



HAL
open science

Données génétiques et indicateurs de biodiversité microbienne pélagique

Benjamin Alric, Christian Jeanthon, Loïs Maignien, Urania Christaki,
Nathalie Simon

► **To cite this version:**

Benjamin Alric, Christian Jeanthon, Loïs Maignien, Urania Christaki, Nathalie Simon. Données génétiques et indicateurs de biodiversité microbienne pélagique. Station biologique de Roscoff, CNRS-UMPC-INSU, Place Georges Tessier, 29680 Roscoff. 2023, pp.78. hal-04495506

HAL Id: hal-04495506

<https://ofb.hal.science/hal-04495506v1>

Submitted on 8 Mar 2024

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.



Convention OFB/Sorbonne Université

Action scientifique

Données génétiques et indicateurs de biodiversité microbienne pélagique dans le cadre de la DCSMM

Rapport final

Benjamin Alric¹, Christian Jeanthon¹, Loïs Maignien², Urania Christaki³,
Nathalie Simon¹

¹Sorbonne Université, CNRS, UMR 7144 ADMM, Station Biologique de Roscoff, Roscoff, France

²Université de Bretagne Occidentale, IFREMER, CNRS, UMR 6197 LM2E, Plouzané, France

³Université Littoral Côte d'Opale, CNRS, Université Lille, UMR 8187 LOG, Wimereux, France

Mai 2023

Citation : Alric B., C. Jeanthon, L. Maignien, U. Christaki & N. Simon. Données génétiques et indicateurs de biodiversité microbienne pélagique dans le cadre de la DCSMM, Rapport Sorbonne Université (SBR) pour l'OFB, mai 2023, 66 pp.

SOMMAIRE

1.	Introduction.....	1
2.	Les indicateurs BEE du descripteur « Habitats pélagiques »	3
2.1.	Indicateur <i>PH1</i>	3
2.2.	Indicateur <i>PH2</i>	4
2.3.	Indicateur <i>PH3</i>	6
3.	Positionnement du projet INDIGENE dans les travaux d'évaluation de l'état écologique du descripteur « Habitats pélagiques » et le programme de surveillance associé	7
4.	Caractéristiques des stations d'échantillonnage.....	9
5.	Action 1 : Traitement, préparation et mise en forme des tableaux de données disponibles ...	10
5.1.	Données de biodiversité obtenues par microscopie	11
5.2.	Données de biodiversité génétique (obtenues par metabarcoding).....	11
5.3.	Paramètres environnementaux.....	13
5.4.	Disponibilité des données	15
6.	Action 2 : Mise en évidence des patrons temporels de diversité et lien avec les facteurs environnementaux.....	15
6.1.	Variabilité temporelle des paramètres environnementaux	16
6.2.	Structures temporelles des communautés microbiennes planctoniques.....	19
6.2.1.	<i>Méthodologie</i>	19
6.2.2.	<i>Exemple pour la station SOMLIT-Astan à Roscoff</i>	21
6.2.3.	<i>Exemple pour les stations B8 et H42 à la Marina du Château dans le port de Brest</i>	25
6.2.4.	<i>Conclusions sur les enveloppes de variation des communautés microbiennes.....</i>	28
7.	Action 3 : Test des indicateurs de composition sur des données génétiques, liens avec les pressions.....	28
7.1.	Méthodologie sur le calcul des indicateurs de changements <i>PH1</i> et <i>PH3</i>.....	28
7.1.1.	<i>Indicateur <i>PH1</i></i>	28
7.1.2.	<i>Indicateur <i>PH3</i></i>	29
7.1.3.	<i>Fréquence d'échantillonnage</i>	32
7.1.4.	<i>Liens indicateur <i>PH3</i> et variables environnementales</i>	33
7.2.	Résultats de l'indicateur <i>PH1</i>	36
7.3.	Résultats sur l'indicateur <i>PH3</i>	42
7.3.1.	<i>Dynamique temporelle de la diversité alpha (indices de Hulburt et de Menhinick).....</i>	42
7.3.2.	<i>Dynamique temporelle de la diversité bêta (LCBD).....</i>	44
7.3.3.	<i>Indicateurs <i>PH3</i> estimé à partir des données de microscopie et de metabarcoding</i>	46

7.3.4.	<i>Evaluation de l'indicateur PH3 selon les classes de qualités, comparaison entre les données microscopiques et de metabarcoding</i>	47
7.3.5.	<i>Evaluation de la sensibilité de l'indicateur PH3 en fonction du nombre d'échantillon considéré</i> 52	
7.3.6.	<i>Lien de l'indicateur PH3 avec l'environnement</i>	54
8.	Action 4 : Protocoles, pipelines et conclusion pertinence des données génétiques	60
9.	Références bibliographiques	63

LISTE DES FIGURES

- Figure 1.** Localisation géographique des quatre stations ateliers (points rouges) dédiées au projet INDIGENE. Pour la station à la Marina du Château dans le port de Brest, le gradient de contamination par les métaux lourds s'établit du ponton le plus interne au ponton le plus externe du port. 9
- Figure 2.** Périodes considérées, fréquences d'échantillonnage et type de données disponibles pour les stations étudiées dans le cadre de ce projet. Pour la Marina du Château, deux sites (B8 et H42) ont été étudiés. 11
- Figure 3.** Variations mensuelles des paramètres environnementaux à la station SOMLIT-Astan de 2009 à 2016. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, COP : Carbone organique particulaire, $\delta^{13}C$: variation relative du rapport $^{13}C/^{12}C$, $\delta^{15}N$: variation relative du rapport $^{15}N/^{14}N$, MES : Matières en suspension, NH_4^+ : Ammonium, NO_2^- : Nitrites, NO_3^- : Nitrates, NOP : Azote organique particulaire, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates. 16
- Figure 4.** Variations mensuelles des paramètres environnementaux aux stations B8 et H42 de 2019 à 2020. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, NO_2^- : Nitrites, NO_3^- : Nitrates, Pheo : Phéophytine, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates. 17
- Figure 5.** Variations annuelles des paramètres environnementaux à la station SOMLIT-Astan de 2009 à 2016. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, COP : Carbone organique particulaire, $\delta^{13}C$: variation relative du rapport $^{13}C/^{12}C$, $\delta^{15}N$: variation relative du rapport $^{15}N/^{14}N$, MES : Matières en suspension, NH_4^+ : Ammonium, NO_2^- : Nitrites, NO_3^- : Nitrates, NOP : Azote organique particulaire, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates. 18
- Figure 6.** Schéma résumant les différentes étapes d'analyse en lien avec l'action 2 du projet. 20
- Figure 7.** Similarités des communautés phytoplanctoniques (analyse RDA) dans les échantillons mensuels de 2009 à 2016 à la station SOMLIT-Astan pour les données de **microscopie**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 22
- Figure 8.** Similarités des communautés de protistes (analyse RDA) dans les échantillons mensuels de 2009 à 2016 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 22
- Figure 9.** Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2012 à 2016 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 18S, fraction de taille 0.2-3 μm)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 23
- Figure 10.** Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2009 à 2011 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 16S, fraction de taille 0.2-3 μm)**. a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 24

- Figure 11.** Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2009 à 2012 à la station SOMLIT-Astan pour les données de **metabarcoding** (ARNr 16S, fraction de taille >3 μm). a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 24
- Figure 12.** Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **microscopie**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier..... 25
- Figure 13.** Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding** (ARNr 18S, fraction de taille >3 μm). a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier... 26
- Figure 14.** Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding** (ARNr 18S, fraction de taille 0.2-3 μm). a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier... 27
- Figure 15.** Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding** (ARNr 16S, fraction de taille 0.2-3 μm). a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier. 27
- Figure 16.** Proposition de seuils pour l'évaluation de l'indicateur PH3 (d'après Duflos et al. 2018). ... 30
- Figure 17.** Représentation schématique de la méthode d'estimation des valeurs de référence utilisées dans le calcul des EQRs des trois métriques constitutives de l'indice PH3. a) Approche proposée par Duflos et al. (2018) dans le cadre du 1^{er} cycle d'évaluation. b) Approche proposée et utilisée dans le rapport. δ : Hulburt, D : indice de Menhinick, LCBD : Contribution Locale à la Bêta Diversité. Les points violets correspondent aux valeurs de référence pour chacune des métriques (i.e., valeurs maximales pour $100 - \delta$ et D et valeurs minimales pour le LCBD). 31
- Figure 18.** Représentation schématique de la procédure de dégradation des séries temporelles pour estimer la sensibilité de l'indicateur PH3 à la fréquence d'acquisition des données. Six scénarios de dégradations ont été considérés, en réalisant un ré-échantillonnage aléatoire des données des séries temporelles. Des tirages aléatoires ont été réalisés : 1) un mois par saison, 2) deux mois par saison, 3) 11 mois par an, 4) une saison par an, 5) deux saisons par an, et 6) trois saisons par an. A partir de ces ré-échantillonnages, l'indicateur PH3 a été calculé et comparé ensuite aux valeurs obtenues à partir de séries temporelles complètes. 32
- Figure 19.** Illustration de la matrice ($n \times m$) de distance cumulée entre deux séries temporelles Q et C, ainsi qu'une fenêtre de déformation w et le chemin P minimisant le coût de déformation entre les deux séquences..... 35
- Figure 20.** Enveloppe calculée à partir de l'abondance du couple fonctionnel (**diatomées-dinophycées**) estimée par **microscopie** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période

- d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge). 37
- Figure 21.** a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-dinophycées**) estimé par **microscopie** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des diatomées et des dinophycées. 37
- Figure 22.** Enveloppe calculée à partir de l'abondance des couples fonctionnels (**diatomées-dinophycées**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge). 38
- Figure 23.** a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-dinophycées**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des diatomées et des dinophycées. 38
- Figure 24.** Enveloppe calculée à partir de l'abondance des couples fonctionnels (**diatomées-cercozoaires**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge). 40
- Figure 25.** a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-cercozoaires**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des diatomées et des cercozoaires. 40
- Figure 26.** Enveloppe calculée à partir de l'abondance des couples fonctionnels (**hétérotrophes-autotrophes**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge). 41
- Figure 27.** a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**hétérotrophes-autotrophes**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des hétérotrophes et des autotrophes. 41

Figure 28. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station SOMLIT-Astan à Roscoff pour la période 2009-2016 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).....	42
Figure 29. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station SOMLIT-C à Wimereux pour la période 2016-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).....	43
Figure 30. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station B8 de la Marina du Château pour la période 2019-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).....	43
Figure 31. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station H42 de la Marina du Château pour la période 2019-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).....	44
Figure 32. Variations temporelles de la LCBD à la station SOMLIT-Astan à Roscoff pour la période 2009-2016 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite) .	44
Figure 33. Variations temporelles de la LCBD à la station SOMLIT-C à Wimereux pour la période 2016-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite) .	45
Figure 34. Variations temporelles de la LCBD à la station B8 de la Marina du Château pour la période 2019-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).....	45
Figure 35. Variations temporelles de la LCBD à la station H42 de la Marina du Château pour la période 2019-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).....	45
Figure 36. Relation entre les valeurs de l'indicateur PH3 inféré à partir des données de microscopie et de metabarcoding (ARNr 18S, fraction de taille >3 µm) . a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris correspond à une date d'échantillonnage de la station SOMLIT-Astan à Roscoff. La relation entre les deux formes de PH3 a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman. .	46
Figure 37. Relation entre les valeurs de l'indicateur PH3 inféré à partir des données de microscopie et de metabarcoding (ARNr 18S, fraction de taille >3 µm) . a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris correspond à une date d'échantillonnage de la station SOMLIT-C à Wimereux. La relation entre les deux formes de PH3 a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman. .	47
Figure 38. Relation entre les valeurs de l'indicateur PH3 inféré à partir des données de microscopie et de metabarcoding (ARNr 18S, fraction de taille >3 µm) . a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris (clair et foncé) correspond à une date d'échantillonnage de la station de la Marine du Château à Brest. Les points gris clair correspondent aux dates d'échantillonnage du site B8 et les points gris foncé correspondent aux dates d'échantillonnage du site H42. La relation entre les deux formes de PH3 a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman.....	47
Figure 39. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de microscopie et les données de metabarcoding (ARNr 18S, fraction de	

taille >3 µm) sur l'ensemble de la série temporelle (de 2009 à 2016) à la station SOMLIT-Astan à Roscoff. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding et deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose..... 49

Figure 40. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de **microscopie** et les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)** sur l'ensemble de la série temporelle (de 2015 à 2020) à la station SOMLIT-C à Wimereux. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding et deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose..... 50

Figure 41. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de microscopie et les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)** sur l'ensemble de la série temporelle (de 2019 à 2020) à la station de la Marine du Château à Brest. Les résultats présentent les données des deux sites B8 et H42 combinés. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding et deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose. 51

Figure 42. Sensibilité de l'indicateur PH3 à la fréquence d'acquisition des données. L'indicateur PH3 est inféré à partir des données de **microscopie** de la station SOMLIT-Astan à Roscoff. En ordonnées, la sensibilité est exprimée comme la valeur absolue de la différence entre les valeurs de PH3 observées (i.e., inférées à partir du jeu de données brutes) et les valeurs de PH3 attendues (i.e., inférées à partir des jeux de données dégradés). Six scénarios de dégradation ont été effectués

par ré-échantillonnage aléatoire des données des séries temporelles. Les détails sur les procédures de dégradations des données figurent dans le paragraphe 7.1.3 (voir notamment dans la figure 18).....	53
Figure 43. Sensibilité de l'indicateur PH3 à la fréquence d'acquisition des données. L'indicateur PH3 est inféré à partir des données de metabarcoding (ARNr 18S, fraction de taille >3 µm) de la station SOMLIT-Astan à Roscoff.	54
Figure 44. Lien entre l'indicateur PH3 et les paramètres environnementaux pour la station SOMLIT-Astan à Roscoff. La similarité entre la série temporelle de l'indicateur PH3 (données de microscopie) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de metabarcoding (ARNr 18S, fraction de taille >3 µm) . Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Les valeurs estimées du lien entre le PH3 et les variables environnementales sont ainsi supérieures pour un calcul à partir des données de metabarcoding lorsque la différence est supérieure à 0. A l'inverse, les valeurs estimées du lien entre le PH3 et les variables environnementales sont supérieures pour un calcul à partir des données de microscopie lorsque la différence est inférieure à 0.	56
Figure 45. Lien entre l'indicateur PH3 et les paramètres environnementaux pour la station SOMLIT-C à Wimereux. La similarité entre la série temporelle de l'indicateur PH3 (données de microscopie) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de metabarcoding (ARNr 18S, fraction de taille >3 µm) . Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. ...	57
Figure 46. Lien entre l'indicateur PH3 et les paramètres environnementaux pour le site B8 à la Marina du château à Brest. La similarité entre la série temporelle de l'indicateur PH3 (données de microscopie) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de metabarcoding (ARNr 18S, fraction de taille >3 µm) . Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Al : Aluminium, As : Arsenic, Co : Cobalt, Cr : Chrome, Cu : Cuivre, Fe : Fer, Mn : Manganèse, Ni : Nickel, Pb : Plomb, Ti : Titane, V : Vanadium, Zn : Zinc.	58
Figure 47. Lien entre l'indicateur PH3 et les paramètres environnementaux pour le site B8 à la Marina du château à Brest. La similarité entre la série temporelle de l'indicateur PH3 (données de microscopie) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de metabarcoding (ARNr 18S, fraction de taille >3 µm) . Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Al : Aluminium, As : Arsenic, Co : Cobalt, Cr : Chrome, Cu : Cuivre, Fe : Fer, Mn : Manganèse, Ni : Nickel, Pb : Plomb, Ti : Titane, V : Vanadium, Zn : Zinc.	59

LISTE DES TABLES

Tableau 1. Paires de formes de vie phytoplanctoniques considérées pour l'indicateur PH1 (Ostle et al. 2017).....	5
Tableau 2. Données de metabarcoding utilisées dans le cadre du projet INDIGENE. A noter que différentes amorces de séquençage ont été utilisées selon le compartiment biologique considéré : ¹ TAREuk454FWD1 et TAREukREV3 (Stoeck et al. 2010), ² amorces 515F(Y)/926R (Parada et al. 2016).....	12
Tableau 3. Paramètres environnementaux in situ. T° : température (°C), S : salinité (PSU), O ₂ : oxygène (ml. L ⁻¹), NH ₄ ⁺ : Ammonium (mol.L ⁻¹), NO ₃ ⁻ : Nitrate (μmol.L ⁻¹), NO ₂ ⁻ : Nitrite (μmol.L ⁻¹), PO ₄ ³⁻ : Phosphate (μmol.L ⁻¹), SiO ₄ ²⁻ : Silice (μmol.L ⁻¹), COP : Carbone Organique Particulaire (μg.L ⁻¹), NOP : Azote Organique Particulaire (μg.L ⁻¹), MES : Matière en suspension (mg.L ⁻¹), δ ¹³ C : rapport isotopique (¹³ C/ ¹² C) du carbone (‰), δ ¹⁵ N : rapport isotopique (¹⁵ N/ ¹⁴ N) de l'azote (‰), Chl-a : Chlorophylle a (μg.L ⁻¹). Pour les stations adossées au SOMLIT, les données sont disponibles sur le site du SOMLIT (http://www.somlit.fr/).....	14
Tableau 4. Paramètres continentaux. Q : débits rivières (m ³ .s ⁻¹), NH ₄ ⁺ : Ammonium (mol.L ⁻¹), NO ₃ ⁻ : Nitrate (μmol.L ⁻¹), NO ₂ ⁻ : Nitrite (μmol.L ⁻¹), PO ₄ ³⁻ : Phosphate (μmol.L ⁻¹), SiO ₄ ²⁻ : Silice (μmol.L ⁻¹), Al : Aluminium (μmol.L ⁻¹), As : Arsenic (μmol.L ⁻¹), Co : Cobalt (μmol.L ⁻¹), Cr : Chrome (μmol.L ⁻¹), Cu : Cuivre (μmol.L ⁻¹), Fe : Fer (μmol.L ⁻¹), Mn : Manganèse (μmol.L ⁻¹), Ni : Nickel (μmol.L ⁻¹), Pb : Plomb (μmol.L ⁻¹), Ti : Titane (μmol.L ⁻¹), V : Vanadium (μmol.L ⁻¹), Zn : Zinc (μmol.L ⁻¹). L'ensemble de ces données est accessible au lien suivant : https://drive.google.com/drive/folders/16kuaXFrCnus5DKLUYeTYHUV5n-Ft2RpP	14
Tableau 5. Accessibilité des données de metabarcoding, de comptages microscopiques et des paramètres environnementaux in situ et continentaux pour les quatre stations du projet INDIGENE.	15
Tableau 6. Estimation des différences entre le site B8 et le site H42 dans les valeurs mesurées de chaque paramètre environnemental à l'aide d'un test des rangs signés de Wilcoxon. En gras, les valeurs significatives.	17
Tableau 7. Test de la tendance monotonique (test de Mann-Kendall) des séries temporelles des valeurs annuelles moyennes des différents paramètres environnementaux mesurés à la station SOMLIT-Astan de 2009 à 2016. En gras, les valeurs significatives.....	19
Tableau 8. Conclusions sur la pertinence d'utiliser les données de metabarcoding pour la surveillance des habitats pélagiques.....	60

LISTE DES ABBRÉVIATIONS ET ACRONYMES

BEE : Bon État Écologique

DCE : Directive Cadre sur l'Eau

DCSMM : Directive Cadre Stratégique pour le Milieu Marin

EQR : Ecological Quality Ratio (Ratio de qualité écologique)

GdG : Golfe de Gascogne

ICES : International Council for the Exploration of the Sea (CIEM en français)

ICG-COBAM : Intersessional Correspondence Group – Coordination of Biodiversity Assessment and Monitoring (Groupe de travail OSPAR sur la biodiversité)

MC : Mers Celtiques

MMN : Manche-Mer du Nord

MO : Méditerranée Occidentale

OSPAR : Convention des Mers Régionales « Oslo-Paris »

PCR : Amplification en chaîne par polymérase

PH1 : Indicateur « Pelagic Habitats » 1

PH2 : Indicateur « Pelagic Habitats » 2

PH3 : Indicateur « Pelagic Habitats » 3

PHYTOBS : Réseau d'Observation du Phytoplancton (<https://www.phytobs.fr/>)

SNO : Service National d'Observation

SOMLIT : Service d'Observation du Milieu LITtoral (<https://www.somlit.fr/>)

SRM : Sous-Région Marine

WG-PME : Groupe de travail « Phytoplankton and Microbial Ecology »

1. Introduction

Sous l'action des activités humaines et des changements globaux qu'elles induisent, les écosystèmes aquatiques sont confrontés à une gamme sans précédent de pressions environnementales (Sala et al. 2000, Halpern et al. 2008). L'impact de ces pressions environnementales sur les écosystèmes aquatiques est une préoccupation sociétale, dans la mesure où elles affectent les niveaux de biodiversité qu'ils abritent (Pereira et al. 2012, Pimm et al. 2014), ce qui compromet leur fonctionnement (Hautier et al. 2015, Duffy et al. 2017) et *in fine* la fourniture de services écosystémiques pour le bien-être des populations humaines (Worm et al. 2006, Walsh et al. 2016). La vulnérabilité des écosystèmes aquatiques, et notamment des écosystèmes marins, à ces changements environnementaux a conduit à la mise en place de la Directive Cadre Stratégique pour le Milieu Marin (2008/56/CE), en complément des directives européennes antérieures, notamment la Directive Cadre sur l'Eau (2000/60/CE). Pour répondre aux objectifs de ces directives, chaque Etat membre de l'Union Européenne s'engage à mettre en place les stratégies nécessaires pour atteindre et/ou maintenir le Bon Etat Ecologique (BEE) des écosystèmes aquatiques. Ces stratégies recouvrent la définition du bon état des différentes composantes de biodiversité et de leur environnement, des pressions anthropiques qui les impactent, d'un programme de surveillance, d'objectifs environnementaux (qui ciblent la réduction de pressions) et la mise en place d'un plan d'action (par la définition de mesures) pour atteindre le BEE (une phase d'évaluation permet de statuer sur l'atteinte ou non de BEE).

Les habitats pélagiques marins abritent l'ensemble des organismes animaux et végétaux vivant en plein eau et se déplaçant essentiellement de manière passive, sans rapport direct avec les habitats benthiques. Ces organismes planctoniques sont des intégrateurs des changements des conditions environnementales (Kirby & Beaugrand 2009). Outre leur variabilité naturelle saisonnière, des modifications de leur distribution, de leur abondance et de leur biomasse peuvent se produire de manière abrupte et refléter l'impact d'événements extrêmes (e.g., tempête, Vincent et al. 2002) ou de manière graduelle en réponse à des changements s'exerçant à plus long terme (e.g., changement global ; Beaugrand 2012). En conséquence, l'évaluation de la qualité écologique des habitats pélagiques marins se base en partie sur la surveillance spatio-temporelle de différents compartiments biologiques planctoniques et leurs réponses aux facteurs environnementaux naturels et d'origine anthropique. Les directives européennes cadre sur l'eau (DCE) et stratégie pour le milieu marin (DCSMM) requièrent l'évaluation de l'état écologique des communautés planctoniques. Les communautés phytoplanctoniques sont ainsi un des « éléments de qualité biologique » de la DCE et les « Habitats pélagiques » constituent une composante du Descripteur 1 – Biodiversité de la DCSMM. Ces évaluations s'appuient sur des indicateurs et sur des programmes de surveillance permettant d'acquérir les données requises pour leur calcul. Le choix des compartiments biologiques planctoniques à surveiller est conditionné par leur place dans les réseaux trophiques et leur rôle fonctionnel. Ainsi, les évaluations et programmes de surveillance mis en œuvre doivent reposer sur les compartiments phytoplanctonique, zooplanctonique et microbien. Les communautés phytoplanctoniques constituent des éléments essentiels des réseaux trophiques marins. En tant que principaux producteurs de matière organique, la structure des communautés phytoplanctoniques ainsi que leur variabilité spatio-temporelle conditionnent fortement celles de leurs prédateurs zooplanctoniques (e.g., dinoflagellés, ciliés, copépodes). Les organismes zooplanctoniques sont à la fois des consommateurs primaires et proies de choix pour les niveaux trophiques

supérieurs tels que les poissons. Ils occupent ainsi une position centrale dans les réseaux trophiques pélagiques comme lien dans les flux d'énergie et de matière entre la base des chaînes trophiques et les prédateurs supérieurs (Ware & Thompson 2005). Les organismes microbiens pélagiques (ou microorganismes pélagiques) comprennent tous les organismes unicellulaires ou coloniaux de taille microscopique. Ce compartiment comprend donc, des eucaryotes¹ phototrophes, hétérotrophes ou mixotrophes, des procaryotes (bactéries phototrophes ou non et archées) et les virus. Les organismes phytoplanctoniques (qui comprennent des eucaryotes unicellulaires et des bactéries phototrophes) sont donc un compartiment microbien. Les microorganismes pélagiques jouent un rôle majeur dans les principales fonctions écosystémiques des habitats pélagiques marins (e.g., la séquestration du carbone, la régulation du climat et le cycle des nutriments ; Ducklow 2008, Falkowski et al. 2008, Hutchins & Fu 2017).

Les directives européennes doivent s'appuyer sur les travaux menés dans le cadre de la Commission européenne et de l'ICES, des conventions de mers régionales OSPAR et de Barcelone, pour répondre aux exercices de définition du BEE de l'ensemble des descripteurs (Guérin et al. 2012). Dans le cadre de la convention OSPAR, le groupe de travail de l'ICG-COBAM formé d'experts internationaux a notamment proposé trois indicateurs (*PH1*, *PH2*, *PH3*) pour renseigner le critère D1C6 de la DCSMM (seul critère considéré actuellement pour l'évaluation de l'état écologique des habitats pélagiques). Ces indicateurs « Habitats pélagiques » visent à décrire la structure et le fonctionnement des communautés planctoniques en couvrant des aspects complémentaires de leur dynamique. L'indicateur *PH1* porte sur les changements entre groupes fonctionnels du plancton, l'indicateur *PH2* décrit les variations de biomasse ou d'abondance du plancton et l'indicateur *PH3* s'intéresse aux changements dans la biodiversité du plancton. Ces indicateurs sont des indicateurs de surveillance (selon Bedford et al. 2018), c'est-à-dire que les variations temporelles observées dans la structure, l'abondance, la biomasse et/ou les rapports fonctionnels traduisent les effets des conditions environnementales naturelles ou liées à des pressions anthropiques dominantes. Ainsi, ces indicateurs agissent comme des signaux d'alarme à partir desquels des actions de gestion (principe de précaution) et/ou de recherche (études des liens état-pression) peuvent être entreprises. La mise en œuvre du programme de surveillance « Habitats pélagiques » au cours du 1^{er} cycle d'évaluation DCSMM a débuté par la prise en compte du compartiment phytoplanctonique, et dans une moindre mesure celle du compartiment zooplanctonique. La surveillance de nombreux groupes d'organismes qui composent les communautés microbiennes n'a pas été initiée durant le 1^{er} cycle d'évaluation. Pour l'heure, l'intégration des communautés procaryotes dans la DCSMM a essentiellement considéré les bactéries pathogènes pour l'homme (e.g., *Escherichia coli*, *Vibrio cholerae*) et leurs incidences sanitaires en lien avec le Descripteur 9. L'une des orientations du 2^{ème} cycle du programme « Habitats pélagiques » est donc de tester des dispositifs candidats permettant la mise en place d'une stratégie de surveillance du compartiment microbien dans son ensemble. Cette action a pour objectifs i) d'identifier et de référencer les données disponibles (accessibles et bancarisées) pouvant alimenter les indicateurs existants pour l'évaluation de l'état écologique du descripteur « Habitats pélagiques », ii) de réaliser les tests numériques pour le calcul des indicateurs concernés, et de contribuer aux développements méthodologiques sur les indicateurs (seuils, méthodes d'intégration) en lien avec les groupes de travail européens

¹ Ces organismes sont aussi parfois regroupés sous le nom de protistes ou microeucaryotes.

(e.g., ICG-COBAM, WG-PME de l'ICES), et iii) d'intégrer la composante bactérienne dans les évaluations, en développant un/des indicateur(s) pertinent(s).

2. Les indicateurs BEE du descripteur « Habitats pélagiques »

Dans le cadre de la DCSMM, la France a choisi de baser l'évaluation du BEE pour les habitats pélagiques marins sur des mesures d'abondance et/ou de biomasse ainsi que de composition spécifique des communautés planctoniques, alimentant le calcul des trois indicateurs (*PH1*, *PH2*, *PH3*) proposés par le groupe de travail de l'ICG-COBAM de la convention OSPAR. Les évaluations de l'élément de qualité biologique « phytoplancton » reposent sur un indicateur composé de trois indices (biomasse, abondance et composition), l'indice de composition étant en cours de développement. La commission européenne encourage l'articulation des évaluations et des programmes de surveillance des différentes directives. Les travaux proposés dans ce rapport s'appuient sur les développements conduits dans le cadre de la convention OSPAR et de la DCSMM, et ont vocation à alimenter les réflexions en cours pour les deux directives (i.e., la DCE et la DCSMM).

2.1. Indicateur PH1

L'indicateur *PH1* est basé sur la méthode du Plankton Index (PI ; Tett et al. 2008), une approche graphique utilisée pour le calcul d'indicateurs adaptés aux organismes planctoniques (Gowen et al. 2011, Gowen et al. 2013, McQuatters-Gollop et al. 2019). Le PI est un indice de composition qui trouve son fondement dans les théories développées par Margalef (1978) sur les successions écologiques. Cet indicateur se base sur un ratio d'abondance ou de biomasse entre deux formes de vie (ou groupes fonctionnels) planctoniques suivies au cours des cycles saisonniers. Sur le principe, le PI permet de décrire l'évolution saisonnière de ces formes de vie et de détecter les réponses des communautés planctoniques aux modifications des conditions environnementales, en quantifiant l'écart entre une succession d'états observés et une succession d'états de référence. L'estimation de ces écarts permet d'évaluer le retour à l'état de référence ou l'éloignement de celui-ci. Des écarts importants renseignent d'un évènement inhabituel (s'il est ponctuel), ou d'un éventuel changement de régime (si un nouveau cycle saisonnier s'établit), pouvant être liés à des perturbations de type changements environnementaux suite à des changements globaux ou des pressions anthropiques. Dans les deux cas, ces changements peuvent avoir des répercussions sur le fonctionnement de l'écosystème considéré. Une enveloppe de référence est théoriquement définie sur la base de données (abondance ou biomasse) acquises pour deux formes de vie sur la période antérieure à la période d'évaluation. En pratique, les données traduisant l'état de référence sont projetées sur un plan bidimensionnel pour former une enveloppe sur laquelle sont ensuite projetées celles acquises dans le cadre de la période d'évaluation et traduisant alors l'état observé. La trajectoire entre l'état de référence et l'état observé reflète i) la variabilité cyclique et à moyen terme (cycles saisonniers, successions d'espèces et variabilité interannuelle) et ii) la variabilité à long terme pouvant résulter des pressions environnementales. A l'origine du développement du PI, seuls les ratios entre diatomées et dinoflagellés ont été considérés, mais par la suite la méthode a été appliquée à d'autres couples de formes de vie (Tableau 1).

Cet indicateur renseigne également l'indicateur FW5 et contribue ainsi à renseigner les critères du descripteur 4 « Réseaux trophiques » de la DCSMM, D4C2 (équilibre de l'abondance entre guildes trophiques, concernant certaines paires de groupes fonctionnels) et D4C3 (répartition par taille des individus au sein d'une guildes trophique, concernant la paire phytoplancton >20µm et <20µm). Par ailleurs, les paires fonctionnelles « dinoflagellés potentiellement nuisibles/toxiques sur dinoflagellés totaux » et « dinoflagellés nuisibles/toxiques sur les diatomées totales » sont liées au critère D5C3 du descripteur 5 « Eutrophisation » qui vise à estimer l'étendue spatiale et la durée des efflorescences d'algues nuisibles et/ou toxiques.

2.2. Indicateur PH2

L'indicateur *PH2* repose sur l'analyse de séries temporelles d'un proxy de la biomasse totale du phytoplancton (ici la concentration en chlorophylle *a*) ou de l'abondance totale des copépodes (Duflos et al. 2018). Les copépodes ont été choisis comme proxy pour suivre l'abondance totale du zooplancton car ils représentent la sous-classe la plus abondante et la plus ubiquiste du zooplancton dans les milieux marins. Ces organismes peuvent représenter jusqu'à 80-95% de la biomasse totale et 70-95% de l'abondance du mésozooplancton (Turner 2004). La dynamique du plancton présentant généralement un cycle saisonnier marqué, le *PH2* se calcule à un pas de temps mensuel sous la forme d'anomalies par rapport à un cycle saisonnier de référence préétabli. Des ajustements méthodologiques ont été apportés lors du 1^{er} cycle d'évaluation du descripteur « Habitats pélagiques » par rapport aux développements réalisés dans le cadre de la convention OSPAR (Duflos et al. 2018). Tout d'abord, le cycle saisonnier de référence est calculé sur une période de référence (définie par défaut comme étant la période antérieure à la période d'évaluation) et non plus sur l'ensemble de la série temporelle. Ensuite, l'interprétation du *PH2* repose sur l'observation des tendances présentées par les anomalies par rapport au cycle saisonnier de référence et de leurs sommes cumulées sur les périodes de référence et d'évaluation et non plus sur la catégorisation des anomalies en événements extrêmes/importants/faibles.

