

Thèse de Master
Année 2018-2019
Master Sciences de la Mer

Typologie globale de la distribution tridimensionnelle des communautés de plancton vues par imagerie *in situ*

Thelma PANAÏOTIS

Réalisé sous la responsabilité de Jean-Olivier IRISSON et Lars STEMMANN

Laboratoire d'Océanographie de Villefranche (UMR 7093)

181 Chemin du Lazaret 06230 Villefranche-sur-Mer

A qui de droit,

L'administration du Master Sciences de la Mer de l'Université Sorbonne Université, est l'unique dépositaire des manuscrits servant à valider les Thèses de Master. Nous ne diffusons des copies que sur demande écrite, après avoir sollicité l'accord de l'auteur et du ou des responsables du stage qui a conduit à l'écriture de la Thèse. En apposant la mention ci-dessous dans la première page du manuscrit (page de couverture intérieure), nous nous engageons à ne pas diffuser le manuscrit et à transmettre toute demande à l'auteur et au responsable de stage qui la traitera indépendamment.

Alain NIERGA

Gestionnaire de la formation

SORBONNE UNIVERSITE
FACULTE SCIENCES et INGENIERIE
MASTER SCIENCES DE LA MER
Tour 46 - 46/00 - Etage 1 - BC 6
4 place JUSSIEU
75252 PARIS Cedex 05

Avertissement/Disclaimer

Cette Thèse de Master a été réalisée de manière indépendante, pour remplir une partie des conditions d'obtention du titre de Master Sciences de la Mer de l'Université Sorbonne Université. Aucune autre assistance ou source que celles mentionnées n'ont été utilisées. Aucune donnée ne pourra être prise de ce manuscrit sans l'approbation de l'auteur et du responsable de stage.

This Master Thesis has been independently composed for the partial fulfillment of the title of Master of Science in Marine Sciences at University «Sorbonne Université ». No other than the indicated aid and sources have been used. It has not been presented to any other examination board. No data can be taken out of this manuscript without prior approval of the author and the thesis-promoter.

Table des matières

1	Introduction	1
1.1	Biogéographies des océans	1
1.2	Distribution du plancton	3
2	Matériel et méthodes	4
2.1	Collecte des données	4
2.2	Prétraitement des données	6
2.2.1	Données biologiques	6
2.2.2	Données environnementales	8
2.3	Définition d'une partition verticale de l'océan	8
2.4	Étude de la distribution globale du plancton	9
2.5	Comparaison avec des régionalisations de l'océan mondial	10
3	Résultats	13
3.1	Effet des cycles circadien et saisonnier	13
3.2	Distribution spatiale des communautés planctoniques	14
3.2.1	Couche épipélagique	14
3.2.2	Couche mésopélagique supérieure	17
3.2.3	Couche mésopélagique inférieure	20
3.2.4	Couche bathypélagique	21
3.2.5	Structuration des communautés avec la profondeur	21
3.3	Comparaison avec d'autres motifs de partition de l'océan mondial	22
4	Discussion	23
4.1	Biais potentiels	23
4.1.1	Saisonnalité et rythme circadien	23
4.1.2	Efficacité de détection des organismes	24
4.1.3	Précision de la taxonomie	26
4.1.4	Distribution des profils	26
4.2	Structure générale des communautés de plancton	27
4.2.1	Couche épipélagique	27
4.2.2	Couches mésopélagique et bathypélagique	28
4.2.3	Homogénéisation en profondeur	29
4.3	Une distribution contrainte par les conditions régionales	30
5	Perspectives	30

Remerciements

J'adresse tout d'abord mes remerciements à Rodolphe LEMÉE, Directeur du LOV, pour son accueil dans le laboratoire, ainsi que Jean-Olivier IRISSON, Responsable de l'équipe COMPLEX, pour l'accueil au sein de son équipe et son co-encadrement pour ce stage de deuxième année de Master. Merci de m'avoir guidée tout au long de ces mois et pour la confiance accordée. Je suis très reconnaissante d'avoir la chance de continuer à travailler au sein de cette équipe pour les prochaines années dans la cadre d'une thèse dirigée par Sakina Dorothee AYATA et Jean-Olivier IRISSON.

J'adresse également mes remerciements à Lars STEMMANN, qui a co-encadré ce stage avec Jean-Olivier, pour son aide précieuse dans l'interprétation des résultats et la rédaction de la discussion.

Je remercie Tristan BIARD pour son aide dans l'identification des rhizaires, ainsi que ses conseils pour l'amélioration de ce travail.

Je remercie également Léo LACOUR pour m'avoir fourni les données satellites.

Je tiens aussi à remercier Marc PICHERAL pour le développement des outils qui nous permettent d'étudier et de comprendre un peu plus les océans chaque jour.

Merci à Laure, Salomé et Virginie pour la bonne ambiance qui a toujours régné dans notre bureau, ainsi que les gâteaux et les bons moments qui y ont été partagés.

Merci à l'ensemble des membres de l'équipe COMPLEX pour leurs conseils, les agréables moments autour d'un café et la très bonne ambiance générale qui règne au sein du laboratoire.

Merci à l'équipe de volley du CLAS de Villefranche-sur-Mer pour tous ces intenses moments sportifs entrecoupés de moments de détente.

Un très grand merci à l'équipe du restaurant de la station, toujours souriante, où ce fût un plaisir d'aller manger tous les midis.

Merci également à toutes les personnes qui m'ont entourée pendant la durée de ce stage et surtout, merci à toi, Julie, de m'avoir tant soutenue ces dernières semaines.

Liste des tables et figures

Liste des tables

1	Liste des groupes retenus pour l'analyse.	7
2	Définition des saisons productives et non productives	10
3	Variance expliquée par les régionalisations	23

Liste des figures

1	Carte de deux régionalisations testées	11
2	Composition planctonique du jeu de données	14
3	Carte de la concentration moyenne du plancton dans la couche 0-200 mètres.	15
4	ACP des stations de la couche épipélagique	15
5	Carte des groupes planctoniques de la couche épipélagique	16
6	Composition des groupes planctoniques de la couche épipélagique	17
7	ACP des stations de la couche mésopélagique supérieure	18
8	Carte des groupes planctoniques de la couche mésopélagique supérieure	18
9	Composition des groupes planctoniques de la couche mésopélagique supérieure	19
10	Valeurs propres centrées des ACP des différentes couches	22

Liste des symboles et abréviations

Symbole	Description	Unité
AOU	Utilisation apparente de l'oxygène	$\mu\text{mol.kg}^{-1}$
BBP	Coefficient de rétrodiffusion des particules	m^{-1}
[Chl a] _{surf}	Concentration en chlorophylle a de surface	mg.m^{-3}
DCM	Profondeur du maximum de chlorophylle	m
snow vol	Biovolume de neige marine	$\text{mm}^3.\text{m}^{-3}$
snow conc	Concentration en neige marine	m^{-3}
epi	Profondeur de la couche épipélagique	m
halo	Profondeur de la halocline	m
K_{dPAR}	Coefficient d'atténuation du PAR	m^{-1}
MLD	Profondeur de la couche de mélange	m
oxy	Concentration en oxygène dissous	mL.L^{-1}
PAR	Rayonnement photosynthétique actif	$\mu\text{mol.m}^{-2}.\text{s}^{-1}$
bulk vol	Biovolume de particules	$\text{mm}^3.\text{m}^{-3}$
bulk conc	Concentration en particules	m^{-3}
PIC	Carbone inorganique dissous	mol.m^{-3}
POC	Carbone organique dissous	mol.m^{-3}
pycno	Profondeur de la pycnocline	m
sal	Salinité	
sigma	Densité potentielle	
strat	Indice de stratification	
temp	Température	$^{\circ}\text{C}$
thermo	Profondeur de la thermocline	m
Z_{eu}	Profondeur de la couche euphotique	m

Abréviation	Description
ACP	Analyse en composantes principales
CAH	Classification ascendante hiérarchique
CTD	Conductivity temperature depth
RDA	Analyse de redondance
UVP	Underwater Vision Profiler

1 Introduction

1.1 Biogéographies des océans

Les océans sont classiquement divisés verticalement en couches épipélagique, mésopélagique et bathypélagique (ainsi qu'abyssopélagique et hadopélagique, plus profondes encore). La couche épipélagique correspond à la couche où l'intensité de la lumière est suffisante pour la production primaire et au maximum de biomasse planctonique (Behrenfeld and Boss, 2014); elle est habituellement considérée comme s'étendant de la surface à 200 m de profondeur. La couche mésopélagique s'étend de 200 à 1000 m et comprend la zone disphotique où la lumière peut être utilisée pour la prédation, mais n'est pas suffisante pour la photosynthèse (Robinson et al., 2010). Elle correspond aussi à la couche dans laquelle la matière produite en surface est reminéralisée au cours de sa sédimentation. Finalement, la couche bathypélagique située en deçà de 1000 mètres correspond à une zone de faible biomasse et forte concentration en nutriments (Reygondeau et al., 2018). Nos connaissances sur les écosystèmes diminuent avec la profondeur en raison des difficultés techniques d'explorer les zones profondes.

Parallèlement à cette zonation verticale, plusieurs partitions horizontales des océans ont été proposées. La plus connue est celle de Longhurst, qui délimite 56 provinces pouvant être regroupées en 4 biomes (Longhurst, 1995, 2010). Cette régionalisation se base essentiellement sur les données biogéochimiques (p. ex. température de surface, profondeur de la couche de mélange). Par la suite d'autres régionalisations ont été proposées, telles que les *Large Marine Ecosystems* (Sherman, 2005) basés sur quatre critères écologiques : la bathymétrie, l'hydrographie, la productivité et les communautés écologiques. Les *Marine Ecoregions of the World* proposées par Spalding et al. (2007) sont quant à elles basées sur la répartition des communautés. Elles ont été établies dans l'objectif d'évaluer les efforts mis en oeuvre dans la conservation des écosystèmes. Comme les *Large Marine Ecosystems*, les *Marine Ecoregions of the World* concernent uniquement les zones côtières et de plateau continental, où se concentrent les activités humaines ainsi que les actions liées à la conservation. D'autres travaux ont cherché à établir une régionalisation intégrant des données biotiques et abiotiques, en se basant à la fois des données d'aires de répartition d'espèces et de forçages physiques (Spalding et al., 2012).

Une autre vision synthétique des propriétés de l'océan couramment considérée est sa divi-

sion en bandes latitudinales. Là aussi, cette “régionalisation” est basée sur des critères environnementaux : la température de surface et l’ensoleillement sont deux variables environnementales importantes qui changent fortement avec la latitude.

Ces régionalisations ne concernent cependant que la couche épipélagique et ne sont pas forcément adaptées pour décrire les couches plus profondes, moins connues et plus difficilement échantillonnables (Spalding et al., 2012, Costello et al., 2018). D’autres régionalisations ont ainsi été proposées spécifiquement pour la couche mésopélagique. Reygondeau et al. (2018) propose une partition horizontale de la couche mésopélagique, tout en définissant des frontières verticales alternatives non pas fixes, comme précédemment, mais variables en fonction des conditions du milieu : lumière, densité et flux de carbone. Cette régionalisation est basée sur des critères environnementaux et délimite 13 provinces biogéochimiques mésopélagiques. Une autre partition de la couche mésopélagique est proposée par Sutton et al. (2017) en se basant sur des critères biotiques et environnementaux. La couche mésopélagique (200-1000 m) est divisée en 33 écorégions, dont 20 océaniques et 13 plutôt néritiques.

Ainsi, plusieurs partitions horizontales de l’océan global ont été proposées pour les couches épipélagique et mésopélagique. Cependant ces partitions sont essentiellement basées sur la biogéochimie des océans et souvent focalisées sur les zones côtières. En effet, peu de données sont disponibles sur la distribution des organismes océaniques à large échelle, en particulier loin de côtes. Pourtant, ces zones, par leur étendue, sont primordiales pour le fonctionnement des cycles biogéochimiques (Emerson et al., 1997) et sont maintenant considérées comme de nouvelles cibles de conservation (p. ex. initiative *Biodiversity Beyond National Jurisdiction* des Nations Unies¹) et de placement d’aires marines protégées (p. ex. efforts pour créer une aire marine protégée hauturière dans l’Océan Austral²). Par ailleurs la façon dont les régionalisations citées reflètent ou non la distribution des organismes n’a pas encore été étudiée d’une façon intensive. Quelques travaux ont porté sur des groupes généralistes tels que le zooplancton (Beaugrand et al., 2002, Woodd-Walker et al., 2002), les bactéries (Li et al., 2004) ou des assemblages plus larges (Tittensor et al., 2010, Chaudhary et al., 2016). Il est donc important de décrire la distribution de la biodiversité à ces échelles et en trois dimensions, ainsi que d’examiner si les régionalisations existantes reflètent correctement la distribution des organismes.

1. <https://www.un.org/bbnj/>

2. <https://www.diplomatie.gouv.fr/en/french-foreign-policy/sustainable-development-environment/events/article/environment-failure-of-projects-to-create-marine-protected-areas-in-the>

1.2 Distribution du plancton

Les organismes planctoniques sont définis comme des organismes qui ne peuvent lutter contre les courants. Ils sont présents dans l'ensemble des bassins océaniques. Le plancton regroupe des organismes très divers au niveau taxonomique (Sunagawa et al., 2015, de Vargas et al., 2015) et cette diversité se retrouve également au niveau de la taille : de quelques micromètres pour les bactéries à plusieurs mètres pour certains siphonophores. Les organismes planctoniques sont généralement séparés en phytoplancton, autotrophe, et zooplancton, hétérotrophe.

Le plancton joue un rôle essentiel dans les cycles biogéochimiques, notamment le cycle du carbone à travers la pompe biologique (Longhurst and Glen Harrison, 1989, Banse, 1995), en permettant le stockage d'une quantité de carbone équivalente aux émissions annuelles (Volk and Hoffert, 1985). Il est également à la base du réseau trophique océanique (Ware and Thomson, 2005, Beaugrand et al., 2010, Falkowski, 2012) et assure le transfert de matière organique depuis les producteurs primaires vers les grands prédateurs marins dont certaines espèces ont une importance patrimoniale ou économique (Frederiksen et al., 2006). De plus la moitié de la production de dioxygène est générée par le phytoplancton (Behrenfeld et al., 2001). Les organismes planctoniques sont très sensibles aux changements qui affectent les masses d'eau au sein desquelles ils sont insérés ; cela en fait donc de bonnes sentinelles. En effet, leur répartition est largement contrainte par les facteurs environnementaux, notamment la température, l'oxygène, les nutriments et la lumière (Hays et al., 2005). Ainsi la diversité planctonique est liée à la latitude avec plus de diversité à l'équateur qu'aux pôles (Rutherford et al., 1999, Rombouts et al., 2009, Tittensor et al., 2010, Miyamoto et al., 2014), tandis que la biomasse planctonique est plus élevée aux pôles qu'à l'équateur (Ikeda, 1985). L'ensemble de ces caractéristiques rendent l'étude de la distribution globale de la diversité planctonique pertinente pour des question de biogéochimie mais aussi de conservation et de gestion.

Peu de données quantitatives sont disponibles sur la répartition des organismes planctoniques en milieu hauturier à l'échelle globale du fait que les techniques ne sont pas adaptées pour la résolution des échelles de temps et d'espace nécessaires à la biogéographie. Historiquement, le plancton était échantillonné en utilisant des filets et des pompes. Or ces méthodes sont susceptibles d'abîmer ou de détruire les organismes fragiles et restent chronophages pour

l'identification des organismes. De nouveaux outils d'imagerie *in situ* permettent maintenant de générer de grandes quantités de données sur les concentrations et biovolumes de neige marine et plancton avec une fine résolution temporelle et spatiale (Dennett et al., 2002, Stemmann et al., 2008a,b, Guidi et al., 2015, Biard et al., 2016). Ces outils sont essentiels pour l'étude de groupes taxonomiques fragiles, par exemple les rhizaires, dont la contribution à la biomasse planctonique a pendant longtemps été sous-estimée (Biard et al., 2016). Ces méthodes d'imagerie permettent également d'obtenir des informations individuelles sur certains traits fonctionnels (p. ex. taille, transparence), desquels on peut déduire d'autres variables d'intérêt écologique et biogéochimique comme le biovolume et la biomasse.

Parmi ces outils, l'Underwater Vision Profiler (UVP) permet de quantifier la distribution verticale des particules et des organismes planctoniques > 600 μm (Picheral et al., 2010). Son efficacité a été démontrée pour l'estimation du flux particulaire (Guidi et al., 2015) ainsi que pour l'étude de la distribution du zooplancton de grande taille (Stemmann et al., 2008a, Forest et al., 2012, Guidi et al., 2015).

