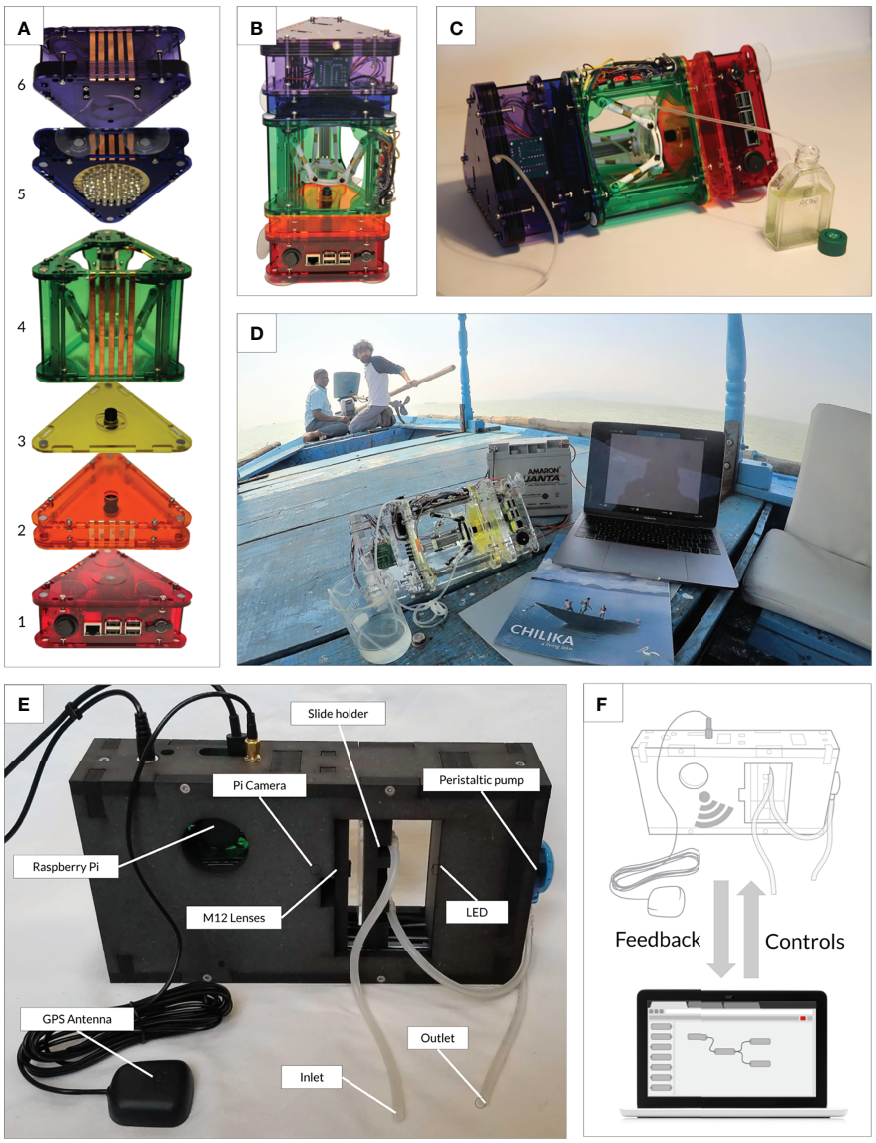


FIGURE 1 | Comparison of the modular (v.1) and monolithic (v.2) PlanktoScope designs. **(A)** Modular stackable flow-through microscope design (bottom to top): Computational/imaging sensor (1), tube lens (2), objective lens (3), delta stage for sample manipulation and focus including flow cell mount (4), illumination (5), pump (6). The platform can be re-assembled and is held together by the alignment of fixed magnets. **(B)** The PlanktoScope v.1 can be used in vertical configuration for static imaging or **(C)** horizontal configuration for flow through imaging. **(D)** Deployment of the PlanktoScope v.1 on board a traditional fishing boat in lake Chilika (Orissa, India), operating autonomously on a 12V car battery. **(E)** Monolithic portable PlanktoScope v.2 with fixed flow-through configuration. **(F)** PlanktoScope is controlled *via* smartphone or laptop allowing real-time feedback during data collection and processing.

2.2 Content of the Modules

The bill of materials (BOM) to assemble a single PlanktoScope v.1 is about $200. The BOM for PlanktoScope v.2 is about $500 (**Supplementary Material Table 1**).

2.2.1 Flow-through Strategies

PlanktoScope v.1 is equipped with a peristaltic pump module (**Figure 1A6**) composed of a stack of 5 acrylic layers forming a closed chamber inside which 3 "rollers" can spin around the motor axis compressing a tube along the internal wall. The speed of the motor and the diameter of the compressed tube determine the flow rate which is about 3 ml/min at maximum speed. The compact PlanktoScope v.2 uses off-the-shelf peristaltic pumps for flow. Many are available in a 10mm x10mm form factor. Common 12V versions provide reliable flow rates of several ml/min. Several models can be easily incorporated by small modifications to the laser cut mount on the 3D model and connected to the other port of the Adafruit Stepper Motor HAT controlling the stage. In both designs, a continuous flow mode and a stop-flow mode can be used. In continuous mode, the peristaltic pump is continuously rotating at a low flow rate while the camera is taking images at a given frame rate. Since Pi Cameras are based on a rolling shutter,

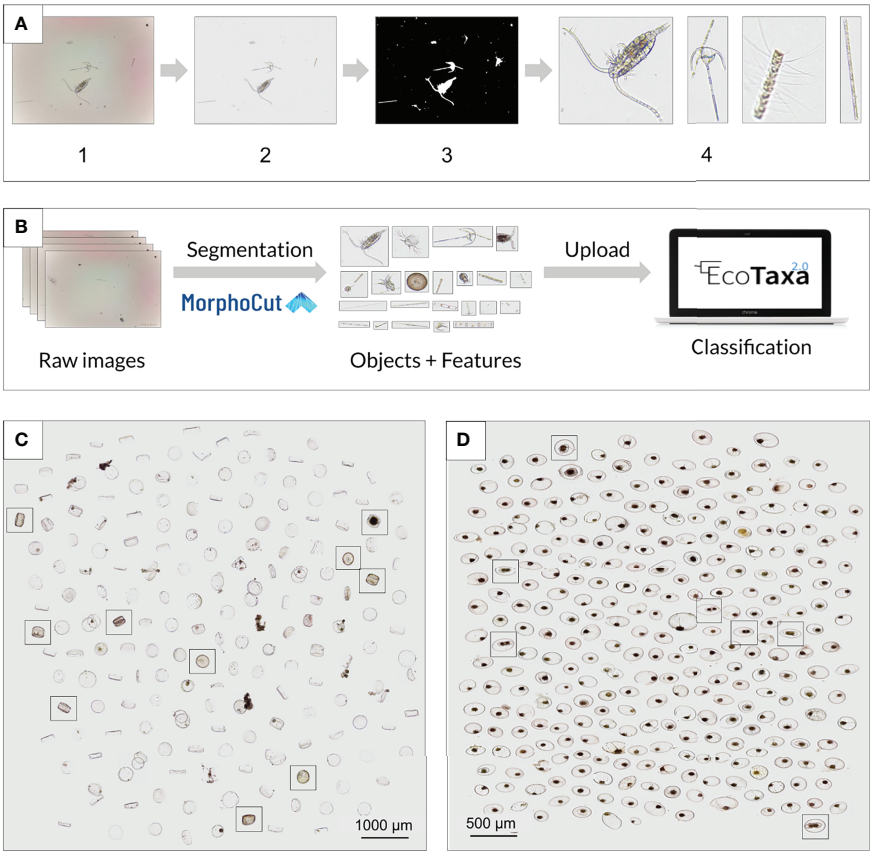


FIGURE 2 | Image processing pipeline for fluidic analysis. **(A)** Workflow used to segment the objects imaged in a single frame and extract features. From the raw images (1) acquired in fluidic mode, MorphoCut applies a running median to approximate the background image (2) based on 5 frames; using OpenCV, a Canny Edge Detection is performed, followed by dilation, closing, and erosion functions (3); from the binary image, MorphoCut extracts the vignette/ROI for each object (4), together with a suite of mathematical image descriptors. **(B)** The Raw images and segments from MorphoCut along with the objects and a table containing all the measured features/metadata can then be directly uploaded on EcoTaxa for classification. **(C)** and **(D)** Non-destructive continuous monitoring of lab cultures using a PlanktoScope allows for cell morphology to be observed at single-cell resolution. **(C)** *Coscinodiscus wailesii* cultures were monitored over a period of 6 hours. Simple montages allow the user to easily quantify living or dead cells at different time points. **(D)** *Pyrocystis noctiluca* cultures were monitored over a period of 6 hours during their night-to-day transition. Dividing cells are easily identifiable.

the imaged objects undergo a morphological deformation when imaged under continuous flow. In addition, peristaltic pumps have a pulsed flow which is difficult to characterize, making post-acquisition correction difficult. Planktoscope uses stop-flow, where rotation of the pump is stopped when each image is taken. The objects are thus stationary when imaged, canceling any morphological deformation due to flow or motion blur, thus allowing quantitative analysis. This enables a longer exposure, increasing the resolution and reducing the need for powerful illumination. This lower frame-rate strategy enables the capture of the full camera sensor for a larger field of view than *via* the continuous mode while maintaining high throughput. However, as the cost of high quality cameras continues to fall, we envision modifications with global shutter sensors or strobed illumination to further improve image acquisition.

2.2.2 Stage and Focus

The PlanktoScope v.1 includes a module combining the focusing and exploring functions (**Figure 1A4**). This linear delta design, used in some 3D printers, uses 3 vertical independent linear stepper motors that hold a platform, each with 2 arms. Each stepper is driven by an A4988 driver powered with 9V and controlled by a common Arduino mini pro present in the module. To control the location of the sample maintained by the platform, an inverse kinematic is necessary to transform an X/Y/Z desired displacement in a delta motion. Here, the code embedded in the Arduino was simplified to control the focus by moving the three stepper motors simultaneously. This Arduino has a defined I2C address allowing the Raspberry Pi to iteratively set a new focal position. The platform made of two separable magnetic bodies can host a broad range of sample holders: a slide, a petri dish, an optical chamber, or a flow cell. Focusing is made possible by controlling 3 independent drivers wired to simultaneously move the stepper motors up or down. The travel distance of the platform measures about 3.2cm with a step size of 0.15µm. This allows fine control of movement to accurately track and image micron-sized objects. For the price of about $30, this represents an affordable way to construct a motorized XYZ stage. In PlanktoScope v.2,

the stage is actuated by two parallel synchronized Allegro linear stepper motors on only the Z-axis for changing focus. Both stepper motors are connected to the same port on the Adafruit Stepper Motor HAT. The flow-cell can be actuated for fine focus using 2 synchronous linear stepper motors offering a step size of 0.15µm on a travel distance measuring about 2.5cm on a single axis.

2.2.3 Illumination

In the PlanktoScope v.1, the illumination module is built of 5 concentric rings composed of 1, 6, 12, 24, and 32 white ultra-bright LEDs having a narrow-angle of 17°. The light intensity of each ring can be tuned separately to offer a broad range of illumination modes. Two main modes are (*i*) pure dark-field where the two external rings are used (**Supplementary Material Figure 5A**) and (*ii*) pure bright-field where the most central LEDs are used (**Supplementary Material Figure 5B**). In the following results, we opted to use the maximum light intensity of the central LED to maximize the depth of field in the flow cell. The compact PlanktoScope uses a single ultra-bright LED at a constant intensity with a narrow angle of 15° enabling bright-field illumination and providing a nearly collimated light source. This achieves a large depth of field for imaging plankton communities with a large size variance. This single white LED (5 mm LTW2S - 17000 mcd) is connected to the stepper board and can be toggled in the user interface.

2.2.4 Optical Modules

The optical train is defined by two inverted S-mount lenses (M12 lenses) that are both encapsulated in different detachable modules. The two modules have been designed to enable a rapid change of each M12 lens used as a couple. The alignment is set by the insertion holes cut and positioned by the laser cutter machine. The distances of the M12 lenses to each other and the sensor are defined by rotating the M12 lenses in the holes tapped using an M12x0.5 hand thread tap from Thorlabs. This optical train remains the same on both versions of the PlanktoScope.

2.2.5 Power, Computational, and Sensor Modules

The PlanktoScope v.1 is directly powered through one multi-functional module dedicated to the computation and sensor (**Figure 1A1**). It receives 12V either by a regular AC power adapter for lab experiments or a battery for field deployment. A custom BUS made of 6 electronic wires dedicated to power the other modules provide 12V, 5V, and Ground wires. The three other wires consist of the I2C, SDA, SCL, and a dedicated Ground enabling the exchange of data between the different modules. The camera sensor is a Pi Camera v2.1 embedded in the module. It is positioned facing up to collect the image coming from above. Under this module the user on one side of the PlanktoScope are 3 suction cups allowing the user to fix the instrument on flat surfaces and improve its vertical and horizontal stability for field experiments (e.g., inside a boat). The PlanktoScope v.2 utilizes the USB-C connector of the Raspberry Pi 4 to power itself, and the Pi HAT (Yahboom Cooling Fan HAT) is mounted on top of it to cool the Raspberry Pi and provide operational feedback to the user *via* 3 RGB LEDs. A ribbon cable connects the Raspberry Pi/Fan HAT to two other HATs, an Adafruit Stepper Motor HAT and the Adafruit Ultimate GPS HAT. The Stepper Motor HAT is powered *via* a DC Power Jack Socket to 12V 1A power. The GPS HAT uses an antenna allowing for a better GPS signal when in the field. Note that in this design, the entire GPIO of the Raspberry becomes the BUS and connects the Raspberry Pi to other physical modules that can be changed, replaced, or upgraded.

2.3 User/machine Interface and Software Architecture

By utilizing the headless configuration for the Raspberry Pi, we removed the need for a dedicated monitor, mouse, and keyboard, enabling control of the instrument from any device able to access a web browser over a WiFi connection (**Figure 1F**). This strategy enables any user to immediately interact with the device without OS or software compatibility issues. The user can then access a browser-based dashboard powered by Node-RED for remote control of the system; acquisition settings, interactive collection of the metadata, as well as rapid state modification of the actuators.

The software architecture (**Supplementary Material Figure 3**) is based on existing programs and python libraries, such as Node-RED (https://nodered.org/) for the Graphical User Interface and the first layer of the programming interface, MorphoCut (https://github.com/morphocut/morphocut) for handling the image processing from the raw images to the online platform, and EcoTaxa (https://ecotaxa.obs-vlfr.fr/) for plankton images classification and annotation.

The back-end of the GUI is also based on Node-RED, a flow-based development tool for visual programming which is provided by default on any Raspberry Pi software suite. Node-RED provides a web browser-based flow editor, which can be used to create JavaScript-based applications. Elements of applications can be saved or shared for re-use. The strategy makes it more accessible to those with limited experience in scripting. This visually modifiable program can easily be shared through a JavaScript Object Notation (.json) text file.

3 METHODS

3.1 Image Workflow and Image Processing

3.1.1 Image Workflow Performed with the PlanktoScope V.1

For the first batch of acquisitions (**Figures 4.1–4.6**, **5**, and **Supplementary Material Figures 5A, B**, **6**), the optical configuration was a 16 mm focal length for the tube lens and a 12 mm focal length for the objective lens. The sensor mode was set to 1080p and the field of view (FOV) was then measured at 2,880 µm wide and 1,620 µm high. The flow cell used was a rectangle-shaped borosilicate glass capillary (VitroTubes), 5000 µm wide, 500 µm deep internally, and 5 cm long. The volume imaged in one frame is about 2.3 µL. Since the capillary width is larger than the FOV width, the whole volume passed in the

capillary is about 4 μL per imaged frame (= FOV height * Cell width * Cell depth). The acquisition was done using a frame-rate set at 8 frames/second, which corresponds to a volume of 1.12 ml imaged per minute. We took 2000 frames per sample, at 5 minutes total, the volume imaged was 5.6 mL per sample. The image processing workflow for this batch was a custom pipeline. Using Numpy, we realized for each frame an average image from 20 frames around the considered frame (10 frames before and 10 frames after) and we subtracted this average image to the current frame using OpenCV. The cleaned frames are then processed with basic Dilation/Closing/Erosion operations in OpenCV. The binary image obtained served to detect the objects in each frame and extract the region of interest along with simple measured features provided by OpenCV such as equivalent diameter, Euler number, extent, area, filled area, major axis length, minor axis length, orientation, perimeter, and solidity. From all the segmented objects, we manually selected the objects most likely to correspond to living organisms to avoid terrigenous sediment abundant in the explored coastal sites. The current segmentation pipeline performed on the instrument is broad pertaining to objects of interest, though parameters for segmentation can be modified in the code depending on the needs of the user. Raw images can also be easily transferred and processed with any custom pipeline off the machine.

3.1.2 Image Workflow Performed with the PlanktoScope V.2

For the second batch of acquisitions (**Figures 2**, **3**, **4.7** and **Supplementary Material Figures 5C–E**), the optical configuration was made using 25mm for the tube lens and 16 mm for the objective lens. The sensor mode was set to full sensor (3280 × 2464 pixels) and the field of view measured 2 300 μm wide and 1 730 μm high. In the v.2 version, the sensor is rotated 90° in comparison to the version v.1. For the camera sensor reference, the direction of the flow is from right to left rather than top to bottom. For this optical configuration and a flow cell with a channel height of 200 μm, the volume imaged in one frame is about 0.8 μL (= FOV width * FOV height * FlowCell depth). The acquisition for both versions was done using a stop flow method which consists of stopping the pump flow and then the flow when acquiring an image. The frame rate is about 1 frame/second, which corresponds to a volume of ~48 μL imaged per minute. The sample was passed through a filter (Überstrainer, PluriSelect inc.) to remove large objects that can clog the capillaries. The image processing workflow for this batch was done using MorphoCut and Ecotaxa as described in *3.1.3* below. For the acquisition shown in **Figure 3**, the extracted vignettes uploaded into *Ecotaxa* can be consulted with all their associated metadata @: https://ecotaxa.obs-vlfr.fr/prj/2748.

3.1.3 Image Processing

The raw images are stored on the Pi after collection and can be automatically processed on the Planktoscope by MorphoCut, a python-based library designed to handle large volumes of imaging data (https://github.com/morphocut/morphocut). Several operations are applied to the raw images (**Figure 2A1**) acquired in fluidic mode. MorphoCut first applies a running median to approximate the background image (**Figure 2A2**) based on 5 frames, a Canny Edge Detection *via* OpenCV is performed, followed by dilation, closing, and erosion functions (**Figure 2A3**) also from OpenCV. From the binary image, MorphoCut extracts the region of interest (ROI) for each present object (**Figure 2A4**). MorphoCut then extracts, using Scikit-image (van der Walt et al., 2014) 32 keys mathematical image descriptors. Each ROI is then stored, along with contextual metadata defined by the user on the Graphical User Interface (GUI). This way, large data sets can be compressed at sea by storing only relevant ROIs and data tables. Finally, all data outputs are zipped in a compressed file and formatted for being imported to the EcoTaxa server (**Figure 2B**). Ecotaxa is a web-interfaced database, which combines supervised machine learning with collaborative visual inspections/classification by taxonomy experts to classify and assign taxonomy to plankton from environmental plankton image datasets. This creates a uniform data format already utilized by plankton researchers worldwide.

3.2 Optical Characterization

Since the optical configuration is made of two reversed M12 lenses, serving respectively as objective lens and tube lens, we choose five different M12 lenses (**Table 1**) based on their compatibility with the chosen camera sensor (Pi Camera v2.1, Sony IMX219, 8MP, sensor area 3.68x2.76mm imaging area, pixel size 1.12x1.12um). As changing the focal length of each lens changes the effective magnification of the image projected on the sensor, we wished to see how each combination enables exploration of objects spanning different size ranges. Pairing and characterization of lens pairs with different effective focal lengths (f) were performed to establish the actual resolution experimentally. We tested the optical performance of each 25 possible configurations by imaging the USAF 1951 resolution test chart. The illumination was set to use only the central LED which represents an illumination existing in both versions. The PiCamera was set to take a picture with 1080p corresponding to a 1920 x 1080px frame. For each optical configuration, a ruler was imaged to calculate, *via* FiJi which is a "batteries-included" distribution of ImageJ (Broeke et al., 2015), the actual size of the field of view from which we can deduce the optical magnification for each pair of M12 lenses. The lateral resolution of each optical configuration was then calculated from the size of the field of view and the width in pixels of the image. The pixel size was deduced from the optical magnification. We found the combination of tube lens with f25mm, and objective lens f16mm provides a large field of view, good depth of focus, and ability to resolve a wide range of planktonic organisms.

To calculate resolution with a 1951 USAF Resolution Target, we found the smallest separable groups and elements for each image of the Resolution Target taken under all the 25 optical configurations. To document the optical characteristics, we calculated the resolution in lp/mm using the following equation ("Edmund Optic" n.d.):

$$Resolution\left[\frac{lp}{mm}\right] = 2^{\left(GroupNumber + \frac{ElementNumber - 1}{6}\right)} \quad (1)$$

To convert lp/mm to microns (μm), simply take the reciprocal of the lp/mm resolution value and multiply by 1000:

$$Resolution\left[\mu m\right] = \frac{1000\,\frac{\mu m}{mm}}{Resolution\left[\frac{lp}{mm}\right]} \quad (2)$$

3.2.1 PlanktoScope Benchmarking

We benchmarked the PlanktoScope by comparing it to the FlowCam (Sieracki et al., 1998) using identical plankton samples. Microplankton samples were collected from subsurface coastal waters in January 2020 in Villefranche Sur Mer, France, by towing a 20μm mesh size plankton net from a kayak. The samples were immediately brought back to the laboratory, split into equal parts after gentle mixing, and imaged alive on both the PlanktoScope v.2 and a Flowcam configured with similar magnification (**Figures 3A, B**). For Flowcam acquisition, a model Benchtop B2 Series equipped with a 4X lens was used. Prior to image acquisition, the sample was passed through a 200μm filter (Überstrainer, PluriSelect inc.) to remove large objects that can clog the capillaries. Samples were imaged on auto-trigger mode (no fluorescence trigger) by passing the sample through a 300μm width glass capillary. Raw images were recorded and processed through ZooProcess according to standardized procedures (Gorsky et al., 2010). Manuals for Flowcam use, including the methodology used, Zooprocess, and Ecotaxa are available at https://sites.google.com/view/piqv. The extracted vignettes were uploaded to ecotaxa and can be consulted with all their associated metadata @: https://ecotaxa.obs-vlfr.fr/prj/2740. For both instruments, the total abundance of organisms, as well as normalized biovolume size, Normalized Biomass Size Spectra (NB-SS) (Platt and Denman 1977) were calculated to evaluate their respective capacity to count and size plankton biodiversity. Both instruments provided enough resolution to allow quantitative taxonomic classification of plankton samples down to the genus, and often species level (**Figure 3C**). Furthermore, the similar NB-SS spectra generated (**Figure 3E**) indicate comparable capacities to measure and count planktonic populations.

3.2.2 Plankton Sampling

The v.1 has been deployed at seven locations representing different ecosystems throughout the planet. The same sampling protocol (except for the Comau Fjord and Palo Alto Baylands Nature Preserve, see below) was performed using a 20 μm mesh plankton net with a diameter of 30 cm, and a 10-minute surface tow at 2 knots. The samples were filtered with a 500 μm mesh sieve to remove larger particles. In Comau Fjord we used a horizontal water sampler from LaMotte (CODE 1087) to sample the vertical distribution of micro-plankton from 0-10 meters below the water's surface. From the 1,200 mL samples collected for every depth, we conserved 15ml and imaged 5.6ml *via* 2000 frames. The salinity at every depth was measured using a hand refractometer from Atago. For the Palo Alto Baylands Nature Preserve, since the site is shallow and quite turbid, the sample was collected directly using a 50mL falcon tube from the subsurface and also filtered using the 500 μm mesh sieve. All samples were collected during the daytime.

The v.2 was first used in a lab context to realize testing on morphological diversity of cultured *Pyrocystis noctiluca* (LB 2504) and *Coscinodiscus wailesii* (CCMP2513) strains (**Figure 2**). Samples were passed directly in the instrument without preliminary concentration. To remove aggregated cells, we placed a mesh filter (Überstrainer, PluriSelect inc.) in between the culture and the field of view. For *Pyrocystis noctiluca*, we used a mesh filter of 200 μm and a μ-Slide I Luer with a channel height of 200 μm. For *Coscinodiscus wailesii*, we used a mesh filter of 500 μm and a μ-Slide I Luer with a channel height of 600 μm. We further tested the v.2 at Villefranche-sur-Mer, France, using plankton samples collected in front of the marine station by towing a 20 μm mesh, 30 cm diameter plankton net for 10 minutes from a kayak.

4 RESULT

4.1 An Open, Modular, and Miniaturized Imaging Platform for Plankton Ecology

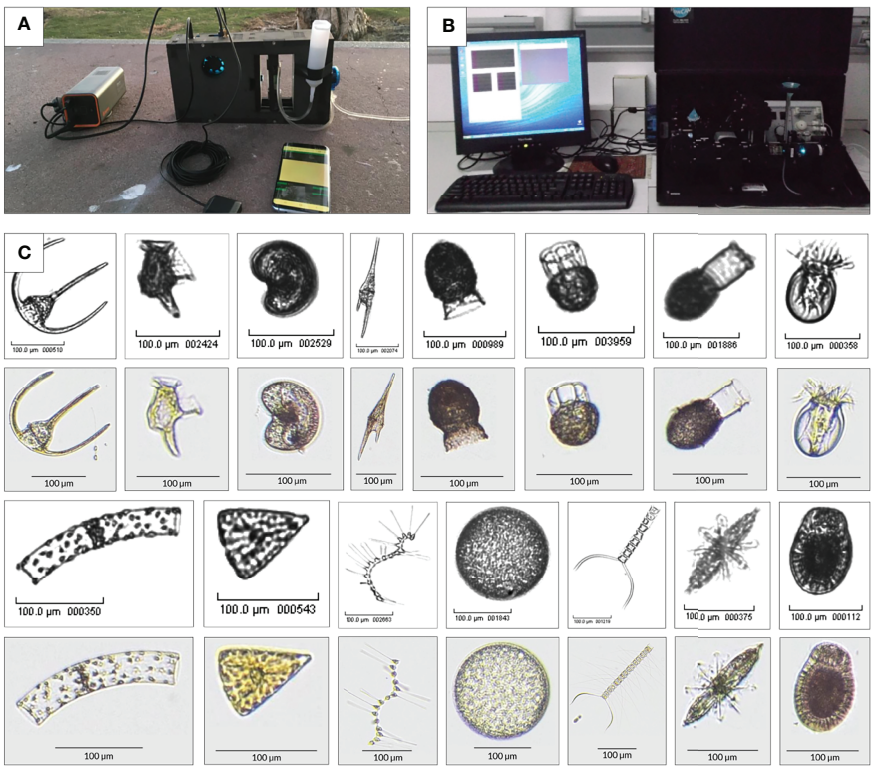
The PlanktoScope was developed in two configurations: v.1, a modular, compartmentalized configuration maximizing multi-functionality and adaptability, and v.2, a compact version designed for rapid assembly, portability, and standardization. Both versions achieve an optical magnification of 1.3X and a pixel size of 0.9μm/px. The travel distance of the specimen stage is about 3.2cm with a step size of 0.15μm to comply with a large range of lens working distances and sample mounting strategies. For a framerate of 8 frames per second and a 500μm thick flow cell, we can image a volume at 0.1ml/min. The components are off-the-shelf and readily accessible from numerous vendors at a low cost to enable replication. The corresponding open-software

TABLE 1 | M12 Lens matrix.

Size of the field of view width x height (mm)		f-number of the objective lens				
		6	8	12	16	25
f-number of the tube lens	6	6.36 x 3.58	4.79 x 2.69	3.25 x 1.83	2.50 x 1.41	1.56 x 0.88
	8	8.31 x 4.68	6.44 x 3.62	4.27 x 2.40	3.32 x 1.87	2.06 x 1.16
	12	12.25 x 6.89	9.37 x 5.27	6.36 x 3.58	4.92 x 2.77	3.07 x 1.73
	16	15.62 x 8.79	11.69 x 6.58	7.90 x 4.44	6.19 x 3.48	3.90 x 2.19
	25	25.00 x 14.06	18.68 x 10.51	14.41 x 6.98	9.80 x 5.51	6.09 x 3.43

strategy utilizes existing libraries for image processing and a flow-based visual programming platform to allow users to rapidly customize acquisition and processing steps. Image segmentation can be toggled for automatic processing after an acquisition sequence, allowing the user to efficiently inspect objects extracted from large volumes, even at low abundance.

The fully modular PlanktoScope v.1 is based on six triangular units (**Figure 1A**), each being a separate functional layer that



Instrument	Flowcam (x4)	PlanktonScope (25/16)
Method	Triggered by optical signal	Flow-stop
Number images	5 432	4 232
Duration (min)	~16	~66
Number objects	10 010	41 000
Volume imaged (ml)	1.93	7.2
Concentration (obj/ml)	~5 180	~5 690
Rapidity acquisition (obj/minute)	~ 625	~ 620
Price (USD)	57 000 to 92 500	452

FIGURE 3 | One-to-one comparison of PlanktoScope with the Flowcam. **(A)** Ultra-portable configuration of the PlanktoScope operated on a cell phone charger and controlled through a user interface on a smartphone. **(B)** Typical setup of a FlowCam on a laboratory bench. **(C)** One-to-one comparison of the same sample (plankton tow, Villefranche/Mer, France) passed through a PlanktoScope and Flowcam (version Benchtop B2 Series). Representative images were chosen from the two data sets (monochromatic images, Flowcam; color images, PlanktoScope) - first row from left to right: *Ceratium* spp., *Dinophysis caudata, Peridiniales* spp., *Ceratium furca, Codonaria* spp., *Dictyocysta* spp., *Codonellopsis* spp., *Undellidae* spp. Second row from left to right: *Guinardia* spp., *Licmophora* spp., *Asterionellopsis* spp., *Coscinodiscophyceae* spp., *Chaetoceros* spp., *Acantharea, unknown* sp. **(D)** Table comparing efficiencies for both trigger-based optical image collection (Flowcam) and flow-stop based wide field of view imaging and computational segmentation (PlanktoScope). When normalized for the total number of objects detected, PlanktoScope performed equally well compared to FlowCam. **(E)** Comparison of total planktonic organisms (objects) sampled with different collection methods and analyzed with different optical/imaging methods as a function of the size of organisms (expressed as equivalent spherical diameter; ESD). Total organism biovolume per size class was expressed as Normalized Biovolume Size Spectra (NBSS) by dividing the total biovolume within a size class by the biovolume interval of the considered size class. NBSS is representative of the number of organisms within a size class. The same plankton net sample was run through a Flowcam and a PlanktoScope v.2. All data are raw counts and converted to biovolume using ellipsoidal calculations. The low count at the smaller size range of each observation corresponds to an underestimation of an object's number due to both the limited capabilities of each imaging device for small objects and net under sampling for small objects utilizing the plankton tow.

FIGURE 4 | Testing of PlanktoScope at seven field sites around the world **(A)**, with sampling and imaging done directly in the field **(B)** for most samples. Composite montages were made to display the objects identified with the highest frequency in each ecosystem, creating a visual representation of local biodiversity. (1) Palo Alto Baylands Nature Preserve (USA) - 1: *Tracheloraphis,* 2: *Tracheloraphis,* 3, 6, 9, 10, 18–23: *Ciliate,* 4: *Unidentified,* 5: Pennate diatom, 7: *Pyrocystis* sp., 8: *Gyrosigma* sp., 11: Pennate diatom, 12: Unidentified, 13: Pennate diatom, 14: *Navicula* sp., 15: Unidentified, 16: *Amphiprora gigantea,* 17: *Enchelyodon.* (2) Monterey Bay (USA) – 1–9: Unidentified, 10–12: Pennate diatom, 13: Centric diatom, 14–17, 20–22: *Odontella longicruris,* 18, 19, 23: Unidentified diatom, 24–26: Armored dinoflagellate (*Protoperidinium*)?, 27, 28: Unidentified Dinoflagellate, 29: *Ornithocercus.* (3) Isla Secas (Panama) - 1: *Nitzschia longissima,* 2, 3, 5, 8: Unidentified, 4: Centric diatom, 6: Copepod fecal pellet, 7: Ciliate, 9: Copepod, 10: Crustacean larvae, 11: Calanoid copepod. (4) Comau Fjord (Chile) – 1–6: Unidentified, 7: Unarmored dinoflagellate, 8, 9: Unidentified Dinoflagellate (resting cyst), 10: *Prorocentrum compressum,* 11: *Dinophysis* sp., 12: *Protoperidinium* sp., 13, 14: *Ditylum brightwellii*, 15: *Detonula pumila*, 16–21: Ciliate, 22–24: *Lepidodinium chlorophorum*, 25–30: *Gyrodinium* sp. (5) Isla Magueyes (Puerto-Rico) - 1: Copepod larva, 2: Nauplius larva, 3, 8: *Chaetoceros* sp., 4, 10, 17: *Oscillatoria* sp., 5, 7: *Eucampia zodiacus,* 6: *Coscinodiscus* sp., 9: Unidentified, 11: Calanoid copepod, 12: *Ceratium furca,* 13: *Ceratium* sp., 14: *Ceratium lineatum,* 15: *Pyrocystis* sp., 16: Unidentified, 18: *Proboscia alata*. (6) Chilika Lake (India) – 1–10, 13-24: Unidentified, 11: Crustacean larva, 12: Nauplius larva, 25: Ciliate. (7) Villefranche/Mer (France) - 1: *Trichodesmium,* 2: Copepoda, 3: Nauplii, 4: Egg, 5: *Rhabdonella,* 6: *Cyttarocylis,* 7: *Undellidae,* 8: *Codonaria,* 9: *Ciliophora,* 10: *Codonellopsis,* 11: *Dictyocysta,* 12: *Chaetoceros,* 13: *Asterionellopsis,* 14: *Bacteriastrum,* 15: Pennate chain, 16, 17: *Licmophora,* 18: *Striatella,* 19: *Rhizosolenia,* 20: *Coscinodiscophyceae*, 21: *Bacillariophyceae,* 22: *Guinardia,* 23: *Dictyochophyceae*, 24: *Acantharea*, 25, 26: Rhizaria, 27, 28: Acantharea, 29: Foraminifera, 30: *Peridinales,* 31: *Pyrocystis, 32, 34*: *Neoceratium,* 33: *Neoceratium ranipes,* 35: *Neoceratium fusus*, 36: *Neoceratium furca,* 37: Dinophyceae, 38: *Neoceratium pentagonum,* 39, 40: Protoperidinium, 41: *Dinophysis caudata,* 42: *Ornithocercus quadratus,* 43: *Ceratocorys.*

couple together through shared optical and electronic paths: 1) a single board computer coupled to its camera sensor, 2-3) two reversed M12 lenses (an objective lens and a tube lens) separated in two modules, 4) a motorized stage and focus delta platform for sample manipulations, 5) independent programmable rings of LEDs for sample illumination, and 6) a peristaltic pump.

The modules are connected mechanically, electronically, and optically, enabling simple re-configuration (**Figure 1B**). The instrument can be used either in the lab (**Figure 1C**) or in the field (**Figure 1D**). The compact Planktoscope v.2 (**Figure 1E**) is a simplification of the platform focusing on robust, flow-through plankton image acquisition in potentially rough field conditions, e.g., sailing boats. A single-board computer controlling a focus actuator holding the flow cell, a single LED for bright-field illumination, a peristaltic pump, and a GPS are all connected through stable electrical wiring.

Both designs use a laser-cut framework and are parametric, enabling the use of different thicknesses of the chosen material. These range from acrylic and recycled plastic, to wood, metal, or fiberboard. This machining strategy allows rapid design iteration and enables a precise yet flexible low-cost method for aligning and spacing optical components.

In the modular PlanktoScope v.1, three magnets are incorporated into the corners of the interface between modules (**Figure 1A**) enabling both proper alignment of the six units and quick reconfiguration. The microscope can be used in vertical or horizontal configurations, placed upright or inverted, depending on the need or constraints of the experimenter. For example, a vertical mode enables manual exploration of a static sample that can be placed on a glass slide, flow-cell, petri dish, or optical chambers (**Figure 1B**). The delta stage enables tracking of an organism with high precision or a quick survey of the sample holder area. A horizontal mode allows automated, continuous imaging of liquid samples passing through the flow cell at a predefined rate (**Figure 1C**). On the other hand, the compact, flow-through PlanktoScope v.2 uses a minimal structure to position and align the components, to increase robustness and stability for field deployment or *in-situ* installation. Modularity is still maintained by allowing the lenses and flow-cell to be quickly interchanged. While the modular version requires 10 h for the machining, soldering, and assembly, the compact version drastically reduces the build complexity enabling a complete machining/assembly in less than 4 h.

Both prototypes are based on a Raspberry Pi single-board computer that controls the electronics, acquires and processes the images, and serves as the user/machine interface (**Figure 1F**). The magnetic coupling of the modular PlanktoScope enables electronic connectivity through the contact of copper ribbons that connect each module at their interface to form a custom BUS for Inter-Integrated Circuit (I^2C) connection and power. The different independent microcontrollers, here Arduinos, receive queries as actuators and send logs as sensors back to the Raspberry Pi. The compact version utilizes Pi HATs (Hardware Attached on Top) for both assembling and deploying code. The Pi HATs enable the rapid addition of numerous off-the-shelf specialized boards. Three HATs are utilized to serve different functions: one for cooling the CPU of the Raspberry Pi and providing visual feedback, one for controlling the focus stage and the pump, and a third HAT supporting a GPS for geolocalization of the images. Thanks to the massive community built around Raspberry Pi, hundreds of other possibilities exist for new modules and more functions built on top of this platform. Both instruments can be powered through either standard wall Alternative Current (AC) power or from battery cells for field deployment (**Figure 5A**). For an acquisition frequency of 0.5 Hz and a standard Lithium-ion or polymer battery of 20,000 mAh, the compact version can collect continuously for more than 8 hours.

Both versions of the PlanktoScope utilize a Raspberry Pi camera sensor. The Pi Camera V2.1 uses a Sony sensor with a still resolution of 8Mp, and a sensor imaging area of 3.68 x 2.76mm for $25. The $40 HQ Pi camera with an imaging area of 6.287mm x 4.712 mm and a 12Mp resolution can easily be incorporated. We used the high-performance and frugal lenses in a compact form factor 'M12' (corresponding to the metric tapping dimension) for magnification, building upon existing successful strategies for constructing low-cost microscopes (Switz et al., 2014). Two M12 lenses were conjugated to construct a reconfigurable solution to project the image of a microscopic object to a camera sensor. By using different focal lengths for both lenses, measuring the size of the field of view, and calculating the resolution of each combination, we obtained a comparative matrix of 25 different optical configurations from low (0.3X) to high (4X) magnification (**Supplementary Material Figure 5C**). These offer a pixel size from 4.5µm to 0.3µm (**Supplementary Material Figure 5D**) and a measured resolution from 15.6µm to 1.9µm (**Supplementary Material Figure 5E**). Since each version allows magnetic swapping of both lenses, all the described optical configurations are readily interchangeable.

4.2 Proof of Concept in Both Laboratory and Field Conditions

4.2.1 Monitoring and Phenotyping Lab Cultures

The PlanktoScope is designed with rapid adaptability in mind, so it can be transported quickly from designated use in the field to controlled data collection in a laboratory setting. We used the continuous flow mode to image monocultures of unicellular eukaryotes and benchmark the PlanktoScope's ability to function as a lab culture monitoring system. First, a culture of the diatom *Coscinodiscus wailesii* was passed through the system to monitor viability over time. Processed images provided straightforward classification and quantification of dead and living cells (**Figure 2C**). Second, the large transparent dinoflagellate *Pyrocystis noctiluca*, an organism that exhibits morphological changes linked to circadian cycles (Seo and Lawrence, 2000), was imaged in flow mode across the day-to-night transition. We could observe various cell morphologies (**Figure 2D**), including cell-cycle states, and built a diagram of temporal phenotypes. As circadian clocks function as major drivers of behavior in most marine life (Seo and Fritz, 2001), such controlled continuous monitoring is a source of informative, non-invasive, and easily accessible information on any cultured strain, providing valuable morphological, physiological, and behavioral data to improve culture and experimental conditions.

FIGURE 5 | PlanktoScope assessment of micro-plankton biodiversity along a vertical gradient in a Chilean fjord. **(A)** The extreme salinity gradient was measured in the Patagonian Comau fjord. Using sampling at discrete depths and the PlanktoScope, we attempted to describe correlations between salinity and plankton abundance/morphology across depths. Samples were collected with a Niskin bottle at different depths and imaged under a PlanktoScope. The plot depicts a vertical snapshot of an ecosystem from 0m (surface water) to 10m (depth) with the number of identified objects and equivalent diameter (10 to 76 µm) as a function of depth (0 to 10m). The measurement of the elongation per equivalent diameter is based on 94,262 objects detected in total. Color bar represents the aspect ratio from purple (small aspect ratio) to green (large aspect ratio). **(B)** Display of distribution of objects with a mean size of 54 µm as a function of aspect ratio, equivalent diameter, and depths (0 to 10m). **(C)** Illustration of objects with a gradual aspect ratio from 1.2 to 6.4 marked in **(B)**.

4.2.2 Field Plankton Ecology

The PlanktoScope platform has been primarily designed for field deployment by both experienced and citizen scientists. It is ultra-portable and battery-powered (**Figure 3A**), able to be transported and used for the duration of a cruise or deployed in the field with the use of a dedicated power supply. We tested the PlanktoScope's robustness, simplicity of use, and capability to acquire high-quality and reproducible data during seven field trips worldwide (**Figure 4A**). By generating panels of the most frequently extracted objects (**Figures 4.1–4.7**), we show how the instrument can rapidly provide qualitative plankton biodiversity surveys of any water body.

We next leveraged the PlanktoScope's portability combined with a quantitative sampling strategy to tackle an ecological question in a Patagonian fjord. The Comau Fjord in southern Chile receives 5 m of rain per year per square meter, with numerous freshwater rivers and streams feeding into the saltwater bay. This provokes a vertical salinity gradient that evolves seasonally with sporadic weather led events such as wind and rain (Buskey and Hyatt, 2006; León-Muñoz et

al., 2018). In April 2019, a typical vertical stratified salinity gradient was visible, with the salinity gradually increasing from 25‰ at the surface to a plateau of 32‰ below the pycnocline at a 5-meter depth (**Figure 5A**). The PlanktoScope was used to investigate whether this salinity gradient corresponded to stratified planktonic communities. We observed a > 10-fold increase in the number of images in the meter of saline water immediately beneath the pycnocline. This number falls off around 9 m below the surface, or around 4 meters below the pycnocline (**Figure 5A**). We then attempted to extract geometrical characteristics from that dataset that could be ecologically informative. By measuring the ratio between the maximal and minimal length of detected objects (i.e., the aspect ratio, quantifying elongation) across plankton size fractions and depth (**Figure 5A**), we observed a relation between the increasing number of detected objects and their aspect ratio. Most of the detected objects collected below the pycnocline and with an equivalent diameter between 36µm and 64µm had an aspect ratio of around 4, indicating the presence of a large population of elongated objects. By further exploring the vertical stratification of objects within the size fraction 53µm to 55µm, where the aspect ratio is on average about 5.9 at 6 meters, we found many elongated plankton (**Figures 5B, C**). This is consistent with previous observations that organisms living at depth in higher nutrient concentrations favor elongated morphologies with a higher aspect ratio [(Colin) Reynolds, 1988; Bauer et al., 2013; Ryabov et al., 2020]. Mining visual data and combining morphological attributes such as these with geochemical measurements can help better describe the regional microbial ecology.

5 CONCLUSION AND PERSPECTIVE

The basis of the largest ecosystem on Earth, the Ocean, lies within the planktonic organisms invisible to the naked eye. Today, we need to magnify not only this hidden world but also our approach. An unfortunate but common reality that limits long-term surveys of planktonic communities is the high cost and erratic funding situations associated with marine research. The significant resources involved in maintaining oceanographic vessels and instruments often cripple the ability to fund studies through time, and even more in low to middle-income countries. As 40% of the world's population lives along coastlines, monitoring these ecosystems remains incredibly important. There are still relatively few long-term time or spatial series that visually document planktonic ecosystems, and a clear tendency to quantify biodiversity and associated ecosystem services using costly and complex protocols based on high-throughput DNA sequencing. These same metagenomic studies have unveiled the massive biodiversity of micro-eukaryotes (de Vargas et al., 2015) with mostly unknown functions Carradec et al., 2018 in planktonic ecosystems. These organisms, essentially protists, are often more complex than metazoans in terms of cell structures, symbiotic interactions, and behavior Gavelis et al., 2015, (Vincent et al., 2018), properties that cannot readily be inferred from genomic data Keeling, 2019. To quantify and understand the role of micro-eukaryotic complexity in ecosystem functions, it is critical to develop instruments allowing their high-throughput imaging worldwide.

The PlanktoScope is a low cost, versatile, and high-resolution digital microscope designed to enable professional and citizen scientists to perform large scale surveys of planktonic life. We have demonstrated here its capacity to monitor morphology and physiology of eukaryotic cells in culture, or to quantify fundamental features of micro-plankton communities in a coastal water column directly in the field in Chile, a country where single blooms have created losses of over 800 million dollars locally, leading to major public health crisis (Mardones et al., 2021). Further demonstration of the PlanktoScope v.2's efficiency for quantitative imaging is presented in the same issue (Mériguet et al., this issue), showing how the PlanktoScope and Flowcam provided comparable data while sampling along a basin-scale transect of the Tara schooner from Lorient (Britanny, France) to Punta Arenas (Chile). The current version of the PlanktoScope is limited in the plankton size range (50-200µm) it can recognize and quantify, however its fundamental modularity and relative simplicity make it possible to implement future new modules to analyze smaller or larger plankton.

The foundation of PlanktoScope lies in the principles of open-source hardware and software, combined with an open yet cohesive community of engineers, makers, researchers, and citizens in daily contact with sea-water (i.e., seatizen of the 'Plankton Planet' initiative, see de Vargas et al., 2022). Current trends in affordable electronics and distributed manufacturing, together with computer vision and automated image processing, make it possible to put instruments' manufacturing and data collection in the hands of thousands of users across the planet. Therefore, we have also launched web tools to share, develop, and replicate the PlanktoScope globally. Instructions to order and/or manufacture the different components and assemble them into a functional instrument are available @ www.planktoscope.org. The PlanktoScope community shares experiences, technical advice, and ideas for new developments @Slack (https://www.planktoscope.org/join). Between May 2020 and Dec 2021, over 286 individuals representing a large spectrum of activities from 28 countries (**Figure 6C**) have engaged in this community. While the canonical version(s) of the PlanktoScope are being and will be developed and deployed for global standard measures plankton life by the Plankton Planet team (see de Vargas et al. 2022), we know of at least 32 functional instruments replicated, and sometimes modified, by colleagues around the world (**Figure 6**).

Deploying a high-throughput frugal microscope platform on a global scale will bring light to the habitats under-surveyed by the large and more infrequent research cruises. Connecting these platforms with a network of climate researchers, ecologists, citizen scientists, and many others across the planet will bring further relevance to each individual measurement and build global capacity to explore our microscopic world. Since cost remains one of the key barriers to engagement in science, we intend to use "frugal science" to greatly enhance affordable approaches to scientific inquiry.

FIGURE 6 | Documenting community replication of the Planktoscope. **(A)** Images of Planktoscopes built and implemented by the community from 2020-2021. First row (left to right), built by: Salima Rafai, Laboratoire Interdisciplinaire de Physique, CNRS - Université Grenoble Alpes; Guillaume Le Guen. Konk Ar Lab, Le Temps des Sciences and Saint Brieuc Factory; Ana Fernandez Carrera, Biological Oceanography, Leibniz Institute for Baltic Sea Research Warnemünde (IOW); Dyche Mullins, Mullins Lab, University of California - Second row (left to right): Andrian Gajigan, School of Ocean and Earth Science and Technology, the University of Hawaii at Manoa; Bronwyn Lira Dyson, Experimental Limnology, Leibniz Institute of Freshwater Ecology and Inland Fisheries (IGB); Guillaume Bourdin, School of Marine Sciences, University of Maine; PlanktoSquad, Dalhousie University - Third row (left to right): Stewart Plaistow, Institute of Integrative Biology, University of Liverpool; Alex Barth, Department of Biological Sciences, University of South Carolina; Macci Wigginton, Ocean & Earth Sciences, Old Dominion University; Yefim Radomyselskiy, Department of Physics, City University of New York - Queens College. **(B)** Field deployments of the Planktoscope by community members. From left to right: v.2 on a NSF science-cruise in 2021; v.2 by a river bed in Northwest France; v.2.5 onboard *Tara* during a cross Atlantic cruise (see Mériguet et al. this issue); v2.5 on a small sailboat off the coast of Southeast France. **(C)** Planktoscope community across the world as of December 2021.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

TP: design of the PlanktoScopes hardware, work on the software, data analysis, community building and management, wrote the initial draft of the manuscript. AL: sample acquisition, data analysis, community building and management, wrote the initial draft of the manuscript FL: planktoscope development and testing (both hardware and software), data analysis and interpretation, field work. HL: hardware development and help with project conception. DG: hardware development. SC: study design, hardware development CV: conception and supervision of the work, provided initial funding, wrote the manuscript. MP: conception, design, and supervision of the work, field work, data analysis, provided supplementary funding All authors revised and edited the manuscript, and approved the final version sent for publication.

FUNDING

This work was initially funded by the 'Plankton Arts' grant from the Fondation d'Entreprise Total to CV and MP. Additional funding came from the Simons Foundation (AL), the Schmidt Futures, Moore Foundation, NSF Center for Cellular Construction (NSF STC award DBI 1548297), CZ BioHub, and HHMI-Gates Faculty Fellows Program (MP), the New Zealand Royal Society Strategic Seeding Fund (16-CAW-008-CSG), the Institut Universitaire de France (FL), and the French Government "Investissements d'Avenir" program OCEANOMICS (ANR-11-BTBR- 0008) (CV, SC) and EMBRC-France (ANR-10-INBS-02).

ACKNOWLEDGMENTS

We acknowledge all members of PrakashLab and Plankton Planet core team, as well as Romain Bazile, Anna Oddone, Simon-Martin Schröder, Daniel Elnatan, B.B. Cael, Rainer Kiko, Marie Walde, Guillaume Bourdin, Aurélie Labarre, and Flora Vincent for insightful comments and suggestions throughout this work. We thank members of the EMBRC platform PIQv for image analysis, as well as Jorge Mardones, Lara Zamora (IFOP, Chile), and Gurudeep Rastogi (Wetland Research and Training Center, Chilika Development Authority) for support during field testing in Chile and on the lake Chilika. This article is contribution number 2 of Plankton Planet.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.949428/full#supplementary-material

REFERENCES

(Colin) Reynolds, C. S. (1988). *Functional Morphology and the Adaptive Strategies of Freshwater Phytoplankton*. Cambridge Univ. Press. United Kingdom. 442 pp

Bauer, B., Sommer, U. and Ursula, G. (2013). High Predictability of Spring Phytoplankton Biomass in Mesocosms at the Species, Functional Group and Community Level. *Freshw. Biol.* 58 (3), p588–596. doi: 10.1111/j.1365-2427.2012.02780.x

Broeke, J., Perez, J. M. M. and Javier, P. (2015). *Image Processing With ImageJ* (Packt Publishing Ltd). United Kingdom.

Buskey, E. J. and Hyatt, C. J. (2006). Use of the FlowCAM for Semi-Automated Recognition and Enumeration of Red Tide Cells (Karenia Brevis) in Natural Plankton Samples. *Harmful Algae*. 5 (6), p685–692. doi: 10.1016/j.hal.2006.02.003

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018). A Global Ocean Atlas of Eukaryotic Genes. *Nat. Commun.* 9 (1), 373. doi: 10.1038/s41467-017-02342-1

Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., et al. (2021). Environmental Vulnerability of the Global Ocean Epipelagic Plankton Community Interactome. *Sci. Advances August*. 7 (35), peabg1921. doi: 10.1126/sciadv.abg1921

Cybulski, J. S., Clements, J. and Manu Prakash, (2014). Foldscope: Origami-Based Paper Microscope. *PLoS One* 9 (6), e98781. doi: 10.1371/journal.pone.0098781

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015). Eukaryotic Plankton Diversity in the Sunlit Ocean.*Science* 348, 6237. 1261605.

de Vargas, C., Le Bescot, N., Pollina, T., Henry, N., Romac, S., Colin, S., et al. (2022) Plankton Planet: a frugal, cooperative measure of aquatic life at the planetary scale. *Frontiers*.

Duarte, C. M. (2015). Seafaring in the 21St Century: The Malaspina 2010 Circumnavigation Expedition. *Limnology Oceanography Bull.* 24, 11–14. doi: 10.1002/lob.10008

Edmund Optic *Exploring 1951 USAF Resolution Targets*. Available at: https://www.edmundoptics.fr/knowledge-center/application-notes/testing-and-detection/choosing-the-correct-test-target/ (Accessed July 20, 2020). ldquo] [rdquon.d.

Efatmaneshnik, M. and Ryan, M. j. (2016). On Optimal Modularity for System Construction. *Complexity*. 21 (5), 176–189. doi: 10.1002/cplx.21646

Fenchel, T. (1988). Marine Plankton Food Chains. *Annu. Rev. Ecol. Systematics*. 1, 19–38. doi: 10.1146/annurev.ecolsys.19.1.19

Field, C. B. (1998). Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science*. 281 (5374), p237–240. doi: 10.1126/science.281.5374.237

Gavelis, G. S., Hayakawa, S., White, R. A., 3rd, Gojobori, T., Suttle, C. A., Keeling, P. J., et al. (2015). Eye-Like Ocelloids Are Built From Different Endosymbiotically Acquired Components. *Nature* 523 (7559), 204–207.

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital Zooplankton Image Analysis Using the ZooScan Integrated System. *J. Plankton Res.* 32 (3), 285–303. doi: 10.1093/plankt/fbp124

Guidi, L., Chaffron, S., Lucie, B., Damien, E., Abdelhalim, L., Simon, R., et al. (2016). Plankton Networks Driving Carbon Export in the Oligotrophic Ocean. *Nature* 532 (7600), 465–470.

Henson, S. A., Sanders, R. and Esben, M. (2012). Global Patterns in Efficiency of Particulate Organic Carbon Export and Transfer to the Deep Ocean. *Global Biogeochemical Cycles*. 26 (1), pGB1028. doi: 10.1029/2011gb004099

Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., De Vargas, C., Raes, J., et al. (2011). A Holistic Approach to Marine Eco-Systems Biology. *PLoS Biol.* 9 (10), e1001177.

Kautsky, U., Saetre, P., Berglund, S., Jaeschke, B., Nordén, S., Brandefelt, J., et al. (2016). The Impact of Low and Intermediate-Level Radioactive Waste on Humans and the Environment Over the Next One Hundred Thousand Years. *J. Environ. Radioactivity* 151 Pt 2, 395–403.

Keeling, P. J. (2019). Combining Morphology, Behaviour, and Genomics to Understand the Evolution and Ecology of Microbial Eukaryotes. *Philos. Trans. R. Soc. London. Ser. B Biol. Sci.* 374 (1786), 20190085. doi: 10.1098/rstb.2019.0085

León-Muñoz, J., Urbina, M. A., Garreaud, René and Iriarte, JoséL. (2018). Hydroclimatic Conditions Trigger Record Harmful Algal Bloom in Western Patagonia (Summer 2016). *Sci. Rep.* 8 (1), 1330.