Dans sa configuration actuelle, cet indicateur est en lien avec le critère D5C2 (chlorophylle *a* dans la colonne d'eau) du descripteur « Eutrophisation » et repose sur l'acquisition des mêmes données relatives à la biomasse phytoplanctonique à partir des concentrations en chlorophylle *a*. Par ailleurs, l'indicateur FW6 (indicateur candidat OSPAR réseaux trophiques) est une adaptation du *PH2*-zooplancton considérant les biomasses (au lieu des abondances) et les classes de taille. Une fois la taille des copépodes prise en compte, le *PH2* sera alors en lien avec le critère D4C3.

Tableau 1. Paires de formes de vie phytoplanctoniques considérées pour l'indicateur PH1 (Ostle et al. 2017)

Intitulé	Groupes phytoplanctoniques (numérateur/dénominateur)	Critère de sélection	Niveau de confiance de l'indice	Phénomène écologique représenté
Life Form Pair « Group »	Diatomées	Toutes diatomées	Élevé	La dominance des dinoflagellés peut représenter un indicateur d'eutrophisation et/ou de changements dans les réseaux trophiques (perturbations)
	Dinoflagellés	Tous dinoflagellés		
Life Form Pair « Size »	Phytoplancton de grande taille (micro)	Cellules > 20µm	Élevé , même si difficulté à visualiser et compter les cellules nanoplanctoniques et même si les cellules picoplanctoniques sont ignorées (non détectées en microscopie photonique traditionnelle)	Indicateur basé sur la taille, indiquant l'efficacité du transfert énergétique vers les niveaux trophiques supérieurs
	Phytoplancton de plus petite taille (nano)	Cellules < 20µm		
Life Form Pair « Feed »	Diatomées Dinoflagellés autotrophes et mixotrophes		Faible (actuellement) car difficulté à distinguer les dinoflagellés mixotrophes. Niveau pouvant être amélioré à l'avenir si recours à un proxy (e.g., déduire les hétérotrophes de l'abondance totale des dinoflagellés)	Un changement au sein des producteurs primaires pourrait indiquer des phénomènes d'eutrophisation
Life Form Pair « Habitat »	Diatomées pélagiques Diatomées thycopélagiques		Élevé	Indicateur de perturbation benthique et des événements de resuspension
Life Form Pair « Tox »	Diatomées potentiellement nuisibles/toxiques vs. Diatomées		Faible si identification au niveau du genre, car un même genre peut inclure à la fois des espèces toxiques ou nuisibles et non-toxiques ou non-nuisibles	Changement au sein de la communauté phytoplanctonique avec des événements nuisibles/toxiques (de types bloom) pouvant impacter les autres niveaux trophiques
	Dinoflagellés potentiellement nuisibles/toxiques vs. Dinoflagellés	Toutes les espèces de la liste IOC-UNESCO ²	Élevé si identification à l'espèce	

² <http://www.marinespecies.org/hab/>

2.3. Indicateur PH3

Les assemblages phytoplanctoniques sont généralement caractérisés par quelques espèces communes et de nombreuses espèces rares, et présentent une variabilité de composition (i) temporelle, caractérisée par une succession saisonnière et (ii) spatiale, caractérisée par exemple par un gradient latitudinal (Reynolds 2006). Les causes de succession sont la réactivité des espèces à des conditions environnementales favorables tout au long de l'année, telles que les changements saisonniers de température, le mélange/la stratification de la colonne d'eau, la charge en nutriments et la disponibilité en lumière (Chalar 2009). D'autres processus agissent sur des périodes allant de quelques jours à plusieurs semaines, comme les événements météorologiques (e.g., vent, pluie, nébulosité) et les événements hydrologiques (e.g., upwelling/downwelling, prélèvement hydrologique, fluctuations du niveau d'eau). A long terme, les assemblages de phytoplancton réagissent également au changement climatique et aux processus à grande échelle tels que la variabilité des schémas de circulation atmosphérique (e.g., l'oscillation nord-atlantique). Le but de l'évaluation à l'aide d'indices de diversité n'est pas seulement de caractériser la composition des communautés, mais également de pouvoir détecter des changements dans la structure des communautés sur une base saisonnière et annuelle. Considérer simultanément différents indices se focalisant principalement sur le nombre d'espèces (ou la richesse), la dominance des espèces au sein des communautés ainsi que l'amplitude des variations de la structure entre les communautés (e.g., au cours du temps) fournissent une image holistique de la structure des communautés. L'indicateur *PH3* repose donc sur une approche multi-métrique combinant le calcul d'indices de diversité alpha et bêta (Duflos et al. 2018, Rombouts et al. 2019). Cet indicateur permet de décrire la structure et la diversité des communautés planctoniques et informe sur les phénomènes de dominance et de perturbation dans un continuum temporel. Pour le phytoplancton, la caractérisation de la diversité alpha, qui représente la diversité à l'échelle locale (Gaston & Spicer 2004), est basée sur un indice de diversité et un indice de dominance. L'indice de Menhinick (D) a été choisi comme indice de diversité, pour évaluer la richesse spécifique, du fait de sa forte sensibilité aux changements des conditions environnementales (Rombouts et al. 2019). Pour ce qui est de l'évaluation de la dominance des espèces, le choix s'est porté sur l'indice de Hulburt (δ), présentant l'avantage d'une relative facilité d'interprétation dans un contexte d'évaluation (Facca et al. 2014). Concernant la diversité bêta, qui correspond à la variation de composition en espèces entre des sites ou entre des événements temporels, il a été choisi de s'intéresser à la composante temporelle. Pour cela, la diversité bêta est considérée selon le calcul de l'indice de la Contribution Locale à la Diversité Bêta (LCDB, Legendre & De Cáceres 2013). L'indice LCDB utilise la variance de la distribution des espèces entre différents sites (ou différentes dates d'analyse pour une série temporelle) pour renseigner l'hétérogénéité entre les communautés planctoniques et donc dans quelle mesure chaque observation (site, date) contribue à la diversité bêta. L'intégration de ces trois indices à l'échelle de l'indicateur *PH3* se fait par l'estimation de leurs ratios de qualité écologique (EQR) respectifs et l'agrégation de ceux-ci par le calcul de moyenne.

Cet indicateur est en lien avec le critère D4C1 qui se réfère à la diversité des guildes trophiques et avec le critère D5C3 en cas de dominances fortes liées à la prolifération d'espèces phytoplanctoniques nuisibles et/ou toxiques.

3. Positionnement du projet INDIGENE dans les travaux d'évaluation de l'état écologique du descripteur « Habitats pélagiques » et le programme de surveillance associé

Le développement d'outils d'évaluation et de surveillance permettant de mettre en évidence les changements d'état écologique du compartiment microbien, et en particulier ceux conduisant à des dysfonctionnements attribuables à des modifications des conditions environnementales liées aux activités anthropiques (e.g., pollutions chimiques, eutrophisation), figure clairement dans les besoins des gestionnaires des milieux marins en charge de la mise en œuvre des politiques publiques (i.e., la DCE, la DCSMM, mais également les documents de gestion des aires marines protégées). L'étude taxinomique des communautés planctoniques, par microscopie optique, dépend du niveau d'expertise mobilisée et devient impossible dès lors que les espèces sont de trop petite taille (i.e., lorsque la résolution optique est insuffisante), et lorsqu'elles sont cryptiques ou en très faibles abondances. Ainsi, les méthodes de microscopie optique sont inadéquates pour le suivi de la diversité de nombreux organismes microbiens comme les virus, les archées et bactéries, mais aussi les eucaryotes pico- ou nano-planctoniques. L'approche par comptage et/ou observation est également très chronophage, ce qui limite son déploiement à grande échelle ou à haute fréquence temporelle.

Les nouvelles technologies de séquençage à haut débit et les méthodes moléculaires de metabarcoding (séquençage de marqueurs amplifiés à partir de l'ADN extrait d'une communauté complexe) sont actuellement les outils les plus performants pour caractériser la diversité des assemblages planctoniques microbiens (Zimmerman et al. 2015). De plus, plusieurs études ont suggéré la pertinence de ce type d'approches pour le calcul d'indices de polluosensibilité ou d'indices de qualité écologique pour les milieux d'eau douce (e.g., Cordier et al. 2017, Rivera et al. 2020). Ces approches semblent donc des outils prometteurs pour la surveillance des habitats pélagiques marins. Par exemple, les données semi-quantitatives obtenues par metabarcoding pourraient alimenter l'indicateur *PH1* dans la mesure où elles rendent compte de l'abondance relative d'un groupe fonctionnel/taxinomique donné par rapport à un autre (en termes de nombre de séquences). L'accès à une résolution plus fine de la diversité par les méthodes de metabarcoding permettrait également d'acquérir des données alimentant directement le calcul de l'indicateur *PH3* relatif aux changements dans la diversité planctonique. Toutefois, le fait que les données de metabarcoding soient semi-quantitatives et exprimées en représentativité des taxons (en proportion du nombre total compté) ne permet pas pour le moment d'alimenter directement l'indicateur *PH2* basés sur des données quantitatives de biomasse.

Plusieurs étapes sont nécessaires avant de pouvoir utiliser ce type de données de manière opérationnelle pour l'évaluation du BEE des habitats pélagiques marins. Tout d'abord, une comparaison des méthodes moléculaires et des méthodes classiques de comptage (e.g., microscopie optique) est indispensable avant de pouvoir utiliser des données moléculaires dans un cadre réglementaire, pour la surveillance mais aussi pour l'évaluation (Goodwin et al. 2017, Hering et al. 2018). Les techniques moléculaires et les protocoles d'échantillonnage et d'analyse (pipelines analytiques pour le traitement bio-informatique des données moléculaires) doivent également être standardisés et harmonisés avant de pouvoir être utilisés en routine. Enfin, les développements numériques doivent s'intéresser à traduire les résultats des analyses moléculaires en indicateurs et métriques pouvant être utilisés pour les évaluations réglementaires du BEE. Dans ce cas précis, il est nécessaire dans un premier

temps de tester, voire adapter, les indicateurs déjà mis en place pour ces évaluations du BEE des habitats pélagiques marins, et notamment les indicateurs *PH1* et *PH3*.

Ces travaux préliminaires doivent s'appuyer tout d'abord sur des données existantes avant d'envisager une stratégie de déploiement opérationnelle.

Dans ce contexte le projet INDIGENE fait office d'étude pilote avec pour objectif de tester la pertinence des données de diversité de type metabarcoding des compartiments microbiens eucaryotes et/ou procaryotes pour l'évaluation du BEE des habitats pélagiques marins. La pertinence des données de métabarcodes est évaluée en utilisant des données morphologiques (i.e., comptage du microphytoplancton par microscopie optique) à titre de comparaison. Le projet se décline en quatre actions :

- La première action a consisté à bancaiser l'ensemble des données, de nature différentes, mobilisables pour le projet.
- La seconde action avait pour objectif de mettre en évidence les enveloppes de variation de la diversité des communautés microbiennes (procaryotes et eucaryotes), sur la base de données moléculaires et morphologiques, et d'estimer le rôle des facteurs environnementaux sur la diversité de ces compartiments biologiques.
- La troisième action a consisté à calculer les indicateurs utilisés classiquement dans la surveillance des habitats pélagiques (i.e., en particulier les indicateurs *PH1* et *PH3*) à partir des données de metabarcoding et à tester leur comportement, en particulier en lien avec les pressions environnementales.
- La quatrième action a été de fournir un cadre analytique standardisé et reproductible permettant le calcul d'indicateurs de composition des communautés microbiennes pélagiques marines et la description de la stratégie de surveillance associée.

Le projet INDIGENE repose sur quatre stations ateliers dédiées (Figure 1), dont trois sont issues des réseaux de surveillance nationaux SOMLIT et PHYTOBS (labélisés comme SNO) : la station SOMLIT-Astan à Roscoff, la station SOMLIT-C à Wimereux et la station SOMLIT-Portzic (Saint-Anne-du-Portzic) à Brest. La quatrième station est localisée à la Marina du Château dans le port de Brest et se caractérise par trois sites d'échantillonnage (B8, F32, H42) positionnés le long d'un gradient de contamination chimique par les métaux lourds. Cette dernière est le site d'étude retenu dans le cadre du projet de recherche PHYTOPORT (projet incitatif de l'UMR 7144) porté par Dominique Davoult et Nathalie Simon, qui est également en charge du projet INDIGENE.

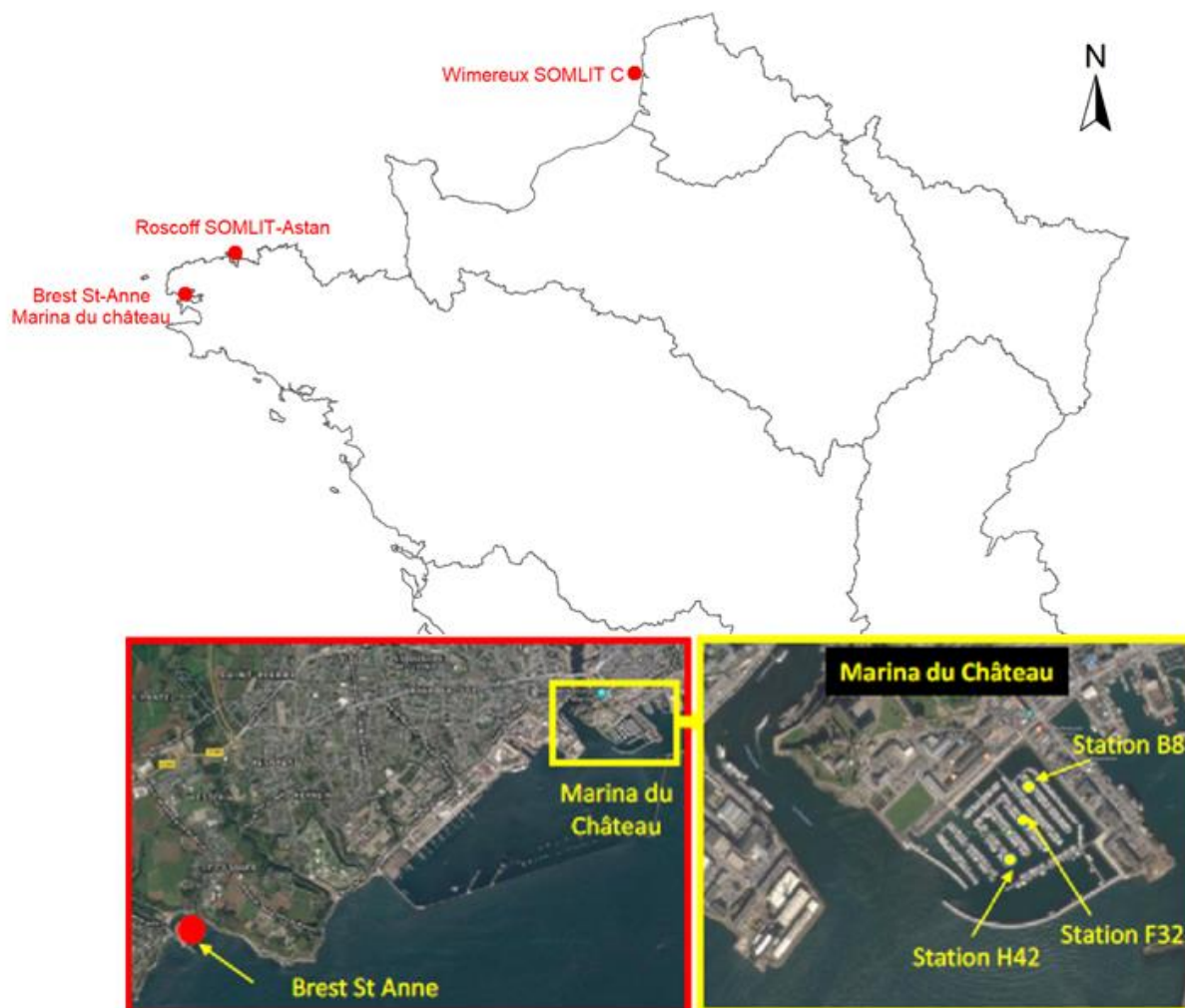


Figure 1. Localisation géographique des quatre stations ateliers (points rouges) dédiées au projet INDIGENE. Pour la station à la Marina du Château dans le port de Brest, le gradient de contamination par les métaux lourds s'établit du ponton le plus interne au ponton le plus externe du port.

4. Caractéristiques des stations d'échantillonnage

Les quatre stations retenues pour le projet se répartissent sur deux sous-régions marines, la sous-région marine Manche/Mer du Nord pour la station SOMLIT-C à Wimereux et la sous-région marine Mers Celtiques pour la station SOMLIT-Astan à Roscoff et les deux stations situées à Brest (i.e., SOMLIT-Portzic et à la Marina du Château). Ces sous-régions marines sont constituées de plusieurs paysages marins, chacun correspondant à des masses d'eau homogènes du point de vue de leur fonctionnement hydrologique et hydrodynamique. La station SOMLIT-C à Wimereux, la station SOMLIT-Astan à Roscoff et les deux stations à Brest font partie de paysages marins différents (Tew-Kai et al. 2020).

Les trois stations adossées au SOMLIT font l'objet de suivis depuis plusieurs années voire décennies, avec notamment l'acquisition de 13 paramètres hydrologiques et des abondances du plancton (<https://www.somlit.fr/>, <https://www.phytobs.fr/>).

- **La station SOMLIT-C** (50°40'75"N, 1°31'17"E), **proche de Wimereux**, est située à 1 mile nautique des côtes et présente une profondeur d'environ 25 m. Cette station est considérée comme représentative de la zone côtière de la Manche Orientale, tout en étant caractérisée par des eaux côtières séparées des eaux du large et formant un fleuve marin côtier (Brylinski et al. 1991). En Manche Orientale, où le régime marégraphique est mégatidal et suscite des courants alternatifs et parallèles à la côte, le phénomène de « fleuve côtier » se caractérise par la création d'une masse d'eau côtière permanente dont l'individualité (dessalure, turbidité, richesse en phytoplancton, teneur en polluants) est sans cesse entretenue par les apports fluviaux qui s'échelonnent de la baie de Seine au Pas-de-Calais (Brylinski et al. 1991). La station SOMLIT-C est également fortement soumise aux apports côtiers présentant un déséquilibre en sels nutritifs disponibles et notamment un excès d'azote et une limitation en silicate (e.g., Peperzak et al. 1998, Schapira et al. 2008, Grattepanche et al. 2011).
- **La station SOMLIT-Astan** (48°46'40"N, 3°56'15"W) est située **au large de Roscoff** à 2.5 miles nautiques de la côte, au nord-est de l'île de Batz. A ce point, la hauteur de la colonne d'eau est de 60 m. Cette station est considérée comme représentative de la Manche Occidentale (en termes de structure et de profondeur moyenne de la colonne d'eau). Elle est très peu soumise à la pollution anthropique et l'eutrophisation (Guilloux et al. 2013).
- **La station SOMLIT-Portzic** (14°21'32"N, 4°33'07"W), localisée au niveau du goulet d'entrée de **la rade de Brest**, à l'interface avec la mer d'Iroise, présente une profondeur moyenne de 9 m. La rade de Brest est soumise à un régime marégraphique macrotidal induisant un fort mélange des eaux et réduisant l'influence des deux fleuves (l'Aulne et l'Elorn) qui s'y jettent (Le Pape et al. 1996).

Les trois sites localisés dans la Marina du Château du port de Brest sont considérés comme étant distribués le long d'un gradient de pollution par les hydrocarbures et les métaux lourds selon un gradient de confinement dans le port (Kenworthy et al. 2018).

5. Action 1 : Traitement, préparation et mise en forme des tableaux de données disponibles

Pour le projet INDIGENE, les périodes temporelles sélectionnées correspondent à celles pour lesquelles les laboratoires impliqués possédaient des jeux de données de metabarcoding. La figure 2 résume les périodes considérées et l'ensemble des données mises à disposition ainsi que les fréquences d'échantillonnages.

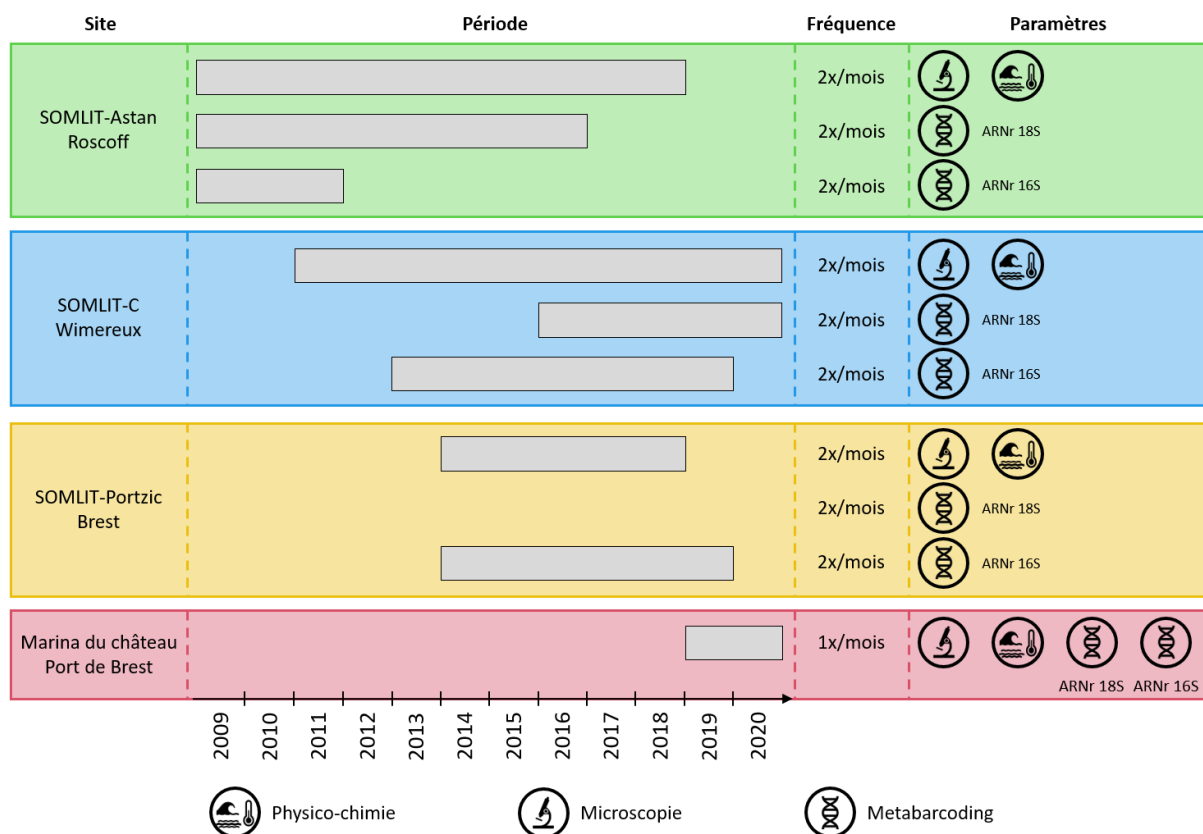


Figure 2. Périodes considérées, fréquences d'échantillonnage et type de données disponibles pour les stations étudiées dans le cadre de ce projet. Pour la Marina du Château, deux sites (B8 et H42) ont été étudiés.

5.1. Données de biodiversité obtenues par microscopie

Les données de comptage des taxons phytoplanctoniques (cellules $\geq 10 \mu\text{m}$) obtenues par microscopie aux sites des SNO SOMLIT et PHYTOBS ont été extraites des bases de données PELAGOS (<https://pelagos.sb-roscoff.fr/>) et/ou PHYTOBS (<https://www.phytobs.fr/>). Pour les sites de la Marina du Château, les données de comptage obtenues dans le cadre du projet PHYTOPORT ont été utilisées. Les périodes considérées ainsi que les fréquences d'échantillonnage, identiques à celles pour le metabarcoding, sont résumées dans le Tableau 2.

5.2. Données de biodiversité génétique (obtenues par metabarcoding)

Pour utiliser des données de metabarcoding issues de séquençage à haut débit, un traitement bio-informatique est nécessaire. La seconde étape de cette action a donc consisté à réaliser ce traitement metabarcoding pour l'ensemble des stations suivies dans le projet (Tableau 2). Les séquences brutes ont été obtenues par séquençage à haut-débit (technologie Illumina MiSeq) grâce à l'utilisation d'amorces spécifiques ciblant des régions variables du gène codant pour l'ARN 18S pour les eucaryotes et 16S pour les procaryotes. L'objectif était d'obtenir, pour chaque échantillon de plancton une liste de séquences uniques ainsi que pour chacune de ces séquences une abondance de « lectures ». Pour répondre à cet objectif, nous avons mis en place un pipeline d'analyse bio-informatique qui comprend une étape de « filtrage » permettant de se débarrasser du « bruit » et en particulier des séquences présentant des

erreurs introduites notamment lors des étapes d'amplification par PCR et de séquençage. Le pack logiciel « open source » DADA2 (Callahan et al. 2016) a été utilisé à cette fin. DADA2 permet l'obtention pour chaque échantillon d'une liste de séquences nommées ASV (pour Amplicon Sequence Variant). Cette stratégie visant à filtrer les erreurs et conserver les séquences uniques au lieu de regrouper les séquences en OTUs (Operational Taxonomic Units) a une résolution plus fine et permet de comparer facilement des séquences entre différentes études (Caruso et al. 2019). Un effort conséquent a été alloué pour automatiser le pipeline d'analyses bio-informatiques et proposer un guide d'utilisation.

Ce guide disponible sous la forme d'une page github, est accessible au lien suivant : https://github.com/benalric/MetaB_pipeline, et permet aux utilisateurs de mettre en pratique le pipeline tout en ayant accès aux codes sources.

Tableau 2. Données de metabarcoding utilisées dans le cadre du projet INDIGENE. A noter que différentes amorces de séquençage ont été utilisées selon le compartiment biologique considéré : ¹TAREuk454FWD1 et TAREukREV3 (Stoeck et al. 2010), ²amorces 515F(Y)/926R (Parada et al. 2016).

Site	Compartiment biologique	Marqueurs génétiques	Période, fréquence	Table d'ASVs obtenue ³
Wimereux (SOMLIT-C)	Procaryotes	16S (V3-V4)	2013-2019 (2x/mois)	Oui
	Eucaryotes (>3 µm)	18S (V4) ¹	2015-2020 (2x/mois)	Oui
Roscoff (SOMLIT-Astan)	Procaryotes	16S (V4-V5) ²	2009-2011 (2x/mois)	Oui
	Eucaryotes (0.2-3, >3 µm)	15S (V4) ¹	2009-2016 (2x/mois)	Oui
Brest (SOMLIT-Portzic)	Procaryotes (0.2, 3, 10 µm)	16 (V4-V5) ²	2014-2019 (2x/mois)	Oui
Marina du Château à Brest (sites B8, F32, H42)	Procaryotes	16S (V4-V5) ²	02/2019-03/2020 (mensuel)	Oui
	Eucaryotes	18S (V4) ¹	02/2019-03/2020 (mensuel)	Oui

³ Les différentes tables d'ASVs sont accessibles au lien suivant : <https://drive.google.com/drive/folders/1rR6jtb66b6ZsrfkD1Fv4XrA2U9humhG>

5.3. Paramètres environnementaux

La troisième étape de l'action 1 a consisté à faire un état des lieux des données contextuelles disponibles pour les différentes stations, avec pour objectif de mobiliser des données existantes sans coûts supplémentaires pour le programme de surveillance « Habitats pélagiques ». Pour les stations SOMLIT, nous avons interrogé la base SOMLIT (<https://www.somlit.fr/>) afin d'extraire les données relatives aux paramètres physico-chimiques mesurés *in situ* (Tableau 3).

Le suivi de ces différents paramètres physico-chimiques est réalisé selon des protocoles d'échantillonnage et d'analyse standardisés, consultables sur le site du SOMLIT.

Les sous-régions marines Manche/Mer du Nord et Mer Celtique étant soumises à des apports continentaux, nous avons également mobilisé les données de débits (<https://www.hydro.eaufrance.fr>) et de concentrations en nutriments (<https://naiades.eaufrance.fr>) des principaux affluents à l'origine de ces apports afin d'avoir des indicateurs de paramètres forçant d'origine anthropique (Tableau 4). Les eaux de la Manche Orientale, où se trouve la station SOMLIT de Wimereux, sont influencées par les apports continentaux provenant du sud notamment de la Somme ($35 \text{ m}^3 \cdot \text{s}^{-1}$, 245 km) et de la Canche ($12 \text{ m}^3 \cdot \text{s}^{-1}$, 100 km). Ces deux fleuves présentent respectivement un bassin versant couvrant une surface de 5560 km² et 1391 km², lesquels sont majoritairement agricoles (environ 80% de leur surface). La station SOMLIT de Roscoff est influencée par les apports continentaux venant principalement de la Penzé ($2.8 \text{ m}^3 \cdot \text{s}^{-1}$, 40 km) et l'Horn ($0.65 \text{ m}^3 \cdot \text{s}^{-1}$, 30 km), deux petits fleuves côtiers se jetant dans la Manche à l'est de la station. Ces deux fleuves côtiers se caractérisent par des bassins versants majoritairement agricole (environ 80% de leur surface) couvrant respectivement 210 km² et 94 km². La rade de Brest reçoit les apports continentaux de l'Aulne ($26 \text{ m}^3 \cdot \text{s}^{-1}$, 144 km) et de l'Elorn ($6 \text{ m}^3 \cdot \text{s}^{-1}$, 56 km), deux fleuves dont les bassins versants majoritairement agricoles (environ 80% de la surface) couvrent respectivement 1875 km² et 385 km². En parallèle, une marée moyenne compte pour environ 40 % du volume de la rade.

Les données des paramètres environnementaux acquis dans le cadre du projet PHYTOPORT pour les trois sites de la Marina du Château à Brest (Tableau 3) ont été obtenues en collaboration avec Laure Arsenieff, responsable des échantillonnages dans le cadre du projet. Pour ces trois sites, les mesures de concentrations en nutriments (NO_3^- , NO_2^- , SiO_4^{2-}) ont été réalisées selon les mêmes protocoles que pour les stations SOMLIT. Cependant, les profils de température de 0 à 2 m ont été obtenus à l'aide d'une sonde CTD YSI CastAway et non par une sonde CTD Sea-bird SBE19+ utilisée pour les stations SOMLIT. Pour les sites de la Marina du Château, située à l'entrée de la Penfeld à proximité des ports commerciaux et militaires, les concentrations en polluants (métaux lourds) mesurées dans le cadre du projet PHYTOPORT ont aussi été prises en compte (Tableau 4).

Tableau 3. Paramètres environnementaux in situ. T° : température ($^\circ\text{C}$), S : salinité (PSU), O_2 : oxygène (ml.L^{-1}), NH_4^+ : Ammonium (mol.L^{-1}), NO_3^- : Nitrate ($\mu\text{mol.L}^{-1}$), NO_2^- : Nitrite ($\mu\text{mol.L}^{-1}$), PO_4^{3-} : Phosphate ($\mu\text{mol.L}^{-1}$), SiO_4^{2-} : Silice ($\mu\text{mol.L}^{-1}$), COP : Carbone Organique Particulaire ($\mu\text{g.L}^{-1}$), NOP : Azote Organique Particulaire ($\mu\text{g.L}^{-1}$), MES : Matière en suspension (mg.L^{-1}), $\delta^{13}\text{C}$: rapport isotopique ($^{13}\text{C}/^{12}\text{C}$) du carbone (‰), $\delta^{15}\text{N}$: rapport isotopique ($^{15}\text{N}/^{14}\text{N}$) de l'azote (‰), Chl- a : Chlorophylle a ($\mu\text{g.L}^{-1}$). Pour les stations adossées au SOMLIT, les données sont disponibles sur le site du SOMLIT (<http://www.somlit.fr/>).

