RESEARCH ARTICLE

Ecology and Evolution $_{\text{Open Access}}$ WILEY

Identification of plankton habitats in the North Sea

Rene-Marcel Plonus | Jens Floeter

Institute of Marine Ecosystem and Fishery Science, University of Hamburg, Hamburg, Germany

Correspondence
Rene-Marcel Plonus, Institute of Marine Ecosystem and Fishery Science, University of Hamburg, Hamburg, Germany.
Email: rene-marcel.plonus@uni-hamburg.de

Abstract

The definition of an ecological niche makes it possible to anticipate the responses of a species to changing environmental conditions. Broad tolerance limits and a paucity of readily observable niches in the pelagic zone make it difficult to anticipate responses of the plankton community related to anthropogenic or environmental changes. Plankton distributions are closely linked to climate change and shape the seascape for higher trophic levels, so monitoring plankton distributions and defining ecological niches will help to understand and predict ecosystem responses. Here we apply a machine learning autoencoder and a density-based clustering algorithm to high-frequency datasets sampled with a ROTV Triaxus in the North Sea. The results indicate that in this highly dynamic environment, local hydrography prevents niche-based separation of plankton species at the sub-mesoscale, despite the availability of different habitats. Plankton patches were associated with naturally occurring frontal systems and anthropogenically induced upwelling-downwelling dipoles in the vicinity of offshore wind farms (OWFs).

KEYWORDS
habitat maps, machine learning, North Sea, plankton distributions, plankton–habitat associations

JEL CLASSIFICATION
Community ecology, Ecosystem ecology, Spatial ecology

1 | INTRODUCTION

The concept of an "ecological niche" was first applied by Grinnell (1917) and referred to the abiotic demands of a species toward its environment and the behavioral adaptations of the species to the same (Grinnellian niche). A formal definition of the concept followed in Hutchinson (1957), describing an ecological niche as an 'n-dimensional hypervolume', where each dimension is influenced by a different environmental parameter. A further specification by Hutchinson (1957) was the distinction between fundamental and realized ecological niches. While the fundamental niche represents the total area that allows a population to survive,

the realized niche describes the area where a species dominates over competitors. The latter refers more to the "Eltonian niche", which focuses on the interaction of different species (Elton, 1927), especially in the more modern differentiation of the two (Dehling & Stouffer, 2018). The original understanding of the Eltonian niche as stated by Hutchinson (1957) assumes that co-occurring species occupy different niches that do not intersect. The paucity of readily observable physical niches in the pelagic zone (Behrenfeld et al., 2021) and a seemingly unstructured environment (Martin et al., 2021) let to the formulation of the 'paradox of plankton' by Hutchinson (1961): the co-existence of relatively many species in an apparently homogeneous environment, even though species

richness tends to increase with habitat heterogeneity (Lapointe & Bourget, 1999; MacArthur & MacArthur, 1961). However, niche-based models do not explain redundancy (Leibold & McPeek, 2006) or the existence of functional groups or traits (Barton et al., 2013; Dehling & Stouffer, 2018), both of which have been observed in plankton communities and are generally considered positive aspects of biodiversity (Leibold & McPeek, 2006).

The term "plankton" encompasses a diverse group of organisms in the oceans, including pytho- as well as zooplankton. Here we will use it to refer to the zooplankton component only. Traditional methods in plankton ecology have been time-consuming and thus prevented the up-scale to pan-oceanic observations (Irisson et al., 2022). This paucity of data and the inconsistency in sampling methods and scales have limited our understanding of the factors and processes determining the abundance or diversity of plankton (Lombard et al., 2019). New optical sampling methods have emerged in the last decade which produce a wealth of information (Irisson et al., 2022), but their scientific use was limited by the concepts and methods applicable to the huge amount of data they generate (Alvarez-Berastegui et al., 2014; Irisson et al., 2022; Lombard et al., 2019; North et al., 2016). In addition, the traditional view of turbulence-homogenized plankton communities has been challenged by recent studies, suggesting that biological and physical processes create a structured realm at scales down to a few centimeters (Basterretxea et al., 2020; Mitchell et al., 2008).

One way to handle these data is habitat maps, which link bio-physically distinct areas to specific species communities (Harris & Baker, 2012). However, due to the high variability of relevant spatial and temporal scales in the pelagic environment (Hinchey et al., 2008; Thompson et al., 2016) it is a daunting task to accurately determine pelagic habitats and identify associated plankton communities.

That is where machine learning excels. Several studies have recently created a link between global plankton datasets and associated physical variables (Busseni et al., 2020; Cael et al., 2021; Drago et al., 2022; Sonnewald et al., 2020). But even though machine learning has already successfully brought insights into plankton ecology in the past, fully automated predictions can still only be trusted for the most abundant species (Irisson et al., 2022; Plonus, Conradt, et al., 2021). Fortunately, plankton communities are usually highly diverse (Siegel, 1998) but dominated by a few, very abundant taxa (Fuhrman, 2009).

For example, there are 31 species known to have pluteus larvae, the larval form of echinoderms, in the North Sea (Laakmann et al., 2016). However, *Echinocardium cordatum* (Pennant, 1777) is the most abundant one, probably due to higher winter temperatures (Kirby et al., 2007), complemented by *Amphiura* spp. (Lindley et al., 1995). Since the 1980s, pluteus larvae have become even more abundant than copepods at times (Kirby et al., 2007; Lindley et al., 1995), which usually numerically dominate the zooplankton community. An extensive overview of copepod species in the North Sea was provided by Fransz et al. (1991). Even though there are several different species, four of them provide 85% of the biomass, namely *Acartia clausi* (Giesbrecht, 1889), *Centropages hamatus*

(Liljeborg, 1853), *Temora longicornis* (Müller, 1785), and *Pseudocalanus elongatus* (Boeck, 1865) (Beaugrand et al., 2002; Hickel, 1975). The by far most abundant one is *P. elongatus* (Fransz et al., 1991). Second in abundance only to copepods (Landry et al., 1994), appendicularians are of special importance for the vertical particle fluxes in the world's oceans (Winder et al., 2017). Due to their affinity for higher temperatures and tolerance toward more acidic conditions, it is likely that their importance will increase under global warming (Winder et al., 2017). The most abundant species in European coastal waters are *Oikopleura longicauda* Vogt, 1854, *Oikopleura dioica* Fol, 1872, *Oikopleura fusiformis* Fol, 1872, and *Fritillaria borealis* Lohmann, 1896 (Lopez-Urrutia et al., 2005). Although all species tolerate the range of temperatures and salinities observed in the North Sea (Lopez-Urrutia et al., 2005), *O. dioica* dominates the other species at temperatures below 20°C (Lombard et al., 2010).

Plankton communities shape the seascape for higher trophic levels (Bertrand et al., 2014) and are considered sentinels of ocean changes (Barton et al., 2013; Benedetti et al., 2021; Cael et al., 2021; Drago et al., 2022), which makes it an important task to understand and monitor their spatio-temporal variation (Friedland et al., 2020; Hays et al., 2005; McGill et al., 2006; McGinty et al., 2018) and their responses to changing environmental conditions.

Here, we map plankton communities to physically distinct habitats using a fully automated method applied to a high-frequency in-situ dataset sampled in the North Sea. Our approach takes advantage of machine learning speed to pry information from a wealth of data and makes the information therein accessible to human researchers. Using a fully connected Autoencoder (AE) and a density-based clustering algorithm we generate habitat maps from physical and biological variables, sampled with a Remotely Operated Towed Vehicle (ROTV) and investigate plankton–habitat interactions to reveal driving factors behind the compositions of the local plankton communities.

2 | MATERIALS AND METHODS

2.1 | Data acquisition and preparation

Physical and biological oceanographic measurements were recorded on different North Sea surveys with the RV Heincke (Knust et al., 2017) using a MacArtney TRIAXUS ROTV, complemented by a video plankton recorder (VPR). The TRIAXUS was towed behind the research vessel in an undulating fashion between the sea surface and bottom. A detailed description of the TRIAXUS and the associated sampling procedure can be found in Plonus, Conradt, et al. (2021) and Plonus, Vogl, and Floeter (2021). General data handling was accomplished with R 4.4.0 (RCoreTeam, 2020) and the tidyverse packages "purrr", "tibble", "dplyr", "ggplot2", "rstatix", and "tidyr" (Wickham et al., 2019). The map was generated using 'ggplot2' and "marmap" (Pante & Simon-Bouhet, 2013) in R (Figure 1). For our analyses, we selected the following parameters to determine bio-physical niches in the pelagic realm: temperature (°C), salinity (psu), oxygen ($\mu mol\,L^{-1}$), and chlorophyll-a (RFU). Furthermore, we

FIGURE 1 Sampling transects from HE466 (T1–T6, yellow) and HE446 (T1–T2, orange). Red polygons: Wind turbines. Water depth ranges from 0–10 m (white) to 40–50 m (dark blue). A map of the cruises HE429 and HE534 was provided in Plonus, Vogl, and Floeter (2021).
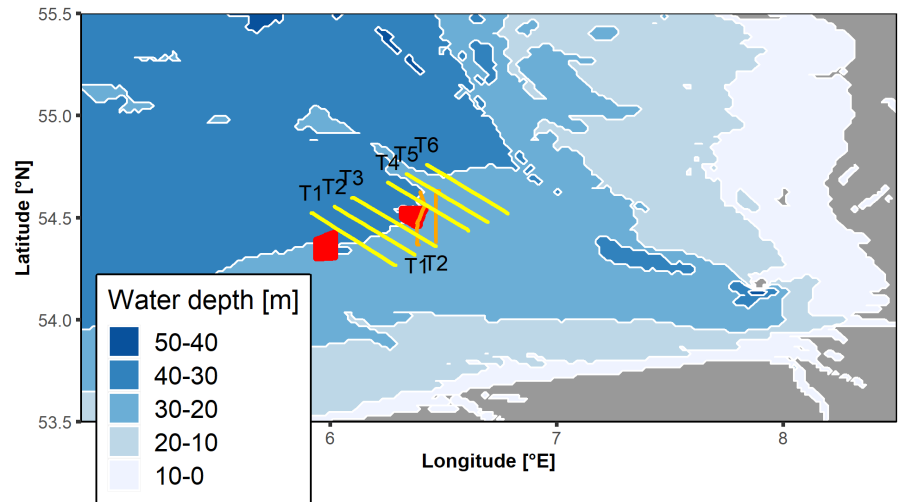
had sufficient VPR-derived density data (N L^{-1}) available for the plankton groups "Appendicularia", "Copepoda", "Dinoflagellates", "Marine snow", and "Pluteus larvae".

Transect diagrams were generated using Ocean Data View (ODV, Schlitzer, 2020) with the embedded spatial interpolation software DIVA (Troupin et al., 2012) and exported as grids with a resolution of ~25 m length × 1 m depth. Abiotic measurements were normalized and rescaled to range from −1 to 1. All plankton density values and chlorophyll a measurements were transformed using the natural logarithm of $x + 1$. The Euclidean distances in the multidimensional space defined by the plankton densities and the chlorophyll a concentration between each grid cell and the top left grid cell of each transect were calculated (Harris et al., 2020). Those distances were normalized and rescaled as described above. This was necessary since deep learning models generally perform better with homogeneous, small values (Bishop, 1995).

The exported grids for each selected parameter and the above-calculated distance measure were stacked and transformed into feature-vectors where each grid cell became one vector with 4 features (1 parameter = 1 feature). In our definition, a pelagic microhabitat with a spatial extent of ~25 m × 1 m corresponds to one of those feature-vectors, which reflect the niche space at this point (Colwell & Rangel, 2009).

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

2.2 | Model description

The AE consisted of four fully connected layers in the Encoder and Decoder, respectively. The Decoder used the transposed weights of the Encoder in reversed order, e.g., the weights of the first Encoder-Layer were shared with the last Decoder-Layer. The first layer of the Encoder inflated the 4-dimensional feature-vector to a 200-dimensional feature-vector, which was reduced to a 100-, 50-, and 4-dimensional feature-vector by the following layers (4–200–100–50–4). The Decoder did the same in reverse (4–50–100–200–4). The batchsize (number of input microhabitats that are processed simultaneously) was set to 100 and the learning rate followed a sawtooth-like undulating scheme, initialized at 5e^{-8}. Each input feature-vector corresponded to one microhabitat and included one measurement of each parameter. The model was trained using data recorded during the research cruises HE429 (July 2014) and HE534 (June 2019). The best performance was achieved using a limited training set of only 50.000 randomly selected micro-habitats over 250 epochs while reserving the remaining ~200.000 microhabitats for validation. During one Epoch each micro-habitat is presented to the model exactly once.

2.3 | Microhabitat segregation

The final model was tested using data generated on the research cruise HE466 (June 2016). Compared to the human-validated plankton densities from HE429 and HE534, the plankton densities for HE466 were estimated exclusively in an automated manner using the threshold procedure proposed by Faillettaz et al. (2016). By applying the trained Encoder only, we created 4-dimensional representations of the original input. In the following, we will refer to the processing of the micro-habitats by the Encoder as "projection". Microhabitats with similar characteristics were projected closer to each other than micro-habitats with different characteristics. HDBSCAN calculates the Euclidean distance to build clusters from the, in this case, 4-dimensional inputs (McInnes et al., 2017). We refer to the clustered micro-habitats as macrohabitat (MH). MHs were labeled manually. For more information regarding the clustering, see Plonus, Vogl, and Floeter (2021).

2.4 | Identification of key parameters

While very deep architectures can easily become some kind of "black box", it was still possible to trace individual inputs in our relatively shallow model which on top consisted only of fully connected Dense-Layers and did not rely on convolutional filters. Fully connected layers take an n-dimensional input which is multiplied by a pre-defined number "X" of weight sets of length n, resulting in a weight matrix of $n \times X$ weights. Basically, the output of a Dense-Layer is the dot product of the input and the weight matrix fed into some nonlinear activation function, and thus the weights ultimately define the relative influence of a given input upon the final output. Weights close to zero results in minor changes in the output even if the input variable varies a lot while increasing weights (positive as well as negative) can facilitate major changes in the output. Thus, analyzing the final weights after completion of the training phase revealed the relative influence each input variable had upon a specific output dimension of the Encoder, regardless of the analyzed transect. A similar approach has been used in Drago et al. (2022) to evaluate the importance of single nodes in a random forest algorithm.

2.5 | Sensitivity analysis

Sensitivity analysis (SA) estimates the importance of an input variable for the model output. We applied a global SA following Sobol (2001) to estimate the importance of each output dimension (D1–D4) of our AE. In a global SE, all parameters are varied simultaneously, allowing not only the contributions of individual parameters to be assessed but also the contribution of their interactions to the variability of the model output.

HDBSCANs prediction method performed poorly on the randomly generated inputs for the SA, likely due to the "curse of dimensionality" (Bellman & Dreyfus, 2015). Randomly generated points were unlikely to be close to HDBSCAN's pre-estimated "core points" and were therefore usually classified as outliers. Thus, we trained a support vector machine (SVM) to separate the clusters identified by HDBSCAN and to predict the randomly alternating data for the SA. SVMs define boundaries between existing clusters by maximizing the distance between the boundary and each adjacent cluster. Because HDBSCAN forms clusters from spatially distinct groups, SVMs are uniquely qualified to assign randomly generated data points to the closest cluster despite the "curse of dimensionality". The SVM was trained with the e1071-package in R (Meyer et al., 2023) and the SA was performed with the SALib-library in Python (Iwanaga et al., 2022).

2.6 | Plankton–habitat associations

Plankton–habitat associations were investigated with the R-package "shar" (Hesselbarth, 2021) which is based on the methods in Plotkin
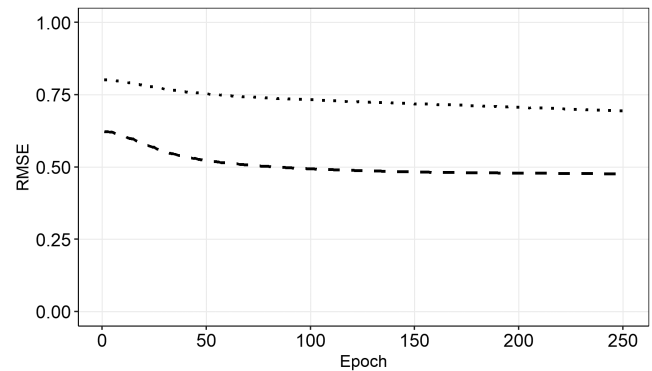

FIGURE 2 History of the model training. Root-mean-squared error (RMSE—dotted) and validation RMSE (dashed).

et al. (2000) and Harms et al. (2001). Basically, the expected abundance of a given species is estimated based on bootstrapped randomized habitat maps. Above (or below) a certain threshold a positive (or negative) association between the habitat and the species is assumed. We used 100 randomized habitat maps for each of the original habitat maps and a significance level of 0.01 for the analysis. For comparability, we used the same 100 randomized habitat maps for each of the three investigated groups, namely "Appendicularia", "Copepods", and "Pluteus". We reduced the resolution of our original feature-vectors ($25\,m \times 1 \times$) to $100\,m \times 1\,m$ due to computational costs.

3 | RESULTS

3.1 | Model training

Each training epoch took ~5 s using a graphic card with 768 gpu-cores and we trained the model for 250 epochs. The final training (Tr) and validation (Val) Root Mean Squared Error (RMSE) of our model were $RMSE_{Tr} \sim 0.69$ and $RMSE_{Val} \sim 0.48$ (Figure 2). As the RMSE were summed over the four features which ranged from −1 to 1, this equals a deviation of ~9% during training and ~6% during validation between the reconstructed and original parameters.

3.2 | Identification of key parameters

Temperature had the highest impact upon the output dimension 1 (D1) with a ~46% share on the final outcome, followed by salinity with ~36% and the Euclidean distance (~17%). Variations in oxygen had likely no effect on D1 since the oxygen components were weighted down to a share of only ~1%. D2 was dominated by the Euclidean distance (~45%). Temperature and salinity had an equal share of ~23% and oxygen had again the least influence but with ~9% more than on D1. D3 was affected by all input variables similarly with shares between 19% and 29%. D4 was most sensible toward salinity (~39%) followed by temperature (~27%), oxygen (~20%), and finally

the Euclidean distance (~14%). Temperature was twice as influential on D1 than on any other output dimension. Variations in oxygen had nearly no effect in D1 but increasingly so on D2, D4, and D3. Overall, salinity was the most influential variable but was never as dominant as temperature or Euclidean distance on a single output dimension. Similar to temperature on D1, the Euclidean distance was twice as influential on D2 as on any other output dimension (Table 1).

3.3 | Sensitivity

The most sensitive dimension was D3 in more than half of the cases (Figure 3). D3 was also the most balanced output dimension regarding the influence of different input parameters. The other dimensions were equally important, with D2 never being the most sensitive dimension.

3.4 | Characterization of macro-habitats

The habitat map of HE466 T5 (Figure 4) features everything one could expect from a tidal mixing front as it was described in Hill et al. (1993). At the beginning of the transect is a stratified water column. The surface layer (MH4) is outlined by the 240 μmol L^{-1} isoline for oxygen, with higher concentrations within the MH. The bottom layer (MH6) is characterized by temperatures below 12.25°C. MH7 probably resembles the area of density-driven circulation and is framed by the 13°C isoline for temperature. Toward the end of the transect (MH5), there is a fully mixed water column with salinity below 33.95 psu. The area left out by those isolines belongs to MH13, which shows the typical characteristics of an along front jet, namely reduced temperatures and upward doming of bottom front isolines.

At the beginning of the transect T1 is a stratified water column with the surface layer this time limited to temperatures above 13.5°C (MH4). The downwelling pole (MH6) is outlined by the 34.3 psu salinity isoline, with higher salinities within the MH. In contrast, the upwelling pole (MH7) is characterized by oxygen concentrations below 225 μmol L^{-1}. There is no fully mixed water column toward the end of the transect, however, what is clustered into MH5 has mostly pluteus larvae densities above 6 individuals per micro-habitat (ind mh^{-1}). The presence of an along front jet as in T5 (MH 13) could explain the accumulation of particles observed in MH5. However, the temperatures do not fully support this (Figure 5). Although it is

possible that the transect did not fully cross the tidal mixing front and this could affect the predictive ability of the model, T1 also shows strong evidence of an upwelling–downwelling dipole (Floeter et al., 2022), with the upwelling pole aligned closely with the surface temperature minimum.

As can be seen in Figure 4, the boundaries of global macrohabitats tend to fit the observed parameter isolines very well. It is noticeable that different parameters contribute to the respective isolines, in T1 there is even an MH that is most closely related to the increased abundance of pluteus larvae.

3.5 | Plankton–habitat associations

The plankton–habitat associations for all groups and habitats are presented in Table S1. The group "appendicularia" had no general positive or negative association with any of the MHs that occurred in four or more transects. There was a negative association for the groups "pluteus" and "copepods" with MH6 and MH7, which represent the stratified waters of the deeper North Sea below the thermocline. The analysis indicated a generally positive association with the surface mixed layer and the fully mixed coastal area for both groups as well.

Parameter ranges of MHs with a positive association with any specific group of organisms did not indicate a distinct physical niche (Figure 6). An exception was the positive association of copepods with MHs of higher temperatures and a negative association with MHs of lower temperatures. Furthermore, the analysis indicated an affinity of appendicularia toward the areas of lower oxygen concentrations.

4 | DISCUSSION

4.1 | Model inputs

The model had difficulty interpreting plankton abundance, most likely due to the skewness of the data. The patchiness of plankton produces many zeros and only relatively few high abundances. Using the Euclidean distance had two main advantages. First, the distance was no longer biased toward zero. Secondly, the Euclidea distance combined the data on the resident plankton communities, using chlorophyll a as a proxy for phytoplankton abundance, into one variable. In theory, the model can be trained with any number of species, groups, or even functional groups, if applied to data that allows such distinctions.

TABLE 1 Influence of each input variable upon an output variable.

Dimension/variable	Temperature	Oxygen	Salinity	Euclidean
D1	0.462	0.007	0.359	0.171
D2	0.234	0.091	0.225	0.449
D3	0.194	0.288	0.291	0.228
D4	0.273	0.197	0.386	0.144

Note: Output variables (dimensions) in rows and input variables in columns.

Furthermore, the feature vectors we use describe the Eltonian niche space (Dehling & Stouffer, 2018) at a specific location and the final macro-habitat is therefore basically a cluster of locations with very similar nic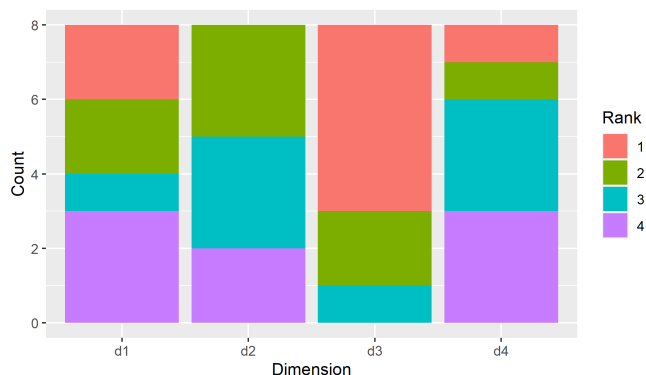he spaces. This makes our method uniquely qualified to study species-habitat interactions and ecological niches, which is of utmost importance considering the ongoing climatic changes (Benedetti et al., 2021; Busseni et al., 2020; Cael et al., 2021). This would require a more detailed dataset than the one used in the current investigation, but there are advantages in using a less detailed dataset as applied in this study. The fully automated data pipeline makes it possible to get an initial near real-time assessment of the pelagic habitats and associated species communities currently present in the study area and therefore helps monitor ongoing changes.

FIGURE 3 Count of Sobol sensitivity rankings for all transects by dimension. Fixing the dimension with the highest rank (1) would result in the greatest average reduction in output variability.

4.2 | Biophysical characteristics of macro-habitats

Transects T1 to T6 of HE466 (Figures S1–S6) were not only in the immediate vicinity of an offshore wind farm (OWF) but partly also crossed a tidal mixing front (Hill et al., 1993). While in T1–T3 all MHs had mean salinities >34 psu, there was a decreasing trend in salinity in the transects T4–T6, indicating a shift from North Sea to coastal water (Lee, 1980). Only T5 had an MH that spanned the entire water column with a mean salinity <34 psu, indicating a fully mixed habitat
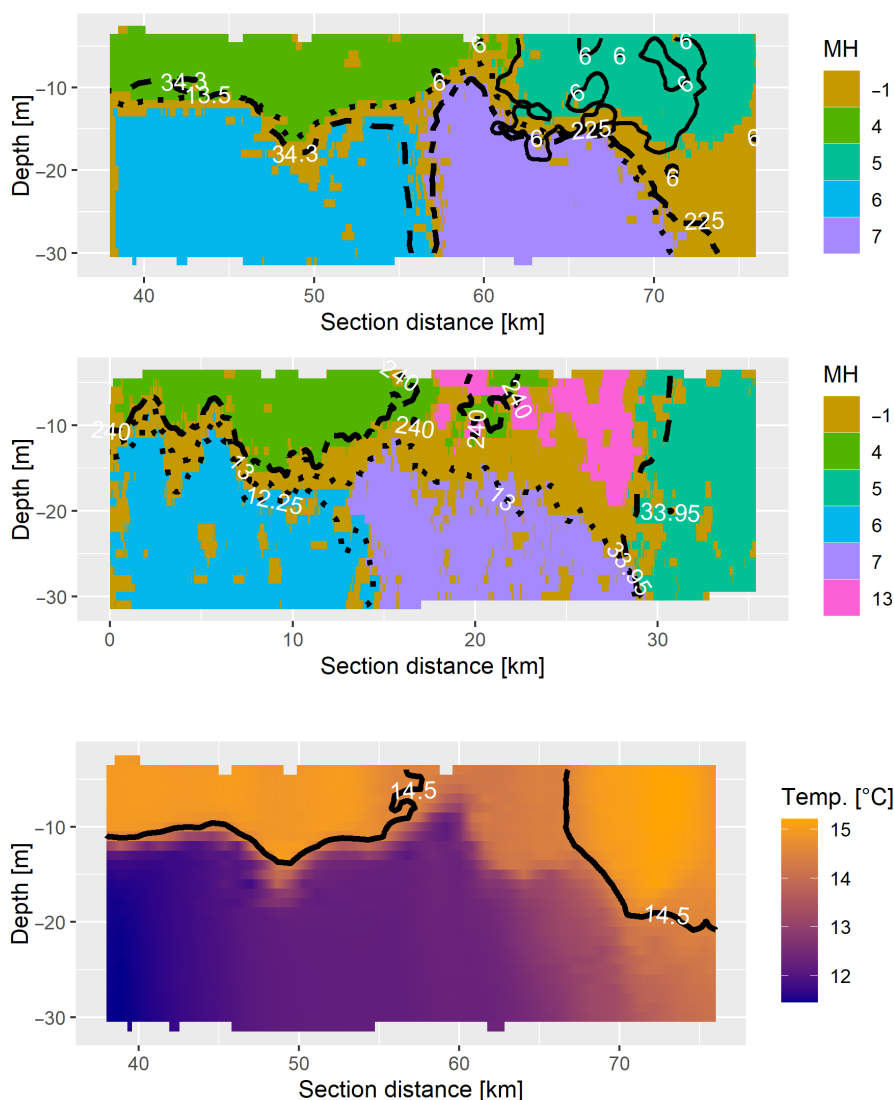

FIGURE 4 Habitat maps for T1 and T5 of haul 5 from HE466. Dotted: Temperature isolines [°C], dashed: Salinity isolines [psu], dot-dashed: Oxygen isolines [μmol L^{-1}], solid: Pluteus larvae abundance [ind/micro-habitat]. Each global macro-habitat is represented by a different color. Brown areas denote the transition areas (originally labeled −1).

FIGURE 5 Map of temperature for HE466 haul 5 T1. The black line is the 14.5°C temperature isoline.
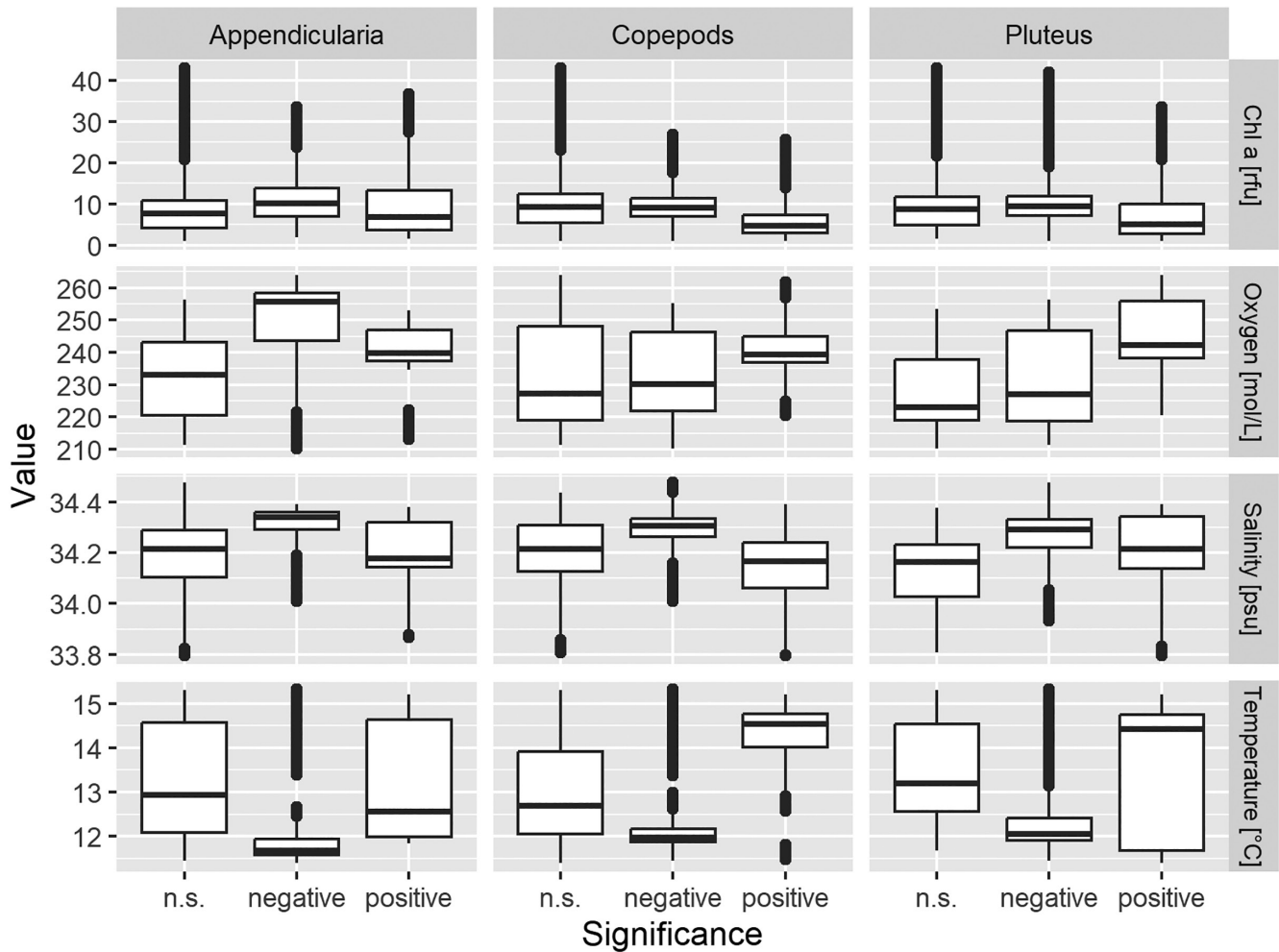
FIGURE 6 Overview for temperature [°C], salinity [psu], oxygen [mol/L], and chlorophyll a [rfu] in areas with a positive, negative, or not significant (n.s.) association with Appendicularia, Copepods, or Pluteus larvae. Positive associations indicate an increased abundance of the respective group in the underlying area and therefore the parameter ranges indicate the optimal niche space for that plankton group.

with characteristics of coastal water. T4 and T6 did not extend into the mixed coastal water.

In general, the model produced a vertical segregation along the thermocline. That is already an important feature since the mixed-layer depth is a good indicator of the productivity and biodiversity of plankton communities (Barton et al., 2013; Brun et al., 2015). Especially the distinct upwelling-downwelling pattern of the dipole in T1 of HE466 shows the ability of our model to accurately track the surface mixed layer.

The horizontal segregations are probably related to changing velocities. Those also induce some small changes in the physical characteristics, which is what our model was able to detect. The velocity patterns of dipoles (Broström, 2008) and fronts (Hill et al., 1993) are well described in the literature and closely match the habitat maps we generated.

4.3 | Plankton niches

On a global scale machine learning has successfully been applied to reevaluate existing datasets and link plankton communities to distinct physical habitats (Busseni et al., 2020; Drago et al., 2022;

Panaïotis et al., 2023; Sonnewald et al., 2020), thereby defining what has been called ecological provinces (Sonnewald et al., 2020). Especially the data from the CPR survey has frequently been used to additionally show the changing distribution patterns of a variety of planktic species in response to climate change (Barton et al., 2013; Beaugrand et al., 2009; Benedetti et al., 2021). The data used in this study covers the much smaller sub-mesoscale, where imaging and machine learning can help to improve spatial resolution (Irisson et al., 2022). Our analysis did not indicate niche segregation between the three studied plankton groups, since multiple groups were regularly found within the same MH. Furthermore, the MHs positively associated with specific groups showed no distinctly different physical properties from other areas. There are several potential reasons for the lack of readily observable niches. The lack of taxonomic details could mask the segregation of morphologically similar species of any of the groups into distinct niches. For example, the calanoid copepods *Calanus finmarchicus* Gunnerus, 1770 and *Calanus helgolandicus* Claus, 1863 both contributed significantly to copepod biomass during summer (Jónasdóttir & Koski, 2011). However, the cold-boreal *C. finmarchicus* (Planque & Fromentin, 1996) was restricted

to the cooler bottom waters while *C. helgolandicus* generally prefers warmer waters (Planque & Fromentin, 1996) and occupied the upper surface-mixed layer. The preferred temperatures for *C. finmarchicus* of <9°C (Jónasdóttir & Koski, 2011) were clearly exceeded in the present sampling area (>12°C) and considering the general northward drift of plankton distributions (Barton et al., 2013; Cael et al., 2021; Martin et al., 2021), and of *C. finmarchicus* specifically (Wilson et al., 2016), it is unlikely that this particular example caused the negative findings in this study. In general, we think it reasonable that despite the great biodiversity of the North Sea, it is possible to investigate spatial niches using VPR images since, as we argued in our introduction, the different taxa and communities are usually dominated by specific species.

Prey selection is another possible mechanism for niche diversity (Cleary et al., 2016; Fransz et al., 1991), which is not revealed using VPR images or images at all. This however would also require different species to be very abundant to mask any possible findings on the sub-mesoscale. Of course, there is the possibility that niche segregation does occur for the less common species, but those processes probably act on even smaller scales (Basterretxea et al., 2020; Mitchell et al., 2008) and are not in the scope of this investigation.

Next to temperature, which is frequently identified as a major force behind plankton distribution patterns (Busseni et al., 2020; Lindegren et al., 2020; Wilson et al., 2016; e.g. Benedetti et al., 2021; Drago et al., 2022; Houliez et al., 2021), the local hydrography is a primary driver of plankton distributions and biodiversity (Beaugrand et al., 2001; Beaugrand & Ibañez, 2002; Sonnewald et al., 2020; Swalethorp et al., 2015). Since plankton is, per definition, subjected to the currents that shape the pelagic environment, the strong velocities associated with upwelling-downwelling dipoles (Broström, 2008) and frontal jets (Hill et al., 1993) have the potential to superimpose behavioral niche segregation (Hidalgo et al., 2012). We believe that this is the most likely explanation for the lack of physically distinct niches in our relatively short transects. A similar result was observed with an Underwater Vision Profiler 6 (Panaïotis et al., 2024) and by examining a picophytoplankton dataset using machine learning (Chen et al., 2020).

In areas less dominated by strong currents like the vast, oligotrophic parts of the Atlantic and Pacific Ocean, machine learning approaches as presented in this study can improve our understanding of how physical gradients shape plankton communities (Chen et al., 2020; Greer et al., 2023). And optical datasets are especially useful in investigating the understudied gelatinous grazers (Greer et al., 2023), which will likely become more important under increasing temperatures (Winder et al., 2017). Thus, optical sampling methods and machine learning will be important tools to investigate the changes in our oceans.

AUTHOR CONTRIBUTIONS

Rene-Marcel Plonus: Formal analysis (lead); methodology (equal); writing – original draft (lead); writing – review and editing (equal). **Jens Floeter:** Methodology (supporting); resources (lead); supervision (lead); writing – review and editing (equal).

CONFLICT OF INTEREST STATEMENT

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.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Dryad at https://doi.org/10.5061/dryad.34tmpg4s4.

ORCID

Rene-Marcel Plonus https://orcid.org/0000-0003-3175-596X

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous distributed systems.* https://www.tensorflow.org/

Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M., López-Jurado, J. L., Tintoré, J., & Alemany, F. (2014). Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution. *PLoS One, 9*(10), e109338. https://doi.org/10.1371/journal.pone.0109338

Barton, A. D., Pershing, A. J., Litchman, E., Record, N. R., Edwards, K. F., Finkel, Z. V., Kiørboe, T., & Ward, B. A. (2013). The biogeography of marine plankton traits. *Ecology Letters, 16*(4), 522–534. https://doi.org/10.1111/ele.12063

Basterretxea, G., Font-Muñoz, J. S., & Tuval, I. (2020). Phytoplankton orientation in a turbulent ocean: A microscale perspective. *Frontiers in Marine Science, 7*, 185. https://doi.org/10.3389/fmars.2020.00185

Beaugrand, G., & Ibañez, F. (2002). Spatial dependence of calanoid copepod diversity in the North Atlantic Ocean. *Marine Ecology Progress Series, 232*, 197–211. https://doi.org/10.3354/meps232197

Beaugrand, G., Ibañez, F., & Lindley, J. (2001). Geographical distribution and seasonal and diel changes in the diversity of calanoid copepods in the North Atlantic and North Sea. *Marine Ecology Progress Series, 219*, 189–203. https://doi.org/10.3354/meps219189

Beaugrand, G., Ibañez, F., Lindley, J. A., & Reid, P. C. (2002). Diversity of calanoid copepods in the North Atlantic and adjacent seas: Species associations and biogeography. *Marine Ecology Progress Series, 232*, 179–195. https://doi.org/10.3354/meps232179

Beaugrand, G., Luczak, C., & Edwards, M. (2009). Rapid biogeographical plankton shifts in the North Atlantic Ocean. *Global Change Biology, 15*(7), 1790–1803. https://doi.org/10.1111/j.1365-2486.2009.01848.x

Behrenfeld, M. J., O'Malley, R., Boss, E., Karp-Boss, L., & Mundt, C. (2021). Phytoplankton biodiversity and the inverted paradox. *ISME*

Communications, 1(1), 52. https://doi.org/10.1038/s43705-021-00056-6

Bellman, R. E., & Dreyfus, S. E. (2015). *Applied dynamic programming*. Princeton University Press.

Benedetti, F., Vogt, M., Elizondo, U. H., Righetti, D., Zimmermann, N. E., & Gruber, N. (2021). Major restructuring of marine plankton assemblages under global warming. *Nature Communications*, 12(1), 5226. https://doi.org/10.1038/s41467-021-25385-x

Bertrand, A., Grados, D., Colas, F., Bertrand, S., Capet, X., Chaigneau, A., Vargas, G., Mousseigne, A., & Fablet, R. (2014). Broad impacts of fine-scale dynamics on seascape structure from zooplankton to seabirds. *Nature Communications*, 5(1), 6239. https://doi.org/10.1038/ncomms6239

Bishop, C. M. (1995). *Neural networks for pattern recognition*. Oxford University Press.

Broström, G. (2008). On the influence of large wind farms on the upper ocean circulation. *Journal of Marine Systems*, 74(1), 585–591. https://doi.org/10.1016/j.jmarsys.2008.05.001

Brun, P., Vogt, M., Payne, M. R., Gruber, N., O'Brien, C. J., Buitenhuis, E. T., Le Quéré, C., et al. (2015). Ecological niches of open ocean phytoplankton taxa. *Limnology and Oceanography*, 60(3), 1020–1038. https://doi.org/10.1002/lno.10074

Busseni, G., Caputi, L., Piredda, R., Fremont, P., Hay Mele, B., Campese, L., Scalco, E., de Vargas, C., Bowler, C., d'Ovidio, F., Zingone, A., Ribera d'Alcalà, M., & Iudicone, D. (2020). Large scale patterns of marine diatom richness: Drivers and trends in a changing ocean. *Global Ecology and Biogeography*, 29(11), 1915–1928. https://doi.org/10.1111/geb.13161

Cael, B. B., Dutkiewicz, S., & Henson, S. (2021). Abrupt shifts in 21st-century plankton communities. *Science Advances*, 7(44), eabf8593. https://doi.org/10.1126/sciadv.abf8593

Chen, B., Liu, H., Xiao, W., Wang, L., & Huang, B. (2020). A machine-learning approach to modeling picophytoplankton abundances in the South China Sea. *Progress in Oceanography*, 189, 102456. https://doi.org/10.1016/j.pocean.2020.102456

Chollet, F. (2015). *Keras*. GitHub. https://github.com/fchollet/keras

Cleary, A. C., Durbin, E. G., Rynearson, T. A., & Bailey, J. (2016). Feeding by Pseudocalanus copepods in the Bering Sea: Trophic linkages and a potential mechanism of niche partitioning. *Deep Sea Research Part II: Topical Studies in Oceanography*, 134, 181–189. https://doi.org/10.1016/j.dsr2.2015.04.001

Colwell, R. K., & Rangel, T. F. (2009). Hutchinson's duality: The once and future niche. *Proceedings of the National Academy of Sciences*, 106, 19651–19658. https://doi.org/10.1073/pnas.0901650106

Dehling, D. M., & Stouffer, D. B. (2018). Bringing the Eltonian niche into functional diversity. *Oikos*, 127(12), 1711–1723. https://doi.org/10.1111/oik.05415

Drago, L., Panaïotis, T., Irisson, J.-O., Babin, M., Biard, T., Carlotti, F., Coppola, L., Guidi, L., Hauss, H., Karp-Boss, L., & Lombard, F. (2022). Global distribution of zooplankton biomass estimated by in situ imaging and machine learning. *Frontiers in Marine Science*, 9, 894372. https://doi.org/10.3389/fmars.2022.894372

Elton, C. S. (1927). *Animal ecology*. Macmillan Co.

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 in Oceanography*, 15-16, 60–77. https://doi.org/10.1016/j.mio.2016.04.003

Floeter, J., Pohlmann, T., Harmer, A., & Möllmann, C. (2022). Chasing the offshore wind farm wind-wake-induced upwelling/downwelling dipole. *Frontiers in Marine Science*, 9, 884943. https://doi.org/10.3389/fmars.2022.884943

Fransz, H. G., Colebrook, J. M., Gamble, J. C., & Krause, M. (1991). The zooplankton of the North Sea. *Netherlands Journal of Sea Research*, 28(1), 1–52. https://doi.org/10.1016/0077-7579(91)90003-J

Friedland, K. D., Bachman, M., Davies, A., Frelat, R., McManus, M. C., Morse, R., Pickens, B. A., Smoliński, S., & Tanaka, K. (2020). Machine

learning highlights the importance of primary and secondary production in determining habitat for marine fish and macroinvertebrates. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 31, 1482–1498. https://doi.org/10.1002/aqc.3527

Fuhrman, J. A. (2009). Microbial community structure and its functional implications. *Nature*, 459(7244), 193–199. https://doi.org/10.1038/nature08058

Greer, A. T., Schmid, M. S., Duffy, P. I., Robinson, K. L., Genung, M. A., Luo, J. Y., Panaïotis, T., Briseño-Avena, C., Frischer, M. E., Sponaugle, S., & Cowen, R. K. (2023). In situ imaging across ecosystems to resolve the fine-scale oceanographic drivers of a globally significant planktonic grazer. *Limnology and Oceanography*, 68(1), 192–207. https://doi.org/10.1002/lno.12259

Grinnell, J. (1917). The niche-relationships of the California thrasher. *The Auk*, 34(4), 427–433. https://doi.org/10.2307/4072271

Harms, K. E., Condit, R., Hubbell, S. P., & Foster, R. B. (2001). Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *Journal of Ecology*, 89(6), 947–959. https://doi.org/10.1111/j.1365-2745.2001.00615.x

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2

Harris, P. T., & Baker, E. K. (2012). 1 - why map benthic habitats? In P. T. Harris & E. K. Baker (Eds.), *Seafloor geomorphology as benthic habitat* (pp. 3–22). Elsevier. https://doi.org/10.1016/B978-0-12-385140-6.00001-3

Hays, G., Richardson, A., & Robinson, C. (2005). Climate change and marine plankton. *Trends in Ecology and Evolution*, 20, 337–344. https://doi.org/10.1016/j.tree.2005.03.004

Hesselbarth, M. H. k. (2021). Shar: An 'r' package to analyze species-habitat associations using point pattern analysis. *Journal of Open Source Software*, 6(68), 3811. https://doi.org/10.21105/joss.03811

Hickel, W. (1975). The mesozooplankton in the wadden sea of Sylt (North Sea). *Helgoländer Wissenschaftliche Meeresuntersuchungen*, 27(3), 254–262. https://doi.org/10.1007/BF01611694

Hidalgo, M., Gusdal, Y., Dingsør, G. E., Hjermann, D., Ottersen, G., Stige, L. C., Melsom, A., & Stenseth, N. C. (2012). A combination of hydrodynamical and statistical modelling reveals non-stationary climate effects on fish larvae distributions. *Proceedings of the Royal Society B: Biological Sciences*, 279(1727), 275–283. https://doi.org/10.1098/rspb.2011.0750

Hill, A. E., James, I., Linden, P., Matthews, J., Prandle, D., Simpson, J., Gmitrowicz, E., Smeed, D. A., Lwiza, K. M. M., Durazo, R., Fox, A. D., & Bowers, D. G. (1993). Dynamics of tidal mixing fronts in the North Sea. *Philosophical Transactions of The Royal Society B: Biological Sciences*, 343, 431–446. https://doi.org/10.1098/rsta.1993.0057

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

Houliez, E., Lefebvre, S., Dessier, A., Huret, M., Marquis, E., Bréret, M., & Dupuy, C. (2021). Spatio-temporal drivers of microphytoplankton community in the Bay of Biscay: Do species ecological niches matter? *Progress in Oceanography*, 194, 102558. https://doi.org/10.1016/j.pocean.2021.102558

Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415–427. https://doi.org/10.1101/SQB.1957.022.01.039

Hutchinson, G. E. (1961). The paradox of the plankton. *The American Naturalist*, 95, 137–145.

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. *Annual Review of Marine Science*, 14(1), 277–301. https://doi.org/10.1146/annurev-marine-041921-013023

Iwanaga, T., Usher, W., & Herman, J. (2022). Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environmental Systems Modelling*, 4, 18155. https://doi.org/10.18174/sesmo.18155

Jónasdóttir, S. H., & Koski, M. (2011). Biological processes in the North Sea: Comparison of *Calanus helgolandicus* and *Calanus finmarchicus* vertical distribution and production. *Journal of Plankton Research*, 33(1), 85–103. https://doi.org/10.1093/plankt/fbq085

Kirby, R. R., Beaugrand, G., Lindley, J. A., Richardson, A. J., Edwards, M., & Reid, P. C. (2007). Climate effects and benthic–pelagic coupling in the North Sea. *Marine Ecology Progress Series*, 330, 31–38.

Knust, R., Nixdorf, U., & Hirsekorn, M. (2017). Research vessel HEINCKE operated by the Alfred-Wegener-institute. *Journal of Large-Scale Research Facilities*, 3, A120. https://doi.org/10.17815/jlsrf-3-164

Laakmann, S., Boos, K., Knebelsberger, T., Raupach, M. J., & Neumann, H. (2016). Species identification of echinoderms from the North Sea by combining morphology and molecular data. *Helgoland Marine Research*, 70(1), 18. https://doi.org/10.1186/s10152-016-0468-5

Landry, M., Peterson, W., & Fagerness, V. (1994). Mesozooplankton grazing in the Southern California bight. I. Population abundances and gut pigment contents. *Marine Ecology Progress Series*, 115, 55–71. https://doi.org/10.3354/meps115055

Lapointe, L., & Bourget, E. (1999). Influence of substratum heterogeneity scales and complexity on a temperate epibenthic marine community. *Marine Ecology Progress Series*, 189, 159–170.

Lee, A. J. (1980). Chapter 14 North Sea: Physical oceanography. In F. T. Banner, M. B. Collins, & K. S. Massie (Eds.), *Elsevier oceanography series (The North-West European Shelf Seas: The Sea Bed and the Sea in Motion II. Physical and Chemical Oceanography, and Physical Resources)* (pp. 467–493). Elsevier. https://doi.org/10.1016/S0422-9894(08)71359-X

Leibold, M. A., & McPeek, M. A. (2006). Coexistence of the niche and neutral perspectives in community ecology. *Ecology*, 87(6), 1399–1410. https://doi.org/10.1890/0012-9658(2006)87[1399:COTNAN]2.0.CO;2

Lindegren, M., Thomas, M. K., Jónasdóttir, S. H., Nielsen, T. G., & Munk, P. (2020). Environmental niche separation promotes coexistence among ecologically similar zooplankton species—North Sea copepods as a case study. *Limnology and Oceanography*, 65(3), 545–556. https://doi.org/10.1002/lno.11322

Lindley, J. A., Gamble, J. C., & Hunt, H. G. (1995). A change in the zooplankton of the central North Sea (55° to 58° n): A possible consequence of changes in the benthos. *Marine Ecology Progress Series*, 119(1), 299–303.

Lombard, F., Boss, E., Waite, A. M., Vogt, M., Uitz, J., Stemmann, L., Sosik, H. M., Schulz, J., Romagnan, J. B., Picheral, M., & Pearlman, J. (2019). Globally consistent quantitative observations of planktonic ecosystems. *Frontiers in Marine Science*, 6, 196. https://doi.org/10.3389/fmars.2019.00196

Lombard, F., Legendre, L., Picheral, M., Sciandra, A., & Gorsky, G. (2010). Prediction of ecological niches and carbon export by appendicularians using a new multispecies ecophysiological model. *Marine Ecology Progress Series*, 398, 109–125. https://doi.org/10.3354/meps08273

Lopez-Urrutia, A., Harris, R. P., Acuña, J. L., Båmstedt, U., Flood, P. R., Fyhn, H. J., Gasser, B., Gorsky, G., Irigoien, X., & Martinussen, M. B. (2005). *A comparison of appendicularian seasonal cycles in four contrasting european coastal environments* (pp. 255–276). Contemporary Publishing International.

MacArthur, R. H., & MacArthur, J. W. (1961). On bird species diversity. *Ecology*, 42(3), 594–598. https://doi.org/10.2307/1932254

Martin, K., Schmidt, K., Toseland, A., Boulton, C. A., Barry, K., Beszteri, B., Brussaard, C. P. D., Clum, A., Daum, C. G., Eloe-Fadrosh, E.,

Fong, A., Foster, B., Foster, B., Ginzburg, M., Huntemann, M., Ivanova, N. N., Kyrpides, N. C., Lindquist, E., Mukherjee, S., ... Mock, T. (2021). The biogeographic differentiation of algal microbiomes in the upper ocean from pole to pole. *Nature Communications*, 12(1), 5483. https://doi.org/10.1038/s41467-021-25646-9

McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, 21(4), 178–185. https://doi.org/10.1016/j.tree.2006.02.002

McGinty, N., Barton, A. D., Record, N. R., Finkel, Z. V., & Irwin, A. J. (2018). Traits structure copepod niches in the North Atlantic and Southern Ocean. *Marine Ecology Progress Series*, 601, 109–126. https://doi.org/10.3354/meps12660

McInnes, L., Healy, J., & Astels, S. (2017). Hdbscan: Hierarchical density based clustering. *Journal of Open Source Software*, 2(11), 205. https://doi.org/10.21105/joss.00205

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2023). *Misc Functions of the Department of Statistics, ProbabilityTheory Group (Formerly: E1071), TU Wien*. Retrieved February 17, 2023, from https://www.semanticscholar.org/paper/Misc-Functions-of-the-Department-of-Statistics%2C-TU-Meyer-Dimitriadou/a0816a1a3a91d22602cdcbbfdb4d3bb4abb922a5

Mitchell, J. G., Yamazaki, H., Seuront, L., Wolk, F., & Li, H. (2008). Phytoplankton patch patterns: Seascape anatomy in a turbulent ocean. *Journal of Marine Systems*, 69(3), 247–253. https://doi.org/10.1016/j.jmarsys.2006.01.019

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

Panaïotis, T., Babin, M., Biard, T., Carlotti, F., Coppola, L., Guidi, L., Hauss, H., Karp-Boss, L., Kiko, R., Lombard, F., McDonnell, A. M. P., Picheral, M., Rogge, A., Waite, A. M., Stemmann, L., & Irisson, J. O. (2023). Three major mesoplanktonic communities resolved by in situ imaging in the upper 500 m of the global ocean. *Global Ecology and Biogeography*, 32(11), 1991–2005. https://doi.org/10.1111/geb.13741

Panaïotis, T., Poteau, A., Riquier, É. D., Catalano, C., Courchet, L., Motreuil, S., Coppola, L., Picheral, M., & Irisson, J. O. (2024). Temporal evolution of plankton and particles distribution across a mesoscale front during the spring bloom. *Limnology and Oceanography*. https://doi.org/10.1002/lno.12566

Pante, E., & Simon-Bouhet, B. (2013). Marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLOS ONE*, 8(9), e73051. https://doi.org/10.1371/journal.pone.0073051

Planque, B., & Fromentin, J.-M. (1996). Calanus and environment in the eastern North Atlantic. I. Spatial and temporal patterns of *C. finmarchicus* and *C. helgolandicus*. *Marine Ecology Progress Series*, 134(1), 101–109.

Plonus, R.-M., Conradt, J., Harmer, A., Janßen, S., & Floeter, J. (2021). Automatic plankton image classification—Can capsules and filters help cope with data set shift? *Limnology and Oceanography: Methods*, 19(3), 176–195. https://doi.org/10.1002/lom3.10413

Plonus, R.-M., Vogl, S., & Floeter, J. (2021). Automatic segregation of pelagic habitats. *Frontiers in Marine Science*, 8, 754375. https://doi.org/10.3389/fmars.2021.754375

Plotkin, J. B., Potts, M. D., Leslie, N., Manokaran, N., Lafrankie, J., & Ashton, P. S. (2000). Species-area curves, spatial aggregation, and habitat specialization in tropical forests. *Journal of Theoretical Biology*, 207(1), 81–99. https://doi.org/10.1006/jtbi.2000.2158

RCoreTeam. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Schlitzer, R. (2020). *Ocean data view*. Retrieved March 30, 2021, from https://odv.awi.de/

Siegel, D. A. (1998). Resource competition in a discrete environment: Why are plankton distributions paradoxical? *Limnology and*

Oceanography, 43(6), 1133–1146. https://doi.org/10.4319/lo.1998.43.6.1133

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1), 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6

Sonnewald, M., Dutkiewicz, S., Hill, C., & Forget, G. (2020). Elucidating ecological complexity: Unsupervised learning determines global marine eco-provinces. *Science Advances*, 6(22), eaay4740. https://doi.org/10.1126/sciadv.aay4740

Swalethorp, R., Malanski, E., Dalgaard Agersted, M., Gissel Nielsen, T., & Munk, P. (2015). Structuring of zooplankton and fish larvae assemblages in a freshwater-influenced Greenlandic fjord: Influence from hydrography and prey availability. *Journal of Plankton Research*, 37(1), 102–119. https://doi.org/10.1093/plankt/fbu099

Thompson, A. F., Lazar, A., Buckingham, C., Naveira Garabato, A. C., Damerell, G. M., & Heywood, K. J. (2016). Open-ocean submesoscale motions: A full seasonal cycle of mixed layer instabilities from gliders. *Journal of Physical Oceanography*, 46(4), 1285–1307. https://doi.org/10.1175/JPO-D-15-0170.1

Troupin, C., Barth, A., Sirjacobs, D., Ouberdous, M., Brankart, J.-M., Brasseur, P., Rixen, M., Alvera-Azcárate, A., Belounis, M., Capet, A., Lenartz, F., Toussaint, M. E., & Beckers, J. M. (2012). Generation of analysis and consistent error fields using the data interpolating variational analysis (DIVA). *Ocean Modelling*, 52-53, 90–101. https://doi.org/10.1016/j.ocemod.2012.05.002

Van Rossum, G., & Drake, F. L. (2009). *Python 3 reference manual*. CreateSpace.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., ... Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. https://doi.org/10.21105/joss.01686

Wilson, R. J., Heath, M. R., & Speirs, D. C. (2016). Spatial modeling of *Calanus finmarchicus* and *Calanus helgolandicus*: Parameter differences explain differences in biogeography. *Frontiers in Marine Science*, 3, 157. https://doi.org/10.3389/fmars.2016.00157

Winder, M., Bouquet, J., Rafael Bermúdez, J., Berger, S. A., Hansen, T., Brandes, J., Sazhin, A. F., Nejstgaard, J. C., Bå, U., Jakobsen, H., Dutz, J., Frischer, M., Troedsson, C., & Thompson, E. M. (2017). Increased appendicularian zooplankton alter carbon cycling under warmer more acidified ocean conditions. *Limnology and Oceanography*, 62(4), 1541–1551. https://doi.org/10.1002/lno.10516

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.