Lombard, F., Boss, E., Anya, M., Meike, W., Uitz, V. J., Stemmann, L., Sosik, H. M., et al. (2019). Globally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Mardones, J. I., Paredes, J., Godoy, M., Suarez, R., Norambuena, L., Vargas, V., et al. (2021). Disentangling the Environmental Processes Responsible for the World's Largest Farmed Fish-Killing Harmful Algal Bloom: Chile 2016. *Sci. Total Environ.* 766, 144383. doi: 10.1016/j.scitotenv.2020.144383

Martini, S., Larras, F., Boyé, A., Faure, E., Aberle, N., Archambault, P., et al. (2021). "Functional Trait-based Approaches as A Common Framework for Aquatic Ecologists." *Limno. Oceanog.* 66 (3), 965−994.

Ruuskanen, M. O., Sommeria-Klein, G., Havulinna, A. S., Niiranen, T. J. and Lahti, L. (2021). "Modelling Spatial Patterns in Host-Associated Microbial Communities." *Environ. Microbiol.* 23, (5), 2374−2388.

Ryabov, A., Kerimoglu, O., Litchman, E., Olenina, I., Roselli, L., Basset, A., et al. (2020). Shape Matters: The Relationship Between Cell Geometry and Diversity in Phytoplankton. *Biorxiv*. 24 (4), p847−861. doi: 10.1101/2020.02.06.937219

Seo, K. S., and Lawrence, F. (2000). "Cell Ultrastructural Changes Correlate with Circadian Rhythms in Pyrocystis Lunula (Pyrrophyta)." *J. Phycology* 36 (2), 351−358.

Seo, K. S., and Lawrence, F. (2001). "Evidence for Sexual Reproduction in the Marine Dinoflagellates, Pyrocystis Noctiluca and Pyrocystis Lunula (Dinophyta)." *J. Phycology* 37.4, 530-535.

Sieracki, C. K., Sieracki, M. E. and Yentsch, C. S. (1998)"An Imaging-in-Flow System for Automated Analysis of Marine Microplankton." *Mar. Ecol.* Progress Series 168, 285-296.

Sosik, H. M., and Olson, R. J. (2007). Automated Taxonomic Classification of Phytoplankton Sampled With Imaging-in-Flow Cytometry. *Limno. Ocean. Methods* 5 (6), 204−216.

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and Function of the Global Ocean Microbiome. *Science* 348 (6237), 1261359.

Sunagawa, S., Acinas, S. G., Bork, P., Bowler, C., Eveillard, D., Gorsky, G., et al. (2020). Tara Oceans: Towards Global Ocean Ecosystems Biology." *Nat. Rev. Microbiol.* 18 (8), 428−445.

Switz, N. A., D'Ambrosio, M.V. and Fletcher, D. A. (2014). Low-cost Mobile Phone Microscopy With A Reversed Mobile Phone Camera Lens. *PloS one* 9 (5), e95330.

van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). "*Scikit-Image: Image Processing in Python.*" *PeerJ* 2, e453.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., (2004) Environmental Genome Shotgun Sequencing of the Sargasso Sea. *Science* 304, 5667. 66−74

Vincent, F. J., Colin, S., Romac, S., Scalco, E., Bittner, L., Garcia, Y., et al. (2018). The Epibiotic Life of the Cosmopolitan Diatom Fragilariopsis Doliolus on Heterotrophic Ciliates in the Open Ocean. *ISME J.* 12 (4), 1094−1108.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

ORIGINAL RESEARCH
published: 09 August 2022
doi: 10.3389/fmars.2022.894372

Global Distribution of Zooplankton Biomass Estimated by *In Situ* Imaging and Machine Learning

Laetitia Drago[1], Thelma Panaïotis[1], Jean-Olivier Irisson[1], Marcel Babin[2], Tristan Biard[3], François Carlotti[4,5], Laurent Coppola[1,6], Lionel Guidi[1], Helena Hauss[7], Lee Karp-Boss[8], Fabien Lombard[1,9], Andrew M. P McDonnell[10], Marc Picheral[1], Andreas Rogge[11], Anya M. Waite[12], Lars Stemmann[1*†] and Rainer Kiko[1*†]*

[1] Sorbonne Université, Laboratoire d'Océanographie de Villefranche-sur-mer, Villefranche-sur-mer, France, [2] Takuvik International Research Laboratory, Québec Océan, Laval University (Canada) - Centre National de la Recherche Scientifique (CNRS), Département de biologie and Québec-Océan, Université Laval, QC, Canada, [3] Laboratoire d'Océanologie et de Géosciences (LOG), Univ. Littoral Côte d'Opale, Univ. Lille, Centre National de la Recherche Scientifique (CNRS), UMR 8187, Wimereux, France, [4] Département Ecologie Marine et BIOdiversité (EMBIO), M.I.O. Institut Méditerranéen d'Océanologie Bâtiment Méditerranée, Marseille, France, [5] Laboratoire d'Océanographie Physique et Biologique (LOPB), case 901 13288, Marseille, France, [6] Sorbonne Université, Centre National de la Recherche Scientifique (CNRS), OSU STAMAR, Paris, France, [7] Department Ocean Ecosystems Biology, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, [8] School of Marine Sciences, University of Maine, Orono, ME, United States, [9] Institut Universitaire de France (IUF), Paris, France, [10] Oceanography Department, University of Alaska Fairbanks, Fairbanks, AK, United States, [11] Section Benthopelagic Processes, Alfred Wegener Institute Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany, [12] Ocean Frontier Institute and Oceanography Department, Dalhousie University, Halifax, NS, Canada

OPEN ACCESS

Edited by:
Sophie G. Pitois,
Fisheries and Aquaculture Science (CEFAS), United Kingdom

Reviewed by:
Dong Sun,
Ministry of Natural Resources, China
Kohei Matsuno,
Hokkaido University, Japan

***Correspondence:**
Laetitia Drago
laetitia.drago@imev-mer.fr
Lars Stemmann
lars.stemmann@imev-mer.fr
Rainer Kiko
rainer.kiko@imev-mer.fr

†These authors share last authorship

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 11 March 2022
Accepted: 05 May 2022
Published: 09 August 2022

Citation:
Drago L, Panaïotis T, Irisson J-O, Babin M, Biard T, Carlotti F, Coppola L, Guidi L, Hauss H, Karp-Boss L, Lombard F, McDonnell AMP, Picheral M, Rogge A, Waite AM, Stemmann L and Kiko R (2022) Global Distribution of Zooplankton Biomass Estimated by In Situ Imaging and Machine Learning. Front. Mar. Sci. 9:894372. doi: 10.3389/fmars.2022.894372

Zooplankton plays a major role in ocean food webs and biogeochemical cycles, and provides major ecosystem services as a main driver of the biological carbon pump and in sustaining fish communities. Zooplankton is also sensitive to its environment and reacts to its changes. To better understand the importance of zooplankton, and to inform prognostic models that try to represent them, spatially-resolved biomass estimates of key plankton taxa are desirable. In this study we predict, for the first time, the global biomass distribution of 19 zooplankton taxa (1-50 mm Equivalent Spherical Diameter) using observations with the Underwater Vision Profiler 5, a quantitative *in situ* imaging instrument. After classification of 466,872 organisms from more than 3,549 profiles (0-500 m) obtained between 2008 and 2019 throughout the globe, we estimated their individual biovolumes and converted them to biomass using taxa-specific conversion factors. We then associated these biomass estimates with climatologies of environmental variables (temperature, salinity, oxygen, etc.), to build habitat models using boosted regression trees. The results reveal maximal zooplankton biomass values around 60°N and 55°S as well as minimal values around the oceanic gyres. An increased zooplankton biomass is also predicted for the equator. Global integrated biomass (0-500 m) was estimated at 0.403 PgC. It was largely dominated by Copepoda (35.7%, mostly in polar regions), followed by Eumalacostraca (26.6%) Rhizaria (16.4%, mostly in the intertropical convergence zone). The machine learning approach used here is sensitive to the size of the training set and generates reliable predictions for abundant groups such as Copepoda (R2 ≈ 20-66%) but not for rare ones (Ctenophora, Cnidaria, R2 < 5%). Still, this study

offers a first protocol to estimate global, spatially resolved zooplankton biomass and community composition from *in situ* imaging observations of individual organisms. The underlying dataset covers a period of 10 years while approaches that rely on net samples utilized datasets gathered since the 1960s. Increased use of digital imaging approaches should enable us to obtain zooplankton biomass distribution estimates at basin to global scales in shorter time frames in the future.

Keywords: global zooplankton, *in situ* imaging, biomass, machine learning, underwater vision profiler (UVP), spatial distribution, boosted regression trees (BRT), habitat modeling

1 INTRODUCTION

1.1 Zooplankton

Present in all the oceans of the globe, zooplankton corresponds to organisms adrift in the water. They represent a great taxonomic diversity and sizes, ranging from a few micrometers to several meters (de Vargas et al., 2015; Karsenti et al., 2011; Stemmann and Boss, 2012). Zooplankton play a central role in the carbon cycle as they contribute to the biological pump that drives the export of photosynthetically fixed organic carbon from the surface to the intermediate and deep oceans (Longhurst and Glen Harrison, 1989; Turner, 2002; Turner, 2015; Steinberg and Landry, 2017). As a major link between primary producers and higher trophic levels (Ikeda, 1985), zooplankton have central ecological and biogeochemical roles, with associated socio-economic interests. This socio-economic impact of plankton can be positive, such as their role as food source for fish (Lehodey et al., 2006; van der Lingen et al., 2006) or as an indicator of water quality (Suthers et al., 2019). It can also be negative, as e.g. jellyfish blooms that can impact various human activities such as aquaculture and fishing (Richardson et al., 2009).

1.2 Spatial Distribution of Zooplankton and Its Biomass

Zooplankton organisms are sensitive to environmental conditions and are thus considered sentinels of ocean changes. Their distribution is finely governed by the interactions between physical [i.e., temperature (Steinberg and Landry, 2017), currents, light (Hays et al., 2005), pressure] and chemical constraints [nutrients, oxygen (Steinberg and Landry, 2017)], but also by biological interactions (e.g. predator-prey, symbiosis, parasitism and commensalism). The dependence of zooplankton on environmental variables leads to very clear global scale patterns even at coarse taxonomic levels (Lucas et al., 2014; Biard et al., 2016). On a global scale, zooplankton diversity is higher at the equator and decreases towards the poles (Rombouts et al., 2009; Ibarbalz et al., 2019). Conversely, zooplankton biomass tends to be low in the tropics and increase with latitude with large seasonal fluctuations in temperate and polar regions (Ikeda, 1985; Moriarty et al., 2012; Soviadan et al., 2022). Although a global quantitative assessment of zooplankton biomass and functional groups is needed (e.g. to be incorporated in biogeochemical and ecological models), it is often hampered by the heterogeneity of sampling methods and the uneven distribution of observations, causing high uncertainty in biomass estimates (Moriarty et al., 2012; Moriarty and O'Brien, 2013; Le Quéré et al., 2016).

1.3 The Study of Zooplankton and Its Difficulties

Assessments of the global distribution of zooplankton organisms are often based on regional datasets, obtained with heterogeneous sampling tools traditionally biased towards non-gelatinous taxa (Lucas et al., 2014), and combined using different standardization procedures (Moriarty et al., 2012; Moriarty and O'Brien, 2013; Buitenhuis et al., 2013). Consequently, the global distribution of only a few zooplankton groups that generally can be well sampled using plankton nets, e.g. crustaceans, have been well studied (Rombouts et al., 2009; Buitenhuis et al., 2013). Indeed, some zooplankton taxa are known to be fragile (cnidarians, ctenophores, rhizarians, etc.) and their destruction by plankton nets as well as their poor preservation in fixatives (Beers and Stewart, 1970) resulted in an underestimation of their biomass and their ecological role in marine ecosystems (Lucas et al., 2014; Biard et al., 2016). In this context, non-intrusive *in situ* methods using imaging (Remsen et al., 2004; Cowen and Guigand, 2008; Sun et al., 2008; Stemmann et al., 2008; Schulz et al., 2010; Picheral et al., 2010; Grossmann et al., 2015) and video (Davis et al., 1992; Davis et al., 2005; Hoving et al., 2019) instruments have been developed (Lombard et al., 2019). Among the different systems, only the Underwater Vision Profiler (UVP) version 4 and 5 have been widely used for plankton on a global scale which allowed comparisons of abundance patterns with the Longhurst (1995) provinces of the ocean (Stemmann et al., 2008; Biard et al., 2016). Since 2008, the creation and expansion of such a global dataset could be executed with the UVP5 thanks to numerous participating teams around the world and the wide commercialization of this *in situ* imaging tool. In this study, we used data from the UVP5, an *in situ* imaging system designed to detect, measure and quantify the distribution of zooplankton organisms and marine particles (Picheral et al., 2010). This instrument, designed for the study of particle size spectra in the ocean (Stemmann et al., 2002; Guidi et al., 2009) was also previously used to obtain plankton data at a high spatial resolution (Forest et al., 2012) and to study fragile organisms (Biard et al., 2016; Stukel et al., 2018; Christiansen et al., 2018; Biard and Ohman, 2020). However, even with the progressive increase in the spatio-temporal density of observations allowed by the use of imaging instruments, the unevenness in the distribution of observations remains,

preventing large scale biomass estimations. Such global observations could nevertheless serve as the basis for large scale estimations through the use of interpolation or extrapolation methods, including statistical habitat models.

1.4 Statistical Habitat Models

Habitat modeling is a machine learning tool to estimate the abundance of a taxon at a location where an observation is missing: instead of interpolating between nearby observation points based on geographical distance, the environmental conditions (i.e. the habitat) are used to inform the estimation. Statistically, a regression analysis can be used to define the relationship between the abundance (or presence) of a taxon at observation sites and the environmental variables at those sites (Guisan and Zimmermann, 2000; Elith and Leathwick, 2009). Then, continuous maps of those environmental variables can be used to predict continuous maps of the taxon's abundance (or presence), by applying the regression.

The objective of this work was the development of a method to estimate zooplankton biomass on a global scale and to study the spatial distribution of zooplankton in relation to its habitat. To obtain such a global view we used global data from the UVP5 *in situ* imaging system. In most cases, it is difficult to identify the imaged organisms to species level. We therefore applied the habitat modeling approach to broader taxonomic groups. We first estimated the individual biovolume and biomass of organisms classified in 25 broad taxonomic groups, within a global *in situ* imaging dataset. We then applied the habitat model methodology to each taxonomic group and built models using different regional and vertical partitions of the data. We separated data of the epipelagic (0-200 m depth layer) from the upper mesopelagic (200 to 500 m depth layer). We also used a global partitioning to separate data from low latitudes (40°S to 40°N) from the remaining high latitude data. We hypothesize that these partitions should allow us to separate subgroups within those broad taxa, which occupy different horizontal and/or vertical habitats. Finally, we used the models' output to estimate the global marine zooplankton biomass distribution in the top 500 m of the water column.

In situ imaging observations with UVP5 have been widely used during the past decade to study zooplankton in the global ocean. Biard et al. (2016) used 694 stations from the UVP5 dataset to reveal that Rhizaria were strongly underestimated in previous studies. Here, we use an updated version of this dataset, now including 3,549 stations to study the biomass distribution of Copepoda, Rhizaria and several other groups of planktonic organisms in the 1.02-50 mm size range. We hypothesize that the total biomass of zooplankton is distributed according to regional production characteristics, associated with climatic and hydrological patterns, showing overall a high biomass in high latitudes and lower values in the subtropical gyres (Ikeda, 1985; Moriarty et al., 2012).

2 MATERIALS AND METHODS

2.1 Plankton Data Collection and Processing

2.1.1 Global Plankton Imaging With the UVP5

UVP5 data (**Figure 1**) were compiled from all oceans, covering a 10 year period (2008-2018). A detailed description of the operation of the UVP5 is given in Picheral et al. (2010). All particles large than ≈ 100 μm in Equivalent Spherical Diameter (ESD) were measured and counted, but only images of particles (zooplankton and aggregates) larger than ≈ 600 μm ESD were kept by the UVP5 for further processing because smaller objects contained too few pixels to be identifiable. Acquisition of metadata (geographic location, date, etc.) and processing of all 8.46 million images (95% being detritus) were carried out by the ZooProcess software which provided information on 42 morphological features associated with each object (area,

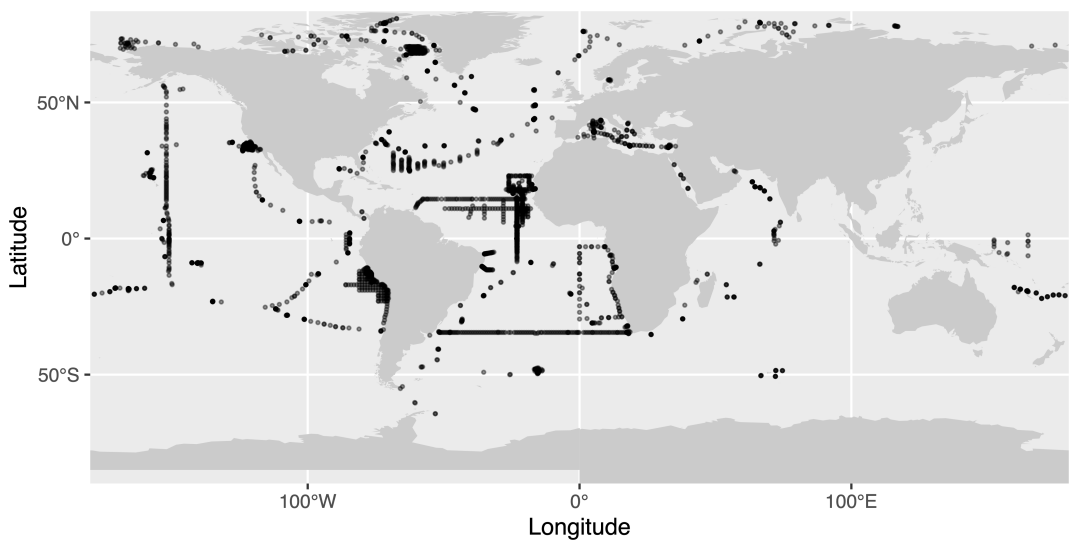


FIGURE 1 | Map of the UVP5 dataset used in this study. Transparency was used to illustrate the density of points on the map.

major and minor axis, etc.). The results were imported into EcoTaxa (Picheral et al., 2017), an application which allows a taxonomic classification of images *via* supervised learning algorithms, followed by manual validation (Irisson et al., 2022). As 61% of the profiles have a maximum depth ≤500 m, only images of organisms between 0-500 m were kept and the overall estimates of biomass were restricted to this depth range. To ensure that profiles were representative, a filter was also applied to only keep profiles that covered at least 80% of the layer of interest.

2.1.2 Image Classification and Size Range Covered

Living organisms were separated from detritus (aggregates, fibers, fecal pellets) as well as artifacts (e.g. bubbles) and classified according to their taxonomic identity. Recognition and sorting of organisms can be a source of bias depending on the levels of perception and experience of the people who perform them. Several cognitive factors biases such as boredom, fatigue or a classification biased towards the most used groups have been presented by Culverhouse (2007) and Culverhouse et al. (2014). To reduce the risk of poor identification, a shared UVP5 taxonomic guide was used to homogenize image sorting into 119 taxonomic groups. The image data were thereafter grouped into 25 broader taxonomic groups (**Table S1**), and a subset of the resulting dataset was checked for homogeneity of sorting within these groups. A minimum of 51 images and a maximum of 10% of all images were extracted from each group and were independently checked after the assembly of the final data set. The maximum error or uncertainty rate per taxon was 9.8% and a vast majority of taxa were under 2.5%. We checked the classification and if accuracy was <95%, we rechecked the categories to assure proper sorting. In addition, only fully validated profiles were used for this analysis. The resulting global data set consisted of 466,872 images from 3,549 stations. Under-sampled groups with less than 500 images in the dataset which could not be used for a global study were not included in the analysis.

We computed the organisms' size spectrum to detect the size range within which the UVP5 can be used to properly quantify their distribution. The concentration of objects in the ocean is expected to decrease with size; when this is computed as a normalized size spectrum, the relationship is expected to be linear (Forest et al., 2012). A peak in the size spectrum at the lower size range generally reflects the minimum size of efficient detection by *in situ* imaging while high variability in the large size range reflects the poor ability to detect rare large objects (Stemmann and Boss, 2012). With that in mind, the spectrum was linear for the size range 1.02-50 mm and organisms outside this range were not included in the analysis since large mobile fauna (including large crustaceans) are likely to be undersampled and small zooplankton organisms close to the UVP5's threshold of detection are difficult to identify. This size range selection ensures that the data used in this study was properly quantified by the UVP5.

2.1.3 Individual Biomass Estimation

To avoid errors due to incorrect ellipse fits (around appendages of organisms rather than their body, ellipse fitted to non-ellipsoidal organisms, etc.), we chose the spheroid method: it is based on the area (**Table 1**), which is more consistently measured by the image analysis performed in ZooProcess.

For Rhizaria, biovolume (mm^3) to carbon (mgC) conversions were done using factors from the literature (**Figure S1** and **Table S2**). For other groups, the conversion from individual volume to individual wet weight assumed a density of 1 g cm^{-3} (Kiørboe, 2013). Then the conversion from individual wet weight to individual biomass in carbon units (mgC) was calculated using taxon-specific linear conversion factors from McConville et al. (2016); when several conversion factors were available for a taxon, their median was used for each group. To take into account differences in density of some parts of the organisms, the Appendicularia group was actually split into Appendicularia_body and Appendicularia_house, whereby the "body" group contains images with only the animal and the "house" group contains the house and the animal. For the images labeled Appendicularia_house, we used the relationship of house diameter (major axis) to Appendicularia trunk length from Lombard and Kiørboe (2010). We then converted this body size equivalent into carbon weight using the corresponding relationship from Lombard et al. (2009). For the images labeled Appendicularia_body, we converted the biovolume of the organism into carbon weight using the corresponding relationship from Lombard et al. (2009). Two groups also have been created to separate the Collodaria into solitary Collodaria and colonial Collodaria. This choice was done based on the fact that solitary Collodaria are smaller than colonial ones and have a different vertical distribution (Faillettaz et al., 2016). For solitary collodarians with a dark central capsule (subgroup of solitary Collodaria) described in Biard et al. (2016), the estimation of carbon (0.189 $mgC\ mm^{-3}$) by Mansour et al., (2021) was done on the capsule of the organisms. As Zooprocess measures the area of the whole organism, we determined the ratio $\frac{area\ \ whole\ \ organism}{area\ \ central\ \ capsule} = 0.713$ and applied this factor to avoid overestimation of carbon biomass for this group. For the rest of the collodarians, the estimation of Mansour et al., (2021) was directly applied.

2.2 Environmental Data Collection and Processing

In order to develop relationships between regional characteristics of the environment (**Figures S2–4**) and observed biomass, climatologies from the World Ocean Atlas (WOA) (Garcia et al., 2019) were used for temperature (in °C), salinity, oxygen (converted from $\mu mol\ kg^{-1}$ to kPa for better physiological interpretation), and macronutrients (nitrate, phosphate and silicate in µmol kg^{-1}). We selected the data sets defined on a 1° horizontal grid, over the 0-500 m depth range, and with a monthly temporal resolution.

TABLE 1 | Methods of calculating individual biovolume with area (mm2); ESD, the equivalent spherical diameter equivalent (mm); major, the major axis (mm) of the best fit ellipse; minor, the minor axis (mm) of the most suitable ellipse.

Method	Formula
Spheroid	$\frac{4}{3} \times \pi \times (\frac{ESD}{2})^3$ with ESD = $2 \times \sqrt{\frac{Area}{\pi}}$
Ellipsoid	$\frac{4}{3} \times \pi \times \frac{major}{2} \times (\frac{minor}{2})^2$

Temporal coverage was from 2005 to 2017 for salinity and temperature and 1955 to 2017 for the other variables. We also used monthly averaged surface chlorophyll-a data (Chl *a* in mg m^{-3}) resolved to 1/24° from 2005 to 2017 from the Copernicus database (OCEANCOLOUR_GLO_CHL_L4_REP_OBSERVATIONS_009_082) as well as bathymetry data from NOAA (Amante and Eakins, 2009) with a spatial resolution of 10 minutes; both were re-gridded to a 1° grid. Finally, distance to coast was computed by calculating the distance of all 1°×1° cells to the closest cell associated to land using the raster package (Hijmans, 2021). To obtain annual climatologies, when relevant, each monthly variable was averaged over its time period of coverage.

This environmental data was then matched to the UVP5 data on the 1°×1° grid. Since the 1°x1° grid used by WOA does not necessarily follow the contour line of the coast perfectly, some UVP5 profiles could not be directly matched to the environmental grids. This is mostly the case where e.g. the coast is situated in a 45 degree angle to latitude or longitude, thereby creating triangle shaped areas that are not covered by the rectangular grid. For profiles that lie in such corners of the grid, we used the environmental values of the closest neighboring 1°×1° WOA cell. In the epipelagic world model, 3,002 points did have a direct match while 156 points did not have a direct match. Out of these 156 points, 14 were not in a neighboring 1°×1° WOA cell and were removed from the model input. For the mesopelagic, 2,172 did have a direct match, while 104 points had a match in a neighboring grid cell and 2 points did not and were removed from the model input. Maps that show the close vicinity of non-matching points to adjacent WOA cells are shown in **Supplementary Figure 5**.

To assess whether we are able to describe various environmental conditions with the UVP5 samples, we compared the distributions of each variable in the worldwide WOA dataset and in the subset matched to UVP5 profiles (**Figures S6, S7**). Although the geographical coverage is not homogeneous (**Figure 1**), the coverage of environmental conditions is good and warrants the use of habitat models.

2.3 Habitat Modeling

The steps of this process are summarized in **Figure 2**.

2.3.1 Modeling Tools

In this work we used boosted regression trees (BRTs) to predict the biomass of different zooplankton groups as they show different advantages over other commonly used machine

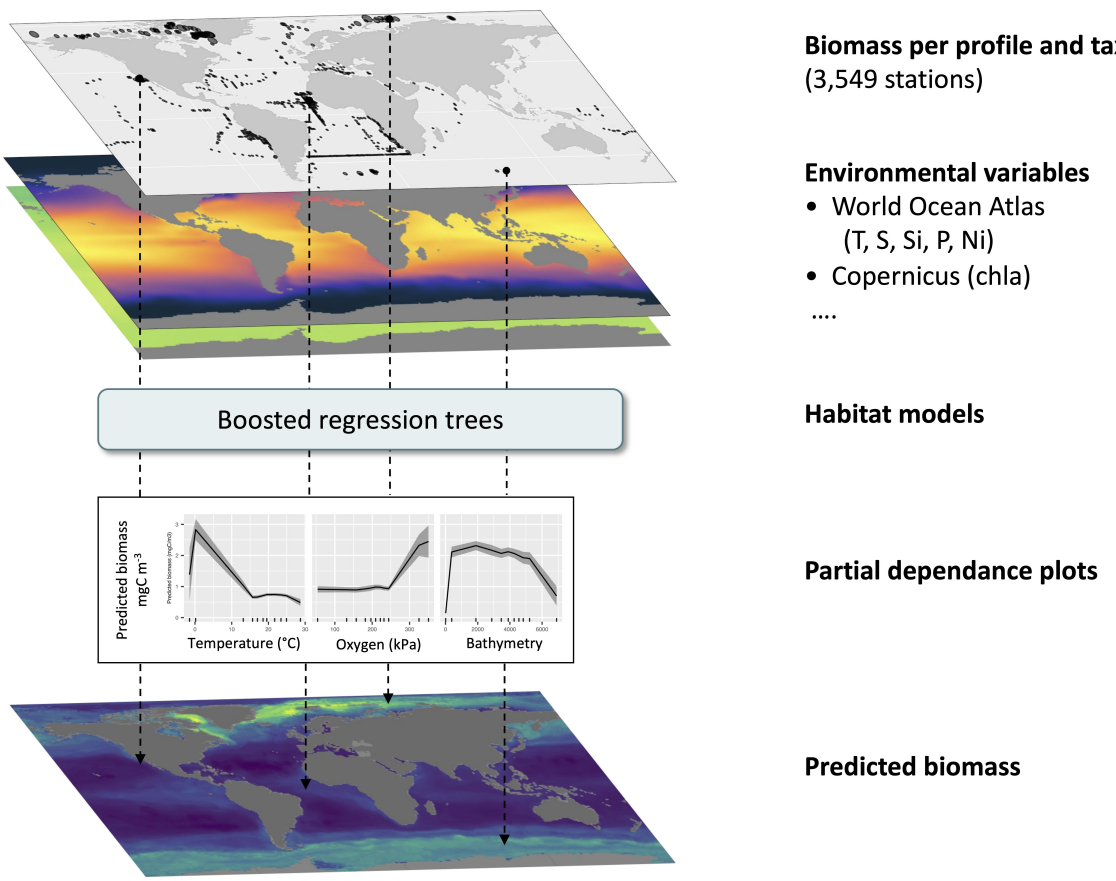


FIGURE 2 | Methodology followed from data selection to prediction of global biomass.

learning approaches for the nature of our dataset and intended application (Elith and Graham, 2009). This ensemble method uses regression trees, models that link a response (here biomass) to predictors (environmental variables) by successive dichotomous separations (Breiman et al., 1984; Hastie et al., 2001). Regression trees automatically select the relevant explanatory variables, can deal with categorical or continuous inputs, are not sensitive to the distribution of the continuous ones, can represent relations of arbitrary form and naturally include interactions among explanatory variables (Elith et al., 2006). With so-called surrogate splits, they can also deal with missing values in the explanatory variables. They are therefore very convenient to use, but their predictive power is often limited and they have difficulties to capture smooth relationships. Boosting is a way to overcome these drawbacks (Schapire, 2003). It is based on the fact that it is easier to find many rough rules of thumb than to find a single, highly accurate prediction rule (Schapire, 2003). BRTs combine many short regression trees in succession, each new tree being adjusted to consider the observations poorly predicted by the previous ones (Elith et al., 2006; Leathwick et al., 2006; Elith et al., 2008). This improves predictive performance and the smoothness of the prediction (Leathwick et al., 2006). In addition, only a random subset of the input data is used to fit each tree and this stochastic component reduces the variance of the final model ensemble (Friedman, 2002).

Boosted regression trees (BRTs) have an ability to handle a large number of variables and - other than Generalised Linear Models (GLMs, Nelder and Wedderburn (1972)) or Generalised Additive Models (GAMs, Hastie and Tibshirani (1986); De'ath (2007); Elith et al. (2008)) - do not seek to fit one single model portraying the relationship of the response variable (here biomass) and its predictors (environmental variables). Various recent studies (González Carman et al., 2019; Chen et al., 2020; Hu et al., 2021) have compared BRTs results to other modeling tools such as GAMs, GLMs, Random Forests (RFs), Maximum Entropy modeling (Phillips et al., 2006; Elith and Graham, 2009) or neural networks and have obtained better predictive performance with BRTs. Other studies (Zhang et al., 2018; Son et al., 2018) used complementary GAMs and BRTs to study the effects of explanatory variables. However, BRTs could be slower than RFs (Chen et al., 2020) and training parameters need to be chosen carefully to avoid overfitting (Leathwick et al., 2006; Elith and Graham, 2009). BRTs were chosen over RFs because of their capacity to reduce both the bias and the variance of model results (Hastie et al., 2001). BRTs are also less sensitive to the effect of extreme outliers and the inclusion of irrelevant predictors (Leathwick et al., 2006). This makes them suitable for plankton datasets, as sometimes very high plankton biomass values do occur during blooms (Brodeur et al., 2018; Pettitt-Wade et al., 2020). BRTs also have the ability to handle sharp discontinuities which is not the case of the GAMs (Elith et al., 2008). This is important when modeling taxa which can have a narrow habitat.

In addition, in regression trees, the loss function, used to determine which dichotomous split to perform, can be changed to be adapted to the distribution of residuals. Here we explored the classic mean squared error, which assumed a somewhat normal distribution of the residuals, as well as a Tweedie loss adapted to zero-inflated data (Zhou et al., 2019), and a Poissonian loss, which considered data as discrete counts, also including many zeros. To use the Poisson loss, the biomass was scaled so that the value of the 1% quantile was ≥ 1 and then rounded to the nearest integer; the inverse scaling was performed after prediction. This later approach proved to produce the best fits and more robust models in a few test taxa and all models were therefore fitted with Poisson loss. The models and statistics were computed using the xgboost package (Chen et al., 2021) in R version 4.1.2 (R Core Team, 2021).

2.3.2 Spatial Partitioning of the Data

Individual biomass values derived from UVP5 images and environmental data measured at various layers were both averaged over a depth range of interest and matched geographically, on the 1°×1° grid. Biomass values matched to the same 1° pixel, and therefore associated to exactly the same environmental data, were averaged.

We hypothesized that an association between biomass and environment investigated at a fine scale could be more efficiently learned by the model because is contains less noise, so we divided the data vertically between the epipelagic (0-200 m) and mesopelagic (200-500 m) zones and also tried a finer partition, into 100 m depth bins between 0 and 500 m. Evaluating separate models for each layer could allow to focus on finer subgroups within our quite coarse taxonomic units (some species being mostly present in one of the layers) and therefore define biomass-habitat relationships at a finer, more relevant biological level.

For the same reason, we also built models on subsets of data partitioned geographically. Indeed, polar copepods have a different thermal niche compared to tropical ones (Rombouts et al., 2009; McGinty et al., 2021). So, in addition to a model fitted on the global dataset (world), we trained models on data from the region between 40°S and 40°N (low latitude) and from the data collected outside of this latitudinal band (high latitude). Out of the 3,549 profiles composing the UVP5 dataset, 2,837 are located between 40°S and 40°N and 712 were done outside of this latitudinal band.

2.3.3 Data Splits for Model Training, Assessment and Evaluation

For each taxon in each spatial partition, the data was split to distribute 80% of it in a training and validation sets, on which the model was fitted and assessed, and 20% to a test set, on which predictive performance was evaluated. This split was stratified according to the deciles of biomass in the data, to ensure that both the learning and test sets contained low and high biomass points.

To choose model hyperparameters (i.e. parameters of the model adjustment algorithm) and to evaluate the variability in the prediction due to the constitution of the training set, each 80% portion set was resampled through five-fold cross validation repeated 20 times [i.e. 100 resamples; (Hastie et al., 2001)]. For

each cross-validation fold, the model was actually trained on four folds and validated on the last one. The splits into the five folds were also stratified according to the deciles of biomass, for the same reason invoked above.

2.3.4 Selection of Hyperparameters and Model Evaluation

To extract as much information from the data, while avoiding overfitting, various combinations of hyperparameters were tested for each model (Elith et al., 2006). They included: 1) the learning rate per tree determining the contribution of each tree to the ensemble model (0.05, 0.08 and 0.1 were tested); 2) the maximum depth of a tree (2, 4 and 8 were tested); 3) the minimum number of elements per leaf (which also limits the depth of the trees; 1, 3 and 5 were used); 4) the number of trees used for the prediction (values up to 600 were tested). For each combination, the model was fitted to the training set and evaluated on the validation set of each of the 100 resamples; the loss was then averaged over the 100 resamples. The best set of hyperparameters is usually the one for which this average loss is minimal. The differences around that minimum are often small and not always meaningful; to be sure to avoid overfitting, we applied an early stopping criterion whereby the increase in the number of trees was stopped when the error did not decrease by more than 1% after adding 10 trees.

Once the best set of hyperparameters had been chosen, the relevance of the corresponding model was quantified by the Pearson correlation between the observed biomass data in the test set and the predicted biomass, where prediction is the average of the predictions of the 100 models fitted to the resamples. This metric captures the model's ability to correctly represent general trends and patterns in the data set and is one way to compute the R^2. The significance of this correlation can also be tested and quantified with a *p*-value. These metrics can be readily compared across the various spatial partitions of the data because they represent the skill of the models on an independent data set, not the quality of the fit to the training data (like the way the R^2 is usually computed). To compare the worldwide and regional approaches fairly, it is important to focus on the same regional subset. To this effect, two additional R^2 were computed for the global model: on the test data located inside the 40°S-40° N latitudinal band and on those outside of it (world low latitude and world high latitude).

2.3.5 Effect of Environmental Variables

To identify which environmental variables drive the change of biomass in each specific model, the percentage of variance explained by each variable was calculated as the sum of the effects of the variable at each node of each tree where it was used. To describe the shape of the effect of each variable, univariate partial dependence plots were computed as the average ± standard deviation marginal effect of the variable in the 100 resamples. Practically, the variable of interest was set at a given value at all training points, the other variables were left at their original values, the average biomass predicted over all points was computed, for each resample; then the mean and standard deviation of those averages were computed across resamples. Finally, the variable was set to another value and so on. To describe the full range of each variable, the partial dependence was estimated at 10% quantile.

2.3.6 Extrapolation to the Globe

To obtain global maps of predicted biomass, the regression between UVP5 biomass data and environmental variables was applied to all points in the corresponding partition of the world, in depth and space. Because 100 models were fitted to the resamples of the training data, the standard deviation of biomass among the 100 predictions (σ_b) can be computed in addition to the mean (m_b), and the coefficient of variation (CV), defined as $CV = \frac{\sigma_b}{m_b}$, then gives an indication of the uncertainty of the model predictions.

To get a robust estimate of global zooplankton biomass in the 1.02 mm to 50 mm size range, we chose to be conservative (i.e. *ad minima*): only the taxonomic groups in the global partition for which the correlation between predicted and observed biomass was significant were used. The surface area of each 1°×1° cell was computed using the following formula:

$$A = \frac{\Pi}{180} \times R^2 \times (\sin(lat_S) - sin(lat_N)) \times 10^6$$

with the area A in m^2, the south and north latitudinal limits of the cell in radians and $\mathcal{R}$, the earth radius (6,378.137 km). For each group used, the biomass was integrated over the relevant layer in each 1°×1° cell by the following calculation

$$\hat{b}_t = \hat{b} \times A \times l$$

where $\hat{b}$ is the estimated biomass in $mgC.m^{-3}$, A in m^2 is defined above, l is the layer thickness in m and therefore $\hat{b}_t$ is the total biomass in *mgC*. Finally, the global *ad minima* zooplankton biomass estimate was computed by adding up the biomass for all selected groups and the 0-200 and 200-500 m depth layer.

3 RESULTS

3.1 Model Comparison

We estimated model performance on the worldwide UVP5 dataset and on a spatial partition of the dataset in low (inside 40°N and 40°S) and high latitudes (outside of the 40°N-40°S latitudinal band) as well as on different depth layers. We hypothesized that a finer data selection might enable the respective model to learn the regional or depth specific habitat more appropriately. Yet, this also meant fitting models to fewer data points. In the end, we find that no clear trend emerges from the relevant comparisons (**Figure 3**): global models are better in 13 comparisons and partitioned models are better in 14 comparisons, whereas for 11 comparisons no clear decision can be made. Comparisons can only be made within a given depth layer between the same regional partitions (e.g. world low latitude only containing the data predicted by the global model between 40°N-40°S vs low latitude; world high latitude only containing data north of 40°N and south of 40°S from the global model vs high latitude).

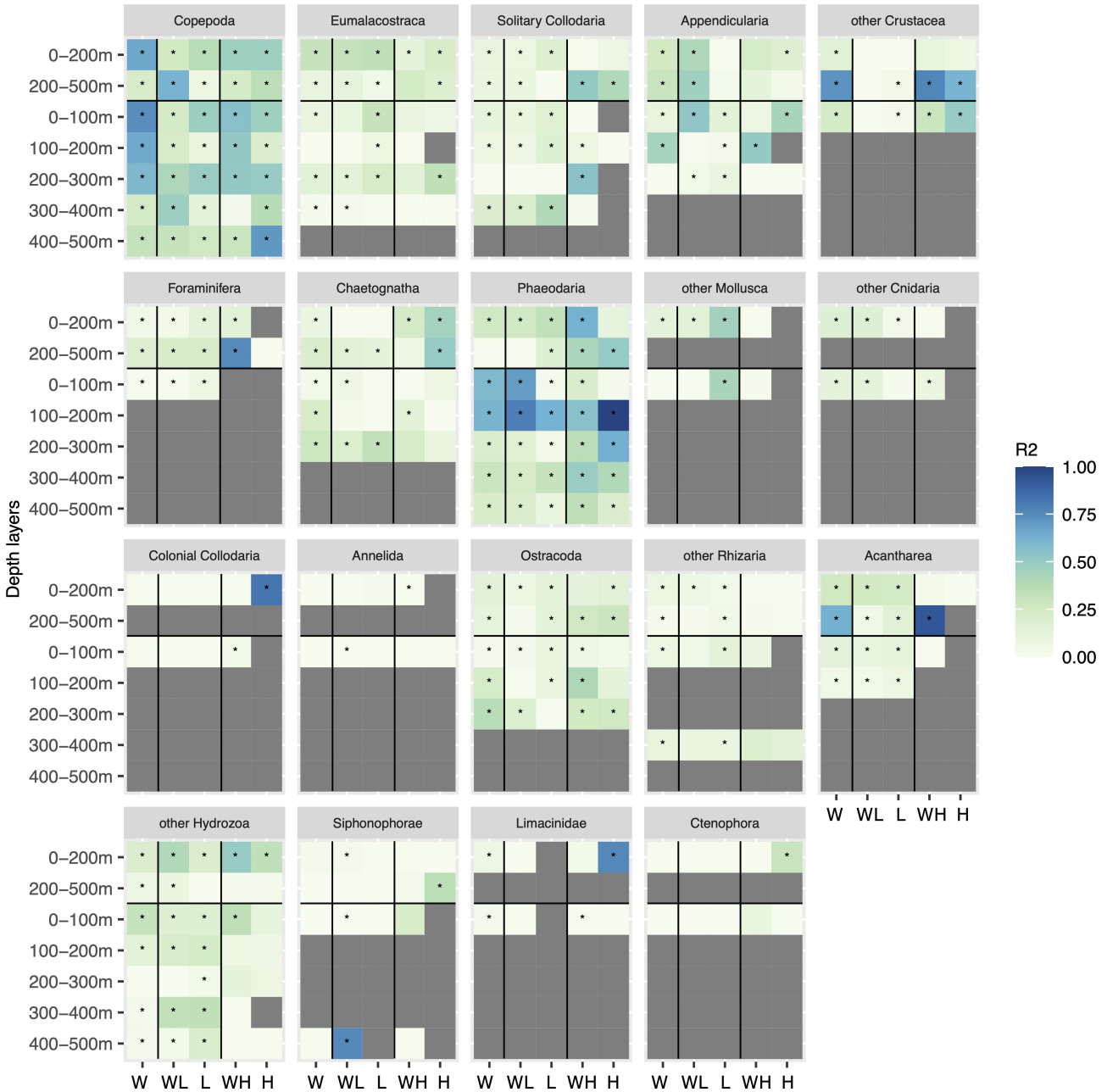


FIGURE 3 | Heatmap of the models' R^2 between observed and predicted biomass for all zooplankton groups arranged from the most important in terms of biomass (Copepoda) to the least important (Limacinidae) in the different depth layers. The regions correspond to: W for world (model run on all data); WL for world low (data between 40°N and 40°S from the world model); L for low latitude (model run between 40°N and 40°S); WH for world high (data outside of 40N and 40S from the world model); H for high latitude (model run outside of 40°N and 40°S). The stars indicate significant results (p-value < 0.05) obtained with the Pearson correlation test.

For some groups such as Annelida and some Mollusca, the high latitude model could not be computed (symbolized by a grey cell) either because they were considered as rare (< 500 images in the layer modeled) or because the model could not learn the link between biomass and environment for this group. However, for other taxa such as Copepoda, solitary Collodaria or Phaeodaria, high and low latitude models are generally better than the world model, as indicated by a higher R^2 value (**Figure 3**). In the epipelagic layer, for Copepoda, the R^2 of world low latitude is 0.26 vs 0.37 in the low latitude model. For the mesopelagic, low latitude has an R^2 of 0.07, lower than the one for world low latitude (0.62). For Appendicularia in the

epipelagic layer, the best R^2 values are obtained in the world low latitude (0.41) and world high latitude (0.24) models respectively compared to low latitude (0.01) and high latitude (0.19).

As for the vertical 100 m-bin layers partition, we obtained the best results overall with the global model. The finer vertical definition also gives better results for multiple other groups such as Appendicularia, Phaeodaria and Ostracoda between 0 and 300 m. In most cases, only the top 100 m layer model worked for this 100 m vertical partition. Overall, the most consistently good choice, when considering all taxa, is a worldwide model fitted separately to the epipelagic (0-200 m) and mesopelagic (200-500 m) layers. This is therefore the configuration retained for the total, global biomass estimate. In **Figure 3**, taxa are arranged in decreasing order of global biomass in the epipelagic layer. For the top five taxa [Copepoda ($R^2 = 0.66$), Eumalacostraca ($R^2 = 0.31$), solitary Collodaria ($R^2 = 0.10$), Appendicularia ($R^2 = 0.26$) and other Crustacea ($R^2 = 0.15$)], the correlation between true and predicted biomass is significant (p-value < 0.05) in the epipelagic worldwide model. In the mesopelagic layer, the correlations for all five groups are also significant (p-value < 0.05 with respective R^2 of 0.22, 0.10, 0.09, 0.30 and 0.72).

3.2 Group-Wise Contribution to Global Zooplankton Biomass

Figure 4 shows the biomass per group predicted for the three spatial partitions and divided into the epi- (0-200 m) and mesopelagic (200-500 m) layers. For the worldwide model, the dominant groups in terms of biomass in the epipelagic were Copepoda (0.083 ± 0.020 PgC), Eumalacostraca (0.058 ± 0.017 PgC) and solitary Collodaria (0.038 ± 0.008 PgC) (**Figure 4**). Among the groups displaying a significant correlation (p-value < 0.05) between true and predicted biomass (and therefore retained for the global estimate), crustaceans (Copepoda, Eumalacostraca, other Crustacea and Ostracoda) represented 68.4% (0.157 PgC) of the biomass in this layer; Rhizaria (solitary Collodaria, Foraminifera, Phaeodaria, other Rhizaria and Acantharea) made up 20.6% (0.047 PgC); but the Cnidaria (other Cnidaria and other Hydrozoa) represented only 0.56% of the global zooplankton biomass (0.0013 PgC). In other words, Crustacea and Rhizaria together made up ~89.1% of the biomass predicted in the epipelagic layer. In the deeper mesopelagic layer, Copepoda (0.061 ± 0.016 PgC) were still the dominant group in terms of biomass, followed by Eumalacostraca (0.049 ± 0.014 PgC) and other Crustaceans (0.017 ± 0.001 PgC) combined. Crustacea (Copepoda, Eumalacostaca, other Crustacea and Ostracoda) represented 0.129 PgC, equivalent to 74.4% of this layer's biomass, while Rhizaria (Foraminifera, solitary Collodaria, other Rhizaria and Acantharea) totaled 0.014 PgC, representing 10.1%, equivalent to most of the remaining biomass in the layer. When combining the results from these two layers, Copepoda represented 44.4% of the global integrated biomass, followed by Eumalacostraca (15.6%), solitary Collodaria (13.1%) and other Crustacea (11.2%). More broadly, Crustacea (Copepoda, Eumalacostraca, other Crustacea and Ostracoda) represented 0.222 PgC or 71.3% of the biomass predicted over 0-500 m, while Rhizaria (Foraminifera, solitary Collodaria, other Rhizaria and Acantharea) made up 0.019 PgC or 10.8% of biomass.

Copepoda were particularly dominant in high latitudes, especially in the epipelagic layer. In the low latitude model, solitary Collodaria contributed most in the epipelagic, followed by Eumalacostraca, Copepoda and Foraminifera. Eumalacostraca dominated biomass in the mesopelagic layer in low latitudes followed by Copepoda and Foraminifera.

3.3 Spatial Distribution Patterns and Occupied Habitat

Presenting the global distribution patterns of all zooplankton groups is beyond the scope of this paper. Instead, we focus on the results for the three groups contributing most to the total global biomass (Copepoda, Eumalacostraca and Solitary Collodaria) as well as on Phaeodaria and Acantharea, Rhizarians that were shown to be important contributors to zooplankton biomass that are underestimated by net-based sampling (Biard et al., 2016). The predicted fields for all modeled groups will be made available in the GitHub repository linked in the data availability statement upon publication of the article.

3.3.1 Copepoda

Copepoda is one of the best predicted groups in the epipelagic ($R^2 = 0.66$), likely because it is the most abundant. The structuring environmental variables were different for the epi- (**Figures S8A, B**) and mesopelagic layers (**Figures S8C, D**): temperature (33%) and oxygen (19%) for the former and temperature (29%), bathymetry (19%) and chlorophyll *a* (15%) for the latter. The highest copepod biomass in the top 200m was found in high latitudes (**Figure 5A**), where water temperature is low and oxygen concentrations are relatively high. In the mesopelagic layer (**Figure 5B**), high copepod biomass was associated with shallow coastal and cold water masses. The patterns of distribution predicted by the global models were similar in both layers (**Figures 5A, B**), with the highest predicted biomass values in the Baffin Bay, Labrador Sea and Greenland Sea as well as at the Southern Ocean polar front region. The lowest predicted biomass was predicted at oceanic gyres and in the Arctic, north of 80°N. For both layers, the highest values of the coefficient of variation (**Figure 5C**) were found north of Canada and Greenland, as well as south of 60°S, especially for the epipelagic layer. These high values depict disagreement among the 100 models fitted to the data resamples and therefore inform on the uncertainty of the model in these zones. Caution is therefore advised regarding the interpretation of the very low values of biomass predicted in those regions. In the northern hemisphere, except for the Arctic ocean, the values of the coefficient of variation were rather low at locations where either low or high biomass values were predicted. In the southern hemisphere, model predictions varied relatively strongly at the level of the Antarctic polar front (**Figures 5C, D**).

3.3.2 Eumalacostraca

Eumalacostraca contains mostly vignettes of euphausiids, amphipods and decapods. They were predicted globally with an R^2 of 0.31 for the epi- and 0.1 for the mesopelagic layer, both

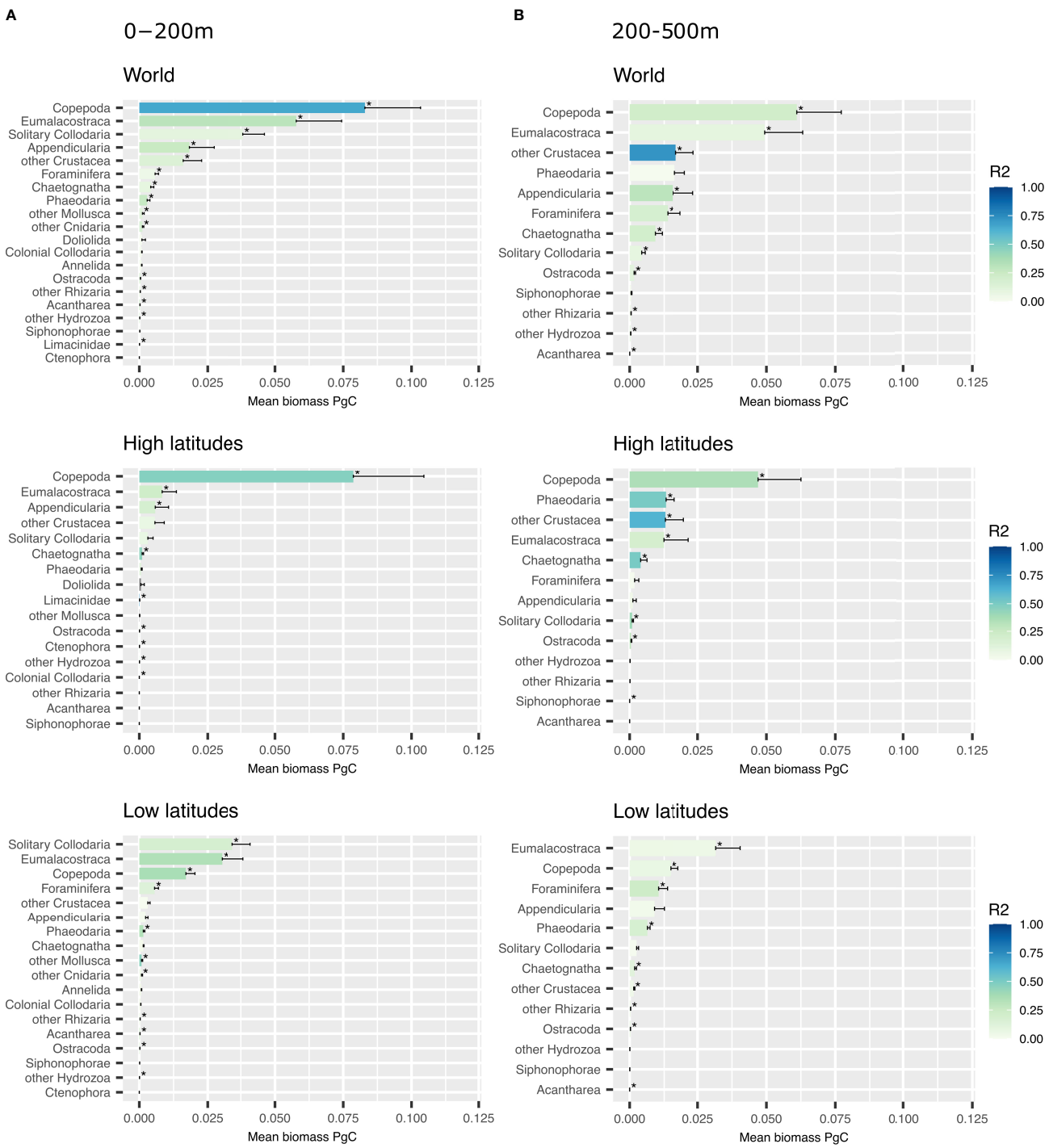


FIGURE 4 | Barplots showing the mean biomass predicted in PgC at 0-200 m **(A)** and 200-500 m **(B)** depth for each group ranked from highest to lowest biomass in 3 types of models: world, outside 40°N-40°S and inside 40°N-40°S. Error bars correspond to upper interval of the biomass estimation's standard deviation. The stars indicate a significant result (p-value < 0.05) obtained with the Pearson correlation test.

with significant p-values (p-value < 0.05; **Figure 3**). In the epipelagic, high biomass of these organisms was associated with high concentrations of phosphate (22%) and low concentrations of silicate (17%) (**Figures S9A, B**). In the mesopelagic layer, the distribution of this group was associated with low concentrations of silicate (16%), bathymetry (15%) and high chlorophyll α (15%) (**Figures S9C, D**). In terms of spatial distribution, high biomass is predicted in eastern boundary

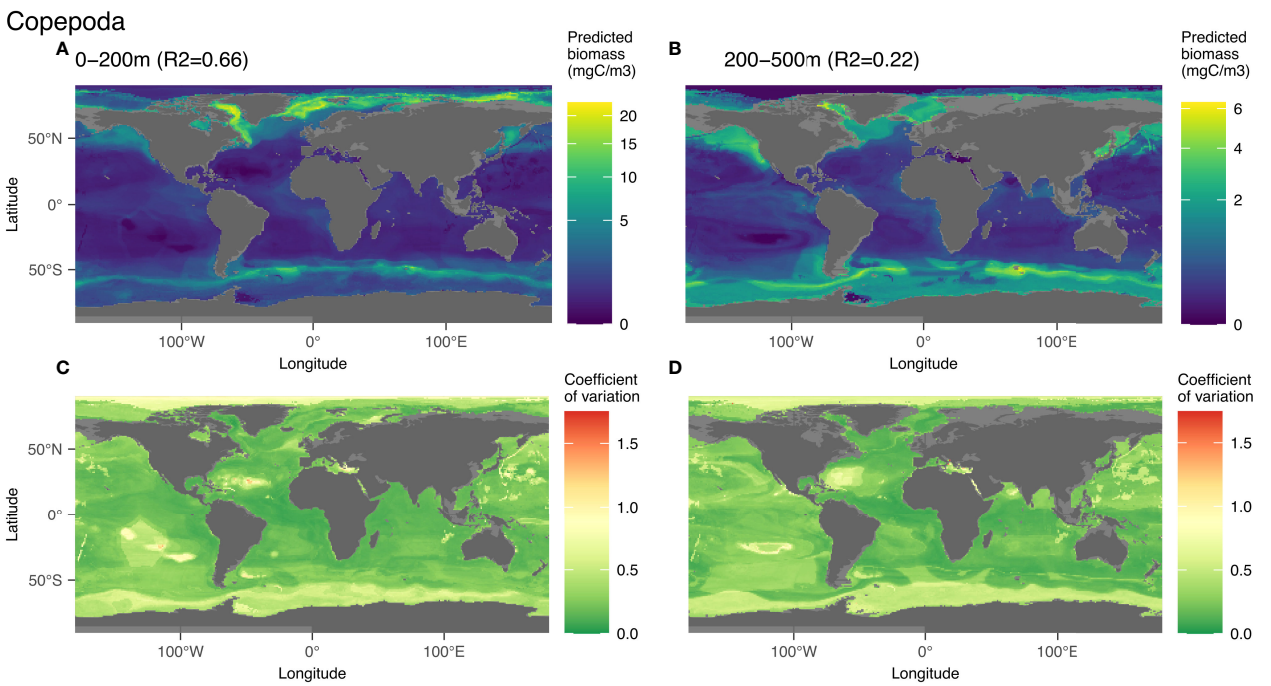


FIGURE 5 | Map of the mean biomass (color scale is log-transformed) of Copepoda as predicted by the model on 0-200 m **(A)** 200-500 m data **(B)** as well as the coefficient of variation for the 0-200 m model **(C)** and 200-500 m one **(D)**. The color scale for the coefficient of variation has the same range for **Figures 5**–**9**.

