

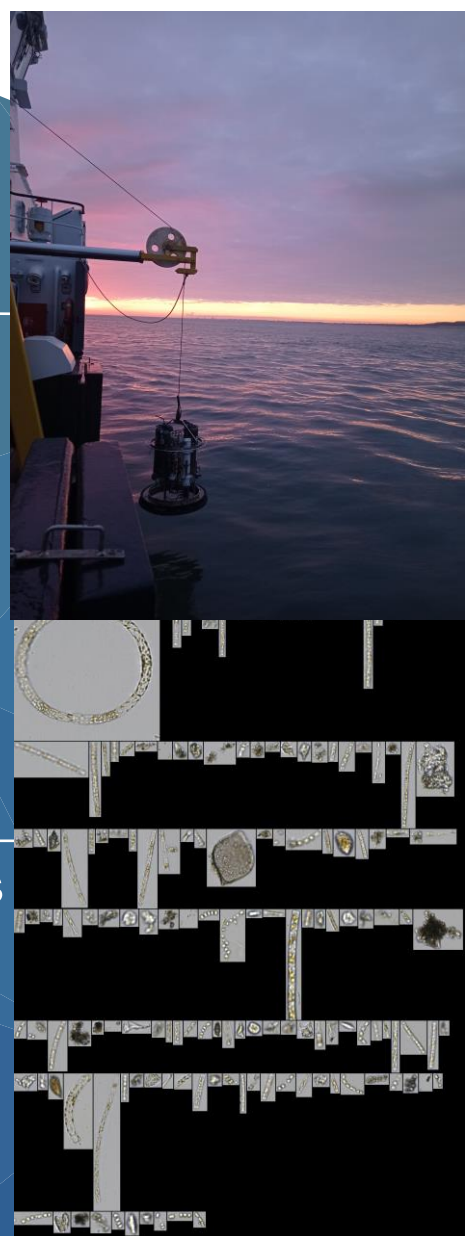
Rapport de stage individuel

5^{ème} année

Contribution à l'amélioration de
l'approche d'imagerie
automatisée pour l'étude du
phytoplancton marin

Laboratoire d'Océanologie et Géosciences
– UMR 8187

32 avenue Foch - 62930 Wimereux



Tuteur entreprise :
Luis Felipe Artigas
Enseignant-chercheur



Florine Verhaeghe
IMA

Tuteur académique :
Pierre Peeters



2021-2022

REMERCIEMENTS

Je tiens à remercier mon encadrant de stage **Luis Felipe Artigas** pour avoir fait en sorte que mon stage puisse avoir lieu. Merci pour cette expérience qui est, je l'espère, le début d'un épanouissement professionnel et personnel. Merci également à mon tuteur pédagogique, **Pierre Peeters**, qui m'a fait confiance pour ce stage de dernière année d'ingénieur.

Merci au personnel du Laboratoire Océanologique et de Géosciences pour m'avoir accueilli quotidiennement et au Directeur, **Hubert Loisel**, pour avoir accepté ce stage spontané. Ce stage s'inscrit au sein du Projet CPER IDEAL que je remercie pour les financements, tout comme les projets support dont le CPER IDEAL et le projet H2020 JERICO S3 sans lesquels ce stage n'aurait pas pu exister.

Un grand merci aux collaborateurs, **Guillaume Wacquet** pour son aide et ses traitements des données sans lesquels je n'aurai pu travailler, aux marins du « Sepia II », **Christophe** et **Noël**, et au technicien supérieur, **Vincent Cornille**, avec lesquels j'ai pu passer que de bons moments en mer lors des échantillonnages de la radiale. Je suis également reconnaissante à l'équipe du VLIZ à bord du « Simon Stevin » de m'avoir acceptée pendant cette campagne de 4 jours qui a été très enrichissante.

Merci à **Manon Ryckman** et **Eva Gruson**, stagiaire durant cette période, pour leur travail et relai au sein du laboratoire et sur l'échantillonnage.

Un dernier remerciement, et pas des moindres, pour les membres de l'équipe Phyto-CytoFluIm. Merci à **Alexandre Epinoux**, pour sa gentillesse et pour m'avoir appris à aimer le vrai café fait avec amour en Vésuve. Je m'en sors grandi. Je remercie **Arnaud LoucharT**, pour sa bienveillance et ses conseils avisés qui m'ont aidé tout au long de mon stage. Merci pour tes passages dans le « bureau du fond » et ta facilité à réchauffer nos chaises de bureau. Gros Big Up à **Clémentine Gallot** et **Zéline Hubert**, pour ces moments et ces souvenirs, de rire, de pleurs, de fatigue, voix cassées et bien plus, merci d'avoir fait que tolérer ma présence dans ce bureau emblématique. J'ai la chance de vous avoir, pour le pire et le meilleur, merci. Evidemment, je remercie l'homme sans lequel le « bureau des -ine » et les kévinettes ne pourraient exister, **Kévin Robache**, merci de m'avoir subi et soutenu durant ces longues heures aux laboratoires avec des blagues plus que douteuses. Sans oublier **Perrine Salvi**, une présence dans la team phyto bien appréciée.

Merci à ma famille qui ont été présent lors de mon stage et à mes amis qui ont accepté de visiter Boulogne-sur-Mer.

Table des matières

REMERCIEMENTS	2
Liste des Abréviations - Glossaire.....	7
1 Introduction.....	1
1.1 Présentation de la structure.....	1
1.2 Contexte	2
1.2.1 Intérêt pour ce stage	2
1.2.2 Projets intégrant le stage.....	2
1.3 Objectif du stage	3
1.3.1 Missions et problématiques abordées	3
1.4 Sujet du stage	3
1.4.1 Importance du suivi phytoplanctonique	3
1.4.2 Histoire de l'imagerie planctonique	4
1.4.1 Présentation du site d'étude.....	4
1.4.2 Paramètres environnementaux.....	5
1.4.3 Localisation des données de suivis.....	7
2 Méthodologie	8
2.1 Outils de travail, le FlowCam.....	8
2.1.1 Fonctionnement	8
2.1.2 Données en sortie.....	9
2.1.3 Concept de Transfer Learning et Deep Learning.....	9
2.2 Savoirs et compétences mis en œuvre.....	10
2.2.1 Echantillonnage	10
2.2.2 Prise en main de l'appareil et technique de laboratoire.....	11
2.2.1 Taxonomie	12
2.2.2 Travail d'analyse automatisé.....	12
2.3 Amélioration du protocole d'analyse au sein du LOG.....	12
2.3.1 Gain de temps	12
2.3.2 Renforcement de la banque d'images	13
2.4 Gestion et traitements de données	13
2.4.1 Configuration du FlowCam pour l'acquisition de la donnée	13
2.4.2 Création du set d'apprentissage.....	14
2.4.3 Outils d'évaluation du set d'apprentissage.....	14
2.5 Analyse des données temporelles et spatiales	15
3 Résultats.....	15
3.1 Validation du set d'apprentissage.....	15

3.1.1	Grossissement 100	15
3.1.2	Grossissement 40	19
3.1.3	Comparaison des apprentissages entre les deux grossissements	22
3.2	Résultats des analyses sur le point R1 de la radiale DYPHYRAD du printemps 2021 à l'été 2022.	23
3.2.1	L'analyse des particules du FlowCam en x4	23
3.2.2	L'analyse des particules du FlowCam en x10	23
3.3	Résultats des analyses temporelles des communautés phytoplanctoniques.....	24
3.3.1	Evolution saisonnière de l'abondance phytoplanctonique sur le point R1 en 2021 et 2022	24
4	Discussion	27
4.1	Méthode	27
4.2	Résultats	28
4.3	Avantages et inconvénients entre les deux grossissements	28
5	Conclusion	29
5.1	Synthèse du travail	29
5.2	Perspectives.....	29
5.3	Bilan des compétences et connaissances.....	30
	BIBLIOGRAPHIE	31

Table des figures

Figure 1: Salinité saisonnière a 5 m de profondeur (sources : climatologie mensuelle manche - golfe de Gascogne boby clim_v2.3, 2010).....	6
Figure 2: Carte des suivis et campagnes effectués en 2022.....	8
Figure 3: fonctionnement du flowcam®, source : fluidimaging, Guide d'utilisation (Wacquet, 2017)...	9
Figure 4: Sepia II (a) et bouteille Niskin sur la CTD (b)	10
Figure 5: Fonctionnement du FlowCam au laboratoire	11
Figure 6: Matrice de confusion du 1 ^{er} set à l'objectif X10 au FlowCam	18
Figure 7: exemple de vignettes au x10 (a. <i>Phaeocystis globosa</i> Cox, b. <i>Phaeocystis globosa</i> , c. <i>Chaetoceros socialis</i> , d. <i>Chaetoceros spp</i> , e. <i>Rhizosolenia imbricata</i> , f. <i>Leptocylindrus danicus</i>)	18
Figure 8: Matrice de confusion du 2 ^e set à l'objectif X10 au FlowCam	19
Figure 9: Matrice de confusion du 1 ^{er} set a l'objectif x4 au flowcam	21
Figure 10: Exemple de vignette au X4 (a. <i>Leptocylindrus danicus</i> , b. <i>Pseudo-nitzschia</i> , c. <i>Rhizosolenia imbricata</i>)	21
Figure 11: Matrice de confusion du 2e set a l'objectif x4 au flowcam.....	22
Figure 12: Graphique représentant la probabilité d'appartenance (à gauche pour le set x4, à droite le set x10)	22
Figure 13: Graphique des particules passées sur le FlowCam au grossissement 40 de mars 2021 à juillet 2022 au point R1	23
Figure 14: Graphique des particules passées sur le FlowCam au grossissement 100 de mars 2021 à juillet 2022 au point R1	24
Figure 15: Graphique de l'abondance absolue du phytoplancton au grossissement 40 de mars 2021 à juillet 2022 au point R1	25
Figure 16: Graphique de l'abondance relative du phytoplancton au grossissement 40 de mars 2021 à juillet 2022 au point r1	25
Figure 17: Graphique de l'abondance absolue du phytoplancton au grossissement 100 de mars 2021 à juillet 2022 au point r1	26
Figure 18: Graphique de l'abondance relative du phytoplancton au grossissement 100 de mars 2021 à juillet 2022 au point r1	26

Table des annexes

Annexe 1 : carte des aires marines protégées de la sous-region marine manche orientale-mer du nord	39
Annexe 2 : Synthèse des enjeux socio-économiques forts et majeurs de la façade maritime Manche Est – mer du Nord	40
Annexe 3: Marnage en Manche (Sources : modèle de marée CSTFRANCE du SHOM, ESRI, 2010).	42
Annexe 4: Circulation moyenne en Manche (Sources : IGN, ESRI, SHOM, OSPAR, 2010).	42
Annexe 5: Station MAREL, digue Carnot à Boulogne-sur-Mer	43
Annexe 6: Tableau des analyses effectuées par point pour le suivi DYPHYRAD	43
Annexe 7: Protocole FlowCam (Source : Elise Caillard).....	45
Annexe 8: Classification des groupes pour le set d'apprentissage en 2019	48
Annexe 9: Contexte du FlowCam pour le X4	2
Annexe 10: Contexte du FlowCam pour le X10.....	3
Annexe 11: Répartition des échantillonnages par station sur le FlowCam depuis 2019	4
Annexe 12: Espèces ne pouvant figurer dans les set d'apprentissage.....	5
Annexe 13: Groupe du 1 ^{er} set d'apprentissage	7
Annexe 14: Evaluation des set d'apprentissage par validation croisée et prédiction pour l'objectif X10 (1 ^e set à gauche et 2 ^e set à droite)	9

Annexe 15: Evaluation des set d'apprentissage par validation croisée et prédiction pour le x4 (1e set à gauche et 2e set à droite).....	10
Annexe 17: Cycle de vie de <i>Phaeocystis</i> (Rousseau et al., 1994).	10
Annexe 16: Groupement pour les vignettes en x4.....	5
Annexe 17: Exemple de vignettes x4 (a. <i>Asterionellopsis glacialis</i> , b. <i>Meuniera membranacea</i> , c. <i>Cerataulina pelagica</i> , d. <i>Thalassionema nitzschioides</i> , e. <i>Ditylum brightwellii</i> , f. <i>Phaeocystis globosa</i> colonie, g. <i>Dictyocha speculum</i>)	11

Liste des Abréviations - Glossaire

CNRS: Centre National de la Recherche Scientifique, 1

CTD: Conductivity Temperature Depth, 6

DCSMM: Directive-cadre stratégie pour le milieu marin, 5

DYPHYRAD: Dynamique du PHYtoplancton en RADiale de la Baie de Saint Jean, 7

FlowCam: Flow Cytometer And Microscope, 7

IRD: Institut de Recherche pour le Développement, 1

JERICO: Joint European Research Infrastructure of Coastal Observatories, 7

LOG: Laboratoire d'Océanologie et Géosciences, 1

MARCO: Recherches Marines et Littorales en Côte d'Opale, 13

MAREL: Mesures Automatisées en Réseau pour l'Environnement Littoral, 7

MREN: Maison de la Recherche en Environnement Naturel, 1

REPHY: Réseau de Surveillance du Phytoplancton et des Phycotoxines, 4

SMW: Station Marine de Wimereux, 1

SOMLIT: Service d'Observation en Milieu Littoral, 7

SRN: Suivi Régional des Nutriments du littoral de la côte d'opale, 7

ULCO: Université du Littoral Côte d'Opale, 1

UMR: Unité Mixte de Recherche, 1

1 Introduction

1.1 Présentation de la structure

Le laboratoire d'Océanologie et Géosciences (LOG) est une Unité Mixte de Recherche (UMR 8187) qui regroupe différentes institutions telles que le Centre National de la Recherche Scientifique (CNRS), l'Université de Lille, l'Université du Littoral Côte d'Opale (ULCO) et, récemment, l'Institut de Recherche pour le Développement (IRD). Avec le BPA-ICV (Biochimie des Produits Aquatiques), le LOG fait partie du pôle de recherche « Pôle Mer et Littoral, Transformations et Enjeux » (*Laboratoires*, s. d.).

Le LOG est implanté sur trois sites dans la Région des Hauts-De-France : au Campus de l'Université de Lille à Villeneuve-d'Ascq (Bâtiment SN5), à Dunkerque (Maison de la Recherche en Environnement Industriel, ULCO), et à Wimereux : à la Station Marine (SMW, U Lille) et à la Maison de la Recherche en Environnement Naturel (MREN ULCO). C'est au sein de cette dernière que j'ai effectué mon stage.

L'UMR 8187 LOG est composée de 63 enseignants chercheurs, 57 doctorants et 22 ingénieurs et administratifs. Cet effectif permet d'avoir une interdisciplinarité au sein de la structure et de constituer 6 équipes de recherches : INTEREST (Ecologie INTERTidiale et ESTuarienne), TELHYD (TELétection et HYDrodynamique), GEOLIT (GEOsciences du LITtoral), GEOSSED (Messages Minéralogiques et GÉOchimiques dans les SÉDiments), TIMES (Tectonique des Marges et des Bassins Sédimentaires) et l'équipe ECOP² (ECOlogie PELagique Planctonique) dont j'ai fait partie, encadré par Luis Felipe Artigas, Maître de conférences et responsable de l'équipe.

Les objectifs de l'équipe ECOP² sont d'étudier et de comprendre la structure, la dynamique, les interactions et le fonctionnement des écosystèmes pélagiques marins, à partir de l'étude du compartiment planctonique (LOG, 2022). Les échelles d'observations spatiale et temporelle de ce compartiment vont de la micro-échelle (seconde, μm) à la méso-échelle (décennie, dizaines à centaines de kilomètres), sur les écosystèmes côtiers tempérés (Manche, Mer du Nord mais aussi Baltique et détroit scandinaves, Golfe de Gascogne, et Méditerranée). Des missions sont également opérées dans l'arc Atlantique tempéré à tropical ainsi que de l'Océan Austral et zones polaires.

Cette équipe est composée de 10 titulaires (maître de conférences, professeur des universités et assistante ingénieure), 4 doctorants, 11 post-doctorants et contractuels (ingénieure d'étude, assistant ingénieur, ATER) (LOG, 2022).

Ces études se concentrent sur trois axes :

- La structure et la variabilité des communautés planctoniques combinant des études *in situ/in vivo* à l'aide d'outils microscopiques, moléculaires et capteurs optiques/imagerie (semi-) automatisés à haute résolution.
- La diversité et les interactions au sein des communautés intégrant les parasites, les blooms d'algues nuisibles et copépodes afin de prendre en compte la santé animale et la biotechnologie par le biais de bio-indicateur.
- Les stocks et flux de matière au sein des écosystèmes pélagiques (bactéries au méso zooplancton) en étudiant les cycles biogéochimiques et leur déterminisme.

Mon stage s'est inséré sur le premier axe au sein de la sous-équipe « Phyto-CytoFlulm » composée du responsable d'équipe, Felipe Artigas, deux post-doctorants, une ingénieure d'étude et une doctorante, ainsi que de stagiaires de plusieurs diplômes (Master 1 et 2, BTS) présents pendant mon stage, en collaboration avec d'autres chercheurs, doctorants, ingénieurs et techniciens du LOG et du LER IFREMER Boulogne sur Mer.

1.2 Contexte

1.2.1 Intérêt pour ce stage

Ce stage, en laboratoire de recherche en océanologie, s'inscrit dans mon projet professionnel d'étudier les écosystèmes marins. La structure d'accueil me permet de mettre à profit mes qualités d'ingénieur dans un laboratoire de recherche.

Grâce à ce stage de fin d'études, j'enrichis mes compétences en laboratoire, acquises lors de mon cycle technologique, sur divers appareils de mesures du phytoplancton (imagerie, cytométrie en flux, fluorimétrie). Ces compétences sont couplées à de la recherche de terrain, du milieu marin avec l'échantillonnage à l'ordinateur pour le traitement, l'analyses et l'interprétation des résultats. J'étoffe ainsi mon Curriculum Vitae tout en travaillant dans un environnement que j'affectionne.

Ce stage pluridisciplinaire, à la fois en laboratoire, en mer et en bureau avec du traitement de données et de l'analyse sous R, en relation étroite avec d'autres stagiaires de plusieurs diplômes, des techniciens, ingénieurs et chercheurs locaux et internationaux, m'apportera des avantages dans la finalisation de mon parcours d'ingénieur mais également sur le plan personnel étant dans une région qui m'est familière.

Ce stage, à l'origine d'une candidature spontanée, témoigne de mon intérêt aussi bien sur le plan personnel que sur le plan professionnel, pour le contexte d'étude et les missions effectuées.

1.2.2 Projets intégrant le stage

Ce stage intègre le projet régional du Contrat Plan Etat Région (CPER) Hauts de France « IDEAL » (approche Intégrée des DEFIS maritimes et Littoraux) sur la plateforme Observation, Géomatique et Télédétection, qui assure la continuité du précédent CPER MARCO, et qui a pour ambition de maintenir et étoffer les plateformes techniques et, plus précisément, de coordonner les actions sur le suivi de la dynamique phytoplanctonique pour une meilleure caractérisation de l'état des écosystèmes côtiers.

Il s'inscrit dans les Work Packages WP4 (Pilot Supersites for innovative coastal monitoring), WP5 (Best operational practices), WP6 (Data management), WP7 (Technical and Artificial Intelligence developments) du projet Européen H2020 - JERICO S3 (Joint European Research Infrastructure of Coastal Observatories: Science, Service, Sustainability) (*JERICO-S3 Project Overview | JERICO Research Infrastructure*, 2021). Ce projet est la continuité de deux autres éditions JERICO-FP7 (2011-2014) et JERICO-NEXT (2015-2019) pour observer et surveiller l'environnement côtier marin au sein d'une infrastructure de recherche rassemblant les observatoires côtiers en Europe, JERICO-RI.

Les variabilités des communautés de phytoplancton peuvent avoir lieu à de petites échelles spatio-temporelles échappant aux suivis mensuels ou bi-mensuels des observatoires en place ou méthodes d'analyse traditionnelles (microscopie, analyses pigmentaires). Des méthodes automatisées sont déployées depuis plus d'une dizaine d'années en Manche Orientale (Guiselin, 2010) combinant la fluorimétrie *in vivo* multispectrale, la cytométrie en flux de type "pulse shape-recording" et l'imagerie au sein du LOG. Mon travail vient compléter celui initié par Eléonore Delecroix (2019) sur un appareil d'imagerie en couleurs (FlowCAM 8100) acquis en 2019, avec l'aide et collaboration de Guillaume Wacquet, ingénieur de recherche à Ifremer. Cette collaboration a permis de développer le système d'analyse automatisé de G. Wacquet sur une ancienne version du FlowCam (Grosjean & Wacquet, 2016). Il permettra d'améliorer les performances de l'analyse d'image automatisée sur les nouveaux capteurs acquis dans le cadre du CPER MARCO et en cours d'acquisition du CPER IDEAL.

1.3 Objectif du stage

Les collaborations au sein de mon stage, multiples et diverses (chercheur, ingénieur, taxonomiste, technicien), m'ont permis de prendre en main la méthode et assister à l'entièreté du processus, de la collecte pour l'analyse et caractérisation de la diversité et dynamique phytoplanctonique au travers d'un outil d'analyse : le FlowCAM.

Mon objectif au cours de ce stage a été **l'amélioration de l'étude du phytoplancton marin à partir de l'imagerie automatisée appliqué au suivi des blooms phytoplanctoniques en eaux côtières** (Manche et Mer du Nord) afin d'émettre les perspectives pour une meilleure surveillance pouvant permettre une meilleure gestion à venir.

1.3.1 Missions et problématiques abordées

La collecte et étude du phytoplancton (méthodes de référence), une fois déduit le coût important des sorties en mer et le personnel associé nécessaire, reste une méthode simple, avec une grande part du travail à réaliser lors des analyses en laboratoire. L'analyse manuelle demandant du temps et une spécialisation conséquente en taxonomie, les échantillons de phytoplancton peuvent demeurer longtemps sans être traités. L'imagerie automatisée représente un progrès technologique dans l'efficacité du traitement à une résolution spatiale et temporelle élevée.

Ma problématique a été d'optimiser l'automatisation des analyses de diversité taxonomique du phytoplancton dans le but de pouvoir traiter plus rapidement et efficacement les échantillons des différents suivis au sein du LOG.

Pour mener à bien cet objectif, mes missions principales ont été l'acquisition de données, l'optimisation des protocoles et des méthodes en cours. J'ai pu travailler sur la valorisation des données d'imagerie automatisée acquises sur la zone d'étude (depuis la Baie de Somme à Dunkerque) depuis 2019. J'ai également participé à la collecte de données lors d'une campagne océanographique internationale transfrontalière en Manche Orientale et Mer du Nord, en collaboration avec le VLIZ (Vlaams Instituut voor de Zee, l'Institut Marine des Flandres).

1.4 Sujet du stage

1.4.1 Importance du suivi phytoplanctonique

Par définition, le phytoplancton représente des microorganismes aquatiques chlorophylliens dérivant dans les masses d'eaux (Jeffrey et al., 1997). Il est responsable d'au moins 50% de la production primaire nette globale (Field et al., 1998). Sa plasticité lui permet de coloniser tout type de milieu, de ce fait il est présent dans l'ensemble des écosystèmes aquatiques, qu'ils soient d'eau douce, saumâtre ou salé (Masmoudi, 2014).

Le phytoplancton marin englobe différents organismes tels que les diatomées, les dinoflagellés et d'autres cellules flagellées et non-flagellées (Choudhury & Pal, 2014; Kraberg et al., 2010). Identifier et comprendre les interactions de cette communauté devient primordial pour prédire l'influence du changement climatique sur les écosystèmes aquatiques (Chisholm, 1992; Winder & Sommer, 2012).

Le phytoplancton joue un rôle fondamental en tant qu'organismes photosynthétiques. En utilisant la photosynthèse oxygénique, il est capable de convertir la matière inorganique en matière organique grâce à l'effet de la lumière (P. Falkowski et al., 2003). Le carbone organique restant de ce processus et de l'utilisation d'une partie de la matière organique par le phytoplancton, appelé production primaire nette (Lindeman, 1942), est rendu disponible pour les organismes hétérotrophes. Ces

échanges, absorption du dioxyde de carbone et production du dioxygène atmosphérique (Boyce et al., 2010; Choudhury & Pal, 2014), sont essentiels au sein du cycle biogéochimique de planète. Le phytoplancton est donc à la base des réseaux alimentaires (Lévêque, 2006) et est considéré comme un acteur majeur de la pompe biologique des océans (Lévy & Bopp, 2007; Rocha & Passow, 2014).

Le suivi du phytoplancton est important en tant qu'indicateur du bon état des masses d'eau, sentinelles de la qualité des eaux, des cycles des éléments, de la production à la base des réseaux alimentaires et des ressources vivantes (F. Artigas, *comm. pers.*) Cependant ces microorganismes aquatiques peuvent représenter une nuisance pour les autres organismes et pour le recyclage de la matière (Hess, 2020). Ils peuvent être responsables d'événements toxiques ou nuisibles pour les organismes et/ou pour la santé humaine dans le cas de conditions physico-chimiques favorables à leur prolifération (Chiswell et al., 2015), c'est le bloom phytoplanctonique. Leur surveillance est cruciale afin de détecter les espèces nuisibles pour la faune marine et pour les consommateurs de coquillages. C'est le rôle qu'assure le Réseau de Surveillance du Phytoplancton et des Phycotoxines (REPHY) (Ifremer REPHY, 2018).

1.4.2 Histoire de l'imagerie planctonique

L'analyse du phytoplancton se faisait essentiellement sous microscope inversé avec la méthode d'Utermöhl par décantation d'un échantillon dans une chambre de sédimentation (Utermöhl, 1958). Malgré que cette méthode soit normalisée (AFNOR, 2006) et répandue pour l'estimation de l'abondance et de la composition taxinomique du phytoplancton marin et d'eau douce (IFREMER, 2014; INRAE, 2009; OFB, 2018), elle est associée à de nombreuses heures en laboratoire et une expertise et investissement considérable. Ce travail du planctonologiste est chronophage et fastidieux. Suite à ce constat, l'imagerie a été développée pour l'analyse du plancton afin d'offrir un gain de temps et minimiser les biais qui peut découler de la méthode par microscopie (Ali et al., 2014; Grosjean & Wacquet, 2016).

Les premières collections d'images du phytoplancton ont été réalisées *in situ* en 1954 (Nishizawa et al., 1954). A partir des années 90, la photographie a laissé place à la technologie des vidéos notamment avec l'enregistreur vidéo de plancton le VPR (Video Plankton Recorder). Suite à cette avancée, l'ingénierie des systèmes d'analyse du plancton n'a cessé d'évoluer avec notamment l'apparition des imageurs en flux de laboratoire comme le FlowCam (1998). L'imagerie a également été ajoutée aux autres appareils pour coupler plusieurs techniques d'analyses phytoplanctoniques. C'est le cas pour l'EurOPA en tant qu'instrument modulaire pour la reconnaissance des espèces en combinaison avec les caractéristiques optiques de la cytométrie en flux (Rutten et al., 2005). Le CytoSense, présent au LOG associe également la cytométrie avec l'imagerie.

