Science of the Total Environment 952 (2024) 175793

Contents lists available at ScienceDirect

Science of the Total Environment

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

Predictors of long-term variability in NE Atlantic plankton communities

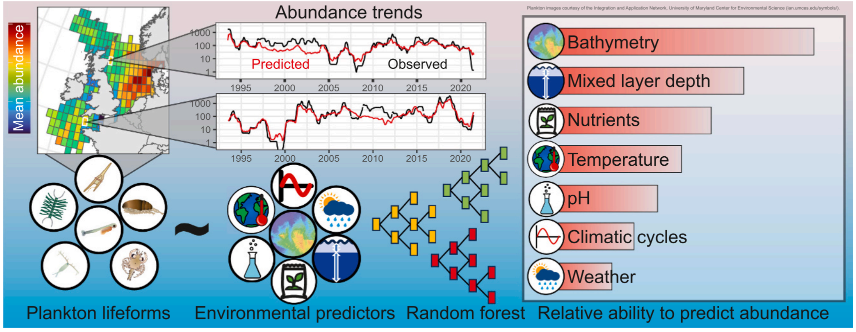
Matthew M. Holland[a,*], Angus Atkinson[b], Mike Best[c], Eileen Bresnan[d], Michelle Devlin[e], Eric Goberville[f], Pierre Hélaouët[g], Margarita Machairopoulou[d], Matthew Faith[a], Murray S.A. Thompson[e], Abigail McQuatters-Gollop[a]

[a] *Marine Conservation Research Group, University of Plymouth, Drake Circus, Plymouth PL4 8AA, UK*
[b] *Plymouth Marine Laboratory, Prospect Place, The Hoe, Plymouth PL1 3DH, UK*
[c] *Environment Agency, Quay House, Floor 6, 2 East Station Road, Fletton Quays, Peterborough PE2 8YY, UK*
[d] *Marine Directorate of the Scottish Government, 375 Victoria Road, AB11 9DB Aberdeen, Scotland, UK*
[e] *Centre for Environment, Fisheries and Aquaculture Science (Cefas), Pakefield Road, Lowestoft NR33 0HT, UK*
[f] *Unité Biologie des Organismes et Ecosystèmes Aquatiques (BOREA), Muséum National d'Histoire Naturelle, CNRS, IRD, Sorbonne Université, Université de Caen Normandie, Université des Antilles, Paris, France*
[g] *The Marine Biological Association (MBA), The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK*

HIGHLIGHTS

- Human-induced pressures drive rapid marine changes, impacting plankton communities.
- Random forest analysis ranked environmental variables to predict plankton abundance.
- Best predictors were bathymetry, and mixed layer depth and potential food groups.
- Nitrogen-to-phosphorus also ranked high, indicating links to nutrient imbalance.
- Warming-induced stratification and nutrient starvation are key drivers of change.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Olga Pantos

Keywords:
Continuous plankton recorder
OSPAR
Plankton
Pelagic habitats
Climate change
Nutrients

ABSTRACT

Anthropogenic pressures such as climate change and nutrient pollution are causing rapid changes in the marine environment. The relative influence of drivers of change on the plankton community remains uncertain, and this uncertainty is limiting our understanding of sustainable levels of human pressures. Plankton are the primary energy resource in marine food webs and respond rapidly to environmental changes, representing useful indicators of shifts in ecosystem structure and function. Categorising plankton into broad groups with similar characteristics, known as "lifeforms", can be useful for understanding ecological patterns related to environmental change and for assessing the state of pelagic habitats in accordance with the EU Marine Strategy Framework Directive and the OSPAR Commission, which mandates protection of the North-East Atlantic. We analysed 29 years of Continuous Plankton Recorder data (1993–2021) from the North-East Atlantic to examine how trends in plankton lifeform abundance changed in relation to one another and across gradients of environmental change associated with human pressures. Random forest models predicted between 57 % and 80 % of

---

* Corresponding author.
E-mail address: `matt.holland@plymouth.ac.uk` (M.M. Holland).

https://doi.org/10.1016/j.scitotenv.2024.175793
Received 10 May 2024; Received in revised form 5 August 2024; Accepted 23 August 2024

the variability in lifeform abundance, based on data not used to train the models. Observed variability was mainly explained by trends in other lifeforms, with mainly positively correlated trends, indicating bottom-up control and/or shared responses to environmental variability were prevalent. Longitude, bathymetry, mixed layer depth, the nitrogen-to-phosphorus ratio, and temperature were also significant predictors. However, contrasting influences of environmental drivers were detected. For example, small copepod abundance increased in warmer conditions whereas meroplankton, large copepods and fish larvae either decreased or were unchanged. Our findings highlight recent changes in stratification, reflected by variation in mixed layer depth, and imbalanced nutrient ratios are affecting multiple lifeforms, impacting the North-East Atlantic plankton community. To achieve environmental improvements in North-East Atlantic pelagic habitats, it is crucial that we continue to address climate change and reduce nutrient pollution.

1. Introduction

Changes in the abundance and composition of plankton, combined with climate change and nutrient pollution, indicate widespread shifts in pelagic habitat conditions (Holland et al., 2023c; Louchart et al., 2023a, 2023b). However, we still need to identify the key drivers of change to focus resources effectively and achieve the environmental improvements advocated internationally. In the EU and the UK, the state of the marine environment is regularly assessed under national and international frameworks to achieve Good Environmental Status (GES) for specific descriptors. One such framework is the OSPAR (Oslo and Paris) Convention, an international agreement among 16 contracting parties (15 nations and the European Union) dedicated to sustainably managing human activities to protect the North-East Atlantic's shared marine environment. The OSPAR Convention (McQuatters-Gollop et al., 2022), the United Kingdom Marine Strategy (UKMS) (McQuatters-Gollop et al., 2019a), and the European Union's Marine Strategy Framework Directive (MSFD) (Magliozzi et al., 2021; McQuatters-Gollop et al., 2015) all use plankton as indicators to monitor and assess biodiversity in pelagic habitats. Under the MSFD, descriptor D1C6 outlines objectives for achieving GES in pelagic habitats, stating: "*The condition of the habitat type, including its biotic and abiotic structure and its functions (e.g., its typical species composition and their relative abundance, absence of particularly sensitive or fragile species or species providing a key function, size structure of species), is not adversely affected due to anthropogenic pressures.*" (European Commission, 2017). The uncertainty surrounding the primary causes of change in the environmental status of pelagic habitats — and how best to recover them once degraded — represents a critical knowledge gap that could hinder progress towards achieving GES.

Biodiversity indicators are used to monitor changes in the status of ecosystem health to help inform decisions on sustainable marine ecosystem management (e.g., to inform progress towards achieving GES) (Magliozzi et al., 2021; McQuatters-Gollop et al., 2022; McQuatters-Gollop et al., 2019b). Plankton typically have short generation times and respond rapidly to environmental change, particularly to alterations in temperature (Chivers et al., 2017), nutrient levels (Van De Poll et al., 2013), and in some cases ocean acidity (Alvarez-Fernandez et al., 2018), making them sensitive biodiversity indicators of environmental shifts (Bedford et al., 2020b; Di Pane et al., 2022; McQuatters-Gollop et al., 2015). Biodiversity indicators for plankton are well-suited for providing early warning signals of important changes in ecosystem structure and function, including shifts in food web dynamics and disruptions to nutrient cycling (Arrigo, 2005). Studying plankton communities in this context offers a crucial tool for early detection and assessment of environmental changes, providing critical insights to inform conservation and management efforts (Bedford et al., 2020b; Holland et al., 2023a; McQuatters-Gollop et al., 2019a).

Identifying significant changes in the plankton community by studying individual species is challenging because many species are adapted to specific ecological and hydrodynamic conditions (McQuatters-Gollop et al., 2019a). Regional biodiversity assessments span a variety of hydrodynamic regimes, which result in the highly patchy distribution of plankton species across space (Van Leeuwen et al., 2016). Moreover, individual species exhibit considerable temporal variability, as short generation times for some taxa lead to fluctuations in species abundance, ultimately affecting community structure. Assessing species individually is also inadequate for evaluating overall changes in the plankton community, as no single species can represent the state of the entire community. Additionally, clear relationships between individual species and ecosystem functions are rarely evident across large spatial scales (Tett et al., 2013).

Plankton lifeforms are groupings of species that perform similar important functional roles in their environment (Tett et al., 2008). Categorising plankton into "lifeforms" — broad groups defined by shared ecological, trophic, morphological, or taxonomic traits (McQuatters-Gollop et al., 2019b) — can enhance our understanding of how changes in their abundance might impact higher trophic levels (Atkinson et al., 2024) and ecosystem services (Bedford et al., 2020b). Since lifeforms are defined by shared functional traits, variation in their abundance or biomass should be more closely related to ecosystem structure and function versus individual taxa (McQuatters-Gollop et al., 2019a). Lifeforms are also better suited than individual taxa for regional assessments because they help to smooth out much of the spatial and temporal variability inherent to individual species distributions (Bedford et al., 2020b). The lifeform approach also supports intercomparison of datasets from different institutions, with varying levels of taxonomic resolution, making it a valuable tool for regional assessments (McQuatters-Gollop et al., 2019a).

Considering the role of plankton in food webs, even moderate changes in the quantity or availability of important plankton lifeforms driven by environmental variability can generate measurable impacts across marine ecosystems. For example, large copepods are key prey for fish (Van Deurs et al., 2015) and changes in their abundance relative to small copepods can affect how energy flows across food webs because species interactions are largely determined by body size (Brose et al., 2019; Brose et al., 2006). Changes in the abundances of plankton lifeforms over short time scales can have immediate consequences for higher trophic levels (e.g., variation in zooplankton abundance closely associated with biomass of sandeel larvae; Frederiksen et al., 2006), whereas gradual change over decades can indicate long-term shifts in important aspects of ecosystem functioning (Di Pane et al., 2022). Categorising plankton as lifeforms can also facilitate the measurement of ecological patterns associated with warming and nutrient pollution (Bedford et al., 2020b). For example, changes in nutrient availability can impact the structure of the phytoplankton community (Schmidt et al., 2020), which can impact copepod abundance (Makareviciute-Fichtner et al., 2020). Declining abundance of copepods subsequently impacts the growth and survival of planktivorous fish populations (Olin et al., 2022). Despite the large environmental changes currently impacting temperate marine ecosystems, the relative influence of environmental variables on plankton lifeforms remains uncertain.

The recent OSPAR assessment of pelagic habitats indicator PH1/FW5 "Changes in Phytoplankton and Zooplankton Communities" (Holland et al., 2023a) linked trends in plankton lifeform abundance with trends in environmental variables at the scale of sub-regional spatial assessment units, or the set of COMP4 polygons used to delineate internally consistent pelagic habitat types for OSPAR eutrophication and pelagic habitats assessments (Enserink et al., 2019; Graves et al., 2023). The assessment concluded that recent changes in the abundance of plankton lifeforms were most closely linked to rising temperatures in offshore areas in the North-East Atlantic, and to variation in nutrient concentrations in some coastal areas (Holland et al., 2023a). This study builds on that work by exploring the plausible links between environmental variability and long-term changes in plankton communities in the North-East Atlantic at regional scale (Holland et al., 2023c).

Due to the inherent patchiness of plankton distribution in both space and time, monitoring and environmental data required to detect significant changes must be geographically and temporally extensive. To address this, we use the Continuous Plankton Recorder (CPR) Survey, which employs plankton sampling instruments towed behind ships of opportunity (Richardson et al., 2006). The CPR Survey provides the most comprehensive record of marine plankton biodiversity on Earth (Vezzulli et al., 2022). From 1993 to 2021, during which concurrent environmental data from the NEMO-ERSEM ecosystem model were available (Madec et al., 2017), the CPR survey dataset includes over 58,000 observations from across the North-East Atlantic.

The aims of this study were to enhance the understanding of the primary drivers of change in plankton lifeforms: first, by assessing whether changes in the abundance of a selection of key plankton lifeforms are principally linked to temperature, nutrients, other plankton lifeforms, or climate cycles; second, by studying whether lifeforms have unique responses to environmental change (e.g., contrasting responses to temperature); and third, by assessing whether models based on trends in other plankton lifeforms and environmental variables can accurately predict trends in lifeform abundance.

2. Materials and methods

2.1. Spatial scale of the analysis

The North-East Atlantic, encompassing the North Sea, Celtic Seas and the Bay of Biscay and Iberian Coast (Fig. 1), is characterised by large environmental gradients in temperature, salinity, nutrient availability, and light, which affect the abundance and composition of the plankton assemblage (Kléparski et al., 2021). Variations in nutrient availability, driven by complex oceanographic processes and riverine discharge, create localised hotspots of plankton activity (Xu et al., 2020), while changing light availability throughout the year affects seasonal blooms (Llope et al., 2009). Anthropogenic pressures like nutrient pollution (McQuatters-Gollop et al., 2009; McQuatters-Gollop et al., 2007) and fishing (Frederiksen et al., 2006; Lindegren et al., 2018; Thompson et al., 2020) add additional complexity to the dynamics of this region.

