



Université de Toulon
ECOLE DOCTORALE 548 Mer et Sciences

Ecole d'ingénieurs de l'université de Toulon, SeaTech

LABORATOIRE D'INFORMATIQUE ET DES SYSTEMES (LIS)/UMR CNRS 7020

Thèse présentée pour obtenir le grade universitaire de docteur

Disciplines : Automatique, signal, productique, robotique

Marie LAMOURET

Titre de la thèse : “ Traitements automatisés des données
acoustiques issues de sondeurs multifaisceaux pour la
cartographie des fonds marins. ”

Soutenue le 23/03/2022 devant le jury :

Caroline FOSSATI	Prof. des Universités, École Centrale Marseille	Rapporteur
Ali MANSOUR	Professeur, ENSTA-Bretagne	Rapporteur
Jérôme MARS	Prof. des Universités, Université Grenoble Alpes	Examinateur
Guillaume MATTE	Ingénieur de recherche & développement, IxBlue	Examinateur
Pierre BOISSERY	Expert mer, Agence de l'eau Rhône Méditerranée Corse	Examinateur
Nadège THIRION-MOREAU	Prof. des Universités, Université de Toulon	Directeur de thèse
Christophe VIALA	Président, Seaviews	Encadrant de thèse



Cette oeuvre est mise à disposition selon les termes de la [Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 3.0 France](#).

Résumé

Le sondeur multifaisceaux (SMF) est l'une des technologies d'acoustique sous-marine les plus avancées pour l'étude des fonds et de la colonne d'eau. Il requiert une réelle expertise pour son déploiement sur le terrain ainsi que pour l'élaboration de cartographies à partir des différentes données acquises. Ces traitements sont souvent chronophages en raison de la quantité de données acquises et demandent à être automatisés pour alléger le travail à l'hydrographe. C'est ce sur quoi portent les travaux réalisés durant cette thèse.

Après des rappels sur des notions d'acoustique sous-marine, le fonctionnement du SMF est décrit et les types de données manipulées tout au long des traitements sont présentés. Le manuscrit s'articule ensuite autour de deux thématiques : la cartographie bathymétrique et la cartographie biocénotique. Les développements sont intégrés dans les logiciels de l'entreprise Seaviews pour laquelle les travaux sont réalisés. Ils répondent à des besoins particuliers de l'entreprise.

En ce qui concerne la cartographie bathymétrique, la donnée bathymétrique doit être préalablement triée pour écarter les sondes aberrantes et éviter qu'elles ne pénalisent la précision topographique. Ce tri d'innombrables sondes est une tâche que réalisent les hydrographes, assistés aujourd'hui d'outils numériques. Nous proposerons une méthode statistique rapide pour trier les sondes tout en réalisant une carte de profondeurs marines. Ce qui amène à se demander si les images de la colonne d'eau acquises également par le sondeur ne seraient pas exploitables pour déduire une bathymétrie exempte d'aberration. Nous testerons cette hypothèse à l'aide de l'apprentissage profond (*deep learning*) et en particulier par des réseaux de neurones convolutifs qui ont permis des progrès considérables en vision par ordinateur.

La cartographie des habitats marins (les biocénoses) est un travail de classification de la nature des fonds à partir des données acoustiques du SMF en concordance avec les espèces vivant sur les lieux. La société Seaviews a développé une méthode de préparation des données SMF pour l'analyse des habitats. Nous nous orientons vers des méthodes de classification des habitats, à partir de ces données, par des techniques d'apprentissage automatique (*machine learning*). Plusieurs méthodes sont mises en place et testées, puis une zone d'étude est choisie pour évaluer et comparer les résultats des différentes approches.

Mots clés : Sondeurs multifaisceaux ; Acoustique sous-marine ; Bathymétrie ; Analyse de données de la colonne d'eau ; Cartographie des habitats marins ; Réseaux de neurones convolutifs ; Apprentissage supervisé ; Apprentissage automatique

Abstract

Among underwater acoustic technologies, multibeam echo sounder (MBES) is one of the most advanced tool to study and map the underwater floors and the above water column. Its deployment on-site requires expertise so as the whole data processing to map the information. These processing are very time-consuming due to the massive quantity of recorded data and thus needs to be automatised to shorten and alleviate the hydrographer's task. This PhD research works focus on the automatisation of the current activities in Seaviews society.

After some reminders on the underwater acoustic sciences, the MBES operating is described as well the produced data that will be manipulated throughout the developments. This document presents two thematics : bathymetric (depths) and marine habitats mapping. The developments are integrated into the Seaviews' software in the aim to be used by all the employees.

About seafloor depths mapping, the bathymetric sounding has to be sorted to avoid that the outlier errors distort the results. Sorting the uncountable measures is cumbersome but necessary, although the hydrographers are today happily computed-assisted. We propose a fast statistical method to exclude the outliers while mapping the information. This leads to wonder if the water column imagery would be workable to deduce the bathymetry without failure. We will test this hypothesis with some technics of deep learning, especially with convolutional neural networks.

The marine habitats mapping is a seabed nature classification according to the local life. Seaviews has worked on a way to prepare MBES data and habitats analysis. Concerning the method of classification itself, we move towards machine learning technics. Several methods are implemented and assessed, and then an area is chosen to evaluate and compare the results.

Keywords : Underwater acoustic ; Multibeam echosounder ; Bathymetry ; Water column data analysis ; Marine habitats mapping ; Convolutive Neural Networks ; Supervised learning ; Machine learning ;

Remerciements

Je voudrais tout d'abord remercier Nadège Thirion-Moreau, ma directrice, d'avoir accepté de me chapeauter durant ces trois années de thèse, alors qu'elle m'avais déjà bien assez supporté avec mes râleries en école d'ingénieur. Je la remercie vraiment d'avoir toujours été disponible et encourageante, d'avoir réussi à me suivre alors que j'ai beaucoup été en entreprise et de la patience qu'elle a eu pour relire mes travaux.

Je remercie mes encadrants de m'avoir accueillie à Seaviews et de m'avoir fait confiance. Nommément, je remercie Christophe Viala pour son soutien technique dans les développements logiciel et son expertise en acoustique qui ont toujours permis d'aboutir à des solutions et à de bonnes idées. Je remercie Arnaud Abadie pour son support moral d'ancien thésard et de l'apport en biologie sous-marine dans les travaux et articles. Plus globalement, un grand merci à toute l'équipe, Christophe, Arnaud et Léo Roques, pour ces trois années de travail dans la bonne humeur au bureau comme sur le terrain. Mais également, de m'avoir amené au monde de la plongée sous-marine loisir et professionnelle et de m'avoir fait découvrir, en autre et surtout, la Corse par des visites atypiques (ports de commerce, émissaire d'eaux usées, recensement de corps morts de mouillage, etc.);). Je tiens également à remercier Alice Nicolas pour son stage de "débroussaillage" en méthodes de *machine learning*, sans qui je n'aurais pas pu aborder la thématique de classification des fonds durant ces trop courtes années.

Je souhaite également remercier l'Agence de l'Eau Rhône-Méditerranée-Corse pour son soutien dans les projets de R&D de Seaviews, et en particulier, Pierre Boissery, pour les discussions toujours intéressantes et pour son regard général sur les enjeux dans lesquels s'inscrivent ces projets.

Je souhaite également remercier les membres de mes comités de suivi de thèse, Lucie Somaglino et Jérôme Mars, pour leur disponibilité et leur bienveillance. De même, je remercie les membres de mon jury, Caroline Fossati, Ali Mansour, Guillaume Matte et Jérôme Mars, qui ont accepté d'évaluer ce travail écrit et oral.

Enfin et bien sûr, un grand merci à mon entourage familial et amical. À mes parents, qui à la Pagnol, m'ont toujours souhaité de bonnes missions par :

“Quand tu vas commencer à mesurer le fond de la mer, fais bien attention de ne pas trop te pencher, et de ne pas tomber par dessus bord et là où ça sera trop profond, laisse un peu mesurer les autres.”

À mon compagnon avec son indéfectible soutien moral tout au long de cette thèse et qui a réussi à entretenir mon niveau de plongée même quand l'eau est froide pour moi. À son père qui devant en avoir assez de la réponse "oui, ma thèse avance, plus qu'elle ne recule" m'a dors et déjà décerné la mention *summa cum laude* ! Et finalement, à Alice et Aimy, la meilleure cuvée de stagiaires, qui m'ont fait passé un merveilleux été à La Ciotat entre plage, escalade et repas gras malgré la rédaction de ce manuscrit !

Table des matières

Résumé	iii
Abstract	iv
Remerciements	v
Liste des figures	xii
Liste des tableaux	xiv
Liste des notations	xv
Introduction Générale	1
I Généralités	5
Introduction	6
1 Notions de systèmes acoustiques sous-marins	8
1.1 Propagation des ondes acoustiques en milieu marin	8
1.2 Composition d'un système acoustique	14
1.3 L'équation du sonar	16
1.4 Mesures hydrographiques par sondeurs acoustiques	17
2 Système d'acquisition de données au sondeurs multifaisceaux	21
2.1 Sondeurs multifaisceaux : fonctionnement & principales caractéristiques	21
2.2 Instrumentation associée : montage de Seaviews	26
2.3 Données issues des Sondeurs multifaisceaux	29
Conclusion	34
II Cartographie bathymétrique	35
Introduction	36
3 Traitements des sondes bathymétriques	37
3.1 État de l'art du filtrage des fausses sondes bathymétriques	37
3.2 Filtrage spatial automatique des sondes bathymétriques	41
3.3 Application sur une zone très bruitée	44
3.4 A propos de la méthode	48

4 Extraction de la bathymétrie dans l'imagerie de la colonne d'eau	50
4.1 Traitement d'image par réseau de neurones	51
4.2 Mise en place d'une méthode d'apprentissage profond	60
4.3 Premiers résultats	63
4.4 Évaluation des objectifs	66
Conclusion	72
III Classification des fonds et des habitats marins	73
Introduction	74
5 Etat de l'art de la classification des fonds par acoustique	76
5.1 Étapes de la chaîne de classification des fonds marins	76
5.2 Méthodes de classification automatique	83
6 Étude biocénotique à Centuri	94
6.1 Zone d'étude et acquisition des données	94
6.2 Résultats des classifications	96
6.3 Discussion des méthodes et résultats	106
Conclusion	111
Conclusion Générale	112
Annexes : Activités biologiques dans la colonne d'eau	114
A Measuring fish activities as additional environmental data during a hydrographic survey with a multi-beam echo sounder	114
A.1 Data acquisition and seabed mapping	116
A.2 Data processing	118
A.3 Discussion	124
B Coupling the map of marine habitats and fish accumulation zones : a three dimensional spatial approach for the management of halieutic resources.	127
B.1 Materials and Methods	128
B.2 Results	130
B.3 Discussion	132
Bibliographie	136

Table des figures

1	Principaux mots clés pour qualifier les activités de la société Seaviews.	1
1.1	Exemple de profils de vitesse du son et de température en mer Méditerranée. Le profil d’hiver a été réalisé en baie de Saint-Florent (Haute-Corse) en mars 2020. Le profil d’été a été réalisé en baie de La Ciotat (Bouches-du-Rhône) en août 2019. . .	13
1.2	Phénomène de réflexion et de réfraction d’une onde sur une interface plane (les couches d’eau de célérités croissantes) et de diffusion sur les interfaces rugueuses (fond et surface). Illustration des déformations et multi-trajectoires d’un rayon. . .	14
1.3	Diagrammes de directivité d’une antenne linéaire, de dimension L , à différentes fréquences d’émission de longueur d’onde λ	15
1.4	Rôle d’un <i>Time Varying Gain</i> dans la compensation du signal acoustique.	18
1.5	Principe du sonar à balayage latéral (sur un seul côté) et allure du signal reçu. D’après Lurton (1998).	19
1.6	Épave de sous-marin de 45 m de long par 24-30m de profondeur, posé kiosque en l’air sur un fond de sable. A gauche, imagerie acoustique de 500 pings obtenue avec le SMF de Seaviews. A droite, la carte bathymétrique, résultante de plusieurs transects, au SMF du même sous-marin pour aider à la visualisation.	20
2.1	A gauche, le faisceau d’un sondeur monofaisceau est large et dirigé vers le nadir. A droite, les faisceaux d’un SMF sont fins et dispersés angulairement sur une fauchée perpendiculaire au navire.	21
2.2	Antenne linéaire avec et sans retard temporel entre capteurs, respectivement sans et avec dépointage de l’onde).	22
2.3	Diagrammes de directivité d’une antenne non dépointée (à gauche) et dépointée de 30° (à droite)	23
2.4	Croisement des faisceaux d’émission et de réception.	23
2.5	Méthodes de détection du fond dans le signal rétrodiffusé par le fond.	24
2.6	Tête (à gauche) et interface (à droite) du sondeur multifaisceaux de la marque R2Sonic version 2022. La tête est composée d’une antenne de réception rectangulaire (à gauche) et d’une antenne d’émission cylindrique (à droite). Le module d’interface sert à la création des impulsions et aux réglages des paramètres. (Source : R2Sonic)	25
2.7	Base et rover pour du RTK GNSS	27
2.8	Attitude d’un navire : mouvement de roulis, tangage, cap. Mouvement de pilonnement.	28
2.9	Célérimètre Valeport miniSVS composé d’un transducteur pour la génération d’OA dans la partie cylindrique et d’une cage avec réflecteur calibré (à droite) pour la mesure de la célérité. (Source : Valeport)	28
2.10	Navire océanographie de Seaviews embarquant deux opérateurs et le SMF	29
2.11	Passage au dessus d’un bloc rocheux de 8m par 5.5m sur un fond plat. A gauche : ping sous forme de fauchée perpendiculaire au navire, les 1024 sondes sont réparties sur ± 10 m autour du nadir (ligne noire verticale). A droite : ensembles de pings consécutifs, le bloc a été coloré en rose artificiellement pour faciliter son identification.	30
2.12	L’ensemble des sondes sont distribuées dans des mailles carrées de différentes tailles, de gauche à droite, 50 cm, 10 cm et 5 cm.	31

2.13	Image de réflectivité d'une roche sur sable près de Vendres (Hérault). Les sédiments ont une moins forte réflectivité que les roches.	32
2.14	Image water column avec l'épave d'un voilier, en baie de La Ciotat par 42m de profondeur.	33
2.15	Deux exemples de bancs de poissons. A gauche, un banc de poissons au dessus d'une roche à 30m de profondeur. A droite, un banc en plein eau à 10m de la surface et du fond.	33
2.16	Travaux sous-marins de Seaviews. En haut à gauche, échappement de bulles d'un plongeur. En haut à droite, balisage sous-marin contenant de l'air, donc visible par acoustique. En bas à gauche, bout tenu en surface par un flotteur, quelques poissons à droite.	33
3.1	Les sondes vertes sont considérées comme valides par le masque d'inclusion asservi (cadre rouge). Les sonde en rouge sont éliminées.	39
3.2	A gauche, le masque asservi (cadre rouge) ne suffit pas à rejeter toutes les sondes aberrantes. A droite, le masque d'exclusion (cadre bleu) neutralise les sondes en bleu et le masque asservi se réadapte automatiquement.	39
3.3	Lecture de données issues du sondeur multifaisceaux (SMF), de la centrale d'attitude, du GPS, du célérimètre (section 2.2) dans <i>ViewSMF</i> et chargement dans un modèle numérique de terrain (MNT) dans <i>ViewMap</i>	41
3.4	A gauche, une projection cylindrique tangente à l'équateur. Les projections cylindriques peuvent également être tangentes aux pôles ou à n'importe quel autre point de la Terre. A droite, une projection conique tangente où le sommet du cône est centré sur le pôle nord. D'autres projections coniques existent avec des variations sur l'emplacement du sommet ou encore avec un cône sécant (et non tangent) à la Terre. (source des images : Wikipedia)	42
3.5	Exemple d'un affichage à l'écran d'un MNT : des couleurs artificielles permettent une visualisation synthétique des valeurs. D'un point de vue structure numérique, un MNT ressemble à un tableau bidimensionnel de valeurs, même si la forme réelle est quelconque.	43
3.6	A gauche, seule les mailles valides sont encore visibles, les autres sont grisées en attente de correction. Au milieu et à droite, durant la correction, les sondes des mailles voisines d'une maille à état valide sont triées et passent à un état valide. Ce MNT est corrigé rapidement en deux passes.	43
3.7	Carte de situation du canyon d'étude dans l'Anse du Mugel à La Ciotat (France)..	44
3.8	Modèle numérique de terrain et vue 3D créés à partir de la donnée bathymétrique brute.	45
3.9	Modèle numérique de terrain et vue 3D créés à partir des données bathymétriques filtrées manuellement.	46
3.10	Résultats de la sélection des mailles valides selon la résolution du MNT. Le cadre rouge correspond au contour de la zone en figure 3.8 et 3.9, il donne un aperçu de la quantité de mailles brutes non valides.	47
3.11	Résultats du filtrage automatique sur le MNT à différentes résolutions.	48
3.12	Vues 3D des résultats de la figure 3.11.	49
4.1	Deux exemples de WCI où le fond de la mer est bien visible mais dont une partie des sondes bathymétriques est erronée (en bleu).	50
4.2	A gauche, schéma d'un neurone biologique. A droite, le neurone formel de Mcculloch and Pitts (1943)	51
4.3	Exemple d'un Perceptron multicouche composé d'une couche d'entrée, de deux couches cachées et d'une couche de sortie. Chaque couche est totalement connectée avec la précédente et la suivante.	52
4.4	Réseau LeNet-5 proposé par LeCun et al. (1998) pour la reconnaissance de chiffres manuscrits.	54

4.5	Couche de convolution sur une image d'entrée augmentée d'une marge de 1 pixel avec un filtre de taille 3x3 et un chevauchement de 1 pixel.	55
4.6	Couche de réduction avec un <i>Max-Pooling</i> de filtre de 2x2 et stride 2.	56
4.7	Modules du résidu à gauche et d'inception à droite, issus respectivement de He et al. (2015) et de Szegedy et al. (2015).	58
4.8	Représentation des cinq formes de CNN les plus rencontrées. Les différentes couches ou modules sont représentées par des rectangles dont les dimensions reflètent le nombre et la taille des caractéristiques calculées à chaque étape. Ces dimensions donnent l'allure des réseaux sans être réellement quantitatives.	59
4.9	Un rayon acoustique se propage selon une trajectoire connue (flèches). A différents temps de propagation, on évalue la différence de profondeur entre celle atteinte par le rayon et celle observée au niveau de la maille du MNT où le rayon viendrait se projeter. C'est le temps t_3 qui correspond le mieux au moment où le signal rencontre le fond dans cet exemple.	60
4.10	Ensemble des éléments du projet d'apprentissage profond sur les données <i>water column</i> pour en déduire la bathymétrie. Les éléments verts sont issus de la librairie MxNet, tandis qu'en noir ce sont des codes que nous avons développé au niveau de la chaîne de traitement global.	63
4.11	Diversité des formes de fonds - fonds plus ou moins plats, reliefs plus ou moins marqués, digues, etc. - et de réglages du sondeur - ouverture angulaire, inclinaison, position plus ou moins haute du fond dans les WCI - composant le jeu d'apprentissage.	63
4.12	Suivi des valeurs de <i>loss</i> et de RMSE à la fin de chaque <i>epoch</i> de l'apprentissage.	64
4.13	Trois exemples de WCI avec une bathymétrie prédite. Sans complication comme à gauche, la prédiction est correcte. En présence de bruit (à droite) ou d'une activité dans la colonne d'eau (en bas), la prédiction perd en précision.	64
4.14	Écart entre un MNT de référence (avec la bathymétrie issue du sondeur nettoyée des fausses sondes) et un MNT dont la bathymétrie est déduite des images colonne d'eau. Les couleurs chaudes désignent les mailles de MNT où la bathymétrie prédite est plus faible que la réalité, tandis que les couleurs froides désignent une bathymétrie prédite plus profonde. La couleur médiane, grise, désigne les mailles où les bathymétries estimées sont comparables.	65
4.15	Un fond plat correctement prédit à gauche. Au milieu et à droite, les deux mauvaises prédictions de fonds plats sont difficilement explicables...	67
4.16	Les reliefs sont prédits avec plus ou moins de précision.	67
4.17	A gauche et au milieu, respectivement le MNT de référence (avec la bathymétrie issue du sondeur) et un MNT dont la bathymétrie est déduite par RNA appliqué aux images colonne d'eau.	67
4.18	Exemples de prédictions de la digue ainsi que des blocs épars.	68
4.19	En haut, la représentation 3D du MNT de référence. Au milieu, ce même MNT avec la bathymétrie déduite. En bas, différence des deux MNT.	69
4.20	A gauche, représentation du CNN <i>from scratch</i> testé, le code couleur est celui de la figure 4.8. A droite, prédiction d'un fond plat en présence d'un banc de poissons.	70
5.1	Exemple de post-traitements sur une zone carrée de 300m de côté à maille de 50 cm. Les erreurs sont retirées, les trous remplis et les frontières entre objets lissées.	79
5.2	Représentation d'un arbre de décision où chaque classe est définie par rapport à des valeurs issues des données de bathymétrie, de rugosité et de réflectivité ainsi qu'aux descriptions des systèmes de classification des habitats. Les valeurs sont données à titre d'exemple.	82
5.3	De gauche à droite, les différentes étapes pour réaliser une partition de données par la méthode des <i>K</i> -moyennes. Les centres sont représentés par des étoiles dont l'initialisation est prise au hasard.	84
5.4	A gauche, deux lunes mal partitionnées par une méthode des <i>K</i> -moyennes à deux classes. A droite, les deux lunes sont découpées en plusieurs classes bien distinctes.	85

5.5	A gauche, différents séparateurs linéaires (en bleu) discriminant les éléments rouge et vert. A droite, le séparateur maximisant la marge entre les deux classes, les deux éléments cerclés de noir sont les vecteurs supports du séparateur.	87
5.6	A gauche, un SVM linéaire échoue à traiter deux lunes non linéairement séparables. A droite, les deux lunes sont partagées par une frontière lisse quitte à laisser quelques éléments du mauvais côté.	88
5.7	Exemple d'arbre de décision construit à partir des descriptions des classes de la figure 5.2	90
5.8	Exemple du fonctionnement d'une forêt d'arbres décisionnels : plusieurs sous-jeux sont extraits d'un jeu de données d'apprentissage et participent à la construction de plusieurs modèles de décisions.	90
5.9	Trois arbres sont créés avec des variantes du jeu d'apprentissage, leurs résultats sont différents et comportent quelques signes de surapprentissage (séparation pour un unique élément). Leur moyenne produit un résultat où le surapprentissage est réduit.	91
5.10	Résultats de classification avec le Bayésien naïf. Le résultat s'approche de ceux du SVM linéaire (figure 5.6 gauche) et du K -moyenne à deux partitions (figure 5.4 gauche). Cependant les éléments mal classés ont un taux de confiance sur la prédiction plus faible que les autres.	92
6.1	Carte de situation de l'île méditerranéenne de la Corse en France (en haut à gauche) et la localisation de la commune de Centuri au Cap Corse (en bas à gauche). A droite, zone d'étude autour du port de Centuri et de la réserve naturelle de l'îlot de Capense.	94
6.2	A gauche, image aérienne de la BD ORTHO. A droite, Bathymétrie de la zone sous forme d'un MNT de résolution 50 cm, fusion de la bathymétrie acoustique et lidar.	95
6.3	A gauche, la carte de réflectivité des fonds issue du SMF montre la part de la zone acquise en acoustique (le reste correspond à de trop petits fonds). Au milieu, la carte de pente issue de la bathymétrie (acoustique et lidar). A droite, la carte de rugosité issue de la bathymétrie (acoustique et lidar).	96
6.4	Photographies de vérités terrain réalisées par caméra tractée (les trois photos de gauche) et par plateforme (photos de gauche et du bas).	97
6.5	Résultats de classification manuelle par interprétation des différentes données disponibles.	98
6.6	Résultats des K -moyennes, avec $K=3$, pour différents lots de données.	99
6.7	Evolution de l'inertie intra-classe et du temps de calcul avec l'augmentation du nombre de classes demandé à la méthode des K -moyennes.	101
6.8	Composition du jeu d'apprentissage.	102
6.9	Résultats du classifieur bayésien normal pour différents lots de données.	103
6.10	Résultats de la méthode des forêts aléatoires pour différents lots de données et une résolution de 50 cm.	105
6.11	Résultats des SVM pour différents lots de données et une résolution de 50 cm ou 1 m.	107
6.12	Estimation de la portée des applications en fonction du degré de développement des méthodes.	113
A.1	Exemples of WCI with different targets. Seabed is represented as a dark horizontal line. Top : sailing yacht (wreck, mast and spreaders) ; middle : bubbles from diver(s) ; bottom : school of fishes.	116
A.2	Location of the "Les Pierres" study site (red frame) on the French Mediterranean coast in the La Ciotat bay, near the "Ile Verte" island.	117
A.3	Bathymetric 3D survey of the south-west part of the "Ile Verte" at La Ciotat bay (France), represented with a colour gradient. Red areas have a depth lower than 15m, yellow ones are between 15m and 30m depth, seabeds in blue are at least 45m depth. The zone corresponds to a 500m \times 500m area.	118
A.4	Data from the acquisition survey mapped. A : the acoustic backscatter map, "sonar image", the strength values are a no-dimension intensity. B : the computed roughness, with metric values. C : classification of the seabed nature.	119

A.5	Block diagram of the global data processing chain applied to MBES. The first block “Acquisition and formatting data” refers to the section A.1. The block depicted by a dotted red line represents the successive stages needed to exploit the WCI, whereas the blocks behind arrows are the two produced maps. *ROI : region of interest.	120
A.6	An example of water column ping including fishes (indicated by the two grey circles) and some bottom persistent interference. The bathymetric points are marked by green points. The two different minimum slant ranges (symbolised by a yellow arc) are drawn (calculation based on the SMSR (top) or the GMSR (bottom)). The WCI had been thresholded at -70 dB above the MSR. Detected targets appear in white (fishes and a part of the interference).	121
A.7	Left : fish accumulation areas on the site “Les Pierres” represented by the value of the maximal biomass of every ping (summer 2016 survey). Right : fish classification on the site “Les Pierres” categorising fishes detected during the summer 2016 survey in four classes (big or small individuals in large or small fish schools).	122
A.8	WCI exhibiting fish school for different values of the chosen threshold : 70 dB (top), 80 dB (middle), 85 dB (bottom).	124
B.1	Study site (red frame) in the Bay of La Ciotat.	129
B.2	MBES data products. A , Bathymetry ; B Backscatter imagery, C , Rugosity.	130
B.3	3D representation of the seabed topology with marine habitats and details of the three batches studied and the reefs names for the study of the fish density.	131
B.4	Temperature profile in June 2016 (black) ; August 2016 (red) ; June 2019 (blue).	132
B.5	Scatter plots representing fishes in A , June 2016 ; B August 2016 ; C , June 2019. Each blue dot represents a detected fish whatever it size.	133
B.6	Local fish densities in fishes/m ² in A , June 2016 ; B August 2016 ; C , June 2019. The red value corresponds to the mean number of fishes/m ²	133

Liste des tableaux

1.1	Exemples de pressions acoustiques extrêmes exprimées en pascal et en décibels. . .	8
1.2	Quelques coefficients approximatifs de l’amortissement α obtenu par application de la formule de François et Garrisson pour différentes fréquences utilisées en ASM. .	10
1.3	Répartition des applications issues de l’ASM selon leurs fréquences. En bleu, les applications avec des fréquences “basses”, en deçà de la dizaine de kHz. En jaune, la plage de fréquence la plus répandue dans les systèmes acoustiques. En orange, les fréquences les plus hautes, peu répandues. D’après Lurton (1998)	11
1.4	Exemples de bruits qui participent au bruit de fond. Les bruits sont répartis par fréquence et par force. En orange, les sources anthropiques; en vert, les sources biologiques et en bleu, les sources abiotiques.	12
3.1	Quantité approximative de sondes par maille selon leur emplacement dans le MNT.	46
3.2	Matrice de confusion ramené en pourcentages des bonnes et mauvaises sondes selon le filtrage manuel de référence et le filtrage automatique.	47
5.1	Diversité des méthodes d’apprentissage automatique utilisées sur des données acoustiques (et lidar) pour produire des cartes d’habitats marins. (UW : Université du Wisconsin-Madison). Méthode supervisée (S), méthode non-supervisée (NS) . . .	79
5.2	Principaux avantages et inconvénients des méthodes de partitionnement les plus connues.	83
5.3	Exemple de noyaux utilisés par les SVM	88
6.1	Temps de calcul d’une classification par la méthode des K -moyennes pour différentes résolutions.	100
6.2	Étude du temps de convergence des K -moyennes pour différentes valeurs de K . .	101
6.3	Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode du classifieur bayésien et une résolution de 50 cm	102
6.4	Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires, une résolution de 50 cm et les paramètres ($N_{arbre} = 200, prof_max = 2$)	104
6.5	Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires et une résolution de 50 cm avec une variation de la profondeur des arbres.	104
6.6	Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires et une résolution de 50 cm avec une augmentation du nombre d’arbre et une variation de la profondeur des arbres.	105
6.7	Temps de calcul, paramètres et exactitude des phases d’entraînement et de prédiction avec la méthode des SVM et une résolution de 50 cm ou 1 m.	106
6.8	“Transferts” d’apprentissage d’une plus basse à une plus haute résolution.	106
6.9	Bilan des types de données manipulées par chacune des applications traitées dans ce document.	112
6.10	Bilan des types de méthodes choisies pour proposer traiter chacune des applications.	113

A.1 Comparison of the B_{Av} and B_{Max} values versus the MSR variant. Values obtained on the WCI given in Fig. A.6	121
A.2 Various WCI targets discriminated according to their coordinates and intensity.	125
B.1 Areas covered by each marine habitats and their proportion.	131
B.2 Number of fish detections and proportions per depth categories.	132
B.3 Functional compartments of the Mediterranean ecosystems investigated by the coupling of marine habitat maps and fish detection in the water column.	135

Liste des notations

Sigles et acronymes

Ne sont listés que ceux qui servent plus d'une fois et par ordre d'apparition.

ASM	Acoustique Sous-Marine
SMF	Sondeur MultiFaisceaux
IA	Intelligence Artificielle
DL	<i>Deep Learning</i>
ML	<i>Machine Learning</i>
OEM	Onde Électro-Magnétique
OHI	Organisation Hydrographique Internationale
ASDIC	<i>Anti-Submarine Detection Investigation Commitee</i>
SONAR	<i>SOund Navigation And Ranging</i>
OA	Onde Acoustique
PT	Perte par transmission
NE	Niveau d'Émission
GD	Gain de Directivité
GT	Gain de Traitement
TS	<i>Target Strength</i>
NB	Niveau de Bruit
ID	Indice de Détection
TVG	<i>Time Varying Gain</i>
SBL	Sonar à Balayage Latéral
GNSS	Géolocalisation et Navigation par Système de Satellites
GPS	<i>Global Positionning Systems</i>
RTK	<i>Real Time Kinematic</i>
IGN	Institut National de l'information Géographique et forestière
RGP	Réseau GNSS Permanent
MNT	Modèle Numérique de Terrain
SIG	Système d'Information Géographique
WCI	<i>Water Colonne Image</i>
RNA	Réseau de Neurones Artificiels
MLP	<i>Multi-Layer Perceptron</i>
CNN	<i>Convolutional Neural Network</i>
FC	[couche de neurones] <i>fully connected</i>
ReLU	<i>Rectified Linear Unit</i>
RMSE	<i>Root Mean Square Error</i>
EUNIS	<i>European Nature Information System</i>
DCSMM	Directive-Cadre Stratégie pour le Milieu Marin
SVM	Support Vector Machine
AD	Arbres de Décisions
FA	Forêts Aléatoires d'arbres de décisions

Abréviations

Elles concernent des noms de fonctions mathématiques.

\log	logarithme en base 10
\log_2	logarithme en base 2
\sin	sinus
\cos	cosinus
\exp	exponentielle
\tanh	tangente hyperbolique
$\arg \min$	argument minimum
$\arg \max$	argument maximum

Symboles

De manière générale :

(X, Y, Z)	repère 3 dimensions
(x, y, z)	un point dans le repère précédent
t	temps
θ, ϕ	angles
N, n	nombre d'éléments, indice d'un élément
M, m	nombre de caractéristiques, indice d'une caractéristique
\mathbf{W}, \mathbf{w}	Matrice et vecteur de poids
\mathbf{B}, \mathbf{b}	Matrice et vecteur de biais

Liste des symboles relatifs aux équations d'acoustique sous-marine :

p	pression acoustique
v	célérité du son
T, S	température, salinité
κ	nombre d'onde
λ	longueur d'onde
ρ	masse volumique
χ	compressibilité d'un fluide
α	amortissement du milieu
j	nombre complexe
π	constante d'Archimède valant 3.14159...
f_0	fréquence fondamentale
S_u	surface
I_t, P_w	intensité, puissance
D_{ir}	fonction de directivité
Q	facteur de directivité
A_m, B_p	amplitude, bande passante
τ	durée d'impulsion
S_p	spectre d'un ping
m_{od}	modulation
a	capteur d'une antenne
R	distance entre deux objets
δt	retard temporel
Ω	onde résultante

Liste des symboles relatifs aux équations des réseaux de neurones :

q	paramètre quelconque
η	taux d'apprentissage
ψ	fonction d'activation
$\mathbf{d}, d_1, d_2, \dots$	vecteur d'une donnée et ses caractéristiques
\mathbf{E}, \mathbf{e}	matrice et vecteur d'étiquettes de données
\mathbf{O}, \mathbf{o}	matrice et vecteur de la couche de sortie du réseau
\mathbf{h}	vecteur d'une couche cachée du réseau
\mathbf{c}	vecteur d'une couche quelconque du réseau
ℓ	fonction de perte
i	indice d'une couche
r, s	indices sur des neurones
g	gradient de la fonction de perte
u	itération de l'apprentissage
ζ, v	réglages de la méthode de descente de gradient adaptatif

Liste des symboles relatifs aux équations du machine learning :

f, g	fonctions quelconques
A_0, \dots, A_6	coefficients d'une fonction linéaire
$S_{x,y}$	somme croisée
K, k	nombre de classes et indice sur les classes
\mathcal{D}	distance au sens d'une norme ou d'une projection orthogonale
Π	partition ou classes de données
μ, σ	moyenne et écart type des données
ν_0, \dots, ν_N	paramètres lagrangiens
H_p	hyperplan
ξ	variable ressort
ε	variable de régulation
γ, β	taille et biais d'un noyau
U, V	événements
Σ_k	matrice de covariance

Opérateurs

L	lagrangien
\mathbb{R}	ensemble des nombres réels
\mathbb{N}	ensemble des entiers naturels
$f \circ g$	fonctions composées
$f'(x)$	dérivée de f selon x
∂	opérateur différentiel
Δ	opérateur Laplacien
$\sum_{n=1}^N$	somme sur les éléments de 1 à N
$\{\dots\}$	ensemble d'éléments
$ \dots $	valeur absolue
$\ \dots\ _{1,2}$	norme 1 ou 2
$\mathbf{w} \cdot \mathbf{x}$	produit scalaire des vecteurs \mathbf{w} et \mathbf{x}
\mathbf{w}^T	transposée du vecteur \mathbf{w}
$P(U)$	probabilité de l'évènement U
$P(U \wedge V)$	probabilité conjointe de U et V
$P(U V)$	probabilité conditionnelle de U sachant V

Introduction Générale

Contexte

Cette thèse CIFRE¹, financée dans le cadre du contrat ANRT² 2018/0588, a été réalisée au sein de l'entreprise Seaviews (située à La Ciotat) en partenariat avec le Laboratoire d'Informatique et des Systèmes, UMR CNRS 7020 (LIS) de l'Université de Toulon.

Cette entreprise est un bureau d'étude en environnement marin utilisant essentiellement des outils d'acoustique sous-marine (ASM) et de photogrammétrie pour proposer des services à ses clients. En parallèle de ces services et dans l'optique de pouvoir en proposer de nouveaux plus innovants, une part de l'activité de Seaviews est également consacrée à de la recherche et à du développement de solutions logicielles, de méthodes de traitement, d'analyse des données ainsi qu'à la mise au point de nouveaux matériels d'acquisition (Fig. 1). Pour cela, possédant ses propres instruments et navire, Seaviews maîtrise toute la chaîne de mesure afin de réaliser des acquisitions. Elle développe également ses propres suites logicielles lui permettant de traiter les données acquises et d'aller jusqu'à l'édition de produits (cartographies, rapports, plans, etc.).

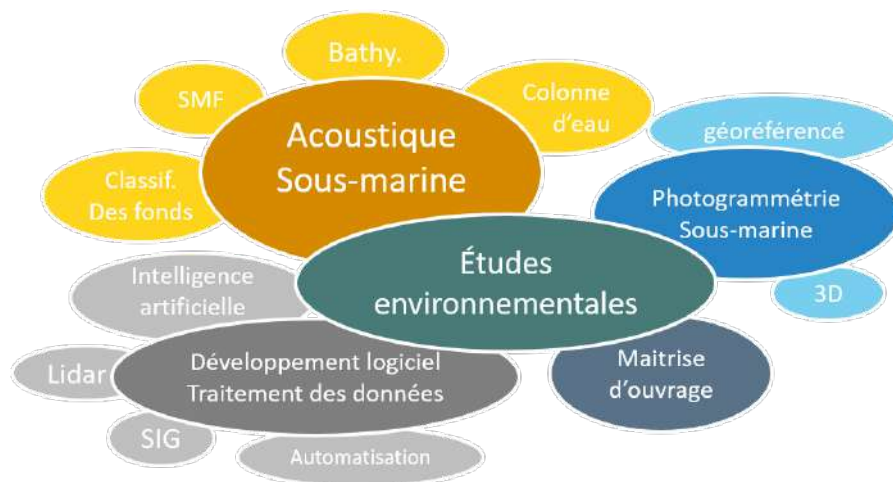


FIGURE 1 – Principaux mots clés pour qualifier les activités de la société Seaviews.

Les travaux de cette thèse répondent à des besoins d'amélioration de ces logiciels par une plus grande automatisation de traitements réalisés jusqu'alors de manière manuelle ou semi-automatique. L'automatisation doit permettre d'une part, d'aller plus loin dans l'analyse des données et dans l'exploitation des résultats et d'autre part, de rendre les traitements plus robustes et plus efficaces afin de les proposer comme de nouveaux services. Le LIS vient épauler ces évolutions par ses compétences en signal-image et en sciences des données. En particulier, les besoins se concentrent sur l'automatisation du traitement des données issues de sondeurs acoustiques multifaisceaux dans le cadre d'applications hydrographiques.

1. Convention Industrielle de Formation par la Recherche
2. Association Nationale de la Recherche et de la Technique

Les sondeurs multifaisceaux (SMF) sont des systèmes acoustiques capables de couvrir efficacement, par insonification, de très larges surfaces de fonds marins. Depuis leurs premières commercialisations dans les années 1970, la quantité d'information acquise n'a cessé d'augmenter en plus de se diversifier. D'appareils à mesure des profondeurs (la bathymétrie), ils sont progressivement devenus des instruments capables d'imager les fonds (par réflectivité) et la colonne d'eau. Ils jouent dorénavant un rôle important aussi bien en hydrographie, qu'en halieutique, qu'en ingénierie côtière et offshore ou enfin qu'en sciences environnementales. Si les temps de levés en mer ont été considérablement réduits grâce à l'utilisation des SMF, les temps de traitement se sont en revanche considérablement allongés. En outre, la manipulation des données requiert un personnel expert qui lors des analyses n'est pas à l'abri d'erreurs et d'interprétation subjective. Du fait de ces trois raisons - temps, erreurs et subjectivité -, il est donc intéressant de développer des méthodes qui tendent vers une plus grande automatisation.

Aujourd'hui, le terme d'automatisation peut évoquer deux idées. La première est un processus qui a lieu sans la moindre intervention humaine. L'outil opère tous les cas de figure sans changement de réglages car le travail à réaliser s'est standardisé. Dans ce cas-là, l'opérateur n'a plus besoin d'être expert. Il s'assure surtout du bon déroulement des opérations. Dans notre cas, chaque expérimentation en mer est unique et c'est plutôt l'outil qui doit s'adapter avec le minimum d'intervention humaine. L'opérateur doit cependant rester expert pour la validation des résultats après traitements. Cette deuxième idée d'outil qui s'adapte intelligemment à la tâche à faire, fait écho à de "l'informatique cognitive", une discipline à mi-chemin entre l'intelligence artificielle (IA), les sciences des données et les sciences cognitives. Ce domaine emploie des outils issus de l'apprentissage automatique (*Machine Learning*, ML), de l'apprentissage profond (*Deep Learning*, DL), mais aussi de raisonnement et de logique pour s'atteler à des tâches mettant en œuvre de grandes quantités de données éventuellement non-structurées, avec des hétérogénéités, des ambiguïtés, des évolutions, voire des données manquantes dans un but d'assister un opérateur dans la prise de décision.

Contributions

Durant ces trois années de thèse, trois applications majeures ont été abordées au fur et à mesure des besoins de l'entreprise. Parmi celles-ci, le traitement des données bathymétriques pour de la cartographie et de la classification des habitats marins est présenté dans ce document (voir section suivante). L'étude de l'activité biologique dans l'imagerie colonne d'eau a donné lieu à deux articles et à des présentations orales, il a donc été choisi de ne pas la redévelopper ici ; les articles publiés sont fournis en annexes [A](#) et [B](#) et les résultats sont discutés dans la conclusion générale.

Ces thématiques ont impliqué d'intégrer de nouvelles méthodes dans les logiciels de Seaviews, d'en modifier significativement certaines, voire de repenser le processus de traitement des données. Les nouveautés viennent de "l'informatique cognitive" : méthode statistique basée sur des règles expertes et sur de l'expérience pour du traitement bathymétrique, méthodes issues du *machine learning* pour la classification des habitats, méthodes à base de réseaux de neurones et d'apprentissage profond pour la proposition d'un traitement bathymétrique alternatif. La méthode pour les études de l'activité dans la colonne d'eau a subi un important remodelage ; essentiellement manuelle (et très longue), le processus s'est automatisé à travers un enchaînement de traitements (seuillage, détection, segmentation, identification) à paramétrer manuellement avant de démarrer les calculs qui se font en tâche de fond. La lecture de ce manuscrit peut donc se faire sous deux angles de vue : celui de l'hydrographe pour les aspects pratiques d'utilisation des données et d'interprétation des résultats obtenus, et celui du "data scientist" pour le choix et la mise en application de méthodes sur des données expérimentales réelles issues de campagnes d'acquisition en mer.

Les méthodes nouvellement intégrées sont comparées en termes d'efficacité (robustesse vis-à-vis d'erreurs, vitesse de traitement, etc.) et en termes d'intelligibilité des résultats par rapport à ceux obtenus au moyen des méthodes déjà en place dans les logiciels. Elles sont accessibles par l'ensemble de l'équipe de Seaviews qui utilise tous les jours ces logiciels. Elles permettent soit un gain de temps dans les traitements, soit une analyse plus avancée des données, soit une exploitation de données plus complexes.

Organisation du document

Ce document est organisé en trois parties indépendantes.

La Partie I a pour but d'introduire les connaissances générales nécessaires pour aborder chacune des parties suivantes. Pour cela, le chapitre 1 permet de rappeler les notions utiles en acoustique sous-marine pour poser l'équation du sonar actif, base de tout système acoustique en milieu aquatique. Le chapitre 2 est entièrement dédié au SMF. Le début s'attarde sur les spécificités de son fonctionnement, reposant sur la formation de voies et les faisceaux croisés. La suite reprend la chaîne de mesure complète au travers de la présentation du matériel utilisé par la société Seaviews : un SMF ne peut être utilisé seul, il est associé à d'autres instruments qui lui confèrent toute sa précision. Enfin, les types de données acquises et les principales applications sont rappelées et détaillées. Cette présentation introduit finalement les trois parties qui suivent.

La Partie II concerne le traitement des données bathymétriques. La bathymétrie est la mesure de la profondeur de l'eau et son acquisition est la principale utilisation d'un SMF afin de concevoir, par exemple, des cartes bathymétriques. L'un des traitements les plus répandus, car nécessaire, est l'élimination des erreurs présentes dans ces données. Le chapitre 3 fait un état des lieux des méthodes qui ont été proposées pour effectuer cette tâche ainsi que l'existant dans les logiciels de Seaviews. Dans la suite du chapitre, une méthode complémentaire d'élimination de ces fausses sondes est proposée et intégrée dans les logiciels. Cette méthode emprunte des éléments du traitement statistique pour les appliquer au domaine spatial où les données sont projetées dans un modèle numérique de terrain - format cartographique adapté à la manipulation des données géographiques.

Dans le chapitre 4, nous expérimentons une nouvelle manière d'accéder à la donnée bathymétrique en "évitant les erreurs régulières de mesures". Nous choisissons d'exploiter les données d'imagerie de la colonne d'eau. Dans celles-ci, le fond est clairement visible et évident pour un humain, mais la grande variabilité de ce type d'images rend l'estimation du fond plus difficilement réalisable par des machines utilisant des méthodes de détection classiques. C'est pourquoi, nous proposons l'utilisation des méthodes issues du *deep learning* au travers des réseaux de neurones artificiels. Ces réseaux sont entraînés avec des images de la colonne d'eau dans lesquelles nous avons préalablement correctement étiqueté la position du fond. Plusieurs jeux de données expérimentales ont été composés et l'étiquetage a été réalisé de manière automatique. L'approche proposée est innovante car elle exploite les données de la colonne d'eau relativement peu acquises et exploitées et repose sur l'utilisation du DL, technique en plein essor mais assez peu mature au niveau certaines applications.

La Partie III est consacrée à la détermination de la nature des fonds et plus particulièrement à l'identification et à la classification des habitats marins. A Seaviews, les développements techniques ont porté sur le traitement des données pour s'intéresser à la morphologie des fonds en complément de l'imagerie sonar et, jusqu'alors, la méthode de classification était manuelle et chronophage. Ces travaux sont résumés au chapitre 5. Un premier besoin portait sur l'amélioration de la technique de classification. Pour cela, le chapitre 5 introduit donc nos travaux d'intégration de méthodes relevant du *machine learning*, c'est-à-dire des méthodes automatiques à apprentissage artificiel. Puisque les méthodes de ML sont des techniques plus éprouvées que celles du *deep learning*, une sélection est faite parmi les plus réputées et permet ainsi d'aborder plusieurs types d'apprentissages, supervisés et non-supervisés, et plusieurs approches, telles que l'approche logique et l'approche statistique. Un second besoin portait sur l'extension de ces méthodes à des données complémentaires comme l'imagerie aérienne et les mesures bathymétriques par lidar aéroporté. Ces données ne sont opérationnelles qu'en tout petits fonds, mais sont pertinentes pour ces applications car leurs résolutions s'améliorent et leur disponibilité s'accroît. Elles sont exploitées dans le chapitre 6 où une zone d'étude est choisie pour tester les méthodes d'apprentissage sur les données acoustiques couplées aux données optiques et évaluer les performances des différents traitements.

Enfin nous terminerons ce manuscrit par une conclusion où seront récapitulés les travaux réalisés et par des perspectives quant aux suites possibles à ce travail.

Publications

Revue internationale

C. Viala, M. Lamouret, A. Abadie, “Seafloor classification using a multibeam echo sounder : A new rugosity index coupled with a pixel-based process to map Mediterranean marine habitats”. *Applied Acoustics* (Elsevier, JCR, IF 2.639), Vol. 179, August 2021, 10.1016/j.apacoust.2021.108067. [Lien vers l'article](#).

M. Lamouret, P. Boissery, C. Viala, N. Thirion-Moreau, A. Abadie, “Coupling the map of marine habitats and fish accumulation zones : a three dimensional spatial approach for the management of halieutic resources”, *Vie et Milieu - Life and Environment*, 2020, 70 (3-4), pp. 175-182. Voir Annexe B ou le [lien vers l'article](#).

Conférences internationales

M. Lamouret, A. Abadie, P. Boissery, N. Thirion-Moreau and C. Viala, “Coupling the map of marine habitats and fish accumulation zones : a three dimensional spatial approach for the management of halieutic resources”, in *Proc. GECOMARS 2020*, Marseille, France, 4-5 February 2020. [Lien vers le support de présentation](#).

M. Lamouret, A. Abadie, N. Thirion-Moreau, P. Boissery and C. Viala, “Measuring fish biomass as additional environmental data during a hydrographic survey with a multi-beams echo-sounder”, in *Proc. IEEE International Conference OCEANS*, pp. 1-10, Marseille, France 17-20 June 2019. Voir Annexe A ou le [lien vers l'article](#) ainsi que [lien vers le poster](#).

Conférences nationales

M. Lamouret, A. Abadie, P. Boissery, N. Thirion-Moreau, C. Viala, “Cartographie simultanée des ressources halieutiques et des habitats marins avec un sondeur multifaisceaux”, colloque merIGéo 2020 (De l’océan à la côte : l’information géographique en mouvement), Bordeaux, France, mars 2020 reporté du 25 au 26 novembre 2020. [Lien vers le support de présentation](#).

M. Lamouret, A. Abadie, C. Viala, P. Boissery, “Cartographie des ressources halieutiques avec des données d’opportunités de sondeur multifaisceaux”, colloque Journées Techniques de l’AFHy, Le Havre, France, 26 et 27 novembre 2019. [Lien vers le support de présentation](#).

Rapports d’études

C. Viala, M. Lamouret, “Bref état de l’art de l’instrumentation acoustique halieutique d’une bouée”, Rapport d’études. Seaviews. 20p. 2020.

Encadrement

Co-encadrement du stage de fin d’études de cycle ingénieur et de Master 2 de Alice Nicolas de mars à août 2021, ayant donné lieu à la rédaction d’un mémoire de fin d’études : “Automatisation de la classification des fonds marins à partir de données multiples (sondeur multifaisceaux, lidar, données d’opportunité)”.

Première partie

Généralités

Introduction

Que cela soit dans le milieu terrestre ou spatial, une place prépondérante est donnée aux ondes électromagnétiques (OEM). Nous percevons le monde qui nous entoure d’abord par nos yeux, grâce à la lumière visible. Et, avec différentes technologies, nous exploitons un plus large spectre : ondes radio en télécommunications, micro-ondes dans nos cuisines, rayons X en médecine, rayons gamma en astrophysique, etc. Toutefois, ces ondes sont inopérantes pour nous permettre d’appréhender le milieu marin. A cause de la forte conductivité de l’eau salée, les OEM sont atténuées très rapidement sous la surface de l’eau, et ne pénètrent pas plus loin qu’une quarantaine de mètres au grand maximum. Passé une centaine de mètres, l’ambiance est crépusculaire. Le son, en revanche, se propage admirablement mieux sous l’eau que dans l’air. En effet, l’eau offre un support de propagation propice à ces vibrations mécaniques, le son se transmet plus vite et plus loin. En somme, les sons compensent “nos yeux” sous la surface.

Cela fait maintenant un siècle que la mer est sur écoute. L’acoustique sous-marine (ASM) a pris une place primordiale dans l’observation, la compréhension et l’exploitation des mers. Les applications de l’ASM peuvent se ranger en quatre grandes classes d’activités :

- Détection et localisation de cibles (ou d’obstacles). Les cibles peuvent prendre tous types de formes, qu’il s’agisse de sous-marins, de drones, de mines dans le domaine militaire, de poissons dans le domaine de la pêche, de réserves de pétrole ou de gaz souterraines dans le domaine de la sismique et de la géophysique interne, etc.
- Estimation et mesure de grandeurs caractéristiques : des fonds (profondeur, reliefs, nature des fonds), du milieu marin (courants, couches d’eau), du vivant (suivi et surveillance des cétaqués, estimation de biomasse), ou encore de cibles diverses (vitesse, dimensions), etc.
- Transmission de l’information. Cela peut prendre la forme de communications entre navires, du recueil de données acquises, de téléopération de robots sous-marins.
- Prédiction et prévention de risques (corrosion des navires, fissuration de pipelines, etc.)

Cette thèse se concentre sur des travaux de mesures, par ASM, de la profondeur et de la nature des fonds et des activités dans la colonne d’eau. Ces mesures sont réalisées dans des objectifs de cartographie du milieu marin à destination de la sécurité de navigation des navires ou à destination d’études environnementales. Ces activités appartiennent au domaine de l’hydrographie, plus ancien que l’ASM, dont la définition actuelle de l’Organisation Hydrographique Internationale (OHI) est :

“L’hydrographie est la branche des sciences appliquées traitant du mesurage et de la description des éléments physiques des océans, des mers, des zones côtières, des lacs et des fleuves, ainsi que de la prédiction de leur changement dans le temps, essentiellement dans l’intérêt de la sécurité de la navigation et pour l’appui de toutes les autres activités maritimes, incluant le développement économique, la sécurité et la défense, la recherche scientifique et la protection environnementale.”

L’évolution rapide des technologies au XXe siècle a permis à différents systèmes acoustiques d’exister et de révolutionner irrémédiablement l’hydrographie en élargissant son champ d’activités. L’effet piézoélectrique a trouvé sa première application grâce à Paul Langevin, à Constantin Chilowsky et à la Marine de Toulon qui mettent au point, à la fin de la Première Guerre mondiale, un détecteur de sous-marins par ultrasons, l’ASDIC (Anti-Submarine Detection Investigation Committee). Durant l’entre-deux-guerres, cette même équipe transforme l’ASDIC en système de sondage

ultrasonore pour mesurer les profondeurs des mers (Lelong; 2001). La Royal Navy s'équipe lourdement en ASDIC et joue un rôle important lors de la Bataille de l'Atlantique (1939-1945) par sa recherche active de sous-marins allemands. Avec l'entrée en guerre des États-Unis (1941) lors de la Seconde Guerre mondiale, les ASDIC changent de nom pour devenir des SONAR (SOund Navigation And Ranging) et de considérables recherches sont réalisées sur la propagation des ondes sonores en milieu aquatique ainsi que sur les performances des systèmes d'acoustique actifs et passifs. L'effort de recherche est maintenu durant toute la Guerre Froide, notamment en acoustique passive pour la détection des sous-marins américains et russes (Lurton; 1998). Du côté de l'acoustique active, les sonars à balayage latéral pour l'archéologie, la recherche d'épaves et de mines sont mis au point dans les années 1950 et commercialisés dans les années 1970. En parallèle, dans les années 1960 démarrent les recherches sur les sondeurs multifaisceaux (SMF) (Tucker; 1961); les sondeurs permettant de mesurer très précisément un grand nombre de sondes bathymétriques sur de larges fauchées. Autour du même concept, un autre M. Tucker, Tucker (1960) propose d'augmenter la précision de la mesure bathymétrique en affinant les faisceaux (augmentation de la directionnalité). Cependant, la position des pieds des faisceaux fins a besoin d'être connue alors que le navire est en mouvement sur la mer. Tucker and Hendersen (1960) proposent alors une stabilisation électronique de l'émission et de la réception des faisceaux, plutôt qu'un montage mécanique de stabilisation du sondeur. Enfin, le premier système de SMF pour la mesure des grands fonds fait son apparition en 1970 (Glenn; 1970). Des SMF pour de plus faibles profondeurs se commercialisent rapidement, le Bo'Sun (Burke and Robson; 1975) pour des fonds de moins de 1000 mètres et le Bathymetric Swath Survey System (McCaffrey; 1981) pour de la bathymétrie portuaire. Les capacités de calcul et la miniaturisation de l'électronique à partir des années 1980 permettent encore l'amélioration des SMF. Ils acquièrent plus de sondes, insonifient les plus petits fonds, les faisceaux s'affinent, la cartographie des données s'accélère. Depuis, les SMF sont également capables d'acquérir, en même temps que la bathymétrie, une imagerie de réflectivité du fond de qualité comparable à celle des sondeurs à balayage latéral et l'imagerie de la colonne d'eau à très haute résolution.

Fonctionnels dans tous types de profondeurs et répondant à de nombreuses applications du fait de la diversité des données acquises, les SMF sont devenus de véritables couteaux suisses de l'hydrographie moderne. Ils sont, cependant, plus complexes à mettre en oeuvre que les systèmes plus anciens car ils nécessitent une instrumentation associée (positionnement par satellites, célérimètre, centrale inertielle, etc.), un stockage numérique pour les grandes quantités de données qu'ils permettent d'acquérir et un personnel expert pour le traitement et l'exploitation des données.

Un certain nombre de connaissances sont nécessaires pour bien comprendre les données issues des SMF qui seront manipulées par la suite, c'est pourquoi cette première partie se décompose en deux chapitres. Le [chapitre 1](#) regroupe les généralités sur l'ASM, parmi lesquelles, les notions de propagation des ondes sonores et les principes de base des systèmes acoustiques, ainsi qu'une présentation de mesures à l'aide de deux systèmes classiques. Le [chapitre 2](#) se concentre, quant à lui, sur les sondeurs multifaisceaux, aussi bien en terme de fonctionnement que des données qu'ils permettent d'acquérir.

Chapitre 1

Notions de systèmes acoustiques sous-marins

Le principe fondamental d'un système acoustique utilisé en milieu aquatique repose sur "l'équation du sonar". Cette formule met en jeu des éléments d'acoustique sous-marine et de traitement d'antenne, éléments qui sont rappelés dans ce chapitre. Puis, deux principaux sondeurs sont présentés pour introduire le SMF dans le chapitre suivant.

1.1 Propagation des ondes acoustiques en milieu marin

Afin d'appréhender les notions utiles de propagation des ondes acoustiques (OA) sous l'eau, ce chapitre est divisé en trois parties. La première est consacrée à des généralités sur les OA. La seconde s'intéresse au choix des fréquences des OA utilisées en fonction d'une application donnée et à l'impact de la propagation de cette onde en milieu marin. Enfin, la dernière partie se concentre sur les trajectoires complexes de ces ondes dans le milieu marin.

1.1.1 Caractéristiques générales des ondes acoustiques

Pression acoustique Une onde acoustique est induite par une vibration mécanique. Elle prend la forme d'un mouvement de compression-dilatation qui se transmet d'élément à élément. Ces éléments font partie d'un ensemble, le support de propagation de l'onde, qui possède une certaine élasticité. Cette vibration n'induit pas de transport de matière mais un mouvement sur place des éléments autour de leur position d'équilibre. Ce mouvement particulière se caractérise par sa vitesse et par son amplitude ou *pression acoustique* p mesurée en pascal (Pa).

Du fait de la dynamique des pressions rencontrées (Tab. 1.1), une échelle logarithmique est plus souvent employée que le pascal. La conversion en décibels (dB) s'obtient par le logarithme décimal du rapport de la pression avec une pression de référence, p_{ref} (Lurton; 1998) :

$$p_{dB} = 20 \log \left(\frac{p}{p_{ref}} \right) \quad \text{avec} \quad p_{ref} = \begin{cases} 1 \mu\text{Pa} & \text{dans l'eau} \\ 20 \mu\text{Pa} & \text{dans l'air} \end{cases}$$

Types de valeurs	Minimale	Maximale
Exemples	Faible bruit de fond en mer	Clic de cachalot longue portée
Valeur en dB	20	230
Valeur en μPa	10	$3.1 \cdot 10^{11}$

TABLE 1.1 – Exemples de pressions acoustiques extrêmes exprimées en pascal et en décibels.

Équation d'onde Une onde acoustique se propageant dans un gaz ou un liquide obéit aux lois de la mécanique des fluides. On peut appliquer l'équation de d'Alembert sur la pression acoustique p d'une onde acoustique dans le milieu caractérisé par un repère cartésien (x, y, z) en fonction du temps t (Lurton; 2002) :

$$\Delta p = \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} = \frac{1}{v^2(x, y, z)} \frac{\partial^2 p}{\partial t^2}$$

avec $v(x, y, z)$, la célérité locale de l'onde, Δ l'opérateur Laplacien et ∂ l'opérateur différentiel.

On cherche les solutions harmoniques de l'équation d'onde à partir d'une onde sinusoïdale de fréquence fondamentale f_0 , de type $p(x, y, z, t) = p(x, y, z) \exp(2j\pi f_0 t)$. On obtient alors l'équation de Helmholtz :

$$\Delta p + \kappa^2(x, y, z)p = 0 \text{ avec le nombre d'onde } \kappa(x, y, z) = \frac{2\pi f_0}{v(x, y, z)}.$$

En considérant une source ponctuelle p_0 dans un milieu isotrope, la célérité locale devient une constante, $v(x, y, z) = v$ et l'équation de d'Alembert admet alors comme solution une onde sphérique d'équation :

$$p(R, t) = \frac{p_0}{R} \exp\left(j2\pi f_0 \left(t - \frac{R}{v}\right)\right) \quad (1.1)$$

où R est la distance sphérique à la source. L'amplitude de pression de l'onde décroît en $\frac{1}{R}$.

Intensité et puissance L'intensité acoustique I_t est le flux d'énergie (en W/m^2) passant un point donné, obtenue par le produit de la pression acoustique par la vitesse particulaire. La puissance P_w , mesurée en W , correspond à l'intensité reçue par une surface S_u , soit (Lurton; 1998) :

$$I_t = \frac{p_0^2}{2\rho v} \quad \text{et} \quad P_w = \frac{p_0^2 S_u}{2\rho v}$$

avec ρ la masse volumique du milieu.

Pertes par divergence géométrique Avec la propagation sphérique de l'onde, l'énergie acoustique s'étale sur une surface S_u qui croît avec le carré de l'éloignement R , $S_u = 4\pi R^2$. L'énergie se conservant, c'est donc l'intensité qui diminue avec l'augmentation de cette surface. En s'intéressant aux intensités I_{t_1} et I_{t_2} d'une source ponctuelle p_0 reçues respectivement aux distances R_1 et R_2 , avec $R_1 < R_2$, le rapport de ces intensités donne :

$$\frac{I_{t_2}}{I_{t_1}} = \left(\frac{R_1}{R_2}\right)^2.$$

Ainsi, la pression acoustique décroît en $1/R$, tandis que l'intensité acoustique décroît en $1/R^2$.

Ces pertes sont souvent aussi exprimées en décibels par les pertes de transmissions, PT, en utilisant une distance R et une distance de référence R_0 :

$$PT = 20 \log\left(\frac{R}{R_0}\right) \quad \text{où} \quad R_0 = 1 \text{ m.}$$

1.1.2 A chaque fréquence, son application

Pertes par amortissement En plus des pertes d'intensité dues à la divergence géométrique de la propagation sphérique de l'onde, il faut tenir également compte du caractère dissipatif du milieu. Une grande part de l'intensité de l'onde sonore est absorbée par la viscosité du milieu et par des réactions chimiques (type relaxations de molécules). La pression acoustique décroît de façon exponentielle avec la distance parcourue et l'équation d'onde sphérique Eq. 1.1 devient :

$$p(R, t) = \frac{p_0}{R} \exp(-j\alpha' R) \exp\left(j2\pi f_0 \left(t - \frac{R}{v}\right)\right). \quad (1.2)$$

Le paramètre α' quantifie l'absorption par le milieu mais n'est pas utilisé en tant que tel. On lui préfère le paramètre α , en dB/m ou dB/km, lié à α' par $\alpha = 20\alpha' \log(e)$. Son intérêt permet de facilement considérer la distance maximale que peut parcourir l'onde avant d'être trop atténuée.

Le paramètre α peut être évalué empiriquement avec l'onde (la fréquence) et le milieu (température, salinité, pression hydrostatique). Mais plusieurs formules ont été établies pour lier α à la fréquence de l'onde selon certaines validités (salinité, profondeur, gamme de fréquence, etc.). Celle de [Francois and Garrison \(1982a\)](#) et ([Francois and Garrison; 1982b](#)), est valable pour des fréquences de 10 à 500 kHz dans de l'eau à température ne dépassant pas les 20°C et à salinité comprise entre 30 et 35 ‰ (Tab. 1.2). Quelles que soient les formules, plus les fréquences sont élevées, plus les ondes sont fortement amorties par le milieu.

Fréquences (kHz)	10	38	70	100	200	500
α (dB/km)	1	12	25	34	52	120

TABLE 1.2 – Quelques coefficients approximatifs de l'amortissement α obtenu par application de la formule de François et Garrison pour différentes fréquences utilisées en ASM.

Bilan des pertes lors de la propagation Ces deux types de pertes d'intensité lors de la transmission - pertes géométriques et pertes par amortissement -, se cumulent. Pour un signal effectuant un aller-retour sur une cible à distance R de la source, les pertes de transmission sont alors données par :

$$2PT = 40\log(R) + 2\alpha R.$$

Fréquences utilisées, portées approximatives et applications A l'instar des gammes de fréquences (ou longueur d'ondes) utilisées par les OEM, un large registre de fréquences est également utilisé par les nombreuses applications de l'ASM. Les fréquences s'échelonnent du Hz au MHz, ce qui correspond à des longueurs d'onde comprises entre environ 1.5 km (à 1 Hz) et 1.5 mm (à 1 MHz). Si les systèmes acoustiques à hautes fréquences voient leurs espaces d'investigation réduits à cause de l'amortissement, il faut rappeler que plus hautes sont les fréquences, meilleures sont les résolutions des signaux et plus finement sont insonifiés les objets sur le passage de l'onde. A l'inverse, les systèmes utilisant des basses fréquences ne sont que peu limités dans l'espace mais ne permettent pas la collecte précise d'information sur les objets plus petits que la longueur d'onde utilisée.

Ainsi, à chaque fréquence correspond donc un type particulier d'application. Jusqu'aux fréquences d'environ 10 kHz, celles-ci concernent des activités civiles et militaires ayant besoin de plusieurs dizaines de kilomètres de portée :

- Les recherches fondamentales en océanographie et en géophysique pour une meilleure compréhension du fonctionnement des océans et de la planète,
- Les activités de prospection pétrolière dans les profondeurs des sous-sols (utilisation d'acoustique dite "sismique"),
- Les systèmes acoustiques militaires pour la surveillance des mers du globe,
- Les activités d'étude des premières couches des sous-sols marins en vue de la préparation de chantiers (en vue d'implantation de plate-forme off-shore par exemple),

Au-delà de la dizaine de kHz, les portées sont de l'ordre de la centaine de mètres au kilomètre. Les systèmes fonctionnant dans ces plages servent en :

- Hydrographie avec les sondeurs monofaisceau et multifaisceaux et les sonars latéraux,
- Halieutique avec les sondeurs et sonars d'aide à la pêche,
- Communications et positionnement sous-marin (à la place des OEM inadaptées),
- Imagerie acoustique avec des "caméras acoustiques" pour surveiller des cages d'aquacultures par exemple.

Un récapitulatif des applications par gammes de fréquences est proposé dans le tableau 1.3. Les travaux de cette thèse sont réalisés avec un sondeur hydrographique de fréquences pouvant être choisies entre 170 et 450 kHz.

TABLE 1.3 – Répartition des applications issues de l’ASM selon leurs fréquences. En bleu, les applications avec des fréquences “basses”, en deçà de la dizaine de kHz. En jaune, la plage de fréquence la plus répandue dans les systèmes acoustiques. En orange, les fréquences les plus hautes, peu répandues. D’après [Lurton \(1998\)](#)

Fréquences (en kHz)	0.1	1	10	100	1000
Distances (en mer, en km)	1000	100	10	1	0.1
Sismique, Géophysique					
Océanographie acoustique					
Activités militaires					
Sondeurs de sédiments					
Communications					
Halieutique					
Bathymétrie					
Caméra acoustique					

Un milieu bruyant Enfin, il faut aussi tenir compte du bruit ambiant et de bruits ponctuels présents à toutes les fréquences d’écoute (tableau 1.4). Puisque le son se propage facilement et que le milieu est plein de bruiteurs, le bruit est un véritable pot-pourri où intervient, entre autres :

- Le bruit du trafic maritime et le bruit des activités humaines (dragage, forage, activité acoustique, etc.). Il est surtout présent près des ports et des grandes lignes maritimes. Avec des fréquences comprises dans des ordres de grandeur allant de 10 Hz à 1kHz. Il perturbe les systèmes acoustiques très basses fréquences,
- L’agitation de surface de la mer en fonction du vent et/ou de la pluie. Cela touche une large plage de fréquences utiles, de l’ordre de 100 Hz à 10 kHz,
- Les organismes vivants, qu’ils soient petits ou géants. Les cétacés sont capables de produire des sons de l’ordre de 10 Hz à 200 kHz. Ils restent cependant des sources ponctuelles de bruit,
- Le bruit thermique dû à l’agitation moléculaire. Il concerne les fréquences au-delà de 100 kHz.

Pour simplifier les modèles d’ASM, il est possible de modéliser le bruit ambiant comme suivant une loi gaussienne, mais ce choix est peu réaliste. En général, le niveau moyen de bruit est variable dans le temps et dans l’espace avec des bruiteurs impulsifs.

Le sondeur hydrographique de Seaviews, travaillant dans la gamme 170-450 kHz, est peu gêné par les bruits. Le bruit thermique est de faible niveau par rapport au niveau des signaux du sondeur. La présence d’autres sondeurs de fréquences proches peuvent toutefois affecter la qualité d’émission-réception lors de levés.

1.1.3 Rebonds dans les chemins de propagation des ondes acoustiques

Plusieurs éléments de propagation des ondes sont à combiner pour comprendre les chemins empruntés sous l’eau. Il faut, d’une part, tenir compte de la célérité du son qui évolue avec les variations du milieu et d’autre part, les frontières du milieu (surface et fond) qui agissent comme des miroirs.

La célérité des ondes acoustiques dépend des caractéristiques du milieu dans lequel elles se propagent. La célérité c peut être déduite connaissant la masse volumique ρ et la compressibilité χ d’un milieu fluide, $v = \sqrt{1/\chi\rho}$ ([Etter; 1998](#)). Mais la célérité dans l’eau de mer est plus généralement calculée avec des formules empiriques faisant intervenir la température, la salinité et la pression à travers la profondeur. Une des formules est proposée par [Clay and Medwin \(1977\)](#) :

$$v = 1449.2 + 4.6T(z) - 0.055T^2(z) + 0.00029T^3(z) + (1.34 - 0.001T(z))(S - 35) + 0.016z \quad (1.3)$$

TABLE 1.4 – Exemples de bruits qui participent au bruit de fond. Les bruits sont répartis par fréquence et par force. En orange, les sources anthropiques ; en vert, les sources biologiques et en bleu, les sources abiotiques.

Fréquences (kHz) Niveau de bruit (dB)	0.1	1	10	100	1000
250	Activités terrestres				
200	Acoustique sismique Dragage, Forage, ..			Sondeurs hydrographiques	
150	Baleines				
100	Sous-Marins		Dauphins		
		Pluies	Poissons, crevettes		
50	Bruit de navigation	Bruit des vagues		Bruit thermique	

où z est la profondeur en m , T la température de l'eau en $^{\circ}C$, S la salinité en ‰. Une valeur moyenne courante est d'environ 1500 m/s, avec des variations généralement comprises entre 1450 et 1550 m/s.

Variations de célérité et profil bathycélérimétrique Si la température et la salinité influencent la célérité, alors il est évident que cette vitesse fluctue en fonction de la géographie puisque la température de l'eau varie avec les saisons et les latitudes et que des mers sont réputées plus douces ou plus salées. En effet, la salinité moyenne mondiale est d'environ 35 ‰ alors que celle de la mer Méditerranée est à 38.5 ‰ (mer quasiment fermée avec forte évaporation) et celle de la mer Baltique est moitié moindre (grands apports en eau douce). Plus régionalement encore, la salinité diminue aux embouchures des fleuves (Etter; 1998).