Les nouvelles machines d'imagerie permettent d'analyser les échantillonnages sans préparation particulière sur du vivant. Ainsi, en plus d'éliminer la perte d'information due à un traitement sur des échantillons fixés au lugol (Choi & Stoecker, 1989; Menden-Deuer et al., 2001), le FlowCam permet de traiter plus d'échantillons *in vivo* dans les heures qui suivent la collecte. Avec l'utilisation du FlowCam, les données peuvent être collectées sur une zone d'étude plus grande ou à une plus grande fréquence que les suivis traditionnels, apportant ainsi une meilleure représentativité de la dynamique du phytoplancton.

1.4.1 Présentation du site d'étude

1.4.1.1 Manche orientale et Mer du Nord

Située au nord de la France, la Manche est une mer épicontinentale de l'océan Atlantique qui sépare la France du Royaume-Uni (PINOT, s. d.). Le détroit du pas de Calais, de Douvres (GB) à Calais (FR), forme la délimitation entre la Manche et la Mer du Nord. Cette dernière borde 7 pays dont la Grande-

Bretagne, la Belgique et la France sur 50km, les Pays-Bas, l'Allemagne, le Danemark et la Norvège (Perrodon & Pinot, s. d.).

1.4.1.1 Enjeux

La Directive-cadre stratégie pour le milieu marin (DCSMM) s'applique à la sous-région marine Manche-mer du Nord. Depuis 2008, elle développe une approche écosystémique du milieu marin, afin de maintenir ou rétablir un bon fonctionnement des écosystèmes dans une démarche de développement durable (Ministère de l'Écologie, du Développement durable et de l'Énergie, 2015). Cette sous-région regroupe 28 815 km² soit 0,28 % des eaux françaises et 7,72% de la superficie des eaux présentes en France métropolitaine (Ministère de l'Écologie, du Développement durable, & des Transports et du Logement, 2012).

Les enjeux de cette zone sont d'importance environnementale (Annexe 1) notamment avec la présence du Parc naturel marin des Estuaires picards et de la mer d'Opale (PNR EPMO), des sites reconnus Natura 2000 pour la Directive « Oiseaux » et la Directive « Habitats faune flore ». Elle présente également sur les côtes des Réserves naturelles OSPAR (convention Oslo-Paris) et Réserves naturelles nationales et régionales comme la Baie de Somme qui a été classée comme zone humide d'importance internationales (RAMSAR) (Guinand, 2012) (Nord, 2018).

De par sa localisation, les enjeux économiques y sont importants. En effet, la Manche forme le rail montant et descendant formant un des trafics maritimes le plus intenses du Monde (Carpentier et al., 2005) pour de nombreuses activités (tourisme, loisir, ports internationaux et fret ainsi que l'exploitation de ressources vivantes ou non).

Ainsi, la façade maritime Manche Est - Mer du Nord présente de forts enjeux socio-économiques et environnementaux (Annexe 2)(Direction Inter-régionale de la Mer - Manche Est - Mer du Nord, 2019) pour lesquelles un suivi phytoplanctonique est d'autant plus important.

1.4.2 Paramètres environnementaux

1.4.2.1 Courantologie de la Manche orientale et Mer du Nord

L'étude de la répartition du phytoplancton doit tenir compte de l'étude des courants. La Manche est sujet à différents processus physiques qui créent un hydrodynamique particulier. Dans un premier temps, les plages de la côte Manche Orientale sont appelées « macrotidales » du fait de leur morphologie qui entraîne des amplitudes de marée significatives (supérieur à 4m, Annexe 3) et inférieur à 8m (Levoy et al., 2000; Stépanian, 2003). Ces marées modulent la profondeur et la hauteur des vagues mais également la vitesse des courants.

Le marnage, différence de hauteur entre la pleine mer et la basse mer (Annexe 3Annexe 4**Erreur ! Source du renvoi introuvable.**), est un signal périodique lié aux positions du Soleil et de la Lune relativement à la Terre (IHO, 2011). Cependant, il subit quelques variations d'une journée à l'autre en fonction de différents phénomènes physiques qui ont lieu en Manche et Mer du Nord :

- La marée semi-diurne est « *un type de marée pour laquelle les composantes diurnes sont négligeables devant les composantes semi-diurnes. Il y a alors deux pleines mers et deux basses mers d'importances sensiblement égales par jour. Ce type de marée est prépondérant en Atlantique et en Manche* ». Ce phénomène donne lieu aux marnages les plus importants des côtes métropolitaines (Lazure & Desmare, 2012). Ces marées donnent lieu à des courants de flot (lors du montant des eaux) vers le nord-est et des courants de jusant (lors du perdant des eaux).
- Courant résiduel de marée, cycles vives-eaux / mortes-eaux (période de 14 jours) « *si le Soleil et la Lune se trouvent à angle droit par rapport à la Terre, ce qui arrive tous les quinze jours, leur force d'attraction se contrarient, donnant lieu à des marées de faible amplitude : ce sont les mortes eaux.*

Lorsque la Lune et le Soleil sont alignés avec la Terre, leurs forces s'additionnent, ce qui provoquent les marées de vives eaux » (Brosche, 1982). Les courants résiduels s'orientent de l'Atlantique vers la Mer du Nord (Annexe 4). L'ordre de grandeur de ces courants est compris entre 5 et 10 cm/s dans le Pas de Calais (Salomon & Breton, 1993)(Bailly Du Bois & Dumas, 2005).

A noter que le vent peut avoir un impact sur les courants, notamment lorsque les vents viennent du Nord où des inversions peuvent avoir lieu (Ellien et al., 2004).

1.4.2.1 Bathymétrie

La variabilité des courants et le marnage, vus précédemment, sont influencés par la bathymétrie et la morphologie de la côte. La bathymétrie est intéressante à mentionner pour caractériser notre site d'étude car elle va influencer la mesure des paramètres physico/chimiques effectuée sur les suivis. Ces paramètres (conductivité, température, salinité, profondeur et la transmittance) sont mesurés sur la sonde CTD (*Conductivity Temperature Depth*) en surface et en profondeur (1m avant le fond)(Milieu Marin France, 2022).

1.4.2.2 « Fleuve côtier »

Le phénomène hydrologique en Manche orientale a été étudié par Brylinski et al. en 1991 en raison de son importance dans les flux entre la Manche orientale et la Mer du Nord. La dénomination « fleuve côtier » vient des multiples points d'embouchures fluviales qui génèrent une masse d'eau côtière légèrement dessalée, de 3 à 5 milles de large, dérivant vers le littoral. Ainsi, de la baie de Seine et plus particulièrement de la baie de Somme au cap Gris-Nez (Figure 1), une zone frontale influencée par la marée est caractérisée par sa dessalure, sa fraîcheur, sa turbidité (Brylinski et al., 1984) et sa teneur en phytoplancton et en pollution (Quisthoudt et al., 1987). La Mer du Nord est sous l'influence principale du delta Rhin-Escaut-Meuse qui dessale ponctuellement l'eau de mer au gré du débit, des courants et du vent (Vandermeirsch, 2012).

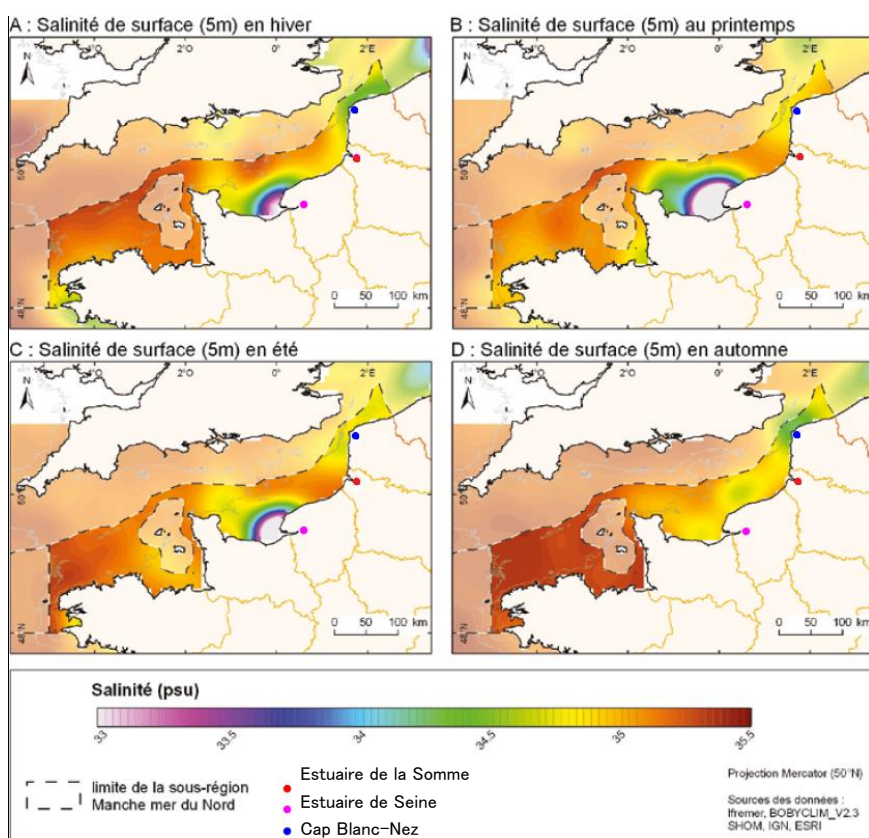


FIGURE 1: SALINITE SAISONNIERE A 5 M DE PROFONDEUR (SOURCES : CLIMATOLOGIE MENSUELLE MANCHE - GOLFE DE GASCOGNE BOBYCLIM_v2.3, 2010).

1.4.3 Localisation des données de suivis

1.4.3.1 Suivis

Les suivis permettent l'acquisition de données régulières depuis plusieurs années afin de caractériser et comprendre la dynamique des communautés phytoplanctoniques. Ainsi le LOG participe à l'analyse les échantillons récoltés par le navire « Sépia II » pour différents services, projets et réseaux d'observation (Figure 2) :

- **DYPHYRAD** (Dynamique du PHYtoplancton en RADiale de la Baie Saint Jean) qui s'étend au large d'Ambleteuse et l'embouchure de la Slack, fait l'objet d'un échantillonnage hebdomadaire depuis 2008 sur 9 stations (suivis ayant été initiés en 2002 à plus basse fréquence et sur un nombre réduit de points). Cette radiale est composée de neuf points (R4, R3', R3, R2', R2, R1', R1, R0', R0) sur une dizaine de kilomètres (Artigas et al., 2016). L'acquisition des données sur cette radiale a débuté en 2008 et s'est consolidée dans le cadre du projet INTERREG IVA « 2 Mers » DYMAPHY (Développement d'un système d'observation dynamique pour la détermination de la qualité des eaux marines, basé sur l'analyse du phytoplancton ; 2010-2014).

- **SRN REPHY** (Suivi Régional des Nutriments du littoral de la côte d'opale, Réseau de surveillance du phytoplancton et des phycotoxines d'IFREMER) avec trois radiales, dans et au large de la Baie de Somme (Mer2, Atso, Mimer, Bif), au large de Boulogne-sur-Mer (BI3, BI2, BI1) et au large de Dunkerque (DK4, DK3, DK1). Le suivi SRN est fait de manière mensuelle par IFREMER depuis 1992 et, depuis 2017, les prélèvements sont faits deux fois par mois (Devreker & Lefebvre, 2022; IFREMER, 2017).

- **SOMLIT** (Service d'Observation en Milieu Littoral au large du Cap d'Alprech (au sud de Boulogne-sur-Mer). Ce Service National d'Observation (SNO) de l'Infrastructure de Recherche Littorale et Côtière (IR-ILICO) est réalisé depuis 1997 tous les 15 jours au niveau de deux stations (côte « SC » et large « SL ») dans des conditions de pleine mer de vives eaux. Comme pour les points principaux de DYPHYRAD, l'échantillonnage permet la réalisation des profils verticaux et de prélèvements discrets en sub-surface et au fond de la colonne d'eau (OASU, 2018).

Ces deux suivis font partie du SNO du microphytoplancton PYTOBS porté par l'Ifremer, le CNRS et les Universités Marines (*Phytobs*, 2018) dans l'infrastructure de recherche IR-ILICO.

- **MAREL** (Mesures Automatisées en Réseau pour l'Environnement Littoral) est une station fixe qui a été construite en 2004 au bout de la digue Carnot dans la rade de Boulogne-sur-Mer (*MAREL Carnot*, s. d.) (Derot, 2014) (Annexe 5).

Elle intègre le SNO de l'IR ILICO « COAST-HF » pour la surveillance des eaux côtières de France métropolitaine à haute fréquence. En plus des données hydrologiques et biogéochimiques, des mesures biologiques sont proposées comme celles effectuées à haute fréquence grâce aux cytomètres en flux automatisés déployés in situ. De plus, des mesures ponctuelles sont effectuées chaque semaine depuis 2021 pendant les périodes de déploiement des capteurs automatisés.

1.4.3.2 Campagnes océanographiques

Ponctuellement l'équipe effectue des campagnes à plus large couverture spatiale en collaboration avec d'autres organismes nationaux et internationaux. Ainsi, en mai 2022 j'ai pu participer à la campagne JERICO-S3/LifeWatch organisée par le Vlaams Instituut voor de Zee (VLIZ) à bord du navire océanographique « Simon Stevin ».

Cette campagne de quatre jours a pour but d'inter comparer le cytomètre et le FlowCam avec celui du VLIZ ainsi que d'obtenir des données d'abondance et de diversité phytoplanctonique printanière en Manche orientale et au sud de la Mer du Nord. Vingt-six stations ont été effectuées réparties dans les eaux françaises, anglaises et néerlandaises (Figure 2).

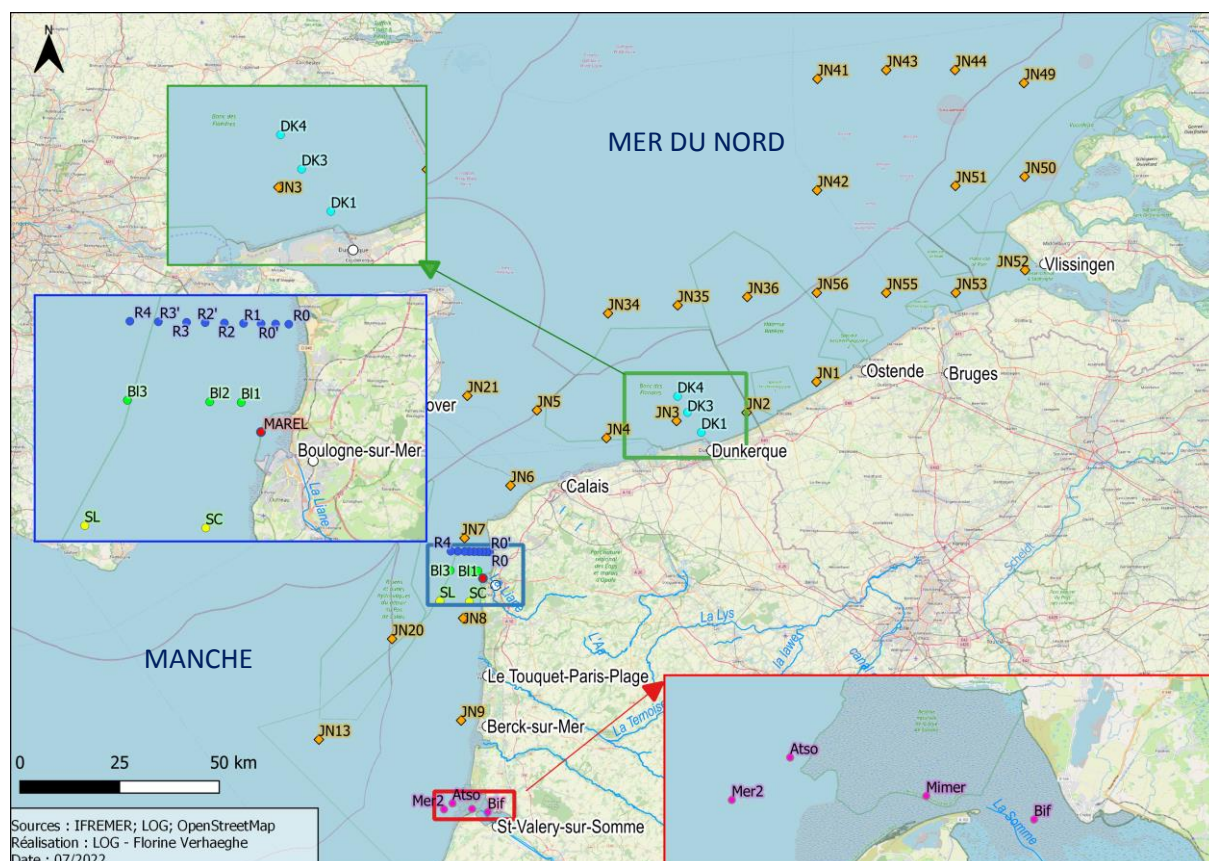


FIGURE 2: CARTE DES SUIVIS ET CAMPAGNES EFFECTUES EN 2022

2 Méthodologie

2.1 Outils de travail, le FlowCam

Le FlowCam® (*Flow-Imaging microscope*) a été créé par Mike Sieracki, Chris Sieracki et Charlie Yensch au *Bigelow Laboratory for Ocean Sciences* dans le Maine (US) en 1996. L'entreprise se nomme désormais *Yokogawa Fluid Imaging Technologies* suite à son acquisition au Japon. Le principe du FlowCam repose sur un système d'imagerie en flux de particules combinant les fonctionnalités d'un cytomètre de flux d'imagerie et d'un microscope.

2.1.1 Fonctionnement

La caméra du FlowCam génère des images haute résolution (1280x960 pixels) des particules dont la taille dépend de la combinaison entre la magnification et l'épaisseur de la cellule de flux (2X/600µm, 4X/300µm, 10X/100µm ou 20X/50µm). L'analyse de ces particules par imagerie dynamique utilise ainsi des images numériques pour mesurer la taille (de 2µm à 2000µm) et la forme de chaque particule.

L'échantillon contenant les particules passe à travers l'optique de la cellule de flux et la caméra capture des milliers d'images de particules par seconde (Figure 3) (Wacquet, 2017).

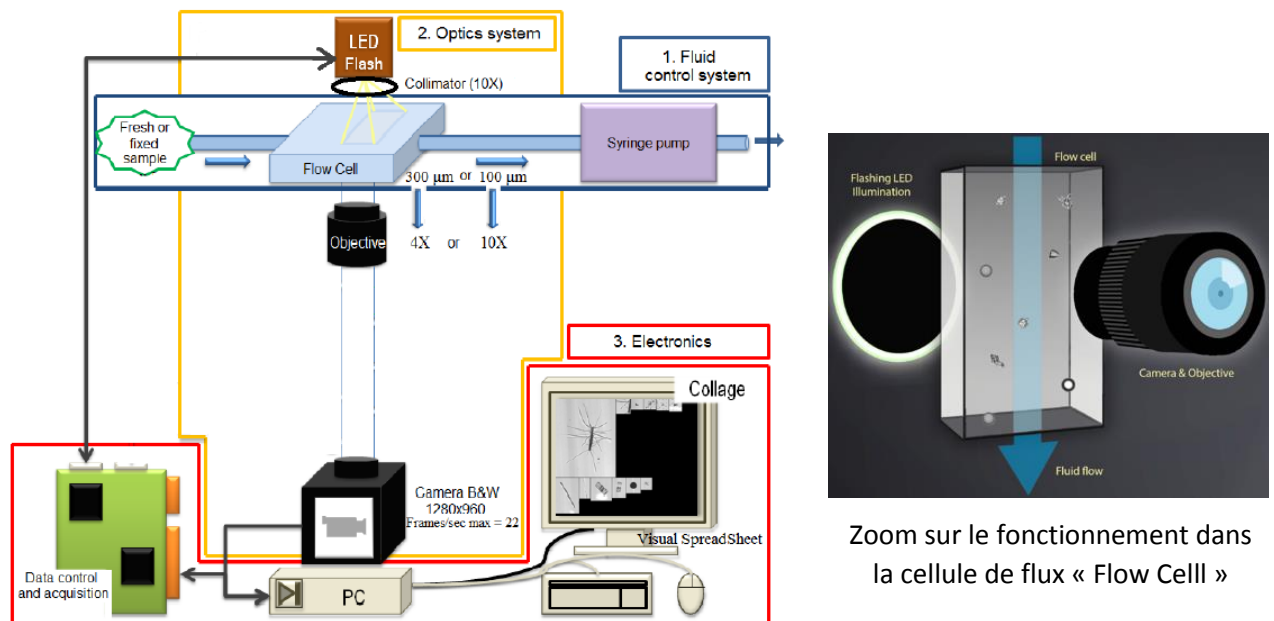


FIGURE 3: FONCTIONNEMENT DU FLOWCAM®, SOURCE : FLUIDIMAGING, GUIDE D'UTILISATION (WACQUET, 2017)

L'appareil est composé d'un système de contrôle de l'écoulement avec une pompe seringue adaptée au système optique (cellule de flux, objectif connecté à une caméra, LED) pour le contrôle du débit.

La détection de la particule sur le FlowCam 8100 se fait à l'aide du mode « Autolmage » (ou auto-déclenchement) sur le logiciel Visual SpreadSheet. Le fonctionnement repose sur la capture de la vision de la caméra à intervalle de temps régulier (défini par l'utilisateur) (Buskey & Hyatt, 2006).

Le logiciel « Visual SpreadSheet » permet d'avoir accès aux commandes électroniques du FlowCam® comme l'accès aux paramètres d'acquisition (contexte de numérisation, auto-stop), contrôle de l'appareil (montage, démontage, Autofocus) et à la gestion des fichiers et données collectés.

2.1.2 Données en sortie

Les données collectées sont stockées dans l'ordinateur interne au FlowCam. Le dossier brut, après chaque échantillon passé, contient ainsi les vignettes sauvegardées dans des « collages » afin de préserver la mémoire en format .tiff ainsi que les notes et résumés de l'acquisition des images.

2.1.3 Concept de Transfer Learning et Deep Learning

Pour traiter ces images de manière automatisée, un modèle d'apprentissage supervisé, appelé réseau de neurones (Neural Network, NN), est utilisé. Ici, le concept repose sur l'utilisation des images obtenues pour entraîner le NN qui apprend une approximation de la relation par un ajustement itératif de ses paramètres (Asch, 2022). En cas d'erreur, le programme « remonte le cours du réseau » pour faire varier les paramètres, de sorte que le résultat de sortie se rapproche du résultat attendu (Kassel, 2022). Le réseau est un graphe avec une architecture (couches, nœuds (« neurones »), activations). Ici le NN utilisé est un réseau convolutionnel (CNN), LeCun, pour le traitement d'image, pré-entraîné sur la base de données ImageNet. Il s'agit du modèle VGG16. Le Deep learning travaille directement sur les images du collage obtenu.

« L'idée du réseau est que le monde visuel est « compositionnel », c'est à dire qu'un objet est composé de parties, qui sont composées de sous parties [...] ça conduit à imaginer des structures qui sont aussi hiérarchiques, dans lesquelles chaque couche détecte des combinaisons de motifs présents sur la couche précédente »

Yann LeCun (2017)

2.2 Savoirs et compétences mis en œuvre

2.2.1 Echantillonnage

Dans cette partie, seront présentés les réseaux d'observation et suivi auxquels j'ai participé.

L'échantillonnage et mesures de ces suivis se fait à bord du navire de la Flotte Océanographique Française le « Sépia II » (Figure 4). Ce navire de pêche (type « *Fishing* ») mesure 13 mètres de long et 5 mètres de large (IFREMER, 2019). Une sonde et des bouteilles Niskin sont en général utilisées pour les échantillonnages. La CTD et la bouteille Niskin sont fixées à un treuil pour prélever à une profondeur connue de la colonne d'eau.



FIGURE 4: SEPIA II (A) ET BOUTEILLE NISKIN SUR LA CTD (B)

Les échantillons récoltés sont par la suite conservés au frais pour une analyse sur du vivant réalisée au laboratoire dans la journée (environ 3-5h après collecte). Les analyses laboratoires des échantillons sur la radiale DYPHYRAD en fonction des points principaux et secondaires sont détaillés en Annexe 6.

Lors de la campagne océanographique à bord du Navire Océanographique « Simon Stevin » (36 mètres de long et 9,4 mètres de large (VLIZ, 2012)), les données (abondance et diversité zoo/phyto, biologie moléculaire, nutriments pigments) ont été collectées sur une station fixe pendant 24h. Les mesures ont été effectuées en continu et prélevé toutes les heures grâce à une rosette de 6 CTD fixées à un treuil. Dans un second temps, du 3 au 5 mai, les prélèvements ont été effectués quasi toutes les heures avec une coupure de 22 heures à 6 heures du matin. Des mesures en continu ont été prises lors des transits entre station et le soir (total de 26 stations) (Figure 2).

Les données sont, suivant les appareils disponibles, directement acquises sur le navire lors des campagnes contrairement aux données de suivi qui sont obtenues au laboratoire dans les cinq heures suivant la sortie (sortie en mer le matin et retour au laboratoire à midi).

Ainsi, durant les multiples échantillonnages, j'ai collaboré avec le technicien du LOG, les marins du Sépia II et du Simon Stevin, des scientifiques et chercheurs flamands. J'ai pu constater les différentes techniques et moyens mis en œuvre sur les deux types de navires et sur une sortie de quelques heures contre une campagne de 4 jours.

2.2.2 Prise en main de l'appareil et technique de laboratoire

Après la collecte des échantillons, l'acquisition de données s'effectue sur différentes machines au laboratoire. J'ai pu me former auprès d'une Ingénieure d'Etudes sur ces différentes techniques d'acquisition (cytométrie en flux de type CytoSense, fluorimétrie multispectrale (Fluoroprobe) et monospectrale, imagerie (FlowCAM) et filtration pour analyses des concentrations en chlorophylle *a*). Une formation hygiène et sécurité est obligatoire afin d'avoir les bonnes pratiques au laboratoire et elle a été suivie auprès du Correspond Hygiène et sécurité du LOG et de ressources en ligne. Enfin, j'ai suivi des cours en ligne organisés par le réseau I/ITAPINA (*Imagine/Imaging the Atlantic - a Pelagic Imaging Network Approach*) et l'Université de Sao Paulo (*Advances in Marine Pelagic Ecology*, s. d.) sur les approches d'imagerie automatisée pour l'observation et étude du plancton.

