



HAL
open science

Contributions en méthodes bioacoustiques multiéchelles : spécifiques, populationnelles, individuelles et comportementales

Marion Poupard

► **To cite this version:**

Marion Poupard. Contributions en méthodes bioacoustiques multiéchelles : spécifiques, populationnelles, individuelles et comportementales. Acoustique [physics.class-ph]. Université de Toulon, 2020. Français. NNT : 2020TOUL0015 . tel-03287502

HAL Id: tel-03287502

<https://theses.hal.science/tel-03287502>

Submitted on 15 Jul 2021

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

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



UNIVERSITÉ DE TOULON, UFR SCIENCES & TECHNIQUES
École doctorale n°548 : Mer & Sciences

THÈSE

présentée par

Marion Poupard

Soutenue le 9 Décembre 2020

pour obtenir le grade de docteur de l'UNIVERSITÉ DE TOULON

Disciplines : Automatique, signal, productique, robotique ;
Bioacoustique.

Contributions en Méthodes Bioacoustiques Multiéchelles :

Spécifiques, Populationnelles, Individuelles et Comportementales

Thèse dirigée par :

M. Hervé Glotin	Pr. Lab. Informatique & Systèmes, CNRS, Univ. Toulon (directeur)
M. Thierry Soriano	Pr. Lab. Conception Systèmes Mécaniques & Robotiques, Univ. Toulon (co-directeur)
M. Thierry Lengagne	CR HDR Lab. Ecologie des Hydrosystèmes Naturels & Anthropisés, CNRS, Univ. Lyon I (co-encadrant)

Rapporteurs :

M. Yvan Simard	DR Ins. Sciences de la Mer, Rimouski, Univ. Québec, Canada
M ^{me} Clémentine Vignal	Pr. Ins. Ecologie & Sciences de l'Environnement de Paris, CNRS, Univ. Sorbonne

Examineurs :

M ^{me} Renata Sousa-Lima	Pr. Lab. of Bioacoustics, Univ. Rio Grande do Norte, Brésil
M ^{me} Aurélie Célérier	MC HDR Centre Ecologie Fonctionnelle & Evolutive, CNRS, Univ. Montpellier II
M. Gianni Pavan	Pr. Cent. Interdisci. Bioacustica & Ricerche Ambientali, Univ. Pavia, Italie

Invitée :

M ^{me} Anne-Laure Bedu	Responsable de la société Biosong
---------------------------------	-----------------------------------

'La sagesse,
c'est d'avoir des rêves suffisamment grands
pour ne pas les perdre de vue lorsqu'on les poursuit.'
Oscar Wilde

Remerciements

Premièrement, je tiens à remercier **Clémentine Vignal** et **Yvan Simard** d'avoir accepté de faire parti de mon jury de thèse et d'évaluer ce manuscrit.

Puis, je tiens à remercier mon directeur de thèse, **Hervé Glotin** pour la confiance qu'il m'a accordée malgré mes connaissances et mon bagage de biologiste, pour m'avoir guidée, encouragée et conseillée tout au long de ces trois années de thèse. Également pour son soutien infailible et régulier sur la rédaction de ce manuscrit.

Je remercie **Thierry Soriano** pour ses précieux conseils et sa patience pour avoir co-encadré une thèse qui n'était pas associée directement à son domaine de recherche. Merci pour ses précieux conseils durant la rédaction de ce manuscrit.

Un grand merci à **Thierry Lengagne**, mon encadrant de thèse qui a pris le temps de mettre en place les acquisitions de données pour les oiseaux, je le remercie pour ses compétences inspirantes en ornithologie et pour son aide précieuse pour la construction du plan de thèse et pour sa relecture.

Je remercie **Olivier Adam** et **Alexis Joly** pour avoir accepté de faire parti de mon comité de thèse et de m'avoir guidée tout le long de ce travail.

Je remercie l'entreprise BIOSONG, **Anne-Laure Bedu** et **Allain Bougrain-Dubourg** pour le co-financement de cette thèse mais également pour toutes les réunions de travail qui ont été très riches et constructives.

Je remercie le pole INPS de l'Université de Toulon pour le co-financement de cette thèse.

Des sincères remerciements à **François et Véronique Sarano** pour leurs précieuses inspirations concernant l'étude des cachalots et leur protection.

Je remercie les deux stagiaires en développement (**Rachid** et **Avotra**) pour avoir participé à la création du prototype d'application mobile Birdify, également **Jean-Marc** pour m'avoir aidée à créer les fiches d'identification.

Un grand merci aux membres de l'équipe DYNI, notamment **Julie** et **Franck**, **Ricard**, **Pascale** et **Jan** pour leurs précieux conseils en traitement du signal et bioacoustique.

Cette thèse à été le fruit de collaborations scientifiques enrichissantes, pour cela j'aimerais adresser un immense merci.

- Aux responsables de la station de recherche Orcalab : **Helena Symonds** et **Paul Spong** pour avoir accepté cette collaboration. Merci de m'avoir accueillie chaleureusement durant deux mois sur l'Île d'Hanson, d'avoir mis à disposition votre matériel et votre connaissance précieuse et unique des orques de Colombie-Britannique.
- A l'ensemble de l'équipe Sea Proven, pour avoir mis à disposition le drone Sphyrna et le trimaran "Solar". Merci pour votre patience et votre aide pour la construction de l'antenne acoustique, pour les jours d'attente de la venue d'UN cachalot. Merci **Fabien** pour ton investissement impressionnant (7j/7, H24) pendant les deux missions. Merci **Pauline** et **Marina** pour ce soutien à bord du One cat. Merci à **Pierrick** et **JB** pour leurs précieuses connaissances du drone et d'avoir fait preuve de patience durant les enregistrements en l'absence de cétacé.. Merci à l'ensemble de l'équipage de nous avoir soutenus pendant les deux missions.
- Au bureau d'étude Aquasearch et particulièrement **Benjamin de Montgolfier** qui m'a tout simplement fait découvrir la bioacoustique. Merci de m'avoir permis d'acquérir les enregistrements sur le terrain et les premières analyses des sifflements. Merci à **Aurore** pour son soutien affaibli sur le traitement des sifflements de dauphins de Martinique.
- A **Aurélie Celerier** pour avoir accepté de collaborer et d'intégrer un protocole acoustique pour l'étude des baleines à bosse. Merci d'avoir répondu présente pour la participation de ce Jury de thèse.

- A toute l'équipe et les acteurs du projet CARIMAM, un projet passionnant et inspirant pour de futures études en bioacoustique.
- Merci à l'ensemble de l'équipe Corsecare pour leur soutien et leurs actions en bioacoustique, également pour la protection du littoral Corse.

Merci à **Renata Sousa-Lima** pour sa bonne humeur et ses connaissances durant ses discussions scientifiques, et bien sûr d'avoir accepté de faire parti du jury de thèse.

Merci à **Gianni Pavan** pour ses conseils lors des workshops organisés par l'équipe DYNI et d'avoir accepté de faire parti de ce jury de thèse.

Un immense merci à **Marie Roch** pour nos précieux échanges, particulièrement au début de ce travail.

Une thèse, ce n'est pas uniquement un travail de terrain, mais c'est aussi beaucoup de temps en "post-traitement" pour l'analyse des enregistrements.

J'adresse donc un TRÈS grand merci à mes collègues et amis **Maxence, Paul, Julien, Kevin, Guillaume** et **Vincent** pour m'avoir épaulée durant ces trois ans, et d'avoir contribué à l'amélioration de mes compétences (quasi inexistantes au début de la thèse) en informatique.

Merci à l'ensemble des doctorants (**Aurélie, Caroline, Mélanie, Elora** et tous les autres) et les non permanents du bâtiment X pour m'avoir encouragée durant ces trois ans. Merci pour toutes ces nouvelles amitiés partagées.

Un grand merci à l'ensemble des écovolontaires de la station de recherche Orcalab, particulièrement **Megan** ("my dive buddy") pour son aide précieuse lors des plongées difficiles à Orcalab (eau à 8 degrés, avec du courant et 5m de visibilité). Merci à **Emy, Alicia** et **Chenuye** pour leur fous-rires, **Momoko** et **Tomoko** pour le partage de leurs connaissances infaillibles des vocalises des orques de Colombie Britannique.

Merci pour le soutien de mes amies proches de Savoie et mes amis de La Rochelle. Merci à **Alice** pour ses précieux conseils de motivation et son encouragement quotidien.

Puis, je tiens à remercier particulièrement ma famille :

mon père pour m'avoir fait découvrir et aimer le milieu marin,

ma mère pour son soutien moral quotidien et sa capacité à lire entre les lignes...

Julien pour sa capacité à me changer les idées et sa bienveillance de grand frère,

mes grands parents, ceux encore de ce monde, et ceux partis trop tôt pour avoir cru en mes convictions...

Enfin, je tiens à remercier mon double, ma jumelle **Sophie**, pour tout simplement d'être là dans les bons comme dans les moments difficiles de cette thèse.

Merci aux oiseaux français, aux Orques de Colombie-britannique, aux Cachalots de Méditerranée, aux Dauphins de Martinique et aux Baleines à bosse de Madagascar...

Glossaire

- **ACP** (Analyse en Composante Principale) : Algorithme de réduction de dimension. Méthode linéaire.
- **AUC** (Area Under the Curve) : Aire sous la courbe ROC qui mesure la performance du modèle.
- **BNP clustering** (Bayesian Nonparametric Clustering) : Algorithme de modèle de mélange permettant de modéliser des sous classes parmi une population.
- **CNN** (Convolutional Neural Network) : Réseau de neurones convolutif.
- **Courbe ROC** (Receiver Operating Characteristic) : Courbe représentant les performances d'un modèle de classification binaire (ou de détection), le taux de vrais positifs en fonction du taux de faux positifs.
- **CR** (Call Rate) : Taux d'appel (vocalises, sifflements etc..) par minute.
- **DBSCAN** (Density-Based Spatial Clustering of Applications with Noise) : Algorithme de clustering basée sur la densité estimée des clusters pour effectuer le partitionnement.
- **FFT** (Fast Fourier Transform) : Transformation de Fourier rapide, algorithme de calcul de la transformation de Fourier discrète.
- **Filtre TK** (Teager-Kaiser filter) : Filtre conservant les pics d'énergie.
- **GPU** (Graphical Processing Unit) : Désigne le processeur dédié au traitement des données graphiques pouvant réaliser des calculs parallèles et donc utilisé pour les calculs des CNN.
- **ICI** (Inter Click Interval) : Temps entre les clics (d'échololocation ou codas).
- **IPI** (Inter Pulse Interval) : Temps entre les pulses au sein d'un clic.
- **KDE** (Kernel Density Estimation) : Estimation par noyau : méthode non-paramétrique d'estimation de la densité de probabilité d'une variable aléatoire.
- **MFCC** (Mel-Frequency Cepstral Coefficient) : Représentation court terme du cepstre, en échelle Mel.
- **NAS** (Network Attached Storage) : Appareil de stockage numérique autonome.
- **NMI** (Normalized Mutual Information) : Mesure d'information mutuelle normalisée entre deux distributions.
- **PPS** (Pulse Par Seconde) : Signal électrique satellite précis à la nanoseconde de forme carré transmettant une information de temps à un système (ici nos cartes son pour les synchroniser).
- **RSB** (Rapport Signal sur Bruit) : Indicateur de la qualité du signal analysé.
- **t-SNE** (t-distributed Stochastic Neighbor Embedding) : Technique de réduction de dimension pour la visualisation de données. Méthode non linéaire.
- **TBC** (Taux de Bonne Classification) : Calcul de la performance d'un modèle de classification ou de détection.
- **TDoA** (Time difference of arrival) : Temps de différence d'arrivée (ici entre deux microphones/hydrophones).

Table des matières

1	Introduction Générale	9
1.1	La communication animale	9
1.1.1	La production du son par l'émetteur	12
1.1.2	Le canal de transmission	14
1.1.3	Réception du son	16
1.2	La bioacoustique : une ouverture pour l'étude des espèces animales	18
1.2.1	Des caractéristiques physiques aux caractéristiques biologiques	18
1.2.2	Détection d'espèces : présence/absence	20
1.2.3	Détection de populations	22
1.2.4	Détection d'individus	24
1.2.5	Qualité globale des écosystèmes : Écoacoustique	27
1.3	Problématiques et objectifs de la thèse	29
1.3.1	Organisation de la thèse et objectifs	30
1.3.1.1	Bioacoustique spécifique	30
1.3.1.2	Bioacoustique populationnelle	30
1.3.1.3	Bioacoustique individuelle	30
1.3.1.4	Bioacoustique comportementale : première approche	31
1.3.2	Lecture du manuscrit	33
2	Protocoles d'expérimentations	35
2.1	Présentation des modèles d'étude	35
2.1.1	Les espèces avifauniques	35
2.1.2	Le Cachalot (<i>Physeter macrocephalus</i>)	36
2.1.3	L'Orque (<i>Orcinus orca</i>)	36
2.1.4	Le Dauphin tacheté pantropical (<i>Stenella attenuata</i>)	38
2.1.5	La Baleine à bosse (<i>Megaptera novaeangliae</i>)	39
2.2	Stratégies et techniques d'enregistrements	41
2.2.1	Matériel utilisé pour la détection d'espèces	43
2.2.1.1	Stratégie d'écoute pour la détection d'espèces avifauniques	44

2.2.1.2	Stratégie d'écoute pour la détection de Cachalots (<i>Physeter macrocephalus</i>)	47
2.2.2	Matériel utilisé pour la détection de population	50
2.2.2.1	La station d'enregistrement à Orcalab	51
2.2.3	Matériel utilisé pour l'identification individuelle	53
2.2.3.1	Mise en place d'une antenne fixe pour la localisation individuelle sous-marine	54
2.2.3.2	Mise en place d'une antenne mobile pour la localisation individuelle sous-marine	57
2.2.3.3	Mise en place d'une antenne fixe pour la localisation individuelle en milieu terrestre	59
2.2.4	Matériel pour la bioacoustique comportementale	60
2.2.4.1	Matériel pour la bioacoustique comportementale des Dauphins tâchetés pantropicaux	61
2.2.4.2	Matériel pour la bioacoustique comportementale des Baleines à bosse	63
3	Méthodes pour l'analyse des signaux	67
3.1	Etat de l'art des méthodes supervisées et non supervisées	67
3.1.1	Les méthodes supervisées	67
3.1.2	Les méthodes non supervisées	70
3.2	Méthode pour la détection d'espèces	73
3.2.1	Méthode pour la détection d'espèces avifauniques	73
3.2.1.1	État de l'art	73
3.2.1.2	Le modèle de classification	74
3.2.2	Méthode pour la détection d'Orques (<i>Orcinus orca</i>) à Orcalab	78
3.2.3	Méthode pour la détection de Cachalots (<i>Physeter macrocephalus</i>) sur la bouée BOMBYX	81
3.3	Méthode pour la détection de population	86
3.3.1	Méthode pour la détection de clans acoustiques chez les Orques de Colombie-Britannique	86
3.3.1.1	Méthode supervisée	88
3.3.1.2	Méthode non supervisée	89

3.4	Méthode pour la détection individuelle	90
3.4.1	Méthode pour le suivi individuel d'Orques	90
3.4.1.1	Annotation des enregistrements	91
3.4.1.2	Estimation des azimuts des vocalises	91
3.4.1.3	Analyse semi-supervisée des vocalises	92
3.4.2	Méthode pour la détection individuelle de Cachalot	93
3.4.2.1	La détection de clics	93
3.4.2.2	Estimation du délai d'arrivée	94
3.4.2.3	Localisation et ICI	96
3.4.3	Méthode pour la localisation aviaire	96
3.5	Méthode pour la bioacoustique comportementale	96
3.5.1	Méthode pour la bioacoustique du Dauphin tacheté pantropical	96
3.5.2	Méthode pour la bioacoustique des Baleines à bosse	100
4	Résultats : validation des concepts et tendances bioacoustiques	102
4.1	Détection d'espèces : présence/absence	102
4.1.1	Analyse de la présence d'espèces avifauniques : application Birdify	102
4.1.2	Analyse de la présence d'Orques à Orcalab	106
4.1.3	Analyse de la présence de Cachalots sur la bouée Bombxy	109
4.2	Détection de population	118
4.2.1	Classification de clans d'Orques	118
4.3	Analyse de mouvements individuels	120
4.3.1	Localisation d'Orques pour l'identification individuelle sonore	121
4.3.1.1	Premier cas d'étude (30 août)	123
4.3.1.2	Deuxième cas d'étude (31 août)	126
4.3.1.3	Troisième cas d'étude (11 août)	130
4.3.2	Localisation de Cachalot depuis le drone Sphyrna	134
4.3.3	Identification individuelle aviaire	136
4.4	Résultats en bioacoustique comportementale	138
4.4.1	Impact du trafic maritime sur le Dauphin tacheté pantropical	138
4.4.2	Réponse acoustique des Baleines à bosse en fonction d'un stimulus chimique	142
5	Discussion générale	144
5.1	Avancées méthodologiques et bioacoustiques	144
5.1.1	Avancées méthodologiques	147
5.1.2	Tendances bioacoustiques	150

5.2	Comparaisons : signaux, milieux et méthodes	152
5.2.1	Les types de signaux étudiés	152
5.2.1.1	Les signaux stationnaires	153
5.2.1.2	Les signaux transitoires	154
5.2.2	La nature du canal de transmission : terrestre et marin	155
5.2.2.1	Le milieu terrestre	156
5.2.2.2	Le milieu marin	158
5.2.3	Comparaisons des apports pour les méthodes supervisées et non supervisées en bioacoustique	159
5.3	La bioacoustique appliquée sur l'ensemble des échelles	164
5.4	L'importance de l'échelle temporelle d'analyse	166
5.5	Représentation synthétique des questionnements en bioacoustique	167
5.6	Conclusions	170
5.7	Perspectives	171
5.7.1	Perspectives méthodologiques	171
5.7.2	Perspectives en bioacoustique	172
5.7.2.1	Le projet CARI'MAM	173
Annexes		176
A.	Liste des espèces aviaires étudiées	176
B.	Réponse fréquentielle des hydrophones utilisés	179
C.	Exemple d'un passage de deux matriline (A54 et A50)	188
D.	Schéma représentant les dimensions des différentes antennes mises en place.	189
E.	Schéma représentant la normalisation des enregistrements réalisés sur Bombyx pour la détection de Cachalots	190
F.	Métriques des performances en détection et classification	191
Liste des publications de l'auteur		213
	Articles de conférence	213
	Articles de Revue	214
	Chapitre dans ouvrage	214
	Rapports scientifiques	214
Table des figures		215
Liste des tableaux		222
	Résumé	225
	Abstract	226

Chapitre 1

Introduction Générale

1.1 La communication animale

Depuis les travaux de Darwin ou encore de Tinbergen (1948), la communication animale est au centre de beaucoup de disciplines comme la physique, la chimie, l'écologie, l'éthologie ou encore la bioacoustique.

Un rugissement de lion, le chant d'une baleine ou d'une cigale font partie de la communication, mais également la couleur rouge vive d'une araignée sauteuse, la danse des abeilles ou encore la sécrétion chimique de phéromones par les insectes.

L'étude de la communication animale peut maintenant être au service de la conservation des populations et au maintien de la biodiversité. La communication peut être un outil pour accéder aux informations telles que l'activité de l'individu, son rythme de vie, ses interactions avec son écosystème et ses congénères.

La communication repose sur l'émission d'informations appelées **signaux**. Ces derniers peuvent être divisés en deux grandes familles qui sont les signaux **transitoires** et **stationnaires**.

Les signaux transitoires sont très courts au niveau temporel (quelques ms) et large en fréquence. On peut prendre pour exemple les clics émis par les cétacés. A l'inverse les signaux stationnaires sont plus longs (quelques secondes) et moins larges en fréquences (exemple : sifflements et vocalisations).

Pour les transitoires, un faible décalage temporel provoque de grandes variations fréquentielles au sein du signal, tandis les signaux stationnaires restent stables en fréquence. Le signal est produit de manière volontaire et possède assurément un émetteur et un récepteur. La communication peut être mise en place grâce au codage de l'information dans ce signal, il y a donc uniformisation de celle ci au sein d'une même espèce.

La définition de la communication animale repose sur la transmission du signal d'un individu à un autre individu via un canal de transmission Shannon (1948) ; Green and Marler (1979). Le schéma en figure 1.1 résume les grands principes de la communication animale.

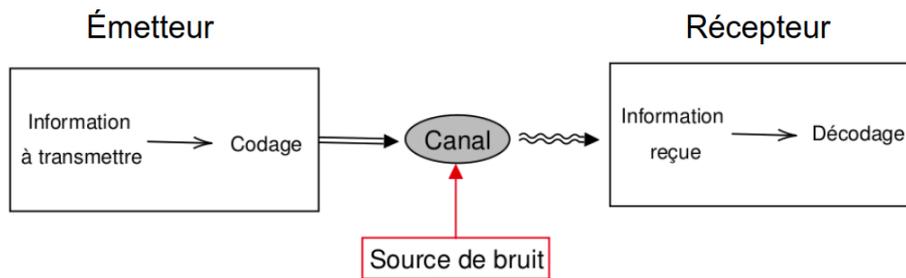


FIG. 1.1 – Schéma global de la structure de la communication animale selon Shannon (1948).

Le point de départ de toute communication est une information à transmettre, un message. C'est l'émetteur qui va coder ce message en signal, puis utiliser un canal adapté pour la bonne transmission de ce dernier.

Durant sa propagation dans le canal de transmission, le signal est altéré de plusieurs façons. Le canal va être soumis à une ou plusieurs sources de bruits qui vont plus ou moins modifier le signal envoyé par l'émetteur. L'altération du signal dépend également des propriétés de celui-ci. Enfin, ce signal va être détecté par un récepteur qui va décoder le message à partir du signal reçu. Durant les étapes de codage, de transmission, de décodage et d'interprétation, le support de l'information subit de nombreuses transformations. C'est principalement pour cette raison que la communication animale regroupe plusieurs disciplines scientifiques.

Le signal envoyé peut être transmis sous différentes modalités.

La première est la lumière relative à la **communication visuelle**.

Ce type de communication est très utilisé. Certains comportements sont développés dans le but de transmettre un signal visuel comme le développement de parades nuptiales ou de postures particulières.

Certaines espèces peuvent également communiquer grâce à certains signaux qu'elles envoient grâce à des pigments de couleurs particuliers. Prenons l'exemple de certains amphibiens, reptiles ou poissons qui ont la capacité de changer de couleur au cours leur vie. La diversité des organes d'expression et de réception visuelle est très impressionnante chez les animaux.

L'exemple le plus simpliste est physique : les ondes lumineuses perceptibles diffèrent entre les groupes taxonomiques ainsi de nombreux oiseaux et insectes peuvent percevoir les ultraviolets tandis que de nombreux reptiles détectent les longueurs d'ondes infrarouges. Mais cette modalité a ses limites.

En effet, les individus utilisant cette communication doivent être proches les uns des autres afin de transmettre une information précise. On peut prendre l'exemple des manifestations visuelles de défense ou d'agressivité de certains prédateurs (loups, ou félins).

Le deuxième type de modalité est la **communication chimique** qui repose sur la transmission d'information entre deux individus via l'envoi de certaines substances chimiques (Alberts, 1992). Cette modalité est la plus ancienne méthode de communication. En effet, les organismes unicellulaires ont la capacité de choisir différents éléments chimiques nécessaires à leur développement (Taga and Bassler, 2003). La communication chimique se fait par le transfert de molécules sur toute la distance entre l'émetteur et le récepteur. Ce mouvement physique peut

être achevé grâce à trois types de transmissions : le flux de courant (air, eau), par diffusion (différence de concentration entre deux systèmes) ou par contact direct entre les individus. Pour beaucoup d'espèces la reproduction est basée sur une communication chimique via l'expression de phéromones.

Certains oiseaux utilisent des signaux chimiques pour se reconnaître entre eux. C'est le cas de la Pétrel bleue (*Halobaena caerulea*) qui développe des signatures individuelles chimiques pour se différencier des autres individus (Célérier et al., 2011).

Il existe également des modalités de communication de courte portée comme le toucher, les stimuli hydrodynamiques, la communication par ondes vibratoires ou électriques. Ces trois types de modalités nécessitent soit un contact entre les individus soit le contact avec un fluide placé entre eux.

Les requins et les raies possèdent des molécules possédant des capacités d'électroréception qui ont pour but de maintenir les distances entre les individus (Kajiura and Holland, 2002). Certains mammifères marins (phoques, lions de mer et otaries) possèdent des vibrisses faciales indispensables à la transmission de l'information via des ondes vibratoires (Marshall et al., 2006). Enfin, la dernière modalité de communication qui fera l'objet de cette thèse est la **modalité acoustique**. Cette dernière est très utilisée chez les humains avec la modalité visuelle notamment. En effet les humains utilisent ce type de communication quotidiennement et beaucoup d'espèces animales disposent de la capacité à émettre et percevoir des sons. Notre planète est constituée de matière ou de fluides qui sont essentiels pour la propagation du son. Les espèces vont utiliser ces matières pour transmettre une information sous la forme d'une onde acoustique. Cette propagation peut être de courte portée (quelques centimètres pour les insectes, Romer and Lewald (1992)), mais également de longue portée (plusieurs centaines de kilomètres pour les mammifères marins (Naguib and Wiley, 2001)).

Le signal peut donc prendre la forme d'une diversité de modalités qui sont la modalité visuelle, chimique, acoustique ou vibratoire. Ces modalités sont utilisées à courte portée (toucher, quelques centimètres) jusqu'à plusieurs dizaines de kilomètres (signaux acoustiques). Elles peuvent évoluer spatialement et temporellement.

Il est possible pour certaines espèces d'utiliser plusieurs modalités en même temps : c'est la **multimodalité sensorielle** (Partan and Marler, 2005). Ces espèces peuvent utiliser différents sens (vue, toucher, chémoréception etc...) en même temps pour transmettre une information.

Prenons l'exemple de certaines espèces d'anoures qui utilisent la multimodalité sensorielle pour sélectionner un partenaire sexuel (acoustique et visuel (Gomez et al., 2011)) ou encore la multimodalité sensorielle chez les chimpanzés pour communiquer des informations notamment sur l'alimentation (Leavens et al., 2010). de Jong et al. (2018) ont réussi à associer l'émission sonore d'un poisson avec une posture particulière (communication visuelle).

La communication repose sur des modalités sensorielles liées à l'individu (capacité physique) mais aussi liées au milieu et au mode de vie des espèces.

Je vais développer ci-après les trois étapes de la communication animale acoustique, qui sont : la production du son par l'émetteur, sa transmission via le canal de transmission et enfin la réception du son par le récepteur.

Je vais détailler plus particulièrement le cas des oiseaux et des cétacés, familles très actives en terme de communication sonore et qui feront donc l'objet de cette thèse.

1.1.1 La production du son par l'émetteur

La faune peut produire des sons pour l'interaction au sein d'un groupe, pour localiser ses congénères, ou encore pour se défendre des prédateurs. Beaucoup d'espèces ont été découvertes grâce à leurs émissions sonores. L'émission sonore peut être de plusieurs types. Le premier est la production de vibrations. Elle est beaucoup utilisée pour des espèces de petites tailles, ne communiquant que sur de petites distances. Il existe quatre grandes catégories de mouvements qui peuvent entraîner une vibration :

- mouvement d'un corps solide sur un autre corps solide (percussion, très utilisé chez les arthropodes)
- mouvement d'une partie du corps pour créer des ondes dans un fluide ou à l'interface entre deux fluides (exemple : l'araignée d'eau)
- mouvement d'un fluide contre une partie du corps (cordes vocales).

Grâce à ces différents mécanismes, les espèces produisent des sons plus ou moins aigus et plus ou moins forts. La figure 1.2 illustre les productions sonores de 0.1Hz à 1GHz.

L'espèce humaine est capable d'entendre entre 20Hz à 20 kHz (figure 1.2, ce qui ne correspond qu'à une infime partie de l'ensemble des sons produits par les écosystèmes (biotiques et abiotiques).

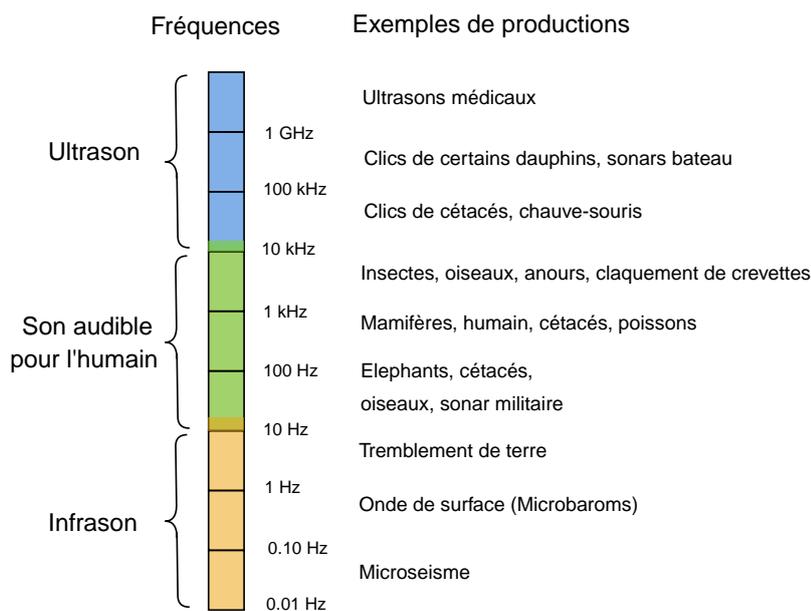


FIG. 1.2 – Exemple de production acoustique de 0.1Hz à 1GHz en fonction des capacités de l'oreille humaine (Bradbury et al., 1998).

Cette thèse portera essentiellement sur les productions acoustiques des oiseaux et des cétacés. Dans un premier temps, je développerai les modalités de l'émission sonore chez les oiseaux.

Les oiseaux sont connus pour être des chanteurs aguerris. C'est pour cette raison que beaucoup de scientifiques s'y intéressent (Kumar, 2003 ; Todt and Naguib, 2000). En effet, bon nombre de bioacousticiens travaillent sur les oiseaux, car ce sont des espèces faciles d'accès et qui peuvent produire une variété très importante de sons.

En effet, les oiseaux utilisent des signaux vocaux quotidiennement pour divers objectifs, comme le maintien des couples, les interactions parents/jeunes ou encore la cohésion d'un groupe. Ces signaux peuvent être très complexes et peuvent durer dans le temps (cf Fig.1.3).

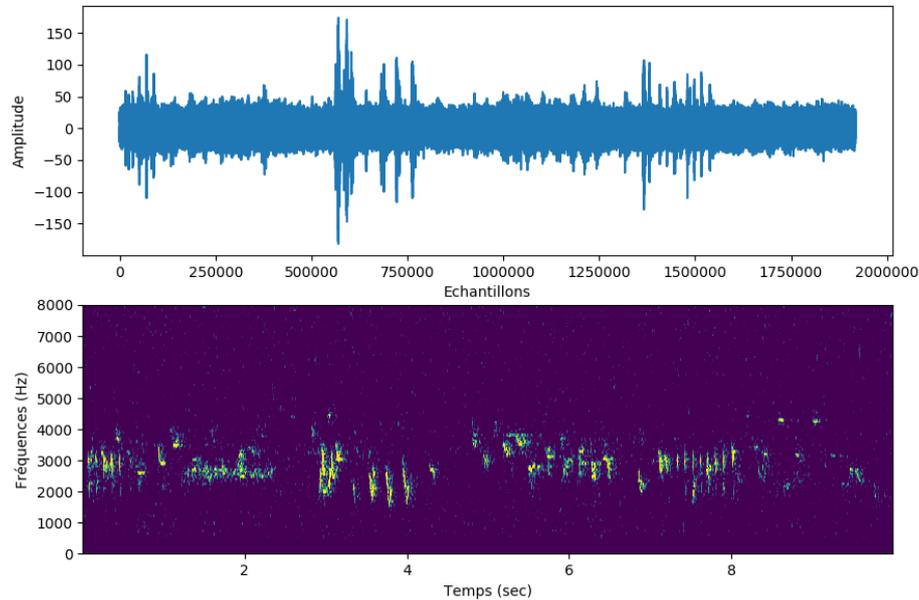


FIG. 1.3 – Forme d'onde et spectrogramme de 8 secondes de chants d'oiseaux enregistrés au jardin du Rayol (Var) en 2019 par l'équipe DYNI (micro et carte JASON de SMIoT).

La production acoustique chez les oiseaux peut se caractériser par deux principaux signaux : les cris et les chants (Catchpole and Slater, 2003). Les cris sont très courts, simples et peuvent être produits toute l'année par le mâle et la femelle alors que les chants sont produits généralement par les mâles en période de reproduction, sont plus longs et très complexes. Les cris sont essentiels pour communiquer quotidiennement, pour informer la présence d'un prédateur tandis que les chants ont pour principale fonction d'attirer les femelles avant la formation d'un couple. L'avifaune est donc spécialisée dans la production du son (Suthers, 2004).

Pour cela, tous les oiseaux possèdent un organe spécialisé appelé la Syrinx qui se situe dans la trachée, au fond de la poitrine (Elemans et al., 2015). Sa fonction et sa position sont identiques chez toutes les espèces d'oiseaux, mais sa structure diffère énormément d'une espèce à une autre (Suthers, 2004).

Lors de l'expiration, l'air exerce une pression sur cette membrane, qui la fait vibrer, et produit des ondes sonores.

L'air peut provenir de chacune des deux bronches (droite et gauche), et donc l'animal peut altérer l'usage des notes produites de droite, et de gauche (Zollinger et al., 2008).

Le bec de l'animal va jouer également un rôle important dans l'émission et la propagation du signal (Fletcher and Tarnopolsky, 1999).

Chez les mammifères marins, le système sensoriel leur permet de recevoir et d'interpréter des sons dans l'eau. En effet, les cétacés doivent se repérer, capturer des proies, communiquer, éviter les prédateurs dans un milieu où parfois la visibilité est limitée à quelques mètres. La propagation du son dans le milieu aquatique est 4.5 fois plus rapide que la propagation dans l'air, et offre beaucoup moins d'atténuation en fonction de sa distance d'émission. Cette caractéristique a provoqué une incidence sur la morphologie et la physiologie des mammifères marins dans la production et réception du son. Le son est utilisé à différentes échelles spatiales, de quelques mètres pour la cohésion d'un groupe, et de plusieurs kilomètres pour certaines espèces (Rossing, 2007). Les cétacés, particulièrement les odontocètes utilisent le son quotidiennement et sont capables de produire une large gamme de fréquences pour la communication et l'écholocation.

Les odontocètes peuvent produire différents types de sons, qui sont également très adaptés au milieu aquatique : les vocalises/sifflements, les séries dense de clics ("burst") et les clics.

La source réelle et le fonctionnement de la production des vocalises/sifflements/burst chez les odontocètes sont encore controversés. Actuellement, beaucoup de scientifiques prônent la théorie suivante : la production du son proviendrait des lèvres phoniques ou museau de singe (Frankel, 2009).

Les clics eux, sont produits principalement pour la recherche de proies et parfois pour la communication (Cachalot). Tous les odontocètes émettent des clics pour chasser, c'est l'écholocation. L'animal émet un signal acoustique, ce signal atteint sa proie et l'animal analyse l'écho de celui-ci. Le système d'écholocation peut être résumé en trois étapes principales : la transmission, la réception et la décision. Ces trois étapes sont réalisées très rapidement. L'émission du clic consiste en sa production, son amplification et sa capacité à traverser le milieu. Sa réception se fait grâce au système auditif et le traitement de l'information (décision) se fait grâce au système nerveux périphérique et central (Rossing, 2007). Chez la plupart des cétacés, les clics sont produits par des sacs d'air dans le conduit nasal situé sous l'évent, une fois produit le son est envoyé vers le melon (avant de la tête) qui contient des lipides spéciaux qui vont moduler le son et le rendre très directif (Au, 1993).

Chez les cachalots, les clics sont également produits dans le conduit nasal à l'avant de la tête de l'animal, puis sont projetés à l'arrière de la tête rebondissant sur un sac d'air et sont renvoyés vers l'avant traversant le "junk".

1.1.2 Le canal de transmission

La vibration, ou pulsation produite par les capacités physiques de l'animal va avoir une influence sur le milieu dans lequel il se trouve. En effet, chaque fluide qui nous entoure (eau ou air) est composé d'atomes qui interagissent entre eux.

En vibrant, une membrane génère une perturbation régulière qui se propage dans le fluide en faisant déplacer ses atomes les uns après les autres, formant ainsi une onde sonore. Plus la vibration de la membrane est puissante, plus les atomes vibrent avec une grande amplitude et plus le volume du son est élevé. L'animal peut également agir sur la fréquence de la vibration. Plus celle-ci est rapide, plus le son perçu est aigu. L'oreille humaine est capable de percevoir des vibrations entre 20 et 20 000 fois par seconde (20 Hz à 20 kHz).

Les sons produits avec une fréquence inférieure à 20Hz sont appelés infrasons et les sons produits avec une fréquence supérieure à 20 000 Hz sont appelés ultrasons. L'environnement a donc une réelle importance dans la transmission sonore. Le premier facteur influençant cette transmission est l'état de la matière dans laquelle le son se propage. Le son peut se propager dans des gaz (air), des liquides (eau) ou des solides (roche, branche, tronc). Ici, nous nous focaliserons davantage sur la propagation du son dans l'air ou dans l'eau.

Au sein de ces deux milieux, certains paramètres vont influencer la bonne propagation du son. Ces paramètres peuvent être la température, la pression, le vent, le courant ou la présence d'obstacles.

Les signaux peuvent donc parfois être des indicateurs des caractéristiques physiques du milieu. En effet, les espèces sont obligées d'adapter leur communication afin d'optimiser la réussite de transfert des messages. Nous pouvons citer ici quelques exemples, comme certains cétacés qui peuvent émettre des chants à plusieurs centaines de kilomètres pour communiquer avec un congénère (Payne and Webb, 1971). L'éléphant est capable également de produire des sons inférieurs à 20 Hz pour communiquer avec à plusieurs dizaines de kilomètres (O'Connell-Rodwell et al., 2000).

Lorsque les signaux acoustiques se propagent, ils subissent diverses transformations comme (O'Connell-Rodwell et al., 2000 ; Wiley, 1982) :

- la réduction de l'intensité sonore (Bardeli et al., 2010). Les ondes sonores vont perdre en énergie en fonction du temps et de la distance qu'elles parcourent, c'est la dissipation. Plus l'espace actif (l'espace dans lequel le récepteur est en mesure de discriminer l'information transmise par le signal) est grand, plus le signal sera réduit en intensité. Cette atténuation du signal émis peut être anticipée par les espèces émettrices (Marten and Marler, 1977).
- le contenu fréquentiel du signal. Certaines fréquences (basses fréquences) ont tendance à être utilisées pour des communications de longue portée, car elles sont moins déformées lors de leur propagation (Chapuis, 1971).

