# Understanding annual plankton succession and its modifications in the context of past, present and future climate change

Loïck Kléparski

# Thèse de Doctorat

*Mention Sciences agronomiques et écologiques*
*Spécialité Sciences de la Mer – Biologie et écologie*

présentée à *l'Ecole Doctorale en Sciences Technologie et Santé (ED 585)*

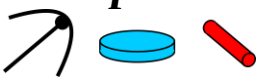
**de l'Université du Littoral Côte d'Opale**

par


## Loïck Kléparski


pour obtenir le grade de Docteur de l'Université du Littoral Côte d'Opale

***Comprendre la succession planctonique annuelle et ses modifications dans le contexte des changements climatiques passés, contemporains et futurs***



***Understanding annual plankton succession and its modifications in the context of past, present and future climate change***

**Soutenue le 30 septembre 2022 après avis des rapporteurs, devant le jury d'examen :**

| | |
|---|---|
| **Pr Lars Stemmann, HDR, Sorbonne Université** | **Rapporteur** |
| **Pr Philippe Koubbi, HDR, Sorbonne Université** | **Rapporteur/Président** |
| **Dr Lawrence Sheppard, Research fellow, MBA** | **Examinateur** |
| **Dr Christophe Luczak, MCF HDR, Université de Lille** | **Examinateur** |
| **Dr Grégory Beaugrand, Directeur de recherche HDR, CNRS** | **Directeur de thèse** |
| **Dr Clare Ostle, Research fellow, MBA** | **Co-encadrante** |
| **Pr Martin Edwards, Research fellow, PML** | **Invité** |

*The highest function of ecology is the understanding of consequences.*

Frank Herbert, Dune

## Publications

***From this PhD:***

**Kléparski, L.**, Beaugrand, G., Edwards, M., 2021. Plankton biogeography in the North Atlantic Ocean and its adjacent seas: Species assemblages and environmental signatures. Ecology and Evolution 11, 5135–5149. https://doi.org/10.1002/ece3.7406

**Kléparski, L.**, Beaugrand, G., Edwards, M., Schmitt, F. G., Kirby, R. R., Breton, E., Gevaert, F., Maniez, E., 2022. Morphological traits, niche-environment interaction and temporal changes in diatoms. Progress in Oceanography 201, 102747. https://doi.org/10.1016/j.pocean.2022.102747

**Kléparski, L.**, Beaugrand, G., 2022. The species chromatogram, a new graphical method to represent, characterise and compare the ecological niches of different species. Ecology and Evolution 12, e8830. https://doi.org/10.1002/ece3.8830

**Kléparski, L.**, Beaugrand, G., Kirby, R. R., 2022. How do plankton species coexist in an apparently unstructured environment? Biology Letters 18, 20220207. https://doi.org/10.1098/rsbl.2022.0207

**Kléparski, L.**, Beaugrand, G., Edwards, M., Ostle, C., 2022. Climate change and phytoplankton phenology in the North Atlantic. Under review in Global Change Biology.

Caracciolo, M., Beaugrand, G., Hélaouët, P., Gevaert, F., Edwards, M., Lizon, F., **Kléparski, L.**, Goberville, E., 2021. Annual phytoplankton succession results from niche-environment interaction. Journal of Plankton Research 43, 85–102. https://doi.org/10.1093/plankt/fbaa060

Edwards, M., Beaugrand, G., **Kléparski, L.**, Helaouët, P., Reid, P.C., 2022. North East Atlantic climate variability and diatom. Communications Earth & Environment 3, 162. https://doi.org/10.1038/s43247-022-00492-9

***Other publications:***

Beaugrand, G., Balembois, A., **Kléparski, L.**, Kirby, R., 2022. Fisheries management, solving the dichotomy of fishing and climate. Accepted in Communication Biology.

Beaugrand, G., **Kléparski, L.**, Luczak, C., Goberville, E., Kirby, R. R. 2022. A new general theory of Island Biogeography. Under review.

Breton, E., Christaki, U., Sautour, B., Demonio, O., Skouroliakou, D. I., Beaugrand, G., Seuront, L., **Kléparski, L.**, Poquet, A., Nowaczyk, A., Crouvoisier, M., Ferreira, S., Pecqueur, D., Salmeron, C., Brylinski, J. M., Lheureux A., Goberville, E., 2021. Seasonal variations in the biodiversity, ecological strategy and specialization of diatoms and copepods in a coastal system with *Phaeocystis* blooms: the key role of trait trade-offs. Frontiers in Marine Science 8, 656300. https://doi.org/10.3389/fmars.2021.656300

Breton, E., Goberville, E., Sautour, B., Ouadi, A., Skouroliakou, D. I., Seuront, L., Beaugrand, G., **Kléparski, L.**, Crouvoisier, M., Pecqueur, D., Salmeron, C., Cauvin, A., Poquet, A., Garcia, N., Gohin, F., Christaki, U. 2022. Phytoplankton community responses to environmental change and effects on ecosystem functioning in a coastal system undergoing unbalanced nutrient reduction: a trait-based approach. Frontiers in Marine Science 9, 914475. https://doi.org/10.3389/fmars.2022.914475

Kirby, R. R., Beaugrand, G., **Kleparski, L.**, Goodall, S., Lavender, S., 2021. Citizens and scientists collect comparable oceanographic data: measurements of ocean transparency from the Secchi Disk study and science programmes. Scientific reports 11, 15499. https://doi.org/10.1038/s41598-021-95029-z

# Acknowledgments

First of all, I would like to thank the Marine Biological Association (MBA, Plymouth, UK) and the Region Hauts-de-France for giving me the opportunity to perform this PhD. I also would like to thank Pr Hubert Loisel and Pr William Wilson for allowing me to work in the Laboratory of Oceanography and Geosciences (LOG) and the Marine Biological Association (MBA).

I am also grateful to my co-supervisors Pr Martin Edwards and Dr Clare Ostle for their precious advice. I would also like to thank some of my co-authors with whom I had the great opportunity to collaborate with: Dr Richard R. Kirby, Dr Mariarita Caracciolo, Dr Elsa Breton, Dr François G. Schmitt and Dr François Gevaert. I hope we can continue to work together in the future.

I would like to thank the owners and crews that have towed the CPRs on a voluntary basis for over 80 years contributing to one of the world's largest and longest ongoing ecological experiments. Without these early pioneers of citizen science and broad-scale volunteer monitoring projects this unique ecological dataset would never have been financially or logistically viable.

I am grateful to the evaluation members who accepted to follow my PhD work (Dr Eric Goberville, Pr Benoit Sautour, Dr Christophe Luczak and Dr François G. Schmitt). I also thank my PhD reviewers, Pr Lars Stemmann, Pr Philippe Koubbi, Dr Christophe Luczak and Dr Lawrence Sheppard.

Je voudrais aussi remercier les personnes qui à Wimereux m'ont aidé à lever la tête de mon écran et/ou dans la réalisation de ce travail. Je voudrais donc dire merci à Monica, Jean-Charles, Marine, Guillaume, Clothilde, Lola, Jérôme, Dewi, Eric, Hélène, Lucdivine, Christelle, Christine et Gislaine. Un grand merci aussi à Philippe et Mickaël pour leur aide précieuse. Je voudrais aussi remercier les stagiaires que j'ai eu l'occasion d'encadrer (officiellement ou non) au cours de ces trois années : Emeline, Hélène, Daniel, Laurie, Léa et Jérémy.

Je voudrais surtout remercier le Dr Grégory Beaugrand pour m'avoir à nouveau fait confiance. Après toutes ces années, il m'est difficile d'écrire de nouveau des remerciements. J'ai énormément appris de vous et j'espère ne pas avoir été un trop mauvais étudiant.

Un grand merci aussi à Emilien, Esteban, Thomas, Benjamin et Aurélien qui m'ont permis de bien me vider la tête à plusieurs reprises. Et bien entendu, un très grand merci à ma famille, sans qui je ne serais pas allé jusqu'au bout de cette quête. Un grand merci à mes frères et sœurs, c'est-à-dire les Michels, Toto et les Bernards, pour toutes les sorties et soirées tout au long de ces trois années (et aussi pour m'avoir accueilli à plusieurs reprises chez vous…). Et, bien entendu, un très grand merci à mes parents, qui m'ont soutenu et supporté sans faillir tout au long de cette aventure, dans les bons comme dans les mauvais jours.

# Table of contents

# General introduction

During James Cook's first expedition on the *Endeavor*, from 1768 to 1771, Joseph Banks made the first known report on phytoplankton cells in the open ocean. This "sea sawdust", as called by the sailors of the time, were in fact colonies of *Trichodesmium*. More than one hundred years later, aboard the SMS *National*, Victor Hensen from Kiel University led the first expedition devoted to what was later termed "plankton" (Fogg, 1990; Karlusich et al., 2020). This German scientist was also one of the first to highlight what is today widely accepted, i.e. the role of phytoplankton as a food basis for aquatic animals (Egerton, 2012).

Although the first observations of phytoplankton cells can be traced back to the invention of microscope during the XVII$^{th}$ century, it was not until the XIX$^{th}$ century that the role of phytoplankton in aquatic ecosystems was recognized (Fogg, 1990; Karlusich et al., 2020). Today, thanks to the progress in genomics and to the geological records, we know that the evolutionary history of eukaryotic phytoplankton started 1.5 billion years ago, when the first photosynthetic eukaryote emerged after the engulfment of a cyanobacterium by a heterotrophic eukaryote (Delwiche, 1999; Simon et al., 2009). This cyanobacterium was probably converted into a plastid and the first members of this new lineage were themselves engulfed by other heterotrophic cells, an event that is thought to be at the origin of the major phytoplanktonic clades during the Middle Triassic that still dominate the extratropical oceans today, i.e. the diatoms, the dinoflagellates and the coccolithophores (Falkowski et al., 2004). Their subsequent evolution associated with the opening of the North Atlantic Ocean, has durably modified the chemical composition of the Earth atmosphere (Katz et al., 2005).

Nowadays, phytoplankton organisms are a key component of marine pelagic ecosystems because they produce almost half of the net primary production (NPP) at a global scale by photosynthesis (Field et al., 1998). This NPP is then transferred to higher trophic levels and ensure the supply of endosomatic energy to heterotrophic organisms. Phytoplankton species undergo strong seasonal variations in abundance, which generate Annual Phytoplankton Successions (APSs) that change little on a year-to-year basis. In the North-East Atlantic, the recurring successional pattern consists in the dominance of diatoms and *Phaeocystis* in spring, followed by coccolithophores, dinoflagellates and nanoflagellates in summer and sometimes a last diatom peak in autumn (Atkinson et al., 2018; McQuatters-Gollop et al., 2007; Widdicombe et al., 2010). This seasonal dynamics affects the whole marine ecosystem and determines the seasonal variations into community composition, influencing the life cycle and migratory behaviour of many species such as copepods or fish (Beaugrand et al., 2003a; Cushing, 1969; Edwards and Richardson, 2004). For example, the survival of haddock larvae and the hatching time of shrimps have been found to be related with the timing of the spring phytoplankton outburst (Koeller et al., 2009; Platt et al., 2003). Furthermore, these massive population peaks sink toward the deep ocean, and contribute to the export of carbon through the biological pump (Henson et al., 2012;

Passow and Carlson, 2012; Smetacek, 1985). Therefore, the study of APS is crucial for our understanding of pelagic ecosystems functioning and services in the context of climate change.

## Phytoplankton phenology

APS is the consequence of species' phenology, which is defined as "*the schedule of recurring biological events in plants and animals*"(Beaugrand and Kirby, 2018).

## Phenological indices

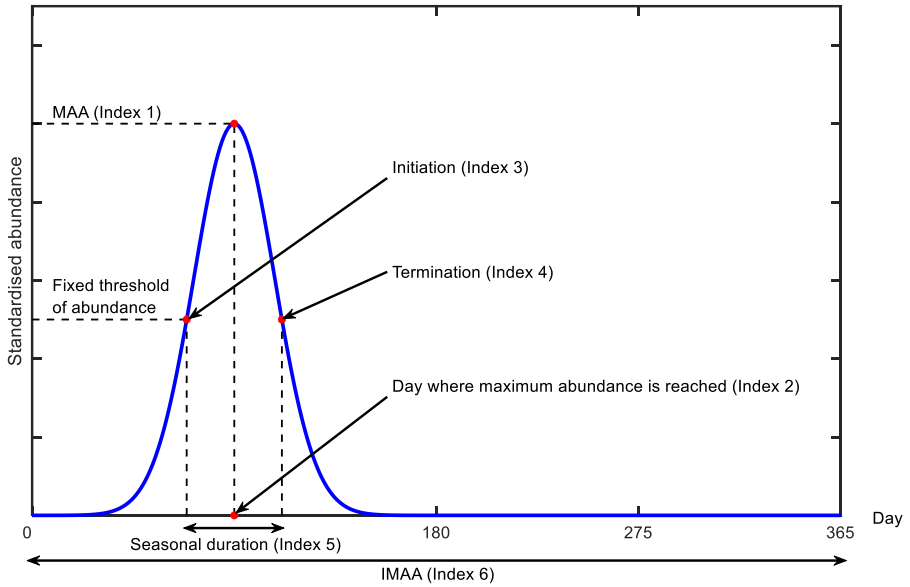



*Figure 1. Theoretical diagram explaining graphically the phenological indices. A species phenology can be described by six indices: the Maximum Annual Abundance (MAA; index 1), the day where MAA is reached (index 2), the initiation (index 3) and the termination (index 4) of the Seasonal Reproductive Period (SRP), i.e. the first and the last days where the abundance is above a given threshold, the seasonal duration (index 5), i.e. the number of consecutive days between the initiation and the termination, and the Integrated Mean Annual Abundance (IMAA; index 6).*

Phenology can be characterised by six different indices (Racault et al., 2012): (1) the Maximum Annual Abundance (MAA), (2) the day where MAA is reached, (3) the initiation and (4) the termination day of the Seasonal Reproductive Period (SRP), i.e. the first and the last days where the abundance of that species is above a given threshold of abundance, (5) the overall seasonal duration, i.e. the number of consecutive days between the initiation and the termination, and (6) the Integrated Mean Annual Abundance (IMAA). The six indices are summarised in Figure 1. The term "bloom" is also regularly used in the literature and here referred as the massive increase of the phytoplankton biomass, attributed to one or more species (Longhurst, 1998; Racault et al., 2012).

## Hypotheses explaining annual phytoplankton succession

Many theories have attempted to explain annual phytoplankton succession through the examination of the mechanisms enabling bloom generation and especially the spring bloom. The first hypothesis, known as the Critical Depth Hypothesis (Sverdrup, 1953), is attributed to H. U. Sverdrup, although it has been formulated by many authors throughout the XX$^{th}$ century (e.g. Gran and Braarud; see Behrenfeld and Boss, 2018). This hypothesis is based on the principle that the balance between population growth by photosynthesis and loss by respirations (assumed to be constant) is controlled by Mixed Layer Depth (MLD) and irradiance along the water column. From the end of autumn, wind-induced mixing causes a deepening of the MLD, which allows the rise of the nutrients concentration from the deep sea toward the surface but also mixed the phytoplankton from the surface down to the pycnocline. As irradiance is low and the MLD is deep, the cells spend most of their time in the dark and their integrated gross rate cannot balance their loss. In spring, the wind diminishes and temperature of the upper water column rises, which induced a shoaling of the MLD. When the MLD becomes shallower than the depth where the loss by respiration strictly balances the gross by photosynthesis (defined has the critical depth; Figure 2), a bloom becomes possible and species populations start to grow. The establishment of a vernal stratification prevents the cell to exit the euphotic zone and the high quantities of both light and nutrients allow a high population growth rate by photosynthesis, which lasts until nutrients diminish. This model has been revisited many times, however.

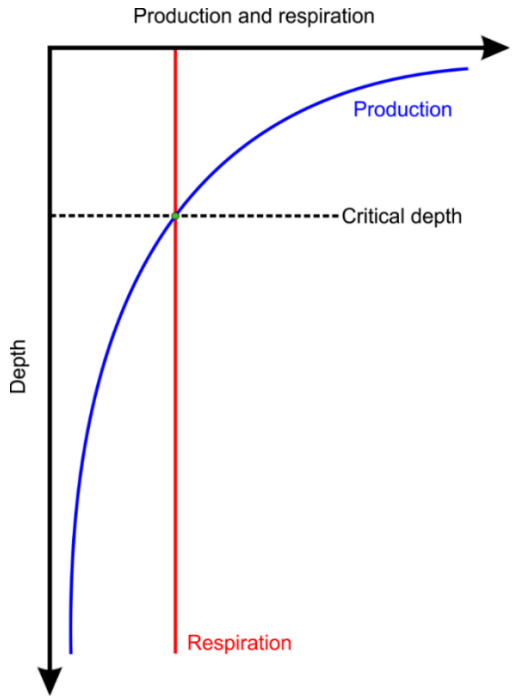



*Figure 2. Simplified diagram displaying the variation with depth of the production by photosynthesis (blue lines) and the loss by respiration (red line). The critical depth (where the production balanced the respiration) is indicated by the green dot. (Adapted from Sverdrup, 1953).*

The critical depth hypothesis has been questioned because of the observation that the phytoplankton population can start to grow well before the shoaling of the MLD (Behrenfeld, 2010). This led M. J. Behrenfeld to propose a new framework, first called the Dilution Recoupling Hypothesis and subsequently renamed the Disturbance Recovery Hypothesis (Behrenfeld, 2014, 2010; Behrenfeld and Boss, 2018). Based on what is observed during dilution experiments, the main assumption of this theory is that there is a strong coupling between phytoplankton production by photosynthesis and loss by grazing, sinking and parasitism. In autumn, when the photoperiod starts to decrease, the phytoplankton population growth rate and the cell concentration within the water column decline. Associated with a deepening of the MLD, this phenomenon causes a dilution of the organisms (i.e. the phytoplanktonic cells, the grazers and the parasites) along the water column, which further relaxes the pressure exerted on phytoplankton by their predators. In winter, while the MLD is deepening, the depth-integrated biomass starts to increase but the cell concentration remains constant because of the dilution in the water column. The encountering rate between phytoplankton and predators (i.e. grazers and parasites) remains low and the cells accumulate in the water column. In late winter and early spring, the shoaling of the MLD associated with a longer photoperiod, induce an acceleration of the division rate, which further increases the concentration of the cells in the upper part of the water column and generate some species populations outburst, i.e. the spring bloom.

## Environmental control of phytoplankton phenology

Large-scale studies have highlighted that phytoplankton phenological patterns are not uniform among oceanic basins, a strong variability being observed with latitudes, suggesting an environmental control of phenology and the resulting local APS. The patterns can be summarised as follows: at high latitudes (above 30°), MAA is important and the duration short, while between 30°S and 30°N, the MAA is lower and the durations longer (Racault et al., 2012). Therefore, high-latitude regions are characterised by fast and intense blooms with high biomass, while in the tropics and subtropics, the seasonality is almost absent and the biomass remains low (Racault et al., 2012). These latitudinal differences are caused by the distinct mechanisms controlling the timing of the SRP: in high-latitude regions, the SRP is initiated by an increase in the photoperiod and therefore a latitude-dependent progression is observed from the equator to the poles (Boyce et al., 2017; Racault et al., 2012). Between the tropics, as light is not a limiting factor, the SRP is initiated when nutrients are supplied in the euphotic zone by increased mixing and no latitudinal-dependent progression is observed (Boyce et al., 2017; Racault et al., 2012). However, some local variability has been detected in the North-East Atlantic Ocean, the SPR being initiated in the North Sea and then spreading outwards (McQuatters-Gollop et al., 2007).

# Plankton biodiversity and species coexistence

## Large scale plankton biodiversity patterns

The great biodiversity exhibited by planktonic organisms has fascinated marine biologists for a long time (Haeckel, 1904; Hardy, 1958, 1926) and it has become more apparent from data collected by the *Tara* Ocean programme that there is even more biodiversity than originally envisioned (De Vargas et al., 2015). At a global scale, plankton biodiversity is not evenly distributed in space but is rather well ordered through the combined actions of climatic, hydrological and topographical factors. Among climatic factors, temperature is probably the most important one, its gradient between the equator and the poles being associated with the latitudinal variations in plankton biodiversity (Righetti et al., 2019; Rombouts et al., 2010, 2009; Rutherford et al., 1999; Thomas et al., 2012). Temperature is important because of its influence at the physiological level, affecting individual growth and reproduction, but also because of its influence on the structure of the water column. Indeed, the thermal difference between the equator and the poles provides the energy for wind-induced mixing which controls nutrients supply in the euphotic zone, a key factor for phytoplankton ecology (Falkowski and Oliver, 2007). Temperature also defines thermal boundaries that separate distinct communities and ecosystems (Beaugrand et al., 2008). Hydrological and topographical factors also influence biodiversity (Barnard et al., 2004; Barton et al., 2010; Beaugrand et al., 2001; Longhurst, 1998) because they define the two main components of a pelagic habitat, i.e. the substrate- and stable-biotope components, respectively (van der Spoel, 1994). The first component is related to oceanic currents, which enable the dispersal of organisms, and the second component is linked with topography, defining neritic and oceanic habitats (van der Spoel, 1994). For example, the dinoflagellate *Ceratium arcticum* is found in the cold Labrador Current while the temperate diatom *Paralia sulcata* occurs above continental shelves (Barnard et al., 2004).

## Plankton coexistence

Mechanisms sustaining plankton biodiversity remain debated because it is assumed that plankton species "*[…] coexist in a relatively isotropic or unstructured environment all competing for the same sort of materials*", a paradigm known as the "paradox of the plankton" (Hutchinson, 1961). Although it has been suggested that plankton species may share their environmental requirements and coexist (Ghilarov, 1984), many theoretical and laboratory studies have tried to resolve the paradox and have shown that the number of coexisting plankton species at equilibrium in a community cannot exceed the number of limiting resources (Levin, 1970; Mac Arthur and Levins, 1964; Rothhaupt, 1988; Tilman, 1977). Consequently, coexistence of a large number of species on a small number of limiting resources is only possible when other mechanisms are involved such as (i) temporal heterogeneity (i.e. a continuously changing pelagic environment due to the seasonal changes in hydro-meteorological forcing) and its influence on the structure of the water column (Bracco et al., 2000; Hutchinson, 1961; Levins, 1979; Richerson et al., 1970), (ii) spatial heterogeneity in the

relative resource supply rate (Tilman, 1980)), (iii) spatial subdivision (Tilman, 1994), (iv) chaotic variations in species abundance induced by competition (Huisman and Weissing, 1999) or (v) difference in the array of resources limiting species growth (Petersen, 1975; Tilman, 1977).

## Plankton and climate change

### Ocean warming

The rise in atmospheric greenhouse gas concentrations (e.g. $CO_2$, $CH_4$ and $NO_2$) is at the origin of a global climate change that may reach a magnitude rarely attained in the geological history of our planet, modifying the environment and threatening the composition and the arrangement of ecosystems (IPCC, 2021). The oceanic hydrosphere has already been altered because it has absorbed 20-30% of the total anthropogenic emissions of $CO_2$ and almost 90% of the heat added to the climate system (IPCC, 2019; Zanna et al., 2019). Therefore, since the second half of the $XX^{th}$ century, the oxygen content of the open ocean has been reduced and the stratification of the oceans has increased, altering nutrient cycles (IPCC, 2019; Sallée et al., 2021).

In the coming century, significant changes in the oceans are predicted by Earth System Models (ESMs), with an increase of Sea Surface Temperature (SST) and stratification, and a decline in pH, dissolved oxygen and nitrate concentration in the euphotic zone (Bopp et al., 2013; IPCC, 2019; Kwiatkowski et al., 2020). A rise of 1.42 ± 0.32°C or 3.47 ± 0.78°C is predicted for SST under the high-mitigation and high-emission scenarios, respectively, and a decline of 0.52 ± 0.23 $mmol.m^{-3}$ or 1.06 ± 0.45 $mmol.m^{-3}$ for nitrate concentration under the high-mitigation and high-emission scenarios, respectively (Kwiatkowski et al., 2020). However, large spatial variabilities are observed with, for example, a more important predicted warming in the Northern Hemisphere, especially in the Arctic and North Pacific Oceans, associated with a minor cooling or a limited warming, depending on the scenario, of the subpolar North Atlantic; this already observed phenomenon has been termed the “warming hole” (Keil et al., 2020; Kwiatkowski et al., 2020). The reinforced stratification, associated with a shoaling of the MLD, are expected to strongly diminish nitrate supply in the North Atlantic and North Pacific Oceans (Kwiatkowski et al., 2020).

### Marine species responses to climate change

Marine species may exhibit a variety of responses to climate change (Beaugrand et al., 2002b; Edwards and Richardson, 2004; Richardson and Schoeman, 2004). Beaugrand and Kirby (2018) have proposed a hierarchical sequence of responses that are briefly summarised here. The first response of an individual to a change in the thermal environment is a modification of their behaviour, i.e. that individuals try to exploit the microheterogeneity of their habitat to find more suitable environmental conditions. For example, some pelagic species go deeper to escape adverse conditions in the surface (Perry et al., 2005). When individuals reach the limits of their behavioural plasticity, they may adjust their physiology, i.e. acclimatization such as a lower activity, growth and development and may exhibit polyphenism

(Beaugrand and Kirby, 2018). When behavioural and physiological changes are no longer possible, phenological adjustments occur, i.e. a shift in the initiation and/or termination of the SRP and when phenological plasticity becomes impossible, local annual abundance decreases and the overall pattern at a large scale may be perceived as a biogeographical shift. If the environmental change is too strong, a local extirpation may be observed.

## Observed changes

Changes in the phenology, abundance and biogeography of plankton have already been documented. Globally, an earlier phenology is observed for both phyto and zooplankton (Friedland et al., 2018; Poloczanska et al., 2016, 2013) but changes are more complex when looking at a finer taxonomic resolution and at smaller spatial scales. In the North-East Atlantic, the phenology of dinoflagellates has been advanced of about thirty days while the phenology of diatoms has not, even if a large inter-species variability is observed for this group (Chivers et al., 2020; Edwards and Richardson, 2004). In contrast, phenology of North Red Sea phytoplankton has been delayed and its duration shortened (Gittings et al., 2018). A global decline in phytoplankton biomass has been documented (Boyce et al., 2010), although local variabilities are observed in some parts of the North-East Atlantic (Reid et al., 1998), where the abundance of diatoms is rising and the abundance of dinoflagellates are declining (Hinder et al., 2012). Northward shifts are also documented, with large variability observed among diatoms and dinoflagellates (Chivers et al., 2017). A decrease in the mean size of zooplankton calanoid copepods has been observed (Beaugrand et al., 2010), being the result of a northward shift in warm-water species (smaller size) associated to a decline in cold-water species (i.e. larger size) (Beaugrand et al., 2002b; Edwards et al., 2021). Hence, shifts in the spatio-temporal distribution of both phyto and zooplankton has altered the composition of the planktonic communities and the resulting new interactome (IPCC, 2019; Winder and Schindler, 2004). As a result, fishery production has declined, although the catch of warm-water fish has increased (IPCC, 2019). (It should be noted that at the same time fish has been strongly impacted by overexploitation (Pauly et al., 1998).) The flux of carbon exported toward the deep ocean has also been modified (Brun et al., 2019). At high latitudes around the pole of the Northern Hemisphere, the reorganization of communities is further amplified by sea-ice melting, which enables species to start to move from an oceanic basin to another, e.g. *Neodenticula seminae* from the Pacific to the Atlantic Ocean (Reid et al., 2007).

## Predicted changes

Because of the increased stratification that will alter nutrient supply in the euphotic zone, a global decline in NPP is expected, although large inter-ESM variability is observed, with some models predicting a decline while others predicting a slight increase (Kwiatkowski et al., 2020). NPP is projected to decrease in the North Atlantic and the western equatorial Pacific, while it is projected to increase in the polar oceans

due to the retreat of sea ice (Kwiatkowski et al., 2020). The reorganisation of the phytoplanktonic community is expected to continue, with an extension of the species that are best adapted to highly stratified low-nutrient waters (Dutkiewicz et al., 2013). Thus, in the North Atlantic, a north-eastward shift in the spatial distribution of diatoms and dinoflagellates is predicted (Barton et al., 2016), as well as a progressive replacement of diatoms by diazotroph (Henson et al., 2021), which participates to the global predicted decline in NPP (Bopp et al., 2005; Marinov et al., 2010). Dinoflagellates are supposed to expand toward the southern ocean (Henson et al., 2021). As a results, plankton biodiversity will be reduced in tropical areas and increased in temperate and polar regions (Beaugrand and Kirby, 2018; Henson et al., 2021; Thomas et al., 2012). Earlier phenology are also expected for phytoplanktonic species, with an advance of 0.5-1 month at a global scale (Henson et al., 2013), although large regional variability are expected with an earlier phenology at high latitudes and a later phenology at mid latitudes (Asch et al., 2019; Henson et al., 2018; Yamaguchi et al., 2022). Phytoplankton phenological shifts might alter fish recruitment, especially for the species that spawn in geographically fixed areas, further impacting fishery catch and productivity that are already altered by other anthropogenic pressures (Asch et al., 2019; IPCC, 2019).

## Projecting species responses to climate change

Projected phenological shifts rely mainly on the analysis of variables such as chlorophyll and NPP, which cannot reproduce the subtle changes that are observed in the phytoplanktonic communities (Asch et al., 2019; Henson et al., 2013; Yamaguchi et al., 2022). To improve our anticipation of the changes that may arise in the next decades, we need to better understand how communities naturally form and are reorganised and what are the underlying mechanisms driving species responses to environmental variability.

### Ecological niche

To be found in a given place, an organism needs to be adapted to the local environmental conditions. This control exerted by the environment is well summarised by the Becking/Beijerinck's law which states that "everything is everywhere, but the environment selects" (De Wit and Bouvier, 2006). The selection occurs because species performances are modulated by the environmental conditions (Shelford, 1913). When conditions are optimal, growth and reproduction are maximal and high abundance are observed (Figure 3; Brown, 1984; Helaouët and Beaugrand, 2009). When the conditions are less favourable, performances decline and the growth, reproduction and feeding are progressively altered and lower abundance are observed, until death (Figure 3; Brown, 1984; Helaouët and Beaugrand, 2009). This paradigm is called the ecological niche of a species and was fully conceptualised by G. E. Hutchinson (Hutchinson, 1957). The niche is therefore defined as the set of environmental conditions enabling a species to grow and reproduce, which determine its response to environmental variability. The niche can be assimilated to a p-dimensional hypervolume, where p ideally represents all the environmental variables

(i.e. both biotic and abiotic)(Hutchinson, 1957). Other concepts of the niche were earlier proposed (e.g. Grinnellian and Eltonian) but were thought to be less operational to our opinion because they envisioned the niche as an attribute of the environment and not of the species (Colwell and Rangel, 2009; Elton, 1927; Grinnell, 1917; Pulliam, 2000).

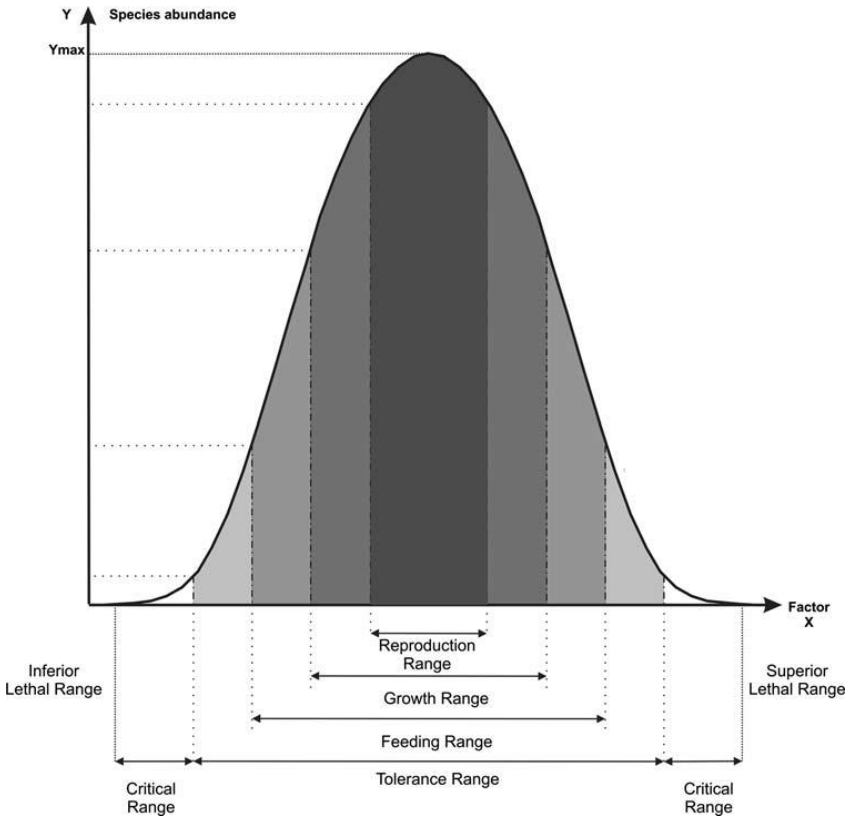



***Figure 3.*** *Theoretical relationship between a species abundance and the variations of a single environmental variable. (Reproduced from Helaouët and Beaugrand, 2009).*

According to Gause's principle, two species cannot coexist in the same geographical area if they possess the same niche because one species excludes the other through competition (Gause, 1934). This competitive exclusion principle, associated with the fact that all the combinations of p-environmental conditions are not permanently represented in the realised environment at the same time (Jackson and Overpeck, 2000), have led to a refinement of the niche concept, i.e. the fundamental, the potential and the realised niches (Hutchinson, 1978; Jackson and Overpeck, 2000). The fundamental niche represents all the combination of the p-environmental variables in which a species can theoretically occurs and the part of those combinations that actually occur in the real world is called the potential niche. The potential niche is subsequently modulated by biotic interaction (i.e. competition or facilitation) and dispersal abilities, which give the realised niche (Bruno et al., 2003; Hutchinson, 1978; Pulliam, 1988).

Many methods have been developed to characterize the multidimensional ecological niche of a species (Beaugrand and Helaouët, 2008; Blonder et al., 2018; Broennimann et al., 2012; Junker et al., 2016). Based on presence or abundance data,

those methods aim at quantifying niche hypervolumes along various environmental dimensions. By doing so, the degree of overlapping among species niches can be assessed and various ecological and evolutionary hypotheses can be examined, such as the role of competition, speciation and habitat filtering in community assembly (Blonder, 2018; Cardillo and Warren, 2016; Cornwell et al., 2006). Furthermore, because of the reciprocal correspondence between the niche space and the real physical space, called "Hutchinson's duality", those methods also enable the reconstruction of past, present and future spatio-temporal species distributions (Colwell and Rangel, 2009; Jackson and Overpeck, 2000). Hence, the niche is a key underlying mechanism to understand the reorganization of annual succession as well as phenological shifts in the context of climate change.

## Functional traits

Functional traits are also important to understand and predict species responses to environmental variability, their spatio-temporal distribution reflecting the distribution of optimal environmental conditions (Ackerly, 2003). A trait is defined as a "*measurable property of organisms, usually measured at the individual level and used comparatively across species*" and a functional trait is defined as a trait "*that strongly influences organismal performance*" (McGill et al., 2006).

Throughout their evolution, phytoplankton species have developed various traits that influence different ecological functions (Litchman and Klausmeier, 2008). For example, diatoms and dinoflagellates have developed resting stages that enable them to survive when environmental conditions become unfavourable (McMinn and Martin, 2013; McQuoid and Hobson, 1996) and also mechanisms that allow them to control their position within the water column to optimize nutrient and light harvesting (Kemp and Villareal, 2018; Miller, 2004). Major phytoplanktonic groups have also developed distinct strategies of resource uptake, with various associated traits and trade-offs, which might explain their distributions and biodiversity in the oceans (Litchman et al., 2007).

The role of phytoplankton functional traits in understanding and predicting APS was first highlighted by Ramon Margalef in his well-known mandala (Margalef, 1978). According to him, major phytoplanktonic taxa can be sorted along gradients of nutrient concentration and turbulence, from diatoms that thrive in high-nutrient low-stratified waters to dinoflagellates that prosper in high-stratified low-nutrient waters; many ESMs predict diatom responses to climate change based on this framework (Bopp et al., 2005; Marinov et al., 2010; Tréguer et al., 2018). More recently, Edwards and colleagues have shown that observed APS in the western English channel can be well reconstructed by means of traits measured in the laboratory and characterising light and nitrate utilisation (K. F. Edwards et al., 2013).

Morphological traits (i.e. cell size and shape) might also be important to explain phenology and APS because they are known to affect phytoplankton biodiversity, resource uptake, vertical motion and grazing resistance (Litchman and Klausmeier, 2008; Ryabov et al., 2021). Morphological traits influence resource uptake

capabilities because nutrient absorption takes place through the action of transport proteins located in the cell surface (e.g. transmembrane pumps). A cell with a larger surface therefore hosts more transport proteins, which optimize nutrient absorption when their concentrations are low; in turn, this cell needs to possess more materials and energy, which is costly (Karp-Boss and Boss, 2016; Naselli-Flores and Barone, 2011). On the contrary, a smaller cell hosts less transport proteins but has a greater surface/volume ratio, which enables nutrients to reach faster any organelles in the cell interior and therefore optimizes physiological reactions (Naselli-Flores and Barone, 2011). A spherical cell has also the lowest nutrient flux compared with other shapes and therefore, the deviation from sphericity confers to a cell a competitive advantage; e.g. elongated shapes experiment the highest nutrient supply (Grover, 1989; Karp-Boss and Boss, 2016; Pahlow et al., 1997).

Morphological traits also modify suspension in the water column, which is thought to be a primary evolutionary target of phytoplanktonic species (Naselli-Flores et al., 2021). Indeed, most of them have developed spines and setae that directly increase the frictional drag or help turn the longest cell dimension perpendicular to the settling axis (Padisak et al., 2003; Smayda, 1970). In the same way, deviation from sphericity enables them to modify sinking velocity and it has been shown that having an elongated shape slows down sinking speed in non-turbulent conditions (Padisak et al., 2003) while it can increase it in turbulent conditions (Clifton et al., 2018). Size is also important in regulating buoyancy: e.g. a disk of 5µm diameter has a smaller sinking rate than a cylinder of same diameter but a disk of 500µm diameter has a higher sinking rate than a cylinder of 500µm diameter (Smayda, 1970).

Morphological traits are also supposed to have evolved in response to predation and might therefore influence interaction with grazers (Karp-Boss and Boss, 2016; Smetacek, 2001). It has been shown that grazing by zooplankton can modify the frequency distribution of some phytoplanktonic shapes, because of a behavioural or mechanical selection by grazers (Sournia, 1982). For example, spherical cells are more easily ingested by copepods than elongated cells, which need to be reoriented (Karp-Boss and Boss, 2016), although cell reorientation by copepods is possible but requires low turbulence to be efficient (Visser and Jonsson, 2000).

Hence, the consideration of both the niche concept and associated functional traits might improve our understanding as well as our predictive capabilities of the consequences of climate change on species phenology and APS.

## The Continuous Plankton Recorder (CPR) survey

To better understand future shifts in phenology and how APS might be altered, we need observations. To do so, long time series with a relatively fine temporal resolution are required, in order to (i) separate the climate-induced trend from natural variability and (ii) have a more precise estimation of the phenological indices (Henson et al., 2018, 2010). However, such long-term datasets are rare and/or spatially constrained to coastal waters (e.g. Station L4 in the English Channel and Point B in the Mediterranean Sea)(Edwards et al., 2010).

The Continuous Plankton Recorder (CPR) survey is a long-term plankton monitoring programme hold by the Marine Biological Association of Plymouth, UK. The CPR survey is the longest and most extensive programme of that kind in the world, covering the North Atlantic Ocean and its adjacent seas (Figure 4)(Richardson et al., 2006). The sampling of phytoplankton and zooplankton is operated by a high-speed plankton recorder, towed behind voluntary merchant ships, call "ships of opportunity" and operating at a depth of approximately 7-10 m (Hays and Warner, 1993; Warner and Hays, 1994). The sampling device was first designed by Alister Hardy during the Discovery expedition in the Southern Ocean in 1925-1927 (Dolan, 2022; Reid et al., 2003). After some modifications brought to the machine, the programme started in the North Sea in 1931 (Reid et al., 2003). After a break during World War II, zooplankton sampling has remained unchanged since 1948 and phytoplankton sampling since 1958 (Warner and Hays, 1994).

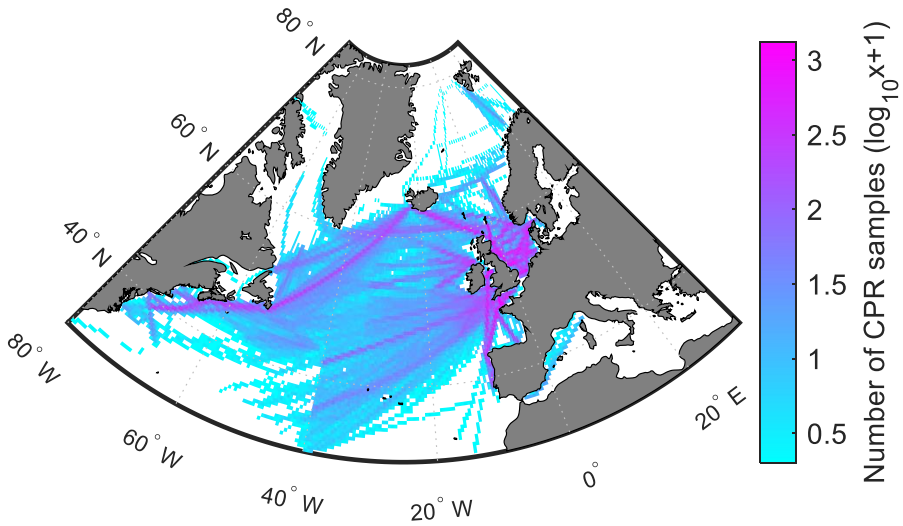



*Figure 4. Spatial distribution of the samples collected by the CPR survey. Colours denote the $log_{10}$ of number of CPR samples in a 0.5° by 0.5° geographical cell. Blue colour denotes a low number of samples and purple a high number.*

The sampling mechanism is described in Figure 5. Water enters through the front opening of the CPR by a 1.27 cm² square entrance aperture and organisms are filtered by a 270 µm mesh silk band, after passing in a tunnel which expand to cross-sectional dimension of 5 cm x 10 cm; this diminishes water pressure on the organisms and minimizes their damage. The mesh silk band is recovered by a second silk band, called the covering silk, as a "sandwich"; sampled organisms are therefore better protected. The whole is finally stored in a tank containing 4% formaldehyde. The silk moves through a propeller action, which turns because of the CPR movement through the water. Silk movement is proportional to ship speed (Reid et al., 2003). Back to the laboratory, the band is unloaded and species are identified. Band is cut into sections of 10 nautical miles, corresponding to ~3 $m^3$ of seawater filtered (Jonas et al., 2004). Further information is provided by the ship crew, e.g. navigational data from the tow such as position of the boat, time of CPR deployment and retrieve, position and ship speed. This information allows the identification of the location of each CPR sample.

Bands are randomly allocated to analysts in a way that they cannot receive consecutive bands from the same tow, in order to minimize analytical bias originating from the experience or the subjectivity of the analyst (Richardson et al., 2006).

The analysis is divided into four steps. **Step 1** is an assessment of the green coloration performed for each CPR sample. This analysis is made by comparison with a four-level standard colour chart: 0 (no colour), 1 (very pale green), 2 (pale green), 3 (green). This estimation is called the Phytoplankton Colour Index (PCI) (Batten et al., 2003) and is considered to be an estimation of the phytoplankton biomass. But as our analysis focussed on species abundance, PCI was not used in this work. The three further steps have a distinct microscope procedure to identify each species in each sample at the highest taxonomic level (Figure 6); up to 500 taxa are identified by the three corresponding procedures (Batten et al., 2003; Richardson et al., 2006; Warner and Hays, 1994). Dinoflagellates, diatoms and copepods are often well preserved, so there are identified at a species level whereas other groups such as chaetognaths and larvaceans are identified to higher taxonomic levels because there are not always well preserved and their identification is more difficult under light microscope (Richardson et al., 2006). **Step 2** identifies and counts phytoplankton cells. On each CPR sample, a cross section of twenty points centred on one silk mesh is defined and there is a semi-quantitative record and an identification of each phytoplankton cell in these twenty points under high magnification (x450) corresponding to a 295 µm field of view. **Step 3** identifies small zooplankton species under x54 magnification corresponding to a 2.06 mm field of view. It is a staggered microscope "traverse" examination of the covering and filtering silks (because some organisms can be transferred between the two silks), with the count and identification of each zooplankton cell ≤2 mm in the field of view of 2.06 mm. **Step 4** identifies all zooplankton species greater than ~2 mm. Large organisms are removed from the silk to be viewed separately. In contrast to Steps 2 and 3, all individuals on the silks (filtering and covering silks) are counted. It is "a compromise between precision of enumeration and speed of processing" (Batten et al., 2003). In this work, we used the three types of abundance data: phytoplankton, zooplankton traverse and zooplankton eye-count.

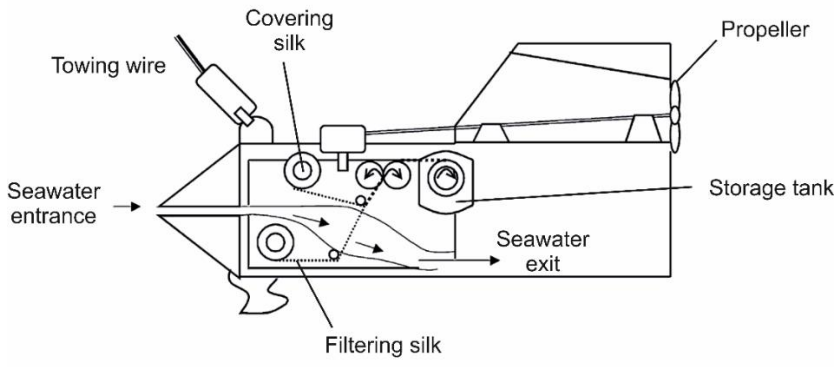



*Figure 5. Diagram of the CPR machine. (Adapted from Batten et al., 2003 and Warner and Hays, 1994).*

It is important to notice that the phytoplankton abundance data assessed by the CPR survey are only semi-quantitative (Richardson et al., 2006). This important

limitation is caused by the relatively large size of silk, which results in an undersampling of small phytoplankton species. However, the CPR collects a consistent fraction of the *in situ* abundance of each taxon and it is therefore considered to be robust when seasonal and inter-annual patterns are examined (Richardson et al., 2006). It is also important to note that the CPR survey is the only scientific monitoring programme of that kind in the world, with no equivalent existing programme. It covers an important time scale from 1948 to present, still active, with a large spatial scale covering the whole North Atlantic Ocean and its adjacent seas such as the Channel, the Celtic Sea and the North Sea. It is "the most extensive long-term survey of marine organisms in the world" (Reid et al., 2003).

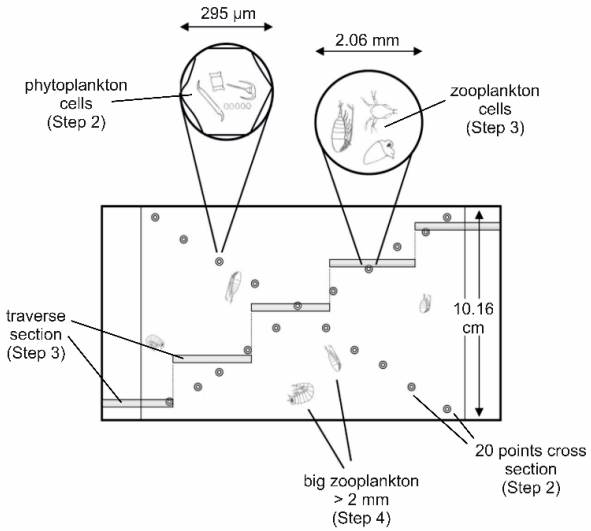



*Figure 6. Diagram of step 2-4 leading to the identification of plankton. Step 2 is phytoplankton identification. Step 3 is traverse zooplankton identification (generally zooplankton length < 2mm). Step 4 is "eyecount" zooplankton identification (generally zooplankton length > 2 mm). (Adapted from Batten et al., 2003).*

## The MacroEcological Theory on the Arrangement of Life

Recently, a new theory called the Macroecological Theory on the Arrangement of Life (METAL) has been proposed. The METAL theory attempts to explain how biodiversity, from the individual to the community level, is organised in space and time and how it responds to environmental changes (Beaugrand, 2015a). The theory uses the concept of the ecological niche *sensu* Hutchinson as a macroscopic brick (Hutchinson, 1957) to connect a large number of observed biogeographical, ecological and paleoecological phenomena. As the niche is determined by the genome, its use enables the consideration of underlying intraspecific processes that are difficult to model (e.g. genetic or molecular processes). The main assumption of METAL is that the niche-environment interaction is a fundamental interaction that propagates from the species to the community and ecosystem level. The strength of METAL is to consider that even though ecosystems are complex adaptive systems, their basic organization and sensitivity can be predicted from simple rules and that a significant proportion of

their changes are deterministic and therefore predictable. The main applications of METAL range from the individual organisational level to the species and community level. More information can be found in https://biodiversite.macroecologie.climat.cnrs.fr.

To reconstruct spatio-temporal changes in the arrangement of community, METAL generates a large number of pseudo-species (i.e. virtual species), which are characterised by a unique niche (in agreement with the competitive exclusion principle). Those pseudo-species are then allowed to colonise a given area or the entire world, so long as they can withstand changes in the environmental regime. The examination of reconstructed pseudo-communities have shown that METAL can reproduce (i) past, present and futures biodiversity patterns (Beaugrand et al., 2015; Zacaï et al., 2021), (ii) abrupt community shifts (Beaugrand et al., 2019a) and (iii) large-scale biogeographical patterns (Beaugrand et al., 2020). This approach has also provided evidence that the arrangement of life is constrained by a mathematical influence that originates from the niche-environment interaction and which limits the number of niches that are available in a given place (Beaugrand et al., 2018). This constraint has been termed the “chessboard of life”.

## Aims of this PhD project

The main objectives of this PhD are (i) to improve our understanding of species phenology at a species level and the resulting APS at a community level, (ii) to identify key underlying mechanisms that control biodiversity, phenology and species coexistence and (iii) to model the past, present and future changes in the phenology of diatoms and dinoflagellates and the resulting APS. Although we first aimed at using our approach at a global scale using data from the Global Alliance of Continuous Plankton Recorder Surveys (GACS), we have finally focused our investigations in the North Atlantic Ocean where the amount of CPR data in space and time was likely to facilitate our analyses and the development of our methods. Nevertheless, we think that the numerical procedures developed in this work may be applied in future studies using GACS data.

This PhD is divided into three chapters. In the first, we describe the seasonal variations of phytoplanktonic taxa well sampled by the CPR survey in the North Sea by means of a METAL-based modelling approach we propose to call PhenoShift and we demonstrate that species phenology results from the niche-environment interaction. We then show that phytoplankton functional traits are connected to the niche and therefore also explain the seasonal variations in species abundance. More precisely, we demonstrate that diatoms with similar cell shape have also similar niches and phenologies, i.e. that oblates (flattened cells) dominate the spring and autumn periods whereas prolates (elongated cells) dominate the summer period. We therefore establish a salient link between functional traits, a species’ niche and phenology. It follows that APS (i.e. annual reorganisation of communities) results from the niche-environment interaction, a fundamental prediction of the METAL theory (Beaugrand and Kirby, 2018).

In the second chapter, we examine how plankton biodiversity is organised in the North Atlantic Ocean. We show that the region is characterised by large spatial coenoclines (i.e. spatial reorganisation of communities) induced by the niche-environment interaction. We also develop a new method, called an “environmental chromatogram” (a species or a community chromatogram when applied at a species or a community level, respectively). A species chromatogram gives a graphical summary of the niche by representing together abundance gradients along various environmental dimensions. This method can be used to characterise and display the multidimensional ecological niches of different species and also to compare the niches of different species by means of an index that quantifies the degree of niche overlapping. We use this index to demonstrate that plankton belonging to the same trophic guild have sufficiently distinct niches to coexist in an impermanent and heterogeneous environment in space and time.

Finally, we apply the METAL-based model PhenoShift in conjunction with projections from six ESMs involved in the Coupled Model Intercomparison Project Phase 6 (CMIP6)(Eyring et al., 2016) to project the past, present and future phenological shifts of three phytoplanktonic groups well sampled by the CPR survey in three key areas of the North Atlantic Ocean, i.e. oblate and prolate diatoms and dinoflagellates. We show that the phenology of oblates is likely to continue to shrink and their abundance to decline whereas the phenology of prolates and dinoflagellates will expand and their abundances rise. Our results suggest a climate-induced reorganisation of the phytoplanktonic community in space and time that will affect the rhythm of generation of endosomatic energy, which may have strong consequences on ecosystem functioning and services.

# ***Chapter 1.*** *Annual phytoplankton succession, phenology, niche and functional traits.*

*Summary*

In this chapter, we investigate the observed Annual Phytoplankton Succession (APS) in the North Sea and identify the underlying mechanisms explaining species phenology in the pelagic environment (Caracciolo et al., 2021). We specifically examine the annual variations in the abundance of 81 species/taxa belonging to five different taxonomic groups, i.e. 53 diatoms, 24 dinoflagellates, one silicoflagellate, two prymnesiophyceae and one cyanophyceae. To do so, we first characterize APS inferred from data collected by the CPR survey by applying a Principal Component Analysis (PCA). We identify five temporal periods, each being characterized by a specific plankton assemblage. Then, we reconstruct the niche of each species along five environmental dimensions by means of an approach based on the MacroEcological Theory on the Arrangement of Life (METAL). We test different models of growing complexity from two to five environmental dimensions to investigate whether an increasing number of ecological parameters can improve our reconstruction of species phenology and APS. Our results show that (i) APS is well modelled by the METAL models using three-dimensional niches based on Sea Surface Temperature (SST), Photosynthetically Active Radiations (PAR) and nitrate concentrations, (ii) APS results from the niche-environment interaction, (iii) this phenomenon should be investigated at a species level and (iv) the use of the most relevant environmental variables appears to be more important than modelling complexity.

Following these results, we investigate whether phytoplankton functional traits influence a species' niche and its phenology and whether they can be used to better understand and reconstruct APS (Kléparski et al., 2022b). To do so, we focus on diatoms and examine their morphological traits, which are major traits affecting nutrients acquisition, light capture, grazing resistance and vertical motion. We attribute to each diatom species/taxa a mean cell height and diameter (information on cell size was retrieved for 45 of the 53 diatoms). According to their mean cell size, 45 diatoms are then divided into three groups: i.e. (i) oblates (i.e. species with a mean cell diameter higher than their mean cell height), (ii) prolates (i.e. species with a mean cell diameter smaller than their mean cell height) and (iii) oblates/prolates for the remaining species/taxa, meaning that they could be both. Annual Diatom Succession (ADS) is identified by means of a PCA and the niches of the 45 species/taxa are characterised along various environmental dimensions (i.e. SST, nutrients concentration, PAR, euphotic and mixed layer depth). Our results show that diatom species with similar morphological traits have similar phenologies and niches, oblate diatoms being mostly abundant in spring and autumn whereas prolate diatoms are dominant during the late spring, summer, autumn period. We show that cell elongation gives a competitive advantage to prolates in the stratified low-nutrient

waters that are found in summer because it enhances their buoyancy without modifying their surface/volume ratio (i.e. their ability to absorb nutrients). Therefore, diatom cell shapes might have evolved as an adaption to viscosity and turbulence and confers to a diatom a niche and a place in the sequence of APS. As a result, we show that oblates and prolates have distinct long-term changes in abundance. The distinction between both groups might drastically change our understanding of marine ecosystem functioning and our predictions on the consequences of future climate changes on the global carbon cycle.

## 1.1. Annual phytoplankton succession results from niche-environment interaction

Mariarita Caracciolo, Grégory Beaugrand, Pierre Helaouët, François Gevaert, Martin Edwards, Fabrice Lizon, Loïck Kléparski and Eric Goberville.

*Published in Journal of Plankton Research. Supplementary Files are displayed in section 1.1.6. Article is available at* https://doi.org/10.1093/plankt/fbaa060

*Abstract*

Annual plankton succession has been investigated for many decades and hypotheses ranging from abiotic to biotic mechanisms have been proposed to explain this recurrent pattern. Here, using data collected by the Continuous Plankton Recorder (CPR) survey and models originating from the MacroEcological Theory on the Arrangement of Life (METAL), we investigate Annual Phytoplankton Succession (APS) in the North Sea at a species level. Our results show that this phenomenon can be well predicted by models combining photosynthetically active radiation, temperature and macro-nutrients. Our findings suggest that APS originates from the interaction between species' ecological niche and annual environmental fluctuations at a community level. We discuss our results in the context of traditional hypotheses formulated to explain this recurrent pattern in the marine field.

*Keywords*

Annual plankton succession, phenology, ecological niche, environment, plankton, Continuous Plankton Recorder (CPR), METAL theory.

### 1.1.1. Introduction

Annual Plankton Succession is defined as the recurrent pattern of species abundance, observed during the annual cycle (Cushing, 1959; Romagnan et al., 2015; Sommer et al., 2012; Winder and Cloern, 2010). In temperate and polar biomes, phytoplankton abundance varies from periods of proliferation in spring and autumn to periods of decline in summer and winter. In subtropical and tropical waters where, seasonal changes in solar radiation and temperature are less prominent, plankton abundance is more stable at an annual scale (Dakos et al., 2009).

Annual Phytoplankton Succession (APS) has been widely described in marine ecosystems, leading to a variety of potential explanations often based on mechanisms such as bottom-up (i.e. nutrients availability, Metaxas and Scheibling, 1996) to top-down control (i.e. grazing by organisms such as zooplankton and fish)(Atkinson et al., 2018; Behrenfeld, 2010; Chiswell, 2011; Chiswell et al., 2015; Huisman et al., 1999; Margalef, 1978; Romagnan et al., 2015; Smyth et al., 2014; Sommer et al., 2012, 1986; Sverdrup, 1953). Gilbert et al. (2012), Romagnan et al. (2015) and Barton et al. (2015) have provided evidence for a strong influence of the physical environment on phytoplankton dynamics, suggesting a bottom-up control of annual succession. A substantial impact of species interaction, such as grazing that imposes a top-down control (e.g. mesozooplankton species on protists), has also been suggested in the western part of the English Channel (Fileman et al., 2010; Kenitz et al., 2017).

The seasonal cycles of irradiance, temperature and stratification - and associated changes in pycnocline, thermocline and halocline - are known to be closely related to the onset of phytoplankton growth (Longhurst, 1998), and nutrients influence the phytoplankton bloom extent (Sommer et al., 2012). Although APS starts typically by the onset of the spring bloom in most extra-tropical regions, winter is a key period during which the ingredients needed to trigger the start of phytoplankton proliferation are prepared (Sommer et al., 2012). During winter, wind intensifies the mixing of the upper layers and an increase in convection leads to a deep mixing involving an augmentation in nitrate, phosphate and silicate surface concentration (Falkowski and Oliver, 2007; Mann and Lazier, 1996), while diluting phytoplankton in the water column (Behrenfeld, 2010).

The main objective of this study is to reconstruct APS using models of increasing complexity - generated from the MacroEcological Theory on the Arrangement of Life (METAL) (Beaugrand, 2015a) - that consider a set of environmental parameters known to influence marine phytoplankton dynamics such as temperature, photosynthetically active radiation and macro-nutrients. The METAL framework unifies behavioural, physiological, phenological biogeographic and long-term community shifts and consequently allows one to model how communities form and how they are altered by environmental fluctuations, including climatic variability and global climate change (Beaugrand et al., 2014, 2013a, 2010; Beaugrand and Kirby, 2018). The strength of this approach is to consider that basic organisation and sensitivity of communities can be predicted from simple founding principles. Even though ecosystems are complex adaptive systems (Levin and Lubchenco, 2008), a significant proportion of the spatial and temporal adjustments of marine communities

are deterministic, which opens the way to testable predictions. In this study, our objectives are to test whether APS is related to the interaction between the ecological niche (sensu Hutchinson (Hutchinson, 1957)) of species and seasonal environmental fluctuations, and to identify the key ecological variables of the niche that control annual phytoplankton dynamics.

Using data from the Continuous Plankton Recorder (CPR) survey (Reid et al., 2003), we characterise APS in an area of the North Sea ranging from 1°E to 4°E and from 54°N to 56°N (Figure S1). The area of consideration is close to the Flamborough Frontal structure, which separates seasonally thermally stratified water to the north and tidally mixed water to the south (Huthnance, 1991; Pingree et al., 1978). We model APS using METAL, comparing observed and predicted patterns, and we investigate how natural environmental fluctuations drive phytoplankton seasonality from initiation to termination. Finally, we discuss our results in the context of APS, including the spring bloom (Romagnan et al., 2015; Sommer et al., 2012; Widdicombe et al., 2010).

## 1.1.2. Materials and methods

### 1.1.2.1. Biological data

Biological data originated from the Continuous Plankton Recorder (CPR, www.cprsurvey.org/data/our-data/) survey, a marine monitoring programme currently operated by the Marine Biological Association (MBA) that samples the North Atlantic Ocean and its adjacent seas since 1946 on a routine monthly basis and at ~7-10 m depth (Reid et al., 2003). Extensively used in the literature, this dataset has allowed researchers to (i) investigate APS (e.g. Barton et al., 2015; Colebrook, 1982, 1979; Zhai et al., 2013), (ii) characterise pelagic biodiversity (e.g. Barnard et al., 2004; Beaugrand et al., 2002), (iii) document distributional, phenological and physiological responses of marine species to climate change (e.g. Beaugrand et al., 2009; Beaugrand and Kirby, 2018; Helaouët and Beaugrand, 2009; Thackeray et al., 2016) and (iv) anticipate the consequences of global warming in the pelagic realm (e.g. Beaugrand et al., 2015; Reid et al., 1998).

Here, we focused our analyses on the phytoplankton community of the North Sea (1-4°E and 54-56°N; Figure S1) and considered 90 species/taxa commonly monitored by the CPR survey over the period 1958-2016 (Table 1). The area was selected for both a regular sampling effort over the study period and its geographical location far from the coast. For each species/taxa, we calculated a daily climatology of species abundances, based on sixty years of sampling.

### 1.1.2.2. Environmental data

Nutrient data originated from the World Ocean Atlas 2013 V2 provided by the NOAA National Centers for Environmental Information (NCEI), Silver Spring, Maryland, USA (www.nodc.noaa.gov/OC5/woa13/woa13data.html)(Garcia et al., 2014). It is a scientifically quality-controlled database of historical in situ surface and subsurface phosphate ($\mu mol.l^{-1}$), silicate ($\mu mol.l^{-1}$) and nitrate ($\mu mol.l^{-1}$) measures. Monthly means

are provided on a 3D grid of 1° latitude by 1° longitude, by 37 depth levels. We calculated the average nutrient concentrations in the area over the first 20m water depth. From nitrate and phosphate concentration, we calculated the N/P ratio (Redfield, 1958) known to influence APS (Falkowski et al., 2000).

We used Photosynthetically Active Radiation (PAR; Einstein.$m^{-2}$.$day^{-1}$), solar radiation spectrum in the wavelength range of 400-700 nm, as a proxy of the level of energy that can be assimilated by photosynthetic organisms (Asrar et al., 1989). Data were provided by the Giovanni online data system, developed and maintained by the NASA GES DISC (gdata1.sci.gsfc.nasa.gov/daac-bin/G3/gui.cgi?instance_id=ocean_month). A monthly climatology of PAR, at a spatial resolution of 9 km, was calculated by compiling Sea-viewing Wide Field-of-View Sensor (SeaWiFS) data from 2009 to 2012.

We also examined annual changes in the Mixed Layer Depth (MLD in meters; Figure S2), using data from the Global Ocean Physical Reanalysis product (GLOBAL_REANALYSIS_ PHY_001_030) provided by the Copernicus Marine Environment Monitoring service (CMEMS; https://marine.copernicus.eu/)).

We assessed the thermal environment of the 90 phytoplankton species over our region of interest using Sea Surface Temperature (SST) from the Optimum Interpolation (OI), which is based on both in situ and satellite observations (see Reynolds et al., 2002 for a full description of the OI analysis). While nutrients and PAR were only available at a monthly scale, daily SST allowed us to calculate daily species thermal preferendum: we first calculated daily SSTs on a 1° by 1° grid from January 1982 to December 2017, and then averaged data for the region ranging from 1°E to 4°E and from 54°N to 56°N (Figure S1). Annual environmental changes are shown in Figure 8.

### 1.1.2.3. The MacroEcological Theory on the Arrangement of Life (METAL)

The MacroEcological Theory on the Arrangement of Life (METAL) is a theory that explains how life is arranged and how changing environmental conditions alter biological arrangements in space and time at different ecological levels (e.g. species, community, ecosystem), allowing predictions to be tested (Beaugrand, 2015a). METAL posits that many ecological (e.g. phenology), biogeographic (e.g. LBGs) and climate-change biology patterns (e.g. phenological and biogeographical shifts) originate from the fundamental niche-environment interaction and unifies a large number of patterns observed in biogeography and ecology at different organisational levels (e.g. spatial range, Rapoport's rule, phenology, latitudinal biodiversity gradients, formation and alteration of species assemblages) and in climate change biology (e.g. phenological shifts, year-to-year to decadal changes in species abundance, range shift, biodiversity shifts, community alteration, abrupt community shifts)(Beaugrand, 2015a; Beaugrand et al., 2020, 2019a, 2018, 2015, 2014, 2013a; Beaugrand and Kirby, 2018).

The theory uses the concept of the ecological niche sensu Hutchinson (Hutchinson, 1957) as a macroscopic elementary brick to understand how species fluctuate in time and space and how communities form and are altered by

environmental fluctuations, including climate change (Figure 7). All species have an ecological niche, which means that they operate within a range of ecological conditions that are suitable for growth and reproduction. The environment acts by selecting species that have the most suitable niche. It follows that this mechanism determines the place where a species lives (i.e. spatial distribution), time when it is active (i.e. phenology) and how individual density fluctuates from short to long time-scales. Locally however, the absence of a species may be explained by species interaction and random processes, such as those discussed in the Unified Neutral Theory of Biodiversity and Biogeography (Hubbell, 2001). The ecological niche, measured by the abundance plotted as a function of some key ecological factors throughout the spatial range of a species, integrates all genetic variations that affect biological traits and physiological processes. More information on METAL can be found in (Beaugrand, 2015a; Beaugrand et al., 2020).

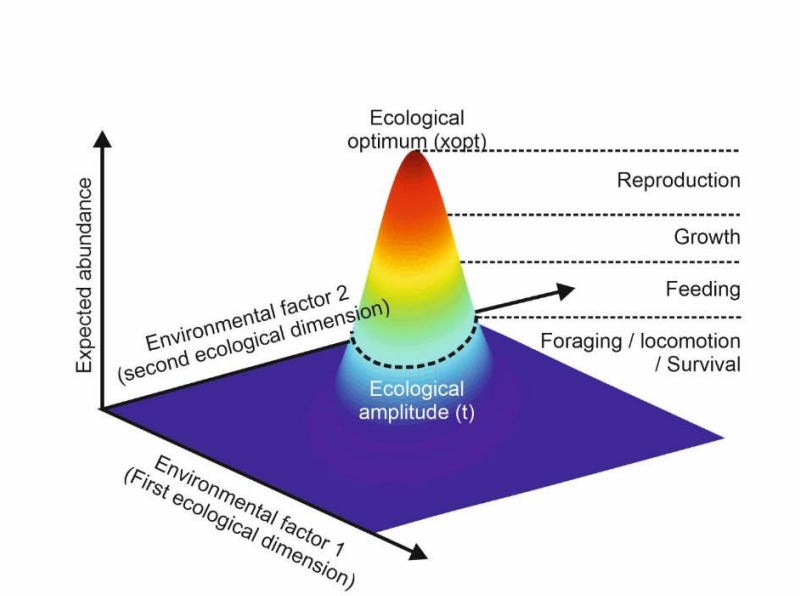



*Figure 7. Representation of a hypothetical 2-dimensional ecological niche of a species. The ecological amplitude is an estimation of the breadth of the species' niche and the ecological optimum is the combination of the environmental factors that are optimal for growth and reproduction. Here, the niche is based on a bi-normal distribution with a correlation coefficient fixed to 0.6 between the two environmental factors. At the optimal part of the niche, all biological processes are operational. When ecological conditions become less favorable, this successively impairs reproduction, growth, feeding, and survival. Ecological dimensions (i.e. the number of ecological variables) varied between 2 and 5 in this study. $x_{opt}$ and t (parameters of Equations 1 and 2) are indicated on the figure.*

### 1.1.2.4. Summary of the numerical procedures

In this paper, we specifically test whether APS originates from the interaction between the ecological niche of a species and environmental fluctuations. First (Step 1 hereafter), we examined APS in the North Sea by means of a Principal Component Analysis (PCA) using data from the CPR survey. APS has been regularly investigated by applying this multivariate technique (Beaugrand et al., 2000a; Colebrook, 1984, 1979). A total of 81 species belonging to different taxonomic groups (e.g. diatoms, dinoflagellates) was considered in this analysis. Secondly (Step 2), we created a large

pool of (Gaussian) ecological niches using a model from the METAL theory and a growing number of ecological dimensions up to five (i.e. environmental parameters, Figure S2). Six environmental parameters were used (e.g. SST, PAR, nitrate, silicate, phosphate, nitrate/phosphate ratio). Thirdly (Step 3), we made 84 simulations building for each a large number of hypothetical ecological niches based on uni- to 5-dimensional niches. All possible niches were therefore built. The maximum abundance of each niche was fixed to 1 and decreasing abundance around the optimum was a function of the environmental amplitude. Estimations of seasonal changes in pseudo-species abundance were calculated at a daily scale by performing a cubic interpolation of the 1-5D niches with the corresponding environmental variables. This calculation step is fully described in Beaugrand and Kirby, 2018 (their Figure 4). We detailed this stage of the procedure in Figure 9 and Figure 10 using simple examples. Fourthly (Step 4), we compared modelled and observed annual patterns in phytoplankton abundance at the species level by means of correlation analyses and the computation of Minimum Assessment Errors (MAE). To consider a possible bias induced by temporal autocorrelation, we tested both correlations and MAEs with null models.

Step 1: Examination of APS from the CPR survey. We first removed species with average annual abundances < 0.5 in the study area, leading to the selection of 81 phytoplankton species (Table 1) for which we estimated average daily abundances for the period 1946-2015. To minimise short-term fluctuations and to reduce the noise inherent to these data, we applied twice an order-6 symmetrical moving average on each daily time series (Legendre and Legendre, 1998).

A standardised PCA (Jolliffe, 1986) was applied on the correlation matrix [81 phytoplankton species x 365 days] and the significant principal components (PCs) were examined to identify changes in annual plankton succession. Species were then sorted according to their phenology by using normalised eigenvectors, i.e. linear correlation values with the corresponding PCs higher than |0.5| (Table 1). Only significant axes (PCs) that best explained APS were considered (Figure 9). We tested the significance of the first three axes by using a broken-stick distribution based on 100,000 simulations (Beaugrand et al., 2019a; Frontier, 1976). Phytoplankton species were then clustered into five groups: Bacillariophyceae, Dinophyceae, Primnesiophyceae, Dictyochophyceae and Cyanophyceae. We reported here the three PCs for which we found correlations between dominant phytoplankton species and principal components (Table 1).

Step 2: Generation of pseudo-species using models from the METAL theory. We modelled patterns of APS using METAL. First, and using each environmental parameter (i.e. nutrients, PAR and SST), we generated a pool of uni-dimensional niches (i.e. niches with only one ecological dimension) based on a Gaussian model (Gauch et al., 1974; Ter Braak, 1996) to calculate species abundance:

$$A = ce^{-\left(\frac{\left(x-x_{opt}\right)^2}{2t^2}\right)} \quad (1)$$

where A is the abundance of a species as a function of the value of a given environmental parameter x; c the maximum abundance of a pseudo-species with c being fixed to 1 (Beaugrand, 2015b); $x_{opt}$, the environmental optimum (e.g. the condition for which a given species reaches the highest level of abundance; Figure 7) and t the ecological amplitude (i.e. the environmental range where a species can occur; Figure 7) of a pseudo-species (Table S2). One dimension of the niche (or one ecological dimension; Figure 7) is represented by the whole range of values of an ecological variable.

Multi-dimensional niches (i.e. niches with more than one ecological dimension) were modelled as follows:

$$A = ce^{-\frac{1}{2}\left[\left(\frac{x_1 - x_{opt1}}{t_1}\right)^2 + \cdots + \left(\frac{x_n - x_{opt}}{t_n}\right)^2\right]} \quad (2)$$

with 2≤n≤5 ecological dimensions, $x_1$ to $x_n$, the values of the environmental parameters, $x_{opt1}$ to $x_{optn}$, the optimum values of $x_1$ to $x_n$, and $t_1$ to $t_n$, the ecological amplitudes of $x_1$ to $x_n$. c is the maximum abundance of a pseudo-species with c being fixed to 1 (Beaugrand, 2015b). We chose a multiplicative model to ensure that one dimension can exclusively control the abundance of a pseudo-species. For example, if the interaction between the silicate niche and the environment leads to a nil or (low) abundance, the abundance will remain nil (or low) whatever the values of the other ecological dimensions. Simulated ecological niches had five ecological dimensions at maximum. We used six ecological variables to build the different niches: (i) SST, (ii) PAR, (iii) nitrate, (iv) phosphate, (v) silicate, and (vi) the N/P ratio. When the N/P ratio was considered, neither nitrate nor phosphate concentrations were included in the niches to avoid possible bias related to multicollinearity. To examine the sensitivity of our analyses to low PAR conditions, three minimum values were considered (Table S2): 1 (termed "PARa"), 10 ("PARb") and 20 ("PARc") $E.m^{-2}.day^{-1}$.

Step 3: Model simulations. We performed simulations using all possible environmental combinations, from one to five ecological dimensions, leading to a total of 84 runs: 16 simulations based on uni-dimensional niches, 23 simulations based on two-dimensional niches, 29 simulations based on three-dimensional niches, 13 simulations based on four-dimensional niches and 3 simulations based on five-dimensional niches (Table S1). The characteristics (optimum and ecological amplitude) of all niches are presented in Table S2. For example, we defined 7 optimum values for temperature (i.e. values corresponding to the highest abundance for a given pseudo-species) from 0 to 36°C - by increment of 6°C - and 4 ecological amplitudes from 1 to 10°C - by increment of 3°C -, leading to the creation of 28 (7 x 4) virtual (or pseudo-) niches.

The core principle of METAL is to generate a large number of pseudo-niches (i.e. simulated niches, we say niche hereafter) and pseudo-species (i.e. virtual species)(Beaugrand et al., 2020) to examine whether the niche-environment interaction (here the annual phytoplankton succession) is responsible for the generation of spatio-temporal patterns in APS. For each simulation, a large number of

niches (from 21 to 15,431,472) was created (Table S2). To determine the total number of niches per simulation, we multiplied the number of niches generated for a given dimension by the number generated for all other ecological dimensions. For example, for a run based on temperature, PAR and nitrate (i.e. a three-dimensional run), the total number of niches was 28 x 21 x 27 = 15,876 ecological niches (with 28 niches for SST, 21 for PARc and 27 for nitrate; Table S2). When all ecological dimensions were considered (a five-dimensional simulation), the total number of niches was 28 (SST) x 27 (PARa) x 27 (nitrate) x 28 (silicate) x 27 (phosphate) = 15,431,472 ecological niches (Table S2). Optimum and ecological amplitude of all niches are presented in Table S1.

To test whether niche resolution - function of the increments (Table S2) - affected our analyses, we compared two extreme cases of uni-dimensional models (i.e. we named them low and high resolutions hereafter) using each ecological variable. The term "bis" was added to identify high-resolution niches (Table S2 and S3). Because of the high number of categories generated in the high-resolution case (e.g. 144,648,000 categories for simulations based on temperature, PARa and phosphate) and the resulting high computation time, we only performed high-resolution analyses for uni-dimensional models.

Finally, annual estimations of pseudo-species abundances were assessed by performing a cubic interpolation of the 1-5D niches with the corresponding environmental variables. Four runs on APS are closely examined as examples: three uni-dimensional runs based on either (i) SST, (ii) PAR or (iii) nitrate (Figure 10) and (iv) one three-dimensional run based on SST, PAR and nitrate (Figure 11). To reveal the simulated patterns in APS, we also applied a standardised PCA and used the PCs to sort the different pseudo-species in the same way as we did for observed APSs (Step 1). We did one PCA per simulation. In the case of these examples, we therefore performed four PCAs: one for the simulation based exclusively on thermal niches (Figure 10a-c), one for the simulation based exclusively on PAR niches (Figure 10d-f), one for the simulation based exclusively on nitrate niches (Figure 10g-i), and a last one for the simulation based on the 3-dimensional (SST, PAR and nitrate) niches (Figure 11). We only examined significant PCs for which we found correlations > |0.5| between pseudo-species and principal components.

Step 4: Comparisons of modelled and observed seasonal patterns. Comparisons between modelled and observed annual patterns in phytoplankton abundance were not based upon regression analyses but assessed by the Pearson's correlation coefficients (Figure 12 and Figure 13) and the Mean Absolute Error (MAE; Figure 13) that measures the average magnitude of the errors in a set of predictions without considering their direction. Equation (3) represents the absolute differences between predictions and observations, divided by the number of differences to be tested (with all individual differences having equal weight):

$$MAE = \sum_{i=1}^{n} \frac{|x_i - y_i|}{n} \qquad (3)$$

with n the number of differences to be tested, $X_i$ is prediction i and $Y_i$ is observation i. The MAE is a negatively-oriented score; the lower values being related to the strongest correlations.

Pearson's correlation coefficients and MAEs were calculated for each run between all observed and modelled daily patterns in (pseudo-) species abundance, leading to a correlation or MAE matrix [species x pseudo-species]. We then identified the highest positive correlations and the lowest MAE values. For each run and species, and using the average correlation and MAE values, we therefore obtained two vectors. To graphically depict the relationships, the daily normalised (between 0 and 1) pseudo-species abundances - that showed the highest correlation with observed species - were plotted against daily observed species abundances (Figure 12).

To consider a possible bias induced by temporal autocorrelation, we tested both correlations and MAEs with null models. First, we randomly generated a number of daily time series corresponding to the total number of pseudo-species generated for each run. The number of time series was small for 1D simulations but became important for an increasing number of dimensions. The procedure was repeated 1000 times and the average Pearson correlation and MAE values were calculated for each simulation. To consider temporal autocorrelation, we generated two million of time series and kept the first 1000 with a 30-order (i.e. 30 days/~one-month autocorrelation for daily time series) autocorrelation higher than average 30-order autocorrelation found in observed daily time series. We represented the results in a diagram that exhibited the observed average correlation for each run and the 1000 correlations found using the null model with (red) and without (blue) autocorrelation (Figure 13). For each combination of environmental variables (i.e. 84 runs), we calculated the probability of significance of each correlation (Table S3) and used contour diagrams to identify (i) the most important environmental parameters and (ii) the number of dimensions to accurately reconstruct APS. This allowed us to highlight the number of species that exhibits the highest correlations in each run (Figure 14).

## 1.1.3. Results

### 1.1.3.1. Seasonal changes in environmental parameters in the North Sea

Temperature exhibited a minimum at the beginning of March and a maximum at the end of July-August (Figure 8). PAR showed minimum and maximum values in December-January and June, respectively. The highest concentrations in nitrate, phosphate and silicate were observed in winter and reached their lowest concentrations from the end of spring to the end of summer. The MLD reaches the sea floor of the studied area every winter (Figure S2), shoals in March to reach the shallowest values (i.e. close to 12 m) between April and September.

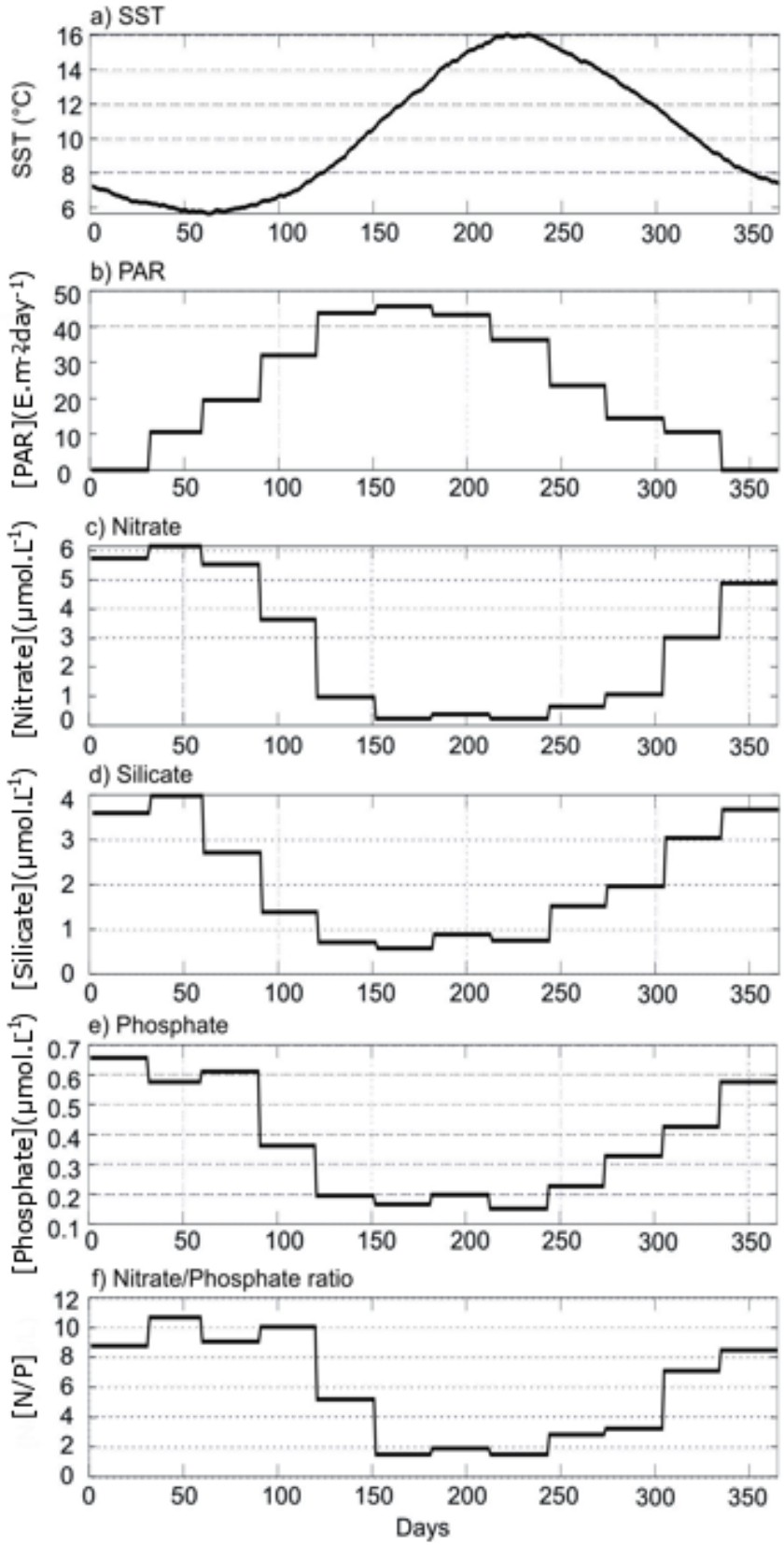



*Figure 8. Annual changes in the environmental parameters considered in this study. (a) Sea Surface Temperature (SST), (b) Photosynthetically Active Radiation (PAR), (c) Nitrate, (d) Silicate, (e) and Phosphate concentrations, and (f) Nitrate/Phosphate (N/P) ratio. Note that SST is at a daily resolution whereas other parameters are at a monthly one (see Materials and Methods).*

### 1.1.3.2. Observed APS

We examined APS based on CPR plankton data by means of a standardised PCA (Figure 9). The first three principal components (PCs) were used because they were significant using a broken-stick model with 100,000 simulations (PC1: 26.86%, PC2: 18.06%, PC3: 12.22%; PCs were above the threshold percentage of 6.14, 4.91 and 4.29%, respectively). The principal components allowed us to differentiate five periods, each being characterised by a species assemblage: (i) an early-spring stage (8 species negatively correlated to PC1; left part of Figure 9), (ii) a spring stage (22 species positively correlated to PC2; Figure 9c), (iii) a widespread summer stage (28 species positively related to PC1; Figure 9a), (iv) a late summer/beginning of autumn stage (13 species negatively correlated to PC3; Figure 9d) and (v) an autumn stage (8 species negatively correlated to PC1; right part of Figure 9b). The summer stage (Figure 9a) was characterised by the highest species richness, but showed a low proportion of diatoms in comparison to both spring or autumn stages. Silicoflagellates were also present (Table 1).

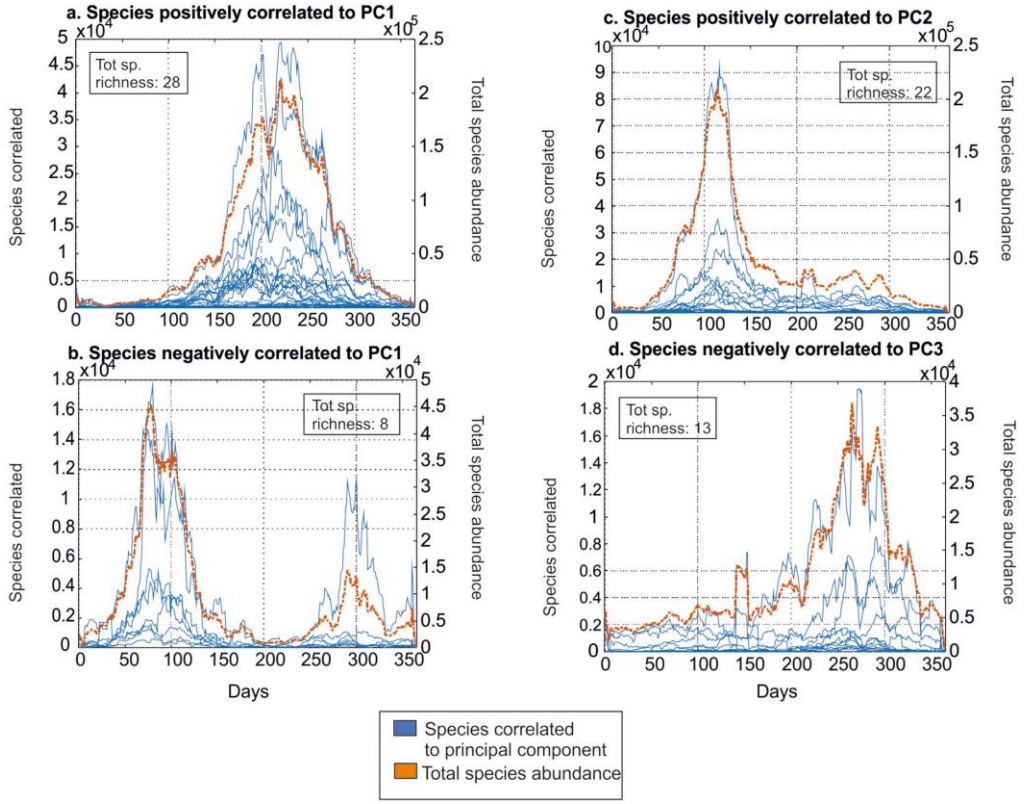



*Figure 9. Annual succession of phytoplankton sorted by PCA. (a) Species positively and (b) negatively correlated with the first principal component (PC1). (c) Species positively correlated with PC2. (d) Species negatively correlated with PC3. Only annual changes in phytoplankton species with normalised eigenvectors negatively (<-0.5) or positively (>0.5) correlated to a corresponding principal component were represented. See Table 1 for information on species and their relations to the PCs. Tot. Sp. Richness: total species richness.*

*Table 1. List of phytoplankton species and their correlation with the first three principal components (PCs). List of phytoplankton species considered in our study area (see Figure S1). The first three PCs considered in Figure 9 and their eigenvalues are reported here. A cross indicates a significant correlation (> |0.5|) between a species and a principal component. Some species were not correlated. The percentage of explained variance per principal component is indicated into brackets. The seasonal cycles of each phytoplankton species are represented on Figure 12 (see species numbers, first column of this table, for correspondence).*

| Species number | Phytoplankton species | PC1 (26.86%) | PC2 (18.06%) | PC3 (12.22%) |
|---|---|---|---|---|
| *1* | *Paralia sulcata* | | | X |
| *2* | *Skeletonema costatum* | X | X | |
| *3* | *Thalassiosira spp.* | | X | |
| *4* | *Dactyliosolen antarcticus* | X | | |
| *5* | *Rhizosolenia styliformis* | | X | |
| 6 | *Rhizosolenia hebetata semispina* | | X | |
| 7 | *Chaetoceros(Hyalochaete) spp.* | | X | |
| 8 | *Chaetoceros(Phaeoceros) spp.* | | X | |
| 9 | *Odontella sinensis* | | | X |
| 10 | *Thalassiothrix longissima* | | X | |
| 11 | *Thalassionema nitzschioides* | X | | |
| 22 | *Asteromphalus spp.* | | | |
| 23 | *Bacteriastrum spp.* | X | | X |
| 24 | *Bellerochea malleus* | X | | X |
| 25 | *Biddulphia alternans* | | | X |
| 26 | *Odontella aurita* | X | | |

| Species number | Phytoplankton species | PC1 (26.86%) | PC2 (18.06%) | PC3 (12.22%) |
|---|---|---|---|---|
| 27 | *Odontella granulata* | | | |
| 28 | *Odontella regia* | X | X | |
| 29 | *Odontella rhombus* | | | |
| 30 | *Cerataulina pelagica* | | | |
| 31 | *Coscinodiscus concinnus* | | X | |
| 32 | *Coscinodiscus spp. (Unidentified)* | X | X | |
| 33 | *Ditylium brightwellii* | X | X | |
| 34 | *Eucampia zodiacus* | | X | X |
| 35 | *Fragilaria spp.* | X | | |
| 36 | *Guinardia flaccida* | | | |
| 37 | *Gyrosigma spp.* | | X | |
| 38 | *Leptocylindrus danicus* | X | | |
| 39 | *Navicula spp.* | | X | |
| 40 | *Cylindrotheca closterium* | | X | |
| 41 | *Rhaphoneis amphiceros* | | | |
| 42 | *Rhizosolenia bergonii* | | | |
| 43 | *Rhizosolenia setigera* | | X | |
| 44 | *Stephanopyis spp.* | | | X |
| 48 | *Nitzschia spp. (Unidentified)* | | | |
| 49 | *Odontella mobiliensis* | | | X |

| Species number | Phytoplankton species | PC1 (26.86%) | PC2 (18.06%) | PC3 (12.22%) |
|---|---|---|---|---|
| 64 | *Proboscia alata* | X | | |
| 65 | *Leptocylindrus mediterraneus* | X | | X |
| 66 | *Proboscia inermis* | | | |
| 67 | *Asterionellopsis glacialis* | | X | |
| 68 | *Ephemera planamembranacea* | | | |
| 69 | *Pseudo-nitzschia delicatissima comple* | | X | |
| 70 | *Pseudo-nitzschia seriata complex* | | X | |
| 72 | *Guinardia delicatula* | | | |
| 73 | *Dactyliosolen fragilissimus* | | X | |
| 74 | *Guinardia striata* | X | | |
| 76 | *Lauderia annulata* | | X | |
| 77 | *Bacillaria paillifera* | X | | |
| 78 | *Corethron hystrix* | X | | |
| 79 | *Proboscia curvirostris* | | X | |
| 80 | *Proboscia indica* | X | | X |
| 81 | *Rhizosolenia imbricata* | X | | |
| 75 | *Helicotheca tamesis* | | | |
| 47 | *Silicoflagellates* | X | | X |
| 12 | *Ceratium fusus* | X | | |

| Species number | Phytoplankton species | PC1 (26.86%) | PC2 (18.06%) | PC3 (12.22%) |
|---|---|---|---|---|
| 13 | *Ceratium furca* | X | | |
| 14 | *Ceratium lineatum* | X | | |
| 15 | *Ceratium tripos* | X | | |
| 16 | *Ceratium macroceros* | X | | X |
| 17 | *Ceratium horridum* | X | | |
| 18 | *Ceratium longipes* | X | | |
| 19 | *Ceratium arcticum* | X | | |
| 20 | *Dinoflagellate cysts (Total)* | X | | X |
| 21 | *Polykrikos schwartzii cysts* | X | | |
| 50 | *Ceratium arietinum* | | | |
| 51 | *Ceratium bucephalum* | | | |
| 52 | *Ceratium buceros* | | | |
| 53 | *Ceratium carriense* | | | |
| 54 | *Ceratium hexacanthum* | X | | |
| 55 | *Ceratium massiliense* | X | | |
| 56 | *Ceratium minutum* | | | X |
| 57 | *Ceratium teres* | | | |
| 58 | *Dinophysis spp. Total* | X | | |
| 59 | *Oxytoxum spp.* | | | |
| 60 | *Protoperidinium spp.* | X | | |
| 61 | *Pronoctiluca pelagica* | | | |

| Species number | Phytoplankton species | PC1 (26.86%) | PC2 (18.06%) | PC3 (12.22%) |
|---|---|---|---|---|
| 62 | *Prorocentrum spp. Total* | X | | |
| 63 | *Noctiluca scintillans* | X | | |
| 45 | *Phaeocystis pouchetii* | | | |
| 46 | *Coccolithaceae (Total)* | X | | |
| 71 | *Trichodesmium spp.* | | | |

### 1.1.3.3. Modelled annual phytoplankton succession

We reconstructed APS by using models of growing complexity (i.e. by considering a growing number of niche dimensions) including all combinations of SST, PAR, nitrate, phosphate, silicate and N/P ratio (a total of 84 simulations). Here, we focused on four examples of modelled APS (Figure 10 and Figure 11). We used the same procedure (standardised PCA) to sort pseudo-species phenology and characterise annual succession. We retained the first two PCs for simulations based exclusively on thermal niches (81.05% and 15.10% of the explained variance; Figure 10a-c), for simulations based exclusively on PAR niches (42.98% and 27.66%; Figure 10d-f) and for simulations based exclusively on nitrate niches (91.84% and 5.10%; Figure 10g-i). The first four PCs (32.30%, 20.90%, 14.82%, and 11.05%) were kept for simulations based on SST, PAR and nitrate (Figure 11), all PCs being significant using the broken-stick distribution based on 100,000 simulations (thresholds of significance were of 6.14, 4.91, 4.29 and 3.88% for PC1-PC4).

The first simulations, based on thermal niches only, showed two main phases of high phytoplankton abundance (also representative of a high species richness) in summer (Figure 10a) and winter (Figure 10b) and two minor phases in spring and autumn (Figure 10c). The winter phase of high abundance did not correspond to any observed patterns (Figure 9 versus Figure 10). The second simulations, based on PAR only, showed several peaks of high phytoplankton abundance in spring, summer and autumn (Figure 10d-f). These patterns were close to observed patterns of annual succession (Figure 9), suggesting an important role of PAR in the modulation of APS. The third simulations - based on nitrate only (Figure 10g-i) - showed an important winter peak in phytoplankton abundance not detected in the observations (Figure 9 versus Figure 10). Considering nitrate only was therefore not sufficient to reconstruct APS. The fourth simulations, that combined SST, PAR and nitrate (Figure 11a-e), were more efficient to reproduce the APS observed in the CPR data, especially during the late-summer phase (Figure 11 versus Figure 9). The relationships between modelled and observed APS is thoroughly examined below.

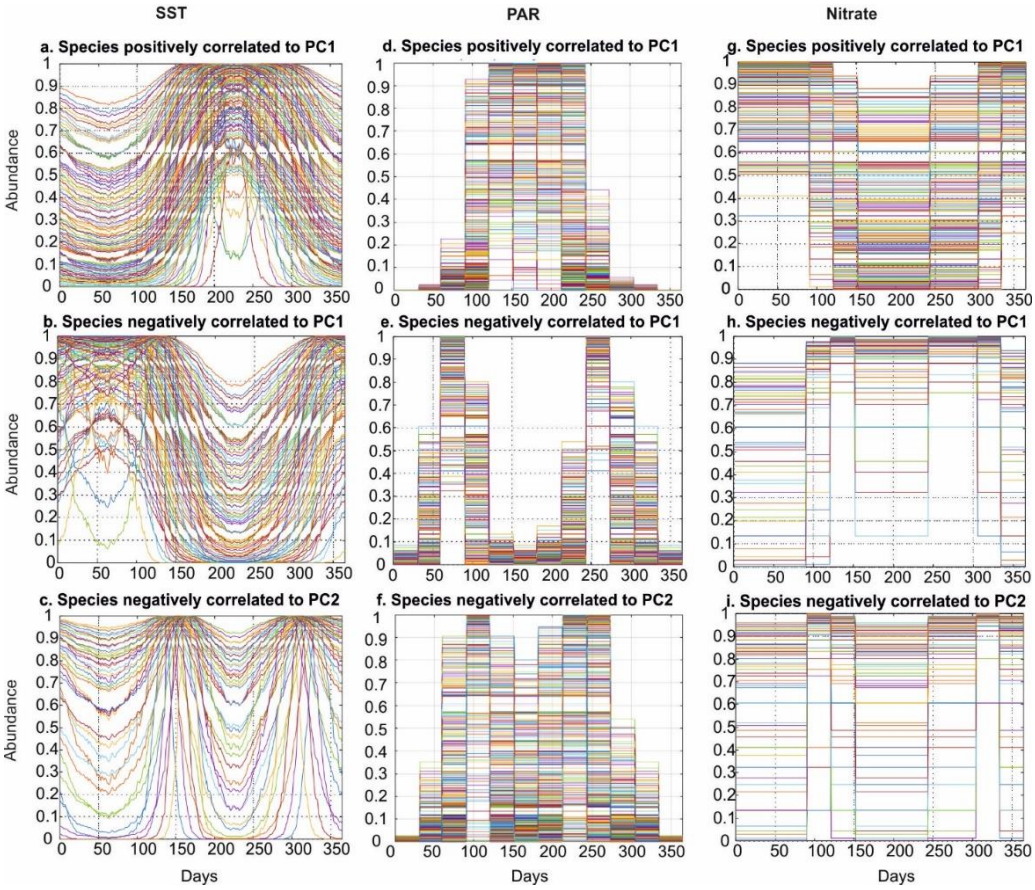



*Figure 10. Reconstructed annual plankton succession from a one-dimensional model based on SST (Sea Surface Temperature, left panels), PAR (Photosynthetically Active Radiation, middle panels) and Nitrate (right panels). A PCA was performed on the relative pseudo-species abundances to identify the most important seasonal phytoplankton abundance patterns. Only modelled plankton seasonal changes, related substantially negatively or positively (i.e. normalised eigenvectors >|0.5|) to the principal components (PCs) are shown. SST (a-c): species (a) positively and (b) negatively correlated to PC1, (c) species negatively correlated to PC2. SST: individual pseudo-species abundance is on the left vertical axis. PAR (d-f): species (d) positively and (e) negatively correlated to PC1, (f) species negatively correlated to PC2. Nitrate (g-i): species (g) positively and (h) negatively correlated to PC1, (i) species negatively correlated to PC2. Relative individual pseudo-species abundances generated from METAL are on the left vertical axis.*

### 1.1.3.4. Reconstruction of species seasonal patterns

We calculated the Pearson's correlation coefficients between observed and modelled phytoplankton abundances for the 84 simulations we performed; we remind here that our simulations were characterised by a growing number of ecological dimensions - ranging from one to five - and that all combinations of environmental parameters were tested. We selected the best correlations and examined graphically the relationships (Figure 12): for most of phytoplankton species or taxa (e.g. *Skeletonema costatum* and *Thalassiosira* spp.), pseudo-species reproduced observed seasonal patterns well, while marginal discrepancies were sometimes observed for some species (e.g. *Paralia sulcata* and *Dactyliosolen antarcticus*; Figure 12). All phytoplankton groups were well modelled (see the colour curves in Figure 12).

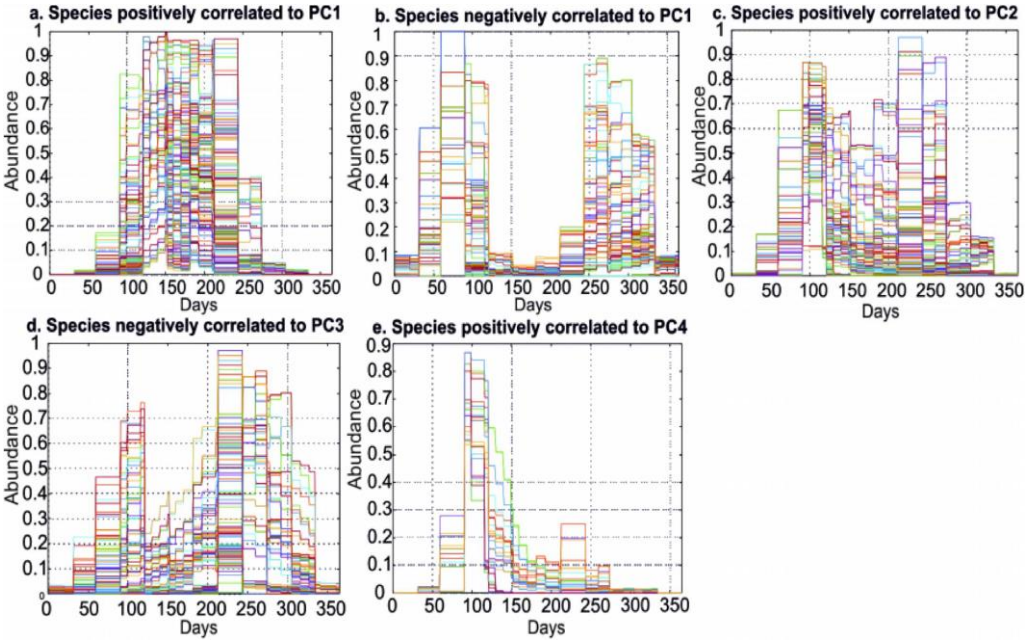



*Figure 11. Reconstructed annual plankton succession from a three-dimensional run based on SST, PAR and Nitrate. A PCA was performed on relative individual pseudo-species abundances to identify the most important seasonal patterns in phytoplankton abundance. Only predicted plankton seasonal changes related substantially negatively or positively (i.e. normalised eigenvectors >|0.5|) to the principal components (PCs) are shown. Species (a) positively and (b) negatively correlated to PC1. (c) Species negatively correlated to PC2. (d) Species negatively correlated to PC3. (e) Species positively correlated to PC4. Individual pseudo-species abundance is on the left vertical axis.*

### 1.1.3.5. Identification of the number of ecological dimensions to reconstruct APS

To identify the number of ecological dimensions to use for reconstructing APS well, we calculated - considering all our simulations - the average of the best correlations and MAEs between observed and modelled phytoplankton abundances (Table S3 and Figure 13). We tested the robustness of correlations and MAEs (e.g. possible bias related to temporal autocorrelation) using null models. While MAE values were sometimes significant for 1D simulations (Figure 13), APS was better reproduced when at least three dimensions were considered (Figure 13). Not all correlations were significant for models based on three or more ecological dimensions; considering five dimensions did not improve the percentage of explained variance (i.e. model quality). This result emphasises that using relevant environmental variables is more important than increasing model complexity.

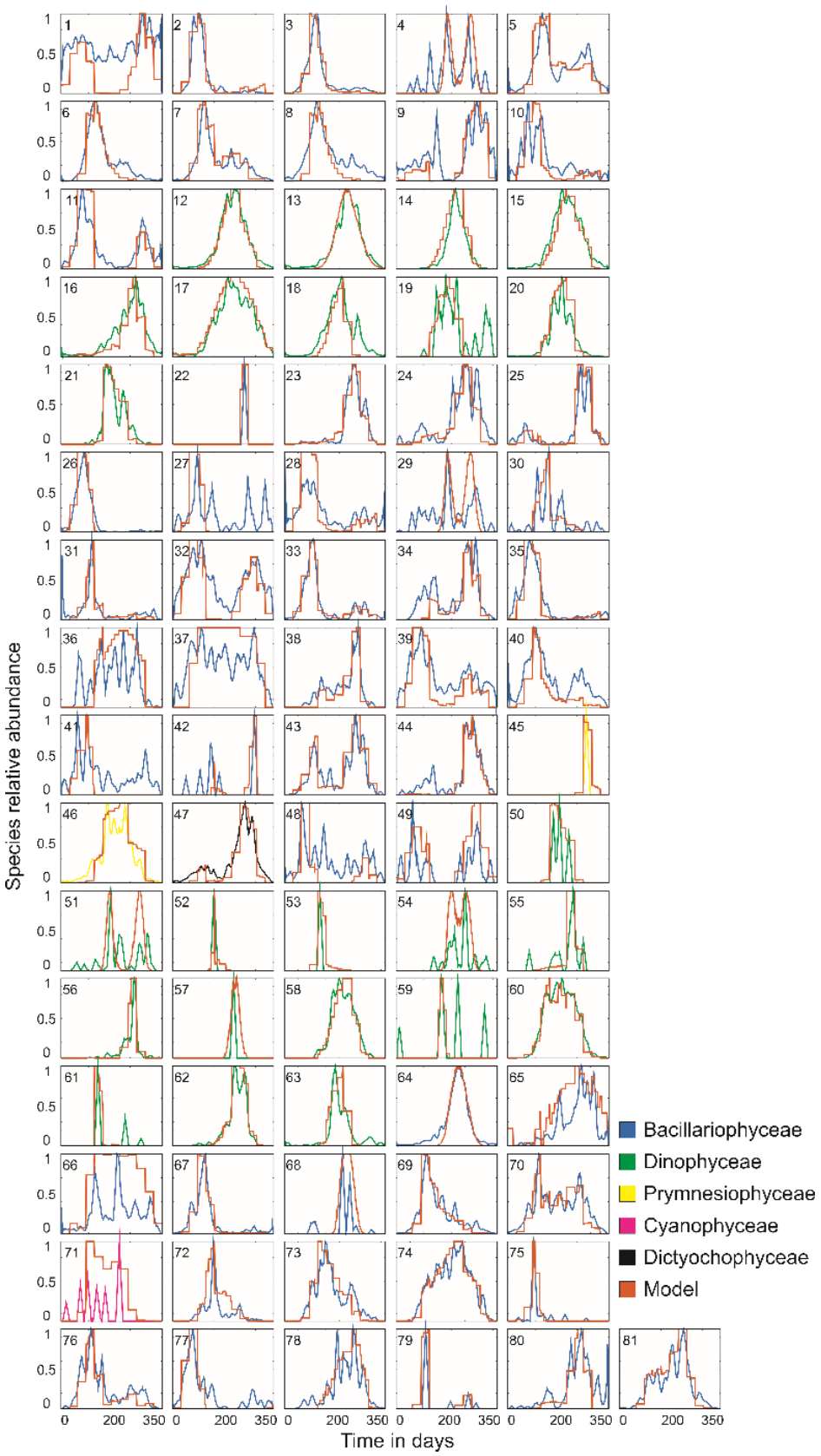



*Figure 12. Seasonal patterns in standardised observed and simulated phytoplankton species abundances. Relative abundances of species sampled by the CPR survey (each taxonomic class being identified by a specific colour, see legend) plotted together with relative abundances of pseudo-species reconstructed using METAL (orange). See Table 1 for species names and taxonomic class.*

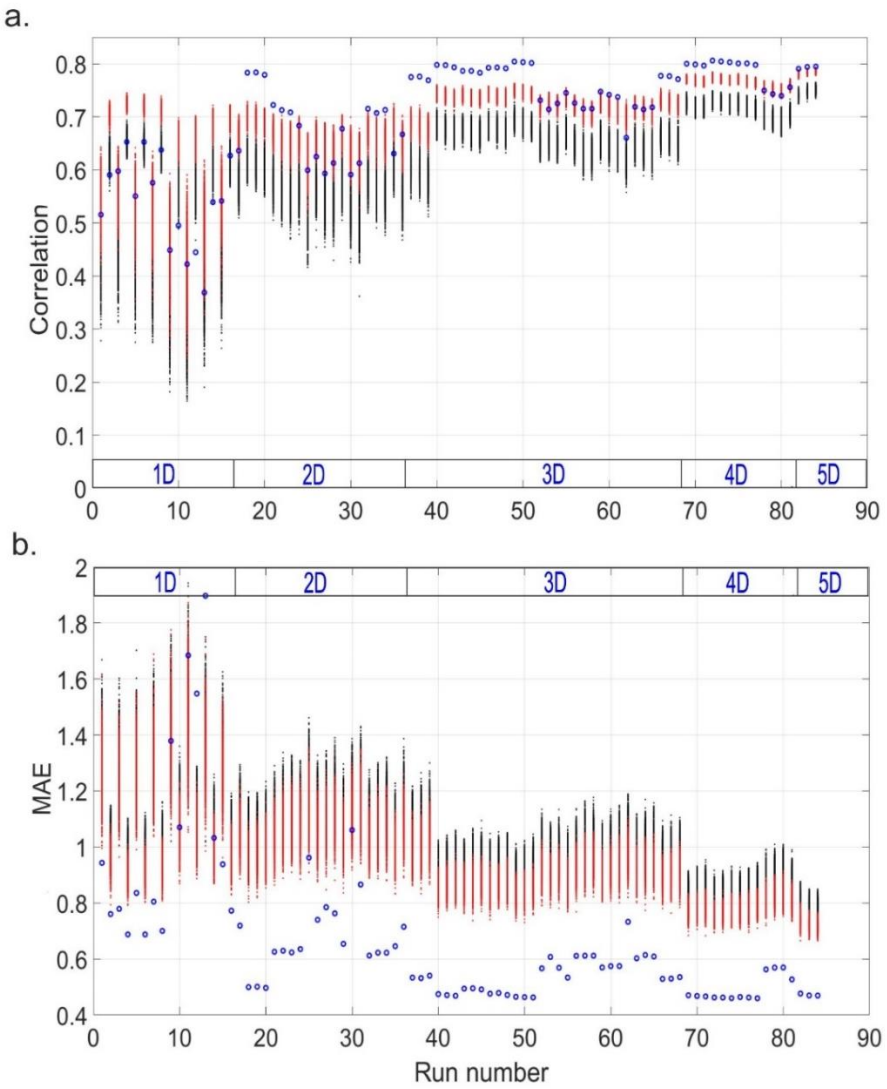



*Figure 13 Average correlation (a) and Mean Absolute Error (MAE) (b) for each run used to reconstruct annual phytoplankton succession from uni-dimensional (1D) to 5-dimensional (5D) models. The average value (blue circle) was based on the best correlations (a) or MAEs (b) assessed between observed species and (simulated) pseudo-species. Black and red points show the results of the same calculations based on a null model with (red) and without (black) consideration of temporal autocorrelation.*

### 1.1.3.6. Identification of key environmental variables to reconstruct APS

We then identified the most relevant environmental variables that best reproduce APS and the seasonal patterns that result of species phenology (Figure 14). Uni-dimensional models (1D, Runs 1-16) explained poorly observed seasonal changes in species abundance, with the exception of Run 2 that was exclusively based on SST (Figure 14a). For Run 2, eight species showed their highest correlations between observed and modelled seasonal patterns (Figure 14a). Two-dimensional models (Runs 17-39) also explained poorly species seasonal patterns and only 3 species exhibited their highest correlations when simulations were based on both temperature and PAR (Figure 14a, Table S1). Better results were achieved when models were based on three or more ecological dimensions: three-, four- and five-dimensional models showed 29 (Runs 40-68), 25 (Runs 69-81) and 14 (Runs 82-84) highest correlations between observed and modelled seasonal patterns, respectively (Figure 14a). Note that Run 51,

based on SST, N/P and PARc (i.e. a minimum value of PAR=20 $E.m^{-2}.day^{-1}$), exhibited 10 highest correlations.