A la vue de l'équation Eq. 1.3, on note également que la célérité dépend aussi de l'immersion. Elle prend en compte la profondeur z , car la pression hydrostatique fait croître la célérité d'environ 16 m/s par kilomètre. Dans un même temps, la température T tend à décroître avec la profondeur. Cette décroissance de température n'est pas linéaire, car il y a généralement dans les eaux de surface une couche où l'eau est brassée (courant, vent, apports extérieurs d'eau). Cette couche d'eau, qui peut être plus chaude grâce au soleil, se distingue des eaux profondes plus stables. L'interface de ces eaux s'appelle la thermocline et sa profondeur n'est pas la même partout dans le monde, elle se situe à un millier de mètres en Atlantique alors qu'en Méditerranée, elle se situe vers 200m. La forme des profils de célérité est donc assez spécifique à une mer ou à un océan.

La figure 1.1 correspond à deux profils de célérité et de température typiques en mer Méditerranée, profil d'hiver en bleu et profil d'été en rouge. Les profils ont été réalisés en mer Méditerranée (Salinité à 38), dans des zones littorales, donc avec moins de 50m de profondeur (pas d'influence de la pression sur la célérité). L'hiver, l'eau est à température homogène de la surface au fond, aux alentours de 14 à 16°C, comme c'est le cas sur le profil bleu. Tandis qu'en été, les eaux de surface sur les côtes françaises peuvent monter à 25°C et une thermocline est souvent présente à quelques dizaines de mètres sous la surface faisant tomber la température de l'eau jusqu'à 10°C en deçà. Sur le profil rouge, une thermocline est présente à environ 13m sous la surface, la température de l'eau chute de 25°C à 19°C en quelques mètres.

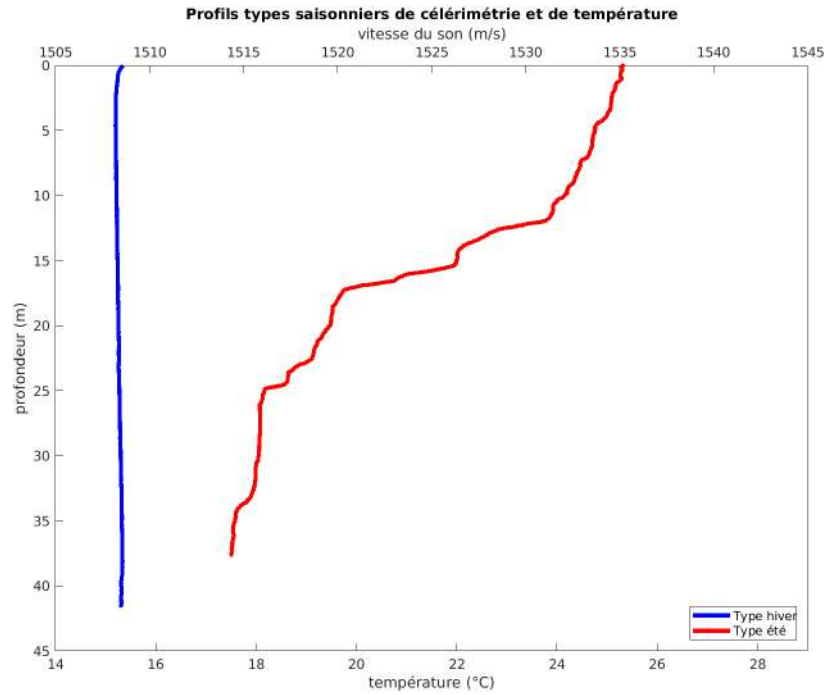


FIGURE 1.1 – Exemple de profils de vitesse du son et de température en mer Méditerranée. Le profil d’hiver a été réalisé en baie de Saint-Florent (Haute-Corse) en mars 2020. Le profil d’été a été réalisé en baie de La Ciotat (Bouches-du-Rhône) en août 2019.

Du fait de l’inhomogénéité spatiale des célérités, des phénomènes de réfractions et de réflexions surviennent à l’interface des différentes couches d’eau. Considérant une onde plane rencontrant l’interface, une réflexion spéculaire se fait dans le premier milieu et une réfraction est engendrée dans le second milieu selon un angle θ_2 donné par la loi de Snell-Descartes (Fig. 1.2 flèches vertes, oranges et jaunes) :

$$\frac{\sin \theta_1}{v_1} = \frac{\sin \theta_2}{v_2}.$$

Il se produit alors une division/duplication de l’onde qui continue à se propager selon deux chemins. Ces phénomènes se reproduisent à chaque changement des paramètres de l’eau de mer. L’évolution progressive de la célérité (avec l’immersion) ou discontinue (avec une thermocline) provoque une déviation de l’onde : l’onde ne se propage pas en ligne droite.

Réflexions et diffusion sur le fond et la surface Le fond et la surface de la mer occasionnent des réflexions supplémentaires à celles liées aux couches de la colonne d’eau. Mais celles-ci diffèrent un peu. La réfraction dans l’air et dans les fonds sont négligeables grâce à la grande différence d’impédance entre les milieux (à part en basse fréquence pour l’étude des sous-sols). En revanche, les surfaces étant rugueuses, il se produit des diffusions de l’onde dans toutes les directions : l’onde réfléchi spéculairement conserve ses caractéristiques mais subit des pertes d’intensité avec sa duplication lors des réfractions et/ou diffusion (Fig. 1.2 flèches roses).

En résumé La propagation des OA dans le milieu marin n’a rien de comparable à celle en milieu aérien. Elle implique :

- Une célérité variable dans le temps, dans l’espace et selon la profondeur,
- Un amortissement différent selon les fréquences utilisées,
- Des rebonds et des déformations de trajectoires aux interfaces ; le signal réceptionné est un signal multi-trajets composé d’un trajet direct suivi d’une multitude d’échos atténués ;
- Une perturbation par le bruit ambiant que l’on ne peut anticiper.

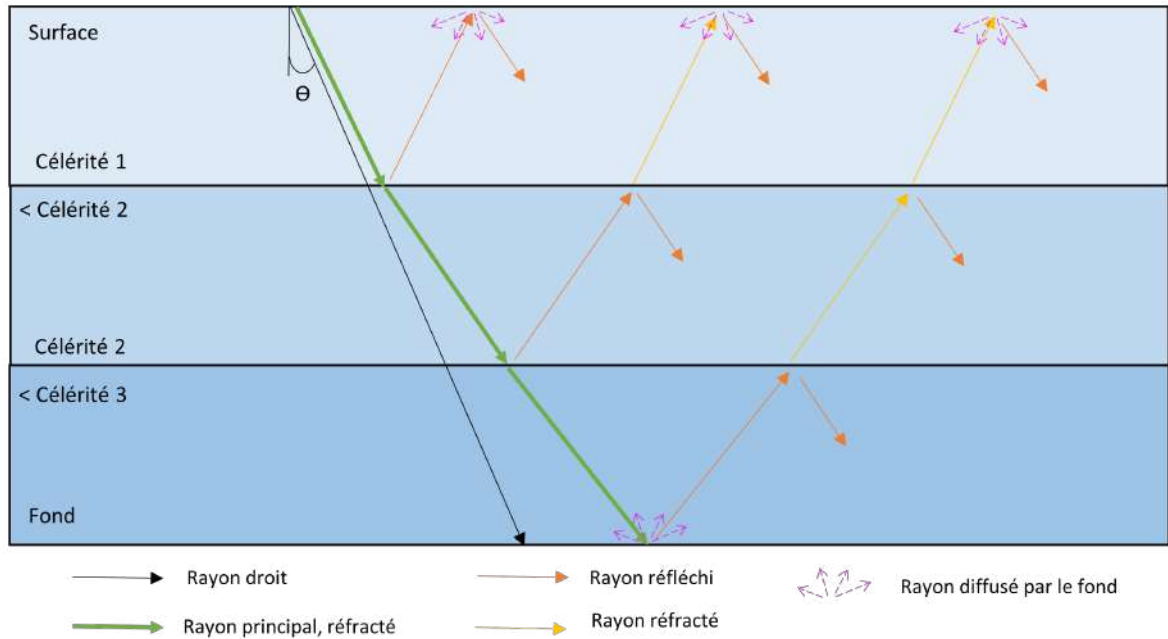


FIGURE 1.2 – Phénomène de réflexion et de réfraction d’une onde sur une interface plane (les couches d’eau de célérités croissantes) et de diffusion sur les interfaces rugueuses (fond et surface). Illustration des déformations et multi-trajectoires d’un rayon.

Tout cela impose une adaptation spécifique du matériel d’ASM selon les besoins.

1.2 Composition d’un système acoustique

La chaîne de mesure globale est généralement composée des éléments suivants que nous allons maintenant rappeler.

1.2.1 Transducteur électro-acoustique

Les transducteurs électro-acoustiques sont des dispositifs qui transforment l’énergie électrique en énergie acoustique, et inversement. Ces conversions sont réalisées grâce aux propriétés piézo-électriques de certains matériaux comme les céramiques synthétiques (dites *piézo-céramiques*). L’application d’un champ électrique provoque une déformation du matériau en convertissant les modulations du champ d’origine, et réciproquement. Les contractions et extensions du matériau dans l’eau engendrent une OA.

Les caractéristiques propres aux transducteurs sont nombreuses. Voici les principales qui seront nécessaires dans la suite :

- Le Niveau d’Émission (NE) acoustique délivré par le transducteur. Il dépend de la puissance électrique fournie et de propriétés du transducteur. Typiquement, NE est de l’ordre de 180 à 240 dB/ μ Pa/1m.
- La fréquence nominale d’utilisation, généralement définie par la résonance des matériaux.
- La bande passante : plus on exploite la résonance du transducteur, plus sa bande passante est faible. Un compromis est à trouver entre les bons rendements des fréquences de résonance et la largeur de bande passante.
- La directivité. Elle traduit la capacité à privilégier un secteur angulaire pour l’émission et la réception. Une antenne est d’autant plus directive que ses dimensions sont grandes devant la longueur d’onde.

1.2.2 Directivité d'une antenne

L'intérêt d'avoir une antenne directive plutôt qu'une antenne omni-directionnelle est simple : on s'intéresse en général à une cible et non à l'espace marin entier (comme en tomographie). La directivité d'une antenne s'obtient, soit en utilisant un transducteur dont la dimension est grande devant la longueur d'onde, soit en couplant correctement plusieurs transducteurs.

En émission, la directivité concentre l'énergie acoustique dans un secteur angulaire donné, l'intensité est donc plus forte que si elle avait été émise dans toutes les directions. En réception, elle permet de s'affranchir d'une partie des trajets multiples et du bruit ambiant.

La fonction de directivité d'un transducteur pour une fréquence choisie décrit l'énergie transmise dans une direction préférentielle de l'espace donnée par les angles θ et ϕ en coordonnées sphériques. On la note $D_{ir}(\theta, \phi)$. Son calcul dépend de l'agencement (linéaire, circulaire, rectangulaire, etc.) du ou des transducteurs. L'onde de pression acoustique générée par les transducteurs possède une directivité régie par une loi de la forme suivante (Clay and Medwin; 1977) :

$$p(R, \theta, \phi) = D_{ir}(\theta, \phi)p(R, t)$$

où $p(R, t)$ est l'équation 1.2.

Le diagramme de directivité est le tracé de la fonction de directivité, il permet de s'intéresser à la distribution angulaire du rayonnement de l'antenne. Cette distribution se caractérise par des lobes pointant dans des directions privilégiées, dont certaines sont bien plus favorisées : le lobe le plus fort par rapport à ses voisins est le lobe principal. Le lobe principal est caractérisé par son ouverture à -3 dB (notée $2\theta_3$ sur la fig. 1.3). Les lobes de plus faibles amplitudes sont les lobes secondaires, souvent considérés comme indésirables, ils sont à minimiser au profit du lobe principal. Une antenne de grande dimension devant les longueurs d'onde utilisées permet d'affiner le lobe principal et d'augmenter la directivité mais multiplie le nombre de lobes secondaires (fig. 1.3).

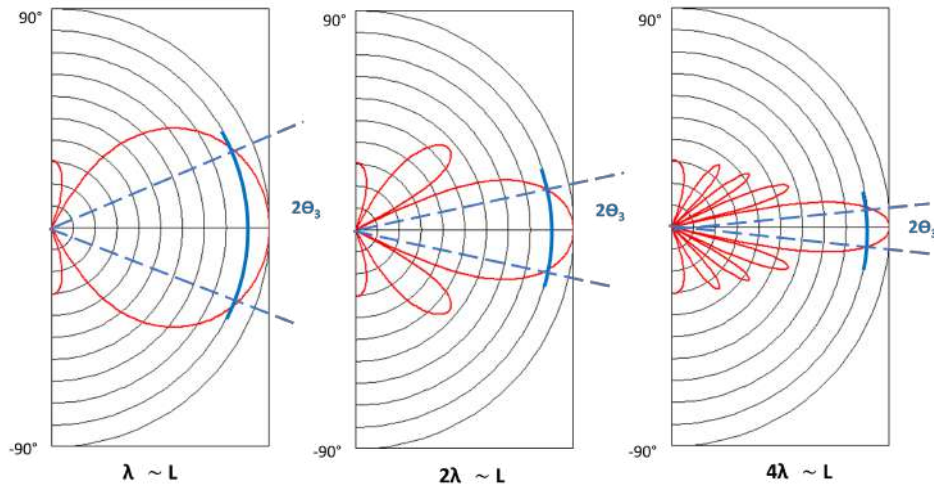


FIGURE 1.3 – Diagrammes de directivité d'une antenne linéaire, de dimension L , à différentes fréquences d'émission de longueur d'onde λ .

Gain de directivité La directivité intervient dans l'équation du sonar sous forme d'un gain de directivité GD . Il quantifie le "gain spatial d'énergie" d'une antenne de directivité $D_{ir}(\theta, \phi)$ par rapport à une antenne omnidirectionnelle. Il s'exprime en décibels par $GD = 10 \log Q$, où Q est

le facteur de directivité évaluant l'intensité acoustique reçu dans une direction par rapport à une intensité d'une source omnidirectionnelle (Clay and Medwin; 1977).

$$Q = \frac{I_t(R, \theta, \phi)}{I_{t,moy}(R)} = 4\pi \left[\iint D_{ir}^2(\theta, \phi) d\theta d\phi \right]^{-1}.$$

1.2.3 Signaux émis

Deux principaux types de signaux sont émis par les systèmes d'ASM, le ping et le chirp. Leur introduction permet de présenter le gain de traitement GT , dernier élément utile à l'équation du sonar. Ce gain correspond à l'amélioration du rapport signal à bruit de l'entrée à la sortie de la chaîne. Il est évalué, en décibels, par le produit de la bande passante B_p et la durée τ de l'impulsion émise :

$$GT = 10 \log(B_p \tau).$$

Le ping est une impulsion monochromatique et est le plus simple et le plus courant des signaux émis en ASM. Il peut s'agir d'une sinusoïde de fréquence f_0 émise durant une courte durée τ :

$$s(t) = \begin{cases} A_m \cos(2\pi f_0 t) & \forall t \in [-\tau/2, \tau/2] \text{ et une amplitude } A_m \in \mathbb{R} \\ 0 & \text{sinon.} \end{cases}$$

La largeur de bande à -3 dB vaut $\delta f_{-3dB} \approx 1/\tau$, ce qui veut dire que GT est proche de 0 dB et qu'il n'y a pas de gain du fait de l'utilisation de cette forme d'impulsion. En revanche, les pings sont simples à émettre et à traiter en réception.

Le chirp est une impulsion modulée en fréquence. La modulation m_{od} peut être linéaire, toujours dans un cas sinusoïdal :

$$s(t) = \begin{cases} A_m \cos(2\pi t (f_0 + m_{od} \frac{t-\tau}{2})) & \forall t \in [-\tau/2, \tau/2] \text{ et une amplitude } A_m \in \mathbb{R} \\ 0 & \text{sinon.} \end{cases}$$

La fréquence varie linéairement durant τ de $f_0 - m_{od}\tau/2$ à $f_0 + m_{od}\tau/2$, soit une largeur de bande de $B_p = m_{od}\tau$. Ainsi, le gain de traitement tend à augmenter avec la durée de l'impulsion et/ou avec la largeur de bande.

Si le ping suffit pour de nombreuses applications, il ne permet pas d'obtenir des informations sur la nature de la cible, contrairement au chirp.

1.3 L'équation du sonar

L'équation du sonar pose un bilan d'énergie pour estimer si une cible peut être détectée considérant les paramètres du milieu et du système acoustique. Très sommairement, elle correspond à :

$$\text{Signal} - \text{Altérations} > \text{Seuil}.$$

L'équation se décline en deux variantes, soit avec un sonar passif utilisant le bruit rayonné de la cible, soit avec un sonar actif utilisant la rétro-diffusion sur la cible d'un signal émis. On ne s'intéresse dans cette thèse qu'à la version active.

La cible Dans l'équation, la cible est représentée par son indice de cible ou plus communément le *Target Strength (TS)*. Le TS traduit la capacité de la cible à renvoyer l'énergie acoustique vers le système. Le TS d'une cible peut être estimé en décibels par $TS = 10 \log(I_{t,r}/I_{t,i})$, où $I_{t,i}$ est l'intensité interceptée par la cible et $I_{t,r}$ est l'intensité rétro-diffusée. Cet indice est fonction de la nature de la cible et de sa dimension (devant le faisceau) mais aussi du signal intercepté (angle et fréquence).

Le signal est fonction de caractéristiques du système d'ASM, c'est-à-dire au niveau d'émission NE , à son gain de directivité GD et à son gain de traitement GT .

Les altérations comprennent les pertes de transmissions $2PT$ et le bruit ambiant NB . Les réverbérations dues aux trajets multiples sont corrigées par le gain de directivité.

L'équation du sonar actif en tant que généralité aux systèmes d'ASM actifs peut ainsi se résumer au niveau de l'équation suivante :

$$NE - 2PT + TS - NB + GD + GT > ID.$$

L'indice de détection, ID , est un seuil qui permet de déterminer si la configuration est suffisante pour permettre une détection. Cet indice peut valoir 10 dB. Or la détection n'est que la première étape de la prise d'information sur une cible, elle est rarement le but en soit de la mesure : détecter un signal ne signifie pas encore qu'un objet particulier, comme une mine ou un poisson, ait été reconnu (d'autres traitements complémentaires sont nécessaires). Pour compléter ce premier chapitre, la dernière section est consacrée au sondeur monofaisceau utilisé dans le cadre de mesures bathymétriques et halieutiques. Son fonctionnement est relativement simple et permet d'expliquer la caractérisation (nature, position, vitesse) d'une cible, au-delà de la détection.

1.4 Mesures hydrographiques par sondeurs acoustiques

1.4.1 Sondeur bathymétrique

Un sondeur monofaisceau mesure la hauteur d'eau, sous un navire, en étudiant le temps de trajet aller-retour des signaux sonores rétro-diffusés sur le fond marin. Cela en fait le système acoustique le plus simple pour la mesure bathymétrique : il est constitué d'un unique faisceau (unique émetteur et récepteur) pointé au nadir sous le navire. Son faisceau est assez large, on peut parler d'un cône, et possède une ouverture $2\theta_3$ de 5 à 30° selon les modèles.

Malgré leur simplicité, il existe des dizaines de sondeurs monofaisceaux aux paramètres différents répondant à différents besoins. Concernant la mesure bathymétrique uniquement, les paramètres vont varier selon les profondeurs à atteindre.

Dans ce cas de figure, le paramètre le plus important est certainement la fréquence des pings. Les gammes sont comprises de 10 à 500 kHz, ce qui permet de mesurer des plus petits aux plus grands fonds. Les fréquences en deçà de 100 kHz servent à insonifier les grands fonds océaniques grâce à leur portée, tandis qu'au delà elles sont destinées à étudier des profondeurs de plateaux continentaux jusqu'aux eaux littorales. Les longueurs d'onde des fréquences utilisées impliquent un dimensionnement adapté des sondeurs : l'antenne doit être au moins aussi grande que la longueur d'onde. Pour des mesures portuaires ou littorale, le sondeur peut tenir dans la main : il est donc installable sur des navires de taille de type semi-rigide ou kayak.

Les durées d'émission de ping sont également ajustées selon les profondeurs pour gagner en portée : d'environ 1 ms pour les grands fonds à des durées de l'ordre de 0.05 ms pour les très petits fonds. En plus de jouer sur la portée du signal, la durée d'émission conditionne la résolution du signal : plus la durée est courte, moindre est la portée, mais meilleure est la résolution et ainsi la précision de la profondeur mesurée.

Enfin, la cadence d'émission des pings est adaptée aux profondeurs à étudier et aux durées d'émission : il s'agit de n'émettre un nouveau ping que lorsque la réception est finie car soit émetteur et récepteur sont confondus soit l'émission sature la réception. Ces cadences vont varier d'un ping toutes les quelques secondes à plusieurs dizaines de pings par secondes. Cela implique qu'en petits fonds, la quantité de données acquises est bien plus grande qu'en grands fonds. Pour obtenir une densité de points comparable, il faut alors augmenter le nombre de passages sur la zone à lever car la vitesse de navigation est généralement déjà faible.

En réception, le signal subit d'abord un traitement non-cohérent à l'aide d'un filtre passe-bande autour de la fréquence d'émission puis d'une quadrature et d'une intégration de l'énergie sur la durée du ping (Lurton; 1998), ceci dans le but d'améliorer le rapport signal à bruit. Ensuite, un TVG (*Time Varying Gain*) est appliqué, c'est un traitement classique en acoustique qui consiste à corriger les pertes d'amplitude du signal lors de sa transmission selon le $2PT$ (fig. 1.4). Ainsi le TS du fond est indépendant de la distance à la source, sinon il pourrait être devenu trop faible et ne pas être détecté. Enfin, la détection du fond est réalisée par la recherche du maximum d'amplitude dans le signal.

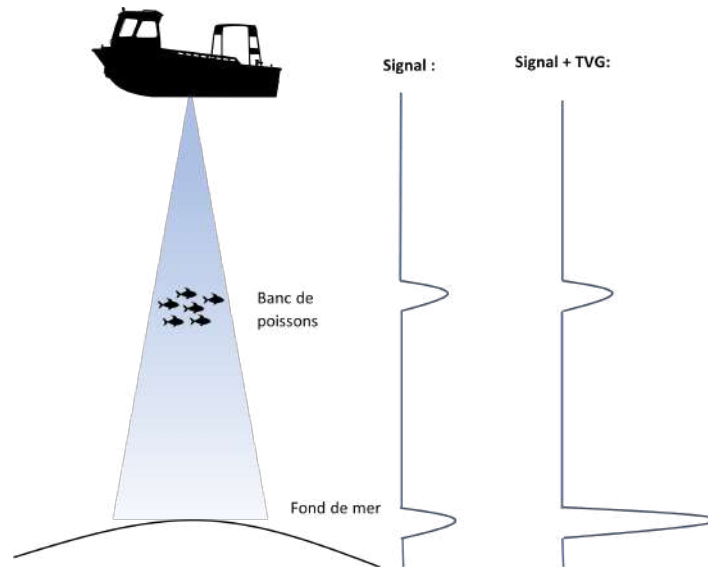


FIGURE 1.4 – Rôle d'un *Time Varying Gain* dans la compensation du signal acoustique.

Le sondeur monofaisceau est l'instrument acoustique le plus simple pour réaliser des mesures bathymétriques de part sa conception mécanique et sa chaîne de traitement de l'information. Cependant, il souffre de deux points faibles majeurs. D'une part, il n'insonifie qu'une petite zone à la verticale du navire : il n'y a pas de levée exhaustive des fonds avec ce type de sondeur. D'autre part, l'ouverture du faisceau ne permet pas des mesures fines des éléments du fond ou de distinguer deux objets proches. L'instrument suivant pallie ces défauts

1.4.2 Imagerie des fonds par sonar à balayage latéral

Comme rapidement énoncé dans l'introduction de cette partie, le sonar à balayage latéral (SBL) ou *side-scan sonar* est un système post-seconde guerre mondiale dont les applications sont tournées vers de l'imagerie acoustique des fonds essentiellement pour de la détection d'objet. Cela en fait un sondeur bien différent du sondeur monofaisceau et une transition avec le prochain chapitre centré sur le SMF.

Une première différence fondamentale avec le précédent système est que le SBL n'est pas un sondeur de coque mais un sondeur remorqué (dit le "poisson") derrière la navire et évoluant à mi-hauteur dans la colonne d'eau. Ensuite, contrairement au sondeur monofaisceau, ce sondeur ne s'intéresse pas ou peu au fond à son nadir mais plutôt au fond de part et d'autre de sa route qu'il insonifie par des rayons rasants (figure 1.5). Pour cela, ce sonar est équipé de deux antennes distinctes pour insonifier à babord et à tribord qui émettent un faisceau large transversalement à la route et fin longitudinalement. Cette bande rétrodiffuse une partie de l'onde acoustique incidente et les antennes reconstruisent le signal sous forme d'une ligne. Avec l'avancement du sonar, l'assemblage des lignes forme une image des fonds où sont visibles les irrégularités (figure 1.6). Les images formées sont en niveau de de gris représentant l'amplitude du signal reçu.

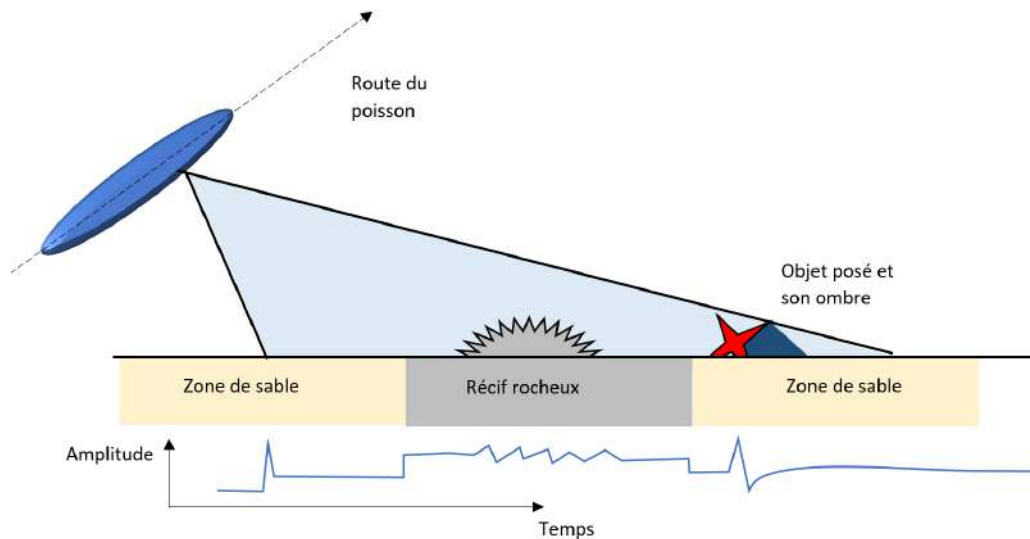


FIGURE 1.5 – Principe du sonar à balayage latéral (sur un seul côté) et allure du signal reçu. D'après [Lurton \(1998\)](#).

Les signaux reçus (figure 1.5) sont plus complexes que ceux d'un monofaisceau atteignant le fond en un point donné. Ici, la partie du faisceau proche de la verticale atteint en premier le fond et forme un écho très fort en amplitude qui n'est pas exploité pour former l'image acoustique (la bande blanche centrale de la figure 1.6). Cependant, il permet de déterminer l'altitude du poisson par rapport au fond. La partie la plus oblique du faisceau atteint le fond avec un retard temporel à la partie verticale du faisceau. Les amplitudes vont varier selon la nature des fonds, des reliefs et des objets présents ([Blondel; 2009](#)). Les objets posés sur le fond avec une forte élévation créent des zones d'ombre pour les faisceaux incidents, cela se traduit par une faible amplitude précédée d'un fort écho. Ce phénomène d'ombre est un réel avantage pour la détection (épave, mine, archéologie, etc.). Ensuite, il faut reconstituer une carte des fonds avec une correction géométrique, c'est-à-dire savoir sur quelle forme de support l'image acoustique doit venir se plaquer. La correction géométrique généralement choisie est une hypothèse de fond plat avec une éventuelle pente quand ce levé n'est pas complété d'une mesure bathymétrique.

Les antennes utilisées sont de géométrie rectangulaire allongées dans le sens longitudinal du poisson, créant une directivité de plusieurs dizaines de degrés dans le sens transversal et fine de l'ordre du degré dans le sens longitudinal ([Lurton; 1998](#)). Les résolutions sont donc inhomogènes et indépendantes dans les deux directions ([Ollivier; 1995](#)). Dans le sens de la route du poisson, la résolution des images est définie par la largeur du faisceau quand celui-ci rencontre le fond (fonction du degré d'ouverture et de la distance oblique avec le fond) mais également par la cadence des pings couplée à la vitesse du navire. Cela implique qu'un objet proche du poisson est plus résolu qu'un objet en bout de fauchée. La résolution perpendiculaire au poisson est définie, comme pour le sondeur monofaisceau, par la durée d'émission du ping mais il faut aussi tenir compte de l'angle d'incidence : la résolution est meilleure avec l'augmentation de l'angle d'incidence.

Les SBL peuvent être utilisés par faibles et grands fonds, leurs caractéristiques sont adaptées selon l'utilisation. Aux deux extrêmes, on peut distinguer les sondeurs longues et courtes portées ([Ollivier; 1995](#)). Les premiers emploient des fréquences de l'ordre de la dizaine de kiloHertz pour une portée de 20 km, ces sondeurs peuvent être tractés jusqu'à 1000 m de profondeur car utilisés pour les très grandes profondeurs. Les seconds emploient des fréquences de l'ordre de la centaine de kiloHertz pour une portée d'une centaine de mètres, ces sondeurs peuvent être tractés jusqu'à une centaine de mètres de profondeurs.

La force du SBL est d'être monté sur un poisson rasant les fonds, échappant aux bruits de surface et subissant moins les mouvements de surface. Cela constitue également un défaut. Les

mouvements de surface sont transmis - en très atténués - au poisson par l'intermédiaire du câble et provoquent des distorsions de l'image acoustique. La solution est d'embarquer un appareil de mesure de ces mouvements pour appliquer une correction en post-traitement ou encore mieux en temps réel (Ollivier; 1995). Mais le véritable souci vient de la grande imprécision de positionnement du poisson sous l'eau au bout de son câble : le câble de part son poids peut ne pas être totalement tendu, le poisson peut être déporté de l'axe longitudinal du navire, etc. Les objets détectés peuvent être situés à plusieurs mètres, au minimum, de la réalité.

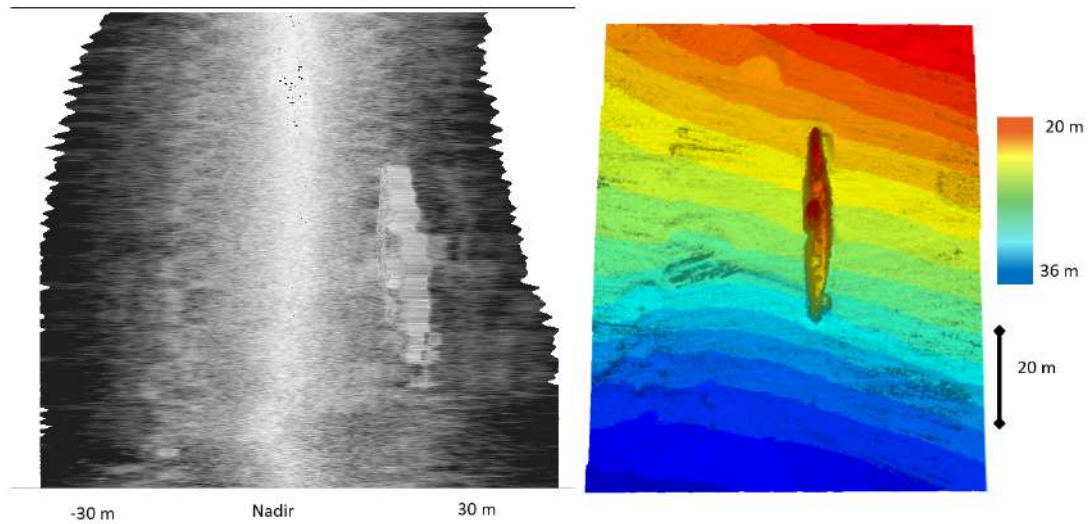


FIGURE 1.6 – Épave de sous-marin de 45 m de long par 24-30m de profondeur, posé kiosque en l'air sur un fond de sable. A gauche, imagerie acoustique de 500 pings obtenue avec le SMF de Seaviews. A droite, la carte bathymétrique, résultante de plusieurs transects, au SMF du même sous-marin pour aider à la visualisation.

Le sondeur multifaisceaux peut être vu comme un instrument intermédiaire entre le sondeur monofaisceau et le sondeur à balayage latéral car il combine d'une part la précision bathymétrique du monofaisceau et d'autre part la résolution des images acoustiques. Cependant, le SMF est bien souvent jugé comme supérieur car il résout le problème de résolution variable dans l'espace, la trop simple hypothèse du fond plat et les imprécisions de positionnement.

Chapitre 2

Systeme d'acquisition de données au sondeurs multifaisceaux

Ce chapitre, dédié aux sondeurs acoustiques multifaisceaux, contient une partie théorique, celle du fonctionnement et de la chaîne d'acquisition, et une partie appliquée décrivant le matériel effectivement utilisé à Seaviews ainsi que les données qu'il permet d'acquérir.

2.1 Sondeurs multifaisceaux : fonctionnement & principales caractéristiques

Les SMF fonctionnent différemment des sondeurs monofaisceau. Lors de mesures réalisées au sondeur monofaisceau, le ping est composé d'un unique et large faisceau émis au nadir sous le navire (verticalement), alors qu'à l'inverse, un sondeur multifaisceaux reçoit plusieurs faisceaux fins répartis le long d'une fauchée perpendiculaire au bateau (Fig. 2.1). Ces faisceaux pointent vers le fond de la mer avec chacun un angle différent, ce dépointage du nadir (le nadir équivaut à 0°) est réalisé par de la formation de voie dont nous rappellerons le principe au niveau de la sous-section 2.1.1. Tandis que la finesse des faisceaux est obtenue par l'intersection de faisceaux émis et reçus (système à faisceaux croisés), qui fera l'objet de la sous-section 2.1.2.

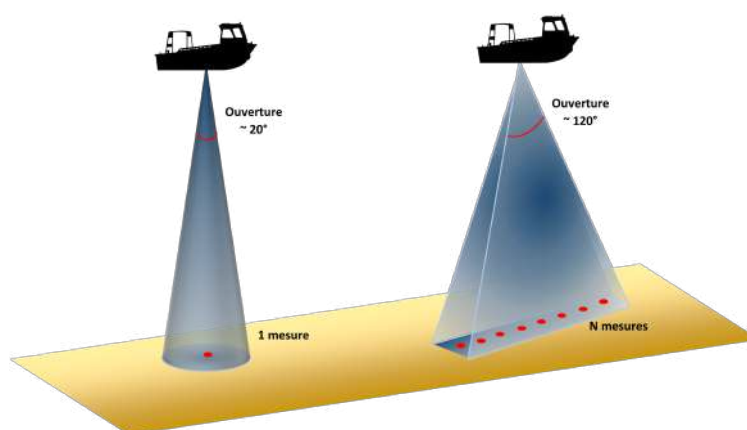


FIGURE 2.1 – A gauche, le faisceau d'un sondeur monofaisceau est large et dirigé vers le nadir. A droite, les faisceaux d'un SMF sont fins et dispersés angulairement sur une fauchée perpendiculaire au navire.

2.1.1 Formation de voie

On aurait pu simplement imaginer qu'un sondeur multifaisceaux ne soit qu'un assemblage de sondeurs monofaisceaux indépendants dont les angles d'émission auraient été réalisés par un montage ou une rotation mécanique. Cela ne peut fonctionner ainsi, car les pings de chaque individu interféreraient les uns avec les autres dans l'écoute des échos. Pour éviter cette cacophonie, la formation de voie joue le rôle de chef d'orchestre.

La formation de voie ou *beamforming* en anglais est une technique de traitement du signal appliquée dans le cadre d'un réseau de capteurs (antenne). Les signaux sont contrôlés pour interférer - en phase et en amplitude - de manière à provoquer une sélectivité spatiale (Van Veen and Buckley; 1988). Ce contrôle peut être obtenu en appliquant un retard temporel entre les capteurs.

Cas d'une antenne linéaire La figure 2.2 représente une antenne linéaire de N capteurs a_n omnidirectionnels régulièrement espacés de la distance R . Si aucun retard n'est appliqué (Fig. 2.2 gauche), l'interférence entre capteurs produit une onde plane sans dépointage. Si un retard est produit entre capteurs (Fig. 2.2 droite), tels que a_n est légèrement en avance sur a_{n+1} , alors le résultat est une onde plane dépointée de θ_r du nadir. Le retard δt_n appliqué à chaque capteur pour obtenir θ_r est :

$$\delta t_n = (n - 1) \frac{R \sin(\theta_R)}{v}$$

avec v , la célérité du son dans le milieu marin. L'onde acoustique résultante Ω_{θ_R} est la moyenne des signaux de pression p_n des capteurs :

$$\Omega_{\theta_R}(t) = \frac{1}{N + 1} \sum_{n=1}^N p_n(t + \delta t_n)$$

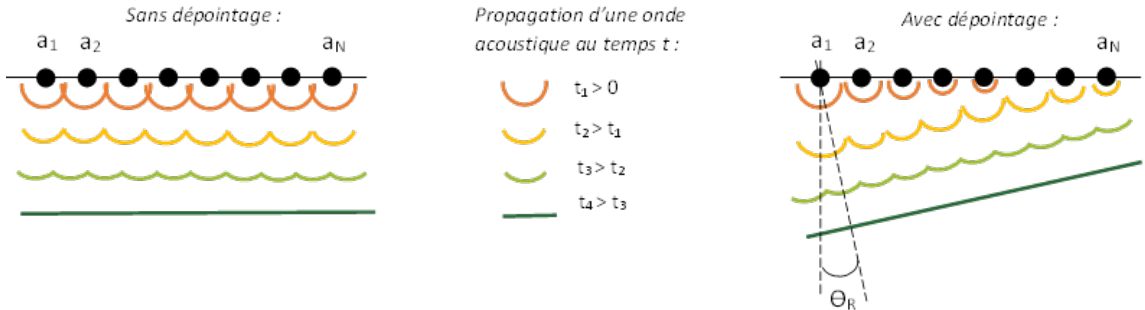


FIGURE 2.2 – Antenne linéaire avec et sans retard temporel entre capteurs, respectivement sans et avec dépointage de l'onde).

Diagramme de directivité Dans le cas d'un sondeur monofaisceau, le lobe principal est pointé vers l'angle 0° , le nadir. Pour un faisceau de SMF dépointé de 30° , le lobe principal pointe vers l'angle 30° et ce lobe est plus "épaté" que sans dépointage (Fig. 2.3). La dégradation de la largeur du lobe principal rend l'antenne moins directive et s'obtient par le facteur $1/\cos(\theta_R)$ (soit 1.15 pour un angle de 30°). Les faisceaux extrêmes des fauchées sont susceptibles d'avoir des mesures de moindre qualité à cause de cette dégradation du lobe principal. Un compromis entre une orientation mécanique des capteurs et l'utilisation de la formation de voie est alors à faire. Une solution, comme pour le SMF présenté ci-après, est l'installation des capteurs sur un cylindre, on parle alors d'antenne circulaire.

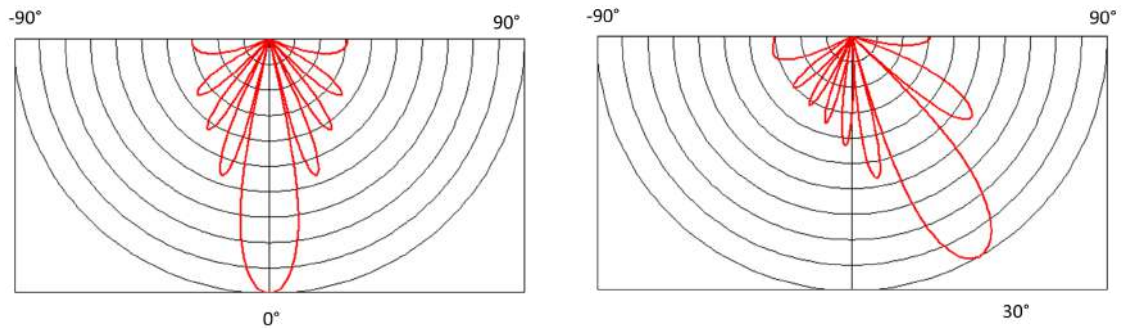


FIGURE 2.3 – Diagrammes de directivité d’une antenne non dépointée (à gauche) et dépointée de 30° (à droite)

2.1.2 Systèmes à faisceaux croisés

A la différence d’un sondeur monofaisceau ou d’un sonar à balayage latéral composé d’une antenne réalisant l’émission et la réception, un SMF se compose de deux antennes, l’une pour l’émission et la seconde pour la réception. Généralement, - c’est d’ailleurs le cas du SMF utilisé par Seaviews -, les deux antennes sont en configuration “croix de Mills”, c’est-à-dire perpendiculaires l’une par rapport à l’autre avec l’émission dans le sens longitudinal du navire. Cette installation permet d’obtenir une finesse d’ouverture dans la directivité des faisceaux de l’ordre du degré.

L’émission et la réception assurent des rôles différents :

- **l’émission** crée un faisceau large transversalement (150°) au navire et étroit longitudinalement (quelques degrés).
- **la réception** écoute l’écho dans un faisceau, large longitudinalement (20°) au navire et étroit transversalement (quelques degrés). La formation de voie est réalisée à la réception.

La combinaison de l’émission et de la réception permet d’obtenir un faisceau étroit de quelques degrés sous forme d’une pastille insonifiée carrée (fig. 2.4).

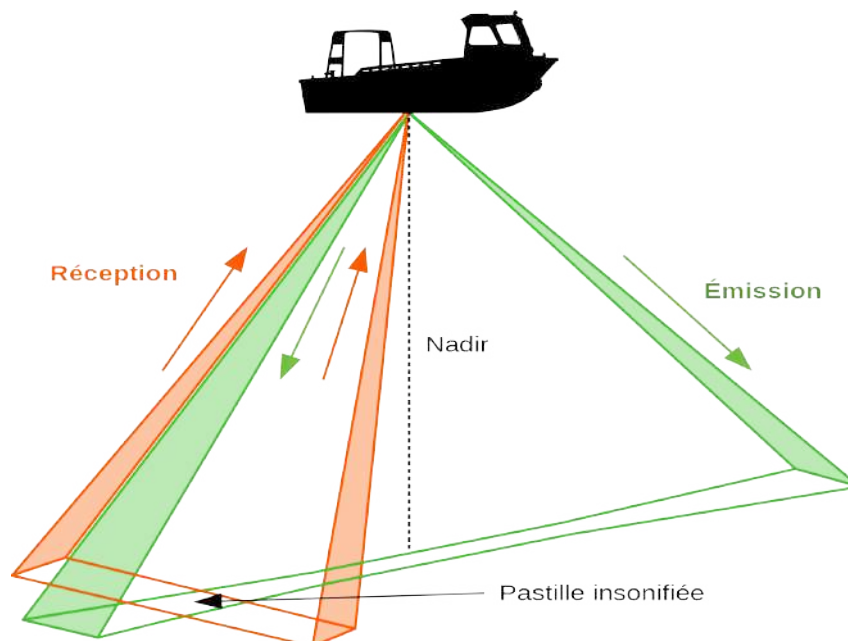


FIGURE 2.4 – Croisement des faisceaux d’émission et de réception.

La détection du fond dans le signal en réception peut se réaliser de deux manières, soit par l'étude de l'amplitude soit par l'étude de la phase, pour chaque voie formée. La recherche par amplitude se fait dans une fenêtre temporelle centrée sur le pic du fond, la détection correspond au barycentre de ce pic (fig. 2.5 gauche). Cette méthode est plus précise pour les voies proches du nadir car aux grands angles la fenêtre temporelle est élargie (le lobe est dégradé) et le pic de moins forte amplitude.

En revanche, la détection par la phase (interférométrie) est plus adaptée aux grands angles qu'aux voies trop verticales. L'interférométrie utilise la différence entre le déphasage mesuré, $\phi(\text{mesure})$, et estimé, $\phi(\theta)$ de deux faisceaux de l'antenne de réception. Le déphasage est estimé par $2\pi R \sin \theta / \lambda$, où R est la distance entre les capteurs responsables des faisceaux, θ est l'angle de la voie où l'on cherche le fond et λ la longueur d'onde des signaux. L'instant de détection correspond au changement de signe de cette différence de déphasage (fig. 2.5 droite). Cette méthode a l'avantage d'offrir une résolution de détection plus fine que la largeur du faisceau et est la base de méthodes hautes résolutions pour obtenir plusieurs sondes par voie (Rønhovde et al.; 1999).

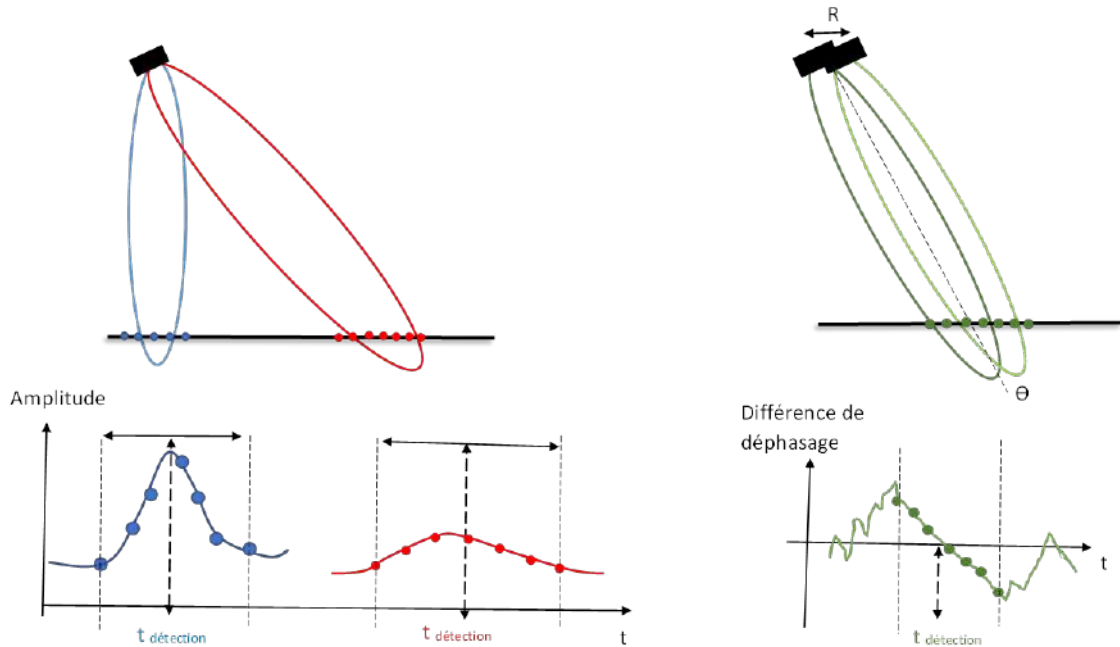


FIGURE 2.5 – Méthodes de détection du fond dans le signal rétrodiffusé par le fond.

2.1.3 Principales caractéristiques d'un SMF

Le SMF utilisé à Seaviews est le **R2Sonic 2022**. C'est un sondeur en croix de Mills, de configuration en "T" (Fig. 2.6). L'émetteur a une forme cylindrique et le récepteur est une antenne plane. Les deux antennes sont installées sur un même support, cela en fait un sondeur compact d'encombrement de l'ordre de 50*30*30 cm pour un poids de 11 kg.

Un SMF est généralement caractérisé par les critères suivants :

La fréquence acoustique est la fréquence de l'impulsion émise par le sondeur. Son choix impacte sur la portée du signal, la taille des antennes du système et la résolution spatiale de l'information acquise.

La portée du signal est très fortement liée à l'amortissement du son dans l'eau. Plus la fréquence de l'onde sonore est élevée, plus la portée est courte. La plage des fréquences utilisées s'étend de



FIGURE 2.6 – Tête (à gauche) et interface (à droite) du sondeur multifaisceaux de la marque R2Sonic version 2022. La tête est composée d'une antenne de réception rectangulaire (à gauche) et d'une antenne d'émission cylindrique (à droite). Le module d'interface sert à la création des impulsions et aux réglages des paramètres. (Source : R2Sonic)

la dizaine de kHz (étude des grands fonds jusqu'aux plaines abyssales) au MégaHerz (études des zones littorales). Plus la fréquence est basse, plus la longueur d'onde est grande, plus l'antenne doit être grande. Enfin, plus la fréquence est élevée, et plus l'antenne est grande devant la longueur d'onde et meilleure est la résolution spatiale.

Le R2Sonic fonctionne sur une gamme de fréquences pouvant aller de 170 à 450 kHz, réglable par pas de 10 kHz. Ces fréquences permettent donc des acquisitions jusqu'à 400 m de profondeur.

Nombre et largeurs des faisceaux. Le nombre de faisceaux varie selon les systèmes. D'une dizaine avec les premiers SMF, les faisceaux se comptent maintenant en centaines ; **le R2Sonic 2022 comporte 256 faisceaux.**

La finesse des faisceaux dépend de la fréquence choisie. Plus la fréquence est haute, plus le faisceau est fin et donc meilleure est la résolution spatiale. **A 200 kHz, les faisceaux du R2Sonic sont larges de $2^\circ \times 2^\circ$, tandis qu'à 450 kHz, les faisceaux atteignent une finesse de $0.9^\circ \times 0.9^\circ$. La quasi totalité des acquisitions à Seaviews sont réalisées à 450 kHz.**

Ouverture angulaire et répartition des sondes. L'ouverture angulaire est celle de la fauchée, c'est-à-dire l'angle entre les faisceaux les plus dépointés du nadir. Les premiers SMF n'avaient qu'une ouverture d'une centaine de degrés, les plus récents peuvent avoir **des ouvertures jusqu'à 160° , comme c'est le cas du R2Sonic.** Une ouverture de 90° autour du nadir (soit $[-45^\circ : 45^\circ]$) permet d'avoir une fauchée deux fois plus large que la hauteur d'eau sous le bateau ($2 \tan 45$), tandis qu'une ouverture de 150° permet d'avoir une fauchée sept fois plus large que la hauteur d'eau sous le bateau. Cependant, les mesures bathymétriques par les faisceaux les plus extrêmes sont moins précises que celles proches du nadir ; **Seaviews opte pour une ouverture réduite à $120\text{-}140^\circ$ (soit une fauchée 3 à 5 fois plus large que la hauteur d'eau).**

Le choix de répartition des sondes le long de la fauchée est laissé à l'opérateur. Il dispose du choix entre une répartition équiangulaire (angle constant entre deux faisceaux consécutifs) ou une répartition équidistante (les angles entre faisceaux sont adaptés pour que la distance entre les pieds de faisceaux soit constante). **La répartition équidistante est préférée à Seaviews pour conserver une répartition plus homogène des sondes sur la zone à lever et éviter un appauvrissement en sondes sur les bords de la zone.**

Durée d'un ping et cadence d'émission Comme la fréquence de l'impulsion, la durée d'émission d'un ping conditionne la portée et la résolution du signal : plus la durée est courte, plus la portée est faible et la résolution meilleure. Selon les SMF et les profondeurs à mesurer, les durées typiques sont comprises entre 10 μs (pour une zone littorale) et 10 ms (pour les plus grandes

profondeurs). Le **R2Sonic 2022** est le plus généralement réglé pour émettre des pings de 15 à 40 μ s.

La cadence d'émission correspond au nombre de pings que l'on peut émettre par seconde. Cette cadence est fonction de la durée d'émission et du temps que met le signal rétro-diffusé à revenir au sondeur. Sur le R2Sonic, elle est réglée indirectement par l'opérateur qui adapte la durée d'acquisition d'un ping par rapport à ce qu'il a en retour sur les écrans. Par exemple, **pour un fond à 10 m sous le bateau et une ouverture de 120°, la cadence est de 60 pings/s (le maximum de R2Sonic). Alors qu'avec 50 m de profondeur et une ouverture de 140°, la cadence tombe à 8 pings/s.**

Le R2Sonic 2022 est donc un SMF à vocation d'étude des zones littorales. Sa compacité lui permet d'être installé sur des petits navires (semi-rigides) pour la mesure de très faibles profondeurs (plages, ports).

2.2 Instrumentation associée : montage de Seaviews

2.2.1 Géo-positionnement

2.2.1.1 Quelques généralités

Petit point de vocabulaire : Les techniques de géo-positionnement modernes utilisent les constellations de satellites artificiels, le **GNSS** : Géolocalisation et Navigation par un Système de Satellites. Par abus de langage, on parle plus généralement de GPS (*Global Positioning Systems*), même si en toute rigueur les GPS concernent l'utilisation de la constellation de satellites américains opérationnelles pour les civils depuis 1995. Depuis, l'Union Européenne, la Russie et la Chine disposent de leur propre constellation, respectivement Galileo, GLONASS et Beidou.

Avant l'ère des satellites : Les méthodes de positionnement par GNSS remplacent en hydrographie le positionnement par sextant en haute mer et par radioguidage à l'aide de théodolites en milieu côtier.

Fonctionnement des GNSS : Un ensemble de satellites (au minimum quatre) fournit à un récepteur portable et mobile les informations qui vont lui permettre de déduire sa position (latitude, longitude, hauteur ellipsoïdale), sa vitesse de déplacement et la date et l'heure.

La déduction est faite en mesurant la distance émetteurs-récepteur avec le temps de parcours des signaux et la connaissance de la trajectoire des émetteurs sous forme d'almanach et d'éphémérides transmis par les émetteurs. Avec un unique satellite, les informations ne permettent que de positionner un récepteur sur une sphère de centre le satellite. Avec un second et un troisième, la position s'affine car les sphères se recoupent en un unique point : c'est la méthode de la trilatération (comme la triangulation mais en utilisant les distances entre objets et non les angles). Un quatrième satellite permet de déduire la date et l'heure très précisément car les satellites sont équipés d'horloges atomiques tandis que les récepteurs ne sont équipés que d'oscillateurs à quartz. On parle alors de méthode par multilatération.

La précision avec les récepteurs classiques est de l'ordre métrique et peut être améliorée par des techniques de RTK (*Real Time Kinematic*).

2.2.1.2 Méthode et matériel à Seaviews

La méthode RTK GNSS est la méthode de GNSS la plus précise actuellement. Elle utilise la phase des ondes porteuses des signaux émis par les constellations de satellites et pas seulement le contenu binaire des signaux. La précision de cette méthode permet d'atteindre un positionnement centimétrique.

En pratique, un système RTK se compose d'une station fixe appelée "base" placée à une position connue précisément et d'un ensemble de récepteurs mobiles (les "rover"). La base compare en

permanence sa position donnée par les satellites visibles avec la position connue où elle est installée puis diffuse aux rover le correctif à appliquer (Fig. 2.7).

En France, un réseau GNSS permanent (RGP) de bases est maintenu par l'IGN (Institut National de l'information Géographique et forestière) sur tout le territoire et permet de s'affranchir d'avoir à installer sa propre base. Mais si la zone d'étude est à plus d'une vingtaine de kilomètres d'une base RGP, il devient nécessaire d'installer sa propre base. Deux solutions sont alors possibles : installer une base sur un repère géodésique de l'IGN (position connue) ou l'installer à une position non connue. Dans le second cas, il est impératif que le ciel soit dégagé pour que l'antenne puisse voir un maximum de satellites et que la base reste en place plusieurs heures pour que celle-ci moyenne sa position dans le temps à l'aide des signaux des satellites et des éventuelles bases du réseau RGP.

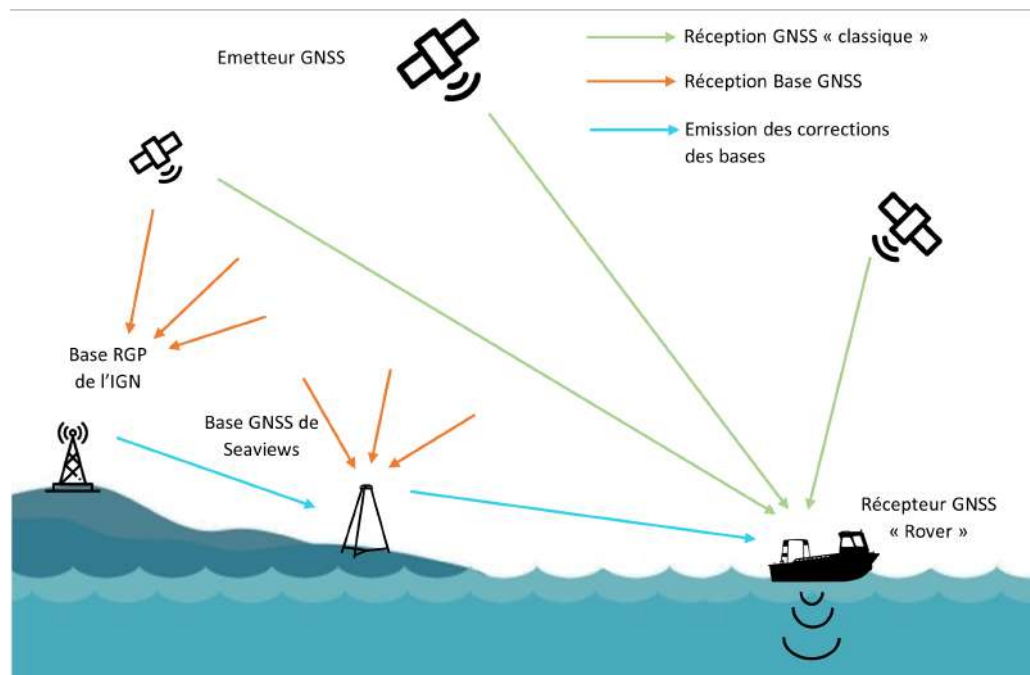


FIGURE 2.7 – Base et rover pour du RTK GNSS

Le matériel de positionnement se compose alors d'une base à terre et de deux antennes sur le navire. La base est une antenne Trimble AG25 et d'un module de calcul de multilatération, Trimble BX982. Les deux antennes sont des AeroAntenna Inc. AT1675-540, elles servent à la réception des données satellites et base mais aussi au calcul de cap du navire : l'une est installée à l'arrière du navire et l'autre au centre de gravité.

L'utilisation de GPS donnant Z en hauteur ellipsoïdale permet de s'affranchir des variations de marées, qui sont de toutes manières presque inexistantes en mer Méditerranée (marnage de 40 cm).

2.2.2 Mesure des mouvements du porteur

Même par temps calme, un navire subit les mouvements de la mer ce qui rend imprécise la position des mesures acoustiques sous l'eau. La mesure des mouvements du navire permet de connaître à tout moment l'orientation des faisceaux du sondeur. Il faut pour cela estimer le roulis, le tangage, le cap et le pilonnement du navire (Fig. 2.8).

On parle de l'attitude d'un navire pour désigner ses trois rotations (roulis, tangage et cap ou lacet) autour de son centre de gravité. Le roulis désigne les mouvements de rotation tribord-babord du navire, le tangage désigne les mouvements d'avant en arrière et le cap désigne les mouvements autour de son axe vertical. Le pilonnement, lui, est un mouvement de translation vertical du navire.

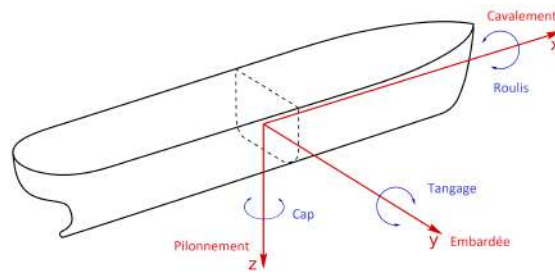


FIGURE 2.8 – Attitude d'un navire : mouvement de roulis, tangage, cap. Mouvement de pilonnement.

Les mouvements peuvent être mesurés avec une centrale d'attitude couplée au GPS. Le R2Sonic 2022 est fourni avec la centrale Applanix I2NS : le module de calcul est couplé avec l'interface du R2Sonic et la centrale est un boîtier étanche à fixer au plus proche du sondeur. Les antennes GPS du navire sont combinées au R2Sonic et à l'I2NS. De plus, cette centrale d'attitude est aussi inertielle : la centrale est capable d'intégrer les mouvements du navire pendant un laps de temps où les réceptions GPS décrochent. L'ensemble GNSS RTK et Applanix I2NS permet d'atteindre une précision de positionnement de ± 0.078 m en XY et en Z de $\pm(0.009 + 0.0066 * z)$ avec z la profondeur (en mètre).

2.2.3 Célérité de surface et profil de célérité

Le matériel comprend aussi deux célérimètres Valeport miniSVS (Fig. 2.9). Le premier est installé au côté du sondeur, il mesure la célérité en surface nécessaire à la formation de voie du SMF. Le second est installé au bout d'une bobine de 50m de câble (RS232) permettant de le faire descendre à la main en profondeur pour réaliser un profil de célérité (la bathycélérimétrie) et déceler une éventuelle thermocline (voir Fig. 1.1).

Le Valeport miniSVS est capable de mesurer simultanément la célérité, la température et la pression. La célérité est mesurée avec une OA de fréquence 2.5 MHz avec une précision de 0.02 m/s.



FIGURE 2.9 – Célérimètre Valeport miniSVS composé d'un transducteur pour la génération d'OA dans la partie cylindrique et d'une cage avec réflecteur calibré (à droite) pour la mesure de la célérité. (Source : Valeport)

On distingue trois types d'instruments pour la mesure de la célérité, la sonde XBT (Expandable bathythermograph), le capteur CTD (Conductivity-Temperature-Depth) et le célérimètre. La sonde XBT est une sonde perdable mesurant la température de l'eau durant sa chute dans la colonne d'eau, l'information remonte par un câble lui aussi perdable. La sonde XBT peut être utilisée sans immobiliser le navire mais souffre d'un biais dans la déduction de la célérité. Le capteur CTD demande l'arrêt du navire mais mesure précisément les trois paramètres de salinité, de température et de pression utiles au calcul de la célérité par une des équations empiriques (telle que celle en Eq. 1.3). Enfin, déployé à l'arrêt, le célérimètre mesure directement la vitesse du son dans l'eau par une onde acoustique parcourant une distance connue. C'est l'instrument le plus précis mais aussi le plus coûteux.

2.2.4 Navigation

Le navire océanographique de l'entreprise est le Seaviews One, un semi-rigide de 6 m transportable sur route par remorque. La coque en aluminium a été conçue spécialement avec un puits au centre de gravité pour y recevoir le SMF R2Sonic 2022 et la centrale inertielle Applanix I2NS. Les deux antennes GPS sont installées sur des mâts avec des emplacements prévus pour ces derniers.

Deux opérateurs prennent place à bord : l'un au pilotage du bateau, le second au réglage du SMF. Le pilotage peut se faire à la main mais aussi et surtout avec l'auto-pilote Raymarine ACU 200, indispensable pour réaliser des trajectoires droites à faible vitesse (de l'ordre de 5 noeuds). L'auto-pilote est asservi au logiciel maison ViewMap Acq; l'opérateur trace dans le logiciel les radiales de lever acoustique à effectuer et l'auto-pilote les enchaîne. Seule la vitesse de déplacement est laissée à l'appréciation de l'opérateur, selon l'état de la mer, la profondeur (et le ping rate), les éventuels dangers de la zone, etc.

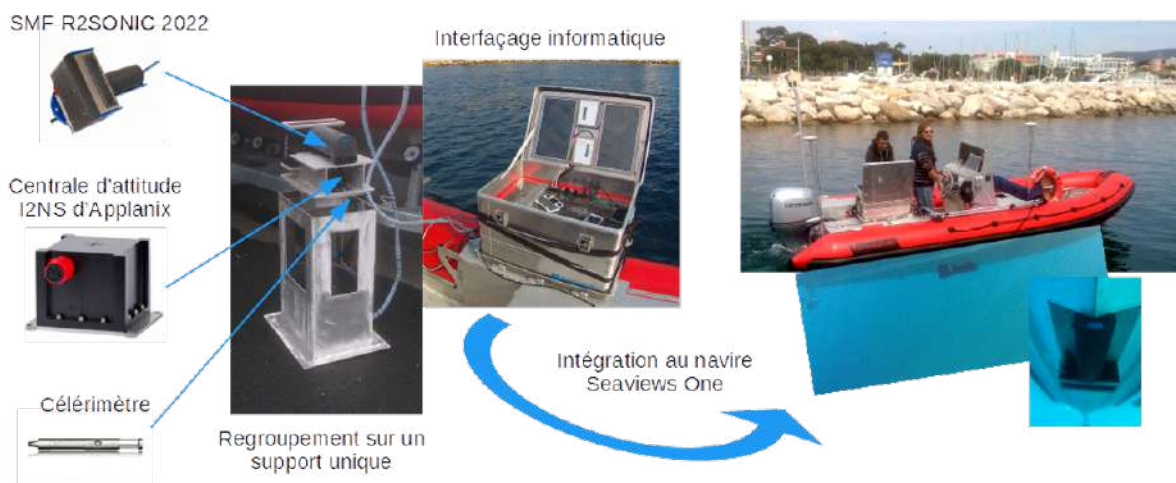


FIGURE 2.10 – Navire océanographie de Seaviews embarquant deux opérateurs et le SMF

L'ensemble des instruments est branché à la partie technique permettant les traitements informatiques (voir Fig. 2.10). Ils sont sous la supervision du second opérateur. Il effectue les réglages en temps réel du sondeur lors de l'acquisition : augmentation du range et de la puissance avec la profondeur, dépointage de la fauchée pour lever des fonds verticaux (digues, quais), élargissement ou rétrécissement de l'ouverture de la fauchée, etc. Il peut aussi demander au pilote de revenir sur une zone pour lever plus finement un fond complexe et éviter ainsi une zone blanche.

2.3 Données issues des Sondeurs multifaisceaux

2.3.1 Données bathymétriques

Qu'est-ce que c'est ? : Les données bathymétriques sont aussi appelées des sondes bathymétriques. Un sondeur multifaisceaux produit autant de sondes par ping qu'il possède de faisceaux. Dans le cas du R2Sonic 2022, 256 sondes sont donc estimées par ping mais une option de densification permet d'en obtenir 1024 par ping : lors de la détection du fond dans le faisceau, une interférométrie est réalisée pour évaluer quatre sondes supplémentaires. Les pings se présentent comme des lignes plus ou moins continues, comme en figure 2.11 où le navire est passé au dessus d'un bloc rocheux détaché d'une digue.

On peut distinguer trois à quatre niveaux de traitement des sondes. Le premier niveau fait référence aux sondes brutes issues du SMF. Le second niveau correspond aux sondes géo-référencées, corrigées en attitude et en célérité. Avec le logiciel développé au sein de SeaViews, l'exploitation des données bathymétriques commence au niveau deux quand toutes les données annexes sont

présentes (figure 2.11). Le troisième niveau correspond aux premières exploitations dont le tri des sondes erronées. Le quatrième niveau correspond à un produit fini, comme une carte sous forme de modèle numérique de terrain.

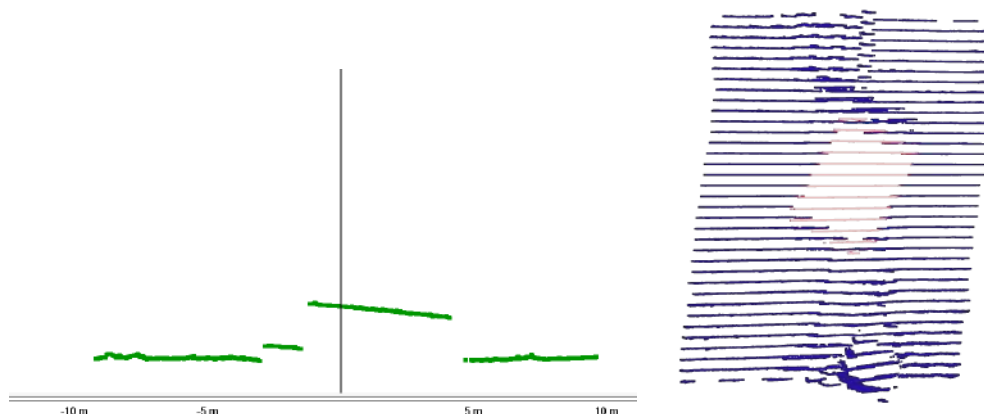


FIGURE 2.11 – Passage au dessus d’un bloc rocheux de 8m par 5.5m sur un fond plat. A gauche : ping sous forme de fauchée perpendiculaire au navire, les 1024 sondes sont réparties sur ± 10 m autour du nadir (ligne noire verticale). A droite : ensembles de pings consécutifs, le bloc a été coloré en rose artificiellement pour faciliter son identification.

A quoi ça sert ? : Les sondes bathymétriques servent principalement à la réalisation de cartographies des fonds, c’est-à-dire des cartes de topologie sous-marine. Ces cartes ont plusieurs emplois : elles peuvent servir à la prospection quand une zone est totalement inconnue, à la sécurité de navigation (port ou danger isolé), à suivre les mouvements de bancs de sable, etc. Ces cartes peuvent aussi être à la base de calculs de volumes de substrats à draguer dans un port (calculs de cubatures) ou de suivis de travaux sous-marins. La donnée bathymétrique peut aussi être dérivée sous forme de pente bathymétrique pour s’intéresser à la morphologie des fonds marins.

Comment traiter les données pour produire des cartes ? : Pour réaliser une carte bathymétrique, le principal travail que nous réalisons sur les données à Seaviews est de trier les données aberrantes. C’est un problème assez classique en statistique mais qui diffère des autres cas d’applications de part la quantité de sondes acquises en une journée (de l’ordre de plusieurs dizaines de millions) et de part la nature des erreurs (raisons physiques, bruits, répartition non gaussienne, etc.). Ce tri est nécessaire pour augmenter la précision des cartes : des normes (S-44 de l’OHI (2020)) régissent la précision demandée en fonction des besoins.

Les données encore très nombreuses après tri sont généralement mises sous forme d’un modèle numérique de terrain (MNT), réalisé par les logiciels de Seaviews, avec une dimension de maille adaptée à la taille de la zone d’étude, aux profondeurs et aux besoins. Le bloc de la figure 2.11 est présenté ci-dessous en MNT avec différentes résolutions, 50 cm, 10 cm et 5 cm (figure 2.12) ; parce que la zone est peu profonde (9-12m) et que l’intérêt se porte sur les blocs posés au fond, les mailles de dimension de l’ordre de 10 cm sont plus adaptées à la représentation. En deçà il n’y a pas plus d’information mais davantage de bruit.

2.3.2 Donnée de réflectivité

Qu’est-ce que c’est ? : Les données de réflectivité, dites *backscatter strength*, sont une forme d’imagerie des fonds marins obtenue par l’amplitude du signal rétrodiffusé par le fond. Selon la nature du substrat, physique et géométrique, le signal est modifié ; les sols mous absorbent plus l’énergie que les sols durs et selon la “rugosité” le signal est plus ou moins diffusé. Il est ainsi possible de différencier les roches, des vases et les sédiments à différentes granulométries (fig. 2.13).