L'efficacité de la communication peut également dépendre du bruit ambiant. Le degré pour lequel le bruit pourrait masquer la communication dépend de son amplitude, de sa fréquence, de sa distribution spatiale et temporelle. La source du bruit ambiant peut être non biologique (vents, pluie), biologique (bruit ambiant des animaux la nuit) ou d'origine anthropique.

Dans le milieu terrestre et marin, le bruit ambiant a tendance à devenir de plus en plus important (Brumm and Slabbekoorn, 2005 ; Klump and Shalter, 1984 ; Ryan and Brenowitz, 1985 ; Haren, 2007). En milieu terrestre, le vent peut générer du bruit entre 100 et 1000 Hz jusqu'à 70 dB pour les vents forts, tandis que la pluie peut atteindre 6 kHz à 35 dB (Lengagne and Slater, 2002).

Comme les autres sons, le bruit se propage également dans l'eau sans difficulté, particulièrement pour les basses fréquences. En effet, l'activité sismique de la planète peut produire des sons entre 10 et 100 Hz pouvant être transmis à des milliers de kilomètres. Puis de 100 à 5000 Hz le bruit provient particulièrement des vagues et du vent de surface (Richardson et al., 2013). Les humains sont en majorité responsables du bruit ambiant dans les océans, particulièrement causé par le trafic maritime (Simmonds et al., 2004). En effet, le bruit des hélices des moteurs peut nuire à la bonne communication de certaines espèces, ceci est dû au masquage des fréquences. Afin de lutter contre ces nuisances sonores, les espèces ont différents choix :

- augmenter l'amplitude de leur signal (effet Lombard, Manabe et al. (1998))
- augmenter la durée du signal (Brumm et al., 2004)
- augmenter le taux de répétition du signal (Payne and Payne, 1998 ; Lengagne et al., 1999)
- changer la fréquence d'émission (La Manna et al., 2016).

Le bruit a donc une réelle influence sur le canal de transmission, et peut être une forte contrainte pour la survie des individus. La nuisance sonore peut avoir des influences sur l'émission, la réception et enfin la mémorisation du signal.

1.1.3 Réception du son

Pour pouvoir détecter ces vibrations sonores, les cétacés posséderaient une membrane spécialisée appelée le tympan située à l'avant de l'oreille moyenne (comme chez l'humain). Le déplacement des atomes d'air ou d'eau va entraîner la vibration de cette membrane. L'oreille moyenne comporte trois os : le marteau, l'enclume et l'étrier, formant la chaîne ossiculaire. Cette chaîne relie le tympan à l'oreille interne. Les vibrations du tympan vont se propager et être amplifiées via cette chaîne d'os, faisant le lien entre le tympan et la cochlée. Cette dernière est constituée de tissus et de liquides. Le liquide va vibrer et entraîner le mouvement des cils des cellules ciliées.

En fonction des fréquences émises certaines cellules vont être activées et d'autres non. Ces cellules sont placées de manière très ordonnée sur la cochlée. Celles situées sur la base sont activables avec des hautes fréquences, tandis que la partie apicale (extrémité cellulaire vers le milieu extérieur) code pour les basses fréquences (tonotopie). Cette activation va provoquer la génération d'un signal électrique qui sera envoyé jusqu'au cerveau afin de lui transmettre cette information, c'est le décodage de l'information (Nummela, 2009 ; Fleischer, 1973).

Une des différences constatées avec les humains est l'absence de pavillon externe. Depuis plusieurs dizaines d'années les scientifiques essaient de percer le mystère de la transmission sonore vers le système tympanique. Certains affirment que les sons sont reçus via la mandibule inférieure, celle ci serait composée de systèmes graisseux qui transmettraient l'onde sonore (Brill et al., 1988). D'autres équipes émettent l'hypothèse qu'il existerait d'autres tissus adipeux nécessaires à la réception sonore (Perrin et al., 2009). Il existerait un canal interne adipeux permettant de conduire les ondes sonores jusqu'au tympan et que les cétacés posséderaient des adaptations supplémentaires aux autres mammifères terrestres.

Les cétacés et particulièrement les dauphins peuvent être sensibles à des sons atteignant 100 kHz (Ridgway, 2000), ils possèdent donc des adaptations physiologiques leur permettant d'entendre des sons plus ou moins distants et de larges bandes de fréquences. Le premier audiogramme a été réalisé par Johnson (1967) et d'autres espèces ont été étudiées jusqu'à nos jours comme le Dauphin bleu et blanc (*Stenella coeruleoalba*) qui peut entendre jusqu'à 160 kHz.

La figure 1.4 représente les audiogrammes des différentes espèces de mammifères marins et notamment des cétacés. Ils sont sensibles à des fréquences allant au delà de 100 kHz.

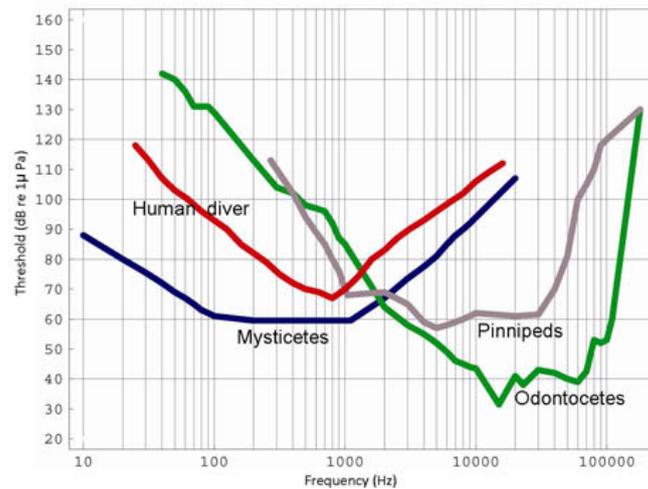


FIG. 1.4 – Exemple d'audiogramme de différents mammifères marins (odontocètes, mysticètes, pinnipèdes) et des humains (NOAA, 1998).

Concernant les oiseaux, leurs oreilles sont très bien adaptées à la réception du son aérien. Ils sont spécialisés pour des fréquences entre 1 kHz et 5 kHz et peuvent entendre des sons de 0.25 kHz allant jusqu'à environ 10 kHz (Dooling et al., 2000).

Ils ne possèdent pas de pavillon externe (excepté pour les rapaces diurnes). L'oreille externe consiste uniquement à un très fin conduit auditif caché sous le plumage. Les oiseaux chanteurs possèdent deux membranes tympaniques, une externe (qui sert de protection) et une interne. Le tympan est relié à un seul os appelé columelle qui est relié à la cochlée. Cette dernière a globalement la même forme que celle des mammifères et est composée de structures sensorielles (cellules ciliées). Ces cellules, une fois activées vont également transmettre une information au cerveau sur la structure du son qui a été reçu. Dans les années 70, (Nottebohm et al., 1976) ont démontré que les oiseaux (chanteurs) possèdent des aires cérébrales spécialisées dans le traitement, la production, la mémorisation et l'apprentissage des chants. Des études comparées avec l'audition des mammifères existent en terme de capacités cognitives (Deecke, 2006).

La réception des sons est primordiale chez beaucoup de familles, elle permet d'identifier le locuteur, d'éviter les prédateurs ou encore d'identifier les proies.

La notion de mémorisation et d'apprentissage est donc indispensable pour la reconnaissance d'une structure sonore.

Les espèces vont donc utiliser leur chaîne de perception sonore pour détecter, identifier, évaluer et localiser la source d'émission. Cette définition de la communication acoustique repose donc sur l'émission d'un message sous forme d'une onde sonore (codage) qui évolue via un canal de transmission et qui sera décodé grâce à l'oreille et le cerveau du receveur.

1.2 La bioacoustique : une ouverture pour l'étude des espèces animales

Les interactions au sein d'un écosystème sont complexes géographiquement et temporellement : les individus évoluent au sein de celui-ci en gardant un équilibre essentiel à sa survie. La vulnérabilité de ces écosystèmes face aux diverses pressions anthropiques n'est pas négligeable. Il est donc nécessaire de réaliser des études et des suivis de leur évolution.

La surveillance par acoustique est une méthode efficace pour étudier et classifier un écosystème. Ainsi, nous pouvons définir la bioacoustique comme une science qui étudie la production, l'émission et la réception du son pour tout être vivant. Afin d'obtenir des informations sur l'état de santé des espèces et des populations, les écologues utilisent différents types d'acquisitions de données comme les observations visuelles, les méthodes de captures-recaptures (CMR), ou encore des analyses génétiques. L'arrivée récente de la bioacoustique propose une alternative à ces méthodes qui peuvent être complexes et onéreuses.

La bioacoustique présente plusieurs avantages : premièrement, la visibilité des espèces n'est pas nécessaire pour leur étude (conditions nocturnes, profondeur). Puis, il est possible de détecter des espèces acoustiquement à des distances beaucoup plus lointaines que visuellement. Enfin, l'émergence des nouvelles technologies permet à la bioacoustique de développer de nouvelles méthodes d'analyses des signaux de manière automatique. Elle permet d'étudier diverses problématiques comme la détection, la classification, la localisation et le suivi des sources sonores, mais également l'estimation des densités de population et l'impact de l'Homme sur les espèces.

Cette deuxième partie introduira les différentes échelles d'études de la bioacoustique qui sont :

- les caractéristiques biologiques des espèces
- le suivi d'espèces
- le suivi des populations
- le suivi d'individus
- l'écoacoustique.

1.2.1 Des caractéristiques physiques aux caractéristiques biologiques

Comme présenté dans la partie 1.1.1 "Production du son par l'émetteur", les espèces produisant des sons possèdent des capacités physiques bien adaptées.

Il est donc possible d'établir un lien entre l'émission acoustique et les capacités physiques/biologiques de l'animal.

Premièrement, les caractéristiques globales du son émis dépendent de la structure de l'appareil phonatoire de l'espèce et il y a donc une composante phylogénétique à cela. Les espèces d'une même famille peuvent avoir développé des compétences phonatoires particulières.

Lorsque nous faisons des analyses plurispécifiques des signaux acoustiques, la phylogénie est souvent prise en compte (McCracken and Sheldon, 1997 ; Gingras et al., 2013).

Ensuite, certaines caractéristiques sonores nécessitent des compétences particulières et il existe alors une variance intra-spécifique comme par exemple l'utilisation de celle-ci par les femelles pour sélectionner certains mâles. Ce lien peut être mécanique ou "automatique" (lien très fréquent entre la fréquence du signal et la taille de l'émetteur). Par exemple le timbre de la voix

n'est pas contrôlé mais peut nous donner une information sur le genre (Féminin ou Masculin) de la personne, quel que soit le message.

Différents scientifiques ont déjà associé les émissions acoustiques avec certains paramètres, c'est le cas de Voituron et al. (2012) qui a fait le lien entre les productions acoustiques des rainettes vertes avec la grosseur de l'animal et ses compétences énergétiques. Les femelles arrivent à sélectionner les mâles de grosse taille, ne consommant que peu d'oxygène, uniquement avec leur émission sonore.

Desprat et al. (2015) ont prouvé que l'expression sonore chez la rainette verte était liée à la quantité de testostérone chez les mâles. Chez les oiseaux également, la performance acoustique va dépendre des capacités physiques de l'animal. En effet, la taille du bec et de la gorge vont influencer les émissions sonores (Fletcher and Tarnopolsky, 1999). Grâce aux émissions il est donc possible d'identifier ces capacités physiologiques.

Ces affirmations peuvent également être appliquées pour la faune marine et plus particulièrement pour le Cachalot.

Le Cachalot, comme tout odontocète produit des clics pour chasser, mais également pour communiquer (coda). Les clics transitoires des Cachalots possèdent une structure multipulsée liée aux résonances de l'émission dans la tête de l'animal (Zimmer et al., 2005).

Le Cachalot produit une impulsion initiale à l'avant de la tête (P₀) ensuite, cette impulsion se propage à travers sa tête (vers l'arrière) et est ensuite renvoyée vers l'avant (P₁). Il est donc possible de définir l'Intervalle Inter-Pulse (IPI) qui correspond au temps de parcours du son à l'intérieur de la tête de l'animal. Il existe donc plusieurs lois allométriques estimant la taille de la tête, et donc la longueur du Cachalot en fonction de l'IPI (Gordon, 1991 ; Growcott et al., 2011).

La figure 1.5 montre la structure multipulsée d'un clic de Cachalot ainsi que l'IPI.

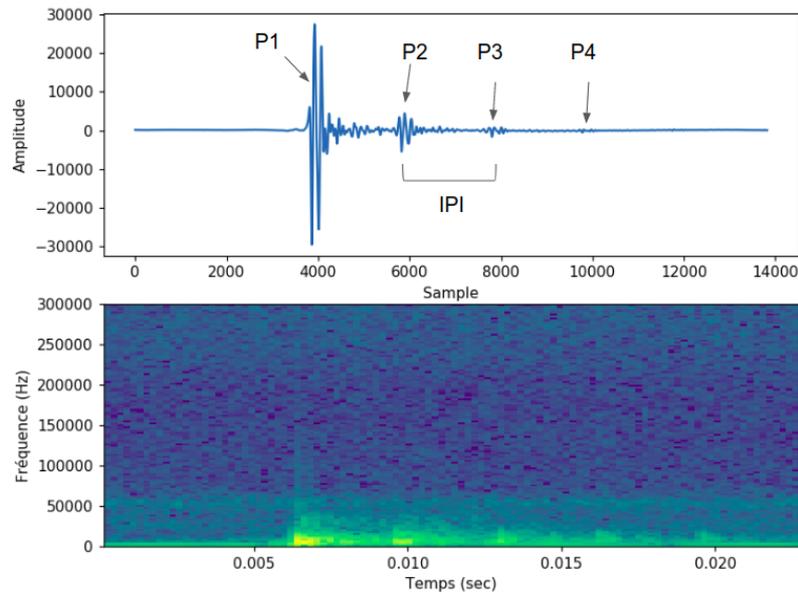


FIG. 1.5 – Représentation d’un clic de Cachalot enregistré à l’île Maurice en 2019. Haut : forme d’onde. Bas : spectrogramme (fenêtre de Hanning, taille de la FFT 256) . Protocole réalisé par l’équipe DYNI (Université de Toulon), carte son JASON SMIoT et Hydrophone C75 Cetacean Research.

Grâce à des relations allométriques, comme celle du Cachalot, il est possible d’identifier la taille de l’animal avec la production acoustique de celui-ci. Pour les animaux produisant des vocalises (présence d’harmoniques), on peut estimer la taille de l’animal avec la fréquence fondamentale (f_0). Garcia et al. (2017) ont démontré que la position de f_0 était corrélée à la taille de l’animal chez plusieurs espèces de primates. Ceci est également valable pour l’humain avec des corrélations significatives entre la taille des hommes avec la position de f_0 (Rendall et al., 2005).

Mais cette corrélation peut aussi être liée à des compétences particulières et volontaires de l’émetteur. C’est la théorie du handicap de Zahavi 1975 démontrant qu’un signal est d’autant plus fiable pour un récepteur qu’il est coûteux pour son émetteur (Zahavi, 1975). Prenons l’exemple du Paon bleu, une femelle (le récepteur) va sélectionner le mâle en fonction de la forme et de la couleur de sa queue. En cas de prédation, plus la queue du paon est belle et imposante, plus il sera difficile d’échapper à ses prédateurs.

Ces lois biologiques peuvent être ainsi utilisées en bioacoustique pour inférer certaines caractéristiques des émetteurs (morphologiques notamment).

La bioacoustique peut donc servir à définir certaines caractéristiques physiques, morphologiques, de l’animal. Ces caractéristiques vont différer en fonction des émissions sonores mais également en fonction des groupes taxonomiques étudiés.

1.2.2 Détection d’espèces : présence/absence

La bioacoustique peut servir à la surveillance des espèces. En effet, certaines espèces sont difficiles d’accès et le manque de données les concernant nuit à leur protection (Laiolo, 2010 ; Blumstein et al., 2011). Les actions de conservation peuvent avoir différents objectifs :

- surveillance des espèces commerciales (Rountree et al., 2006)
- amélioration des connaissances sur l'écologie des espèces et leur distribution (Mellinger et al., 2007)
- lutte contre les extinctions des espèces et la réduction de la biodiversité.

Les méthodes traditionnelles pour étudier la biodiversité et identifier les espèces présentes sont relativement coûteuses, limitées dans l'espace et dans le temps (Parker III, 1991 ; Sueur et al., 2008). Mais surtout, elles ne peuvent être réalisées en continu, contrairement à la bioacoustique. Tout comme les captures vidéos des espèces dans leur milieu naturel, les enregistrements peuvent être utilisés pour identifier la répartition des espèces. Pour cela, il est nécessaire de mettre en place des méthodes de détection adaptées en fonction de l'espèce étudiée.

A ce stade, il est important de souligner qu'un enregistrement dépourvu de signaux ne prouve pas l'absence d'espèce. En effet, beaucoup de taxons peuvent évoluer dans leur écosystème en silence.

La fréquence d'utilisation des signaux acoustiques par les espèces peut évoluer temporellement et spatialement, ce qui rend les détections parfois difficiles.

Nous pouvons prendre l'exemple de Jaramillo-Legorreta et al. (2017) qui a réalisé une étude sur le Marsouin du pacifique (*Phocoena sinus*), mammifère marin le plus menacé des océans (30 individus en novembre 2016). Cette espèce est soumise à diverses pressions anthropiques, notamment les captures accidentelles dans les filets dérivants. Les suivis visuels de cette population endémique sont impossibles car ils passent peu de temps en surface. Mais ce petit cétacé produit des clics très reconnaissables pouvant être détectés automatiquement.

Il a donc été mis en place un réseau d'hydrophones avec un détecteur de clics. Il a été constaté la baisse de 80% des émissions acoustiques en 4 ans. Cette étude a permis d'apporter des données pour la mise en place d'une loi mexicaine pour la protection de cette espèce en déclin. Ce cas concret a illustré l'utilité de la bioacoustique dans le cadre de la protection des espèces, ceci n'est pas un cas isolé.

Les relevés bioacoustiques peuvent être réalisés durant quelques secondes jusqu'à plusieurs années. En effet, il est possible d'identifier la présence d'espèces acoustiquement actives sur une échelle de temps très longue (Pettorelli et al., 2013 ; Sousa-Lima et al., 2018 ; Jeliaskov et al., 2016) où l'accès des espèces est limité (Riede, 1993). Il est donc impossible d'analyser manuellement les enregistrements (Chesmore, 2004). En effet, les analyses manuelles peuvent être chronophages et dépendent beaucoup des connaissances de l'annotateur. Un biais de l'annotation est donc inévitable (Fitzpatrick et al., 2009). De plus, peu d'experts peuvent identifier en toute certitude les animaux sur la base d'un signal acoustique.

Pour beaucoup de taxons, des méthodes automatiques ou semi-automatiques peuvent être utilisées, le tableau 1.1 donne quelques exemples d'études réalisées.

L'approche automatique peut être limitée par le fait que la plupart des utilisateurs (écologues, biologistes) ne possèdent pas de compétences suffisantes en informatique pour développer des algorithmes automatiques d'identification d'espèces. C'est pour cela que le décloisonnement des sciences entre l'informatique et l'écologie est essentiel.

La bioacoustique est une excellente approche pour la détection d'espèces via leur communication. En effet, les stations acoustiques peuvent étudier les espèces présentes 24h/24, tous les jours de l'année dans plusieurs habitats simultanément (Acevedo and Villanueva-Rivera, 2006 ;

TAB. 1.1 – Exemples d'études où l'espèce a été identifiée automatiquement par l'acoustique pour différents taxons.

Taxons	Bibliographie
Orthoptères	Chesmore (2004) ; Riede (1998)
Cétacés	Deecke and Janik (2006) ; Rickwood and Taylor (2008) ; Kimura et al. (2010) ; Morrissey et al. (2006) ; Murray et al. (1998) ; Giraudet and Glotin (2006) ; Roch et al. (2011) ; Halkias et al. (2013)
Oiseaux	McIlraith and Card (1995) ; Selouani et al. (2005) ; Priyadarshani et al. (2018b) ; Glotin et al. (2013b,c) ; Morfi et al. (2019) ; Stowell et al. (2019)
Amphibiens	Shin et al. (2012) ; Camacho et al. (2011) ; Grigg et al. (1996)
Mammifères terrestres	Reby et al. (1997) ; Clemins and Johnson (2002) ; Suter et al. (2017)
Chiroptères	Hill and Hüppop (2008) ; Skowronski and Harris (2006) ; Mac Aodha et al. (2018)
Poissons	Malfante et al. (2018)

Lammers et al., 2008).

1.2.3 Détection de populations

Beaucoup d'écologues, de biologistes et de gestionnaires s'intéressent à la répartition de certaines populations ou peuplements au sein d'une aire géographique précise. Avant d'exposer le rôle de la bioacoustique pour étudier l'évolution des populations, il est nécessaire de définir ce qu'est une population en écologie.

“Une population désigne l'ensemble des individus d'une même espèce qui occupent simultanément un écosystème et qui échangent librement leurs gènes dans un processus reproductif” (Duquet and Thiollay, 1993). Les individus au sein de cette population suivent un certain équilibre entre l'arrivée de nouveaux individus (maintien du patrimoine génétique) et le maintien d'un nombre d'individu suffisant. Afin d'étudier les populations, les scientifiques utilisent régulièrement des études génétiques très lourdes et longues se focalisant uniquement sur un petit groupe d'animaux. La biologie de la conservation repose sur le suivi de la dynamique des populations, s'appuyant sur l'analyse du mouvement, des stratégies de reproduction, et des stratégies comportementales de cette population (McGregor, 2000). En effet, la compréhension des moteurs du changement des populations est essentielle pour la mise en oeuvre des stratégies de conservation. Les changements démographiques au sein de cette population peuvent être utilisés comme indice pour anticiper de futures mesures de protection.

Il est donc possible de collecter des enregistrements d'individus au sein d'une population précise. En effet, si les groupes d'individus sont isolés les uns des autres, il peut exister des variations distinctes dans les émissions de ces groupes : ce sont les dialectes (Catchpole and Slater, 2003 ; Lengagne et al., 2020). Cette notion de “dialecte” a déjà été mise en évidence pour un bon nombre de taxons. Le premier, le plus étudié est celui des oiseaux (Krebs and Kroodsma, 1980). En effet, les espèces aviaires sont spécialisées dans la production sonore, mais également dans leur mémorisation (Kroodsma, 1974). Beaucoup d'études se sont focalisées sur l'évolution des chants mais également sur l'aspect de transmission culturelle de ces derniers (Mundinger, 1975). L'étude de Joachim and Lauga (1996) a mis en évidence une cartographie des dialectes, plus

précisément il a pu faire un lien entre l'éloignement géographique et acoustique des signaux chez le Pinson des arbres (*Fringilla coelebs*). La continuité temporelle et spatiale des groupes de chants a également été observé chez le Créadion rounoir (*Philesturnus carunculatus*) (Jenkins, 1978). Chez les moineaux, Marler (1987) a démontré que des oiseaux isolés possédaient des chants moins complexes que des oiseaux mobiles. Mais l'acquisition des dialectes chez les oiseaux est très complexe, et particulièrement difficile à identifier, car elle dépend de nombreux paramètres dont :

- la capacité des jeunes à apprendre les différenciations acoustiques
- le temps d'apprentissage (au moins deux premiers mois de la vie) (Marler and Peters, 1982), variable en fonction des espèces
- la qualité de l'apprentissage chez le parent (Clayton, 1988)
- la complexité acoustique du chant
- la capacité des jeunes mâles à émettre ce qu'ils ont appris (lié à la testostérone).

Chez les mammifères marins la notion de dialecte est également fréquente (Conner, 1982). Les différents dialectes ont également été utilisés pour étudier les populations chez la Baleine à bosse (Winn et al., 1981). Des variations géographiques d'émissions ont été également observées chez les pinnipèdes comme le Phoque de Weddell (Thomas and Stirling, 1983) et le Phoque barbu (Risch et al., 2007). Les mécanismes culturels de transmission sont notamment très importants lorsque les populations suivent une hiérarchie matriarcale, c'est le cas des Orques de la Colombie-Britannique (Canada). Les Orques fréquentant ces eaux sont divisées en deux communautés (Nord et Sud), celles-ci sont divisées en clans acoustiques, eux-mêmes divisés en pods. Tous les pods ont des caractéristiques distinctives dans leurs vocalisations. Il existe donc un dialecte précis pour chacun des pods (Ford, 1991).

L'analyse des dialectes chez les cétacés ne se résume pas uniquement à l'analyse des vocalisations, mais également à l'analyse de certains clics de communication émis notamment chez le Cachalot. En effet, les Cachalots produisent des codas pour communiquer (succession de plusieurs clics). Rendell and Whitehead (2003) ont comparé les codas produits par différents groupes de Cachalots à travers le monde. Ils ont identifié 6 clans acoustiques (1 dans les Caraïbes et 5 dans le Pacifique).

Les codas des Cachalots méditerranéens sont également bien distincts (Pavan et al., 2000).

L'acoustique peut donc répondre à certaines problématiques, notamment la détection de mouvements d'une population pour des espèces acoustiquement actives. La notion de dialecte passe par certains principes de reconnaissance.

Ces principes peuvent reposer sur des phénomènes biologiques permettant à un individu de pouvoir identifier les membres de sa population.

Nous pouvons supposer que plus le nombre de classes (populations) est élevé, plus le signal émis sera complexe. Les mécanismes de reconnaissance des populations peuvent être divers mais ils présentent tous le même processus, résumé Fig.1.6.

Un individu acquiert tout au long de sa vie les capacités à différencier les membres de sa population grâce au développement d'une "liste" de cibles possibles" (1). Puis lorsqu'il reçoit une information d'un congénère (2), il va comparer cette information avec les éléments de sa "liste" (3). Ensuite, en comparant le signal reçu avec ceux de sa mémoire, il va pouvoir prendre une décision pour attribuer l'individu à sa population (4). La dernière étape consiste à répondre à

son interlocuteur (5). Ces cinq étapes évoluent dans le temps, et les individus améliorent leurs mécanismes de reconnaissance au cours de leur vie (Bradbury et al., 1998).

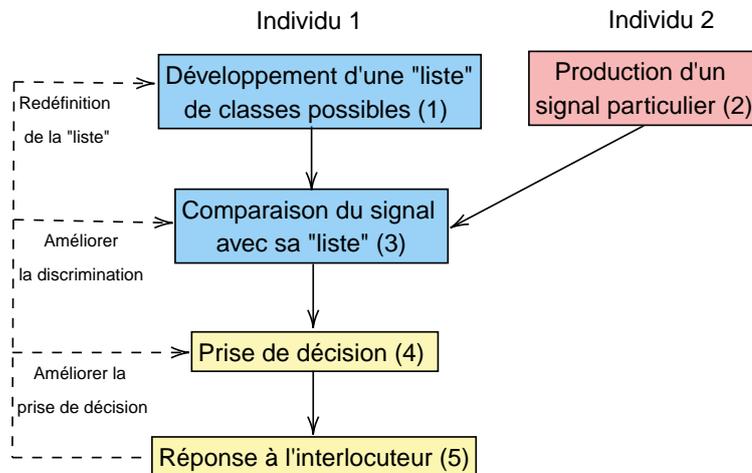


FIG. 1.6 – Représentation des processus indispensables à la reconnaissance d'appartenance à une population (Bradbury et al., 1998).

Ces processus sont donc mis en place quotidiennement pour l'utilisation des dialectes au sein d'une population. Les émissions sonores au sein d'une population peuvent servir à la chasse coopérative, au maintien des groupes ou encore à la reproduction. Ces émissions sont donc de véritables indices bioacoustiques afin d'évaluer la présence et le comportement de certaines populations sur une aire géographique.

Il est important de signaler que beaucoup d'études sont réalisées sur l'identification de patterns acoustiques liés aux populations. Mais cela ne signifie pas que ces patterns sont utilisés par les animaux et possèdent un quelconque rôle biologique. Peu d'études ont réellement montré que les individus étaient capables de reconnaître les différentes populations sur la base de ces invariants (Mortega et al., 2014 ; Lengagne et al., 2020 ; Nicholls, 2008 ; Levréro et al., 2009).

Les causes de ces variations géographiques de ces signaux acoustiques peuvent être nombreuses, comme des barrières géographiques (Perry and Terhune, 1999), des variations environnementales (Nicholls and Goldizen, 2006), différentes pressions de sélections (Slater, 1986) ou des processus aléatoires.

La dérive génétique et culturelle peut également jouer un rôle dans la différenciation des signaux (Irwin et al., 2008).

1.2.4 Détection d'individus

L'échelle d'analyse en bioacoustique peut être encore plus précise que l'étude classique des espèces et des populations, et intégrer une notion de suivi individuel au sein d'un groupe d'animaux.

L'individu produit donc un signal distinctif où la différence interindividuelle est élevée et la différence intra-individuelle est faible, de sorte que chaque individu possède son propre signal. Deux mécanismes peuvent contribuer à la reconnaissance d'un individu au sein d'un groupe

(Laiolo, 2010).

Le premier est la distinction des types d'émission. L'exemple le plus concret est la voix humaine. Nous pouvons reconnaître un membre de notre famille avec son timbre de voix (Sheng et al., 2017). L'émetteur ne produit pas ce signal de manière volontaire, mais il est quand même reconnu. Ce phénomène dépend uniquement des paramètres physiologiques de l'émetteur. Mais parfois ce caractère distinctif n'est pas suffisant, et les espèces sont contraintes d'utiliser un deuxième mécanisme à savoir la production de signatures individuelles émises uniquement dans le but d'une identification (Russo, 2001 ; McCowan and Reiss, 2001).

L'illustration de ce phénomène est l'attribution d'un prénom chez l'humain afin de s'identifier et de s'interpeller.

Dans les deux mécanismes, un phénomène d'apprentissage est nécessaire, afin d'apprendre et de différencier les membres du groupe. Au sein d'un groupe, les individus peuvent émettre certains types de signaux afin de maintenir la communication intra-groupe. Lorsque le groupe est stable et de taille suffisante, différents processus peuvent être mis en place afin que chaque individu possède une signature individuelle. En effet, l'identification individuelle au sein d'un groupe favorise le développement de la stabilité de celui-ci (Barnard and Burk, 1979). Beaucoup de scientifiques ont déjà mis en évidence l'existence de signatures individuelles (Snowdon and Cleveland, 1980 ; Charrier and Harcourt, 2006 ; Jouventin and Aubin, 2002 ; Lengagne et al., 2001).

Prenons l'exemple de colonies d'oiseaux marins (plusieurs milliers) où les parents doivent aller chasser au large, puis reviennent à terre afin de nourrir leur(s) progéniture(s). Les poussins âgés de quelques semaines ne possèdent aucune différenciation au niveau visuel pour permettre aux parents de les reconnaître. Le poussin et ses parents sont donc obligés d'émettre des signaux particuliers afin de se distinguer dans la colonie, c'est le cas du manchot Royal (Lengagne et al., 2001). Chez différentes espèces d'oiseaux, le concept de reconnaissance individuelle a été mis en évidence notamment dans des contextes de reproduction. En effet, certaines espèces sont monogames, il faut donc que les deux oiseaux appariés arrivent à s'identifier. Pour cela ils doivent apprendre certaines caractéristiques individuelles pouvant être acoustiques (Beer, 1971). Sharp and Hatchwell (2005) ont réussi à mettre en évidence la reconnaissance individuelle chez l'Orite à longue queue, en associant un type d'appel à un individu. Cette espèce utiliserait donc ces types d'appels particuliers pour se reconnaître entre eux.

Le contexte de reconnaissance entre deux individus repose donc sur des concepts étudiés au sens large du terme (Reeve, 1989).

Chez les anoues, l'identification individuelle n'a pas encore été clairement démontrée (Bee and Gerhardt, 2002), mais différentes études ont prouvé l'existence de signaux acoustiques pour la reconnaissance de l'identité sexuelle (Vignal and Kelley, 2007). Lors d'une reconnaissance individuelle, le signaleur (l'émetteur) est reconnu par des signaux de reconnaissance acoustique unique, le récepteur apprend ces signaux et les utilise pour identifier le signaleur lors d'interactions futures. Les receveurs développent des capacités de discrimination pour identifier le destinataire cible. Ces capacités reposent sur un équilibre de décisions qui maximise les avantages d'assigner un signal à un individu et également minimise l'erreur qui peut être faite.

On peut définir deux principaux types d'émetteurs : les émetteurs désirables et indésirables pour le récepteur, voir Fig.1.7. Ce graphique représente la dissimilarité entre le phénotype de

l'émetteur (exemple une vocalise signature) et la "liste" possible du récepteur (Johnstone, 1997).

Si l'émetteur arrive à envoyer son signal et se faire reconnaître, il sera classé comme émetteur désirable. Pour cela il faut que la dissimilarité entre le phénotype de l'émetteur et la "liste" du récepteur soit faible.

Si la dissemblance perçue tombe en dessous du seuil de réponse du récepteur (ligne rouge), le signaleur est « accepté » comme souhaitable et suscite une réponse favorable ; sinon, le signaleur est « rejeté », c'est-à-dire qu'il suscite une réponse défavorable.

Le receveur doit ajuster son seuil d'acceptation du signal (ligne rouge).

Il existe un recouvrement entre les deux courbes, donc certains émetteurs indésirables peuvent être acceptés et inversement.

Dans la zone bleue, le récepteur adopte réponse favorable à un signaleur indésirable. Tandis que dans la zone violette il "refuse" un individu qui aurait dû être accepté.

En d'autres termes, il est impossible pour le récepteur de minimiser simultanément les probabilités de fausses alarmes et détection manquée.

Au cours de la vie de l'animal, le recouvrement des courbes se réduit et le taux d'erreur aussi. (Johnstone, 1997).

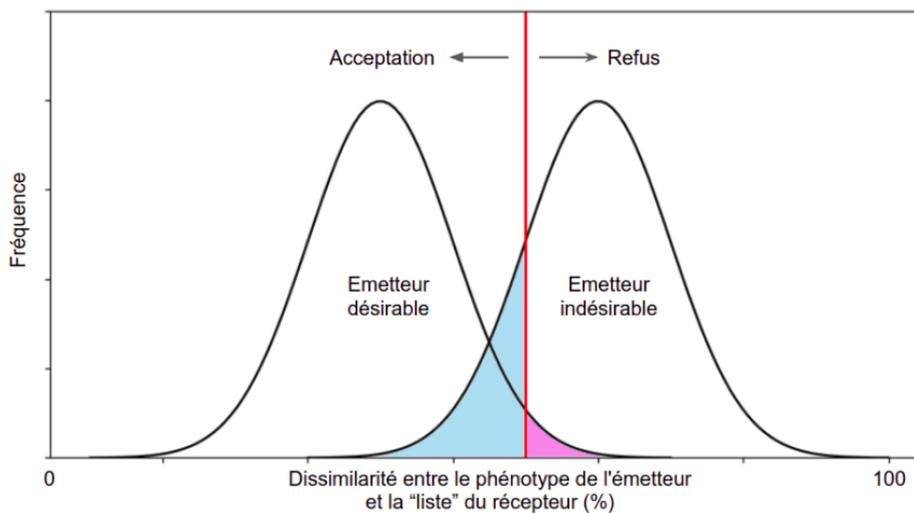


FIG. 1.7 – Schéma tiré de Johnstone (1997) résumant les mécanismes de reconnaissance et la distribution des fréquences des deux types d'individus reconnus ou non.

Chez les cétacés, l'espèce la plus étudiée en captivité ou en liberté est le Grand dauphin (*Tursiops truncatus*). En effet, cette espèce vit en communauté avec des groupes qui peuvent atteindre plus de 100 individus. Pour maintenir la cohésion d'un groupe, ils utilisent des "signaux signatures" qui ont la forme de sifflements afin de déclarer leur présence dans le groupe. Donc au sein du groupe, chaque individu est "étiqueté" (Janik et al., 2006).

Les individus peuvent également imiter l'étiquette (sifflement) d'un autre individu pour s'adresser à lui. Ce signal est développé par les jeunes durant la première année de vie de l'animal. Le petit va créer ce sifflement en écoutant ses congénères et en ajoutant une spécificité particulière (Janik et al., 1994, 2006 ; Sayigh et al., 1990 ; Smolker et al., 1993). La figure 1.8 montre les sifflements signatures appartenant à deux individus différents. La stabilité de ces signaux est

très importante pour favoriser une bonne identification.

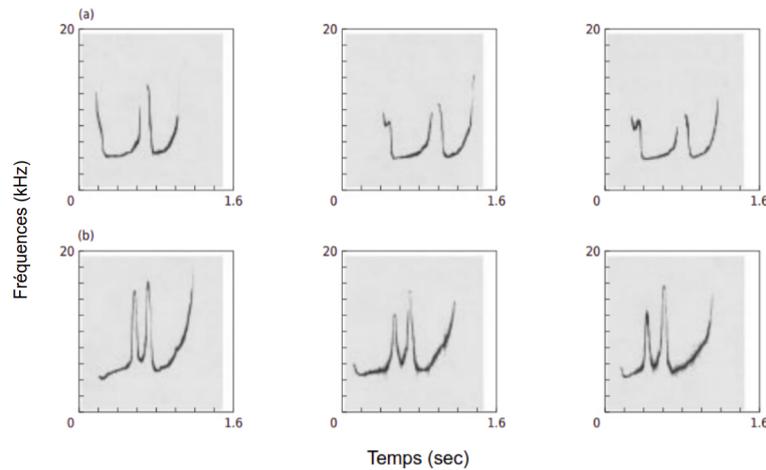


FIG. 1.8 – Représentation de 6 spectrogrammes appartenants à deux individus d'un même groupe (*Turisops truncatus*). a) : un mâle, b) : une femelle (les harmoniques ont été supprimées). Figure tirée de Janik and Slater (1998).

Il n'y a pas que chez le Dauphin commun chez qui les scientifiques ont mis en évidence la présence de signatures individuelles. Shapiro (2006) a travaillé sur deux Narvals de la mer de Baffin, et a démontré qu'il existait potentiellement des différences dans les signaux émis entre deux individus. Cette hypothèse a été également validée chez le Dauphin d'Électre (Kaplan et al., 2014) et le Dauphin commun (Fearey et al., 2019).

Nous pouvons donc utiliser ces types de signaux pour étudier un individu en particulier chez les espèces d'oiseaux et chez les cétacés.

1.2.5 Qualité globale des écosystèmes : Écoacoustique

Une discipline récente a vu le jour ces dernières années : l'écoacoustique. C'est un moyen d'étudier les signaux émis de façon plus globale à l'échelle de la biodiversité. Elle repose sur des principes de base de l'écologie comme les interactions existantes entre différents êtres vivants au sein d'un écosystème. L'échelle d'étude de l'écoacoustique est donc très vaste.