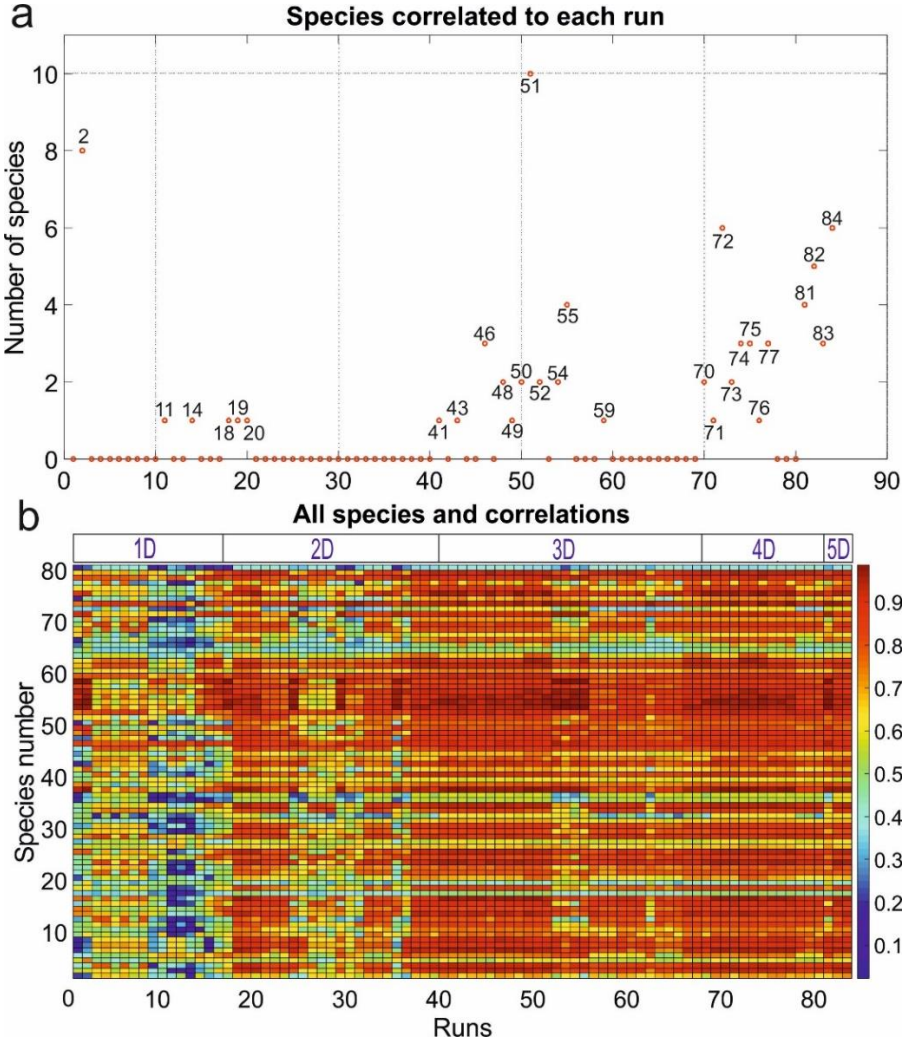



*Figure 14. Identification of the key environmental parameters for reconstructing annual phytoplankton succession. (a) Number of phytoplankton species exhibiting their highest correlation for each model (run). See Table S1 for the correspondence between run numbers and environmental combinations of variables. (b) Highest correlation for a given phytoplankton species and run. The colorbar shows the linear correlation value.*

We also examined the correlations between each simulated and observed seasonal patterns for all species and runs (Figure 14b). Even if the best results were achieved for models based on SST only (Run 2), results were similar when three or more dimensions were included (Figure 14b). Low correlations generally appeared when the triplet SST/PAR/macro-nutrient was not used (Figure 14b and Table S1, e.g. Runs 53-56), revealing that this combination was important to model species seasonal patterns.

## 1.1.4. Discussion

### 1.1.4.1. Annual phytoplankton succession

The application of the Plankton Ecology Group (PEG) model in lakes and subsequently in the marine realm (Sommer et al., 2012, 1986) has suggested that (i)

physics (light and stratification) controls the start and the end of the phytoplankton growth season, (ii) grazing by metazoan plankton results in a clear water phase, (iii) nutrients define the carrying capacity of phytoplankton, (iv) food limitation determines zooplankton abundance and (v) fish predation determines zooplankton size structure.

While grazing may have a substantial influence on phytoplankton (Fileman et al., 2010; Kenitz et al., 2017; Kivi et al., 1993), its non-consideration in our analyses did not prevent us from accurately reconstructing species phenology. Here, we show that annual plankton succession in the North Sea - including the spring bloom - may originate from the niche-environment interaction with a key role of bottom-up processes in shaping APS, as observed by Romagnan and colleagues (Romagnan et al., 2015) for the Mediterranean Sea. Our results suggest that PAR, and to a lesser extent SST, are important for the initiation of the spring bloom, macro-nutrients for the end of the spring bloom and both SST and macro-nutrients for the development of APS. All these parameters (i.e. light, nutrients and temperature) are seen as master parameters controlling photosynthesis in physiological studies (Geider et al., 1997; Longhurst, 1998; McMinn and Martin, 2013; Ras et al., 2013).

In previous works, we suggested that large-scale patterns in biodiversity emerged from the niche-environment interactions that propagate from the species to the community level (Beaugrand, 2015a; Beaugrand et al., 2020, 2013a; Beaugrand and Kirby, 2018). While APS has been frequently investigated at the group level (e.g. plankton functional type, plankton ecology groups or categories), we show that - even within a given ecological or taxonomic group - species reacts to environmental fluctuations individually through the niche-environment interaction, conforming themselves to the principle of species individuality (Whittaker, 1975).

By investigating APS at a species scale, we detected four main microphytoplanktonic successions in the North Sea (see Table 1 for the species list). The first assemblage is composed of species that exhibited their highest abundance at the beginning of spring and a second less important peak in autumn (PC1 in Figure 9b and Figure 12). This microphytoplanktonic assemblage, generally composed of large diatoms such as *Thalassionema nitzschioides* and *Ditylum brightwellii* (Table 1 and Figure 12), was primarily controlled by PAR and nutrients availability. PAR is an essential parameter limiting photosynthesis with a well-known influence on species growth rate (Eppley and Sloan, 1966) that mainly acts in polar regions (McMinn and Martin, 2013), but also in lower latitude areas such as the North Sea (Peeters et al., 1993). The first assemblage is also psychrophilic, reaching its highest (lowest) abundance when temperature is lowest (highest) (Figure 8). When PAR is highest and when PAR or nutrients concentration is lowest (Figure 8), the assemblage is not detected, which is consistent with a positive influence of nutrients on both growth rate and primary production (Goldman, 1980; Longhurst, 1998). Although not considered in our simulations (because of data availability), turbulence, mixing and high SST variability - environmental conditions that characterise early spring and autumn - may also influence positively the first assemblage which is more adapted to this environment than dinoflagellates (Beaugrand et al., 2010; Holligan et al., 1980; Margalef, 1978). In winter, PAR (or the number of daily light hours) and temperature,

to a lesser extent, limit diatom growth; deep-water column mixing combined to an absence of biological production enable nutrients to increase at the surface.

The second assemblage, less psychrophile than the first one and which encompasses species such as *Chaetoceros* spp. or *Coscinodiscus concinnus*, occurs generally between April and June at a time when temperature and PAR increase, and silicate - and nitrate and phosphate, but to a lesser extent - concentrations diminish (Figure 8, Figure 9c and Figure 12).

The third assemblage, mainly composed of dinoflagellates (e.g. *Ceratium fusus* and *C. furca*) and of some small diatoms (e.g. *Guinardia striata* and *G. flaccida*), is observed in oligotrophic conditions and when temperature and PAR are high (Figure 8, Figure 9a and Figure 12). Silicate depletion played an important role in the change in dominance observed between the second and third assemblage. In a mesocosm experiment, silicate deficiency was assumed to be the cause of the strong reduction in large spring bloom diatoms and the replacement by flagellates (Jacobsen et al., 1995). Small diatoms need less silicic acid for their skeleton and have a higher surface to volume ratio which increases nutrient absorption (Miller, 2004). Dinoflagellates occur in areas when temperatures are warm, SST variability is low and the water column is well stabilised (Beaugrand et al., 2010; Margalef, 1978).

The fourth assemblage is composed of late-summer/autumn warm-temperate species (e.g. the diatoms *Bellerochea malleus* and *Biddulphia alternans*; Figure 9d and Figure 12) with a northern distributional limit in the North Sea (e.g. *Bellerochea malleus*) (Barnard et al., 2004). This assemblage occurs when temperature is high and when nutrients concentration tends to increase.

#### 1.1.4.2. The spring bloom

Our study also provides evidence for a strong environmental control of the initiation, development and termination phases of the spring bloom. The integration of PAR, and to a lesser extent SST, in the simulations can simply explain the initiation of the spring bloom in the North Sea. Average light intensity in the mixed layer is known to govern the timing of the spring bloom (Legendre, 1990; Riley, 1967). This is especially the case in the shallow regions of high latitudes (Eilertsen, 1993; Reid et al., 1990; Shaw and Purdie, 2001). Smyth and colleagues (Smyth et al., 2014) have provided evidence that oceanic net heat flux strongly affects ecosystem dynamics and have also conveyed that the spring bloom started in the western part of the English Channel (Station L4, Plymouth) when net heat flux becomes positive. Because net heat flux is highly positively correlated with irradiance and PAR (Beaugrand, 2015a), a strong control of PAR on the initiation of the spring bloom may be expected.

The physical structure of the sea water strongly changes at the time of spring bloom initiation, and many works have suggested that it exerts a strong control, although the debate remains active on the exact types of physical processes that may play a critical role (Atkinson et al., 2018). Our biological model, however, suggests that APS results from the interaction between niche of species and annual environmental fluctuations. Parameters such as those related to vertical mixing would only affect the

abundance by influencing sinking rate and vertical distribution, r = μ-l with r the net specific biomass accumulation rate, μ the phytoplankton growth rate and l a loss term influenced by sinking and vertical mixing (and other processes such as grazing, respiration and parasitism) (Behrenfeld, 2010; Chiswell et al., 2015). In this study, we concentrate on μ at the species level and have not implemented any loss rate in our models. We think this lack of complexity in this shallow region cannot affect our conclusions on the primary control of APS.

Our models suggest that the limitation in macro-nutrients is a key factor for bloom termination. To model the end of the spring bloom, we did not have to consider the influence of grazing in regulating phytoplankton communities and the exhaustion of surface macro-nutrients could explain alone bloom termination (Fischer et al., 2014). Large seasonal changes in atmospheric forcing and ocean surface conditions shape, to a large degree, the seasonal cycles of phytoplankton biomass, but also the relative abundance of phytoplankton species (Barton et al., 2015). Investigating the oceanic region of the North Atlantic, Beaugrand (Beaugrand, 2015a), showed that phytoplankton and zooplankton seasonal fluctuations were closely related (his figure 5.28), suggesting a bottom-up control. More recently, by focussing on a region with approximately the same bathymetry than ours, Atkinson and colleagues (Atkinson et al., 2018) demonstrated that both the increase and termination of the spring bloom were encapsulated by zooplankton, providing strong evidence against a top-down control.

In the pelagic ecosystem of the North Atlantic, diatom blooms end with depletion of silicate and are progressively replaced by slower growing dinoflagellates. Although the succession between diatoms and dinoflagellates is well explained by macro-nutrients and temperature in our simulations, it is also known - since Margalef (1979) - that water column stability is a key factor to explain the succession between these two functional groups. Dinoflagellates are more sensitive than diatoms to turbulence (Karp-Boss et al., 2000). They can undergo significant vertical migrations to nutrient-rich areas but cannot reproduce when turbulence is too high (Estrada and Berdalet, 1997). In contrast, diatoms can continue cell division and the photosynthetic energy products are used to synthesize fatty acid that are converted to energy when cells are exported below the euphotic zone; fatty acid can be considered as a buoyancy regulator (Amato et al., 2017). It is possible that mixing and turbulence are not required in our models because temperature is a proxy of mixing and turbulence conditions in the North Sea (Sharples et al., 2006). Confirmation of our results should be searched in regions that experience different sequences of environmental conditions.

### 1.1.4.3. Uncertainties and potential caveats related to our approach

As with all studies based on modelling and data analysis, it is sometimes difficult to identify primary factors and putative mechanisms at work. In this paper, we have primarily focussed on physical parameters of high biological relevance (temperature, PAR and nutrients) (Behrenfeld and Boss, 2014; Brown et al., 2004; Eilertsen and Degerlund, 2010; McMinn and Martin, 2013; Peeters et al., 1993). We

know that these parameters have a key biological role. However, temperature could also be a correlate for another physical parameter of primary importance such as MLD and depth of the euphotic zone (Beaugrand, 2015a). The role of vertical mixing is primarily to increase nutrient concentrations in surface and to influence phytoplankton sinking rate and vertical distribution (Behrenfeld and Boss, 2014; Chiswell, 2013). In the field, natural systems are more complex than models and all parameters act in synergy.

The niche-environment interaction is certainly more unpredictable in the field than in our modelling approach for two main reasons. First, while the fundamental niche (sensu Hutchinson) was estimated here, the environment - through random meteorological conditions - may influence the realised niche of microalgae species. Second, phytoplankton community before and/or during the growth of a given species may alter species realised niche by competition for resources that lead to competitive exclusion (Barton et al., 2010). For example, the trait-based approach of Breton and colleagues (Breton et al., 2017) suggests that competitive exclusion prevails during *Phaeocystis* spp. blooms in the eastern English Channel.

It is well-known that the underwater light available for photosynthesis (PAR) is a key environmental variable for primary production (Capuzzo et al., 2018, 2015, 2013; Cole and Cloern, 1987; Foden et al., 2010; MacIntyre et al., 2000). Light field in the water column depends in turn on phytoplankton biomass (self-shading), inorganic suspended particulate materials, colored dissolved organic materials and water itself (IOCCG, 2000). Recent works on light quality have also revealed the important role of spectral irradiance on phytoplankton succession (Lawrenz and Richardson, 2017). In this study, we used surface PAR data that originated from a climatology. All phytoplankton species can perform photo-regulation or photo-acclimation (i.e. the first occurs at time scales of minutes and the second takes place in a few hours or a day) to limit photo-inhibition in high light surface waters or optimise both light harvesting and Calvin cycle activity in the water column (Dubinsky and Stambler, 2009; Lavaud, 2007; MacIntyre et al., 2000). In addition, photo-acclimation processes can be conducted on different kinetic models and time scales (Cullen and Lewis, 1988), according to environmental conditions and functional phytoplankton groups (MacIntyre et al., 2000). Even if photosynthesis performances between different species remain poorly documented (Goss and Lepetit, 2015; Suggett et al., 2015), they can induce a competitive effect between species at a given time.

### 1.1.5. Conclusions

Our study suggests that APS may result from the niche-environment interaction and that APS must be investigated at the species level to accurately explore and understand ecological patterns and processes. Our models provide evidence that sharp temporal environmental gradients may be responsible for the strong annual shifts in microphytoplanktonic composition in the North Sea; this occurs when an environmental factor becomes rapidly favourable (e.g. increasing PAR at the end of winter) or limiting (e.g. diminution of macro-nutrients at the end of spring). We

identify the three key parameters that are the best predictors of the succession: (i) temperature, (ii) PAR and (iii) macro-nutrients. There is a clear effect of temperature on APS with a cline from cold-water species in early spring to warm-water species in late summer. By enabling the initiation of the spring bloom and ending the second bloom in autumn, PAR exerts a pivotal role. Macro-nutrients are critical at the end of the spring bloom and their increases in autumn trigger a secondary bloom which then becomes rapidly limited by conditions in PAR and temperature. Mixing is an important process by which macro-nutrients increase in the euphotic zone.

*Acknowledgements*

The CPR Survey is an internationally funded charity that operates the CPR programme. The CPR survey operations and routes are funded by a funding consortium from the UK, USA, Canada and Norway. Within the UK, government organisations DEFRA and NERC contribute to core operations. Part of this research was funded by the Centre National de la Recherche Scientifique (CNRS). This research was funded as part of the ANR TROPHIK.

## 1.1.6. Supplementary Files

**Supplementary Figures.**

**Figure S1. Location of the study area.** The geographical boundary of the rectangle (black box) is 54-56°N and 1-4°E.

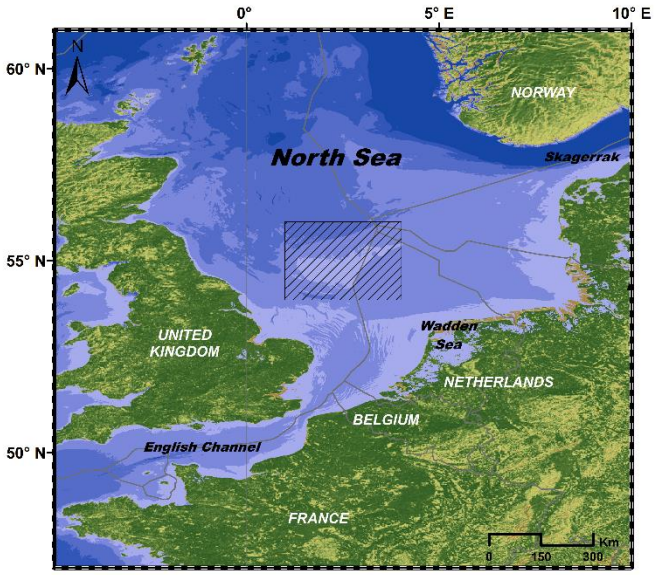
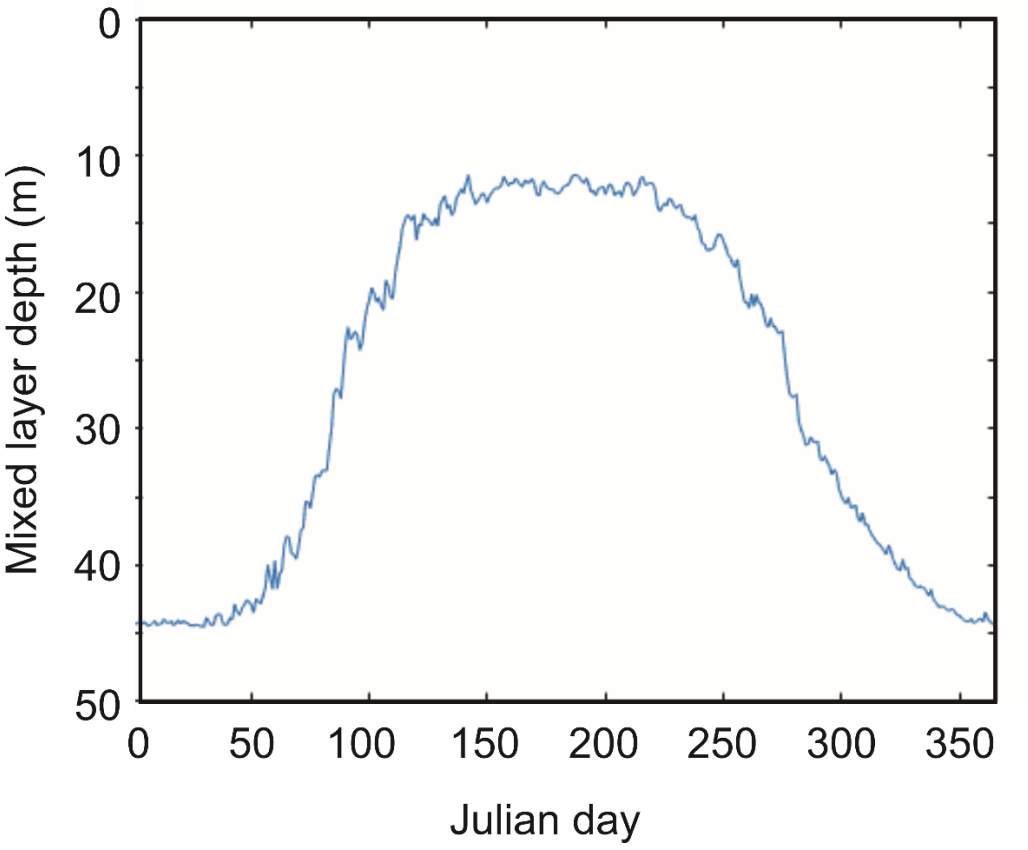


**Figure S2. Mean annual changes in the Mixed Layer Depth (MLD, in meters) in our studied area (54-56°N and 1-4°E).**




**Supplementary Tables.**

**Table S1. Model simulations.** Information on the 84 runs based on all possible combinations of environmental parameters from one (1D; uni-dimensional) to five (5D; 5-dimensional) variables. T: Sea Surface Temperature, Tbis: Sea Surface Temperature using a higher number of niches, PARa,b,c: Photosynthetically Active Radiation (the letters represent 3 different measures used to calculate the optimum values, $E.m^{-2}.day^{-1}$; see text and Table S2 for details), N: Nitrate ($\mu mol.L^{-1}$), S: Silicate ($\mu mol.L^{-1}$), P: Phosphate ($\mu mol.L^{-1}$), N/P: Nitrate/Phosphate ratio (see Table S3 for a better understanding of the selected SST and PAR values). For each run, computation time required for building pseudo-species and calculating species abundances is reported.

| Run n° | Dimensions | Variable | Computation time (hh:mm:ss) |
|---|---|---|---|
| 1 | 1D | T | 00:00:05 |
| 2 | 1D | T bis | 00:00:08 |
| 3 | 1D | PARa | 00:00:13 |
| 4 | 1D | PARa bis | 00:00:17 |
| 5 | 1D | PARb | 00:00:21 |
| 6 | 1D | PARb bis | 00:00:25 |
| 7 | 1D | PARc | 00:00:05 |
| 8 | 1D | PARc bis | 00:00:08 |
| 9 | 1D | N | 00:00:05 |
| 10 | 1D | N bis | 00:00:08 |
| 11 | 1D | S | 00:00:05 |
| 12 | 1D | S bis | 00:00:08 |

| Run n° | Dimensions | Variable | Computation time (hh:mm:ss) |
|---|---|---|---|
| 13 | 1D | P | 00:00:05 |
| 14 | 1D | P bis | 00:00:08 |
| 15 | 1D | N/P | 00:00:05 |
| 16 | 1D | N/P bis | 00:00:08 |
| 17 | 2D | T and N | 00:02:00 |
| 18 | 2D | T and PARa | 00:05:00 |
| 19 | 2D | T and PARb | 00:03:30 |
| 20 | 2D | T and PARc | 00:02:00 |
| 21 | 2D | N and PARa | 00:05:00 |
| 22 | 2D | N and PARb | 00:03:30 |
| 23 | 2D | N and PARc | 00:02:00 |
| 24 | 2D | T and S | 00:02:00 |
| 25 | 2D | N and S | 00:02:00 |
| 26 | 2D | S and PARa | 00:05:00 |
| 27 | 2D | S and PARb | 00:03:30 |
| 28 | 2D | S and PARc | 00:02:00 |
| 29 | 2D | T and P | 00:02:00 |
| 30 | 2D | N and P | 00:02:00 |
| 31 | 2D | S and P | 00:02:00 |
| 32 | 2D | P and PARa | 00:05:00 |
| 33 | 2D | P and PARb | 00:03:30 |
| 34 | 2D | P and PARc | 00:02:00 |
| 35 | 2D | T and N/P | 00:02:00 |
| 36 | 2D | S and N/P | 00:02:00 |
| 37 | 2D | N/P and PARa | 00:05:00 |
| 38 | 2D | N/P and PARb | 00:18:00 |
| 39 | 2D | N/P and PARc | 00:15:00 |
| 40 | 3D | T, N and PARa | 00:20:00 |
| 41 | 3D | T, N and PARb | 00:18:00 |
| 42 | 3D | T, N and PARc | 00:15:00 |
| 43 | 3D | T, S and PARa | 00:20:00 |
| 44 | 3D | T, S and PARb | 00:18:00 |
| 45 | 3D | T, S and PARc | 00:15:00 |
| 46 | 3D | T, P and PARa | 00:20:00 |
| 47 | 3D | T, P and PARb | 00:18:00 |
| 48 | 3D | T, P and PARc | 00:15:00 |
| 49 | 3D | T, N/P and PARa | 00:20:00 |
| 50 | 3D | T, N/P and PARb | 00:18:00 |
| 51 | 3D | T, N/P and PARc | 00:15:00 |
| 52 | 3D | T, N and S | 00:15:00 |
| 53 | 3D | T, N and P | 00:15:00 |
| 54 | 3D | T, S and P | 00:15:00 |
| 55 | 3D | T, S and N/P | 00:15:00 |
| 56 | 3D | N, S and PARa | 00:20:00 |
| 57 | 3D | N, S and PARb | 00:18:00 |
| 58 | 3D | N, S and PARc | 00:15:00 |
| 59 | 3D | N, P and PARa | 00:20:00 |
| 60 | 3D | N, P and PARb | 00:18:00 |
| 61 | 3D | N, P and PARc | 00:15:00 |
| 62 | 3D | N, S and P | 00:15:00 |
| 63 | 3D | S, P and PARa | 00:20:00 |

| Run n° | Dimensions | Variable | Computation time (hh:mm:ss) |
|---|---|---|---|
| 64 | 3D | S, P and PARb | 00:18:00 |
| 65 | 3D | S, P and PARc | 00:15:00 |
| 66 | 3D | S, N/P and PARa | 00:20:00 |
| 67 | 3D | S, N/P and PARb | 00:18:00 |
| 68 | 3D | S, N/P and PARc | 00:15:00 |
| 69 | 4D | T, N, S and PARa | 96:00:00 |
| 70 | 4D | T, N, S and PARb | 72:00:00 |
| 71 | 4D | T, N, S and PARc | 60:00:00 |
| 72 | 4D | T, N/P, PARa and S | 120:00:00 |
| 73 | 4D | T, N/P, PARb and S | 96:00:00 |
| 74 | 4D | T, N/P, PARc and S | 72:00:00 |
| 75 | 4D | T, N, P and PARa | 36:00:00 |
| 76 | 4D | T, N, P and PARb | 36:00:00 |
| 77 | 4D | T, N, P and PARc | 24:00:00 |
| 78 | 4D | S, N, P and PARa | 60:00:00 |
| 79 | 4D | S, N, P and PARb | 60:00:00 |
| 80 | 4D | S, N, P and PARc | 48:00:00 |
| 81 | 4D | T, N, S and P | 48:00:00 |
| 82 | 5D | T, N, S, P and PARc | 480:00:00 |
| 83 | 5D | T, N, S, P and PARb | 984:00:00 |
| 84 | 5D | T, N, S, P and PARa | 1176:00:00 |

**Table S2. Environmental variables used for the calculation of pseudo-species abundances and respective optimum and tolerance values.** For each environmental parameter the table shows the different range of optimum values and ecological amplitudes defined for niche construction. We used several resolutions (first column) to calculate the environmental niche. For runs ended by "bis", the resolution was improved to examine model sensitivity related to the number of points used to calculate the niche. To examine the sensitivity of our analysis to PAR, three categories (a, b, c) were determined by selecting different minimum values (see text). When more than one factor was considered, the number of niches was multiplied by each ecological dimension to obtain the total number of niches (see text).

| Environmental variable | Optimum | Number of optimum values | Tolerance | Number of ecological amplitudes | Number of niches |
|---|---|---|---|---|---|
| **Temperature**= -2, -1, 0,…,44 | 0, 6, 12,…, 36 | 7 | 1, 4, 7, 10 | 4 | 28 |
| **Temperature bis**= -2, 1.99, 1.98,…,44 | 0, 1, 2,…, 40 | 41 | 1, 2, 3,…, 40 | 10 | 410 |
| **PARa**=0, 1, 2,…, 70 | 1, 9, 17,…, 70 | 9 | 1, 5, 9 | 3 | 27 |
| **PARa bis**= 0, 0.25, 0.50,…, 70 | 1, 2, 3,…, 70 | 70 | 1, 2, 3,…, 9 | 9 | 630 |
| **PARb**= 0, 1, 2,…, 70 | 10, 18, 26,…, 70 | 8 | 1, 5, 9 | 3 | 24 |
| **PARb bis**= 0, 0.25, 0.5,…, 70 | 10, 11, 12,…, 70 | 60 | 1, 2, 3,…, 9 | 9 | 540 |

| Environmental variable | Optimum | Number of optimum values | Tolerance | Number of ecological amplitudes | Number of niches |
|---|---|---|---|---|---|
| **PARc**= 0, 1, 2,..., 70 | 20, 28, 36,…, 70 | 7 | 1, 5, 9 | 3 | 21 |
| **PARc bis**= 0, 0.25, 0.5,…, 70 | 20, 21, 22,…, 70 | 50 | 1, 2, 3,…, 9 | 9 | 450 |
| **Nitrate**= 0, 1, 2,…,43 | 1, 6, 11,…, 41 | 9 | 1, 4, 7 | 3 | 27 |
| **Nitrate bis**= 0, 0.01, 0.02,…, 43 | 1, 2, 3,…, 41 | 41 | 1, 2, 3,…, 7 | 7 | 287 |
| **Silicate**= 0, 1, 2,…127 | 1, 20, 39,…, 126 | 7 | 1, 6, 11, 16 | 4 | 28 |
| **Silicate bis**= 0, 0.1, 0.2,…, 127 | 1, 4, 7,…, 126 | 42 | 1, 2, 3,…, 16 | 16 | 672 |
| **Phosphate**= 0, 0.1, 0.2,…, 3.8 | 0.1, 0.4, 3.5 | 3 | 0.1, 0.3, 0.5 | 3 | 27 |
| **Phosphate bis**= 0, 0.01, 0.02,…, 3.8 | 0.1, 0.2, 0.3,…, 3.5 | 35 | 0.1, 0.13, 0.16,…, 0.5 | 16 | 560 |
| **N/P**= 0, 0.2, 0.4,…, 25 | 0, 4, 8,…, 24 | 7 | 1, 2, 3, 4 | 4 | 28 |
| **N/P bis**= 0, 0.01, 0.02,…, 25 | 0, 1, 2,…, 24 | 25 | 1, 1.25, 1.5,…, 4 | 16 | 400 |

**Table S3. Statistics calculated for the different runs, using several combinations of environmental parameters.** For each combination of environmental parameters, the number of species, mean and maximum correlation values and probability values that result from the application of null models for both the Pearson correlation and MAE, with and without consideration of temporal autocorrelation, are reported. NM1 = probability values that result from the application of null models for the Pearson correlation without autocorrelation, NM2 = probability values that result from the application of null models for the Pearson correlation with autocorrelation, NM3 = probability values that result from the application of null models for the MAE without autocorrelation and NM4 = probability values that result from the application of null models for the MAE with autocorrelation.

| Run n° | Variable | Mean correlation | Species correlated | Highest correlations | NM1 | NM2 | NM3 | NM4 |
|---|---|---|---|---|---|---|---|---|
| 1 | T | 0.5156 | 0 | - | 2.8 | 83.8 | 0 | 0 |
| 2 | T bis | 0.5906 | 8 | 0.7676 | 98.3 | 100 | 0 | 0 |
| 3 | PARa | 0.5976 | 0 | - | 0 | 15.7 | 0 | 0 |
| 4 | PARa bis | 0.6527 | 0 | - | 59.4 | | 0 | 0 |
| 5 | PARb | 0.5506 | 0 | - | 0.1 | 100 | 0 | 0 |
| 6 | PARb bis | 0.6527 | 0 | - | 33.9 | 49.3 | 0 | 0 |
| 7 | PARc | 0.5757 | 0 | - | 0 | 100 | 0 | 0 |
| 8 | PARc bis | 0.6376 | 0 | - | 42.9 | 12.3 | 0 | 0 |
| 9 | N | 0.4488 | 0 | - | 3.1 | 66.6 | 24.5 | 46.5 |
| 10 | N bis | 0.4951 | 0 | - | 99.7 | 100 | 5.4 | 55.8 |
| 11 | S | 0.4224 | 1 | 0.6937 | 3.9 | 52 | 88.4 | 94.7 |
| 12 | S bis | 0.4447 | 0 | - | 100 | 100 | 100 | 100 |
| 13 | P | 0.3687 | 0 | - | 61.3 | 99 | 100 | 100 |
| 14 | P bis | 0.5392 | 1 | 0.6318 | 99.8 | 100 | 7.6 | 77.5 |

| Run n° | Variable | Mean correlation | Species correlated | Highest correlations | NM1 | NM2 | NM3 | NM4 |
|---|---|---|---|---|---|---|---|---|
| 15 | N/P | 0.5414 | 0 | - | 0.2 | 62.1 | 0 | 0 |
| 16 | N/P bis | 0.6270 | 0 | - | 25.7 | 100 | 0 | 0 |
| 17 | T and N | 0.6362 | 0 | - | 0.1 | 98.8 | 0 | 0 |
| 18 | T and PARa | 0.7833 | 1 | 0.8790 | 0 | 0 | 0 | 0 |
| 19 | T and PARb | 0.7838 | 1 | 0.8344 | 0 | 0 | 0 | 0 |
| 20 | T and PARc | 0.7791 | 1 | 0.7589 | 0 | 0 | 0 | 0 |
| 21 | N and PARa | 0.7224 | 0 | - | 0 | 0 | 0 | 0 |
| 22 | N and PARb | 0.7128 | 0 | - | 0 | 0 | 0 | 0 |
| 23 | N and PARc | 0.7085 | 0 | - | 0 | 0 | 0 | 0 |
| 24 | T and S | 0.6835 | 0 | - | 0 | 3.9 | 0 | 0 |
| 25 | N and S | 0.5993 | 0 | - | 0.1 | 77.7 | 0 | 0.7 |
| 26 | S and PARa | 0.6251 | 0 | - | 0.2 | 98.6 | 0 | 0 |
| 27 | S and PARb | 0.5933 | 0 | - | 3.2 | 99.9 | 0 | 0 |
| 28 | S and PARc | 0.6129 | 0 | - | 0.2 | 93.1 | 0 | 0 |
| 29 | T and P | 0.6776 | 0 | - | 0 | 28.9 | 0 | 0 |
| 30 | N and P | 0.5912 | 0 | - | 0.9 | 98.5 | 1.5 | 19.1 |
| 31 | S and P | 0.6127 | 0 | - | 0.1 | 66.7 | 0 | 0 |
| 32 | P and PARa | 0.7153 | 0 | - | 0 | 0 | 0 | 0 |
| 33 | P and PARb | 0.7073 | 0 | - | 0 | 0.1 | 0 | 0 |
| 34 | P and PARc | 0.7130 | 0 | - | 0 | 0 | 0 | 0 |
| 35 | T and N/P | 0.6308 | 0 | - | 8.4 | 100 | 0 | 0 |
| 36 | S and N/P | 0.6670 | 0 | - | 0 | 17.4 | 0 | 0 |
| 37 | N/P and PARa | 0.7751 | 0 | - | 0 | 0 | 0 | 0 |
| 38 | N/P and PARb | 0.7759 | 0 | - | 0 | 0 | 0 | 0 |
| 39 | N/P and PARc | 0.7690 | 0 | - | 0 | 0 | 0 | 0 |
| 40 | T, N and PARa | 0.7981 | 0 | - | 0 | 0 | 0 | 0 |
| 41 | T, N and PARb | 0.7975 | 1 | 0.8276 | 0 | 0 | 0 | 0 |
| 42 | T, N and PARc | 0.7934 | 0 | - | 0 | 0 | 0 | 0 |
| 43 | T, S and PARa | 0.7866 | 1 | 0.9613 | 0 | 0 | 0 | 0 |
| 44 | T, S and PARb | 0.7866 | 0 | - | 0 | 0 | 0 | 0 |

| Run n° | Variable | Mean correlation | Species correlated | Highest correlations | NM1 | NM2 | NM3 | NM4 |
|---|---|---|---|---|---|---|---|---|
| 45 | T, S and PARc | 0.7831 | 0 | - | 0 | 0 | 0 | 0 |
| 46 | T, P and PARa | 0.7925 | 3 | 0.9384 | 0 | 0 | 0 | 0 |
| 47 | T, P and PARb | 0.7931 | 0 | - | 0 | 0 | 0 | 0 |
| 48 | T, P and PARc | 0.7915 | 2 | 0.9162 | 0 | 0 | 0 | 0 |
| 49 | T, N/P and PARa | 0.8043 | 1 | 0.8913 | 0 | 0 | 0 | 0 |
| 50 | T, N/P and PARb | 0.8031 | 2 | 0.7807 | 0 | 0 | 0 | 0 |
| 51 | T, N/P and PARc | 0.8016 | 10 | 0.8644 | 0 | 0 | 0 | 0 |
| 52 | T, N and S | 0.7313 | 2 | 0.8896 | 0 | 7.9 | 0 | 0 |
| 53 | T, N and P | 0.7142 | 0 | - | 0 | 97.8 | 0 | 0 |
| 54 | T, S and P | 0.7253 | 2 | 0.9691 | 0 | 34.2 | 0 | 0 |
| 55 | T, S and N/P | 0.7452 | 4 | 0.8548 | 0 | 4.7 | 0 | 0 |
| 56 | N, S and PARa | 0.7261 | 0 | - | 0 | 24.3 | 0 | 0 |
| 57 | N, S and PARb | 0.7154 | 0 | - | 0 | 19.6 | 0 | 0 |
| 58 | N, S and PARc | 0.7154 | 0 | - | 0 | 21.5 | 0 | 0 |
| 59 | N, P and PARa | 0.7475 | 1 | 0.4239 | 0 | 0.1 | 0 | 0 |
| 60 | N, P and PARb | 0.7415 | 0 | - | 0 | 0.1 | 0 | 0 |
| 61 | N, P and PARc | 0.7374 | 0 | - | 0 | 0 | 0 | 0 |
| 62 | N, S and P | 0.6605 | 0 | - | 0.1 | 100 | 0 | 0 |
| 63 | S, P and PARa | 0.7189 | 0 | - | 0 | 53.8 | 0 | 0 |
| 64 | S, P and PARb | 0.7141 | 0 | - | 0 | 53.4 | 0 | 0 |
| 65 | S, P and PARc | 0.7180 | 0 | - | 0 | 14.8 | 0 | 0 |
| 66 | S, N/P and PARa | 0.7774 | 0 | - | 0 | 0 | 0 | 0 |
| 67 | S, N/P and PARb | 0.7767 | 0 | - | 0 | 0 | 0 | 0 |
| 68 | S, N/P and PARc | 0.7710 | 0 | - | 0 | 0 | 0 | 0 |

| Run n° | Variable | Mean correlation | Species correlated | Highest correlations | NM1 | NM2 | NM3 | NM4 |
|---|---|---|---|---|---|---|---|---|
| 69 | T, N, S and PARa | 0.8002 | 0 | - | 0 | 0 | 0 | 0 |
| 70 | T, N, S and PARb | 0.7988 | 2 | 0.8240 | 0 | 0 | 0 | 0 |
| 71 | T, N, S and PARc | 0.7966 | 1 | 0.9490 | 0 | 0 | 0 | 0 |
| 72 | T, N/P, PARa and S | 0.8062 | 6 | 0.8390 | 0 | 0 | 0 | 0 |
| 73 | T, N/P, PARb and S | 0.8045 | 2 | 0.8274 | 0 | 0 | 0 | 0 |
| 74 | T, N/P, PARc and S | 0.8030 | 3 | 0.7431 | 0 | 0 | 0 | 0 |
| 75 | T, N, P and PARa | 0.8008 | 3 | 0.8920 | 0 | 0 | 0 | 0 |
| 76 | T, N, P and PARb | 0.8008 | 1 | 0.7896 | 0 | 0 | 0 | 0 |
| 77 | T, N, P and PARc | 0.7980 | 3 | 0.8575 | 0 | 0 | 0 | 0 |
| 78 | S, N, P and PARa | 0.7497 | 0 | - | 0 | 90.8 | 0 | 0 |
| 79 | S, N, P and PARb | 0.7431 | 0 | - | 0 | 96 | 0 | 0 |
| 80 | S, N, P and PARc | 0.7397 | 0 | - | 0 | 97.1 | 0 | 0 |
| 81 | T, N, S and P | 0.7558 | 4 | 0.8849 | 0 | 76.5 | 0 | 0 |
| 82 | T, N, S, P and PARc | 0.7906 | 5 | 0.7922 | 0 | 0 | 0 | 0 |
| 83 | T, N, S, P and PARb | 0.7941 | 3 | 0.8741 | 0 | 0 | 0 | 0 |
| 84 | T, N, S, P and PARa | 0.7948 | 6 | 0.8378 | 0 | 0 | 0 | 0 |

## 1.2. Morphological traits, niche-environment interaction and temporal changes in diatoms

Loïck Kléparski, Grégory Beaugrand, Martin Edwards, François G. Schmitt, Richard R. Kirby, Elsa Breton, François Gevaert and Emeline Maniez

*Published in Progress in Oceanography. Supplementary Files are displayed in section 1.2.5. Article is available at* https://doi.org/10.1016/j.pocean.2022.102747

*Abstract*

Annual phytoplankton succession is a key ecological phenomenon that drives marine species' life cycles and energy flows within marine ecosystems. Identifying processes that control annual succession is critical to anticipate climate-induced environmental perturbations of this phenomenon and the consequences upon ecosystem functioning. Here, we demonstrate that diatoms in the North Sea undergo strong morphological changes throughout the year and that species with similar phenology possess comparable morphological traits (e.g. cell elongation) and ecological niches. The spring and autumn periods appear to be dominated by oblates (flattened cells) whereas the summer period is dominated by prolate (elongated cells). Elongation of the cell shape enhances buoyancy and confers a selective advantage in stratified low-nutrient waters typical of summer without changing a diatoms' surface area/volume ratio or its ability to absorb nutrients. Diatom shape thus appears to have evolved as a key adaptation to a specific environment and confers upon a species its specific niche and phenology, and therefore its place in the sequence of annual succession. As a result, shape influences temporal changes in the abundance of diatoms and their putative response to environmental forcing. We suggest that biogeochemical and earth-system models should include diatom cell shape as a parameter in order to improve model predictions and help our understanding of the consequences of climate change on marine ecosystems.

*Keywords*

Annual diatom succession, cell elongation, ecological niche, morphological traits, phenology, North Sea, CPR survey.

## 1.2.1. Introduction

Diatoms are a major phytoplankton clade that have evolved during the Middle Triassic when sea levels were rising and continental margins were flooding (Falkowski et al., 2004). Diatoms' subsequent increase in diversity during the mid-Cenozoic Era strongly influenced the global carbon cycle (Katz et al., 2005). Today, diatoms account for about 40% of total marine primary production and almost 40% of total particulate organic carbon exported to the deep ocean annually (Jin et al., 2006; Nelson et al., 1995; Tréguer et al., 2018). Annually, the abundance of diatoms rises and falls on a seasonal basis following a predictable pattern named "Annual Diatom Succession" (called hereafter ADS).

In the North Sea, ADS is characterised by the succession of various species throughout the year, with species such as *Skeletonema costatum* and *Thalassiothrix longissima* dominating in spring, *Guinardia striata* and *G. flaccida* in summer and *Poroboscia indica* and *Bidulphia alternans* in autumn (Caracciolo et al., 2021). ADS therefore determines the pulses of energy that influence the dynamics of the whole marine ecosystem, the life cycle of many zooplankton and fish being coupled with the peaks in primary production or species dominance (Cushing, 1990; Platt et al., 2003). Recently, anthropogenic climate change has begun to alter diatom phenology and biogeography, and this is likely to have strong consequences on trophic interaction and ecosystem functioning (Chivers et al., 2017; Edwards and Richardson, 2004).

To predict the future consequences of climate change on marine ecosystems, we need to i) better understand how diatom communities naturally form and reorganise at multiple time-scales and ii) to identify the key elements and processes that control phenology. Morphological traits (i.e. cell size and shape) are known to play an important role in diatom succession because they influence resource uptake, buoyancy and predation, which are all properties that may explain seasonal changes in species dominance (Behrenfeld, 2010; Caracciolo et al., 2021; Karp-Boss and Boss, 2016; Margalef, 1978; Naselli-Flores et al., 2021; Naselli-Flores and Barone, 2011). The niche of a species (i.e. the set of environmental parameters that enable a species to grow, maintain and reproduce) is also of paramount importance as it controls the range of environmental conditions that a species can tolerate (Hutchinson, 1957). Here, we investigate the relationships between species' morphological traits (size and shape), ecological niches, and phenology with the aim of identifying the key elements and processes responsible for ADS.

## 1.2.2. Materials and Methods

### 1.2.2.1. Biological data

We used abundance data of 45 diatom species/taxa commonly sampled by the Continuous Plankton Recorder (CPR) survey in the North Sea (51°N, 60°N, -3°E, 9°E; Figure S1) between 1958 and 2017 (Reid et al., 2003). This long-term plankton monitoring programme has collected plankton on a monthly basis in the extratropical regions of the North Atlantic Ocean and its adjacent seas since 1946. The sampling machine is a high-speed plankton recorder towed behind voluntary merchant ships at

a depth of approximately 7 m (Beaugrand et al., 2003; Reid et al., 2003) (see Text S1 for an overview of the CPR limitations and Text S2 for CPR data availability)(Beaugrand et al., 2003; Reid et al., 2003). Each value of abundance corresponds to a number of cells per CPR sample, which corresponds on average to 3 $m^3$ of seawater filtered (Jonas et al., 2004). We calculated a daily climatology of the abundance of 45 diatoms sampled by the CPR survey. To account for all CPR samples collected between 1958 and 2017, we based our climatology on 366 days. We applied twice a 6-order simple moving average on the matrix (366 days x 45 species/taxa) and standardized the abundance data between 0 and 1 for each species, following the procedure described in Caracciolo and colleagues (Caracciolo et al., 2021).

#### 1.2.2.2. Environmental data

Text S2 provides key references and internet sites for each environmental dataset we used. Sea Surface Temperature (SST; °C) originated from the NOAA OI SST V2 high resolution dataset provided by the Earth System Research Laboratory, Physical Science Division, Boulder Colorado, USA. Mixed Layer Depth (MLD; m) was provided by the Copernicus Marine Environment Monitoring service (CMEMS) and originated from the Global Ocean Physical Reanalysis product (GLOBAL_REANALYSIS_ PHY_001_030). We used the MLD as a proxy of turbulence level (the deeper the MLD is, the more active the wind-induced turbulence is likely to be) and stratification (a lower MLD is in general associated with a higher stratification, which prevents nutrients recycling to the surface).

Phosphate, silicate and nitrate concentrations ($mmol.m^{-3}$) were provided by the CMEMS and came from the Global Ocean Biogeochemistry Hindcast (GLOBAL_REANALYSIS_BIO_001_029). Photosynthetically Active Radiation (PAR; mole /$m^2$/day) was provided by Globcolour Hermes. PAR was estimated from satellites data. Euphotic depth of sea water (m) was provided by the NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Euphotic depth was calculated with Lee algorithm. Bathymetry (m) was provided by the British Oceanographic Data Centre (BODC). Distance to nearest coastlines (km) was provided by the Nasa's Ocean Biology Processing Group, Goddard Space Flight Center.

We calculated a daily climatology for each environmental parameter in the North Sea (Figure 16b-h and Figures S1 and S2a-c), based on the time period 1997-2017. The climatology of the euphotic depth of sea water was based on the time period 2002-2017. Here also, we based our climatologies on 366 days to match the diatom climatologies. We assumed that the difference of time periods between biological (1958-2017) and environmental (1997-2017) datasets did not alter our conclusions because the seasonal changes observed here, are much more important than decadal changes (see Figure 18). The use of the full time period for CPR data was important to have a more reliable daily climatology and we know that daily variability is much higher than decadal variability, an assumption that has been checked for copepod biodiversity assessed from the CPR data (Beaugrand et al., 2003).

The kinematic viscosity (in $m^2.s^{-1}$) was assessed from SST daily climatology using equations from Pilson (Pilson, 2013) and assuming a salinity of 35 (i.e. mean salinity of the North Sea).

Data on monthly long-term changes in SST (°C), mean surface downward short-wave radiation flux ($W.m^{-2}$), and northward and eastward wind components ($m.s^{-1}$) originated from the ERA5 dataset. This monthly gridded dataset, which covered the period 1958-2017, were provided by the Copernicus climate change service. Data on mean surface downward short-wave radiation flux (solar radiations) were used as a proxy of PAR.

### 1.2.2.3. Morphological traits

#### *1.2.2.3.1. Cell shape*

We gathered together information on the geometrical shape of each species from the Global Diatoms Database (Leblanc et al., 2012). Six different shapes were distinguished: cylinder, cylinder + 2 half spheres, rectangular box, prism on an elliptic base, prism on a parallelogram base and prism on a triangular base (Table S1).

#### *1.2.2.3.2. Oblates and prolates*

We also retrieved information on minimum and maximum cell dimensions (diameter/height or length/width/height depending upon cell shapes). For cylindrical cells, we defined height as the dimension perpendicular to cell diameter. For other geometrical shapes, we first converted them to a cylinder of equivalent size: we calculated the area of the slice perpendicular to cell height (as defined in the database) and estimated the diameter of the equivalent disk (i.e. the disk of equivalent surface). Based on these cell dimensions, we divided diatoms into three groups: (i) oblates (i.e. minimum height < minimum diameter and maximum height < maximum diameter), (ii) prolates (i.e. minimum height > minimum diameter and maximum height > maximum diameter) and (iii) oblates/prolates for the remaining species/taxa, meaning that they can be both oblates and prolates (Table S2).

#### *1.2.2.3.3. Other morphological indices*

Based on mean cell height ($h$) and diameter ($d$) (Table S2), we estimated the area ($s$), the volume ($v$) and the area/volume ratio ($sv^{-1}$) of each species by assimilating their shape to a cylinder:

$$s_{cylinder} = \pi d \left(\frac{d}{2} + h\right) \quad (1)$$

$$v_{cylinder} = \frac{\pi}{4} d^2 h \quad (2)$$

We assessed an index of cell morphology, which had the advantage to be dimensionless:

$$\delta = msv^{-1} \quad (3)$$

with $m$ the cell maximum linear dimension (i.e. the mean cell diameter or height for oblates and prolates, respectively). $\delta$ describes cell attenuation (i.e. reduction of one of the dimensions of the cell volume or cell flattening) and its departure from a sphere (Naselli-Flores et al., 2021).

We estimated the maximum area of light interception normalized by volume (β). For oblates, we calculated the ratio between disk area and volume as follows:

$$\beta_{oblates} = \frac{\pi\left(\frac{d}{2}\right)^2}{v} \quad (4)$$

For prolates, we calculated the ratio between rectangle area (projection of a cylinder on a surface) and volume:

$$\beta_{prolates} = \frac{dh}{v} \quad (5)$$

### 1.2.2.4. Predation

We used abundance data of 74 small (≤ 2 mm) and 196 large (> 2 mm) copepods species/taxa collected by the CPR survey in the North Sea (Figure S1). We summed the total abundance for small/large copepods in each CPR sample between 1958-2017 and then calculated a daily abundance climatology for small/large copepods. We therefore obtained two vectors of 366 days (one vector for small and another for large copepods) with the daily mean abundance for small/large copepods. As above, we applied twice a simple 6-order moving average and standardized the abundance data between 0 and 1 (Figure S2c).

### 1.2.2.5. Phylogenetical classification

Information on the phylogenetical classification of the 45 diatoms (family, order, superorder, subclasses) was retrieved from the WOrld Register of Marine Species (WORMS) (http://www.marinespecies.org/index.php). (Figure S3 and Table S3).

### 1.2.2.6. Analyses

#### *1.2.2.6.1. Principal Component Analysis (PCA)*

We characterized Annual Diatom Succession (ADS) by applying a standardized PCA on a matrix of daily abundance climatology (366 days) x diatom species (45 species) (Figure 15a). We examined the first three axes (principal components and eigenvectors) of the PCA because a test based on a broken-stick distribution (Legendre and Legendre, 1998) revealed that the first three axes of the PCA were significant, i.e. they were above average eigenvalues calculated by the broken-stick model (Table S4). Seasonal changes in diatom abundance were examined by the first three principal components (PCs) and relationships between species were investigated by using the first three normalized eigenvectors, which represented the correlation of the 45

diatoms with the first three PCs. We tested the relationships between the first three principal components and the daily climatology for SST, euphotic depth, PAR, MLD, mean small/large copepods abundance, viscosity, silicate, phosphate and nitrate concentration by means of a Spearman rank correlation coefficient. The significance of the correlation coefficients was tested by applying a Montecarlo test with 100,000 simulations (Figure 17 and Table S5).

#### *1.2.2.6.2. Clustering*

We applied a cluster analysis based on the first three normalized eigenvectors of the PCA. First, we calculated a squared matrix of Euclidean distance using the first three eigenvectors. Then we applied a Ward linkage algorithm (Legendre and Legendre, 1998). The resulting dendrogram was cut using a threshold of Euclidean distance of 8; this threshold was chosen because it enabled the identification of three species groups that closely matched the seasonal patterns exhibited by the first three PCs (Figure 15a, Figure 17 and Figures S2d-f, S4 and S5).

#### *1.2.2.6.3. Relationships between phenology and morphological traits*

We examined the relationships between the species groups and morphological traits in the space defined by the first three eigenvectors: (i) geometrical shapes (Figure 15b), (ii) oblates/prolates (Figure 15c), (iii) height-diameter difference (Figure 15d), (iv) area/volume ratio (Figure 15e), (v) δ (Figure 15f), (vi) mean cell diameter (Figure S6a) and (vii) mean cell height (Figure S6b).

We tested the differences in the mean of the morphological indices (height-diameter (Table S6), mean cell diameter (Table S7), mean cell height (Table S8), area/volume ratio (Table S9) and δ (Table S10)) among species groups by means of a Kruskal-Wallis test and post-hoc tests.

#### *1.2.2.6.4. Comparison between morphological traits based on cylindrical and spheroidal shapes*

In our analyses, we assimilated diatom shapes to cylinders. To test whether our conclusions were affected by this assumption, we also compared our results with those obtained by assimilating diatom shapes to spheroids. Based on mean cell height ($h$) and diameter ($d$) (Table S2), we re-estimated the area ($s$), the volume ($v$) and the area/volume ratio ($sv^{-1}$) of each species by assimilating their shape to spheroids as follows:

$$s_{spheroid} = \frac{\pi d}{2}\left(d + \frac{h^2}{\sqrt{h^2-d^2}}\sin^{-1}\frac{\sqrt{h^2-d^2}}{d}\right) \qquad (6)$$

$$v_{spheroid} = \frac{\pi}{6}d^2h \qquad (7)$$

We examined the relationships between morphological traits (area, volume, area/volume ratio and δ) when cells were based on cylinders (see equations 1-3) and spheroids by means of Pearson correlation coefficients (Figure S7a-d).

We also investigated the relationships between species groups and morphological traits estimated with spheroidal shape in the space defined by the first three eigenvectors using (i) area/volume ratio and (ii) δ (Figure S7e-h).

Finally, we tested the differences in the mean of the morphological indices among species groups, based on spheroidal shapes for area/volume ratio (Table S11) and δ (Table S12), by means of a Kruskal-Wallis test and post-hoc tests.

#### *1.2.2.6.5. Relationships between morphological indices and the environment*

We also examined the relationships between $m$ and area/volume ratio (Figure 15g-h), PAR and β (Figure S8), MLD and $\delta$ (Figure S9), kinematic viscosity and $\delta$ (Figure S10) and between copepod abundance and height-diameter difference (Figure S11).

We calculated the Spearman rank correlation coefficient between $m$ and area/volume ratio (Figure 15g), PAR and β (Figure S8), MLD and $\delta$ (Figure S9), kinematic viscosity and $\delta$ (Figure S10) and between copepod abundance and height-diameter difference (Figure S11). The correlations were tested by means of a Montecarlo test using 100,000 permutations (Jackson and Somers, 1989).

#### *1.2.2.6.6. Phenology and niche*

We estimated the optimum and amplitude (breadth) of the niche using nine ecological dimensions: SST, euphotic depth, PAR, MLD, silicate phosphate and nitrate concentration, bathymetry and distance to nearest coastlines. To do so, we estimated the weighted mean as a proxy for optimum ($X_{opt}$) and the weighted standard deviation as a proxy for the amplitude ($\sigma$) (Figure 16 and Figure S12). These proxies were calculated for each ecological dimension:

$$X_{opt} = \frac{\sum_{i=1}^{t} w_i x_i}{\sum_{i=1}^{t} w_i} \quad (8)$$

$$\sigma = \frac{\sum_{i=1}^{t} w_i \left(x_i - x_{opt}\right)^2}{\frac{t-1}{t}\sum_{i=1}^{t} w_i} \quad (9)$$

where $w$ is the abundance of the diatom species, $x$ the value of the environmental parameter or copepod abundance and $t$ the number of days in the climatology (for SST, euphotic depth, PAR, MLD, silicate phosphate and nitrate concentration, and small/large copepod abundance) or the total number of CPR samples when bathymetry or distance to nearest coastlines was used.

We examined the relationships between diatom niche optimum/amplitude, their day of maximum abundance and the daily climatology for SST, euphotic depth, PAR, MLD, silicate phosphate and nitrate concentration (Figure 16b-h). We also investigated the relationships between diatom niche optimum/amplitude and height-diameter difference (Figure S12).

We tested the niche separation between oblates and prolates by means of a Krukal-Wallis test performed on niche optima and amplitude (Tables S13-S14). The relationships between species niche optima/amplitudes and species height-diameter

difference was examined by applying a Spearman rank correlation coefficient and tested with a Montecarlo test using 100,000 simulations (Figure S12)(Jackson and Somers, 1989).

#### *1.2.2.6.7. Long-term changes in the abundance of oblates and prolates*

We investigated long-term monthly changes in the mean abundance of oblates and prolates collected by the CPR survey between 1958 and 2017 (Figure 18a). Oblate abundance in April and October and prolate abundance in August were smoothed by means of a second-order simple moving average (Figure 18b-d). The relationships between species abundance (in April and October for oblates and August for prolates) and the long-term changes in SST, wind and solar radiations were examined by applying a Spearman rank correlation coefficient and tested with a Montecarlo test using 100,000 simulations (Table S15)(Jackson and Somers, 1989).

## 1.2.3. Results and Discussion

We first characterised ADS by means of a Principal Component Analysis (PCA) performed on a matrix of daily abundance climatology (366 days) x diatom species (45 species; Methods). The abundance matrix originated from data collected by the Continuous Plankton Recorder (CPR) survey between 1958 and 2017 in the North Sea (Figure S1). We retained the first three Principal Components (PCs), which were significant after applying a broken-stick test (Methods and Table S4). The first PC (26% of the total variance) showed a strong peak in spring at a time of weak stratification and low temperature, reflecting the spring bloom, and a smaller peak in late autumn at a time of decreasing temperature and stratification breakdown (Caracciolo et al., 2021) (Figure 15a). The second PC (25%) exhibited a similar pattern with two peaks, the lowest in spring and the highest in autumn, again reflecting both spring and autumn blooms. The third PC (15%) showed high positive values from late spring to the end of summer, identifying diatoms that were dominant during the stratified low-nutrients period at a time of higher zooplankton abundance (Figure S2a-c). We therefore identify three key periods in ADS, confirming a pattern seen in temperate seas (Caracciolo et al., 2021; Colebrook, 1986, 1984).

We examined species phenology by studying the first three normalised eigenvectors, which represent the correlation between each species and the first three PCs (Figure 15b-f). A cluster analysis, performed on the first three eigenvectors, allowed us to identify three assemblages (Figures S2d-f and S4), each belonging to one of the phenological patterns we identified previously (Figure 15b-f): (i) a spring-dominated group related to PC1, (ii) an autumn-dominated group related to PC2 and (iii) a Late-Spring Summer Autumn (LSSA) group related to PC3 (Figures S2, S4-S5).

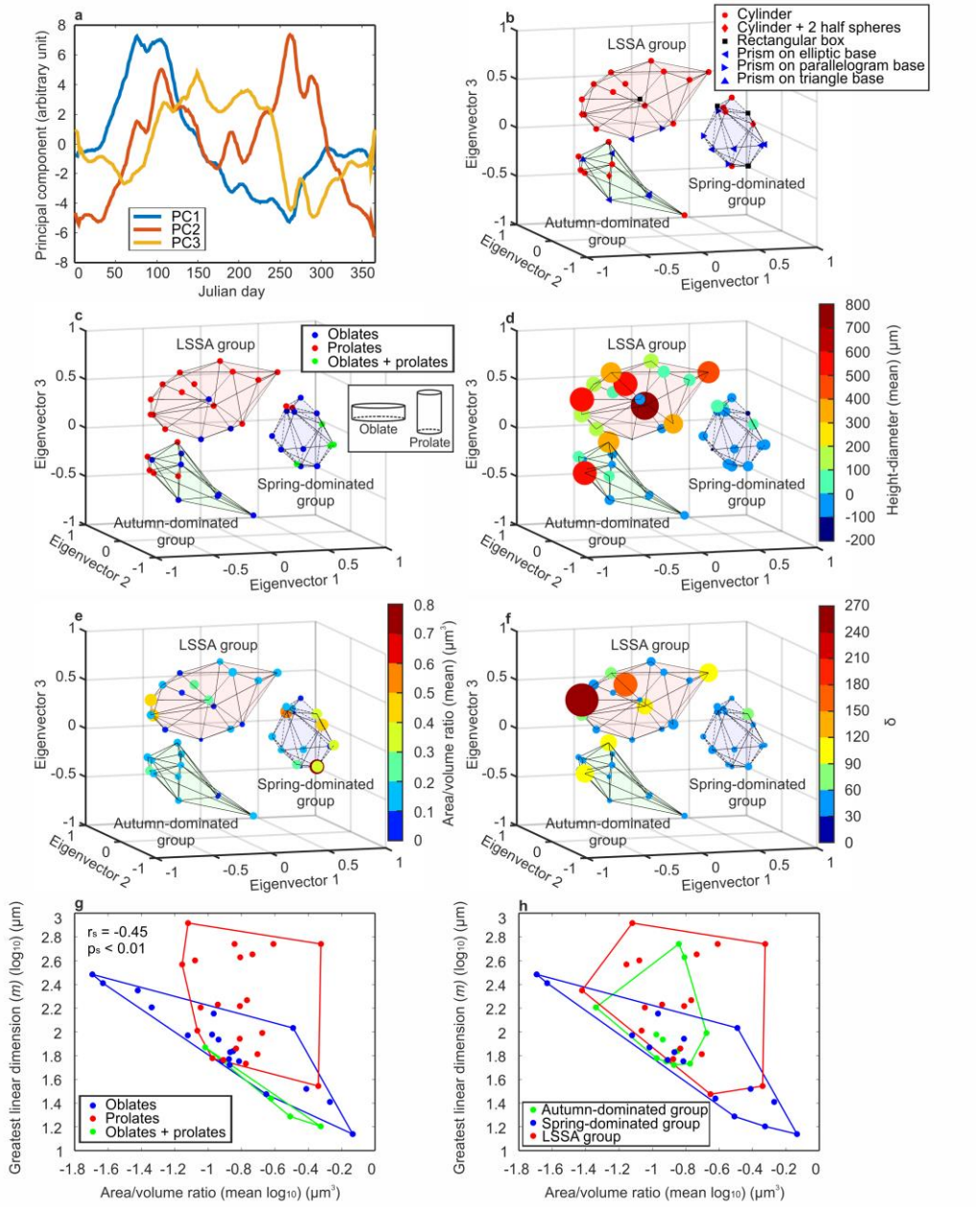



*Figure 15. Annual Diatom Succession (ADS) and morphological traits. (a) ADS characterized by a PCA applied on a species abundance matrix 366 days x 45 species. The first three principal components (PCs) are shown. (b-f) Relationships between the first three eigenvectors and morphological traits. Relationships between species groups and (b) morphological traits, (c) oblates and prolates, (d) height-diameter differences, (e) area/volume ratio and (f) δ. Blue area delineates the space occupied by the spring-dominated group, green area the space occupied by the autumn-dominated group and red area the space occupied by the LSSA group. (d-f) Size and colour of the bullets are scaled with (d) mean cell height-diameter difference, (e) mean area/volume ratio and (f) δ. (g-h) Relationships between cell greatest linear dimension (m)(log10) and mean area/volume ratio (log10). Spearman rank correlation coefficient (rs) and its associated probability (ps) are displayed on the top left.*

We next studied relationships between species phenology and morphological traits. Cell shape and size are major traits affecting nutrient acquisition, light capture, grazing resistance, and vertical motion (Clifton et al., 2018; Karp-Boss and Boss, 2016; Litchman and Klausmeier, 2008; Naselli-Flores et al., 2021; Naselli-Flores and Barone, 2011; Padisak et al., 2003). To do so, we compiled data on cell size and geometrical

shape from the Global Diatom Database (Leblanc et al., 2012) for 45 species/taxa (Methods). Spring-dominated and autumn-dominated groups comprised a mixture of species of different shapes (cylindrical, rectangular box and prism) whereas the LSSA group mainly comprised cylindrical diatoms (Figure 15b and Table S1). By assuming all diatom shapes to be cylindrical (note that we also performed the same analyses by assimilating diatom shapes to spheroids, see below), we separated diatoms as oblates (i.e., cell diameter > cell height, flattened cell) and prolates (cell diameter < cell height, elongated cell). We identified 20 oblates and 21 prolates, the remaining being identified as oblates/prolates (i.e. species/taxa that can be both during their life cycle; Methods and Table S2). The spring- and autumn-dominated groups mainly comprised oblates (62.5% and 58% in the spring- and autumn-dominated group, respectively; Figure 15c and Table S1) and the LSSA group mostly comprised prolates (82%; Figure 15c and Table S1).

Following this result, we investigated the variations of the mean cell height-diameter difference between groups more deeply (Figure 15d). Highest and lowest positive differences characterized the LSSA group and the spring-dominated group, respectively. Values were more variable from the autumn-dominated group. A Kruskal-Wallis test revealed significant differences and post-hoc tests showed that differences occurred between LSSA and spring/autumn-dominated groups, but not between spring- and autumn-dominated groups (Table S6). Interestingly, negative differences (shown in blue on the colour bar) were smaller than positive ones (from green to red; Figure 15d), suggesting that cell height rather than diameter is modified during ADS; this result is confirmed by insignificant differences in cell diameter and significant differences in cell height among groups (Figure S6 and Tables S7-S8). Previous studies have suggested that the size of phytoplankton is rarely greater than 300 µm (Naselli-Flores and Barone, 2011) because of fundamental structural constraints imposed by natural selection, i.e. limited minimum thickness and maximum diameter. For example, to avoid turbulence-induced damage and to benefit from the viscosity of water, planktonic cells have to remain smaller than the Kolmogorov scale (i.e. micro eddies) (Naselli-Flores et al., 2021; Reynolds, 1998) and the area/volume ratio needs to remain relatively low to ensure chemical exchanges between the cell and its environment (Raven et al., 2005).

Phytoplankton exhibit different cell shapes that modify their area/volume ratio to optimize nutrient exchange and buoyancy with their surrounding environment (Karp-Boss and Boss, 2016; Pahlow et al., 1997). Energy flux occurs through the cell surface while metabolism is related to the cell volume, which explains why the ratio between surface area and volume should be high, although the building of a large surface has also a negative metabolic cost (Brown et al., 2004). We did not find any significant seasonal changes in the ratio between surface area and volume amongst our three groups (Figure 15e and Table S9). Similar results have been documented in tropical lakes and coastal marine waters where the area/volume ratio was preserved through morphological changes (Lewis, 1976; Stanca et al., 2013). We tested whether elongation (i.e. an increase in the maximum linear dimension or $m$) could be a more relevant parameter. To do so, we used an index of cell morphology $\delta$ (i.e. $m$ times the

area/volume ratio; Methods)(Naselli-Flores et al., 2021). In contrast to the area/volume ratio (Figure 15e), we found significant differences in cell morphology (δ) among the three groups (Figure 15f and Table S10). Assimilating diatom shapes to spheroids led to similar conclusions (Figure S7 and Tables S11-S12). A closer look on $m$ versus $sv^{-1}$ allowed a better understanding of why there was no significant change in the area/volume ratio (Figure 15g-h). The two figures suggest that the dominance of prolates in the LSSA group during the stratified low-nutrients period does not affect the area/volume ratio in contrast to oblates where changes in size do influence the area/volume ratio. These results suggest that cell elongation is a way of altering shape without changing the area/volume ratio that would otherwise affect nutrient exchange (Karp-Boss and Boss, 2016; Pahlow et al., 1997).

Further, we examined whether phylogeny (from family to subclass level) was responsible for the observed differences in traits. We found no evidence for a phylogenetic effect at the family, order and superorder levels. This was not so at the subclass level; although the LSSA and autumn-dominated groups were mainly composed of *Coscinodiscophycidae,* the spring-dominated group was characterized by a diversity of subclasses (Figure S3 and Table S3). This analysis, therefore, shows that differences in morphological traits are also influenced by phylogeny. Other studies have suggested that major morphological traits that drive strategies in nutrient uptake are conserved over evolutionary time at a high phylogenetic level, a phenomenon known as trait conservatism (Litchman et al., 2007). Trait conservatism results from stabilizing selection and niche conservatism, i.e. species keeping the shape of their niche unaltered over time because of their propensity to remain in places where the environment is highly suitable for them (Ackerly, 2003). (Not all phytoplankton traits exhibit the same level of conservatism, some being more labile than others. (Litchman et al., 2010; Schwaderer et al., 2011)).

Three other hypotheses could also explain the morphological changes we observed: (i) light capture optimization (Naselli-Flores and Barone, 2011), (ii) buoyancy (Naselli-Flores et al., 2021; Naselli-Flores and Barone, 2011; Padisak et al., 2003) and (iii) response to predation (Smetacek, 2001). We investigated relationships between maximum area of light interception normalized by volume (β) and Photosynthetically Active Radiation (PAR) (Methods, Figure S8). We found no clear relationship between PAR and β. Noticeably, β diminished when maximum PAR increased, which is in agreement with physiology (Reynolds, 1989). However, small values of β were observed for a large range of PARs, which suggests that changes in β are not modulated by PAR for most species considered here. We acknowledge that physiological processes by which the cell can adapt to different light intensity are complex, however (e.g. photosynthetic pigments, see Text S3)(Reynolds, 1989).

Many other studies have provided evidence that variations in cell shape might result from an adaptation to turbulence through a modification in sinking velocity (Clifton et al., 2018; Margalef, 1978; Naselli-Flores et al., 2021; Naselli-Flores and Barone, 2011; Padisak et al., 2003). We tested this hypothesis by investigating the relationship between Mixed Layer Depth (MLD, used as a proxy for wind-induced turbulence intensity) and $\delta$ (Figure S9, Methods). This analysis revealed a negative

relationship between maximum MLD values and $\delta$. While oblates were found for a large range of MLDs, prolates dominated low-turbulence waters. We therefore concluded that diatom elongation observed during ADS may result mainly from an adaptation to buoyancy, although resource acquisition and predation probably also play a role (Karp-Boss and Boss, 2016; Pahlow et al., 1997; Sommer, 1998).

Studies have also found negative relationships between temperature and phytoplankton sinking velocity through an alteration in viscosity (Smayda, 1970; Sournia, 1982). From Stokes' equation, and considering a cell of 20 µm of diameter, sinking velocity increases of 4% for each increase of 1°C because of a decrease in viscosity (Smayda, 1970). However, sinking velocity depends on cell shapes (Padisak et al., 2003; Smayda, 1970; Sournia, 1982). We therefore tested the relationship between kinematic viscosity and δ and found a weak but significant negative relationship (Figure S10; Methods). This result suggests that prolates may also occur in stratified warm waters because their shapes slow down sinking velocity (Padisak et al., 2003). We acknowledge that we have not considered all processes that can impact phytoplankton buoyancy. First, many species produce organic substances that increase the viscosity of the surrounding environment and that may then trigger cell aggregation, sometimes associated with minerals and particles, which may ultimately increase sinking velocity (Smayda, 1970; Smetacek, 1985). Second, phytoplankton cells can also develop protuberances such as spines and setae that directly increase the frictional drag or help turn the longest cell dimension perpendicular to the settling axis (Padisak et al., 2003; Smayda, 1970). Third, some species form chains that may affect sinking speed, although contrasting results have been observed for different species with either an increase or a decrease in buoyancy (Padisak et al., 2003; Smayda, 1970). Lastly, we have not considered intra-specific variations in size and seasonal morphological plasticity, which can be important in many species (Ligowski et al., 2012).

Lastly, we investigated whether there is a relationship between diatom shape and predation. We examined the relationship between the abundance of small (≤2mm)/large(>2mm) copepods and morphological shape (Methods, Figure S11). Although correlation cannot imply causation, we found a nonlinear positive correlation between zooplankton abundance for small and large copepods and height-diameter difference. This result is in agreement with works that have suggested that changes in diatom shape may be influenced by zooplankton predation (Karp-Boss and Boss, 2016; Naselli-Flores and Barone, 2011; Smetacek, 2001), although other works have shown that elongated cells can be ingested by copepods when turbulence is low (Visser and Jonsson, 2000). We caution that defence mechanisms in diatoms are diverse, ranging from physiological (allelopathy) to phenomena such as colony formation and escape response (Pančić and Kiørboe, 2018) and that all mechanisms involved could not be realistically tested in this study.

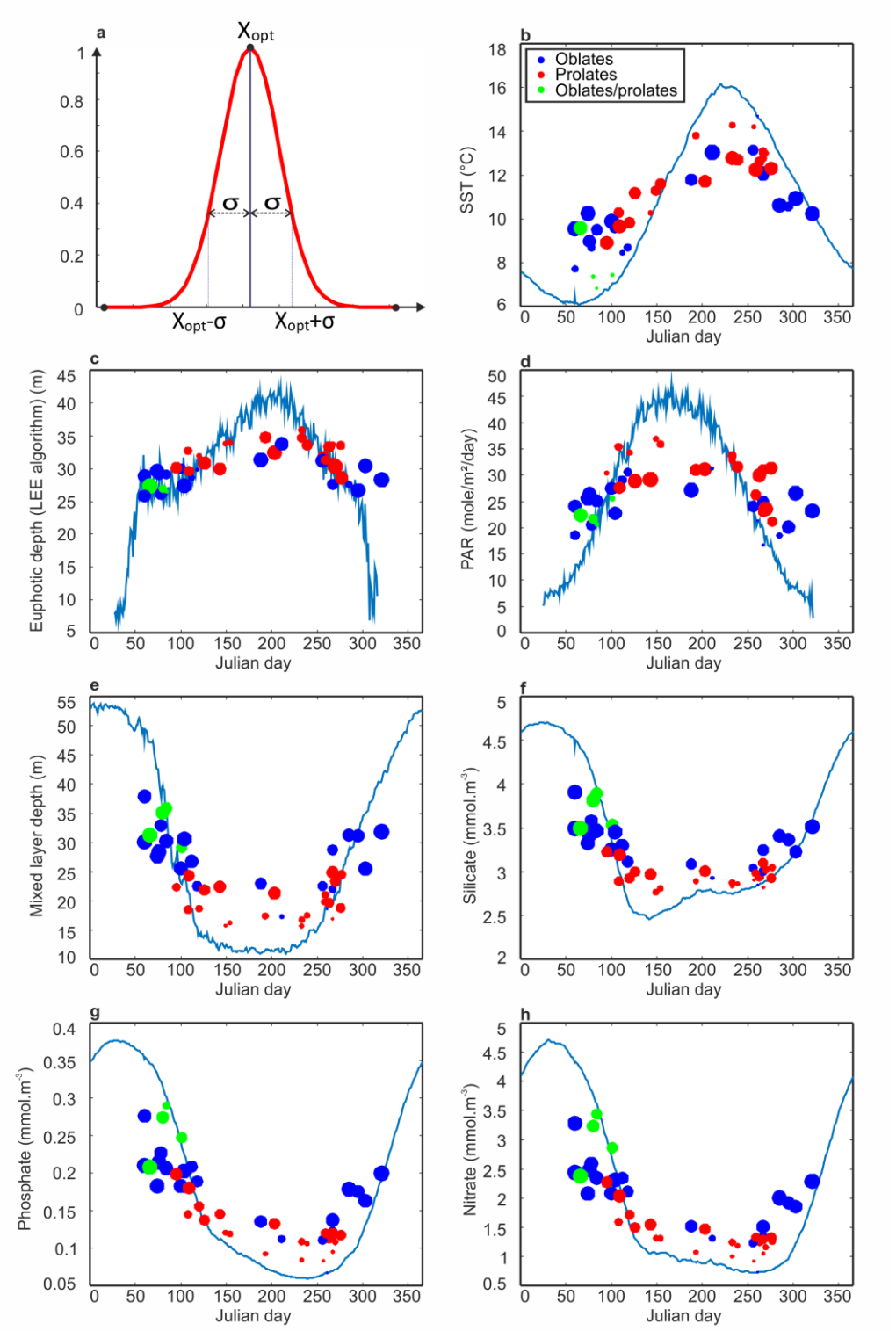



*Figure 16. Relationships between day of maximum abundance and ecological optima/amplitude of each diatom and environmental parameter. (a) Theoretical ecological niche with Xopt representing species niche optimum and σ denoting niche amplitude assessed by the weighted standard deviation. The Y axis corresponds to the species abundance and the X axis to an hypothetical environmental parameter. Relationships between day of maximum abundance and niche optimum/amplitude for (b) SST, (c) euphotic depth, (d) PAR, (e) mixed layer depth (MLD), (f) silicate, (g) phosphate and (h) nitrate concentration. Blue lines correspond to the North Sea daily climatology of the parameter considered in the panel. The niche optimum can be read on the ordinate axis. The size of the bullet is a function of species' rank of the weighted standard deviation; larger bullets are associated with larger niche amplitude and therefore ranks. Oblates are in blue, prolates in red and oblates/prolates in green.*

Recent works have suggested that annual plankton succession results from the niche-environment interaction (Caracciolo et al., 2021). We therefore explored the relationships between cell shape and ecological niche for the 45 species/taxa. Here, the ecological niche was characterised by its optimum and amplitude (i.e. niche breadth) (Hutchinson, 1957); these niche parameters were assessed for each species/taxa and nine environmental parameters (Methods; Figure 16a). Our results suggest that morphological traits influence species' niche of diatoms and that oblates and prolates had significant differences in ecological optima for all niche dimensions. Amplitudes were also distinct for four niche dimensions: nutrients (nitrate, phosphate and silicate) and MLD (Figure 16b-h, Tables S13-S14). Although the relationship was less obvious for sea surface temperature and euphotic depth (Figure 16b-c), optima and amplitudes plotted against daily climatology of each parameter showed a clear dominance of prolates during low-nutrients, turbulence and PAR conditions (Figure 16d-h). In addition, oblates were significantly more euryoecious (i.e. larger niche amplitude) than prolates with respect to nutrients and MLD (Figure 16e-h, Figure S12 and Table S14). Prolates are therefore K-strategists (i.e. equilibrium species), adapted to stable environment (i.e. predictable environment with low perturbations) with respect to nutrients, turbulence and stratification (low MLD) whereas oblates are more r-strategists (i.e. opportunist species), adapted to a more unstable environment (i.e. unpredictable) (Southwood et al., 1974). In other words, prolates (K-strategist) thrive in stratified water with low turbulence and constant low-nutrient concentration whereas oblates (r-strategists) thrive in turbulent waters in the absence of stratification with more variable levels of nutrients concentration (Figure 16). Because oblates and prolates have distinct niches, this analysis demonstrates that morphological traits influence species niche and phenology. We summarize the relationships between diatom morphological traits, phenology and the environment in Figure 17 (see also Table S5).

Finally, we investigated long-term monthly changes in the abundance of oblates and prolates between 1958 and 2017. Our results show that both groups exhibit distinct patterns of temporal changes (Figure 18a), with oblates having a constant and high level of abundance in April (Figure 18b) but a more variable abundance in October with blooms of decreasing intensity during the 1960s and 1970s, followed by a strong increase after *circa* 1980 (Figure 18c). Similar patterns are also observed for prolates in August but with an abundance level below the maximum reached in 1960 (Figure 18d). The period of low abundance found for oblates and prolates between 1978 and 1982 was probably the consequence of a cold hydro-meteorological anomaly observed in the North Sea (Edwards et al., 2002). We subsequently investigated whether these long-term changes in abundance were correlated with long-term changes in Sea Surface Temperature (SST), wind and solar radiations (Table S15). We found a nonlinear positive correlation between oblate abundance in October and SST and a negative correlation with eastward wind as well as wind intensity. A negative correlation was also found between oblate abundance in April and eastward wind. No significant correlations with northward wind and solar radiations were found. In the Northeast Atlantic region, long-term changes in SST/wind

variations have led to a marked increase in diatoms, associated with a diminution in dinoflagellates (Hinder et al., 2012).

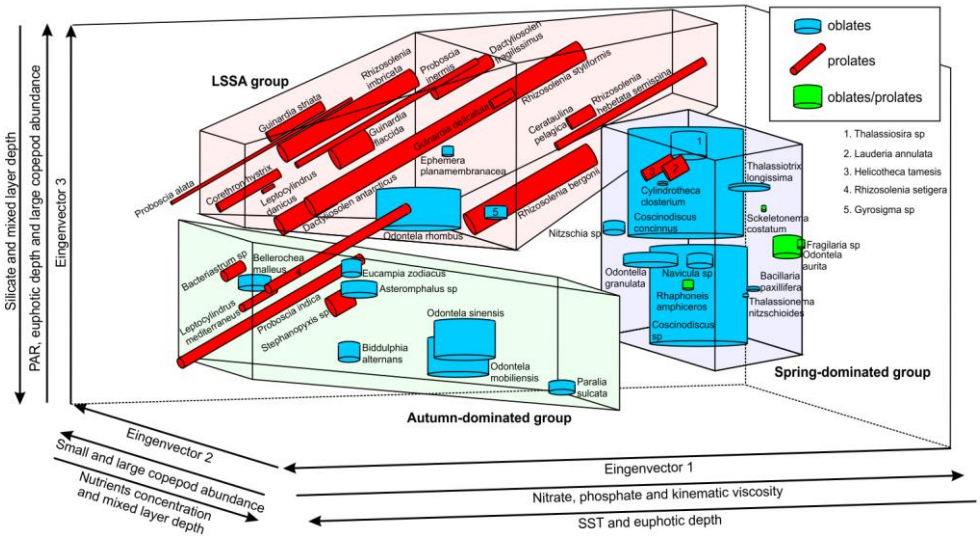



*Figure 17. Summary of the relationships between the first three eigenvectors, diatom shapes and the environment. Cell shapes are represented by cylinders scaled with species height and diameter (Table S2). Oblates are in blue, prolates in red and oblates/prolates in green. Species names are those recorded in the CPR dataset. Blue area delineates the space occupied by the spring-dominated group, green area the space occupied by the autumn-dominated group and red area the space occupied by the LSSA group. Correlation between the first three principal components and the environmental parameters are displayed in Table S5. More information on the principal components can be found in Figure 15a.*

Because oblates and prolates exhibit specific long-term changes, our results suggest they may respond to different environmental forcing. However, it is widely assumed that diatom contribution to primary production and carbon export will decline in the mid and low latitudes because ocean warming will increase stratification and decrease nutrients supply to the surface ocean (Bopp et al., 2013, 2005; Dutkiewicz et al., 2013; Fu et al., 2016). Such predictions are made because of an oversimplification of Margalef's mandala (Margalef, 1978) assimilating diatoms to a single Plankton Functional Type (PFT) that thrives when nutrients and turbulence are high (Fu et al., 2016; Kemp and Villareal, 2018; Tréguer et al., 2018). Therefore, current state-of-the-art biogeochemical models only consider oblates in their simulations and not prolates, but our results clearly show that diatoms, through morphological adaptations, are able to exploit a wider range of environment, which is also confirmed by (i) a *Tara* Oceans expedition study that has found comparable diatom diversity in oligotrophic and coastal regions (Malviya et al., 2016) and by (ii) fossil records of the Mediterranean sea, where sapropels (one of the most carbon-rich sediments) formation is associated with highly stratified conditions and high diatom production and export (Kemp and Villareal, 2013). Hence, we suggest that biogeochemical and earth-system models should consider diatom morphology by defining two PFTs (i.e. oblates and prolates) in their settings to improve their predictions on the consequences of climate change on primary production and carbon export.

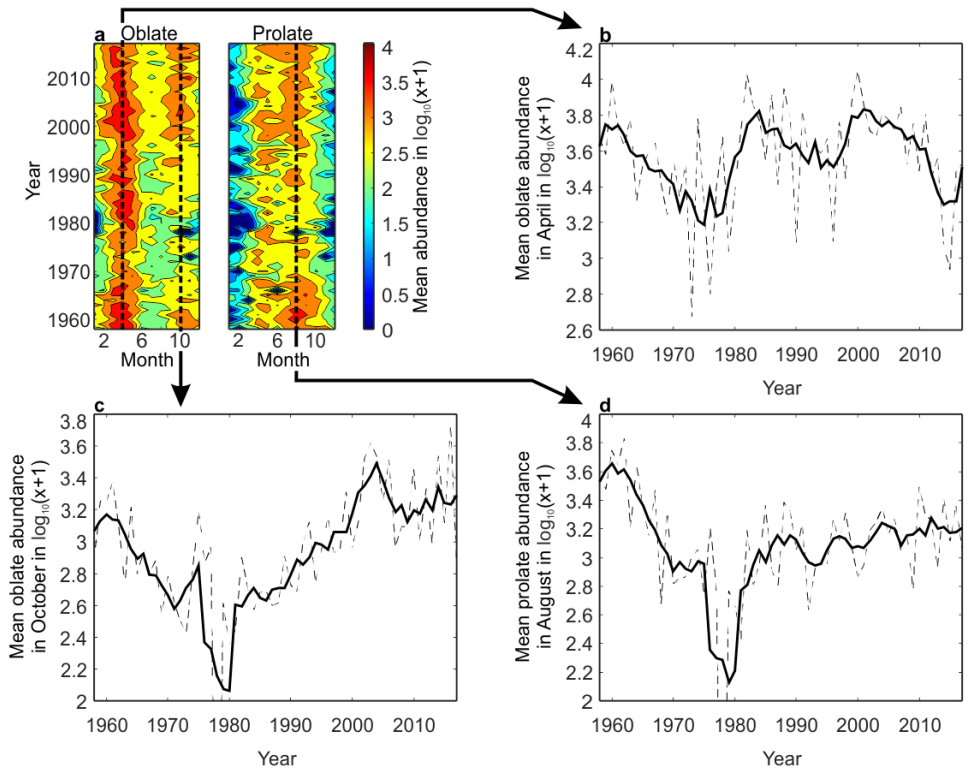



*Figure 18. Long-term changes in oblate and prolate abundance (i.e. number per CPR sample) between 1958 and 2017. (a) Long-term monthly changes in total oblate and prolate abundance. Dashed lines highlight the months that were considered in the examination of long-term changes in oblates or prolates on panels (b), (c) and (d). (b-d) Long-term changes in the abundance of oblates in (b) April and (c) October and (d) of prolates in August. Dashed lines represent the long-term changes in diatom abundance from panel (a) and black bold lines show the smoothed long-term changes after applying a second-order simple moving average. Note that the exceptional nil abundance value observed for oblates and prolates in 1978 is below the y axis in the panels c and d.*

## 1.2.4. Conclusions

By characterizing ADS in the North Sea with a trait-based approach, we have shown that this phenomenon is controlled by the niche-environment interaction that results from seasonal changes in morphological traits. Although four different hypotheses have been proposed to explain changes in cell shapes (Karp-Boss and Boss, 2016; Litchman and Klausmeier, 2008; Naselli-Flores et al., 2021; Naselli-Flores and Barone, 2011; Smetacek, 2001), our results suggest that cell adaptation to turbulence and viscosity is of primary importance. Cell elongation gives a competitive advantage to diatoms in low-turbulence and low-viscous waters because it diminishes sinking velocity (Padisak et al., 2003). Our results have also shown that elongation does not affect surface/volume ratio and therefore the ability of cells to absorb nutrients (Figure 15g-h). We did not find any evidence that elongation might enhance light capture despite some suggestions that it might (Naselli-Flores and Barone, 2011). Finally, although some studies have suggested that zooplankton can feed on large chains and elongated cells (Djeghri et al., 2018; Visser and Jonsson, 2000), the positive correlation between copepod abundance and elongation that we found suggests predation may

also be an important selective agent. Our study suggests that prolates and oblates are key adaptations to specific environments, conferring to species a specific niche. The resulting niche-environment interaction explains phenology and therefore the precise position of a species in the sequence of ADS. Understanding the relationships between morphological traits and species niche can help us to better anticipate community shifts and their functional implications for the ecosystems in the context of global climate change (Litchman et al., 2012). To conclude, as both oblates and prolates exhibit different long-term changes in abundance, suggesting that they might also have different responses to climate forcing, we encourage biogeochemical and earth-system modelers to implement these two diatom groups to improve the assessment of marine primary production and carbon export.

*Acknowledgements*

We would like to thank the owners and crews that have towed the CPRs on a voluntary basis for over 80 years contributing to one of the world's largest and longest ongoing ecological experiments. Without these early pioneers of citizen science and broad-scale volunteer monitoring projects this unique ecological dataset would never have been financially or logistically viable. The authors also want to thanks Eric Lecuyer, Dimitra-Ioli Skouroliakou, Monica Michel-Rodriguez and Lucie Courcot for their helpful comments during the design of this study.

## 1.2.5. Supplementary Files

**Supplementary Text**

**Text S1. CPR limitations**

The Continuous Plankton Recorder (CPR) survey is a long-term plankton-monitoring programme, operated currently by the Marine Biological association of Plymouth. The CPR is the longest and most extensive programme of that kind in the world, started in 1931, with minor modifications to the mechanism after 1948. After a break during World War II, zooplankton sampling has not been changed since 1948, and phytoplankton since 1958 (Warner and Hays, 1994). The machine is a high-speed plankton recorder, towed behind voluntary merchant ships, called "ships of opportunity", operating at a depth of approximately 7 m, filtering phytoplankton and zooplankton (Hays and Warner, 1993).

A sampling road effect may affect phytoplankton abundance data. Usually phytoplankton cells are smaller than zooplankton species so their abundance is greater; this is an expression of Damuth's law (Damuth, 1981), which links body size of an organism to its species population density by an inverse relationship, i.e. smaller organisms tend to have a greater population density.

The CPR survey currently records ~1000 plankton entities (many to species level) in routine taxonomic analysis dating back over multi-decades. Due to the mesh size of CPR silks, many plankton species are only semi-quantitatively sampled owing to the small size of the organisms. In the case of phytoplankton there is thus a bias towards recording larger armoured flagellates and chain-forming diatoms and that smaller species abundance estimates from cell counts will probably be underestimated in relation to other water sampling methods. However, the proportion of the population that is retained by the CPR silk reflects the major changes in abundance, distribution and specific composition, i.e. the percentage retention is roughly

constant within each species even with very small-celled species (Edwards et al., 2006). Potential under estimation of zooplankton abundances has recently been thoroughly statistically explored by Hélaouët *et al.* (Helaouët et al., 2016) who found that while the CPR survey does underestimate abundance in some cases the CPR survey does give a correct picture of both temporal (i.e. seasonal and diel scales) and spatial (i.e. regional to basin-scale) changes in zooplankton taxa. The Hélaouët et al. (2016) study also showed that while the CPR sampling is restricted to the surface waters (~10 m in depth) the seasonal and diel patterns of abundance of *C. finmarchicus* were positively correlated to patterns of abundance to a depth of 100 m.

The CPR is not a perfect sampling mechanism by any means and it will underestimate components of the plankton, e.g. large plankton like fish larvae and delicate gelatinous plankton. This has been well documented and users of the data are advised of the CPR's limitations (Richardson et al., 2006). It is also widely recognised that all plankton sampling systems have their own limitations and nuances and all underestimate abundance to some degree and that the varying mechanisms are not always directly comparable (Owens et al., 2013). It is not known whether certain taxa such as krill are capable of gear avoidance and in some cases this might be a bias for some fast moving taxa, however, since CPR sampling is conducted at high-speeds (up to 20 knots), gear avoidance is considered to be fairly minimal (Owens et al., 2013). A detailed study has been conducted on flow rate and ship speed on CPR sampling (Jonas et al., 2004) given that the speed of the ships has, in some circumstances, increased since the 1960s, which may impact sample efficiencies. However, no significant correlation was found between the long-term changes in the speed of the ships and two commonly used indicators of plankton variability: the Phytoplankton Colour and the Total Copepods indices. This absence of relationship may indicate that the effect found is small in comparison with the influence of hydroclimatic forcing (Jonas et al., 2004). For further details on the technical background, methods, consistency, and comparability of CPR sampling, see Batten et al. (Batten et al., 2003).

It is important to note that the CPR survey is the only scientific monitoring programme of that kind in the world, with no equivalent existing program. It covers an important time scale from 1948 to present, still active, with a large spatial scale covering the whole North Atlantic Ocean and its adjacent seas such as the Channel, the Celtic Sea and the North Sea. It has become "*the most extensive long-term survey of marine organisms in the world*" (Reid et al., 2003).

**Text S2. CPR data availability and key references and/or internet sites for each environmental dataset we used.**

All Continuous Plankton Recorder data is freely available on request by contacting the Marine Biological Association (MBA), United Kingdom. Data requests (Dan Lear: dble@mba.ac.uk).

Sea surface temperature (°C) originated from the NOAA OI SST V2 high resolution dataset provided by the Earth System Research Laboratory, Physical Science Division, Boulder Colorado, USA (Reynolds et al., 2007).
https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html

Mixed layer depth (m) was provided by the Copernicus Marine Environment Monitoring service (CMEMS) and originated from the Global Ocean Physical Reanalysis product (GLOBAL_REANALYSIS_ PHY_001_030). We consider the Mixed Layer Depth (MLD) as an indicator of the turbulence level. In the database considered, the MLD is estimated using a classical threshold method on the density profile (De Boyer Montégut, 2004).

https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_PHY_001_030

Phosphate, silicate and nitrate concentrations ($mmol.m^{-3}$) were provided by the CMEMS and came from the Global Ocean Biogeochemistry Hindcast (GLOBAL_REANALYSIS_BIO_001_029).
https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_REANALYSIS_BIO_001_029

Photosynthetically Active Radiation (PAR, $mole/m^2/day$) was provided by Globcolour Hermes. PAR was estimated from satellites data (Frouin et al., 2003). http://hermes.acri.fr/index.php

Euphotic depth of sea water (m) was provided by the NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Euphotic Depth Data; 2018 Reprocessing. NASA OB.DAAC, Greenbelt, MD, USA. doi: data/10.5067/AQUA/MODIS/L3M/ZLEE/2018.
Euphotic depth was calculated with Lee algorithm (Lee et al., 2007; Lee, 2005).
https://oceancolor.gsfc.nasa.gov/data/10.5067/AQUA/MODIS/L3M/ZLEE/2018/

Bathymetry (m) was provided by the British Oceanographic Data Centre (BODC). doi:10.5285/836f016a-33be-6ddc-e053-6c86abc0788e
https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/836f016a-33be-6ddc-e053-6c86abc0788e/

Distance to nearest coastlines (km) was provided by the Nasa's Ocean Biology Processing Group, Goddard Space Flight Center. Produced by Rick Stumpf (NOAA, National Ocean Service) and Norman Kuring at the NASA Ocean Biology Processing Group in June 2009, based on concept of Eric Leuliette of NOAA Satellite Altimetry Lab.
https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/

Long-term changes in SST (°C), mean surface downward short-wave radiation flux ($W.m^{-2}$), and northward and eastward wind components ($m.s^{-1}$) were provided by the Copernicus climate change service (https://climate.copernicus.eu/) and originated from the ERA5 dataset of monthly averaged data on single levels. We used the data from 1979 to present (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means?tab=overview) and the preliminary version of the data between 1950 and 1978 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means-preliminary-back-extension?tab=overview). Data on mean surface downward short-wave radiation flux (solar radiations) were used as a proxy of PAR.

**Text S3. Adaptations to variable light intensity**

Although an increase in the maximal area of light interception normalized by volume (β) might be a way to adapt for low light intensity, other physiological processes, not tested in the present study, have been reported. Here we describe four other processes by which diatom cells may adjust to variable light intensity.

First, diatoms can redistribute light within the cell with their frustule to optimize chloroplast exposure. Several studies have shown that the optical properties of diatom frustules, which has been assimilated to a biological photonic crystal, can modulate incoming incident light inside the living cell, resulting in a redistribution, optimization and filtering of incident sunlight to the photosynthetic apparatuses in the chloroplasts (De Tommasi, 2016; Fuhrmann et al., 2004; Goessling et al., 2018; Noyes et al., 2008).

Second, diatom have also specific light-harvesting systems, referred to as fucoxanthin-chlorophyll proteins (FCPs) to collect and convert light energy, with a particularly high efficiency

of energy transduction and photoprotection due to a high concentration of carotenoids (Büchel, 2019).

Third, diatoms have seggregated photosynthetic complex (photosystems I and II) located inside the loosely stacked tylakoid membranes in the chloroplast. The separation of the two photosystems enables an improved light utilisation and their interconnection allows a rapid equilibration of electron carriers that optimizes photosynthesis (Flori et al., 2017).

Fourth, most of centric diatoms are polyplastidic and can adjust the location of the chloroplasts inside the cell to adapt to different light intensities (Mann, 1996).

The optical properties of diatom frustules, the adjustment of the location of chloroplast inside the cell, the tylakoid architecture and the light harvesting systems, are mechanisms by which diatoms may adapt to variable light intensity, which perhaps explain why we found many diatoms with low β values (Figure S8).

**Supplementary Figures**

**Figure S1.** Location of the studied area (51 to 60°N, 3°W to 9°E).

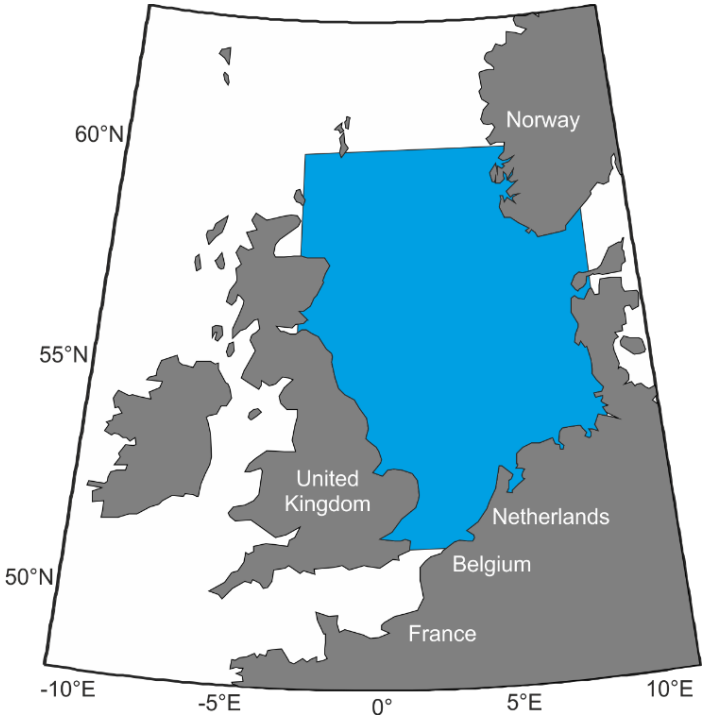



**Figure S2. Daily changes in Photosynthetically Active Radiation (PAR), nitrate, silicate, phosphate, Mixed Layer Depth (MLD), (small and large) copepod and diatom abundance.** (**a**) Daily changes in PAR, (**b**) nitrate (red line), silicate (blue) and phosphate (black), (**c**) MLD (blue), mean small (≤2mm; black) and large (>2mm; red) copepod abundance. Daily changes in the abundance of species of the (**d**) spring-dominated group, (**e**) autumn-dominated group and (**f**) Late-Spring Summer Autumn (LSSA) group. Coloured lines reflect daily changes at the species level. The black bold lines on panels **d-f** correspond to daily mean abundance of all species of a group. Species composition of each group is displayed in Figure 17 and Figures S4-5. Here we used the MLD as a proxy of turbulence level (the deeper the MLD is, the more active the wind-induced turbulence is likely to be) and stratification (a lower MLD is in general associated with a higher stratification, which prevent nutrients reloading at the surface).

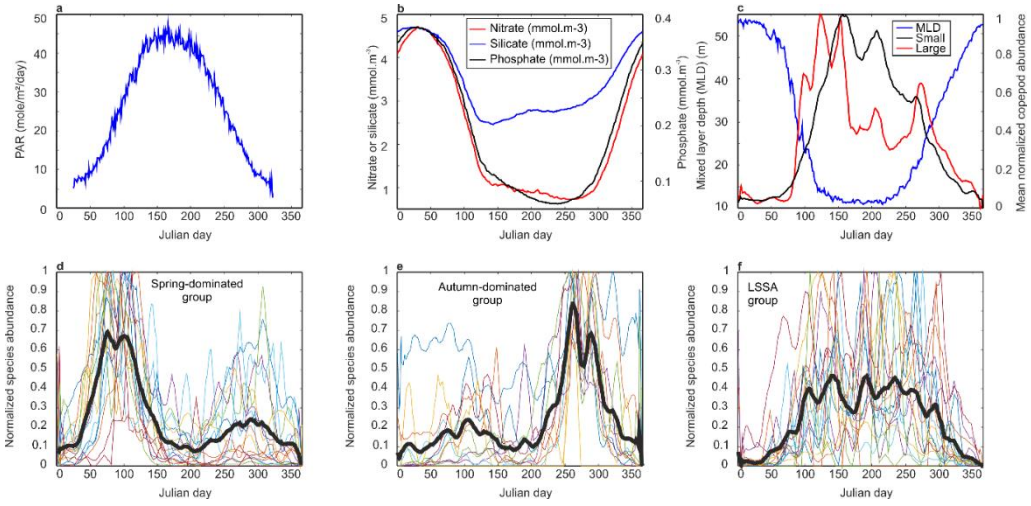


**Figure S3. Relationships between the first three eigenvectors and species subclasses.** Oblates are in blue, prolates in red and oblate/prolates in green. The three groups were defined by a cluster analysis (see Figure S4). The volume occupied by the spring-dominated group is in blue, the autumn-dominated group in green and the LSSA group in red.

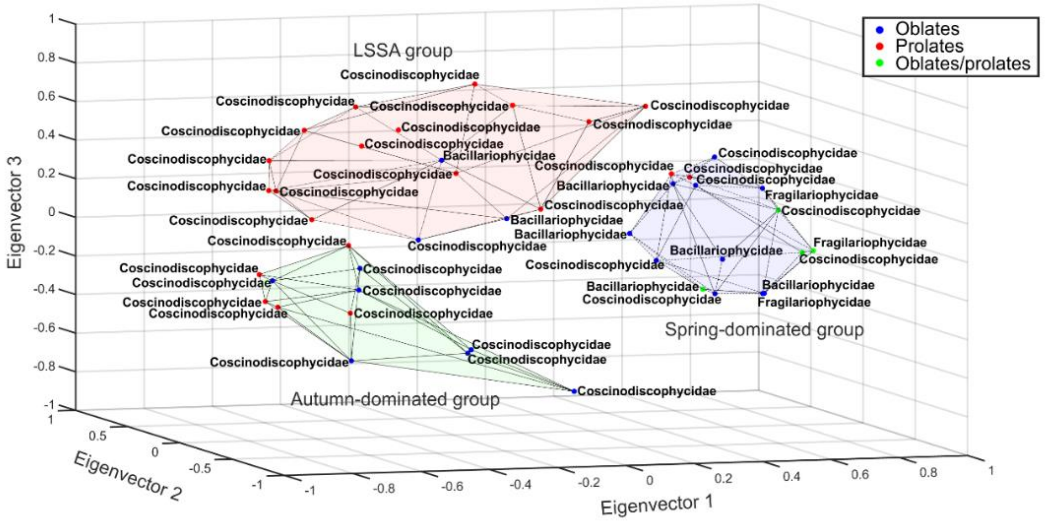



**Figure S4. Dendrogram based on a cluster analysis performed on the first three eigenvectors, using a Ward algorithm applied on a squared matrix of Euclidean distance.** A total of three groups was identified at a cut-off level of 8. The black dashed line denotes the position of the cut-off level. Colours enable the distinction of the three groups: (1) spring-dominated group (blue), (2) LSSA group (red) and (3) autumn-dominated group (green). Species names are those recorded in the CPR dataset.

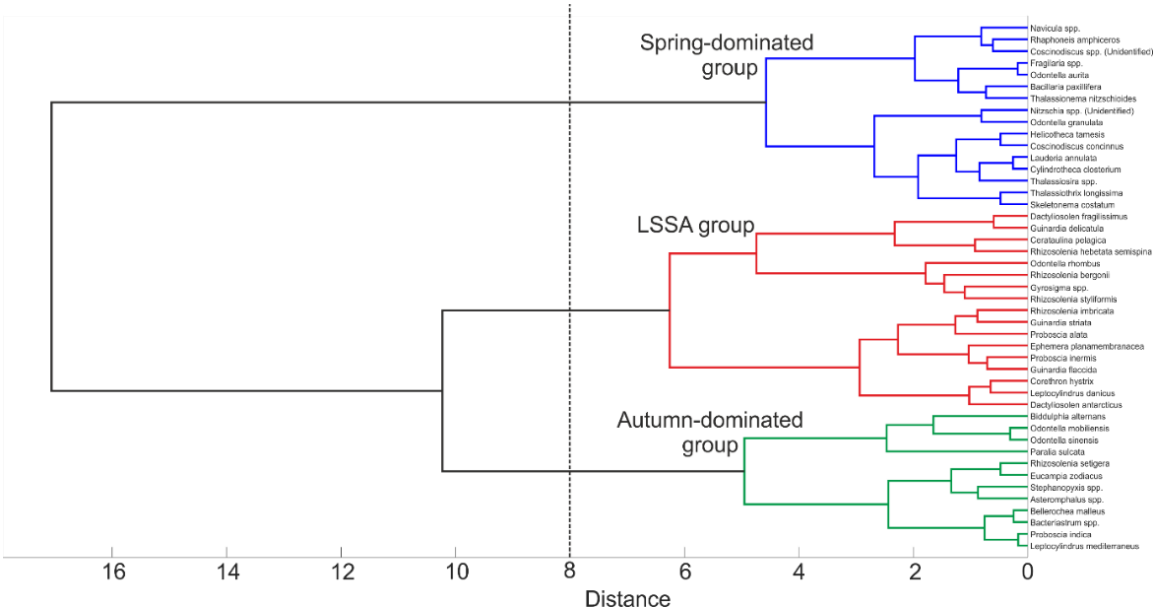


**Figure S5. Species composition of the three species groups defined in the space of the first three eigenvectors of the PCA.** Oblates are in blue, prolates in red and oblate/prolates in green. The volume occupied by the spring-dominated group is in blue, the autumn-dominated group in green and the LSSA group in red. Species names are those recorded in the CPR dataset.

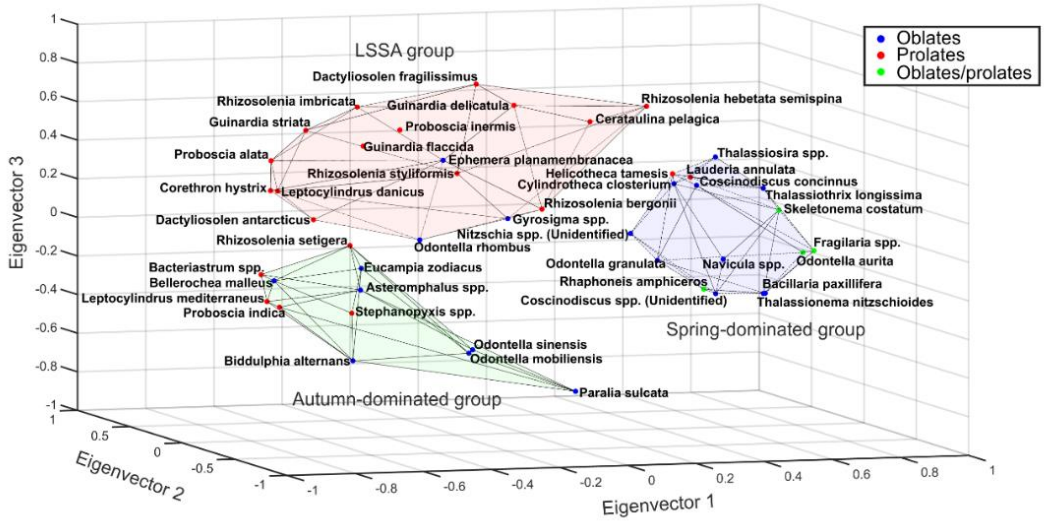



**Figure S6. Relationships between the first three eigenvectors and morphological traits.** (**a**) Mean cell diameter and (**b**) mean cell height. The three groups were defined by a cluster analysis (see Figure S4). The volume occupied by the spring-dominated group is in blue, the autumn-dominated group in green and the LSSA group in red.