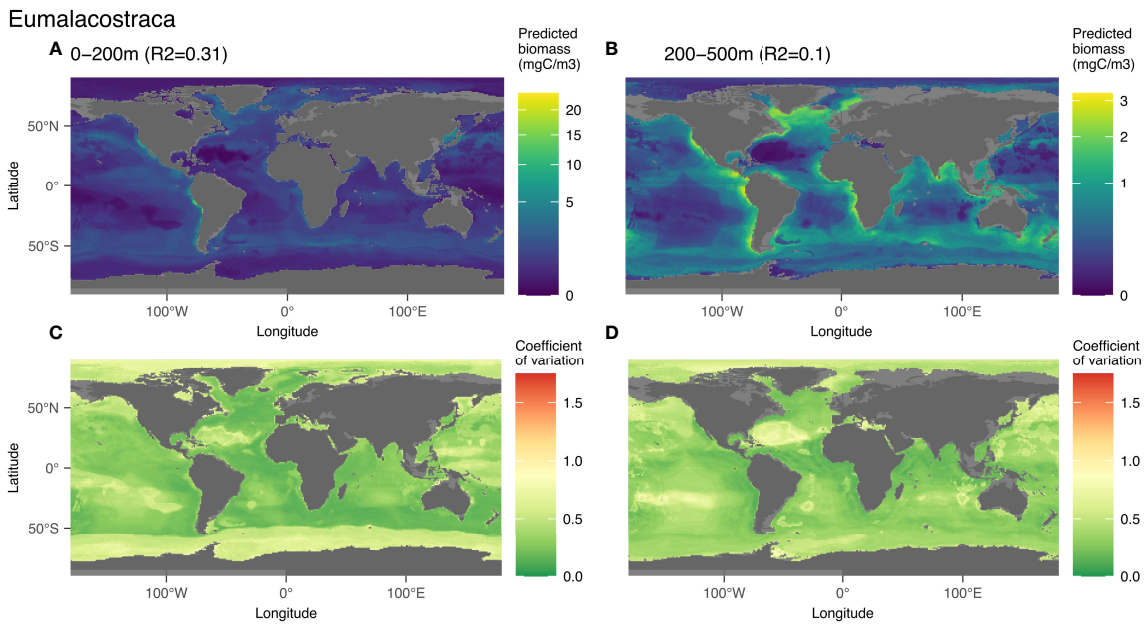


FIGURE 6 | Map of the mean biomass (color scale is log-transformed) of Copepoda as predicted by the model on 0-200 m **(A)** 200-500 m data **(B)** as well as the coefficient of variation for the 0-200 m model **(C)** and 200-500 m one **(D)**. The color scale for the coefficient of variation has the same range for **Figures 5**–**9**.

currents, especially in the Peruvian and Californian upwelling systems. Low biomass is predicted in high latitudes and in the oceanic gyres, especially in the North Atlantic. Similar patterns were predicted in the mesopelagic layer, but with lower biomass values. The model uncertainties are highest in the zones of low biomass (high latitudes and oceanic gyres).

3.3.3 Solitary Collodaria

Solitary Collodaria were predicted globally with an R^2 of 0.1 for the epi- and 0.09 for the mesopelagic layer, both with significant p-values (p-value < 0.05; **Figure 3**). In the epipelagic, the distribution of solitary Collodaria were mainly associated with low salinity (21%, between 35 and 37) and bathymetry (14%) (**Figures S10A, B**). In the mesopelagic, high abundances of this group were associated with distance to shore (18%) and high chlorophyll *a* (17%)(**Figures S10C, D**). In this layer, 65% of the biomass was predicted at less than 1,000 km from the coast. Solitary collodaria were mainly located between 50°N and 50°S, in a rather diffuse manner (**Figure 7**) with maximum biomass predicted at the equator. In the intertropical region, the highest biomass was found in the epipelagic zones of productive areas such as the upwelling regions off the western coast of Africa (Cape Verde and Angola) and of the eastern boundary of the Pacific Ocean (Peru and California). The model also predicted high biomass in the Mediterranean Sea. The importance of the environmental variable "distance to coast" in the learning process created unusual patterns in the prediction map such as a hexagonal shape in the Pacific Ocean. North of 50°N and south of 50°S, environments that are typically characterized by water masses with low salinity (1st most structuring variable in the epipelagic) and high nitrate (4th variable), the predicted biomass was rather low especially in the epipelagic layer.

3.3.4 Phaeodaria

For this group, the worldwide epipelagic model was statistically significant (p-value < 0.05; **Figure 8**) with an R^2 of 0.27, but the mesopelagic model was not (p-value > 0.05; **Figure 3**). Therefore, only the 0-200 m layer is displayed (**Figure 8**). In this layer, Phaeodaria was one of the best predicted groups (**Figure 3**) especially in the upper 200m. The predicted epipelagic distribution of Phaeodaria is associated with low values of salinity (38%) followed by bathymetry (11%), surface chlorophyll *a* (10%), oxygen and temperature (8% each) (**Figures S11A, B**). This is visualised on the map of global prediction (**Figure 8A**) on which high biomass was mainly predicted in the Californian upwelling (characterized by low salinity, cold and coastal waters), with lower biomass north of the upwelling up to the Gulf of Alaska. High biomass values were also predicted in the Bay of Bengal and Adaman Sea. The coefficient of variation in zones of high biomass is very low, providing strong confidence in this pattern. The lowest predicted biomass for this group are found in oceanic gyres and high latitudes of the northern hemisphere.

3.3.5 Acantharea

The group Acantharea was predicted with low total biomass (**Figure 4**). This group was well predicted in the world model fitted with the epi- (R^2 = 0.26) and mesopelagic (R^2 = 0.63) layers (**Figure 9**). In the epipelagic layer, nitrate (18%), salinity (15%) and phosphate (12%) were the main driving variables (**Figures S12A, B**). In the mesopelagic layer, the link between biomass and environment (**Figure 9B**) was defined by the influence of several variables: silicate (19%), phosphate (12%) followed by chlorophyll *a* (12%) (**Figures S12C, D**). The highest epipelagic

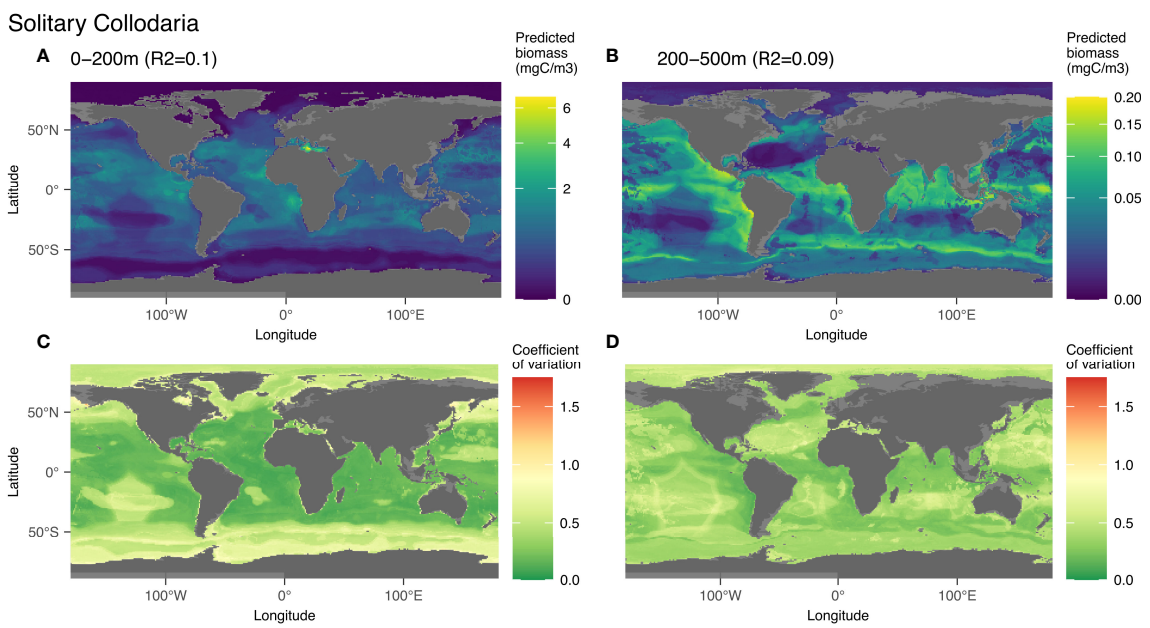


FIGURE 7 | Map of the mean biomass (color scale is log-transformed) of solitary Collodaria as predicted by themodel on 0-200 m **(A)** 200-500 m data **(B)** as well as the coefficient of variation for the 0-200 m model **(C)** and 200-500 m one **(D)**. The color scale for the coefficient of variation has the same range for **Figures 5**–**9**.

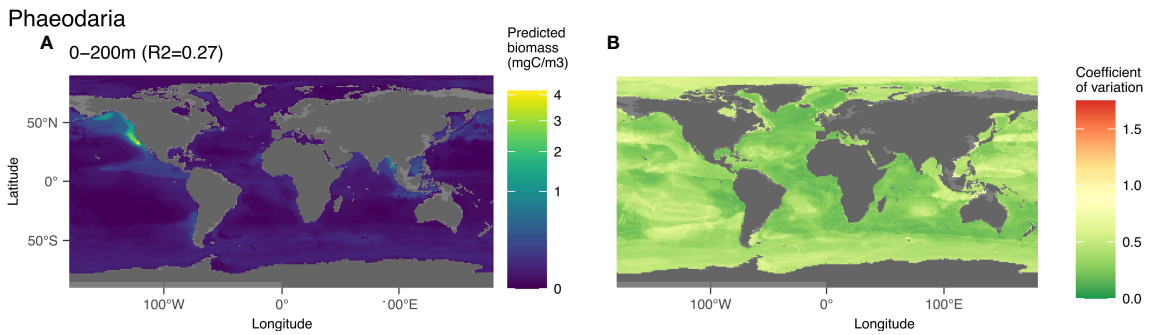


FIGURE 8 | Map of the mean biomass (color scale is log-transformed) of Phaeodaria as predicted by the model on 0-200 m **(A)**, as well as the coefficient of variation for the 0-200 m model **(B)**. In the map of predicted biomass, 12 cells in the California upwelling presented a value between 3 and 6 mgC m^{-3} and were represented here in yellow to observe the distribution of this group on a global scale. The color scale for the coefficient of variation has the same range for **Figures 5**–**9**.

biomass (**Figure 9A**) was predicted in the intertropical range, in productive areas such as the upwellings off the West coast of Africa (Cape Verde, Angola) and America (Peru and California). These high biomass patches are associated with a salinity around 35 as the 2nd most structuring variable, as well as with high nitrate and phosphate concentrations (respectively 1st and 3rd). Intermediate biomass values were predicted mostly between 50°N and 50°S in a diffuse way, except in the oceanic gyres where the predicted biomass was lowest. The largest uncertainty was present in the Southern and Artic Oceans, Bering Sea and Gulf of Alaska where low biomass values were predicted (**Figure 9C**). In the mesopelagic layer, biomass was predicted to be 16.7-times lower overall (**Figure 9B**), with highest values found in the Gulf of Alaska and the Bering Sea. Intermediate biomass values were predicted for the upwelling regions and the Southern Ocean. In this layer, the high biomass estimates correspond with low coefficient of variation values (**Figure 9D**).

3.4 *In Situ* Imaging Compared to Net Based Sampling

The latitudinal biomass distribution of Copepoda and Rhizaria obtained by combining the predictions of global models for the epi- and mesopelagic is shown in **Figure 10**. It is compared against data (interpolated on 0-500 m) from the Tara Oceans mission (Pesant et al., 2015; Soviadan et al., 2022) acquired using 300 μm multinet samples and ZooScan (Gorsky et al., 2010). To make the comparison meaningful, we only selected organisms in the ZooScan samples with an ESD >1 mm. For Copepoda, the values observed by the UVP5 and the nets reveal a similar latitudinal pattern between 70°N and 60°S. The trend computed on the output of the models shows lower biomass between 40°N and 40°S compared to Tara observations. For Rhizaria, the highest biomass was found in the UVP5 observations and models around the equator. Generally, almost no Rhizaria were observed in nets whereas they were consistently observed with the UVP5.

3.5 Global Zooplankton Biomass Distribution

The biomass integrated over 0-500 m was predicted to be maximal at around 60°N and 55°S, with values decreasing both north and south of these two latitudes (**Figure 11**). The lowest values of biomass were predicted north of 80°N and in the Weddell Sea as well as in the oceanic gyres (especially in the southern hemisphere). We also observed an increase of the predicted biomass around the equator. The highest biomass values were predicted between 50 and 80°N, in coastal waters of the Labrador Sea and Baffin Bay, as well as in the Greenland Sea. Relatively high biomass was predicted around these locations as well as in the Gulf of Alaska, Bering Sea and Sea of Okhotsk. A band of high biomass was predicted between 40 and 50°S, a region associated with the Arctic polar front.

Finally, by summing only the predictions that significantly correlated with observations, we can get to a first robust, conservative, global biomass estimate of zooplankton biomass based on UVP5 *in situ* imaging. As not all groups could be included in this computation, we refer to the following numbers as biomass *ad minima*. With that in mind, the zooplankton biomass estimated by the models was 0.229 PgC for the epipelagic, and 0.173 PgC for the mesopelagic. Thus, the estimated biomass for the upper 500m of the ocean is to 0.403 PgC.

4 DISCUSSION

4.1 Sensitivity of Model Prediction to Partitioning

In this study, we explored whether a partitioning approach would improve model performance through the use of different horizontal and vertical divisions of our dataset. The aim of using partitioned models was to test if we could model local taxa that otherwise would be mixed within the coarse taxonomic definition imposed by the dataset. The R^2 computed on the models' output show a high variability across groups, layers and regional combinations. Overall, when

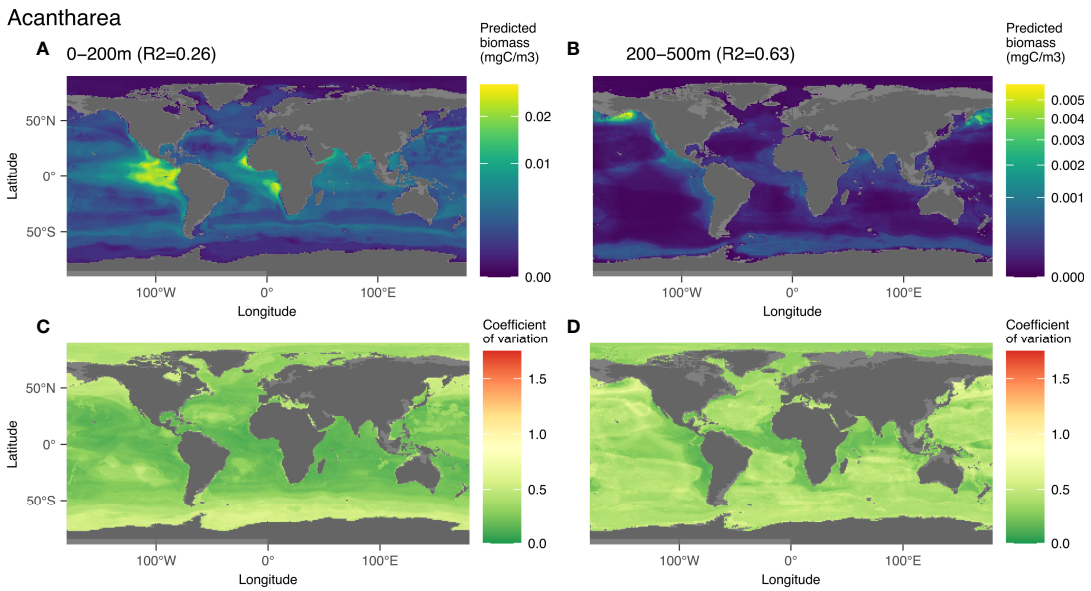


FIGURE 9 | Map of the mean biomass (color scale is log-transformed) of Acantharea as predicted by the model on 0-200 m **(A)**, 200-500 m data **(B)**, as well as the coefficient of variation for the 0-200 m model **(C)** and 200-500 m one **(D)**. The color scale for the coefficient of variation has the same range for **Figures 5**–**9**.

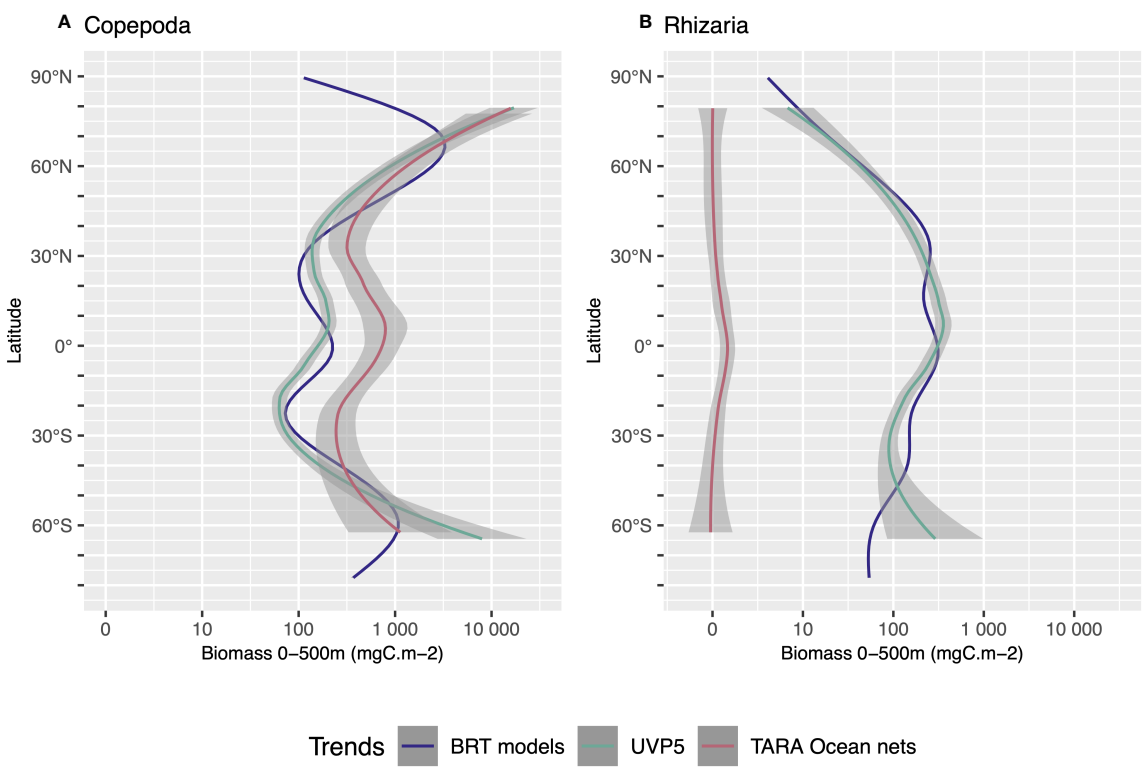


FIGURE 10 | Comparison of the latitudinal distribution of biomass mgC m^{-2}) integrated over 0-500m depth between our models' estimation and the results from the Tara Ocean multinet (300 mm mesh size), for Copepoda **(A)** and Rhizaria **(B)**. Trends were obtained by using Loess regression on: BRT models (blue line) using the global model outputs for Copepoda or Rhizaria (summed across 0-200 m and 200-500 m depth); UVP5 data (green line) using the biomass as seen by the UVP5 between 0-500m; TARA Ocean net data (red line) using the sampling points between 0-500m. The Shaded areas represent the 95% confidence interval of the Loess fit.

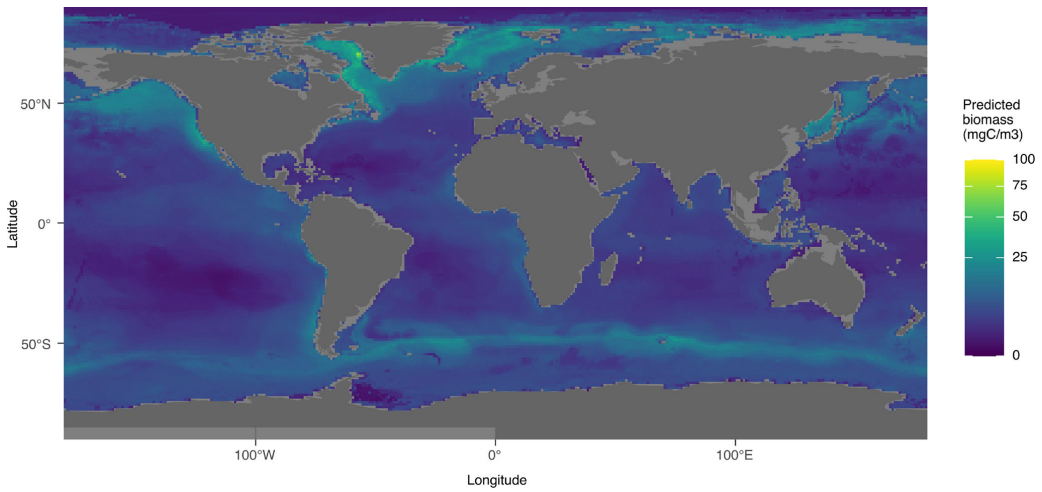


FIGURE 11 | Distribution map of the predicted minimum global biomass between 0 and 500m using taxa which obtained a significant result (p-value < 0.05) in Pearson test between the predicted biomass and the biomass calculated from UVP5 data.

comparing each partitioned model to the same zone in the global model, the global and the partitioned models had similar performance. The reduction in dataset size might be the explanation why in many cases global models perform better than the smaller partitioned models. The high latitude dataset contains 712 UVP5 profiles, the low latitude 2,837 and the world 3,549 data points. Another drawback of the partitioned models could be that some groups might have an environmental habitat associated with regions on both sides of the limits of the two models (here 40°N or 40°S). A vertical resolution that consists of two layers (0-200 and 200-500m) provided the best results (**Figure 3**) compared to a finer depth separation. The reduction of data per model with a finer depth layer resolution probably made it impossible for some models to learn the association between a group's biomass distribution and the associated habitat properties, either because the model could not learn this association or because the group was considered rare (< 500 images). If enough data are available, however, a finer vertical model might perform better, because it better delimits the vertical habitat structure. This seems to be the case for the Phaeodaria for which models with 100 m resolution obtained higher R^2 results, especially for those between 0 and 300 m depth.

4.2 Group-Wise Contribution to Global Zooplankton Biomass

Globally, in the 1.02 - 50 mm size range, we observed up to four zooplankton groups dominating each region and layer (**Figure 4**), mainly Crustacea (Copepoda, Eumalacostraca, other Crustacea) and Rhizaria (solitary Collodaria, Phaeodaria, Foraminifera). The dominance by copepods was expected: they are known to be a central trophic link in marine ecosystems (Steinberg and Landry, 2017) and their dominance was already shown in several studies (Turner, 2004; Forest et al., 2012; Dai et al., 2016). Rhizaria were also presented as substantial participants in the global zooplankton biomass by Biard et al. (2016) with Phaeodaria and Collodaria being the most important contributors to rhizarian biomass. In addition, Rhizaria were previously shown to play an important role in the biological carbon pump by intercepting (Stukel et al., 2018; Stukel et al., 2019) but also generating particle flux (Lampitt et al., 2009). In contrast, gelatinous predators such as Chaetognatha and other Cnidaria (other Cnidaria, other Hydrozoa, Siphonophorae) can be well predicted but their predicted biomass is low. This might be due to different reasons, ranging from their low carbon content (McConville et al., 2016), their size range which can exceed the specific range of the UVP5 (1.02 - 50 mm), their lower abundance reducing the probability of observation in the rather small volume of the UVP5 and the reduced capacity of the UVP5 to image them due to their transparency. Other instruments, such as the pelagic *in situ* observation system (PELAGIOS, Hoving et al. (2019)), the Zooglider (Ohman, 2019) or the *In Situ* Ichthyoplankton Imaging System (ISIIS, Cowen and Guigand (2008)) might be more adapted to study these organisms, thanks to their larger sampling volumes or different image approach.

4.3 Distribution Patterns and Occupied Habitats

4.3.1 Copepoda

Copepoda biomass was predicted to be highest in high latitudes in both epi- and mesopelagic layers of the global models. The lowest values were predicted at the gyres and an increase of biomass was observed centered at the equator. In the global models, temperature always appeared within the top three environmental factors explaining the distribution of copepods (except for 0-100 m model where it appeared 4th), which is in agreement with previous work suggesting that surface temperature and thermal tolerance of marine ectotherms, including copepods, are important constraints for their distribution and abundance (Beaugrand et al., 2009; Sunday

et al., 2012). We also predict significant Copepoda biomass centered at 50°S in the Southern Ocean, at the location of the strongest horizontal gradient of temperature within the epipelagic layer. This geographic pattern is in agreement with earlier observations of high Copepoda occurrence along the Polar front (Pinkerton et al., 2020). Hence, despite a low number of UVP5 profiles in this latitudinal band, the model is able to retrieve this fundamental pattern. Higher values of the coefficient of variation (**Figure 5C**) are found in the Arctic Ocean, as well as south of 60°S. More data from these regions could help to further reduce the uncertainty of our models.

4.3.2 Eumalacostraca

The distribution of the predicted Eumalacostraca biomass showed high values in coastal areas mainly on the eastern boundary currents of the Atlantic and Pacific Oceans and low values at high latitudes and at the locations of the oceanic gyres. Due to the low image resolution, a finer taxonomic resolution than Eumalacostraca (mostly euphausiids, decapods and amphipods) is not possible for UVP5 vignettes. Euphausiids are well known for their ability to escape standard oceanographic plankton nets (Brinton, 1967; Wiebe et al., 1982; Sameoto et al., 1993) and even low noise gliders (Guihen et al. 2022). This behavior might also be dependent on the species and stage development while the UVP5 mostly detects small Eumalacostraca (≤ 50 mm) for which taxonomic identification is not possible. Nevertheless, as Euphausiids are the second most abundant crustacean taxon after copepods (Castellanos et al., 2009), they may compose a large fraction of the biomass in this group. They are described as widely distributed in high numbers in the world ocean between 0-300 m with the exception of the eastern Canadian Arctic and the Arctic Ocean (Castellanos et al., 2009). This is consistent with our predictions of higher biomass in the epipelagic zone (0.058 PgC) compared to the mesopelagic (0.049 PgC), and low values predicted for the Arctic Ocean. The high Eumalacostraca biomass predicted in the North Atlantic also consistent with other observations that reported high abundances of krill in this region (Edwards et al., 2021). *Euphausia superba* and *Euphausia mucronata* have been respectively described as keystone species of the Antarctic and the Humboldt Current System (Antezana, 2010). The comparatively low values of biomass predicted in the Antarctic in the epipelagic layer (**Figure 6A**) might be too low, as *Euphausia superba* is known to show a patchy distribution (Siegel, 2005; Siegel, 2016). Since we only have very few samples from the Antarctic Ocean, we probably under-sampled this region and specifically krill. The high coefficient of variation in this region seems to reflect this problem. Overall, our observations and models likely underestimate the abundances of Euphausiids and of Eumalacostraca, due to their escape behaviors, the comparatively small sampling volume of the system and the low sample size in the Southern Ocean.

4.3.3 Solitary Collodaria

Global models in epi- and mesopelagic layers predicted a widespread distribution of solitary Collodarians between 50°N and 50°S, from oligotrophic to eutrophic zones. Their distribution can be explained by the selective advantage of their mixotrophy, since all collodarian species live in symbiosis with photosynthetic microalgae (Suzuki and Not, 2015; Biard et al., 2016). Consistently with the models' prediction of solitary Collodaria as the third most important group in terms of global biomass in 0-200 m, it has been shown by Biard et al. (2016) that Collodaria contribute most to the biomass of the Rhizaria between 0-100 m.

4.3.4 Phaeodaria

The distribution of Phaeodaria shows a latitudinal pattern with three peaks in biomass, at 50°N (with high biomass values at the level of the subarctic gyres), at 5°N and at 60°S. These three peaks were not observed by Biard et al. (2016). The highest values being predicted in the subarctic gyre are consistent with Steinberg et al. (2008) who estimated their mean biomass there as 5.5% (range 2.7–13%) of the metazoan biomass sampled using a MOCNESS (Wiebe et al., 1985). The distribution of this group in the epipelagic (high biomass in coastal regions especially around the Californian upwelling and low biomass in the gyres conditions) could be related to food availability which might not be abundant enough in the open ocean. In the models' output, this group only accounted for to ~ 1.2% of the global biomass in the epipelagic. This is consistent with previous work describing these organisms as being distributed in water below 150-200 m (Stemmann et al., 2008; Nakamura and Suzuki, 2015; Boltovskoy et al., 2017; Biard and Ohman, 2020). The high (R^2 = 0.50) and low latitude (R^2 = 0.39) models for the mesopelagic layer reveal similar patterns as the ones shown for the epipelagic layer in **Figure 8**. This pattern of high biomass predicted in the North Pacific can be put in perspective with a previous study (Ikenoue et al., 2019) which highlighted Phaeodaria in the Western North Pacific as one of the major carriers of carbon in the twilight zone (200-1000 m (Buesseler and Boyd, 2009)), with an organic carbon standing stock reaching its highest value at depths between 200-500 m. A maximum in abundance of Phaeodaria was observed in the lower epipelagic or mesopelagic zone in the Sea of Japan by Nakamura et al. (2013) as well as in the Antarctic beneath the sea ice with similar abundances as the North Atlantic and Pacific (Morley and Stepien, 1984). In the regional mesopelagic predictions, the mean biomass in the Sea of Japan is not particularly high, but it reached higher values in the Southern Ocean.

4.3.5 Acantharea

Here, we present results on large Acantharea only, but it should be kept in mind that most species are smaller than 600 μm (Biard et al., 2016). Most Acantharea species are associated with symbiotic algae (Michaels, 1991) which could explain the rapid observed biomass decline with depth. Indeed, the biomass predicted is 16.7-times lower in the mesopelagic ($1.36\ 10^{-5}$ PgC) compared to the epipelagic layer ($2.27\ 10^{-4}$ PgC). These mixotrophs are present throughout the world oceans (Suzuki and Not, 2015) and commonly distributed in intertropical latitudes (Bottazzi and Andreoli, 1982) mostly in the surface with an abundance rapidly declining below 20-50 m depth (Michaels, 1988). The model confirmed this biomass

diminution in the epi- and mesopelagic layers (**Figure 9**). We also observed latitudinal patterns with the highest biomass in intertropical areas consistent with these previous studies. The highest biomass of Acantharea predicted by the mesopelagic global model in the Gulf of Alaska coincides with a large number of organisms imaged by the UVP5. This is surprising knowing the above described distribution patterns. More observations from this region are required to clarify whether this was a temporally limited occurrence or whether it represents a region of permanent abundance maxima. The predicted biomass in Antarctic waters in this depth layer is also surprising. Acantharea are marine planktonic unicellular eukaryotes in the Rhizaria group and produce a mineral skeleton made of strontium sulfate (Michaels, 1991; Decelle and Not, 2015). The surprisingly high abundance at high latitudes might be important for studies done on the strontium biogeochemical cycle (Bernstein et al., 1987; Decelle et al., 2013).

4.4 Comparison Between Net Sampling and *In Situ* Imaging

The integrated global predicted biomass is dominated by Copepoda (35.7%), Eumalacostraca (26.6%) and Rhizaria (16.4%). Because of their important contribution to the predicted global biomass, the distribution map of total biomass *ad minima* (**Figure 11**) reflects in part the major distribution patterns of these three groups: polar waters are dominated by Copepoda and intertropical waters are dominated by mixotrophic Rhizaria. Eumalacostraca follows the predicted distribution of zooplankton with 3 peaks of biomass at 60°N (55°N for zooplankton), at the equator and at 45°S (55°S for zooplankton). The comparison of the models' output with data from the Tara Ocean expedition, obtained with a 300 μm mesh size multinet (Pesant et al., 2015; Soviadan et al., 2022) shows a good agreement for the latitudinal patterns of Copepod biomass. Net data is estimated to be higher than biomass estimated from UVP5 data in the intertropical latitude range for this group. Results in the high latitudes regions with strong seasonality and sea ice cover should be taken with caution as no data was available in the UVP5 dataset in winter for these latitudes. For Rhizaria, we observe that at most locations the biomass estimated by the nets is zero, while the UVP5 images suggest a considerable biomass in this group (**Figure 10**). In the TARA Ocean multinet samples, only Acantharea, Foraminifera and Phaeodaria are sometimes detected, while Collodaria are consistently absent from these samples. Indeed, Collodaria and Acantharea are poorly sampled by nets and are not well preserved in plankton samples fixed with regular fixatives such as formaldehyde (Suzuki and Not, 2015). Yet, solitary Collodaria are predicted as the 3rd most important group in terms of biomass in the upper 200 m of the global model. Our results show that *in situ* imaging is far more suitable for the study of this group and all other fragile plankton groups. As described above, several important zooplankton groups are generally well modeled, allowing us to combine the taxon-specific models to yield a global estimate of zooplankton biomass in the 1.02 to 50 mm size range. Previous studies (**Table 2**) have computed such global zooplankton biomass obtained largely (Hatton et al., 2021) or completely (Moriarty et al., 2012; Moriarty and O'Brien, 2013; Buitenhuis et al., 2013) from net collected organisms. These studies also used a proportionality method for estimating the global biomass presented in **Table 2** by multiplying the median value of biomass with the surface of the ocean and the studied depth. Our predictions are within the same order of magnitude — but at the lower limit — of these compilations if one combines their meso- and macrozooplankton biomass estimates. We refrain from a more detailed comparison due to the difference in size studied (here 1.02 - 50 mm ESD — equivalent to 765 μm to 37.5 mm meshsize according to Nichols and Thompson (1991)'s 3/4 law of mesh selection — compared to ≥ 200 μm for the cited meso- and macrozooplankton studies), sampling methods and depth covered (Buitenhuis et al., 2013). Contrary to the complementary use of nets and Zooscan, such as with the TARA dataset, these previous studies are based on data obtained through methods which do not allow to split the organisms based on fixed criteria (size, area of the organism or taxonomy). One would expect a large contribution to biomass in the 200 to 765 μm mesh size range (Gallienne, 2001; Hwang et al., 2007).

4.5 Global Zooplankton Biomass Distribution

The distribution of the global integrated biomass (0-500 m) *ad minima* follows the patterns described by Ikeda (1985), Moriarty et al. (2012) and Hatton et al. (2021) which correspond to a latitudinal distribution of the biomass with high values north of 55°N and south of 55°S. Relatively higher values of biomass are predicted around the equator (15°N-15°S). The benefit of our work and of compiled datasets such as the ones used in Moriarty et al. (2012); Moriarty and O'Brien (2013), Buitenhuis et al. (2013) and Hatton et al. (2021) is that they bring together numerous single transects and allow to have an integrated view

TABLE 2 | Comparison of global biomass estimates in the literature.

Study	Size range (mesh size)	Depth	Global estimates (PgC)
Moriarty et al., 2012	≥2 mm	0-350 m	0.02
Moriarty and O'Brien, 2013	≥200 µm	0-200	0.19
Buitenhuis et al., 2013	≥200 µm	Integrated	0.33-0.59
Buitenhuis et al., 2013	≥2 mm	0-500 m	0.22-1.52
Hatton et al., 2021	≥200 µm	0-200 m	0.53-31.57
Hatton et al., 2021	≥2 mm	0-200 m	0.02-2.64
This study	≥765 µm - 37.5 mm	0-200 m	0.229

Please note that we have converted the size range we cover with the UVP5 (1.02-50 mm ESD) to meshsize using the empirical Nichols and Thompson (1991)'s 3/4 law of mesh selection.

of global zooplankton distribution. The results depicted in **Figure 11** in the Southern Ocean are consistent with a recent study done with BRTs (Pinkerton et al., 2020) showing that the highest environmental suitability for zooplankton was located between the Subantarctic Front and the southern limit of the Antarctic Circumpolar Current with a lower suitability north and south of this band. The spatial distribution of plankton biomass thus shows the importance of oceanographic hydrodynamics leading to oligotrophy in central gyres and mesotrophy in areas of high latitudes and equatorial and coastal upwellings. Zooplankton plays a crucial role in fisheries e.g. in the Humboldt Current System which harbors the largest fishery in the world and most economically important fish species, supported by the upwelling of Peru (Chavez et al., 2008). Peruvian anchovies and sardines obtain most of their energy from zooplankton (van der Lingen et al., 2009).

4.6 Conclusions and Outlook

In summary, our results show, for the first time, that spatial patterns and global biomass of key zooplankton groups can be calculated using a machine learning method (BRT) to extrapolate individual zooplankton biomass estimates from sparse UVP5 observation. They also highlight the important contribution of Rhizaria (predicted mainly in the intertropical range) and Copepoda (predicted mainly in high latitudes) to the global estimate of zooplankton biomass. Within the size range covered, Copepoda contributes 35.7%, Eumalacostraca 26.6% and Rhizaria 16.4% to global zooplankton biomass. This suggests that it is especially crucial to extend work on the fragile Rhizaria, which are comparatively little studied. As a biogeographical study, our aim was not to represent proximal mechanisms that drive the distribution of zooplankton, or to describe seasonal or transient (e.g. mesoscale) features, but rather to represent the global distribution patterns of biomass according to general properties of the water masses. This method worked well in general as seen in **Figure 3** for at least 3 of the combinations of regions and depths. It made it possible to model 19 groups of zooplankton and obtain corresponding maps with the relative importance of the environmental variables used for the model. The WOA climatologies used in this study compile data of salinity and temperature (2005-2017) and other variables (1955-2017). The temporal coverage of the latter being much coarser, we hope to use more constrained nutrient datasets in our future work as they become available.

The zooplankton biomass predictions based on UVP5 datasets presented here are important for global biogeochemical modeling of pelagic ecosystems because they usually lack zooplankton observations to constrain their development (Stemmann and Boss, 2012; Buitenhuis et al., 2013; Séférian et al., 2020). A current trend is to add a more realistic representation of plankton in ecosystem models to better predict future ecosystem states and ocean conditions and to inform sustainable management strategies for climate mitigation at global scale (Séférian et al., 2020). The UVP5, the newly developed UVP6 (Picheral et al., 2021) and other commercialized *in situ* systems, provided that they are inter-calibrated (Lombard et al., 2019), will continue to be used in the foreseeable future, increasing data availability. Still, the bottleneck lies in the classification of the massive amount of images which still require human validation, but new algorithms to recognise plankton types and traits are expected (Irisson et al., 2022). The further anticipated expansion of image datasets will enable the quantitative assessment of rare groups that were not well predicted here. In addition, the deployment of the UVP6 on autonomous platforms will also help to sample certain areas that are difficult to access at certain times of the year such as polar regions in winter. The large dataset used in this study spans 10 years of data collection and can be compared to the COPEPOD database collected since about 1960. The possibilities given by imaging systems could hence help to reach a useful amount of data in a much smaller time frame. It would be interesting to use other imaging system's data sets such as the ones presented by Lombard et al. (2019) to reconstruct the wider size spectrum of these groups in terms of biomass. To have a better understanding of the vertical habitat of zooplanktonic groups, we highly recommend that UVP5 and 6 profiles should be done to at least 1,000 m when the bathymetry allows it. Long term inter annual data acquisition is also highly recommended. This will enable us to monitor global zooplankton biomass changes at pace with the speed of global change.

DATA AVAILABILITY STATEMENT

The inputs and outputs of the world models for 0-200 m and 200-500 m were uploaded to the GitHub repository: https://github.com/dlaetitia/Global_zooplankton_biomass_distribution.git. The code used for the models and the post treatment of their outputs was also made available in the same GitHub repository. The dataset of environmental data from World Ocean Atlas is available at https://www.ncei.noaa.gov/products/world-ocean-atlas. The surface chlorophyll a data is available on the Copernicus website at https://resources.marine.copernicus.eu/productdetail/ OCEANCOLOUR_GLO_CHL_L4_REP_OBSERVATIONS_009_082/

AUTHOR CONTRIBUTIONS

LD, RK, LS, JO-I and TP developed the study's concept; RK, LS, LD, FC, FL, MB, TB, LC, LG, HH, LK-B, AD, MP, AR, AW, contributed to data acquisition; RK, LS, LD, JO-I, TP contributed substantially to the data analysis; LD led the code development with major assistance by JO and guidance by TP and RK; LD created all figures and drafted the manuscript; All authors contributed substantially to drafting the manuscript; All authors approved the final submitted manuscript.

FUNDING

LD, RK and LS received support by the European Union project TRIATLAS (European Union Horizon 2020 program, grant agreement 817578) and the Sorbonne Université through the Ecole doctorale 129. LD was supported in the beginning of this project by the Laboratoire d'Océanographie de Villefranche-sur-

mer (LOV). LS was supported in the initial phase of the development by the CNRS/Sorbonne University Chair VISION. RK furthermore acknowledges support *via* a Make Our Planet Great Again grant from the French National Research Agency (ANR) within the Programme d'Investissements d'Avenir (reference ANR-19-MPGA-0012). AR was funded by the PACES II (Polar Regions and Coasts in a Changing Earth System) program of the Helmholtz Association and the INSPIRES program of the Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research. Some of the observations used in this study were funded by NASA grants #NNX15AE67G.

ACKNOWLEDGMENTS

The sampling effort on which this paper is based was enabled by the dedicated cruise leaders and participants who helped in the creation of the UVP5 dataset. We thank them as well as all the people who have participated in the classification of the huge amount of UVP5 images. We are grateful for the ship time provided by the respective institutions and programs. Furthermore, we would like to thank our colleagues from the Laboratoire d'Oceanographie de Villefranche (LOV) as well as Marina Levy and Fabio Benedetti for their support and precious guidance. We also thank Yawouvi Dodji Soviadan for his help with the Tara Oceans mission data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.894372/full#supplementary-material

REFERENCES

Amante, C., and Eakins, B. W. (2009). *ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. NOAA Technical Memorandum NESDIS NGDC-24*. National Geophysical Data Center, NOAA. doi: 10.7289/V5C8276M [2021.01.19].

Antezana, T. (2010). Euphausia Mucronata: A Keystone Herbivore and Prey of the Humboldt Current System. *Deep. Sea. Res. Part II.: Top. Stud. Oceanogr.* 57, 652–662. doi: 10.1016/j.dsr2.2009.10.014

Beaugrand, G., Luczak, C., and Edwards, M. (2009). Rapid Biogeographical Plankton Shifts in the North Atlantic Ocean. *Global Change Biol.* 15, 1790–1803. doi: 10.1111/j.1365-2486.2009.01848.x

Beers, J. R., and Stewart, G. L. (1970). The Preservation of Acantharians in Fixed Plankton Samples. *Limnol. Oceanogr.* 15, 825–827. doi: 10.4319/lo.1970.15.5.0825

Bernstein, R. E., Betzer, P. R., Feely, R. A., Byrne, R. H., Lamb, M. F., and Michaels, A. F. (1987). Acantharian Fluxes and Strontium to Chlorinity Ratios in the North Pacific Ocean. *Science* 237, 1490–1494. doi: 10.1126/science.237.4821.1490

Biard, T., and Ohman, M. D. (2020). Vertical Niche Definition of Test-Bearing Protists (Rhizaria) Into the Twilight Zone Revealed by *in Situ* Imaging. *Limnol. Oceanogr.* 65, 2583–2602. doi: 10.1002/lno.11472

Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Hauss, H., et al. (2016). *In Situ* Imaging Reveals the Biomass of Giant Protists in the Global Ocean. *Nature* 532, 504–507. doi: 10.1038/nature17652

Boltovskoy, D., Anderson, O. R., and Correa, N. M. (2017). "Radiolaria and Phaeodaria," in *Handbook of the Protists*. Eds. J. M. Archibald, A. G. Simpson and C. H. Slamovits (Cham: Springer International Publishing), 731–763. doi: 10.1007/978-3-319-28149-0_19

Bottazzi, E. M., and Andreoli, M. G. (1982). Distribution of adult and juvenile Acantharia (Protozoa Sarcodina) in the Atlantic Ocean. *Journal of Plankton Research*, 4 (4), 757–777. doi: 10.1093/plankt/4.4.757

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees* (Wadsworth & Brooks. Cole Statistics/Probability Series).

Brinton, E. (1967). Vertical Migration and Avoidance Capability of Euphausiids in the California Current. *Limnol. Oceanogr.* 12, 451–483. doi: 10.4319/lo.1967.12.3.0451._eprint:https://onlinelibrary.wiley.com/doi/pdf/10.4319/lo.1967.12.3.0451

Brodeur, R., Perry, I., Boldt, J., Flostrand, L., Galbraith, M., King, J., et al. (2018). *An Unusual Gelatinous Plankton Event in the NE Pacific: The Great Pyrosome Bloom of 2017*, Vol. 26. 6. Available at: https://www.researchgate.net/publication/322665037_An_unusual_gelatinous_plankton_event_in_the_NE_Pacific_The_Great_Pyrosome_Bloom_of_2017

Buesseler, K. O., and Boyd, P. W. (2009). Shedding Light on Processes That Control Particle Export and Flux Attenuation in the Twilight Zone of the Open Ocean. *Limnol. Oceanogr.* 54, 1210–1232. doi: 10.4319/lo.2009.54.4.1210

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., et al. (2013). MAREDAT: Towards a World Atlas of MARine Ecosystem DATa. *Earth Syst. Sci. Data* 5, 227–239. doi: 10.5194/essd-5-227-2013

Castellanos, I., Suárez-Morales, E., and Álvaro Morales-Ramírez, Ã. (2009). "Euphausiids," in *Marine Biodiversity of Costa Rica, Central America*, vol. 86. Eds. I. S. Wehrtmann and J. Cortés (Dordrecht: Springer Netherlands), 199–207. doi: 10.1007/978-1-4020-8278-818

Chavez, F. P., Bertrand, A., Guevara-Carrasco, R., Soler, P., and Csirke, J. (2008). The Northern Humboldt Current System: Brief History, Present Status and a View Towards the Future. *Prog. Oceanogr.* 79, 95–105. doi: 10.1016/j.pocean.2008.10.012

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., et al. (2021). *Xgboost: Extreme Gradient Boosting*.

Chen, B., Liu, H., Xiao, W., Wang, L., and Huang, B. (2020). A Machine-Learning Approach to Modeling Picophytoplankton Abundances in the South China Sea. *Prog. Oceanogr.* 189, 102456. doi: 10.1016/j.pocean.2020.102456

Christiansen, S., Hoving, H., Schütte, F., Hauss, H., Karstensen, J., Körtzinger, A., et al. (2018). Particulate Matter Flux Interception in Oceanic Mesoscale Eddies by the Polychaete Poeobius Sp. *Limnol. Oceanogr.* 63, 2093–2109. doi: 10.1002/lno.10926

Cowen, R. K., and Guigand, C. M. (2008). *In Situ* Ichthyoplankton Imaging System (*I* SIIS): System Design and Preliminary Results: *In Situ* Ichthyoplankton Imaging System. *Limnol. Oceanogr.: Methods* 6, 126–132. doi: 10.4319/lom.2008.6.126

Culverhouse, P. F. (2007). Human and Machine Factors in Algae Monitoring Performance. *Ecol. Inf.* 2, 361–366. doi: 10.1016/j.ecoinf.2007.07.001

Culverhouse, P. F., Macleod, N., Williams, R., Benfield, M. C., Lopes, R. M., and Picheral, M. (2014). An Empirical Assessment of the Consistency of Taxonomic Identifications. *Mar. Biol. Res.* 10, 73–84. doi: 10.1080/17451000.2013.810762

Dai, L., Li, C., Yang, G., and Sun, X. (2016). Zooplankton Abundance, Biovolume and Size Spectra at Western Boundary Currents in the Subtropical North Pacific During Winter 2012. *J. Mar. Syst.* 155, 73–83. doi: 10.1016/j.jmarsys.2015.11.004

Davis, C. S., Gallager, S. M., and Solow, A. R. (1992). Microaggregations of Oceanic Plankton Observed by Towed Video Microscopy. *Science* 257, 230–232. doi: 10.1126/science.257.5067.230

Davis, C. S., Thwaites, F. T., Gallager, S. M., and Hu, Q. (2005). A Three-Axis Fast-Tow Digital Video Plankton Recorder for Rapid Surveys of Plankton Taxa and Hydrography: New Video Plankton Recorder. *Limnol. Oceanogr.: Methods* 3, 59–74. doi: 10.4319/lom.2005.3.59

De'ath, G. (2007). Boosted Trees for Ecological Modeling and Prediction. *Ecology* 88, 243–251. doi: 10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2

Decelle, J., Martin, P., Paborstava, K., Pond, D. W., Tarling, G., Mahé, F., et al. (2013). Diversity, Ecology and Biogeochemistry of Cyst-Forming Acantharia (Radiolaria) in the Oceans. *PLoS One* 8, e53598. doi: 10.1371/journal.pone.0053598

Decelle, J., and Not, F. (2015). "Acantharia," in *eLS, 1 edn* (John Wiley & Sons, Ltd (Wiley), 1–10. doi: 10.1002/9780470015902.a0002102.pub2

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* 348, 1261605.5 – 126160511. doi: 10.1126/science.1261605

Edwards, M., Hélaouët, P., Goberville, E., Lindley, A., Tarling, G. A., Burrows, M. T., et al. (2021). North Atlantic Warming Over Six Decades Drives Decreases in Krill Abundance With No Associated Range Shift. *Commun. Biol.* 4, 644. doi: 10.1038/s42003-021-02159-1

Elith, J., and Graham, C. H. (2009). Do They? How do They? WHY do They Differ? On Finding Reasons for Differing Performances of Species Distribution Models. *Ecography* 32, 66–77. doi: 10.1111/j.1600-0587.2008.05505.x

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., et al. (2006). Novel Methods Improve Prediction of Species' Distributions From Occurrence Data. *Ecography* 29, 129–151. doi: 10.1111/j.2006.0906-7590.04596.x

Elith, J., and Leathwick, J. R. (2009). Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annu. Rev. Ecol. Evol. Syst.* 40, 677–697. doi: 10.1146/annurev.ecolsys.110308.120159

Elith, J., Leathwick, J. R., and Hastie, T. (2008). A Working Guide to Boosted Regression Trees. *J. Anim. Ecol.* 77, 802–813. doi: 10.1111/j.1365-2656.2008.01390.x

Faillettaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K., and Irisson, J.-O. (2016). Imperfect Automatic Image Classification Successfully Describes Plankton Distribution Patterns. *Methods Oceanogr.* 15-16, 60–77. doi: 10.1016/j.mio.2016.04.003

Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., et al. (2012). Size Distribution of Particles and Zooplankton Across the Shelf-Basin System in Southeast Beaufort Sea: Combined Results From an Underwater Vision Profiler and Vertical Net Tows. *Biogeosciences* 9, 1301–1320. doi: 10.5194/bg-9-1301-2012

Friedman, J. H. (2002). Stochastic Gradient Boosting. *Comput. Stat Data Anal.* 38, 367–378. doi: 10.1016/S0167-9473(01)00065-2

Gallienne, C. P. (2001). Is Oithona the Most Important Copepod in the World's Oceans? *J. Plankt. Res.* 23, 1421–1432. doi: 10.1093/plankt/23.12.1421

Garcia, H. E., Boyer, T.P., Baranova, O. K., Locarnini, R. A., Mishonov, A.V., Grodsky, A., et al (2019). *World Ocean Atlas 2018: Product Documentation*. A. Mishonov, Technical Editor.

González Carman, V., Piola, A., O'Brien, T. D., Tormosov, D. D., and Acha, E. M. (2019). Circumpolar Frontal Systems as Potential Feeding Grounds of Southern Right Whales. *Prog. Oceanogr.* 176, 102123. doi: 10.1016/j.pocean.2019.102123

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital Zooplankton Image Analysis Using the ZooScan Integrated System. *J. Plankt. Res.* 32, 285–303. doi: 10.1093/plankt/fbp124

Griffith, A. W., and Gobler, C. J. (2020). Harmful Algal Blooms: A Climate Change Co-Stressor in Marine and Freshwater Ecosystems. *Harmf. Algae.* 91, 101590. doi: 10.1016/j.hal.2019.03.008

Grossmann, M. M., Gallager, S. M., and Mitarai, S. (2015). Continuous Monitoring of Near-Bottom Mesoplankton Communities in the East China Sea During a Series of Typhoons. *J. Oceanogr.* 71, 115–124. doi: 10.1007/s10872-014-0268-y

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., et al. (2009). Effects of Phytoplankton Community on Production, Size, and Export of Large Aggregates: A World-Ocean Analysis. *Limnol. Oceanogr.* 54, 1951–1963. doi: 10.4319/lo.2009.54.6.1951

Guihen, D., Brearley, J. A., and Fielding, S. (2022). Antarctic Krill Likely Avoid Underwater Gliders. *Deep. Sea. Res. Part I.: Oceanogr. Res. Pap.* 179, 103680. doi: 10.1016/j.dsr.2021.103680

Guisan, A., and Zimmermann, N. E. (2000). Predictive Habitat Distribution Models in Ecology. *Ecol. Model.* 135, 147–186. doi: 10.1016/S0304-3800(00)00354-9

Hastie, T., and Tibshirani, R. (1986). General Additive Models 3, 297–318. doi: 10.1214/ss/1177013604

Hastie, T., Tibshirani, R., and Friedman, J. (2001). *The Elements of Statistical Learning. 2nd ed.* (Springer).

Hatton, I. A., Heneghan, R. F., Bar-On, Y. M., and Galbraith, E. D. (2021). *The Global Ocean Size-Spectrum From Bacteria to Whales. Preprint, Ecology*. doi: 10.1101/2021.04.03.438320

Hays, G., Richardson, A., and Robinson, C. (2005). Climate Change and Marine Plankton. *Trends Ecol. Evol.* 20, 337–344. doi: 10.1016/j.tree.2005.03.004

Hijmans, R. J. (2021). *Raster: Geographic Data Analysis and Modeling*.