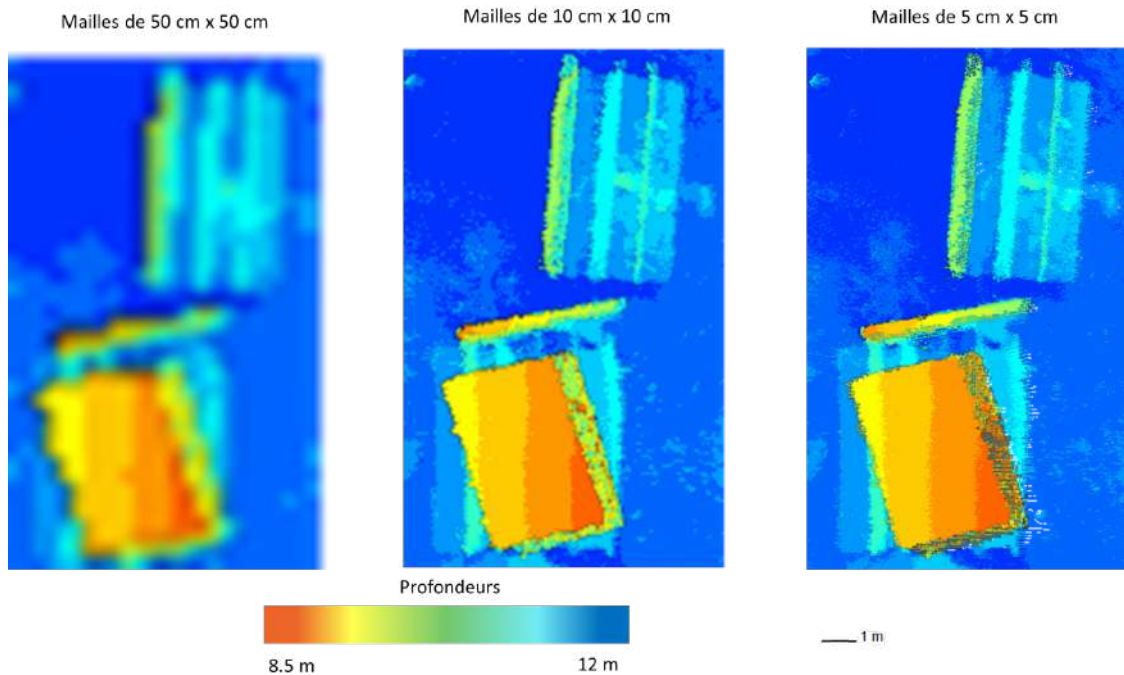


FIGURE 2.12 – L'ensemble des sondes sont distribuées dans des mailles carrées de différentes tailles, de gauche à droite, 50 cm, 10 cm et 5 cm.

Cependant, les valeurs de réflectivité vont varier avec la fréquence des pings et avec l'angle d'incidence. Il existe des méthodes de calibration.

A quoi ça sert ? : Les images de réflectivité peuvent servir brutes pour la détection d'objets artificiels (épaves, mines, canalisations, etc.). Mais elles sont aussi utilisées pour réaliser des cartes de classification des fonds ou des habitats marins.

Comment les traiter ? : Les images de réflectivité ne suffisent souvent pas à elles-seules à réaliser une carte des habitats benthiques, elles sont généralement couplées aux données bathymétriques et à des données de vérité terrain (photos et vidéos acquises lors de plongées etc.) pour l'identification des habitats. La bathymétrie et la réflectivité sont alors dérivées pour obtenir des caractéristiques (dureté, rugosité, forme, etc.) soit en étudiant le signal acoustique autour de la détection du fond soit en étudiant les MNT de bathymétrie et de réflectivité.

2.3.3 Données de la colonne d'eau

Qu'est-ce que c'est ? : Les données ou les images de la colonne d'eau, dites *water column imagery* (WCI) sont les échos de réflectivité sur tout le trajet des faisceaux de la surface jusqu'à la fin de réception du ping (et non juste autour du fond comme pour la réflectivité). Chaque ping produit alors une image (une donnée bi-dimensionnelle) avec autant de colonnes que de faisceaux et avec un nombre de lignes en fonction du temps d'acquisition (écho-range) et de la fréquence d'échantillonnage du signal. Les images sont volumineuses puisque dans le cas du R2Sonic 2022, avec un fond plat à 20 m de profondeur et une ouverture de fauchée de 130°, l'image contient 256 colonnes et environ 7000 lignes. Celle de l'épave du voilier, fig 2.14, par 42m de profondeur atteint 15000 lignes, dont les deux derniers tiers ne contiennent pas d'information notable.

A quoi ça sert ? : Ces images permettent de s'intéresser à tout ce qui est présent entre le fond et la surface : épaves posées sur le fond (fig 2.14), bancs de poissons (fig. 2.15), planctons,

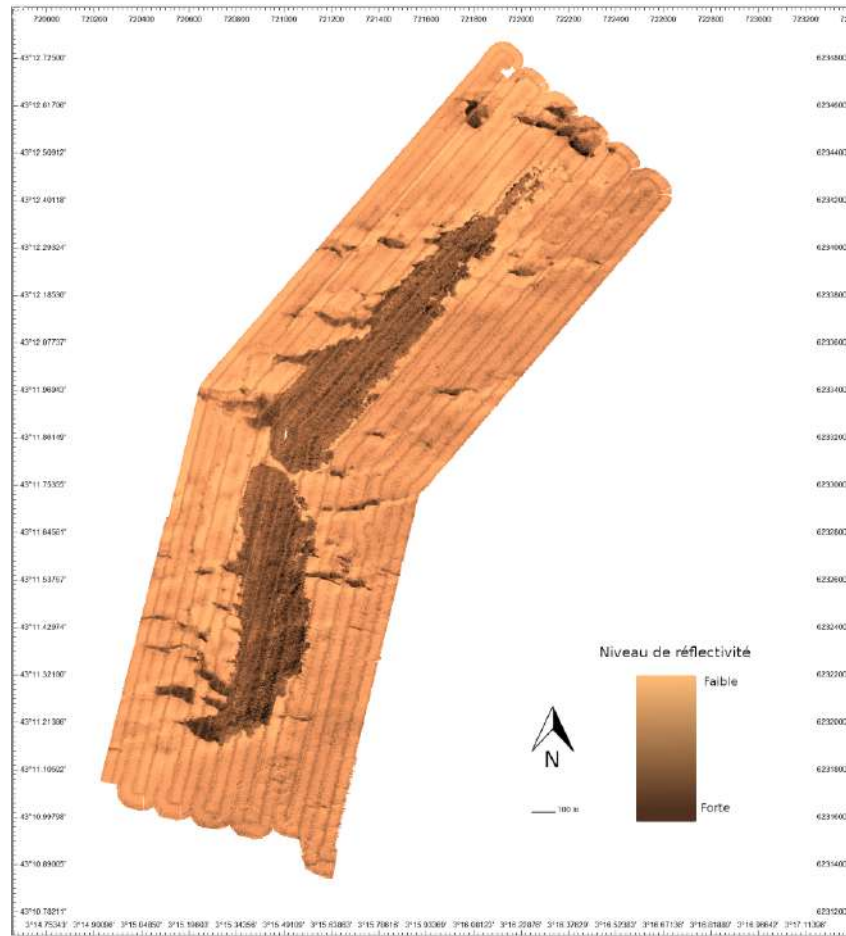


FIGURE 2.13 – Image de réflectivité d'une roche sur sable près de Vendres (Hérault). Les sédiments ont une moins forte réflectivité que les roches.

cordage (fig. 2.16 en bas), bouée (fig. 2.16 droite), échappement de bulles d'un plongeur (fig. 2.16 gauche), couche d'eau, etc. De très nombreux domaines d'application ont été trouvés aux images de la colonne d'eau. Notamment dans le domaine de l'halieutique où les estimations de biomasses permettent d'évaluer les stocks de pêche. De même, les monts hydrothermaux sont repérables à l'aide de leurs échappements gazeux, l'étude des bulles dans la colonne d'eau permet d'estimer leur débit par exemple (Colbo et al.; 2014).

Comment : La principale difficulté est l'acquisition des énormes quantités d'images de la colonne d'eau, surtout par petits fonds (grande cadence de pings) : les images en taille brute satureront les acquisitions et les espaces de stockage. De plus, seule une partie de l'image est peu bruitée, le reste contient les interférences des faisceaux qui rebondissent sur le fond.

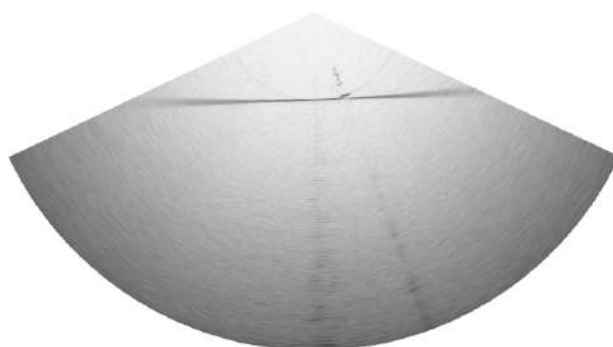


FIGURE 2.14 – Image water column avec l'épave d'un voilier, en baie de La Ciotat par 42m de profondeur.



FIGURE 2.15 – Deux exemples de bancs de poissons. A gauche, un banc de poissons au dessus d'une roche à 30m de profondeur. A droite, un banc en plein eau à 10m de la surface et du fond.

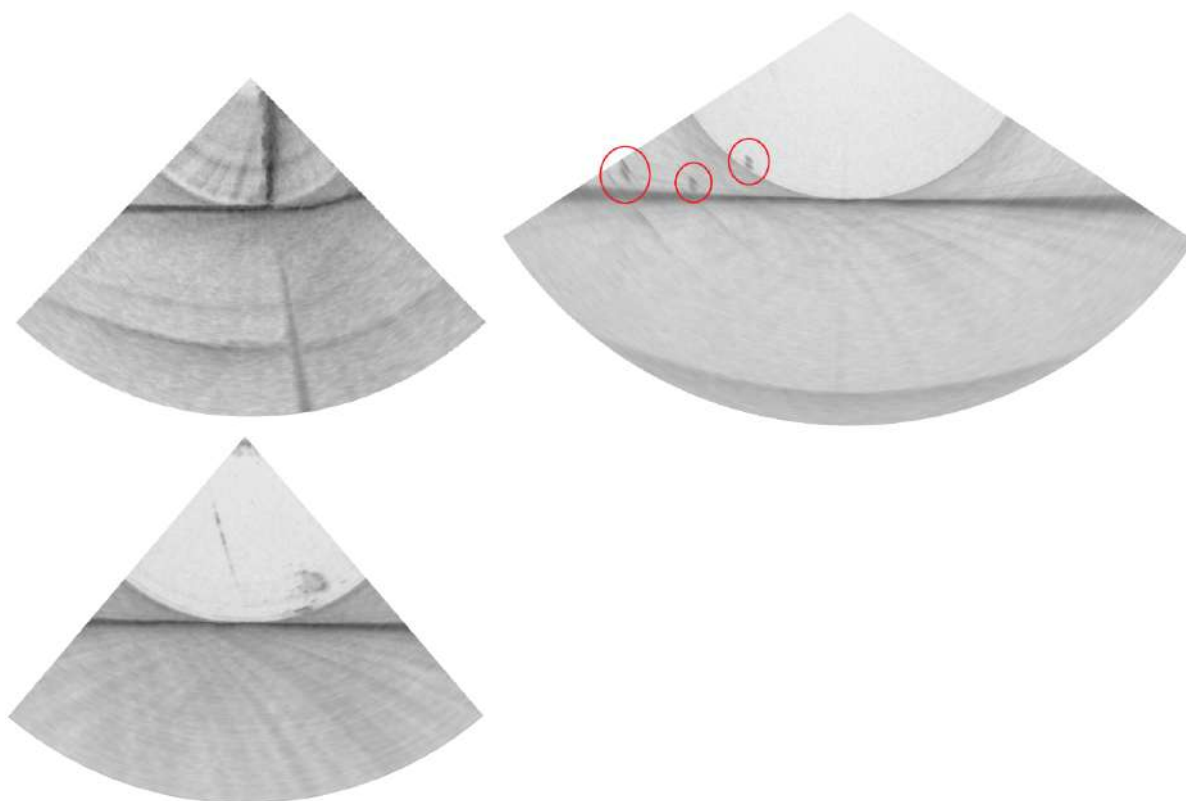


FIGURE 2.16 – Travaux sous-marins de Seaviews. En haut à gauche, échappement de bulles d'un plongeur. En haut à droite, balisage sous-marin contenant de l'air, donc visible par acoustique. En bas à gauche, bout tenu en surface par un flotteur, quelques poissons à droite.

Conclusion

Cette première partie, en deux temps, nous a permis de présenter :

- Les principales notions théoriques nécessaires à la connaissance et à la maîtrise des sondeurs multi-faisceaux et des données dont ils permettent l'acquisition. Les rappels en acoustique sous-marine permettent de bien comprendre les contraintes du milieu auxquelles nous allons être confrontés : prise en compte des paramètres physiques, portée des signaux émis, trajectoires des ondes acoustiques. Tandis que les rappels en traitement d'antenne permettent de comprendre le principe général du fonctionnement des instruments acoustiques utilisés : directivité des signaux, principaux réglages en fonction de l'environnement.
- Les outils utilisés et les données réelles manipulées tout au long de ces travaux de thèse. La chaîne complète d'acquisition et de traitement est intégralement maîtrisée par Seaviews, il est donc intéressant de présenter les instruments, les utilisations quotidiennes du matériel et les difficultés lors de l'utilisation de systèmes aussi complexes. L'ensemble des données acoustiques est traité dans la suite de logiciels développés en interne au sein de Seaviews. Seront surtout mentionnés dans la suite, le logiciel de gestion des données SMF ping à ping *ViewSMF* et le logiciel d'information géographique, *ViewMap*. Il existe des logiciels commercialisés pour l'acquisition, le traitement et la mise en forme des données SMF, tels que Hypack ou Qinsy (QPS) ; Seaviews ne dispose pas de licence pour ce type de logiciels.

La suite du manuscrit s'organise en fonction des applications que nous allons traiter et des méthodes qui en découlent.

La partie **II** concerne des applications de cartographie bathymétrique où l'examen des données est nécessaire pour trier les erreurs. D'une part, nous utiliserons la redondance des sondes bathymétriques pour détecter et neutraliser les erreurs aberrantes dans un MNT. D'autre part, nous utiliserons l'imagerie de la colonne d'eau pour accéder à l'information bathymétrique qui s'y trouve.

La partie **III** concerne des applications de cartographie de la nature des fonds. L'imagerie de réflectivité du fond est utilisée avec des données dérivées de la bathymétrie pour nous permettre de nous intéresser à la nature et à la morphologie des fonds. En opportunité, des images aériennes et de bathymétrie lidar peuvent venir en complément. Il devient en effet de plus en plus courant de mêler des données issues de divers systèmes pour l'étude des fonds car l'acoustique seule n'est pas toujours suffisante.

Deuxième partie

Cartographie bathymétrique

Introduction

En mesure bathymétrique par acoustique, la tendance est à l'augmentation de la quantité de données acquises et ce pour trois raisons principales. La multiplication du nombre de faisceaux d'un sondeur est le facteur le plus évident. Les premiers SMF des années 1970-1980 possédaient une vingtaine de faisceaux, aujourd'hui c'est au moins dix fois plus, quand il ne s'agit pas d'un millier de sondes par fauchées (déduction de plusieurs sondes par faisceau). Ensuite, les sondeurs se sont miniaturisés et ont pu être embarqués sur de plus petits navires, leurs fréquences de fonctionnement se sont largement élevées (de 36KHz à +400kHz) pour permettre de mesurer les faibles profondeurs et la fréquence d'émission des pings a elle aussi crû pour détailler au mieux ces fonds (jusqu'à 60Hz avec le R2SONic 2022). Enfin, la technologie des SMF s'est largement diffusée en 50 ans dans le monde et à différentes catégories de professionnels (pêche commerciale, recherche scientifique, ingénierie environnementale, travaux sous-marins, etc.). Ce net accroissement en données est clairement visible dans les projets de mutualisation des données hydrographiques comme *European Marine Observation and Data Network* (EMODnet) ou *The General Bathymetric Chart of the Oceans* (GEBCO) au niveau mondial pour les rendre accessibles à tous.

Toutefois, une part non négligeable des sondes acquises est erronée et un tri préalable de ces sondes est impératif pour générer les cartes marines assurant la sécurité des navigants. La qualité du jeu de sondes retenue influe directement sur la qualité des produits cartographiques. Ainsi, avec des enjeux tels que le transport maritime ou la pose de câble de télécommunications, ces produits sont régis par des normes comme la norme S-44 (6^e édition) de OHI (2020) portant sur la réalisation de levés hydrographiques. Avec les évolutions des SMF, le taux d'erreurs s'est affaibli et les techniques de tri des sondes se sont diversifiées. Manuelles dans les premiers temps et déjà extrêmement chronophages avec des quantités moindres de données, la norme S44 de l'OHI recommande maintenant des outils d'automatisation. Il existe donc d'ores et déjà une littérature dense sur le sujet ainsi que des méthodes intégrées dans des logiciels commercialisés. Elle sera évoquée au chapitre 3.

Seaviews n'échappe pas à cette règle et effectue ce tri à l'aide de ses logiciels développés en interne. La technique employée relève du filtrage par masques interactifs qui sélectionne les sondes supposées aberrantes et attend la validation d'un opérateur. Ce travail reste chronophage et répétitif car les pings sont traités un à un. Nous proposons donc dans le chapitre 3 une méthode automatisée et rapide effectuant un tri spatialisé des sondes, plutôt que temporel.

D'autre part, Seaviews acquiert systématiquement les images de la colonne d'eau (WCI), lesquelles s'apparentent à des données brutes des pings. Celles-ci servent durant le tri en cas de doute sur une ou des sondes à garder ou à neutraliser et sont conservées ensuite pour d'autres applications (cf. les annexes A et B). La pratique du filtrage en gardant un œil sur ces images montrent qu'elles ne présentent pas les erreurs aberrantes de la bathymétrie. Il est alors tentant de se servir des WCI pour en extraire une nouvelle série de sondes bathymétriques sans défaut. C'est l'approche qui sera développée au chapitre 4 par l'utilisation des techniques modernes de réseaux de neurones à apprentissage profond.

Chapitre 3

Traitements des sondes bathymétriques

Dans ce chapitre, nous dressons un rapide panorama des principales méthodes employées pour trier les erreurs bathymétriques, présentons la méthode classiquement utilisée à Seaviews et proposons une méthode complémentaire à la première. Ce chapitre permet aussi d'introduire la manipulation de ces données ping-à-ping et de manière spatiale pour la suite du document.

3.1 État de l'art du filtrage des fausses sondes bathymétriques

La détection d'erreurs aberrantes (*outlier data*) n'est pas un problème récent ni cantonné à la bathymétrie par SMF. C'est un sujet récurrent qui dépasse largement le domaine hydrographique et s'étend à bien d'autres d'applications tel que par exemple, dans la pose d'un diagnostic médical (Deneshkumar et al.; 2014), dans la détection de fraude fiscale (Savić et al.; 2022) ou d'intrusion dans un système (Beulah and Punithavathani; 2015). Il existe alors une littérature générale et régulièrement mise à jour, telle que Aggarwal (2016), et une plus spécialisée au domaine hydrographique.

Le contenu de cette section se base sur deux revues, Artilheiro (1998) et Le Deunf et al. (2020), et sur les connaissances acquises à Seaviews. La première revue, avec une vingtaine d'années de recul sur l'utilisation des SMF, fait un point sur les procédures d'acquisition et de traitement des données bathymétriques issues de SMF. La seconde, très récente, propose une classification des méthodes automatiques de tri des fausses sondes.

3.1.1 Caractérisation des sondes erronées et des méthodes de tri

Une définition, assez générale, du *outlier* et souvent utilisée est celle de Hawkins (1980) :

“An outlier is an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism.”

Cette définition permet de différencier ces erreurs “grossières” des erreurs systématiques et des erreurs aléatoires (Artilheiro; 1998). L'erreur systématique est définie par l'OHI (2020) comme “composante de l'erreur de mesurage qui reste constante ou varie de manière prévisible” et doit être évitée par une procédure de calibration du SMF et des autres instruments. L'erreur aléatoire est définie comme étant un “bruit dans un mesurage causé par des facteurs qui varient d'une mesure à l'autre”, elle n'est pas ou peu compensée mais il faut en tenir compte pour estimer l'incertitude de mesure. Dans la suite de ce manuscrit, ce sont essentiellement les erreurs aberrantes qui sont visées.

Les sondes aberrantes viennent d'erreurs de détection du fond marin dans l'écho du signal. Plusieurs causes externes sont bien connues comme :

- des objets présents dans la colonne d'eau (bouts, bancs denses de poisson, forêts de laminaires, bulles d'air etc.),
- des profils de célérité complexes (fort gradient de température ou de salinité entre deux couches d'eau),
- des réflexions multiples des rayons acoustiques,
- la présence d'autres instruments acoustiques utilisant les mêmes plages fréquentielles dans la zone d'étude.

A cela, il faut ajouter des causes internes venant du matériel, effet des lobes secondaires et sensibilité des faisceaux aux grands angles, sans parler de dysfonctionnements occasionnels des instruments. Une partie de ces erreurs est évitée par le réglage en temps réel du sondeur lors de l'acquisition : l'ordinateur de contrôle du SMF permet le réglage de ses paramètres d'acquisition (signal d'émission, ouverture angulaire du ping, etc.), mais aussi d'une "fenêtre" de recherche préférentielle du fond dans le signal (R2Sonic; 2014). Mais ces réglages ne remplacent en aucun cas le filtrage post-acquisition des sondes.

D'après Artilheiro (1998), les méthodes de filtrage doivent répondre aux attentes suivantes. Il est préférable que les méthodes neutralisent des sondes plutôt qu'elles ne les suppriment définitivement, c'est-à-dire que le traitement soit facilement réversible sans détruire le jeu initial de sondes. Ensuite, les méthodes doivent être suffisamment efficace pour que le nettoyage ne soit pas plus long que le temps d'acquisition, ce qui exclut d'emblée les méthodes totalement manuelles. Enfin, il faut que les méthodes restent faciles d'utilisation et intelligibles. Ce dernier point n'est plus aussi important aujourd'hui car les méthodes se sont complexifiées et doivent pouvoir traiter de très grandes quantités de données avec peu d'intervention.

Artilheiro (1998) et Le Deunf et al. (2020) proposent chacun un classement des méthodes, complémentaire l'un par rapport l'autre. Une synthèse peut être la suivante :

- différencier d'abord l'analyse subjective de l'objective. L'analyse subjective correspond essentiellement à l'utilisation d'outils visuels pour repérer et invalider les sondes aberrantes, tandis que l'analyse objective est tournée vers des méthodes automatiques, dont les résultats sont reproductibles indépendamment d'un expert. Cette deuxième analyse est préférée, mais la première permet une validation finale car la responsabilité de la décision repose sur l'expert.
- discerner les approches orientées données de celles orientées modélisation de surface, plus récentes (début des années 2000). Bien que ces deux approches soient bien représentées dans la littérature, les approches orientées données sont plus anciennes et nombreuses (Le Deunf et al.; 2020) ; les approches par surface ne sont pas mentionnées dans Artilheiro (1998).
- distinguer le type de supervision. Les méthodes sont actuellement quasiment toutes non-supervisées, c'est-à-dire que les algorithmes sont capables de s'ajuster aux données pour détecter les erreurs sans *a priori* sur celles-ci. A l'inverse, une méthode supervisée utilise des résultats pour s'ajuster au problème et être capable de généraliser sur de nouvelles données.

3.1.2 Exemples de méthodes par catégorie

3.1.2.1 Filtrage par masques interactifs

Le tri des sondes bathymétrique est effectué dans le logiciel *ViewSMF* de Seaviews, spécifique à la lecture, l'écriture et au traitement des fichiers SMF (et de position, attitude, etc.). L'interface du logiciel comprend (entre autres) un affichage de la donnée sous forme de ping - soit uniquement les sondes bathymétriques (cf figure 2.11 à gauche), soit uniquement l'image de la colonne d'eau d'un ping (comme sur les figures de 2.14 à 2.16), soit les deux superposés (comme les figures 3.1 et 3.2) - et une fenêtre avec les masques et leurs réglages. Plus d'une dizaine de masques sont disponibles, mais deux d'entre eux sont utilisés systématiquement : un masque d'inclusion asservi et un masque d'exclusion. Le premier se base les sondes valides du ping précédent pour déterminer une zone de recherche où, s'il y en a, les sondes seront également valides et, en dehors de cette zone, les sondes sont à neutraliser sur accord avec l'opérateur (figure 3.1). Comme ce masque ne suffit pas toujours (figure 3.2 à gauche), le second masque est une zone rectangulaire de taille modulable

que l'utilisateur déplace à la souris, il neutralise les sondes dans le rectangle (cadre bleu sur la figure 3.2 à droite). En cas de doute sur une proposition d'un masque, l'opérateur peut s'aider de l'image de la colonne d'eau car elle permet la visualisation complète du signal écho.

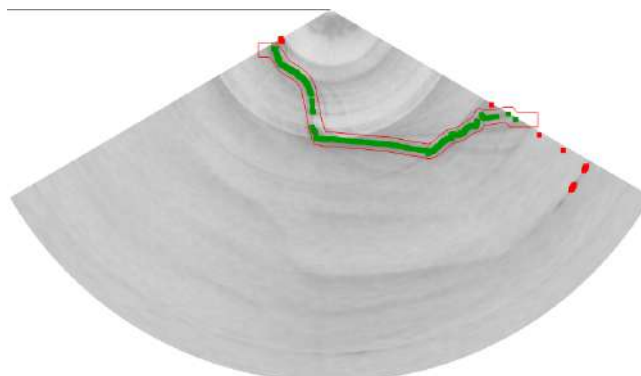


FIGURE 3.1 – Les sondes vertes sont considérées comme valides par le masque d'inclusion asservi (cadre rouge). Les sondes en rouge sont éliminées.

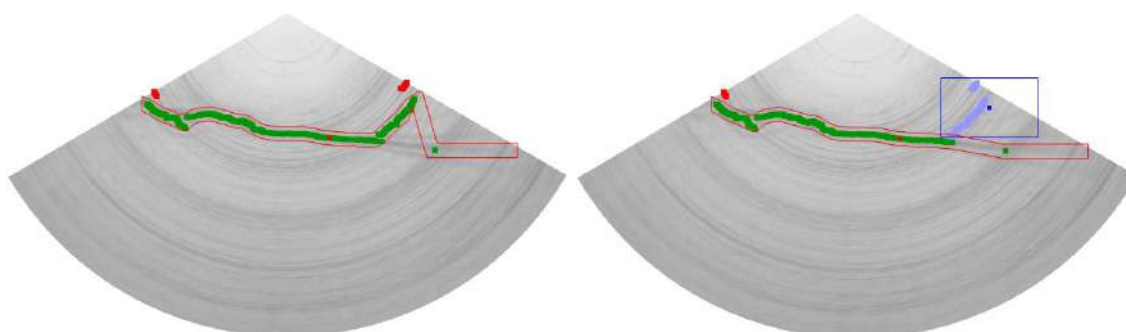


FIGURE 3.2 – A gauche, le masque asservi (cadre rouge) ne suffit pas à rejeter toutes les sondes aberrantes. A droite, le masque d'exclusion (cadre bleu) neutralise les sondes en bleu et le masque asservi se réadapte automatiquement.

Le traitement ping-à-ping est donc un long travail car les jeux de données contiennent de l'ordre de 10 000 pings par hectare et bien plus en cas de petits fonds. Le traitement peut être accéléré en faisant défiler les pings jusqu'à ce qu'un masque éjecte une sonde et demande une validation à l'opérateur.

3.1.2.2 Méthodes à approche orientée données

Le principe du *Hydrographic Data Cleaning System* (HDCS) est proposé par Ware et al. (1992) et est intégré dans la suite logiciel *HIPS and SIPS* de Teledyne Caris. L'algorithme utilise une moyenne et un écart-type mobiles pondérés calculés pour toutes les sondes d'un sous-ensemble sélectionné par un opérateur. Un champ d'influence décroissant autour de la sonde d'étude, par une pondération équivalente à l'inverse de la distance à cette sonde, limite la quantité de sondes voisines entrant en compte dans les calculs. Ces moyennes et écart-types permettent de classer les sondes en huit niveaux pour ne conserver que celles qui correspondent à un bon niveau de confiance et d'élargir ou de restreindre à un niveau proche en cas de besoin (par exemple, si les données resservent dans une autre application avec un ordre selon l'OHI (2020) moins contraignant).

Du et al. (1996) ont bâti une méthode de partitionnement des données (*data clustering*) sur l'expertise des hydrographes à prendre les décisions. Le processus de filtrage est réalisé en deux étapes successives, en étudiant un lot de 10x10 sondes (10 faisceaux contigus sur 10 pings consécutifs),

puis avec un lot réduit à 3x3 sondes. Pour chaque lot, un histogramme des valeurs des sondes est tracé de manière à identifier des valeurs aberrantes, trop éloignées du groupe des bonnes valeurs. Enfin, chaque sonde est comparée avec les six sondes voisines du même ping par un test statistique (type test de Dixon) et rejetée si elle est trop éloignée des extrêmes.

La méthode *Combined Uncertainty and Bathymetry Estimator* (Calder and Mayer; 2003), dite CUBE, a été largement utilisée par des organismes étatiques d'hydrographie et est disponible dans plusieurs logiciels commerciaux dont ceux de Teledyne Caris. Plutôt que d'être appliqué sur des petites quantités de sondes, la méthode fonctionne sur des sondes bathymétriques projetées dans un modèle maillé bathymétrique (voir section 3.2.1). A chaque nœud de maille, est estimée une profondeur *vraie*, à partir des sondes voisines, et son indice de confiance. Pour améliorer cet indice, un algorithme est mis en oeuvre afin de proposer plusieurs hypothèses de ce qui est pourrait être une profondeur plus plausible. L'hypothèse retenue dépend du contexte, en se référant aux nœuds voisins sans ambiguïté d'hypothèse. Sur des fonds chaotiques (rocheux par exemples), l'algorithme n'est plus si robuste, ni autant capable de trier les hypothèses et une intervention humaine devient alors nécessaire. Pour pallier ce défaut, une version améliorée de CUBE avec des modèles à multi-résolution de mailles évite des imbroglios dans les hypothèses (Calder and Rice; 2011).

3.1.2.3 Méthodes à approche orientée surface

Le principe repose sur la modélisation - locale ou globale - des fonds et est capable de rejeter les erreurs grâce à la comparaison de la modélisation avec la réalité. La modélisation globale recherche une surface de tendance et le modèle peut être élaboré car il doit rendre compte de la variabilité morphologique des fonds. *A contrario*, la modélisation locale peut être réalisée par une forme polynomiale de degré compris entre 0 et 3 (Le Deunf et al.; 2020), mais nécessite de partitionner convenablement les fonds selon leurs variations.

Debese et al. (2012) réalisent une modélisation locale par une fonction polynomiale. Pour rendre robuste la régression face aux *outliers*, un M-estimateur est employé, il s'agit d'une version plus robuste de la technique des moindres carrés : moindres-carrés repondérés (*Iteratively Reweighted Least Squares*). Une fois la fonction polynomiale ajustée, les erreurs aberrantes se voient affecter une pondération très faible par rapport aux autres sondes, ce qui permet de les identifier. Et, afin d'avoir une partition adaptée à des fonds variables, la zone globale se subdivise de manière récursive en quatre cellules équivalentes (arbre quaternaire) jusqu'à avoir un contenu de cellule modélisable par une fonction polynomiale.

Bottelier et al. (2005) utilise une technique d'interpolation robuste pour produire une modélisation globale d'une zone d'étude. La technique d'interpolation estime une valeur de profondeur en recensant les valeurs voisines et en les pondérant par une fonction de covariance gaussienne. Un M-estimateur est également employé pour diminuer l'influence des *outliers* (poids plus faible) par itérations. Une fois la surface de tendance obtenue, celle-ci est comparée aux données réelles (par ping ou par groupe de pings), et les sondes trop éloignées de l'estimation sont alors rejetées.

3.1.2.4 Méthode à traitement supervisé

Jusqu'à présent ce sont les méthodes automatiques non-supervisées qui ont été présentées, car elles concernent la grande majorité des méthodes existantes. Un traitement supervisé sous-entend que l'opérateur soit capable de fournir un jeu de données déjà triées à un algorithme qui va pouvoir en déduire ce qu'il faut faire sur des nouvelles données non triées. L'intérêt est qu'une fois que l'algorithme est ajusté, le traitement est effectué très rapidement, durablement et en autonomie.

Stephens et al. (2020) proposent d'entraîner par apprentissage profond un réseau de neurones (à convolutions, cf chapitre 4) à partir de sondes bathymétriques sous forme de nuages de points en trois dimensions. Pour cela, plusieurs jeux de données bathymétriques sont triés une première fois (semi-)manuellement pour alimenter l'apprentissage. Ensuite, les sondes géo-référencées en nuage de points sont réparties dans des "voxels", c'est-à-dire dans des cubes d'une taille standardisée, que le réseau prend en entrée. Cette méthode, très récente, présente des résultats encourageants. Il est probable que d'autres études par apprentissages profonds suivront rapidement.

3.2 Filtrage spatial automatique des sondes bathymétriques

La méthode de filtrage interactif de Seaviews est une technique opérant mode temporel (traitement ping-à-ping). Elle est peu automatisée, au regard des nombreuses méthodes mises au point et réalisant des traitements en mode spatial (sur un sous-ensemble de sondes ou sur toutes les sondes). Nous proposons une nouvelle méthode de filtrage, automatisée et intégrée dans le logiciel de SIG (Système d'Information Géographique) de Seaviews, *ViewMap*. Elle se déroule en trois grandes étapes, chacune détaillée ci-après :

- la mise sous un format de modèle numérique de terrain (MNT) des sondes bathymétriques ;
- le tri des sondes dans les mailles du MNT ;
- l'inspection et la sauvegarde du résultat.

Un modèle numérique de terrain est une représentation de la topographie d'une surface sous une structure de données adaptée aux calculs par ordinateur. Les données sont géopositionnées et organisées au niveau de la structure élémentaire du MNT, à savoir, la "maille".

3.2.1 Modèle numérique de terrain

La première étape consiste à charger les données bathymétriques de *ViewSMF* vers *ViewMap* dans un modèle numérique de terrain. En effet, dans la configuration des logiciels de Seaviews, il n'est pas prévu de lire directement les pings dans le SIG, car il faut d'abord les coupler à l'aide de *ViewSMF* aux données de célerimétrie, d'attitude et de positionnement (figure 3.3). Par ailleurs, dans *ViewMap*, il faut spécifier deux caractéristiques pour créer un MNT : le type de projection cartographique et la taille de la maille.

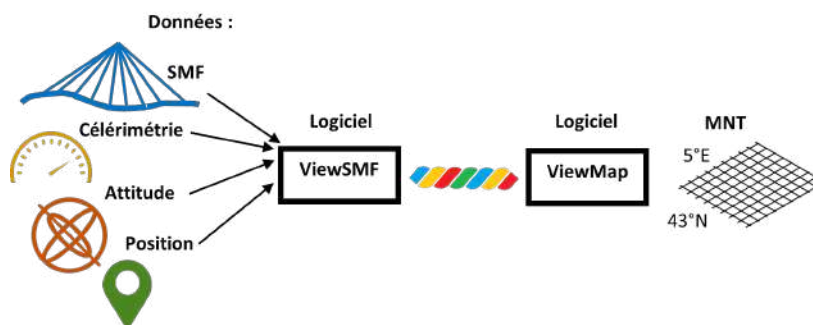


FIGURE 3.3 – Lecture de données issues du sondeur multifaisceaux (SMF), de la centrale d'attitude, du GPS, du célerimètre (section 2.2) dans *ViewSMF* et chargement dans un modèle numérique de terrain (MNT) dans *ViewMap*.

La projection cartographique est un mode de représentation d'une surface non plane (comme la Terre) sur une surface plane (telle une carte papier) impliquant donc des distorsions lors de la "mise à plat". Il existe alors plusieurs types de projection ayant chacun ses avantages pour des usages particuliers. Par exemple, les cartes de navigation maritime emploient (ou employaient) la projection cylindrique de Mercator (figure 3.4 à gauche), conservant les angles utiles à la navigation par les astres mais déformant les surfaces ; le Groenland apparaît plus grand que l'Amérique du Sud. Les cartes de France emploient une projection conique de Lambert (figure 3.4 à droite), la projection Lambert 93. Elle conserve également les formes et les angles mais induit des déformations de distances (de l'ordre métrique) selon les latitudes. **Dans tout ce manuscrit, l'ensemble des MNT produits et des cartes réalisées utilisent une projection Lambert 93.**

Le maillage d'un MNT fait partie de ses caractéristiques intrinsèques. Si plusieurs formes de mailles existent, telle que mailles hexagonales ou triangulaires, seules les mailles carrées sont utilisées à Seaviews. Elles permettent une structuration des données, proche d'un tableau, plus simple qu'avec une autre forme géométrique. L'autre caractéristique qui nous importe est la taille de la maille, c'est-à-dire la résolution du MNT. Comme les MNT définis dans *ViewMap* ne possèdent

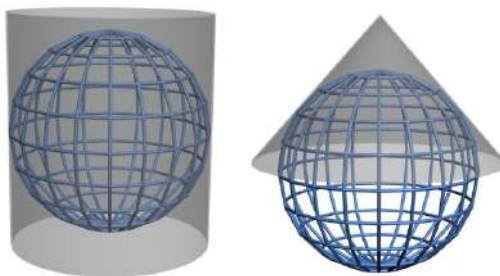


FIGURE 3.4 – A gauche, une projection cylindrique tangente à l'équateur. Les projections cylindriques peuvent également être tangentes aux pôles ou à n'importe quel autre point de la Terre. A droite, une projection conique tangente où le sommet du cône est centré sur le pôle nord. D'autres projections coniques existent avec des variations sur l'emplacement du sommet ou encore avec un cône sécant (et non tangent) à la Terre. (source des images : Wikipedia)

qu'une résolution unique pour l'ensemble des mailles, il convient donc d'adapter la taille selon les besoins - cartographie générale d'une zone ou cartographie fine d'une digue ou d'une canalisation - ou selon l'étendue de la zone car une résolution élevée implique de nombreuses mailles et donc de plus longs calculs. Usuellement, à Seaviews, les tailles de mailles employées vont varier entre 0.2 m pour une cartographie fine à 1 m pour une zone de plusieurs kilomètres carrés.

Les sondes ont une position géographique très précise grâce à l'ensemble des données acquises (figure 3.3), elles sont projetées dans les mailles du MNT. Une fois l'ensemble chargé, les mailles contiennent un nombre de sondes qui varie selon l'emplacement de la maille dans le MNT (au bord ou au centre), selon la densité des données (profondeur d'étude, réglage du sondeur). **Pour la visualisation, on choisit généralement de ne représenter la maille que par une unique valeur telle que la moyenne des sondes présentes - c'est le choix fait au niveau de l'ensemble des MNT en illustration dans ce manuscrit. Les valeurs de profondeurs sont données avec une valeur positive.** Mais pour les traitements à suivre, les sondes des mailles sont accessibles.

Structure d'un MNT Le code du logiciel *ViewMap* est doté d'un grand nombre de structures pour la manipulation des MNT selon les opérations à réaliser. Nous utilisons pour la méthode décrite ci-après un format dit "brut" car permettant un accès facile à l'ensemble des données. Il s'agit en réalité d'une imbrication de structures sur trois niveaux (ligne, maille, sonde sur la figure 3.5). Ce MNT brut est un *deque* de lignes, c'est-à-dire une liste à double entrées (*double-ended queue*) : il est possible d'ajouter ou d'enlever des lignes au début ou à la fin de cette liste. Ces lignes sont indépendantes les unes des autres : elles n'ont pas la même origine, pas la même longueur (voir figure 3.5). Une ligne est également un *deque* de mailles. Une maille est un vecteur de sondes. Si une maille doit recevoir des sondes supplémentaires, elles sont ajoutées en fin de liste car l'ordre n'est pas représentatif. Enfin, une sonde contient plusieurs informations : sa valeur de profondeur et ses origines (fichier, numéro de ping, numéro de sonde dans le ping, angle d'émission) et un booléen sur sa validité. De ce fait, toutes les sondes - qu'elles soient valides ou non - sont projetées dans le MNT et en cas d'invalidité, ces sondes ne participent pas au calcul des moyennes des mailles ou à tout autre type de traitement.

3.2.2 Traitement maille par maille

La méthode que nous allons maintenant décrire ne supprime pas irrémédiablement les sondes mais les neutralise au travers d'un booléen de validité. Deux étapes sont réalisées successivement. La première consiste à repérer les mailles dont les sondes semblent valides pour qu'elles participent à la correction des mailles voisines dans lesquelles des erreurs sont présentes.

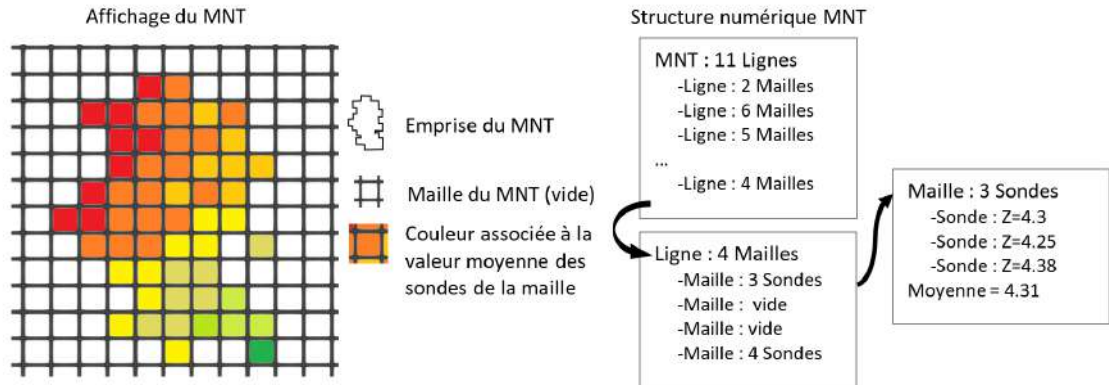


FIGURE 3.5 – Exemple d’un affichage à l’écran d’un MNT : des couleurs artificielles permettent une visualisation synthétique des valeurs. D’un point de vue structure numérique, un MNT ressemble à un tableau bidimensionnel de valeurs, même si la forme réelle est quelconque.

Déterminer les mailles valides Durant cette phase, l’ensemble des mailles sont passées en revue. En tout premier lieu, les sondes “empilées” sans ordre dans la maille sont ordonnées en profondeurs croissantes en vue des traitements à venir. Selon le contenu de la maille et sa résolution, celle-ci est invalidée si : (i) elle contient moins de 10 sondes *OU* (ii) l’écart type des sondes dépasse la résolution de la maille *OU* (iii) si l’étendue dépasse deux fois la résolution. La condition (i) suppose qu’une maille avec aussi peu de sondes se trouve dans une partie de la zone peu insonifiée (bordure de la zone, relief particulier), si ça n’est pas le cas, il est possible que la résolution du MNT ne soit pas adaptée à la densité des sondes. Les conditions (ii) et (iii) sont des indicateurs de dispersion des sondes au sein d’une maille. La comparaison à la résolution de la maille est un choix empirique, mais dans une taille de maille la variabilité est assez limitée : même un fond très en pente atteint rarement une pente de 1 pour 1 et le cas d’un relief accidenté mérite une analyse plus profonde de la maille.

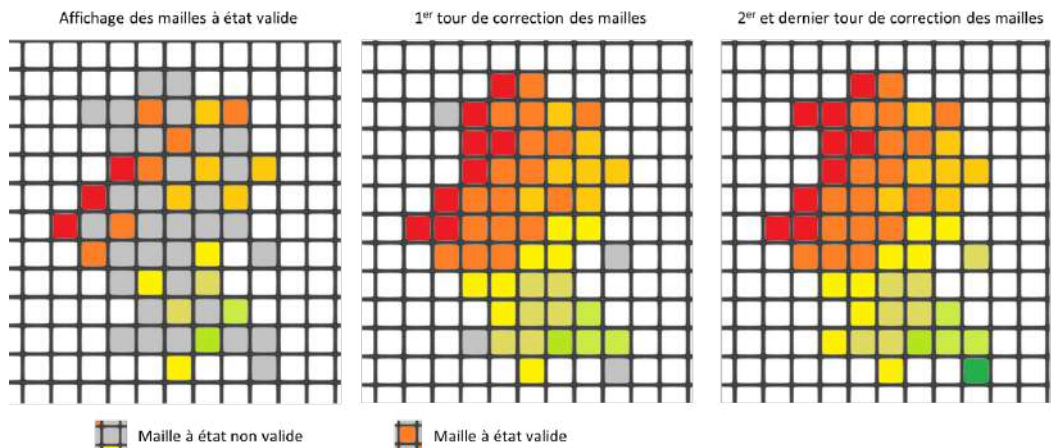


FIGURE 3.6 – A gauche, seule les mailles valides sont encore visibles, les autres sont grisées en attente de correction. Au milieu et à droite, durant la correction, les sondes des mailles voisines d’une maille à état valide sont triées et passent à un état valide. Ce MNT est corrigé rapidement en deux passes.

Corriger les mailles non valides Durant cette phase de nettoyage, les mailles du MNT vont être passées en revue de nombreuses fois jusqu’à ce qu’elles prennent toutes un état valide. En

effet, la correction se fait par propagation de proche en proche : quand une maille à état valide est rencontrée, ses voisines contiguës (jusqu'à 8) sont interrogées sur leur état de validité et en cas de réponse négative, un tri des sondes est opéré. La moyenne des profondeurs des sondes de la maille valide, notée μ_m , sert de repère pour traiter la maille voisine.

La manière d'effectuer le tri dépend ensuite du contenu de la maille.

- Si celle-ci contient peu de sondes (moins de 10), la décimation n'a lieu que si l'écart type dépasse 0.7. Dans ce cas, les sondes extrêmes sont neutralisées une à une jusqu'à passer sous 0.7 d'écart type. Le choix de la sonde extrême à neutralisée est déterminé par son éloignement à la profondeur μ_m de la maille de référence.

Dans le cas où une maille contient plus de sondes, la décimation peut être plus forte.

- Si cette maille a un écart type dépassant 0.5, il y a une assez grande probabilité pour qu'il y ait deux groupes de valeurs dans les sondes. Auquel cas, on cherche une séparation entre les deux groupes, on garde le groupe le plus proche de μ_m et si besoin, on neutralise des sondes extrêmes jusqu'à faire tomber l'écart type du groupe sous 0.7.
- Dans le cas, où l'écart type est déjà inférieur à 0.5, on enlève les sondes en dehors des quartiles de 1 à 3 à condition que cela fasse baisser l'écart type des sondes restantes. En fin de traitement, la maille prend un état valide et au prochain parcours du MNT, elle servira à corriger ses voisines.

Dans le cas où l'un de ses trois points n'a pas permis d'améliorer le filtrage de la maille par rapport à sa voisine considérée comme valide, l'ensemble de ses sondes sont invalidées.

3.2.3 Vérification

Deux solutions visuelles et rapides sont possibles, elles permettent de chercher les erreurs restantes mais la correction supplémentaire devra se faire à l'aide du filtrage par masques dans *ViewSMF*. La première, dans *ViewMap*, est d'afficher un ombrage des reliefs dans le MNT. Celui-ci, même fixé exagérément fort, met en avant les détails, les reliefs et les erreurs restantes. De même, l'autre solution consiste à exporter le MNT sous forme d'un modèle 3D. Celui-ci est lu dans un autre logiciel (comme Meshlab), manipulé sous tous les angles, la lumière orientée pour raser les surfaces, les erreurs restantes se présentent sous forme de pics alors très visibles.

Néanmoins, la solution la plus rigoureuse est de vérifier à l'aide des masques interactifs les pings un par un. Dans ce cas, il est possible d'utiliser le masque "asservi" et de faire défiler jusqu'à trouver une sonde en dehors.

3.3 Application sur une zone très bruitée

3.3.1 Définition de la zone d'étude et filtrage manuel

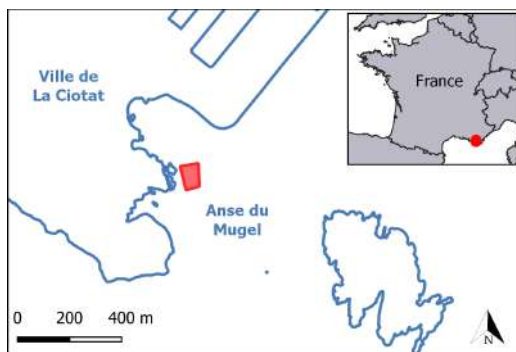


FIGURE 3.7 – Carte de situation du canyon d'étude dans l'Anse du Mugel à La Ciotat (France).

Dans cet exemple, la zone d'étude se situe dans l'anse du Mugel à La Ciotat, représentée en rouge sur la figure 3.7. Cette zone est un site de plongée sous-marine avec un petit canyon entre

deux roches remontant jusqu'à quelques mètres sous la surface sur un fond sableux en pente. Ce site a été levé en 2016, il s'agit de l'une des premières acquisitions de Seaviews avec le sondeur R2Sonic et les pings sont particulièrement bruités (voir celui de la figure 3.8 à gauche). Les raisons de cet état sont de fait (i) une puissance d'émission un peu trop faible couplée à une trop grande ouverture angulaire des pings, (ii) une portée trop courte (les sondes devraient être situées dans la moitié haute des WCI) et (iii) les fenêtres préférentielles de recherche du fond dans le signal n'ont pas été utilisées. Il s'agit d'un des jeux de données de Seaviews contenant le plus de fausses sondes.

L'acquisition a été réalisée en une heure au moyen d'une quinzaine de transects (de passages sur la zone). Le linéaire total de navigation est de 800 m sur une zone de 4800 m² (un rectangle d'environ 60 m par 75 m); il y a une forte donc redondance dans l'acquisition des 20 000 pings. Cette forte redondance ne suffit pas à lisser les erreurs les plus fortes, comme en témoignent le MNT et la vue 3D de la figure 3.8, et donne un aspect très rugueux aux fonds plus plats (couleur verte-turquoise correspondant à 10-11m de profondeur).

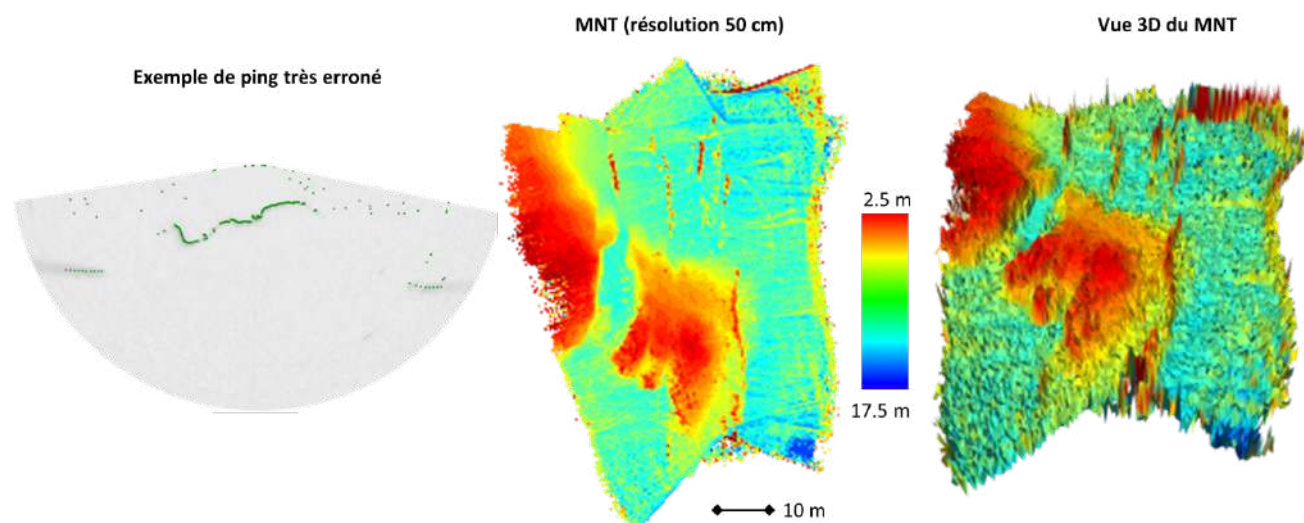


FIGURE 3.8 – Modèle numérique de terrain et vue 3D créés à partir de la donnée bathymétrique brute.

Du fait de l'extrême mauvais état des sondes, le filtrage manuel a nécessité six heures de travail. Le masque "asservi" a été utilisé autant que possible mais les boîtes d'exclusion et d'inclusion des sondes ont dû être largement employées - quasiment à chaque ping. Le résultat est présenté sous forme d'un MNT en figure 3.9, il est considéré comme le résultat de référence. La comparaison visuelle des MNT bruts et filtrés prouvent l'importance du filtrage; les fonds sableux sont plus profonds une fois nettoyés des erreurs (les couleurs de turquoise à bleu roi correspondent à 10-15m de profondeur)! La vue 3D ne présente plus les pics ni l'aspect rugueux.

3.3.2 Résultats par le filtrage automatique

Précédemment, les deux MNT - brut et filtré - avaient été représentés avec une résolution de 50 cm, tout à fait acceptable par rapport à la taille de la zone, des dimensions des reliefs et à la densité de sondes acquises. Comme le filtrage automatique est sensible à la quantité et à l'écart type des sondes dans les mailles, il convient de trouver une résolution adaptée : nous testons donc des résolutions de 1 m, 50 cm, 25 cm et 10 cm. Les mailles du centre de la zone sont naturellement plus denses que celles du bord. A la vue du tableau 3.1, la densité des sondes semble permettre de pouvoir accéder à une très haute résolution (10 cm).

Pour ces quatre résolutions, l'étape de sélection des mailles valides est drastique : très peu de mailles sont jugées bonnes au regard des conditions énoncées à la section précédente et plus la résolution est faible, moins il y a de mailles valides. Pour les trois résolutions - 1 m, 50 cm, 25 cm -, les mailles valides sont concentrées au même endroit, en haut à gauche, alors que pour la plus

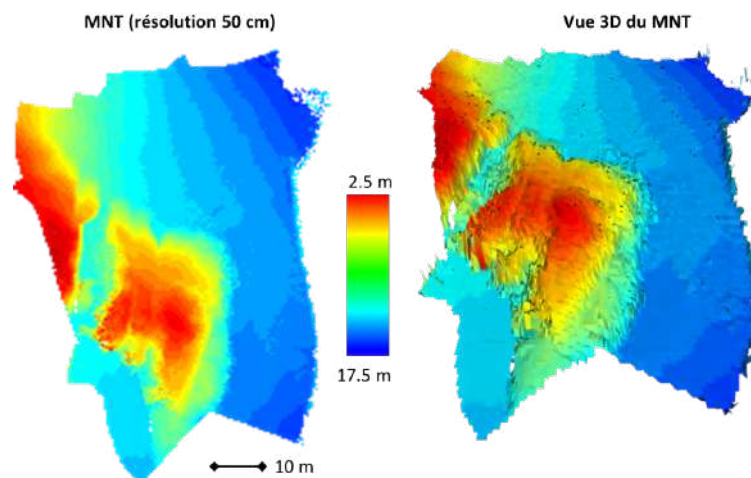


FIGURE 3.9 – Modèle numérique de terrain et vue 3D créés à partir des données bathymétriques filtrées manuellement.

TABLE 3.1 – Quantité approximative de sondes par maille selon leur emplacement dans le MNT.

Résolution	Centre		Bords	
	maximum	en moyenne	en moyenne	minimum
1 m	5000	1000	100	<10
50 cm	1200	200	50	<10
25 cm	400	80	20	<10
10 cm	60	15	<10	<10

grande résolution, les mailles sont davantage dispersées (figure 3.10).

Les résultats du nettoyage automatique des mailles sont fournis au niveau de la figure 3.11 pour les MNT vus du dessus et en figure 3.12 pour les vues 3D. Par rapport au MNT de référence (figure 3.9), le résultat obtenu à résolution de 25 cm est celui qui présente le moins d'erreur dans la sélection des sondes à garder, suivi par le résultat à résolution de 50 cm où quelques erreurs subsistent en bas à gauche sous forme de lignes diagonales : cela témoigne d'une maille mal corrigée qui a transmis son anomalie à certaines mailles voisines. Le côté vertical droit du MNT présente également des mailles mal corrigées (les couleurs jaune-vertes devraient être bleus foncés) : il s'agit de la zone la moins couverte car elle est la plus profonde et uniquement insonifiée par les grandes angles des fauchées, ce sont donc les mailles avec les plus faibles quantité de sondes. Ce pourtour très peu densément sondé pour une résolution de 10 cm, est le plus neutralisé après filtrage par rapport aux autres résolutions ; la zone conservée est donc plus réduite mais avec peu ou pas d'erreurs grossières. Pour le résultat à résolution de 10 cm, une série d'erreurs a été propagée au niveau de la remontée centrale comme pour le résultat à résolution de 1 m. Outre ces erreurs restantes, les quatre résultats ont une répartition des couleurs bleues (profondeurs) équivalente à la référence : le nettoyage des fonds plats est bien réussi au regard des représentations brutes en figure 3.8.

3.3.3 Comparaisons avec le filtrage manuel

Les temps de filtrage automatique sont très rapides : pour les quatre résolutions, la durée d'exécution est inférieure à une minute. La zone est assez petite, elle tient intégralement en mémoire dans l'ordinateur.

On s'intéresse à la matrice de confusion entre le filtrage manuel et automatique. La matrice de confusion est un outil de base pour l'évaluation de nombreux critères. Dans notre cas, elle s'obtient

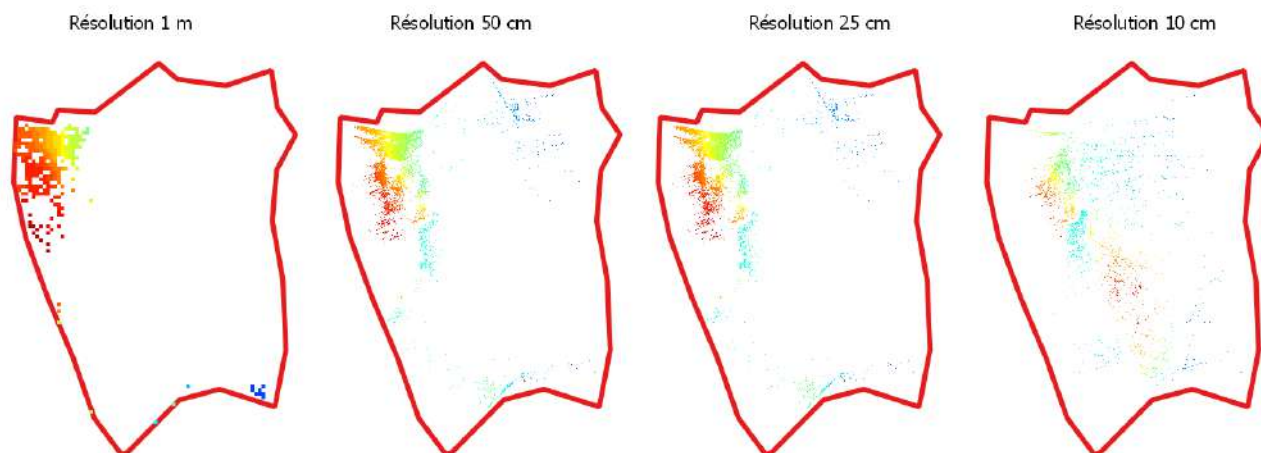


FIGURE 3.10 – Résultats de la sélection des mailles valides selon la résolution du MNT. Le cadre rouge correspond au contour de la zone en figure 3.8 et 3.9, il donne un aperçu de la quantité de mailles brutes non valides.

en remplissant une matrice de taille 2×2 (2 états de sondes, bon ou mauvais) où l'on compare un résultat obtenu avec une vérité. Chaque cellule de la matrice tient compte d'éléments par état. Celle donnée au niveau du tableau 3.2 est valable pour les quatre résolutions ; les pourcentages sont très similaires (à $\pm 1\%$) même si les résultats visuels semblent plus différents.

TABLE 3.2 – Matrice de confusion ramené en pourcentages des bonnes et mauvaises sondes selon le filtrage manuel de référence et le filtrage automatique.

		Filtrage Automatique		Total
		Bonnes sondes	Mauvaises sondes	
Filtrage Manuel	Bonnes sondes	69	2	71
	Mauvaises sondes	13	16	29
Total		82	18	100

De cette matrice, on peut déduire plusieurs mesures d'intérêt comme :

- (i) l'exactitude de la méthode indiquant la quantité de sondes dont l'état est équivalent avec les deux filtrages,
- (ii) l'exactitude aléatoire indique un taux d'accord "au hasard",
- (iii) l'indice de kappa combine ces deux exactitudes pour indiquer un taux d'accord entre les deux filtres.

Appliquées au tableau précédents, ces mesures valent :

$$\begin{aligned}
 \text{(i)} : & \frac{(69 + 16)}{100} && = 85\% \text{ de sondes en accord avec la vérité} \\
 \text{(ii)} : & \frac{69 + 13}{100} \times \frac{69 + 2}{100} + \frac{16 + 13}{100} \times \frac{16 + 2}{100} && = 62\% \text{ de sondes en accord "par hasard"} \\
 \text{(iii)} : & \frac{(i) - (ii)}{1 - (ii)} && = 0.6 \in [-1, 1] \text{ désigne un accord fort.}
 \end{aligned}$$

Ce taux d'accord entre filtres de "seulement" 0.6/1 peut s'expliquer par leur fonctionnement intrinsèque. En effet, lors du filtrage manuel, l'opérateur invalide les sondes douteuses car la redondance est telle qu'il n'y aura pas de "trous" dans le MNT : il n'y pas de mailles blanches dans le MNT de référence en figure 3.9. Tandis que le filtre automatique conserve globalement

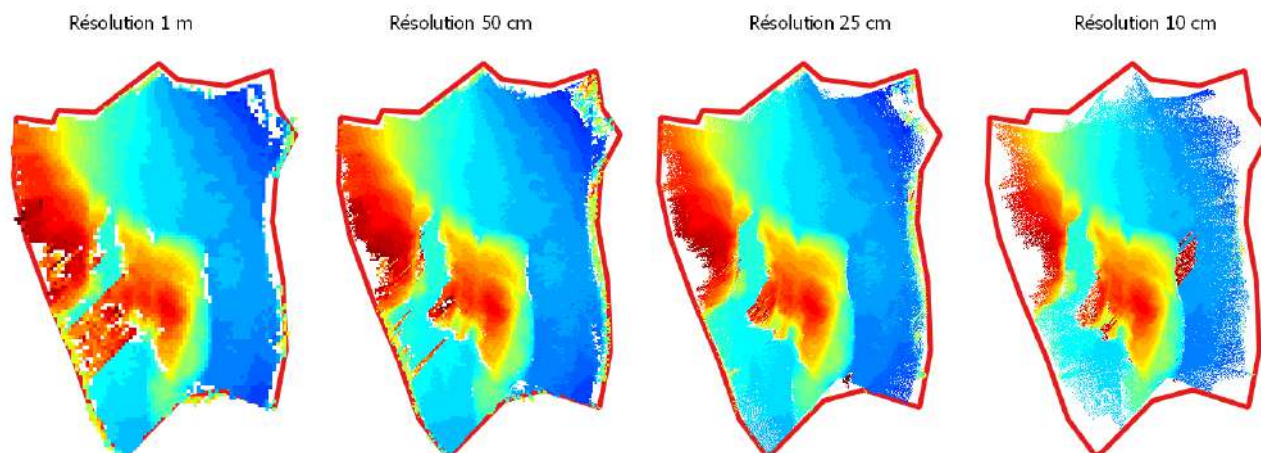


FIGURE 3.11 – Résultats du filtrage automatique sur le MNT à différentes résolutions.

plus de sondes, ici 15% en plus par rapport au filtrage manuel. En revanche, quelques mailles sont complètement invalidées et créent des mailles blanches (les trous dans les vues 3D en figure 3.12).

3.4 A propos de la méthode

La méthode proposée a été mise au point en analysant les sondes aberrantes couramment rencontrées lors du filtrage interactif (dans *ViewSMF*) ainsi que le contenu dans des mailles de MNT dans le logiciel *ViewMap*. Elle n'a pas ou peu été influencée par les méthodes existantes, ni par les méthodes disponibles dans d'autres logiciels : Seaviews ne dispose pas de licence pour ces logiciels. Il est certain, de ce fait, que cette méthode reste très élémentaire face aux méthodes plus modernes citées auparavant. Et pourtant, elle prend une direction commune à d'autres méthodes en :

- réalisant un traitement en mode spatial plutôt que temporel,
- étant complètement automatisée et rapide,
- traitant l'ensemble des données d'un jeu,
- laissant peu de paramètres à régler à un opérateur.

Et d'autre part, elle possède des points communs avec d'autres méthodes comme celle de [Du et al. \(1996\)](#) qui cherche à discerner localement le groupe de sondes aberrantes du groupe des sondes standard, mais également avec [Calder and Mayer \(2003\)](#) qui se sert du voisinage dans le MNT pour décider des sondes à rejeter et de celles à garder.

D'un point de vue résultat, ce filtrage automatique permet de passer en revue rapidement un grand nombre de sondes et de pouvoir rejeter les principales erreurs. Un traitement plus fin est donc laissé à un opérateur expert qui réalise des modifications à l'aide du filtrage par masques interactifs.

La zone étudiée est ici une petite surface, l'application du filtrage automatique est très rapide. En revanche, lorsque la zone devient très grande et/ou que la résolution du MNT est choisie fine, le filtrage peut devenir long à cause des nombreuses itérations sur l'ensemble des mailles du MNT. La solution consiste alors à subdiviser la zone d'étude en plusieurs parties et à traiter indépendamment chaque partie.

La méthode a néanmoins montré certaines limites au fur et à mesure de son utilisation. Si la méthode laisse un seul paramètre à régler à l'opérateur, en l'occurrence la résolution du MNT, il y a en réalité plusieurs paramètres qui devraient être ajustés.

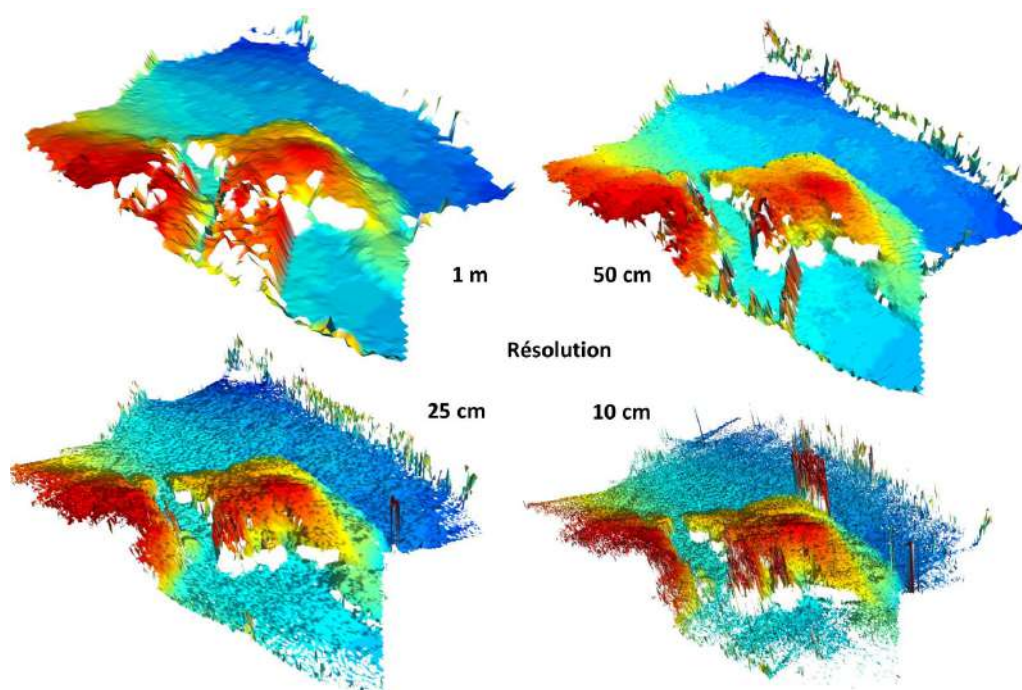


FIGURE 3.12 – Vues 3D des résultats de la figure 3.11.

Par exemple, les conditions de validité d'une maille basées sur des comparaisons entre l'écart type des sondes et la résolution de la maille semblent être appropriés pour cette zone d'étude bruitée et aux reliefs variés. Mais dans le cas d'une zone nécessitant une grande dimension de maille avec des fonds très plats et peu d'erreurs lors de l'acquisition, ces conditions ne permettent pas de détecter les sondes aberrantes noyées dans la densité de bonnes sondes.

Le réglage empirique des valeurs de seuils d'écart-type (0.5 ou 0.7) lors du nettoyage des mailles est lui aussi discutable. Un fort écart-type peut tout aussi bien refléter un fond en très forte pente (voir vertical comme un quai) tout comme une série de mesures avec de véritables erreurs. Ces seuils ne sont donc pas suffisant pour déterminer le nombre de partitions de données à faire. De plus, leurs valeurs deviennent inadaptées pour des fonds très réguliers ou contraire extrêmement chaotiques (telle une digue de rochers amorphes).

Chapitre 4

Extraction de la bathymétrie dans l'imagerie de la colonne d'eau

Les sondes bathymétriques sont issues d'une détection du fond dans le signal d'un faisceau acquis par le sondeur et Les sondes aberrantes sont le résultat d'une erreur de détection plutôt qu'une erreur de signal (cf paragraphe 2.1.2). Dans le chapitre précédent, nous avons mentionné que la méthode par filtrage interactif faisait intervenir l'image de la colonne d'eau (*Water Column Image*, WCI) comme "vérité" afin de décider de garder ou de rejeter des sondes comme sur les deux exemples de la figure 4.1. La WCI est certainement la donnée la plus brute accessible avec un SMF du commerce. Ce chapitre propose d'exploiter la donnée bathymétrique contenue dans les images de la colonne d'eau avec un double objectif : s'affranchir des erreurs de détection du fond et augmenter la précision bathymétrique en bout de fauchées où les sondes sont très souvent bruitées et par conséquent, éliminées. La méthode que nous proposons est fondée sur des réseaux de neurones artificiels (RNA) pour réaliser un apprentissage supervisé de la détection du fond dans ces images.

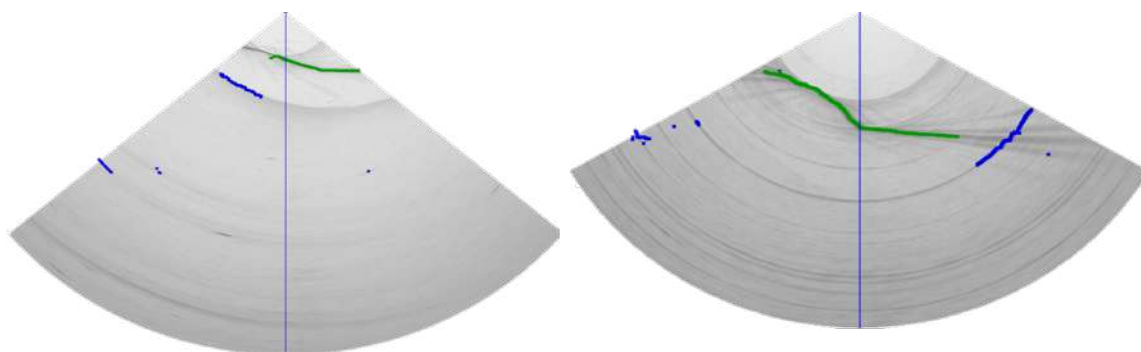


FIGURE 4.1 – Deux exemples de WCI où le fond de la mer est bien visible mais dont une partie des sondes bathymétriques est erronée (en bleu).

Ce chapitre s'articule autour de quatre sections. La première résume les principales notions en matière de réseaux de neurones dans le cadre de l'apprentissage profond (*Deep Learning*, DL) pour du traitement d'images. La seconde détaille les choix réalisés et la mise en place d'un tel processus sur les données WCI. La troisième présente les premiers résultats d'apprentissage obtenus lors de leur application sur des jeux de données sélectionnés. La quatrième discute de la méthode et de son développement ainsi que des résultats obtenus.