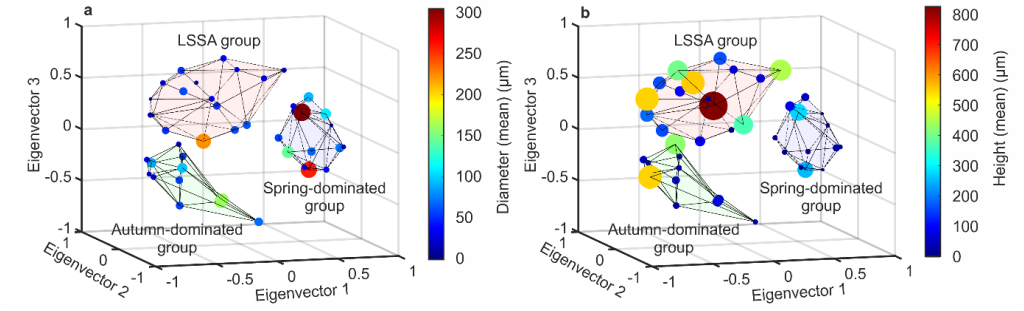


**Figure S7. Comparison between values of diatom morphological traits approximated by cylinders and spheroids.** Comparisons between (**a**) area, (**b**) volume, (**c**) area/volume ratio and (**d**) δ values based on cylinders and spheroids. (**a**-**d**) The value of the Pearson correlation coefficient is displayed on the top left. (**e**-**h**) Relationships between the first three eigenvectors and morphological traits estimated by assimilating their shapes to (**e**, **g**) cylinders and (**f**, **h**) spheroids: (**e**-**f**) area/volume ratio or (**g**-**h**) δ.

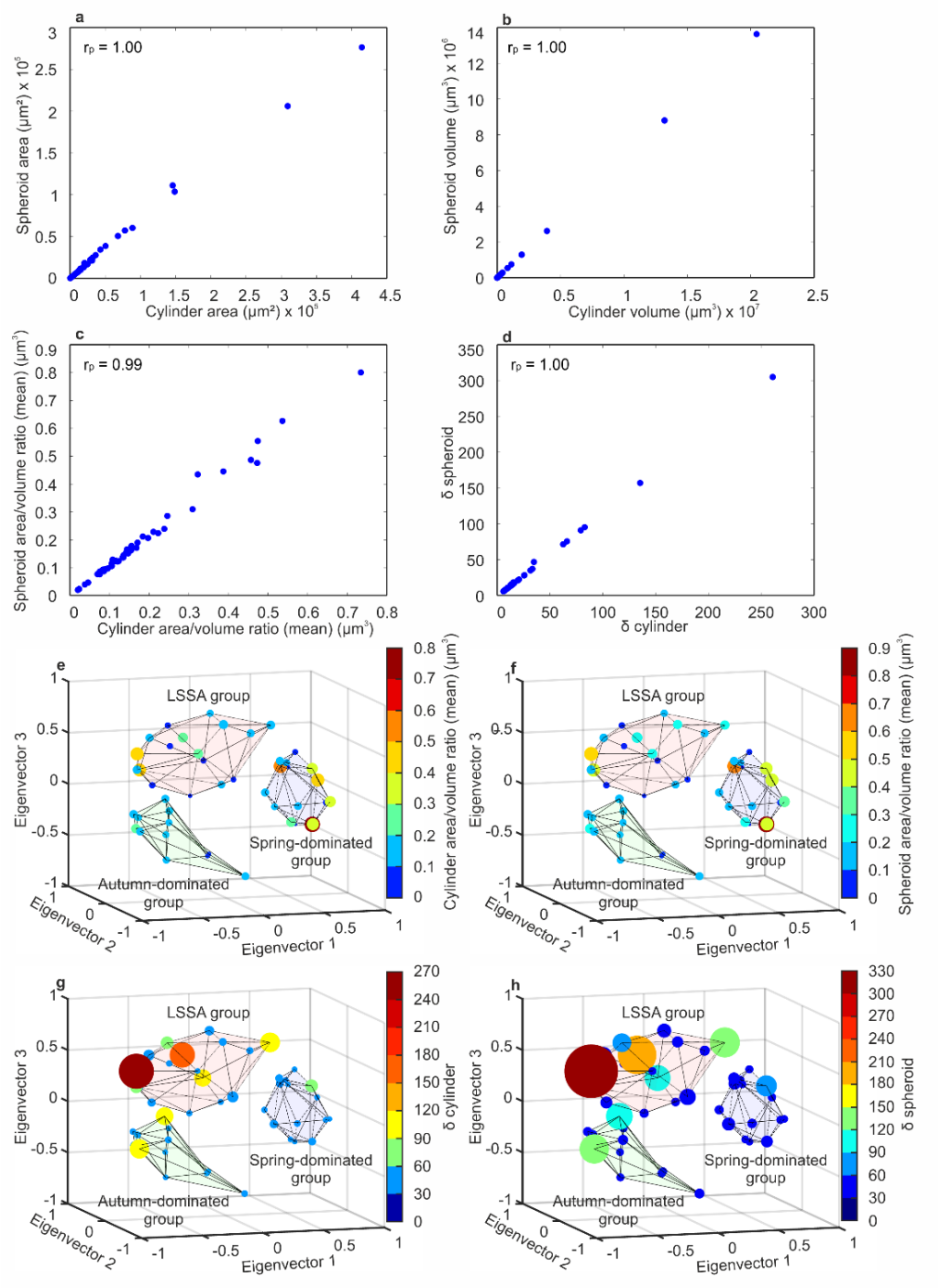


**Figure S8. Relationships between maximum area of light interception normalized by volume (β) and PAR.** (**a**) Oblates are in blue and prolates in red. Spearman rank correlation coefficient ($r_s$) and its associated probability ($p_s$) are displayed on the top right. (**b**) Spring-dominated group is in blue, Autumn-dominated group in green and LSSA group in red. In each panel, a dashed polygon was drawn around all dots to emphasize a potential relationship between PAR and β.

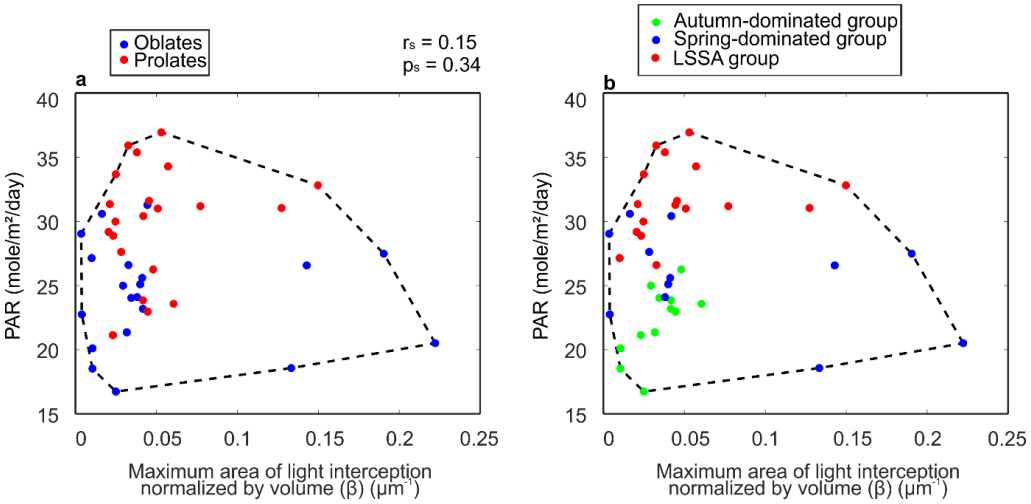



**Figure S9. Relationships between Mixed Layer Depth (MLD) and an index of cell morphology ($\delta$).** (**a**) Oblates are in blue, prolates in red and oblates/prolates in green. Blue lines delimit the space occupied by oblates, red lines by prolates and green lines by oblates/prolates. Spearman rank correlation coefficient ($r_s$) and its associated probability ($p_s$) are displayed on the top right. (**b**) Species belonging to the spring-dominated group are in blue, autumn-dominated group in green and LSSA in red. Blue lines delimit the space occupied by spring-dominated group, green lines by autumn dominated group and red lines by LSSA group.

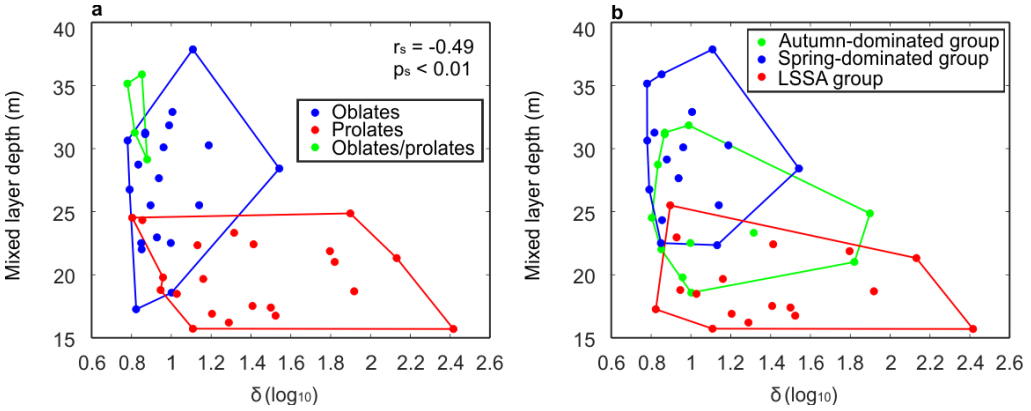


**Figure S10. Relationships between kinematic viscosity and an index of cell morphology ($\delta$).** (**a**) Oblates are in blue, prolates in red and oblates/prolates in green. Blue lines delimit the space occupied by oblates, red lines by prolates and green lines by oblates/prolates. Spearman rank correlation coefficient ($r_s$) and its associated probability ($p_s$) are displayed on the top right. (**b**) Species belonging to the spring-dominated group are in blue, autumn-dominated group in green and LSSA in red. Blue lines delimit the space occupied by spring-dominated group, green lines by autumn dominated group and red lines by LSSA group.

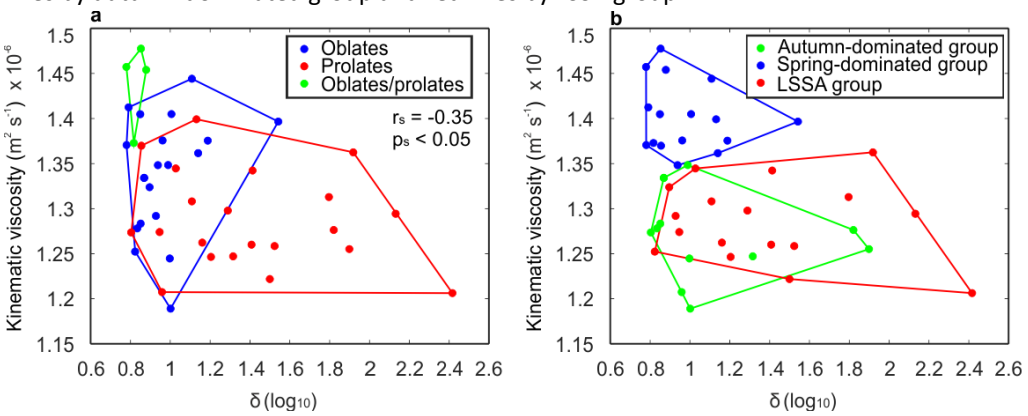



**Figure S11. Relationships between (a-b) small (≤2mm)/(c-d) large (>2mm) copepod abundance and morphological traits (height-diameter differences).** (**a** and **c**) Oblates are in blue, prolates in red and oblates/prolates in green. Spearman rank correlation coefficient ($r_s$) and its associated probability ($p_s$) are displayed on the top right. (**b** and **d**) spring-dominated group is in blue, autumn-dominated group in green and LSSA group in red.

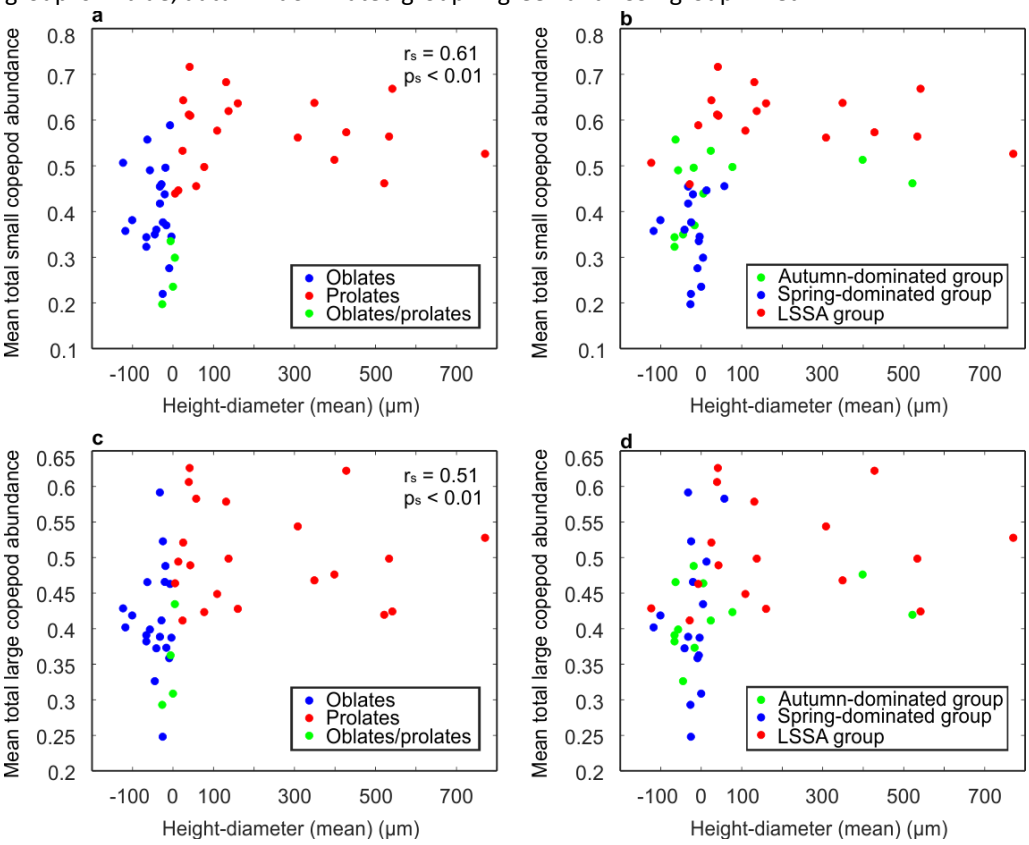


**Figure S12.** ***Caption overleaf.***

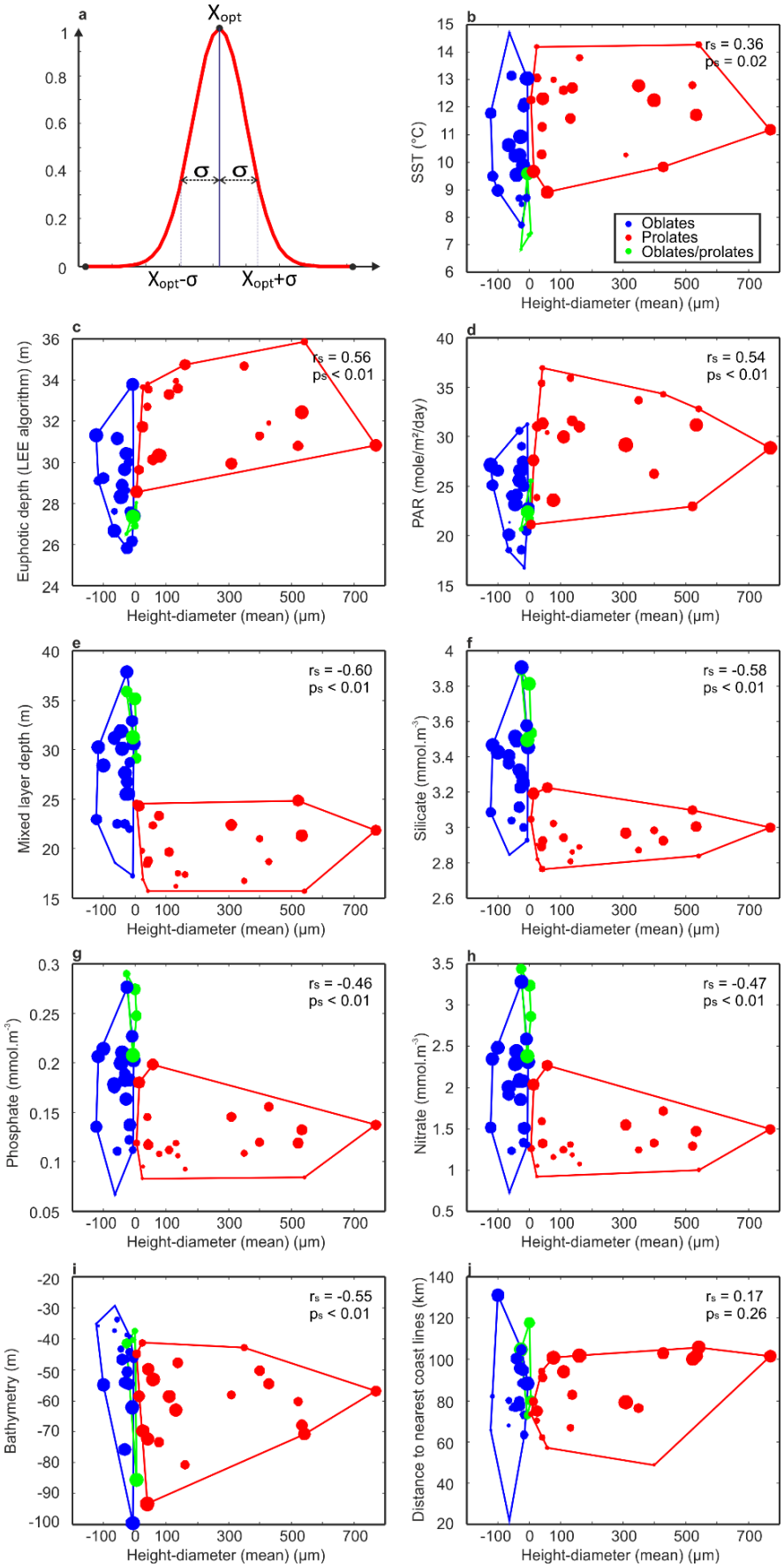


**Figure S12. (*Overleaf.*) Relationships between the ecological niche of species (ecological optima and amplitude) and cell morphological traits (mean cell height-diameter difference).** (**a**) Theoretical ecological niche with $X_{opt}$ representing species niche optimum and σ denoting niche amplitude assessed by the weighted standard deviation. The Y axis corresponds to the species abundance and the X axis to an hypothetical environmental parameter. Relationships between mean cell height-diameter difference and niche optimum/amplitude for (**b**) SST, (**c**) euphotic depth, (**d**) PAR, (**e**) mixed layer depth, (**f**) silicate, (**g**) phosphate, (**h**) nitrate concentration, (**i**) bathymetry and (**j**) distance to nearest coastlines. The niche optimum can be read on the ordinate axis. The size of the bullet is a function of species' rank of the weighted standard deviation; larger bullets are associated with larger niche amplitude and therefore ranks. Oblates are in blue, prolates in red and oblate/prolates in green. Spearman rank's correlation coefficient ($r_s$) and its associated probability ($p_s$) are displayed on the top right of panel **b-j**.

**Supplementary Tables.**

**Table S1. Percentage of geometrical shapes and oblates, prolates and oblates/prolates in each species group revealed by the cluster analysis.**

| | Spring-dominated group | Autumn-dominated group | LSSA group |
|---|---|---|---|
| Cylindrical species | 31% | 58% | 82% |
| Rectangular species | 25% | 0% | 6% |
| Prismatic species | 44% | 42% | 12% |
| Oblates | 62.5% | 58% | 18% |
| Prolates | 12.5% | 42% | 82% |
| Oblates/prolates | 25% | 0% | 0% |

**Table S2. Values of some diatom morphological traits: minimum (min), maximum (max) and mean of cell diameter and height.** Species names are those recorded in the CPR dataset.

| **Species names** | **Min diam (µm)** | **Max diam (µm)** | **Min height (µm)** | **Max height (µm)** | **Mean diam (µm)** | **Mean height (µm)** | **Shape** |
|---|---|---|---|---|---|---|---|
| *Asteromphalus* spp. | 15 | 175 | 5 | 58 | 95 | 31.5 | Oblate |
| *Bacillaria paxillifera* | 21.11 | 45.14 | 5 | 10 | 33.12 | 7.5 | Oblate |
| *Bacteriastrum* spp. | 5 | 56 | 28 | 80 | 30.5 | 54 | Prolate |
| *Bellerochea malleus* | 16.34 | 155.93 | 13 | 45 | 86.13 | 29 | Oblate |
| *Biddulphia alternans* | 28.28 | 84.85 | 20 | 60 | 56.57 | 40 | Oblate |
| *Cerataulina pelagica* | 7 | 60 | 25 | 120 | 33.5 | 72.5 | Prolate |
| *Corethron hystrix* | 12 | 38 | 20 | 350 | 25 | 185 | Prolate |
| *Coscinodiscus concinnus* | 110 | 500 | 60 | 500 | 305 | 280 | Oblate |
| *Coscinodiscus* spp. | 15 | 500 | 7 | 500 | 257.5 | 253.5 | Oblate |
| *Cylindrotheca closterium* | 6.31 | 45.13 | 2.5 | 8 | 25.72 | 5.25 | Oblate |

| Species names | Min diam (µm) | Max diam (µm) | Min height (µm) | Max height (µm) | Mean diam (µm) | Mean height (µm) | Shape |
|---|---|---|---|---|---|---|---|
| *Dactyliosolen antarcticus* | 13 | 90 | 21 | 300 | 51.5 | 160.5 | Prolate |
| *Dactyliosolen fragilissimus* | 8 | 70 | 40 | 300 | 39 | 170 | Prolate |
| *Ephemera planamembranacea* | 25.93 | 33.85 | 15 | 30 | 29.89 | 22.5 | Oblate |
| *Eucampia zodiacus* | 16 | 89.44 | 8 | 60 | 52.72 | 34 | Oblate |
| *Fragilaria* spp. | 7.48 | 31.37 | 3.5 | 35 | 19.43 | 19.25 | Oblate & prolate |
| *Guinardia delicatula* | 8 | 40 | 20 | 110 | 24 | 65 | Prolate |
| *Guinardia flaccida* | 25 | 95 | 45 | 160 | 60 | 102.5 | Prolate |
| *Guinardia striata* | 6 | 50 | 30 | 300 | 28 | 165 | Prolate |
| *Gyrosigma* spp. | 22.3 | 95.75 | 11 | 50 | 59.02 | 30.5 | Oblate |
| *Helicotheca tamesis* | 11.28 | 49.44 | 55 | 120 | 30.36 | 87.5 | Prolate |
| *Lauderia annulata* | 15 | 75 | 20 | 96 | 45 | 58 | Prolate |
| *Leptocylindrus danicus* | 5 | 15 | 20 | 50 | 10 | 35 | Prolate |
| *Leptocylindrus mediterraneus* | 7 | 35 | 21 | 175 | 21 | 98 | Prolate |
| *Navicula* spp. | 7.74 | 127.28 | 2.5 | 50 | 67.51 | 26.25 | Oblate |
| *Nitzschia* spp. | 2.39 | 110.77 | 1.5 | 47 | 56.58 | 24.25 | Oblate |
| *Odontella aurita* | 14.14 | 134.16 | 15 | 80 | 74.15 | 47.5 | Oblate & prolate |
| *Odontella granulata* | 140.57 | 144.83 | 25 | 25 | 142.7 | 25 | Oblate |
| *Odontella mobiliensis* | 53.67 | 268.33 | 40 | 150 | 161 | 95 | Oblate |
| *Odontella rhombus* | 83.67 | 363.32 | 50 | 150 | 223.49 | 100 | Oblate |
| *Odontella sinensis* | 53.67 | 268.33 | 40 | 150 | 161 | 95 | Oblate |
| *Paralia sulcata* | 8 | 130 | 3 | 45 | 69 | 24 | Oblate |
| *Proboscia alata* | 2 | 15 | 100 | 1000 | 8.5 | 550 | Prolate |
| *Proboscia indica* | 3 | 54 | 100 | 1000 | 28.5 | 550 | Prolate |
| *Proboscia inermis* | 12 | 21 | 100 | 1000 | 16.5 | 550 | Prolate |
| *Rhaphoneis amphiceros* | 15.14 | 39.89 | 18 | 25 | 27.52 | 21.5 | Oblate & prolate |
| *Rhizosolenia bergonii* | 9 | 115 | 210 | 530 | 62 | 370 | Prolate |
| *Rhizosolenia hebetata semispina* | 4.5 | 40 | 150 | 750 | 22.25 | 450 | Prolate |

| Species names | Min diam (µm) | Max diam (µm) | Min height (µm) | Max height (µm) | Mean diam (µm) | Mean height (µm) | Shape |
|---|---|---|---|---|---|---|---|
| *Rhizosolenia imbricata* | 2 | 100 | 150 | 650 | 51 | 400 | Prolate |
| *Rhizosolenia setigera* | 3 | 50 | 125 | 725 | 26.5 | 425 | Prolate |
| *Rhizosolenia styliformis* | 4 | 105 | 150 | 1500 | 54.5 | 825 | Prolate |
| *Skeletonema costatum* | 3 | 20 | 2 | 30 | 11.5 | 16 | Oblate & prolate |
| *Stephanopyxis* spp. | 20 | 90 | 20 | 100 | 55 | 60 | Prolate |
| *Thalassionema nitzschioides* | 3.91 | 23.67 | 1 | 8 | 13.79 | 4.5 | Oblate |
| *Thalassiosira* spp. | 2 | 185 | 2 | 120 | 93.5 | 61 | Oblate |
| *Thalassiothrix longissima* | 41.07 | 174.81 | 7 | 7 | 107.94 | 7 | Oblate |

**Table S3. Percentage of *Bacillariophycidae*, *Coscinodiscophycidae* and *Fragilariophycidae* in each species group revealed by the cluster analysis.**

| Subclass | Spring-dominated group | Autumn-dominated group | LSSA group |
|---|---|---|---|
| *Bacillariophycidae* | 31% | 0% | 12% |
| *Coscinodiscophycidae* | 50% | 100% | 88% |
| *Fragilariophycidae* | 19% | 0% | 0% |

**Table S4. Test of significance of the eigenvalues using a test based on a broken-stick distribution.** When observed eigenvalues (PCA) are above predicted values from a null model based on a broken-stick distribution, they are significant (in bold). This test demonstrates that the first three axes of the PCA are significant.

| | Predicted eigenvalues based on the broken-stick test | Observed eigenvalues (PCA) |
|---|---|---|
| **1** | **9.76** | **11.65** |
| **2** | **7.54** | **11.31** |
| **3** | **6.43** | **6.84** |
| 4 | 5.69 | 2.99 |
| 5 | 5.14 | 2.14 |
| 6 | 4.70 | 1.64 |
| 7 | 4.33 | 1.36 |
| 8 | 4.01 | 1.08 |
| 9 | 3.73 | 1.02 |
| 10 | 3.48 | 0.83 |

**Table S5. Correlations between the first three principal components and the daily climatology of each environmental parameter.** Significant correlations ($r_s \geq 50$ or $r_s \leq -50$ and $p<0.05$) are in bold.

| | $r_s$ (PC1) | $p_s$ (PC1) | $r_s$ (PC2) | $p_s$ (PC2) | $r_s$ (PC3) | $p_s$ (PC3) |
|---|---|---|---|---|---|---|
| Nitrate | **0.76** | **<0.01** | **-0.67** | **<0.01** | -0.16 | <0.01 |
| Phosphate | **0.74** | **<0.01** | **-0.61** | **<0.01** | -0.29 | <0.01 |
| Silicate | 0.33 | <0.01 | **-0.56** | **<0.01** | **-0.58** | **<0.01** |
| PAR | -0.18 | <0.01 | 0.09 | 0.14 | **0.82** | **<0.01** |
| Euphotic depth | **-0.54** | **<0.01** | 0.03 | 0.6 | **0.68** | **<0.01** |
| Mixed layer depth | 0.41 | <0.01 | **-0.52** | **<0.01** | **-0.62** | **<0.01** |
| SST | **-0.88** | **<0.01** | 0.39 | **<0.01** | 0.23 | <0.01 |
| Small copepod abundance | -0.12 | <0.05 | **0.68** | **<0.01** | 0.47 | <0.01 |
| Large copepod abundance | -0.41 | <0.01 | **0.51** | **<0.01** | **0.61** | **<0.01** |
| Kinematic viscosity | **0.88** | **<0.01** | -0.39 | <0.01 | -0.23 | <0.01 |

**Table S6. Test of significance of the height-diameter differences among the three species groups revealed by the cluster analysis.** Significant differences ($p<0.05$) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | **<0.01** |
| **Post-hoc tests** | |
| Autumn-dominated and spring-dominated group | 0.81 |
| Autumn-dominated and LSSA group | **0.03** |
| Spring-dominated and LSSA group | **<0.01** |

**Table S7. Test of significance of the mean cell diameter difference among the three species groups revealed by the cluster analysis.** Significant differences ($p<0.05$) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | 0.19 |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.91 |
| Autumn-dominated and LSSA group | 0.21 |
| Spring-dominated and LSSA group | 0.35 |

**Table S8. Test of significance of the mean cell height difference among the three species groups revealed by the cluster analysis.** Significant differences ($p<0.05$) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-value** |
|---|---|
| All groups | **<0.01** |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.15 |
| Autumn-dominated and LSSA group | 0.24 |
| Spring-dominated and LSSA group | **<0.01** |

**Table S9. Test of significance of the area/volume ratio difference among the three species groups revealed by the cluster analysis.** Significant differences (p<0.05) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | 0.42 |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.39 |
| Autumn-dominated and LSSA group | 0.76 |
| Spring-dominated and LSSA group | 0.77 |

**Table S10. Test of significance of δ difference among the three species groups revealed by the cluster analysis.** Significant differences (p<0.05) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | **<0.01** |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.72 |
| Autumn-dominated and LSSA group | 0.11 |
| Spring-dominated and LSSA group | **<0.01** |

**Table S11. Test of significance of the area/volume ratio (estimated with spheroidal cell shapes) difference among the three species groups revealed by the cluster analysis.** Significant differences (p<0.05) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | 0.45 |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.42 |
| Autumn-dominated and LSSA group | 0.78 |
| Spring-dominated and LSSA group | 0.79 |

**Table S12. Test of significance of δ (estimated with spheroidal cell shapes) difference among the three species groups revealed by the cluster analysis.** Significant differences (p<0.05) from a Kruskal-Wallis test and associated post-hoc tests are in bold.

| **Kruskal-Wallis test** | **P-values** |
|---|---|
| All groups | **<0.05** |
| **Post-hoc test** | |
| Autumn-dominated and spring-dominated group | 0.75 |
| Autumn-dominated and LSSA group | 0.13 |
| Spring-dominated and LSSA group | **<0.05** |

**Table S13. Comparison of mean of different environmental parameters (ecological optima) among the two groups oblates and prolates using a Kruskal-Wallis test.** Significant differences (p<0.05) are in bold.

| **Kruskal-Wallis test** | **P-value** |
|---|---|
| Nitrate | **<0.01** |
| Phosphate | **<0.01** |
| Silicate | **<0.01** |
| PAR | **<0.01** |
| Euphotic depth | **<0.01** |
| Mixed layer depth | **<0.01** |
| SST | **<0.01** |

**Table S14. Comparison of mean of different environmental parameters (ecological amplitudes) among the two groups oblates and prolates using a Kruskal-Wallis test.** Significant differences (p<0.05) are in bold.

| Kruskal-Wallis test | P-value |
|---|---|
| Nitrate | **<0.01** |
| Phosphate | **<0.01** |
| Silicate | **<0.01** |
| PAR | 0.45 |
| Euphotic depth | 0.14 |
| Mixed layer depth | **<0.01** |
| SST | 0.13 |

**Table S15. Nonlinear (Spearman) correlation between long-term changes in oblate/prolate abundance ($log_{10}(x+1)$) and the long-term changes in SST, solar radiations and wind.** Significant correlations (p<0.05) are in bold.

| | Oblate April | | Oblate October | | Prolate August | |
|---|---|---|---|---|---|---|
| | $r_s$ | $p_s$ | $r_s$ | $p_s$ | $r_s$ | $p_s$ |
| SST (°C) | 0.13 | p>0.05 | **0.57** | **P<0.01** | 0.06 | p>0.05 |
| Solar radiations ($W.m^{-2}$) | 0.04 | p>0.05 | 0.13 | p>0.05 | 0.06 | p>0.05 |
| Northward wind components ($m.s^{-1}$) | 0.06 | p>0.05 | -0.01 | p>0.05 | 0.20 | p>0.05 |
| Eastward wind components ($m.s^{-1}$) | **-0.26** | **P<0.05** | **-0.35** | **P<0.05** | 0.18 | p>0.05 |
| Total wind | -0.14 | p>0.05 | **-0.29** | **P<0.05** | 0.08 | p>0.05 |

# ***Chapter 2.*** *Plankton biodiversity, ecological niches and species coexistence*

*Summary*

In this chapter, we examine the spatio-temporal organisation of plankton biodiversity in the North Atlantic Ocean. First, by defining species assemblages based on both phytoplankton and zooplankton species/taxa, we show that the pelagic ecosystems of the North Atlantic Ocean are organised in a complex way (Kléparski et al., 2021). We show that phytoplankton and zooplankton biodiversity are strongly influenced by hydro-climatic conditions and especially the 10°C isotherm. Although no assemblages are identified as being a unique indicator of a specific oceanic region, some are preferentially found in one of the three main biomes of the North Atlantic Ocean, i.e. the Polar, the Westerlies and the Coastal Boundary Zone biomes *sensu* Longhurst (1998). We suggest that the North Atlantic region is characterised by a mosaic of plankton assemblages that have overlapping spatial distribution in many places, generating spatial coenoclines (i.e. gradient of biocoenosis or community). We show that these spatial patterns are the consequences of the ecological niches of the species composing an assemblage and that each assemblage has its own environmental signature (i.e. the combination of the ecological niche of the species composing an assemblage), which drives their spatial distribution. To do so, we develop a new method that we called the "environmental chromatogram", which is based on a graphic that identifies where species of an assemblage aggregate along various environmental gradients figuring the multiple ecological dimensions.

We then adapt the "environmental chromatogram" method at the species organisational level to rapidly characterise the (multidimensional) ecological niche of a species in a plane. The "environmental chromatogram" is therefore renamed a "community chromatogram" when applied at the community/assemblage level, and a "species chromatogram" when applied at the species level. A species chromatogram gives a graphical summary of the niche by representing together abundance gradients with respect to all environmental variables. From a species chromatogram, one can easily quantify the niche optimums and amplitudes and when several chromatograms are examined, the niche of different species can be easily compared. We also propose an index that quantifies the degree of niche overlapping as well as the identification of the most discriminant combinations of environmental variables. We compare our method with the *dynRB* (Junker et al., 2016) and *Hypervolume* (Blonder et al., 2018) R package and show that our estimates of niche overlapping are comparable with both methods. We also implement a Matlab and a R package (called *specieschrom*) that is freely available (Kléparski and Beaugrand, 2022).

Finally, we apply the species chromatogram method to all plankton species recorded quantitatively by the CPR survey. We do so to examine the mechanisms explaining plankton coexistence and specially to revisit the "paradox of the plankton".

Formulated in 1961 by G.E. Hutchinson, this paradox asked whether "*[…] it is possible for a number of species to coexist in a relatively isotropic or unstructured environment all competing for the same sort of materials*"(Hutchinson, 1961). We show that the niches of both phyto and zooplankton species rapidly separate when the number of niche dimensions rises. We also demonstrate that the pelagic environment is not only diverse but also impermanent in space and time. Therefore, the two implicit postulates on which the Hutchinson's paradox is based are incorrect. High niche differentiation leads species to have a unique combination of environmental conditions (i.e. a unique niche) where their performances in terms of growth and reproduction are optimal. Therefore, each species covers a part of the multidimensional environmental space and the segregation along various environmental niche dimensions enables them to coexist in space and time.

## 2.1. Plankton biogeography in the North Atlantic Ocean and its adjacent seas: species assemblages and environmental signatures

Loïck Kléparski, Grégory Beaugrand and Martin Edwards

*Published in Ecology and Evolution. Supplementary Files are displayed in section 2.1.6. Article is available at* https://doi.org/10.1002/ece3.7406

*Abstract*

Plankton biodiversity is a key component of marine pelagic ecosystems. They are at the base of the food web, control the productivity of marine ecosystems and provide many provisioning and regulating ecological services. It is therefore, important to understand how plankton are organised in both space and time. Here we use data of varying taxonomic resolution, collected by the Continuous Plankton Recorder (CPR) survey, to map phytoplankton and zooplankton biodiversity in the North Atlantic and its adjacent seas. We then decompose biodiversity into 24 species assemblages and investigate their spatial distribution using ecological units and ecoregions recently proposed. Finally, we propose a descriptive method, which we call the environmental chromatogram, to characterise the environmental signature of each plankton assemblage. The method is based on a graphic that identifies where species of an assemblage aggregate along an environmental gradient composed of multiple ecological dimensions. The decomposition of the biodiversity into species assemblages allows us to show (i) that most marine regions of the North Atlantic are composed of coenoclines (i.e. gradients of biocoenoses or communities) and (ii) that the overlapping spatial distribution of assemblages is the result of their environmental signatures. It follows that neither the ecoregions nor the ecological units identified in the North Atlantic are characterised by a unique assemblage but instead by a mosaic of assemblages that overlap in many places.

*Keywords*

Plankton, biogeography, macroecology, taxonomic assemblages, North Atlantic Ocean, environmental signature, coenoclines.

### 2.1.1. Introduction

Plankton are a key component of marine pelagic ecosystems controlling their productivity (M. Edwards et al., 2013). Phytoplankton produce by photosynthesis almost half of the oxygen at the global scale (Behrenfeld, 2014). They create endosomatic energy that is progressively channelled through the whole marine food web. Zooplankton ensure the transfer of this energy between phytoplankton and higher trophic levels such as fish. Any changes in the abundance and composition of plankton affect higher trophic levels (M. Edwards et al., 2013; Luczak et al., 2011). Plankton also control a part of carbon exportation in the North Atlantic by a process termed the biological carbon pump (Brun et al., 2019). Plankton are also good indicators of climate change impacts because they are in general not commercially exploited and have a relatively short life cycle.

Understanding the influence of climate change on plankton necessitates having a good understanding of its spatial distribution and its natural annual variations (Reid and Edwards, 2001). However, large-scale plankton monitoring programmes are rare (Richardson and Poloczanska, 2008). In the North Atlantic and some of its adjacent seas, plankton have been investigated by the Continuous Plankton Recorder (CPR) survey on a monthly basis since 1946 (Reid et al., 2003). Recently, the survey has been implicated in the European research project ATLANTOS, gathering 18 countries with the aims of bringing together all the existing Atlantic Ocean observing activities into a more integrated wide observation system ("Concept & Objectives | AtlantOS," 2019).

As part of the ATLANTOS project, a new partition of the North Atlantic Ocean and adjacent seas has been proposed (Beaugrand et al., 2019b). This new partition divides the North Atlantic into 13 ecological units and 40 ecoregions (see Supplementary Information Figures S1 and S2). An ecological unit is a group of observations with a homogeneous environmental regime, similar biodiversity and seasonal variability. It can be subsequently divided into ecoregions that are a group of interconnected observations (Beaugrand et al., 2019b). The partition was primarily based on spatio-temporal fluctuations in plankton biodiversity but also physical data (e.g. bathymetry, SST, mean surface current). This new partition was built at a relatively fine spatial resolution (i.e. 0.5° latitude x 0.5° longitude) in contrast to previous global studies (Longhurst, 1998) and was the result of the analysis of six key plankton groups (dinoflagellates, diatoms, small and large copepods, small and large zooplankton other than copepods). However, the study did not decompose the biodiversity into species groups, being based exclusively on spatial and temporal (i.e. 2-month time periods) changes in taxonomic richness.

The main objective of this study is to decompose the biodiversity into species assemblages and to relate them with the partition proposed by Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). Such a work is needed to provide an information on the biotic composition of the ecological units and ecoregions of the North Atlantic and its adjacent seas. Although a study has identified assemblages of calanoid copepods (108 species or taxa), no joint analyses have been performed on both phytoplankton and zooplankton collected from the CPR survey (Beaugrand et al., 2002a). We identified taxonomic assemblages using all taxa (species, genera, or higher taxonomic resolution)

recorded by the CPR survey. All phytoplankton and zooplankton were considered and combined together to be associated into groups of taxa. First, we gridded spatially the abundance of all taxa for each 2-month period (1948-2016). We assumed that spatial variance was much more pronounced than temporal (i.e. year-to-year to decadal) variance. To minimise the potential influence of year-to-year and decadal variability, we performed data smoothing. Then, we mapped the biodiversity of phytoplankton and zooplankton. We used a cluster analysis to examine the relationships among ~300 plankton taxa that had a level of abundance sufficiently high to be analysed. Then, we examined the spatial distribution and annual changes of each assemblage. We used nine environmental parameters to characterise the environmental signature of each assemblage (i.e. a combination of environmental variable that characterises a species group) using a procedure that leads to the display of what we propose to call an environmental chromatogram (i.e. a graphic that identifies where species of an assemblage aggregate along an environmental gradient composed of multiple ecological dimensions). We prefer the term environmental signature (i.e. the environmental regime where species aggregation within an assemblage is highest) instead of the term ecological niche because the latter is usually restricted to the species level (Chase and Leibold, 2003). Finally, we examined the composition of each ecoregion and ecological unit (*sensu* Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b)) in term of assemblage.

## 2.1.2. Materials and Methods

### 2.1.2.1 Physical data

We used a set of nine physical variables: sea surface temperatures, bathymetry, Monthly mean Downward Solar Radiation Flux at surface, macro-nutrients (nitrate, phosphate and silicate), distance to nearest coastlines, density mixed layer depth and sea surface salinity. Information on data sources can be found in Appendix 1.

### 2.1.2.2 Biological Data, the CPR survey

Biological data originated from the Continuous Plankton Recorder (CPR) survey. It is a long-term plankton-monitoring programme currently operated by the Marine Biological association of Plymouth. The CPR is the longest and most extensive programme of that kind in the world. The machine is a high-speed plankton recorder towed behind voluntary merchant ships, called “ships of opportunity”, that filters phytoplankton and zooplankton at a depth of ~7 m (Hays and Warner, 1993). The taxonomic resolution of the data used in this study is shown in Table S1. More information about the CPR survey (advantages and limitations) can be found in Text S1 and in Section 2.1.4.1 of the Discussion. Information on data sources can be found in Appendix 1.

### 2.1.2.3 Mathematical analyses

Six main analyses, all coded in MATLAB, were performed in this study (Figure S3).

Analysis 1. Spatial regularisation. Spatial sampling by the CPR survey is heterogeneous. Therefore, we first created a spatial grid for every plankton species/taxa sampled by the CPR survey for every two-month period using data collected between 1948 and 2016. The spatial grids were identical to the ones used in Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). For each geographical cell of 0.5° of latitude and 0.5° of longitude from 80.5°W to 9.5°E and from 40.5°N to 65.5°N, we calculated average abundance values of each plankton taxa for every two-month period based on the period 1948-2016. The procedure led to a three-dimensional matrix of 304 taxa x 9231 geographical cells (latitudes x longitudes) x 6 two-month periods.

Analysis 2. Spatial smoothing of the gridded data. Abundance data can be highly variable from a geographical cell to another which can be in part attributed to the CPR sampling (e.g. variable seawater filtered by the machine, ship speed) and year-to-year variability (Jonas et al., 2004). We therefore smoothed spatially all two-month abundance grids for each species/taxa. Smoothing was performed by applying a c-order spatial simple moving average:

$$y_{i,j} = \frac{1}{(2c+1)^2} \sum_{s=i-c}^{i+c} \sum_{t=j-c}^{j+c} x_{s,t} \qquad (1)$$

with $1 + c \leq i \leq k - c$ and $1 + c \leq j \leq l - c$, k the number of latitudes ($k$=51) and l the number of longitudes ($l$=181). $y_{i,j}$ is the smoothed abundance value of a species/taxa at the geographical cell corresponding to latitude $i$ and longitude $j$ and $x$ is the original abundance value for a species/taxa. Threshold selection $c$ depends upon the size of geographical cells (here a cell is 0.5° latitude x 0.5° longitude), the noise inherent to the data and the type and the location of spatial structures. By trial and error, we fixed $c$ to 2 as a compromise between noise reduction and potential numerical artefacts. Many values can be missing in some areas and we fixed to 20 the maximum number of missing values allowed to have an estimation.

Analysis 3. Spatial distribution of the taxonomic richness. The spatial distribution of taxonomic richness was mapped using the 3-dimensional matrix resulting from Analysis 1 and 2. For this analysis, abundance was converted into presence-absence using a threshold of 0 (i.e. abundance > 0 means an occurrence). The use of other thresholds did not affect our conclusions. We then summed all presence for phytoplankton and/or zooplankton species or taxa and smoothed the resulting matrix by applying a first-order spatial triple moving average (c=1) to obtain a map of the spatial distribution of taxonomic richness. The analysis was performed for total (i.e. phytoplankton and zooplankton) taxonomic richness, phytoplankton taxonomic richness and zooplankton taxonomic richness. Among the 304 taxa we used, 149 were phytoplankton, 155 were zooplankton; ~60% of plankton taxa were identified at the species level (Figure 19 and Table S1).

Analysis 4. Identification of taxonomic assemblages. We then decomposed the biodiversity into taxonomic assemblages. We first calculated a squared distance matrix (taxa x taxa) using the Hellinger distance coefficient, which is robust to a high number of double zeros (Legendre and Legendre, 1998):

$$Distance(species1, species2) = \sqrt{\sum_{j=1}^{p}\left[\sqrt{\frac{y_{1j}}{y_{1+}}} - \sqrt{\frac{y_{2j}}{y_{2+}}}\right]^2} \qquad (2)$$

With $y_{1j}$ the abundance of the species/taxa 1 in geographical cell *j*, $y_{1+}$, the total abundance of species/taxa 1 across all geographical cells, $y_{2j}$ the abundance of species/taxa 2 in geographical cell *j*, $y_{2+}$ the total abundance of species/taxa 2 across all geographical cells.

The distance coefficient was calculated between species/taxa on the basis of their patterns of abundance in space (i.e. all geographical cells covering the Atlantic Ocean) and time (i.e. two-month period).

Prior to the use of Hellinger's distance coefficient, abundance data were transformed using the function $log_{10}(x + 1)$, a procedure frequently applied to the CPR data that also clearly limits the Euclidean's distance paradox.

A cluster analysis was subsequently applied using Ward's minimum variance method (Legendre and Legendre, 1998). The resulting dendrogram is in Figure S4. Spatial distribution of all assemblages was subsequently mapped by applying the same procedure to represent the spatial distribution of taxonomic richness (Analysis 3) but at the assemblage level. For each assemblage, we mapped the percentage of species/taxa aggregation (i.e. percentage of co-occurring taxa of a given taxonomic assemblage in a cell) by averaging the maps built for each 2-month period (1948-2016; Figure 20). The procedure was also applied for each two-month period to examine seasonal changes in each taxonomic assemblage (Figure 21).

Taxonomic composition of each assemblage was characterized by the use of pie charts in two ways:

- First, we determined the number of species/taxa belonging to six categories: (i) diatoms, (ii) dinoflagellates, (iii) other phytoplankton, (iv) large copepods, (v) small copepods and (vi) other zooplankton (Figure S5).
- Second, we determined the number of taxa identified at (i) a species level and (ii) a higher taxonomic resolution (Figure S6).

Analysis 5. Relationships between taxonomic assemblages and both ecoregions and ecological units identified by Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). From Analysis 4, we calculated the percentage of geographical cells with a percentage of taxonomic aggregation (i.e. percentage of co-occurring taxa of a given taxonomic assemblage in a cell) higher than 10% (Figure 22 and Figure S7) and 50% (Figures S8 and S9) in all ecoregions (Figures S7 and S8) and ecological units (Figure 22 and Figure S9) *sensu* Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). This analysis was performed to determine the taxonomic assemblage composition of each ecoregion and ecological unit.

Analysis 6. Estimation of the environmental signature of each taxonomic assemblage. We used a set of nine environmental variables (see Physical data section) covering the whole North Atlantic Ocean and adjacent seas in space and time to characterise the environmental signature of each assemblage by means of what we call an environmental chromatogram, i.e. a graphic that identifies where species of an assemblage aggregate along an environmental gradient composed of multiple ecological dimensions. First, the value of each environmental variable was interpolated using the grid we used for plankton data. With the exception of bathymetry and distance to coast, all data were linearly interpolated for each two-month period. We therefore obtained a matrix 9231 geographical cells x 6 by two-month period for each environmental variable; note that for bathymetry and distance to coast, the same values were repeated for each two-month period. By this way, it was possible to relate the abundance of each species/taxa with any environmental variable in space and time. All values of each environmental variable were then standardised between 0 (i.e. lowest value) and 1 (highest value).

We then divided all environmental values between 0 (lowest values) and 1 (highest values) into 100 categories and calculated the abundance of each species/taxa that corresponded to each environmental category between 0 and 1. The choice of the categories (100) resulted from a compromise between the resolution of the chromatogram and the number of observed environmental values (here 9231 x 6 = 55,386 if we include missing values). Because the number of values was high, we chose 100 categories to improve the resolution. The standardisation of each environmental dimension between 0 and 1 allowed (i) their representation inside a two-dimensional space and (ii) the subsequent characterisation of the environmental signature of each assemblage (Figure 23); we propose to call this graphic an environmental chromatogram. The environmental signature of each assemblage was the result of the ecological niche of all species in that assemblage.

We also characterised the environmental signature of all taxonomic assemblages by considering phytoplankton and zooplankton separately (Figure S10). This comparison was made to check whether signatures were identical. By this way we checked the homogeneity of phytoplankton and zooplankton signatures among each species assemblage. We then compared patterns of environmental signature for phytoplankton and zooplankton by means of a correlation analysis. We used a Spearman correlation coefficient tested by a Montecarlo test using 10,000 permutations to estimate the probability (Jackson and Somers, 1989).

## 2.1.3 Results

Figure 19a shows the spatial distribution of taxonomic richness in the North Atlantic Ocean and its adjacent seas based on 304 phytoplankton and zooplankton species/taxa. A strong gradient was observed between the northern and the southern part of the oceanic basin, with an increasing taxonomic richness towards the equator (we used data with different degree of taxonomic resolution, see Table S1). A second gradient was also observed between the western and the eastern sides of the Atlantic Ocean, with an increasing taxonomic richness eastward. Biodiversity was greatest

south of the 10°C isotherm (magenta line in Figure 19) with taxonomic richness varying between 100 and 140 taxa in the south and between 60 and 80 in the north. Higher taxonomic richness occurred in areas characterised by warmer temperatures (e.g. West European Basin) and high current velocity (e.g. the Gulf Stream). Similar patterns in taxonomic richness were observed for phytoplankton (Figure 19b) and zooplankton (Figure 19c) although maximal taxonomic richness was observed in the North Sea for phytoplankton and in the West European Basin for zooplankton. Zooplanktonic taxonomic richness was higher than phytoplanktonic taxonomic richness in the Gulf Stream extension.

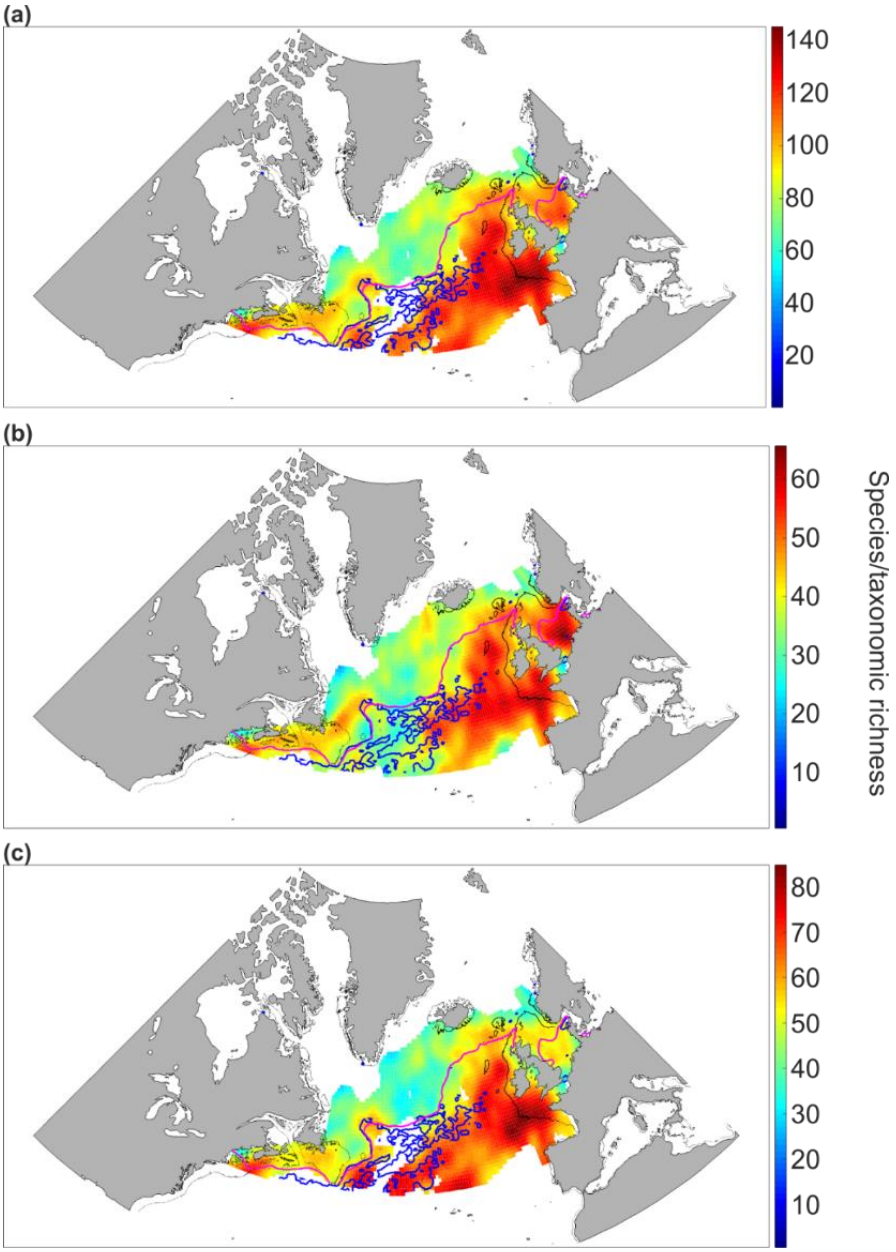



*Figure 19. Spatial distribution of total taxonomic richness (a), phytoplankton taxonomic richness (b) and zooplankton taxonomic richness (c) sampled by the CPR survey in the North Atlantic Ocean. The magenta line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 m.s-1. The black lines denote isobath 200m. We caution that phytoplankton and zooplankton biodiversity maps were based on taxa identified at the species (72.5% for phytoplankton and 47.1% for zooplankton), genus (26.2% and 23.9%) and higher taxonomic (1.3% and 29%) resolutions.*

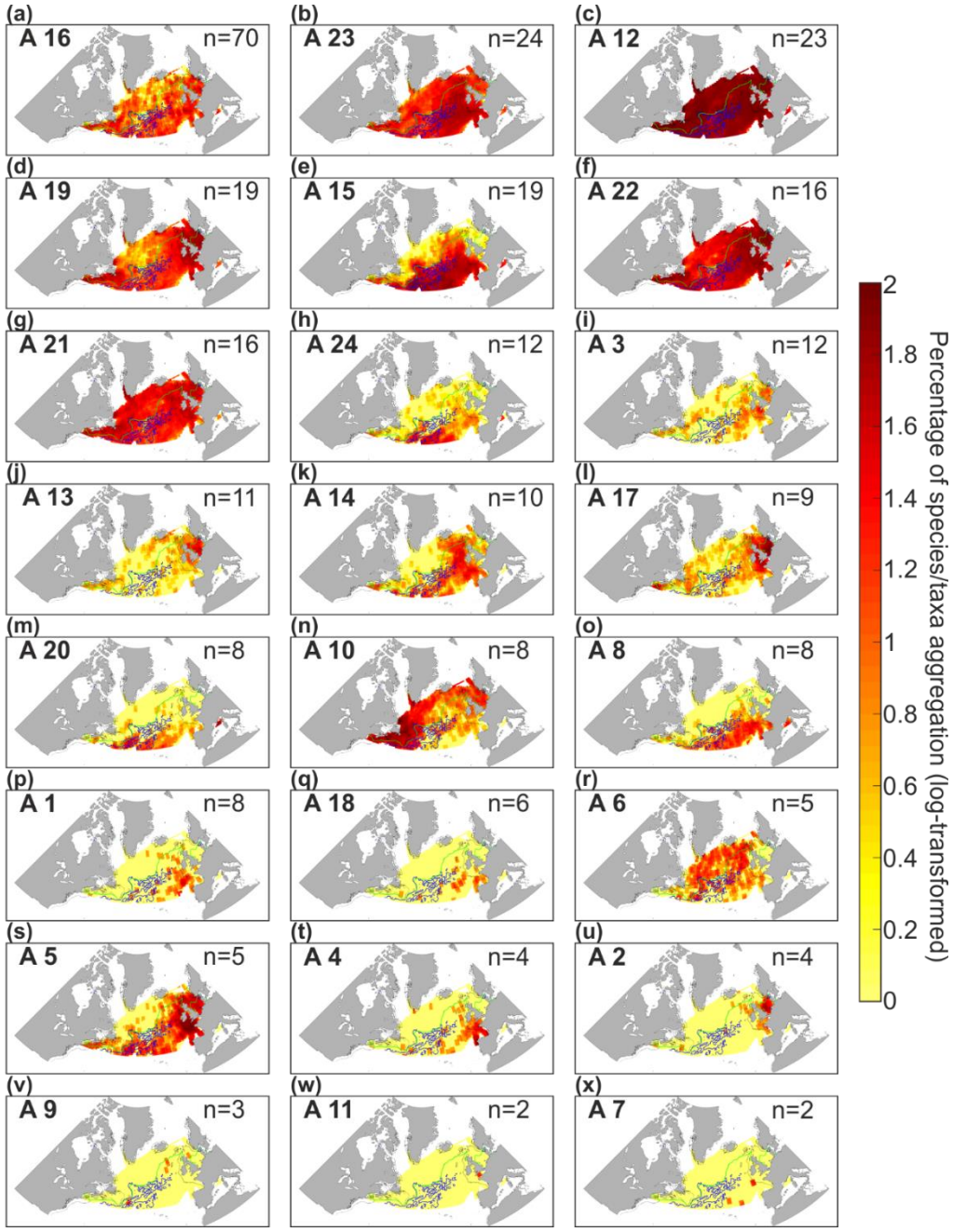



*Figure 20. Mean spatial distribution of taxonomic assemblages identified by the cluster analysis (see Figure S4 and Table S1) and based on an average of 6 2-month period (1948-2016). On the top right of each panel, the number of species/taxa (n) in the assemblage is indicated. The assemblage number, corresponding to the numbers in Figure S4, is displayed in bold on the top left. The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 m.s-1. The black lines denote the isobath 200 m. Panels are classified from a to x by decreasing taxonomic richness.*

Plankton biodiversity was then decomposed into 24 plankton assemblages (Figure 20) by means of a cluster analysis (see Figure S4). We chose a cut-off level of 1.7 in the dendrogram to select a maximum of groups. When thresholds were too high (i.e. >1.7), groups remained too spatially and/or temporally heterogeneous, and when thresholds were too small, we had too many isolated species. Five assemblages (Assemblages 16, 23, 12, 19, 22 and 21, Figure 20a-d and f-g) occurred nearly everywhere in the studied area; they were therefore eurygraphic. All the remaining

assemblages were located south of the 10°C isotherm, except Assemblage 10 (Figure 20n) observed in the northern part of our studied area and Assemblage 6 mainly detected nearly everywhere in the open ocean (Figure 20r). Among the eighteen southern assemblages, three (Assemblages 1, 18 and 4, Figure 20p, q and t) occurred in the West European Basin including the Bay of Biscay and four (Assemblages 3, 13, 17 and 2, Figure 20i, j, l and u) were mainly located over continental shelves. Assemblage 5 (Figure 20s) was located south of the Oceanic Polar Front (OPF) *sensu* Dietrich 1964 (Dietrich, 1964), over continental shelves and in open ocean. The remaining assemblages (i.e. Assemblages 15, 24, 14, 20, 8, 9, 11 and 7) mainly occurred south of the 10°C isotherm in the open ocean (Figure 20e, h, k, m, o, v, w and x).

Annual changes were examined for all assemblages. We only show such changes for Assemblages 8 (Figure 21a-f) and 10 (Figure 21g-l) as examples. Both assemblages exhibited strong seasonal variations in taxonomic aggregation throughout the year. Assemblage 8 was virtually absent from November to April in the surface. From May to October, there was a substantial increase in taxonomic aggregation south of the OPF (see the 10°C isotherm in Figure 21). Assemblage 10 (Figure 21g-l) was observed over continental shelves (American and European) for all 2-month periods. From March to October, this assemblage spread over the oceanic regions north of the 10°C isotherm in the Subarctic Gyre.

Taxonomic composition and resolution of each assemblage are displayed as pie charts (Figures S5-6). These last figures show that 18 out of 24 assemblages (75% of the assemblages) were composed of at least 50% of taxa identified at the species level. Each assemblage was indicative of one or more ecological units *sensu* Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) (hereafter termed EUs). The name of each unit is indicated in Figure S1. For this analysis, we used a threshold of taxonomic aggregation of 10% (i.e. species or taxa that had a percentage of co-occurrence higher than 10% in a given geographical cell for a given assemblage). Other thresholds were tried and did not alter substantially our conclusions when they were fixed below 50% (see Figures S8-9). Four assemblages (Assemblages 23, 12, 22 and 21, Figure 22b-c, f and g) occurred in almost 80% of the geographical cells composing all ecological units. Assemblages 16 (Figure 22a) and 19 (Figure 22d) mainly occurred around the British Isles (e.g. the Cold-Temperate Neritic and the Cold-Temperate Shallow Neritic EUs), although they were also detected in oceanic EUs south of the 10°C isotherm (e.g. the Gulf-Stream Extension, the Northern Sub-tropical and the Oceanic Warm-Temperate EUs (see the nomenclature in Figure S1). Assemblages 15, 24 and 20 (Figure 22e, h and m) were also present in the last three EUs but mostly in the southern ones, i.e. the Gulf Stream Extension and the Northern Sub-Tropical EUs. Assemblages 10 and 6 (Figure 22n and r) were mainly located in northern EUs (e.g. the Polar oceanic EU). Many assemblages (3, 13, 17, 5, and 2 in Figure 22i, j, l, s and u) occurred in EUs covering the continental shelf (e.g. the Cold-Temperate Shallow Neritic, the Cold-Temperate Neritic and the Ocean-Influenced Cold-Temperate Neritic EUs) as well as the shelf-edge for Assemblage 5 (Figure 22s) (i.e. the Diverse and Productive Oceanic and Temperate and the Pseudo-Oceanic Warm-Temperate EUs). Remaining assemblages, i.e. 14, 8, 1, 18, 4, 9, 11 and 7 (Figure 22k, o, p, q, t, v, w and x) occurred mostly in EUs close to or over

the European shelf-edge, e.g. the Diverse and productive Oceanic Temperate EU, the Pseudo-Oceanic Warm Temperate EU and the Mixed Coastal-Oceanic Highly-Seasonally dynamical EU. Therefore, all assemblages were linked to specific EUs of the North Atlantic Ocean and its adjacent seas.

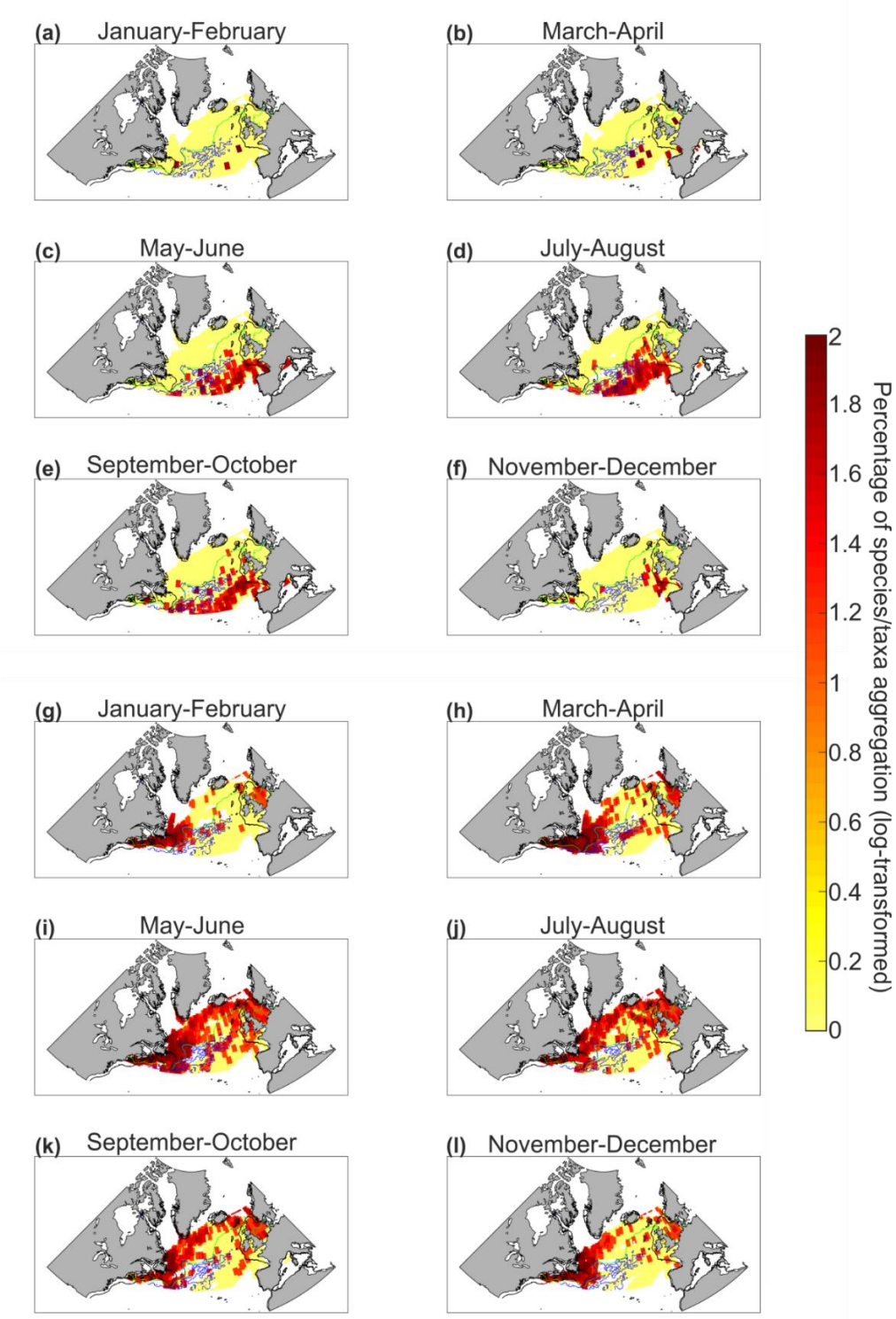



*Figure 21. Seasonal changes in the spatial distribution of Assemblage 8 (panels a-f) and 10 (panels g-l). The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 m.s-1. The black lines denote the isobath 200 m.*

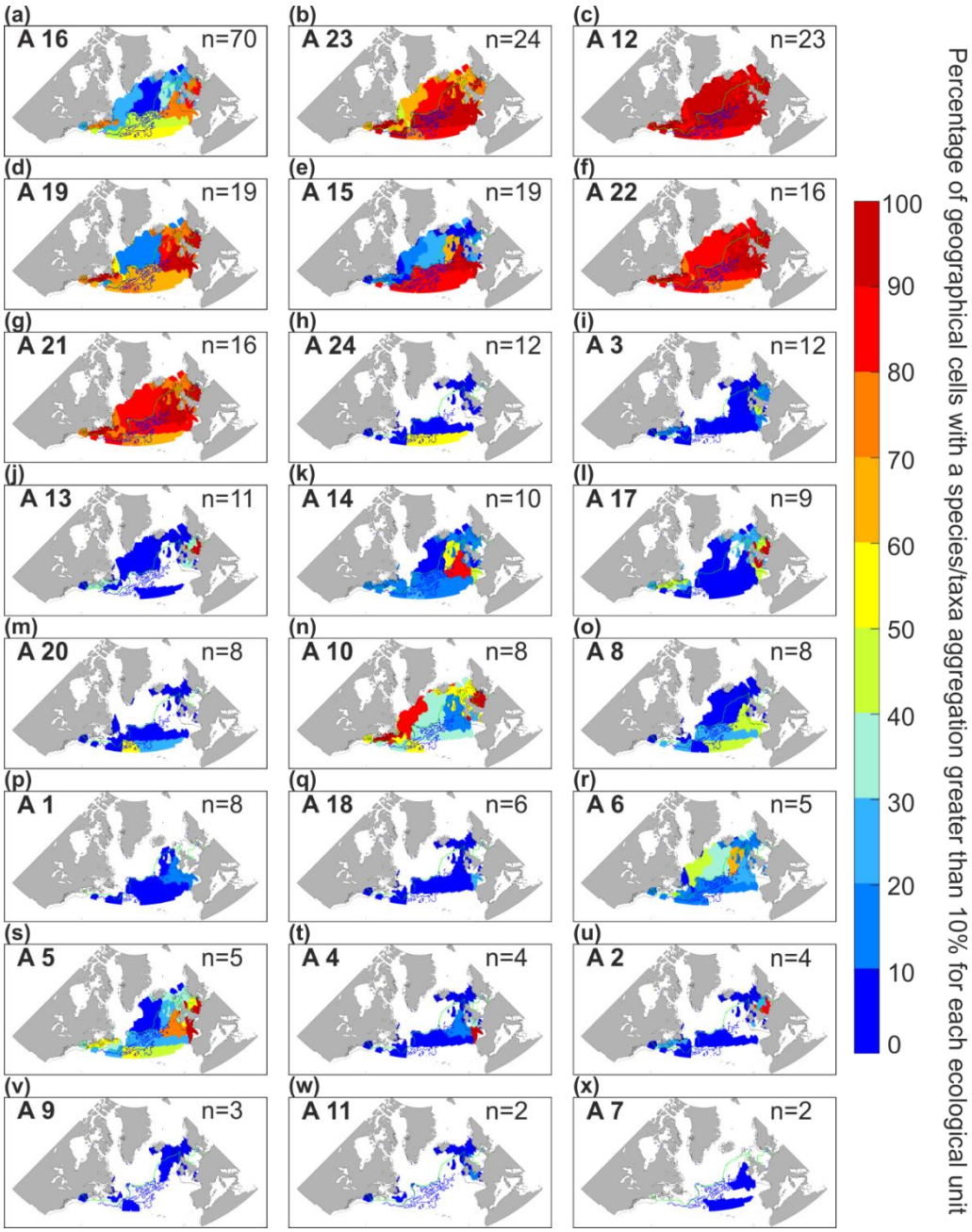



*Figure 22. Percentage of taxonomic aggregation (i.e. number of species or taxa of the same assemblage) greater than 10% in each ecological unit as defined by Beaugrand et al., (2019) (9). Here, the percentage of species/taxa aggregation is used to identify an assemblage characteristic of an ecological unit. On the top right of each panel, the number of species/taxa (n) in the assemblage is indicated. The assemblage number, corresponding to the numbers in Figure S4 and in Figure 20, is displayed in bold on the top left. The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 m.s-1. The black lines denote isobath 200 m. Panels are sorted from a to x by decreasing taxonomic richness.*

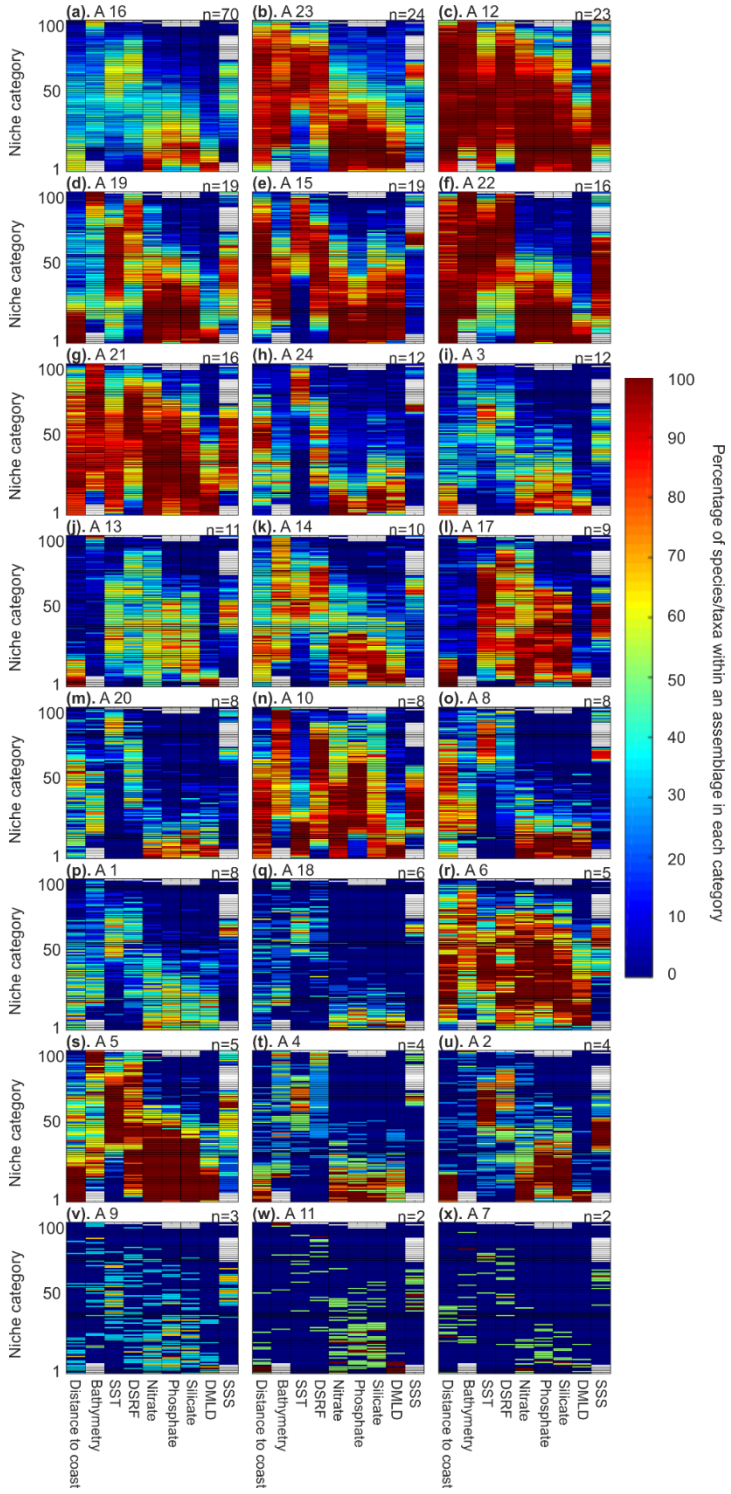



*Figure 23. Environmental signature of the 24 species/taxa for all 24 assemblages. For each figure, from a to x, column corresponds to all the values take by an environmental variable (e.g. distance to coast or silicate). For each variables (column) all values were divided into 100 categories standardised between 0 and 1, bottom categories (0) corresponding to the smallest values taken by an environmental variable. Colour indices denote the percentage of species/taxa of an assemblage found into a category. Red colour indicates that the majority of the species/taxa composing the assemblages are found in these environmental categories. Blue colour indicates that no species/taxa or a few, were found in these environmental categories. Panels are classified from a to x by decreasing taxonomic richness. The number at the top left of each panel corresponds to the assemblage number (see Figure 20 and Figure S4) and the number at the top right (n) indicates the taxonomic richness of each assemblage.*

Similar conclusions were reached for ecoregions, although results were more difficult to interpret as ecological units originate from the aggregation of ecoregions (see Figures S8-9); therefore, the number of ecoregions (40) is higher than the number of ecological units (13) and it is more difficult to find a pattern on a higher number of units.

We characterised the environmental signature of each assemblage by means of what we propose to call an environmental chromatogram (Figure 23). The figure shows the percentage of co-occurring species/taxa of an assemblage in each environmental category standardised between 0 and 1 (0 being the smallest category and 1 the highest). By this way, we represented nine ecological dimensions (i.e. ecological variables) into two. Each assemblage had its own environmental signature. Assemblages 23, 12, 22, 21 and 6 (Figure 23b, c, f, g and r) occurred in a large number of categories for almost all environmental variables; the species/taxa that compose these assemblages were euryoecious. Other assemblages such as number 16, 24, 3, 20, 1 and 18 (Figure 23a, h, i, m, p and q) were observed in a restricted number of environmental categories; they are more stenoecious.

Then we compared the environmental signature of phytoplankton and zooplankton species/taxa composing each assemblage in order to check whether signatures were identical. By this way we verified the homogeneity of phytoplankton and zooplankton signatures among each species assemblage. Expectedly, we found the same environmental signature for phytoplankton and zooplankton of the same assemblage (Figure S10). We also compared the environmental signature of phytoplankton and zooplankton for each assemblage by means of a Spearman correlation coefficient (Table S2 and Figure S10). All correlations were significant ($p < 0.01$; Table S2). With the exception of four assemblages (1, 4, 9 and 11) that had a correlation lower than 0.5, all assemblages had correlations above 0.5. This result confirmed that the environmental signature of phytoplankton and zooplankton was similar for all assemblages, although the strength of these similarities varied from one assemblage to another.

## 2.1.4. Discussion

### 2.1.4.1. Potential limitations

Our study has potential limitations related to the nature of the CPR data and to our methodology. First, the CPR is not a perfect sampling mechanism. It underestimates some components of the plankton, e.g. large plankton like fish larvae and delicate gelatinous plankton. Due to the mesh size (~270 µm) of the CPR silk (Jonas et al., 2004), some organisms are only semi-quantitatively recorded and abundance of small species is probably underestimated when compared to other water sampling methods. Despite this bias, the proportion of the population that is retained by the CPR silk reflects the major changes in abundance, distribution and specific composition, i.e. the percentage retention is roughly constant within each species even with very small-celled species (Edwards et al., 2006).

Second, taxonomic identification of plankton has evolved since 1948, which might have affected our results. Nevertheless, on the 304 species/taxa considered in our study, only 27 underwent a change in taxonomic resolution (Batten et al., 2003), which represented less than 9%. Moreover, the taxonomic names recorded in the database have mostly remained unchanged throughout years (Richardson et al., 2006). Some zooplankton species/taxa can also be recorded twice by the CPR taxonomists during the identification laboratory process. The only species recorded in eyecount and traverse analyses in our study was *Centropages chierchiae* and the two categories were grouped in the same assemblage (i.e. Assemblage 4, Figure 20t). When the same taxon is recorded twice (e.g. Chaetognatha), it is often because it includes different species or developmental stages (e.g. *Euphausiacea* nauplii and *calyptopis*) (Richardson et al., 2006). Because this information was also important for taxonomists and scientists working on the CPR survey, we chose to keep them in our analyses. The choice to keep all categories had no effect on the species assemblages because of the use of the Ward algorithm that minimizes the intra-group variance (Legendre and Legendre, 1998).

Third, sampling by the CPR survey is also restricted to the surface water (~6.5m in depth) (Hays and Warner, 1993), which might affect our perception of how ecosystems are organized. However, a recent study showed that the seasonal and diel patterns in the abundance of *Calanus finmarchicus* at surface were positively correlated to patterns of abundance observed at 100m (Helaouët et al., 2016). Despite these limitations, it has been shown that the CPR gives a correct picture of both temporal (i.e. seasonal and diel scale) and spatial (i.e. regional to basin scale) changes in plankton (Batten et al., 2003; Richardson et al., 2006).

It is widely recognised that all plankton sampling systems have their own limitations and nuances and that all underestimate abundance to some degree. However, it is important to note that the CPR survey is the only scientific monitoring programme of that kind in the world, with no equivalent existing program. It covers an important time scale from 1948 to present, still active, with a large spatial scale covering the whole North Atlantic Ocean and its adjacent seas such as the Channel, the Celtic Sea and the North Sea. The CPR is now "the most extensive long-term survey of marine organisms in the world" (Reid et al., 2003).

Potential limitations also arose because of the methodological choices we made. First, we assumed that the spatial variance was more pronounced than temporal (i.e. year-to-year and decadal) variance; we did not consider year-to-year to decadal variability. This assumption was needed to cover as fully as possible the North Atlantic Ocean and its adjacent seas. The same assumption was made in Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) to propose a new partition of the North Atlantic Ocean. The effect of temporal variability was to inflate local spatial variance. To reduce this effect, we spatially smoothed the data (see Analysis 2 in Materials and Methods) prior to conducting other analyses such as the cluster analysis and the identification of environmental signatures. This assumption had no effect on the identification of species assemblages because the environmental signature of assemblages was based upon their spatial and 2-monthly aggregation for the period 1948-2016. Environmental

signatures should be stable at a decadal scale because of niche conservatism (Crisp et al., 2009, p. 2009). Furthermore, the lack of consideration of year-to-year to decadal variability did not affect our comparison with the ecological units and ecoregions of Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) because we considered the same time period. However, the spatial distribution of species assemblages identified in our study is likely to change at a decadal scale with large-scale hydroclimatic variability (e.g. the Atlantic Multidecadal Oscillation or AMO) (Faillettaz et al., 2019) or global climate change (IPCC, 2013).

Second, we considered different levels of taxonomic resolution (e.g. species, genera, order) when we identified the species assemblages. We thought it was important to consider those different levels inside the same analysis because some taxa enable the clear identification of some key ecoregions. For example, Gammaridae, Cumacea and Mysidacea are mainly found over the continental shelves (Figure 20l). It was also important to consider different developmental stages (e.g., Euphausiacea nauplii and calyptopis) because they are key for ecosystem trophodynamics (Kirby et al., 2008). For example, most eggs and larvae were clustered in Assemblage 19, which was composed of species or taxa mainly abundant south of the Polar Front (Figure 20d). Although the analysis that leads to the identification of species assemblages was not affected by merging different types of entities, this was not so for the maps of biodiversity we present. Our maps therefore display taxonomic richness and not species richness. However, merging different taxonomic entities inside the same analysis did not alter our perception of the spatial difference in biodiversity among regions because patterns observed in this study are close to those observed in other studies that focussed only on calanoid copepods (Beaugrand et al., 2000c), *Ceratium*, diatoms (Beaugrand et al., 2010) and phytoplankton (Righetti et al., 2019).

Third to produce abundance maps, we spatially regularized and smoothed the CPR data for each two-month period. Our procedure gave estimations of gridded abundance (e.g. *Calanus finmarchicus*) similar to that obtained from kriging (Planque et al., 1997), the inverse distance method (Beaugrand et al., 2002a) or spatial regularization (Helaouët and Beaugrand, 2007). In addition, our biodiversity maps gave similar biodiversity patterns to those originated from kriging, the inverse square distance method and spatial regularization using a first order jackknife procedure (Beaugrand, 1999; Beaugrand et al., 2010, 2000c).

### 2.1.4.2. Factors contributing to the large-scale pelagic biodiversity patterns

We provide a map of plankton taxonomic richness based on both phytoplankton and zooplankton at the North Atlantic basin scale. We caution that phytoplankton and zooplankton biodiversity maps were based on taxa identified at the species (72.5% for phytoplankton and 47.1% for zooplankton), genus (26.2% and 23.9%) and higher taxonomic (1.3% and 29%) levels. Similar maps have been shown in Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) but the maps we provide here have an improved spatial resolution (2° x 2° *versus* 0.5° x 0.5°) due to the implementation

of the smoothing algorithm into our procedure of spatial regularization (see Analysis 2 in Materials and Methods).

The examination of the biodiversity maps (Figure 19) has revealed two gradients in taxonomic richness in the North Atlantic: (i) a meridional gradient (south to north) corresponding to the latitudinal biodiversity gradient, and (ii) a zonal (west to east) gradient. Thus, we observed the greatest biodiversity in the south-eastern part of our studied zone and the lowest biodiversity in the Subarctic Gyre. A similar pattern has already been observed for phytoplankton (Righetti et al., 2019) and zooplankton (Beaugrand et al., 2001). Higher taxonomic richness observed south of our studied area coincided with warmer sea surface temperatures and to a lesser extent with oceanic circulation, e.g. the Gulf Stream and its extension the North Atlantic Current. Some studies have investigated the relationships between temperature and plankton biodiversity using three taxonomic groups and found clear nonlinear relationships between mean as well as annual variability in temperature and biodiversity (Beaugrand et al., 2010). We have also seen a strong influence of the OPF (identified here by the 10°C isotherm) on biodiversity. These results tend to confirm the influence of the 9-10°C isotherm on plankton as revealed by Beaugrand *et al.,* 2008 (Beaugrand et al., 2008) using three trophic levels: phytoplankton, zooplankton and fish. These results also reflect the strong biodiversity difference between the Polar and Westerlies-Wind biomes *sensu* Longhurst 1998 (Longhurst, 1998).

Because biodiversity is low in the subarctic gyre and higher south of the OPF, our study suggests that temperature may be an important factor controlling plankton biodiversity in the North Atlantic. During the eighties, Colebrook (Colebrook, 1985, 1982) already suggested a relationship between plankton abundance and temperature. More recently, temperature has been identified as a key driver for both phytoplankton and zooplankton biodiversity (Beaugrand et al., 2010). Phytoplankton biodiversity is three time higher in the tropics than in higher latitudes (Righetti et al., 2019). Such a pattern is related to the Latitudinal Biodiversity Gradient (LBG). The LBG has also been observed for zooplankton (Rombouts et al., 2009).

Other secondary factors may also influence biodiversity patterns at a regional scale. Investigating foraminifera biodiversity in the North Atlantic, Ruddiman 1969 (Ruddiman, 1969) stressed that the biodiversity gradient was virtually erased by the strength of the diverse subtropical North Atlantic gyre. To explain phytoplankton biodiversity patterns in the North Atlantic, Righetti *et al.,* 2019 (Righetti et al., 2019) proposed that phytoplankton biodiversity was influenced by a great species turnover resulting from high seasonal variability in wind stress, turbulence and light limitation. Margalef 1978 (Margalef, 1978) highlighted that there are two important parameters in phytoplankton biology: turbulence that controls sedimentation rates and variance in current velocity that affects β diversity, i.e. the differences between local community. An increase in β diversity may also explain why there is a very high taxonomic richness over the Celtic Sea and the Bay of Biscay at the boundary between the continental shelf and the open ocean. The strong spatial variability in the bathymetry in these areas enables the coexistence of oceanic and neritic (meroplankton and holoplankton) species but also pseudo-oceanic species (i.e. species

that occur in the ocean and over the continental shelf but are mainly abundant along the shelf-edge).

Oceanic circulation has a strong regional or local influence on biodiversity. Many authors have provided evidence that regional biodiversity can be highly influenced by surface currents (Beaugrand et al., 2001; Longhurst, 1998; Ruddiman, 1969; van der Spoel, 1994). The influence of oceanic circulation is crucial at the Atlantic Basin scale and the zonal difference in North Atlantic biodiversity is clearly explained by the warmer North Atlantic Current that flows northwards in the eastern part of the North Atlantic (Beaugrand et al., 2002a). More locally, the Gulf Stream and its northward extension the North Atlantic Current bring more warm-water species polewards. The OPF (Dietrich, 1964) has a major influence by separating the low biodiversity of the Polar biome (and the subarctic gyre) from regions of higher biodiversity in the Westerlies-Wind biome (*sensu* Longhurst 1998 (Longhurst, 1998)). This influence is stronger on warm-water species than on cold-water species. The European shelf-edge current has also a major influence on the biodiversity of the Bay of Biscay and the western regions of the British Isles in modulating local upwelling and warm-water advection northwards (Reid et al., 2001). The Labrador Current is characterised by poor biodiversity and the occurrence of a few species such as *Calanus hyperboreus* and *C. glacialis*. In the North Sea, the Flamborough Front has a strong influence on biodiversity by separating stratified waters in the north (i.e. lower taxonomic richness) from mixed waters to the south (i.e. higher taxonomic richness). In the north-eastern part of Georges Bank (shelf-edge at the south-eastern part of Newfoundland), Flemish Cape is characterized by a lower taxonomic richness (i.e. ~70 species/taxa, Figure 19a) and the Northwest Corner (Worthington, 1976), which is located to the north-eastern part of Newfoundland, has a higher taxonomic richness (i.e. ~100 species/taxa). North of the British Isles, the Faroe Current and the associated Iceland Faroe Front (Read and Pollard, 1992) limit the spatial location of annual maximum taxonomic richness (Figure 19a). Southwest of the British Isles, maximum taxonomic richness recorded in the Southwest European Basin (west of the Bay of Biscay) may also originate from hydrographic processes. The area closely corresponds with the northward spreading of Intermediate Mediterranean Water (Käse and Zenk, 1996). This water mass has a maximum influence at ~1000m depth but its range extends from 600 to 2500 m (Käse and Zenk, 1996). Therefore, the influence of hydrographical features resulting from oceanic circulation and in some cases topographic features exerts a strong secondary influence on plankton biodiversity in the North Atlantic by controlling temperature or by controlling more directly plankton advection. Therefore, plankton biodiversity in the North Atlantic Ocean is mainly driven by temperature, oceanic circulation and bathymetry, which have a more local/regional influence.

#### 2.1.4.3. Decomposition of plankton biodiversity into assemblages: relationships with North Atlantic ecological units

In a similar study on calanoid copepods, Beaugrand *et al.,* 2002 (Beaugrand et al., 2002a) applied the IndVal method (Dufrêne and Legendre, 1997) in order to group

species into indicator assemblages. In our study, the IndVal would not work because the ecoregions used to calculate the indicator values are too heterogeneous (i.e. the size of the ecoregions varied strongly), which would inflate the number of indicator values in smaller ecoregions, a numerical artefact described in De Cáceres *et al.,* 2010 (De Cáceres et al., 2010). That is why we divided plankton biodiversity into assemblages by using a cluster analysis, based on a Hellinger metric distance, jointly considering space and time (2-month period) variability in the abundance of all plankton species/taxa. Therefore, an assemblage is here characterised by species/taxa exhibiting similar spatial and temporal patterns of abundance. This work has allowed us to complete the new biogeographic work of Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) made as part the European programme ATLANTOS.

We identified 24 assemblages, each being characterised by their own degree of eurygraphy (Figure 20); some were truly eurygraphic (Figure 20c) while others were stenographic (Figure 20j); some were oceanic (Figure 20e) while others were more neritic or both; some were located south of the OPF while others were detected north of the front. All together, they form a mosaic of taxonomic assemblages with overlapping spatial distribution in many different locations, a likely consequence of their environmental signatures. The examination of the spatial distribution of the 24 assemblages confirmed the existence of three main biomes in the North Atlantic (Figure 22): the Polar, the Westerlies-Wind and the Continental Shelves Biomes (Longhurst, 1998). Each biome was identified by specific taxonomic assemblages having a specific environmental signature. North of the OPF, the Polar biome was characterised by Assemblages 10 and 6, with *Ceratium articum*, *Calanus glacialis*, *C. hyperboreus and Heterorhabdus norvegicus* (Figure 22n and r). South of the OPF, the Westerlies-Wind biome was characterised by many assemblages such as Assemblages 24, 14, 20 and 8, with *Ceratium teres*, *Candacia ethiopica*, *Euchaeta marina* and *Rhincalanus nasutus* (Figure 22h, k, m and o). Assemblages 13, 17 and 2 (Figure 22j, l and u) were more characteristic of the Continental-shelves biome, with *Odontella regia*, *O. mobiliensis*, *Paralia sulcata*, *Isias clavipes* and *Labidocera wollastoni*.

Although many assemblages were found in a specific biome (*sensu* Longhurst 1998 (Longhurst, 1998)), none were characteristic of a specific ecological unit *sensu* Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). In contrast to the terrestrial realm, lack of strong physical barriers associated to changing hydro-climatic conditions may partly explain this observation (van der Spoel, 1994). Spatial coenoclines (i.e. gradients of biocoenoses or communities) we observed in our study (Figure 20, Figure 21 and Figure 22) can be compared to annual succession observed in ecosystems during the year (Romagnan et al., 2015). These biological gradients in space and time are a perfect illustration of the ecological principle of impermanence resulting from the constant biological adjustments that originate from niche-environment interaction (Beaugrand et al., 2019a). Sharp or gradual environmental gradients interact with the niche of each species within a multidimensional space to generate a variety of biogeographic patterns.

We do not think that the overlapping spatial distribution of some assemblages with the partition proposed by Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b) is due

to differences in spatial resolution. Partitioning the pelagic ocean is difficult because of the absence of geographical barriers and because any partition changes over time from daily to decadal scales (Reygondeau et al., 2013). In addition, because of the principle of competitive exclusion of Gause (Gause, 1934), we should expect all spatial ranges and phenologies to be unique (Caracciolo et al., 2021). The necessary corollary of this principle is that any synthetic partition should not be expected to work for all species or species assemblages. Temporal dynamics and plankton dispersal or expatriation are also mechanisms adding further complexity.

We found no obvious relationships between the spatial distribution of the species assemblages, seasonal patterns and taxonomic composition in most cases (Figure 20 versus Figure S5). We also found no obvious relationships between assemblages and taxonomic resolution (Figure S6). However, assemblages occurring in the same ecological units frequently have the same environmental signature, as Assemblages 13 and 17 (Figure 22j and l & Figure 23j and l). We therefore suggest that the environmental signature of each assemblage drives their spatial and temporal patterns. Assemblages with a large environmental signature occurred in almost all the ecological units and thus had a large spatial distribution. This was the case for Assemblages 12 and 21 (Figure 22c and g and Figure 23c and g) which had large environmental signature and occurred in the thirteen ecological units. Therefore, assemblages composed of euryoecious species were eurygraphic.

The environmental signature of a group was characterised by means of an environmental chromatogram. This procedure has enabled a rapid display of the 9-dimensional environmental signature of the 24 assemblages into a two-dimensional space (Figure 23 and S10). The environmental chromatogram separated plankton assemblages according to their degree of environmental tolerance (i.e. degree of euryoecy) and optima (e.g. degree of thermophily). We suggest that the method could also characterise the (multidimensional) ecological niche (*sensu* Hutchinson (Hutchinson, 1957)) of a species where both the amplitude and the optimum of an environmental factor would be represented as a function of the number of ecological dimensions, with the abundance (instead of a percentage of aggregation) as a third dimension (i.e. the colour of the contour plot). The environmental chromatogram is a new method that enables one to display all multidimensional niches composing an assemblage (there are as many niches as species). Such chromatograms allow a rapid display of the environmental signature of an assemblage. Furthermore, the visual comparison of the chromatograms can immediately reveal the degree of overlapping of two environmental signatures and explain why the spatial distribution of two assemblages overlap. For example, Assemblage 8 (Figure 20o) has a spatial distribution that is included inside Assemblage 5 (Figure 20s), a result that is expected from the environmental chromatogram of the two assemblages (Figure 23o and s). To our knowledge, there is no graphical method enabling a clear representation of a multidimensional ecological niche into more than two dimensions. A radar plot can be used but results are less easy to interpret, especially when a large number of species are considered (Reygondeau et al., 2012). Multivariate analyses have been sometimes used (Helaouët and Beaugrand, 2007) but because each component is a linear

combination of different ecological factors the resulting multidimensional niche is more difficult to interpret.

It has been shown that cold-water species (e.g. *Calanus finmarchicus*, Assemblage 12) have a larger niche breadth than their warmer-water counterparts (e.g. *Calanus helgolandicus*, Assemblage 19) (Sunday et al., 2011). Indeed, in our 24 assemblages, we noticed that few were specific of northern ecological units (e.g. Assemblages 10 and 6 in Figure 22n and r). Among them, many had a large spatial range covering sometimes the whole North Atlantic Ocean and its adjacent seas. We suggest that the spatial range of these assemblages illustrates well the Rapoport's effect, which postulates that high-latitude species have a larger geographical range than mid-latitude species (Stevens, 1989).

### 2.1.5. Conclusions

Spatial patterns in plankton biodiversity are the result of environmental factors acting at both global and local scales (Beaugrand et al., 2000c). We have decomposed CPR-sampled plankton biodiversity into 24 taxonomic assemblages using both phytoplankton and zooplankton. These assemblages are interesting because they characterise specific hydro-climatic conditions and so they can be used as biological indicators, either of a substrate-biotope (water-mass) or a stable-biotope (key area) component (Beaugrand et al., 2002a). Their associated specific environmental signatures have allowed us to better characterize the ecosystems of the North Atlantic Ocean and its adjacent seas and to improve our understanding of the arrangement of plankton biodiversity.

Although some assemblages were characteristic of the three biomes observed in our studied areas, none of them were observed specifically in the ecological units identified by Beaugrand *et al.,* 2019 (Beaugrand et al., 2019b). This important result can in part be explained by the regional complexity of the northern part of the North Atlantic Ocean, which exhibits a pronounced hydro-dynamic variability with interwoven substrate and stable-biotope components. An alternative hypothesis is that it could be a universal feature. All species having a unique ecological niche after the principle of competitive exclusion of Gause (Gause, 1934), the corollary is that they are expected to exhibit distinct spatial distributions. We have clearly shown that each taxonomic assemblage had a unique environmental signature and that therefore Gause's corollary can be extended at the community level. In the same way that coenoclines are observed during annual succession and that community reorganisation takes place all the time from year-to-year to multi-decadal scales, coenoclines may occur everywhere in the ocean. Pelagic ecosystems are therefore likely to be more complex than previously envisioned.

*Acknowlegments*

We thank the CPR survey, the programme AtlantOS and the French National Center for Scientific Research (CNRS).

*Appendix 1*

All Continuous Plankton Recorder data is freely available on request by contacting the Marine Biological Association (MBA), United Kingdom. Data requests (Dan Lear: dble@mba.ac.uk).

Sea Surface Temperature (SST) originated from Centennial *in situ* Observation Based Estimates (COBE) SST2 dataset (https://www.esrl.noaa.gov/psd/) (Hirahara et al., 2014). Units were in degree Celsius.

Bathymetry data originated from the BIO-ORACLE V2.0 dataset (Assis et al., 2017; Tyberghein et al., 2012). (http://www.bio-oracle.org/). Bathymetry was expressed in meters.

Monthly mean Downward Solar Radiation Flux at surface (DSRF) originated from the National Centers for Environmental Prediction (NCEP), National Center for Atmospheric Research (NCAR), provided by the NOAA/OAR/ESRL PSD, Boulder Colorado, USA (https://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.html). We used this variable as a proxy of PAR (Photosynthetically Active Radiation) because it covered the northern part of our studied zone. Unit was in $W/m^2$.

Nutrients data, originated from the World Ocean Database (WOD) (https://www.nodc.noaa.gov/OC5/WOD/pr_wod.html). We used three variables: (i) nitrate, (ii) phosphate and (iii) silicate concentration. Data were expressed in µmol/kg.

Distance to nearest coastline was provided by the University Corporation for Atmospheric Research (UCAR), community programs, Boulder Colorado, USA (http://oos.soest.hawaii.edu/thredds /ncss/dist2coast_1deg_ocean/dataset.html). Units were in kilometres.

Density Mix Layer Depth (DMLD) originated from Monthly Isopycnal and Mixed-layer Ocean Climatology (MIMOC) provided by NOAA (Johnson et al., 2012). Units were in dbar.

Sea Surface Salinity (SSS) originated from NCPE Global Ocean Data Assimilation System (GODAS) dataset, provided by NOAA ESRL, Physical sciences division, Boulder Colorado, USA. Units were in kg/kg. (https://www.esrl.noaa.gov/psd/).

## 2.1.6. Supplementary Files

**Supplementary Text**

**Text S1. CPR limitations.**

The Continuous Plankton Recorder (CPR) survey is a long-term plankton-monitoring programme, operated currently by the Marine Biological association of Plymouth. The CPR is the longest and most extensive programme of that kind in the world, started in 1931, with minor modifications to the mechanism after 1948. After a break during World War 2, zooplankton sampling has not been changed since 1948, and phytoplankton since 1958 (Warner and Hays, 1994). The machine is a high-speed plankton recorder, towed behind voluntary merchant ships, called “ships of opportunity”, operating at a depth of approximately 7 m, filtering phytoplankton and zooplankton (Hays and Warner, 1993).

A sampling road effect may affect phytoplankton abundance data. Usually phytoplankton cells are smaller than zooplankton species so their abundance is greater; this is

an expression of Damuth's law (Damuth, 1981), which links body size of an organism to its species population density by an inverse relationship, i.e. smaller organisms tend to have a greater population density.

The CPR survey currently records ~1000 plankton entities (many to species level) in routine taxonomic analysis dating back over multi-decades. Due to the mesh size of CPR silks, many plankton species are only semi-quantitatively sampled owing to the small size of the organisms. In the case of phytoplankton there is thus a bias towards recording larger armoured flagellates and chain-forming diatoms and that smaller species abundance estimates from cell counts will probably be underestimated in relation to other water sampling methods. However, the proportion of the population that is retained by the CPR silk reflects the major changes in abundance, distribution and specific composition, i.e. the percentage retention is roughly constant within each species even with very small-celled species (Edwards et al., 2006). Potential under estimation of zooplankton abundances has recently been thoroughly statistically explored (Helaouët et al., 2016) who found that while the CPR survey does underestimate abundance in some cases the CPR survey does give a correct picture of both temporal (i.e. seasonal and diel scales) and spatial (i.e. regional to basin-scale) changes in zooplankton taxa in this case the species *Calanus finmarchicus*. The Hélaouët et al. study also showed that while the CPR sampling is restricted to the surface waters ~10 m in depth the seasonal and diel patterns of abundance of *C. finmarchicus* were positively correlated to patterns of abundance to a depth of 100 m.

The CPR is not a perfect sampling mechanism by any means and it will underestimate components of the plankton, e.g. large plankton like fish larvae and delicate gelatinous plankton. This has been well documented and users of the data are advised of the CPR's limitations (Richardson et al., 2006). It is also widely recognised that all plankton sampling systems have their own limitations and nuances and all underestimate abundance to some degree and that the varying mechanisms are not always directly comparable (Owens et al., 2013). However, in the case of krill, as in this study, the CPR may be underestimating the larger species due to the small sampling aperture (12.7 × 12.7 mm) of the CPR mechanism. As estimated in Silva et al. (Silva et al., 2014), ~87% of the larvae and juveniles of krill are captured by the CPR but it is less efficient at catching larger krill, therefore, the CPR krill data may not represent absolute abundances of adult krill but will provide indices of the larval and juvenile stages of the three most common euphausiid species in the North Atlantic as well as the adult stages of the smallest species (*T. longicaudata*). It is likely that the large adult stages of the euphausiid species *M. norvegica* are underestimated in CPR sampling. Taking the northern North Atlantic Ocean as a whole, as well as within the specific area of the present study, the most numerically abundant near-surface species is *T. longicaudata*, making it likely that our findings are representative of real changes in euphausiid populations. However, further north, in areas such as the Barents Sea (not covered in this study), where the species *M. norvegica* can dominate, the CPR sampling is less likely to reflect real changes in the adult euphausiid populations. It is not known whether certain taxa such as krill are capable of gear avoidance and in some cases this might be a bias for some fast moving taxa, however, since CPR sampling is conducted at high-speeds (up to 20 knots), gear avoidance is considered to be fairly minimal (Owens et al., 2013). A detailed study has been conducted on flow rate and ship speed on CPR sampling (Jonas et al., 2004) given that the speed of the ships has, in some circumstances, increased since the 1960s, which may impact sample efficiencies. However, no significant correlation was found between the long-term changes in the speed of the ships and two commonly used indicators of plankton variability: the Phytoplankton Colour and the Total Copepods indices. This absence of relationship may indicate that the effect found is small in

comparison with the influence of hydroclimatic forcing (Jonas et al., 2004). For further details on the technical background, methods, consistency, and comparability of CPR sampling, see (Batten et al., 2003).

It is important to note that the CPR survey is the only scientific monitoring programme of that kind in the world, with no equivalent existing program. It covers an important time scale from 1948 to present, still active, with a large spatial scale covering the whole North Atlantic Ocean and its adjacent seas such as the Channel, the Celtic Sea and the North Sea. It has become "*the most extensive long-term survey of marine organisms in the world*" (Reid et al., 2003).

**Supplementary Figures**

**Figure S1.** Spatial distribution of the ecological units identified by (Beaugrand et al., 2019b).

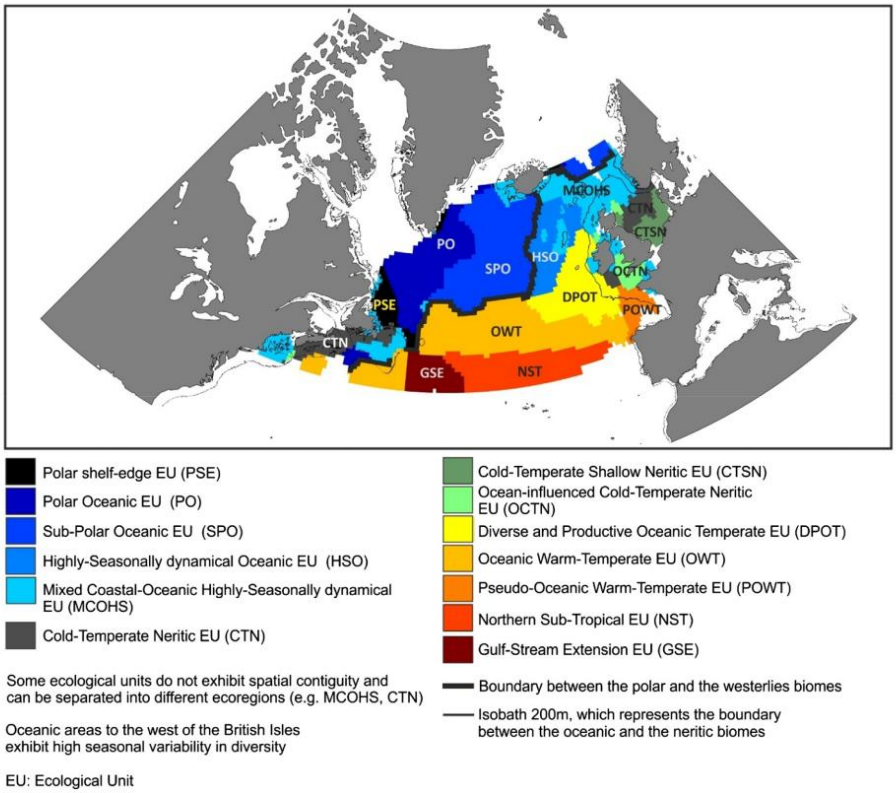


**Figure S2.** Spatial distribution of the ecoregions (panel a) and the ecological units (panel b) of Beaugrand et al., (see also Figure S1). Adapted from (Beaugrand et al., 2019b).