Site	Période	Paramètres
Wimereux (SOMLIT-C)	2013-2020	T° , S , O_2 , pH, NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-} , COP, NOP, MES, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, Chl- a
Roscoff (SOMLIT-Astan)	2009-2016	T° , S , O_2 , pH, NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-} , COP, NOP, MES, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, Chl- a
Brest (SOMLIT-Portzic)	2014-2019	T° , S , O_2 , pH, NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-} , COP, NOP, MES, $\delta^{13}\text{C}$, $\delta^{15}\text{N}$, Chl- a
Marina du Château à Brest (sites B8, F32, H42)	03/2019- 03/2020	T° , S , O_2 , NO_2^- , $\text{NO}_3^- + \text{NO}_2^-$, SiO_4^{2-} , Chl- a

Tableau 4. Paramètres continentaux. Q : débits rivières ($\text{m}^3.\text{s}^{-1}$), NH_4^+ : Ammonium (mol.L^{-1}), NO_3^- : Nitrate ($\mu\text{mol.L}^{-1}$), NO_2^- : Nitrite ($\mu\text{mol.L}^{-1}$), PO_4^{3-} : Phosphate ($\mu\text{mol.L}^{-1}$), SiO_4^{2-} : Silice ($\mu\text{mol.L}^{-1}$), Al : Aluminium ($\mu\text{mol.L}^{-1}$), As : Arsenic ($\mu\text{mol.L}^{-1}$), Co : Cobalt ($\mu\text{mol.L}^{-1}$), Cr : Chrome ($\mu\text{mol.L}^{-1}$), Cu : Cuivre ($\mu\text{mol.L}^{-1}$), Fe : Fer ($\mu\text{mol.L}^{-1}$), Mn : Manganèse ($\mu\text{mol.L}^{-1}$), Ni : Nickel ($\mu\text{mol.L}^{-1}$), Pb : Plomb ($\mu\text{mol.L}^{-1}$), Ti : Titane ($\mu\text{mol.L}^{-1}$), V : Vanadium ($\mu\text{mol.L}^{-1}$), Zn : Zinc ($\mu\text{mol.L}^{-1}$). L'ensemble de ces données est accessible au lien suivant : <https://drive.google.com/drive/folders/16kuaXFrCnus5DKLUYeTYHUV5n-Ft2RpP>.

Site	Affluents	Période	Paramètres
Wimereux (SOMLIT-C)	La Canche, la Somme	2013-2020	Q , NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-}
Roscoff (SOMLIT-Astan)	L'Horn, La Penzé	2009-2016	Q , NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-}
Brest (SOMLIT-Portzic)	L'Aulne, l'Elorn	2014-2019	Q , NH_4^+ , NO_3^- , NO_2^- , PO_4^{3-} , SiO_4^{2-}
Marina du Château à Brest (sites B8, F32, H42)		03/2019- 03/2020	Al, As, Co, Cr, Cu, Fe, Mn, Ni, Pb, Ti, V, Zn

5.4. Disponibilité des données

Les données utilisées (comptages microscopiques ; données de metabarcoding, i.e., abondances relatives des ASVs ; et paramètres environnementaux) sont disponibles dans un dossier partagé (Tableau 5).

Tableau 5. Accessibilité des données de metabarcoding, de comptages microscopiques et des paramètres environnementaux in situ et continentaux pour les quatre stations du projet INDIGENE.

Types de données	Lien internet
Metabarcoding	https://drive.google.com/drive/folders/1rR6jtbb6tb6Zsrfd1Fv4XrA2U9humhG
Microscopie	https://drive.google.com/drive/folders/1mctmQWjBubPI3ArdsAxWYw7jDOTR2xbK
Environnement	https://drive.google.com/drive/folders/16kuaXFrCnus5DKLUYeTYHUV5n-Ft2RpP

Il est important de noter ici, que pour la station SOMLIT-Portzic à Brest, les analyses de metabarcoding ont porté uniquement sur les procaryotes (ARNr 16S) ce qui rend impossible toute comparaison entre les indicateurs de changements (*PH1* et *PH3*) pour les eucaryotes. De plus, seuls les sites B8 et H42 de la station de la Marina du Château à Brest ont fait l'objet de comptages microscopiques, rendant également la comparaison des indicateurs de changements impossible pour le site F32.

6. Action 2 : Mise en évidence des patrons temporels de diversité et lien avec les facteurs environnementaux

Avant d'aborder le calcul d'indicateurs de changement, nous avons travaillé sur la mise en évidence des enveloppes de variation de la diversité microbienne, dont une description détaillée des résultats est disponible dans le rapport d'avancement N°2 du projet (Alric et al. 2021). L'identification des patrons de variation de la diversité dans le temps et dans l'espace et la mise en évidence des paramètres et processus qui façonnent la structure des communautés sont au cœur de nombreux travaux d'écologie (Cormack & Ord 1979). La structure des communautés est intrinsèquement spatiale pour deux raisons. Premièrement, la dynamique des populations et les interactions interspécifiques sont régies par le biais d'organismes qui n'existent et n'interagissent que dans leur voisinage immédiat (Tilman & Kareiva 1997). De tels processus biotiques sont à l'origine d'une autocorrélation spatiale au sein des communautés (Legendre 1993). Notons également que des processus dits neutres, qui ne sont pas fonctionnellement reliés à des changements de conditions environnementales, tels que la dérive écologique et la dispersion aléatoire peuvent engendrer une autocorrélation spatiale au sein des communautés (Diniz-Filho et al. 2012). Deuxièmement, les processus physiques créent une structure spatiale dans les facteurs environnementaux qui, à leur tour, provoquent une dépendance spatiale induite dans les communautés écologiques (Legendre 1993). De manière analogue, une structure temporelle des communautés se manifeste par l'observation de relations (ou l'absence d'indépendance statistique) entre différents intervalles de temps le long d'une série temporelle et découle de mécanismes similaires (Legendre & Gauthier 2014).

6.1. Variabilité temporelle des paramètres environnementaux

Pour l'ensemble des stations considérées, les paramètres environnementaux présentent des **variations saisonnières** marquées (voir par exemple les figures 3 et 4 pour les stations SOMLIT-Astan et pour les deux sites B8 et H42 localisés à la Marina du Château dans le port de Brest). Toutefois, les valeurs mensuelles ou amplitudes de variabilités varient selon les stations. Pour les 2 stations distantes de quelques dizaines de mètre dans le port de Brest, des différences significatives entre certains paramètres ont été observées (Tableau 6).

Pour les séries temporelles considérées, d'importantes variations interannuelles sont identifiées. Au site SOMLIT-Astan, aucune tendance interannuelle significative n'est détectée sur la période de 2009 à 2016 (à l'exception du pH, Figure 5, Tableau 7).

Au vu des résultats obtenus, il apparaît crucial de prendre soigneusement en compte la saisonnalité des variations des paramètres environnementaux lors des choix des analyses réalisées par la suite.



Figure 3. Variations mensuelles des paramètres environnementaux à la station SOMLIT-Astan de 2009 à 2016. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, COP : Carbone organique particulaire, $\delta^{13}\text{C}$: variation relative du rapport $^{13}\text{C}/^{12}\text{C}$, $\delta^{15}\text{N}$: variation relative du rapport $^{15}\text{N}/^{14}\text{N}$, MES : Matières en suspension, NH_4^+ : Ammonium, NO_2^- : Nitrites, NO_3^- : Nitrates, NOP : Azote organique particulaire, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates.

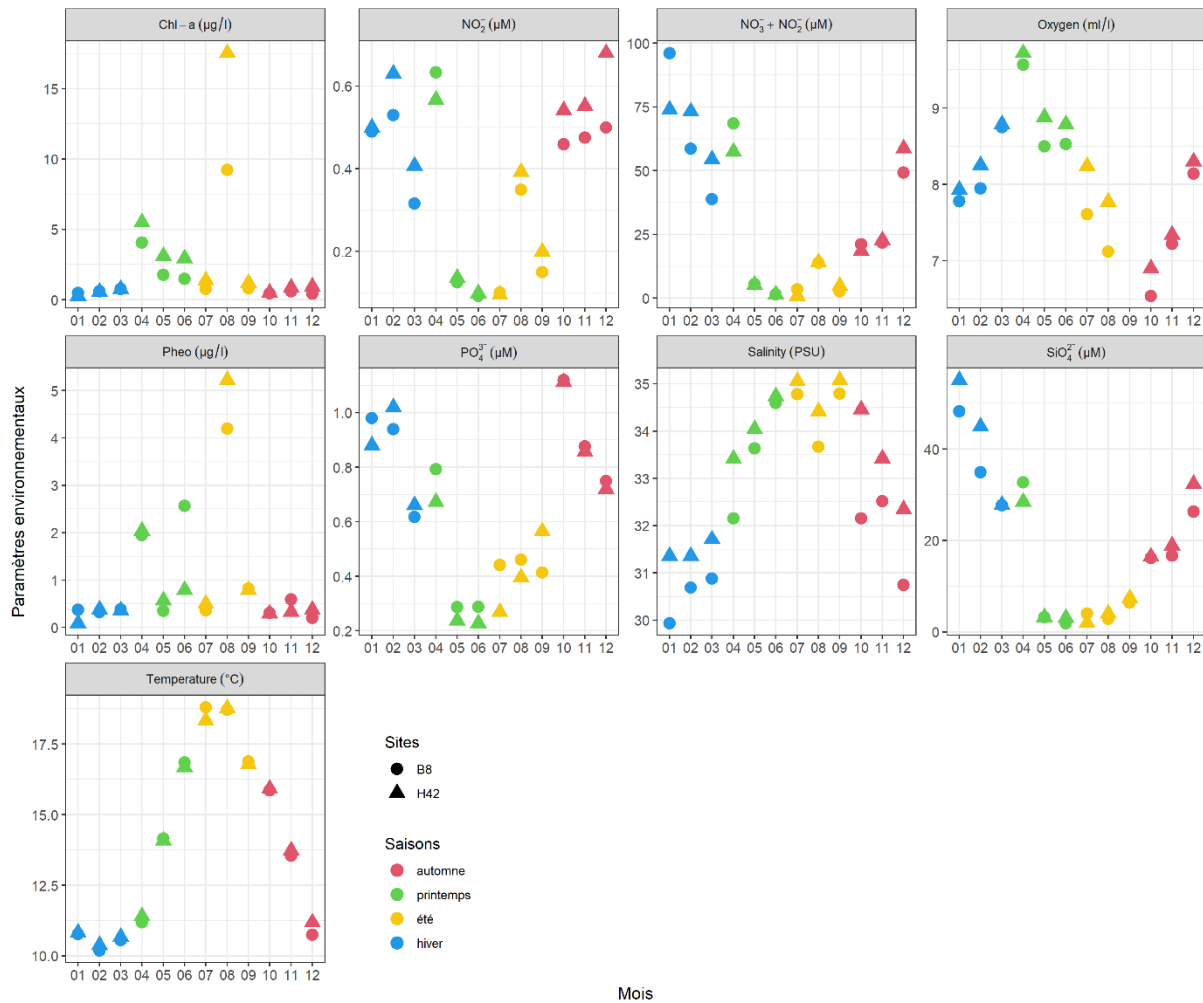


Figure 4. Variations mensuelles des paramètres environnementaux aux stations B8 et H42 de 2019 à 2020. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, NO_2^- : Nitrites, NO_3^- : Nitrates, Pheo : Phéophytine, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates.

Tableau 6. Estimation des différences entre le site B8 et le site H42 dans les valeurs mesurées de chaque paramètre environnemental à l'aide d'un test des rangs signés de Wilcoxon. En gras, les valeurs significatives.

Paramètres environnementaux	Médiane des différences x-y	p
Chl-a	-0.510	0.065
NO_2^-	-0.046	0.038
$NO_2^- + NO_3^-$	-0.099	0.774
Oxygène	-0.280	4.8e-4
Phéophytine	-0.020	1.000
PO_4^{3-}	0.039	0.146
Salinité	-0.789	4.8e-4
SiO_4^{2-}	-1.104	0.038
Température	-0.061	0.388

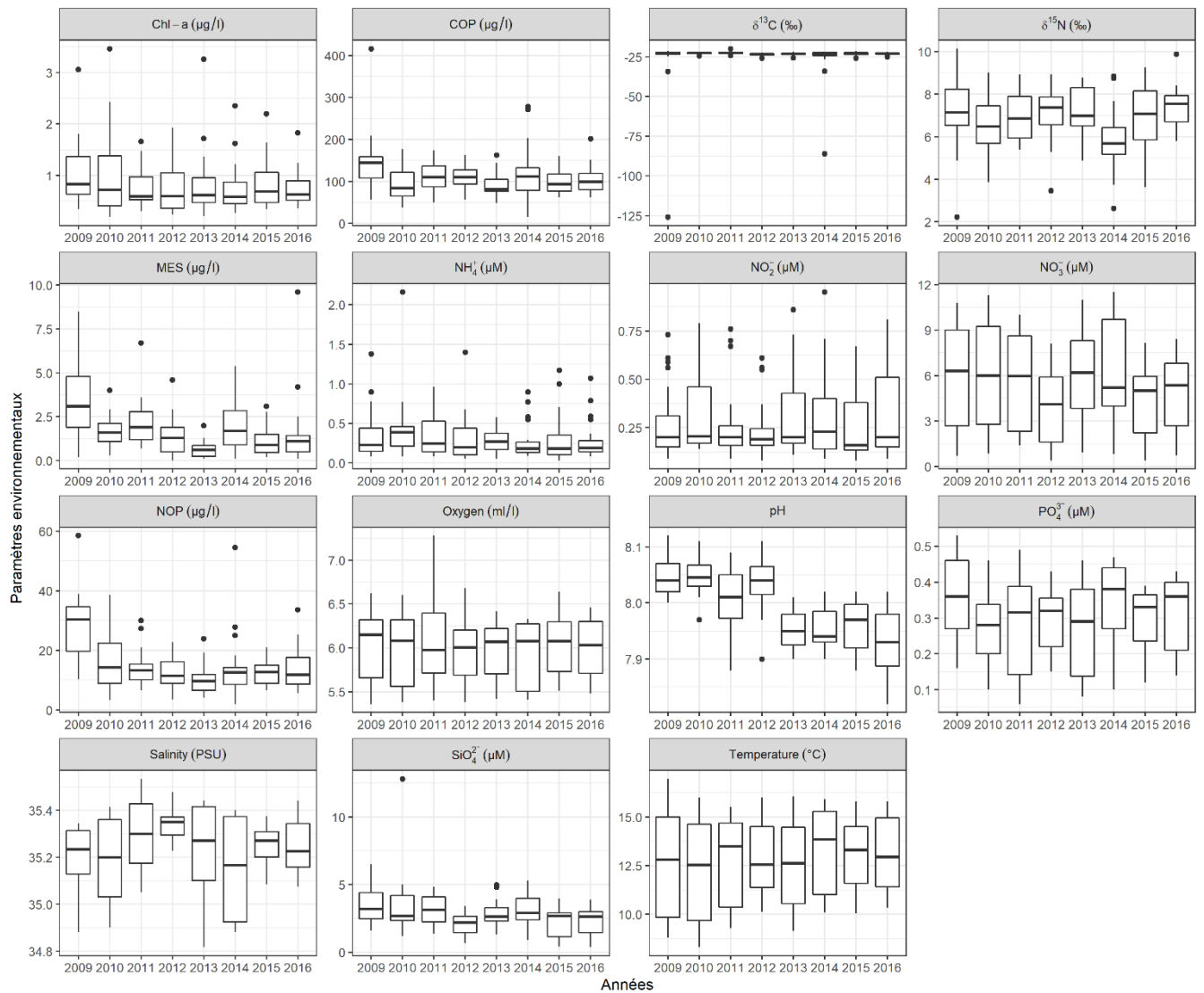


Figure 5. Variations annuelles des paramètres environnementaux à la station SOMLIT-Astan de 2009 à 2016. Les échantillonnages ont été réalisés aux grandes marées de mortes-eaux. Chl-a : Chlorophylle a, COP : Carbone organique particulaire, $\delta^{13}\text{C}$: variation relative du rapport $^{13}\text{C}/^{12}\text{C}$, $\delta^{15}\text{N}$: variation relative du rapport $^{15}\text{N}/^{14}\text{N}$, MES : Matières en suspension, NH_4^+ : Ammonium, NO_2^- : Nitrites, NO_3^- : Nitrates, NOP : Azote organique particulaire, PO_4^{3-} : Phosphate, SiO_4^{2-} : Silicates.

Tableau 7. Test de la tendance monotonique (test de Mann-Kendall) des séries temporelles des valeurs annuelles moyennes des différents paramètres environnementaux mesurés à la station SOMLIT-Astan de 2009 à 2016. En gras, les valeurs significatives.

Paramètres environnementaux	tau	p
Chl- <i>a</i>	-0.500	0.108
COP	-0.214	0.536
$\delta^{13}C$	0.143	0.710
$\delta^{15}N$	0.071	0.901
MES	-0.429	0.173
NH_4^+	-0.500	0.108
NO_2^-	0.286	0.386
NO_3^-	-0.071	0.901
NOP	-0.429	0.173
Oxygène	-0.143	0.710
pH	-0.714	0.019
PO_4^{3-}	0.071	0.901
Salinité	0.071	0.901
SiO_4^{2-}	-0.571	0.063
Température	0.571	0.063

6.2. Structures temporelles des communautés microbiennes planctoniques

6.2.1. Méthodologie

Pour rappel, les jeux de données de biodiversité étudiés dans le cadre du projet INDIGENE sont issus de séries temporelles correspondant à des sites situés dans des paysages hydrologiques distincts (les 3 stations SOMLIT ; Tew-Kai et al. 2020) ainsi qu'à des sites installés sur un gradient de pollution (Marina du Château). Pour chacune des stations, des comptages microscopiques et des données de metabarcoding sont disponibles et seront comparées dans les analyses ci-dessous.

La structure temporelle des communautés associées à chaque station d'échantillonnage a été étudiée suivant le cadre d'analyses statistiques présenté en Figure 5, qui permet de quantifier la part de variance expliquée par les facteurs environnementaux et celle attribuable à l'autocorrélation temporelle.

La stratégie choisie utilise une approche de modélisation multivariée de séries temporelles, qui décompose la structure temporelle globale des communautés en modèles temporels indépendants et spécifiques à différents groupes d'espèces (Figure 6, voir aussi Legendre & Gauthier 2014 et Caracciolo et al. 2022). La méthode est basée sur une analyse de redondance (RDA, implémentée sous la fonction *rda* dans le package-R *vegan*, Oksanen et al. 2019), avec comme variables réponses les abondances d'espèces (microscopie) ou les abondances relatives d'ASVs (metabarcoding) et comme variables explicatives des descripteurs de la

relation temporelle entre les échantillons. Les descripteurs temporels ont été obtenus en utilisant les cartes de vecteurs propres de Moran (MEMs, Dray et al. 2006) en utilisant la fonction *dbmem* du package-R *adespatial* (Dray et al. 2021). Une description détaillée de la méthode figure dans le rapport d'avancement N°2 du projet INDIGENE (Alric et al. 2021).

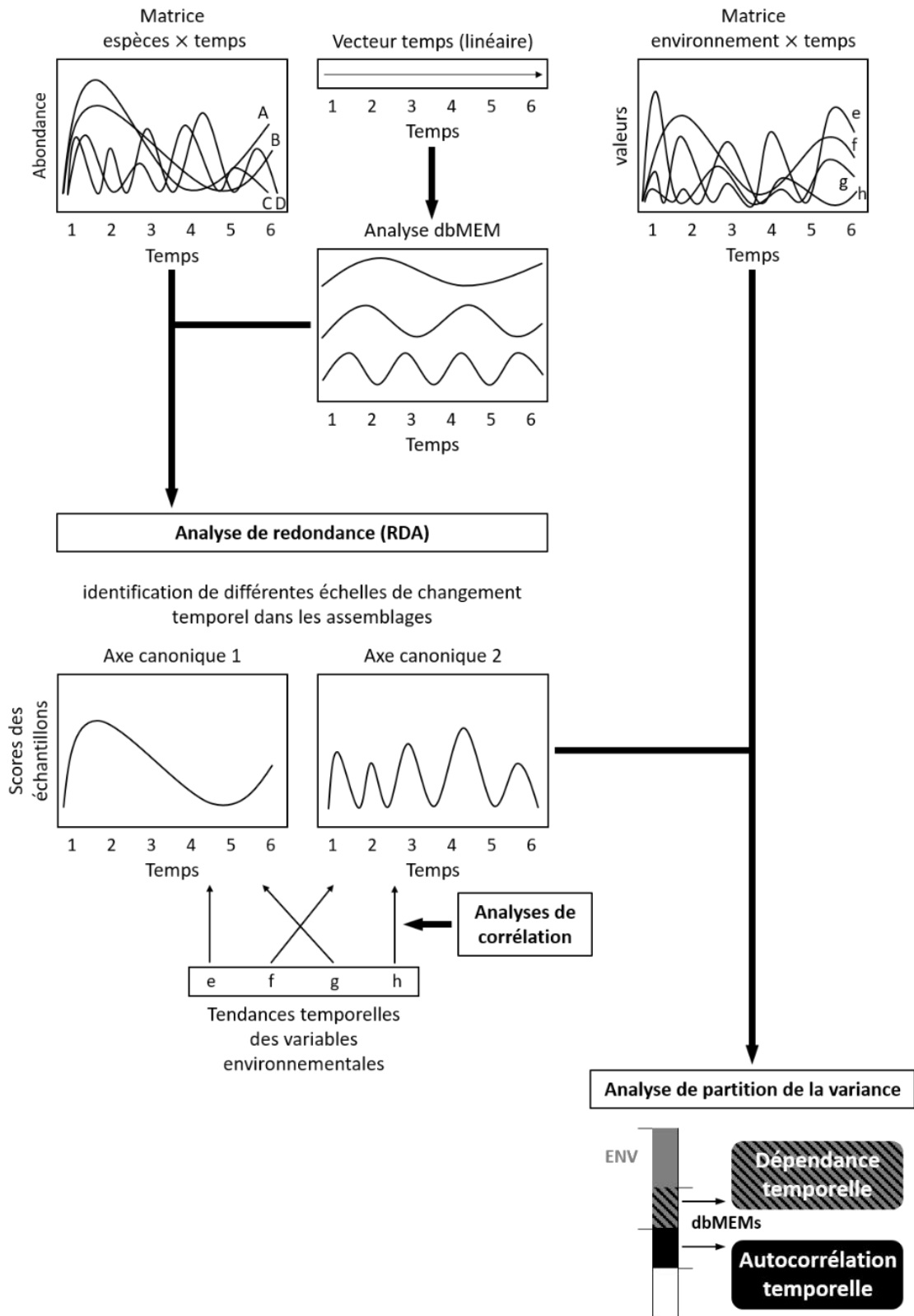


Figure 6. Schéma résumant les différentes étapes d'analyse en lien avec l'action 2 du projet.

6.2.2. Exemple pour la station SOMLIT-Astan à Roscoff

6.2.2.1. Analyse des compartiments phytoplanctoniques et des protistes (eucaryotes, fractions de taille > 3 µm et 0,2-3 µm)

Les modèles parcimonieux (i.e., RDA) basés sur les données de microscopie et de metabarcoding et les vecteurs propres dbMEMs mettent en évidence une dynamique temporelle dans la structure de la communauté de protistes entre 2009 et 2016. Pour les données de microscopie, comme pour les données de metabarcoding de la fraction de taille > 3 µm, les vecteurs propres retenus dans les analyses (dbMEM) expliquent respectivement 51 % et 61 % de la variance totale de la communauté (49 et 47 % après correction du R2 en prenant en compte le nombre de variables explicatives) respectivement pour les jeux de données de microscopie (phytoplancton > 10 µm) et de metabarcoding (protistes > 3 µm).

Les dynamiques temporelles modélisées par les deux premiers axes canoniques sont dans les deux cas caractéristiques d'une dynamique saisonnière de la structure de la communauté (Figures 6a et 7a). Les scores des échantillons sur les premiers axes canoniques de la RDA présentent une distribution sinusoïdale en fonction du temps d'une périodicité de 6 mois (maxima et minima relatifs estivaux et hivernaux pour l'axe 1 et maxima et minima relatifs printaniers et automnaux pour l'axe 2, Figures 7b-c et 8b-c).

La proportion plus faible de variance totale expliquée avec les données de metabarcoding peut s'expliquer en partie par le nombre plus grand de taxons inclus dans le modèle parcimonieux (1176 ASVs), comparé au modèle avec les données de microscopie (93 espèces), ce qui peut entraîner un bruit de fond. Toutefois, le cycle annuel est mieux caractérisé avec les données de metabarcoding avec une variabilité intra-saison moins importante (Figure 8a). En effet, la distribution sinusoïdale des scores des échantillons le long des deux premiers axes canoniques de la RDA en fonction du temps présente une dynamique plus monotone (Figures 8b et c).

Des résultats similaires ont été obtenus pour les communautés de protistes de la fraction de taille 0,2-3 µm (pourcentage de variance expliquée de 65.96%, corrigée à 50.41%, Figure 9) pour lesquels les méthodes de microscopie sont inopérantes. Il est ainsi intéressant de noter que les dynamiques sont très similaires à celles observées pour les plus grandes fractions de taille.

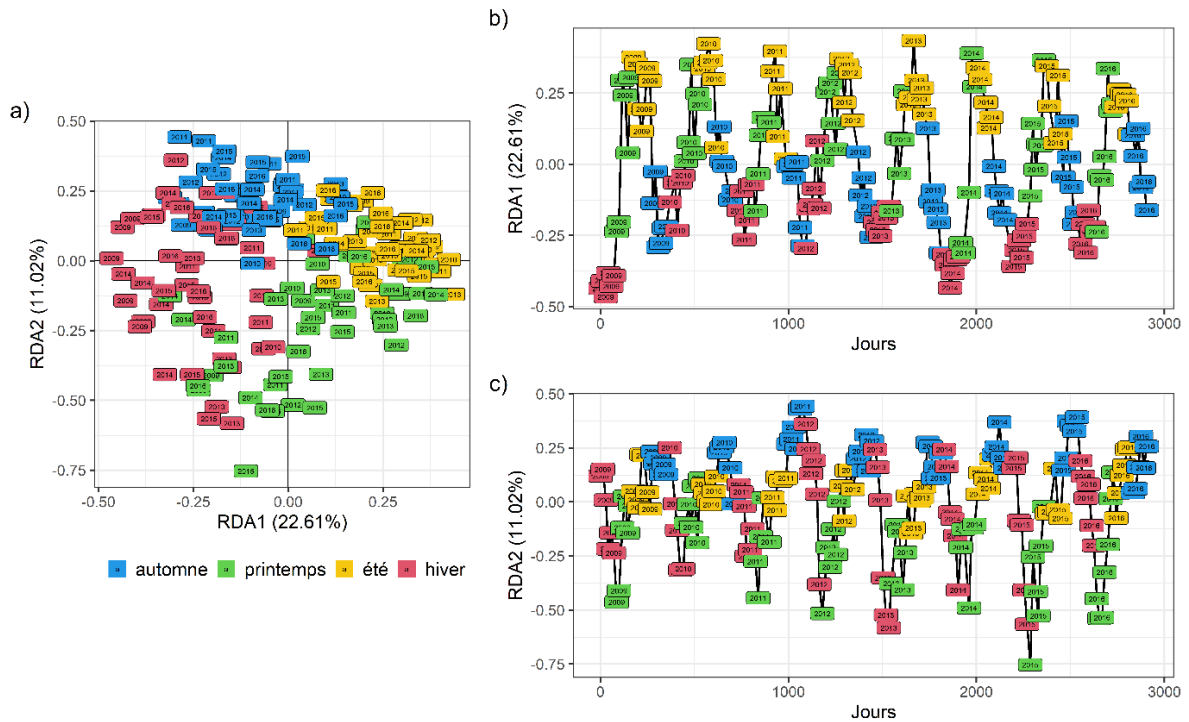


Figure 7. Similarités des communautés phytoplanctoniques (analyse RDA) dans les échantillons mensuels de 2009 à 2016 à la station SOMLIT-Astan pour les données de **microscopie**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

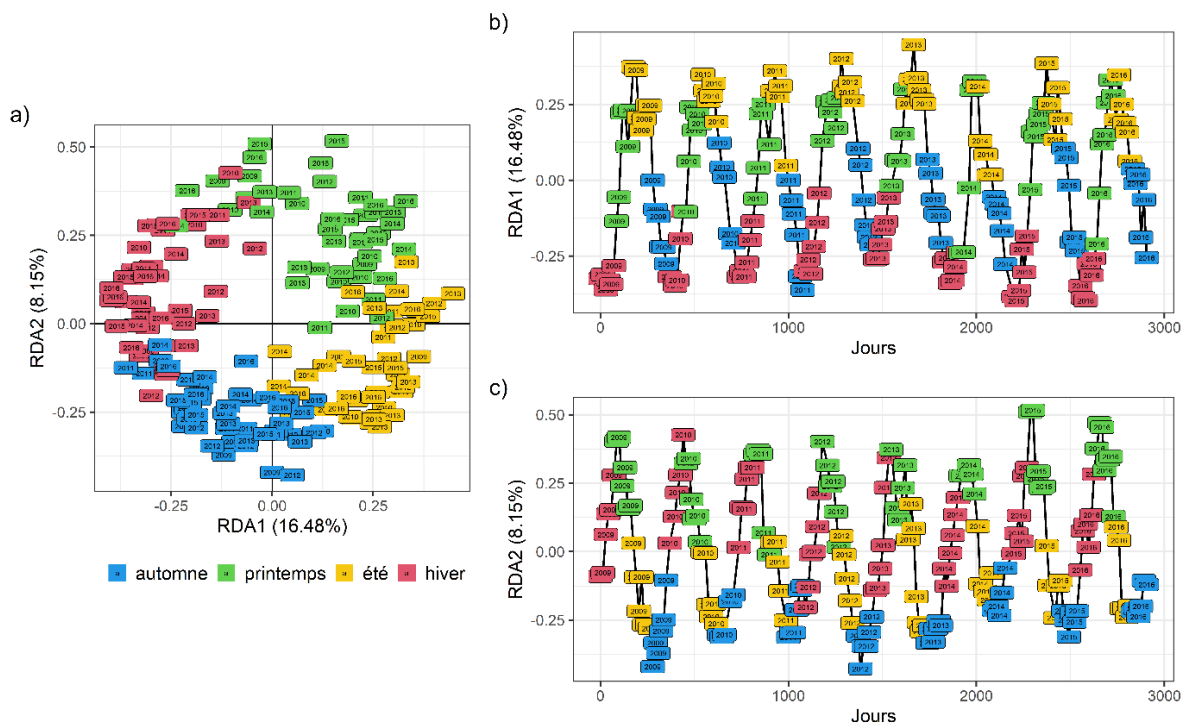


Figure 8. Similarités des communautés de protistes (analyse RDA) dans les échantillons mensuels de 2009 à 2016 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

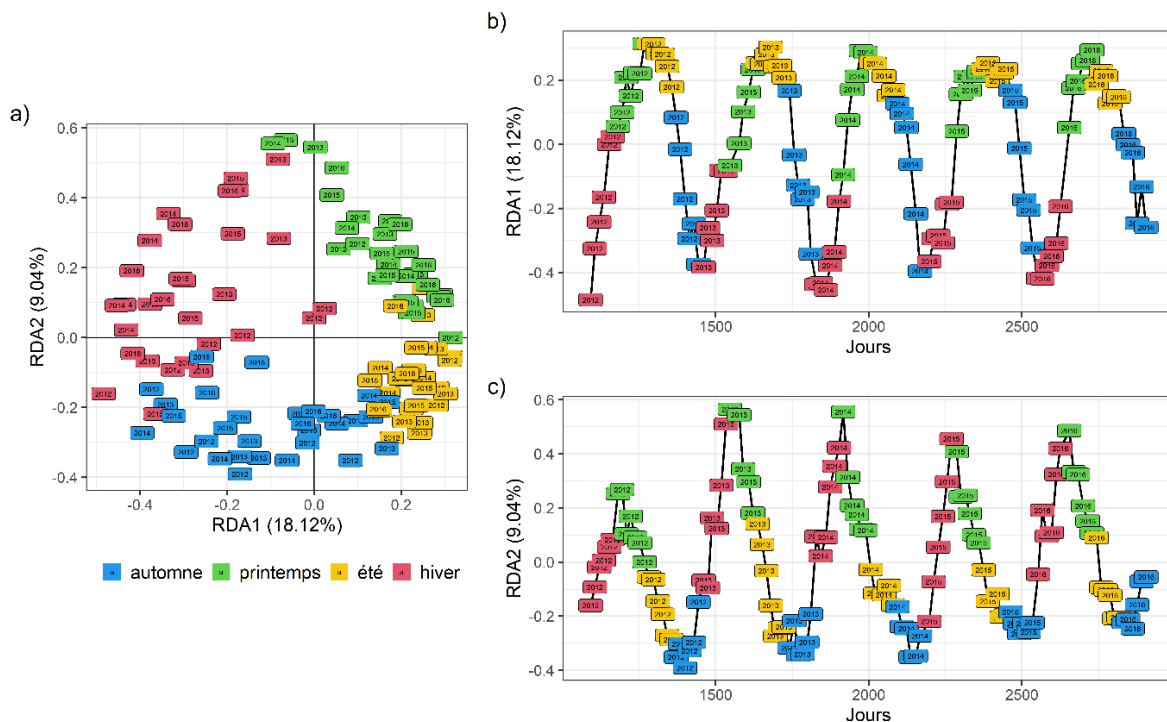


Figure 9. Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2012 à 2016 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 18S, fraction de taille 0.2-3 µm)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

6.2.2.2. Analyse des communautés de procaryotes

De manière analogue à la communauté de protistes, la communauté de procaryotes, quelle que soit la fraction de taille considérée (i.e., 0.2-3 µm ou >3 µm), présente un cycle annuel dans sa structure. Ce cycle annuel est expliqué significativement ($F = 5.57$, $p = 0.001$) à hauteur de 63.61% (corrigé à 52.19%) par 16 (sur 20) vecteurs propres retenus par sélection forward pour la communauté de procaryotes associée à la fraction 0.2-3 µm et significativement ($F = 4.81$, $p = 0.001$) à hauteur de 63.26% (corrigé à 50.10%) par 24 (sur 33) vecteurs propres pour la communauté de procaryotes associée à la fraction >3 µm. Les cycles annuels dans la structure de la communauté de procaryotes associés aux deux classes de taille (Figures 10 et 11) se caractérisent par un regroupement des dates en fonction de la saison sur le plan factoriel formé par les deux premiers axes canoniques de la RDA.