A l'inverse de la bioacoustique, l'écoacoustique peut s'intéresser à plusieurs espèces simultanément et étudier l'écologie des populations, des communautés et des paysages (Sueur and Farina, 2015).

L'écoacoustique utilise des sons biologiques mais également des sons "environnementaux" comme le bruit des vagues ou de la pluie, des sons associés à des paysages (Pijanowski et al., 2011). L'écoacoustique repose principalement sur l'hypothèse de la niche acoustique ("Acoustic Niche Hypothesis" ANH). En écologie, une niche est une position ou un rôle d'une espèce dans le fonctionnement de l'écosystème. On peut donc dire que deux espèces possédant la même niche écologique peuvent être en compétition pour les ressources et devront potentiellement s'affronter pour survivre. L'hypothèse de la niche acoustique suppose donc que si deux espèces produisent des sons en même temps, elles doivent occuper des niches différentes (fréquences) pour optimiser

leur communication (Krause, 1993). Il existerait donc un partitionnement au niveau fréquentiel et temporel de l'expression des signaux afin d'éviter les recouvrements (Sueur, 2002 ; Sinsch et al., 2012 ; Robert et al., 2019).

Ce concept de niche acoustique a été mis en évidence pour des écosystèmes marins, mais également terrestres.

Au et al. (1985) ont analysé les clics d'écholocation des Bélugas, et ont démontré que cette espèce augmente la fréquence et l'amplitude des impulsions des clics lorsque le bruit ambiant (notamment les cliquetis des crevettes) était présent. Cet exemple illustre très bien le concept de la niche acoustique : la plasticité acoustique du Béluga lui permet de changer de niche lorsqu'il est en compétition avec une autre espèce. Les Orques (*Orcinus orca*) et les phoques léopards (*Hydmrga ieptonyx*) sont des prédateurs au sommet de la chaîne alimentaire. Ces deux espèces fréquentent les eaux de l'Antarctique dans la mer de Ross durant la même période (Décembre). Il s'est avéré qu'ils partagent la même niche acoustique. Lorsque les deux espèces partagent le même environnement, les Orques ont développé leur propre niche acoustique au dessus et en dessous des fréquences émises par le Phoque léopard. La modulation en fréquence est donc un moyen d'occuper une niche différente lorsque les phoques sont présents (Mossbridge and Thomas, 1999).

Concernant les espèces aviaires, Stone (2000) a mis en évidence la création de niche acoustique chez certaines espèces d'oiseaux, suite à une anthropisation du milieu. En Nouvelle Zélande, Azar and Bell (2016) ont démontré que les oiseaux introduits possédaient une niche acoustique bien différente que celle des espèces locales.

L'écoacoustique est basée sur l'enregistrement de longues périodes temporelles, mais également sur des échelles géographiques remarquables. Elle doit donc mettre en place des méthodes automatiques adaptées afin d'extraire les informations nécessaires de l'écosystème sonore (Farina, 2013).

L'extraction de certaines informations (présence de certaines espèces) peut être fait automatiquement (Potamitis, 2014), mais chaque écosystème comporte des particularités sonores bien spécifiques qui peuvent entraver l'universalité des tâches de classification (capacité de généraliser) (Sueur and Farina, 2015). Récemment, l'écoacoustique a mis en place différents indices mathématiques pouvant être utilisés dans divers écosystèmes. Ces indices estiment l'amplitude, la régularité, la richesse et l'hétérogénéité d'une communauté acoustique ou d'un paysage sonore (Sueur et al., 2014). Il existe différentes catégories d'indices :

- indice étudiant l'énergie sonore (amplitude)
- indice qui extrait la complexité d'un enregistrement
- indice qui extrait des informations sur la biophonie, géophonie et anthrophonie.

Ces trois catégories recouvrent une vingtaine d'indices, qui peuvent être complémentaires. Dans cette thèse nous n'aborderons pas l'écoacoustique de manière pragmatique pour étudier la communication animale, je vais m'orienter principalement sur les espèces/populations/individus et non sur l'écosystème en général. Je développerai les avantages/inconvénients de l'intégration de ces nouvelles méthodes dans l'étude de la communication animale.

1.3 Problématiques et objectifs de la thèse

La partie précédente a montré que la bioacoustique peut être réalisée sur différentes échelles : l'échelle spécifique, populationnelle et individuelle.

L'objectif de cette thèse est de mettre en place des outils depuis la pose du matériel d'acquisition jusqu'à l'analyse des données pour l'ensemble des échelles présentées. Différents moyens et dispositifs acoustiques seront présentés et associés à différents cas d'études. Les méthodes proposées dans cette thèse se concentrent essentiellement sur la corrélation des signaux, et quelques classifications supervisées ou non supervisées.

Les problématiques sont posées sur l'ensemble d'une séquence acoustique et comportementale qui peut durer de quelques minutes (réunion d'un individu avec son groupe) à plusieurs heures (séquence d'appels de Baleines à bosse pour la reproduction), jusqu'à plusieurs jours (fréquentation d'un site pour la chasse des Cachalots).

Quatre questions seront récurrentes tout le long de cette thèse : "quoi?", "qui?", "où?" et "pourquoi?" :

- " Quoi ? " : quel est le type de signal émis? Plus généralement, cela peut être une vocalise, un sifflement, un clic ou encore un buzz. Ces derniers sont des signaux, car ils sont émis volontairement par l'émetteur. Pour y répondre nous verrons quels matériels et méthodes sont nécessaires.
- " Qui ? " : Cette deuxième question est relative à l'individu qui code (volontairement ou non) des informations dans son émission afin qu'il soit identifiable en tant qu'individu. Cette information peut donc être volontaire comme les signatures individuelles ou involontaire comme le timbre (fréquence fondamentale). Pour répondre à cette question, il faut envisager également du matériel et des méthodes qui seront encore plus poussées.
- " Où ? " : Cette question se décline en plusieurs sous-questions. Prenons le cas où un émetteur envoie un signal à un récepteur. La première utilisation serait l'écholocation. L'émetteur envoie un signal pour localiser son récepteur. La deuxième possibilité serait de répondre à la question "où" pour laquelle l'émetteur envoie un signal afin que ses congénères puisse le localiser. Cette question "où" est donc double (passive ou active). Pour y répondre il faudra en général plusieurs hydrophones et des méthodes de traitement du signal pour localiser la source.
- " Pourquoi ? " : cette dernière question concerne les comportements volontaires des individus (chasser, socialiser) ou des comportements involontaires (facultés génétiques ou phénotypiques). C'est donc l'étude des causes qui ont déclenché un comportement particulier.

Dans ce travail de thèse, nous essayerons de répondre à ces différentes questions qui seront associées à chaque échelle d'étude.

Les cas d'étude menés dans cette thèse, sur différentes espèces et milieux, permettront de mettre en évidence des méthodes génériques propices au développement de la bioacoustique, d'en tirer les grands principes, et de les mettre en perspective notamment dans le cadre de données massives (plusieurs téraoctets).

1.3.1 Organisation de la thèse et objectifs

Dans ce travail de thèse je propose d'analyser différents corpus d'enregistrements afin de valider les méthodes proposées en répondant à différentes problématiques en bioacoustique. Je vais baser mes démonstrations sur différentes échelles d'études décrites dans la partie précédente.

1.3.1.1 Bioacoustique spécifique

La première échelle, qui est la plus large est la détection d'espèces au sein d'un écosystème par bioacoustique. En se basant sur les méthodes existantes, je vais proposer plusieurs méthodes pour la détection d'espèces avifauniques et sous-marines.

Le résultat issu de chaque méthode sera discuté.

Le premier protocole de détection d'espèces sera appliqué à la détection d'espèces d'oiseaux sur le sol français. Pour cela, je présenterai les espèces considérées, le modèle de détection mis au point et enfin les applications et l'utilisation d'un tel modèle pour les citoyens et son utilisation en bioacoustique.

Mes recherches seront également basées sur deux autres protocoles sous-marins. Premièrement une bouée sous-marine (BOMBYX) mise en place au large de Toulon de 2014 à fin 2018, permettra d'étudier les Cachalots avec un détecteur semi-automatique. J'en présenterai les spécificités et les résultats. Puis la détection d'espèces s'articulera également sur la détection d'Orques en Colombie-Britannique (Canada).

Ces trois protocoles, sous-marins ou terrestre, seront comparés et discutés afin d'identifier la méthode de détection d'espèces optimale en fonction de la cible étudiée.

1.3.1.2 Bioacoustique populationnelle

La deuxième échelle d'étude est l'échelle au niveau de la population, donc sur une recherche de dialecte acoustique.

Pour cela, je me baserai sur l'exemple des Orques de Colombie-Britannique regroupées en plusieurs populations bien définies. L'objectif de ce protocole sera d'identifier automatiquement les différents pods qui fréquentent ces eaux. Je discuterai ces résultats pour décrire les modes de communication au sein d'une population.

Je comparerai mes résultats avec des travaux sur l'étude des dialectes sous-marins mais également terrestres. J'essaierai de répondre sur la potentialité de l'apport des études des dialectes dans le but d'améliorer les connaissances en bioacoustique.

1.3.1.3 Bioacoustique individuelle

L'échelle d'étude s'affinera dans la troisième partie de la thèse qui portera sur l'utilisation de la bioacoustique pour décrire la communication interindividuelle. Mes propos seront fondés sur deux protocoles sous-marins et un protocole terrestre. Le premier protocole sous-marin portera sur l'identification individuelle d'Orques au Canada, dans le but d'étudier la communication intra-groupe de ces animaux. Je présenterai le protocole mis en place, les résultats de l'identification individuelle et également les perspectives de cette étude.

Le deuxième protocole mis en place pour l'identification individuelle sous-marine se déroule en Méditerranée. Il porte sur le suivi individuel de Cachalots grâce à leur clic d'écholocation. Ce

protocole a été mis en place durant l'été 2018 au large de Toulon.

Puis l'identification individuelle terrestre portera sur la localisation d'oiseaux via une antenne multi-microphones sur le campus de l'Université de Toulon.

Ces trois protocoles pour le suivi individuel seront mis en comparaison, ainsi que les espèces étudiées. Je présenterai également les intérêts de ces identifications individuelles pour de futurs études en écologie comportementale.

1.3.1.4 Bioacoustique comportementale : première approche

La dernière échelle d'étude travaillée est la bioacoustique "comportementale" (sans détection acoustique d'espèce, ni de population, ni d'individus).

Pour cela, les espèces, populations et individus ne seront pas détectés acoustiquement, mais un détecteur de suivi de signaux stationnaires sera mis en place. L'objectif de cette échelle est de mettre en corrélation des comportements avec des émissions acoustiques.

J'ai choisi d'appliquer une seule méthode sur deux cas d'étude : les Dauphins tachetés pantropicaux et la Baleine à bosse.

Ces deux espèces sont associées à une problématique en écologie comportementale que je développerai dans le chapitre Matériel.

Cette échelle d'étude est transverse, car elle peut être appliquée au niveau spécifique, populationnel et individuel.

J'analyserai les avantages et les inconvénients des méthodes utilisées et surtout ces nouveaux champs d'applications pouvant être appliqués au niveau spécifique, populationnel, individuel et comportemental.

Ce travail reposera donc sur 4 principales échelles d'études, associées à des protocoles précis. Les échelles ainsi que les cas d'étude sont présentés Fig.1.9.

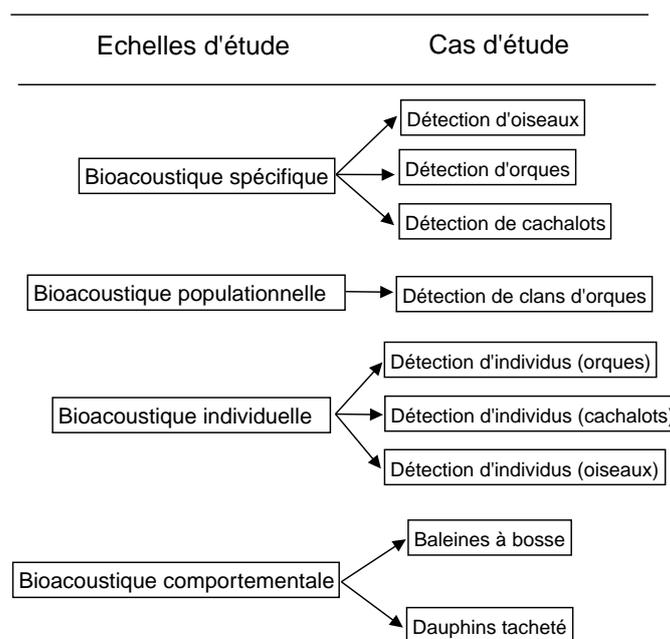


FIG. 1.9 – Schéma représentant les différentes échelles d'analyses ainsi que leurs cas d'étude associés.

En résumé cette thèse tentera de répondre à cette question :

Comment la bioacoustique peut étudier la faune sur différentes échelles d'études : l'échelle spécifique, l'échelle populationnelle, l'échelle individuelle et l'échelle comportementale ?

Ce travail fera l'objet d'une première partie qui présentera le matériel et les méthodes utilisés pour ces 4 échelles d'études. Je décrirai donc les protocoles acoustiques mis en place suivi des méthodes de traitement des signaux.

Puis la troisième partie exposera les premiers résultats, en terme de détection d'espèces, de suivi de population, de suivi individuel et enfin en terme de bioacoustique comportementale.

Chacun de ces protocoles a produit des premiers résultats intéressants concernant les espèces, populations ou individus mais également intéressants en terme d'amélioration de nos techniques actuelles pour le développement de la bioacoustique.

Les résultats nous permettront premièrement de **valider nos protocoles et nos méthodes d'analyses**, et de poser de nouvelles perspectives en bioacoustique.

Ils seront donc présentés au niveau acoustique et en écologie comportementale, mais également discutés en fonction des méthodes utilisées. Les méthodes employées au cours de cette thèse seront critiquées afin qu'elles puissent être améliorées dans de travaux futurs. Je présenterai donc les perspectives écologiques et méthodologiques de ces études.

L'ensemble de ces différentes étapes est décrit dans le schéma 1.10. La première étape est de fournir des outils et des dispositifs acoustiques novateurs pour la bioacoustique. Ces différentes ressources ont été sélectionnées pour répondre à des problématiques en bioacoustique. Puis, une fois le matériel et les méthodes sélectionnés, ils seront appliqués sur des protocoles opportunistes associés à différents cas d'étude. C'est à dire que les concepts développés dans les parties précédentes concernant les méthodes et le matériel seront validés dans le chapitre Résultat. Ce dernier aura deux principaux objectifs : le premier sera de valider les protocoles d'acquisitions proposés, et le deuxième sera d'analyser les tendances bioacoustiques dégagées par ces protocoles.

Puis la dernière étape de mon raisonnement sera basée sur les perspectives naissantes des différents cas d'étude. En fonction des tendances dégagées, de futures études seront envisagées pour faire de ses tendances, des résultats bioacoustiques robustes. La deuxième partie de mes perspectives portera sur les différentes avancées méthodologiques nécessaires pour l'amélioration des différents systèmes proposés.

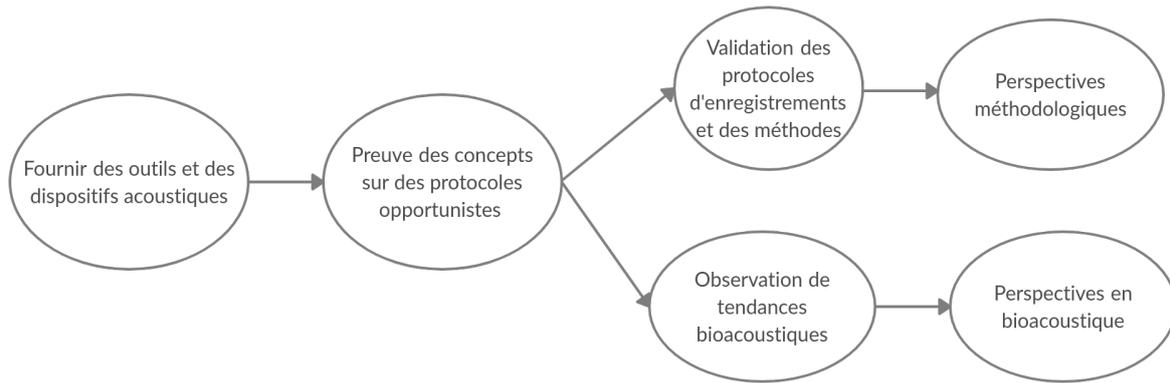


FIG. 1.10 – Schéma représentant les différentes étapes réalisées dans cette thèse.

1.3.2 Lecture du manuscrit

Ce manuscrit est écrit de la manière classique : Introduction, Matériel, Méthodes, Résultats et Discussion.

Dans chacun de ces chapitres, chacune des échelles d'analyse est présentée : spécifique, populationnelle, individuelle et comportementale. Et, à l'intérieur de chacune des échelles des cas d'étude sont développés (Fig.1.9). La lecture de ce manuscrit peut donc se faire en suivant les chapitres principaux.

Mais elle peut se faire également en suivant les cas d'étude analysés au cours de la thèse. La figure 1.11 représente les chapitres matériels, méthodes et résultats avec leurs échelles d'analyses et leurs cas d'étude associés. Il est donc possible de lire le manuscrit cas d'étude par cas d'étude. Cette méthode de lecture est indiquée sous forme de carrés de même couleurs dans le schéma 1.11.

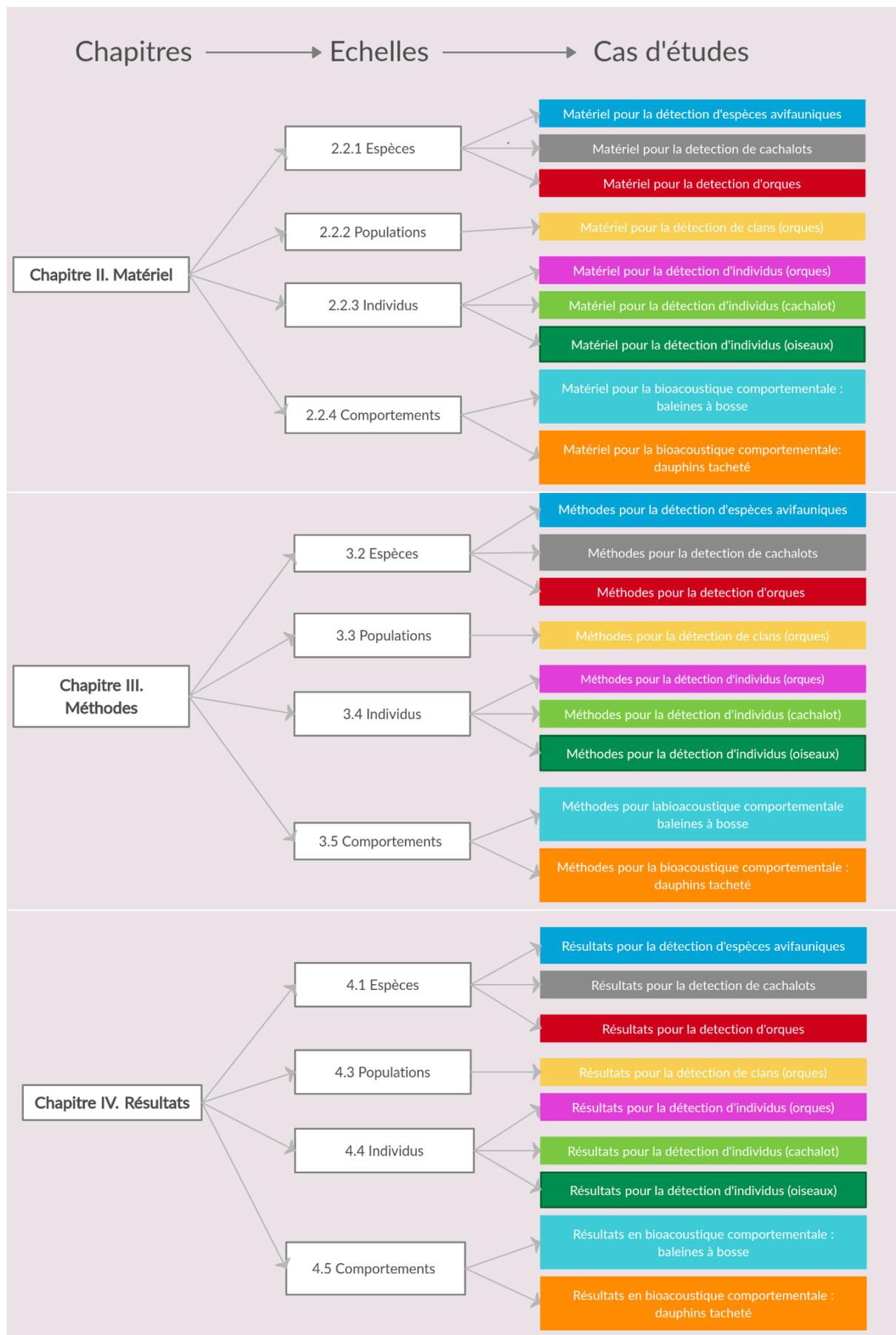


FIG. 1.11 – Schéma représentant les chapitres Matériels, Méthodes et Résultats ainsi que les différentes options de lecture du manuscrit (couleurs).

Chapitre 2

Protocoles d'expérimentations

Ce chapitre va décrire les différents modèles d'étude (espèces) utilisés. Puis nous présenterons le matériel nécessaire pour la bioacoustique spécifique, populationnelle, individuelle et comportementale.

Pour chaque partie (espèce, population, individu, comportement) un état de l'art de la bibliographie sur le matériel utilisé est présenté.

2.1 Présentation des modèles d'étude

2.1.1 *Les espèces avifauniques*

Le premier protocole au niveau spécifique repose sur la classification de plusieurs espèces d'oiseaux à partir de leur signaux acoustiques, et le troisième protocole pour la détection d'individus repose sur la localisation d'oiseaux. Comme décrit dans l'introduction, la production acoustique chez les oiseaux peut se caractériser par deux principaux signaux : les cris et les chants (Catchpole and Slater, 2003). Les cris sont très courts, simples et peuvent être produits toute l'année par le mâle et la femelle, tandis que les chants sont produits généralement par les mâles en période de reproduction et sont plus longs et plus complexes. Dans certaines zones tropicales, les femelles peuvent également produire des chants pour la reproduction (Slater and Mann, 2004). Ces deux types de signaux permettent la détection de l'espèce.

L'Europe, et plus particulièrement la France, est dotée d'un habitat favorable pour plus de 950 espèces d'oiseaux (Mullarney et al., 2009), certaines sont très courantes tandis que d'autres sont extrêmement rares. [BirdLife International](#) a constitué une liste d'espèces (530) fréquentant la zone européenne et a évalué l'état de conservation de ces espèces. Plus de 13% de cette liste est menacé d'extinction, et 10% sont en danger critique.

C'est pour cette raison qu'il est fondamental de mettre en place des surveillances acoustiques et de développer l'étude de cette biodiversité vulnérable. Les populations sont globalement en danger à cause de différentes menaces : le réchauffement climatique (Brown et al., 1999), les pratiques agricoles (Benton et al., 2002), la perte d'habitat (Owens and Bennett, 2000), ou encore le braconnage (Barca et al., 2016).

Nous avons convenu, avec Thierry Lengagne (ornithologue), et nos collaborateurs (Allain Bougrain-dubourg et Anne-Laure Bedu), de sélectionner une cinquantaine d'espèces d'oiseaux les plus fréquentes sur le territoire français afin de classifier automatiquement ces espèces. La liste des espèces est présentée en Annexe A. Cette liste d'espèces, ainsi que leur fiche d'identification est également disponible sur le site [BIRDIFY](#). La bioacoustique spécifique reposera sur la classification automatique de cette cinquantaine d'espèces d'oiseaux. Le protocole d'identification

individuelle en milieu terrestre repose également sur les oiseaux. Les espèces que nous avons essayé de localiser faisaient également partie de la liste présentée en Annexe A.

2.1.2 Le Cachalot (*Physeter macrocephalus*)

Le premier modèle d'étude sous-marin est le Cachalot dans le sanctuaire Pelagos. Ce sanctuaire est un espace maritime au coeur de la Méditerranée pour la protection des mammifères marins. La mégafaune au sein du sanctuaire est représentée par deux espèces : le Rorqual commun (*Balaenoptera physalus*) et le Cachalot (*Physeter macrocephalus*). Ce dernier est le plus grand des odontocètes (jusqu'à 18m) et est cosmopolite (Whitehead, 2003) (Fig.2.1).

Il est classé comme vulnérable par l'UICN à cause de la réduction de la population mondiale due à la chasse commerciale. Mais dernièrement, d'autres menaces pèsent sur les Cachalots comme la pollution chimique, le bruit anthropique ou encore les collisions avec les navires. En effet, ces dernières ont été amplifiées à cause du nombre de bateaux de plus en plus important (Laist et al., 2001).

Le Cachalot est un grand plongeur et peut évoluer autour de 1000 m de profondeur (Watwood et al., 2006 ; Wahlberg, 2002 ; Poupard et al., 2019d) avec des maximums à -3000 m (Jaquet et al., 2000), pour chasser principalement des céphalopodes. Il est donc très difficile de les étudier car ils passent plus de 90% de leur temps dans les abysses. La bioacoustique est donc une méthode très bien adaptée pour étudier cette espèce. Afin de capturer ses proies dans un milieu totalement obscur, le Cachalot utilise l'écholocation. Il localise ses proies en émettant un son très directif (clics) et écoute son écho. L'intervalle entre chaque clic est usuellement entre 0.5 et 1 sec (Jaquet et al., 2001). Différents clics peuvent être produits : les clics d'écholocation (chasse) et les clics de socialisation (codas) pour la communication.

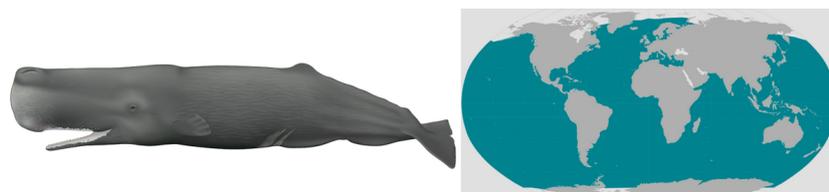


FIG. 2.1 – Représentation du Cachalot (*Physeter macrocephalus*) associée à sa répartition géographique. Source : NOAA Fisheries (2011a).

2.1.3 L'Orque (*Orcinus orca*)

La deuxième espèce marine étudiée est l'Orque (*Orcinus orca*). L'Orque est un mammifère marin qui vit dans les océans du monde entier (Fig.2.2). Chez ces animaux, la communication est vitale afin de pouvoir échanger, chasser et se reproduire (Council et al., 2003 ; Burham et al., 2016).



FIG. 2.2 – Représentation de l'Orque (*Orcinus orca*) associée à sa répartition géographique.
Source : NOAA Fisheries (2011b).

Elles peuvent produire une grande variété de sons, y compris des clics, des sifflements et des vocalises (Ford, 1991). Les appels ou vocalisations sont primordiaux pour la communication, et un pod (groupe d'animaux) peut partager entre 7 et 17 appels différents. L'activité vocale dépend également de l'activité du groupe (Filatova et al., 2013). De nombreuses études acoustiques ont déjà prouvé que les vocalisations jouent un rôle important dans la vie sociale du groupe (Bain, 1986 ; Filatova et al., 2013, 2010). Nos études se concentrent sur les Épaulards Résidents du Nord (NRKW), communauté composée de différents pods et matriline (Bigg et al., 1990).

Trois écotypes distincts d'Orques vivent dans les eaux de la Colombie-Britannique : résidents, nomades (transients/Bigg's) et les hauturiers (offshores) (Hoelzel and Dover, 1991 ; Dalheim et al., 2008). Les Orques nomades sont des mangeurs de mammifères marins tandis que les résidents chassent principalement le saumon (Ford et al., 1998). La population des hauturiers est la moins étudiée car ces animaux sont beaucoup plus mobiles et leur aire de répartition est beaucoup plus grande.

Le niveau le plus élevé d'organisation sociale des épaulards résidents est appelée la Communauté (Fig.2.3). Il existe deux communautés appelées communauté des résidents du sud et communauté des résidents du Nord.

La communauté nordique est composée de 15 pods différents et fréquente les eaux entre l'île de Vancouver et le sud-est de l'Alaska. Les pods sont des groupes d'Orques apparentés qui vivent et voyagent fréquemment ensemble. Ils partagent également le même dialecte acoustique. La plus petite unité de socialisation au sein des pods est la matriliné, qui relie les individus par la descendance maternelle (Ford, 2009). La communication au sein de ces pods est très développée et un dialecte très spécifique est utilisé par l'ensemble des animaux du groupe. Certains types d'appels peuvent être partagés par plusieurs pods, tandis que certains sont spécifiques à un seul pod.

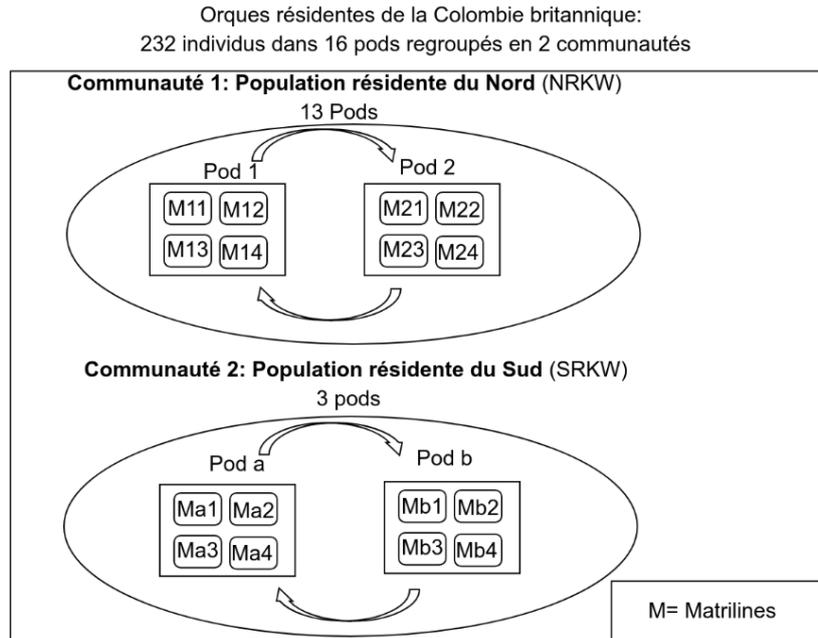


FIG. 2.3 – Organisation des Orques résidentes de la Colombie-Britannique. Les résidents sont regroupés en 2 communautés (Sud et Nord). Ces communautés sont composées de différents pods. Et chacun des pods est constitué de matriline qui sont composées d'individus liés de manière générationnelle (Ford et al., 1987).

Les dialectes au sein du pods sont donc probablement essentiels pour le maintien de la cohésion sociale au sein du groupe (Ford, 1991).

2.1.4 Le Dauphin tacheté pantropical (*Stenella attenuata*)

La première étude en bioacoustique comportementale de ce travail permet la compréhension des modalités de communication d'une espèce de dauphin localisée en Martinique en présence de whale watching.

Cette espèce, la plus fréquemment observée est le Dauphin tacheté pantropical : *Stenella attenuata*, qui est un dauphin cosmopolite dans les eaux tropicales et chaudes, de la famille des Delphinidae. Cette espèce est facilement identifiable, avec ses tâches plus claires sur le dos, son bec et son corps très fin (Fig.2.4).

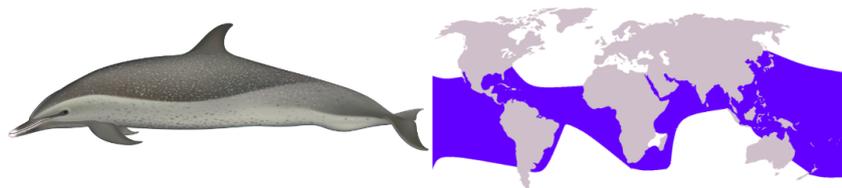


FIG. 2.4 – Représentation du Dauphin tacheté pantropical (*Stenella attenuata*) associée à sa répartition géographique. Source : American Cetacean Society (2017).

Très sociaux, ces dauphins vivent en groupes, qui peuvent atteindre plusieurs milliers d'indi-

vidus (Perrin, 2009). Comme beaucoup d'autres espèces de dauphins, cette espèce peut produire des sifflements, des clics et des bursts (Lammers et al., 2003). Les sifflements produits ont une fréquence minimum de 8.2 kHz et une fréquence maximum de 19.7 kHz (Oswald et al., 2003). Mais ces études ont été réalisées sur des animaux du Pacifique, il est donc possible que la population en Martinique ne possède pas ces mêmes caractéristiques. Un exemple de quelques secondes est en Fig.2.5.

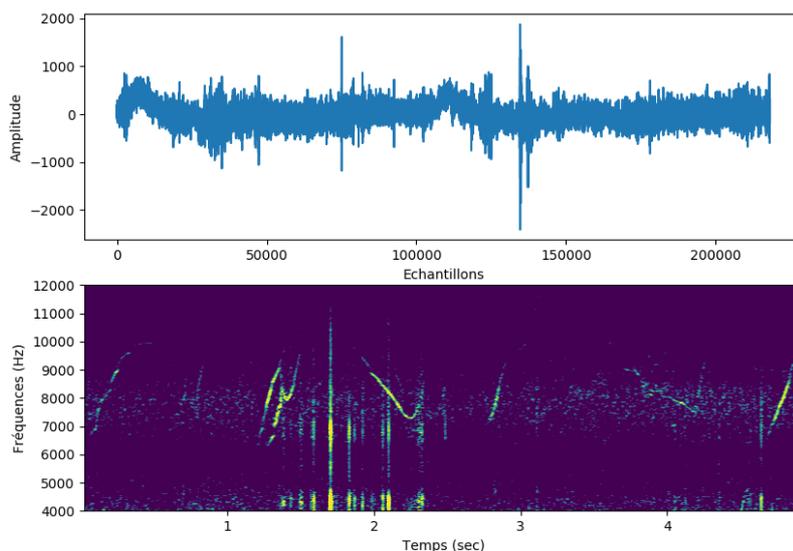


FIG. 2.5 – Forme d'onde (haut) et spectrogramme (bas) de sifflements de Dauphins tachetés pantropicaux (*Stenella attenuata*) en Martinique. Spectrogramme : fenêtre de Hanning de taille 256 échantillons.

Dans les eaux des Caraïbes, cette espèce est communément trouvée dans une zone avec une bathymétrie entre -600 et -2500 m, et est très accessible par les scientifiques et les bateaux de whale watching. Identifiée entre 3 et 72 km des côtes, sa niche écologique est superposée avec plusieurs espèces de cétacés notamment le Dauphin à long bec (*S. longirostris*) et le Dauphin tacheté de l'Atlantique (*S. frontalis*) (Barragán-Barrera et al., 2019).

2.1.5 La Baleine à bosse (*Megaptera novaeangliae*)

Le dernier modèle d'étude sera utilisé en bioacoustique comportementale, il concerne la Baleine à bosse *Megaptera novaeangliae*. Ce mysticète cosmopolite (Fig.2.6) peut se nourrir de plancton (krill) et également de petits poissons pélagiques.

Les Baleines à bosse sont des baleines migratrices. En hiver, ces animaux se trouvent dans des zones de reproduction à basse latitude et durant l'été dans des zones d'alimentation à haute latitude (Cerchio and Dahlheim, 2001). Elles peuvent donc emprunter des routes de migration de plus de 8000 km (Horton et al., 2011). Actuellement, on considère qu'il reste environ 35000 Baleines à bosse. Cette espèce, très emblématique peut vivre plus de 50 ans, et atteindre 25

tonnes pour les mâles.

Les Baleines à bosse produisent des sons différents et complexes (Fig.2.7), en particulier des sons sociaux. Cette communication acoustique comprend ceux générés par des comportements de surface particuliers (sauts) et des vocalises.

Ces types de signaux ont des fonctions sociales et comportementales spécifiques selon la composition du groupe et la médiation des interactions sociales (Dunlop et al., 2008). Ces sons généralement produits dans les zones de reproduction peuvent également être présents dans les aires de nutrition (Thompson et al., 1977).

Les fréquences des vocalisations émises par les Baleines à bosse sont comprises entre 30 Hz et 20 000 Hz (Payne and McVay, 1971), et peuvent se succéder pendant plusieurs heures (chants, thèmes...). Ces successions de chants peuvent varier dans le temps et l'espace (Payne and Payne, 1985).

Il est particulièrement difficile d'identifier un répertoire de vocalisations de ce cétacé, car il dépend du contexte, de la variation géographique et de la transmission culturelle. Ainsi, les chants des Baleines à bosse sont très variables et acoustiquement complexes (Cholewiak et al., 2013). Beaucoup de scientifiques travaillent sur cette espèce impliquant une forte multidisciplinarité. La bioacoustique (Sousa-Lima et al., 2018 ; Adam et al., 2013), le traitement du signal, la génétique (Engel et al., 2008), mais également les observations visuelles (Frankel et al., 1995) sont au coeur de la recherche pour cette espèce.

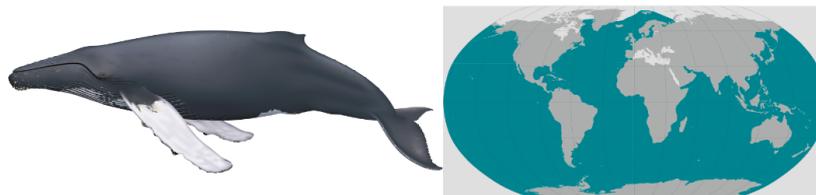


FIG. 2.6 – Représentation de la Baleine à bosse (*Megaptera novaeangliae*) associée à sa répartition géographique (Source : [American Cetacean Society](#), [National Oceanic and atmospheric administration](#))

Parmi toutes les espèces de cétacés, les Baleines à bosse sont les plus communes à Madagascar. A partir du mois de Mai, elles arrivent des eaux de l'Antarctique pour se reproduire à proximité des côtes malgaches (Best et al., 1998). Durant cette période, elles ne sont donc pas supposées s'alimenter et les interactions entre les individus sont nombreuses. Il n'est pas rare d'observer des groupes de plusieurs individus composés en général d'au moins une femelle et son petit (Mobley Jr and Herman, 1985). Les zones côtières de Madagascar sont un habitat idéal pour le développement des nouveaux nés et pour la reproduction (Trudelle et al., 2016).