Hoving, H.-J., Christiansen, S., Fabrizius, E., Hauss, H., Kiko, R., Linke, P., et al. (2019). The Pelagic *In Situ* Observation System (PELAGIOS) to Reveal Biodiversity, Behavior, and Ecology of Elusive Oceanic Fauna. *Ocean. Sci.* 15, 1327–1340. doi: 10.5194/os-15-1327-2019

Hu, C., Xu, J., Li, X., Shi, Z., and Li, R. (2021). Environmental Regulations on Bacterial Abundance in the South China Sea Inferred From Regression Models. *Sci. Tot. Environ.* 774, 146315. doi: 10.1016/j.scitotenv.2021.146315

Hwang, J.-S., Kumar, R., Dahms, H.-U., Tseng, L.-C., and Chen, Q.-C. (2007). Mesh Size Affects Abundance Estimates of Oithona Spp. (Copepoda, Cyclopoida). *Crustaceana* 80, 827–837. doi: 10.1163/156854007781363169

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., et al. (2019). Global Trends in Marine Plankton Diversity Across Kingdoms of Life. *Cell* 179, 1084–1097. doi: 10.1016/j.cell.2019.10.008

Ikeda, T. (1985). Metabolic Rates of Epipelagic Marine Zooplankton as a Function of Body Mass and Temperature. *Mar. Biol.* 85, 1–11. doi: 10.1007/BF00396409

Ikenoue, T., Kimoto, K., Okazaki, Y., Sato, M., Honda, M. C., Takahashi, K., et al. (2019). Phaeodaria: An Important Carrier of Particulate Organic Carbon in the Mesopelagic Twilight Zone of the North Pacific Ocean. *Global Biogeochem. Cycle.* 33, 1146–1160. doi: 10.1029/2019GB006258

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. (2022). Machine Learning for the Study of Plankton and Marine Snow From Images. *Annu. Rev. Mar. Sci.* 14, annurev–marine–041921–013023. doi: 10.1146/annurev-marine-041921-013023

Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., De Vargas, C., Raes, J., et al. (2011). A Holistic Approach to Marine Eco-Systems Biology. *PLoS Biology* 9, e1001177.1–5. doi: 10.1371/journal.pbio.1001177

Kiørboe, T. (2013). Zooplankton Body Composition. *Limnol. Oceanogr.* 58, 1843–1850. doi: 10.4319/lo.2013.58.5.1843

Lampitt, R. S., Salter, I., and Johns, D. (2009). Radiolaria: Major Exporters of Organic Carbon to the Deep Ocean. *Global Biogeochem. Cycle.* 23. doi: 10.1029/2008GB003221

Leathwick, J., Elith, J., Francis, M., Hastie, T., and Taylor, P. (2006). Variation in Demersal Fish Species Richness in the Oceans Surrounding New Zealand: An Analysis Using Boosted Regression Trees. *Mar. Ecol. Prog. Ser.* 321, 267–281. doi: 10.3354/meps321267

Lehodey, P., Alheit, J., Barange, M., Baumgartner, T., Beaugrand, G., Drinkwater, K., et al. (2006). Climate Variability, Fish, and Fisheries. *J. Climate* 19, 5009–5030. doi: 10.1175/JCLI3898.1

Le Quéré, C., Buitenhuis, E. T., Moriarty, R., Alvain, S., Aumont, O., Bopp, L., et al. (2016). Role of Zooplankton Dynamics for Southern Ocean Phytoplankton Biomass Andglobal Biogeochemical Cycles. *Biogeosciences* 13, 4111–4133. doi: 10.5194/bg-13-4111-2016

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). EnglishGlobally Consistent Quantitative Observations of Planktonic Ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Lombard, F., and Kiørboe, T. (2010). Marine Snow Originating From Appendicularian Houses: Age-Dependent Settling Characteristics. *Deep. Sea. Res. Part I.: Oceanogr. Res. Pap.* 57, 1304–1313. doi: 10.1016/j.dsr.2010.06.008

Lombard, F., Renaud, F., Sainsbury, C., Sciandra, A., and Gorsky, G. (2009). Appendicularian Ecophysiology I. *J. Mar. Syst.* 78, 606–616. doi: 10.1016/j.jmarsys.2009.01.004

Longhurst, A. (1995). Seasonal Cycles of Pelagic Production and Consumption. *Prog. Oceanogr.* 36, 77–167. doi: 10.1016/0079-6611(95)00015-1

Longhurst, A. R., and Glen Harrison, W. (1989). The Biological Pump: Profiles of Plankton Production and Consumption in the Upper Ocean. *Prog. Oceanogr.* 22, 47–123. doi: 10.1016/0079-6611(89)90010-4

Lucas, C. H., Jones, D. O. B., Hollyhead, C. J., Condon, R. H., Duarte, C. M., Graham, W. M., et al. (2014). Gelatinous Zooplankton Biomass in the Global

Oceans: Geographic Variation and Environmental Drivers: Global Gelatinous Biomass. *Global Ecol. Biogeogr.* 23, 701–714. doi: 10.1111/geb.12169

Mansour, J. S., Norlin, A., Llopis Monferrer, N., L'Helguen, S., and Not, F. (2021). Carbon and Nitrogen Content to Biovolume Relationships for Marine Protist of the Rhizaria Lineage (Radiolaria and Phaeodaria). *Limnol. Oceanogr.* 66, 1703–1717. doi: 10.1002/lno.11714

McConville, K., Atkinson, A., Fileman, E. S., Spicer, J. I., and Hirst, A. G. (2016). Disentangling the Counteracting Effects of Water Content and Carbon Mass on Zooplankton Growth. *J. Plankt. Res.* 39 (2), 246–256. doi: 10.1093/plankt/fbw094

McGinty, N., Barton, A. D., Record, N. R., Finkel, Z. V., Johns, D. G., Stock, C. A., et al. (2021). Anthropogenic Climate Change Impacts on Copepod Trait Biogeography. *Global Change Biol.* 27, 1431–1442. doi: 10.1111/gcb.15499

Michaels, A. F. (1988). Vertical Distribution and Abundance of Acantharia and Their Symbionts. *Mar. Biol.* 97, 559–569. doi: 10.1007/BF00391052

Michaels, A. F. (1991). Acantharian Abundance and Symbiont Productivity at the VERTEX Seasonal Station. *J. Plankt. Res.* 13, 399–418. doi: 10.1093/plankt/13.2.399

Moriarty, R., Buitenhuis, E. T., Le Quéré, C., and Gosselin, M.-P. (2012). Distribution of Known Macrozooplankton Abundance and Biomass in the Global Ocean. *Earth Syst. Sci. Data Discuss.* 5, 187–220. doi: 10.5194/essdd-5-187-2012

Moriarty, R., and O'Brien, T. D. (2013). Distribution of Mesozooplankton Biomass in the Global Ocean. *Earth Syst. Sci. Data* 5, 45–55. doi: 10.5194/essd-5-45-2013

Morley, J., and Stepien, J. (1984). Siliceous Microfauna in Waters Beneath Antarctic Sea Ice. *Mar. Ecol. Prog. Ser.* 19, 207–210. doi: 10.3354/meps019207

Nakamura, Y., Imai, I., Yamaguchi, A., Tuji, A., and Suzuki, N. (2013). Aulographis Japonica Sp. Nov. (Phaeodaria, Aulacanthida, Aulacanthidae), an Abundant Zooplankton in the Deep Sea of the Sea of Japan. *Plankt. Bentho. Res.* 8, 107–115. doi: 10.3800/pbr.8.107

Nakamura, Y., and Suzuki, N. (2015). "Phaeodaria: Diverse Marine Cercozoans of World-Wide Distribution," in *Marine Protists*. Eds. S. Ohtsuka, T. Suzaki, T. Horiguchi, N. Suzuki and F. Not (Tokyo: Springer Japan), 223–249. doi: 10.1007/978-4-431-55130-0_9

Nelder, J. A., and Wedderburn, R. W. M. (1972). *Journal of the Royal Statistical Society. Series A (General)* 135(3):370–384 (15 pages) doi: 10.2307/2344614

Nichols, J., and Thompson, A. (1991). Mesh Selection of Copepodite and Nauplius Stages of Four Calanoid Copepod Species. *J. Plankt. Res.* 13, 661–671. doi: 10.1093/plankt/13.3.661

Ohman, M. D. (2019). A Sea of Tentacles: Optically Discernible Traits Resolved From Planktonic Organisms in Situ. *ICES. J. Mar. Sci.* 76, 1959–1972. doi: 10.1093/icesjms/fsz184

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015). Open Science Resources for the Discovery and Analysis of Tara Oceans Data. *Sci. Data* 2, 150023. doi: 10.1038/sdata.2015.23

Pettitt-Wade, H., Pearce, T., Kuptana, D., Gallagher, C. P., Scharffenberg, K., Lea, E. V., et al. (2020). Inuit Observations of a Tunicata Bloom Unusual for the Amundsen Gulf, Western Canadian Arctic. *Arct. Sci.* 6, 340–351. doi: 10.1139/as-2020-0018

Phillips, S. J., Anderson, R. P., and Schapire, R. E. (2006). Maximum Entropy Modeling of Species Geographic Distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

Picheral, M., Catalano, C., Brousseau, D., Claustre, H., Coppola, L., Leymarie, E., et al. (2021). The Underwater Vision Profiler 6: An Imaging Sensor of Particle Size Spectra and Plankton, for Autonomous and Cabled Platforms. *Limnol. Oceanogr.: Methods* 20, 115–129. doi: 10.1002/lom3.10475. lom3.10475.

Picheral, M., Colin, S., and Irisson, J.-O. (2017). *EcoTaxa, a Tool for the Taxonomic Classification of Images*.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The Underwater Vision Profiler 5: An Advanced Instrument for High Spatial Resolution Studies of Particle Size Spectra and Zooplankton: Underwater Vision Profiler. *Limnol. Oceanogr.: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462

Pinkerton, M. H., Décima, M., Kitchener, J. A., Takahashi, K. T., Robinson, K. V., Stewart, R., et al. (2020). Zooplankton in the Southern Ocean From the Continuous Plankton Recorder: Distributions and Long-Term Change. *Deep. Sea. Res. Part I.: Oceanogr. Res. Pap.* 162, 103303. doi: 10.1016/j.dsr.2020.103303

R Core Team (2021). *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing).

Remsen, A., Hopkins, T. L., and Samson, S. (2004). What You See is Not What You Catch: A Comparison of Concurrently Collected Net, Optical Plankton Counter, and Shadowed Image Particle Profiling Evaluation Recorder Data From the Northeast Gulf of Mexico. *Deep. Sea. Res. Part I.: Oceanogr. Res. Pap.* 51, 129–151. doi: 10.1016/j.dsr.2003.09.008

Richardson, A. J., Bakun, A., Hays, G. C., and Gibbons, M. J. (2009). The Jellyfish Joyride: Causes, Consequences and Management Responses to a More Gelatinous Future. *Trends Ecol. Evol.* 24, 312–322. doi: 10.1016/j.tree.2009.01.010

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., and Legendre, L. (2009). Global Latitudinal Variations in Marine Copepod Diversity and Environmental Factors. *Proc. R. Soc. B.: Biol. Sci.* 276, 3053–3062. doi: 10.1098/rspb.2009.0742

Sameoto, D., Cochrane, N., and Herman, A. (1993). Convergence of Acoustic, Optical, and Net-Catch Estimates of Euphausiid Abundance: Use of Artificial Light to Reduce Net. *Can. J. Fish. Aquat. Sci.* 50, 334–346. doi: 10.1139/f93-039

Schapire, R. E. (2003). "The Boosting Approach to Machine Learning: An Overview," in *Nonlinear Estimation and Classification*. Eds. D. D. Denison, M. H. Hansen, C. C. Holmes, B. Mallick and B. Yu (New York, NY: Springer), 149–171. Lecture Notes in Statistics. doi: 10.1007/978-0-387-21579-2_9

Schulz, J., Barz, K., Ayon, P., Ludtke, A., Zielinski, O., Mengedoht, D., et al. (2010). Imaging of Plankton Specimens With the Lightframe on-Sight Keyspecies Investigation (LOKI) System. *J. Eur. Optic. Soc.: Rapid Publicat.* 5, 10017s. doi: 10.2971/jeos.2010.10017s

Séférian, R., Berthet, S., Yool, A., Palmieri, J., Bopp, L., Tagliabue, A., et al. (2020). Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6. *Curr. Climate Change Rep.* 6, 95–119. doi: 10.1007/s40641-020-00160-0

Siegel, V. (2005). Distribution and Population Dynamics of Euphausia Superba: Summary of Recent Findings. *Pol. Biol.* 29, 1–22. doi: 10.1007/s00300-005-0058-5

V. Siegel (Ed.) (2016). "Biology and Ecology of Antarctic Krill," in *Advances in Polar Ecology* (Cham: Springer International Publishing). doi: 10.1007/978-3-319-29279-3

Son, D., Cho, H., and Lee, E. J. (2018). Determining Factors for the Occurrence and Richness of Submerged Macrophytes in Major Korean Rivers. *Aquat. Bot.* 150, 82–88. doi: 10.1016/j.aquabot.2018.07.003

Soviadan, Y. D., Benedetti, F., Brandão, M. C., Ayata, S.-D., Irisson, J.-O., Jamet, J. L., et al. (2022). Patterns of Mesozooplankton Community Composition and Vertical Fluxes in the Global Ocean. *Prog. Oceanogr.* 200, 102717. doi: 10.1016/j.pocean.2021.102717

Steinberg, D. K., Cope, J. S., Wilson, S. E., and Kobari, T. (2008). A Comparison of Mesopelagic Mesozooplankton Community Structure in the Subtropical and Subarctic North Pacific Ocean. *Deep. Sea. Res. Part II.: Top. Stud. Oceanogr.* 55, 1615–1635. doi: 10.1016/j.dsr2.2008.04.025

Steinberg, D. K., and Landry, M. R. (2017). Zooplankton and the Ocean Carbon Cycle. *Annu. Rev. Mar. Sci.* 9, 413–444. doi: 10.1146/annurev-marine-010814-015924

Stemmann, L., and Boss, E. (2012). Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. *Annu. Rev. Mar. Sci.* 4, 263–290. doi: 10.1146/annurev-marine-120710-100853

Stemmann, L., Gorsky, G., Marty, J.-C., Picheral, M., and Miquel, J.-C. (2002). Four-Year Study of Large-Particle Vertical Distribution (0–1000 M) in the NW Mediterranean in Relation to Hydrology, Phytoplankton, and Vertical Flux. *Deep. Sea. Res. Part II.: Top. Stud. Oceanogr.* 49, 2143–2162. doi: 10.1016/S0967-0645(02)00032-2

Stemmann, L., Youngbluth, M., Robert, K., Hosia, A., Picheral, M., Paterson, H., et al. (2008). Global Zoogeography of Fragile Macrozooplankton in the Upper 100–1000 M Inferred From the Underwater Video Profiler. *ICES. J. Mar. Sci.* 65, 433–442. doi: 10.1093/icesjms/fsn010

Stukel, M. R., Biard, T., Krause, J., and Ohman, M. D. (2018). Large Phaeodaria in the Twilight Zone: Their Role in the Carbon Cycle: Phaeodarian Ecology in the Twilight Zone. *Limnol. Oceanogr.* 63, 2579–2594. doi: 10.1002/lno.10961

Stukel, M. R., Ohman, M. D., Kelly, T. B., and Biard, T. (2019). The Roles of Suspension-Feeding and Flux-Feeding Zooplankton as Gatekeepers of Particle Flux Into the Mesopelagic Ocean in the Northeast Pacific. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00397

Sun, H., Benzie, P., Burns, N., Hendry, D., Player, M., and Watson, J. (2008). Underwater Digital Holography for Studies of Marine Plankton. *Philos. Trans. R. Soc. A.: Math. Phys. Eng. Sci.* 366, 1789–1806. doi: 10.1098/rsta.2007.2187

Sunday, J. M., Bates, A. E., and Dulvy, N. K. (2012). Thermal Tolerance and the Global Redistribution of Animals. *Nat. Climate Change* 2, 686–690. doi: 10.1038/nclimate1539

Suthers, I., Rissik, D., and Richardson, A. (2019). *Plankton: A Guide to Their Ecology and Monitoring for Water Quality* (Csiro Publishing). Google-Books-ID: e8uPDwAAQBAJ). Available at: https://www.researchgate.net/publication/265624443_Plankton_A_Guide_to_Their_Ecology_and_Monitoring_for_Water_Quality

Suzuki, N., and Not, F. (2015). "Biology and Ecology of Radiolaria," in *Marine Protists*. Eds. S. Ohtsuka, T. Suzaki, T. Horiguchi, N. Suzuki and F. Not (Tokyo: Springer Japan), 179–222. doi: 10.1007/978-4-431-55130-0_8

Turner, J. T. (2002). Zooplankton Fecal Pellets, Marine Snow and Sinking Phytoplankton Blooms. *Aquat. Microbial. Ecol.* 27, 57–102. doi: 10.3354/ame027057

Turner, J. T. (2004). The Importance of Small Planktonic Copepods and Their Roles in Pelagic Marine Food Webs. *Zool. Stud.* 43, 255–266.

Turner, J. T. (2015). Zooplankton Fecal Pellets, Marine Snow, Phytodetritus and the Ocean's Biological Pump. *Prog. Oceanogr.* 130, 205–248. doi: 10.1016/j.pocean.2014.08.005

van der Lingen, C., Bertrand, A., Bode, A., Brodeur, R., Cubillos, L., Espinoza, P., et al. (2009). Trophic Dynamics of Small Pelagic Fish 333–403.

van der Lingen, C., Hutchings, L., and Field, J. (2006). Comparative Trophodynamics of Anchovy *Engraulis Encrasicolus* and Sardine *Sardinops Sagax* in the Southern Benguela: Are Species Alternations Between Small Pelagic Fish Trophodynamically Mediated? *Afr. J. Mar. Sci.* 28, 465–477. doi: 10.2989/18142320609504199

Wiebe, P. H., Boyd, S. H., Davis, M., and Cox, J. L. (1982). Avoidance of Towed Nets by the Euphausiid Nematoscelis Megalops. *Fish. Bull.* 80, 75–91. doi: 10.1080/03610918.2020.1772302

Wiebe, P. H., Morton, A. W., Bradley, A. M., Backus, R. H., Craddock, J. E., Barber, V., et al. (1985). New Development in the MOCNESS, an Apparatus for Sampling Zooplankton and Micronekton. *Mar. Biol.* 87, 313–323. doi: 10.1007/BF00397811

Zhang, M., Straile, D., Chen, F., Shi, X., Yang, Z., Cai, Y., et al. (2018). Dynamics and Drivers of Phytoplankton Richness and Composition Along Productivity Gradient. *Sci. Tot. Environ.* 625, 275–284. doi: 10.1016/j.scitotenv.2017.12.288

Zhou, H., Yang, Y., and Qian, W. (2019). Tweedie Gradient Boosting for Extremely Unbalanced Zero-Inflated Data. *arXiv*:1811.10192 [stat] ArXiv: 1811.10192. doi: 10.1080/03610918.2020.1772302

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor SP declared a shared affiliation with the author FL at the time of review.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 17 August 2022
DOI 10.3389/fmars.2022.936972

OPEN ACCESS

EDITED BY
Anne Chenuil,
Centre National de la Recherche Scientifique (CNRS), France

REVIEWED BY
Dirk Steinke,
University of Guelph, Canada
Owen S. Wangensteen,
The Arctic University of Norway, Norway

*CORRESPONDENCE
Colomban de Vargas
vargas@sb-roscoff.fr
Emmanuel Boss
emmanuel.boss@maine.edu
Manu Prakash
manup@stanford.edu

SPECIALTY SECTION
This article was submitted to Marine Ecosystem Ecology, a section of the journal Frontiers in Marine Science

RECEIVED 05 May 2022
ACCEPTED 11 July 2022
PUBLISHED 17 August 2022

CITATION
de Vargas C, Le Bescot N, Pollina T, Henry N, Romac S, Colin S, Haëntjens N, Carmichael M, Berger C, Le Guen D, Decelle J, Mahé F, Poulain J, Malpot E, Beaumont C, Hardy M, Guiffant D, Probert I, Gruber DF, Allen A, Gorsky G, Follows MJ, Pochon X, Troublé R, Cael BB, Lombard F, Boss E, Prakash M and the Plankton Planet core team (2022) *Plankton Planet*: a frugal, cooperative measure of aquatic life at the planetary scale. *Front. Mar. Sci.* 9:936972. doi: 10.3389/fmars.2022.936972

Plankton Planet: A frugal, cooperative measure of aquatic life at the planetary scale

Colomban de Vargas[1,2,3]*, Noan Le Bescot[1,2], Thibaut Pollina[2,4], Nicolas Henry[1,2,3], Sarah Romac[1,2,3], Sébastien Colin[2,5], Nils Haëntjens[2,6], Margaux Carmichael[2], Calixte Berger[2], David Le Guen[2], Johan Decelle[7], Frédéric Mahé[8], Julie Poulain[3,9], Emmanuel Malpot[10], Carole Beaumont[11], Michel Hardy[12], Damien Guiffant[2], Ian Probert[1], David F. Gruber[13], Andrew E. Allen[14,15], Gabriel Gorsky[2,16], Michael J. Follows[17], Xavier Pochon[18,19], Romain Troublé[2,20], B. B. Cael[2,21], Fabien Lombard[2,16,22], Emmanuel Boss[2,6]*, Manu Prakash[2,4]* and the Plankton Planet core team

[1]Sorbonne Université, CNRS, Station Biologique de Roscoff, UMR7144, ECOMAP, Roscoff, France, [2]Plankton Planet Non-Governmental Organization (NGO), Station Biologique de Roscoff, Roscoff, France, [3]Research Federation for the study of Global Ocean Systems Ecology and Evolution, FR2022/Tara GOSEE, Paris, France, [4]Department of Bioengineering, Stanford University, Stanford, CA, United States, [5]BioOptics Facility, Max Planck Institute for Biology Tübingen, Tübingen, Germany, [6]School of Marine Sciences, University of Maine, Orono, ME, United States, [7]Laboratoire de Physiologie Cellulaire et Végétale, Université Grenoble Alpes, CNRS, CEA, INRA, Grenoble, France, [8]CIRAD, UMR BGPI, Biologie et Génétique des Interactions Plante-Parasites, Montpellier, France, [9]Génomique Métabolique, Genoscope, Institut Francois-Jacob, CEA, CNRS, Université d'Evry, Université Paris-Saclay, Evry, France, [10]Moana New Zealand, Cawthron Aquaculture Park, Nelson, New Zealand, [11]On board 'Folligou', Lorient, France, [12]On board 'Taravana', Lorient, France, [13]Baruch College and the Graduate Center, Department of Natural Sciences, City University of New York, New York, NY, United States, [14]J. Craig Venter Institute, Microbial and Environmental Genomics, La Jolla, CA, United States, [15]Scripps Institution of Oceanography, University of California, La Jolla, CA, United States, [16]Sorbonne Universite', CNRS, Laboratoire d'Oce' anographie de Villefranche, Villefranche-sur-mer, France, [17]Department of Earth, Atmosphere and Planetary Sciences, Massachusetts Institute of Technology(MIT), Cambridge, MA, United States, [18]Biosecurity Group, Cawthron Institute, Nelson, New Zealand, [19]Institute of Marine Science, University of Auckland, Auckland, New Zealand, [20]Tara Ocean Foundation, Paris, France, [21]Ocean Biogeochemistry and Ecosystems, National Oceanography Centre, Southampton, United Kingdom, [22]Institut Universitaire de France (IUF), Paris, France

In every liter of seawater there are between 10 and 100 billion life forms, mostly invisible, called marine plankton or marine microbiome, which form the largest and most dynamic ecosystem on our planet, at the heart of global ecological and economic processes. While physical and chemical parameters of planktonic ecosystems are fairly well measured and modeled at the planetary scale, biological data are still scarce due to the extreme cost and relative inflexibility of the classical vessels and instruments used to explore marine biodiversity. Here we introduce 'Plankton Planet', an initiative whose goal is to engage the curiosity and creativity of researchers, makers, and mariners to (*i*) co-develop a new generation of cost-effective (frugal) universal scientific instrumentation to measure the genetic and morphological diversity of marine microbiomes in context, (*ii*) organize their systematic deployment

through coastal or open ocean communities of sea-users/farers, to generate uniform plankton data across global and long-term spatio-temporal scales, and (*iii*) setup tools to flow the data without embargo into public and explorable databases. As proof-of-concept, we show how 20 crews of sailors were able to sample plankton biomass from the world surface ocean in a single year, generating the first seatizen-based, planetary dataset of marine plankton biodiversity based on DNA barcodes. The quality of this dataset is comparable to that generated by *Tara* Oceans and is not biased by the multiplication of samplers. The data unveil significant genetic novelty and can be used to explore the taxonomic and ecological diversity of plankton at both regional and global scales. This pilot project paves the way for construction of a miniaturized, modular, evolvable, affordable and open-source citizen field-platform that will allow systematic assessment of the eco/morpho/genetic variation of aquatic ecosystems and microbiomes across the dimensions of the Earth system.

KEYWORDS

planetary biology, citizen oceanography, DNA metabarcoding, plankton, seatizens, sailors, frugal science

Introduction

The need for global, long-term surveys of plankton life

The ocean contains 97% of all water on our planet. In every liter of seawater there are between 10 and 100 billion, mostly invisible planktonic life forms. These form a continuous global ecosystem that generates approximately half of planetary oxygen, sustains the large majority of marine life, and regulates atmospheric CO_2 and climate. Understanding and modeling the structure, dynamics, and evolution of global plankton populations is critical for predicting the future of our biosphere and learning how to live in symbiosis with our spaceship, the Earth.

Plankton populations comprise, like in terrestrial biomes, organisms from across the tree of life (viruses, bacteria, archaea, protists, and animals) which interact in complex networks (Lima-mendez et al., 2015; Guidi et al., 2016) that are also shaped by ocean currents and associated physico-chemical environmental parameters (Richter et al., 2019; Logares et al., 2020) – i.e. the *seascape* (Pittman, 2018). But in contrast to terrestrial ecosystems, there are no plants and trees in the plankton - primary production is driven by a large and ancient diversity of photosynthetic bacteria and protists (called phytoplankton) - and the pelagic ecosystem is much more dynamic in terms of both organism life cycle and strategies (e.g. mixotrophy) (Falkowski, 2012) and transport (advection and mixing). The self-organization of local plankton biota into complex ecosystems (Follows et al., 2007) determines their impact on the carbon cycle. For instance, in some regions and seasons, blooms of relatively large cells with mineral components lead to a vigorous sinking flux of organic carbon into the deep sea (Decelle et al., 2013; Durkin et al., 2016). Overall, these fundamental properties of the plankton ecosystem make it arguably the most reactive and proactive compartment of the biosphere to climate change and pollution. Changes in plankton communities can have dramatic effects on global biogeochemical cycles (e.g. Falkowski et al., 2008), climate (e.g. Kwon et al., 2009; Buesseler et al., 2020), and major human societal and economic activities (i.e. fisheries, aquaculture, tourism, etc. Beaugrand and Kirby, 2010; Morgan et al., 2010).

Current models aimed at predicting global ocean ecological changes (e.g. Follows and Dutkiewicz, 2011; Ward et al., 2014) are fairly well constrained in terms of physics and chemistry, but are by comparison heavily oversimplified and unrealistic in terms of biology. In fact, they simply lack good quality, high resolution data on the nature and dynamics of oceanic plankton biodiversity at a planetary scale. While quantitative global data on ocean physics and biogeochemistry are abundantly available by satellites (e.g. Stanley Wilson et al., 2019), *in situ* floats (e.g. Roemmich et al., 2019; Claustre et al., 2020) sail drones (Vazquez-Cuervo et al., 2019), as well as research (e.g. Sloyan et al., 2019) and citizen (e.g. (Simoniello et al., 2019) vessels, standardized biological data are still scarce due to the challenge of sampling and assessing complex communities of fragile plankton in a harmonized and comparable manner (Lombard et al., 2019). Despite the availability of a century's worth of

recorded data on oceanic plankton (Richardson et al., 2006; Buitenhuis et al., 2013, and see for instance https://www.st.nmfs.noaa.gov/copepod/), some areas of the ocean (including much of the south and tropical Pacific Ocean) are almost devoid of any biological observations and we do not yet have a consistent and informed vision of the global distribution and variation of plankton communities.

The largest and longest homogenous survey of plankton life has been conducted using the Continuous Plankton Recorder (CPR), a visionary plankton-scroll instrument created by Sir Alister Hardy in 1931 (Reid et al., 2003), and towed since then behind ferries and cargo ships, particularly in the North Atlantic, North Pacific, and Southern Ocean south of Australia (Batten et al., 2019). The CPR database is currently the only basin-wide standardized historical record of ocean plankton life, and it has given rise to keystone studies describing and modeling basin-scale dynamics of plankton community over time and global climate change (e.g. Beaugrand et al., 2002; Beaugrand and Kirby, 2010; Chivers et al., 2017). However, CPR data also has drawbacks: (*i*) the instrument mainly recovers zooplankton >300µm, thus missing most marine microbiome diversity (see below), (*ii*) it is rather destructive for soft or gelatinous taxa, (*iii*) data analyses rely on taxonomic experts identifying and counting a restricted number of morpho-taxa, (*iv*) the formalin preservation of the silk severely limits subsequent light microscopy observations for some groups and analysis of nucleic acids in general, and (*v*) data come from ships navigating over a restricted number of (mostly northern) commercial routes. CPR data are therefore semi-quantitative and taxonomically and geographically limited.

Over the past 20 years, the revolution in environmental DNA/RNA sequencing has stimulated a new era of global-scale ocean cruises led by molecular and cellular biologists, notably the 'Global Ocean Sampling' (GOS - Venter et al., 2004), *Tara* Oceans (Karsenti et al., 2011), and Malaspina (Duarte, 2015) expeditions. Interestingly, the first two expeditions were private or semi-private enterprises undertaken by sailing boats. The *Tara* Oceans sailing expedition (2009 – 2013) was the longest and most comprehensive: its team developed an eco-systems biology strategy to explore plankton diversity from genes to communities, from viruses to animals, and across coarse but planetary spatial and seasonal scales (Sunagawa et al., 2020). The combination of standardized DNA metabarcoding (De Vargas et al., 2015; Ibarbalz et al., 2019), metagenomic (Sunagawa et al., 2015; Gregory et al., 2019; Zayed et al., 2022) and metatranscriptomic (Carradec et al., 2018; Salazar et al., 2019) datasets is unveiling the basic structure of plankton taxonomic and metabolic diversity (Bork et al., 2015; Sunagawa et al., 2020), generating hypotheses about its interactions (Chaffron et al., 2021), dynamics in the seascape (Richter et al., 2019), and role in emerging ecosystem functions such as the carbon pump (Guidi et al., 2016; Caputi et al., 2019). Notably, the *Tara* Oceans team discovered that the great majority of plankton biodiversity is found in organismal size fractions <100µm, and above all in eukaryotes rather than viruses or prokaryotes (De Vargas et al., 2015; Carradec et al., 2018).

Given the massive local (e.g. plastics, pollutants, Tornero and Hanke, 2016; Jacquin et al., 2019) and global (e.g. deoxygenation, warming, acidification, freshening, ocean circulation changes (Hays et al., 2005; Pörtner et al., 2019) anthropogenic pressures on our ocean, we urgently need global plankton surveys that merge the spatio-temporal sampling power of the CPR with the systems-biology approach of *Tara* Oceans. Only application of a comprehensive and homogenous measure of plankton life from micro- to meso- to global ocean scales (Follows et al., 2007; Lévy et al., 2018; Lombard et al., 2019) will provide the data necessary not only to unveil fundamental principles of ecology and evolution of marine life at the ecosystem level of organization, but also to feed mathematical models of the ocean system that integrate physical, chemical, and biological processes

'Seatizen' oceanography to change scale

A long-term measure of the ocean microbiome across planetary scales is hindered by the extreme cost and limited logistical flexibility of classical oceanographic research vessels and instruments, together with the current impossibility to use autonomous samplers (e.g. floats) to sample the biocomplexity of plankton and generate high-quality data from it. In this context, the thousands of citizen sailing boats (Lauro et al., 2014), professional sailing yachts, the >50,000 cargo ships (https://www.ics-shipping.org/shipping-facts/shipping-and-world-trade) and about the same number of global fishing vessels (https://globalfishingwatch.org/datasets-and-code/vessel-identity/) which are navigating the world ocean every day represent an outstanding opportunity.

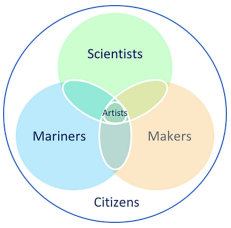
With the miniaturization of sequencing (e.g. Urban et al., 2020) and imaging (Cybulski et al., 2014) devices, the power of cloud computing and artificial intelligence, and the possibilities offered by a participative science engaging actors and volunteers across disciplines and societal/economic activities (Garcia-Soto et al., 2017; Simoniello et al., 2019), six years ago we - a team of researchers, engineers, makers and sailors from France, the US, and New Zealand - started to develop a frugal approach and effective protocols to sample the world ocean plankton for the production of high quality eco/morpho/genetic data in collaboration with recreational and professional mariners. We created '*Plankton Planet*' (P2, https://planktonplanet.org), an international initiative that develops a cooperative, frugal, and sustainable global measure of plankton to assess the biodiversity and health of the world's open and coastal oceans (see P2 vision & mission, Box 1).

10.3389/fmars.2022.936972

BOX 1 Plankton Planet Vision & Mission.

Vision: To harness the creativity of mariners, makers, and researchers for a cooperative, long-term and global measure of aquatic invisible life, toward understanding of our blue planet's biodiversity, evolution, and health for sustainable living in a symbiophere.

Mission: To co-develop a suite of user-friendly and cost-effective tools to collect, measure, and share consistent samples and data from the aquatic microbiome at a planetary scale, providing critical new knowledge on plankton morphology, genetics, and ecology, that will be universally accessible.



We strongly believe that mariners, makers, and researchers share a passionate curiosity and will to explore and discover their environment, such that assembling these communities will generate the necessary synergies to achieve our objectives. Our practical goal is to co-construct a new generation of frugal and open-source yet robust scientific instruments and protocols, which will allow all interested sea-users and sea-farers to collect comparable eco-morpho-genetic data on plankton diversity and abundance at a planetary scale.

In this paper, we present the first steps of this cost-effective, eco-friendly, agile, and society-engaging approach. We show how 20 pioneer crews of citizen sailors – the 'planktonauts'-, equipped with a simple kit to sample plankton for DNA-metabarcoding, were able to help generate a scientifically sound, planetary dataset of plankton biodiversity in less than a year. We finally discuss recent developments (see also Pollina et al., 2022), as well as the 'Plankton Planet' perspectives toward planetary-scale deployment of integrated, affordable and portable 'field-AquaScopes' for long-term assessment of aquatic life and ecosystems at an unprecedented level of sensitivity.

Materials and methods

Between November 2014 and January 2016, with a total budget of $70,000, we provided proof-of-concept for the P2 vision and mission. Our primary goal was to demonstrate that, with the goodwill of mariners, we can sample ocean waters from the entire planet and obtain high quality plankton DNA data for global ecological analyses. We designed a general functional strategy (Supplementary Figure 1) based on robust methods linking affordable and user-friendly instruments for on-board citizen plankton sampling to cutting-edge DNA sequencing and bioinformatic pipelines developed previously in *Tara* Oceans (De Vargas et al., 2015).

Frugal and global plankton sampling by citizen-sailors ('planktonauts')

The P2 PlanktoKit

We first assembled a simple plankton sampling kit including a small net to collect plankton (>20µm) and a manual pumping system to rapidly transfer the freshly collected plankton onto a filter membrane (Figure 1). Of note, this sampling protocol does not rely on toxic chemicals or electricity, which are typically required in a regular laboratory. In order to avoid the need for high-energy storage of the filter membranes (plankton samples) in freezers on board, as well as complex frozen-shipping to the lab, the planktonauts were asked to place the filter on a clean aluminum cove and carefully heat-dry the filter membranes in a pan on the boat gas-cooker, store the dried plankton samples in labeled zip-lock bags with granular desiccants, and send them to the lab *via* standard postal services (Figure 2, Supplementary Material).

Heat-dried plankton DNA preservation

The protocol for plankton DNA preservation by heating and desiccation was first tested during two field campaigns along the French Atlantic coast and compared with gold-standard cryo-fixation and preservation. Six citizen crews sampled plankton at different locations (Supplementary Figure 2A) using the P2 protocol (Figures 1, 2). The concentrated plankton samples were equally divided into two subsamples: one was flash-frozen in liquid nitrogen as classically performed on oceanographic vessels for genetics analyses, while the other was heat-dried as implemented in P2 (Figure 2). On land, samples were preserved in a -80°C freezer (flash-frozen samples) and at room temperature (heat-dried samples) for a couple of months before total DNA extraction and sequencing of 1.3 ± 0.07 million V9 SSU rDNA amplicons per sample (see Sup. Mat. for details). Bioinformatic clustering of the plankton communities (defined by types and abundance of rDNA Operational Taxonomic Units (OTUs, see Sup. Mat.) using different dissimilarity indices indicated that the sub-samples preserved by desiccation and flash-freezing systematically clustered together (Supplementary Figures 2B, C). Plankton communities segregated then first by sampling location, then by their distance to fresh-water input within each bay, irrespective of preservation method. These results proved that the heat drying and subsequent desiccation preservation method do not alter the community composition as measured by DNA metabarcoding and can be applied globally.

Empowering pioneer planktonauts to sample the world oceans.

With modest seed-funding, we were able to assemble a first set of 20 PlanktoKits, and our primary strategy was to maximize

FIGURE 1
The original Plankton Planet sampling kit. Picture on the left: the P2 plankton net (25cm diameter, 20µm mesh size) with a 2kg weight and a small float to maintain it at a maximum depth of 3m. Picture on the right: the manual vacuum pumping system used to transfer plankton from the cod-end of the net onto a 10µm filter membrane. The total cost for one kit approximates $700 but will be largely reduced through tinkering of home-made parts for mass production.

geographic coverage and sampling conditions. We promoted the idea amongst the French sailing community and encountered great enthusiasm to participate. The antique proverb –'*There are three sorts of people: the living, the dead, and those who sail the sea*' is profound: mariners are natural engineers, explorers, and planet-lovers. Word of mouth is powerful in their close-knit community and, limited by the low number of sampling kits available, we soon had to start declining requests. The 27 selected pioneer yacht crews, who we call 'planktonauts' (Figure 3A, and https://planktonplanet.org/the-planktonauts/) represented a wide variety of boats and sailing modes, from multi-year expeditions around the world (e.g. *Race4Water*, *Taravana*, *Folligou*), to family cruises (e.g. *Manevai*, *Nika*, or *Zigomar* - see kidsforsea.over-blog.com/), recreation sailing in the same zone over the year (e.g. *Suhail*), explorers of the poles (e.g. *Vagabond* or *Podorange*), or participants in a New Zealand yachting rally across the South Pacific Islands (Figure 3B, bottom-right insert).

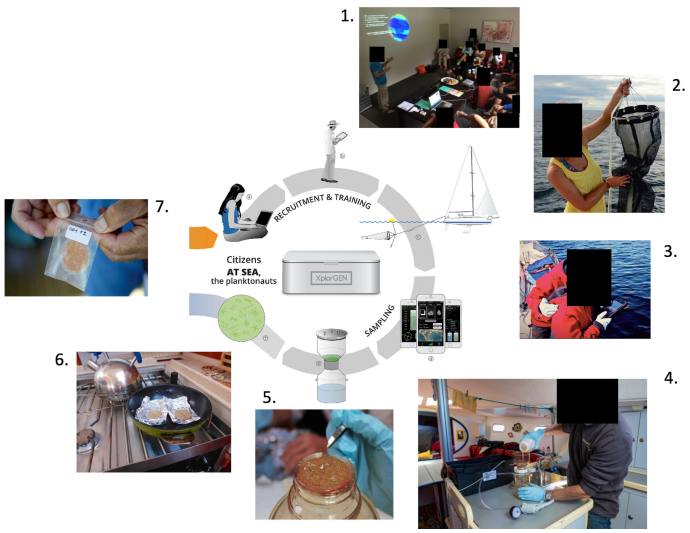


FIGURE 2
Pictures illustrating various steps of the sampling protocol for planktonauts at sea. 1. Training of a group of planktonauts in Auckland with Dr. Pochon. 2. After 15-20min of towing at a maximum speed of 2 knots, the plankton net is recovered on board, here onboard *Tethys*. 3. Recording of contextual parameters, here ocean color using the *Hydrocolor App* on board *Zigomar*. Note that 3 families participated in the pilot project, showing that even kids can realize parts of the protocol (see for instance: https://vimeo.com/219660346); 4. Pouring of the concentrated plankton from the net cod-end into the manual vacuum-pump, onboard *Taravana*. 5. Manual recovery of the 10µm filter membrane full of plankton biomass. 6. Gentle drying of the filter membrane in a pan on the boat (*Taravana*) gas cooker. 7. Storage and labeling of dried plankton samples into zip lock plastic bags before shipping to the laboratory by regular mail.

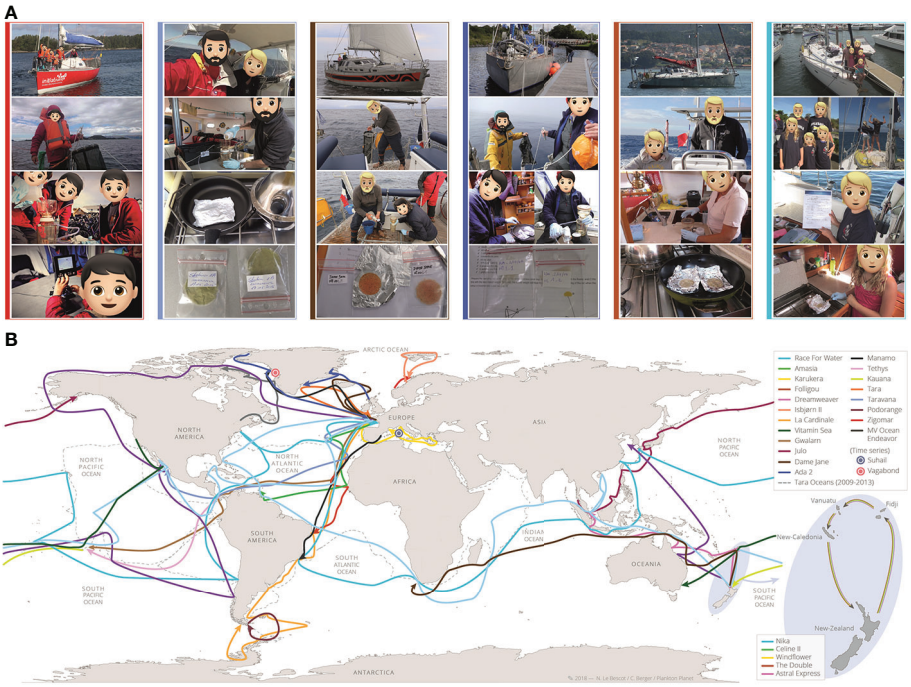


FIGURE 3
Planktonauts and their routes across the world ocean. **(A)** Examples of pictures sent by the planktonauts illustrating their actions at sea: towing and recovering the P2 plankton net, filtering, drying, and storing plankton samples, recording contextual data and observing plankton through a microscope. From left to right columns: N. Fabry and family on board *Zigomar*; M. and A. Hardy on board *Taravana*; J-M. and B. Viant on board *Dame Jane*; I. Autissier and crew on board *Ada2*; C. and D. Beaumont on board *Folligou*; C. McIntyre and family on board *Nika*. **(B)** Routes of the main 20 (of 27) pioneer planktonaut crews recruited during the pilot phase of the project, selected to maximize the geographic coverage and oceanographic and sampling conditions. Note the 5 boats from New Zealand who participated to a rally (May to November 2015) organized in collaboration with the 'Island Cruising Association (http://www.islandcruising.co.nz).

The planktonauts were trained to perform the P2 protocol (Figure 2), individually or in small groups, and regular internet dialogues were established with them during their voyage to answer their questions and follow their progress. They regularly sent movies (e.g. https://vimeo.com/219660346, https://vimeo.com/164511514) and pictures (Figure 3A) of their actions at sea, allowing us to improve our training protocols and outreach. Their feedback on all steps of the protocol at sea (Figure 2) have been key to identify the bottlenecks and challenges to overcome for the implementation stage of the scientific program (see Discussion section).

After 15-20 min of net towing at a speed of ~2 knots (Supplementary Figure 1C), 500mL of concentrated seawater were split and manually filtered onto two replicate 10µm polycarbonate filter-membranes (diameter 47mm), dried at 70° C for 5 min in a pan (on the boat gas-cooker), and then stored at room temperature in a hermetic Ziploc bag with granular desiccants. Upon arrival in a port, samples were shipped to the Roscoff Marine Laboratory (Brittany, France) in a simple envelope *via* regular mail. In the laboratory, the samples were stored at -80°C and information related to samples was archived in a database together with contextual data.

Samples processing, DNA data generation and analyses

Total plankton DNA extraction and rDNA metabarcoding

DNA extraction was performed using a protocol modified from the *NucleoSpin Plant Midi* kit (Macherey-Nagel). One replicate filter of each sample was cut in small pieces and incubated for 2h at 56°C with 3.6mL of the lysis buffer PL1 and 250µL of proteinase K. The lysate was transferred to a large capacity *NucleoSpin* Filter (*DNA Midi kit*) and centrifuged for 10 min at 1,500 g. The eluate was transferred to a new tube and 1 volume of PC buffer was added. The mixture was loaded into the appropriate spin column and washed 3 times with the DNA wash solution. Total DNA was finally eluted twice with 150 µl of DNA elution buffer, and stored in sterile microtubes at -20°C. The amount of recovered DNA was quantified by dsDNA-specific fluorimetry using a Qubit 2.0 Fluorometer with Qubit dsDNA Broad Range and High Sensitivity Assays (ThermoFisher Scientific, Waltham, MA). The DNA quality was double-checked in a subset of samples by running 1 µl on 1.2% agarose gel for 45 min at 120V.

To address general questions of eukaryotic biodiversity over extensive taxonomic and ecological scales, the hyper-variable loop V9 of the Small Sub-Unit (SSU) ribosomal (r) RNA gene was targeted for the generation of amplicons by Polymerase Chain Reaction (PCR). This barcode presents a combination of advantages: (*i*) it is universally conserved in length and simple in secondary structure, thus allowing relatively unbiased PCR amplification across eukaryotic lineages followed by *Illumina*s sequencing, (*ii*) it includes both stable and highly-variable nucleotide positions over evolutionary time frames, allowing discrimination of taxa over a significant phylogenetic depth, (*iii*) it is extensively represented in public reference databases across the eukaryotic tree of life, allowing taxonomic assignment amongst all known lineages.

The first 184 DNA extracts were processed at the Roscoff Marine Laboratory. PCR amplification of the V9 region was performed using 28-bases tagged PCR primers 1389f 5'-*CTTTCCCTACACGACGCTCTTCCGATCT*TTGTACACAC CGCCC -3' and 1510r 5'-*GGAGTTCAGACGTGTGC TCTTCCGATCT*CCTTCYGCAGGTTCACCTAC -3'. The PCR mixture (25µL final volume) contained 10ng of template with 0.35µM final concentrations of each primer, 3% of DMSO and 2X of GC buffer Phusion Master Mix (Finnzymes). Amplifications were conducted following the PCR program: initial denaturation step at 98°C for 30 sec, followed by 25 cycles of 10sec at 98°C, 30sec at 57°C, 30sec at 72°C, and a final elongation step at 72°C for 10 min. Each sample was amplified in triplicate to get enough amounts of amplicons. Results from amplification were checked by running 1 µl of PCR product on a 1.2% agarose gel for 45 min at 120V. All replicates from each sample were then pooled and sent to the GeT Genotoul sequencing platform (Toulouse, France) for library preparation, and loading on 1 lane of Illumina sequencing HiSeq3000 per PCR product. 32 additional DNA extracts were PCR amplified and sequenced by the CEA Genoscope (Evry, France). Amplifications of the V9 from SSU rDNA were conducted with Phusion® High-Fidelity DNA Polymerase (Finnzymes) using the PCR primers 1389f 5'-TTGTACACACCGCCC -3' and 1510r 5'-CCTTCYGCAGGTTCACCTAC -3'. The PCR mixture (25µL final volume) contained 10ng of template with 0.35µM final concentrations of each primer, 3% of DMSO and 2X of GC buffer Phusion Master Mix (Finnzymes). Amplifications were conducted following the PCR program: initial denaturation step at 98°C for 30 sec, followed by 25 cycles of 10sec at 98°C, 30sec at 57°C, 30sec at 72°C, and a final elongation step at 72°C for 10 min. Each sample was amplified in triplicate to get enough amounts of amplicons. PCR products were pooled after amplification and cleaned using AMPure XP beads using a DNA/beads ratio adapted to the length of the amplicon (1,8 Vol). Amplicon lengths were verified using a high-throughput LabChip GX microfluidic capillary electrophoresis system (Perkin Elmer, Waltham, MA, USA), and quantified with a Fluoroskan instrument. A negative control was included in each PCR experiment, as well as a positive control specific to the targeted gene marker. All libraries were prepared using the NEBNext DNA Modules Products and NextFlex DNA barcodes with 100 ng of purified PCR product as input. PCR products were end-repaired, A-tailed at the 3'end, and ligated to Illumina-compatible adaptors using the NEBNext DNA Modules and NextFlex DNA barcodes using a Biomek FX Laboratory Automation Workstation liquid handler (Beckman Coulter Genomics, Danvers, MA, USA), able to perform up to 96 reactions in parallel. After a 1x AMPure XP clean up, the ligated products were amplified using the Kapa Hifi HotStart NGS library Amplification kit, followed by 1x AMPure XP purification. All libraries prepared using the Biomek FX Laboratory Automation Workstation were quantified first by PicoGreen in 96-well plates. Library profiles were assessed using a high throughput microfluidic capillary electrophoresis LabChip GX system (Perkin Elmer, Waltham, MA, USA) and qPCR with the KAPA Library Quantification Kit for Illumina Libraries on an MXPro instrument. Libraries were loaded on 1 lane of Illumina sequencing HiSeq4000 with 20% of PhiX DNA spike-ins (to minimize the impact on the run quality of the low nucleotide diversity at the beginning of the reads, due to the presence of the primer sequences used for amplification), in order to obtain between 3 and 12 millions of paired-end reads 2x150 bp per sample.

Bioinformatic data processing

In order to compare P2 DNA metabarcoding data to the primary *Tara* Oceans (*TO*) eukaryotic metabarcoding dataset (De Vargas et al., 2015), we first merged raw reads from both datasets and applied the following bioinformatics steps. Paired Illumina™ MiSeq reads from the 214 P2 samples and the 883 samples from *Tara* Oceans (2009-2012) were assembled with vsearch v2.7.1 (Rognes et al., 2016) using the command fastq_mergepairs and the option fastq_allowmergestagger. Demultiplexing and primer clipping were performed with cutadapt v1.9 (Martin, 2011) enforcing a full-length match for sample tags and allowing a 2/3-length partial match for forward and reverse primers. Only reads containing both primers were retained. For each trimmed read, the expected error was estimated with vsearch's command fastq_filter and the option *eeout*. Each sample was then dereplicated, i.e. strictly identical reads were merged, using vsearch's command derep_fulllength, and converted to FASTA format. To prepare for clustering, samples were pooled and submitted to another round of dereplication with vsearch. Files containing expected error estimations were also dereplicated to retain only the lowest expected error for each unique sequence. Clustering was performed with Swarm v2.2.2 (Mahé et al., 2015), using a local threshold of one difference and the *fastidious* option. The

representative sequences of each molecular operational taxonomic unit (OTU) were then searched for chimeras with the vsearch's command uchime_denovo (Edgar et al., 2011). In parallel, the OTU representative sequences were pairwise compared to a custom version (https://doi.org/10.5281/zenodo.3768951) of the Protist Ribosomal Reference database PR2 (Guillou et al., 2013), using a global pairwise alignment approach (usearch_global vsearch's command), and taxonomically assigned to their best hit (https://github.com/frederic-mahe/stampa/). In case of ties, the sequence is assigned to the last common ancestor of the references. OTUs with a score below 80% similarity were considered as unassigned. This custom reference database is an update of the V9_PR2 reference database used for the taxonomic assignation of the *Tara* Oceans metabarcodes (De Vargas et al., 2015). A schematic representation of the bioinformatic pipeline used herein together with quantitative details is available in the Figure 6 of the Supplementary Material.

Clustering results, expected error values, taxonomic assignments and chimera detection results were used to build a raw OTU table. Up to that point, reads that could not be merged, reads without tags or primers, reads shorter than 32 nucleotides and reads with uncalled bases ("N") were eliminated. To create the "cleaned" OTU table, additional filters were applied to retain only: non-chimeric OTUs, OTUs with an expected error per nucleotide below 0.0002, and OTUs containing more than 3 reads or seen in 2 samples. The final OTU table, analyzed in this study, integrated the 214 P2 samples together with the 386 *Tara* Oceans samples collected in surface waters for 4 organismal size fractions (0.8-5, 5-20, 20-180, 180-2000 µm). This table, with no taxonomic filtration, was subsampled (rarefied) at the minimum number of reads observed for a sample (313,539 reads) for comparison of OTU richness between samples. Based on the rarefied OTU table, an OTU is considered as unique to P2 if it is present in at least one P2 sample and not present in any of the 386 *Tara* Oceans surface samples. Only OTUs assigned to Eukaryota with more than 80% of similarity were considered for ecological analyses (*e.g.* alpha and beta diversity, taxonomic composition). The plankton samples from '*Objectif Plancton*' used to compare the preservation methods were processed independently using the same bioinformatic pipeline. Raw data were deposited at the European Nucleotide Archive (ENA) under the project ids PRJEB53961 and PRJEB53911, and OTU tables along with corresponding contextual data were also made available in Zenodo @: https://doi.org/10.5281/zenodo.6778240.

Results

Geographical coverage and cost of P2-pilot samples

Despite the relatively low number of samples recovered per boat (2 to 18), the widespread routes of many planktonauts (Figure 3B) yielded plankton samples from 258 sites across the world surface ocean in less than 1 year (Figure 4B). This is remarkable compared to the few tens of spatio-temporally relatively restricted stations usually sampled during classical oceanographic cruises, or even to the 147 sites sampled by the schooner *Tara* in 3 years during her first circumglobal expedition (Bork et al., 2015) (Figure 4B). Furthermore, the total cost of ~200 US$ per sample (including the price of the kit, training, sampling, shipping, and DNA extractions), which could be significantly reduced with increasing sampling frequency and decreasing sampling gears' costs, is one to several orders of magnitude lower than sample cost on oceanographic vessels. The running budget of an open oceanographic research vessel is on the order of 30,000 US $/day. Therefore, a single research vessel traveling at 10 knots and stopping for one hour at each sampling station would have taken 8.6 months - and thus a minimum cost of 7.9 million US$ - to cover 56% of the total P2 sites sampled in slightly more than a year by the planktonauts (Supplementary Figure 3). Note that the average cost to equip and run a sailing boat for a transoceanic route is ~25,000 US$; all together, volunteer planktonauts thus offered ~90% of the cost of field work in sailing charges. This being said, we acknowledge that planktonauts are logistically constrained and will never be able to gather the sort of comprehensive bathymetric, oceanographic and physico-chemical data that oceanographic vessels routinely collect. P2 is not a substitute, but an invaluable complement to the intensive research cruise manned by experts.

High-quality data for global plankton biodiversity and ecology

18 months after the launch of P2, we had extracted total DNA from 214 plankton samples collected by 27 boats, PCR amplified rDNA metabarcodes from each sample, and generated a total of 453 million rDNA reads to assess the diversity of eukaryotic plankton (>20 µm) in the explored surface water masses. The methods used for DNA extraction, sequencing, and analyses were essentially developed in *Tara* Oceans (De Vargas et al., 2015; Alberti et al., 2017); the metabarcode used (V9 SSU rDNA) has proven successful to measure the ecological diversity of total eukaryotic plankton (De Vargas et al., 2015), focus eco-evolutionary analyses on specific groups (Malviya et al., 2015; Cabello et al., 2016; Flegontova et al., 2016; Mordret et al., 2016), reconstruct plankton ecological networks (Lima-mendez et al., 2015), or revisit plankton macro-ecological patterns (Ser-Giacomi et al., 2018; Ibarbalz et al., 2019), or biogeochemical processes (Guidi et al., 2016).