Le FlowCAM est un appareil disposé sur une paillasse au laboratoire. Afin d'acquérir les données, il faut précédemment effectuer le *focus* pour avoir des images nettes. L'objectif est sur un rail qui s'adapte en fonction de l'objectif choisi et au cours de la mise au point. On dispose l'échantillon dans l'entonnoir fixé à un cône. La pompe seringue (Figure 5, n°1) aspire l'échantillon qui passe dans la cellule de flux (Figure 5, n°2). Les photos en couleur de chaque particule sont prises au moment de la traversée du tube en téflon. Le liquide est ensuite évacué par des tuyaux en silicone (Figure 5, n°3) jusqu'à un réservoir « poubelle ». Ce réservoir est ensuite vidé dans les poubelles du laboratoire appropriées. Le protocole détaillé appliqué concernant le sujet de mon stage, le FlowCAM, se trouve en Annexe 7.

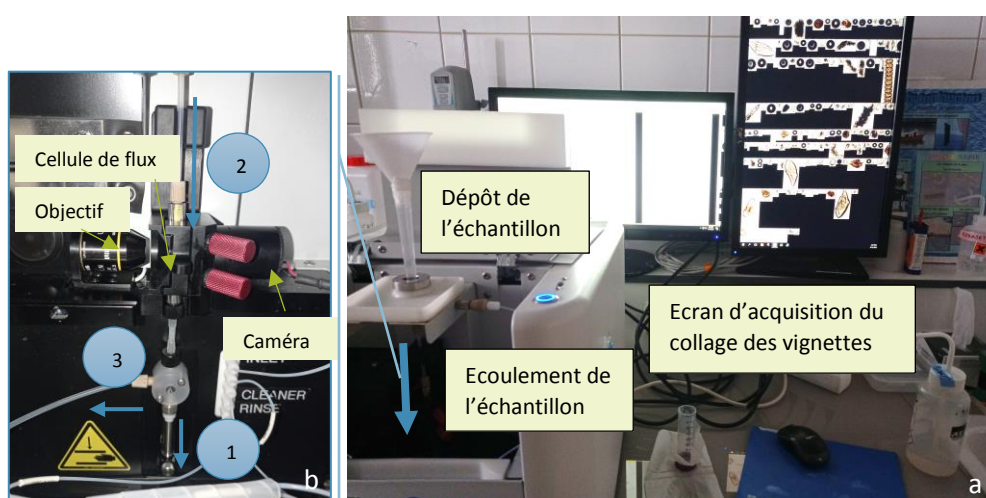


FIGURE 5: FONCTIONNEMENT DU FLOWCAM AU LABORATOIRE

L'utilisation des objectifs x4 et x10 pour le passage des échantillons dépend de la taille des particules présentes dans l'eau. Lors des blooms de colonie (la taille pouvant aller de 10 μm à 3 mm suivant la forme de *Phaeocystis globosa* (Lefebvre & Delpech, 2004) par exemple) l'objectif x4 (40) était privilégié afin d'avoir la capture de la plupart des colonies dans leur ensemble. Le grossissement X10 (100) est plus pertinent à employer afin d'avoir une meilleure résolution de l'image rendant possible l'identification au niveau de l'espèce.

Le temps des deux mesures est également à prendre en compte lors du passage des échantillons (5 min pour le x4 et 20 min pour le x10 environ). Ainsi durant le printemps le x4 est utilisé en priorité, le x10 est passé par la suite si le temps le permet (environ 5/7 minutes pour le démontage et la mise au point). En été, automne et hiver le x10 est utilisé, le x4 peut être utilisé ponctuellement. L'utilisation

du x10 associé à la cellule de 100µm en période de blooms ou dans des eaux chargées en matières en suspensions (MES) implique l'emploi d'un pré-filtre de 100µm (depuis 2022, précédemment 150µm). Cette méthode est fortement préconisée afin d'éviter le blocage dans la cellule de flux du fait de sa faible épaisseur.

Du fait de cette conception, un nettoyage est recommandé. Entre chaque échantillon, l'utilisation du « flush » peut suffire. Il est accessible avec le contrôle de la pompe et permet de passer du liquide (ici de l'eau purifiée appelée « milli-Q ») suivant un nombre de cycles déterminé par l'utilisateur. Après avoir passé les billes de calibration pour le focus et avant le démontage le nettoyage peut se faire à partir de l'onglet « Autofocus » afin de visualiser la cellule de flux et contrôler le débit. L'emploi du CitraJet (détergent liquide) peut être utilisé lorsque la cellule de flux a des saletés qui ne peuvent pas être enlevé à l'eau Milli-Q.

2.2.1 Taxonomie

Afin de me former sur la reconnaissance du phytoplancton durant la période de mon stage, j'ai utilisé des outils d'apprentissage divers. Trois livres ont été mis à disposition (Kraberg et al., 2010; Larink & Westheide, 2011; Not, 2011), j'ai également eu accès à un diaporama fait en interne du LOG sur les communautés de phytoplanctons se trouvant dans la Manche orientale et la Baie Sud de la Mer du Nord par Elsa Breton (Ingénieure de Recherche). Je me suis également aidée de deux sites principaux (*biodiversity data provider*; *EOS - Phytoplankton Encyclopedia Project*). De plus, j'ai pu collaborer avec un expert taxonomiste en phytoplancton Fernando Gómez et avec mon encadrant, qui m'ont aidé dans la validation de mon set d'apprentissage.

2.2.2 Travail d'analyse automatisé

Le logiciel R a été d'une grande ressource pour l'automatisation de l'organisation des données et de leur traitement graphique. J'ai pu développer ma connaissance sur ce logiciel notamment grâce aux collègues du groupe « Phyto CytoFluIm » de l'Equipe ECOP² et à diverses sites et blogs internet.

De même pour la compréhension du fonctionnement de l'intelligence artificielle, j'ai pu bénéficier des explications de Guillaume Wacquet (Ingénieur de Recherche IFREMER) via une réunion et des mails échangés ainsi que des cours de doctorant, assisté par Zéline Hubert (Asch, 2022; Kassel, 2022).

Deux scripts ont été créés afin de découper les vignettes et les trier dans les dossiers associés à la date et aux stations des échantillons. La classification s'est faite manuellement (cette démarche de mon stage sera détaillée dans la partie 2.4.2 Création du set d'apprentissage) et les librairies obtenues ont été envoyées à Guillaume Wacquet afin d'utiliser son package « EcoTransLearn » (en cours de publication) basé sur un réseau de neurones qui permet d'évaluer le set d'apprentissage (ici les vignettes de 2019) et de lancer les analyses sur les années suivantes (ici 2021 et 2022).

2.3 Amélioration du protocole d'analyse au sein du LOG

2.3.1 Gain de temps

Ces scripts, de découpages des vignettes à partir du collage du FlowCam et rangement de ces vignettes, permettront de traiter les fichiers du FlowCam plus rapidement et facilement afin d'optimiser le temps pour l'analyse de ces données par la suite.

Au cours de ce stage, j'ai organisé et mis à jour les données acquises sur le FlowCAM depuis 2019 en créant un Excel accessible en fonction des dates, stations et grossissements. Ce fichier m'a servi lors de la sélection des données pour l'analyse temporelle. La mise à jour régulière de ce fichier, en plus des carnets de suivi du laboratoire, permettra d'avoir un regard sur l'ensemble des données acquises jusqu'à ce jour.

2.3.2 Renforcement de la banque d'images

Mon objectif de stage étant d'optimiser le protocole d'analyse des données du FlowCAM, l'amélioration s'est essentiellement faite grâce au renforcement des *librairies* effectuées en 2019 par Eléonore Delecroix (2019). Ces *librairies* étaient constituées de vignettes du FlowCAM en noir et blanc car les configurations du machine learning n'étaient alors associées qu'à la spécificité des caractéristiques obtenues en noir et blanc et provenaient du suivi de la radiale DYPHYRAD. Les dossiers étaient classés suivant différents groupes (Annexe 8).

En utilisant la même base, j'ai rajouté des vignettes en couleurs de l'ensemble des suivis (DYPHYRAD, SRN, Somlit) effectués en 2019. Ce renforcement permet donc d'ajouter une plus-value au set d'apprentissage avec des catégories plus largement représentées et la création de nouveaux groupes.

2.4 Gestion et traitements de données

2.4.1 Configuration du FlowCam pour l'acquisition de la donnée

Le LOG a acquis le FlowCAM 8100 opérationnel en 2019 dans le cadre du projet CPER MARCO. Les paramètres utilisés pour les analyses ont été définis au cours d'un stage en 2019. Au cours de ce stage différents essais pour la configuration des « proches voisins » des particules ont été effectués. Cette configuration (Tableau 1, ligne « Proches voisins ») a été la plus optimisée afin de réussir à capturer l'ensemble des colonies, composées de plusieurs individus séparés, souvent tronquées. Le diamètre minimal des particules défini à 20µm permet la comparaison des particules sur les deux grossissements et d'éliminer les particules trop petites pour être identifiées. Le mode d'acquisition choisi est « l'AutoImage » qui permet la capture d'un nombre d'image à intervalle de temps régulier contrairement au mode « Trigger » qui déclenche la caméra en fonction de la détection d'une valeur de fluorescence. Ce FlowCAM n'est pas muni de laser, le mode « AutoImage » est donc choisi avec une acquisition d'image sur 5 ou 2 mL pour le x4 et x10 respectivement. Le « stop conditions » permet d'avoir un volume échantillonné connu pour chaque station analysée et avoir des données comparables entre elles. Pour chaque analyse, l'association grossissement/cellule du flux/pompe (et éventuellement pré-filtre) doit être respecté afin d'avoir un objectif adapté à l'épaisseur de la cellule de flux et un débit qui permet la bonne numérisation de l'échantillon. Ainsi le FlowCAM est calibré comme suit (Tableau 1) pour l'ensemble des données obtenues depuis 2019. L'ensemble des paramètres (contexte) de l'appareil est détaillés en Annexe 9 et Annexe 10.

TABLEAU 1: METHODE APPLIQUEE SUR LE FLOWCAM POUR L'IMAGERIE DU PHYTOPLANCTON

Paramètres\Grossissement	40	100
Préparation des échantillons	Echantillons frais	
Pré-filtration	X	100µm (blooms coloniaux)
Version VisualSpreadsheet	4.15.1	
Mode d'acquisition	AutoImage	
Proches voisins	60µm	20µm
Diamètre minimal de particules	20µm	20µm
Epaisseur cellule de flux	300µm	100µm

Pompe	5mL	1mL
Stop conditions	5mL de fluide imagé	2mL de fluide imagé

454 échantillons ont été passés sur le FlowCAM dans les deux objectifs de mars 2021 à juillet 2022. La répartition des échantillons suivant les stations est présenté en Annexe 11.

2.4.2 Création du set d'apprentissage

Le set d'apprentissage sert à alimenter l'intelligence artificielle d'images afin qu'elle puisse identifier et « apprendre » les images en fonction des groupes taxonomiques définis. La création d'un set d'apprentissage pour l'analyse du FlowCam réside dans le fait de découper les vignettes du collage (donnée de sortie) afin de constituer un dossier. Cette création de vignette a été effectué automatiquement sous R avec un code modifié (permettant de découper plus de vignettes) de Guillaume Wacquet et son code « ImgProcess ». Le dossier contient ainsi les vignettes correspondant au capture du FlowCam pour l'année d'intérêt servant au set d'apprentissage (ici 2019).

Les images ont pu être triées manuellement suivant la taxonomie la plus fine possible (jusqu'à l'espèce). A l'aide des ressources mises à disposition détaillé en partie 2.2.1 ci-dessus, j'ai constitué les bibliothèques grâce à la comparaison d'images et sélection de critères caractéristiques. J'ai formé des catégories où j'avais des doutes pour revenir dessus. Ce processus a fait l'objet de plusieurs échanges au sein de l'équipe lors de réunions ayant permis d'avoir un regard extérieur en comparant une image d'espèce certifiée (dans les livres ou site web) et une image d'espèce dont j'avais des doutes. J'ai pu bénéficier de l'aide à distance de Fernando Gómez, un taxonomiste spécialiste dans le phytoplancton, qui été alors à bord d'un navire avec mon encadrant de stage. La transmission des dossiers de doutes et d'espèces indéterminés a pu se faire deux fois pendant leur mission en mer avec des points fait avant et après avec Felipe Artigas. Une cinquantaine de dossiers puis une vingtaine ont pu ainsi être étudiés, pour certains solutionnés mais, au cours de ces discussion, comparaison et décisions itératives, les doutes n'ont pas pu être à 100% résolus.

Les catégories ont dues être ajustées et parfois groupées afin d'éviter un sous-apprentissage (minimum de vignettes par catégorie a été fixé à 30 par G.Wacquet) (Tableau 2, Annexe 12). Certaines espèces sous-représentées, et ne pouvant pas être groupées, n'ont pas pu figurer dans le set d'apprentissage Annexe 13. Les groupes constituant le premier set d'apprentissage sont présentés en Annexe 14. Un deuxième set a été créé à la suite des résultats du premier dans le but d'améliorer son apprentissage en prenant compte des résultats de la matrice de confusion

Au total, 9998 images pour le 1^{er} training set et 8196 vignettes pour le 2^{ème} training set ont été analysées pour l'apprentissage au grossissement 40 (objectif x4). Cette différence est due à la modification et suppression de catégories (ex. : « détritux fond » a été supprimé et diverses formes de *Phaeocystis globosa* ont été regroupées afin de former deux catégories : cellules individuelles et coloniales. Concernant le grossissement 100 (objectif x10), 11560 vignettes et 11573 vignettes ont été utilisées respectivement pour le 1^{er} set et 2^{ème}. Ces données viennent de l'ensemble des suivis effectués au cours de l'année 2019 (DYPHYRAD, SRN, SOMLIT).

2.4.3 Outils d'évaluation du set d'apprentissage

Une fois le set d'apprentissage créé, il est soumis à des outils de vérification et d'évaluation statistique sous R. L'algorithme Random Forest donne le taux d'erreur, total et par classe, du set d'apprentissage. Si le taux d'erreur est faible alors les groupes sont bien différenciés et sont identifiables les uns par rapport aux autres. Pour vérifier ces taux d'erreur et donc mesurer le degré de performance du

« machine learning » (algorithme qui apprend par lui-même en fonction des données qu'on lui donne) (datascientest, 2021), un tableau de contingence, appelé matrice de confusion, est utilisé.

Pour calculer la matrice de confusion, le set d'apprentissage est divisé. Une partie (60%) est utilisée pour l'entraînement du « machine learning », ce sont les données tests. L'autre partie (40%) est utilisée pour la validation de l'apprentissage. La table indique en diagonales toutes les particules correctement identifiées et les erreurs en dehors de cette diagonale.

Une validation croisée est obtenue afin de tester les performances de l'outil de reconnaissance en utilisant uniquement le set d'apprentissage. Cette méthode mélange le set d'apprentissage aléatoirement en 10 sous-ensembles de même effectif. L'apprentissage est ensuite effectué sur 9/10 de ce set d'apprentissage et la reconnaissance sur le 1/10 restant. L'opération est effectuée 10 fois avec, à chaque fois, une fraction différente servant au test, celle-ci étant donc toujours exclue du set d'apprentissage réellement utilisé (*comm. pers.* G.Wacquet, Arlot, 2020).

2.5 Analyse des données temporelles et spatiales

Par soucis de temps, le choix des données à analyser s'est porté sur un point stratégique de la radiale DYPHYRAD : R1 (Guiselin, 2010) de 2021 à 2022 pour réaliser une analyse temporelle. Ce point a été le plus échantillonné sur ces deux années et est situé dans les eaux côtières (environ 3km de la côte). L'année 2020 présentant d'importantes lacunes temporelles au printemps du fait de la pandémie de COVID, n'a finalement pas été traitée.

Pour apporter une dimension spatiale aux résultats de la méthode optimisée, les données de la campagne océanographique JERICO/LifeWatch (VLIZ) devaient également être exploitées. Par manque de temps, elle ne sera traitée qu'en septembre.

Le traitement de l'ensemble de ces données est acquis sous la forme d'un tableur Excel et traité graphiquement sous le logiciel R.

3 Résultats

3.1 Validation du set d'apprentissage

L'évaluation des sets d'apprentissage s'est faite sur les deux grossissements deux fois par Guillaume Wacquet. Les résultats de ces apprentissages sont présentés ci-dessous.

3.1.1 Grossissement 100

Banque d'images

Les groupements effectués sont présentés dans le Tableau 2 ci-dessous. Nous pouvons voir que différentes espèces de la diatomée coloniale du genre *Chaetoceros* identifiées (Cf 1^{ère} colonne) sont assemblées au rang de genre des *Chaetoceros* (Cf 2^{ème} colonne). D'autres espèces ou genres, ne comptant pas assez d'individus, sont rassemblées avec d'autres espèces au rang de famille ou d'embranchement (Cf Famille Gymnodiniaceae ou Embranchement Ciliates par exemple) ou supprimées (Annexe 13). Les *Pseudo-nitzschia* et les *Cylindrotheca closterium* sont semblables morphologiquement et font parties de la famille des Bacillariaceae.

TABLEAU 2 : GROUPEMENT POUR LES VIGNETTES EN X10

Diatomées	Nb de vignettes	Rang : Groupement
<i>Actinoptychus</i>	6	Ordre : Coscinodiscales
<i>Coscinodiscus spp</i>	38	
<i>Cerataulina pelagica</i>	111	Espèce
<i>Chaetoceros sp</i>	348	Genre : Chaetoceros spp
<i>Chaetoceros eibenii</i>	3	
<i>C.concortus</i>	1	
<i>C.curvisetus</i>	21	
<i>C.affinis</i>	6	
<i>C.danicus</i>	11	
<i>Chaetoceros socialis</i>	154	
<i>Ditylum brightwellii</i>	132	Espèce
<i>G. flaccida</i>	129	Espèce
<i>G.delicatula</i>	42	Genre : Guinardia
<i>G.fragilissima</i>	2	
<i>G.striata</i>	32	
<i>Leptocylindrus danicus</i>	1921	Espèce
<i>Pseudo-nitzschia</i>	464	Famille: Bacillariaceae
<i>Cylindrotheca closterium</i>	2	
<i>Rhizosolenia imbricata</i>	219	Espèce
<i>Rhizosolenia pungens</i>	45	Espèce
<i>Raphoneis amphiceros</i>	85	Espèce
<i>Navicula spp</i>	61	Ordre : Naviculales
<i>Meuniera membranacea</i>	17	
<i>Pleurosigma</i>	16	
<i>Thalassionema nitzschioides</i>	55	Espèce
<i>Skeletonema</i>	16	Ordre : Thalassiosirales
<i>Lauderia annulata</i>	10	
<i>Thalassiosira sp</i>	4	
<i>Thalassiosira rotula</i>	11	
<i>Thalassiosira punctigera</i>	3	
<i>Detonula sp.</i>	1	
Haptophytes		Rang : Groupement
<i>Phaeocystis globosa multiple</i>	665	Espèce
<i>P. coloniales</i>	15	Espèce coloniales
<i>P. ind.</i>	524	Espèce individu
Dinoflagellés	Nb de vignettes	Rang : Groupement
<i>Gyrodinium sp</i>	32	Famille : Gymnodiniaceae
<i>G.spirale</i>	50	
<i>K.glaucum</i>	27	
<i>Torodinium sp</i>	6	
<i>Amphidinium</i>	1	
<i>Gymnodinium</i>	2	
Dinoflagellates	32	Embranchement : Dynophyta
Dino clair	4	
<i>Protoperidinium</i>	15	Famille : Protoperidiniaceae
<i>Diplopsalis spp</i>	41	
Ciliés		Rang : Groupement
Ciliés spp	105	Embranchement : Ciliates
<i>Laboa strobila</i>	8	
<i>Lohmanniella oviformis</i>	3	

Mesodinium	1	
Tiarina sp.	3	

Training set

Ci-après les tableaux (Tableau 3) récapitulant les espèces pour les 2 apprentissages effectués suivant les groupes précédemment formés. 60% des vignettes de chaque groupe ont servies au « training » (vert) et 40% ont servies à la validation (bleu). Une limite à 2000 images (**) a été fixé pour éviter le sur-apprentissage.

TABEAU 3: TABLEUX RECAPITULATIFS DES GROUPES CONSTITUANTS LES SETS D'APPRENTISSAGE POUR L'OBJECTIF X10.

1er set du 23 juin		Train	Test
Bacillariaceae	462	278	184
Billes	2000**	1200	800
Bulles	483	290	193
Cerataulina_pelagica	111	67	44
Chaetoceros_socialis	154	92	62
Chaetoceros_spp	390	234	156
Ciliates	120	72	48
Coscinodiscales	44	26	17
Ditylum_brightwellii	132	80	52
Dynophyta	42	26	16
Eau	123	74	49
Flou	216	130	86
Fond	2000**	1200	800
Guinardia_flaccida	129	78	51
Guinardia_spp	74	45	29
Gymnodiniaceae	114	70	44
Leptocylindrus_danicus	1919	1150	769
Naviculales	98	59	39
Phaeocystis_globosa	675	404	271
Phaeocystis_globosa_Cox	198	119	79
Phaeocystis_globosa_Ind	524	314	210
Protoperidiniaceae	42	26	16
Rhaphoneis_amphiceros	85	51	34
Rhizosolenia_imbricata	219	132	87
Rhizosolenia_pungens	45	27	18
Thalassionema_nitzschoides	55	33	22
Thalassiosirales	45	27	18
Vide	1062	637	425
Total	11560	6941	4619

2e set du 19 juillet		Train	Test
Bacillariaceae	462	278	184
Billes	2000**	1200	800
Bulles	483	290	193
Cerataulina_pelagica	109	66	43
Chaetoceros_spp	544	326	218
Ciliates	120	72	48
Coscinodiscales	43	25	17
Ditylum_brightwellii	132	80	52
Dynophyta	41	26	15
Eau	123	74	49
Flou	216	130	86
Fond	2000**	1200	800
Guinardia_flaccida	129	78	51
Guinardia_spp	76	46	30
Gymnodiniaceae	115	70	45
Leptocylindrus_danicus	1919	1153	766
Naviculales	98	59	39
Phaeocystis_globosa	675	404	271
Phaeocystis_globosa_Col	211	127	84
Phaeocystis_globosa_Ind	524	314	210
Protoperidiniaceae	42	26	16
Rhaphoneis_amphiceros	85	51	34
Rhizosolenia_imbricata	219	132	87
Rhizosolenia_pungens	45	27	18
Thalassionema_nitzschoides	55	33	22
Thalassiosirales	45	27	18
Vide	1062	637	425
Total	11573	6952	4621

La précision et les erreurs commises lors de l'apprentissage par le modèle VGG16 sont présentées suivant les 2 graphiques (Annexe 15) pour les 2 sets construits.

La courbe « Accuracy » montre le pourcentage de vignettes correctement classifiées. Pour les 2 sets, le score est similaire avec une validation sur l'ensemble de « training » (validation croisée) et sur l'ensemble de validation (prédiction à partir du set test) de 0.9 soit proche de 100%.

Le graphique « Loss » est obtenu suivant la mise à jour du modèle en fonction de ces erreurs entre les données réelles et prédites. Il constitue la courbe d'apprentissage d'optimisation. Sur les 2 sets, les courbe du « training » et de la validation suivent la même tendance et se rapproche du 0. Le tracé de la perte de validation diminue jusqu'à un point de stabilité et présente un petit écart avec la perte d'apprentissage.

La matrice de confusion faite pour chaque set est présentée ci-dessous (Figure 6 et Figure 8). Elle rend compte des confusions du modèle de classification lorsqu'il fait des prédictions (Brownlee, 2016). Les

lignes de la matrice correspondent aux vignettes prédites donc attendu venant du jeu de données et les colonnes correspondent à la classe réelle.

Actual // Predicted	01	03	05	07	09	11	13	15	17	19	21	23	25	27		
Phaeocystis_globosa 01	204	12	26	6		1		1	1		2		2	1	15	01
Phaeocystis_globosa_Cox 02	46	5	11	3			1					1			12	02
Chaetoceros_socialis 03	3	2	11	35	5	1	1			1		1			2	03
Chaetoceros_spp 04	7		1	121	13	1	3		1		2	1			5	04
Leptocylindrus_danicus 05	3	2	6	734	3	4	3	9	1				1	1	2	05
Rhizosolenia_imbricata 06	1		2	46	29	1	3	2	3							06
Rhizosolenia_pungens 07				5	1	10	1	1								07
Bacillariaceae 08	4			21	4	1	148	2						1	3	08
Naviculales 09	1	1	2	2		14	6	1	1	3			2		3	09
Guinardia_spp 10		1	2	10			10	4			1	1				10
Cerataulina_pelagica 11				11	1		1	5	18	4	2		1	1		11
Guinardia_flaccida 12						1	1	9	40							12
Dynophyta 13	1		1		1			7	3		1	1			1	13
Ciliates 14			7	1			1	8	17	5	6			1	1	14
Ditylum_brightwellii 15			1	2	2	1		4	6		8	23	1	4		15
Rhaphoneis_amphiceros 16	2		2	1	1		1		2		19	3	1		1	16
Gymnodiniaceae 17		2	4	2		1		3	4	1	18	5	1	1	1	17
Protoperidiniaceae 18			1					4	1	1	2	4	1			18
Coscinodiscus 19						1		3	2	1		8	1		1	19
Thalassiosirales 20	4		1				2		5	1	1	1	2		1	20
Phaeocystis_globosa_ind 21		10	4		1					2				2	1	21
Thalassionema_nitzschoides 22	3		4	1	2	1			2	2		7				22
Billes 23													796	4		23
Bulles 24										1			26	162	1	24
Eau 25			1								3	1	3	28	5	25
Flou 26					2				1					15	64	26
Fond 27	3				1	7		1						4	784	27
Vide 28	2				1						2		5	9	16	28
	01	03	05	07	09	11	13	15	17	19	21	23	25	27		

FIGURE 6: MATRICE DE CONFUSION DU 1^{ER} SET A L'OBJECTIF X10 AU FLOWCAM

D'après cette matrice du 1^{er} set, les principales confusions sont entre « *Phaeocystis globosa Cox* » et « *Phaeocystis globosa* », « *Chaetoceros socialis* » et « *Chaetoceros spp* », « *Rhizosolenia imbricata* » et « *Leptocylindrus danicus* » (Figure 7). Ces confusions s'expliquent par une ressemblance entre les cellules (e. et f.) mais également par une mauvaise identification.