4.1 Traitement d'image par réseau de neurones

Les RNA que nous connaissons aujourd'hui réalisent des tâches pointues de classification, détection, segmentation, régression, etc. pour des domaines d'application variés tels que le bio-médical (Shahid et al.; 2019), la sécurité et la défense (Svenmarck et al.; 2018) ou encore le divertissement (Dhawale et al.; 2020). Ils sont issus d'au moins 80 ans de développement dans des domaines variés tels que les mathématiques pour la neurophysiologie, la logique, les sciences des données, les sciences cognitives. Seules les notions essentielles et leurs déclinaisons pour le traitement d'images sont présentées ci-dessous.

4.1.1 Notions générales sur les réseaux de neurones artificiels

4.1.1.1 De la modélisation d'un neurone aux premières machines neuronales

L'intelligence artificielle (IA) connexionniste est issue d'un ensemble de travaux dont les prémices remontent aux années 1940 (pour un panorama historique de l'IA, voir (Cardon et al.; 2018)). La pierre fondatrice du domaine est celui de Mcculloch and Pitts (1943) qui propose un modèle de neurone formel s'inspirant du fonctionnement simplifié d'un neurone biologique. Un neurone reçoit des informations de la part d'autres neurones par ses dendrites et en réalise une somme pondérée (ou tout autre type d'opération simple) dans le soma (Fig. 4.2). Si la somme dépasse un seuil, le soma déclenche une activation communiquée à d'autres neurones à l'aide de l'axone dont les extrémités se terminent par des synapses (Fig. 4.2). Mathématiquement, ce modèle s'exprime de la manière suivante :

$$o = \psi \left(\sum_{n=1}^N w_n d_n + b \right) = \begin{cases} 0 & \text{si } \sum w_n d_n + b < 0 \\ 1 & \text{sinon.} \end{cases} \quad (4.1)$$

avec N informations d'entrée, notées $d_n \in \mathbb{R}, \forall n \in [1, N]$, pondérées par les poids, w_n et par un biais, b . La fonction ψ est la fonction d'activation, ici la fonction de Heaviside. L'information de sortie du neurone est notée o (*output*).

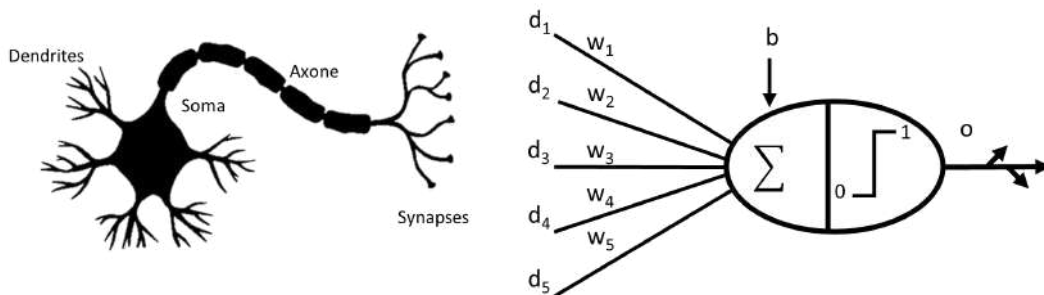


FIGURE 4.2 – A gauche, schéma d'un neurone biologique. A droite, le neurone formel de Mcculloch and Pitts (1943)

Peu après, le conséquent travail de recherche de Hebb (1949) est publié proposant une théorie d'apprentissage et de mémorisation d'un réseau de neurones (formels et biologiques). La règle de Hebb suppose qu'un apprentissage se réalise par l'activation répétée d'un neurone par un autre, c'est-à-dire qu'un chemin préférentiel s'établit entre éléments. Cette règle complète le modèle du neurone formel pour adapter les différents poids suivant les liens renforcés ou réduits entre neurones.

De ces modèles mathématiques sont apparues plusieurs applications, dont la plus connue est le Perceptron de Rosenblatt (1958). C'est un classifieur binaire à apprentissage supervisé composé d'un unique neurone : des données étiquetées selon deux classes (il s'agit d'un problème binaire) sont fournies à ce neurone durant une phase d'entraînement. Lors de l'entraînement, chaque donnée

produit un ajustement des poids et l'entraînement perdure jusqu'à ce que le taux d'erreur entre la prédiction et l'étiquette soit suffisamment faible - quitte à ce que le neurone passe en revue plusieurs fois les mêmes données. L'ajustement des poids fait intervenir un taux d'apprentissage, η , pris entre 0 et 1, et l'erreur entre la prédiction o et l'étiquette de la donnée e ,

$$\forall n \in [1, N], w_n = w_n - \eta(e - o)d_n. \quad (4.2)$$

En début d'apprentissage, il est possible d'initialiser les w_n à 0 ou avec des valeurs tirées au hasard. Un taux d'apprentissage choisi trop grand peut entraîner la divergence de l'algorithme, tandis que les temps d'apprentissage peuvent devenir très longs avec des taux trop petits.

Cependant, l'engouement retombe à la fin des années 1960 car d'une part les puissances de calcul de l'époque ne sont pas suffisantes pour les problèmes complexes et d'autre part ces systèmes ne peuvent résoudre que des problèmes simples (Minsky and Papert; 1969).

4.1.1.2 Les origines du connexionnisme moderne

A la fin des années 1980, plusieurs travaux font revenir le connexionnisme à la mode. D'une part, les architectures à partir du Perceptron se complexifient et se diversifient et d'autre part, une nouvelle manière d'ajuster les poids révolutionne leur entraînement.

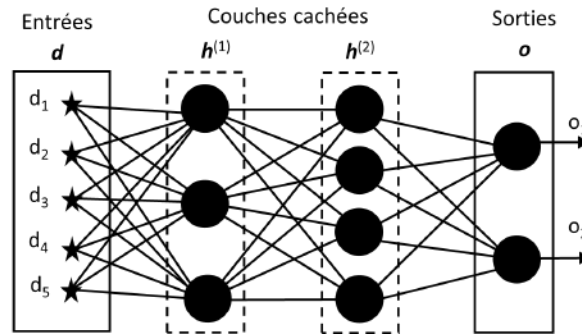


FIGURE 4.3 – Exemple d'un Perceptron multicouche composé d'une couche d'entrée, de deux couches cachées et d'une couche de sortie. Chaque couche est totalement connectée avec la précédente et la suivante.

Parmi les architectures de cette décennie, le Perceptron multicouche (*Multi-Layer Perceptron*, MLP) est l'un des plus répandus, encore aujourd'hui. Il se compose d'au moins une couche cachée de neurones (notée \mathbf{h} pour *hidden layer*) glissée entre les entrées et le/les neurones de sortie, et chaque neurone d'une couche est relié à tous les neurones de la couche suivante (couche dite totalement connectée ou *fully connected layer*) (Fig. 4.3). La sortie d'un MLP s'exprime à l'aide de fonctions composées des différentes couches décrites par l'équation 4.1. En prenant l'exemple de la figure 4.3 et en associant les données d'entrée \mathbf{d} à la couche numéro 0 et la couche de sortie \mathbf{o} à la 3^e couche, on obtient :

$$\begin{aligned} \mathbf{o} &= \psi \left(\mathbf{W}^{(3)} \cdot \mathbf{h}^{(2)} + \mathbf{b}^{(3)} \right) \\ \mathbf{h}^{(2)} &= \psi \left(\mathbf{W}^{(2)} \cdot \mathbf{h}^{(1)} + \mathbf{b}^{(2)} \right) \\ \mathbf{h}^{(1)} &= \psi \left(\mathbf{W}^{(1)} \cdot \mathbf{d} + \mathbf{b}^{(1)} \right) \\ (\mathbf{o} \circ \mathbf{h}^{(2)} \circ \mathbf{h}^{(1)})(\mathbf{d}) &= \psi \left(\mathbf{W}^{(3)} \cdot \psi \left(\mathbf{W}^{(2)} \cdot \psi \left(\mathbf{W}^{(1)} \cdot \mathbf{d} + \mathbf{b}^{(1)} \right) + \mathbf{b}^{(2)} \right) + \mathbf{b}^{(3)} \right) \end{aligned} \quad (4.3)$$

$$\text{où la fonction d'activation est une sigmoïde définie par : } \psi(x) = \frac{1}{1 + e^{-x}} \quad (4.4)$$

où les $\mathbf{b}^{(i)}$ sont les vecteurs de biais des neurones de la couche $i \in [1, 3]$ et les $\mathbf{W}^{(i)}$ sont les matrices de poids des neurones de la couche i par rapport aux éléments de la couche précédente. Ainsi, pour la couche numéro 1, couche cachée $\mathbf{h}^{(1)}$, on peut noter :

$$\text{l'entrée } \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \\ d_4 \\ d_5 \end{bmatrix}, \text{ les biais } \mathbf{b}^{(1)} = \begin{bmatrix} b_1^{(1)} \\ b_2^{(1)} \\ b_3^{(1)} \end{bmatrix} \text{ et les poids } \mathbf{W}^{(1)} = \begin{bmatrix} \mathbf{w}_1^{(1)} \\ \mathbf{w}_2^{(1)} \\ \mathbf{w}_3^{(1)} \end{bmatrix} = \begin{bmatrix} w_{1,1}^{(1)} & \dots & w_{1,5}^{(1)} \\ w_{2,1}^{(1)} & \dots & w_{2,5}^{(1)} \\ w_{3,1}^{(1)} & \dots & w_{3,5}^{(1)} \end{bmatrix}$$

$$\text{de telle manière que } \mathbf{h}^{(1)} = \begin{bmatrix} h_1^{(1)} \\ h_2^{(1)} \\ h_3^{(1)} \end{bmatrix} = \begin{bmatrix} \psi \left(\sum_{n=1}^5 w_{1,n}^{(1)} x_n + b_1^{(1)} \right) \\ \psi \left(\sum_{n=1}^5 w_{2,n}^{(1)} x_n + b_2^{(1)} \right) \\ \psi \left(\sum_{n=1}^5 w_{3,n}^{(1)} x_n + b_3^{(1)} \right) \end{bmatrix}$$

Ces MLP sont très longs à entraîner car ils comptent un grand nombre d'hyperparamètres ; celui de la figure 4.3 contient déjà 44 poids et biais. La technique de rétropropagation du gradient, popularisée par Rumelhart et al. (1986), y apporte une solution. Cette méthode intervient après la passe *en avant* du calcul dans le réseau (équation 4.3). Elle calcule le gradient de l'erreur de chaque neurone d'un réseau en le parcourant à l'envers puis corrige les poids et les biais proportionnellement à leur contribution à l'erreur. Pour cela, on définit une fonction de perte (*Loss function*) ℓ qui calcule la différence entre la sortie \mathbf{o} et l'étiquette \mathbf{e} de la donnée ayant M caractéristiques. La différence peut s'exprimer au moyen de l'erreur quadratique moyenne. La fonction de perte s'écrit alors :

$$\ell = \frac{1}{M} \sum_{m=1}^M (y_m - e_m)^2. \quad (4.5)$$

Puis, chaque poids et biais est mis à jour, à la manière de l'équation 4.2, par :

$$w_{r,s}^{(i)} = w_{r,s}^{(i)} - \eta \frac{\partial \ell}{\partial w_{r,s}^{(i)}} \text{ et } b_r^{(i)} = b_r^{(i)} - \eta \frac{\partial \ell}{\partial b_r^{(i)}} \quad (4.6)$$

avec i l'indice de la couche, r l'indice du neurone de la couche i et s l'indice du neurone de la couche $i - 1$.

Pour calculer la dérivée partielle de la fonction de perte pour un paramètre $q_{r,(s)}^{(i)}$ (poids ou biais) du neurone r de la couche i (couche cachée ou de sortie), $c_r^{(i)}$, on utilise une formule de la chaîne pour la décomposer en trois termes :

$$\frac{\partial \ell}{\partial q_{r,(s)}^{(i)}} = \frac{\partial \ell}{\partial c_r^{(i)}} \frac{\partial c_r^{(i)}}{\partial \mathbf{w}_r^{(i)} \cdot \mathbf{c}^{i-1} + b_r^{(i)}} \frac{\partial \mathbf{w}_r^{(i)} \cdot \mathbf{c}^{i-1} + b_r^{(i)}}{\partial q_{r,(s)}^{(i)}}. \quad (4.7)$$

Le deuxième et le troisième terme se simplifient par :

$$\frac{\partial c_r^{(i)}}{\partial \mathbf{w}_r^{(i)} \cdot \mathbf{c}^{i-1} + b_r^{(i)}} = \psi'(\mathbf{w}_r^{(i)} \cdot \mathbf{c}^{i-1} + b_r^{(i)})$$

$$\text{et } \frac{\partial \mathbf{w}_r^{(i)} \cdot \mathbf{c}^{i-1} + b_r^{(i)}}{\partial q_{r,(s)}^{(i)}} = \begin{cases} c_s^{(i-1)} & \text{si le paramètre est un poids} \\ 1 & \text{si le paramètre est un biais.} \end{cases}$$

Le premier terme, appelé gradient local, est à développer par récursivité de la formule de la chaîne sur le neurone $c_r^{(i)}$ jusqu'à la sortie \mathbf{o} . A chaque étape, on retrouve les trois termes à développer et à simplifier. En fin de développement, il reste le terme $\frac{\partial \ell}{\partial \mathbf{o}} = \frac{2}{M}(\mathbf{o} - \mathbf{e})$.

Cette technique est aujourd'hui appliquée à une grande partie des RNA. Elle connaît un certain nombre de variantes. La fonction de perte doit être adaptée au type de problème : erreur quadratique moyenne ou erreur absolue moyenne pour un problème de régression, entropie croisée ou fonction de vraisemblance pour un problème de classification. Les corrections de paramètres au niveau de l'équation 4.6 peuvent également faire intervenir un facteur d'inertie (*momentum*) (Polyak; 1964) qui prend en compte les gradients antérieurs pour éviter la convergence vers des minimum locaux du gradient, comme l'optimisation avec le *gradient accéléré* de Nesterov (Nesterov; 1983). Il existe également des méthodes utilisant un taux d'apprentissage adaptatif, comme *Ada-Grad* de Duchi et al. (2011). Mais de manière assez large, la rétropropagation qualifie l'ensemble des méthodes de mise à jour des paramètres à partir d'un calcul de gradient.

4.1.2 Application à la vision par ordinateur

4.1.2.1 L'apparition des réseaux de neurones à convolution

Les neurosciences avancent en parallèle des développements en IA connexionniste. Les études de Hubel and Wiesel (1959) portant sur le fonctionnement du cortex visuel montrent que les neurones de ce cortex ont chacun un *champ récepteur local* dans une scène. L'ensemble de ces neurones couvrent le champ visuel avec éventuellement de la redondance (chevauchement des champs locaux). De plus, les neurones sont organisés en niveaux selon la largeur de leur champ et selon des sensibilités différentes pour des motifs - simples ou complexes - de la scène et ces niveaux communiquent entre eux (Géron; 2017).

Cette structure du cortex visuel est convertie en un RNA, le *Neocognitron* de Fukushima et al. (1983). Ce réseau se compose de neuf couches, avec une alternance de couches d'extraction de caractéristiques puis de moyennage pour lisser les erreurs de déformation des motifs. Ce réseau est entraîné sur la reconnaissance de chiffres manuscrits.

La rétropropagation du gradient est appliquée à un réseau s'inspirant du *Neocognitron* par LeCun et al. (1989). Ce réseau très connu s'appelle *LeNet-5* et résout également la reconnaissance de chiffres manuscrits, dont la base d'entraînement, *MNIST*, elle aussi célèbre, est réalisée avec 60 000 chiffres issus des codes postaux de courriers. A la manière de son prédécesseur, les premières couches sont une alternance de convolution (extraction de caractéristiques) et de sous-échantillonnage, mais les dernières couches sont totalement connectées comme un MLP et réalisent l'opération de classification (figure 4.4). Ce schéma en deux phases est un classique des architectures de réseaux de neurones à convolution (*Convolutional Neural Network*, CNN). Les détails des principales couches et modules des CNN récents sont présentés dans la section suivante.

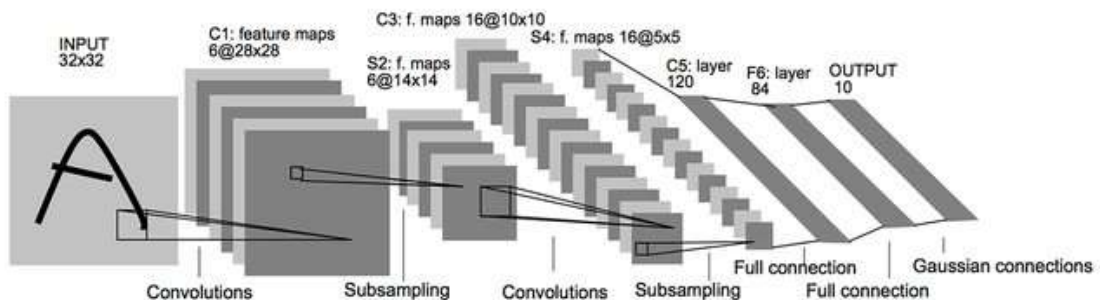


FIGURE 4.4 – Réseau LeNet-5 proposé par LeCun et al. (1998) pour la reconnaissance de chiffres manuscrits.

Les réseaux de neurones à convolution sont l'une des approches de DL les plus populaires pour traiter des problèmes de vision par ordinateur - détection, segmentation, classification, etc. - pour des données de type images couleurs (trois canaux Rouge-Vert-Bleu) ou des tenseurs de données.

Le terme de “convolution” signifie que le réseau effectue des opérations linéaires de convolution au sein de certaines couches à la place de multiplications matricielles.

4.1.2.2 Types de couches

Le neurone formel est l'élément de base d'un MLP. Dans un CNN, ce sont les différentes couches que l'on empile qui sont les briques élémentaires de ce type de réseaux. Un très grand nombre de couches existent, mais certaines d'entre elles sont présentes dans tous les réseaux modernes : convolution, sous-échantillonnage, activation, couche totalement connectée, régularisation et normalisation.

La couche de convolution est la couche la plus représentative des CNN. Elle permet l'extraction de caractéristiques dans une image à l'aide de filtres (dits noyaux ou *kernel*) parcourant l'image. Les filtres sont carrés de tailles impaires (classiquement 3, 5, 7 ou 11) et décroissantes au fur et à mesure des successions de convolution car les images intermédiaires se réduisent en dimension avec la profondeur du traitement. Cependant, les couches d'entrées calculent moins d'images intermédiaires que les couches plus profondes ; il y a une compensation entre la taille du filtre et leur nombre. Les filtres sont donc définis par leur nombre et leur taille.

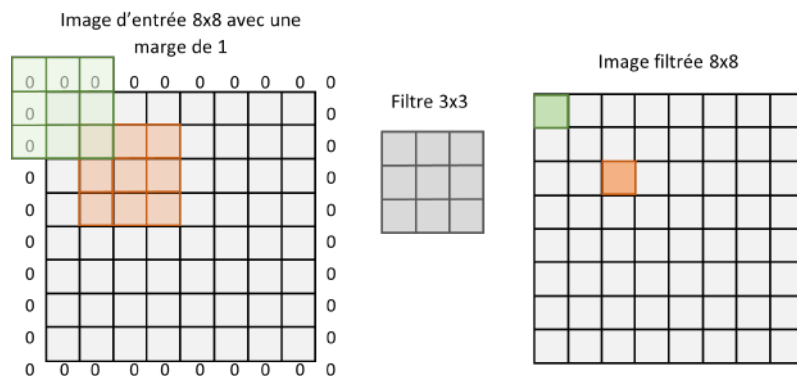


FIGURE 4.5 – Couche de convolution sur une image d'entrée augmentée d'une marge de 1 pixel avec un filtre de taille 3x3 et un chevauchement de 1 pixel.

Deux autres réglages complètent le calcul de convolution : la gestion de la marge (*padding*) et du pas de chevauchement (*stride*). Un pas de chevauchement de 1 pixel signifie que le filtre se décale d'un pixel en un pixel. Une marge nulle signifie que l'image n'est pas élargie avec des valeurs à 0 ; les filtres, du fait leur taille, ne peuvent couvrir les pixels sur les bords et l'image filtrée est plus petite. Pour conserver une image filtrée de la même taille que celle d'entrée, la marge doit être adaptée à la taille du filtre et au chevauchement comme sur le schéma de la figure 4.5.

Les couches de convolution apportent plusieurs avantages au traitement des images. Une couche de convolution comporte assez peu de neurones - et donc de paramètres à régler - par rapport aux couches de MLP. La première couche de convolution du réseau *LeNet-5* possède 8 filtres de taille 5x5 à appliquer à une image à un canal (image en niveau de gris), cela fait $((5 \times 5 \times 1) + 1 \text{ biais}) \times 8 = 208$ paramètres. Cette quantité n'est pas dépendante de la taille des images en entrée. D'autre part, ces calculs permettent de conserver les corrélations d'information entre pixels voisins.

La couche de sous-échantillonnage, dite aussi de “mise en commun” ou *pooling*, permet de réduire la dimension des images intermédiaires. Les réglages d'une couche de pooling sont la taille du filtre - quasiment toujours 2x2 - et la fonction appliquée - très généralement le maximum ou éventuellement la moyenne. Il n'y a pas de marge ou de chevauchement du filtre, comme sur le schéma de la figure 4.6. Le pooling intervient généralement après une ou plusieurs convolutions

et avant des couches totalement connectées pour réduire le nombre de paramètres induits par les connections. Ce sous-échantillonnage par le maximum permet également la réduction du bruit présent dans les images intermédiaires ; seules les fortes valeurs des motifs sont conservées dans la suite.

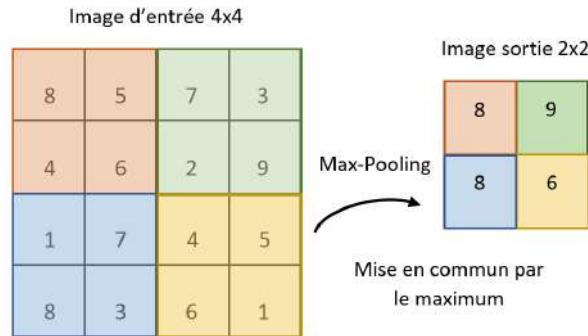


FIGURE 4.6 – Couche de réduction avec un *Max-Pooling* de filtre de 2x2 et stride 2.

La couche entièrement connectée, dite aussi *fully connected* (FC), est une couche empruntée au MLP. Une ou plusieurs couches FC sont empilées en fin de réseau avant la sortie. Ces couches permettent l'apprentissage, parfois non-linéaire, des combinaisons de caractéristiques. Une forte réduction de dimension des données d'entrée par l'empilement de couches convolution-pooling est importante car ce sont les couches FC qui engendrent le plus de paramètres à ajuster. La première couche FC du réseau *LeNet-5* contient 120 neurones reliés aux 16 images intermédiaires de taille 5x5, cela représente donc $((5 \times 5 \times 16) + 1 \text{ biais}) \times 120 = 48\,120$ paramètres.

La couche d'activation intervient en sortie de couche de convolution et de FC. Elle applique une fonction non-linéaire qui améliore l'efficacité du traitement. Cette fonction doit être dérivable pour la rétropropagation du gradient. Dans l'équation 4.1 du neurone formel, il est question d'une fonction de Heaviside ; mais, bien que dérivable, sa dérivée est toujours nulle, ce qui ne permet pas une rétropropagation du gradient. Ainsi, les fonctions d'activation les plus couramment utilisées au sein d'un réseau sont les fonctions *tangente hyperbolique* (*tanh*) ou *sigmoïde* qui ramènent l'espace d'entrée \mathbb{R} dans, respectivement, $[-1,1]$ ou $[0,1]$:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad \text{et} \quad \text{sigmoïde}(x) = \frac{1}{1 + e^{-x}}.$$

Cependant, la dérivée de ces fonctions est assez longue à calculer. Elles sont donc passées de mode au profit de la fonction *unité linéaire rectifiée* ou ReLu (Krizhevsky et al.; 2012), qui $\forall x \in \mathbb{R}$ possède une image dans $[0, +\infty[$ telle que $\text{ReLu}(x) = \max(0, x)$.

En général, les couches cachées utilisent toutes la même fonction d'activation et seule la couche de sortie est constituée d'une autre fonction. Pour des problèmes de classification de données, c'est la fonction *softmax* qui est employée, elle traduit une valeur en une probabilité d'appartenance à une classe. Dans des problèmes de régression, on préférera une fonction d'activation de type linéaire.

La couche de régularisation est plus communément appelée *DropOut*. Uniquement employée pour la phase d'entraînement, son action est de "désactiver" temporairement et aléatoirement un pourcentage de neurones (et leurs sorties) sur une couche donnée. C'est une des solutions pour éviter les problèmes de surapprentissage (Srivastava et al.; 2014), c'est-à-dire l'incapacité d'un réseau à généraliser l'apprentissage. Depuis son invention par Hinton et al. (2012), le DropOut est majoritairement utilisé entre des couches FC avec un taux de 50% d'extinction de neurones.

La couche de normalisation est aussi connue sous le nom de *Batch Normalisation* introduit par Ioffe and Szegedy (2015). Les données, avant tout passage dans un RNA, sont normalisées, c'est-à-dire que pour à tous les éléments des données est retranchée la moyenne de l'ensemble du jeu et la différence est divisée par l'écart type. Lors du passage dans un RNA, les données transformées perdent leur normalisation et changent de distribution par rapport à la donnée initiale. La couche de normalisation appelée après une convolution ou une FC calcule la moyenne et l'écart type des valeurs du *batch* de données en cours pour les appliquer. Cette étape répétée à chaque fois que nécessaire dans le réseau permet de transférer des couches moins profondes au plus profondes les modifications sans atténuation et ainsi, permet d'accélérer l'apprentissage.

4.1.2.3 Modules classiques et macro-réseaux

Les briques élémentaires qui viennent d'être présentées ne sont pas empilées aléatoirement mais composent des modules qui sont répétés plusieurs fois au sein de réseaux. De 2010 à 2017, c'est le challenge annuel (ILSVRC¹) de reconnaissance d'animaux dans la base de photos *ImageNet* (1000 classes avec plus d'un million d'images) qui a fait avancer rapidement les innovations en matière de CNN.

Le “**tri-couche**” **convolution-activation-pooling** est un classique des réseaux tels que *LeNet-5* (LeCun et al.; 1998), *AlexNet* (Krizhevsky et al.; 2012) et VGG² de (Simonyan and Zisserman; 2015).

Dans le réseau *LeNet-5* (figure 4.4), le module est [**Convolution - Sigmoid - Average Pooling**] répété deux fois et le noyau de convolution est grand (taille de 5x5) devant les dimensions de l'image d'entrée (taille de 32x32). Les tailles des images intermédiaires sont donc réduites rapidement tout en ayant un nombre de paramètres (de l'ordre de 12 000) convenable par rapport à la taille de la base d'entraînement (60 000 images).

Le réseau *AlexNet* est le gagnant 2012 de l'ILSVRC et le premier gagnant notable de cette compétition. Dans ce réseau, on retrouve le module [**Convolution- ReLu - Max Pooling**] deux fois, avec une taille de noyaux de 11 puis de 5 alors que les images d'entrée sont de bien plus grandes dimensions, 224x224. Puis, c'est un nouveau module qui traite finement les motifs de haut niveau : [[**Convolution - Relu**]**x3 - Max Pooling**] avec des petits noyaux de convolution (3x3). Ce réseau popularise l'activation ReLu et le pooling par le maximum plutôt que par la moyenne. Cette même architecture est également gagnante en 2013 avec le réseau *ZFNet* (Zeiler and Fergus; 2014) en apportant une modification sur la taille du noyau à la première convolution, le noyau de 11x11 est réduit à une taille 7x7 car un trop grand noyau passe à côté d'informations. L'autre modification entre *AlexNet* et *ZFNet* tient à la manière dont sont réalisés les calculs. Pour *AlexNet*, les opérations sont réalisées sur deux unités distinctes de calcul qui ne croisent les caractéristiques qu'une fois au milieu du réseau puis au niveau des FC (figure 4.8 à gauche). Ce n'est pas le cas avec *ZFNet* (figure 4.8, le deuxième depuis la gauche).

L'intuition derrière ces deux réseaux est que l'empilement de plusieurs couches de convolutions à petits noyaux est au moins aussi efficace qu'une convolution avec un grand noyau et qu'il y a moins de paramètres à ajuster. C'est ce sur quoi s'appuie l'un des gagnants de 2014, le réseau *VGG* de Simonyan and Zisserman (2015). Le module [[**Convolution - Relu**]**x2 ou 3 - Max Pooling**] est répété cinq fois dans le réseau avec une unique taille de noyau de convolution (3x3). En effet, deux convolutions 3x3 correspondent à une convolution 5x5 avec 38% de paramètres en moins, et trois correspondent à l'utilisation d'un noyau 7x7 avec 81% de paramètres en moins.

Le **module d'inception** est proposé par Szegedy et al. (2015) dans le réseau *GoogLeNet*, grand gagnant de 2014. Ce module parallélise quatre calculs - un max-pooling et trois convolutions - avant de concaténer les différents résultats (figure 4.7 à droite). Les trois convolutions sont réalisées avec trois tailles de noyau, 1x1, 3x3 et 5x5, et des marges adaptées pour que les filtres soient de tailles

1. ImageNet Large Scale Visual Recognition Challenge
2. Visual Geometry Group

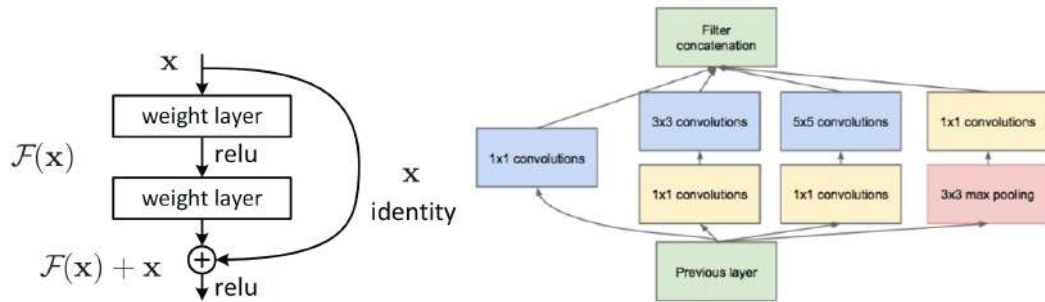


FIGURE 4.7 – Modules du résidu à gauche et d’inception à droite, issus respectivement de He et al. (2015) et de Szegedy et al. (2015).

identiques pour permettre une concaténation. De même pour le max-pooling. L’idée derrière ces quatre opérations est d’aller chercher les motifs à différentes échelles.

Les convolutions par les noyaux de 3x3 et 5x5 et le max-pooling sont précédées ou suivies d’une couche de convolution à noyau de taille 1x1 dite “goulot de rétrécissement” (*bottleneck*) ou projection. Cette convolution permet de diminuer le nombre de caractéristiques de la donnée d’entrée pour éviter une explosion du nombre de poids et biais par la nouvelle convolution. Par exemple, au 4^e inception, la donnée d’entrée se compose de 512 caractéristiques de dimension 14x14 pixels. La convolution 1x1 calcule 160 caractéristiques. La convolution 3x3 calcule 224 caractéristiques après un goulot ayant ramené à 112 caractéristiques. La convolution 5x5 calcule 64 caractéristiques après un rétrécissement à 24 caractéristiques. La projection ramène le nombre de caractéristiques de 512 à 64 après le 3x3 max-pooling. Cela fait un total de $160+224+64+64 = 512$ caractéristiques à la concaténation avec environ 437000 paramètres au lieu de 2 millions sans goulot.

Le résidu dans le réseau *ResNet* de He et al. (2015), gagnant de ILSVRC 2015, correspond à des “connexions de raccourcis” entre une couche et le résultat quelques couches après (figure 4.7 à gauche), c’est-à-dire à ajouter la valeur d’entrée x à la valeur de sortie des couches, $F(x) + x$. Cela doit forcer le réseau à modéliser le résidu, $F(x) - x$, plutôt que $F(x)$. L’intérêt est double, car au début de l’apprentissage les poids sont très mal ajustés voire proches de zéro et les $F(x)$ sont aussi proches de zéro dans tout le réseau; il y a une certaine inertie au démarrage et les couches plus profondes ne peuvent commencer à s’ajuster que si les moins profondes ont amorcé leur convergence. Ainsi, en rappelant régulièrement x dans le réseau, les $F(x) + x$ sont proches de x et cela accélère l’entraînement au démarrage et permet aux couches même profondes de s’ajuster plus vite.

Dans les différentes versions de *ResNet* (de 18 à 152 couches), les connexions de raccourcis sont réalisées sur deux ou trois convolutions successives avec des paramètres identiques (taille du noyau et nombre de filtres), comme sur la figure 4.8. D’autre part, *ResNet* diffère des autres réseaux car les changements de taille des images intermédiaires sont effectués plutôt avec les paramètres des couches de convolution (pas de chevauchement de 2 pixels) qu’avec du pooling. Basé sur une architecture relativement simple et puissante, le nombre de poids et de biais à régler est d’environ 11 millions pour *ResNet18*.

Au cours de ces 20 dernières années, les architectures des CNN ont beaucoup évolué et amélioré leurs capacités à résoudre des problèmes. On observe un approfondissement des réseaux, avec plus d’une centaine de couches pour *Resnet152* contre une dizaine pour *AlexNet*, et la possibilité de manipuler plus d’une centaine de millions de paramètres (*VGG16* ou 19) même si la tendance actuelle est à la réduction (7 millions pour *GoogLeNet*, 11 millions pour *ResNet18*).

Dans la section suivante, nous allons voir comment transposer ces notions sur un problème bien différent de la reconnaissance d’un sujet présent dans une photographie. Outre l’architecture d’un réseau, la préparation des données demeure une phase très importante, voire essentielle. La

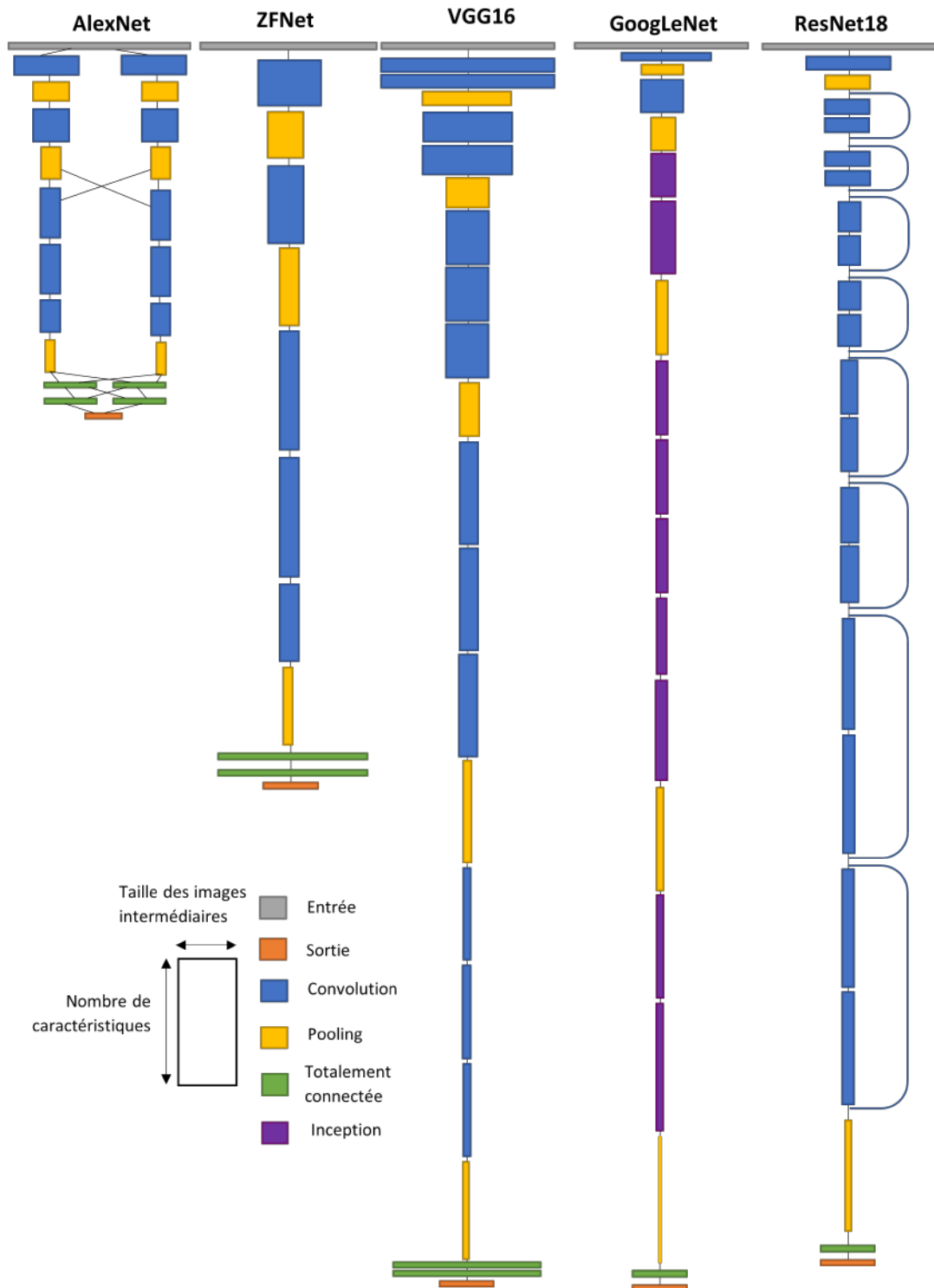


FIGURE 4.8 – Représentation des cinq formes de CNN les plus rencontrées. Les différentes couches ou modules sont représentées par des rectangles dont les dimensions reflètent le nombre et la taille des caractéristiques calculées à chaque étape. Ces dimensions donnent l'allure des réseaux sans être réellement quantitatives.

sélection des réglages autour du réseau l'est tout autant pour éviter les écueils du sur- ou du sous-apprentissage, de l'extinction ou de l'explosion du gradient.

4.2 Mise en place d'une méthode d'apprentissage profond

La réalisation d'un projet utilisant de l'apprentissage profond est un processus assez long comprenant plusieurs étapes distinctes. Nous allons les présenter au travers de deux sections, dont la première est dédiée à la préparation des données et la seconde aux algorithmes de calcul.

4.2.1 La préparation des données

4.2.1.1 Labelliser les images WC

Dans notre problème, le label d'une image WC doit correspondre à un vecteur de 256 valeurs bathymétriques - une valeur par faisceau de ping. L'idée naïve est d'utiliser en tant que telle la donnée bathymétrique acquise par le sondeur en parallèle des WCI, mais cette approche ne permettrait pas les améliorations attendues aux grands angles ou l'élimination des erreurs de détection. L'idée proposée est donc de labelliser les WCI à l'aide du modèle numérique de terrain (MNT) construit à partir des données bathymétriques. Pour cela, la première étape est donc de réaliser un MNT avec des données bathymétriques nettoyées de leurs fausses sondes - par exemple avec la méthode décrite au chapitre précédent. Ensuite, pour un faisceau donné, on cherche où se situe le fond dans le signal : en connaissant la trajectoire du rayon acoustique dans ce faisceau, on détermine le temps de propagation mis par le rayon depuis son émission afin qu'il coïncide avec la surface du MNT réalisé (figure 4.9). Cette recherche, réalisée à l'aide d'un algorithme de dichotomie, est effectuée pour tous les faisceaux de tous les pings à labelliser.

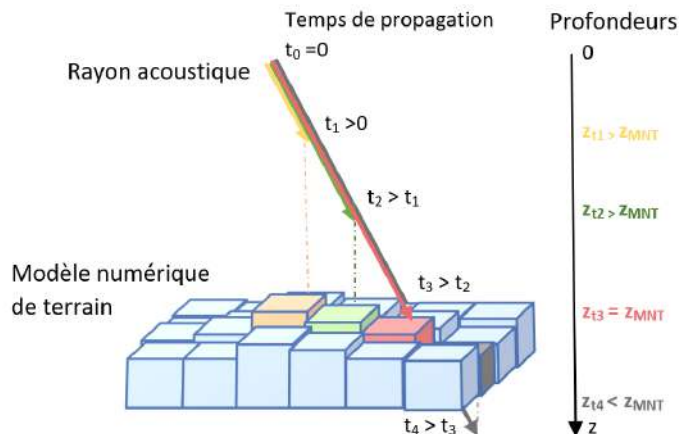


FIGURE 4.9 – Un rayon acoustique se propage selon une trajectoire connue (flèches). A différents temps de propagation, on évalue la différence de profondeur entre celle atteinte par le rayon et celle observée au niveau de la maille du MNT où le rayon viendrait se projeter. C'est le temps t_3 qui correspond le mieux au moment où le signal rencontre le fond dans cet exemple.

4.2.1.2 Redimensionner les images WC

Usuellement les exemples les plus connus de CNN emploient des images carrées de dimensions proches de 200 pixels de côté car cela est plus optimal en terme de calcul (que cela soit en temps et en nombre de paramètres) qu'une image rectangulaire de grandes dimensions. C'est le cas des images de ImageNet, les images sont réduites en dimensions puis découpées pour obtenir une forme carrée de 224x224 pixels.

Dans notre cas, les images *water column* sont “rectangulaires” avec 256 colonnes et un nombre de lignes qui peut varier de 1000 à 20000 selon la profondeur de la mer. L’entraînement d’un RNA avec des données d’entrée de taille variable est possible mais plus compliqué. Pour respecter le côté standard des images carrées, nous appliquons un redimensionnement afin de ramener les WCI à une taille de 256x256. Le sous-échantillonnage est réalisé colonne par colonne dans l’image à l’aide d’un filtre à réponse impulsionnelle finie de type passe-bas à phase linéaire et d’une fenêtre de Kaiser-Bessel.

4.2.1.3 Data Iterator

En général, les données ne sont pas fournies telles quelles au CNN mais au travers d’un “gestionnaire” qui construit des lots de données d’apprentissage et de test (figure 4.10). En effet, si les données sont fournies une par une au CNN et qu’une rétropropagation a lieu pour chaque donnée, les temps de calcul deviennent extrêmement long. A l’inverse, si la rétropropagation a lieu après la revue de l’ensemble des données, l’apprentissage peinera à se faire. La solution est donc de fournir au CNN des données par lots, *batch*, et ainsi de faire une rétropropagation sur un petit nombre de données. La taille du batch est à ajuster entre les deux extrêmes précédents et la mémoire vive disponible, en tenant compte du fait qu’un RNA occupe déjà une partie de cette mémoire avec un grand nombre d’hyperparamètres.

Les données sont vues plusieurs fois lors de l’apprentissage. On dit qu’une *epoch* a été effectuée quand l’ensemble des données ont été balayées. A la fin d’une epoch (ou dans d’autres circonstances), le gestionnaire fournit un lot de données de test pour évaluer l’apprentissage. Il est conseillé de mélanger les données avant d’entamer une nouvelle *epoch*, passer les mêmes *batch* dans le même ordre pouvant entraîner du surapprentissage.

4.2.2 Les choix algorithmiques

4.2.2.1 Choisir une librairie

Au cours des dernières années, 5 à 10 librairies de Deep Learning ont occupé le haut de classement, parmi lesquelles Tensorflow, Keras, Caffe, Apache MXnet, PyTorch, Microsoft Cognitive Toolkit. Généralement, Tensorflow est conseillé car massivement utilisé et documenté. Mais plusieurs raisons ont motivé le choix de Apache MXNet. D’abord, l’un des critères est que la librairie est écrite et interfacée en C++ ce qui devrait permettre l’intégration du projet aux logiciels de Seaviews. Or, les librairies sont majoritairement tournées vers le langage Python, le choix est donc réduit. Ensuite, les évolutions très rapides font apparaître de nouvelles librairies, certaines fusionnent et d’autres ne sont plus maintenues. Pour pallier ces aléas, le système *Open Neural Network Exchange* (ONNX) permet de basculer d’un environnement à l’autre (transfert de modèle et d’apprentissage) pour les librairies compatibles. La librairie MXNet l’est et cela a entériné notre choix.

4.2.2.2 Autour du réseau de neurones

Autour du réseau de neurones lui-même et de son entraînement, se pose une série de questions. La première est peut-on faire du “**transfert learning**”, c’est-à-dire utiliser un RNA déjà entraîné sur un sujet pour qu’il apprenne plus rapidement le nôtre ? Pour cela, il faut partir d’un RNA qui accomplit une tâche similaire et formater les données selon l’entrée de ce réseau. Ce sont deux points qui posent problème dans notre situation. Malgré les immenses ressources d’internet, il n’y a pas ou peu de CNN de type régression appliqués à des images acoustiques, médicales ou sismologiques (signaux multidimensionnels). De plus, si l’entrée des CNN est de dimension 224x224, l’adaptation des WCI de 256 à 224 colonnes passe par une décimation ou une fusion des signaux colonnes mais également par la re-labellisation de l’ensemble du jeu de données. De ce fait, nous avons choisis de “**partir de zéro**” pour nos apprentissages, même si cela n’est généralement pas recommandé (Géron; 2017).

L'initialisation des paramètres est donc à prévoir quand l'apprentissage part de zéro. Une initialisation aléatoire des poids et biais est préférable à une initialisation à 1 ou à 0 et le tirage peut suivre des lois classiques (normale, uniforme) ou être spécifiquement pensé pour des neurones d'un réseau. C'est le cas de la méthode *Xavier* (Glorot and Bengio; 2010) qui propose d'initialiser les poids d'une couche i (les biais sont initialisés à 0) avec une loi uniforme sur $[-1/\sqrt{N_{i-1}}, 1/\sqrt{N_{i-1}}]$ avec N_{i-1} le nombre de neurones à la couche précédente. Dans MXNet, nous utilisons une variante dans laquelle la loi uniforme est sur $[-3/\sqrt{0.5(N_{i-1} + N_{i+1})}, 3/\sqrt{0.5(N_{i-1} + N_{i+1})}]$, ce qui prend en compte le nombre de neurones dans la couche précédente et suivante. L'intérêt de la méthode *Xavier* est qu'elle conserve la même variance dans les poids d'une couche à l'autre et évite que le gradient, lors de la rétropropagation, tendent vers 0 ("vanishing gradient") ou vers l'infini ("exploding gradient").

L'ajustement des paramètres est réalisé à chaque rétropropagation du gradient de la fonction *loss*, qui elle est calculée après le passage d'un batch de données dans le réseau. La fonction de perte choisie pour un problème de régression est la *L2Loss* qui fait la somme sur la taille du batch, b_s , des carrés des différences entre les M caractéristiques de sorties des matrices d'étiquettes, E , et de valeurs prédites, O :

$$\ell = \frac{1}{2b_s} \sum_{n=1}^{b_s} \sum_{m=1}^M (O_{n,m} - E_{n,m})^2.$$

Un *optimizer* modifie ensuite les paramètres de manière à réduire la fonction *loss*. Celui utilisé, *AdaDelta* (Zeiler; 2012) est une amélioration de *AdaGrad* qui peine à atteindre la fin de la convergence car le taux d'apprentissage est diminué trop rapidement. *AdaDelta* n'est pas la seule variante de *AdaGrad* mais il semblerait qu'elle soit la plus rapide à converger d'après Ruder (2016). Précédemment, dans l'équation 4.6 d'ajustement des paramètres du RNA, le même taux d'apprentissage η était appliqué à l'ensemble des paramètres. Dans *AdaGrad*, ce taux est adapté à chaque paramètre q et à chaque itération (indicée u) par l'accumulation des gradients des itérations précédentes, notée $Sg(u)$:

$$q_{u+1} = q_u - \Delta q_u \quad \text{où} \quad \Delta q_u = \frac{\eta g_u}{\sqrt{Sg(u) + \zeta}} \quad \text{avec} \quad Sg(u) = \sum_{v=0}^u g_v^2 \quad \text{et} \quad g_u = \frac{\partial \ell_u}{\partial q_u}.$$

L'accumulation prévoit une petite valeur ζ pour éviter une division par zéro. C'est cette accumulation de tous les gradients depuis le début de l'entraînement qui entrave la convergence au long terme car le taux d'apprentissage devient de plus en plus petit. La solution apportée par *AdaDelta* est de limiter cette accumulation aux gradients récents par une moyenne glissante pondérée :

$$Sg(u) = \Upsilon Sg(u-1) + (1 - \Upsilon) g_u^2 \quad \text{avec} \quad Sg(0) = 0 \quad \text{et} \quad \Upsilon < 1.$$

De plus, le taux d'apprentissage η est remplacé par l'accumulation des ajustements du paramètre $Sa(u)$, ce qui donne :

$$q_{t+1} = q_t - \frac{\sqrt{Sa(u-1) + \zeta}}{\sqrt{Sg(u) + \zeta}} g_u \quad \text{avec} \quad Sa(u) = \Upsilon Sa(u-1) + (1 - \Upsilon) \Delta q_u^2 \quad \text{où} \quad Sa(0) = 0.$$

Les réglages de cette méthode, (ζ, Υ) peuvent conserver les valeurs par défaut de MxNet, à savoir $(1e^{-5}, 0.9)$; au contraire d'un taux d'apprentissage dont le réglage est délicat (Géron; 2017).

L'évaluation au fil de l'apprentissage est assuré par une mesure d'écart entre la prédiction et le label. Dans notre cas, l'évaluation est réalisée à chaque *batch* de données et mesure la racine de l'écart quadratique moyen (RMSE), c'est-à-dire :

$$RMSE = \sqrt{\frac{1}{b_s} \sum_{n=0}^{b_s} \sum_{m=1}^M (O_{n,m} - E_{n,m})^2}.$$

Cette évaluation est également faite à la fin d'une *epoch* sur un lot de données issus d'un jeu de validation pour suivre les progrès de l'apprentissage.

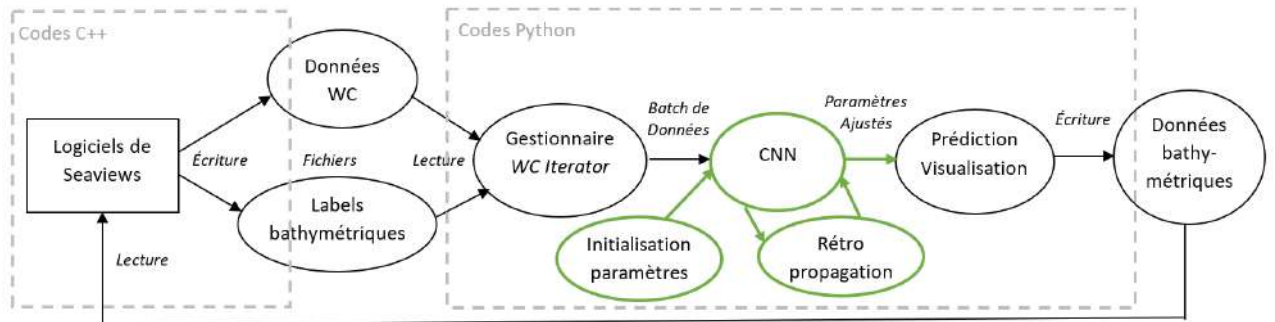


FIGURE 4.10 – Ensemble des éléments du projet d'apprentissage profond sur les données *water column* pour en déduire la bathymétrie. Les éléments verts sont issus de la librairie MxNet, tandis qu'en noir ce sont des codes que nous avons développés au niveau de la chaîne de traitement global.

La figure 4.10 résume les notions abordées dans cette partie de mise en place d'un apprentissage supervisé. L'utilisation du *Deep Learning* ne se limite donc pas à la simple construction d'un réseau de neurones.

4.3 Premiers résultats

4.3.1 Description de l'apprentissage

Le CNN utilisé est un ResNet50, constitué donc de 50 couches (18 au niveau de la figure 4.8). Pour l'appliquer à notre problème, seule la dernière couche (celle de sortie) est modifiée. Initialement, cette couche est une couche totalement connectée à 1000 valeurs (classification pour ImageNet à 1000 types d'animaux) avec une activation *softmax*, elle reste totalement connectée mais à 256 valeurs de sortie avec une activation linéaire.

Le jeu de données d'apprentissage se compose de 32 fichiers totalisant environ 56000 WCI labellisées issues de diverses campagnes d'acquisition. La composition ainsi réalisée fait intervenir plusieurs formes de fonds à différentes profondeurs et avec des réglages de sondeurs variés (quelques exemples sont fournis dans la figure 4.11). De plus, certaines WCI sont bruitées ou présentent une activité au-dessus du fond (banc de poissons) amenant encore des variations.

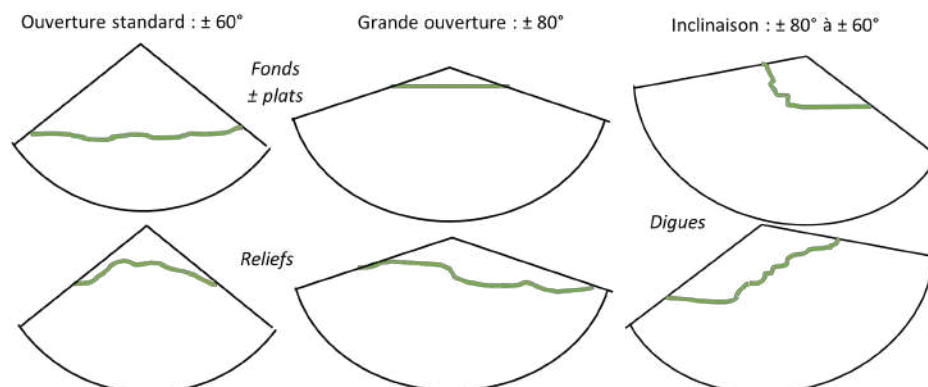


FIGURE 4.11 – Diversité des formes de fonds - fonds plus ou moins plats, reliefs plus ou moins marqués, digues, etc. - et de réglages du sondeur - ouverture angulaire, inclinaison, position plus ou moins haute du fond dans les WCI - composant le jeu d'apprentissage.

L'entraînement a été effectué avec des *batch* de 32 WCI durant 500 *epochs* (soit environ 875000

prises à jour des paramètres), et a duré un mois. La courbe de suivi d'apprentissage est en figure 4.12. L'apprentissage a été arrêté par choix ; faire durer l'apprentissage jusqu'à ce que la fonction de coût atteigne zéro est assez illusoire et augmente le risque de sur-apprentissage. Au bout d'un mois, la décision a été prise d'arrêter l'entraînement car les valeurs d'évaluation sur le jeu d'apprentissage décroissaient de plus en plus lentement tandis que celles sur le jeu de validation stagnaient autour des mêmes valeurs lors des 300 dernières *epochs*.

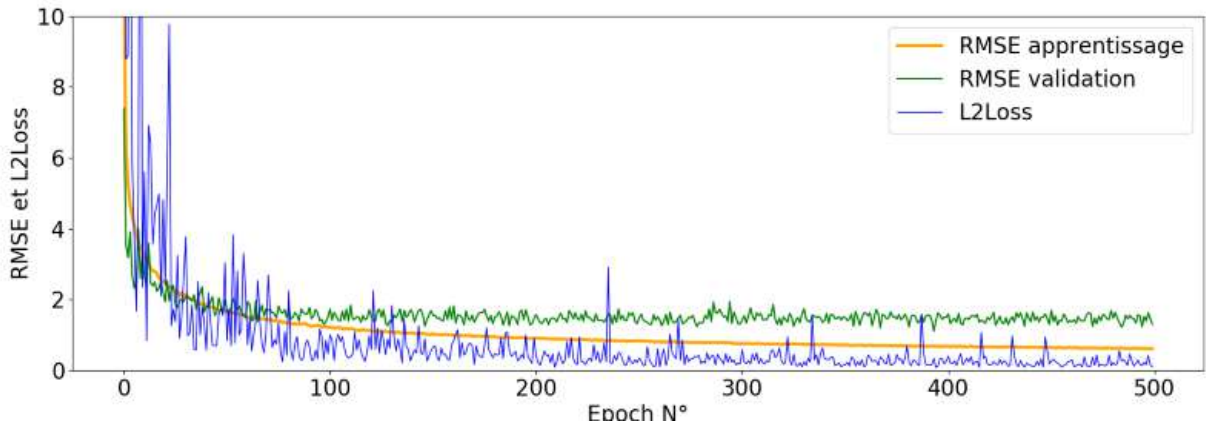


FIGURE 4.12 – Suivi des valeurs de *loss* et de RMSE à la fin de chaque *epoch* de l'apprentissage.

4.3.2 Évaluation sur fonds plats

La première application du réseau est réalisée sur une petite zone aux fonds très plats dont les profondeurs varient lentement de 17 à 25 m comme sur la figure 4.13 à gauche. Le jeu se compose de six transects tête-bêche espacés de 20 m et dont la portée d'un ping est de 40 à 50 m de part et d'autre du navire, il y a donc beaucoup de recouvrement d'un transect à l'autre. Le contenu des images *water column* est assez monotone et il n'y a pas de changement de réglages du sondeur lors de l'acquisition (ouverture angulaire ou longueur de l'image). Cependant quelques WCI sont manquantes, bruitées ou contiennent un banc dense de poissons.

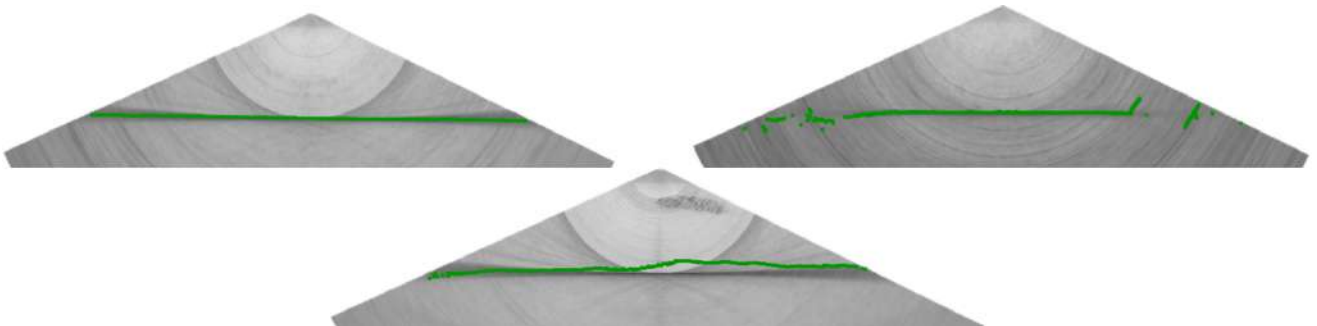


FIGURE 4.13 – Trois exemples de WCI avec une bathymétrie prédite. Sans complication comme à gauche, la prédiction est correcte. En présence de bruit (à droite) ou d'une activité dans la colonne d'eau (en bas), la prédiction perd en précision.

Sur l'ensemble de ce jeu, il y a en moyenne 0.15 m d'erreur moyenne sur un ping de 256 sondes bathymétriques. L'erreur moyenne est calculée pour un ping par $erreur = \sum_{m=1}^{256} |y_m - e_m| / 256$. Les pings bruités sont parfois difficilement lisibles même pour un opérateur, ce qui explique que l'erreur moyenne sur un ping puisse atteindre 1 m en moyenne. En revanche, lorsque le navire passe

au-dessus d'un banc de poissons, l'image reste lisible mais le CNN peine à prédire la bathymétrie et peut avoir jusqu'à 1.80 m d'erreur moyenne (figure 4.13).

La bathymétrie prédite est projetée dans un MNT pour être comparée avec celle issue du sondeur (figure 4.14). Majoritairement gris, le MNT d'écart montre que les prédictions sont bonnes. Seul l'extrémum positif (rouge) est présent et est situé au niveau des pings bruités et des bancs de poissons. Chaque transect a une différence positive côté (+10 à +30 cm) sur tout le côté bâbord et une différence négative (-10 à -30 cm) sur une partie du côté tribord - ces ondulations ne sont pourtant pas visibles sur les pings eux-mêmes. Mais la moyenne de l'ensemble des transects comme en grande partie ces variations. Seule la ligne extérieure dont l'extrémité bâbord n'est pas recouverte présente une bande orangée (+10 à +30cm) n'impactant que pas la zone d'intérêt (cadre noir dans la figure 4.14). Ainsi, le centre de la zone en gris (± 10 cm d'écart) valide l'ordre le plus stricte, l'ordre exclusif de l'OHI (2020), qui demanderait une précision verticale totale de plus ou moins 20cm pour ce type de profondeur.

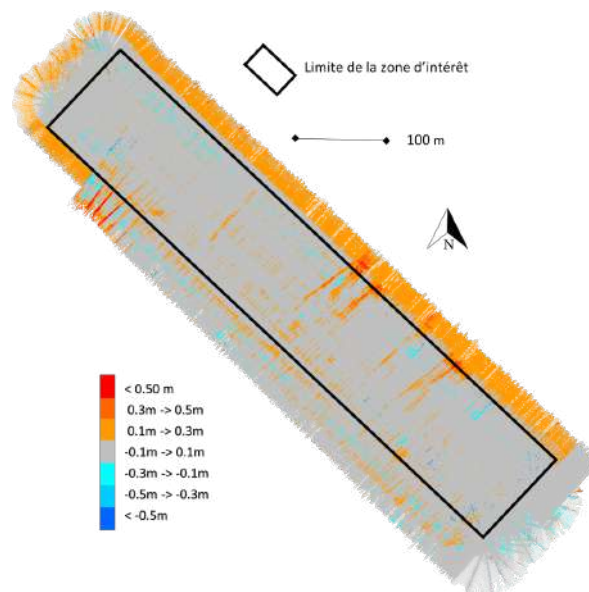


FIGURE 4.14 – Écart entre un MNT de référence (avec la bathymétrie issue du sondeur nettoyée des fausses sondes) et un MNT dont la bathymétrie est déduite des images colonne d'eau. Les couleurs chaudes désignent les mailles de MNT où la bathymétrie prédite est plus faible que la réalité, tandis que les couleurs froides désignent une bathymétrie prédite plus profonde. La couleur médiane, grise, désigne les mailles où les bathymétries estimées sont comparables.

4.3.3 Évaluation sur fond à récifs rocheux

Ce second jeu de données comporte des fonds plats de 10 à 60 m de profondeurs ainsi que des récifs rocheux s'élevant de plusieurs mètres. La diversité des formes de fonds à prédire est le facteur intéressant de ce jeu. Les douze transects, toujours tête-bêche, ne se recouvrent quasiment pas ce qui est dû à une faible ouverture des pings ($\pm 50^\circ$).

Sur l'ensemble du jeu de données, l'erreur moyenne sur un ping est de 0.4m. Les fonds plats sont plutôt bien prédits, comme c'est le cas de la WCI de droite sur la figure 4.15. Cependant, une portion complète d'un transect avec des fonds plats par 60m de profondeur a été mal prédite (voir zone rouge du MNT de droite de la figure 4.17 et les deux autres WCI de la figure 4.15) : les erreurs sont plus importantes que sur le jeu de fond plats précédents. Les reliefs, eux, ont eu plus de difficultés à être retrouvés surtout lorsqu'ils sont escarpés. Les bancs de poissons, pourtant régulièrement présents, n'ont pas affecté la qualité de l'estimation faite.

Au niveau du MNT bathymétrique, le résultat est visuellement comparable au MNT de référence

(MNT de gauche et au milieu de la figure 4.17), les différentes pierres et tombants rocheux conservent bien leurs formes malgré le manque de précision des prédictions observé sur les images précédentes. La différence des deux MNT (figure 4.17 à droite) fait ressortir des couleurs plus extrêmes qu'avec le jeu de données précédent ; il n'y a presque pas de recouvrement entre les transects permettant de moyenniser les erreurs de chacun.

4.3.4 Évaluation sur une digue

Ce dernier jeu de données correspond à une digue en travaux où des grands blocs sont également présents au pied de la digue. Les quatre transects tête-bêche couvrent la digue de 1 m de profondeur à 11 m où les blocs épars s'élèvent jusqu'à 2 m au-dessus du fond. Ce suivi demande une grande précision, le MNT en vue 3D est alors réalisé avec des mailles fines de 10x10 cm (contre 50x50cm précédemment) et présenté en haut de la figure 4.19.

Dans l'ensemble, les prédictions sont bonnes - il n'y a pas de résultats aberrants tels que ceux des figures 4.14 et 4.15 - mais manquent d'exactitude sur le chaos de bloc empilés ou passent parfois à côté des blocs épars (figure 4.18). L'erreur moyenne sur un ping est de 0.30m. Le MNT prédit en vue 3D donne un aspect très lissé à la digue et les blocs épars sont assez peu visibles (figure 4.19 au milieu). La différence des deux MNT (en bas de la figure 4.19) met en évidence la digue et les blocs tandis que le fond plats au pied de la digue - en gris - est proche de la réalité.

4.4 Évaluation des objectifs

4.4.1 De longs développements

L'idée de cette approche était de penser différemment le problème du "filtrage des sondes bathymétriques aberrantes" inhérent aux acquisitions acoustiques de nombreuses données par sondeur multifaisceaux. En effet, le filtrage consiste en une mise à l'écart par des décisions plus ou moins subjectives des données bathymétriques calculées par le sondeur sur des données brutes non-accessibles. Les données brutes qui s'en approchent le plus sont certainement les images de la colonne d'eau et celles-ci présentent un écho du fond bien identifiable. On peut voir sur les deux WCI de la figure 4.1 que les sondes en bleu sont jugées mauvaises et se situent sur des échos forts autres que le fond lui-même - qui devrait être l'écho le plus fort. Ainsi, la détection du fond dans ces images pourrait produire une bathymétrie exempte de ce type d'erreurs car le ping est traité dans son ensemble plutôt que faisceau par faisceau.

Les "miracles" produits par le *Deep Learning* incitent à aborder notre problème de détection par le biais de ces méthodes modernes. Même si, d'une part, pour les non-spécialistes, l'intelligence artificielle reste, à la fois, fascinante et nébuleuse en raison des nombreuses fictions traitant du sujet mais aussi en raison de son côté "inabordable" ou réservé à une minorité. D'autre part, les librairies informatiques proposent des produits assez "clés en main" pour les publics initiés aux langages informatiques. La réalité se situe entre ces deux extrêmes. La prise en main des outils dans un environnement Python sur des exemples fournis par des tutoriels s'est révélée d'une grande facilité. En revanche, l'intégration de la librairie dans les logiciels de Seaviews pour un développement en C++ est resté infructueux. Nous avons alors exporté nos données à partir de fonctions écrites dans les logiciels en C++ (figure 4.10) afin de travailler en Python. La mise en place de la méthode d'apprentissage adaptée aux données WCI résulte ensuite de laborieux essais.

4.4.2 CNN *from scratch*

Parmi les essais d'apprentissage, les premiers ont été réalisés avec des CNN créés de toutes pièces et entraînés sur des jeux de WCI à fonds plats, considérés comme plus simples que le jeu de la section des résultats. Les réseaux construits s'apparentent au réseau LeNet-5 (figure 4.4), avec des variantes portant sur la fonction d'activation (ReLU au lieu de sigmoïde) et le pooling (maximum au lieu de moyenne) avec en plus l'utilisation de couches *DropOut* et de *Batch Normalisation*. Le "tri-couche" devient alors [Convolution - Batch Normalisation - ReLu - Max Pooling] et

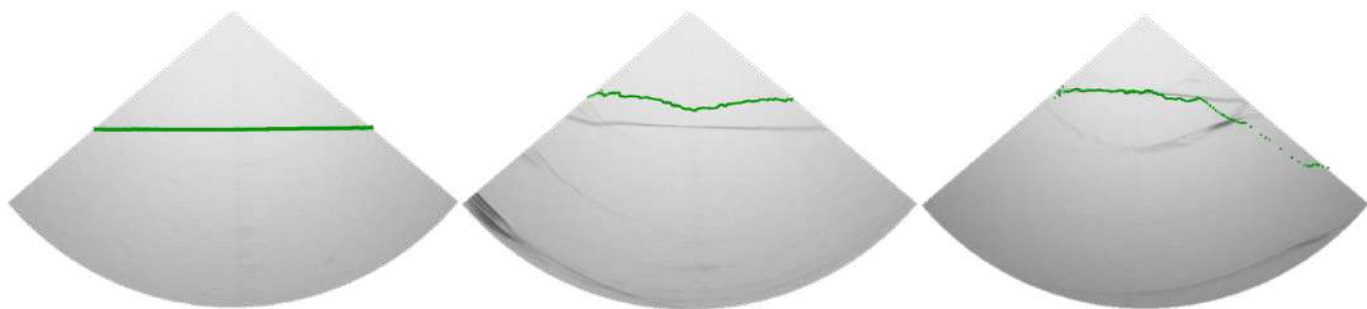


FIGURE 4.15 – Un fond plat correctement prédit à gauche. Au milieu et à droite, les deux mauvaises prédictions de fonds plats sont difficilement explicables...

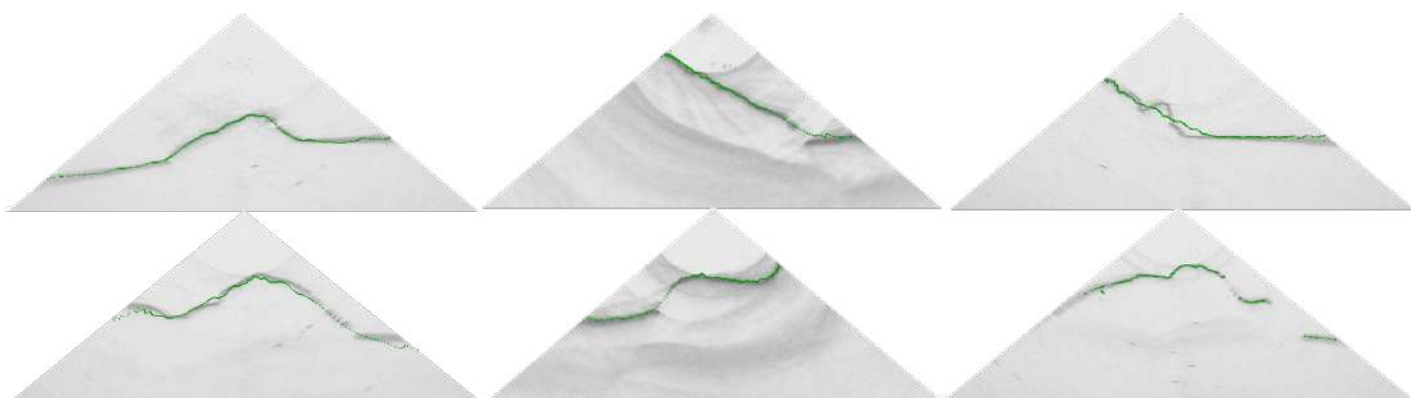


FIGURE 4.16 – Les reliefs sont prédits avec plus ou moins de précision.