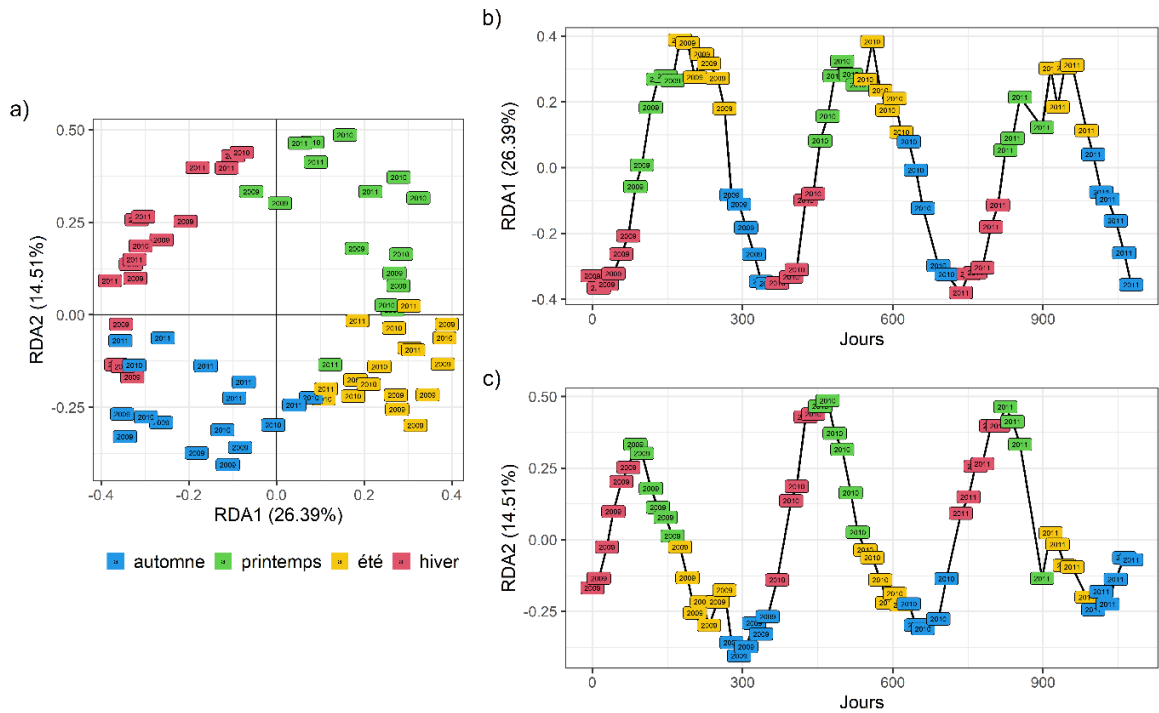


Figure 10. Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2009 à 2011 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 16S, fraction de taille 0.2-3 µm)**. a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

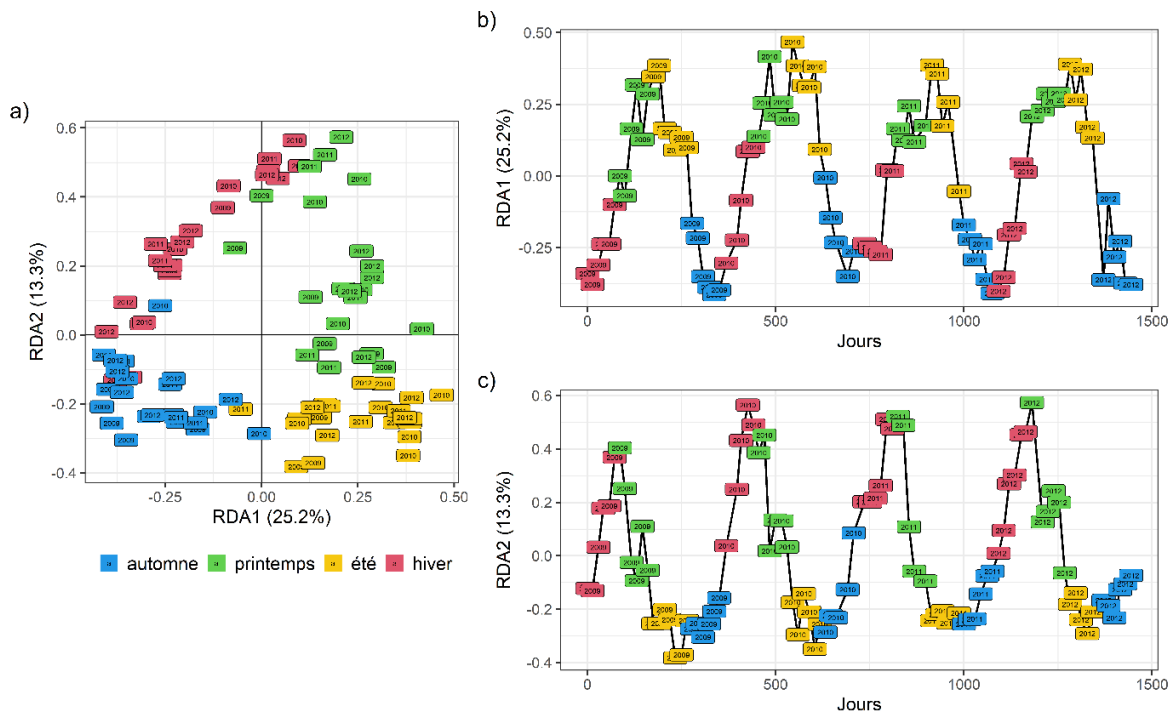


Figure 11. Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2009 à 2012 à la station SOMLIT-Astan pour les données de **metabarcoding (ARNr 16S, fraction de taille >3 µm)**. a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

6.2.3. Exemple pour les stations B8 et H42 à la Marina du Château dans le port de Brest

Le modèle parcimonieux basé sur les données de microscopie et sur 5 des 6 vecteurs propres explique un faible pourcentage (3.72 % après correction) de la variance totale de la communauté de protistes et ne met pas en évidence de dynamique temporelle significative ($F = 1.16$, $p = 0.224$) dans la structure de cette communauté entre 2019 et 2020 (Figure 12a). Cela peut s'expliquer par le fait que l'on ne dispose que d'une année de suivi et de peu de données de microscopie. Notons toutefois que les communautés printanières et estivales se distinguent des communautés automnales et hivernales en terme de composition sur le plan factoriel formé par les deux premiers axes canoniques de la RDA (Figures 12b et c).

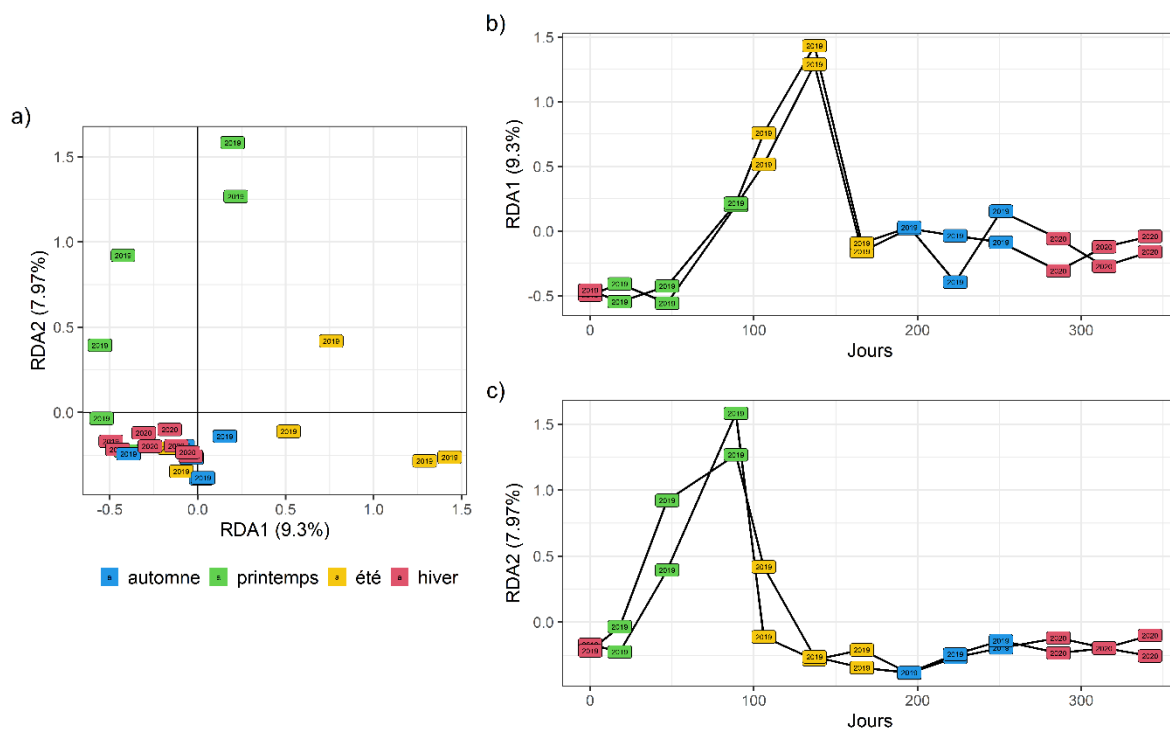


Figure 12. Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **microscopie**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

A l'inverse, le modèle parcimonieux basé sur les données de metabarcoding (ARNr 18S, fraction de taille > 3 μ m) et sur 5 des 6 vecteurs propres explique un pourcentage plus élevé (47.23% après correction) de la variance totale de la communauté de protistes et met en évidence une dynamique temporelle significative ($F = 4.73$, $p = 0.001$) dans la structure de cette communauté entre 2019 et 2020 (Figure 13a). Nous retrouvons un cycle annuel comparable à celui observé pour la communauté de protistes de la grande fraction de taille à la station SOMLIT-Astan, mais moins marqué (Figures 13b et c).

Nous retrouvons aussi un cycle annuel (Figure 14) dans les changements de structure de la communauté de protistes de la petite fraction de taille (ARNr 18S, fraction 0.2-3 μ m). Le modèle parcimonieux basé sur 4 (sur 5) vecteurs propres explique significativement ($F = 3.19$, $p = 0.001$) 31.32% (après correction) de la variance totale de cette communauté.

Un patron similaire, dans la dynamique temporelle de la structure de la communauté de procaryotes est également à noter (Figure 15). Toutefois, le modèle parcimonieux basé sur 4 (sur 5) vecteurs propres n'explique significativement ($F = 2.83$, $p=0.001$) que 27.66% (après correction) de la variance totale de cette communauté.

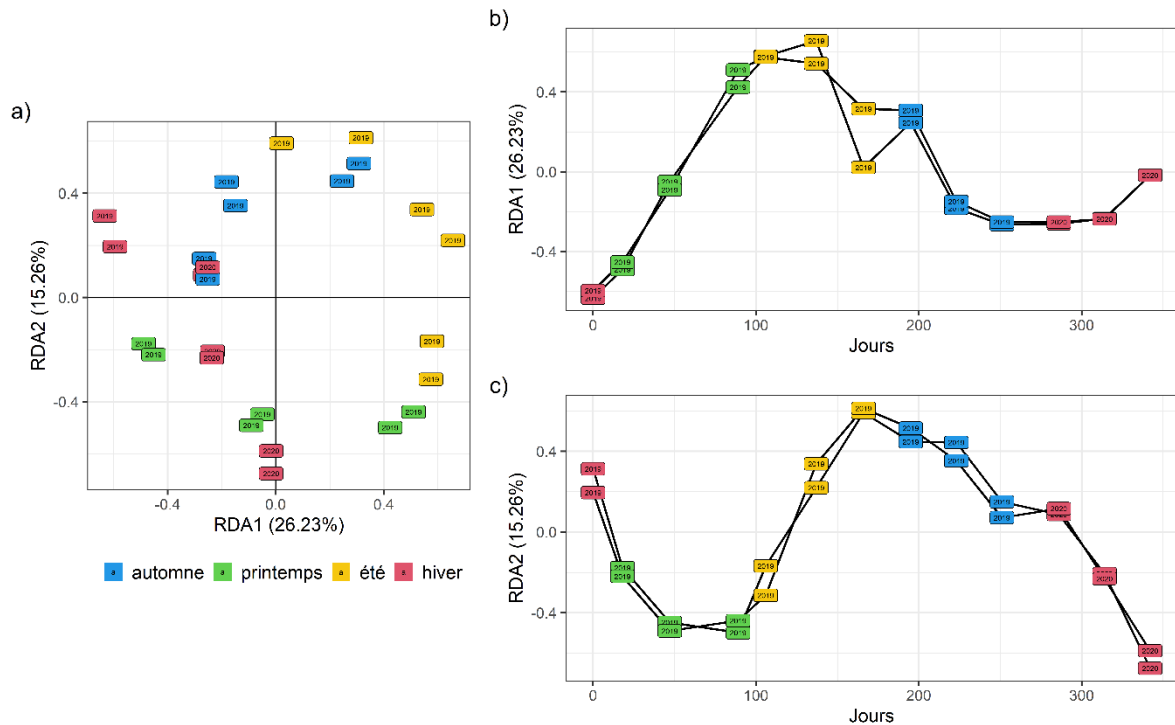


Figure 13. Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

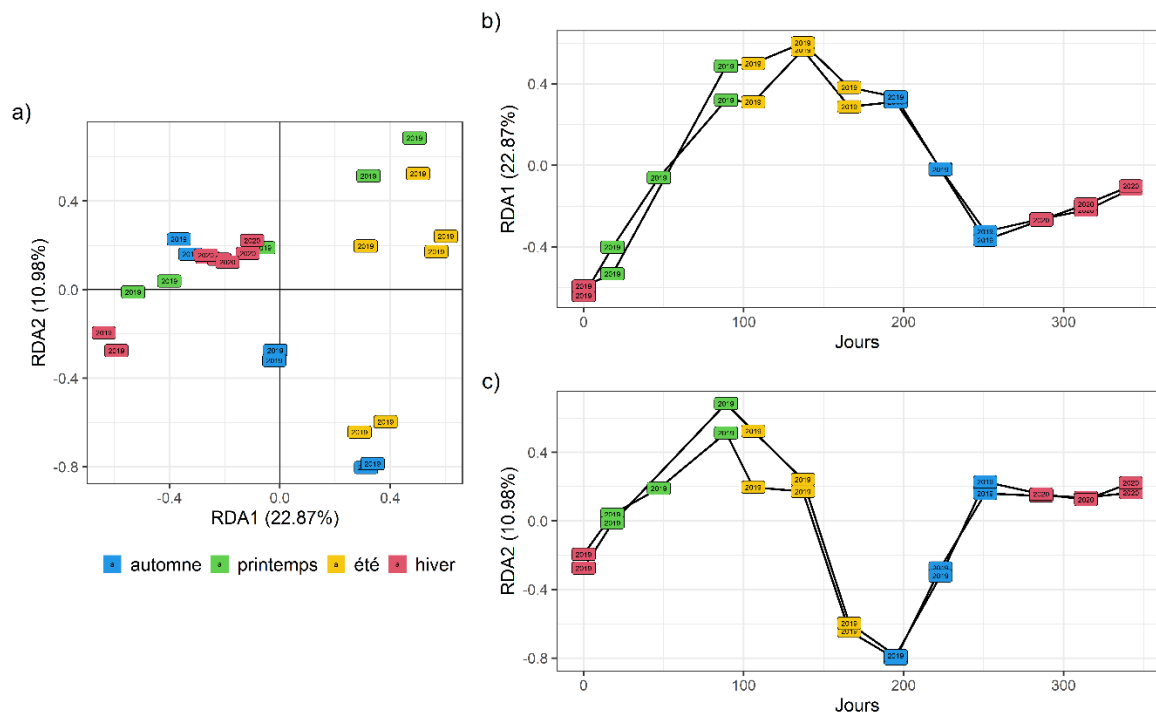


Figure 14. Similitude de la communauté de protistes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding (ARNr 18S, fraction de taille 0.2-3 μ m)**. a) Cycle annuel de la communauté de protistes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

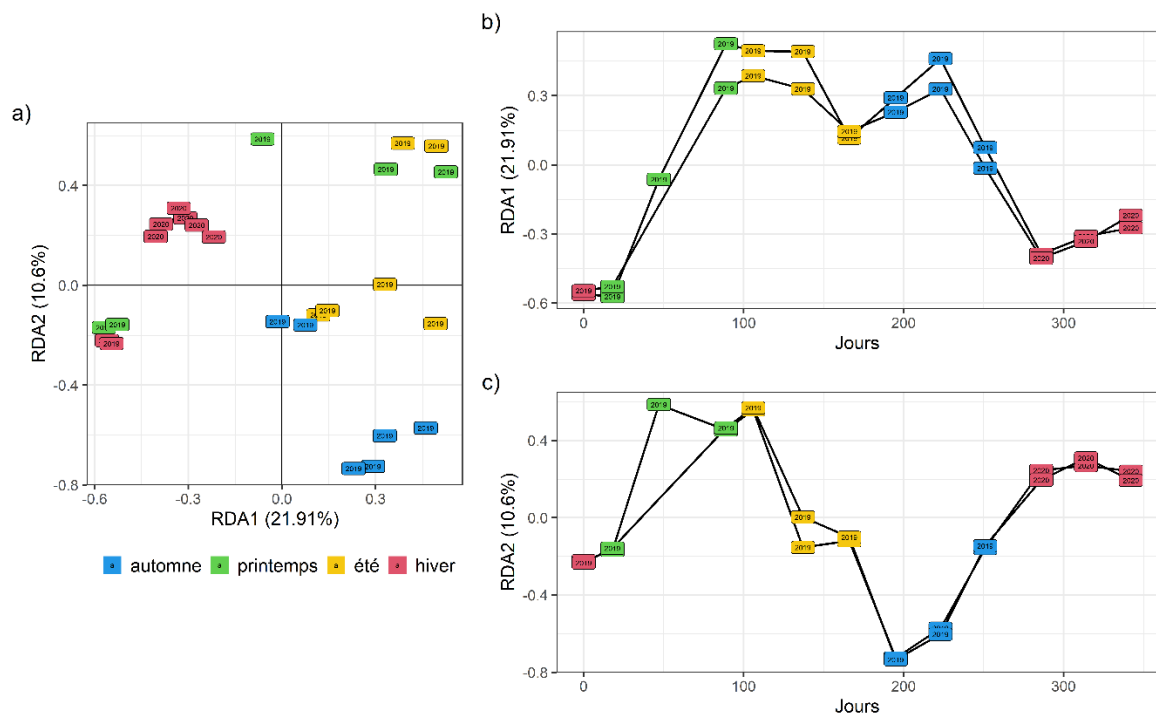


Figure 15. Similitude de la communauté de procaryotes (analyse RDA) dans les échantillons mensuels de 2019 à 2020 aux stations B8 et H42 localisées à la Marina du Château dans le port de Brest pour les données de **metabarcoding (ARNr 16S, fraction de taille 0.2-3 μ m)**. a) Cycle annuel de la communauté de procaryotes obtenu par la projection des scores des échantillons mensuels sur le plan factoriel de la RDA. b) et c) Décomposition des axes de la RDA révélant un modèle saisonnier.

6.2.4. Conclusions sur les enveloppes de variation des communautés microbiennes

Le travail réalisé permet **de confirmer l'importance de la prise en compte des patrons temporels saisonniers pour analyser les communautés microbiennes procaryotes et eucaryotes des habitats pélagiques marins**, quelles que soient les fractions de taille analysées ainsi que la nature des données utilisées (i.e., microscopie, metabarcoding).

Les données de metabarcoding permettent d'accéder aux patrons de diversité des procaryotes et des protistes de petite taille et les enveloppes de variations temporelles obtenues à partir de ces données apparaissent mieux caractérisées que celles obtenues par les données de comptage.

7. Action 3 : Test des indicateurs de composition sur des données génétiques, liens avec les pressions

L'intégration du compartiment microbien dans l'évaluation et la surveillance des habitats pélagiques marins nécessite de réaliser des tests numériques afin de déterminer si les données de metabarcoding peuvent être utilisées pour le calcul des indicateurs de changement (*PH1* et *PH3*) ou si des adaptations sont à apporter sur ces derniers.

7.1. Méthodologie sur le calcul des indicateurs de changements *PH1* et *PH3*

Les indicateurs *PH1* et *PH3* sont actuellement des indicateurs de changements d'état des communautés écologiques (qualifiés aussi « d'indicateurs de surveillance ») et non des indicateurs renseignant de l'atteinte ou non du BEE par les habitats pélagiques marins. Pour aboutir à des indicateurs BEE opérationnels, des développements conceptuels et numériques ont été mis en place lors du 2^{ème} cycle d'évaluation du descripteur « Habitats pélagiques » afin d'améliorer l'opérationnalité méthodologiques des indicateurs. Ceux-ci consistent notamment à préciser la période de référence (pour l'indicateur *PH1*) et à définir des seuils éventuels pour chaque indicateur (i.e., *PH1* et *PH3*), en lien avec les données de pressions environnementales.

7.1.1. Indicateur *PH1*

L'indicateur *PH1* renseigne sur l'évolution temporelle de deux formes de vie (ou groupes fonctionnels) et permet de détecter les réponses des communautés planctoniques aux modifications des conditions environnementales, ceci en quantifiant l'écart entre une succession d'états observés et une succession d'états de référence. L'obtention d'un indicateur *PH1* robuste repose sur la définition d'une période de référence pertinente pour son opérationnalité et son interprétation en vue d'une évaluation du BEE. Idéalement, la période de référence doit tenir compte d'un ensemble d'observations dans un écosystème marin donné dont on sait avec certitude qu'il est en BEE. Une analyse comparative de la réponse des indicateurs aux échelles temporelles d'évaluation considérées a montré l'intérêt de prendre en compte des échelles contemporaines récentes (e.g., 6 ans avant la période d'évaluation correspondant à la durée d'un cycle d'évaluation) pour répondre aux besoins des directives (Bedford et al. 2020). Dans notre cas, les séries temporelles de données de chaque station étaient trop courtes (inférieures à 5 ans pour certaines) ou ne couvraient pas l'intégralité de la période d'évaluation antérieure, pouvant être utilisée comme période de

référence dans le calcul de l'indicateur. Selon les recommandations de Duflos et al. (2018), le cycle de référence a donc été calculé sur une période minimale de 3 ans.

L'indicateur *PH1* se calcule de la manière suivante :

$$PH1 = \frac{\text{nouveaux points dans et hors de l'enveloppe de référence}}{\text{nombre total de nouveaux points}} \quad (1)$$

La valeur du *PH1* correspond à la proportion de points de l'enveloppe d'un état observé compris dans l'enveloppe de référence. Un *PH1* de 1 correspond à un état identique à la référence, et un *PH1* de 0 correspond à un changement total de l'état. Dans l'Intermediate Assessment d'OSPAR 2017, un seuil de 10% a été donné pour définir la proportion de points qui sont tolérés en dehors de l'enveloppe de référence. Pour déterminer si l'état observé est significativement différent de l'état de référence, un test du Khi-deux (unilatéral avec un degré de liberté de 1) est effectué. Le package « PI2016_Oct_beta » et le guide associé (Tett 2016) ont été utilisés pour effectuer le regroupement des espèces en groupes fonctionnels, les calculs de ratios, la détermination de l'enveloppe de référence et les tests statistiques nécessaires au calcul de l'indicateur *PH1*. Le script d'analyse étant développé sous MATLAB, une plateforme de calcul numérique nécessitant une licence, et non disponible en langage Open Source (type R) rend difficile son utilisation. Nous avons pu calculer l'indicateur *PH1* grâce à l'accès à une licence MATLAB via le portail de Sorbonne Université.

7.1.2. Indicateur *PH3*

L'indicateur *PH3* est un indicateur multi-métrique reposant sur le calcul d'indices de diversité des communautés planctoniques permettant de décrire les changements dans la structure et la diversité des communautés planctoniques le long d'un continuum temporel. L'approche est basée sur l'utilisation combinée d'indices de diversité alpha (i.e., l'indice de Hulbert et l'indice de Menhinick) et de diversité bêta (l'indice LCBD).

L'indice d'Hulbert se calcule de la manière suivante :

$$\delta = 100 \times (n_1 + n_2) / N \quad (2)$$

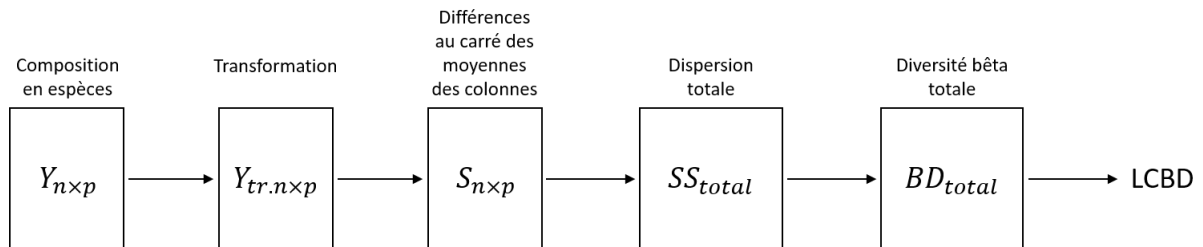
où *N* correspond à l'abondance de toutes les espèces de l'échantillon et *n*₁ et *n*₂ sont respectivement l'abondance des deux espèces les plus abondantes dans l'échantillon. Afin de relier l'indice d'Hulbert à la qualité écologique de l'eau, la valeur de 100 – δ est utilisée, conformément à la théorie classique selon laquelle des phénomènes de dominance et des changements dans la composition de la communauté se produisent dans les zones impactées par des changements dans les conditions environnementales (Howarth et al. 2000, Facca et al. 2014).

L'indice de Menhinick est estimé de la manière suivante :

$$D = \frac{S}{\sqrt{N}} \quad (3)$$

où *S* est le nombre total d'espèces de l'échantillon et *N* est l'abondance totale de toutes les espèces dans l'échantillon.

La Contribution Locale à la Bêta Diversité (indice LCBD ; Legendre P. & De Cáceres 2013), une mesure de la singularité d'un échantillon (e.g., une date au sein d'une série temporelle), se calcule en transformant les données d'abondance puis en effectuant la suite des calculs présentés dans le schéma ci-dessous :



L'intégration des trois indices au niveau de l'indicateur PH3 nécessite une transformation des valeurs brutes de chaque indice, selon l'estimation d'EQRs.

Les EQRs sont déterminés pour chaque indice de la façon suivante :

$$EQR = \frac{\text{valeur de l'indice pour la période d'évaluation}}{\text{valeur de référence}} \quad (4)$$

pour les indices de Menhinick et 100 – Hulburt,

$$EQR = \frac{\text{valeur de référence}}{\text{valeur de l'indice pour la période d'évaluation}} \quad (5)$$

pour l'indice LCBD.

Duflos et al. (2018) ont proposé d'estimer des EQRs associés à une grille de lecture (Figure 16).

Très éloigné des conditions de référence	Assez éloigné	Intermédiaire	Assez proche	Très proche des conditions de référence
0 ≤ EQR ≤ 0,2 richesse faible, dominance forte, structure atypique	0,21 ≤ EQR ≤ 0,4	0,41 ≤ EQR ≤ 0,6	0,61 ≤ EQR ≤ 0,8	0,81 ≤ EQR ≤ 1 richesse forte, dominance faible, structure proche de la structure de référence.

Figure 16. Proposition de seuils pour l'évaluation de l'indicateur PH3 (d'après Duflos et al. 2018).

La grille de lecture actuelle a été estimée de manière empirique, avec des valeurs allant de 0 à 1, en faisant l'hypothèse qu'un EQR égal à 1 reflète des conditions similaires à celles qui ont été définies par défaut comme référence pour les communautés planctoniques. Cette grille permet donc de définir cinq classes de qualité allant d'une communauté planctonique très éloignée (en termes de composition) des conditions de référence ($0 \leq EQR \leq 0.2$) à une communauté planctonique très proche des conditions de référence ($0.81 \leq EQR \leq 1$). Le calcul des EQRs ainsi que la délimitation des bornes de la grille de lecture nécessitent d'être approfondis et validés au cours du 2^{ème} cycle d'évaluation. **A ce titre, nous proposons un**

ajustement méthodologique pour l'estimation des valeurs de référence nécessaires au calcul des EQRs. A l'origine, pour un type de masse d'eau donné, la valeur la plus élevée des indices Menhinick et $100 - \text{Hulburt}$ et la valeur la plus faible pour l'indice LCBD, prises sur l'ensemble de la série temporelle, étaient choisies comme valeurs de référence par défaut (Duflos et al. 2018). Selon les résultats obtenus dans l'action 2 du projet et présentés dans la section 6.2 de ce rapport, **il nous a semblé important de considérer la saisonnalité dans le calcul des valeurs de référence.** Ainsi pour chaque station, la série temporelle de chaque indice a été partitionnée selon les quatre saisons (i.e., hiver, printemps, été, automne) et une valeur de référence a été déterminée pour chacune d'elles. Ces valeurs de référence ont ensuite été utilisées pour calculer les EQRs de chaque indice à l'échelle saisonnière (Figure 17). Concernant l'estimation des bornes de la grille de lecture, ce point n'a pas été abordé dans le cadre de ce rapport. Toutefois, des tests préliminaires ont été réalisés afin d'estimer des seuils de changement des communautés microbiennes en lien avec chacun des paramètres environnementaux disponibles. Bien que les résultats ne soient pas concluants pour l'instant, l'approche utilisée basée sur la méthode TITAN (Baker & King 2010) semble être prometteuse et pourrait être une piste à approfondir à l'avenir.

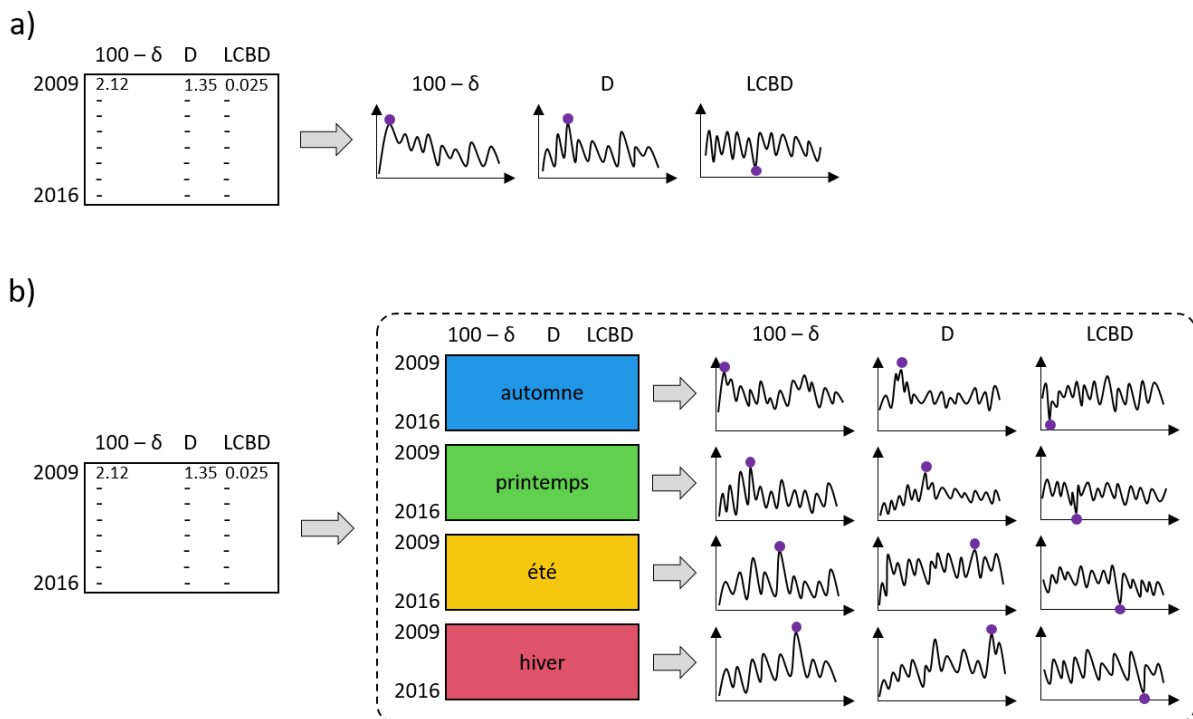


Figure 17. Représentation schématique de la méthode d'estimation des valeurs de référence utilisées dans le calcul des EQRs des trois métriques constitutives de l'indice PH3. a) Approche proposée par Duflos et al. (2018) dans le cadre du 1^{er} cycle d'évaluation. b) Approche proposée et utilisée dans le rapport. δ : Hulburt, D : indice de Menhinick, LCBD : Contribution Locale à la Bêta Diversité. Les points violets correspondent aux valeurs de référence pour chacune des métriques (i.e., valeurs maximales pour $100 - \delta$ et D et valeurs minimales pour le LCBD).