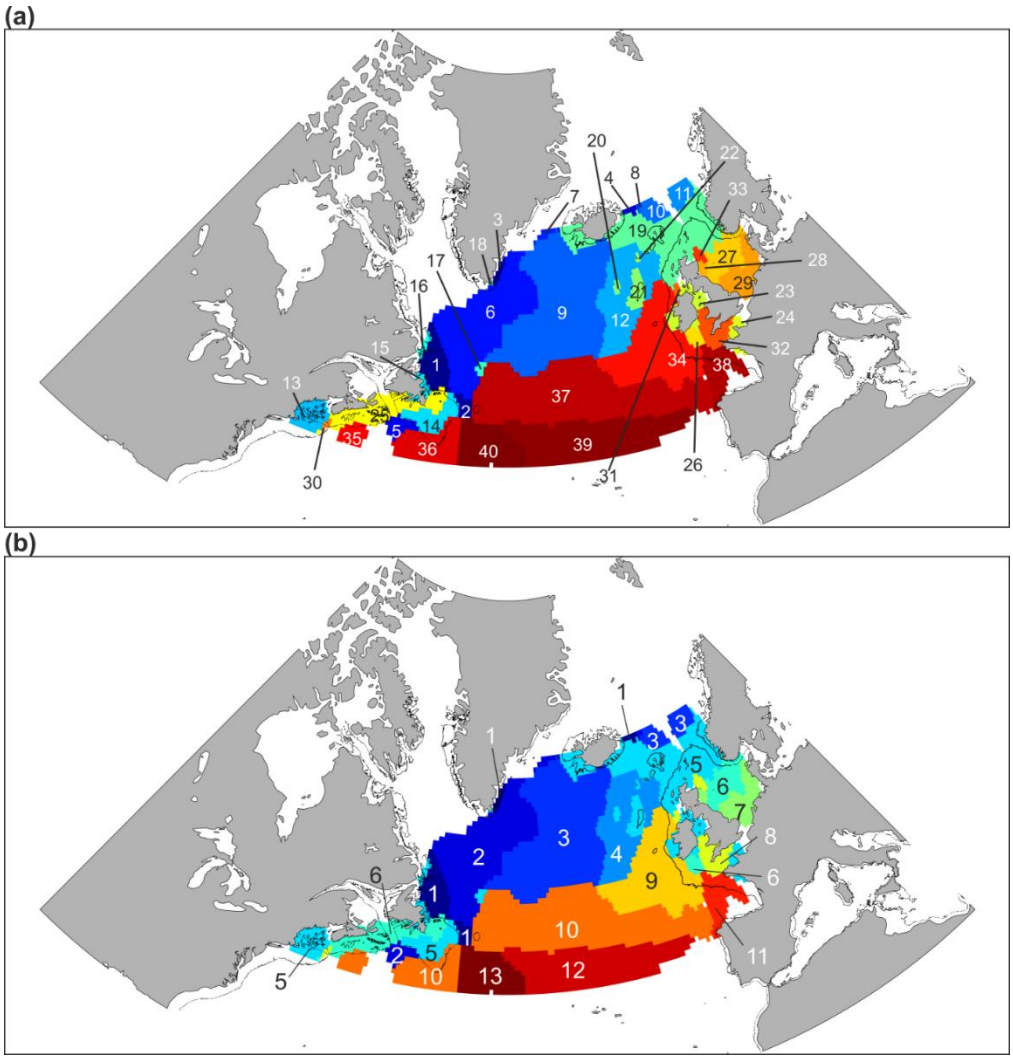


**Figure S3.** Sketch diagram that summarises all analyses performed in this study.

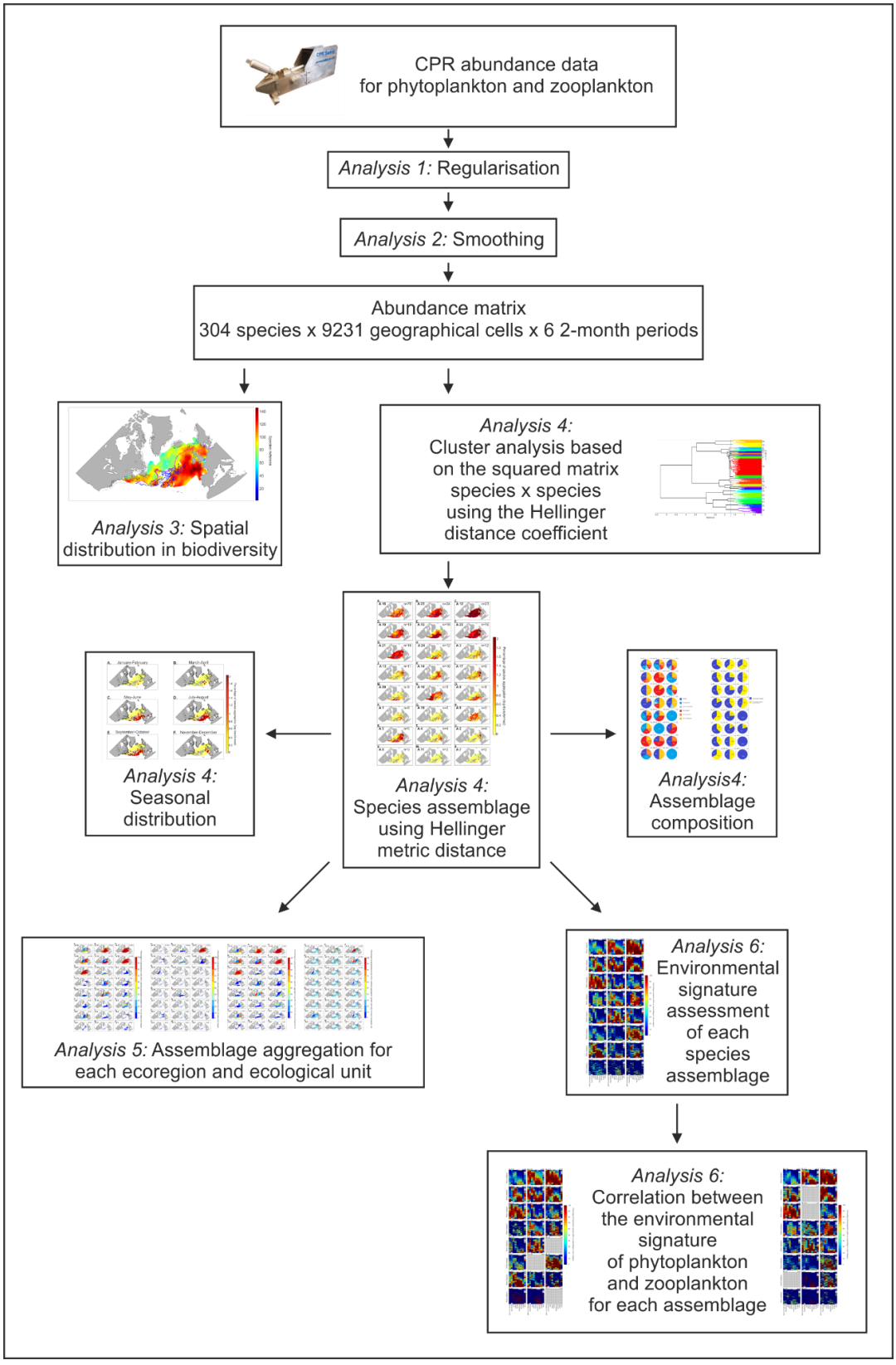


**Figure S4.** Dendrogram based on a cluster analysis performed on 304 species (or taxa) x 9231 geographical cells x 6 2-month period matrix using Ward's algorithm applied on squared matrix of Hellinger distance. A total of 24 assemblages was found at the cut-off level of 1.7. Numbers on the right correspond to the assemblage number displayed in Figure 20. The grey dashed line denotes the position of the cut off level. Colours enable the distinction of a taxonomic assemblage.

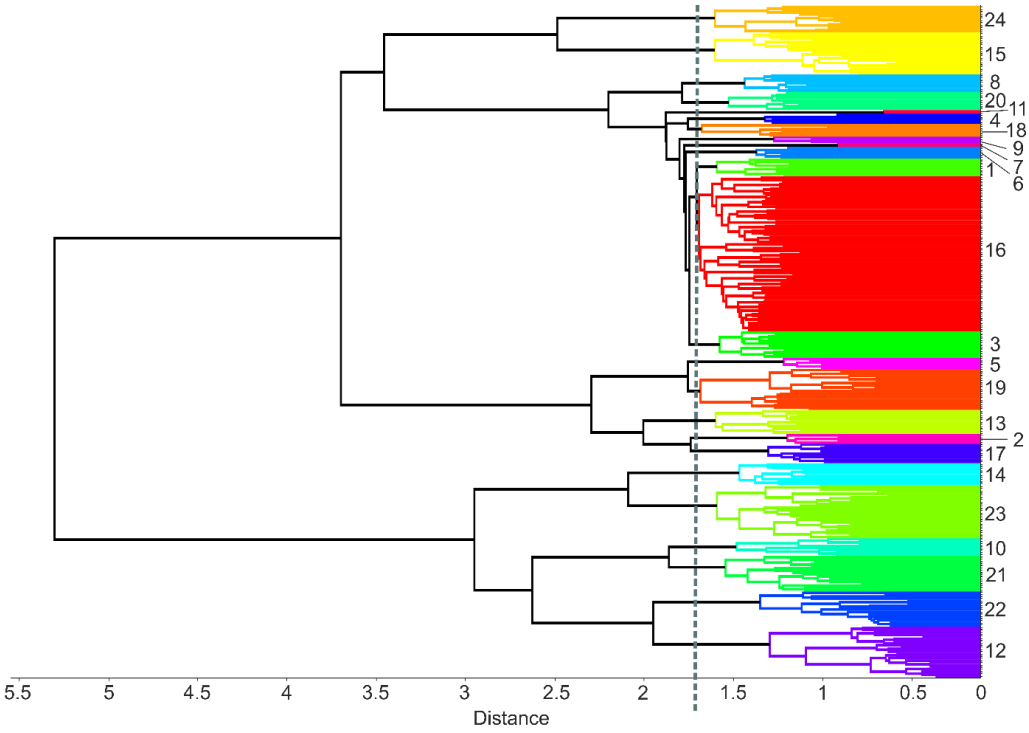


**Figure S5.** Taxonomic composition of the 24 assemblages in Figure 20 expressed as a percentage. Each colour corresponds to one of the taxonomic groups: diatoms, dinoflagellates, other phytoplankton, large copepods, small copepods and other zooplankton. Panels are classified from a to x by decreasing taxonomic richness. The number at the top left of each panel corresponds to the assemblage number (see Figure 20 and Figure S4) and the number at the top right (n) indicates the number of species (or taxa) of each assemblage.

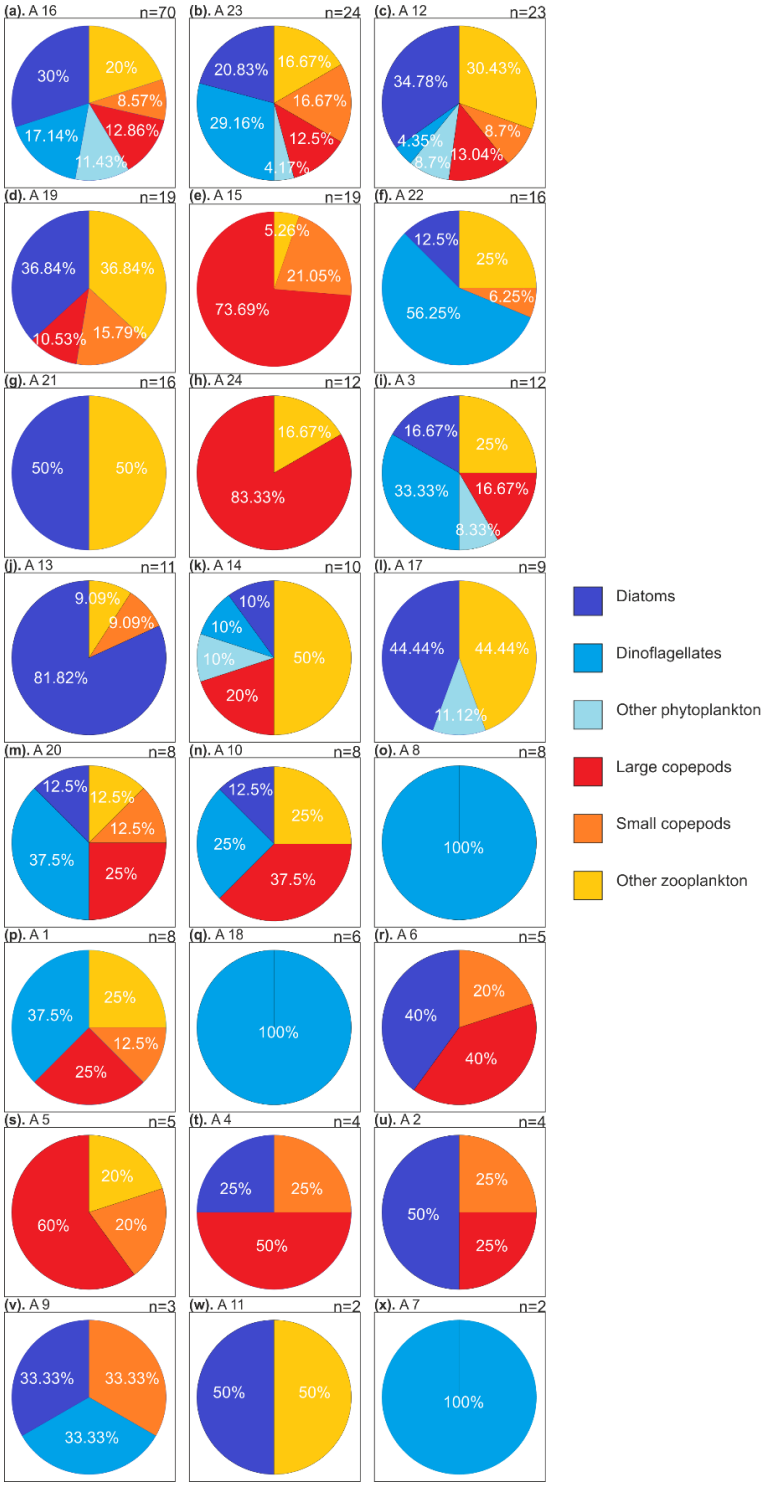


**Figure S6.** Assemblage composition considering only two taxonomic levels: species and coarser taxonomic resolution. Blue colour corresponds to the percentage of taxonomic units identified at the species level and yellow colour corresponds to the percentage of taxonomic units identified at a higher than the species level. As in Figure 20, panels (i.e. assemblages) are classified by decreasing taxonomic richness. The number at the top left of each panel corresponds to the assemblage number (see Figure 20 and Figure S4) and the number at the top right (n) indicates the taxonomic richness of each assemblage.

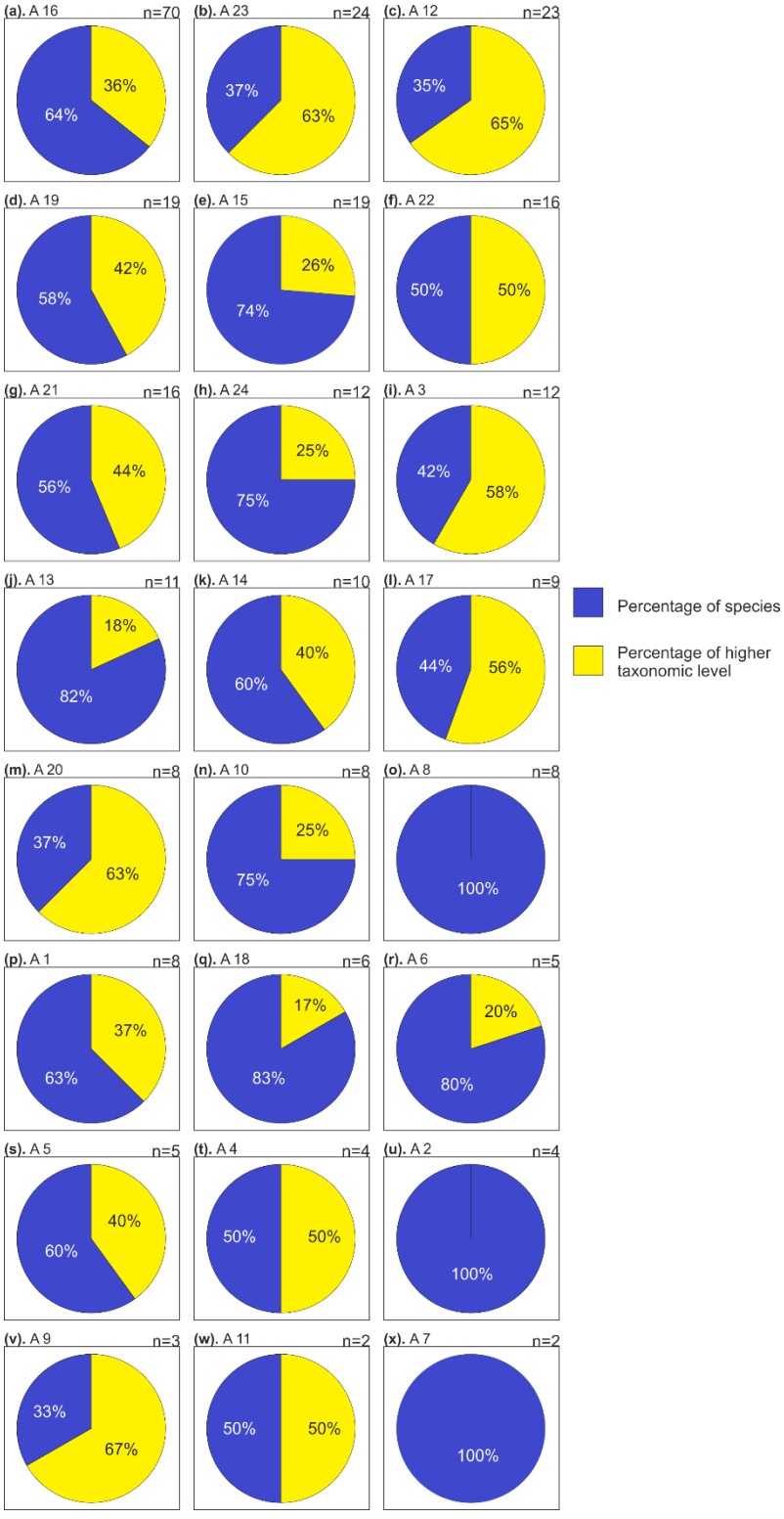


**Figure S7.** Percentage of taxonomic aggregation (i.e. number of species or taxa of the same assemblage) greater than 10% in each ecoregion as defined by (Beaugrand et al., 2019b). Here, the percentage of species/taxa aggregation is used to identify an assemblage characteristic of an ecoregion. On the top right of each panel, the number of species/taxa (n) in the assemblage is indicated. The assemblage number, corresponding to the numbers in Figure S4 and in Figure 20, is displayed in bold on the top left. The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 $m.s^{-1}$. The black lines denote isobath 200 m. Panels are sorted from a to x by decreasing taxonomic richness.

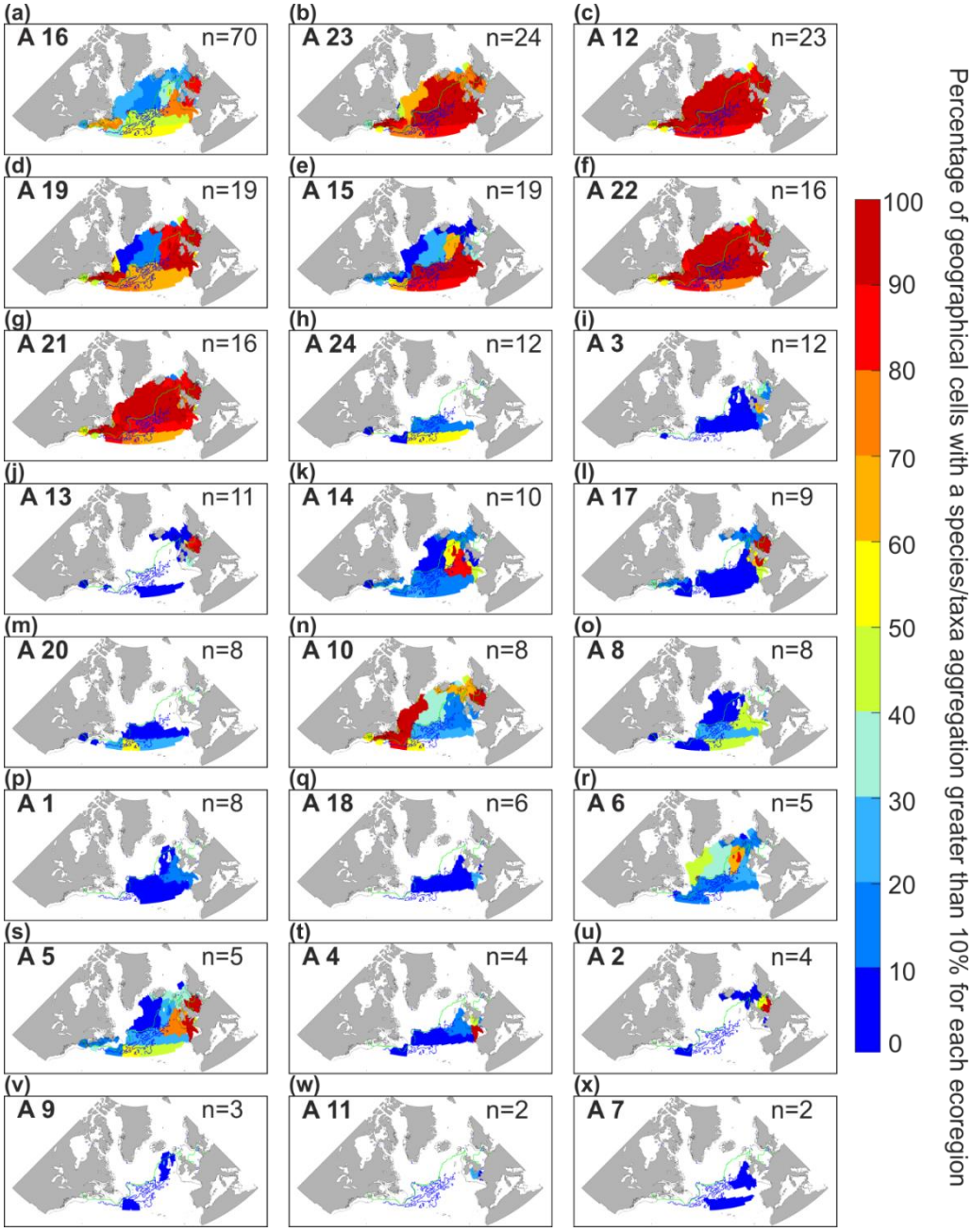


**Figure S8.** Percentage of taxonomic aggregation (i.e. number of species or taxa of the same assemblage) greater than 50% in each ecoregion as defined by (Beaugrand et al., 2019b). Here, the percentage of species/taxa aggregation is used to identify an assemblage characteristic of an ecoregion. On the top right of each panel, the number of species/taxa (n) in the assemblage is indicated. The assemblage number, corresponding to the numbers Figure S4 and in Figure 20, is displayed in bold on the top left. The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 $m.s^{-1}$. The black lines denote isobath 200 m. Panels are sorted from a to x by decreasing taxonomic richness.

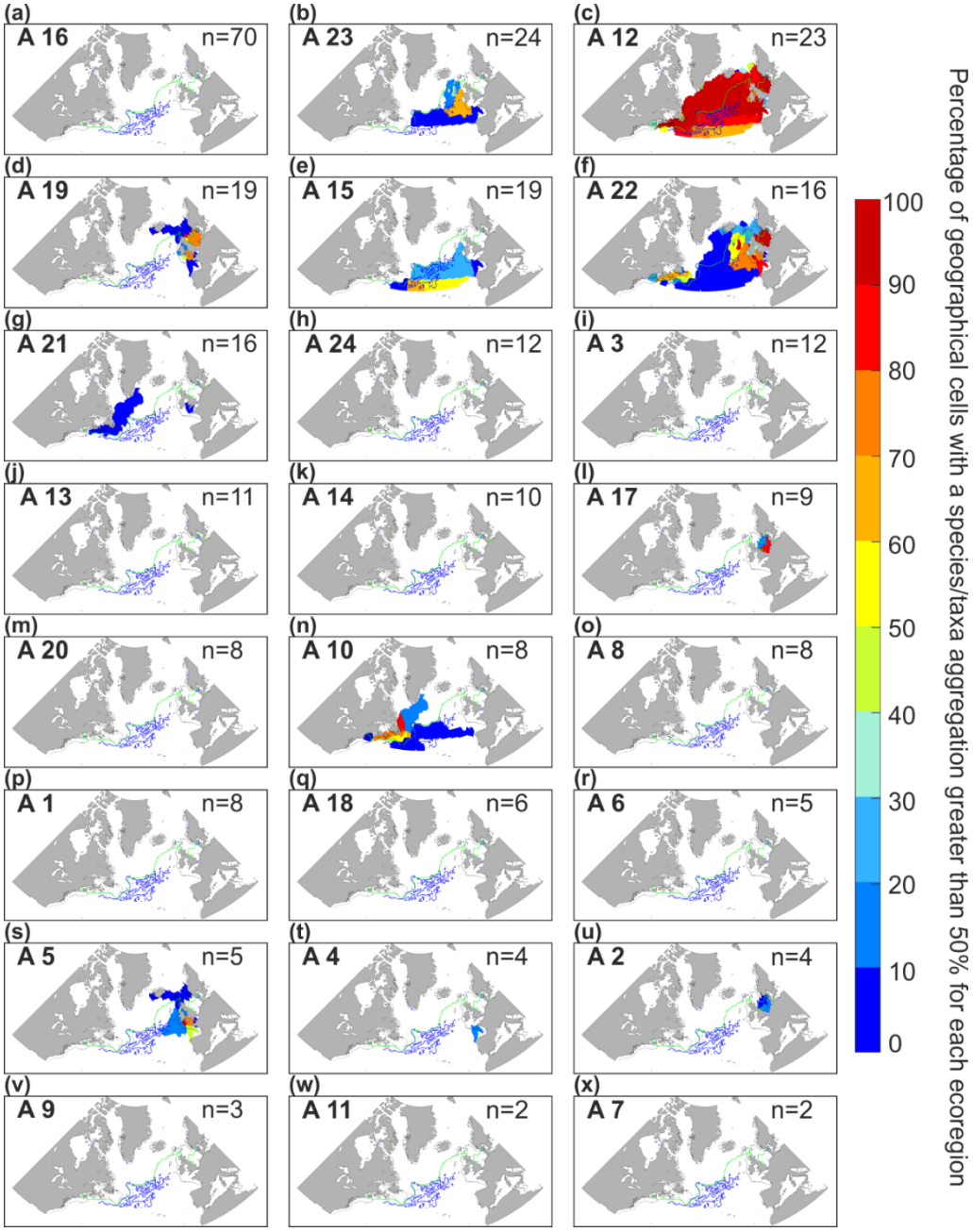


**Figure S9.** Percentage of taxonomic aggregation (i.e. number of species or taxa of the same assemblage) greater than 50% in each ecological unit as defined by (Beaugrand et al., 2019b). Here, the percentage of species/taxa aggregation is used to identify an assemblage characteristic of an ecological unit. On the top right of each panel, the number of species/taxa (n) in the assemblage is indicated. The assemblage number, corresponding to the numbers in Figure S4 and in Figure 20, is displayed in bold on the top left. The green line corresponds to isotherm 10°C. The blue lines correspond to current velocities from 0.5 to 2 $m.s^{-1}$. The black lines denote isobath 200 m. Panels are sorted from a to x by decreasing taxonomic richness.

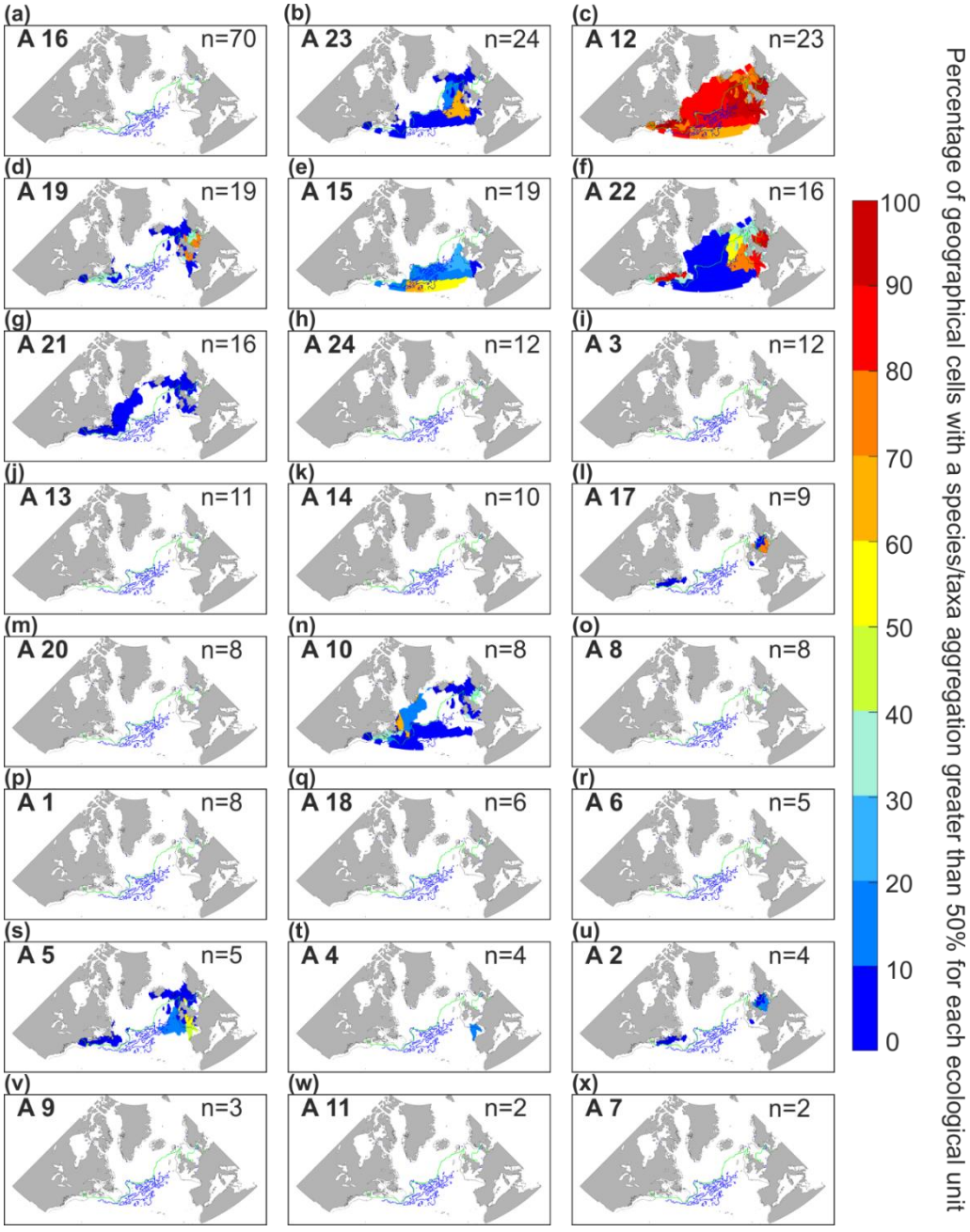
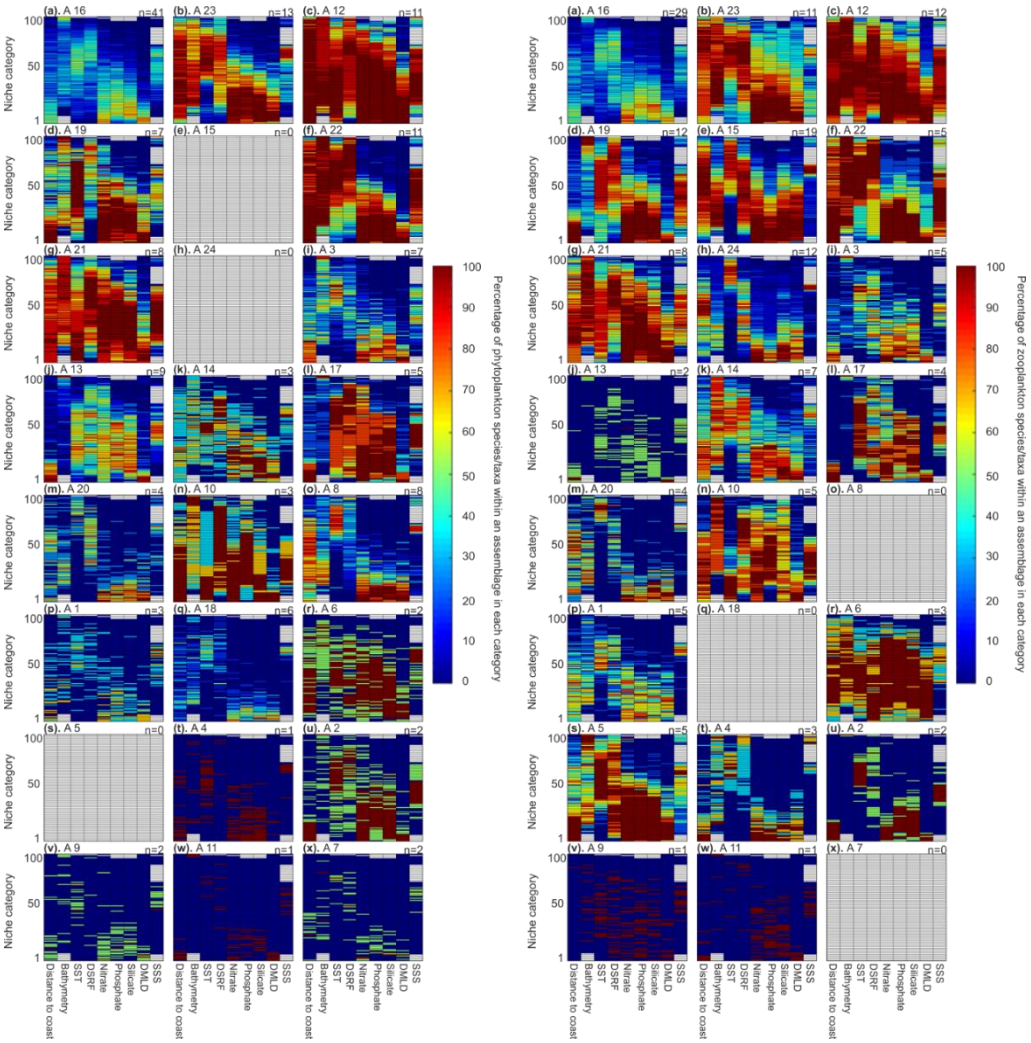


**Figure S10.** Environmental signature of phytoplankton species/taxa (left panel) and zooplankton species/taxa (right panel) for all 24 assemblages. For each figure, from a to x, column corresponds to the entire values taken by an environmental variable (e.g. distance to coast or silicate). For each variable (column) all values were divided into 100 categories standardised between 0 and 1, bottom categories (0) corresponding to the smallest values taken by an environmental variable. Colour indices denote the percentage of phytoplankton species/taxa (on left) or zooplankton species/taxa (on right) of an assemblage found into a category. Red colour indicates that the majority of the zooplankton or phytoplankton species/taxa composing the assemblage are found in these environmental categories. Blue colour indicates that no (or a few) zooplankton or phytoplankton species/taxa were found in these environmental categories. Panels are classified from a to x by decreasing taxonomic richness. The number at the top left of each panel corresponds to the assemblage number (see Figure 20 and Figure S4) and the number at the top right (n) indicates the taxonomic richness of each assemblage.



**Supplementary Tables.**

**Table S1.** List of species/taxa composing each assemblage.

| Assemblage | Species/taxa | |
|---|---|---|
| 1 | *Ceratium platycorne* | Dinoflagellates |
| 1 | *Ceratium vultur* | Dinoflagellates |
| 1 | *Phalacroma* spp. | Dinoflagellates |
| 1 | *Candacia longimana* | Copepoda |
| 1 | *Diacria* spp. | Gastropoda |
| 1 | *Lepas cypris* | Cirripedia |
| 1 | *Paraeuchaeta gracilis* | Copepoda |
| 1 | *Ctenocalanus vanus* | Copepoda |
| 2 | *Bellerochea malleus* | Diatom |
| 2 | *Guinardia flaccida* | Diatom |
| 2 | *Labidocera wollastoni* | Copepoda |
| 2 | *Isias clavipes* | Copepoda |
| 3 | *Pyrophacus* spp. | Dinoflagellates |
| 3 | *Biddulphia alternans* | Diatom |
| 3 | *Pterosperma* spp. (Total) | Prymnesiophyceae |
| 3 | *Cystodinium* spp. | Dinoflagellates |
| 3 | *Corythodinium* spp. | Dinoflagellates |
| 3 | *Gyrodinium* spp. | Dinoflagellates |
| 3 | *Prorocentrum micans* | Dinoflagellates |
| 3 | *Branchiostoma lanceolatum* | Cephalochordata |
| 3 | *Euterpina acutifrons* | Copepoda |
| 3 | *Zoothamnium pelagicum* | Ciliophora |
| 3 | Spindelei | Trematoda |
| 3 | Clytemnestridae | Copepoda |
| 4 | *Detonula pumila* | Diatom |
| 4 | *Calanoides carinatus* | Copepoda |
| 4 | *Centropages chierchiae* eyecount | Copepoda |
| 4 | *Centropages chierchiae* traverse | Copepoda |
| 5 | *Candacia armata* | Copepoda |
| 5 | *Echinoderm* post larvae | Echinodermata |
| 5 | *Subeucalanus crassus* | Copepoda |
| 5 | *Paraeuchaeta hebes* | Copepoda |
| 5 | *Corycaeus* spp. | Copepoda |
| 6 | *Rhizosolenia acuminata* | Diatom |
| 6 | *Proboscia curvirostris* | Diatom |
| 6 | *Aetideus armatus* | Copepoda |

| Assemblage | Species/taxa | |
|---|---|---|
| 6 | *Heterorhabdus norvegicus* | Copepoda |
| 6 | *Scolecithricella* spp. | Copepoda |
| 7 | *Ceratium buceros* | Dinoflagellates |
| 7 | *Ceratium lunula* | Dinoflagellates |
| 8 | *Ceratium candelabrum* | Dinoflagellates |
| 8 | *Ceratium carriense* | Dinoflagellates |
| 8 | *Ceratium declinatum* | Dinoflagellates |
| 8 | *Ceratium extensum* | Dinoflagellates |
| 8 | *Ceratium massiliense* | Dinoflagellates |
| 8 | *Ceratium pentagonum* | Dinoflagellates |
| 8 | *Ceratium teres* | Dinoflagellates |
| 8 | *Ceratium trichoceros* | Dinoflagellates |
| 9 | *Pyrocystis* spp. | Dinoflagellates |
| 9 | *Guinardia cylindrus* | Diatom |
| 9 | *Microcalanus* spp. | Copepoda |
| 10 | *Ceratium longipes* | Dinoflagellates |
| 10 | *Ceratium arcticum* | Dinoflagellates |
| 10 | *Fragilaria* spp. | Diatom |
| 10 | *Calanus glacialis* | Copepoda |
| 10 | *Calanus hyperboreus* | Copepoda |
| 10 | *Metridia longa* | Copepoda |
| 10 | *Parafavella gigantea* | Ciliophora |
| 10 | *Ptychocylis* spp. | Ciliophora |
| 11 | *Neocalyptrella robusta* | Diatom |
| 11 | Pycnogonida | Arthropoda |
| 12 | *Thalassiosira* spp. | Diatom |
| 12 | *Rhizosolenia styliformis* | Diatom |
| 12 | *Chaetoceros(Hyalochaete)* spp. | Diatom |
| 12 | *Chaetoceros(Phaeoceros)* spp. | Diatom |
| 12 | *Thalassiothrix longissima* | Diatom |
| 12 | *Thalassionema nitzschioides* | Diatom |
| 12 | *Ceratium fusus* | Dinoflagellates |
| 12 | Coccolithaceae (Total) | Prymnesiophyceae |
| 12 | Silicoflagellates | Prymnesiophyceae |
| 12 | *Pseudo-nitzschia delicatissima complex* | Diatom |
| 12 | *Pseudo-nitzschia seriata complex* | Diatom |
| 12 | *Calanus finmarchicus* | Copepoda |
| 12 | Hyperiidea (Total) | Malacostraca |

| Assemblage | Species/taxa | |
|---|---|---|
| 12 | Euphausiacea Adult | Malacostraca |
| 12 | Euphausiacea Juvenile | Malacostraca |
| 12 | *Paraeuchaeta norvegica* | Copepoda |
| 12 | *Para-Pseudocalanus* spp. | Copepoda |
| 12 | *Oithona* spp. | Copepoda |
| 12 | Calanus Total Traverse | Copepoda |
| 12 | Foraminifera (Total) | Protozoa |
| 12 | Appendicularia | Urochordata |
| 12 | Tintinnida Total | Ciliophora |
| 12 | Thecosomata (North Atlantic) | Gastropoda |
| 13 | *Actinoptychus* spp. | Diatom |
| 13 | *Odontella aurita* | Diatom |
| 13 | *Odontella granulata* | Diatom |
| 13 | *Odontella regia* | Diatom |
| 13 | *Odontella mobiliensis* | Diatom |
| 13 | *Rhizosolenia pungens* | Diatom |
| 13 | *Asterionellopsis glacialis* | Diatom |
| 13 | *Bleakeleya notata* | Diatom |
| 13 | *Bacillaria paxillifera* | Diatom |
| 13 | Ammodytidae | Actinopterygii |
| 13 | *Hemicyclops aberdonensis* | Copepoda |
| 14 | *Phaeocystis pouchetii* | Prymnesiophycea |
| 14 | *Ceratium compressum* | Dinoflagellates |
| 14 | *Nitzschia bicapitata* | Diatom |
| 14 | *Rhincalanus nasutus* | Copepoda |
| 14 | *Eucalanus hyalinus* | Copepoda |
| 14 | *Pneumodermopsis ciliata* | Gastropoda |
| 14 | Thaliacea | Urochordata |
| 14 | Salpidae (Total) | Urochordata |
| 14 | Doliolidae | Urochordata |
| 14 | Cnidaria tissue | Cnidaria |
| 15 | *Neocalanus gracilis* | Copepoda |
| 15 | *Nannocalanus minor* | Copepoda |
| 15 | *Euchirella rostrata* | Copepoda |
| 15 | *Euchaeta acuta* | Copepoda |
| 15 | *Pleuromamma abdominalis* | Copepoda |
| 15 | *Pleuromamma borealis* | Copepoda |
| 15 | *Pleuromamma gracilis* | Copepoda |

| Assemblage | Species/taxa | |
|---|---|---|
| 15 | *Centropages bradyi* | Copepoda |
| 15 | *Heterorhabdus papilliger* | Copepoda |
| 15 | *Pleuromamma piseki* | Copepoda |
| 15 | *Pleuromamma xiphias* | Copepoda |
| 15 | *Sapphirina* spp. | Copepoda |
| 15 | *Undeuchaeta plumosa* | Copepoda |
| 15 | Sergestidae | Malacostraca |
| 15 | *Mesocalanus tenuicornis* | Copepoda |
| 15 | *Calocalanus* spp. | Copepoda |
| 15 | *Lucicutia* spp. | Copepoda |
| 15 | *Mecynocera clausi* | Copepoda |
| 15 | *Oncaea* spp. | Copepoda |
| 16 | *Asteromphalus* spp. | Diatom |
| 16 | *Cerataulina pelagica* | Diatom |
| 16 | *Climacodium frauenfeldianum* | Diatom |
| 16 | *Detonula confervacea* | Diatom |
| 16 | *Leptocylindrus danicus* | Diatom |
| 16 | *Rhaphoneis amphiceros* | Diatom |
| 16 | *Planktoniella* sol | Diatom |
| 16 | *Stephanopyxis* spp. | Diatom |
| 16 | *Surirella* spp. | Diatom |
| 16 | *Pachysphaera* spp. | Chlorodendrophyceae |
| 16 | *Amphisolenia* spp. | Dinoflagellates |
| 16 | *Ceratium arietinum* | Dinoflagellates |
| 16 | *Ceratium bucephalum* | Dinoflagellates |
| 16 | *Ceratium lamellicorne* | Dinoflagellates |
| 16 | *Ceratium pulchellum* | Dinoflagellates |
| 16 | *Pronoctiluca pelagica* | Dinoflagellates |
| 16 | *Ptychodiscus noctiluca* | Dinoflagellates |
| 16 | *Ornithocercus* spp. | Dinoflagellates |
| 16 | *Glenodinium* spp. | Dinoflagellates |
| 16 | *Gymnodinium* spp. | Dinoflagellates |
| 16 | *Katodinium* spp. | Dinoflagellates |
| 16 | *Hexasterias problematica* | Prasinophyceae |
| 16 | *Diploneis* spp. | Diatom |
| 16 | *Nitzschia sigma rigida* | Diatom |
| 16 | *Actiniscus pentasterias* | Dinoflagellates |
| 16 | *Nitzschia longissima* | Diatom |

| Assemblage | Species/taxa | |
|---|---|---|
| 16 | *Neodenticula seminae* | Diatom |
| 16 | *Dinophysis acuminata* | Dinoflagellates |
| 16 | *Dinophysis acuta* | Dinoflagellates |
| 16 | *Dinophysis caudata* | Dinoflagellates |
| 16 | *Dinophysis norvegica* | Dinoflagellates |
| 16 | *Dinophysis tripos* | Dinoflagellates |
| 16 | *Bacterosira bathyomphala* | Diatom |
| 16 | *Prorocentrum rostratum* | Dinoflagellates |
| 16 | *Prorocentrum dentatum* | Dinoflagellates |
| 16 | *Phalacroma rotundatum* | Dinoflagellates |
| 16 | *Podosira stelligera* | Diatom |
| 16 | *Pseudosolenia calcar-avis* | Diatom |
| 16 | *Helicotheca tamesis* | Diatom |
| 16 | *Mediopyxis helysia* | Khakista |
| 16 | *Membraneis* spp. | Diatom |
| 16 | *Scaphocalanus echinatus* | Copepoda |
| 16 | *Euchaeta media* | Copepoda |
| 16 | *Paracandacia bispinosa* | Copepoda |
| 16 | *Rhincalanus cornutus* | Copepoda |
| 16 | *Undeuchaeta major* | Copepoda |
| 16 | *Labidocera aestiva* | Copepoda |
| 16 | *Paedoclione doliiformis* | Gastropoda |
| 16 | Isopoda (Total) | Malacostraca |
| 16 | *Cavolinia* spp. | Gastropoda |
| 16 | *Clio* spp. | Gastropoda |
| 16 | *Pneumodermopsis paucidens* | Gastropoda |
| 16 | *Oxygyrus* spp. | Gastropoda |
| 16 | *Peraclis* spp. | Gastropoda |
| 16 | Stomatopoda | Malacostraca |
| 16 | *Alteutha* spp. | Arthropoda |
| 16 | *Lucifer* spp. | Malacostraca |
| 16 | Pipefish | Fish |
| 16 | *Siphonostomatoida* | Copepoda |
| 16 | *Parathalestris croni* | Copepoda |
| 16 | *Penilia avirostris* | Cladocera |
| 16 | *Favella serrata* | Ciliophora |
| 16 | *Acartia danae* | Copepoda |
| 16 | *Acrocalanus* spp. | Copepoda |

| Assemblage | Species/taxa | |
|---|---|---|
| 16 | *Lubbockia* spp. | Copepoda |
| 16 | *Parapontella brevicornis* | Copepoda |
| 16 | *Tortanus discaudatus* | Copepoda |
| 16 | *Acartia longiremis* | Copepoda |
| 16 | Euphausiacea eggs | Malacostraca |
| 16 | Rotifer eggs | Rotifera |
| 17 | *Paralia sulcata* | Diatom |
| 17 | *Odontella sinensis* | Diatom |
| 17 | *Gyrosigma* spp. | Diatom |
| 17 | *Noctiluca scintillans* | Dinoflagellates |
| 17 | *Coscinodiscus wailesii* | Diatom |
| 17 | Gammaridea | Malacostraca |
| 17 | Cumacea | Malacostraca |
| 17 | Mysidacea | Malacostraca |
| 17 | Caprelloidea | Malacostraca |
| 18 | *Ceratium belone* | Dinoflagellates |
| 18 | *Ceratium gibberum* | Dinoflagellates |
| 18 | *Ceratium inflatum* | Dinoflagellates |
| 18 | *Ceratium karstenii* | Dinoflagellates |
| 18 | *Histioneis* spp. | Dinoflagellates |
| 18 | *Goniodoma polyedricum* | Dinoflagellates |
| 19 | *Ditylum brightwellii* | Diatom |
| 19 | *Eucampia zodiacus* | Diatom |
| 19 | *Rhizosolenia setigera* | Diatom |
| 19 | *Guinardia delicatula* | Diatom |
| 19 | *Dactyliosolen fragilissimus* | Diatom |
| 19 | *Guinardia striata* | Diatom |
| 19 | *Lauderia annulata* | Diatom |
| 19 | *Calanus helgolandicus* | Copepoda |
| 19 | Decapoda larvae (Total) | Malacostraca |
| 19 | Fish eggs (Total) | Fish |
| 19 | *Anomalocera patersoni* | Copepoda |
| 19 | *Temora longicornis* | Copepoda |
| 19 | *Centropages hamatus* | Copepoda |
| 19 | Cyphonautes | Bryozoa |
| 19 | Echinoderm larvae | Echinodermata |
| 19 | Cirripede larvae (Total) | Cirripedia |
| 19 | *Tintinnopsis* spp. | Ciliophora |

| Assemblage | Species/taxa | |
|---|---|---|
| 19 | Bivalvia *larvae* | Mollusca |
| 19 | *Pseudocalanus* spp. Adult Atlantic | Copepoda |
| 20 | *Hemiaulus* spp. | Diatom |
| 20 | *Ceratium setaceum* | Dinoflagellates |
| 20 | *Ceratocorys* spp. | Dinoflagellates |
| 20 | *Cladopyxis* spp. | Dinoflagellates |
| 20 | *Centropages violaceus* | Copepoda |
| 20 | *Copilia* spp. | Copepoda |
| 20 | *Atlanta* spp. | Heteropoda |
| 20 | *Temora stylifera* | Copepoda |
| 21 | *Skeletonema costatum* | Diatom |
| 21 | *Rhizosolenia hebetata semispina* | Diatom |
| 21 | *Coscinodiscus concinnus* | Diatom |
| 21 | *Navicula* spp. | Diatom |
| 21 | *Cylindrotheca closterium* | Diatom |
| 21 | *Proboscia inermis* | Diatom |
| 21 | *Ephemera planamembranacea* | Diatom |
| 21 | *Corethron hystrix* | Diatom |
| 21 | *Tomopteris* spp. | Annelida |
| 21 | *Clione limacina* | Gastropoda |
| 21 | Fish larvae | Fish |
| 21 | Ostracoda | Ostracoda |
| 21 | Cephalopoda larvae | Cephalopoda |
| 21 | *Halosphaera* spp. | Prymnesiophyceae |
| 21 | *Euphausiacea calyptopis* | Malacostraca |
| 21 | *Euphausiacea nauplii* | Malacostraca |
| 22 | *Ceratium furca* | Dinoflagellates |
| 22 | *Ceratium lineatum* | Dinoflagellates |
| 22 | *Ceratium tripos* | Dinoflagellates |
| 22 | *Ceratium macroceros* | Dinoflagellates |
| 22 | *Ceratium horridum* | Dinoflagellates |
| 22 | *Dinophysis* spp. Total | Dinoflagellates |
| 22 | *Protoperidinium* spp. | Dinoflagellates |
| 22 | *Prorocentrum* spp. Total | Dinoflagellates |
| 22 | *Scrippsiella* spp. | Dinoflagellates |
| 22 | *Proboscia alata* | Diatom |
| 22 | *Rhizosolenia imbricata* | Diatom |
| 22 | Chaetognatha eyecount | Chaetognatha |

| Assemblage | Species/taxa | |
|---|---|---|
| 22 | *Centropages typicus* | Copepoda |
| 22 | *Podon* spp. | Cladocera |
| 22 | *Evadne* spp. | Cladocera |
| 22 | Chaetognatha Traverse | Chaetognatha |
| 23 | *Dactyliosolen antarcticus* | Diatom |
| 23 | *Bacteriastrum* spp. | Diatom |
| 23 | *Rhizosolenia bergonii* | Diatom |
| 23 | *Ceratium azoricum* | Dinoflagellates |
| 23 | *Ceratium hexacanthum* | Dinoflagellates |
| 23 | *Ceratium minutum* | Dinoflagellates |
| 23 | *Gonyaulax* spp. | Dinoflagellates |
| 23 | *Oxytoxum* spp. | Dinoflagellates |
| 23 | *Podolampas* spp. | Dinoflagellates |
| 23 | *Leptocylindrus mediterraneus* | Diatom |
| 23 | *Prorocentrum* spp. ('Exuviaella' type) | Dinoflagellates |
| 23 | *Trichodesmium* spp. | Prymnesiophycea |
| 23 | *Proboscia indica* | Diatom |
| 23 | *Metridia lucens* | Copepoda |
| 23 | *Pleuromamma robusta* | Copepoda |
| 23 | *Clausocalanus* spp. | Copepoda |
| 23 | *Dictyocysta* spp. | Ciliophora |
| 23 | Harpacticoida Total Traverse | Copepoda |
| 23 | Metridia Total traverse | Copepoda |
| 23 | Radiolaria Total | Protozoa |
| 23 | *Acantharia* spp. | Protozoa |
| 23 | Radiolaria non-Acantharian | Protozoa |
| 23 | *Paracalanus* spp. | Copepoda |
| 23 | *Microsetella* spp. | Copepoda |
| 24 | *Miracia efferata* | Copepoda |
| 24 | *Urocorycaeus* spp. | Copepoda |
| 24 | *Candacia bipinnata* | Copepoda |
| 24 | *Candacia ethiopica* | Copepoda |
| 24 | *Candacia pachydactyla* | Copepoda |
| 24 | *Euchaeta marina* | Copepoda |
| 24 | *Scolecithrix danae* | Copepoda |
| 24 | *Undinula vulgaris* | Copepoda |
| 24 | *Paracandacia simplex* | Copepoda |
| 24 | Siphonophora | Cnidaria |

| Assemblage | Species/taxa | |
|---|---|---|
| 24 | *Lepas nauplii* | Cirripedia |
| 24 | *Corycaeus speciosus* | Copepoda |

**Table S2.** Spearman correlation coefficient (and its probability) between the environmental signature of each assemblage of phytoplankton and zooplankton. na: not applicable.

| Assemblage number | r | probability |
|---|---|---|
| 1 | 0.43 | 0 |
| 2 | 0.70 | 0 |
| 3 | 0.66 | 0 |
| 4 | 0.38 | 0 |
| 5 | na | na |
| 6 | 0.58 | 0 |
| 7 | na | 0 |
| 8 | na | na |
| 9 | 0.17 | 0 |
| 10 | 0.73 | 0 |
| 11 | 0.29 | 0 |
| 12 | 0.71 | 0 |
| 13 | 0.55 | 0 |
| 14 | 0.70 | 0 |
| 15 | na | na |
| 16 | 0.87 | 0 |
| 17 | 0.85 | 0 |
| 18 | na | na |
| 19 | 0.79 | 0 |
| 20 | 0.68 | 0 |
| 21 | 0.76 | 0 |
| 22 | 0.90 | 0 |
| 23 | 0.85 | 0 |
| 24 | na | na |

## 2.2. The species chromatogram, a new graphical method to represent, characterise and compare the ecological niches of different species

Loïck Kléparski and Grégory Beaugrand

*Published in Ecology and Evolution. Supplementary Files are displayed in section 2.2.6. Article is available at* https://doi.org/10.1002/ece3.8830

*Abstract*

The ecological niche *sensu* Hutchinson is defined as the set of environmental conditions allowing a species to grow, maintain and reproduce. This conception of the niche, which is assimilated to a p-dimensional hypervolume, with p representing all environmental variables, has been widely applied in ecology. However, displaying the niche hypervolume has proved challenging when more than three environmental dimensions are considered simultaneously. We propose a simple method (implemented in the *specieschrom* R package) that displays the full multidimensionality of the ecological niche of a species into a two-dimensional space by means of a graphic we call species chromatogram. This method gives a graphical summary of the niche by representing together abundance gradients with respect to all environmental variables. A chromatogram enables niche optimums and breaths to be rapidly quantified, and when several chromatograms are examined (one per species), rapid comparisons can be made. From our chromatograms, we proposed a procedure that quantifies niche optimum and breadth as well as niche overlapping (index D) and the identification of the most discriminant combination of environmental variables. We apply these analyses on eight planktonic species collected by the Continuous Plankton Recorder (CPR) survey in the North Atlantic Ocean using ten environmental variables. We display their full multidimensional niches and quantify their niche optimums and breadths along each dimension. We also compare our index D with other indices by means of *hypervolume* and *dynRB* R packages. By catching the full complexity of the niche, species chromatograms allow many different niche properties to be rapidly assessed and compared among species from niche optimums and breadths to the identification of the most relevant environmental parameters and the degree of niche overlapping among species. Species chromatograms may be seen as species' fingerprint and may also allow a better identification of the mechanisms involved in species assembly.

*Keywords*

Niche breadth, degree of niche overlapping, ecological niche, gradient analysis, hypervolume, niche optimum.

## 2.2.1. Introduction

Throughout the twentieth century, various definitions of the concept of ecological niche have been proposed. The first was formulated in 1917 by Joseph Grinnell, who defined the niche as the place occupied by a species in an environment (Grinnell, 1917). Ten years later, in 1927, Charles Elton proposed a more functional concept, the niche being seen as the role of a species in the food chain and its influence on the environment (Elton, 1927). These two conceptions envisioned the niche as an attribute of the environment and not as a property of a species, the niche being the place or the role that a species plays within a community and/or an ecosystem (Colwell and Rangel, 2009; Pulliam, 2000).

In 1957, Evelyn Hutchinson proposed a new concept of the niche, envisioned here as a species property (Hutchinson, 1957). He defined the niche as a set of environmental variables enabling a species to grow, maintain and reproduce. According to Hutchinson, the niche of a species can be viewed as a p-dimensional hypervolume in which each environmental combination enables a species to exist indefinitely (i.e. the species fundamental niche), this hypervolume being subsequently modulated by species interactions (i.e. the realized niche)(Hutchinson, 1978). This way to define the niche was in line with the Law of Tolerance, which states that a species is limited by its range of tolerance for environmental factors (Shelford, 1913).

A corollary of this new concept is that two species with the same niche in the same location cannot coexist, a statement known as the principle of competitive exclusion (Gause, 1934; Hutchinson, 1978). Therefore, each species of a community has a unique niche and the niche-environment interaction determines the place where a species lives and when it is active (Beaugrand, 2015a). Hence, the niche is a powerful tool to explain major biogeographical patterns at the species and even at the community levels (Beaugrand et al., 2020) because of the reciprocal correspondence, called Hutchinson's duality, between the niche space and the real physical space (Colwell and Rangel, 2009). Hutchinson's niche concept has been used to assess species and community responses to climate change in both space and time (Araújo and Guisan, 2006; Goberville et al., 2015; Thuiller et al., 2009)

However, the clear representation of the multidimensionality of the niche is challenging because of the difficulty for human to handle a space beyond more than three dimensions. Mathematicians have developed tools to solve this problem, e.g. Schlegel's diagrams, which enable the projection of a 4-dimensional hypercube (i.e. a tesseract) into a 3-dimensional space, in other words the representation of a p-dimensional polytope into a p-1 dimensional space. In ecology, indirect and direct gradient analyses have been applied but these techniques have some limitations due to normality assumption, the lack of explanatory power of the components or inherent complexity (Beaugrand et al., 2000b; Ter Braak and Prentice, 1988). Most of the time, ecologists manage dimensionality by seeking to summarize the information in a limited set of dimensions. To do so, they use multivariate analyses (e.g. Principal Component Analysis (PCA)) to characterise and display the niche (e.g. Broennimann et al., 2012). The Outlying Mean Index (OMI) is another technique that is also applied to characterise some properties of the niche (e.g. niche breadth) and assess which

environmental factors are the most structuring in a community (Dolédec et al., 2000). However, interpreting the outputs of those techniques is often challenging because the resulting components that are used to display the niche are typically a linear combination of different environmental dimensions and some variables can contribute to more than one principal component.

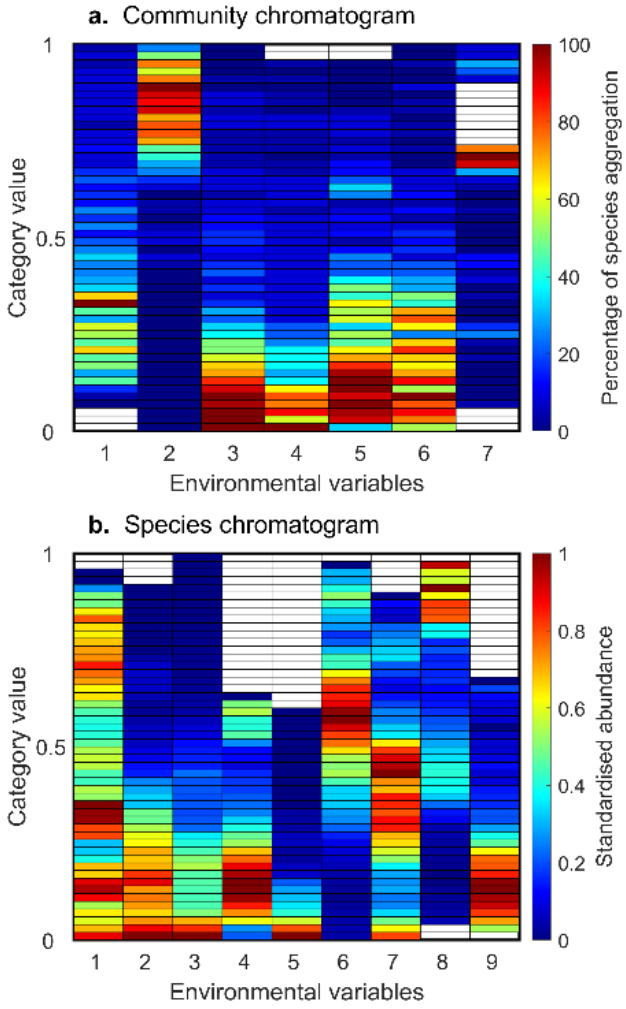



*Figure 24. Chromatogram of a hypothetical community (a) and a virtual species (b). (a) A community chromatogram shows where species of an assemblage aggregate along multiple ecological dimensions. Each column represents an environmental gradient divided into α categories (see Materials and Methods), from the lowest values taken by an environmental variable (bottom categories) to the highest (top categories). The colour in a category denotes the percentage of species of an assemblage, between 0 and 100%. Blue colour indicates that no or few species of an assemblage are found in a category and red colour indicates that the majority of the species composing an assemblage are found in a category. In this hypothetical example, large bands of high aggregation (see Table S1 for a definition) are observed from dimensions 3 to 5 and narrow bands for dimension 7. (b) A species chromatogram displays the multidimensional niche of a species into a two-dimensional space. Each column represents an environmental gradient divided into α categories, from the lowest (bottom categories) to the highest values taken by an environmental variable (top categories). The colour in a category denotes the standardised abundance of a species between 0 and 1. Blue colour in a category means that the species has a nil or low abundance in a category and red colour means that the species has a high abundance in a category. In this hypothetical example, large coloured bands of high abundance are observed for environmental dimension 1 and narrower bands for environmental dimensions 5 and 9.*

The niche hypervolume can also be represented by a set of two-dimensional pair plots of all possible combinations of the p-environmental variables defining the niche hyperspace. However this method leads to a vast amount of figures for a single species when the number of environmental dimensions is large (Blonder et al., 2014). Recently, Kléparski et al., (2021) proposed a new method called "the environmental chromatogram" to represent graphically the environmental signature of plankton assemblages in the North Atlantic Ocean, with colour bands representing the percentage of species aggregation within an assemblage or a community along multiple environmental gradients (Kléparski et al., 2021)(Figure 24a). The method allowed the authors to rapidly display the optimal environmental conditions in which an assemblage was found. Applied at a community/assemblage level, we propose to call such a graphic a community chromatogram from now on (see Table S1 for a full definition of the terms used in this paper).

In this study, we adapt this method at a species level to characterise graphically the ecological niche of a species by projecting the multidimensional space into a plane. Here, the resulting graphic is termed a species chromatogram (Figure 24b and Table S1). From species chromatograms, we propose a way to measure (i) niche optimum and (ii) breadth, (iii) to quantify the degree of niche overlapping among species and (iv) to identify the most discriminant combinations of environmental variables in term of niche differentiation. We apply the procedures on real species, i.e. four phytoplankton and four zooplankton species/taxa routinely sampled by the Continuous Plankton Recorder (CPR) survey in the North Atlantic Ocean. Finally, using 14 pseudo-species, we compared our estimation of niche overlapping against values obtained by means of the *hypervolume* (Blonder et al., 2018, 2014) and *dynRB* (Junker et al., 2016) R packages.

## 2.2.2. Materials and Methods

### 2.2.2.1. Materials

Plankton abundance data came from the Continuous Plankton Recorder (CPR) survey (Batten et al., 2003). It is a long-term plankton monitoring programme currently operated by the Marine Biological Association of the United Kingdom. Started in 1931, the programme has sampled plankton on a monthly basis in the North Atlantic Ocean and its adjacent seas. The CPR machine is a high-speed plankton recorder towed behind voluntary merchant ship, called "ship of opportunity", and operating at a depth of approximately ~7-10 m (Hays and Warner, 1993; Warner and Hays, 1994). We chose four diatoms and four copepods to test whether the robustness of our methods did not vary with taxonomic group. For each taxon, we chose species with known different spatial distribution (Barnard et al., 2004). Selected diatoms were *Paralia sulcata* (neritic tychopelagic species), *Skeletonema costatum* (neritic), *Rhizosolenia styliformis* (eurygraph) and *R. bergonii* (oceanic). Chosen copepods were *Temora longicornis* (temperate neritic species), *Clausocalanus* spp. (warm-temperate oceanic), *Calanus finmarchicus* (subarctic oceanic) and *Calanus helgolandicus* (pseudo-oceanic temperate). We used data collected between 1998 and 2018 in the North Atlantic

Ocean and its adjacent seas (Helaouët, 2021). This time period was preferred to correspond to the period covered by the environmental datasets described below.

Mass concentration of chlorophyll-a in sea water ($mg.m^{-3}$), nitrate, phosphate and silicate concentration ($mmol.m^{-3}$) data originated from the Global Ocean Biogeochemistry Hindcast (GLOBAL_REANALYSIS_BIO_001_029) and were provided by the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu). Daily means were provided on a 0.25° resolution grid and along 75 depth levels from 0 to 5500m. The dataset covers the time period from 1993 to present and is regularly updated.

Sea water potential temperature (°C), salinity (no unit) and Mixed Layer Depth (MLD, m) data originated from the Global Ocean Ensemble Physics Reanalysis (GLOBAL_REANALYSIS_PHY_001_031) and were provided by the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu). Daily means were provided on a 0.25° resolution grid along 75 depth levels from 0 to 5500m. The dataset covers the time period from 1993 to present and is regularly updated.

Euphotic depth data (m) data originated from the Global ocean low and mid trophic levels biomass content hindcast (GLOBAL_MULTIYEAR_BGC_001_033) provided by the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu). Daily means were provided on a 0.083° resolution grid, covering the time period 1998-2020.

Photosynthetically Active Radiations clear sky in surface (PAR, in $J.m^{-2}$) originated from the ERA interim dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (https://www.ecmwf.int/). Hourly means were provided on a 0.25° resolution grid, covering the time period 1998-2018. Daily PAR was estimated by summing all the values corresponding to a given day and were subsequently converted into $E.m^{-2}.day^{-1}$.

$K_d$ (PAR) data originated from the Glob Colour project (http://www.globcolour.info/). The product merges together all the daily data from satellites (MODIS, SeaWIFS, VIIRS) available for each parameter, from September 1997 to present and on a 4km resolution spatial grid. It provides daily means for each parameter. As the data can be very holey because of cloud cover and sun glint effect during the winter season, missing $K_d$ (PAR) values were first spatio-temporally interpolated and the remaining missing data (i.e. the one above 45°N in winter) where interpolated with chlorophyll-a data according to the relationships presented in Morel et al., 2007.

PAR in depth was finally estimated from the Beer-Lambert law (Swinehart, 1962):

$$I_Z = I_0 e^{-K_d \times Z} \qquad (1)$$

with $I_0$ the PAR in surface and $Z$ the depth (from 0 to 100 m).

Bathymetry (m) came from GEBCO Bathymetric Compilation Group 2019 (The GEBCO_2019 Grid - a continuous terrain model of the global oceans and land). Data are provided by the British Oceanographic Data Centre, National Oceanography

Centre, NERC, UK. doi:10/c33m. (https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/836f016a-33be-6ddc-e053-6c86abc0788e/). To work on the same spatial grid, $K_d$(PAR), euphotic depth and bathymetry were interpolated on a 0.25° latitude x 0.25° longitude grid.

We used data collected between 1998 and 2018 (i.e. from 1rst January 1998 to 31rst December 2018) in order to work on a common time period with respect to all biological and environmental datasets. All data were subsequently arranged on a grid covering the North Atlantic Ocean (100°W-10°E and 35°N-65°N). The dimension of all matrices (a total of ten matrices, each matrix corresponding to an environmental variable) was 121 latitudes x 441 longitudes x 7670 days. By means of nearest-neighbour interpolation (Wackernagel, 1995), we attributed to each CPR sample a value for each of the ten chosen environmental variables at a depth of 8 m (except for bathymetry), a value included in the range of sampling depth of the CPR instrument (Batten et al., 2003; Hays and Warner, 1993). CPR samples with at least one missing value along a single environmental dimension were discarded from the analysis.

### 2.2.2.2. Sketch of the method:

The method described below has been implemented in a R package (*specieschrom*, available on Github: https://github.com/loick-klpr/specieschrom.git). It is also available as Matlab functions (https://github.com/loick-klpr/Species-chromatogram-with-Matlab.git).

#### *2.2.2.2.1. Assessment of the species chromatogram*

The complete procedure to build a species chromatogram was composed of four main steps (see Figure 25):

**Step 1**: matrix **X** (with the spatio-temporal coordinates of n samples) and the corresponding vector **Y** (with the abundance of a species in the n samples) are built. Then, from p gridded environmental datasets, the values of the environmental variables (e.g. temperature) are assessed by nearest neighbour interpolation at the spatio-temporal coordinates stored in Matrix **X** (Wackernagel, 1995). This step enables the arrangement of a new matrix, Matrix **Z** (n samples by p environmental variables).

**Step 2**: matrix **Z** is standardised between 0 (the lowest value of an environmental variable) and 1 (the highest) as follows:

$$Z^*_{(i,j)} = \frac{Z_{i,j} - min(Z_j)}{max(Z_j) - min(Z_j)} \qquad (2)$$

where $Z^*_{(i,j)}$ the matrix of standardised environmental values for sample $i$ and environmental variable $j$. Standardisation is applied simultaneously along each environmental variable and for all species so as to niche dimensions could be compared from one chromatogram to another.

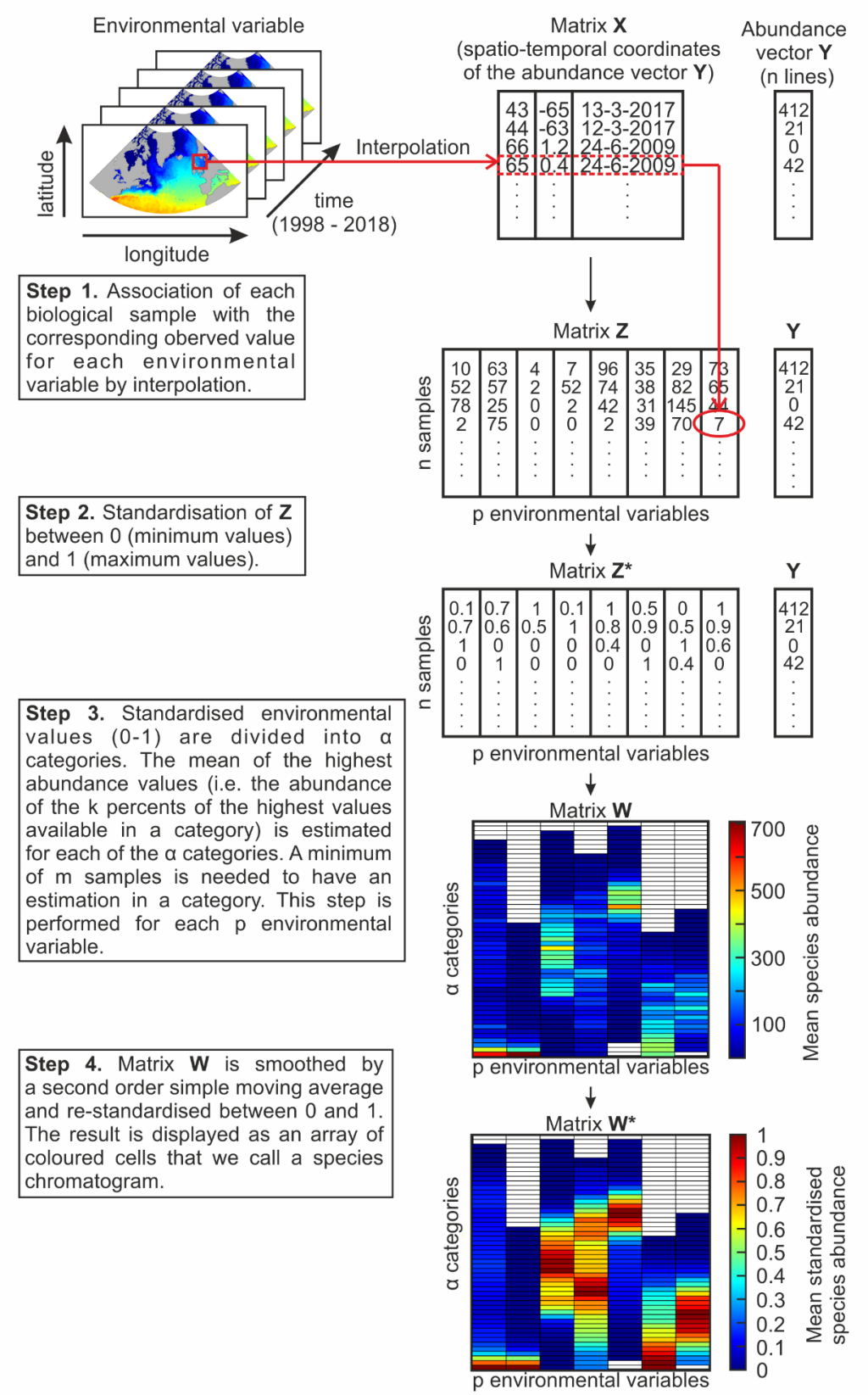



*Figure 25. Sketch diagram summarizing the different steps leading to the representation of the multidimensional ecological niche of a species as a chromatogram.*

**Step 3**: p-standardised environmental gradients are defined between 0 and 1 and divided into α equidistant categories, leading to Matrix $\mathbf{W}_{\alpha,p}$ (α categories by p environmental variables). Each sample in Matrix **Z*** is assigned to one of the α categories along a corresponding environmental gradient in **W**. If more than m samples are available within a category, an estimation of maximum abundance is calculated. This calculation is done by assessing the mean of the highest abundance values only, i.e. the abundance of the k percents of the highest abundance values available in that category. This threshold is implemented to account for the high number of nil or low abundance in a category due to adverse environmental conditions in other dimensions; in other words, environmental conditions can be suitable in a given dimension but unsuitable in others. Such a choice is in agreement with ecological niche theory (Brown, 1984). At the end of the procedure, each column in **W** corresponds to the average of the highest abundance of a species observed along an environmental dimension (e.g. temperature, PAR) from the lowest (bottom categories in the chromatograms) to the highest (top categories) environmental values. The niche of the species is therefore displayed by the location of its abundance in the categories along each environmental gradient.

**Step 4**: for each column in **W** (i.e. each environmental gradient), a second-order simple moving average is applied to reduce the noise in the mean abundance sometimes observed in the chromatograms from a category to another (see Figures S1-2 *versus* S3-4 for an example with CPR data). Then mean abundances are standardised between 0 (nil abundance) and 1 (highest abundance) for each environmental variable. Here, standardisation is performed as follows:

$$W^*_{(i,j)} = \frac{W_{(i,j)}}{max(W_j)} \qquad (3)$$

where $W^*_{(i,j)}$ the standardised value of Matrix **W** for sample $i$ and environmental variable $j$ (see Figures S3-4 *versus* S5-6 for an example with the CPR data).

Matrix **W*** is finally displayed as an array of coloured cells, which we call a species chromatogram, i.e. a graphic that shows how species abundance is distributed along multiple environmental gradients. (Species chromatogram can be built by means of the *chromato_env16* function available in the *specieschrom* R package.) All terms used to characterise the chromatograms are defined in Table S1.

#### *2.2.2.2.2. Niche optimum and breadth*

Niche optimum along each environmental dimension are assessed for each species and each dimension (i.e. each chromatogram; Table 2 and Table 3). To do so, we assume that highest species abundances are observed when environmental conditions are optimal (Brown, 1984; Helaouët and Beaugrand, 2009). For each niche dimension (i.e. each of the p environmental gradients/variables), the categories of the chromatogram where species abundance is maximal are identified. To provide a more precise estimation, the optimum is assessed by averaging the values of the environmental variable (Matrix **Z**) associated with the samples used to assess the

species abundance within the selected category (i.e. the k % of the samples with the highest values available in a category).

Niche breadth (Table S1) *E* is assessed for each niche dimension p by calculating the percentage of categories with an abundance higher than or equal to threshold T, with 0≤T≤1:

$$E = \frac{U_{max} - U_{min}}{U^*} \times 100 \qquad (4)$$

with $U_{max}$ the highest category with an abundance ≥ T and $U_{min}$ the lowest category with an abundance ≥ T for a given environmental dimension p; the difference in the numerator represents the number of contiguous categories with abundance ≥ T. $U^*$ is the number of categories with an estimation (i.e. non-missing values). Only the categories with an abundance > T are considered when T=0. For each given environmental dimension, we fill categories between the smallest and the highest, assuming a unimodal (continuous) niche in agreement with ecological niche theory (Brown, 1984; Hutchinson, 1957).

Average niche breadth ($E_T$, Table S1) is estimated as follows:

$$E_T = \frac{\sum_{i=1}^{p} E_i}{p} \qquad (5)$$

where p is the number of environmental dimensions. (Niche optimums and breadths can be assessed by means of the *opti_eury_niche2* function available in the R package *specieschrom*).

#### *2.2.2.2.3. Degree of niche overlapping*

The degree of niche overlapping (Table S1) between two species of the same taxonomic group is assessed by means of index $D$, which is estimated by calculating the ratio of the part of the hypervolume of the niche common to the two species $V_{S1,S2}$ on the total volume filled by the sum of the hypervolume of the two niches $V_{S1}$ (species 1) and $V_{S2}$ (species 2). Index $D$ is calculated as follows:

$$D(s_1, s_2) = \frac{100 \times V_{S1,S2}}{V_{S1} + V_{S2} - V_{S1,S2}} \qquad (6)$$

with $V_{S1} = \prod_{i=1}^{p} \beta_i$, $V_{S2} = \prod_{i=1}^{p} \gamma_i$ and $V_{S1,S2} = \prod_{i=1}^{p} \theta_i$ with $\beta_i$ and $\gamma_i$ the number of category higher or equal to T for species 1 and species 2 respectively, and $\theta_i$ the number of common categories in the species chromatogram of species 1 and 2, with a joint standardised abundance value (between 0 and 1) higher than or equal to threshold T. Only the categories with an abundance > T are considered when T=0. p is the number of environmental dimensions. When there was no overlap among the two species' niches, $V_{S1,S2} = 0$ and $D = 0\%$. When the two species' niches are identical, $V_{S1} = V_{S2} = V_{S1,S2}$ , so $D = 100\%$. For this analysis, we assume that the niche has the shape of a p-dimensional orthotope (i.e. the generalization of a rectangle in higher

dimensions or hyperrectangle). Therefore, prior to the calculation of index $D$, we also fill vacant categories between the smallest and the highest selected categories for a given environmental dimension assuming a unimodal (continuous) niche in agreement with ecological niche theory (Brown, 1984; Hutchinson, 1957). We warn that value of $D$ might be biased when the smallest or the highest category (or both) is not well identified. The advantage of our index is that it is not influenced by niche asymmetry.

Index $D$ can be calculated for all species of a taxonomic group and all combinations of dimensions ranging from 1 to p. By calculating the average of all values of the matrix, we can identify the most discriminant combinations of environmental dimensions, i.e. the combinations of environmental dimensions that play an important role in term of niche differentiation for the group of species under investigation. Results can be sorted for niches based on a growing number of dimensions from 1 to p. (Niche overlapping among species can be assessed by means of the *combina_niche3* function available in the *specieschrom* R package.)

### 2.2.2.3. Example with real data

In this study, we used the species chromatogram to display and characterise the multidimensional niche of eight plankton species/taxa into a two-dimensional space. To do so, we used the abundance data of four diatoms (i.e. *Paralia sulcata*, *Skeletonema costatum*, *Rhizosolenia styliformis* and *R. bergonii*) and four copepods (i.e. *Temora longicornis*, *Clausocalanus* spp., *Calanus finmarchicus* and *C. helgolandicus*) collected by the Continuous Plankton Recorder (CPR) survey between 1998 and 2018 in the North Atlantic Ocean and its adjacent seas (35-65°N and 100°W-10°E). For diatoms, we used nine environmental dimensions (p=9): bathymetry (in m), nitrate, phosphate and silicate concentrations (mmol.m$^{-3}$), mixed layer depth (MLD, m), temperature (°C), photosynthetically active radiations (PAR, E.m$^{-2}$.day$^{-1}$), salinity (no unit) and euphotic depth (m) (Figure 26). For copepods, we used seven dimensions (p=7): bathymetry (m), MLD (m), temperature (°C), PAR, salinity (no unit), chlorophyll-a concentration (mg.m$^{-3}$) and euphotic depth (m) (Figure 27).

A total of 90,527 CPR samples was used (the repartition of the CPR samples in each category of the chromatograms are displayed in Figure S7). We chose α=50 categories for each column of the species chromatogram and thresholds of m=1 (as an example to illustrate the method; Figure S1-6) and m=20 samples (for deep analysis, Figure 26 and Figure 27 and subsequent tables). To handle with the high proportion of nil abundance observed in many CPR samples, we fixed k to 5%.

We estimate niche breadth per dimension or average niche breath (all dimensions) and index D for the four diatoms and the four copepods by selecting five thresholds: T=0, 0.05, 0.1, 0.25 and 0.5 (Table 4, Table 5, Table 6 and Table 7 and Tables S2-17). Spearman rank correlation coefficients were calculated between niche breadth values obtained for T=0.25 and T=0, 0.05, 0.1 and 0.5. Correlation were tested by means of a Monte Carlo test using 10,000 simulations (Jackson and Somers, 1989).

### 2.2.2.4. Comparison with other methods

To test the validity of a method, simulated rather than real data should be used because the former has known distributions and overlaps whereas the latter might be affected by unknown biases and sampling error (Broennimann et al., 2012). Therefore, to test our approach, we generated seven pseudo-species (i.e. virtual species) with a three-dimensional niche using the following equation (Yan and Hunt, 1999):

$$A(p_i) = c\left(\frac{p_{max}-p_i}{p_{max}-p_{opt}}\right)\left(\frac{p_i-p_{min}}{p_{opt}-p_{min}}\right)^{\left(\frac{p_{opt}-p_{min}}{p_{max}-p_{opt}}\right)} \quad (7)$$

with $A$ the abundance of a pseudo-species along an environmental gradient $p_i$, $c$ the maximal abundance (here 1), $p_{opt}$ the niche optimum along $p_i$ and $p_{min}$ and $p_{max}$ the amplitude (i.e. niche breadth). Abundances were estimated along three hypothetical environmental dimensions (i.e. $p_1$ from 0 to 25, $p_2$ from 0 to 40 and $p_3$ from 1 to 0). A matrix of 100 samples by 3 environmental dimensions was obtained and estimated abundances in each sample and along each environmental dimension were aggregated with an additive model. Each pseudo-species was duplicated (therefore we considered 7x2=14 pseudo-species) to assess the reliability of the overlapping estimation when two niches were identical.

Degrees of niche overlapping between the 14 pseudo-species were estimated with index D (using α=50, k=5, m=1 and T=0) and with the functions available in the *hypervolume* and *dynRB* R packages using their default parameter settings (Blonder et al., 2018; Junker et al., 2016). Basically, the method developed in the *hypervolume* package uses a hyperellipse random sampling algorithm to generate a uniform random set of points around each observation of the dataset (i.e. a matrix of n samples by p-environmental dimensions). Then a function describing the niche hypervolume is assessed on these points by means of a Gaussian Kernel Density Estimation (Gaussian KDE) or a one-class Support Vector Machine (one-class SVM, i.e. an algorithm based on machine learning). The method used in the *dynRB* package is based on an improvement of the concept of multivariate range boxes (Hutchinson, 1957), i.e. a finite number of nested standardised range boxes enclosing a decreasing quantile range of the data are generated and then used to assess a volume and an overlap between each pair of species, the results being subsequently aggregated along all niche dimensions.

Prior to the estimation with both R packages, the dataset (100 samples by 3 dimensions) were standardised according to Equation 2. With the *hypervolume* package, niche overlapping was assessed by means of a Jaccard similarity coefficient on niches delineated with (i) a Gaussian KDE (using a Silverman bandwidth estimator; Figure 28a) or (ii) a one-class SVM (Figure 28b). With the *dynRB* package, niche overlapping was assessed by mean of dynamic range boxes on niches where highly correlated environmental dimensions were (i) kept (Figure 28c) or (ii) replaced with principal components (Figure 28d). As the overlaps estimated with this package are asymmetric (i.e. niche overlap of pseudo-species $s_1$ on pseudo-species $s_2$ is different from the niche overlap of $s_2$ on $s_1$) we converted the estimations as follow:

$$D_{dynRB}(s_1, s_2) = \frac{(vol(s_1) \times port(s_2, s_1)) \times 100}{vol(s_1) + vol(s_2) - (vol(s_1) \times port(s_2, s_1))} \quad (8)$$

With $D_{dynRB}(s_1, s_2)$ the symmetric niche overlap between pseudo-species $s_1$ and $s_2$, $vol(s_1)$ the niche volume of pseudo-species $s_1$, $vol(s_2)$ the niche volume of pseudo-species $s_2$ and $port(s_2, s_1)$ the average portion of the niche of $s_1$that is covered by $s_2$. We calculated the Spearman rank correlation coefficients between index D and the overlaps estimated with the other methods. The correlations were tested by means of a Monte Carlo test using 10,000 permutations (Jackson and Somers, 1989)

## 2.2.3. Results

### 2.2.3.1. The species chromatograms

Species chromatograms were performed for diatoms and copepods using a threshold m=20 CPR samples to have more reliable estimates of abundance per category (Figure 26 and Figure 27). In this case, the threshold k=5% meant that we performed the average of at least four highest abundance values (0.05 x 20 samples = 4), which decreased the between category variability (Figure 26 and Figure 27 *versus* Figures S5-S6). The chromatograms allowed a rapid characterisation of the niche of these planktonic species. When all environmental dimensions were considered together, all species had a distinct chromatogram (Figure 26 and Figure 27). A visual inspection between patterns exhibited by the species chromatograms (Figure 26 and Figure 27) and patterns in the number of CPR samples (Figure S7) suggests that variation in sampling effort among categories did not influence chromatograms substantially.

Because results of a chromatogram are self-understandable, we only highlighted a few key patterns. The diatoms *Paralia sulcata* and *Skeletonema costatum,* as well as the copepod *Temora longicornis* and to a lesser extent *Calanus helgolandicus,* were found in shallow regions (i.e. neritic species) in contrast to the diatoms *Rhizosolenia styliformis* and *R. bergonii* (oceanic regions) and the copepods *Clausocalanus* spp. and *Calanus finmarchicus* (Figure 26 and Figure 27). *R. bergonii* and *Clausocalanus* spp. were thermophilic species, with high abundance towards the top categories for temperature. They also displayed great abundance for high values of PAR, salinity and euphotic depth. In addition, *R. bergonii* showed high abundance for low values of nutrients concentration (Figure 26d and Figure 27b). Their chromatogram therefore suggests that *R. bergonii* and *Clausocalanus* spp. are oceanic species and that *R. bergonii* is adapted to oligotrophic waters (Figure 26d and Figure 27b).

Some species such as *R. styliformis* and *Clausocalanus* spp. displayed large colour bands of high abundance along many environmental dimensions (i.e. a colour band is an aggregation of more or less continuous categories along an environmental dimension of the chromatogram; Table S1 and Figure 26c Figure 27b). These large colour bands observed along some dimensions revealed large niche breadth with respect to the dimensions. Other species such as *R. bergonii* and *T. longicornis*

exhibited narrower bands of high abundance (e.g. silicate for *R. bergonii* and MLD for *T. longicornis*) and therefore narrower niche breadth (Figure 26d and Figure 27a).

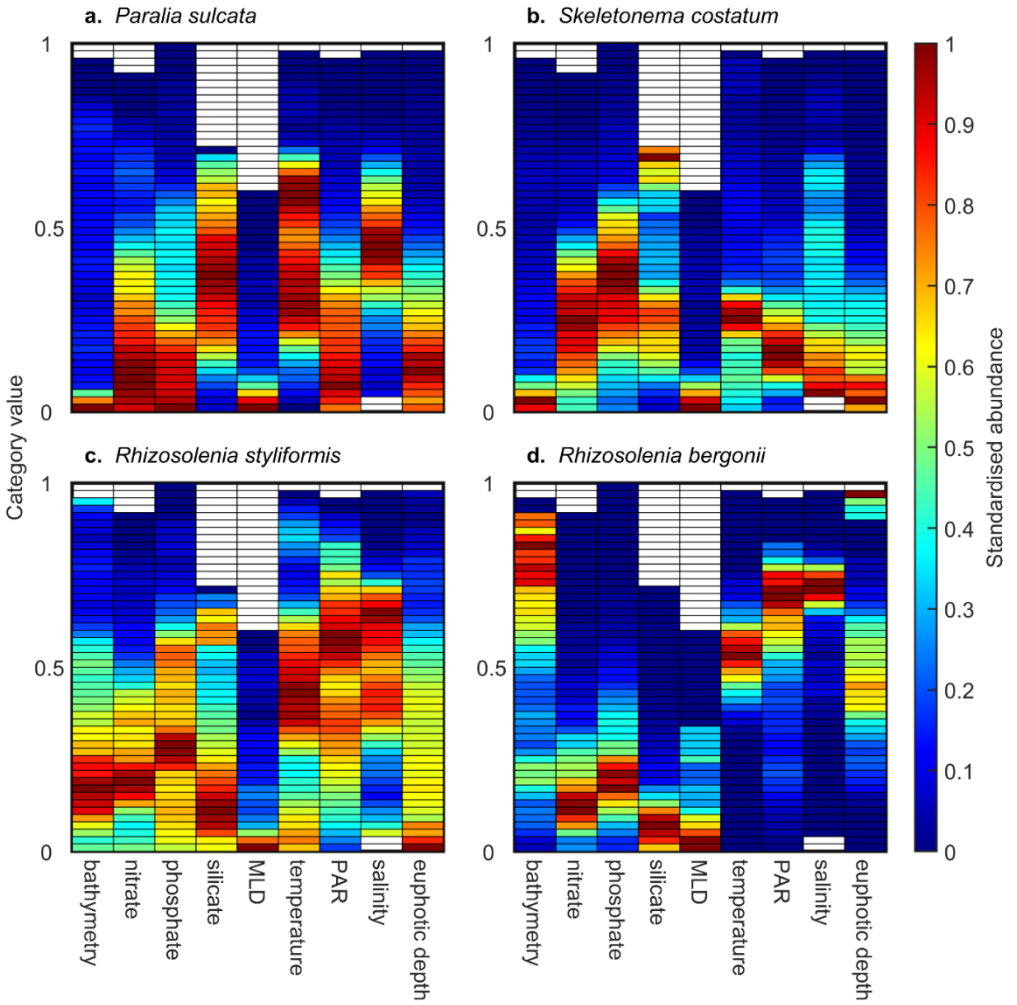



*Figure 26. Species chromatograms of four diatom species. Species chromatogram of (a) Paralia sulcata, (b) Skeletonema costatum, (c) Rhizosolenia styliformis and (d) Rhizosolenia bergonii. In a-d, each column represents the species abundance along nine environmental dimensions (i.e. bathymetry, nitrate, phosphate and silicate concentration, MLD, temperature, PAR, salinity and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue.*

Some species had complementary chromatograms along some dimensions, e.g. *P. sulcata* and *R. styliformis* for bathymetry (Figure 26a *versus* c) or *S. costatum* and *R. bergonii* for temperature, PAR and salinity (Figure 26b *versus* d). Although two species may have similar chromatograms with respect to some dimensions (e.g. *T. longicornis* and *C. helgolandicus* along the bathymetric dimensions; Figure 27a and d), the same species may be separated by other environmental dimensions (e.g. *T. longicornis* and *C. helgolandicus* along the dimension salinity; Figure 27a and d).

Therefore, the species chromatograms can rapidly characterise the full multidimensional complexity of the niche and allows species niche comparisons to be made rapidly.

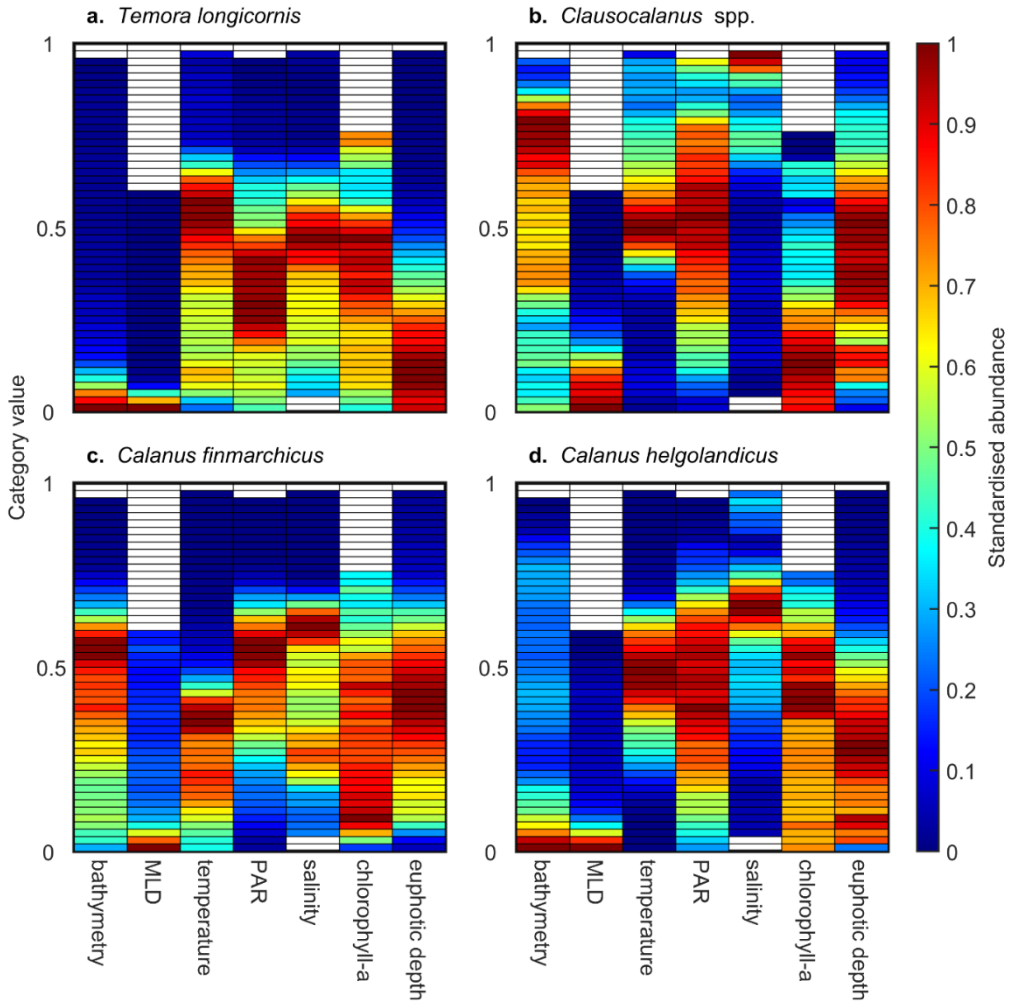



*Figure 27. Species chromatograms of four copepods. Species chromatogram of (a) Temora longicornis, (b) Clausocalanus spp., (c) Calanus finmarchicus and (d) Calanus helgolandicus. In a-d, each column represents the mean species abundance along seven environmental dimensions (i.e. bathymetry, MLD, temperature, PAR, salinity, chlorophyll-a concentration and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue.*

### 2.2.3.2. Estimates of niche optimums

Among diatoms, *P. sulcata* had the lowest optimum for bathymetry, PAR, nitrate and phosphate concentrations but the highest for temperature, suggesting that the diatom was abundant over continental shelves during warm periods (Table 2, Figure 26a). *S. costatum* had the highest optimum for nutrients concentration and MLD but the lowest for temperature and salinity, conditions indicative of the spring bloom over cold temperate continental shelves (Caracciolo et al., 2021)(Table 2, Figure 26b).

The lowest optimum for euphotic depth was found for *R. styliformis* (Table 2, Figure 26c). The diatom *R. bergonii* had the highest optimum for bathymetry, PAR, salinity and euphotic depth and the lowest for silicate, some features indicative of oligotrophic waters, which are characteristic of the open ocean during warm-stratified periods (Table 2, Figure 26d).

Among zooplankton, *T. longicornis* had the lowest niche optimum for bathymetry, MLD, PAR, salinity and euphotic depth but the highest for temperature and chlorophyll-a concentration (Table 3, Figure 27a). The lowest thermal optimum was found for *C. finmarchicus*, a subarctic oceanic species, which also had the highest optimum for PAR and the lowest for chlorophyll-a concentration (Table 3 and Figure 27c). A low optimum was also observed with respect to chlorophyll-a concentration for *Clausocalanus* spp., a warm-temperate oceanic genus. This copepod had optimums for higher values of bathymetry, MLD, salinity and euphotic depth (Table 3 and Figure 27b). *C. helgolandicus* had intermediate optimums for all the variables (Table 3 and Figure 27d).

*Table 2. Niche optimum assessed from the species chromatogram for each variable and diatom. The ecological niche of each species is displayed in Figure 26.*

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (m)** | 43.86 | 164.87 | 1024.75 | 4924.60 |
| **Nitrate (mmol.m$^{-3}$)** | 1.53 | 4.24 | 3.23 | 1.86 |
| **Phosphate (mmol.m$^{-3}$)** | 0.00 | 0.40 | 0.29 | 0.17 |
| **Silicate (mmol.m$^{-3}$)** | 5.24 | 9.54 | 1.79 | 1.26 |
| **MLD (m)** | 18.29 | 18.34 | 17.82 | 17.71 |
| **Temperature (°C)** | 16.45 | 6.31 | 11.67 | 14.66 |
| **PAR (E.m$^{-2}$.day$^{-1}$)** | 2.34 | 5.02 | 19.04 | 23.03 |
| **Salinity (no unit)** | 33.18 | 29.48 | 35.03 | 35.58 |
| **Euphotic depth (m)** | 20.48 | 12.97 | 10.93 | 102.88 |

*Table 3. Niche optimum assessed from the species chromatogram for each variable and copepod. The ecological niche of each species/taxa is displayed in Figure 27.*

| | ***Temora longicornis*** | ***Clausocalanus* spp.** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (m)** | 45.01 | 4716.65 | 3166.20 | 81.85 |
| **MLD (m)** | 15.60 | 18.27 | 17.57 | 16.84 |
| **Temperature (°C)** | 15.85 | 13.47 | 9.28 | 13.46 |
| **PAR ($E.m^{-2}.day^{-1}$)** | 9.70 | 17.70 | 18.36 | 13.01 |
| **Salinity (no unit)** | 33.37 | 38.01 | 34.66 | 35.21 |
| **Chlorophyll-a ($mg.m^{-3}$)** | 2.08 | 0.59 | 0.41 | 1.83 |
| **Euphotic depth (m)** | 18.25 | 58.92 | 51.31 | 38.22 |

### 2.2.3.3. Estimates of niche breadth

Niche breadth was assessed for each species based on five different thresholds of abundance T. Only results for T=0.25 (Table 4 and Table 5) are described in details here; results with other thresholds are shown in Tables S2-9. Among diatoms, *R. styliformis* had the highest average niche breadth ($E_T$=73.02%; Table 4 and Figure 26c). Among studied copepods, *Clausocalanus* spp. was the most euryoecious (average niche breadth $E_T$=68.42%; Table 5 and Figure 27b). In contrast, *R. bergonii* and *T. longicornis* had the narrowest average niche breadth, with $E_T$=42.21% and $E_T$=53.36% respectively (Table 4 and Table 5 and Figure 26d and Figure 27a). Niche breadth sometimes exhibited very different values among diatoms or copepods for the same dimensions. For example, although niche breadth for bathymetry was ~6% for *P. sulcata* and 12.5% for *T. longicornis*, niche breadth was 100% and ~92% for *R. styliformis* and *Clausocalanus* spp., respectively (Table 4and Table 5 and Figure 26a, Figure 27a, Figure 26c and Figure 27b). Within a species, e.g. *R. bergonii*, large niche breadth (e.g. bathymetry) could be found for a niche dimension whereas narrow niche breadth (e.g. salinity) could be observed for another (Table 4 and Figure 26d). Among copepods, *C. helgolandicus* had the narrowest niche for temperature but one of the largest for PAR (Table 5 and Figure 27d).

Altering threshold T did affect estimates of niche breadth but this effect was small when the modification on T was moderate (Tables S2-9). For diatoms, Spearman rank correlation between niche breadth based on T=0.25 and T=0, 0.05, 0.1 and 0.5 was 0.45 (degree of freedom df=36, $p<0.05$), 0.68 (df=36, $p<0.01$), 0.82 (df=36, $p<0.01$) and 0.88 (df=36, $p<0.01$) respectively. For copepods, Spearman rank correlation between niche breadth based on T=0.25 and T=0, 0.05, 0.1 and 0.5 was 0.69 (df=28,

p<0.01), 0.69 (df=28, p<0.01), 0.84 (df=28, p<0.01) and 0.85 (df=28, p<0.01) respectively.

*Table 4. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four diatoms based on a threshold of abundance T=0.25. The mean niche breadth (E_T) for all dimensions and each species is also displayed in bold. The ecological niche of each diatom is shown in Figure 26.*

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 6.25 | 10.42 | 100.00 | 81.25 |
| **Nitrate (%)** | 54.35 | 52.17 | 54.35 | 34.78 |
| **Phosphate (%)** | 58.00 | 58.00 | 64.00 | 36.00 |
| **Silicate (%)** | 86.11 | 88.89 | 97.22 | 25.00 |
| **MLD (%)** | 16.67 | 16.67 | 26.67 | 56.67 |
| **Temperature (%)** | 59.18 | 34.69 | 91.84 | 26.53 |
| **PAR (%)** | 50.00 | 31.25 | 85.42 | 35.42 |
| **Salinity (%)** | 48.94 | 68.09 | 74.47 | 14.89 |
| **Euphotic depth (%)** | 44.90 | 34.69 | 63.27 | 69.39 |
| **$E_T$ (%)** | **47.15** | **43.87** | **73.02** | **42.21** |

### 2.2.3.4. Niche differentiation

Finally, we investigated which sets of environmental variables were the most discriminant in term of niche differentiation among species/taxa of the same taxonomic group. As for niche breadth, the degree of niche overlapping $D$ was assessed for five different thresholds of abundance T. Results for T=0.25 are shown in details here (Table 6 and Table 7) and results based on other thresholds are shown in Tables S10-17.

When only one environmental dimension was considered, the most discriminant variable, i.e. the variable that allowed niche overlapping to be the smallest was, for diatoms (for T=0.25), bathymetry ($D$=26.32%), followed by PAR (32.97%), temperature (34.56%), euphotic depth (42.35%), salinity (42.51%), silicate (54.24%), MLD (55.15%), phosphate (76.94%) and finally nitrate (81.11%). For zooplankton species/taxa, the most discriminant variable was salinity ($D$=34.17%),

followed by temperature (42.02%), bathymetry (49.91%), MLD (53.61%), euphotic depth (68.49%), PAR (69.03%) and chlorophyll-a (92.5%).

*Table 5. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four copepods based on a threshold of abundance T=0.25. The mean niche breadth (E_T) for all dimensions and each species/taxa is also displayed in bold. The ecological niche of each copepod is shown in Figure 27.*

| | ***Temora longicornis*** | ***Clausocalanus spp.*** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 12.50 | 91.67 | 70.83 | 83.33 |
| **MLD (%)** | 10.00 | 33.33 | 20.00 | 13.33 |
| **Temperature (%)** | 69.39 | 59.18 | 48.98 | 44.90 |
| **PAR (%)** | 68.75 | 91.67 | 54.17 | 79.17 |
| **Salinity (%)** | 65.96 | 34.04 | 61.70 | 61.70 |
| **Chlorophyll-a (%)** | 100.00 | 89.47 | 97.37 | 94.74 |
| **Euphotic depth (%)** | 46.94 | 79.59 | 65.31 | 59.18 |
| **$E_T$ (%)** | **53.36** | **68.42** | **59.77** | **62.34** |

We then examined the effect of the combinations of environmental variables on index $D$. Expectedly, for diatoms or copepods, we found that the lowest mean degree of niche overlapping was reached when the number of niche dimensions considered was highest (Table 6 and Table 7 and Tables S10-17). For diatoms, when only two environmental variables were considered, bathymetry and temperature was the combination of variables that allowed to best separate their niches (i.e. lowest degree of niche overlapping, Table 6). For zooplankton species/taxa, this was bathymetry and salinity (Table 7). For diatoms, when four environmental variables were considered, the lowest value of $D$ was found with the combination of variables bathymetry, temperature, PAR and salinity (Table 6). For zooplankton, the lowest value of $D$ was reached for the combination of bathymetry, MLD, temperature and salinity (Table 7).

The use of other thresholds T could lead to the detection of other combinations of variables but in general there was a high consistency in the combinations from T=0 to T=0.5 (Table 6 and Table 7 and Tables S10-17).

*Table 6. Mean degree of niche overlapping for the four diatoms based on a threshold of abundance T=0.25. The first column displays the number of dimensions considered simultaneously, columns 2 to 10 display the combinations of dimensions (i.e. 1=bathymetry, 2=nitrate, 3=phosphate, 4=silicate, 5=MLD, 6=temperature, 7=PAR, 8=salinity and 9=euphotic depth). The last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each species is displayed in Figure 26.*

| **Number of dimensions** | **Combinations** | | | | | | | | | **Index D (%)** |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 1 | | | | | | | | | 26.32 |
| **2** | 1 | 6 | | | | | | | | 9.20 |
| **3** | 1 | 6 | 8 | | | | | | | 4.79 |
| **4** | 1 | 6 | 7 | 8 | | | | | | 3.15 |
| **5** | 1 | 6 | 7 | 8 | 9 | | | | | 2.49 |
| **6** | 1 | 4 | 6 | 7 | 8 | 9 | | | | 2.34 |
| **7** | 1 | 4 | 5 | 6 | 7 | 8 | 9 | | | 2.24 |
| **8** | 1 | 2 | 4 | 5 | 6 | 7 | 8 | 9 | | 2.17 |
| **9** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 2.14 |

### 2.2.3.5. Comparison of species chromatograms with other methods

The three-dimensional niches of 14 pseudo-species were examined by means of (i) species chromatograms and (ii) sets of pair plots from the *hypervolume* package (Blonder et al., 2018). A visual comparison of the figures revealed that both procedures gave similar results (Figures S8 *versus* S9-S10). In our hypothetical examples, both methods easily identified which pseudo-species had overlapping or non-overlapping niches (e.g. pseudo-species 8 *versus* 11 and 7 *versus* 14, respectively) and enabled the comparison of the niche breadth of all pseudo-species along all niche dimensions (e.g. pseudo-species 5 had a larger niche breadth than pseudo-species 4). However, because the procedure used in the *hypervolume* method does not consider abundance, optimums cannot be visually identified which can make interpretations difficult. For example, according to the *hypervolume* display (Figures S9-10), pseudo-species 1 and 9 seemed to have very similar niches but the species chromatogram display revealed that they had different niche optimums (Figure S8).

We also compared the degree of niche overlapping estimated from our index D with the *hypervolume* or *dynRB* packages (Figure 28). Although the relationships

were not always linear (Figure 28b-c), we found comparable degree of niche overlapping with both methods (Spearman correlation $r_s$>0.9 and $p_s$<0.01), even when the procedures used to delineate the niche were different (Figure 28a *versus* b) or when the dimensionality was reduced by means of a PCA (Figure 28c *versus* d). However, we noticed that niche overlapping estimated by means of the *hypervolume* method never reached 100%, even when both pseudo-species had the same niche (Figure 28a-b and Figure S8), an issue that has already been reported elsewhere (Junker et al., 2016).

*Table 7. Mean degree of niche overlapping for the four copepods based on a threshold of abundance T=0.25. The first column displays the number of dimensions considered simultaneously, columns 2 to 8 display the combinations of dimensions (i.e. 1=bathymetry, 2=MLD, 3=temperature, 4=PAR, 5=salinity, 6=chlorophyll-a concentration and 7=euphotic depth) and the last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each copepod is shown in Figure 27.*

| **Number of dimensions** | **Combinations** | | | | | | | **Index D (%)** |
|---|---|---|---|---|---|---|---|---|
| **1** | 5 | | | | | | | 34.17 |
| **2** | 1 | 5 | | | | | | 18.05 |
| **3** | 1 | 3 | 5 | | | | | 10.05 |
| **4** | 1 | 2 | 3 | 5 | | | | 6.38 |
| **5** | 1 | 2 | 3 | 5 | 7 | | | 5.01 |
| **6** | 1 | 2 | 3 | 4 | 5 | 7 | | 4.17 |
| **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 4.07 |

## 2.2.4. Discussion

The chromatography is a physico-chemical method used to separate the different components of a mixture. This mixture, dissolved in a fluid, is allowed to travel in a system including a fixed stationary phase. The mixture migrates along papers or polymers at a velocity that depends upon the characteristics of the molecules, which enable them to be separated. Some methods show the results under the form of a diagram, called a chromatogram, with different coloured bands, each reflecting a different component of the fluid. (We recall that the Greek etymology of the word “chromatography” means “to write in colour”.) Although very different, our procedure leads to a graphic that can be called a species chromatogram, by analogy

with the classical physico-chemical method. In a species chromatogram, coloured bands spread along different environmental gradients for each dimension of the niche.

The species chromatogram method summarises rapidly the niche of a species and enables rapid comparisons to be made. Comparison between chromatograms is possible because of the double standardisation between 0 and 1, i.e. (i) the standardisation of each environmental dimension considering all species involved in a study (a continuous and unitless environmental dimension is an essential prerequisite for comparing different hypervolumes in an Euclidean space (Blonder, 2018)) and (ii) the standardisation of the abundance for all categories of a given environmental dimension. From a chromatogram, one can identify niche optimums and breaths with respect to all niche dimensions. A rapid quantification of the difference among niches can also be undertaken, which is important to evaluate the degree of niche overlapping among species. Least and not last, the method allows combinations of environmental dimensions that minimise niche overlapping to be identified.

