ORIGINAL RESEARCH article

Front. Mar. Sci., 09 August 2022

Sec. Ocean Observation

Volume 9 - 2022 | https://doi.org/10.3389/fmars.2022.894372 (https://doi.org/10.3389/fmars.2022.894372)

This article is part of the Research Topic

Technological Advances for Measuring Planktonic Components of the Pelagic Ecosystem: An Integrated Approach to Data Collection and Analysis

View all 14 Articles (/research-topics/25191#articles)

Global Distribution of Zooplankton Biomass Estimated by *In Situ* Imaging and Machine Learning

Laetitia Drago (https://www.frontiersin.org/people/u/1476606)[1*], Thelma Panaïotis (https://www.frontiersin.org/people/u/1498153)[1], Jean-Olivier Irisson[1], Marcel Babin (https://www.frontiersin.org/people/u/557586)[2], Tristan Biard (https://www.frontiersin.org/people/u/738204)[3], François Carlotti (https://www.frontiersin.org/people/u/718065)[4,5], Laurent Coppola[1,6], Lionel Guidi[1], Helena Hauss (https://www.frontiersin.org/people/u/320569)[7], Lee Karp-Boss (https://www.frontiersin.org/people/u/613088)[8], Fabien Lombard[1,9]

[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

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

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).

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

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, 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

Figure 1

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g001.jpg)

FIGURE 1 Map of the UVP5 dataset used in this study. Transparency was used to illustrate the density of points on the map.

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

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.

Table 1


www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-t001.jpg)

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.

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 used. The capsule of

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. 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

FIGURE 2 Methodology followed from data selection to prediction of global biomass.

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 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

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

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

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:

$$A = \frac{\Pi}{180} \times \mathscr{R}^2 \times (\sin(lat_S) - sin(lat_N)) \times 10^6$$

with the area A in m^2, the south and north latitudinal limits of the cell in radians and $\mathscr{R}$, 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.

(/articles/10.3389/fmars.2022.894372/pdf)

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

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

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.

Figure 4 (/articles/10.3389/fmars.2022.894372/pdf)

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.

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

Figure 5

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g005.jpg)

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

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g006.jpg)

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.

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 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 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

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.

Figure 7


www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g007.jpg)

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.

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

Figure 8

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g008.jpg)

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.

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 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).

Figure 9

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-

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.

Figure 10

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.

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

Figure 11

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-g011.jpg)

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.

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 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

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

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

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

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

Table 2

 www.frontiersin.org (https://www.frontiersin.org/files/Articles/894372/fmars-09-894372-HTML/image_m/fmars-09-894372-t002.jpg)

TABLE 2 Comparison of global biomass estimates in the literature.

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 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%

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_distribu-

(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.

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.

manufacturer, is not guaranteed or endorsed by the publisher.

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 (https://www.frontiersin.org/articles/10.3389/fmars.2022.894372/full#supplementary-material)

References

Amante C., 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].

CrossRef Full Text (https://doi.org/10.7289/V5C8276M) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+Amante&author=B.%20W.+Eakins&publication_year=2009&book=ETOPO1+1+Arc-Minute+Global+Relief+Model:+Procedures,+Data+Sources+and+Analysis.+NOAA+Technical+Memorandum+NESDIS+NGDC-24&)

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

(/articles/10.3389/fmars.2022.894372/pdf)

Ocean. *Global Change Biol.* 15, 1790–1803. doi: 10.1111/j.1365-2486.2009.01848.x

CrossRef Full Text (https://doi.org/10.1111/j.1365-2486.2009.01848.x) | Google Scholar (http://scholar.google.com/scholar_lookup?author=G.+Beaugrand&author=C.+Luczak&author=M.+Edwards&publication_year=2009&title=Rapid%20Biogeographical%20Plankton%20Shifts%20in%20the%20North%20Atlantic%20Ocean&journal=Global+Change+Biol.&volume=15&pages=1790)

Beers J. R., 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

CrossRef Full Text (https://doi.org/10.4319/lo.1970.15.5.0825) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20R.+Beers&author=G.%20L.+Stewart&publication_year=1970&title=The%20Preservation%20of%20Acantharians%20in%20Fixed%20Plankton%20Samples&journal=Limnol.+Oceanogr.&volume=15&pages=825)

Bernstein R. E., Betzer P. R., Feely R. A., Byrne R. H., Lamb M. F., 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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/17816791/) | CrossRef Full Text (https://doi.org/10.1126/science.237.4821.1490) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.%20E.+Bernstein&author=P.%20R.+Betzer&author=R.%20A.+Feely&author=R.%20H.+Byrne&author=M.%20F.+Lamb&author=A.%20F.+Michaels&publication_year=1987&title=Acantharian%20Fluxes%20and%20Strontium%20to%20Chlorinity%20Ratios%20in%20the%20North%20Pacific%20Ocean&journal=Science&volume=237&pages=1490)

Biard T., 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

CrossRef Full Text (https://doi.org/10.1002/lno.11472) | Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Biard&author=M.%20D.+Ohman&publication_year=2020&title=Vertical%20Niche%20Definition%20of%20Test-Bearing%20Protists%20%28Rhizaria%29%20Into%20the%20Twilight%20Zone%20Revealed%20by%20in%20Situ%20Imaging&journal=Limnol.+Oceanogr.&volume=65&pages=2583)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/27096373/) | CrossRef Full Text (https://doi.org/10.1038/nature17652) | Google Scholar (http://scholar.google.com/scholar_lookup?

CrossRef Full Text (https://doi.org/10.1007/978-3-319-28149-0_19) | Google Scholar (http://scholar.google.com/scholar_lookup?author=D.+Boltovskoy&author=O.%20R.+Anderson&author=N.%20M.+Correa&publication_year=2017&title=Radiolaria%20and%20Phaeodaria&book=Handbook+of+the+Protists&pages=731)

Bottazzi E. M., 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

CrossRef Full Text (https://doi.org/10.1093/plankt/4.4.757) | Google Scholar (http://scholar.google.com/scholar_lookup?author=E.%20M.+Bottazzi&author=M.%20G.+Andreoli&publication_year=1982&title=Distribution%20of%20adult%20and%20juvenile%20Acantharia%20%28Protozoa%20Sarcodina%29%20in%20the%20Atlantic%20Ocean&journal=Journal+of+Plankton+Research&volume=4&pages=757)

Breiman L., Friedman J. H., Olshen R. A., Stone C. J. (1984). *Classification and Regression Trees* (Wadsworth & Brooks. Cole Statistics/Probability Series).

Google Scholar (http://scholar.google.com/scholar_lookup?author=L.+Breiman&author=J.%20H.+Friedman&author=R.%20A.+Olshen&author=C.%20J.+Stone&publication_year=1984&book=Classification+and+Regression+Trees&)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/32336785/) | CrossRef Full Text (https://doi.org/10.4319/lo.1967.12.3.0451._eprint:https://onlinelibrary.wiley.com/doi/pdf/10.4319/lo.1967.12.3.0451) | Google Scholar (http://scholar.google.com/scholar_lookup?author=E.+Brinton&publication_year=1967&title=Vertical%20Migration%20and%20Avoidance%20Capability%20of%20Euphausiids%20in%20the%20California%20Current&journal=Limnol.+Oceanogr.&volume=12&pages=451)

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 (https://www.researchgate.net/publication/322665037_An_unusual_gelatinous_plankton_event_in_the_NE_Pacific_The_Great_Pyrosome_Bloom_of_2017)

Google Scholar (http://scholar.google.com/scholar_lookup?author=R.+Brodeur&author=I.+Perry&author=J.+Boldt&author=L.+Flostrand&author=M.+Galbraith&author=J.+King&publication_year=2018&book=An+Unusual+Gelatinous+Plankton+Event+in+the+NE+Pacific:+The+Great+Pyrosome+Bloom+of+2017&volume=26&pages=6)

20on%20Processes%20That%20Control%20Particle%20Export%20and%20Flux%20Attenuation%20in%20the%20Twilight%20Zone%20of%20the%20Open%20Ocean&journal=Limnol.+Oceanogr.&volume=54&pages=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