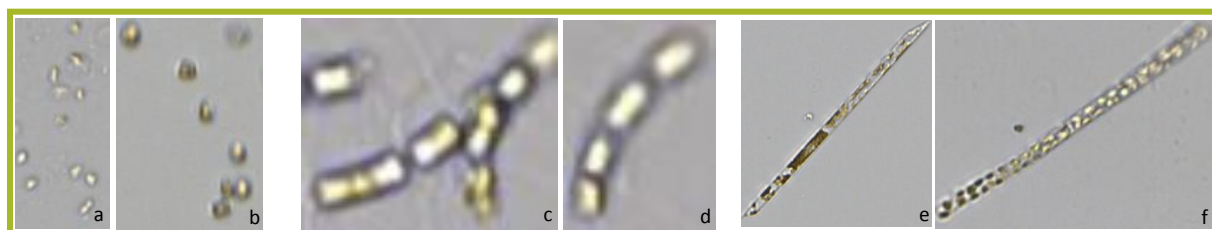


FIGURE 7: EXEMPLE DE VIGNETTES AU X10 (A. *PHAEOCYSTIS GLOBOSA COX*, B. *PHAEOCYSTIS GLOBOSA*, C. *CHAETOCEROS SOCIALIS*, D. *CHAETOCEROS SPP*, E. *RHIZOSOLENIA IMBRICATA*, F. *LEPTOCYLINDRUS DANICUS*)

Suite aux modifications dans le but d'améliorer l'apprentissage, les principaux changements effectués entre les deux sets ont été réalisés sur la bibliothèque d'images de l'espèce *Cerataulina pelagica* et *Leptocylindrus danicus*, *Chaetoceros socialis*, *Phaeocystis globosa* coloniale. Ces changements restent minoritaires du fait du temps accordé et des résultats plutôt satisfaisants dans l'ensemble de l'apprentissage.

TABLEAU 4: TABLEAUX RECAPITULATIFS DES GROUPES CONSTITUANTS LES SETS D'APPRENTISSAGE POUR LE X4.

1er set du 23 juin		Train	Test
Cerataulina_spp	110	66	44
Chaetoceros_curvisetus	208	125	83
Chaetoceros_socialis_chaine	109	66	43
Chaetoceros_socialis_pelote	62	37	25
Chaetocerotaceae	397	238	159
Ciliophora	113	68	45
Coscinodiscales	32	20	12
Cylindrotheca_closterium_Phaeo	166	100	66
Detritus	1132	680	452
Detritus_fond	2000**	1200	800
Ditylum_brightwellii	77	47	30
Eau	37	23	14
Fibre	62	36	26
Flou	406	244	162
Fond	79	48	31
Frustules_cassees_diatomees	47	29	18
Guinardia_delicatula	138	83	55
Guinardia_flaccida	219	132	87
Guinardia_striata	379	228	151
Gymnodiniaceae	119	72	47
Leptocylindrus_danicus	1382	830	552
Naviculales	104	62	42
Odontella_regia	118	71	47
Paralia_sulcata	64	38	26
Phaeocystis_globosa	577	346	231
Phaeocystis_globosa_Col	193	116	77
Phaeocystis_globosa_G_Cox	102	62	40
Phaeocystis_globosa_P_Cox	258	155	103
Prorocentrum_spp	44	27	17
Protoperidiniaceae	73	44	29
PseudoNitzschia_spp	243	146	97
Rhaphoneis_amphiceros	65	40	25
Rhizosolenia_imbricata	444	266	178
Rhizosolenia_pungens	69	42	27
Sphere	186	112	74
Thalassionema_nitzschoides	37	23	14
Thalassiosirales	84	51	33
Zooplankton	63	38	25
Total	9998	6011	3987

2e set du 19 juillet		Train	Test
Ciliophora	114	69	45
Detritus	1132	680	452
Asterionellopsis_glacialis	31	19	12
C.curvisetus	236	142	94
C.socialis_Chaine	109	65	44
C.socialis_Pelote	65	39	26
Cerataulina_sp	112	67	45
Chaetocerotaceae	370	222	148
Coscinodiscales	32	19	13
Cylindrotheca_closterium_et_Phaeo	171	103	68
Ditylum_brightwellii	77	46	31
G.delicatula	138	83	55
G.flaccida	221	133	88
G.striata	379	227	152
Leptocylindrus_danicus	1382	829	553
Meuniera_membranacea	55	33	22
Naviculales	50	30	20
Paralia_sulcata	65	39	26
P-nitzschia_spp	244	146	98
R.imbricata	447	268	179
Rhaphoneis_amphiceros	65	39	26
Rhizosoleniaceae	85	51	34
Thalassionema_nitzschoides	37	22	15
Thalassiosirales	88	53	35
Triceratiaceae	145	87	58
Eau	37	22	15
Fibre	65	39	26
Gymnodiniaceae	122	73	49
P.globosa	1099	659	440
P.globosa_Col	121	73	48
Prorocentrum_sp	44	26	18
Protoperidiniaceae	73	44	29
Flou	407	244	163
Fond	79	47	32
Frustules_cassées_de_diatomées	48	29	19
Sphere	186	112	74
Zooplankton	65	39	26
Total	8196	4918	3278

La précision et les erreurs commises lors de l'apprentissage par le modèle VGG16 sont présentées suivant les 2 graphiques (Annexe 17) pour les 2 set construits.

La courbe « Accuracy » montre le pourcentage de vignettes correctement classifiées. Pour les 2 sets, le score est similaire avec une validation sur l'ensemble de « training » (validation croisée) de 0.78 et sur l'ensemble de validation (prédiction à partir du set test) d'environ 71%.

Le graphique « loss » est obtenu suivant la mise à jour du modèle en fonction de ces erreurs entre les données réelles et prédites. Il constitue la courbe d'apprentissage d'optimisation. Sur les 2 sets, les courbe du « training » et de la validation suivent la même tendance et se rapproche du 0. Le tracé de la perte de validation diminue jusqu'à un point de stabilité et présente un écart de 0.4 avec la perte d'apprentissage.

La matrice de confusion faite pour chaque set est présentée (Figure 9 et Figure 11)). Elle rend compte des confusions du modèle de classification lorsqu'il fait des prédictions (Brownlee, 2016). Les lignes de la matrice correspondent aux vignettes prédites donc attendu venant du jeu de données et les colonnes correspondent à la classe réelle.

Actual // Predicted	01	03	05	07	09	11	13	15	17	19	21	23	25	27	29	31	33	35	37
PseudoNitzschia_spp 01	22	53	9	5	1	1	5	1											
Leptocylindrus_danicus 02	9	47	20	1	14		1	10	5	1	9							2	6
Rhizosolenia_imbricata 03	2	7	85			1	1	3	3		12							1	1
iaeocystis_globosa_P_Cox 04				88	5	51	1	1			1								
Detritus_fond 05	1	2	3	74	13		6	9	14		1						2		
Phaeocystis_globosa 06			10	57	41	5	5	6	2	5									
Irotheca_closterium_Phao 07	1	3	2	15	3	5	17	1	7			2	3				6		
Phaeocystis_globosa_Col 08			1	1	46	5	18		3		2	1							
Chaetoceros_curviretus 09	6		5			20	48												3
Chaetoceroaceae 10	1	19	1	12	12	1	4	79	9	1	5	3		1	1		3	1	
Sphere 11						53	16	5											
Detritus 12			13	6		1	4	2	38	22	4	1			1	3	3	1	
Flou 13			53	1				8	98									1	1
Guinardia_flaccida 14						1	1	2	1	89	7	3	1						
Naviculales 15	1	4				6	7	1	2	7	3	2		1	3	1	1		
Guinardia_delicatula 16	11	7	1			1	7	1	1	2	15	4	2	1					
Cerataulina_spp 17	11	2				2	1	1	4	1	7	10	1					1	3
Thalassiosirales 18	1	1	1	1		3	1	1		3	11	7			2				
iaeocystis_globosa_G_Cox 19					3						86								
iaetoceros_socialis_chaine 20	1	3	4	2		3	3	1	1	1	1	14	7				1		
haetoceros_socialis_pelote 21			2	1		1	1	1			5	3	8					1	1
Gymnodiniaceae 22					1	1	1	8		2			25	2	1	2	2		
Paralia_sulcata 23	1						13		1	2			2	4	1	1			
Prorocentrum_spp 24						1	5		1			1	1	1	5	1	1		
Ciliophora 25			1				14	1		1	1	1	1	10	7		1	3	1
Protoperidiniaceae 26											4	1	3	19	1				
Coscinodiscals 27							1	3		1				1	5	1			
Odontella_regia 28					1			2	1	1	2			4		34			1
ialassionema_nitzschoides 29		1	1		1	1	9									1			
Rhizosolenia_pungens 30	1	9	10						2							5			
Rhaphoneis_amphiceros 31			1	3		1	3	6	2			1	1		1		6		
stules_cassees_diatomees 32						1	6						1				9		
Eau 33				1					2									10	
Fond 34																		30	
Ditylum_brightwellii 35							3	1	2	1	1	1		1		1			16
Zooplankton 36											1	2							4
Fibre 37	2					1		1				1		1				1	2
Guinardia_striata 38	3					6	2	1		2	2	1	3	1	2		2	1	1

FIGURE 9: MATRICE DE CONFUSION DU 1^{ER} SET A L'OBJECTIF X4 AU FLOWCAM

Les principales confusions faites lors du premier set d'apprentissage concernent surtout l'espèce *Leptocylindrus danicus* notamment avec le genre *Pseudo-nitzschia* et *Rhizosolenia imbricata* (Figure 10). L'espèce *Phaeocystis globosa* est confondue avec ses différentes formes (individuelle et coloniale). Les groupes non vivants comme « Détritus fond » et « Flou » augmentent également les erreurs du fait de leur ressemblance.

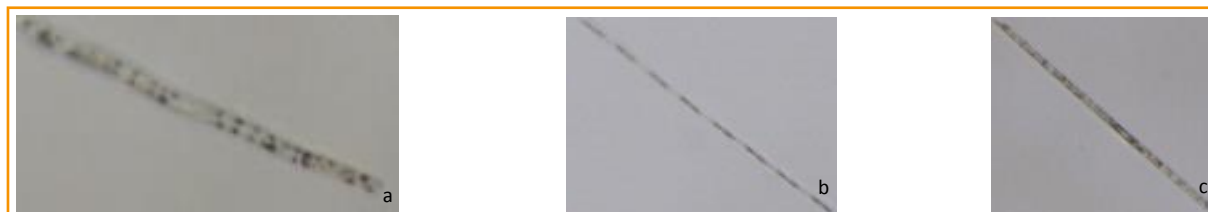


FIGURE 10: EXEMPLE DE VIGNETTE AU X4 (A. *LEPTOCYLINDRUS DANICUS*, B. *PSEUDO-NITZSCHIA*, C. *RHIZOSOLENIA IMBRICATA*)

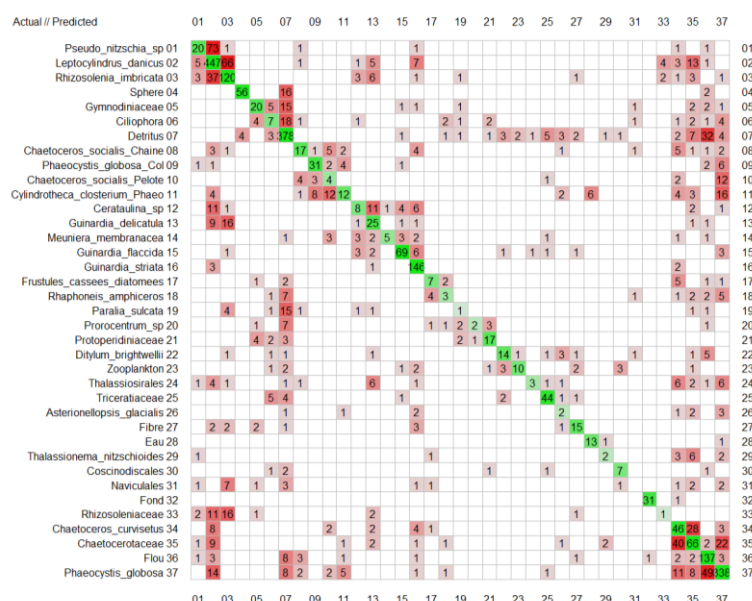


FIGURE 11: MATRICE DE CONFUSION DU 2E SET A L'OBJECTIF X4 AU FLOWCAM

Suite aux modifications dans le but d'améliorer l'apprentissage, les principaux changements effectués entre les deux sets ont été la suppression de la catégorie « detritus fond », l'ajout de *Asterionellopsis glacialis* (Annexe 18) avec la technique de rotation d'image, c'est-à-dire d'enregistrer l'image sous différentes orientations (Wallès, 2020) et de l'espèce *Meuniera membranacea* (Annexe 18). La famille Triceratiaceae a été créé afin de rassembler le genre *Trieres* dont les espèces étaient peu représentées. Deux formes de *Phaeocystis globosa* (G.Cox et P.Cox) présent initialement dans le set de 2019 ont été rassemblées dans la catégorie plus large des *Phaeocystis globosa*.

Pour le 2^e set, les erreurs sont globalement entre *Phaeocystis globosa* et la catégorie des vignettes « Flou » qui sont des particules trop floues pour être identifiées. Les détritits sont également confondus avec cette catégorie. On retrouve une confusion mutuelle entre *Chaetoceros curvisetus* et la famille pour laquelle elle appartient (Chaetocerotaceae). L'espèce *Leptocylindrus danicus* cause toujours les confusions citées au 1^{er} set.

3.1.3 Comparaison des apprentissages entre les deux grossissements

En se focalisant sur les derniers sets (qui sont utilisés par la suite pour les analyses), on remarque que l'apprentissage se fait plus facilement sur le grossissement 100. La probabilité d'image correctement classée rend compte de ces résultats (Figure 12).

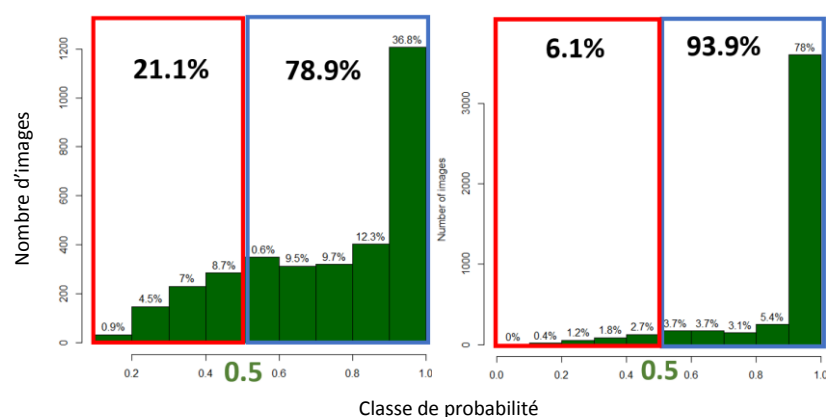


FIGURE 12: GRAPHIQUE REPRESENTANT LA PROBABILITE D'APPARTENANCE (A GAUCHE POUR LE SET X4, A DROITE LE SET X10)

Pour le x4, 80% des vignettes ont une probabilité supérieure à 50% d’être correctement classées tandis que 94% des vignettes au x10 ont cette probabilité. Seulement 36,8% des vignettes au x4 ont une probabilité d’appartenance entre 90 et 100% contre 78% au grossissement 100.

3.2 Résultats des analyses sur le point R1 de la radiale DYPHYRAD du printemps 2021 à l’été 2022.

3.2.1 L’analyse des particules du FlowCam en x4

L’analyse de l’ensemble de l’échantillon permet d’intégrer la totalité de l’échantillon qui passe dans le FlowCAM (Figure 13). Ce graphique résume l’abondance absolue en fonction des dates des échantillons R1 passés sur le grossissement 40. On observe que la catégorie « non vivant » (comprenant les billes, bulles, eau, flou, fond, vide, sphère, fibre, détritux) est majoritaire dans un échantillon. Plus de 7000 particules ont été détectés sur le FlowCAM au printemps.

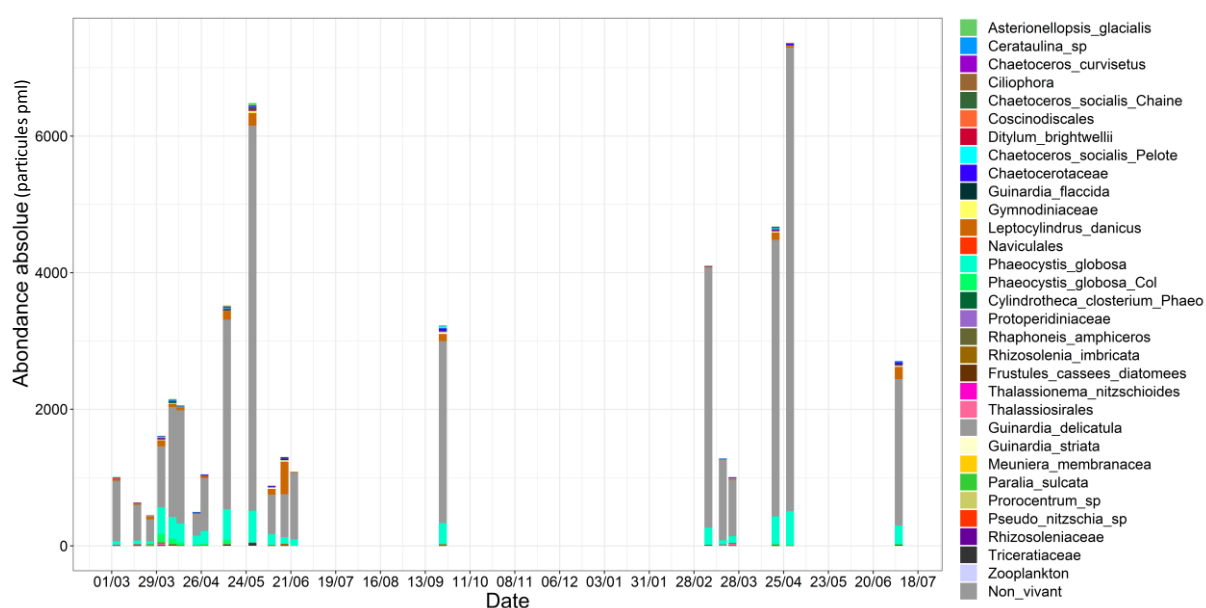


FIGURE 13: GRAPHIQUE DES PARTICULES PASSEES SUR LE FLOWCAM AU GROSSISSEMENT 40 DE MARS 2021 A JUILLET 2022 AU POINT R1

Les échantillons, au grossissement 40, sont essentiellement passés en périodes de blooms (au printemps) et peu sont passés en été et en hiver dû à la priorisation du grossissement 100 pour une meilleure visibilité des particules.

On peut également observer une augmentation de l’abondance des *Phaeocystis globosa* individuel et colonial sur la même période les deux années (au printemps). L’observation des organismes vivants sera faite ci-après grâce à la suppression des particules non vivantes.

3.2.2 L’analyse des particules du FlowCam en x10

Les particules non vivantes suivant leur apprentissage correspondent, sur le x 10, au billes, bulles, eau, flou, fond, vide. Ces particules sont moins présentes dans ce grossissement que dans les échantillons du x 4. Les groupes d’organismes vivants sont plus représentés dans ce grossissement (Figure 14).

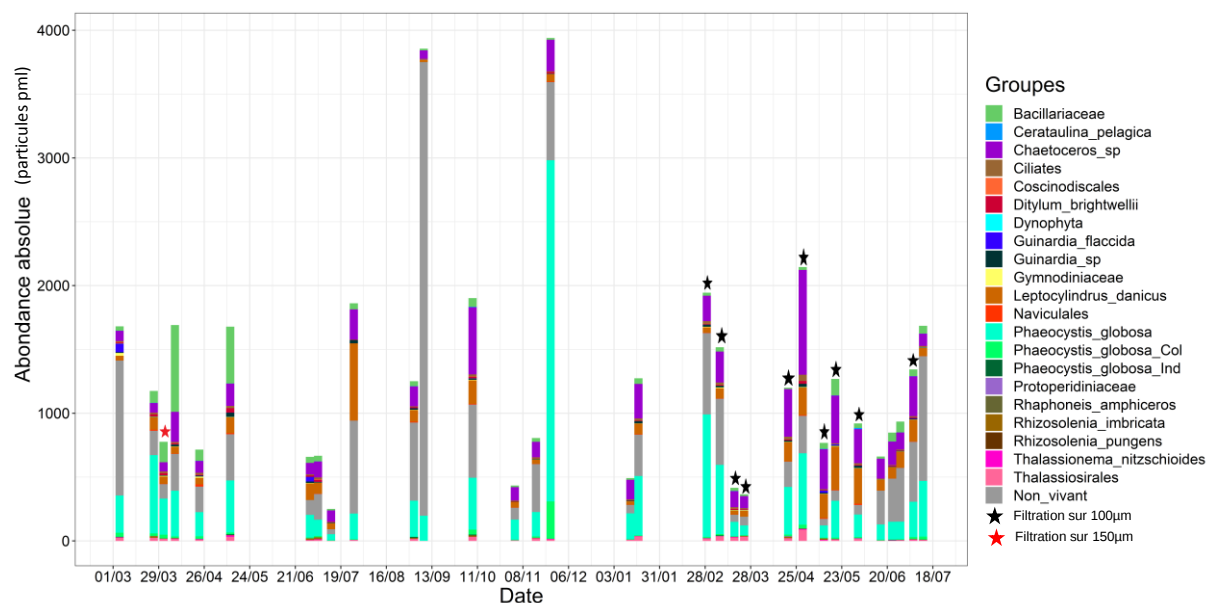


FIGURE 14: GRAPHIQUE DES PARTICULES PASSEES SUR LE FLOWCAM AU GROSSISSEMENT 100 DE MARS 2021 A JUILLET 2022 AU POINT R1

Dans l'ensemble, moins de particules sont échantillonnées sur ce grossissement par rapport au x4 (moins de 4000 particules pour le x10 contre 7000 pour le x4). Les pré-filtrations recensées à l'aide des rapports et du cahier de laboratoire sont représentées d'une étoile.

3.3 Résultats des analyses temporelles des communautés phytoplanctoniques

Afin de mieux observer la dynamique du phytoplancton, les groupes inertes ont été supprimés des graphiques.

3.3.1 Evolution saisonnière de l'abondance phytoplanctonique sur le point R1 en 2021 et 2022

3.3.1.1 Grossissement x4

Pour rappel, le grossissement 40 est utilisé afin de mieux observer les colonies phytoplanctoniques. En effet sous ce grossissement, les échantillons ne sont pas filtrés et la combinaison cellule de flux/grossissement permet d'avoir une meilleure vision de la totalité de la particule.

Dans l'ensemble, le nombre d'organismes augmente au printemps (Figure 15) avec un pic en cumulé à plus de 800 particules par millilitres fin mai.

Le graphique de l'abondance relative (Figure 16) permet de refléter la représentativité des organismes sur l'ensemble de l'échantillon. On remarque ici que l'espèce *Phaeocystis globosa* est largement représentée au printemps, sa forme coloniale suit également cette dynamique. On remarque également la présence de *Leptocylindrus danicus* en mars et juin. La famille des Thalassiosirales apparaît en mars sur les deux années étudiées.

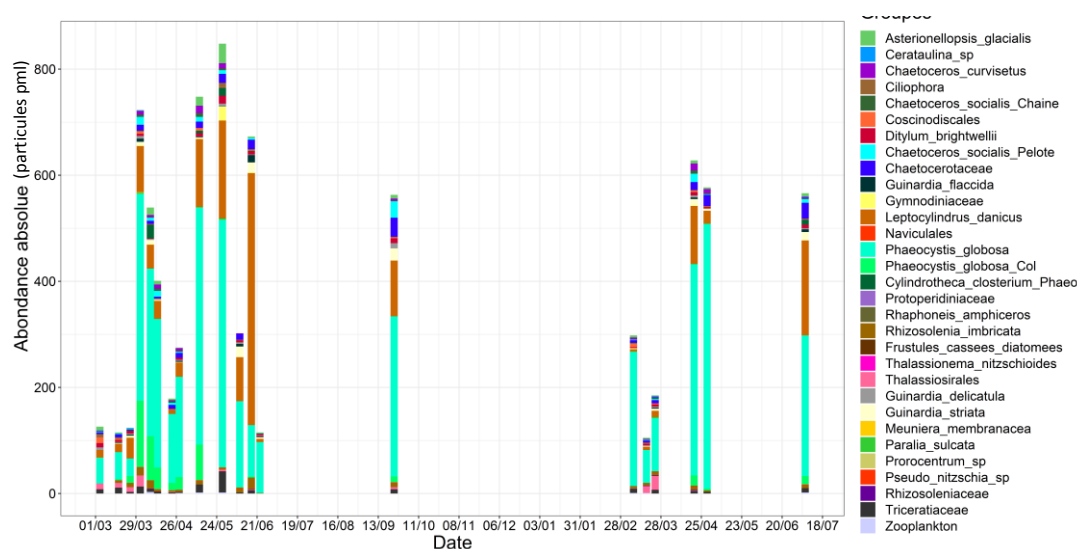


FIGURE 15: GRAPHIQUE DE L'ABONDANCE ABSOLUE DU PHYTOPLANCTON AU GROSSISSEMENT 40 DE MARS 2021 A JUILLET 2022 AU POINT R1

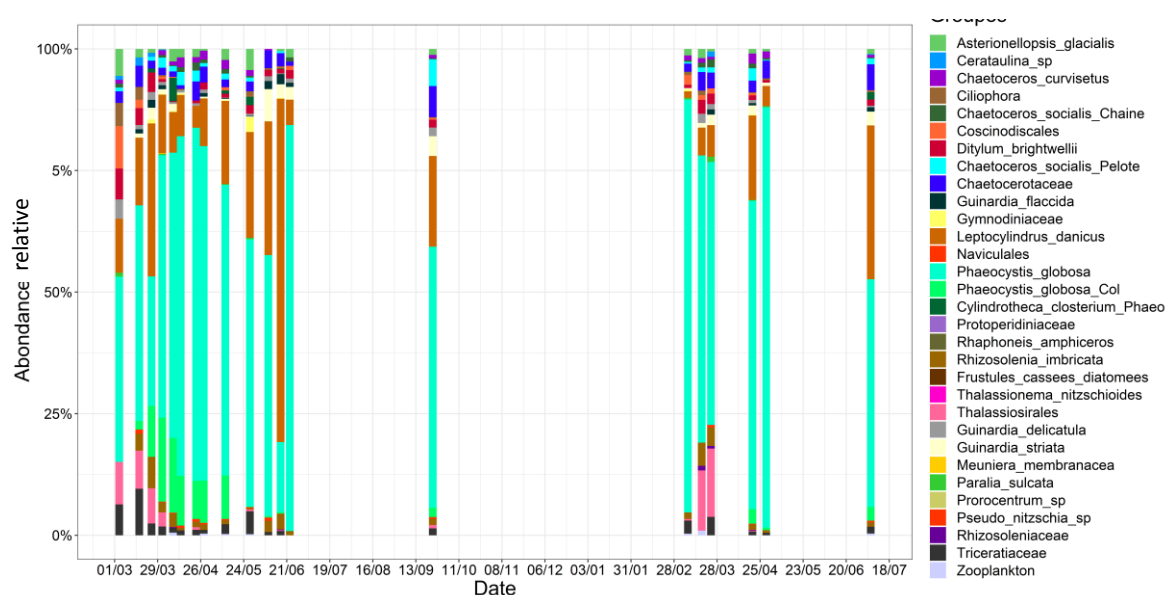


FIGURE 16: GRAPHIQUE DE L'ABONDANCE RELATIVE DU PHYTOPLANCTON AU GROSSISSEMENT 40 DE MARS 2021 A JUILLET 2022 AU POINT R1

3.3.1.2 Grossissement x10, dynamique des cellules individuelles phytoplanctoniques

Les individus de *Phaeocystis globosa* sont, d'après le graphique des abondance absolue (Figure 17), présents tout au long de l'année. La présence de ces organismes varie au cours de l'année, on observe une diminution en été et une augmentation fin hiver, début du printemps (février/mars).