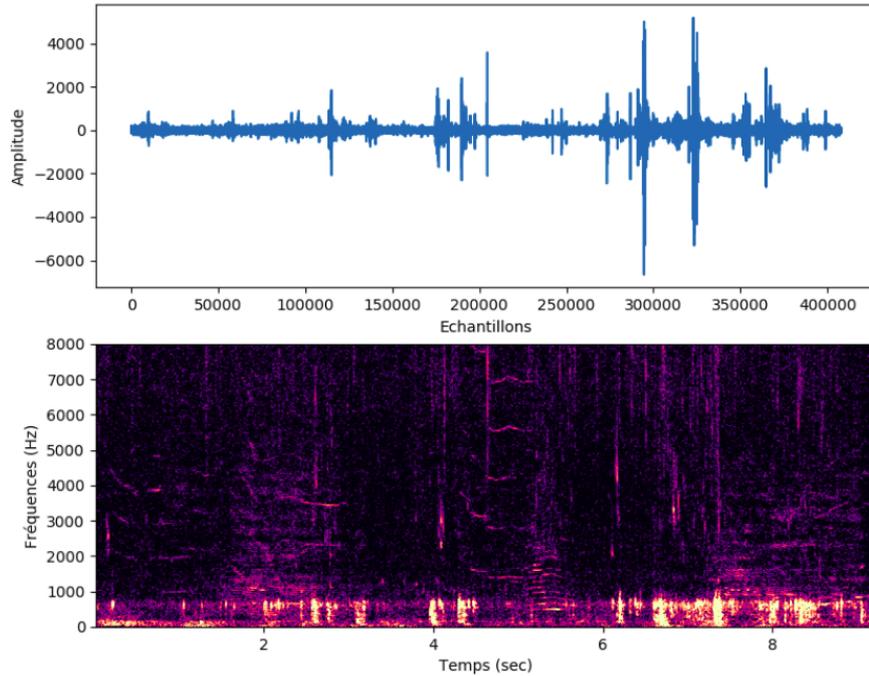


FIG. 2.7 – Spectrogramme et forme d’onde de signaux de Baleine à bosse (*Megaptera novaeangliae*).

La commission baleinière a séparé l’espèce en différentes populations, représentant des stocks d’individus différents. La population de Madagascar représente le stock C3 et est en 2015 représentée par 8854 individus (Jackson et al., 2015 ; Rosenbaum et al., 2009).

Cette première partie du Chapitre Matériel a décrit les différents modèles d’études utilisés dans cette thèse, ainsi que leurs émissions acoustiques.

La partie suivante va décrire comment ces espèces ont été enregistrées et dans quels objectifs.

2.2 Stratégies et techniques d’enregistrements

Avant de détailler plus précisément les protocoles proposés dans cette thèse, il est indispensable de définir le matériel de base en bioacoustique. Ces études reposent sur le PAM (“Passive Acoustic Monitoring”).

Celles-ci développent donc la surveillance acoustique des écosystèmes et ces études ont un point commun : l’utilisation d’un microphone ou d’un hydrophone.

Le principe général d’un microphone ou hydrophone est la transformation des ondes sonores en signaux électriques. Cet appareil va donc convertir les variations de pressions en variations de tensions électriques (transduction).

Cette conversion se fait grâce aux propriétés piézoélectriques de certains matériaux qui vont, sous l’action d’une contrainte mécanique (onde sonore), se polariser électriquement. Cette variation de tension électrique est amplifiée grâce à un amplificateur électronique.

Chaque microphone ou hydrophone dispose d’une réponse en fréquence particulière, définie par l’intervalle de fréquence qu’il est capable de capter. Cette caractéristique doit être prise

en compte pour l'enregistrement de différentes espèces produisant des signaux à différentes fréquences. L'ensemble des réponses en fréquences pour les hydrophones/microphones utilisés dans ce travail sont présentées en annexe B.

Pendant l'enregistrement, ce signal électrique doit être converti grâce à un Convertisseur Analogique Numérique (CAN) à l'intérieur d'un enregistreur. Le CAN va donc convertir la tension électrique en une valeur numérique. Pendant cet enregistrement numérique, l'amplitude du signal électrique est échantillonnée à une certaine fréquence (nombre d'échantillons par seconde), nommée la Fréquence d'Échantillonnage (Fe).

L'enregistrement dépend également du nombre de bits qui encodent les niveaux d'amplitudes possibles pouvant être mesurés. En effet, le signal dit analogique, porte des valeurs continues, tandis que le signal numérique est un signal discontinu quantifié numériquement. L'enregistrement dépend de deux paramètres essentiels : la fréquence d'échantillonnage et le nombre de bits. Il est essentiel de bien choisir la fréquence d'échantillonnage en fonction des espèces visées grâce au théorème de Nyquist-Shannon : la fréquence d'échantillonnage doit être supérieure au double de la fréquence maximale émise par la cible (Shannon, 1949). Par exemple, pour enregistrer des sifflements de dauphins autour de 16 kHz, il faut enregistrer à une fréquence d'échantillonnage minimum de 32 kHz mais en pratique de 44 kHz en tenant compte des repliements spectraux au bord de la fréquence de Nyquist.

La capacité de stockage des enregistrements est donnée par carte (μ)SD, ou un disque dur. Nous avons travaillé avec des volumes de 500Go ou de 1 To.

Les données sont ensuite archivées sur des périphériques de stockage NAS (Network Attached Storage) de plus d'une centaine de To. Souvent la première mission du bioacousticien est de visualiser le son à l'aide du logiciel de visualisation (exemple : Audacity) afin de valider sa qualité. L'ensemble de ces étapes est résumé Fig.2.8.

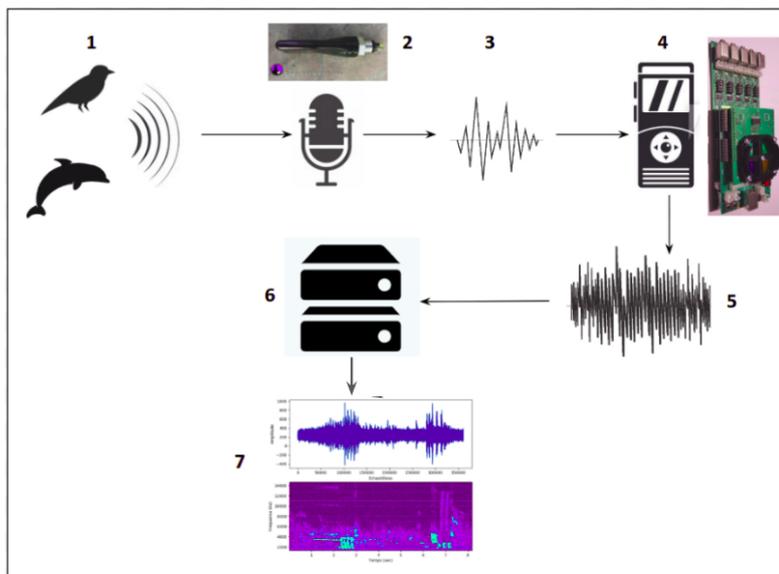


FIG. 2.8 – Schéma représentant une acquisition globale en bioacoustique. 1 : Émission acoustique de la faune, provoquant une onde sonore. 2 : Cette onde sonore est captée par un microphone/hydrophone. 3 : Transformation de l'onde sonore en un signal électrique transmis à un enregistreur (ou carte son) (4). 5 : Conversion de la tension électrique en valeurs numériques et enregistrement. 6 : Archivage sur des disques durs (ex : NAS). La dernière étape est la validation de la qualité des enregistrements.

La partie suivante développe différents types de protocoles mis en place en fonction des échelles d'observations : l'échelle spécifique, populationnelle, individuelle, et l'échelle comportementale. Je vais donc décrire les protocoles utilisés en état de l'art mais également les protocoles qui ont été mis en place au cours de mes recherches.

Chaque protocole est étudié et déterminé pour répondre à une problématique associée à une échelle d'étude.

2.2.1 Matériel utilisé pour la détection d'espèces

La détection d'espèces au sein d'un écosystème est une tâche très courante en bioacoustique. Les stratégies de déploiement du matériel dépendent donc principalement de la cible étudiée. Les scientifiques doivent dans un premier temps, déterminer les espèces à détecter pour ensuite mettre en place le système d'écoute approprié.

Il existe ainsi un équilibre à respecter entre **maximiser l'efficacité des enregistrements** pour détecter tous les signaux appartenant à une espèce, et **optimiser le temps d'enregistrement**, c'est donc un compromis entre la qualité et la quantité des données.

Le premier compromis est la durée de pose, c'est à dire combien de temps le capteur va enregistrer au sein de l'écosystème. Cette période peut durer quelques minutes, jusqu'à plusieurs années. Puis le deuxième paramètre fondamental est la fréquence d'échantillonnage (F_e). Plus la F_e sera importante plus les fréquences détectées seront élevées, et ainsi la qualité des

enregistrements sera optimisée. En revanche, une grande Fe implique une grande consommation d'énergie et surtout une capacité de stockage importante. Il est également possible de réaliser des enregistrements avec des temps de pause. L'intervalle entre ces temps de pause et d'enregistrement est également un paramètre à considérer.

En effet, si les temps de pauses sont importants, l'acquisition pourra durer beaucoup plus de temps (économie de batterie et stockage). Enfin, le dernier paramètre à considérer est la capacité de stockage des enregistrements. Plus cette capacité sera grande plus l'acquisition pourra durer dans le temps, mais les gros supports de stockage (disques durs) consomment d'avantage d'énergie.

Évidemment, le choix de ces paramètres diffère en fonction des objectifs et de l'espèce étudiée mais peut s'appliquer pour des écosystèmes terrestres et marins.

Différents protocoles ont été mis en place pour la détection d'espèces en bioacoustique. Le premier porte sur la détection d'espèces avifauniques depuis des téléphones mobiles.

Puis la détection d'espèces s'appliquera sur deux protocoles sous-marins. Le premier est la détection de Cachalots depuis la bouée BOMBYX en Méditerranée.

Le deuxième cas sous-marin concerne la détection d'Orques en Colombie-Britannique. Pour ce cas d'étude, le matériel était identique à la détection de population (clan acoustique d'Orques). C'est donc dans la partie suivante qu'il sera présenté ("Matériel utilisé pour la détection de population").

2.2.1.1 Stratégie d'écoute pour la détection d'espèces avifauniques

Les méthodes classiques pour détecter acoustiquement les oiseaux, depuis plusieurs années sont les "points d'écoutes" où un observateur humain doit rester sur le terrain et inventorier les espèces qu'il voit et qu'il entend (Rosenstock et al., 2002). Cette méthode présente beaucoup d'avantages, le premier est qu'il n'est pas nécessaire d'utiliser un matériel particulier mais l'observateur, lui doit être très expérimenté afin de minimiser les erreurs d'observations.

Le principal inconvénient de ce type de protocole est sa répétabilité. En effet, il paraît difficile d'utiliser toujours le même observateur afin de limiter les biais de l'échantillonnage. Afin de limiter ces erreurs, les relevés bioacoustiques autonomes sont donc de plus en plus utilisés pour la détection d'espèces aviaires (Blumstein et al., 2011), particulièrement les unités d'enregistrements autonomes (ARU, Autonomous Recording Units) (Dawson and Efford, 2009 ; Marques et al., 2013).

Shonfield and Bayne (2017) ont comparé les méthodes de détections habituelles (écoute de la faune par des scientifiques) avec des méthodes acoustiques autonomes (ARU). Sur 21 études, 11 ont obtenu des performances égales entre les écoutes et les ARU, 6 ont eu des performances plus élevées avec un protocole d'écoute humaine, et 4 ont eu des performances supérieures avec un protocole acoustique autonome.

Donc, actuellement, nous pouvons confirmer que les performances peuvent être similaires entre des enregistrements et des points d'écoutes. Beaucoup d'études ont également montré que ces deux méthodes pourraient être complémentaires (Borker et al., 2015 ; Klingbeil and Willig, 2015 ; Van Wilgenburg et al., 2017). Mais globalement les chercheurs ont admis l'efficacité des enregistreurs autonomes pour la détection d'espèces, particulièrement lorsque la richesse spécifique

est élevée (Hobson et al., 2002).

Puis, plusieurs enregistrements simultanés devrait permettre d'étudier une répartition spatiale tandis que plusieurs enregistrements dans le temps donneront accès à une dynamique temporelle de l'espèce. En terme de matériel, il existe une diversité d'enregistreurs et de microphones efficaces pour enregistrer les espèces et ce, à des fins scientifiques ou destinées au grand public.

En effet, la communauté ornithologique Française est de plus en plus importante. Des sites collaboratifs ont été mis en place afin d'aider le grand public à identifier visuellement et acoustiquement les oiseaux fréquentant leur jardin, ou les lieux publics.

Des associations comme la LPO (Ligue de Protection des Oiseaux) incite la population à réaliser des enregistrements de la faune. La LPO met en place des suivis collaboratifs qui ont pour objectifs de suivre l'état des populations d'oiseaux, de mettre en place des cartes de répartitions et de conserver ces espèces (LPO, 1912).

Les suivis acoustiques des espèces aviaires font partie des objectifs de la LPO. Certains projets collaboratifs sont de véritables banques d'enregistrements cosmopolites, comme [XenoCanto](#). Cette plateforme regroupe les expertises des ornithologues du monde entier, et accumule des milliers d'enregistrements annotés (Vellinga and Planqué, 2015).

Dans ce contexte collaboratif, un des objectifs de cette thèse est de mettre en place une application téléphonique pour la reconnaissance d'espèces d'oiseaux. L'objectif sera de créer une application mobile pour la reconnaissance de chants d'oiseaux permettant l'accès aux informations encyclopédiques de l'espèce d'oiseau identifiée. Pour cela, la première étape est de sélectionner un type de microphone qui pourrait être utilisé pour les citoyens. Nous avons deux choix, le premier était d'utiliser les microphones des téléphones portables, le deuxième aurait été de fournir un microphone aux utilisateurs.

1 Test du matériel pour la détection d'espèces avifauniques

Nous avons donc rassemblé plusieurs téléphones portables et un microphone de la plateforme technologique SMIoT (plateforme technologique élaborée par l'IMN2P et le LIS à l'UTLN) afin de comparer leur performance.

La stratégie était d'émettre par 2 enceintes les sons d'oiseaux et d'enregistrer avec des téléphones portables (5 exemplaires différents) et avec un enregistreur relié à un micro SMIoT pour comparer la qualité de réception sur les différents micros et cartes.

Différents effets ont donc été analysés :

- L'effet « **téléphone** » puisque la bande son diffusée par haut-parleur a été enregistrée avec 5 téléphones différents (+ 1 micro SMIoT comme référence). Le fait d'enregistrer avec un téléphone d'une marque X ou Y affecte-il la qualité des sons et la fiabilité de la classification ?
- L'effet **distance** : à quelle distance le son est-il perçu par un téléphone ? Nous avons enregistré une séquence à 12.5 m, 25 m, 50 m et 100 m en milieu ouvert.
- L'effet **angulaire** : quelles sont les conséquences d'un non alignement de l'oiseau chanteur avec l'axe du micro. Tests réalisés avec haut-parleur aligné, une distance angulaire de 45° ou 90° et pour chaque cas une distance de 10 m, 25 m 50 m entre haut-parleur et le téléphone.
- L'effet **milieu** va il affecter la qualité des sons ? Autrement dit, quelle est la distance

à laquelle signal pourra être enregistré correctement par un téléphone et, in fine, être traité correctement par un algorithme de classification ?

Tests réalisés en forêt, en milieu ouvert et avec une haie entre le téléphone et le haut-parleur.

- L'effet **du vent** . Le vent perturbe t-il beaucoup la prise de son ? Mais comment cela affecte-il l'efficacité de la classification ? Une bonnette permet-elle d'améliorer l'enregistrement ? Tests réalisés sans vent, avec une faible brise et un vent fort, avec et sans bonnette.

A l'aide de mon co-encadrant, Thierry Lengagne, nous avons mené des expériences à Lyon au printemps 2018 afin de répondre à ces problématiques. Une application de "dictaphones" a été téléchargée pour cinq téléphones de marques différentes (Sony, Samsung, Wiko, Iphone, Motorola). L'ensemble des téléphones ainsi que le micro SMiOT est présenté Fig.2.9.



FIG. 2.9 – Téléphones utilisés ainsi que l'enregistreur type "Zoom". Le cinquième téléphone est utilisé pour prendre la photo. Photographie tirée du rapport de recherche Poupard et al. (2018c).

En ce qui concerne l'émission, nous avons construit un signal de 12,5 minutes contenant des espèces seules, puis par couples à différents niveaux d'énergie (Fig.2.10) émise à l'aide d'une enceinte (Ohms Monacor SPH-60x).

Les espèces émises étaient les suivantes : la Mésange à longue queue, l'Étourneau sansonnet, la Mésange charbonnière, la Mésange bleue, le Pigeon ramier, la Tourterelle turque, la Fauvette à tête noire, le Rouge-gorge, le Rouge-queue à front blanc et le Troglodyte mignon.

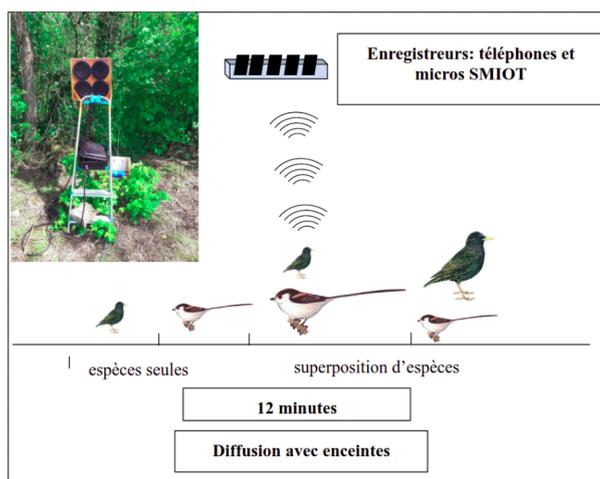


FIG. 2.10 – Schéma représentant les émissions sonores, des espèces seules, puis des couples à différents niveaux d'énergie ainsi que la photographie de l'enceinte émettrice.

Les résultats de ce protocole nous ont permis de valider ou non l'efficacité des microphones des téléphones pour la détection d'espèces aviaires. Les résultats de ce protocole sont présentés dans le chapitre "Résultats" (4.1.1).

2.2.1.2 Stratégie d'écoute pour la détection de Cachalots (*Physeter macrocephalus*)

La détection d'espèces se fait aussi en milieu sous marin. La communauté scientifique est depuis 30 ans en mesure d'écouter, d'enregistrer, et de stocker le son de certains cétacés. Deux principales catégories d'enregistrements sont utilisées : les systèmes d'enregistrements autonomes et les hydrophones câblés embarqués (Mellinger et al., 2007). Les hydrophones embarqués sont typiquement déployés depuis un navire, leur installation n'est donc pas permanente. Le principal inconvénient de ce système est le bruit occasionné par le bateau de recherche scientifique, par le câble de l'hydrophone et par la variation de la pression de l'eau sur le capteur.

Il faut donc mettre en place des protocoles particuliers afin d'éloigner au maximum l'hydrophone du moteur du bateau tout en restant éloigné de la surface. Ce type de bruit est le plus souvent en basse fréquence, il est donc possible par la suite de filtrer les bruits d'eau et du bateau (Sayigh et al., 1993). Plus récemment, le développement de systèmes d'enregistrements autonomes (ARS, Autonomous underwater sound recorders) ouvre de grandes perspectives afin d'étudier les espèces sur le long terme. Ces systèmes permettent d'enregistrer les espèces de manière totalement autonome (stockage et alimentation interne) pendant des jours, des semaines ou encore plusieurs mois.

Ces enregistreurs autonomes présentent un avantage considérable pour l'enregistrement des écosystèmes, notamment dans leur efficacité lorsqu'il faut enregistrer des écosystèmes difficile d'accès, comme les régions polaires, ou autres régions éloignées (Oswald et al., 2004). Ces systèmes sont donc indépendants de la présence de l'utilisateur qui peut nuire à la détection d'espèces. En effet la présence humaine et le bruit anthropique peuvent gêner les animaux et provoquer l'apparition d'un biais de détection. Le principal inconvénient de ce type de système

est que l'analyse des enregistrements en temps réel est impossible. Les données sont récupérées, puis analysées en laboratoire. De nombreux laboratoires tentent de réaliser leur propre système d'enregistrement afin de pouvoir réaliser les protocoles en totale autonomie (Calupca et al., 2000 ; Wiggins, 2003 ; Fourniol et al., 2018). Ils adaptent également les paramètres d'enregistrement comme la fréquence d'échantillonnage (Fe).

Prenons l'exemple de certains mysticètes, qui produisent des vocalises inférieures à 500Hz, la Fe peut donc être faible (1kHz) (Wiggins, 2003). Tandis que la plupart des odontocètes nécessite une Fe d'au moins 48kHz (Oswald et al., 2004). Il est également possible de faire varier le temps d'enregistrement en programmant des temps de pause (exemple : 5 min d'enregistrement et 30min de pause). Ce type de protocole a l'avantage d'économiser du stockage, de la batterie et du traitement d'analyse.

Il faut également prendre en compte le profil de propagation du son de la zone d'étude (Sousa-Lima et al., 2013). Prenons l'exemple des Baleines à bosse pouvant fréquenter des zones peu profondes (inférieures à 100 m), les enregistreurs peuvent être directement posés sur le socle sableux. Puis certains paramètres anthropiques doivent également être pris en compte comme le trafic maritime (ferry qui peuvent endommager les bouées), les activités de pêche et touristiques.

La détection d'espèce sous-marine dans le cadre d'études en bioacoustique portera sur la mise en place d'une bouée sous-marine au large de Toulon (Bombyx).

La Bouée BOMBYX a été posée début 2014, par l'équipe DYNI de l'université de Toulon en collaboration avec le MIO (l'Institut Méditerranéen d'Océanologie) et le Parc National de Port-Cros. Elle est située à 42 ° N et 6 °19 E à 25 m sous la surface, 275 m au dessus du fond, en haut d'une falaise de 2 km de profondeur.

La station est équipée d'un hydrophone HN1 Neptune et d'un hydrophone HTI (compagnie High Tech, réponse en fréquence présentée en Annexe B) ainsi qu'un enregistreur provenant de la société OSEAN. L'échantillonnage était de 50 kHz Fe sur 2 voies en 24 bits. Les enregistrements ont été stockés sur les NAS de l'équipe DYNI.

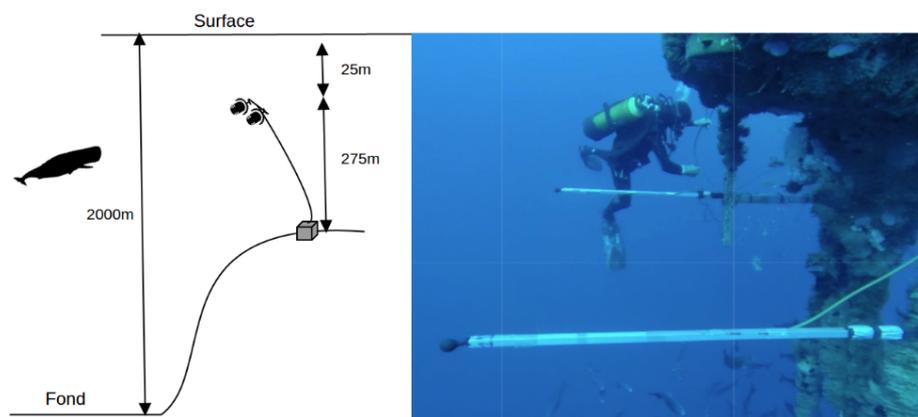


FIG. 2.11 – Gauche : Schéma de la station d'acquisition BOMBYX au large de Port-Cros. Droite : installation des 2 hydrophones.

L'orientation de la bouée est relativement stable et son axe prend la direction de 230 degrés

grâce au courant Ligure (Rougier et al., 2015) (cf Fig.2.12). L'écart entre les deux hydrophones était de 1.83m.

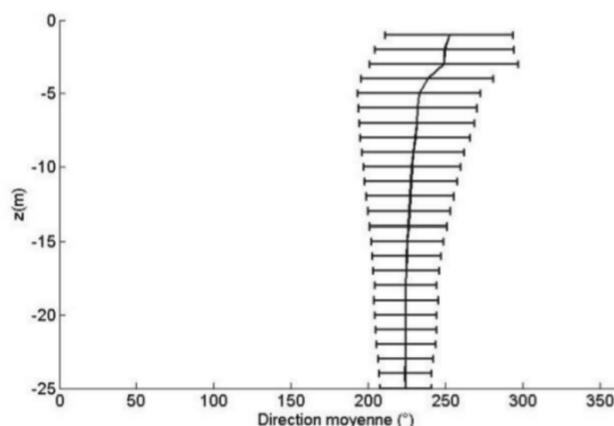


FIG. 2.12 – Profil vertical de la direction moyenne des vitesses en 2015, les barres horizontales représentent l'écart-type. Figure tirée de Rougier et al. (2015).

La première session d'enregistrement de BOMBYX fut en juillet 2015, et la dernière en décembre 2018. Le récapitulatif des sessions est présenté dans le tableau 2.1. En fonction des sessions, différentes stratégies d'enregistrements ont été utilisées. La colonne "échantillonnage" présente le pourcentage en temps des heures d'enregistrements, allant de 16 à 100%.

TAB. 2.1 – Récapitulatif des enregistrements réalisés avec BOMBYX. La session B ne sera pas utilisée dans ce manuscrit (pas d'accès aux enregistrements mais déjà traitée dans le rapport Glotin et al. (2017)).

Sessions	Date Début	Date fin	durée rec (min)	durée pause (min)	échantillonnage (%)	durée totale
Session A	06/07/15	16/07/15	10	20	33	43h30
Session B	07/07/16	23/09/16	5	25	16	316h
Session C	17/11/16	03/04/17	1	5	16	340h
Session D	10/11/17	16/10/17	5	15	25	26h25
Session E	21/12/17	10/04/18	5	15	25	661h
Session F	08/06/18	11/11/18	5	15	25	626h50
Session G	28/11/18	26/12/18	14	0	100	682h
Total						2695h45 min (env 112 jours)

L'objectif de cette bouée est donc de détecter les Cachalots fréquentant la zone. Les deux matériels pour la détection d'espèces ont été présentés dans cette partie : l'un pour la

détection d'espèces d'oiseaux, l'autre pour la détection de Cachalots. Je présenterai les méthodes de détections automatiques et semi-automatiques pour la détection d'espèces dans le chapitre Méthodes (Partie 3.2).

2.2.2 Matériel utilisé pour la détection de population

Comme précisé dans l'introduction, la détection acoustique de populations repose sur la détection de dialectes particuliers. Ces derniers sont répartis la plupart du temps en différentes régions géographiques. Ces variations sont beaucoup moins documentées pour les mammifères que pour les oiseaux.

Les espèces avifauniques, peuvent être détectées au niveau populationnelle avec deux principaux concepts :

le premier est l'existence de différences dans la structure des signaux émis par des animaux d'une même espèce vivant sur des sites différents (Nottebohm, 1972). Il existe donc des barrières géographiques où les populations ne se mélangent pas. Le deuxième est la présence de dialectes différents pour deux populations qui évoluent dans une même aire géographique. La mise en place de matériel pour la détection de dialecte peut donc reposer sur deux principales stratégies :

- poser des capteurs à des endroits géographiques différents afin de comparer les signaux entre les régions pour une seule espèce,
- poser un/plusieurs capteurs au sein d'une seule région et essayer de trouver des différences au sein des signaux pour différentes populations (évoluant dans la même aire géographique).

Ces deux types de stratégies peuvent être appliqués pour des problématiques terrestres et marines.

Les espèces avifauniques sont beaucoup plus documentées que les mammifères en terme de dialecte. Ces attributions de dialectes auraient les mêmes évolutions que les dialectes ou accents chez les humains (Podos and Warren, 2007). Cette diversité de signaux reposerait donc sur la flexibilité et capacité chez les individus pour innover, apprendre mais également repose sur des transmissions génétiques et des facteurs écologiques (Cavalli-Sforza et al., 1988).

L'occupation géographique de la population est donc un élément à considérer pour anticiper la pose des capteurs. Cette taille correspond donc à une unité de cohésion sociale (Payne, 1978) qui peut être anticipée avant la mise en place d'une étude. Si la zone de dialecte est trop grande, les résidents "locaux" peuvent avoir du mal à se différencier. Podos and Warren (2007) a regroupé une quantité importante d'études concernant l'évolution des dialectes chez les oiseaux et a hiérarchisé les dialectes en fonction de leur taille. L'échelle des dialectes pourrait donc être visualisée en 4 catégories : microgéographique (1), petite (2), moyenne (3) et large (4). La catégorie 1 doit comporter jusqu'à 10 individus dans une aire géographique inférieure à 2 km, et la dernière catégorie doit être supérieure à 100 km.

En terme de microphones, il existe énormément de systèmes autonomes qui peuvent être efficaces (Priyadarshani et al., 2018a). En effet, la mise en évidence des dialectes nécessite des enregistrements clairs et précis, avec peu de bruit afin d'identifier les paramètres acoustiques qui seront propres au dialecte. Les bioacousticiens doivent considérer :

- La qualité de l'habitat (ouvert, fermé etc..)
- L'espèce étudiée

- Bruits naturels (exemple d'un cours d'eau)
- Bruits anthropiques (exemple d'un trafic routier).

Différents microphones peuvent être utilisés comme les microphones canon très directifs, qui ont la capacité de cibler l'espèce intéressée. Nous pouvons également utiliser un microphone équipé d'un système réflecteur parabolique afin d'amplifier l'espèce cible.

Le matériel disponible est souvent conçu pour l'enregistrement de voix humaines et non de la faune. La Fig.2.13 représente quelques exemples d'enregistreurs utilisés couramment dans la recherche de dialecte avifaunique.

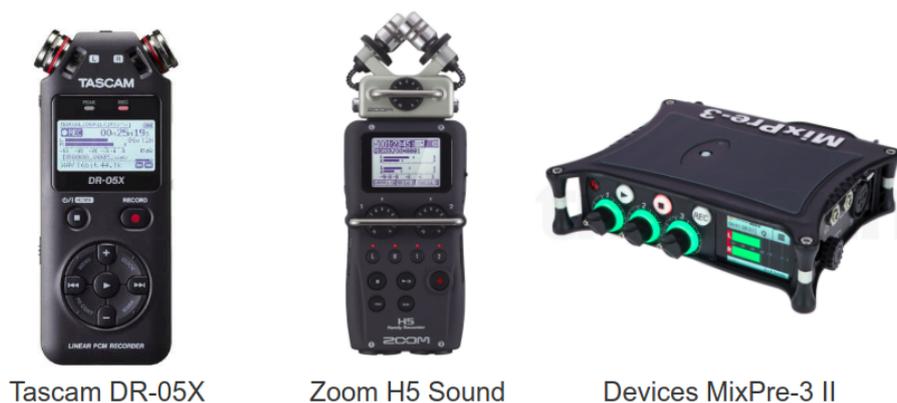


FIG. 2.13 – Enregistreurs les plus couramment utilisés pour l'acquisition de sons d'oiseaux.

Chez les mammifères, la notion de dialecte a notamment été mise en évidence chez certaines espèces comme les primates (Mitani et al., 1992), les Orques (Ford, 1991), et les Cachalots (Weilgart and Whitehead, 1997).

Afin de mettre en évidence des dialectes chez une espèce de dauphin, Campbell (2004) a enregistré des animaux provenant de 3 régions très éloignées. Et afin d'obtenir des échantillons représentatifs pour chacune des régions, ils ont enregistré les animaux dans différentes conditions comportementales. Le matériel pour les 3 régions était identique afin de minimiser les différences liées aux caractéristiques techniques du matériel utilisé (courbe de réponse en fréquence, rapport signal/bruit, directivité, etc). Cette même stratégie a été utilisée chez le chimpanzé pour la mise en évidence de dialecte chez deux populations isolées (Mitani et al., 1992).

Afin d'illustrer mes propos sur la mise en évidence des dialectes acoustiques, un protocole a été mis en place au Canada et sera développé dans la partie suivante.

2.2.2.1 La station d'enregistrement à Orcalab

Le matériel présenté dans cette partie s'applique également pour la détection d'espèces appliquée aux Orques de Colombie-Britannique.

La station de recherche Orcalab a été créée par le docteur Paul Spong en 1970 et est maintenue avec Helena Symonds. Elle est située au Nord de l'Île de Vancouver, sur l'île de Hanson. Elle a pour principal objectif d'étudier les Orques dans leur milieu naturel.

Depuis 20 ans, Orcalab a mis en place une station multihydrophone composée de 5 capteurs (Fig.2.15) recouvrant plus de 50 km² d'océan afin d'enregistrer différents pods fréquentant la

zone (particulièrement la communauté Nord résidente). Les différents hydrophones sont posés à quelques mètres sous la surface et enregistrent les animaux sans les perturber et les gêner par d'éventuels bruits de bateaux. Les Orques sont présents pendant l'été. Orcalab recrute donc des écovolontaires afin de monitorer les sons enregistrés, mais également de surveiller les caméras présentes au dessus de chaque hydrophone (présentes depuis 2018).

Depuis 2015, l'équipe DYNI a mis en place sur site, l'enregistrement en continu de tous les hydrophones, la compression et leur envoi en continu sur nos serveurs, afin d'observer et de modéliser les activités acoustiques des espèces de cétacés fréquentant la zone sur une large échelle temporelle et spatiale. L'architecture du système est montrée Fig.2.14.

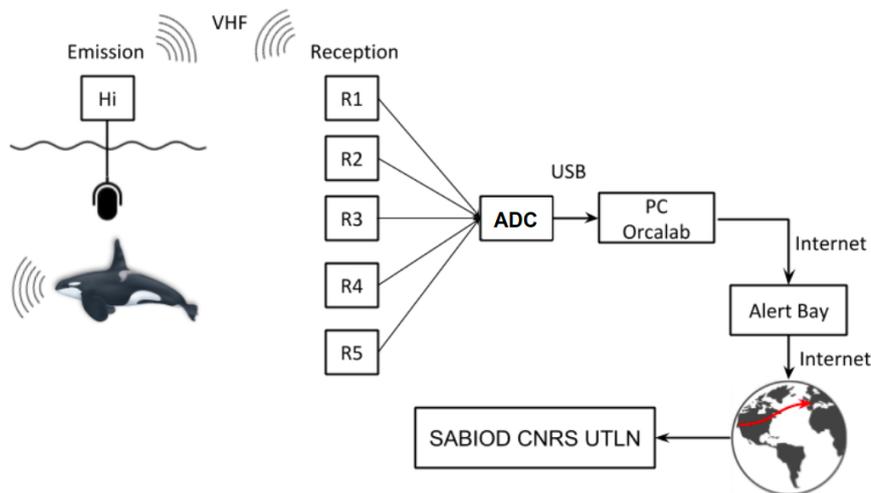


FIG. 2.14 – Représentation de l'acquisition des données, des enregistrements à Orcalab jusqu'à l'archivage sur les serveurs DYNI à Toulon.

Les hydrophones enregistrent le paysage sonore en continu, puis ils sont transmis à la station Orcalab en temps réel via VHF (Radio Haute Fréquence). Chaque signal analogique est reçu par un récepteur à Orcalab, puis numérisé par un convertisseur analogique-numérique Presonus (ADC) et envoyé à un ordinateur portable DELL Latitude Linux. Les enregistrements sont ensuite séparés en segments de 2 minutes comprenant les 5 canaux (avec compression sans perte .flac). La fréquence d'échantillonnage est de 22050Hz avec une résolution de 16bits. Chaque segment est ensuite envoyé via internet en temps-réel sur un NAS (serveur de stockage en réseau) de 300 To à l'Université de Toulon. Un stockage tampon de 7 jours sur l'ordinateur portable local permet de faire face aux interruptions de la connexion internet locale. La transmission de données est automatiquement poursuivie dès que la connexion est stabilisée.

Au total, de Juillet 2015 à ce jour (Novembre 2020) ; environ 50To de sons (environ 14 500 h) ont été archivés sur notre serveur.

Il est donc impossible de traiter ces données manuellement, l'automatisation des traitements est indispensable pour ce jeu de données et mes recherches s'y sont en parties consacrées.

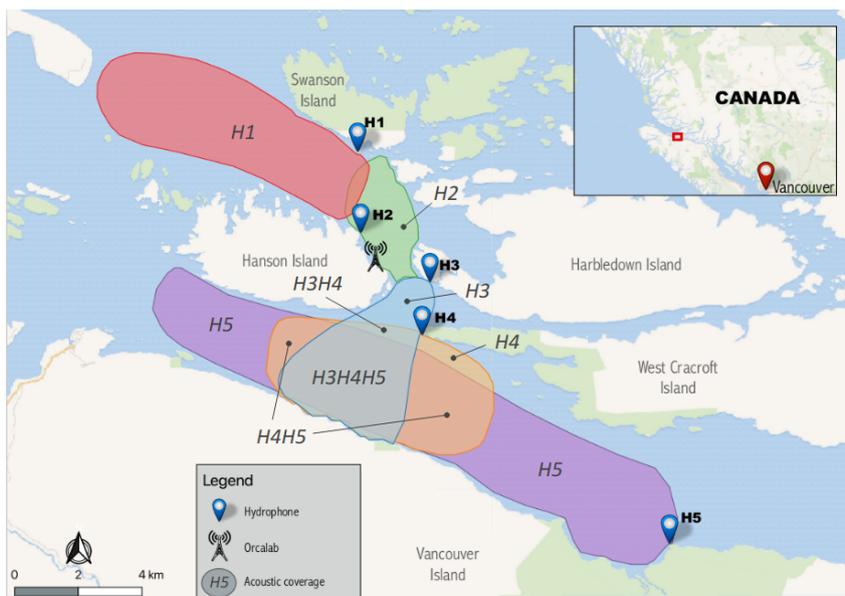


FIG. 2.15 – Carte de la zone et des plages d'écoutes des 5 hydrophones. Les repères de H1 à H5 en gras indiquent l'emplacement des hydrophones. Les zones de détections acoustiques estimées sont indiquées en couleurs (estimées par les experts d'Orcalab).

L'objectif donc de ce protocole est double : le premier est d'identifier les Orques présentes (bioacoustique spécifique), mais également l'identification des clans acoustiques (bioacoustique populationnelle). Donc les enregistrements réalisés à Orcalab seront utilisés dans la partie "Détection d'espèces" mais également dans la partie "Détection de populations".

2.2.3 Matériel utilisé pour l'identification individuelle

La troisième échelle d'analyse est le niveau individuel.

L'identification individuelle acoustique repose sur deux stratégies : identification grâce à la localisation précise de l'individu, ou identification grâce à une signature individuelle précise et reconnue. Ces deux stratégies peuvent être utilisées dans des écosystèmes terrestres et marins, mais elles ne nécessitent pas le même matériel d'enregistrement.

Pour la détection de signatures individuelles chez les espèces avifauniques, le matériel peut être constitué uniquement d'un microphone, d'un enregistreur, d'une capacité de stockage et d'énergie suffisante (Jouventin et al., 1999 ; Lengagne et al., 2001 ; Mathevon et al., 2003).