Spatial resolution is an important factor which can influence results of many analyses (Scott et al., 2023). For this analysis, we selected a resolution that captured a sufficient number of CPR samples per cell, while maintaining the fine-scale complexity of the gridded environmental data. We assessed four levels of resolution for our analysis: 0.5, 1, 2 and 4°. We calculated the mean number of CPR samples per grid cell (± SE) to be 0.5°: 67 ± 4, 1°: 225 ± 19, 2°: 657 ± 88 and 4°: 2312 ± 459. We ultimately selected 1° spatial resolution to balance capturing fine scale environmental gradients, while maintaining statistically robust sample sizes.

To assess model predictions of lifeform abundance, we examined two specific grid cells as comparative case studies, which we have named Channel Well Mixed (CWM) and Northern North Sea (NNS) (indicated in Fig. 1) to correspond with two of the spatial assessment units used in the recent OSPAR assessment of pelagic habitats (OSPAR, 2023). These two regions were selected due to their ample survey coverage and because they represent two distinct pelagic habitat types (Philippart et al., 2011), with CWM characterised by permanently mixed conditions in a coastal area of the English Channel, and NNS by permanently mixed and intermittently stratified conditions for a large shelf area of the northern North Sea (van Leeuwen et al., 2015).

2.2. Plankton monitoring data

We used plankton abundance monitoring data from the extensive pan-European Continuous Plankton Recorder (CPR) survey (Richardson et al., 2006). We used an extraction from the CPR dataset bounded between 13° W and 11° E and between 44 and 63° N. Plankton samples were collected and fixed autonomously in situ using a standardised mechanical sampling apparatus with 270 μm mesh size (Richardson et al., 2006) which is towed along regular routes at an approximate depth of ~7 m (Hays and Warner, 1993) using ships of opportunity

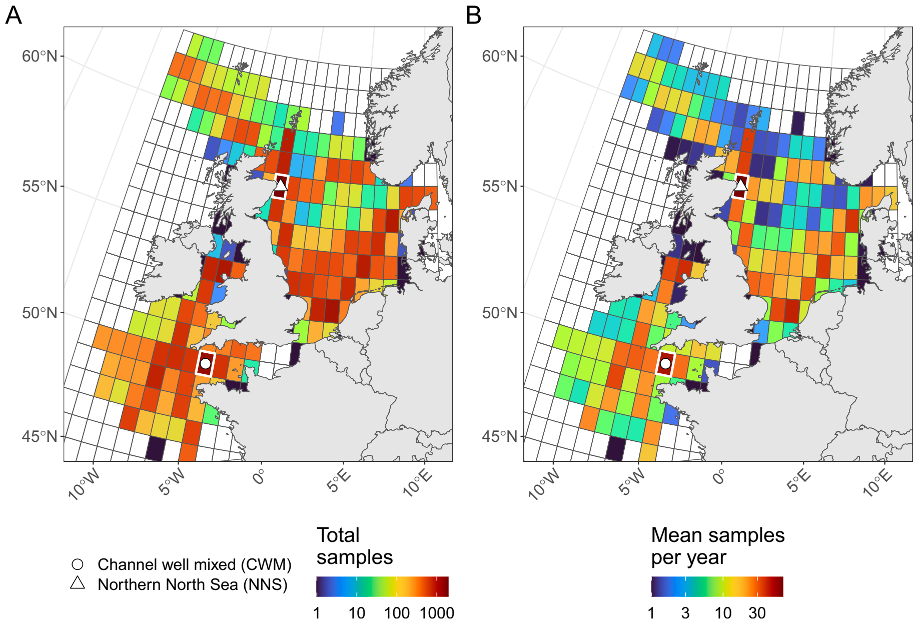


Fig. 1. The total number of CPR samples (A) and mean number of samples per year (B) from 1993 to 2021 overlaid across a 1° square grid. Unfilled grid cells were poorly sampled by the CPR and excluded from this analysis (i.e., >20 % of months missing). Two grid cells (indicated with white symbols) within the Channel Well Mixed (CWM; circle) and Northern North Sea (NNS; triangle) regions represent spatial areas where plankton lifeform abundance was investigated as a case study (as described in the Materials and methods).

(McQuatters-Gollop et al., 2015). Each CPR sample represents a transect of 10 nautical miles, constituting ~3 m^{-3} of seawater filtered (Richardson et al., 2006). Due to hazards associated with deploying towed gear in shallow water, CPR sampling is typically restricted to offshore waters and is thus not representative of conditions in coastal waters. CPR data are available for a large proportion of the North-East Atlantic, North Sea and Celtic Seas. Because plankton distribution is patchy in both space and time, CPR data are best analysed at an aggregated monthly time scale (Richardson et al., 2006). While CPR data from 1958 to 2021 were available, our analysis was restricted to samples collected between 1993 and 2021 (Fig. 1) due to the availability of concurrent model hind-cast environmental data for this period only.

2.3. Plankton lifeforms

To define the plankton lifeforms, taxa were categorised based on shared traits, including taxonomic grouping, body size, and habitat (McQuatters-Gollop et al., 2019a; Tett et al., 2008). All lifeforms were defined based on our master taxa list of functional traits (https://doi.mba.ac.uk/data/3260/1) (Ostle et al., 2021). For simplicity, we use the term "functional traits" to encompass various types of traits, including ecological, trophic, morphological, and taxonomic. This list was populated with functional trait information for each taxon or group (e.g., egg, larvae, post-larvae) counted in the CPR dataset. All taxa have been manually assigned functional traits from expert taxonomist knowledge and scientific literature. Every entry in the master taxa list has also been assigned a unique Aphia ID via the World Register of Marine Species (WoRMS), to provide a means of matching functional traits to plankton datasets; all taxa recorded in the CPR Survey are represented in the current master list. These classifications were used to calculate the total abundance of each lifeform for each sample (i.e., individuals per 3 m^{-3} CPR sample), pooling members of the same lifeform regardless of life stage (e.g., copepodites and adult copepods) (McQuatters-Gollop et al., 2019a). A detailed list of these classifications can be found in the Supplementary materials (Table S 1).

To understand the key environmental drivers of change in plankton lifeforms, we selected six lifeforms based on the high confidence in their classification (McQuatters-Gollop et al., 2019a) and because they shared no common taxa with one another. For each lifeform, the total number of unique categories (n) counted in the CPR Survey is as follows: diatoms ($n = 58$), dinoflagellates ($n = 36$), large copepods ($n = 51$), small copepods ($n = 23$), meroplankton ($n = 13$), and fish larvae plus fish eggs ($n = 2$).

Diatoms and dinoflagellates represent the phytoplankton community component of this analysis (McQuatters-Gollop et al., 2019a). Both lifeforms contribute to pelagic food webs and are functionally distinct due to differences in motility (dinoflagellates propel themselves with flagella), nutrient composition (e.g., diatoms have a silicious cell wall), and trophic strategies (diatoms are autotrophic, whereas dinoflagellates can be mixotrophic). Diatoms and dinoflagellates also tend to exhibit different patterns in seasonality, so their importance in supporting the food web varies depending on the time of year. Their low trophic level means that they are particularly sensitive to changes in water quality and can respond quickly to variation in nutrient concentration. It is important to note that the CPR tends to be selective in its sampling of diatoms and dinoflagellates, with tendency to preferentially capture large or spikey taxa, therefore many of the numerically dominant taxa typically caught in bottle samples are under-represented (Richardson et al., 2006). The CPR also tends to under-represent athecate dinoflagellates, which are damaged by the formalin preservative (Batten et al., 2018). However, sampling methods have been broadly consistent across time and space, so intercomparisons remain valid (McQuatters-Gollop et al., 2015).

For this analysis, we defined large copepods as those with typical adult body length (total length) $\geq$2 mm, and small copepods as those <2 mm (McQuatters-Gollop et al., 2019a). These measurements were based on current scientific literature and did not account for developmental stages. Small copepods are typically the numerically dominant zooplankton lifeform and are thus important for the transfer of energy from phytoplankton to higher trophic levels (e.g., larval and planktivorous fish). Large copepods represent lipid-rich, but less abundant, important prey items for juvenile and planktivorous fish (Capuzzo et al., 2018; Lynam et al., 2017).

Meroplankton are a diverse group containing the planktonic larvae of benthic organisms. For the current study, meroplankton included polychaetes, bivalves, cephalopods, cirripedes, decapods, stomatopods, echinoderms, bryozoans, lancelets, and nematode parasites. Meroplankton spend only the juvenile portion of their lifecycle in the pelagic ecosystem, before recruiting to benthic habitats. For this analysis we did not include any cnidarians with benthic polyps or larvae of benthic fish in the meroplankton lifeform.

Fish larvae and eggs represent the reproductive output of broadcast spawning benthic and pelagic fish, which may travel for considerable time and distance before hatching. This lifeform represents part of the next trophic level, as fish larvae rely heavily on both large and small copepods. For Atlantic cod, the recruitment success of a year-class is highly dependent on whether copepods are in high abundance during the period when larval fish deplete their yolk sacs (Endo et al., 2022). The higher trophic position of fish may limit the influence of environmental variability and they might be more affected by bottom-up effects driven by duration of overlap between larval fish and copepod abundance (Kristiansen et al., 2011).

2.4. Time-series construction

All data analysis was performed using R programming language (R Core Team, 2020). We decided to assess plankton lifeform abundance rather than biomass to avoid the added uncertainty and complexity of estimating biomass and to maintain consistency with our previous works (Holland et al., 2023a; Holland et al., 2023c) upon which the current study builds.

Abundance values of each lifeform within each sample were grouped by grid cell intersection to divide data spatially. Total lifeform abundance values were averaged per month within each grid cell before being $\log_{10}(x + 1)$ transformed to stabilise the exponential growth and decline typical of plankton time-series, to reduce the influence of extreme values, and to account for zero values. For months when there were no samples within a particular grid cell, gaps were filled by extracting a mean value from an inverse distance weighted interpolated surface generated from $\log_{10}(x + 1)$ transformed values from nearby cells, using a maximum search radius of 250 km and a minimum of five neighbouring samples. While this relatively large search radius was necessary for filling gaps in the time-series, it will inevitably cause some boundaries between separate water masses to become muddled. However, a radius of this size is very commonly used when interpolating CPR data (Beaugrand and Reid, 2012; Brun et al., 2016; Edwards et al., 2021; Pitois and Fox, 2006). The interpolation was evaluated by plotting against mean abundance values calculated from CPR samples and generating linear models to assess coefficients.

Following methods described in Bedford et al. (2020b), which combined spatial and temporal interpolation techniques, linear interpolation was performed to fill additional gaps of three months or fewer within the time-series for each grid cell. Maps displaying the total number of months derived from interpolated data are provided in the appendix (Fig. S 1). Following these steps, grid cells missing lifeform abundance values for >20 % of months across the full time-series (12 months × 29 years = 348 months) were excluded from the analysis.

2.5. Regional scale links to environmental variables and between lifeforms

We used random forest regression (Cutler et al., 2007; Prasad et al., 2006) to assess the strength of relationship between predictor variables and lifeform abundance (Fig. 2). Random forest regression generates a multitude of decision trees during training and outputs the mean prediction among the ensemble of individual trees. This method is particularly useful for assessing the relative importance among a large number of predictor variables because it can capture non-linear relationships between predictor and response variables. It can also handle high-dimensional data with correlated predictors with less tendency for overfitting and it provides a measure of variable importance, which indicates how much each predictor contributes to the predictive accuracy of the model.

2.5.1. Assessment of variable importance

The 'ranger' R package (Wright and Ziegler, 2015), a fast implementation of random forest modelling, was applied to generate all models, using the default setting of 500 trees. To maintain comparability of the six models, we did not conduct any hyperparameter tuning (e.g., number of trees, number of variables to consider at each split), since the primary focus was to evaluate variable importance, rather than generating optimised prediction models. For each model we evaluated the same set of environmental variables, and the abundance of the five other lifeforms (i.e., excluding the lifeform represented by the response variable) as predictors. Relative variable importance was evaluated for each model to determine the best predictors of lifeform abundance trends at the regional scale. Relative importance of the predictor variables in each model was evaluated based on the homogeneity in the variance of the response. Variable importance is determined based on the decrease in impurity that each variable contributes when making splits in the decision trees. The variable importance score measures how much the inclusion of a particular variable improves predictive accuracy or reduces variability in the predicted values. Variable importance scores can be assessed relative to each other within models but are not comparable across models. It is important to note that observing high importance of any variable is simply an indicator that it co-varies predictably with lifeform abundance. Although ranking variable importance in machine learning models, such as random forest, does not identify causation (e.g., Prosperi et al., 2020), the output can be highly informative for future modelling efforts and to help prioritise new research questions.

2.5.2. Environmental data processing

In situ data were not available for all environmental variables at each CPR sampling location and time, so we used gridded variables from modelled and remotely sensed environmental data, resolved at monthly resolution, and capable of capturing environmental change over the period of this dataset. We used the NOAA 1° resolution global International Comprehensive Ocean-Atmosphere Data Set (ICOADS; Freeman et al., 2017), the Copernicus Atlantic- European North West Shelf- Ocean Biogeochemistry Reanalysis, hindcast from the 0.06° (~7 km) resolution European Regional Seas Ecosystem Model (NEMO-ERSEM; Butenschön et al., 2016; Madec et al., 2017) as well as Bathymetry Product tiles from EMODnet, resolved at 1/16 × 1/16 arc minutes (~115 m grid; Thierry et al., 2019) (Table S 2). We also used North Atlantic Oscillation (NAO) and Atlantic Multidecadal Oscillation (AMO) data from NOAA's Climate Prediction Centre and Physical Sciences Laboratory, respectively. This analysis was constrained to the period 1993–2021 to match the historical extent available from the NEMO-ERSEM hindcast, which represents the most comprehensive environmental dataset currently available for the region.