L'intégration finale des trois indices au niveau de l'indicateur *PH3* se réalise par le calcul de la moyenne de leurs EQR :

$$PH3 = \sum [EQR(D), EQR(100 - \delta), EQR(LCBD)] \quad (6)$$

où *D* est l'indice de Menhinick, δ est l'indice de Hulburt et LCBD est la Contribution Locale à la Diversité Bêta.

7.1.3. Fréquence d'échantillonnage

Un autre aspect méthodologique soulevé pour le 2^{ème} cycle d'évaluation est de tester la sensibilité des indicateurs à la fréquence d'acquisition des données. Il s'agit de comparer les valeurs d'indicateurs calculés à partir de séries temporelles dégradées (e.g., en ne considérant que certaines dates d'échantillonnage) par rapport à des séries temporelles complètes. En pratique, la sensibilité de l'indicateur *PH3* à la fréquence d'acquisition des données a été testée, pour chaque station, par ré-échantillonnage aléatoire des dates des séries temporelles selon six scénarios prenant en compte un mois par saison, deux mois par saison, 11 mois par an, une saison par an, deux saisons par an, et trois saisons par an (Figure 18).

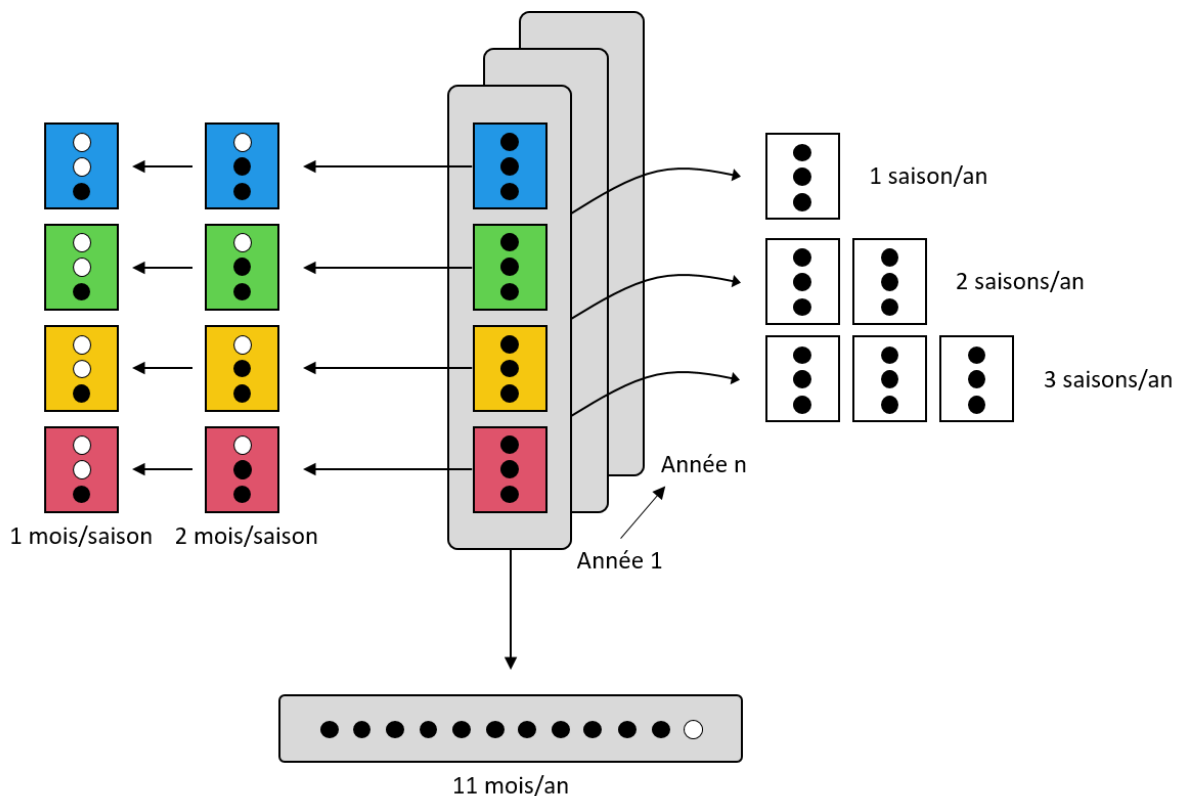


Figure 18. Représentation schématique de la procédure de dégradation des séries temporelles pour estimer la sensibilité de l'indicateur *PH3* à la fréquence d'acquisition des données. Six scénarios de dégradations ont été considérés, en réalisant un ré-échantillonnage aléatoire des données des séries temporelles. Des tirages aléatoires ont été réalisés : 1) un mois par saison, 2) deux mois par saison, 3) 11 mois par an, 4) une saison par an, 5) deux saisons par an, et 6) trois saisons par an. A partir de ces ré-échantillonnages, l'indicateur *PH3* a été calculé et comparé ensuite aux valeurs obtenues à partir de séries temporelles complètes.

7.1.4. Liens indicateur PH3 et variables environnementales

Afin de passer d'indicateurs de changement d'état des communautés écologiques (en l'occurrence le PH3) vers des indicateurs BEE opérationnels, il est nécessaire **d'évaluer également le lien entre la dynamique temporelle de ces indicateurs et celle des paramètres environnementaux, susceptibles de varier notamment sous l'effet de pressions anthropiques**. Les données de séries temporelles contiennent des informations non seulement sur les valeurs des données, mais aussi sur l'ordre de ces valeurs. De nombreuses méthodes d'analyse regroupent ces données en points temporels discrets ou en statistiques descriptives globales, dans l'idée de supprimer toute structure temporelle. Les données d'écologie sont fondamentalement spatiales et/ou temporelles (Wolkovich et al. 2014). Par conséquent, de nouvelles méthodes d'analyse des séries temporelles développées dans d'autres domaines pourraient permettre un traitement plus nuancé de ce type de données en écologie, réduisant ainsi la perte d'informations dans les données de séries temporelles en raison de l'agrégation. Une récente étude a mis en évidence l'intérêt de la méthode de déformation dynamique temporelle (Dynamic Time Warping, DTW) dans le traitement de données écologiques de séries temporelles (Hegg et al. 2021). La méthode DTW, inspirée de la distance de Levenshtein, et également appelée distance d'édition, a été introduite initialement dans le domaine de la reconnaissance vocale (Sakoe & Chiba 1971, Sakoe & Chiba 1978). Cette méthode est un moyen de comparer deux séquences temporelles qui ne se synchronisent pas parfaitement. Elle permet de trouver l'alignement global optimal entre deux séquences, en déformant l'axe du temps de manière itérative afin d'associer chaque élément de chaque séquence à au moins un élément de l'autre séquence et ceci en minimisant les coûts d'association. Le coût d'une association correspond à la distance entre les deux éléments. Supposons deux séries temporelles, Q et C , de longueur n et m respectivement :

$$Q = [q_1, q_2, \dots, q_i, \dots, q_n] \quad (7)$$

$$C = [c_1, c_2, \dots, c_j, \dots, c_m] \quad (8)$$

Pour aligner ces deux séquences à l'aide de l'algorithme DTW, une matrice $n \times m$ est d'abord construite (Figure 19).

Chaque élément $p_s = (i, j)$ de la matrice correspond à un alignement de q_i et c_j . Un alignement est doté d'un coût de déformation par une fonction de distance $d(q_i, c_j)$, qui dans ce cas correspond à une distance euclidienne :

$$d(q_i, c_j) = (q_i - c_j)^2 \quad (9)$$

L'algorithme DTW estime alors un chemin P à travers la matrice qui minimise la somme de cette distance. Le chemin P peut être défini comme :

$$P = [p_1, p_2, \dots, p_k] \quad (10)$$

où chaque élément p_k correspond à un point $(i, j)_k$ de la matrice.

Le chemin doit alors minimiser la fonction de coût :

$$DTW(Q, C) = \min \left\{ \sqrt{\sum_{i=1}^K p_k} \right. \quad (11)$$

et doit respecter les conditions suivantes :

1. Condition limite : cette contrainte garantit que le chemin de déformation commence par les points de départ des deux séquences et se termine par leurs points terminaux.

$$i_1 = 1, i_k = n \text{ et } j_1 = 1, j_k = m \quad (12)$$

2. Condition de monotonie : cette contrainte préserve l'ordre temporel des points (sans remonter dans le temps).

$$i_{t-1} \leq i_t \text{ et } j_{t-1} \leq j_t \quad (13)$$

3. Condition de continuité : cette contrainte limite les transitions de chemin vers des points adjacents dans le temps (sans sauter dans le temps).

$$i_t - i_{t-1} \leq 1 \text{ et } j_t - j_{t-1} \leq 1 \quad (14)$$

4. Etat de la fenêtre de déformation (w) : c'est une contrainte optionnelle à partir de laquelle les points admissibles peuvent être limités pour se situer dans une fenêtre de déformation de largeur donnée (un entier positif).

$$\forall p_k \in P, |i_t - j_t| \leq w \quad (15)$$

Afin de trouver le chemin optimal, la technique de programmation dynamique suivante peut être utilisée. A partir du point $(1, 1)$, la distance cumulée Δ à chaque point de la matrice est calculée par l'équation réursive suivante :

$$\Delta(i, j) = d(q_i, c_j) + \min \begin{cases} \Delta(i-1, j-1), \\ \Delta(i, j-1), \\ \Delta(i-1, j) \end{cases} \quad (16)$$

Une fois que toutes les distances cumulées sont calculées, le chemin de déformation optimal peut être trouvé en commençant au point (n, m) et en revenant en arrière dans la matrice, en prenant la plus petite valeur de distance à chaque fois et comprise dans la fenêtre de déformation.

Dans la matrice $(n \times m)$, chaque mouvement a sa propre signification :

- Un mouvement horizontal représente la suppression. Cela signifie que la séquence Q a accéléré par rapport à la séquence C pendant cet intervalle de temps.
- Un déplacement vertical représente une insertion. Cela signifie que la séquence Q a décéléré par rapport à la séquence C pendant cet intervalle de temps.

- Un mouvement en diagonale représente une correspondance. Pendant cette période de temps, les séquences Q et C sont synchrones.

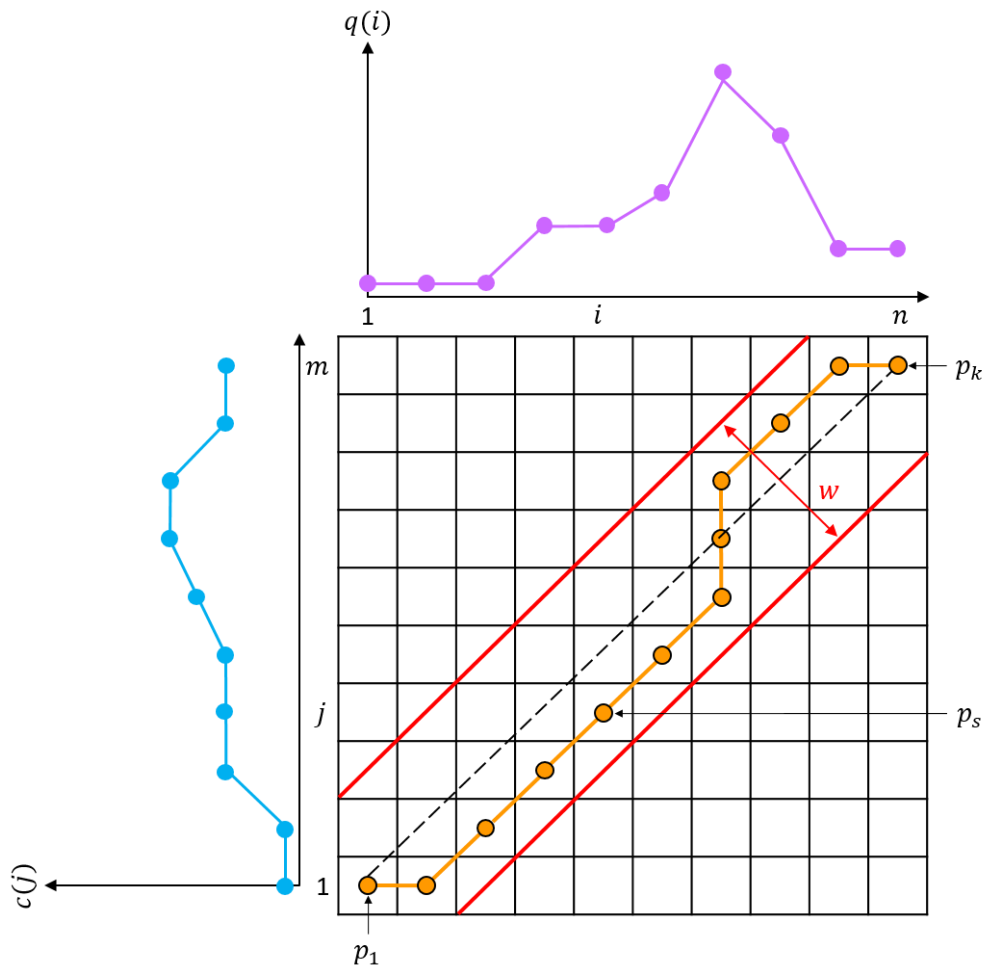


Figure 19. Illustration de la matrice ($n \times m$) de distance cumulée entre deux séries temporelles Q et C , ainsi qu'une fenêtre de déformation w et le chemin P minimisant le coût de déformation entre les deux séquences.

7.2. Résultats de l'indicateur PH1

La station SOMLIT-Astan à Roscoff a été choisie comme station de test pour illustrer l'utilisation de l'indicateur *PH1* du fait de la complétude (i.e., absence de dates d'échantillonnage manquantes) et de la longueur (9 ans, permettant de considérer 3 années dans la période de référence) des séries temporelles disponibles pour les données obtenues par microscopie et par metabarcoding.

L'indicateur *PH1* a été calculé, dans un premier temps pour le couple de formes de vie diatomées/dinoflagellés, en considérant la période de 2009 à 2011 comme période de référence et la période de 2012 à 2016 comme période d'analyse. Ici, les dinoflagellés sont représentés uniquement par des individus de la classe des dinophycées. Le *PH1* global pour la période 2012-2016 est de 0.35 pour les données de microscopie (Figure 20) indiquant un changement significatif (probabilité binomiale < 0.01), avec 65% des observations de cette période de temps se situant en dehors de l'enveloppe de référence définie par les trois premières années d'observation (2009 à 2011). Parmi ces observations, des dates appartenant aux quatre saisons sont retrouvées, avec toutefois une majorité au printemps et en été. Les changements semblent être dus à l'augmentation du nombre de diatomées et, à la fois, à la hausse et à la baisse du nombre de dinophycées. En examinant chaque année séparément depuis 2012, les cinq années de la période d'analyse sont significativement différentes avec une tendance à la baisse, ce qui suggère que la structure de la communauté phytoplanctonique (décrite à travers les comptages microscopiques) continue de changer au fil du temps (Figure 21a). La projection des valeurs médianes annuelles de chaque forme de vie dans l'espace d'état 2D des diatomées et des dinophycées montre qu'il existe une variabilité interannuelle importante (Figure 21b). Bien qu'un changement de communautés phytoplanctoniques ne signifie pas, en soi, un changement négatif (même si la dominance des dinoflagellés peut, dans certains cas représenter un indicateur d'eutrophisation et/ou de changements dans les réseaux trophiques), les résultats de cette analyse du *PH1* calculé à partir des données de microscopie suggèrent que la communauté phytoplanctonique à la station SOMLIT-Astan, concernant ces deux formes de vie, connaît des modifications continues en termes de structure et de composition.

Des résultats similaires ont été obtenus avec les données de metabarcoding (Figure 22), avec toutefois un changement global (bien que significatif) caractérisé par un nombre de dates en dehors de l'enveloppe de référence moins important ($PH1 = 0.47$, signifiant que 53% des observations se situent en dehors de l'enveloppe de référence). L'examen individuel des années de 2012 à 2016 montre une tendance générale à la baisse des valeurs de *PH1* par rapport à la période de référence, avec des différences significatives excepté pour l'année 2014 (Figure 23a). La projection des valeurs médianes annuelles de chaque forme de vie dans l'espace d'état 2D des diatomées et des dinophycées montre également une variabilité interannuelle, mais moins marquée que pour les données de microscopie (Figure 23b). Toutefois, une période de référence limitée à 3 ans peut générer des incertitudes plus importantes sur le calcul de l'indicateur que lorsque de plus longues séries temporelles (e.g., 6 à 10 ans) sont utilisées. Il est donc, à l'heure actuelle, difficile de savoir si les variations temporelles observées reflètent un réel changement d'état du milieu ou une variabilité interannuelle.

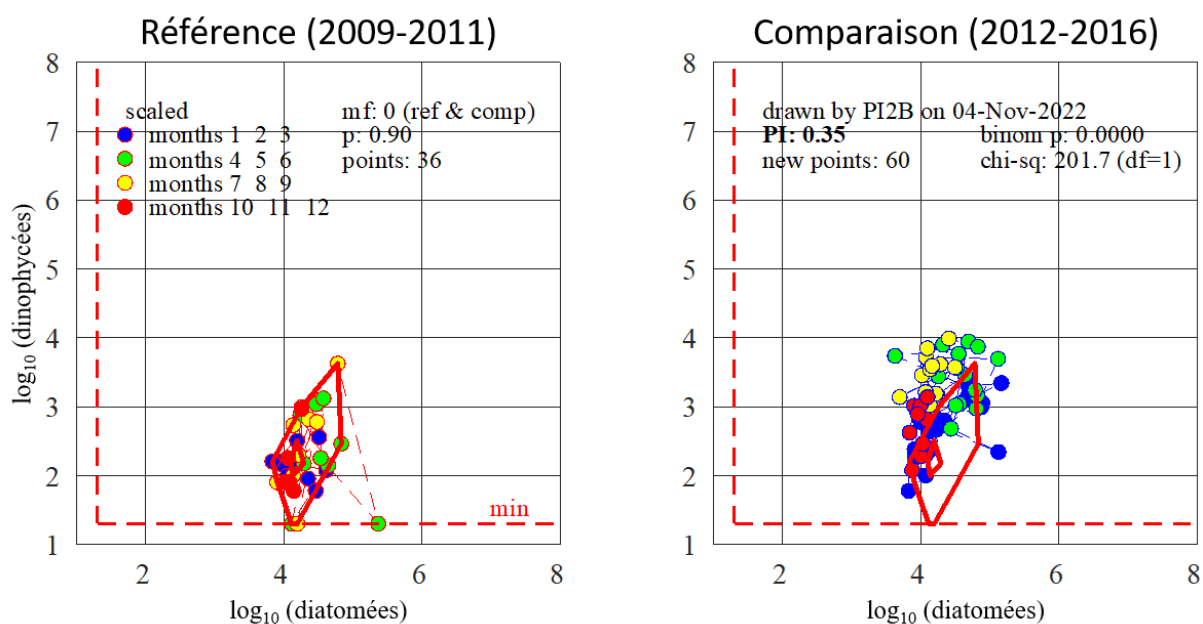


Figure 20. Enveloppe calculée à partir de l'abondance du couple fonctionnel (**diatomées-dinophycées**) estimée par **microscopie** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge).

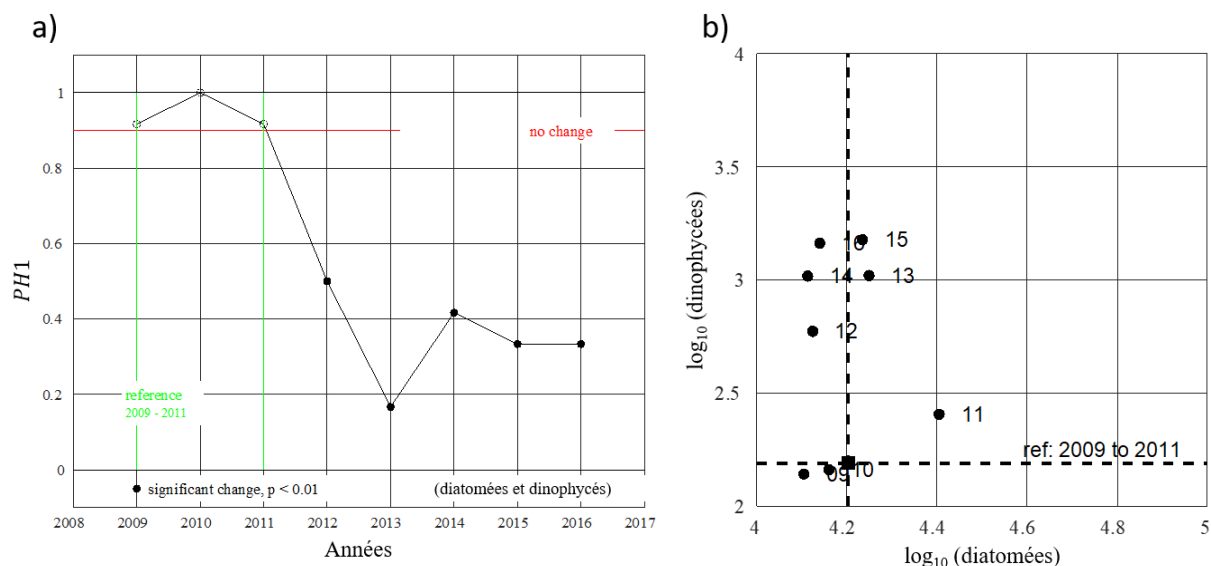


Figure 21. a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-dinophycées**) estimé par **microscopie** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance log₁₀ des diatomées et des dinophycées.

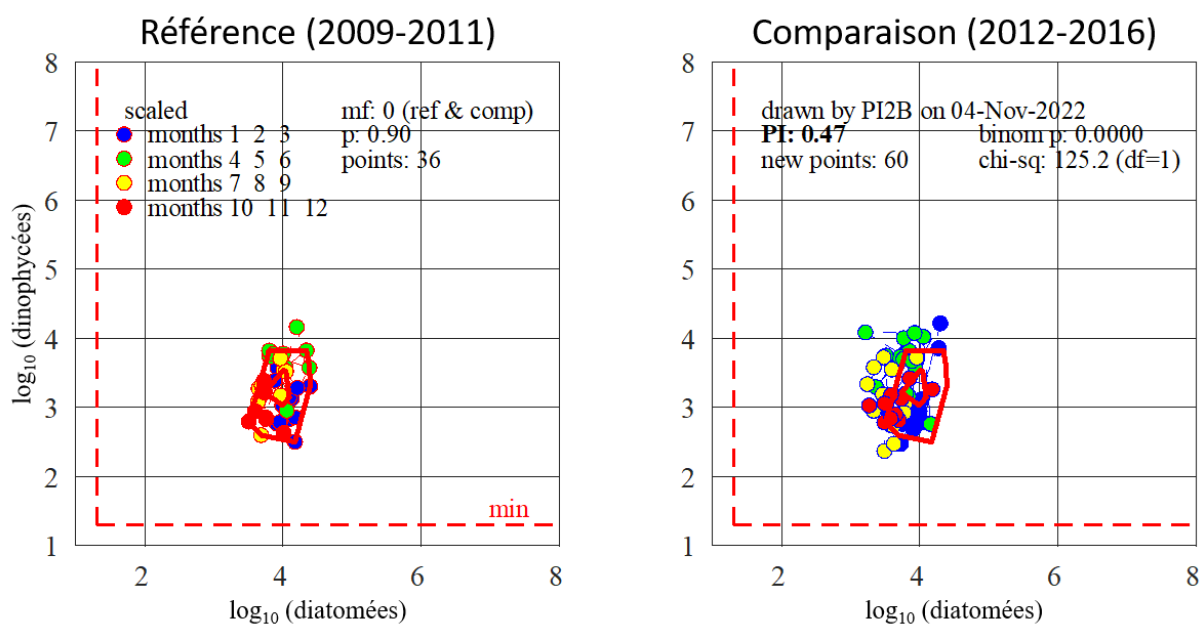


Figure 22. Enveloppe calculée à partir de l'abondance des couples fonctionnels (**diatomées-dinophycées**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 µm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge).

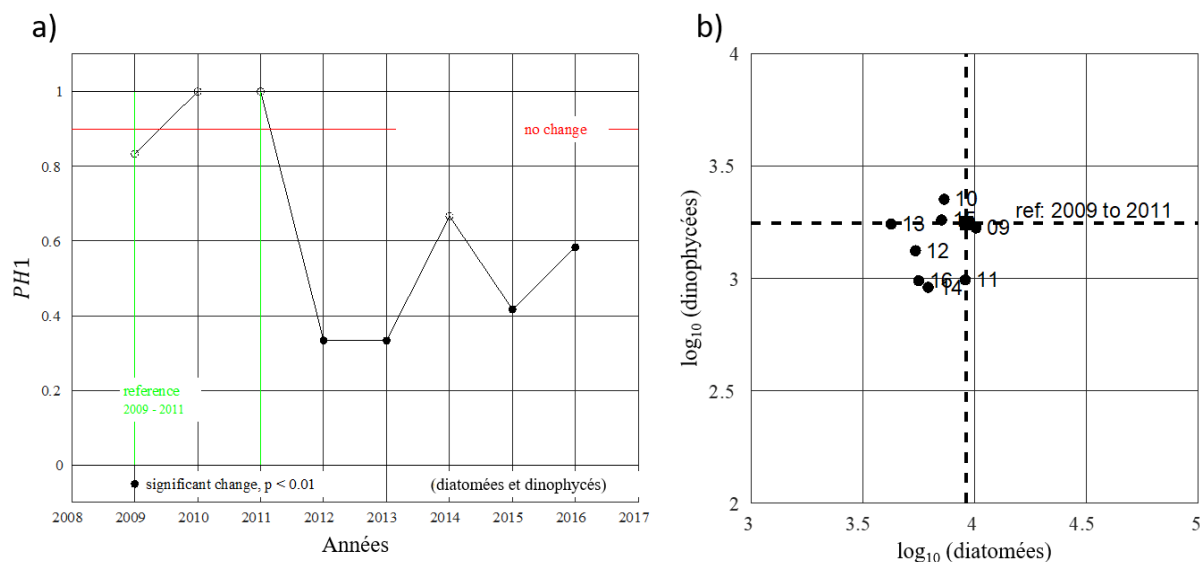


Figure 23. a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-dinophycées**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 µm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des diatomées et des dinophycées.

Le déploiement de l'approche par metabarcoding permet d'accéder à toutes les classes de taille du compartiment microbien, et notamment au pico- et nanoplancton dont la diversité n'est pas accessible via la microscopie. Cette approche permet aussi d'accéder aux compartiments microbiens mixotrophes ou hétérotrophes, aux parasites et symbiotes qui sont très rarement comptabilisés en microscopie. Il est donc possible de calculer l'indice *PH1* sur de très nombreuses paires de formes de vie non accessibles par la microscopie. Ainsi, outre le rapport diatomées-dinoflagellés, initialement calculé avec les données obtenues par microscopie (Figures 20 à 23), le rapport diatomées-cercozoaires (les cercozoaires comprennent de nombreux parasites planctoniques) et le rapport hétérotrophes-autotrophes peuvent par exemple être testés avec des données de metabarcoding (Figures 24 à 27).

D'après les analyses réalisées avec les données de metabarcoding, les rapports entre les abondances des diatomées et des cercozoaires étaient plus stables dans le temps ($PH1 = 0.53$, ce qui signifie que 47% des points sont hors de l'enveloppe de référence, Figure 24) en comparaison du couple de formes de vie diatomées-dinoflagellés ($PH1 = 0.35$, Figure 20). De plus, l'enveloppe de référence du couple diatomées-cercozoaires est moins étendue dans le plan factoriel signifiant une plus faible variabilité temporelle dans les rapports diatomées/cercozoaires. A l'échelle annuelle, une tendance générale à la baisse des valeurs de *PH1* est également observée. Comme précédemment observé pour le rapport diatomées/dinoflagellés, l'année 2014 apparaît comme étant la seule année à ne pas présenter de changement significatif par rapport à la période de référence. Les dates qui se trouvent à l'extérieur de l'enveloppe de référence, et qui présentent donc un changement significatif, sont majoritairement des dates printanières et estivales. Par ailleurs, une variabilité interannuelle dans les abondances des diatomées et des cercozoaires est toutefois à noter (Figure 25b). Des résultats similaires sont observés pour le couple hétérotrophes-autotrophes avec un indicateur *PH1* global égal à 0.57 (Figure 26) et une tendance générale à la baisse (Figure 27a) mais une stabilité temporelle encore plus prononcée, marquée notamment par une faible variabilité interannuelle dans les abondances des hétérotrophes et des autotrophes (Figure 27b).

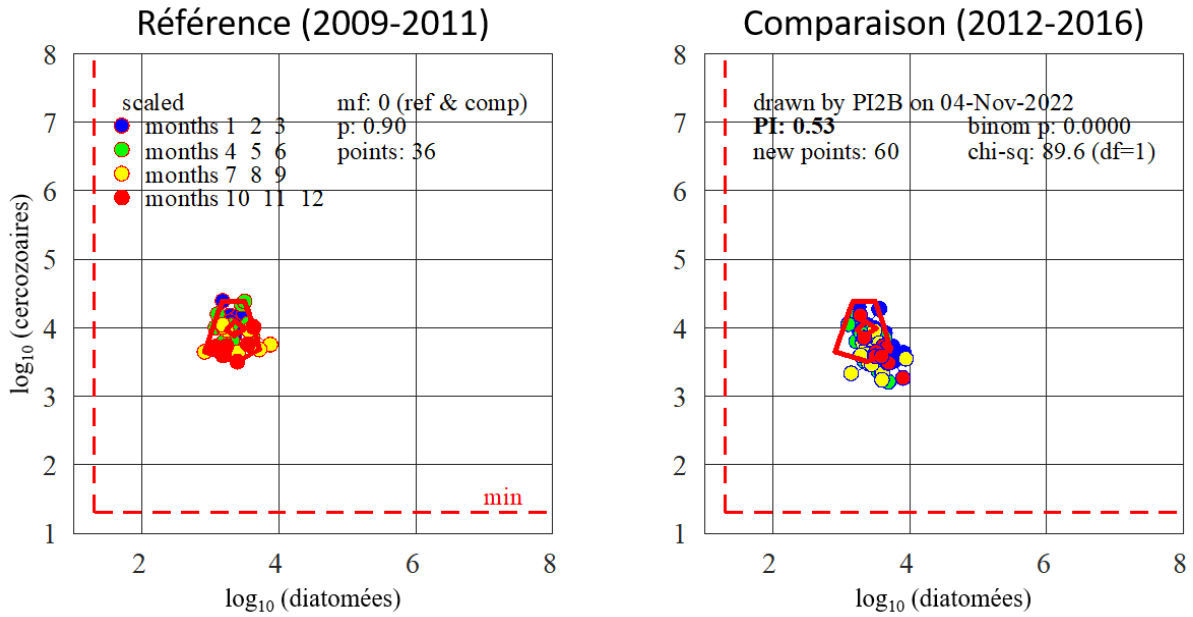


Figure 24. Enveloppe calculée à partir de l'abondance des couples fonctionnels (**diatomées-cercozoaires**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge).