Integration of P2 data into *Tara* Oceans data

Tara Oceans (*TO*) data are today a gold standard for ocean plankton ecology, and we first merged the P2 data into the global *TO* metabarcoding dataset for quality checks and comparison of

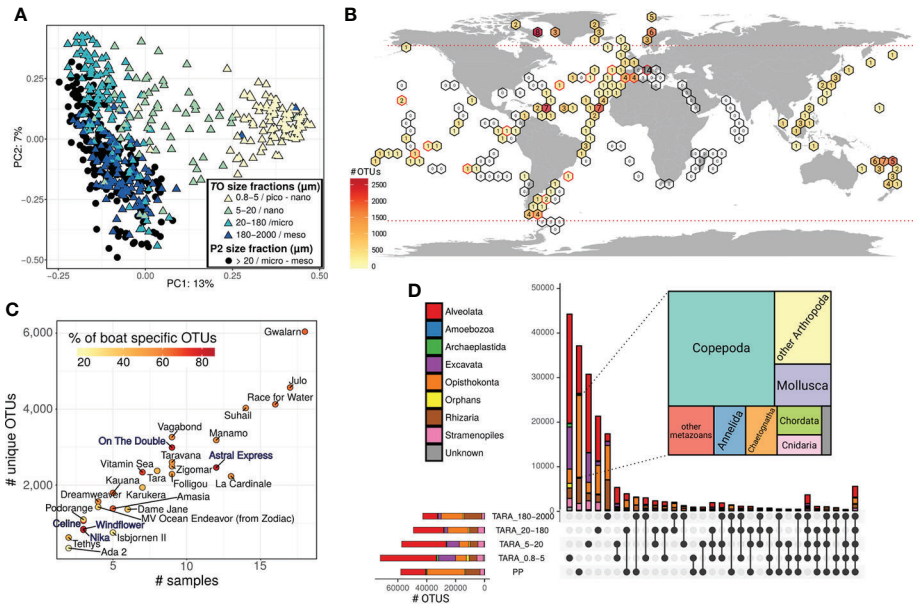


FIGURE 4
The extent, novelty, and community composition of P2 metabarcoding data. **(A)** Grouping of P2 (black dots) and *TO* (colored triangles) plankton communities from surface water according to taxonomic compositional similarity (PCA of Hellinger standardized abundances). Colors correspond to the different plankton size-fractions sampled in *TO*. **(B)** Geographic distribution and novelty of P2 sequenced samples as compared to *TO* samples. Sampled sites are aggregated in pre-defined geographic area (hexagons) for readability. Hexagons containing P2 samples are filled with a color gradient corresponding to the number of rDNA OTUs that were not detected in TO (see colored scale). Number inside the colored hexagons indicate the number of samples collected in the area. Empty hexagons are area with only *TO* samples. Hexagons with red-line borders are area with both P2 and *TO* samples. Horizontal dotted red lines indicate the Northern and Southern 60 Degree latitudes; only samples comprised between these lines were kept for ecological comparative analyses between P2 and TO data. **(C)** Novel plankton diversity (rDNA OTU) uncovered by each boat. Each dot represents a boat, with its position along the X and Y axes corresponding, respectively, to the number of samples collected, and the number of OTUs recovered by the boat that were not detected in TO. The color-gradient indicates the % of novel P2 OTUs that are unique to the particular boat. Note the higher values for the New Zealander crews (blue characters). **(D)** UpSetR plot displaying the taxonomic richness, divided by eukaryotic super-groups, shared between P2 (>20µm) and *TO* size-fractionated samples. The horizontal bars show each individual complete dataset, while vertical bars correspond to the number of OTUs shared between particular datasets (intersections given by the dots under the vertical bars). The tree map shows the taxonomic composition of the 18,430 unique Opisthokonta OTU unveiled in P2, with a dominance (46%) of copepods.

content. As P2 samples were collected by different sailors, sometimes in harsh conditions, we first checked for the presence of obvious biases, such as the sequencing of bacterial contaminants. After taxonomic assignment of the rDNA reads, we found that 1.9% of Plankton Planet reads were assigned to prokaryotes whereas prokaryotes represented 0.5% of *Tara* Oceans reads. This difference is explained by the presence of few outlier samples in P2, typically explained by on-board major processing errors (see Supplementary Figure 4). The median percentages of prokaryotic reads, 0.17% and 0.11% for TO and P2 samples respectively, are, however, comparable (and expected given the very large taxonomic spectrum of the eukaryotic PCR primers used that also amplify prokaryotic genes).

We then compared plankton community composition between P2 and *TO* samples. Principal component analysis confirmed the primary influence of organism size on community structuring (De Vargas et al., 2015), with the pico-nanoplankton (0.8-5 µm) displaying stronger cohesiveness than micro- (20-180 µm) and meso- (180-2000 µm) planktonic communities (Figure 4A). P2 samples clearly fell within the range of variability of *TO* micro- and meso-plankton samples, reflecting the P2 protocol that uses a 20µm mesh-size net and does not apply any pre-filtration. P2 and *TO* micro/meso-plankton size fractions samples spread together along the second axis of the PCA whose variance (7%) is the result of multiple physico-chemical (dispersal and mixing, variations in temperature, light, nutrients, etc) and biological (species interactions, life cycles, behavior, acclimation/adaptation) processes.

Despite the overall similarity between P2 and *TO* micro/meso-plankton samples (Figure 4A), the P2 sampling effort did unveil significant novelty in global plankton diversity. All P2 samples produced rDNA OTUs that had not been reported in the *TO* world ocean survey (Figure 4B), and the number of discovered OTUs clearly increased with sampling effort, both per area (Figure 4B) and per boat (Figure 4C). Each planktonaut

10.3389/fmars.2022.936972

crew unveiled between 62 and 3,907 unique OTUs, unseen by either *Tara* or any other P2 boats, with a discovery rate partly explained by the eccentricity of the sampled areas relative to *TO* sampling sites. Note for instance the high percentage of boat-specific OTUs (>75%) in the New Zealand area (Figure 4C).

We further dug into the composition of plankton diversity unveiled by P2 versus *TO* samples. P2 samples yielded a total of 57,994 OTUs, as compared to 158,716 for *TO* surface ocean samples. Phylogenetic breakdown of the rDNA data (Figure 5) shows the overall taxonomic similarity, in both abundance and richness, between data from P2 and the *TO* micro- and meso-planktonic size fractions. While 8,220 OTUs (mostly metazoans and alveolates) were shared exclusively between P2 and the *TO* larger (>20 µm) organismal size fractions, only 1,981 were common with the smaller (<20 µm) *TO* plankton size fractions (Figure 4D). This emphasizes the consistency of the organisms harvested by P2 and *TO* in comparable plankton size-fractions, and provides organismal data support for our principal component analysis (Figure 4A). Remarkably, however, more than half of all P2 OTUs (37,163) were actually not seen in any plankton size fractions of the *TO* circumglobal dataset (Figure 4D). These correspond mostly (~78%) to Alveolata and Opisthokonta, with an overwhelming majority (98%) of relatively large metazoans (copepods, other arthropods, mollusks, see insert in Figure 4D). This major difference is most likely explained by the fact that the P2 sampling protocol does not involve an upper-size filtration while the *TO* larger size-fraction was constrained by a sieving at 2mm (micro-plankton: 180-2000 µm), resulting in a more complete survey of larger plankton in P2.

In order to check the value of P2 data at finer-grained taxonomic resolution, we explored the phylogenetic and organismal size-fraction distribution of P2 and *TO* OTUs assigned to a well-known phytoplankton group, the dinoflagellate order Peridiniales (Figure 6). Most Peridiniales OTUs observed in *TO* were also found in P2, except a few taxa particularly abundant in the piconano- and nano- size fractions (<20 µm). The relatively large, microplanktonic taxa from the genus *Protoperidinium* were particularly well represented in the P2 dataset, with four OTUs unseen in the *TO* dataset. Note also the presence of many *Blastodinium* and *Brandtodinium* OTUs, which are well-known parasites (Skovgaard et al., 2012) and photosymbionts (Probert et al., 2014) of respectively copepods and radiolarians, and are thus part of the meso- and macro-plankton in their symbiotic stage. Overall, these analyses confirm the quality of the P2 data to assess plankton diversity from the OTU to the community level.

Insights into multiscale plankton ecology using P2 data

We finally used P2 metabarcoding data to explore the consistency of macroecological patterns at both global and local scales, across the data collected by the different boats and planktonauts who sampled plankton independently. At the world ocean scale, the P2 data displayed an increase in alpha diversity (Shannon index) from both poles to the tropics

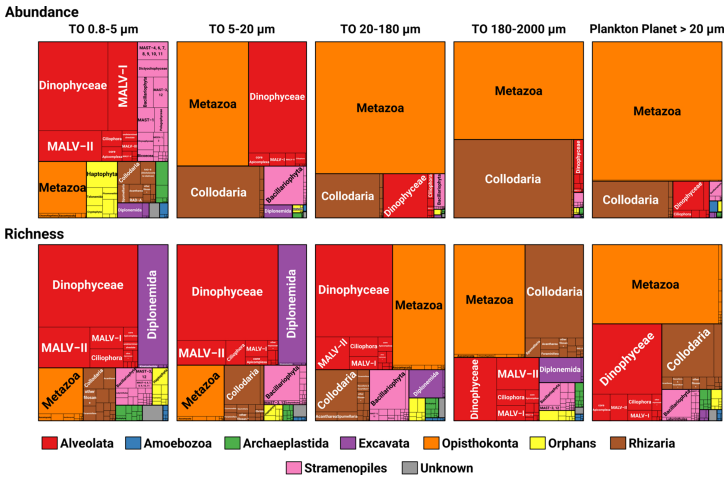


FIGURE 5
Phylogenetic breakdown of the *Tara* Oceans and Plankton Planet global ocean metabarcoding datasets at the eukaryotic supergroup and 'taxogroup' levels. All V9 rDNA reads and OTUs with genetic similarity to a eukaryotic reference sequence ≥80% were retained and taxonomically assigned. The tree-maps display the relative abundance (upper part) and richness (lower part) of the different taxonomic groups in TO plankton size fractions and P2. The category 'Orphans' contains the known but phylogenetically uncertain deep-branching lineages (i.e. Haptophyta, Telonemida, Picomonadida, Katablepharidida, Cryptophyta, Centrohelida and Apusozoan); 'Unknown' corresponds to OTUs assigned to two different supergroups.

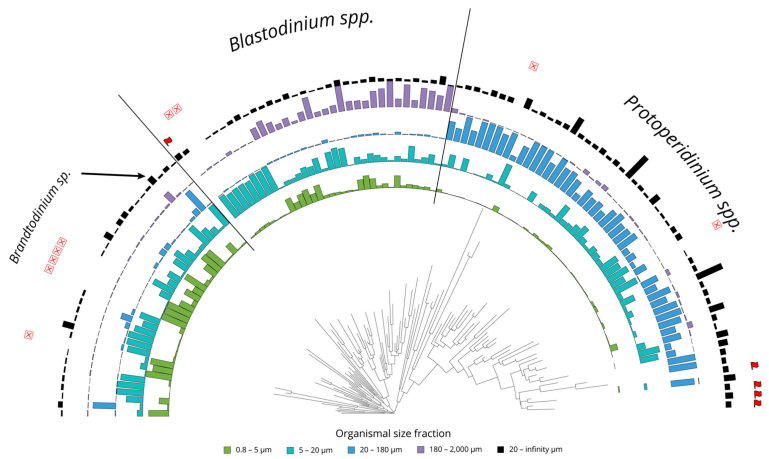


FIGURE 6
Phylogenetic and plankton size fraction distribution of the most abundant OTUs assigned to Peridiniales dinoflagellates in both the P2 and TO metabarcoding datasets. The 100 most abundant Peridiniales OTUs in TO and the 50 most abundant in P2 were selected, together representing 119 unique OTUs. Colored bars indicate the proportion of mean relative abundance of each OTU amongst the four TO plankton size fractions. Black bars indicate the mean abundance (log transformed and scaled to 1) of each OTU amongst all P2 samples. Flag and red crosses symbols indicate OTUs observed respectively only in P2 and only in TO.

(Figure 7A). On a more regional scale, the five planktonaut crews from New Zealand collected enough samples to highlight a biogeographical pattern in this area, with no apparent bias linked to the multiplication of samplers. Clustering of the plankton communities based on their OTU composition into three groups fitted with the three legs of the navigation loop: New Zealand - Fiji; Fiji - New Caledonia, New Caledonia - New Zealand (Figure 7B).

Contextual environmental data

Given the importance of the abiotic environment in structuring planktonic ecosystems (Richter et al., 2019; Logares et al., 2020), we also included the measure of basic environmental parameters at each P2 sampling site. Besides UTC date/time for each sampling event, the planktonauts were asked to record surface water temperature using the water temperature sensor of their boat and/or mercury thermometer (Figure 7A), and broad-band spectral water reflectance using the Hydrocolor App (Leeuw and Boss, 2018). *In situ* temperature data were compared with those from the NASA's MODIS-Aqua satellite, and a good match – on average better than 1°C – was observed. We also developed automated procedures to extract remote sensing ocean color data at each sampling site (https://github.com/OceanOptics/getOC), providing bulk information on chlorophyll-a concentration, as well as suspended dissolved and particulate materials in the target water. In the future, these can not only assist planning of sampling stations at sea in near-real time, but also allow analyses linking ocean color to phytoplankton functional type measured by DNA metabarcoding and the PlanktoScope (see below and Pollina et al., 2022). On the other hand, the Hydrocolor App has been compared to commercial instrumentation and found to correlate well (Yang et al., 2018); it is hoped it will help relate water reflectance to *in-situ* plankton communities, increasing further the utility of remote sensed ocean color.

Discussion

Operational PlanktoKit integrating eco/morpho/genetic data

In the pilot stage of 'Plankton Planet' described herein, we have witnessed the enormous desire of sailors to sample ocean life during their voyages around the globe. We have also generated what is, to our knowledge, the first citizen-based consistent planetary dataset to explore plankton biodiversity and ecology. The DNA metabarcoding data generated from the samples collected by the 27 crews of planktonauts (*i*) displayed a quality comparable to that of the *Tara* Ocean metabarcoding dataset, (*ii*) unveiled significant genetic novelty in zooplankton, and (*iii*) uncovered consistent plankton ecological patterns at both global and local scales. At the world ocean scale, the full dataset displayed a Latitudinal Diversity Gradient that is well known from terrestrial and marine ecosystems (Willig et al., 2003), and has been previously recorded in more restricted planktonic groups (e.g. (Fuhrman et al., 2008; Dolan et al., 2016; Boltovskoy and Correa, 2017) and across plankton kingdoms (Ibarbalz et al., 2019). In the New-Zealand area, the metabarcoding data clustered into three 'eco-

10.3389/fmars.2022.936972

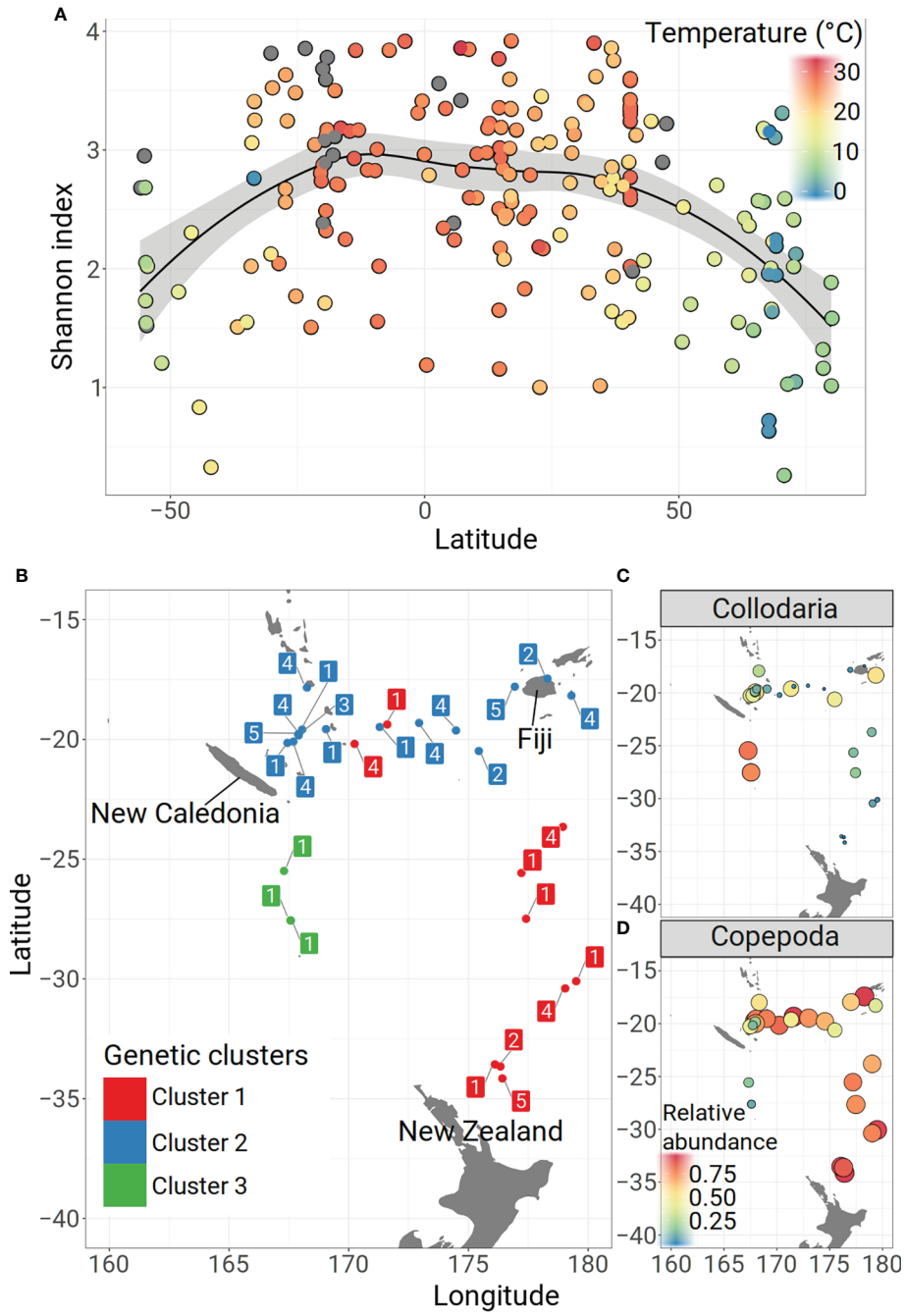


FIGURE 7
P2 insights into plankton ecology. **(A)** *Global ocean latitudinal diversity gradient*. Scatter plot showing the evolution of the Shannon index (alpha diversity) of all P2 samples (colored dots) along latitudes. The color gradient is related to the sea surface temperature measured by the planktonauts during plankton sampling. LOESS regression curve[42] fitting the data. **(B–D)** *Plankton biogeography at the scale of a navigation loop* (the 'kiwi loop', around New Zealand, Fiji and New Caledonia, Figure 2B). **(B)** The samples from this loop (colored dots) were clustered into three groups based on their OTU composition (Jaccard distance) using the Partitioning Around Medoids (PAM) algorithm [43]. See Sup. Figure 5 for details. Two samples from the *Nika* boat, considered as outliers (and collected very close to the coast), were excluded from the analysis. Squared numbers close to each sample site correspond to the sampling boat (1: Astral Express; 2: Celine; 3: Nika; 4: On the Double; 5: Windflower); the colors represent the genetic clusters. **C, D** Relative abundances of the two most abundant taxonomic groups, the collodarians **(C)** and the copepods **(D)**. The color gradient and the size of the circles both correspond to the relative abundance.

genetic' regions along the navigation loop: (*i*) a first portion characterized by high abundances of copepods (Figure 7D) and important richness of indicator OTUs (Dufrene and Legendre, 1997) amongst Marine Alveolates (MALV), Acantharia, Dinophyceae and also copepods (Supplementary Figure 5A); (*ii*) a second portion from Fiji to New Caledonia with lower diversity of indicator OTUs and lower genetic homogeneity (Supplementary Figure 5B), which may be due to the variability of habitats characterizing this region full of islands; (*iii*) a third portion (genetic cluster 3, Figure 7B) characterized by low abundances of copepods and high abundances of collodarians (Figure 7C), with indicator OTUs belonging to collodarians and spumellarians, diatoms (Bacillariophyta), copepods and choanoflagellates. All boats have sampled plankton in at least two different eco-genetic regions, reinforcing both the robustness of the protocol and the accuracy of the emerging plankton biogeography.

However, along the way, we also identified critical challenges to overcome.

Sampling total plankton at sailing-speed

In the P2 pilot project, the planktonauts were asked to maneuver their boats at a speed of less than 2 knots when towing the 20µm-mesh size plankton net. This requests uncomfortable sailing operations impacting the cruising speed, and it was identified as the primary limiting factor for denser sampling. We have therefore been working on the design of new miniaturized *high-speed nets* (von Ammon et al., 2020; Mériguet et al., 2022), inspired by our successful experience during the *Tara* Pacific expedition (Gorsky et al., 2019), as well as simple manual pumping system allowing aspiration of seawater at cruising speed, followed by filtration through a small net system installed on board. The latter device has the advantage of being able to collect pristine water that can also be used to extract DNA/RNA from the smaller plankton size fractions enriched in bacteria, archaea and viruses. Indeed, long-term monitoring of marine plankton systems will ultimately require a sampling protocol covering the 8 orders of plankton organismal size-magnitude, from viruses to animals (Lombard et al., 2019; Sunagawa et al., 2020), in order to assess both top-down and bottom-up ecological mechanisms shaping the ecosystem.

Collecting morphological and behavioral plankton data at sea

The revolution in environmental DNA/RNA sequencing provides the power to comprehensively assess plankton taxonomic and metabolic diversity (Sunagawa et al., 2020). However, meta-omics data convey relatively poor information on the shapes, structures and behaviors of organelles, cells and organisms, which may well be the primary drivers for the self-organization of contemporary ecosystems and their emergent functions (Karsenti, 2012). Fundamental phenotypic mechanisms such as symbioses *sensu lato*, selective feeding, vertical motions and migrations (Krishnamurthy et al., 2019), and many other unknown complex cellular and organismal behaviors, need to be identified and quantified *in the context of the seascape* if we ever want to reach a mechanistic pheno-genomic understanding of ecosystem patterning and functioning in the ocean.

Over the last few years, we have thus developed the *PlanktoScope* (see companion papers by Pollina et al., 2022 and Mériguet et al., 2022), an affordable, miniaturized, modular and evolvable, open source imaging platform for citizen oceanography. For a cost of less than 1000 US$ in parts, the *PlanktoScope* allows both quantitative imaging of microplankton communities through a fluidic module before their storage for total DNA/RNA extraction and genetic analyses in the lab, and high-quality imaging/filming of individual cells or organisms under different types of illumination. Planktonic organisms are notably fragile, versatile, and ephemeral, and the scarcity of live images and movies of plankton *at sea* constitutes arguably the major knowledge-gap in oceanography, today's oceanographers spending most of their time behind their office computer screens or analyzing in the lab fixed and thus highly-altered plankton samples. Beyond generating quantitative phenotypic data complementary to DNA metabarcoding, the *PlanktoScope* can trigger an emotional shock in each planktonaut discovering the beauty of the marine microbiome, changing their view of the ocean forever and further driving their will to explore and preserve the invisible life thriving under their boat.

Toward seatizen field-Aquascopes for perennial cooperative monitoring of the global aquatic microbiome

Today we have developed all parts to assemble a stable P2 PlanktoKit 1.0 allowing any engaged seatizen to generate consistent plankton eco/morpho/genetic data directly in the field. The PlanktoKit 1.0, which costs less than $3,000 in parts, consists of a cruising-speed plankton collection system (von Ammon et al., 2020), a PlanktoScope (Pollina et al., 2022), a new cost-effective plankton DNA collection-kit we call the '*Lamprey*', and a tablet computer to drive the PlanktoScope, record all contextual and meta-data, and visualize and record plankton images and movies. We have shown that these frugal tools deployed from sailing boats can generate consistent, planetary scale genetic (this paper) and morphological (Mériguet et al., 2022) data, paving the way for the Plankton Planet vision (Box 1). Deployment of the first kits on key navigation loops and routes will start in 2023 in the Atlantic oceans as part of the European research program AtlantEco. We believe that implementation of a perennial and self-sustainable survey of the world ocean surface plankton could happen by

2030, depending on our capacity to organize the synergies between seatizens (crews of all kinds of boats - pleasure and racing yachts, commercial, fishing, and defense vessels - fishermen and aquaculturists, managers), makers and engineers, and reseachers. This would bring critical new knowledge and outputs for human societies sharing the blue planet (Box 2).

At the core of our long-term vision, the principles applied to development of the *PlanktoScope* (Pollina et al., 2022) should be extended to co-construct a field-AquaScope for seatizen exploration of aquatic eco-systems, and the concept could be easily applied for the exploration of other planetary biomes. We foresee the field-AquaScope as a miniaturized platform that can host modules and sensors to measure the critical biological, but also (bio)chemical, and physical parameters of any aquatic eco-system. Each module (including the *PlanktoScope*) is co-developed by small international and interdisciplinary teams of (*i*) researchers who define the scientific specifications to address fundamental question in global ecology, (*ii*) mariners and citizen samplers who assess practical user-constraints linked to specific field conditions (e.g. *on-board* sailing boats or cargo-ships), and (*iii*) makers and engineers who design and construct the frugal modules to measure given parameters. The Plankton Planet ecosystem will provide overall coordination to ensure that all modules can be fully integrated into the miniaturized field-laboratory and generate good-quality homogenous data that are shared in public databases.

Over the last century, societies have spent phenomenal amounts of money to discover the properties of the infinitely small (atoms and cells) and the infinitely large (weather, planets and stars, and the universe). But it is in between these extremes that we find arguably the most complex object in the known universe: eco-systems, i.e. self-organized and evolving life networks that interact with physico-chemical processes at micro- to planetary scales. Ecosystems have shaped the Earth atmosphere and biogeochemical cycles, they currently buffer climate change while providing sources of food and medicine, they will determine the future habitability of our planet and the fate of humanity. Due to their extreme complexity, integrating biology, chemistry, and physics, ecosystems have long escaped holistic quantitative assessment. Today the tools and methods exist, in particular automated sequencing and imaging, as well as artificial intelligence and massive computing to collect and integrate the layers of big data needed to understand eco-systems. The major challenge is to develop the new generation of affordable tools that can be deployed in a systematic manner across the spatio-temporal dimension of the Earth system. Plankton Planet proposes a coherent and frugal approach to overcome this challenge, at the interface of science and society, using one of the simplest biomes – marine waters- as a case study. Clearly being at the onset of planetary biology, we hope that Plankton Planet will contribute to reaching a profound understanding of our habitat in the decades to come, and to learning how to live in synergy with the biosphere.

BOX 2 Perspective outputs of Plankton Planet.

Oceanography 3.0 ecosystem

- An international fleet of planktonauts to act as sentinels and the collective consciousness of the biological health of our oceans.
- An evolvable toolkit of affordable, scientifically relevant instruments for seatizen-based assessment of aquatic (marine and freshwater) biodiversity and ecosystems.
- An ever-growing cryo-bank of global ocean DNA samples, archiving the memory of our changing oceans for future generations and technologies.
- A continuous flow of standardized ocean microbiome imaging and genetic data at the planetary scale, available for fundamental and applied science, policy-makers, and education.

Blue-sky science

- Unique long-term monitoring of the distribution and evolution of global plankton biodiversity in our fast-changing ocean; including analysis of the impacts of warming, acidification, and de-oxygenation.
- Novel understanding of abundance, structures, functions, and behaviors of plankton life *via* the billions of images and movies of planktonic organisms generated at sea using *PlanktoScopes*.
- Incorporation of high-resolution global-scale biological data into efforts to model the dynamics of ocean ecosystems and ocean-climate interactions.

Applied science & policies

- Detection of invasive, toxic, or economically relevant species at the planetary scale.
- Evaluation of the health of oceanic regions based on species content, species richness, and trends in time.
- Assessment of ocean biogeographic zones ("seascapes") for optimal design of marine protected areas in the high seas (e.g. ecosystems with high capacity for carbon pumping).
- Linking of data on plankton communities (including fish gametes and larvae, and eDNA) to fish-catch data for robust prediction of fish stocks.

Education through 'Plankton & Arts'

- Collective awareness of the ocean microbiomes and their planetary impact, both directly (through the *PlanktoScope*) and indirectly *via* the shared images, movies, and 3D prints of plankton. Educational tools will notably be distributed by planktonauts in remote countries and islands whose communities interact with and depend on the marine ecosystem.

The plankton planet core team

The 'Plankton Planet' core team includes: Romain Bazile (Seatizen for Plankton Planet, SBR, Roscoff, France), Emmanuel Boss (University of Maine, Orono, USA), Guillaume Bourdin (University of Maine, Orono, USA), BB Cael (NOC, Southampton, UK), Roberto Casati (CNRS IJN, Paris, France), Sébastien Colin (MPI Tübingen, Germany), Colomban de Vargas (CNRS SBR, Roscoff, France), Gabriel Gorsky (Sorbonne Université, LOV, Villefranche/Mer, France), Damien Guiffant (Seatizen for Plankton Planet, SBR, Roscoff, France), Nils Haentjens (University of Maine, Orono, USA), Nicolas Henry (CNRS SBR, Roscoff, France), Adam Larson (Stanford University, Stanford, USA), Noan Le Bescot (Seatizen for Plankton Planet, SBR, Roscoff, France), Fabien Lombard (Sorbonne Université, LOV, Villefranche/Mer, France), Gilles Mirambeau (Sorbonne Université, OOB, Banyuls/Mer, France), Clémentine Moulin (Tara Ocean Foundation, Paris, France), Anna Oddone (Seatizen for Plankton Planet, SBR, Roscoff, France), Manu Prakash (Stanford University, Stanford, USA), Christophe Prazuck (Institut de l'Ocean, Sorbonne Université, Paris, France), Vincent Raimbault (CNRS LAAS, Toulouse, France), Clara Trellu (Seatizen for Plankton Planet, SBR, Roscoff, France), Romain Troublé (Tara Ocean Foundation, Paris, France).

Data availability statement

All raw and secondary data produced and analysed in this study are publicly available. Raw sequencing data were deposited at the European Nucleotide Archive (ENA) under the project ids PRJEB53961 and PRJEB53911, and OTU tables along with corresponding contextual data are available in Zenodo: https://doi.org/10.5281/zenodo.6778240.

Author contributions

CdV: Founded Plankton Planet, conception and design of the work, data analysis and interpretation, wrote the manuscript, supervised the study NLB: study design, manuscript edition, field work TP: study design, logistical support NHe: data analysis SR: study design, data production SC: study design, constructive comments NHa: data analysis MC: field work, logistics CalB: study design, logistics DLG: instrument design JD: constructive comments, logistics FM: data analysis JP: data production EM: logistical support CalB: field sampling MH: field sampling the P2 core team: critical review of the work and manuscript DG: administration, logistics IP: critical revision of the article DFG: study design AA: study design GG: constructive comments MF: study design XP: study design RT: logistical support BC: study design FB: study design EB: design and supervision of the work MP: design and supervision of the work All authors revised and edited the manuscript, and approved the version sent for publication

Funding

The Plankton Planet pilot project was made possible by seed-funding from the Richard Lounsbery Foundation. Additional funding came from the French Government "Investissements d'Avenir" program OCEANOMICS (ANR-11-BTBR- 0008) and National Research Agency grant CORALGENE (ANR-17-CE02-0020), the 'Plankton Arts' grant from the Fondation d'Entreprise Total, the New Zealand Royal Society Dumont-D'Urville (DDU-CAW1501) and Strategic Seeding (16-CAW-008-CSG) Funds, the Cawthron Institute, as well as the Okeanos, Tara Ocean, and Schmidt Futures Foundations.

Acknowledgments

We are deeply grateful to the pioneer planktonauts: E. Abadie on board *Manevai*, I. Autissier on board *Ada 2*, R. Barnaud on board *Kauana*, C. Beaumont on board *Folligou*, E. Brossier on *Vagabond*, K. & K. Brownie on board *Astral Express*, V. Chirié on board *Dreamweaver*, I. Corrias on board *Attila 3*, Q. d'Avout on board *Vitamin Sea*, B. de Ravignan on board *La Cardinale*, M. Denis on board *Isbjørn II*, E. Di Iorio on board *Suhail*, B. Dumontet on board *Ainez*, E. Dupuis on board *Vaihere*, N. Fabbri on board *Zigomar*, Y. Gladu on board *Tethys*, J. Goulias on board *Gwalarn*, M. Hardy & A. Schnarwiler on board *Taravana*, P. Laparre on board *Amasia*, A. Marchandise on board *Karukera*, L. Marie on board *Vagabond*, J. & L. Martin on board *Windflower*, C., L., G., O. & R. McIntyre on board *Nika*, M. & C. Stephens on board *The Double*, B. Monégier on board *Podorange*, D. & S. Patterson on board *Celine*, Lorena Piana on board *Giulia*, A. Schmid on board *Julo*, M. Simeoni on board *Race for Water*, Students on Ice on board the MV Ocean Endeavor, Tara's crew on board *Tara*, S. Valcke on board *Manamo*, & JM. Viant on board *Dame Jane*. We thank also Pierre Mollo, Francoise Gaill, Éric Karsenti, Paul Falkowski, Étienne Bourgois, Catherine Chabaud, Roland Jourdain, Isabelle Autissier, Raphaela le Gouvello, Daniel Richter, Chris Cornelisen and Gilles Boeuf, for their continuous advice and support. This article is contribution number 1 of Plankton Planet.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.936972/full#supplementary-material

References

Alberti, A., Poulain, J., Engelen, S., Labadie, K., Romac, S., Ferrera, I., et al. (2017). Viral to metazoan marine plankton nucleotide sequences from the Tara oceans expedition. *Sci. Data* 4, 1–20. doi: 10.1038/sdata.2017.93

Batten, S. D., Abu-Alhaija, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., et al. (2019). A global plankton diversity monitoring program. *Front. Mar. Sci.* 6 (APR). doi: 10.3389/fmars.2019.00321

Beaugrand, G., and Kirby, R. R. (2002). Reorganization of north Atlantic marine copepod biodiversity and climate. *Science* 296 (5573), 1692–1694. doi: 10.1126/science.1071329

Beaugrand, G., and Kirby, R. R. (2010). Climate, plankton and cod. *Global Change Biol.* 16 (4), 1268–1280. doi: 10.1111/j.1365-2486.2009.02063.x

Boltovskoy, D., and Correa, N. (2017). Planktonic equatorial diversity troughs: Fact or artifact? latitudinal diversity gradients in radiolaria. *Ecology* 98 (1), 112–124. doi: 10.1002/ecy.1623

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., Wincker, P., et al. (2015). Tara Oceans studies plankton at planetary scale. *Science* 348 (6237), 873. doi: 10.1126/science.aac5605

Buesseler, K. O., Boyd, P. W., Black, E. E., and Siegel, D. A. (2020). Metrics that matter for assessing the ocean biological carbon pump. *Proc. Natl. Acad. Sci.* 117 (18), 201918114. doi: 10.1073/pnas.1918114117

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., et al. (2013). MAREDAT: Towards a world atlas of MARine ecosystem DATa. *Earth System Sci. Data* 5 (2), 227–239. doi: 10.5194/essd-5-227-2013

Cabello, A. M., Cornejo-Castillo, F. M., Raho, N., Blasco, D., Vidal, M., Audic, S., et al. (2016). Global distribution and vertical patterns of a prymnesiophyte–cyanobacteria obligate symbiosis. *ISME J.* 10 (3), 693–706. doi: 10.1038/ismej.2015.147

Caputi, L., Carradec, Q., Eveillard, D., Kirilovsky, A., Pelletier, E., Pierella Karlusich, J. J., et al. (2019). Community-level responses to iron availability in open ocean plankton ecosystems. *Global Biogeochemical Cycles* 33 (3), 391–419. doi: 10.1029/2018GB006022

Carradec, Q., Pelletier, E., Da Silva, C., Alberti, A., Seeleuthner, Y., Blanc-Mathieu, R., et al. (2018). A global ocean atlas of eukaryotic genes (SupMat). *Nat. Commun.* 9 (1), 1–13. doi: 10.1038/s41467-017-02342-1

Chaffron, S., Delage, E., Budinich, M., Vintache, D., Henry, N., Nef, C., et al. (2021). Environmental vulnerability of the global ocean epipelagic plankton community interactome. *Sci. Adv.* 7 (35), 1–16. doi: 10.1126/sciadv.abg1921

Chivers, W. J., Walne, A. W., and Hays, G. C. (2017). Mismatch between marine plankton range movements and the velocity of climate change. *Nat. Commun.* 8, 1–8. doi: 10.1038/ncomms14434

Claustre, H., Johnson, K. S., and Takeshita, Y. (2020). Observing the global ocean with biogeochemical-argo. *Annu. Rev. Mar. Sci.* 12 (1), 23–48. doi: 10.1146/annurev-marine-010419-010956

Cybulski, J. S., Clements, J., and Prakash, M. (2014). Foldscope: Origami-based paper microscope. *PLoS One* 9 (6), e98781. doi: 10.1371/journal.pone.0098781

Decelle, J., Martin, P., Paborstava, K., Pond, D. W., Tarling, G., Mahé, F., et al. (2013). Diversity, ecology and biogeochemistry of cyst-forming acantharia (Radiolaria) in the oceans. *PLoS One* 8 (1), e53598. doi: 10.1371/journal.pone.0053598

De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science* 348 (6237), 1261605–1/11. doi: 10.1007/s13398-014-0173-7.2

Dolan, J. R., Yang, E. J., Kang, S. H., and Rhee, T. S. (2016). Declines in both redundant and trace species characterize the latitudinal diversity gradient in tintinnid ciliates. *ISME J.* 10 (9), 2174–2183. doi: 10.1038/ismej.2016.19

Duarte, C. M. (2015). Seafaring in the 21st century: The malaspina 2010 circumnavigation expedition. *Limnology Oceanography Bull.* 24 (1), 11–14. doi: 10.1002/lob.10008

Dufrene, M., and Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecol. Monogr.* 67 (3), 345–366. doi: 10.2307/2963459

Durkin, C. A., Van Mooy, B. A.S., Dyhrman, S. T., and Buesseler, K. O. (2016). Sinking phytoplankton associated with carbon flux in the Atlantic ocean. *Limnology Oceanography* 61 (4), 1172–1187. doi: 10.1002/lno.10253

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Falkowski, P. (2012). The power of plankton. *Science* 483, 17–20. doi: 10.1038/483S17a

Falkowski, P. G., Fenchel, T., and Delong, E. F. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science* 320 (5879), 1034–1039. doi: 10.1126/science.1153213

Flegontova, O., Flegontov, P., Malviya, S., Audic, S., Wincker, P., de Vargas, C., et al. (2016). Extreme diversity of diplonemid eukaryotes in the ocean. *Curr. Biol.* 26 (22), 3060–3065. doi: 10.1016/j.cub.2016.09.031

Follows, M. J., and Dutkiewicz, S. (2011). Modeling diverse communities of marine microbes. *Annu. Rev. Mar. Sci.* 3, 427–451. doi: 10.1146/annurev-marine-120709-142848

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* 315 (5820), 1843–1846. doi: 10.1126/science.1138544

Fuhrman, J. A., Steele, J. A., Hewson, I., Schwalbach, M. S., Brown, M. V., Green, J. L., et al. (2008). A latitudinal diversity gradient in planktonic marine bacteria. *Proc. Natl. Acad. Sci.* 105 (22), 7774–7778. doi: 10.1073/pnas.0803070105

Garcia-Soto, C, van der Meeren, G. I., Busch, J. A., Delany, J., Domegan, C., Dubsky, K., et al. (2017). *Advancing citizen science for coastal and ocean research. position p.* Eds. V. Kellett, J. Delany and N McDonough. Position Paper 23 of the European Marine Board, Ostend, Belgium. 112pp.

Gorsky, G., Bourdin, G., Lombard, F., Pedrotti, M. L., Audrain, S., Bin, N., et al. (2019). Expanding Tara oceans protocols for underway, ecosystemic sampling of the ocean-atmosphere interface during Tara pacific expedition 2016–2018). *Front. Mar. Sci.* 6 (November). doi: 10.3389/fmars.2019.00750

Gregory, A. C., Zayed, A. A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et al. (2019). Marine DNA viral macro- and microdiversity from pole to pole. *Cell* 177 (5), 1109–1123.e14. doi: 10.1016/j.cell.2019.03.040

Guidi, L., Chaffron, S., Bittner, L., Eveillard, D., Larhlimi, A., Roux, S., et al. (2016). Plankton networks driving carbon export in the oligotrophic ocean. *Nature* 532, 465–470. doi: 10.1038/nature16942

Guillou, L., Bachar, D., Audic, S., Bass, D., Berney, C., Bittner, L., et al. (2013). The protist ribosomal reference database (PR2): A catalog of unicellular eukaryote small Sub-unit rRNA sequences with curated taxonomy. *Nucleic Acids Res.* 41 (D1), 597–604. doi: 10.1093/nar/gks1160

Hays, G. C., Richardson, A. J., and Robinson, C. (2005). Climate change and marine plankton. *Trends Ecol. Evol.* 20 (6 SPEC. ISS.), 337–344. doi: 10.1016/j.tree.2005.03.004

Ibarbalz, F. M., Henry, N., Brandão, M. C., Martini, S., Busseni, G., Byrne, H., et al. (2019). Global trends in marine plankton diversity across kingdoms of life. *Cell* 179 (5), 1084–1097.e21. doi: 10.1016/j.cell.2019.10.008

Jacquin, J., Cheng, J., Odobel, C., Pandin, C., Conan, P., Pujo-Pay, M., et al. (2019). Microbial ecotoxicology of marine plastic debris: A review on colonization and biodegradation by the 'plastisphere. *Front. Microbiol.* 10 (APR). doi: 10.3389/fmicb.2019.00865

Karsenti, E. (2012). A journey from reductionist to systemic cell biology aboard the schooner Tara. *Mol. Biol. Cell* 23 (13), 2403–2406. doi: 10.1091/mbc.E11-06-0571

Karsenti, E., Acinas, S. G., Bork, P., Bowler, C., de Vargas, C., Raes, J., et al. (2011). A holistic approach to marine eco-systems biology. *PloS Biol.* 9 (10), 7–11. doi: 10.1371/journal.pbio.1001177

Krishnamurthy, D., Li, H., Rey du Cambournac, F. B.P., Larson, A., and Prakash, M. (2019). Scale-free vertical tracking microscopy: Towards bridging scales in biological oceanography. *bioRxiv*, 610246. doi: 10.1101/610246

Kwon, E. Y., Primeau, F., and Sarmiento, J. L. (2009). The impact of remineralization depth on the air-sea carbon balance. *Nat. Geosci.* 2 (9), 630–635. doi: 10.1038/ngeo612

Lauro, F. M., Senstius, S. J., Cullen, J., Neches, R., Jensen, R. M., Brown, M. V., et al. (2014). The common oceanographer: Crowdsourcing the collection of oceanographic data. *PloS Biol.* 12 (9), e1001947. doi: 10.1371/journal.pbio.1001947

Leeuw, T., and Boss, E. (2018). The HydroColor app: Above water measurements of remote sensing reflectance and turbidity using a smartphone camera. *Sensors (Switzerland)* 18 (1), 1–15. doi: 10.3390/s18010256

Lévy, M., Franks, P. J. S., and Smith, K. S. (2018). The role of submesoscale currents in structuring marine ecosystems. *Nat. Commun.* 9 (1), 1–16. doi: 10.1038/s41467-018-07059-3

Lima-mendez, G., Faust, K., Henry, N., Colin, S., Carcillo, F., Chaffron, S., et al. (2015). Determinants of community structure in the global plankton interactome. *Sci. (New York N.Y.)* 348 (6237), 1262073. doi: 10.1126/science.1262073

Logares, R., Deutschmann, I. M., Junger, P. C., Giner, C. R., Krabberød, A. K., Schmidt, T. S.B., et al. (2020). Disentangling the mechanisms shaping the surface ocean microbiota. *Microbiome* 8 (1), 55. doi: 10.1186/s40168-020-00827-8

Lombard, F., Boss, E., Waite, A. M., Uitz, J., Stemmann, L., Sosik, H. M., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* 6 (MAR). doi: 10.3389/fmars.2019.00196

Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M., et al. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ* 3, e1420. doi: 10.7717/peerj.1420

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Bittner, L., et al. (2015). Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci.* 348, E1516–E1525. doi: 10.1073/pnas.1509523113

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17 (1), 10. doi: 10.14806/ej.17.1.200

Mordret, S., Romac, S, Henry, N, Colin, S, Carmichael, M, Berney, C, et al. (2016). The symbiotic life of symbiodinium in the open ocean within a new species of calcifying ciliate (Tiarina sp.). *ISME J.* 10, 1424–1436. doi: 10.1038/ismej.2015.211

Mériguet, Z., Oddone, A., Le Guen, D., Pollina, T., Basile, R., Moulin, C., et al (2022). Basin-scale underway quantitative survey of surface microplankton using affordable collection and imaging tools deployed from Tara. *Front. Mar. Sci.* 1–12. doi: 10.3389/fmars.2022.916025

Morgan, K. L., Larkin, S. L., and Adams, C. M. (2010). Red tides and participation in marine-based activities: Estimating the response of southwest Florida residents. *Harmful Algae* 9 (3), 333–341. doi: 10.1016/j.hal.2009.12.004

Pittman, S. J. (2018). Seascape ecology. Ed. S. J. Pittman Wiley. (Oxford).

Pollina, T., Larson, A. G., Lombard, F., Hongquan, L., Le Guen, D., Colin, S., et al. (2022). PlanktoScope: Affordable modular imaging platform for citizen oceanography. *Front. Mar. Sci* 9, 949428. doi: 10.3389/fmars.2022.949428.

Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., Zhai, P., M, T., Poloczanska, E., et al. (2019). *IPCC special report on the ocean and cryosphere in a changing climate* (Cambridge University Press, Cambridge, UK and New York, NY, USA) 755 p. doi: 10.1017/9781009157964

Probert, I., Siano, R., Poirier, C., Decelle, J., Biard, T., Akihiro, T., et al. (2014). Brandtodinium gen. nov. (Dinophyceae), a dinoflagellates commonly found in symbiosis with polycystines radiolarians. *J. Phycology* 50 (2), 388–399. doi: 10.1111/jpy.12174

Reid, P. C., Colebrook, J. M., Matthews, J. B.L., and Aiken, J. (2003). The continuous plankton recorder: concepts and history, from plankton indicator to undulating recorders. *Prog. Oceanography* 58 (2–4), 117–173. doi: 10.1016/j.pocean.2003.08.002

Richardson, A. J., Walne, A. W., John, A. W.G., Jonas, T. D., Lindley, J. A., Sims, D. W., et al. (2006). Using continuous plankton recorder data. *Prog. Oceanography* 68, 27–74. doi: 10.1016/j.pocean.2005.09.011

Richter, D. J., Watteaux, R., Vannier, T., Leconte, J., Frémont, P., Reygondeau, G., et al. (2019). Genomic evidence for global ocean plankton biogeography shaped by large-scale current systems. *bioRxiv*, 867739. doi: 10.1101/867739

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K. S., King, B., Moum, J., et al. (2019). On the future of argo: A global, full-depth, multi-disciplinary array. *Front. Mar. Sci.* 6 (JUL). doi: 10.3389/fmars.2019.00439

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584. doi: 10.7717/peerj.2584

Salazar, G., Paoli, L., Alberti, A., Huerta-Cepas, J., Ruscheweyh, H. J., Cuenca, M., et al. (2019). Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* 179 (5), 1068–1083.e21. doi: 10.1016/j.cell.2019.10.014

Ser-Giacomi, E., Zinger, L., Malviya, S., de Vargas, C., Karsenti, E., Bowler, C., et al. (2018). "Ubiquitous abundance distribution of non-dominant plankton across the world's ocean," *Nat. Ecol. Evol.* 2, 1243–1249. doi: 10.1038/s41559-018-0587-2.

Simoniello, C., Jencks, J., Lauro, F. M., Loftis, J. D., Weslawski, J. M., Deja, K., et al. (2019). Citizen-science for the future: Advisory case studies from around the globe. *Front. Mar. Sci.* 6:1–15. doi: 10.3389/fmars.2019.225

Skovgaard, A., Karpov, S. A., and Guillou, L. (2012). The parasitic dinoflagellates blastodinium spp. inhabiting the gut of marine, planktonic copepods: Morphology, ecology, and unrecognized species diversity. *Front. Microbiol.* 3 (AUG). doi: 10.3389/fmicb.2012.00305

Sloyan, B. M., Wanninkhof, R., Kramp, M., Johnson, G. C., Talley, L., Tanhua, T., et al. (2019). The global ocean ship-base hydrographic investigations program (GO-SHIP): A platform for integrated multidisciplinary ocean science. *Front. Mar. Sci.* 6 (JUL). doi: 10.3389/fmars.2019.00445

Stanley Wilson, W., Lindstrom, E. J., and Willis, J. K. (2019). "Satellite oceanography-history and introductory concepts," in *Encyclopedia of ocean sciences*, *3rd3rd* (Silver Spring, MD, United States: Elsevier Ltd). doi: 10.1016/B978-0-12-409548-9.10816-4

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348 (6237), 1–10. doi: 10.1126/science.1261359

Sunagawa, S., Acinas, S. G., Bork, P., and Bowler, C. (2020). Tara Oceans: towards global ocean eco-systems biology. *Nat. Rev. Microbiol.* 18 (8), 428–445 doi: 10.1038/s41579-020-0364-5

Tornero, V., and Hanke, G. (2016). Chemical contaminants entering the marine environment from sea-based sources: A review with a focus on European seas. *Mar. pollut. Bull.* 112 (1–2), 17–38. doi: 10.1016/j.marpolbul.2016.06.091

Urban, L., Holzer, A., Baronas, J. J., Hall, M. B., Braeuninger-Weimer, P., Scherm, M. J., et al. (2020). Freshwater monitoring by nanopore sequencing. *bioRxiv*. doi: 10.1101/2020.02.06.936302

Vazquez-Cuervo, J., Gomez-Valdes, J., Bouali, M., Miranda, L. E., Van der Stocken, T., Tang, W., et al. (2019). Using saildrones to validate satellite-derived sea surface salinity and sea surface temperature along the California/Baja coast. *Remote Sens.* 11 (17), 1–18. doi: 10.3390/rs11171964

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., et al. (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 304 (5667), 66–74. doi: 10.1126/science.1093857

von Ammon, U., Jeffs, A., Zaiko, A., van der Reis, A., Goodwin, D., Beckley, L. E., et al. (2020). A portable cruising speed net: Expanding global collection of Sea surface plankton data. *Front. Mar. Sci.* 7 (December). doi: 10.3389/fmars.2020.615458

Ward, B. A., Dutkiewicz, S., and Follows, M. J. (2014). Modelling spatial and temporal patterns in size-structured marine plankton communities: Top-down and bottom-up controls. *J. Plankton Res.* 36 (1), 31–47. doi: 10.1093/plankt/fbt097

Willig, M. R., Kaufman, D. M., and Stevens, R. D. (2003). Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annu. Rev. Ecology Evolution Systematics* 34 (1), 273–309. doi: 10.1146/annurev.ecolsys.34.012103.144032

Yang, Y., Cowen, L. L. E., and Costa, M. (2018). Is ocean reflectance acquired by citizen scientists robust for science applications? *Remote Sens.* 10 (6), 835. doi: 10.3390/rs10060835

Zayed, A. A., Wainaina, J. M., Dominguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., et al. (2022). Cryptic and abundant marine viruses at the evolutionary origins of earth's RNA virome. *Science* 376:156–162. doi: 10.1126/science.abm5847

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fmars.2022.867695

OPEN ACCESS

EDITED BY
Fabien Lombard,
Sorbonne Universités, France

REVIEWED BY
Eric Orenstein,
Monterey Bay Aquarium Research Institute (MBARI), United States
Jean-Olivier Irisson,
Sorbonne Universités, France

*CORRESPONDENCE
Kaisa Kraft
kaisa.kraft@syke.fi

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 01 February 2022
ACCEPTED 05 August 2022
PUBLISHED 02 September 2022

CITATION
Kraft K, Velhonoja O, Eerola T, Suikkanen S, Tamminen T, Haraguchi L, Ylöstalo P, Kielosto S, Johansson M, Lensu L, Kälviäinen H, Haario H and Seppälä J (2022) Towards operational phytoplankton recognition with automated high-throughput imaging, near-real-time data processing, and convolutional neural networks. *Front. Mar. Sci.* 9:867695. doi: 10.3389/fmars.2022.867695

Towards operational phytoplankton recognition with automated high-throughput imaging, near-real-time data processing, and convolutional neural networks

Kaisa Kraft[1]*, Otso Velhonoja[1], Tuomas Eerola[2], Sanna Suikkanen[1], Timo Tamminen[1], Lumi Haraguchi[1], Pasi Ylöstalo[1], Sami Kielosto[1], Milla Johansson[3], Lasse Lensu[2], Heikki Kälviäinen[2], Heikki Haario[2] and Jukka Seppälä[1]

[1]Finnish Environment Institute, Marine Research Centre, Helsinki, Finland, [2]Computer Vision and Pattern Recognition Laboratory, School of Engineering Science, Lappeenranta-Lahti University of Technology LUT, Lappeenranta, Finland, [3]Finnish Meteorological Institute, Helsinki, Finland

Plankton communities form the basis of aquatic ecosystems and elucidating their role in increasingly important environmental issues is a persistent research question. Recent technological advances in automated microscopic imaging, together with cloud platforms for high-performance computing, have created possibilities for collecting and processing detailed high-frequency data on planktonic communities, opening new horizons for testing core hypotheses in aquatic ecosystems. Analyzing continuous streams of big data calls for development and deployment of novel computer vision and machine learning systems. The implementation of these analysis systems is not always straightforward with regards to operationality, and issues regarding data flows, computing and data treatment need to be considered. We created a data pipeline for automated near-real-time classification of phytoplankton during remote deployment of imaging flow cytometer (Imaging FlowCytobot, IFCB). Convolutional neural network (CNN) is used to classify continuous imaging data with probability thresholds used to filter out images not belonging to our existing classes. The automated data flow and classification system were used to monitor dominating species of filamentous cyanobacteria on the coast of Finland during summer 2021. We demonstrate that good phytoplankton recognition can be achieved with transfer learning utilizing a relatively shallow, publicly available, pre-trained CNN model and fine-tuning it with community-specific phytoplankton images (overall F1-score of 0.95 for test set of our labeled image data complemented with a 50% unclassifiable image portion). This enables both fast training and low computing resource requirements for model deployment making it easy to modify and applicable in wide range of situations. The system performed well when used to classify a natural phytoplankton community over different seasons (overall F1-score 0.82

for our evaluation data set). Furthermore, we address the key challenges of image classification for varying planktonic communities and analyze the practical implications of confused classes. We published our labeled image data set of Baltic Sea phytoplankton community for the training of image recognition models (~63000 images in 50 classes) to accelerate implementation of imaging systems for other brackish and freshwater communities. Our evaluation data set, 59 fully annotated samples of natural communities throughout an annual cycle, is also available for model testing purposes (~150000 images).

KEYWORDS

IFCB, near-real-time classification, phytoplankton imaging, automated data processing, imaging flow cytometry (IFC), convolutional neural network, CNN, operational observations

1 Introduction

The role of oceans and coastal seas in the global climate is well recognized, with phytoplankton playing a key role in organic carbon fluxes (Moigne 2019). At the same time, changes in the marine environment related to climate change affect the abundance and diversity of phytoplankton (Hutchins and Fu, 2017; Righetti et al., 2019), which is also likely to affect ecosystem functioning. Phytoplankton communities consist of hundreds of species of microorganisms with generation times in the order of hours (Reynolds, 2006). As phytoplankton community dynamics reflect changes in environmental forcing, growth traits of competing species and multiple food web interactions, a high-frequency characterization of those communities is required to improve both ecological studies and monitoring.