Chez les cétacés, et particulièrement les odontocètes, la mise en évidence de signature individuelle a été beaucoup étudiée pour des animaux en captivité comme en liberté (Duarte de Figueiredo and Simão, 2009 ; Smolker et al., 1993). En terme de matériel, ces études ont nécessité qu'un seul hydrophone. Différentes méthodes ont été utilisées afin de détecter ces signatures.

La deuxième méthode d'identification individuelle nécessite la localisation des animaux. Une fois le signal localisé, il peut être attribué à un animal précis.

Différentes méthodes peuvent être utilisées afin de localiser une cible, mais la plus commune nécessite plusieurs capteurs (Janik et al., 2000 ; Miller and Tyack, 1998 ; Filatova et al., 2006 ; Gedamke et al., 2001). Ces études comportent donc au moins 2 hydrophones, pouvant être fixés au sol ou tractés derrière un navire de recherche. Récemment, les scientifiques ont mis au

point des méthodes et surtout du matériel pour poser des hydrophones directement sur l'animal. En effet, les tags acoustiques sont de plus en plus compacts et peuvent détecter une position précise de l'animal pendant un temps plus ou moins long (Johnson et al., 2009 ; Witteveen et al., 2008 ; Madsen et al., 2002). Ce type de méthode ne sera pas approfondie dans cette thèse car ces systèmes peuvent provoquer un changement de comportement de l'animal et donc nuire à l'analyse.

Afin d'étudier les avantages et les inconvénients de ces deux méthodes (tractée et fixe) deux protocoles sous-marins ont été développés durant ma thèse pour la localisation individuelle dans le cadre d'études en bioacoustique.

La première a eu lieu en Méditerranée pour la localisation de Cachalots, et la deuxième au Canada pour la localisation d'Orques.

Puis nous avons mis en place un protocole terrestre pour la localisation individuelle d'espèces aviaires.

2.2.3.1 Mise en place d'une antenne fixe pour la localisation individuelle sous-marine

L'objectif principal du système que nous avons mis en place à la station de recherche Orcalab (OL) était d'identifier les vocalises émises individuellement par les Orques devant la station.

Pour cela, nous avons construit une antenne acoustique permettant l'enregistrement des individus devant la station de recherche Orcalab (OL) et de relier ces données aux observations visuelles faites depuis la terre à OL.

L'objectif était l'estimation des délais d'arrivée (TDoAs) des vocalises, donc leur azimut, et l'attribution de labels multimodales pour chacune des vocalises :

le type d'appel, l'azimut et l'identité de l'Orque émettrice.

Le réseau de quatre hydrophones était une antenne tétraédrique de 12 kg (Fig.2.16) composée de quatre hydrophones SQ26-01 avec une sensibilité entre 10Hz et 50 kHz que j'ai placé à l'aide des membres d'Orcalab à 25 m de profondeur devant la station de recherche. L'ouverture maximale de l'antenne est de 78,5 cm. Les dimensions de l'antenne sont précisées en annexe D et les caractéristiques des hydrophones en Annexe B.



FIG. 2.16 – L'antenne à quatre hydrophones, avec le câble 12 V (vert) connecté à la station de recherche Orcalab.

Nous avons alimenté directement ce système depuis la station via un câble de 75 m que nous avons préparé sur site et protégé des risques d'érosion avec un simple tuyau d'arrosage.

Le système comprend la carte son autonome haute résolution basse consommation JASON (Fig.2.17), ayant une capacité d'enregistrement de jusqu'à 5 canaux non compressés à une fréquence d'échantillonnage allant jusqu'à 2 MHz, 16 bits.

Cette carte peut être accompagnée d'un déclencheur acoustique de faible puissance (Fourniol et al., 2018), mais ceci n'étant pas indispensable à OL car JASON était allumée manuellement à partir de la station lorsque des Orques étaient vues.



FIG. 2.17 – Carte d'acquisition Jason (SMIoT), 5 x 2 MHz. Le luxmètre n'a pas été utilisé pour ce protocole.

Dans cette étude, nous avons enregistré à une fréquence d'échantillonnage de 192 kHz sur quatre canaux. Cette haute Fe nous a permis d'augmenter la précision de l'angle d'arrivée estimé et donc d'accroître l'efficacité de la séparation des appels de façon individuelle. L'antenne a été placée à une profondeur de 24m pointant vers l'Est, déterminé à l'aide d'un compas. En raison de sa densité, il n'y avait aucune inquiétude concernant sa stabilité, même lors de forts courants (> 1 m/s).

La possibilité d'alimenter l'appareil depuis la terre nous permettait de limiter notre enregistrement au moment où les Orques passaient devant Orcalab et de sauver la capacité de stockage. L'antenne a été retirée de l'eau lorsque le stockage 512 Go uSD était plein.

Lorsque des Orques étaient repérées dans le Blackney Pass (BP), situé en face d'Orcalab, l'antenne était alimentée par un simple bouton marche / arrêt depuis le laboratoire et il était éteint lorsque les Orques n'étaient plus visibles.

Lors de chaque passage d'animaux, plusieurs observateurs de Orcalab ont travaillé ensemble afin d'identifier les groupes et les individus présents (voir figure 2.18). Pour faciliter la collecte visuelle des données, les observateurs de OL ont utilisé plusieurs lunettes et jumelles, y compris une jumelle "Big Eyes" dédiée à ce protocole ainsi que des caméras fixes et vidéos pour l'identification en temps réel et post-traitement des Orques.



FIG. 2.18 – Protocole d'acquisition visuel à Orcalab. Crédit Y. Rose.

Afin de suivre visuellement les animaux, nous utilisons les azimuts des jumelles "Big Eyes" et le nom de chaque point de référence visuel devant orcalab (voir figure 2.19). Chaque site (exemple "Green House") correspondait à un azimut précis (15 degrés) et nous pouvions donc suivre visuellement en temps réel les individus.

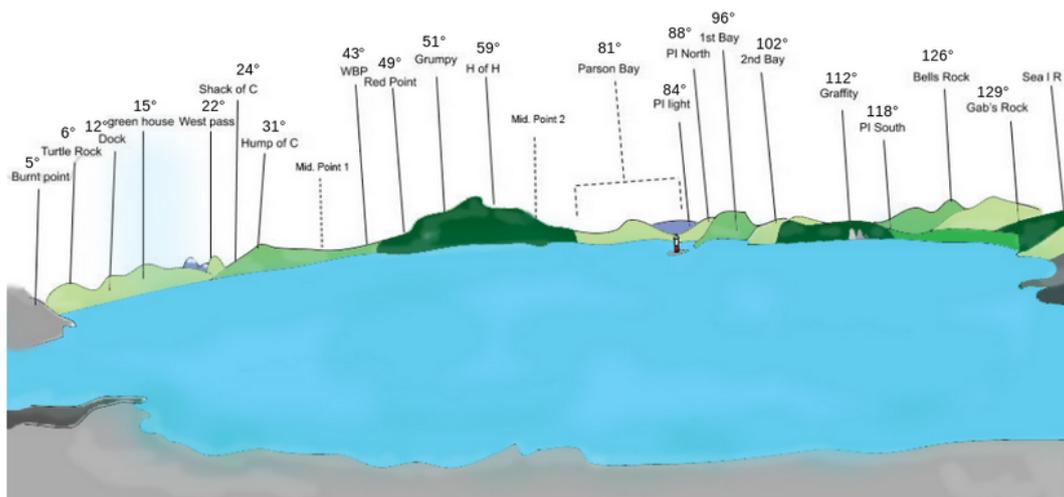


FIG. 2.19 – Représentation de la vue du Blackney Pass depuis la plateforme d'Orcalab.

Selon le passage des bateaux et des Orques observées visuellement, la portée d'écoute de l'antenne était d'environ 2km, donc similaire à la portée visuelle (voir Fig.2.20).

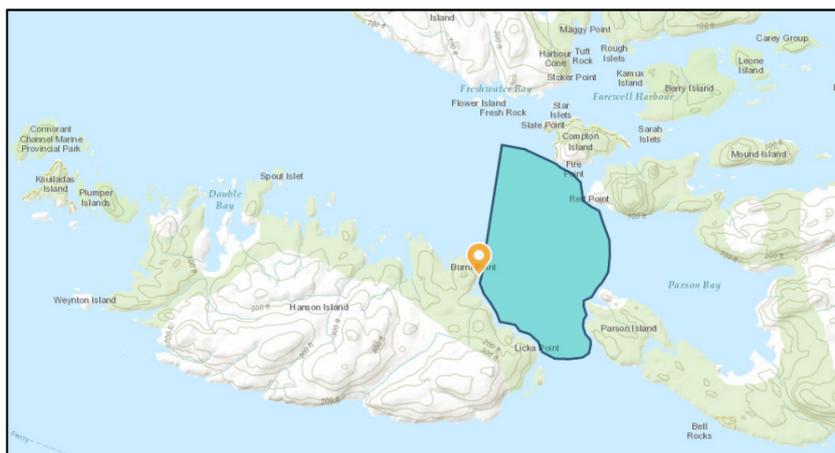


FIG. 2.20 – Carte des plages visuelles et d'écoutes, couvrant une zone d'environ 5km² avec une possible estimation azimutale audio et visuelle des passages des Orques.

Ce protocole a donc été mis en place durant l'été 2019, où j'ai pu séjourner à la station de recherche Orcalab. A chaque passage d'Orques, le même protocole était respecté :

- Mise en tension de l'antenne depuis Orcalab
- Note de la date et de l'heure (à la seconde près)
- Helena Symonds identifiait en temps réel la position de chaque individu dans le groupe. J'enregistrais ses commentaires à l'aide d'un dictaphone
- Une personne dans la station suivait les Orques à l'aide des caméras
- Une personne prenait des photographies pour confirmer la présence de certains individus après les passages
- Deux autres personnes participaient également à l'identification des individus en temps réel

Après chaque passage, l'ensemble des identifications avec le temps était rapporté sur un document. Le nombre de bateaux présents, la vitesse et le comportement des individus étaient relevés.

L'analyse des enregistrements sera présentée dans le chapitre Méthode (partie 3.4.1).

2.2.3.2 Mise en place d'une antenne mobile pour la localisation individuelle sous-marine

Les projets Sphyrna Odyssey 2018 et 2019-2020 ont été réalisés en collaboration avec la société SeaProven et l'équipe DYNI. Expert en robotique marine, l'entreprise SeaProven a mis au point le Sphyrna qui est le plus grand drone de surface civil au monde. Sa conception et son design le rendent autonome sur de grandes distances. Les formes de ce bateau sont inspirées de coques polynésiennes (Fig.2.21), il mesure plus de 16m et est réalisé en matériaux composites. Les caractéristiques du Sphyrna sont les suivantes :

- Tonnage : 4,81 tonnes
- Taille : Longueur : 16,75 m / Largeur : 4 m
- Vitesse maximale théorique : 12 noeuds
- Vitesse de mission : entre 3 et 5 noeuds
- Déplacement : 1,4 tonnes

- Tirant d'eau : 20 cm.



FIG. 2.21 – Drone Sphyrna en mer (Crédit P. Cottaz).

Les hydrophones étaient reliés à la carte son Jason (Fig.2.17). Cette carte d'acquisition acoustique permet l'enregistrement de 5 voies simultanément avec une fréquence d'échantillonnage très élevée (plus de 1 MHz). Pour ce protocole la fréquence d'échantillonnage était de 300 kHz, sur 16bits x 5 hydrophones.

La carte est équipée d'un accéléromètre 3D, d'une centrale inertielle 3D et d'un compas 3D synchrones à l'acquisition acoustique pour connaître l'assiette et le cap précis de l'antenne acoustique tous les 10 millisecondes.

En mer, l'ensemble des enregistrements étaient stockés sur des disques durs, puis copiés sur les NAS de l'université de Toulon.

La position des hydrophones a été réfléchié par l'ensemble de l'équipe suivant le succès des calculs de trajectoires de Cachalots avec un cube d'hydrophones de 2 m de côté (Bénard-Caudal et al., 2010). L'antenne profilée a été dessinée en 2018. J'ai placé les cinq hydrophones sous le drone à quai.

H1 à H4 sont 2 Cetacean Res. CR75 et 2 CR57 placés sous la quille à une profondeur de -2 m, espacés de 35 à 70 cm. H5 est un CR75 à la poupe du drone sur la coque, à une profondeur de -0,2 m (Fig.2.22).

Les distances entre chaque hydrophone sont décrites en annexe D et les réponse en fréquence en Annexe B.

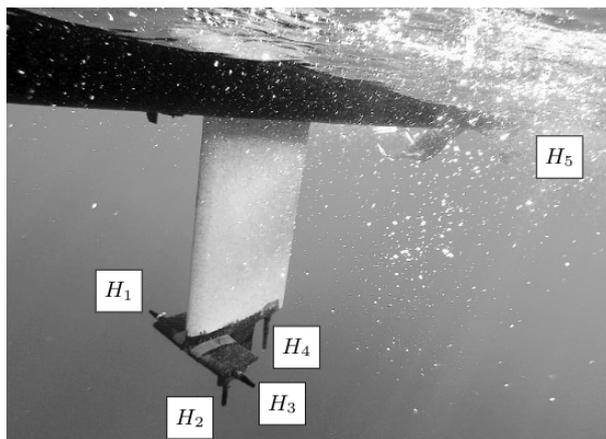


FIG. 2.22 – Disposition du réseau d'hydrophones 4 + 1. Les distances interhydrophones sous la quille sont de 35, 59, 59, 63, 63, 70cm. L'hydrophone H5 est placé à l'arrière, à 7 m de distance de H4

J'ai réalisé, avec le directeur de la mission (F. de Varenne), la maintenance des hydrophones (nettoyages afin d'éviter la présence de masques biologiques sur les capteurs comme des algues, bivalves etc..).

L'équipe scientifique ainsi que des skippers suivaient le drone à distance grâce à un trimaran mis à disposition pour l'étude.

La mission s'est déroulée durant l'été 2019, au moins deux scientifiques étaient à bord afin de vérifier le système, choisir les zones d'écoutes et mettre en place le protocole acoustique. Les enregistrements étaient réalisés lorsque le drone se trouvait en dérive afin d'éviter les bruits du moteur et de flux d'eau sur les hydrophones.

Nous avons donc prospecté durant plusieurs semaines au large de Toulon à la rencontre de Cachalots.

Les étapes pour l'analyse des signaux sont décrites dans le chapitre Méthode (partie 3.4.2).

2.2.3.3 Mise en place d'une antenne fixe pour la localisation individuelle en milieu terrestre

Lors de cette thèse, nous avons essayé de mettre en place un protocole pour la localisation d'espèces avifauniques. Ce protocole a été mis en place à La Garde, non loin de l'Université. L'objectif était de compter le nombre de sources sur un enregistrement composé de plusieurs oiseaux. Pour cela, nous disposons de quatre voies (96 kHz 16 bits), d'un enregistreur Tascam avec deux micros internes et de deux autres micros externes (Tbone LC 97 TWS). La distance maximale entre les microphones était de 21cm. La figure 2.23 illustre le matériel utilisé.



FIG. 2.23 – Protocole réalisé pour la détection de source en milieu terrestre.

L'objectif de ce protocole était de créer un jeu de données pour le challenge LifeClef 2019 où les challengers auraient trouvé le bon nombre de sources avec un seul des quatre microphones. Pour cela, nous avons laissé l'antenne enregistrer l'écosystème pendant 3 h. Puis nous avons mis en point un jeu de données avec l'ornithologue Thierry Lengagne qui a annoté des fichiers afin de compter automatiquement le nombre d'individus dans chacun d'entre eux. Notre jeu de données était donc de 14 enregistrements (de 3 à 20 secondes) :

- 2 enregistrements sans oiseau
- 6 enregistrements avec 1 individu
- 4 enregistrements avec 2 individus
- 2 enregistrements avec 4 individus.

L'objectif était donc de mettre en place des méthodes d'analyses afin de compter le nombre d'individus sur ces 14 enregistrements (Méthodes présentées dans la partie 3.4.3).

2.2.4 Matériel pour la bioacoustique comportementale

Pour cette quatrième échelle d'étude, l'objectif est de détecter des signaux biologiques et de les mettre en corrélation avec des comportements particuliers. Aucune espèce, population, ou individu n'est détectée **acoustiquement**. Un support visuel sera donc indispensable.

Pour cela, en terme de matériel, il est nécessaire d'acquérir des données qualifiables de multimodales. Deux grands types de données vont être fondamentales pour la bioacoustique comportementale :

- L'acquisition acoustique avec au moins 1 hydrophone/microphone
- L'acquisition de données dites "environnementales" influençant les signaux émis.

Ce dernier type d'acquisition peut être très variable en fonction de la problématique étudiée. Pour cela, il est nécessaire de définir le matériel nécessaire pour l'étude comportementale associée à une émission acoustique.

Pour les cétacés beaucoup d'études ont mis en place des protocoles précis pour fusionner des données acoustiques et visuelles, particulièrement pour le Grand dauphin (*Tursiops truncatus*) (Gridley et al., 2016 ; Herzing, 2000).

Dans Hawkins and Gartside (2010), les scientifiques ont associé des évènements et des comportements à des émissions acoustiques. Pour cela ils ont du relever les comportements de surface des individus, la taille et la composition des groupes. Afin de qualifier les comportements qu'ils ont regroupés en 4 classes différentes : mouvement, socialisation, repos et chasse. La même méthode a été utilisée pour l'étude de Gridley et al. (2016).

Chez les oiseaux également, l'association d'un comportement avec une émission acoustique peut être faite. Certaines études ont mis en évidence des chants associés à un comportement de reproduction (Anderson et al., 2008) ou à un comportement de défense (Magrath et al., 2010).

Cette partie en bioacoustique comportementale a fait l'objet de deux cas pratiques sous-marins. Le premier se focalisera sur l'impact des comportements de surface sur l'émission des sifflements des Dauphins tachetés pantropicaux (*Stenella attenuata*). Et le deuxième portera sur l'influence de stimuli chimiques sur l'émission acoustique des Baleines à bosse (*Megaptera novaeangliae*).

2.2.4.1 Matériel pour la bioacoustique comportementale des Dauphins tachetés pantropicaux

Cette étude permet la compréhension des modalités de communication d'une espèce de dauphin localisée en Martinique en présence de whale watching.

Le whale watching est une activité touristique récente qui a pour objectif l'observation des mammifères marins en milieu naturel. L'espèce la plus fréquemment observée en Martinique, est le Dauphin tacheté pantropical : *Stenella attenuata*, qui est un dauphin cosmopolite dans les eaux tropicales et chaudes (Shirihai, 2007).

La Martinique est un site de prédilection pour étudier les conséquences du whale watching sur les animaux car dans cette zone la faune marine est très abondante et diversifiée. En effet, plus de 25 espèces de mammifères marins sont présentes. Mais cette île Française située dans les petites Antilles (position GPS : 14 48'03 N ; 061 07'70 W), encadrée par la mer des Caraïbes (à l'Ouest) et par l'océan Atlantique (à l'Est) est également concernée par la pression anthropique liée au tourisme.

Le tourisme maritime dans la mer des Caraïbes augmente (Mayol et al., 2015) : actuellement 25 compagnies de whale watching sont actives sur les côtes de Martinique.

L'objectif de cette étude est la mise en place d'une nouvelle méthode d'analyse basée sur les signaux bioacoustiques, qui permettra, en comparant la diversité des sifflements émis par des groupes de dauphins (*Stenella attenuata*), de définir les impacts des perturbations anthropiques et notamment celles du whale watching sur les populations de delphinidés. Afin de mener à bien ce projet, les signaux sonores émis par une population de dauphins tachetés pantropicaux observés en mer des Caraïbes (Martinique) ont été utilisés.

Les enregistrements ont été récoltés entre 2013 et 2017 sur différentes embarcations de tourisme baleinier. Les observations ont été réalisées le matin, sur la côte Ouest de la Martinique. Le matériel utilisé était un hydrophone (H2a-XLR, Aquarian Audio Products) connecté à un enregistreur et à un casque audio (Fig.2.24). Les enregistrements ont été réalisés à 15 m de profondeur, en mono avec une fréquence d'échantillonnage de 44.1kHz.



FIG. 2.24 – Présentation de la zone d'étude et du matériel. 1 : Hydrophone (dans sa protection), 2 : Enregistreur, 3 : Casque audio.

Lorsqu'un groupe de cétacés était repéré, une approche respectueuse des individus était réalisée, selon la « Charte d'approche des mammifères marins en Martinique ». Le bateau ne s'approchait pas à moins de 50 m des cétacés, mais laissait les individus s'approcher en toute liberté. Les enregistrements ont tous été réalisés lorsque le moteur était éteint, en dérive. Chaque enregistrement des individus débutait 15 minutes après la première observation, afin de leur permettre de se familiariser avec la présence des scientifiques et reprendre une activité normale. En plus des enregistrements effectués, des données environnementales étaient relevées :

- La date, le début et la fin de l'observation
- Les coordonnées géographiques
- Le nombre total d'individus ainsi que le nombre d'adultes et de jeunes
- Le nombre de bateaux de whales watching présents
- Les conditions climatiques : état de la mer (calme ou agitée) et le vent (faible, moyen ou fort)
- Le comportement des individus (Tableau 2.2). Les différents états comportementaux ont été répartis en cinq classes : le repos, la socialisation, la chasse, le déplacement et l'évitement.

Ce dernier comportement a été attribué aux individus qui stoppaient leurs activités et essayaient de fuir la zone. Ces classes et leurs descriptions sont présentées dans le tableau 2.2. Nous assignions un certain comportement si la majorité du groupe le suivait.

TAB. 2.2 – Présentation des différents comportements et de leurs descriptions associées.

Comportements	Caractéristiques
Repos	Vitesse de nage lente, direction constante, groupes rapprochés, intervalle de plongée constant
Chasse	Chasse & Vitesse de nage rapide, direction constante, présence d'oiseaux
Socialisation	Événement comportemental interactif, sauts, poursuites, contacts, intervalle de plongée irrégulier, comportement de reproduction
Déplacement	Direction constante, vitesse de nage moyenne et rapide
Évitement	Comportement de fuite, différents groupes, intervalle de plongée irrégulier

Ce protocole nous a permis de mettre en corrélation des comportements d'individus avec des émissions acoustiques.

La détection des signaux (ici des sifflements) sera développée dans le chapitre Méthode (partie 3.5.1).

2.2.4.2 Matériel pour la bioacoustique comportementale des Baleines à bosse

Cette deuxième étude concerne les Baleines à bosse. Plus particulièrement, elle va étudier la réaction acoustique des Baleines à bosse face à deux produits chimiques introduits dans le milieu : l'extrait de krill et le DMS (sulfure de diméthyle).

Les interactions chimiques dans le milieu marin entre la faune et la flore sont courantes. Ce type d'interaction est utile pour maintenir des structures communautaires stables, en particulier pour la connexion entre les prédateurs et les proies. En fait, la stabilité d'une population repose sur des interactions proie-prédateur, et ces dernières peuvent dépendre d'une interaction chimique (Rittschof et al., 1990). L'environnement marin est particulièrement efficace pour la diffusion et la propagation du signal chimique (voir Partie Introduction).

Pour détecter ces signaux, les animaux font appel à deux sens : l'odorat ou le goût. Des études génétiques et anatomiques sur les cétacés ont prouvé que l'odorat ou le goût sont présents, en particulier chez les Mysticètes (McGowen et al., 2008 ; Kishida et al., 2015).

L'objectif de cette étude est donc de déterminer si les Baleines ont développé une réponse comportementale et acoustique à différents stimuli chimiques diffusés à travers d'un appareil flottant. Ce protocole a été réalisé par le CEFE (Centre d'Écologie et Fonctionnel Évolutif), notamment sous la direction d'Aurélié Célérier et de Bertrand Bouchard en collaboration avec l'équipe DYNI pour le volet acoustique. Ils se sont concentrés uniquement sur l'analyse des comportements des baleines. L'équipe DYNI a proposé d'ajouter une surveillance acoustique pour compléter ce protocole car nous imaginions une réponse acoustique possible à ses stimuli et une variation de communication acoustique suite à la présence de nourriture dans la colonne d'eau.

Cette recherche utilise du DMS (sulfure de diméthyle) et des extraits de krill comme stimuli chimiques.

Le premier est un composé volatil produit dans des zones à forte productivité primaire connu pour être perçue par plusieurs espèces d'oiseaux de mer, de pinnipèdes et de tortues. L'extrait de krill est obtenu par hydrolyse enzymatique du krill (*Euphausia superba*), c'est une poudre partiellement soluble, rouge à brun foncé. La taille de la poudre est inférieure à 100 microns. Le krill est la principale proie des Baleines à bosse en Antarctique. Notre hypothèse est la suivante : en présence de stimuli, les baleines n'ont pas la même communication par rapport à un environnement neutre. Si les résultats sont positifs, cela confirmerait que la gustation et/ou l'olfaction peuvent jouer un rôle important dans la détection des proies des Baleines à bosse. Les enregistrements analysés ont été effectués à Madagascar (Fig.2.25) au cours de l'hiver 2016 (juin et juillet). Cette méthode est basée sur 30 minutes d'observation du comportement des Baleines à bosse (distance 200-300 m).

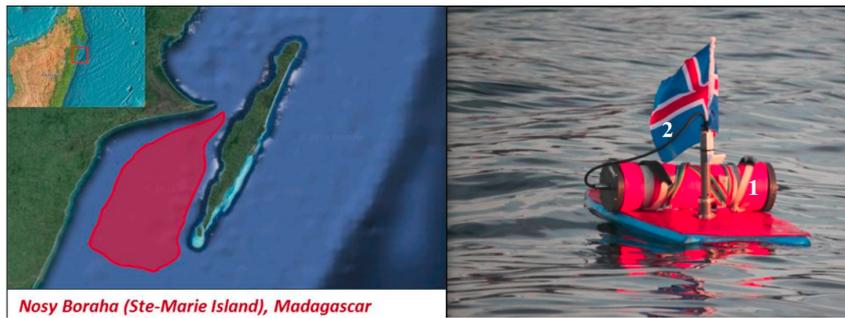


FIG. 2.25 – Gauche : Carte de la zone d'étude. Droite : flotteur contenant une batterie, un enregistreur (1), le câble des hydrophones (2) préparés par l'équipe DYNI (Bouchard et al., 2019).

L'attractivité de ces composés a été comparée à celle d'un solvant témoin : l'argile (témoin du krill) and CTL (témoin du DMS) respectivement.

Quatre kilogrammes de produit sont versés à la surface à 200-300 mètres d'un groupe de baleines. Une tâche rouge brique appelée « zone de stimulus » (Fig.2.26) se forme et se dissout ensuite progressivement dans l'eau sur une surface d'environ 50 mètres de diamètre, imitant une odeur, un goût et une couleur de banc de krill. Des caméras marines flottantes et des hydrophones ont été placés dans la zone (Fig.2.25).

Après cela, les opérateurs s'éloignent (200 m de distance) et le moteur est arrêté pour éviter toute perturbation pendant l'expérience. Pendant 20 minutes, les opérateurs ont analysé les comportements des baleines avec des paires de jumelles et l'hydrophone a enregistré toutes les vocalisations (le modèle de l'hydrophone est un CR57, avec une fréquence d'échantillonnage de 96kHz, 24bits, enregistreur TASCAM DR40).

Après 20-30 minutes d'observation, le flotteur est récupéré et les baleines sont observées pendant 30 minutes supplémentaires, pour s'assurer que l'expérience n'a pas d'impact à long terme sur les comportements. Le stimulus "contrôle" est l'argile pour le krill et le CTL pour le DMS.

Nous avons utilisé 120 minutes d'enregistrements pour cette analyse : 30 minutes avec du krill dans la colonne d'eau, 30 minutes avec de l'argile, 30 minutes avec du DMS et 30 minutes avec du CTL.



FIG. 2.26 – Gauche : Photographie des deux stimuli : Clay (argile) et krill. Droite : Photographie de la zone de stimuli.

Ces deux expériences sur l'influence des comportements sur le Dauphin tacheté et sur l'influence des stimuli chimiques sur les Baleines à bosse sont les deux études qui illustrent l'échelle d'étude de la bioacoustique comportementale.

Cette partie a présenté le matériel nécessaire pour analyser les différentes échelles d'études décrites dans l'introduction :

- La bioacoustique spécifique
- La bioacoustique populationnelle
- La bioacoustique individuelle
- La bioacoustique comportementale.

Pour chaque échelle d'étude, du matériel spécifique a été mis en place et des espèces ont été sélectionnées comme modèle d'étude (cf Tab. 2.3).

TAB. 2.3 – Tableau récapitulatif du matériel utilisé et des paramètres des enregistrements pour chaque cas d'étude.

Échelle	Cas d'étude	Écosystème	Capteur	Nombre de capteurs	Fe	Réponse fréquentielle	Mobile/fixe	Profondeur
Espèce								
	Détection d'espèces avifauniques	Terrestre	Microphones des téléphones portables	1	44kHz		mobile	
	Détection d'Orques	Marin	Hydrophones SQ26, SQ 28	6	44 kHz	20Hz à 30 kHz	fixe	-2m
	Détection de Cachalots	Marin	Neptune Sonar LTD,D140	2	50 kHz	100 Hz 40 kHz	fixe	-25 m
Population								
	Détection de clans d'Orques	Marin	Hydrophones SQ26, SQ 28	6	44kHz	20Hz à 30 kHz	fixe	-2m
Individu								
	Détection d'individus (Orques)	Marin	Hydrophones SQ26	4	192 kHz	20 Hz à 45 kHz	fixe	-25 m
	Détection d'individus (Cachalots)	Marin	3 CR75 et 2 CR57	5	300 kHz	3 Hz à 250 kHz 8 Hz à 100 kHz	mobile	-40cm à -3m
	Détection d'individus (oiseaux)	Terrestre	2 micros interne (Tascam) et 2 Tbone LC 97 TWS	4	96 kHz	20 Hz - 40 kHz	fixe	
Comportementale								
	Bioacoustique comportementale Dauphin tacheté	Marin	H2a-XLR, Aquarian Audio Products	1	48.1 kHz	2 Hz à 30 kHz	mobile	-15 m
	Bioacoustique comportementale Baleine à bosse	Marin	CR57	1	96 KHz	8 Hz à 100 kHz	mobile	-1m

La figure 2.27 regroupe l'ensemble des protocoles d'expérimentations associés à chaque échelle d'analyse ainsi que leur problématiques.

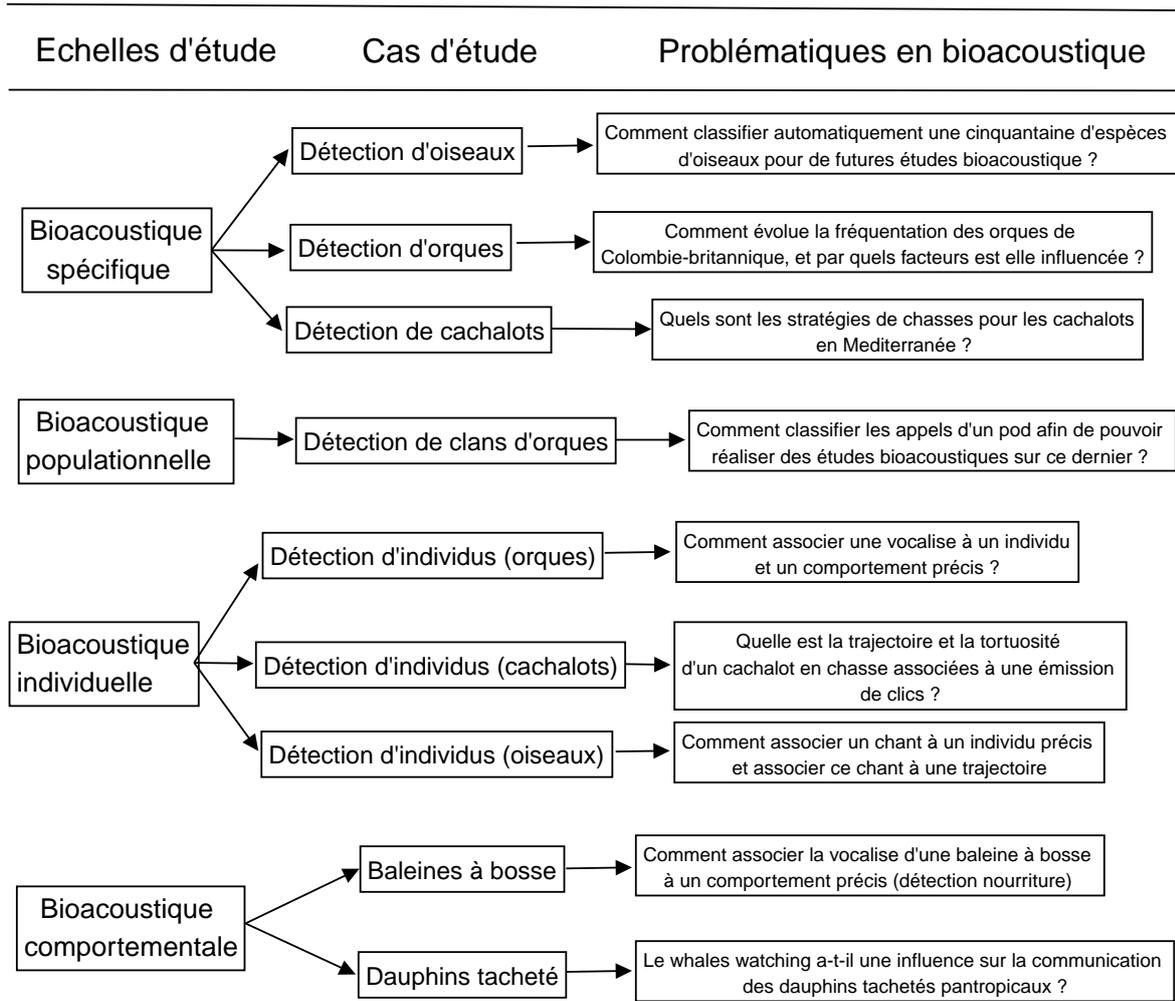


FIG. 2.27 – Résumé des différentes problématiques en bioacoustique associées à chaque échelle d'analyse.

Chapitre 3

Méthodes pour l'analyse des signaux

Avant de décrire les méthodes utilisées dans ce travail de thèse, un état de l'art est nécessaire sur l'utilisation de certaines méthodes en bioacoustique. Notamment les avantages et inconvénients de certaines approches seront développés et une comparaison entre les méthodes supervisées et non supervisées sera réalisée.

La première étape pour extraire des informations des enregistrements est la représentation du son. Il existe donc différentes représentations du signal. La première est la forme d'onde (oscillogramme), qui se base sur l'évolution de l'intensité sonore en fonction du temps.

Généralement, une échelle logarithmique pour mesurer les amplitudes sonores est utilisée (décibel).

Puis, il est possible d'afficher le son sous forme de représentation spectrale, c'est qui va permettre de visualiser la composition fréquentielle de l'enregistrement. Il y aura donc les fréquences en abscisse et l'amplitude en ordonnée. Cette représentation va nous permettre de connaître la répartition des fréquences "pures" qui composent ce son.

Puis, la troisième représentation utilisée par de nombreux bioacousticiens est la représentation temps-fréquence : le spectrogramme. L'objectif est de tracer la répartition énergétique du son en fonction du temps et des fréquences (Beecher, 1988).

En fonction de la tâche à réaliser certaines représentations vont être plus efficaces que d'autres. Une fois que le choix de la représentation du son est effectué, il faut extraire les signaux provenant des différentes espèces.

3.1 Etat de l'art des méthodes supervisées et non supervisées

Afin de comprendre nos choix des différentes approches pour les problématiques de détection et/ou classification d'espèces, de populations et d'individus, nous résumons les approches supervisées et non supervisées utilisées par la suite.

3.1.1 *Les méthodes supervisées*

Les méthodes supervisées consistent en la classification d'objets dans des classes avec des données d'entraînement connues a priori (LeCun et al., 2015). Donc l'utilisateur connaît les classes avant de créer son modèle. Les méthodes supervisées peuvent être divisées en deux catégories : la classification et la régression.

La **régression** consiste à un ensemble de méthodes d'apprentissages nous permettant de prédire une variable issue d'une fonction (linéaire ou non) basée sur une distribution. Par exemple, pour une régression linéaire, le principe est de placer une droite Y sur un nuage de point.

Cette droite va être exploitée pour effectuer des prédictions pour des caractéristiques encore non observées (Goodfellow et al., 2016).

Tandis que la **classification** fait référence à l'attribution d'une classe déjà définie à un objet.

Il existe différentes méthodes de classification supervisée :

- Les réseaux neurones
- Les K-NN (Méthodes des K plus proches voisins)
- Les arbres de décisions
- Les forêts d'arbres décisionnels
- Les machines à vecteurs de support
- Les classifications naïves bayésiennes

L'ensemble de ces méthodes ne sera pas décrit dans cet état de l'art. J'expliquerai les grands principes des réseaux neurones, qui sont utilisés pour la bioacoustique au niveau spécifique et populationnel.

Pour la classification supervisée avec des réseaux neurones, des exemples annotés servent à entraîner une fonction de prédiction. Dans notre cas, les exemples annotés peuvent être des sons d'oiseaux ou d'Orques. Ces exemples servent donc à entraîner un modèle afin qu'il prédise l'attribution d'une classe à un nouvel objet.

Les principales étapes de la classification supervisée sont illustrées en Fig.3.2. La première étape est donc de détenir des données labellisées.

L'objectif du réseau de neurones, ici à convolution, est d'attribuer une classe à chaque image présentée (ici un spectrogramme).

L'apprentissage considère par parties le spectrogramme en le divisant en petits champs réceptifs de quelques pixels.

Le CNN va donc comparer l'ensemble des spectrogrammes fragments par fragments avec des caractéristiques précises. Le modèle va extraire les caractéristiques de chaque fragment grâce à une succession de filtres créant de nouvelles images appelées cartes de convolutions. Chaque pixel de la caractéristique est multiplié par la valeur de ce même pixel dans l'image.

Ensuite, on additionne toutes ces réponses et on divise ce résultat par le nombre total de pixels dans le fragment (moyenne par fragment). Ce processus est répété pour l'ensemble du spectrogramme, et également pour l'ensemble des caractéristiques.

Ces convolutions sont répétées successivement de couche en couche pour obtenir un vecteur de caractéristiques appelé code CNN, donné en entrée de l'étage de classification composé de couches entièrement connectées ('perceptron multicouche' ou MLP). Ce dernier combine les caractéristiques du code CNN pour classer le spectrogramme. La sortie de ce modèle comporte un neurone par classe, qui estime la probabilité d'appartenance à une classe.