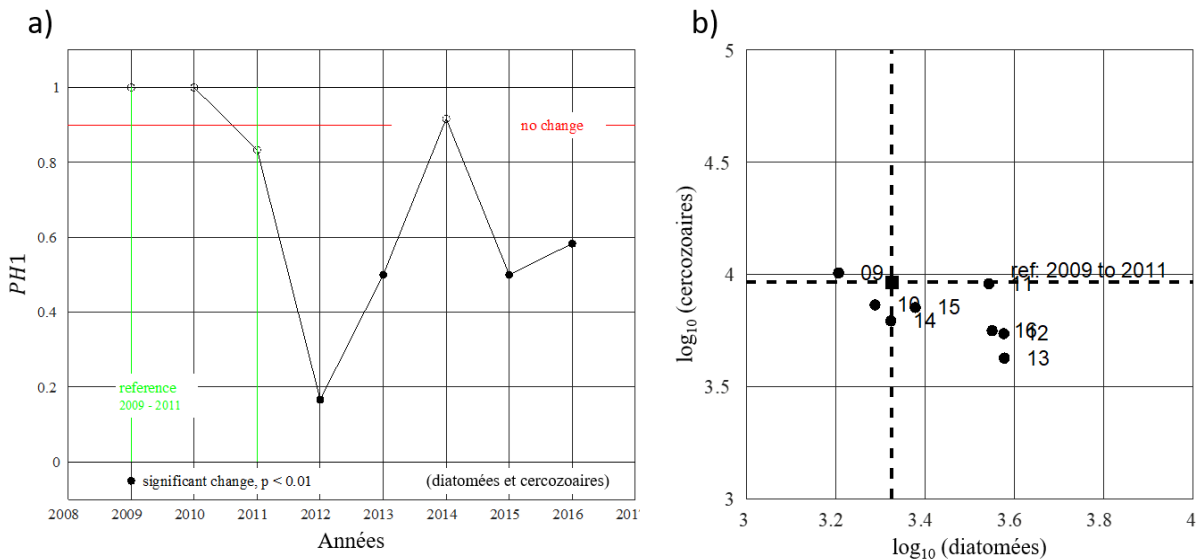


Figure 25. a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**diatomées-cercozoaires**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des diatomées et des cercozoaires.

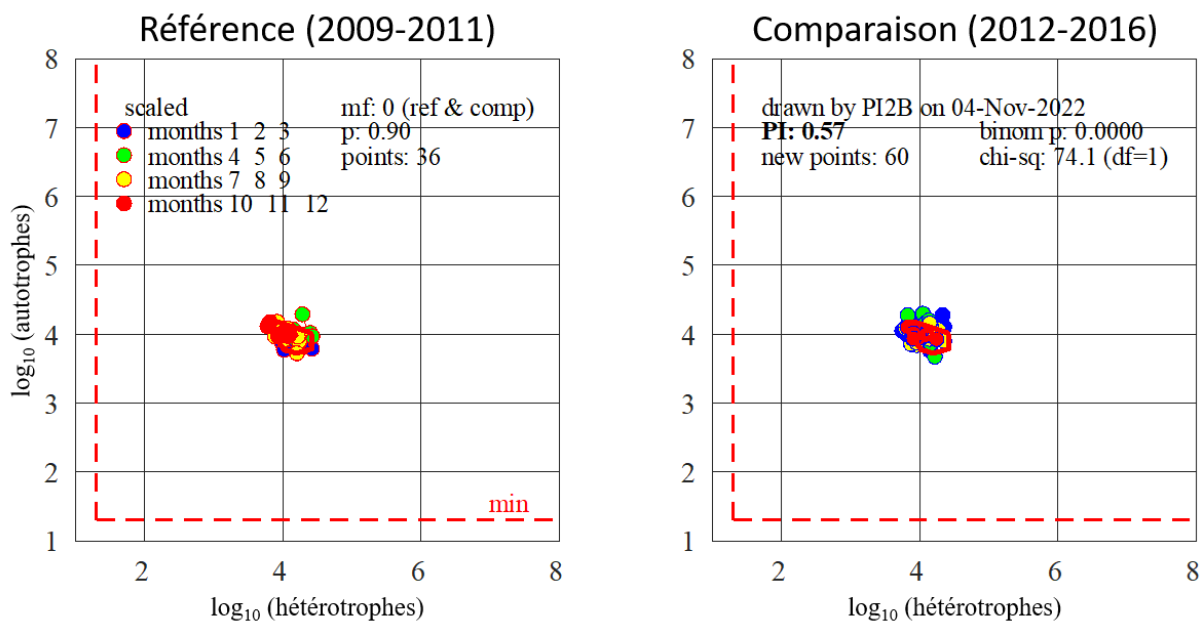


Figure 26. Enveloppe calculée à partir de l'abondance des couples fonctionnels (**hétérotrophes-autotrophes**) estimée par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Le graphique de gauche montre l'enveloppe de référence (enveloppe convexe en rouge) estimée à partir des dates d'échantillonnage (point de couleurs) sur une période de référence de 3 ans (de 2009 à 2011). Le graphique de droite représente la projection des dates d'échantillonnage de la période d'évaluation (de 2012 à 2016) sur l'enveloppe de référence. Chaque point est coloré en fonction de la saison (hiver : bleu, printemps : vert, été : jaune, automne : rouge).

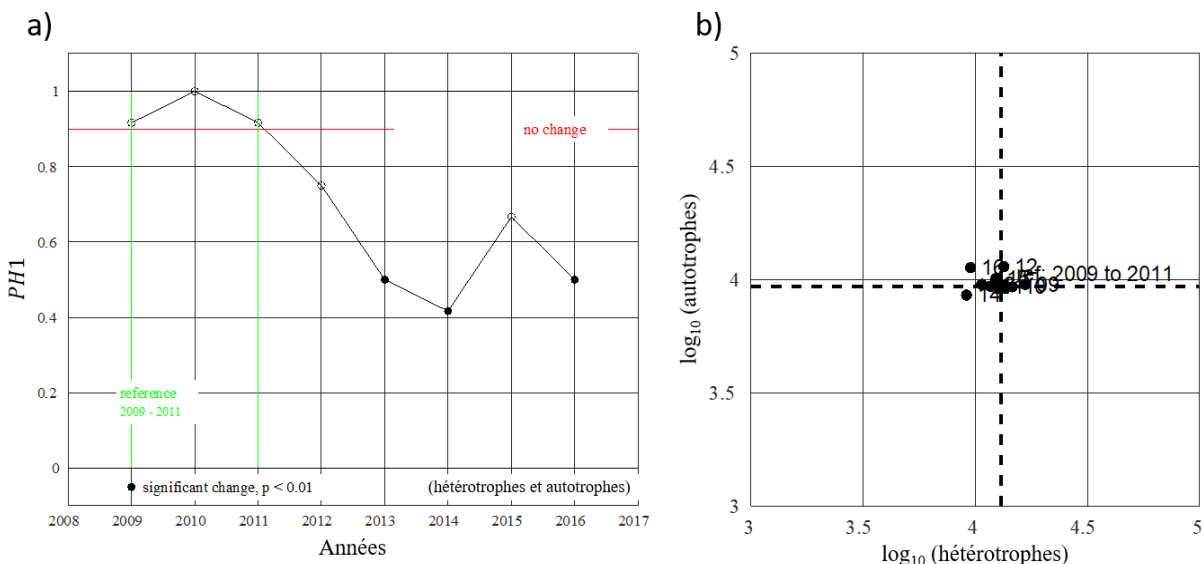


Figure 27. a) Série chronologique des valeurs de l'indice PH1 pour le couple fonctionnel (**hétérotrophes-autotrophes**) estimé par **metabarcoding (ARNr 18S, fraction de taille >3 μm)** pour la station SOMLIT-Astan à Roscoff. Les cercles pleins représentent une valeur PH1 significativement différente de la période de référence (ligne horizontale rouge). b) Diagramme d'espace d'état 2D des valeurs médianes annuelles de l'abondance \log_{10} des hétérotrophes et des autotrophes.

7.3. Résultats sur l'indicateur PH3

7.3.1. Dynamique temporelle de la diversité alpha (indices de Hulburt et de Menhinick)

La diversité alpha des communautés phytoplanctoniques est décrite par un indice de richesse (Menhinick) et un indice de dominance (Hulburt). Les variations temporelles de ces indices ont été estimées à partir des données acquises par microscopie ou par metabarcoding pour la station de Roscoff (Figure 28), la station de Wimereux (Figure 29) et les deux stations de la Marina du Château (Figures 30 et 31). **Ces indices de diversité montrent d'importantes variations intra- et interannuelles quel que soit le type de données (i.e., microscopie ou metabarcoding) et la station considérée.** Sur les périodes temporelles disponibles pour chaque station, l'indice 100 – Hulburt varie entre 0.587 et 76.190 (Wimereux : [0.587, 76.190], Roscoff : [9.151, 73.342], Marina du Château : [2.553, 50.082]) pour les données de microscopie et entre 4.955 et 88.340 (Wimereux : [4.955, 82.842], Roscoff : [36.790, 88.340], Marina du Château : [15.259, 74.845]) pour les données de metabarcoding. L'indice de Menhinick varie entre 0.002 et 0.486 (Wimereux : [0.002, 0.221], Roscoff : [0.042, 0.486], Marina du Château : [0.002, 0.036]) pour les données de microscopie et entre 0.056 et 1.266 (Wimereux : [0.056, 1.266], Roscoff : [0.281, 1.011], Marina du Château : [0.108, 0.416]) pour les données de metabarcoding. **D'une manière générale, pour l'ensemble des stations sur les périodes temporelles considérées, les épisodes de forte richesse (fortes valeurs de l'indice de Menhinick) sont associés à de faibles dominances (fortes valeurs de 100 – Hulburt).**

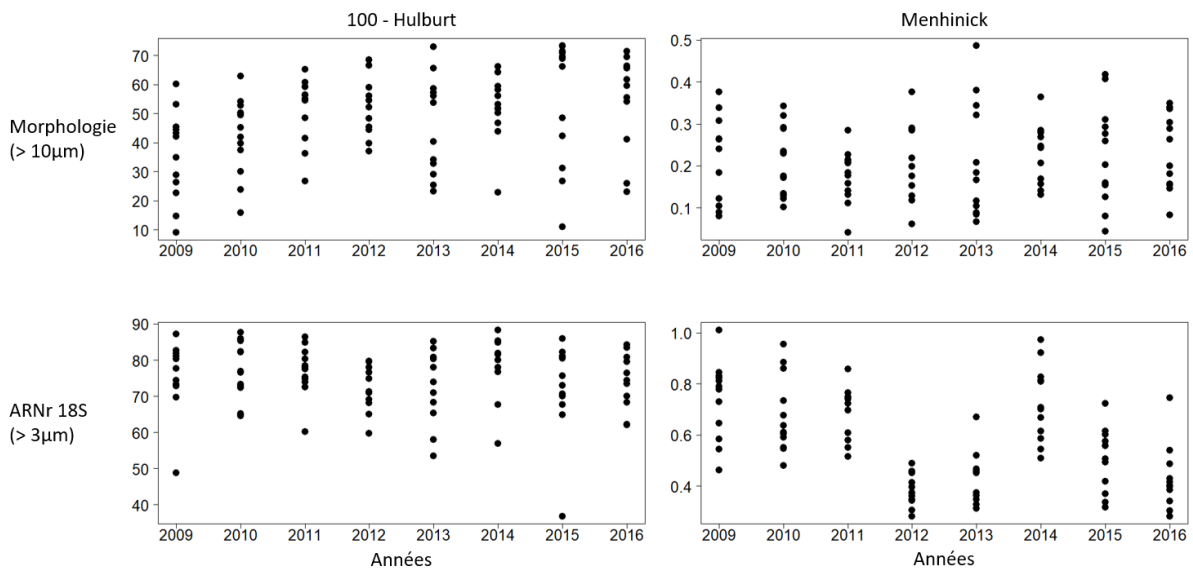


Figure 28. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station SOMLIT-Astan à Roscoff pour la période 2009-2016 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).

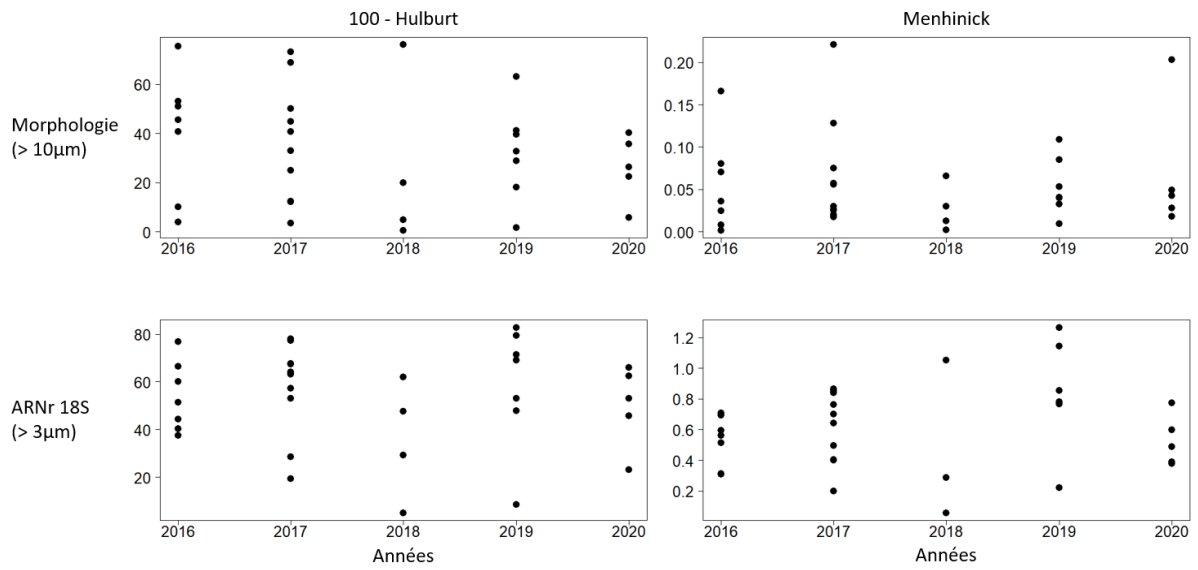


Figure 29. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station SOMLIT-C à Wimereux pour la période 2016-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).

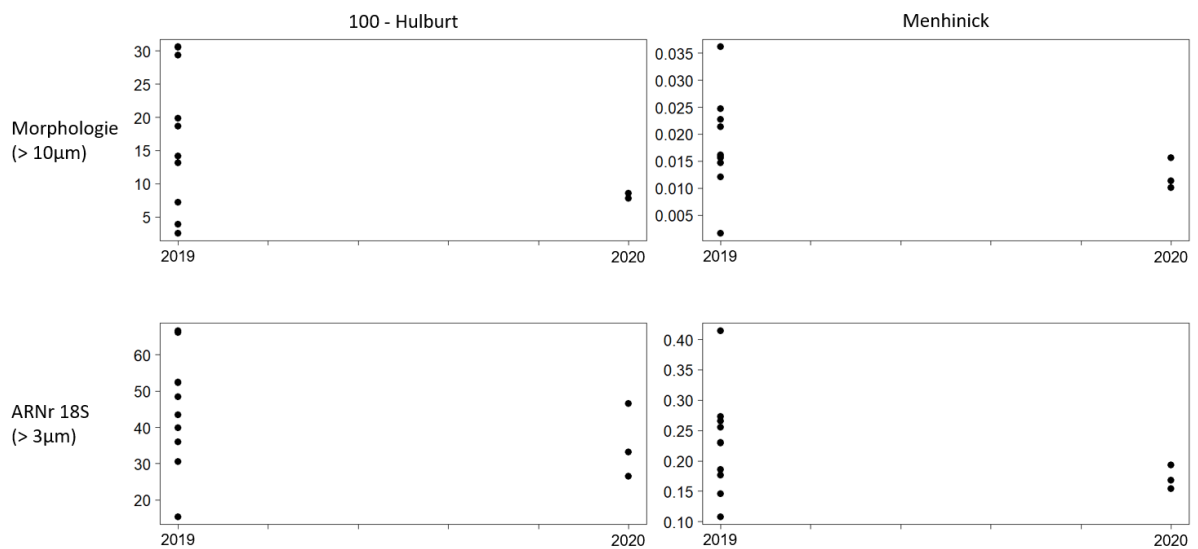


Figure 30. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station B8 de la Marina du Château pour la période 2019-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).

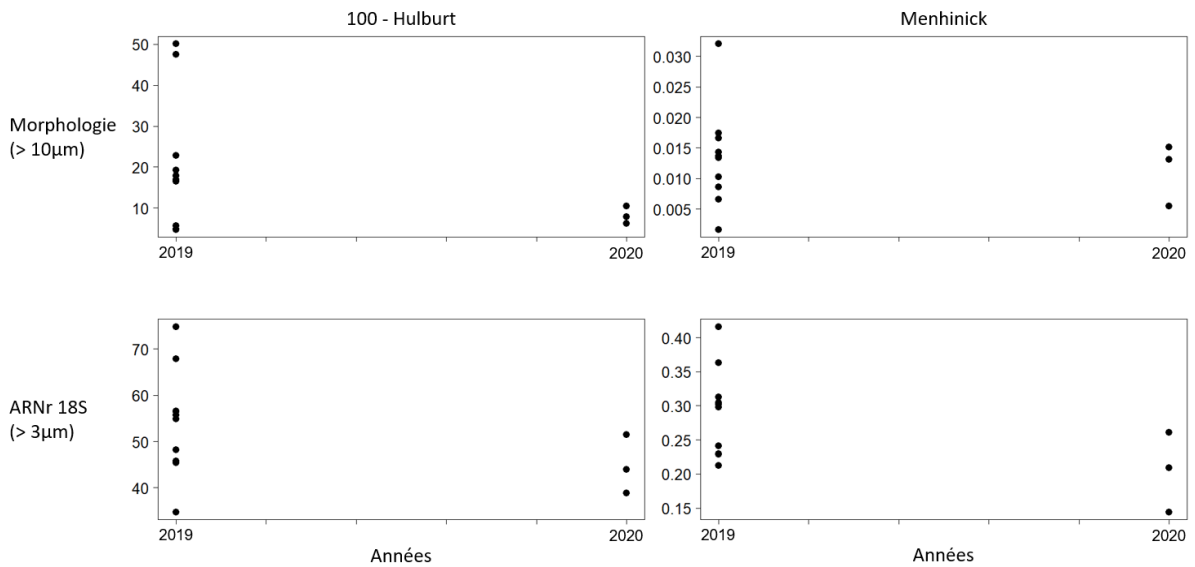


Figure 31. Variations temporelles de l'indice 100 – Hulburt et de l'indice de Menhinick à la station H42 de la Marina du Château pour la période 2019-2020 estimés à partir des données de microscopie (en haut) et de metabarcoding (en bas).

7.3.2. Dynamique temporelle de la diversité bêta (LCBD)

La LCBD renseigne sur l'hétérogénéité des communautés planctoniques et permet d'identifier des épisodes temporels où la structure des communautés est atypique par rapport au reste de la série. Les valeurs mensuelles de LCBD ont été calculées à partir des données de microscopie et de metabarcoding pour la station de Roscoff (Figure 32), la station de Wimereux (Figure 33) et les deux stations de la Marina du Château (Figures 34 et 35). **Cet indice montre d'importantes variations intra- et interannuelles quel que soit le type de données (i.e., microscopie ou metabarcoding) et la station considérée.** La LCBD varie entre 0.004 et 0.373 (Wimereux : [0.022, 0.041], Roscoff : [0.004, 0.022], Marina du Château : [0.013, 0.373]) pour les données de microscopie et entre 0.005 et 0.126 (Wimereux : [0.020, 0.041], Roscoff : [0.005, 0.019], Marina du Château : [0.034, 0.126]) pour les données de metabarcoding.

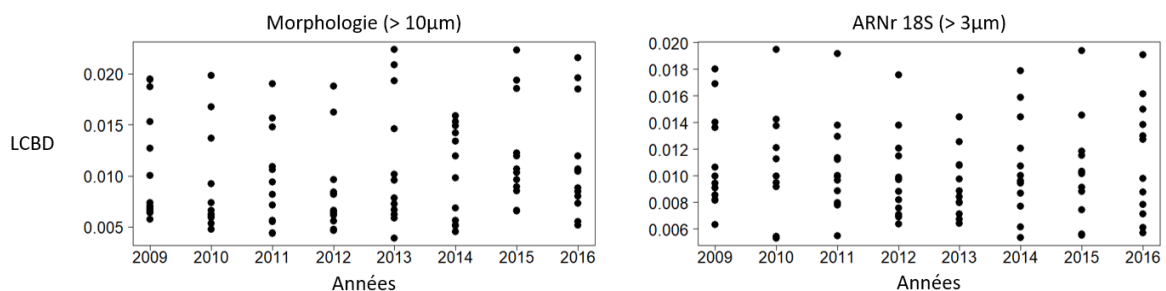


Figure 32. Variations temporelles de la LCBD à la station SOMLIT-Astan à Roscoff pour la période 2009-2016 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).

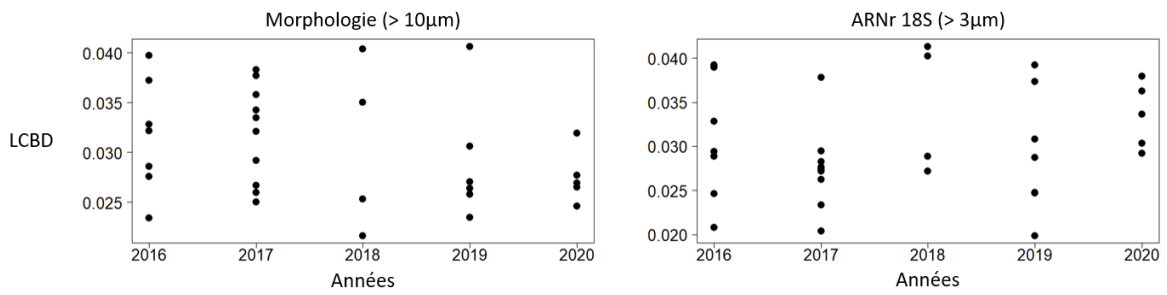


Figure 33. Variations temporelles de la LCBD à la station SOMLIT-C à Wimereux pour la période 2016-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).

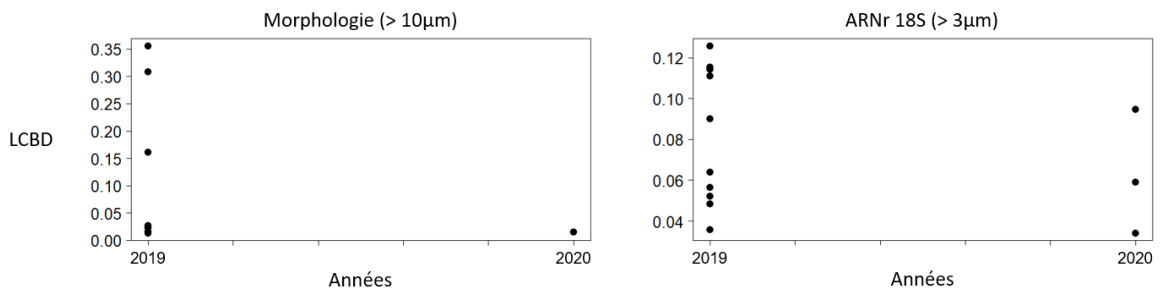


Figure 34. Variations temporelles de la LCBD à la station B8 de la Marina du Château pour la période 2019-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).

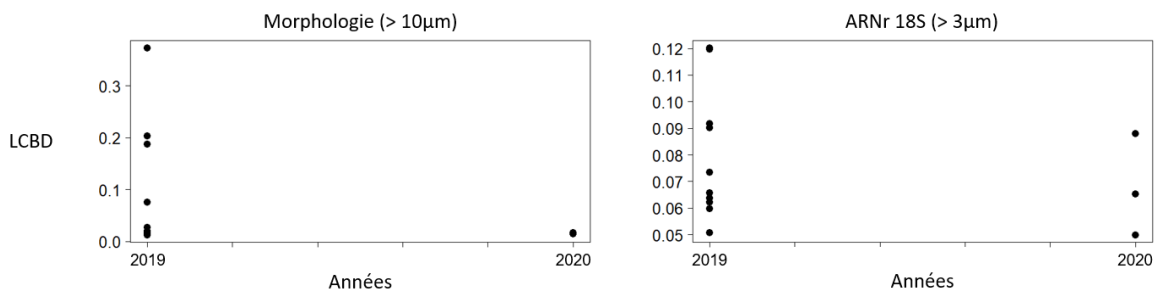


Figure 35. Variations temporelles de la LCBD à la station H42 de la Marina du Château pour la période 2019-2020 estimée à partir des données de microscopie (à gauche) et de metabarcoding (à droite).

7.3.3. Indicateurs PH3 estimé à partir des données de microscopie et de metabarcoding

Pour les stations SOMLIT à Roscoff et à Wimereux et les deux stations de la Marina du Château, l'indicateur *PH3* a pu être calculé à partir des données de microscopie et de metabarcoding pour le compartiment eucaryote. Les méthodes de calcul du *PH3* sont décrites dans le paragraphe 7.1.2 (voir aussi l'équation (6)). Pour rappel, le ***PH3* correspond à la somme des EQRs calculés à partir de 3 indices : 100 –Hulburt, Menhinick et LCBD** (avec des EQRs calculés par rapport à des valeurs de référence obtenues par saison, voir la figure 16 du paragraphe 7.1.2).

Pour chaque série temporelle (correspondant à une station donnée), les valeurs de l'indicateur *PH3* obtenues à partir des données de microscopie et de metabarcoding ont été comparées (Figures 36, 37, 38). **Pour la station SOMLIT-Astan** (pour laquelle la série temporelle continue de 8 années avec prélèvements bimensuels réguliers permet une analyse sur 96 dates reportées en moyennes mensuelles) **les valeurs de *PH3* calculées à partir des deux types de jeux de données sont corrélées significativement**. Il apparaît par ailleurs que le *PH3* calculé à partir des données de metabarcoding est généralement plus élevé que celui calculé à partir des données de microscopie et ceci quel que soit le niveau taxonomique considéré (Figure 36). **Pour les données acquises aux stations de Wimereux et de la marina du Château, les mêmes conclusions peuvent être déduites lorsque l'on observe le lien entre les valeurs de *PH3* calculé à partir des deux types de jeux de données et que les niveaux de corrélation sont estimés (Figures 37 et 38)**.

Rappelons que les résultats obtenus pour le *PH3* à partir des données de metabarcoding se rapportent à la communauté de microeucaryotes dans son ensemble (grâce à l'utilisation d'amorces eucaryotes universelles) et pas seulement aux taxons distingués en microscopie. De plus, les données de metabarcoding sont obtenues de manière reproductible (de l'extraction de l'ADN à la production des ASVs).

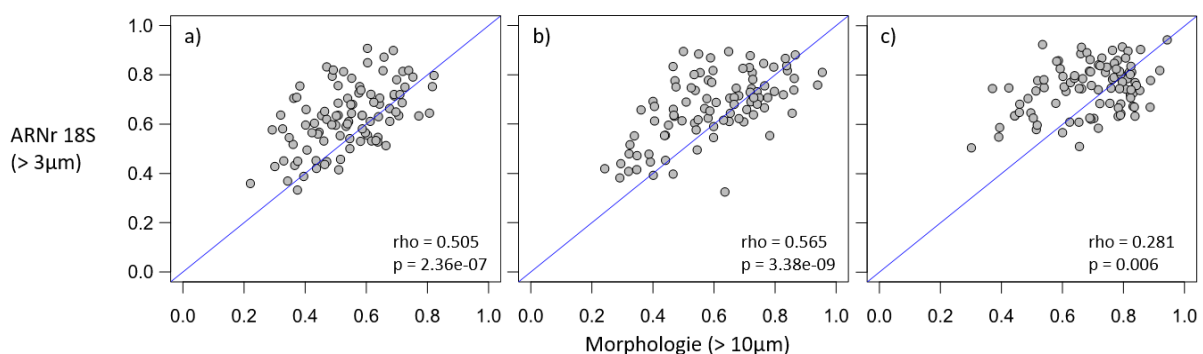


Figure 36. Relation entre les valeurs de l'indicateur *PH3* inféré à partir des données de **microscopie** et de **metabarcoding** (**ARNr 18S, fraction de taille >3 µm**). a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris correspond à une date d'échantillonnage de la station SOMLIT-Astan à Roscoff. La relation entre les deux formes de *PH3* a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman.

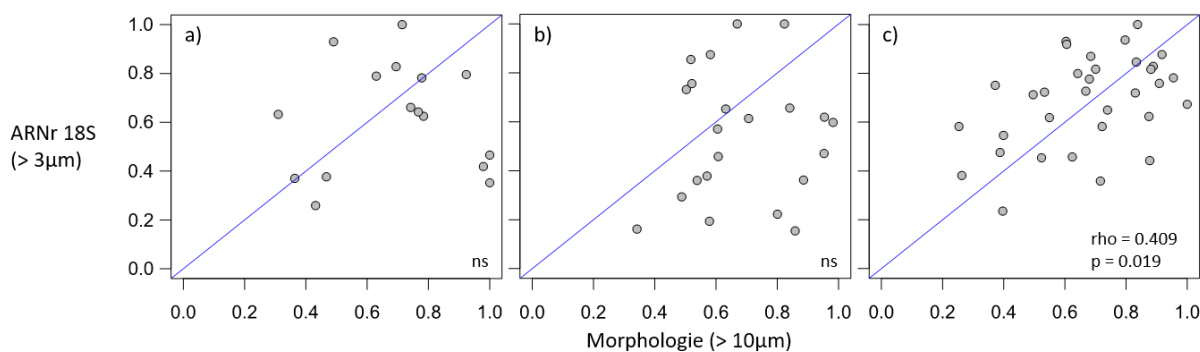


Figure 37. Relation entre les valeurs de l'indicateur PH3 inféré à partir des données de **microscopie** et de **metabarcoding** (ARNr 18S, fraction de taille >3 µm). a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris correspond à une date d'échantillonnage de la station SOMLIT-C à Wimereux. La relation entre les deux formes de PH3 a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman.

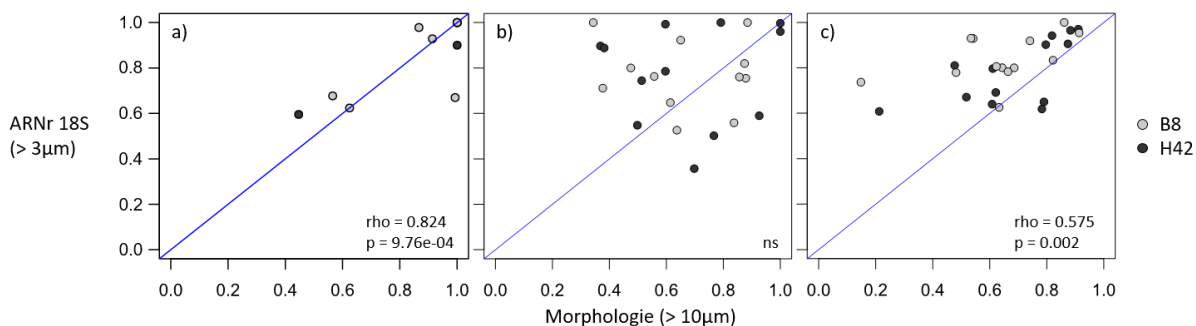


Figure 38. Relation entre les valeurs de l'indicateur PH3 inféré à partir des données de **microscopie** et de **metabarcoding** (ARNr 18S, fraction de taille >3 µm). a) La relation est estimée en considérant uniquement les espèces retrouvées dans les deux types de jeux de données. b) La relation est estimée en considérant les genres communs aux deux types de jeux de données. c) La relation est estimée en considérant la globalité des deux jeux de données. Chaque point gris (clair et foncé) correspond à une date d'échantillonnage de la station de la Marine du Château à Brest. Les points gris clair correspondent aux dates d'échantillonnage du site B8 et les points gris foncé correspondent aux dates d'échantillonnage du site H42. La relation entre les deux formes de PH3 a été estimée par le calcul du coefficient de corrélation non paramétrique de Spearman.

7.3.4. Evaluation de l'indicateur PH3 selon les classes de qualités, comparaison entre les données microscopiques et de metabarcoding

Dans le cadre des directives européennes, et notamment de la DCSMM, il est intéressant de traduire les indices de changement de l'indicateur *PH3* en classes de qualité pour les jeux de données acquises en microscopie et par metabarcoding. **Nos résultats montrent que la qualification des échantillons prélevés le long des séries temporelles est, pour les deux types de données, très proche (même classe) ou proche (classes contiguës) pour 89%, 91% et 81% des échantillons, respectivement pour les jeux de données SOMLIT-Astan (Roscoff), SOMLIT-C (Wimereux) et de la Marina du Château (Figures 39, 40 et 41).** Très peu d'échantillons présentaient une mauvaise correspondance de qualification entre les deux types de données, à savoir un écart d'au moins deux classes de qualité (Astan : 11%, Wimereux : 9%, Marina du

Château : 15%), et un seul échantillon de la station B8 à la Marina du Château présentait un écart de trois classes de qualité (Figure 41). Deux tendances se dégagent pour les échantillons présentant une mauvaise correspondance dans les classes de qualité du PH3 entre la microscopie et le metabarcoding. Aux stations de Roscoff et de Wimereux, cette mauvaise correspondance des classes de qualité est bidirectionnelle, avec des cas de figure où les données de metabarcoding attribuent à un échantillon une classe de qualité supérieure à celle estimée avec les données de microscopie et inversement. Aux stations de la Marina du Château la mauvaise correspondance est unidirectionnelle, avec uniquement le cas de figure où les données de metabarcoding attribuent une classe de qualité supérieurs à celle obtenue par microscopie.