To follow and understand these changes at appropriate spatial and temporal scales, and to provide data for ecosystem modeling in predicting future responses, sustained observations of phytoplankton diversity are required. Traditional methods of phytoplankton community research using light microscopy results in a bottleneck, due to the constraints of acquiring community composition information on these small organisms, which require laborious sample preparation and microscopic identification. Recent frameworks for Essential Ocean Variables and Essential Biodiversity Variables underline the need to develop and improve automated observing technologies for phytoplankton, combined with open solutions for data handling (Miloslavich et al., 2018; Muller-Karger et al., 2018).

Recent technological advances have led to the emergence of automated and semi-automated imaging instruments for plankton studies, with steadily improving image resolution and output rates. One of the most promising methods for observing nano to mesoscale aquatic organisms is imaging flow cytometry. The Imaging FlowCytobot (IFCB) (Olson and Sosik, 2007) is among the most frequently used imaging flow cytometers for phytoplankton (covering a size range of approximately 10 to 150 µm Equivalent Spherical Diameter, or ESD) and its usefulness in phytoplankton ecology has been demonstrated by several studies (e.g. Laney and Sosik, 2014; Harred and Campbell, 2014; Anglès et al., 2019; Fischer et al., 2020). It has also been popular in Harmful Algal Bloom (HAB) studies or as an early warning detection of rare but toxic species (Campbell et al., 2010; 2013, Harred and Campbell, 2014; Henrichs et al., 2021; Kraft et al., 2021). The IFCB can produce up to tens of thousands of images per hour (Olson and Sosik, 2007), yielding real-time big data. The use of this type of new instrument opens new horizons for exploring planktonic systems (Lombard et al., 2019).

However, this creates a new bottleneck as it is impossible for a human to screen millions of images. Analyzing this big data calls for computer vision and machine learning methods capable of producing interoperable data across platforms and systems. As reviewed by Irisson et al. (2022) automatic plankton image classification traditionally starts with the extraction of manually engineered image features which are then used to train a classifier, typically either a Support Vector Machine (SVM) (Cortes and Vapnik, 1995) or a Random Forest (RF) (Breiman, 2001). The main problem with this approach is in finding image features which are both general and provide good delineation between the classes. Recent progress in both computer vision techniques and computing resources has made it possible to learn relevant image features directly from the images themselves, through deep learning (LeCun et al., 2015). Recent papers using deep learning techniques for plankton identification, especially Convolutional Neural Networks (CNNs), have shown them to be an attractive choice

for automation of the process (e.g. Luo et al., 2018; Dunker et al., 2018; Lumini and Nanni, 2019; Kerr et al., 2020; Lumini et al., 2020; Guo et al., 2021; Henrichs et al., 2021).

Transfer learning i.e., pre-training the model with a large data set of generic images and fine-tuning it to the target data set, is a common method used with CNNs. Thus, multiple platforms distribute pre-trained generic CNN models. Consequently, choices for CNN architectures and training procedures are numerous (Lumini and Nanni, 2019). However, applying CNN techniques to plankton image recognition is not straightforward due to the differing distribution of the training and target data and the multitude of CNN architectures to choose from. In addition, CNN-based methods are usually trained using data sets with hundreds, or even thousands of example images per class, which is often difficult to obtain in practice, especially in new locations (Dai et al., 2017). Furthermore, "data set shift" (i.e. the change in distribution of data across classes between the training data set and reality) is an important issue when deploying machine learning models (Moreno-Torres et al., 2012). Data set shift is highly relevant to plankton applications due to factors such as seasonal changes in community composition. This underlines the importance of assembling a diverse training data set, over time and space (González et al., 2017). Additionally, the size of CNN models becomes an important factor in moving towards automated/semi-automated plankton classification for real-time observations, determining the computational capacity needed.

High-throughput imaging coupled with efficient deep learning techniques will be one of the key game changers in the ecological research of phytoplankton. As with other branches of science using big data, the key challenges in plankton imaging are validation of data quality, integration of different data sources, defining common vocabularies of metadata and sharing of data and technology solutions to create reliable, acceptable and timely products (Muller-Karger et al., 2018; Lombard et al., 2019). In their review, Lombard et al. (2019) list a set of challenges and priorities for emerging phytoplankton detection technologies. One of their main recommendations is collaboration between experts and exchange with other disciplines, such as modeling. Phytoplankton imaging is also recognized as one of the main emerging technologies of coastal observation research systems for the provision of data to various stakeholders (Farcy et al., 2019). This study helps solve some of these technological challenges and improve the applicability of phytoplankton image recognition systems.

Our aim in this study is to address some of the fundamental challenges in the implementation of automated/semi-automated phytoplankton classification for real-time plankton image observations, using the Baltic Sea phytoplankton community as an example. This environment is one of Earth's largest brackish water habitats, with an especially challenging mix of phytoplankton species of both freshwater and marine origins, including many small-sized species (Olli et al., 2019). Such a data set, collected from a completely new type of habitat, has a different species composition to those from previous studies on plankton image classification. This poses a challenge to the implementation of automated recognition systems.

We created a data pipeline that allows near-real-time automated classification of individual plankton organisms using a CNN, throughout a remote deployment of an IFCB. We demonstrate its operationality by monitoring the filamentous cyanobacteria of the Baltic Sea, which are an important phytoplankton group due to their harmful summer blooms. We used a relatively shallow, openly available, pre-trained CNN model and fine-tuned it to plankton images from brackish waters. We demonstrate that, through this simple transfer-learning approach, one can achieve good classification accuracy. This makes our approach applicable to a wide range of users with low resources for model deployment. We further address the practical implications of the classifier performance by discussing the highest confusions among the classes.

2 Materials and methods

2.1 Sampling system

The IFCB (McLane Research Laboratories, Inc., U.S.) is an *in situ* automated submersible imaging-in-flow cytometer developed to image planktonic organisms (Olson and Sosik, 2007). The instrument can be used with either scatter or chlorophyll *a* fluorescence as a trigger, the latter being used more often for phytoplankton detection. Sheath fluid is used to force the particles to flow through the middle of the flow cell, improving the focus of the images and enabling excellent quality. The instrument has an image resolution of roughly 3.5 pixels per µm. According to the manufacturer, it captures images of suspended particles in the range of 10 to 150 µm (ESD), but in practice particles ranging from ~5 µm ESD to filaments ~300 µm in length have been captured (Kraft et al., 2021). The limiting factors are camera resolution, to get identifiable images, on the lower end and a 150-µm mesh at the instrument inlet that prevents it from clogging, together with the size of the field of view. However, the size range needed for quantitative observations is likely restricted to 10 to 80 µm (ESD) (Lombard et al., 2019). The IFCB processes a 5-mL sample every ~20 minutes, collecting up to ~30 000 images per hour.

The Marine Research Centre of the Finnish Environment Institute (SYKE) has had an IFCB deployed at the Utö Atmospheric and Marine Research Station (59°46.84' N, 21° 22.13' E) sporadically since 2017, now deployed continuously since early 2020 (see the detailed description of the station in Laakso et al., 2018 and the deployment setup in Kraft et al., 2021). Water is pumped continuously for the station's flow-through measurements, from 250 m offshore, with an

underwater pump (Grundfos SP3A-9N) through a 50-mm black PE tube lying at the sea bottom, at a depth of 23 m. The inlet for water sampling is located at a depth of ~5 m, representing the near-surface layer. The time it takes for the water to reach the cabin is approximately 5-6 minutes. Water is distributed through several flow-through sensors (including the IFCB), after reaching the inside of the station building (Laakso et al., 2018). The IFCB is currently operated with a chlorophyll *a* trigger to prevent the imaging of detritus and other non-living material.

2.2 Labeled image data sets

2.2.1 Image data set for model training and testing

To implement an automated image recognition system for Baltic Sea phytoplankton, a labeled image data set is required for training a classifier and testing its performance. Our labeled image data set, referred to as SYKE-plankton_IFCB_2022, consists of approximately 63 000 images belonging to 50 different phytoplankton taxa, defined, identified and verified by expert taxonomists (Figure 1). Due to differences in the features of the organisms visible in the images, which form the basis of the identification, some classes have been determined to the species level while others have been determined at a higher taxonomic level. The 50 classes represent the most common phytoplankton species/groups present in the Gulf of Finland and the Northern Baltic Proper. The taxonomy follows the Checklist of Baltic Sea Phytoplankton Species (Hällfors, 2004) and the nomenclature of the World Register of Marine Species (WoRMS Editorial Board, 2021). The data set SYKE-plankton_IFCB_2022 is publicly available at: http://doi.org/10.23728/b2share.abf913e5a6ad47e6baa273ae0ed6617a.

The SYKE-plankton_IFCB_2022 data set was collected in the Baltic Sea on different occasions, to cover spatio-temporal variations in plankton communities. In 2016 and 2019 water samples (n=52) were collected using the Alg@line ferrybox systems of M/S Finnmaid and Silja Serenade (Ruokanen et al., 2003; Kaitala et al., 2014) and analyzed in the laboratory with the IFCB. In 2017 and 2018 data were collected at the Utö station over the deployment periods, with the continuous set up of the IFCB, followed by the sporadic selection and labeling of a set of samples (n=62) (Figure 2).

Images of natural phytoplankton communities reflect their wide morphological diversity, resulting in large variations in size and aspect ratios of the images, with images ranging from tens to hundreds of pixels vertically and tens to more than one thousand pixels horizontally. The samples were labeled using a tool created by Sosik et al. (https://github.com/hsosik/ifcb-analysis/wiki/Instructions-for-manual-annotation-of-images). Some samples were labeled so that all identifiable regions of interest (ROIs) were assigned to a class and some samples were labeled only partially, to expand the labeled sets of some classes. The data set, therefore, does not represent real-life proportions among classes, however, the number of images per class still reflects, to some extent, their prevalence in natural populations.

The IFCB produces a non-negligible amount of images that are difficult or even impossible to identify with certainty. To train a CNN model, only images that can be reliably labeled should be used, to avoid mislabeled images which would negatively affect the training process. However, for testing the performance of the model, the unidentifiable part of the samples should be considered for calculating realistic performance metrics. The SYKE-plankton_IFCB_2022 data set was divided using stratified sampling, into training, validation and test sets (60%, 20% and 20% respectively). The training set, referred to here as Training Data, was used exclusively for training the model. The validation set had two purposes. First, it was used to monitor the model's accuracy during training, which is what it is usually used for, and in this sense is referred to as a validation set. After the model training was complete, the validation set was complemented with an equal number of unclassifiable images (50-50%) to make it more representative of image data from natural phytoplankton communities (including detritus and other unidentifiable images). The validation set complemented by the unidentifiable images (Validation Data) was used to determine class-specific thresholds which will be explained in section 2.3.1 Probability filtering of unclassifiable images using thresholds. For the same reason as with the Validation Data, the test set was similarly complemented with equal numbers of unclassifiable images (50-50%). The test set with unclassifiable images (Test Data) was used to calculate the final, unbiased estimation of the model's performance. The difference between dominant and rare taxa in the SYKE-plankton_IFCB_2022 data set manifests itself as a large imbalance in the number of images per class: it varies from 19 (*Amylax triacantha*) to 12 280 (*Dolichospermum* sp./*Anabaenopsis* sp.).

2.2.2 Image data set for performance evaluation

As previously explained, correct evaluation of model performance in classifying natural samples requires test data to contain difficult-to-classify images. That is, their features fit several different classes (so-called borderline images), as is the case with multiple images in natural samples. For this purpose, 59 samples were selected from 2021 when the IFCB was continuously deployed in Utö; first as one per week throughout the year complemented with samples from specific seasons to target specific classes. All samples were manually labeled in their entirety so that each image was assigned to one of the 50 classes or as "unclassified" (unable to be assigned to any of the existing classes). The image labeling was done with a custom graphical tool in a Jupyter Notebook, utilizing

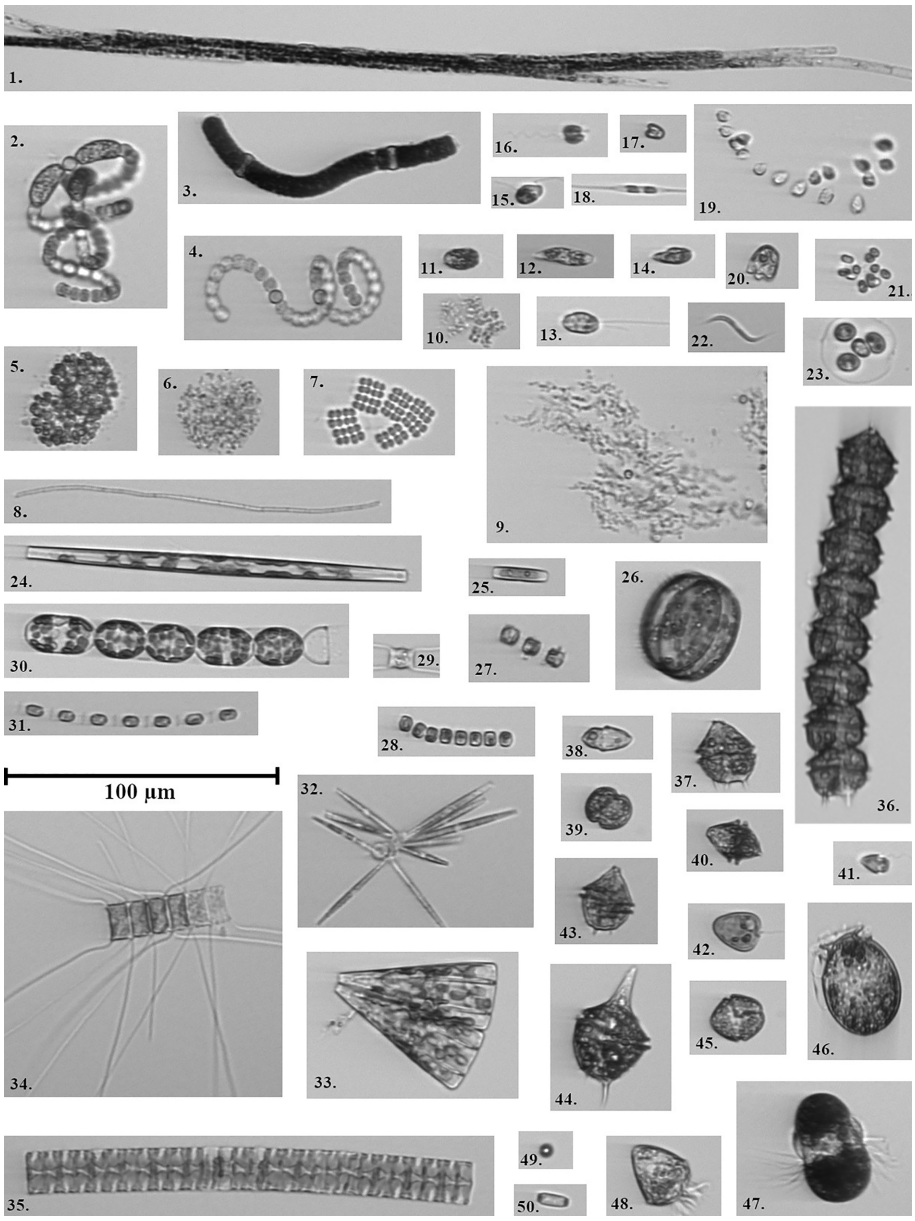


FIGURE 1
Example images representing the classes. 1) *Aphanizomenon flosaquae*, 2) *Dolichospermum* sp./*Anabaenopsis* sp. coiled, 3) *Nodularia spumigena*, 4) *Dolichospermum* sp./*Anabaenopsis* sp., 5) *Snowella* sp./*Woronichinia* sp., 6) Chroococcales, 7) *Merismopedia* sp., 8) Oscillatoriales, 9) *Aphanothece paralleliformis*, 10) *Chroococcus* sp., 11) *Eutreptiella* sp., 12) Euglenophyceae, 13) Cryptomonadales, 14) Cryptophyceae/*Teleaulax* sp., 15) *Katablepharis remigera*, 16) *Pseudopedinella* sp., 17) *Pyramimonas* sp., 18) *Ceratoneis closterium*, 19) *Uroglenopsis* sp., 20) *Cymbomonas tetramitiformis*, 21) Chlorococcales, 22) *Monoraphidium contortum*, 23) *Oocystis* sp., 24) Pennales thin, 25) Pennales thick, 26) Centrales, 27) *Thalassiosira levanderi*, 28) *Cyclotella choctawhatcheeana*, 29) *Chaetoceros* sp. single, 30) *Melosira arctica*, 31) *Skeletonema marinoi*, 32) *Nitzschia paleacea*, 33) *Licmophora* sp., 34) *Chaetoceros* sp., 35) *Pauliella taeniata*, 36) *Peridiniella catenata* chain, 37) *Peridiniella catenata* single, 38) Gymnodiniales, 39) *Gymnodinium* like cells, 40) *Heterocapsa triquetra*, 41) *Heterocapsa rotundata*, 42) *Prorocentrum cordatum*, 43) *Gonyaulax verior*, 44) *Amylax triacantha*, 45) Dinophyceae, 46) *Dinophysis acuminata*, 47) *Mesodinium rubrum*, 48) Ciliata, 49) Beads, 50) Heterocyte.

the model's predictions to speed up the process. This data set, SYKE-plankton_IFCB_Utö_2021 (Evaluation Data), is publicly available at http://doi.org/10.23728/b2share.7c273b6f409c47e98a868d6517be3ae3.

2.3 The CNN model

The neural network model used in this study is based on a pre-trained ResNet-18 (He et al., 2016) and fine-tuned with the

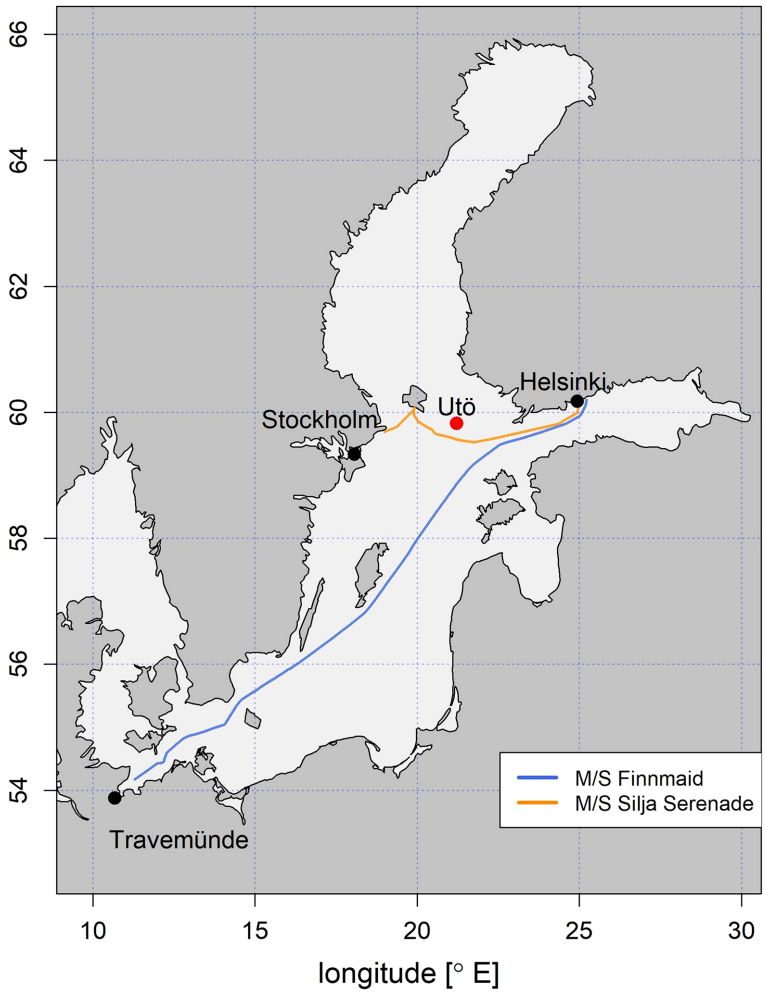


FIGURE 2
Map showing the location of Utö Atmospheric and Marine Research Station, and the points along the Alg@line routes of M/S Finnmaid and M/S Silja Serenade, from where the manually annotated samples were collected.

SYKE-plankton_IFCB_2022 Baltic Sea phytoplankton image data set, described above. The pre-trained model was obtained from TorchVision, which is part of the PyTorch project (Paszke et al., 2019). TorchVision models are pre-trained on the ImageNet data set (Deng et al., 2009), which consists of RGB images of 1000 classes, such as *fire truck* and *Golden Retriever*. The head of the pre-trained ResNet-18, i.e., the last fully connected linear layer, was replaced with three new linear layers, while the rest of the network layers were only fine-tuned. The new layers were initialized with the default PyTorch initialization for linear layers, which was in more detail a uniform distribution between $-\sqrt{k}$ and $\sqrt{k}$, where $k = 1/\text{in\_features}$.

To improve the performance of unseen images (not present in the Training Data), avoid overfitting and reduce class imbalance, first, random oversampling was done for the smaller classes in the Training Data so that each class contained a minimum of 100 training images. Secondly, some simple image augmentations were used for all classes in the Training Data (including the resampled images): horizontal and vertical flip, translation, zoom and brightness change. However, all augmentations were done sparingly since images generated by the IFCB are quite homogeneous. More specifically, translation was done only on the shorter side of the original image and none of the original pixels were clipped, the zoom range was 0.6 to 1.4, the rotation range was -10 to 10 degrees, and the range of brightness change was 0.95 to 1.1. Another approach to address class imbalance would be to provide class-specific weights to the optimizer (see e.g. cost-sensitive learning, Thai-Nghe et al., 2010). However, to avoid reducing the generalizability of the model to data sets with different class proportions, this method was not used.

Images were resized to 180×180 pixels since 180 is the mean width of the training images. To preserve the original image aspect ratio, the mode (i.e., most frequent) pixel value of each image was used as padding in its resized form. Pixel values were scaled between 0 and 1 for each image, a process known as min-max normalization, to avoid any training overhead caused by unnecessarily large integer values. Although the IFCB images are grayscale, because the original ResNet was trained with RGB images, three color channels were used.

A categorical cross-entropy loss was used with Adam as the optimizer function (Kingma and Ba, 2014). To further improve model training, a custom learning rate schedule was used. This schedule consisted of three steps. At each step, the number of trainable layers was increased, and the learning rate was decreased. Step 1 lasted for 5 epochs, where only the last linear layers were trained with a learning rate of 0.01. Step 2 lasted for the next 10 epochs, where the training of the last convolutional layer was started with a learning rate of 0.001, and the learning rate of the linear layers was decreased to 0.005. Step 3 lasted from epoch 16 onward, where the rest of the base layers were trained with a 0.0001 learning rate, the last base layer was trained with 0.001, and the head layers were trained with 0.0025. The training was stopped when the loss value on the validation set did not decrease for 12 epochs. The average training time was one hour on a single NVIDIA Tesla P100 GPU.

Class-specific recall, precision and F1-score were calculated for the classification results to describe the class-specific performance of the model. The weighted average F1-score was calculated to describe the entire model performance since global accuracy is a flawed metric for class-imbalanced data (Hossin and Sulaiman, 2015). Weighted average F1-score was chosen since we were evaluating the classification model from an operational point of view, in which case, the common classes and therefore more abundant ones should be given more weight. The computation involves True positive (TP), False positive (FP) and False negative (FN) numbers. Recall quantifies how well classes are identified and is computed as the proportion of successful identifications in a class. Precision quantifies how well other classes are rejected and is computed as the proportion of positive identifications that were correct. F1-score expresses the balance between recall and precision.

$$\text{Recall} = (\text{TP}/(\text{TP} + \text{FN}))$$

$$\text{Precision} = (\text{TP}/(\text{TP} + \text{FP}))$$

$$\text{F1} - \text{score} = (2*((\text{precision}*\text{recall})/(\text{precision} + \text{recall})))$$

2.3.1 Probability filtering of unclassifiable images using thresholds

As explained before, not all images captured by IFCB are classifiable due to a lack of characteristic features for example. We chose not to create classes for those. Therefore, a filtering method was needed to remove those images when the CNN was deployed. For each image, a classifier produces prediction scores for all classes in the training data. Prediction scores can be considered the probability of correctly classifying an image and the highest prediction score represents the winning class. By assigning a threshold, which the prediction score must exceed, they can then be used to filter out images with low classification probabilities. The threshold is not universal but class-specific. Therefore a unique probability threshold was estimated for each class, and only images with at least one class probability above any assigned thresholds were assigned to a class. Filtering the data is a proven method to treat low probability classifications (Faillettaz et al., 2016; Luo et al., 2018).

The final layer in our CNN model uses a softmax activation function, which outputs a normalized probability distribution over the classes. Since the probability distribution coming from a softmax can be quite extreme, i.e., one class has most of the probability mass, the outputs from the layer before the softmax were scaled down. Scaling was done by multiplication by the natural logarithm of 1.3. This has the same effect as changing the base of the exponents in the softmax function from e to 1.3, however, it is easier to scale the outputs rather than modify the softmax itself. The value of 1.3 was determined by manually testing different values. In short, the conversion is: softmax (out×ln(1.3)), where out = the outputs of the layer before softmax. This conversion introduced more smoothness in the class probabilities while maintaining their order (and therefore the classification). Smoothness made it easier to set class probability thresholds. A figure illustrating the effect of scaling on selected ROIs can be found in the supplementary material (Supplementary 3).

Ideally, thresholds would be assigned with a data set representing a species distribution similar to that of a natural community. However, the community composition changes with the seasons and species dynamics differ from year to year. Therefore, acquiring an ideal data set for threshold determination, which represents natural distribution covering all common species, is laborious. To start the implementation of the classifier we used the Validation Data to determine thresholds. The Validation Data was run through the classifier and precision, recall and F1-score were calculated. The threshold was varied and the value yielding the highest F1-score was chosen for a given class. The chosen thresholds were tested by running the Test Data through the classifier, and precision, recall and F1-score were calculated. Images below the thresholds were still considered when calculating performance metrics: correct images of a class which reached the assigned threshold limit were considered as TP and incorrect images FP; correct images below the threshold limit were considered as FN and incorrect images TN.

The code implementing the model described above can be found at https://github.com/veot/syke-pic.

2.4 Data transfer and services

First, the IFCB data is stored on the instrument's hard drive. Then the IFCB is connected to the Utö station's inner network, through which the data flows to the Finnish Meteorological Institute (FMI) file server *via* optical fibre, where it is temporarily stored. From there onwards it is transferred to a cloud object storage service Allas, provided by the Finnish IT Center for Science (CSC). Allas is based on CEPH object storage technology, allowing to easily share data to other services within the CSC's computing platform - much like Amazon S3 in AWS. The subsequent data analysis (described below) is done on a Linux virtual machine with 6 vCPUs and 16 GB of memory (the number of resources required for computing also image biovolumes, described below), also provided by CSC. Significantly fewer resources are needed when running the CNN-classifier alone.

2.5 Near-real-time data analysis

To use the generated IFCB images and the CNN classifier for near-real-time phytoplankton monitoring, a basic data pipeline was established. The near-real-time data pipeline and classification system were taken into use at the beginning of summer 2021. The entire data transfer pipeline results in a total delay of about two hours from the image capture to the point when the image is classified. The classification is performed automatically *via* the above-described CNN model as soon as a new batch of IFCB data is updated to Allas, on an hourly schedule, and the data is classified into the 50 classes. In addition to image classification also image-specific biovolumes are computed. A method developed by Moberg and Sosik (2012) is used for computing the biovolumes of the objects (phytoplankton) in the images taken with IFCB. More detailed descriptions of available MATLAB-based tools and open access codes can be found at https://github.com/hsosik/ifcb-analysis. For comparison of the biovolume estimates with those obtained *via* traditional phytoplankton monitoring methods of the Baltic Sea area (HELCOM, 2017), the biovolumes are converted to biomass (µg L^{-1}) assuming a plasma density of 1 g cm^{-3} (CEN, 2015). Finally, the biovolume/biomass information is combined with the classifications resulting in a usable form of class-specific biovolume/biomass per L in a sample and the hourly mean is calculated.

2.6 Evaluation of the near-real-time classifier system

To assess how well the model classified natural samples using the thresholds determined using the Validation Data, a total of 59 samples (a total of approx. 20 hours of data) were selected from data collected with IFCB between January to December 2021 (the Evaluation Data). First the selection targeted one sample per week, but due to the scarcity of some classes additional samples were selected from expected seasons to find images of the scarce classes. As proposed by González et al. (2017) for proper performance validation a set of samples should have sufficient variability. We attempted to ensure this by selecting samples from different seasons which also covered transition phases. Selected samples were manually inspected: all classifications were assessed (confirmed or corrected) and all identifiable images which fell below the thresholds were labeled. The unidentifiable images left without an assigned class were considered unclassified. Unclassified images are still accounted for in the total community biomass with the assumption that when chlorophyll *a* is used as a trigger the majority of imaged particles should be living material. The TP, FP and FN were counted and consequently precision, recall and F1-score were calculated for each class. Class-specific metrics were calculated based on the thresholds determined using the Validation Data, so images below the thresholds were still taken into account: correct images within a class to reach the assigned threshold were considered TP and incorrect images FP; correct images below the threshold were considered FN and incorrect images TN.

3 Results

3.1 CNN classifier performance

The first step in implementing a near-real-time analysis of plankton communities is to establish a suitable recognition model. Overall classification performance of the Test Data using CNN was high (F1-score 0.95), and the network was able to identify many common species of the Baltic Sea phytoplankton community. The class-specific precision, recall and F1-score were between 0.85 and 1 in over half of the classes, but some of the classes had much lower values (0.4-0.6) (Table 1).

With classes having the largest training sets (> 1000 images), all the metrics (precision, recall and F1-score) were between 0.94 - 1. With some classes such as Euglenophyceae, *Dinophysis acuminata, Peridiniella catenata* chain, *Ceratoneis closterium, Nitzschia paleacea, Monoraphidium contortum* and *Cymbomonas tetramitiformis* all the metrics were > 0.9, although the Training Data contained < 200 images. Classes with the poorest performance (all metrics < 0.7) were *Aphanothece paralleliformis, Pseudopedinella* sp., *Pyramimonas* sp., Chlorococcales and Beads (calibration). These all contained low numbers of training images (17 – 227, discounting data augmentation) except for *Pyramimonas* sp. (734 images). The largest differences between precision and recall were found with

TABLE 1 The table represents the class-specific classification metrics for the Test Data and for the Evaluation Data (Pr, precision; Re, Recall; F1, F1-score; N, number of images).

Class/taxonomic group	Training Data	Validation set	Test Data				Evaluation Data			
	N	Threshold	N	Pr	Re	F1	N	Pr	Re	F1
Cyanophyceae (0.94)										
Dolichospermum sp./*Anabaenopsis* sp.	7368	0.38	2456	0.98	0.99	0.98	790	0.88	0.96	0.92
Aphanizomenon flosaquae	4193	0.24	1398	0.97	1.00	0.98	1849	0.87	0.98	0.92
Oscillatoriales	2664	0.31	888	0.99	1.00	0.99	3893	0.98	0.98	0.98
Snowella sp./*Woronichinia* sp.	1770	0.63	590	0.99	0.97	0.98	42	0.64	0.69	0.67
Dolichospermum sp./*Anabaenopsis* sp. *coiled*	1502	0.41	501	0.93	0.96	0.95	70	0.74	0.99	0.85
Chroococcus sp.	496	0.61	166	0.90	0.94	0.92	2	ND	ND	ND
Nodularia spumigena	101	0.32	34	0.80	0.94	0.86	62	0.80	0.84	0.82
Chroococcales	85	0.73	29	0.75	0.93	0.83	793	1.00	0.51	0.68
Merismopedia sp.	59	0.63	19	0.79	0.79	0.79	2	ND	ND	ND
Aphanothece paralleliformis	17	0.80	6	0.57	0.67	0.62	9	ND	ND	ND
Cryptophyceae (0.89)										
Cryptophyceae/Teleaulax sp.	4098	0.53	1366	0.96	0.97	0.96	16952	0.97	0.90	0.93
Cryptomonadales	428	0.37	142	0.79	0.82	0.81	525	0.65	0.58	0.61
Euglenophyceae (0.76)										
Eutreptiella sp.	1348	0.43	450	0.95	0.94	0.94	1678	0.90	0.76	0.83
Euglenophyceae	61	0.24	21	0.90	0.90	0.90	18	0.28	0.39	0.33
Dinophyceae (0.75)										
Heterocapsa triquetra	1966	0.39	655	0.98	0.97	0.97	2267	0.91	0.95	0.93
Dinophyceae	860	0.40	286	0.88	0.94	0.91	1562	0.83	0.45	0.59
Peridiniella catenata single	539	0.52	180	0.89	0.97	0.93	222	0.75	0.81	0.78
Heterocapsa rotundata	368	0.56	123	0.84	0.90	0.87	2609	0.85	0.74	0.79
Prorocentrum cordatum	166	0.47	55	0.87	0.82	0.84	0	ND	ND	ND
Dinophysis acuminata	130	0.68	44	0.98	0.91	0.94	17	0.79	0.65	0.71
Peridiniella catenata chain	116	0.70	38	0.97	1.00	0.99	89	0.99	0.87	0.92
Gymnodinium like cells	95	0.44	31	0.76	0.52	0.62	102	0.59	0.25	0.36
Gymnodiniales	41	0.29	14	0.92	0.86	0.89	38	0.78	0.74	0.76
Gonyaulax verior	13	0.32	5	0.57	0.80	0.67	1	ND	ND	ND
Amylax triacantha	11	0.34	4	0.60	0.75	0.67	3	ND	ND	ND
Bacillariophyceae (0.86)										
Skeletonema marinoi	2477	0.46	825	1.00	0.99	0.99	7402	0.99	0.94	0.97
Thalassiosira levanderi	1522	0.63	508	0.95	0.95	0.95	2008	0.87	0.68	0.77
Chaetoceros sp. *chain*	829	0.51	277	0.93	0.95	0.94	693	0.76	0.77	0.76
Pennales thin	469	0.29	156	0.96	0.99	0.97	334	0.61	0.84	0.71
Centrales	288	0.51	96	0.98	0.89	0.93	92	0.77	0.68	0.72
Chaetoceros sp. *single*	128	0.12	42	0.85	0.98	0.91	571	0.75	0.60	0.67
Pennales thick	126	0.37	42	0.93	0.88	0.90	1088	0.72	0.85	0.78
Pauliella taeniata	71	0.62	24	1.00	0.96	0.98	56	0.96	0.86	0.91
Cyclotella choctawhatcheeana	61	0.47	21	0.89	0.81	0.85	199	0.92	0.57	0.71
Licmophora sp.	44	0.43	15	1.00	0.80	0.89	78	0.88	0.77	0.81
Nitzschia paleacea	39	0.40	13	0.92	0.92	0.92	4	ND	ND	ND
Ceratoneis closterium	27	0.41	9	1.00	1.00	1.00	75	0.68	0.91	0.78
Melosira arctica	26	0.30	8	0.73	1.00	0.84	58	0.85	0.91	0.88
Chrysophyceae (0.51)										
Uroglenopsis sp.	310	0.88	103	0.89	0.83	0.86	134	0.50	0.66	0.57
Pseudopedinella sp.	227	0.76	76	0.69	0.67	0.68	579	0.81	0.46	0.59

(Continued)

TABLE 1 Continued

Class/taxonomic group	Training Data	Validation set	Test Data				Evaluation Data			
	N	Threshold	N	Pr	Re	F1	N	Pr	Re	F1
Chlorophyta (0.34)										
Pyramimonas sp.	734	0.95	245	0.57	0.41	0.48	8422	0.88	0.32	0.47
Oocystis sp.	505	0.50	169	0.88	0.93	0.90	161	0.91	0.89	0.90
Monoraphidium contortum	196	0.69	66	0.98	0.98	0.98	439	0.99	0.96	0.97
Cymbomonas tetramitiformis	119	0.44	40	0.90	0.90	0.90	4	ND	ND	ND
Chlorococcales	57	0.48	19	0.43	0.47	0.45	45	0.81	0.29	0.43
Other										
Katablepharis remigera	32	0.36	11	0.83	0.45	0.59	4	ND	ND	ND
Ciliophora (0.76)										
Mesodinium rubrum	679	0.44	227	0.96	0.95	0.96	560	0.91	0.86	0.88
Ciliata	146	0.39	48	0.89	0.88	0.88	288	0.93	0.49	0.64
Additional classes										
Heterocyte	158	0.88	52	0.76	0.75	0.76	318	0.72	0.51	0.60
Beads	75	0.90	25	0.44	0.64	0.52	0	ND	ND	ND
Unclassifiable			12600				94028			

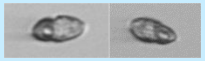
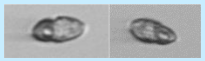
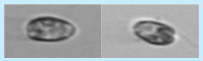
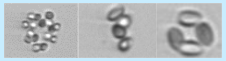
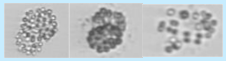
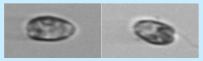
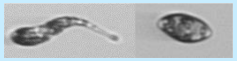
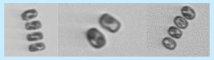
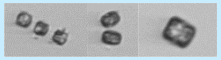
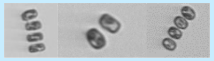
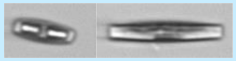
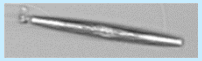
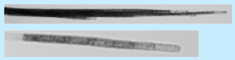
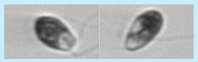
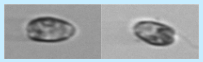
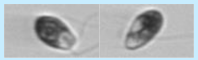
Classes are organized by the size of the training set to different taxonomic groups. N is the number of images per class in the Training Data, Test Data and Evaluation Data. Threshold is the class-specific threshold used for the classification, determined with the Validation Data. ND in the Evaluation Data means "Not Determined", the metrics were not calculated for classes with < 10 images. The group level results are presented in brackets after each taxonomic group (the percentage of how big portion of the images belonging to that group were correctly classified to the group).

the classes, *Gymnodinium* like cells (0.76, 0.52), *Gonyaulax verior* (0.57, 0.8), *Melosira arctica* (0.73, 1), *Katablepharis remigera* (0.83, 0.45) and Beads (0.44, 0.64). With *Gymnodinium* like cells and *Katablepharis remigera* precision was higher than recall, meaning rejection of images not belonging to the class was higher than recognition of images which did belong to the class. For the classes *Gonyaulax verior, Melosira arctica* and Beads, there was no issue in recognizing the images belonging to the class, however, there was the problem of a high proportion of false positives. Metrics for classes of filamentous cyanobacteria (an important group in the Baltic Sea) were all ≥ 0.93 except for the class, *Nodularia spumigena*, which had the poorest performance (0.8 – 0.94). It is important to note that *Nodularia spumigena*'s training set had a considerably smaller number of images (only 101, compared to *Aphanizomenon flosaquae*: 4193, *Dolichospermum* sp./*Anabaenopsis* sp.: 7368, *Dolichospermum* sp./*Anabaenopsis* sp. coiled: 1502) (Table 1).

When applying the classification system (the Evaluation Data) overall performance dropped, but remained fairly high for natural samples (F1-score 0.82). All class-specific classification metrics are presented in Table 1. For the classes with > 1000 images in the Training Data, the change in F1-score was ≤ 0.1 except for the classes, *Snowella* sp./*Woronichinia* sp., *Eutreptiella* sp. and *Thalassiosira levanderi*. For classes with < 200 images in the Training Data score decreased ≤ 0.1 for *Nodularia spumigena, Peridiniella catenata* chain, *Pauliella taeniata, Licmophora* sp., *Melosira arctica* and *Monoraphidium contortum*. The F1-score of *Melosira arctica* increased along with a larger number of images, on which the evaluation was based (8 in the Test Data and 58 in the Evaluation Data). The results for classes with < 10 images in their Evaluation Data will not be presented as these rare occurrences would not help in analyzing the model performance. However, classes with < 10 images in their Test Data were presented, as some classes had much higher number of images in the Evaluation Data (*Melosira arctica* and *Ceratoneis Closterium*). Other classes with > 10 images in the Evaluation Data, the poorest performance (metrics < 0.7) was found for classes *Snowella* sp./*Woronichinia* sp., Cryptomonadales, *Gymnodinium* like cells, *Uroglenopsis* sp. and Heterocyte. The recall and F1-score (0.29 – 0.59) were low for the classes *Pseudopedinella* sp., *Pyramimonas* sp. and Chlorococcales but precision was relatively high indicating a poor function in class recognition and that class-specific thresholds should be adjusted. Class recognition performance of filamentous cyanobacteria (*Aphanizomenon flosaquae*, *Dolichospermum* sp./*Anabaenopsis* sp., *Dolichospermum* sp./*Anabaenopsis* sp. coiled, *Nodularia spumigena)* was relatively high. This was also true for natural samples (F1-scores 0.82 – 0.92) (Table 1).

The majority of classification problems, of course, occurred between classes which resembled one another and were typically from closely related taxa. The highest confusion amongst CNN model results, when probability filtering thresholds were not used, were within different classes of dinoflagellates and between species-level classes and higher taxonomic-level classes belonging to the same order (Table 2). *Gymnodinium* like cells

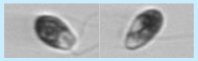
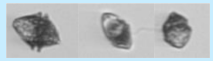
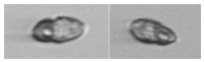
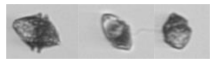
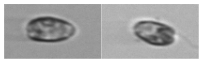
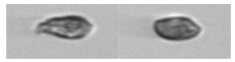
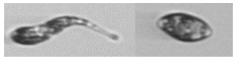
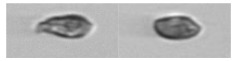
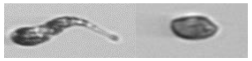
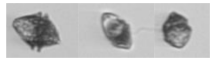
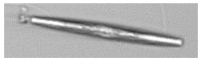
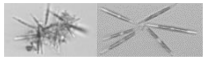
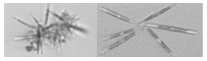
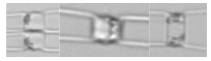
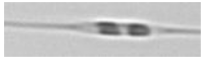
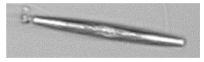
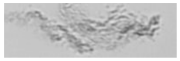
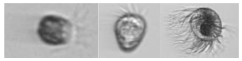
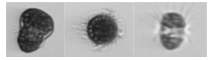
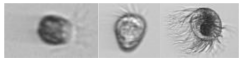
TABLE 2 Class pairs with the highest number of inter-class classification errors.

Group	Actual label A		A -> B (%)		Predicted label B	Group
Dino	*Gymnodinium* like cells		35 8		Dinophyceae	Dino
Dino	*Amylax triacantha**		25 33		*Peridiniella catenata* single	Dino
Dino	*Gonyaulax verior**		20		*Peridiniella catenata* single	DIno
Dino	Gymnodiniales		7		*Eutreptiella* sp.	Eugleno
Dino	Gymnodiniales		7		Cryptomonadales	Crypto
Chloro	Chlorococcales		26		*Oocystis* sp.	Chloro
Chloro	Chlorococcales		11		*Snowella* sp. / *Woronichinia* sp.	Cyano
Crypto	Cryptomonadales		10		*Eutreptiella* sp.	Eugleno
Eugleno	Euglenophyceae		10		*Eutreptiella* sp.	Eugleno
Diatom	*Cyclotella choctawhatcheeana*		14		*Thalassiosira levanderi*	Diatom
Diatom	*Cyclotella choctawhatcheeana*		5		*Chaetoceros* sp. chain	Diatom
Diatom	Pennales thick		5		Pennales thin	Diatom
Diatom	*Chaetoceros* sp. single		7		*Chaetoceros* sp. chain	Diatom
Cyano	*Nodularia spumigena*		6		*Aphanizomenon flosaquae*	Cyano
Other	*Katablepharis remigera*		9		Cryptomonadales	Crypto
Other	*Katablepharis remigera*		9		Dinophyceae	Dino

(Continued)

10.3389/fmars.2022.867695

TABLE 2 Continued

Group	Actual label A		A -> B (%)		Predicted label B	Group
Other	*Katablepharis remigera*		9		*Heterocapsa triquetra*	Dino
Dino	Gymnodiniales		5		*Heterocapsa triquetra*	Dino
Crypto	Cryptomonadales		5		*Cryptophyceae/ Teleaulax* sp.	Crypto
Eugleno	Euglenophyceae		6		*Cryptophyceae/ Teleaulax* sp.	Crypto
Eugleno	Euglenophyceae		6		*Heterocapsa triquetra*	Dino
Diatom	Pennales thin		8		*Aphanizomenon flosaquae*	Cyano
Diatom	*Nitzschia paleacea**		25		*Chaetoceros* sp. chain	Diatom
Diatom	*Nitzschia paleacea**		25		*Chaetoceros* sp. single	Diatom
Diatom	*Ceratoneis closterium*		7		Pennales thin	Diatom
Cyano	*Aphanothece paralelliformis**		22		*Aphanizomenon flosaquae*	Cyano
Ciliata	*Ciliata*		5		*Mesodinium rubrum*	Ciliata
Ciliata	Ciliata		6		*Chaetoceros* sp. chain	Diatom

The table contains each class (Actual label A) with higher than 5% confusion to another class, as well as the terminal classes (Predicted label B). The middle column contains the portions of images in class A that were incorrectly classified to class B. The results highlighted in blue are for the test set of SYKE-plankton_IFCB_2022 (without unclassifiable images and probability filtering with thresholds) and the rest of the results are for the Evaluation Data. The cells with two confusion values are: top: the test set, below: the Evaluation Data. The abbreviations of the phytoplankton groups are: Chloro, Chlorophyta; Crypto, Cryptophyceae; Cyano, Cyanophyceae; Eugleno, Euglenophyceae; Dino, Dinophyceae; Diatom, Bacillariophyceae; Ciliata, Ciliophora. The example images were randomly selected and not necessarily misclassified. Classes with * included <10 images.

10.3389/fmars.2022.867695

were confused (35%) for Dinophyceae and placed within a higher taxonomic branch, *Amylax triacantha* and *Gonyaulax verior* (with 4 and 5 images in the Test Data respectively) were confused (25% and 20% respectively) for *Peridiniella catenata* single (539 training images). Chlorococcales was also considerably confused for *Oocystis* sp. (26%). From filamentous cyanobacteria, only 6% of *Nodularia spumigena* was confused with *Aphanizomenon flosaquae* (Table 2). A confusion matrix with all confused classes is provided as supplementary material (Supplementary 1).

Confusion was lower among classes in the Evaluation Data when probability filtering with thresholds was applied. However, several images were left unclassified, as a delicate balance between TP and FN must be achieved for threshold assignation. A class-specific confusion matrix for the Evaluation Data, including those left unclassified, is provided as supplementary material (Supplementary 2). Similar, to the Test Data without filtering, the highest confusion among classes in the Evaluation Data was mainly between classes of close taxonomic relation. The highest confusion occurred (> 15%) between classes with < 10 images of data. Therefore, drawing any conclusion should be done very scarcely. What can be said reliably, is that classes with a small amount of training data are easily confused with classes similar in morphological appearance. Otherwise, the confusion rates were very moderate (5 – 8%) (Table 2).

When looking at confusion at the level of broader taxonomic groups there was practically no confusion between different groups (Figure 3). However, the proportion of images left unclassified due to probability filtering with thresholds varied greatly. Groups with the best identification rates, and with the least unclassified images, were Cyanophyceae, Cryptophyceae and Bacillariophyceae (6%, 11% and 13% respectively). For the classes, Euglenophyceae, Dinophyceae and Ciliophora, a reasonable proportion of images were left unclassified (22%, 24% and 19% respectively). Chrysophyceae and Chlorophyta had the highest proportion of unclassified images (49% and 65% respectively) (Figure 3).

3.2 Implementation of a near-real-time phytoplankton community information system

The operability and utility of the near-real-time data processing pipeline were used in the summer of 2021, as a demo, for up-to-date information on the abundance of the three bloom-forming cyanobacteria taxa of the Baltic Sea (with approx. 2h delay between sampling and online publication of classified results) (Figure 4). A simple visualization of the cyanobacteria situation was created and published online ("Cyanobacteria biomass" in https://swell.fmi.fi/hab-info/) to ensure public accessibility of the prevailing situation. The visualization shows a continuously updated graph containing information on the biomass of three main bloom-forming cyanobacteria taxa. This biomass graph was used, as an indicator of the predominant taxa off the coast of Finland, in SYKE's weekly cyanobacterial reports in summer 2021 (https://www.syke.fi/en-US/Current/Algal_reviews).

Dolichospermum sp./*Anabaenopsis* sp. started to bloom in late June, quickly achieving high biomass (peak in ~500 µg L^{-1} on July 2) followed by a quick drop within a few days. While *Dolichospermum* sp./*Anabaenopsis* sp. biomass was on the decline, *Aphanizomenon flosaquae* biomass started to increase, reaching its peak (~400 µg L^{-1}) within approximately five days (on July 5). *A. flosaquae* achieved a lower biomass peak but was spread over a longer period than that of *Dolichospermum* sp./*Anabaenopsis* sp. A secondary and smaller peak (~150 µg L^{-1}) appeared later in July (19^{th}/20^{th}) and was caused by both *Dolichospermum* sp./*Anabaenopsis* sp. and *A. flosaquae*. *A. flosaquae* formed a third peak at the end of July reaching a biomass of ~400 - 500 µg L^{-1}. *Nodularia spumigena* was detected, sporadically more abundant in some samples but did not exhibit a consistent biomass increase (Figure 5A). Although it was the cyanobacterial bloom season, the total filamentous cyanobacteria biomass constituted only approx. a third of the total phytoplankton community biomass. Simultaneously to the decline of the first cyanobacteria peak, the phytoplankton community's total biomass increased (Figure 5B). The third cyanobacteria peak achieved a similar magnitude as the second peak, but with differing community composition, demonstrating the importance of obtaining more detailed, higher-resolution information on community composition (Figures 5A, B).

4 Discussion

Recently plankton imaging systems have become numerous, diverse and widely used (Lombard et al., 2019). The classification of plankton images has become popular resulting in multiple publications and classification algorithms, often focusing on CNN applications (see e.g. Dunker et al., 2018; Luo et al., 2018; Lumini and Nanni, 2019; Kerr et al., 2020; Lumini et al., 2020; Guo et al., 2021; Henrichs et al., 2021; and the references therein). This popularity is due to the great need for efficient solutions for automated analysis and data flows of the vast amounts of image data produced, and underlines the importance of the quality of the data products (Muller-Karger et al., 2018).

However, papers on plankton classification often report only the classification performance and do not account for its practical implications on aquatic research, such as the effect of confused classes and classification processes from an operational point of view (e.g. Orenstein and Beijbom, 2017; Bureš et al., 2021). In transitioning to more frequent use of these new instruments it is important to focus on the steps needed for

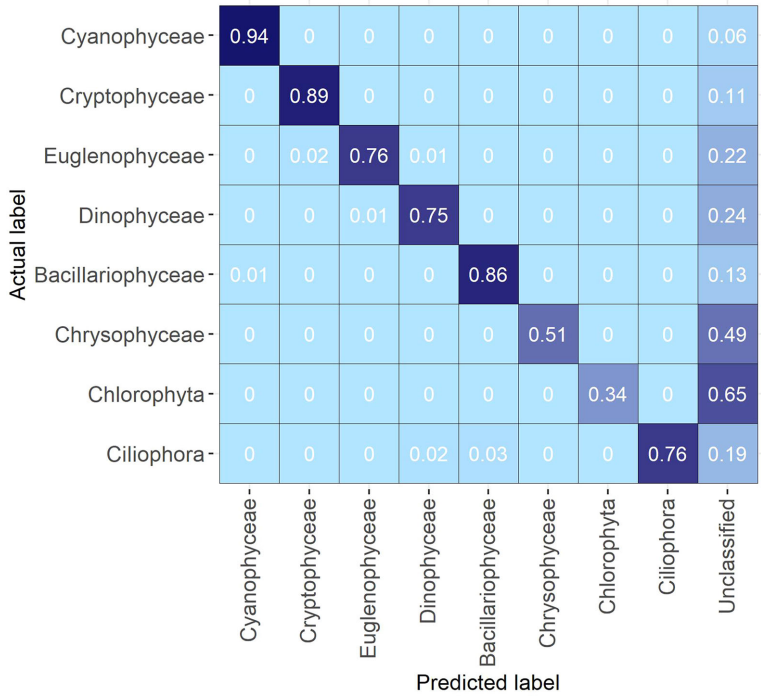


FIGURE 3
Confusion matrix of the Evaluation Data aggregated on broader taxonomic group level.

operationality and reference them with traditional light microscopy (Haraguchi et al., 2017; Kraft et al., 2021) as well as combining the two methods, as different methods confer different advantages. Between-sensor studies are scarce, nevertheless, in future, they will be sorely needed.

We demonstrated in this paper the functionality of CNN in classifying IFCB images from the Baltic Sea. Furthermore, we developed a framework for near-real-time image classification, which is sorely needed in HAB observations and also supports the future development of operational modelling and remote sensing applications. We provided a practical example with near-real-time observations of the summer cyanobacteria blooms, a reoccurring nuisance for users of the Baltic Sea. The species composition, timing and magnitude of the blooms are

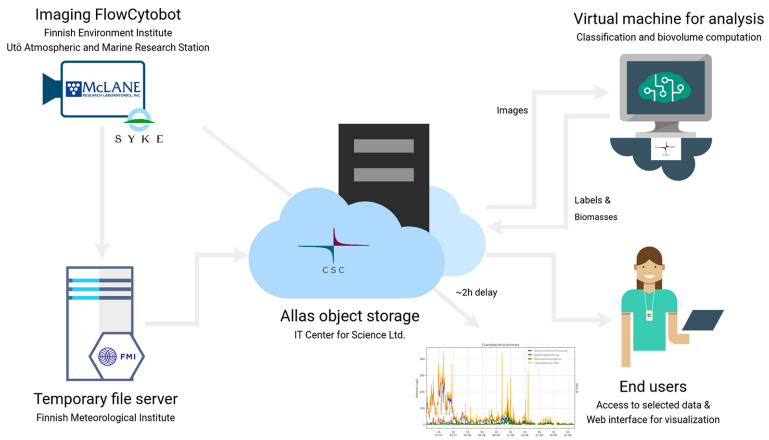


FIGURE 4
Scheme of the automated data flow and the subsequent data processing pipeline.

difficult to predict as controlling factors are still something of a puzzle (Kahru and Elmgren, 2014; Kownacka et al., 2018; Kahru et al., 2020). The summer of 2021, for example, was atypical (Figure 5) and the highest biomass peaks were only half that of the peaks recorded during an intensive bloom in 2018 (Kraft et al., 2021). The three major bloom-forming taxa in the Baltic Sea are *Nodularia spumigena, Aphanizomenon flosaquae* and *Dolichospermum* spp. (Niemistö et al., 1989; Stal et al., 2003; Olofsson et al., 2020), of which only *A. flosaquae* is not known to be toxic. Therefore, the separation of these three taxa in the Baltic Sea environment is highly important, which was already achieved (Table 1, Supplementary 2).

Though this paper is not about the study of ecological phenomena, two observations highlighting its future potential are worth mentioning. First, even during the filamentous cyanobacteria bloom peaks, their biomass was a third of the total phytoplankton biomass (Figure 5B). Second, while the total cyanobacteria biomass was of the same order of magnitude during the second and third peaks, the species composition differed, with the second peak consisting of approximately equal parts of two taxa, *Dolichospermum* sp./*Anabaenopsis* sp. and *Aphanizomenon flosaquae*, and the third peak almost solely of the latter. These observations, as well as the variability of the overall phytoplankton species composition, will be considered in a more ecologically focused follow-up study. However, it demonstrates that the utilization of new automated methods, such as imaging flow cytometry, plays a key role in deepening our understanding of these bloom processes (Kraft et al., 2021). Yet, these measurements must be made in conjunction with physical and biogeochemical observations, using the same observation platforms, such as the one at Utö (Laakso et al., 2018; Honkanen et al., 2021; Kraft et al., 2021).