Le groupement des Bacillariaceae, qui regroupe les *Cylindrotheca closterium* et *Pseudo-Nitzschia*, sont plus abondants au printemps (mars/avril/mai). On remarque que ces individus étaient plus présents au printemps 2021 que cette année. Les individus du genre *Chaetoceros* sont également présents tout au long de l'année. Les espèces de la famille de Thalassiosirales sont plus représentés dans les échantillons du mois de mars 2022 (Figure 18).

L'abondance absolue des organismes passés au grossissement 100 est d'environ 800 à 1000 particules par millilitre. On observe un pic de plus de 3000 particules enregistrées le 25 novembre qui correspond à un artefact dû à l'enregistrement de la cellule de flux vide et la confusion avec l'espèce *Phaeocystis globosa*.

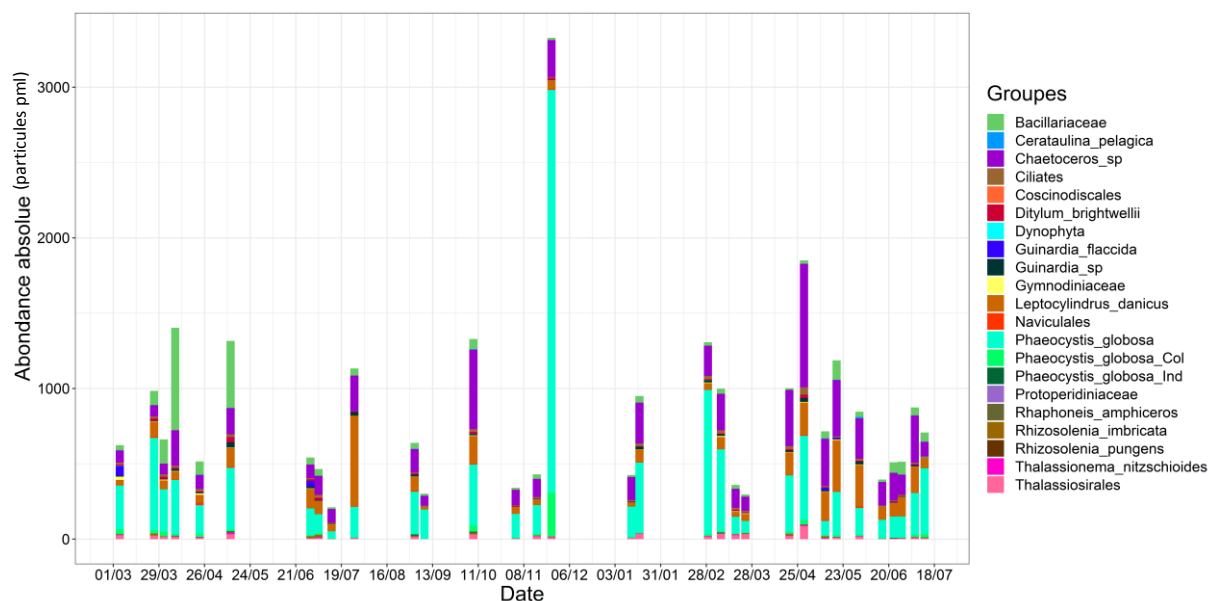


FIGURE 17: GRAPHIQUE DE L'ABONDANCE ABSOLUE DU PHYTOPLANKTON AU GROSSISSEMENT 100 DE MARS 1 A JUILLET 2022 AU POINT R1

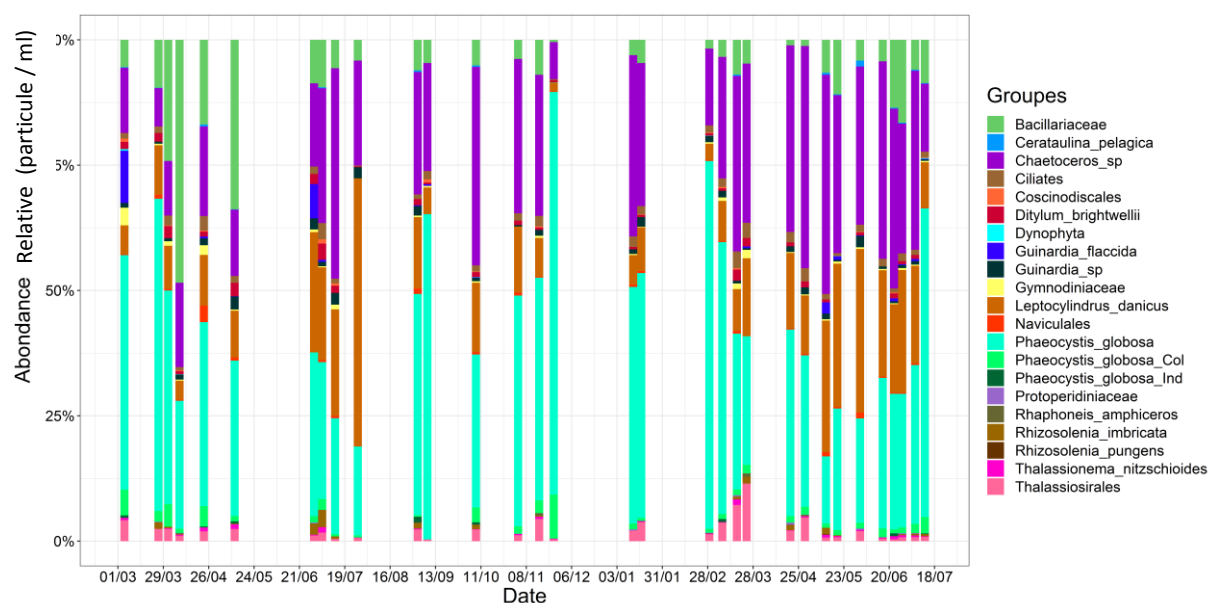


FIGURE 18: GRAPHIQUE DE L'ABONDANCE RELATIVE DU PHYTOPLANKTON AU GROSSISSEMENT 100 DE MARS 2021 A JUILLET 2022 AU POINT R1

3.3.1.3 Comparaison des résultats entre le grossissement 40 et 100

Le traitement de ces résultats au deux grossissements montre une certaine complémentarité. En effet, afin d'avoir une vue d'ensemble de la dynamique du phytoplancton sur ces 2 années, il est essentiel de prendre en compte les colonies qui sont représentées au x4 non pré-filtré et les individus phytoplanctoniques plus visibles au x10.

Dans l'ensemble, les résultats montrent qu'il y a une régularité sur l'objectif x10 sur l'année contrairement à l'objectif x4 qui présente des lacunes dû aux changements d'objectifs.

L'abondance des organismes vivants suivant la classification et l'apprentissage de réseaux de neurones est plus élevé au x10. C'est-à-dire que les abondances absolues observées pour les deux

grossissements sont plus élevées (entre 500 et 1000) pour le grossissement 100 que pour le 40 (entre 200 et 700).

Les résultats de l'objectif 10 semblent également donner des informations plus précises quant aux communautés avec de plus grandes variations entre celle-ci. Ce dernier constat peut être biaisé par l'utilisation de mailles de filtration différentes rendant la comparaison peu significative.

4 Discussion

4.1 Méthode

La méthode de la création du set d'apprentissage est limitée dans son efficacité. En effet, la création d'une bibliothèque (*library*) demande beaucoup de temps suivant la qualité (le nombre de vignettes) que l'on souhaite. Dans le cadre de mon stage, la contrainte du temps a réduit le nombre de vignettes collectées. Des espèces ont été sous-représentées (<30) du fait du manque d'images sur l'année choisie (2019) ou de leur mauvaise qualité qui ne permettait pas une identification assurée.

Pour éviter ce sous-apprentissage, défini par G.Wacquet, et homogénéiser la répartition des vignettes dans chaque catégorie (au moins 100 images par catégorie), le groupement des espèces en fonction de leur taxonomie a été choisi. Les espèces ne pouvant être regroupées et n'excédant pas 30 images ont été supprimées. Ce groupement engendre une perte de précision dans les futures analyses. La famille, l'ordre jusqu'à l'embranchement a dû être choisie afin de regrouper certaines espèces minoritaires dans les bibliothèques créées. La perte d'information due au manque de données, et au regroupement, est certaine mais modérée car cela concerne des espèces qui ne sont pas souvent présentes en Manche orientale et sud Mer du Nord (Devreker & Lefebvre, 2022; Lamy, 2006).

La méthode de création de la bibliothèque représente un biais de la part du créateur dans l'attribution des catégories de chaque espèce. Les confusions peuvent exister d'autant plus quand le créateur n'est pas taxonomiste. Cependant, le biais généré par l'Homme peut être supérieur à celui généré par une machine dû aux biais cognitifs que nous subissons (Culverhouse, 2007; Culverhouse et al., 2003). L'amélioration de la méthode de formation de la machine sera donc bénéfique pour la surveillance de l'environnement grâce à l'augmentation de la fréquence d'analyse par rapport aux comptages manuels.

Un autre biais existe concernant la manipulation au laboratoire. Suivant les années et les utilisateurs du FlowCAM (stagiaires, ingénieurs, doctorants) la méthode d'acquisition des données peut changer. En effet nous avons pu voir dans la répartition des analyses qu'il y a eu des changements de protocoles. Les protocoles concernant la pré-filtration des échantillons au grossissement 100 reste subjectif suivant l'état de l'échantillon et sa densité en particules. L'utilisation de la maille de 100 μm pour la préfiltration représente un biais. Le choix de la filtration sur une maille de 100 μm à différents moments de l'année pourrait être discutable. La filtration sur du 100 μm a été choisie avant tout pour empêcher le colmatage des particules à l'entrée de la cellule de flux de 100 μm (objectif x10). Cependant, la filtration représente une perte de données qui n'est pas visible directement. Dans cet inconnu, présent sur la maille de 100 μm , on suggère qu'il y a toutes les particules « à risque », supérieures à 100 μm , mais une analyse de la maille > 100 μm devrait être réalisée afin de quantifier les cellules < 100 μm qui n'ont pas pu passer. En 2021, le choix sur la préfiltration avait été fait sur une maille de 150 μm pour éviter le colmatage mais également pour discriminer les phytoplanctons proches des 200 μm qui sont une limite statistique du microplancton (20-200 μm) avant les grosses colonies. Ce choix s'est également porté car l'IFCB utilise 150 μm pour imager le phytoplancton. La filtration en 2021 n'était pas non plus aussi régulière (ou non signaler) qu'en 2022.

La comparaison de la méthode utilisée au LOG et dans les autres laboratoires utilisant le FlowCAM reste limitée. En effet, d'après l'article d'Owen et al. (2022), qui dresse un rapport des méthodes pour l'étude du phytoplancton avec l'utilisation du FlowCAM, peu de méthodes sont correctement répertoriées. Ainsi, quelques travaux peuvent se rapprocher de nos échantillonnages (modèle du FlowCAM, version VS, préservation de l'échantillon, préfiltration, concentration, objectif et cellule de flux) sans pour autant être comparables sur tous les points (objectif, cellule, mode de filtration etc.).

4.2 Résultats

L'acquisition des données par le FlowCAM en lui-même peut également engendrer des biais sur les résultats suite aux problèmes de fuites dues au colmatage qui peuvent avoir lieu. Ces problèmes peuvent engendrer une perte d'information (volume prélevé incomplet ou vide) sur l'échantillon passé ou des duplicatas et de biaiser la numérisation.

D'autres aléas techniques existent concernant le remplacement de matériel défectueux comme ce fut le cas pour la seringue 5mL utilisée avec l'objectif x4 et la cellule de flux 300µm. Le délai de livraison a été d'environ 2 mois. La période a pu atténuer la gravité du manque d'information en x4 étant à l'approche de l'été et que le grossissement 100 est principalement utilisé à cette époque.

La périodicité des échantillonnages (hebdomadaire) peut varier suivant les conditions météorologiques engendrant des lacunes dans les données. Cependant, les principales périodes d'intéressements du suivi phytoplanctonique (printemps/été/automne) sont météorologiquement plus clémentes entraînant des échantillonnages réguliers.

La présence en un exemplaire du FlowCAM rend son utilisation limitée et réfléchie. Lors de campagne océanographique de plusieurs jours, le choix de sa présence à bord est au détriment des suivis hebdomadaire ou bimensuels au laboratoire.

D'autre part, les résultats des analyses suite au set d'apprentissage sont à prendre avec précaution. Les résultats peuvent être biaisés par les confusions du *Deep learning*, ou réseaux de neurones. L'exemple des échantillons du 25 novembre 2021 le démontre pour l'objectif 10. On a des abondances assez élevées pour *Phaeocystis globosa*, qui correspondent à des images de fond avec des impuretés sur la cellule de flux. Cela peut s'expliquer par le fait que ces vignettes (de l'espèce ou des images de fond) ne sont peut-être pas assez représentées dans le training set.

La différence d'abondance entre les deux grossissements est une addition entre la concentration de chaque échantillon passé ou non en filtration et la reconnaissance du modèle de neurone pour chaque particule. C'est-à-dire entre le mode d'acquisition de la donnée et le mode d'analyse entre les deux grossissements il existe une différence significative qui se traduit par une variation de l'abondance du phytoplancton. Cela est dû à la différence de concentration entre les grosses particules présentes dans l'eau qui passent dans au x4 mais filtré au x10 notamment. D'autres part, le choix du diamètre minimal de 20µm pour les particules détectables discrimine certaines cellules de *Phaeocystis* libre.

4.3 Avantages et inconvénients entre les deux grossissements

L'utilisation des deux grossissements permet d'avoir une plus grande observation de l'abondance de l'ensemble des communautés phytoplanctoniques. Les groupements au x4 et x10 se complètent. Le grossissement 40 permet d'avoir une large représentativité des groupes phytoplanctoniques coloniaux (*Phaeocystis globosa*) mais également de diatomées longues (*Chaetoceros sp*, *Rhizozolenia sp.*, *Guinardia sp*). Ils ont pu être laissées jusqu'à l'espèce dans le set d'apprentissage. Le grossissement 100 permet quant à lui d'obtenir des vignettes avec des détails plus approfondis afin d'obtenir une résolution d'image suffisante pour identifier de manière précise certaines espèces. Cette différence de résolution est visible sur l'apprentissage du set par le *machine learning* qui reflète également les difficultés de l'Homme à traiter les images obtenues par le grossissement 40.

Le temps d'acquisition de la donnée sur les deux objectifs varie. L'utilisation d'une cellule de flux plus étroite et d'une pompe plus faible pour un grossissement 100 influe sur le débit et donc le passage de l'échantillon. Pour l'objectif x4, le temps pour prélever 5mL est de 5 minutes contre 20 min pour 2mL d'échantillon passé à l'objectif x10. A cela s'ajoute le temps de nettoyage, montage et démontage des objectifs/pompe/cellule de flux comprenant la calibration adéquate. L'utilisation simultanée de ces deux objectifs à des fins de comparaison et de complémentarité de la donnée reste donc limité pour l'efficacité de la méthode automatisée.

5 Conclusion

5.1 Synthèse du travail

Mon travail autour de la problématique de comment automatiser l'analyse des données du FlowCAM pour l'étude de la dynamique du phytoplancton a permis l'organisation et le bilan des données recueillies par le FlowCAM 8100 depuis 2019. J'ai pu constituer une bibliothèque d'images en couleur suivant un travail taxonomique réalisé en collaboration avec un taxonomiste et mon responsable de stage. Ce travail a permis la comparaison des deux grossissements (40 et 100) sur le passage des échantillons au LOG pour déterminer le suivi de la dynamique spatiale et temporelle du phytoplancton. Il a permis d'aboutir à des prémices de résultats sur un point de la radiale DYPHYRAD pendant 5 saisons pour percevoir les fluctuations des communautés grâce à l'application d'une méthode novatrice de classification basée sur les réseaux de neurones et une collaboration entre deux laboratoires (le LOG et le LER/BL IFRFEMER).

5.2 Perspectives

Initialement, ce stage a été motivé afin de mettre en place un protocole d'analyse automatisé à l'aide du modèle de réseaux de neurones, développé par un spécialiste de la classification du signal, Guillaume Wacquet (IR IFREMER). En effet les données en sortie du FlowCAM restaient, jusqu'alors, non exploitables directement ou de façon peu performante. Le travail élaboré lors de ce stage a permis d'avoir une base pour l'analyse des données du FlowCAM. Les résultats qui en découlent permettent d'avoir une perspective sur les démarches à poursuivre afin d'établir un protocole opérationnel post-échantillonnage dans le but d'optimiser le temps d'analyse.

Dans un premier temps, obtenir des bibliothèques d'images plus grandes et homogènes est primordial afin d'améliorer le set d'apprentissage. Cette démarche permettrait de minimiser les confusions et renforcer la fiabilité des données d'analyses. Pour ce faire, des monocultures de diverses communautés phytoplanctoniques pourraient optimiser le renforcement des bibliothèques de par la certitude des espèces présentes. Le but étant ici de collecter des vignettes d'espèces acquises sur le FlowCAM, mettre à profit les collaborations au sein du LOG et au sein du SNO PhytOBS au niveau national permettrait de créer une bibliothèque participative suivant le modèle du logiciel collaboratif ECOTAXA® ou le réseau d'imagerie avec des anciennes versions du FlowCAM par l'IFREMER et l'Université de Mons (Neaud-Masson, 2015) et le LOG (Wacquet, 2017).

A plus petite échelle de temps, la suite du stage servira à améliorer le set d'apprentissage en supprimant les catégories à « problème » comme la catégorie « Flou » et en créant une catégorie pour les impuretés. Concernant la résolution des catégories phytoplanctoniques, dissocier au mieux les catégories d'espèces qui se confondent (*Leptocylindrus danicus*, *Rhizosolenia imbricata*) en incluant uniquement des vignettes de « bonne qualité » pourraient être une solution. Cette proposition reste à moduler suivant le nombre de vignettes présentes dans les catégories.

Quant au protocole appliqué au FlowCAM, revenir sur une pré-filtration de 150µm serait souhaitable pour diminuer les biais. Cette pré-filtration sur du 150µm sera utilisée lors de blooms de *Phaeocystis*

globosa avec couplage sur un passage x4 et ainsi avoir l'ensemble de la gamme phytoplanctonique spécifique de la zone répertoriée avec les deux training set effectués

Des analyses spatiales de la campagne océanographique à bord du navire « Simon Stevin » pourront également être traitées et comparées directement avec des données d'abondances réalisées en simultané avec le Cytosense et de biomasse chlorophyllienne par groupes spectraux au moyen du Fluoroprobe, pour une présentation aux équipes du VLIZ le 30 août. Cet échange pourra donner lieu à une comparaison de leur données d'imagerie réalisées sur une version antérieure du FlowCam (FlowCam VS) *a posteriori* de la mission pour la plupart des échantillons.

Les analyses du CytoSense, possédant un imageur, pourront être menées sur les données du stage de master 1 de Manon Ryckman (abondances du phytoplancton après construction d'un set d'apprentissage) et être comparées avec les résultats du FlowCam.

Le FlowCAM est utilisé lors des blooms d'algues toxiques (HAB), concernant *Phaeocystis globosa* pour sa grande production massive de matière organique mais également d'autres espèces (dont les toxiques ou potentiellement toxiques) comme certaines espèces de diatomées (*Pseudo-nitzschia*) ou des dinoflagellés. Bauman et al., 2010 en octobre et novembre 2008, ont étudié un bloom de dinoflagellé tropical dans le golfe d'Oman, Buskey & Hyatt, 2006, quant à eux, ont utilisé le FlowCAM pour la surveillance des HAB avec l'enregistrement automatique des informations sur la taille et la fluorescence (Broutin et al., 2011). Les données récoltées en Manche Orientale pourront être analysées avec des paramètres environnementaux afin de comprendre les forçages et contrôle de la dynamique du phytoplancton et pouvoir participer au réseau de surveillance en temps réel des algues toxiques.

5.3 Bilan des compétences et connaissances

Pour conclure sur cette expérience au laboratoire d'Océanologie et de Géosciences durant ces 20 semaines, j'ai pu me former à différents protocoles que ce soit en laboratoire, sur des machines, ou sur le terrain, en bateau pour la collecte des échantillons, ainsi qu'au bureau avec la familiarisation des codes sous R. J'ai acquis une connaissance sur le phytoplancton marin notamment sur sa taxonomie.

Au sein de mon stage, j'ai pu collaborer avec des collègues de différents statuts (stagiaire, ingénieur, doctorant, post-doctorant et chercheur). L'organisation du laboratoire m'a permis de travailler en équipe, pour l'acquisition de la donnée. J'ai pu prendre part aux réflexions de l'équipe lors de réunions bimensuelles sur le travail de chacun.

Du fait des circonstances (absence en mission pendant un mois et demi puis une dizaine de jours de mon encadrant), j'ai pu développer mon sens de l'organisation et du travail en autonomie en gardant une communication à distance avec mon tuteur professionnel.

Mon travail m'a appris la patience, dans une certaine mesure, avant de pouvoir traiter des résultats (dépendant d'une collaboration extérieure au laboratoire) et m'a permis d'avoir accès à des cours en ligne à l'international sur le sujet de mon stage.

BIBLIOGRAPHIE

- Advances in Marine Pelagic Ecology : Imaging an Ocean of Change Open Course – AA-MARINET | All-Atlantic Marine Research Infrastructure Network. (s. d.). Consulté 28 juillet 2022, à l'adresse <https://www.aa-mari.net/itapina-news/advances-in-marine-pelagic-ecology-imaging-an-ocean-of-change-open-course/>
- AFNOR. (2006). NF EN 15204. Afnor EDITIONS, 41.
- Ali, N., Wacquet, G., Didry, M., Hamad, D., Artigas, L. F., & Grosjean, P. (2014). Utilisation conjointe de FlowCAM / ZooPhytoImage et de la cytométrie en flux. Premiers résultats et perspectives. Action 9. FlowCam ZooPhytoImage. Livrable n° 4. Rapport final, 23 septembre 2014. <https://archimer.ifremer.fr/doc/00363/47442/>
- Arlot, S. (2020). Validation croisée. https://www.imo.universite-paris-saclay.fr/~arlot/enseign/2020-21_M2/2-validation_croisee.pdf
- Artigas, L. F., Didry, M., Prevost, E., Bonato, S., & Lefebvre, A. (2016). Premières propositions pour un indice de composition du phytoplancton, basé sur les résultats de cytométrie en flux. 66.
- Asch, M. (2022). Traitement de données : Apprentissage Supervisé—NN.
- Bailly Du Bois, P., & Dumas, F. (2005). Fast hydrodynamic model for medium- and long-term dispersion in seawater in the English Channel and southern North Sea, qualitative and quantitative validation by radionuclide tracers. *Ocean Modelling*, 9(2), 169 – 210. <https://doi.org/10.1016/j.ocemod.2004.07.004>
- Bauman, A. G., Burt, J. A., Feary, D. A., Marquis, E., & Usseglio, P. (2010). Tropical harmful algal blooms : An emerging threat to coral reef communities? *Marine Pollution Bulletin*, 60(11), 2117-2122. <https://doi.org/10.1016/j.marpolbul.2010.08.015>
- Biodiversity data provider. (s. d.). Consulté 9 juin 2022, à l'adresse https://planktonnet.awi.de/index.php?contenttype=image_details&itemid=63412#content

- Boyce, D., Lewis, M., & Worm, B. (2010). Global phytoplankton decline over the past century. *Nature*, 466, 591–596. <https://doi.org/10.1038/nature09268>
- Brosche, P. (1982). Tidal Friction and the Earth Rotation Proceedings of a Workshop Held at the Centre for Interdisciplinary Research (ZiF) of the University of Bielefeld. <https://imusic.co/books/9783540120117/p-brosche-1982-tidal-friction-and-the-earth-s-rotation-ii-proceedings-of-a-workshop-held-at-the-centre-for-interdisciplinary-research-zif-of-the-university-of-bielefeld-september-28-october-3-1981-paperback-book>
- Broutin, M., Caffier, G., Madi, F., & Artigas, L. F. (s. d.). Synthèse bibliographique sur les techniques de suivi de l'abondance, biomasse et diversité du phytoplancton en eaux marines.
- Broutin, M., Caffier, G., Madi, F., & Artigas, L. F. (2011). Synthèse bibliographique sur les techniques d'étude de la dynamique et de la diversité du phytoplancton en eaux marines—Convention 2010—Action 1. <https://archimer.ifremer.fr/doc/00036/14730/>
- Brownlee, J. (2016, novembre 17). What is a Confusion Matrix in Machine Learning. Machine Learning Mastery. <https://machinelearningmastery.com/confusion-matrix-machine-learning/>
- Brylinski, J. M., Lagadeuc, Y., Gentilhomme, V., Dupont, J. P., Lafite, R., Dupeuble, P. A., Huault, M. F., & Auger, Y. (1991). Le fleuve cotier : Un phenomene hydrologique important en Manche orientale. Exemple du Pas-de-Calais. *Oceanologica Acta*, Special issue. <https://archimer.ifremer.fr/doc/00268/37874/>
- BRYLINSKI, J.-M., DUPONT, J., & BENTLEY, D. (1984). Conditions hydrobiologiques au large du Cap Gris-Nez (France) : Premiers résultats. Conditions hydrobiologiques au large du Cap Gris-Nez (France): premiers résultats, 7(3), 315–322.
- Buskey, E. J., & Hyatt, C. J. (2006). Use of the FlowCAM for semi-automated recognition and enumeration of red tide cells (*Karenia brevis*) in natural plankton samples. *Harmful Algae*, 5(6), 685–692. <https://doi.org/10.1016/j.hal.2006.02.003>
- Carpentier, A., Coppin, F., Dauvin, J.-C., Desroy, N., Dewarumez, J.-M., Eastwood, P. D., Ernande, B., Harrop, S., Kemp, Z., Koubbi, P., Williams, N. L., Lefebvre, A., Lemoine, M., Loots, C., S. Martin,

- C., Meaden, G. J., Ryan, N., Tan, L., Vaz, S., & Walkey, M. (2005). CHARM : Atlas des Habitats des Ressources Marines de la Manche Orientale.
<https://archimer.ifremer.fr/doc/00000/2350/>
- Chisholm, S. W. (1992). Phytoplankton Size. In P. G. Falkowski, A. D. Woodhead, & K. Vivirito (Éds.), *Primary Productivity and Biogeochemical Cycles in the Sea* (p. 213 - 237). Springer US.
https://doi.org/10.1007/978-1-4899-0762-2_12
- Chiswell, S., Calil, P., & Boyd, P. (2015). Spring blooms and annual cycles of phytoplankton : A unified perspective. *Journal of Plankton Research*, 37. <https://doi.org/10.1093/plankt/fbv021>
- Choi, J. W., & Stoecker, D. K. (1989). Effects of Fixation on Cell Volume of Marine Planktonic Protozoa. *Applied and Environmental Microbiology*, 55(7), 1761 - 1765.
<https://doi.org/10.1128/aem.55.7.1761-1765.1989>
- Choudhury, A., & Pal, R. (2014). An Introduction to Phytoplanktons : Diversity and Ecology.
https://doi.org/10.1007/978-81-322-1838-8_2
- Culverhouse, P. F. (2007). Human and machine factors in algae monitoring performance. *Ecological Informatics*, 2(4), 361-366. <https://doi.org/10.1016/j.ecoinf.2007.07.001>
- Culverhouse, P. F., Williams, R., Reguera, B., Herry, V., & González-Gil, S. (2003). Do experts make mistakes? A comparison of human and machine identification of dinoflagellates. *Marine Ecology Progress Series*, 247, 17-25. <https://doi.org/10.3354/meps247017>
- datascientest, T. (2021, février 16). Comment lire et exploiter une matrice de confusion ? Formation Data Science | DataScientest.com. <https://datascientest.com/matrice-de-confusion>
- Derot, J. (2014). Utilisation des données de MAREL Carnot pour la compréhension des mécanismes des extrêmes dans la qualité des eaux à Boulogne-sur-Mer [These de doctorat, Littoral].
<https://www.theses.fr/2014DUNK0375>
- Devreker, D., & Lefebvre, A. (2022). Résultats de la mise en oeuvre des réseaux REPHY et SRN. Zones côtières de la Manche orientale et de la baie sud de la Mer du Nord. Bilan de l'année 2021.
<https://archimer.ifremer.fr/doc/00506/61790/>

- Direction Inter-régionale de la Mer - Manche Est - Mer du Nord. (2019, septembre). Stratégie de Façade Maritime adoptée. <https://www.dirm.memn.developpement-durable.gouv.fr/strategie-de-facade-maritime-adoptee-r436.html>
- Ellien, C., Thiebaut, E., Dumas, F., Salomon, J., & Nival, P. (2004). A modelling study of the respective role of hydrodynamic processes and larval mortality on larval dispersal and recruitment of benthic invertebrates : Example of *Pectinaria koreni* (Annelida : Polychaeta) in the Bay of Seine (English Channel). *Journal of Plankton Research*, 26. <https://doi.org/10.1093/plankt/fbh018>
- EOS - Phytoplankton Encyclopedia Project. (s. d.). Consulté 9 juin 2022, à l'adresse https://www.eoas.ubc.ca/research/phytoplankton/at_a_glance/chain_forming_diatoms.htm
- Falkowski, P., Laws, E., Barber, R., & Murray, J. (2003). Phytoplankton and Their Role in Primary, New, and Export Production. https://doi.org/10.1007/978-3-642-55844-3_5
- Field, C. B., Behrenfeld, M. J., Randerson, J. T., & Falkowski, P. (1998). Primary Production of the Biosphere : Integrating Terrestrial and Oceanic Components. *Science*, 281(5374), 237 -240. <https://doi.org/10.1126/science.281.5374.237>
- Grosjean, P., & Wacquet, G. (2016). Optimisation de l'identification et du dénombrement du microphytoplancton avec le système couplé de numérisation et d'analyse d'images FlowCAM – Zoo/PhytoImage (système innovant). <https://archimer.ifremer.fr/doc/00389/49990/>
- Guingand, A. (2012). Protection de l'environnement littoral et marin. Sous-région marine Manche-Mer du Nord. Evaluation initiale DCSMM.
- Guiselin, N. (2010). Etude de la dynamique des communautés phytoplanctoniques par microscopie et cytométrie en flux, en eaux côtière de la Manche orientale [These de doctorat, Littoral]. <https://www.theses.fr/2010DUNK0258>
- Hess, P. (2020). Harmful Algae 2018 – From Ecosystems to Socioecosystems. Proceedings of the 18th International Conference on Harmful Algae, 21-26 October 2018, Nantes, France. <https://archimer.ifremer.fr/doc/00651/76285/>

- IFREMER. (2014). Protocole d'échantillonnage et analyse en laboratoire. 3.
- IFREMER. (2017). Suivi Régional des Nutriments. L'Ifremer en Manche Mer du Nord.
<https://wwz.ifremer.fr/manchemerdunord/Environnement/LER-Boulogne-sur-Mer/Surveillance-et-Observation/Suivi-Regional-des-Nutriments>
- IFREMER. (2019). Sépia II. Flotte océanographique française opérée par l'Ifremer.
<https://www.flotteoceanographique.fr/Nos-moyens/Navires-engins-et-equipements-mobiles/Navires-de-station/Sepia-II>
- Ifremer REPHY. (2018). Environnement et Ressources des Pertuis Charentais.
<https://wwz.ifremer.fr/lerpc/Activites-et-Missions/Surveillance/REPHY>
- IHO. (2011). Manuel d'Hydrographie, 1ère édition.
- INRAE. (2009). Fiche de comptage du phytoplancton (méthode Utermöhl). Cemagref.
- Jeffrey, S. W., Mantoura, R. F. C., Wright, S. W., International Council of Scientific Unions, & Unesco (Éds.). (1997). Phytoplankton pigments in oceanography : Guidelines to modern methods. UNESCO Publishing.
- JERICO-S3 Project Overview | JERICO Research Infrastructure. (2021, avril 21). <https://www.jerico-ri.eu/projects/jerico-s3/>
- Kassel, G. (2022). Focus sur les réseaux neuronaux.
- Kraberger, A., Baumann, M., & Dürselen, Claus.-D. (2010). Coastal Phytoplankton : Photo Guide for Northern European Seas. In Verlag Dr. Friedrich Pfeil.
- Laboratoires. (s. d.). Consulté 26 juillet 2022, à l'adresse <https://www.univ-littoral.fr/laboratoires/>
- Lamy, D. (2006). Dynamique bactérienne en Manche orientale—Relations avec les poussées de *Phaeocystis globosa*. [Theses, Université du Littoral Côte d'Opale]. <https://tel.archives-ouvertes.fr/tel-00105902>
- Larink, O., & Westheide, W. (2011). Coastal Plankton Photo Guide for European Seas (2nd edition).
- Lazure, P., & Desmare, S. (2012). ÉTAT PHYSIQUE ET CHIMIQUE Caractéristiques physiques. 9.

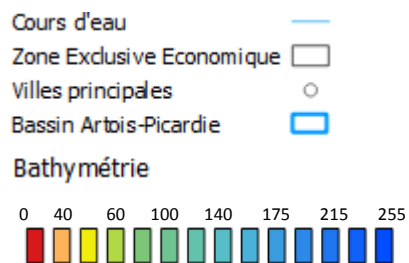
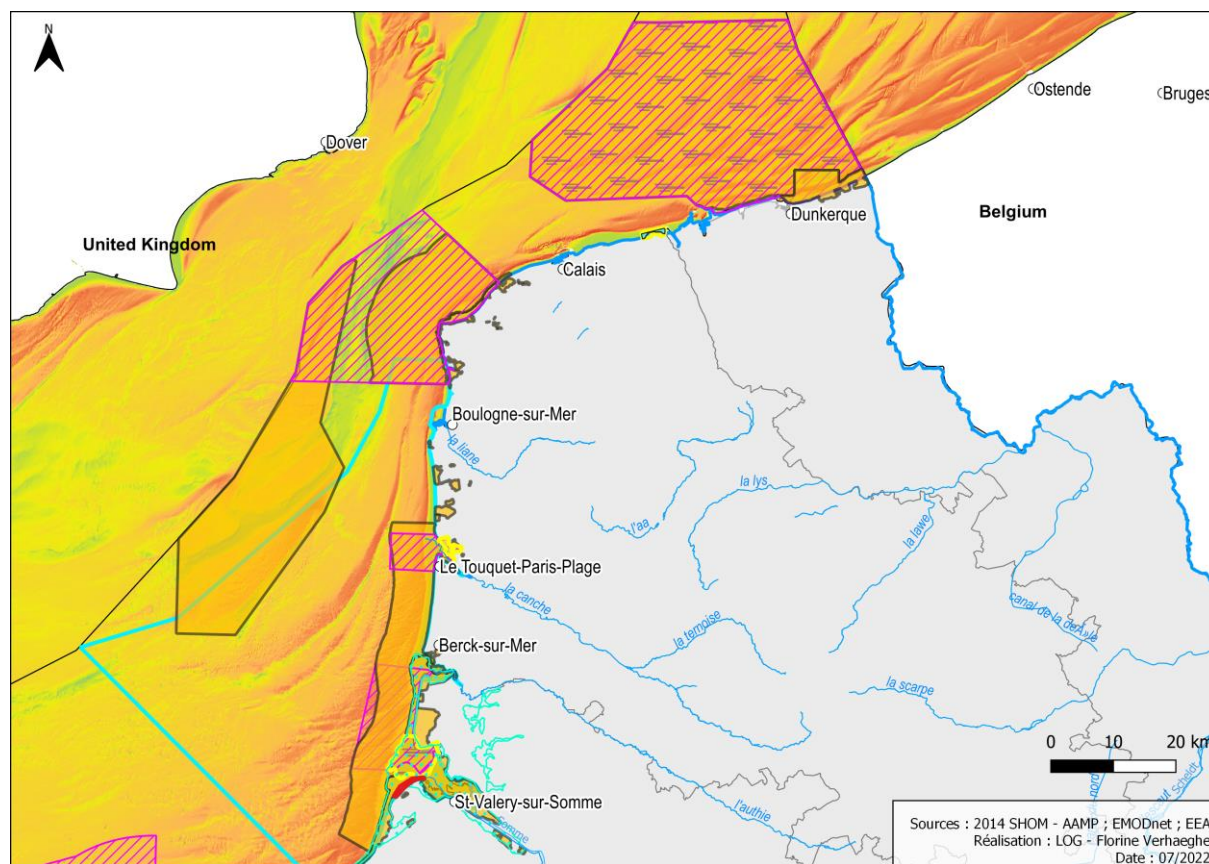
- Lefebvre, A., & Delpech, J.-P. (2004). Le bloom de *Phaeocystis* en Manche orientale—Nuisances socio-économiques et / ou écologiques ? Ifremer.
- Lévêque, C. (2006). Les poissons des eaux continentales africaines : Diversité, écologie, utilisation par l'homme. IRD Editions.
- Levoy, F., Anthony, E. J., Monfort, O., & Larsonneur, C. (2000). The morphodynamics of megatidal beaches in Normandy, France. *Marine Geology*, 171(1), 39 - 59.
[https://doi.org/10.1016/S0025-3227\(00\)00110-9](https://doi.org/10.1016/S0025-3227(00)00110-9)
- Lévy, M., & Bopp, L. (2007). Turbulences dans l'océane. 3.
- Lindeman, R. L. (1942). The Trophic-Dynamic Aspect of Ecology. *Ecology*, 23(4), 399 - 417.
<https://doi.org/10.2307/1930126>
- LOG. (2022). Présentation équipe ECOP2. <https://log.cnrs.fr/presentation-equipe-ecop2>
- MAREL Carnot. (s. d.). L'Ifremer en Manche Mer du Nord. Consulté 7 juillet 2022, à l'adresse <https://www.ifremer.fr/manchemerdunord/Environnement/LER-Boulogne-sur-Mer/Surveillance-et-Observation/MAREL-Carnot>
- Masmoudi, S. (2014). Dynamique du phytoplancton et caractérisation physiologique et moléculaire de trois espèces autotrophes de la saline de Sfax(Tunisie), un milieu extrémophile. 269.
- Menden-Deuer, S., Lessard, E. J., & Satterberg, J. (2001). Effect of preservation on dinoflagellate and diatom cell volume and consequences for carbon biomass predictions. *Marine Ecology Progress Series*, 222, 41-50. <https://doi.org/10.3354/meps222041>
- Milieu Marin France. (2022, février 9). Nature et intégrité des fonds marins. Milieu Marin France.
<https://www.milieumarinfrance.fr/Nos-rubriques/Etat-du-milieu/Nature-et-integrite-des-fonds-marins>
- Ministère de l'Écologie, du Développement durable, & des Transports et du Logement. (2012). Stratégie national pour la création et la gestion des aires marines protégées.
- Ministère de l'Écologie, du Développement durable et de l'Énergie (Éd.). (2015). Stratégie nationale de création et de gestion des aires marines protégées. 24.

- Neaud-Masson, N. (2015). Catalogue illustré des images d'un set d'apprentissage réalisé à l'aide de l'outil FlowCAM – ZooPhytoImage utilisé dans le cadre du REPHY. Images issues de numérisations de phytoplancton au grossissement 4X, d'échantillons d'eau, pour le set d'apprentissage global Manche-Atlantique. <https://archimer.ifremer.fr/doc/00340/45157/>
- Nishizawa, S., Fukuda, M., & Inoue, N. (1954). PHOTOGRAPHIC STUDY OF SUSPENDED MATTER AND PLANKTON IN THE SEA. 6.
- Nord, D. I. de la M.-M. E.-M. du. (2018, avril 11). Document synthétique SFM. <https://www.dirm.memn.developpement-durable.gouv.fr/document-synthetique-sfm-a953.html>
- Not, F. (2011). Le plancton—Encyclopédie illustrée (Y. Nakamura, Trad.). Société Japonaise de Plancton.
- OASU. (2018, décembre 14). Wimereux | Somlit. <https://www.somlit.fr/wimereux/>
- OFB. (2018, octobre 17). Éléments biologiques—Plans d'eau—Méthodes de surveillance état écologique | The OFB Technical Portal. <https://professionnels.ofb.fr/en/node/398>
- Owen, B. M., Hallett, C. S., Cosgrove, J. J., Tweedley, J. R., & Moheimani, N. R. (2022). Reporting of methods for automated devices : A systematic review and recommendation for studies using FlowCam for phytoplankton. *Limnology and Oceanography: Methods*, 20(7), 400 - 427. <https://doi.org/10.1002/lom3.10496>
- PERRODON, A., & PINOT, J.-P. (s. d.). NORD MER DU. Encyclopædia Universalis. Consulté 15 juin 2022, à l'adresse <http://www.universalis-edu.com/encyclopedie/mer-du-nord/>
- Phytobs. (2018). PHYTOBS. <https://www.phytobs.fr/Phytobs>
- PINOT, J.-P. (s. d.). MANCHE. Encyclopædia Universalis. Consulté 15 juin 2022, à l'adresse <http://www.universalis-edu.com/encyclopedie/manche/>
- Quisthoudt, C., Bentley, D., & Brylinski, J. M. (1987). Discontinuité hydrobiologique dans le détroit du Pas-de-Calais. *Journal of Plankton Research*, 9(5), 995 - 1002. <https://doi.org/10.1093/plankt/9.5.995>

- Rocha, C., & Passow, U. (2014). The Biological Pump. Treatise on Geochemistry. <https://doi.org/10.1016/B978-0-08-095975-7.00604-5>
- Rutten, T. P. A., Sandee, B., & Hofman, A. R. T. (2005). Phytoplankton monitoring by high performance flow cytometry : A successful approach? *Cytometry Part A*, 64A(1), 16 - 26. <https://doi.org/10.1002/cyto.a.20106>
- Salomon, J.-C., & Breton, M. (1993). An atlas of long-term currents in the channel. *Oceanologica Acta*, 16(5-6), 439-448.
- Stépanian, A. (2003). Evolution morphodynamique d'une plage macrotidale à barres : Omaha beach (Normandie). 284.
- Utermöhl, H. (1958). Methods of collecting plankton for various purposes are discussed. *SIL Communications*, 1953-1996, 9(1), 1-38. <https://doi.org/10.1080/05384680.1958.11904091>
- Vandermeirsch, F. (2012). ÉTAT PHYSIQUE ET CHIMIQUE Caractéristiques physiques. 8.
- VLIZ. (2012). Technical specifications. <https://www.vliz.be/en/technical-specifications>
- Wacquet, G. (2017). Amélioration de l'étude et du suivi des espèces phytoplanctoniques dominantes dans la Manche orientale, en focalisant à la fois sur les petites cellules et les espèces coloniales. 48.
- Walles, T. J. W. (2020). Mesozooplankton Studies Through Automation and Machine Learning. Technische Universität Berlin.
- Winder, M., & Sommer, U. (2012). Phytoplankton response to a changing climate. *Hydrobiologia*, 698(1), 5-16. <https://doi.org/10.1007/s10750-012-1149-2>

ANNEXES

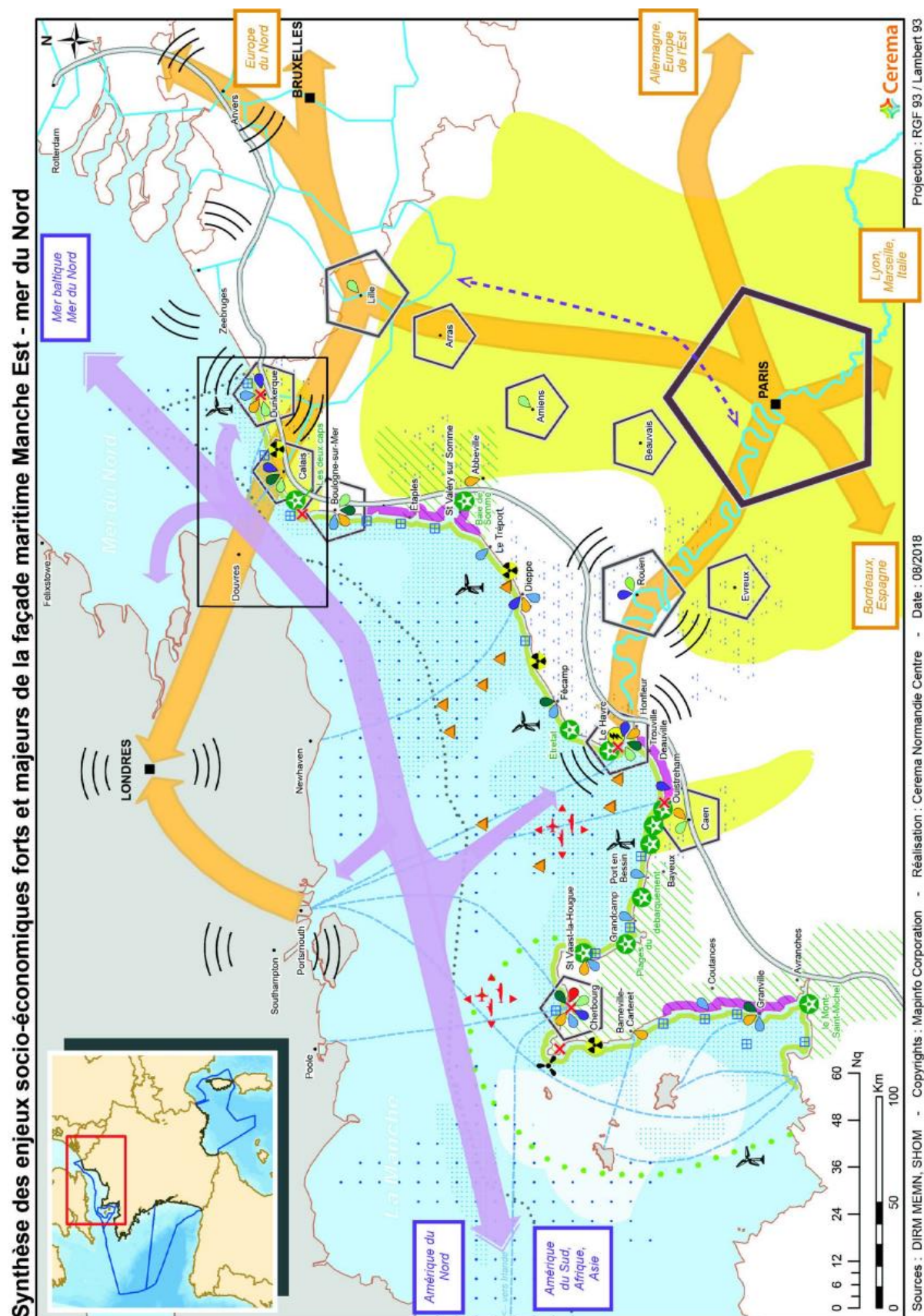
ANNEXE 1 : CARTE DES AIRES MARINES PROTEGEES DE LA SOUS-REGION MARINE MANCHE ORIENTALE-MER DU NORD



Aires Marines Protégées

- Parc naturel marin
- Domaine public maritime (Conservatoire du littoral)
- Réserve naturelle nationale
- Zone de protection spéciale (N2000, DO)
- Zone humide d'importance internationale (Ramsar)
- Zone spéciale de conservation (N2000, DHFF)
- Zone marine protégée de la convention OSPAR

ANNEXE 2 : SYNTHÈSE DES ENJEUX SOCIO-ECONOMIQUES FORTS ET MAJEURS DE LA FAÇADE MARITIME MANCHE EST – MER DU NORD

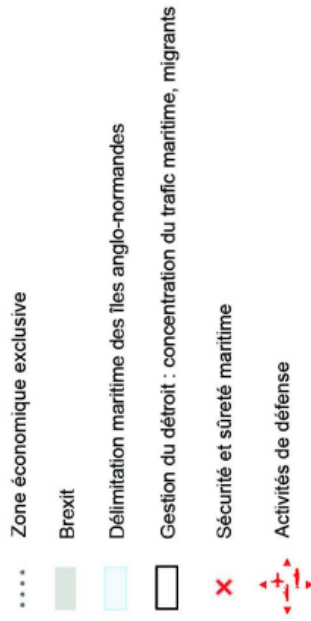


L'espace transmanche : une façade maritime reliée à l'économie mondiale

Enjeux économiques internationaux et transfrontaliers

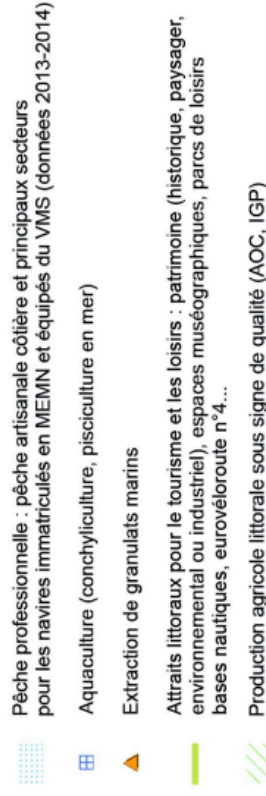


Gouvernance, défense et sécurité maritime



Les spécificités économiques de l'interface terre-mer en Manche Est - mer du Nord

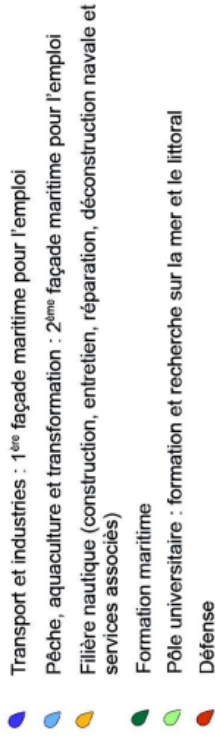
Gestion durable des ressources marines et littorales



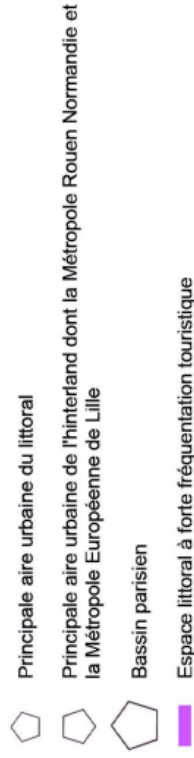
Énergies marines renouvelables - Énergies terrestres non renouvelables (Site projet ou à l'étude)