L'architecture du MLP est inspirée du cerveau. Son principe est que les neurones sont tous connectés entre eux. L'ensemble de ces neurones est réparti en trois familles : la première est la famille des entrées ('input'), puis la famille des couches cachées ('hidden layers'), enfin celle des sorties. Le nombre de couches de neurones est variable. A chaque neurone une fonction d'activation (ex : sigmoid, ReLU) transfère l'information d'un neurone aux autres. L'information (code CNN) est donnée à l'entrée du modèle, puis transmise de neurones en neurones à travers les couches, et aboutit à une prédiction d'appartenance par classe (par exemple une espèce). Les informations sont apprises par rétropropagation de l'erreur.

En fonction des entraînements, les poids de connections entre neurones sont adaptés par rétropropagation de l'erreur, afin de minimiser la fonction d'erreur entre le label et la prédiction proposée par le modèle.

L'architecture globale du modèle est montrée en figure 3.1.

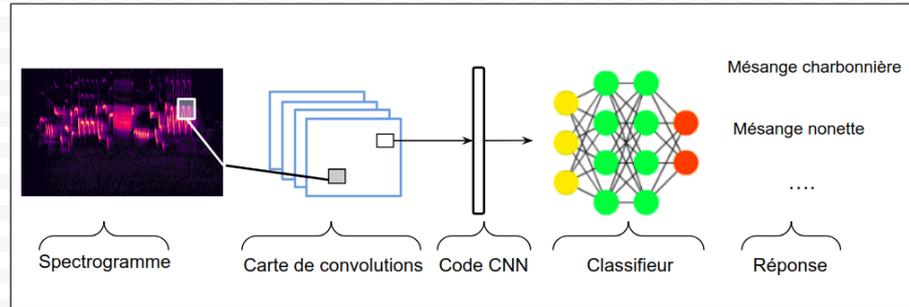


FIG. 3.1 – Résumé des étapes globales d'un CNN.

Le modèle est entraîné sur des données de référence et donnera une prédiction de présence d'espèce par fichier. Ces prédictions nous permettent d'établir des statistiques sur la présence/absence d'une espèce (ou population) sur une échelle temporelle importante.

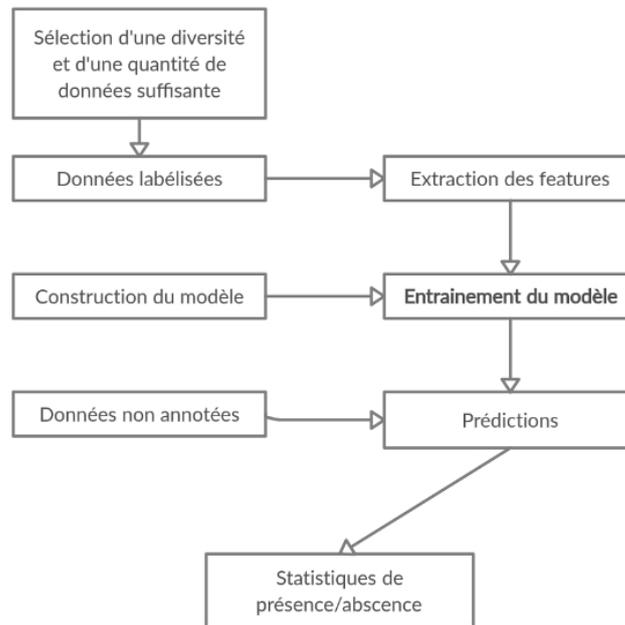


FIG. 3.2 – Résumé des étapes globales de la classification supervisée.

Une fois le modèle choisi, un long travail de sélection des hyper-paramètres va être nécessaire. Chaque modèle va présenter un certain nombre de paramètres qu'il va falloir ajuster avant et pendant l'entraînement.

Puis, une fois les hyperparamètres fixés et les données labellisées sélectionnées, différents entraînements sont nécessaires afin d'optimiser la qualité de notre modèle. Pour cela, il faut

diviser notre jeu de données en deux ensembles : en général, 80% des données sont destinées à l'entraînement du modèle, et 20% au test du modèle. Généralement, cette attribution ('split') se fait de manière aléatoire.

Il est également possible de réaliser une validation "croisée", c'est à dire de sélectionner plusieurs ensembles de validations, afin d'obtenir une estimation plus robuste (Hastie et al., 2009).

L'évaluation d'un modèle de classification se donne en taux de bonne classification (TBC) qui est le cardinal des échantillons bien classés sur le total des échantillons. C'est la somme de la diagonale de la matrice de confusion sur sa somme totale. La matrice de confusion permet aussi de finement analyser la probabilité qu'une classe soit confondue avec une autre.

La performance d'un détecteur est défini par sa sensibilité (probabilité de classer un segment positif si il est réellement positif), et sa spécificité (probabilité de classer un segment négatif, s'il est vraiment négatif) (voir Annexe F).

Une fois le modèle entraîné, il a un score sur ses données d'entraînement et sur ses données de test. Si les scores d'entraînement sont très bons, tandis qu'ils sont mauvais sur le test, c'est que le modèle est en sur-apprentissage. Globalement, cela signifie que le modèle a appris par coeur les données de l'entraînement, et ses prédictions sur les données de tests sont erronées. A l'inverse, le modèle peut ne pas apprendre les caractéristiques de chaque classe ('underfit'). L'un des avantages des méthodes supervisées comme le Deep Learning est la possibilité de faire du "transfer learning", consistant à utiliser les connaissances d'un premier entraînement pour une autre tâche de classification. Nous verrons dans cette thèse la mise en oeuvre d'une telle stratégie. Le modèle qui sera utilisé pour la détection automatique de vocalises d'Orques a été préalablement construit pour une tâche de classification d'oiseaux (Grill and Schlüter, 2017). L'architecture ainsi que les hyperparamètres du modèle ont été extraits pour la détection d'Orques.

Le principal inconvénient de la classification supervisée est la nécessité de beaucoup de données annotées pour son entraînement (Purwins et al., 2019).

3.1.2 Les méthodes non supervisées

Les méthodes de classification non supervisées, visent à regrouper les éléments à l'intérieur de classes (les plus homogènes possibles) mais sans connaître les classes à priori.

Trois approches peuvent être utilisées dans ce type de classification : soit les algorithmes sont basés sur des distances entre les éléments, soit sur des probabilités (densité), ou alors sur une approche hiérarchique. Concernant les mesures des distances entre les échantillons, on peut utiliser des distances barycentriques ou euclidiennes.

Pour l'approche barycentrique, l'algorithme K-means, prend en compte tous les points d'un cluster, et mesure la distance d'un point à la moyenne des points de son cluster. Puis il minimise la somme des carrés des distances (MacQueen et al., 1967).

La deuxième approche en terme de distance est le DBSCAN ou HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) traitant également la densité à l'intérieur des clusters. Les éléments appartiennent à un même cluster s'ils sont assez proches et si le nombre minimal d'éléments dans un cluster est respecté. Cet algorithme écarte tous les éléments qui ne sont pas dans des clusters (à l'inverse de K-means).

La dernière approche utilisée en terme de distance est la classification dite "hiérarchique" (Classification Ascendante Hiérarchique CAH). L'algorithme crée à partir des distances entre chaque paire d'individus un arbre de classification (dendrogramme). On l'appelle hiérarchique, car elle produit des classes de plus en plus grandes. Ce regroupement hiérarchique crée une hiérarchie entre les groupes, sous une forme arborescente. Tous les groupes sont interconnectés entre eux. Les groupes en haut de l'arbre sont très différents, tandis que les groupes en bas de l'arborescence sont similaires.

La deuxième approche de classification non supervisée est basée sur des probabilités. L'objectif de ces approches, est de chercher l'homogénéité à l'intérieur de certains clusters en respectant des modèles de lois probabilistes (Classification bayésienne (Ghahramani, 2015)).

Différentes bibliothèques sont disponibles en langage Python pour ces opérations de clustering, comme la bibliothèque Scikit-learn qui est très efficace (Pedregosa et al., 2011), (voir figure 3.3).

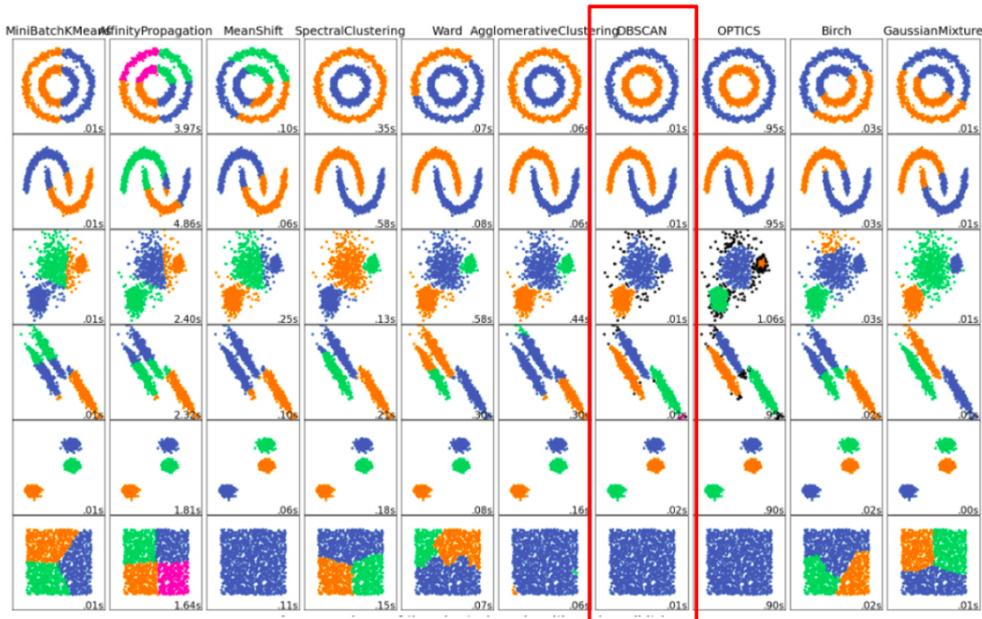


FIG. 3.3 – Comparaisons des principaux algorithmes de clustering proposés par la bibliothèque Scikit-learn (Halkidi et al., 2001). DBSCAN, en 7ème colonne, partitionne correctement pratiquement tous les ensembles.

Le dernier champ d'étude en classification non supervisée concerne la réduction de dimensionnalité. On présente ici deux méthodes classiques permettant de projeter des données issues d'un espace de grande dimension dans un espace de plus petite dimension (en général, deux ou trois dimensions).

Pour réduire la dimensionnalité d'un jeu de données, deux possibilités sont envisageables : soit enlever des dimensions (supprimer celles qui sont les moins informatives, expliquant le moins de variance de données), soit combiner les dimensions pour obtenir des dimensions plus pertinentes.

Parmi ces dernières, l'analyse en composante principale (ACP) est la plus connue. L'ACP crée un jeu de données artificiel avec un nombre de dimensions égal à celui de départ. Mais les

premières dimensions de ce nouveau jeu de données projetées concentre le maximum d'informations. C'est un algorithme de réduction de dimensionnalité linéaire. C'est à dire que l'algorithme ne préserve pas les distances locales des données à haute dimension.

En revanche, certaines méthodes de réduction de dimensionnalité sont non linéaires comme la t-distributed Stochastic Neighbor Embedding (t-SNE) (Van der Maaten and Hinton, 2012 ; Hinton and Roweis, 2002).

La t-SNE se base sur la similarité et la dissimilarité des données : il crée une probabilité de distribution sur des paires d'objets à haute dimension. Les objets (ici les formes de vocalises) qui se ressemblent le plus, auront une forte probabilité d'être rapprochés sur la représentation, tandis que des objets différents sont éloignés. La figure 3.4 représente la principale différence entre l'ACP et la t-SNE. Les deux points liés sont proches dans l'espace euclidien (en pointillé la distance euclidienne) mais éloignés si nous considérons le jeu de données entier (en trait plein la distance "en fonction données").

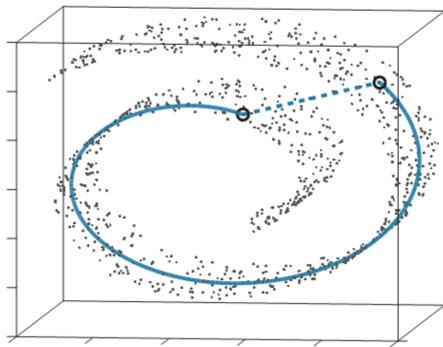


FIG. 3.4 – Représentation théorique des distances utilisées en ACP (tiret) et t-SNE (continu).

La dernière méthode de réduction de dimensionnalité est le Uniform Manifold Approximation (UMAP), qui est également une méthode non linéaire (McInnes et al., 2018).

Ces méthodes non supervisées sont présentées dans la figure 3.5.

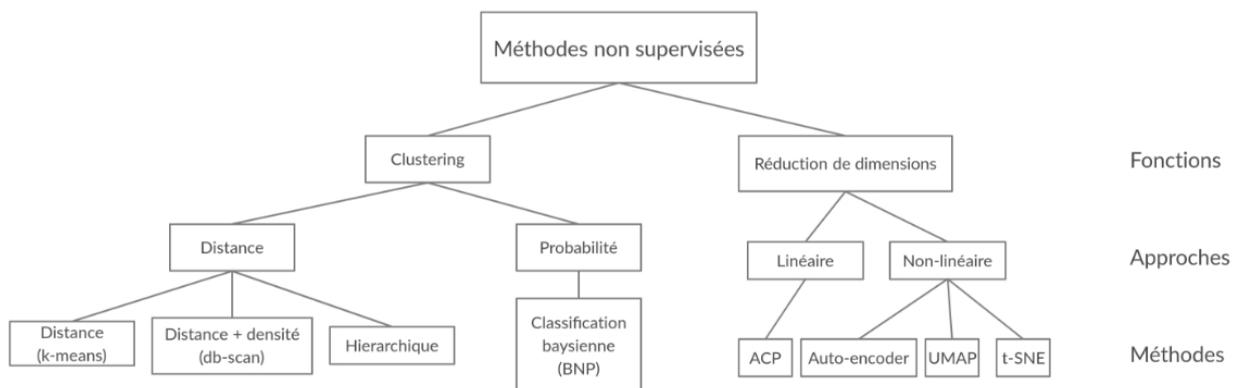


FIG. 3.5 – Résumé des méthodes non supervisées.

Pour l'ensemble des méthodes, le regroupement automatique doit produire des groupes avec :

- Une forte similarité intra-groupe (ressemblance à l'intérieur des groupes).
- Une faible similarité inter-groupe (dissemblance entre les groupes).

Dans cette thèse différentes méthodes supervisées et non supervisées ont été utilisées.

La sélection de ces méthodes s'est faite grâce à différents tests et à la bibliographie existante en bioacoustique.

Dans ce travail, les deux types de méthodes ont été utilisés (supervisée et non supervisée). Ces deux méthodes sont dites automatiques, la machine, après un apprentissage défini, est capable d'extraire des informations nécessaires (Machine Learning). En effet, l'automatisation des méthodes en bioacoustique est indispensable à cause des larges quantités de données à traiter (Au and Lammers, 2016). Le Machine learning en bioacoustique est de plus en plus présent et Bianco et al. (2019) insiste sur le fait d'intégrer des informations environnementales afin d'améliorer les performances des différents modèles.

Les parties suivantes vont décrire l'ensemble des méthodes utilisées pour les différentes échelles d'études : spécifique, populationnelle, individuelle et comportementale.

3.2 Méthode pour la détection d'espèces

Une fois le matériel choisi, les enregistrements effectués et stockés il faut analyser nos données. Cette étape primordiale peut être un véritable obstacle pour les scientifiques à cause de la grandeur du jeu de donnée.

En effet, les performances technologiques de stockages n'ont fait qu'augmenter ces dernières années ce qui représente un avantage considérable pour les sciences comme la bioacoustique. Le "Big data" est devenu très courant, et d'actualité. Il regroupe beaucoup de disciplines comme la sociologie, la médecine, l'économie mais également la bioacoustique (De Mauro et al., 2016 ; Koga, 2015 ; Kays et al., 2020).

Récemment, beaucoup de travaux ont été réalisés pour l'identification de grands groupes acoustiques, comme les poissons, les cétacés (Pourhomayoun et al., 2013 ; Dugan et al., 2015) ou encore le bruit des navires (Das et al., 2013). Ces différents articles utilisent des méthodes de "Machine learning" et plus particulièrement de "Deep Learning". Ces différents modèles sont relativement récents, et permettent d'éviter les annotations manuelles des enregistrements, et donc le traitement de grand volumes de données. On peut donc différencier les méthodes en deux grandes classes qui sont les méthodes automatiques et les méthodes manuelles. Ce chapitre va présenter différentes méthodes qui ont été utilisées pour la détection d'espèces. Ces méthodes seront qualifiées d'automatiques ou de semi-automatiques.

3.2.1 Méthode pour la détection d'espèces avifauniques

3.2.1.1 État de l'art

Le premier cas d'étude pour la détection d'espèces s'applique à la détection d'espèces aviaires. La détection et classification d'oiseaux est récemment devenue une tâche d'actualité notamment dans le développement de nouvelles méthodes d'apprentissage. Il y a plus de cinq ans Borker et al. (2014) ont montré que le coût de suivi automatique versus manuel est similaire du

fait du temps de mise au point des méthodes automatiques. De même Duan et al. (2013) montrent que les premiers détecteurs bioacoustiques (Raven ; Song Scope) nécessitent une phase d'optimisation par l'utilisateur, ce qui nuit à leur efficacité. En plus de cet investissement important, les classifications obtenues étaient peu performantes. Par exemple, concernant le logiciel Song Scope, Buxton and Jones (2012) montrent une sensibilité relativement faible de 65% de bonne reconnaissance pour seulement 3 espèces d'oiseaux.

Puis des compétitions internationales sur des tests publiés publiquement ont permis de progresser rapidement. La première compétition sur ce domaine a été lancée en 2013 au sein de la conférence ICML (International Machine Learning Society), puis NIPS (Neural Information Processing Systems) et ICML2014. Ces conférences de référence en machine learning, ont lancé les défis sur une trentaine d'espèces puis près de 90 par la suite (Glotin et al., 2013c,a ; Dufour et al., 2013). In fine les performances étaient de l'ordre de 80% de TBC mais les fichiers de tests avaient été sélectionnés pour leur clarté, donc à relativement fort Rapport Signal sur Bruit (RSB) .

Il fallait alors envisager des modèles testés sur des cas d'usage plus réels, c'est à dire à RSB moyen voir faible. Le challenge international annuel nommé 'BirdClef' lancé en 2014, a poursuivi cet effort. Il a rassemblé des dizaines de compétiteurs chaque année depuis 6 ans, travaillant sur des fichiers de test difficiles (enregistrements à RSB moyen ou plusieurs espèces par fichier test). Ces challenges ont permis de cerner un type de modèle de plus en plus performant, le CNN, avec un virage net en sa faveur en 2016 (Goëau et al., 2016). Les challenges qui ont suivi montrent alors une amélioration des performances (Joly et al., 2015, 2017 ; Sevilla and Glotin, 2017). En 2018 et 2019 les modèles sont évalués dans BirdClef sur plus d'un millier d'espèces avec des performances d'environ 72% de TBC, pour des tests à RSB non sélectionnés.

Cependant, un double constat peut être établi. D'une part, si on élargit trop le nombre de classes, les performances ne permettent pas une application réelle et utile sur le terrain pour l'ornithologue ou le citoyen, avec au moins 85% de TBC. D'autre part dans le cas où le paysage acoustique est trop complexe avec des recouvrement d'espèces, les performances chutent, comme le montre les scores sur les paysages du Pérou qui ont jusqu'à 2 espèces chantant simultanément sur chaque 5 secondes de test. Testés dans BirdClef 2018 et 2020, les scores sont d'environ 8% de TBC seulement (Kahl et al., 2020).

3.2.1.2 Le modèle de classification

Grâce à nos collaborations¹, nous avons eu accès aux corpus de la base de données de [Xeno-canto](#), qui est un site web dédié au partage de différents sons d'oiseaux du monde entier. J'ai donc sélectionné manuellement des dizaines d'enregistrements de Xeno-canto afin qu'il soient bien labellisés, et qu'ils ne soient pas trop bruités, ni que plusieurs espèces y chantent en même temps. Il est préférable que chaque classe soit représentée par un nombre d'enregistrements conséquents et égal entre les classes, car un déséquilibre entre les classes affecterait la généralisation du modèle Kahl et al. (2017).

Une fois que les espèces cibles ont été sélectionnées, le modèle de reconnaissance de ces espèces, que l'on nomme DynetBird, un réseau neuronal à convolution, fut appris. Ce modèle

1. Nous remercions Willem-Pier Vellinga et Alexis Joly pour leur collaboration.

fut construit par un ingénieur au sein de l'équipe DYNI. L'architecture de ce modèle CNN est inspirée de Kahl et al. (2017), présentée en figure 3.6, et précisée dans notre rapport de recherche (Poupard et al., 2018b). J'ai effectué les entraînements en ajustant les données et les paramètres du modèle.

Model 2
9 weighted Layers
~5500s per Epoch
Conv1, 32x7x7, Stride 1
MaxPooling, Size 2
Conv2, 32x5x5, Stride 1
MaxPooling, Size 2
Conv3, 64x5x5, Stride 1
MaxPooling, Size 2
Conv4, 128x3x3, Stride 1
MaxPooling, Size 2
Conv5, 512x3x3, Stride 1
MaxPooling, Size 2
Conv6, 1024x3x3, Stride 1
MaxPooling, Size 2
DenseLayer, 512 Units
Dropout, p=0,5
DenseLayer, 512 Units
Dropout, p=0,5
DenseLayer, 1500 Units
Softmax Output

FIG. 3.6 – Architecture du modèle utilisé. Image tirée de Kahl et al. (2017).

Les étapes de construction de notre modèle sont :

- Extraction des spectrogrammes de tous les enregistrements (512x256 pixels)
- Sélection des spectrogrammes contenant des sons d'oiseaux (Sprengel et al., 2016 ; Lasseck, 2013)
- Augmentation des données (vertical roll : décalage en fréquence des signaux ; Gaussian Noise : ajout de bruit d'intensité aléatoire ; Noise sample : ajout de bruit issu d'échantillons vides en signaux d'oiseau)
- Sélection de la meilleure architecture du modèle
- Entraînement du modèle et sélection de la meilleure architecture.

Différents entraînements ont été réalisés. Leurs optimisations se faisaient sur le score de la matrice de confusion. Certaines espèces étaient difficiles à classer notamment car elles imitent d'autres espèces.

Les différents entraînements de notre modèle sont décrits dans le tableau 3.2. Le jeu de données initial comportait 35 espèces (appelées classes). Les enregistrements utilisés pour l'entraînement et le test provenaient de deux sources : la base de données web de Xeno-Canto, et le jeu de données du challenge bioacoustique ICML 2013 (Glotin et al., 2013a).

Mon travail a donc été de sélectionner précisément les données sur le site de Xeno-Canto, afin que les enregistrements soient de bonne qualité, et qu'il n'y ai pas plusieurs espèces au sein de ces derniers. Le tableau 3.1 décrit l'ensemble du jeu de données utilisé.

L'entraînement 1 était composé de 15 min par espèce, pour les 35 espèces. Mais le nombre de

fichier par classe était très variable : pour quelques classes il n'y avait qu'un seul fichier (de 15 min) pour l'entraînement.

Donc le modèle était bien entraîné uniquement sur quelques individus. Il était donc difficile de bien généraliser sur la classe. J'ai donc sélectionné manuellement sur le site de Xeno-Canto un autre jeu de données pour compléter l'entraînement 1. L'entraînement 2 était donc composé au moins de 10 enregistrements différents, et ce pour les 35 classes 3.1. Le dernier entraînement (3) était composé de l'entraînement 2 avec une sélection d'enregistrements de 15 nouvelles espèces (au moins 20 enregistrements différents/classe).

Concernant les différents tests, le premier provenait du Challenge de classification d'oiseaux de la conférence ICML 2013 (Glotin et al., 2013a) et était composé de 30 secondes d'enregistrements par espèce. Puis j'ai sélectionné le Test 1 et 2 sur le site de Xeno-Canto (différents des entraînements).

TAB. 3.1 – Tableau représentant les différents jeux de données utilisés pour la classification aviaire.

Jeux de données	Source	Description
Entraînement 1	Xeno-Canto	15 minutes d'enregistrement/espèce pour 35 espèces
Entraînement 2	Xeno-Canto	Sélection d'au moins 10 enregistrements différents/espèces, pour 35 espèces
Entraînement 3	Xeno-Canto	50 espèces : Entraînement 2+ sélection d'au moins 20 enreg/esp pour les 15 nouvelles
Test 1	Train data for ICML 2013 - ML4B - Machine Learning for Bioacoustics	30 secondes d'enreg/esp pour 35 espèces
Test 2	Xeno-Canto	Sélection de sons récents de Xeno canto de 8 à 10sec pour les 15 nouvelles
Test 3	Xeno-Canto	Sélection de sons récents de Xeno-Canto, pour les 52 espèces fois 2, de 8 à 10sec

Différents entraînements ont été réalisés, toujours avec la même architecture de modèle, mais en faisait varier les données d'entraînements et les données des tests. Ils sont décrits dans le tableau 3.2 ainsi que leurs résultats.

TAB. 3.2 – Tableau représentant les différents runs effectués pour la classification aviaire avec leur Taux de Bonne Classification (TBC).

Entraînements	Jeux de données Entraînement	Jeux de données Test	Nbr de classes	Nbr de fichiers Entraînement	Nbr de fichiers/classe	TBC
1	Entraînement 1	Test 1	35	1253	1 à 129	54%
2	Entraînement 1	Test 2	35	1605	16 à 129	71%
3	Entraînement 1+ Test 1	Test 2	35	1640	17 à 130	77%
4	Entraînement 2	Test 2	35	937	16 à 40	86%
5	Entraînement 3	Test 3	52	1286	16 à 40	88%

Au total, 5 entraînements ont été effectués. On peut voir que le TBC augmente au fur et à mesure de ces derniers. Il passe de 54% de TBC pour le premier entraînement (avec 35 espèces) à 88% de TBC sur le dernier (5).

Le test final (3) était composé d'1h40 au total d'enregistrements (104 échantillons).

Notre modèle a donc été testé sur 52 espèces, et possède un TBC de 88% (= somme de la diagonale / nombre de tests). La matrice de confusion Fig.3.7 résume les erreurs commises du modèle sur ce test. Le jeu de données test était composé de 2 enregistrements tests par classe (de 5 à 20sec d'enregistrements).

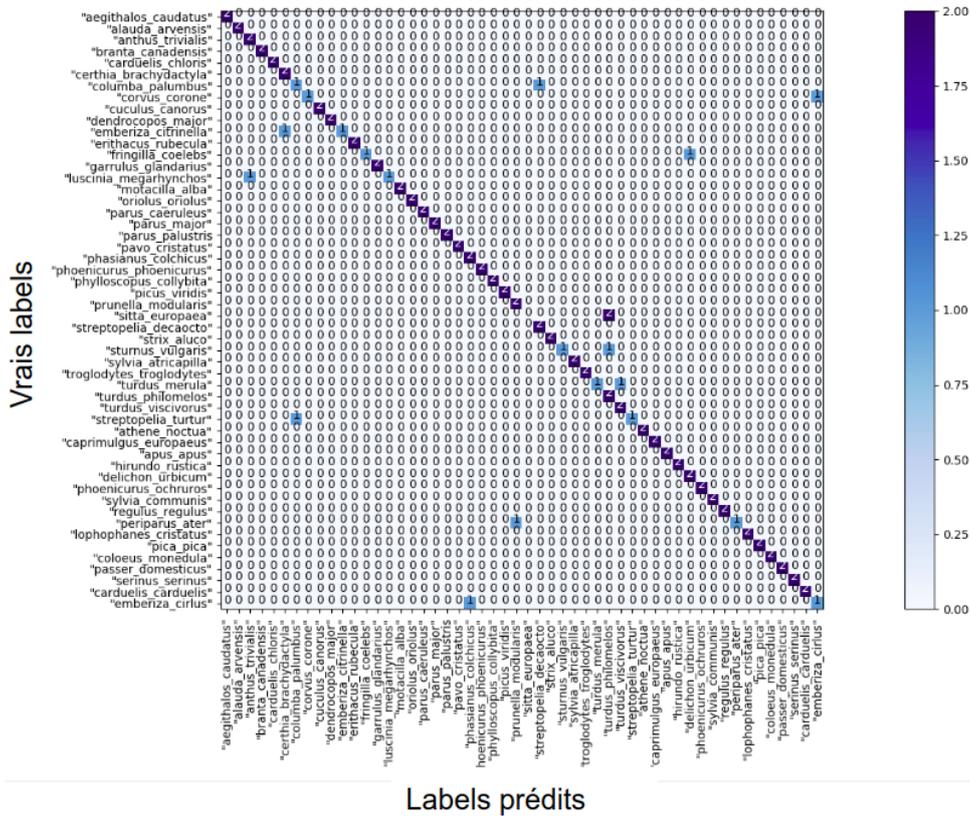


FIG. 3.7 – Matrice de confusion du test set pour la tâche de classification de 52 espèces d’oiseaux, avec un TBC de 88%.

Pour ce dernier test, 41 espèces ont été trouvées dans les 2 cas, 10 une fois sur 2, et une n’a pas été trouvée.

Cette matrice de confusion mesure finement la qualité de notre modèle. Par exemple la première espèce *Aegithalos Caudatus* (Mésange à longue queue) est toujours bien classée, tandis que l’espèce *Collumba Palumbus* (Pigeon ramier) est une fois bien classifiée et une fois confondu avec *Streptopelia decaocto* (Tourterelle turque).

Le modèle a donc été validé et nous l’avons testé dans de vraies conditions d’enregistrements, partie présentée dans le chapitre "Résultat" (partie 4.1.1).

3.2.2 Méthode pour la détection d’Orques (*Orcinus orca*) à Orcalab

Comme énoncé dans l’introduction, une des principales fonctions de la bioacoustique est la détection d’espèces au sein de leur environnement. Cette partie présente la méthode utilisée pour la détection d’Orques à la station de recherche Orcalab (Matériel présenté dans la partie 2.2.2.1).

Au total, plus de 50 To de son ont été stockés sur les serveurs de l’équipe DYNI.

Il est donc impossible de mettre en place des méthodes manuelles pour la détection des vocalises.

Depuis plusieurs années, beaucoup de scientifiques se penchent sur la tâche de détection automatique d’espèces sous-marines au sein d’enregistrements acoustiques. Beaucoup d’approches

ont utilisé les réseaux neurones ('deep learning') pour répondre à certaines problématiques (Bermant et al., 2019 ; Luo et al., 2019 ; Liu et al., 2018).

Pour cette tâche de détection d'Orques, nous avons également utilisé une méthode supervisée avec un modèle de deep learning.

Dans notre cas de détection, le modèle est entraîné uniquement sur deux classes : Orques ou bruit. Pour cela, il est nécessaire d'obtenir des données d'entraînement afin d'apprendre au modèle la reconnaissance de ces deux classes. L'acquisition de données annotées peut être un véritable obstacle pour cette discipline, c'est pour cela qu'un "extracteur automatique d'évènements" a été conçu pour ce protocole.

Les étapes de cet algorithme sont :

- Calcul du spectrogramme avec une fenêtre de Tukey de 1024 échantillons.
- Binarisation du spectrogramme, en comparant avec la médiane sur la bande de fréquences et la bande de temps : si la valeur d'un pixel est supérieure à 3 fois la médiane plus 3 fois l'écart-type de sa ligne, et plus de 3 fois la médiane plus 3 fois l'écart type de sa colonne, il est mis à 1, sinon à 0.
- Appliquer des filtres (closing and dilatation filter) pour éliminer le bruit.
- Trouver les composants connectés et supprimer les petites composantes et les pixels isolés.
- Regrouper les clusters restants (moins de 0.2 sec) et supprimer les clusters verticaux (bruits).

Toutes ces étapes nous ont permis de sélectionner un grand nombre d'évènements acoustiques. Cet algorithme a permis de détecter plus de 3500 évènements, dont 872 ont été attribués à des Orques après vérification manuelle.

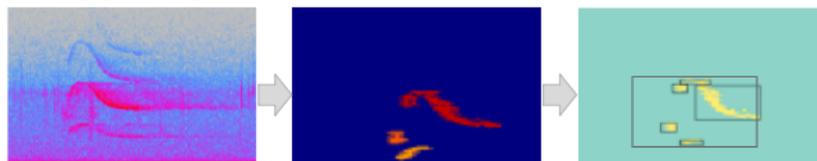


FIG. 3.8 – Principales étapes de l'extraction des évènements acoustiques : Binarisation et détection des composants connectés. Le spectrogramme montre des fréquences de 0 à 6,5 kHz pendant 2,5 sec.

Grâce à cet algorithme, nous avons pu construire un ensemble de données composé de 872 exemples de vocalises et de 6801 échantillons de bruits (bateaux, pluie etc..). Nous avons réparti au hasard ces données avec 60% pour le training du modèle, 20% pour sa validation et 20% pour le test.

Puis, nous avons entraîné un CNN (Grill and Schlüter, 2017) conçu à l'origine pour une tâche de détection d'oiseaux. Dans l'article de Grill and Schlüter (2017), deux modèles ont été testés, nous avons choisi de sélectionner le modèle "Sparrow". Ce modèle avait de très bonnes performances pour la détection de chants d'oiseaux, qui sont des espèces possédant les mêmes caractéristiques fréquentielles et temporelles que les vocalises d'Orques. Avant de décrire rapidement l'architecture du modèle, certaines étapes ont été effectuées :

- Calcul du spectrogramme avec une taille de fenêtre de 1024 échantillons avec une F_e de

22050 Hz.

- Banque de filtre à l'échelle mel (10 kHz) : découpage du spectrogramme en bande de fréquence de hauteur variant en échelle log.
- Calcul des Log de l'énergie pour chaque bande (dB).

Ces paramètres sont envoyés à l'entrée du modèle.

La topologie du modèle de Grill and Schlüter (2017) a donc été utilisée mais également les paramètres d'apprentissages. C'est ce qu'on appelle l'apprentissage par transfert (transfer learning) (Pan and Yang, 2009). Le modèle a donc été optimisé pour une tâche de classification d'oiseaux, mais son architecture et ses paramètres sont transférés pour la tâche de détection d'Orque.

Le modèle est présenté en figure 3.9 et résume son architecture globale, composé de ses 8 couches convolutionnelles et 2 périodes de pooling.

Le résultat du modèle présente une prédiction locale de la présence ou non d'Orques dans l'enregistrement.

NETWORK ARCHITECTURE OF
sparrow SUBMISSION

Input	1×701×80
Conv(3×3)	32×699×78
Conv(3×3)	32×697×76
Pool(3×3)	32×232×25
Conv(3×3)	32×230×23
Conv(3×3)	32×228×21
Conv(3×19)	64×226×3
Pool(3×3)	64×75×1
Conv(9×1)	256×67×1
Conv(1×1)	64×67×1
Conv(1×1)	1×67×1
GlobalMax	1

FIG. 3.9 – Image tirée de Grill and Schlüter (2017) représentant l'architecture du modèle utilisé pour la détection d'Orques.

Les performances du modèle sont présentées dans le tableau 3.3. Le taux de bonne classification pour l'entraînement était de 97%, pour la validation 96% et pour le test de 97%. Ces scores étant plutôt encourageants, nous avons décidé de sélectionner cette architecture pour la tâche de détection d'Orques.

TAB. 3.3 – Performances du modèle d'apprentissage pour la détection d'Orques sur le test.

	TBC (%)	Aire sous la courbe (AUC)
Training	97	0.88
Validation	96	0.89
Test	97	0.89

La fonction d'efficacité du détecteur est montrée en figure 3.10, avec en ordonnée le taux de vrais positifs et en abscisse le taux de faux positifs. Donc la spécificité est traduite en abscisse

et la sensibilité en ordonnée. Un seuil de 0.9 a été choisi afin de maximiser le taux de vrais positifs et de minimiser le taux de faux positifs.

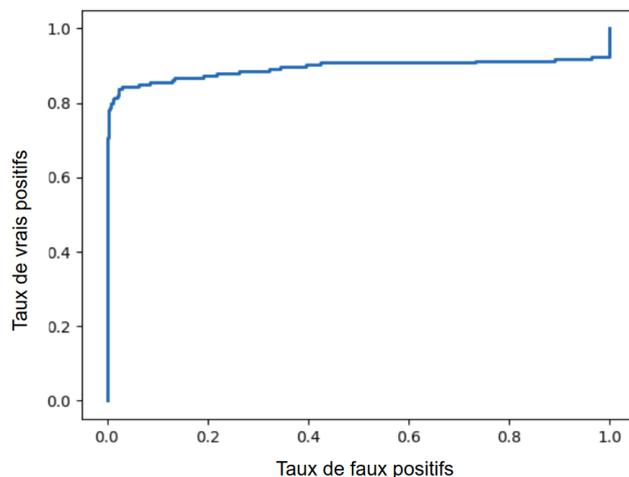


FIG. 3.10 – Représentation de la courbe "ROC" (Receiver Operating Characteristic) du détecteur acoustique d'Orques.

Une fois le modèle sélectionné, il a été lancé sur plusieurs années d'enregistrements. Les résultats sont présentés dans la partie 4.1.2.

3.2.3 Méthode pour la détection de Cachalots (*Physeter macrocephalus*) sur la bouée BOMBYX

Le dernier cas d'étude pour la détection d'espèces a été mis en place avec la bouée Bombyx pour la détection de Cachalots.

La bouée BOMBYX posée au large de Port-Cros en Méditerranée a été présentée dans la partie Matériel. Au total, près de 100 jours ont été enregistrés, ce qui représente 2,3 téraoctets de données. Il paraît donc impossible de mettre en place des méthodes manuelles pour la détection d'espèces (dans ce cas précis, le Cachalot).

Un détecteur de clics a été mis en place pour cette étude. Les étapes de ce détecteur sont décrites ci dessous.

Nous avons corrélé le signal avec une période d'un sinus de 12.5 kHz. Puis nous avons appliqué un filtre Teager-Kaiser (Kandia and Stylianou, 2006).

Une fois le filtre réalisé, les maximums locaux ont été récupérés dans une fenêtre de 20 ms (double de l'intervalle entre les pulses d'un clic de 10ms (Abeille et al., 2014)). Enfin, la dernière étape est de filtrer les bruits restants, en récupérant uniquement les maximums au dessus de trois écarts-types (Pukelsheim, 1994) de l'énergie logarithmique médiane de toutes les détections. L'ensemble de ces étapes est illustré figure 3.11.