Ainsi cet outil permet de mener des études à l'interface entre la biodiversité et la biogéochimie, à un niveau diversité taxinomique intermédiaire entre ce qui est permis par les observations fines de prélèvements au filet et les modèles biogéochimiques intégrant généralement moins de 10 groupes planctoniques fonctionnels. Le premier objectif de ce travail est de décrire la composition des communautés planctoniques vues par l'UVP, d'examiner sa corrélation avec l'environnement biogéochimique immédiat et de diviser ces écosystèmes en grands types. Ensuite, nous avons évalué la capacité de diverses régionalisations existantes, basées sur la biogéochimie, à décrire ces variations de communautés afin de juger de leur pertinence dans le domaine écologique également.

2 Matériel et méthodes

2.1 Collecte des données

Les données utilisées proviennent de 25 campagnes océanographiques menées entre 2008 et 2018, couvrant l'ensemble des océans (Figure S1 en annexe) et lors desquelles ont été réali-

sés des profils verticaux utilisant l'UVP5. Cet instrument d'imagerie sous-marine *in situ* photographie les éléments situés dans une tranche d'eau éclairée par deux faisceaux de LED, correspondant à un volume de 1,02 L par image (Picheral et al., 2010). Les images de tous les objets d'un diamètre supérieur à 600 μm sont enregistrées pour être analysées ultérieurement. En parallèle, tous les objets d'un diamètre supérieur à 60 μm (nommés "particules" ensuite) sont mesurés individuellement pour la taille.

L'UVP5 est déployé sur une rosette équipée d'une sonde CTD effectuant un profil vertical et qui récolte donc simultanément des données de température ($^{\circ}\text{C}$) et conductivité ($\text{mS}\cdot\text{cm}^{-1}$) ainsi que, fréquemment, de concentration en chlorophylle a ($\text{mg}\cdot\text{m}^{-3}$) mesurée par fluorescence et d'oxygène ($\text{mL}\cdot\text{L}^{-3}$ ou $\mu\text{mol}\cdot\text{kg}^{-1}$). Des produits des satellites dits "couleurs de l'eau" ont été extraits pour compléter les profils de ces quelques variables. Ces données sont moyennées sur 8 jours et sur une surface de $0,5^{\circ}$ par $0,5^{\circ}$ centrée sur le lieu d'échantillonnage de l'UVP. Les variables retenues sont :

- concentration en chlorophylle a de surface ($[\text{Chl a}]_{\text{surf}}$, $\text{mg}\cdot\text{m}^{-3}$), un indicateur de la biomasse phytoplanctonique
- coefficient de rétrodiffusion des particules (BBP, m^{-1}) à la longueur d'onde 443 nm, indicateur de la concentration des particules en suspension dans l'eau
- rayonnement photosynthétique actif (PAR, $\mu\text{mol}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$), correspondant au flux quotidien moyen de photons dans le domaine du visible (400-700 nm)
- coefficient d'atténuation du PAR (K_{dPAR} , $\text{m}\cdot^{-1}$) (Morel et al., 2007)
- concentration en carbone organique particulaire (POC, $\text{mol}\cdot\text{m}^{-3}$)
- concentration en carbone inorganique particulaire (PIC, $\text{mol}\cdot\text{m}^{-3}$), qui correspond à la concentration en carbonate de calcium.

Cela constitue ainsi un jeu de données biotiques et abiotiques, où les données de concentration des organismes planctoniques sont associées à un ensemble de données environnementales.

2.2 Prétraitement des données

2.2.1 Données biologiques

Les images de chaque objet, générées par l'UVP5, sont importées sur l'application web Eco-Taxa où elles sont classées à l'aide d'un algorithme d'apprentissage supervisé (Gorsky et al., 2010). Les objets sont classés en fonction de leur appartenance à un groupe : groupe taxonomique, neige marine, artefact ou éventuellement objet non identifié. Cette classification doit néanmoins être validée par un opérateur. J'ai participé à ce travail en validant 256 000 images issues de trois campagnes. Un guide de référence pour la classification des images de l'UVP5 a été mis en place afin que la taxonomie soit homogène entre des projets validés par des opérateurs différents. Parmi les campagnes sélectionnées, seuls les profils validés à 100% sont conservés pour l'analyse ; cela représente 3167 profils comprenant 7 782 988 objets dont 700 358 sont des organismes vivants classés taxonomiquement. Du fait de différences de précision dans la classification taxonomique entre différentes campagnes, il est nécessaire de regrouper certains taxa, pour obtenir un plus petit dénominateur commun homogène entre toutes les campagnes. Par la suite, d'autres regroupements entre taxa sont effectués au vu de leur comportement similaire dans les analyses préliminaires. La liste définitive contient 28 taxa (Table 1, Figure S2). Certains objets dont la classification était visiblement erronée mais qu'il était difficile de re-classer correctement sur l'ensemble du jeu de données, notamment des trichodesmiums situés à plus de 500 m de profondeur, sont supprimés. Pour les trichodesmiums, cela correspond à 711 organismes sur un total de 193 714, soit une part négligeable.

La classification et la mesure de chaque objet permettent de calculer des concentrations (L^{-1}) et biovolumes ($mm^3 \cdot m^{-3}$) de plancton par tranche de 5 m pour chaque taxon présent dans cette tranche en distinguant les tranches pour lesquelles il n'y a pas de données des tranches dans lesquelles aucun organisme planctonique d'un diamètre supérieur à $600 \mu m$ n'a été détecté. Ces mêmes variables sont calculées pour la neige marine (objets supérieurs à $600 \mu m$ de diamètre identifiés comme des agrégats) et les "particules" (ensemble des objets de diamètre supérieur à $60 \mu m$, comprenant neige marine et plancton). Ces deux derniers sont traités comme des variables environnementales, quantifiant respectivement la richesse en neige marine et la richesse globale en objets, vivants ou non.

Table 1 – Liste des groupes retenus pour l’analyse. Les numéros d’images renvoient aux imagerie de la Figure S2 en annexe.

Groupe	Organismes classés comme	Image
Acantharea	Acantharea	1
Actinopterygii	Actinopterygii	2
Annelida	Annelida	3
Appendicularia	Appendicularia ou Appendicularia_house	4
Cephalopoda	Cephalopoda	5
Chaetognatha	Chaetognatha	6
Cnidaria_others	Cnidaria à l’exception des Hydrozoa	7
Collodaria	Collodaria	8
Copepoda	Copepoda ou Copepoda_like	9
Crustacea_others	Crustacea à l’exception des Copepoda, Ostracoda et Eumalacostraca	10
Ctenophora	Ctenophora	11
Doliolida	Doliolida	12
Eumalacostraca	Eumalacostraca	13
Foraminifera	Foraminifera	14
Gymnosomata	Gymnosomata	15
Hydrozoa_others	Hydrozoa à l’exception des Narcomedusae et Siphonophorae	16
Limacinidae	Limacinidae	17
Mollusca_others	Mollusca à l’exception des Gymnosomata, Limacinidae, Cavolina et Creseis	18
Narcomedusae	Narcomedusae	19
Nostocales	Nostocales	20
Ostracoda	Ostracoda	21
Phaeodaria	Phaeodaria	22
Pyrosoma	Pyrosoma	23
Rhizaria_others	Rhizaria à l’exception des Acantharea, Collodaria, Foraminifera et Phaeodaria	24
Salpida	Salpida	25
Siphonophorae	Siphonophorae	26
Thecosomata_cavo_or_creseis	Cavolina ou Creseis	27
Trichodesmium	puffs (Trichodesmium en étoile) ou tufts (Trichodesmium en bâtonnet)	28

2.2.2 Données environnementales

Les données aberrantes codées par des valeurs négatives sont éliminées puis les données brutes de température, salinité et concentration en oxygène dissous sont utilisées pour calculer la densité potentielle (σ) et l'utilisation apparente de l'oxygène (AOU, $\mu\text{mol.kg}^{-1}$). Les données sont nettoyées en éliminant les valeurs extrêmes en utilisant un écart médian à la médiane sur une fenêtre mobile (Leys et al., 2013). Les profils avec plus de 20% de données manquantes sont éliminés puis les données sont interpolées sur un pas de 1 m.

Ces données de résolution verticale au mètre sont utilisées pour dériver un ensemble de variables synthétiques de la forme de chaque profil. Les profondeurs de la thermocline, de la halocline et de la pycnocline sont calculées comme la profondeur du maximum de l'écart-type de la variable pertinente (température, salinité, densité) calculé sur une fenêtre mobile de taille 5 m. La profondeur de la couche de mélange (MLD) est déterminée par la profondeur à laquelle la densité présente un écart de $0,03 \text{ kg.m}^{-3}$ par rapport à la densité de référence mesurée sur la couche 0-5 m en surface (de Boyer Montégut et al., 2004). Sont également calculées la profondeur du maximum profond de chlorophylle (DCM), la profondeur de la zone euphotique (Z_{eu}) à partir du profil intégré de chlorophylle a (Morel and Maritorena, 2001) et l'indice de stratification, qui correspond au quotient de la densité mesurée entre les cinq premiers mètres de la colonne d'eau et une tranche de profondeur de référence, fixée ici entre 245 et 250 m de profondeur qui est la plupart du temps en dessous de la pycnocline.

Les données précises au mètre sont ensuite moyennées par tranche de 5 m pour correspondre aux tranches des données de concentration de plancton.

Dans le cas où une donnée est manquante (jusqu'à 25% de données manquantes pour les produits satellites), elle est remplacée par la moyenne de la variable concernée.

2.3 Définition d'une partition verticale de l'océan

Les données de concentration de plancton et les données environnementales sont moyennées selon une partition verticale de l'océan en 4 couches : épipélagique, mésopélagique supérieure, mésopélagique inférieure et bathypélagique. La frontière entre les couches épi et

mésopélagique est généralement fixée à 200 m. Ici, une définition dynamique représentant mieux la définition fonctionnelle et les variations locales d'épaisseur de cette couche est utilisée (Reygondeau et al., 2018) : profondeur maximale entre la pycnocline et le bas de la couche euphotique. La définition classique de la frontière méso-bathypélagique à 1000 m de profondeur est conservée. Cependant, de nombreux profils UVP s'arrêtent à 500 m et ne sont pas représentatifs de la couche mésopélagique dans son ensemble. La couche mésopélagique a donc été divisée en couches mésopélagique supérieure et mésopélagique inférieure dont la limite se situe à 500 m. Pour chaque couche, seuls les profils couvrant au moins 80% de l'épaisseur de la couche ont été conservés dans un souci de représentativité. L'ensemble des données biotiques et environnementales sont moyennées par couche pour chaque profil.

2.4 Étude de la distribution globale du plancton

Dans un premier temps, la structure des communautés est décrite en lien avec l'environnement ; puis afin de synthétiser géographiquement les grands types de communautés de plancton, les stations avec des communautés similaires sont regroupées.

Tout d'abord les concentrations de plancton moyennées sont transformées par la méthode d'Hellinger pour diminuer l'effet des fortes abondances et se focaliser sur la *composition* des communautés plutôt que sur les groupes dominants en concentration (Borcard et al., 2018). Une analyse en composantes principales (ACP) est effectuée sur ces données afin d'en synthétiser l'information, sur laquelle sont projetées les variables environnementales qui ont été centrées et réduites après avoir été log n+1 transformées pour les données de particules et de neige marine afin de les rapprocher de la normalité. Une classification ascendante hiérarchique (CAH) avec lien de Ward (Legendre and Legendre, 2012) est ensuite effectuée sur les coordonnées des stations sur les 5 premiers axes de l'ACP, capturant ainsi une grande part de variance et réduisant le bruit dans les données. L'arbre généré est coupé pour séparer les branches principales, déterminant ainsi le nombre de groupes.

Les biais potentiellement induits par le regroupement de profils effectués de jour et de nuit et à différentes saisons sont testés par le calcul de la part de variance de la composition des communautés planctoniques expliquée par ces variables. Pour cela, une analyse de redondance (RDA) est effectuée en utilisant les données de concentrations de plancton transformées par

Table 2 – Définition de saisons productives et non productives en fonction des zones latitudinales.

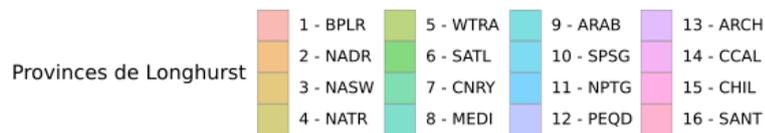
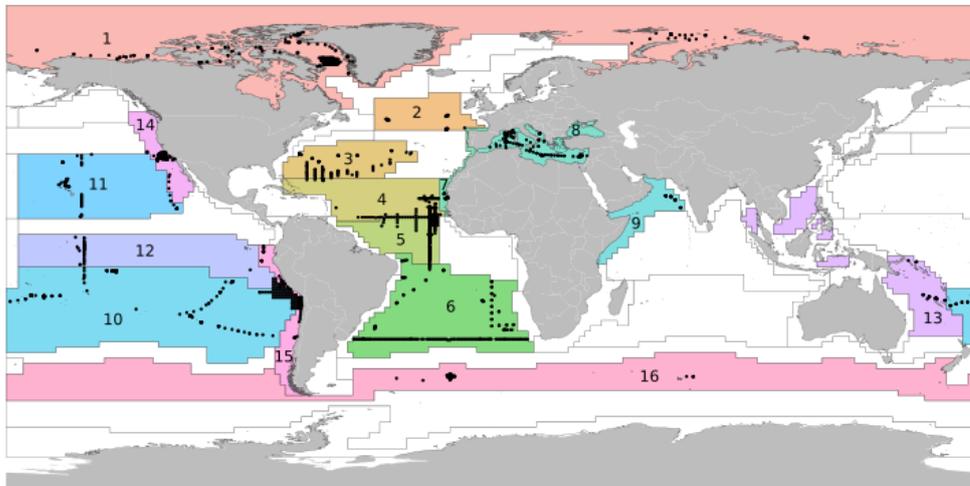
Zone latitudinale	Intervalle de latitude	Mois											
		J	F	M	A	M	J	J	A	S	O	N	D
Polaire nord	90°N - 66,5°N	0	0	0	0	0	1	1	1	0	0	0	0
Tempéré nord	66,5°N - 23,5°N	0	0	1	1	1	0	0	0	1	1	0	0
Tropical	23,5°S - 23,5 °S	0	0	0	0	0	0	0	0	0	0	0	0
Tempéré sud	23,5°S - 66,5°S	1	1	0	0	0	1	1	1	0	0	1	1
Polaire sud	66,5°S - 90°S	1	1	1	1	1	0	0	0	1	1	1	1

Hellinger en variable réponse et la variable binaire jour/nuit en variable explicative, cela dans la couche épipélagique et dans la couche mésopélagique. L'effet des saisons est vérifié de la même manière, dans la couche épipélagique uniquement, en construisant une variable binaire séparant la saison productive de la saison non-productive ; ces deux saisons sont définies en fonction du mois d'échantillonnage et de la latitude (Table 2) (Lalli and Parsons, 1997).

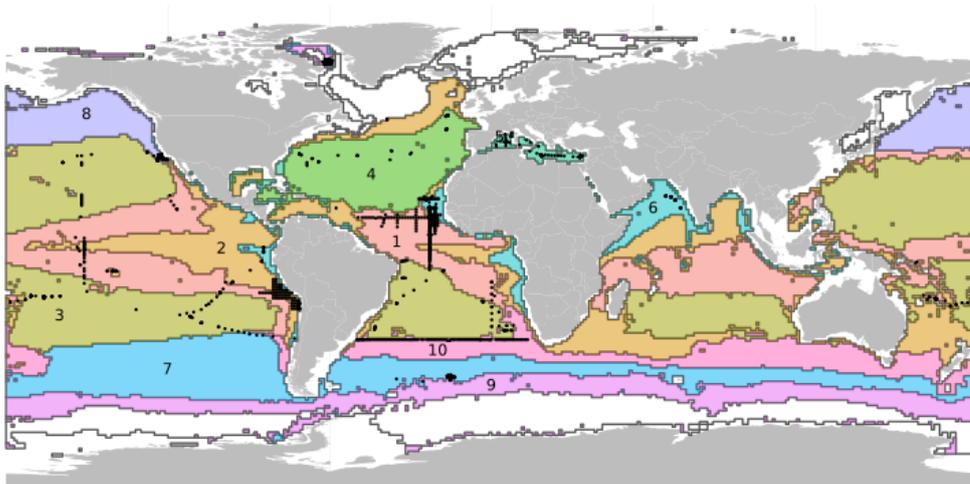
2.5 Comparaison avec des régionalisations de l'océan mondial

La part de variance de la composition de la communauté planctonique expliquée par diverses régionalisations de l'océan mondial est calculée pour estimer à quel point ces régionalisation sont pertinentes pour représenter les processus régissant l'écologie planctonique. Ceci est fait en suivant le même principe que le test de l'effet jour/nuit et saisonnier ci-dessus : une analyse de redondance (RDA) avec les concentrations de plancton transformées par Hellinger en variable réponse et la variable catégorielle contenant les noms de régions de la régionalisation choisie en variable explicative. Le test est fait dans chaque couche séparément.