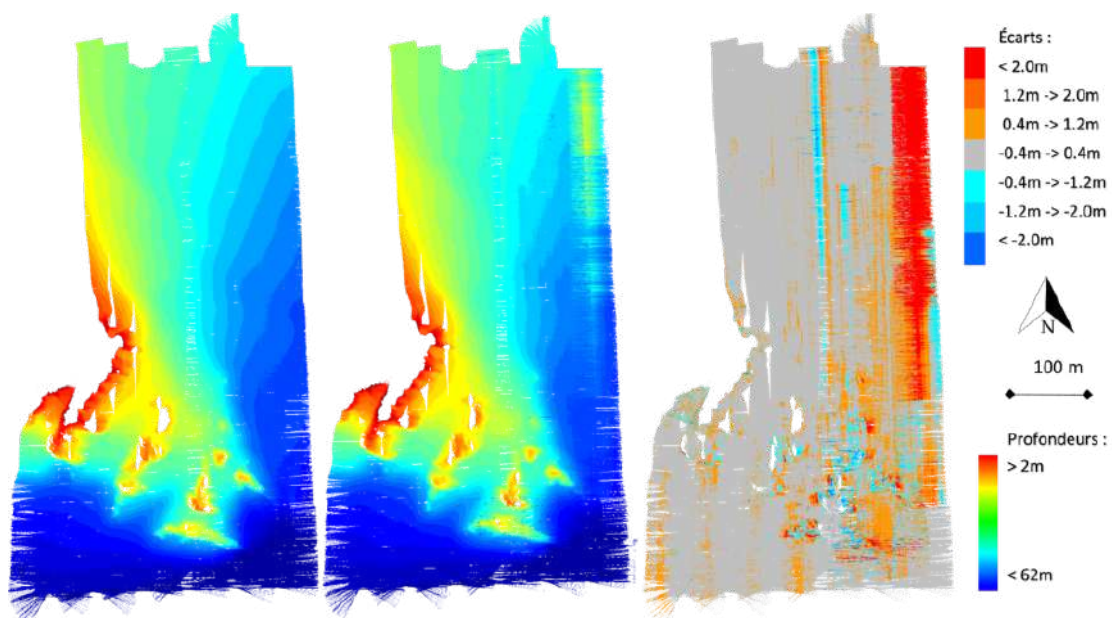


FIGURE 4.17 – A gauche et au milieu, respectivement le MNT de référence (avec la bathymétrie issue du sondeur) et un MNT dont la bathymétrie est déduite par RNA appliqué aux images colonne d'eau.

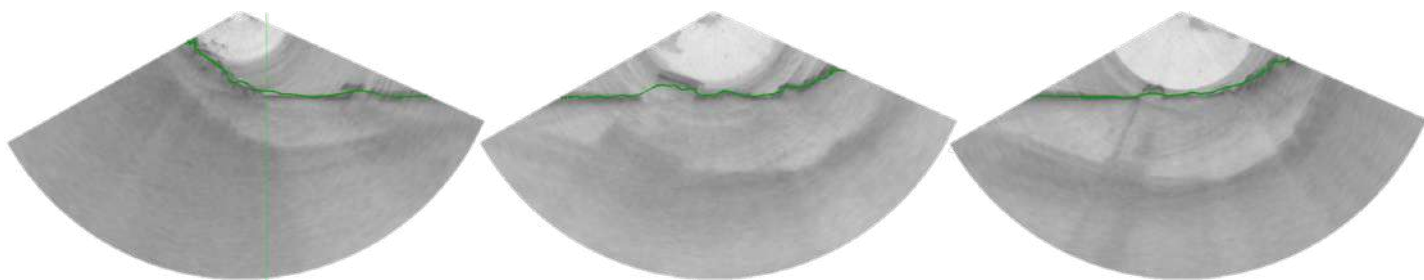


FIGURE 4.18 – Exemples de prédictions de la digue ainsi que des blocs épars.

un *DropOut* de 40% est inséré avant les FC. Mais les règles de construction d'un réseau (nombres de couches, de caractéristiques, etc.) s'approchent plus des astuces de cuisine grappillées sur des forums spécialisés plutôt qu'à des protocoles validés. Finalement, le précepte qui a guidé la construction est que le nombre d'hyperparamètres à entraîner doit être proche de la quantité de données disponible pour l'apprentissage. Cela a conduit à avoir des couches de convolution à grand noyaux, des *max-pooling* réduisant drastiquement les dimensions des caractéristiques et un goulot de rétrécissement au niveau des couches totalement connectées (figure 4.20). Les résultats sur des fonds plats sont très corrects (et moins perturbés par les bancs de poissons comme sur la figure 4.13 à gauche), mais entraînés sur des fonds plus variés, les prédictions ne sont plus aussi probantes.

Par la suite, quelques-unes des architectures présentées en figure 4.8 ont été essayées sur différents jeux de données. Les résultats encourageants ont permis d'envisager des entraînements longs sur de grands jeux de données variées.

4.4.3 Retours sur les résultats

Le jeu de données employé a été pensé pour contenir de nombreuses variations et rendre l'apprentissage le plus robuste possible face à des configurations peu communes. Mais il est difficile de représenter chaque cas, puisque le nombre de possibilités - forme des fonds et profondeurs, réglages de sondeur, présence de bruit - n'est pas facilement dénombrable.

Deux principaux avantages sont idéalement attendus de la bathymétrie extraite de la WCI. Le premier est qu'il n'y ait plus de filtrage des sondes aberrantes à réaliser. Au vu des différents résultats, cet objectif est atteint partiellement. Dans les résultats à fonds plats, les erreurs les plus importantes se sont produites en présence de bruit ou d'animation dans la colonne d'eau (figure 4.13 au milieu et à droite), ce qui laisse supposer que le jeu d'apprentissage ne contient pas encore assez de ce type d'images pour rendre l'entraînement robuste. En revanche, dans les résultats avec les récifs, les erreurs les plus aberrantes sont celles illustrées en figure 4.15 et n'ont pas trouvé d'explication.

La seconde attente concerne les sondes aux grands angles des pings. Il est courant de ne pas conserver les sondes les plus extrêmes en raison d'une baisse de précision et d'une forte occurrence d'aberrations. Lors des levés en mer, les transects sont alors resserrés pour obtenir du recouvrement entre les fauchées. Cela implique de réaliser des transects supplémentaires, d'avoir plus de données à traiter, et donc, un coût de fonctionnement plus élevé. Ici, les résultats obtenus ne semblent pas moins précis aux grands angles qu'au centre, sauf en cas de bruit qui rendent illisible la WCI.

Mais si ces deux points sont en bonne voie, celle de la précision de prédiction n'est pas réellement satisfaisante. Les fonds qui visiblement ont une prédiction cohérente présente un différentiel au moins décimétrique avec la réalité, tandis que ceux avec du relief peuvent présenter des prédictions "lissées" notamment sur les formes les plus dures et courtes. Ces prédictions lissées peuvent être dues au sous-échantillonnage - parfois très fort - des WCI qui détruit une partie des hautes fréquences de l'image.

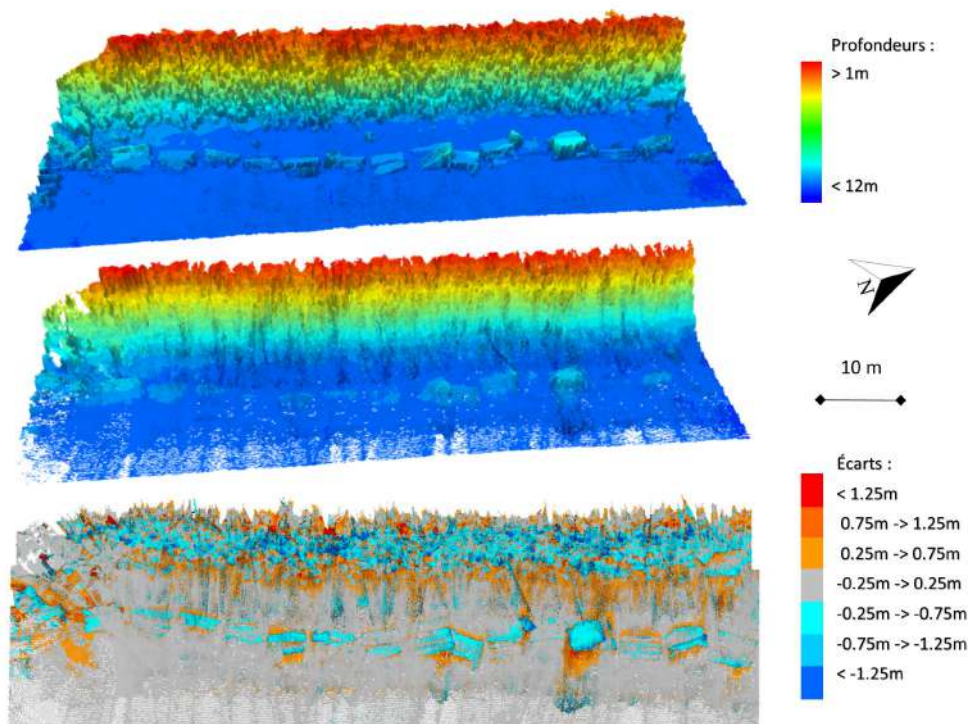


FIGURE 4.19 – En haut, la représentation 3D du MNT de référence. Au milieu, ce même MNT avec la bathymétrie déduite. En bas, différence des deux MNT.

De plus, l'erreur moyenne sur un ping selon les jeux de données est assez trompeuse. Toutes les WCI sont réduites à une taille de 256 échantillons par faisceau, mais les longueurs d'origine sont diverses. Ainsi, dans le cas des fonds très plats, avec environ 7000 échantillons à l'origine, les 15 cm d'erreur moyenne sur un ping correspondent à une précision d'un demi-échantillon sur les 256. Tandis, dans le cas du second jeu, les WCI brutes ont une longueur de 15000 échantillons, les 40 cm d'erreur moyenne correspond à une précision au trois quart d'un échantillon sur les 256. C'est pourquoi, il est difficile d'augmenter la précision de prédiction quand celle-ci a déjà atteint la résolution de la donnée. Pour une amélioration, il faut recommencer les entraînements avec des WCI moins sous-échantillonnées, l'idéal étant de travailler à pleine dimension.

4.4.4 Des variantes et des perspectives

Améliorer la précision La suite logique consiste, maintenant, à réaliser de nouveaux apprentissages sur des images plus résolues. Il n'est pas vraiment envisageable de recommencer les entraînements avec des images de plus en plus grandes car d'une part les apprentissages à plus basses résolutions ne sont pas transférables à ceux à plus hautes et d'autre part les WCI à pleine résolution peuvent avoir des longueurs variables de 1000 à 20000 échantillons. Cependant, une prédiction précise peut être réalisée en deux étapes. La première est celle réalisée ; une prédiction du fond est réalisée avec une précision relative et détermine la zone de recherche du fond dans l'image WCI à pleine résolution. Le second apprentissage - ou peut être traitement plus classique - se fait à nouveau sur une image de dimension 256x256 recadrée autour du fond précédemment prédit.

Traitements à 1, 2, 3 dimensions Les sondes bathymétriques classiques sont calculées faisceau par faisceau, le calcul se fait donc à une dimension. Le fait d'avoir travaillé à partir d'images (des

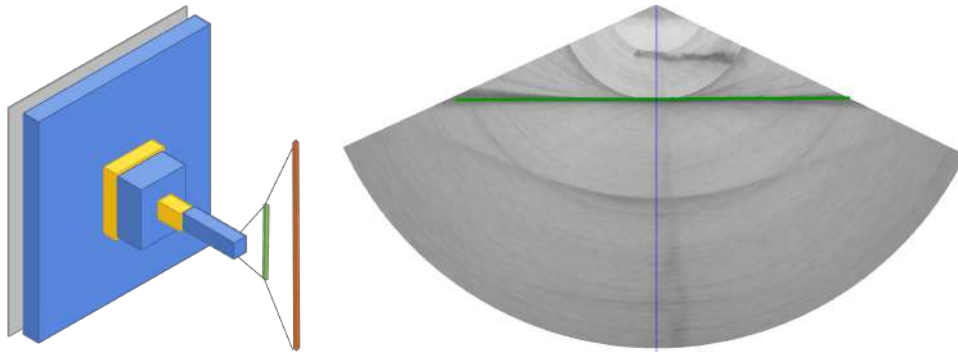


FIGURE 4.20 – A gauche, représentation du CNN *from scratch* testé, le code couleur est celui de la figure 4.8. A droite, prédiction d'un fond plat en présence d'un banc de poissons.

signaux en 2-dimensions) et avec des labellisations issues de MNT semblent avoir apporté une notion de continuité dans la série de sondes bathymétriques à prédire pour les pings. Cette notion n'existe pas dans la détection classique du fond (voir les points bleus complètement isolés de la figure 4.1). C'est donc un réel plus. Même si, la discontinuité existe quand un obstacle crée une "ombre" (voir figure 1.5), des sondes peuvent être prédites et éviter les "zones blanches".

Ici, ce sont les CNN du deep learning qui ont été testés - les CNN sont adaptés au traitement de l'image -, mais les réseaux de type LSTM (long short term memory) appartenant aux réseaux de neurones récurrents et adapté au traitement de la donnée séquentielle (écriture manuelle, parole) peuvent être envisagés pour traiter une série de pings consécutifs puisque le fond possède une contiguité dans le ping mais aussi d'un ping à l'autre. Cela serait donc un traitement en 3D qui s'apparente à du *tracking* d'objet.

Super résolution Les réseaux utilisés produisent 256 valeurs de sortie, autant que de valeurs que dans les labels et correspondant à une valeur par faisceau (une valeur par colonne de l'image). Ainsi, au lieu de prédire des coordonnées (x, y) dans l'image ou (*angle du faisceau, distance à l'origine*), on peut se contenter de la distance à l'origine. De ce fait, les labels et les prédictions sont facilement comparables valeurs à valeurs (il n'y pas de souci de prédiction de la seconde coordonnée) et également facilement visualisables sur les images. Cependant, cette bathymétrie prédite est moins dense que celle originelle car le sondeur à 256 faisceaux déduit 1024 sondes. Néanmoins, il est tout à fait possible de prédire plus de valeurs que dans dans les labels en modifiant la couche de sortie du réseau. Il faut ensuite adapter les fonctions de *Loss* et d'évaluation pour comparer deux vecteurs de longueurs différentes. Quant à la manière d'afficher plus de points pour autant de colonnes, il faudra peut-être prédire le couple de coordonnées car les sondes peuvent ne pas être équidistantes ou équiréparties.

En bref

Aborder un problème avec des images couleurs par du *Deep learning* est aujourd'hui devenu accessible. En effet, les librairies spécialisées se multiplient, laissent le choix des langages informatiques (C++, Python, Java, Matlab, etc.) et sont fonctionnelles avec des configurations matérielles variées. De même, de nombreux réseaux de neurones pré-entraînés sont disponibles pour réaliser du *transfer learning*. En cas de base de données de faible envergure, la *Data Augmentation* dédoublement les données avec de l'ajout de bruit, de flou ou des modification de translation ou rotation. La labellisation des données adaptée au problème à résoudre reste en général la principale charge de "l'apprenti sorcier" (*Data scientist, Data Analyst*).

Sortir des sentiers battus est autrement plus difficile car nos données multidimensionnelles, les images de la colonne d'eau, ne sont pas des images au sens classique du terme et le problème à résoudre, la détection du fond, est bien différent des problèmes classiquement traités. Il faut alors

partir de zéro (*from scratch*) aussi bien pour l'apprentissage que des choix de construction d'un réseau, de fonction de *Loss* et de réglages de la vitesse d'apprentissage au travers de l'*optimizer*. Les tâtonnements sont alors longs et laborieux car il n'y a pas (encore?) de protocole à suivre tel une recette de cuisine. Il est souvent reproché au *deep learning* la faible intelligibilité des données au cours de leur passage dans les RNA, les réseaux s'approchant de boîtes noires, ce phénomène est exacerbé pour des apprentissages depuis zéro. Cependant, à l'inverse du précédent paragraphe, nous disposons d'une quantité faramineuse de WCI car elles sont acquises systématiquement et d'une méthode de labellisation automatique permettant de composer efficacement des jeux de données.

Les résultats présentés sont stupéfiants au regard de l'opacité des RNA mais pas totalement satisfaisants pour un aboutissement ayant requis plusieurs mois de mise en place et de tâtonnements. La poursuite des développements doit se faire, au minimum, pour savoir s'il est possible d'obtenir une précision de prédiction équivalente à la bathymétrie classique. Ensuite, au-delà de l'aspect prospectif, il faut s'interroger sur les finalités à donner à cette méthode : pour quelles raisons remplacerait-elle ou compléterait-elle la bathymétrie classique? Peut-elle être transférée à des WCI acquises avec d'autres sondeurs acoustique? Peut-elle être embarquée directement dans le sondeur pour s'affranchir de la très lourde acquisition des WCI?

Conclusion

Lors de la réalisation d'une cartographie bathymétrique, un tri des sondes est nécessaire pour écarter les erreurs aberrantes qui n'ont rien de tangibles. Si les SMF modernes produisent moins d'erreurs que leurs ancêtres, ils acquièrent considérablement plus de données. Le besoin d'automatisation du tri n'est donc pas nouveau pour ramener les temps de traitement en deçà des temps d'acquisition. Cependant, la validation finale revient toujours à un expert qui doit pouvoir retoucher, le cas échéant, l'état des sondes.

Dans le chapitre 3, nous avons proposé une méthode automatisée de tri des sondes aberrantes. Elle s'ajoute à la longue liste des diverses méthodes existantes où elle s'apparente aux approches orientées données en empruntant des éléments aux techniques statistiques et de partitionnement. Sans être aussi évoluée que les autres méthodes, celle-ci est intégrée aux logiciels de Seaviews et permet de traiter efficacement tout un jeu de données dans un modèle numérique de terrain. Les erreurs principales sont identifiées. La validation et les retouches sont laissées au soin de l'opérateur.

Dans le chapitre 4, nous avons étudié la possibilité de réaliser autrement une cartographie bathymétrique ; en déduisant l'information bathymétrique des images de la colonne d'eau. Ce traitement est réalisé par un réseau de neurones artificiels qui apprend de manière supervisée à extraire l'information. La part importante de ce travail a consisté à préparer un jeu de données et à faire de bons choix algorithmiques. Les premiers résultats sont encourageants et il est crucial de continuer dans cette direction : les bathymétries prédites manquent encore de précision mais valident les objectifs, à savoir contenir moins d'erreurs grossières et présenter des fauchées valides aux grands angles.

Troisième partie

Classification des fonds et des habitats marins

Introduction

La classification des fonds marins correspond à une partition artificielle des fonds en différentes entités. Les activités telles que la navigation, la géologie marine ou l'écologie sont susceptibles de produire des cartes des fonds divisées en entités bien différentes : risque de talonnage (de toucher le fond) selon marées, nature des substrats - sables, vases, graviers - et leur dureté, détection d'herbiers et de roches, etc. Ce sont les applications en écologie marine qui sont traitées dans cette partie. Il est alors question d'une double analyse, celle de l'identification de la nature des substrats et celle de la caractérisation des habitats marins présents pour produire une carte de classification des habitats benthiques. L'analyse ne se limite donc pas uniquement à l'identification purement physique des matériaux recouvrant les fonds mais s'étend aux espèces vivants dans un même environnement (le biotope). En Europe, les habitats - naturels et artificiels, terrestres et marins - sont inventoriés et décrits à différentes échelles de précision dans la classification EUNIS (European Nature Information System) (Gayet et al.; 2018).

La production de cartes des habitats marins participent au management des écosystèmes et à la gestion des ressources naturelles (Cogan et al.; 2009). Depuis 2008, ce management est devenu une directive européenne, la Directive-Cadre Stratégie pour le Milieu Marin (DCSMM) (Long; 2011), avec pour objectifs d'assurer le bon état écologique du milieu marin et l'amélioration de l'état de conservation de la biodiversité marine. En France, les agences de l'Office Français pour la Biodiversité sont responsables de la gestion des Aires Marines Protégées et encouragent les études d'amélioration des connaissances et de suivi.

Classification par systèmes acoustiques Les premières classifications des fonds ont pu être réalisées à l'aide des sonars à balayage latéral (SBL) dès les années 1970. Ceux-ci fournissent une image de réflectivité du fond témoignant des différences physiques des matériaux présents (Blondel; 2009). Ils permettent, entre autres, la détection de mines et d'épaves ou bien des études de discrimination de sédiments - vases, sables, graviers, galets - (Brown et al.; 2002).

L'analyse des échos issus des systèmes acoustiques monofaisceaux a conduit à l'apparition de méthodes de discrimination des fonds (*Système Acoustique de Classification Automatique des Fonds* ou SACLAF) ou en anglais, *Acoustic Ground Discrimination System* (AGDS)). Les méthodes Roxann, Echoplus et QTC View s'intéressent au premier et au second échos pour déterminer, respectivement, la rugosité et la dureté des fonds (Wilding et al.; 2003). Mises à contribution dans différents domaines, elles ont permis de trouver des épaves (Lawrence; 2001), d'évaluer la densité de laminaires (Mielck et al.; 2014) ou de distinguer différents herbiers marins (Sabot et al.; 2002).

Cependant, les systèmes monofaisceaux possèdent l'inconvénient de la couverture si on les compare à l'utilisation d'un SBL ou de sondeurs multifaisceaux (SMF). Comme leur faisceau est un cône d'ouverture typiquement de l'ordre 10-30°, il n'y a pas de levés exhaustifs des fonds. Et, de plus, une augmentation de profondeur s'accompagne d'une augmentation du pied de faisceau, ce qui réduit la résolution des caractéristiques releves par ces seuls systèmes. Les combinaisons de différents systèmes acoustiques ont été exploitées, comme le présente le travail de revue dans (Brown et al.; 2011). Les SACLAF ont pu être couplés avec des SBL pour des comparaisons de performance sur la discrimination de sédiments (Collier and Brown; 2005), mais également avec des SMF pour la cartographie précise de récifs rocheux dans la Manche (Diesing et al.; 2009). Des combinaisons encore plus complexes, intégrant système bathymétrique (SMF ou monofaisceau), SACALF, SBL et sondeur de sédiments ont aussi été testés pour étudier finement des habitats tels

que les récifs de coralligène en mer Méditerranée (Georgiadis et al.; 2008) ou des récifs couverts de moules en Irlande (Lindenbaum et al.; 2008).

Depuis quelques années, les acquisitions ont montré que les images de réflectivité des SMF ont atteint un niveau de qualité comparable à celles acquises par des SBL (Le Bas and Huvenne; 2009), (Brown and Blondel; 2009). Le SMF peut donc être utilisé comme seul outil acoustique pour classifier les fonds (McGonigle et al.; 2009). C'est cette solution qui nous intéresse dans cette partie. Les méthodes sont toutefois encore en développement et on ne recense pas, pour l'instant, de méthodes commerciales comme Roxann.

Complément par moyens optiques Aujourd'hui, l'étude des tous petits fonds - jusqu'à une vingtaine de mètres - peut également être réalisée sur des très grandes échelles à l'aide d'images aériennes et/ou satellitaires comme cela se fait en classification des habitats terrestres. Leur détournement en milieu marin soulève des difficultés comme l'atténuation des couleurs qui peut être corrigée par mesures, entre autre, de la matière organique et des particules en suspension (Minghelli-Roman and Dupouy; 2014). Ces calculs sont nécessaires pour retrouver la signature des végétaux sous-marins, notamment les fréquences du proche infra-rouge absorbées dès le(s) premier(s) mètre(s) d'eau (Minghelli et al.; 2021). Néanmoins, dans notre cas, elles apportent un complément d'information aux systèmes acoustiques dans leur classification des fonds. De même, les techniques de lidar permettent d'avoir une donnée bathymétrique de 0 à 30 m maximum et peut être traitée comme de la bathymétrie acoustique pour conduire à une classification des habitats (Zavalas et al.; 2014). Ces données peuvent remplacer ou compléter l'acquisition acoustique par très petits fonds (Wang et al.; 2018), où il existe un réel risque de talonnement et où l'acquisition exhaustive par SMF devient très longue.

Méthodes de classification des données De même que les moyens instrumentaux sont variés pour produire des cartes des habitats marins, les méthodes utilisées pour exploiter les données sont tout aussi diverses. On différencie d'abord les méthodes par interprétations expertes, mêlant connaissances et données, des méthodes se basant uniquement sur le contenu des données. Les premières sont chronophages, subjectives et non répétitives : tout le raisonnement est fait par des experts qui s'aident des outils numériques pour segmenter l'espace selon les habitats à représenter (Kostylev et al.; 2001). Parmi les méthodes basées sur les données, on différencie les méthodes s'intéressant aux signaux acoustiques des méthodes avec une approche image pour les données issues des SMF et des SBL (Diesing et al.; 2016). L'approche signal est celle des AGDS pour les sondeurs monofaisceaux. Très peu de méthodes ont été développées pour les SMF car la compensation de l'angle des faisceaux dans la déformation du signal est complexe et très variable (profondeur, type de substrat, réglage du sondeur, etc.). L'approche image est celle que nous allons développer dans cette partie, les méthodes de classification issues de l'apprentissage machine ou *Machine Learning* (ML) sont appropriées pour ce type de données.

Le contenu de cette partie s'organise en deux chapitres. Le chapitre 5 constitue un état de l'art sur les processus de classification des fonds marins par SMF. Dans une première section est rappelée la chaîne de traitement pour obtenir une carte des habitats benthiques, ce qui permet de présenter les données utilisées ainsi que les principales méthodes employées. Puis, la méthodologie de Seaviews est expliquée, les données sont exploitées à l'aide d'une méthode relativement peu automatisée. Dans une seconde section, une sélection de méthodes de classification issues du ML est proposée pour être mises en œuvre dans la suite logicielle ViewMap de Seaviews. Leur fonctionnement est d'abord décrit, puis nous expliquons comment ces méthodes sont paramétrées. Le chapitre 6 propose une mise en application des méthodes sur des données expérimentales réelles. La zone d'étude et les données acquises sont présentées. Elles serviront de test pour la méthode manuelle de Seaviews ainsi que pour ces méthodes davantage automatisées. Les résultats obtenus sont analysés et comparés pour permettre de fixer les prochaines perspectives d'amélioration de cette approche au sein de la suite logicielle de Seaviews.

Chapitre 5

Etat de l'art de la classification des fonds par acoustique

5.1 Étapes de la chaîne de classification des fonds marins

Il n'existe pas de méthode universelle pour réaliser des cartes de classification des fonds marins en raison des multiples approches et des diverses combinaisons de capteurs et de données. Cependant, on peut quand même dégager un schéma commun constitué de grandes étapes pour réaliser une classification :

1. Une mise en forme des données depuis leur acquisition jusqu'à leur exploitation comprenant le géoréférencement des données SMF, la fusion de données issues de différentes sources, une éventuelle recalibration des valeurs, la gestion des données manquantes ou sous différents formats, etc. ;
2. Une analyse des différentes données acoustiques et une extraction des caractéristiques principales ;
3. La classification elle-même, utilisant divers techniques et algorithmes ;
4. Un éventuel post-traitement pour finaliser les résultats ;
5. Une évaluation des performances selon les résultats obtenus.

Ce principe de traitement est détaillé dans cette section. Nous verrons ensuite la méthode utilisée par Seaviews ainsi que les points à améliorer.

5.1.1 Traitements préliminaires des données

5.1.1.1 Mise en forme des données

Cette première étape de mise en forme des données SMF acquises (bathymétriques et réflectivité) consiste tout d'abord à combiner les données acoustiques aux données des autres instruments, puis à neutraliser les pings dont le géopositionnement est mauvais ainsi que les sondes aberrantes. A ce stade, les données bathymétriques peuvent être projetées dans un modèle numérique de terrain (MNT) dans lequel la dimension des mailles est adaptée à la zone d'étude, aux profondeurs, et aux objets d'intérêts.

Dans un second temps, on applique un TVG adapté sur les données de réflectivité pour corriger au maximum les effets d'angles. Cependant, la correction d'un TVG n'est pas toujours suffisante : il reste en général les traces de navigation, lignes très marquées au nadir ainsi que les deux extrémités de fauchées où le bruit est le plus fort. Un exemple d'image de réflectivité est proposé dans la première partie en figure 2.13. [Lamarche et al. \(2011\)](#) proposent de nettoyer ces traces par un calcul faisant intervenir des paramètres issus de la réflectivité. De plus, il peut être intéressant de réaliser une calibration de ces données pour que les valeurs soient standardisées d'une zone à

l'autre pour un même type de fond. Le logiciel SonarScope de l'IFREMER (2007) le permet en post-acquisition. Enfin, les données de réflectivité peuvent être projetées dans un MNT, on parle d'une mosaïque. L'idéal est de conserver la même dimension de maille de MNT ainsi que le type de projection pour la bathymétrie et la réflectivité. Ces premières étapes sont souvent réalisées dans les logiciels spécialisés pour le traitement des données acoustiques des SMF (Hypack, logiciels de QPS, etc.).

En plus des données acoustiques, le levé s'accompagne, en général, de *vérités terrain* réalisées au moyen d'appareils photo, de caméras vidéos tractées, de plongeurs, de prélèvements, etc. (van Rein et al.; 2009). Ces données-là sont ponctuelles et ne couvrent qu'une toute petite surface de la zone d'étude. Elles ont pour but de faire l'inventaire des habitats et des substrats présents, puis d'identifier la signature acoustique qu'ont ces fonds pour pouvoir ainsi associer un habitat à chaque classe. Les données de vérité terrain doivent avoir une position géographique suffisamment précise pour ne pas engendrer d'erreur d'association substrat-signature. Ces données non-acoustiques sont manipulées dans d'autres logiciels, de type Système d'Information Géographique (SIG) tels qu'un des plus connus, ArcGis développé par la société Esri.

5.1.1.2 L'extraction des caractéristiques

Cette seconde étape consiste à extraire des données acoustiques, bathymétrique et de réflectivité, des données secondaires reflétant différentes informations spécifiques. En tant que telles, la bathymétrie et la réflectivité sont déjà des éléments classifiants. A très large échelle, selon les profondeurs, on peut distinguer les estuaires des côtes ou les côtes des zones de grandes profondeurs. De même, une réflectivité calibrée permet l'identification des substrats. Cependant, les algorithmes de classification demandent généralement en entrée des caractéristiques plutôt que des données brutes.

Selon les revues méthodologiques de Brown et al. (2011) et Diesing et al. (2016), la bathymétrie permet de déduire des informations morphologiques des fonds par des calculs tels que la pente, la rugosité, l'orientation, la variabilité de terrain, etc, tandis que la réflectivité permet de déduire la variance ainsi que plusieurs indices de textures (rugosité, matrices de co-occurrence, etc.). Les calculs sont souvent réalisés au moyen de produits de convolution sur les mailles des MNT de bathymétrie et de réflectivité dans des logiciels de SIG comme ENVI ou ArcGis. Ainsi, à partir de ces deux informations brutes, on obtient *artificiellement* n -couches d'information 2D qui vont permettre d'établir une partition de la zone d'étude en classes.

Des approches utilisant la réponse angulaire du signal de réflectivité autour du fond ont aussi été testées pour produire des signatures acoustiques caractérisant les fonds. Parmi celles-ci, on trouve les *Angular Responses Analysis* (Fonseca et al.; 2009) ou l'écho-intégration sur les faisceaux proches du nadir d'un SMF (Abadie et al.; 2017). Ces approches sont assez peu répandues car la compensation de l'angle des faisceaux dans la déformation du signal est complexe et très variable selon les profondeurs, la nature des substrats, les réglages du sondeur, etc.

Des dizaines d'indices peuvent donc être produits à partir des données brutes du SMF, certains étant redondants ou peu informatifs. Une étape intermédiaire de sélection des caractéristiques les plus pertinentes peut être faite avant la classification par une analyse en composantes principales ou en étudiant les valeurs d'auto-corrélation entre les indices pour réduire le nombre de données produites et alléger les calculs de classification.

Par exemple, dans Hasan et al. (2014), les auteurs extraient six dérivées de la bathymétrie, six dérivées de la mosaïque de réflectivité et quatre caractéristiques de la réponse angulaire des signaux de réflectivité. Cela fait un total de dix-huit caractéristiques en comptant les données brutes. Par la suite, lors de la classification, les auteurs essaient plusieurs combinaisons de caractéristiques pour déterminer lesquelles sont les plus discriminantes et les plus redondantes ; ce qui se vérifie par des calculs d'auto-corrélation entre caractéristiques.

5.1.2 La classification et son évaluation

5.1.2.1 Méthodes de classification

La première question à se poser pour réaliser une classification est l'échelle à laquelle elle est réalisée. Les données étant représentées sous forme de MNT dont l'unité est la maille - de dimension définie précédemment-, une classification maille par maille semble une approche naturelle. Mais, d'une part, plus les MNT sont grands et résolus, plus le nombre de maille est grand et plus les calculs sont longs, et d'autre part, la maille n'a aucune réalité géographique (Diesing et al.; 2016). C'est pourquoi, une approche basée objet peut être préférable, et dans ce cas, une étape de segmentation des données en objets est à réaliser avant une classification (Hay and Castilla; 2008). L'avantage supplémentaire de la segmentation est la réduction du bruit (type tavelure ou *speckle*) qu'il peut y avoir d'une maille à l'autre car les objets ont alors des valeurs plus homogènes.

L'emploi de méthodes objectives pour réaliser une classification est un choix décrit comme préférable car cela rend les résultats plus reproductibles et plus généralisables (Anderson et al.; 2008). Les méthodes objectives incluent les outils d'analyse exploratoire des données et des méthodes de classification par apprentissage automatique. Les outils d'analyse aident un opérateur à la sélection des caractéristiques (analyse en composantes principales, analyse de redondance, analyse discriminante) ou à la séparation en classe des données (positionnement multidimensionnel) (Simard and Stepnowski; 2007). A l'inverse, les méthodes d'apprentissage automatique réalisent leur tâche en autonomie; la machine "apprend un raisonnement" à partir de données via des outils issus des mathématiques, des statistiques et de l'informatique et en amont, l'opérateur prépare les données puis il évalue en aval les résultats.

On distingue deux approches principales parmi ces méthodes : l'apprentissage supervisé et l'apprentissage non-supervisé. L'apprentissage supervisé de classement (classification en anglais) consiste à ranger les données par catégories déjà constituées. Le terme de "classification" en français désigne l'art de concevoir des critères de classement des données traduit par "clustering" en anglais (ou "partitionnement"). Cela correspond à un apprentissage non-supervisé. Dans la pratique, l'apprentissage supervisé est utilisé quand on connaît *a priori* la catégorie de certaines données. La méthode établit alors un modèle pour généraliser ce qu'elle déduit des exemples catégorisés pour l'appliquer sur de nouvelles données. Dans le cas d'apprentissage non-supervisé, la méthode découvre par elle-même la structure des données et propose une manière de les organiser en catégories homogènes.

Une grande partie des résultats de classification des fonds marins que l'on trouve dans la littérature scientifique est obtenue avec un nombre restreint de logiciels de SIG (ENVI, ArcGis) ou de logiciels de calcul (Matlab, R) et de bibliothèques (LIBSVM, mlpack, OpenCV) qui intègrent les méthodes les plus répandues (Tableau 5.1). L'emploi des méthodes non-supervisées est en général réservé à une phase exploratoire des données, par exemple pour estimer si le nombre de classes à considérer est cohérent avec les habitats *a priori* répertoriés ou parce qu'il y a trop peu de vérités terrain pour réaliser un apprentissage supervisé. Les catégories créées par ces méthodes peuvent être difficiles à interpréter car elles sont le résultat d'un raisonnement de la machine complètement différent du nôtre. C'est pourquoi, dans ce contexte, l'apprentissage supervisé est plus souvent rencontré car les classes sont constituées d'après l'expertise de l'opérateur et les vérités terrain et cela facilite l'évaluation des résultats (voir paragraphe suivant). Selon les études et comparaisons de méthodes du tableau 5.1, il semblerait que les méthodes de "forêts aléatoires" et de "séparateurs à vastes marges" soient les plus efficaces.

5.1.2.2 Post-traitements

Même si une méthode de classification réalise un bon découpage des données en différentes catégories, le résultat n'est généralement pas parfait et le réajustement des paramètres de la méthode n'apporte pas toujours d'amélioration. Cependant l'application de post-traitements classiques d'image permet bien souvent de peaufiner le rendu (figure 5.1). Par exemple, le filtrage du bruit ou des pixels isolés améliore les résultats en réaffectant la classe majoritaire du voisinage et

TABLE 5.1 – Diversité des méthodes d'apprentissage automatique utilisées sur des données acoustiques (et lidar) pour produire des cartes d'habitats marins. (UW : Université du Wisconsin-Madison). Méthode supervisée (S), méthode non-supervisée (NS)

Références	Méthodes (logiciel ou librairie)	Type de méthode
Marsh and Brown (2009)	Cartes Auto-Adaptatives (Matlab)	NS, réseau de neurones artificiels
Ierodiaconou et al. (2011)	Maximum de vraisemblance (Envi) QUEST (logiciel de UW) et CRUISE (Envi)	S, inférence de lois de probabilités S, arbres de décisions
Lucieer and Lamarche (2011)	K -moyennes "floues"	NS, variante des K -moyennes (cf 5.2.1.1)
Hasan et al. (2012)	Forêts aléatoires (Matlab) Séparateur à Vastes Marges (LIBSVM) QUEST (logiciel de UW) Maximum de vraisemblance (Arcgis)	S, cf 5.2.2.2 S, cf 5.2.2.1 S, arbre de décisions S, inférence de lois de probabilités
Stephens and Diesing (2014)	K plus proches voisins (R) Séparateurs à Vastes Marges (R) Forêts aléatoires (R) Perceptron multi-couches (R) Classifieur bayésien naïf (R)	S, déduction de classes selon les voisins S, cf 5.2.2.1 S, cf 5.2.2.2 S, réseau de neurones artificiels S, cf 5.2.2.3
Calvert et al. (2015)	ISO Cluster (Arcgis) Maximum de vraisemblance (Arcgis)	NS, variante des K -moyennes (cf 5.2.1.1) S, inférence de lois de probabilités
Wang et al. (2018)	Séparateurs à Vastes Marges et variantes	S, cf 5.2.2.1
Held and Schneider von Deimling (2019)	Forêts aléatoires (mlpack)	S, cf 5.2.2.2

simplifie visuellement l'aspect (Lu and Weng; 2007). De la même manière, les objets (ensemble contigu de mailles de même classe) peuvent être reclassés selon leur voisinage quand ceux-ci ont une surface trop petite (Costa and Battista; 2013). Enfin, un lissage des limites réduit le bruit des contours des objets et connecte en une même entité les objets très proches spatialement. L'intérêt de ce type de post-traitements est de généraliser les résultats et d'alléger le raster (l'image matricielle) produit.

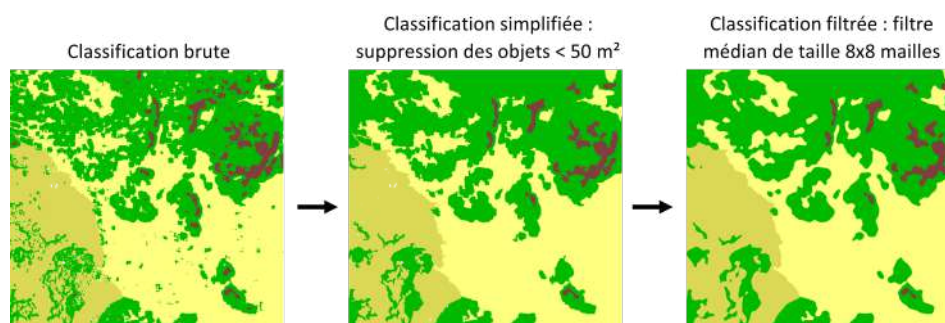


FIGURE 5.1 – Exemple de post-traitements sur une zone carrée de 300m de côté à maille de 50 cm. Les erreurs sont retirées, les trous remplis et les frontières entre objets lissées.

Enfin, selon les informations à représenter et les connaissances (expertises et vérités terrain), un opérateur peut être amené à fusionner ou à diviser des classes, voire à reprendre manuellement la classification et le contour de certains objets (Costa and Battista; 2013; Lu and Weng; 2007). En effet, l'introduction des algorithmes appliqués aux données est bornée aux informations qui s'y trouvent alors que l'injection d'autres informations physiques (température de l'eau, hydrodynamisme, saisonnalité, turbidité, etc.) ou d'expertise peut se faire manuellement et participe tout autant à la distinction des habitats benthiques (Brown et al.; 2011).

5.1.2.3 Évaluation

Sur le principe, évaluer les performances d'un modèle est simple, cela revient à comparer des résultats obtenus au moyen d'une méthode de calcul avec une vérité (précédente classification, vérités terrain). En pratique, il peut être assez difficile d'évaluer un résultat de classification pour diverses raisons. D'abord qualitativement, le rendu peut sembler bon en général et peu précis dans le détail ou bon pour certaines classes mais pas pour d'autres (Diesing et al.; 2016). Ensuite, la difficulté peut venir du fait qu'il n'y a pas de support de comparaison ou pas assez de connaissances réelles des fonds (manque de vérités terrain).

Quand on possède un support de comparaison, il existe de nombreux de critères pour des évaluations quantitatives. Ces critères permettent d'avoir rapidement une idée générale de la qualité des résultats au travers d'une valeur et de pouvoir déterminer si tel ou tel réglage d'une méthode est meilleur ou non. Le critère le plus majoritairement utilisé est l'exactitude qui est le rapport entre le nombre d'éléments bien classés par rapport au nombre total d'éléments, ce rapport est généralement donné sous forme d'un pourcentage. Meilleur est le résultat de classification, plus l'exactitude tend vers 100%.

Enfin, il est aussi intéressant d'estimer techniquement une méthode plutôt que ses résultats. On tiendra compte alors, par exemple, de la complexité algorithmique (temps de calcul et utilisation de mémoire), de la robustesse au bruit dans les données et de la sensibilité vis à vis du paramétrage ou de l'initialisation (DeFries and Chan; 2000).

5.1.3 Méthode Seaviews

La méthode de classification des fonds et des habitats marins développée et utilisée à Seaviews est décrite dans l'article Viala et al. (2021). Les traitements sont intégralement réalisés dans le logiciel maison de SIG, ViewMap et peuvent être appliqués aux données SMF et aux données bathymétriques issues de lidar aéroporté.

5.1.3.1 Données utilisées

Après préparation des données SMF R2Sonic, les données de bathymétrie et de réflectivité sont projetées dans des modèles numériques de terrain, généralement en projection Lambert 93, de même résolution. La bathymétrie est déclinée en deux autres grandeurs - pente et rugosité - également mises sous forme d'un MNT de projection et de résolution identique à la bathymétrie. La rugosité caractérise les irrégularités des fonds tandis que la pente caractérise l'inclinaison des fonds par rapport à l'horizontal.

L'indice de rugosité est calculé au niveau des sondes (sous forme de nuage de points) en s'intéressant au résidu métrique issu d'une régression linéaire par la méthode des moindres carrés. Pour toute sonde (non aberrante) du jeu de données, on effectue une régression linéaire en prenant en compte les sondes voisines : soit i) toutes les sondes dans un rayon d'un mètre soit ii) les dix sondes les plus proches. Il faut donc avoir une bonne densité de sondes et avoir bien éliminer les sondes aberrantes pour éviter qu'elles ne biaisent le calcul de rugosité. Ces deux choix de valeurs empiriques (rayon d'un mètre ou dix sondes) sont discutables : elles sont le reflet de la densité de sondes généralement acquises (acquisition avec recouvrement, profondeurs inférieures à 100m) et des résolutions de MNT typiquement utilisées (de 1m ou 50cm).

Soit un nuage de N points tels que $\{(x_1, y_1, z_1), \dots, (x_N, y_N, z_N)\}$ soient les coordonnées de ces points dans le système (X, Y, Z) . L'indice de rugosité de ce nuage se calcule par le résidu des moindres carrés d'une régression linéaire de Z sur X, Y . La fonction linéaire qui s'approche au mieux du nuage est $\hat{z}_n = A_0 + A_1x_n + A_2y_n, \forall n \in [1, N]$ où les paramètres (A_0, A_1, A_2) sont choisis pour minimiser la différence entre z_n et \hat{z}_n . Le résidu au carré s'exprime alors par :

$$Res^2 = \frac{1}{N} \sum_{n=1}^N (z_n - (A_0 + A_1x_n + A_2y_n))^2.$$

Les paramètres optimaux sont obtenus de la manière suivante :

$$\begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix} = \begin{pmatrix} N & S_x & S_y \\ S_x & S_{x^2} & S_{xy} \\ S_y & S_{xy} & S_{y^2} \end{pmatrix}^{-1} \begin{pmatrix} S_z \\ S_{xz} \\ S_{yz} \end{pmatrix} \quad \text{avec} \quad \begin{matrix} S_x = \sum_{n=1}^N x_n & S_y = \sum_{n=1}^N y_n & S_z = \sum_{n=1}^N z_n \\ S_{xy} = \sum_{n=1}^N x_n y_n & S_{xz} = \sum_{n=1}^N x_n z_n & S_{yz} = \sum_{n=1}^N y_n z_n \end{matrix}$$

La pente bathymétrique est obtenue par calcul au niveau des sondes contenues dans une maille du MNT bathymétrique interpolé. Dans un MNT interpolé, une maille n'a pas la même valeur au centre ou sur le bord pour assurer une continuité d'une maille à une autre. L'interpolation bilinéaire établit une fonction f sur les deux dimensions x et y de la grille du MNT pour connaître la valeur en tout point. Pour cela, en chaque maille, on cherche les quatre constantes (A_3, A_4, A_5, A_6) à partir des valeurs des quatre mailles voisines (système de quatre équations à quatre inconnues) :

$$f(x, y) = A_3 + A_4x + A_5y + A_6xy.$$

Une fois ces coefficients connus pour toutes les mailles, la pente se calcule en chaque maille de centre (x, y) comme étant la norme de la dérivée dans les deux directions :

$$\begin{aligned} \text{pente}(x, y) &= \sqrt{\left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2} \\ &= \sqrt{(A_4 + A_6y)^2 + (A_5 + A_6x)^2} \end{aligned}$$

Enfin, les données de vérités terrain sont, si possible, acquises après l'analyse des données acoustiques : les objets suspects en forme et valeur sur les données de réflectivité et de rugosité font l'objet d'une investigation à l'aide d'une caméra tractée, d'un appareil photo descendu par un bout lesté, voire de plongeurs. En zone peu profonde, l'utilisation des images aériennes les plus récemment acquises, telles que la BD Ortho (résolution 20 cm) et Ortholite V2 peuvent permettre l'identification de substrats.

5.1.3.2 Méthode manuelle de classification

Il faut dans un premier temps établir un inventaire des habitats présents à l'aide de l'ensemble des données - données de vérité terrain, données acoustiques (et/ou lidar) et images aériennes disponibles - mais aussi des connaissances du milieu. Cependant, seules les données sous forme de MNT - bathymétrie, pente, rugosité et réflectivité - participent à la classification, les autres données ne servent qu'à aider l'opérateur dans son interprétation et dans ses corrections.

Pour chaque classe de fond marin à représenter, l'opérateur doit cibler les valeurs des MNT qui décrivent au mieux les habitats. Ces valeurs sont définies par l'analyse visuelle des MNT. Ces descriptions et valeurs sont variables d'une zone d'étude à une autre. Assez communément sur les côtes de la mer Méditerranée, on peut trouver ces types de fonds que l'on peut décrire de cette manière :

- une faible rugosité pour des fonds de sédiments avec une très faible réflectivité pour des sédiments fins et homogènes et une réflectivité un peu plus élevée pour des sédiments plus grossiers ;
- une très forte rugosité pour les fonds rocheux ;
- des valeurs intermédiaires et assez variables de rugosité et de réflectivité correspondant à des herbiers sous-marins ou à des forêts d'algues.

Ensuite, les données de vérité terrain et les connaissances du milieu (descriptions des habitats EUNIS et géographie de la zone d'étude) entrent en jeu pour affiner le niveau de description des habitats, en reprenant les exemples de fonds précédents (voir également la figure 5.2) :

- Les espèces qui composent les herbiers doivent être repérées par les données de vérité terrain. En mer Méditerranée française, les herbiers sont majoritairement composés de *Posidonia oceanica* présents jusqu'à une profondeur variable de 30 à 40 m de profondeur (selon la clarté des eaux), au delà il peut s'agir de matte de *P. oceanica* (réseaux de racines) ou d'un autre type d'herbier.

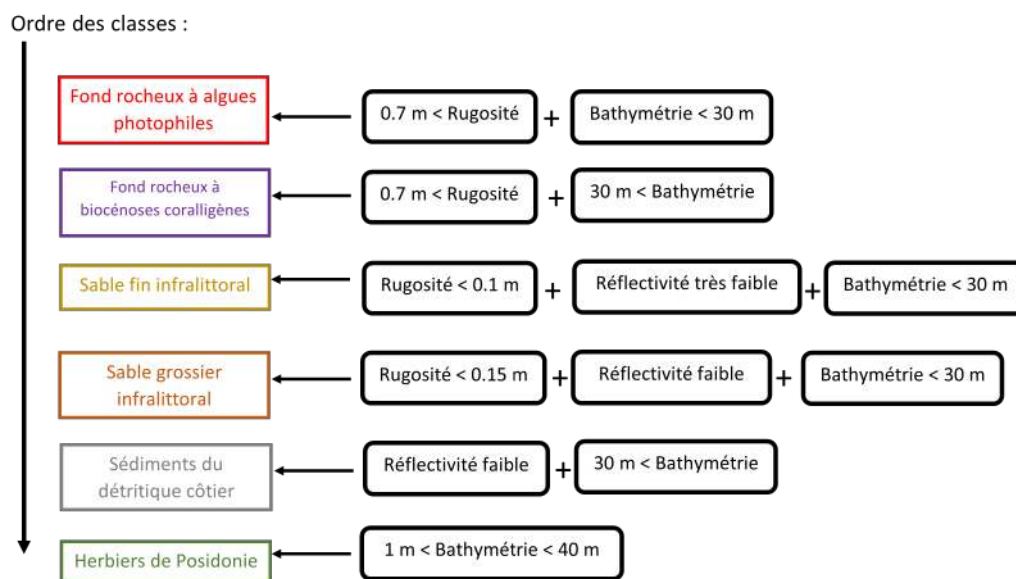


FIGURE 5.2 – Représentation d'un arbre de décision où chaque classe est définie par rapport à des valeurs issues des données de bathymétrie, de rugosité et de réflectivité ainsi qu'aux descriptions des systèmes de classification des habitats. Les valeurs sont données à titre d'exemple.

- Pour les fonds rocheux, il faut chercher à savoir s'ils sont soumis à l'action des vagues ou de la marée ou s'ils sont profonds (se référer au MNT de bathymétrie et aux marées et à la courantologie de la zone d'étude), puis chercher par quelles espèces marines ils sont colonisés (se référer aux données de vérité terrain).
- Les sédiments sont fragmentés en une multitude d'habitats différents qu'il faut caractériser plus finement à partir de i) leur composition (granulométrie, part de vase ou de cailloutis, débris de coquillage, présence d'algues et d'animaux) qui peut s'évaluer au travers de la réflectivité et de données de vérité terrain et ii) de leurs caractéristiques d'emplacement (profondeur, hydrodynamisme, lumière) qui s'évaluent à partir de la bathymétrie et de la géographie de la zone d'étude.

Une fois que l'utilisateur a caractérisé chacune des classes à partir des valeurs des MNT, il faut établir un ordre d'agencement des classes pour définir des priorités entre elles. L'ensemble peut être vu comme un arbre de décision que le logiciel applique à l'ensemble des mailles de la zone pour leur attribuer une classe. Dans l'exemple de la figure 5.2, il faut placer la classe des herbiers de Posidonie en dernier car ceux-ci sont uniquement décrits par la bathymétrie ; placés en priorité dans l'arbre, ils prendraient le pas sur quasiment toutes les autres classes.

En général, le premier arbre construit demande des corrections car les valeurs décrivant une classe ne sont pas identiques d'un bout à l'autre de la zone à traiter ou parce que des classes empiètent sur d'autres. Dans ce cas, l'arbre peut être remanié en réajustant les valeurs et les ordres des classes. La zone peut aussi être fragmentée en sous-zones pour appliquer différemment l'arbre. Les reprises peuvent également être manuelles car il faut représenter un habitat non visible sur les données acoustiques mais existant dans les données de vérité terrain (comme des herbiers très peu denses tel que des herbiers de Cymodocée). C'est un travail long et minutieux. Enfin, Les résultats sont simplifiés en supprimant les objets trop petits pour produire une carte finale des habitats.

La méthode est donc une interprétation des données par un utilisateur qui construit un arbre de décision. Ce travail devient long et complexe avec l'augmentation du nombre de classes et de la surface à traiter. Il est intéressant alors de s'orienter vers des méthodes de classification automatiques plus efficaces qu'un opérateur quand le problème se complique.

5.2 Méthodes de classification automatique

5.2.1 Méthode de partitionnement des K -moyennes

Les méthodes de partitionnement, dites non-supervisées, organisent les données en catégories “homogènes” au regard des caractéristiques qui servent à leur analyse, selon des critères de similarités (distance) entre individus. C’est ensuite à l’opérateur de déduire un sens physique pour chaque catégorie. Plusieurs grandes familles de méthodes permettent cette organisation non-supervisée des données dont les centres mobiles (K -moyennes) (Müller and Guido; 2018), les regroupements hiérarchiques (Müller and Guido; 2018), les approches par distribution de probabilité (Sharma et al.; 2019) ou encore les approches par densité (Müller and Guido; 2018). Nous avons choisi de ne présenter que la méthode des centres mobiles, et en particulier celle des K -moyennes, pour des raisons d’efficacité sur de grands jeux de données (expérimentales) ainsi que pour sa facilité de réglage (Tableau 5.2).

TABLE 5.2 – Principaux avantages et inconvénients des méthodes de partitionnement les plus connues.

Famille de méthodes	Avantages	Inconvénients
Centres mobiles	Très répandue et simple à comprendre, Peu de paramètres et efficace sur des grands jeux de données réelles	Faiblesse vis-à-vis des classes ayant des structures complexes
Regroupements Hiérarchiques	Simple à comprendre, Peu de paramètres	Temps de calcul longs, Gourmand en mémoire
Approches par densité de données	Corrige une partie des faiblesses des centres mobiles	Réglage des paramètres assez délicats
Approches par inférence des lois de probabilités	Efficace sur des données dont la loi de distribution est connue	Peu efficace sur des données expérimentales sans loi de distribution

5.2.1.1 Fonctionnement des K -moyennes

La méthode des K -moyennes (K -means) est une méthode de partitionnement des données qui répond à un problème d’optimisation combinatoire : obtenir la meilleure partition en minimisant les écarts entre individus d’une même classe et à maximiser les écarts entre les différentes catégories. Le principe se décompose en quatre étapes (voir aussi figure 5.3) :

- Initialisation : Choix de K individus de départ dans le jeu de données de N individus ($N \gg K$). Ces K individus forment les K valeurs moyennes initiales et donc les K classes initiales.
- Assignment : Les $N - K$ individus restants sont associés à l’une des K classes, en l’occurrence celle la plus proche au sens d’une distance de similarité.
- Mise à jour : La moyenne pour chaque classe est calculée à partir des individus nouvellement associés. Ces K moyennes remplacent les moyennes précédemment calculées.
- Répétition des étapes d’assignment et de mise à jour jusqu’à ce qu’il n’y ait plus de mouvements d’individus d’une classe à une autre (critère d’arrêt).

Pourtant simple, cet algorithme laisse déjà plusieurs choix quant à la manière de fixer la valeur de K , à l’initialisation des K classes, à la mesure de distance pour associer les individus aux classes et au critère d’arrêt.

Avant l’initialisation, il faut fixer la valeur pour K correspondant au nombre de classes à réaliser. La valeur de K peut être fixée d’avance par un opérateur car l’on sait *a priori* combien de catégories sont à obtenir mais dans un problème plus complexe, la valeur de K peut faire partie des inconnues à estimer en laissant alors l’algorithme optimiser au mieux les partitionnements. Dans ce

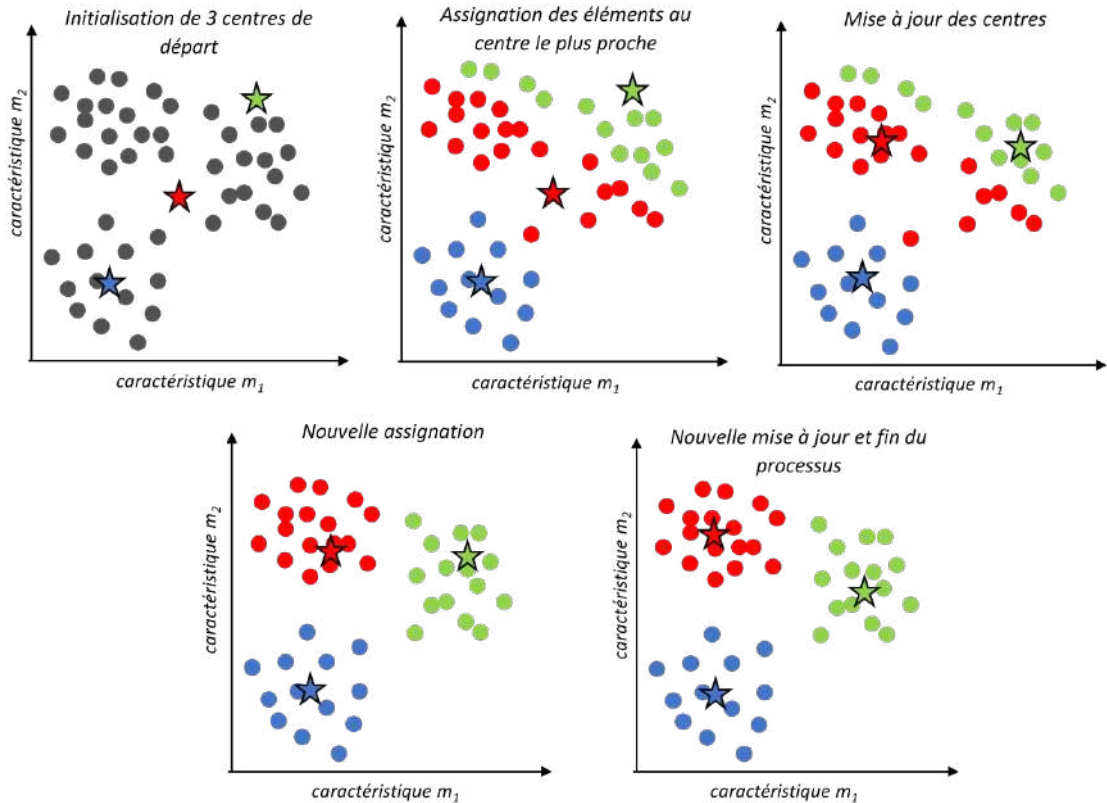


FIGURE 5.3 – De gauche à droite, les différentes étapes pour réaliser une partition de données par la méthode des K -moyennes. Les centres sont représentés par des étoiles dont l’initialisation est prise au hasard.

cas, l’algorithme est relancé plusieurs fois avec différentes valeurs de K , la meilleure optimisation permettant alors d’estimer la valeur de K (Thorndike; 1953).

Ensuite, pour le choix des individus au départ, plusieurs possibilités sont envisageables. Le tirage aléatoire des individus est le plus simple mais cela peut entraîner une convergence plus longue (un grand nombre d’itérations) ou à l’inverse, ils peuvent être fixés par un opérateur pour réduire le nombre d’itérations. Il existe également des algorithmes d’initialisation comme K -moyennes++ (Arthur and Vassilvitskii; 2007).

L’assignation d’un élément (une donnée) \mathbf{d} , à M caractéristiques notées (d_1, \dots, d_M) , à l’une des K classes s’effectue par minimisation de la distance entre cet élément et les centres des classes $\boldsymbol{\mu}_k, \forall k \in [1, K]$ correspondant à un vecteur de M moyennes. Plusieurs définitions de distance, au sens d’une norme, \mathcal{D} existent et conviennent, par exemple la distance de Manhattan (utilisant la norme 1),

$$\mathcal{D}_{\text{Manhattan}} = \|\mathbf{d} - \boldsymbol{\mu}_k\|_1 = \sum_{m=1}^M |d_m - \mu_{k,m}|$$

ou encore la distance Euclidienne (utilisant la norme 2),

$$\mathcal{D}_{\text{Euclidienne}} = \|\mathbf{d} - \boldsymbol{\mu}_k\|_2 = \sqrt{\sum_{m=1}^M |d_m - \mu_{k,m}|^2}.$$

Enfin, le critère d’arrêt, “jusqu’à ce qu’il n’y ait plus de mouvements d’individus d’une classe à une autre”, n’est que peu applicable car il existe un risque de boucle infinie avec quelques individus

oscillant d'une classe à une autre selon une légère variation des moyennes de chaque classe. Un critère plus quantitatif est la minimisation de l'inertie intra-classe des K classes :

$$\min Inertie_{intra-classe} = \arg \min_{\Pi} \sum_{k=1}^K \sum_{\mathbf{d} \in \Pi_k} \mathcal{D}(\mathbf{d}, \boldsymbol{\mu}_k)$$

où l'ensemble des éléments sont assignés à des partitions Π_k telles que,

$$\Pi_k = \{\mathbf{d} : \mathcal{D}(\mathbf{d}, \boldsymbol{\mu}_k) \leq \mathcal{D}(\mathbf{d}, \boldsymbol{\mu}_{k^*}), \forall k^* = 1, \dots, K\}.$$

Dans les **généralités**, nous avons énoncé que l'optimum est atteint lorsque les écarts entre individus d'une même classe étaient minimisés (inertie intra-classe) ET que les écarts entre classes étaient maximisés (inertie inter-classes). Cependant, le critère de l'inertie intra-classe suffit car les écarts entre tous les individus (inertie globale, constante au cours des itérations) n'est autre que la somme des inerties intra et inter classe; maximiser l'une revient à minimiser l'autre, et inversement :

$$Inertie_{globale} = Inertie_{intra-classe} + Inertie_{inter-classe} = constante \quad \forall \text{ iteration}$$

Enfin, ce critère peut être complété par un nombre d'itérations maximum pour éviter de boucler à l'infini.

La faiblesse principale des K -moyennes (voir tableau 5.2) réside dans son manque d'efficacité sur des structures complexes de données à partitionner comme des classes d'objets non linéairement séparables, des classes à différentes densités d'éléments ou avec des éléments principalement dispersés selon une dimension (Müller and Guido; 2018). Le jeu de données des *deux lunes* est un exemple synthétique à deux classes non linéairement séparables et avec une orientation privilégiée des données. Les K -moyennes avec une valeur de $K = 2$ ne permet pas d'obtenir un partitionnement convenable (figure 5.4 gauche). Cependant, augmenter le nombre de classes jusqu'à $K = 9$ dans cette situation, permet d'avoir des classes sans mélange de données des lunes (figure 5.4 droite). Ainsi, lors d'un post-traitement, on peut regrouper les classes dans deux ensembles différenciant bien les deux lunes.

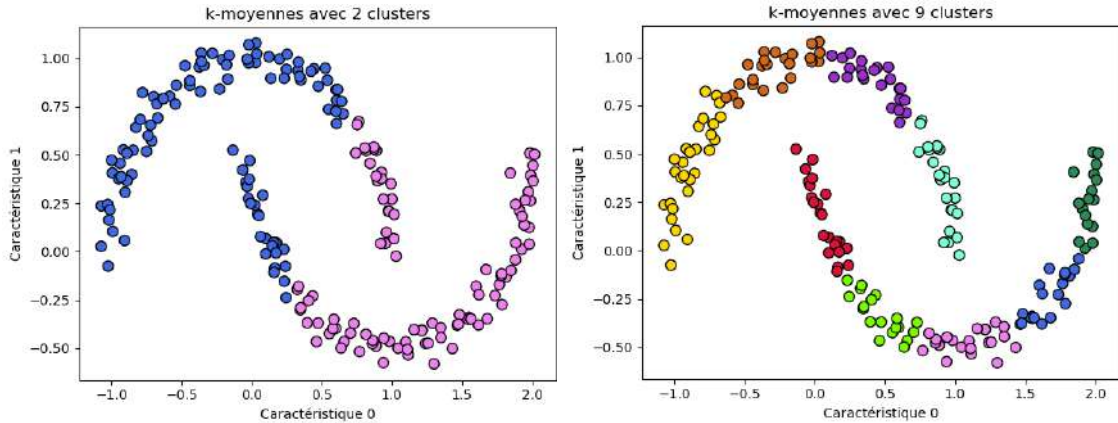


FIGURE 5.4 – A gauche, deux lunes mal partitionnées par une méthode des K -moyennes à deux classes. A droite, les deux lunes sont découpées en plusieurs classes bien distinctes.

5.2.1.2 Implémentation de la méthode dans ViewMap

Puisque très accessible d'un point de vue programmation, l'algorithme a été implémenté dans le logiciel ViewMap (en C++) pour travailler directement sur le format de données des MNT avec autant de caractéristiques que désirées.

Une Interface Homme Machine laisse à l'opérateur la liberté de fixer :

- la valeur de K ,
- le type d'initialisation, soit aléatoire soit en plaçant des points sur les MNT,
- les caractéristiques à prendre en compte pour le calcul ainsi qu'une éventuelle pondération pour donner plus de poids à certaines caractéristiques.

Les données manipulées par la méthode ont des amplitudes et des plages de valeurs très différentes ; des entiers de 0 à 255 par pas de 1 pour les images aériennes tandis que la pente est représentée par des nombres flottants (précision décimale de 6) s'étendant de 0 à 1m environ. Ces disparités faussent l'attribution des individus aux classes en donnant plus de crédit aux caractéristiques ayant de fortes valeurs. C'est pourquoi, il est préférable de centrer-normaliser toutes les caractéristiques. De même, les individus ayant des valeurs manquantes (mises à zéro ou à Not-a-Number) ne sont pas à prendre en compte pour éviter de biaiser les attributions.

A la fin du calcul, un affichage précise le nombre d'itérations qui a été nécessaire ainsi que l'inertie intra-classe minimale obtenue. Cela permet à l'opérateur de comparer l'efficacité de la méthode en faisant varier les réglages (nombre de classes, initialisation, sélection des caractéristiques).

5.2.2 Classement automatique

Les méthodes reposant sur un apprentissage supervisé sont intégrées dans ViewMap au travers de l'utilisation de la librairie de vision par ordinateur, OpenCV (version 4.5.2), disponible en langage C++. Les méthodes de Séparateurs à Vastes Marges ou Machine à Vecteurs de Support (SVM), de Forêts Aléatoires (FA) d'Arbres Décisionnels et de classifieurs bayésiens sont trois méthodes au fonctionnement très différent auxquelles nous avons choisi de nous intéresser. Elles sont présentées en détails dans les sous-sections ci-dessous.

5.2.2.1 Séparateurs à Vastes Marges

Les SVM (en anglais, *Support Vector Machine*) sont utilisés pour la résolution de problèmes de régression et de classification. Ce sont des techniques qui généralisent les classifieurs linéaires en reposant sur la notion de marge maximale. Une des versions les plus simples consiste en un classifieur binaire (problème à deux classes) et linéaire à hyperplan séparateur, c'est-à-dire qu'elle établit une fonction f qui à un élément \mathbf{d} composé de M caractéristiques fait correspondre la prédiction o (*output*) telle que

$$o = f(\mathbf{d}) = \text{sgn}(\mathbf{w}^T \cdot \mathbf{d} + b)$$

où \mathbf{w} est un vecteur à M poids, b un biais scalaire et sgn est la fonction signe qui renvoie $\{-1, 1\}$ selon le signe de $\mathbf{w}^T \cdot \mathbf{d} + b$. La prédiction o prend alors l'une des deux valeurs

$$\begin{cases} \mathbf{w}^T \cdot \mathbf{d} + b \geq 0 & \Rightarrow o = 1 \\ \mathbf{w}^T \cdot \mathbf{d} + b < 0 & \Rightarrow o = -1. \end{cases}$$

En particulier, $\mathbf{w}^T \cdot \mathbf{d} + b = 0$ est un hyperplan séparateur entre les deux classes de données, les données au-dessus sont étiquetées à 1 tandis que celles au-dessous sont à -1. L'entraînement de ce type de classifieur passe par la recherche des valeurs de \mathbf{w} et b qui permettent la discrimination des deux classes, cela correspond à chercher les hyperplans séparateurs. Souvent, de nombreuses solutions sont possibles comme c'est le cas sur la figure 5.5 gauche.

Les SVM diffèrent des autres classifieurs linéaires dans le fait que l'hyperplan doit satisfaire une marge maximale entre l'hyperplan et les éléments les plus proches (figure 5.5 droite). Ces derniers sont appelés les *vecteurs supports* car ils conditionnent le choix du séparateur, les autres éléments des classes étant sans influence. Cette contrainte de marge maximale est justifiée par la théorie de Vapnik-Chervonenkis (Hastie et al.; 2001).

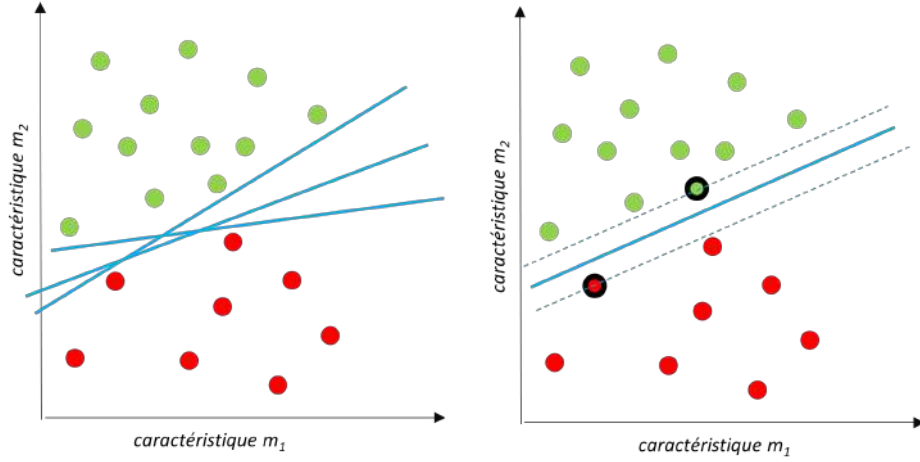


FIGURE 5.5 – A gauche, différents séparateurs linéaires (en bleu) discriminant les éléments rouge et vert. A droite, le séparateur maximisant la marge entre les deux classes, les deux éléments cerclés de noir sont les vecteurs supports du séparateur.

Cette marge entre un élément \mathbf{d} et l'hyperplan \mathcal{H}_p est une distance \mathcal{D} qui s'exprime par la projection orthogonale de l'élément sur l'hyperplan de vecteur normal \mathbf{w} (le vecteur de poids) (Hastie et al.; 2001) :

$$\mathcal{D}(\mathbf{d}, \mathcal{H}_p) = \frac{|\mathbf{w}^T \cdot \mathbf{d} + b|}{\|\mathbf{w}\|_2}.$$

Si l'on considère un jeu d'apprentissage composé de N paires d'éléments $(\mathbf{d}^{(n)}, e_n), \forall n \in [1, N]$ les données, $\mathbf{d}^{(n)} \in \mathbb{R}^M$ et les étiquettes associées, $e_n \in \{-1, 1\}$, alors il existe un unique hyperplan optimal tel que :

$$\arg \max_{\mathbf{w}, b} \left\{ \frac{1}{\|\mathbf{w}\|} \min_n (|\mathbf{w}^T \cdot \mathbf{d}^{(n)} + b|) \right\} \text{ sous la contrainte que } \mathbf{w} \text{ soit non nul.}$$

Ce problème peut se résoudre à l'aide des multiplicateurs de Lagrange en introduisant des contraintes. Une contrainte simple peut être de dire que les vecteurs supports ont une marge de 1 à \mathcal{H}_p . Ce qui se traduit pour l'un des vecteurs supports de chaque classe 1 et -1, notés respectivement $(\mathbf{d}^-, e^- = -1)$ et $(\mathbf{d}^+, e^+ = 1)$ par :

$$\begin{cases} \mathbf{w}^T \cdot \mathbf{d}^- + b = 1 & \Rightarrow e^- (\mathbf{w}^T \cdot \mathbf{d}^- + b) = 1 \\ \mathbf{w}^T \cdot \mathbf{d}^+ + b = -1 & \Rightarrow e^+ (\mathbf{w}^T \cdot \mathbf{d}^+ + b) = 1 \end{cases} \quad (5.1)$$

D'autre part, cette contrainte implique que pour toutes les paires d'éléments, la distance à \mathcal{H}_p est d'au moins 1 : $e_n (\mathbf{w}^T \cdot \mathbf{d}_n + b) \geq 1, \forall n \in [1, N]$. On introduit alors le lagrangien \mathcal{L} tel que (Hastie et al.; 2001)

$$\mathcal{L}(\mathbf{w}, b, \nu_1, \nu_2, \dots, \nu_N) = \frac{1}{\|\mathbf{w}\|} - \sum_{n=1}^N \left(\nu_n * (e_n (\mathbf{w}^T \cdot \mathbf{d}^{(n)} + b) - 1) \right), \text{ avec les } \nu_n \in \mathbb{R}$$

sous les contraintes de minimiser \mathbf{w} et b et de maximiser les ν_n .