Les filières maritimes de la façade maritime



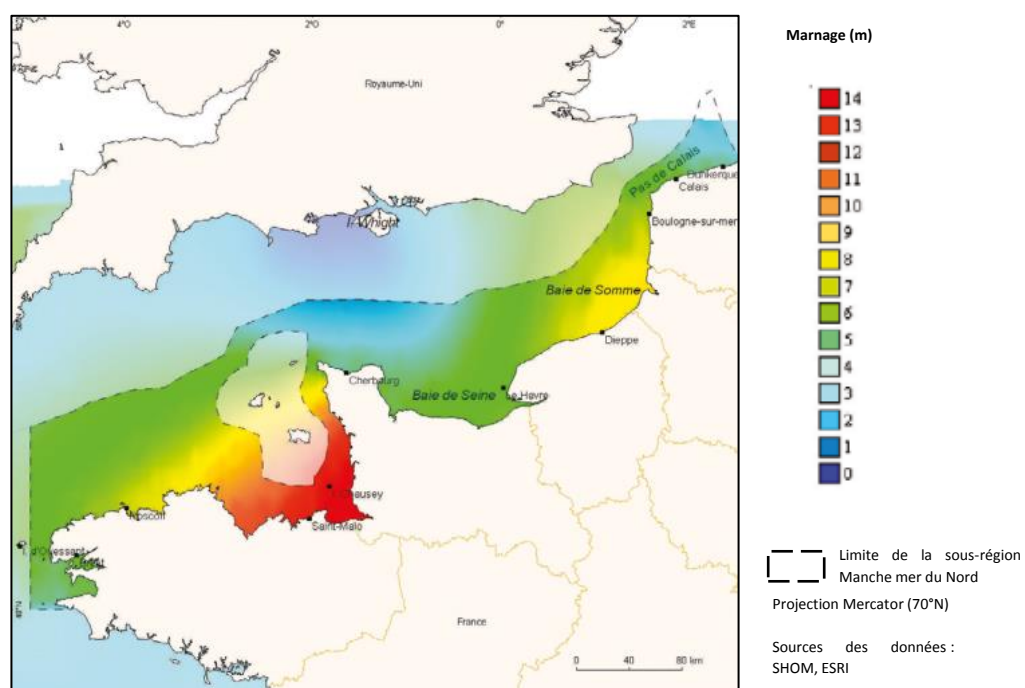
Les marchés de la façade maritime



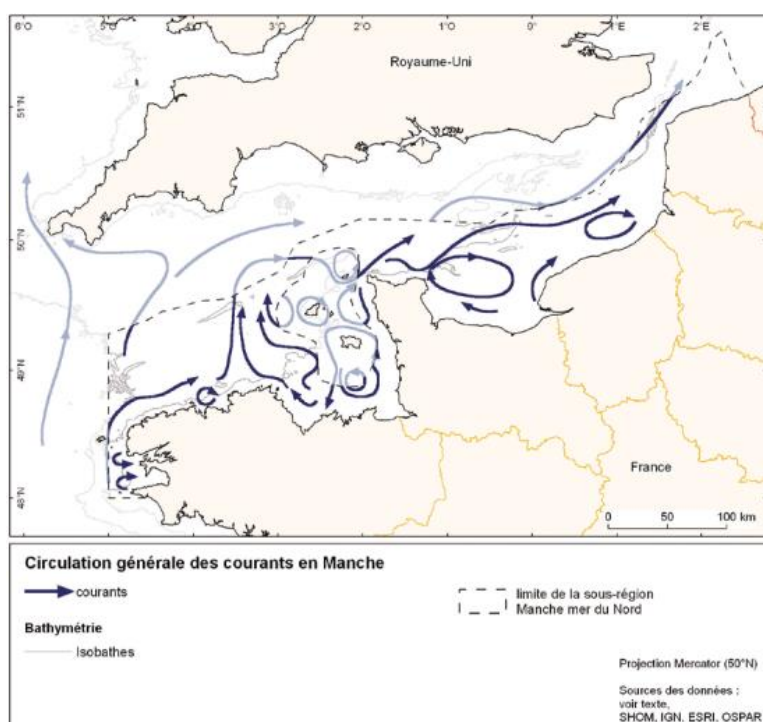
Les axes majeurs de la façade maritime



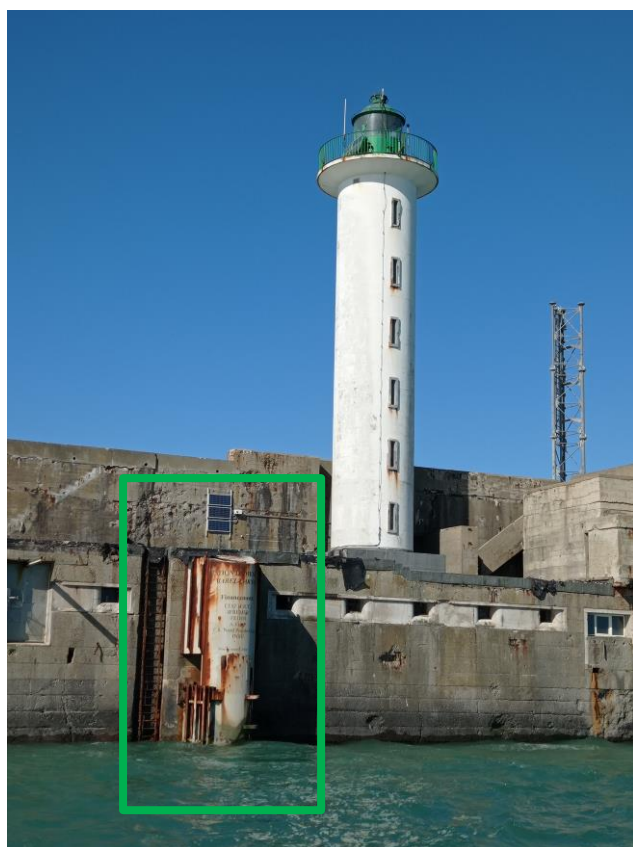
ANNEXE 3: MARNAGE EN MANCHE (SOURCES : MODELE DE MAREE CSTFRANCE DU SHOM, ESRI, 2010).



ANNEXE 4: CIRCULATION MOYENNE EN MANCHE (SOURCES : IGN, ESRI, SHOM, OSPAR, 2010).



ANNEXE 5: STATION MAREL, DIGUE CARNOT A BOULOGNE-SUR-MER



ANNEXE 6: TABLEAU DES ANALYSES EFFECTUEES PAR POINT POUR LE SUIVI DYPHYRAD

	FluoroProbe (groupes spectraux)	Cytomètre (groupe fonctionnel)	FlowCam (caractérisation taxonomique)	Fluorimètre Chl a in vivo	Concentration Chl a in vitro	HPLC (analyse pigmentaire)
R4	x	x	x	x	xxx	x
R4 (fond)	x	x		x	x	
R3'	x	x		x	x	
R3	x	x	x	x	xxx	x
R3 (fond)	x	x		x	x	
R2'	x	x		x	x	
R2	x	x	x	x	xxx	x
R2 (fond)	x	x		x	x	

R1'	x	x		x	x	
R1	x	x	x	x	xxx	x
R1 (fond)	x	x		x	x	
R0'	x	x		x	x	
R0	x	x	x	x	xxx	x
R0 (fond)	x	x		x	x	

ANNEXE 7: PROTOCOLE FLOWCAM (SOURCE : ELISE CAILLARD)

1. Vérification des connexions internes et flacons

- Ouvrir les capots supérieurs, puis inférieurs, et vérifier les connexions des tuyaux :
- Pour les flacons de rinçage, de nettoyage et de déchets
- Pour l'évacuation au-dessus de la cellule de flux



Le FlowCAM contient 3 flacons : un pour le nettoyage, un pour le rinçage et un pour le recueil des déchets. Le flacon de nettoyage doit être rempli avec une solution de nettoyage adaptée (« CitraJet ») pour supprimer les particules encore présentes dans le système. Le flacon de rinçage doit être rempli avec de l'eau déionisée ou distillée.

- Remplir les flacons de nettoyage et de rinçage
- Vider le flacon de recueil des déchets
- Vérifier que les tuyaux soient bien raccordés.

2. Mise sous tension du FlowCAM

- ❖ **Mettre sous tension le FlowCAM** à en appuyant sur le bouton à l'arrière de l'appareil.
- ❖ **Démarrer le FlowCAM** en appuyant sur le bouton situé sur le dessus (à droite), puis allumer les ordinateurs et **lancer le programme VisualSpreadSheet** et **choisir l'objectif** utilisé (4X, 10X ou 20X)



3. Assemblage Objectif / Cellule de Flux / Pompe

- L'Objectif 4X doit être utilisé avec une cellule de flux calibrée de 300 ou 600µm (lors de bloom), et avec la pompe seringue de 5 ml.
- L'objectif 10X doit être utilisé avec une cellule de flux calibrée de 100µm, et avec la pompe seringue de 1 ml.
- L'objectif 20X doit être utilisé avec une cellule de flux calibrée de 50µm, et avec la pompe seringue de 0,5 ml.

3.1 Mise en place du système optique

Dans VisualSpreadSheet, sélectionner **Setup > Change Objective**. Le bloc devant accueillir l'objectif se déplace alors vers la gauche. Une fois cette opération terminée, plusieurs actions doivent être réalisées :

1. Dévisser l'objectif en place avec précaution.
2. Visser le nouvel objectif (ne pas trop serrer)
3. Cliquer sur **OK** dans la boîte de dialogue.
4. Sélectionner la magnification et cliquer de nouveau sur **OK**.

L'objectif se visse dans le **sens des aiguilles** d'une montre et se **dévisse** dans le **sens antihoraire**. L'objectif est vissé jusqu'à la butée.

Le système optique du FlowCAM doit être le plus propre possible. **Ne jamais toucher directement la cellule de flux et l'objectif** afin d'éviter toutes traces de doigts.



3.2 Mise en place de la pompe seringue

Pour changer la pompe seringue :



1. Sélectionner **Setup > Pump** dans VisualSpreadSheet
2. Cliquer sur le bouton **Change**
3. Cliquer sur **Yes** à la question de changement de pompe, et suivre les instructions à l'écran.
4. Tourner la vis de blocage du piston 3 fois dans le sens antihoraire. Cliquer sur **OK**.
5. Dévisser la pompe de la valve et retirer la.
6. Visser la nouvelle pompe à la valve, et centrer le piston dans la seringue. Cliquer sur **OK**.
7. Serrer fermement la vis de blocage du piston. Cliquer sur **OK**.
8. Sélectionner la taille de la nouvelle pompe dans la fenêtre qui s'affiche et cliquer sur **OK**.
9. Nettoyer l'ancienne pompe (retirer complètement le piston) et ranger là.

3.3 Mise en place de la cellule de flux

1. Connecter le tube supérieur (le plus long, avec les connecteurs marrons) au-dessus de la cellule de flux (une tâche de couleur indique le haut de la cellule de flux). Serrer fortement jusqu'à entendre un « clic ».
2. Connecter le tube inférieur (le plus court, avec les connecteurs noirs) en dessous de la cellule de flux. Serrer fortement jusqu'à entendre un « clic ».



3. Desserrer les deux vis à molette rouges du dispositif de retenue de la cellule de flux et retirer le.
4. Placer la tâche de couleur vers vous, et connecter le tuyau supérieur au port d'entrée de l'échantillon.
5. Connecter le tuyau inférieur au port de la pompe seringue
6. Placer la cellule de flux dans son bloc et insérer les tuyaux supérieur et inférieur dans leur fente respective.
7. Placer le dispositif de retenue de la cellule de flux dans la fente de sorte que celle-ci soit maintenue en place et serrer les deux vis à molette.



4 Réglage du focus

Pour le réglage automatique du focus, utiliser les billes suivantes :

- Billes pour faible grossissement pour les cellules de flux 300µm ou 600µm (environ 50µm de diamètre)
- Billes pour fort grossissement pour la cellule de flux 100µm (environ 15µm de diamètre)
- Billes pour la cellule de flux 50µm (environ 10µm de diamètre).

Pour la mise au point automatique :

1. Dans la barre d'outils, cliquer sur l'outil **Setup and Focus**
2. Si nécessaire, aligner la cellule de flux sur la position correcte à l'aide de la vis de positionnement située à gauche de la cellule de flux. La cellule de flux peut être déplacée à gauche et à droite (remarque : les lignes rouges doivent être positionnées à l'intérieur des lignes grises, qui représentent les parois de la cellule de flux).



3. Cliquer sur le bouton **Autofocus**. L'assistant de mise au point automatique s'ouvre. Cliquer sur **Next** et suivre les instructions à l'écran. Cliquer sur Best Focus Image.
4. Placer un embout de pipette 5000 dans l'orifice d'introduction de l'échantillon situé en haut du FlowCAM et ajouter au moins 500µl de billes. Cliquer sur **Next**. La pompe aspire l'échantillon dans la cellule de flux et le processus de mise au point automatique débute.
5. Lorsque le processus de mise au point automatique est terminé, un message apparaît pour indiquer que l'opération a abouti. Cliquer sur **Start Flush** pour vider la solution du système. Si échec de l'autofocus, remettre 400µl de billes et recommencez la manipulation.
6. Fermer la fenêtre **Setup** et **Focus Mode**.



5. Réglages du contexte d'acquisition

Pour régler le contexte d'acquisition du FlowCAM :

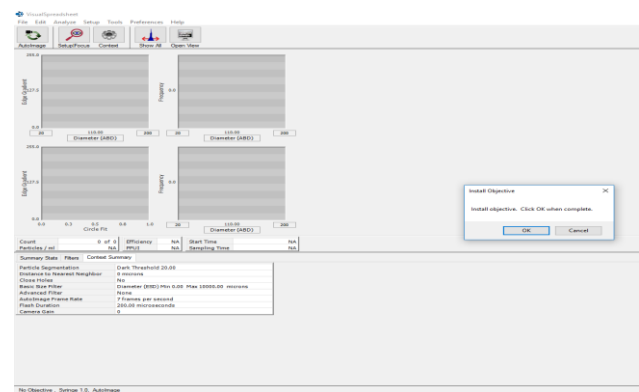
1. Dans la barre d'outils, cliquer sur **Context**
2. La boîte de dialogue **Context** s'ouvre
3. Cliquer sur l'onglet **Load**, et cliquer sur **Load a context file**.
4. Sélectionner le fichier de contexte en fonction de l'objectif (dans le répertoire **Documents/VisualSpreadSheet/context**) :

DYPHYRAD_context_4X.2019-05-02.ctx

DYPHYRAD_context_10X.2019-05-02.ctx

6. Numérisation des échantillons

- Remplir l'entonnoir avec l'échantillon
- Ouvrir le répertoire **Desktop\DYPHYRAD\Year**
- Créer un répertoire portant la **date du jour** (ex : 2020-09-01). Dans ce répertoire seront sauvegardés les fichiers issus de la numérisation.
- Lancer la numérisation en cliquant sur le bouton **Autoimage** dans la barre d'outils.
- Remplir les champs demandés dans la section **Notes**. Pour le nom de l'échantillon, utiliser le format suivant :
 - ☐ Pour une numérisation au 4X avec une cellule de flux de 300µm : XX.yyyy-mm-dd.300A4X.rr
 - ☐ Pour une numérisation au 10X avec une cellule de flux de 100µm : XX.yyyy-mm-dd.100A10X.rr
- Cliquer sur **OK**
- Une fenêtre de sauvegarde s'ouvre. Nommer le fichier de numérisation en respectant le format ci-dessus.
- L'enregistrer dans le répertoire créé en cliquant sur **Save**
- Laisser les fenêtres ouvertes jusqu'à la fin de l'acquisition. Sinon, mettre fin à l'acquisition au moment choisi en cliquant sur Pause dans la fenêtre de la caméra, puis fermer cette fenêtre.
- Avant de numériser un autre échantillon, attendre que le système se nettoie et se rince automatiquement



10X :

- « Autres »
 - Cryptophyceae
 - *Diplopsalis*
 - *Gyrodinium*
 - *K.glaucum*
 - *Phaeocystis globosa*
 - *Prorocentrum* sp
 - *Torodinium*
 - *Tripos*
- « Bulles »
- « Cilies »
 - *Brachionus* sp
 - *Cilies* spp
 - *Tiarina*
- « Diatomees »
 - *Actinoptycus* sp
 - *Cerataulina pelagica*
 - *Chaetoceros* spp
 - *Coscinodiscus* sp
 - *Ditylum brightwellii*
 - *Eucampia* sp
 - *Guinardia*
 - *Helicotheca tamesis*
 - *Lauderia annulata*
 - *Leptocylindrus danicus*
 - *Meuniera membranacea*
 - *Navicula* spp
 - *Odontella*
 - *Paralia sulcata*
 - *Pleurosigma* sp
 - *Pseudo-nitzschia* spp
 - *Rhaphoneis ampiceros*
 - *Rhizosolenia*
 - *Thalassionema nitzschioides*
 - *Thalassiosira* spp
- « Zoo »
 - *Mytilus edulis*
 - Nauplius
 - Oeufs

4X :

- « Autres »
 - *Diplopsalis*
 - *Gyrodinium*
 - *Phaeocystis globosa*
 - *Prorocentrum* sp
 - *Tripos*
- « Bulle »
- « Cilies »
 - Cilies
 - *Laboea* sp
 - *Tiarina* sp
- « Diatomees »
 - *Actinoptycus* sp
 - *Brockmanniella*
 - *Cerataulina pelagica*
 - *Chaetoceros* spp
 - *Coscinodiscus* sp
 - *Ditylum brightwellii*
 - *Eucampia zodiacus*
 - *Guinardia*
 - *Lauderia annulata*
 - *Leptocylindrus danicus*
 - *Meuniera membranacea*
 - *Navicula* spp
 - *Odontella*
 - *Paralia sulcata*
 - *Pleurosigma* sp
 - *Pseudo-nitzschia* spp
 - *Rhaphoneis ampiceros*
 - *Rhizosolenia*
 - *Thalassionema nitzschioides*
 - *Thalassiosira* spp
- « Pollen »
- « Zoo »
 - Apendicularia
 - Copepode
 - Larve
 - *Mytilus edulis*
 - Nauplius
 - Oeufs

ANNEXE 9: CONTEXTE DU FLOWCAM POUR LE X4

[Software]

SoftwareName=VisualSpreadsheet
SoftwareVersion=4.15.1
SoftwareBetaFlag=0
CameraMagnification=4
GumStixFirmwareVersion=23
SerialNo=10439

[General]

TechnicianName=
LocationName=
LocationLatitude=
LocationLongitude=
RunStartTime=2019-05-02 15:56:51
RunEndTime=2019-05-02 16:00:32
TimeZone=

[CameraBehavior]

AutoImageRate=12
FlashDuration=600.00
FlashAmplitude=24
FlashDelay=500.00
RecoveryTime=1.00
LaserCycleTime=1250.00
CameraTriggerDuration=100.00
PVDBRBC=1000.00
PVDBRBE=1000.00

[Vertex]

ListDir=C:\VisualSpreadsheet

[Fluorescence]

Ch1Measure=1
Ch1Trigger=1
Ch1Threshold=15000
Ch1HighThreshold=7000000
Ch2Measure=1
Ch2Trigger=1
Ch2Threshold=15000
Ch2HighThreshold=7000000
QDAC_PMTSensitivity1=1435
QDAC_PMTSensitivity2=1435
QDAC_LaserPower=3072
QDAC_FlashAmplitude=24
PMTSampleWidth=0
MinTriggerPulseWidth=48
ScatterDetector=0
UseLogScale=0
SkipMultiParticleImages=0
PMTGain=0
SmallParticleFilter=0

[Laser]

Enabled=1

[Camera]

AutoTriggerFlag=1
CameraName=SenTech 1.3 Mp
FrameCount=10

Initialization=
Shutter=18
Gain=0
WhiteBalanceU=2150
WhiteBalanceV=2150
Brightness=0
Gamma=0
Hue=2290
Saturation=300
Sharpness=0
RedGain=0
GreenGain=0
BlueGain=0
ColorGain=0
AutoAdjustIntensity=0
AutoIntensityTarget=140
AutoIntensityShutterMin=6
AutoIntensityShutterMax=16
AutoIntensityGainMin=0
AutoIntensityGainMax=680
AutoIntensityBrightnessMin=0
AutoIntensityBrightnessMax=2047
AutoIntensityFlashDurationMin=1.00
AutoIntensityFlashDurationMax=100.00
AutoAdjustWhiteBalance=0
ImageWidth=1200
ImageHeight=1920

[RunTermination]

MaxParticles=1000000
StopOnMaxParticles=0
MaxRunTimeMinutes=60
MaxRunTimeSeconds=0
StopOnMaxRunTime=0
MaxImageFiles=100
StopOnMaxImageFiles=0
MaxCameraImages=1000
StopOnMaxCameraImages=0
MaxFluidVolume=5.0000
StopOnFluidVolume=1
StopOnUserOnly=0

[CountCalibration]

CalibrationFactorPPML=1.2000
bStandardVolumeCorrection=1

[Files]

RecalibrationIntervalMinutes=30

[CaptureReport]

AutoExportList=1
AutoExportListSummary=1
AutoPrintSummary=0
AutoClassify=0
AutoExportClassificationSummary=0

ClassificationTemplate=

[Fluid]

FlowCellType=FC300FV
FlowCellDepth=300
FlowCellWidth=1500
UseManualVolume=0
ManualVolumeML=20.0000
FluidVolumeImaged_ml=0.0000
UseDeadVolumeCalculation=0
DeadVolumeTubeLength_cm=0.00
DeadVolumeTubeDiameter_cm=0.16
UseDilutionCalculation=0
SampleDilution=1
PumpFlowRate=1.8000
PrimingMethod=0
PumpToFlowCellTubeLength_cm=0.00
ValveToCleanerTubeLength_cm=0.00
PumpErrorBudgetVolume=0.0100
PumpErrorBudgetMultiplier=1.6000
PumpFlowCellLength_cm=3.5000
PumpFlowCellFraction=0.5000
CalibrationConstant=1.8005
FringeSize=4.5000

[CaptureRegion]

AcceptableLeft=190
AcceptableRight=1010
AcceptableTop=1
AcceptableBottom=1919
ImagePixelBorder=2
SaveImageFiles=1
SaveBinaryImageFiles=0
SaveRawFiles=0
MaxNumRawFiles=1000
FlatFieldCount=10

[CaptureParameters]

DistanceToNeighbor=20
ThresholdDark=20.0000
ThresholdLight=20.0000
CloseHoles=0
CaptureDarkOrLightPixels=0
MinESD=20.0000
MaxESD=10000.0000
SimpleCaptureFilterDiameter=1
UseAdvancedFilter=0
FilterName=C:\Program
Files\VisualSpreadsheet\FactoryDefault
s\Filter - Aspect Ratio 0.9-1.0.flr

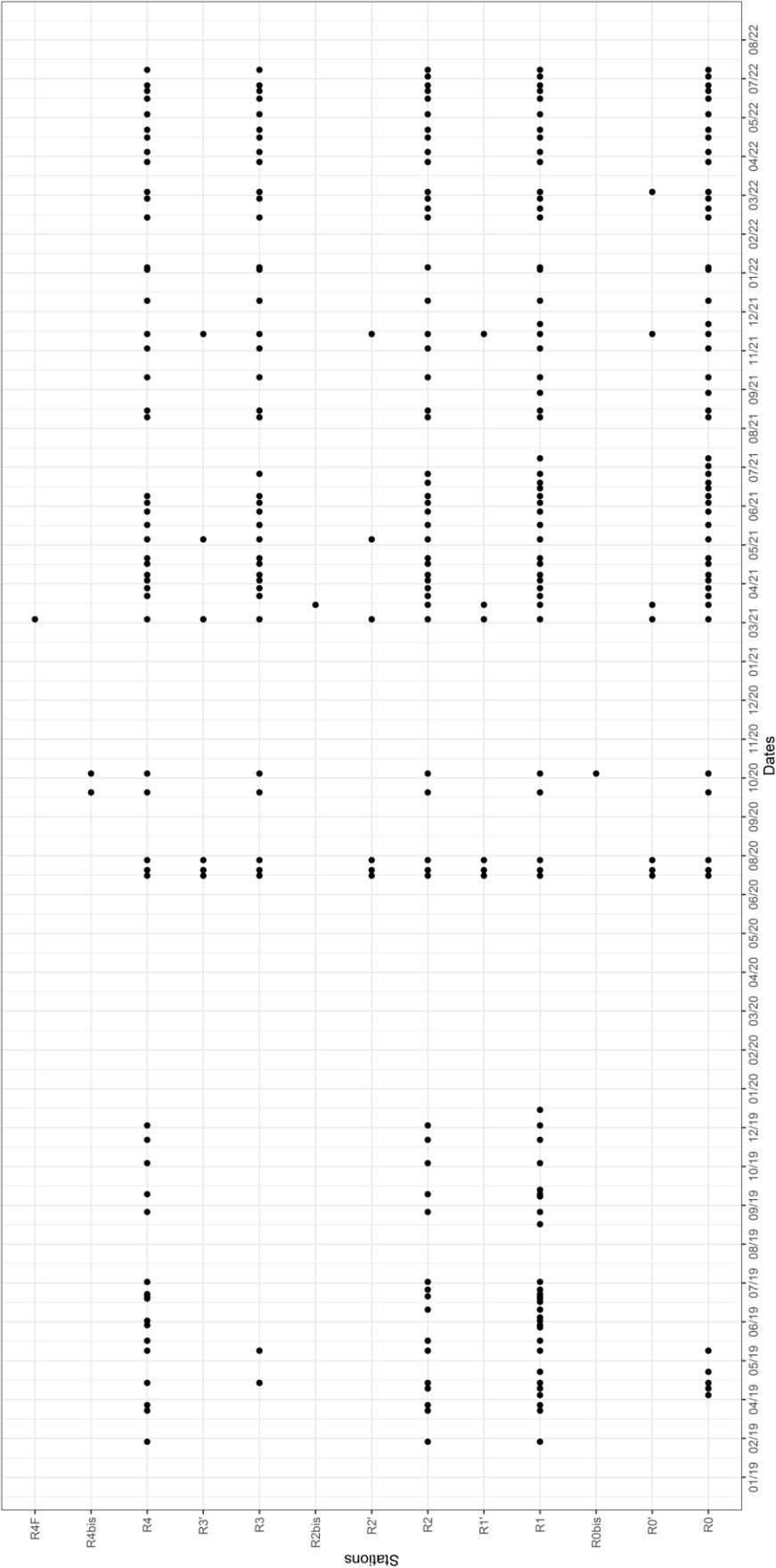
[RollingCal]