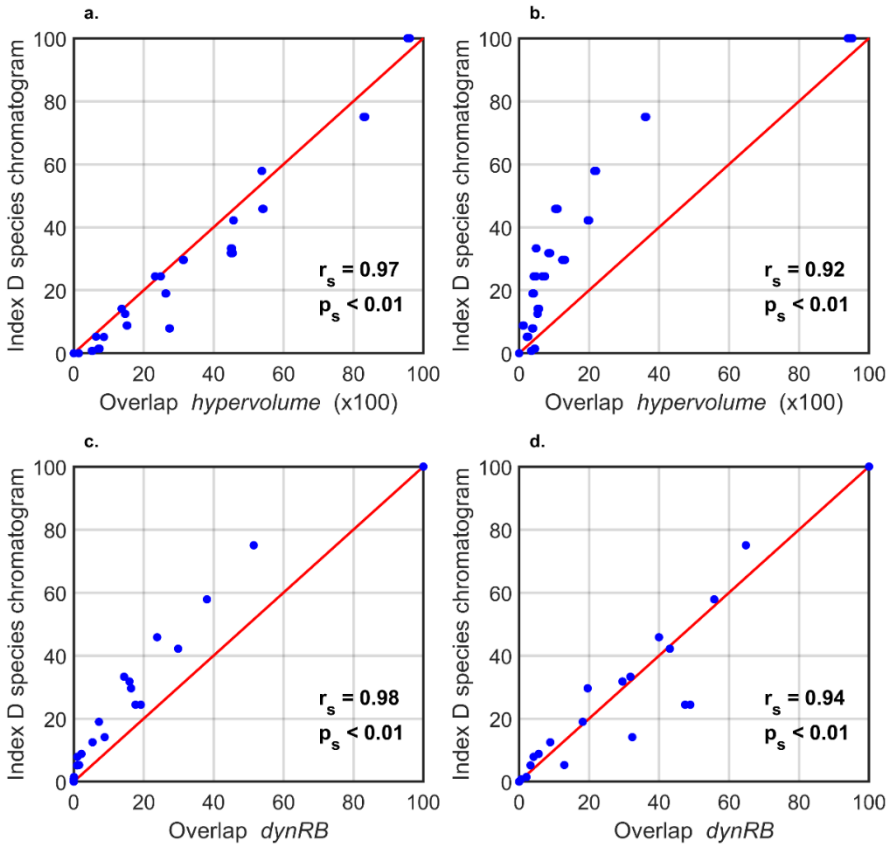



*Figure 28. Relationships between the indices of niche overlapping estimated from the species chromatograms (index D) and the indices originating from (a-b) the hypervolume (Blonder et al., 2018) and (c-d) the dynRB (Junker et al., 2016) R packages. Comparison between niche overlapping of 14 pseudo-species assessed by means of index D and (a-b) the Jaccard similarity coefficient and (c-d) dynamic range boxes. In a and b, pseudo-species niches hypervolume were delineated with (a) a Gaussian KDE and (b) a one-class SVM. In c and d, highly correlated environmental dimensions have been (c) kept or (d) replaced with principal components before niche hypervolume estimation. Red line displayed the y=x relationship. Spearman rank correlation and its associated probability are displayed at the bottom right of each panel.*

Other methods have been proposed to represent the niche of a species. Among them, the simplest and most efficient is perhaps the one used in the *hypervolume* R package (Blonder et al., 2014), which consists in a set of pair plots for all dimensions that define the space where the hypervolume belongs to. However, as $p \times (p-1)/2$ combinations of dimensions are possible and because each variable is represented according to another, it would have led here for a single species to 21 figures when p=7 and 36 when p=9. Ordination methods can also be applied to characterise the ecological niche of a species (Dolédec et al., 2000; Ter Braak and Prentice, 1988). Among them, a Principal Component Analysis (PCA) has already been used to represent the niche of *C. finmarchicus* and *C. helgolandicus* based on more than ten environmental variables in the North Atlantic Ocean (Helaouët and Beaugrand, 2007). Principal Components (PCs) enabled the reduction of the number of environmental dimensions and species abundance was then represented in a space defined by the different PCs (three in Helaouët & Beaugrand, 2007). However, PCs are linear combinations of environmental factors, and therefore the resulting assessment of the multidimensional niche is difficult to interpret because the weight of each environmental dimension in the PCs is not so easy to understand and some variables can be represented in more than one PC. Furthermore, as in many multivariate techniques applied at a species or a community level, the PCs may sometimes explain a small fraction of the variance (Ter Braak and Prentice, 1988).

In contrast, a species chromatogram enables a simple characterisation of the niche of a species and allows niches to be compared. Multidimensional niches are not summarised by creating composite variables and all the environmental dimensions are used (even if correlated) to display a species niche. Therefore, niche holes (i.e. unoccupied part of the niche which are difficult to detect but indicative of important ecological and evolutionary processes (Blonder, 2016)) can be easily identified. Differences and similarities among niches can be visually assessed, allowing the examination of three ecological phenomena: (i) Hutchinson's duality (Colwell and Rangel, 2009), (ii) environmental filtering (Zobel, 1997) and (iii) niche complementarity (i.e. the niche differentiation effect) (Tilman, 1999). Last and not least, the role and contribution of each environmental dimension to a species niche can be easily assessed.

Our method has some limitations. First, the value of threshold T influences the estimation of the niche breadth and the degree of niche overlapping among species and different results can be observed for different thresholds (Table 4, Table 5, Table 6 and Table 7 and Tables S2-17). Although a few differences were found, results remained quite consistent especially for copepods. Estimating the degree of overlapping and niche breadth is difficult especially in the pelagic environment because of the absence of strong physical barriers (van der Spoel, 1994). Therefore, the realized niche can be larger than the fundamental niche because of species dispersal (also called species expatriation) (Pulliam, 2000). The application of our numerical procedures on terrestrial data may show less variability for different values of T.

Second, large variability in the abundance estimates of each category of the chromatogram can also occur (Figure S1). This large variability has two main causes. The first cause is related to the fact that the abundance of a species within a category of a given environmental dimension is also influenced by the range of conditions that also occurs in other dimensions. Having a nil abundance in a category corresponding to optimal conditions for a particular dimension is possible when other dimensions have unsuitable environmental conditions. That is why we calculated the average abundance corresponding to the k% of the highest values available within a given category. The second cause is more inherent to the CPR survey. The CPR machine samples ~3 $m^3$ of seawater but the range of filtered water can vary between 2 and 5 $m^3$ depending on ship speed and plankton concentration in the water column (Jonas et al., 2004). Variation in seawater filtered may have severe consequences for abundance estimation.

Third, empty (white) categories were observed for some dimensions of the chromatogram, e.g. silicate, MLD or euphotic depth (Figure 26 and Figure 27). These empty categories were due to an insufficient number of CPR samples and was reduced when threshold m diminished from 20 to 1 sample(s) although outliers appeared in the same time (Figures S5 and S6 *versus* Figure 26 and Figure 27). However, empty categories could still be observed along some dimensions because some environmental conditions are rarely observed in the North Atlantic sector, i.e. fundamental environmental conditions are not always realised (Jackson and Overpeck, 2000). The smoothing of the data also exacerbated the number of missing categories by altering the location of some white categories in the chromatograms (see the chlorophyll-a dimension in Figures S2 *versus* S4). Adjusting the order (i.e. degree of smoothing) of the simple moving average may be necessary if the method is applied to other datasets.

Last, we assimilated the shape of the multidimensional niche to a hyperrectangle in order to estimate the degree of niche overlapping (index D). This assumption agrees with the definition of the niche *sensu* Hutchinson (1957), which supposed an equal probability of persistence in each point composing the fundamental niche, even if suboptimal conditions should be observed near the boundaries. More complicated shapes are observed with the realised niche because of the distortions and the modulations generated by biotic interactions that create unoccupied spaces e.g. niche holes (Blonder, 2016; Soberón and Peterson, 2020). Therefore, prior to overlapping estimation, we filled the unoccupied categories along each dimension, which assimilated niche shape to a hyperrectangle, an assumption that was tested through a comparison we performed with the *hypervolume* and *dynRB* packages (Figure 28).

During the last decade, many approaches have been developed to characterise a niche hypervolume, each with their own assumptions and drawbacks (Blonder, 2018; Qiao et al., 2017, 2015; Soberon and Nakamura, 2009). However, some methods are more easily applicable than other. For example, the *dynRB* package has been developed to provide reliable results using default parameter settings (Junker et al., 2016). Although it has been recently updated, the *hypervolume* method requires

expert knowledge and might therefore be misused because of the numerous assumptions underlying its application (Blonder et al., 2018, 2017; Qiao et al., 2017). Furthermore, as the method uses an algorithm which generates a random set of points, it cannot assess total niche overlapping (i.e. an overlapping of 100%) between two pseudo-species having the same niche (Figure 28a-b). In contrast, a species chromatogram is simple and easily understandable. The method does not require the selection of multiple thresholds and underlying functions. Our technique is therefore reproducible by a broad range of ecologists and might be straightforwardly adaptable to various datasets and conditions.

### 2.2.5. Conclusions

The species chromatogram is a simple and visually appealing method enabling a clear and rapid representation of the (multidimensional) ecological niche of a species into a two-dimensional space. The method thereby allows one to characterise the full multidimensional complexity of the niche of a species. The niche is displayed with p-standardised gradients figuring the continuous variation of each environmental variable, defining the niche axes from the lowest to the highest value taken by each variable, each gradient being divided into α equidistant categories filled with species' abundance. A species chromatogram can be seen as a species' fingerprint, summarizing its environmental requirements. Although we only used quantitative variables, semi-quantitative or qualitative variables can also be selected. A precise estimation of niche optimums and breadths along each dimension is also possible from a chromatogram. In addition, the quantification of the degree of niche overlapping can be made, which rapidly identifies the most discriminant combinations of environmental variables that minimise niche overlapping among species. Although we applied the method on marine plankton species, the species chromatogram can also be applied to terrestrial data.

## 2.2.6. Supplementary Files

**Supplementary Figures.**

**Figure S1. Species chromatograms of four diatom species without the application of the second-order simple moving average and the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Paralia sulcata*, (**b**) *Skeletonema costatum*, (**c**) *Rhizosolenia styliformis* and (**d**) *Rhizosolenia bergonii*. In **a**-**d**, each column represents the mean species abundance along nine environmental dimensions (i.e. bathymetry, nitrate, phosphate and silicate concentration, MLD, temperature, PAR, salinity and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (without standardisation between 0 and 1) in each category. High abundance values are in red and low values in blue.

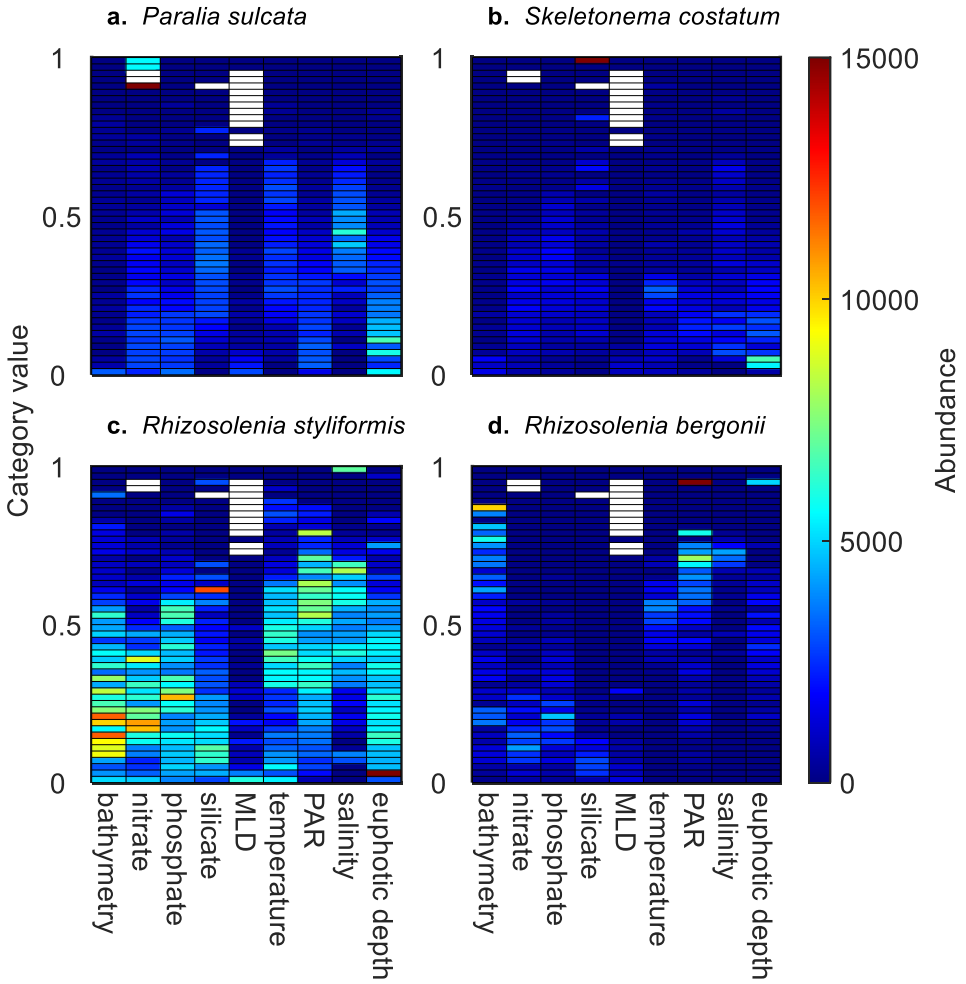


**Figure S2. Species chromatograms of four copepods without the application of the second-order simple moving average and the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Temora longicornis*, (**b**) *Clausocalanus* spp., (**c**) *Calanus finmarchicus* and (**d**) *Calanus helgolandicus*. In **a**-**d**, each column represents the mean species abundance along seven environmental dimensions (i.e. bathymetry, MLD, temperature, PAR, salinity, chlorophyll-a concentration and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (without standardisation between 0 and 1) in each category. High abundance values are in red and low values in blue.

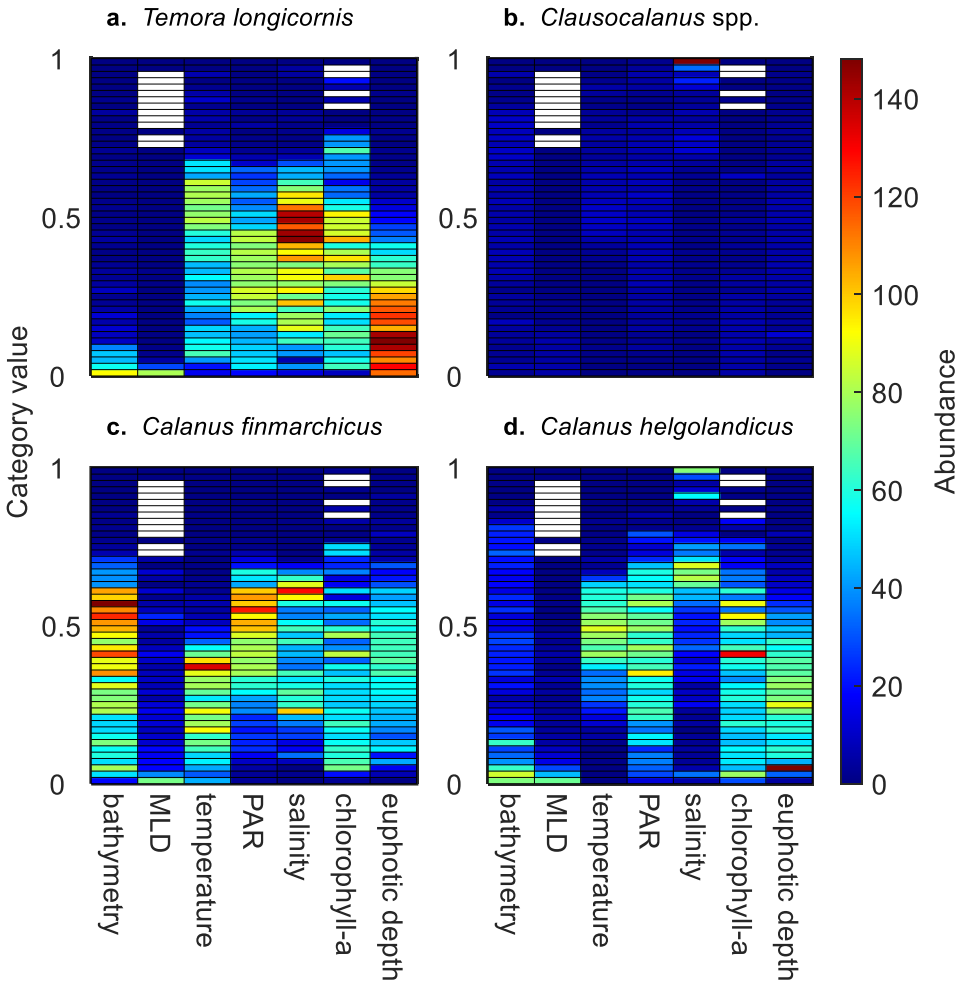


**Figure S3. Species chromatograms of four diatom species with the application of the second-order simple moving average but without the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Paralia sulcata*, (**b**) *Skeletonema costatum*, (**c**) *Rhizosolenia styliformis* and (**d**) *Rhizosolenia bergonii*. In **a**-**d**, each column represents the mean species abundance along nine environmental dimensions (i.e. bathymetry, nitrate, phosphate and silicate concentration, MLD, temperature, PAR, salinity and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (without standardisation between 0 and 1) in each category. High abundance values are in red and low values in blue.

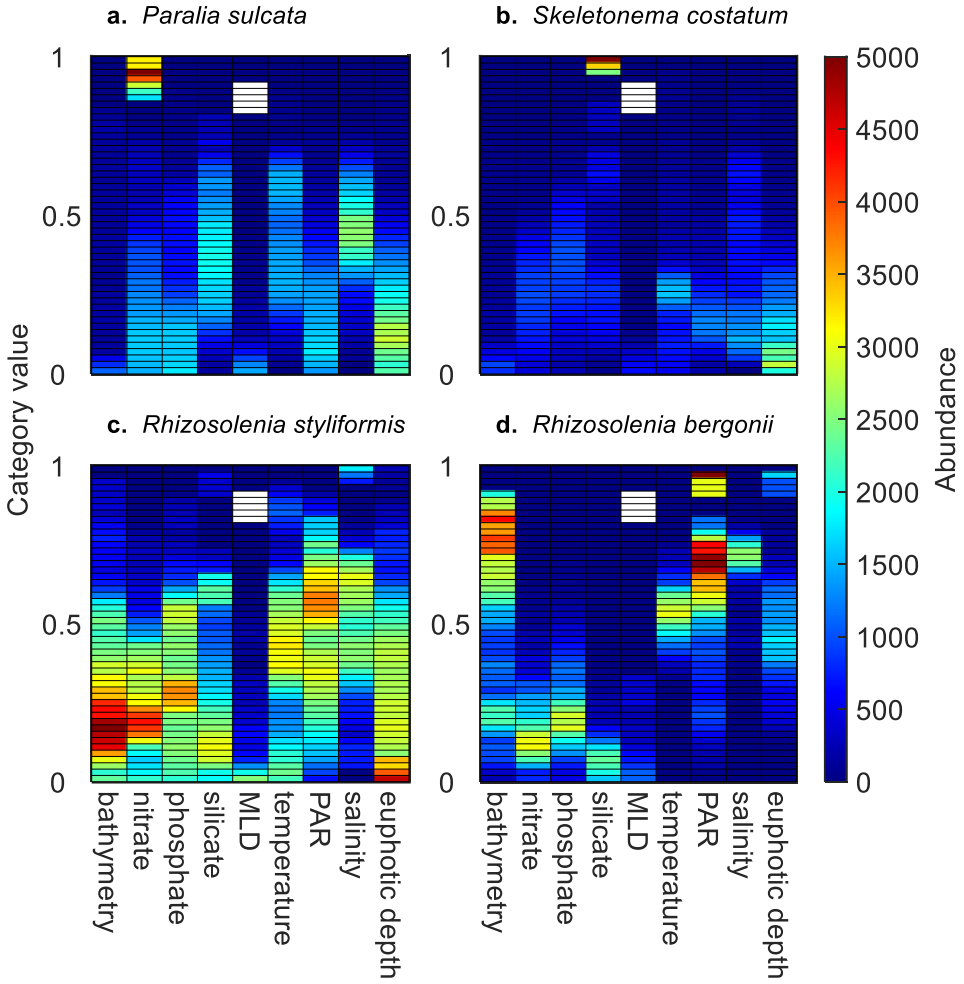


**Figure S4. Species chromatograms of four copepods with the application of the second-order simple moving average but without the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Temora longicornis*, (**b**) *Clausocalanus* spp., (**c**) *Calanus finmarchicus* and (**d**) *Calanus helgolandicus*. In **a**-**d**, each column represents the mean species abundance along seven environmental dimensions (i.e. bathymetry, MLD, temperature, PAR, salinity, chlorophyll-a concentration and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (without standardisation between 0 and 1) in each category. High abundance values are in red and low values in blue.

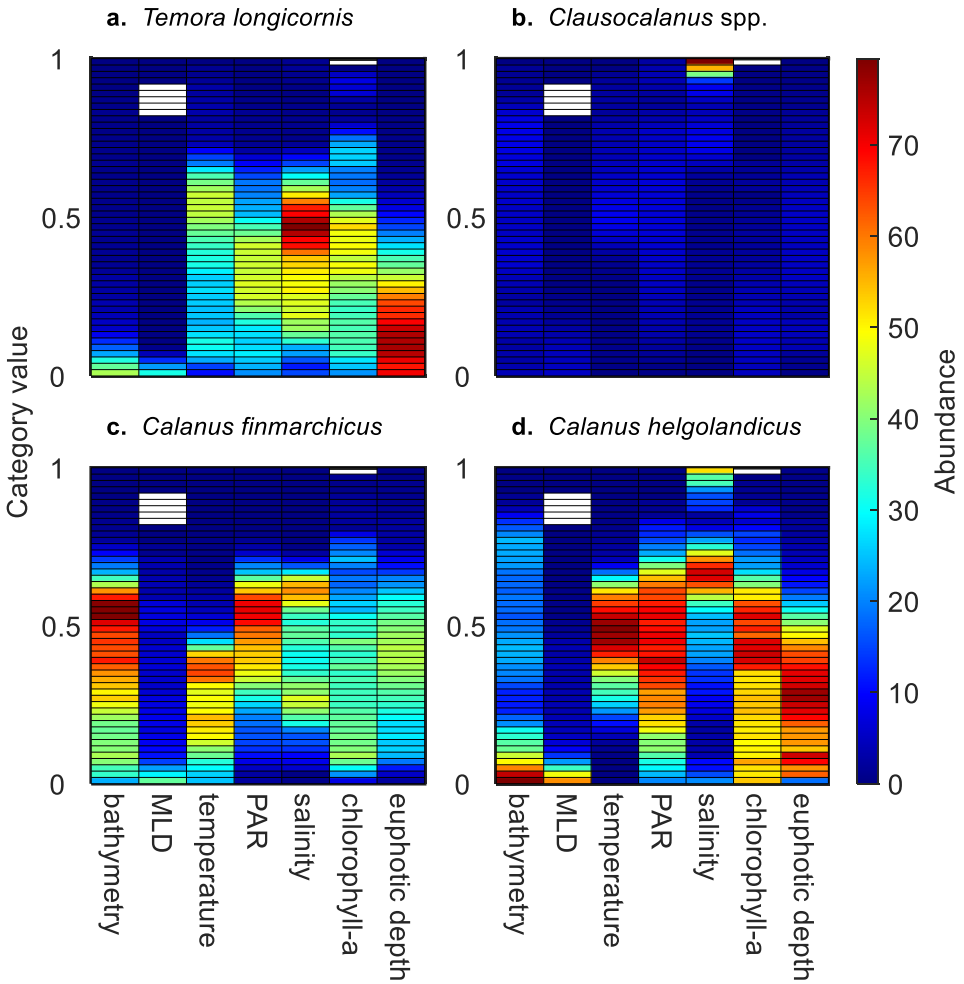


**Figure S5. Species chromatograms of four diatom species with the application of the second-order simple moving average and the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Paralia sulcata*, (**b**) *Skeletonema costatum*, (**c**) *Rhizosolenia styliformis* and (**d**) *Rhizosolenia bergonii*. In **a**-**d**, each column represents the species abundance along nine environmental dimensions (i.e. bathymetry, nitrate, phosphate and silicate concentration, MLD, temperature, PAR, salinity and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (standardised between 0 and 1) in each category. High abundance values are in red and low values in blue.

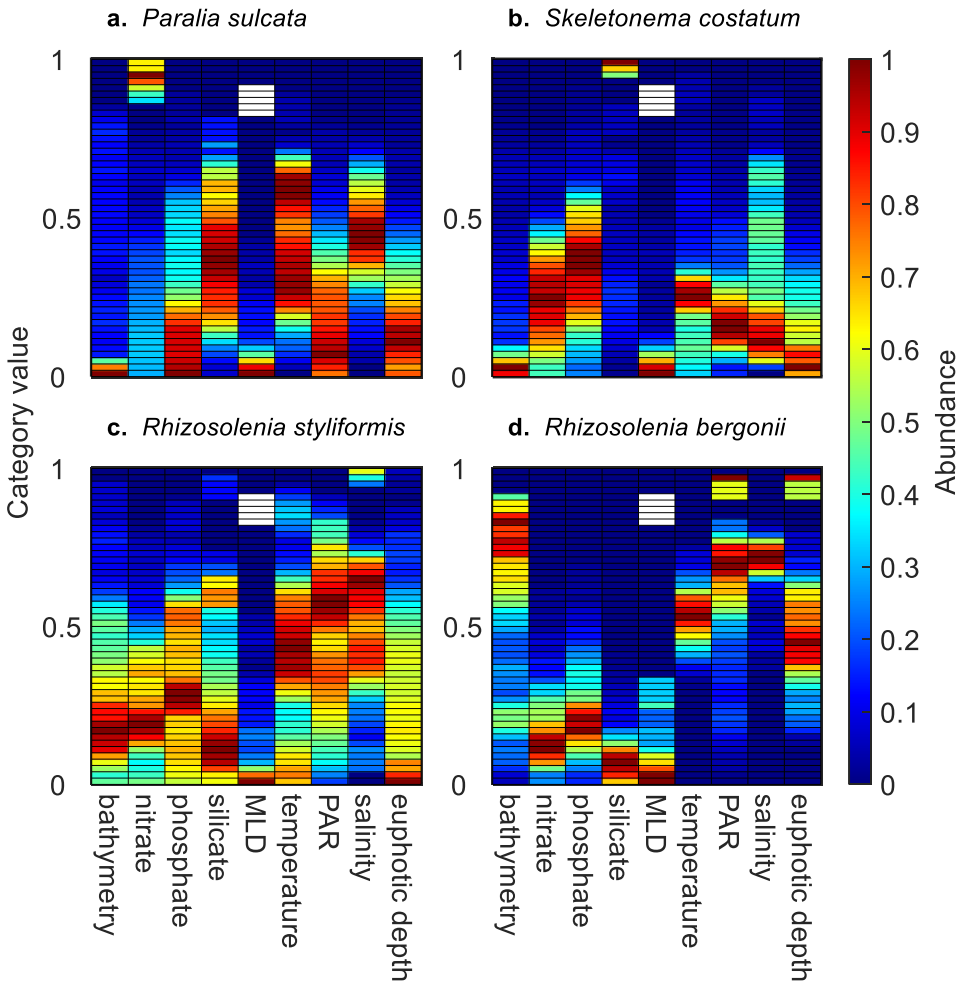


**Figure S6. Species chromatograms of four copepods with the application of the second-order simple moving average and the standardisation of the abundance between 0 and 1.** Species chromatograms for (**a**) *Temora longicornis*, (**b**) *Clausocalanus* spp., (**c**) *Calanus finmarchicus* and (**d**) *Calanus helgolandicus*. In **a**-**d**, each column represents the mean species abundance along seven environmental dimensions (i.e. bathymetry, MLD, temperature, PAR, salinity, chlorophyll-a concentration and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 CPR sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (standardised between 0 and 1) in each category. High abundance values are in red and low values in blue.

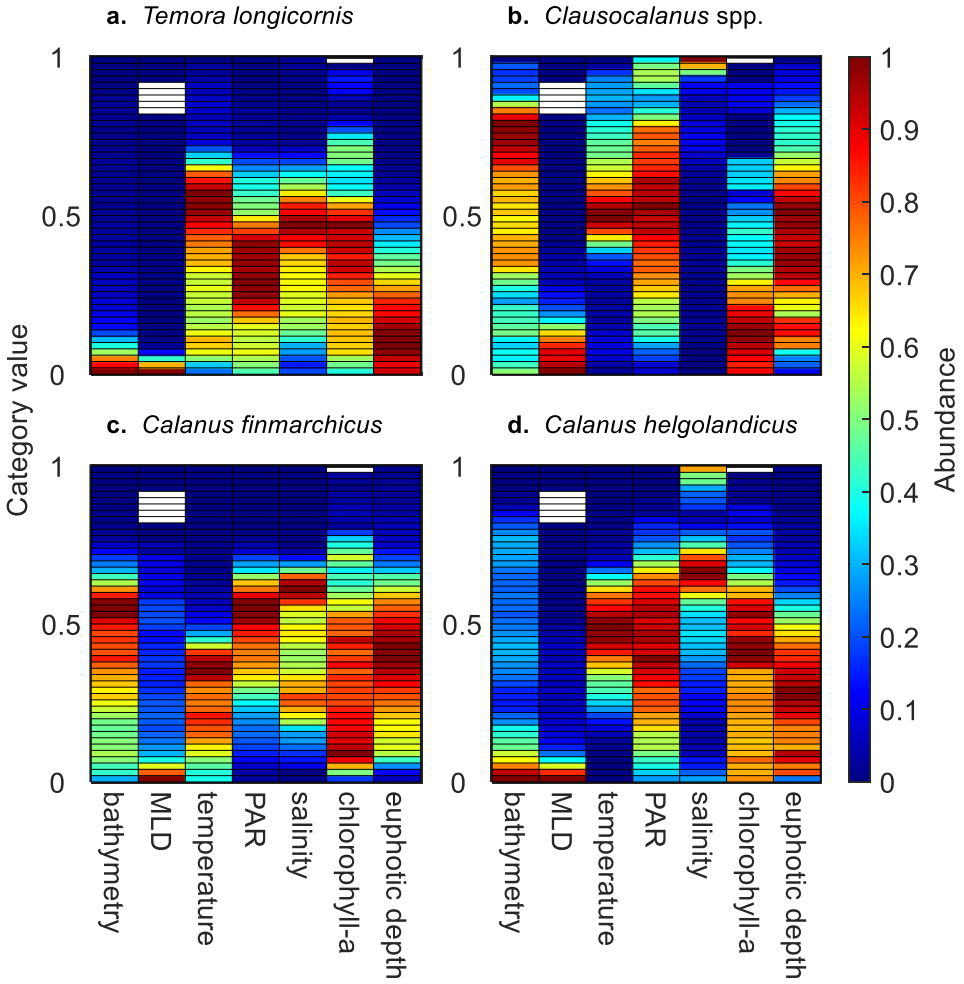


**Figure S7. Total number of samples ($log_{10}$ transformed) available in each category.**

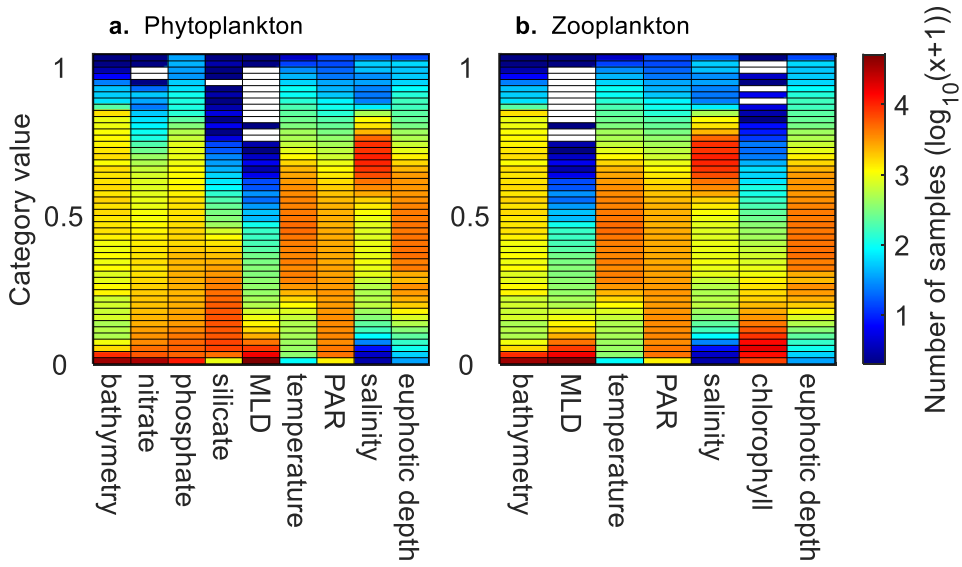



**Figure S8. Species chromatograms of the 14 pseudo-species.** Each panel displays the three-dimensional ecological niche of a pseudo-species (i.e. virtual species). In **a**-**n**, each column represents the mean pseudo-species abundance along three environmental dimensions (i.e. $p_1$, $p_2$ and $p_3$). Pseudo-species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 1 sample was available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours represent the mean species abundance (standardised between 0 and 1) in each category. High abundance values are in red and low values in blue.

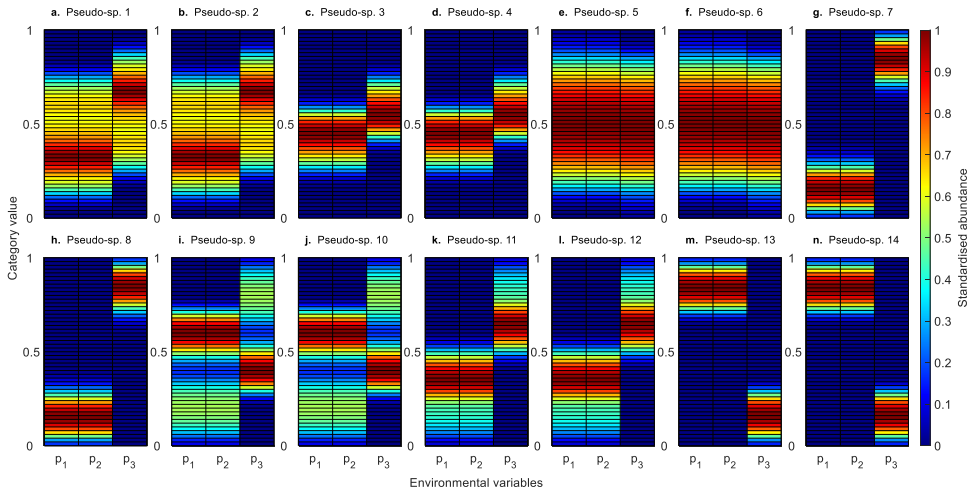


**Figure S9. Three-dimensional ecological niches of the 14 pseudo-species display with the *hypervolume* package.** Each panel displays the three-dimensional ecological niche of a pseudo-species as a set of pair plots of the pseudo-species abundance repartition along the three environmental dimensions $p_1$, $p_2$ and $p_3$. Niche hypervolumes were delineated by means of a Gaussian kernel density estimation.

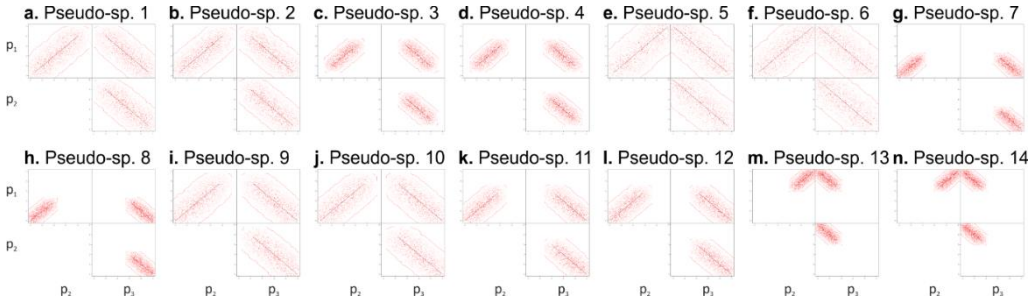
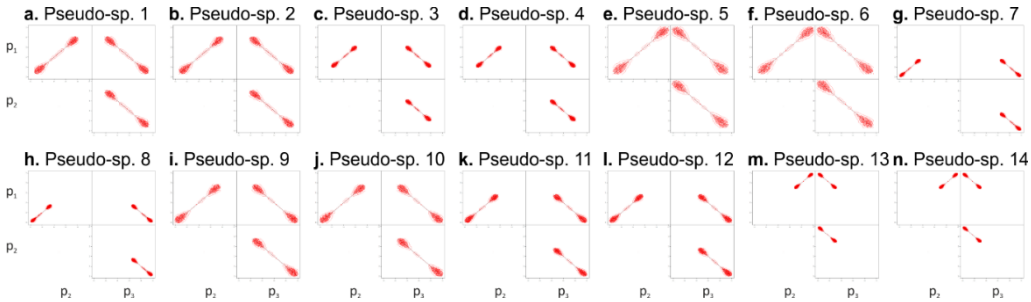



**Figure S10. Three-dimensional ecological niches of the 14 pseudo-species display with the *hypervolume* package.** Each panel displays the three-dimensional ecological niche of a pseudo-species as a set of pair plots of the pseudo-species abundance repartition along the three environmental dimensions $p_1$, $p_2$ and $p_3$. Niche hypervolumes were delineated by means of a one-class support vector machine.




**Supplementary Tables.**

**Table S1. Glossary of the technical terms used in this paper.**

| Term | Definition | Reference |
|---|---|---|
| **Ecological niche** | Although different definitions of the niche have been proposed, the concept is here defined as the set of environmental conditions allowing a species to growth, maintain and reproduce. The niche is therefore assimilated to a p-dimensional hypervolume (with p corresponding to the number of environmental variables) in which each environmental combination enables a species to exist indefinitely (i.e. the species fundamental niche). | (Hutchinson, 1978, 1957) |
| **Community chromatogram** | A graphic, originally called environmental chromatogram, that displays where species of an assemblage aggregate along multiple environmental dimensions. | (Kléparski et al., 2021) |

| Term | Definition | Reference |
|---|---|---|
| **Species chromatogram** | A graphic that displays the (multidimensional) ecological niche of a species into a two-dimensional space. | - |
| **Environmental signature** | The set of environmental conditions where a species assemblage (or a community) is typically found. | - |
| **Niche optimum** | The average environmental value where a species has its highest abundance along one niche dimension. | - |
| **Niche breadth** | A measure of the environmental range where a species is abundant. In our study, niche breadth was assessed by the number of categories of the species chromatogram where abundance was higher or equal to a threshold T along an environmental gradient, with $0<T\leq1$. Five values of T were tested: T=0, T=0.05, T=0.1, T=0.25 and T=0.5. | - |
| **Average niche breadth** | The average of niche breadth based on all environmental dimensions of the species chromatogram. | - |
| **Degree of niche overlapping** | The percentage of the niche hypervolume common to different species. | - |
| **Eury/stenography** | Terms used to characterise the spatial distribution of a species. A species with a large spatial distribution is eurygraphic whereas a species with a narrow spatial distribution is stenographic. | - |
| **Environmental gradient** | This term refers to the continuous variation of an environmental variable from its lowest to its highest value. | - |
| **Environmental variable** | The state of a biotic (e.g. chlorophyll-a concentration) or abiotic (e.g. sea water temperature) parameter that can be quantitative or qualitive. Only quantitative variables were used in this study. | - |
| **Environmental dimension** | An axis of the niche defined by an environmental variable. | - |
| **Coloured band** | In a community chromatogram, an aggregation of more or less continuous categories with a significant percentage of aggregation. In a species chromatogram, more or less continuous categories with an abundance above a threshold along an environmental dimension. | - |

**Table S2. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four diatoms based on a threshold of abundance T=0.** The mean niche breadth ($E_T$) for all dimensions and each species is also displayed in bold. The ecological niche of each diatom is shown in Figure 26.

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 100.00 | 95.83 | 100.00 | 95.83 |
| **Nitrate (%)** | 91.30 | 84.78 | 100.00 | 89.13 |
| **Phosphate (%)** | 92.00 | 92.00 | 100.00 | 94.00 |
| **Silicate (%)** | 100.00 | 100.00 | 97.22 | 80.56 |
| **MLD (%)** | 86.67 | 83.33 | 93.33 | 60.00 |
| **Temperature (%)** | 97.96 | 97.96 | 97.96 | 85.71 |
| **PAR (%)** | 91.67 | 87.50 | 95.83 | 89.58 |
| **Salinity (%)** | 100.00 | 93.62 | 87.23 | 51.06 |
| **Euphotic depth (%)** | 100.00 | 87.76 | 100.00 | 89.80 |
| **$E_T$ (%)** | **95.51** | **91.42** | **96.84** | **81.74** |

**Table S3. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four diatoms based on a threshold of abundance T=0.05.** The mean niche breadth ($E_T$) for all dimensions and each species is also displayed in bold. The ecological niche of each diatom is shown in Figure 26.

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 87.50 | 72.92 | 100.00 | 95.83 |
| **Nitrate (%)** | 84.78 | 76.09 | 84.78 | 52.17 |
| **Phosphate (%)** | 68.00 | 78.00 | 90.00 | 56.00 |
| **Silicate (%)** | 97.22 | 100.00 | 97.22 | 44.44 |
| **MLD (%)** | 60.00 | 66.67 | 90.00 | 56.67 |
| **Temperature (%)** | 81.63 | 97.96 | 97.96 | 40.82 |
| **PAR (%)** | 75.00 | 62.50 | 95.83 | 77.08 |
| **Salinity (%)** | 74.47 | 89.36 | 80.85 | 42.55 |
| **Euphotic depth (%)** | 59.18 | 51.02 | 100.00 | 83.67 |
| **$E_T$ (%)** | **76.42** | **77.17** | **92.96** | **61.03** |

**Table S4. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four diatoms based on a threshold of abundance T=0.1.** The mean niche breadth ($E_T$) for all dimensions and each species is also displayed in bold. The ecological niche of each diatom is shown in Figure 26.

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 85.42 | 35.42 | 100.00 | 91.67 |
| **Nitrate (%)** | 80.43 | 56.52 | 67.39 | 43.48 |
| **Phosphate (%)** | 62.00 | 62.00 | 72.00 | 50.00 |
| **Silicate (%)** | 91.67 | 94.44 | 97.22 | 33.33 |
| **MLD (%)** | 40.00 | 23.33 | 53.33 | 56.67 |
| **Temperature (%)** | 69.39 | 63.27 | 95.92 | 32.65 |
| **PAR (%)** | 62.50 | 52.08 | 93.75 | 72.92 |
| **Salinity (%)** | 72.34 | 72.34 | 78.72 | 19.15 |
| **Euphotic depth (%)** | 53.06 | 44.90 | 81.63 | 83.67 |
| **$E_T$ (%)** | **68.53** | **56.03** | **82.22** | **53.73** |

**Table S5. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four diatoms based on a threshold of abundance T=0.5.** The mean niche breadth ($E_T$) for all dimensions and each species is also displayed in bold. The ecological niche of each diatom is shown in Figure 26.

| | ***Paralia sulcata*** | ***Skeletonema costatum*** | ***Rhizosolenia styliformis*** | ***Rhizosolenia bergonii*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 4.17 | 6.25 | 37.50 | 77.08 |
| **Nitrate (%)** | 45.65 | 39.13 | 36.96 | 21.74 |
| **Phosphate (%)** | 26.00 | 40.00 | 60.00 | 16.00 |
| **Silicate (%)** | 72.22 | 83.33 | 91.67 | 19.44 |
| **MLD (%)** | 10.00 | 13.33 | 10.00 | 16.67 |
| **Temperature (%)** | 51.02 | 12.24 | 65.31 | 16.33 |
| **PAR (%)** | 37.50 | 20.83 | 62.50 | 25.00 |
| **Salinity (%)** | 34.04 | 19.15 | 72.34 | 12.77 |
| **Euphotic depth (%)** | 32.65 | 20.41 | 51.02 | 61.22 |
| **$E_T$ (%)** | **34.81** | **28.30** | **54.14** | **29.58** |

**Table S6. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four copepods based on a threshold of abundance T=0.** The mean niche breadth ($E_T$) for all dimensions and each species/taxa is also displayed in bold. The ecological niche of each copepod is shown in Figure 27.

| | ***Temora longicornis*** | ***Clausocalanus spp.*** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 97.92 | 100.00 | 97.92 | 97.92 |
| **MLD (%)** | 100.00 | 93.33 | 100.00 | 93.33 |
| **Temperature (%)** | 100.00 | 100.00 | 100.00 | 95.92 |
| **PAR (%)** | 97.92 | 100.00 | 100.00 | 100.00 |
| **Salinity (%)** | 100.00 | 97.87 | 89.36 | 100.00 |
| **Chlorophyll-a (%)** | 100.00 | 100.00 | 100.00 | 100.00 |
| **Euphotic depth (%)** | 100.00 | 100.00 | 100.00 | 100.00 |
| **$E_T$ (%)** | **99.40** | **98.74** | **98.18** | **98.17** |

**Table S7. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four copepods based on a threshold of abundance T=0.05.** The mean niche breadth ($E_T$) for all dimensions and each species/taxa is also displayed in bold. The ecological niche of each copepod is shown in Figure 27.

| | ***Temora longicornis*** | ***Clausocalanus spp.*** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 33.33 | 100.00 | 79.17 | 89.58 |
| **MLD (%)** | 13.33 | 50.00 | 100.00 | 76.67 |
| **Temperature (%)** | 100.00 | 95.92 | 59.18 | 59.18 |
| **PAR (%)** | 75.00 | 100.00 | 72.92 | 87.50 |
| **Salinity (%)** | 72.34 | 91.49 | 72.34 | 97.87 |
| **Chlorophyll-a (%)** | 100.00 | 89.47 | 100.00 | 100.00 |
| **Euphotic depth (%)** | 59.18 | 100.00 | 85.71 | 73.47 |
| **$E_T$ (%)** | **64.74** | **89.55** | **81.33** | **83.47** |

**Table S8. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four copepods based on a threshold of abundance T=0.1.** The mean niche breadth ($E_T$) for all dimensions and each species/taxa is also displayed in bold. The ecological niche of each copepod is shown in Figure 27.

| | ***Temora longicornis*** | ***Clausocalanus spp.*** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 20.83 | 100.00 | 77.08 | 89.58 |
| **MLD (%)** | 13.33 | 43.33 | 100.00 | 23.33 |
| **Temperature (%)** | 73.47 | 65.31 | 53.06 | 53.06 |
| **PAR (%)** | 72.92 | 95.83 | 68.75 | 87.50 |
| **Salinity (%)** | 70.21 | 38.30 | 72.34 | 80.85 |
| **Chlorophyll-a (%)** | 100.00 | 89.47 | 100.00 | 100.00 |
| **Euphotic depth (%)** | 53.06 | 100.00 | 73.47 | 69.39 |
| **$E_T$ (%)** | **57.69** | **76.03** | **77.81** | **71.96** |

**Table S9. Niche breadth (ecological niche breadth) assessed from the species chromatogram of the four copepods based on a threshold of abundance T=0.5.** The mean niche breadth ($E_T$) for all dimensions and each species/taxa is also displayed in bold. The ecological niche of each copepod is shown in Figure 27.

| | ***Temora longicornis*** | ***Clausocalanus spp.*** | ***Calanus finmarchicus*** | ***Calanus helgolandicus*** |
|---|---|---|---|---|
| **Bathymetry (%)** | 8.33 | 89.58 | 60.42 | 10.42 |
| **MLD (%)** | 6.67 | 26.67 | 10.00 | 10.00 |
| **Temperature (%)** | 63.27 | 26.53 | 38.78 | 32.65 |
| **PAR (%)** | 54.17 | 81.25 | 37.50 | 62.50 |
| **Salinity (%)** | 48.94 | 6.38 | 51.06 | 17.02 |
| **Chlorophyll-a (%)** | 94.74 | 42.11 | 73.68 | 86.84 |
| **Euphotic depth (%)** | 34.69 | 63.27 | 55.10 | 51.02 |
| **$E_T$ (%)** | **44.40** | **47.97** | **46.65** | **38.64** |

**Table S10. Mean degree of niche overlapping for the four diatoms based on a threshold of abundance T=0.** The first column displays the number of dimensions considered simultaneously, columns 2 to 10 display the combinations of dimensions (i.e. 1=bathymetry, 2=nitrate, 3=phosphate, 4=silicate, 5=MLD, 6=temperature, 7=PAR, 8=salinity and 9=euphotic depth). The last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each species is displayed in Figure 26.

| Number of dimensions | Combinations | | | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 8 | | | | | | | | | 73.03 |
| **2** | 5 | 8 | | | | | | | | 61.31 |
| **3** | 5 | 8 | 9 | | | | | | | 55.55 |
| **4** | 2 | 5 | 8 | 9 | | | | | | 50.92 |
| **5** | 2 | 4 | 5 | 8 | 9 | | | | | 47.56 |
| **6** | 2 | 4 | 5 | 7 | 8 | 9 | | | | 45.29 |
| **7** | 2 | 3 | 4 | 5 | 7 | 8 | 9 | | | 43.46 |
| **8** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | | 41.99 |
| **9** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 40.88 |

**Table S11. Mean degree of niche overlapping for the four diatoms based on a threshold of abundance T=0.05.** The first column displays the number of dimensions considered simultaneously, columns 2 to 10 display the combinations of dimensions (i.e. 1=bathymetry, 2=nitrate, 3=phosphate, 4=silicate, 5=MLD, 6=temperature, 7=PAR, 8=salinity and 9=euphotic depth). The last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each species is displayed in Figure 26.

| Number of dimensions | Combinations | | | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 9 | | | | | | | | | 59.61 |
| **2** | 6 | 9 | | | | | | | | 42.65 |
| **3** | 6 | 8 | 9 | | | | | | | 33.05 |
| **4** | 6 | 7 | 8 | 9 | | | | | | 26.02 |
| **5** | 5 | 6 | 7 | 8 | 9 | | | | | 20.85 |
| **6** | 3 | 5 | 6 | 7 | 8 | 9 | | | | 17.36 |
| **7** | 1 | 3 | 5 | 6 | 7 | 8 | 9 | | | 15.09 |
| **8** | 1 | 2 | 3 | 5 | 6 | 7 | 8 | 9 | | 13.28 |
| **9** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 12.02 |

**Table S12. Mean degree of niche overlapping for the four diatoms based on a threshold of abundance T=0.1.** The first column displays the number of dimensions considered simultaneously, columns 2 to 10 display the combinations of dimensions (i.e. 1=bathymetry, 2=nitrate, 3=phosphate, 4=silicate, 5=MLD, 6=temperature, 7=PAR, 8=salinity and 9=euphotic depth). The last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each species is displayed in Figure 26.

| Number of dimensions | Combinations | | | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 8 | | | | | | | | | 55.19 |
| **2** | 1 | 8 | | | | | | | | 32.07 |
| **3** | 1 | 8 | 9 | | | | | | | 21.63 |
| **4** | 1 | 5 | 6 | 8 | | | | | | 14.64 |
| **5** | 1 | 5 | 6 | 8 | 9 | | | | | 10.05 |
| **6** | 1 | 5 | 6 | 7 | 8 | 9 | | | | 7.21 |
| **7** | 1 | 2 | 5 | 6 | 7 | 8 | 9 | | | 5.89 |
| **8** | 1 | 2 | 4 | 5 | 6 | 7 | 8 | 9 | | 5.17 |
| **9** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 4.67 |

**Table S13. Mean degree of niche overlapping for the four diatoms based on a threshold of abundance T=0.5.** The first column displays the number of dimensions considered simultaneously, columns 2 to 10 display the combinations of dimensions (i.e. 1=bathymetry, 2=nitrate, 3=phosphate, 4=silicate, 5=MLD, 6=temperature, 7=PAR, 8=salinity and 9=euphotic depth). The last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each species is displayed in Figure 26.

| Number of dimensions | Combinations | | | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 8 | | | | | | | | | 13.61 |
| **2** | 1 | 8 | | | | | | | | 0.92 |
| **3** | 1 | 8 | 9 | | | | | | | 0.25 |
| **4** | 1 | 6 | 8 | 9 | | | | | | 0.07 |
| **5** | 1 | 6 | 7 | 8 | 9 | | | | | 0.02 |
| **6** | 1 | 4 | 6 | 7 | 8 | 9 | | | | 0.01 |
| **7** | 1 | 3 | 4 | 6 | 7 | 8 | 9 | | | 0.00 |
| **8** | 1 | 2 | 3 | 4 | 6 | 7 | 8 | 9 | | 0.00 |
| **9** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0.00 |

**Table S14. Mean degree of niche overlapping for the four copepods based on a threshold of abundance T=0.** The first column displays the number of dimensions considered simultaneously, columns 2 to 8 display the combinations of dimensions (i.e. 1=bathymetry, 2=MLD, 3=temperature, 4=PAR, 5=salinity, 6=chlorophyll-a concentration and 7=euphotic depth) and the last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each copepod is shown in Figure 27.

| Number of dimensions | Combinations | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|
| **1** | 5 | | | | | | | 93.62 |
| **2** | 2 | 5 | | | | | | 89.65 |
| **3** | 2 | 3 | 5 | | | | | 87.85 |
| **4** | 1 | 2 | 3 | 5 | | | | 86.98 |
| **5** | 1 | 2 | 3 | 4 | 5 | | | 86.13 |
| **6** | 1 | 2 | 3 | 4 | 5 | 6 | | 86.13 |
| **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 86.13 |

**Table S15. Mean degree of niche overlapping for the four copepods based on a threshold of abundance T=0.05.** The first column displays the number of dimensions considered simultaneously, columns 2 to 8 display the combinations of dimensions (i.e. 1=bathymetry, 2=MLD, 3=temperature, 4=PAR, 5=salinity, 6=chlorophyll-a concentration and 7=euphotic depth) and the last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each copepod is shown in Figure 27.

| Number of dimensions | Combinations | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|
| **1** | 2 | | | | | | | 41.55 |
| **2** | 2 | 3 | | | | | | 30.70 |
| **3** | 1 | 2 | 3 | | | | | 23.04 |
| **4** | 1 | 2 | 3 | 5 | | | | 18.96 |
| **5** | 1 | 2 | 3 | 5 | 7 | | | 15.77 |
| **6** | 1 | 2 | 3 | 4 | 5 | 7 | | 13.66 |
| **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 13.30 |

**Table S16. Mean degree of niche overlapping for the four copepods based on a threshold of abundance T=0.1.** The first column displays the number of dimensions considered simultaneously, columns 2 to 8 display the combinations of dimensions (i.e. 1=bathymetry, 2=MLD, 3=temperature, 4=PAR, 5=salinity, 6=chlorophyll-a concentration and 7=euphotic depth) and the last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each copepod is shown in Figure 27.

| Number of dimensions | Combinations | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|
| **1** | 2 | | | | | | | 36.96 |
| **2** | 2 | 5 | | | | | | 17.37 |
| **3** | 1 | 2 | 5 | | | | | 10.78 |
| **4** | 1 | 2 | 3 | 5 | | | | 6.75 |
| **5** | 1 | 2 | 3 | 5 | 7 | | | 5.44 |
| **6** | 1 | 2 | 3 | 4 | 5 | 7 | | 4.87 |
| **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 4.77 |

**Table S17. Mean degree of niche overlapping for the four copepods based on a threshold of abundance T=0.5.** The first column displays the number of dimensions considered simultaneously, columns 2 to 8 display the combinations of dimensions (i.e. 1=bathymetry, 2=MLD, 3=temperature, 4=PAR, 5=salinity, 6=chlorophyll-a concentration and 7=euphotic depth) and the last column displays index D associated with the combination of environmental dimensions. D=0% when species niches are different and D=100% when species niches are identical; the higher the number of dimensions, the lower the value of index D. Only the combinations of environmental variables that minimise values of index D are displayed. The ecological niche of each copepod is shown in Figure 27.

| Number of dimensions | Combinations | | | | | | | Index D (%) |
|---|---|---|---|---|---|---|---|---|
| **1** | 5 | | | | | | | 16.40 |
| **2** | 1 | 5 | | | | | | 1.16 |
| **3** | 1 | 3 | 5 | | | | | 0.79 |
| **4** | 1 | 3 | 5 | 7 | | | | 0.53 |
| **5** | 1 | 3 | 4 | 5 | 7 | | | 0.40 |
| **6** | 1 | 2 | 3 | 4 | 5 | 7 | | 0.33 |
| **7** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0.31 |

## 2.3. How do plankton species coexist in an apparently unstructured environment?

Loïck Kléparski, Grégory Beaugrand and Richard R. Kirby

*Published in Biology Letters. Supplementary Files are displayed in section 2.3.5. Article is available at* https://dx.doi.org/10.1098/rsbl.2022.0207

*Abstract*

In a paper entitled the paradox of plankton, Hutchinson asked "*how it is possible for a number of species to coexist in a relatively isotropic or unstructured environment all competing for the same sorts of materials*". Particularly relevant for phytoplankton, this paradox was based on two implicit, and perhaps naive, postulates, i.e. (i) that all plankton species have similar requirements and (ii) that the marine environment is relatively homogeneous in space and time. A number of hypotheses, based on purely theoretical or experimental studies, have been proposed to solve this conundrum, ranging from spatio-temporal environmental heterogeneity to biotic chaotic variability. Here, we characterize the ecological niche of 117 plankton species belonging to three different taxonomic groups and show that all species have a niche sufficiently distinct to ensure coexistence in a structured marine environment. We also provide evidence that pelagic habitats are, unsurprisingly, more diverse in space and time than Hutchinson imagined, the marine environment being neither unstructured nor stable in space and time. We therefore conclude that the niche theory, and its corollary the principle of competitive exclusion, apply as much for the plankton as for other forms of life, be they terrestrial or marine.

*Keywords*

Plankton paradox, niche theory, biodiversity, pelagic environment, principle of competitive exclusion.

## 2.3.1. Introduction

Marine plankton is composed of both uni- and multicellular organisms covering a wide range of size, from less than 0.2 µm (virioplankton) to a few meters (e.g. the giant jellyfish *Nemopilema nomurai*)(Kirby, 2010). They are at the basis of the food web and play a key role in carbon dioxide regulation, influencing the composition of the Earth's atmosphere (Cushing, 1990; Kirby, 2010; Passow and Carlson, 2012; Platt et al., 2003). The great diversity they exhibit has fascinated marine biologists for a long time (Haeckel, 1904; Hardy, 1958, 1926) and it is now apparent that there is even more diversity than originally realised (De Vargas et al., 2015; Goetze, 2003). The mechanisms sustaining the plankton biodiversity remain debated, however, because it is assumed that plankton species coexist in an apparently homogeneous environment with few niches (*sensu* Hutchinson realised niche i.e. the set of environmental conditions modulated by biotic interactions, enabling a species to growth and reproduce (Hutchinson, 1957)) thought to be available (Behrenfeld et al., 2021; Hutchinson, 1961).

Although it has been suggested that plankton species may share niches and coexist (Ghilarov, 1984), many theoretical and laboratory studies have tried to resolve the "paradox of the plankton"(Hutchinson, 1961) and have shown that the number of coexisting plankton species at equilibrium in a community cannot exceed the number of limiting resources (Levin, 1970; Mac Arthur and Levins, 1964; Rothhaupt, 1988; Tilman, 1977). Consequently, coexistence of a large number of species on a small number of limiting resources is only possible when other mechanisms are involved such as (i) temporal heterogeneity (i.e. a continuously changing pelagic environment due to the seasonal changes in hydro-meteorological forcing) and its influence on the structure of the water column (Bracco et al., 2000; Hutchinson, 1961; Levins, 1979; Richerson et al., 1970), (ii) spatial heterogeneity in the relative resource supply rate (Tilman, 1980)), (iii) spatial subdivision (Tilman, 1994), (iv) chaotic variations in species abundance induced by competition (Huisman and Weissing, 1999) or (v) difference in the array of resources limiting species growth (Petersen, 1975; Tilman, 1977). The tenet that underlies the paradox of the plankton —that the plankton environment is unstructured and so different plankton share the same niche, and therefore the question of whether niche separation might promote plankton coexistence— has received little attention, however.

In this article, we examine how niche separation can enable plankton coexistence. Niche separation, which is often estimated by the degree of overlapping (i.e. the opposite of niche separation)(Beaugrand et al., 2013b), results from the principle of competitive exclusion (Gause, 1934; Hutchinson, 1978). By allowing a better exploitation of the available resources, this mechanism has been shown to enable terrestrial plant species coexistence (Silvertown, 2004; Tilman, 1999). Here, we examine the two postulates on which the paradox of the plankton is based, i.e. (i) do plankton species share niches and (ii) is the oceanic environment unstructured. To do so, we estimate the degree of niche overlapping of 117 plankton species belonging to three taxonomic groups (diatoms, dinoflagellates and calanoid copepods). We show that species' niches of the three taxonomic groups are well separated and that the

environment is much more structured in space and time than Hutchinson envisioned (Hutchinson, 1961). We conclude that these two mechanisms jointly ensure plankton species coexistence.

## 2.3.2. Materials and Methods

### 2.3.2.1. Biological data

Plankton abundance data came from the Continuous Plankton Recorder (CPR) survey (Text S1)(Batten et al., 2003). It is a long-term plankton monitoring programme using a high-speed plankton recorder sampling at a depth of ~7-10m. Started in 1931, the CPR programme has since sampled plankton on a monthly basis in the North Atlantic Ocean and its adjacent seas (Hays and Warner, 1993; Warner and Hays, 1994). We used the data collected between 1998 and 2018 in the North Atlantic Ocean and its adjacent seas (Helaouët, 2021). This time period was chosen because it corresponds to the period covered by the environmental datasets we used (see Sections b and c). Species that were found in less than 100 CPR samples were removed from the analysis so that the total number of species used was 117; this reflected abundance data for 71 phytoplankton (i.e. 45 diatoms and 26 dinoflagellates) and 46 calanoid copepod taxa.

### 2.3.2.2. Environmental data

We used a set of nine different environmental variables to characterise the niches of the 117 species. The chosen variables were: bathymetry (in m), nitrate, phosphate and silicate concentrations ($mmol.m^{-3}$), Mixed Layer Depth (MLD, m), temperature (°C), Photosynthetically Active Radiations (PAR, $E.m^{-2}.day^{-1}$), salinity (no unit) and chlorophyll-a concentration ($mg.m^{-3}$). Detail descriptions and references for each dataset are provided in Text S2.

### 2.3.2.3. Data pre-processing

We used the environmental data collected between 1998 and 2018 to work on a common time period with respect to all datasets. All data were subsequently arranged on a grid covering the North Atlantic Ocean (100°W-10°E and 35°N-65°N). By means of nearest-neighbour interpolation (Wackernagel, 1995), we attributed to each CPR sample a value for each of the nine chosen environmental variables at a depth of 8 m (except for bathymetry), a value included in the range of sampling depth of the CPR instrument (Batten et al., 2003; Hays and Warner, 1993). All the samples with a missing value along a single dimension were discarded prior to the analyses. We therefore used a total of 90,527 CPR samples.

### 2.3.2.4. Niche separation

To characterise and compare the multidimensional niches of the 117 planktonic species, we used the recently proposed “species chromatogram” method (Kléparski et al., 2021; Kléparski and Beaugrand, 2022); this technique uses a chromatogram to display the multidimensional niche of a species in a two-dimensional

space. To do so, p-standardised environmental gradients (one per environmental variable, e.g. temperature or nutrients), representing the p-niche environmental dimensions, are defined between 0 (the lowest value taken by an environmental variable) and 1 (the highest). Each gradient is divided into α equidistant categories (here α=50 for all chromatograms). For each category along each dimension (i.e. gradient), the mean species abundance is estimated by averaging the values of the k% of the CPR samples (associated to that category by means of the nearest-neighbour interpolation, see the previous section) with the highest abundance values (here k= 5%), if at least m samples are present in that category (here m=20 CPR samples). The mean abundances along each dimension are then standardised between 0 and 1, and smoothed by means of a second-order simple moving average to alleviate the noise associated with the CPR data. Finally, the niche is displayed as an array of coloured cells, which is called a species chromatogram (Figure S1-12)(Kléparski and Beaugrand, 2022).

As each environmental gradient is globally standardised between 0 and 1, the niches of different species can be compared and a degree of niche overlapping (called index D) can be estimated among species (Kléparski and Beaugrand, 2022). This index is estimated by considering the categories with an abundance greater than or equal to a threshold T along each dimension, with $0<T\leq1$. They are then used to calculate the niche hypervolume, the shape of the niche being assimilated to a hyperrectangle. D is finally assessed for each pair of species by calculating the ratio of the hypervolume of the niche common to the two species on the total volume filled by the two species. The method is fully described in Kléparski & Beaugrand (Kléparski and Beaugrand, 2022).

Here, D was estimated for each pair of plankton species and for each combination of 1 to p-environmental dimensions (i.e. 1 to p-dimensions considered simultaneously) and five thresholds T: 0, 0.1, 0.25, 0.5 and 0.75. For phytoplankton (i.e. diatoms and dinoflagellates), we chose eight environmental dimensions (p=8): bathymetry, nitrate, phosphate and silicate concentrations, MLD, temperature, PAR and salinity (Figure 29 and Figure S1-8). For copepods, we used six dimensions (p=6): bathymetry, MLD, temperature, PAR, salinity and chlorophyll-a concentration (Figure 29 and Figure S9-12). D was calculated among species belonging to the same taxonomic group, i.e. among diatoms (45 species), dinoflagellates (26 species), phytoplankton (45 diatoms + 26 dinoflagellates = 71 species) or calanoid copepods (46 species).

All species chromatograms on which index D is based are displayed in Figures S1-12. We have synthetized the information of all species chromatograms by means of a community chromatogram (Kléparski et al., 2021; Kléparski and Beaugrand, 2022). To do so, species' abundance was first replaced by 1 when its standardised abundance was higher than or equal to 0.5 in a category (Kléparski et al., 2021)(see Figure 29a-d). Then, we estimated the number of co-occurring species in each category of the chromatogram and we converted it into percentage. Therefore, a community chromatogram displays the percentage of co-occurring species in each environmental category.

#### 2.3.2.5. Habitat diversity

Habitat diversity of the pelagic environment was estimated with respect to each environmental variable (but bathymetry that is constant) by means of the Shannon entropy (Magurran, 2004). In other words, we quantified the diversity of values that is taken by an environmental variable for each day or in each geographical cell. To do so, eight environmental matrices (121 latitudes x 441 longitudes x 24 depth levels (except for MLD) x 7670 days) were used and standardised between 0 (i.e. the lowest value of an environmental variable at the scale of the North Atlantic) and 1 (i.e. the highest). The environmental values taken by each matrix were then divided into 50 categories and the habitat diversity $H(S)$ (expressed in bit) was calculated as follows:

$$H(S) = -\sum_{i=1}^{50} S_i \times log_2(S_i) \qquad (1)$$

With $S_i$ the probability that the value of the $i^{th}$ category was observed, i.e. its relative frequency. The estimate of the Shannon entropy was performed in two ways for each year between 1998 and 2018: (i) at a daily scale using all geographical cells and depth levels between 0 and 20 m and (ii) at a spatial scale using all the days and depth levels between 0 and 20m. The resulting annual cycles were then averaged (1998-2018).

### 2.3.3. Results

#### 2.3.3.1. Niche separation

Species chromatograms are all displayed in Figure S1-12. Their examination revealed that each species has a unique chromatogram and therefore a specific niche, although many overlaps were observed along all niche dimensions. The species chromatograms were summarised by taxonomic group under the form of a community chromatogram, where species abundance was replaced by the percentage of co-occurring species in an environmental category (Figure 29a-d). At this organisational level, niche separation occurred between diatoms and dinoflagellates along five niche dimensions (bathymetry, silicate, temperature, PAR and salinity; Figure 29b-c). Similarly, a clear niche separation can be seen between phytoplankton and copepods with respect to salinity, although the two groups have overlapping chromatograms for bathymetry, temperature and PAR (Figure 29a and d). Hence, each group seems to have distinct environmental requirements, although niche overlaps are the rule.

The calculation of the degree of niche overlapping (i.e. Index D) confirmed that species niches were rapidly separated when the number of environmental dimensions of the niche rose (Figure 29e-h). Niche overlapping diminished rapidly below 30% for the four taxonomic groups (Figure 29e-h). We therefore investigated whether T (i.e. the threshold of abundance between 0 and 1) influenced our perception of niche separation. When T increased from 0 to 0.75, overlapping rapidly reached a value below 20% and even close to 0% when T becomes higher than 0.25. Consequently, the niches of the four groups were almost totally separated towards the optimal part of their niche (i.e. the part of the niche where a species has its highest abundance (Brown,

1984; Helaouët and Beaugrand, 2009)), although some overlapping was still observed in the suboptimal parts (i.e. the parts where a species has low abundance, i.e. when T<0.25) (Figure 29e-h), as might be the expected outcome of competition and natural selection (Begon et al., 2005).

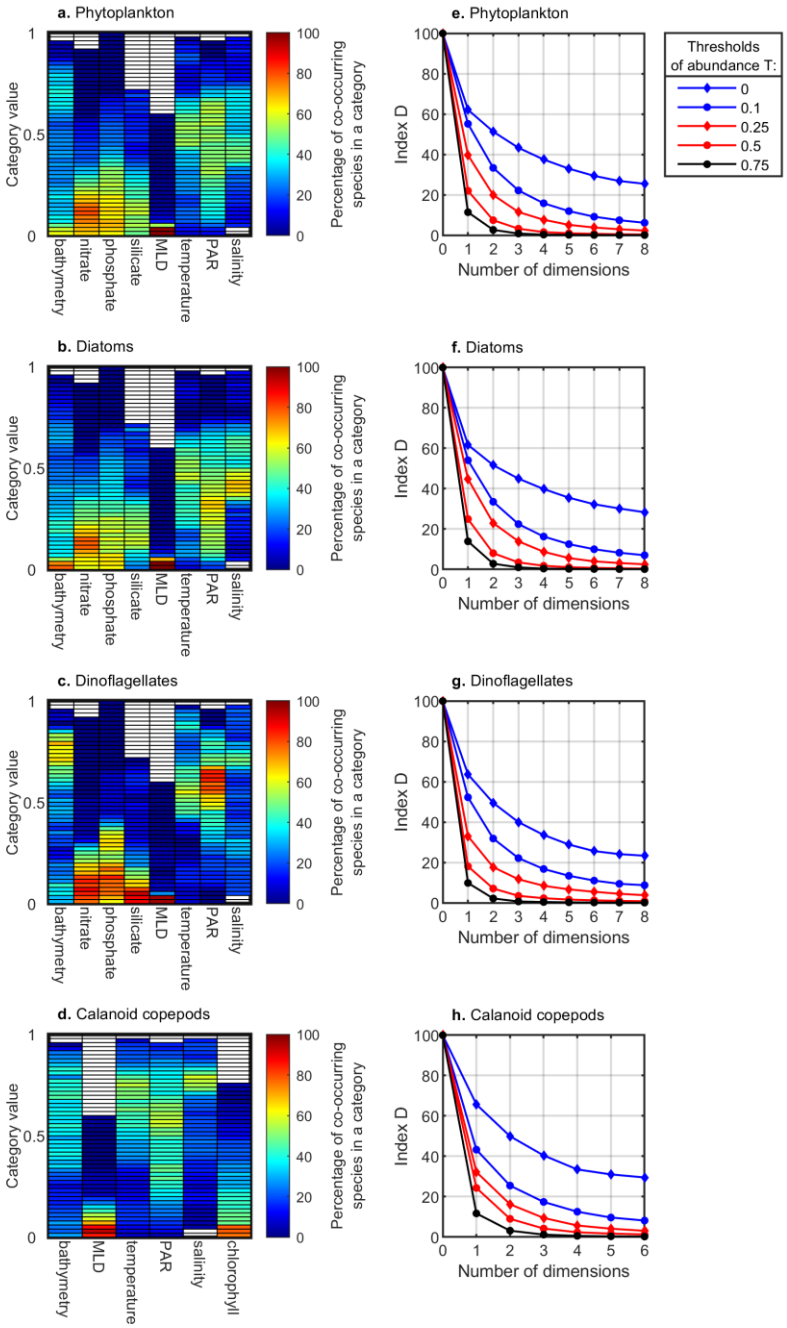



*Figure 29. Community chromatograms and changes in the degree of niche overlapping for four groups of plankton. Community chromatogram of (a) phytoplankton, (b) diatoms, (c) dinoflagellates and (d) calanoid copepods. In a-d each column represents an environmental dimension. The Y axis corresponds to the 50 environmental categories standardised between 0 and 1, i.e. it represents all the values taken by an environmental variable from the lowest (bottom category) to the highest (top category). Colours denote*

*the percentage of co-occurring species in each category. High co-occurrence values are in red and low values in blue. Categories where no CPR samples are available are shown in white. Panels e-f represent the changes in the degree of niche overlapping for (e) phytoplankton, (f) diatoms, (g) dinoflagellates and (h) calanoid copepods. Niche overlapping has been estimated by means of a species chromatogram (i.e. index D, M&M section). Overlapping was assessed by considering simultaneously an increasing number of environmental dimensions, from 1 to 8 for phytoplankton, diatoms and dinoflagellates, and from 1 to 6 for calanoid copepods. Overlapping among species was based on five thresholds of abundance T=0 (blue lines with diamond), 0.1 (blue lines with dot), 0.25 (red lines with diamond), 0.5 (red lines with dot) and 0.75 (black lines with dot).*

### 2.3.3.2. Habitat diversity

We quantified habitat diversity in the North Atlantic Ocean by means of the Shannon entropy. The analysis was performed in two ways. First, based on different depth levels (0-20 m) for all geographical cells of the North Atlantic, we found two different daily patterns in habitat diversity: the first pattern revealed higher diversity with respect to nutrients, MLD and chlorophyll-a in Spring (plots in Figure 30a, b, e, f and h) and the second pattern revealed high diversity in Summer for temperature, salinity and PAR (plots in Figure 30c, d and g). Second, based on depth levels (0-20m) and days, we identified spatial patterns in habitat diversity. High diversity was found in the northern parts of the North Atlantic Ocean for nutrients, MLD and chlorophyll-a (maps in Figure 30a, b, e, f and h) whereas high habitat diversity was found along continental shelves and in the southern part of the North Atlantic Ocean for temperature and salinity (maps in Figure 30c-d). PAR showed high habitat diversity everywhere in the North Atlantic (Figure 30g). Our results therefore confirm that not only the environment is diverse in space and time but that it is also highly impermanent at the same scales (Figure 30).

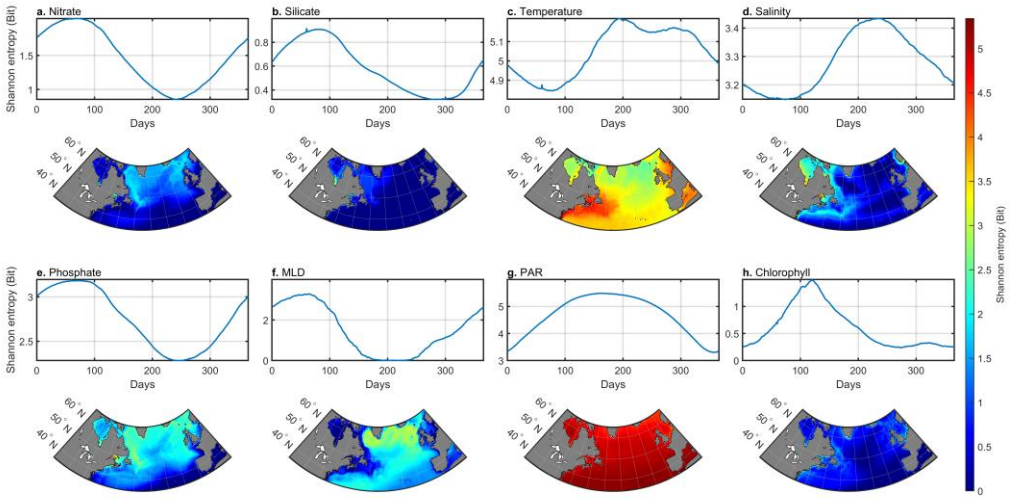



*Figure 30. Habitat diversity in the North Atlantic Ocean. Habitat diversity for (a) nitrate, (b) silicate, (c) temperature, (d) salinity, (e) phosphate, (f) MLD, (g) PAR and (h) chlorophyll-a. Changes in temporal diversity are shown by the plots and changes in spatial diversity by the maps below the plots. Habitat diversity was assessed by means of the Shannon entropy estimated on the annual changes of each environmental variable (see M&M section). High values of Shannon entropy (red colour in the maps) mean that habitats diversity is high and low values (blue colour in the maps) that habitat diversity is small.*

### 2.3.4. Discussion

It is known that diatoms and dinoflagellates have distinct environmental requirements (Irwin et al., 2012). Here, our results are going farther and show that the niches of both phyto and zooplankton species rapidly separate when the number of niche dimensions rises (Figure 29). They also demonstrate that the pelagic environment is not only diverse but also impermanent in space and time (Figure 30), a result that has been shown by many studies since the original paper of Hutchinson (Beaugrand et al., 2019b; Keerthi et al., 2021; Longhurst, 1998). Therefore, the two implicit postulates on which the Hutchinson's paradox was based are incorrect. High niche differentiation leads species to have a unique combination of environmental conditions (i.e. a unique niche) where their performances in terms of growth and reproduction are optimal (Tilman, 1999). Therefore, each species covers a part of the multidimensional environmental space and the segregation along various environmental niche dimensions enables them to coexist in space and time, a mechanism that has been suggested to enable terrestrial plant coexistence (Silvertown, 2004; Tilman, 1999). Furthermore, we have shown that niche separation can occur along a single unique dimension in a multivariable space. For example, two phytoplankton species may still coexist if they share the same niche for a nutrient but utilise a different light wavelengths (Burson et al., 2019). In other words, to simply resolve the plankton paradox we just need to fully understand marine environmental heterogeneity as it applies to the plankton.

Because of the duality between the niche space and the real physical space (i.e. Hutchinson's duality (Colwell and Rangel, 2009)), niche separation (Figure 29) can be related with the diversity of habitats we have observed (Figure 30). Environmental variability is an important mechanism known to promote species coexistence because of a mechanism known as the "storage effect" (Adler et al., 2006; Descamps-Julien and Gonzalez, 2005). This mechanism has three requirements: (i) the species of a given community have to develop strategies to survive during unfavourable environmental conditions, (ii) these species have distinct responses to environmental variability (i.e. distinct niches) and (iii) this variability affects competition, meaning that the niche-environment interaction modulates competition (i.e. a species is more resistant to competition when its environment is more in adequation with its niche and inversely)(Beaugrand and Kirby, 2016; Chesson, 2000).

Expectedly, niche separation is maximal at the optimal part of a species' niche (Figure 29e-h). Therefore, this separation alleviates competition and prevents competitive exclusion. Niche separation takes place through evolution via an alteration in life history strategies. Plankton have developed various strategies, like the development of pigments, cysts, spores, resting eggs or lipidic bags, that enable them to survive during adverse environmental conditions (Jónasdóttir et al., 2015; McMinn and Martin, 2013; McQuoid and Hobson, 1996; Miller, 2004; Stomp et al., 2004). Some strategies (e.g. diatom morphological traits) even explain phenology and annual plankton succession (Breton et al., 2021; Kléparski et al., 2022b). Our study therefore shows that the niche theory, and its corollary the principle of competitive exclusion, applies for the plankton as for other groups, be they terrestrial or marine.

## 2.3.5. Supplementary Files

**Supplementary Text.**

**Text S1.**

The Continuous Plankton Recorder (CPR) survey is a long-term plankton-monitoring programme hold by the Marine Biological association of Plymouth. The CPR is the longest and most extensive program of that kind in the world, started in 1931, with minor modification to the mechanism since 1948. After a break during World War 2, zooplankton sampling has not been changed since 1948, and phytoplankton since 1958 (Warner and Hays, 1994). The machine is a high-speed plankton recorder, towed behind voluntary merchant ships, call "ships of opportunity", operating at a depth of approximately 7-10 m, filtering phytoplankton and zooplankton (Hays and Warner, 1993; Warner and Hays, 1994).

Water enters through the front opening of the CPR by a 1.27 $cm^2$ square entrance aperture, and organisms are filtered by a 270 µm mesh silk band, after passing in a tunnel which expand to cross-sectional dimension of 5 cm x 10 cm; this diminishes water pressure on the organisms and minimizes their damage. The mesh silk band is recovered by a second silk band, call the covering silk, as a "sandwich"; sampled organisms are therefore better protected. The whole is finally stored in a tank containing 4% formaldehyde. The silk moves through a propeller action, which turns because of the CPR movement through the water. Silk movement is proportional to ship's speed (Reid et al., 2003). Back to the laboratory, the band is unloaded and species are identified. Band is cut into sections of 10 nautical miles, corresponding to ~3 $m^3$ of seawater filtered (Jonas et al., 2004). Further information is provided by the ship crew, e.g. navigational data from the tow such as position of the boat, time of CPR deployment and retrieve, position and time of change of ship's speed. This information allows the identification of the location of each CPR sample. Bands are randomly allocated to analysts in a way that they cannot receive consecutive bands from the same tow, in order to minimize analytical bias originating from the experience or the subjectivity of the analyst. The analysis is divided into four steps.

First, green coloration is estimated for each CPR sample. This is made by comparison with a four-level standard colour chart: 0 (no colour), 1 (very pale green), 2 (pale green), 3 (green). This estimation is called the Phytoplankton Colour Index (PCI) (Batten et al., 2003) and is an estimation of phytoplankton biomass. But as our analysis focussed on species abundance, PCI was not used in this work.

The three further steps have a distinct microscope procedure to identify each cell of each sample at the highest taxonomic level; up to 500 taxa are identified by the three corresponding procedures (Batten et al., 2003; Warner and Hays, 1994). Dinoflagellates and copepods are often well preserved, so there are identified at a species level whereas other groups such as chaetognaths and larvaceans are identified to higher taxonomic levels because there are not always well preserved and their identification is more difficult under light microscope.

Step 2 identifies and counts phytoplankton cells. On each CPR sample, a cross section of twenty points centred on one silk's mesh is defined and there is a semi-quantitative record and an identification of each phytoplankton cell in these twenty points under high magnification (x450) corresponding to a 295 µm field of view.

Step 3 identifies small zooplankton species under x54 magnification corresponding to a 2.06 mm field of view. It is a staggered microscope "traverse" examination of the covering and filtering silks (because some organisms can be transferred between the two silks), with the count and identification of each zooplankton cells ≤2 mm in the field of view of 2.06 mm.

Step 4 identifies all zooplankton species greater than ~2 mm. Large organisms are removed from the silk to be viewed separately. In contrast to Steps 2 and 3, all individuals on the silks (filtering and covering silks) are counted. It is "a compromise between precision of enumeration and speed of processing" (Batten et al., 2003). Here we used the three types of abundance data: phytoplankton, zooplankton traverse and zooplankton eye-count, between 1998 and 2018.

It is important to note that the CPR survey is the only scientific monitoring programme of that kind in the world, with no equivalent existing program. It covers an important time scale from 1948 to present, still active, with a large spatial scale covering the whole North Atlantic Ocean and its adjacent seas such as the Channel, the Celtic Sea and the North Sea. Now it is "*the most extensive long-term survey of marine organisms in the world*" (Reid et al., 2003).

**Text S2.**

Mass concentration of chlorophyll-a in sea water ($mg.m^{-3}$), nitrate, phosphate and silicate concentration ($mmol.m^{-3}$) data originated from the Global Ocean Biogeochemistry Hindcast (GLOBAL_REANALYSIS_BIO_001_029) and were provided by the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu). Daily means were provided on a 0.25° resolution grid and along 75 depth levels from 0 to 5500 m. The dataset covers the time period from 1993 to present and is regularly updated.

Sea water potential temperature (°C), salinity (no unit) and Mixed Layer Depth (MLD, m) data originated from the Global Ocean Ensemble Physics Reanalysis (GLOBAL_REANALYSIS_PHY_001_031) and were provided by the Copernicus Marine Environment Monitoring Service (CMEMS) (http://marine.copernicus.eu). Daily means were provided on a 0.25° resolution grid along 75 depth levels from 0 to 5500m. The dataset covers the time period from 1993 to present and is regularly updated.

Photosynthetically Active Radiations clear sky in surface (PAR, in $J.m^{-2}$) originated from the ERA interim dataset provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) (https://www.ecmwf.int/). Hourly means were provided on a 0.25° resolution grid, covering the time period 1998-2018. Daily PAR was estimated by summing all the values corresponding to a given day and were subsequently converted into $E.m^{-2}.day^{-1}$.

$K_d$ (PAR) data originated from the Glob Colour project (http://www.globcolour.info/). The product merges together all the daily data from satellites (MODIS, SeaWIFS, VIIRS) available for each parameter, from September 1997 to present and on a 4km resolution spatial grid. As the data can be very holey because of cloud cover and sun glint effect during the winter season, missing $K_d$ (PAR) values were first spatio-temporally interpolated and the remaining missing data (i.e. the one above 45°N in winter) where interpolated with chlorophyll-a data according to the relationships presented in (Morel et al., 2007).

PAR in depth was finally estimated from the Beer-Lambert law (Swinehart, 1962): $I_Z = I_0 e^{-K_d \times Z}$, with $I_0$ the PAR in surface and $Z$ the depth (from 0 to 100 m).

Bathymetry (m) came from GEBCO Bathymetric Compilation Group 2019 (The GEBCO_2019 Grid - a continuous terrain model of the global oceans and land). Data are provided by the British Oceanographic Data Centre, National Oceanography Centre, NERC, UK. (https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/836f016a-33be-6ddc-e053-6c86abc0788e/). To work on the same spatial grid $K_d$ (PAR) and bathymetry were interpolated on a 0.25° latitude x 0.25° longitude grid.

**Supplementary Figures.**

**Figure S1. Species chromatograms of nine diatoms.** Multidimensional niches of (**a**) *Paralia sulcata*, (**b**) *Skeletonema costatum*, (**c**) *Dactyliosolen antarcticus*, (**d**) *Rhizosolenia styliformis*, (**e**) *Rhizosolenia hebetata semispina*, (**f**) *Odontella sinensis*, (**g**) *Thalassiothrix longissima*, (**h**) *Thalassionema nitzschioides* and (**i**) *Bellerochea malleus*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

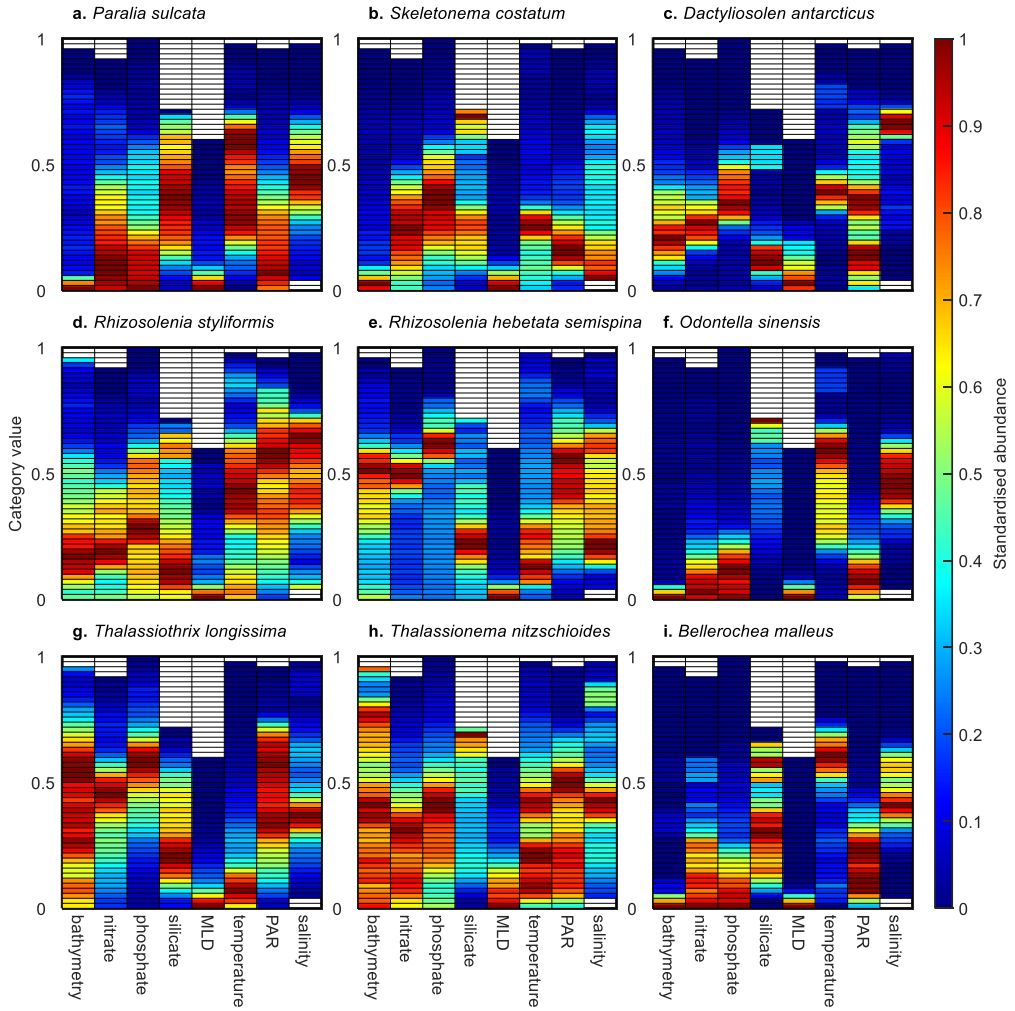


**Figure S2. Species chromatograms of nine diatoms.** Multidimensional niches of (**a**) *Biddulphia alternans*, (**b**) *Odontella aurita*, (**c**) *Odontella granulata*, (**d**) *Odontella regia*, (**e**) *Odontella rhombus*, (**f**) *Coscinodiscus concinnus*, (**g**) *Detonula confervacea*, (**h**) *Ditylum brightwellii* and (**i**) *Eucampia zodiacus*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

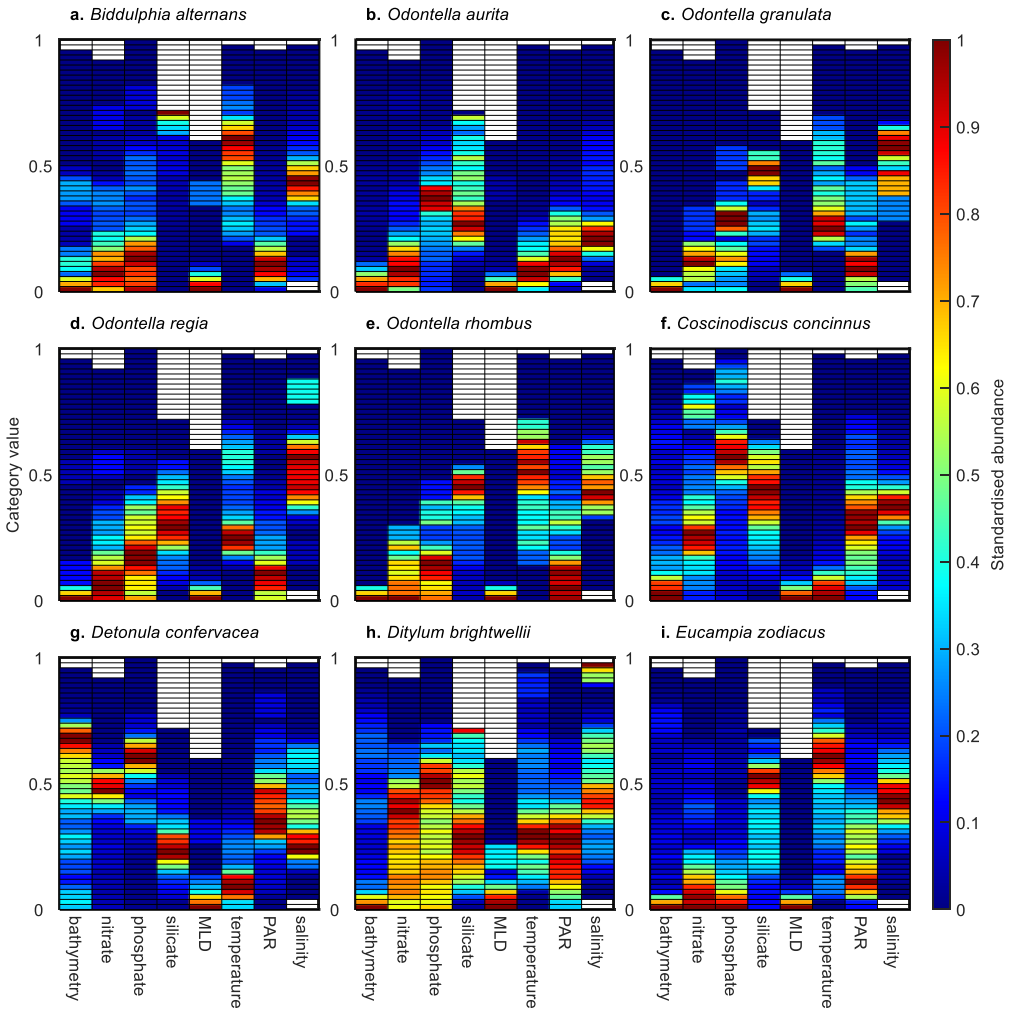


**Figure S3. Species chromatograms of nine diatoms.** Multidimensional niches of (**a**) *Guinardia flaccida*, (**b**) *Leptocylindrus danicus*, (**c**) *Cylindrotheca closterium*, (**d**) *Rhaphoneis amphiceros*, (**e**) *Rhizosolenia acuminata*, (**f**) *Rhizosolenia bergonii*, (**g**) *Rhizosolenia setigera*, (**h**) *Hexasterias problematica* and (**i**) *Coscinodiscus wailesii*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

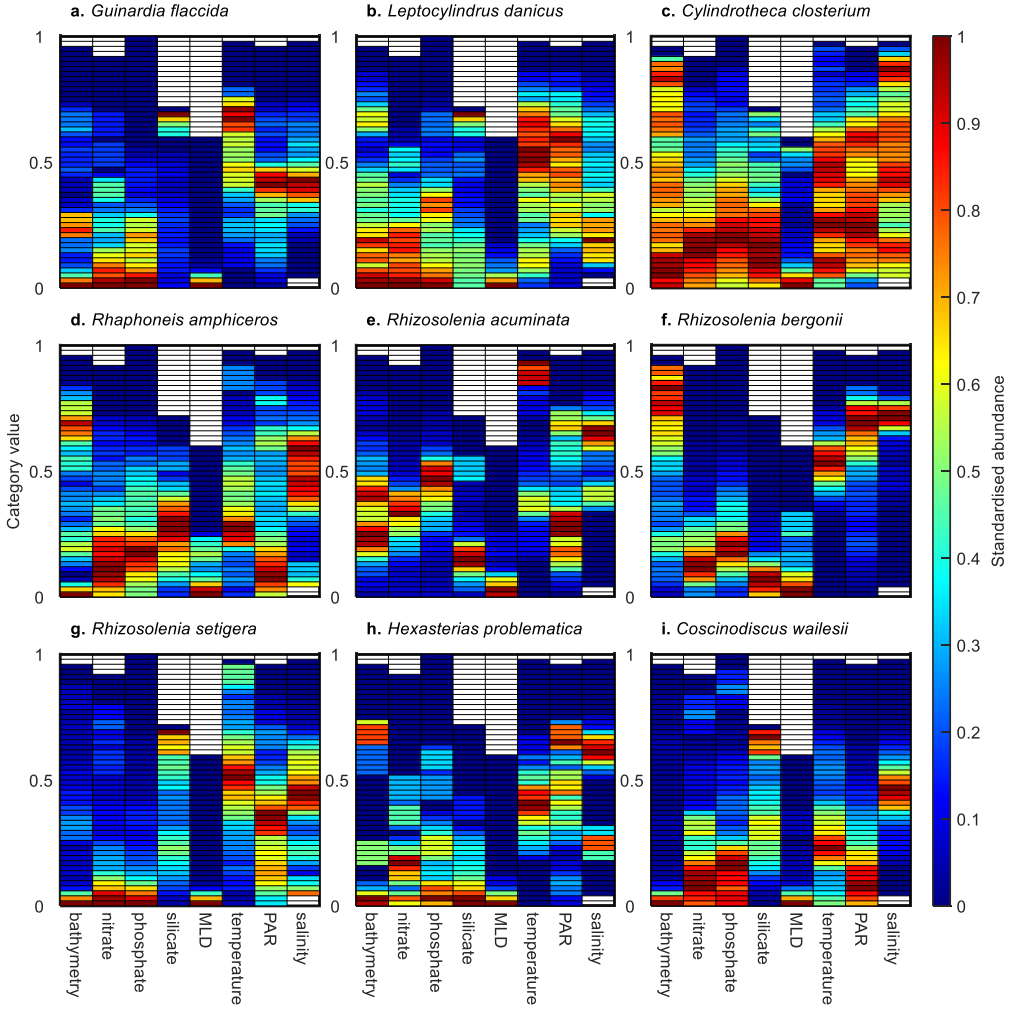


**Figure S4. Species chromatograms of nine diatoms.** Multidimensional niches of (**a**) *Nitzchia longissima*, (**b**) *Proboscia alata*, (**c**) *Leptocylindrus mediterraneus*, (**d**) *Proboscia inermis*, (**e**) *Asterionellopsis glacialis*, (**f**) *Ephemera planamembranacea*, (**g**) *Pseudo-nitzschia delicatissima*, (**h**) *Pseudo-nitzschia seriata* and (**i**) *Podosira stelligera*. Niches are displayed by means of species chromatograms. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

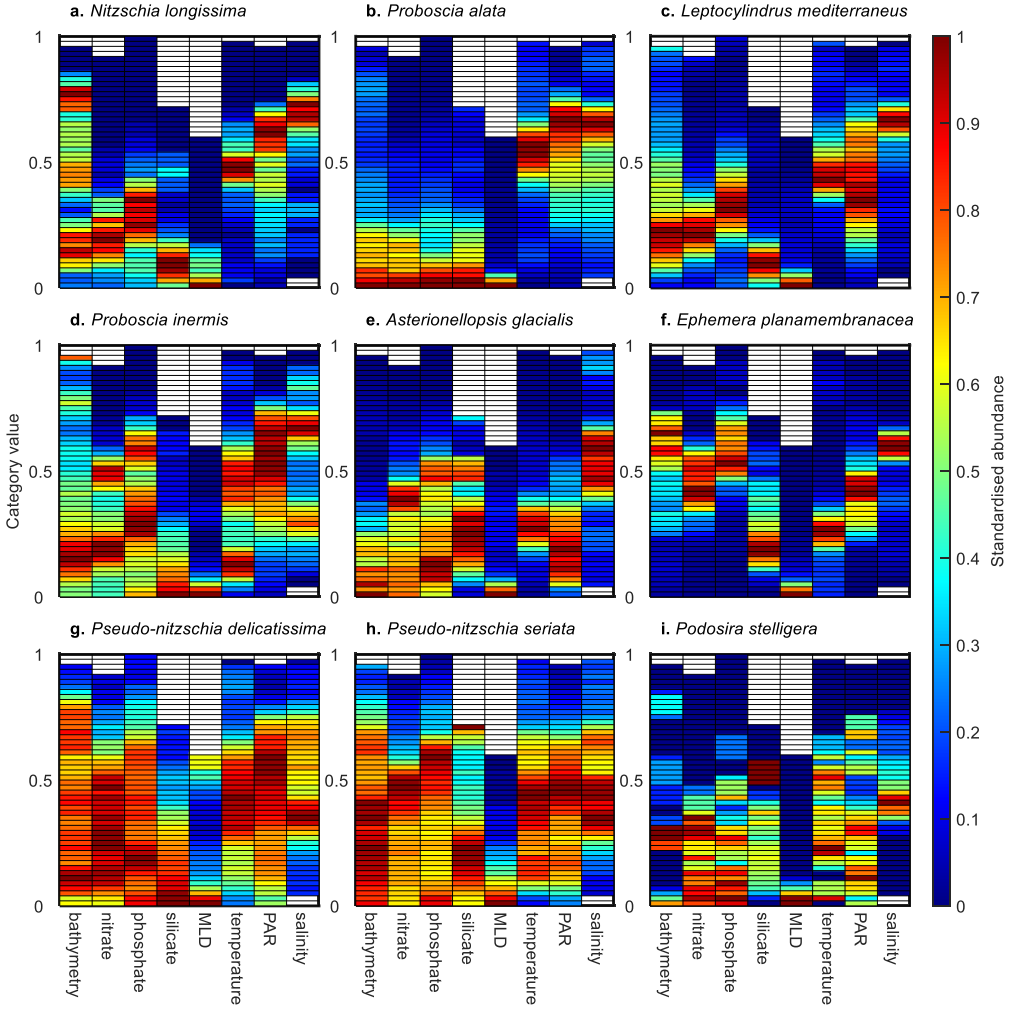


**Figure S5. Species chromatograms of nine diatoms.** Multidimensional niches of (**a**) *Pseudosolenia calcar-avis*, (**b**) *Guinardia delicatula*, (**c**) *Dactyliosolen fragilissimus*, (**d**) *Guinardia striata*, (**e**) *Lauderia annulata*, (**f**) *Bacillaria paxillifera*, (**g**) *Corethron hystrix*, (**h**) *Proboscia indica* and (**i**) *Rhizosolenia imbricata*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

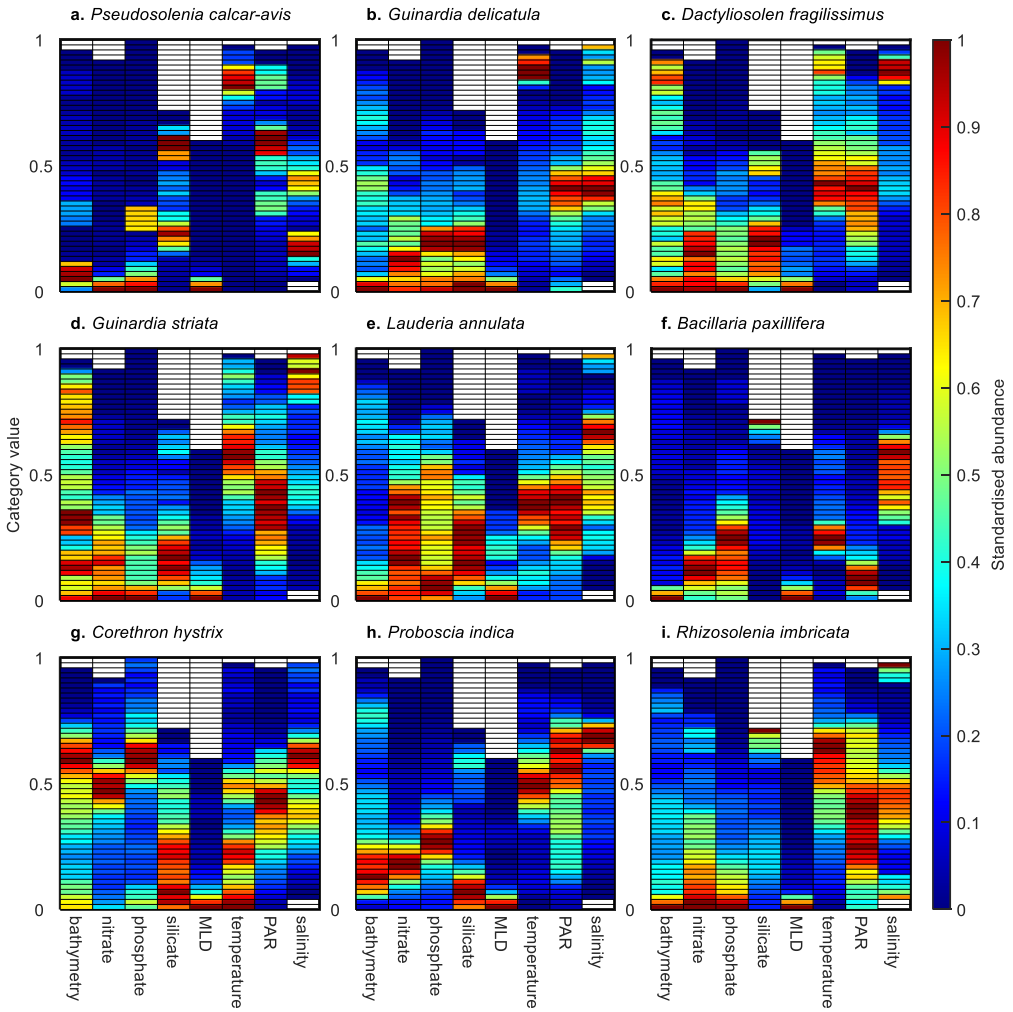


**Figure S6. Species chromatograms of nine dinoflagellates.** Multidimensional niches of (**a**) *Ceratium fusus*, (**b**) *Ceratium furca*, (**c**) *Ceratium lineatum*, (**d**) *Ceratium tripos*, (**e**) *Ceratium macroceros*, (**f**) *Ceratium horridum*, (**g**) *Ceratium longipes*, (**h**) *Ceratium arcticum* and (**i**) *Ceratium arietinum*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

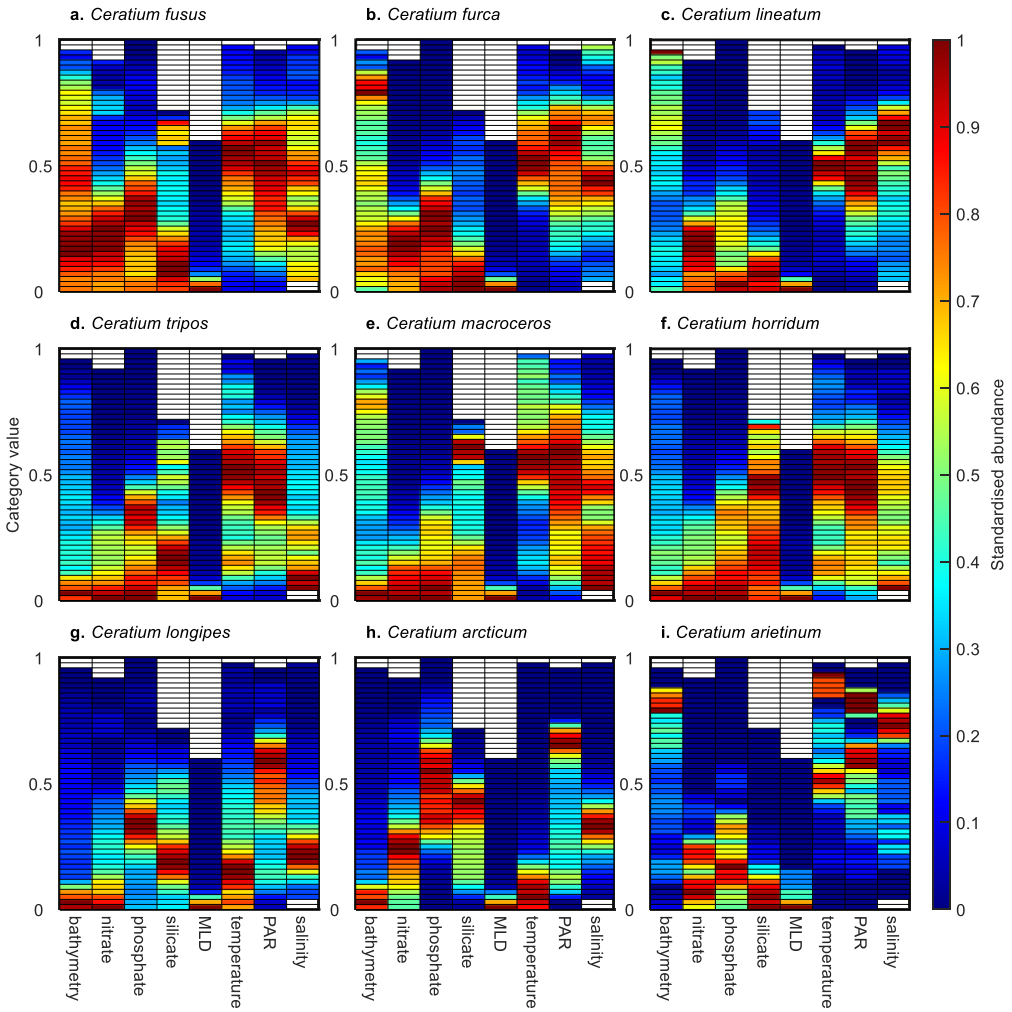


**Figure S7. Species chromatograms of nine dinoflagellates.** Multidimensional niches of (**a**) *Ceratium azoricum*, (**b**) *Ceratium bucephalum*, (**c**) *Ceratium candelabrum*, (**d**) *Ceratium carriense*, (**e**) *Ceratium compressum*, (**f**) *Ceratium declinatum*, (**g**) *Ceratium extensum*, (**h**) *Ceratium gibberum* and (**i**) *Ceratium hexacanthum*. In **a**-**i**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

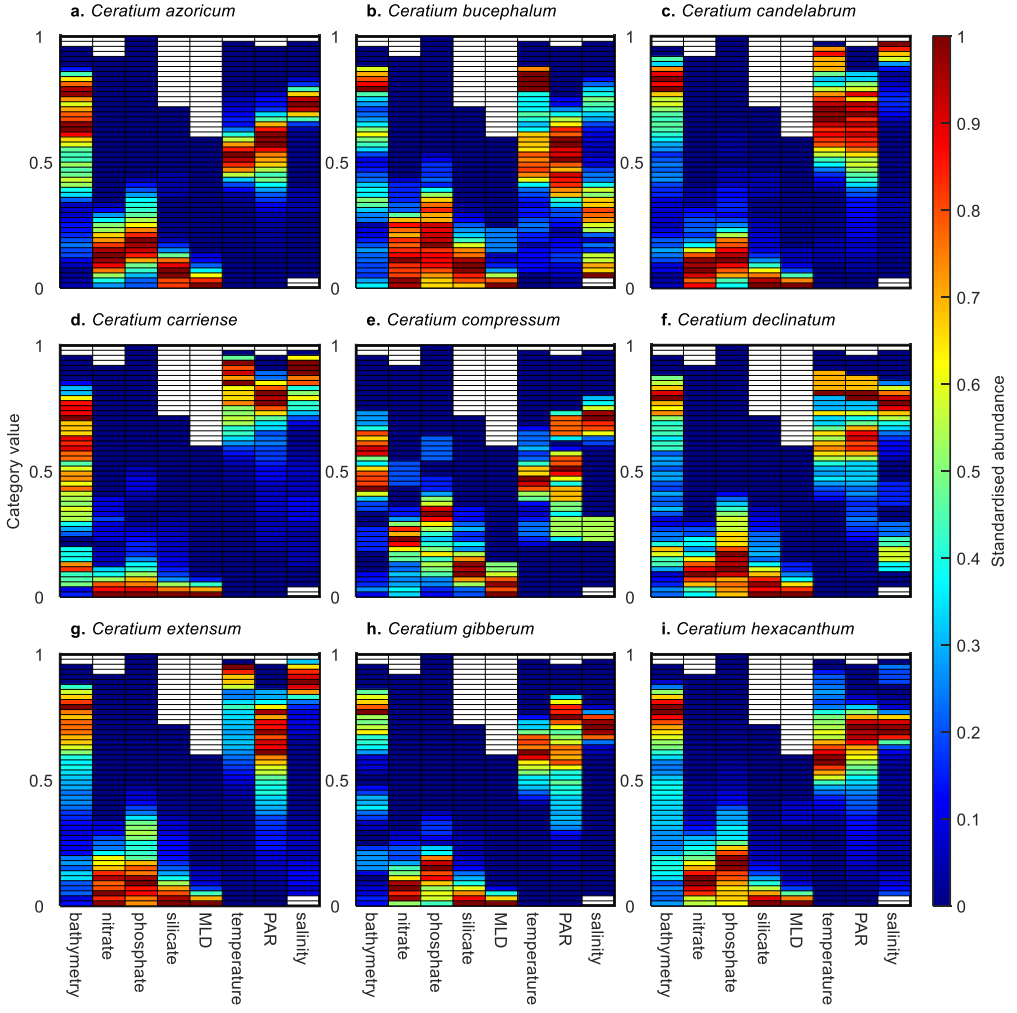


**Figure S8. Species chromatograms of eight dinoflagellates.** Multidimensional niches of (**a**) *Ceratium lamellicorne*, (**b**) *Ceratium massiliense*, (**c**) *Ceratium minutum*, (**d**) *Ceratium pentagonum*, (**e**) *Ceratium teres*, (**f**) *Ceratium trichoceros*, (**g**) *Ceratium pentasterias* and (**h**) *Ceratium polyedricum*. In **a**-**h**, each column corresponds to the dimension of the niche (i.e. bathymetry, nitrate, phosphate, silicate, MLD, temperature, PAR and salinity). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

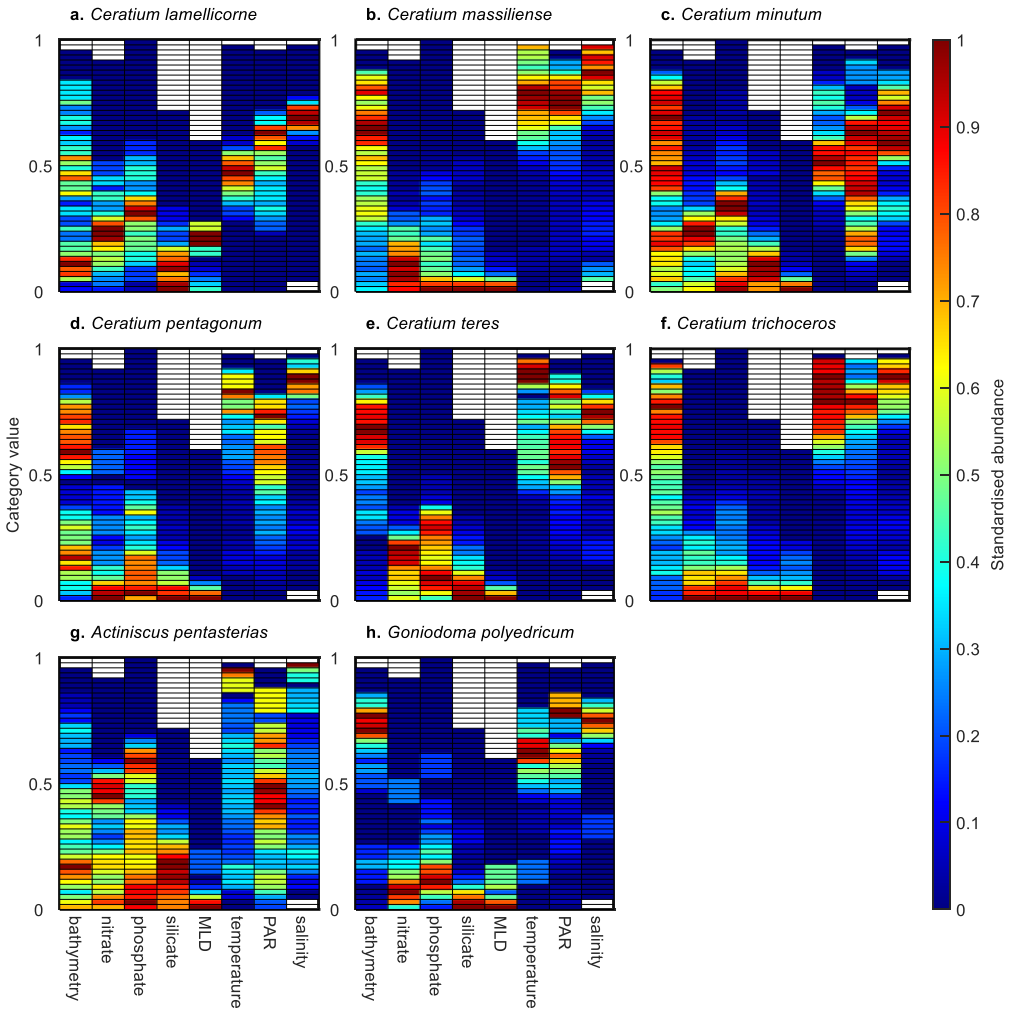


**Figure S9. Species chromatograms of twelve copepods.** Multidimensional niches of (**a**) *Calanus finmarchicus*, (**b**) *Calanus helgolandicus*, (**c**) *Calanus glacialis*, (**d**) *Calanus hyperboreus*, (**e**) *Neocalanus gracilis*, (**f**) *Nannocalanus minor*, (**g**) *Calanoides carinatus*, (**h**) *Rhincalanus nasutus*, (**i**) *Euchirella rostrata*, (**j**) *Euchaeta acuta*, (**k**) *Metridia lucens* and (**l**) *Metridia longa*. In **a**-**l**, each column corresponds to the dimension of the niche (i.e. bathymetry, MLD, temperature, PAR, salinity and chlorophyll-a concentration). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

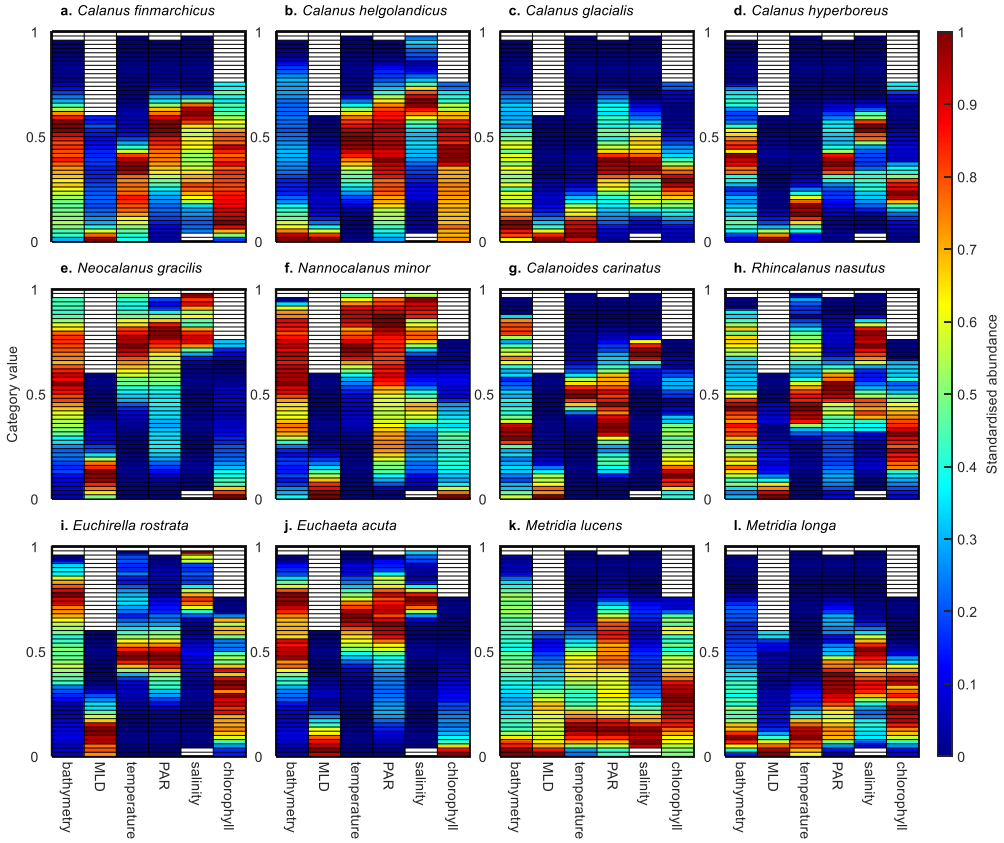


**Figure S10. Species chromatograms of twelve copepods.** Multidimensional niches of (**a**) *Pleuromamma robusta*, (**b**) *Pleuromamma abdominalis*, (**c**) *Pleuromamma borealis*, (**d**) *Pleuromamma gracilis*, (**e**) *Candacia armata*, (**f**) *Labidocera wollastoni*, (**g**) *Aetideus armatus*, (**h**) *Anomalocera patersoni*, (**i**) *Candacia bipinnata*, (**j**) *Candacia ethiopica*, (**k**) *Candacia pachydactyla* and (**l**) *Centropages bradyi*. In **a**-**l**, each column corresponds to the dimension of the niche (i.e. bathymetry, MLD, temperature, PAR, salinity and chlorophyll-a concentration). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

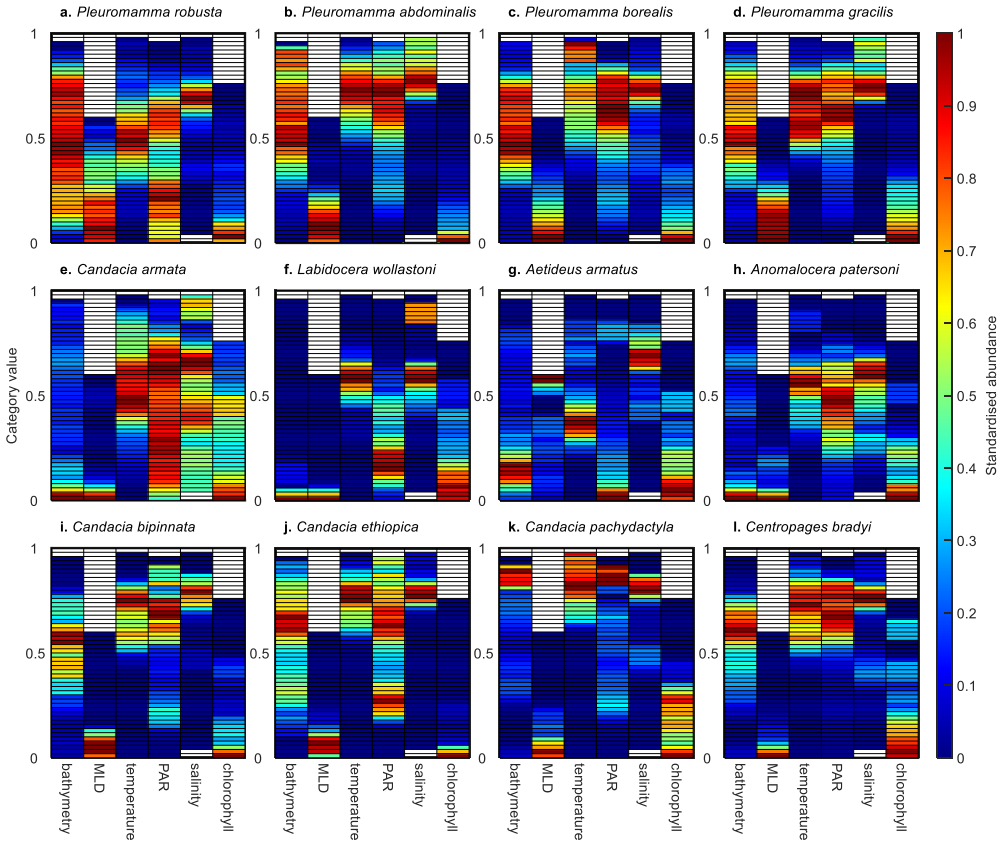


**Figure S11. Species chromatograms of twelve copepods.** Multidimensional niches of (**a**) *Centropages violaceus*, (**b**) *Euchaeta marina*, (**c**) *Heterorhabdus norvegicus*, (**d**) *Heterorhabdus papilliger*, (**e**) *Pleuromamma piseki*, (**f**) *Pleuromamma xiphias*, (**g**) *Rhincalanus cornutus*, (**h**) *Scolecithrix danae*, (**i**) *Undeuchaeta plumosa*, (**j**) *Undinula vulgaris*, (**k**) *Paracandacia simplex* and (**l**) *Mesocalanus tenuicornis*. In **a**-**l**, each column corresponds to the dimension of the niche (i.e. bathymetry, MLD, temperature, PAR, salinity and chlorophyll-a concentration). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.