Before digging into the ecology behind these phenomena, there are still a few practical aspects to be considered from an operational point of view. First is the classification model performance. The level of performance must be adequate to enable utilization of the results and verification of these results needs to be done for natural samples to ensure adequate performance during operational use. Second is the implications of confused classes. Some confusion doesn't mean the results are unusable, but a proper aggregation level needs to be selected, otherwise, the results of only certain classes that meet the user's criteria should be used. Third, some practical decisions on how to deal with the large number of difficult-to-assign images should be made. There is a lot of work to be done before the plankton classification problem is solved and data collected during the meantime needs to be harnessed while development continues.

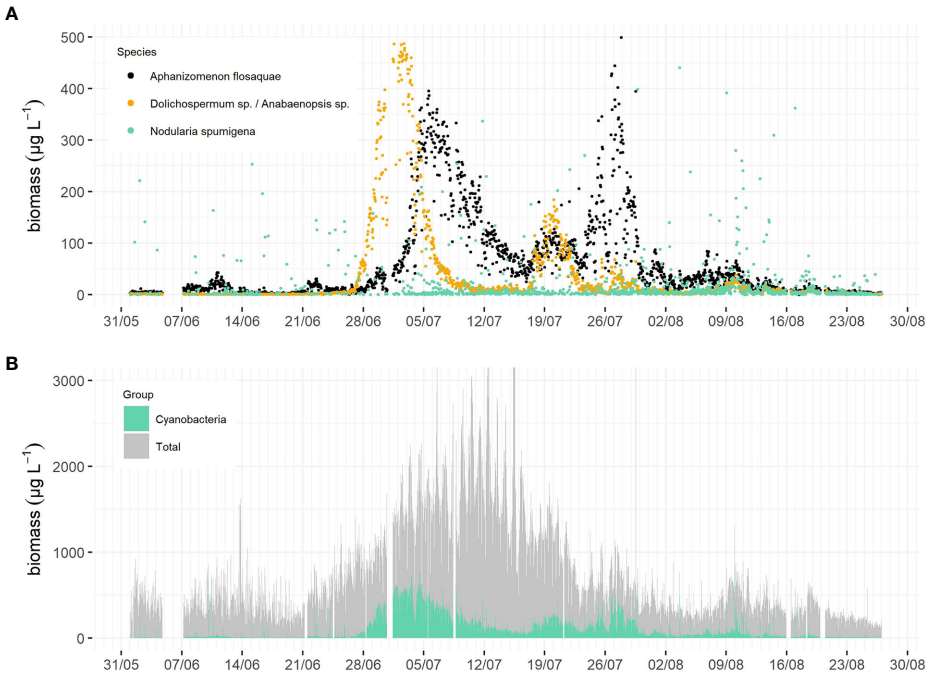


FIGURE 5
The biomasses of the three bloom forming filamentous cyanobacteria taxa of the Baltic Sea in summer 2021, black = *Aphanizomenon flosaquae*, orange = *Dolichospermum* sp./*Anabaenopsis* sp., green = *Nodularia spumigena* **(A)**. The total phytoplankton community biomass (grey) and the total filamentous cyanobacteria biomass (green) from the same period **(B)**. The data was classified with the automated CNN model in near-real time.

4.1 Classification model performance

Overall, the CNN model used for classification in this study performed very well, although there was some class-specific variation in the classification scores (Table 1). The overall F1-score for the Test Data was 0.95. This can be considered highly accurate and is on par with the results obtained in recent phytoplankton studies (Lumini and Nanni, 2019; Lumini et al., 2020; Walker and Orenstein, 2021). There was a drop in performance when using the classifier in operational mode (the overall F1-score for the Evaluation Data was 0.82). Recht et al. (2019) found that the accuracy of models tends to drop even when tested with data created to match the training data's distribution profile. This is due to human labeling subjectivity which makes it impossible to produce the same distribution. They concluded that the models are insufficient for generalization to more difficult images e.g. absence or deficiency of necessary features in the image. In our case, this drop may partly be due to the different distribution of training images compared to the target data and is partly explained by the inferior performance of some classes. Similar to the conclusions of Recht et al. (2019), the drop in our case is probably largely due to the large number of so-called borderline cases in natural samples, which make them difficult to classify. This makes the decision on where to draw the line difficult and leads to a high number of both FP and FN.

The drop, when applying the classifier during operational use, was in either precision or recall, or both. In many cases, the drop was larger with precision than recall meaning that the thresholds applied to those classes should be adjusted downward. Also, in many cases, the drop in recall was higher than that of precision meaning that thresholds for those classes should be adjusted upward. This proves the importance of threshold selection and although our results indicate that Validation Data is adequate for setting initial threshold values, it does not provide optimal thresholds as the borderline images are missing from the Validation Data. However, threshold adjustment is done cumulatively based on operational data such as is collected at Utö and they need to be adjusted and refined as data and information is accumulated. After adjusting the threshold values, it is not a heavy task to compute the previous time series using the new threshold values since the predictions have already been done.

In our study, we included some classes with only ~20-70 images per class and still reached relatively high classification metrics with the Test Data for some of those classes (Table 1). This is partly explained by the efficient data augmentation methods, as supported by the results of e.g. Correa et al. (2017) and partly by the unique morphology/appearance compared to other classes (Figure 1). However, those classes (except *Pauliella taeniata* and *Melosira arctica*) experienced a noticeable drop or didn't contain sufficient occurrences for the proper estimation of performance when classifying the Evaluation Data (Table 1). In the case of *Ceratoneis closterium, Licmophora* sp. and perhaps *Cyclotella choctawhatcheeana* this may also be a need for adjusting the class-specific thresholds since there was a relatively high difference between their precision and recall. Although it is possible to increase the training set of classes with low numbers of images using data augmentation, the classification results cannot be considered reliable when they are based on only few images e.g.< 10 images.

The classification score improved when more images were available in the Training Data, with a greater improvement for classes containing > 2000 images. All classes with large training sets had an F1-score > 0.9 for both the Test Data and the Evaluation Data. Although more images in the training set seem to be advantageous, there were also several classes (*Nodularia spumigena, Peridiniella catenata* chain, *Pauliella taeniata, Licmophora* sp., *Melosira arctica* and *Monoraphidium contortum*) with relatively good F1-scores (> 0.8) even with< 200 images in the training set. This would imply that distinguishable features (e.g. a specific shape) strongly influence the successful identification of specific classes amongst those with less labeled images (Figure 1, Table 1).

4.2 Class-specific confusion and its practical implications

Characteristic features of an organism tend to lead to a more accurate classification of images, however, many common Baltic Sea phytoplankton species, such as dinoflagellates, do not have obvious distinguishable features in their IFCB images which could be used to differentiate them. Consequently, those cases which were most confused were among classes of dinoflagellates (Table 2). Confusion within classes closely related taxonomically, such as classes on a higher and lower level of the same taxonomic hierarchy or different species belonging to the same order (*Oocystis* sp. and Chlorococcales, *Eutreptiella* and Euglenophyceae, *Mesodinium rubrum* and Ciliata, *Cyclotella choctawhatcheeana* and *Thalassiosira levanderi, Nodularia spumigena* and *Aphanizomenon flosaquae)* were usually due to those classes being very similar in appearance (Table 2). The same holds for other types of flagellates, e.g., classes Cryptomonadales and Euglenophyceae. Similar results have also been found by Sosik and Olson (2007), but it is difficult to compare our findings to the literature as class-specific confusion is usually not presented, let alone discussed. In many cases of confusion, the class differentiated with fewer training images was confused with a class differentiated by a large training image set with a close morphological resemblance. This emphasizes the power of larger training sets (Table 1, Table 2). This is a common problem with imbalanced data sets as the trained classifiers tend to be biased towards more numerous classes (Johnson and Khoshgoftaar, 2019).

Confusion of classes does not always create a major problem. In some cases, it is sufficient to simply achieve group-level identification. However, this must be carefully evaluated for research topics which require species discrimination. If the focus is on the identification of phytoplankton functional groups, it may be sufficient to determine which groups of plankton, e.g. dinoflagellates or diatoms, dominate the community. However, group-level differentiation is insufficient if we are interested in determining the biodiversity or whether toxic species are present. Group-level identification means that the results of classes closely related taxonomically, such as the *Eutreptiella* sp. and Euglenophyceae, can be united without having a practical impact. Looking at our classification results on a broader taxonomic group level it is evident that the results are reliable, at least for some groups. For the groups, Cyanophyceae, Cryptophyceae, and Bacillariophyceae 86 – 94% of the images were correctly classified, and for Euglenophyceae, Dinophyceae and Ciliophora 75 – 76% were correctly classified. For Cryptophyceae and Chlorophyta, less than 50% of the images were correctly identified and there was high uncertainty limiting the ability to make conclusions about their presence and abundance (Table 1, Figure 3). Confusion between groups was minor and incorrectly classified images were usually assigned to the unclassified group due to thresholding (Figure 3). It is always best to determine community composition down to the lowest taxonomical level possible. This is also desirable when using automated classification systems, especially when it is possible to identify them visually.

4.3 Towards operationality

Though often considered superior to other methods, CNNs are still not widely utilized for classifying and analyzing natural phytoplankton data sets. While the deep feature extraction outperforms handcrafted features, the latter performs well for several phytoplankton groups originally classified with both SVM and RF-based classifiers (Sosik and Olson, 2007; Laney and Sosik, 2014; Anglès et al., 2015; Bueno et al., 2017; Anglès et al., 2019; Fischer et al., 2020; Kraft et al., 2021). Currently, most ecological studies using phytoplankton data sets collected with an IFCB base their classification on the features and method developed by Sosik and Olson (2007) using RF instead of SVM. One reason for this is that CNNs typically require a long training time, high computing power and many training images, requiring more time and effort to establish an operational classification system. Additionally, publicly available codes and workflow are accessible for the RF-based classifier system, thus making it easier for biologically oriented groups to begin establishing new classifier systems (https://github.com/hsosik/ifcb-analysis).

CNNs require a notable amount of computational resources especially if they are trained from scratch. We showed that even relatively shallow CNN model requiring only a quite basic level of computing power with a small number of training data (26 classes out of 50 contained less than 300 training images) performs well. Previous studies support these findings (Bureš et al., 2021). During operational mode (the Evaluation Data), classification performance dropped drastically for many of the classes with few training images. However, 12 of those also achieved an F1-score of 0.7-0.97. This suggests not all classes require extremely large training sets, speeding up the process.

Creating the training sets in and of itself is laborious. It is therefore impractical to create classes for images of all small objects of similar shapes, such as different types of detritus and other types of difficult-to-assign objects. The creation of classes for such images will lead to too many variations in appearance. This will result in the matching of such images to other classes with similar features causing confusion. Hence we chose not to classify such images, but used thresholds instead. However, there is some benefit to creating classes for certain types of detritus as it allows them to be filtered out from the total community biomass. On the other hand, this leads to questions such as when should a phytoplankton cell be considered detritus, considering that all images have been triggered by a certain level of chlorophyll *a*. Therefore, all the collected images cannot be assigned to a specific class, but need to be filtered out.

A common approach in filtering of difficult-to-assign images is to apply thresholding to the class probabilities. Unfortunately, this approach is impractical, since thresholds need to be tailored to each class (see e.g. Luo et al., 2018). However, using thresholds is presently one of the most common methods when classifying natural samples (Sosik and Olson, 2007; Laney and Sosik, 2014; Anglès et al., 2015; Bueno et al., 2017; Anglès et al., 2019; Fischer et al., 2020; Kraft et al., 2021). In addition, the use of probability thresholds with CNNs is not straightforward due to the softmax function in the network architecture which converts neuron activations into class probabilities. This forces the network to assign a high probability to a specific class from the training set even when the input image is from a novel class. This makes it impossible to spot images which do not belong to existing classes because they are assigned to the wrong class creating false positives with high probabilities. We solved this problem by smoothing out the probability distribution making it easier to use this approach.

Classifier systems also tend to struggle with open-class problems, i.e., when it is applied to novel data whose classes are not featured in the training data (e.g. new species). It is impossible to make training sets for all possible new classes in advance and they would need to be distinguished from the classification results. Therefore, as has been noted in other studies, there is a need to solve this open-class problem. Exploration of different solutions such as anomaly detection (Pu et al., 2021), metric learning (Teigen et al., 2020), and hard negative mining (Walker and Orenstein, 2021) are ongoing.

4.4 Future perspectives

Utilizing new instruments which provide high-frequency information on plankton communities, combined with data analysis using CNNs are powerful tools for the investigation of plankton community dynamics. However, these methods require an entirely new way of both transferring and managing the data, as well as ensuring data quality. The possibilities provided by these new methods are only fully exploited when setting up a real-time data flow and analysis.

The data pipeline we created would have been more difficult to build without a proper service provider. We use an existing optical fibre infra for part of the data transfer but in most cases, a powerful cloud service is the most convenient solution. Here, the importance of the accessibility of these services with regards to both data transfer and storage for different fields of science along with technical developments is highlighted. The next step is to connect the different data pipelines to national and global level data repositories, broadening the accessibility and findability of different data, bearing in mind all the FAIR-data principles (Wilkinson et al., 2016). This also includes sharing large manually labeled image data sets, making it possible to adapt the new methods more quickly to a broader range of users. Additionally, data sets used to assess model performance should be more widely shared for testing purposes of new machine learning methods. However, validation of different image data sets is important as manual labeling is prone to human error. Additionally, it is often the more inexperienced taxonomists who carry out the manual image labeling tasks even if expert taxonomists would have been involved in the creation of the classes and identification of example images (Irisson et al., 2022).

Sharing labeled data sets is fundamental to the rapid development and implementation of classifier systems as this is the most laborious part of their set-up. The creation of a model library with pre-trained CNN models of different plankton communities could also aid the more widespread adoption of these new methods. According to Orenstein and Beijbom (2017) the best classification performance was achieved with a model originally trained with a general image repository and fine-tuned for plankton images. Models, already fine-tuned for different plankton communities, could be adopted into use for communities with similar species compositions and further fine-tuned to the target data with only a moderate amount of training data and computing resources. This would be useful because of the lack of machine learning expertise and the lack of availability of computational resources among plankton researchers as well as reducing the amount of training data needed in the final stage. It would also make it easier to test previously developed methods on different data and to find the most suitable solutions for different types of data sets. To get towards this the EcoTaxa (Picheral et al., 2017) has been created. However, it is a tool for storing, browsing and classifying slightly smaller data sets and is not targeted to large data sets (e.g. minimum of tens of millions of images per year as at Utö) produced by operational use (Irisson et al., 2022).

5 Conclusions

Novel automated microscopic imaging solutions, like imaging flow cytometry, combined with automated data flow and analysis systems take us a step towards real-time plankton community information. This is especially important for harmful algal bloom observations, such as the filamentous cyanobacteria in the Baltic Sea. Nevertheless, high-frequency community information will also be important in model development and remote sensing data validation. Thus, the development of these systems underlines the importance of data flow and analysis infrastructure as well as principles of open science. Collecting large, annotated image data sets requires a lot of work and creating efficient and functioning data pipelines and classification systems requires a substantial amount of coding. Sharing image data sets and classification models vastly ease the implementation of these systems and would accelerate the exploration of the vast number of plankton data sets already collected within a multitude of monitoring programs and research projects around the world.

Multiple studies have shown that CNNs function well in the classification of plankton. We also achieved high classification accuracy with transfer learning and relatively shallow CNN architecture. Moreover, our method was able to adequately classify natural samples making our approach suitable for operational use. Some issues in the utilization of automatic classification methods, such as CNNs, remain due to them struggling with the open-class problem. During the search for more sophisticated solutions, the use of probability thresholds can enable the filtering of images not belonging to those classes. However, this does not solve the problem of detecting and identifying new species. Although the use of thresholds is quite tedious and time-consuming, at the moment it is still the most commonly used solution. Some of the workload can be reduced with the use of validation and test sets of the labeled image data set to set proper thresholds and evaluate their suitability. However, the ideal method of setting thresholds would be by use of a data set consisting of images from different seasons and locations as well as multiple years. This can be achieved by gradually fine-tuned the thresholds while compiling data. High classification confusion is often related to close taxonomic affiliations, which is not an issue if the goal is to determine the dynamics of larger functional groups rather than the determination of species-specific dynamics. Our study represents a step forward in the development of automated, fully operational, near-real-time classification system which can ultimately help to uncover novel insights in plankton ecology.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below:

Data set 1: http://doi.org/10.23728/b2share.abf913e5a6ad47e6baa273ae0ed6617a, Eudat b2share data repository, Record number: ABF913E5A6AD47E6BAA273AE0ED6617A, Data set name: SYKE-phytoplankton_IFCB_2022.

Data set 2: http://doi.org/10.23728/b2share.7c273b6f409c47e98a868d6517be3ae3,Eudat b2share data repository, Record number: 7C273B6F409C47E98A868D6517BE3AE3,Data set name: SYKE-phytoplankton_IFCB_Utö_2021.

Author contributions

OV designed and wrote the code and established the data pipeline and near-real-time classifier system with the guidance of KK and JS. OV and KK drew the figures and tables. PY, SK, and MJ executed the practicalities at Utö station, with regards to data transfer and creation of the public HAB information webpage. KK and LH validated the Evaluation Data set. KK wrote the manuscript with the help of all authors. All authors contributed to the article and approved the submitted version.

Funding

The study utilized SYKE and FMI marine research infrastructure as a part of the national FINMARI RI consortium. The study was partly funded by Tiina and Antti Herlin Foundation (personal grant for KK), Academy of Finland project FASTVISION (grant no. 321980 and 321991), Academy of Finland project FASTVISION-plus (grant no. 339355 and 339612), JERICO-S3 project, funded by the European Commission's H2020 Framework Programme under grant agreement No. 871153, and PHIDIAS project, funded by the European Union's Connecting Europe Facility under grant agreement INEA/CEF/ICT/A2018/1810854.

Acknowledgments

We thank Ismo and Brita Willström for help with sensor maintenance at Utö. We acknowledge CSC for providing high-performance computing and cloud computing platforms for the study. We thank Danielle Bansfield for checking the manuscript language. We also thank the reviewers for providing valuable comments on the manuscript which helped vastly to improve it.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.867695/full#supplementary-material

References

Anglès, S., Jordi, A., and Campbell, L. (2015). Responses of the coastal phytoplankton community to tropical cyclones revealed by high-frequency imaging flow cytometry. *Limnol. Oceanogr.* 60, 1562–1576. doi: 10.1002/lno.10117

Anglès, S., Jordi, A., Henrichs, D. W., and Campbell, L. (2019). Influence of coastal upwelling and river discharge on the phytoplankton community composition in the northwestern gulf of Mexico. *Progr. Oceanogr.* 173, 26–36. doi: 10.1016/j.pocean.2019.02.001

Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

Bueno, G., Deniz, O., Pedraza, A., Ruiz-Santaquiteria, J., Salido, J., Cristóbal, G., et al. (2017). Automated diatom classification (Part a): Handcrafted feature approaches. *Appl. Sci.* 7, 753. doi: 10.3390/app7080753

Bureš, J., Eerola, T., Lensu, L., Kälviäinen, H., and Zemčík, P. (2021). "Plankton recognition in images with varying size" in *Proceedings of the international conference on pattern recognition* (ICPR). *Workshops Challenges*, 110–120. doi: 10.1007/978-3-030-68780-9_11

Campbell, L., Olson, R. J., Sosik, H. M., Abraham, A., Henrichs, D. W., Hyatt, C. J., et al (2010). First harmful Dinophysis (Dinophyceae, Dinophysiales) bloom in the US revealed by automated imaging flow cytometry. *J. Phycol.* 46, 66–75. doi: 10.1111/j.1529-8817.2009.00791.x

Campbell, L., Henrichs, D. W., Olson, R. J., and Sosik, H. M. (2013). Continuous automated imaging-in-flow cytometry for detection and early warning of Karenia brevis blooms in the Gulf of Mexico. *Environ. Sci. Pollut. Res.* 20, 6896–6902. doi: 10.1007/s11356-012-1437-4

CEN (2015) *DIN EN 16695 water quality – guidance on the estimation of phytoplankton biovolume: English version EN 16695, 2015*. Available at: https://standards.iteh.ai/catalog/standards/cen/bcc87031-164e-45b9-933a-7db83d4658f4/en-16695-2015 (Accessed July 9, 2020).

Correa, I., Drews, P., Botelho, S., de Souza, M. S., and Tavano, V. M. (2017). "Deep learning for microalgae classification," in *Proceedings of the 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 20–25. Available at: https://ieeexplore.ieee.org/abstract/document/8260609

Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Dai, J., Yu, Z., Zheng, H., Zheng, B., and Wang, N. (2017). A hybrid convolutional neural network for plankton classification, (eds) C. S. Chen, J. Lu and K. K. Ma In: *Computer Vision – ACCV 2016 Workshops. ACCV 2016. Lecture Notes in Computer Science()*. vol 10118, Cham: Springer. doi: 10.1007/978-3-319-54526-4_8

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). "Imagenet: A large-scale hierarchical image database," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 248–255. Available at: https://ieeexplore.ieee.org/abstract/document/5206848

Dunker, S., Boho, D., Wäldchen, J., and Mäder, P. (2018). Combining high-throughput imaging flow cytometry and deep learning for efficient species and life-cycle stage identification of phytoplankton. *BMC Ecol.* 18, 51. doi: 10.1186/s12898-018-0209-5

Faillettaz, R., Picheral, M., Luo, J. Y., Guigand, C., Cowen, R. K., and Irisson, J. O. (2016). Imperfect automatic image classification successfully describes plankton distribution patterns. *Methods Oceanogr.* 15, 60–77. doi: 10.1016/j.mio.2016.04.003

Farcy, P., Durand, D., Charria, G., Painting, S. J., Tamminen, T., Collingridge, K., et al. (2019). Towards a European coastal observing network to provide better answer to science and to societal challenges; the JERICO/JERICO-NEXT research infrastructure. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00529

Fischer, A. D., Hayashi, K., McGaraghan, A., and Kudela, R. M. (2020). Return of the "age of dinoflagellates" in Monterey bay: Drivers of dinoflagellate dominance examined using automated imaging flow cytometry and long-term time series analysis. *Limnol. Oceanogr.* 65, 2125–2141. doi: 10.1002/lno.11443

González, P., Álvarez, E., Díez, J., López-Urrutia, Á., and del Coz, J. J. (2017). Validation methods for plankton image classification systems. *Limnol. Oceanogr. Methods* 15, 221–237. doi: 10.1002/lom3.10151

Guo, B., Nyman, L., Nayak, A. R., Milmore, D., McFarland, M., Twardowski, M. S., et al. (2021). Automated plankton classification from holographic imagery with deep convolutional neural networks. *Limnol. oceanogr. Methods 19*, 21–36. doi: 10.1002/lom3.10402

Hällfors, G. (2004). Checklist of Baltic Sea phytoplankton species (including some heterotrophic protistan groups). *Baltic Sea Environ. Proc.* 95, 210.

Haraguchi, L., Jakobsen, H., Lundholm, N., and Carstensen, J. (2017). Monitoring natural phytoplankton communities: A comparison between traditional methods and pulse-shape recording flow cytometry. *Aquat. Microb. Ecol.* 80, 77–92. doi: 10.3354/ame01842

Harred, L. B., and Campbell, L. (2014). Predicting harmful algal blooms: A case study with *Dinophysis ovum* in the gulf of Mexico. *J. Plankton Res.* 36, 1434–1445. doi: 10.1093/plankt/fbu070

HELCOM (2017)"Monitoring of phytoplankton species composition, abundance and biomass." In: *Manual for marine monitoring in the COMBINE programme of HELCOM*. Available at: https://helcom.fi/media/publications/Manual-for-Marine-Monitoring-in-the-COMBINE-Programme-of-HELCOM.pdf (Accessed December 17, 2021).

Henrichs, D. W., Anglès, S., Gaonkar, C. C., and Campbell, L. (2021). Application of a convolutional neural network to improve automated early warning of harmful algal blooms. *Environ. Sci. pollut. Res.* 28, 28544–28555. doi: 10.1007/s11356-021-12471-2

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. doi: 10.1109/CVPR.2016.90

Honkanen, M., Müller, J. D., Seppälä, J., Rehder, G., Kielosto, S., Ylöstalo, P., et al. (2021). The diurnal cycle of pCO 2 in the coastal region of the Baltic Sea. *Ocean Sci.* 17, 1657–1675. doi: 10.5194/os-17-1657-2021

Hossin, M., and Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. knowledge Manage. process (IJDKP)*. 5, 1. doi: 10.5121/ijdkp.2015.5201

Hutchins, D., and Fu, F. (2017). Microorganisms and ocean global change. *Nat. Microbiol.* 2, 17058. doi: 10.1038/nmicrobiol.2017.58

Irisson, J. O., Ayata, S. D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. (2022). Machine learning for the study of plankton and marine snow from images. *Ann. Rev. Mar. Sci.* 14, 277–301. doi: 10.1146/annurev-marine-041921-013023

Johnson, J. M., and Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *J. Big Data* 6, 1–54. doi: 10.1186/s40537-019-0192-5

Kahru, M., and Elmgren, R. (2014). Multidecadal time series of satellite-detected accumulations of cyanobacteria in the Baltic Sea. *Biogeosciences* 11, 3619–3633. doi: 10.5194/bg-11-3619-2014

Kahru, M., Elmgren, R., Kaiser, J., Wasmund, N., and Savchuk, O. (2020). Cyanobacterial blooms in the Baltic Sea: Correlations with environmental factors. *Harmful Algae* 92, 101739. doi: 10.1016/j.hal.2019.101739

Kaitala, S., Kettunen, J., and Seppälä, J. (2014). Introduction to special issue: 5th ferrybox workshop–celebrating 20 years of the alg@ line. *J. Mar. Syst.* 140, 1–3. doi: 10.1016/j.jmarsys.2014.10.001

Kerr, T., Clark, J. R., Fileman, E. S., Widdicombe, C. E., and Pugeault, N. (2020). Collaborative deep learning models to handle class imbalance in FlowCam plankton imagery. *IEEE Access* 8, 170013–170032. doi: 10.1109/ACCESS.2020.3022242

Kingma, D. P., and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv*. [arXiv:1412.6980]

Kownacka, J., Busch, S., Göbel, J., Gromisz, S., Hällfors, H., Höglander, H., et al. (2018). Cyanobacteria biomass 1990-2018. HELCOM Baltic Sea environment fact sheets 2018. Available at: https://helcom.fi/wp-content/uploads/2020/06/BSEFS-Cyanobacteria-biomass.pdf. [Accessed August 19, 2022]

Kraft, K., Seppälä, J., Hällfors, H., Suikkanen, S., Ylöstalo, P., Anglès, S., et al. (2021). First application of IFCB high-frequency imaging-in-flow cytometry to investigate bloom-forming filamentous cyanobacteria in the Baltic Sea. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.594144

Laakso, L., Mikkonen, S., Drebs, A., Karjalainen, A., Pirinen, P., and Alenius, P. (2018). 100 years of atmospheric and marine observations at the Finnish utö island in the Baltic Sea. *Ocean Sci.* 14, 617–632. doi: 10.5194/os-14-617-2018

Laney, S. R., and Sosik, H. M. (2014). Phytoplankton assemblage structure in and around a massive under-ice bloom in the chukchi Sea. *Deep-Sea Res. II* 105, 30–41. doi: 10.1016/j.dsr2.2014.03.012

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Lombard, F., Boss, E., Waite, A. M., Uitz, J., Stemmann, L., Sosik, H. M., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

Lumini, A., and Nanni, L. (2019). Deep learning and transfer learning features for plankton classification. *Ecol. Inform.* 51, 33–43. doi: 10.1016/j.ecoinf.2019.02.007

Lumini, A., Nanni, L., and Maguolo, G. (2020). Deep learning for plankton and coral classification. *Appl. Comput. Inform.* Available at: https://www.emerald.com/insight/content/doi/10.1016/j.aci.2019.11.004/full/html

Luo, J. Y., Irisson, J.-O., Graham, B., Guigand, C., Sarafraz, A., Mader, C., et al. (2018). Automated plankton image analysis using convolutional neural networks. *Limnol. Oceanogr. Methods* 16, 814–827. doi: 10.1002/lom3.10285

Miloslavich, P., Bax, N. J., Simmons, S. E., Klein, E., Appeltans, W., Aburto-Oropeza, O., et al. (2018). Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. *Glob. Change Biol.* 24, 2416–2433. doi: 10.1111/gcb.14108

Moberg, E. A., and Sosik, H. M. (2012). Distance maps to estimate cell volume from two-dimensional plankton images. *Limnol. oceanogr. Methods* 10, 278–288. doi: 10.4319/lom.2012.10.278

Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., and Herrera, F. (2012). A unifying view on dataset shift in classification. *Pattern Recognit.* 45, 521–530. doi: 10.1016/j.patcog.2011.06.019

Muller-Karger, F. E., Miloslavich, P., Bax, N. J., Simmons, S., Costello, M. J., Sousa Pinto, I., et al. (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. doi: 10.3389/fmars.2018.00211

Niemistö, L., Rinne, I., Melvasalo, T., and Niemi, Å. (1989). Blue-green algae and their nitrogen fixation in the Baltic Sea in 1980, 1982 and 1984. *Meri* 17, 3–59.

Olli, K., Ptacnik, R., Klais, R., and Tamminen, T. (2019). Phytoplankton species richness along coastal and estuarine salinity continua. *Am. Nat.* 194, E41–E51. doi: 10.1086/703657

Olofsson, M., Suikkanen, S., Kobos, J., Wasmund, N., and Karlson, B. (2020). Basin-specific changes in filamentous cyanobacteria community composition across four decades in the Baltic Sea. *Harmful Algae* 91, 101685. doi: 10.1016/j.hal.2019.101685

Olson, R. J., and Sosik, H. M. (2007). A submersible imaging-in-flow instrument to analyze nano-and microplankton: Imaging FlowCytobot. *Limnol. oceanogr. Methods* 5, 195–203. doi: 10.4319/lom.2007.5.195

Orenstein, E. C., and Beijbom, O. (2017). "Transfer learning and deep feature extraction for planktonic image data sets," in *Proceedings of 2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. 1082–1088. Available at: https://ieeexplore.ieee.org/abstract/document/7926708

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). "Pytorch: An imperative style, high-performance deep learning library," Eds H. Wallach In *Advances in Neural Information Processing Systems 32*. 8024–8035. (Neural Information Processing Systems, 2019)

Picheral, M., Colin, S., and Irisson, J.-O. (2017). EcoTaxa, a tool for the taxonomic classification of images. Available at: https://ecotaxa.obs-vlfr.fr/ [Accessed August 19, 2022]

Pu, Y., Feng, Z., Wang, Z., Yang, Z., and Li, J. (2021). "Anomaly detection for *In situ* marine plankton images," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. 3654–3664. doi: 10.1109/ICCVW54120.2021.00409

Recht, B., Roelofs, R., Schmidt, L., and Shankar, V. (2019). "Do ImageNet classifiers generalize to ImageNet?," in *Proceedings of the 36th International Conference on Machine Learning (PMLR)* 97, 5389–5400. Available at: http://proceedings.mlr.press/v97/

Reynolds, C. S. (2006). *The ecology of phytoplankton* (Cambridge: Cambridge University Press). doi: 10.1017/CBO9780511542145

Righetti, D., Vogt, M., Gruber, N., Psomas, A., and Zimmermann, N. E. (2019). Global pattern of phytoplankton diversity driven by temperature and environmental variability. *Sci. Adv.* 5, eaau6253. doi: 10.1126/sciadv.aau6253

Ruokanen, L., Kaitala, S., Fleming, V., and Maunula, P. (2003). Alg@ line–joint operational unattended phytoplankton monitoring in the Baltic Sea. *Elsevier Oceanogr. Ser.* 69, 519–522. doi: 10.1016/S0422-9894(03)80083-1

Sosik, H. M., and Olson, R. J. (2007). Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* 5, 204–216. doi: 10.4319/lom.2007.5.204

Stal, L. J., Albertano, P., Bergman, B., von Bröckel, K., Gallon, J. R., Hayes, P. K., et al. (2003). BASIC: Baltic Sea cyanobacteria. an investigation of the structure and dynamics of water blooms of cyanobacteria in the Baltic Sea–responses to a changing environment. *Cont. Shelf Res.* 23, 1695–1714. doi: 10.1016/j.csr.2003.06.001

Teigen, A. L., Saad, A., and Stahl, A. (2020). Leveraging similarity metrics to in-situ discover planktonic interspecies variations or mutations. *Global Oceans 2020: Singapore–US. Gulf Coast, 2020*, pp. 1–8. doi: 10.1109/IEEECONF38699.2020.9388998

Thai-Nghe, N., Gantner, Z., and Schmidt-Thieme, L. (2010). "Cost-sensitive learning methods for imbalanced data," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*. Available at: https://ieeexplore.ieee.org/abstract/document/5596486

Walker, J. L., and Orenstein, E. C. (2021). "Improving rare-class recognition of marine plankton with hard negative mining," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*. 3672–3682. Available at: https://ieeexplore.ieee.org/abstract/document/9607849

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). Comment: the FAIR guiding principles for scientific data management and stewardship. *Sci. Data* 3, 1–9. doi: 10.1038/sdata.2016.18

WoRMS Editorial Board (2021). World register of marine species. Available at: https://www.marinespecies.org/ [Accessed August 19, 2022]

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 16 November 2022
DOI 10.3389/fmars.2022.986206

OPEN ACCESS

EDITED BY
Tim Wilhelm Nattkemper,
Bielefeld University, Germany

REVIEWED BY
Bank Beszteri,
Universität Duisburg-Essen, Germany
Jason Stockwell,
University of Vermont, United States

*CORRESPONDENCE
Sarah L. C. Giering
s.giering@noc.ac.uk

SPECIALTY SECTION
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

RECEIVED 04 July 2022
ACCEPTED 26 September 2022
PUBLISHED 16 November 2022

CITATION
Giering SLC, Culverhouse PF, Johns DG, McQuatters-Gollop A and Pitois SG (2022) Are plankton nets a thing of the past? An assessment of *in situ* imaging of zooplankton for large-scale ecosystem assessment and policy decision-making. *Front. Mar. Sci.* 9:986206. doi: 10.3389/fmars.2022.986206

Are plankton nets a thing of the past? An assessment of *in situ* imaging of zooplankton for large-scale ecosystem assessment and policy decision-making

Sarah L. C. Giering [1*], Phil F. Culverhouse [2], David G. Johns [3], Abigail McQuatters-Gollop [4] and Sophie G. Pitois [5]

[1]Ocean BioGeoscience, National Oceanography Centre, Southampton, United Kingdom, [2]Plankton Analytics Ltd., Plymouth, United Kingdom, [3]Marine Biological Association, Plymouth, United Kingdom, [4]School of Biological and Marine Sciences, University of Plymouth, Plymouth, United Kingdom, [5]Centre for Environment Fisheries and Aquatic Sciences (Cefas), Lowestoft, United Kingdom

Zooplankton are fundamental to aquatic ecosystem services such as carbon and nutrient cycling. Therefore, a robust evidence base of how zooplankton respond to changes in anthropogenic pressures, such as climate change and nutrient loading, is key to implementing effective policy-making and management measures. Currently, the data on which to base this evidence, such as long time-series and large-scale datasets of zooplankton distribution and community composition, are too sparse owing to practical limitations in traditional collection and analysis methods. The advance of *in situ* imaging technologies that can be deployed at large scales on autonomous platforms, coupled with artificial intelligence and machine learning (AI/ML) for image analysis, promises a solution. However, whether imaging could reasonably replace physical samples, and whether AI/ML can achieve a taxonomic resolution that scientists trust, is currently unclear. We here develop a roadmap for imaging and AI/ML for future zooplankton monitoring and research based on community consensus. To do so, we determined current perceptions of the zooplankton community with a focus on their experience and trust in the new technologies. Our survey revealed a clear consensus that traditional net sampling and taxonomy must be retained, yet imaging will play an important part in the future of zooplankton monitoring and research. A period of overlapping use of imaging and physical sampling systems is needed before imaging can reasonably replace physical sampling for widespread time-series zooplankton monitoring. In addition, comprehensive improvements in AI/ML and close collaboration between zooplankton researchers and AI developers are needed for AI-based taxonomy to be trusted and fully adopted. Encouragingly, the adoption of cutting-edge technologies for zooplankton research may provide a solution to maintaining the critical

taxonomic and ecological knowledge needed for future zooplankton monitoring and robust evidence-based policy decision-making.

KEYWORDS

in situ imaging, artificial intelligence/machine learning, taxonomy, digital samples, ecosystem assessment, long-term monitoring, zooplankton

Introduction

Zooplankton biodiversity contributes to multiple ecosystem services such as carbon and nutrient cycling, as well as the role of plankton in the marine food web. Understanding how plankton communities respond to changes in anthropogenic pressures, such as climate change and nutrient loading, is key to implementing effective management measures. The new generation of policy initiatives explicitly recognises the role that plankton biodiversity plays in delivering a variety of ecosystem services. These legislations, such as the United Nations Sustainable Development Goals (UN General Assembly, 2015), the Convention on Biological Diversity Aichi Targets (Convention on Biological Diversity, 2011), and the upcoming Post-2020 Global Biodiversity Framework (Convention on Biological Diversity, 2021), focus on a holistic view of biodiversity including the value of zooplankton. In Europe, for example, the Marine Strategy Framework Directive (Directive (EC) 2008/56, 2008) aims to achieve Good Environmental Status of marine waters, with plankton representing pelagic habitats in the legislation and implementation (European Commission, 2008; OSPAR, 2017; Bedford et al., 2018; McQuatters-Gollop et al., 2019; McQuatters-Gollop et al., 2022). European Union Member States are therefore required to monitor and assess the state of plankton, and, if needed, to implement management measures to achieve Good Environmental Status for pelagic habitats. Consequently, a comprehensive understanding of plankton communities is critically needed to inform a robust evidence base for supporting decision-making for marine management. Establishing a robust understanding of the relationships between anthropogenic pressures and zooplankton, however, depends on consistent time-series datasets, which are limited in number and spatial scale (McQuatters-Gollop et al., 2015; Zingone et al., 2015; McQuatters-Gollop et al., 2017). These gaps mean that policymakers have limited evidence on which to base decisions about enacting management measures related to plankton and the ecosystem services they provide.

Even though plankton in European waters are better sampled than those in many other parts of the world (O'Brien et al., 2017), gaps in this evidence base exist due to both lack of sampling and lack of knowledge of plankton dynamics and pressure-state relationships (McQuatters-Gollop et al., 2022). Zooplankton sampling is historically more limited than phytoplankton sampling, resulting in more numerous knowledge gaps around changes in zooplankton communities and the consequent effects on the marine food web and ecosystem services (McQuatters-Gollop et al., 2015). The UK's fixed-point monitoring programme, for example, has 11 phytoplankton sampling stations but only four of these also sample zooplankton; these stations are supplemented by phytoplankton sampling by the Environment Agency, but to date there has been virtually no inshore zooplankton sampling (Bedford et al., 2020). For larger spatial coverage, the Continuous Plankton Recorder [CPR, a towed net system (Batten et al., 2003)] provides a wealth of taxonomic data for both zooplankton and phytoplankton, particularly in UK and northern European waters, as well as parts of the North Atlantic, Pacific basins, Southern Ocean, and Australian waters (Figure 1). Yet, coverage for zooplankton data is still highly inconsistent, and wide expanses of coastlines and oceans are not covered at all (Figure 1).

A promising way to fill these gaps in spatial coverage is through the rapid advance of automated sampling systems and plankton imaging capabilities. Numerous commercial and custom-built plankton imaging systems are available (see reviews by Lombard et al., 2019; Giering et al., 2020a), and global roll-outs of zooplankton imaging platforms to match the Argo float global network for physical ocean parameters are starting (Lombard et al., 2019; Picheral et al., 2021).

While the technical abilities now exist to collect data continuously and at fine resolution (Lombard et al., 2019), a major bottleneck is - besides image storage and access - the processing and interpretation, specifically the taxonomic classification of zooplankton images (MacLeod et al., 2010; Orenstein et al., 2022). An obvious avenue to tackle the growing number of plankton images is the use of artificial intelligence (AI) and machine learning tools (ML) for the taxonomic classification of plankton. To date, AI/ML for plankton has been used primarily to aid human-based classification by presorting the images, because their 'predicted' taxonomic classifications can be highly variable (Gorsky et al., 2010). Tools available to the community that facilitate such AI/ML-augmented manual classification exist, such as EcoTaxa (Picheral et al., 2017) and MorphoCluster (Schröder et al., 2020). Yet, the reliance on human

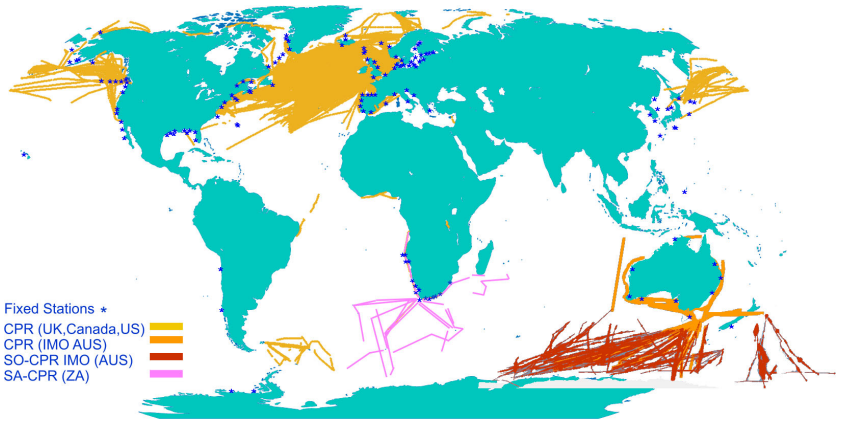


FIGURE 1
World map overlaid with fixed net sampling stations and Continuous Plankton Recorder tracks (adapted from source: MBA CPR map 1958-2020 (Batten et al., 2019), with data from NOAA Copepod database (O'Brian and Oakes, 2020), Australian IMOS database 1993-2021 (Re3Data.org, 2021) and South African CPR 2005-2021 (Huggett, pers. comm.).

verification limits the speed with which plankton images can be used for science.

The move to a global plankton-monitoring network hence heavily depends on the automation of taxonomic classification. But challenges with the needed fully-automated taxonomy (because of the large amounts of data) exist, such as questions about whether AI/ML can achieve a taxonomic resolution that scientists trust. We here develop a roadmap for future zooplankton monitoring for policy and management, specifically for the role of imaging and AI/ML, based on community consensus. To do so, we determined current perceptions by the zooplankton research and monitoring community about the use of imaging and AI/ML for zooplankton monitoring, with particular focus on their experience and trust in imaging and AI/ML to produce reliable taxonomic data. Specifically, we assessed the questions:

- Do zooplankton scientists think that images can ever replace physical samples to generate monitoring data?
- Do zooplankton scientists think that an AI can ever replace a human taxonomist in the role of identifying zooplankton?

We recommend the next steps for obtaining robust zooplankton data for large-scale ecosystem assessment and policy decision-making.

We use the term artificial intelligence (AI) to mean *the use of computer algorithms to make decisions.* In context, AI typically performs data analysis tasks done by humans such as identifying organisms from images. Machine learning (ML) denotes the method of training AI whereby *the algorithm improves ('learns') based on experience and use of data.* In context, ML may be carried out on images already labelled by humans ('training data').

Community survey on future of zooplankton monitoring

To obtain a broad sample of responses, we developed a questionnaire in English using JotForm (Supplementary Material). The survey was distributed between November 2021 and January 2022 using social media and through the authors' professional and personal networks, resulting in 179 complete responses. The final survey used a mixed-methods approach of 34 closed-answer questions. The first part of the survey used classification questions designed to provide an overview of the respondents' background (age, gender identity, location, education). The remainder of the questionnaire was designed to profile respondents' experience with zooplankton taxonomy, plankton imaging and AI/ML (qualifications, training, level of expertise, etc.), and their perceptions and trust in plankton imaging and AI/ML for zooplankton taxonomy. The latter was assessed using a series of 5-point Likert scale questions.

All respondents completed the survey themselves and gave their permission to use the results. Individuals were not identifiable from the data provided. All participants were 18 years of age or older. The survey described in this paper was reviewed and approved by the Ethics Committee of the National Oceanography Centre, UK.

Survey analysis

Quantitative data were analysed in R v4.0.2 (R Core Team, 2018). The level of expertise for zooplankton taxonomy, zooplankton imaging, and AI/ML was calculated as the sum of three questions (years of experience, skill level self-assessment, and frequency of training). Likert data were analysed using the 'Likert' function from the Likert package in R. Correlations were explored using simple

linear regression. The general bias towards or against imaging and AI/ML was calculated as follows. Each question was scored from 1 ('strongly disagree') to 5 ('strongly agree') on positive questions, and 1 ('strongly agree') to 5 ('strongly disagree') on reverse questions. If a participant was neutral, they would have scored 18 for imaging or AI/ML (6 questions all answered with neutral = 3). Consequently, a trust score of >18 indicates a favourable disposition towards the technology, while a score of < 18 indicates a negative disposition.

Participant demographic

We collected 179 complete responses. The participant gender distribution showed a near-equal gender balance in the field (55% male and 43% female) with the majority of respondents between 30-39 and 40-49 years old (31 and 26%, respectively) (Figure 2A). Globally, the survey reached participants working in 42 countries. The highest number of participants were from the UK and United States (33 and 25 participants, respectively), followed by Japan (11), Australia (10), Germany (9) and Canada (8) (Figure 2C). This distribution likely reflects funding support and activities in zooplankton monitoring and research as well as network connections both within the community and with the authors, and the use of language (English only).

The participants' expertise in zooplankton taxonomy was well spread with a slight bias towards intermediate and advanced (median of 3.2 on a scale from 1-Novice to 5-Expert) (Figure 2B). Overall, the participants had less expertise in zooplankton imaging (median 2.3), and least experience in AI/ML (median 1.8) with the majority identifying themselves as novices in this field (Figure 2B). This spread of expertise likely reflects that the field of AI/ML for zooplankton monitoring and research is relatively young and emerging compared to the field of zooplankton taxonomy.

Community consensus

Imaging for zooplankton monitoring and research

When asked about their perceptions on the use of imaging for zooplankton monitoring and research, the participants showed strong consensus that images can provide meaningful information (80%) and have clear advantages over net samples (68%) (Figure 3A). Conversely, participants agreed that images cannot provide the same level of information as physical samples (70%) and physical samples will always be required (72%) (Figure 3). No clear consensus emerged on whether physical samples are preferable (39% neither agreed nor disagreed). Finally, the survey suggested a consensus that time series can be continued with image samples once the technology has evolved sufficiently (62%) (Figure 3A). Overall, the majority of

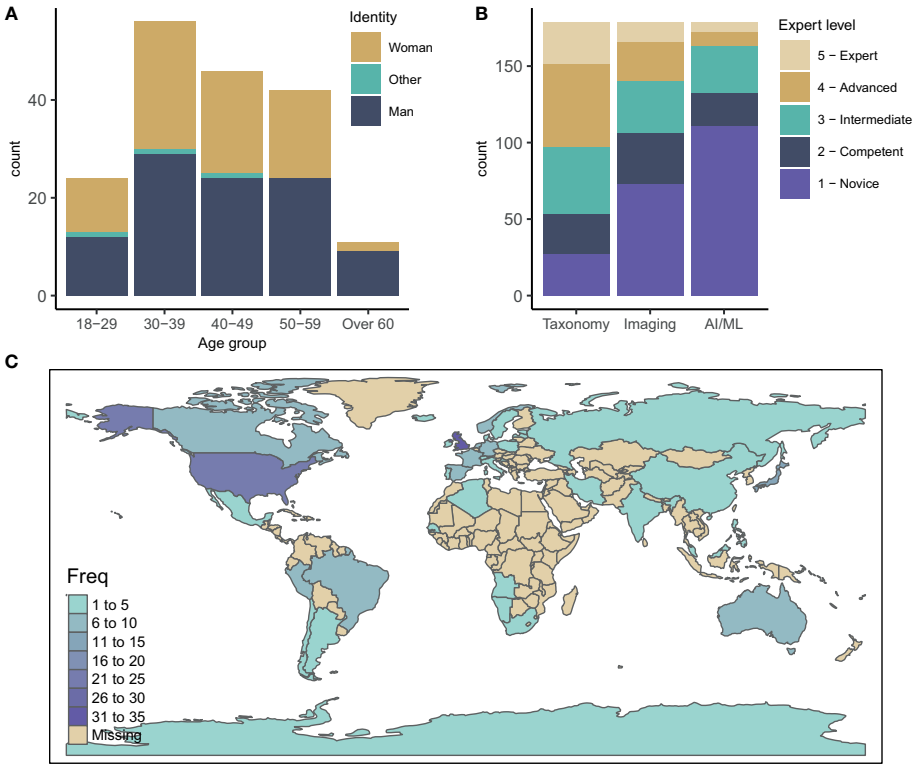


FIGURE 2
(A) Age and gender distribution of survey participants. **(B)** Expertise of participants in the three fields of zooplankton taxonomy, zooplankton imaging and AI/ML for zooplankton monitoring and research. **(C)** Geographic distribution of survey participants.

participants was optimistic about the use of images for plankton monitoring: 53% of participants responded positively towards images (with 12% being neutral and 35% having a negative disposition) (Figure 4A).

AI/ML for zooplankton taxonomy

When questioned about the potential of AI/ML for zooplankton taxonomy, respondents showed a strong consensus that AI/ML can help to analyse zooplankton data faster than current methods (79%; Figure 5A). However, a strong consensus that AI/ML is limited in its abilities and will always require human guidance and quality control was also evident (83%). When asked whether AI/ML will ever be as good as human taxonomists, which we assessed using both a positive and a reverse statement, no clear consensus was evident. Participants disagreed with the statement that AI/ML would be unbiased and more reliable than humans in identifying images (41%; Figure 5A). Finally, the participants strongly disagreed with the statement that human taxonomists will not be required in future once AI/ML has been trained sufficiently (84%). Indeed, the consensus on this statement was strongest when compared across all 12 questions. Overall, trust in AI/ML for correct taxonomic classification was low: 50% of the participants responded negatively towards AI/ML (with 13% being neutral and 37% having a positive disposition) (Figure 4B).

Perceived trustworthiness of AI for zooplankton taxonomy

A scientist's perception is likely influenced by their experience, and we observed clear patterns of this dependency in our survey results: The more respondents were experienced in zooplankton taxonomy, the less they trusted the use of zooplankton images and AI/ML for accurate taxonomy (Figures 6A, D); a significant negative trend was evident between taxonomy expertise and trust in images ($p < 0.001$, $R2 = 0.10$, $n = 179$) and AI/ML ($p < 0.001$, $R2 = 0.10$, $n = 179$). The expertise level in zooplankton imaging had no significant influence on perception of imaging and AI/ML for zooplankton monitoring (for both: $p > 0.13$, $R2 < 0.1$, $n = 179$). Across all imaging expertise levels, respondents were marginally positive towards imaging (median trust scores ≥ 18; Figure 6B). Conversely, the participants were marginally negative towards AI/ML (median trust scores ≤ 18; Figure 6C). Finally, very few of the survey participants were experienced in AI/ML (Figures 6C, D). While novices in this field were undecided on the usefulness of images and AI/ML, the experts tended to be optimistic about the use of zooplankton images (Figure 6C) though less optimistic about the use of AI/ML (Figure 6F). These trends were heavily influenced by single opinions because of the small number of participants who identified as advanced and expert users in the field of AI/ML for zooplankton research.

The survey participants perceived humans to be good taxonomists with 70% of the participants judging humans to identify >80% of the zooplankton specimens accurately. Sixty-six percent of the participants who rated both AI/ML and human accuracy in identifying zooplankton (103 out of 159) rated AI/ML skill lower than human skill. Only 21% of the participants thought they were similar, and 13% thought that AI/ML was more accurate than humans (Figure 7). Overall, the accuracy of humans was perceived to be significantly better (average rating of 80-90%) than that of AI/ML (average rating 70-80%; paired Wilcoxon test: $p < 0.001$, $n = 156$).

The participants believed that AI/ML could reasonably identify a copepod to family (28% of participants) or genus level (33% of participants) Figure 8A. For gelatinous zooplankton, the consensus appeared to be that AI/ML could reasonably identify gelatinous zooplankton to family level (33% of participants) (Figure 8B). For both questions (identifying copepods and gelatinous zooplankton), we also asked the participants whether their opinion was mostly influenced by their understanding of the image quality, the capability of AI/ML, or both in equal measures. The participants based their predictions primarily on their understanding of image quality alone or equally both on their understanding of image quality and AI/ML.

Towards a road map

Are images the future?

Our survey results indicated a strong community consensus that images (i.e. digital samples) are a valuable tool for plankton monitoring with clear advantages over physical net samples (Figure 3A), likely reflecting the financial and logistical constraints associated with net sampling. Traditional nets require human-centric, platform-based deployments (usually off a ship) and are hence very limited in their spatiotemporal resolution. The physical samples are stored, often in hazardous chemical preservatives, and shipped to a laboratory for analysis, leading to logistical challenges and considerable delays between sample collection and data availability. Image samples, in contrast, are stored digitally, which offers - amongst other advantages - the ability to share images easily for, e.g., quality control and additional taxonomic classifications by other researchers. Our survey supports the notion that moving towards automated routine image-based sampling combined with image analysis is key to increasing the quantity of zooplankton data to obtain the spatiotemporal coverage required for robust decision-making.

Yet, the survey results also revealed a strong community consensus that, despite the logistical constraints of collecting physical samples, physical samples cannot be replaced by images entirely and will always be required (Figure 3). The reason for the ongoing need for physical samples is likely twofold. First, deeper taxonomic analyses still require physical samples as, at this stage, microscopes offer the often required higher resolution and, importantly, allow the user to investigate each specimen in multiple dimensions and with different exposures. For example, species of the same genus may be morphologically almost indistinguishable bar minute differences in body structures (Fleminger and Hulsemann, 1977; Frost, 1989; Wilson et al., 2015). Considering current technology, such detailed taxonomic information is unlikely from *in situ* images in the foreseeable future. This notion is also reflected in our survey, where participants revealed low confidence that image quality is sufficient to resolve copepod and gelatinous zooplankton at the species level (Figure 8). Second, physical samples are required for information that cannot be obtained from images, such as biochemical and molecular analyses, which have the potential to greatly advance our understanding of zooplankton biodiversity, ecology and connectivity (Lenz et al., 2021).

While both types of samples (physical and digital) have their advantages, the biggest gain can likely be made when both are used strategically in conjunction (Figure 9). We could leverage the existing monitoring strategies and enhance these through imaging. Ships Of Opportunity have been used by the CPR survey since 1931 to collect physical plankton samples (Batten et al., 2003). Initiatives are now in progress to fit CPRs with holographic camera systems, allowing the simultaneous match-up of *in situ* imaging data with CPR physical samples (Johns, pers. comm.). In addition, when the physical CPR samples are analysed under the microscope in the laboratory, taxonomists are asked to take an image of each specimen (Johns, pers. comm.). For physical samples, specimens can be imaged before being analysed, e.g., for biochemical composition (Giering et al., 2019). Bench-top instruments for net sample imaging include ZooScan (Grosjean et al., 2004; Picheral et al., 2010) and FlowCam (Detmer et al., 2019). Hence, all physical samples could also be imaged, potentially providing high-quality taxonomic training datasets and additional information on how to translate images into biochemical parameters.