CrossRef Full Text (https://doi.org/10.5194/essd-5-227-2013) | Google Scholar (http://scholar.google.com/scholar_lookup?author=E.%20T.+Buitenhuis&author=M.+Vogt&author=R.+Moriarty&author=N.+Bednar%C5%A1ek&author=S.%20C.+Doney&author=K.+Leblanc&publication_year=2013&title=MAREDAT%3A%20Towards%20a%20World%20Atlas%20of%20MARine%20Ecosystem%20DATa&journal=Earth+Syst.+Sci.+Data&volume=5&pages=227)

Castellanos I., Suárez-Morales E., Álvaro Morales-Ramírez Ã. (2009). "Euphausiids," in *Marine Biodiversity of Costa Rica, Central America*, vol. 86. Eds. Wehrtmann I. S., Cortés J. (Dordrecht: Springer Netherlands), 199–207. doi: 10.1007/978-1-4020-8278-818

CrossRef Full Text (https://doi.org/10.1007/978-1-4020-8278-818) | Google Scholar (http://scholar.google.com/scholar_lookup?author=I.+Castellanos&author=E.+Su%C3%A1rez-Morales&author=%C3%83.+%C3%81lvaro%20Morales-Ram%C3%ADrez&publication_year=2009&title=Euphausiids&book=Marine+Biodiversity+of+Costa+Rica,+Central+America&volume=86&pages=199)

Chavez F. P., Bertrand A., Guevara-Carrasco R., Soler P., 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

CrossRef Full Text (https://doi.org/10.1016/j.pocean.2008.10.012) | Google Scholar (http://scholar.google.com/scholar_lookup?author=F.%20P.+Chavez&author=A.+Bertrand&author=R.+Guevara-Carrasco&author=P.+Soler&author=J.+Csirke&publication_year=2008&title=The%20Northern%20Humboldt%20Current%20System%3A%20Brief%20History%2C%20Present%20Status%20and%20a%20View%20Towards%20the%20Future&journal=Prog.+Oceanogr.&volume=79&pages=95)

Chen T., He T., Benesty M., Khotilovich V., Tang Y., Cho H., et al. (2021). *Xgboost: Extreme Gradient Boosting*.

Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Chen&author=T.+He&author=M.+Benesty&author=V.+Khotilovich&author=Y.+Tang&author=H.+Cho&publication_year=2021&book=Xgboost:+Extreme+Gradient+Boosting&)

Chen B., Liu H., Xiao W., Wang L., Huang B. (2020). A Machine Learning Approach to Modeling

Flux Interception in Oceanic Mesoscale Eddies by the Polychaete Poeobius Sp. *Limnol. Oceanogr.* 63, 2093–2109. doi: 10.1002/lno.10926

CrossRef Full Text (https://doi.org/10.1002/lno.10926) | Google Scholar (http://scholar.google.com/scholar_lookup?author=S.+Christiansen&author=H.+Hoving&author=F.+Sch%C3%BCtte&author=H.+Hauss&author=J.+Karstensen&author=A.+K%C3%B6rtzinger&publication_year=2018&title=Particulate%20Matter%20Flux%20Interception%20in%20Oceanic%20Mesoscale%20Eddies%20by%20the%20Polychaete%20Poeobius%20Sp&journal=Limnol.+Oceanogr.&volume=63&pages=2093)

Cowen R. K., 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

CrossRef Full Text (https://doi.org/10.4319/lom.2008.6.126) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.%20K.+Cowen&author=C.%20M.+Guigand&publication_year=2008&title=In%20Situ%20Ichthyoplankton%20Imaging%20System%20%28I%20SIIS%29%3A%20System%20Design%20and%20Preliminary%20Results%3A%20In%20Situ%20Ichthyoplankton%20Imaging%20System&journal=Limnol.+Oceanogr.:+Methods&volume=6&pages=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

CrossRef Full Text (https://doi.org/10.1016/j.ecoinf.2007.07.001) | Google Scholar (http://scholar.google.com/scholar_lookup?author=P.%20F.+Culverhouse&publication_year=2007&title=Human%20and%20Machine%20Factors%20in%20Algae%20Monitoring%20Performance&journal=Ecol.+Inf.&volume=2&pages=361)

Culverhouse P. F., Macleod N., Williams R., Benfield M. C., Lopes R. M., Picheral M. (2014). An Empirical Assessment of the Consistency of Taxonomic Identifications. *Mar. Biol. Res.* 10, 73–84. doi: 10.1080/17451000.2013.810762

CrossRef Full Text (https://doi.org/10.1080/17451000.2013.810762) | Google Scholar (http://scholar.google.com/scholar_lookup?author=P.%20F.+Culverhouse&author=N.+Macleod&author=R.+Williams&author=M.%20C.+Benfield&author=R.%20M.+Lopes&author=M.+Picheral&publication_year=2014&title=An%20Empirical%20Assessment%20of%20the%20Consistency%20of%20Taxonomic%20Identifications&journal=Mar.+Biol.+Res.&volume=10&pages=73)

Dai L., Li C., Yang G., 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., Solow A. R. (1992). Microaggregations of Oceanic Plankton Observed by Towed Video Microscopy. *Science* 257, 230–232. doi: 10.1126/science.257.5067.230

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/17794756/) | CrossRef Full Text (https://doi.org/10.1126/science.257.5067.230) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.%20S.+Davis&author=S.%20M.+Gallager&author=A.%20R.+Solow&publication_year=1992&title=Microaggregations%20of%20Oceanic%20Plankton%20Observed%20by%20Towed%20Video%20Microscopy&journal=Science&volume=257&pages=230)

Davis C. S., Thwaites F. T., Gallager S. M., 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

CrossRef Full Text (https://doi.org/10.4319/lom.2005.3.59) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.%20S.+Davis&author=F.%20T.+Thwaites&author=S.%20M.+Gallager&author=Q.+Hu&publication_year=2005&title=A%20Three-Axis%20Fast-Tow%20Digital%20Video%20Plankton%20Recorder%20for%20Rapid%20Surveys%20of%20Plankton%20Taxa%20and%20Hydrography%3A%20New%20Video%20Plankton%20Recorder&journal=Limnol.+Oceanogr.:+Methods&volume=3&pages=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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/17489472/) | CrossRef Full Text (https://doi.org/10.1890/0012-9658(2007)88[243:BTFEMA]2.0.CO;2) | Google Scholar (http://scholar.google.com/scholar_lookup?author=G.+De%E2%80%99ath&publication_year=2007&title=Boosted%20Trees%20for%20Ecological%20Modeling%20and%20Prediction&journal=Ecology&volume=88&pages=243)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/23326463/) | CrossRef Full Text (https://doi.org/10.1371/journal.pone.0053598) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Decelle&author=P.+Martin&author=K.+Paborstava&author=D.%20W.+Pond&author=G.+Tarling&author=F.+Mah%C3%A9&publication_year=2013&title=Diversity%2C%20Ecology%20and%20Biogeochemistry%20of%20Cyst-Forming%20Acantharia%20%28Radiolaria%29%20in%20the%20Oceans&journal=PLoS+One&volume=8&pages=e53598)

(/articles/10.3389/fmars.2022.894372/pdf)

CrossRef Full Text (https://doi.org/10.1126/science.1261605) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+de%20Vargas&author=S.+Audic&author=N.+Henry&author=J.+Decelle&author=F.+Mahe&author=R.+Logares&publication_year=2015&title=Eukaryotic%20plankton%20diversity%20in%20the%20sunlit%20ocean&journal=Science&volume=348&pages=1261605.5)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/34059795/) | CrossRef Full Text (https://doi.org/10.1038/s42003-021-02159-1) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.+Edwards&author=P.+H%C3%A9laou%C3%ABt&author=E.+Goberville&author=A.+Lindley&author=G.%20A.+Tarling&author=M.%20T.+Burrows&publication_year=2021&title=North%20Atlantic%20Warming%20Over%20Six%20Decades%20Drives%20Decreases%20in%20Krill%20Abundance%20With%20No%20Associated%20Range%20Shift&journal=Commun.+Biol.&volume=4&pages=644)