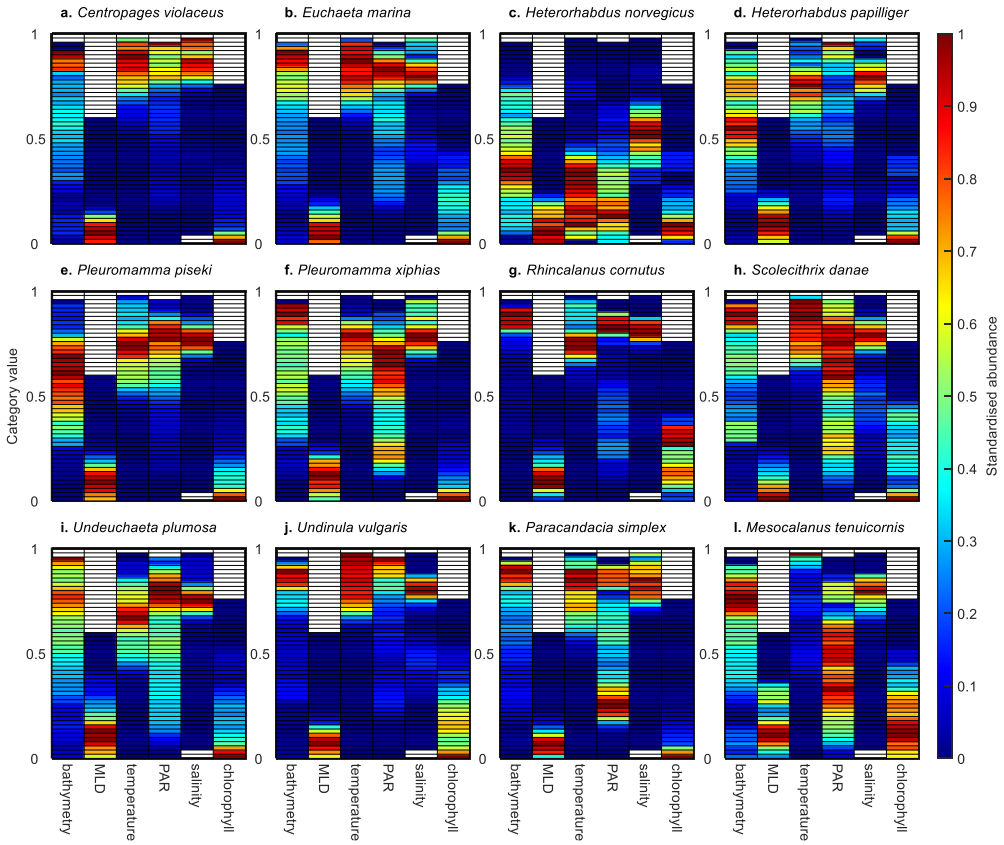
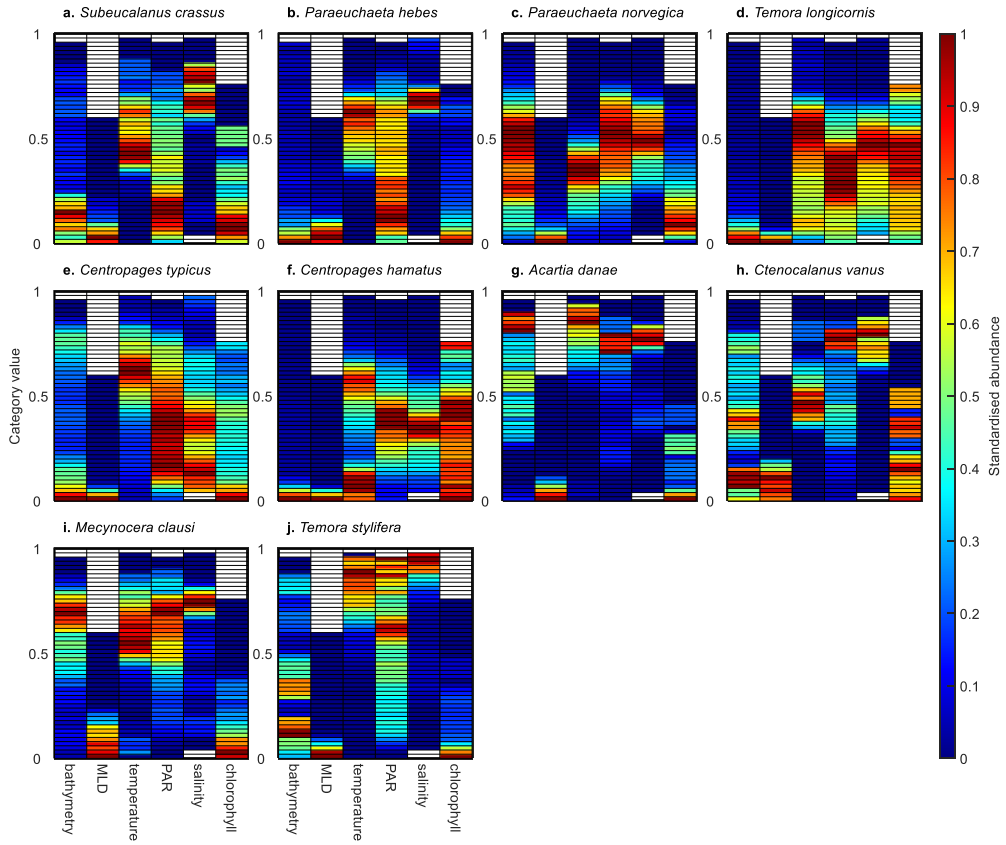


**Figure S12. Species chromatograms of ten copepods.** Multidimensional niches of (**a**) *Subeucalanus crassus*, (**b**) *Paraeuchaeta hebes*, (**c**) *Paraeuchaeta norvegica*, (**d**) *Temora longicornis*, (**e**) *Centropages typicus*, (**f**) *Centropages hamatus*, (**g**) *Acartia danae*, (**h**) *Ctenocalanus vanus*, (**i**) *Mecynocera clausi* and (**j**)*Temora stylifera*. In **a**-**j**, each column corresponds to the dimension of the niche (i.e. bathymetry, MLD, temperature, PAR, salinity and chlorophyll-a concentration). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5 % of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue. Categories where no CPR samples are available are shown in white.



# ***Chapter 3.*** *Consequences of climate change on phytoplankton phenology and abundance*

*Summary*

In this last chapter, we examine the predicted and observed effects of climate change on phytoplankton abundance and phenology. First, by means of an approach based on METAL (see Chapter 1), we investigate whether the phenology of diatoms (including both oblate and prolate) and dinoflagellates is altered by ocean warming (Kléparski et al., 2022a). To do so, we use projections from six Earth System Models and two warming scenarios. We designed six indices to characterise the long-term phenological changes: (1) Maximum Annual Abundance (MAA), (2) the day where MAA is reached, (3) initiation and (4) termination of the Seasonal Reproductive Period (SRP), i.e. the first and the last days where the abundance is ≥50% of MAA, (5) seasonal duration, i.e. the number of consecutive days where abundance is ≥ 50% of MAA, and (6) Integrated Mean Annual Abundance. Our results show that the phenology of the three groups is altered in three oceanic systems, i.e. the North Sea, the North-East Atlantic and the Labrador Sea, and that changes are more important under the high-emission scenario. The phenology of oblates is predicted to shrink and their abundance to decline whereas the phenology of prolates and dinoflagellates is expected to expand and their abundance to rise. These projections agree with theoretical expectations and bring new evidence that may revisit our understanding and anticipation of the consequences of phytoplankton shifts on the biological carbon pump. They also suggest that future ESMs should increase their level of complexity by distinguishing more diatom groups, i.e. oblates and prolates, and also dinoflagellates.

In a second study, we examine the observed impact of climate change on diatoms in the eastern part of the North Atlantic Ocean (Edwards et al., 2022). In this article led by Pr. Martin Edwards, our results show that diatom abundance is positively correlated with temperature and wind intensity in the northern North Sea and the Islandic basin whereas it is negatively correlated in the Celtic Sea and the Bay of Biscay. Therefore, diatom population is rising in the northern part of the studied area while it declines in the southern parts. Although this study does not distinguish between both types of diatoms, results agree with the phenological shifts we predicted for oblates, i.e. that the decline of their abundance is more pronounced in the North Sea than in the North-East Atlantic.

# 3.1. Climate change and phytoplankton phenology in the North Atlantic

Loïck Kléparski, Grégory Beaugrand, Martin Edwards and Clare Ostle

*Under review in Global Change Biology. Supplementary Files are displayed in section 3.1.5.*

*Abstract*

Significant phenological shifts induced by climate change are predicted within the phytoplanktonic community (Asch et al., 2019; Henson et al., 2018, 2013). However, predictions from current Earth System Models (ESMs) understandably rely on simplified community responses that do not consider evolutionary strategies manifested as various life phenotypes and trait groups. Here, we use a species-based modelling approach, combined with large-scale plankton observations, to investigate past, contemporary and future phenological shifts in diatoms (grouped by their morphological traits) and dinoflagellates in three key areas of the North Atlantic Ocean (North Sea, North-East Atlantic and Labrador Sea) from 1850-2100. Our study reveals that the groupings exhibit coherent and different shifts in phenology and abundance throughout the North Atlantic Ocean. The phenology of large flattened diatoms is predicted to shrink and their abundance to decline, whereas the phenology of slow-sinking elongated diatoms and of dinoflagellates is expected to expand and their abundance to rise, which may alter carbon export in this important sink region. Our results show new understanding of the possible consequences of phytoplankton shifts on the biological carbon pump, suggesting that future ESMs should increase their level of complexity to include ecological strategies adapted by various phytoplankton groups.

*Keywords*

Diatoms, oblates, prolates, dinoflagellates, phenology, annual phytoplankton succession, climate change

## 3.1.1. Introduction

Current Earth System Models (ESMs) are predicting a decline in phytoplankton abundance (Bopp et al., 2013; Dutkiewicz et al., 2013; IPCC, 2019; Kwiatkowski et al., 2020) and anticipate major phenological shifts (Asch et al., 2019; Henson et al., 2018, 2013; Yamaguchi et al., 2022) that may lead to trophic desynchronization and community reorganisation (Edwards and Richardson, 2004; IPCC, 2019; Winder and Schindler, 2004; Yamaguchi et al., 2022). Phenological shifts in phytoplankton are expected because of ocean warming and its strengthening influence on water column stability, which is expected to diminish nutrients supply in the euphotic zone, favouring small phytoplankton (e.g. nanoplankton and picoplankton) at the expense of diatoms (Bopp et al., 2005; Marinov et al., 2010), a major group responsible for 40% of total marine primary production (Field et al., 1998; Tréguer et al., 2018).

Although shifts in phytoplankton phenology are now widely observed among marine and freshwater ecosystems (Friedland et al., 2018; Poloczanska et al., 2013; Thackeray et al., 2016), the examination of diatom seasonality in some regions of the North Atlantic suggests a relative stability (Chivers et al., 2020; Edwards and Richardson, 2004). This apparent paradox may originate from the range of strategies that have been developed by diatoms and enable them to occur in diverse environments (Kemp and Villareal, 2018, 2013). Among those strategies, diatom morphological traits might have an important influence on species phenology (Kléparski et al., 2022b). Indeed, it has recently been shown that diatom cell shape has evolved as a key adaptation that confers to a species a specific phenology, oblate (i.e. flattened) diatoms being dominant in well-mixed, nutrients-rich, low stratified waters whereas prolate (i.e. elongated) diatoms dominate in the stratified low-nutrients waters (Kléparski et al., 2022b).

Most projected phenological shifts (Asch et al., 2019; Henson et al., 2018, 2013; Yamaguchi et al., 2022), which rely on the analyses of variables such as chlorophyll concentration or net primary production, originating from ESMs that only consider a few phytoplanktonic types (only one for diatoms and none for dinoflagellates), cannot anticipate subtle changes in the phytoplankton community (Kemp and Villareal, 2018; Séférian et al., 2020; Tréguer et al., 2018). Here we applied a species-based modelling approach, combined with large-scale plankton observations, to investigate past, contemporary and future long-term changes (1850-2100) in the phenology of oblates, prolates and dinoflagellates by using six ESMs and two scenarios (a medium and a high emission scenario, i.e. SSP245 and SSP585).

## 3.1.2. Methods

### 3.1.2.1. Biological data

Phytoplankton data came from the Continuous Plankton Recorder (CPR) survey. The CPR is a long-term plankton monitoring programme that has collected plankton on a monthly basis in the North Atlantic Ocean and its adjacent seas since 1946. The sampling machine is a high-speed plankton recorder towed behind voluntary merchant ships (called “ships of opportunity”) at a depth of approximately

7m (Beaugrand et al., 2003a; Reid et al., 2003). As phytoplankton sampling remained unchanged since 1958 (Warner and Hays, 1994) and historical climatic simulations ended in 2014 (see section 2 below), we used the abundance data collected for the diatoms and the dinoflagellates between 1958 and 2014. Each value of abundance corresponds to a number of cells per CPR sample, which corresponds to ~3 $m^3$ of seawater filtered (Jonas et al., 2004). Species that were first identified after 1958 were discarded from the analyses. In order to account for the climatic variability that was observed in the North Atlantic sector during the period (Beaugrand et al., 2019a; Edwards et al., 2002), we calculated a daily mean abundance climatology (based on 365 days) for each species for three 18-year periods: 1958-1976, 1977-1995 and 1996-2014 (a total of 365 days x 3 = 1095 days). Climatologies were estimated in three distinct oceanic regions of the North Atlantic: (i) the North Sea (51°N, 60°N, -3°E, 9°E; Figure 32a), (ii) the North-East Atlantic (61°N, 63.5°N, -21°E, -8°E; Figure 33a) and (iii) the Labrador Sea (48°N, 60°N, -55°E, -40°E; Figure 34a). Climatologies were calculated in a given region for the species/taxa that were present in more than 100 CPR samples. A total of 44 species/taxa were therefore selected. Climatologies (for the three periods) were smoothed in each region by means of a double 6-order simple moving average and subsequently standardised between 0 and 1, applying the approach of Caracciolo and colleagues (Caracciolo et al., 2021). Standardisation was performed as follows:

$$A^*_{(i,j)} = \frac{A_{ij}}{max(A_j)} \qquad (1)$$

With $A^*_{(i,j)}$ the standardised abundance of species $j$ on day $i$, $A_{ij}$ the abundance of species $j$ on day $i$ and $max(A_j)$ the maximal abundance of species $j$ in the three regions (i.e. North Sea, North-East Atlantic and Labrador Sea) between 1958 and 2014. We then estimated the 90$^{th}$ percentile of the abundance (P90) of each species using the non-standardised daily climatologies in the three oceanic regions.

Species were divided into three groups: oblate and prolate diatoms and dinoflagellates, which are known to have distinct environmental requirements (Irwin et al., 2012; Kléparski et al., 2022b). The oblate group gathered diatoms (species or taxa) that have a mean cell diameter greater than their mean cell height and the prolate group the diatoms (species or taxa) that have a mean cell diameter smaller than their mean cell height. Information on diatom cell shapes were retrieved from Kléparski *et. al* 2022 (Kléparski et al., 2022b) (Figure 17). Mean monthly abundances between 1958 and 2014 were estimated for each group and in each oceanic region (Figure S1a-i). Data were subsequently smoothed by means of a first-order moving average and then standardised between 0 and 1 by using the 90$^{th}$ percentile of the monthly abundance of each taxonomic group ($P90_m$):

$$N^*_m = \frac{N_m}{P90_m} \qquad (2)$$

With $N_m^*$ the standardised monthly abundance of a taxonomic group for month $m$, $N_m$ the monthly abundance for month $m$ and $P90_m$ the 90$^{th}$ percentile of the monthly abundance of the taxonomic group. In rare cases, standardised monthly abundance above 1 were fixed to 1. We chose to use the 90$^{th}$ percentile instead of the maximum abundance value because the latter would have been too sensitive to outliers that may originate from multiple causes (e.g. exceptional abundance, patchiness or abundance misestimated from only 3m$^3$ of seawater filtered).

### 3.1.2.2. CMIP6 climate simulations 1850-2100

Species responses to environmental variability was modelled by using three environmental variables known to influence phytoplankton phenology at high latitudes (Beaugrand and Kirby, 2018; Boyce et al., 2017; Caracciolo et al., 2021; Lewandowska and Sommer, 2010; Miller, 2004): i.e. Sea Surface Temperature (SST; °C), Surface Downwelling Shortwave Radiation (SDSR; W.m$^{-2}$) and dissolved nitrate concentration (mol.m$^{-3}$). Climate projections for the three variables originated from the Coupled Model Intercomparison Project Phase 6 (CMIP6)(Eyring et al., 2016) and were provided by the Earth System Grid Federation (ESGF). We used the Shared Socioeconomic Pathways (SSP) 2-4.5 and 5-8.5 corresponding respectively to a low, a medium and a high radiative forcing by 2100 (2.5 W.m$^{-2}$ and 8.5 W.m$^{-2}$)(O'Neill et al., 2017, 2016). The simulations of six different ESMs (i.e. CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-LR, NorESM2-LM and UKESM1-0-LL) covering the time period 1850-2014 (historical simulation) and 2015-2100 (future projections for the two SSP scenarios) were used. All the data were interpolated into daily on a 0.5° by 0.5° regular grid. Key references (i.e. DOI and dataset version) are provided in Text S1. Long-term forecasted changes of the three variables, for the six models, the two scenarios and the three regions are displayed in Figures S6-8.

### 3.1.2.3. The MacroEcological Theory on the Arrangement of Life (METAL)

A framework from the METAL theory was used to investigate past, contemporary and future phenological changes of the three taxonomic groups. METAL is a theory that attempts to explain how biodiversity, from the individual to the community level, is organised in space and time and how it responds to environmental changes (Beaugrand, 2015a). Based on the concept of the ecological niche *sensu* Hutchinson (i.e. the set of conditions enabling a species to growth and reproduce)(Hutchinson, 1957), one fundamental assumption of METAL is that the niche-environment interaction is a fundamental interaction that enables one to unify and predict a large number of ecological and biogeographical phenomena, as well as the effect of climate-induced environmental changes on individuals, species and communities (Beaugrand, 2015a; Beaugrand and Kirby, 2018, 2016). Recently, it has been shown that a framework originating from METAL could also explain phytoplankton phenology and the resulting annual plankton succession in the North Sea (Caracciolo et al., 2021). By creating a local pseudo-community, composed of fictive species (i.e. pseudo-species), it was possible to reconstruct at a species level

phytoplankton phenology and at a community level the annual plankton succession (Caracciolo et al., 2021). We therefore applied the same framework here to explore how diatom and dinoflagellate phenology might be modified in the coming decades with global climate change.

### 3.1.2.4. Niche characterisation

We first estimated daily climatologies (based on 365 days) of the three environmental variables described above (i.e. SST, RSDS and nitrate concentration) for three time periods (1958-1976, 1977-1995 and 1996-2014 i.e. 365 days x 3 = 1095 days), three oceanic regions (North Sea, North-East Atlantic and Labrador Sea) and based on six ESMs (i.e. CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-LR, NorESM2-LM and UKESM1-0-LL). Then, we generated a set of 1,755,000 pseudo-species (i.e. fictive species) and we calculated their abundance along the daily climatologies by using a three-dimensional Gaussian niche (Caracciolo et al., 2021):

$$B = ce^{-\frac{1}{2}\left[\left(\frac{x_1 - x_{opt}}{t_1}\right)^2 + \ldots + \left(\frac{x_n - x_{optn}}{t_n}\right)^2\right]} \qquad (3)$$

With $B$ the abundance of a pseudo-species along a given environmental gradient $x$, $c$ the maximum abundance of a pseudo-species (here $c$ was fixed to 1), $x_1$ to $x_n$ the environmental gradients, with n=3 (SST, SDSR and nitrate concentration). $x_{opt1}$ to $x_{optn}$ are the ecological niche optima along $x_1$to $x_n$ and $t_1$ to $t_n$ the niche amplitude along $x_1$to $x_n$. Niche optima were defined for SST from 0 to 25°C every 1°C, for RSDS from 0 to 400 W.m$^{-2}$ every 40 W.m$^{-2}$ and for nitrate concentration from 0.001 to 0.05 mol.m$^{-3}$ every 0.005 mol.m$^{-3}$. Niche amplitudes were defined for SST from 1 to 10°C every 2°C and for nitrate concentration from 0.001 to 0.015 mol.m$^{-3}$ every 0.001 mol.m$^{-3}$. Amplitude for RSDS were defined as followed: 1, 3, 5, 8, 10, 15, 20, 50 and 100 W.m$^{-2}$. Each pseudo-species was therefore defined by a unique combination of optima and amplitudes along the three environmental gradient, i.e. by a unique niche, according to the principle of competitive exclusion (Gause, 1934; Hutchinson, 1978).

We compared the modelled and observed (CPR) abundance at a species and a group (oblates, prolates and dinoflagellates) scale. At a species level, we calculated the Mean Absolute Error (MAE) between the daily abundance climatology (based on 365 days x 3 temporal periods = 1095 days) of each species and each of the 1,755,000 pseudo-species we created, for each oceanic region and ESM. Some species exhibited erratic short peaks that were difficult to explain in some region. To prevent our model to be influenced by these events that might be related to misidentification, CPR silk contamination or expatriation, we only summed MAEs from the regions where a given species had a standardised abundance greater than 0.1 during more than 90 days (the use of different thresholds did not alter our conclusions substantially). Finally, for each ESM, we chose a pseudo-species to represent an observed species when it has the lowest MAE in the three regions (Figure 31a-c). Therefore, the 44 species may be characterised by different pseudo-species from one ESM to another. This allowed us to examine inter-ESM variability (Figures S6-8). The relationships between predicted

and observed daily abundance climatologies of each species in each region were examined by means of a Spearman rank correlation (Figure 31d-f).

### 3.1.2.5. Long-term changes in the abundance of the three taxonomic groups

Daily abundances of the 44 species were estimated from Equation 3 between 1850 and 2100, using environmental variables assessed from the six models and for the two scenarios (Section 4 above) (Figures S3-5). To assess the mean abundance of each group, we weighted species abundance by using their P90 (Section 1) and this to account for species contemporary difference in abundance within the three groups (oblates, prolates, and dinoflagellates). We chose to use the 90$^{th}$ percentile instead of the maximum abundance value because the latter would have been too sensitive to outliers that may originate from multiple causes (e.g. exceptional abundance, patchiness or abundance misestimated from only 3m$^3$ of seawater filtered). Predicted abundances were finally standardised between 0 (the lowest abundance) and 1 (the highest) for each taxonomic group, all ESMs and scenarios as follows (Figure 32b-g, Figure 33b-g and Figure 34b-g and Figures S3-5):

$$D_m^* = \frac{D_{(m)} - min(D)}{max(D) - min(D)} \qquad (4)$$

With $D_m^*$ the standardised predicted abundance of a group for a given ESM, scenario and month $m$, $D_m$ the predicted abundance for month $m$ and $min(D)$ and $max(D)$ the minimal and maximal predicted abundance, respectively. Finally, the relationships between predicted and observed standardised monthly abundances were quantified for 1958-2014 (i.e. overlapping period of observed and predicted data) by using the Spearman correlation coefficient (Figure S1 and Table S2).

### 3.1.2.6. Long-term phenological changes of the three taxonomic groups.

We examined phenological shifts of the three taxonomic groups between 1850 and 2100, for the two scenarios and the six ESMs in the three regions. Average daily abundances of each group were smoothed by means of a 15-order moving average (i.e. time scale covering a month) and then, for each year between 1850 and 2100, we estimated six phenological indices: i.e. (1) Maximum Annual Abundance (MAA), (2) the day where MAA is reached, (3) initiation and (4) termination of the Seasonal Reproductive Period (SRP), i.e. the first and the last days where the abundance is ≥50% of MAA, (5) seasonal duration, i.e. the number of consecutive days where abundance is ≥ 50% of MAA, and (6) Integrated Mean Annual Abundance (IMAA)(Figure S2). The six indices were calculated for each taxonomic group, all models, scenarios and North Atlantic regions. When a bimodality was observed (i.e. two population peaks within a single year) we only considered the peak with the highest abundance. An annual average of the results was then calculated for 1850-2100 for each scenario and North Atlantic region (Figure 32, Figure 33 and Figure 34). In addition, we focussed on three decades: 1850-1859, 2000-2009 and 2090-2099. For these decades, we assessed the

inter-model average and the associated standard deviation of each index (mean ± uncertainty; Tables S3-11)(Kwiatkowski et al., 2020).

### 3.1.3. Results

We first examined observed phenological variations in the daily abundance of 44 phytoplanktonic species (Table S1), including oblate and prolate diatoms and dinoflagellates using data originating from the Continuous Plankton Recorder (CPR) survey collected between 1958 and 2014 in three distinct regions of the North Atlantic Ocean: (i) the North Sea, (ii) the North-East Atlantic and (iii) the Labrador Sea (see Methods). To account for the environmental variability that was observed in these regions during that period (Beaugrand et al., 2019a; Edwards et al., 2002), we estimated daily climatologies of the standardised abundance of each species for three 18-year periods : 1958-1976, 1977-1995 and 1996-2014 (365 days x 3 = 1095 days).

We used an approach based on the MacroEcological Theory on the Arrangement of Life (METAL) to reconstruct long-term changes in the abundance of the three phytoplankton groups (Methods)(Beaugrand, 2015a; Caracciolo et al., 2021). The use of METAL enabled us to convert the monthly CPR observations at a daily scale, which is more adapted to reconstruct the long-term phenological changes from 1850 to 2100 and examine their trends (Henson et al., 2018). We generated a set of 1,755,000 pseudo-species (i.e. virtual species), each being characterised by a unique ecological niche (i.e. the set of environmental conditions enabling a species to grow and reproduce (Hutchinson, 1957)). The niche was based on three environmental variables known to influence phenology at high latitudes (i.e. Sea Surface Temperature or SST, expressed in °C; Surface Downwelling Shortwave Radiation or SDSR, $W.m^{-2}$; and dissolved nitrate concentration, expressed in $mol.m^{-3}$)(Beaugrand and Kirby, 2018; Boyce et al., 2017; Caracciolo et al., 2021; Lewandowska and Sommer, 2010; Miller, 2004). We then chose the pseudo-species that best reconstructed long-term daily changes in the abundance of the 44 chosen phytoplankton species/taxa during the three time periods and regions of the North Atlantic (Methods). Six ESMs (i.e. CNRM-ESM2-1, GFDL-ESM4, IPSL-CM6A-LR, MPI-ESM1-2-LR, NorESM2-LM and UKESM1-0-LL, all involved in the Coupled Model Intercomparison Project Phase 6, CMIP6 (Eyring et al., 2016)) were used to characterise the ecological niche of the 44 species.

Model performance was first assessed by visually comparing the observed and modelled daily abundance of each species for the three periods and North Atlantic regions (Animation S1-3). The phenology of most species was well reconstructed by our approach and the Mean Absolute Error (MAE) between the observed and modelled abundance was generally low (below 0.1 for diatoms and below 0.2 for dinoflagellates; Figure 31a-c; see Methods). We then estimated the Spearman rank correlation coefficient between observed and predicted patterns, which revealed that for most species, correlations were highly positive (Figure 31d-f). A few low correlations were observed, however, indicating that for some species our approach did not work well in some North Atlantic regions. However, their weight in the subsequent analyses was low, as shown by the value of their P90 (i.e. the $90^{th}$ percentile of their abundance, Methods, Figure 31d-f); the higher the P90, the higher

the abundance of a species in a North Atlantic region and therefore the greater its weight in subsequent analyses at the group level (Figure 32, Figure 33 and Figure 34).

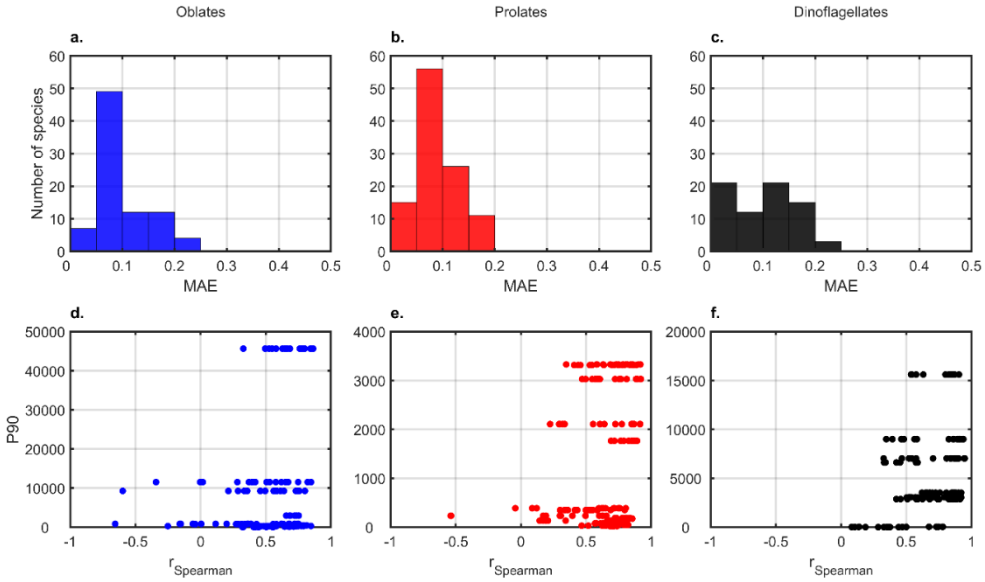



*Figure 31. Model performance. (a-c) Distribution of the MAEs calculated between observed and predicted abundances of each species/taxa for (a) oblates (blue bars), (b) prolates (red bars) and (c) dinoflagellates (black bars) in the three oceanic systems. (d-f) Relationships between the Spearman rank correlation (rSpearman) calculated between the observed and the predicted abundances, and the corresponding P90 values for (d) oblates (blue dots), (e) prolates (red dots) and (f) dinoflagellates (black dots) in the three oceanic regions. Comparison between the abundance predicted by the models forced by the six ESMs and the abundance observed by the CPR are shown in Animations S1-3. Mean monthly observed and predicted standardised abundances of the three taxonomic groups in the three oceanic systems are shown in Figure S1.*

Modelled species abundances were finally aggregated at a monthly scale for each group (oblates, prolates and dinoflagellates) by weighting the abundance of each species according to the value of its P90. A monthly scale was first chosen for comparison of the modelling approach with the in situ monthly observations collected by the CPR survey, whom sampling is carried out at a monthly interval (Richardson et al., 2006). Long-term monthly changes in the resulting mean abundance of each group were well-reconstructed by our approach in the three regions and all were positively correlated significantly (Figure S1 and Table S2). Our results were statistically more significant in the North Sea and the North-East Atlantic than in the Labrador Sea (Spearman rank correlations coefficient between 0.7 and 0.83 in the North Sea, 0.64 and 0.73 in the North-East Atlantic and 0.5 and 0.58 in the Labrador Sea; Table S2).

We then used the six ESMs and two scenarios, i.e. a medium (SSP245) and a high (SSP585) emission scenarios (O'Neill et al., 2017, 2016), to calculate the mean daily abundance of each group between 1850 and 2100 in the three North-Atlantic regions to examine more precisely their phenological shifts (see Methods). An historical perspective is important to better understand the magnitude of the present and future shifts (Beaugrand et al., 2015). Long-term phenological shifts were assessed by calculating six indices: (1) Maximum Annual Abundance (MAA), (2) the day where MAA is reached, (3) initiation and (4) termination of the Seasonal Reproductive Period

(SRP), i.e. the first and the last days where the abundance is ≥50% of MAA, (5) seasonal duration, i.e. the number of consecutive days where abundance is ≥ 50% of MAA, and (6) Integrated Mean Annual Abundance (IMAA; Figure S2). Seasonal dilatation and contraction is an extension or a shrinkage of SRP, respectively (Beaugrand and Kirby, 2018). The six indices were calculated for each taxonomic group, all models, scenarios and North Atlantic regions. An annual average of the results was subsequently calculated for 1850-2100 for each scenario and North Atlantic region (Figure 32, Figure 33 and Figure 34). In addition, we focussed on three decades: 1850-1859, 2000-2009 and 2090-2099. For these decades, we assessed the inter-model average and the associated standard deviation of each index (mean ± uncertainty; Tables S3-11)(Kwiatkowski et al., 2020).

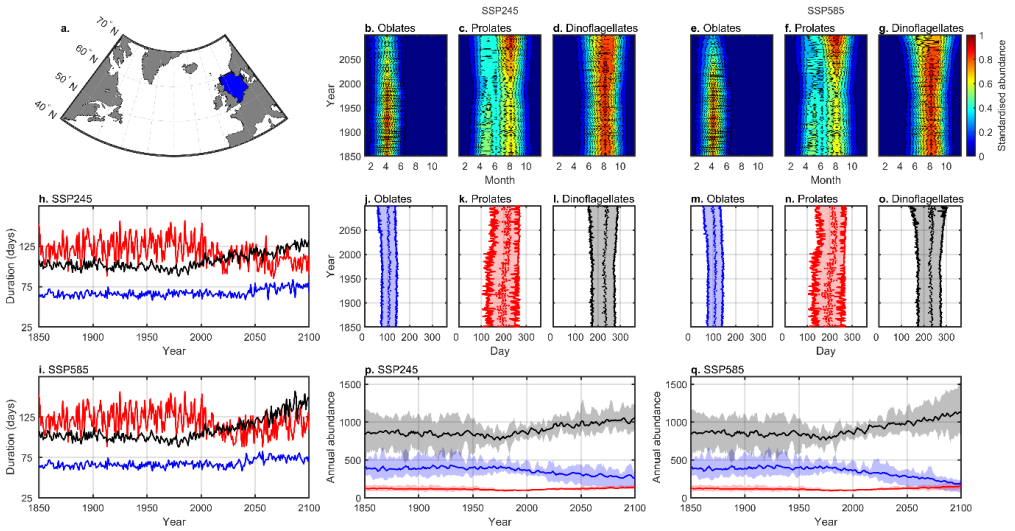



*Figure 32. Modelled long-term phenological changes of diatoms and dinoflagellates in the North Sea (1850-2100). (a) Spatial distribution of CPR sampling. (b-g) Long-term monthly changes in the standardised abundance of (b and e) oblates, (c and f) prolates and (d and g) dinoflagellates for scenarios (b-d) SSP245 and (e-g) SSP585 (Figure S3). Each panel was standardised between 0 and 1. Colours denote the mean monthly standardised abundance of a taxonomic group. (h-i) Long-term annual changes in phenology duration for scenarios (h) SSP245 and (i) SSP585. (j-o) Long-term changes in the initiation and termination of the Seasonal Reproductive Period (SRP) for (j and m) oblates, (k and n) prolates and (l and o) dinoflagellates. In each panel, the first and last tick line denotes the initiation and the termination day of the SRP. Shading denotes the duration of the SRP and the dotted line displays the day where Maximum Annual Abundance (MAA) is reached. (p-q) Long-term changes in Integrated Mean Annual Abundance (IMAA; cell per CPR sample) for scenarios (p) SSP245 and (q) SSP585. The shade denotes the minimum and maximum mean abundance estimated from the environmental variables originating from the six ESMs. In h-q oblates are in blue, prolates in red and dinoflagellates in black. The meaning of the six phenological indices is summarised in Figure S2.*

In the North Sea (Figure 32a), modelled long-term daily phenological changes exhibited similar patterns for both SSP scenarios, with a shift in the diatom community composition, although changes were highest under SSP585 (Figure 32b-q). Oblate diatoms are expected to decline, the MAA decreasing from 1671.1 ± 494.7 cells per CPR sample in 1850-1859 to 1176.9 ± 573.2 or 785.1 ± 347.7 cells per CPR sample in 2090-2099 for scenarios SSP245 and SSP585, respectively (Figure 32b and e and Figure S3 and Table S3). In contrast, prolate MAA rose from 302 ± 53.3 cells per CPR sample in 1850-1859 to 360.5 ± 77.5 or 392.6 ± 115.5 cells per CPR sample in 2090-2099 for

SSP245 and SSP585, respectively (Figure 32c and f and Figure S3 and Table S4). No significant changes in the duration of the SRP was found for both groups (Figure 32h-i and Tables S3-4), although they exhibit an earlier and a later (both initiation and termination) phenology, respectively (Figure 32j-k and m-n and Tables S3-4). These phenological shifts were accompanied by a shift in their IMAA, our results suggesting that oblates may experience a strong decrease from 390 ± 92.8 cells per CPR sample in 1850-1859 to 283.8 ± 97.4 or 192 ± 59 cells per CPR sample in 2090-2099 for scenarios SSP245 and SSP585, respectively. An opposite pattern is predicted for prolates, although it may be less important (from 122.8 ± 32.7 cells per CPR sample in 1850-1859 to 134.5 ± 18.6 or 148.8 ± 30.1 cells per CPR sample in 2090-2099 for SSP245 and SSP585, respectively; Figure 32p-q and Tables S3-4).

Modelled changes in dinoflagellates exhibited a phenological dilatation. Their MAA remained constant throughout the period, although they decreased at the very end of the century under SSP585 from 2853.9 ± 336.3 cells per CPR sample in 1850-1859 to 2599.3 ± 553.2 in 2090-2099 (Figure 32d and g and Figure S3 and Table S5). The duration of their SRP is expected to rise throughout the XXI$^{st}$ century (from a mean of 103 ± 16 days in 1850-1859 to 107 ± 6 days in 2000-2009 and to 127 ± 13 or 143 ± 25 days in 2090-2099 under SSP245 and SSP585 respectively; Figure 32h-i and Table S5) with both an earlier initiation and a later termination of their SRP (Figure 32l and o and Table S5), associated with an increase in their IMAA (Figure 32p-q and Table S5).

In the North-East Atlantic region (Figure 33a), phenological shifts were also predicted by our approach, the magnitude of which depending upon warming intensity (Figure 33b-q). Oblate MAA decreased after the end of the XX$^{th}$ century whereas prolate MAA rose *circa* 2050 under both scenarios (Figure 33b-c and e-f, Figure S4 and Tables S6-7). Oblates underwent a phenological contraction from 76 ± 22 days in 2000-2009 to 69 ± 14 or 64 ± 15 days in 2090-2099 for SSP245 and SSP585, respectively (Figure 33h-i and Table S6), accompanied by an earlier initiation in their SRP (Figure 33j and m and Table S6). In contrast, prolates exhibited a strong phenological dilatation from 91 ± 6 days in 2000-2009 to 110 ± 11 or 119 ± 22 days in 2090-2099 for scenarios SSP245 and SSP585, respectively (Figure 33h-i and Table S7), with an earlier initiation and a later termination of their SRP (Figure 33k and n and Table S7). At the end of the century, they reached earlier their MAA in the year, i.e. around day number 204 ± 16 or 197 ± 19 in 2090-2099 for scenarios SSP245 and SSP585, respectively, in comparison to day 215 ± 4 in 1850-1859 (Figure 33k and n and Table S7). Oblate IMAA remained constant under scenario SSP245 but decreased under SSP585. Prolate IMAA slightly increased under both scenarios but remained below the value observed for oblates (Figure 33p-q and Tables S6-7).

Dinoflagellates also underwent a phenological dilatation, associated with an increase of their MAA throughout the XXI$^{st}$ century (Figure 33d and g and Figure S4 and Table S8), a longer duration (Figure 33h-i and Table S8) caused by an earlier initiation (around day number 215 ± 4 in 1850-1859 and 204 ± 16 or 197 ± 19 for SSP245 and 585, respectively) and a later termination (around day number 254 ± 9 in 1850-1859 and 262 ± 5 or 268 ± 9 in 2090-2099 for scenarios SSP245 and SSP585, respectively, Table S8) of their SRP. Their IMAA rose above those of oblates during the first half of

the XXI$^{st}$ century, although there was a high inter-model variance (Figure 33p-q and Table S8).

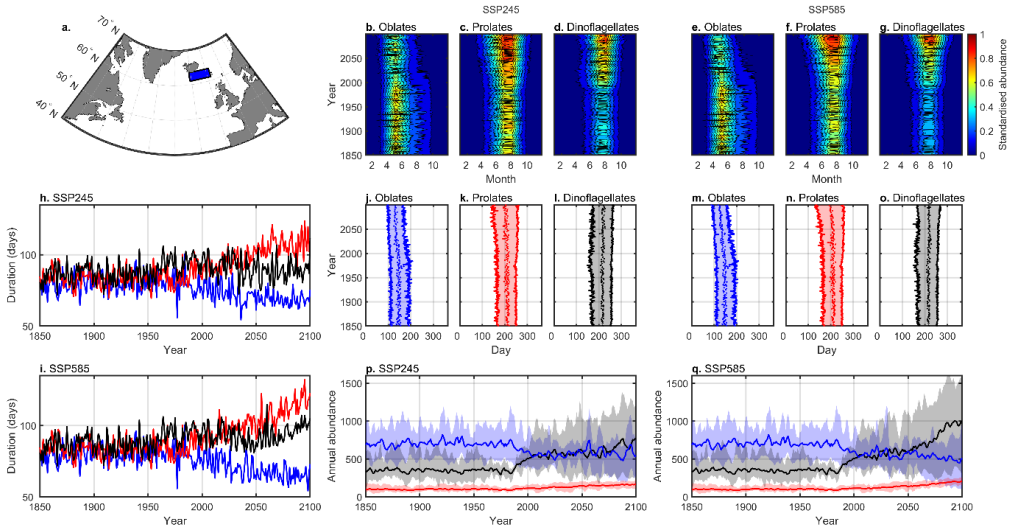



*Figure 33. Modelled long-term phenological changes of diatoms and dinoflagellates in the North-East Atlantic (1850-2100). (a) Spatial distribution of CPR sampling. (b-g) Long-term monthly changes in the standardised abundance of (b and e) oblates, (c and f) prolates and (d and g) dinoflagellates for scenarios (b-d) SSP245 and (e-g) SSP585 (Figure S4). Each panel was standardised between 0 and 1. Colours denote the mean monthly standardised abundance of a taxonomic group. (h-i) Long-term annual changes in phenology duration for scenarios (h) SSP245 and (i) SSP585. (j-o) Long-term changes in the initiation and termination of the Seasonal Reproductive Period (SRP) for (j and m) oblates, (k and n) prolates and (l and o) dinoflagellates. In each panel, the first and last tick line denotes the initiation and the termination day of the SRP. Shading denotes the duration of the SRP and the dotted line displays the day where Maximum Annual Abundance (MAA) is reached. (p-q) Long-term changes in Integrated Mean Annual Abundance (IMAA; cell per CPR sample) for scenarios (p) SSP245 and (q) SSP585. The shade denotes the minimum and maximum mean abundance estimated from the environmental variables originating from the six ESMs. In h-q oblates are in blue, prolates in red and dinoflagellates in black. The meaning of the six phenological indices is summarised in Figure S2.*

In the Labrador Sea (Figure 34a), distinct phenological shifts were found, also depending upon the SSP scenarios (Figure 34b-q). Under scenario SSP245, oblates maintained a high MAA (i.e. 3770.1 ± 769.1 cells per CPR sample in 2000-2009 *versus* 4055.5 ± 545.4 cells per CPR sample in 2090-2099; Figure S5 and Table S9) although the duration of their SRP slightly diminished from 87 ± 11 days in 2000-2009 to 84 ± 13 days in 2090-2099 (Figure 34h and j and Table S9). Prolates maintained both their MAA and IMAA throughout the period 1850-2100 (Figure 34c and p and Figure S5 and Table S10), the most significant changes occurring in the duration of their SRP (from 97 ± 31 days in 2000-2009 to 107 ± 36 in 2090-2099; Figure 34h and Table S10) accompanied by an earlier phenology (Figure 34k and Table S10). Dinoflagellates should exhibit a large increase of their MAA, especially after 2050, from 910 ± 428.4 cells per CPR sample in 2000-2009 to 1914.4 ± 1125.2 in 2090-2099 (Figure 34d and Table S11) without a significant increase in the duration of their SRP (Figure 34h and Table S11) or a change in initiation or termination (Figure 34l and Table S10). Their IMAA slightly increased but remained above that of oblates (Figure 34p and Tables S8 and 10).

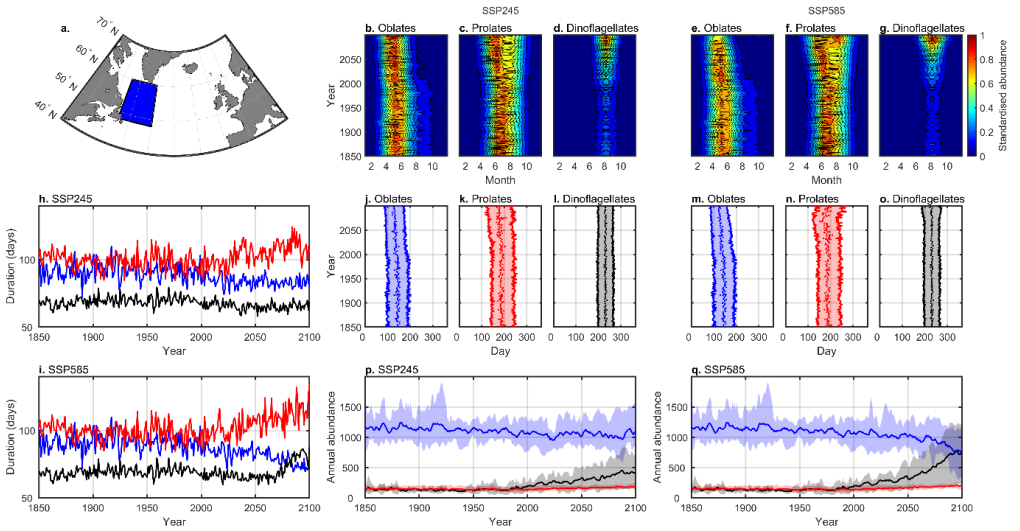



*Figure 34. Modelled long-term phenological changes of diatoms and dinoflagellates in the Labrador Sea (1850-2100). (a) Spatial distribution of CPR sampling. (b-g) Long-term monthly changes in the standardised abundance of (b and e) oblates, (c and f) prolates and (d and g) dinoflagellates for scenarios (b-d) SSP245 and (e-g) SSP585 (Figure S5). Each panel was standardised between 0 and 1. Colours denote the mean monthly standardised abundance of a taxonomic group. (h-i) Long-term annual changes in phenology duration for scenarios (h) SSP245 and (i) SSP585. (j-o) Long-term changes in the initiation and termination of the Seasonal Reproductive Period (SRP) for (j and m) oblates, (k and n) prolates and (l and o) dinoflagellates. In each panel, the first and last tick line denotes the initiation and the termination day of the SRP. Shading denotes the duration of the SRP and the dotted line displays the day where Maximum Annual Abundance (MAA) is reached. (p-q) Long-term changes in Integrated Mean Annual Abundance (IMAA; cell per CPR sample) for scenarios (p) SSP245 and (q) SSP585. The shade denotes the minimum and maximum mean abundance estimated from the environmental variables originating from the six ESMs. In h-q oblates are in blue, prolates in red and dinoflagellates in black. The meaning of the six phenological indices is summarised in Figure S2.*

Under SSP585, our approach predicted a shift from a system dominated by oblates to a system dominated by dinoflagellates (Figure 34b-q). The oblates exhibited a decrease of their MAA from 3770.1 ± 769.1 cells per CPR sample in 2000-2009 to 3184.4 ± 727.2 in 2090-2099 (Figure 34e and Figure S5 and Table S9), accompanied by a decrease in the duration of their SRP from 86.9 ± 10.5 days in 2000-2009 to 72.8 ± 8.3 days in 2090-2099 (Figure 34i and Table S9), and an earlier initiation (Figure 34m and Table S9). Their IMAA sharply decreased, especially during the second half of the XXI$^{st}$ century, from 1077.1 ± 178.6 cells per CPR sample in 2000-2009 to 777.9 ± 245.5 in 2090-2099 (Figure 34q and Table S9). No changes were found for prolate MAA (Figure 34f and Figure S5 and Table S10) but their SRP lasted longer, from 97 ± 31 days in 2000-2009 to 115 ± 35 in 2090-2099 (Figure 34i and n and Table S10). Their IMAA was low under this scenario (Figure 34p-q and Table S10). Dinoflagellate shifts were amplified, their MAA increased substantially, from 910 ± 428.4 cells per CPR sample in 2000-2009 to 2898.2 ± 1052.0 in 2090-2099 (Figure 34g and Figure S5 and Table S11). These changes were more prominent at the end of the XXI$^{st}$ century, from 66 ± 12 days in 2000-2009 to 82 ± 18 in 2090-2099 (Figure 34i and Table S11), resulting in a phenological dilatation (Figure 34o and Table S11). Their IMAA tripled between the beginning and the end of the XXI$^{st}$ century, from 204.5 ± 66.9 cells per CPR sample in 2000-2009 to 738.8 ± 319.3 in 2090-2099 (Table S11), resulting in a shift from oblates to dinoflagellates. As the long-term phenological changes of the three groups were

more difficult to reconstruct in this region (Table S2), our predictions should be considered with degree of caution.

### 3.1.4. Discussion

Phenological shifts predicted by our model were relatively moderate for 1850-1950 in the three regions of the North Atlantic (Figure 32, Figure 33 and Figure 34). Major changes have started to occur from 1950 onwards when the rate of ocean warming have increased, which led to an increase in water column stratification (IPCC, 2019). Modelled shifts are in agreement with theoretical predictions, i.e. that spring species (i.e. oblates) experience a phenological contraction and an earlier phenology whereas late summer species (i.e. prolates and dinoflagellates) exhibit a phenological dilatation with both an earlier initiation and a later termination (Beaugrand and Kirby, 2018). The phenological contraction of oblates, associated with a decline in abundance, is indicative of species/taxa that reach the limits of their phenological plasticity (Beaugrand and Kirby, 2018), probably because an earlier phenology in spring is impossible as winter photoperiod and light levels limits primary production at high latitudes (Boyce et al., 2017; Caracciolo et al., 2021). Therefore, a biogeographical shift of these diatoms toward higher latitudes is expected and would be in agreement with the increased primary production projected by ESMs in the Arctic region (Kwiatkowski et al., 2020). On the contrary, the phenological expansion of prolates and dinoflagellates is indicative of species/taxa that are adapted to the reinforced stratified and nutrient-depleted conditions in the future North Atlantic Ocean (Beaugrand and Kirby, 2018; Kwiatkowski et al., 2020). The elongated cell shape of prolates reduces their sinking speed while not affecting their capacity to uptake nutrients, enabling them to remain in the well-lit upper part of the water column (Kléparski et al., 2022b; Padisak et al., 2003). Furthermore, it has been shown that some taxa belonging to this taxonomic group (e.g. *Rhizosolenia*), were able to passively migrate between the surface layers and the nutricline, enabling them to harvest nutrients more efficiently (Kemp and Villareal, 2018, 2013). Similarly, the dinoflagellates have flagella which enable them to move and actively exploit nutrients in the water column. They are also capable of switching nutritional strategies to mixotrophy when less nutrients are available (Miller, 2004).

Diatoms and dinoflagellates have been historically separated into two distinct functional types (Kemp and Villareal, 2018; Margalef, 1978). According to Margalef's mandala (Margalef, 1978), diatoms thrive when the water column is poorly stratified and turbulent and when nutrients concentration is relatively high. In contrast, dinoflagellates thrive in stratified conditions and lower nutrients concentration. From Margalef's framework (Margalef, 1978), a decline in primary production and carbon export (therefore the biological pump (Passow and Carlson, 2012)) is anticipated in the North Atlantic because ocean warming will enhance water column stratification and reduce nutrients, which are believed to negatively affect diatoms (Bopp et al., 2005; Marinov et al., 2010). Our results nuance this prediction because prolates may indeed persist in a more stratified ocean, although to an abundance level that might not compensate the large diminution in the abundance experienced by oblates, which are

responsible of massive blooms in spring generating high carbon export through the generation of marine snow (Raven and Waite, 2004; Smetacek, 1985). Therefore, it is unlikely that the slight increase in prolates will overcome the diminution in carbon exportation induced by the diminution in oblate abundance. Furthermore, the transfer efficiency of carbon to the deep ocean is reduced with smaller sized slow-sinking cells, and therefore the shift from large round oblates to smaller elongated prolates will also decrease the rate of carbon export (Henson et al., 2022). However, our results also show that dinoflagellate abundance will rise in the three North Atlantic regions. Although the contribution of this group to the biological pump is poorly understood, a study has shown that dinoflagellate bloom sometimes coincides with a rise in carbon export at an annual scale in the North-East Atlantic at 3000m (Henson et al., 2012). It has also been shown that the consideration of mixotrophic organisms (e.g. dinoflagellates) within biogeochemical models can increase the flux of carbon export (Ward and Follows, 2016). Therefore, this group might alleviate to some extent the consequences of the reduction of oblates on the biological pump in the North Atlantic.

ESMs used as part of CMIP5 include a few functional types (i.e. a group of species that have a similar trophic and/or biogeochemical function) to represent the phytoplanktonic community, often with a single grouping for diatoms (Séférian et al., 2020; Tréguer et al., 2018). The last generations of ESMs in CMIP6 have improved the representation of phytoplankton ecology by including more functional types (e.g. diazotrophs for some models), which has enhanced model performances, however diatoms are still represented by a single grouping (Séférian et al., 2020). Results from model inter-comparisons highlight the great sensitivity of models to the types of functional groups they include, as well as the ecological traits they represent (Fu et al., 2016; Taucher and Oschlies, 2011). Dutkiewicz and colleagues have recently used morphological traits (i.e. cell size), biogeochemical functions and thermal tolerances to define more than 350 phytoplankton types (Dutkiewicz et al., 2020). Their results showed that model predictions were altered by the number as well as the nature of phytoplankton traits, each ones controlling a distinct aspect of the species' responses (Dutkiewicz et al., 2020). They concluded that the addition of traits, such as cell shape, might improve model predictions (Dutkiewicz et al., 2020). *In situ* studies have also highlighted the importance of phytoplankton composition for carbon exportation (Dybwad et al., 2021; Henson et al., 2012). However, such models with hundreds of phytoplanktonic type cannot realistically be used to realise centuries scaled climatic simulations because of computational cost. Hence, the distinction of oblate and prolate diatoms as well as dinoflagellates in ESM models may be a simple way to improve model predictions because these three groups have different phenologies and responses to climate change, which may affect our current understanding of the consequences of phytoplankton shifts on the biological carbon pump (Karp-Boss and Boss, 2016; Kléparski et al., 2022b; Litchman and Klausmeier, 2008; Naselli-Flores et al., 2021; Ryabov et al., 2021).

*Acknowledgments*:

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. Funding that supports data collection has come from a number of contracts since inception of the Continuous Plankton Recorder survey, the current funded projects include: the UK Natural Environment Research Council, Grant/Award Number: NE/R002738/1 and NE/M007855/1; EMFF; Climate Linked Atlantic Sector Science, Grant/Award Number: NE/ R015953/1, DEFRA UK ME-5308, NSF USA OCE-1657887, DFO CA F5955-150026/001/HAL, NERC UK NC-R8/H12/100, Horizon 2020: 862428 Atlantic Mission and AtlantECO 862923, IMR Norway, DTU Aqua Denmark and the French Ministry of Environment, Energy, and the Sea (MEEM).

## 3.1.5. Supplementary Files

**Supplementary Figures.**

**Figure S1. Comparisons between observed and predicted mean monthly abundance of the three taxonomic groups (oblates, prolates and dinoflagellates).** (**a**-**i**) Long-term changes in the monthly abundance of (**a**-**c**) oblates, (**d**-**f**) prolates and (**g**-**i**) dinoflagellates in the North Sea (**a**, **d** and **g**), the North-East Atlantic (**b**, **e** and **h**) and the Labrador Sea (**c**, **f** and **i**). (**j**-**r**) Predicted long-term changes in the monthly abundance of (**j**-**l**) oblates, (**m**-**o**) prolates and (**p**-**r**) dinoflagellates in the North Sea (**j**, **m** and **p**), the North-East Atlantic (**k**, **n** and **q**) and the Labrador Sea (**l**, **o** and **r**). Colours denote the mean monthly observed or predicted abundance of a taxonomic group. The abundance in each panel was standardised between 0 and 1 for comparison purpose (see Methods). Spearman correlation coefficients between observed and predicted abundance are displayed in Table S1.

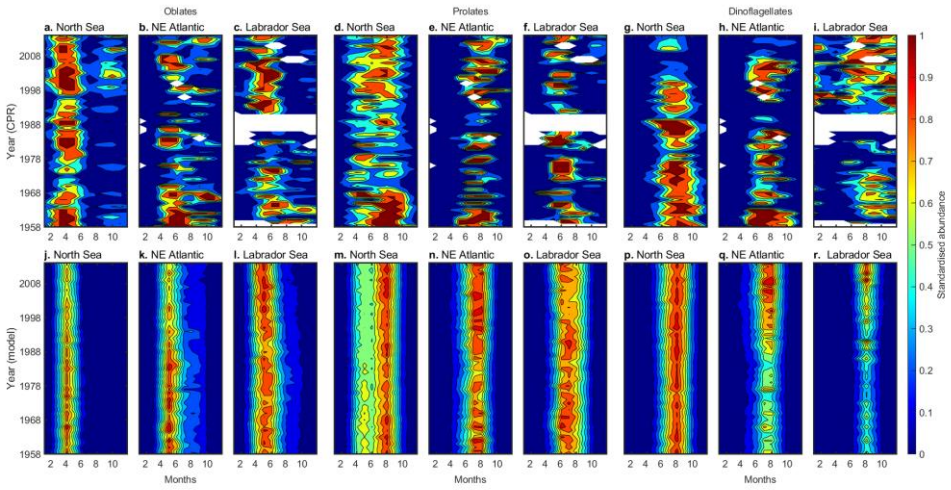


**Figure S2. Theoretical diagram explaining graphically the six phenological indices used in this study.** Six phenological indices were defined: the Maximum Annual Abundance (MAA; index 1), the day where MAA is reached (index 2), the initiation (index 3) and the termination (index 4) of the Seasonal Reproductive Period (SRP), i.e. the first and the last days where the abundance is ≥50% of MAA, the seasonal duration (index 5), i.e. the number of consecutive days where abundance is ≥ 50% of MAA, and the Integrated Mean Annual Abundance (IMAA; index 6).

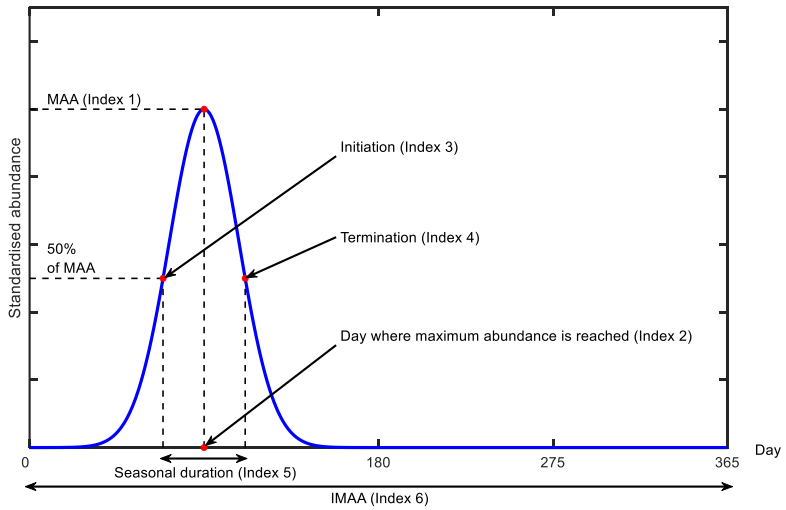



**Figure S3. Long-term changes in the monthly abundance of oblates, prolates and dinoflagellates in the North Sea (1850-2100).** Long-term monthly changes in the abundance of (**a**-**b**) oblates, (**c**-**d**) prolates and (**e**-**f**) dinoflagellates for scenarios (**a**, **c** and **e**) SSP245 and (**b**, **d** and **f**) SSP585. Colours denote the mean monthly predicted abundance of a taxonomic group. Each panel was standardised between 0 and 1 for comparison purpose (see Methods).

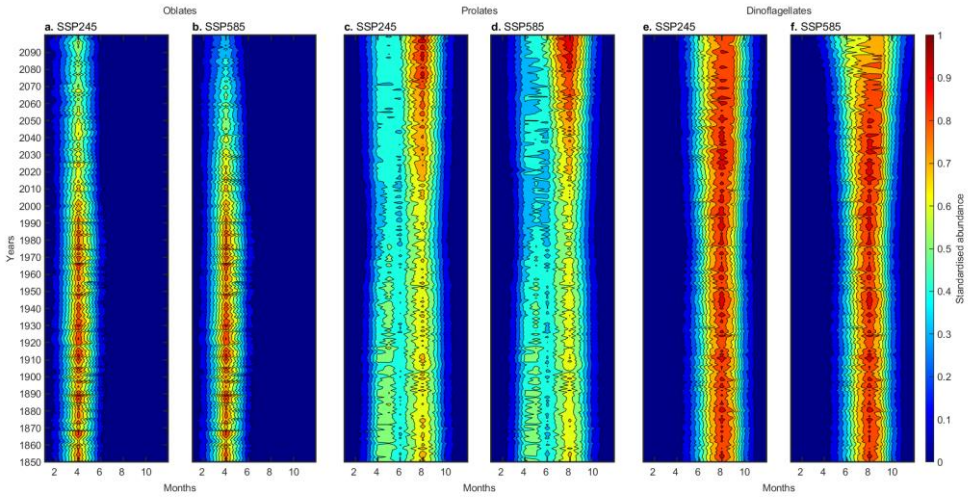


**Figure S4. Long-term changes in the monthly abundance of oblates, prolates and dinoflagellates in the North-East Atlantic (1850-2100).** Long-term monthly changes in the abundance of (**a**-**b**) oblates, (**c**-**d**) prolates and (**e**-**f**) dinoflagellates for scenarios (**a**, **c** and **e**) SSP245 and (**b**, **d** and **f**) SSP585. Colours denote the mean monthly predicted abundance of a taxonomic group. Each panel was standardised between 0 and 1 for comparison purpose (see Methods).

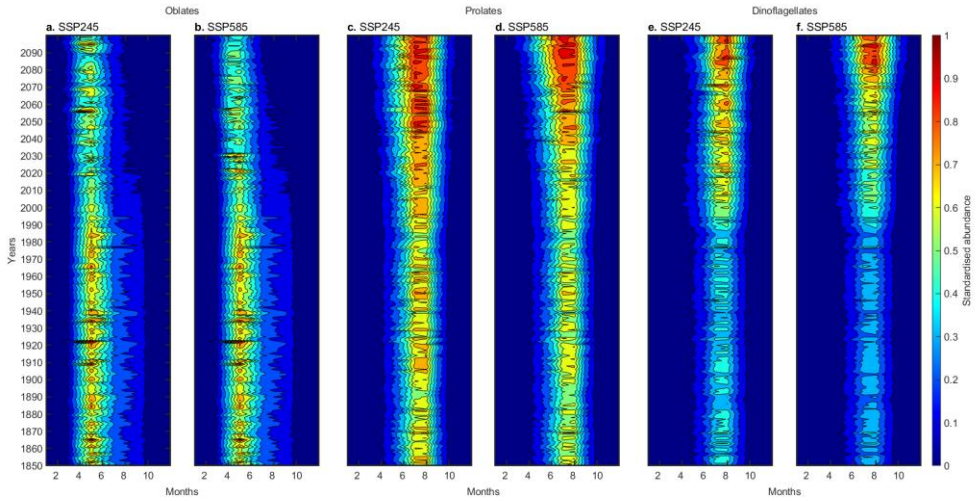



**Figure S5. Long-term changes in the monthly abundance of oblates, prolates and dinoflagellates in the Labrador Sea (1850-2100).** Long-term monthly changes in the abundance of (**a**-**b**) oblates, (**c**-**d**) prolates and (**e**-**f**) dinoflagellates for scenarios (**a**, **c** and **e**) SSP245 and (**b**, **d** and **f**) SSP585. Colours denote the mean monthly predicted abundance of a taxonomic group. Each panel was standardised between 0 and 1 for comparison purpose (see Methods).

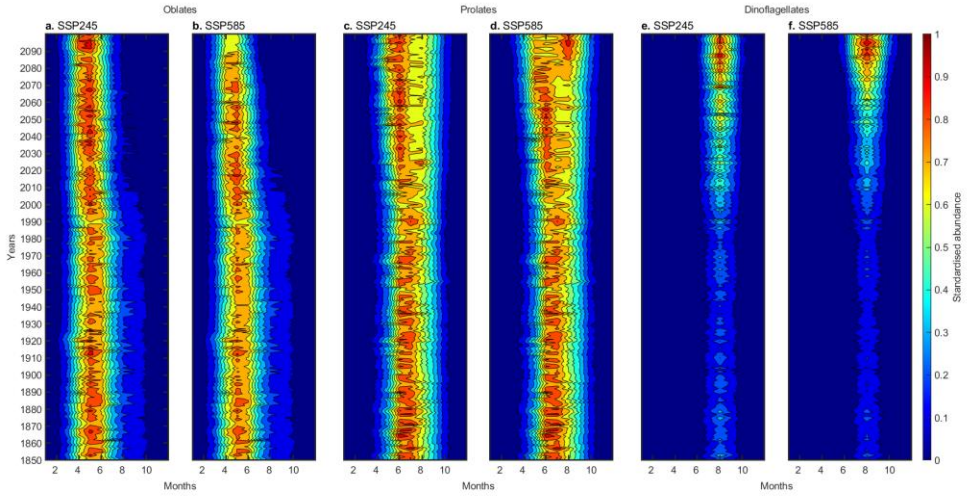


**Figure S6. Long-term forecasted changes in SST, SDSR and nitrate concentrations in the North Sea for the six ESMs and the two scenarios (1850-2100). Long-term forecasted changes** for (**a**) SST, (**b**) SDSR and (**c**) nitrate concentration. Projection for CNRM-ESM2-1 (blue), IPSL-CM6A-LR (red), MPI-ESM1-2-LR (black), GFDL-ESM4 (cyan), UKESM1-0-LL (magenta) and NorESM2-LM (green). Historical (1850-2014) projections are displayed by solid lines and scenarios SSP245 and 585 are displayed by dashed, dotted and dash-dot lines, respectively.

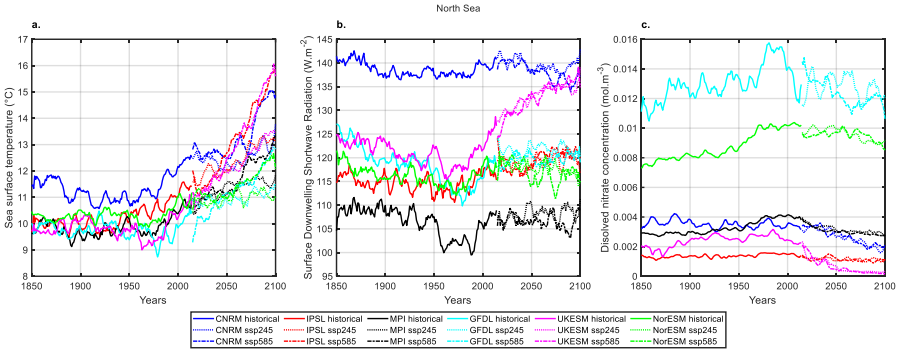



**Figure S7. Long-term forecasted changes in SST, SDSR and nitrate concentrations in the North-East Atlantic for the six ESMs and the two scenarios (1850-2100). Long-term forecasted changes** for (**a**) SST, (**b**) SDSR and (**c**) nitrate concentration. Projection for CNRM-ESM2-1 (blue), IPSL-CM6A-LR (red), MPI-ESM1-2-LR (black), GFDL-ESM4 (cyan), UKESM1-0-LL (magenta) and NorESM2-LM (green). Historical (1850-2014) projections are displayed by solid lines and scenarios SSP245 and 585 are displayed by dashed, dotted and dash-dot lines, respectively.

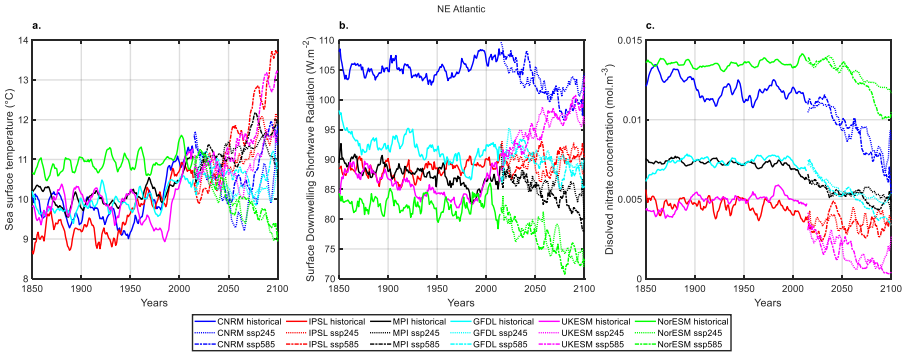
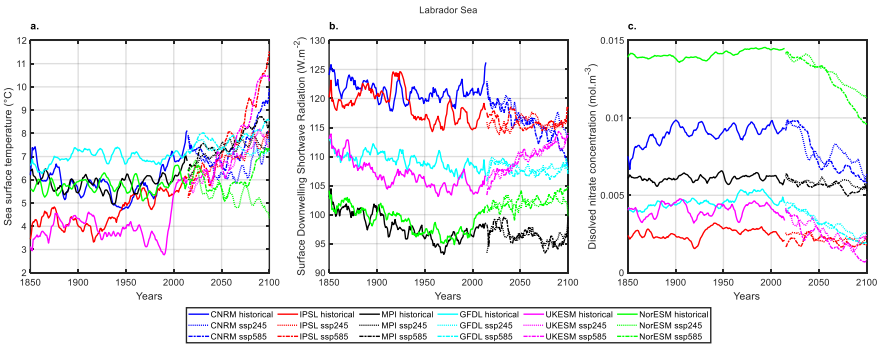



**Figure S8. Long-term forecasted changes in SST, SDSR and nitrate concentrations in the Labrador Sea for the six ESMs and the two scenarios (1850-2100). Long-term forecasted changes** for (**a**) SST, (**b**) SDSR and (**c**) nitrate concentration. Projection for CNRM-ESM2-1 (blue), IPSL-CM6A-LR (red), MPI-ESM1-2-LR (black), GFDL-ESM4 (cyan), UKESM1-0-LL (magenta) and NorESM2-LM (green). Historical (1850-2014) projections are displayed by solid lines and scenarios SSP245 and 585 are displayed by dashed, dotted and dash-dot lines, respectively.



**Supplementary Tables.**

**Table S1. List of the CPR species/taxa that have been used in our analyses.** Species/taxa name are displayed in the first column, their taxonomic group (i.e. oblate and prolate diatoms and dinoflagellates) in the second and the region where they have been modelled in the third. Diatoms are divided according to their cell shape in oblate (i.e. flattened cell) and prolate (elongated cell).

| Species or taxa name | Taxonomic group | Region |
|---|---|---|
| *Bacillaria paxillifera* | oblates | North Sea |
| *Bacteriastrum spp.* | prolates | North Sea and Labrador Sea |
| *Bellerochea malleus* | oblates | North Sea |
| *Ceratium arcticum* | dinoflagellates | Labrador Sea |
| *Ceratium bucephalum* | dinoflagellates | North Sea |
| *Ceratium furca* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium fusus* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium hexacanthum* | dinoflagellates | North Sea |
| *Ceratium horridum* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium lineatum* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium longipes* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium macroceros* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Ceratium minutum* | dinoflagellates | North Sea |
| *Ceratium tripos* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Corethron hystrix* | prolates | North Sea and Labrador Sea |
| *Coscinodiscus concinnus* | oblates | North Sea and Labrador Sea |
| *Cylindrotheca closterium* | oblates | North Sea, North-East Atlantic and Labrador Sea |
| *Dactyliosolen antarcticus* | prolates | North-East Atlantic |
| *Dactyliosolen fragilissimus* | prolates | North Sea |
| *Ephemera planamembranacea* | oblates | Labrador Sea |
| *Eucampia zodiacus* | oblates | North Sea |
| *Guinardia delicatula* | prolates | North Sea |
| *Guinardia flaccida* | prolates | North Sea |
| *Guinardia striata* | prolates | North Sea |

| Species or taxa name | Taxonomic group | Region |
|---|---|---|
| *Gyrosigma spp.* | oblates | North Sea |
| *Lauderia annulata* | prolates | North Sea |
| *Leptocylindrus danicus* | prolates | North Sea |
| *Leptocylindrus mediterraneus* | prolates | North Sea, North-East Atlantic and Labrador Sea |
| *Navicula spp.* | oblates | North Sea and Labrador Sea |
| *Odontella granulata* | oblates | North Sea |
| *Odontella rhombus* | oblates | North Sea |
| *Odontella sinensis* | oblates | North Sea |
| *Proboscia alata* | prolates | North Sea, North-East Atlantic and Labrador Sea |
| *Proboscia indica* | prolates | North Sea and North-East Atlantic |
| *Proboscia inermis* | prolates | North Sea, North-East Atlantic and Labrador Sea |
| *Protoperidinium spp.* | dinoflagellates | North Sea, North-East Atlantic and Labrador Sea |
| *Rhizosolenia hebetata semispina* | prolates | North Sea, North-East Atlantic and Labrador Sea |
| *Rhizosolenia imbricata* | prolates | North Sea and North-East Atlantic |
| *Rhizosolenia setigera* | prolates | North Sea |
| *Rhizosolenia styliformis* | prolates | North Sea, North-East Atlantic and Labrador Sea |
| *Stephanopyxis spp.* | prolates | North Sea |
| *Thalassionema nitzschioides* | oblates | North Sea, North-East Atlantic and Labrador Sea |
| *Thalassiosira spp.* | oblates | North Sea, North-East Atlantic and Labrador Sea |
| *Thalassiothrix longissima* | oblates | North Sea, North-East Atlantic and Labrador Sea |

**Table S2. Spearman rank correlation coefficient between long-term changes in observed and predicted monthly abundance of the three taxonomic groups (oblates, prolates and dinoflagellates).** All correlations were significant ($p<0.01$). Long-term monthly changes in observed and predicted abundance are displayed in Figure S1.

| | North Sea | North-East Atlantic | Labrador Sea |
|---|---|---|---|
| **Oblates** | 0.7 | 0.64 | 0.55 |
| **Prolates** | 0.79 | 0.73 | 0.58 |
| **Dinoflagellates** | 0.83 | 0.72 | 0.5 |

**Table S3. Mean values (based on six ESMs) of the phenological indices of oblates for three decades and two scenarios in the North-Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 74 | 5 | 139 | 8 | 65 | 8 | 107 | 4 | 1671.1 | 494.7 | 390.7 | 92.8 |
| **2000-09** | 74 | 6 | 141 | 9 | 67 | 10 | 109 | 5 | 1482.6 | 372.5 | 347.0 | 40.8 |
| **2090-99 SSP245** | 60 | 29 | 136 | 11 | 76 | 23 | 105 | 9 | 1176.9 | 573.2 | 283.8 | 97.4 |
| **2090-99 SSP585** | 57 | 29 | 131 | 12 | 74 | 21 | 97 | 20 | 785.1 | 347.7 | 192.0 | 59.0 |

**Table S4. Mean values (based on six ESMs) of the phenological indices of prolates for three decades and two scenarios in the North-Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 130 | 34 | 252 | 33 | 123 | 26 | 194 | 40 | 302.0 | 53.3 | 122.8 | 32.7 |
| **2000-09** | 141 | 21 | 265 | 5 | 123 | 20 | 218 | 9 | 274.6 | 39.3 | 104.0 | 4.9 |
| **2090-99 SSP245** | 154 | 33 | 260 | 9 | 106 | 28 | 213 | 23 | 360.5 | 77.5 | 134.5 | 18.6 |
| **2090-99 SSP585** | 143 | 36 | 261 | 34 | 119 | 42 | 213 | 39 | 392.6 | 115.5 | 148.8 | 30.1 |

**Table S5. Mean values (based on six ESMs) of the phenological indices of dinoflagellates for three decades and two scenarios in the North-Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 173 | 8 | 276 | 9 | 103 | 16 | 235 | 8 | 2853.9 | 336.3 | 856.2 | 187.3 |
| **2000-09** | 170 | 2 | 277 | 5 | 107 | 6 | 232 | 5 | 2873.1 | 220.6 | 886.2 | 49.8 |
| **2090-99 SSP245** | 160 | 7 | 287 | 9 | 127 | 13 | 239 | 7 | 2819.2 | 328.2 | 1017.9 | 124.8 |
| **2090-99 SSP585** | 150 | 18 | 293 | 8 | 143 | 25 | 235 | 14 | 2599.3 | 553.2 | 1109.4 | 260.1 |

**Table S6. Mean values (based on six ESMs) of the phenological indices of oblates for three decades and two scenarios in the North-East Atlantic.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 113 | 20 | 190 | 39 | 78 | 22 | 215 | 4 | 2730.8 | 1143.9 | 685.1 | 176.1 |
| **2000-09** | 107 | 9 | 183 | 27 | 76 | 22 | 212 | 6 | 2242.4 | 1006.6 | 556.2 | 156.4 |
| **2090-99 SSP245** | 98 | 9 | 167 | 20 | 69 | 14 | 204 | 16 | 2464.9 | 780.8 | 611.9 | 229.2 |
| **2090-99 SSP585** | 93 | 10 | 157 | 18 | 64 | 15 | 197 | 19 | 1927.4 | 987.9 | 480.6 | 265.9 |

**Table S7. Mean values (based on six ESMs) of the phenological indices of prolates for three decades and two scenarios in the North-East Atlantic.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 215 | 4 | 250 | 5 | 83 | 11 | 215 | 4 | 371.5 | 100.8 | 99.2 | 39.8 |
| **2000-09** | 212 | 6 | 255 | 6 | 91 | 6 | 212 | 6 | 393.7 | 76.6 | 113.8 | 29.7 |
| **2090-99 SSP245** | 204 | 16 | 258 | 10 | 110 | 11 | 204 | 16 | 488.8 | 112.9 | 160.9 | 33.5 |
| **2090-99 SSP585** | 197 | 19 | 256 | 17 | 119 | 22 | 197 | 19 | 562.1 | 99.1 | 198.7 | 33.9 |

**Table S8. Mean values (based on six ESMs) of the phenological indices of dinoflagellates for three decades and two scenarios in the North-East Atlantic.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 215 | 4 | 254 | 9 | 83 | 21 | 215 | 4 | 1305.6 | 487.7 | 342.6 | 122.5 |
| **2000-09** | 212 | 6 | 258 | 9 | 96 | 34 | 212 | 6 | 1901.0 | 663.6 | 536.5 | 156.7 |
| **2090-99 SSP245** | 204 | 16 | 262 | 5 | 89 | 14 | 204 | 16 | 2441.7 | 1408.0 | 693.2 | 406.8 |
| **2090-99 SSP585** | 197 | 19 | 268 | 9 | 101 | 18 | 197 | 19 | 3035.0 | 1297.9 | 973.7 | 460.3 |

**Table S9. Mean values (based on six ESMs) of the phenological indices of oblates for three decades and two scenarios in the Labrador Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 100 | 7 | 191 | 13 | 91 | 6 | 148 | 15 | 3701.6 | 413.9 | 1149.1 | 190.2 |
| **2000-09** | 96 | 4 | 183 | 12 | 87 | 11 | 136 | 8 | 3770.1 | 769.1 | 1077.1 | 178.6 |
| **2090-99 SSP245** | 89 | 8 | 172 | 19 | 84 | 13 | 128 | 17 | 4055.5 | 545.4 | 1092.3 | 245.5 |
| **2090-99 SSP585** | 86 | 8 | 159 | 16 | 73 | 8 | 122 | 12 | 3184.4 | 727.2 | 777.9 | 245.5 |

**Table S10. Mean values (based on six ESMs) of the phenological indices of prolates for three decades and two scenarios in the Labrador Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 142 | 21 | 528 | 205 | 103 | 34 | 188 | 18 | 528.0 | 205.3 | 150.8 | 40.5 |
| **2000-09** | 145 | 19 | 494 | 158 | 97 | 31 | 187 | 23 | 493.7 | 157.9 | 140.2 | 35.9 |
| **2090-99 SSP245** | 126 | 16 | 546 | 94 | 107 | 36 | 177 | 32 | 545.7 | 94.5 | 180.3 | 27.2 |
| **2090-99 SSP585** | 137 | 34 | 543 | 48 | 115 | 35 | 187 | 37 | 543.5 | 47.9 | 199.3 | 35.1 |

**Table S11. Mean values (based on six ESMs) of the phenological indices of dinoflagellates for three decades and two scenarios in the Labrador Sea.** For each index (i.e. the initiation and termination of the Seasonal Reproductive Period (SRP), the seasonal duration, the day where the Maximum Annual Abundance (MAA) is reached, the MAA and the Integrated Mean Manual Abundance (IMAA)) the means over the decade (i.e. 1850-1859, 2000-2009 and 2090-2099 for the two SSP scenarios) and the six ESMs are displayed in the column on the left and the associated standard deviations (SD, i.e. the inter-model uncertainties) on the right.

| | SRP initiation (day of year) | | SRP termination (day of year) | | Seasonal duration (day) | | Day MAA (day of year) | | MAA (cells per CPR sample) | | IMAA (cells per CPR sample) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** | **mean** | **SD** |
| **1850-59** | 198 | 7 | 267 | 6 | 69 | 9 | 233 | 3 | 640.9 | 328.2 | 146.8 | 62.6 |
| **2000-09** | 196 | 7 | 263 | 8 | 66 | 12 | 229 | 5 | 910.0 | 428.4 | 204.5 | 66.9 |
| **2090-99 SSP245** | 199 | 2 | 264 | 4 | 66 | 5 | 230 | 2 | 1914.4 | 1125.1 | 412.1 | 252.2 |
| **2090-99 SSP585** | 188 | 9 | 270 | 9 | 82 | 18 | 228 | 3 | 2898.2 | 1052.0 | 738.8 | 319.3 |

**Supplementary Animations**

**Animation S1. Comparison between the abundance predicted by the models forced by the six ESMs (blue lines) and the abundance observed by the CPR (black lines) in the North Sea.** Note the distinct scale between the y-axis on the left and on the right.

**Animation S2. Comparison between the abundance predicted by the models forced by the six ESMs (blue lines) and the abundance observed by the CPR (black lines) in the North-East Atlantic.** Note the distinct scale between the y-axis on the left and on the right.

**Animation S3. Comparison between the abundance predicted by the models forced by the six ESMs (blue lines) and the abundance observed by the CPR (black lines) in the Labrador Sea.** Note the distinct scale between the y-axis on the left and on the right.

## 3.2. Climate variability and multi-decadal diatom abundance in the Northeast Atlantic

Martin Edwards, Gregory Beaugrand, Loïck Kléparski, Pierre Hélaouët and Philip C. Reid

*Published in Communication Earth and Environment. Supplementary Files are displayed in section 3.2.5. Article is available at https://doi.org/10.1038/s43247-022-00492-9*

*Abstract*

Diatoms are important contributors to marine primary production and the ocean carbon cycle. In the North Atlantic and its adjacent seas primary production is driven by diatoms that transfer a significant part of the produced energy to higher trophic levels and carbon to the deep ocean. Anthropogenic warming and climate variability will likely have important consequences for the productivity and spatial dynamics of these eukaryotic phytoplankton. Using multidecadal diatom abundance data (>60 years) for large areas of the North Atlantic and the North Sea, we show significant spatial and temporal correlations over these scales between diatoms and climate variability. In the NE Atlantic a general multidecadal trend is established where climate warming is increasing diatom populations in northerly systems, but decreasing populations in more southerly systems. We discover major phase shifts in diatom abundance synchronous with multi-decadal trends in Atlantic climate variability that occurred after the mid-1990s.

## 3.2.1. Introduction

At a planetary scale, the eukaryotic phytoplankton group collectively known as diatoms contribute approximately one-fifth of all of Earth's photosynthesis (Armbrust, 2009) and up to 30-40% of global marine primary production each year (Mann, 1999). They are particularly important in boreal systems where they dominate seasonal phytoplankton production and occupy key ecosystem roles (Rynearson, 2013; Smetacek, 1999). Diatoms and the role of nutrients and ocean biogeochemistry on their biological patterns have been at the forefront of intense studies (Allen, 2005; Boyd et al., 2010; Hátún et al., 2017). Excluding mathematical models (Bopp et al., 2005), least known, however, is the role of climate on in situ phytoplankton populations particularly at macro and multidecadal scales. How climate variability and diatom abundance interact at these scales is paramount to our understanding of the effects and prediction of the future impact of climate change on pelagic communities as well as having important future consequences to carbon sequestration (the biological carbon pump) and North Atlantic productivity.

Due to costs in maintaining consistent measurements over time and logistical complexities, multidecadal observations are rare. Here we used one of the few ecological time-series at this scale available derived from the Continuous Plankton Recorder (CPR) Survey to investigate spatial and temporal patterns of diatoms over 60 years in the NE Atlantic and adjacent seas. We relate these spatial and temporal patterns to the changing climatic conditions in the region, which include climate warming and wind patterns as well as an index that reflects Atlantic Multidecadal Variability (AMV).

## 3.2.2. Methods

### 3.2.2.1. Biological data

The Continuous Plankton Recorder (CPR) Survey is a long-term, sub-surface marine plankton monitoring programme consisting of a network of CPR transects towed monthly across the major geographical regions of the North Atlantic. It has been operating in the North Sea since 1931 with some standard routes existing with a virtually unbroken monthly coverage back to 1946 (Batten et al., 2003; Reid et al., 2003). The CPR survey is recognised as the longest sustained and geographically most extensive marine biological survey in the world (Edwards et al., 2010). The dataset comprises a uniquely large record of marine biodiversity covering ~1000 taxa over multidecadal periods. The survey determines the abundance and distribution of phytoplankton and zooplankton (including fish larvae) in our oceans and shelf seas. Using ships of opportunity from ~30 different shipping companies, it obtains samples at monthly intervals on ~50 trans-ocean routes. In this way the survey autonomously collects biological and physical data from ships covering ~20,000 km of the ocean per month, ranging from the Arctic to the Southern Ocean.

The CPR is a high-speed plankton recorder that is towed behind 'ships of opportunity' through the surface layer of the ocean (~10 m depth) (Warner and Hays, 1994). Water passes through the recorder and plankton are filtered by a slow-moving

silk (mesh size 270µm). A second layer of silk covers the first and both are reeled into a tank containing 4% formaldehyde. Upon returning to the laboratory, the silk is unwound and cut into sections corresponding to 10 nautical miles and approximately 3 $m^3$ of filtered sea water (Richardson et al., 2006). There are four separate stages of analysis carried out on each CPR sample, with each focusing on a different aspect of the plankton: (1) overall chlorophyll (the phytoplankton colour index; PCI); (2) larger phytoplankton cells (phytoplankton) in the case of this study we have just analysed diatoms which are all on the large size compared with smaller flagellates and coccolithophores that are also recorded; (3) smaller zooplankton (zooplankton "traverse"); and (4) larger zooplankton (zooplankton "eyecount"). The collection and analysis of CPR samples have been carried out using a consistent methodological approach, coupled with strict protocols and Quality Assurance procedures since 1958, making the CPR survey the longest continuous dataset of its kind in the world. Of the more than 200 phytoplankton species recorded by the Survey we analysed a sub-set of data consisting of over 49 diatom species/taxonomic entities. The selection of taxa was based on their frequency of occurrence, consistency of recorded taxa across the whole time-series (1958-2017). Rare species were removed during the selection to make the dataset statistically more robust. Practically, we chose diatoms that had an annual abundance > 0 for at least 50 years.

#### 3.2.2.2. Potential biases and limitations

The CPR survey currently records ~1000 plankton entities (many to species level) in routine taxonomic analysis with strict Quality Assurance protocols in operation for its sampling and plankton counting procedures since 1958. Due to the mesh size of CPR silks, many plankton species are only semi-quantitatively sampled owing to the small size of the organisms. In the case of phytoplankton there is thus a bias towards recording larger armoured flagellates and chain-forming diatoms and that smaller species abundance estimates from cell counts will probably be underestimated in relation to other water sampling methods. However, the proportion of the population that is retained by the CPR silk reflects the major changes in abundance, distribution and specific composition, i.e. the percentage retention is roughly constant within each species even with very small-celled species down to 10µm (Edwards et al., 2006). In certain circumstances clogging due to increased loading associated with mucilage and high densities of plankton may contribute to the number of small plankton captured. Despite these sporadic occurrences the CPR has still been shown to capture a consistent fraction of the in-situ abundance of the species assayed (Richardson et al., 2006). A similar potential under estimation of zooplankton abundances has recently been thoroughly statistically explored and found that while the CPR survey does underestimate abundance in some cases it provides a realistic quantification of both temporal (i.e. seasonal and diel scales) and spatial (i.e. regional to basin-scale) changes in zooplankton taxa (In this case for the species Calanus finmarchicus) (Helaouët et al., 2016). The study also showed that while the CPR sampling is restricted to surface waters ~10 m in depth the seasonal and diel patterns

of abundance of *C. finmarchicus* were positively correlated to patterns of abundance to a depth of 100 m (Helaouët et al., 2016).

While all sampling and analytical tools for measuring plankton have their own individual strengths and weaknesses, the CPR Survey has had perhaps the most well documented and transparent examination of its consistency and compatibility of any ecological time series over time (Batten et al., 2003; Edwards et al., 2010). Perhaps the most obvious limitation of CPR sampling is its underestimation of components of the plankton, e.g. large plankton like fish larvae and delicate gelatinous plankton. This has been well documented and users of the data are advised of the CPR's potential biases (Richardson et al., 2006). It is also widely recognised that all plankton sampling systems have their own limitations and nuances and all underestimate abundance to some degree and that the varying mechanisms are not always directly comparable (Owens et al., 2013). A detailed study has been conducted on flow rate and ship speed on CPR sampling (Jonas et al., 2004) given that the speed of the ships has, in some circumstances, increased since the 1960s, which may impact sample efficiencies. However, no significant correlation was found between the long-term changes in the speed of the ships and two commonly used indicators of plankton variability: Phytoplankton Colour and Total Copepods indices. This absence of relationship may indicate that the effect found is small in comparison with the influence of hydroclimatic forcing (Jonas et al., 2004). For further details on the technical background, methods, consistency, and comparability of CPR sampling, see (Batten et al., 2003).

#### 3.2.2.3. Hydroclimatic indices

The North Atlantic experiences climate variability on multidecadal scales often referred to as Atlantic Multidecadal Variability (AMV)(O'Reilly et al., 2019). The AMV has been operating over the last few centuries as an observed phenomenon and has also been identified at millennial timescales using paleoclimatic proxy data and models. However, identifying the mechanisms behind the AMV, be they externally forced or driven by internal dynamics or a combination of both is still widely debated (Mann et al., 2020; O'Reilly et al., 2019; Qin et al., 2020). Here we use the Atlantic Multidecadal Oscillation (AMO) index to reflect this low frequency variability in the North Atlantic. The AMO is a large-scale oceanic phenomenon and a source of a natural variability in the range of 0.4°C in the oceanic regions of the North Atlantic (Enfield et al., 2001). The AMO is also detected in paleoclimatic reconstructions with a cycle ranging from 60 to 100 years (Gray et al., 2004). The index is calculated from the Kaplan SST dataset which is updated monthly and represents an index of the North Atlantic temperatures; we used the unsmoothed version of the index (http://www.esrl.noaa.gov/psd/data/timeseries/AMO/). The AMO has been related to various biological shifts in the North Atlantic sector (M. Edwards et al., 2013).

The North Atlantic Oscillation (NAO) index is a basin scale alternation of atmospheric mass over the North Atlantic between the high pressures centred on the subtropical Atlantic and low pressures around Iceland and is a prominent mode of variability with important impacts from the polar to subtropical Atlantic and

surrounding landmasses. The index is calculated on the surface sea-level pressure difference between the Subtropical (Azores) High and the Subpolar Low. (https://www.ncdc.noaa.gov/teleconnections/nao/).

We also used Northern Hemisphere temperature anomalies (HadCRUT4). This climatic index was developed by the Climatic Research Unit (University of East Anglia) and the Hadley Centre (UK Met Office) (Jones et al., 1999). CRUTEM4 and HadSST3 are the land and ocean components of this dataset, respectively (https://crudata.uea.ac.uk/cru/data/temperature/). Global temperature anomalies have been frequently correlated with long-term changes in biological and ecological systems (Beaugrand et al., 2002b).

#### 3.2.2.4. Climatic data (wind and sea surface temperature)

We used wind intensity (W wind), the zonal (U wind, from west to east) and meridional (V wind, from south to north) components of the wind to investigate potential relationships between wind intensity and direction and diatom abundance. Wind intensity and direction influence the structure of the water column and are known to be important environmental parameters for diatoms (Alvain et al., 2005). Wind intensity and direction (monthly basis and spatial resolution of 2.5° latitude x 2.5° latitude) was obtained from the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR), who have run a reanalysis project using a state-of-the-art analysis/forecast system to perform data assimilation using past data from 1948 to the present (Kalnay et al., 1996). For the analysis we constructed a spatially gridded (2.5° latitude x 2.5° longitude) dataset for the northeast Atlantic region (40.5°N to 65.5°N and 30.5°W to 10.5°E) and the North Sea (52°N to 60°N and 2°W to 10°E), corresponding to the period 1958-2017.

Sea surface temperature (SST) was extracted from the Extended Reconstructed Sea Surface Temperature (ERSST)_V5 data-set, which is derived from a reanalysis of the most recently available International Comprehensive Ocean-Atmosphere Data Set (ICOADS) SST data. Improved statistical methods were used to produce a stable monthly reconstruction from relatively sparse data (Huang et al., 2017). For this analysis we constructed a spatially gridded (2° latitude x 2° longitude) annual average SST dataset for the Northern Hemisphere.

#### 3.2.2.5. Spatialised standardised principal component analysis on total diatoms in the northeast Atlantic

We calculated the sum of the abundance of all diatoms recorded semi-quantitatively by the CPR survey (264,837 samples). Because data from the CPR survey are not regularly spaced in time and space, we interpolated the data for total diatoms on a 2-monthly basis between 1958 and 2017 using an inverse squared distance method with a search radius of 150 km and a number of values for calculating the weighted mean ranging between 3 and 20 (Beaugrand et al., 2012; Lam, 1983). The boundary of the studied region ranged from 40.5°N to 65.5°N and from 30.5°W to 9.5°E for every 1° latitude x 1° longitude; inside the area, we kept geographical cells that had

at least 50 non-zero estimates of total diatom abundance. Annual means were subsequently calculated for every geographical cell. We kept cells that had at least 30 years of estimates of total diatom abundance.

A spatialised standardised Principal Component Analysis (PCA) was performed on the matrix 1066 geographical cells (26 latitudes x 41 longitudes) x 60 years. The matrix of standardised eigenvectors 1066 x 1066 geographical cells represented the linear correlations with the principal components that had a dimension of 60 years (Beaugrand et al., 2012). We applied a broken-stick model to test the significance of the different axes of the PCA using 10,000 simulations (Beaugrand et al., 2019a).

#### 3.2.2.6. Standardised principal component analysis on species of diatoms in the North Sea

We calculated the abundance of all diatom species (including total diatom abundance) occurring in the North Sea (52°N to 60°N and 2°W to 10°E) on a monthly basis for the period 1958-2017. Long-term monthly changes in total diatoms were examined first in relation to the hydro-climatic indices (AMO, NAO and NHT). Data were then annually averaged and a standardised principal component (PCA) was performed on the matrix 60 years x 49 species/taxa. This analysis therefore focussed on all species or taxa of diatoms, which was not investigated on the first PCA that jointly considers spatial and temporal changes in total diatom abundance. Here also, we applied a broken-stick model to test the significance of the different axes of the PCA using 10,000 simulations (Beaugrand et al., 2019a).

#### 3.2.2.7. Relationships between changes occurring in the northeast Atlantic and the North Sea

We correlated the long-term changes in the first principal components (PCs) originating from the PCA performed at the scale of the North Atlantic with changes in diatom abundance at the species or taxonomic level. For this analysis, we represented changes in diatom species in addition to a PC when the PC-species correlation was higher than 0.7 (i.e. percentage of explanation >49%).

#### 3.2.2.8. Maps of correlations

We calculated maps of correlation between W, U, V wind and SST and the first PC originating from the PCA performed on total diatoms at the northeast Atlantic scale. For this analysis, we applied an order-2 simple moving average to alleviate the influence of short-term higher frequency variability. This analysis was conducted for the time period 1958-2017.

#### 3.2.2.9. Correlation analyses between variables

Correlations were calculated to assess the strength of the relationships between total diatom abundance and climate in the North Sea for different seasons. The probability of significance of the coefficient of correlation, called $p_{ACF}$, was adjusted to correct for temporal autocorrelation

### 3.2.3. Results

We first started to examine long-term changes in total diatom abundance from January 1958 to December 2017 (see methods) (Warner and Hays, 1994) in the North Sea where sampling by the CPR survey has been most dense. For reasons of statistical robustness, for the analysis of phytoplankton at the macroscale level we used diatom abundance for the whole NE Atlantic but for individual species trends we only used data from the North Sea region. For this analysis, we included a total of 49 species/taxa. Annual diatom abundance (Figure 35) is dominated by the spring bloom with some further seasonal peaks occurring in autumn following the classic diatom pattern found in temperate and boreal marine environments (Smetacek, 1999). Peak diatom abundance occurs between March-May followed by a decline during the summer months while the water-column becomes stratified and is then followed by a floral renaissance in production in the autumn when the water column becomes vertically mixed again. Long-term trends in total diatom abundance show a sharp decadal decline during the 1960s which continued to a very low point in the late 1970s associated with anomalously cold and low-salinity hydro-climatic conditions before increasing again in the mid-1980s (Allen et al., 2020; Edwards et al., 2002). After recording low values in the early 1990s, diatom abundance substantially increased again in the mid-1990s and has continued to remain high since this phase shift in abundance (Figure 35, all trends statistically significant). While diatom phenology has remained fairly temporally static over time (Edwards and Richardson, 2004), in particular the timing of the spring bloom, there have been changes to the amplitude in abundance over the seasonal cycle (Figure 35). Particularly noticeable is the sharp decline in the autumnal bloom of diatoms that occurred during the 1960s. Over the last 20 years there appears to have been an increase in diatoms across the seasonal growth period with a return to increased autumnal growth and more abundance dispersed throughout the seasonal cycle including the summer period. In particular, the taxa *Thalassiosira* spp., *Rhizosolenia imbricata* and *Pseudo-nitzschia seriata complex* have increased their seasonal spread and autumnal bloom formation over the last few decades. The autumn bloom is less well studied than its vernal cousin but is a significant contributor to overall annual phytoplankton production (Wihsgott et al., 2019).

Subsequently, we examined individually species trends over time of the most common 49 species/taxa recorded by the CPR Survey (Figure 36). A Principal Component Analysis (PCA) was performed on the long-term trends of the diatom species/taxa at the scale of the North Sea. Where we determined the first (Figure 36a) and second (Figure 36b) principal components (PCs, 26.14 and 11.01% of the total variance, respectively) and their significance using a broken-stick test (see Methods). The first PC clearly shows a large decline in the 1960s with some notable low values in the late 1970s and early 1990s before a shift in the time-series after the mid-1990s. Examples of the highest species/taxa correlations with the first principal component were mainly neritic spring diatom species which included *Paralia sulcata* (the only tychopelagic species), *Bellerochea malleus*, *Biddulphia alternans*, *Thalassionema nitzschioides*, *Asterionellopsis glacialis*, *Pseudo-nitzschia delicatissima complex*,

*Pseudo-nitzschia seriata complex*, *Rhizosolenia imbricata* and *Rhizosolenia setigera*. The second PC (Figure 36b) shows a clear positive linear trend through time. Highest positive species/taxa correlations for this PC were with *Gyrosigma* spp. and *Rhaphoneis amphiceros*, and high negative correlations included *Chaetoceros (Phaeoceros)* spp., *Chaetoceros (Hyalochaete)* spp., and *Odontella regia*. Most of the species associated with a positive PC2 are autumnal bloom forming species that are found in shallow neritic waters of the North Sea.

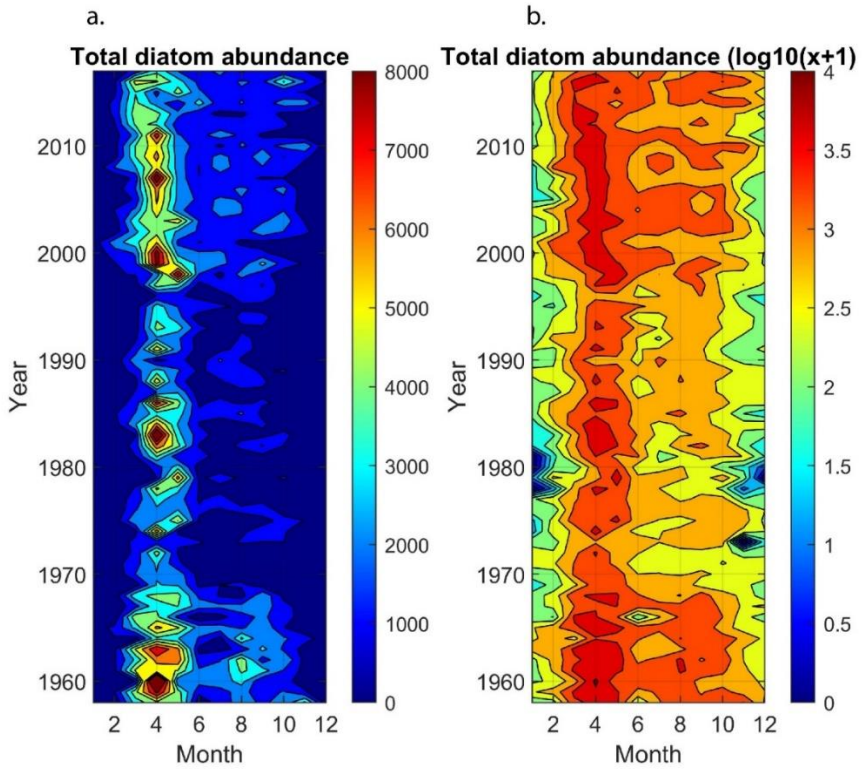



*Figure 35. Total monthly diatom abundance in the North Sea from 1958-2017. (a), total diatom (49 taxa) counts per month. (b), total diatom counts per month log transformed log10(x+1).*

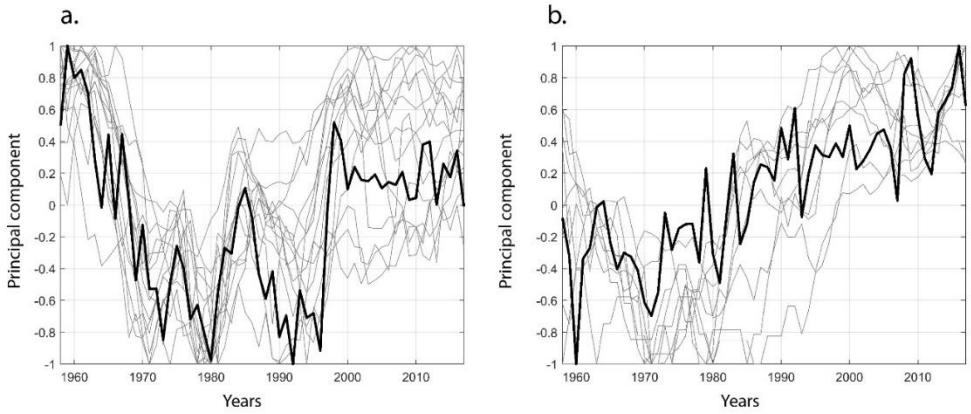



*Figure 36. Long term trends of all diatom species by using a Principal Component Analysis in the North Sea from 1958-2017. a, Principal Component 1 (bold line) and associated 49 species trends (light lines). Standardised abundance of all species related to the first PC (species with a correlation with PC1 higher than 0.7). b, Principal Component 2 (bold line) and associated 49 species trends (light lines). Standardised abundance of all species related to the second PC (species with a correlation with PC2 higher than 0.7). Species were selected on the basis of the examination of normalised eigenvectors.*

We then examined the relationships between long-term trends in diatom abundance and Atlantic Multidecadal Variability (AMV) at the scale of the Northeast (NE) Atlantic by using the Atlantic Multidecadal Oscillation (AMO) index and Northern Hemisphere Temperature (NHT) anomalies (see Methods). Patterns of changes in total diatom abundance were similar to those observed in the North Sea (Figure 36a versus Figure 37b). The positive normalised eigenvector 1 (Figure 37a) shows that the trend observed in the first PC (28.92% of the total variance) reflected changes that took place in most regions of the Northeast Atlantic (Figure 37a-b). The multidecadal trend between the diatom PC1 and the AMO index is remarkably similar with positive correlations found throughout the NE Atlantic ($r=0.71$, $p_{ACF}<0.05$; Figure 37b). Long-term trends in the AMO index and the diatom PC1 show a strong decline in the early part of the time-series, low values between the 1970s to the 1990s and a sharp shift in both time-series occurring in the mid-1990s (Figure 37b). High values have remained since this phase shift up to the current time-period examined.

Long-term changes in the second PC (8.46% of the total variance) reflects an opposition between a negative trend in total diatom abundance in the southern part of the NE Atlantic and a positive trend in the northeastern part (Figure 37c-d). Both time-series PC2 and NHT anomalies show a strong upward trend in values from the 1960s to the present ($r=0.93$, $p_{ACF}<0.05$). Because of the spatial patterns observed in the second normalised eigenvector, strong positive correlations between diatom abundance and the NHT are mainly spatially found in the shallow regions of the North Sea and off southern Norway (Figure 37c). In contrast, negative correlations between diatom abundance and NHT are geographically found in the Bay of Biscay.

To investigate the relationship between wind and temperature in detail we examined spatio-temporal correlations between diatom abundance and various components of wind and Sea Surface Temperature (SST) in the NE Atlantic. Figure 38 shows the correlation maps between total diatom abundance and SST (Figure 38a), U wind (zonal or westerlies) (Figure 38b), V wind (meridional or southerlies) (Figure 38c) and W wind (wind intensity) (Figure 38d) for the NE Atlantic over the sixty year time-period (see Methods). There is a positive correlation between temperature and diatoms found in the northern regions of the NE Atlantic such as the northern North Sea and the Icelandic basin region (Figure 38a). Negative correlations between total diatom abundance and SST are found in the Celtic Sea and Bay of Biscay (Figure 38a). Negative correlations are found between U wind and diatoms in the oceanic region south of Iceland (Figure 38b). Positive correlations are found between V wind and total diatom abundance in the central North Sea and in some central oceanic regions of the NE Atlantic (Figure 38c). Finally, the overall wind intensity is negatively correlated with diatom abundance with this signal mainly found in the central North Sea. To examine these temporal correlations in more detail we focused on correlation coefficients between these various climatic indices and the total diatom abundance found in the North Sea between different seasons. Very low correlations (excluding the AMO) are found between all climate variables and spring diatom abundance. A significant positive correlation is found between summer diatoms and the AMO. During autumn

the abundance of diatoms in the North Sea are significantly correlated with temperature (AMO and SST) and negatively correlated with U wind and wind intensity.

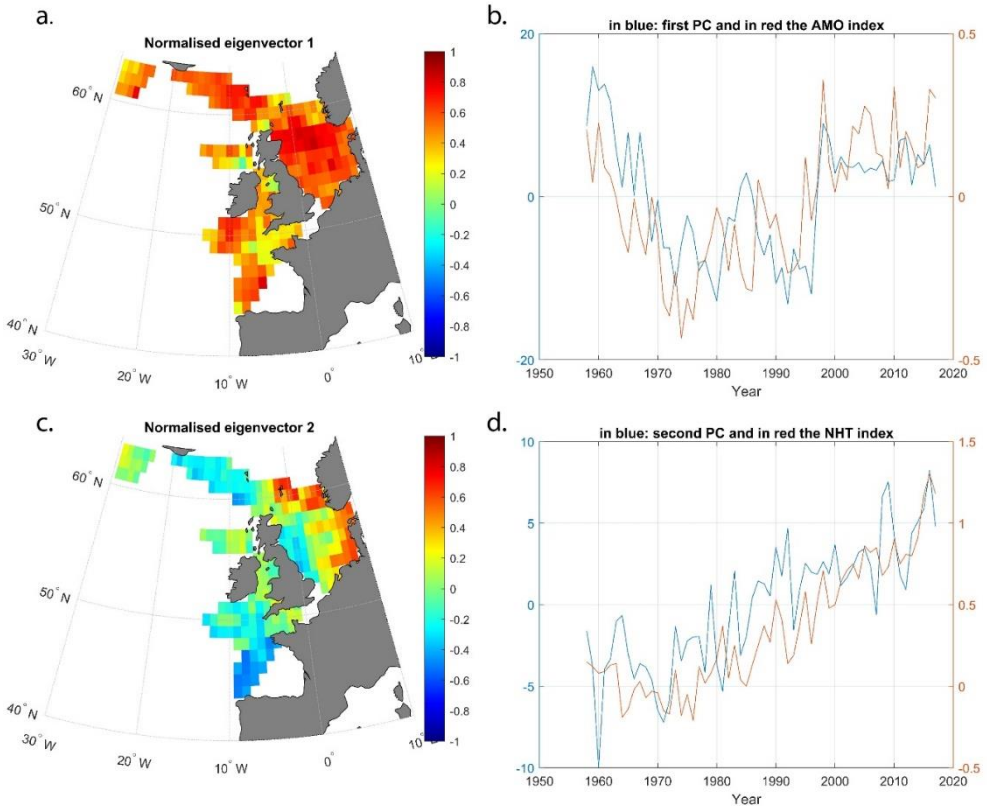



*Figure 37. Long-term (1958-2017) changes in diatom abundance as revealed by a PCA performed at the scale of the Northeast Atlantic and relationships with the Atlantic Multidecadal Oscillation (AMO) and Northern Hemisphere Temperatures (NHT). (a), First normalised eigenvector showing the spatial pattern of correlations between spatial changes in total diatom abundance and the first principal component (PC). (b), long-term relationship between the first PC of diatom abundance (blue) and the AMO index (red). (c), Second normalised eigenvector showing the spatial pattern of correlations between spatial changes in total diatom abundance and NHT. (d), long-term relationship between the second PC of diatom abundance (blue) and NHT (red).*

## 3.2.4. Discussion

In this study, we have established a strong relationship between multidecadal variability in the Atlantic and the abundance of diatoms in the NE Atlantic and have also established a significant relationship between trends in neritic diatoms in the North Sea and Northern Hemisphere Temperature. In the temperate and boreal North Atlantic where CPR sampling is most prolific there have been a significant change in climate due to natural variability and global warming over the duration of this 60-year study. At the global level climate warming of pelagic surface waters is thought to increase the density gradient between the surface layer and the underlying nutrient-rich waters, in turn enhancing the stability of the water-column and hence restricting nutrient availability. Access to the principal nutrient (nitrate) that limits phytoplankton growth has been found to be negatively related to temperatures globally (Kamykowski and Zentara, 2005, 1986). In another global analysis using satellite derived chlorophyll data it has been shown that a strong inverse relationship exists between Sea Surface Temperatures (SST) and chlorophyll concentration (López-Urrutia et al., 2006). In

terms of feedback mechanisms on Earth's climate, it is speculated that phytoplankton driven biological carbon pumps will be less efficient in a warmer world due to changes mainly in phytoplankton composition. These changes favour small flagellates (Bopp et al., 2005) to the detriment of diatoms, and less overall nutrient mixing due to increased stratification. It is also predicted that warmer temperatures would shift the metabolic balance between production and respiration in the world's oceans towards an increase in respiration thus reducing the capacity of the oceans to capture CO2 (López-Urrutia et al., 2006). However, in the case of our study, the manifestation of climate variability in the NE Atlantic is less than homogenous with distinct spatial and temporal patterns emerging with diatom abundance increasing in some areas and declining in others. In the NE Atlantic, a prevalent multidecadal trend has been established where climate warming is increasing diatom populations in more northerly regions (e.g. Icelandic Basin), but deceasing populations in more southerly regions (e.g. Bay of Biscay).

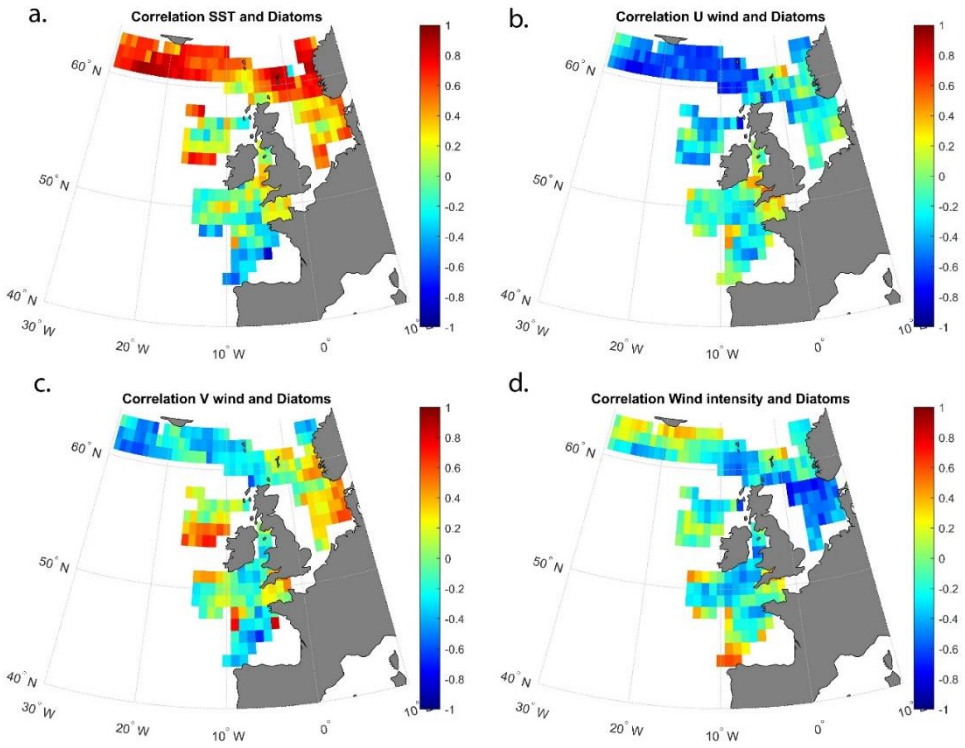



*Figure 38. Long-term spatial and temporal (1958-2017) correlations between total diatom abundance and climate variables in the NE Atlantic. (a), diatoms and SST, (b), diatoms and U wind (westerlies), (c), diatoms and V wind (southerlies), d, diatoms and W wind (intensity).*

Over the whole NE Atlantic there has been an increase in phytoplankton biomass during spring and autumn (where diatoms dominate) with increasing temperatures in cooler regions but a decrease in phytoplankton biomass in warmer regions (Richardson and Schoeman, 2004). This is possibly a trade-off between increased phytoplankton metabolic rates caused by warming temperatures in cooler regions but conversely a decrease in nutrient supply in warmer regions. This pattern is probably only applicable to oceanic regions of the NE Atlantic such as the Bay of Biscay and the Icelandic Basin and not as important to diatoms in well-mixed neritic waters

such as those found in regions of the North Sea (as nutrient replenishment is not cut off). Our study also shows that autumnal diatom abundance is positively correlated with SST and NHT. Regional climate warming in some areas of the North Sea has been linked to an increase in certain diatoms that are associated with Harmful Algal Blooms (HABs), in particular the taxa *Pseudo-nitzschia seriata complex* (Edwards et al., 2006; Hinder et al., 2012). Diatom growth in such well mixed areas may be enhanced by temperature as these regions are not inhibited by stratification and hence nutrient availability (see Figure 37b).

Although it is difficult to tease out absolute mechanisms between climate and phytoplankton due to the nature of macroscale correlative studies such as this analysis, our study does shows that the AMO has a very strong influence on diatoms in the NE Atlantic at the multidecadal scale. During a negative phase of the AMO, total diatom abundance is reduced despite the fact that temperature decreases and that wind increases. We observe the opposite effect during a warm phase of the AMO and show that the influence of the AMO mainly manifests itself during the end of summer and in autumn for the North Sea regions. Higher temperatures and less wind mixing during the autumn months could enhance diatom growth in these well mixed regions. It also appears that since the positive phase shift in the AMO beginning the mid-1990s to the present time has substantially tightened the relationship between diatom abundance and climate warming (AMO and NHT). These findings are important because mechanisms will differ between different regional areas of the world with varying hydro-climatic regimes. This study illustrates the complexity of the responses of phytoplankton and their functional types to climate variability as regional nuances will not always be represented well by homogenised global models.

*Acknowledgements*

An international consortium of government agencies from Canada, Norway, UK and USA funds the Continuous Plankton Recorder Survey. M.E was supported by NERC grants UK Changing Arctic programme DIAPOD and UK NERC CLASS. We would like to thank the owners and crews that have towed the CPRs on a voluntary basis for over 80 years contributing to one of the world's largest and longest ongoing ecological experiments. Without these early pioneers of citizen science and broad-scale volunteer monitoring projects this unique ecological dataset would never have been financially or logistically viable.

## 3.2.5. Supplementary Files

**Supplementary Tables.**

**Table S1. Correlations coefficients between the 49 taxa of diatoms examined and their Principal Components (PC1-PC5).** Correlations greater than 0.5 and -0.5 highlighted in bold.

| Taxa | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| *Paralia sulcata* | **0.86** | -0.01 | -0.01 | 0.10 | 0.19 |
| *Skeletonema costatum* | 0.47 | -0.36 | -0.46 | 0.13 | -0.31 |
| *Thalassiosira* spp. | 0.48 | 0.04 | -0.13 | -0.12 | **-0.51** |
| *Dactyliosolen antarcticus* | 0.09 | -0.30 | -0.12 | -0.08 | 0.07 |
| *Rhizosolenia styliformis* | 0.16 | -0.31 | 0.03 | 0.01 | 0.06 |
| *Rhizosolenia hebetata semispina* | 0.10 | -0.33 | -0.13 | 0.44 | -0.14 |
| *Chaetoceros(Hyalochaete)* spp. | 0.42 | **-0.76** | 0.02 | -0.06 | -0.08 |
| *Chaetoceros(Phaeoceros)* spp. | 0.32 | **-0.69** | -0.04 | 0.19 | 0.18 |
| *Odontella sinensis* | **0.65** | 0.13 | 0.08 | -0.33 | 0.30 |
| *Thalassiothrix longissima* | -0.07 | -0.01 | **0.52** | -0.08 | -0.26 |
| *Thalassionema nitzschioides* | **0.61** | -0.22 | 0.11 | -0.05 | 0.24 |
| *Bacteriastrum* spp. | **0.76** | 0.17 | 0.00 | -0.18 | -0.01 |
| *Bellerochea malleus* | **0.73** | 0.17 | 0.05 | 0.22 | -0.01 |
| *Biddulphia alternans* | **0.53** | 0.45 | -0.17 | -0.10 | 0.35 |
| *Odontella aurita* | **0.51** | -0.10 | 0.10 | 0.02 | -0.06 |
| *Odontella granulata* | -0.03 | 0.18 | 0.27 | 0.37 | -0.03 |
| *Odontella regia* | 0.08 | **-0.69** | 0.34 | 0.33 | 0.14 |
| *Odontella rhombus* | 0.28 | -0.33 | 0.36 | 0.40 | 0.00 |
| *Cerataulina pelagica* | 0.39 | -0.30 | 0.02 | -0.43 | -0.16 |
| *Coscinodiscus concinnus* | 0.43 | -0.28 | 0.23 | -0.26 | 0.09 |
| *Coscinodiscus* spp. (Unidentified) | **0.54** | **0.61** | 0.04 | 0.26 | 0.14 |
| *Ditylum brightwellii* | **0.74** | 0.14 | -0.30 | -0.01 | -0.04 |
| *Eucampia zodiacus* | 0.23 | -0.06 | **0.50** | -0.26 | 0.22 |
| *Fragilaria* spp. | **0.57** | -0.19 | -0.12 | 0.40 | 0.05 |
| *Guinardia flaccida* | 0.32 | 0.00 | 0.34 | -0.09 | 0.32 |
| *Gyrosigma* spp. | 0.18 | **0.55** | -0.06 | 0.27 | -0.05 |
| *Leptocylindrus danicus* | 0.50 | -0.28 | 0.22 | 0.10 | -0.37 |
| *Navicula* spp. | -0.02 | -0.45 | 0.39 | 0.40 | 0.17 |
| *Cylindrotheca closterium* | **0.69** | -0.22 | -0.03 | 0.06 | -0.17 |
| *Rhaphoneis amphiceros* | **0.56** | **0.72** | 0.17 | 0.17 | -0.01 |

| Taxa | PC1 | PC2 | PC3 | PC4 | PC5 |
|---|---|---|---|---|---|
| *Rhizosolenia setigera* | **0.73** | 0.39 | -0.08 | -0.18 | -0.09 |
| *Stephanopyxis* spp. | 0.18 | 0.45 | 0.28 | -0.01 | -0.06 |
| *Nitzschia* spp. (Unidentified) | **0.67** | 0.21 | -0.27 | 0.18 | 0.18 |
| *Odontella mobiliensis* | 0.41 | 0.12 | -0.24 | 0.09 | 0.49 |
| *Proboscia alata* | 0.38 | **-0.55** | 0.09 | -0.32 | 0.15 |
| *Leptocylindrus mediterraneus* | 0.46 | -0.28 | -0.14 | -0.28 | -0.35 |
| *Proboscia inermis* | 0.02 | 0.12 | 0.49 | -0.08 | -0.17 |
| *Asterionellopsis glacialis* | **0.72** | -0.17 | 0.03 | 0.20 | -0.08 |
| *Pseudo-nitzschia delicatissima complex* | **0.70** | -0.14 | -0.30 | 0.07 | -0.15 |
| *Pseudo-nitzschia seriata complex* | **0.88** | 0.01 | -0.02 | -0.03 | 0.01 |
| *Guinardia delicatula* | **0.56** | 0.35 | 0.11 | -0.16 | -0.29 |
| *Dactyliosolen fragilissimus* | **0.56** | 0.33 | 0.23 | 0.04 | -0.46 |
| *Guinardia striata* | **0.63** | -0.04 | 0.25 | -0.39 | 0.21 |
| *Helicotheca tamesis* | 0.19 | 0.09 | 0.43 | 0.13 | -0.44 |
| *Lauderia annulata* | **0.64** | 0.03 | -0.03 | 0.05 | 0.18 |
| *Bacillaria paxillifera* | **0.68** | 0.01 | 0.06 | 0.39 | 0.02 |
| *Corethron hystrix* | 0.29 | -0.23 | **-0.59** | 0.06 | -0.10 |
| *Proboscia indica* | 0.08 | -0.24 | -0.13 | -0.47 | -0.11 |
| *Rhizosolenia imbricata* | **0.80** | -0.06 | 0.15 | -0.21 | 0.12 |

**Table S2: Correlations coefficients between seasonal diatom abundance (spring, summer and autumn) and climatic variables.** *: correlation significantly different to 0 ($p_{ACF}$<0.05) after a test accounting for temporal autocorrelation. Atlantic Multidecadal Oscillation (AMO), North Atlantic Oscillation (NAO), Sea Surface Temperature (SST).

| | Spring diatoms | Summer diatoms | Autumn diatoms |
|---|---|---|---|
| **AMO** | 0.3396 | **0.5373*** | **0.4949*** |
| **NAO** | -0.19 | -0.05 | -0.15 |
| **SST** | 0 | 0.13 | **0.5*** |
| **W wind (intensity)** | -0.22 | 0.21 | **-0.44*** |
| **U wind (westerlies)** | -0.08 | 0.29 | **-0.43*** |
| **V wind (southerlies)** | 0.14 | 0.25 | 0.08 |

# 4. General discussion

## 4.1. Main results

The results of this work have provided evidence that the concept of the ecological niche is central to our understanding of the spatio-temporal distribution of plankton. The corollary of these results is that the concept can be used to anticipate how species and communities will be altered in the context of anthropogenic climate change. The METAL theory posits that the niche-environment interaction is fundamental to our understanding of the arrangement of species, communities and biodiversity in space and time (Beaugrand, 2015a). According to this theory, a species' niche is a summary of the expression of the genome that determines enzymatic reactions, norms of reaction, physiology as well as behavioural and phenotypic plasticity (Beaugrand and Kirby, 2018). Until now however, the METAL theory has not been applied to model (i) species phenology and (ii) to investigate APS or changes in community composition in space. Figure 39 displays a synthesis of the main results of this PhD. Plankton biodiversity is composed of many communities and coenoclines in the North Atlantic Ocean and its adjacent seas. Using data from the CPR survey, we have shown that a species' niche explains the alteration of the biocoenoses in space (Figure 1). We have also provided evidence that the niche explains well a species' phenology as well as the annual rearrangement of communities during APS, therefore the alteration of the biocoenoses in time. Species' traits are also fundamental and the niche summarises the different traits of a species, which are the result of genetic expression. Various life strategies and traits have been developed by phytoplankton, which in turn affect their niches and confer to each species a phenology and a distinct place in the sequence of APS. For example, oblate and prolate diatoms and dinoflagellates exhibit distinct phenologies; oblates occur mainly during spring and to a lesser extent in autumn whereas prolates and dinoflagellates dominate during period of high-water column stratification. The understanding of the key role played by the niche allowed us to investigate past, present and future long-term changes (1850-2100) in phenology and abundance using the PhenoShift model. Although the intensity of the changes varies according to greenhouse gas emission scenarios, we found coherent patterns of changes in the three key regions of the North Atlantic we investigated (North Sea, North-East Atlantic and Labrador Sea); oblates are predicted to decline while prolates and dinoflagellates are expected to become more prominent, which may affect ecosystem trophodynamics and carbon export.

## 4.2. Environmental control of species succession

In the first chapter, we showed that the METAL-based model we now call PhenoShift can well reconstruct APS when three-dimensional gaussian niches with as environmental dimensions SST, nutrient concentrations (e.g. nitrate) and PAR are used (Caracciolo et al., 2021). These three variables are known to influence phytoplankton phenology substantially (Boyce et al., 2017; Racault et al., 2012).

Temperature controls metabolic processes (e.g. enzymatic reaction), reproduction, locomotion, feeding and species interaction (Beaugrand, 2015a; Brown et al., 2004; Kirby and Beaugrand, 2009) but also the structure of the water column and therefore the supply of nutrients in the euphotic zone, which are essential for organic matter production through photosynthesis (Boyce et al., 2017; Falkowski and Oliver, 2007; Racault et al., 2012). PAR influences photosynthesis and modulates phenology; the parameter is also responsible for bloom initiation in many regions of the North Atlantic (Boyce et al., 2017; Racault et al., 2012). Nutrients are a fundamental constituent of proteins and nucleic acids and are essential for the synthesis of organic matter (Miller, 2004).

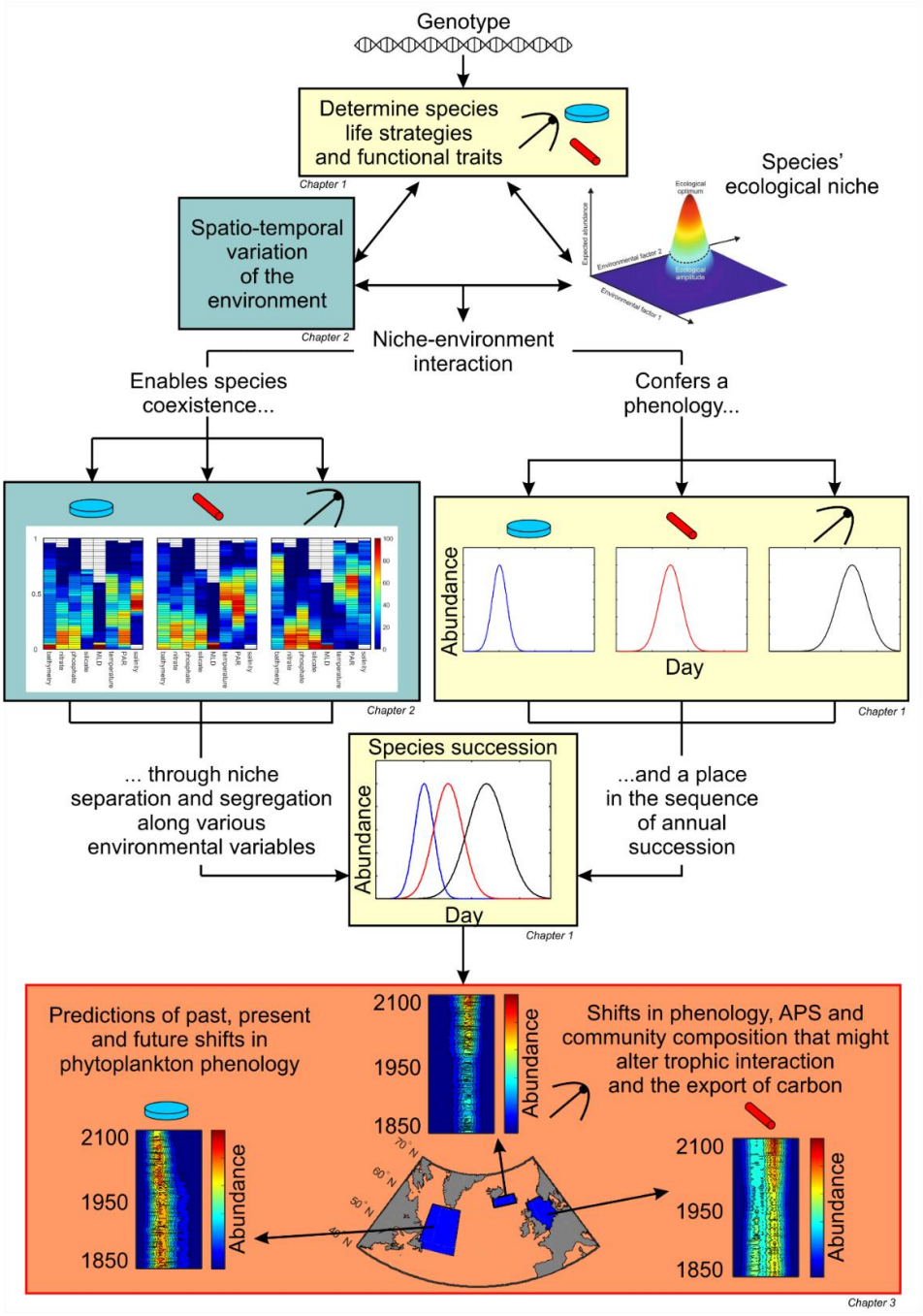



*Figure 39. Conceptual scheme summarizing the main results of this PhD.*

Our results enable us to better understand the spring bloom (Caracciolo et al., 2021). They have confirmed what was previously shown by mesocosm and modelling experiments, i.e. that PAR and temperature are important for spring bloom initiation in high latitudes (Lewandowska and Sommer, 2010; Sommer et al., 2012). Our work also suggests that nutrients depletion caused the end of the spring bloom. Modelling studies, as the ones conducted by the Plankton Ecology Group (PEG) on freshwater lakes, considered that grazing pressure (and not nutrients) is the main cause of the end of the spring bloom (Sommer et al., 2012), which agrees with the framework proposed by the Disturbance-Recovery hypotheses (Behrenfeld, 2010). However, our approach correctly reconstructs phytoplankton succession although it does not consider species interaction. Therefore, bottom-up control might be more important than top-down control in spring bloom termination, which is confirmed by observations conducted in the western English Channel where peaks in phytoplankton biomass are encapsulated by zooplankton peaks (Atkinson et al., 2018).

## 4.3. Phytoplankton life strategy, niche and phenology

It has been suggested that phytoplankton life forms have evolved in response to a changing environment (Margalef, 1978). Our results suggest that life forms (e.g. cell shape) affect a species' niche and confer to a species a distinct phenology and therefore a place in the sequence of APS (Caracciolo et al., 2021; Kléparski et al., 2022b). As all organisms, diatom and dinoflagellate strategies (i.e. the values taken by various functional traits and the associated trade-off between them) have been progressively developed in response to the interaction between their genome and the environment, an interaction that has also affected continuously their niche. Dinoflagellates have adapted themselves to a stratified low-nutrient environment and have developed, through evolution, a flagellum that allows them to actively move within the water column to harvest nutrients. Furthermore, their mixotrophic diet enables them to switch to heterotrophy when nutrients are absent and photosynthesis becomes impossible (Litchman and Klausmeier, 2008; Miller, 2004).

Diatoms have developed different cell shapes enabling them to occur in two distinct environments, i.e. in the low-stratified high-nutrients waters in spring and autumn (i.e. oblate cells) and the stratified low-nutrient waters found from late spring to the beginning of autumn (i.e. prolate cells) (Figure 40). We have provided evidence that diatom adaptation to stratification is achieved through an elongation of their cell shape, which enables them to diminish their sinking velocity without impacting their ability to absorb nutrients. Indeed, elongated cells have a slower sinking velocity in low turbulence waters (Clifton et al., 2018; Padisak et al., 2003), allowing them to remain in the well-lit upper part of the water column, but they also experiment a higher-nutrient flux because of a low surface/volume ratio (Karp-Boss and Boss, 2016; Pahlow et al., 1997). It is known that phytoplanktonic cells minimise this ratio to maximise the absorption of nutrients by increasing their surface while optimizing cell metabolism, by diminishing cell volume (Karp-Boss and Boss, 2016). However, in vertical shears, elongated cells tend to align with the flow and to aggregate in downwelling region,

which increases their sedimentation (Clifton et al., 2018). Therefore, changes in the dominance of diatom shape throughout the year are explained by the modifications in the environmental regime and more specifically by the magnitude of turbulence, shears direction and nutrient supply, which select species that are the most adapted. A summary of the relationships between diatom cell shapes and the fluctuations of the environment is displayed in Figure 40. The figure clearly shows that both kinds of diatoms are well distinguished by the temperature, water column structure and nutrients. It should be noted that other mechanisms independent of cell shape have been developed by the diatoms to regulate their vertical position in the water column. For example, some species are able to regulate the solute content of their protoplasm to modify their cell density and achieve large migrations from the deep chlorophyll maximum to the surface, in order to actively mine nutrients and light (Kemp and Villareal, 2018). The diatom frustule is also important in regulating buoyancy and some authors have suggested that this hard mineral shell has evolved as a ballast to regulate sinking rate, enabling them to sink toward the deep ocean when environmental conditions at surface become unsuitable and/or remove cells that have been contaminated by parasites from the healthy population (Raven and Waite, 2004; Smetacek, 1985).

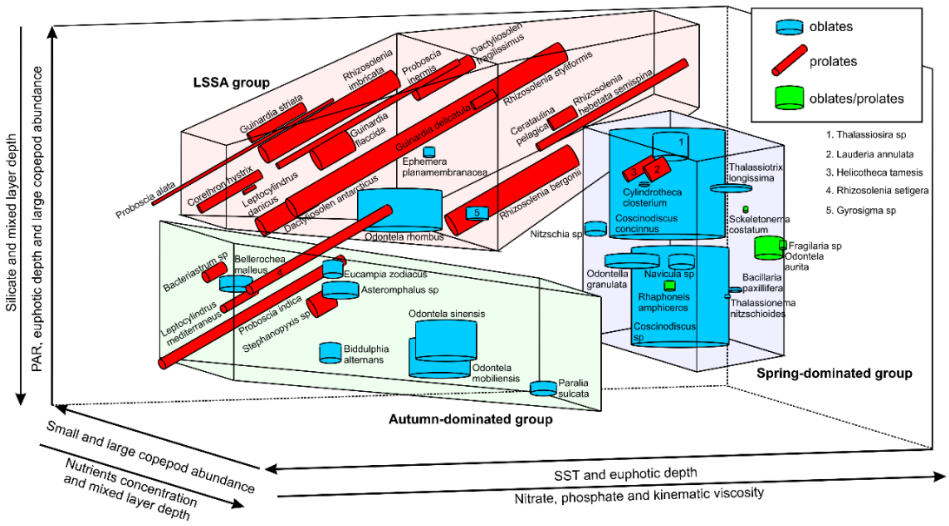



*Figure 40. Summary of the relationships between diatom shapes and the environment. Cell shapes are represented by cylinders scaled with species height and diameter (see Chapter 1). Oblates are in blue, prolates in red and oblates/prolates in green. Species names are those recorded in the CPR dataset. Blue area delineates the space occupied by the spring-dominated group, green area the space occupied by the autumn-dominated group and red area the space occupied by the Late Spring, Summer, Autumn group (LSSA). (Adapted from Kléparski et al., 2022a).*

Other hypotheses have also been proposed to explain morphological changes in diatom community. The relationship we found between zooplankton abundance and cell elongation (Kléparski et al., 2022b), suggests that grazing pressure might impact cell shape. Grazing by zooplankton is known to modify the frequency distribution of some phytoplanktonic shapes, because of a behavioural or mechanical selection of their preys (Sournia, 1982). Spherical cells are thus more easily ingested

by copepods than elongated cells, which need to be reoriented (Karp-Boss and Boss, 2016), but cell reorientation by copepods has been shown to be possible under low turbulence conditions (Visser and Jonsson, 2000). Therefore, the dominance of elongated shapes during summer, when stratification is important and turbulence low (Figure 40), cannot be explained by an effect of grazing pressure. However, other defence mechanisms have been developed by diatoms, such as silica shell or colony formation (Pančić and Kiørboe, 2018), which might explain why we found an effect of predation in the APS of the Eastern English Channel (Breton et al., 2021).

It has also been suggested that changes in shape dominance can be a strategy to optimize light capture, the chloroplasts being better exposed in an elongated cell (Naselli-Flores and Barone, 2011). However, the diatoms have developed various strategies to optimize their light capture without altering cell shape. For example, their frustule can be assimilated to a biological photonic crystal, which modulates incoming incident light inside the living cell, resulting in a redistribution, optimization and filtering of incident sunlight to the photosynthetic apparatuses in the chloroplasts (De Tommasi, 2016; Fuhrmann et al., 2004; Goessling et al., 2018; Noyes et al., 2008). Diatoms have also developed specific light-harvesting systems, referred to as fucoxanthin-chlorophyll proteins (FCPs) to collect and convert light energy, with a particularly high efficiency of energy transduction and photoprotection due to a high concentration of carotenoids (Büchel, 2019). They also have segregated photosynthetic complex (photosystems I and II) located inside the loosely stacked thylakoid membranes of their chloroplasts. The separation of the two photosystems enabling an improved light utilisation and their interconnection allowing a rapid equilibration of electron carriers that optimizes photosynthesis (Flori et al., 2017). Finally, most of centric diatoms are polyplastidic and can adjust the location of their chloroplasts inside the cell to adapt to different light intensities (Mann, 1996). Hence, all those mechanisms enable diatoms to adapt to variable light intensity, without the need to alter their cell shapes.

## 4.4. Niche differentiation, species coexistence, and community rearrangement in space and time

It is well-known that diatoms and dinoflagellates have distinct environmental requirements as a group (Irwin et al., 2012). In the second chapter, we have demonstrated, by means of a new method proposed as part of this PhD (i.e. the species chromatogram), that the difference also occurs at a species level and that species of the two groups have themselves distinct niches (Kléparski et al., 2022c). In other words, each species has a unique combination of environmental conditions where its performances in term of growth and reproduction are optimal. Each species therefore covers a part of the multidimensional environmental space and the spatio-temporal diversity of habitats enable their coexistence because of their segregation along various environmental gradients (Kléparski et al., 2022b). A similar mechanism is thought to enable species coexistence in terrestrial plants (Silvertown, 2004; Tilman, 1999). We also found that niche differentiation is highest in the optimal part of a

species' niche, alleviating competition and preventing competitive exclusion (Kléparski et al., 2022c). Therefore, when a species is in the optimal part of its niche, high biomass can be generated (Helaouët and Beaugrand, 2009) and the succession of different environmental regimes throughout the year generates a succession of species with different life strategies and niches. Hence, APS can be assimilated to a temporal rearrangement of communities (i.e. coenoclines or gradients of biocoenoses). Interestingly, we noticed that plankton biodiversity is also organised in spatial coenoclines, resulting from the effect of oceanic circulation, topography and temperature (Kléparski et al., 2021). Coenoclines may occur everywhere in the ocean but there are some places where they are more likely, e.g. boundaries between biomes or provinces and boundaries between the continental selves and the ocean (Kléparski et al., 2021).

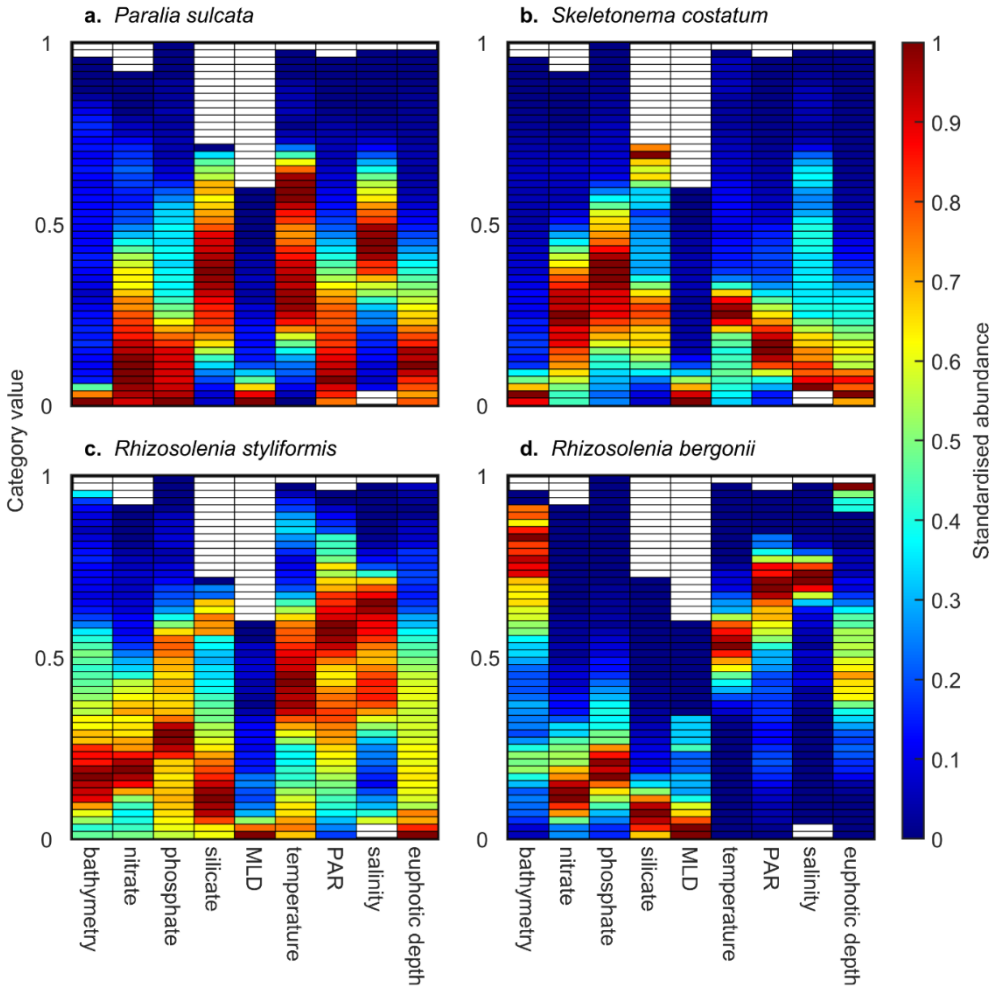



*Figure 41. Species chromatograms of four diatom species. Species chromatogram of (a) Paralia sulcata, (b) Skeletonema costatum, (c) Rhizosolenia styliformis and (d) Rhizosolenia bergonii. In a-d, each column represents the species abundance along nine environmental dimensions (i.e. bathymetry, nitrate, phosphate and silicate concentration, MLD, temperature, PAR, salinity and euphotic depth). Species abundance in each category (colour in the cells) was assessed by estimating the abundance of the 5% of the highest values available in a category if at least 20 CPR samples were available in that category. The Y axis corresponds to the 50 categories standardised between 0 and 1. This axis represents all values taken by an environmental variable between 0 and 1 from the lowest (bottom category) to the highest (top category). Colours denote the species abundance standardised between 0 and 1 in each category. High abundance values are in red and low values in blue.*

Niche differentiation and environmental gradients are important to understand the dynamics of plankton biodiversity in space and time and this PhD has provided evidence that niche differentiation and environmental gradients conspire together to ensure species coexistence. Separation along a single environmental dimension of the niche is sufficient to allow species coexistence and reflects evolutionary constrains that select specific traits (Brown, 1984). For example, it has been shown that two phytoplanktonic species can coexist while having similar requirements for nutrients but a differential use of the light spectrum (Burson et al., 2019). The visual examinations of the species chromatograms for diatoms also confirmed this hypothesis: two species can have a similar chromatogram along a given dimension but can diverge along another (Kléparski and Beaugrand, 2022). For example, *Paralia sulcata* and *Skeletonema costatum*, which both occur in the North Sea during the same season (Barnard et al., 2004; Caracciolo et al., 2021), have similar requirements for bathymetry and euphotic depth, but diverge along the nutrient, PAR and salinity dimensions (Figure 41a-b). In the same way, *Rhizosolenia styliformis* and *R. bergonii*, which both occur in the same region of the North Atlantic Ocean but not at the same time (Barnard et al., 2004; Caracciolo et al., 2021) have overlapping chromatograms along the nutrient and temperature dimensions but not along the bathymetry and the euphotic depth dimensions (Figure 41c-d).

Understanding how plankton species coexist has always fascinated ecologists (Bracco et al., 2000; Huisman and Weissing, 1999; Hutchinson, 1961; Petersen, 1975). Species coexistence is promoted by environmental variability through a mechanism known as the “storage effect”, which has three requirements (Adler et al., 2006; Chesson, 2000; Descamps-Julien and Gonzalez, 2005). The first is that the species of a given community have to develop strategies to survive during unfavourable environmental conditions, as the cysts and spores that have been developed by diatoms and dinoflagellates (McMinn and Martin, 2013; McQuoid and Hobson, 1996). The frustule of diatoms is even though to have evolved in order to serve as a ballast to increase the sinking speed when conditions become unsuitable in surface (Raven and Waite, 2004; Smetacek, 1985). The second is that these species should exhibit distinct responses to environmental variability, which is achieved for plankton through niche differentiation. The last is that environmental variability affects competition, meaning that the niche-environment interaction modulates competition, i.e. a species is more resistant to competition when its environment is more in adequation with its niche and inversely (Beaugrand and Kirby, 2018). Hence, the “storage effect” also applied to plankton.

## 4.5. Phenology and APS under climate change

As we said earlier, contemporary patterns in species succession we observed in the North Atlantic consist of the dominance of oblate diatoms in spring, followed by prolates and dinoflagellates during the stratification period and sometimes a last peak of oblates in autumn (Caracciolo et al., 2021; Kléparski et al., 2022b; McQuatters-Gollop et al., 2007). Using the PhenoShift model in three key areas of the North Atlantic

(i.e. Labrador Sea, North-East Atlantic and North Sea; Figure 42), we showed that species phenology and succession will be altered, especially if global warming continues unabated. Figure 4 synthetises the long-term changes in abundance predicted by METAL between 1850 and 2100, for oblates, prolates and dinoflagellates in the three regions and for two greenhouse gas emission scenarios, i.e. a medium (SSP245) and a high (SSP585) emissions scenario. A decline in oblate abundance combined with a narrowing of their phenology is predicted with both scenarios in the three regions (Figure 4b, e, h, k, n and q), associated to an expansion of prolate and dinoflagellate phenology and abundance (Figure 4c-d, f-g, i-j, l-m, o-p and r-s). The phenological shifts predicted by the PhenoShift model were relatively moderate for 1850-1950 in the three regions of the North Atlantic (Figure 4). Major changes started to occur from 1950 onwards (Figure 42) when the rate of ocean warming increased, which led to an increase in water column stratification (IPCC, 2019).

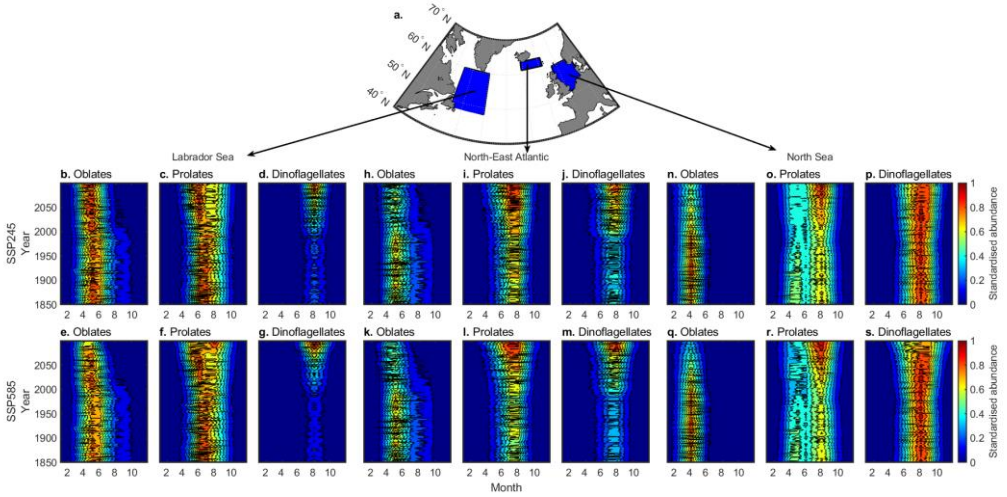



*Figure 42. Long-term changes (1850-2100) in the monthly abundance of oblates, prolates and dinoflagellates in three areas of the North Atlantic Ocean. (a) spatial distribution of the CPR samples in the three chosen areas. Monthly changes in (b-g) the Labrador Sea, (h-m) the North-East Atlantic and (n-s) the North-Sea. Long-term monthly changes in the standardized abundance of (b, e, h, k, n and q) oblates, (c, f, i, l, o and r) prolates and (d, g, j, m, p and s) dinoflagellates using scenarios (b-d, h-j and n-p) SSP245 and (e-g, k-m and q-s) SSP585. Colours denote the mean monthly predicted (standardized) abundance of a taxonomic group. Each panel was standardised between 0 and 1 for comparison purpose.*

Changes projected by METAL are in agreement with theoretical expectations, i.e. spring species (oblates) will experiment a contraction of their phenology associated with a declining abundance, while summer species (i.e. prolates and dinoflagellates that occur during period of stratification) will exhibit a phenological dilatation with both earlier and later termination (Figure 4)(Beaugrand and Kirby, 2018). The phenological contraction of oblates, associated with a decline in abundance, might be caused by the shortening of photoperiod and light levels in winter, which are known to prevent primary production at high latitudes (Boyce et al., 2017; Caracciolo et al., 2021). Therefore, a biogeographical shift of these diatoms toward higher latitudes is expected and would be in agreement with the increased primary production projected by ESMs in the Arctic region (Kwiatkowski et al., 2020) and the rise in diatoms we observed in the northern regions of the North-East Atlantic (Edwards et al., 2022). On

the contrary, the phenological expansion of prolates and dinoflagellates is indicative of species/taxa that are adapted to the enhanced stratification and nutrient-depleted conditions in the future North Atlantic Ocean (Beaugrand and Kirby, 2018; Kwiatkowski et al., 2020).

Because of the shifts in phytoplankton phenology, changes in APS are expected to desynchronise species interaction and trigger trophic mismatch (Beaugrand et al., 2003a; Cushing, 1990; Edwards and Richardson, 2004). Such alterations of the trophic links have already been observed in freshwater lakes (e.g. Winder and Schindler, 2004) but remain difficult to anticipate in the marine environment because each species has a specific response to environmental variability, which is driven by its life strategy and niche (Beaugrand and Kirby, 2018). Furthermore, it has also been shown that only seasonally heterogeneous environmental changes can lead to a desynchronization in species interaction (e.g. warmer temperature in spring but not in summer)(Straile et al., 2015). Results from the METAL theory suggest that trophic mismatch might be limited because "*interacting species have at least a part of their ecological niche in common, providing that no other important ecological dimensions or habitat requirements differ among them*" (Beaugrand and Kirby, 2018). Indeed, species are expected to track the changes along the environmental variable for which they have the narrowest tolerance (Ackerly, 2003). Therefore, in the pelagic environment, one could expect a mismatch among species that are not linked to the same habitat components (i.e. the stable- and substrate-biotope components; van der Spoel, 1994). For example, it has been shown that a shift in phytoplankton phenology is more likely to affect the fish that spawn in geographically fixed area than those for which their spawning grounds move according to the environmental changes (Asch et al., 2019).

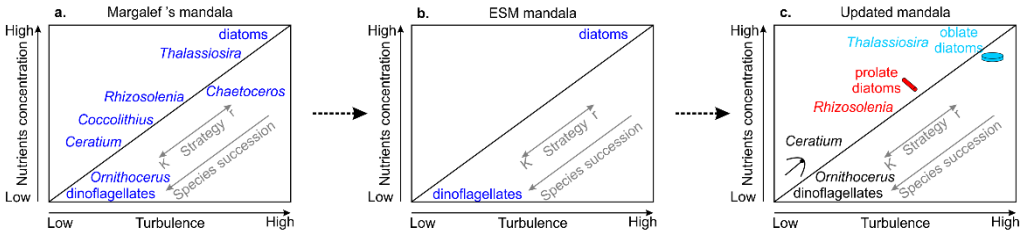



*Figure 43. (a) Margalef' mandala (adapted from Margalef, 1978), (b) version of the mandala that is used by ESMs to represent the diatoms and (c) an updated version according to our results on diatom morphological traits.*

Changes in phytoplankton community composition might also strongly affect the biological pump and carbon exportation (Passow and Carlson, 2012). ESMs predict a decline in primary production and carbon exportation in the North Atlantic because ocean warming will enhance water column stratification and diminish nutrient supply in the euphotic zone, which are believed to negatively affect diatoms (Bopp et al., 2005; Fu et al., 2016; Kwiatkowski et al., 2020; Marinov et al., 2010). However, the geological records and genomic studies have shown that diatoms are able to survive and even to sustain high biomass production in stratified conditions. For example, the formation of Mediterranean sapropels, some of the most carbon rich sediment, has

been shown to result in part from a high production of Rhizosolenid (i.e. prolate, see Figure 2) diatoms during periods of highly stratified conditions (Kemp and Villareal, 2013) and genomic studies conducted as part of the *Tara* Ocean Expedition observed similar diatom diversity in coastal and oligotrophic waters (Malviya et al., 2016).

Some authors have therefore suggested that the predicted decline of the diatoms results from the choice made by modelling teams to use a simplified version of Margalef's mandala (Kemp and Villareal, 2018; Kléparski et al., 2022b; Tréguer et al., 2018). Figure 43a displays the mandala as described by Margalef and Figure 43b the mandala as it is currently applied as part of ESMs. One can see that ESMs only consider diatoms that are in the upper right corner (i.e. the r-strategist oblate diatoms that thrive when turbulence and nutrients are important) and not the *Rhizosolenia* (i.e. K-strategist prolates that thrive when turbulence and nutrients are lower) (see Figure 40 and Figure 43a *versus* b). The framework proposed by Margalef can therefore be updated according to our results on diatom morphological traits (Figure 43c). In this updated version, dinoflagellates and prolate diatoms are K-strategists (i.e. equilibrium species), adapted to stable environments (i.e. predictable environments with low perturbations) with respect to nutrients, low turbulence and stratification (low MLD) whereas oblates are r-strategists (i.e. opportunist species), adapted to a more unstable environment (i.e. unpredictable; Figure 43c)(Kléparski et al., 2022b; Margalef, 1978; Southwood et al., 1974). Hence, the consideration of our revisited mandala (Figure 43c) may be a simple way to improve model predictions because oblate and prolate diatoms and dinoflagellates have different phenologies and responses to climate change, which affect our current understanding of the consequences of ocean warming on phytoplankton community structure and the biological carbon pump.

# 5. Conclusions and perspectives

By investigating phenology and the annual phytoplankton succession by means of the Phenoshift model in the North Atlantic and its adjacent seas, we have provided evidence that these phenomena are controlled by the niche-environment interaction. More precisely, a species' genome and its interaction with the environment determine species functional traits and life strategies, which in turn confers to a species a niche, a phenology and a place in the sequence of the annual phytoplankton succession. We show that diatoms have adapted to two distinct environments through an elongation of their cell shapes, oblate species being dominant when turbulence and nutrients are important, whereas prolates dominate when stratification is important and nutrients concentration low. We also demonstrate that planktonic species have a niche sufficiently distinct to enable their coexistence in a structured and impermanent marine environment. Plankton are therefore segregated along various environmental gradients and each species covers a specific part of the multidimensional environmental space available. Therefore, when a species is in the optimal part of its niche, it can generate high abundance (i.e. a bloom) and the succession of various environmental combinations determines the sequence of species appearance throughout the year. Hence, the phytoplankton community is reorganised permanently and the APS can be seen as revealing temporal coenoclines. Based on this framework, we showed that climate-induced changes in the environmental regime will alter species phenologies and the composition of the phytoplankton community, oblates being expected to exhibit a phenological contraction and a decline in annual abundance, shifting their range polewards, whereas prolates and dinoflagellates are expected to experience a phenological dilatation and an increase in annual abundance. These shifts might strongly affect ecosystem trophodynamics and carbon exportation.

We have identified key mechanisms involved in the APS of some key areas of the North Atlantic Ocean and its adjacent seas. It would be interesting to examine whether or not our results could be extended to other marine regions. It would be important to apply our approach to the data collected as part of the Global Alliance of the CPR survey (GACS) programme, especially in areas around Australia and the North Pacific Ocean where CPR data have been collected for many years; the Australian and North Pacific CPR surveys have monitored plankton since 2007 and 2002, respectively. By means of the Phenoshift model we could reproduce phenology and annual phytoplankton succession and investigate their past, present and future changes in the context of natural climatic variability and anthropogenic climate change. Furthermore, as the data on cell size and shape we used are available at a global scale (Leblanc et al., 2012), we could investigate potential changes in oblate and prolate diatoms in the North Pacific Ocean and Australian waters.

A second question is whether the biological pump will be affected by climate-mediated phenological shifts and associated changes in abundance in the North Atlantic Ocean? In the third chapter, we did not quantify carbon exportation but the

carbon cell content of diatoms and dinoflagellates can be retrieved from various datasets (e.g. Barton et al., 2013; Leblanc et al., 2012). Thereby, it might be possible to estimate the consequences of phytoplankton changes for carbon exportation, although we recognise other mechanisms such as remineralisation processes in the euphotic zone might complicate our analyses.

A third question is whether the phenological shifts we predict will affect higher trophic levels (i.e. copepods and fish). To answer this question, the Phenoshift model can also be applied to characterise the annual succession of zooplankton species/taxa. By doing so, we should be able to investigate their past present and future changes in phenology and abundance and quantify putative alterations in carbon exportation.

# Synthèse en français

Le phytoplancton est une composante clé des écosystèmes marins, produisant par photosynthèse près de la moitié de la production primaire nette (PPN) globale (Field et al., 1998) et assurant ainsi l'apport d'énergie endosomatique vers les niveaux trophiques supérieurs tels que les copépodes et les poissons. Les espèces phytoplanctoniques subissent de façon saisonnière de fortes variations de leur abondance et se succèdent en fonction de leur phénologie (leur rythme de reproduction) tout au long de l'année. Dans l'Atlantique Nord-Est, le schéma de succession consiste en une dominance des diatomées et des espèces du genre *Phaeocystis* au printemps, puis des coccolithophores vers la fin du printemps, suivi des dinoflagellés et des nanoflagellés en été avec parfois un dernier pic de diatomées à l'automne (Atkinson et al., 2018; McQuatters-Gollop et al., 2007; Widdicombe et al., 2010). Cette dynamique annuelle affecte l'ensemble de l'écosystème marin et détermine la composition saisonnière des communautés, de par son influence sur le cycle de vie et le comportement migratoire de nombreuses espèces telles que les copépodes et les poissons (Beaugrand et al., 2003a; Cushing, 1969; Edwards and Richardson, 2004). Par exemple, la survie des larves d'églefin et le temps d'éclosion des crevettes sont liés aux efflorescences printanières du phytoplancton (Koeller et al., 2009; Platt et al., 2003). De plus, ces pics massifs de population sédimentent dans les profondeurs des océans et contribuent ainsi à l'exportation de carbone (Henson et al., 2012; Passow and Carlson, 2012; Smetacek, 1985). Par conséquent, l'étude de la succession phytoplanctonique annuelle est cruciale pour notre compréhension du fonctionnement des écosystèmes pélagiques et de leurs services d'approvisionnement et de régulation dans le contexte du dérèglement climatique.

Des études à grande échelle ont mis en évidence que les patrons phénologiques du phytoplancton ne sont pas uniformes d'un bassin océanique à l'autre, une forte variabilité latitudinale étant observée, suggérant un contrôle de ce phénomène et de la succession annuelle par des facteurs environnementaux. Les différents patrons observés peuvent être résumés comme suit : aux latitudes élevées (au-dessus de 30°), des pics de population très intenses mais aussi très brefs sont observés, tandis qu'entre 30°S et 30°N les pics sont plus atténués mais durent plus longtemps (Racault et al., 2012). Ainsi, les régions de hautes latitudes se caractérisent par des efflorescences rapides et intenses avec une biomasse élevée, alors que dans les régions tropicales et subtropicales, la saisonnalité est quasi-absente et la biomasse reste faible (Racault et al., 2012). Ces différences latitudinales sont causées par des mécanismes d'initiation distincts de la période saisonnière de reproduction : dans les régions de haute latitude, la saison de reproduction est initiée par une augmentation de la photopériode, ce qui explique pourquoi on observe une progression latitudinale des efflorescences depuis l'équateur vers les pôles (Boyce et al., 2017; Racault et al., 2012). Entre les tropiques, la lumière n'étant pas un facteur limitant, la saison de reproduction est initiée lorsque des nutriments sont apportés dans la zone euphotique par un mélange accru de la colonne d'eau (Boyce et al., 2017; Racault et al., 2012).

Cependant, une certaine variabilité locale a été détectée dans l'océan Atlantique Nord-Est, le bloom printanier s'initiant en mer du Nord puis se propageant vers les régions situées au nord (McQuatters-Gollop et al., 2007).

L'augmentation des concentrations atmosphériques en gaz à effet de serre (comme le $CO_2$, le $CH_4$ et le $NO_2$) est à l'origine d'un changement global du climat qui devrait atteindre une ampleur rarement égalée dans l'histoire géologique de notre planète, modifiant l'environnement et menaçant la composition et l'agencement des écosystèmes (IPCC, 2021). L'hydrosphère océanique a déjà été affectée par ces changements car elle a absorbé entre 20 et 30 % des émissions anthropiques totales de $CO_2$ et près de 90 % du réchauffement (IPCC, 2019; Zanna et al., 2019). Ainsi, depuis la seconde moitié du XX$^{e}$ siècle, la teneur en oxygène globale dans l'océan a diminué et la stratification des océans s'est accrue, altérant le cycle des nutriments (IPCC, 2019; Sallée et al., 2021). Au cours du siècle à venir, des changements importants sont prédits par les modèles du système Terre, avec une augmentation globale de la température de surface de la mer et de la stratification, et une baisse du pH, de l'oxygène dissous et de la concentration en nitrate dans la zone euphotique (Bopp et al., 2013, p. 2013; IPCC, 2019; Kwiatkowski et al., 2020). Cependant, de grandes variabilités spatiales sont observées dans les projections de ces modèles, avec par exemple, un réchauffement plus important de l'hémisphère Nord, en particulier dans les océans Arctique et Pacifique Nord, associé à un refroidissement mineur ou à un réchauffement limité, selon les scénarios, du gyre subarctique (Keil et al., 2020; Kwiatkowski et al., 2020). La stratification renforcée, associée à un mélange moins important de la colonne d'eau, devrait fortement diminuer l'apport de nitrate dans les océans Atlantique Nord et Pacifique Nord (Kwiatkowski et al., 2020).

Les réponses des espèces marines face à ces changements sont diverses (Beaugrand et al., 2002b; Edwards and Richardson, 2004; Richardson and Schoeman, 2004). La séquence hiérarchique des réponses typiquement observées a été décrite par Beaugrand et Kirby (Beaugrand and Kirby, 2018); elle est brièvement résumée ici. La première réponse d'un individu à un changement dans son environnement thermique sera de modifier son comportement, c'est-à-dire d'essayer d'exploiter la micro-hétérogénéité de l'habitat pour trouver des conditions environnementales plus favorables. Par exemple, certaines espèces pélagiques descendent plus en profondeur pour échapper aux conditions défavorables en surface (Perry et al., 2005). Lorsque les individus d'une espèce atteignent les limites de leur plasticité comportementale, ils vont ajuster leur physiologie, c'est-à-dire s'acclimater, ce qui va avoir des effets sur leur activité, leur croissance et leur développement (Beaugrand and Kirby, 2018). Quand les changements comportementaux et physiologiques ne sont plus possibles, des ajustements phénologiques se produisent, c'est-à-dire des décalages dans l'initiation et/ou la fin de la saison de reproduction. Lorsque la plasticité phénologique devient impossible, l'abondance annuelle locale commence à diminuer, menant à ce qui est perçu à grande échelle comme un déplacement biogéographique. Si le changement environnemental est trop fort, une disparition locale peut être observée.

Des changements dans la phénologie, l'abondance et la biogéographie du plancton sont déjà documentés. Globalement, une phénologie plus précoce est

observée à la fois pour le phytoplancton et le zooplancton (Friedland et al., 2018; Poloczanska et al., 2016, 2013) mais ces changements sont beaucoup plus complexes lorsqu'on les étudie à une résolution taxonomique plus fine et à des échelles spatiales plus petites. Ainsi, dans l'Atlantique Nord-Est, la phénologie des dinoflagellés a été avancée d'une trentaine de jours alors que la phénologie des diatomées ne l'a pas été, même si une grande variabilité inter-espèces est observée pour ce groupe (Chivers et al., 2020; Edwards and Richardson, 2004). En revanche, la phénologie du phytoplancton du Nord de la mer Rouge a été retardée et sa durée raccourcie (Gittings et al., 2018). Un déclin global de la biomasse phytoplanctonique a aussi été documenté (Boyce et al., 2010), bien que des variabilités locales soient observées ici aussi dans certaines parties de l'Atlantique du Nord-Est (Reid et al., 1998), où l'abondance des diatomées augmente mais pas celle des dinoflagellés qui sont en déclin (Hinder et al., 2012). Des déplacements vers le pôle sont également documentés, avec une grande variabilité observée parmi les diatomées et les dinoflagellés (Chivers et al., 2017). Une diminution de la taille moyenne des copépodes calanoïdés a également été observée (Beaugrand et al., 2010), résultant d'un déplacement vers le nord des espèces caractérisant les eaux chaudes (de taille plus petite) associé à un déclin des espèces caractérisant les eaux plus froides (de plus grande taille) (Beaugrand et al., 2002b; Edwards et al., 2021). Par conséquent, ces changements dans la distribution spatio-temporelle du phytoplancton et du zooplancton ont modifié la composition des communautés planctoniques et le nouvel interactome qui en résulte (IPCC, 2019; Winder and Schindler, 2004), affectant la production halieutique (IPCC, 2019). (À noter que dans le même temps le necton a été fortement impacté par la surexploitation (Pauly et al., 1998).) Le flux de carbone exporté vers l'océan profond a également été modifié (Brun et al., 2019). Aux hautes latitudes autour du pôle de l'hémisphère nord, la réorganisation des communautés est encore amplifiée par la fonte de la glace de mer, ce qui permet à certaines espèces de commencer à se déplacer d'un bassin océanique à un autre, par exemple la diatomée de l'océan Pacifique *Neodenticula seminae* qui se retrouve aussi maintenant dans l'océan Atlantique (Reid et al., 2007).

En raison de la stratification accrue qui modifiera l'apport en nutriments dans la zone euphotique, une baisse globale de la PPN est attendue. Il est important cependant de noter que tous les modèles ne sont pas d'accord à ce sujet, certains prédisant une baisse tandis que d'autres prédisent une légère augmentation (Kwiatkowski et al., 2020). La PPN devrait diminuer dans l'Atlantique Nord et le Pacifique équatorial occidental, tandis qu'elle devrait augmenter dans les océans polaires en raison du retrait de la glace de mer (Kwiatkowski et al., 2020). La réorganisation de la communauté phytoplanctonique devrait se poursuivre, avec une expansion des espèces les mieux adaptées aux eaux fortement stratifiées et pauvres en nutriments (Dutkiewicz et al., 2013). Ainsi, dans l'Atlantique Nord, un déplacement vers le nord-est de la distribution spatiale des diatomées et des dinoflagellés est envisagé (Barton et al., 2016), ainsi qu'un remplacement progressif des diatomées par les diazotrophes (Henson et al., 2021), participant au déclin global prévu de la PPN (Bopp et al., 2005; Marinov et al., 2010). De leur côté, les dinoflagellés sont censés s'étendre vers l'océan Austral (Henson et al., 2021). En conséquence, la biodiversité du

plancton sera réduite dans les zones tropicales et augmentée dans les régions tempérées et polaires (Beaugrand and Kirby, 2018; Henson et al., 2021; Thomas et al., 2012). Une phénologie plus précoce est également attendue pour les espèces phytoplanctoniques, avec une avancée de 0,5 à 1 mois à l'échelle mondiale (Henson et al., 2013), bien qu'une grande variabilité régionale soit attendue ici aussi avec une phénologie plus précoce aux latitudes élevées et une phénologie plus tardive aux latitudes moyennes (Asch et al., 2019; Henson et al., 2018; Yamaguchi et al., 2022). Les changements phénologiques du phytoplancton pourraient aussi modifier le recrutement des poissons, en particulier pour les espèces qui fraient dans des zones géographiques fixes, impactant davantage les captures et la productivité de la pêche qui sont déjà altérées par d'autres pressions anthropiques (Asch et al., 2019; IPCC, 2019).

Il est donc important de comprendre comment et pourquoi les espèces planctoniques répondent aux changements de leur environnement. Pour se trouver en un lieu donné, un organisme doit être adapté aux conditions environnementales locales. Ce contrôle exercé par l'environnement est bien illustré par la loi de Becking/Beijerinck qui stipule que « tout est partout, mais l'environnement sélectionne » (De Wit and Bouvier, 2006). La sélection se produit parce que les performances d'une espèce sont modulées par les conditions environnementales (Shelford, 1913). Lorsque les conditions sont optimales, la croissance et la reproduction sont maximales et une forte abondance est observée (Brown, 1984; Helaouët and Beaugrand, 2009). Lorsque les conditions sont moins favorables, les performances diminuent et la croissance, la reproduction et l'alimentation sont progressivement altérées et une abondance moindre est observée, jusqu'à la mort (Brown, 1984; Helaouët and Beaugrand, 2009). Ce paradigme est appelé la niche écologique d'une espèce et a été conceptualisé par G. E. Hutchinson (Hutchinson, 1957). La niche est donc définie comme l'ensemble des conditions environnementales permettant à une espèce de croître et de se reproduire, déterminant sa réponse à la variabilité environnementale. La niche peut être assimilée à un hypervolume à p dimension, où p représente idéalement toutes les variables environnementales (i.e. biotiques et abiotiques) (Hutchinson, 1957). D'autres concepts de niche ont également été proposés (par exemple le niche grinnellienne et eltonienne) mais ont été jugés moins opérationnels ici car ils envisageaient la niche comme un attribut de l'environnement et non de l'espèce (Colwell and Rangel, 2009; Elton, 1927; Grinnell, 1917; Pulliam, 2000). Selon le principe de Gause, deux espèces ne peuvent coexister dans la même aire géographique si elles possèdent la même niche car une espèce devrait exclure l'autre par compétition (Gause, 1934). Ce principe d'exclusion compétitive, associé au fait que toutes les combinaisons des p variables environnementales ne sont pas représentées en permanence dans l'environnement réalisé (Jackson and Overpeck, 2000), ont conduit à affiner le concept de niche, en définissant les niches dites fondamentales, potentielles et réalisées (Hutchinson, 1978; Jackson and Overpeck, 2000). La niche fondamentale représente toutes les combinaisons de p variables environnementales dans lesquelles une espèce peut théoriquement se retrouver et la partie de ces combinaisons qui se retrouvent

véritablement dans le monde réel est appelée la niche potentielle. La niche potentielle est ensuite elle-même modulée par les interactions biotiques (c'est-à-dire la compétition ou la facilitation) et les capacités de dispersion, ce qui donne la niche réalisée (Bruno et al., 2003; Hutchinson, 1978; Pulliam, 1988). De nombreuses méthodes ont été développées pour caractériser la niche multidimensionnelle d'une espèce (Beaugrand and Helaouët, 2008; Blonder et al., 2018; Broennimann et al., 2012; Junker et al., 2016). Basées sur des données de présence ou d'abondance, ces méthodes visent à quantifier les hypervolumes des niches le long de différents paramètres environnementaux. Ce faisant, le degré de chevauchement entre les niches peut être évalué et diverses hypothèses écologiques et évolutives peuvent être examinées, telles que le rôle de la compétition, de la spéciation et du filtrage par l'habitat dans l'assemblage des communautés (Blonder et al., 2018; Cardillo and Warren, 2016; Cornwell et al., 2006). De plus, du fait de la correspondance réciproque entre l'espace de la niche et l'espace physique réel, appelée « la dualité d'Hutchinson », ces méthodes permettent également de reconstruire les distributions spatio-temporelles passées, présentes et futures des espèces (Colwell and Rangel, 2009; Jackson and Overpeck, 2000). Par conséquent, la niche est un mécanisme clé pour comprendre la réorganisation de la succession annuelle ainsi que les changements phénologiques dans le contexte du dérèglement climatique.

Les traits fonctionnels des espèces sont aussi des éléments importants pour comprendre et prédire leurs réponses aux changements environnementaux, leurs distributions spatio-temporelles reflétant la distribution des conditions optimales pour une espèce donnée (Ackerly, 2003). Un trait est défini comme une "propriété mesurable des organismes, généralement mesurée au niveau de l'individu et utilisée de manière comparative entre les espèces" et un trait fonctionnel est défini comme un "trait qui influence fortement la performance de l'organisme" (McGill et al., 2006). Au cours de leur évolution, les organismes phytoplanctoniques ont développé divers traits influençant différentes fonctions écologiques (Litchman and Klausmeier, 2008). Par exemple, les diatomées et les dinoflagellés ont développé des stades de repos qui leur permettent de survivre lorsque les conditions environnementales deviennent défavorables (McMinn and Martin, 2013; McQuoid and Hobson, 1996) mais aussi des mécanismes pour contrôler leur position dans la colonne d'eau et optimiser leur exposition à la lumière (Kemp and Villareal, 2018, p. 2018; Miller, 2004). Les principaux groupes phytoplanctoniques ont également développé des stratégies distinctes d'absorption des ressources, avec divers traits et compromis associés, qui pourraient expliquer la distribution et la biodiversité des organismes phytoplanctoniques dans les océans (Litchman et al., 2007). Le rôle des traits fonctionnels dans la compréhension et la prédiction de la succession annuelle a été souligné par Ramon Margalef dans son célèbre mandala (Margalef, 1978). Selon lui, les principaux taxons phytoplanctoniques sont séparés par leurs besoins en termes de nutriments et de turbulence, depuis les diatomées qui se développent dans les eaux peu stratifiées et riches en nutriments aux dinoflagellés qui prospèrent dans les eaux pauvres en nutriments hautement stratifiées; de nombreux modèles du système Terre prédisent la réponse des diatomées au changement climatique sur la base de ce travail (Bopp et al., 2005;

Marinov et al., 2010; Tréguer et al., 2018). Plus récemment, Edwards et collègues ont montré que la succession phytoplanctonique annuelle observée dans la Manche occidentale pouvait être reconstituée au moyen de traits mesurés en laboratoire et caractérisant l'utilisation de la lumière et des nitrates (K. F. Edwards et al., 2013).

Les traits morphologiques (c'est-à-dire la taille et la forme des cellules) pourraient également être importants pour expliquer la phénologie et la succession annuelle, car ils sont connus pour affecter la biodiversité du phytoplancton, l'absorption des ressources, le mouvement vertical et la résistance face aux prédateurs (Litchman and Klausmeier, 2008; Ryabov et al., 2021). Les traits morphologiques influencent les capacités d'absorption des ressources car l'absorption des nutriments se fait par l'action de protéines de transport situées à la surface des cellules (par exemple, les pompes transmembranaires). Une cellule de plus grande surface héberge donc plus de protéines de transport, ce qui optimise l'absorption des nutriments lorsque leurs concentrations est faible mais à son tour, cette cellule doit posséder plus de matériaux et d'énergie, ce qui est coûteux (Karp-Boss and Boss, 2016; Naselli-Flores and Barone, 2011). Au contraire, une cellule plus petite héberge moins de protéines de transport mais possède un rapport surface/volume plus important, ce qui permet aux nutriments d'atteindre plus rapidement tous les organites à l'intérieur de la cellule et donc d'optimiser les réactions physiologiques (Naselli-Flores and Barone, 2011). Une cellule sphérique a également le flux de nutriments le plus faible par rapport aux autres formes et, par conséquent, l'écart par rapport à la sphéricité confère à une cellule un avantage concurrentiel. Par exemple, les formes plus allongées expérimentent un apport nutritif plus important (Grover, 1989; Karp-Boss and Boss, 2016; Pahlow et al., 1997). Les traits morphologiques modifient également la suspension dans la colonne d'eau, ce qui est considérée comme étant un objectif évolutif très important chez les espèces phytoplanctoniques (Naselli-Flores et al., 2021). En effet, la plupart d'entre elles ont développé des épines et des soies qui augmentent directement la traînée de frottement ou aident à tourner la dimension cellulaire la plus longue perpendiculairement à l'axe de sédimentation (Padisak et al., 2003; Smayda, 1970). De la même manière, la déviation de la sphéricité leur permet de modifier la vitesse de sédimentation, cette dernière étant ralentie en conditions non turbulentes lorsque la forme est allongée (Padisak et al., 2003) alors qu'elle est augmentée en conditions turbulentes (Clifton et al., 2018). La taille est également importante dans la régulation de la flottabilité : par exemple un disque de 5 µm de diamètre va couler moins vite qu'un cylindre de même diamètre mais un disque de 500 µm de diamètre coulera plus rapidement qu'un cylindre de 500 µm de diamètre (Smayda, 1970). Les traits morphologiques sont également supposés avoir évolués en réponse à la prédation et pourraient donc influencer l'interaction avec les prédateurs (Karp-Boss and Boss, 2016; Smetacek, 2001). Il a été montré que la prédation par le zooplancton peut modifier la fréquence de certaines formes phytoplanctoniques, du fait d'une sélection comportementale ou mécanique par le zooplancton (Sournia, 1982). Par exemple, les cellules sphériques sont plus facilement ingérées par les copépodes que les cellules allongées, qui doivent être réorientées (Karp-Boss and Boss, 2016), bien que les copépodes soient en mesure de réorienter les cellules quand

la turbulence est faible (Visser and Jonsson, 2000). Par conséquent, la prise en compte à la fois du concept de niche et des traits fonctionnels associés pourrait nous aider à mieux comprendre et prédire les conséquences du dérèglement climatique sur la phénologie des espèces phytoplanctoniques et la succession annuelle qui en résulte.

Dans le cadre de ce travail, nous avons utilisé les données collectées par le programme Continuous Plankton Recorder (CPR) et développé un modèle (le modèle PhenoShift) dans le cadre de la théorie METAL (MacroEcological Theory on the Arrangement of Life). Le CPR est un programme de suivi à long terme du plancton géré par la Marine Biological Association (MBA) de Plymouth, au Royaume-Uni (Batten et al., 2003; Reid et al., 2003). Le CPR est le programme le plus long et le plus étendu de suivi du plancton marin au monde, couvrant l'océan Atlantique Nord et ses mers adjacentes depuis 1958 jusqu'à nos jours (Richardson et al., 2006). L'échantillonnage du phytoplancton et du zooplancton est opéré par une machine tractée derrière des navires marchands, appelés "navires d'opportunité", et opérant à une profondeur d'environ 7-10 m (Hays and Warner, 1993; Warner and Hays, 1994).

La théorie « MacroEcological Theory on the Arrangement of Life » (METAL) a pour but d'expliquer comment la biodiversité, du niveau individuel au niveau communautaire, s'organise dans l'espace et dans le temps et comment elle répond aux changements environnementaux (Beaugrand, 2015a). La théorie utilise le concept de niche écologique au sens d'Hutchinson comme une brique macroscopique élémentaire (Hutchinson, 1957) pour relier un grand nombre de phénomènes biogéographiques, écologiques et paléoécologiques. La niche étant déterminée par le génome, son utilisation permet de prendre en compte des processus intraspécifiques sous-jacents difficiles à modéliser (comme les processus génétiques ou moléculaires). L'hypothèse principale de METAL est que l'interaction niche-environnement est une interaction fondamentale qui se propage depuis le niveau organisationnel des individus jusqu'au niveau de la communauté et de l'écosystème. La force de METAL est de considérer que même si les écosystèmes sont des systèmes adaptatifs complexes, leur organisation de base et leur sensibilité peuvent être prédites à partir de règles simples et qu'une part importante de leurs changements sont déterministes et donc prévisibles. Les principales applications de METAL vont du niveau organisationnel individuel au niveau de l'espèce et de la communauté. Pour reconstituer les changements spatio-temporels dans l'arrangement des communautés, METAL génère un grand nombre de pseudo-espèces (c'est-à-dire des espèces virtuelles), qui sont caractérisées par une niche unique (en accord avec le principe d'exclusion compétitive). Ces pseudo-espèces sont alors autorisées à coloniser une région donnée tant qu'elles peuvent résister aux changements du régime environnemental. L'examen des pseudo-communautés reconstituées a montré que METAL peut reproduire (i) la biodiversité passée et contemporaine (Beaugrand et al., 2015; Zacaï et al., 2021), (ii) certains changements abrupts de communauté (Beaugrand et al., 2019a) et (iii) des patrons biogéographiques à grande échelle tel que la règle de Rapoport (Beaugrand et al., 2020).

Les principaux objectifs de cette thèse sont donc (i) d'améliorer notre compréhension de la phénologie des espèces phytoplanctoniques et de la succession

annuelle qui en résulte au niveau de la communauté, (ii) d'identifier les principaux mécanismes sous-jacents qui contrôlent la biodiversité, la phénologie et la coexistence des espèces et (iii) modéliser les évolutions passées, présentes et futures de la phénologie des diatomées et des dinoflagellés et la succession annuelle qui en résulte. Bien que nous ayons d'abord cherché à appliquer notre approche à l'échelle mondiale en utilisant les données de la « Global Alliance of Continuous Plankton Recorder Surveys » (GACS), nous avons finalement concentré nos investigations dans l'océan Atlantique Nord où la quantité de données CPR dans l'espace et le temps était la plus importante, ce qui a facilité le développement de nos méthodes et nos analyses. Néanmoins, nous pensons que les procédures numériques développées dans le cadre de ce travail peuvent être appliquées dans de futures études utilisant les données GACS.

Dans le premier chapitre, nous avons étudié la succession phytoplanctonique annuelle observée en mer du Nord et nous avons identifié les mécanismes sous-jacents expliquant la phénologie des espèces dans l'environnement pélagique (Caracciolo et al., 2021). Nous avons spécifiquement examiné les variations annuelles de l'abondance de 81 espèces/taxons appartenant à cinq groupes taxonomiques différents, soit 53 diatomées, 24 dinoflagellés, un silicoflagellé, deux prymnésiophycées et une cyanophycée. Pour ce faire, nous avons tout d'abord caractérisé la succession annuelle observée dans les données recueillies par le CPR en appliquant une Analyse en Composantes Principales (ACP). Nous avons ainsi identifié cinq périodes temporelles distinctes, chacune étant caractérisée par un assemblage planctonique spécifique. Ensuite, nous avons reconstruit la niche de chaque espèce le long de cinq dimensions environnementales au moyen d'une approche basée sur METAL. Nous avons testé différents modèles de complexité croissante (de deux à cinq dimensions environnementales) pour déterminer si un nombre croissant de paramètres environnementaux pouvaient améliorer notre reconstruction de la phénologie des espèces et de la succession annuelle. Nos résultats ont montré que (i) la succession annuelle est bien modélisée par les modèles METAL utilisant des niches tridimensionnelles basées sur la température de surface de la mer, les rayonnements photosynthétiquement actifs et les concentrations en nitrate, (ii) que la succession annuelle résulte de l'interaction niche-environnement, (iii) que ce phénomène doit être étudié au niveau de l'espèce et (iv) que l'utilisation des variables environnementales les plus pertinentes apparaît plus importante que la complexité du modèle.

Suite à ces résultats, nous avons cherché à savoir si certains traits fonctionnels pouvaient influencer la niche d'une espèce phytoplanctonique et sa phénologie et donc être utilisés pour mieux comprendre et reconstruire la succession annuelle (Kléparski et al., 2022b). Pour ce faire, nous nous sommes concentrés sur les diatomées et avons examiné leurs traits morphologiques, qui sont des traits majeurs affectant l'acquisition des nutriments, la capture de la lumière, la résistance aux prédateurs et le mouvement vertical dans la colonne d’eau. Nous avons attribué à chaque espèce/taxon de diatomée une hauteur et un diamètre moyens de cellule (les informations sur la taille des cellules ont été récupérées pour 45 des 53 diatomées).

Selon leurs dimensions cellulaires moyennes, les 45 diatomées ont ensuite été séparées en trois groupes : les diatomées oblates (les espèces qui ont un diamètre cellulaire moyen supérieur à leur hauteur moyenne), (ii) les prolates (les espèces qui ont un diamètre cellulaire moyen inférieur à leur hauteur moyenne) et (iii) les oblates/prolates pour les espèces/taxons restants, ce qui signifie qu'ils peuvent être les deux. La succession annuelle des diatomées a ensuite été étudiée au moyen d'une Analyse en Composante Principale (ACP) et les niches des 45 espèces/taxons ont été caractérisées le long de diverses dimensions environnementales (température de surface de la mer, concentration en nutriments, rayonnements photosynthétiquement actifs, profondeur de la couche euphotique et de mélange). Nos résultats ont montré que les diatomées avec des traits morphologiques similaires ont des phénologies et des niches similaires, les oblates étant principalement abondantes au printemps et en automne tandis que les prolates sont dominantes à la fin du printemps, en été et en automne. Nous avons montré que l'allongement de la cellule donne un avantage concurrentiel aux prolates dans les eaux stratifiées pauvres en nutriments que l'on trouve en été car elle améliore leur flottabilité sans modifier leur rapport surface/volume (c'est-à-dire leur capacité à absorber les nutriments). Par conséquent, les formes de cellule des diatomées pourraient avoir évoluées en tant qu'adaptation à la viscosité et à la turbulence, leur conférant une niche et une place dans la séquence de la succession annuelle du plancton. Nous avons aussi démontré que les oblates et les prolates ont des changements d'abondance à long terme distincts. L'utilisation de ces deux groupes pourrait donc changer radicalement notre compréhension du fonctionnement des écosystèmes marins et nos prédictions sur les conséquences des futurs changements climatiques sur le cycle mondial du carbone.

Dans le second chapitre, nous avons examiné l'organisation spatio-temporelle de la biodiversité planctonique dans l'océan Atlantique Nord. Premièrement, en définissant des assemblages d'espèces basés à la fois sur des espèces/taxons phytoplanctoniques et zooplanctoniques, nous avons montré que les écosystèmes pélagiques de l'océan Atlantique Nord sont organisés de manière complexe (Kléparski et al., 2021). Nous avons aussi démontré que la biodiversité planctonique est fortement influencée par les conditions hydro-climatiques et notamment l'isotherme 10°C. Bien qu'aucun assemblage ne fût identifié comme étant un indicateur unique d'une région océanique spécifique, certains se trouvaient préférentiellement dans l'un des trois principaux biomes de l'océan Atlantique Nord, à savoir les biomes polaire, tempéré et côtier (au sens de Longhurst) (Longhurst, 1998). Ainsi, la région de l'Atlantique Nord est caractérisée par une mosaïque d'assemblages planctoniques qui ont des distributions spatiales chevauchantes dans de nombreux endroits, générant des cœnoclines spatiales (c'est-à-dire des gradients de biocénoses ou de communautés). Nos résultats ont montré que ces modèles spatiaux sont la conséquence des niches écologiques des espèces composant les assemblages et que chaque assemblage a sa propre signature environnementale (c'est-à-dire la combinaison des niches écologiques des espèces composant un assemblage), qui détermine sa distribution spatiale. Cela nous a amené à développer une nouvelle méthode que nous avons appelé dans un premier temps le « chromatogramme

environnemental » et qui est basée sur un graphique qui identifie où les espèces d'un assemblage s'agrègent le long de divers gradients environnementaux figurant les multiples dimensions environnementales. Dans un second temps, nous avons adapté cette méthode du « chromatogramme environnemental » au niveau organisationnel de l'espèce pour caractériser rapidement la niche écologique (multidimensionnelle) d'une espèce dans un plan. Le « chromatogramme environnemental » fût donc renommé « chromatogramme communautaire » lorsqu'il est appliqué au niveau de la communauté/de l'assemblage, et « chromatogramme d'espèce » lorsqu'il est appliqué au niveau de l'espèce. Un chromatogramme d'espèce donne un résumé graphique de la niche en représentant ensemble des gradients d'abondance par rapport à plusieurs variables environnementales. À partir d'un chromatogramme d'espèce, on peut facilement quantifier les optimums et les amplitudes de niche et lorsque plusieurs chromatogrammes sont examinés, la niche de différentes espèces peut être facilement comparée. Nous avons également proposé un indice pour quantifier le degré de chevauchement des niches et identifier les combinaisons les plus discriminantes de variables environnementales. La comparaison de notre méthode avec les packages R *dynRB* (Junker et al., 2016) et *Hypervolume* (Blonder et al., 2018) a montré que nos estimations de chevauchement de niche sont comparables avec ces deux méthodes. La méthode du chromatogramme d'espèce a été implémentée en Matlab et sous la forme d'un package R (appelé specieschrom) qui est librement disponible (Kléparski and Beaugrand, 2022).

L'application de cette méthode du chromatogramme d'espèces à toutes les espèces planctoniques enregistrées par le CPR nous a ainsi permis d'examiner les mécanismes expliquant la coexistence du plancton et surtout de revisiter le « paradoxe du plancton ». Formulé en 1961 par G.E. Hutchinson, ce paradoxe se demande comment "[…] il est possible qu'un certain nombre d'espèces coexistent dans un environnement relativement isotrope ou non structuré, toutes en compétition pour le même type de ressources" (Hutchinson, 1961). Notre analyse a permis de démontrer que les niches des espèces phytoplanctoniques et zooplanctoniques se séparent rapidement lorsque le nombre de dimensions environnementales utilisées pour les caractériser augmente. Nous avons également démontré que l'environnement pélagique est non seulement diversifié mais aussi impermanent dans l'espace et dans le temps. Par conséquent, les deux postulats implicites sur lesquels repose le paradoxe d'Hutchinson sont incorrects. La différenciation élevée des niches conduit les espèces à avoir une combinaison unique de conditions environnementales (c'est-à-dire une niche unique) où leurs performances en termes de croissance et de reproduction sont optimales. Ainsi, chaque espèce couvre une partie de l'espace environnemental multidimensionnel et la ségrégation le long des différentes dimensions de la niche leur permet de coexister dans l'espace et dans le temps.

Dans le dernier chapitre, nous avons examiné les effets prédits et observés du changement climatique sur l'abondance et la phénologie du phytoplancton. Premièrement, au moyen d'une approche basée sur METAL, nous avons examiné si la phénologie des diatomées (oblates et prolates) et des dinoflagellés était altérée par le réchauffement des océans (Kléparski et al., 2022a). Pour ce faire, nous avons utilisé les

projections de six modèles du système Terre pour deux scénarios de réchauffement. Nous avons également conçu six indices pour caractériser les changements phénologiques à long terme : (1) l'abondance annuelle maximale (AAM), (2) le jour où l'AAM est atteint, (3) l'initiation et (4) la fin de la période de reproduction saisonnière, c'est-à-dire le premier et le dernier jour où l'abondance est ≥ 50 % de l'AAM, (5) la durée saisonnière, c'est-à-dire le nombre de jours consécutifs où l'abondance est ≥ 50 % de l'AAM, et (6) l'abondance annuelle moyenne intégrée. Nos résultats ont montré que la phénologie des trois groupes va être altérée dans trois régions océaniques, à savoir la mer du Nord, l'Atlantique du Nord-Est et la mer du Labrador, les changements étant plus importants pour le scénario à fortes émissions. La phénologie des oblates devrait donc se raccourcir et leur abondance diminuer tandis que la phénologie des prolates et des dinoflagellés devrait s'allonger et leur abondance augmenter. Ces projections sont conformes aux attentes théoriques et apportent de nouvelles preuves susceptibles de revisiter notre compréhension et notre anticipation des conséquences des changements dans la phénologie et l'abondance du phytoplancton sur la pompe à biologique du carbone. Ils suggèrent également que les futurs modèles du système Terre devraient augmenter leur niveau de complexité en considérant différents groupes de diatomées adaptés, ou non, aux conditions stratifiées, c'est-à-dire les oblates et les prolates, ainsi que les dinoflagellés.

Dans une seconde et dernière étude, nous avons aussi examiné l'impact observé de la variabilité climatique naturelle et du dérèglement climatique sur les diatomées de la partie orientale de l'océan Atlantique Nord (Edwards et al., 2022). Dans cet article dirigé par le Pr. Martin Edwards, nos résultats ont montré que l'abondance des diatomées est positivement corrélée à la température et à l'intensité du vent dans le nord de la mer du Nord et le bassin insulaire alors qu'elle est négativement corrélée dans la mer Celtique et le golfe de Gascogne. Par conséquent, la population des diatomées augmente dans la partie nord de la zone étudiée alors qu'elle diminue dans les parties sud. Bien que cette étude ne fasse pas de distinction entre les diatomées oblates et prolates, les résultats concordent avec les changements phénologiques que nous avions prédits pour les oblates, c'est-à-dire que la baisse de leur abondance est plus prononcée en mer du Nord que dans l'Atlantique du Nord-Est.

# References

Ackerly, D.D., 2003. Community Assembly, Niche Conservatism, and Adaptive Evolution in Changing Environments. International Journal of Plant Sciences 164, S165–S184. https://doi.org/10.1086/368401

Adler, P.B., HilleRisLambers, J., Kyriakidis, P.C., Guan, Q., Levine, J.M., 2006. Climate variability has a stabilizing effect on the coexistence of prairie grasses. Proceedings of the National Academy of Sciences 103, 12793–12798. https://doi.org/10.1073 pnas.0600599103

Allen, J.T., 2005. Diatom carbon export enhanced by silicate upwelling in the northeast Atlantic. Nature 437, 728–732.

Allen, S., Henson, S., Hickman, A., Beaulieu, C., Doncaster, P.C., Johns, D.G., 2020. Interannual stability of phytoplankton community composition in the North-East Atlantic. Marine Ecology Progress Series 655, 43–57.

Alvain, S., Moulin, C., Dandonneau, Y., Bréon, F.M., 2005. Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. Deep Sea Research Part I: Oceanographic Research Papers 52, 1989–2004.

Amato, A., Dell'Aquila, G., Musacchia, F., Annunziata, R., Ugarte, A., Maillet, N., Carbone, A., Ribera d'Alcalà, M., Sangres, R., Iudicone, R., Ferrante, M., 2017. Marine diatoms change their gene expression profile when exposed to microscale turbulence under nutrient replete conditions. Scientific Reports 3826, 1–11.

Araújo, M.B., Guisan, A., 2006. Five (or so) challenges for species distribution modelling. Journal of Biogeography 33, 1677–1688.

Armbrust, E.V., 2009. The life of diatoms in the world's oceans. Nature 459, 185–192. https://doi.org/10.1038/nature08057

Asch, R.G., Stock, C.A., Sarmiento, J.L., 2019. Climate change impacts on mismatches between phytoplankton blooms and fish spawning phenology. Global Change Biology 25, 2544–2559. https://doi.org/10.1111/gcb.14650

Asrar, G., Myneni, R.B., Li, Y., Kanemasu, E.T., 1989. Measuring and modeling spectral characteristics of a tallgrass prairie. Remote Sensing of Environment 27, 143–155.

Assis, J., Tyberghein, L., S., B., Verbruggen, H., Serrão, E.A., Clerck, O., 2017. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography 27, 277–84.

Atkinson, A., Polimene, L., Fileman, E.S., Widdicombe, C.E., McEvoy, A.J., Smyth, T.J., Djeghri, N., Sailley, S.F., Cornwell, L.E., 2018. Comment. What drives plankton seasonality in a stratifying shelf sea? Some competing and complementary theories. Limnology and Oceanography 63, 2877–2884. https://doi.org/10.1002/lno.11036

Barnard, B., Batten, S.D., Beaugrand, G., Buckland, C., Conway, D.V.P., Edwards, M., Finlayson, J., Gregory, L.W., Halliday, N.C., John, A.W.G., Johns, D.G., Johnson, A.D., Jonas, T.D., Lindley, J.A., Nyman, J., Pritchard, P., Reid, P.C., Richardson, A.J., Saxby, R.E., Sidey, J., Smith, M.A., Stevens, D.P., Taylor, C.M., Tranter, P.R.G., Walne, A.W., Wootton, M., Wotton, C.O.M., Wright, J.C., 2004. Continuous Plankton Records: Plankton Atlas of the North Atlantic Ocean (1958-1999). II. Biogeographical charts. Marine Ecology Progress Series Supplement, 11–75.

Barton, A.D., Dutkiewicz, S., Flierl, G., Bragg, J., Follows, M.J., 2010. Patterns of Diversity in Marine Phytoplankton. Science 327, 1509–1511. https://doi.org/10.1126/science.1184961

Barton, A.D., Finkel, Z.V., Ward, B.A., Johns, D.G., Follows, M.J., 2013. On the roles of cell size and trophic strategy in North Atlantic diatom and dinoflagellate communities.

Limnology and Oceanography 58, 254–266. https://doi.org/10.4319/lo.2013.58.1.0254

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. Proceedings of the National Academy of Sciences 113, 2964–2969. https://doi.org/10.1073/pnas.1519080113

Barton, A.D., Lozier, M.S., Williams, R.G., 2015. Physical controls of variability in North Atlantic phytoplankton communities. Limnology and Oceanography 60, 181–197.

Batten, S.D., Clark, R., Flinkman, J., Hays, G., John, E., John, A.W.G., Jonas, T., Lindley, J.A., Stevens, D.P., Walne, A., 2003. CPR sampling: the technical background, materials and methods, consistency and comparability. Progress in Oceanography 58, 193–215. https://doi.org/10.1016/j.pocean.2003.08.004

Beaugrand, G., 2015a. Marine biodiversity, climatic variability and Global Change, Earthscan. Routledge, London.

Beaugrand, G., 2015b. Theoretical basis for predicting climate-induced abrupt shifts in the oceans. Philosophical Transactions of the Royal Society B: Biological Sciences 370, 20130264. https://doi.org/10.1098/rstb.2013.0264

Beaugrand, G., 1999. Le programme Continuous Plankton Recorder (CPR) et son application à l'étude des changements spatio-temporels de la biodiversité pélagique en Atlantique nord et en mer du Nord. Océanis 25, 417–433.

Beaugrand, G., Brander, K.M., Lindley, J.A., Souissi, S., Reid, P.C., 2003a. Plankton effect on cod recruitment in the North Sea. Nature 426, 661–664. https://doi.org/10.1038/nature02164

Beaugrand, G., Conversi, A., Atkinson, A., Cloern, J., Chiba, S., Fonda-Umani, S., Kirby, R.R., Greene, C., Goberville, E., Otto, S.A., Reid, P.C., Stemmann, L., Edwards, M., 2019a. Prediction of unprecedented biological shifts in the global ocean. Nature Climate Change 9, 10.

Beaugrand, G., Edwards, M., Brander, K., Luczak, C., Ibanez, F., 2008. Causes and projections of abrupt climate-driven ecosystem shifts in the North Atlantic: Causes and projections of abrupt climate-driven ecosystem shifts. Ecology Letters 11, 1157–1168. https://doi.org/10.1111/j.1461-0248.2008.01218.x

Beaugrand, G., Edwards, M., Hélaouët, P., 2019b. An ecological partition of the Atlantic Ocean and its adjacent seas. Progress in Oceanography 173, 86–102. https://doi.org/10.1016/j.pocean.2019.02.014

Beaugrand, G., Edwards, M., Legendre, L., 2010. Marine biodiversity, ecosystem functioning, and carbon cycles. Proceedings of the National Academy of Sciences 107, 10120–10124. https://doi.org/10.1073/pnas.0913855107

Beaugrand, G., Edwards, M., Raybaud, V., Goberville, E., Kirby, R.R., 2015. Future vulnerability of marine biodiversity compared with contemporary and past changes. Nature Climate Change 5, 695–701. https://doi.org/10.1038/nclimate2650

Beaugrand, G., Goberville, E., Luczak, C., Kirby, R.R., 2014. Marine biological shifts and climate. Proceedings of the Royal Society B: Biological Sciences 281, 20133350. https://doi.org/10.1098/rspb.2013.3350

Beaugrand, G., Helaouët, P., 2008. Simple procedures to assess and compare the ecological niche of species. Marine Ecology Progress Series 363, 29–37. https://doi.org/10.3354/meps07402

Beaugrand, G., Ibañez, F., Lindley, J.A., 2003b. An overview of statistical methods applied to CPR data. Progress in Oceanography 58, 235–262. https://doi.org/10.1016/j.pocean.2003.08.006

Beaugrand, G., Ibañez, F., Lindley, J.A., 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., Philip, C., Reid, P.C., 2002a. 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., Ibañez, F., Reid, P.C., 2000a. Long-term and seasonal fluctuations of plankton in relation to hydroclimatic features in the English Channel, Celtic Sea and Bay of Biscay. Marine Ecology Progress Series 200, 93–102.

Beaugrand, G., Ibañez, F., Reid, P.C., 2000b. Spatial, seasonal and long-term fluctuations of plankton in relation to hydroclimatic features in the English Channel, Celtic Sea and Bay of Biscay. Marine Ecology Progress Series 200, 93–102. https://doi.org/10.3354/meps200093

Beaugrand, G., Kirby, R.R., 2018. How Do Marine Pelagic Species Respond to Climate Change? Theories and Observations. Annual Review of Marine Science 10, 169–197. https://doi.org/10.1146/annurev-marine-121916-063304

Beaugrand, G., Kirby, R.R., 2016. Quasi-deterministic responses of marine species to climate change. Climate Research 69, 117–128. https://doi.org/10.3354/cr01398

Beaugrand, G., Kirby, R.R., Goberville, E., 2020. The mathematical influence on global patterns of biodiversity. Ecology and Evolution 10, 6494–6511. https://doi.org/10.1002/ece3.6385

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

Beaugrand, G., Luczak, C., Goberville, E., Kirby, R.R., 2018. Marine biodiversity and the chessboard of life. PLOS ONE 13, e0194006. https://doi.org/10.1371/journal.pone.0194006

Beaugrand, G., Mackas, D., Goberville, E., 2013a. Applying the concept of the ecological niche and a macroecological approach to understand how climate influences zooplankton: Advantages, assumptions, limitations and requirements. Progress in Oceanography 111, 75–90. https://doi.org/10.1016/j.pocean.2012.11.002

Beaugrand, G., McQuatters-Gollop, A., Edwards, M., Goberville, E., 2012. Long-term responses of North Atlantic calcifying plankton to climate change. Nature Climate Change 3, 263–267.

Beaugrand, G., Reid, P.C., Ibanez, F., Lindley, J.A., Edwards, M., 2002b. Reorganization of North Atlantic Marine Copepod Biodiversity and Climate. Science 296, 1692–1694. https://doi.org/10.1126/science.1071329

Beaugrand, G., Reid, P.C., Ibañez, F., Planque, B., 2000c. Biodiversity of North Atlantic and North Sea calanoid copepods. Marine Ecology Progress Series 204, 299–303. https://doi.org/10.3354/meps204299

Beaugrand, G., Rombouts, I., Kirby, R.R., 2013b. Towards an understanding of the pattern of biodiversity in the oceans: The pattern of biodiversity in the oceans. Global Ecology and Biogeography 22, 440–449. https://doi.org/10.1111/geb.12009

Begon, M., Townsend, C.R., Harper, J.L., 2005. Ecology: From Individuals to Ecosystems, 4th Edition, Wiley-Blackwell. ed.

Behrenfeld, M.J., 2014. Climate-mediated dance of the plankton. Nature Climate Change 4, 880–887. https://doi.org/10.1038/nclimate2349

Behrenfeld, M.J., 2010. Abandoning Sverdrup's Critical Depth Hypothesis on phytoplankton blooms. Ecology 91, 977–989. https://doi.org/10.1890/09-1207.1

Behrenfeld, M.J., Boss, E.S., 2018. Student's tutorial on bloom hypotheses in the context of phytoplankton annual cycles. Global Change Biology 24, 55–77. https://doi.org/10.1111/gcb.13858

Behrenfeld, M.J., Boss, E.S., 2014. Resurrecting the ecological underpinnings of ocean plankton blooms. Annual Review of Marine Science 6, 167–194.

Behrenfeld, M.J., O'Malley, R., Boss, E., Karp-Boss, L., Mundt, C., 2021. Phytoplankton biodiversity and the inverted paradox. ISME Communications 1, 52. https://doi.org/10.1038/s43705-021-00056-6

Blonder, B., 2018. Hypervolume concepts in niche- and trait-based ecology. Ecography 41, 1441–1455. https://doi.org/10.1111/ecog.03187

Blonder, B., 2016. Do Hypervolumes Have Holes? The American Naturalist 187, E93–E105. https://doi.org/10.1086/685444

Blonder, B., Lamanna, C., Violle, C., Enquist, B.J., 2017. Using *n* -dimensional hypervolumes for species distribution modelling: A response to Qiao et al. (2016). Global Ecology and Biogeography 26, 1071–1075. https://doi.org/10.1111/geb.12611

Blonder, B., Lamanna, C., Violle, C., Enquist, B.J., 2014. The n-dimensional hypervolume. Global Ecology and Biogeography 23, 595–609. https://doi.org/10.1111/geb.12146

Blonder, B., Morrow, C.B., Maitner, B., Harris, D.J., Lamanna, C., Violle, C., Enquist, B.J., Kerkhoff, A.J., 2018. New approaches for delineating n-dimensional hypervolumes. Methods in Ecology and Evolution 9, 305–319. https://doi.org/10.1111/2041-210X.12865

Bopp, L., Aumont, O., Cadule, P., Alvain, S., Gehlen, M., 2005. Response of diatoms distribution to global warming and potential implications: A global model study. Geophysical Research Letters 32, L19606. https://doi.org/10.1029/2005GL023653

Bopp, L., Resplandy, L., Orr, J.C., Doney, S.C., Dunne, J.P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., Vichi, M., 2013. Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. Biogeosciences 10, 6225–6245. https://doi.org/10.5194/bg-10-6225-2013

Boyce, D.G., Lewis, M.R., Worm, B., 2010. Global phytoplankton decline over the past century. Nature 466, 591–596. https://doi.org/10.1038/nature09268

Boyce, D.G., Petrie, B., Frank, K.T., Worm, B., Leggett, W.C., 2017. Environmental structuring of marine plankton phenology. Nature Ecology & Evolution 1, 1484–1494. https://doi.org/10.1038/s41559-017-0287-3

Boyd, P.W., Strzepek, R., Fu, F., Hutchins, D.A., 2010. Environmental control of open-ocean phytoplankton groups: Now and in the future. Limnology and Oceanography 55, 1353–1376.

Bracco, A., Provenzale, A., Scheuring, I., 2000. Mesoscale vortices and the paradox of the plankton. Proceedings of the Royal Society B: Biological Sciences 267, 1795–1800. https://doi.org/10.1098/rspb.2000.1212

Breton, E., Christaki, U., Bonato, S., Didry, M., Artigas, L.F., 2017. Functional trait variation and nitrogen use efficiency in temperate coastal phytoplankton. Marine Ecology Progress Series 563, 35–49.

Breton, E., Christaki, U., Sautour, B., Demonio, O., Skouroliakou, D.I., Beaugrand, G., Seuront, L., Kléparski, L., Poquet, A., Nowaczyk, A., Crouvoisier, M., Ferreira, S., Pecqueur, D., Salmeron, C., Brylinski, J., Lheureux, A., Goberville, E., 2021. Seasonal Variations in the Biodiversity, Ecological Strategy, and Specialization of Diatoms and Copepods in a

Coastal System With *Phaeocystis* Blooms: The Key Role of Trait Trade-Offs. Frontiers in Marine Science 8, 656300. https://doi.org/10.3389/fmars.2021.656300

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M.J., Randin, C., Zimmermann, N.E., Graham, C.H., Guisan, A., 2012. Measuring ecological niche overlap from occurrence and spatial environmental data: Measuring niche overlap. Global Ecology and Biogeography 21, 481–497. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Brown, J.H., 1984. On the relationship between abundance and distribution of species. The American Naturalist 124, 255–279. https://doi.org/10.1086/284267

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M., West, G.B., 2004. Toward a metabolic theory of ecology. Ecology 85, 1771–1789. https://doi.org/10.1890/03-9000

Brun, P., Stamieszkin, K., Visser, A.W., Licandro, P., Payne, M.R., Kiørboe, T., 2019. Climate change has altered zooplankton-fuelled carbon export in the North Atlantic. Nature Ecology & Evolution 3, 416–423. https://doi.org/10.1038/s41559-018-0780-3

Bruno, J.F., Stachowicz, J.J., Bertness, M.D., 2003. Inclusion of facilitation into ecological theory. Trends in Ecology & Evolution 18, 119–125.

Büchel, C., 2019. How diatoms harvest light. Science 365, 447–448. https://doi.org/10.1126/science.aay3036

Burson, A., Stomp, M., Mekkes, L., Huisman, J., 2019. Stable coexistence of equivalent nutrient competitors through niche differentiation in the light spectrum. Ecology 100, e02873. https://doi.org/10.1002/ecy.2873

Capuzzo, E., Lynam, C., Barry, J., Stephens, D., Forster, R.M., Greenwood, N., McQuatters-Gollop, A., Silva, T., 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. Global Change Biology 24, 352–364.

Capuzzo, E., Painting, S.J., Forster, R.M., Greenwood, N., Stephens, D.T., Mikkelsen, O.A., 2013. Variability in the sub-surface light climate at ecohydrodynamically distinct sites in the North Sea. Biogeochemistry 113, 85–103.

Capuzzo, E., Stephens, D., Silva, T., Barry, J., Forster, R.M., 2015. Decrease in water clarity of the southern and central North Sea during the 20th century. Global Change Biology 21, 2206–2214.

Caracciolo, M., Beaugrand, G., Hélaouët, P., Gevaert, F., Edwards, M., Lizon, F., Kléparski, L., Goberville, E., 2021. Annual phytoplankton succession results from niche-environment interaction. Journal of Plankton Research 43, 85–102. https://doi.org/10.1093/plankt/fbaa060

Cardillo, M., Warren, D.L., 2016. Analysing patterns of spatial and niche overlap among species at multiple resolutions. Global Ecology and Biogeography 25, 951–963. https://doi.org/10.1111/geb.12455

Chase, J.M., Leibold, M.A., 2003. Ecological niches: linking classical and contemporary approches. The University of Chicago Press, Chicago.

Chesson, P., 2000. Mechanisms of Maintenance of Species Diversity. Annual Review of Ecology and Systematics 31, 343–366. https://doi.org/10.1146/annurev.ecolsys.31.1.343

Chiswell, S.M., 2013. Comment on “Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom.” Global Biogeochemical Cycles 27, 1291–1293.

Chiswell, S.M., 2011. Annual cycles and spring blooms in phytoplankton: don’t abandon Sverdrup completely. Marine Ecology Progress Series 443, 39–50. https://doi.org/10.3354/meps09453

Chiswell, S.M., Calil, P.H.R., Boyd, P.W., 2015. Spring blooms and annual cycles of phytoplankton: a unified perspective. Journal of Plankton Research 37, 500–508.

Chivers, W.J., Edwards, M., Hays, G.C., 2020. Phenological shuffling of major marine phytoplankton groups over the last six decades. Diversity and Distributions 26, 536–548. https://doi.org/10.1111/ddi.13028

Chivers, W.J., Walne, A.W., Hays, G.C., 2017. Mismatch between marine plankton range movements and the velocity of climate change. Nature Communications 8, 14434. https://doi.org/10.1038/ncomms14434

Clifton, W., Bearon, R.N., Bees, M.A., 2018. Enhanced sedimentation of elongated plankton in simple flows. IMA Journal of Applied Mathematics 83, 743–766. https://doi.org/10.1093/imamat/hxy024

Cole, B.E., Cloern, J.E., 1987. An empirical model for estimating phytoplankton productivity in estuaries. Marine Ecology Progress Series 36, 299–305.

Colebrook, J.M., 1986. Environmental influences on long-term variability in marine plankton. Hydrobiologia 142, 309–325.

Colebrook, J.M., 1985. Sea surface temperature and zooplankton, North Sea, 1948 to 1983. Journal du Conseil International pour l'Exploration de la Mer 42, 179–185.

Colebrook, J.M., 1984. Continuous plankton records: relationships between species of phytoplankton and zooplankton in the seasonal cycle. Marine Biology 83, 313–323.

Colebrook, J.M., 1982. Continuous Plankton Records: seasonal variations in the distribution and abundance of plankton in the North Atlantic Ocean and the North Sea. Journal of Plankton Research 4, 435–462.

Colebrook, J.M., 1979. Continuous Plankton Records: seasonal cycles of phytoplankton and copepods in the North Atlantic Ocean and the North Sea. Marine Biology 51, 23–32.

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

Concept & Objectives | AtlantOS [WWW Document], 2019. URL https://www.atlantos-h2020.eu/project-information/conceptobjectives/ (accessed 5.17.19).

Cornwell, W.K., Schwilk, D.W., Ackerly, D.D., 2006. A trait-based test for habitat filtering: convex hull volume. Ecology 87, 1465–1471. https://doi.org/10.1890/0012-9658(2006)87[1465:ATTFHF]2.0.CO;2

Crisp, M.D., Arroyo, M.T.K., Cook, L.G., Gandolfo, M.A., Jordan, G.J., McGlone, M.S., Weston, P.H., Westoby, M., Wilf, P., Linder, H.P., 2009. Phylogenetic biome conservatism on a global scale. Nature 458, 754–756. https://doi.org/10.1038/nature07764

Cullen, J.J., Lewis, M.R., 1988. The kinetics of algal photoadaptation in the context of vertical mixing. Journal of Plankton Research 10, 1039–1063.

Cushing, D.H., 1990. Plankton Production and Year-class Strength in Fish Populations: an Update of the Match/Mismatch Hypothesis. Advances in Marine Biology 26, 249–293. https://doi.org/10.1016/S0065-2881(08)60202-3

Cushing, D.H., 1969. The regularity of the spawning season of some fishes. ICES Journal of Marine Science 33, 81–92.

Cushing, D.H., 1959. The seasonal variation in oceanic production as a problem in population dynamics. ICES Journal of Marine Science 24, 455–464.

Dakos, V., Beninca, E., van Nes, E.H., Philippart, C.J.M., Scheffer, M., Huisman, J., 2009. Interannual variability in species composition explained as seasonally entrained chaos. Proceedings of the Royal Society B: Biological Sciences 276, 2871–2880.

Damuth, J., 1981. Population density and body size in mammals. Nature 290, 699–700. https://doi.org/10.1038/290699a0

De Boyer Montégut, C., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. Journal of Geophysical Research 109, C12003. https://doi.org/10.1029/2004JC002378

De Cáceres, M., Legendre, P., Moretti, M., 2010. Improving indicator species analysis by combining groups of sites. Oikos 119, 1674–1684. https://doi.org/10.1111/j.1600-0706.2010.18334.x

De Tommasi, E., 2016. Light Manipulation by Single Cells: The Case of Diatoms. Journal of Spectroscopy 2016, 1–13. https://doi.org/10.1155/2016/2490128

De Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Jaillon, O., Lima-Mendez, G., Luke, J., Morard, R., Mulot, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., Karsenti, E., 2015. Eukaryotic plankton diversity in the sunlit ocean. Science 348, 1261605. https://doi.org/10.1126/science.1261605

De Wit, R., Bouvier, T., 2006. "Everything is everywhere, but, the environment selects"; what did Baas Becking and Beijerinck really say? Environmental Microbiology 8, 755–758. https://doi.org/10.1111/j.1462-2920.2006.01017.x

Delwiche, C.F., 1999. Tracing the Thread of Plastid Diversity through the Tapestry of Life. The American Naturalist 154, S164–S177. https://doi.org/10.1086/303291

Descamps-Julien, B., Gonzalez, A., 2005. Stable coexistence in a fluctuating environment: an experimental demonstration. Ecology 86, 2815–2824.

Dietrich, G., 1964. Oceanic Polar Front Survey in the North Atlantic. Research in Geophysics 2, 291–308.

Djeghri, N., Atkinson, A., Fileman, E.S., Harmer, R.A., Widdicombe, C., McEvoy, A.J., Cornwell, L., Mayor, D.J., 2018. High prey-predator size ratios and unselective feeding in copepods: A seasonal comparison of five species with contrasting feeding modes. Progress in Oceanography 165, 63–74. https://doi.org/10.1016/j.pocean.2018.04.013

Dolan, J.R., 2022. Pioneers of plankton research: Alister Hardy (1896–1985). Journal of Plankton Research 44, 477–485. https://doi.org/10.1093/plankt/fbac033

Dolédec, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche Separation in Community Analysis: A New Method. Ecology 81, 2914–2927.

Dubinsky, Z., Stambler, N., 2009. Photoacclimation processes in phytoplankton: mechanisms, consequences, and applications. Aquatic Microbial Ecology 56, 163–176.

Dufrêne, M., Legendre, P., 1997. Species Assemblages and Indicator Species: The Need for a Flexible Asymmetrical Approach. Ecological Monographs 67, 345. https://doi.org/10.2307/2963459

Dutkiewicz, S., Cermeno, P., Jahn, O., Follows, M.J., Hickman, A.E., Taniguchi, D.A.A., Ward, B.A., 2020. Dimensions of marine phytoplankton diversity. Biogeosciences 17, 609–634. https://doi.org/10.5194/bg-17-609-2020

Dutkiewicz, S., Scott, J.R., Follows, M.J., 2013. Winners and losers: Ecological and biogeochemical changes in a warming ocean. Global Biogeochemical Cycles 27, 463–477. https://doi.org/10.1002/gbc.20042

Dybwad, C., Assmy, P., Olsen, L.M., Peeken, I., Nikolopoulos, A., Krumpen, T., Randelhoff, A., Tatarek, A., Wiktor, J.M., Reigstad, M., 2021. Carbon Export in the Seasonal Sea Ice Zone North of Svalbard from Winter to Late Summer. Frontiers in Marine Science 7, 525800. https://doi.org/10.3389/fmars.2020.525800

Edwards, K.F., Litchman, E., Klausmeier, C.A., 2013. Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. Ecology Letters 16, 56–63. https://doi.org/10.1111/ele.12012

Edwards, M., Beaugrand, G., Hays, G.C., Koslow, J.A., Richardson, A.J., 2010. Multi-decadal oceanic ecological datasets and their application in marine policy and management. Trends in Ecology & Evolution 25, 602–610. https://doi.org/10.1016/j.tree.2010.07.007

Edwards, M., Beaugrand, G., Helaouët, P., Alheit, J., Coombs, S., 2013. Marine Ecosystem Response to the Atlantic Multidecadal Oscillation. PLoS ONE 8, e57212. https://doi.org/10.1371/journal.pone.0057212

Edwards, M., Beaugrand, G., Kléparski, L., Helaouët, P., Reid, P.C., 2022. Climate variability and multi-decadal diatom abundance in the Northeast Atlantic. Communications Earth & Environment 3, 162. https://doi.org/ttps://doi.org/10.1038/s43247-022-00492-9

Edwards, M., Beaugrand, G., Reid, P.C., Rowden, A.A., Jones, M.B., 2002. Ocean climate anomalies and the ecology of the North Sea. Marine Ecology Progress Series 239, 1–10. https://doi.org/10.3354/meps239001

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. Communications Biology 4, 644. https://doi.org/10.1038/s42003-021-02159-1

Edwards, M., Johns, D.G., Leterme, S.C., Svendsen, E., Richardson, A.J., 2006. Regional climate change and harmful algal blooms in the northeast Atlantic. Limnology and Oceanography 51, 820–829. https://doi.org/10.4319/lo.2006.51.2.0820

Edwards, M., Richardson, A.J., 2004. Impact of climate change on marine pelagic phenology and trophic mismatch. Nature 430, 881–884. https://doi.org/10.1038/nature02808

Egerton, F.N., 2012. Roots of Ecology, antiquity to Haeckel. University of California Press, Berkeley, Los Angeles, London.

Eilertsen, H.C., 1993. Spring blooms and stratification. Nature 363, 24.

Eilertsen, H.C., Degerlund, M., 2010. Phytoplankton and light during the northern high-latitude winter. Journal of Plankton Research 32, 899–912. https://doi.org/10.1093/plankt/fbq017

Elton, C., 1927. Animal ecology., Sidgwick and Jackson. ed. London.

Enfield, D.B., Mestas-Nuñez, A.M., Trimble, P.J., 2001. The Atlantic Multidecadal Oscillation and its relation to rainfall and river flows in the continental U.S. Geophysical Research Letters 28, 2077–2080.

Eppley, R.W., Sloan, P.R., 1966. Growth rates of marine phytoplankton: correlation with light absorption by cell chlorophyll a. Physiologia Plantarum 19, 47–59.

Estrada, M., Berdalet, E., 1997. Phytoplankton in a turbulent world. Scientia Marina 61, 125–140.

Eyring, V., Bony, S., Meehl, G.A., Senior, C.A., Stevens, B., Stouffer, R.J., Taylor, K.E., 2016. Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development 9, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Faillettaz, R., Beaugrand, G., Goberville, E., Kirby, R.R., 2019. Atlantic Multidecadal Oscillations drive the basin-scale distribution of Atlantic bluefin tuna. Science Advances 5, eaar6993. https://doi.org/10.1126/sciadv.aar6993

Falkowski, P.G., Katz, M.E., Knoll, A.H., Quigg, A., Raven, J.A., Schofield, O., Taylor, F.J.R., 2004. The Evolution of Modern Eukaryotic Phytoplankton. Science 305, 354–360. https://doi.org/10.1126/science.1095964

Falkowski, P.G., Oliver, M.J., 2007. Mix and match: how climate selects phytoplankton. Nature Reviews Microbiology 5, 813–819. https://doi.org/10.1038/nrmicro1751

Falkowski, P.G., Scholes, R.J., Boyle, E., Canadell, J., Canfield, D., Elser, J., Gruber, N., Hibbard, K., Högberg, P., Linder, S., Mackensie, F.T., Moore, B., Pedersen, T., Rosenthal, Y., Seitzinger, S., Smetacek, V., Steffen, 2000. The global carbon cycle: a test of our knowledge of Earth as a system. Science 290, 291–296.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P.G., 1998. Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. Science 281, 237–240.

Fileman, E., Petropavlovsky, A., Harris, R., 2010. Grazing by the copepods Calanus helgolandicus and Acartia clausi on the protozooplankton community at station L4 in the Western English Channel. Journal of Plankton Research 32, 709–724.

Fischer, A.D., Moberg, E.A., Alexander, H., Brownlee, E.F., Hunter-Cevera, K.R., Pitz, K.J., Rosengard, S.Z., Sosik, H.M., 2014. Sixty years of Sverdrup: A retrospective of progress in the study of phytoplankton blooms. Oceanography 27, 222–235.

Flori, S., Jouneau, P.-H., Bailleul, B., Gallet, B., Estrozi, L.F., Moriscot, C., Bastien, O., Eicke, S., Schober, A., Bártulos, C.R., Maréchal, E., Kroth, P.G., Petroutsos, D., Zeeman, S., Breyton, C., Schoehn, G., Falconet, D., Finazzi, G., 2017. Plastid thylakoid architecture optimizes photosynthesis in diatoms. Nature Communications 8, 15885. https://doi.org/10.1038/ncomms15885

Foden, J., Devlin, M.J., Mills, D.K., Malcolm, S.J., 2010. Searching for undesirable disturbance: an application of the OSPAR eutrophication assessment method to marine waters of England and Wales. Biogeochemistry 106, 157–175.

Fogg, G.E., 1990. Our perceptions of phytoplankton: an historical sketch the first Founders' Lecture. British Phycological Journal 25, 103–115. https://doi.org/10.1080/00071619000650101

Friedland, K.D., Mouw, C.B., Asch, R.G., Ferreira, A.S.A., Henson, S.A., Hyde, K.J.W., Morse, R.E., Thomas, A.C., Brady, D.C., 2018. Phenology and time series trends of the dominant seasonal phytoplankton bloom across global scales. Global Ecology and Biogeography 27, 551–569. https://doi.org/10.1111/geb.12717

Frontier, S., 1976. Etude de la décroissance des valeurs propres dans une analyse en composantes principales: comparaison avec le modèle du bâton brisé. Journal of Experimental Marine Biology and Ecology 25, 67–75.

Frouin, R., Franz, B.A., Werdell, P.J., 2003. The SeaWiFS PAR product, in: Hooker, S.B., Firestone, E.R. (Eds.), Algorithm Updates for the Fourth SeaWiFS Data Reprocessing. pp. 46–50.

Fu, W., Randerson, J.T., Moore, J.K., 2016. Climate change impacts on net primary production (NPP) and export production (EP) regulated by increasing stratification and phytoplankton communitystructure in the CMIP5 models. Biogeosciences 13, 5151–5170. https://doi.org/10.5194/bg-13-5151-2016

Fuhrmann, T., Landwehr, S., El Rharbi-Kucki, M., Sumper, M., 2004. Diatoms as living photonic crystals. Applied Physics B 78, 257–260. https://doi.org/10.1007/s00340-004-1419-4

Garcia, H.E., Locarnini, R.A., Boyer, T.P., Antonov, J.I., Baranova, O.K., Zweng, M.M., Reagan, J.R., Johnson, D.R., 2014. World ocean atlas 2013: Dissolved inorganic nutrients

(phosphate, nitrate, silicate). NOAA Atlas NESDIS 76, 25. https://doi.org/10.7289/V5J67DWD

Gauch, H.G., Chase, J.M., Whittaker, R.H., 1974. Ordination of vegetation samples by Gaussian species distribution. Ecology 55, 1382–1390.

Gause, G.F., 1934. The struggle for existence, Williams and Wilkins. ed. Baltimore.

Geider, R., MacIntyre, H., Kana, T., 1997. Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a: carbon ratio to light, nutrient-limitation and temperature. Marine Ecology Progress Series 148, 187–200. https://doi.org/10.3354/meps148187

Ghilarov, A.M., 1984. The Paradox of the Plankton Reconsidered; Or, Why Do Species Coexist? Oikos 43, 46–52. https://doi.org/10.2307/3544244

Gittings, J.A., Raitsos, D.E., Krokos, G., Hoteit, I., 2018. Impacts of warming on phytoplankton abundance and phenology in a typical tropical marine ecosystem. Scientific Reports 8, 2240. https://doi.org/10.1038/s41598-018-20560-5

Goberville, E., Beaugrand, G., Hautekeete, N.C., Piquot, Y., Luczak, C., 2015. Uncertainties in species distribution projections and general circulation models. Ecology and Evolution 5, 1100–1116. https://doi.org/10.1002/ece3.1411

Goessling, J.W., Su, Y., Cartaxana, P., Maibohm, C., Rickelt, L.F., Trampe, E.C.L., Walby, S.L., Wangpraseurt, D., Wu, X., Ellegaard, M., Kühl, M., 2018. Structure-based optics of centric diatom frustules: modulation of the *in vivo* light field for efficient diatom photosynthesis. New Phytologist 219, 122–134. https://doi.org/10.1111/nph.15149

Goetze, E., 2003. Cryptic speciation on the high seas; global phylogenetics of the copepod family Eucalanidae. Proceedings of the Royal Society B: Biological Sciences 270, 2321–2331. https://doi.org/10.1098/rspb.2003.2505

Goldman, J.C., 1980. Physiological processes, nutrient availability, and the concept of relative growth rate in marine phytoplankton ecology, in: Falkowski, P.G. (Ed.), Primary Productivity in the Sea. Plenum, New York, pp. 179–194.

Goss, R., Lepetit, B., 2015. Biodiversity of NPQ. Journal of Plant Physiology 172, 13–32.

Gray, S.T., Graumlich, L.J., Betancourt, J.L., Pederson, G.T., 2004. A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation since 1567 A.D. Geophysical Research Letters 31, 2–5.

Grinnell, J., 1917. The niche-relations of the California thrasher. Auk 34, 427–433.

Grover, J.P., 1989. Influrence of cell shape and size on algal competitive ability. Journal of Phycology 25, 402–405. https://doi.org/10.1111/j.1529-8817.1989.tb00138.x

Haeckel, E., 1904. Kunstformen der Natur. Bibliographilchen Institust, Leipzig, Wien.

Hardy, A.C., 1958. The open sea: the world of plankton. Collins, London.

Hardy, A.C., 1926. A new method of plankton research. Nature 118, 630.

Hátún, H., Azetsu-Scott, K., Somavilla, R., Rey, F., Johnson, C., Mathis, M., Mikolajewicz, U., Coupel, P., Tremblay, J.-É., Hartman, S., Pacariz, S.V., Salter, I., Ólafsson, J., 2017. The subpolar gyre regulates silicate concentrations in the North Atlantic. Scientific Reports 7, 14576. https://doi.org/10.1038/s41598-017-14837-4

Hays, G.C., Warner, A.J., 1993. Consistency of Towing Speed and Sampling Depth for the Continuous Plankton Recorder. Journal of the Marine Biological Association of the United Kingdom 73, 967–970. https://doi.org/10.1017/S0025315400034846

Helaouët, P., 2021. Marine Biological Association of the UK (MBA) (2021): Continuous Plankton Recorder data. The Archive for Marine Species and Habitats Data (DASSH). https://doi.org/10.17031/1708

Helaouët, P., Beaugrand, G., 2009. Physiology, Ecological Niches and Species Distribution. Ecosystems 12, 1235–1245. https://doi.org/10.1007/s10021-009-9261-5

Helaouët, P., Beaugrand, G., 2007. Macroecology of *Calanus finmarchicus* and *C. helgolandicus* in the North Atlantic Ocean and adjacent seas. Marine Ecology Progress Series 345, 147–165. https://doi.org/10.3354/meps06775

Helaouët, P., Beaugrand, G., Reygondeau, G., 2016. Reliability of spatial and temporal patterns of *C. finmarchicus* inferred from the CPR survey. Journal of Marine Systems 153, 18–24. https://doi.org/10.1016/j.jmarsys.2015.09.001

Henson, S.A., Cael, B.B., Allen, S.R., Dutkiewicz, S., 2021. Future phytoplankton diversity in a changing climate. Nature Communications 12, 5372. https://doi.org/10.1038/s41467-021-25699-w

Henson, S.A., Cole, H., Beaulieu, C., Yool, A., 2013. The impact of global warming on seasonality of ocean primary production. Biogeosciences 10, 4357–4369. https://doi.org/10.5194/bg-10-4357-2013

Henson, S.A., Cole, H.S., Hopkins, J., Martin, A.P., Yool, A., 2018. Detection of climate change-driven trends in phytoplankton phenology. Global Change Biology 24, e101–e111. https://doi.org/DOI: 10.1111/gcb.13886

Henson, S.A., Lampitt, R., Johns, D., 2012. Variability in phytoplankton community structure in response to the North Atlantic Oscillation and implications for organic carbon flux. Limnology and Oceanography 57, 1591–1601. https://doi.org/10.4319/lo.2012.57.6.1591

Henson, S.A., Laufkötter, C., Leung, S., Giering, S.L.C., Palevsky, H.I., Cavan, E.L., 2022. Uncertain response of ocean biological carbon export in a changing world. Nature Geoscience 15, 248–254. https://doi.org/10.1038/s41561-022-00927-0

Henson, S.A., Sarmiento, J.L., Dunne, J.P., Bopp, L., Lima, I., Doney, S.C., John, J., Beaulieu, C., 2010. Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. Biogeosciences 7, 621–640. https://doi.org/10.5194/bg-7-621-2010

Hinder, S.L., Hays, G.C., Edwards, M., Roberts, E.C., Walne, A.W., Gravenor, M.B., 2012. Changes in marine dinoflagellate and diatom abundance under climate change. Nature Climate Change 2, 271–275. https://doi.org/10.1038/nclimate1388

Hirahara, S., Ishii, M., Fukuda, Y., 2014. Centennial-scale sea surface temperature analysis and its uncertainty. 27, 57–75.

Holligan, P.M., Maddock, L., Dodge, J.D., 1980. The distribution of dinoflagellates around the British Isles in July 1977: a multivariate analysis. Journal of the Marine Biological Association of the United Kingdom 60, 851–867.

Huang, B., Thorne, P.W., Banzon, V.F., Boyer, T., Chepurin, G., Lawrimore, J.H., Menne, M.J., Smith, T.M., Vose, R.S., Zhang, H.M., 2017. Extended reconstructed sea surface temperature, version 5 (ERSSTv5): Upgrades, validations, and intercomparisons. Journal of Climate 30, 8179–8205.

Hubbell, S.P., 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton.

Huisman, J., Oostveen, P., Weissing, F.J., 1999. Critical depth and critical turbulence: two different mechanisms for the development of phytoplankton blooms. Limnology and Oceanography 44, 1781–1787.

Huisman, J., Weissing, F.J., 1999. Biodiversity of plankton by species oscillations and chaos. Nature 402, 407–410.

Hutchinson, G.E., 1978. An introduction to population ecology, Yale University Press. ed. New Haven.

Hutchinson, G.E., 1961. The paradox of the plankton. The American Naturalist 95, 137–145. https://doi.org/10.1086/282171

Hutchinson, G.E., 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22, 415–427.

Huthnance, J.M., 1991. Physical oceanography of the North Sea. Ocean & Coastal Management 16, 199–231.

IOCCG, 2000. Remote sensing of ocean colour in coastal and optically-complex waters. (No. 3), Reports of the International Ocean-Colour Coordination Group. Dartmooth, Canada.

IPCC, 2021. Summary for policymakers, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3–32.

IPCC, 2019. Technical Summary, in: Roberts, D.C., Pörtner, H.O., Masson-Delmotte, V., Zhai, P., Poloczanska, E., Mintenbeck, K., Tignor, M., Alegria, A., Nicolai, M., Okem, A., Petzold, J., Rama, B., Weyer, N.M. (Eds.), IPCC Special Report on the Ocean and Cryosphere in a Changing Climate. Cambridge University Press, Cambridge, UK and New York, NY, US, pp. 39–69.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press. ed. Cambridge, United Kingdom and New York, NY, USA.

Irwin, A.J., Nelles, A.M., Finkel, Z.V., 2012. Phytoplankton niches estimated from field data. Limnology and Oceanography 57, 787–797. https://doi.org/10.4319/lo.2012.57.3.0787

Jackson, D.A., Somers, K.M., 1989. Are probability estimates from the permutation model of Mantel's test stable? Canadian Journal of Zoology 67, 766–769. https://doi.org/10.1139/z89-108

Jackson, S.T., Overpeck, J.T., 2000. Responses of plant populations and communities to environmental changes of the late Quaternary. Paleobiology 26, 194–220. https://doi.org/10.1017/S0094837300026932

Jacobsen, A., Egge, J.K., Heimdal, B.R., 1995. Effects of increased concentration of nitrate and phosphate during a springbloom experiment in mesocosm. Journal of Experimental Marine Biology and Ecology 187, 239–251. https://doi.org/10.1016/0022-0981(94)00183-E

Jin, X., Gruber, N., Dunne, J.P., Sarmiento, J.L., Armstrong, R.A., 2006. Diagnosing the contribution of phytoplankton functional groups to the production and export of particulate organic carbon, $CaCO_3$, and opal from global nutrient and alkalinity distributions. Global Biogeochemical Cycles 20, GB2015. https://doi.org/10.1029/2005GB002532

Johnson, G.C., Schmidtko, S., Lyman, J.M., 2012. Relative contributions of temperature and salinity to seasonal mixed layer density changes and horizontal density gradients. Journal of Geophysical Research 117, C04015. https://doi.org/10.1029/2011JC007651

Jolliffe, I.T., 1986. Principal Component Analysis, series in statistics. Springer-Verlag New York Inc. Springer, New York, NY.

Jonas, T.D., Walne, A.W., Beaugrand, G., Gregory, L., Hays, G.C., 2004. The volume of water filtered by a Continuous Plankton Recorder sample: the effect of ship speed. Journal of Plankton Research 26, 1499–1506. https://doi.org/10.1093/plankt/fbh137

Jónasdóttir, S.H., Visser, A.W., Richardson, K., Heath, M.R., 2015. Seasonal copepod lipid pump promotes carbon sequestration in the deep North Atlantic. Proceedings of the National Academy of Sciences 112, 12122–12126. https://doi.org/10.1073/pnas.1512110112

Jones, P.D., New, M., Parker, D.E., Martin, S., Rigor, I.G., 1999. Surface air temperature and its changes over the past 150 years. Reviews of Geophysics 37, 173–199.

Junker, R.R., Kuppler, J., Bathke, A.C., Schreyer, M.L., Trutschnig, W., 2016. *Dynamic range boxes* – a robust nonparametric approach to quantify size and overlap of *n* -dimensional hypervolumes. Methods in Ecology and Evolution 7, 1503–1513. https://doi.org/10.1111/2041-210X.12611

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woolen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., Joseph, D., 1996. The NCEP/NCAR 40-Year reanalysis project. Bulletin of the American Meteorological Society 77, 437–471.

Kamykowski, D., Zentara, S.J., 2005. Changes in world ocean nitrate availability through the 20th century. Deep Sea Research Part I: Oceanographic Research Papers 52, 1719–1744.

Kamykowski, D., Zentara, S.J., 1986. Predicting plant nutrient concentrations from temperature and sigma-T in the upper kilometer of the world ocean. Deep Sea Research Part I: Oceanographic Research Papers 33, 89–105.

Karlusich, J.J.P., Ibarbalz, F.M., Bowler, C., 2020. Exploration of marine phytoplankton: from their historical appreciation to the omics era. Journal of Plankton Research fbaa049. https://doi.org/10.1093/plankt/fbaa049

Karp-Boss, L., Boss, E., 2016. The Elongated, the Squat and the Spherical: Selective Pressures for Phytoplankton Shape, in: Glibert, P.M., Kana, T.M. (Eds.), Aquatic Microbial Ecology and Biogeochemistry: A Dual Perspective. Springer International Publishing, Cham, pp. 25–34. https://doi.org/10.1007/978-3-319-30259-1_3

Karp-Boss, L., Boss, E., Jumars, P.A., 2000. Motion of dinoflagellates in simple shear flow. Limnology and Oceanography 45, 1594–1602.

Käse, R.H., Zenk, W., 1996. Structure of the Mediterranean water and meddy characteristics in the Northeastern Atlantic, in: Krauß, W. (Ed.), The Warmwatersphere of the North Atlantic Ocean. Bornträger, Berlin, Germany, pp. 365–395.

Katz, M.E., Wright, J.D., Miller, K.G., Cramer, B.S., Fennel, K., Falkowski, P.G., 2005. Biological overprint of the geological carbon cycle. Marine Geology 217, 323–338.

Keerthi, M.G., Lévy, M., Aumont, O., 2021. Intermittency in phytoplankton bloom triggered by modulations in vertical stability. Scientific reports 11, 1285. https://doi.org/10.1038/s41598-020-80331-z

Keil, P., Mauritsen, T., Jungclaus, J., Hedemann, C., Olonscheck, D., Ghosh, R., 2020. Multiple drivers of the North Atlantic warming hole. Nature Climate Change 10, 667–671. https://doi.org/10.1038/s41558-020-0819-8

Kemp, A.E.S., Villareal, T.A., 2018. The case of the diatoms and the muddled mandalas: Time to recognize diatom adaptations to stratified waters. Progress in Oceanography 167, 138–149. https://doi.org/10.1016/j.pocean.2018.08.002

Kemp, A.E.S., Villareal, T.A., 2013. High diatom production and export in stratified waters – A potential negative feedback to global warming. Progress in Oceanography 119, 4–23. https://doi.org/10.1016/j.pocean.2013.06.004

Kenitz, K., Visser, A., Mariani, P., Andersen, K.H., 2017. Seasonal succession in zooplankton feeding traits reveals trophic trait coupling. Limnology and Oceanography 62, 1184–1197.

Kirby, R.R., 2010. Ocean drifters: a secret world beneath the waves. Studio Cactus Book, Singapore.

Kirby, R.R., Beaugrand, G., 2009. Trophic amplification of climate warming. Proceedings of the Royal Society B: Biological Sciences 276, 4095–4103. https://doi.org/10.1098/rspb.2009.1320

Kirby, R.R., Beaugrand, G., Lindley, J.A., 2008. Climate-induced effects on the meroplankton and the benthic-pelagic ecology of the North Sea. Limnology and Oceanography 53, 1805–1815. https://doi.org/10.4319/lo.2008.53.5.1805

Kivi, K., Kaitala, S., Kuosa, H., Kuparinen, J., Leskinen, E., Lignell, R., Marcussen, B., Tamrninen, T., 1993. Nutrient limitation and grazing control of the Baltic planktonic community during annual succession. Limnology and Oceanography 38, 893–905.

Kléparski, L., Beaugrand, G., 2022. The species chromatogram, a new graphical method to represent, characterise and compare the ecological niches of different species. Ecology and Evolution 12, e8830. https://doi.org/10.1002/ece3.8830

Kléparski, L., Beaugrand, G., Edwards, M., 2021. Plankton biogeography in the North Atlantic Ocean and its adjacent seas: Species assemblages and environmental signatures. Ecology and Evolution 11, 5135–5149. https://doi.org/10.1002/ece3.7406

Kléparski, L., Beaugrand, G., Edwards, M., Ostle, C., 2022a. Climate change and phytoplankton phenology in the North Atlantic. In prep.

Kléparski, L., Beaugrand, G., Edwards, M., Schmitt, F.G., Kirby, R.R., Breton, E., Gevaert, F., Maniez, E., 2022b. Morphological traits, niche-environment interaction and temporal changes in diatoms. Progress in Oceanography 201, 102747. https://doi.org/10.1016/j.pocean.2022.102747

Kléparski, L., Beaugrand, G., Kirby, R.R., 2022c. How do plankton species coexist in an apparently unstructured environment? Biology Letters 18. https://doi.org/10.1098/rsbl.2022.0207

Koeller, P., Fuentes-Yaco, C., Platt, T., Sathyendranath, S., Richards, A., Ouellet, P., Orr, D., Skúladóttir, U., Wieland, K., Savard, L., Aschan, M., 2009. Basin-scale coherence in phenology of schrimps and phytoplankton in the North Atlantic Ocean. Science 324, 791–793. https://doi.org/10.1126/science.1170987

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J.R., Dunne, J.P., Gehlen, M., Ilyina, T., John, J.G., Lenton, A., Li, H., Lovenduski, N.S., Orr, J.C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C.A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., Ziehn, T., 2020. Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections. Biogeosciences 17, 3439–3470. https://doi.org/10.5194/bg-17-3439-2020

Lam, N.S.N., 1983. Spatial interpolation methods: A review. The American Cartographer 10, 129–150.

Lavaud, J., 2007. Fast regulation of photosynthesis in diatoms: Mechanisms, evolution and ecophysiology. Functional Plant Science and Biotechnology 1, 267–287.

Lawrenz, E., Richardson, T.L., 2017. Differential effects of changes in spectral irradiance on photoacclimation, primary productivity and growth in *Rhodomonas salina* (cryptophyceae) and *Skeletonema costatum* (bacillariophyceae) in simulated black water environment. Journal of Phycology 53, 1241–1254.

Leblanc, K., Arístegui, J., Armand, L., Assmy, P., Beker, B., Bode, A., Breton, E., Cornet, V., Gibson, J., Gosselin, M.-P., Kopczynska, E., Marshall, H., Peloquin, J., Piontkovski, S., Poulton, A.J., Quéguiner, B., Schiebel, R., Shipe, R., Stefels, J., van Leeuwe, M.A., Varela, M., Widdicombe, C., Yallop, M., 2012. A global diatom database – abundance, biovolume and biomass in the world ocean. Earth System Science Data 4, 149–165. https://doi.org/10.5194/essd-4-149-2012

Lee, Z., Weidemann, A., Kindle, J., Arnone, R., Carder, K.L., Davis, C., 2007. Euphotic zone depth: Its derivation and implication to ocean-color remote sensing. Journal of Geophysical Research 112, C03009. https://doi.org/10.1029/2006JC003802

Lee, Z.-P., 2005. A model for the diffuse attenuation coefficient of downwelling irradiance. Journal of Geophysical Research 110, C02016. https://doi.org/10.1029/2004JC002275

Legendre, L., 1990. The significance of microalgal blooms for fisheries and for the export of particulate organic carbon in oceans. Journal of Plankton Research 12, 681–699.

Legendre, P., Legendre, L., 1998. Numerical Ecology. Elsevier, Amsterdam - Lausanne - New York -Oxford - Shannon - Singapore - Tokyo.

Levin, S.A., 1970. Community Equilibria and Stability, and an Extension of the Competitive Exclusion Principle. The American Naturalist 104, 413–423. https://doi.org/10.1086/282676

Levin, S.A., Lubchenco, J., 2008. Resilience, robustness, and marine ecosystem-based management. Bioscience 58, 27–32.

Levins, R., 1979. Coexistence in a Variable Environment. The American Naturalist 114, 765–783. https://doi.org/10.1086/283527

Lewandowska, A., Sommer, U., 2010. Climate change and the spring bloom: a mesocosm study on the influence of light and temperature on phytoplankton and mesozooplankton. Marine Ecology Progress Series 405, 101–111. https://doi.org/10.3354/meps08520

Lewis, W.M., 1976. Surface/Volume Ratio: Implications for Phytoplankton Morphology. Science 192, 885–887. https://doi.org/10.1126/science.192.4242.885

Ligowski, R., Jordan, R.W., Assmy, P., 2012. Morphological adaptation of a planktonic diatom to growth in Antarctic sea ice. Marine Biology 159, 817–827.

Litchman, E., de Tezanos Pinto, P., Klausmeier, C.A., Thomas, M.K., Yoshiyama, K., 2010. Linking traits to species diversity and community structure in phytoplankton. Hydrobiologia 653, 15–28. https://doi.org/10.1007/s10750-010-0341-5

Litchman, E., Edwards, K.F., Klausmeier, C.A., Thomas, M.K., 2012. Phytoplankton niches, traits and eco-evolutionary responses to global environmental change. Marine Ecology Progress Series 470, 235–248. https://doi.org/doi: 10.3354/meps09912

Litchman, E., Klausmeier, C.A., 2008. Trait-Based Community Ecology of Phytoplankton. Annual Review of Ecology, Evolution, and Systematics 39, 615–639.

Litchman, E., Klausmeier, C.A., Schofield, O.M., Falkowski, P.G., 2007. The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. Ecology Letters 10, 1170–1181. https://doi.org/10.1111/j.1461-0248.2007.01117.x

Longhurst, A., 1998. Ecological geography of the sea. Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto.

López-Urrutia, A., San Martin, E., Harris, R.P., Irigoien, X., 2006. Scaling the metabolic balance of the oceans. Proceedings of the National Academy of Sciences 103, 8739–44.

Luczak, C., Beaugrand, G., Jaffré, M., Lenoir, S., 2011. Climate change impact on Balearic shearwater through a trophic cascade. Biology Letters 7, 702–705. https://doi.org/10.1098/rsbl.2011.0225

Mac Arthur, R., Levins, R., 1964. Competition, habitat selection and character displacement in a patchy environment. Proceedings of the National Academy of Sciences 51, 1207–1210.

MacIntyre, H.L., Kana, T.M., Geider, R.J., 2000. The effect of water motion on short-term rates of photosynthesis by marine phytoplankton. Trends in plant science 5, 12–17. https://doi.org/10.1016/S1360-1385(99)01504-6

Magurran, A.E., 2004. Measuring biological diversity. Blackwell Science Ltd.

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., Wincker, P., Iudicone, D., de Vargas, C., Bittner, L., Zingone, A., Bowler, C., 2016. Insights into global diatom distribution and diversity in the world's ocean. Proceedings of the National Academy of Sciences 113, 1516–1525. https://doi.org/10.1073/pnas.1509523113

Mann, D.G., 1999. The species concept in diatoms. Phycologia 38, 437–495.

Mann, D.G., 1996. Chloroplast morphology, movements and inheritance in diatoms., in: Chaudhary, B.R., Agrawal, S.B. (Eds.), Cytology, Genetics and Molecular Biology of Algae. SPB Academic Publishing., Amsterdam, The Netherlands, pp. 249–274.

Mann, K.H., Lazier, J.R.N., 1996. Dynamics of marine ecosystems: biological-physical interactions in the oceans, 2nd ed. Blackwell Science, Incorporated, Oxford.

Mann, M.E., Steinman, B.A., Miller, S.K., 2020. Absence of internal multidecadal and interdecadal oscillations in climate model simulations. Nature Communications 11, 1–9.

Margalef, R., 1979. Functional morphology of organisms involved in red tides, as adapted to decaying turbulence. Toxic dinoflagellate blooms 1, 89–94.

Margalef, R., 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. Oceanologica acta 1, 493–509.

Marinov, I., Doney, S.C., Lima, I.D., 2010. Response of ocean phytoplankton community structure to climate change over the 21st century: partitioning the effects of nutrients, temperature and light. Biogeosciences 7, 3941–3959. https://doi.org/10.5194/bg-7-3941-2010

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

McMinn, A., Martin, A., 2013. Dark survival in a warming world. Proceedings of the Royal Society B: Biological Sciences 280, 20122909. https://doi.org/10.1098/rspb.2012.2909

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. Marine Ecology Progress Series 339, 301–306. https://doi.org/10.3354/meps339301

McQuoid, M.R., Hobson, L.A., 1996. Diatom resting stages. Journal of Phycology 32, 889–902. https://doi.org/10.1111/j.0022-3646.1996.00889.x

Metaxas, A., Scheibling, R.E., 1996. Top-down and bottom-up regulation of phytoplankton assemblages in tidepools. Marine Ecology Progress Series 145, 161–177.

Miller, C.B., 2004. Biological Oceanography. Blackwell Science Ltd.

Morel, A., Huot, Y., Gentili, B., Werdell, P.J., Hooker, S.B., Franz, B.A., 2007. Examining the consistency of products derived from various ocean color sensors in open ocean

(Case 1) waters in the perspective of a multi-sensor approach. Remote Sensing of Environment 111, 69–88. https://doi.org/10.1016/j.rse.2007.03.012

Naselli-Flores, L., Barone, R., 2011. Invited Review - Fight on Plankton! or, Phytoplankton Shape and Size as Adaptive Tools to Get Ahead in the Struggle for Life. Cryptogamie, Algologie 32, 157–204. https://doi.org/10.7872/crya.v32.iss2.2011.157

Naselli-Flores, L., Zohary, T., Padisák, J., 2021. Life in suspension and its impact on phytoplankton morphology: an homage to Colin S. Reynolds. Hydrobiologia 848, 7–30. https://doi.org/10.1007/s10750-020-04217-x

Nelson, D.M., Tréguer, P., Brzezinski, M.A., Leynaert, A., Quéguiner, B., 1995. Production and dissolution of biogenic silica in the ocean: Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. Global Biogeochemical Cycles 9, 359–372. https://doi.org/10.1029/95GB01070

Noyes, J., Sumper, M., Vukusic, P., 2008. Light manipulation in a marine diatom. Journal of Materials Research 23, 3229–3235. https://doi.org/10.1557/JMR.2008.0381

O'Neill, B.C., Kriegler, E., Ebi, K.L., Kemp-Benedict, E., Riahi, K., Rothman, D.S., van Ruijven, B.J., van Vuuren, D.P., Birkmann, J., Kok, K., Levy, M., Solecki, W., 2017. The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. Global Environmental Change 42, 169–180. https://doi.org/10.1016/j.gloenvcha.2015.01.004

O'Neill, B.C., Tebaldi, C., van Vuuren, D.P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J., Lowe, J., Meehl, G.A., Moss, R., Riahi, K., Sanderson, B.M., 2016. The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. Geoscientific Model Development 9, 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

O'Reilly, C.H., Zanna, L., Woollings, T., 2019. Assessing external and internal sources of Atlantic multidecadal variability using models, proxy data, and early instrumental indices. Journal of Climate 32, 7727–7745.

Owens, N.J.P., Hosie, G.W., Batten, S.D., Edwards, M., Johns, D.G., Beaugrand, G., 2013. All plankton sampling systems underestimate abundance: Response to "Continuous plankton recorder underestimates zooplankton abundance" by J.W. Dippner and M. Krause. Journal of Marine Systems 128, 240–242. https://doi.org/10.1016/j.jmarsys.2013.05.003

Padisak, J., Soroczki-Pinter, E., Rezner, Z., 2003. Sinking properties of some phytoplankton shapes and the relation of form resistance to morphological diversity of plankton – an experimental study. Hydrobiologia 500, 243–257.

Pahlow, M., Riebesell, U., Wolf-Gladrow, D.A., 1997. Impact of cell shape and chain formation on nutrient acquisition by marine diatoms. Limnology and Oceanography 42, 1660–1672. https://doi.org/10.4319/lo.1997.42.8.1660

Pančić, M., Kiørboe, T., 2018. Phytoplankton defence mechanisms: traits and trade-offs: Defensive traits and trade-offs. Biological Reviews 93, 1269–1303. https://doi.org/10.1111/brv.12395

Passow, U., Carlson, C.A., 2012. The biological pump in a high CO2 world. Marine Ecology Progress Series 470, 249–271. https://doi.org/10.3354/meps09985

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., Torres, T.J., 1998. Fishing down marine food webs. Science 279, 860–863.

Peeters, J.C.H., Haas, H.A., Peperzak, L., Vries, I., 1993. Nutrients and light as factors controlling phytoplankton biomass on the Dutch Continental Shelf (North Sea) in 1988-1990. (No. DGW-93.004). Rijkswaterstaat Tidal Waters division, Middelburg.

Perry, A.I., Low, P.J., Ellis, J.R., Reynolds, J.D., 2005. Climate change and distribution shifts in marine fishes. Science 308, 1912–1915.

Petersen, R., 1975. The Paradox of the Plankton: An Equilibrium Hypothesis. The American Naturalist 109, 35–49. https://doi.org/10.1086/282972

Pilson, M.E.Q., 2013. An Introduction to the Chemistry of the Sea., second edition. ed. Cambridge University Press.

Pingree, R.D., Holligan, P.M., Mardell, G.T., 1978. The effects of vertical stability on phytoplankton distributions in the summer on the northwest European Shelf. Deep Sea Research 25, 1011–1028. https://doi.org/10.1016/0146-6291(78)90584-2

Planque, B., Hays, G.C., Ibañez, F., Gamble, J.C., 1997. Large scale spatial variations in the seasonal abundance of *Calanus finmarchicus*. Deep Sea Research Part I: Oceanographic Research Papers 44, 315–326.

Platt, T., Fuentes-Yaco, C., Frank, K.T., 2003. Spring algal bloom and larval fish survival. Nature 423, 398–399.

Poloczanska, E.S., Brown, C.J., Sydeman, W.J., Kiessling, W., Schoeman, D.S., Moore, P.J., Brander, K., Bruno, J.F., Buckley, L.B., Burrows, M.T., Duarte, C.M., Halpern, B.S., Holding, J., Kappel, C.V., O'Connor, M.I., Pandolfi, J.M., Parmesan, C., Schwing, F., Thompson, S.A., Richardson, A.J., 2013. Global imprint of climate change on marine life. Nature Climate Change 3, 919–925. https://doi.org/10.1038/nclimate1958

Poloczanska, E.S., Burrows, M.T., Brown, C.J., García Molinos, J., Halpern, B.S., Hoegh-Guldberg, O., Kappel, C.V., Moore, P.J., Richardson, A.J., Schoeman, D.S., Sydeman, W.J., 2016. Responses of Marine Organisms to Climate Change across Oceans. Frontiers in Marine Science 3. https://doi.org/10.3389/fmars.2016.00062

Pulliam, H.R., 2000. On the relationship between niche and distribution. Ecology Letters 3, 349–361. https://doi.org/10.1046/j.1461-0248.2000.00143.x

Pulliam, H.R., 1988. Sources, sinks, and population regulation. The American Naturalist 132, 652–661.

Qiao, H., Escobar, L.E., Saupe, E.E., Ji, L., Soberón, J., 2017. A cautionary note on the use of hypervolume kernel density estimators in ecological niche modelling: Kernel density methods in ecological niche modelling. Global Ecology and Biogeography 26, 1066–1070. https://doi.org/10.1111/geb.12492

Qiao, H., Soberón, J., Peterson, A.T., 2015. No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. Methods in Ecology and Evolution 6, 1126–1136. https://doi.org/10.1111/2041-210X.12397

Qin, M., Dai, A., Hua, W., 2020. Quantifying contributions of internal variability and external forcing to atlantic multidecadal variability since 1870. Geophysical Research Letters 47, 1–11.

Racault, M., Le Quéré, C., Buitenhuis, E., Sathyendranath, S., Platt, T., 2012. Phytoplankton phenology in the global ocean. Ecological Indicators 14, 152–163. https://doi.org/10.1016/j.ecolind.2011.07.010

Ras, M., Steyer, J.P., Bernard, O., 2013. Temperature effect on microalgae: a crucial factor for outdoor production. Reviews in Environmental Science and Bio/Technology 12, 153–164.

Raven, J.A., Waite, A.M., 2004. The evolution of silicification in diatoms: inescapable sinking and sinking as escape? New Phytologist 162, 45–61. https://doi.org/10.1111/j.1469-8137.2004.01022.x

Raven, P., Johnson, G., Losos, J., Singer, S., 2005. Biology. Mc Graw-Hill.

Read, J.F., Pollard, R.T., 1992. Water masses in the region of the Iceland-Faeroes front. Journal of Physical Oceanography 22, 1365–1378.

Redfield, A., 1958. The biological control of chemical factors in the environment. American Scientist 46, 205–221.

Reid, P., Holliday, N., Smyth, T., 2001. Pulses in the eastern margin current and warmer water off the north west European shelf linked to North Sea ecosystem changes. Marine Ecology Progress Series 215, 283–287. https://doi.org/10.3354/meps215283

Reid, P.C., Colebrook, J.M., Matthews, J.B.L., Aiken, J., Continuous Plankton Recorder Team, 2003. The Continuous Plankton Recorder: concepts and history, from Plankton Indicator to undulating recorders. Progress in Oceanography 58, 117–173. https://doi.org/10.1016/j.pocean.2003.08.002

Reid, P.C., Edwards, M., 2001. Plankton And Climate, in: Encyclopedia of Ocean Sciences. Elsevier, pp. 2194–2200. https://doi.org/10.1006/rwos.2001.0267

Reid, P.C., Edwards, M., Hunt, H.G., Warner, A.J., 1998. Phytoplankton change in the North Atlantic. Nature 391, 546–546. https://doi.org/10.1038/35290

Reid, P.C., Johns, D.G., Edwards, M., Starr, M., Poulin, M., Snoeijs, P., 2007. A biological consequence of reducing Arctic ice cover: arrival of the Pacific diatom Neodenticula seminae in the North Atlantic for the first time in 800 000 years. Global Change Biology 13, 1910–1921. https://doi.org/10.1111/j.1365-2486.2007.01413.x

Reid, P.C., Lancelot, W.W.C., Gieskes, E., Hagmeier, E., Weickart, G., 1990. Phytoplankton of the North Sea and its dynamics: a review. Netherlands Journal of Sea Research 26, 295–331.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., Maury, O., 2013. Dynamic biogeochemical provinces in the global ocean. Global Biogeochemical Cycles 27, 1046–1058. https://doi.org/10.1002/gbc.20089

Reygondeau, G., Maury, O., Beaugrand, G., Fromentin, J.M., Fonteneau, A., Cury, P., 2012. Biogeography of tuna and billfish communities. Journal of Biogeography 39, 114–129. https://doi.org/10.1111/j.1365-2699.2011.02582.x

Reynolds, C.S., 1998. Plants in motion: Physical - biological interaction in the plankton, in: Imberger, J. (Ed.), Coastal and Estuarine Studies. American Geophysical Union, Washington, D. C., pp. 535–560. https://doi.org/10.1029/CE054p0535

Reynolds, C.S., 1989. Physical determinants of phytoplankton succession, in: Sommer, U. (Ed.), Plankton Ecology Succession in Plankton Communities. Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo, p. 377.

Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C., Wang, W., 2002. An improved in situ and satellite SST analysis for climate. Journal of Climate 15, 1609–1625.

Reynolds, R.W., Smith, T.M., Liu, C., Chelton, D.B., Casey, K.S., Schlax, M.G., 2007. Daily High-Resolution-Blended Analyses for Sea Surface Temperature. Journal of Climate 20, 5473–5496. https://doi.org/10.1175/2007JCLI1824.1

Richardson, A.J., Poloczanska, E.S., 2008. OCEAN SCIENCE: Under-Resourced, Under Threat. Science 320, 1294–1295. https://doi.org/10.1126/science.1156129

Richardson, A.J., Schoeman, D.S., 2004. Climate Impact on Plankton Ecosystems in the Northeast Atlantic. Science 305, 1609–1612.

Richardson, A.J., Walne, A.W., John, A.W.G., Jonas, T.D., Lindley, J.A., Sims, D.W., Stevens, D., Witt, M., 2006. Using continuous plankton recorder data. Progress in Oceanography 68, 27–74. https://doi.org/10.1016/j.pocean.2005.09.011

Richerson, P., Armstrong, R., Goldman, C.R., 1970. Contemporaneous Disequilibrium, a New Hypothesis to Explain the "Paradox of the Plankton." Proceedings of the National Academy of Sciences 67, 1710–1714. https://doi.org/10.1073/pnas.67.4.1710

Righetti, D., Vogt, M., Gruber, N., Psomas, A., Zimmermann, N.E., 2019. Global pattern of phytoplankton diversity driven by temperature and environmental variability. Science Advances 5, eaau6253. https://doi.org/10.1126/sciadv.aau6253

Riley, G.A., 1967. The plankton of estuaries. Estuaries 83, 316–326.

Romagnan, J.-B., Legendre, L., Guidi, L., Jamet, J.-L., Jamet, D., Mousseau, L., Pedrotti, M.-L., Picheral, M., Gorsky, G., Sardet, C., Stemmann, L., 2015. Comprehensive Model of Annual Plankton Succession Based on the Whole-Plankton Time Series Approach. PLOS ONE 10, e0119219. https://doi.org/10.1371/journal.pone.0119219

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., Legendre, L., 2010. A multivariate approach to large-scale variation in marine planktonic copepod diversity and its environmental correlates. Limnology and Oceanography 55, 2219–2229. https://doi.org/10.4319/lo.2010.55.5.2219

Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., Legendre, L., 2009. Global latitudinal variations in marine copepod diversity and environmental factors. Proceedings of the Royal Society B: Biological Sciences 276, 3053–3062. https://doi.org/10.1098/rspb.2009.0742

Rothhaupt, K.O., 1988. Mechanistic resource competition theory applied to laboratory experiments with zooplankton. Nature 333, 660–662.

Ruddiman, W.F., 1969. Recent planktonic foraminifera: dominance and diversity in North Atlantic surface sediments 164, 1164–1167.

Rutherford, S., D'Hondt, S., Prell, W., 1999. Environmental controls on the geographic distribution of zooplankton diversity. Nature 400, 749–753. https://doi.org/10.1038/23449

Ryabov, A., Kerimoglu, O., Litchman, E., Olenina, I., Roselli, L., Basset, A., Stanca, E., Blasius, B., 2021. Shape matters: the relationship between cell geometry and diversity in phytoplankton. Ecology Letters 24, 847–861. https://doi.org/10.1111/ele.13680

Rynearson, T.A., 2013. Major contribution of diatom resting spores to vertical flux in the sub-polar North Atlantic. Deep Sea Research Part I: Oceanographic Research Papers 82, 60–71.

Sallée, JB., Pellichero, V., Akhoudas, C., Pauthenet, E., Vignes, L., Schmidtko, S., Garabato, A.N., Sutherland, P., Kuusela, M., 2021. Summertime increases in upper-ocean stratification and mixed-layer depth. Nature 591, 592–598. https://doi.org/10.1038/s41586-021-03303-x

Schwaderer, A.S., Yoshiyama, K., de Tezanos Pinto, P., Swenson, N.G., Klausmeier, C.A., Litchman, E., 2011. Eco-evolutionary differences in light utilization traits and distributions of freshwater phytoplankton. Limnology and Oceanography 56, 589–598. https://doi.org/10.4319/lo.2011.56.2.0589

Séférian, R., Berthet, S., Yool, A., Palmiéri, J., Bopp, L., Tagliabue, A., Kwiatkowski, L., Aumont, O., Christian, J., Dunne, J., Gehlen, M., Ilyina, T., John, J.G., Li, H., Long, M.C., Luo, J.Y., Nakano, H., Romanou, A., Schwinger, J., Stock, C., Santana-Falcón, Y., Takano, Y., Tjiputra, J., Tsujino, H., Watanabe, M., Wu, T., Wu, F., Yamamoto, A., 2020. Tracking Improvement in Simulated Marine Biogeochemistry Between CMIP5 and CMIP6. Current Climate Change Reports 6, 95–119. https://doi.org/10.1007/s40641-020-00160-0

Sharples, J., Ross, O.N., Scott, B.E., Greenstreet, S.P.R., Fraser, H., 2006. Inter-annual variability in the timing of stratification and the spring bloom in the northwestern North Sea. Continental Shelf Research 26, 733–751.

Shaw, P.J., Purdie, D.A., 2001. Phytoplankton photosynthesis-irradiance parameters in the near-shore UK coastal waters of the North Sea: temporal variation and environmental control. Marine Ecology Progress Series 216, 83–94.

Shelford, V.E., 1913. Animal communities in a temperate America., University of Chicago Press. ed. Chicago.

Silva, T., Gislason, A., Licandro, P., Marteinsdóttir, G., Ferreira, A.S.A., Gudmundsson, K., Astthorsson, O.S., 2014. Long-term changes of euphausiids in shelf and oceanic habitats southwest, south and southeast of Iceland. Journal of Plankton Research 36, 1262–1278. https://doi.org/10.1093/plankt/fbu050

Silvertown, J., 2004. Plant coexistence and the niche. Trends in Ecology & Evolution 19, 605–611. https://doi.org/10.1016/j.tree.2004.09.003

Simon, N., Cras, A.-L., Foulon, E., Lemée, R., 2009. Diversity and evolution of marine phytoplankton. Comptes Rendus Biologies 332, 159–170. https://doi.org/10.1016/j.crvi.2008.09.009

Smayda, T., 1970. The suspension and sinking of phytoplankton in the sea. Oceanography and Marine Biology - An Annual Review 8, 353–414.

Smetacek, V., 2001. A watery arms race. Nature 411, 745.

Smetacek, V., 1999. Diatoms and the ocean carbon cycle. Protist 150, 25–32.

Smetacek, V., 1985. Role of sinking in diatom life-history cycles: ecological, evolutionary and geological significance. Marine Biology 84, 239–251.

Smyth, T.J., Allen, I., Atkinson, A., Bruun, J.T., Harmer, R.A., Pingree, R.D., Widdicombe, C.E., Somerfield, P.J., 2014. Ocean Net Heat Flux Influences Seasonal to Interannual Patterns of Plankton Abundance. PLOS ONE 9, e98709.

Soberon, J., Nakamura, M., 2009. Niches and distributional areas: Concepts, methods, and assumptions. Proceedings of the National Academy of Sciences 106, 19644–19650. https://doi.org/10.1073/pnas.0901637106

Soberón, J., Peterson, A.T., 2020. What is the shape of the fundamental Grinnellian niche? Theoretical Ecology 13, 105–115. https://doi.org/10.1007/s12080-019-0432-5

Sommer, U., 1998. Silicate and the functional geometry of marine phytoplankton. Journal of Plankton Research 20, 1853–1859. https://doi.org/10.1093/plankt/20.9.1853

Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J.J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J.C., Mooij, W.M., van Donk, E., Winder, M., 2012. Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. Annual Review of Ecology, Evolution, and Systematics 43, 429–448. https://doi.org/10.1146/annurev-ecolsys-110411-160251

Sommer, U., Gliwicz, Z., Lampert, W., Duncan, A., 1986. The Peg-Model of seasonal succession of planktonic events in fresh waters. Archiv für Hydrobiologie 106, 433–471.

Sournia, A., 1982. Form and function in marine phytoplankton. Biological Reviews 57, 347–394.

Southwood, T.R.E., May, R.M., Hassell, M.P., Conway, G.R., 1974. Ecological strategies and population parameters. The American Naturalist 108, 791–804.

Stanca, E., Cellamare, M., Basset, A., 2013. Geometric shape as a trait to study phytoplankton distributions in aquatic ecosystems. Hydrobiologia 701, 99–116. https://doi.org/10.1007/s10750-012-1262-2

Stevens, G.C., 1989. The Latitudinal Gradient in Geographical Range: How so Many Species Coexist in the Tropics. The American Naturalist 133, 240–256. https://doi.org/10.1086/284913

Stomp, M., Huisman, J., de Jongh, F., Veraart, A.J., Gerla, D., Rijkeboer, M., Ibelings, B.W., Wollenzien, U.I.A., Stal, L.J., 2004. Adaptive divergence in pigment composition promotes phytoplankton biodiversity. Nature 432, 104–107. https://doi.org/10.1038/nature03044

Straile, D., Kerimoglu, O., Peeters, F., 2015. Trophic mismatch requires seasonal heterogeneity of warming. Ecology 96, 2794–2805. https://doi.org/10.1890/14-0839.1

Suggett, D.J., Goyen, S., Evenhuis, C., Szabo, M., Pettay, D.T., Warner, M.E., Ralph, P.J., 2015. Functional diversity of photobiological traits within the genus Symbiodinium appears to be governed by the interaction of cell size with cladal designation. The New phytologist 208, 370–381.

Sunday, J.M., Bates, A.E., Dulvy, N.K., 2011. Global analysis of thermal tolerance and latitude in ectotherms. Proceedings of the Royal Society B: Biological Sciences 278, 1823–1830. https://doi.org/10.1098/rspb.2010.1295

Sverdrup, H.U., 1953. On Conditions for the Vernal Blooming of Phytoplankton. ICES Journal of Marine Science 18, 287–295. https://doi.org/10.1093/icesjms/18.3.287

Swinehart, D.F., 1962. The Beer-Lambert law. Journal of Chemical Education 39, 333–335.

Taucher, J., Oschlies, A., 2011. Can we predict the direction of marine primary production change under global warming? Geophysical Research Letters 38, L02603. https://doi.org/10.1029/2010GL045934

Ter Braak, C.J.F., 1996. Unimodal models to relate species to environment. Agricultural Mathematics group, Wageningen, The Netherlands.

Ter Braak, C.J.F., Prentice, I.C., 1988. A Theory of Gradient Analysis, in: Advances in Ecological Research. Elsevier, pp. 271–317. https://doi.org/10.1016/S0065-2504(08)60183-X

Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Helaouet, P., Johns, D.G., Jones, I.D., Leech, D.I., Mackay, E.B., Massimino, D., Atkinson, S., Bacon, P.J., Brereton, T.M., Carvalho, L., Clutton-Brock, T.H., Duck, C., Edwards, M., Elliott, J.M., Hall, S.J.G., Harrington, R., Pearce-Higgins, J.W., Høye, T.T., Kruuk, L.E.B., Pemberton, J.M., Sparks, T.H., Thompson, P.M., White, I., Winfield, I.J., Wanless, S., 2016. Phenological sensitivity to climate across taxa and trophic levels. Nature 535, 241–245. https://doi.org/10.1038/nature18608

Thomas, M.K., Kremer, C.T., Klausmeier, C.A., Litchman, E., 2012. A Global Pattern of Thermal Adaptation in Marine Phytoplankton. Science 338, 1085–1088. https://doi.org/10.1126/science.1224836

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD - A platform for ensemble forecasting of species distributions. Ecography 32, 369–373.

Tilman, D., 1999. The ecological consequences of changes in biodiversity: a search for general principles. Ecology 80, 1455–1474.

Tilman, D., 1994. Competition and biodiversity in spatially structured habitats. Ecology 75, 2–16.

Tilman, D., 1980. Resources: A Graphical-Mechanistic Approach to Competition and Predation. The American Naturalist 116, 362–393. https://doi.org/10.1086/283633

Tilman, D., 1977. Resource competition between planktonic algae: an experimental and theoretical approach. Ecology 58, 338–348.

Tréguer, P., Bowler, C., Moriceau, B., Dutkiewicz, S., Gehlen, M., Aumont, O., Bittner, L., Dugdale, R., Finkel, Z., Iudicone, D., Jahn, O., Guidi, L., Lasbleiz, M., Leblanc, K., Levy, M., Pondaven, P., 2018. Influence of diatom diversity on the ocean biological carbon pump. Nature Geoscience 11, 27–37. https://doi.org/10.1038/s41561-017-0028-x

Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., DeClerck, O., 2012. Bio-ORACLE: A global environmental dataset for marine species distribution modelling. 21, 272–281.

van der Spoel, S., 1994. The basis for boundaries in pelagic biogeography. Progress in Oceanography 34, 121–133. https://doi.org/10.1016/0079-6611(94)90005-1

Visser, A.W., Jonsson, P.R., 2000. On the reorientation of non-spherical prey particles in a feeding current. Journal of Plankton Research 22, 761–777.

Wackernagel, H., 1995. Multivariate geostatistics. An introduction with applications. Springer-Verlag, Berlin, Heidelberg.

Ward, B.A., Follows, M.J., 2016. Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. Proceedings of the National Academy of Sciences 113, 2958–2963. https://doi.org/10.1073/pnas.1517118113

Warner, A.J., Hays, G.C., 1994. Sampling by the continuous plankton recorder survey. Progress in Oceanography 34, 237–256. https://doi.org/10.1016/0079-6611(94)90011-6

Whittaker, R.H., 1975. Communities and ecosystems, 2nd ed. Macmillan, New York.

Widdicombe, C.E., Eloire, D., Harbour, D., Harris, R.P., Somerfield, P.J., 2010. Long-term phytoplankton community dynamics in the Western English Channel. Journal of Plankton Research 32, 643–655. https://doi.org/10.1093/plankt/fbp127

Wihsgott, J.U., Sharples, J., Hopkins, J.E., Woodward, E.M.S., Hull, T., Greenwood, N., Sivyer, D.B., 2019. Observations of vertical mixing in autumn and its effect on the autumn phytoplankton bloom. Progress in Oceanography 177, 102059. https://doi.org/10.1016/j.pocean.2019.01.001

Winder, M., Cloern, J., 2010. The annual cycles of phytoplankton biomass. Philosophical Transactions of the Royal Society B: Biological Sciences 365, 3215–3226.

Winder, M., Schindler, D.E., 2004. Climate change uncouples trophic interactions in an aquatic ecosystem. Ecology 85, 2100–2106. https://doi.org/10.1890/04-0151

Worthington, L.V., 1976. On the North Atlantic circulation. The Hopkins University Press.

Yamaguchi, R., Rodgers, K.B., Stein, K.J., Timmermann, A., Schlunegger, S., Bianchi, D., Dunne, J.P., Slater, R.D., 2022. Trophic level decoupling drives future change in phytoplankton bloom phenology. Nature Climate Change 12, 469–476. https://doi.org/10.1038/s41558-022-01353-1

Yan, W., Hunt, L.A., 1999. An Equation for Modelling the Temperature Response of Plants using only the Cardinal Temperatures. Annals of Botany 84, 607–614. https://doi.org/10.1006/anbo.1999.0955

Zacaï, A., Monnet, C., Pohl, A., Beaugrand, G., Mullins, G., Kroeck, D.M., Servais, T., 2021. Truncated bimodal latitudinal diversity gradient in early Paleozoic phytoplankton. Science Advances 7, eabd6709. https://doi.org/10.1126/sciadv.abd6709

Zanna, L., Khatiwalab, S., Gregory, J.M., Ison, J., Heimbach, P., 2019. Global reconstruction of historical ocean heat storage and transport. Proceedings of the National Academy of Sciences 116, 1126–1131. https://doi.org/10.1073/pnas.1808838115

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., Walne, A., 2013. The response of phytoplankton to climate variability associated with the North Atlantic Oscillation. Deep Sea Research Part II: Topical Studies in Oceanography 93, 159–168.

Zobel, M., 1997. The relative role of species pools in determining plant species richness: an alternative explanation of species coexistence? Tree 12, 266–269.

# Publications from this PhD

Journal of
Plankton Research
academic.oup.com/plankt

*J. Plankton Res.* (2021) 43(1): 85–102. First published online December 22, 2020 doi:10.1093/plankt/fbaa060

ORIGINAL ARTICLE

# Annual phytoplankton succession results from niche-environment interaction

MARIARITA CARACCIOLO[1,*], GRÉGORY BEAUGRAND[2,3,4,*], PIERRE HÉLAOUËT[4], FRANCOIS GEVAERT[3], MARTIN EDWARDS[4,5], FABRICE LIZON[3], LOÏCK KLÉPARSKI[3,4] AND ERIC GOBERVILLE[6]

[1]SORBONNE UNIVERSITÉ, CNRS, STATION BIOLOGIQUE DE ROSCOFF, UMR 7144, ECOMAP, PLACE GEORGES TEISSIER, 29680 ROSCOFF, FRANCE, [2]CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS), UNIVERSITÉ DE LILLE, UNIVERSITÉ LITTORAL CÔTE D'OPALE, UMR 8187, LOG, LABORATOIRE D'OCÉANOLOGIE ET DE GÉOSCIENCES, F 62930 WIMEREUX, FRANCE, [3]UNIVERSITÉ DE LILLE, CNRS, UNIV. LITTORAL CÔTE D'OPALE, UMR 8187, LOG, LABORATOIRE D'OCÉANOLOGIE ET DE GÉOSCIENCES, F 62930 WIMEREUX, FRANCE, [4]MARINE BIOLOGICAL ASSOCIATION, CITADEL HILL, PLYMOUTH PL1 2PB, UK, [5]SCHOOL OF BIOLOGICAL AND MARINE SCIENCES, UNIVERSITY OF PLYMOUTH, DRAKE CIRCUS, PLYMOUTH, PL4 8AA, UNITED KINGDOM AND [6]UNITÉ BIOLOGIE DES ORGANISMES ET ÉCOSYSTÈMES AQUATIQUES (BOREA), MUSÉUM NATIONAL D'HISTOIRE NATURELLE, SORBONNE UNIVERSITÉ, UNIVERSITÉ DE CAEN NORMANDIE, UNIVERSITÉ DES ANTILLES, CNRS, IRD, CP53, 61, RUE BUFFON 75005 PARIS, FRANCE

*CORRESPONDING AUTHORS: mariarita.caracciolo@sb-roscoff.fr (MARIARITA CARACCIOLO) AND gregory.beaugrand@univ-lille.fr (GREGORY BEAUGRAND)

Received February 29, 2020; editorial decision November 12, 2020; accepted November 16, 2020

Corresponding editor: Lisa Campbell

Annual plankton succession has been investigated for many decades with hypotheses ranging from abiotic to biotic mechanisms being proposed to explain these recurrent patterns. Here, using data collected by the Continuous Plankton Recorder (CPR) survey and models originating from the MacroEcological Theory on the Arrangement of Life, we investigate Annual Phytoplankton Succession (APS) in the North Sea at a species level. Our results show that this phenomenon can be predicted well by models combining photosynthetically active radiation, temperature and macronutrients. Our findings suggest that APS originates from the interaction between species' ecological niches and the annual environmental fluctuations at a community level. We discuss our results in the context of traditional hypotheses formulated to explain this recurrent pattern in the marine field.

KEYWORDS: annual plankton succession; phenology; ecological niche; environment; plankton; continuous plankton recorder (CPR); METAL theory

available online at academic.oup.com/plankt

Progress in Oceanography 201 (2022) 102747

Contents lists available at ScienceDirect

Progress in Oceanography

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

# Morphological traits, niche-environment interaction and temporal changes in diatoms

Loïck Kléparski [a,b,*], Grégory Beaugrand [a,*], Martin Edwards [c,d], François G. Schmitt [a], Richard R. Kirby [e,f], Elsa Breton [a], François Gevaert [a], Emeline Maniez [a]

[a] Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France
[b] Marine Biological Association, Citadel Hill, Plymouth PL1 2PB, United Kingdom
[c] Plymouth Marine Laboratory, Prospect Place, Plymouth PL13DH, United Kingdom
[d] University of Plymouth, School of Biological and Marine Sciences, Drake Circus, Plymouth, United Kingdom
[e] The Secchi Disk Foundation, Kiln Cottage, Gnaton, Yealmpton PL8 2HU, United Kingdom
[f] Ronin Institute, Montclair, NJ 07043, USA

ARTICLE INFO

*Keywords:*
Annual diatom succession
Cell elongation
Ecological niche
Morphological traits
Phenology
North Sea
CPR survey

ABSTRACT

Annual phytoplankton succession is a key ecological phenomenon that drives marine species' life cycles and energy flows within marine ecosystems. Identifying processes that control annual succession is critical to anticipate climate-induced environmental perturbations of this phenomenon and the consequences upon ecosystem functioning. Here, we demonstrate that diatoms in the North Sea undergo strong morphological changes throughout the year and that species with similar phenology possess comparable morphological traits (e.g. cell elongation) and ecological niches. The spring and autumn periods appear to be dominated by oblates (flattened cells) whereas the summer period is dominated by prolates (elongated cells). Elongation of the cell shape enhances buoyancy and confers a selective advantage in stratified low-nutrient waters typical of summer without changing a diatom's surface area/volume ratio or its ability to absorb nutrients. Diatom shape thus appears to have evolved as a key adaptation to a specific environment and confers upon a species its specific niche and phenology, and therefore its place in the sequence of annual succession. As a result, shape influences temporal changes in the abundance of diatoms and their putative response to environmental forcing. We suggest that biogeochemical and earth-system models should include diatom cell shape as a parameter in order to improve model predictions and help our understanding of the consequences of climate change on marine ecosystems.

## 1. Introduction

Diatoms are a major phytoplankton clade that have evolved during the Middle Triassic when sea levels were rising and continental margins were flooding (Falkowski et al., 2004). Diatoms' subsequent increase in diversity during the mid-Cenozoic Era strongly influenced the global carbon cycle (Katz et al., 2005). Today, diatoms account for about 40% of total marine primary production and almost 40% of total particulate organic carbon exported to the deep ocean annually (Jin et al., 2006; Nelson et al., 1995; Tréguer et al., 2018). Annually, the abundance of diatoms rises and falls on a seasonal basis following a predictable pattern named "Annual Diatom Succession" (called hereafter ADS).

In the North Sea, ADS is characterised by the succession of various species throughout the year, with species such as *Skeletonema costatum* and *Thalassiothrix longissima* dominating in spring, *Guinardia striata* and *G. flaccida* in summer and *Poroboscia indica* and *Bidulphia alternans* in autumn (Caracciolo et al., 2021). ADS therefore determines the pulses of energy that influence the dynamics of the whole marine ecosystem, the life cycle of many zooplankton and fish being coupled with the peaks in primary production or species dominance (Cushing, 1990; Platt et al., 2003). Recently, anthropogenic climate change has begun to alter diatom phenology and biogeography, and this is likely to have strong consequences on trophic interactions and ecosystem functioning (Chivers et al., 2017; Edwards and Richardson, 2004).

* Corresponding authors at: Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France (L. Kléparski).
*E-mail addresses:* loick.kleparski@hotmail.fr (L. Kléparski), gregory.beaugrand@univ-lille.fr (G. Beaugrand).

https://doi.org/10.1016/j.pocean.2022.102747
Received 17 May 2021; Received in revised form 5 December 2021; Accepted 26 January 2022
Available online 1 February 2022

Received: 9 February 2021 | Accepted: 20 February 2021

DOI: 10.1002/ece3.7406

ORIGINAL RESEARCH

Ecology and Evolution Open Access WILEY

# Plankton biogeography in the North Atlantic Ocean and its adjacent seas: Species assemblages and environmental signatures

Loïck Kléparski[1,2] | Grégory Beaugrand[1] | Martin Edwards[3]

[1]CNRS, UMR 8187 – LOG – Laboratoire d'Océanologie et de Géosciences, Univ. Littoral Côte d'Opale, Univ. Lille, Wimereux, France

[2]Continuous Plankton Recorder (CPR) Survey, The Marine Biological Association, Plymouth, UK

[3]Marine Institute, Plymouth University, Plymouth, UK

**Correspondence**
Loïck Kléparski, CNRS, UMR 8187 – LOG – Laboratoire d'Océanologie et de Géosciences, Univ. Littoral Côte d'Opale, Univ. Lille, F-62930 Wimereux, France.
Email: loick.kleparski@hotmail.fr

## Abstract

Plankton biodiversity is a key component of marine pelagic ecosystems. They are at the base of the food web, control the productivity of marine ecosystems, and provide many provisioning and regulating ecological services. It is therefore important to understand how plankton are organized in both space and time. Here, we use data of varying taxonomic resolution, collected by the Continuous Plankton Recorder (CPR) survey, to map phytoplankton and zooplankton biodiversity in the North Atlantic and its adjacent seas. We then decompose biodiversity into 24 species assemblages and investigate their spatial distribution using ecological units and ecoregions recently proposed. Finally, we propose a descriptive method, which we call the environmental chromatogram, to characterize the environmental signature of each plankton assemblage. The method is based on a graphic that identifies where species of an assemblage aggregate along an environmental gradient composed of multiple ecological dimensions. The decomposition of the biodiversity into species assemblages allows us to show (a) that most marine regions of the North Atlantic are composed of coenoclines (i.e., gradients of biocoenoses or communities) and (b) that the overlapping spatial distribution of assemblages is the result of their environmental signatures. It follows that neither the ecoregions nor the ecological units identified in the North Atlantic are characterized by a unique assemblage but instead by a mosaic of assemblages that overlap in many places.

KEYWORDS

biogeography, coenoclines, environmental signature, macroecology, North Atlantic Ocean, plankton, taxonomic assemblages

## 1 | INTRODUCTION

Plankton are a key component of marine pelagic ecosystems controlling their productivity (Edwards et al., 2013). Phytoplankton produce by photosynthesis almost half of the oxygen at the global scale (Behrenfeld, 2014). They create endosomatic energy that is progressively channeled through the whole marine food web. Zooplankton ensure the transfer of this energy between phytoplankton and higher trophic levels such as fish. Any changes in the abundance and composition of plankton affect higher trophic levels (Edwards et al., 2013; Luczak et al., 2011). Plankton also control a part of carbon exportation in the North Atlantic by a process termed

Received: 14 March 2022 | Revised: 29 March 2022 | Accepted: 30 March 2022

DOI: 10.1002/ece3.8830

RESEARCH ARTICLE

Ecology and Evolution Open Access WILEY

# The species chromatogram, a new graphical method to represent, characterize, and compare the ecological niches of different species

Loïck Kléparski[1,2] | Grégory Beaugrand[1]

[1]UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, Université du Littoral Côte d'Opale, CNRS, Université de Lille, Wimereux, France

[2]Marine Biological Association, Plymouth, UK

**Correspondence**
Loïck Kléparski, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, Université du Littoral Côte d'Opale, CNRS, Université de Lille, Wimereux F-62930, France.
Email: loick.kleparski@hotmail.fr

## Abstract

The ecological niche *sensu* Hutchinson is defined as the set of environmental conditions allowing a species to grow, maintain, and reproduce. This conception of the niche, which is assimilated to a *p*-dimensional hypervolume, with *p* representing all environmental variables, has been widely applied in ecology. However, displaying the niche hypervolume has proved challenging when more than three environmental dimensions are considered simultaneously. We propose a simple method (implemented in the *specieschrom* R package) that displays the full multidimensionality of the ecological niche of a species into a two-dimensional space by means of a graphic we call species chromatogram. This method gives a graphical summary of the niche by representing together abundance gradients with respect to all environmental variables. A chromatogram enables niche optimums and breaths to be rapidly quantified, and when several chromatograms are examined (one per species), rapid comparisons can be made. From our chromatograms, we proposed a procedure that quantifies niche optimum and breadth as well as niche overlapping (index *D*) and the identification of the most discriminant combination of environmental variables. We apply these analyses on eight planktonic species collected by the Continuous Plankton Recorder (CPR) survey in the North Atlantic Ocean using 10 environmental variables. We display their full multidimensional niches and quantify their niche optimums and breadths along each dimension. We also compare our index *D* with other indices by means of *hypervolume* and *dynRB* R packages. By catching the full complexity of the niche, species chromatograms allow many different niche properties to be rapidly assessed and compared among species from niche optimums and breadths to the identification of the most relevant environmental parameters and the degree of niche overlapping among species. Species chromatograms may be seen as species' fingerprint and may also allow a better identification of the mechanisms involved in species assembly.

**KEYWORDS**
degree of niche overlapping, ecological niche, gradient analysis, hypervolume, niche breadth, niche optimum

BIOLOGY LETTERS

royalsocietypublishing.org/journal/rsbl

Research

Cite this article: Kléparski L, Beaugrand G, Kirby RR. 2022 How do plankton species coexist in an apparently unstructured environment? *Biol. Lett.* **18**: 20220207. https://doi.org/10.1098/rsbl.2022.0207

Received: 28 April 2022
Accepted: 30 June 2022

**Subject Areas:**
ecology, ecosystems, environmental science

**Keywords:**
plankton paradox, niche theory, biodiversity, pelagic environment, principle of competitive exclusion

**Author for correspondence:**
Loïck Kléparski
e-mail: loick.kleparski@hotmail.fr

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.6080899.

THE ROYAL SOCIETY PUBLISHING

**Marine biology**

# How do plankton species coexist in an apparently unstructured environment?

Loïck Kléparski[1,2], Grégory Beaugrand[1] and Richard R. Kirby[3,4]

[1]Laboratoire d'Océanologie et de Géosciences, University Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 – LOG, Wimereux F-62930, France
[2]Marine Biological Association, Citadel Hill, Plymouth PL1 2PB, UK
[3]The Secchi Disk Foundation, Kiln Cottage, Gnaton, Yealmpton PL8 2HU, UK
[4]Ronin Institute, Montclair, NJ 07043, USA

LK, 0000-0002-7536-2941; GB, 0000-0002-0712-5223; RRK, 0000-0002-9867-4454

In a paper entitled *The paradox of the plankton*, Hutchinson asked 'how it is possible for a number of species to coexist in a relatively isotropic or unstructured environment all competing for the same sorts of materials' (Hutchinson 1961 *Am. Nat.* **95**, 137–145 (doi:10.1086/282171)). Particularly relevant for phytoplankton, this paradox was based on two implicit, and perhaps naive, postulates, i.e. (i) that all plankton species have similar requirements and (ii) that the marine environment is relatively homogeneous in space and time. A number of hypotheses, based on purely theoretical or experimental studies, have been proposed to solve this conundrum, ranging from spatio-temporal environmental heterogeneity to biotic chaotic variability. Here, we characterize the ecological niche of 117 plankton species belonging to three different taxonomic groups and show that all species have a niche sufficiently distinct to ensure coexistence in a structured marine environment. We also provide evidence that pelagic habitats are, unsurprisingly, more diverse in space and time than Hutchinson imagined, the marine environment being neither unstructured nor stable in space and time. We, therefore, conclude that the niche theory, and its corollary the principle of competitive exclusion, apply as much for the plankton as for other forms of life, be they terrestrial or marine.

## 1. Introduction

Marine plankton is composed of both uni- and multicellular organisms covering a wide range of sizes, from less than 0.2 µm (virioplankton) to a few metres (e.g. the giant jellyfish *Nemopilema nomurai*) [1]. They are at the basis of the food web and play a key role in carbon dioxide regulation, influencing the composition of the Earth's atmosphere [1–4]. The great diversity they exhibit has fascinated marine biologists for a long time [5–7] and it is now apparent that there is even more diversity than originally realized [8,9]. The mechanisms sustaining plankton biodiversity remain debated, however, because it is assumed that plankton species coexist in an apparently homogeneous environment with few niches (*sensu* Hutchinson realized niche i.e. the set of environmental conditions modulated by biotic interactions, enabling a species to growth and reproduce [10]) thought to be available [11,12].

Although it has been suggested that plankton species may share niches and coexist [13], many theoretical and laboratory studies have tried to resolve the 'paradox of the plankton' [11] and have shown that the number of coexisting plankton species at equilibrium in a community cannot exceed the number of limiting resources [14–17]. Consequently, coexistence of a large number of species on a small number of limiting resources is only possible when other mechanisms are involved such as (i) temporal heterogeneity (i.e. a continuously changing pelagic environment due to the seasonal changes in hydro-meteorological forcing) and its influence on the structure of the water column

# Climate change and phytoplankton phenology in the North Atlantic

Loïck Kléparski [1,2 *], Grégory Beaugrand [1 *], Martin Edwards[3, 4], Clare Ostle[2]

[1] Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France

[2] Marine Biological Association, Citadel Hill, Plymouth PL1 2PB, United Kingdom.

[3] Plymouth Marine Laboratory, Prospect Place, Plymouth PL13DH, United Kingdom.

[4] University of Plymouth, School of Biological and Marine Sciences, Drake Circus, Plymouth, United Kingdom.

**Key-words:** diatoms, oblates, prolates, dinoflagellates, phenology, annual phytoplankton succession, climate change

**Abstract**

Significant phenological shifts induced by climate change are predicted within the phytoplanktonic community[1–3]. However, predictions from current Earth System Models (ESMs) understandably rely on simplified community responses that do not consider evolutionary strategies manifested as various life phenotypes and trait groups. Here, we use a species-based modelling approach, combined with large-scale plankton observations, to investigate past, contemporary and future phenological shifts in diatoms (grouped by their morphological traits) and dinoflagellates in three key areas of the North Atlantic Ocean (North Sea, North-East Atlantic and Labrador Sea) from 1850-2100. Our study reveals that the groupings exhibit coherent and different shifts in phenology and abundance throughout the North Atlantic Ocean. The phenology of large flattened diatoms is predicted to shrink and their abundance to decline, whereas the phenology of slow-sinking elongated diatoms and of dinoflagellates is expected to expand and their abundance to rise, which may alter carbon export in this important sink region. Our results show new understanding of the possible consequences of phytoplankton shifts on the biological carbon pump, suggesting that future ESMs should increase their level of complexity to include ecological strategies adapted by various phytoplankton groups.

## 1. Introduction

Current Earth System Models (ESMs) are predicting a decline in phytoplankton abundance[4–7] and anticipate major phenological shifts[1–3,8] that may lead to trophic desynchronization and community reorganisation[7–10]. Phenological shifts in phytoplankton are expected because of ocean warming and

communications
earth & environment

ARTICLE

https://doi.org/10.1038/s43247-022-00492-9 OPEN

# Climate variability and multi-decadal diatom abundance in the Northeast Atlantic

Martin Edwards[1✉], Gregory Beaugrand[2], Loïck Kléparski[3,4], Pierre Hélaouët[3] & Philip C. Reid[3]

Diatoms are important contributors to marine primary production and the ocean carbon cycle. In the North Atlantic and its adjacent seas primary production is driven by diatoms that transfer a significant part of the produced energy to higher trophic levels and carbon to the deep ocean. Anthropogenic warming and climate variability will likely have important consequences for the productivity and spatial dynamics of these eukaryotic phytoplankton. Using multidecadal diatom abundance data (>60 years) for the Northeast Atlantic and the North Sea, we show significant spatial and temporal correlations over these scales between diatoms and climate variability. A general multidecadal trend is established where climate warming is increasing diatom populations in northerly systems but decreasing populations in more southerly systems. We discover major phase shifts in diatom abundance synchronous with multi-decadal trends in Atlantic climate variability that occurred after the mid-1990s.

[1] Plymouth Marine Laboratory, Prospect Place, Plymouth PL1 3DH, UK. [2] Centre National de la Recherche Scientifique, Laboratoire d'Océanologie et de Géosciences' UMR LOG CNRS 8187, Station Marine, Université des Sciences et Technologies, de Lille 1 - Lille 1 BP 80, 62930 Wimereux, France. [3] The Marine Biological Association (MBA), The Laboratory, Citadel Hill, Plymouth PL12PB, UK. [4] Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, F-62930 Wimereux, France. ✉email: maed@pml.ac.uk

# Other publications

frontiers in Marine Science

ORIGINAL RESEARCH
published: 06 September 2021
doi: 10.3389/fmars.2021.656300

# Seasonal Variations in the Biodiversity, Ecological Strategy, and Specialization of Diatoms and Copepods in a Coastal System With *Phaeocystis* Blooms: The Key Role of Trait Trade-Offs

*Elsa Breton[1]*, Urania Christaki[1], Benoit Sautour[2], Oscar Demonio[3], Dimitra-Ioli Skouroliakou[1], Gregory Beaugrand[1], Laurent Seuront[1,4,5], Loïck Kléparski[1,6], Adrien Poquet[1,7], Antoine Nowaczyk[8], Muriel Crouvoisier[1], Sophie Ferreira[9], David Pecqueur[10], Christophe Salmeron[10], Jean-Michel Brylinski[1], Arnaud Lheureux[8] and Eric Goberville[3]*

[1] Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 LOG, Wimereux, France, [2] Univ. Bordeaux, CNRS, UMR 5805 EPOC, Rue Geoffroy Saint Hilaire – Bâtiment, Pessac, France, [3] 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, [4] Department of Marine Resources and Energy, Tokyo University of Marine Science and Technology, Tokyo, Japan, [5] Department of Zoology and Entomology, Rhodes University, Grahamstown, South Africa, [6] Marine Biological Association, Citadel Hill, Plymouth, United Kingdom, [7] Univ. Côte d'Azur, CNRS, INSERM, IRCAN, Medical School of Nice, Nice, France, [8] Univ. Bordeaux, CNRS, UMR 5805 EPOC, Station Marine d'Arcachon, Arcachon, France, [9] Observatoire Océanologique de Banyuls s/mer, FR 3724 – Laboratoire Arago – SU/CNRS, Banyuls-sur-Mer, France, [10] Univ. Bordeaux, CNRS, OASU, UMS 2567 POREA, Allée Geoffroy Saint-Hilaire, Pessac, France

Although eutrophication induced by anthropogenic nutrient enrichment is a driver of shifts in community composition and eventually a threat to marine biodiversity, the causes and consequences on ecosystem functioning remain greatly unknown. In this study, by applying a trait-based approach and measuring niche breadth of diatoms and copepods, the drivers and underlying mechanisms of the seasonal species succession of these ecological communities in a coastal system dominated in spring by *Phaeocystis* blooms were explored. It is suggested that the seasonal succession of diatoms and copepods is the result of several trade-offs among functional traits that are controlled by the seasonal abiotic and biotic pressure encountered by the plankton communities. The results of this study highlight that a trade-off between competition and predator, i.e., weak competitors are better protected against predation, plays an important role in promoting plankton species richness and triggers the *Phaeocystis* bloom. As often observed in eutrophicated ecosystems, only the biotic homogenization of the copepod community and the shift in the diet of copepods toward *Phaeocystis* detrital materials have been detected during the *Phaeocystis* bloom. The diatom and copepod communities respond synchronously to fluctuating resources and biotic conditions by successively selecting species with specific traits. This study confirms the key role of competition and predation in controlling annual plankton succession.

**Keywords: diatoms, copepods, *Phaeocystis*, biodiversity, functional traits, seasonality, trade-off**

OPEN ACCESS

***Edited by:***
*Maria Moustaka-Gouni, Aristotle University of Thessaloniki, Greece*

***Reviewed by:***
*Jun Sun, China University of Geosciences, China*
*Fabio Benedetti, ETH Zürich, Switzerland*
*Benjamin Alric, Centre National de la Recherche Scientifique (CNRS), France*

****Correspondence:***
*Elsa Breton*
*elsa.breton@univ-littoral.fr*

***Specialty section:***
*This article was submitted to Aquatic Microbiology, a section of the journal Frontiers in Marine Science*

***Received:*** *20 January 2021*
***Accepted:*** *30 July 2021*
***Published:*** *06 September 2021*

***Citation:***
*Breton E, Christaki U, Sautour B, Demonio O, Skouroliakou D-I, Beaugrand G, Seuront L, Kléparski L, Poquet A, Nowaczyk A, Crouvoisier M, Ferreira S, Pecqueur D, Salmeron C, Brylinski J-M, Lheureux A and Goberville E (2021) Seasonal Variations in the Biodiversity, Ecological Strategy, and Specialization of Diatoms and Copepods in a Coastal System With Phaeocystis Blooms: The Key Role of Trait Trade-Offs. Front. Mar. Sci. 8:656300. doi: 10.3389/fmars.2021.656300*

www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

# Citizens and scientists collect comparable oceanographic data: measurements of ocean transparency from the Secchi Disk study and science programmes

Richard R. Kirby[1]✉, Gregory Beaugrand[2], Loick Kleparski[2,5], Susie Goodall[3] & Samantha Lavender[4]

Marine phytoplankton accounts for approximately 50% of all photosynthesis on Earth, underpins the marine food chain and plays a central role in the Earth's biogeochemical cycles and climate. In situ measurements of ocean transparency can be used to estimate phytoplankton biomass. The scale and challenging conditions of the ocean make it a difficult environment for in situ studies, however. Here, we show that citizen scientists (seafarers) using a simple white Secchi Disk can collect ocean transparency data to complement formal scientific efforts using similar equipment. Citizen scientist data can therefore help understand current climate-driven changes in phytoplankton biomass at a global scale.

The ocean is a difficult environment to access for in situ study due to its scale, remoteness and challenging conditions. Although ocean science research is vital for our sustainable future[1], scientific research lags current, climate-driven ocean changes[2]. The ocean's phytoplankton account for approximately 50% of all photosynthesis on Earth[3,4] and temporal and spatial changes in the phytoplankton can influence marine productivity[5], weather[6] and climate[7,8]. Monitoring the phytoplankton is therefore essential as an early indicator of regional and global ecosystem change[9–12].

Around 44% of the human population lives within 150 km of the coast[13] and a number go to sea for work and recreation. The seafaring public often visits the same locations, whether as sailors on short day trips, commercial fishermen accessing fishing grounds, or offshore yachtsmen/women whose passages follow common routes dictated by the season, prevailing winds and currents[14]. Therefore, the seafaring public provides an opportunity to collect oceanographic data over varying spatial and temporal scales to contribute to scientific efforts and many now participate in marine citizen science[15,16].

The global Secchi Disk study (http://www.secchidisk.org)[17] engages seafarers to use a Secchi Disk[18] to collect in situ data on ocean transparency that can be used to estimate phytoplankton biomass[19]. The Secchi Disk study uses a simple, 30 cm diameter white Secchi Disk that is weighted and attached to a tape-measure, and lowered vertically into the water from a boat's side. The depth (m) below the surface when the Secchi Disk disappears from sight is the Secchi depth ($Z_{SD}$), which measures ocean transparency. When the bathymetry is > 25 m depth and the distance > 1 km from shore, the primary influence upon ocean transparency is phytoplankton pigments and their breakdown products and therefore, $Z_{SD}$ estimates phytoplankton biomass in the water column; re-suspended sediments and dissolved organic matter from rivers further reduce transparency and introduce optical errors in shallower water and closer inshore[12].

Marine scientists have used Secchi Disks to measure ocean transparency since 1865[20] and archives of $Z_{SD}$ represent one of the longest-running, spatially extensive global marine datasets[21]. Recently, the Secchi Disk has fallen from widespread use among marine scientists[22] due to spectrophotometric determination of chlorophyll

[1]The Secchi Disk Foundation, Kiln Cottage, Gnaton, Yealmpton PL8 2HU, Devon, UK. [2]UMR 8187 - LOG - Laboratoire d'Océanologie et de Géosciences, CNRS, Univ. Lille, Univ. Littoral Côte d'Opale, 59000 Lille, France. [3]14 Osprey Close, Southampton SO40 4XJ, Hampshire, UK. [4]Pixalytics Ltd, Plymouth Science Park, 1 Davy Rd, Plymouth PL6 8BX, Devon, UK. [5]The Marine Biological Association, Citadel Hill, The Hoe, Plymouth PL1 2PB, Devon, UK. ✉email: richard.kirby@planktonpundit.org

communications biology

ARTICLE

https://doi.org/10.1038/s42003-022-04100-6 OPEN

# Addressing the dichotomy of fishing and climate in fishery management with the FishClim model

Grégory Beaugrand[1✉], Alexis Balembois[1], Loïck Kléparski[1,2] & Richard R. Kirby[3,4]

The relative influence of fishing and Climate-Induced Environmental Change (CIEC) on long-term fluctuations in exploited fish stocks has been controversial[1,2,3] because separating their contributions is difficult for two reasons. Firstly, there is in general, no estimation of CIEC for a pre-fishing period and secondly, the assessment of the effects of fishing on stocks has taken place at the same time as CIEC[4]. Here, we describe a new model we have called FishClim that we apply to North Sea cod from 1963 to 2019 to estimate how fishing and CIEC interact and how they both may affect stocks in the future (2020-2100) using CMIP6 scenarios[5]. The FishClim model shows that both fishing and CIEC are intertwined and can either act synergistically (e.g. the 2000-2007 collapse) or antagonistically (e.g. second phase of the gadoid outburst). Failure to monitor CIEC, so that fisheries management immediately adjusts fishing effort in response to environmentally-driven shifts in stock productivity, will therefore create a deleterious response lag that may cause the stock to collapse. We found that during 1963-2019, although the effect of fishing and CIEC drivers fluctuated annually, the pooled influence of fishing and CIEC on the North Sea cod stock was nearly equal at ~55 and ~45%, respectively. Consequently, the application of FishClim, which quantifies precisely the respective influence of fishing and climate, will help to develop better strategies for sustainable, long-term, fish stock management.



[1] Univ. Littoral Côte d'Opale, CNRS, Univ, Lille, UMR 8187 LOG, F-62930 Wimereux, France. [2] Marine Biological Association, Citadel Hill, Plymouth PL1 2PB, UK. [3] The Secchi Disk Foundation, Kiln Cottage, Gnaton, Yealmpton, Devon PL8 2HU, UK. [4] Ronin Institute, Montclair, NJ 07043, USA. ✉email: gregory.beaugrand@univ-lille.fr

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fmars.2022.914475



OPEN ACCESS

EDITED BY
Alberto Basset,
University of Salento, Italy

REVIEWED BY
Kyle Edwards,
University of Hawaii at Manoa,
United States
Márcio Silva de Souza,
Federal University of Rio Grande, Brazil

*CORRESPONDENCE
Elsa Breton
elsa.breton@univ-littoral.fr

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

RECEIVED 06 April 2022
ACCEPTED 15 September 2022
PUBLISHED 06 October 2022

CITATION
Breton E, Goberville E, Sautour B, Ouadi A, Skouroliakou D-I, Seuront L, Beaugrand G, Kléparski L, Crouvoisier M, Pecqueur D, Salmeron C, Cauvin A, Poquet A, Garcia N, Gohin F and Christaki U (2022) Multiple phytoplankton community responses to environmental change in a temperate coastal system: A trait-based approach. *Front. Mar. Sci.* 9:914475. doi: 10.3389/fmars.2022.914475



# Multiple phytoplankton community responses to environmental change in a temperate coastal system: A trait-based approach

Elsa Breton[1*], Eric Goberville[2], Benoit Sautour[3], Anis Ouadi[3], Dimitra-Ioli Skouroliakou[1], Laurent Seuront[1,4,5], Gregory Beaugrand[1], Loïck Kléparski[1,6], Muriel Crouvoisier[1], David Pecqueur[7], Christophe Salmeron[7], Arnaud Cauvin[1], Adrien Poquet[1,8], Nicole Garcia[9], Francis Gohin[10] and Urania Christaki[1]

[1]Univ. Littoral Côte d'Opale, CNRS, Univ. Lille, UMR 8187 LOG, Wimereux, France, [2]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, [3]Univ. Bordeaux, CNRS, UMR 5805 EPOC, Rue Geoffroy Saint Hilaire – Bâtiment B18N, Pessac, France, [4]Department of Marine Resources and Energy, Tokyo University of Marine Science and Technology, Tokyo, Japan, [5]Department of Zoology and Entomology, Rhodes University, Grahamstown, South Africa, [6]Continuous Plankton Recorder (CPR) Survey, Marine Biological Association, Plymouth, United Kingdom, [7]Observatoire Océanologique de Banyuls s/mer, FR 3724 - Laboratoire Arago - SU/CNRS, avenue Pierre Fabre Bât B, Banyuls sur Mer, France, [8]Univ. Côte d'Azur, CNRS, INSERM, IRCAN, Medical School of Nice, Nice, France, [9]Aix Marseille Univ., Université de Toulon, CNRS, IRD, MIO, Marseille, France, [10]Laboratoire d'écologie pélagique, IFREMER, DYNECO PELAGOS, CS 10070-29280 Plouzané, Brittany, France

The effect of environmental change in structuring the phytoplankton communities of the coastal waters of the Eastern English Channel was investigated by applying a trait-based approach on two decades (1996-2019) of monitoring on diatoms and *Phaeocystis*. We show that phytoplankton species richness in an unbalanced nutrient supply context was influenced by wind-driven processes, ecological specialization for dissolved inorganic phosphorous, temporal niche differentiation, and a competition-defense and/or a growth-defense trade-off, a coexistence mechanism where weak competitors (i.e., slower growing) are better protected against predation. Under the influence of both environmental perturbations (e.g., wind-driven processes, freshwater influence, unbalanced nutrient levels) and biotic interactions (e.g., competition, predation, facilitation), phytoplankton species exhibited specific survival strategies such as investment on growth, adaptation and tolerance of species to environmental stresses, silicification and resource specialization. These strategies have led to more speciose communities, higher productivity, functional redundancy and stability in the last decade. Our results revealed that the unbalanced nutrient reduction facilitated *Phaeocystis* blooms and that anthropogenic climate warming and nitrate reduction may threaten the diatom communities of the eastern English Channel in a near future. Our

# Toward a better understanding of Island Biogeography

**Gregory Beaugrand[1], Loick Kléparski[1,2], Christophe Luczak[1], Eric Goberville[3], Richard R Kirby[4,5]**

[1]*CNRS, Univ. Lille, Univ. Littoral Côte d'Opale, UMR 8187, LOG, Laboratoire d'Océanologie et de Géosciences, F 62930 Wimereux, France*

[2]*Marine Biological Association, The Continuous Plankton Recorder (CPR) survey, The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK.*

[3]Muséum National d'Histoire Naturelle, UMR 7208 BOREA, Sorbonne Université, CNRS, UCN, UA, IRD, Paris, France

[4]The Secchi Disk Foundation, Kiln Cottage, Gnaton, Yealmpton, Devon PL8 2HU, UK.

[5] Ronin Institute, Montclair, NJ 07043, USA

**Corresponding author:** Gregory.beaugrand@univ-lille.fr

**04/02/2022, 12.00**

# Abstract

The spatial arrangement of biodiversity and annual plankton succession are key phenomena influencing ocean biogeochemistry (e.g. the biological pump) and the life cycle of many species. Biodiversity and phenological shifts induced by climate change might alter species succession and lead to trophic desynchronization and community reorganisation in space and time. The aim of this PhD is to improve our understanding of plankton biodiversity and phenology by identifying factors and processes that control them and by modelling the annual plankton succession in the context of global climate change. To do so, we used an approach based on the MacroEcological Theory on the Arrangement of Life (METAL) and observations collected by the Continuous Plankton Recorder (CPR) survey.

To understand how annual plankton succession and species phenology will be altered in the context of global warming, it is important to identify what parameters and processes control these phenomena. Therefore, in the first part of this PhD, we describe the seasonal variations of major phytoplanktonic taxa in the North Sea and demonstrate that species' phenology results from the interaction between species' niche and the environment. We also show that diatoms with similar cell shape have also similar phenology and niches, i.e. that oblates (flattened cells) dominate the spring and autumn periods whereas prolates (elongated cells) dominate the summer period. We therefore establish a salient link between functional traits, the niche and the phenology.

In the second part, we examine the spatio-temporal organisation of plankton biodiversity in the North Atlantic Ocean and show that this region is characterised by large spatial coenoclines (i.e. gradient of biocoenosis or community) induced by the niche-environment interaction. We also develop a new method, called a "species chromatogram", that gives a graphical summary of the niche by representing together abundance gradients along various environmental dimensions. This method can be used to characterise and display the multidimensional ecological niches of different species and also to compare the niches of different species by means of an index that quantify the degree of niche overlapping. We use this index to demonstrate that the plankton belonging to the same trophic guilds have sufficiently distinct niches to coexist in an impermanent and heterogeneous environment in space and time.

Finally, we apply a model based on METAL and the projections of six Earth System Models involved in the Coupled Model Intercomparison Project Phase 6 (CMIP6) to project the past, present and future phenological changes of three phytoplanktonic groups, i.e. oblate and prolate diatoms and dinoflagellates, in three key areas of the North Atlantic Ocean. We show that the phenology of oblates is likely to continue to shrink and their abundance to decline whereas the phenology of prolates and dinoflagellates will expand and their abundance rise. Our results suggest a climate-induced reorganisation of the phytoplanktonic community in space and time that will affect the rhythm of generation of endosomatic energy, which may have strong consequences on ecosystem functioning and services.

# Résumé

L'arrangement spatial de la biodiversité et la succession annuelle du plancton sont des phénomènes clés influençant la biogéochimie des océans (par exemple la pompe biologique du carbone) et le cycle de vie de nombreuses espèces. La biodiversité et les changements phénologiques induits par le dérèglement climatique pourraient altérer la succession des espèces et conduire à une désynchronisation trophique ainsi qu'à une réorganisation des communautés dans l'espace et dans le temps. L'objectif de cette thèse est d'améliorer notre compréhension de la biodiversité et de la phénologie du plancton en identifiant les facteurs et processus qui les contrôlent, et en modélisant la succession annuelle du plancton dans le contexte du dérèglement climatique. Pour ce faire, nous avons utilisé une approche basée sur la théorie METAL (« MacroEcological Theory on the Arrangement of Life ») et les observations recueillies par le programme de suivi du plancton marin appelé « Continuous Plankton Recorder » (CPR).

Pour comprendre comment la succession planctonique annuelle et la phénologie des espèces seront affectées par le réchauffement climatique, il est important d'identifier quels paramètres et processus contrôlent ces phénomènes. Ainsi, dans la première partie de cette thèse, nous décrivons les variations saisonnières des principaux taxons phytoplanctoniques en mer du Nord et démontrons que la phénologie des espèces résulte de l'interaction entre la niche des espèces et les fluctuations de l'environnement. Nous montrons également que les diatomées avec une forme de cellule similaire ont également une phénologie et des niches similaires ; les oblates (cellules aplaties) dominent les périodes du printemps et de l'automne tandis que les prolates (cellules allongées) dominent la période estivale. Nous établissons donc un lien saillant entre les traits fonctionnels, la niche et la phénologie.

Dans une deuxième partie, nous examinons l'organisation spatio-temporelle de la biodiversité planctonique dans l'océan Atlantique Nord et montrons que cette région est caractérisée par de grandes cœnoclines spatiales (gradient de biocénose ou communauté) induites par l'interaction niche-environnement. Nous développons également une nouvelle méthode, appelée « chromatogramme spécifique », qui donne un résumé graphique de la niche en représentant ensemble des gradients d'abondance selon différentes dimensions environnementales. Cette méthode peut être utilisée pour caractériser et afficher les niches écologiques multidimensionnelles des espèces et comparer les niches au moyen d'un indice qui quantifie leur degré de chevauchement. Nous utilisons cette méthode pour démontrer que les espèces planctoniques appartenant aux mêmes guildes trophiques ont des niches suffisamment distinctes pour coexister dans un environnement impermanent et hétérogène dans l'espace et dans le temps.

Enfin, nous appliquons un modèle basé sur la théorie METAL et les projections de six modèles de système Terre impliqués dans le projet d'intercomparaison des modèles couplés phase 6 (CMIP6) pour projeter les changements phénologiques passés, contemporains et futurs des diatomées oblates et prolates ainsi que les dinoflagellés, dans trois zones clés de l'océan Atlantique Nord. Nous montrons que la durée de la saison d'abondance des oblates va poursuivre sa diminution et que leur abondance va probablement encore décliner. Au contraire, la phénologie des prolates et des dinoflagellés va s'étendre et leur abondance augmenter. Nos résultats suggèrent que le dérèglement climatique va modifier l'organisation des communautés phytoplanctoniques dans l'espace et le temps et affectera le rythme de génération d'énergie endosomatique, ce qui aura probablement de fortes conséquences sur le fonctionnement des écosystèmes et les services écosystémiques.