Numeric variables representing longitude and latitude were included in our analysis to account for spatial variation in the abundance of plankton lifeforms between the Atlantic and North Sea (Thompson et al., 2021) which could not be attributed to the other environmental variables we tested, using the coordinates of the centroid for each grid cell.

For ICOADS and NEMO-ERSEM data, we calculated the mean of monthly mean gridded values (modelled and remotely sensed) extracted within each grid cell. For NOAA Atlantic Multidecadal Oscillation (AMO) and North Atlantic Oscillation (NAO) data, monthly values were applied identically across all assessment units since these variables have basin-scale influence likely to cover the entire assessment region. For latitude, longitude and bathymetry, location-specific values were applied identically across all months since these variables have no temporal context. To ensure that land-based measurements did not influence mean values for each grid cell (e.g., bathymetry >0 m), all land areas were subtracted from the grid prior to extracting the environmental data.

2.5.3. Decomposing time-series

All time-series, including lifeform abundance, were transformed to remove seasonality and extract the long-term trends for each grid cell by conducting additive time-series decomposition, using the 'decompose' function in R. This function extracts the trend in a time-series by applying a 12-month symmetrical moving window to generate a mean value for each month. This step caused the first and last six months of data from each time-series to be excluded. Our analysis was only concerned with long-term variation, thus this seasonality was removed despite accounting for the largest portion of variation in abundance for a large number of species (Colebrook, 1982). The 12-month symmetrical moving window also smooths over the bi-modal annual cycle typical of many phytoplankton taxa (Nogueira et al., 2012). However, the approach of pooling taxa with similar characteristics into lifeforms already eliminates much of the variation attributed to the separate taxa (Bedford et al., 2020a). While this smoothing method can be influenced by missing months in the time-series, our spatial and temporal

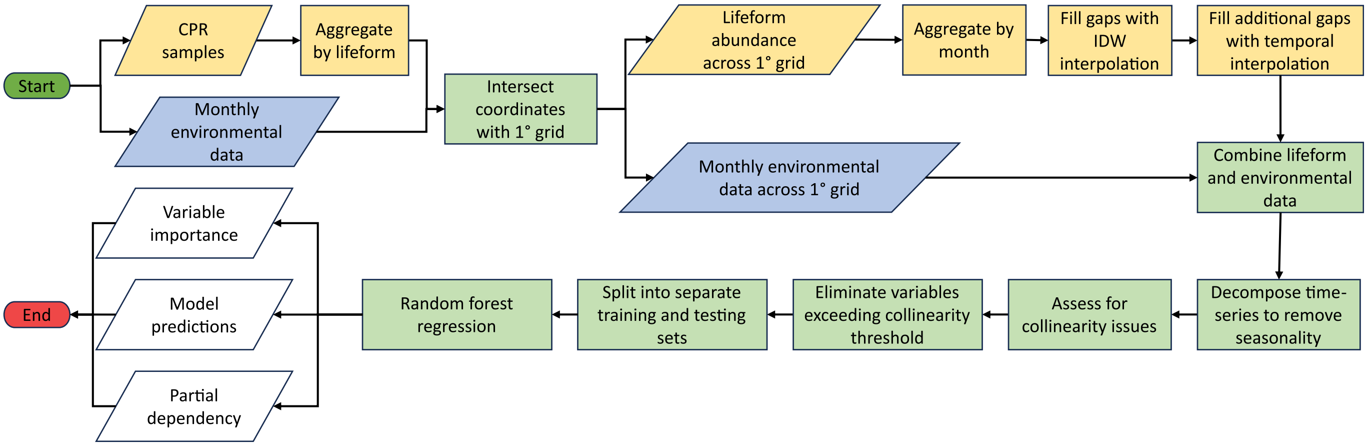


Fig. 2. Flowchart describing how the plankton (yellow) and environmental (blue) data were processed prior to being combined (green) and used to generate the six random forest regression models.

interpolation approach ensured that the number of missing months was minimised.

2.5.4. *Handling collinearity among predictors*

Although random forest regressions are generally robust to collinearities among predictors, multicollinearities can reduce the reliability of predictor importance rankings (Nicodemus et al., 2010). We therefore took steps to minimise the potential influence of collinearity. De-seasonalised predictor variables were assessed for collinearity using Spearman correlation (Fig. S 2) and any pairwise negative or positive correlations >0.7 required the exclusion of one variable in the pair to avoid distortion of model estimation and predictive performance (Dormann et al., 2013) (see Table S 2). A predictor was a more likely to be excluded over others if it demonstrated high correlation with multiple other predictors. In cases where there was no clear choice for which variable to eliminate, we prioritised retaining variables with the clearest direct ecological link to plankton dynamics. The following variables were excluded from all models for collinearity issues, with correlation coefficients (ρ) provided in parentheses: latitude, for strong negative correlation with temperature ($\rho = -0.9$); salinity for strong negative correlation with nitrate ($\rho = -0.9$) and chlorophyll ($\rho = -0.7$); light attenuation for strong positive correlation with chlorophyll ($\rho = 0.8$); primary productivity for strong positive correlation with chlorophyll ($\rho = 0.9$); oxygen for moderate negative correlation with temperature ($\rho = -0.7$); and phosphate for moderate positive correlation with nitrate ($\rho = 0.7$). Correlation among the remaining predictors did not exceed $\rho = 0.6$.

2.5.5. *Data partitioning for model training and testing*

Lifeform and environmental variable time-series for each assessment unit were then split into separate training and testing sets. For each time-series, a discrete period representing 30 % of the temporal range, and randomly positioned along the time-series, was extracted to serve as a testing, or hold-out, set to validate model predictive accuracy on unseen data. We used a discrete period of consecutive months for the testing data instead of randomly sampling months across each time-series. This approach was chosen because the data represent the trend component of each time-series, making adjacent samples temporally autocorrelated and non-independent. For each grid cell, testing data covered 100 consecutive months, equivalent to 8.3 years. The remaining 70 % of each time-series was used to train the models. Missing values in the environmental dataset were imputed separately for the training and testing datasets using the R package 'missRanger' (Mayer and Mayer, 2019). This approach predicts missing values by generating a separate random forest model to predict each environmental variable, using the remaining environmental variables as predictors.

2.5.6. *Model evaluation*

The predictive accuracy of each model was assessed in terms of the R^2 and root mean squared error (RMSE) of predictions on the out-of-bag (OOB) data, and on the testing data. OOB data represents the subset of training data that was not used in the construction of each individual tree, while models were completely naive to the testing data. Large differences in the R^2 of observed versus predicted values between OOB and testing data would indicate the variables we examined were poor predictors of current trends in lifeform abundance. The direction and consistency of long-term trends in lifeform abundance and environmental predictors were assessed across the study area using the non-parametric Kendall trend test from the R package 'EnvStats' (Millard, 2013) on the de-seasonalised mean values (Bedford et al., 2020b; Holland et al., 2023c).

2.5.7. *Relative variable importance ranking*

To interrogate the ability of predictor variables to predict the abundance of lifeforms, variable importance scores were ranked according to their relative position among the other predictors to generate a relative variable abundance ranking which was comparable across models, with 1 indicating the predictor variable with highest importance, and 0.05 indicating the lowest importance for a model with 20 predictors. These were then displayed as pairwise comparisons on a heatmap to identify patterns in relationship between predictor and response variables across multiple models. This approach was adapted from the methodology used to assess links between trends in pelagic habitats indicators and environmental variables for the OSPAR Pelagic Habitats Thematic Assessment 2023 (Holland et al., 2023a, 2023b).

2.5.8. *Spatial pattern analysis and partial dependency*

To assess spatial patterns in environmental variables and in the distribution of lifeforms, we calculated mean values per grid cell for the de-seasonalised environmental predictors and for the six lifeforms. We also compared the amount of variability attributed to spatial and temporal factors in each predictor. To assess temporal variability, we calculated the standard deviation across the time-series for each grid cell. For spatial variability, we computed the standard deviation across the grid for each month. We then compared the distribution of standard deviation values using one-way *t*-tests.

Finally, to investigate relationships between lifeform abundance and predictor variables in each model, partial dependency plots were generated for the five lifeform abundance predictors, and for brevity the environmental predictors with the top five greatest variable importance scores. Partial dependency plots (PDPs) illustrate how lifeform abundance changes across the range of a single predictor variable, while keeping other predictors constant (i.e., at their mean). PDPs can be useful for identifying trends, peaks, or valleys in linear and non-linear relationships. However, PDPs assume independence between predictors and therefore cannot account for interactions. PDPs also represent average relationships across the full training dataset, so they may not accurately reflect individual observations.

3. Results

3.1. *Spatial distribution of lifeform abundance*

Spatial patterns in lifeform distribution indicated that diatom and dinoflagellate abundance were concentrated along the southern coast of the North Sea, with hotspots of very high abundance off the Flemish coast for diatoms and in the eastern North Sea for dinoflagellates (Fig. 3). Large and small copepods exhibited distinct distributions, with small copepods concentrated in a pattern similar to diatoms, while large copepods were found in the offshore shelf areas of the northern North Sea and Celtic Sea. Meroplankton were most abundant across the North Sea and Celtic Sea, reflecting the location of the continental shelf, while fish larvae/eggs were primarily concentrated along the southern coast of the North Sea and along the east coast of England and Scotland.

3.2. *Relative importance of predictor variables*

Our random forest models captured between 94 % and 96 % of the variability (OOB R^2) in lifeform abundances, indicating that the selected predictor variables captured substantial variation in the abundance of each. Bathymetry and longitude were consistently ranked as the environmental variables with high importance across most models (Fig. 4, Fig. 5). Temperature, mixed layer depth, chlorophyll, nitrate, and the nitrogen-to-phosphorus ratio (Nitrogen:Phosphorus) were ranked consistently high across models, while current and wind velocity, cloud cover and climate oscillations (i.e., AMO, NAO) ranked low.

See Supplementary results for descriptive spatial and temporal patterns in lifeform abundance and environmental predictors, long-term trends in lifeform abundance and environmental predictors, accuracy of the spatial interpolation, and the comparison of spatial versus temporal variability.

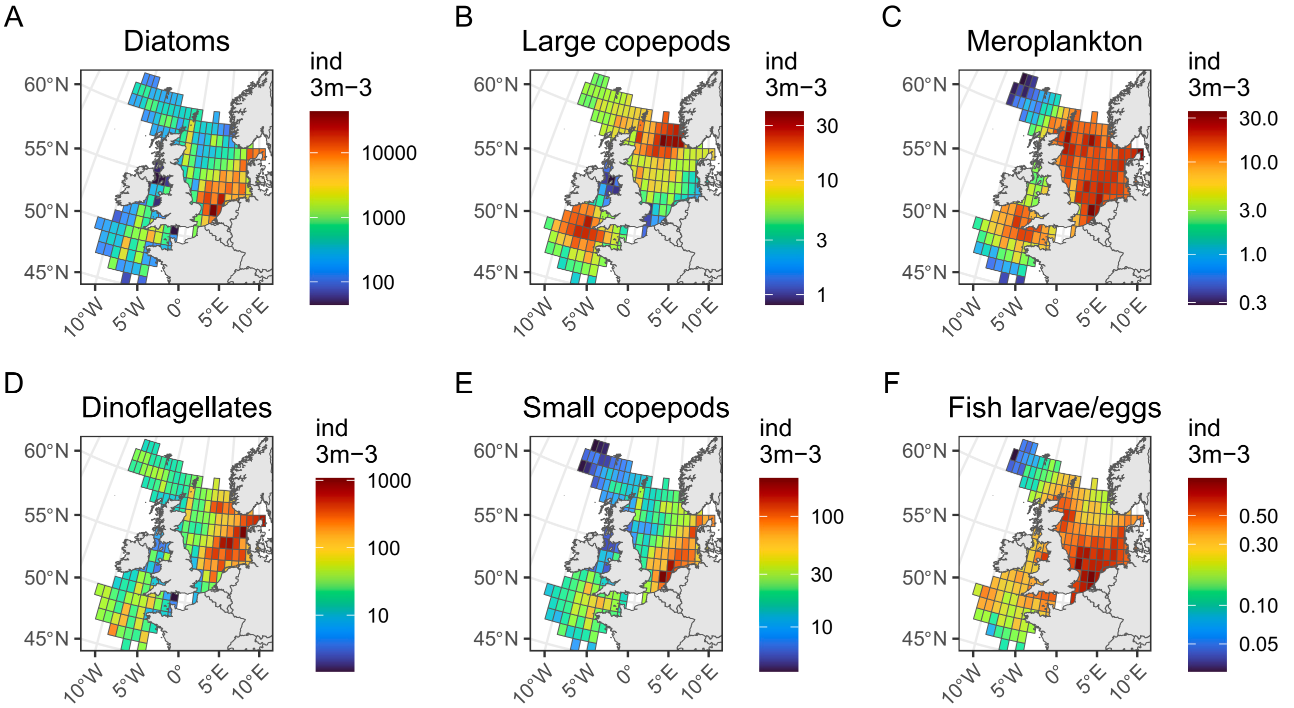


Fig. 3. The mean distribution of abundance values for the six plankton lifeforms across the time-series of the environmental data (1993–2021).

3.3. Different lifeforms have distinct relationships with predictors

Partial dependency plots for environmental predictors indicate that lifeforms have unique responses to environmental changes, although there are some shared responses to individual predictors and positive correlations with one another (Fig. 6, Fig. 7). For instance, the abundance of large copepods declined in shallower areas, whereas meroplankton and fish larvae/eggs increased (Fig. 6), yet their abundances were all positively related to one another (Fig. 7). Small copepods responded positively to temperature, whereas meroplankton abundance declined above 12.5 °C. Diatoms and fish larvae/eggs both had negative relationships with pH, while meroplankton and fish larvae/eggs were negatively related to mixed layer depth. Diatoms and fish larvae/eggs both had positive relationships with the nitrogen-to-phosphorus ratio, whereas small and large copepods had contrasting relationships with nitrate— positive and negative, respectively.

The high OOB R^2 scores across all models and the consistently high variable importance attributed to the abundance of other lifeforms suggest that these independent lifeforms co-vary in predictable patterns relative to one another, despite sharing no common taxa. Small copepods, diatoms, dinoflagellates and meroplankton demonstrated strong positive correlations with one another, suggesting that the abundance of these groups is affected by similar drivers and/or the plankton assemblage is dominated by bottom-up food web processes. Similarly, fish larvae/eggs and meroplankton also appear closely linked, with high relative variable importance and strong positive partial dependency relationships for corresponding predictors in the two models.

3.4. Predicting change in lifeforms

The random forest models performed well at predicting lifeform abundance for OOB values across the range of 234 months or 19.5 years of data used for training the models. Small differences in the R^2 of observed versus predicted data between the models' fitted data (OOB data) and data the models were naïve to (test data) can indicate whether the examined variables were suitable predictors of trends in lifeform abundance (i.e., OOB R2 – Test R2). The anticipated difference between OOB and Test R2 values ranged from 0.16 for meroplankton to 0.36 for fish larvae/eggs (Fig. 4). Despite this large difference in predictive ability, meroplankton and fish larvae/eggs were relatively good predictors of one another. The high test R^2 for meroplankton indicates that relationships present in the training data remained relatively consistent across the data used for the training period. For fish larvae/eggs the high variability in the response was poorly captured by the model.

Across the training time-series for the two case-study assessment units, models were able to predict variation at near-annual scale, as evidenced by how closely predicted values tracked the general trajectory of the observed training data (e.g., large copepods and small copepods in CWM and NNS), but were less well-suited to capturing month-to-month variation present in more dynamic time-series (e.g., dinoflagellates and fish larvae/eggs in NNS) (Fig. 8). Across all models, the accuracy of predictions on training data ranged from 0.89 to 0.95 for CWM, and from 0.89 to 0.93 for NNS. For predictions on testing data, these values ranged from 0.001 to 0.56 for CWM, and from 0.001 to 0.45 for NNS. A few very low test R^2 scores (e.g., 0.001) indicate that for some grid cells lifeform abundance did not always behave predictably, however, the reasonably high test R^2 scores for predictions across the full spatial domain of each model (see Fig. 4) suggest that predictability was only poor for a minority of grid cells (Fig. S 3). For most time-series, predictions on testing data followed the net increasing or decreasing trend in lifeform abundance across the duration of testing data. This observation suggests that the random forest regression models performed better at predicting long-term abundance trends versus predicting month-to-month variation, even for time-series that had seasonality removed. This was particularly evident for diatoms and dinoflagellates at CWM and for dinoflagellates, large copepods and meroplankton at NNS.

Comparing R^2 and RMSE values, there was no clear difference in predictive ability between the two assessment units. There were also no obvious spatial patterns in predictability across the domain of the models (Fig. S 4), however, for some lifeforms test prediction accuracy was marginally greater in the centre of the North Sea and Celtic Seas, where survey coverage was also greatest (see Fig. 1). These results illustrate the important differences in predictability on data used to train

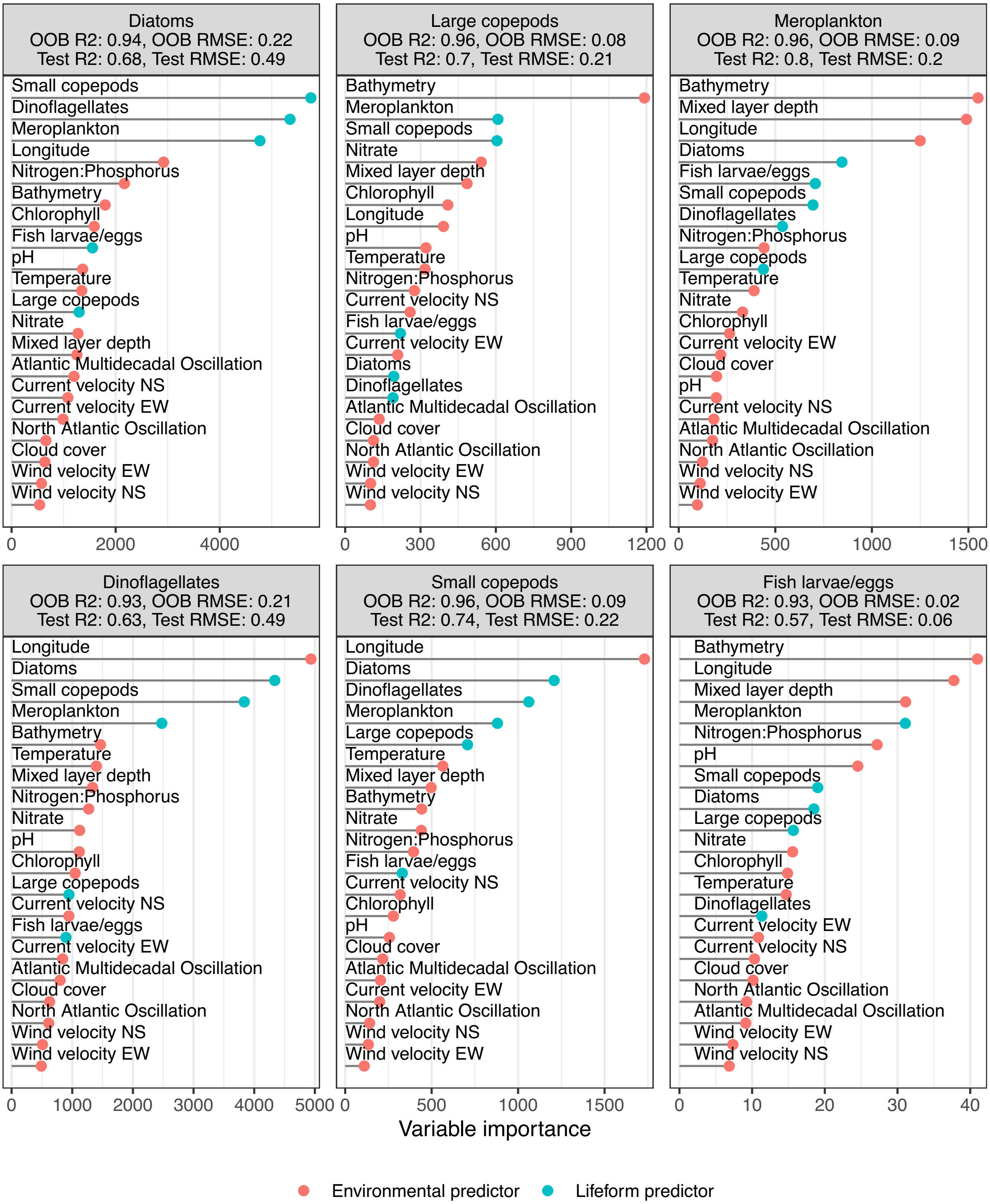


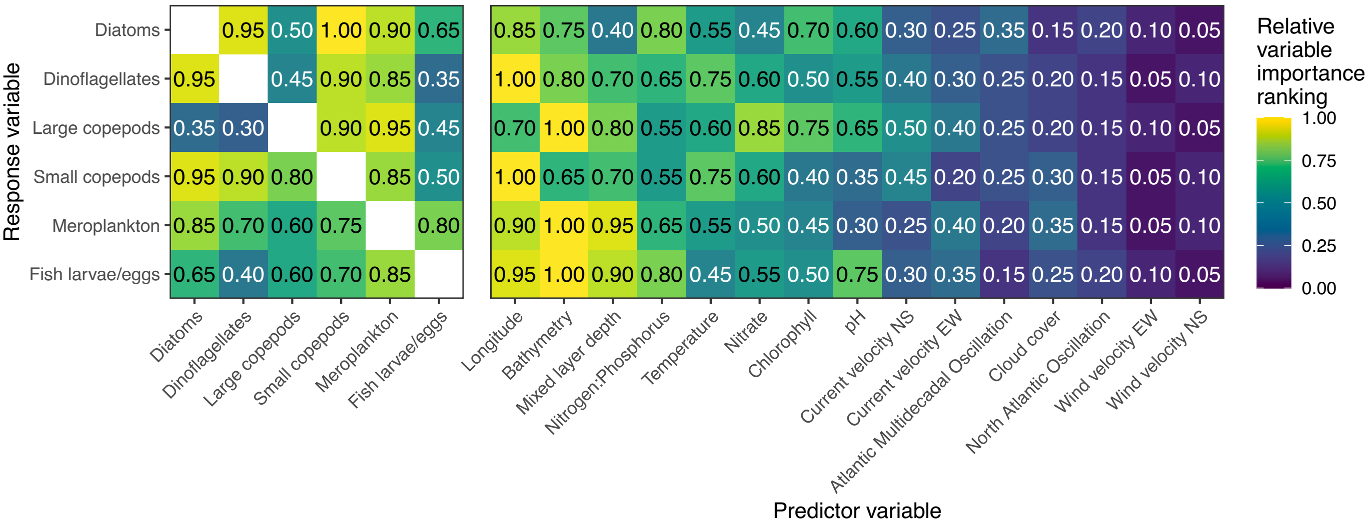
Fig. 4. Mean variable importance for the six plankton lifeforms at the regional scale, as determined by the impurity in the variance of the response. Out of Bag (OOB) R^2 (R squared) and RMSE (Root mean squared error) values for each random forest model are indicated above each facet title. OOB R^2 indicates the mean prediction accuracy on each training sample, using only the trees that did not have this sample in their bootstrap sample. OOB RMSE indicates the root mean square prediction error of the same data. Predictor variables are coloured according to whether they represent the abundance of other lifeforms (Lifeform predictor) or environmental data (Environmental predictor). Descriptions of each environmental predictor can be found in Table S 2.

the models versus unseen test data, and the high degree of variability in lifeform time-series among lifeforms and across the study region.

In summary, our results show that changes in the abundance of plankton lifeforms are largely positively related to the abundance of other plankton lifeforms. However, distinct and even contrasting relationships of lifeforms with environmental change gradients, including those associated with climate change and nutrient pollution, were evident. We also show that the abundance of plankton lifeforms can be predicted using our models.

4. Discussion

This study found that bathymetry, longitude, temperature, mixed layer depth, chlorophyll, and nutrient ratios were all important



Response variable	Diatoms	Dinoflagellates	Large copepods	Small copepods	Meroplankton	Fish larvae/eggs	Longitude	Bathymetry	Mixed layer depth	Nitrogen:Phosphorus	Temperature	Nitrate	Chlorophyll	pH	Current velocity NS	Current velocity EW	Atlantic Multidecadal Oscillation	Cloud cover	North Atlantic Oscillation	Wind velocity EW	Wind velocity NS
Diatoms		0.95	0.50	1.00	0.90	0.65	0.85	0.75	0.40	0.80	0.55	0.45	0.70	0.60	0.30	0.25	0.35	0.15	0.20	0.10	0.05
Dinoflagellates	0.95		0.45	0.90	0.85	0.35	1.00	0.80	0.70	0.65	0.75	0.60	0.50	0.55	0.40	0.30	0.25	0.20	0.15	0.05	0.10
Large copepods	0.35	0.30		0.90	0.95	0.45	0.70	1.00	0.80	0.55	0.60	0.85	0.75	0.65	0.50	0.40	0.25	0.20	0.15	0.10	0.05
Small copepods	0.95	0.90	0.80		0.85	0.50	1.00	0.65	0.70	0.55	0.75	0.60	0.40	0.35	0.45	0.20	0.25	0.30	0.15	0.05	0.10
Meroplankton	0.85	0.70	0.60	0.75		0.80	0.90	1.00	0.95	0.65	0.55	0.50	0.45	0.30	0.25	0.40	0.20	0.35	0.15	0.05	0.10
Fish larvae/eggs	0.65	0.40	0.60	0.70	0.85		0.95	1.00	0.90	0.80	0.45	0.55	0.50	0.75	0.30	0.35	0.15	0.25	0.20	0.10	0.05

Fig. 5. Heatmap of pairwise relative variable importance ranking for all predictors from 1 (highest relative importance) to 0.05 (lowest relative importance). Environmental predictor variables are ordered in terms of mean relative variable importance.

predictors of plankton lifeform abundance. However, lifeforms exhibited a range of responses to these predictors, highlighting the complexity of interactions between physical and chemical ocean conditions and plankton dynamics. Plankton lifeforms also showed predictable covariation in abundance, with strong positive correlations among certain groups, such as small copepods, diatoms, and dinoflagellates, as well as meroplankton and fish larvae/eggs. This suggests that shared environmental drivers influence variations in their abundance. This covariation underscores the intricate interconnectedness of pelagic habitats and the importance of considering multiple lifeform groups in biodiversity assessments. The physical, chemical, and ecological variables we examined were effective predictors of plankton lifeform abundance across space and time, indicating that trends in these variables align with trends in lifeform abundance documented in Holland et al. (2023c) and Bedford et al. (2020b).

4.1. *Changes in pelagic habitat conditions*

The changes in plankton lifeform abundance for the North-East Atlantic presented here, which have been found to be associated with climate change and nutrient pollution, are indicative of widespread deteriorating pelagic habitat conditions associated with anthropogenic pressures (Dailianis et al., 2018). Our models provide information on where along gradients of environmental change we can anticipate change in lifeform abundance (Fig. 6), and thereby help improve understanding of where and how much anthropogenic pressures need to be reduced to achieve environmental ambitions set by OSPAR, UKMS and MSFD. For example, we can anticipate changes in the community composition and size distribution of plankton, primarily driven by dinoflagellates, small copepods and meroplankton, as mean annual temperatures increase from 11 to 13 °C (Fig. 6). Additionally, both direct and indirect effects (i.e., through food webs) of changes in nutrient concentrations, especially where riverine inputs are high, are expected to influence these dynamics (Fig. S 5).

The region covered by this study is experiencing an increase in thermal stratification intensity and a prolonged duration of the summer stratified period due to climate change (Bopp et al., 2013), driving decreases in phytoplankton productivity (Edwards et al., 2022; Schmidt et al., 2020; Steinacher et al., 2010). Although the mixed layer depth is not shoaling uniformly across the study region (Somavilla et al., 2017), it is likely that increases in stratification intensity are impacting the abundance of plankton lifeforms, with additional impacts caused by rising water temperatures (Winder and Sommer, 2012), driving poleward shifts in the distribution of plankton taxa (Barton et al., 2016).

The North Sea and Celtic Seas have experienced increased phosphate limitation for coastal ecosystems (Breton et al., 2022) as a result of nutrient management efforts commencing in the 1980s (Lenhart and Große, 2018). This has led to a strong offshore gradient from phosphorus to nitrogen limitation in the North Sea, which may lead to changes in the growth, species composition and nutritional quality of the phytoplankton community (Burson et al., 2016). Our analysis revealed that the ratio of nitrogen-to-phosphorus, and to a lesser extent nitrate alone, were important predictors of plankton lifeform abundance. It is possible that imbalanced nutrient ratios are influencing the abundance of these lifeforms, through direct effects on nutrients available to diatoms and indirect effects on meroplankton and fish larvae/eggs (i.e., through changes in the supply of phytoplankton). In support of our findings, a recent study using non-CPR data concluded that diatom communities in the eastern Channel are being impacted by nutrient reductions and climate change-driven warming (Breton et al., 2022). Given this body of evidence and studies which show that aquatic trophic transfer efficiency can be impaired by warming-induced nutrient stress (Atkinson et al., 2024; Barneche et al., 2021; Capuzzo et al., 2018; Du Pontavice et al., 2020), future work could attempt to quantify how changes in nutrient concentrations may interact with climate change.

The importance of longitude and bathymetry as primary drivers of variation in most lifeforms is challenging to interpret. Both variables exhibit spatial variation but remain constant over time, indicating that their correlation with lifeform abundance is purely spatial. In the region we assessed, both longitude and bathymetry follow a gradient from off-shelf to on-shelf conditions, coinciding with gradients in nutrient supply, tidal mixing, salinity (from freshwater inflow) and benthic-pelagic coupling (Emeis et al., 2015). The distribution of meroplankton and fish larvae/eggs is strongly influenced by proximity to on-shelf spawning adult populations (Highfield et al., 2010). Meroplankton, in particular, arise from the connectivity between benthic and pelagic habitats, which tends to align with bathymetric gradients (Hidalgo et al., 2014). Gradients in longitude and bathymetry also intersect with the European shelf-edge current and the Flamborough Front, which act as barriers to phyto- and zooplankton by controlling temperature and advection, thereby influencing biodiversity patterns (Kléparski et al., 2021).

4.2. *Plankton lifeforms covary with each other*

Variable importance and partial dependency relationships suggest that diatoms, dinoflagellates, and small copepods, as well as

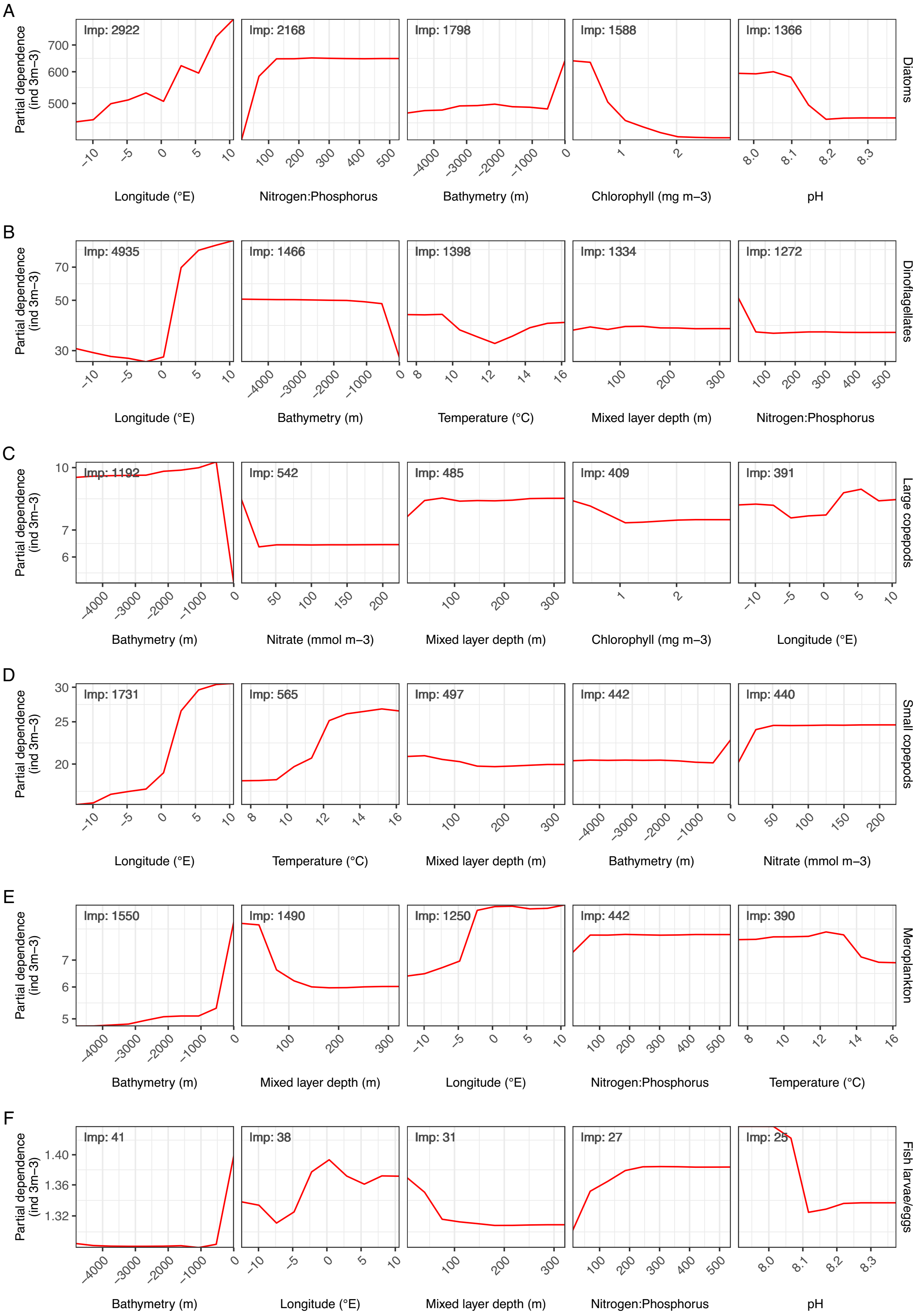


Fig. 6. Partial dependency plots for the six plankton lifeforms, indicated on the right of each row of plots, display the top five highest-ranked environmental predictor variables, ordered by variable importance score (displayed as “Imp” inside each subplot).

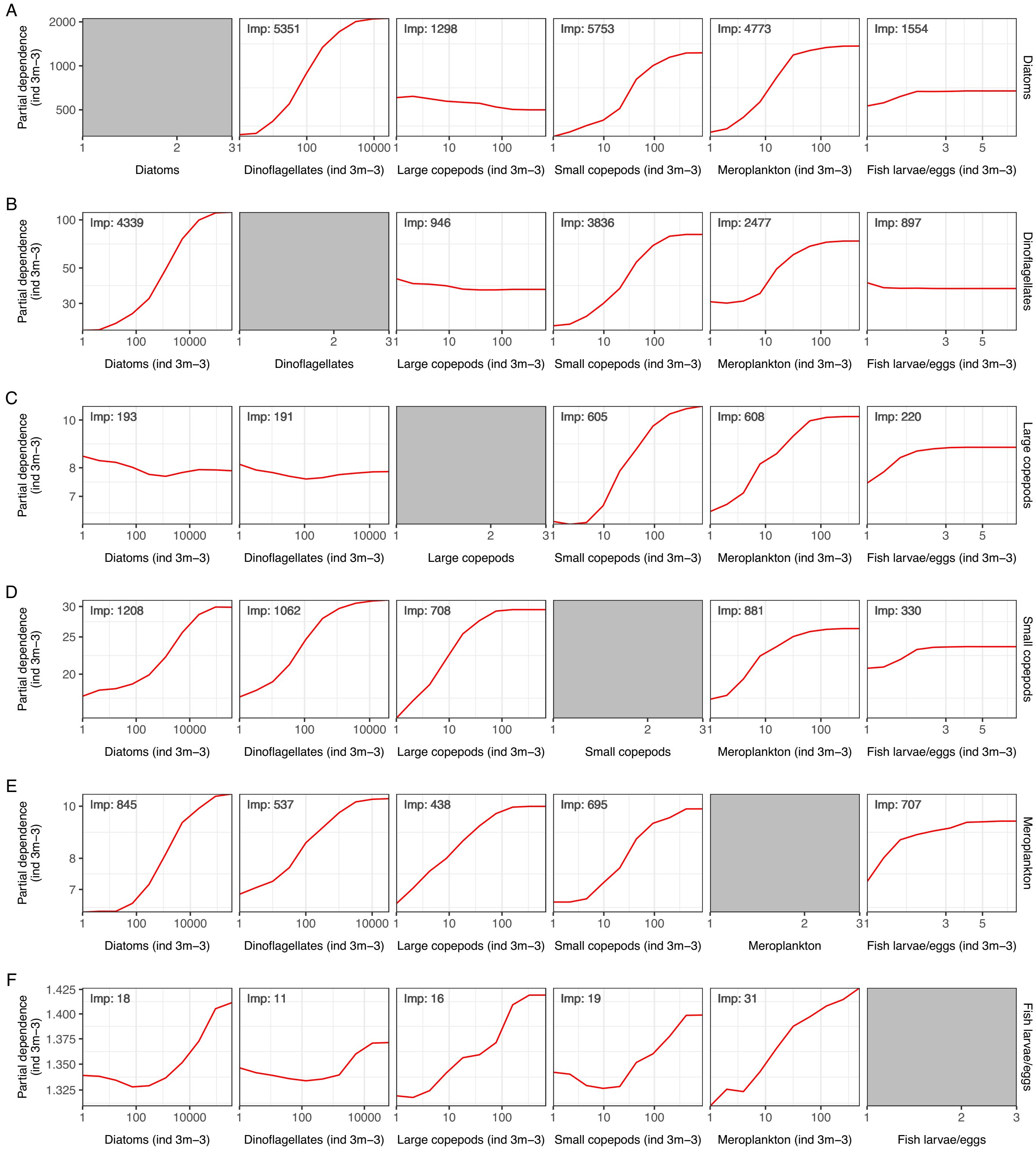
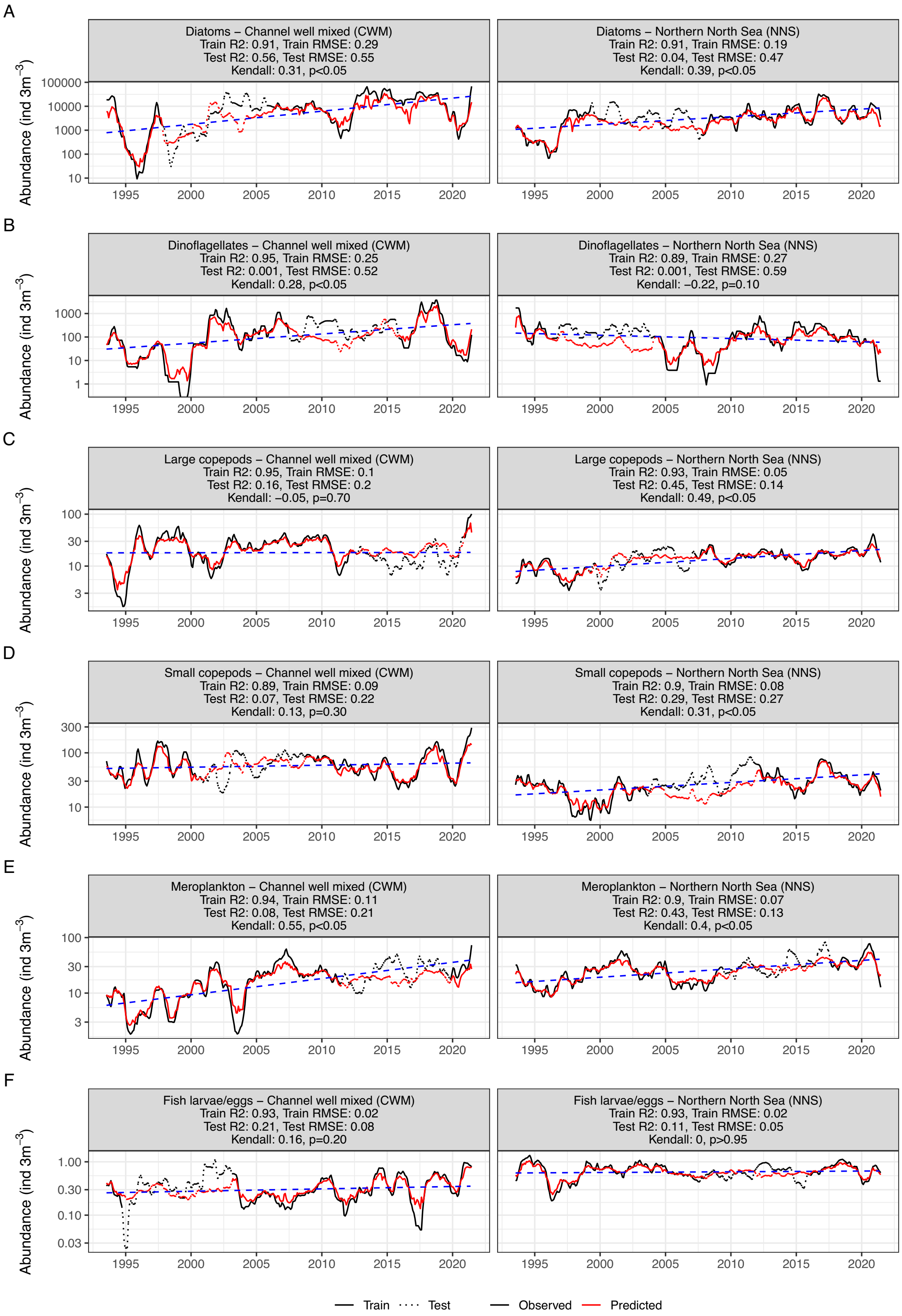


Fig. 7. Partial dependency plots for the six plankton lifeforms, indicated on the right of each row of plots, displaying the five lifeform abundance predictor variables. “Imp” indicates the variable importance score for the corresponding predictor.

meroplankton and fish larvae/eggs, are two ensembles of lifeforms which co-vary in a very similar manner across space and time. Despite their overlap in spatial distribution (Fig. 3), the analysis of abundance trend direction indicated that diatoms have largely increased in abundance while dinoflagellates have decreased across much of the region (see Fig. S 6). Positive correlation among lifeforms suggests that they respond in similar ways to environmental variation, or that there is some dependency among them. Copepods and meroplankton have similar trophic levels and thus likely compete for resources. In our analysis, meroplankton were dominated by echinoderm larvae, while large and small copepods were dominated by *Calanus* and *Acartia* spp., respectively. All three groups primarily feed on phytoplankton, although many, including *Acartia* spp. (Kiørboe et al., 1996) and *C. finmarchicus* (Nejstgaard et al., 1997), also consume microzooplankton. Thorpe et al.



A

Diatoms – Channel well mixed (CWM)
Train R2: 0.91, Train RMSE: 0.29
Test R2: 0.56, Test RMSE: 0.55
Kendall: 0.31, p<0.05

Diatoms – Northern North Sea (NNS)
Train R2: 0.91, Train RMSE: 0.19
Test R2: 0.04, Test RMSE: 0.47
Kendall: 0.39, p<0.05

B

Dinoflagellates – Channel well mixed (CWM)
Train R2: 0.95, Train RMSE: 0.25
Test R2: 0.001, Test RMSE: 0.52
Kendall: 0.28, p<0.05

Dinoflagellates – Northern North Sea (NNS)
Train R2: 0.89, Train RMSE: 0.27
Test R2: 0.001, Test RMSE: 0.59
Kendall: –0.22, p=0.10

C

Large copepods – Channel well mixed (CWM)
Train R2: 0.95, Train RMSE: 0.1
Test R2: 0.16, Test RMSE: 0.2
Kendall: –0.05, p=0.70

Large copepods – Northern North Sea (NNS)
Train R2: 0.93, Train RMSE: 0.05
Test R2: 0.45, Test RMSE: 0.14
Kendall: 0.49, p<0.05

D

Small copepods – Channel well mixed (CWM)
Train R2: 0.89, Train RMSE: 0.09
Test R2: 0.07, Test RMSE: 0.22
Kendall: 0.13, p=0.30

Small copepods – Northern North Sea (NNS)
Train R2: 0.9, Train RMSE: 0.08
Test R2: 0.29, Test RMSE: 0.27
Kendall: 0.31, p<0.05

E

Meroplankton – Channel well mixed (CWM)
Train R2: 0.94, Train RMSE: 0.11
Test R2: 0.08, Test RMSE: 0.21
Kendall: 0.55, p<0.05

Meroplankton – Northern North Sea (NNS)
Train R2: 0.9, Train RMSE: 0.07
Test R2: 0.43, Test RMSE: 0.13
Kendall: 0.4, p<0.05

F

Fish larvae/eggs – Channel well mixed (CWM)
Train R2: 0.93, Train RMSE: 0.02
Test R2: 0.21, Test RMSE: 0.08
Kendall: 0.16, p=0.20

Fish larvae/eggs – Northern North Sea (NNS)
Train R2: 0.93, Train RMSE: 0.02
Test R2: 0.11, Test RMSE: 0.05
Kendall: 0, p>0.95

Fig. 8. Observed (black line) versus predicted (red line) values for two grid cells of interest: Channel Well Mixed (CWM), and Northern North Sea (NNS) for the six lifeform abundance models. The solid (black) line component of each time-series indicates data used to train the model (i.e., training data), while the dotted (black) component of each time-series indicates values left out of model fitting (i.e., testing data). Linear trendlines (blue line) were calculated from observed data. Train R^2 indicates the accuracy of model prediction on the training data, while Test R^2 indicates the same statistic generated from prediction on testing data for the assessment units (CWM or NNS). Train and Test RMSE indicate root mean squared error of predictions on the training and testing data for the two assessment units. Kendall trend test coefficient and *p*-value are also indicated.

(2022) concluded that high abundances of small copepods in the North Sea were influenced by increased chlorophyll-*a* concentration through bottom-up control. It is probable that variation in the abundance of phytoplankton lifeforms, through warming and changes in nutrient concentration (Di Pane et al., 2022), may influence the abundance of small copepods and meroplankton consumers (Marques et al., 2024). The key take home message from our correlations between lifeforms is that the abundance of these groups in the plankton assemblage appears dominated by bottom-up food web processes, with no compelling evidence of top-down control from zooplankton on phytoplankton lifeforms, for instance.

It is important to note that our analysis did not include key carnivores other than fish larvae and carnivorous copepods; carnivores such as chaetognaths, siphonophores, ctenophores, and cnidarians did not fall into the lifeforms examined here. Consequently, top-down control within the plankton community cannot be justifiably ruled out. These groups, particularly gelatinous zooplankton, are poorly represented in the CPR Survey due to their non-quantitative capture (Baxter et al., 2010), which prevented their inclusion in our study. There has been a suggested increase in the dominance of carnivorous zooplankton globally, often at the expense of copepods (Gibbons and Richardson, 2009; Heneghan et al., 2021). This underscores the need for better representation of these organisms in monitoring programs to enhance biodiversity assessments of pelagic habitats (Long et al., 2021).

4.3. Predicting lifeform abundance

In many cases, the direction of change in the abundance of plankton lifeforms has been consistent in the long-term, with most lifeforms in this region demonstrating net decreasing trends, except in the case of diatoms, meroplankton and fish larvae/eggs (Bedford et al., 2020b; Holland et al., 2023c). While 29 years (1993 to 2021) is not a particularly long duration in the context of environmental change, it represents a very long time in terms of plankton lifecycles. The duration of the analysis is a core strength of this study, reinforcing the observation that pelagic habitats in the North-East Atlantic are changing (Holland et al., 2023c).

Fish larvae/eggs demonstrated the greatest decrease in model predictive accuracy between the training and testing data, indicating that trends in their abundance were particularly difficult to predict, relative to the other lifeforms. This low predictability was likely influenced by the highly patchy distribution of fish larvae/eggs in space and time (Jansen et al., 2012). This observation might also relate to the fact that nursery areas are mainly located in coastal environments where the CPR does not sample (Bils et al., 2012; van der Veer et al., 2024). It is possible that there were periods in time when large changes occurred across multiple grid cells, making lifeform abundance during that period particularly difficult to predict. Thus, random assignment of testing period data in each time-series undoubtedly influenced the accuracy of individual predictions (i.e., on each grid cell), however this likely balances out when all grid cells in the study region are considered. It is also important to note that this analysis was largely based on modelled environmental data, and in some cases remotely sensed variables. The uncertainty arising from discrepancies between hydrodynamic and biogeochemical model outputs and the validation data used to train these models must also be considered in the results.

4.4. Future work

Consistently high variable importance and strong monotonic relationships with longitude in most models provide evidence that some of the spatial variability in lifeform abundance was linked to geographical and/or oceanographic characteristics that were not explored in this study. The range of longitude we examined corresponds with environmental variables known to more directly impact plankton (e.g., salinity, nutrient supply, chlorophyll, oceanic fronts) (Emeis et al., 2015), but could also be an important proxy for biological processed such as dispersal and habitat connectivity (Ayata et al., 2010; Bendtsen et al., 2023). Similarly, bathymetry correlates with a host of possible drivers of plankton assemblage structure such as light levels, oxygen concentration and resource supply. Future work could attempt to attribute this variability to direct drivers to establish whether important temporal changes are currently going undetected. It will also be important to explore how relationships between lifeform abundance and environmental variables change at different spatial scales (Graves et al., 2023; Scott et al., 2023) and to dig deeper into potential links with climate change (Bedford et al., 2020b; Holland et al., 2023c).

Our analysis focused solely on trends in plankton lifeform abundance. Future work should apply our methodology to study trends in biomass, as this would better represent energy flow in pelagic food webs. Efforts are underway to facilitate such studies by converting abundance data into estimated biomasses. It is important to note that while abundance is directly measured, biomass is secondarily estimated.

Finally, future work should attempt to incorporate seasonality, since several authors have now shown that trends in plankton are driven by summer events when warmer sea surface temperature drives particularly high demands from nutrient-limited food to support growth (Capuzzo et al., 2018; Edwards et al., 2022; Schmidt et al., 2020).

4.5. Policy implications and final conclusions

Basing this study on plankton lifeforms has allowed us to directly link variations in the marine environment with biodiversity indicators that are important for regional marine policy, due to the fundamental role of plankton at the base of marine ecosystems (McQuatters-Gollop et al., 2019a). Focusing on groups of taxa that perform similar ecosystem functions (i.e., lifeforms) rather than on individual taxa, has facilitated an understanding of broad ecological processes and increased comparability with other work using plankton lifeforms (Bedford et al., 2020b; Holland et al., 2023a; Holland et al., 2023c; McQuatters-Gollop et al., 2019a). Considering the vast diversity of plankton species, grouping them by functional role also supports more manageable assessments of ecosystem health and intercomparison between different marine regions, which is crucial for broad-scale policies like OSPAR and MSFD. Our work highlights the critical need to address climate change and nutrient pollution to meet international ambitions set by OSPAR, the EU MSFD and the UKMS for achieving Good Environmental Status and restoring biodiversity where it has been degraded.

By analysing nearly three decades of data from the CPR Survey, this study provides robust evidence of how environmental variability impacts the plankton community. This long-term perspective is essential for understanding the effects of climate change and nutrient pollution on marine biodiversity. Our findings emphasise the need for comprehensive monitoring of plankton and environmental variables to gain a better understanding of the state of pelagic habitats and to support effective marine conservation and management. Identifying the specific drivers of change in the plankton community is essential for addressing broader ecosystem health and achieving Good Environmental Status as mandated by international frameworks.

CRediT authorship contribution statement

Matthew M. Holland: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Angus Atkinson:** Writing – review & editing, Methodology, Investigation. **Mike Best:** Writing – review & editing, Resources, Project administration, Investigation, Funding acquisition. **Eileen Bresnan:** Writing – review & editing. **Michelle Devlin:** Writing – review & editing. **Eric Goberville:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Pierre Hélaouët:** Writing – review & editing, Methodology, Investigation. **Margarita Machairopoulou:** Writing – review & editing.

Matthew Faith: Writing – review & editing, Methodology. **Murray S.A. Thompson:** Writing – review & editing, Writing – original draft, Methodology, Investigation, Conceptualization. **Abigail McQuatters-Gollop:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to acknowledge the contributions from our networks of pelagic habitats experts who have enabled and progressed this work, to the OSPAR ICG-COBAM and BDC members and convenors, and to the OSPAR secretariat. We are grateful for the funding received from the UK Department for Environment, Food & Rural Affairs (Defra) through the marine arm of their Natural Capital and Ecosystem Assessment (NCEA) programme (NC34 Pelagic program-“PelCap”) and from the European maritime and Fisheries Fund for the project: “North-east Atlantic project on biodiversity and eutrophication assessment integration and creation of effective measures (NEA PANACEA)”, financed by the European Union's DG ENV/MSFD 2020, under agreement No. 110661/2020/839628/SUB/ENV.C.2. The marine NCEA programme is leading the way in supporting Government ambition to integrate natural capital approaches into decision making for the marine environment. Find out more at https://www.gov.uk/government/publications/natural-capital-and-ecosystem-assessment-programme. AMG would like to thank the United Kingdom National Environmental Research Council (UKRI-NERC) for support through the NERC Knowledge Exchange Fellowship Scheme (NE/R002738/1); MSAT was supported via the Natural Environment Research Council grant NE/V017039/1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.175793.

References

Alvarez-Fernandez, S., Bach, L.T., Taucher, J., Riebesell, U., Sommer, U., Aberle, N., Brussaard, C., Boersma, M., 2018. Plankton responses to ocean acidification: the role of nutrient limitation. Prog. Oceanogr. 165, 11–18.

Arrigo, K.R., 2005. Marine microorganisms and global nutrient cycles. Nature 437 (7057), 349–355.

Atkinson, A., Rossberg, A.G., Gaedke, U., Sprules, G., Heneghan, R.F., Batziakas, S., Grigoratou, M., Fileman, E., Schmidt, K., Frangoulis, C., 2024. Steeper size spectra with decreasing phytoplankton biomass indicate strong trophic amplification and future fish declines. Nat. Commun. 15 (1), 381.

Ayata, S.-D., Lazure, P., Thiébaut, É., 2010. How does the connectivity between populations mediate range limits of marine invertebrates? A case study of larval dispersal between the Bay of Biscay and the English Channel (North-East Atlantic). Prog. Oceanogr. 87 (1–4), 18–36.

Barneche, D.R., Hulatt, C.J., Dossena, M., Padfield, D., Woodward, G., Trimmer, M., Yvon-Durocher, G., 2021. Warming impairs trophic transfer efficiency in a long-term field experiment. Nature 592 (7852), 76–79.

Barton, A.D., Irwin, A.J., Finkel, Z.V., Stock, C.A., 2016. Anthropogenic climate change drives shift and shuffle in North Atlantic phytoplankton communities. Proc. Natl. Acad. Sci. 113 (11), 2964–2969.

Batten, S.D., Raitsos, D.E., Danielson, S., Hopcroft, R., Coyle, K., McQuatters-Gollop, A., 2018. Interannual variability in lower trophic levels on the Alaskan Shelf. Deep-Sea Res. II Top. Stud. Oceanogr. 147, 58–68.

Baxter, E., Walne, A., Purcell, J., McAllen, R., Doyle, T., 2010. Identification of Jellyfish from Continuous Plankton Recorder Samples. Jellyfish Blooms: New Problems and Solutions, pp. 193–201.

Beaugrand, G., Reid, P.C., 2012. Relationships between North Atlantic salmon, plankton, and hydroclimatic change in the Northeast Atlantic. ICES J. Mar. Sci. 69 (9), 1549–1562.

Bedford, J., Johns, D.G., McQuatters-Gollop, A., 2020a. Implications of taxon-level variation in climate change response for interpreting plankton lifeform biodiversity indicators. ICES J. Mar. Sci. 77 (7–8), 3006–3015.

Bedford, J., Ostle, C., Johns, D.G., Atkinson, A., Best, M., Bresnan, E., Machairopoulou, M., Graves, C.A., Devlin, M., Milligan, A., 2020b. Lifeform indicators reveal large-scale shifts in plankton across the North-West European shelf. Glob. Chang. Biol. 26 (6), 3482–3497.

Bendtsen, J., Sørensen, L.L., Daugbjerg, N., Lundholm, N., Richardson, K., 2023. Phytoplankton diversity explained by connectivity across a mesoscale frontal system in the open ocean. Sci. Rep. 13 (1), 12117.

Bils, F., Kanstinger, P., Kloppmann, M.H., Peck, M.A., 2012. Habitat partitioning by fish larvae among coastal, offshore and frontal zones in the southern North Sea. Aquat. Biol. 15 (3), 237–250.

Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Seferian, R., 2013. Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. Biogeosciences 10 (10), 6225–6245.

Breton, E., Goberville, E., Sautour, B., Ouadi, A., Skouroliakou, D.-I., Seuront, L., Beaugrand, G., Kléparski, L., Crouvoisier, M., Pecqueur, D., 2022. Multiple phytoplankton community responses to environmental change in a temperate coastal system: a trait-based approach. Front. Mar. Sci. 9.

Brose, U., Jonsson, T., Berlow, E.L., Warren, P., Banasek-Richter, C., Bersier, L.-F., Blanchard, J.L., Brey, T., Carpenter, S.R., Blandenier, M.-F.C., 2006. Consumer–resource body-size relationships in natural food webs. Ecology 87 (10), 2411–2417.

Brose, U., Archambault, P., Barnes, A., Bersier, L., Boy, T., Canning-Clode, J., Conti, E., Dias, M., Digel, C., Dissanayake, A., 2019. Predator traits determine food-web architecture across ecosystems. Nat. Ecol. Evol. 3, 919–927.

Brun, P., Kiørboe, T., Licandro, P., Payne, M.R., 2016. The predictive skill of species distribution models for plankton in a changing climate. Glob. Chang. Biol. 22 (9), 3170–3181.

Burson, A., Stomp, M., Akil, L., Brussaard, C.P., Huisman, J., 2016. Unbalanced reduction of nutrient loads has created an offshore gradient from phosphorus to nitrogen limitation in the N orth S ea. Limnol. Oceanogr. 61 (3), 869–888.

Butenschön, M., Clark, J., Aldridge, J.N., Allen, J.I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., 2016. ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. Geosci. Model Dev. 9 (4), 1293–1339.

Capuzzo, E., Lynam, C.P., Barry, J., Stephens, D., Forster, R.M., Greenwood, N., McQuatters-Gollop, A., Silva, T., van Leeuwen, S.M., Engelhard, G.H., 2018. A decline in primary production in the North Sea over 25 years, associated with reductions in zooplankton abundance and fish stock recruitment. Glob. Chang. Biol. 24 (1), e352–e364.

Chivers, W.J., Walne, A.W., Hays, G.C., 2017. Mismatch between marine plankton range movements and the velocity of climate change. Nat. Commun. 8 (1), 1–8.

Colebrook, J., 1982. Continuous plankton records: seasonal variations in the distribution and abundance of plankton in the North Atlantic Ocean and the North Sea. J. Plankton Res. 4 (3), 435–462.

Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88 (11), 2783–2792.

Dailianis, T., Smith, C.J., Papadopoulou, N., Gerovasileiou, V., Sevastou, K., Bekkby, T., Bilan, M., Billett, D., Boström, C., Carreiro-Silva, M., 2018. Human activities and resultant pressures on key European marine habitats: an analysis of mapped resources. Mar. Policy 98, 1–10.

Di Pane, J., Wiltshire, K.H., McLean, M., Boersma, M., Meunier, C.L., 2022. Environmentally induced functional shifts in phytoplankton and their potential consequences for ecosystem functioning. Glob. Chang. Biol. 28 (8), 2804–2819.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B., Leitão, P.J., 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36 (1), 27–46.

Du Pontavice, H., Gascuel, D., Reygondeau, G., Maureaud, A., Cheung, W.W., 2020. Climate change undermines the global functioning of marine food webs. Glob. Chang. Biol. 26 (3), 1306–1318.

Edwards, M., Hélaouët, P., Goberville, E., Lindley, A., Tarling, G.A., Burrows, M.T., Atkinson, A., 2021. North Atlantic warming over six decades drives decreases in krill abundance with no associated range shift. Commun. Biol. 4 (1), 644.

Edwards, M., Beaugrand, G., Kléparski, L., Hélaouët, P., Reid, P.C., 2022. Climate variability and multi-decadal diatom abundance in the Northeast Atlantic. Commun. Earth Environ. 3 (1), 1–8.

Emeis, K.-C., van Beusekom, J., Callies, U., Ebinghaus, R., Kannen, A., Kraus, G., Kröncke, I., Lenhart, H., Lorkowski, I., Matthias, V., 2015. The North Sea—a shelf sea in the Anthropocene. J. Mar. Syst. 141, 18–33.

Endo, C.A.K., Stige, L.C., Skogen, M.D., Ciannelli, L., Vikebø, F.V., 2022. Two decades of match-mismatch in Northeast Arctic cod–feeding conditions and survival. Front. Mar. Sci. 9, 767290.

Enserink, L., Blauw, A., van der Zande, D., Markager, S., 2019. Summary Report of the EU Project ‘Joint Monitoring Programme of the Eutrophication of the North Sea with Satellite data’ (Ref: DG ENV/MSFD Second Cycle/2016).

European Commission, 2017. Commission decision (EU) 2017/848 of 17 May 2017 laying down criteria and methodological standards on good environmental status of marine waters and specifications and standardised methods for monitoring and assessment, and repealing decision 2010/477/EU. Off. J. Eur. Union L 125, 43–74.

Frederiksen, M., Edwards, M., Richardson, A.J., Halliday, N.C., Wanless, S., 2006. From plankton to top predators: bottom-up control of a marine food web across four trophic levels. J. Anim. Ecol. 75 (6), 1259–1268.

Freeman, E., Woodruff, S.D., Worley, S.J., Lubker, S.J., Kent, E.C., Angel, W.E., Berry, D. I., Brohan, P., Eastman, R., Gates, L., 2017. ICOADS release 3.0: a major update to the historical marine climate record. Int. J. Climatol. 37 (5), 2211–2232.

Gibbons, M.J., Richardson, A.J., 2009. Patterns of jellyfish abundance in the North Atlantic. In: Jellyfish Blooms: Causes, Consequences, and Recent Advances: Proceedings of the Second International Jellyfish Blooms Symposium, held at the Gold Coast, Queensland, Australia, 24–27 June, 2007. Springer.

Graves, C., Best, M., Atkinson, A., Bear, B., Bresnan, E., Holland, M., Johns, D., Machairopoulou, M., McQuatters-Gollop, A., Mellor, A., 2023. At what scale should we assess the health of pelagic habitats? Trade-offs between small-scale manageable pressures and the need for regional upscaling. Ecol. Indic. 154, 110571.

Hays, G., Warner, A., 1993. Consistency of towing speed and sampling depth for the continuous plankton recorder. J. Mar. Biol. Assoc. U. K. 73 (4), 967–970.

Heneghan, R., Everett, J., Blanchard, J., Sykes, P., Richardson, A., 2021. Climate-driven Zooplankton Shifts Could Cause Global Declines in Food Quality for Fish.

Hidalgo, M., Reglero, P., Álvarez-Berastegui, D., Torres, A.P., Álvarez, I., Rodriguez, J. M., Carbonell, A., Zaragoza, N., Tor, A., Goñi, R., 2014. Hydrographic and biological components of the seascape structure the meroplankton community in a frontal system. Mar. Ecol. Prog. Ser. 505, 65–80.

Highfield, J.M., Eloire, D., Conway, D.V., Lindeque, P.K., Attrill, M.J., Somerfield, P.J., 2010. Seasonal dynamics of meroplankton assemblages at station L4. J. Plankton Res. 32 (5), 681–691.

Holland, M., Louchart, A., Artigas, L.F., Mcquatters-Gollop, A., 2023a. Changes in phytoplankton and zooplankton communities common Indicator assessment changes in phytoplankton and zooplankton communities. In: OSPAR, 2023: The 2023 Quality Status Report for the Northeast Atlantic.

Holland, M., Louchart, A., Artigas, L.F., Mcquatters-Gollop, A., 2023b. OSPAR CEMP guideline for common indicator: PH1/FW5 change in plankton communities (the 2023 quality status report for the Northeast Atlantic, issue). https://www.ospar.org/documents?v=39001.

Holland, M.M., Louchart, A., Artigas, L.F., Ostle, C., Atkinson, A., Rombouts, I., Graves, C.A., Devlin, M., Heyden, B., Machairopoulou, M., Bresnan, E., Schilder, J., Jakobsen, H.H., Lloyd-Hartley, H., Tett, P., Best, M., Goberville, E., McQuatters-Gollop, A., 2023c. Major declines in NE Atlantic plankton contrast with more stable populations in the rapidly warming North Sea. Sci. Total Environ. 165505 https://doi.org/10.1016/j.scitotenv.2023.165505.

Jansen, T., Kristensen, K., Payne, M., Edwards, M., Schrum, C., Pitois, S., 2012. Long-term retrospective analysis of mackerel spawning in the North Sea: a new time series and modeling approach to CPR data. PloS One 7 (6), e38758.

Kiørboe, T., Saiz, E., Viitasalo, M., 1996. Prey switching behaviour in the planktonic copepod *Acartia tonsa*. Mar. Ecol. Prog. Ser. 143, 65–75.

Kléparski, L., Beaugrand, G., Edwards, M., 2021. Plankton biogeography in the North Atlantic Ocean and its adjacent seas: species assemblages and environmental signatures. Ecol. Evol. 11 (10), 5135–5149.

Kristiansen, T., Drinkwater, K.F., Lough, R.G., Sundby, S., 2011. Recruitment variability in North Atlantic cod and match-mismatch dynamics. PloS One 6 (3), e17456.

Lenhart, H.-J., Große, F., 2018. Assessing the effects of WFD nutrient reductions within an OSPAR frame using trans-boundary nutrient modeling. Front. Mar. Sci. 5, 447.

Lindegren, M., Van Deurs, M., MacKenzie, B.R., Worsoe Clausen, L., Christensen, A., Rindorf, A., 2018. Productivity and recovery of forage fish under climate change and fishing: North Sea sandeel as a case study. Fish. Oceanogr. 27 (3), 212–221.

Llope, M., Chan, K.-S., Ciannelli, L., Reid, P., Stige, L.C., Stenseth, N.C., 2009. Effects of environmental conditions on the seasonal distribution of phytoplankton biomass in the North Sea. Limnol. Oceanogr. 54 (2), 512–524.

Long, A., Haberlin, D., Lyashevska, O., Brophy, D., O'Hea, B., O'Donnell, C., Scarrott, R., Lawton, C., Doyle, T., 2021. Interannual variability of gelatinous mesozooplankton in a temperate shelf sea: greater abundance coincides with cooler sea surface temperatures. ICES J. Mar. Sci. 78 (4), 1372–1385.

Louchart, A., Artigas, L.F., Holland, M., Mcquatters-Gollop, A., 2023a. PH2 change in phytoplankton biomass and zooplankton abundance (the 2023 quality status report for the Northeast Atlantic, issue). https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/indicator-assessments/changes-plankton-biomass-abundance/.

Louchart, A., Artigas, L.F., Holland, M., Mcquatters-Gollop, A., 2023b. PH3 changes in plankton diversity (the 2023 quality status report for the Northeast Atlantic, issue). https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/indicator-assessments/changes-plankton-diversity/.

Lynam, C.P., Llope, M., Möllmann, C., Helaouët, P., Bayliss-Brown, G.A., Stenseth, N.C., 2017. Interaction between top-down and bottom-up control in marine food webs. Proc. Natl. Acad. Sci. 114 (8), 1952–1957.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., Chanut, J., Clementi, E., Coward, A., Delrosso, D., 2017. NEMO ocean engine. http://hdl.handle.net/2122/13309.

Magliozzi, C., Druon, J.-N., Palialexis, A., Artigas, L., Boicenco, L., González-Quirós, R., Gorokhova, E., Heyden, B., McQuatters-Gollop, A., Varkitzi, I., 2021. Pelagic habitats under MSFD D1: current approaches and priorities: an overview of approaches towards D1C6 assessment (9276309888). Publications Office of the European Union. https://oceanrep.geomar.de/id/eprint/57641.

Makareviciute-Fichtner, K., Matthiessen, B., Lotze, H.K., Sommer, U., 2020. Decrease in diatom dominance at lower Si: N ratios alters plankton food webs. J. Plankton Res. 42 (4), 411–424.

Marques, R., Otto, S., Di Pane, J., Boersma, M., Meunier, C., Wiltshire, K.H., Möllmann, C., Renz, J., 2024. Response of the meso-and macro-zooplankton community to long-term environmental changes in the southern North Sea. ICES J. Mar. Sci. 81 (3), 526–539.

Mayer, M., Mayer, M.M., 2019. Package 'missRanger'. R Package.

McQuatters-Gollop, A., Raitsos, D.E., Edwards, M., Attrill, M.J., 2007. Spatial patterns of diatom and dinoflagellate seasonal cycles in the NE Atlantic Ocean. Mar. Ecol. Prog. Ser. 339, 301–306.

McQuatters-Gollop, A., Gilbert, A.J., Mee, L.D., Vermaat, J.E., Artioli, Y., Humborg, C., Wulff, F., 2009. How well do ecosystem indicators communicate the effects of anthropogenic eutrophication? Estuar. Coast. Shelf Sci. 82 (4), 583–596.

McQuatters-Gollop, A., Edwards, M., Helaouët, P., Johns, D.G., Owens, N.J., Raitsos, D. E., Schroeder, D., Skinner, J., Stern, R.F., 2015. The continuous plankton recorder survey: how can long-term phytoplankton datasets contribute to the assessment of good environmental status? Estuar. Coast. Shelf Sci. 162, 88–97.

McQuatters-Gollop, A., Atkinson, A., Aubert, A., Bedford, J., Best, M., Bresnan, E., Cook, K., Devlin, M., Gowen, R., Johns, D.G., 2019a. Plankton lifeforms as a biodiversity indicator for regional-scale assessment of pelagic habitats for policy. Ecol. Indic. 101, 913–925.

McQuatters-Gollop, A., Mitchell, I., Vina-Herbon, C., Bedford, J., Addison, P.F., Lynam, C.P., Geetha, P., Vermeulan, E.A., Smit, K., Bayley, D.T., 2019b. From science to evidence–how biodiversity indicators can be used for effective marine conservation policy and management. Front. Mar. Sci. 6, 109.

McQuatters-Gollop, A., Guérin, L., Arroyo, N.L., Aubert, A., Artigas, L.F., Bedford, J., Corcoran, E., Dierschke, V., Elliott, S.A.M., Geelhoed, S.C.V., Gilles, A., González-Irusta, J.M., Haelters, J., Johansen .F., M., Lynam, C.P., Niquil, N., Meakins, B., Mitchell, I., Vina-Herbon, C., 2022. Assessing the state of marine biodiversity in the Northeast Atlantic. Ecol. Indic. 141, 109148 https://doi.org/10.1016/j.ecolind.2022.109148.

Millard, S.P., 2013. EnvStats: An R Package for Environmental Statistics. Springer Science & Business Media.

Nejstgaard, J.C., Gismervik, I., Solberg, P.T., 1997. Feeding and reproduction by Calanus finmarchicus, and microzooplankton grazing during mesocosm blooms of diatoms and the coccolithophore Emiliania huxleyi. Mar. Ecol. Prog. Ser. 147, 197–217.

Nicodemus, K.K., Malley, J.D., Strobl, C., Ziegler, A., 2010. The behaviour of random forest permutation-based variable importance measures under predictor correlation. BMC Bioinformatics 11, 1–13.

Nogueira, E., González-Nuevo, G., Valdés, L., 2012. The influence of phytoplankton productivity, temperature and environmental stability on the control of copepod diversity in the North East Atlantic. Prog. Oceanogr. 97, 92–107.

Olin, A.B., Banas, N.S., Johns, D.G., Heath, M.R., Wright, P.J., Nager, R.G., 2022. Spatio-temporal variation in the zooplankton prey of lesser sandeels: species and community trait patterns from the continuous plankton recorder. ICES J. Mar. Sci. 79 (5), 1649–1661.

OSPAR, 2023. Pelagic habitat thematic assessment (The 2023 quality status report for the Northeast Atlantic, issue). https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/thematic-assessments/pelagic-habitats/.

Ostle, C., Paxman, K., Graves, C.A., Arnold, M., Artigas, L.F., Atkinson, A., Aubert, A., Baptie, M., Bear, B., Bedford, J., 2021. The plankton lifeform extraction tool: a digital tool to increase the discoverability and usability of plankton time-series data. Earth Syst. Sci. Data 13 (12), 5617–5642.

Philippart, C.J., Anadón, R., Danovaro, R., Dippner, J.W., Drinkwater, K.F., Hawkins, S. J., Oguz, T., O'Sullivan, G., Reid, P.C., 2011. Impacts of climate change on European marine ecosystems: observations, expectations and indicators. J. Exp. Mar. Biol. Ecol. 400 (1–2), 52–69.

Pitois, S.G., Fox, C.J., 2006. Long-term changes in zooplankton biomass concentration and mean size over the Northwest European shelf inferred from continuous plankton recorder data. ICES J. Mar. Sci. 63 (5), 785–798.

Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9, 181–199.

Prosperi, M., Guo, Y., Sperrin, M., Koopman, J.S., Min, J.S., He, X., Rich, S., Wang, M., Buchan, I.E., Bian, J., 2020. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nat. Mach. Intell. 2 (7), 369–375.

R Core Team, 2020. R: A language and environment for statistical computing. In: In (Version 4.0.2) R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richardson, A., Walne, A., John, A., Jonas, T., Lindley, J., Sims, D., Stevens, D., Witt, M., 2006. Using continuous plankton recorder data. Prog. Oceanogr. 68 (1), 27–74.

Schmidt, K., Birchill, A.J., Atkinson, A., Brewin, R.J., Clark, J.R., Hickman, A.E., Johns, D.G., Lohan, M.C., Milne, A., Pardo, S., 2020. Increasing picocyanobacteria success in shelf waters contributes to long-term food web degradation. Glob. Chang. Biol. 26 (10), 5574–5587.

Scott, J., Pitois, S., Creach, V., Malin, G., Culverhouse, P., Tilbury, J., 2023. Resolution changes relationships: optimizing sampling design using small scale zooplankton data. Prog. Oceanogr. 210, 102946.

Somavilla, R., González-Pola, C., Fernández-Diaz, J., 2017. The warmer the ocean surface, the shallower the mixed layer. H ow much of this is true? J. Geophys. Res. Oceans 122 (9), 7698–7716.

Steinacher, M., Joos, F., Frölicher, T., Bopp, L., Cadule, P., Cocco, V., Doney, S., Gehlen, M., Lindsay, K., Moore, J., 2010. Projected 21st century decrease in marine productivity: a multi-model analysis. Biogeosciences 7 (3), 979–1005.

Tett, P., Carreira, C., Mills, D.K., van Leeuwen, S., Foden, J., Bresnan, E., Gowen, R.J., 2008. Use of a phytoplankton community index to assess the health of coastal waters. ICES J. Mar. Sci. 65 (8), 1475–1482.

Tett, P., Gowen, R., Painting, S., Elliott, M., Forster, R., Mills, D., Bresnan, E., Capuzzo, E., Fernandes, T., Foden, J., 2013. Framework for understanding marine ecosystem health. Mar. Ecol. Prog. Ser. 494, 1–27.

Thierry, S., Dick, S., George, S., Benoit, L., Cyrille, P., 2019. EMODnet Bathymetry a compilation of bathymetric data in the European waters. In: OCEANS 2019, Marseille. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/OCEANSE.2019.8867250.

Thompson, M.S., Pontalier, H., Spence, M.A., Pinnegar, J.K., Greenstreet, S.P., Moriarty, M., Hélaouët, P., Lynam, C.P., 2020. A feeding guild indicator to assess environmental change impacts on marine ecosystem structure and functioning. J. Appl. Ecol. 57 (9), 1769–1781.

Thompson, M.S., Couce, E., Webb, T.J., Grace, M., Cooper, K.M., Schratzberger, M., 2021. What's hot and what's not: making sense of biodiversity 'hotspots'. Glob. Chang. Biol. 27 (3), 521–535.

Thorpe, R.B., Arroyo, N.L., Safi, G., Niquil, N., Preciado, I., Heath, M., Pace, M.C., Lynam, C.P., 2022. The response of North Sea ecosystem functional groups to warming and changes in fishing. Front. Mar. Sci. 9.

Van De Poll, W., Kulk, G., Timmermans, K., Brussaard, C., Van Der Woerd, H., Kehoe, M., Mojica, K., Visser, R., Rozema, P., Buma, A., 2013. Phytoplankton chlorophyll a biomass, composition, and productivity along a temperature and stratification gradient in the northeast Atlantic Ocean. Biogeosciences 10 (6), 4227–4240.

van der Veer, H.W., Witte, J.I., Flege, P., van der Molen, J., Poiesz, S.S., 2024. Spatial and temporal variability in larval connectivity of North Sea plaice *Pleuronectes platessa* between spawning grounds and coastal European nurseries. Mar. Ecol. Prog. Ser. 733, 111–127.

Van Deurs, M., Jørgensen, C., Fiksen, Ø., 2015. Effects of copepod size on fish growth: a model based on data for North Sea sandeel. Mar. Ecol. Prog. Ser. 520, 235–243.

van Leeuwen, S., Tett, P., Mills, D., van der Molen, J., 2015. Stratified and nonstratified areas in the North Sea: long-term variability and biological and policy implications. J. Geophys. Res. Oceans 120 (7), 4670–4686.

Van Leeuwen, S., Le Quesne, W., Parker, E., 2016. Potential future fisheries yields in shelf waters: a model study of the effects of climate change and ocean acidification. Biogeosciences 13 (2), 441–454.

Vezzulli, L., Martinez-Urtaza, J., Stern, R., 2022. Continuous plankton recorder in the omics era: from marine microbiome to global ocean observations. Curr. Opin. Biotechnol. 73, 61–66.

Winder, M., Sommer, U., 2012. Phytoplankton response to a changing climate. Hydrobiologia 698, 5–16.

Wright, M.N., Ziegler, A., 2015. Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. arXiv preprint arXiv:1508.04409.

Xu, X., Lemmen, C., Wirtz, K.W., 2020. Less nutrients but more phytoplankton: long-term ecosystem dynamics of the southern North Sea. Front. Mar. Sci. 7, 662.