Puis, on calcule les différentielles de \mathcal{L} par rapport à ses $N + 2$ variables (\mathbf{w}, b et les N valeurs de ν_n) et l'on cherche quand elles s'annulent. Ces résolutions amènent à trouver l'hyperplan défini par (\mathbf{w}, b) .

Les équations décrivent un cas facilement séparable linéairement. Bien souvent, pour que le problème soit solvable par une frontière linéaire, il faut accepter que des éléments soient mal

classés par rapport à leur étiquette e_i . Cette adaptation des SVM est appelée marge souple (Hastie et al.; 2001). Pour cela, on introduit pour chaque élément, une distance de “laxisme” par rapport à la marge ou des variables ressorts, $\xi_n \geq 0, \forall n \in [1, N]$ telle que l'équation 5.1 devienne :

$$e_n(\mathbf{w}^T \cdot \mathbf{d}^{(n)} + b) \geq 1 - \xi_n, \forall n \in [1, N].$$

Pour pénaliser ce “laxisme”, une variable de régularisation ε est ajoutée et le problème de recherche de la marge devient alors :

$$\text{Minimiser } \frac{1}{\|\mathbf{w}\|} + \varepsilon \sum_{n=1}^N \xi_n \text{ sous les contraintes } e_n(\mathbf{w}^T \cdot \mathbf{d}^{(n)} + b) \geq 1 - \xi_n.$$

Cependant, un SVM linéaire n'est pas performant sur un problème avec des classes non-linéairement séparables même avec une marge souple (figure 5.6 gauche). Si ce problème ne l'est pas dans l'espace des données à N dimensions, il l'est peut être dans un espace à $N + 1$ ou $N + 2$ (voire davantage) dimensions. La dimension peut être “artificiellement” augmentée en prenant, par exemple, le carré ou le cube de certaines caractéristiques ou encore le produit de caractéristiques. Toutefois, augmenter la dimension est très lourd en terme de temps de calcul quand on ne sait pas si les nouvelles caractéristiques vont aider ou non à la discrimination. Alors, plutôt que d'utiliser une fonction de redescription d'espace sur un élément \mathbf{d} , on emploie plutôt l'*astuce du noyau* qui consiste à calculer des distances (sous forme de produits scalaires) entre des éléments qui ont le plus d'influence sur le choix de l'hyperplan (les éléments vecteurs supports). Il existe plusieurs noyaux proposés dans la librairie OpenCV (voir le tableau 5.3). Cependant c'est le noyau à base radiale (*Radial Basis Function*) qui est le plus populaire pour des classes non-linéairement séparables (figure 5.6 droite). Il est même présenté comme un bon premier choix dans le guide de la librairie LIBSVM (Chang and Lin; 2011), incluse dans OpenCV.

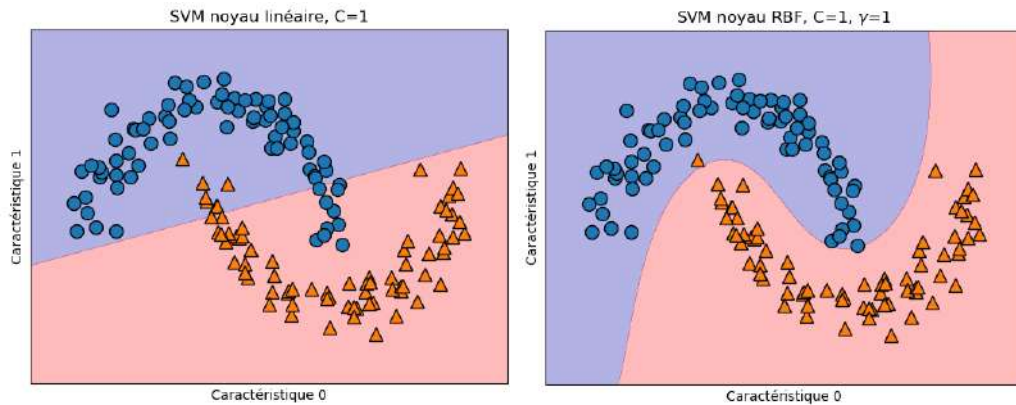


FIGURE 5.6 – A gauche, un SVM linéaire échoue à traiter deux lunes non linéairement séparables. A droite, les deux lunes sont partagées par une frontière lisse quitte à laisser quelques éléments du mauvais côté.

TABLE 5.3 – Exemple de noyaux utilisés par les SVM

Nom du noyau	Fonction	Paramètres
Noyau linéaire	$(\mathbf{d}^\circ, \mathbf{d}^\bullet) \mapsto \mathbf{d}^{\circ T} \cdot \mathbf{d}^\bullet$	$\mathbf{d}^\circ, \mathbf{d}^\bullet$ sont deux éléments
Noyau à base radiale	$(\mathbf{d}^\circ, \mathbf{d}^\bullet) \mapsto \exp(-\gamma \mathbf{d}^\circ - \mathbf{d}^\bullet ^2)$	largeur du noyau $\gamma > 0$
Noyau sigmoïde	$(\mathbf{d}^\circ, \mathbf{d}^\bullet) \mapsto \tanh(\gamma \mathbf{d}^{\circ T} \cdot \mathbf{d}^\bullet + b)$	un biais $b \in \mathbb{R}$
Noyau polynomial	$(\mathbf{d}^\circ, \mathbf{d}^\bullet) \mapsto (\gamma \mathbf{d}^{\circ T} \cdot \mathbf{d}^\bullet + b)^\diamond$	degré du polynôme $\diamond \in \mathbb{N}$

Enfin, les SVM sont aussi capables de traiter des problèmes multi-classes (nombre de classes supérieur à 2) en ramenant le problème global à ensemble de problèmes binaires en faisant soit du un-contre-un soit du un-contre-tous. L'approche un-contre-tous consiste à chercher N séparateurs pour N en confrontant une par une les classes au reste des autres classes : pour $N = 4$, cela fait quatre problèmes binaires. C'est l'approche utilisée dans OpenCV pour les SVM. L'autre approche, un-contre-un est moins répandue et cherche un plus grand nombre de séparateurs, $(N - 1)!$ pour N classes, en confrontant une classe à une autre.

Implémentation dans ViewMap Outre le choix du noyau (qui peut se résumer à prendre par défaut le noyau à base radiale) et le réglage des paramètres du noyau, il y a aussi le réglage du paramètre de régularisation ε dans le SVM proposé par OpenCV. Cette régularisation fixe le compromis entre un modèle simple (peu précis) et un modèle où tous les éléments d'apprentissage sont bien classés au risque de trop complexifier le modèle et de le rendre non-généralisable sur de nouvelles données. En pratique cela se traduit par la part d'éléments d'apprentissage qui seront mal classés exprès pour choisir un hyperplan plus optimal. Cela permet de construire un modèle plus simple en laissant de l'autre côté de la frontière les éléments aux valeurs aberrantes. Plus la régulation est choisie grande et plus l'importance donnée aux éléments est forte pour que tous soient correctement classés. Au contraire, un ε plus petit donne plus de souplesse au modèle pour qu'il s'ajuste à une majorité d'éléments et limite la complexité. Le paramètre γ du noyau à base radiale est l'inverse de la largeur du noyau qui se traduit par la portée de l'influence de chaque élément. Un γ petit engendre un grand rayon de portée et produit un modèle plus simple aux frontières plus lisses, tandis qu'avec une plus grande valeur, la portée est réduite et les frontières se centrent autour des éléments.

Les résultats des SVM restent très sensibles aux réglages des paramètres qui ne peuvent pas être fixés identiquement pour tous les problèmes. Leurs réglages sont donc faits automatiquement à travers d'une grille de recherche proposée par la librairie.

Comme pour la méthode des K -moyennes, les SVM demandent des données normalisées aux valeurs comprises entre 0 et 1 pour éviter que les caractéristiques à grandes valeurs n'écrasent celles à plus faibles valeurs. Une standardisation est donc préalablement appliquée à tout le jeu de données.

5.2.2.2 Forêts aléatoires d'arbres décisionnels

Les Arbres de Décisions (AD) sont un outil d'aide à la prise de décision représentant les choix possibles (ici les classes) sous la forme d'un arbre. Ce type d'arbre se construit sous la forme d'une séquence ordonnée de questions amenant à une décision. Chaque question est posée sur l'une des caractéristiques des données et forme un noeud dans l'arbre, les réponses coupent le jeu de données en deux parties. La caractéristique retenue pour un noeud donné est celle qui sépare au mieux les données de différentes classes. Il y a répétition de questions-réponses jusqu'à ce que les dernières réponses soient composées de données de la même classe et forme une feuille au bout de la ramification. Un exemple est proposé en figure 5.7 où un arbre est construit d'après les classes et les contraintes de la figure 5.2.

Lors d'un apprentissage supervisé, les arbres deviennent souvent très profonds et complexes, et peuvent conduire à du surapprentissage : le modèle est trop ajusté aux données d'apprentissage et n'est plus capable de généralisation sur de nouvelles données. Le paramétrage des AD va permettre de limiter le surapprentissage en fixant, entre autre, la profondeur des arbres ou bien la quantité minimale de données non partageables dans une feuille (technique dite d'élagage d'arbre) (Müller and Guido; 2018). Lors de la construction d'un arbre, est mesurée l'importance d'une caractéristique qui reflète le fait qu'elle contienne des informations aidant à la décision.

Les Forêts Aléatoires (FA) d'Arbres Décisionnels ou, en anglais, *random forest classifier*, font parties des méthodes d'apprentissage de type arbre de décisions. Ces forêts construisent de multiples arbres de décision à partir de jeux d'apprentissage à chaque fois différents et constitués

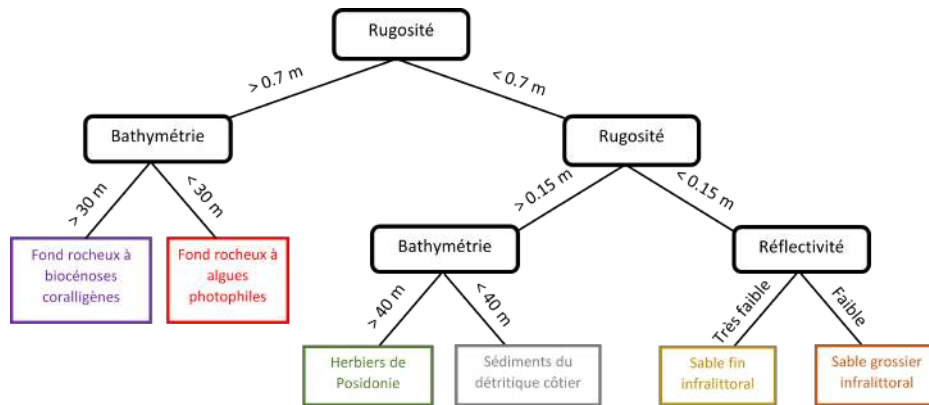


FIGURE 5.7 – Exemple d’arbre de décision construit à partir des descriptions des classes de la figure 5.2

aléatoirement (figure 5.8). L'utilisateur fournit un jeu de données de N éléments avec M caractéristiques. Pour chaque arbre à créer, un jeu de N éléments est constitué en tirant aléatoirement les éléments avec remise ; ce jeu d'apprentissage possède la même taille N que celui fournit par l'utilisateur mais avec des éléments redondants et absents. Ensuite, lors de la construction de l'arbre, à chaque noeud, sont d'abord sélectionnés au hasard un sous-ensemble de $M_1 < M$ caractéristiques avant de garder la plus discriminante. Une fois qu'un grand nombre d'arbres ont été réalisés, ils sont "moyennés" pour former un arbre général, cela noie les erreurs commises par chacun et évite le surapprentissage propre à tout arbre simple (Tufféry; 2019). Les FA sont donc plus facilement utilisables que les AD. Un exemple reprenant la configuration des deux lunes est proposé en figure 5.9.

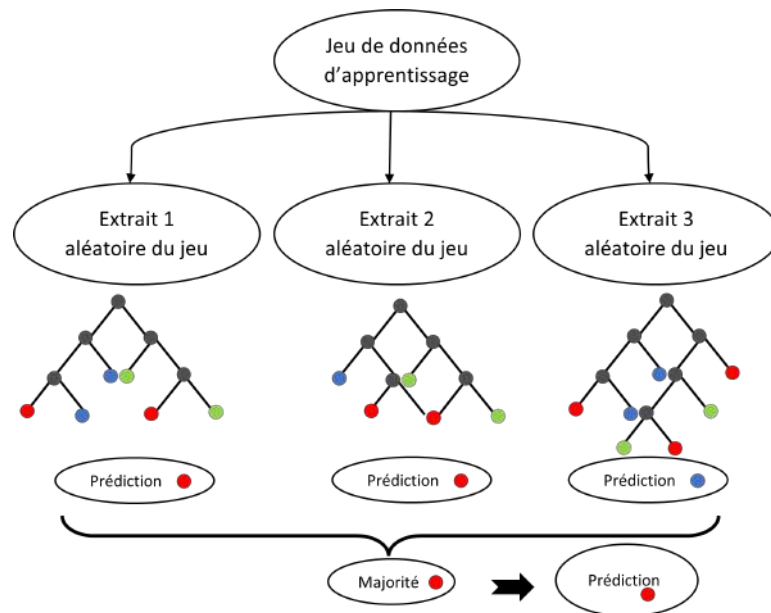


FIGURE 5.8 – Exemple du fonctionnement d’une forêt d’arbres décisionnels : plusieurs sous-jeux sont extraits d’un jeu de données d’apprentissage et participent à la construction de plusieurs modèles de décisions.

Implémentation dans ViewMap Les FA ont plus de paramètres à régler que les SVM : le nombre d’arbres à réaliser, le nombre M_1 de caractéristiques à sélectionner à chaque noeud et

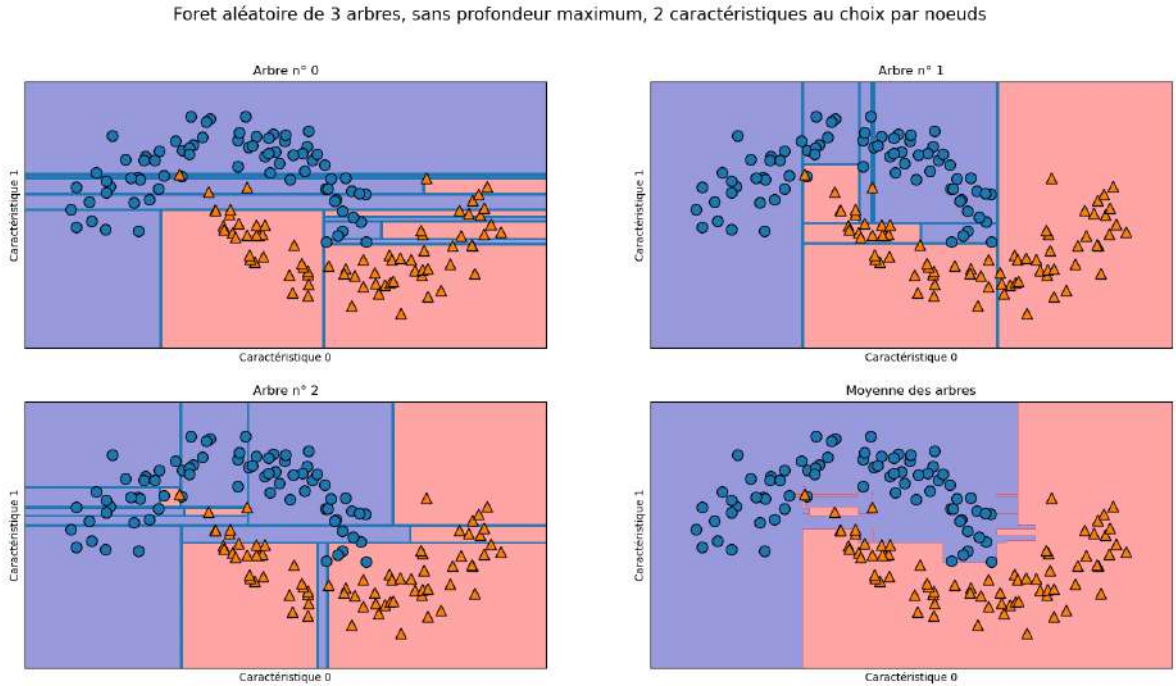


FIGURE 5.9 – Trois arbres sont créés avec des variantes du jeu d’apprentissage, leurs résultats sont différents et comportent quelques signes de surapprentissage (séparation pour un unique élément). Leur moyenne produit un résultat où le surapprentissage est réduit.

les paramètres des AD tels que la profondeur maximum de chaque arbre ou le nombre minimum d’éléments par feuille. Cependant, leur réglage n’est pas aussi sensible que pour les SVM. Un grand nombre d’arbres augmente le temps de calcul mais également la robustesse des résultats du fait de l’aspect aléatoire des différents tirages, ce nombre est à fixer aussi grand que possible dans la limite des capacités de l’ordinateur et de la patience de l’utilisateur (Müller and Guido; 2018). Le nombre M_1 peut être plus délicat à choisir : un trop petit M_1 ne laisse que peu de choix pour choisir la caractéristique la plus discriminante et peut faire que l’arbre devienne profond avant d’être fini, un M_1 trop proche de M fait que les arbres seront trop ressemblants. Une valeur de compromis valable pour toute configuration peut être de $M_1 = \sqrt{M}$ (Tufféry; 2019) qui est le choix fait au niveau d’OpenCV.

A la différence des K -moyennes et des SVM, les arbres et donc les FA n’ont pas besoin que les données soient recalibrées ou standardisées. Ils continuent de bien fonctionner avec des caractéristiques ayant des échelles de valeurs différentes.

5.2.2.3 Classifieurs bayésiens

Ce type de classifieurs repose sur le théorème de Bayes (de Thomas Bayes, mathématicien du 18^e siècle). Ce théorème décrit la probabilité d’un évènement à partir d’autres évènements déjà réalisés. Soient deux évènements, U et V , la probabilité de U sachant V s’exprime par :

$$P(U|V) = \frac{P(V|U) P(U)}{P(V)} \text{ ou encore } P(U|V) = \frac{P(U \wedge V)}{P(V)} \text{ avec } P(V) \neq 0.$$

A partir de ce théorème, il est possible de construire un modèle statistique de prédiction des classes en fonction des caractéristiques, également capable de préciser le taux de confiance de la prédiction (figure 5.10). Pour chaque élément du jeu d’apprentissage, leurs caractéristiques sont analysées individuellement pour la construction du modèle. Pour un élément \mathbf{d} ayant les M caractéristiques $\mathbf{d} : (d_1, \dots, d_M)$ que l’on cherche à classer parmi les K classes (Π_1, \dots, Π_K) , la

probabilité d'appartenir à la classe Π_k pour cet individu est donné par (Fleury et al.; 2021) :

$$P(\Pi_k \wedge d_1 \wedge \dots \wedge d_M) = P(d_M | \Pi_k \wedge d_1 \wedge \dots \wedge d_{M-1}) \cdot P(d_{M-1} | \Pi_k \wedge d_1 \wedge \dots \wedge d_{M-2}) \dots P(d_2 | \Pi_k \wedge d_1) \cdot P(d_1 | \Pi_k) \cdot P(\Pi_k). \quad (5.2)$$

Les probabilités calculées sont ensuite organisées par un ensemble de règles de décision représentées sous forme de graphe orienté acyclique où chaque noeud correspond à une caractéristique, on parle alors de **réseaux bayésiens**. Le fait que les graphes soient acycliques signifie qu'il n'y a pas de boucle, les graphes dessinent plutôt une hiérarchie entre les noeuds. Les relations entre chaque noeud sont définies avec des arcs orientés avec des notions de parents-enfants, de couples, d'ancêtres-descendants entre les différentes caractéristiques (Fleury et al.; 2021).

Les classifieurs bayésiens reposent sur une hypothèse forte permettant de simplifier l'équation précédente où toutes les probabilités impliquées dans le produit ne sont généralement pas connues.

Par exemple, le classifieur que l'on rencontre le plus souvent dans les librairies d'apprentissage automatique est le **classifieur bayésien naïf**. Son côté "naïf" est dû à l'hypothèse sur laquelle il fonde la construction d'un modèle : l'indépendance mutuelle entre toutes les caractéristiques du jeu de données. Cette hypothèse simplifie alors l'équation car tous les d_i sont indépendants entre eux et l'équation 5.2 devient alors (Sharma et al.; 2019) :

$$P(\Pi_k \wedge d_1 \wedge \dots \wedge d_M) = P(d_M | \Pi_k) \cdot P(d_{M-1} | \Pi_k) \dots P(d_2 | \Pi_k) \cdot P(d_1 | \Pi_k) \cdot P(\Pi_k).$$

Les probabilités se calculent plus simplement car chacun des éléments devient connu. L'équation est calculée pour chacune des classes et est ensuite combinée avec un ensemble de règles de décision pour sélectionner la classe qui a la probabilité la plus élevée. D'un point de vue pratique, ce classifieur naïf est performant même quand l'hypothèse d'indépendance n'est que peu vérifiée (Müller and Guido; 2018).

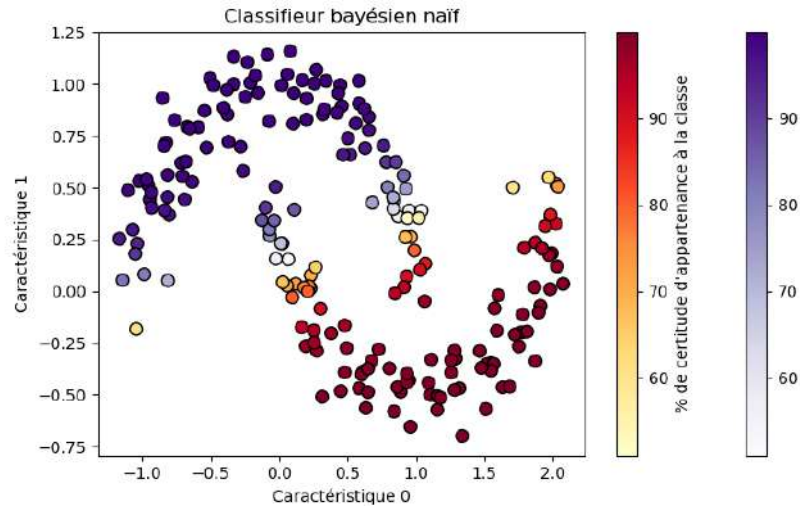


FIGURE 5.10 – Résultats de classification avec le Bayésien naïf. Le résultat s'approche de ceux du SVM linéaire (figure 5.6 gauche) et du K -moyenne à deux partitions (figure 5.4 gauche). Cependant les éléments mal classés ont un taux de confiance sur la prédiction plus faible que les autres.

La librairie OpenCV intègre un **classifieur bayésien normal** n'ayant pas besoin de caractéristiques mutuellement indépendantes. Par contre, son hypothèse se fonde sur une distribution gaussienne (ou distribution normale) des caractéristiques, telle que (Hastie et al.; 2001) :

$$P(d_1 \wedge \dots \wedge d_M | \Pi_k) = \frac{1}{\sqrt{(2\pi)^M |\Sigma_k|}} \exp \left(-\frac{1}{2} (d - \mu_k)^T \sum_K^{-1} (d - \mu_k) \right)$$

où il faut calculer à part les matrices de covariance Σ_k et leur déterminant $|\Sigma_k|$, pour les K classes, entre les M caractéristiques, et les vecteurs moyennes μ_k pour chaque classe. Cette hypothèse sur la distribution est moins forte que celle de l'indépendance des variables. Il n'y a pas de paramètres à régler pour la construction d'un modèle avec le classifieur bayésien normal d'OpenCV.

Dans le chapitre suivant, sont mises en œuvre ces différentes méthodes sur un jeu de données réelles. Nous conservons les étapes 1 et 2 de Seaviews, qui sont respectivement la préparation des données et l'extraction des caractéristiques. La phase de résultats se concentre sur l'application de ces méthodes - K -moyennes, SVM, FAD et classifieur Bayésien normal - en gardant en tête que ces résultats doivent être reproductibles par le personnel de Seaviews, familiarisé avec les logiciels mais pas avec ces nouvelles méthodes et leurs réglages. Le post-traitement ne sera que peu employé car il s'agit de garder les résultats les plus bruts possibles afin d'établir une comparaison objective entre les méthodes. L'analyse des performances sera réalisée au niveau de la discussion.

Chapitre 6

Étude biocénotique à Centuri

6.1 Zone d'étude et acquisition des données

Le site d'étude de 2.5 km² borde le littoral autour du Port de Centuri, en Corse. Il correspond au cadre rouge sur la fig. 6.1. L'îlot de Capense constitue un abri contre les vents d'ouest pour de nombreux navires de plaisance qui mouillent dans un herbier dense de posidonies. Afin de réaliser des études environnementales approfondies, la première étape consiste en la réalisation d'une carte des habitats benthiques.

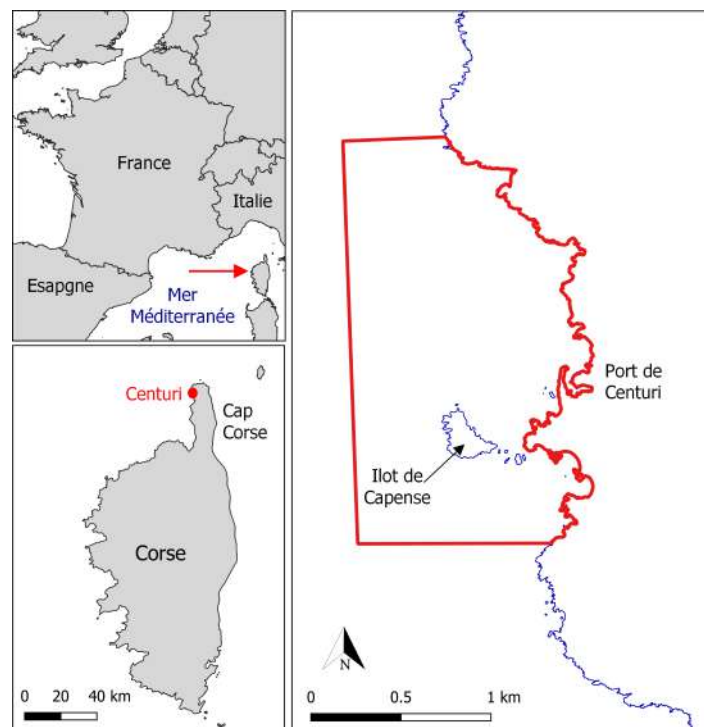


FIGURE 6.1 – Carte de situation de l'île méditerranéenne de la Corse en France (en haut à gauche) et la localisation de la commune de Centuri au Cap Corse (en bas à gauche). A droite, zone d'étude autour du port de Centuri et de la réserve naturelle de l'îlot de Capense.

La zone est majoritairement couverte d'herbiers de Posidonie continus sur fond rocheux et sur fond de sables fins. Le pourtour de l'îlot de Capense et la côte sont composés de rochers couverts d'algues photophiles. A plus grande profondeur, les sédiments deviennent plus grossiers, il s'agit de détritique côtier. Une partie des plages dans le sud de la zone est constituée de galets et le

port possède une digue avec un enrochement artificiel. Ces habitats sont également visibles sur les images aériennes de la BD Ortho 2016 (produit par l'IGN) de résolution 20 cm (fig. 6.2, partie gauche) : on distingue les zones de sable des herbiers jusqu'à une profondeur d'une quinzaine de mètres, les roches à algues sur les pourtours des zones émergées et les herbiers en couleurs plus foncées.

Les levés acoustiques au SMF ont été réalisés du 21 au 23 mai 2021 avec le matériel décrit dans la section 2.2 pour les plus grandes profondeurs, de 5 à 40 m. Les plus petites profondeurs, de 0 à 5m, sont des données lidar aéroporté acquises en 2017-2018 avec un capteur HawkEye 3 par le SHOM (Service Hydrographique et Océanographique de la Marine) pour le programme Litto3D. Les données bathymétriques sont fusionnées dans un MNT de résolution 50 cm et de projection géographique Lambert 93 (fig. 6.2, partie droite). Cette fusion de données issues de différents capteurs réduit les temps d'acquisitions très longs par tous petits fonds. Fezzani et al. (2019) propose également de la fusion de données bathymétriques acquises par différents sondeurs acoustiques pour construire un MNT exhaustif de la zone en réduisant les temps de levés.

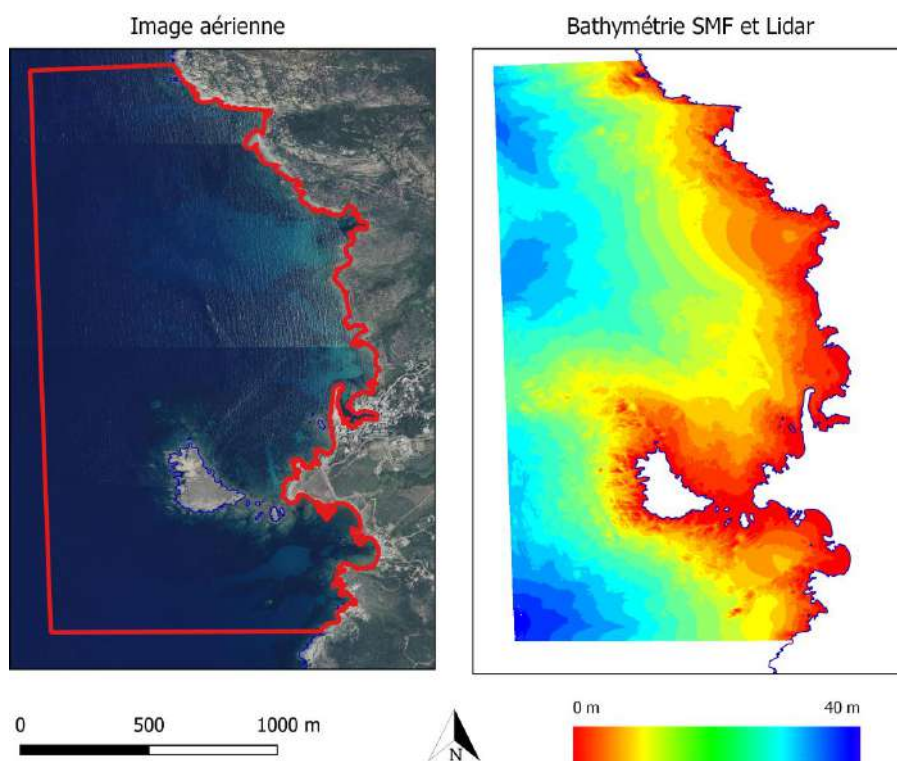


FIGURE 6.2 – A gauche, image aérienne de la BD ORTHO. A droite, Bathymétrie de la zone sous forme d'un MNT de résolution 50 cm, fusion de la bathymétrie acoustique et lidar.

La donnée de réflectivité n'est disponible que pour la partie de la zone levée au SMF et de plus, les traces de navigation rendent difficiles l'exploitation de cette donnée. Elle est illustrée à titre indicatif au niveau de la figure 6.3 sur la gauche. Les dérivées de la bathymétrie - pente et rugosité - sont calculées sur l'ensemble de la zone. Présentées en figure 6.3, elles sont assez proches en contenu et en valeurs (même échelle de fausses couleurs). Cependant, la pente rend plus visible la délimitation des différents substrats (couleur bleu foncé par opposition du reste) jusqu'aux grandes profondeurs.

Des données de vérité terrain ont été acquises par une caméra tractée ainsi que par des nageurs. La caméra est attachée à dépresseur en forme d'aile tractée par le navire Seaviews One et réalise

des vidéos des fonds le long d'un transect de navigation. La profondeur du dépresseur s'ajuste avec la vitesse du navire : à petite vitesse, l'aile coule et la caméra filme plus proche du fond, en augmentant la vitesse, l'aile remonte et les vidéos possèdent un plus grand angle de vue. Son utilisation est réservée pour des fonds de 10 à 25m de profondeur afin de repérer des roches et les limites d'herbiers de Posidonie. La position des images acquises est métrique car on ne sait pas exactement où se trouve l'aile derrière le navire. Les nageurs, eux, sont équipés d'une plateforme flottante embarquant un système GPS couplé à un appareil photographique : les photographies sont synchronisées à l'heure du GPS et leurs métadonnées contiennent la position. Les nageurs réalisent des transects dans les plus faibles profondeurs (moins de 10m). Leur attention se concentre sur les limites entre substrats (rochers, sédiments, herbiers de Posidonie), sur la présence d'espèces protégées (herbiers de Cymodocée, grandes nacres, etc.) ou remarquables (raies, grands poissons et crustacés, etc.) et sur la présence de macro-déchets. Un résumé des vérités terrain est proposé en figure 6.4.

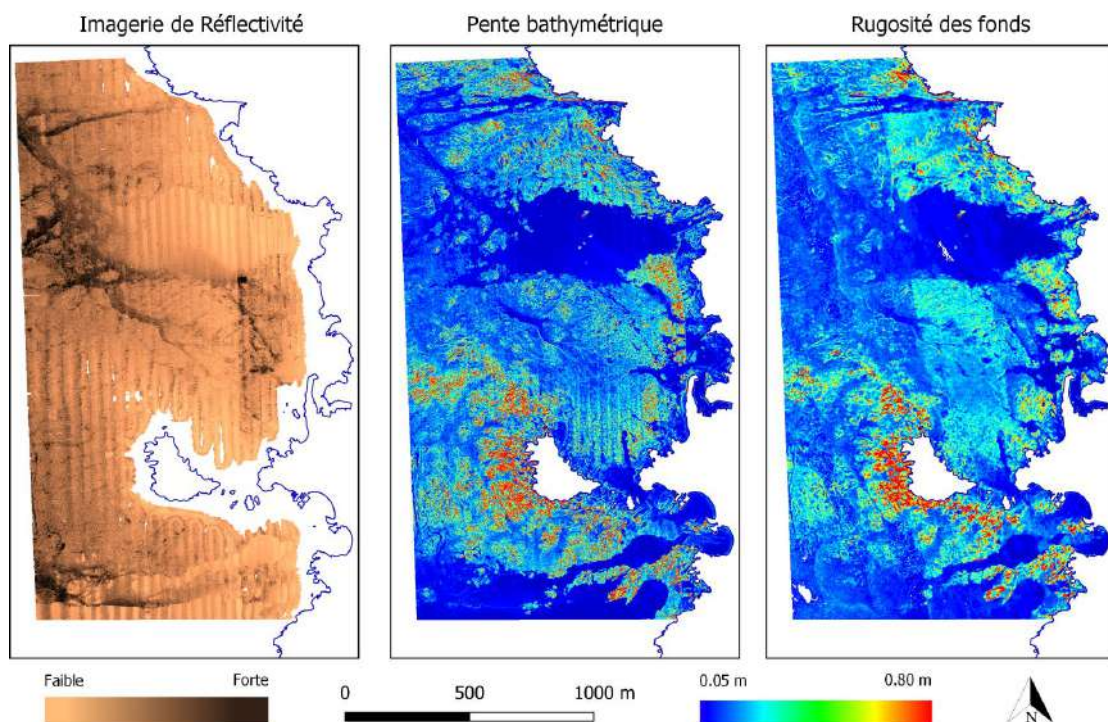


FIGURE 6.3 – A gauche, la carte de réflectivité des fonds issue du SMF montre la part de la zone acquise en acoustique (le reste correspond à de trop petits fonds). Au milieu, la carte de pente issue de la bathymétrie (acoustique et lidar). A droite, la carte de rugosité issue de la bathymétrie (acoustique et lidar).

6.2 Résultats des classifications

NB : Les résultats produits dans cette section sont des choix de classes et de représentations spécifiques pour ce travail de thèse. Ce ne sont pas les cartes et résultats qui ont été livrés au client.

Les résultats de classification manuelle sont présentés en premier car leur élaboration passe par une analyse des données. Cette analyse et interprétation est utile au lecteur pour comprendre et pour pouvoir comparer les résultats obtenus avec les autres méthodes. En comparaison d'un résultat établi sous le contrôle total d'un opérateur, est proposé celui obtenu par la méthode des *K*-moyennes où seule l'interprétation finale des classes est laissée à l'opérateur. Ensuite, viennent

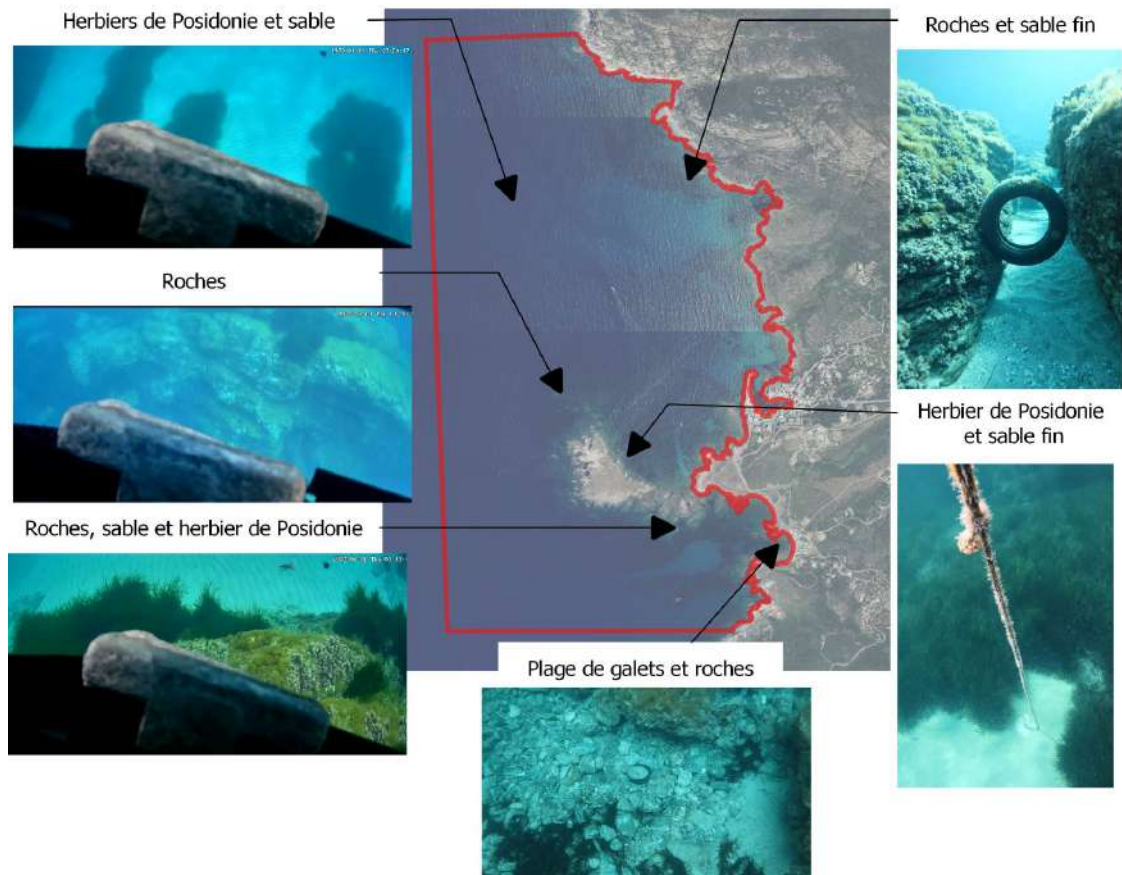


FIGURE 6.4 – Photographies de vérités terrain réalisées par caméra tractée (les trois photos de gauche) et par plateforme (photos de gauche et du bas).

les résultats obtenus au moyen de méthodes basées sur de l'apprentissage supervisé, dans lesquelles l'opérateur intervient en amont de l'apprentissage automatique.

6.2.1 Classification par interprétation

Pour la réalisation d'un arbre de décision construit manuellement dans Viewmap, les habitats suivants sont retenus selon la classification EUNIS (Gayet et al.; 2018) : (i) sables fins infralittoraux (EUNIS A5.235), (ii) sédiments du détritique côtier (EUNIS A5.46), (iii) herbiers de Posidonie (EUNIS A5.535), (iv) rochers à algues photophiles (EUNIS A3.33), (v) plages de galets (EUNIS A2.111) et (vi) enrochements artificiels ou digue portuaire (EUNIS J2.53).

Les sédiments (i) et (ii) se distinguent essentiellement par les profondeurs de répartition, le (i) est présent de 2 à 25 m tandis que le (ii) commence au-delà. Ils sont particulièrement bien visibles et identifiables sur les données de pente grâce à leurs très faibles valeurs. Il faut analyser la rugosité et la réflectivité pour les différencier, le détritique côtier est plus grossier donc plus rugueux et possède une plus forte réflectivité. En revanche, les zones de détritique côtier que l'on repère sur la pente sont moins visibles sur la rugosité. Les tâches de sables sont également visibles sur l'image aérienne, ce sont les tâches bleues turquoise.

Les herbiers sont présents à toutes les profondeurs dans cette zone. Ils sont caractérisés par une pente ou une rugosité plus fortes que les sédiments (i) et (ii) mais bien plus faibles que les

rochers dont les valeurs sont les plus élevées. La raison est simple : les sondes ne sont pas mesurées sur le sol lui-même mais soit sur la canopée des herbiers soit dans l'épaisseur de la Posidonie. Il en résulte des variations décimétriques de bathymétrie et donc une pente et une rugosité plus marquées qu'avec les sédiments (i) et (ii). Sur l'image de réflectivité, ils sont caractérisés par des valeurs intermédiaires entre le sable fin et le détritique côtier. Enfin, sur l'image aérienne, les herbiers apparaissent en couleurs foncées et les rochers de manières plus claires le long des côtes.

Les plages de galets sont présentes sur des surfaces limitées et se confondent avec le sable fin sur les images de pente et de rugosité. Tandis que sur l'image aérienne, les galets se confondent, en terme de couleurs, avec les rochers.

L'enrochement artificiel de la digue correspond aussi de surface très limitée. Il se confond avec des rochers car il présente une forte pente et rugosité à cause de l'organisation des blocs rocheux ou de béton. Il en est de même sur l'image aérienne, car les rochers de la digue sont couverts des mêmes algues.

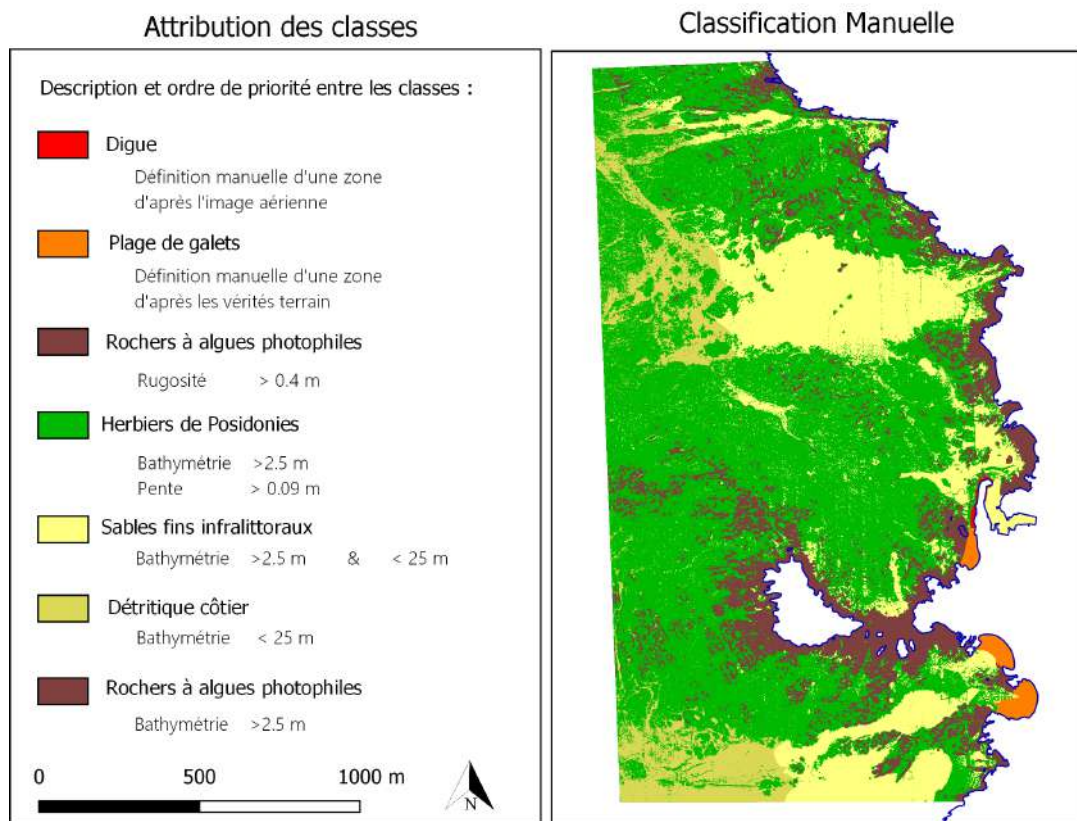


FIGURE 6.5 – Résultats de classification manuelle par interprétation des différentes données disponibles.

Ces analyses permettent à un opérateur d'associer mentalement une signature des fonds à un type d'habitat, puis de construire un arbre de décision dans le logiciel ViewMap. Plusieurs arbres sont possibles pour obtenir le résultat mentalement envisagé, cependant il est préférable de réaliser l'arbre le plus simple possible. Celui de la figure 6.5 à gauche se focalise sur la description des rochers et des herbiers de Posidonie, ensuite les autres classes peuvent être décrites avec moins de contraintes. Cependant, dans ce cas de figure, les rochers ne peuvent être décrits en une fois, on complète les petits fonds par des rochers en fin d'arbres. Ainsi, de cette manière, toutes les mailles sont attribuées à une classe dans le résultat de la carte en figure 6.5 à droite. Les classes supplémentaires de la digue et plages de galets sont représentées manuellement.

Ce résultat est un travail de plusieurs journées pour trouver la meilleure description des classes ainsi que l'ordre des priorités dans l'arbre. Cependant, il est impossible d'être sûr que cette classification manuelle soit totalement vraie. Elle servira néanmoins de support de comparaison pour les classifications automatiques à venir. Pour s'assurer, au moins visuellement, que cette carte est correcte, il est possible de la comparer avec les résultats du rapport [STARESO and ANDROMEDE-OCEANOLOGIE \(2012\)](#) dans lequel les données acoustiques utilisées proviennent d'un sondeur à balayage latéral.

6.2.2 Exploration des données par la méthode des K -moyennes

Avec la méthode non-supervisée des K -moyennes, on cherche à obtenir un partitionnement qui s'approche du résultat précédent. Pour la classification manuelle, nous avons retenu six habitats où, les sédiments (i) et (ii) sont différenciés d'après le système de classification EUNIS et non d'après les données et où la digue et les plages de galets sont obtenues manuellement. Pour cette exploration, nous ne cherchons à retrouver que trois habitats simplifiés : (i + ii + v) sédiments et galets, (iii) herbiers de Posidonie et (iv + v) substrats durs (rochers et digue) couverts d'algues photophiles. Comme les classes composées par l'algorithme ne seront pas nécessairement celles envisagées, le code couleur du résultat sera pris différent de la classification manuelle. Après l'obtention des résultats, l'opérateur analyse les classes réalisées et les rapproche des classes désirées.

Outre la valeur de K à initialiser à l'appel de la méthode des K -moyennes, se pose la question des caractéristiques à utiliser. Une première série d'essais est effectuée avec K égal à 3, autant de classes que d'habitats recherchés et trois sélections de caractéristiques.

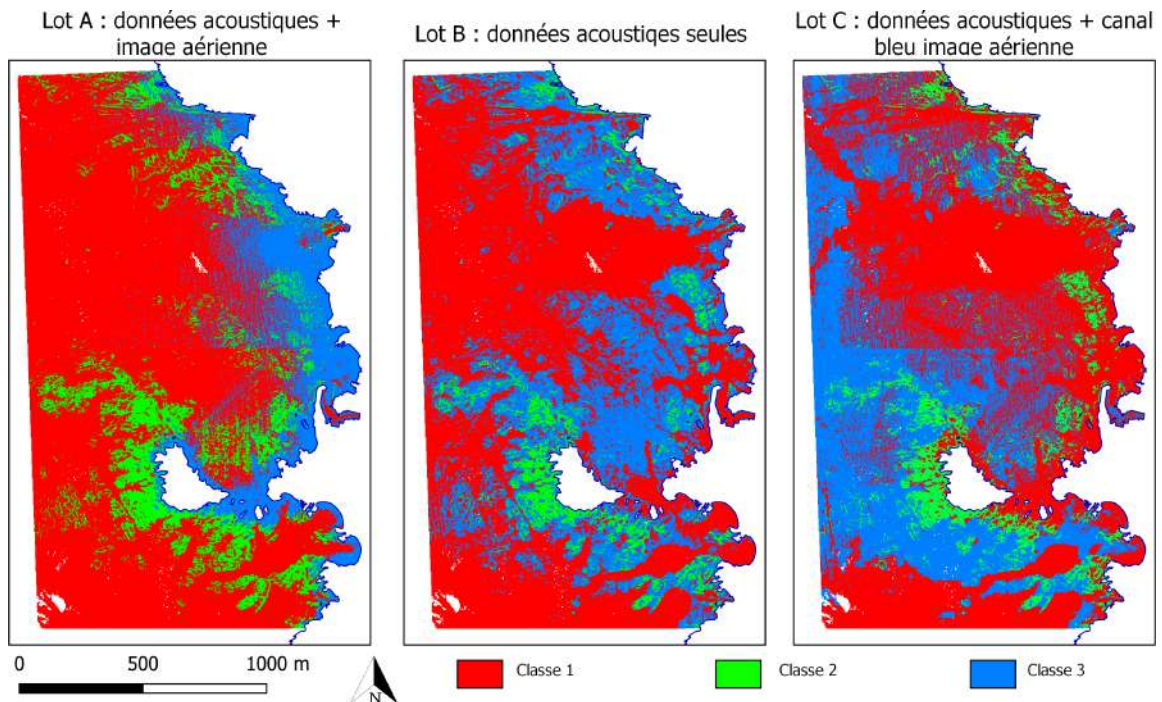


FIGURE 6.6 – Résultats des K -moyennes, avec $K=3$, pour différents lots de données.

Le premier lot de données, le lot A, est composé des données acoustiques (bathymétrie, rugosité et pente) et de l'image aérienne (rouge, vert, bleu). Seule l'image de réflectivité n'est pas utilisée car elle ne couvre pas toute la zone. Les résultats donnés à la figure 6.6 gauche font apparaître en classe 2 (verte) ce qui se rapproche le plus des rochers (iv) de la classification manuelle. Tandis que les classes 1 et 3 (rouge et bleue) sont plus difficilement interprétables mais leur séparation correspond aux couleurs de l'image aérienne, les zones les plus claires (et les moins profondes) sont en bleu et les plus foncées en rouge. De plus, les reflets du soleil sur l'image aérienne, sous forme

de petites tâches blanches, induisent des erreurs de classification ; les reflets sont classés comme les zones les plus claires, donc en bleu plutôt qu'en rouge.

Le lot B est quant à lui constitué uniquement des données acoustiques (bathymétrie, rugosité et pente), avec donc moitié moins de caractéristiques (3 au lieu de 6) que le lot A. Les résultats (fig. 6.6 milieu) semblent assez proches de la classification manuelle jusqu'à mi-profondeur : en rouge, les zones correspondent correctement aux zones de sédiments et galets (i + ii + v) et en bleu aux herbiers de Posidonie. Cependant, à plus grande profondeur, la classe rouge devient trop majoritaire par rapport à la classe bleue. La classe verte, correspondant aux rochers, reste encore assez juste.

Pour le lot C, nous avons réintroduit une partie de l'information de l'image aérienne dans la mesure où seul le canal bleu est utilisé avec une pondération de 3 et nous avons conservé les données acoustiques. Le bleu est la couleur qui est absorbée le moins vite par la hauteur d'eau et grâce à la clarté des eaux corses, les zones de sable sont visibles plus profondément - quasiment jusqu'au plus profond de la zone. Le lot C est alors composé comme le lot A à moitié de données acoustiques (bathymétrie, rugosité, pente) et à moitié de données optiques (canal bleu compté trois fois). Cette fois-ci, les classes rouge et bleu (fig. 6.6 droite), que l'on peut encore associer aux sédiments et galets (i + ii + v) et aux herbiers de Posidonie, semblent bons à toutes les profondeurs au regard de la classification manuelle. Cependant, il y a toujours les reflets du soleil sur le canal bleu qui induisent des erreurs entre les classes 1 et 3. D'autre part, l'image aérienne est en réalité composée d'au moins deux photos, dont l'assemblage se voit nettement comme une ligne horizontale au milieu de la zone, cette démarcation est plus discrète sur la figure 6.6 partie gauche.

La classification manuelle est de résolution 50 cm comme les autres MNT qui ont servi à sa construction. Cependant, les calculs des K -moyennes comportent de multiples itérations sur l'ensemble des mailles : les calculs sont donc très longs à pleine résolution (50 cm) car la zone compte 9.5 millions de mailles, comme en témoigne le tableau 6.1. Les résultats de la figure 6.6 sont produits à une résolution de 2 m, ce qui divise par 16 le nombre de mailles (environ 600 000 mailles) et ramène les temps de calculs de l'ordre d'une demie minute.

TABLE 6.1 – Temps de calcul d'une classification par la méthode des K -moyennes pour différentes résolutions.

	Résolution		
	0.5 m	1 m	2 m
Lot A	606 s	137 s	27 s
Lot B	266 s	65 s	19 s
Lot C	882 s	127 s	26 s

La question de savoir si le nombre de classes considérées ($K = 3$) est la bonne initialisation pour les lots de données a été également étudiée. Le nombre de classes à réaliser, K , a été fixé précédemment par connaissance *a priori* des habitats alors que l'inertie intra-classe peut être plus basse avec une autre valeur de K .

Pour cela, le lot B a été testé pour des valeurs K de 1 à +10. Le nombre d'itérations ainsi que le temps de calcul augmentent avec la valeur de K et à l'inverse, l'inertie intra-classe diminue (voir le tableau 6.2 et la courbe associée en figure 6.7). L'inertie intra-classe diminue lorsque K augmente jusqu'à ce que K soit équivalent au nombre d'éléments à trier : lorsque qu'il n'y a plus qu'un élément par classe, l'inertie au sein de la classe est nulle. Cependant, cette diminution n'est pas linéaire et dessine en général "un coude" : l'inertie diminue fortement avec des petits K puis plus doucement ensuite à partir d'un "coude" qui indique alors la valeur idéale de K .

Cette méthode du coude conduit dans cet exemple à K idéal à 4 où la classe 1 (en rouge, les sédiments) est scindée en 2 partitions selon l'isobathe 18 m. Ce résultat est à notre avantage puisque dans les autres méthodes, nous différencions le sable fin infralittoral du détritique côtier

TABLE 6.2 – Étude du temps de convergence des K -moyennes pour différentes valeurs de K

Nombre de classes	1	2	3	4	5	6	7	8	9	10	15
Nombre d'itérations	4	38	39	49	74	84	198	178	152	205	284
Temps de calcul (s)	1	15	15	20	30	35	84	82	71	96	138
Inertie Intra-classe ($\times e^{11}$)	4,4	3,6	2,9	2,6	2,5	2,3	2,1	1,8	1,7	1,6	1,4

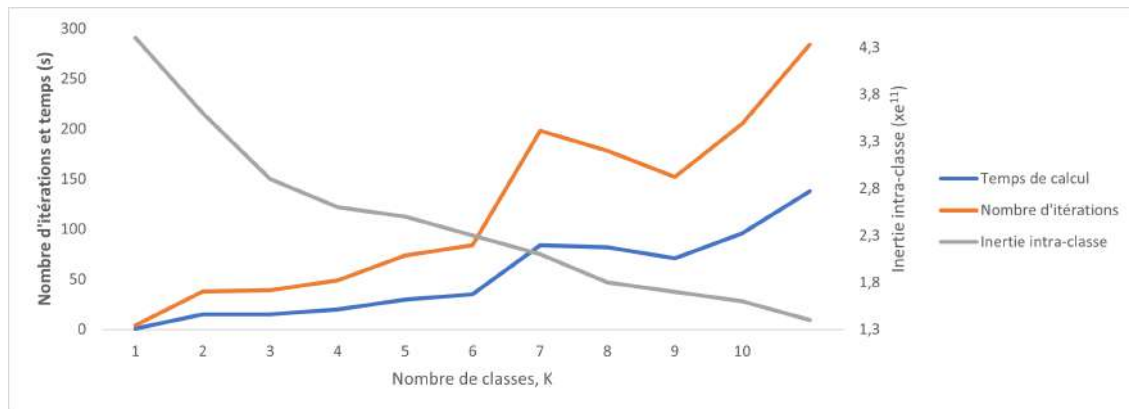


FIGURE 6.7 – Evolution de l'inertie intra-classe et du temps de calcul avec l'augmentation du nombre de classes demandé à la méthode des K -moyennes.

dont les limites varient d'une vingtaine à une trentaine de mètres de profondeurs (limite à 25m dans l'arbre de décision manuel). Cette séparation à 18 m n'est pas une "réalité physique" du terrain, mais dépend plutôt de la valeur moyenne des profondeurs de la zone. Cette méthode a également été testée sur le lot C, ce qui conduit aussi à $K=4$, cependant c'est alors la classe type rocher qui se scinde en deux et cela n'est pas lié à une réalité physique du terrain. Pour que les sédiments se scindent en deux avec le lot C, il faut augmenter K jusqu'à 7.

6.2.3 Résultats obtenus avec des méthodes à apprentissage supervisé

L'emploi des méthodes de classification nécessite de fournir à l'algorithme un jeu d'apprentissage composé d'éléments représentatifs des classes. Il est fortement conseillé pour la méthodes des FA que chaque classe du jeu d'apprentissage soit représentée de manière équivalente. Pour des comparaisons cohérentes de résultat, le même jeu d'apprentissage est utilisé pour les trois méthodes supervisées. Ainsi, les quatre grandes classes - rochers à algues photophiles, sables fins infralittoraux, détritique côtier et herbiers de Posidonie - sont représentées, sur la figure 6.8, par une série de sélections (manuelles) choisies d'après la classification manuelle (figure 6.5) et réparties sur l'ensemble de la zone. Les deux classes - digue et galets - sont trop petites au regard des quatre autres et ne sont pas considérées dans cette section.

Pour les trois méthodes à venir, nous considérons les trois lots de caractéristiques proposés dans les résultats des K -moyennes. Le choix de la résolution de la carte de classification se concentre sur 50 cm ou 1 m : les entraînements se faisant sur une petite portion de données, il est plus envisageable de travailler à pleine résolution (50 cm) ou éventuellement à 1m. L'exactitude de la phase d'entraînement est évaluée par rapport aux *labels* des sélections sur une portion de 20% du jeu d'apprentissage mis de côté, tandis que l'exactitude de la prédiction est estimée sur l'ensemble de la zone comparée à la classification manuelle.

6.2.3.1 Méthode du classifieur bayésien normal

Nous commençons par présenter les résultats de classification par la méthode bayésienne car celle-ci ne comporte aucun paramètre à régler. Le tableau 6.3 présente les temps de calcul mis par

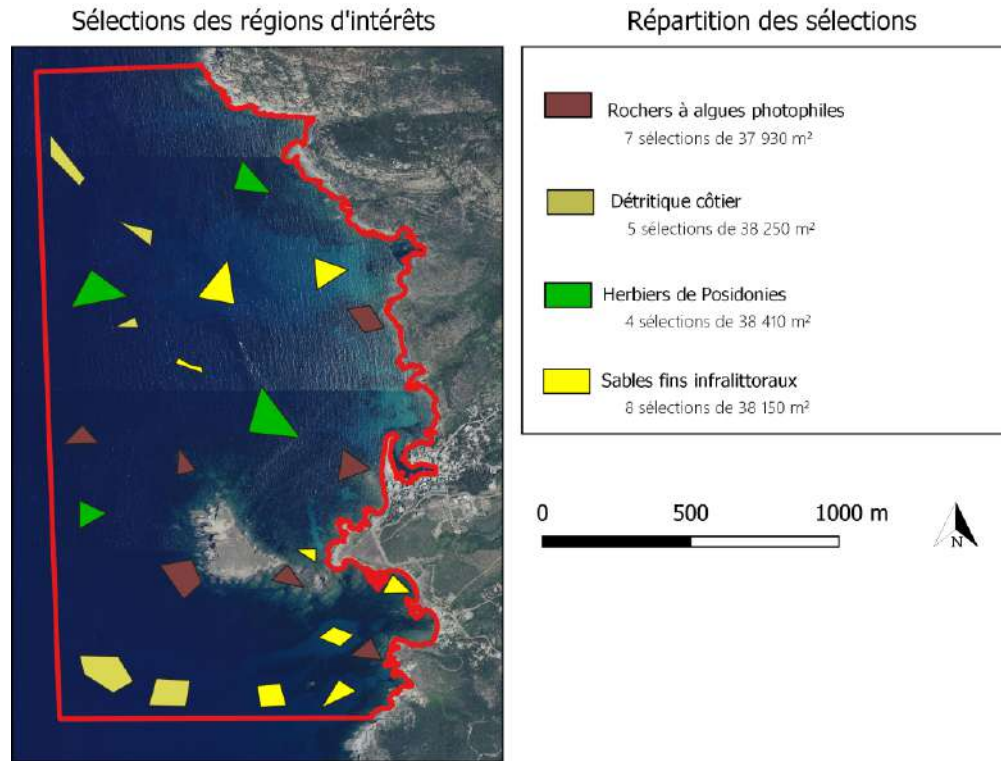


FIGURE 6.8 – Composition du jeu d'apprentissage.

la méthode pour réaliser une classification à pleine résolution 50 cm, les temps sont extrêmement courts : moins d'une minute est nécessaire pour réaliser l'apprentissage puis la prédiction. Nous n'avons donc pas jugé nécessaire de relancer des tests avec une résolution plus grossière. Pour les trois lots, les exactitudes sont très bonnes et tournent autour de 90% à l'entraînement puis de 80% à la prédiction (tableau 6.3). A elles-seules, ces valeurs ne permettent pas de décider du meilleur lot.

TABLE 6.3 – Temps de calcul et exactitude des phases d'entraînement et de prédiction avec la méthode du classifieur bayésien et une résolution de 50 cm

	Entraînement		Prédiction	
	Temps (s)	Exactitude (%)	Temps (s)	Exactitude (%)
Lot A	30	92.5	19	77.7
Lot B	10	87.8	15	80.4
Lot C	14	89.8	16	80.5

Au premier regard, les résultats, présentés en figure 6.9, sont très ressemblants entre eux mais également avec le support de comparaison qui constitue la classification manuelle. Cependant, avec plus d'attention, la carte obtenue avec le lot A fait apparaître les artefacts de l'image aérienne (reflets du soleil et assemblage des photos) en tant que rochers (entre l'îlot et la côte). Et, celle obtenue avec le lot B néglige une partie des herbiers sur la moitié haute gauche de la carte. Visuellement, c'est donc la troisième carte, celle obtenue avec le lot C qui semble la plus comparable avec la classification manuelle, comme ce qui avait été conclu avec la méthode des K -moyennes. Mais, dans les trois cas, ce classifieur augmente - mais peut-être à raison - la quantité de rochers le long de la côte par rapport à la classification manuelle.

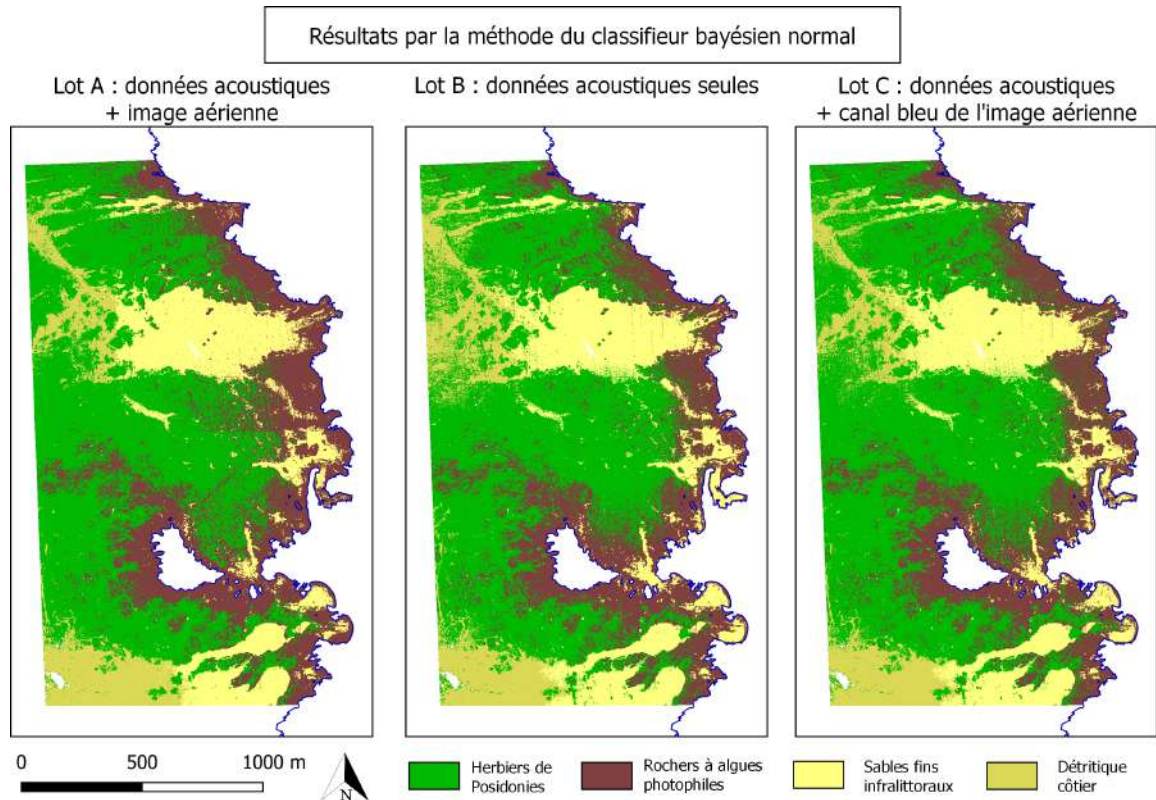


FIGURE 6.9 – Résultats du classifieur bayésien normal pour différents lots de données.

6.2.3.2 Méthode des forêts aléatoires d'arbres décisionnels

La méthode des FA possèdent des paramètres qui peuvent être “réglé une fois pour toute” avec des valeurs par défaut. Ici, nous nous contentons de jouer sur deux paramètres : le nombre d'arbres décisionnels à réaliser (N_{arbre}) et leur profondeur maximale ($prof_max$). Le nombre d'arbres influe sur la capacité du modèle à diluer les erreurs de chaque arbre au travers d'une moyenne de leur ensemble, nous fixons $N_{arbre} = 200$ dans un premier temps. La profondeur maximale d'un arbre est un paramètre de “pré-élagage” qui permet d'éviter le sur-apprentissage dans chacun des 200 arbres, nous le fixons volontairement bas à $prof_max = \lceil \log_2(K) \rceil$ (à la partie entière supérieure du logarithme binaire du nombre de classes), c'est-à-dire à 2 dans notre cas où nous avons 4 classes.

Le tableau 6.4 présente des temps de calcul lors des phases d'entraînement et de prédiction, avec une résolution de 50 cm, plus longs qu'avec le classifieur bayésien tout en restant très courts (moins de 3 minutes au total) - donc ne requérant pas de passer sur une résolution plus grossière. Les exactitudes d'entraînement sont un peu moins bonnes que celles du tableau 6.3 tandis que curieusement, les exactitudes lors de la prédiction sont plus hautes qu'à l'entraînement et meilleures qu'avec le classifieur bayésien. Visuellement, les trois résultats sont très semblables, ils présentent moins de rochers sur les côtes que la classification manuelle et omettent les rochers éloignés de la côte. Les trois autres classes sont bien représentées. Nous présentons le résultat avec le lot B en figure 6.10 à droite .

Nous modifions cette fois-ci la profondeur des arbres en essayant les valeur de K puis $K * 2$, c'est-à-dire 4 puis 8. Les temps de calculs - entraînement et prédiction - sont plus longs puisque le modèle crée des arbres plus complexes. Ces temps sont de l'ordre de 5 à 8 minutes avec ces nouvelles profondeurs (tableau 6.5). Les exactitudes d'entraînement sont cette fois-ci plus élevées que précédemment et celles de prédiction plus basses, tout en restant très bonnes. Il y a lieu de se

TABLE 6.4 – Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires, une résolution de 50 cm et les paramètres ($N_{arbre} = 200, prof_max = 2$).

	Entraînement		Prédiction	
	Temps	Exactitude (%)	Temps	Exactitude (%)
Lot A	2 min 21 s	82.3	25 s	85.5
Lot B	1 min 55 s	82.4	25 s	87.1
Lot C	2 min 10 s	83.7	25 s	83.5

demander si le modèle ne fait pas du sur-apprentissage car bien que l’exactitude d’entraînement augmente avec la profondeur, elle plafonne à 80-81% à la prédiction. Visuellement, les six résultats sont très ressemblants entre eux et ont en commun de faire apparaître plus de roche sur les côtes que la classification manuelle, plus encore qu’avec le classifieur bayésien. Le résultat avec le lot A et une profondeur d’arbre de 8 est présenté en figure 6.10 au milieu. Concernant le lot A, avec les essais sur trois profondeurs différentes, les résultats produits ne font plus apparaître les artefacts de l’image aérienne que l’on retrouvait avec les K -moyennes et le classifieur bayésien.

TABLE 6.5 – Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires et une résolution de 50 cm avec une variation de la profondeur des arbres.

	Paramètres	Entraînement		Prédiction	
	($N_{arbre}, prof_max$)	Temps	Exactitude (%)	Temps	Exactitude (%)
Lot A	(200, 4)	3 min 40 s	92.8	1 min 5 s	80.6
Lot B	(200, 4)	3 min 15 s	92.5	50 s	81.0
Lot C	(200, 4)	3 min 15 s	93.1	58 s	80.4
Lot A	(200, 8)	5 min 15 s	95.6	2 min 50 s	81.3
Lot B	(200, 8)	5 min 17 s	93.4	2 min 21 s	81.7
Lot C	(200, 8)	5 min 5 s	94.8	2 min 40 s	81.3

Enfin, nous augmentons le nombre d’arbres à $N_{arbre} = 500$ et nous refaisons les tests pour les 3 lots de caractéristiques et les 3 profondeurs d’arbre. Les temps et les exactitudes sont reportés dans le tableau 6.6. Finalement, quelles que soient les profondeurs d’arbres, les exactitudes d’entraînement et de prédiction sont similaires à celles des tableaux 6.4 et 6.5 avec des temps de calcul des deux phases qui doublent avec l’augmentation du nombre d’arbres. Visuellement, les résultats sont très proches, voire même identiques, de ceux obtenus avec seulement 200 arbres. Le résultat avec le lot B et une profondeur de 4 est présenté sur la figure 6.10 à droite.

6.2.3.3 Méthode des Séparateurs à Vastes Marges

Au contraire de la méthode des FA, la méthode des SVM ont au moins deux paramètres (selon le noyau utilisé, voir le tableau 5.3) et ceux-ci doivent être adaptés au jeu de données d’apprentissage. Il n’est donc pas question de les fixer une fois pour toute. Nous utilisons dans un premier temps la fonction *trainAuto* d’OpenCv laissant l’algorithme trouver les bons paramètres (γ et ε pour un noyau à base radiale) lors de l’apprentissage. Puisque l’algorithme établit une grille de recherche des paramètres optimaux, la phase d’apprentissage dure beaucoup plus longtemps : de 3 à 4 minutes avec une résolution à 1 m et de 15 minutes avec une résolution de 50 cm. Il semble que les résultats ne soient pas aussi répétitifs qu’avec les précédentes méthodes car d’un essai à

TABLE 6.6 – Temps de calcul et exactitude des phases d’entraînement et de prédiction avec la méthode des forêts aléatoires et une résolution de 50 cm avec une augmentation du nombre d’arbre et une variation de la profondeur des arbres.

	Paramètres (<i>N_arbre</i> , <i>prof_max</i>)	Entraînement		Prédiction	
		Temps	Exactitude (%)	Temps (s)	Exactitude (%)
Lot A	(500, 2)	5 min 44 s	87.7	1 min 22 s	82.6
Lot B	(500, 2)	5 min 14 s	83.2	1 min 5 s	87.1
Lot C	(500, 2)	5 min 6 s	83.6	1 min 11 s	87.2
Lot A	(500, 4)	8 min 25 s	92.8	2 min 51 s	79.7
Lot B	(500, 4)	8 min 8 s	92.0	2 min 13 s	82.0
Lot C	(500, 4)	7 min 47 s	93.1	2 min 32 s	80.8
Lot A	(500, 8)	12 min 46 s	95.7	8 min 50 s	81.1
Lot B	(500, 8)	13 min 16 s	93.3	7 min 33 s	81.7
Lot C	(500, 8)	12 min 52 s	94.7	8 min 40 s	81.1

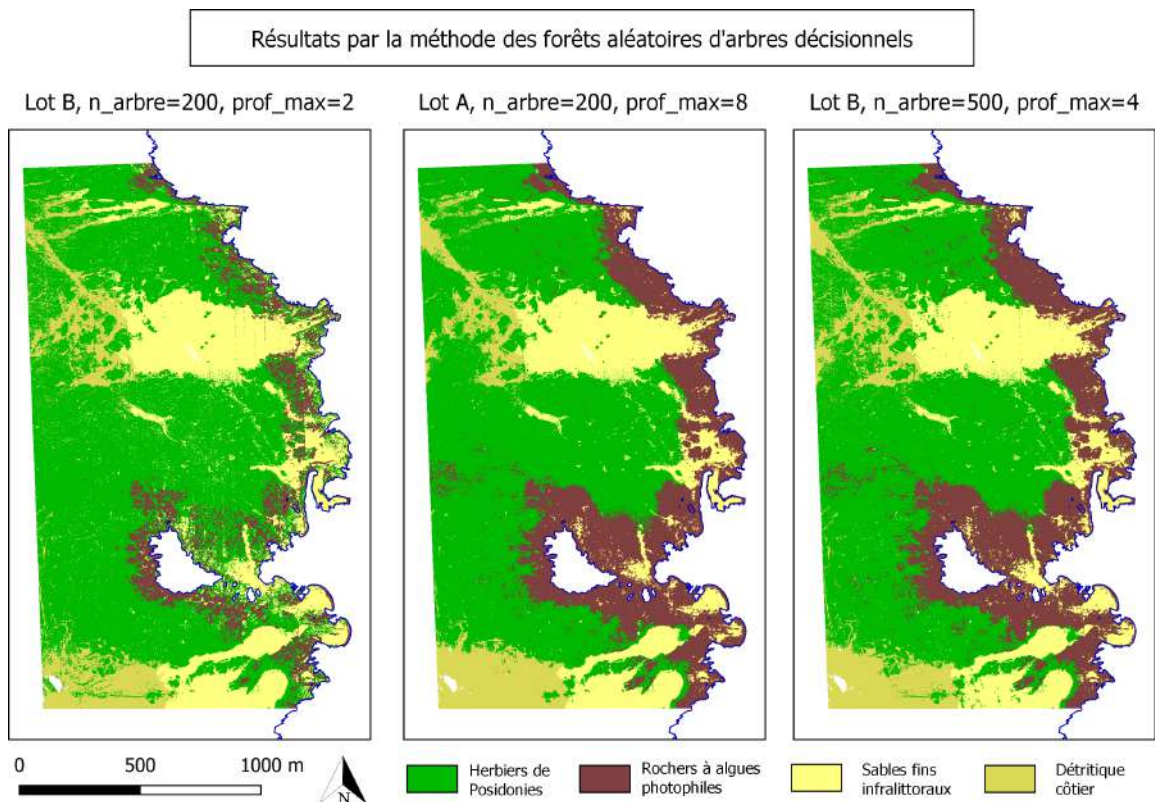


FIGURE 6.10 – Résultats de la méthode des forêts aléatoires pour différents lots de données et une résolution de 50 cm.

l’autre les paramètres, les exactitudes, l’allure globale de la carte varie beaucoup. En témoigne le tableau 6.7. La figure 6.11 présente trois résultats du précédent tableau avec un mauvais résultat et deux bons résultats. Les deux bons résultats diffèrent sur leur quantité de rochers attribués dans la classification : le lot A s’approche des résultats avec le classifieur bayésien tandis que le lot C s’approche plus de la classification manuelle. Dans les deux cas, ces lots négligent une partie des

herbiers de Posidonie dans la partie haute gauche de la carte.

TABLE 6.7 – Temps de calcul, paramètres et exactitude des phases d’entraînement et de prédiction avec la méthode des SVM et une résolution de 50 cm ou 1 m.

	Résolution de la carte	Paramètres	Entraînement		Prédiction		Appréciation visuelle
		(γ, ε) trainAuto	Temps	Exactitude (%)	Temps (s)	Exactitude (%)	
Lot A	1 m	(0.506, 0.5)	3 min 40 s	95.1	1 min 5 s	67.2	Mauvais
Lot A	1 m	(0.506, 0.5)	3 min 40 s	97.7	1 min 2s	71.6	Mauvais
Lot A	1 m	(0.506, 0.5)	3 min 50 s	97.9	1 min 3 s	75.5	Bien
Lot B	1 m	$(1e^{-5}, 0.1)$	2 min 40 s	63.4	1 min 42 s	36.1	Mauvais
Lot B	1 m	$(2.25e^{-3}, 0.5)$	2 min 34 s	90.1	1 min 34 s	74.2	Bien
Lot B	50 cm	$(1.5e^{-4}, 2.5)$	14 min 29 s	79.2	6 min 50 s	69.1	Mauvais
Lot C	1 m	(0.506, 0.5)	2 min 55 s	92.3	49 s	82.2	Bien
Lot C	50 cm	$(1e^{-5}, 0.1)$	15 min 20 s	75.6	6 min 58 s	49.1	Mauvais
Lot C	50 cm	(0.506, 2.5)	15 min 3 s	91.4	3 min 40 s	80.3	Bien

Afin de s’épargner les apprentissages longs avec *trainAuto* à pleine résolution, il est possible d’utiliser les apprentissages réalisés à résolution 1 m et de les appliquer sur les données à résolution 50 cm. Les résultats obtenus par “transfert” sont identiques visuellement et à exactitude très proche du résultat obtenu lors de l’apprentissage (voir tableau 6.8).

TABLE 6.8 – “Transferts” d’apprentissage d’une plus basse à une plus haute résolution.

Apprentissage (Lot, résolution, (γ, ε))	Application (Lot, résolution)	Temps de Prédiction	Exactitude de prédiction (%)
Lot A, 1 m, (0.506, 0.5)	Lot A, 50 cm	4 min 33 s	74.9
Lot B, 1 m, $(2.25e^{-3}, 0.5)$	Lot B, 50 cm	6 min 22 s	73.8
Lot C, 1 m, (0.506, 0.5)	Lot C, 50 cm	3 min 25 s	82.0

6.3 Discussion des méthodes et résultats

6.3.1 Interprétation humaine *versus* raisonnement machine

Le raisonnement appliqué pour une classification manuelle consiste à établir la liste des habitats à représenter puis à analyser en conséquence les données disponibles pour les caractériser. Ce schéma mental est l’inverse de ce qui est réalisé avec une méthode à apprentissage non-supervisé (type *K*-moyennes), la méthode cherche une partition des données, puis l’opérateur essaie de relier les classes aux habitats.

Retour sur la classification manuelle L’analyse des différentes données est assez rapide à établir car le nombre d’habitats, ici, est restreint et qu’ils sont reconnaissables sur les différents MNT disponibles. D’approche simple et intuitive, la construction de l’arbre de décision d’après ces

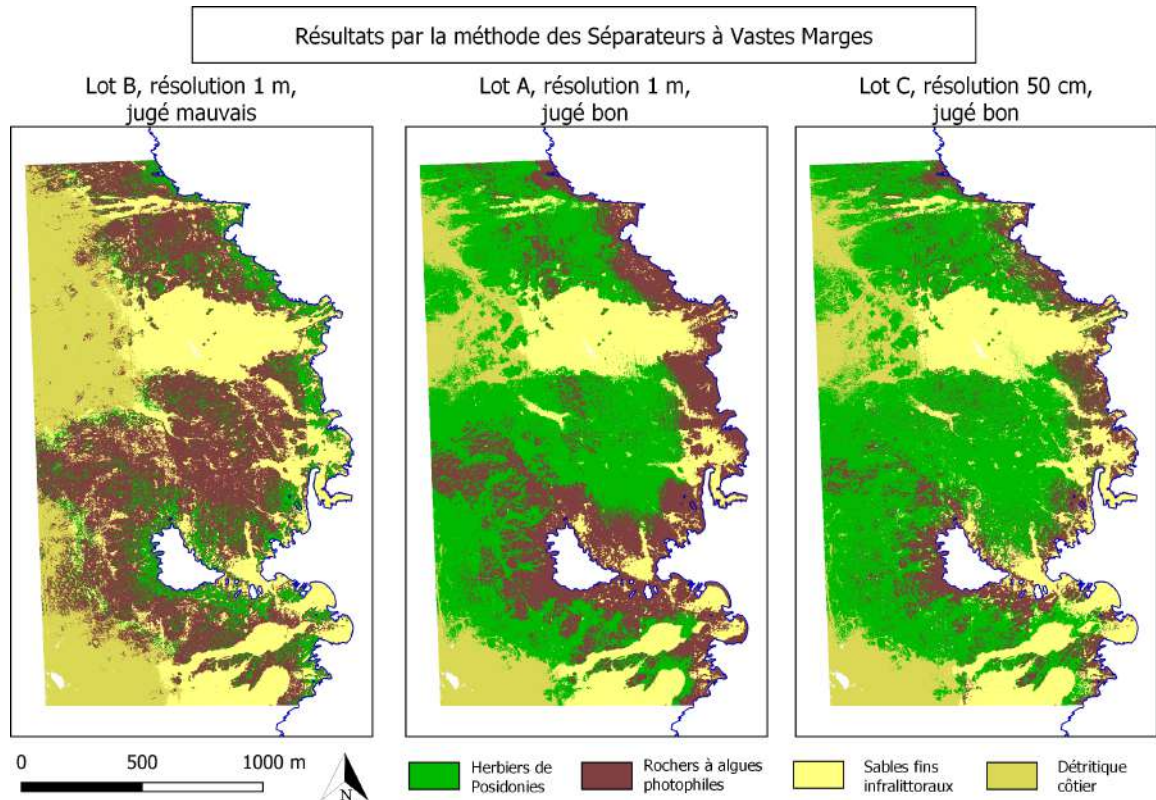


FIGURE 6.11 – Résultats des SVM pour différents lots de données et une résolution de 50 cm ou 1 m.

analyses ne devrait pas devenir un casse tête. Et pourtant, elle le devient rapidement car plusieurs complications entrent en jeu :

- jongler d’abord entre les différentes caractéristiques pour trouver celle qui décrit au mieux une classe donnée,
- trouver ensuite les bonnes plages de valeurs dans les caractéristiques correspondant aux classes en tenant compte des valeurs aberrantes et des disparités de valeurs d’un endroit à un autre dans la zone,
- ne pas sur-contraindre une classe car il s’agit d’un “ET” entre les différentes conditions, ce qui tend à augmenter le nombre de mailles non classées,
- l’ordre d’apparition des classes dans l’arbre modifie significativement le résultat.

Ces quatre considérations peuvent engendrer de nombreux essais d’arbres avant de trouver une solution cohérente pour chaque classe envisagée. Mais, heureusement, un opérateur expérimenté est plus efficace qu’un novice pour obtenir un résultat satisfaisant. Un autre “piège” consiste à chercher la perfection en ajoutant des contraintes, en ajustant les valeurs ou en modifiant l’ordre des classes ne conduit pas à un meilleur résultat global : “le mieux est l’ennemi du bien”. Il vaut mieux alors s’en tenir à cet arbre et effectuer des post-traitements tels que la suppression des objets trop petits et la réaffectation manuelle de classes à certains endroits comme pour la digue et les plages de galets.

Retour sur les K -moyennes Les partitions réalisées avec les K -moyennes (figure 6.6) montrent que la méthode est très sensible au contenu des données et à la composition des lots de données. Les extrêmes (et donc les valeurs aberrantes) dans les différentes données ont plus de chance d’être représentés dans la classification, d’où les rochers (maximum de pente et rugosité) bien découpés et les reflets du soleil (saturation des senseurs optiques) dès que l’image aérienne intervient. On ne

peut donc pas espérer un résultat satisfaisant en fournissant à la méthode toutes les caractéristiques sans tri préalable. De même, un pré-traitement sur les images aériennes pour réduire les reflets du soleil et homogénéiser l'image au niveau des coupures, ce qui devrait permettre d'améliorer significativement le résultat.

Ce type de méthode se concentre uniquement sur les informations contenues dans les données, à l'inverse de la méthode manuelle qui, plus souple, permet d'injecter des informations extérieures. En effet, il faut forcer l'apparition du détritique côtier en scindant la bathymétrie en deux (plus et moins de 25 m de profondeur) ainsi que des galets et de la digue par des zones dessinées manuellement dans la classification manuelle. C'est pourtant bien une réalité physique -le détritique côtier est plus grossier que le sable fin et abrite une faune différente, les galets sont bien différents des zones de sable et les rochers de la digue des rochers naturels - mais cela ne transparaît pas dans les données. Cela confirme l'intérêt de conserver un accès à la classification par des retouches manuelles.

En revanche, sans connaissances *a priori* sur les fonds, cette méthode permet de réaliser une première exploration de la zone : les résultats avec différentes informations ou caractéristiques font ressortir les objets d'intérêts. Dans notre cas, cela nous a surtout permis de vérifier que les habitats désirés contiennent assez d'information pour qu'un algorithme puisse les représenter (carte obtenue avec le lot C dans la figure 6.6). Puis la carte obtenue avec lot A, nous rappelle que les artefacts de l'image aérienne peuvent vraiment altérer les résultats.

6.3.2 Méthodes supervisées : alliance de la vision large de l'opérateur et de l'efficacité de la machine

Les trois méthodes supervisées testées font partie des méthodes les plus performantes de l'apprentissage automatique. Elles ont d'ailleurs été mises en compétition avec d'autres méthodes non présentées dans ce travail et ont présenté des résultats bien supérieur aux méthodes des *K*-plus-proches-voisins, des perceptrons multi-couches et des *spectral angle mapper*. Leurs résultats sont plus que satisfaisants. Il est, malgré tout, assez difficile et inintéressant de chercher *la* meilleure méthode car les performances vont varier d'un type de jeu de données à un autre.

Les perceptrons multi-couches, type de réseau de neurones à propagation directe où les neurones sont totalement connectés, seraient aussi à ajouter à cette liste de méthodes très performantes mais ils sont beaucoup plus long à entraîner et à exécuter. Par rapport à la partie II, où il est question d'apprentissages très longs (en semaines), ici il n'est pas envisageable de dépasser un calcul d'une demie-heure : l'opérateur attend son résultat et relance très certainement, dans la foulée, un autre calcul avec d'autres réglages afin d'obtenir plusieurs cartes pour comparaison. L'objectif étant de garder la meilleure.

Ainsi, les trois méthodes supervisées étudiées sont très rapides (moins de 5 minutes) à condition de choisir une résolution de calcul adaptée à la méthode (ceci est également vrai pour la méthode des *K*-moyennes). On peut donc estimer qu'un opérateur passera une demie-journée à réaliser une classification à partir du moment où les données sont prêtes à cet emploi. Dans ce laps de temps, cet opérateur :

- teste ses données avec la méthode des *K*-moyennes,
- construit un jeu d'apprentissage avec des sélections pour chacun des habitats,
- effectue plusieurs essais de classification en comparant les méthodes et les caractéristiques utilisées,
- réalise éventuellement des corrections manuelles sur la carte voire du post-traitement.

Dans l'ordre des essais de classification, il est certainement intéressant de commencer avec la méthode du classifieur bayésien. Cette méthode sans paramètre est extrêmement rapide : on peut donc tester de multiples combinaisons de caractéristiques et évaluer si le jeu d'apprentissage est suffisamment exhaustif pour fournir un résultat cohérent. Avec de très bons taux d'exactitude, cette méthode peut suffire à elle-seule. Son seul inconvénient est sa sensibilité aux défauts des

données : la carte obtenue avec le lot A dans la figure 6.9 fait apparaître les artefacts de l'image aérienne comme la méthode des K -moyennes.

Ce défaut de sensibilité est moins présent avec la méthode des forêts aléatoires : la moyenne des multiples arbres décisionnels créés rend robuste la méthode. La carte obtenue avec lot A de la figure 6.10 ressemble à celle de droite de la même figure. C'est aussi pour cette raison qu'avec des réglages identiques, les cartes sont très ressemblantes qu'importe le lot de caractéristiques utilisé. Cependant, la qualité du jeu d'apprentissage est un facteur important : ce même moyennage peut faire disparaître une classe sous-représentée dans le jeu d'apprentissage. Quant au paramétrage, il semble que le nombre d'arbres $N_arbres = 200$ soit suffisant dans notre cas d'application car les essais à 500 arbres ne sont pas meilleurs, simplement plus longs à obtenir. En revanche, la formule de profondeur des arbres $prof_max = \lceil \log_2(K) \rceil$ est totalement empirique. Elle permet d'obtenir des résultats plus proches de la classification manuelle, alors que les essais avec des arbres plus profonds laissent supposer du sur-apprentissage : l'entraînement est quasiment parfait (92-97%) mais la prédiction plafonne à 81%. En attendant de futurs essais de classification sur d'autres jeux de données, les paramètres $N_arbres = 200$ et $prof_max = \lceil \log_2(K) \rceil$ sont conservés.

La méthode des Séparateurs à Vastes Marges a posé plus de difficulté que les précédentes méthodes et ce pour trois raisons. Premièrement, il faut se fier à l'ajustement des paramètres par la fonction *trainAuto* car il est laborieux de les trouver soit-même. Deuxièmement, la non-répétitivité des résultats laisse perplexe ; d'un essai à l'autre, avec les mêmes caractéristiques, seul le jeu d'apprentissage est variant car 20% en est choisi aléatoirement pour être mis de côté afin d'évaluer l'entraînement. Et enfin, par rapport aux forêts aléatoires ayant un paramètre de pré-élagage pour limiter le sur-apprentissage, il n'y pas un tel contrôle avec les SVM proposés dans OpenCV. Cela peut expliquer une partie de la non-répétitivité des résultats car avec 90-98% d'exactitude à l'apprentissage il est surprenant d'obtenir des résultats aussi décevants comme ceux de la carte de gauche de la figure 6.11. Néanmoins, à force d'essai, la méthode est aussi capable de produire de très bons résultats comme la carte de droite de la même figure.

Le jeu de données de Centuri (mai 2021) est le premier grand jeu de données complexe qui ait été testé avec ces méthodes supervisées et non-supervisées. Les résultats produits sont très bons au vu du faible recul que l'équipe de Seaviews a sur l'emploi de ces méthodes.

6.3.3 Conclusions et prochains axes de travail

En bref Les différents résultats cartographiques de ce chapitre ont montré l'efficacité des méthodes d'apprentissage automatique. Elles ont réussi à produire des classifications proches de celle obtenue par un arbre de décision basé sur des interprétations des données acoustiques. Mais, ces méthodes automatiques se concentrent sur les données des MNT pour bâtir un modèle de prédiction, elles ignorent toutes les informations extérieures (vérités terrain, connaissances, autres données) dont l'opérateur tient pour sa part compte. Cela implique donc un peu d'édition manuelle pour ajouter les zones de digues et de galets tout comme pour la classification manuelle.

Avec ces méthodes, il est maintenant possible d'intégrer les images aériennes à la classification ; cela n'était pas le cas avec l'arbre de décision. De même, quand elles sont disponibles, les photogrammétries réalisées par Seaviews peuvent remplacer les images aériennes. Quant à la classification des images aériennes ou photogrammétries seules (sans utilisation des données acoustiques), une méthode reposant sur du traitement d'image classique a été ajoutée dans la suite logicielle. Ces résultats ont servi à créer des jeux d'apprentissage de données labélisées pour les méthodes à apprentissage automatique sur les données acoustiques.

Seule une partie du travail a été présentée et résumée dans ce chapitre d'applications. Au total, huit méthodes à apprentissage ont été intégrées dans les logiciels, en plus de la méthode destinée à classifier les images aériennes ou photogrammétries. Les possibilités de réutiliser un ancien modèle ou de créer un modèle d'après une classification déjà réalisée ont été ajoutées au logiciel. Celles-ci sont fonctionnelles mais n'ont encore été que peu exploitées. Les calculs de matrices de corrélations entre caractéristiques et de matrice de confusion entre deux classifications sont également disponibles. Concernant des post-traitements possibles, un filtre convolutif médian a

été ajouté pour lisser les contours des objets. Enfin, plusieurs aspects restent à être complétés ou explorés.

Pré-traitement des données Avant de poursuivre plus en avant avec les méthodes d'apprentissage automatique, la prochaine amélioration va consister à mieux préparer les données. En effet, quelle que soit la méthode (manuelle, supervisée et non-supervisée), les “défauts” dans les données sont reproduits dans les résultats de classification. Cela a été vu avec les images aériennes de la zone de Centuri. Deux options sont possibles, soit se tourner vers une autre base de données d'images sans ces défauts, soit réaliser des prétraitements pour en améliorer l'aspect. De même, le MNT de réflectivité n'a pas du tout été exploité dans cette étude car les traces de navigations visibles au nadir des différents transects perturbent énormément les données. Une correction post-acquisition est en cours de réalisation ; les lignes sont détectées avec une transformée de Hough puis modifiées avec une convolution adaptée sur les mailles du MNT.

La gestion des données manquantes Les différentes données utilisées dans cette étude (bathymétrie, rugosité, pente et image aérienne) couvrent l'ensemble de la zone. L'utilisation future du MNT de réflectivité dans la classification automatique soulève la question de la gestion des données dans les algorithmes à apprentissage. En effet, la donnée de réflectivité n'est présente que sur la partie levée au SMF. Comment peut-on faire sur le reste de la zone ? Certaines méthodes sont capables de gérer ces valeurs manquantes et d'autres non. Quand l'algorithme n'en est pas capable, la solution la plus rencontrée consiste à compléter ces valeurs par une valeur moyenne. Dans notre cas, cela n'est sûrement pas valable. En revanche, il semble plus judicieux d'établir au moins deux modèles : un valable sur toute la zone sans la réflectivité et un second avec la réflectivité sur la zone levée au SMF, puis de fusionner les deux résultats par un vote sur les prédictions.

La question de l'évaluation a été traitée ici en comparant la classification manuelle avec les classifications automatiques. Mais cela implique qu'une classification manuelle, longue et laborieuse, ait été faite. Dans le cas où le temps est compté pour éditer et livrer une carte à un client, il est certain qu'une classification manuelle ne sera pas faite en plus des essais obtenus avec les méthodes à apprentissage automatique. Des critères d'évaluation s'appuyant sur des vérités terrain partielles sont à explorer.

Conclusion

Seaviews réalise des cartes des habitats marins avec son sondeur acoustique multifaisceaux et sa suite de logiciels. Jusqu'à présent, la méthode de classification reposait uniquement sur une interprétation des données (acoustique et lidar) par un opérateur qui construisait un arbre de décision. Les récents développements de Seaviews dans cette direction ont porté sur le développement d'un indice de rugosité des fonds calculé depuis les données bathymétriques (Viala et al.; 2021) permettant de décrire les habitats dans l'arbre de décision. Mais la construction de l'arbre reste une étape souvent longue et complexe. Dans ce contexte, nous avons initié une transition vers des solutions plus objectives et efficaces à l'aide de méthodes issues de l'apprentissage automatique.

Plusieurs méthodes ont été intégrées et essayées. Trois d'entre elles - les forêts aléatoires d'arbres décisionnels, les Séparateurs à Vastes Marges et le classifieur bayésien normal - fournissent de très bons résultats en quelques minutes de calcul. En plus d'être efficaces, ces méthodes ont aussi l'avantage de ne pas avoir à être paramétrées par un opérateur, cela en fait des méthodes réellement automatiques. Ce point est important car le personnel de Seaviews est soit utilisateur (et non développeur) des logiciels soit profane vis-à-vis du fonctionnement de ces méthodes. Ainsi, comme avec la méthode par interprétation, il est attendu de l'opérateur qu'il connaisse les données acoustiques qu'il manipule ainsi que les habitats qu'il veut cartographier. En effet, la qualité du résultat dépend de l'attention que l'opérateur a fourni pour la sélection des zones d'apprentissage pour chaque classe - sans mélange de classe et avec des surfaces équivalentes pour chaque classe - ainsi que les caractéristiques qui servent aux calculs.

Ces méthodes font dorénavant participer les images aériennes aux classifications car celles-ci ne sont pas utilisables dans l'arbre de décision. Une première méthode pour permettre leur classification ou celles des photogrammétriques sans aucune donnée acoustique a aussi été mise au point et permet de s'affranchir d'une édition totalement manuelle. Cela doit permettre de réaliser une première classification d'après des données RGB qui servira ensuite de jeu d'apprentissage pour traiter des données acoustiques.

Conclusion générale

Au cours de ce manuscrit, deux thématiques - cartographie bathymétrique et classification des fonds - ont été considérées par l'intermédiaire de trois applications : le filtrage des sondes bathymétriques, l'extraction de la bathymétrie dans les images de la colonne d'eau et la cartographie des habitats benthiques à l'aide de données hétérogènes. À celles-ci, s'ajoute la cartographie des activités biologiques de la colonne d'eau dont la méthode et un premier résultat sont présentés en annexes.

L'ensemble des types de données utilisées à Seaviews (tableau 6.9) ont pu être manipulées au travers de ces quatre applications ; les données acquises par un sondeur multifaisceaux (de 1 à 3 dans ce tableau), celles calculées post-acquisition (4), des données d'opportunités (5 et 6), et des données acquises en compléments par Seaviews (7 et 8). Ce tableau omet les données de position, de célérimétrie, d'attitude systématiquement acquises et couplées de manière transparente aux autres données.

TABLE 6.9 – Bilan des types de données manipulées par chacune des applications traitées dans ce document.

Données	Applications			
	Tri des sondes aberrantes	Bathymétrie Colonne d'eau	Cartographie Habitats	Cartographie Poissons
1-Sondes bathymétriques	x	x	x	x
2-Réflectivité des fonds			x	
3-Images de la colonne d'eau		x		x
4-Pente, Rugosité			x	
5-Images aériennes			x	
6-Bathymétrie lidar			x	
7-Vidéo tractée			x	x
8-Photogrammétrie			x	

Ainsi, par essence même, les applications ne peuvent être abordées par un moyen identique d'automatisation. D'autant plus, qu'il s'agit de compléter et d'améliorer l'existant dans les logiciels développés par Seaviews et que l'apport se doit d'être "opérationnel" le plus rapidement possible et utilisable par tous. Enfin, il est question que l'automatisation fournisse une aide à la décision pour l'opérateur, en lui accélérant ou simplifiant le travail, plutôt qu'en le remplaçant car la validation d'un résultat reste sous la responsabilité d'un expert. Le tableau 6.10 retrace pour chacune des applications les types de méthodes employées avant et après ce travail. Les méthodes "avant" restent disponibles dans les logiciels.

Aujourd'hui, ces applications - à l'exception du *deep learning* sur les WCI - sont utilisables par les employés de Seaviews, en complément ou à la place des méthodes déjà existantes. La plus opérationnelle est certainement le tri des sondes aberrantes qui permet un gain de temps dans le premier filtrage des données. Mais la portée de cette méthode se limite aux logiciels de Seaviews développés en interne et non-diffusés : c'est l'application traitée avec la plus courte portée.

TABLE 6.10 – Bilan des types de méthodes choisies pour proposer traiter chacune des applications.

Applications	Types de méthodes					
	Manuelle	Interactive	Règles expertes	Semi-automatique	Machine learning	Deep Learning
Tri des sondes aberrantes		Avant	Après			
Bathymétrie colonne d'eau						Après
Cartographie Habitats	Avant				Après	
Cartographie Poisson	Avant	Avant	Après	Après		

A l'opposée, la déduction de l'information bathymétrique dans la WCI est la moins aboutie, une preuve de faisabilité plus complète et une démonstration d'intérêt sont encore à réaliser. Mais d'un point de vue portée, cette méthode pourrait sortir des murs de l'entreprise, devenir un produit à part, et s'appliquer à d'autres SMF (capables de collecter les données WCI). La cartographie des poissons a un objectif trouvé car en tant que nouveau produit elle répond à des besoins de connaissance et de surveillance pour la Directive-Cadre Stratégique pour le Milieu Marin (DCSMM), mais quelques points de fonctionnement sont à apporter pour son opérationnalité (suppression des derniers réglages manuels d'hyperparamètres). La cartographie des habitats benthiques n'est pas un nouveau produit mais l'amélioration d'un produit déjà vendu à Seaviews car de nouvelles données sont prises en compte et la méthode devient automatique et plus facile d'utilisation. Le graphique 6.12 résume la dualité opérationnelle et de rayonnement des applications.

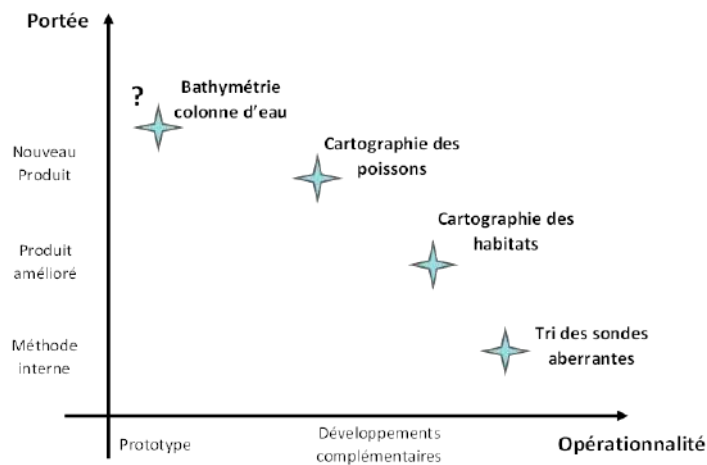


FIGURE 6.12 – Estimation de la portée des applications en fonction du degré de développement des méthodes.

L'Agence de l'Eau Rhône-Méditerranée-Corse soutient d'ores et déjà la poursuite des projets de cartographie des habitats et des poissons. Il est notamment prévu, pour ces deux projets, de continuer à améliorer les développements logiciels, approfondir les applications nouvelles, et de finir de mettre au point la caméra tractée utile aux vérités terrain autant pour les habitats que pour les bancs de poissons. L'utilisation des images de la colonne d'eau doit rester également un point central de la R&D à Seaviews. D'une part, pour ses applications originales - cartographie des poissons et bathymétrie déduite -, mais aussi car elle est l'un des rares organismes à acquérir systématiquement cette donnée si riche en informations.

Annexe A

Measuring fish activities as additional environmental data during a hydrographic survey with a multi-beam echo sounder

Marie LAMOURET ^{1,2,*}, Arnaud ABADIE ¹, Christophe VIALA ¹, Pierre BOISSERY ³, Nadège THIRION-MOREAU ²

1 Seaviews, 603 chemins des Séveriers Nord, 13600 La Ciotat.

2 Université de Toulon, Aix Marseille Université, CNRS, LIS, UMR 7020, Equipe SIIM, Toulon, France.

3 Agence de l'Eau Rhône Méditerranée Corse, 62 La Canebière, 13001 Marseille, France.

* lamouret@seaviews.fr

In processings of *OCEANS 2019 - Marseille, 2019*, doi : 10.1109/OCEANSE.2019.8867095.

Abstrat

The modern multi-beam echo sounders (MBES) are advanced instrumentation for active underwater acoustic surveys that can be boarded on oceanic vessels as well on light crafts. Although their versatility allows scientists to perform various environmental studies, their potential is seldom fully exploited. A single data acquisition cruise is not only able to display the seabed backscatter, but also provide an estimation of the fish activities from an underwater site thanks to water column imagery. This work is aiming at developing some (automatic) signal processing techniques to detect, analyse and classify objects observed in the water column with a focus on fish activities to provide fish accumulation and classification but also some comparative analyses along with the seafloor classification.

Key-words : Multi-beam echo sounder, Water column imaging, Underwater mapping, Fishery, Data sciences.

Introduction

Designed in the late 1970s, multi-beam echo sounders (MBES) perform depth measurements on numerous points along a line, the swath, perpendicular to the ship's heading. According to the angular distribution of the beams, it is possible to collect data along a swath five longer as the depth

under the ship. The developments of MBES over the last decades has led to significant improvement in hydrographic surveys with high resolution wide swaths that can cover a larger seabed area in a shorter time with a higher precision than a single-beam echo sounder. A bathymetry survey becomes thereby efficient and effective with a MBES for covering very large areas for all possible depth ranges. An example of a bathymetric survey rendering realised with a MBES is provided in the Fig. A.3, the zone of 0.8 km² is covered with only 15 km of navigation.

MBES are also able to record simultaneously other data among which backscattering images and water column imagery (Abadie and Viala; 2018). Various applications are derived from the backscatter imagery, since the seafloor natures (e.g. sand, rocky ground, muddy floor) can be determined using-specific backscatter signatures (Hasan et al.; 2014). Based on this theory, detection and inspection have been led on natural and artificial structures (e.g. shipwrecks, archaeology, mines, rocks, reefs) but also classification have been led on marine habitats (e.g. sediments, seagrass meadows, rocky habitats).

The WCI stand for the time series of the acoustic backscatters from all the elements present in each beam during the reception time. In other words, the seabed backscatter is a little part of the WCI, that is the strongest backscatter of a ping, which is truncated from the other previous lower backscatter and from the time series until the end of the receipting time. They also have a lot of applications in various fields among which are the biological application (including fishery) as well as geophysical and oceanographic sectors (Colbo et al.; 2014). In the case of fishery, even if the fish size is small as compared to the water column height, their flesh and swim bladder are strong acoustic scatter that allow to detect individuals and schools (Løvik and Hovern; 1979)(Chu; 2011). An example of a fish school is visible on the third WCI of the Fig. A.1. In the same way, fish schools form acoustic targets studied in the WCI. Their shape, size and morphology have been already considered and catalogued (Innangi et al.; 2016). Individual fish behaviours in a group have been identified (Holmin et al.; 2012) as well as global compartment of schools in front of a predator (Nøttestad et al.; 2002). Evaluation of the fish biomass is a current topic too. By definition, the biomass estimates the weight of fishes in a fixed volume using the target strength. This topic is complex to deal with because of (i) the fishes orientation that modifies the backscatter strength (Cutter and Demer; 2007); (ii) the fishes' avoidance moves on the path of the vessel that may lead to an under-estimation of their biomass (Misund and Aglen; 1992) and finally (iii) the harsh calibration of the backscattering strength with the quantitative fish stock assessment (Gurshin et al.; 2009).

Since the standard bathymetric MBES does not seem fully adapted to advanced fishery surveys, specific MBES have been designed like the ME70 of Simrad (Trenkel et al.; 2008) or the Seapix of iXblue (Mosca et al.; 2015). However, the WCI provided by the standard compact MBES such as the R2Sonic 2022 are sufficiently accurate to be processed for acoustic target detection. Keeping in mind that this standard MBES cannot fulfil the role of equipment specifically designed for fishery monitoring, they are still able to provide additional data on areas of fish accumulation and to classify the detected individuals and fish schools. In the end, it is thus possible to collect, in a single acquisition transect, the bathymetry, the backscatter imagery and the WCI to map the fish accumulations areas. The ultimate aim would be to correlate the detected marine organism with the seafloor nature.

In order to reach this goal, the main purpose of this work is to develop a target detection method dedicated to WCI based on signal and image processing techniques coupled with a statistical analysis of the results. Here, the main targeted application is the mapping of fish accumulation for environmental conservation. Another aim is to provide some preliminary tools for classification and to draw some interpretation of the biological activity. Furthermore, these new developments will allow us to improve ViewSMF and ViewMap (Viala; 2015b). These two specific software - developed by Seaviews - are dedicated to the processing and mapping of MBES data and to the seafloor classification. We intend to add a new functionality to exploit the WCI, namely, an automated target (fishes or whatever) detection tool.

This article is organised as follows. The first section is dedicated to the presentation of the area chosen to illustrate the method. The whole acquisition system set up is also described as well as the formatting techniques used to map the seabed morphology and nature. Then, in the

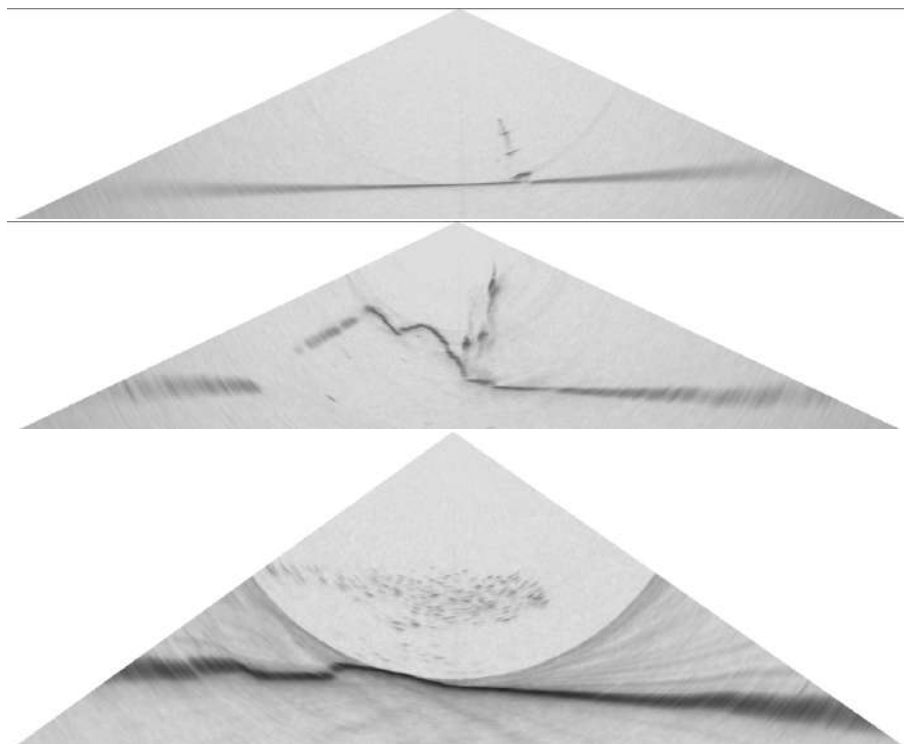


FIGURE A.1 – Examples of WCI with different targets. Seabed is represented as a dark horizontal line. Top : sailing yacht (wreck, mast and spreaders); middle : bubbles from diver(s); bottom : school of fishes.

second section, we detail the processing suggested to handle the water column data for fish activity applications. Example will be provided all along this article to illustrate some results. Finally, a conclusion will be drawn and a discussion will be delineated on methods, on interpretation limits and on further development and analyses.

A.1 Data acquisition and seabed mapping

A.1.1 Areas of interest

The chosen study site is called “Les Pierres”. It is located in the bay of La Ciotat (Provence-Alpes-Côte d’Azur, France), on the east of the island “Ile Verte”, see Fig. A.2. Well-known by the scuba divers, the 0.8 km² zone includes nine sub-zones (falls and reefs) which each presents an important marine fauna and various marine habitats (e.g. seagrass meadow, rocks with algal covers, coralligenous communities) between 10 and 60 m depth. The data sets were collected during spring and summer 2016 with the acquisition system described in the next subsection. During the surveys, bathymetric data, backscatter imagery and WCI were recorded.

A.1.2 Acquisition system

As already mentioned, the MBES used for this acquisition campaign is a R2SonicTM 2022. The R2Sonic swath is composed of 256 beams distributed over a maximal opening angle of 180°, but usually over a 140° opening angle in a more common usage. The acoustic signal, emitted with a chosen frequency of 450kHz during a very short pulse length (15μs), was adapted to map with a high resolution (0.9°×0.9° angular opening) in shallow water (less than 100 m) (R2Sonic; 2014). The MBES was coupled with an inertial navigation system (INS) AplanixTM I2NS and a full Real

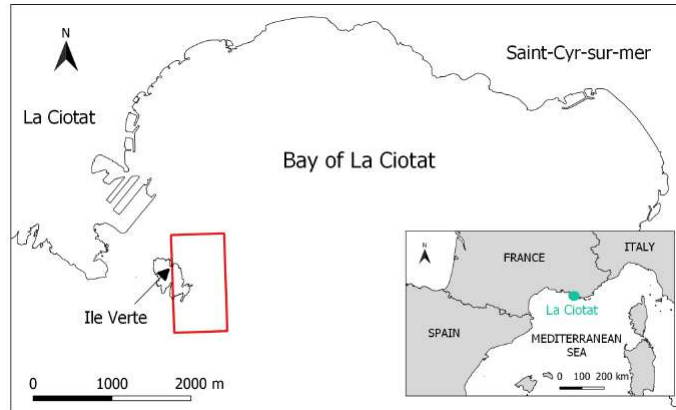


FIGURE A.2 – Location of the “Les Pierres” study site (red frame) on the French Mediterranean coast in the La Ciotat bay, near the “Ile Verte” island.

Time Kinematic (RTK) Global Navigation Satellite System (GNSS). The whole devices provided a positioning precision of 1.0 cm on XY axes and 1.5 cm vertically, a rolling and pitching precision of 0.015° and a precision for heading trajectories of 0.02° .

Two GPS antennas AeroAntenna Inc. AT1675-540 were used for satellite signal reception and navigation heading. Each offset of the whole system (sounder and positioning devices) was carefully measured at the vessel’s conception phase and regularly checked to avoid constant accuracy errors.

The navigation during acquisition was operated by a Raymarine ACU 200 autopilot synchronised with the RTK GNSS using ViewMap, a Geographic Information System (GIS) and navigation software developed by C. Viala (Viala; 2015b) which allows to trace and follow precise trajectories during acoustic data acquisition. The whole navigation system has an accuracy of 0.5 m to follow trajectories. Underwater sound velocity was constantly checked using a Valeport Ltd miniSVS sound velocity sensor mounted on the MBES. Additional underwater sound velocity profiles were performed with another miniSVS to detect the possible presence of a thermocline or fresh water layers impacting on the sound propagation.

The whole acquisition system was finally composed of the R2Sonic 2022 MBES, the Aplanix I2NS INS, the two GPS antennas, the two sound velocity sensors and computers for the navigation, configuration of the MBES and acquisition-storage data. All these devices were boarded on the Seaviews One, a small vessel (6 m long) specifically designed for hydrographic surveys with a central shaft for the sounder, the navigation unit and the sound velocity probe. This light craft allows to follow the shore as close as needed with its low draught.

A.1.3 Mapping the seabed morphology and nature

After each survey, the data collected by this set of acquisition systems are first checked then merged before being mapped.

GNSS positioning data were post-treated with the open source program package RTKLIB. By taking into account the accuracy of each sensor, i.e. the RTK GNSS, the inertial navigation system, the MBES, the sound velocity sensor and the sound velocity profiles leading to horizontal and vertical errors, we obtained a XY accuracy of ± 0.078 m, and a Z total accuracy calculated according to

$$Z_{\text{accuracy}} = 0.009 + 0.0066 \times D \quad (\text{A.1})$$

where Z_{accuracy} (in m) is the vertical error and D (in m) is the depth.

Using the ViewSMF software (Viala; 2015b), outliers and seabed miss-tracking points were also removed from the bathymetric data with the help of some filters.

Backscatter data were corrected in angle of incidence by using the sound velocity constant measurements and profiles. A time variable gain (TVG) was also run to avert noise from ping

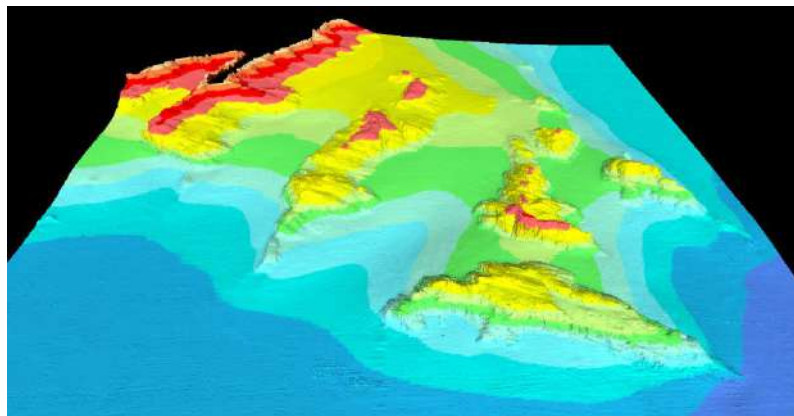


FIGURE A.3 – Bathymetric 3D survey of the south-west part of the “Ile Verte” at La Ciotat bay (France), represented with a colour gradient. Red areas have a depth lower than 15m, yellow ones are between 15m and 30m depth, seabeds in blue are at least 45m depth. The zone corresponds to a $500\text{m} \times 500\text{m}$ area.

overload. Snippets (i.e. time series of data samples per beam) were computed too, in order to reduce the noise generated by the beam spreading on the bottom and thus, to increase the overall resolution of backscatter images.

Once all these pre-processing performed, the bathymetric maps have been realised. A 3D rendering of the south part of the “Les Pierres” site is presented in Fig. A.3. The acoustic backscatters are also mapped (Fig. A.4A) and the roughness of the seabed is evaluated with the bathymetric soundings (Fig. A.4B). According to this three maps (bathymetry, sonar and roughness), an automated classification of the seabed is established (Abadie et al.; 2017). Ground truth data were collected by scuba diving on the each rocky reefs as well on seagrass meadows that had not a high acoustic signature on the rocky substrate. The results are provided on the Fig. A.4C.

A.2 Data processing

A.2.1 Introduction

As already said in the introduction, in the medium to long term, there are several reasons justifying a further exploitation of WCI. First, to endow the software that with develop, with new competitive tools for the detection and study of the fish activities from hydrographic MBES data. Secondly, to be able to realise environmental studies in addition to hydrographic surveys. And finally, to extend the tools devoted to the detection and the analysis of fish activities to any kind of WCI targets.

In order to demonstrate the feasibility and the interest of the aforementioned points, the software suite composed of ViewSMF and ViewMap was endowed with interfaces and extensions to handle this data set in 2017 (Viala and Abadie; 2017), yet, the initial treatments were performed purely manually. At that time, with our software, it was possible to display the map of fish accumulation zones and then to perform fish classification before correlating this information with the seabed. However, this manual method and more precisely the fish classification stage was time-consuming and had to be automated. To produce these maps - either with the manual or semi-automated method -, several successive stages are required :

- The reduction of the WCI to the part of interest (see subsection A.2.2),
- The detection of pings exhibiting any activity (A.2.3),
- The mapping of the fish accumulation zones (A.2.4),
- The manual (A.2.5) or semi-automated (A.2.6) classification of fishes and the mapping of the results.

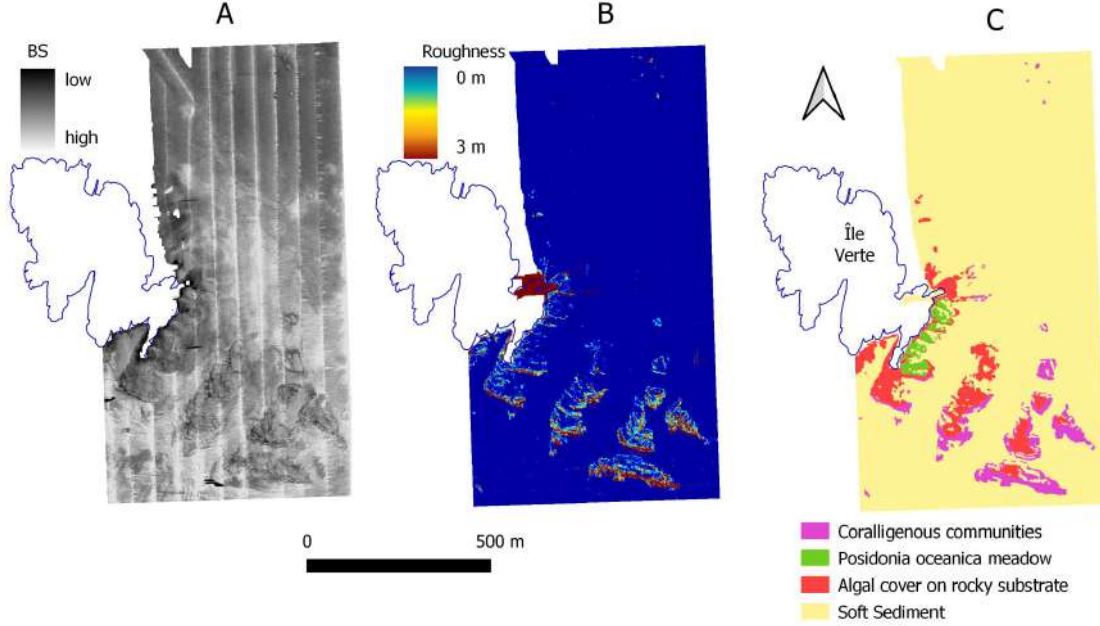


FIGURE A.4 – Data from the acquisition survey mapped. A : the acoustic backscatter map, “sonar image”, the strength values are a no-dimension intensity. B : the computed roughness, with metric values. C : classification of the seabed nature.

The Fig. A.5 is a graphical depiction of the whole processing chain applied to WCI when fish activity studies are considered.

A.2.2 Choice of the useful part of the WCI

The manual approach is not performed on the whole water column images : only their upper part is kept. It corresponds to the portion of the WCI which is not altered by the signal returned from the seafloor in interaction with the beam pattern-specific side-lobes. The limit between these two parts is fixed by the Minimum Slant Range (MSR) which stands for the shortest radial distance between the sonar transducer and the seafloor¹. It is easier to detect and analyse targets above the minimum slant range also known as the “validity arc”, due to a better Signal-to-Noise Ratio (SNR).

For the calculation of the minimum slant range, two variants have been implemented in the ViewSMF software :

- The Standard Minimum Slant Range (SMSR) in (A.2) is determined for every swath by the range of the minimum depth range point :

$$\text{SMSR}(p) = \underset{b \in [1, N_{bm}]}{\text{argmin}} D_{p,r}(b) \quad (\text{A.2})$$

where $\text{SMSR}(p)$ (in m) stands for the SMSR of the p -th ping, $D_{p,r}(b)$ (in m) are the bathymetric points in range distance r (in m) for the b -th beam of the p -th ping and N_{bm} is the beam number,

- The Gliding Minimum Slant Range (GMSR) in (A.3) is determined for every swath by the minimum of the SMSR along the previous-next i neighbour pings :

$$\text{GMSR}(p) = \underset{i \in [\pm 10]}{\text{argmin}} \text{SMSR}(p - i) \quad (\text{A.3})$$

1. For the record, the depth is a measure of the *vertical* distance below a system reference water level, whereas the range is a measure of the distance of the seabed to the MBES by taking into account the incidence angle.

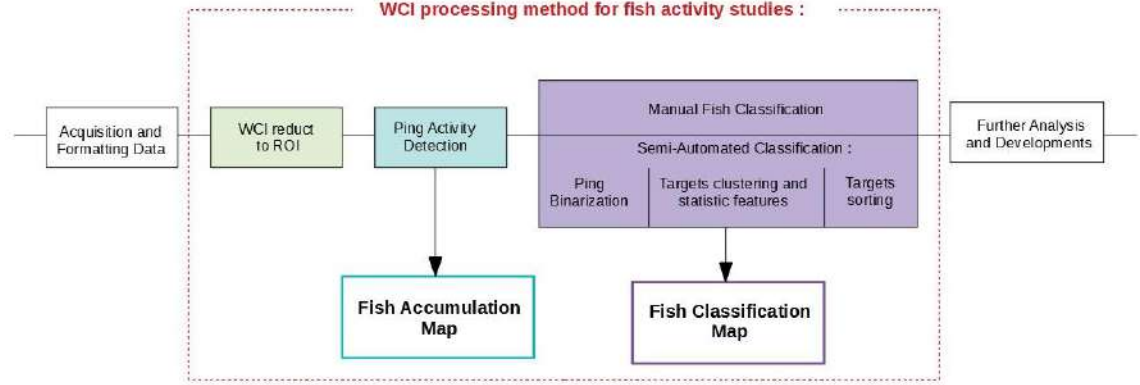


FIGURE A.5 – Block diagram of the global data processing chain applied to MBES. The first block “Acquisition and formatting data” refers to the section A.1. The block depicted by a dotted red line represents the successive stages needed to exploit the WCI, whereas the blocks behind arrows are the two produced maps. *ROI : region of interest.

where $GMSR(p)$ (in m) stands for the GMSR of the p -th ping.

It turns out that these two MSR are meaningful. Most of the time, the SMSR is sufficient and well determined. But, because of the side-lobes effect, a persistence phenomenon due to the bottom shape can appear in the WCI. Generally, it does not provoke any bottom miss-tracking, but some bottom persistent interferences overflow in the validity arc which can lead to a miss-interpretation of targets. That is why the GMSR, by taking into account the previous-next pings, is more robust than the MSR with regards to this possible persistence problem. An example of these two variants of the MSR in the case of a persistent bottom is provided on the Fig. A.6.

A.2.3 Ping activity detection

In order to evaluate the biological activity in a WCI, two indexes have been introduced : the average biomass (denoted by B_{Av}) and the maximal biomass (denoted by B_{Max}). Here, the “biomass” term, in an acoustic meaning, refers to the whole objects and living organisms that sign in a WCI, not the weight of the living organisms. The average biomass given in (A.4) estimates the average value of all the impulse intensity responses (in dB) within the validity arc :

$$B_{Av}(p) = \frac{1}{N_{bm} \times N_{sp}} \sum_{i=1}^{N_{bm}} \sum_{j=1}^{N_{sp}} BS_{i,j}(p), \quad (A.4)$$

where $B_{Av}(p)$ (in dB) is the average biomass for the p -th ping, N_{bm} and N_{sp} are respectively the beam number and the samples number in a beam before the MSR, $BS_{i,j}(p)$ is the validity arc in backscatter strength value.

The value B_{Av} should give an account of all the suspended objects (e.g. fish, zoo-plankton) while the maximal biomass gives only an account of the strongest target signatures in the same zone. The maximal biomass is estimated by storing the n strongest intensity (in dB) samples among all the samples within the validity arc (the hyper parameter n is typically fixed at 100 in (A.5)). From the max-biomass, it is possible to compute the maximum, the minimum and the average among these n values :

$$B_{Max}(p) = \underset{BS' \subset BS, \text{card}(BS')=n}{\text{argmax}} \sum_{bs \in BS'} |bs| \quad (A.5)$$

where $B_{Max}(p)$ (in dB) is the maximal biomass for the p -th ping. The set BS is the backscatter strength values of the validity arc, i.e. $BS = \{BS_{i,j}(p), \forall i \in \{1, \dots, N_{bm}\}, \forall j \in \{1, \dots, N_{sp}\}\}$. The

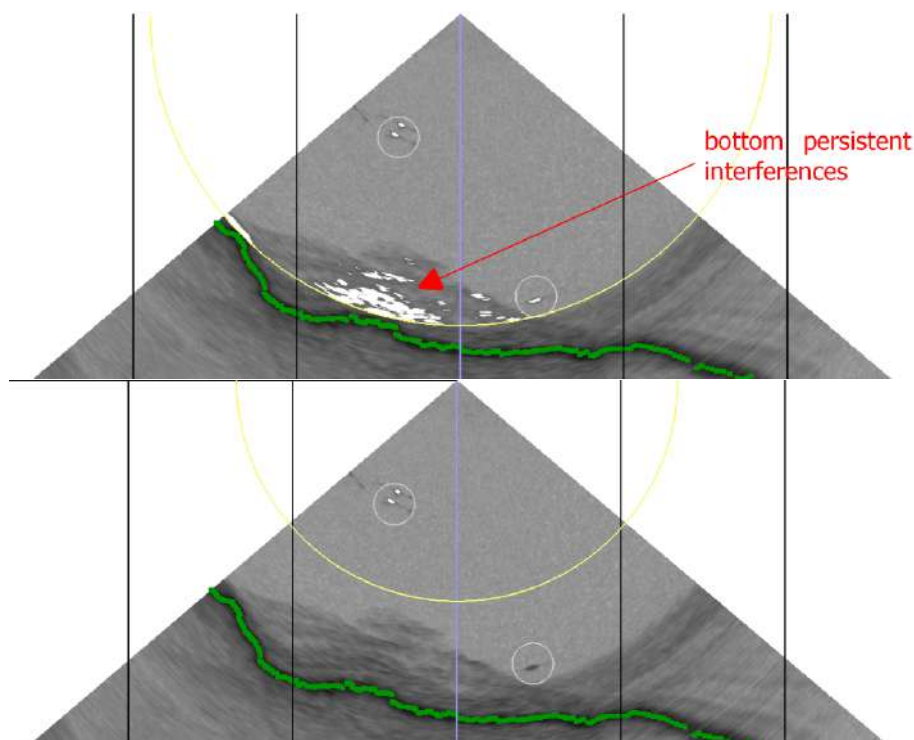


FIGURE A.6 – An example of water column ping including fishes (indicated by the two grey circles) and some bottom persistent interference. The bathymetric points are marked by green points. The two different minimum slant ranges (symbolised by a yellow arc) are drawn (calculation based on the SMSR (top) or the GMSR (bottom)). The WCI had been thresholded at -70 dB above the MSR. Detected targets appear in white (fishes and a part of the interference).

subset of the n higher backscatter strength values of BS is denoted by BS' . Its n elements are stored in a vector denoted by bs .

The biomass values from the ping in Fig. A.6 are reported in the Table A.1.

TABLE A.1 – Comparison of the B_{Av} and B_{Max} values versus the MSR variant. Values obtained on the WCI given in Fig. A.6

Biomass index	MSR : Minimum Slant Range	
	<i>Standard MSR</i>	<i>Gliding MSR</i>
B_{Av} (dB)	-84.3	-87.9
B_{Max} - min (dB)	-60.4	-66.9
B_{Max} - avg (dB)	-58.3	-61.7
B_{Max} - max (dB)	-53.2	-55.4

The difference between the MSR led to some differences in the biomass values. Especially in the B_{Av} because the large persistence near the bottom had higher backscatter intensities than the background noises, that skewed therefore the average, in the SMSR, which should be closer to the background noise value. The B_{Max} index was less affected by the bottom persistence because its value was computed from the n strongest samples that can come either from the fishes or the persistence.

Hence, since at least one target in a WCI is enough to make B_{Max} really higher than B_{Av} , it is quite easy to distinguish between pings exhibiting at least one target and those which do

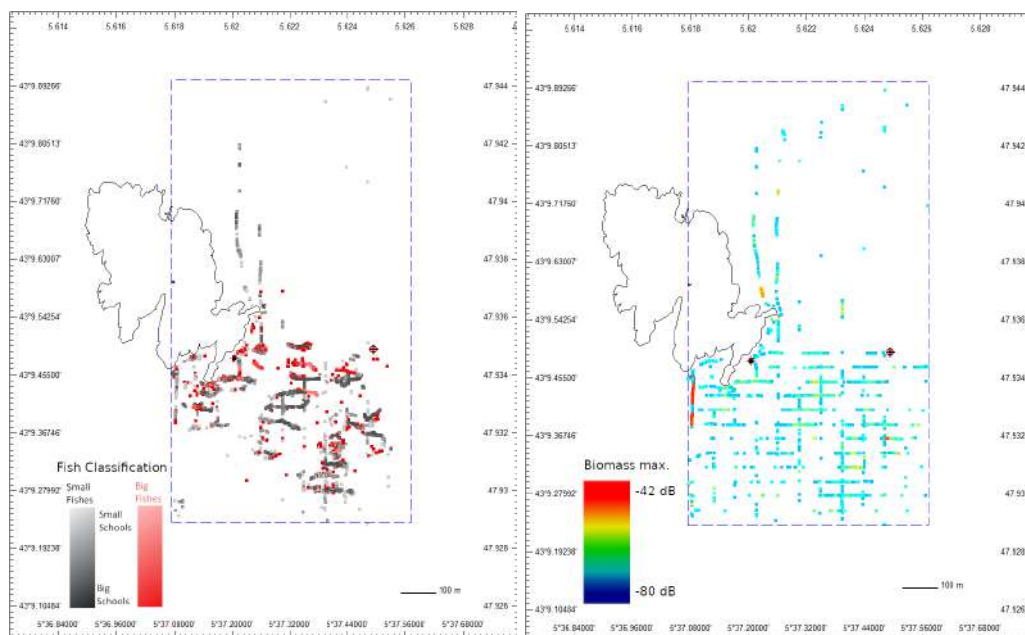


FIGURE A.7 – Left : fish accumulation areas on the site “Les Pierres” represented by the value of the maximal biomass of every ping (summer 2016 survey). Right : fish classification on the site “Les Pierres” categorising fishes detected during the summer 2016 survey in four classes (big or small individuals in large or small fish schools).

not. When there is no target in the WCI, the minimum of B_{Max} is much more close to B_{Av} . So, it is easy to fix a threshold value : 15 dB above B_{Av} seems a good choice. Here, in Fig.A.6, the threshold has been set at -70 dB, highlighting the targets that are present and sign with a higher value computed as the average value of B_{Max} .

A.2.4 Fish accumulation mapping

Once the biomass indexes are computed for each ping of the survey, this information can be charted. Each ping is represented by an intensity colour point on a map and positioned according to the vessel path. Since most paths do not exhibit fish activities, they are not drawn to better emphasise the pings with fish echoes and thus to better delimit the different fish accumulation areas. The fish accumulation map of the summer 2016 survey is given on the left side of Fig. A.7. It tends to show that fishes are visibly concentrated on the south of the site “Les Pierres” where reefs are located.

A.2.5 Manual fish classification

Since WCI provided sufficient details about individuals and fish schools, a classification was attempted. Actually, from one ping to another, fishes swam at different depths, had various sizes and evolved within more or less large schools. The manual classification was established according to these three parameters. The software ViewSMF, dedicated to the MBES data processing, was endowed with an interface to quickly browse the pings including any detected fish and then to assign the targets to a class. Targets were spotted in the frame by circling them (as in the Fig. A.6). Four sizes of circle were available for classifying the size of the schools but also two colours for classifying the fishes in little or big categories. The position of the circle in the frame automatically classified the targets according to their depth - surface, column and near seabed. An example of a result of the manual classification is visible on the right of Fig. A.7.

This first manual method for the mapping of fish accumulation areas and fish classification

proved that studying fish activities in very shallow water with a hydrographic MBES is utterly practicable. However, this manual classification was really fastidious and time-consuming and had to be automated prior to any other improvement of the rest of the method. This is described in the next paragraphs.

A.2.6 Semi-automated fish classification

The automation of any method generally requires more pre-processing to be really efficient. Thus, a semi or completely automated classification still could not be done from the data at the end of the ping activity detection step A.2.3. In fact, this stage had to take a binary decision about the WCI content - empty or not empty (i.e. containing fish target or whatever) - but was not yet designed to detect all the targets. In this subsection, the first step is to binarize properly the WCI with any target, then to cluster the different target(s) before computing for each of them, the statistical features that were used to sort the fishes as previously described. At the end of this semi-automated classification, the results should be roughly equivalent (or better since less subject to assessment errors).

A.2.6.1 Ping binarization

After considering the whole data set, we chose to manually threshold the WCI. In average, the values of the background noise ranged between -90 dB and -100 dB. The strongest backscatter objects have samples which start to be detected at -65 dB, whereas from -80 dB, nearly all the samples of a target are effectively detected. A too high value threshold (like -70 dB) does not seem to be adapted since the targets are still incomplete. On the contrary, a too low value threshold (like -90 dB) might not be adapted either since it may happen that close targets are merged in one single target. Finally, a fixed threshold with a value between the aforementioned cited values is often a good compromise, even if some false alarms may appear and considered as targets like the fishes. An example of different values for the threshold is provided in Fig. A.8.

Thus, at that time, a range of several threshold values adapted to the current survey was evaluated by hand during the data discovery and then fixed as an “hyperparameter” for the rest of the method.

A.2.6.2 Targets clustering and statistical features

Once the WCI thresholded, it is possible to agglomerate the samples into cluster corresponding to the detected targets (even the false alarms). The method simply browses the WCI values until it finds a sample above the threshold value and then search the neighbouring samples which belong to the same cluster (and also the target). For further treatments, only the coordinates - depth range and angle, corresponding respectively to row and column of a classical image - of the clustered samples are stored. From these clusters and their coordinates, some statistical features are computed. The aim of these features is then to sort the targets into two classes : i) detection of interest (like fishes), ii) detection that do not interest us in this application (e.g. wake, noise, scuba diver bubbles) and false alarms. The chosen statistical features are divided into three components : the depth range, the angle of the beam and the backscatter intensity. For each components of every targets, the minimum and maximum value are stored but also the mean and the standard deviation. Thanks to these statistical features, it is then possible to compute the centroid and the size of every targets and then to establish sorting rules.

A.2.6.3 Sorting the targets

Since it is a semi-automated classification, the rules to sort the targets were fixed heuristically. The knowledge of the survey data can allow to express the distinction between targets that interests us - fishes - and those which do not. As an example, fishes can be recognised by their small amount of samples and a round shape (angle and range sprawls are close) when false alarms correspond to very few samples and noise do not exhibit a constant form (persistence : very large patches,

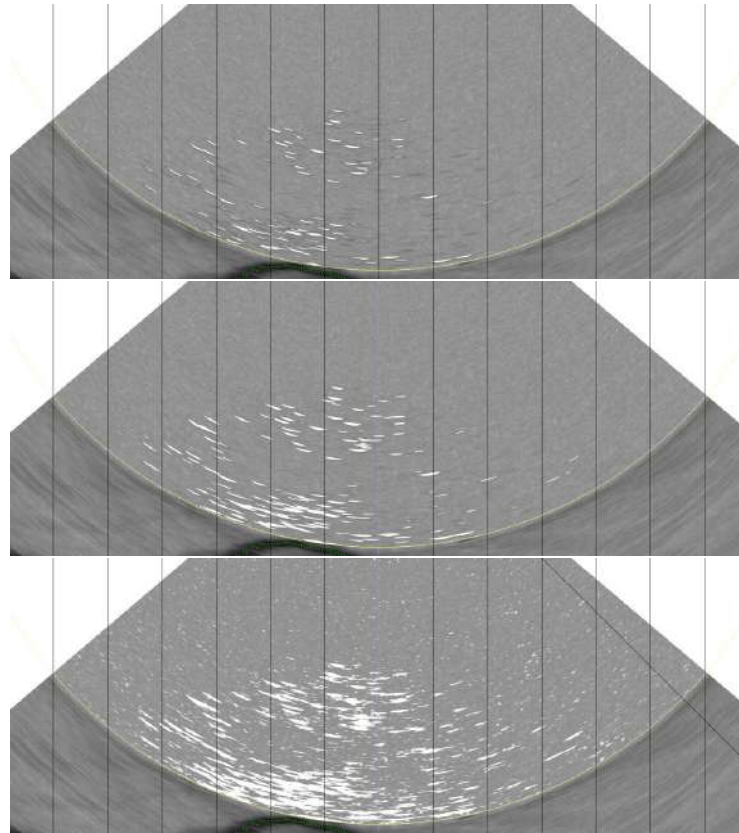


FIGURE A.8 – WCI exhibiting fish school for different values of the chosen threshold : 70 dB (top), 80 dB (middle), 85 dB (bottom).

arc side lobes : an angle sprawl bigger than this in range). The table [A.2](#) of sorting rules for different targets is given below as examples. Afterwards, the fishes can be classified according to different parameters : i) their position in the water column thanks to the computed centroid (range position), ii) their size thanks to the quantity of samples and finally according to iii) the size of the schools taking into account the number of fishes detected in the WCI.

A.3 Discussion

Our main goal was to give an estimation of the fish activity - accumulation areas and classification - during a survey realised with a hydrographic MBES. A manual method was developed in 2017 and implemented in our software through GUIs. It allowed to obtain, as results, the maps of fish accumulation and classification and thereby to assert the feasibility of the proposed approach.

On the WCI collected by a well-resolved MBES in very shallow water, it is currently possible to visually distinguish the fish echoes from other targets echoes (e.g. bubbles, mooring line, wake, bottom persistence, etc.) ([Schneider von Deimling and Weinrebe; 2014](#)). Moreover, the fish schools can be distinguished from the individuals and then, the individuals can be classified according to their size. This evaluation remains therefore essentially qualitative since fish size can not be precisely estimated due to i) the WCI resolution and ii) the fact that the fish signatures do not necessarily reflect the whole fish since it is essentially the swim bladders that sign.

The suggested approach developed in this work provides the fish biomass map (in an acoustical meaning) of the studied areas and enables to identify the sub-areas where fishes are concentrated. Although this type of maps cannot achieve the estimation of the quantity of all the present fishes - i.e. only fishes in the water column, rocky reef fishes can not be detected - it remains sufficient to

*ANNEXE A. MEASURING FISH ACTIVITIES AS ADDITIONAL ENVIRONMENTAL DATA
DURING A HYDROGRAPHIC SURVEY WITH A MULTI-BEAM ECHO SOUNDER*

TABLE A.2 – Various WCI targets discriminated according to their coordinates and intensity.

Features	WCI Targets				
	False alarm	Arc side-lobes	Fish	Vessel Wake	Bottom persistence
Number of samples	very few	-	fixed range of few	a large number	a large number
Centroid of samples	-	-	-	near the surface	near the bottom
Backscatter intensity	very low	low in the edges, higher in center	-	-	-
Angles of samples	very concentrated	very sprawling	concentred	-	-
Ranges of samples	very concentrated	concentred	concentred	-	-

locate the main fish pelagic living zones. The cross-analysis of bathymetry, marine habitats and fish biomass maps have led to retrieve some hardly surprising facts. Over the last decades, numerous studies have highlighted that fish accumulations are more likely located on sites with complex seascapes (such as rocky reefs and seagrass meadows) than on flat deeper zones (Calò et al.; 2013) (Charbonnel et al.; 2002). Likewise, the surveys, realised at different seasons in a year, allow to correlate the water temperature with the size of the fishes and their schools (Palomera et al.; 2007) (Sabatés et al.; 2006). In summer, with a thick thermocline, there were larger schools with big fishes than in winter, where little fishes seem to swim in small schools. This remark allowed to assert the second purpose of this study which was to correlate the fish activities with their environment (seabed morphology and nature).