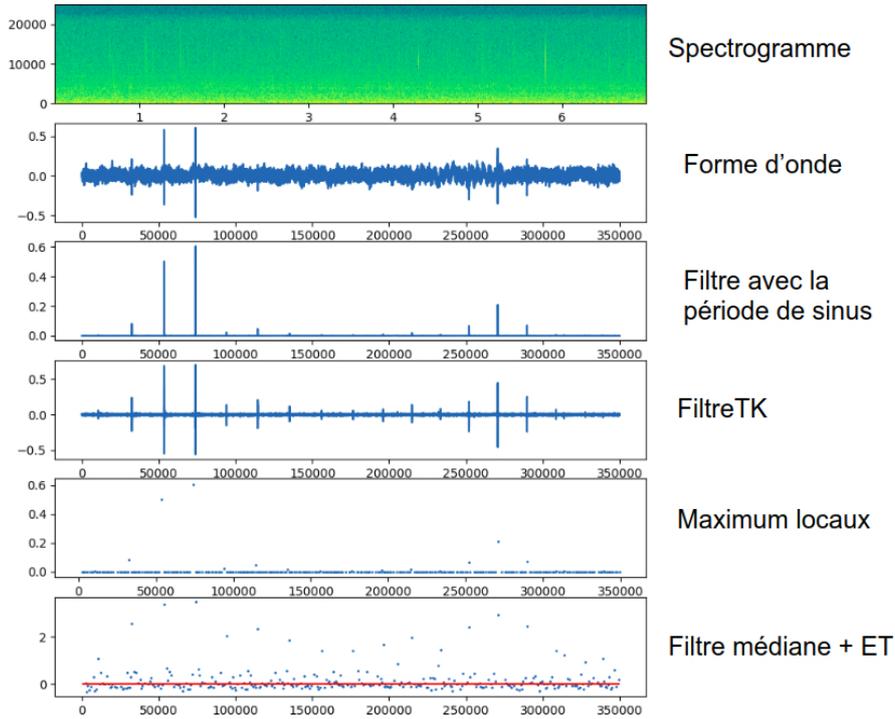


FIG. 3.11 – Représentation des différents filtrages réalisés pour détecter les clics de Cachalot pour un enregistrement de 7 secondes.

Le détecteur décrit ci-dessus ne détectait pas uniquement les clics de Cachalots, mais également d'autres clics d'odontocètes et des navires. En effet, la plupart des clics détectés ne proviennent pas des Cachalots mais du bruit ambiant.

La qualité d'un détecteur est décrite par sa **précision** et son **rappel** (notions décrites en Annexe F).

Notre détecteur va donc produire une certaine quantité de "vrais positifs" correspondant à de véritables clics de Cachalots, mais également une quantité de "faux positifs" qui ne sont pas des clics de Cachalots. L'objectif de ce détecteur était d'obtenir un taux de rappel très haut et une précision faible, afin d'être sûr de ne pas manquer des détections. Nous avons donc beaucoup de "faux positifs" dans notre jeu de détections.

Afin de sélectionner uniquement les clics de Cachalots, les TDoAs sur les deux hydrophones ont été calculés. La figure 3.12 montre un exemple de TDoA d'un clic de Cachalot visualisé sous Audacity. Le clic arrive d'abord sur la voie 1 (H1), puis sur la voie 2, donc l'animal vient de l'Est (voir le positionnement de la bouée dans le chapitre Matériel).

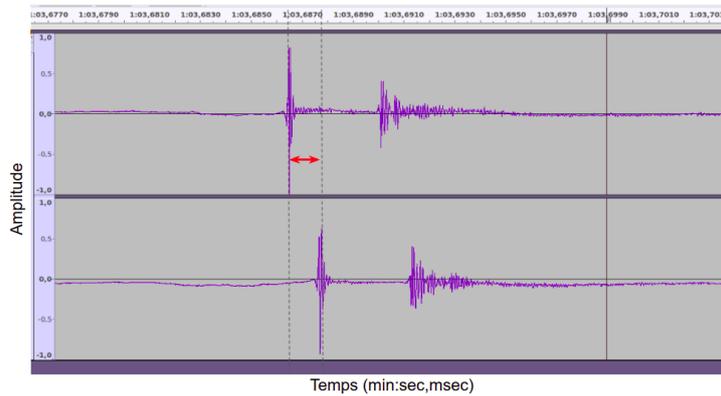


FIG. 3.12 – Exemple d'un clic visualisé sous Audacity, sur les deux voies de BOMBYX. La double flèche rouge représente le TDoA entre les deux voies.

Une interface d'affichage des temps de délais d'arrivée des clics détectés permet alors d'identifier manuellement les trajectoires de Cachalots. Cet outil d'annotation amélioré a été mis en place par l'équipe DYNI (Paul Best) afin de faciliter ces analyses sur une masse de données importante. Il est représenté en figure 3.13. Cet outil affiche l'ensemble des sessions mises bout à bout avec les TDoAs de chacune d'entre elles. Afin de vérifier l'exactitude des pistes, j'avais la possibilité d'écouter les clics à l'intérieur de celles-ci (par sélection manuelle). Un point rouge sur la trajectoire représente un clic qui vient d'être écouté.

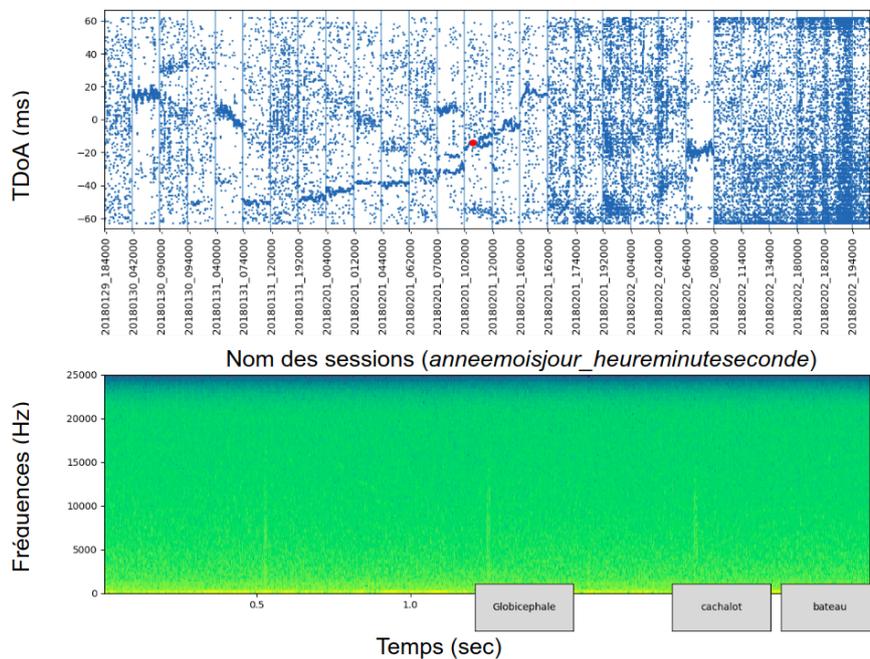


FIG. 3.13 – Représentation de l'outil d'annotation réalisé par l'équipe DYNI. Haut : affichage de plusieurs heures de TDoAs. Bas : Affichage d'un clic en spectrogramme (point rouge) représentant un clic qui vient d'être écouté.

Après avoir annotées "positives" ou non chaque session à la présence de Cachalots, j'ai pu

déterminer le sens de passage et le nombre d'animaux.

Différentes sessions ont été réalisées avec Bombyx. Ces dernières étaient enregistrées à différents recouvrements temporels (ex : 1min d'enregistrement, 5 min de pause), de 16% à 100%. Afin d'évaluer la perte de l'information qu'il pourrait y avoir pour la visualisation de la piste TDoAs, j'ai isolé une trajectoire d'un passage de Cachalot enregistré en continu que j'ai volontairement décimé.

En effet, la dernière session enregistrée sur la bouée BOMBYX a été faite du 26 Novembre 2018 au 26 Décembre 2018. Cette dernière session a été définie comme la session "contrôle" puisque les enregistrements ont été réalisés en continu. L'équipe DYNI a décidé de mettre en place un échantillonnage continu afin de pouvoir comparer les pistes complètes avec les pistes obtenues sur les autres sessions. La figure 3.14 illustre la trajectoire obtenue d'un individu lors de la session en continu. Il a été possible de confirmer qu'il n'y avait qu'un seul individu grâce à l'écoute des intervalles inter-clics (ICI).

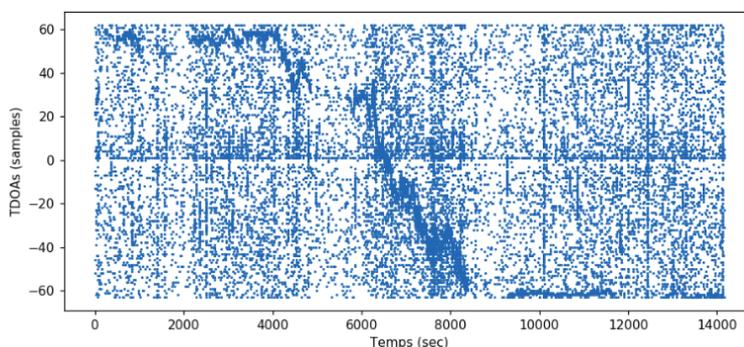


FIG. 3.14 – Affichage de la piste en TDoAs du Cachalot le 25 Décembre 2018.

La continuité de l'enregistrement nous permet aisément d'identifier l'individu et également son sens de déplacement.

Afin d'évaluer la qualité de l'échantillonnage des autres sessions, cet exemple a été utilisé en tant que trajectoire "contrôle". Nous pouvons grâce à cet échantillon continu tester nos hypothèses de décimations, ceci afin d'évaluer la perte d'information en fonction de l'échantillonnage. Différentes stratégies d'échantillonnages ont été réalisées, j'ai donc décimé cette trajectoire avec deux stratégies différentes : 5 min d'enregistrement/15 min de pause, et 1 min/ 5 min de pause. Le résultat de cette décimation est présenté en figure 3.15.

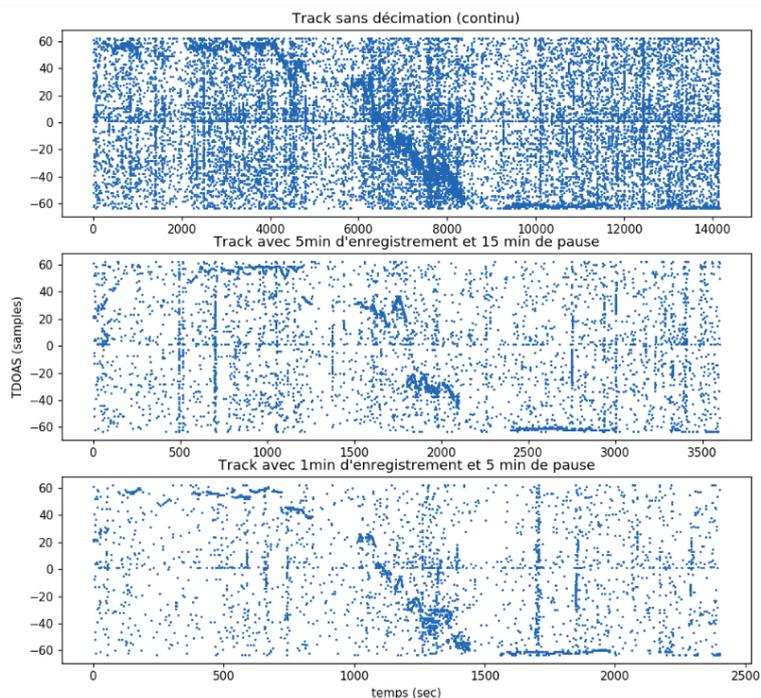


FIG. 3.15 – Exemple de la piste TDoAs d'un Cachalot, avec en haut la piste sur les 4h (continu).
Puis la piste décimée et concaténée à 25%, puis à 16%.

La décimation nous permet de conserver la direction principale de l'animal (60 vers -60 échantillons). Cependant, il paraît difficile d'estimer précisément les temps de plongée et de repos du Cachalot.

Cette étape m'a donc permis de valider la méthode d'extraction des pistes issues du calcul des TDoAs. L'ensemble des étapes est présenté en figure 3.16. Une fois que les sessions ont été annotées positives à la présence de Cachalots, des statistiques descriptives ont été réalisées (voir la partie résultats) et le sens de déplacement des animaux a été également extrait.

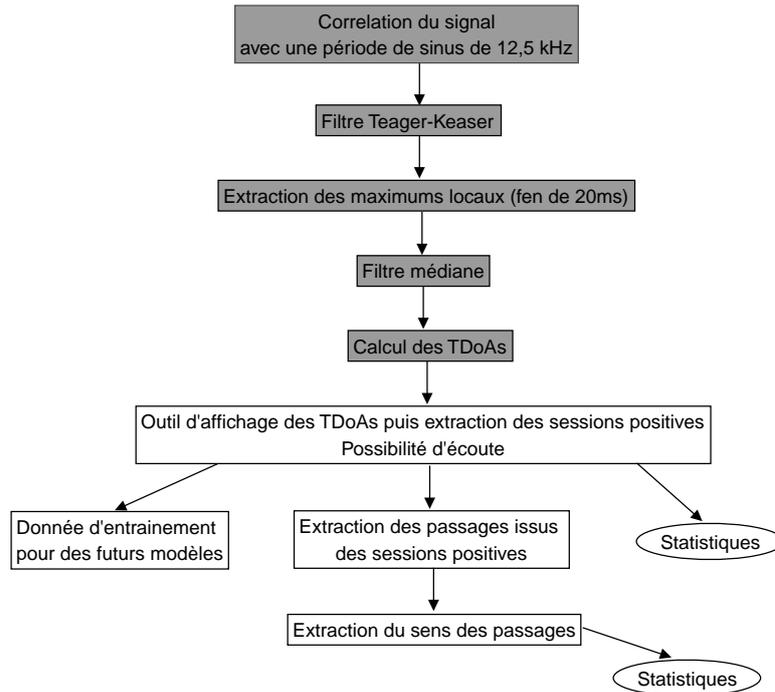


FIG. 3.16 – Diagramme représentant les étapes réalisées pour la détection de Cachalots sur la bouée BOMBYX. L’expertise manuelle pour l’extraction des pistes de Cachalots s’est faite via l’écoute du son, l’analyse de la forme d’onde et de la visualisation du spectrogramme du clic.

Mes annotations manuelles ont été utilisées par les membres de l’équipe DYNI, afin d’entraîner des modèles de Deep learning (CNN) pour détecter de manière totalement automatique la détection Cachalots, notamment pour le projet Européen Feder Marittimo GIAS de bouée anti-collision entre la mégafaune et le trafic maritime.

L’ensemble des méthodes pour la détection d’espèces a été présenté dans cette partie (méthodes pour la détection d’espèces avifauniques, pour les Orques et les Cachalots).

3.3 Méthode pour la détection de population

La détection de population a été appliquée à un seul cas d’étude qui est la détection de pods d’Orques en Colombie-Britannique (Canada).

3.3.1 Méthode pour la détection de clans acoustiques chez les Orques de Colombie-Britannique

Comme décrite dans l’introduction, l’analyse à l’échelle des populations nécessite la détection et l’identification de dialectes particuliers émis lors d’une communication acoustique. Cette partie se concentrera sur les Orques résidentes de Colombie-Britannique. La partie 2.1.3 a décrit cette population particulièrement hiérarchisée.

La figure suivante rappelle l’organisation de cette communauté.

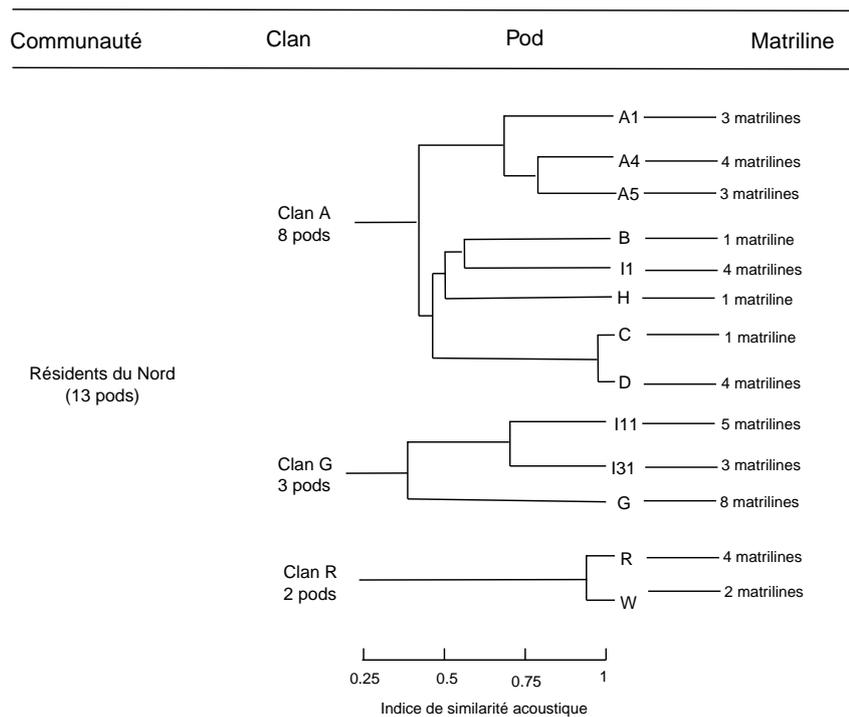


FIG. 3.17 – Diagramme représentant l’organisation hiérarchique de la communauté des résidents du Nord. Figure inspirée de Ford (1991).

Les Orques résidentes du Nord sont composées de 13 pods, divisés en 3 clans différents : le clan A, le clan G et le clan R. Le clan A est composé de 8 pods regroupés en 21 matriline, le clan G en 3 pods regroupés en 16 matriline et enfin le clan R rassemblé en 2 pods et 6 matriline.

Les clans ont été défini par Ford (1989) et ils sont identifiés grâce à la similarité acoustique des vocalises. Cet indice est basé sur le coefficient d’association de Dice (Dice, 1945).

Différents index de similarité peuvent être pris en compte pour comparer les signaux émis. Par exemple, McGregor and Krebs (1982) a utilisé un indice particulier pour comparer les différentes populations de Mésanges charbonnière (*Parus Major*), ou encore Hedwig et al. (2014) a également utilisé l’indice de Dice pour comparer les signaux de deux populations de Gorilles de l’Ouest (*Gorilla gorilla*).

Le coefficient de Dice (ou indice de similarité) est couramment utilisé en écologie pour comparer deux populations.

Ford (1989) à utilisé l’indice suivant :

$$\text{Index de similarité} = \frac{2(N_c + N_s)}{R_1 + R_2} \quad (3.1)$$

Avec N_c le nombre de type d’appels partagés entre deux groupes, N_s , le nombre de subtypes partagés, R_1 la taille du premier répertoire et N_2 la taille du deuxième répertoire. Grâce à cet indice, Ford (1989) a pu hiérarchiser acoustiquement les différents clans, pods et matriline.

Il nous paraît donc judicieux de débiter la tâche de classification par l’identification des clans A G et R. Ces trois clans peuvent produire un certain nombre de vocalises.

- Le clan A : 19 types de vocalises, 37 subtypes

- Le clan G : 15 types de vocalises : 16 subtypes
- Le clan R : 8 types de vocalises

Différentes méthodes supervisées et non supervisées ont été testées pour classifier les clans. Ces méthodes sont décrites dans les parties suivantes.

3.3.1.1 Méthode supervisée

Différents entraînements via un CNN ont été testés afin de différencier les 3 clans.

Nous avons utilisé la même architecture de modèle que pour la détection d'espèces (Orques). L'objectif était donc de classifier automatiquement le clan A, le clan G et le bruit. Nous voulions également classer le clan R mais nous ne disposions pas assez de d'enregistrements pour cette classe. Les transients sont très discrets et ne produisent que très rarement des vocalises, il est donc très complexe d'obtenir beaucoup d'exemples.

Le premier test qui a été réalisé est l'utilisation de l'ensemble des types d'appels pour la classification des clans. J'ai donc regroupé des vocalises appartenant a chaque clan afin d'entraîner le modèle.

La fréquentation des différents clans à Orcalab n'est malheureusement pas équilibrée. Les clans clan A et G fréquentent beaucoup les eaux de Hanson Island tandis qu'il est relativement rare de rencontrer le clan R. Il a donc été impossible de sélectionner assez de vocalises pour ce dernier.

Trois classes étaient donc utilisées : le clan A, le clan G et le bruit.

La première stratégie a été d'utiliser l'ensemble des appels pour la classification. Puis d'utiliser uniquement certains types d'appels produits uniquement par certains clans. Pour entraîner les modèles différents jeux de données ont été extraits :

- **Jeu de données 0 (JDo)** : données de bruit extraites du détecteur d'Orques (détection d'espèces). 5695 exemples de 5 secondes
- **Jeu de données 1** : Annotations faites par les experts d'Orcalab. Annotation de l'été 2017.
Classe 1 (clan A) = 108 passages (1855 enregistrements).
Classe 2 (clan G) : 25 passages (624 enregistrements)
- **Jeu de données 2** : Passages des clan G et A de l'été 2019.
7 passages du clan G (74 enregistrements)
5 passages du clan A (62 enregistrements)
3 passages avec le clan A et G (17 enregistrements)
- **Jeu de données 3** : Base [Orchive](#)
- **Jeu de données 4** : Extraction de deux types d'appels
1250 N04s pour le clan A (issue des jeux de données 1, et 3)
160 N23 pour le clan G (issue des jeux de données 1, et 3)
384 autres appels (issue des jeux de données 1, et 3)

Différents entraînements ont donc été réalisés avec différents jeux de données.

La première stratégie était d'utiliser l'ensemble des types de vocalises. Deux tests ont été mis en place pour chaque entraînement : le premier test pour l'obtention des scores de classification (TBC), et le deuxième test était réalisé sur plusieurs passages entier (en 2019) afin de visualiser l'évolution des probabilités de classification le long du passage

des animaux devant l'hydrophone.

Les différents entraînements sont présentés dans le tableau 3.4.

Le premier entraînement a été effectué sur les jeux de données 0 et 1, avec 3 classes (A, G et bruit). Le TBC était de 64 % sur le test. Nous n'avons donc pas fait de deuxième test et réalisé directement un autre entraînement. L'objectif était d'arriver à 90 % sur les deux tests.

Pour le deuxième entraînement, nous avons d'abord lancé le détecteur d'Orques sur le jeu de données 1. Nous avons récupéré les segments contenant les hautes probabilités de présence d'Orques.

Puis le modèle a été entraîné uniquement sur ces segments. Nous avons donc pu enlever la classe "Bruit". Le TBC était relativement élevé (90 %), j'ai donc lancé ce modèle sur différents passages d'Orques afin de valider ce score. Les résultats de cette manipulation sont présentés dans le chapitre Résultat (partie 4.2.1).

TAB. 3.4 – Récapitulatif des différents entraînements réalisés pour la classification des clans.

	Entraînement	Test 1	Test 2	Classe de bruit	Entraînement/ Test	Remarque	TBC (%)
1	JD0 JD1	JD0 JD1		JD0	80% train 20% test	Tous les types d'appels Sans détecteur préalable	64% sur test 1
2	JD1	JD1	JD2		80% train 20% test	Détecteur Puis classifieur Pas de classe de bruit	90% sur test 1 Sur-apprentissage
3	JD1 JD3	JD1 JD3	JD2		80% train 20% test	Détecteur Puis classifieur Pas de classe de bruit	68% sur test 1
4	JD4	JD4	JD2		80% train 20% test	Détecteur Puis classifieur de type (N23 vs N04)	51% sur test 1

La deuxième stratégie a été de sélectionner uniquement certains types d'appels pour l'entraînement (JD4).

Différents entraînements ont été réalisés pour classer automatiquement le clan A et G. Le TBC des différents modèles a été présenté et le résultat de ces derniers est présenté dans le chapitre Résultat (4.2.1).

3.3.1.2 Méthode non supervisée

La deuxième méthode qui a été utilisée pour la classification de clan est une méthode non supervisée : le clustering.

Pour cela, nous avons lancé le détecteur d'Orques (CNN) sur trois ans (2015, 2016 et 2017).

Puis, afin de décrire les caractéristiques de chaque appel, le pitch (hauteur) a été utilisé (Shapiro and Wang, 2009). Le pitch est une propriété qui décrit la fréquence fondamentale de l'onde de

la parole (Houtsma, 1997).

Comme les humains, les différents appels des Orques possèdent des harmoniques et la mesure du pitch est le contour de la fréquence fondamentale de ces derniers.

Nous avons donc utilisé une librairie en langage Python (Parselmouth Python library) afin d'extraire le pitch de chaque vocalise. Un exemple de quelques secondes est présenté en dans la figure 3.18.

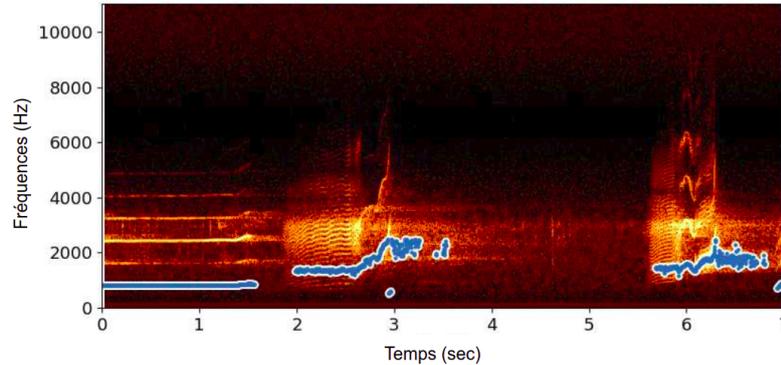


FIG. 3.18 – Exemple d'extraction du pitch tirée de Poupard et al. (2019b).

Puis, une fois le pitch extrait, nous avons récupéré les caractéristiques de chaque vocalise (Maximum, moyenne et minimum des fréquences ; maximum, minimum et moyenne des intensités ; minimum, moyenne et maximum des dérivés des fréquences (vitesse) et delta de la vitesse (accélération)).

L'ensemble de ces paramètres a été clusterisé à l'aide de l'algorithme HDBSCAN, et sont décrits dans Poupard et al. (2019b).

Les clusters trouvés à l'aide du regroupement automatique (méthode non supervisée) correspondaient légèrement aux différents types d'appels connus dans la bibliographie (Ford, 1989).

Mais ces résultats n'étaient pas assez satisfaisants pour classifier les différents pods de la zone. Les résultats de cette méthode non supervisée sont présentés dans Poupard et al. (2019b).

3.4 Méthode pour la détection individuelle

La détection individuelle en bioacoustique repose sur deux cas d'étude sous-marins : la détection individuelle de Cachalots depuis un drone mobile, et la détection individuelle d'Orques depuis une antenne fixe. Elle repose également sur un cas d'étude terrestre : la localisation d'individus (oiseaux) via une antenne multimicrophones.

L'ensemble des méthodes utilisées pour ces trois cas d'étude est présenté dans les parties suivantes.

3.4.1 Méthode pour le suivi individuel d'Orques

La première étude mise en place pour la détection individuelle se concrétise à la station de recherche Orcalab en Colombie-Britannique au Canada. L'ensemble du matériel est décrit dans

le chapitre précédent.

Une fois les enregistrements réalisés, différentes étapes ont été mises au point pour détecter et localiser les vocalises.

3.4.1.1 Annotation des enregistrements

Les extractions manuelles des vocalises ont été effectuées avec le logiciel Audacity pour apporter une forme d'onde et une visualisation en spectrogramme (Audacity-Team, 2018). Helena Symonds, experte des vocalisations d'Orques résidentes du Nord a annoté avec précision trois enregistrements et identifié les types d'appels pour chaque vocalise (voir figure 3.19).

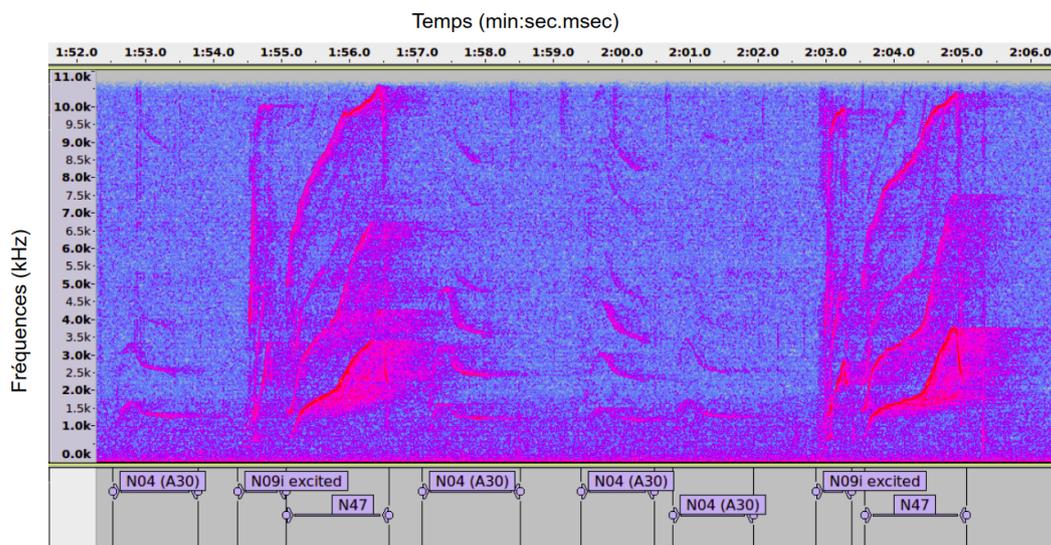


FIG. 3.19 – Exemple d'une annotation de 6 secondes réalisée par Helena Symonds faite sous Audacity.

3.4.1.2 Estimation des azimuts des vocalises

Dans la section précédente, chaque vocalisation a été détectée et annotée. À partir des quatre canaux, nous avons calculé les trois délais d'arrivée indépendants (TDoAs) par inter-corrélations : τ_{21} , τ_{31} , τ_{41} . La fréquence d'échantillonnage élevée (192 kHz) nous a permis de mesurer des TDoAs précis. À une distance relativement éloignée, le sinus de l'angle d'arrivée est égal à l'angle d'arrivée ; pour cette raison, nous estimons l'azimut avec un système comme présenté dans Bénard-Caudal et al. (2010) ; Marquardt (1963).

L'estimation de la distance des appels est possible dans certaines conditions (Glotin et al., 2008, 2014 ; Poupard et al., 2019d), mais en raison des effets acoustiques du canal, nous n'avons pas estimé la distance entre les Orques et l'antenne dans notre protocole.

Cependant, nous montrons que les estimations d'azimut sont exactes pour cette étude. Puis, nous joignons les estimations d'azimut de chaque appel à chaque Orque identifiée en utilisant les positions visuelles notées par les observateurs d'Orcalab. Nous associons des individus ou des groupes d'Orques à une piste de vocalisations en utilisant un catalogue de photo-identification (Towers, 2015). Lorsque les Orques sont très proches les unes des autres (moins d'un mètre,

voire photo annexe C), il est difficile d'associer un appel à un individu, mais pour certains exemples, nous avons pu associer des vocalisations à un individu en particulier.

3.4.1.3 Analyse semi-supervisée des vocalises

Nous avons identifié les vocalises d'individus isolés, puis les avons comparé par MFCC (Mel-Frequency Cepstral Coefficient), comme déjà utilisé dans les études sur les Orques (Logan et al., 2000 ; Ness et al., 2013). Plus le nombre de coefficients MFCC est important, plus la résolution du spectre est grande et moins l'analyse de clustering est générique. Dans Ness et al. (2013), 30 MFCCS ont été utilisées. Ici, nous avons extrait 30 MFCC ainsi que leurs delta et delta-delta (vitesse et accélération).

Nous avons supprimé le premier coefficient MFCC relatif à l'énergie afin d'éviter un biais de l'effet direct de l'appel, car les appels dans l'axe sont plus forts en énergie. Ainsi, nous représentons chaque appel dans 89 (= 29 + 30 + 30) dimensions.

Pour les comparer, nous avons utilisé la réduction de dimensionnalité t-SNE (Van der Maaten and Hinton, 2012).

Il s'agit d'une visualisation efficace et bien adaptée d'un ensemble de données complexe de grande dimension. La t-SNE s'est déjà révélé efficace pour les études en bioacoustique sur des dauphins (chapitre suivant). Elle traite de la similitude et de la dissemblance des données. Elle crée une distribution de probabilités sur des paires d'objets de grande dimension. Des objets similaires ont une forte probabilité d'être proches sur la représentation, tandis que des points dissemblables ont une faible probabilité d'être proches.

La deuxième caractéristique que nous avons extraite est le nombre de vocalises en fonction du temps (call rate : CR). Nous avons ensuite calculé l'estimation de la densité du noyau (KDE) (Terrell and Scott, 1992) sur ce taux, afin d'estimer la densité de probabilité du CR dans le temps. L'ensemble du protocole est résumé dans la figure 3.20.

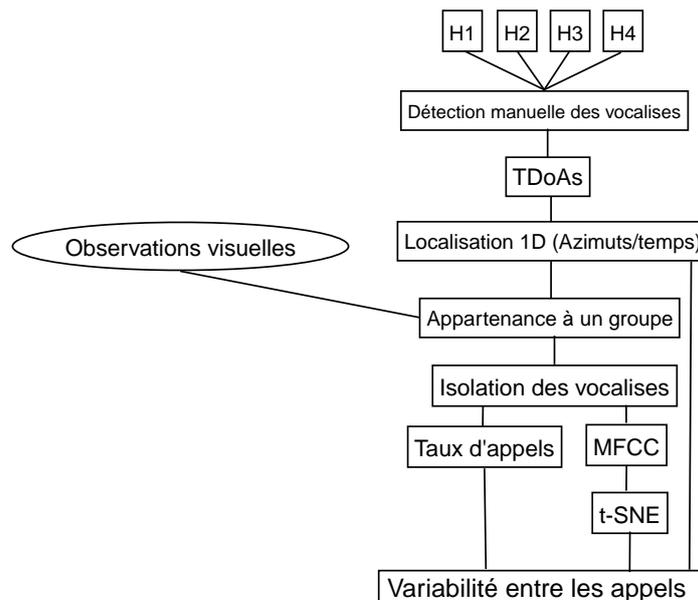


FIG. 3.20 – Architecture de la méthode utilisée pour l'extraction des vocalises pour l'identification individuelle.

L'ensemble des résultats pour ce cas d'étude est décrit dans le chapitre suivant (partie 4.3.1).

3.4.2 Méthode pour la détection individuelle de Cachalot

La deuxième étude sous-marine porte sur la localisation de Cachalots depuis un drone de surface : le Sphyrna. La partie Matériel a présenté l'équipement nécessaire pour cette étude. Nous avons collecté nos données le 12 août 2018 de 11h00 à 13h00, au large du Cap Sicié, au large de Toulon, alors que le drone dérivait à une vitesse moyenne de 0,26 m/s.

Les scientifiques à bord du trimaran qui accompagnait le drone écoutaient régulièrement les enregistrements réalisés à bord du drone. Nous avons donc détecté d'abord manuellement (analyse du spectrogramme) la présence d'un Cachalot.

Une fois les enregistrements réalisés, la totalité de leur traitement a été réalisé en laboratoire quelques semaines après la mission.

L'extraction des trajectoires de l'animal peut être résumée en 3 étapes principales :

- La détection des clics dans chaque canal
- Estimation du temps de délais d'arrivée (TDoAs) de chaque clic entre les canaux
- Reconstruction de la piste en 3 dimensions issue des TDoAs.

3.4.2.1 La détection de clics

La première étape de la détection de clics est une corrélation du signal avec une période d'un sinus de 12,5 kHz. Puis nous avons réalisé un filtre Teager-Kaiser (Kandia and Stylianou, 2006), ainsi que l'extraction des maximums locaux dans une fenêtre de 20 ms (le double de l'intervalle entre les pulses d'un clic de 10 ms (Abeille et al., 2014).

Afin de filtrer les bruits restants, nous récupérons uniquement les maximums au dessus de trois écart-types (Pukelsheim, 1994) de l'énergie logarithmique moyenne de toutes les détections. Ce seuil permet d'enlever les fausses détections. L'ensemble de ces étapes est décrit dans la figure 3.21.

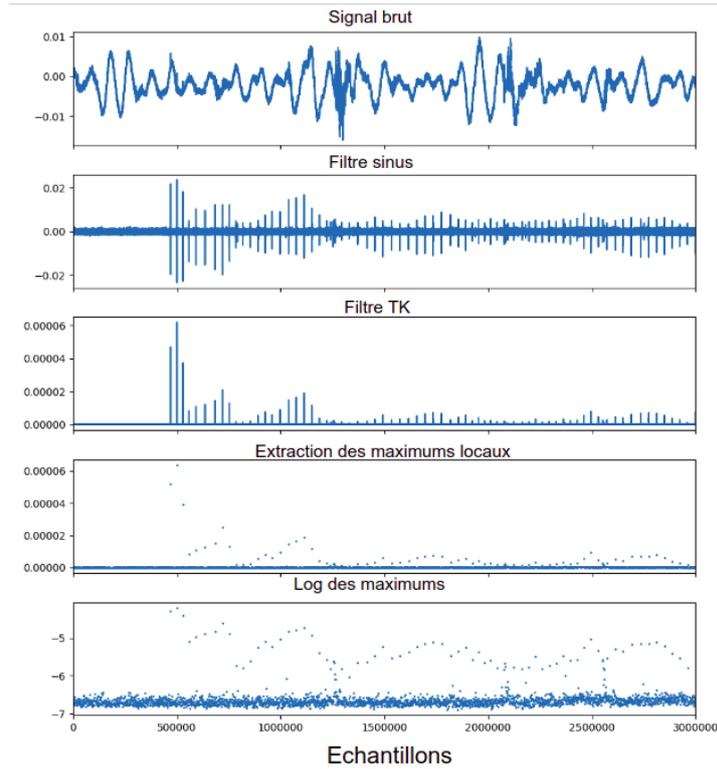


FIG. 3.21 – Représentation des différents filtrages pour la détection de clics de Cachalots enregistrés depuis le drone Sphyrna.

3.4.2.2 Estimation du délai d'arrivée

Une fois les clics détectés, le temps de délai d'arrivée entre les détections est calculé pour chacune des voies.

Pour cela, notons τ_{ij} le TDoA entre les hydrophones H_i et H_j et le signal x_i enregistré par l'hydrophone H_i .

L'estimation des TDoA est réalisée en deux parties :

tout d'abord, nous calculons les trois TDoAs indépendants $\tau_{21}, \tau_{31}, \tau_{41}$ de l'antenne de quille (H_1 à H_4), puis nous calculons le TDoA τ_{15} restant. Cette méthode permet de réduire le temps de calcul et d'estimer le TDoA avec le cinquième hydrophone lors de la deuxième partie.