Pour chaque régionalisation, seules les régions contenant au minimum 50 stations pour la couche épipélagique ou 25 stations pour les autres couches sont conservées, dans un souci de représentativité de chaque région. Les données des autres stations ne sont pas incluses dans cette analyse. Cela limite le nombre de régions utilisées pour chaque régionalisation, ce qui a pour effet d'aboutir à un nombre de régions conservé similaire entre les différentes régionalisation. Ceci est nécessaire car, dans une RDA, la part de variance expliquée par une variable catégorielle est souvent d'autant plus élevée que celle-ci contient un nombre élevé de modalités. Après sélection des régions, chaque régionalisation en contient entre 10 et 15, permettant de comparer la variance expliquée par chaque régionalisation (Table 3).



(a)



(b)

Figure 1 – Carte des Provinces de Longhurst (a) et des provinces biogéochimiques de la zone mésopélagique (b) (Reygondeau et al., 2018). Sont légendées les provinces utilisées dans l'analyse. La carte de l'ensemble des provinces ainsi que leur nom détaillé sont disponibles dans les figures S3 et S4 en annexe. Les stations utilisées sont représentées par des points noirs.

La première régionalisation testée est celle des provinces de Longhurst (Longhurst, 2010) (Figure 1). C'est une régionalisation basée sur la biogéochimie, donc susceptible de contraindre la répartition des communautés de plancton qui en dépendent.

Nous testons ensuite une régionalisation selon des bandes de latitudes de 10° chacune, dont les forçages dominants sont la lumière et la température variant en fonction de la latitude. L'effet de la latitude a déjà été démontré sur la diversité et la biomasse du plancton et pourrait donc également en contraindre la composition des communautés.

La régionalisation proposée par Reygondeau et al. (2018) (Figure 1) pour la couche mésopélagique est également testée pour les couches mésopélagiques supérieure et inférieure. Cette régionalisation est basée sur des données biogéochimiques issues de climatologies annuelles, donc susceptible également d'influencer la répartition des communautés de plancton. Elle est potentiellement plus pertinente pour ces couches profondes que la régionalisation de Longhurst qui considère principalement les processus épipélagique.

Nous testons également un regroupement basé sur les données environnementales locales, contenant les données CTD, les données satellites ainsi que les données de concentration et biovolume de neige marine et particules. Cette "régionalisation" sera représentative de processus à plus petite échelle spatio-temporelle que celles mentionnées ci-dessus, qui reposent sur des climatologies. Pour synthétiser géographiquement les données environnementales de la même manière que les communautés de plancton, ces données sont analysées avec une ACP suivie d'une CAH avec lien de Ward. Cette CAH est appliquée aux coordonnées des stations sur les 5 premiers axes de l'ACP, encore une fois pour réduire le bruit dans les données tout en conservant une part de variance importante. Le nombre de groupes générés par la CAH est choisi de façon à être similaire au nombre de régions dans les autres régionalisations.

Enfin, l'ensemble de ces régionalisations est comparé à un modèle explicatif maximal. Il est construit en calculant un regroupement sur les données de composition planctonique elles-mêmes et en choisissant un nombre de groupes comparable aux régionalisations testées. De la même manière que précédemment, ceci est effectué avec une CAH utilisant le lien de Ward basée sur les coordonnées des stations sur les 5 premiers axes d'une ACP ; cette ACP est également réalisée sur les communautés de plancton (concentrations transformées par Hellinger). La variance expliquée par la RDA utilisant ce regroupement est donc la part maximale de la variance

des communautés planctoniques explicable par une variable catégorielle avec ~ 12 modalités, étant donnée que la variable explicative est construite à partir des variables prédites.

Toutes les analyses ont été effectuées en utilisant le package *vegan* (Oksanen et al., 2007) sur le logiciel R 3.6.0.

3 Résultats

3.1 Effet des cycles circadien et saisonnier

Parmi les profils conservés pour décrire la couche épipélagique, 1633 ont été réalisés de jour, contre 1162 la nuit. La RDA réalisée avec la variable catégorielle jour/nuit est significative ($p < 0,001$), probablement en raison du grand nombre d'observations utilisé ($n = 2795$), mais elle n'explique qu'une part négligeable de la variance dans la composition de la communauté ($R^2_{\text{adj}} = 1,1\%$). Dans la couche mésopélagique supérieure, principalement affectée par les migrations nyctémérales du zooplancton, où 1057 profils ont été réalisés de jour et 806 de nuit, la part de variance expliquée par le cycle circadien est encore plus faible ($R^2_{\text{adj}} = 0,4\%$, $p < 0,001$). Ainsi il n'apparaît pas nécessaire de séparer les profils réalisés de jour ou de nuit, pour étudier la composition des communautés. Cela ne signifie pas que les migrations nyctémérales n'ont pas d'influence sur les concentrations de plancton estimées par l'UVP dans la zone épipélagique mais que leur effet sur la *composition* de la communauté est négligeable.

Concernant l'effet de la saison, dans la couche épipélagique 761 profils ont été réalisés en période qualifiée de productive contre 2034 en période non productive. La RDA réalisée avec la variable catégorielle sur la productivité de la saison est également significative ($p < 0,001$), mais elle n'explique que 3,7% de la variance dans la couche épipélagique. Ainsi pour l'étude de la composition des communautés, l'effet saisonnier semble également négligeable par rapport à l'effet spatial examiné plus bas.

3.2 Distribution spatiale des communautés planctoniques

Les organismes dominant les communautés planctoniques *in situ*, dans la gamme de taille vue par l'UVP (600 μm à quelques cm), sont des trichodesmiums, des copépodes et des rhizaires (phaeodaires, collodaires, acanthaires et autres rhizaires non identifiés) (Figure 2).

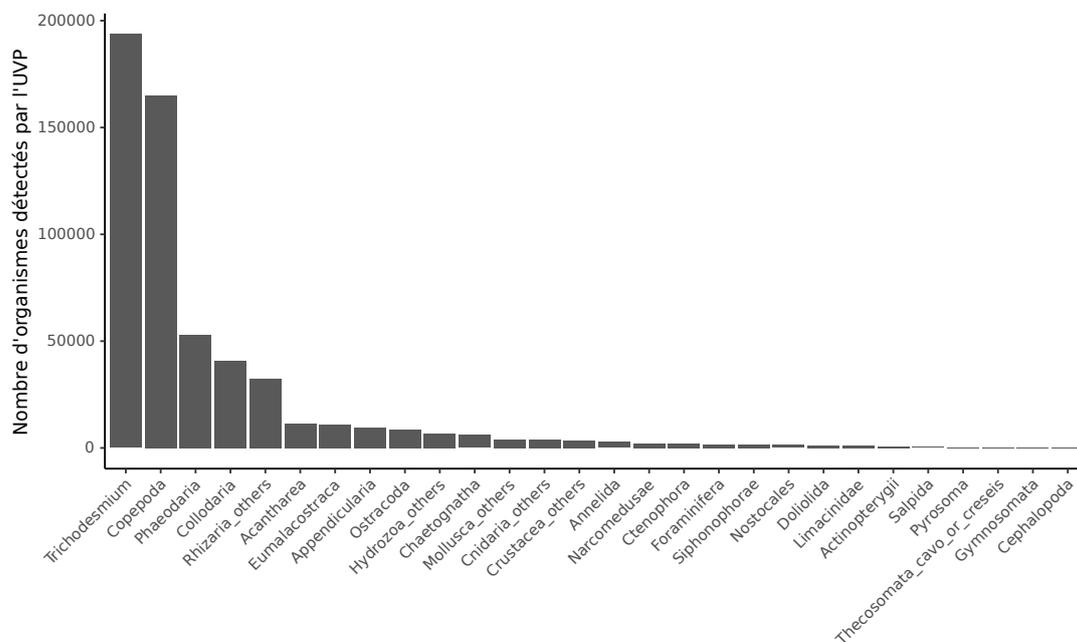


Figure 2 – Nombre d'organismes de chaque taxon détectés par l'UVP pour l'ensemble du jeu de données.

Les concentrations moyennées de l'ensemble du plancton sur les 200 premiers mètres de la colonne d'eau sont plus élevées dans la zone intertropicale de l'océan Atlantique, dans la baie de Baffin ainsi qu'au sud-ouest de l'océan Pacifique et au niveau du courant de Californie (Figure 3). Ce motif dépend essentiellement de la répartition des groupes dominants dans les données issues de l'UVP (Figure 2), c'est-à-dire les trichodesmiums (plutôt intertropical) et copépodes (hautes latitudes) et rhizaires (intertropical et surtout dans les zones d'upwelling).

3.2.1 Couche épipélagique

2795 profils ont été retenus pour la description des communautés de la couche épipélagique. L'ACP capture 49% de variance sur les deux premiers axes (Figure 4) et 68% sur les 4 premiers. L'axe 1 sépare des communautés d'eaux chaudes, riches en trichodesmiums et rhizaires, de communautés d'eaux froides, riches en copépodes. Cette répartition est bien visible sur la carte,

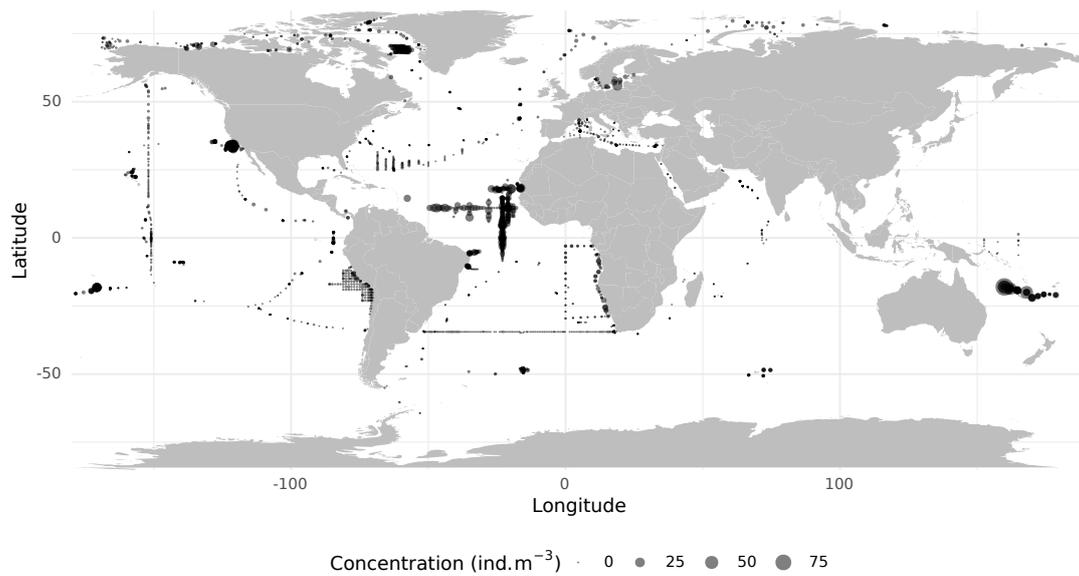


Figure 3 – Carte de la concentration moyenne du plancton dans la couche 0-200 mètres.

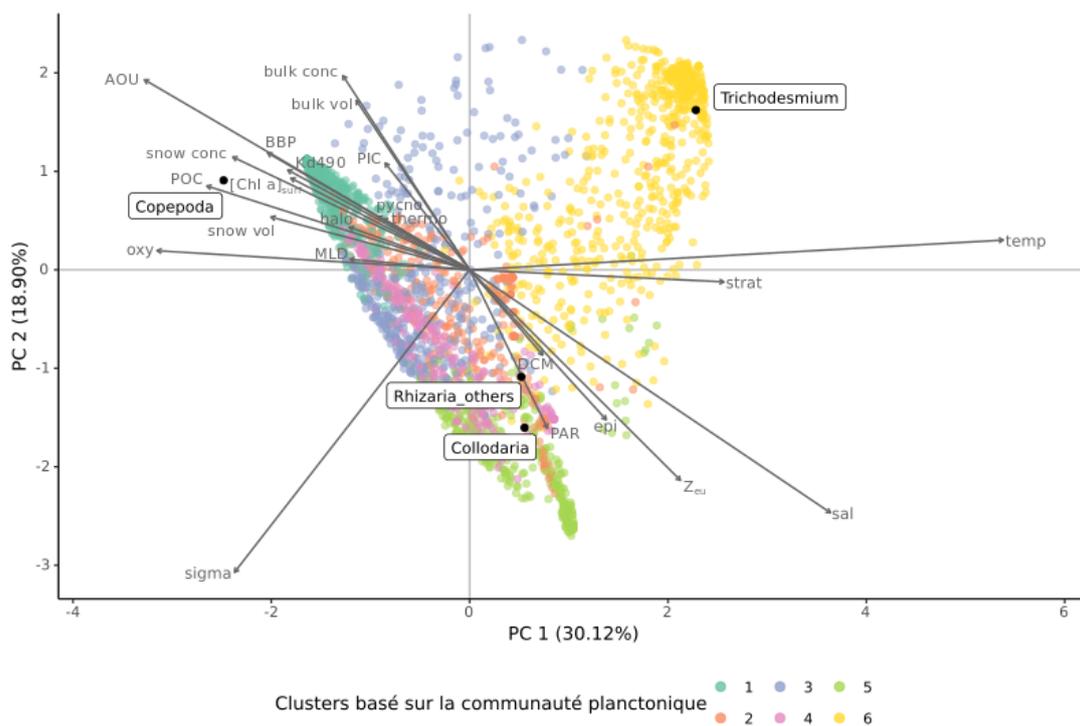


Figure 4 – Projection des stations de la couche épipélagique sur les deux premiers axes de l'ACP colorées en fonction des groupes définis par la CAH ; des catégories taxonomiques et des variables environnementales. Toutes les variables environnementales sont projetées, tandis que seules les catégories taxonomiques dont la norme de la projection sur les deux premiers axes est supérieure au rayon du cercle des contributions équivalentes le sont.

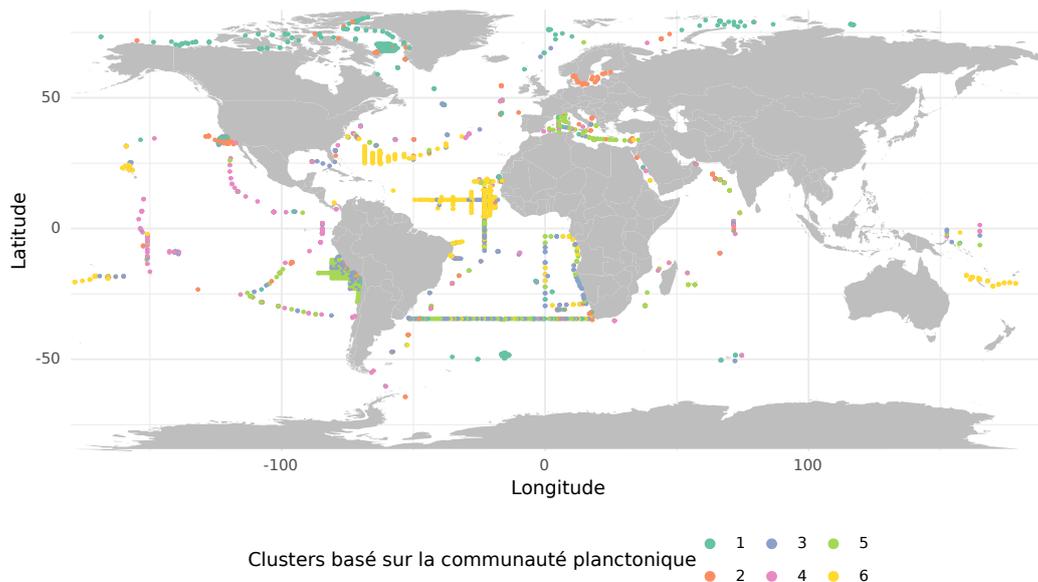


Figure 5 – Carte des stations de la couche épipélagique colorées en fonction des groupes définis par la CAH.

où le groupe 1, riche en copépodes, est essentiellement présent dans les eaux froides polaires et au sud des océans Atlantique et Indien, tandis que les groupes riches en rhizaires et trichodesmiums (groupes 5 et 6) se retrouvent principalement dans la bande intertropicale (Figure 5). La richesse en copépodes est associée à une richesse en neige marine et particules (Figure 4), correspondant à un milieu globalement eutrophe et productif. L'axe 2 sépare les communautés d'eaux chaudes entre celles riches en trichodesmiums (principalement dans l'Atlantique intertropical) de celles riches en collodaires et autres rhizaires (mer Méditerranée, zone d'upwelling au large du Pérou). L'axe 3 permet de distinguer les collodaires des autres rhizaires, ces derniers étant associés à des eaux peu stratifiées tandis que les collodaires se retrouvent plutôt dans les eaux stratifiées). Ainsi on distingue 4 pôles principaux dans la zone épipélagique : les zones riches en trichodesmiums, les zones riches en copépodes, les zones riches en collodaires et les zones riches en autres rhizaires.