Our results also suggest that it is possible to compare the areas with each other. For examples, the fish abundance within or at the edge of a protected marine area can be mapped. This type of study had been realised in 1995 not far from our study site (Carry-le-Rouet, France) by monitoring two types of species (Harmelin et al.; 1995). In this case, acoustical measures could complete visual evaluations. Likewise, it is possible to compare the areas over time. For examples, it would consist of realising a seasonal monitoring to attest that fishes have settled on artificial reefs (Clark and Edwards; 1994) or to compare their attractiveness against the one of natural reefs (Carr and Hixon; 1997).

At that point, the analysis process described previously does not allow any precise abundance evaluation or species identification of fish schools without additional information. It offers benefits nevertheless. As an additional data of hydrographic surveys, it remains a way to quickly acquire fishery data for the first time in a wide area. Once the fish accumulation have been drawn, it can lead to more specific surveys where traditional fish counting methods can be used. In this way, measuring fish accumulation do not replace specific fishery monitoring techniques, but provides additional spatial data through the analysis of biological indexes. For instance, if fish assemblage tracking and counting is now possible with cameras which can be fixed and active for a long time, they are efficient only if they were set up in the interesting places (Boom et al.; 2012) that our method can highlight.

Gaining in precision for abundance evaluation and species identification of fishes is conceivable in the future. It will not be possible without additional information however. For the first task, it requires a calibration of the MBES with targets of known size. This is made possible with underwater cameras or by scuba divers. For the second task, it requires the knowledge of locals – divers, fishermen, environmental administrators – who are familiar with the main local marine species and their habits.

This study has a great potential for current and coming applications as described above, but also a technical potential since the method still remains mostly manual. More Specifically, the manual classification was fastidious and time-consuming. For this reasons, more autonomous tech-

niques were required and thus a semi-automatic classification was presented. This new contribution brought two major advantages : (i) when the manual classification could take day(s) to be properly done, the new one takes a few minutes at worst ; (ii) this allowed to deepen the exploitation of the WCI and prepare the data to be used with more autonomous techniques (like machine learning).

Before going further in the WCI exploitation, other steps of the process need a deeper study. Among others, the number of pings considered for the gliding-MSR and the number of strongest samples kept to compute the B_{Max} value were this time fixed by hand. For the gliding-MSR, the number of ping had to be adapted to the situation since the bottom persistence can liner more or less longer according to the intensity of the original bottom. About the number of samples kept to the maximal biomass ($n=100$), n was fixed with an absolute value, but it should be adapted to the size of the WCI above the minimum slant range which can vary from one ping to another. Finally, the thresholding pre-processing before the clustering should be automatically computed prior to the gliding-SMR or the maximal biomass modification. Actually, with a well-chosen threshold, the ping activity detection and the ping binarization could be done in one stage for a WCI being always truncating with the standard-MSR.

The main difficulties to properly automatically threshold the WCI, from survey to survey and from ping to ping, are raised by the background noises due to the inhomogeneous and randomly distributed sea clutter². Since the noises do not follow a Gaussian distribution and vary with the depth, classical methods such as the Otsu binarisation (adaptive image thresholding method) or global thresholding techniques were not really suitable. Thus, local thresholding techniques may be considered such like adaptive methods with a median kernel, or the use of a wavelet method as proposed in (Karthikeyan and Chandrasekar; 2011), and else combining image and learning methods like in (Hofmann et al.; 2012).

Conclusion

The innovative WCI technique of analysis developed in this work, and its future improvements, have the potential to provide in a near future exhaustive surveys of the seafloor, as well as of the water column, thanks to a single sensor : the multibeam echosounder. Combining high resolution maps of marine habitats with fish assemblages are mandatory to fall in the scope of the global effort to investigate, monitor and evaluate the ecological status of fundamental coastal marine ecosystems. This is especially true in the framework of the European Marine Strategy Framework Directive (MSFD) that aims to apply an ecosystems-based approach in the regulation and management of the marine environment, marine natural resources and marine ecological services. In the Mediterranean Sea, the MSFD targets, among others, the fundamental ecosystems based on *Posidonia oceanica* meadows as well as those relying on coralligenous communities. As demonstrated in this study, compact MBES have the potential to provide a complete spatial evaluation of these habitats, even when they are mixed with other habitats such as rocky and sedimentary substrates. Overall, such an approach allows to decrease the cost of environmental underwater studies by increasing their efficiency in connection with the size of the area exhaustively investigated.

Acknowledgment

This work has been made possible thanks to the Agence de l'Eau RMC which financed the hydrographic surveys in 2016 at La Ciotat and the first research works on the underwater biological activity. Special thanks as well to the Association Nationale de la Recherche et de la Technologie (ANRT) for the "Convention industrielle de formation par la recherche (CIFRE)" service allowing the funding of a tripartite PhD thesis between a PhD student, a society and a research laboratory.

2. The sea clutter is a term used for unwanted echoes in electronic systems, particularly in reference to radars. This term can be extended to sonars since the sea is a noisy environment due to the characteristics of underwater acoustic and to various noise origins (i.e. marine traffic, rain and surface waves, marine organisms, thermal noise) (Lurton; 1998)

Annexe B

Coupling the map of marine habitats and fish accumulation zones : a three dimensional spatial approach for the management of halieutic resources.

Marie LAMOURET ^{1,2}, Pierre BOISSERY ³, Christophe VIALA ¹,
Nadège THIRION-MOREAU ², Arnaud ABADIE ^{1,*}

1 Seaviews, 603 chemins des Séveriers Nord, 13600 La Ciotat.

2 Université de Toulon, Aix Marseille Université, CNRS, LIS, UMR 7020, Equipe SIIM, Toulon, France.

3 Agence de l'Eau Rhône Méditerranée Corse, 62 La Canebière, 13001 Marseille, France.

* abadie@seaviews.fr

Published in *Vie et Milieu - Life and Environment* 70(3-4), 2020.

Abstract

The spatialisation of halieutic data is an essential element to define and create efficient protected and managed areas. Moreover, the distribution of fish schools is not homogeneous in the water column and is strongly linked with marine habitats. It is thus necessary to develop techniques allowing a spatial evaluation of halieutic resources. Multibeam echo sounders (MBES) provide acoustic data of the seafloor and the water column with a high accuracy and resolution. A single acquisition gives the bathymetry, a backscatter mosaic of the sea bottom and an acoustic imagery of the water column. The bathymetric data processing highlights the seafloor rugosity using several metric indices. A semi-automated classification including the depth, the rugosity indices and the backscatter values provides maps of marine habitats which are finally validated with ground truth. Data from the water column are analysed using an algorithm that detects acoustic targets corresponding to fishes. A georeferenced scatter graph of fish schools is thus automatically created. The 3D model of the seafloor obtained from the bathymetry is textured with the map of marine habitats. Points corresponding to fish detection are then added on the 3D model to provide a complete map. Through this process managers can access to a clear visualization of fish accumulations and the key marine habitats within their areas of interest.

Introduction

Marine ecosystems are among the most rich and complex biologic systems of our planet but remain difficult to study when compared with terrestrial habitats (Appeltans et al.; 2012). To integrate and assess marine ecosystems' complexity in governmental efforts to protect them against anthropogenic influences, an ecosystem-based management theory was produced (Slocombe; 1993). More and more applications based on this approach are currently developed to take into account the resilience and robustness of marine systems (Curtin and Prellezo; 2010). These approaches rely on the understanding of ecosystems functioning by associating marine species in functional groups according to their ecosystemic roles (Hooper et al.; 2002). They are especially suitable to respond to the European directives, such as the Marine Strategy Framework Directive (MSFD), aiming to reach a good ecological status of marine areas.

Due to complexity and the difficult access to the marine environment, most of the survey techniques classically used are based on discrete samplings and observations focusing on a single or few functional compartments. A good example of the paradigm evolution concerning the survey of marine ecosystems is the study of *Posidonia oceanica* (L.) Delile seagrass meadows that previously focused on the plant morphology (Pergent-Martini et al.; 2005) while, nowadays, ecosystem-based approaches are developed (Personnic et al.; 2014). Although this shift in paradigm provides a better evaluation of ecosystems status, the data on which it relies are still discrete and random measurements upscaled to large areas. Complementary spatial data are thus required to fulfil an effective ecosystem-based management of the marine environment.

The most effective tool to acquire spatial information on the seafloor and the water column above – with a high resolution and positioning accuracy – is currently the multibeam echo sounder (MBES) (Abadie and Viala; 2018). This type of acoustic probe is able to provide simultaneously bathymetric data, backscatter images and the water column imagery (WCI) on a large swath (increasing with the depth). Recent signal processing methods for bathymetric data allow to generate maps of marine habitats using various rugosity indices (Abadie et al.; 2017). Likewise, innovative algorithms are able to detect various acoustic targets on the WCI among which the fish schools and the individuals that are composing them (Lamouret et al.; 2019).

In order to pave the way for a spatial approach of the ecosystem-base management, we investigated the capacity of MBES to provide exploitable two and three-dimensional information on several functional compartments of key marine ecosystems of the Mediterranean Sea. With this main target in mind, we studied the inter-seasonal and inter-annual characteristics of fish accumulations on various habitats. We also experimented different environmental indicators in an attempt to describe the link between marine habitats' features and fish distribution.

B.1 Materials and Methods

Study site and data acquisition

This study took place in the Bay of La Ciotat in the south of France off the Ile Verte (Fig. B.1) on a site renowned for the richness of its habitats and biodiversity. The site covers an area of 0.83 km². An exhaustive acoustic data acquisition was performed two times in 2016 in June and August to take into account the increase of the water temperature linked with fish observations. These two acquisitions were realised in the same conditions : a two hours work done in the morning to collect the data along the same lines north-south oriented. Another acoustic dataset was obtained in June 2019, allowing an inter-annual comparison.

Acoustic data were acquired by using a R2Sonic 2022 MBES fixed on the hull of a 6 m long survey boat. Position and attitude were recorded by an Applanix I2NS, an inertial system equipped with a RTK GNSS positioning device providing 0.015° roll/pitch precision as well as a horizontal accuracy of 1 cm and a vertical one of 1.5 cm. Acoustic data were acquired at a frequency of 450 kHz with an individual beam width of 0.9° x 0.9° for a maximum swath sector of 160° and 1 024 soundings per swath. Transects were defined prior to the data acquisition and the navigation was operated by a Raymarine ACU 200 autopilot synchronised with the RTK GNSS using ViewMap, a

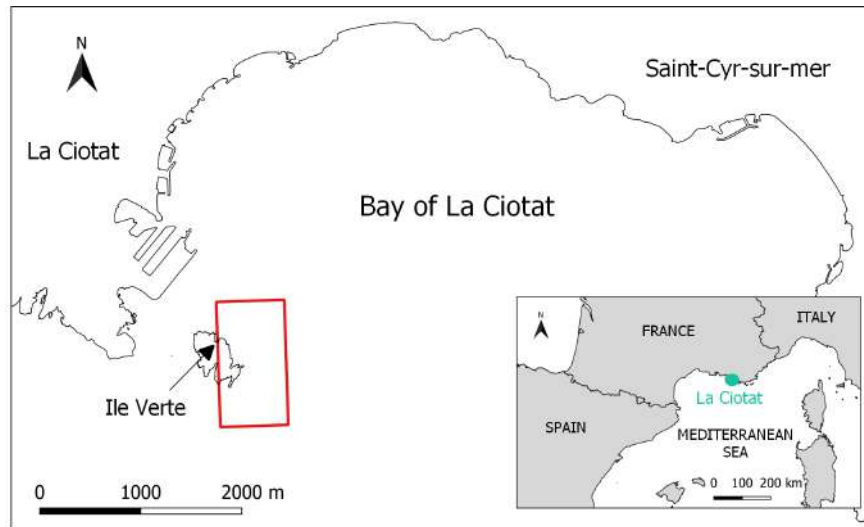


FIGURE B.1 – Study site (red frame) in the Bay of La Ciotat.

Geographic Information System (GIS) and navigation software developed by (Viala; 2015a). The underwater sound velocity was constantly checked using a Valeport Ltd miniSVS sound velocity sensor mounted on the MBES. Additional underwater sound velocity profiles were performed with another miniSVS to detect the possible presence of a thermocline or fresh water layers impacting the sound propagation. Water temperature profiles were computed from the sound velocity data. Ground truth data were performed by scuba diving to validate the seafloor classification.

Acoustic data processing and generation of habitat maps

R2Sonic 2022 bathymetric soundings were processed using the ViewSMF computer program developed by (Viala; 2015a) for the visualisation and processing (automatic or manual) of MBES acoustic data and metadata. False echoes were removed using filters to isolate one or several soundings. A rugosity index, named Bathymetric Automated Treatment for the Classification of the Seafloor (BATCLAS), is then computed from bathymetric soundings to highlight the underwater landscape according to the method developed by Abadie et al. (2017). The noise on the backscatter imagery was reduced using a time variable gain and snippets. A digital elevation model (DEM) encompassing a bathymetric map, the treated backscatter imagery and the BATCLAS index was generated. Finally, data from the DEMs and ground truthing were computed in ViewMap using a decision tree to classify marine habitats and build polygons exploitable in Geographical Information System (GIS) for further analysis. The final maps takes the shape of 2D and 3D maps of marine habitats.

Water Column Processing and Environmental indicators

The WCI was analysed following the technique developed by Lamouret et al. (2019). This method utilises an algorithm that automatically detect and identify acoustic targets corresponding to fishes on the WCI. For each detection, the localisation, the dimensions and the energy are computed and stored in computer files. The halieutic data set is finally exported under the shape of a scatter plot for 2D and 3D analyses, as well as for investigating the relationships between fish distribution and marine habitats.

In order to compare the density between the different dataset, the fish density is computed from the scatter plot for the whole area and for several sub-areas of interest of the marine habitat map. The density is given in fishes/m², corresponding to the number of fishes in a column of 1 m² and of height given by the bathymetry. The density by subzone is then easily comparable through

times on 2D maps. The study area was divided in three main sub-areas : (1) the Ile Verte walls, (2) the rocky reefs, (3) the sedimentary plains. Moreover, the rocky reefs were subdivided one by one.

B.2 Results

Depths varied from a few meters (<10m) on the shallowest cost of the Ile Verte to 66m in the south-east corner (Fig. B.2A). In the extension of the island towards the south-east an uneven seascape was clearly visible, composed of vertical walls and plateaus. The backscatter imagery (Fig. B.2B) highlights the relief seen on bathymetric data. Apart from these irregularities, the seabed was even with two ranges of backscatter intensity : a high value on the northern part and a lower one in the south, indicating two types of sediments with contrasted granularities. The unevenness were highlighted again and well delimited on the rugosity index map (Fig. B.2C). The rest of the seabed appeared smooth, except for small spots in the north east of the zone corresponding to known ship wrecks.

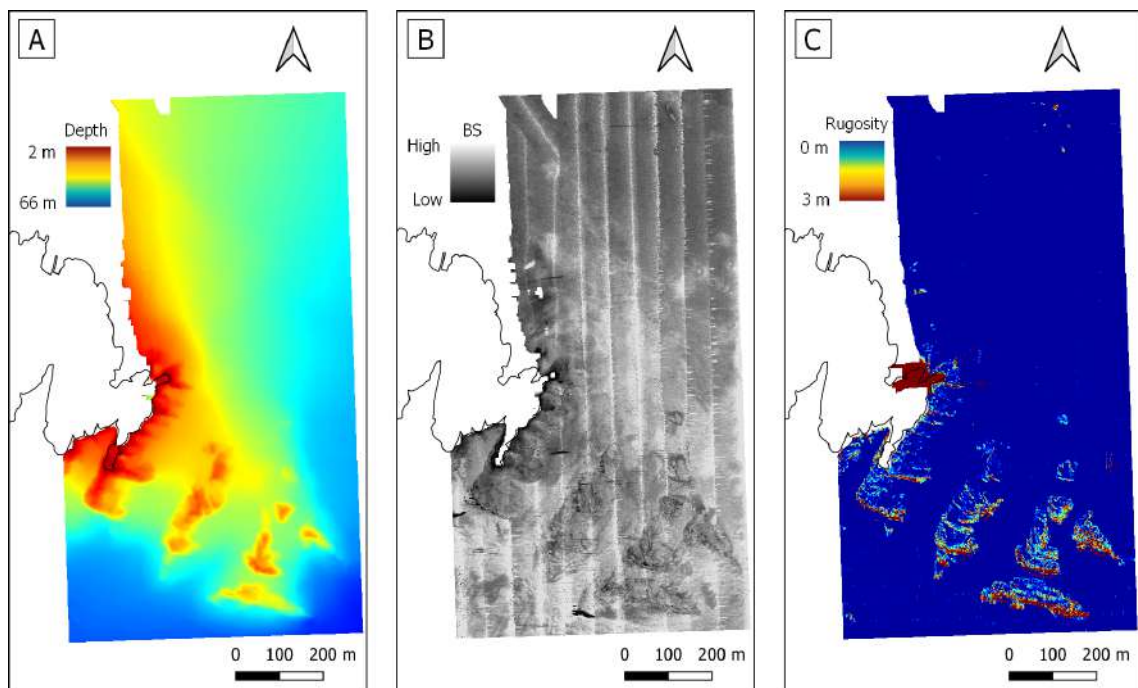


FIGURE B.2 – MBES data products. **A**, Bathymetry ; **B** Backscatter imagery, **C**, Rugosity.

Ground truthing confirmed the existence of the spots that were three very dilapidated wrecks - a wooden trawler and two sailing ships - playing the role of artificial reefs on the sandy plain. Ground truths also established the main habitats, i.e. rocky substrate with algal cover (39 151 m²), *Posidonia oceanica* (L.) Delile meadows (7 225 m²), coralligenous communities (27 838 m²) and soft sediments (624 502 m²; Table B.1). The *P. oceanica* meadows were not found on sediments, but rather on hard substrates (Fig. B.3). Sparse meadows were observed on the top of the rocky reefs too, however they were not represented due to their small size. The coralligenous communities were present on each rocky substrate from around 25 m depth while rocks covered by algal communities were found above this limit.

The temperature profiles of the two June acquisitions did not present a clear thermocline, but rather two main temperature gradients (Fig. B.4). The greatest temperature decreasing of 1.5°C was found in the twenty first metres in June 2016 (from 19.0 °C to 17.5 °C), and in the ten first metres in June 2019 (from 23 °C to 21.5 °C). Beyond 20 m depth, the temperatures went down more slowly to stabilise at 17.2 °C and 19.0 °C in June 2016 and June 2019 respectively. The

ANNEXE B. COUPLING THE MAP OF MARINE HABITATS AND FISH ACCUMULATION ZONES : A THREE DIMENSIONAL SPATIAL APPROACH FOR THE MANAGEMENT OF HALIEUTIC RESOURCES.

TABLE B.1 – Areas covered by each marine habitats and their proportion.

Habitat	Area (m ²)	Proportion (%)
Coralligenous communities	27 838	2.9
<i>P. oceanica</i> meadows	7 225	1.0
Algal cover on rocky substrate	39 151	5.7
Soft sediments	624 502	89,4
Total	698 716	100

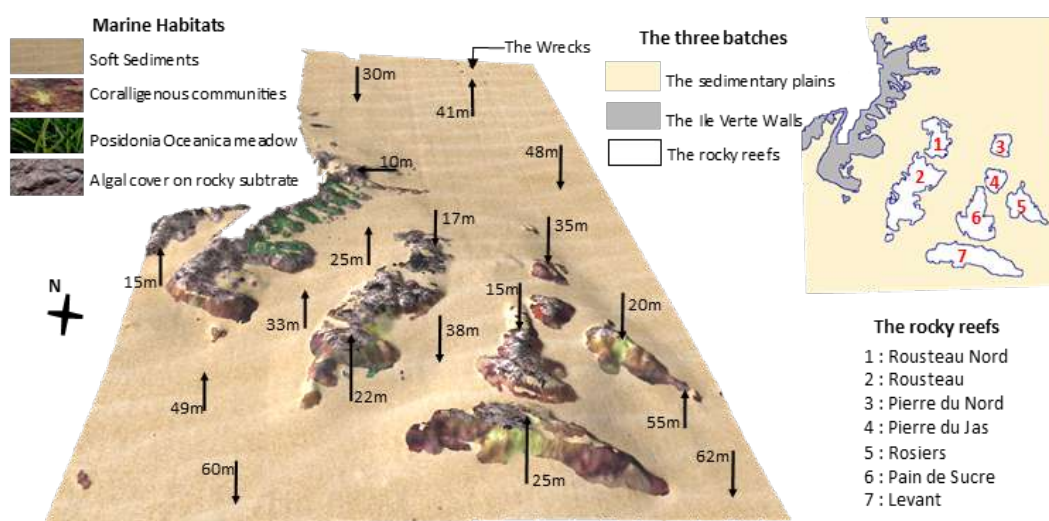


FIGURE B.3 – 3D representation of the seabed topology with marine habitats and details of the three batches studied and the reefs names for the study of the fish density.

August 2016 profile showed a thermocline at 5-10 m depth where temperatures dropped from 25 °C to 21 °C and then declined more slowly to 19 °C deeper (Fig. B.4).

For each acquisition, fish accumulations were well focused on the Ile Verte walls and on the rocky reefs (Fig. B.5). They appeared as large continuous and dense schools rather than numerous medium schools. The accumulation on the wrecks were worth more noteworthy than the surroundings, except in August 2019 where a large fish school was observed (Fig. B.5). With the exception of the north west boundary of the site, there were no major fish schools on the sedimentary plains (Fig. B.5). Numerous single detections were pointed on this habitat however. Fish detection were more numerous in June 2016 (38 498 detections) than in August (23 268 detections), while the number of fishes in August 2019 reached 57 472 detections (Table B.2). Concerning fish densities, soft sediments showed low fish densities (from 0.01 fishes/m² in June and August 2016 to 0.03 fishes/m² in June 2019), however close to the average value (Fig. B.6). On average through the three acquisitions, the Ile Verte walls had a higher fish abundance (0.31, 0.14 and 0.38 fishes/m² in June, August 2016 and June 2019 respectively) than the rocky reefs (0.25, 0.15 and 0.24 fishes/m² in June, August 2016 and June 2019 respectively ; Fig. B.6). Among the reefs, Rousteau Nord and Pierre du Jas showed the highest fish abundances, while the largest reefs – Rousteau and Levant – were the among the poorest (Fig. B.6).

Whatever the time acquisition, the fishes were mainly detected in the twenty first metres, more than 50 % of all detections in general and up to 66.8 % in June 2016 (Table B.2). By averaging the three acquisitions, about a third of detections were situated in the mid-depth waters, between 20 and 40 m. Finally, less than 10 % of the fishes detected were found in the deepest waters of the

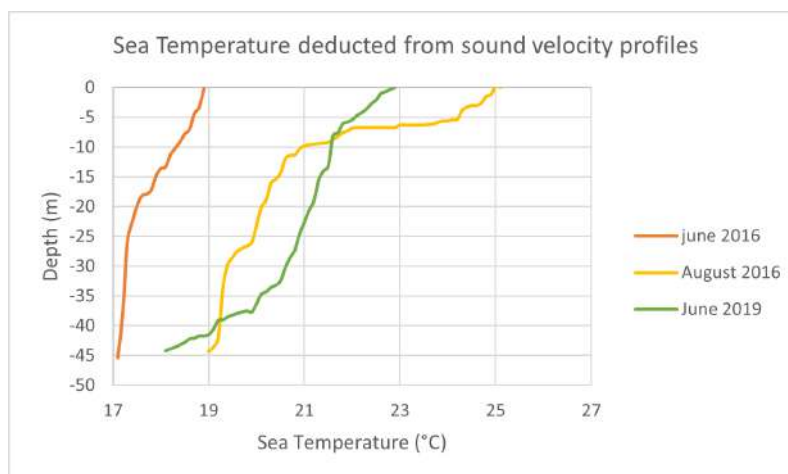


FIGURE B.4 – Temperature profile in June 2016 (black) ; August 2016 (red) ; June 2019 (blue).

TABLE B.2 – Number of fish detections and proportions per depth categories.

Position in the water column and depth range (m)	June 2016		August 2016		June 2019	
Near surface [0, 20[25 695	66,8 %	14 316	61,5 %	29 590	51,5 %
Mid-depth [20, 40[10 892	28,3 %	6 979	30,0 %	23 653	41,1%
Deep [40, max]	1 903	4,9 %	1 971	8,5 %	4 253	7,4 %
Total	38 492		23 268		57 472	

study site (Table B.2).

B.3 Discussion

This work was aiming at studying the capacity of a compact MBES to provide a precise map of marine habitats along with the fish accumulation, with the final purpose of providing a spatial ecosystem based approach to managers and stakeholders.

MBES advantages and operability

The main original outcome of this research effort is to put into light the possibility to have a global spatial review of an area with a single acoustic acquisition. Modern MBES are able to collect both bathymetry, backscatter and WCI without compensation of data quality and quantity. The WCI alone is used in various fields of research, such as biology, archeology, physical oceanology (Colbo et al.; 2014), and reveals its full potential when combined with the other MBES outputs. The versatility of the MBES used in this study (a R2Sonic 2022) is an advantage, on the condition of having the capacity and to properly collect, store and process all the acoustic data, without forgetting GPS and Navigation information. However, this MBES is commonly dedicated to bathymetry and seafloor imagery acquisition, it is not a fishery-dedicated tool. Moreover, the system is adjusted and calibrated for the seafloor measurement and not for the water column observation. This implies several consequences : (1) acoustic noises can alter the WCI and hinder their processing ; (2) only the pelagic fishes can be seen, the demersal and benthic ones are mingled with the seabed noise ; (3) When compared to fishery MBES, the one used in this work is installed under the hull at mid-length of the boat and not in the bow as a forward-looking MBES. Thus, the WCI is a vertical cut of the fish schools, and not a fish school seen as a whole ; (4) as the MBES is used to map the seabed, the acoustic signals are emitted towards the nadir. There is also a good cover

ANNEXE B. COUPLING THE MAP OF MARINE HABITATS AND FISH ACCUMULATION ZONES : A THREE DIMENSIONAL SPATIAL APPROACH FOR THE MANAGEMENT OF HALIEUTIC RESOURCES.

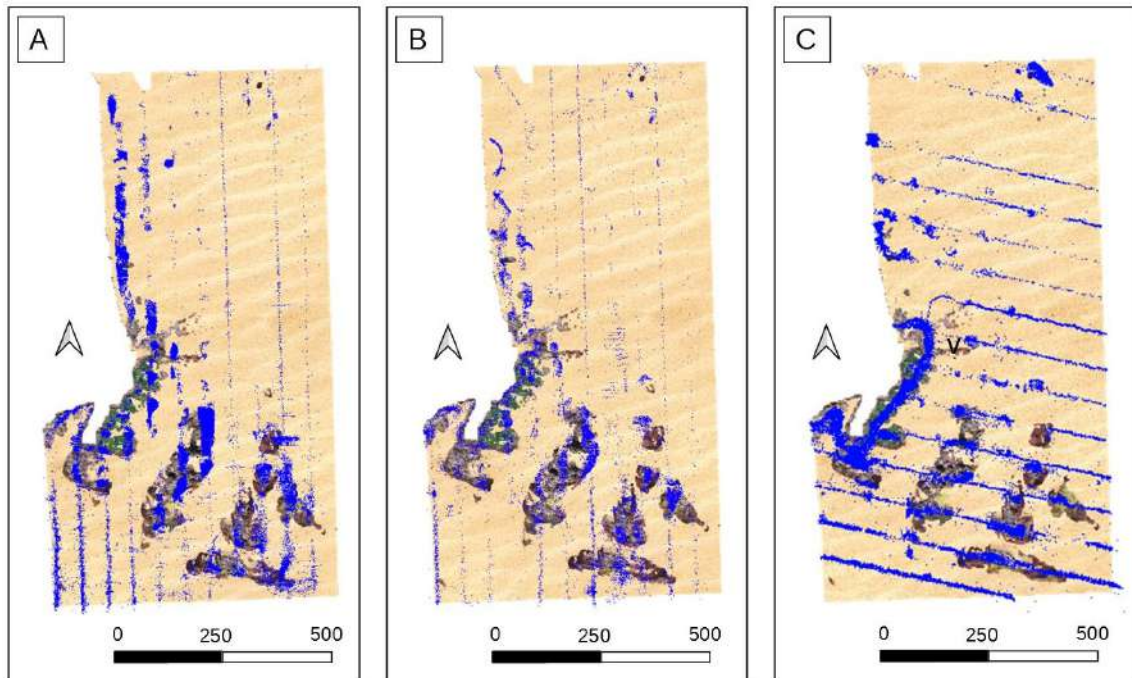


FIGURE B.5 – Scatter plots representing fishes in **A**, June 2016; **B** August 2016; **C**, June 2019. Each blue dot represents a detected fish whatever it size.

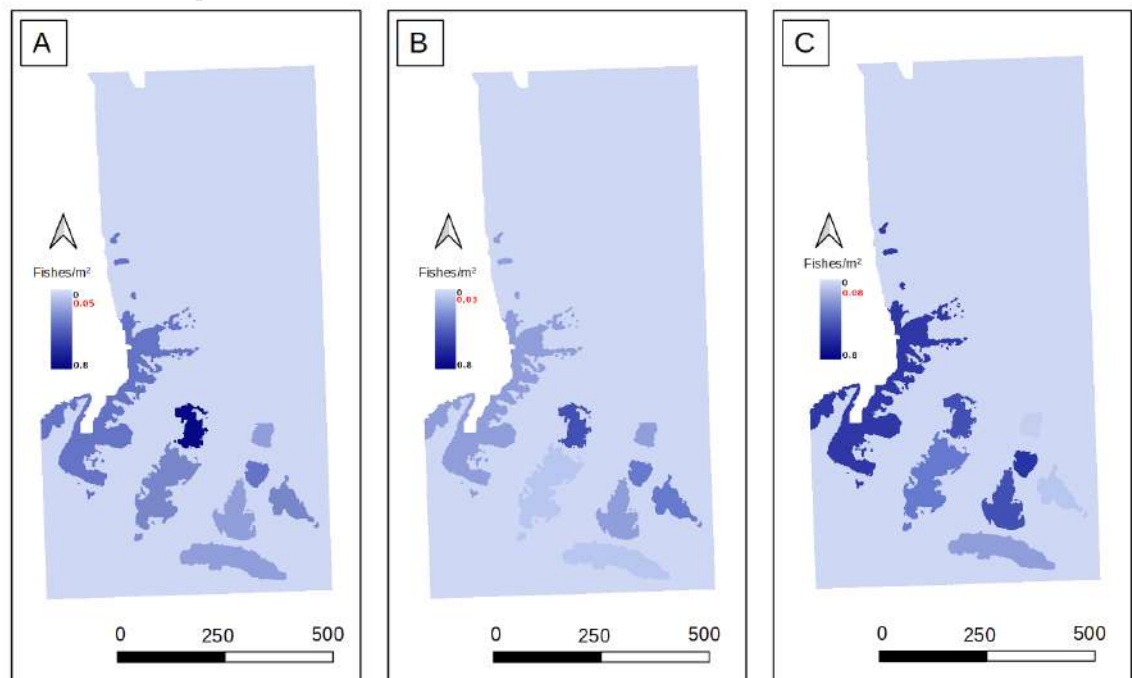


FIGURE B.6 – Local fish densities in fishes/m² in **A**, June 2016; **B** August 2016; **C**, June 2019. The red value corresponds to the mean number of fishes/m².

of the seabed, but the water column is not entirely scanned. This explains why the scatter plot is composed of bands North-South oriented.

The place on the hull and the looking direction of the MBES become important when the avoiding-boat behaviour of schools is considered. In a previous study on fish behaviours regarding

MBES acquisition, [Soria et al. \(1996\)](#) explained that a fish school feels the vessel coming far away. From this moment, a first part of the school avoids laterally the vessel and is not seen. Then, when the disturbance, i.e. the boat, arrives above the school, another part of the school also avoids it laterally and can only be seen on the edge of the WCI. What remains of the school is the little part recorded by the MBES while fishes are avoiding the disturbance by diving. That is why [Soria et al. \(1996\)](#) and [Paramo et al. \(2010\)](#) used a MBES with a 45° tilt from the nadir. Nevertheless, it is hard to say how many fishes are missed in the detection process.

Processing of the WCI

Although the algorithm developed to extract the fish information from the WCI is able to automatically detect fish targets without human intervention, this type of processing has currently several drawbacks. In this line of thought, wherever in the area, several fishes are not detected because they do not pass the filters. Some false alarms exist too. This is not disturbing the scatter plot and it does not influence so much the local density as well. However, it is more troublesome on the vast soft sedimentary seafloor where it seems that too much detections were performed. Thus, all these points are lonely, close to the seabed and around the nadir and could correspond to some acoustic noise. However, when looking the WCI at great depths, these points look like fishes for the algorithm as well as the human eyes. On the one hand, should these detections might be noises, then the sedimentary seafloor are really deserted by pelagic fishes. On the other hand, they could be true detections and these vast areas might be more populated than expected, while remaining very sparsely populated. We decided to display them, at least so that the reader may view the boat trajectories and assess the difficulty to validate fish detections.

Despite all these drawbacks, the processing method is fast enough to provide quick results, meaning with a mid-powerful computer, one hour of acquisition is processed in one hour of computer calculation. Moreover, improvement are under study in order to obtain a higher precision in target detection while decreasing the processing time. Another enhancement under progress concerns the pre-processing of the WCI to reduce the noise or calibrate the background noise.

Contribution to the ecosystem based management

This research work suggests a new approach to evaluate the ecological status of an area of interest with different levels of analysis that can be adapted according the characteristics of the managed zone (e.g. large areas, complex patchwork of marine habitats, extensive seagrass meadows), and the final aim of the study (e.g. MSFD, seascape analysis, halieutic research). It also allows to obtain spatial data for several functional compartments of the Mediterranean ecosystems *P. oceanica* meadows, algae-dominated rock reefs and coralligenous communities (Table B.3), according to the conceptual representation of [Personnic et al. \(2014\)](#), [Thibaut et al. \(2017\)](#) and [Ruitton et al. \(2014\)](#) respectively. Obviously this spatial approach does not replace the qualitative and quantitative assessment made by scuba diving but rather intervenes as a complementary tool allowing cross-validation. It may also be used to produce a first investigation of an area with few data on fish accumulations and benthic habitats for a more efficient scuba diving evaluation later.

One of the limits of a spatial approach through discrete acquisitions relies on the instantaneousness of the maps produced. Moreover, the pelagic fish distribution depends of numerous biotic and abiotic parameters that are virtually impossible to wholly assess such as marine habitats, sea temperature, salinity currents, day period, light intensity ([Saraux et al.; 2014](#)). This difficulty is illustrated in this study where the combination of the sea temperature, the seascape and the marine habitats are not sufficient to explain the fish distribution observed. In fact, if the fish biomass is increasing with sea temperatures, the August 2016 acquisition should have been the survey with the most detections in total and especially within the twenty first metres where temperatures were the highest. On the contrary, this acquisition shows about two times less fish detections than in June 2016 and 2019 (Table B.2). The same observation is made for the June acquisitions, due to the sea surface temperature, the one of 2019 should have been more fish abundant than the one of 2016.

TABLE B.3 – Functional compartments of the Mediterranean ecosystems investigated by the coupling of marine habitat maps and fish detection in the water column.

Ecosystem	Functional compartments	Type of data
<i>Posidonia oceanica</i> meadows	Posidonia leaves	Area covered
	Planktivorous teleosts	Fish number
	Piscivorous teleosts	
	Predatory teleosts	
	Herbivores 1	
Algae-dominated rock reefs	Multicellular photosynthetic organisms	Area covered
	Herbivorous teleosts	Fish number
	Piscivorous teleosts	
	Omnivorous teleosts	
	Invertivorous teleosts	
	Planktivorous teleosts	
Coralligenous communities	Builders	Area covered
	High-level predators	Fish number
	Predatory teleosts	
	Planktivorous teleosts	

Although this first application is promising for an effective evaluation in a management purpose, further developments are still required for a deeper investigation of the link between the marine habitats and the fish accumulations. Moreover, the anthropogenic impacts and abiotic factors must be included in the analysis for a sharper ecological assessment. For instance, the fish densities need to be calculated per habitats and not only by area. Likewise, the vertical dimension should be more exploited with volumetric analysis rather than the only study of the vertical repartition of fish schools. An important research effort is also required to link the size of the WCI acoustic targets with the one of actual fishes to produce an evaluation of the biomass per surface and/or volume (even with a relatively large margin of error). At last, an ecological index can be built on the comparison between different sites (impacted and protected) at various seasons to link anthropogenic impacts with the ecological status.

This research effort clearly highlights the capacity of spatial acoustic data obtained with a MBES to provide quantitative information on the marine habitats and fish distribution. This work aimed at paving the way to further developments to provide managers with effective spatial tools to evaluate the ecological status of key Mediterranean marine ecosystems. If coupled with *in situ* underwater observations, this spatial approach has the potential to give a complete view of underwater key biological systems as never before.

Acknowledgments

This work was made possible thanks to the Agence de l'Eau RMC that financed the hydrographic surveys in 2016 and 2019, as well as the first research works on the analysis of the WCI. The first author acknowledges a PhD fund from the Association Nationale de la Recherche et de la Technologie (ANRT).

Bibliographie

- Abadie, A., Marty, P. and Viala, C. (2017). Batclas index : a new method to identify and map with high resolution natural and artificial underwater structures on marine wind turbine sites, *3rd Wind Energy and Wildlife seminar. Artigues-près-Bordeaux, France* pp. 120–127.
- Abadie, A. and Viala, C. (2018). Le sondeur multifaisceaux en hydrographie : utilisations actuelles et futures., *XYZ* **157** : 17—27.
- Aggarwal, C. C. (2016). *Outlier Analysis*, 2nd edn, Springer Publishing Company, Incorporated.
- Anderson, J., Holliday, D., Kloser, R., Reid, D. and Simard, Y. (2008). Acoustic seabed classification : Current practice and future directions, *Ices Journal of Marine Science - ICES J MAR SCI* **65** : 1004–1011.
- Appeltans, W., Ah Yong, S., Anderson, G., Angel, M., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., Berta, A., Blazewicz-Paszkwowycz, M., Bock, P., Boxshall, G., Boyko, C., Brandao, S., Bray, R., Bruce, N., Cairns, S., Chan, T., Cheng, L., Collins, A., Cribb, T., Curini-Galletti, M., Dandouh-Guebas, F., Davie, P., Dawson, M., De Clerck, O., De Cock, W., De Grave, S., de Voogd, N., Domning, D., Emig, C., Erseus, C., Eschmeyer, W., Fauchald, K., Fautin, D., Feist, S., Fransen, C., Furuya, H., Garcia-Alvarez, O., Gerken, S., Gibson, D., Gittenberger, A., Gofas, S., Gomez-Daglio, L., Gordon, D., Guiry, M., Hernandez, F., Hoeksema, B., Hopcroft, R., Jaume, D., Kirk, P., Koedam, N., Koeneemann, S., Kolb, J., Kristensen, R., Kroh, A., Lambert, G., Lazarus, D., Lemaitre, R., Longshaw, M., Lowry, J., Macpherson, E., Madin, L., Mah, C., Mapstone, G., McLaughlin, P., Mees, J., Meland, K., Messing, C., Mills, C., Molodtsova, T., Mooi, R., Neuhaus, B., Ng, P., Nielsen, C., Norenburg, J., Opresko, D., Osawa, M., Paulay, G., Perrin, W., Pilger, J., Poore, G., Pugh, P., Read, G., Reimer, J., Rius, M., Rocha, R., Saiz-Salinas, J., Scarabino, V., Schierwater, B., Schmidt-Rhaesa, A., Schnabel, K., Schotte, M., Schuchert, P., Schwabe, E., Segers, H., Self-Sullivan, C., Shenkar, N., Siegel, V., Sterrer, W., Stohr, S., Swalla, B., Tasker, M., Thuesen, E., Timm, T., Todaro, M., Turon, X., Tyler, S., Uetz, P., van der Land, J., Vanhoorne, B., van Ofwegen, L., van Soest, R., Vanaverbeke, J., Walker-Smith, G., Walter, T., Warren, A., Williams, G., Wilson, S. and Costello, M. (2012). The magnitude of global marine species diversity, *Current Biology* **22**(23) : 2189–2202.
- Arthur, D. and Vassilvitskii, S. (2007). K-means++ : The advantages of careful seeding, *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, Society for Industrial and Applied Mathematics, USA, p. 1027–1035.
- Artalheiro, F. M. F. (1998). Analysis and procedures of multibeam data cleaning for bathymetric charting, *Technical Report 192*, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, New Brunswick, Canada.
- Beulah, J. and Punithavathani, D. (2015). Outlier detection methods for identifying network intrusions – a survey, **10** : 40488–40496.
- Blondel, P. (2009). *Acoustic signals and data acquisition*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 7–34.

- Boom, B., Huang, P., Beyan, C., Spampinato, C., Palazzo, S., He, J., Beauxis-Aussalet, E., Lin, S.-I., Chou, H.-M., Nadarajan, G., Chen-Burger, J., Van Ossenbruggen, J., Giordano, D., Hardman, L., Lin, F.-P. and Fisher, R. (2012). Long-term underwater camera surveillance for monitoring and analysis of fish populations, *International Conference on Visual observation and Analysis of Animal and Insect Behaviour* p. 4.
- Bottelier, P., Briese, C., Hennis, N., Lindenbergh, R., Pfeifer, N., Renard, P., Demougeot-Renard, H. and Froidevaux, R. (2005). *Distinguishing features from outliers in automatic Kriging-based filtering of MBES data : a comparative study*, pp. 403–414.
- Brown, C., Cooper, K., Meadows, W., Limpenny, D. and Rees, H. (2002). Small-scale mapping of sea-bed assemblages in the eastern english channel using sidescan sonar and remote sampling techniques, *Estuarine, Coastal and Shelf Science* **54**(2) : 263–278.
URL: <https://www.sciencedirect.com/science/article/pii/S0272771401908418>
- Brown, C. J. and Blondel, P. (2009). Developments in the application of multibeam sonar backscatter for seafloor habitat mapping, *Applied Acoustics* **70**(10) : 1242–1247. The Application of Underwater Acoustics for Seabed Habitat Mapping.
URL: <https://www.sciencedirect.com/science/article/pii/S0003682X0800176X>
- Brown, C. J., Smith, S. J., Lawton, P. and Anderson, J. T. (2011). Benthic habitat mapping : A review of progress towards improved understanding of the spatial ecology of the seafloor using acoustic techniques, *Estuarine, Coastal and Shelf Science* **92**(3) : 502–520.
URL: <https://www.sciencedirect.com/science/article/pii/S0272771411000485>
- Burke, R. and Robson, J. (1975). An evaluation of the bo'sun multi-beam sonar system, *The International Hydrographic Review* **52**(2) : 53–69.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/23749>
- Calder, B. R. and Mayer, L. A. (2003). Automatic processing of high-rate, high-density multibeam echosounder data, *Geochemistry Geophysics Geosystems* **4**(6) : 1048.
- Calder, B. R. and Rice, G. (2011). Design and implementation of an extensible variable resolution bathymetric estimator, *In proceeding of the U.S. Hydrographic Conference, Tampa, FL, USA*, p. 15.
- Calò, A., Félix-Hackradt, F., Garcia, J., Hackradt, C., Rocklin, D., Treviño Otón, J. and Charton, J. (2013). A review of methods to assess connectivity and dispersal between fish populations in the mediterranean sea, *Advances in Oceanography and Limnology* **4** : 150–175.
- Calvert, J., Strong, J. A., Service, M., McGonigle, C. and Quinn, R. (2015). An evaluation of supervised and unsupervised classification techniques for marine benthic habitat mapping using multibeam echosounder data, *ICES Journal of Marine Science* **72**(5) : 1498–1513.
- Cardon, D., Cointet, J.-P. and Mazieres, A. (2018). La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle, *Réseaux* **5** : 173.
- Carr, M. and Hixon, M. (1997). Artificial reefs : The importance of comparisons with natural reefs, *Fisheries Magazine* **22**(4) : 28–33.
- Chang, C.-C. and Lin, C.-J. (2011). Libsvm : a library for support vector machines, *ACM Transactions on Intelligent Systems and Technology* **2**(2) : 27.
URL: <https://www.csie.ntu.edu.tw/~cjlin/papers/libsvm.pdf>
- Charbonnel, E., Serre, C., Ruitton, S., Harmelin, J.-G. and Jensen, A. (2002). Effects of increased habitat complexity on fish assemblages associated with large artificial reef units (french mediterranean coast), *ICES Journal of Maritime Sciences* **59** : 208–213.
- Chu, D. (2011). Technology evolution and advances in fisheries acoustics, *Journal of Marine Science and Technology* **19** : 245–252.

- Clark, S. and Edwards, A. (1994). Use of artificial reef structures to rehabilitate reef flats degraded by coral mining in the maldives, *Bulletin of Marine Sciences* **55**(2-3) : 724–744.
- Clay, C. S. and Medwin, H. (1977). *Acoustical Oceanography : Principles and Applications*, John Wiley & sons.
- Cogan, C. B., Todd, B. J., Lawton, P. and Noji, T. T. (2009). The role of marine habitat mapping in ecosystem-based management, *ICES Journal of Marine Science* **66**(9) : 2033–2042.
URL: <https://doi.org/10.1093/icesjms/fsp214>
- Colbo, K., Ross, T., Brown, C. and Weber, T. (2014). A review of oceanographic applications of water column data from multibeam echosounders, *Estuarine, Coastal and Shelf Science* **145** : 41–56.
- Collier, J. and Brown, C. (2005). Correlation of sidescan backscatter with grain size distribution of surficial seabed sediments, *Marine Geology* **214**(4) : 431–449.
URL: <https://www.sciencedirect.com/science/article/pii/S0025322704003196>
- Costa, B. and Battista, T. (2013). The semi-automated classification of acoustic imagery for characterizing coral reef ecosystems, *International Journal of Remote Sensing* **34** : 6389–6422.
- Curtin, R. and Prellezo, R. (2010). Understanding marine ecosystem based management : A literature review, *Marine Policy* **34**(5) : 821 – 830.
- Cutter, G. and Demer, D. (2007). Accounting for scattering directivity and fish behaviour in multibeam-echosounder surveys, *ICES Journal of Marine Science* **64** : 1664–1674.
- Debese, N., Moitié, R. and Seube, N. (2012). Multibeam echosounder data cleaning through a hierarchic adaptive and robust local surfacing, *Computers & Geosciences* **46** : 330–339.
URL: <https://hal-ensta-bretagne.archives-ouvertes.fr/hal-00672672>
- DeFries, R. and Chan, J. C.-W. (2000). Multiple criteria for evaluating machine learning algorithms for land cover classification from satellite data, *Remote Sensing of Environment* **74**(3) : 503–515.
- Deneshkumar, V., Senthamaraiannan, K. and Manikandan, M. (2014). Identification of outliers in medical diagnostic system using data mining techniques, *International journal of statistics and applications* **4** : 241–248.
- Dhawale, C., Dhawale, K. and Dubey, R. (2020). *A Review on Deep Learning Applications*, pp. 21–31.
- Diesing, M., Coggan, R. and Vanstaen, K. (2009). Widespread rocky reef occurrence in the central english channel and the implications for predictive habitat mapping, *Estuarine Coastal and Shelf Science* **83** : 647–658.
- Diesing, M., Mitchell, P. and Stephens, D. (2016). Image-based seabed classification : what can we learn from terrestrial remote sensing?, *ICES Journal of Marine Science* **73**(10) : 2425–2441.
- Du, Z., Wells, D. and Mayer, L. (1996). An approach to automatic detection of outliers in multi-beam echo sounding data, *The Hydrographic Journal* **79** : 19–23.
- Duchi, J., Hazan, E. and Singer, Y. (2011). Adaptive subgradient methods for online learning and stochastic optimization, *Journal of Machine Learning Research* **12** : 2121–2159.
- Etter, P. C. (1998). *Underwater acoustic modeling : princiles, techniques and application*, E & FN Spon.
- Fezzani, R., Zerr, B., Mansour, A., Legris, M. and Vrignaud, C. (2019). Fusion of swath bathymetric data : Application to auv rapid environment assessment, *IEEE Journal of Oceanic Engineering* **44**(1) : 111–120.

- Fleury, G., Gondran, M., Lacomme, P. and Samir, C. (2021). *Découverte du Machine Learning. Les outils de l'apprentissage automatique*, ellipses.
- Fonseca, L., Brown, C., Calder, B., Mayer, L. and Rzhannov, Y. (2009). Angular range analysis of acoustic themes from stanton banks ireland : A link between visual interpretation and multi-beam echosounder angular signatures, *Applied Acoustics* **70**(10) : 1298–1304. The Application of Underwater Acoustics for Seabed Habitat Mapping.
- Francois, R. E. and Garrison, G. R. (1982a). Sound absorption based on ocean measurements : Part 1 :pure water and magnesium sulfate contributions, *Journal of the Acoustical Society of America* **72**(3) : 896–907.
- Francois, R. E. and Garrison, G. R. (1982b). Sound absorption based on ocean measurements : Part 2 :boric acid contribution and equation for total absorption, *Journal of the Acoustical Society of America* **72**(6) : 1879–1890.
- Fukushima, K., Miyake, S. and Ito, T. (1983). Neocognitron : A neural network model for a mechanism of visual pattern recognition, *IEEE Transactions on Systems, Man, and Cybernetics SMC-13*(5) : 826–834.
- Gayet, G., Baptist, F., Maciejewski, L., Poncet, R. and Bensettiti, F. (2018). *Guide de détermination des habitats terrestres et marins de la typologie EUNIS*, Agence française pour la biodiversité.
- Georgiadis, M., Papatheodorou, G., Tzanatos, E., Geraga, M., Ramfos, A., Koutsikopoulos, C. and Ferentinos, G. (2008). Coralligène formations in the eastern mediterranean sea : Morphology, distribution, mapping and relation to fisheries in the southern aegean sea (greece) based on high-resolution acoustics, *Journal of Experimental Marine Biology and Ecology* **368** : 44–58.
- Géron, A. (2017). *Deep Learning avec TensorFlow*, Dunod.
- Glenn, M. F. (1970). Introducing an operational multi-beam array sonar, *The International Hydrographic Review* **47**(1) : 35–39.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/23921>
- Glorot, X. and Bengio, Y. (2010). Understanding the difficulty of training deep feedforward neural networks, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, Vol. 9 of *Proceedings of Machine Learning Research*, PMLR, pp. 249–256.
- Gurshin, C., Jech, J., Howell, W., Weber, T. and Mayer, L. (2009). Measurements of acoustic backscatter and density of captive atlantic cod with synchronized 300-khz multibeam and 120-khz split-beam echosounders, *ICES Journal of Marine Sciences* **66** : 1303–1309.
- Harmelin, J., Bachet, F. and Garcia, F. (1995). Mediterranean marine reserves : Fish indices as tests of protection efficiency, *Marine Ecology* **16** : 233–250.
- Hasan, R. C., Ierodiaconou, D., Laurensen, L. and Schimel, A. (2014). Integrating multibeam backscatter angular response, mosaic and bathymetry data for benthic habitat mapping, *PLOS ONE* **9**(5) : 1–14.
URL: <https://doi.org/10.1371/journal.pone.0097339>
- Hasan, R. C., Ierodiaconou, D. and Monk, J. (2012). Evaluation of four supervised learning methods for benthic habitat mapping using backscatter from multi-beam sonar, *Remote Sensing* **4**(11) : 3427–3443.
- Hastie, T., Tibshirani, R. and Friedman, J. (2001). *The Elements of Statistical Learning*, Springer Series in Statistics, Springer New York Inc., New York, NY, USA.
- Hawkins, D. (1980). *Identification of outliers*, Monographs on applied probability and statistics, Chapman and Hall, London, U.K.

- Hay, G. and Castilla, G. (2008). *Geographic Object-Based Image Analysis (GEOBIA) : A new name for a new discipline*, pp. 75–89.
- He, K., Zhang, X., Ren, S. and Sun, J. (2015). Deep residual learning for image recognition.
- Hebb, D. O. (1949). *The organization of behavior : A neuropsychological theory*, Wiley.
- Held, P. and Schneider von Deimling, J. (2019). New feature classes for acoustic habitat mapping—a multibeam echosounder point cloud analysis for mapping submerged aquatic vegetation (sav), *Geosciences* **9**(5).
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2012). Improving neural networks by preventing co-adaptation of feature detectors, *CoRR* .
URL: <http://arxiv.org/abs/1207.0580>
- Hofmann, M., Tiefenbacher, P. and Rigoll, G. (2012). Background segmentation with feedback : the pixel-based adaptative segmenter, *IEEE Computer society Conference on Computer vision and pattern recognition workshops* .
- Holmin, A., Handegard, N., Korneliussen, R. and Tjostheim, G. (2012). Simulation of multi-beam sonar echos from schooling individual fish in a quiet environment, *The journal of Acoustical Society of America* **132**.
- Hooper, D., Buchmann, N., Degrange, V., Diaz, S., Gessner, M., Grime, P., Hulot, F., Mermillod-Blondin, F., Peer, L., Roy, J., Symstad, A., Solan, M. and Spehn, E. (2002). *Species diversity, functional diversity and ecosystem functioning*, Oxford University Press, pp. 195–208.
- Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurons in the cat’s striate cortex, *Journal of Physiology* **148** : 574–591.
- Ierodiaconou, D., Monk, J., Rattray, A., Laurenson, L. and Versace, V. (2011). Comparison of automated classification techniques for predicting benthic biological communities using hydroacoustics and video observations, *Continental Shelf Research* **31**(2, Supplement) : S28–S38. Geological and Biological Mapping and Characterisation of Benthic Marine Environments.
- IFREMER (2007). Sonarscope.
URL: <https://www.flotteoceanographique.fr/Nos-moyens/Logiciels-de-la-flotte/Analyse-et-traitement-de-l-information/SonarScope>
- Innangi, S., Bonanno, A., Tonielli, R., Gerlotto, F., Innangi, M. and Mazzola, S. (2016). High resolution 3d shapes of fish schools : a new method to use the water column backscatter from hydrographic multi-beam echo sounders, *Applied Acoustics* **111** : 148–160.
- Ioffe, S. and Szegedy, C. (2015). Batch normalization : Accelerating deep network training by reducing internal covariate shift.
- Karthikeyan, K. and Chandrasekar, C. (2011). Speckle noise reduction of medical ultrasound images using bayesshrink wavelet threshold, *International Journal of Computer Application* **22**(9) : 8–14.
- Kostylev, V., Todd, B., Fader, G., Courtney, R., Cameron, G. and Pickrill, R. (2001). Benthic habitat mapping on the scotian shelf based on multibeam bathymetry, surficial geology and sea floor photographs, *Marine Ecology Progress Series* **219** : 121–137.
- Krizhevsky, A., Sutskever, I. and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks, in F. Pereira, C. J. C. Burges, L. Bottou and K. Q. Weinberger (eds), *Advances in Neural Information Processing Systems 25*, Curran Associates, Inc., pp. 1097–1105.

- Lamarche, G., Lurton, X., Verdier, A.-L. and Augustin, J.-M. (2011). Quantitative characterisation of seafloor substrate and bedforms using advanced processing of multibeam backscatter—application to Cook Strait, New Zealand, *Continental Shelf Research* **31**(2, Supplement) : S93–S109.
- Lamouret, M., Abadie, A., Viala, C., Boissery, P. and Thirion-Moreau, N. (2019). Measuring fish activities as additional environmental data during a hydrographic survey with a multi-beam echosounder., *Oceans 2019 Marseille - France MTS*, IEEE, p. 10.
- Lawrence, M.J. and Bales, C. (2001). Acoustic ground discrimination techniques for submerged archaeological site investigations, *Marine Technology Society Journal* **35**(4) : 65–73.
- Le Bas, T. and Huvenne, V. (2009). Acquisition and processing of backscatter data for habitat mapping – comparison of multibeam and sidescan systems, *Applied Acoustics* **70**(10) : 1248–1257. The Application of Underwater Acoustics for Seabed Habitat Mapping.
URL: <https://www.sciencedirect.com/science/article/pii/S0003682X08001825>
- Le Deunf, J., Debese, N., Schmitt, T. and Billot, R. (2020). A review of data cleaning approaches in a hydrographic framework with a focus on bathymetric multibeam echosounder datasets, *Geosciences* **10**(7).
URL: <https://www.mdpi.com/2076-3263/10/7/254>
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition, *Neural Computation* **1** : 541–551.
- LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, pp. 2278–2324.
- Lelong, B. (2001). Paul Langevin et la détection sous-marine, 1914-1929. un physicien acteur de l'innovation industrielle et militaire, *Épistémologiques* **1**(3-4) : 205–232.
- Lindenbaum, C., Bennell, J., Rees, E., McClean, D., Cook, W., Wheeler, A. and Sanderson, W. (2008). Small-scale variation within a *Modiolus modiolus* (Mollusca : Bivalvia) reef in the Irish Sea : I. Seabed mapping and reef morphology, *Journal of the Marine Biological Association of the United Kingdom* **88**(1) : 133–141.
- Long, R. (2011). The marine strategy framework directive : A new European approach to the regulation of the marine environment, marine natural resources and marine ecological services, *Journal of Energy & Natural Resources Law* **29**(1) : 1–44.
- Løvik, A. and Høvern, J. (1979). An experimental investigation of swimbladder resonance in fishes, *The Journal of Acoustical Society of America* **66** : 850–854.
- Lu, D. and Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance, *International Journal of Remote Sensing* **28**(5) : 823 – 870.
- Lucieer, V. and Lamarche, G. (2011). Unsupervised fuzzy classification and object-based image analysis of multibeam data to map deep water substrates, Cook Strait, New Zealand, *Continental Shelf Research* **31**(11) : 1236–1247.
- Lurton, X. (1998). *Acoustique sous-marine : Présentation et applications*, IFREMER.
- Lurton, X. (2002). *An Introduction to Underwater Acoustics : Principles and Applications*, Geophysical Sciences Series, Springer.
- Marsh, I. and Brown, C. (2009). Neural network classification of multibeam backscatter and bathymetry data from Stanton Bank (Area IV), *Applied Acoustics* **70**(10) : 1269–1276. The Application of Underwater Acoustics for Seabed Habitat Mapping.

- McCaffrey, E. K. (1981). A review of the bathymetric swath survey system, *The International Hydrographic Review* **58**(1) : 19–27.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/23587>
- Mcculloch, W. and Pitts, W. (1943). A logical calculus of ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics* **5** : 127–147.
- McGonigle, C., Brown, C., Quinn, R. and Grabowski, J. (2009). Evaluation of image-based multi-beam sonar backscatter classification for benthic habitat discrimination and mapping at stanton banks, uk, *Estuarine, Coastal and Shelf Science* **81**(3) : 423–437.
URL: <https://www.sciencedirect.com/science/article/pii/S027277140800454X>
- Mielck, F., Bartsch, I., Hass, H., Wölfl, A.-C., Bürk, D. and Betzler, C. (2014). Predicting spatial kelp abundance in shallow coastal waters using the acoustic ground discrimination system roxann, *Estuarine, Coastal and Shelf Science* **143** : 1–11.
URL: <https://www.sciencedirect.com/science/article/pii/S027277141400064X>
- Minghelli, A., Vadakke-Chanat, S., Chami, M., Guillaume, M., Migne, E., Grillas, P. and Boutron, O. (2021). Estimation of bathymetry and benthic habitat composition from hyperspectral remote sensing data (biodiversity) using a semi-analytical approach, *Remote Sensing* **13**(10).
- Minghelli-Roman, A. and Dupouy, C. (2014). Correction of the water column attenuation : Application to the seabed mapping of the lagoon of new caledonia using meris images, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **7**(6) : 2619–2629.
- Minsky, M. and Papert, S. (1969). *Perceptrons : An Introduction to Computational Geometry*, MIT Press, Cambridge, MA, USA.
- Misund, O. and Aglen, A. (1992). Swimming behaviour of fish schools in the north sea during acoustic surveying and pelagic trawl sampling, *ICES Journal of Marine Science* **49** : 325–334.
- Mosca, F., Matte, G., Lerda, O., Naud, F., Charlot, D., Rioblanco, M. and Corbières, C. (2015). Scientific potential of a new 3d multibeam echosounder in fisheries and ecosystem research, *Fisheries Research* .
- Müller, A. C. and Guido, S. (2018). *Le Machine Learning avec Python*, O’Reilly Media, Inc.
- Nesterov, Y. (1983). A method for unconstrained convex minimization problem with the rate of convergence $o(1/k^2)$.
- Nøttestad, L., Fernø, A., Mackinson, S., Pitcher, T. and Misund, O. (2002). How whales influence herring school dynamics in a cold-front area of the norwegian sea, *ICES Journal of Marine Science* **59** : 393–400.
- OHI (2020). Standards for hydrographic surveys - 6th edition - special publication no. 44, *Technical report*.
URL: https://iho.int/uploads/user/pubs/Drafts/S-44_Edition_6.0.0-Final.pdf
- Ollivier, F. (1995). *Conception et réalisation d’un sonar latéral bathymétrique à haute résolution avec voies préformées compensées en attitude*, Theses, Université Pierre et Marie Curie (UPMC Paris 6).
URL: <https://hal.sorbonne-universite.fr/tel-01675027>
- Palomera, I., Olivar, M., Salat, J., Sabatés, A., Coll, M., García, A. and Morales-Nin, B. (2007). Small pelagic fish in the nw mediterranean sea : an ecological review, *Prog. Oceanogr* **74** : 377–396.
- Paramo, J., Gerlotto, F. and Oyarzún, C. (2010). Three dimensional structure and morphology of pelagic fish schools issue, *Journal of Applied Ichthyology* **26** : 853 – 860.

- Pergent-Martini, C., Leoni, V., Pasqualini, V., Ardizzone, G., Balestri, E., Bedini, R., Belluscio, A., Belsher, T., Borg, J., Boudouresque, C., Boumaza, S., Bouquegneau, J., Buia, M., Calvo, S., Cebrian, J., Charbonnel, E., Cinelli, F., Cossu, A., Maida, G. D., Dural, B., Francour, P., Gobert, S., Lepoint, G., Meinesz, A., Molenaar, H., Mansour, H., Panayotidis, P., Peirano, A., Pergent, G., Piazzzi, L., Pirrotta, M., Relini, G., Romero, J., Sanchez-Lizaso, J., Semroud, R., Shembri, P., Shili, A., Tomasello, A. and Velimirov, B. (2005). Descriptors of posidonia oceanica meadows : Use and application, *Ecological Indicators* **5**(3) : 213 – 230.
- Personnic, S., Boudouresque, C. F., Astruch, P., Ballesteros, E., Blouet, S., Bellan-Santini, D., Bonhomme, P., Thibault-Botha, D., Feunteun, E., Harmelin-Vivien, M., Pergent, G., Pergent-Martini, C., Pastor, J., Poggiale, J.-C., Renaud, F., Thibaut, T. and Ruitton, S. (2014). An ecosystem-based approach to assess the status of a mediterranean ecosystem, the posidonia oceanica seagrass meadow, *PLOS ONE* **9**(6) : 1–17.
- Polyak, B. (1964). Some methods of speeding up the convergence of iteration methods, *Ussr Computational Mathematics and Mathematical Physics* **4** : 1–17.
- R2Sonic (2014). Sonic 2024/2022 broadband multibeam echosounders, operation manual v5.0, *Technical report*.
- Rosenblatt, F. (1958). The perceptron : A probabilistic model for information storage and organization in the brain, *Psychological Review* pp. 65–386.
- Ruder, S. (2016). An overview of gradient descent optimization algorithms, *CoRR abs/1609.04747*.
URL: <http://arxiv.org/abs/1609.04747>
- Ruitton, S., Personnic, S., Ballesteros, E., Bellan-Santini, D., Boudouresque, C.-F., Chevaldonné, P., Bianchi, C. N., David, R., Feral, J.-P., Guidetti, P., Harmelin, J.-g., Montefalcone, M., Morri, C., Pergent, G., Pergent-Martini, C., Sartoretto, S., Tanoue, H., Thibaut, T., Vacelet, J. and Verlaque, M. (2014). An ecosystem-based approach to assess the status of the mediterranean coralligenous habitat., *RAC/SPA 2nd Mediterranean Symp. on the Conservation of coralligenous and other calcareous bio-concretions*, Vol. 2, RAC/SPA, Portorož , Slovenia, pp. 153–158.
- Rumelhart, D. E., Hinton, G. E. and Williams, R. J. (1986). Learning representations by back-propagating errors, *Nature* **323** : 533–536.
- Rønhovde, A., Yang, L., Taxt, T. and Holm, S. (1999). High-resolution beamforming for multibeam echo sounders using raw em3000 data, *OCEANS '99 MTS/IEEE*, Vol. 2, pp. 923 – 930.
- Sabatés, A., Martín, P., Lloret, J. and Raya, V. (2006). Sea warming and fish distribution : the case of the small pelagic fish, sardinella aurita, in the western mediterranean, *Global Change Biology* **12**(11) : 2209–2219.
- Sabol, B. M., Bales, C., Eddie Melton, R., Chamberlain, R., Doering, P. and Haunert, K. (2002). Evaluation of a digital echo sounder system for detection of submersed aquatic vegetation, *Estuaries* **25** : 133–141.
- Saraux, C., Fromentin, J., Bigot, J., Bourdeix, J., Morfin, M., Roos, D., Beveren, E. V. and Bez, N. (2014). Spatial structure and distribution of small pelagic fish in the northwestern mediterranean sea, *PLoS ONE* **9**.
- Savić, M., Atanasijević, J., Jakovetić, D. and Krejić, N. (2022). Tax evasion risk management using a hybrid unsupervised outlier detection method, *Expert Systems with Applications* p. 116409.
- Schneider von Deimling, J. and Weinrebe, W. (2014). Beyond bathymetry : Water column imaging with multibeam echo sounder systems, *Hydrographische Nachrichten* **31** : 6–10.

- Shahid, N., Rappon, T. and Berta, W. (2019). Applications of artificial neural networks in health care organizational decision-making : A scoping review, *PLOS ONE* **14**(2) : 1–22.
URL: <https://doi.org/10.1371/journal.pone.0212356>
- Sharma, A., Shrimali, V. R. and Beyeler, M. (2019). *Machine Learning for OpenCV 4*, packt.
- Simard, Y. and Stepnowski, A. (2007). Classification methods and criteria, *Acoustic Seabed Classification of Marine Physical and Biological Landscapes ICES CRR* **286** : 61–72.
- Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition, *Conference paper of International Conference on Learning Representations 2015*.
- Slocombe, D. S. (1993). Implementing ecosystem-based management : Development of theory, practice, and research for planning and managing a region, *BioScience* **43**(9) : 612–622.
- Soria, M., Fréon, P. and Gerlotto, F. (1996). Analysis of vessel influence on spatial behaviour of fish schools using a multi-beam sonar and consequences for biomass estimates by echo-sounder, *ICES Jour.Mar. Sci.* **53** : 453–458.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014). Dropout : A simple way to prevent neural networks from overfitting, *Journal of Machine Learning Research* **15**(56) : 1929–1958.
- STARESO and ANDROMEDE-OCEANOLOGIE (2012). Inventaires biologiques et analyse écologique des habitats marins patrimoniaux du site natura 2000 « agriates » - fr 9400570. contrat agence des aires marines protégées / stareso – andromède océanologie, *Technical report*.
- Stephens, D. and Diesing, M. (2014). A comparison of supervised classification methods for the prediction of substrate type using multibeam acoustic and legacy grain-size data, *PLOS ONE* **9**(4) : 1–14.
URL: <https://doi.org/10.1371/journal.pone.0093950>
- Stephens, D., Smith, A., Redfern, T., Talbot, A., Lessnoff, A. and Dempsey, K. (2020). Using three dimensional convolutional neural networks for denoising echosounder point cloud data, *Applied Computing and Geosciences* **5** : 100016.
- Svenmarck, P., Luotsinen, L., Nilsson, M. and Schubert, J. (2018). Possibilities and challenges for artificial intelligence in military applications, *NATO Big Data and Artificial Intelligence for Military Decision Making Specialists' Meeting, Bordeaux, France*.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015). Going deeper with convolutions, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9.
- Thibaut, T., Blanfuné, A., Boudouresque, C. F., Personnic, S., Ruitton, S., Ballesteros, E., Bellan-Santini, D., Bianchi, C. N., Bussotti, S., Cebrian, E., Cheminée, A., Culioli, J.-M., Derrien-Courtel, S., Guidetti, P., Harmelin-Vivien, M., Hereu, B., Morri, C., Poggiale, J.-C. and Verlaque, M. (2017). An ecosystem-based approach to assess the status of mediterranean algae-dominated shallow rocky reefs, *Marine Pollution Bulletin* **117**(1-2) : 311–329.
- Thorndike, R. L. (1953). Who belongs in the family?, *Psychometrika* **18** : 267–276.
- Trenkel, V., Mazauric, V. and Berger, L. (2008). The new fisheries multibeam echosounder me70 : description and expected contribution to fisheries research, *ICES Journal of Marine Sciences* **65** : 645–655.
- Tucker, D. G. (1960). Directional echo sounding : Some possible improvement in equipment and technique, *The International Hydrographic Review* **37**(2) : 43–53.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/26540>

- Tucker, D. G. and Hendersen, J. G. (1960). Automatic stabilization of underwater acoustic beams without mechanical motion of the transducer, *The International Hydrographic Review* **37**(1) : 69–78.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/26569>
- Tucker, M. J. (1961). Beam identification in multiple-beam echo-sounders, *The International Hydrographic Review* **38**(2) : 25–32.
URL: <https://journals.lib.unb.ca/index.php/ihr/article/view/26493>
- Tufféry, S. (2019). *Big Data, Machine Learning et apprentissage profond*, Editions TECHNIP.
- van Rein, H., Brown, C., Quinn, R. and Breen, J. (2009). A review of sublittoral monitoring methods in temperate waters : a focus on scale, *Underwater Technology : The International Journal of the Society for Underwater* **28** : 99–113.
- Van Veen, B. and Buckley, K. (1988). Beamforming : a versatile approach to spatial filtering, *IEEE ASSP Magazine* **5**(2) : 4–24.
- Viala, C. (2015a). Viewmap and viewsmf software.
URL: <https://seaviews.fr/en/services-3/software-development>
- Viala, C. (2015b). Viewmap and viewsmf softwares.
URL: <https://seaviews.fr/en/services-3/software-development>
- Viala, C. and Abadie, A. (2017). Mesure de l'activité biologique par sondeur multifaisceau. study report from seaviews to agence de l'eau rm& c., *Technical report*, Seaviews.
- Viala, C., Lamouret, M. and Abadie, A. (2021). Seafloor classification using a multibeam echo sounder : A new rugosity index coupled with a pixel-based process to map mediterranean marine habitats, *Applied Acoustics* **179** : 9.
- Wang, M., Wu, Z., Yang, F., Ma, Y., Wang, X. H. and Zhao, D. (2018). Multifeature extraction and seafloor classification combining lidar and mbes data around yuazhi island in the south china sea, *Sensors* **18**(11).
- Ware, C., Slipp, L., Wong, K. W., Nickerson, B. G., Wells, D. E., Lee, Y. C., Dodd, D. and Costello, G. (1992). A system for cleaning high volume bathymetry, *International Hydrographic Review* **69**.
- Wilding, T. A., Sayer, M. D. and Provost, P. G. (2003). Factors affecting the performance of the acoustic ground discrimination system RoxAnn™, *ICES Journal of Marine Science* **60**(6) : 1373–1380.
URL: [https://doi.org/10.1016/S1054-3139\(03\)00151-6](https://doi.org/10.1016/S1054-3139(03)00151-6)
- Zavalas, R., Ierodionou, D., Ryan, D., Rattray, A. and Monk, J. (2014). Habitat classification of temperate marine macroalgal communities using bathymetric lidar, *Remote Sensing* **6**(3) : 2154–2175.
- Zeiler, M. D. (2012). Adadelata : An adaptive learning rate method.
- Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks, *Computer Vision – ECCV 2014*, Springer International Publishing, pp. 818–833.