Enabled=1

ANNEXE 10: CONTEXTE DU FLOWCAM POUR LE X10

[Software]	CameraName=SenTech 1.3 Mp	AutoExportClassificationSummary=0
SoftwareName=VisualSpreadsheet	Initialization=	ClassificationTemplate=
SoftwareVersion=4.15.1	Shutter=18	[Fluid]
SoftwareBetaFlag=0	Gain=75	FlowCellType=FC100FV
CameraMagnification=10	WhiteBalanceU=2150	FlowCellDepth=100
GumStixFirmwareVersion=23	WhiteBalanceV=2150	FlowCellWidth=700
SerialNo=10439	Brightness=0	UseManualVolume=0
[General]	Gamma=0	ManualVolumeML=20.0000
TechnicianName=	Hue=2290	FluidVolumeImaged_ml=0.0000
LocationName=	Saturation=300	UseDeadVolumeCalculation=0
LocationLatitude=	Sharpness=0	DeadVolumeTubeLength_cm=0.00
LocationLongitude=	RedGain=0	DeadVolumeTubeDiameter_cm=0.16
RunStartTime=2018-12-17 20:28:26	GreenGain=0	UseDilutionCalculation=0
RunEndTime=2018-12-17 20:32:06	BlueGain=0	SampleDilution=1
TimeZone=	ColorGain=0	PumpFlowRate=0.3000
[CameraBehavior]	AutoAdjustIntensity=0	PrimingMethod=0
AutoImageRate=20	AutoIntensityTarget=140	PumpToFlowCellTubeLength_cm=0.00
FlashDuration=750.00	AutoIntensityShutterMin=6	ValveToCleanerTubeLength_cm=0.00
FlashAmplitude=75	AutoIntensityShutterMax=16	PumpErrorBudgetVolume=0.0050
FlashDelay=650.00	AutoIntensityGainMin=0	PumpErrorBudgetMultiplier=1.6000
RecoveryTime=1.00	AutoIntensityGainMax=680	PumpFlowCellLength_cm=3.5000
LaserCycleTime=1250.00	AutoIntensityBrightnessMin=0	PumpFlowCellFraction=0.5000
CameraTriggerDuration=100.00	AutoIntensityBrightnessMax=2047	CalibrationConstant=0.7300
PVDBRBC=1000.00	AutoIntensityFlashDurationMin=1.00	FringeSize=4.5000
PVDBRBE=1000.00	AutoIntensityFlashDurationMax=100.00	[CaptureRegion]
[Vertex]	AutoAdjustWhiteBalance=0	AcceptableLeft=121
ListDir=C:\VisualSpreadsheet	ImageWidth=1200	AcceptableRight=1100
[Fluorescence]	ImageHeight=1920	AcceptableTop=1
Ch1Measure=1	[RunTermination]	AcceptableBottom=1919
Ch1Trigger=1	MaxParticles=2000	ImagePixelBorder=2
Ch1Threshold=15000	StopOnMaxParticles=0	SaveImageFiles=1
Ch1HighThreshold=7000000	MaxRunTimeMinutes=60	SaveBinaryImageFiles=0
Ch2Measure=1	MaxRunTimeSeconds=0	SaveRawFiles=0
Ch2Trigger=1	StopOnMaxRunTime=0	MaxNumRawFiles=1000
Ch2Threshold=15000	MaxImageFiles=100	FlatFieldCount=10
Ch2HighThreshold=7000000	StopOnMaxImageFiles=0	[CaptureParameters]
QDAC_PMTSensitivity1=1435	MaxCameraImages=1000	DistanceToNeighbor=20
QDAC_PMTSensitivity2=1435	StopOnMaxCameraImages=0	ThresholdDark=20.0000
QDAC_LaserPower=3072	MaxFluidVolume=2.0000	ThresholdLight=20.0000
QDAC_FlashAmplitude=75	StopOnFluidVolume=1	CloseHoles=0
PMTSampleWidth=0	StopOnUserOnly=0	CaptureDarkOrLightPixels=0
MinTriggerPulseWidth=48	[CountCalibration]	MinESD=20.0000
ScatterDetector=0	CalibrationFactorPPML=1.2000	MaxESD=10000.0000
UseLogScale=0	bStandardVolumeCorrection=1	SimpleCaptureFilterDiameter=1
SkipMultiParticleImages=0	[Files]	UseAdvancedFilter=0
PMTGain=0	RecalibrationIntervalMinutes=30	FilterName=C:\Program
SmallParticleFilter=0	[CaptureReport]	Files\VisualSpreadsheet\FactoryDefault
[Laser]	AutoExportList=1	s\Filter - Aspect Ratio 0.9-1.0.fl
Enabled=1	AutoExportListSummary=1	[RollingCal]
[Camera]	AutoPrintSummary=0	Enabled=1
AutoTriggerFlag=1	AutoClassify=0	FrameCount=10
FrameCount=10		

ANNEXE 11: REPARTITION DES ECHANTILLONNAGES PAR STATION SUR LE FLOWCAM DEPUIS 2019



ANNEXE 12: GROUPEMENT POUR LES VIGNETTES EN X4

Diatomées	Nb de vignettes	Rang : Groupement
<i>Asterionellopsis glacialis</i>	31	Espèce
<i>Actinoptychus</i>	6	Ordre : Coscinodiscales
<i>Coscinodiscus spp</i>	26	
<i>Cerataulina pelagica</i>	110	Espèce
<i>Chaetoceros socialis chaîne</i>	109	Espèce
<i>Chaetoceros socialis pelote</i>	62	Espèce
<i>C.curvitus</i>	208	Espèce
<i>Chaetoceros sp</i>	369	Genre : Chaetoceros
<i>C.contortus</i>	2	
<i>C.danicus</i>	22	
<i>C.decipens</i>	4	
<i>Guinardia delicatula</i>	138	Espèce
<i>G. flaccida</i>	219	Espèce
<i>G.striata</i>	379	Espèce
<i>Leptocylindrus danicus</i>	1382	Espèce
<i>Pseudo-nitzschia</i>	243	Espèce
<i>Rhizosolenia imbricata</i>	444	Espèce
<i>Rhizosolenia pungens</i>	69	Famille : Rhizosoleniaceae
<i>Rhizosolenia styliformis</i>	6	
<i>Dactyliosolen fragilissimus</i>	10	
<i>Odontella regia</i>	118	Famille : Triceratiaceae
<i>Odontella rhombus</i>	16	
<i>Odontella sinensis</i>	10	
<i>Raphoneis amphiceros</i>	65	Espèce
<i>Navicula spp</i>	24	Ordre : Naviculales
<i>Pleurosigma</i>	26	
<i>Meuniera membranacea</i>	54	Espèce
<i>Paralia sulcata</i>	64	Espèce
<i>Colonie Cylindrotheca closterium et Phao</i>	166	Espèces coloniales
<i>Lauderia annulata</i>	46	Ordre : Thalassiosirales
<i>Skeletonema sp.</i>	21	
<i>Thalassiosira sp</i>	4	
<i>Thalassiosira rotula</i>	13	
<i>Ditylum brightwellii</i>	77	Espèce
<i>Thalassionema nitzschioides</i>	37	Espèce
Dinoflagellés		Rang : Groupement
<i>Diplopsalis spp</i>	69	Famille : Protoperidiniaceae
<i>Protoperidium sp</i>	4	Famille : Gymnodiniaceae
<i>Gyrodinium spirale</i>	117	
<i>Torodinium</i>	2	Genre : Prorocentrum
<i>Prorocentrum sp</i>	44	
Haptophytes		Rang : Groupement
<i>Phaeocystis globosa multiple</i>	577	Espèce
<i>P. coloniales</i>	193	Stade de l'espèce
Zooplankton		Rang : Groupement
<i>Apendicularia</i>	5	Groupe fonctionnel: Zooplankton
<i>Copepode</i>	4	

Larve	2	
Mytilus edulis	10	
Nauplius	34	
Ciliés		Rang : Groupement
Ciliés	84	Embranchement : Ciliophora
Laboa strobila	26	
Tiarina sp	3	

ANNEXE 13: ESPECES NE POUVANT FIGURER DANS LES SET D'APPRENTISSAGE

❖ Grossissement 100

Espèces minoritaires	Diatoms		Flagellates
Asterionellopsis glacialis	23	Cryptophyceae	6
Corethron	1	Dictyocha sp	7
Diatomees benthiques	3	Minuscula sp	15
Diploneis	2	Prorocentrum sp	4
Eucampia sp	2	Scrippsiella	2
H.tamesis	2		Zooplankton
Paralia sulcata	27	Brachionus sp	1
Dinobryon	4	Mytilus edulis	5
Trieres rhombus	10	Nauplius	8
Trieres sinensis	11	Oeufs	14
Tripus fusus	3		
Tripus lineatus	1		

❖ Grossissement 40

Espèces minoritaires	Nombre de vignettes
Dynobrion	3
Dino claire	2
Oeufs	3
Tripus fusus	6
Tripus lineatus	1
Dictyocha speculum	21

Cryptophyceae	14
<i>Asterionellopsis glacialis</i>	23
<i>Eucampia zodiacus</i>	15
<i>Dactyliosolen fragilissimus</i>	10
<i>R.styliiformis</i>	6
<i>O,rhombus</i>	16
<i>O.sinensis</i>	10

ANNEXE 14: GROUPE DU 1^{ER} SET D'APPRENTISSAGE

❖ Grossissement 100

Diatoms	Sous-groupe	Nb de vignettes	Total groupe	Flagellates	Sous-groupe	Nb de vignettes	Total groupe
<i>Actinoptychus</i>	Ordre : Coscinodiscales	6	44	<i>Phaeocystis globosa</i>	Espèce	665	680
<i>Coscinodiscus spp</i>		38		<i>P. coloniales</i>	Espèce coloniales	15	
<i>Cerataulina pelagica</i>	Espèce	111	111	<i>P. cox.</i>	Espèce coloniales	198	198
<i>Chaetoceros sp</i>	Genre	348	390	<i>P. ind.</i>	Espèce individu	524	524
<i>Chaetoceros eibonii</i>		3		<i>Gyrodinium sp</i>	Famille : Gymnodiniaceae	32	118
<i>C.concortus</i>		1		<i>G.spirale</i>		50	
<i>C.curvisetus</i>		21		<i>K.glaucum</i>		27	
<i>C.affinis</i>		6		<i>Torodinium sp</i>		6	
<i>C.danicus</i>		11		<i>Amphidinium</i>		1	
<i>C.deciapiens</i>	Espèce	175	175	<i>Gymnodinium</i>		2	
<i>Chaetoceros socialis</i>	Espèce en colonie	154	154	Dinoflagellates	Embranchement : Dynophyta	32	36
<i>Ditylum brightwellii</i>	Espèce	132	132	<i>Dino clair</i>		4	
<i>G. flaccida</i>	Espèce	129	129	<i>Protoperidinium</i>	Famille : Protoperidiniaceae	15	56
<i>G.delicatula</i>	Genre	42	76	<i>Diplopsalis spp</i>		41	
<i>G.fragilissima</i>		2					
<i>G.striata</i>		32					
<i>Leptocylindrus danicus</i>	Espèce	1921	1921	Ciliates			
<i>P-nitzschia</i>	Famille: Bacillariaceae	464	466	Ciliés spp		105	120
<i>Cylindrotheca closterium</i>		2		<i>Laboa strobila</i>		8	
<i>Rhizosolenia imbricata</i>	Espèce	219	219	<i>Lohmanniella oviformis</i>		3	
<i>Rhizosolenia pungens</i>	Espèce	45	45	<i>Mesodinium</i>		1	
<i>Raphoneis amphiceros</i>	Espèce	85	85	<i>Tiarina sp</i>		3	
<i>Navicula spp</i>	Ordre : Naviculales	61	94				
<i>Meuniera membranacea</i>		17					
<i>Pleurosigma</i>		16					

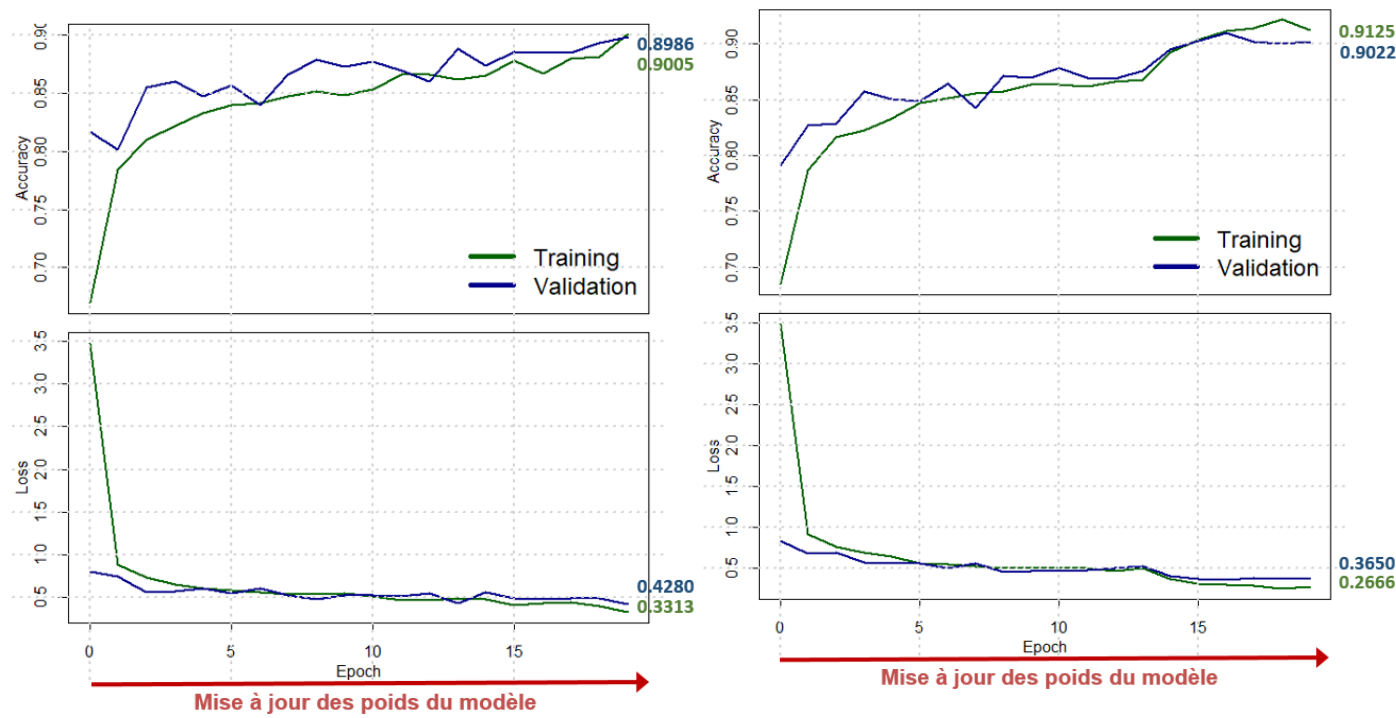
<i>Thalassionema nitzschioides</i>	Espèce	55	55
<i>Skeletonema</i>	Ordre : Thalassiosirales	16	45
<i>Lauderia annulata</i>		10	
<i>Thalassiosira sp</i>		4	
<i>Thalassiosira rotula</i>		11	
<i>Thalassiosira punctigera</i>		3	
<i>Detonula</i>		1	

❖ Grossissement 40

Diatoms	Sous-groupe	Nb de vignettes	Total groupe	Ciliates	Sous-groupe	Nb de vignettes	Total groupe
Actinoptychus	Coscinodiscales	6	32	<i>Ciliés</i>	Ciliophora	84	113
Coscinodiscus spp		25		<i>Laboa strobila</i>		26	
Cerataulina pelagica	Espèce	110	110	<i>Tiarina sp</i>		3	
Chaetoceros socialis chaîne	Espèce	146	146	Flagellates	Sous-groupe	Nb de vignettes	Total groupe
Chaetoceros socialis pelote	Espèce	62	62	<i>Diplopsalis spp</i>	Protoperidiniaceae	69	73
C.curvitus	Espèce	192	192	<i>Protoperidium sp</i>		4	
Chaetoceros sp	Chaetocerotaceae	346	374	<i>G.spirale</i>	Gymnodiniaceae	117	119
C.concortus		2		Torodinium		2	
C.danicus		22		<i>Prorocentrum sp</i>	Genre	43	43
C.decipens		4		<i>Phaeocystis globosa</i>	Espèce	577	577
Guinardia delicatula	Espèce	138	138	<i>P. coloniales</i>	Stade de l'espèce	192	192
G. flaccida	Espèce	219	219	<i>P. cox. G</i>	Stade de l'espèce	102	102
G.striata	Espèce	379	379	<i>P. Cox.P</i>	Stade de l'espèce	243	243
Leptocylindrus danicus	Espèce	1382	1382	Zooplankton	Nb de vignettes	Total groupe	
P-nitzschia	Espèce	242	242	Apendicularia	5	53	
Rhizosolenia imbricata	Espèce	444	444	Copepode	4		
Rhizosolenia pungens	Rhizosoleniaceae	69	85	Larve	2		
Rhizosolenia styliformis		6		Mytilus edulis	10		
Dactyliosolen fragilissimus		10		Nauplius	32		
Odontella regia	Triceratiaceae	118	144				
Odontella rhombus		16					
Odontella sinensis		10					
Raphoneis amphiceros	Espèce	65	65				
Navicula spp	Naviculales	24	103				
Meuniera membranacea		54					
Pleurosigma		25					
Paralia sulcata	Espèce	64	64				

Colonie Cylindrotheca closterium et Phao	Espèces coloniales	166	166
Lauderia annulata	Thalassiosirales	46	82
Skeletonema		20	
Thalassiosira sp		4	
Thalassiosira rotula		12	
Ditylum brightwellii	Espèce	77	77
Thalassionema nitzschioides	Espèce	37	37

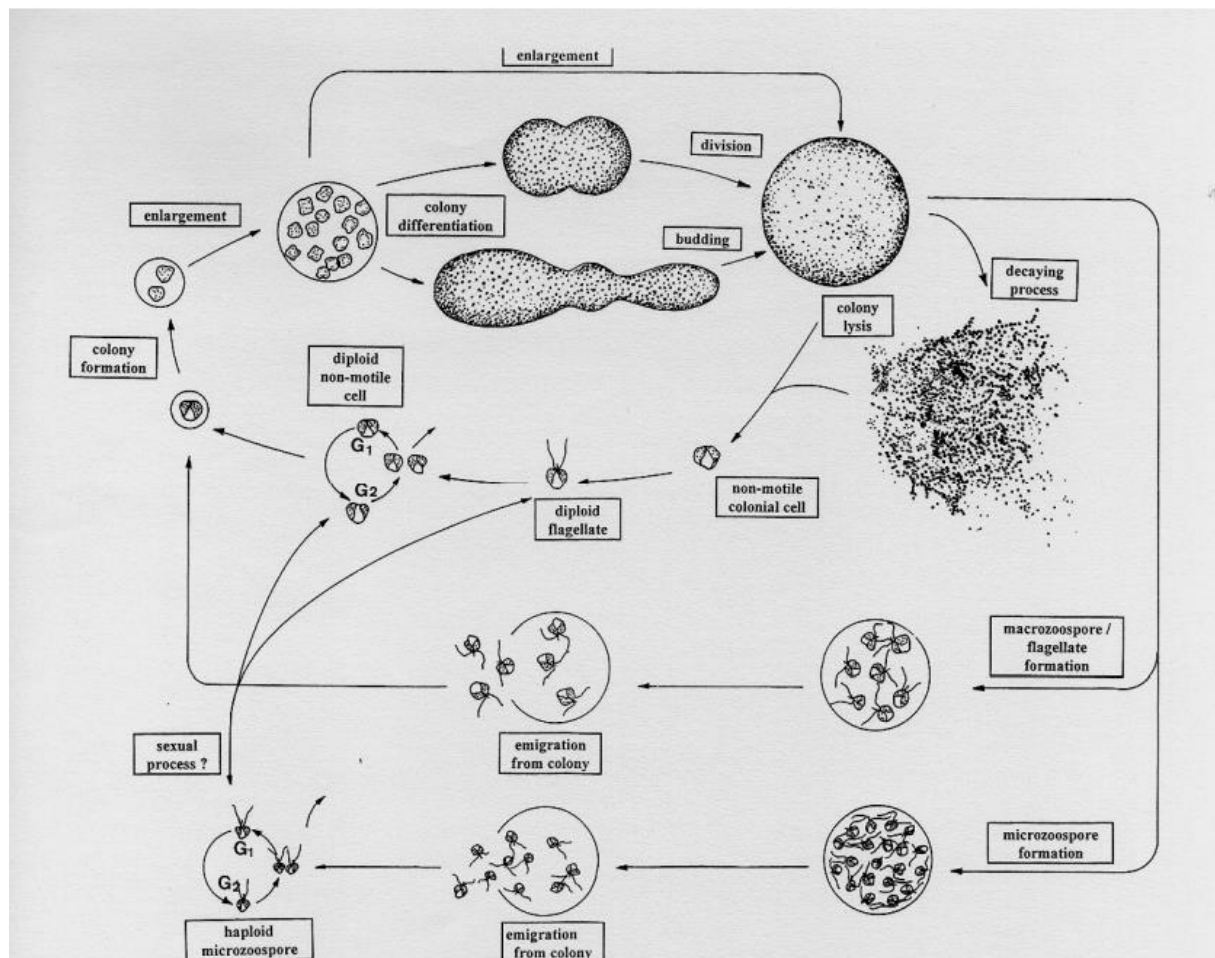
ANNEXE 15: EVALUATION DES SET D’APPRENTISSAGE PAR VALIDATION CROISEE ET PREDICTION POUR L’OBJECTIF X10 (1^è SET A GAUCHE ET 2^è SET A DROITE)



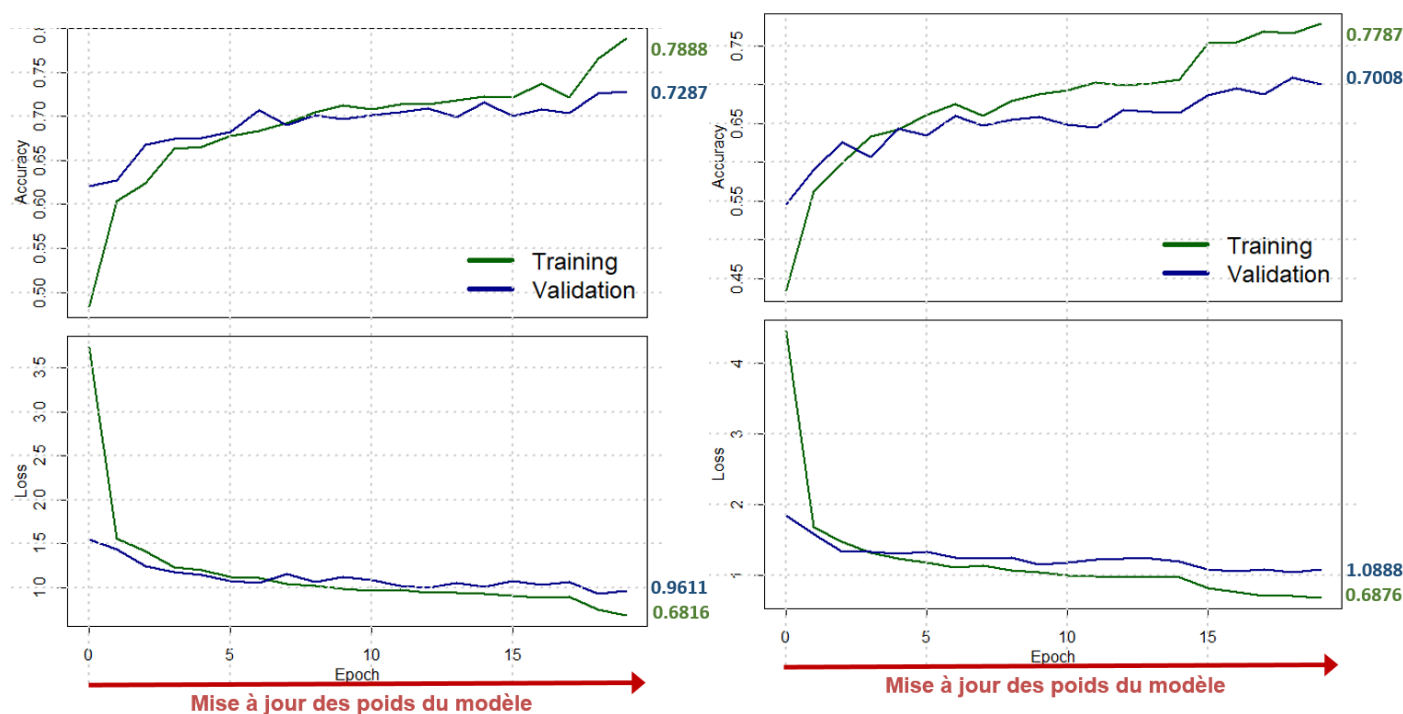
Accuracy ≈ pourcentage de données correctement classifiées
 Loss ≈ distance entre les données réelles et les données prédites

Training Accuracy ≈ score obtenu par validation croisée
 Validation Accuracy ≈ score de prediction obtenu sur le set de test

ANNEXE 16: CYCLE DE VIE DE PHAEOCYSTIS (ROUSSEAU ET AL., 1994).



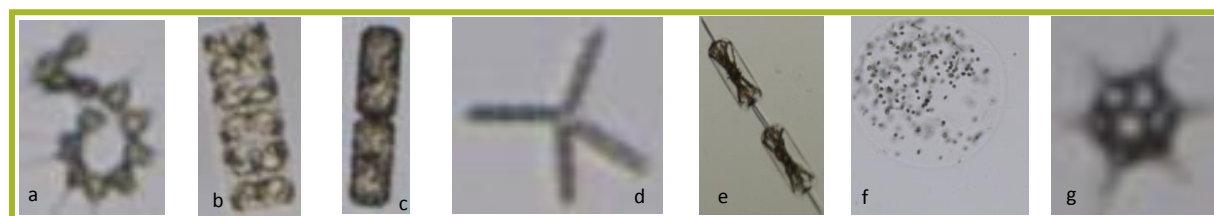
ANNEXE 17: EVALUATION DES SET D'APPRENTISSAGE PAR VALIDATION CROISEE ET PREDICTION POUR LE x4 (1E SET A GAUCHE ET 2E SET A DROITE)



Accuracy ≈ pourcentage de données correctement classifiées
Loss ≈ distance entre les données réelles et les données prédites

Training Accuracy ≈ score obtenu par validation croisée
Validation Accuracy ≈ score de prediction obtenu sur le set de test

ANNEXE 18: EXEMPLE DE VIGNETTES x4 (A. *ASTERIONELLOPSIS GLACIALIS*, B. *MEUNIERA MEMBRANACEA*, C. *CERATAULINA PELAGICA*, D. *THALASSIONEMA NITZSCHOIDES*, E. *DITYLUM BRIGHTWELLII*, F. *PHAEOCYSTIS GLOBOSA* COLONIE, G. *DICTYocha SPECULUM*)



Contribution à l'amélioration de l'approche d'imagerie automatisée pour l'étude du phytoplancton marin avec le FlowCam

Résumé : L'utilisation de méthode d'imagerie automatisée a permis une avancée dans le suivi du phytoplancton à grande échelle. L'amélioration du protocole du FlowCAM afin d'exploiter les données de ce dernier au sein du LOG a été l'objet de mon travail. La création d'un nouveau set d'apprentissage et des analyses grâce à un modèle d'apprentissage novateur basé sur le *deep learning* permet d'obtenir des premiers résultats encourageants sur l'abondance du phytoplancton sur le point R1 de la radiale DYPHYRAD. Les critiques de ce premier travail sont également présentés dans ce rapport.

Mots Clés : FlowCam – Phytoplancton – set d'apprentissage – Manche orientale – Mer du Nord – méthode automatisé

Abstract: The use of automated imaging methods has allowed a breakthrough in the monitoring of phytoplankton on a large scale. The improvement of the FlowCAM protocol in order to exploit its data within the LOG was the subject of my work. The construction and test of a new learning set and the analyses performed by a novel learning and classification approach (Convolutional Neural Network) allows obtaining first encouraging results on phytoplankton abundance on the station « R1 » of the DYPHYRAD radial. The criticisms of this first work are also presented in this report.

Keywords : FlowCam - Phytoplankton - training set - Eastern Channel - North Sea - automated method

Laboratoire d'Océanologie et Géosciences – UMR 8187
32 avenue Foch - 62930 Wimereux

Tuteur entreprise :
Luis Felipe Artigas
Enseignant-chercheur

Tuteur académique :
Pierre Peeters