Ces résultats montrent également, à la station SOMLIT-Astan, que les différences de classement entre les deux types de jeux de données ne varient pas en fonction de la saison. Il est à noter toutefois, qu'aucun échantillon hivernal n'a été classé dans la classe de qualité « éloignée » de l'état de référence avec l'approche de metabarcoding à l'inverse de l'approche par microscopie (Figure 39). **Une telle similarité dans les proportions d'échantillons distribués entre les différentes saisons est observée pour la station SOMLIT-C à Wimereux et tout particulièrement pour les classes de qualité « très proche », « proche » et « éloignée » de l'état de référence (Figure 40).** Concernant la classe de qualité « intermédiaire », aucun échantillon automnal de cette station n'est retrouvé dans cette classe de qualité selon les données de microscopie, et aucun échantillon hivernal selon les données de metabarcoding. **Une correspondance dans les proportions d'échantillons distribués entre les différentes saisons est également observée pour les deux stations à la Marina du Château et en particulier pour les classes de qualité « très proche » et « proche » de l'état de référence (Figure 41).** En effet, selon l'approche de metabarcoding, aucun échantillon n'est classé dans un état « intermédiaire », « éloigné » ou « très éloigné ».

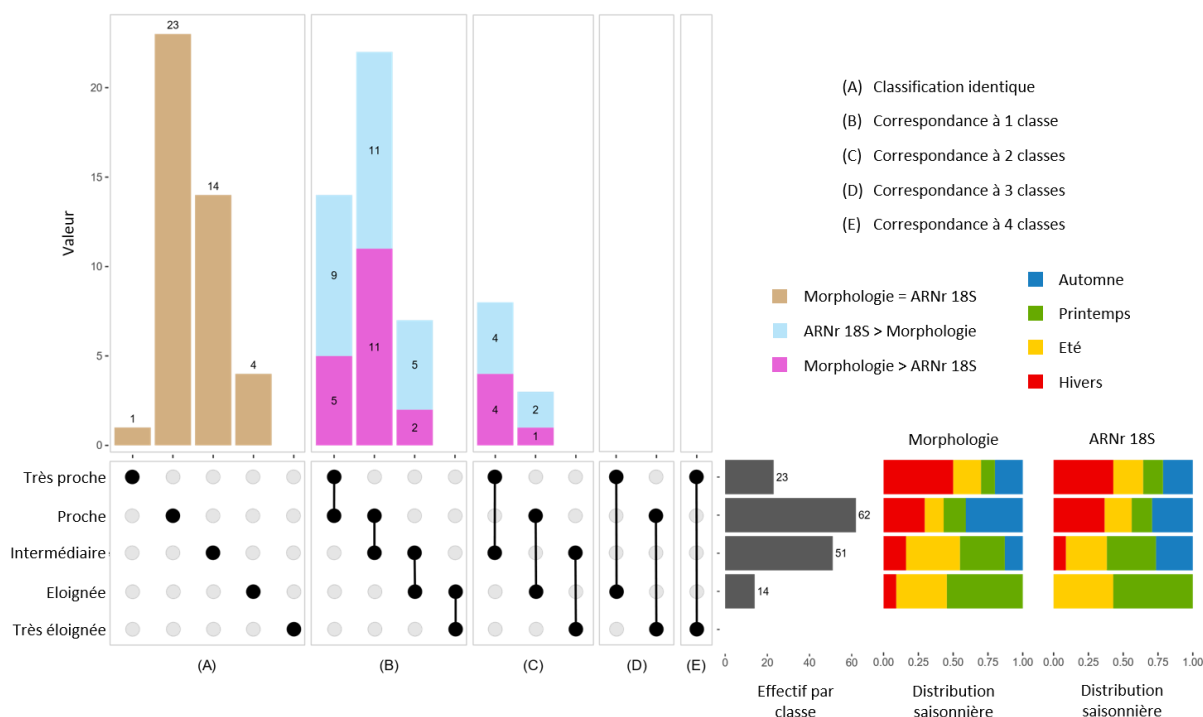


Figure 39. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de **microscopie** et les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)** sur l'ensemble de la série temporelle (de 2009 à 2016) à la station SOMLIT-Astan à Roscoff. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding. Deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose.

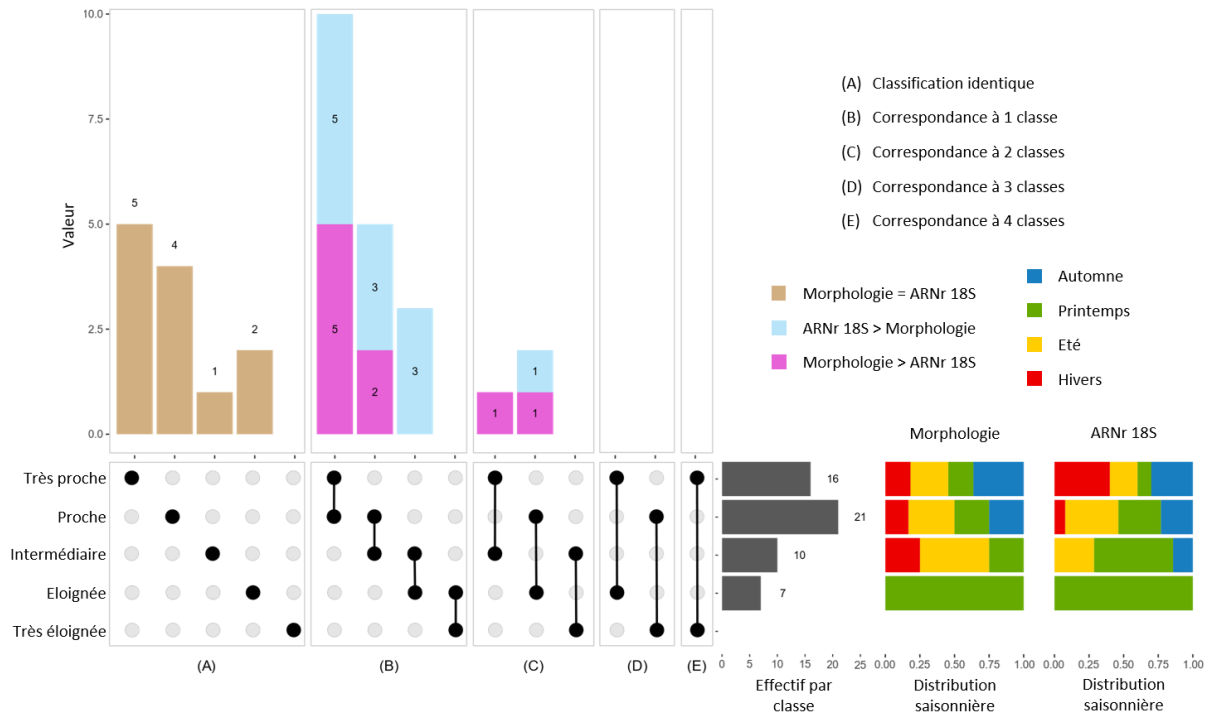


Figure 40. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de **microscopie** et les données de **metabarcoding (ARNr 18S, fraction de taille >3 µm)** sur l'ensemble de la série temporelle (de 2015 à 2020) à la station SOMLIT-C à Wimereux. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding et deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose.

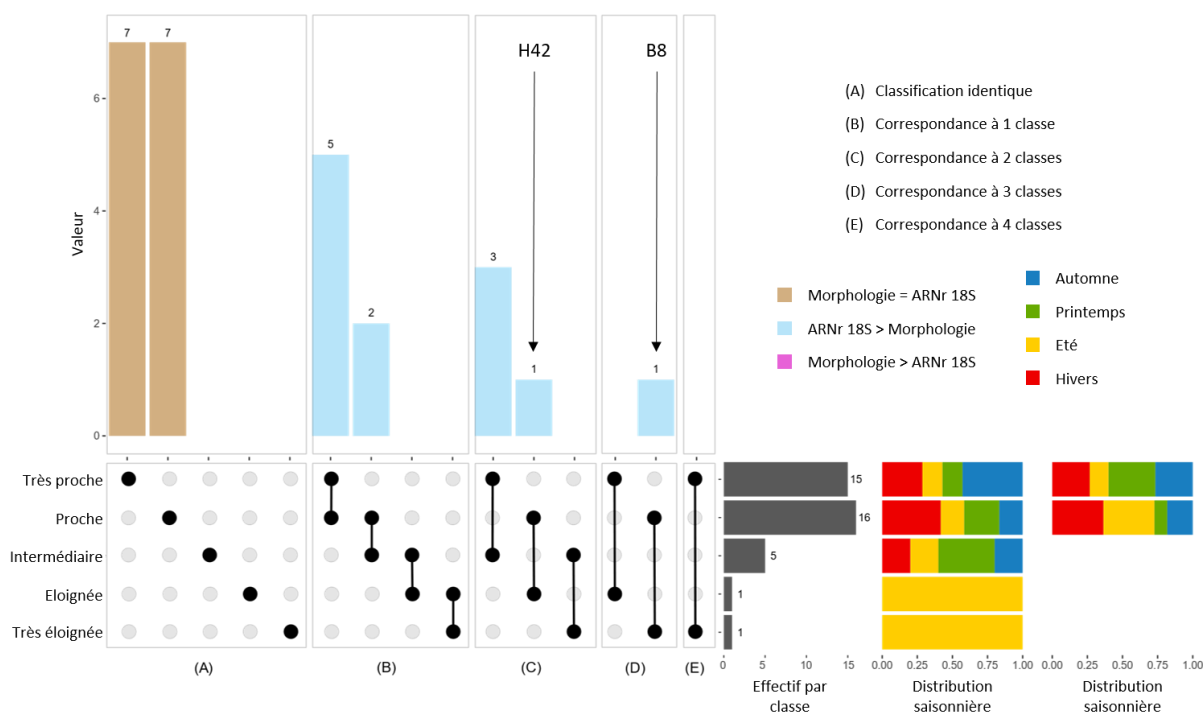


Figure 41. Analyse comparative (UpSet plot) des classes de qualité (i.e., état très éloigné, éloigné, intermédiaire, proche, très proche de l'état de référence) assignées aux valeurs de l'indicateur PH3 calculé avec les données de microscopie et les données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)** sur l'ensemble de la série temporelle (de 2019 à 2020) à la station de la Marine du Château à Brest. Les résultats présentent les données des deux sites B8 et H42 combinés. Dans le panneau inférieur gauche, les points noirs indiquent les interactions entre les classes. Un seul point noir indique une date assignée à la même classe par la microscopie et le metabarcoding et deux points noirs joints indiquent une date assignée à deux classes différentes par la microscopie et le metabarcoding. Le nombre de dates d'échantillonnage dans chaque classe de qualité distribué en fonction des saisons (hiver : rouge, printemps : vert, été : jaune, hiver : rouge) est représenté dans les deux panneaux inférieurs à droite. Le nombre total de dates d'échantillonnage dans chaque classe de qualité est représenté dans le panneau inférieur central. Le nombre de dates d'échantillonnage de chaque interaction est indiqué dans le panneau supérieur. Les cas où la classe de qualité assignée à l'indicateur PH3 calculé à partir des données de metabarcoding est supérieure à celle obtenue à partir des données de microscopie sont indiqués en bleu clair dans le panneau supérieur. La relation inverse est indiquée en rose.

Malgré les résultats des analyses comparatives des valeurs du PH3 obtenues selon les deux types de données (cf. paragraphe 7.3.3), montrant une surestimation globale de l'indice PH3 inféré avec les données de metabarcoding par rapport aux données de microscopie (Figures 36, 37 et 38), **les résultats de ce paragraphe suggèrent que l'évaluation du changement d'état des communautés microbiennes eucaryotes, en termes de classes de qualité, est comparable spatialement et temporellement entre les deux types de données.**

7.3.5. *Evaluation de la sensibilité de l'indicateur PH3 en fonction du nombre d'échantillon considéré*

Pour évaluer l'effet de la fréquence d'acquisition des données sur les valeurs de l'indicateur *PH3*, une analyse de sensibilité de ce dernier à la dégradation des séries temporelles a été réalisée. Pour cela, les valeurs de *PH3* calculées avec le jeu de données original (fréquence bimensuelle à mensuelle) ont été comparées aux valeurs calculées à partir de jeux de données dégradés selon six scénarios (fréquence variant d'un échantillonnage sur un/deux mois par saison à un échantillonnage sur 11 mois par an ou de une/deux/trois saisons par an). Les différences absolues entre les valeurs obtenues avec les données originales et les données dégradées ont été utilisées pour estimer la sensibilité. Les détails concernant les procédures de dégradation utilisées figurent dans le paragraphe 7.1.3 (voir notamment dans la figure 18). **Les valeurs moyennes des différences absolues sont faibles pour les données de microscopie, quel que soit le type de dégradation considéré.** La différence absolue moyenne est maximale (0.08) lorsque seulement la fréquence d'un mois par saison (i.e., 4 mois par an) est utilisée pour calculer l'indicateur *PH3* et diminue de manière continue pour les autres types de dégradation jusqu'à une valeur minimale de 0.01 lorsque les données de 11 mois par an sont utilisées (Figure 42). L'intervalle des valeurs de chaque classe de qualité de l'indicateur *PH3* (Figure 16) est de 0.2. Par conséquent, une sensibilité de 0.08 signifie que dans le pire des cas un échantillon peut être déclassé ou surclassé à la classe de qualité contiguë à l'originale quand le *PH3* est calculé en considérant une fréquence d'un mois par saison dans la série temporelle. **De tels résultats suggèrent une sensibilité moyenne à la dégradation des données relativement faible de l'indicateur *PH3*, quand celui-ci est calculé à partir des données de microscopie.** Toutefois, deux groupes de scénarios de dégradation se distinguent. Le premier groupe, constitué de deux scénarios de dégradation (i.e., 1 mois par saison et 1 saison par an), présente des différences absolues moyennes supérieures à 0.05. Le deuxième groupe, regroupant les quatre autres scénarios de dégradation, présente des différences absolues moyennes inférieures à 0.025. Ces groupes se distinguent également par des gammes de variabilité différentes dans les valeurs de différences absolues mesurées. Globalement, le premier groupe présente une gamme de variabilité des différences absolues plus importante que le deuxième groupe (1^{er} groupe : [0, 0.35[, 2^{ème} groupe : [0, 0.20[). **Cela indique que dans le cas le plus pessimiste de dégradation des données du groupe 1 (i.e., 1 mois par saison, sensibilité maximale supérieure à 0.30) et du groupe 2 (i.e., 2 mois par saison, sensibilité maximale égale à 0.20), un échantillon peut être déclassé ou surclassé respectivement à deux classes de qualité d'écart à l'originale (voir le cas C de la figure 38) et à la classe de qualité contiguë à celle-ci (voir le cas B de la figure 38).**

Des résultats similaires sont observés lorsque le *PH3* est calculé à partir des données de metabarcoding (Figure 43). Ceci concerne la dynamique des différences absolues moyennes entre les différents scénarios de dégradation, la distinction des deux mêmes groupes de scénarios de dégradation en fonction des valeurs des différences absolues moyennes et les gammes de variabilité observée dans ces deux groupes. Pour autant, les valeurs des différences absolues moyennes et les gammes de variation sont moins élevées. **Un tel résultat signifie, que dans le cas le plus pessimiste entre les deux groupes (i.e., 1 mois par saison, sensibilité maximale inférieure à 0.20), un échantillon sera déclassé ou surclassé à la classe de qualité contiguë à l'originale.**

Globalement, ces résultats montrent que les scénarios de dégradations constituant le 2^{ème} groupe semblent être de meilleures stratégies d'échantillonnage, dans l'optique d'une réduction du nombre d'échantillon (donc du coût total) dans le cadre de l'évaluation des habitats marins, pour les données de microscopie et dans une moindre mesure pour les données de metabarcoding. En effet, **l'impact de la dégradation des jeux de données sur le comportement de l'indicateur PH3 est moins marqué lorsque celui-ci est calculé à partir des données de metabarcoding.**

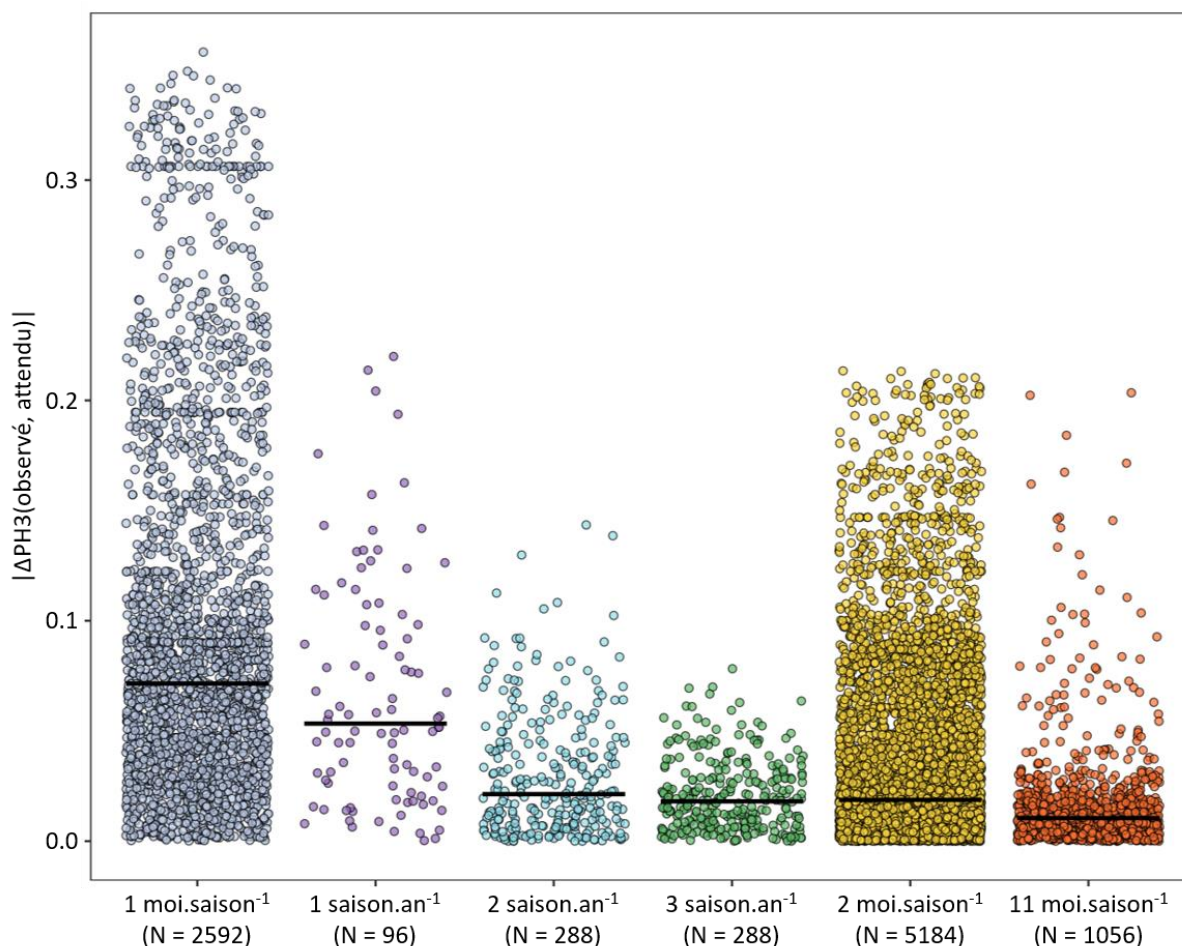


Figure 42. Sensibilité de l'indicateur PH3 à la fréquence d'acquisition des données. L'indicateur PH3 est inféré à partir des données de **microscopie** de la station SOMLIT-Astan à Roscoff. En ordonnées, la sensibilité est exprimée comme la valeur absolue de la différence entre les valeurs de PH3 observées (i.e., inférées à partir du jeu de données brutes) et les valeurs de PH3 attendues (i.e., inférées à partir des jeux de données dégradés). Six scénarios de dégradation ont été effectués par ré-échantillonnage aléatoire des données des séries temporelles. Les détails sur les procédures de dégradations des données figurent dans le paragraphe 7.1.3 (voir notamment dans la figure 18).

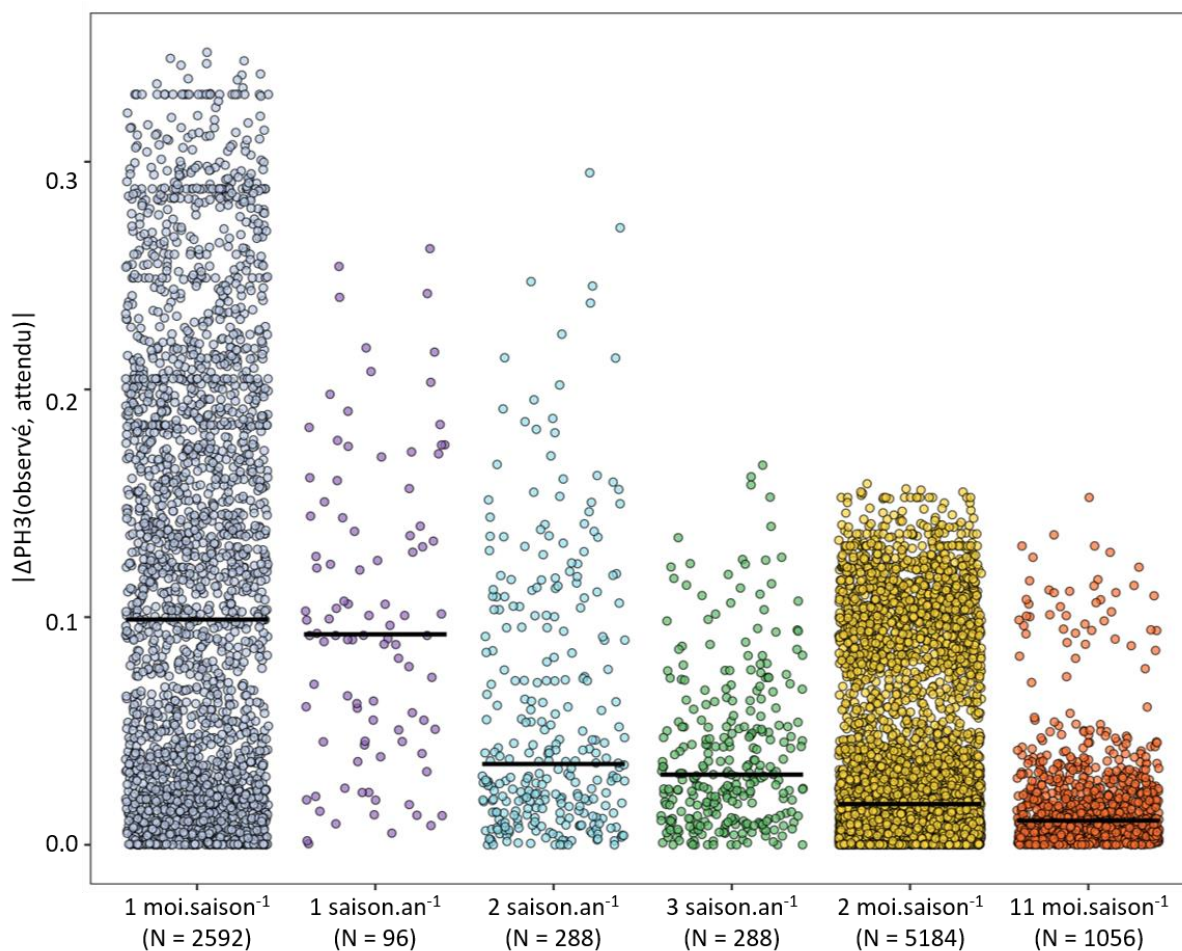


Figure 43. Sensibilité de l'indicateur PH3 à la fréquence d'acquisition des données. L'indicateur PH3 est inféré à partir des données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)** de la station SOMLIT-Astan à Roscoff.

7.3.6. Lien de l'indicateur PH3 avec l'environnement

Pour essayer de comprendre les différences existant entre les valeurs du *PH3* calculé à partir des données de microscopie et de metabarcoding dans les systèmes étudiés, nous avons analysé les liens avec l'environnement. Plus précisément, nous avons cherché à savoir si **les liens entre les variations du *PH3* et les variables environnementales étaient plus forts lorsque le calcul de l'indicateur était effectué à partir des données de metabarcoding par rapport aux données de microscopie**. Pour cela, les différences entre les similarités (estimées par le calcul de la distance de déformation dynamique temporelle) des dynamiques du *PH3* et des variables environnementales pour les 2 types de données ont été calculées (Figures 44, 45, 46 et 47). Les détails concernant la méthode utilisée figurent dans le paragraphe 7.1.4 (voir notamment dans la figure 18).

Il apparaît, pour l'ensemble des sites analysés que :

- La différence entre les valeurs du lien *PH3*/variables environnementales estimées par les deux types de données est généralement faible (min : -0.075, max : 0.050).

- Les valeurs du lien *PH3*/variables environnementales sont plus élevées pour les calculs basés sur les données de metabarcoding pour une majorité des variables environnementales étudiées (77% pour les stations SOMLIT-Astan et SOMLIT-C, 73% et 68% pour les stations B8 et H42 à la Marina du Château, respectivement).
- Les valeurs du lien *PH3*/variables environnementales assimilées à des pressions d'origine anthropique sont généralement plus élevées pour les calculs basés sur les données de metabarcoding. Les variables environnementales assimilées à des pressions d'origine anthropique sont i) les apports continentaux exprimés au travers des mesures de débits (e.g., Flux_Horn, Figure 44) et les concentrations en nutriments en fonction des débits (e.g., NO₃_flux_Horn, Figure 44) et ii) les concentrations en métaux lourds (Figures 46 et 47).

Ces résultats sont intéressants dans la mesure où les changements d'état des communautés microbiennes eucaryotes en lien avec les variables environnementales d'origine anthropique semblent être mieux mis en évidence avec les données de metabarcoding dans des stations de différents paysages marins.

A une échelle locale, et plus particulièrement à la Marina du Château, les patrons du lien *PH3*/variables environnementales sont similaires entre les stations B8 et H42. Un tel résultat peut paraître surprenant dans la mesure où ces deux stations sont positionnées le long d'un gradient de contamination métallique (allant de la station B8 vers la station H42) et des patrons distincts pouvaient être attendus. De plus, les valeurs du lien sont plus fortes pour quatre métaux lourds (Co, Cr, Pb, Ti) lorsque les calculs sont basés sur les données de microscopie. Pour le moment, aucun élément de réponse ne permet d'expliquer ce résultat.

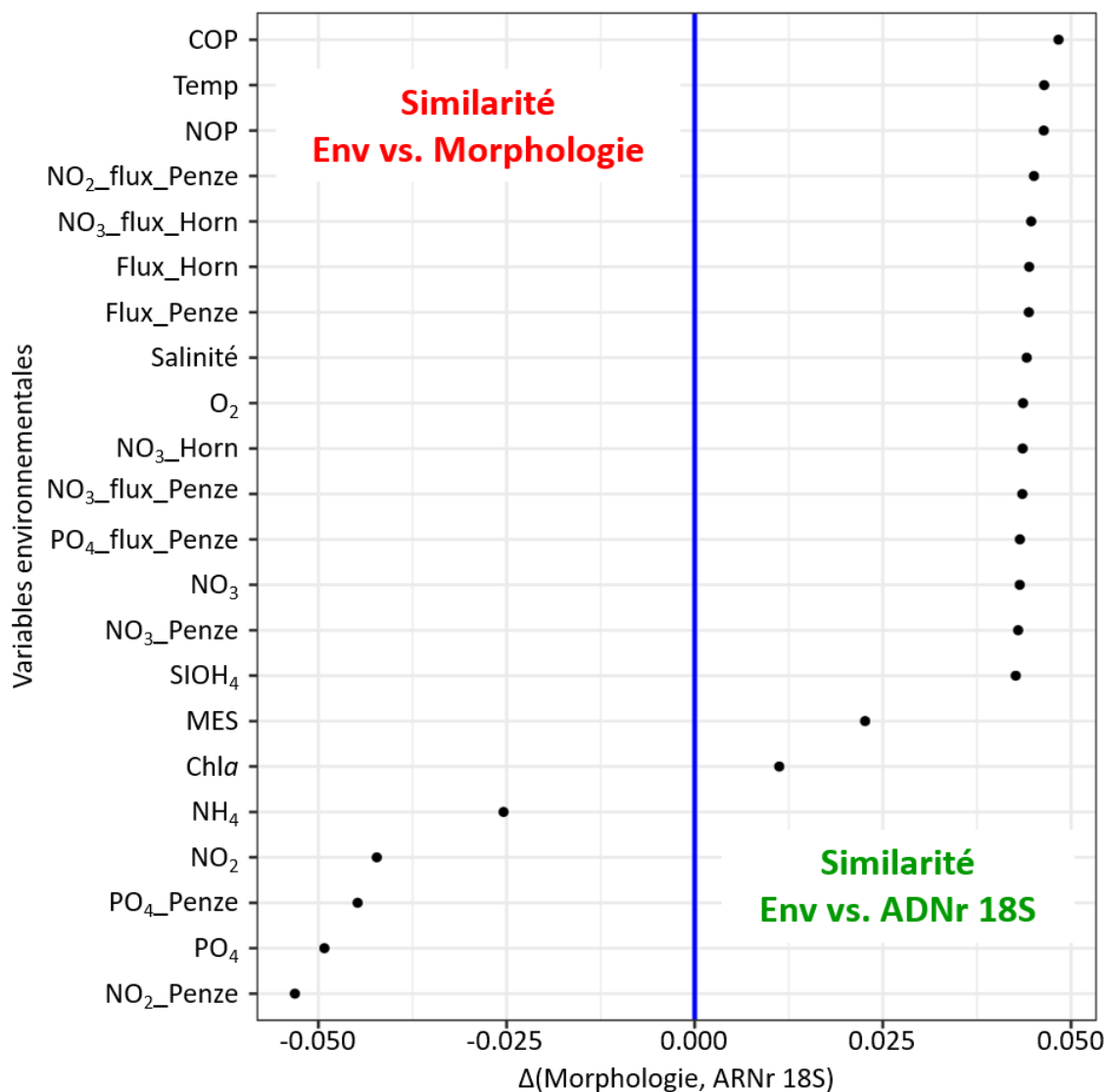


Figure 44. Lien entre l'indicateur PH3 et les paramètres environnementaux pour la station SOMLIT-Astan à Roscoff. La similarité entre la série temporelle de l'indicateur PH3 (données de **microscopie**) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)**. Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Les valeurs estimées du lien entre le PH3 et les variables environnementales sont ainsi supérieures pour un calcul à partir des données de metabarcoding lorsque la différence est supérieure à 0. À l'inverse, les valeurs estimées du lien entre le PH3 et les variables environnementales sont supérieures pour un calcul à partir des données de microscopie lorsque la différence est inférieure à 0.

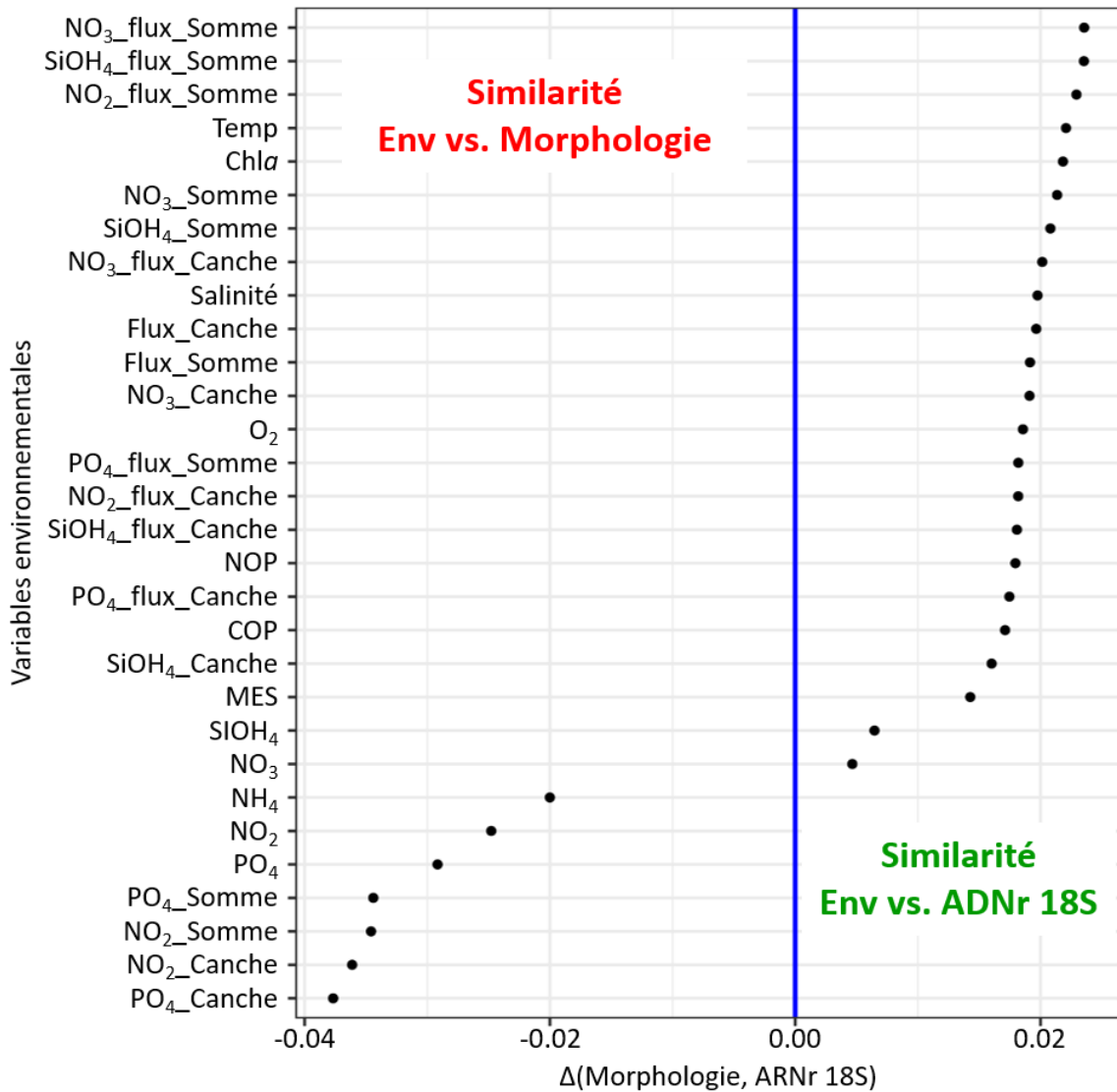


Figure 45. Lien entre l'indicateur PH3 et les paramètres environnementaux pour la station SOMLIT-C à Wimereux. La similarité entre la série temporelle de l'indicateur PH3 (données de **microscopie**) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)**. Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse.

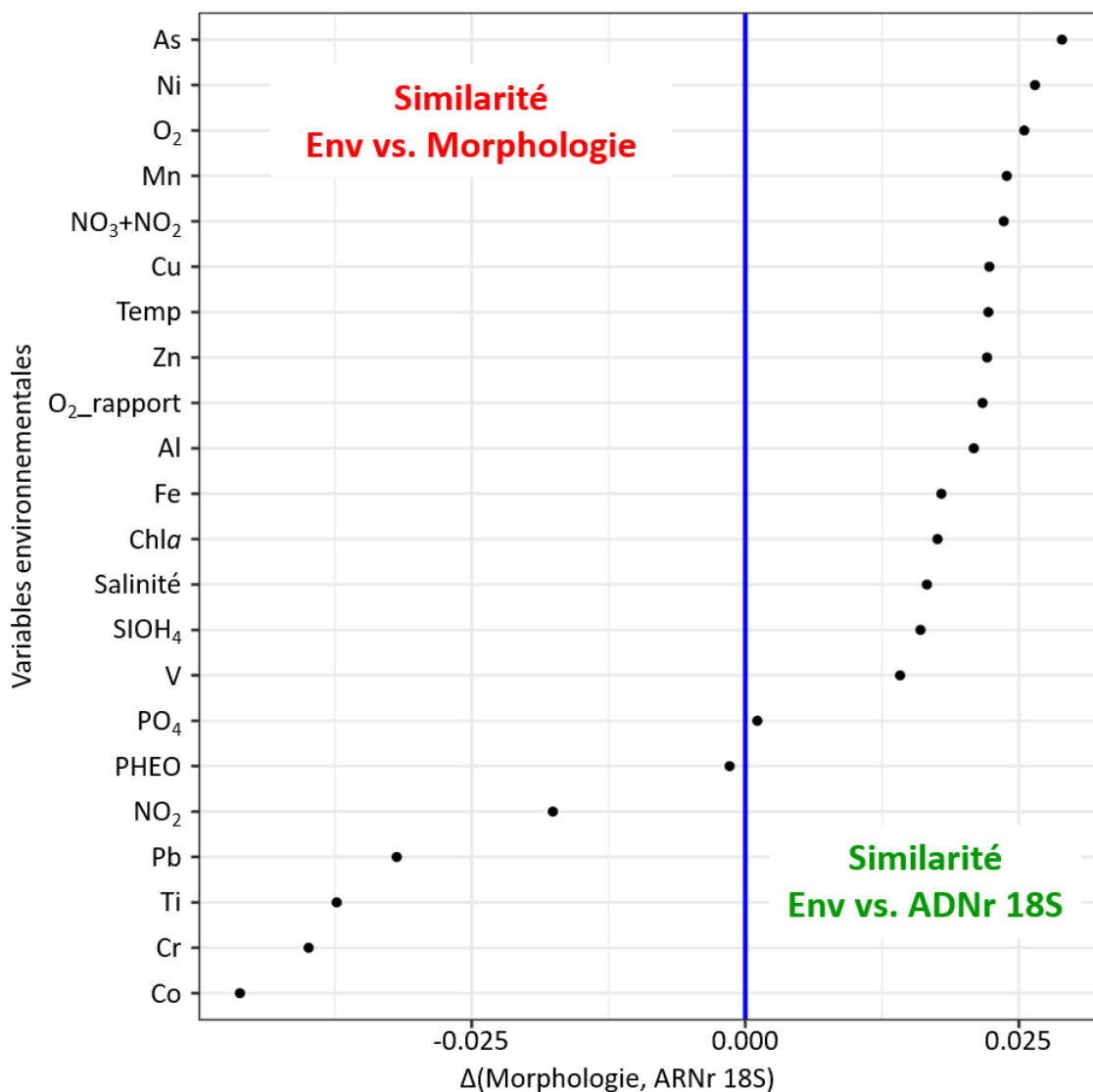


Figure 46. Lien entre l'indicateur PH3 et les paramètres environnementaux pour le site B8 à la Marina du château à Brest. La similarité entre la série temporelle de l'indicateur PH3 (données de **microscopie**) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (*dynamic time warping*). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)**. Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Al : Aluminium, As : Arsenic, Co : Cobalt, Cr : Chrome, Cu : Cuivre, Fe : Fer, Mn : Manganèse, Ni : Nickel, Pb : Plomb, Ti : Titane, V : Vanadium, Zn : Zinc.

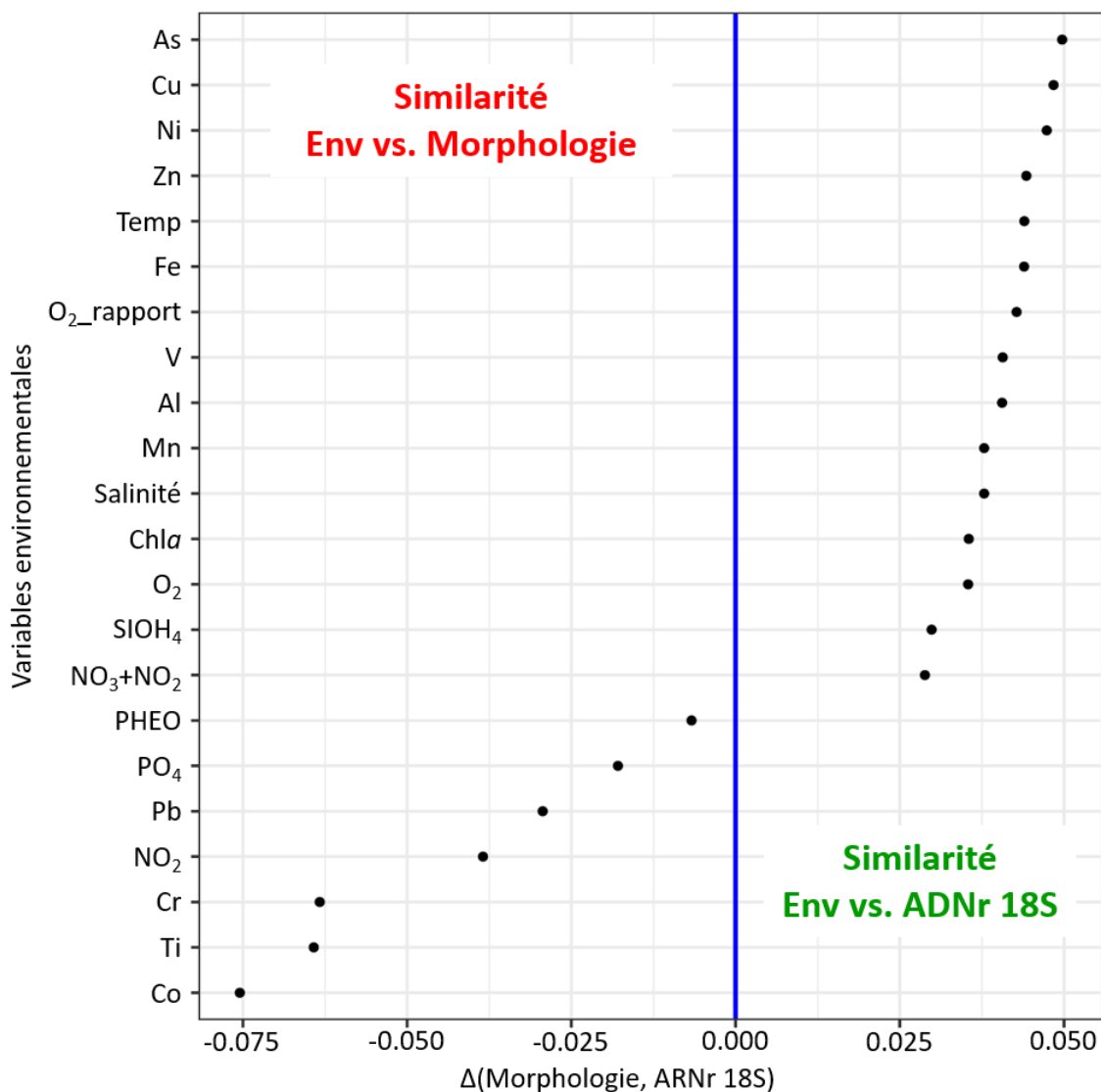


Figure 47. Lien entre l'indicateur PH3 et les paramètres environnementaux pour le site B8 à la Marina du château à Brest. La similarité entre la série temporelle de l'indicateur PH3 (données de **microscopie**) et celle de chaque paramètre environnemental est calculée sur la base de la distance de déformation dynamique temporelle (dynamic time warping). La même chose est réalisée avec l'indicateur PH3 calculé à partir des données de **metabarcoding (ARNr 18S, fraction de taille >3 μm)**. Pour chaque paramètre environnemental la différence entre ces deux distances est présentée en abscisse. Al : Aluminium, As : Arsenic, Co : Cobalt, Cr : Chrome, Cu : Cuivre, Fe : Fer, Mn : Manganèse, Ni : Nickel, Pb : Plomb, Ti : Titane, V : Vanadium, Zn : Zinc.

8. Action 4 : Protocoles, pipelines et conclusion pertinence des données génétiques

L'analyse des résultats (sections 6 et 7) a permis de dégager certaines conclusions quant à la pertinence d'utiliser des données de type metabarcoding dans la surveillance et l'évaluation du BEE des habitats pélagiques. Celles-ci sont regroupées dans le tableau 8, ci-dessous.

Ce tableau montre que les données de microscopie et de metabarcoding sont, de toutes évidences, complémentaires. S'il est possible d'accéder à la diversité de tous les compartiments microbiens via l'utilisation du metabarcoding, les données de comptage microscopique sont actuellement les seules qui permettent l'accès à l'abondance par unité de volume, ainsi qu'à la biomasse, deux paramètres très importants dans le cadre de la surveillance des écosystèmes marins.

Tableau 8. Conclusions sur la pertinence d'utiliser les données de metabarcoding pour la surveillance des habitats pélagiques.

Paramètre considéré	Metabarcoding	Microscopie classique (méthode d'Utermohl, telle qu'utilisée dans les SNO actuels)
Possibilité de prendre en compte tous les groupes fonctionnels et classes de taille	Accès à toute la diversité > 0.2 µm (procaryotes, eucaryotes), via l'utilisation de marqueurs génétiques adaptés	Accès à une fraction de la diversité (cellules de taille ≥ 10 µm) La microscopie ne permet pas d'accéder : <ul style="list-style-type: none"> • A la diversité taxonomique des procaryotes • Plus généralement aux classes de taille les plus petites qui peuvent dominer les communautés • Aux taxons présentant peu de caractères morphologiques distinctifs • A la plupart des microorganismes endosymbiotiques ou endoparasites
Sensibilité	<ul style="list-style-type: none"> • Accès aux taxons « rares » (d'autant plus si les volumes échantillonnés et la profondeur de séquençage sont importants) (Pedrós-Alió 2012) • Accès aux taxons cryptiques ou pseudo-cryptiques 	<ul style="list-style-type: none"> • Accès peu aisé aux taxons « rares » : la plupart du temps, les volumes analysés ne permettent que de détecter les espèces les plus abondantes (Cermeno et al. 2014) • Pas d'accès aux taxons cryptiques, accès peu aisé

		aux taxons pseudo-cryptiques
Quantifier abondance, biomasse, contribution relative.	<p>Méthode dite semi-quantitative</p> <ul style="list-style-type: none"> • Accès à des abondances relatives de nombres de séquences • Pas d'accès aisé aux abondances absolues • Pas d'accès aisé à la biomasse 	<p>Méthode quantitative</p> <ul style="list-style-type: none"> • Accès aux abondances absolues (cellules par unité de volume) mais uniquement pour les taxons les plus abondants • Accès à la biomasse (via des proxys de biomasse par cellule par espèce)
Reproductibilité	<ul style="list-style-type: none"> • Méthode standardisable de l'échantillonnage à la production des tables d'ASV × sites (dates). Les étapes allant de l'échantillon à la production des librairies ASVs et aux données brutes de séquences sont effectuées à l'aide de kits sur des plateformes. Des pipelines d'analyse sont fournis dans le cadre du projet INDIGENE pour toutes les étapes qui permettent à partir des données brutes, d'obtenir des valeurs pour les indicateurs <i>PH1</i> et <i>PH3</i> • Etant donnée la rapidité de l'évolution des technologies associées à la production de séquences ainsi que des outils bioinformatiques, il serait important, si ces technologies sont utilisées pour caractériser les changements dans les communautés, de veiller à conserver des protocoles strictement identiques ou comparables 	<ul style="list-style-type: none"> • Méthode standardisable pour la plupart des étapes • Les étapes d'identification des taxons et de dénombrement font intervenir des analystes experts en taxonomie et donc des possibles biais « observateur ». La formation des experts et le maintien de ces expertises dans les laboratoires est indispensable de manière à limiter ces biais
Indices de diversité alpha	<ul style="list-style-type: none"> • Possibilité de calculer les indices classiques (en prenant les abondances de séquences et ASVs au lieu des abondances cellulaires et taxons) • Possibilité d'accéder à la diversité phylogénétique (non traité dans le projet INDIGENE) 	<ul style="list-style-type: none"> • Possibilité de calculer les indices classiques • Pas de prise en compte des taxons rares (non détectés si le volume échantillonné est faible)
Indices de diversité bêta	<ul style="list-style-type: none"> • Possibilité de calculer les indices classiques (en prenant les abondances de séquences et ASVs au lieu des abondances cellulaires et taxons) • Etant donné la nature compositionnelle des données (abondances relatives) certaines approches posent des problèmes ou sont invalides mathématiquement 	

Calcul des indicateurs préconisés par la DCSMM

- Les tests effectués dans le cadre du projet INDIGENE permettent de valider la possibilité de calculer des indices de diversité bêta. Il est probable que d'autres indices pourraient être testés, et notamment ceux préconisés pour des données compositionnelles
- **PH1 :**
 - Calcul aisé du **PH1** (basé sur des ratios d'abondances et donc adapté à des données compositionnelles) à partir des données d'abondances de séquences pour les 2 groupes choisis
 - Possibilité de calculer le **PH1** pour un très grand nombre de combinaisons de groupes fonctionnels/taxonomiques et ceci pour les procaryotes ou les eucaryotes (parasites /non parasites, autotrophes / hétérotrophes, diatomées pennées / diatomées centrées, ...). Ceci constitue une plus-value très importante des données de metabarcoding. Les tests effectués dans le cadre du projet INDIGENE permettent de valider la faisabilité et la pertinence de cette approche. La plus-value des données de metabarcoding est très clairement la possibilité d'analyser le comportement du **PH1** en prenant en compte de très nombreux couples de groupes fonctionnels ou taxonomiques
 - **PH2 :** Les données de metabarcoding ne sont pas adaptées au calcul du **PH2**
 - **PH3 :** Calcul relativement aisé du **PH3**. Les tests effectués dans le cadre du projet INDIGENE permettent de valider la faisabilité et la pertinence de cette approche lorsque les EQR sont calculés par rapport à des états de référence par saison, ceci pour prendre en compte la variabilité saisonnière. La plus-value des données de metabarcoding est ici la prise en compte du compartiment microbien dans son ensemble et non seulement des fractions identifiables au microscope
- **PH1 :**
 - Calcul aisé du **PH1** (basé sur des ratios d'abondances)
 - Possibilité de calculer le **PH1** pour un nombre moins important de combinaisons de groupes fonctionnels/taxonomiques
 - **PH2 :** Les données de microscopie permettent d'obtenir des valeurs de biomasse grâce à l'utilisation de proxys de biomasse par cellule
 - **PH3 :** Calcul relativement aisé du **PH3**. Cependant, de très nombreuses espèces microbiennes ne sont pas prises en compte

9. Références bibliographiques

- Alric B., C. Jeanthon, L. Maignien, U. Christaki, & N. Simon. (2021). Données génétiques et indicateurs de biodiversité. Rapport d'avancement, Actions 1 et 2. Octobre 2021, 37pp.
- Baker M.E., & King R.S. (2010). A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods in Ecology and Evolution* **1**, 25–37.
- Beaugrand G. (2012). Unanticipated biological changes and global warming. *Marine Ecological Progress Series* **445**, 293–301.
- Bedford J., Johns D., Greenstreet S., & McQuatters-Gollop A. (2018). Plankton as prevailing conditions: A surveillance role for plankton indicators within the Marine Strategy Framework Directive. *Marine Policy* **89**, 109–115.
- Bedford J., Ostle C., Johns D. G., Budria A., & McQuatters-Gollop A. (2020). The influence of temporal scale selection on pelagic habitat biodiversity indicators. *Ecological Indicators* **114**, 106311.
- Brylinski J.M., Lagadeuc Y., Gentilhomme V., Dupont J.P., Lafite R., Dupeuble P.A., Huault M.F., & Auger Y. (1991). Le "fleuve côtier": Un phénomène hydrologique important en Manche orientale. Exemple du Pas-de-Calais. *Oceanologica Acta* **11**, 197–203.
- Callahan B. J., McMurdie P. J., Rosen M. J., Han A. W., Johnson A. J. A., & Holmes S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581–583.
- Caracciolo M., Rigaut-Jalabert F., Romac S., Mahé F., Forsans S., Gac J.-P., Arsenieff L., Manno M., Chaffron S., Cariou T., Hoebeke M., Bozec Y., Goberville E., Le Gall F., Guilloux L., Baudoux A.-C., Vargas C., Not F., Thiébaud E., Henry N., & Simon N. (2021). Seasonal dynamics of marine protist communities in tidally mixed coastal waters. *Molecular Ecology* **31**, 3761–3783.
- Caruso V., Song X., Asquith M., & Karstens L. (2019). Performance of Microbiome Sequence Inference Methods in Environments with Varying Biomass. *mSystems* **4**, e00163-18.
- Cermeno P., Teixeira I.G., Branco M., Figueiras F.G., & Marañón. (2014). Sampling the limits of species richness in marine phytoplankton communities. *Journal of Plankton Research* **36**, 1135–1139.
- Chalar G. (2009). The use of phytoplankton patterns of diversity for algal bloom management. *Limnologia* **39**, 200–208.
- Cordier T., Esling P., Lejzerowicz F., Visco J., Ouadahi A., Martins C., Cedhagen T., & Pawlowski J. (2017). Predicting the ecological quality status of marine environments from eDNA metabarcoding data using supervised machine learning. *Environmental Science and Technology* **51**, 9118-9126.
- Cormack R., & Ord J. (1979). Spatial and temporal analysis in ecology. International Co-operative Publishing House, Fairland, Maryland, USA.
- DCSMM, 2008/56/CE. Directive du Parlement Européen et du Conseil du 17 juin 2008 établissant un cadre d'action communautaire dans le domaine de la politique pour le milieu marin (directive-cadre « stratégie pour le milieu marin »), Chapitre 15, Volume 026, pp. 136-157.
- Diniz-Filho J. A., Siqueira T., Pádua A. A., Rangel T. F., Landeiro V. L., & Bini L. M. (2012). Spatial autocorrelation analysis allows disentangling the balance between neutral and niche processes in metacommunities. *Oikos* **121**, 201-210.
- Dray S. Bauman D., Blancher G., Borcard D., Clappe S., Guenard G. Jombart T., Larocque G., et al. (2021). Adepstail: Multivariate Spatial Analysis. R package version 0.3-14. <https://CRAN.R-project.org/package=adespatial>.
- Dray S. Legendre P., & Peres-Neto P. R. (2006). Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**, 482-493.
- Ducklow H. (2008). Microbial services: Challenges for microbial ecologists in a changing world. *Aquatic Microbial Ecology* **53**, 13-19.
- Duffy J. E., Godwin C. M., & Cardinale B. J. (2017). Biodiversity effects in the wild are common and as strong as key drivers of productivity. *Nature* **549**, 261-264.

- Duflos M., Wacquet G., Aubert A., Rombouts I., Mialet B., Devreker D., Lefebvre A., & Artigas L.F. (2018). Evaluation de l'état écologique des habitats pélagiques en France métropolitaine. Rapport scientifique pour l'évaluation 2018 au titre du descripteur 1 de la DCSMM. 333pp.
- Facca C., Bernardi Aubry F., Socal G., Ponis E., Acri F., Bianchi F., Giovanardi F., & Sfriso A. (2014). Description of a Multimetric Phytoplankton Index (MPI) for the assessment of transitional waters. *Marine Pollution Bulletin* **79**, 145–154.
- Falkowski P. G., Fenchel T., & Delong E. F. (2008). The microbial engines that drive earth's biogeochemical cycles. *Science* **320**, 1034-1038.
- Gaston K.J., & Spicer J.I. (2004). Biodiversity : An Introduction. Wiley-Blackwell, Hoboken.
- Goodwin K.D., Thompson L.R., Duarte B., Kahlke T., Thompson A.R., Marques J.C., & Caçador I. (2017). DNA Sequencing as a Tool to Monitor Marine Ecological Status. *Frontiers in Marine Science* **4**, 107.
- Gowen R.J., McQuatters-Gollop A., Tett P., Best M., Bresnan E., Castellani C., Cook K., Forster R., Scherer C., & McKinney A. (2011). The Development of UK Pelagic (Plankton) Indicators and Targets for the MSFD. A Report of a workshop held at AFBI, Belfast 2-3rd June 2011, 41pp.
- Gowen R.J., Tett P., McQuatters-Gollop A., Forster R., Scherer C., Bresnan E., Cook K., Atkinson A., Best M., Baptie M., Keeble S., & McCullough G. (2013). Part One of the UK Marine Strategy for the Marine Strategy Framework Directive. Plankton indicators: Development of the Life-form and State-Space Method. A report to Defra on a pre-project workshop held at the Agri-Food and Biosciences Institute, Belfast. March 5th- 6th 2013.
- Grattepanche J. D., Breton E., Brylinski J. M., Lecuyer E., & Christaki U. (2011). Succession of primary producers and micrograzers in a coastal system dominated by *Phaeocystis* globose blooms. *Journal of Plankton Research* **33**, 37-50.
- Guérin L., Feunteun E., Gonson C., Laurand S., Lavesque S., Lejart M., & You H. (2012). Méthodes d'évaluation de l'état écologique, caractérisation du « bon état » pour la DCSMM et recommandations pour les travaux futurs. Descripteurs 1 et 2.
- Guilloux L., Rigaut-Jalabert F., Jouenne F., Ristori S., Viprey M., Not F., Vaulot D., & Simon N. (2013). An annotated checklist of Marine Phytoplankton taxa at the SOMLIT-Astan time series off Roscoff (Western English Channel, France): Data collected from 2000 to 2010. *Cahiers de Biologie Marine* **54**, 247-256.
- Halpern B. S., Walbridge S., Selkoe K. A., Kappel C. V., Micheli F., D'Agrosa C., Bruno J. F., Casey K. S., et al. (2008). A global map of human impact on marine ecosystems. *Science* **319**, 948-952.
- Hautier Y., Tilman D., Isbell F., Seabloom E. W., Borer E. T., & Reich P. B. (2015). Anthropogenic environmental changes affect ecosystem stability via biodiversity. *Science* **348**, 336-340.
- Hegg J.C., & Kennedy B.P. (2021). Let's do the time warp again : non-linear time series matching as a tool for sequentially structured data in ecology. *Ecosphere* **12**, e03742.
- Hering D., Borja A., Jones J.I., Pont D., Boets P., Bouchez A., Bruce K., Drakare S., Hänfling B., Kahlert M., Leese F., Meissner K., Mergen P., Reyjol Y., Segurado P., Vogler A., & Kelly M. (2018). Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. *Water Research* **138**, 192–205.
- Howarth R.W., Anderson D.B., Cloern J.E., Elfring C., Hopkinson C.S., Lapointe B., Malone T., Marcus N., McGlathery K., Sharpley A.N., & Walker D. (2000). Nutrient pollution of coastal rivers, bays, and seas. *Isseus in Ecology* **7**, 2–15.
- Hutchins D.A., & Fu F. (2017). Microorganisms and ocean global change. *Nature Microbiology* **2**, 17058.
- Kenworthy J.M., Rolland G., Samadi S., & Lejeune C. (2018). Local variation within marinas: Effects of pollutants and implications for invasive species. *Marine Pollution Bulletin* **133**, 96-106.
- Kirby R.R., & Beaugrand G. (2009). Trophic amplification of climate warming. *Proceedings of the Royal Society B* **276**, 4095–4103.

- Le Pape O., & Menesguen A. (1997). Hydrodynamic prevention of eutrophication in the Bay of Brest (France), a modelling approach. *Journal of Marine Systems* **12**, 171–186.
- Legendre P. & De Cáceres M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters* **16**, 951–963.
- Legendre P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* **74**, 1659-1673.
- Legendre P., & Gauthier O. (2014). Statistical methods for temporal and space – time analysis of community composition data. *Proceedings of the Royal Society B* **281**, 20132728.
- Margalef R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanologica acta* **1**, 493–509.
- McQuatters-Gollop A., Atkinson A., Aubert A., Bedford J., Best M., Bresnan E., Cooke K., Devlin M., Gowen R.J., Johns D.G., Machairopoulou M., McKinney A., Mellor A., Ostle C., Schererh C., & Tett P. (2019). Plankton lifeforms as a biodiversity indicator for regional-scale assessment of pelagic habitats for policy. *Ecological Indicators* **101**, 913–925.
- Oksanen J., Blanchet G., Friendly M., Kindt R., Legendre P., McGlinn D., O’Hara R. B., Simpson G. L., Solymos P., Stevens M. H. H., Szoecs E., & Wagner H. vegan: Community Ecology Package. R package version 2.5-6, 2019. <https://CRAN.R-project.org/package=vegan>.
- Ostle C., Artigas L.F., Aubert A., Budria A., Graham G., Johansen M., Johns D., Padegimas B., Rombouts I., & McQuatters-Gollop A. (2017). Programming outputs for constructing the plankton lifeform indicator from disparate data types. Contribution to the EU Co-financed EcApRHA project (Applying an ecosystem approach to (sub) regional habitat assessments), Deliverable report 1.1.
- Parada A., D. M. Needham, & J. A. Fuhrman. (2016). Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time-series and global field samples. *Environmental Microbiology* **18**, 1403–1414.
- Pedrés-Alió C. (2012). The Rare Bacterial Biosphere. *Annual Review of Marine Science* **4**, 449–466.
- Peperzak L., Colijn F., Gieskes W. W. C., & Peeters J. C. H. (1998). Development of the diatom – Phaeocystis spring bloom in the Dutch coastal zone of the North Sea: the silicon depletion versus the daily irradiance threshold hypothesis. *Journal of Plankton Research* **20**, 517–537.
- Pereira, H. M.; Navarro, L. M.; Martins, I. S. Global biodiversity change: The bad, the good, and the unknown. *Annual Review of Environment and Resources* **37**, 25-50.
- Pimm S. L., Jenkins C. N., Abell R., Brooks T. M., Gittleman J. L., Joppa L. N., Raven P. H., Roberts C. M., & Sexton J. O. (2014). The biodiversity of species and their rates of extinction, distribution, and protection. *Science* **344**, 1246752.
- Reynolds C.S. (2006). *The Ecology of Phytoplankton*. Cambridge University Press.
- Rivera S. F., Vasselon V., Bouchez A., & Rimet F. (2020). Diatom metabarcoding applied to large scale monitoring networks: Optimization of bioinformatics strategies using Mothur software. *Ecological Indicators* **109**, 105775.
- Rombouts I., Simon N., Aubert A., Carriou T., Feunteun E., Guérin L., Hoebeke M., McQuatters-Gollop A., Rigaut-Jalabert F., & Artigas L.F. (2019). Changes in marine phytoplankton diversity: assessment under the Marine Strategy Framework Directive. *Ecological Indicators* **102**, 265-277.
- Sakoe H, & Chiba S. (1971). A dynamic programming approach to continuous speech recognition. Proceeding of International Congress on Acoustics, Budapest, Hungary, Paper 20C-13, 1971.
- Sakoe H., & Chiba S. (1978). dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing* **26**, 43–49.
- Sala O.E., Chapin III F. S., Armesto J. J., Berlow E. L., Bloomfield J., Huenneke L., Jackson R. B., Kinzig A., et al. (2000) Biodiversity: Global biodiversity scenario for the year 2100. *Science* **287**, 1770-1774.
- Schapira M., Vincent D., Gentilhomme V., & Seuront L. (2008). Temporal patterns of phytoplankton assemblages, size spectra and diversity during the wane of a *Phaeocystis globosa* spring bloom in

- hydrologically contrasted coastal waters. *Journal of the Marine Biological Association of the United Kingdom* **88**, 649–662.
- Stoeck T, Bass D, Nebel M, Christen R, Jones MDM, Breiner HW, *et al.* (2010). Multiple marker parallel tag environmental DNA sequencing reveals a highly complex eukaryotic community in marine anoxic water. *Molecular Ecology* **19**, 21–31.
- Tett P. (2016). Guide to the Plankton Index method and software, v.3.992. Scottish Association for Marine Science, Oban. 33pp. + appendices.
- Tett P., Carreira C., Mills D.K., van Leeuwen S., Foden J., Bresnan E., & Gowen R.J. (2008). Use of a Phytoplankton Community Index to assess the health of coastal waters. *ICES Journal of Marine Science* **65**, 1475-1482.
- Tew-Kai E., Quilfen V., Cachera M., & Boutet M. (2020). Dynamic Coastal-Shelf Seascapes to Support Marine Policies Using Operational Coastal Oceanography: The French Example. *Journal of Marine Science and Engineering* **8**, 585.
- Tilman D., & Kareiva P. (1997). Spatial ecology: the role of space in population dynamics and interspecific interactions. Princeton University Press, Princeton, New Jersey, USA.
- Turner J.T. (2004). The importance of small planktonic copepods and their roles in pelagic marine food webs. *Zoological Studies* **43**, 255-266
- Vincent D., Luczak C., & Sautour B. (2002). Effects of a brief climatic event on zooplankton community structure and distribution in Arcachon Bay (France). *Journal of Marine Biological Association of U.K.* **82**, 21–30.
- Walsh J. R., Carpenter S. R., & Vander Zanden M. J. (2016). Invasive species triggers a massive loss of ecosystem services through a trophic cascade. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 4081-4085.
- Ware D.M., & Thomson R.E. (2005). Bottom-Up Ecosystem Trophic Dynamics Determine Fish Production in the Northeast Pacific. *Science* **308**, 1280–1284.
- Wolkovich E.M., Cook B.I., McLauchlan K.K., & Davies T.J. (2014). Temporal ecology in the Anthropocene. *Ecology Letters* **17**, 1365–1379.
- Worm B., Barbier E. B., Beaumont N., Duffy J. E., Folke C., Halpern B. S., Jackson J. B., Lotze H. K. *et al.* (2006). Impacts of biodiversity loss on ocean ecosystem services. *Science* **314**, 787-790.
- Zimmermann J., Glöckner G., Jahn R., Enke N., & Gemeinholzer B. (2015). Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. *Molecular Ecology Resources* **15**, 526-542.