Elith J., 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

CrossRef Full Text (https://doi.org/10.1111/j.1600-0587.2008.05505.x) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Elith&author=C.%20H.+Graham&publication_year=2009&title=Do%20They%3F%20How%20do%20They%3F%20WHY%20do%20They%20Differ%3F%20On%20Finding%20Reasons%20for%20Differing%20Performances%20of%20Species%20Distribution%20Models&journal=Ecography&volume=32&pages=66)

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

CrossRef Full Text (https://doi.org/10.1111/j.2006.0906-7590.04596.x) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Elith&author=C.%20H.+Graham&author=R.%20P.+Anderson&author=M.+Dud%C3%ADk&author=S.+Ferrier&author=A.+Guisan&publication_year=2006&title=Novel%20Methods%20Improve%20Prediction%20of%20Species%E2%80%99%20Distributions%20From%20Occurrence%20Data&journal=Ecography&volume=29&pages=129)

Elith J., 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

(https://doi.org/10.1111/j.1365-2656.2008.01390.x) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Elith&author=J.%20R.+Leathwick&author=T.+Hastie&publication_year=2008&title=A%20Working%20Guide%20to%20Boosted%20Regression%20Trees&journal=J.+Anim.+Ecol.&volume=77&pages=802)

Faillettaz R., Picheral M., Luo J. Y., Guigand C., Cowen R. K., 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

CrossRef Full Text (https://doi.org/10.1016/j.mio.2016.04.003) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.+Faillettaz&author=M.+Picheral&author=J.%20Y.+Luo&author=C.+Guigand&author=R.%20K.+Cowen&author=J.-O.+Irisson&publication_year=2016&title=Imperfect%20Automatic%20Image%20Classification%20Successfully%20Describes%20Plankton%20Distribution%20Patterns&journal=Methods+Oceanogr.&volume=15-16&pages=60)

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

CrossRef Full Text (https://doi.org/10.5194/bg-9-1301-2012) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.+Forest&author=L.+Stemmann&author=M.+Picheral&author=L.+Burdorf&author=D.+Robert&author=L.+Fortier&publication_year=2012&title=Size%20Distribution%20of%20Particles%20and%20Zooplankton%20Across%20the%20Shelf-Basin%20System%20in%20Southeast%20Beaufort%20Sea%3A%20Combined%20Results%20From%20an%20Underwater%20Vision%20Profiler%20and%20Vertical%20Net%20Tows&journal=Biogeosciences&volume=9&pages=1301)

Friedman J. H. (2002). Stochastic Gradient Boosting. *Comput. Stat Data Anal.* 38, 367–378. doi: 10.1016/S0167-9473(01)00065-2

CrossRef Full Text (https://doi.org/10.1016/S0167-9473(01)00065-2) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20H.+Friedman&publication_year=2002&title=Stochastic%20Gradient%20Boosting&journal=Comput.+Stat+Data+Anal.&volume=38&pages=367)

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

CrossRef Full Text (https://doi.org/10.1093/plankt/23.12.1421) | Google Scholar (http://scholar.google.com/scholar_lookup?

author=H.%20E.+Garcia&author=T.P.+Boyer&author=O.%20K.+Baranova&author=R.%20A.+Locarnini&author=A.V.+Mishonov&author=A.+Grodsky&publication_year=2019&book=World+Ocean+Atlas+2018:+Product+Documentation&)

González Carman V., Piola A., O'Brien T. D., Tormosov D. D., 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

CrossRef Full Text (https://doi.org/10.1016/j.pocean.2019.102123) | Google Scholar (http://scholar.google.com/scholar_lookup?author=V.+Gonz%C3%A1lez%20Carman&author=A.+Piola&author=T.%20D.+O%E2%80%99Brien&author=D.%20D.+Tormosov&author=E.%20M.+Acha&publication_year=2019&title=Circumpolar%20Frontal%20Systems%20as%20Potential%20Feeding%20Grounds%20of%20Southern%20Right%20Whales&journal=Prog.+Oceanogr.&volume=176&)

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

CrossRef Full Text (https://doi.org/10.1093/plankt/fbp124) | Google Scholar (http://scholar.google.com/scholar_lookup?author=G.+Gorsky&author=M.%20D.+Ohman&author=M.+Picheral&author=S.+Gasparini&author=L.+Stemmann&author=J.-B.+Romagnan&publication_year=2010&title=Digital%20Zooplankton%20Image%20Analysis%20Using%20the%20ZooScan%20Integrated%20System&journal=J.+Plankt.+Res.&volume=32&pages=285)

Griffith A. W., 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

CrossRef Full Text (https://doi.org/10.1016/j.hal.2019.03.008) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.%20W.+Griffith&author=C.%20J.+Gobler&publication_year=2020&title=Harmful%20Algal%20Blooms%3A%20A%20Climate%20Change%20Co-Stressor%20in%20Marine%20and%20Freshwater%20Ecosystems&journal=Harmf.+Algae.&volume=91&pages=101590)

Grossmann M. M., Gallager S. M., 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

CrossRef Full Text (https://doi.org/10.1007/s10872-014-0268-y) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.%20M.+Grossmann&author=S.%20M.+Gallager&author=S.+Mitarai&publication_year=2015&title=Continuous%20Monitoring%20of%20Near-Bottom%20Mesoplankton%20Communities%20in%20the%20East%20China%20Sea%20During%20a%20

ustre&author=L.+Legendre&publication_year=2009&title=Effects%20of%20Phytoplankton%20Community%20on%20Production%2C%20Size%2C%20and%20Export%20of%20Large%20Aggregates%3A%20A%20World-Ocean%20Analysis&journal=Limnol.+Oceanogr.&volume=54&pages=1951)

Guihen D., Brearley J. A., 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

CrossRef Full Text (https://doi.org/10.1016/j.dsr.2021.103680) | Google Scholar (http://scholar.google.com/scholar_lookup?author=D.+Guihen&author=J.%20A.+Brearley&author=S.+Fielding&publication_year=2022&title=Antarctic%20Krill%20Likely%20Avoid%20Underwater%20Gliders&journal=Deep.+Sea.+Res.+Part+I.:+Oceanogr.+Res.+Pap.&volume=179&)

Guisan A., Zimmermann N. E. (2000). Predictive Habitat Distribution Models in Ecology. *Ecol. Model.* 135, 147–186. doi: 10.1016/S0304-3800(00)00354-9

CrossRef Full Text (https://doi.org/10.1016/S0304-3800(00)00354-9) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.+Guisan&author=N.%20E.+Zimmermann&publication_year=2000&title=Predictive%20Habitat%20Distribution%20Models%20in%20Ecology&journal=Ecol.+Model.&volume=135&pages=147)

Hastie T., Tibshirani R. (1986). General Additive Models 3, 297–318. doi: 10.1214/ss/1177013604

CrossRef Full Text (https://doi.org/10.1214/ss/1177013604) | Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Hastie&author=R.+Tibshirani&publication_year=1986&title=General%20Additive%20Models&volume=3&pages=297)

Hastie T., Tibshirani R., Friedman J. (2001). *The Elements of Statistical Learning*. 2nd ed. (Springer).

Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Hastie&author=R.+Tibshirani&author=J.+Friedman&publication_year=2001&book=The+Elements+of+Statistical+Learning&)

Hatton I. A., Heneghan R. F., Bar-On Y. M., Galbraith E. D. (2021). *The Global Ocean Size-Spectrum From Bacteria to Whales. Preprint, Ecology*. doi: 10.1101/2021.04.03.438320

CrossRef Full Text (https://doi.org/10.1101/2021.04.03.438320) | Google Scholar (http://scholar.google.com/scholar_lookup?author=I.%20A.+Hatton&author=R.%20F.+Heneghan&author=Y.%20M.+Bar-On&author=E.%20D.+Galbraith&publication_year=2021&book=The+Global+Ocean+Size-Spectrum+From+Bacteria+to+Whales.+Preprint,+Ecology&)

Hays G., Richardson A., Robinson C. (2005). Climate Change and Marine Plankton. *Trends Ecol. Evol.* 20,

Google Scholar (http://scholar.google.com/scholar_lookup?author=R.%20J.+Hijmans&publication_year=2021&book=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

CrossRef Full Text (https://doi.org/10.5194/os-15-1327-2019) | Google Scholar (http://scholar.google.com/scholar_lookup?author=H.-J.+Hoving&author=S.+Christiansen&author=E.+Fabrizius&author=H.+Hauss&author=R.+Kiko&author=P.+Linke&publication_year=2019&title=The%20Pelagic%20In%20Situ%20Observation%20System%20%28PELAGIOS%29%20to%20Reveal%20Biodiversity%2C%20Behavior%2C%20and%20Ecology%20of%20Elusive%20Oceanic%20Fauna&journal=Ocean.+Sci.&volume=15&pages=1327)

Hu C., Xu J., Li X., Shi Z., 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

CrossRef Full Text (https://doi.org/10.1016/j.scitotenv.2021.146315) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+Hu&author=J.+Xu&author=X.+Li&author=Z.+Shi&author=R.+Li&publication_year=2021&title=Environmental%20Regulations%20on%20Bacterial%20Abundance%20in%20the%20South%20China%20Sea%20Inferred%20From%20Regression%20Models&journal=Sci.+Tot.+Environ.&volume=774&)

Hwang J.-S., Kumar R., Dahms H.-U., Tseng L.-C., Chen Q.-C. (2007). Mesh Size Affects Abundance Estimates of Oithona Spp. (Copepoda, Cyclopoida). *Crustaceana* 80, 827–837. doi: 10.1163/156854007781363169

CrossRef Full Text (https://doi.org/10.1163/156854007781363169) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.-S.+Hwang&author=R.+Kumar&author=H.-U.+Dahms&author=L.-C.+Tseng&author=Q.-C.+Chen&publication_year=2007&title=Mesh%20Size%20Affects%20Abundance%20Estimates%20of%20Oithona%20Spp.%20%28Copepoda%2C%20Cyclopoida%29&journal=Crustaceana&volume=80&pages=827)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/31730851/) | CrossRef Full Text (https://doi.org/10.1016/j.cell.2019.10.008) | Google Scholar (http://scholar.google.com/scholar_lookup?

(/articles/10.3389/fmars.2022.894372/pdf)

CrossRef Full Text (https://doi.org/10.1007/BF00396409) | Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Ikeda&publication_year=1985&title=Metabolic%20Rates%20of%20Epipelagic%20Marine%20Zooplankton%20as%20a%20Function%20of%20Body%20Mass%20and%20Temperature&journal=Mar.+Biol.&volume=85&pages=1)

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

CrossRef Full Text (https://doi.org/10.1029/2019GB006258) | Google Scholar (http://scholar.google.com/scholar_lookup?author=T.+Ikenoue&author=K.+Kimoto&author=Y.+Okazaki&author=M.+Sato&author=M.%20C.+Honda&author=K.+Takahashi&publication_year=2019&title=Phaeodaria%3A%20An%20Important%20Carrier%20of%20Particulate%20Organic%20Carbon%20in%20the%20Mesopelagic%20Twilight%20Zone%20of%20the%20North%20Pacific%20Ocean&journal=Global+Biogeochem.+Cycle.&volume=33&pages=1146)

Irisson J.-O., Ayata S.-D., Lindsay D. J., Karp-Boss L., 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

CrossRef Full Text (https://doi.org/10.1146/annurev-marine-041921-013023) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.-O.+Irisson&author=S.-D.+Ayata&author=D.%20J.+Lindsay&author=L.+Karp-Boss&author=L.+Stemmann&publication_year=2022&title=Machine%20Learning%20for%20the%20Study%20of%20Plankton%20and%20Marine%20Snow%20From%20Images&journal=Annu.+Rev.+Mar.+Sci.&volume=14&pages=annurev)

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

CrossRef Full Text (https://doi.org/10.1371/journal.pbio.1001177) | Google Scholar (http://scholar.google.com/scholar_lookup?author=E.+Karsenti&author=S.%20G.+Acinas&author=P.+Bork&author=C.+Bowler&author=C.+De%20-Vargas&author=J.+Raes&publication_year=2011&title=A%20Holistic%20Approach%20to%20Marine%20Eco-Systems%20Biology&journal=PLoS+Biology&volume=9&pages=e1001177.1)

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

CrossRef Full Text (https://doi.org/10.4319/lo.2013.58.5.1843) | Google Scholar (http://scholar.google.com/scholar_lookup?

obal+Biogeochem.+Cycle.&volume=23&)

Leathwick J., Elith J., Francis M., Hastie T., 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

CrossRef Full Text (https://doi.org/10.3354/meps321267) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Leathwick&author=J.+Elith&author=M.+Francis&author=T.+Hastie&author=P.+Taylor&publication_year=2006&title=Variation%20in%20Demersal%20Fish%20Species%20Richness%20in%20the%20Oceans%20Surrounding%20New%20Zealand%3A%20An%20Analysis%20Using%20Boosted%20Regression%20Trees&journal=Mar.+Ecol.+Prog.+Ser.&volume=321&pages=267)

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

CrossRef Full Text (https://doi.org/10.1175/JCLI3898.1) | Google Scholar (http://scholar.google.com/scholar_lookup?author=P.+Lehodey&author=J.+Alheit&author=M.+Barange&author=T.+Baumgartner&author=G.+Beaugrand&author=K.+Drinkwater&publication_year=2006&title=Climate%20Variability%2C%20Fish%2C%20and%20Fisheries&journal=J.+Climate&volume=19&pages=5009)

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

CrossRef Full Text (https://doi.org/10.5194/bg-13-4111-2016) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+Le%20Qu%C3%A9r%C3%A9&author=E.%20T.+Buitenhuis&author=R.+Moriarty&author=S.+Alvain&author=O.+Aumont&author=L.+Bopp&publication_year=2016&title=Role%20of%20Zooplankton%20Dynamics%20for%20Southern%20Ocean%20Phytoplankton%20Biomass%20Andglobal%20Biogeochemical%20Cycles&journal=Biogeosciences&volume=13&pages=4111)

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

CrossRef Full Text (https://doi.org/10.3389/fmars.2019.00196) | Google Scholar (http://scholar.google.com/scholar_lookup?author=F.+Lombard&author=E.+Boss&author=A.%20M.+Waite&author=M.+Vogt&author=J.+Uitz&author=L.+Stemmann&publication_year=2019&title=EnglishGlobally%20Consistent%20Quantitative%20Observations%20of%20Planktonic%20Ecosystems&journal=Front.+Mar.+Sci.&volume=6&)

Lombard F., Kiørboe T. (2010). Marine Snow Originating From Appendicularian Houses: Age-Dependent

Lombard F., Renaud F., Sainsbury C., Sciandra A., Gorsky G. (2009). Appendicularian Ecophysiology I. *J. Mar. Syst.* 78, 606–616. doi: 10.1016/j.jmarsys.2009.01.004

CrossRef Full Text (https://doi.org/10.1016/j.jmarsys.2009.01.004) | Google Scholar (http://scholar.google.com/scholar_lookup?author=F.+Lombard&author=F.+Renaud&author=C.+Sainsbury&author=A.+Sciandra&author=G.+Gorsky&publication_year=2009&title=Appendicularian%20Ecophysiology%20I&journal=J.+Mar.+Syst.&volume=78&pages=606)

Longhurst A. (1995). Seasonal Cycles of Pelagic Production and Consumption. *Prog. Oceanogr.* 36, 77–167. doi: 10.1016/0079-6611(95)00015-1

CrossRef Full Text (https://doi.org/10.1016/0079-6611(95)00015-1) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.+Longhurst&publication_year=1995&title=Seasonal%20Cycles%20of%20Pelagic%20Production%20and%20Consumption&journal=Prog.+Oceanogr.&volume=36&pages=77)

Longhurst A. R., 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

CrossRef Full Text (https://doi.org/10.1016/0079-6611(89)90010-4) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.%20R.+Longhurst&author=W.+Glen%20Harrison&publication_year=1989&title=The%20Biological%20Pump%3A%20Profiles%20of%20Plankton%20Production%20and%20Consumption%20in%20the%20Upper%20Ocean&journal=Prog.+Oceanogr.&volume=22&pages=47)

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

CrossRef Full Text (https://doi.org/10.1111/geb.12169) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.%20H.+Lucas&author=D.%20O.%20B.+Jones&author=C.%20J.+Hollyhead&author=R.%20H.+Condon&author=C.%20M.+Duarte&author=W.%20M.+Graham&publication_year=2014&title=Gelatinous%20Zooplankton%20Biomass%20in%20the%20Global%20Oceans%3A%20Geographic%20Variation%20and%20Environmental%20Drivers%3A%20Global%20Gelatinous%20Biomass&journal=Global+Ecol.+Biogeogr.&volume=23&pages=701)

Mansour J. S., Norlin A., Llopis Monferrer N., L'Helguen S., 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

CrossRef Full Text (https://doi.org/10.1002/lno.11714) | Google Scholar (/articles/10.3389/fmars.2022.894372/pdf)

CrossRef Full Text (https://doi.org/10.1093/plankt/fbw094) | Google Scholar (http://scholar.google.com/scholar_lookup?author=K.+McConville&author=A.+Atkinson&author=E.%20S.+Fileman&author=J.%20I.+Spicer&author=A.%20G.+Hirst&publication_year=2016&title=Disentangling%20the%20Counteracting%20Effects%20of%20Water%20Content%20and%20Carbon%20Mass%20on%20Zooplankton%20Growth&journal=J.+Plankt.+Res.&volume=39&pages=246)

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

CrossRef Full Text (https://doi.org/10.1111/gcb.15499) | Google Scholar (http://scholar.google.com/scholar_lookup?author=N.+McGinty&author=A.%20D.+Barton&author=N.%20R.+Record&author=Z.%20V.+Finkel&author=D.%20G.+Johns&author=C.%20A.+Stock&publication_year=2021&title=Anthropogenic%20Climate%20Change%20Impacts%20on%20Copepod%20Trait%20Biogeography&journal=Global+Change+Biol.&volume=27&pages=1431)

Michaels A. F. (1988). Vertical Distribution and Abundance of Acantharia and Their Symbionts. *Mar. Biol.* 97, 559–569. doi: 10.1007/BF00391052

CrossRef Full Text (https://doi.org/10.1007/BF00391052) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.%20F.+Michaels&publication_year=1988&title=Vertical%20Distribution%20and%20Abundance%20of%20Acantharia%20and%20Their%20Symbionts&journal=Mar.+Biol.&volume=97&pages=559)

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

CrossRef Full Text (https://doi.org/10.1093/plankt/13.2.399) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.%20F.+Michaels&publication_year=1991&title=Acantharian%20Abundance%20and%20Symbiont%20Productivity%20at%20the%20VERTEX%20Seasonal%20Station&journal=J.+Plankt.+Res.&volume=13&pages=399)

Moriarty R., Buitenhuis E. T., Le Quéré C., 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

CrossRef Full Text (https://doi.org/10.5194/essdd-5-187-2012) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.+Moriarty&author=E.%20T.+Buitenhuis&author=C.+Le%20Qu%C3%A9r%C3%A9&author=M.-P.+Gosselin&publication_year=2012&title=Distribution%20of%20Known%20Macrozooplankton%20Abu-

ata&volume=5&pages=45)

Morley J., Stepien J. (1984). Siliceous Microfauna in Waters Beneath Antarctic Sea Ice. *Mar. Ecol. Prog. Ser.* 19, 207–210. doi: 10.3354/meps019207

CrossRef Full Text (https://doi.org/10.3354/meps019207) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.+Morley&author=J.+Stepien&publication_year=1984&title=Siliceous%20Microfauna%20in%20Waters%20Beneath%20Antarctic%20Sea%20Ice&journal=Mar.+Ecol.+Prog.+Ser.&volume=19&pages=207)

Nakamura Y., Imai I., Yamaguchi A., Tuji A., 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

CrossRef Full Text (https://doi.org/10.3800/pbr.8.107) | Google Scholar (http://scholar.google.com/scholar_lookup?author=Y.+Nakamura&author=I.+Imai&author=A.+Yamaguchi&author=A.+Tuji&author=N.+Suzuki&publication_year=2013&title=Aulographis%20Japonica%20Sp.%20Nov.%20%28Phaeodaria%2C%20Aulacanthida%2C%20Aulacanthidae%29%2C%20an%20Abundant%20Zooplankton%20in%20the%20Deep%20Sea%20of%20the%20Sea%20of%20Japan&journal=Plankt.+Bentho.+Res.&volume=8&pages=107)

Nakamura Y., Suzuki N. (2015). "Phaeodaria: Diverse Marine Cercozoans of World-Wide Distribution," in *Marine Protists*. Eds. Ohtsuka S., Suzaki T., Horiguchi T., Suzuki N., Not F. (Tokyo: Springer Japan), 223–249. doi: 10.1007/978-4-431-55130-0_9

CrossRef Full Text (https://doi.org/10.1007/978-4-431-55130-0_9) | Google Scholar (http://scholar.google.com/scholar_lookup?author=Y.+Nakamura&author=N.+Suzuki&publication_year=2015&title=Phaeodaria%3A%20Diverse%20Marine%20Cercozoans%20of%20World-Wide%20Distribution&book=Marine+Protists&pages=223)

Nelder J. A., Wedderburn R. W. M. (1972). *Journal of the Royal Statistical Society. Series A (General)* 135(3):370–384 (15 pages) doi: 10.2307/2344614

CrossRef Full Text (https://doi.org/10.2307/2344614) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20A.+Nelder&author=R.%20W.%20M.+Wedderburn&publication_year=1972&journal=Journal+of+the+Royal+Statistical+Society.+Series+A+(General)&volume=135&pages=370)

Nichols J., 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

CrossRef Full Text (https://doi.org/10.1093/plankt/13.3.661) | Google Scholar (http://scholar.google.com/scholar_lookup?

com/scholar_lookup?
author=M.%20D.+Ohman&publication_year=2019&title=A%20Sea%20of%20Tentacles%3A%20Optically%20Discernible%20Traits%20Resolved%20From%20Planktonic%20Organisms%20in%20Situ&journal=ICES.+J.+Mar.+Sci.&volume=76&pages=1959)

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

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/26029378/) | CrossRef Full Text (https://doi.org/10.1038/sdata.2015.23) | Google Scholar (http://scholar.google.com/scholar_lookup?author=S.+Pesant&author=F.+Not&author=M.+Picheral&author=S.+Kandels-Lewis&author=N.+Le%20Bescot&author=G.+Gorsky&publication_year=2015&title=Open%20Science%20Resources%20for%20the%20Discovery%20and%20Analysis%20of%20Tara%20Oceans%20Data&journal=Sci.+Data&volume=2&pages=150023)

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

CrossRef Full Text (https://doi.org/10.1139/as-2020-0018) | Google Scholar (http://scholar.google.com/scholar_lookup?author=H.+Pettitt-Wade&author=T.+Pearce&author=D.+Kuptana&author=C.%20P.+Gallagher&author=K.+Scharffenberg&author=E.%20V.+Lea&publication_year=2020&title=Inuit%20Observations%20of%20a%20Tunicata%20Bloom%20Unusual%20for%20the%20Amundsen%20Gulf%2C%20Western%20Canadian%20Arctic&journal=Arct.+Sci.&volume=6&pages=340)

Phillips S. J., Anderson R. P., Schapire R. E. (2006). Maximum Entropy Modeling of Species Geographic Distributions. *Ecol. Model.* 190, 231–259. doi: 10.1016/j.ecolmodel.2005.03.026

CrossRef Full Text (https://doi.org/10.1016/j.ecolmodel.2005.03.026) | Google Scholar (http://scholar.google.com/scholar_lookup?author=S.%20J.+Phillips&author=R.%20P.+Anderson&author=R.%20E.+Schapire&publication_year=2006&title=Maximum%20Entropy%20Modeling%20of%20Species%20Geographic%20Distributions&journal=Ecol.+Model.&volume=190&pages=231)

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.

PubMed Abstract (https://pubmed.ncbi.nlm.nih.gov/35909413/) | CrossRef Full Text (https://doi.org/10.1002/lom3.10475) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.+Picheral&author=C.+Catalano&author=D.+Brousseau&author=H.+Claustre&author=L.+Coppola&author=E.+Leymarie&publication_year=2021&title=The%20Underwater%20Vision%20Profiler%20

thor=J.-O.+Irisson&publication_year=2017&book=EcoTaxa,+a+Tool+for+the+Taxonomic+Classification+of+Images&)

Picheral M., Guidi L., Stemmann L., Karl D. M., Iddaoud G., 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

CrossRef Full Text (https://doi.org/10.4319/lom.2010.8.462) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.+Picheral&author=L.+Guidi&author=L.+Stemmann&author=D.%20M.+Karl&author=G.+Iddaoud&author=G.+Gorsky&publication_year=2010&title=The%20Underwater%20Vision%20Profiler%205%3A%20An%20Advanced%20Instrument%20for%20High%20Spatial%20Resolution%20Studies%20of%20Particle%20Size%20Spectra%20and%20Zooplankton%3A%20Underwater%20Vision%20Profiler&journal=Limnol.+Oceanogr.:+Methods&volume=8&pages=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

CrossRef Full Text (https://doi.org/10.1016/j.dsr.2020.103303) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.%20H.+Pinkerton&author=M.+D%C3%A9cima&author=J.%20A.+Kitchener&author=K.%20T.+Takahashi&author=K.%20V.+Robinson&author=R.+Stewart&publication_year=2020&title=Zooplankton%20in%20the%20Southern%20Ocean%20From%20the%20Continuous%20Plankton%20Recorder%3A%20Distributions%20and%20Long-Term%20Change&journal=Deep.+Sea.+Res.+Part+I.:+Oceanogr.+Res.+Pap.&volume=162&)

R Core Team (2021). *R: A Language and Environment for Statistical Computing* (Vienna, Austria: R Foundation for Statistical Computing).

Google Scholar (http://scholar.google.com/scholar_lookup?author=R%20Core%20Team&publication_year=2021&book=R:+A+Language+and+Environment+for+Statistical+Computing&)

Remsen A., Hopkins T. L., 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

CrossRef Full Text (https://doi.org/10.1016/j.dsr.2003.09.008) | Google Scholar (http://scholar.google.com/scholar_lookup?author=A.+Remsen&author=T.%20L.+Hopkins&author=S.+Samson&publication_year=2004&title=What%20You%20See%20is%20Not%20What%20You%20Catch%3A%20A%20Comparison%20of%20Concurrently%20Collected%20Net%2C%20Optical%20Plankton%20Counter%2C%20and%20Shadowed%20Im

author=A.%20J.+Richardson&author=A.+Bakun&author=G.%20C.+Hays&author=M.%20J.+Gibbons&publication_year=2009&title=The%20Jellyfish%20Joyride%3A%20Causes%2C%20Consequences%20and%20Management%20Responses%20to%20a%20More%20Gelatinous%20Future&journal=Trends+Ecol.+Evol.&volume=24&pages=312)

Rombouts I., Beaugrand G., Ibañez F., Gasparini S., Chiba S., 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

CrossRef Full Text (https://doi.org/10.1098/rspb.2009.0742) | Google Scholar (http://scholar.google.com/scholar_lookup?author=I.+Rombouts&author=G.+Beaugrand&author=F.+Iba%C5%88ez&author=S.+Gasparini&author=S.+Chiba&author=L.+Legendre&publication_year=2009&title=Global%20Latitudinal%20Variations%20in%20Marine%20Copepod%20Diversity%20and%20Environmental%20Factors&journal=Proc.+R.+Soc.+B.:+Biol.+Sci.&volume=276&pages=3053)

Sameoto D., Cochrane N., 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

CrossRef Full Text (https://doi.org/10.1139/f93-039) | Google Scholar (http://scholar.google.com/scholar_lookup?author=D.+Sameoto&author=N.+Cochrane&author=A.+Herman&publication_year=1993&title=Convergence%20of%20Acoustic%2C%20Optical%2C%20and%20Net-Catch%20Estimates%20of%20Euphausiid%20Abundance%3A%20Use%20of%20Artificial%20Light%20to%20Reduce%20Net&journal=Can.+J.+Fish.+Aquat.+Sci.&volume=50&pages=334)

Schapire R. E. (2003). "The Boosting Approach to Machine Learning: An Overview," in *Nonlinear Estimation and Classification*. Eds. Denison D. D., Hansen M. H., Holmes C. C., Mallick B., Yu B. (New York, NY: Springer), 149–171. Lecture Notes in Statistics. doi: 10.1007/978-0-387-21579-2_9

CrossRef Full Text (https://doi.org/10.1007/978-0-387-21579-2_9) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.%20E.+Schapire&publication_year=2003&title=The%20Boosting%20Approach%20to%20Machine%20Learning%3A%20An%20Overview&book=Nonlinear+Estimation+and+Classification&pages=149)

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

CrossRef Full Text (https://doi.org/10.2971/jeos.2010.10017s) | Google Scholar (http://scholar.google.com/scholar_lookup?

CrossRef Full Text (https://doi.org/10.1007/s40641-020-00160-0) | Google Scholar (http://scholar.google.com/scholar_lookup?author=R.+S%C3%A9f%C3%A9rian&author=S.+Berthet&author=A.+Yool&author=J.+Palmieri&author=L.+Bopp&author=A.+Tagliabue&publication_year=2020&title=Tracking%20Improvement%20in%20Simulated%20Marine%20Biogeochemistry%20Between%20CMIP5%20and%20CMIP6&journal=Curr.+Climate+Change+Rep.&volume=6&pages=95)

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

CrossRef Full Text (https://doi.org/10.1007/s00300-005-0058-5) | Google Scholar (http://scholar.google.com/scholar_lookup?author=V.+Siegel&publication_year=2005&title=Distribution%20and%20Population%20Dynamics%20of%20Euphausia%20Superba%3A%20Summary%20of%20Recent%20Findings&journal=Pol.+Biol.&volume=29&pages=1)

Siegel V. (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

CrossRef Full Text (https://doi.org/10.1007/978-3-319-29279-3) | Google Scholar (http://scholar.google.com/scholar_lookup?author=V.+Siegel&publication_year=2016&title=Biology%20and%20Ecology%20of%20Antarctic%20Krill&book=Advances+in+Polar+Ecology&)

Son D., Cho H., 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

CrossRef Full Text (https://doi.org/10.1016/j.aquabot.2018.07.003) | Google Scholar (http://scholar.google.com/scholar_lookup?author=D.+Son&author=H.+Cho&author=E.%20J.+Lee&publication_year=2018&title=Determining%20Factors%20for%20the%20Occurrence%20and%20Richness%20of%20Submerged%20Macrophytes%20in%20Major%20Korean%20Rivers&journal=Aquat.+Bot.&volume=150&pages=82)

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

CrossRef Full Text (https://doi.org/10.1016/j.pocean.2021.102717) | Google Scholar (http://scholar.google.com/scholar_lookup?author=Y.%20D.+Soviadan&author=F.+Benedetti&author=M.%20C.+Brand%C3%A3o&author=S.-D.+Ayata&author=J.-O.+Irisson&author=J.%20L.+Jamet&publication_year=2022&title=Patterns%20of%20Mesozooplankton%20Community%20Composition%20and%20Vertical%20Fluxes%20in%20the%20Global%20Ocean&journal=Prog.+Oceanogr.&volume=200&pages=102717)

cation_year=2008&title=A%20Comparison%20of%20Mesopelagic%20Mesozooplankton%20Community%20Structure%20in%20the%20Subtropical%20and%20Subarctic%20North%20Pacific%20Ocean&journal=Deep.+Sea.+Res.+Part+II.:+Top.+Stud.+Oceanogr.&volume=55&pages=1615)

Steinberg D. K., Landry M. R. (2017). Zooplankton and the Ocean Carbon Cycle. *Annu. Rev. Mar. Sci.* 9, 413–444. doi: 10.1146/annurev-marine-010814-015924

CrossRef Full Text (https://doi.org/10.1146/annurev-marine-010814-015924) | Google Scholar (http://scholar.google.com/scholar_lookup?author=D.%20K.+Steinberg&author=M.%20R.+Landry&publication_year=2017&title=Zooplankton%20and%20the%20Ocean%20Carbon%20Cycle&journal=Annu.+Rev.+Mar.+Sci.&volume=9&pages=413)

Stemmann L., 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

CrossRef Full Text (https://doi.org/10.1146/annurev-marine-120710-100853) | Google Scholar (http://scholar.google.com/scholar_lookup?author=L.+Stemmann&author=E.+Boss&publication_year=2012&title=Plankton%20and%20Particle%20Size%20and%20Packaging%3A%20From%20Determining%20Optical%20Properties%20to%20Driving%20the%20Biological%20Pump&journal=Annu.+Rev.+Mar.+Sci.&volume=4&pages=263)

Stemmann L., Gorsky G., Marty J.-C., Picheral M., 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

CrossRef Full Text (https://doi.org/10.1016/S0967-0645(02)00032-2) | Google Scholar (http://scholar.google.com/scholar_lookup?author=L.+Stemmann&author=G.+Gorsky&author=J.-C.+Marty&author=M.+Picheral&author=J.-C.+Miquel&publication_year=2002&title=Four-Year%20Study%20of%20Large-Particle%20Vertical%20Distribution%20%280%E2%80%931000%20M%29%20in%20the%20NW%20Mediterranean%20in%20Relation%20to%20Hydrology%2C%20Phytoplankton%2C%20and%20Vertical%20Flux&journal=Deep.+Sea.+Res.+Part+II.:+Top.+Stud.+Oceanogr.&volume=49&pages=2143)

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

CrossRef Full Text (https://doi.org/10.1093/icesjms/fsn010) | Google Scholar (http://scholar.google.com/scholar_lookup?author=L.+Stemmann&author=M.+Youngbluth&author=K.+Robert&author=A.+Hosia&author=M.+Picheral&author=H.+Paterson&publication_year=2008&title=Global%20Zoogeography%20of%20Fragile%20-

tion_year=2018&title=Large%20Phaeodaria%20in%20the%20Twilight%20Zone%3A%20Their%20Role%20in%20the%20Carbon%20Cycle%3A%20Phaeodarian%20Ecology%20in%20the%20Twilight%20Zone&journal=Limnol.+Oceanogr.&volume=63&pages=2579)

Stukel M. R., Ohman M. D., Kelly T. B., 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

CrossRef Full Text (https://doi.org/10.3389/fmars.2019.00397) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.%20R.+Stukel&author=M.%20D.+Ohman&author=T.%20B.+Kelly&author=T.+Biard&publication_year=2019&title=The%20Roles%20of%20Suspension-Feeding%20and%20Flux-Feeding%20Zooplankton%20as%20Gatekeepers%20of%20Particle%20Flux%20Into%20the%20Mesopelagic%20Ocean%20in%20the%20Northeast%20Pacific&journal=Front.+Mar.+Sci.&volume=6&)

Sun H., Benzie P., Burns N., Hendry D., Player M., 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

CrossRef Full Text (https://doi.org/10.1098/rsta.2007.2187) | Google Scholar (http://scholar.google.com/scholar_lookup?author=H.+Sun&author=P.+Benzie&author=N.+Burns&author=D.+Hendry&author=M.+Player&author=J.+Watson&publication_year=2008&title=Underwater%20Digital%20Holography%20for%20Studies%20of%20Marine%20Plankton&journal=Philos.+Trans.+R.+Soc.+A.:+Math.+Phys.+Eng.+Sci.&volume=366&pages=1789)

Sunday J. M., Bates A. E., Dulvy N. K. (2012). Thermal Tolerance and the Global Redistribution of Animals. *Nat. Climate Change* 2, 686–690. doi: 10.1038/nclimate1539

CrossRef Full Text (https://doi.org/10.1038/nclimate1539) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20M.+Sunday&author=A.%20E.+Bates&author=N.%20K.+Dulvy&publication_year=2012&title=Thermal%20Tolerance%20and%20the%20Global%20Redistribution%20of%20Animals&journal=Nat.+Climate+Change&volume=2&pages=686)

Suthers I., Rissik D., 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 (https://www.researchgate.net/publication/265624443_Plankton_A_Guide_to_Their_Ecology_and_Monitoring_for_Water_Quality)

Google Scholar (http://scholar.google.com/scholar_lookup?

(/articles/10.3389/fmars.2022.894372/pdf)

Turner J. T. (2002). Zooplankton Fecal Pellets, Marine Snow and Sinking Phytoplankton Blooms. *Aquat. Microbial. Ecol.* 27, 57–102. doi: 10.3354/ame027057

CrossRef Full Text (https://doi.org/10.3354/ame027057) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20T.+Turner&publication_year=2002&title=Zooplankton%20Fecal%20Pellets%2C%20Marine%20Snow%20and%20Sinking%20Phytoplankton%20Blooms&journal=Aquat.+Microbial.+Ecol.&volume=27&pages=57)

Turner J. T. (2004). The Importance of Small Planktonic Copepods and Their Roles in Pelagic Marine Food Webs. *Zool. Stud.* 43, 255–266.

Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20T.+Turner&publication_year=2004&title=The%20Importance%20of%20Small%20Planktonic%20Copepods%20and%20Their%20Roles%20in%20Pelagic%20Marine%20Food%20Webs&journal=Zool.+Stud.&volume=43&pages=255)

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

CrossRef Full Text (https://doi.org/10.1016/j.pocean.2014.08.005) | Google Scholar (http://scholar.google.com/scholar_lookup?author=J.%20T.+Turner&publication_year=2015&title=Zooplankton%20Fecal%20Pellets%2C%20Marine%20Snow%2C%20Phytodetritus%20and%20the%20Ocean%E2%80%99s%20Biological%20Pump&journal=Prog.+Oceanogr.&volume=130&pages=205)

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.

Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+van%20der%20Lingen&author=A.+Bertrand&author=A.+Bode&author=R.+Brodeur&author=L.+Cubillos&author=P.+Espinoza&publication_year=2009&title=Trophic%20Dynamics%20of%20Small%20Pelagic%20Fish&pages=333)

van der Lingen C., Hutchings L., 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

CrossRef Full Text (https://doi.org/10.2989/18142320609504199) | Google Scholar (http://scholar.google.com/scholar_lookup?author=C.+van%20der%20Lingen&author=L.+Hutchings&author=J.+Field&publication_year=2006&title=Comparative%20Trophodynamics%20of%20Anchovy%20Engraulis%20Encrasicolus%20and%20Sardine

CrossRef Full Text (https://doi.org/10.1080/03610918.2020.1772302) | Google Scholar (http://scholar.google.com/scholar_lookup?author=P.%20H.+Wiebe&author=S.%20H.+Boyd&author=M.+Davis&author=J.%20L.+Cox&publication_year=1982&title=Avoidance%20of%20Towed%20Nets%20by%20the%20Euphausiid%20Nematoscelis%20Megalops&journal=Fish.+Bull.&volume=80&pages=75)

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

CrossRef Full Text (https://doi.org/10.1007/BF00397811) | Google Scholar (http://scholar.google.com/scholar_lookup?author=P.%20H.+Wiebe&author=A.%20W.+Morton&author=A.%20M.+Bradley&author=R.%20H.+Backus&author=J.%20E.+Craddock&author=V.+Barber&publication_year=1985&title=New%20Development%20in%20the%20MOCNESS%2C%20an%20Apparatus%20for%20Sampling%20Zooplankton%20and%20Micronekton&journal=Mar.+Biol.&volume=87&pages=313)

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

CrossRef Full Text (https://doi.org/10.1016/j.scitotenv.2017.12.288) | Google Scholar (http://scholar.google.com/scholar_lookup?author=M.+Zhang&author=D.+Straile&author=F.+Chen&author=X.+Shi&author=Z.+Yang&author=Y.+Cai&publication_year=2018&title=Dynamics%20and%20Drivers%20of%20Phytoplankton%20Richness%20and%20Composition%20Along%20Productivity%20Gradient&journal=Sci.+Tot.+Environ.&volume=625&pages=275)

Zhou H., Yang Y., 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

CrossRef Full Text (https://doi.org/10.1080/03610918.2020.1772302) | Google Scholar (http://scholar.google.com/scholar_lookup?author=H.+Zhou&author=Y.+Yang&author=W.+Qian&publication_year=2019&title=Tweedie%20Gradient%20Boosting%20for%20Extremely%20Unbalanced%20Zero-Inflated%20Data&journal=arXiv&)

Keywords: global zooplankton, *in situ* imaging, biomass, machine learning, underwater vision profiler (UVP), spatial distribution, boosted regression trees (BRT), habitat modeling

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

***Correspondence:** Laetitia Drago, **laetitia.drago@imev-mer.fr (mailto:laetitia.drago@imev-mer.fr)**; Lars Stemmann, **lars.stemmann@imev-mer.fr (mailto:lars.stemmann@imev-mer.fr)**; Rainer Kiko, **rainer.kiko@imev-mer.fr (mailto:rainer.kiko@imev-mer.fr)**

†These authors share last authorship

People also looked at

A Multi-Mode Convolutional Neural Network to reconstruct satellite-derived chlorophyll-a time series in the global ocean from physical drivers (/articles/10.3389/fmars.2023.1077623/full)

Joana Roussillon (https://loop.frontiersin.org/people/2051493/overview), Ronan Fablet (https://loop.frontiersin.org/people/998988/overview), Thomas Gorgues (https://loop.frontiersin.org/people/960137/overview), Lucas Drumetz (https://loop.frontiersin.org/people/1086957/overview), Jean Littaye and Elodie Martinez (https://loop.frontiersin.org/people/759755/overview)

Editorial: Technological advances for measuring planktonic components of the pelagic ecosystem: An integrated approach to data collection and analysis (/articles/10.3389/fmars.2023.1100998/full)

Sophie G. Pitois (https://loop.frontiersin.org/people/334624/overview), Elaine S. Fileman (https://loop.frontiersin.org/people/896192/overview), Mark C. Benfield (https://loop.frontiersin.org/people/728412/overview), Peter H. Wiebe (https://loop.frontiersin.org/people/458870/overview) and Fabien Lombard

Global Observing Needs in the Deep Ocean (/articles/10.3389/fmars.2022.894372/pdf)

(https://loop.frontiersin.org/people/627364/overview), Felix Janssen
(https://loop.frontiersin.org/people/199318/overview), Andrea McCurdy, Henry A. Ruhl, Paul Snelgrove
(https://loop.frontiersin.org/people/132980/overview), Karen I. Stocks
(https://loop.frontiersin.org/people/560911/overview), David Bailey
(https://loop.frontiersin.org/people/492591/overview), Simone Baumann-Pickering
(https://loop.frontiersin.org/people/606504/overview), Chris Beaverson
(https://loop.frontiersin.org/people/688450/overview), Mark C. Benfield, David J. Booth
(https://loop.frontiersin.org/people/430041/overview), Marina Carreiro-Silva
(https://loop.frontiersin.org/people/358956/overview), Ana Colaço
(https://loop.frontiersin.org/people/156224/overview), Marie C. Eblé
(https://loop.frontiersin.org/people/695933/overview), Ashley M. Fowler
(https://loop.frontiersin.org/people/601721/overview), Kristina M. Gjerde
(https://loop.frontiersin.org/people/479531/overview), Daniel O. B. Jones
(https://loop.frontiersin.org/people/370765/overview), K. Katsumata
(https://loop.frontiersin.org/people/689770/overview), Deborah Kelley, Nadine Le Bris
(https://loop.frontiersin.org/people/138106/overview), Alan P. Leonardi
(https://loop.frontiersin.org/people/688307/overview), Franck Lejzerowicz
(https://loop.frontiersin.org/people/60666/overview), Peter I. Macreadie, Dianne McLean
(https://loop.frontiersin.org/people/617062/overview), Fred Meitz
(https://loop.frontiersin.org/people/610298/overview), Telmo Morato
(https://loop.frontiersin.org/people/137790/overview), Amanda Netburn
(https://loop.frontiersin.org/people/335933/overview), Jan Pawlowski
(https://loop.frontiersin.org/people/48493/overview), Craig R. Smith
(https://loop.frontiersin.org/people/142101/overview), Song Sun, Hiroshi Uchida
(https://loop.frontiersin.org/people/656508/overview), Michael F. Vardaro
(https://loop.frontiersin.org/people/657517/overview), R. Venkatesan
(https://loop.frontiersin.org/people/590183/overview) and Robert A. Weller
(https://loop.frontiersin.org/people/603392/overview)

Comparison of an In Situ Imaging Device and Net-Based Method to Study Mesozooplankton Communities in an Oligotrophic System (/articles/10.3389/fmars.2022.898057/full)

Alexander Barth (https://loop.frontiersin.org/people/1381519/overview) and Joshua Stone
(https://loop.frontiersin.org/people/1166827/overview)

(/articles/10.3389/fmars.2022.894372/pdf)

(https://loop.frontiersin.org/people/856580/overview), Sarah Romac (https://loop.frontiersin.org/people/823067/overview), Sébastien Colin (https://loop.frontiersin.org/people/1874325/overview), Nils Haëntjens (https://loop.frontiersin.org/people/814734/overview), Margaux Carmichael, Calixte Berger, David Le Guen, Johan Decelle (https://loop.frontiersin.org/people/831958/overview), Frédéric Mahé (https://loop.frontiersin.org/people/439375/overview), Julie Poulain (https://loop.frontiersin.org/people/858685/overview), Emmanuel Malpot, Carole Beaumont (https://loop.frontiersin.org/people/1801714/overview), Michel Hardy, Damien Guiffant, Ian Probert (https://loop.frontiersin.org/people/184458/overview), David F. Gruber (https://loop.frontiersin.org/people/381863/overview), Andrew E. Allen (https://loop.frontiersin.org/people/47163/overview), Gabriel Gorsky (https://loop.frontiersin.org/people/640145/overview), Michael J. Follows, Xavier Pochon (https://loop.frontiersin.org/people/417500/overview), Romain Troublé, B. B. Cael (https://loop.frontiersin.org/people/1704126/overview), Fabien Lombard (https://loop.frontiersin.org/people/606116/overview), Emmanuel Boss (https://loop.frontiersin.org/people/423697/overview), Manu Prakash (https://loop.frontiersin.org/people/1850258/overview), the Plankton Planet core team , Romain Bazile, Emmanuel Boss, Guillaume Bourdin, BB Cael, Roberto Casati, Sébastien Colin, Colomban de Vargas, Gabriel Gorsky, Damien Guiffant, Nils Haentjens, Nicolas Henry, Adam Larson, Noan Le Bescot, Fabien Lombard, Gilles Mirambeau, Clémentine Moulin, Anna Oddone, Manu Prakash, Christophe Prazuck, Vincent Raimbault, Clara Trellu and Romain Troublé

(/articles/10.3389/fmars.2022.894372/pdf)

(/articles/10.3389/fmars.2022.894372/pdf)