On a broader scale, an extension of the current imaging network is the next logical step, and international initiatives to facilitate such networks have commenced (Lombard et al., 2019; de Vargas et al., 2022). Coverage of CPR lines, ideally coupled with imaging, should be expanded to regions with currently poor coverage such as the South Atlantic and Central and South Pacific (Figure 1). As the CPR instrument has to be lowered into the sea and towed behind, it is not suitable for use on all ships. An alternative method is the FerryBox concept, which uses the ship's pumped water supply (Petersen and Colijn, 2017). While some imaging systems have already been integrated into FerryBoxes (Gannon, 1975), major problems remain with their operation, reliability, size range (too small for large zooplankton), and the development of efficient image processing and classification [https://www.ferrybox.org/]. The Plankton Imager (Pitois et al., 2018; Pitois et al., 2021; Scott et al.,

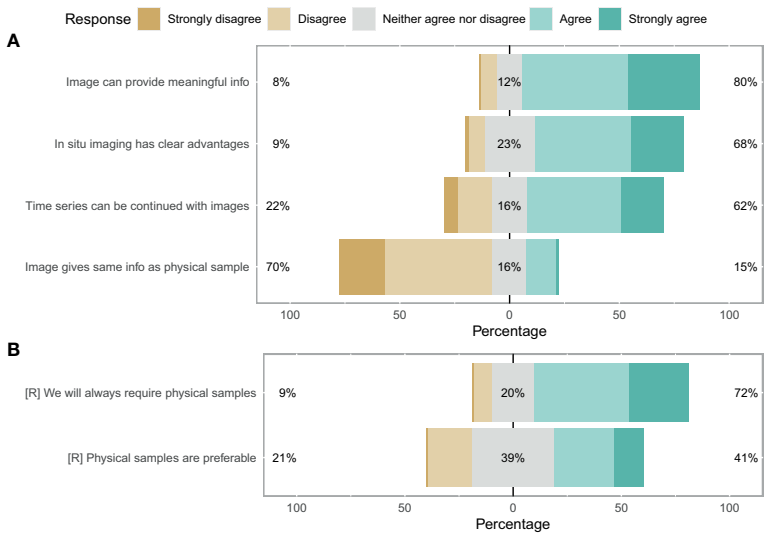


FIGURE 3
Likert plot for image sufficiency. **(A)** Darker green shows consensus favours replacement of physical samples with images. Darker brown means a preference for keeping the system as it is. An equal spread likely indicates no clear consensus. **(B)** Reverse as for **(A)**.

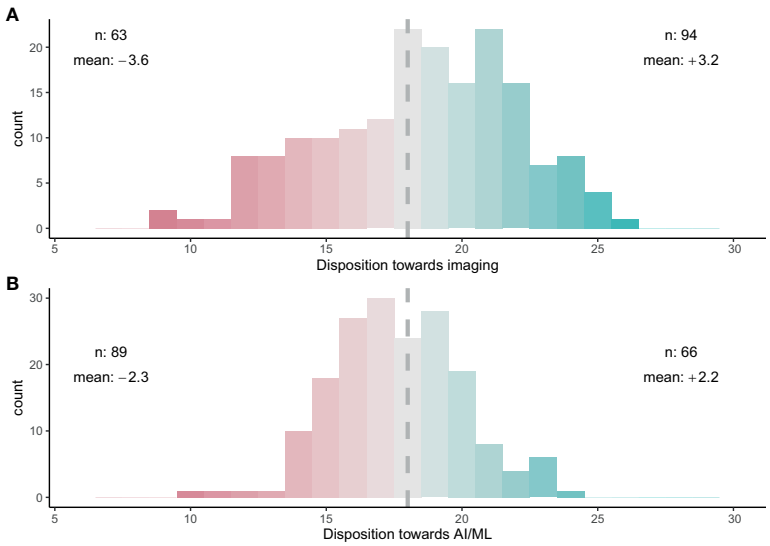


FIGURE 4
Overall trust in **(A)** imaging for zooplankton taxonomy and **(B)** AI/ML for zooplankton taxonomy. A score of 18 indicates a neutral stance towards imaging or AI/ML. A trust score of >18 (green) indicates a favourable disposition towards the technology, while a score of < 18 (red) indicates a negative disposition. 'n' shows the number of participants on either side of 18. 'mean' shows the average score above or below 18, with a higher absolute number indicating a stronger community bias.

2021) can use the same water source as FerryBox, allowing images of the mesozooplankton to be collected at high speed and moderate volume (34 L min^{-1}) offering similar sampling volume to the CPR (300 L [nautical mile]$^{-1}$) (John et al., 2002). For research vessels, camera systems such as the Underwater Vision Profiler (Picheral et al., 2010) could be integrated with water sampling rosettes as standard to improve vertically resolved information on zooplankton. Finally, miniaturised camera systems can be fitted on autonomous vehicles, such as floats and gliders (Picheral et al., 2021).

Zooplankton cover a wide range of diversity of organisms in terms of size, shape, and behaviour. As a result, no plankton

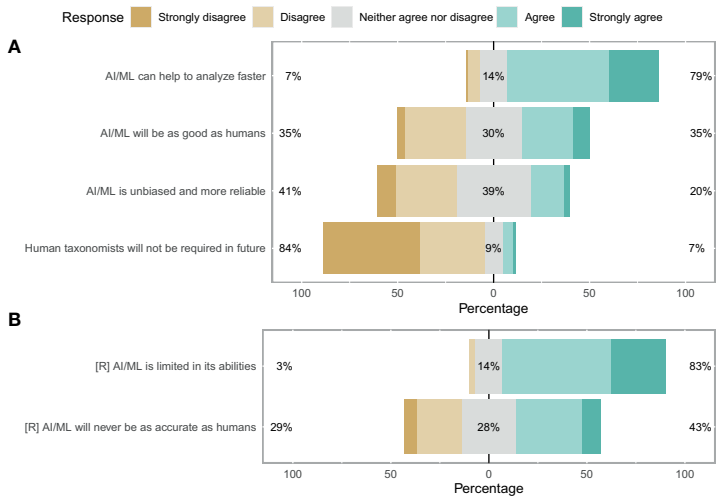


FIGURE 5
Likert plot for taxonomy sufficiency. **(A)** More green means consensus favours the replacement of human taxonomists with AI. More brown means a preference for keeping the system as it is. An equal spread likely indicates no clear consensus. **(B)** Reverse as for a.

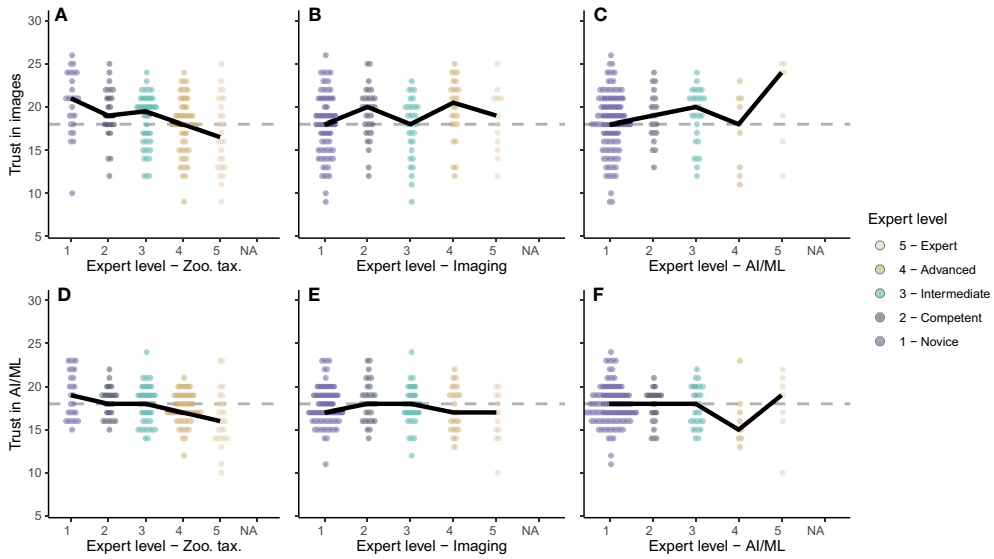


FIGURE 6
(A–C) Trust in images for zooplankton monitoring and **(D–F)** trust in AI/ML for zooplankton classification based on expertise in zooplankton taxonomy **(A, D)**, zooplankton imaging **(B, E)** and AI/ML **(C, F)**. Expert levels range from 1-Novice to 5-Expert as shown in the legend. Dots show the individual scores for each survey participant. Solid black lines connect the medians for each expert level.

sampling system - whether collecting digital or physical samples - can estimate the abundance for all components of the plankton at any given time, and any system will likely be biased towards a specific component of the plankton (Owens et al., 2013). Combining datasets from different plankton sampling systems is hence non-trivial. The selection of a sampler and associated sampling design will determine sampling efficiency and selectivity (Pitois et al., 2016; Pitois et al., 2018). Practical issues associated with the collection of physical zooplankton samples (Sameoto et al., 2000) include: active and passive avoidance of the net (Fleminger and Clutter, 1965; Clutter and Anraku, 1968), net clogging, and plankton patchiness (Wiebe and Benfield, 2003; Skjoldal et al., 2013). Imaging devices will not have to cater for all issues associated with nets, but their efficiency will also be dependent on system avoidance, potential damage to fragile organisms particularly when a pumped system is used (albeit typically less problematic compared to net sampling), and camera performance (Pitois et al., 2018). Comparisons between imaging systems and net samples indicated that sampling caveats affect nets and imaging systems in similar proportions (e.g. (Finlay and Roff, 2004; Nogueira et al., 2004; Basedow et al., 2013; Pitois et al., 2018)). In addition, different image processing routines (Giering et al., 2020b) specific to each instrument can result in images that are not directly comparable. A very important step going forward is hence the inter-calibration of all instruments so that all datasets can be combined (Lombard et al., 2019).

Our survey revealed a consensus that time series can be continued with image samples once the technology has evolved sufficiently, indicating a general optimism about the future of imaging for zooplankton monitoring. As we did not investigate what the participants deemed as 'sufficient', two aspects need to be considered when evaluating this statement: (1) scientific sufficiency of state-of-the-art technologies, and (2) perceived sufficiency. While recent reviews suggest that further technological and methodological developments are needed to meet the scientific needs (e.g. Lombard et al., 2019; Giering et al., 2020), this survey suggests that the zooplankton research community is generally willing to adopt these technologies and methodologies. Yet, a period of overlapping use of imaging systems and physical sampling systems, as well as thorough intercalibration between technologies [e.g. (Lombard et al., 2019; Giering et al., 2020a)], will be needed to establish a statistical correlation between the methods before imaging can reasonably replace physical sampling.

Are future taxonomists human?

Our survey results indicated that the community is less favourable towards AI/ML for zooplankton research than towards imaging (Figure 4), likely reflecting the challenges that accurate zooplankton taxonomy poses. Taxonomists learn from concepts, examples and experience, and apply context metadata knowledge to each classification task. Zooplankton taxonomy has a well-established framework with an extensive base of taxonomic literature, most of which is text-based with hand drawings of the organisms' key features. Reference sheets (for

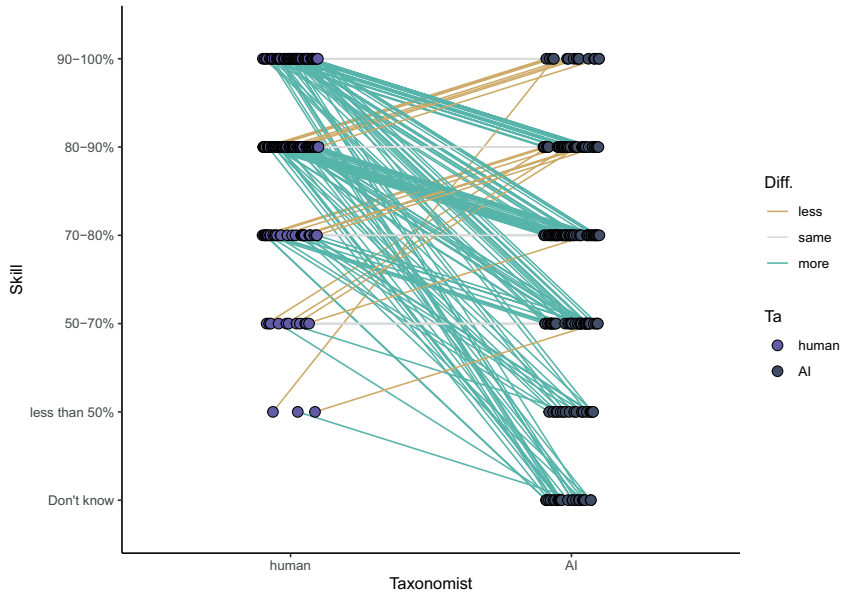


FIGURE 7
Opinions on the level that humans and AI can correctly identify zooplankton. Lines indicate whether a participant thinks humans are more accurate than AI (green), equally accurate (grey) or less accurate (yellow) than AI.

example, the ICES leaflets for marine zooplankton) hold expert keys and drawings that are both distillations and translations of the physical properties of organisms as seen under visual examination (Figure 10). The expert will perform the visual mapping from the hand-drawn "type specimen" to interpret the taxonomic features of the collected specimen, and supplement these using the textual notes on taxonomic descriptions and context metadata (such as size, species distribution, and life history). Together, these information constrain the identification of the collected specimen. The desired confidence in classification often requires a serial search through such taxonomic guides and the call for a second expert opinion. If a taxonomist is not confident in their classification, the specimen is assigned the highest taxonomic level the expert is confident in, or the most probable identification (Choquet et al., 2018). As such, accurate classification of zooplankton is a complex task. This complexity likely explains why researchers with more expertise in zooplankton taxonomy mistrust the use of AI/ML for zooplankton research (Figures 6A, D).

Currently, AI is typically trained on image training datasets produced specifically for the target study, including region and instrument, annotated by the study's primary researchers. With image quality sometimes low (Lombard et al., 2019; Giering et al., 2020a) and identification frequently carried out by non-specialists (Irisson et al., 2022), confidence in human-led annotation can be low. Sixty percent of the survey participants (that answered the question with a rating) have only a moderate level of trust in current zooplankton training datasets. As ML relies on training data quality ('garbage in, garbage out'), low-confidence training datasets pose a problem. Hence, we propose that a sufficiently rigorous process, including consensus of training data classification by multiple experts (akin to 'quality in, quality out'), is needed to facilitate reliable automated classifications that are trusted by the scientific community.

One option is to programme AI to use the same cross-referencing and matching through both visual and textual descriptions as the taxonomy texts (Figure 10). Multimodal approaches, which use both text and image, are now widely applied across a variety of tasks [see reviews by (Baltrušaitis et al., 2019; Chen et al., 2021; Uppal et al., 2022)]. In addition, AI could apply context metadata knowledge to each classification task in a similar way as humans do. For seafloor mapping, for example, the assumption that images captured close to each other are more similar than those taken further apart improves image classification by a factor of two (Yamada et al., 2021). For zooplankton images, the inclusion of context metadata (geometric, hydrographic and geo-temporal information) significantly improves classification accuracy (Ellen et al., 2019).

The survey participants did not agree with the statement that AI/ML is unbiased and more reliable than humans in identifying images (41%), suggesting that human taxonomists are considered reasonably reliable. Yet, expert cognitive biases can contribute to inconsistent performance when manually labelling physical specimens, with inconsistencies in both counting and classification (Culverhouse et al., 2014). For example, repeat analyses of physical net samples, by the same analyst, using microscopy revealed that

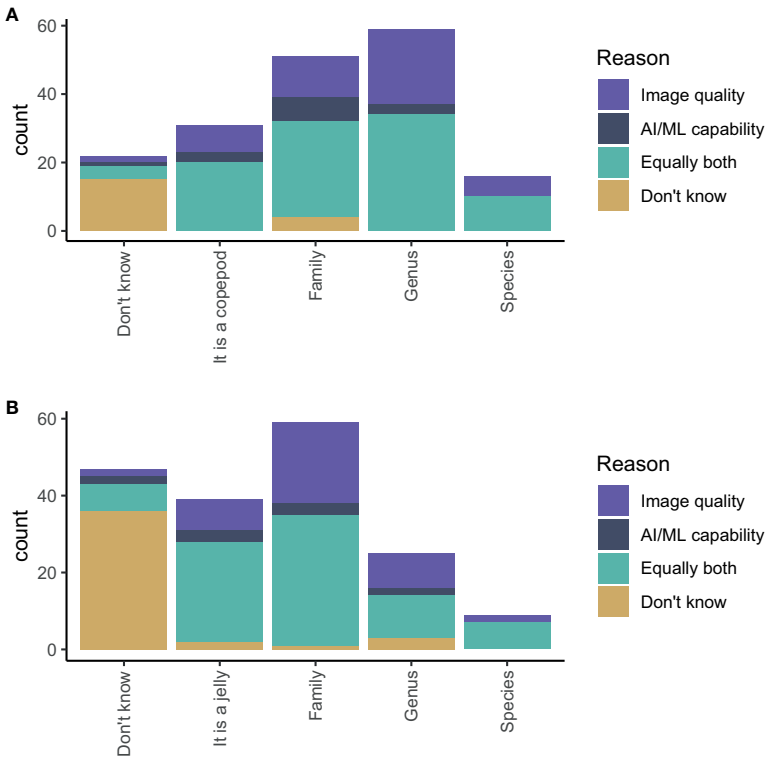


FIGURE 8
Estimates of the highest taxonomic level an AI/ML algorithm can positively identify **(A)** a copepod and **(B)** a gelatinous zooplankton using images generated from an appropriate imaging system. The colours show whether a participant's answer was most influenced by their understanding of image quality (purple), the capability of AI/ML (dark blue), equally both (turquoise), or unknown (yellow).

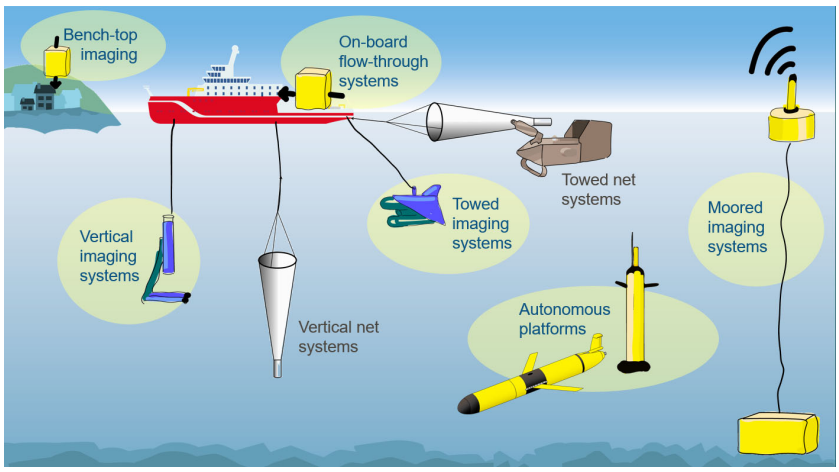


FIGURE 9
Future zooplankton sampling network. The biggest leverage can be made when both physical sampling, such as vertical and towed net systems, and imaging (highlighted by yellow backgrounds) are used strategically in conjunction. Imaging systems fitted to autonomous and moored platforms allow global coverage with reduced reliance on ships.

human-generated repeat counts differed on average by 8%, and taxonomic classification consistency (into 10 broad categories) was on average 75% (Culverhouse et al., 2014). Moreover, differences in counts varied as much as an order of magnitude for the same sample when analysed by different taxonomists, likely owing to psychological factors such as boredom, fatigue and prior expectations (Culverhouse et al., 2014). As such, human-led taxonomy results in non-repeatable outputs, where the same taxonomist at the same location using the same methodology is unlikely to arrive at the same result when repeating sample analysis. Machine learning, in contrast, allows repeatability of analysis results as long as the pre-trained model and weights are used [though the implementation of repeatability needs to be checked prior to model deployment, particularly for deep learning models (Alahmari et al., 2020)]. Yet, self-consistency (i.e. the same person coming to the same conclusion every time) and peer-consistency (i.e. several experts arrive at the same conclusion) (Culverhouse et al., 2003) have received relatively little attention in AI/ML for zooplankton research (Culverhouse et al., 2014).

As current AI-based classifications may be too inaccurate to be used directly for many zooplankton research questions (Irisson et al., 2022), a common practice is to use AI/ML to presort images into classes and then manually verify each AI-based classification; and several commercial and open-source platforms have been designed specifically for this purpose, such as EcoTaxa (Picheral et al., 2017). Alternative strategies are being developed, where unsupervised and supervised classifications alternate to reduce the number of images that a human has to manually verify (e.g. Schröder et al., 2020). The benefit of such workflows is widely accepted by the community, as indicated by our survey results (79% of survey participants agreed that AI/ML can help to analyse images faster; Figure 5A).

AI/ML is still in its infancy and formalised assessment of bias is largely unexplored, which partly explains our survey results that the community is currently undecided whether AI/ML can be as or more accurate as humans for the classification of plankton images (Figures 5, 6D–F). Even if AI will someday be as accurate as human taxonomists, our survey shows a strong community consensus that taxonomists will still be needed in future (84% of participants; Figure 5). While we did not ask specifically why this is the case, several reasons for this judgement are possible. First, the purpose of AI-led classification is to help researchers address scientific questions. Thus, an aspect of scientific quality control will always be required, where a taxonomically literate researcher may perform spot checks and affirm the overall classification as appropriate for the scientific endeavour on hand. Taxonomic experts may further oversee the expansion of current classification algorithms to include newly discovered species or similar amendments to reflect the current scientific knowledge accurately. Second, zooplankton research extends far beyond simply identifying images. Hence, physical samples will continue to play a major role in environmental research (e.g. for biogeochemical analysis or experimental work) and their handling will require expert human taxonomists.

Building human trust in AI/ML for zooplankton research

The survey showed an overall mistrust in the use of AI/ML for zooplankton research, which agrees with reported general attitudes towards AI (Schepman and Rodway, 2020). In their study, Schepman and Rodway (2020) found that participants were positive towards AI and felt comfortable with its use when the application helped humans carry out tasks but did not replace humans or gain autonomy. Conversely, negative feelings were associated with AI applications that involved aspects of human judgement, skill, social understanding or empathy (Schepman and Rodway, 2020). These conclusions can explain some of the trends we observed in our survey, suggesting that researchers consider accurate taxonomy as a difficult skill often relying on judgement based on abstraction and context understanding. While the survey participants felt comfortable with using AI/ML to aid taxonomy (e.g. by presorting images), the replacement of humans with AI was met with scepticism even though research has documented the inaccuracies in human-based taxonomy (Culverhouse et al., 2014).

The reason for the apparent negative perception of AI/ML for zooplankton research is likely founded on a combination of aspects. For those who have not had successful experiences with AI-based classifications and required further taxonomic verification by human taxonomists, trust in automated classification may be weak. Such experiences could explain why 66% of the survey participants that answered the question thought that humans can achieve higher levels of taxonomic accuracy than AI/ML (Figure 7). Yet, ring trials using microscopy on physical samples (community-driven comparison of taxonomic classification across different zooplankton laboratories) show that even highly trained professional zooplankton taxonomists often achieve an identification accuracy of only ~80%, with the identification of copepods to species level posing the biggest challenge (Wootton and Johns, 2019). In contrast, plankton classifiers with an accuracy of >90% have already been developed (Dai et al., 2016; Wang et al., 2018; Ellen et al., 2019; Kerr et al., 2020); though it is seldomly reported whether these classifiers successfully identify key and indicator species, which may be rare (Xue et al., 2018). Another aspect that will influence the trust in AI/ML is previous experience with this technology. Even though we tried to distribute our survey widely, 62% of the participants rated themselves as 'novice' in AI/ML for zooplankton image identification and only 9% rated themselves as 'advanced' or 'expert', reflecting that AI/ML is young in this field and has had limited uptake by the community.

Trust, experience and expertise in AI within the zooplankton research community need to increase for AI-based taxonomy to become fully adopted. Trust is influenced by both the perception of

the technology's competence and emotional factors, and actions to facilitate the adoption of new complex technology, such as AI, need to address both (Hoff and Bashir, 2015; Glikson and Woolley, 2020). Experiments suggested that interaction with AI can significantly increase trust in the observed AI and in future uses (Ullman and Malle, 2017). A visual presence of the AI (rather than an embedded, 'black box' feature) also builds trust (Glikson and Woolley, 2020). Visual presence could include visual interfaces as well as visual representations of the results, such as group collages that can be explored by the researcher. In addition, generating the perception of a 'persona', the use of human-like behaviour, and personalization to the user's needs and preferences can help to build emotional trust (Glikson and Woolley, 2020).

In addition, investing in a good reputation and transparency of how the algorithm works also increases trust in the AI's competence (Glikson and Woolley, 2020). A key step in this process is an increased effort in the development of explainable AI (often referred to as XAI), which provides explanations for the algorithm's decisions and outputs that are understandable for non-AI experts (in this case, a zooplankton researcher). Keystones for explainability include (1) transparency of how the algorithm works, (2) explanation of the underlying rules for the decision ('causality'), (3) quantification of bias that could have originated from shortcomings of the training data or choice in algorithm, and (4) confidence in the reliability of the predictions (Hagras, 2018). XAI has gained attention only in the past decade (Carvalho et al., 2019) but is now considered critical for the widespread adoption of AI (e.g. UK Parliament, 2017). However, how exactly XAI for plankton classification could be implemented to maximize trust and confidence by zooplankton researchers is yet unclear, and an appropriate framework needs to be developed through close collaboration between zooplankton taxonomists and researchers ('users') and AI developers. Finally, matching users' expectations and AI performance by providing clear explanations about the AI's functionality both in terms of how the algorithms work and why they should be used (compared to alternatives) is important. Possible avenues to build cognitive trust thus include demonstration and quantification of reliability of the AI, development of XAI, and close collaboration and dedicated workshops for zooplankton researchers and AI developers.

A new face for zooplankton taxonomy

The exciting developments in cutting-edge information technology for zooplankton research further offer the opportunity of a 'face-lift' for the field of taxonomy. The number of taxonomists has declined worldwide (MacLeod et al., 2010; Culverhouse, 2015; McQuatters-Gollop et al., 2017), and fewer trained taxonomists and plankton analysts are recruited each year to replace the previous generation as it retires (McQuatters-Gollop et al., 2017). One solution to this 'brain drain' in plankton taxonomy could be to engage traditional taxonomists in training in AL/ML. Such engagement would build trust in the new techniques and also enhance the field of taxonomy with innovative, cutting-edge engineering and informatics technologies. The added interdisciplinary flavour could increase interest in the field of plankton taxonomy because the skills used to collect and analyse *in situ* imaging data are globally in demand and widely transferable across many non-scientific sectors such as business, economics, and computing. A starting point for merging taxonomy with engineering and computer sciences could be the development of courses that teach the combined skills of AI/imaging/plankton taxonomy at universities. By teaching these skills together, students may start to recognise the links between taxonomy and technology, helping to rebrand zooplankton taxonomy as 'exciting and relevant' rather than a career 'dead end'. This 'new face' for zooplankton taxonomy and research may provide a solution to securing, into the future, critical taxonomic and ecological knowledge needed for future zooplankton monitoring and robust evidence-based decision-making and policy.

Robust evidence base for decision-making

To enable policymakers to best make informed decisions about enacting management measures, we require a robust evidence base founded on consistent time-series datasets and broad global coverage. Currently, understanding of plankton dynamics, particularly in response to climate change and direct anthropogenic pressures, is limited due to gaps in data coverage or taxonomic mismatches between time-series with different methodologies. The result is a lack of confidence in the evidence base underpinning decision making (McQuatters-Gollop et al., 2015; McQuatters-Gollop et al., 2017). A big challenge here is how to combine different datasets of varying taxonomic levels. An example of how merged datasets can work is the UK's and OSPAR's approach to assessing pelagic habitats in the Northeast Atlantic and the North Sea, which uses flexible indicators that work with a variety of plankton datasets, regardless of differences in sampling method or taxonomic resolution (McQuatters-Gollop et al., 2017; Rombouts et al., 2019; Bedford et al., 2020; McQuatters-Gollop et al., 2022). For example, the "Change in Plankton Communities" indicator applies a plankton lifeform indicator approach that uses functional traits to group plankton taxa into ecologically-relevant lifeform pairs where changes in relative abundance indicate an alteration in ecosystem functioning (McQuatters-Gollop et al., 2019; McQuatters-Gollop et al., 2022). This approach uses taxonomic phytoplankton and zooplankton data that do not need to be refined to the species level. Rather, because of the aggregative nature of lifeforms, data at the order, family, and genus levels can

still inform the indicator. Similarly, the "Change in Plankton Biomass and Abundance" indicator is partially informed by data on copepod abundance (OSPAR, 2017). For these indicators, sufficient information is hence broad zooplankton lifeforms identification (e.g. large and small copepods, meroplankton) and abundance. Zooplankton image data therefore has great potential to contribute to these indicators as our survey results indicated a community consensus that images can reasonably inform on the family and genus level (Figure 8).

To retrieve such information from the growing amount of image data swiftly, however, we will have to come up with a strategy to extract relevant taxonomic and abundance information from the images in an automated way. In the foreseeable future, AI will likely be able to classify and count from image datasets with limited input from human taxonomists. Image data could hence be used at a coarse taxonomic level to provide information on lifeforms, or other easily identifiable zooplankton groups, over large spatial scales, akin to Argo data (Roemmich et al., 2019). Thus, the combined use of imaging and automated classification will likely be appropriate to answer questions that require taxonomic resolution that is consistent with the accuracy of the available AI/ML. With AI automation, data can be analysed on-board, for example on the ship during a survey or on a platform, and sent *via* satellite to provide near-real-time information.

Clear guidelines on quality assurance are required for such a workflow. Algorithms have to follow the FAIR principles (Findability, Accessibility, Interoperability, Reusability) (Hartley and Olsson, 2020). Taxonomic classifications need to link image labels to machine-readable taxonomic trees, such as WoRMS. Data should contain all sources and contributors, including information on both the human and AI who carried out the classifications (i.e. a 'taxonomist ID'). Finally, the accuracy and certainty of all classifications should be clearly documented (e.g. how sure is the algorithm/human about the identification).

Conclusion and roadmap

The ultimate goal is a cost-effective global zooplankton monitoring programme with comprehensive spatiotemporal coverage that can answer scientific questions and contribute to the robust evidence base required to inform decision making for environmental management. Our survey revealed a clear community consensus that net sampling and traditional taxonomy must be retained in future, yet imaging will play an increasingly important part in the future of zooplankton monitoring and research. For imaging, challenges to address will include, besides technical hurdles such as the transfer of large data and image processing speed, the integration of the outputs from both physical and digital sampling methods. A period of overlapping use of imaging systems and physical sampling systems will be needed before imaging can reasonably replace physical sampling for widespread time-series zooplankton monitoring. In addition, improvements in AI/ML are needed for these to be trusted and fully adopted by zooplankton researchers, particularly taxonomists. The key step forward is parallel programmes that complement each other, while efforts are focussed on bringing imaging technologies on par with traditional taxonomy. This long-term goal will no doubt mean overcoming several challenges, and only then can nets for routine monitoring become a thing of the past.

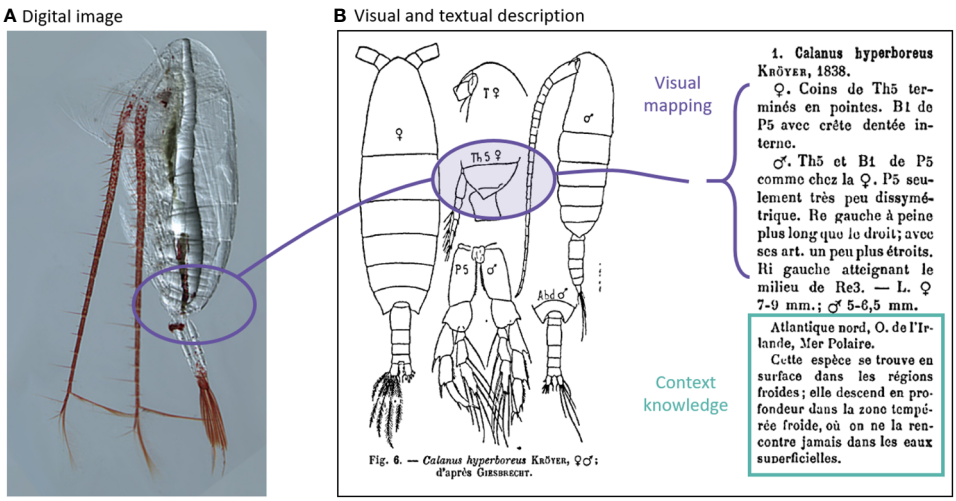


FIGURE 10
Multimodal learning uses both image and text data. **(A)** Image of *Calanus hyperboreus* (Source: Hopcroft at arcodiv.org). **(B)** Taxonomic description of *Calanus hyperboreus* (Rose, 1933).

Based on our discussion above, we recommend the following roadmap:

1. Evidence-based science for decision-making: Use all available plankton datasets to form a robust evidence base for decision-making. Collated and curated datasets will offer unprecedented opportunities to explore differences between collecting instruments. Moreover, large-scale intercomparable datasets can already be used to explore important ecological questions.
2. Technical validation: Enable long-term overlap of imaging and traditional techniques to secure continuity and quality control for high-quality continuous zooplankton monitoring and research.
3. Quality assurance: High-quality robust science demands high levels of self-consistency and peer-consistency. Routines to ensure consistency by humans and AI/ML need to be developed, and the adoption of XAI is required.
4. Interdisciplinary expertise: Invest in training in modern techniques for traditional taxonomists. Support workshops and collaboration between AI/ML and human taxonomists to offer (1) a way of exposing taxonomic experts to AI/ML data and (2) feedback from zooplankton researchers to instrument and AI/ML developers.
5. Capacity building: Invest in retaining taxonomists in the scientific community. Teach combined imaging/AI/taxonomy in university (currently taught independently and traditionally).

Data availability statement

The datasets presented in this article are not readily available because of confidentiality agreements. Requests to access the datasets should be directed to Sarah L. C. Giering (s.giering@noc.ac.uk).

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the National Oceanography Centre, UK. The patients/participants provided their written informed consent to participate in this study.

Author contributions

SG: survey design, data analysis and visualization, writing. PC: manuscript writing. DJ: manuscript writing. AM-G: survey design, writing. SP: manuscript writing. All authors contributed to the article and approved the submitted version.

Funding

SG was supported through the ANTICS project, receiving funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant Agreement No 950212), and through the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 101000858 (TechOceanS). AM-G is supported by the UK Natural Environment Research Council (NE/R002738/1) Knowledge Exchange fellowship scheme.

Acknowledgments

We would like to thank the editor and two anonymous reviewers for their valuable and constructive feedback.

Conflict of interest

PC is a director of Plankton Analytics Ltd., which supplied a Pi-10 plankton imager to CEFAS for evaluation.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest

Publisher's note

Author disclaimer

This output reflects only the author's view and the Research Executive Agency (REA) cannot be held responsible for any use that may be made of the information contained therein.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.986206/full#supplementary-material

SUPPLEMENTARY DATA SHEET 1
Questionnaire on future of zooplankton monitoring.

References

Alahmari, S. S., Goldgof, D. B., Mouton, P. R., and Hall, L. O. (2020). Challenges for the repeatability of deep learning models. *IEEE Access* 8, 211860–211868. doi: 10.1109/ACCESS.2020.3039833

Baltrušaitis, T., Ahuja, C., and Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* 41, 423–443. doi: 10.1109/TPAMI.2018.2798607

Basedow, S. L., Tande, K. S., Norrbin, M. F., and Kristiansen, S. A. (2013). Capturing quantitative zooplankton information in the sea: Performance test of laser optical plankton counter and video plankton recorder in a calanus finmarchicus dominated summer situation. *Prog. Oceanography* 108, 72–80. doi: 10.1016/j.pocean.2012.10.005

Batten, S. D., Clark, R., Flinkman, J., Hays, G., John, E., John, A. W. G., et al. (2003). CPR Sampling: the technical background, materials and methods, consistency and comparability. *Prog. Oceanography* 58, 193–215. doi: 10.1016/j.pocean.2003.08.004

Batten, S.D., Abu-Alhaija, R., Chiba, S., Edwards, M., Graham, G., Jyothibabu, R., et al (2019). A Global Plankton Diversity Monitoring Program. *Frontiers in Marine Science* 6. doi: 10.3389/fmars.2019.00321

Bedford, J., Johns, D., Greenstreet, S., and McQuatters-Gollop, A. (2018). Plankton as prevailing conditions: A surveillance role for plankton indicators within the marine strategy framework directive. *Mar. Policy* 89, 109–115. doi: 10.1016/j.marpol.2017.12.021

Bedford, J., Johns, D. G., and McQuatters-Gollop, A. (2020). Implications of taxon-level variation in climate change response for interpreting plankton lifeform biodiversity indicators. *ICES J. Mar. Sci.* 77, 3006–3015. doi: 10.1093/icesjms/fsaa183

Carvalho, D. V., Pereira, E. M., and Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics* 8, 832. doi: 10.3390/electronics8080832

Chen, W., Wang, W., Liu, L., and Lew, M. S. (2021). New ideas and trends in deep multimodal content understanding: A review. *Neurocomputing* 426, 195–215. doi: 10.1016/j.neucom.2020.10.042

Choquet, M., Kosobokova, K., Kwaśniewski, S., Hatlebakk, M., Dhanasiri, A. K. S., Melle, W., et al. (2018). Can morphology reliably distinguish between the copepods calanus finmarchicus and c. glacialis, or is DNA the only way? *Limnology Oceanography: Methods* 16, 237–252. doi: 10.1002/lom3.10240

Clutter, R. I., and Anraku, M. (1968). "Avoidance of samplers in zooplankton sampling," in *UNESCO Zooplankton sampling, part i. reviews on zooplankton sampling methods monographs on oceanographic methodology 2* (Paris: UNESCO), 57–76.

Convention on Biological Diversity (2011) *Strategic plan for biodiversity 2011-2020, including aichi biodiversity targets. secretariat of the convention on biological diversity*. Available at: https://www.cbd.int/sp/ (Accessed September 20, 2022).

Convention on Biological Diversity (2021) *First draft of the post-2020 global biodiversity. secretariat of the convention on biological diversity*. Available at: https://www.cbd.int/doc/c/abb5/591f/2e46096d3f0330b08ce87a45/wg2020-03-03-en.pdf (Accessed September 20, 2022).

Culverhouse, P. F. (2015). Biological oceanography needs new tools to automate sample analysis. *J. Mar. Biol. Aquaculture* 1, 0–0. doi: 10.15436/2381-0750.15.e002

Culverhouse, P. F., Macleod, N., Williams, R., Benfield, M. C., Lopes, R. M., and Picheral, M. (2014). An empirical assessment of the consistency of taxonomic identifications. *Mar. Biol. Res.* 10, 73–84. doi: 10.1080/17451000.2013.810762

Culverhouse, P., Williams, R., Reguera, B., Herry, V., and González-Gil, S. (2003). Do experts make mistakes? a comparison of human and machine identification of dinoflagellates. *Mar. Ecol. Prog. Ser.* 247, 17–25. doi: 10.3354/meps247017

Dai, J., Wang, R., Zheng, H., Ji, G., and Qiao, X. (2016). ZooplanktoNet: Deep convolutional network for zooplankton classification, in: OCEANS 2016 - Shanghai. *Presented at the OCEANS 2016 - Shanghai*, 1–6. doi: 10.1109/OCEANSAP.2016.7485680

Detmer, T. M., Broadway, K. J., Potter, C. G., Collins, S. F., Parkos, J. J., and Wahl, D. H. (2019). Comparison of microscopy to a semi-automated method (FlowCAM®) for characterization of individual-, population-, and community-level measurements of zooplankton. *Hydrobiologia* 838, 99–110. doi: 10.1007/s10750-019-03980-w

de Vargas, C., Le Bescot, N., Pollina, T., Henry, N., Romac, S., Colin, S., et al. (2022). Plankton planet: A frugal, cooperative measure of aquatic life at the planetary scale. *Front. Mar. Sci.*. 9. doi: 10.3389/fmars.2022.936972

Ellen, J. S., Graff, C. A., and Ohman, M. D. (2019). Improving plankton image classification using context metadata. *Limnology Oceanography: Methods* 17, 439–461. doi: 10.1002/lom3.10324

European Commission (2008) *Marine strategy framework directive 2008/56/EC — European environment agency. official journal of the European union*. Available at: https://www.eea.europa.eu/policy-documents/2008-56-ec (Accessed February 9, 2022).

Finlay, K., and Roff, J. (2004). Radiotracer determination of the diet of calanoid copepod nauplii and copepodites in a temperate estuary. *ICES J. Mar. Sci.* 61, 552–562. doi: 10.1016/j.icesjms.2004.03.010

Fleminger, A., and Clutter, R. I. (1965). Avoidance of towed nets by zooplankton. *Limnology Oceanography* 10, 96–104. doi: 10.4319/lo.1965.10.1.0096

Fleminger, A., and Hulsemann, K. (1977). Geographical range and taxonomic divergence in north Atlantic calanus (C. helgolandicus, c. finmarchicus and c. glacialis). *Mar. Biol.* 40, 233–248. doi: 10.1007/BF00390879

Frost, B. W. (1989). A taxonomy of the marine calanoid copepod genus pseudocalanus. *Can. J. Zool.* 67, 525–551. doi: 10.1139/z89-077

Gannon, J. E. (1975). Horizontal distribution of crustacean zooplankton along a cross-lake transect in lake Michigan. *J. Great Lakes Res.* 1, 79–91. doi: 10.1016/S0380-1330(75)72336-5

Giering, S. L. C., Cavan, E. L., Basedow, S. L., Briggs, N., Burd, A. B., Darroch, L. J., et al. (2020a). Sinking organic particles in the ocean–flux estimates from *in situ* optical devices. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00834

Giering, S. L. C., Hosking, B., Briggs, N., and Iversen, M. H. (2020b). The interpretation of particle size, shape, and carbon flux of marine particle images is strongly affected by the choice of particle detection algorithm. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00564

Giering, S. L. C., Wells, S. R., Mayers, K. M. J., Schuster, H., Cornwell, L., Fileman, E. S., et al. (2019). Seasonal variation of zooplankton community structure and trophic position in the celtic Sea: A stable isotope and biovolume spectrum approach. *Prog. Oceanography* 177, 101943–101943. doi: 10.1016/j.pocean.2018.03.012

Glikson, E., and Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical research. *Acad. Manage. Annals*. 14 (2): 627–660. doi: 10.5465/annals.2018.0057

Gorsky, G., Ohman, M. D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.-B., et al. (2010). Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.* 32, 285–303. doi: 10.1093/plankt/fbp124

Grosjean, P., Picheral, M., Warembourg, C., and Gorsky, G. (2004). Enumeration, measurement, and identification of net zooplankton samples using the ZOOSCAN digital imaging system. *ICES J. Mar. Sci.* 61, 518–525. doi: 10.1016/j.icesjms.2004.03.012

Hagras, H. (2018). Toward human-understandable, explainable AI. *Computer* 51, 28–36. doi: 10.1109/MC.2018.3620965

Hartley, M., and Olsson, T. S. G. (2020). dtoolAI: Reproducibility for deep learning. *Patterns* 1, 100073. doi: 10.1016/j.patter.2020.100073

Hoff, K. A., and Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors* 57, 407–434. doi: 10.1177/0018720814547570

Irisson, J.-O., Ayata, S.-D., Lindsay, D. J., Karp-Boss, L., and Stemmann, L. (2022). Machine learning for the study of plankton and marine snow from images. *Annu. Rev. Mar. Sci.* 14, 277–301. doi: 10.1146/annurev-marine-041921-013023

John, E. H., Batten, S. D., Stevens, D., Walne, A. W., Jonas, T. J., and Hays, G. C. (2002). Continuous plankton records stand the test of time: Evaluation of flow rates, clogging and the continuity of the CPR time-series. *J. Plankton Res.* 24, 941–946. doi: 10.1093/plankt/24.9.941

Kerr, T., Clark, J. R., Fileman, E. S., Widdicombe, C. E., and Pugeault, N. (2020). Collaborative Deep Learning Models to Handle Class Imbalance in FlowCam Plankton Imagery *IEEE*, 170013–170032. doi: 10.1109/ACCESS.2020.3022242

Lenz, P. H., Lieberman, B., Cieslak, M. C., Roncalli, V., and Hartline, D. K. (2021). Transcriptomics and metatranscriptomics in zooplankton: Wave of the future? *J. Plankton Res.* 43, 3–9. doi: 10.1093/plankt/fbaa058

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., et al. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00196

MacLeod, N., Benfield, M., and Culverhouse, P. (2010). Time to automate identification. *Nature* 467, 154–155. doi: 10.1038/467154a

McQuatters-Gollop, A., Atkinson, A., Aubert, A., Bedford, J., Best, M., Bresnan, E., et al. (2019). Plankton lifeforms as a biodiversity indicator for regional-scale assessment of pelagic habitats for policy. *Ecol. Indic.* 101, 913–925. doi: 10.1016/j.ecolind.2019.02.010

McQuatters-Gollop, A., Edwards, M., Helaouët, P., Johns, D. G., Owens, N. J. P., Raitsos, D. E., et al. (2015). The continuous plankton recorder survey: How can long-term phytoplankton datasets contribute to the assessment of good environmental status? *Estuarine Coast. Shelf Sci.* 162, 88–97. doi: 10.1016/j.ecss.2015.05.010

McQuatters-Gollop, A., Guérin, L., Arroyo, N. L., Aubert, A., Artigas, L. F., Bedford, J., et al. (2022). Assessing the state of marine biodiversity in the northeast Atlantic. *Ecol. Indic.* 141, 109148. doi: 10.1016/j.ecolind.2022.109148

McQuatters-Gollop, A., Johns, D. G., Bresnan, E., Skinner, J., Rombouts, I., Stern, R., et al. (2017). From microscope to management: The critical value of plankton taxonomy to marine policy and biodiversity conservation. *Mar. Policy* 83, 1–10. doi: 10.1016/j.marpol.2017.05.022

Nogueira, E., González-Nuevo, G., Bode, A., Varela, M., Morán, X. A. G., and Valdés, L. (2004). Comparison of biomass and size spectra derived from optical plankton counter data and net samples: application to the assessment of mesoplankton distribution along the Northwest and north Iberian shelf. *ICES J. Mar. Sci.* 61, 508–517. doi: 10.1016/j.icesjms.2004.03.018

O'Brien, T. D., Lorenzoni, L., Isensee, K., and Valdés, L. (2017). "What are marine ecological time series telling us about the ocean?," in *A status report* (Paris, France: IOC-UNESCO).

O'Brien, T.D., and Oakes, S.A., 2020. Visualizing and Exploring Zooplankton Spatio-Temporal Variability, in: *Zooplankton Ecology*. (Boca Raton: CRC Press) 292.

Orenstein, E. C., Ayata, S. D., Maps, F., Becker, É. C., Benedetti, F., Biard, T., et al (2022). Machine learning techniques to characterize functional traits of plankton from image data *Limnol Oceanogr*, 67, 1647–1669. doi: 10.1002/lno.12101

OSPAR (2017) *Changes in phytoplankton biomass and zooplankton, intermediate assessment 2017*. Available at: https://oap.ospar.org/en/ospar-assessments/intermediate-assessment-2017/biodiversity-status/habitats/plankton-biomass/ (Accessed March 4, 2022).

Owens, N. J. P., Hosie, G. W., Batten, S. D., Edwards, M., Johns, D. G., and Beaugrand, G. (2013). All plankton sampling systems underestimate abundance: Response to "Continuous plankton recorder underestimates zooplankton abundance" by J.W. dippner and m. Krause. *J. Mar. Syst.* 128, 240–242. doi: 10.1016/j.jmarsys.2013.05.003

Petersen, W., and Colijn, F. (2017). FerryBox whitebook. *EuroGOOS AIBL* 44. doi: 10.25607/OBP-1002

Picheral, M., Catalano, C., Brousseau, D., Claustre, H., Coppola, L., Leymarie, E., et al. (2021). The underwater vision profiler 6: an imaging sensor of particle size spectra and plankton, for autonomous and cabled platforms. *Limnology Oceanography: Methods*. 20, 115–129 doi: 10.1002/lom3.10475

Picheral, M., Colin, S., and Irisson, J.-O. (2017) *EcoTaxa, a tool for the taxonomic classification of images*. Available at: http://ecotaxa.obs-vlfr.fr.

Picheral, M., Guidi, L., Stemmann, L., Karl, D. M., Iddaoud, G., and Gorsky, G. (2010). The underwater vision profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnology Oceanography: Methods* 8, 462–473. doi: 10.4319/lom.2010.8.462

Pitois, S. G., Bouch, P., Creach, V., and van der Kooij, J. (2016). Comparison of zooplankton data collected by a continuous semi-automatic sampler (CALPS) and a traditional vertical ring net. *J. Plankton Res.* 38, 931–943. doi: 10.1093/plankt/fbw044

Pitois, S. G., Graves, C. A., Close, H., Lynam, C., Scott, J., Tilbury, J., et al. (2021). A first approach to build and test the copepod mean size and total abundance (CMSTA) ecological indicator using *in-situ* size measurements from the plankton imager (PI). *Ecol. Indic.* 123, 107307. doi: 10.1016/j.ecolind.2020.107307

Pitois, S. G., Tilbury, J., Bouch, P., Close, H., Barnett, S., and Culverhouse, P. F. (2018). Comparison of a cost-effective integrated plankton sampling and imaging instrument with traditional systems for mesozooplankton sampling in the celtic Sea. *Front. Mar. Sci.* 5. doi: 10.3389/fmars.2018.00005

Re3Data.org (2021). re3data.org: Integrated Marine Observing System Ocean Data Portal; editing status 2021-11-10; re3data.org - Registry of Research Data Repositories. doi: 10.17616/R35057

R Core Team (2018). *R: A language and environment for statistical computing* (R Foundation for Statistical Computing: R Foundation for Statistical Computing). Available at: https://www.r-project.org/.

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., et al. (2019). On the future of argo: A global, full-depth, multi-disciplinary array. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00439

Rombouts, I., Simon, N., Aubert, A., Cariou, T., Feunteun, E., Guérin, L., et al. (2019). Changes in marine phytoplankton diversity: Assessment under the marine strategy framework directive. *Ecol. Indic.* 102, 265–277. doi: 10.1016/j.ecolind.2019.02.009

Rose, M. (1933). *Faune de France 26, copepodes pelagique* (Paris: P. Lechevalier).

Sameoto, D., Wiebe, P., Runge, J., Postel, L., Dunn, J., Miller, C., et al. (2000). "3 - collecting zooplankton," in *ICES zooplankton methodology manual*. Eds. R. Harris, P. Wiebe, J. Lenz, H. R. Skjoldal and M. Huntley (London: Academic Press), 55–81. doi: 10.1016/B978-012327645-2/50004-9

Schepman, A., and Rodway, P. (2020). Initial validation of the general attitudes towards artificial intelligence scale. *Comput. Hum. Behav. Rep.* 1, 100014. doi: 10.1016/j.chbr.2020.100014

Schröder, S.-M., Kiko, R., and Koch, R. (2020) *MorphoCluster: Efficient annotation of plankton images by clustering. arXiv:2005.01595 [cs]*. Available at: http://arxiv.org/abs/2005.01595 (Accessed December 9, 2021).

Scott, J., Pitois, S., Close, H., Almeida, N., Culverhouse, P., Tilbury, J., et al. (2021). *In situ* automated imaging, using the plankton imager, captures temporal variations in mesozooplankton using the celtic Sea as a case study. *J. Plankton Res.* 43, 300–313. doi: 10.1093/plankt/fbab018

Skjoldal, H. R., Wiebe, P. H., Postel, L., Knutsen, T., Kaartvedt, S., and Sameoto, D. D. (2013). Intercomparison of zooplankton (net) sampling systems: Results from the ICES/GLOBEC sea-going workshop. *Prog. Oceanography* 108, 1–42. doi: 10.1016/j.pocean.2012.10.006

UK Parliament and (House of Lords) Artificial Intelligence Committee (2017) *AI In the UK: Ready, willing and able*? Available at: https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm (Accessed June 15, 2022).

Ullman, D., and Malle, B. F. (2017). "Human-robot trust: Just a button press away," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction HRI '17*. Association for Computing Machinery 309–310 (New York, NY, USA: Association for Computing Machinery). doi: 10.1145/3029798.3038423

UN General Assembly (2015) *Transforming our world : the 2030 agenda for sustainable development*. Available at: http://bit.ly/TransformAgendaSDG-pdf (Accessed September 20, 2022).

Uppal, S., Bhagat, S., Hazarika, D., Majumder, N., Poria, S., Zimmermann, R., et al. (2022). Multimodal research in vision and language: A review of current and emerging trends. *Inf. Fusion* 77, 149–171. doi: 10.1016/j.inffus.2021.07.009

Wang, C., Zheng, X., Guo, C., Yu, Z., Yu, J., Zheng, H., et al (2018). Transferred Parallel Convolutional Neural Network for Large Imbalanced Plankton Database Classification, in: *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO). Presented at the 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, 1–5. doi: 10.1109/OCEANSKOBE.2018.8558836

Wiebe, P. H., and Benfield, M. C. (2003). From the hensen net toward four-dimensional biological oceanography. *Prog. Oceanography* 56, 7–136. doi: 10.1016/S0079-6611(02)00140-4

Wilson, R. J., Speirs, D. C., and Heath, M. R. (2015). On the surprising lack of differences between two congeneric calanoid copepod species, calanus finmarchicus and c. helgolandicus. *Prog. Oceanography* 134, 413–431. doi: 10.1016/j.pocean.2014.12.008

Wootton, M., and Johns, D. (2019) *Zooplankton ring test 2018/2019. Plymouth: Marine biological analytical quality control scheme*. Available at: http://www.nmbaqcs.org/scheme-components/zooplankton/reports/.

Xue, Y., Chen, H., Yang, J. R., Liu, M., Huang, B., and Yang, J. (2018). Distinct patterns and processes of abundant and rare eukaryotic plankton communities following a reservoir cyanobacterial bloom. *ISME J.* 12, 2263–2277. doi: 10.1038/s41396-018-0159-0

Yamada, T., Prügel-Bennett, A., and Thornton, B. (2021). Learning features from georeferenced seafloor imagery with location guided autoencoders. *J. Field Robotics* 38, 52–67. doi: 10.1002/rob.21961

Zingone, A., Harrison, P. J., Kraberg, A., Lehtinen, S., McQuatters-Gollop, A., O'Brien, T., et al. (2015). Increasing the quality, comparability and accessibility of phytoplankton species composition time-series data. *Estuarine Coast. Shelf Sci.* 162, 151–160. doi: 10.1016/j.ecss.2015.05.024

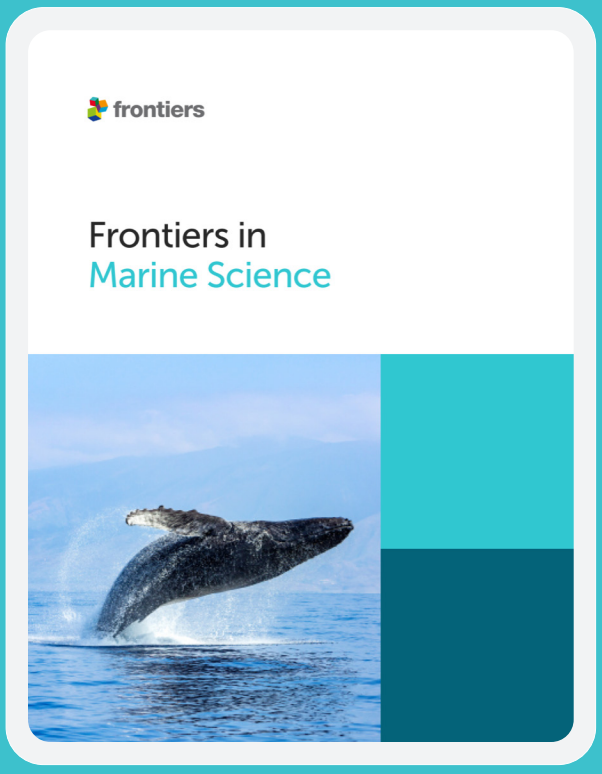
Frontiers in Marine Science

Explores ocean-based solutions for emerging global challenges

The third most-cited marine and freshwater biology journal, advancing our understanding of marine systems and addressing global challenges including overfishing, pollution, and climate change.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact