



HAL
open science

Approche de biorégionalisation pour alimenter les échelles d'évaluation du programme thématique “Habitats pélagiques”

Marion Vilain, Eric Goberville, Dorothée Vincent, Anne Goffart

► To cite this version:

Marion Vilain, Eric Goberville, Dorothée Vincent, Anne Goffart. Approche de biorégionalisation pour alimenter les échelles d'évaluation du programme thématique “Habitats pélagiques”. BOREA; OFB; Université de Liège. 2022. hal-04003137

HAL Id: hal-04003137

<https://hal.science/hal-04003137v1>

Submitted on 23 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Approche de biorégionalisation pour alimenter les échelles d'évaluation du programme thématique «Habitats pélagiques»

M. Vilain¹, E. Goberville¹, D. Vincent² et A. Goffart³

¹Laboratoire BOREA "Biologie des Organismes et des Ecosystèmes Aquatiques" Sorbonne Université, MNHN, CNRS UMR 8067, IRD 207, UCN, UA.

²Office Français de la Biodiversité (OFB) - Site de Brest – Direction Surveillance, Evaluation, Données (DSUED) - Service Ecosystèmes, Connaissances et Usages des Milieux Marins (ECUMM).

³Laboratoire d'Océanologie, UR FOCUS, Université de Liège.

1. Introduction	2
1.1. Contexte réglementaire et objectifs	2
1.2. Biorégionalisation : concept et approches méthodologiques	3
1.3. Zone d'étude et données utilisées	5
2. Stratégie d'analyse	9
2.1. Traitement des données biologiques	9
2.1.1. Etapes préliminaires	9
2.1.2. Gestion de l'échantillonnage irrégulier	10
2.1.3. Réseaux biogéographiques	12
2.2. Liens avec les paramètres environnementaux	12
3. Résultats préliminaires et discussion	14
4. Conclusions et perspectives	20
Références	22

1. Introduction

1.1. Contexte réglementaire et objectifs

La **Directive-Cadre « Stratégie pour le Milieu Marin » (DCSMM)**¹ vise à maintenir ou restaurer un bon fonctionnement des écosystèmes marins (*i.e.* une diversité biologique et des interactions espèces-habitats conservées, des océans dynamiques et productifs) dans le but d'**atteindre le Bon Etat Ecologique (BEE)** ; au niveau biologique, physique, chimique et sanitaire), tout en permettant des activités humaines dans une perspective d'usage durable du milieu marin. La surveillance du milieu marin est structurée en thématiques de pression ou de composante de l'écosystème, qui correspondent à des **descripteurs du BEE**. Chaque descripteur est évalué selon des **critères** composés d'**éléments constitutifs**².

Les Habitats pélagiques sont une composante du **Descripteur 1 - Biodiversité**, défini comme suit : *“La diversité biologique est conservée. La qualité des habitats et leur nombre, ainsi que la distribution et l'abondance des espèces sont adaptés aux conditions physiographiques, géographiques et climatiques existantes.”* (Directive 2008/56/UE). L'atteinte du BEE au titre de la composante Habitats pélagiques est évaluée par le **critère D1C6**³ qui s'assure que les habitats pélagiques ne subissent pas d'effets néfastes dus à des pressions anthropiques. Les habitats pélagiques y sont considérés dans leur globalité en intégrant leurs caractéristiques biotiques, abiotiques et fonctionnelles. Les éléments constitutifs du critère D1C6 sont les **grands types d'habitats pélagiques**, qui correspondent aux eaux à salinité variable, aux eaux côtières, aux eaux du plateau continental et à la haute mer, s'ils sont présents dans la région ou la sous-région. D'autres types d'habitats supplémentaires peuvent être choisis par les Etats membres conformément aux spécifications de la directive relatives aux choix des espèces et des habitats.

Au 2nd cycle de l'évaluation du BEE, les **Unités Géographiques d'Evaluation (UGE)** des habitats pélagiques considérées au titre du Descripteur 1 correspondaient (i) pour les eaux côtières, aux types de masses d'eaux côtières telles que définies dans le cadre de la Directive Cadre sur l'Eau⁴ ; (ii) au-delà des eaux côtières, aux paysages marins définis par Tew-Kai *et al.* (2020) et mis à jour par Quilfen *et al.* (2021).

L'échelle d'évaluation des Habitats pélagiques doit **“refléter les différences biogéographiques dans la composition en espèces du type d'habitat”**. Les paysages marins sont définis à partir de 12 variables océaniques “essentielles” (EOVs, *Essential Ocean Variables*) telles que la température, la salinité ou encore la courantologie. Les indicateurs de surveillance développés pour évaluer le critère D1C6 s'intéressant aux communautés planctoniques, nous avons choisi d'étudier directement la biogéographie des espèces pour alimenter des échelles pertinentes d'évaluation (UGEE). Notre démarche de travail consiste à **(i) définir des biorégions**, c'est-à-dire des entités géographiques se distinguant par le caractère unique de leurs communautés planctoniques, puis à **(ii)**

¹ directive 2008/56/CE du Parlement européen et du Conseil du 17 juin 2008.

² «éléments constitutifs de critères» : les éléments constitutifs d'un écosystème, particulièrement ses éléments biologiques (espèces, habitats et leurs communautés), ou les aspects des pressions exercées sur l'environnement marin (pressions biologiques et physiques, substances, déchets et énergie), évalués pour chaque critère.

³ d'après la version révisée de la décision 2017/848/UE relative aux normes méthodologiques pour l'atteinte du Bon Etat Ecologique.

⁴ Directive 2000/60/CE - cadre pour une politique communautaire dans le domaine de l'eau.

caractériser l'environnement hydro-climatique associé à chaque biorégion obtenue. Afin d'évaluer la cohérence des échelles spatiales d'évaluation, (iii) une comparaison avec les paysages marins pourra être effectuée.

1.2. Biorégionalisation : concept et approches méthodologiques

Un objectif central en **biogéographie** est de **caractériser des zones géographiques se distinguant par le caractère unique de leurs conditions environnementales et/ou climatiques, de leur géomorphologie, ou encore de la distribution spatiale des organismes vivants** (Kreft & Jetz, 2010). La plupart des classifications proposées sont soit fondées sur la structure physique de l'océan, c'est-à-dire la température, la stratification de la colonne d'eau et la circulation océanique (Emery & Meincke, 1986 ; Cushing, 1989), soit sur la distribution spatiale des organismes marins (Beklemishev, 1961 ; McGowan, 1971). Platt *et al.* (1991) ont, quant à eux, proposé de définir des biorégions sur la base de la concentration de chlorophylle *a* de surface.

Les méthodes permettant de définir des unités biogéographiques et leurs subdivisions varient dans la littérature (Robertson & Cramer, 2014). Certaines études se concentrent exclusivement sur les **modèles d'endémisme**⁵ (Briggs, 1974 ; Briggs & Bowen, 2012), d'autres - principalement pour les analyses menées aux échelles régionale (Valentine, 1966 ; Fox & Beckley, 2005) et globale (Spalding *et al.*, 2007 ; Kulbicki *et al.*, 2013) - considèrent le **caractère distinctif général des biotas**. Le caractère distinctif de la faune comprend des informations sur les patrons géographiques de présence de toutes les composantes des faunes régionales et locales, y compris, mais sans s'y limiter, les espèces endémiques qui représentent souvent une composante mineure d'une faune. Le caractère distinctif peut être utilisé pour établir une hiérarchie de relations entre les faunes locales qui définissent les limites entre les unités biogéographiques (Valentine, 1966), sans s'appuyer sur des "règles" arbitraires fondées sur les niveaux d'endémisme (Briggs & Bowen, 2012) ou sur des facteurs externes tels que la température de surface (Kulbicki *et al.*, 2013), et en évitant l'influence des idées préconçues découlant d'évaluations biogéographiques antérieures. Cette approche de similitude permet d'évaluer explicitement l'influence des espèces endémiques en examinant les schémas spatiaux des espèces endémiques et non endémiques régionales (Robertson & Cramer, 2009).

L'approche dominante pour la biorégionalisation a longtemps été fondée sur des **méthodes de classification hiérarchique** appliquées à la **dissimilarité** des compositions (Holt *et al.*, 2013 ; Kreft & Jetz, 2010 ; Procheş & Ramdhani, 2012), méthodes agglomératives qui s'inscrivent en opposition aux méthodes divisives telles que les K-means (Lasram *et al.*, 2015). Cette recherche de hiérarchie s'appuie sur la notion de **distance écologique** entre les observations, les sites, ou encore les espèces : les données à disposition doivent alors être transformées en matrices de distance afin de mesurer cette dissimilarité. Cette approche classique reste cependant fortement débattue dans la littérature (Mouchet *et al.*, 2008), et de nombreuses études soulèvent que les résultats obtenus peuvent être conditionnés par des critères méthodologiques tels que (i) la sélection de la mesure de distance écologique utilisée pour mesurer la dissimilarité (Cormack (1971)

⁵ Une espèce est dite endémique sur une période temporelle considérée si sa distribution est restreinte à une région géographique donnée, qu'elle soit fondée sur des frontières politiques humaines ou des caractéristiques géographiques naturelles (Anderson, 1994; Townsend Peterson & Watson, 1998).

parlant de “*burgeoning bibliography*” et de “*plethora of definitions of similarity*”) - les mesures les plus communément utilisées étant les indices de Jaccard, de Sorensen ou de Gower, la distance euclidienne, le coefficient de corrélation de Pearson, pour n’en nommer que quelques-unes - ou (ii) le choix de la méthode de regroupement : méthode de Ward, liens simples, liens complets, liens moyens, liens médians ou encore centroïdes (Hattab *et al.*, 2015). Bien que de nombreuses approches aient été proposées afin d’évaluer l’influence de ces critères méthodologiques (e.g. construction de classifications consensus, Mouchet *et al.*, 2008 ; utilisation de métriques d’évaluation de la qualité des classifications obtenues, Mérigot *et al.*, 2010) une question cruciale reste en suspens dès lors qu’une classification est obtenue : à quel niveau la classification doit-elle être coupée afin de déterminer un nombre optimal - et cohérent - d’unités biogéographiques (Kelley *et al.*, 1996 ; Guidi *et al.*, 2008)? Ce choix, qui reste subjectif, influence directement le nombre de régions caractérisées par la classification (e.g. Beaugrand *et al.* (2019) ont examiné les 8 premiers niveaux de coupure de leur partitionnement : ils obtiennent entre 2 et 8 régions distinctes). En écologie, les données sont généralement très variables et peuvent contenir des valeurs extrêmes - voire aberrantes - ce qui augmente la difficulté d’identifier le nombre de clusters (ou groupes) à conserver (Guidi *et al.*, 2008).

Récemment, une **approche fondée sur les réseaux biogéographiques** a été introduite par Vilhena & Antonelli (2015) et a été recommandée pour délimiter des régions biogéographiques (Bloomfield *et al.*, 2017 ; Edler *et al.*, 2016 ; Rojas *et al.*, 2017). Cette approche présente plusieurs avantages par rapport aux approches fondées sur les distances. Un avantage majeur est de **limiter les choix méthodologiques** (sélection de la mesure de distance, choix de la méthode de regroupement et du seuil de coupure) en ayant une approche plus intuitive et standardisée. En outre, là où les approches classiques nécessitent une transformation des données qui a tendance à brouiller l’information de départ (abstraction des données dans des matrices de dissimilarité entre les sites), **l’approche par réseaux biogéographiques conserve l’identité des espèces**. Sachant que le renouvellement des espèces a tendance à augmenter avec la distance géographique par rapport à une source, cette approche permet de déterminer des biorégions par des changements réels dans les affinités taxinomiques, et non pas uniquement sur la base de la distance (Vilhena & Antonelli, 2015). L’algorithme de détection des communautés utilisé dans l’approche par réseaux biogéographiques (*Map Equation*, Encart 1) **assigne chaque espèce à une biorégion spécifique**, permettant ainsi de caractériser quelles espèces connectent les sites entre eux et structurent les biorégions (Leroy *et al.*, 2019). Vilhena & Antonelli (2015) ont montré que l’utilisation de la *Map Equation* donne de meilleurs résultats que les méthodes de regroupement classiques lorsqu’elle est comparée à des ruptures biogéographiques connues et bien acceptées. Toutes ces raisons ont conditionné notre choix d’utiliser l’approche par réseaux biogéographiques pour caractériser des biorégions pélagiques.

Encart 1 : Map Equation

La *Map Equation* est un algorithme de regroupement hiérarchique (Rosvall & Bergstrom, 2008) qui détecte les patrons de distribution des communautés dans un réseau (Figure 1). L'algorithme choisit aléatoirement une cellule de grille, ou "site", puis il choisit au hasard une espèce présente dans cette cellule, examine l'aire de répartition géographique de cette espèce et sélectionne au hasard une cellule de grille dans son aire de répartition géographique (Vilhena & Antonelli, 2015). Ce processus est appliqué de manière itérative et exhaustive jusqu'à ce que tous les nœuds du réseau soient affiliés à un cluster, de façon à maximiser la connectivité intra-groupe (*i.e.* ressemblance forte au sein du groupe) et à minimiser la connectivité inter-groupe (*i.e.* différence élevée entre les groupes).

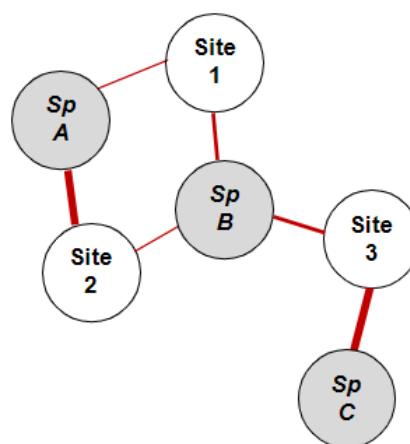


Figure 1 : Schéma d'un réseau bipartite Site - Espèce. Chaque cercle représente un nœud. Un lien rouge entre deux nœuds reflète la présence d'une espèce dans un site, son épaisseur étant proportionnelle à l'abondance de l'espèce à ce site.

1.3. Zone d'étude et données utilisées

Nous avons choisi de développer la méthode de biorégionalisation d'après la distribution des **communautés zooplanctoniques puis, une fois stabilisée, de l'appliquer aux communautés phytoplanctoniques.**

Le zooplancton joue un rôle essentiel dans le **fonctionnement des écosystèmes marins** et dans les **cycles biogéochimiques**. Les organismes zooplanctoniques occupent une **place centrale au sein du réseau trophique pélagique**. Ils comprennent à la fois les brouteurs du phytoplancton et les prédateurs des plus petites tailles du zooplancton (*i.e.* microzooplancton ; Kleppel, 1993; Stoecker & Egloff, 1987), et constituent la proie principale des organismes de niveaux trophiques supérieurs (poissons, Ware & Thompson 2005 ; oiseaux marins, Amélineau *et al.*, 2016 ; méduses et cténaires, Purcell *et al.*, 1994 ; Richardson *et al.*, 2009). Des modifications dans les assemblages d'espèces zooplanctoniques reflètent souvent une réponse intégrée de l'écosystème aux forçages hydro-climatiques (Beaugrand *et al.*, 2002). Nos résultats pourront être comparés à des études biogéographiques réalisées sur les assemblages zooplanctoniques issus des mêmes données pour évaluer la robustesse de la méthode (*e.g.* Beaugrand *et al.*, 2002 ; Beaugrand 2004).

Les **données biologiques** utilisées dans cette étude sont les **données mensuelles d'abondances zooplanctoniques** (organismes compris entre 270 µm et 2 mm) issues du *Continuous Plankton Recorder (CPR) Survey* (Encart 2) et couvrant les secteurs Atlantique, Manche et mer du Nord (Figure 2). Une approche de biorégionalisation nécessitant une **couverture spatiale à large échelle** pour pouvoir détecter des patrons de répartition, nous avons privilégié l'utilisation des données du *CPR Survey* au-delà des sous-régions marines DCMM (*i.e.* unités marines de rapportage). Limiter l'utilisation des données du *CPR Survey* aux sous-régions marines (SRM) réduirait de plus d'un facteur 10 le nombre de données disponibles pour réaliser ce travail.

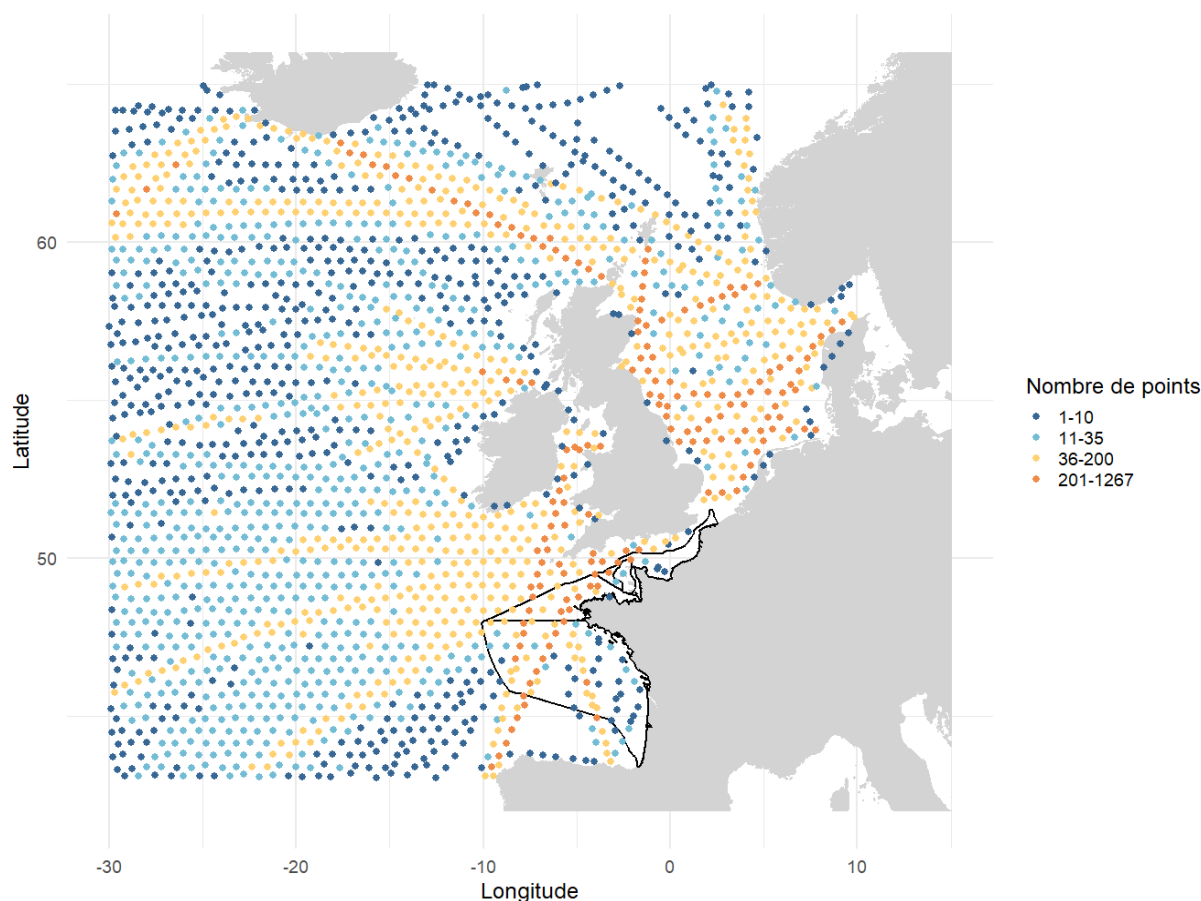


Figure 2 : Couverture spatiale à large échelle et effort d'échantillonnage par le *CPR Survey* au cours de la période 1966-2013. L'effort d'échantillonnage est résumé sur cette période par construction d'une grille hexagonale (taille fixée : 50 cellules en longitude) : les observations contenues dans chaque cellule hexagonale sont comptées et leur localisation est assimilée au centre de masse de la cellule. Le code couleur représente l'intensité de l'échantillonnage (catégories logarithmiques). Les sous-régions marines DCSMM golfe de Gascogne, mers celtiques et Manche - mer du Nord sont délimitées par un trait noir.

Les **variables hydro-climatiques de surface** (température de surface, pression au niveau de la mer) **et atmosphériques** (température de l'air, pression, composantes zonale et méridionale du vent et son intensité) que nous avons utilisées (Tableau I) proviennent des réanalyses du NCEP/NCAR (*National Centers for Environmental Prediction/National Center for Atmospheric Research*) de la NOAA (*National Oceanic and Atmospheric Administration*) qui reposent sur un système d'analyse/prédiction assimilant des informations **mensuelles** couvrant notre période d'étude (*i.e.* 1966-2013). Les **données bathymétriques** (version 2021 ; Tableau I) sont issues du GEBCO (*General Bathymetric Chart of the Oceans*). Des variables additionnelles, telles que la stratification de la colonne d'eau (*Mixed-Layer Depth* ; MLD) ou la salinité seront - à court terme - intégrées afin de caractériser plus finement les biorégions et d'établir, par exemple, un lien critique entre les biorégions obtenues et certains grands types d'habitats (*e.g.* habitats à salinité variable) du critère D1C6.

Tableau I : Variables environnementales mobilisées dans cette étude, couverture et résolution spatiales, temporalité et fréquence de rejeu.

Variable	Unité	Couverture spatiale	Résolution spatiale	Résolution temporelle	Fréquence temporelle
Température de surface (SST)	°C	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Pression au niveau de la mer (SLP)	hPa	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Température de l'air (AIR_TEMP)	°C	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Pression (PRESSURE)	hPa	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Vent zonal (U-WIND)	m.s ⁻¹	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Vent méridional (V-WIND)	m.s ⁻¹	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Intensité du Vent (WIND), calculée d'après les composantes U-WIND et V-WIND.	m.s ⁻¹	90N - 90S 0E - 357,5E	2,5°	1948-2022	mensuelle
Bathymétrie (BATHY)	m	89° 59' 52.5"N - 89° 59' 52.5"S 179° 59' 52.5"W - 179° 59' 52.5"E	15"	version 2021	statique

Encart 2 : Continuous Plankton Recorder Survey

Le CPR Survey, opéré par la *Marine Biological Association* (Plymouth, UK), est un dispositif de surveillance permettant la collecte régulière d'échantillons planctoniques par le biais du *Continuous Plankton Recorder* (CPR ; Figures 3.A, 3.B et 4), instrument d'échantillonnage conçu pour être tracté à une profondeur d'environ 5-10 m par un navire marchand - ou un autre type de navire d'opportunité - au cours de sa navigation (Figure 3.B).

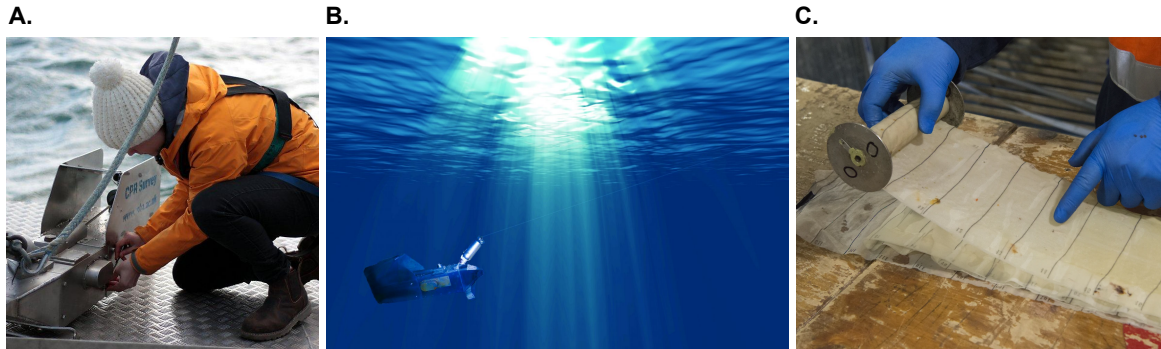


Figure 3 : (A) Scientifique préparant le CPR ©Marine Biological Association ; (B) CPR tracté par un ferry ©Brittany Ferries 2022 ; (C) Soie du CPR avec des échantillons de plancton ©Rowan Butler.

L'eau (Figure 4 - flèche bleue) passe à travers le CPR où le plancton est filtré sur une bande de soie (maille de 270 μm , Figure 4 - flèche verte) qui se déroule lentement, pendant qu'une seconde bande de soie la recouvre (Figure 4 - flèche rouge) afin de piéger les organismes planctoniques. Ces bandes de soie s'enroulent dans un bac de stockage contenant du formol. De retour au laboratoire, la soie est retirée du mécanisme (Figure 3.C) et divisée en échantillons représentant chacun 10 milles nautiques (19 km) de remorquage. Les échantillons sont analysés pour obtenir une estimation de la biomasse phytoplanctonique (PCI = « *Phytoplankton Color Index* »), une identification (jusqu'à l'espèce) des taxons phyto- et zooplanctoniques au microscope et à la loupe binoculaire respectivement, ainsi qu'une estimation (semi-quantitative ; Batten *et al.*, 2003) de leurs abondances. Les méthodes d'échantillonnage et d'analyse du plancton marin par CPR sont inchangées depuis 1948, ce qui permet d'analyser et de comparer, temporellement et spatialement, les dynamiques des communautés planctoniques sur près de 70 ans, sans biais potentiellement imputable à la combinaison d'information issues de différents engins de collecte et/ou de méthodes d'échantillonnage.

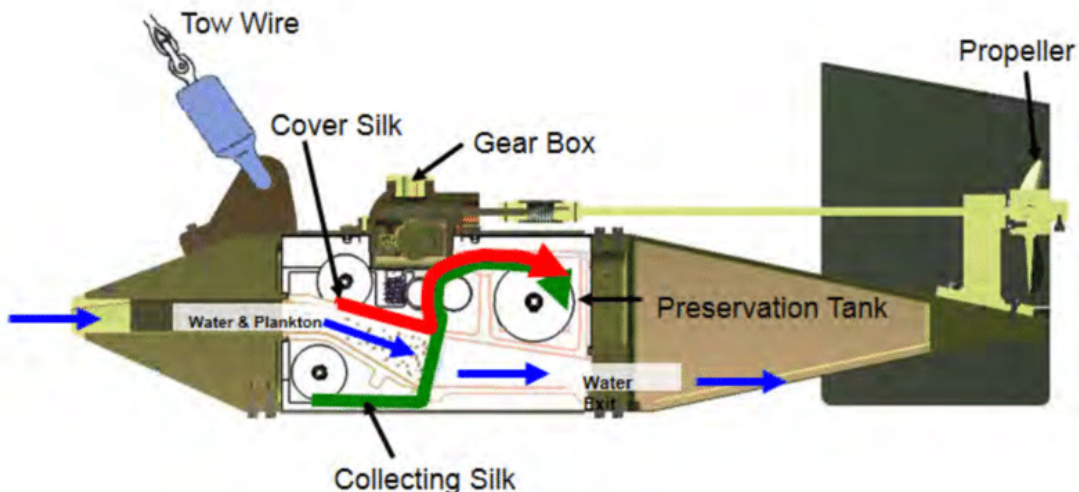


Figure 4 : Fonctionnement interne du *Continuous Plankton Recorder*, issu de Takahashi *et al.*, 2021.

Tow Wire : câble de remorquage ; Cover Silk : soie de recouvrement ; Collecting Silk : soie de collecte ; Gear Box : réducteur de vitesse ; Preservation Tank : bac de stockage ; Propeller : hélice ; Water & plankton : eau et plancton ; Water Exit : sortie d'eau.

2. Stratégie d'analyse

2.1. Traitement des données biologiques

2.1.1. Etapes préliminaires

La première étape du traitement de données consiste à **filtrer taxinomiquement à l'espèce** (Figure 5.A.1) afin de déterminer finement les communautés zooplanctoniques. L'historique du *CPR Survey* révèle un ajout d'espèces identifiées au cours du temps, qui atteste non seulement d'une amélioration de la capacité d'identification des organismes (*e.g.* une plus grande précision taxinomique grâce au développement des méthodes en laboratoire), mais également de l'arrivée de nouvelles espèces depuis une zone non échantillonnée vers une zone couverte par le *CPR Survey* (*e.g.* introduction d'espèces). Ici, nous avons privilégié une vision homogène des communautés d'espèces au cours de la période temporelle 1966-2013, en conservant uniquement les **espèces identifiées pour la première fois avant 1966** (*NB* : nous avons prévu de mener une analyse complémentaire ne considérant pas ce filtre afin de quantifier l'effet de cette sélection *a priori*). Après ces différentes étapes de filtrage, notre jeu de données inclut 16 espèces, dont la majorité appartiennent à la sous-classe des Copepoda (Tableau II).

Tableau II : Liste des espèces retenues pour l'analyse. Les groupes d'appartenance taxinomique et les plus anciennes dates de dénombrement sont précisés.

Espèce	Groupe taxinomique	Compté depuis
<i>Temora longicornis</i>	copépode	1946
<i>Centropages typicus</i>	copépode	1946
<i>Centropages hamatus</i>	copépode	1946
<i>Isias clavipes</i>	copépode	1946
<i>Acartia danae</i>	copépode	1964
<i>Ctenocalanus vanus</i>	copépode	1959
<i>Euterpina acutifrons</i>	copépode	1964
<i>Macrosetella gracilis</i>	copépode	1965
<i>Mecynocera clausi</i>	copépode	1948
<i>Parapontella brevicornis</i>	copépode	1947
<i>Temora stylifera</i>	copépode	1963
<i>Tortanus discaudatus</i>	copépode	1961
<i>Acartia longiremis</i>	copépode	1964
<i>Diaixis hibernica</i>	copépode	1962
<i>Diaixis pygmaea</i>	copépode	1965
<i>Zoothamnium pelagicum</i>	protozoaire cilié	1964

Afin de mettre en place une approche de biorégionalisation, il convient d'agrèger temporellement les données afin de réduire le bruit inhérent à un effet de variabilité spatiale de l'effort d'échantillonnage, et ainsi maximiser le signal lié à la dynamique spatiale des communautés. L'évolution temporelle des biorégions est rapportée pour une **période glissante de 6 ans**, en cohérence avec les cycles de surveillance/évaluation de la DCSMM (Figure 5.A.1).

2.1.2. Gestion de l'échantillonnage irrégulier

La couverture spatiale du *CPR Survey* étant dépendante des routes empruntées par les navires d'opportunité, une différence dans la fréquence d'échantillonnage est à noter (Figure 2) : sur la période 1966-2013, 25% des zones échantillonnées présentent moins de 8 observations (1^{er} quartile) et 75% présentent moins de 63 observations (3^{ème} quartile). Les routes où l'on retrouve le plus d'observations (de 201 à 1267 observations) sont principalement situées en mer du Nord. Pour prévenir un biais potentiel lié à l'échantillonnage, nous avons utilisé une **approche de quadtree** afin de mettre en place une **grille adaptative** : la taille des cellules géographiques est adaptée pour intégrer le nombre requis d'observations (ici $n = 10$ observations minimum par cellule, seuil de sélection arbitraire) : plus les 10 observations sont spatialement proches (forte intensité d'échantillonnage), plus la cellule géographique est "petite". Les zones fortement échantillonnées (e.g. en mer du Nord) ont alors des cellules de plus petite taille, à l'inverse des zones moins échantillonnées (e.g. au large de l'Atlantique Nord) pour lesquelles l'intensité d'échantillonnage est plus éparse.

Un quadtree est une structure de données hiérarchique permettant de partitionner un espace bidimensionnel en le décomposant de manière récursive (i.e. répétée un nombre indéfini de fois, par l'application de la même règle). Une grille régulière initiale d'une taille de cellule donnée est construite et les cellules initiales sont subdivisées hiérarchiquement, de manière récursive, en quatre quadrants de taille égale. Un seuil (fixé par l'utilisateur) dirige la subdivision des cellules en imposant un nombre de points minimum par cellule. Nous avons utilisé le package R AQuadtree, en suivant les paramétrages par défaut, à savoir l'utilisation d'une projection azimutale équivalente de Lambert (qui conserve les surfaces). En accord avec la couverture spatiale du *CPR Survey*, nous avons déterminé une grille initiale sur la base de cellules de 100 km de côté, un nombre de 10 subdivisions à réaliser et un **seuil de 10 observations minimum par cellule**. Chaque cellule de la grille possède un code permettant de l'identifier, et ainsi d'y associer les informations biologiques à traiter. Au sein de chaque cellule géographique, nous avons **ré-échantillonné les données d'abondance** avec le même seuil utilisé pour le processus de quadtree qui est, par définition, le nombre minimum d'observations par cellule. Les **moyennes d'abondance de chaque espèce**, après ré-échantillonnage, sont ensuite calculées, permettant *in fine* l'obtention d'une **matrice biologique de la forme [Cellules géographiques x Espèces x Abondances]** (Figure 5.B.1).

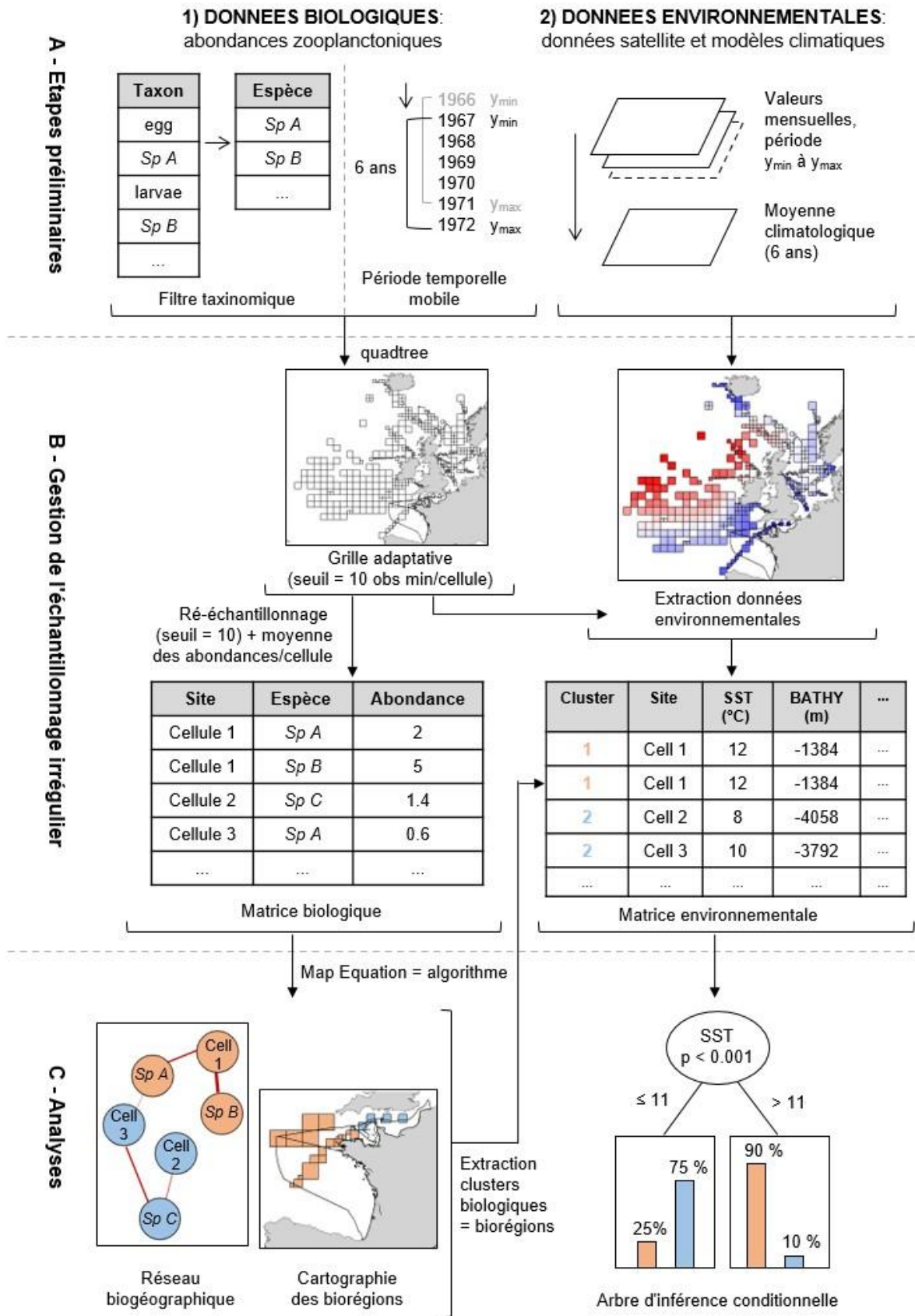


Figure 5 : Schéma décrivant les différentes étapes de travail - de la préparation des données au traitement numérique - et le flux de données dans notre étude.

2.1.3. Réseaux biogéographiques

Nous avons ensuite appliqué la **méthode des réseaux biogéographiques** proposée par Leroy *et al.* (2019 ; package R biogeonetworks ; Figure 5.C.1) sur cette matrice biologique. Un **réseau bipartite (cellule géographique - espèce)** est généré depuis cet ensemble de données. Le réseau est composé (i) de sites (cellules géographiques) et (ii) d'espèces, formant un ensemble de nœuds connectés les uns aux autres par des liens : lorsqu'une espèce est présente dans une cellule géographique, un lien est établi entre l'espèce et la cellule géographique en question, ce lien étant pondéré par la valeur d'abondance associée. Par définition, les liens site-site et espèce-espèce ne sont pas autorisés dans ce type d'analyse, l'objectif étant d'analyser la distribution spatiale des différentes espèces.

Un **algorithme de regroupement hiérarchique (*Map Equation*, Encart 1)** est appliqué sur ce réseau pour détecter les régions. L'algorithme rassemble, en clusters, les nœuds ayant une forte connectivité intra-groupe mais une faible connectivité inter-groupe, ce qui correspond à la définition des régions biogéographiques, *i.e.* des régions d'assemblages distincts de taxons endémiques. Chaque cellule géographique est affiliée à un cluster qui représente une région structurée d'après les communautés zooplanctoniques. L'algorithme est lancé 1000 fois afin de caractériser les clusters optimaux, c'est-à-dire d'assurer une robustesse de l'analyse en limitant toute perturbation induite par une petite modification dans les données ou dans les paramètres choisis. Enfin, et pour visualisation, une couleur est attribuée à chaque cluster et **les résultats sont cartographiés sur la grille adaptative**. A noter que seules les cellules géographiques contenues dans les sous-régions marines DCSMM sont représentées (Figure 6 - haut). En complément, **le réseau est représenté graphiquement sous le logiciel Gephi 0.9.2** (Bastian *et al.*, 2009), avec l'algorithme forceatlas2 qui regroupe les nœuds étroitement interconnectés (groupes de sites et d'espèces de la même région biogéographique) et sépare les groupes de nœuds qui ne sont pas interconnectés (régions biogéographiques distinctes). Cette représentation graphique permet d'analyser et d'explorer le réseau, en particulier pour **déterminer les espèces structurantes de chaque cluster** (*i.e.* biorégion). Pour **résumer l'évolution des clusters** entre plusieurs périodes temporelles, un **diagramme alluvial** est généré (Figure 6 - bas). Cette représentation graphique, fondée sur les flux, permet de mettre en évidence les changements temporels de la structure du réseau : les blocs représentent des groupes (*i.e.* biorégions) de nœuds et sont reliés par des champs de flux connectant les composantes de chaque bloc. La hauteur d'un bloc est proportionnelle à la taille du cluster.

2.2. Liens avec les paramètres environnementaux

Après caractérisation des biorégions, nous avons analysé leurs relations avec les variables hydro-climatiques, **l'objectif étant de caractériser la gamme de valeur de chaque paramètre environnemental structurant les biorégions**.

Dans un premier temps, les données environnementales sont mises en forme : contrôle et ajustement du système de projection des données, calcul de climatologies moyennes sur les périodes mobiles de 6 ans (Figure 5.A.2).

Dans un second temps, pour chaque donnée biologique contenue dans une cellule de la grille adaptative précédemment créée, les données environnementales associées sont extraites puis moyennées par cellule. Nous obtenons alors une **matrice environnementale de la forme [Cellules géographiques x Variables hydro-climatiques]**, regroupant les

variables suivantes : Température de l'air et de surface de l'océan, Bathymétrie, Pression atmosphérique, Pression au niveau de la mer, Vent zonal, Vent méridional et Vent total (Figure 5.B.2 et Tableau I).

Enfin, les clusters biologiques définis par la méthode des réseaux biogéographiques sont associés aux cellules de la matrice environnementale. L'utilisation d'une **approche d'arbres d'inférence conditionnelle** nous permet de quantifier la relation entre les régions biogéographiques préalablement identifiées et les variables environnementales (Figure 5.C.2). Pour résumer, l'algorithme de calcul des arbres d'inférence conditionnelle fonctionne de la façon suivante :

(1) Tester l'hypothèse nulle globale d'indépendance entre l'une des variables d'entrée (*i.e.* les paramètres hydro-climatiques) et la réponse (*i.e.* les régions biogéographiques). S'arrêter si cette hypothèse ne peut être rejetée ou sélectionner la variable d'entrée ayant la plus forte association avec la réponse si l'hypothèse peut être rejetée. Cette association est mesurée par une *p-value* correspondant à un test de l'hypothèse nulle partielle d'une seule variable d'entrée et de la réponse.

(2) Implémenter une division binaire sur la variable d'entrée sélectionnée.

(3) Appliquer récursivement les étapes (1) et (2).

L'arbre d'inférence conditionnelle obtenu permet de mettre en évidence les variables dominantes (avec leur gamme de valeurs) expliquant la structuration des biorégions, ainsi que leur pourcentage d'explication desdites biorégions. Nous avons utilisé la fonction `ctree` du package R `partykit`.

3. Résultats préliminaires et discussion

Nos résultats mettent en évidence 2 biorégions présentes depuis 1966 (données présentées ici à partir de 1976 ; Figure 6) :

- Un groupement représenté en bleu, situé en Manche - mer du Nord et en mers celtiques, structuré majoritairement par *Temora longicornis* et *Centropages hamatus*, deux espèces de copépodes néritiques à affinité boréale.
- Un groupement représenté en jaune, situé en mers celtiques et dans la partie nord du golfe de Gascogne, structuré majoritairement par le copépode néritique eurytherme plus cosmopolite *Centropages typicus*.

Ces biorégions sont constantes jusqu'à la fin des années 80/début des années 90, puis elles se rejoignent pour ne former qu'une seule biorégion (Figure 6, en vert). Ce regroupement des deux biorégions en une seule coïncide temporellement avec le changement de régime global et synchrone qui s'est opéré à l'échelle du globe à la fin des années 80 (Reid *et al.*, 2015) et a provoqué des effets en cascade dans toutes les composantes du système Terre (*i.e.* cryosphère, hydrosphère et biosphère). A partir de cette période, les communautés ne forment plus qu'une unique entité biogéographique et ne se différencient donc plus. Cette biorégion se maintient en Manche - mer du Nord, mers celtiques et dans la partie nord du golfe de Gascogne jusqu'en 2013 (données présentées ici jusqu'en 2005).

A partir de la deuxième moitié des années 90, l'introduction de nouvelles routes au *CPR Survey* permet un échantillonnage dans la partie sud du golfe de Gascogne. Si ces données supplémentaires étendent la biorégion précédemment décrite vers le large, elles permettent également la caractérisation d'une nouvelle biorégion structurée majoritairement par *Temora stylifera*, copépode néritique adapté aux eaux chaudes (Figure 6, en rose).

Nos résultats montrent que **la capacité à caractériser des biorégions est fonction de la couverture spatiale d'échantillonnage**. L'apparition d'une biorégion peut révéler un changement dans les communautés zooplanctoniques (*e.g.* regroupement des biorégions bleue et jaune), mais également mettre en évidence des communautés "nouvelles" qui résultent d'une modification de la stratégie d'échantillonnage (*e.g.* biorégion rose). La mise en place de nouvelles zones d'échantillonnage peut alors permettre d'identifier une structuration des communautés à des échelles où les particularités écologiques locales peuvent induire des communautés spécifiques.

Ces premiers résultats soulignent le **caractère dynamique des communautés zooplanctoniques** (caractérisé ici sur des périodes glissantes de 6 ans) et confirment qu'**il est essentiel d'intégrer cette dynamique aux unités géographiques d'évaluation dans un objectif de suivi de l'évolution temporelle des communautés**. Cela questionne l'utilisation d'unités géographiques statiques (comme les paysages marins) pour mener une évaluation, surtout si le but est d'apprécier des changements dans des communautés intrinsèquement dynamiques. Il est à noter que les biorégions marines ont des frontières mouvantes qui évoluent au cours du temps, en particulier au niveau des zones de transition entre biorégions. Cette variabilité n'indique pas nécessairement un changement dans la structure des biorégions et, par extension, dans les communautés zooplanctoniques.

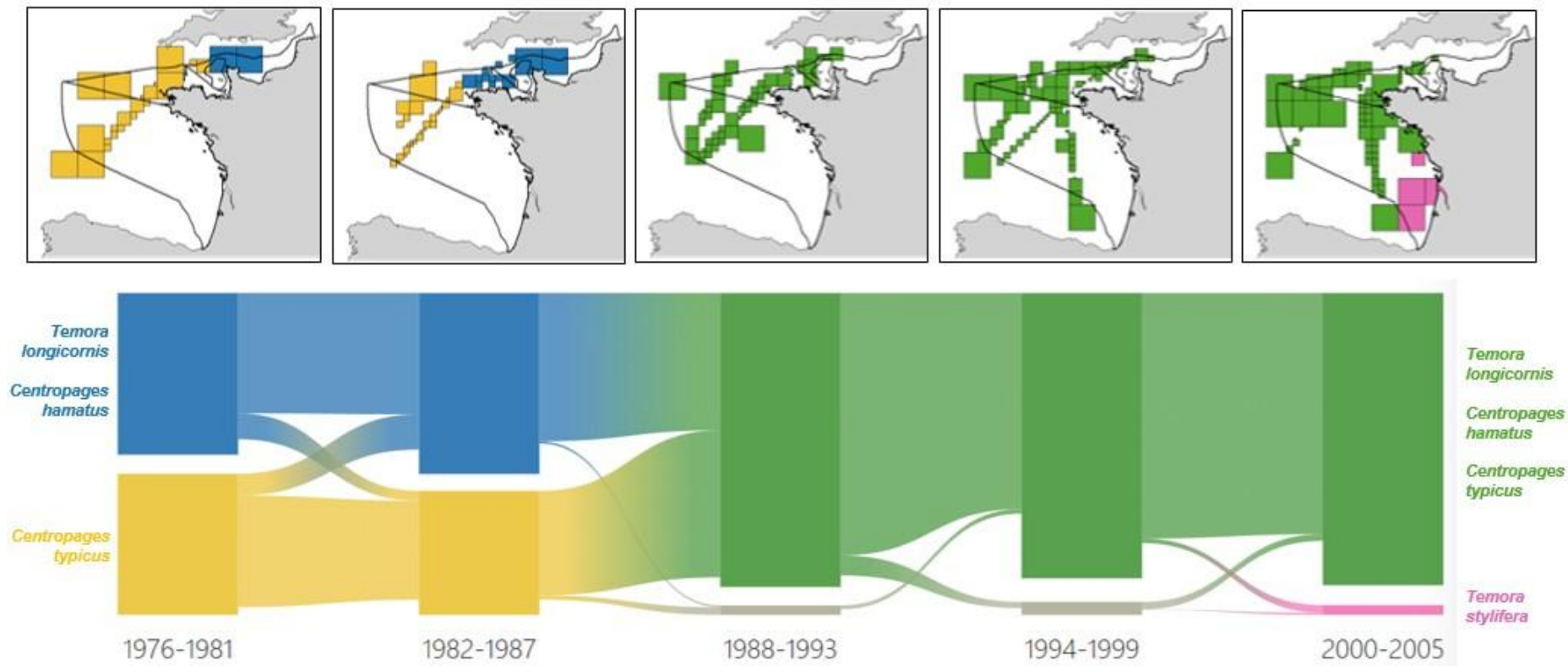


Figure 6 : Représentation spatiale des biorégions obtenues pour 5 périodes de 6 ans consécutives au cours la période 1966-2013 (panels du haut) et diagramme alluvial du flux associé (figure du bas). Les couleurs des biorégions sont associées au diagramme du flux et les espèces dominantes au sein de chaque biorégion sont indiquées. A noter que, si seulement 5 périodes sont représentées ici (pour une question de clarté), l'analyse porte sur un total de 43 périodes (de 6 ans) glissantes sur la période 1966-2013

Une approche en cours (calcul d'arbres d'inférence conditionnelle), nous permet de **détecter les variables hydro-climatiques dominantes** qui discriminent les biorégions et de **quantifier leur gamme de variations**. Les arbres d'inférence conditionnelle ont été construits sur les biorégions à large échelle (au-delà des SRM) et traduisent un contexte environnemental plus large.

Sur la période 1976-1981 (Figure 7), la biorégion située en Manche - mer du Nord et mers celtiques (en bleu, Figure 6), est essentiellement caractérisée par une température de surface inférieure à 11°C et un vent méridional inférieur à 1,5 m.s⁻¹ (prise en compte à la fois de l'intensité et de la direction du vent). La biorégion située en mers celtiques et dans la partie nord du golfe de Gascogne (en jaune, Figure 6) est, quant à elle, principalement caractérisée par une température de surface supérieure à 11°C et une pression atmosphérique supérieure à 1008,8 hPa.

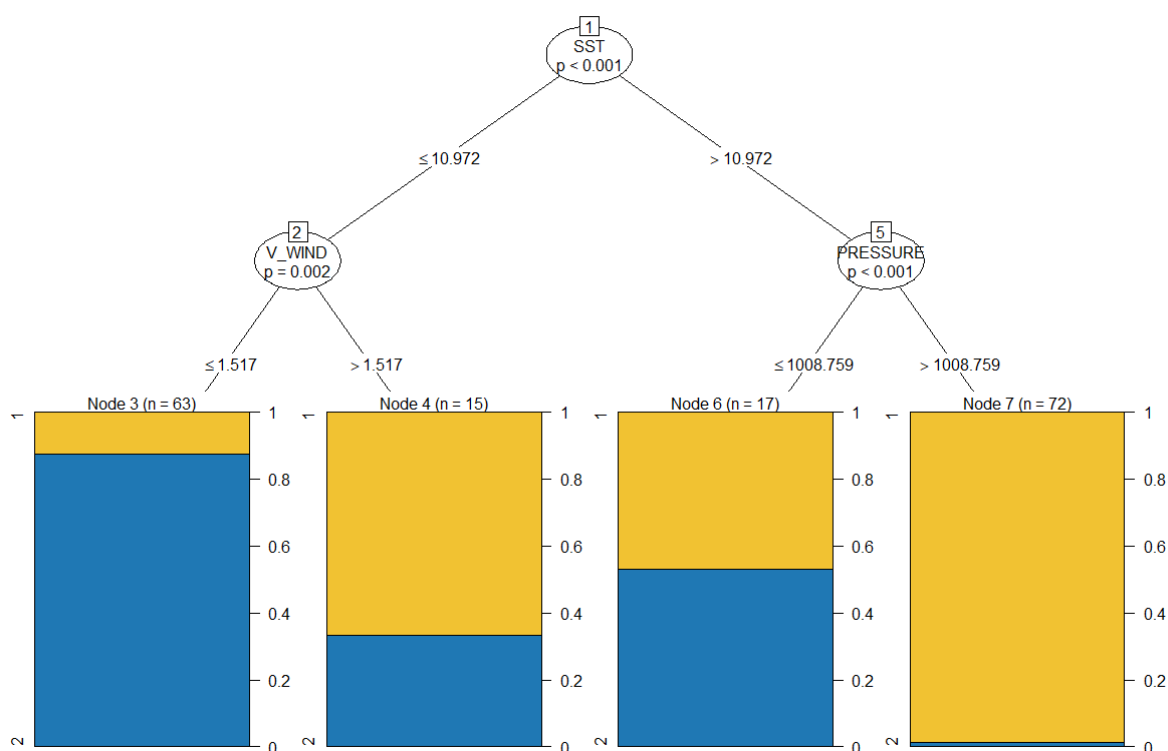


Figure 7 : Arbre d'inférence conditionnelle calculé pour la période 1976-1981. A chaque variable structurante (= nœud) sont associés un numéro de nœud et une valeur de significativité (*p-value*). Les valeurs seuil déclenchant le partitionnement binaire sont indiquées sur les branches. Chaque branche de l'arbre mène, dans la partie inférieure, à un graphique qui représente le pourcentage de la biorégion expliqué par la branche. Le nombre (n) de cellules prises en compte dans le partitionnement est indiqué au-dessus du graphique. Les couleurs correspondent à celles des biorégions (cf. Figure 6).

Sur la période 1982-1987 (Figure 8), période précédant le regroupement des deux biorégions, notre analyse met en évidence que les principales variables structurantes sont la température de surface (toujours avec une première distinction des deux biorégions autour d'une température de 11°C) et l'intensité totale du vent. Dans le contexte du changement de régime global de la fin des années 80, Beaugrand (2004) et Beaugrand *et al.* (2019) ont décrit un changement abrupt des communautés pélagiques en mer du Nord qui s'explique par une modification des processus hydro-météorologiques à grande échelle (augmentation

de la température de surface, changement dans l'intensité et la direction des vents). Cette modification pourrait expliquer le regroupement de nos deux biorégions - fortement structurées par la température de surface et l'intensité du vent - au cours de la période 1988-1993.

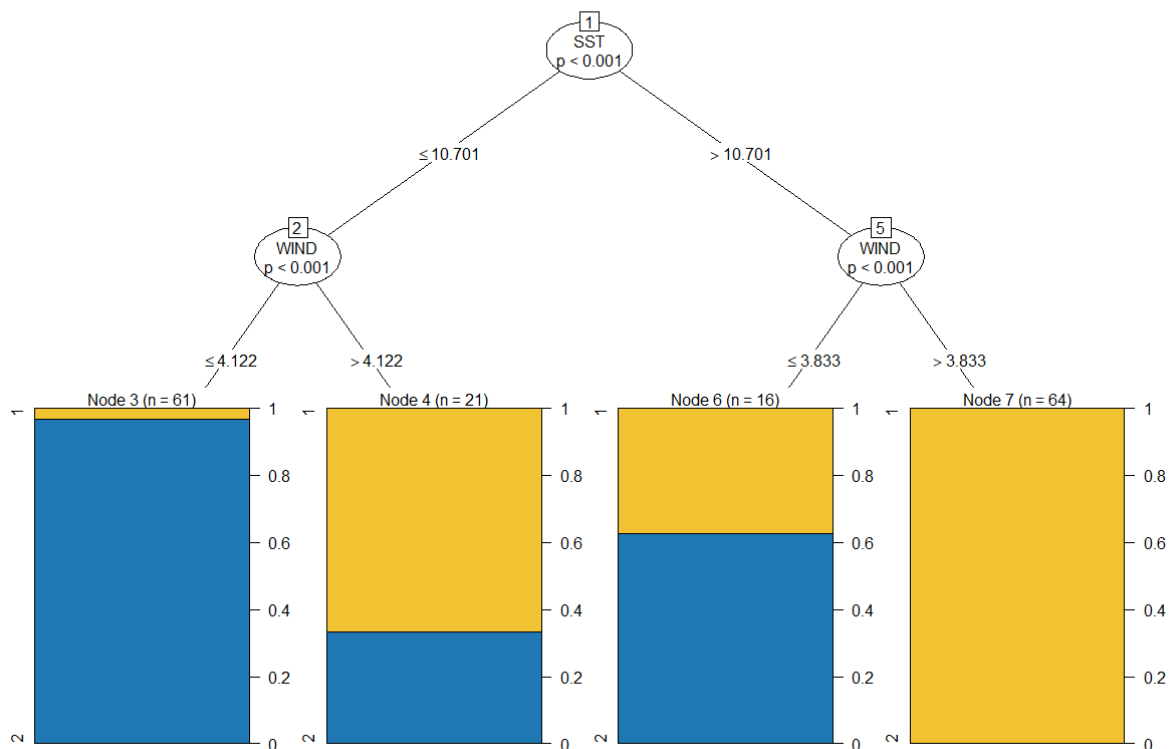


Figure 8 : Arbre d'inférence conditionnelle calculé pour la période 1982-1987. A chaque variable structurante (= nœud) sont associés un numéro de nœud et une valeur de significativité (p -value). Les valeurs seuil déclenchant le partitionnement binaire sont indiquées sur les branches. Chaque branche de l'arbre mène, dans la partie inférieure, à un graphique qui représente le pourcentage de la biorégion expliqué par la branche. Le nombre (n) de cellules prises en compte dans le partitionnement est indiqué au-dessus du graphique. Les couleurs correspondent à celles des biorégions (cf. Figure 6).

Sur la période 2000-2005 (Figure 9), nous retrouvons une structuration essentiellement caractérisée par l'intensité totale du vent et la température de surface. Cette fois-ci, le poids de la variable climatique "vent" est plus important et des valeurs inférieures à $5,2 \text{ m.s}^{-1}$ distinguent 86% des cellules géographiques totales. Quasiment 100% de ces cellules appartiennent à la biorégion située dans la partie sud du golfe de Gascogne et structurée par *Temora stylifera* (en rose, Figure 6). Cette biorégion se distingue donc fortement de la biorégion dominante (en vert, Figure 6) par la variable vent. La proximité des deux biorégions au sud du golfe de Gascogne rend complexe l'interprétation de la discrimination par la température (cf. contexte environnemental des biorégions, Figure 10). Une prochaine étape de ce travail sera de **limiter le bruit dans la délimitation des biorégions**, en utilisant par exemple un algorithme de détection des plus proches voisins, et de vérifier si la discrimination des biorégions au sud du golfe de Gascogne en est statistiquement modifiée.

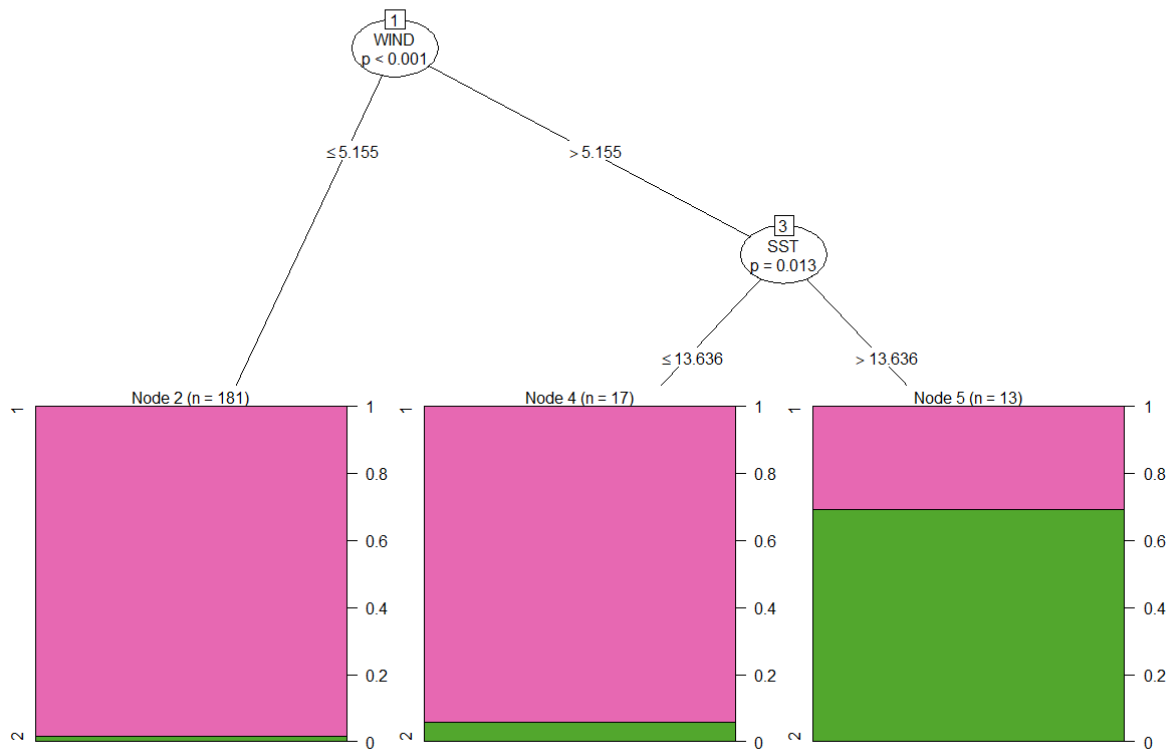


Figure 9 : Arbre d'inférence conditionnelle calculé pour la période 2000-2005. A chaque variable structurante (= nœud) sont associés un numéro de nœud et une valeur de significativité (*p-value*). Les valeurs seuil déclenchant le partitionnement binaire sont indiquées sur les branches. Chaque branche de l'arbre mène, dans la partie inférieure, à un graphique qui représente le pourcentage de la biorégion expliqué par la branche. Le nombre (n) de cellules prises en compte dans le partitionnement est indiqué au-dessus du graphique. Les couleurs correspondent à celles des biorégions (cf. Figure 6).

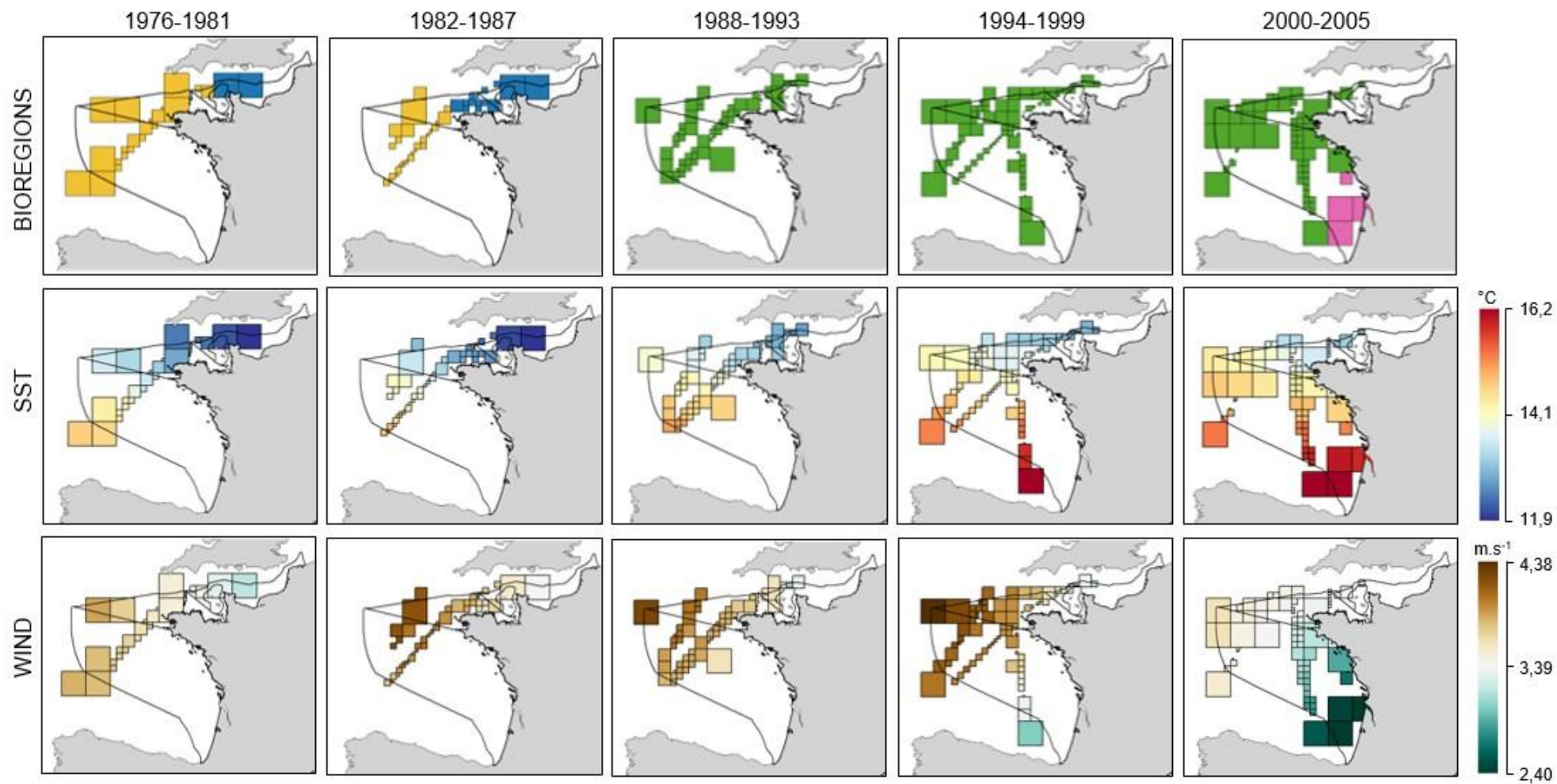


Figure 10 : Représentation spatiale des biorégions et des 2 variables environnementales/climatiques les plus structurantes qui ressortent de notre analyse : température de surface (SST, en °C) et vent (WIND, en m.s⁻¹). A noter que les gammes de valeurs concernent seulement les cellules géographiques contenues dans les SRM, ce qui explique les différences avec les valeurs seuils des arbres d'inférence conditionnelle, construits en considérant les biorégions à large échelle.

4. Conclusions et perspectives

Les premiers résultats de notre étude démontrent la pertinence de l'approche par biorégionalisation pour définir des échelles spatiales dynamiques fondées sur les communautés planctoniques (zooplancton dans cette note) et pouvant être associées aux variables environnementales les structurant.

Suite à cette étude, nous avons identifié des perspectives de travail à court terme qui permettront d'affiner la caractérisation des biorégions en (1) renforçant et stabilisant la méthode, et en (2) diversifiant les variables environnementales et climatiques utilisées.

Nous poursuivrons ce travail en axant nos priorités de recherche sur :

(1) un développement méthodologique. Les pistes prioritaires identifiées étant :

- de mettre en place un algorithme de détection des plus proches voisins pour limiter le bruit. Le résultat attendu est une meilleure délimitation des biorégions et des zones de transitions associées, et aura pour conséquence de stabiliser notre méthode de caractérisation des biorégions ;
- de mettre en place des analyses de sensibilité afin de quantifier l'effet du choix (i) du seuil conditionnant la construction de la grille adaptative (10 observations minimum/cellule dans les résultats présentés dans cette note), et (ii) de la période glissante (ici 6 ans), ainsi que (iii) l'influence du filtre appliqué sur les espèces (sélection *a priori* et prise en compte de l'introduction de nouvelles espèces - par exemple des espèces invasives - en validant par la bibliographie) ;
- de poursuivre et d'affiner l'approche par arbres d'inférence conditionnelle, notamment en effectuant un ré-échantillonnage sur les arbres afin de rendre le résultat encore plus robuste et représentatif de la biorégion ;
- de compléter les données *in situ* (observations biologiques) par des résultats issus de modèles d'habitat. Le résultat attendu est une meilleure délimitation des biorégions et permettra d'augmenter la couverture spatiale de nos biorégions, notamment dans les eaux côtières ;
- d'analyser la coïncidence des biorégions avec les paysages marins, sous réserve de disposer de paysages dynamiques - calculés à minima sur la même temporalité que les biorégions - et de leurs variables structurantes (gamme de variations) ;
- de réfléchir à une stratégie pour adapter les méthodes développées à la SRM Méditerranée occidentale, où les jeux de données de terrain présentent une couverture spatiale et/ou une couverture temporelle réduite(s).

(2) une meilleure prise en compte de la complexité des habitats pélagiques, notamment par une intégration de variables additionnelles. Ce travail sur les données nous permettra :

- d'augmenter la période temporelle étudiée, en intégrant plus particulièrement les périodes d'évaluation DCSMM ;
- d'identifier les variables qui contribuent le plus fortement à la dynamique spatio-temporelle des habitats pélagiques, tout en tentant de discriminer les échelles d'influence (locale, régionale, globale) auxquelles s'exercent les différentes sources de forçages environnementaux et climatiques ;

- d'ajouter de nouvelles espèces à notre analyse - en particulier les organismes zooplanctoniques supérieurs à 2 mm - de façon à mieux intégrer la complexité des communautés présentes ;
- de tester la méthode sur le compartiment phytoplanctonique, afin d'identifier d'éventuelles convergences/divergences avec les biorégions identifiées à partir des communautés zooplanctoniques (en intégrant le spectre de taille complet, cf. perspective précédente).

Les besoins inhérents à l'approche de biorégionalisation nécessitent d'être anticipés pour les travaux à plus long terme. Pour le prochain cycle d'évaluation, les données CPR du chantier collecte devront être élargies au-delà des sous-régions marines DCSMM, suivant les recommandations européennes d'intégrer des dispositifs de suivi conjoints à large échelle. Nos résultats montrent que l'approche de biorégionalisation doit se fonder sur des données ayant une spatialité la plus large possible, au risque d'obtenir une estimation biaisée, induite par le fait de travailler sur un jeu de données réduit. Cela nécessite d'établir un bilan de l'existant et de renforcer les collaborations européennes, à l'image de ce que nous avons initié dans le cadre du projet ABIOMMED. D'un point de vue stratégique, il sera nécessaire de maintenir et d'augmenter la couverture spatiale des dispositifs de surveillance dont les données de terrain sont mobilisables pour la caractérisation des biorégions. A ce titre, la route CPR Bilbao - Plymouth (arrêtée depuis 2017) est la seule route du *CPR Survey* permettant d'échantillonner le sud du golfe de Gascogne ; or l'existence de cette route nous a permis d'identifier une biorégion caractéristique de ces masses d'eaux. Remettre en service cette route apparaît essentiel pour caractériser la dynamique des communautés planctoniques dans cette zone. La mise en place de routes CPR en Méditerranée occidentale constituerait un véritable atout, permettant de développer également des unités géographiques fondées sur la biologie dans cette sous-région marine. Ces perspectives doivent être discutées en termes de surveillance et d'évaluation. Elles permettraient d'aboutir à des unités géographiques comparables à la régionalisation consensus proposée par le projet MERMEX (Ayata *et al.*, 2018) qui intègre des indicateurs globaux des communautés phytoplanctoniques (e.g. dynamique de la concentration de chlorophylle *a* en surface) et la distribution des organismes épipélagiques (Reygondeau *et al.*, 2014).

Références

- Amélineau, F. *et al.* (2016) 'Microplastic pollution in the Greenland Sea: Background levels and selective contamination of planktivorous diving seabirds', *Environmental Pollution*, 219, pp. 1131–1139. doi:10.1016/j.envpol.2016.09.017.
- Anderson, S. (1994) 'Area and Endemism', *The Quarterly Review of Biology*, 69(4), pp. 451–471. doi:10.1086/418743.
- Ayata, S.-D. *et al.* (2018) 'Regionalisation of the Mediterranean basin, a MERMEX synthesis', *Progress in Oceanography*, 163, pp. 7–20. doi:10.1016/j.pocean.2017.09.016.
- Bastian, M., Heymann, S. and Jacomy, M. (2009) 'Gephi: An Open Source Software for Exploring and Manipulating Networks', p. 2.
- Batten, S.D. *et al.* (2003) 'CPR sampling: the technical background, materials and methods, consistency and comparability', *Progress in Oceanography*, 58(2–4), pp. 193–215. doi:10.1016/j.pocean.2003.08.004.
- Beaugrand, G. *et al.* (2002) 'Reorganization of North Atlantic Marine Copepod Biodiversity and Climate', *Science*, 296(5573), pp. 1692–1694. doi:10.1126/science.1071329.
- Beaugrand, G. (2004) 'The North Sea regime shift: Evidence, causes, mechanisms and consequences', *Progress in Oceanography*, 60(2–4), pp. 245–262. doi:10.1016/j.pocean.2004.02.018.
- Beaugrand, G. *et al.* (2019) 'Prediction of unprecedented biological shifts in the global ocean', *Nature Climate Change*, 9(3), pp. 237–243. doi:10.1038/s41558-019-0420-1.
- Beklemishev, C.W. (1961) 'On the spatial structure of plankton communities in dependence of oceanic circulation. Boundaries of ranges of oceanic plankton animals in the North Pacific', *Okeanologia*, 5, 1059–1072.
- Bloomfield, N.J., Knerr, N. and Encinas-Viso, F. (2017) 'A comparison of network and clustering methods to detect biogeographical regions', *Ecography*, 41(1), pp. 1–10. doi:10.1111/ecog.02596.
- Briggs, J.C. (1974) 'Operation of zoogeographic barriers', *Systematic Zoology*, p. 9.
- Briggs, J.C. and Bowen, B.W. (2012) 'A realignment of marine biogeographic provinces with particular reference to fish distributions', *Journal of Biogeography*, p. 19.
- Cormack, R.M. (1971) 'A Review of Classification', *Journal of the Royal Statistical Society. Series A (General)*, 134(3), p. 321. doi:10.2307/2344237.
- Cushing, D.H. (1989) 'A difference in structure between ecosystems in strongly stratified waters and in those that are only weakly stratified', *Journal of Plankton Research*, 11(1), pp. 1–13. doi:10.1093/plankt/11.1.1.
- Edler, D. *et al.* (2016) 'Infomap Bioregions: Interactive Mapping of Biogeographical Regions from Species Distributions', *Systematic Biology*, p. syw087. doi:10.1093/sysbio/syw087.
- Emery, W.J. & Meincke, J. (1986) 'Global water masses: summary and review', *Oceanologica Acta*, 9, pp. 383–391.
- Fox, N.J. and Beckley, L.E. (2005) 'Priority areas for conservation of Western Australian coastal fishes: A comparison of hotspot, biogeographical and complementarity approaches', *Biological Conservation*, p. 12.

- Guidi, L. *et al.* (2008) 'A new procedure to optimize the selection of groups in a classification tree: Applications for ecological data', *Ecological Modelling*, 220(4), pp. 451–461. doi:10.1016/j.ecolmodel.2008.11.006.
- Hattab, T. *et al.* (2015) 'A biogeographical regionalization of coastal Mediterranean fishes', *Journal of Biogeography*, 42(7), pp. 1336–1348. doi:10.1111/jbi.12505.
- Holt, B.G. *et al.* (2013) 'An Update of Wallace's Zoogeographic Regions of the World', *Science*, 339(6115), pp. 74–78. doi:10.1126/science.1228282.
- Kelley, L.A., Gardner, S.P. and Sutcliffe, M.J. (1996) 'An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies', *Protein Engineering, Design and Selection*, 9(11), pp. 1063–1065. doi:10.1093/protein/9.11.1063.
- Kleppel, G. (1993) 'On the diets of calanoid copepods', *Marine Ecology Progress Series*, 99, pp. 183–195. doi:10.3354/meps099183.
- Kreft, H. and Jetz, W. (2010) 'A framework for delineating biogeographical regions based on species distributions: Global quantitative biogeographical regionalizations', *Journal of Biogeography*, 37(11), pp. 2029–2053. doi:10.1111/j.1365-2699.2010.02375.x.
- Kulbicki, M. *et al.* (2013) 'Global Biogeography of Reef Fishes: A Hierarchical Quantitative Delineation of Regions', *PLOS ONE*, 8(12), p. 11.
- Lasram, F.B.R. *et al.* (2015) 'Modeling of Beta Diversity in Tunisian Waters: Predictions Using Generalized Dissimilarity Modeling and Bioregionalisation Using Fuzzy Clustering', *PLOS ONE*. Edited by C.N. Bianchi, 10(7), p. e0131728. doi:10.1371/journal.pone.0131728.
- Leroy, B. *et al.* (2019) 'Global biogeographical regions of freshwater fish species', *Journal of Biogeography*, 46(11), pp. 2407–2419. doi:10.1111/jbi.13674.
- McGowan, J.A. (1971) 'Ocean biogeography of the Pacific. The micropaleontology of the oceans' (ed. by F.A. Riedl), pp. 3–74. *Cambridge University Press*, Cambridge.
- Mérigot, B., Durbec, J.-P. and Gaertner, J.-C. (2010) 'On goodness-of-fit measure for dendrogram-based analyses', *Ecology*, 91(6), pp. 1850–1859. doi:10.1890/09-1387.1.
- Mouchet, M. *et al.* (2008) 'Towards a consensus for calculating dendrogram-based functional diversity indices', *Oikos*, 117(5), pp. 794–800. doi:10.1111/j.0030-1299.2008.16594.x.
- Platt, T., Caverhill, C. and Sathyendranath, S. (1991) 'Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic', *Journal of Geophysical Research*, 96(C8), p. 15147. doi:10.1029/91JC01118.
- Procheş, Ş. and Ramdhani, S. (2012) 'The World's Zoogeographical Regions Confirmed by Cross-Taxon Analyses', *BioScience*, 62(3), pp. 260–270. doi:10.1525/bio.2012.62.3.7.
- Purcell, J.E., White, J.R. and Roman, M.R. (1994) 'Predation by gelatinous zooplankton and resource limitation as potential controls of *Acartia tonsa* copepod populations in Chesapeake Bay', *Limnology and Oceanography*, 39(2), pp. 263–278. doi:10.4319/lo.1994.39.2.0263.
- Quilfen, V. *et al.* (2021) 'Partitioning ocean dynamical patches using modelled essential and derived ocean variables and data mining in the northwestern Mediterranean Sea', 9th EuroGOOS International conference, Shom; Ifremer; EuroGOOS AISBL, May 2021, Brest, France. pp.438-446. hal-03338105v2
- Reid, P.C. *et al.* (2015) 'Global impacts of the 1980s regime shift', *Global Change Biology*, 22(2), pp. 682–703. doi:10.1111/gcb.13106.

- Richardson, A.J. *et al.* (2009) 'The jellyfish joyride: causes, consequences and management responses to a more gelatinous future', *Trends in Ecology & Evolution*, 24(6), pp. 312–322. doi:10.1016/j.tree.2009.01.010.
- Robertson, D. and Cramer, K. (2009) 'Shore fishes and biogeographic subdivisions of the Tropical Eastern Pacific', *Marine Ecology Progress Series*, 380, pp. 1–17. doi:10.3354/meps07925.
- Robertson, D.R. and Cramer, K.L. (2014) 'Defining and Dividing the Greater Caribbean: Insights from the Biogeography of Shorefishes', *PLOS ONE*, 9(7), p. 16.
- Rojas, A. *et al.* (2017) 'Global biogeography of Albian ammonoids: A network-based approach', *Geology*, 45(7), pp. 659–662. doi:10.1130/G38944.1.
- Rosvall, M. and Bergstrom, C.T. (2008) 'Maps of random walks on complex networks reveal community structure', *Proceedings of the National Academy of Sciences*, 105(4), pp. 1118–1123. doi:10.1073/pnas.0706851105.
- Spalding, M.D. *et al.*, (2007) 'Marine Ecoregions of the World : A Bioregionalization of Coastal and Shelf Areas', *BioScience*, 57(7), pp. 573–583.
- Stoecker, D.K. and Egloff, D.A. (1987) 'Predation by *Acartia tonsa* Dana on planktonic ciliates and rotifers', *Journal of Experimental Marine Biology and Ecology*, 110(1), pp. 53–68. doi:10.1016/0022-0981(87)90066-9.
- Takahashi, E.K.T. and Hosie, G.W. (2021) 'Report on the Status and Trends of Southern Ocean Zooplankton based on the SCAR Southern Ocean Continuous Plankton Recorder (SO-CPR) Survey', p. 90.
- Tew-Kai, E. *et al.* (2020) 'Dynamic Coastal-Shelf Seascapes to Support Marine Policies Using Operational Coastal Oceanography: The French Example', *Journal of Marine Science and Engineering*, 8(8), p. 585. doi:10.3390/jmse8080585.
- Townsend Peterson, A. and Watson, D.M. (1998) 'Problems with areal definitions of endemism: the effects of spatial scaling', *Diversity and Distributions*, 4(4), pp. 189–194. doi:10.1046/j.1472-4642.1998.00021.x.
- Valentine, J.W. (1966) 'Numerical analysis of marine molluscan ranges on the extratropical northeastern Pacific shelf', *Limnology and Oceanography*, 11: 198–211.
- Vilhena, D.A. and Antonelli, A. (2015) 'A network approach for identifying and delimiting biogeographical regions', *Nature Communications*, 6(1), p. 6848. doi:10.1038/ncomms7848.
- Ware, D.M. and Thomson, R.E. (2005) 'Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific', *Science*, 308(5726), pp. 1280–1284. doi:10.1126/science.1109049.