Le dendrogramme généré par la CAH a été coupé en 6 groupes (Figure 6). Le groupe 1 contient essentiellement des copépodes. Les stations appartenant au groupe 1 sont situées dans des eaux froides et particulièrement bien oxygénées : océan Arctique et sud des Océans Atlantique et Indien (Îles Kerguelen). Le groupe 2 est composé majoritairement de copépodes et de phaeodaires, qui sont tous les deux des groupes associés à des eaux plutôt froides également. On retrouve les stations du groupe 2 dans la mer Baltique, au large du courant de Cali-

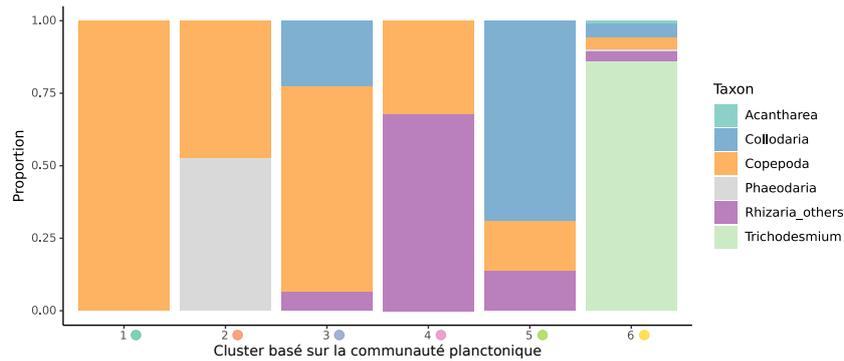


Figure 6 – Abondance relative des groupes taxonomiques dans les groupes définis par la CAH sur la couche épipélagique. Seuls les taxa présents dans au moins 50% des profils sont représentés.

fornie, mais aussi dans l’océan Arctique, c’est-à-dire des eaux froides et productives. Le groupe 3 contient majoritairement des copépodes ainsi que des collodaires. Ces stations sont situées principalement dans le sud de l’océan Atlantique, ainsi qu’au niveau du courant du Benguela sur la côte ouest-africaine ainsi qu’au niveau de la zone d’upwelling du Pérou, dont les eaux sont plutôt froides et peu oxygénées. Le groupe 4 contient une majorité de rhizaires appartenant à d’autres groupes que ceux identifiés précédemment, ainsi que des copépodes. Les stations de ce groupe sont situées dans l’océan Pacifique est, au niveau de la bande intertropicale, dans des eaux plutôt denses. Le groupe 5 est composé principalement de collodaires. On retrouve ces stations au niveau de la mer Méditerranée, au sud de l’océan Atlantique ainsi qu’au large de l’upwelling du Pérou. Le groupe 6 est marqué par la dominance des trichodesmiums. Ces stations se trouvent dans l’océan Atlantique et dans l’ouest de l’océan Pacifique au niveau de la bande intertropicale, correspondant à des eaux chaudes et peu denses.

Ainsi, les groupes générés permettent de définir de grands types de communautés planctoniques épipélagiques. Celles-ci sont structurées autour de pôles définis par les trichodesmiums, les copépodes et les rhizaires. La répartition spatiale des groupes montre clairement que la distribution des communautés est contrôlée par des facteurs environnementaux, notamment la température, mais aussi la disponibilité relative en oxygène, la salinité et la densité.

3.2.2 Couche mésopélagique supérieure

1863 profils ont été retenus pour la description des communautés de la couche mésopélagique supérieure. L’ACP capture 37,8% de variance sur les deux premiers axes (Figure 7) et

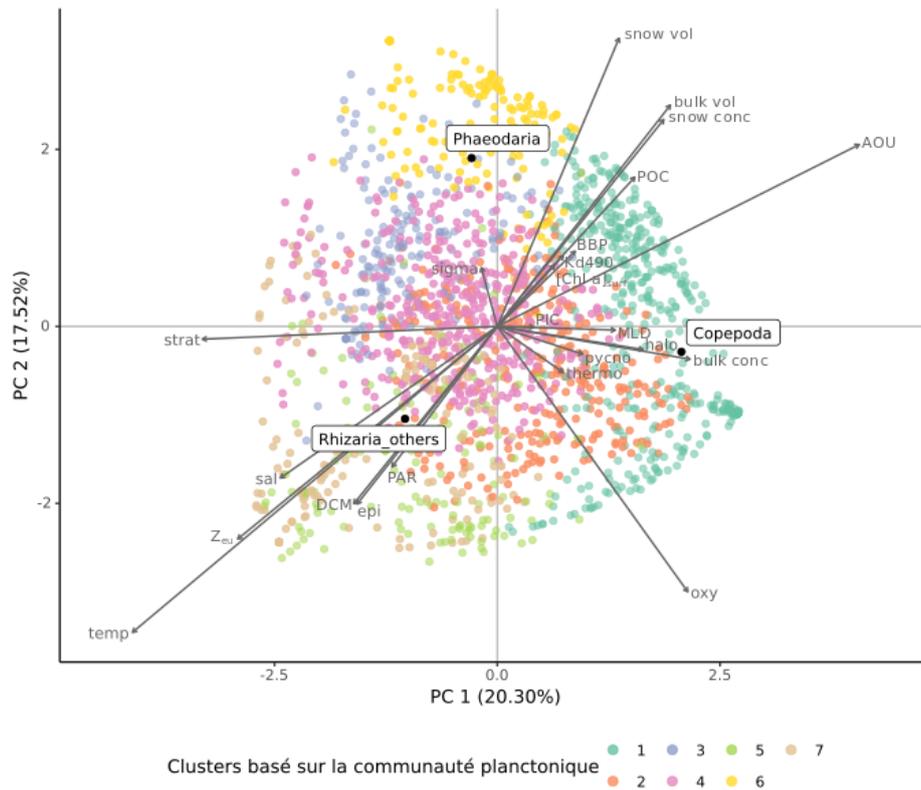


Figure 7 – Projection des stations de la couche mésopélagique supérieure sur les deux premiers axes de l’ACP colorées en fonction des groupes définis par la CAH ; des catégories taxonomiques et des variables environnementales. Toutes les variables environnementales sont projetées, tandis que seules les catégories taxonomiques dont la norme de la projection sur les deux premiers axes est supérieure au rayon du cercle des contributions équivalentes le sont.

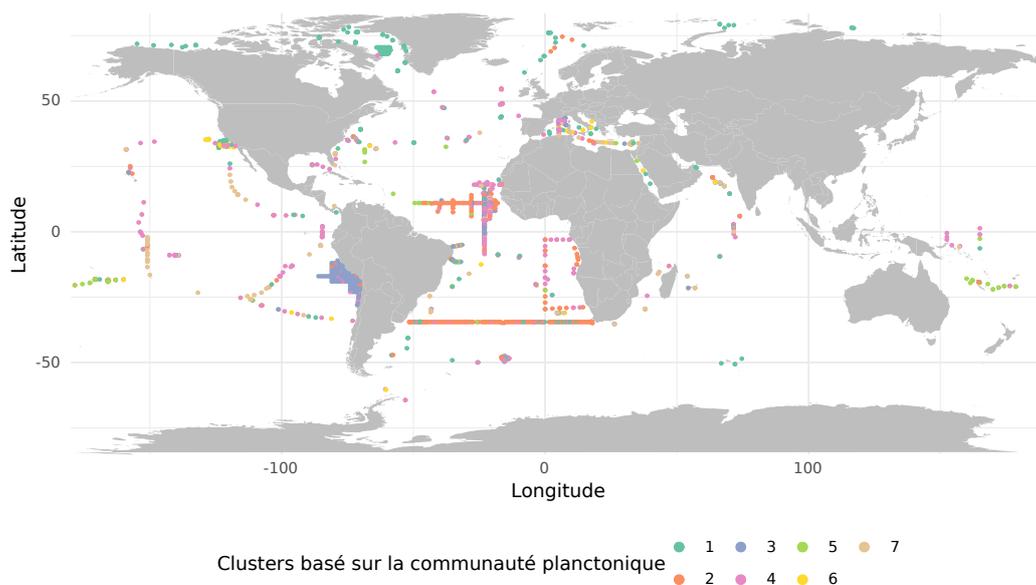


Figure 8 – Carte des stations de la couche mésopélagique supérieure colorées en fonction des groupes définis par la CAH.

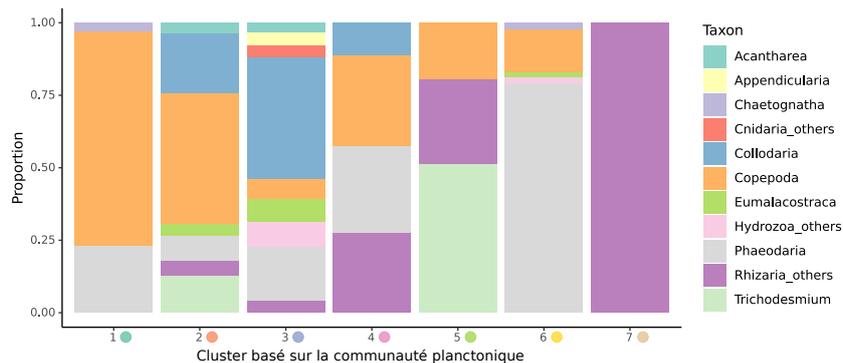


Figure 9 – Abondance relative des groupes taxonomiques dans les groupes définis par la CAH sur la couche mésopélagique supérieure. Seuls les taxa présents dans au moins 50% des profils sont représentés.

60,2% sur les 4 premiers. Le patron général est similaire à celui observé dans la couche épipélagique. L'axe 1 sépare les stations riches en copépodes (groupes 1 et 2) associées à des eaux froides, oxygénées et globalement riches en POC, particules et neige marine des stations riches en rhizaires, associées à des eaux plutôt chaudes, salées et oligotrophes. Cette répartition est bien marquée sur la carte où les stations des groupes 1 et 2 associées à des eaux eutrophes et productives se retrouvent principalement dans l'océan Arctique et au sud de l'océan Atlantique (Figure 8). L'axe 2 sépare les stations en fonction des rhizaires présents : phaeodaires (groupes 3 et 6) ou autres rhizaires (groupes 5 et 7), ces derniers étant associés aux eaux chaudes et oligotrophes.

Le dendrogramme généré par la CAH a été divisé en sept groupes (Figure 9). Le groupe 1, dominé par les copépodes et phaeodaires, se concentre dans la baie de Baffin et dans l'océan Arctique, ce qui correspond à des eaux froides et productives. Le groupe 2, dominé par les copépodes et collodaires et contenant également des trichodesmiums, se retrouve dans au sud de l'océan Atlantique ainsi qu'au niveau de la bande intertropicale. Le groupe 3, plus diversifié que les autres, mais tout de même dominé par les collodaires et phaeodaires, est essentiellement localisé dans la zone d'upwelling située au large du Pérou, dans des eaux appauvries en oxygène. Le groupe 4 contient un assemblage en proportions identiques de copépodes, phaeodaires et autres rhizaires, ainsi que des collodaires. Les stations de ce groupe sont distribuées dans tous les océans entre 50°N et 50°S. Le groupe 5 où se concentrent les trichodesmiums dont la proportion est similaire à la proportion cumulée des copépodes et autres rhizaires, se retrouve principalement au sud-ouest de l'océan Pacifique, ce qui correspond à des eaux chaudes et oligotrophes. Ainsi la distribution des stations des deux groupes caractérisés par la présence de

trichodesmiums dans la couche mésopélagique supérieure (groupes 2 et 5) est très similaire à celle des stations contenant des trichodesmiums dans la couche épipélagique, situées dans la bande intertropicale de l'océan Atlantique ainsi qu'au sud-ouest de l'océan Pacifique. Le groupe 6, dominé par les phaeodaires, se concentre en mer Méditerranée et au large du courant de Californie. Enfin, le groupe 7 est dominé par des rhizaires non identifiés. Ce groupe contient un nombre moindre de stations, qui sont réparties dans la mer Méditerranée et dans la bande intertropicale de l'océan Pacifique, c'est-à-dire des eaux plutôt chaudes et oligotrophes.

La distribution des grands types de communautés décrites dans la couche mésopélagique supérieure apparaît donc comme moins délimitée que dans la couche épipélagique, bien que les conditions environnementales, notamment la température et la richesse du milieu, semblent toujours contraignantes pour la distribution des communautés de plancton dans la couche mésopélagique supérieure. Ces communautés sont structurées par les copépodes et deux groupes de rhizaires : les phaeodaires et des rhizaires non identifiés.

3.2.3 Couche mésopélagique inférieure

1191 profils ont été retenus pour la description des communautés de la couche mésopélagique inférieure. L'ACP capture 38,2% de variance sur les deux premiers axes (Figure S5) et 56,7% sur les 4 premiers. Le patron général est très similaire à celui observé dans la couche mésopélagique supérieure. L'axe 1 sépare les stations riches en rhizaires non identifiés (groupes 4 et 5) de celles riches en copépodes ou en collodaires et phaeodaires. L'axe 2 sépare les stations riches en copépodes (groupes 1 et 3) des stations riches en phaeodaires et collodaires (groupe 2). Les copépodes sont associés à des eaux froides et oxygénées ; alors que les phaeodaires et collodaires sont associés à des eaux riches en particules et neige marine, tandis que les rhizaires non identifiés sont associés à des eaux plutôt chaudes.

Le dendrogramme généré par la CAH a été divisé en six groupes. Les groupes 1 et 3 riches en copépodes se répartissent au sud de l'océan Atlantique pour le groupe 1, dans l'océan Arctique et au large du courant de Californie pour le groupe 3, ce qui correspond à des eaux froides (Figure S6). Le groupe 2 riche en phaeodaires et collodaires est essentiellement situé dans la bande intertropicale dans l'océan Atlantique et au niveau de l'upwelling du Pérou. Les groupes 4 et 5 plutôt riches en rhizaires non identifiés se répartissent plutôt dans l'océan Pacifique.

Le groupe 6, de composition intermédiaire, se retrouve essentiellement dans les eaux oligotrophes de la mer Méditerranée. Les communautés de la couche mésopélagique inférieure sont donc structurées par les mêmes groupes que celles de la couche mésopélagique supérieure : copépodes, rhizaires non identifiés et phaeodaires auxquels semblent associés des collodaires, bien que la distribution spatiale des communautés semble moins marquée que dans la couche mésopélagique supérieure.

3.2.4 Couche bathypélagique

1041 profils ont été retenus pour la description des communautés de la couche mésopélagique inférieure. L'ACP capture 38,0% de variance sur les deux premiers axes (Figure S7) et 60,6% sur les 4 premiers. Le patron général diffère de ceux des couches épipélagique et mésopélagiques. L'axe 1 sépare les stations riches en copépodes (groupes 1, 2 et 4) des eaux oxygénées de stations associées à des eaux chaudes et pauvres en copépodes (groupe 3). L'axe 2 distingue les stations riches en rhizaires non identifiés (groupe 5) associées à des eaux oligotrophes.

Le dendrogramme issu de la CAH a été découpé en 5 groupes. Les groupes 1 et 2 plutôt riches en copépodes sont situés dans la bande intertropicale et au sud de l'océan Atlantique, ainsi qu'à proximité des côtes dans les zones d'upwelling du Pérou et du courant du Benguela (Figure S8). Le groupe 3, de composition intermédiaire, non enrichi en copépodes ou rhizaires et caractérisé par des eaux plus chaudes, est distribué dans tous les bassins océaniques. Les stations du groupe 4, caractérisé par la présence de copépodes et rhizaires non identifiés, sont situées majoritairement dans la bande intertropicale de l'océan Atlantique. Le groupe 5, caractérisé par sa richesse en rhizaires non identifiés, est principalement situé dans l'océan Pacifique.

Ainsi les communautés planctoniques bathypélagiques semblent principalement structurées par deux pôles : les copépodes et les rhizaires non identifiés.

3.2.5 Structuration des communautés avec la profondeur

Les valeurs propres des premiers axes de l'ACP sont plus élevées dans la couche épipélagique que dans les autres couches (Figure 10), expliquant une plus grande part de variance. Ainsi les communautés semblent plus structurées dans la couche épipélagique que dans les autres

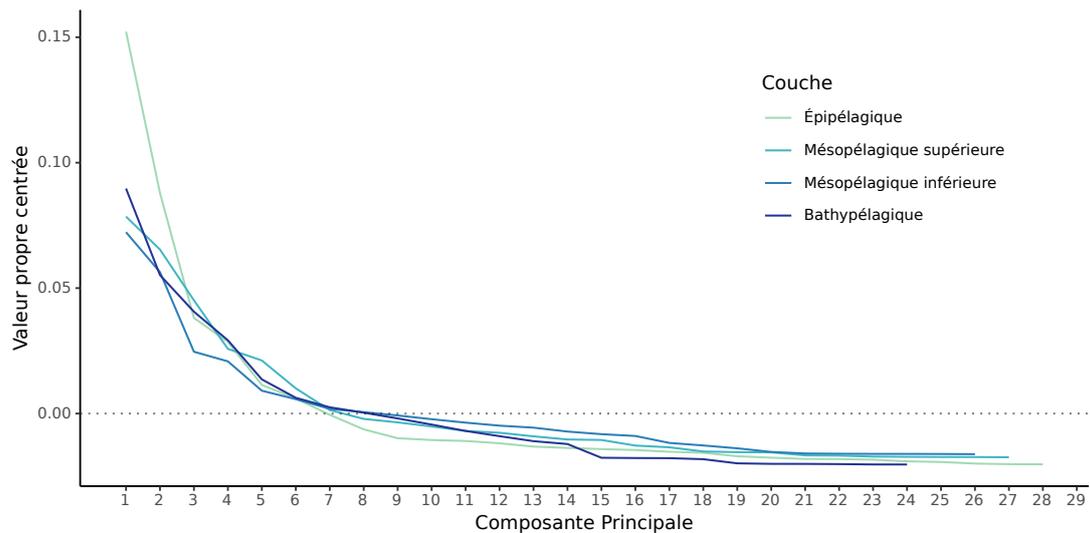


Figure 10 – Valeurs propres des axes des ACP pour les différentes couches corrigées par la moyenne des valeurs propres. Le nombre d’axes diminue pour les couches les plus profondes car certains taxa deviennent totalement absents et ne sont plus comptabilisés.

couches.

3.3 Comparaison avec d’autres motifs de partition de l’océan mondial

Le regroupement effectué directement sur les données de composition de la communauté de plancton (modèle nul) explique toujours une plus grande part de variance que les autres régionalisations (Table 3). Parmi les régionalisations testées, dans toutes les couches verticales, celle de Longhurst explique plus de variance que les autres. Cependant elle n’explique qu’entre un tiers et la moitié de la variance expliquée par le modèle nul. Dans la couche épipélagique, la partition en bandes de latitudes explique plus de variance que le regroupement basé sur les données environnementales locales. Dans les couches mésopélagiques supérieure et inférieure, les trois autres régionalisations (Reygondeau, bandes de latitude et environnement local) expliquent une part similaire de variance. Enfin, dans la couche bathypélagique les regroupements par bande de latitude et basé sur l’environnement expliquent la même part de variance. La proportion de variance expliquée n’est pas comparable entre couches, car le nombre de stations utilisées et de groupes considérés n’est pas le même.

Ainsi, les proportions taxonomiques dans les données mondiales de concentration de plancton dans la couche épipélagique sont mieux expliquées par les provinces biogéochimiques de Longhurst, basées sur des données biogéochimiques.

Table 3 – Proportion de variance expliquée par les différentes régionalisations pour chaque couche. n = nombre de groupes retenus pour chaque régionalisation.

Couche	Nb profils	Modèle nul		Environnement local		Bandes de latitude		Longhurst		Reygondeau	
		n	R ² _{adj}	n	R ² _{adj}	n	R ² _{adj}	n	R ² _{adj}	n	R ² _{adj}
Épipélagique	2238	13	57,0 %	11	19,6%	12	25,2%	14	32,5 %	-	-
Mésopélagique supérieur	1236	14	47,2 %	9	12,5%	10	10,5%	14	17,4 %	10	12,7%
Mésopélagique inférieur	694	9	43,7 %	8	13,7%	10	17,3%	12	22,9 %	9	16,1%
Bathypélagique	659	9	50,6 %	8	13,2%	9	14,8%	11	19,7 %	-	-

4 Discussion

4.1 Biais potentiels

Ce travail est potentiellement soumis à un certain nombre de biais, en raison de la nature même des données et de l'analyse spatiale à grande échelle qui en est faite. Il est possible que des biais liés à la méthodologie et la stratégie d'échantillonnage affectent les résultats. Ceux-ci sont discutés dans cette section.

4.1.1 Saisonnalité et rythme circadien

De nombreux taxa du plancton réalisent une migration verticale nyctémérale d'amplitudes variables (Lampert, 1989, Ohman and Romagnan, 2016). Cependant, nos analyses ne mettent pas en évidence un effet de la migration verticale sur la *composition* des communautés. Plusieurs explications sont envisageables. Tout d'abord les données de concentration ont été modifiées par la transformation d'Hellinger, focalisant ainsi notre analyse sur la composition des *communautés* plutôt que sur la valeur absolue des concentrations. Ainsi si la migration nyctémérale influence les concentrations (souvent plus élevées la nuit dans la couche épipélagique) mais que les groupes dominants ne changent pas (à la résolution taxonomique permise par l'UVP5 au moins), nous ne détecterions pas d'effet. De plus, nous utilisons une définition dynamique de la couche épipélagique, qui a plus de sens écologique mais dont la profondeur varie de quelques dizaines de mètres à plus de 300 m. La migration verticale pourrait donc avoir lieu au sein même de la couche épipélagique définie ici et ne pourrait donc être détectée dans les données moyennées par couche. L'étude de la concentration moyenne dans la couche épipélagique des cinq groupes taxonomiques les plus abondants (copépodes, phaeodaires, acanthaires, eumalacostracés et narcoméduses), au large du courant de Californie révèle que les

copépodes, eumalacostracés et narcoméduses sont effectivement plus abondants la nuit que le jour (Figure S9 en annexe). Ainsi, à une échelle locale et en travaillant sur des concentrations absolues, la migration nyctémérale est détectée pour les organismes dominants. Il semble donc que la première explication soit à favoriser : la migration existe, est détectable sur les concentrations mais n'influence que très marginalement la composition des communautés. La saisonnalité affecte l'abondance du phytoplancton (Cloern and Jassby, 2010) et du zooplancton (Licandro et al., 2012, Mazzocchi et al., 2007, Beaugrand and Kirby, 2018).

Cependant, il est difficile d'estimer le niveau de productivité à partir de nos données. Les augmentations saisonnières de productivité sont marquées par des blooms saisonniers, qui restent des événements soudains, imprévisibles et spatialement restreints pouvant concerner le phytoplancton ou le zooplancton (Graham et al., 2001). D'autres événements de blooms peuvent avoir lieu en dehors des blooms saisonniers, il s'agit de blooms intermittents, en réponse à des conditions environnementales particulièrement favorables (McGillicuddy et al., 2007). D'après les concentrations de trichodesmiums observées dans nos données, il semble qu'un certain nombre de profils réalisés dans la zone intertropicale de l'océan Atlantique ainsi que dans le sud-ouest de l'océan Pacifique l'ont été pendant une période de bloom de trichodesmiums. Cependant, avec la résolution et la couverture spatio-temporelle de ces données, l'effet spatial à grande échelle semble dominer. Le cycle saisonnier affecte les abondances, mais semble avoir peu d'effet sur la composition de la communauté.

D'autre part, les régionalisations testées ici présentent des frontières fixes pour toutes les saisons. La prise en compte de l'effet saisonnier dans la délimitation de provinces biogéochimiques a permis de mettre en évidence des changements dans la répartition des provinces. Cependant ces changements affectent très peu le patron global de la régionalisation (Reygondeau et al., 2013).

4.1.2 Efficacité de détection des organismes

Deux sources de biais sont possibles pour la détection des organismes en raison de la morphologie (taille, détail morphologique comme les antennes) et du comportement des organismes : l'évitement et la mauvaise détection des petits organismes. Il est tout à fait possible que les organismes planctoniques capables de nager puissent percevoir l'arrivée de la rosette

et ainsi l'éviter. Cependant, les bonnes corrélations entre les abondances de copépodes de grande taille estimées par l'UVP et par traits de filets semblent exclure cette possibilité (Forest et al., 2012).

Concernant la taille des organismes, l'UVP5 stocke les images des organismes avec une taille supérieure à 600 μm correspondant à la taille à partir de laquelle il est possible de distinguer les objets. Mais compte tenu de l'orientation ou du comportement des organismes dans leur milieu le succès de reconnaissance est variable pour une même taille. Par exemple dans le cas des copépodes, la détection des antennes est un élément indispensable pour l'identification des petits organismes, les copépodes de grande taille sont donc mieux identifiés que les petits.

Par ailleurs, parmi les grands objets détectés par l'UVP5, une partie importante correspond à des filaments de plusieurs millimètres de long. Ces filaments ont été identifiés comme des diatomées par des prélèvements ponctuels. Mais dans la majorité des cas, ces filaments ne sont pas distinguables uniquement à partir des images et ont été regroupés dans la catégorie fibre détritique, qui contient également les pelotes fécales. Nous avons donc décidé de ne pas inclure ce groupe. L'inclusion de ce groupe aurait pu modifier la structure des communautés qui ressort de nos analyses.

Nos résultats sont donc valables pour une fraction seulement du plancton, dont la taille varie de 600 μm à quelques cm. Cette fraction représente une faible portion de la biomasse planctonique, puisque la biomasse du macrozooplancton d'une taille supérieure à 2 mm (0,02 PgC) (Moriarty et al., 2013) compte pour 10% de la biomasse du zooplancton dont la taille est comprise entre 200 μm et 2 cm (0,19 PgC) (Moriarty and O'Brien, 2013). Toutefois, dans les systèmes arctiques ce sont les copépodes de grande taille qui représentent la plus grande partie de la biomasse (Darnis et al., 2008). Par ailleurs la biomasse représentée par les rhizaires de grande taille (> 600 μm) a été estimée à 0,089 PgC, soit 29% des biomasses combinées du méso et du macrozooplancton (Biard et al., 2016). Même si l'UVP5 ne détecte qu'une fraction des objets planctoniques, la biomasse associée n'est pas négligeable.

4.1.3 Précision de la taxonomie

Bien qu'un guide taxonomique ait été mis en place pour la classification des imagerie générées par l'UVP5, la classification n'est pas parfaitement homogène entre projets. Par ailleurs, la validation manuelle par un opérateur est une source de biais. Quelques erreurs de classifications sont apparues dans les données de certains projets, notamment pour les rhizaires, dont la classification des imagerie UVP reste difficile. En effet lors de certaines campagnes, certains organismes classés dans les collodaires appartenaient en réalité au groupe des phaeodaires. Cela est probablement à l'origine de la présence de collodaires dans les couches mésopélagique inférieure et bathypélagique dans nos analyses, alors que ce sont des organismes habituellement présents dans la couche épipélagique (Biard et al., 2016). Nous avons vérifié l'effet de ces imprécisions de deux façons : en éliminant de l'analyse des campagnes les plus touchés et en regroupant les groupes de collodaires dans lesquels figuraient des phaeodaires avec les rhizaires non identifiés pour ne conserver dans les collodaires que les collodaires coloniaux dont l'identification est aisée. Dans les deux cas, les analyses montrent un signal similaire au signal original. Ainsi, l'effet des erreurs de classification reste mineur pour une analyse globale de la distribution des communautés de plancton.

D'autre part, les différents groupes de rhizaires apparaissent comme des éléments structurant les communautés de plancton. Une raison à l'origine de ce résultat pourrait être la précision de la classification des rhizaires, plus élevée que pour les autres groupes. En effet si d'autres groupes avaient été classés plus précisément, par exemple les copépodes en calanoides, cyclopoïdes et harpacticoides, ce qui n'est pas possible avec les images issues de l'UVP, la communauté aurait peut-être été structurée par les différents groupes de copépodes plutôt que par les différents groupes de rhizaires. Cependant, les groupes de rhizaires identifiés ici (collodaires, phaeodaires, acanthaires et foraminifères) ont des milieux de vie très différents et il n'est donc pas étonnant de les retrouver comme groupes clés pour la structure des communautés de plancton.

4.1.4 Distribution des profils

Les profils UVP5 utilisés n'ont pas une distribution spatio-temporelle homogène. Certaines zones ont été particulièrement échantillonnées (courant de Californie, upwelling du Pérou, mer

Méditerranée) tandis que d'autres sont très peu représentées (océan Indien, océan Austral) (Figure S1). Pour vérifier l'effet de la surreprésentation de certaines régions, nous avons effectué les analyses sur un sous-échantillon des données en fixant un seuil de densité de profils à 50 profils par carré de 2° de côté. Les patrons issus de cette analyse sont identiques à ceux de l'analyse incluant tous les profils, ils semblent donc robustes.

4.2 Structure générale des communautés de plancton

4.2.1 Couche épipélagique

Les communautés épipélagiques sont essentiellement structurées par les proportions de trichodesmiums, de collodaires et de copépodes, des métazoaires. La dominance de ces organismes est conditionnée par la température : les copépodes occupent les eaux froides tandis que les trichodesmiums et rhizaires se retrouvent plutôt dans les eaux chaudes. La présence des copépodes dans la plupart des écosystèmes est déjà bien connue (Gallienne et al., 2001, Turner, 2004, Champalbert et al., 2005). Leur distribution est largement contrainte par les conditions environnementales (Beaugrand and Reid, 2003, Rombouts et al., 2009, Chertoprud et al., 2010). C'est un groupe clé, jouant le rôle de maillon trophique entre le phytoplancton et les niveaux trophiques supérieurs (Rombouts et al., 2009). Par ailleurs cette étude souligne l'importance écologique des rhizaires, récemment mise en évidence par d'autres études basées sur de l'imagerie *in situ* (Dennett et al., 2002, Biard et al., 2016). L'importance écologique de ces organismes fragiles a longtemps été sous-estimée à cause de leur dégradation par les méthodes d'échantillonnage historiques. Plus particulièrement, les rhizaires identifiés dans la couche épipélagique sont essentiellement des collodaires, dont toutes les espèces décrites actuellement vivent en symbiose avec des microalgues photosynthétiques intracellulaires (Biard et al., 2017). Il est donc logique de retrouver ces organismes à proximité de la surface où leurs symbiontes peuvent bénéficier du rayonnement solaire. Enfin les communautés épipélagiques sont également structurées par les concentrations relatives de trichodesmiums. Ces cyanobactéries filamenteuses se retrouvent en grande abondance à la surface de zones oligotrophes tropicales et subtropicales lors de périodes de bloom qui restent cependant relativement rares (Westberry and Siegel, 2006). Ces blooms ont été observés dans la plupart des bassins océaniques (Sellner, 1992). Les trichodesmiums sont capables de fixer le diazote atmosphérique,

et sont ainsi à l'origine un apport majeur d'azote vers les océans (Capone et al., 1997). Ainsi la pauvreté des eaux en azote n'est pas un facteur limitant pour la croissance des trichodesmiums.

4.2.2 Couches mésopélagique et bathypélagique

Dans la couche mésopélagique supérieure (jusqu'à 500 m), les communautés sont structurées par les proportions de copépodes et de rhizaires. Le motif diffère de celui de la couche épipélagique : les trichodesmiums ne sont pas un pôle structurant dans la couche mésopélagique supérieure, mais sont toujours présents ; et les rhizaires structurant la couche mésopélagique supérieure ne sont pas des collodaires, mais des phaeodaires et des rhizaires non identifiés. Comme dans la couche épipélagique, la répartition des communautés de plancton est fortement conditionnée par l'environnement, notamment la température, mais cela est moins marqué que dans la couche épipélagique. Bien que photosynthétiques, des trichodesmiums ont déjà été détectés à plus de 100 m de profondeur (Walsby, 1978, Davis and McGillicuddy, 2006). Ils sont présents ici dans la couche mésopélagique supérieure, car la frontière supérieure de cette dernière est calculée en fonction de la profondeur de la couche euphotique et de la pycnocline. Dans certains cas notre couche mésopélagique supérieure débute dès 14 mètres de profondeur, expliquant l'inclusion des trichodesmiums dans cette couche. Par ailleurs, les trichodesmiums pourraient également être emportés en profondeur lors d'évènements de downwelling (Guidi et al., 2012). De tels évènements résulteraient en la mort des colonies, mais celles-ci pourraient rester identifiable sur les images de l'UVP. Comme dans la couche épipélagique, les rhizaires sont largement présents dans les communautés de la couche mésopélagique supérieure, à la différence qu'on retrouve ici des phaeodaires à la place des collodaires. En effet, alors que les collodaires hébergent des symbiotes photosynthétiques, les phaeodaires sont des organismes hétérotrophes apparemment dépourvus de symbiotes (Kling and Boltovskoy, 1999). Ils se rencontrent essentiellement en profondeur, au-delà de 300 m (Boltovskoy et al., 2016), et constituent la part la plus importante de la biomasse des rhizaires dans la tranche 100-500 m (Biard et al., 2016).

Dans la couche mésopélagique inférieure (500 - 1000 m), le patron est très similaire à celui de la couche mésopélagique supérieure, à la différence qu'il n'y a plus de trichodesmiums. La présence de collodaires à ces profondeurs est inattendue, car ces organismes hébergent des

symbiontes photosynthétiques, cela est probablement causé par une mauvaise classification dans les collodaires d'organismes qui sont en fait des phaeodaires. Cette hypothèse est renforcée par la projection très similaire des collodaires et phaeodaires sur les deux premiers axes de l'ACP.

Enfin dans la couche bathypélagique, les communautés sont structurées par les copépodes et des rhizaires non identifiés dont la répartition est plus homogène que dans les couches moins profondes.

Au final, les communautés de plancton sont structurées par des groupes taxonomiques très éloignés. On aurait pu s'attendre à ce que la structure de ces communautés soit dirigée par des organismes dont la biomasse est importante et dont les organismes seraient principalement de grands métazoaires, notamment du meso et macrozooplancton autre que les copépodes. Or la structure mise en évidence ici pour les organismes de grande taille repose sur un groupe de métazoaires, les copépodes, un groupe de protistes mixotrophes, les rhizaires et un groupe de cyanobactéries, les trichodesmiums. Ces groupes clés pour la structure des communautés de plancton vivent dans des milieux écologiques très différents : les copépodes dominent dans les eaux froides, oxygénées, productives riches en particules ; tandis que les trichodesmiums sont essentiellement présents dans des eaux chaudes, salées, peu oxygénées et oligotrophes. Parmi les rhizaires, les collodaires et phaeodaires se concentrent principalement dans les zones d'upwelling, où les eaux sont plutôt froides et appauvries en oxygène ; les collodaires étant plus présents dans la couche épipélagique tandis que les phaeodaires dominent dans les couches plus profondes.

4.2.3 Homogénéisation en profondeur

Nos analyses mettent en évidence que la communauté de la couche épipélagique est plus structurée que celle des autres couches. Par ailleurs, la distribution spatiale des communautés dans la couche épipélagique est particulièrement marquée par la latitude, reflétant l'importance de la température et de l'ensoleillement, alors que les délimitations sont moins marquées dans les couches profondes. Cela est en accord avec le fait que les conditions environnementales ont tendance à s'homogénéiser en profondeur (Costello and Breyer, 2017, Costello et al., 2018). De la même manière, les discontinuités marquées en surface dans la distribution des or-

ganismes planctoniques sont moins importantes en profondeur (Vecchione et al., 2015). Dans notre étude, l'effet sur la structure de communautés de plancton n'est pas visible entre les couches mésopélagiques et bathypélagique, ce qui peut s'expliquer par le fait que moins de profils étaient disponibles pour l'analyse de ces couches.

4.3 Une distribution contrainte par les conditions régionales

Les données environnementales utilisées pour la construction de notre régionalisation sont des données issues de CTD ainsi que des données satellites hebdomadaires. Ce sont donc des données très locales dans le temps et l'espace. Or nos résultats montrent que les variations dans la distribution des organismes planctoniques sont mieux expliquées par des régionalisations basées sur les données globales. Ce travail met en avant que la composition des communautés de plancton en un endroit localisé dans le temps et l'espace dépend essentiellement des conditions régionales et à plus long terme que des conditions locales et immédiates.

5 Perspectives

La principale perspective à l'issue de ce travail est de le publier dans un journal à comité de lecture. Dans un premier temps, il faudra améliorer la classification des rhizaires dans un certain nombre de projets, ainsi qu'estimer les biomasses des différents groupes pour évaluer l'importance biogéochimique. À plus long terme, l'amélioration des performances d'identification ainsi que le développement de l'approche par les traits plutôt que la taxonomie permettront de mieux caractériser les communautés de plancton. Enfin, l'augmentation de la résolution spatiale et temporelle des données est un autre axe à développer. Notamment, la commercialisation de l'UVP6 pouvant être déployé de manière autonome sur des flotteurs et bénéficiant de la reconnaissance embarquée permettra de générer de nouvelles données de concentrations et de biovolumes des organismes planctoniques et de la neige marine *in situ*.

Références

- Banse, K., 1995. Zooplankton: Pivotal role in the control of ocean production. *ICES J. Mar. Sci.* 52(3-4), 265–277.
- Beaugrand, G., Edwards, M., Legendre, L., 2010. Marine biodiversity, ecosystem functioning, and carbon cycles. *Proc. Natl. Acad. Sci.* 107(22), 10120–10124.
- Beaugrand, G., Ibañez, F., Lindley, J., Philip, C., Reid, P., 2002. Diversity of calanoid copepods in the North Atlantic and adjacent seas: species associations and biogeography. *Mar. Ecol. Prog. Ser.* 232, 179–195.
- Beaugrand, G., Kirby, R.R., 2018. How Do Marine Pelagic Species Respond to Climate Change? Theories and Observations. *Ann. Rev. Mar. Sci.* 10(1), 169–197.
- Beaugrand, G., Reid, P.C., 2003. Long-term changes in phytoplankton, zooplankton and salmon related to climate. *Glob. Chang. Biol.* 9(6), 801–817.
- Behrenfeld, M.J., Boss, E., 2014. Resurrecting the Ecological Underpinnings of Ocean Plankton Blooms. *Ann. Rev. Mar. Sci.* 6(1), 167–194.
- Behrenfeld, M.J., Randerson, J.T., McClain, C.R., Feldman, G.C., Los, S.O., Tucker, C.J., Falkowski, P.G., Field, C.B., Frouin, R., Esaias, W.E., Kolber, D.D., Pollack, N.H., 2001. Biospheric primary production during an ENSO transition. *Science* 291(5513), 2594–2597.
- Biard, T., Bigeard, E., Audic, S., Poulain, J., Gutierrez-Rodriguez, A., Pesant, S., Stemmann, L., Not, F., 2017. Biogeography and diversity of Collodaria (Radiolaria) in the global ocean. *ISME J.* 11(6), 1331–1344.
- Biard, T., Stemmann, L., Picheral, M., Mayot, N., Vandromme, P., Haus, H., Gorsky, G., Guidi, L., Kiko, R., Not, F., 2016. In situ imaging reveals the biomass of giant protists in the global ocean. *Nature* 532(7600), 504–507.
- Boltovskoy, D., Anderson, O.R., Correa, N.M., 2016. Radiolaria and Phaeodaria. In: *Handb. Protists*, pp. 1–33. Springer International Publishing, Cham.
- Borcard, D., Gillet, F., Legendre, P., 2018. Numerical ecology with R. Springer.

- Capone, D.G., Zehr, J.P., Paerl, H.W., Bergman, B., Carpenter, E.J., 1997. *Trichodesmium*, a Globally Significant Marine Cyanobacterium. *Science* (80-). 276(5316), 1221–1229.
- Champalbert, G., Pagano, M., Kouamé, B., Riandey, V., 2005. Zooplankton Spatial and Temporal Distribution in a Tropical Oceanic Area off West Africa. *Hydrobiologia* 548(1), 251–265.
- Chaudhary, C., Saeedi, H., Costello, M.J., 2016. Bimodality of Latitudinal Gradients in Marine Species Richness. *Trends Ecol. Evol.* 31(9), 670–676.
- Chertoprud, E.S., Garlitska, L.A., Azovsky, A.I., 2010. Large-scale patterns in marine harpacticoid (Crustacea, Copepoda) diversity and distribution. *Mar. Biodivers.* 40(4), 301–315.
- Cloern, J.E., Jassby, A.D., 2010. Patterns and Scales of Phytoplankton Variability in Estuarine–Coastal Ecosystems. *Estuaries and Coasts* 33(2), 230–241.
- Costello, M.J., Basher, Z., Sayre, R., Breyer, S., Wright, D.J., 2018. Stratifying ocean sampling globally and with depth to account for environmental variability. *Sci. Rep.* 8(1), 11259.
- Costello, M.J., Breyer, S., 2017. Ocean Depths: The Mesopelagic and Implications for Global Warming. *Curr. Biol.* 27(1), R36–R38.
- Darnis, G., Barber, D.G., Fortier, L., 2008. Sea ice and the onshore–offshore gradient in pre-winter zooplankton assemblages in southeastern Beaufort Sea. *J. Mar. Syst.* 74(3-4), 994–1011.
- Davis, C.S., McGillicuddy, D.J., 2006. Transatlantic abundance of the N₂-fixing colonial cyanobacterium *Trichodesmium*. *Science* 312(5779), 1517–1520.
- de Boyer Montégut, C., Madec, G., Fischer, A.S., Lazar, A., Iudicone, D., 2004. Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109(C12), C12003.
- de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Bescot, N.L., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mulo, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Coordinators, T.O., 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., Weisenbach, J., Wincker, P., Karsenti, E., 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* (80-.). 348(6237), 1261605.
- Dennett, M.R., Caron, D.A., Michaels, A.F., Gallagher, S.M., Davis, C.S., 2002. Video plankton recorder reveals high abundances of colonial Radiolaria in surface waters of the central North Pacific. *J. Plankton Res.* 24(8), 797–805.
- Emerson, S., Quay, P., Karl, D., Winn, C., Tupas, L., Landry, M., 1997. Experimental determination of the organic carbon flux from open-ocean surface waters. *Nature* 389(6654), 951–954.
- Falkowski, P., 2012. Ocean Science: The power of plankton. *Nature* 483(7387), S17–S20.
- Forest, A., Stemmann, L., Picheral, M., Burdorf, L., Robert, D., Fortier, L., Babin, M., 2012. Size distribution of particles and zooplankton across the shelf-basin system in southeast Beaufort Sea: combined results from an Underwater Vision Profiler and vertical net tows. *Biogeosciences* 9(4), 1301–1320.
- Frederiksen, M., Edwards, M., Richardson, A.J., Halliday, N.C., Wanless, S., 2006. From plankton to top predators: bottom-up control of a marine food web across four trophic levels. *J. Anim. Ecol.* 75(6), 1259–1268.
- Gallienne, C., Robins, D., Woodd-Walker, R., 2001. Abundance, distribution and size structure of zooplankton along a 20° west meridional transect of the northeast Atlantic Ocean in July. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 48(4-5), 925–949.
- Gorsky, G., Ohman, M.D., Picheral, M., Gasparini, S., Stemmann, L., Romagnan, J.B., Cawood, A., Pesant, S., Garcia-Comas, C., Prejger, F., 2010. Digital zooplankton image analysis using the ZooScan integrated system. *J. Plankton Res.* 32(3), 285–303.
- Graham, W.M., Pagès, F., Hamner, W.M., 2001. A physical context for gelatinous zooplankton aggregations: a review. In: *Jellyfish Bloom. Ecol. Soc. Importance*, pp. 199–212. Springer Netherlands, Dordrecht.
- Guidi, L., Calil, P.H.R., Duhamel, S., Björkman, K.M., Doney, S.C., Jackson, G.A., Li, B., Church, M.J., Tozzi, S., Kolber, Z.S., Richards, K.J., Fong, A.A., Letelier, R.M., Gorsky, G., Stemmann, L., Karl, D.M., 2012. Does eddy-eddy interaction control surface phytoplankton distribution and

- carbon export in the North Pacific Subtropical Gyre? *J. Geophys. Res. Biogeosciences* 117, G02024.
- Guidi, L., Legendre, L., Reygondeau, G., Uitz, J., Stemmann, L., Henson, S.A., 2015. A new look at ocean carbon remineralization for estimating deepwater sequestration. *Global Biogeochem. Cycles* 29(7), 1044–1059.
- Hays, G.C., Richardson, A.J., Robinson, C., 2005. Climate change and marine plankton. *Trends Ecol. Evol.* 20(6), 337–344.
- Ikeda, T., 1985. Metabolic rates of epipelagic marine zooplankton as a function of body mass and temperature. *Mar. Biol.* 85(1), 1–11.
- Kling, S.A., Boltovskoy, D., 1999. Radiolaria Phaeodaria. *South Atl. Zooplankt.* 1, 231–264.
- Lalli, C.M., Parsons, T.R., 1997. Phytoplankton and Primary Production. In: C.M. Lalli, T. Parsons, eds., *Biol. Oceanogr. An Introd.*, pp. 39–73. Butterworth-Heinemann, Oxford.
- Lampert, W., 1989. The Adaptive Significance of Diel Vertical Migration of Zooplankton. *Funct. Ecol.* 3(1), 21–27.
- Legendre, P., Legendre, L., 2012. *Numerical ecology*. Elsevier.
- Leys, C., Ley, C., Klein, O., Bernard, P., Licata, L., 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* 49(4), 764–766.
- Li, W.K., Head, E.J., Glen Harrison, W., 2004. Macroecological limits of heterotrophic bacterial abundance in the ocean. *Deep Sea Res. Part I Oceanogr. Res. Pap.* 51(11), 1529–1540.
- Licandro, P., Souissi, S., Ibanez, F., Carré, C., 2012. Long-term variability and environmental preferences of calycophoran siphonophores in the Bay of Villefranche (north-western Mediterranean). *Prog. Oceanogr.* 97-100, 152–163.
- Longhurst, A.R., 1995. Seasonal cycles of pelagic production and consumption. *Prog. Oceanogr.* 36(2), 77–167.
- Longhurst, A.R., 2010. *Ecological geography of the sea*. Academic Press.

- Longhurst, A.R., Glen Harrison, W., 1989. The biological pump: Profiles of plankton production and consumption in the upper ocean. *Prog. Oceanogr.* 22(1), 47–123.
- Mazzocchi, M.G., Christou, E.D., Capua, I.D., Fernández de Puellas, M., Fonda-Umani, S., Molinero, J.C., Nival, P., Siokou-Frangou, I., 2007. Temporal variability of *Centropages typicus* in the Mediterranean Sea over seasonal-to-decadal scales. *Prog. Oceanogr.* 72(2-3), 214–232.
- McGillicuddy, D.J., Anderson, L.A., Bates, N.R., Bibby, T., Buesseler, K.O., Carlson, C.A., Davis, C.S., Ewart, C., Falkowski, P.G., Goldthwait, S.A., Hansell, D.A., Jenkins, W.J., Johnson, R., Kosnyrev, V.K., Ledwell, J.R., Li, Q.P., Siegel, D.A., Steinberg, D.K., 2007. Eddy/wind interactions stimulate extraordinary mid-ocean plankton blooms. *Science* 316(5827), 1021–1026.
- Miyamoto, H., Kotori, M., Itoh, H., Nishida, S., 2014. Species diversity of pelagic chaetognaths in the Indo-Pacific region. *J. Plankton Res.* 36(3), 816–830.
- 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 Sens. Environ.* 111(1), 69–88.
- Morel, A., Maritorena, S., 2001. Bio-optical properties of oceanic waters: A reappraisal. *J. Geophys. Res.* 106(C4), 7163–7180.
- Moriarty, R., Buitenhuis, E.T., Le Quéré, C., Gosselin, M.P., 2013. Distribution of known macrozooplankton abundance and biomass in the global ocean. *Earth Syst. Sci. Data* 5(2), 241–257.
- Moriarty, R., O’Brien, T.D., 2013. Distribution of mesozooplankton biomass in the global ocean. *Earth Syst. Sci. Data* 5, 45–55.
- Ohman, M.D., Romagnan, J., 2016. Nonlinear effects of body size and optical attenuation on Diel Vertical Migration by zooplankton. *Limnol. Oceanogr.* 61(2), 765–770.
- Oksanen, J., Kindt, R., Legendre, P., O’Hara, B., Stevens, M.H.H., Oksanen, M.J., Suggests, M., 2007. The vegan package. *Community Ecol. Packag.* 10, 631–637.
- Picheral, M., Guidi, L., Stemmann, L., Karl, D.M., Iddaoud, G., Gorsky, G., 2010. The Underwater Vision Profiler 5: An advanced instrument for high spatial resolution studies of particle size spectra and zooplankton. *Limnol. Oceanogr. Methods* 8(9), 462–473.

- Reygondeau, G., Guidi, L., Beaugrand, G., Henson, S.A., Koubbi, P., MacKenzie, B.R., Sutton, T.T., Fioroni, M., Maury, O., 2018. Global biogeochemical provinces of the mesopelagic zone. *J. Biogeogr.* 45(2), 500–514.
- Reygondeau, G., Longhurst, A.R., Martinez, E., Beaugrand, G., Antoine, D., Maury, O., 2013. Dynamic biogeochemical provinces in the global ocean. *Global Biogeochem. Cycles* 27(4), 1046–1058.
- Robinson, C., Anderson, T.R., Arístegui, J., Carlson, C.A., Frost, J.R., Ghiglione, J.F., Hernández-León, S., Jackson, G.A., Koppelman, R., Quéguiner, B., Ragueneau, O., Rassoulzadegan, F., Robison, B.H., Tamburini, C., Tanaka, T., Wishner, K.F., Zhang, J., 2010. Mesopelagic zone ecology and biogeochemistry – a synthesis. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 57(16), 1504–1518.
- Rombouts, I., Beaugrand, G., Ibañez, F., Gasparini, S., Chiba, S., Legendre, L., 2009. Global latitudinal variations in marine copepod diversity and environmental factors. *Proc. R. Soc. B Biol. Sci.* 276(1670), 3053–3062.
- Rutherford, S., D’Hondt, S., Prell, W., 1999. Environmental controls on the geographic distribution of zooplankton diversity. *Nature* 400(6746), 749–753.
- Sellner, K.G., 1992. Trophodynamics of Marine Cyanobacteria Blooms. In: *Mar. Pelagic Cyanobacteria Trichodesmium other Diazotrophs*, pp. 75–94. Springer Netherlands, Dordrecht.
- Sherman, K., 2005. The large marine ecosystem approach for assessment and management of ocean coastal waters. In: T. Hennessey, J. Sutinen, eds., *Sustain. Large Mar. Ecosyst. Hum. Dimens.*, pp. 3–16. Elsevier.
- Spalding, M.D., Agostini, V.N., Rice, J., Grant, S.M., 2012. Pelagic provinces of the world: A biogeographic classification of the world’s surface pelagic waters. *Ocean Coast. Manag.* 60, 19–30.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. *Marine Ecoregions of the World: A Bioregionalization of Coastal and Shelf Areas*. *Bioscience* 57(7), 573–583.

- Stemmann, L., Hosiá, A., Youngbluth, M.J., Sjøiland, H., Picheral, M., Gorsky, G., 2008a. Vertical distribution (0–1000 m) of macrozooplankton, estimated using the Underwater Video Profiler, in different hydrographic regimes along the northern portion of the Mid-Atlantic Ridge. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 55(1-2), 94–105.
- Stemmann, L., Youngbluth, M.J., Robert, K., Hosiá, A., Picheral, M., Paterson, H., Ibañez, F., Guidi, L., Lombard, F., Gorsky, G., 2008b. Global zoogeography of fragile macrozooplankton in the upper 100-1000 m inferred from the underwater video profiler. *ICES J. Mar. Sci.* 65(3), 433–442.
- Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D.R., Alberti, A., Cornejo-Castillo, F.M., Costea, P.I., Cruaud, C., D’Ovidio, F., Engelen, S., Ferrera, I., Gasol, J.M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B.T., Royo-Llonch, M., Sarmiento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Tara Oceans coordinators, Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M.B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S.G., Bork, P., 2015. Structure and function of the global ocean microbiome. *Science* 348(6237), 1261359.
- Sutton, T.T., Clark, M.R., Dunn, D.C., Halpin, P.N., Rogers, A.D., Guinotte, J., Bograd, S.J., Angel, M.V., Perez, J.A.A., Wishner, K., Haedrich, R.L., Lindsay, D.J., Drazen, J.C., Vereshchaka, A., Piatkowski, U., Morato, T., Błachowiak-Samołyk, K., Robison, B.H., Gjerde, K.M., Pierrot-Bults, A., Bernal, P., Reygondeau, G., Heino, M., 2017. A global biogeographic classification of the mesopelagic zone. *Deep Sea Res. Part I Oceanogr. Res. Pap.* 126, 85–102.
- Tittensor, D.P., Mora, C., Jetz, W., Lotze, H.K., Ricard, D., Berghe, E.V., Worm, B., 2010. Global patterns and predictors of marine biodiversity across taxa. *Nature* 466(7310), 1098–1101.
- Turner, J.T., 2004. The importance of small planktonic copepods and their roles in pelagic marine food webs. *Zool. Stud.* 43(2), 255–266.
- Vecchione, M., Falkenhaus, T., Sutton, T., Cook, A., Gislason, A., Hansen, H.Ø., Heino, M., Miller, P.I., Piatkowski, U., Porteiro, F., Sjøiland, H., Bergstad, O.A., 2015. The effect of the North Atlantic Subpolar Front as a boundary in pelagic biogeography decreases with increasing depth and organism size. *Prog. Oceanogr.* 138, 105–115.

- Volk, T., Hoffert, M.I., 1985. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO₂ changes. carbon cycle Atmos. CO₂ Nat. Var. Archean to Present pp. 99–110.
- Walsby, A., 1978. The properties and buoyancy-providing role of gas vacuoles in *Trichodesmium* Ehrenberg. Br. Phycol. J. 13(2), 103–116.
- Ware, D.M., Thomson, R.E., 2005. Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific. Science (80-.). 308(5726), 1280–1284.
- Westberry, T.K., Siegel, D.A., 2006. Spatial and temporal distribution of *Trichodesmium* blooms in the world's oceans. Global Biogeochem. Cycles 20(4), GB4016.
- Woodd-Walker, R.S., Ward, P., Clarke, A., 2002. Large-scale patterns in diversity and community structure of surface water copepods from the Atlantic Ocean. Mar. Ecol. Prog. Ser. 236, 189–203.

Annexes

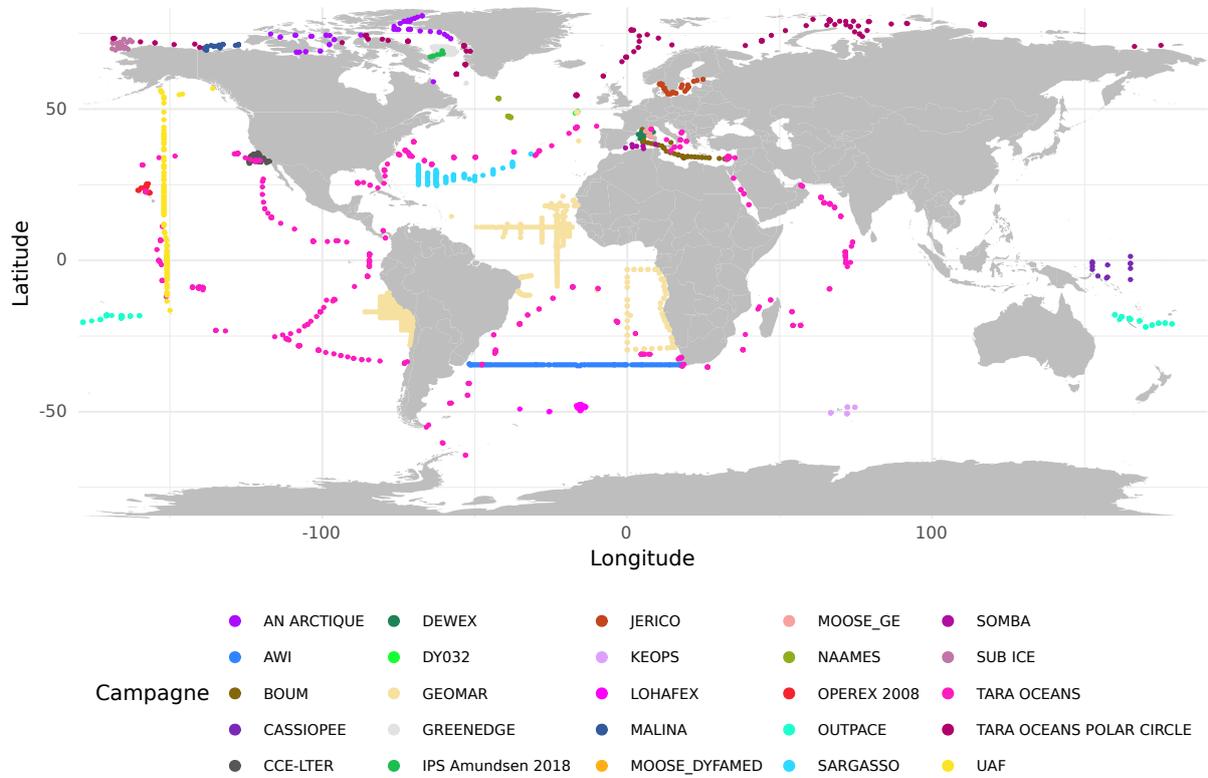


Figure S1 – Carte des stations des campagnes océanographiques incluses dans cette étude.

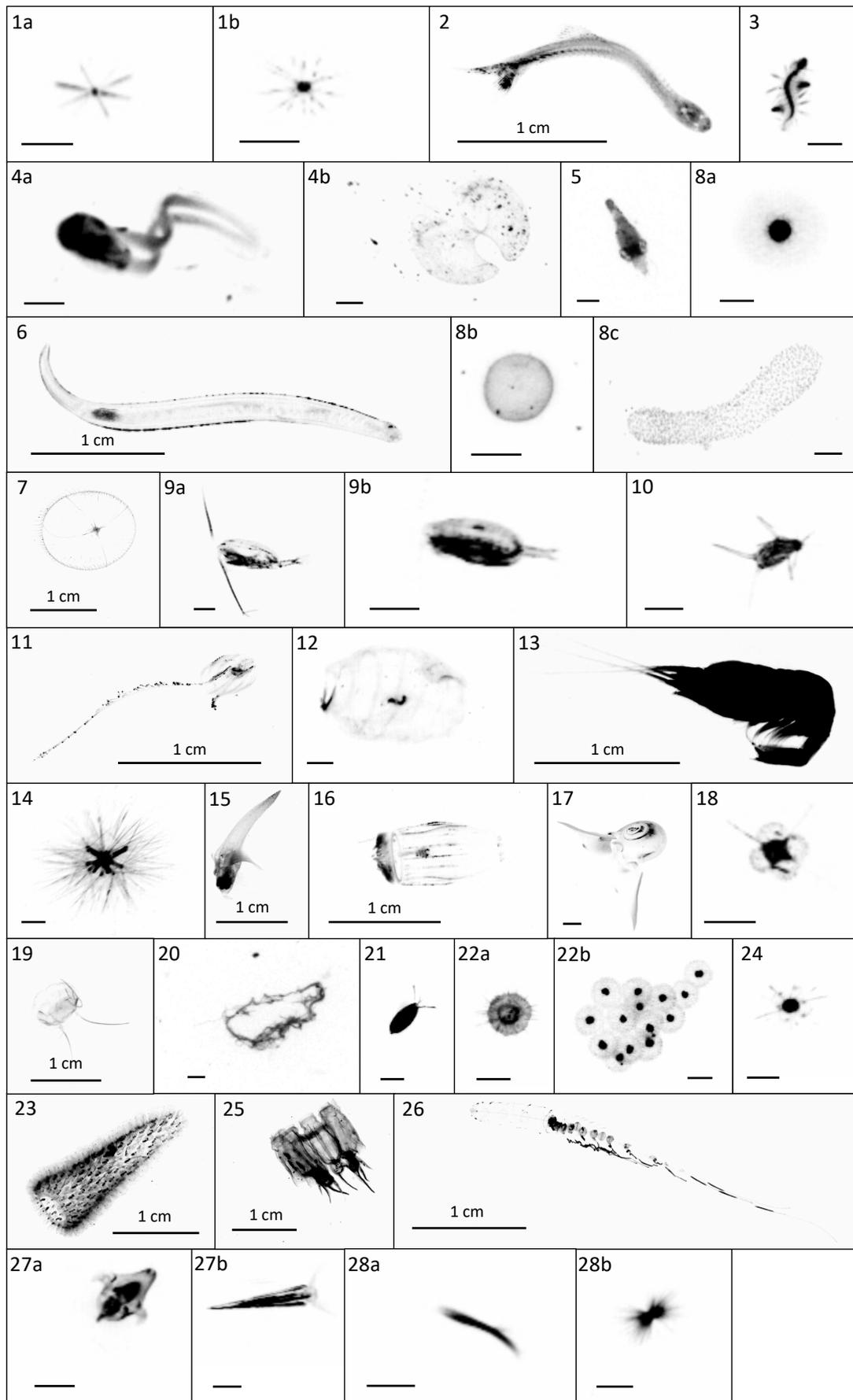
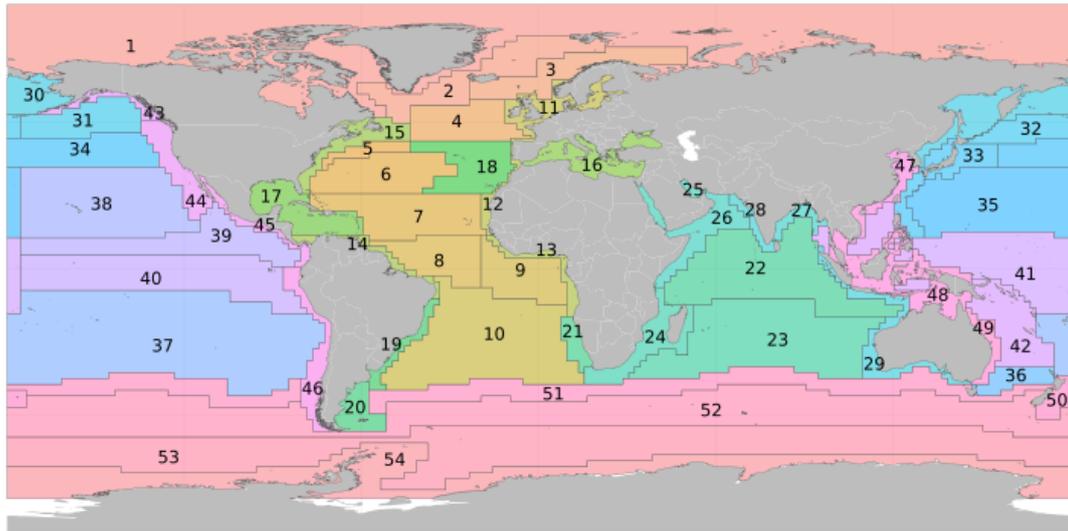


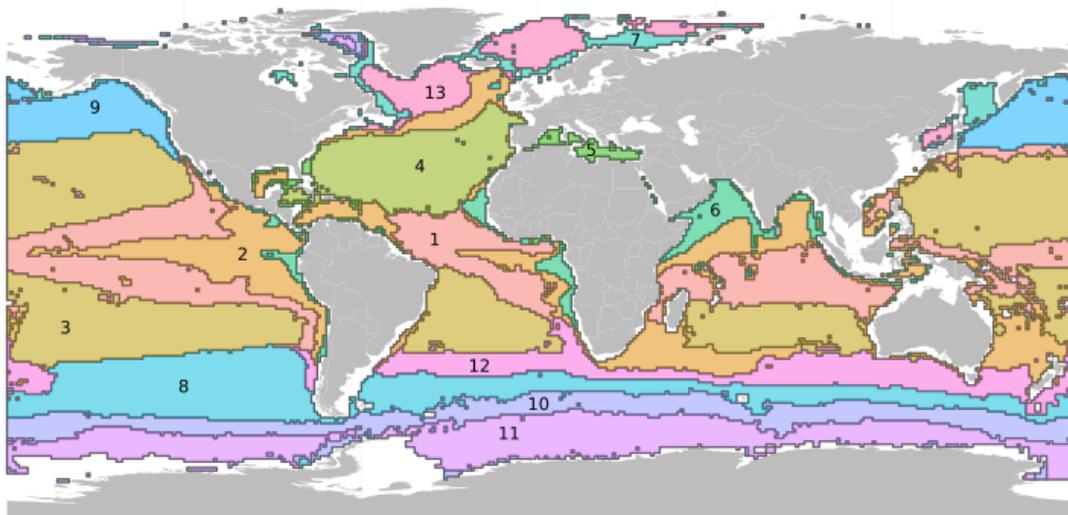
Figure S2 – Exemple de imagettes enregistrées par l’UVP5 pour les groupes taxonomiques retenus. Sauf mention contraire les barres d’erreur représentent 2 mm.



Provinces de Longhurst

- | | |
|--|---|
| 1 - Polar - Boreal Polar Province (POLR) | 28 - Coastal - W. India Coastal Province |
| 2 - Polar - Atlantic Arctic Province | 29 - Coastal - Australia-Indonesia Coastal Province |
| 3 - Polar - Atlantic Subarctic Province | 30 - Polar - N. Pacific Epicontinental Province |
| 4 - Westerlies - N. Atlantic Drift Province (WWDR) | 31 - Westerlies - Pacific Subarctic Gyres Province (East) |
| 5 - Westerlies - Gulf Stream Province | 32 - Westerlies - Pacific Subarctic Gyres Province (West) |
| 6 - Westerlies - N. Atlantic Subtropical Gyral Province (West) (STGW) | 33 - Westerlies - Kuroshio Current Province |
| 7 - Trades - N. Atlantic Tropical Gyral Province (TRPG) | 34 - Westerlies - N. Pacific Polar Front Province |
| 8 - Trades - Western Tropical Atlantic Province | 35 - Westerlies - N. Pacific Subtropical Gyre Province (West) |
| 9 - Trades - Eastern Tropical Atlantic Province | 36 - Westerlies - Tasman Sea Province |
| 10 - Trades - South Atlantic Gyral Province (SATG) | 37 - Westerlies - S. Pacific Subtropical Gyre Province |
| 11 - Coastal - NE Atlantic Shelves Province | 38 - Trades - N. Pacific Tropical Gyre Province |
| 12 - Coastal - Canary Coastal Province (EACB) | 39 - Trades - N. Pacific Equatorial Countercurrent Province |
| 13 - Coastal - Guinea Current Coastal Province | 40 - Trades - Pacific Equatorial Divergence Province |
| 14 - Coastal - Guianas Coastal Province | 41 - Trades - W. Pacific Warm Pool Province |
| 15 - Coastal - NW Atlantic Shelves Province | 42 - Trades - Archipelagic Deep Basins Province |
| 16 - Westerlies - Mediterranean Sea, Black Sea Province | 43 - Coastal - Alaska Downwelling Coastal Province |
| 17 - Trades - Caribbean Province | 44 - Coastal - California Upwelling Coastal Province |
| 18 - Westerlies - N. Atlantic Subtropical Gyral Province (East) (STGE) | 45 - Coastal - Central American Coastal Province |
| 19 - Coastal - Brazil Current Coastal Province | 46 - Coastal - Chile-Peru Current Coastal Province |
| 20 - Coastal - SW Atlantic Shelves Province | 47 - Coastal - China Sea Coastal Province |
| 21 - Coastal - Benguela Current Coastal Province | 48 - Coastal - Sunda-Arafura Shelves Province |
| 22 - Trades - Indian Monsoon Gyres Province | 49 - Coastal - East Australian Coastal Province |
| 23 - Trades - Indian S. Subtropical Gyre Province | 50 - Coastal - New Zealand Coastal Province |
| 24 - Coastal - E. Africa Coastal Province | 51 - Westerlies - S. Subtropical Convergence Province |
| 25 - Coastal - Red Sea, Persian Gulf Province | 52 - Westerlies - Subantarctic Province |
| 26 - Coastal - NW Arabian Upwelling Province | 53 - Polar - Antarctic Province |
| 27 - Coastal - E. India Coastal Province | 54 - Polar - Austral Polar Province |

Figure S3 – Carte de l'ensemble des Provinces de Longhurst (Longhurst, 1995).



Provinces biogéochimiques de la zone mésopélagique

- | | |
|---|--|
| 1 - Tropical Province | 8 - Southern Ocean Temperate Province |
| 2 - Subtropical Province | 9 - Cold and Anoxic Temperate Province |
| 3 - Subtropical Gyres Province | 10 - Subpolar Province |
| 4 - North Atlantic subtropical Gyres Province | 11 - Cold and Oxic Polar Province |
| 5 - Mediterranean and salty water Province | 12 - Southern Subtropical Frontal Province |
| 6 - Tropical Anoxic Water Province | 13 - Subpolar Oxic Northern Province |
| 7 - Subpolar and Nutrient Depleted Province | |

Figure S4 – Carte de l'ensemble des provinces biogéochimiques de la zone mésopélagique (Reygondeau et al., 2018).

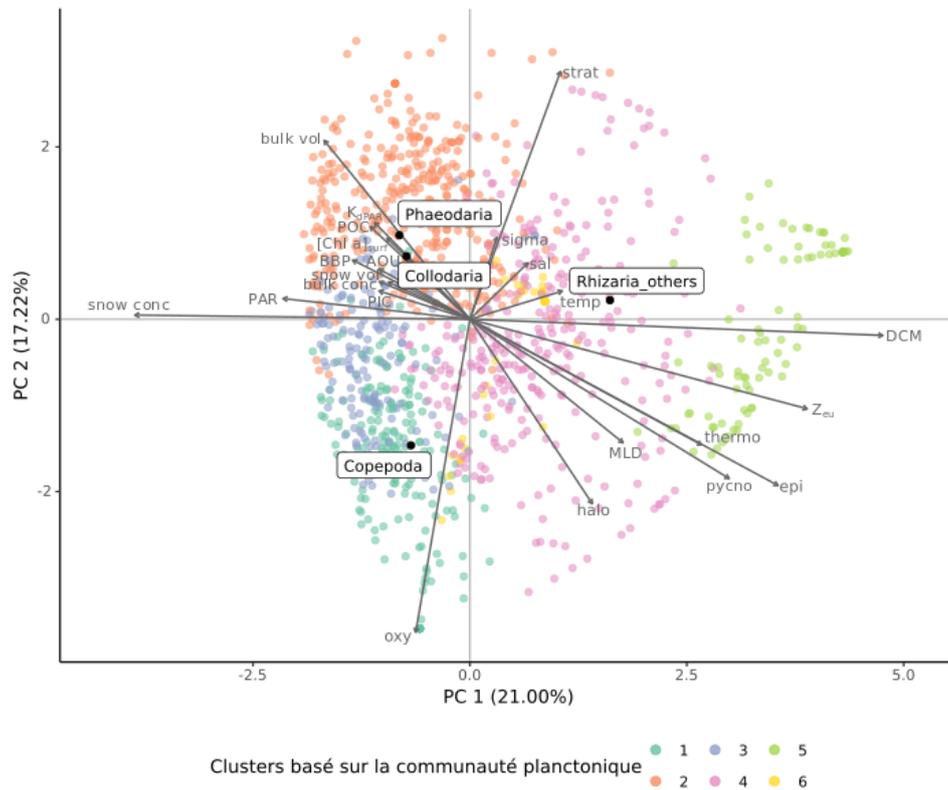


Figure S5 – Projection des stations de la couche mésopélagique inférieure sur les deux premiers axes de l’ACP colorées en fonction des groupes définis par la CAH ; des catégories taxonomiques et des variables environnementales. Toutes les variables environnementales sont projetées, tandis que seules les catégories taxonomiques dont la norme de la projection sur les deux premiers axes est supérieure au rayon du cercle des contributions équivalentes le sont.

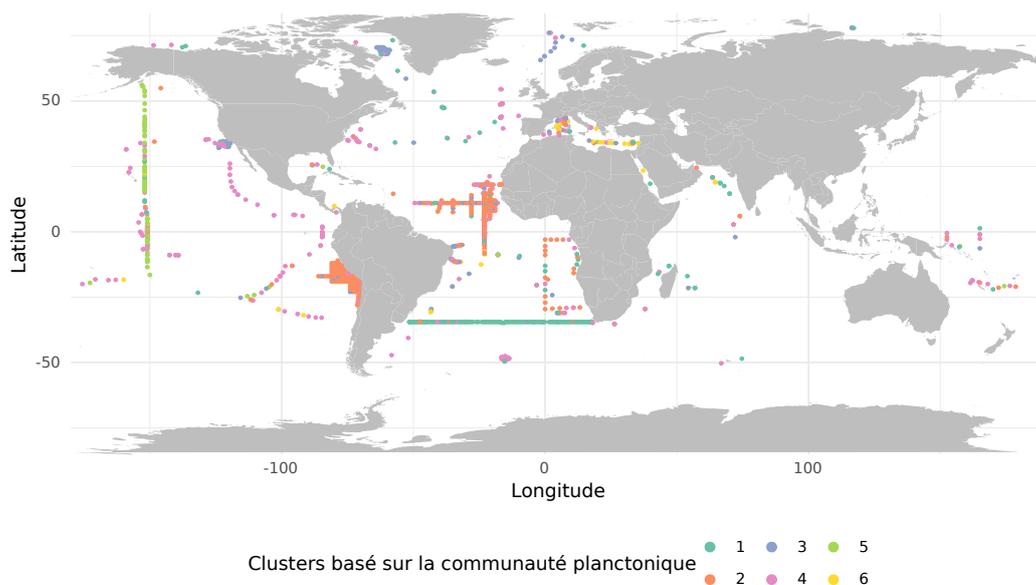


Figure S6 – Carte des stations de la couche mésopélagique inférieure colorées en fonction des groupes définis par la CAH.

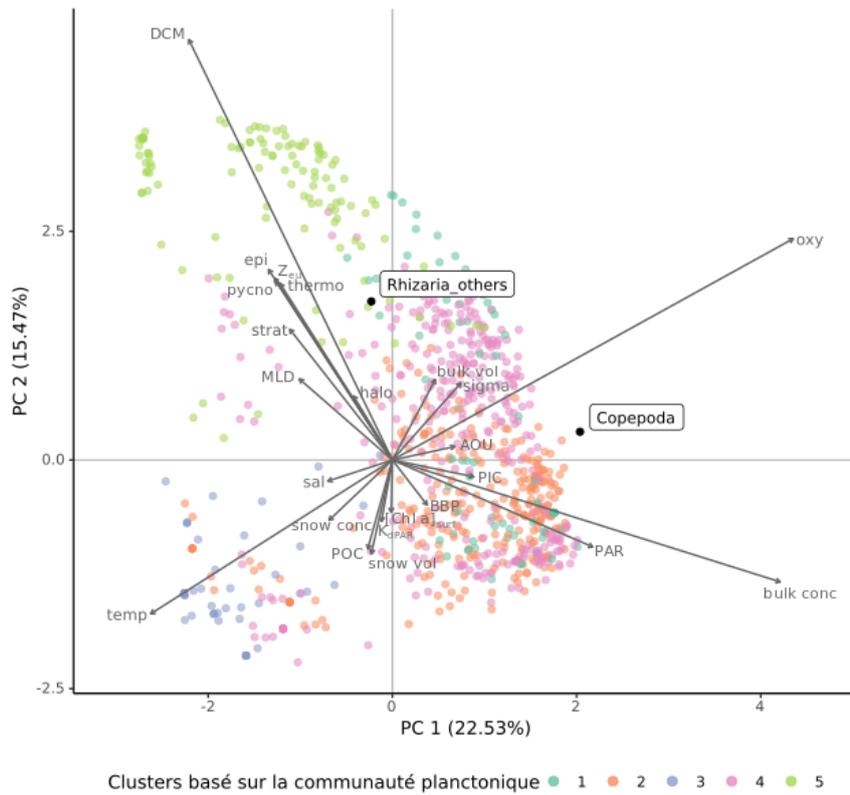


Figure S7 – Projection des stations de la couche bathypélagique sur les deux premiers axes de l’ACP colorées en fonction des groupes définis par la CAH ; des catégories taxonomiques et des variables environnementales. Toutes les variables environnementales sont projetées, tandis que seules les catégories taxonomiques dont la norme de la projection sur les deux premiers axes est supérieure au rayon du cercle des contributions équivalentes le sont.

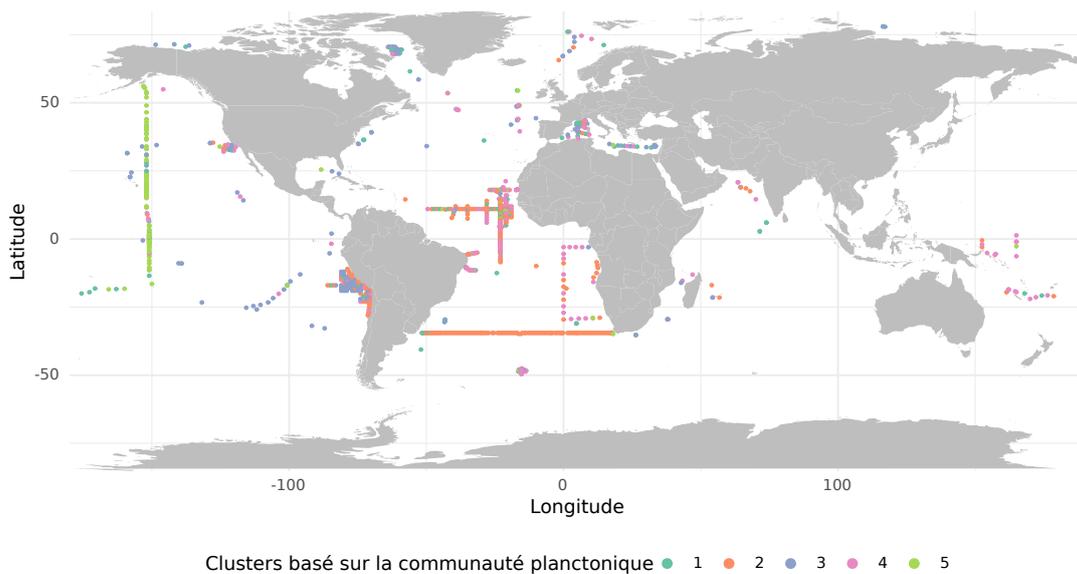


Figure S8 – Carte des stations de la couche bathypélagique colorées en fonction des groupes définis par la CAH.

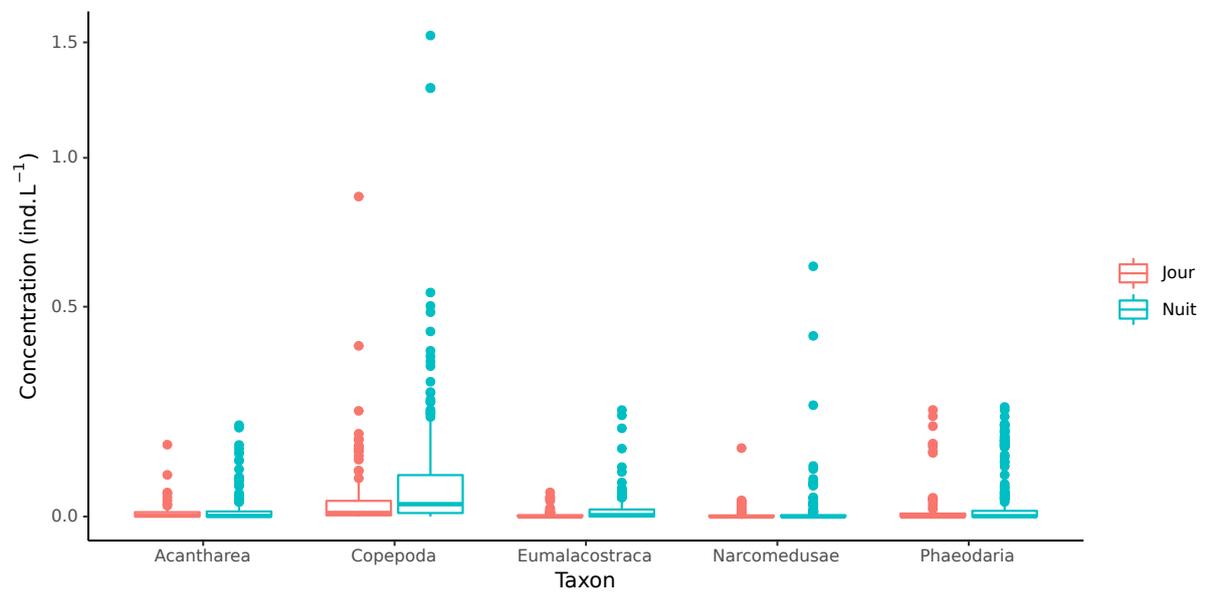


Figure S9 – Concentration moyenne des 5 taxons les plus abondants des campagnes CCE-LTER (au large du courant de Californie) dans la couche épipelagique le jour et la nuit.

Typologie globale de la distribution tridimensionnelle des communautés de plancton vues par imagerie *in situ*

De nombreuses tentatives ont été faites pour partitionner les écosystèmes marins en grands biomes à l'image des écosystèmes terrestres, mais cela reste difficile en raison du manque de connaissances du milieu hauturier. Alors qu'une partition verticale fait aujourd'hui consensus, plusieurs propositions de partitions horizontales, essentiellement basées sur la biogéochimie, ont été avancées. Il reste à déterminer si ces partitions reflètent ou non la distribution des organismes. L'étude de la distribution des communautés de plancton, organismes sensibles aux changements affectant les masses d'eaux dans lesquelles ils sont insérés, est ainsi adaptée pour évaluer la pertinence écologique de ces régionalisations.

En utilisant un outil d'imagerie *in situ* du plancton et de la neige marine couplé à des outils d'enregistrement des conditions environnementales, nous avons étudié la distribution des communautés de plancton en lien avec leur environnement. Ces outils ont permis de générer un grand nombre de données de concentration de plancton sur l'ensemble des bassins océaniques. Ces données ont été analysées avec des méthodes multivariées et d'ordination afin de décrire les communautés de plancton.

Ainsi nous avons pu mettre en évidence des grands types de communautés de plancton pour les différentes couches de l'océan. Ces communautés sont principalement structurées par trois groupes : les trichodesmiums, les copépodes et les rhizaires. La comparaison entre la distribution des communautés de plancton et un ensemble de partitions géographiques des océans montre que les communautés de plancton sont essentiellement contraintes par les conditions environnementales régionales et non par des conditions environnementales locales et immédiates.

Mots clés : communautés de plancton, régionalisation, environnement, imagerie *in situ*

Global typology of the tridimensional distribution of plankton communities seen by *in situ* imagery

There were many attempts at partitioning marine ecosystems in large biomes like terrestrial ecosystems. This remains difficult due to the lack of knowledge of the open ocean. A consensus has been found for a vertical partitioning, and several propositions have been made for a horizontal partitioning, mainly based on biogeochemical signatures. Yet the ability of these partitionings to capture the distribution of organisms is still to explore. The study of the distribution of plankton communities, whose organisms are affected by changes in water masses in which they are included, is well adapted to assess the ecological relevance of these partitionings.

Using an *in situ* imagery tool for plankton and marine snow paired with environment recording sensors, we studied the distribution of plankton communities in relation with their environment. These tools generated a large dataset of planktonic concentrations all around the world. This dataset was analysed with multivariate and ordination methods in order to describe plankton communities.

We were able to distinguish large types of plankton communities in different oceanic layers. These communities are mainly structured by three groups: trichodesmiums, copepods and rhizaria. The comparison between the distribution of these communities and a set of geographical partitionings of oceans show that plankton communities are mainly driven by regional environmental conditions instead of local conditions.

Keywords: plankton communities, ecoregionalisation, environment, *in situ* imagery