Pour la première antenne sous la quille, nous calculons d'abord les corrélations (cross corrélation) entre les 6 paires d'hydrophones possibles, et les rendons positives en soustrayant leurs valeurs minimales. Ensuite, nous recherchons la combinaison de $\tau_{21}, \tau_{31}, \tau_{41}$ qui maximise le produit :

$$\tau_{21}, \tau_{31}, \tau_{41} =$$

$$\arg \max_{\hat{\tau}_{21}, \hat{\tau}_{31}, \hat{\tau}_{41}} \prod_{i=2}^4 \prod_{j=1}^{i-1} (x_i \otimes x_j - \min(x_i \otimes x_j))_{\hat{\tau}_{ij}}, \quad (3.2)$$

où \otimes indique une corrélation croisée et la relation $\tau_{ij} = \tau_{i1} - \tau_{j1}$ réduit la recherche à trois dimensions. Idéalement, cette méthode trouve la combinaison qui correspond aux clics

d'écholocalisation : contrairement au bruit de fond, ils sont cohérents entre les canaux. Pour les τ_{51} restants, nous utilisons la même méthode, mais maintenons $\tau_{21}, \tau_{31}, \tau_{41}$ fixe, ne recherchant que sur une dimension.

Cette étape nous permet de calculer précisément les TDoAs des clics détectés tout en diminuant le temps de calcul.

Avant de localiser les TDoAs, une étape de filtre fut nécessaire. L'antenne de 5 hydrophones (4+1) va nous être d'un grand recours pour éliminer les fausses détections. C'est particulièrement le cinquième canal qui va nous permettre de rejeter les détections parasites. Le nuage de points dans la figure Fig.3.22 représente les maximums des produits des corrélations avec d'un côté $\tau_{21} \dots \tau_{41}$, et de l'autre τ_{51} . Il est possible de voir deux modes :

l'un des clics d'écholocalisation, l'autre du bruit (histogrammes bleus à gauche et en bas). Nous filtrons les détections parasites en seuillant les faibles corrélations au sein de la sous-antenne, et les faibles corrélations avec H_5 (avec des seuils choisis manuellement en fonction des histogrammes).

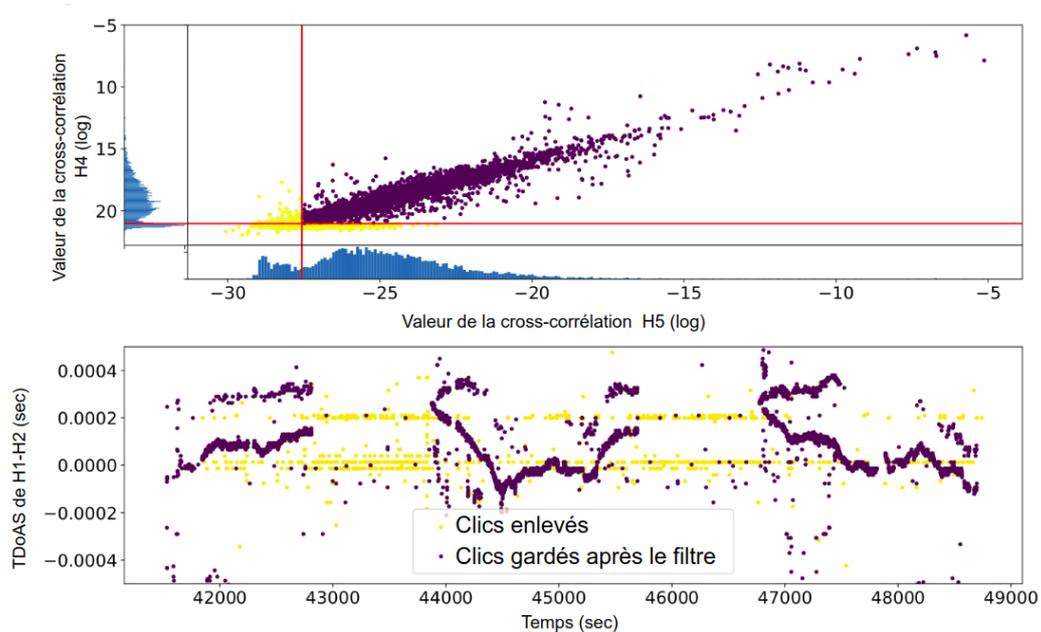


FIG. 3.22 – En haut : filtrage des valeurs de corrélation croisée (H_4 et H_5). En bas : TDoAs entre H_1 et H_2 au fil du temps. Les points jaunes ne sont pas utilisés pour l'analyse.

Nous voyons que cela filtre les clics qui ne semblent pas appartenir à la plongée, mais causés par le bruit du bateau ou d'autres sources (lignes horizontales jaunes sur la Fig.3.22, en bas). Nous voyons également des trains de clics parallèles dus aux réflexions de surface ; leur emplacement calculé est au-dessus du niveau de la mer, nous pouvons facilement les filtrer ultérieurement.

3.4.2.3 Localisation et ICI

À l'aide d'un solveur non linéaire, nous estimons les positions 3D du Cachalot (X, Y et profondeur Z) à partir des TDoAs (Giraudet and Glotin, 2006). Les pistes sont stabilisées en fonction du lacet et du roulis de la MPU (Motion Processing Unit) du navire, et lissées avec une moyenne glissante de plus de 20 s. Le bruit des données brutes est faible en raison de la stabilité du Sphyrna et des bonnes conditions météorologiques. Nous calculons également l'intervalle entre les clics (ICI). Pour supprimer les artefacts du solveur ou les faux positifs, la séquence ICI est lissée avec une moyenne cumulée de 32 secondes.

Cette méthode a donc été appliquée pour une trajectoire de Cachalot qui a duré plusieurs heures présentée dans le chapitre Résultat (partie 4.3.2).

3.4.3 Méthode pour la localisation aviaire

Le chapitre Matériel a présenté comment l'acquisition acoustique a été réalisée. Une fois les enregistrements réalisés, nous disposons de 14 enregistrements contenant de 0 à 4 individus. L'objectif est donc de mettre en place des méthodes pour compter le nombre d'individus présents.

Pour cela nous avons utilisé une détection automatique utilisée dans la partie 3.2.2. Cette algorithme a été conçu pour la construction de données d'entraînement pour la détection automatique d'Orques. Les principales étapes de l'algorithme sont la représentation du signal en spectrogramme, la binarisation du spectrogramme, appliquer des filtres (closing and dilation filter) et regrouper les clusters restants.

Puis les TDoAs de ces événements détectés ont été calculés afin de différencier les individus. Enfin un type de clustering (DBSCAN) a été appliqué afin de regrouper les chants émis en fonction de leur 3 TDoAs indépendants.

3.5 Méthode pour la bioacoustique comportementale

La principale caractéristique pour cette dernière échelle d'étude est qu'il n'y a pas d'identification acoustique au niveau spécifique, populationnelle ou individuelle.

L'objectif est donc de mettre en place un détecteur de signaux qui serait applicable pour différentes espèces et de mettre en corrélation ces signaux avec un comportement. L'acquisition des données de comportements a été présentée dans le Chapitre Matériel.

Les deux parties suivantes vont présenter les méthodes utilisées pour les deux problématiques décrites dans le chapitre précédent.

3.5.1 Méthode pour la bioacoustique du Dauphin tacheté pantropical

La première application en bioacoustique comportementale concerne les Dauphins tachetés pantropicaux, et notamment l'influence de leur comportement sur leurs émissions acoustiques. Une fois que les enregistrements ont été réalisés, il a fallu extraire tous les signaux biologiques (sifflements) de chaque enregistrement. Pour cela, une méthode automatique a été utilisée (Abeille et al., 2012). L'objectif de l'algorithme est d'extraire, de quantifier et de dater tous les sifflements d'un enregistrement. Ce détecteur a été implémenté en 2012 par l'équipe DYNi pour l'étude

DECAV du sanctuaire PELAGOS (Abeille et al., 2012).

Toutes les étapes présentées dans la détection automatique des sifflements ont été réalisées avec l'environnement Matlab. La première étape de cet algorithme est d'afficher un spectrogramme à partir du signal audio original. Puis le détecteur va analyser chaque pixel de ce spectrogramme et le transformer en une matrice de 0 et 1 : c'est la binarisation. Les zéros sont des points qui correspondent à des faibles amplitudes (bruits ambiants) et les 1 des signaux avec de fortes amplitudes. Donc les points sélectionnés possèdent une forte intensité (dB) par rapport au bruit naturel de l'écosystème.

Enfin l'algorithme va retenir les points formant une trajectoire continue de "1". Cette dernière manipulation va permettre d'ignorer les bruits discontinus de l'environnement. L'algorithme exclut donc des bruits tels le son des moteurs ou ceux des vagues et sélectionne les signaux biologiques : les sifflements.

Par la suite, certains paramètres dans l'algorithme ont été optimisés, comme la taille de la fenêtre ou encore les fréquences maximales et minimales de détection. Si ces paramètres ne sont pas optimisés, une partie des sifflements ne sera pas détectée, ou au contraire, l'algorithme peut prendre en compte des bruits de l'écosystème.

Pour l'espèce *Stenella attenuata*, après différents tests, les fenêtres de durée 256 ou 512 points ont été sélectionnées. La durée de ces fenêtres représente le nombre d'échantillons de son utilisé pour le calcul de la Fast Fourier Transform (FFT). Cette durée influence la résolution fréquentielle et temporelle de la détection. Une durée de fenêtre de 256 points (= $256/48000 = 5$ millisecondes) détecte mieux les sifflements rapides et larges en fréquence, tandis que la fenêtre de 512 échantillons acoustiques (de durée deux fois supérieure à la précédente, 10 ms) détecte des sifflements plus longs en temps et plus étroits en fréquence. Ce détecteur permet d'extraire automatiquement et très rapidement tous les sifflements peu visibles, car noyés dans le bruit ambiant, ce qui est un gain de qualité et de temps par rapport à la méthode manuelle qui consiste en l'inspection des images FFT par un expert acousticien. Il est possible, à partir de la matrice binaire, d'extraire automatiquement certains paramètres pour chaque sifflement. Donc pour chaque sifflement, 16 paramètres ont été extraits de manière automatique :

- Fréquence maximale, minimale, médiane, moyenne
- Durée du sifflement
- Vitesse du sifflement (pente), (minimale, maximale, médiane, moyenne et durée)
- Accélération du sifflement (dérivé de la pente), (minimale, maximale, médiane, moyenne et durée)
- Vitesse de l'accélération du sifflement.

La figure 3.23 représente l'extraction automatique de sifflements dans un enregistrement de 12 secondes.

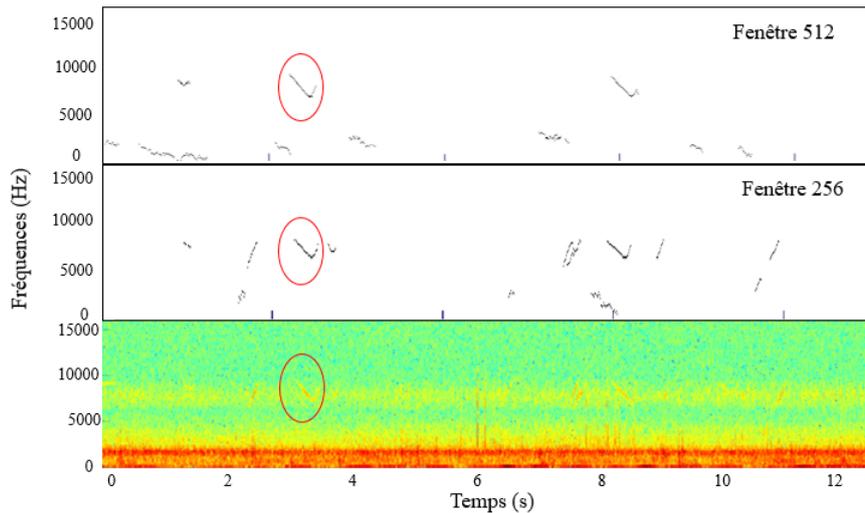


FIG. 3.23 – Spectrogramme de 13 secondes d'un enregistrement contenant des signaux de *Stenella attenuata* et la représentation finale suite à la détection automatique pour les différentes fenêtres.

Une ACP et une t-SNE ont été réalisées à l'aide du langage Python.

La t-SNE est utilisée également pour effectuer une réduction de dimensionnalité, donc pour la visualisation d'un jeu de données à dimensions multiples (Van der Maaten and Hinton, 2012). À l'inverse de l'ACP, c'est une technique non linéaire, mais elle va également permettre une bonne visualisation des données (Hinton and Roweis, 2002).

Le but de ces représentations est d'identifier différents groupes comportementales (différents groupes en fonction du comportement de surface attribué). C'est-à-dire que les sifflements appartenant au même comportement de surface, devraient être proches sur la représentation. Les coordonnées (abscisse et ordonnée) des réductions de dimensionnalité sont sélectionnées pour la réalisation de regroupement automatique (Clustering). Nous avons utilisé le clustering de la librairie Sklearn, le Bayesian Nonparametric Clustering (BNP). Cette méthode va permettre de comparer les groupes comportementaux (vérité) observés visuellement avec la t-SNE et l'ACP et ce nouveau groupe automatique.

Les méthodes de regroupement automatique sont beaucoup utilisées pour la classification (Frasier et al., 2016). Le regroupement automatique permet de mettre en évidence des objets appartenant au même groupe selon un critère spécifique (Lutz and Biernat, 2015). C'est une méthode particulièrement utilisée pour des jeux de données conséquents, comme la bioacoustique, l'analyse d'image ou l'apprentissage.

Le jeu de données est donc rassemblé en 2 distributions : la distribution des groupes statistiques trouvée à l'aide du BNP et les groupes de comportements (groupes comportementales visualisés avec la t-SNE). L'objectif est alors d'identifier la correspondance entre ces deux distributions.

Pour comparer ces deux distributions, nous avons utilisé le NMI. Le NMI (Normalized Mutual Info score) est une mesure de la dépendance mutuelle entre deux distributions, en particulier entre les groupes statistiques (trouvés avec le BNP) et les groupes estimés (groupes comportementaux). Le NMI est une normalisation du score de l'information mutuelle (MI). Il varie entre 0 (pas

d'information mutuelle) et 1 (corrélation parfaite entre ces deux distributions) (Fowlkes and Mallows, 1983). La cohérence des groupes est donc quantifiée à l'aide du NMI (Strehl and Ghosh, 2002).

Il existe deux parties différentes dans le calcul du NMI (Équation 2.2), le numérateur (I) est l'information mutuelle entre X et Y . Le dénominateur dépend de l'entropie de chaque partitionnement. Dans notre cas, X est la distribution du modèle automatique, et Y est le comportement des individus.

$$NMI(X, Y) = \frac{I(X, Y)}{\sqrt{H(X)H(Y)}} \quad (3.3)$$

Ainsi, le NMI compare les groupes comportementaux (observés visuellement avec t-SNE) par rapport aux groupes théoriques (trouvés automatiquement avec le BNP). L'ensemble de ces étapes est décrit dans la figure 3.24.

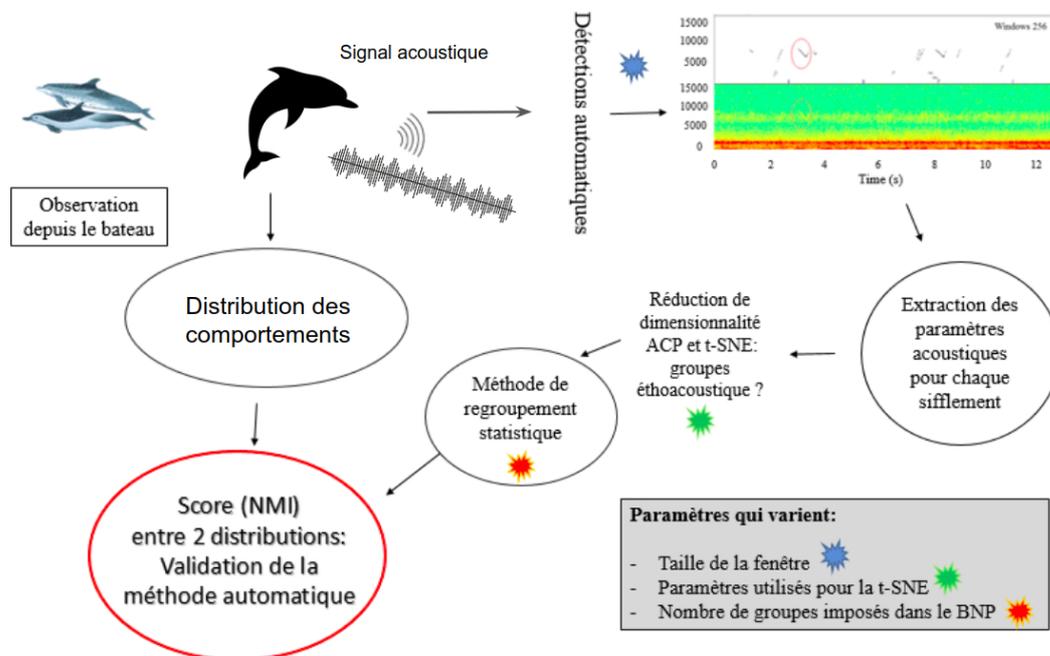


FIG. 3.24 – Architecture de l'analyse pour la bioacoustique comportementale du Dauphin tacheté pantropical.

La méthode en bioacoustique décrite ci-dessus a été mise en place afin qu'elle soit applicable pour différentes espèces produisant des sons stationnaires (vocalises ou sifflements).

Parallèlement à cette analyse, une analyse manuelle a été effectuée afin de valider la chaîne automatique. Pour cela, les mêmes enregistrements ont été affichés sur le logiciel Spectrogram 5.0. (Electronics Lab, USA). J'ai isolé chaque sifflement visuellement sur le spectrogramme et j'ai extrait différents paramètres acoustiques pour chaque sifflement :

- le début, la fin ainsi que la durée du sifflement en millisecondes,
- le nombre des points d'inflexion,
- la fréquence (Hz) en $\frac{1}{4}$, $\frac{1}{2}$ et $\frac{3}{4}$ de la durée du sifflement,

- la fréquence initiale, finale, minimale et maximale.

Une analyse statistique a été réalisée afin de comparer ces paramètres avec les comportements de surface.

Pour cela, les moyennes et écart-types des fréquences initiales, finales, maximales et minimales ont été calculées. La normalité des fréquences a été analysée à l'aide du test de Shapiro-Wilk. Les données de fréquences initiales, finales, minimales et maximales ne suivant pas une distribution normale, l'analyse de ces données a donc été réalisée grâce au test de Kruskal-Wallis. La distribution normale des données de durée des sifflements a été constatée après transformation logarithmique. La variation de la durée du sifflement a été testée par une ANOVA. Si la variation était significative (c'est-à-dire pour $p < 0,05$), un test d'homogénéité des variances était réalisé (test de Levene). En cas de variances homogènes, le test a posteriori de Tuckey a été appliqué. Et dans le cas contraire, c'est le test de Games et Howell qui était choisi.

Cette méthode manuelle va nous permettre de valider ou non la chaîne de méthodes automatiques présentée dans la figure 3.24.

3.5.2 Méthode pour la bioacoustique des Baleines à bosse

L'objectif de cette étude est de déterminer si les baleines ont une réponse comportementale et acoustique à différents stimuli chimiques, diffusés dans la colonne d'eau. La partie matériel a présenté le protocole utilisé pour la mise en place de cette étude.

Une fois les enregistrements réalisés, nous avons analysé 120 minutes d'enregistrement :

- 30min avec du krill dans le milieu
- 30min avec de l'argile dans le milieu
- 30min avec du DMS dans le milieu
- 30min avec du CTL dans le milieu

Afin d'extraire les vocalises de ces enregistrements, nous avons utilisé le détecteur spectral (décrit dans la partie précédente).

Pour cette espèce, nous n'avons besoin que d'une taille de fenêtre : 256 échantillons.

Quinze paramètres acoustiques ont été extraits pour chaque vocalisation (fréquence minimale, maximale, médiane et moyenne, durée de la vocalisation, durée minimale, maximale, médiane et moyenne de la vitesse, durée minimale, maximale, médiane et moyenne de l'accélération). Nous avons projeté ces 15 variables dans un espace à deux dimensions grâce à la t-SNE (Distributed Stochastic Neighbor).

Nous avons utilisé les coordonnées optimales en terme de séparation de chaque vocalisation dans l'espace t-SNE pour construire des clusters bayésiens non paramétriques (BNP) (Bishop, 2006). Cette méthode de clustering s'est avérée efficace pour les analyses bioacoustiques des cétacés (voir partie précédente), permettant la classification des vocalisations avec un haut degré de similitude intraclasse (ressemblance au sein d'un cluster) et un faible degré de similitude interclasse (dissemblance entre les clusters).

Puis, nous avons mesuré la qualité des clusters avec des scores d'Information Mutuelle Normalisée (NMI) afin de quantifier la correspondance (c'est-à-dire la dépendance mutuelle ou l'information mutuelle) entre les variables acoustiques et chimiques.

Le NMI est un scalaire entre 0 (les deux variables ne partagent aucune information) et 1

(correspondance parfaite entre les groupes acoustiques et la condition du stimulus).

La méthode pour cette étude est donc identique à la partie 3.5.1 :

- Extraction des vocalises
- Extraction des paramètres acoustiques
- Réduction de dimensionnalité (t-SNE)
- Clustering (BNP)
- Calcul de l'information mutuelle (NMI).

Le même protocole a été appliqué pour ces deux études : un enregistreur avec un support visuel afin d'identifier visuellement les espèces présentes. Puis un détecteur de signaux stationnaires a été mis en place afin d'extraire les paramètres acoustiques de chaque vocalise ou sifflement. Cette échelle d'étude est donc appelée *bioacoustique comportementale* car nous avons détecté acoustiquement des signaux biologiques (vocalises et sifflements) sans détecter acoustiquement les espèces, populations ou individus. Un support visuel est donc indispensable pour ce protocole. L'ensemble des méthodes mises en place a été décrit dans ce chapitre. Chaque cas d'étude associé à une échelle d'analyse comporte une méthode pour l'analyse des signaux. La figure 3.25 décrit l'ensemble des méthodes utilisées. Les cases vertes représentent les procédés automatiques, et les cases grises les procédés manuels. L'objectif principal était de rendre les procédés les plus automatiques possible à cause de la large quantité de données à traiter.

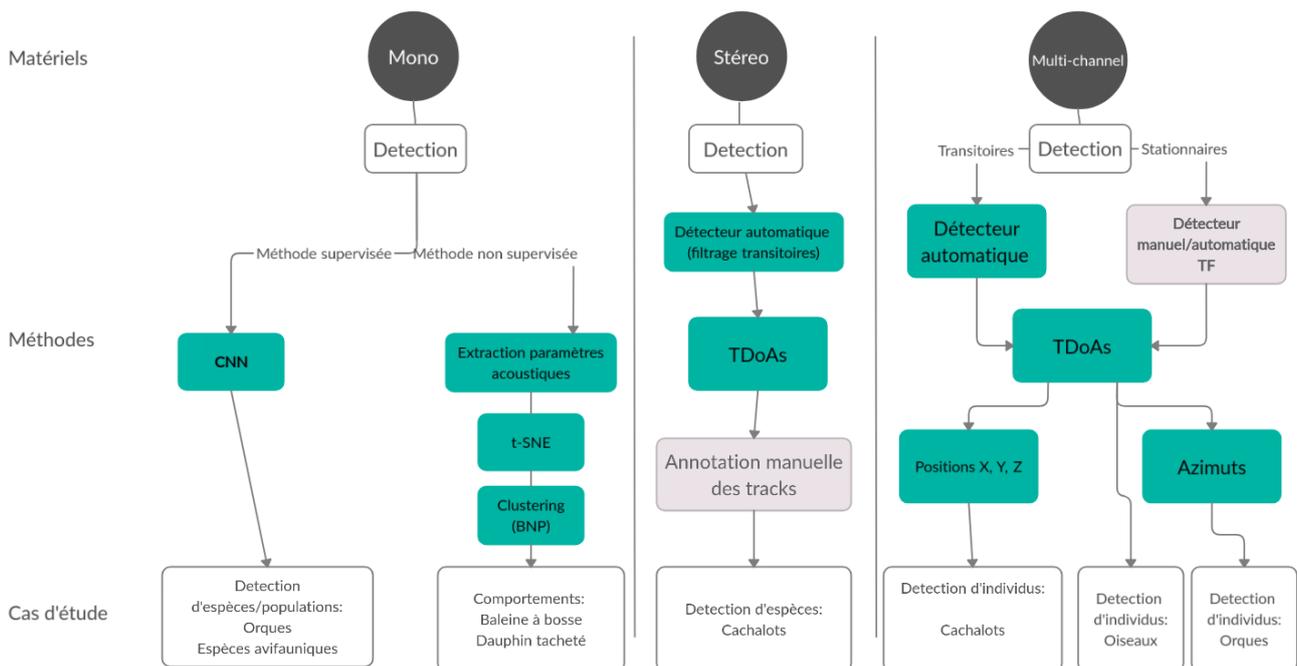


FIG. 3.25 – Représentation de l'ensemble des méthodes utilisées dans cette thèse. Les carrés verts : méthodes automatiques, carrés gris : méthodes manuelles, carrés blancs : cas d'étude.

Chapitre 4

Résultats : validation des concepts et tendances bioacoustiques

Ce chapitre présente les résultats pour chacune des échelles d'étude : bioacoustique spécifique, populationnelle, individuelle et comportementale. Il sera donc divisé en quatre grandes parties au sein duquel les différents résultats des cas d'étude sont présentés. A la fin de chaque cas d'étude, les résultats seront discutés brièvement en terme de bioacoustique. La discussion générale sur la critique des méthodes sera développée dans le chapitre suivant.

4.1 Détection d'espèces : présence/absence

La première échelle d'étude, la plus large, concerne la détection d'espèces. Trois modèles d'études ont été utilisés : les oiseaux (50 espèces), les Orques et les Cachalots.

4.1.1 *Analyse de la présence d'espèces avifauniques : application Birdify*

La première application en terme de détection d'espèces et perspectives en bioacoustique repose sur la classification d'espèces d'oiseaux. Les chapitres précédents ont développé le matériel nécessaire et la méthode utilisée pour cette tâche. Nous avons donc utilisé une méthode d'intelligence artificielle pour la détection et classification d'oiseaux. L'objectif du développement de ce modèle est la mise en ligne d'un outil de classification d'oiseaux disponible sur un appareil mobile dans une perspective pédagogique mais également d'observation des déplacements des espèces à l'échelle du pays. Le deuxième objectif est davantage scientifique et concerne l'évolution des performances d'un tel modèle de classification.

Pour rappel, notre modèle a un taux de bonne classification de 88% sur la classification de 50 espèces d'oiseaux.

Nous avons alors testé les performances de notre modèle dans de vraies conditions environnementales. Comme décrit dans le Chapitre "Matériel", nous avons mis au point un protocole de diffusion acoustique d'espèces via des enceintes, afin d'évaluer la distance à laquelle le son est perçu par le téléphone et bien classé par notre modèle.

La figure 4.1 représente 15 secondes de son de Troglodyte mignon enregistrées par un Iphone, à différentes distances d'émissions. Plus la distance est grande, plus les signaux sont confondus avec le bruit. C'est cet effet que nous essayons de mesurer.

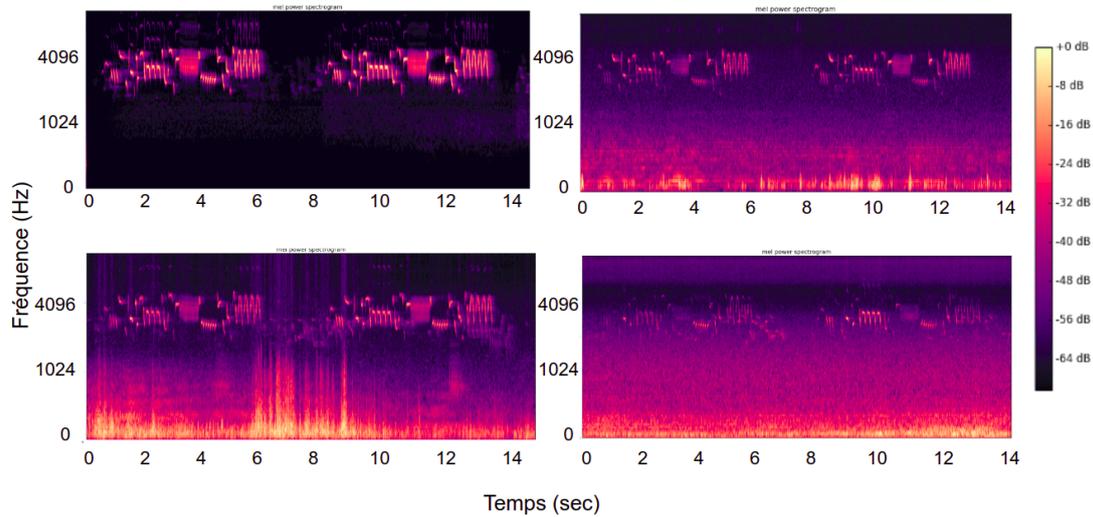


FIG. 4.1 – Spectrogrammes (échelle Mel) représentant les différences de dB entre les sons reçus. Haut Gauche le spectrogramme initial du son émis. Bas gauche : le sons reçu à 10m. Haut droit : son reçu à 50m, Bas droit : son perçu à 100m. Le son émis était 15 secondes de Troglodyte mignon enregistrées avec un iphone 4.

La figure suivante (Fig.4.2) représente la performance du modèle en fonction des téléphones (4 modèles + un enregistreur Zoom), de l'angle d'écoute et de la distance pour le Troglodyte mignon (*Troglodytes troglodytes*). Cette figure nous montre que le téléphone le plus performant est le Wiko et le Samsung. Comme attendu, plus la distance d'émission augmente, plus les performances chutent et atteignent 0 pour une distance d'émission de 100m. Lorsque le micro est à 25 m de la source, la probabilité est toujours supérieure à 80%, et ce, pour les 5 appareils. A 100 m de distance, seul le Wiko est capable d'identifier l'espèce. L'alignement entre le micro et l'oiseau peut également influencer la détection. Les probabilités sont moins bonnes avec un angle de 90° que lorsqu'il est en face du micro, à une distance de 25 m. Mais la performance va également évoluer en fonction de l'espèce à classifier.

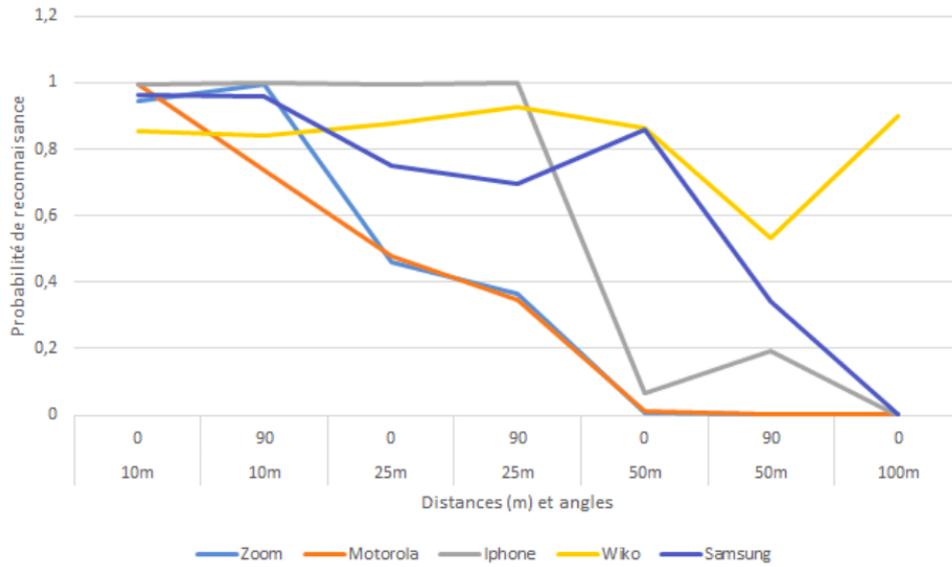


FIG. 4.2 – Probabilité de reconnaissance du Troglodyte mignon (*Troglodytes troglodytes*) en fonction du téléphone, de la distance d’émission et de l’angle d’écoute.

La figure 4.3 illustre avec quelle espèce le Troglodyte est confondu lorsqu’on augmente la distance et l’angle d’émission. Jusqu’à 50m, la prédiction est toujours correcte (Troglodyte en vert), puis à 50 m de distance, le modèle commence à confondre avec cinq espèces différentes notamment le Grimpereau des jardins (*Certhia brachydactyla*).

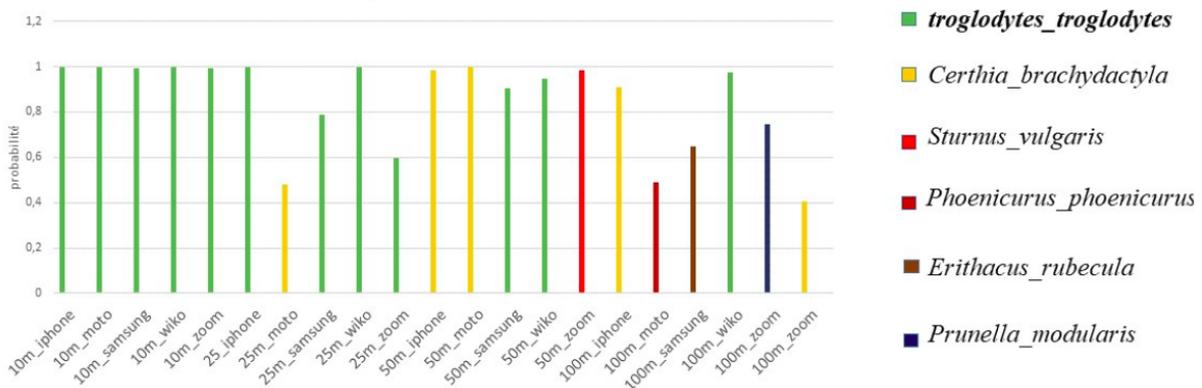


FIG. 4.3 – Évolution des probabilités maximales du modèle en fonction de la distance d’émission en émettant le chant du Troglodyte mignon.

L’expérience sur le terrain montre que les micros des téléphones sont suffisants par rapport à l’efficacité de notre modèle car les probabilités de reconnaissance restent hautes pour plusieurs modèles classiques jusqu’à 25 m de distance quelque soit l’orientation du téléphone, et 50 m pour un téléphone bien orienté (pour les modèles Wiko et Samsung).

Il ne sera donc pas nécessaire d’intégrer un microphone externe pour faire fonctionner l’application et assurer un enregistrement homogène quel que soit les espèces.

Dans un premier temps, l’application fonctionnerait uniquement via internet, les sons seront

transférés sur un serveur en GPU de notre laboratoire qui calculera les probabilités de reconnaissance puis retransmettra l'information à l'utilisateur via l'application.

Un prototype d'application a été développé par un stagiaire en développement WEB que j'ai co-encadré, la prochaine étape est donc de tester ce prototype à plus grande échelle avec des ornithologues opérationnels.

Afin d'évaluer la concurrence des applications de reconnaissance d'oiseaux existantes, j'ai testé les plus connues : Birdup, Cuicui ChirpOMatic, Bird Song et enfin Birnet. Pour cela, le protocole consiste pour toutes les applications, y compris la nôtre, de jouer un banc d'essai espèce par espèce, avec le haut parleur d'un ordinateur (PC Gigabyte) devant le téléphone (à 10 cm environ), et d'envoyer ce son en requête à l'application. Si l'application fait une erreur de classification, alors le test est renouvelé pour une seconde chance, avec un casque audio (Panasonic RP-HD10) placé à 5 cm du téléphone. Les téléphones utilisés sont le SAMSUNG A3 pour les applications Android et l'Iphone 5 pour IOS.

Ces expériences sont menées dans une pièce sans aucun autre bruit. Nous avons vérifié que tous les sons étaient bien enregistrés par l'application (pas de saturation). Lorsque l'application exigeait uniquement une partie du spectrogramme (le cas pour BirdNet et BirdGenie), la partie la plus claire était sélectionnée par nous-même.

Afin de comparer les performances des applications, nous avons calculé le taux de bonne classification (TBC). Le résultat de notre modèle sur le banc d'essai de cette étude (non inclus bien sûr en aucune sorte dans les données d'apprentissage ni de validation de notre modèle), donne un TBC de 86% en considérant la 1ère réponse seulement (maximum des probabilités). L'intégralité de cette étude de comparaison entre les applications existantes et notre modèle est présentée dans le rapport d'étude (Poupard et al., 2018d).

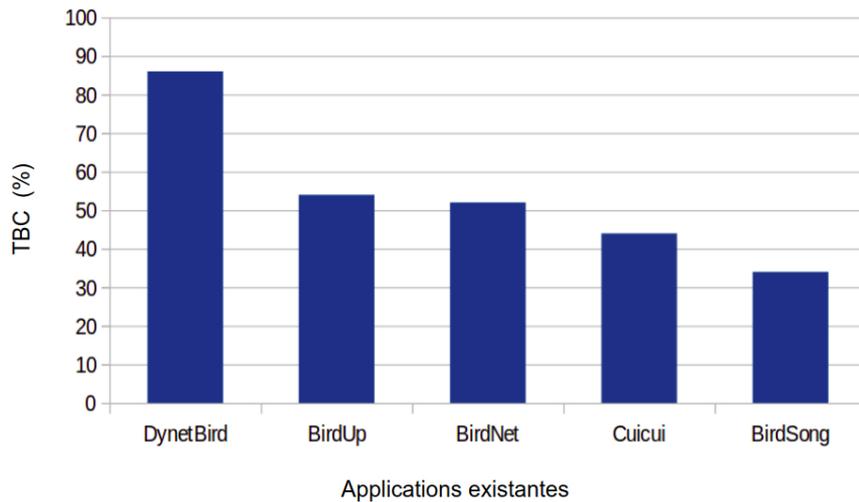


FIG. 4.4 – Graphique récapitulatif des scores de TBC en fonction des différents modèles disponibles sur le web comparé à notre modèle DynetBird. Il y a une différence importante entre elles (toutes ont un TBC entre 50 et 33%) et la nôtre à 86%. Voir rapport de recherche Poupard et al. (2018d).

La figure 4.4 présente le bilan des applications classées par TBC. Nous apportons donc par cette étude la preuve par mesures objectives, sur plus de 250 tests de reconnaissance effectués, que notre modèle est le plus performant pour une application sur des espèces en France métropolitaine. Il est le seul à satisfaire le critère de qualité imposé afin de satisfaire au mieux les usagers (i.e. plus de 80% de reconnaissance). Il est important de souligner que ces tests ont été réalisés en août 2018, et qu'il est donc probable que ces applications aient été améliorées ou modifiées.

4.1.2 Analyse de la présence d'Orques à Orcalab

L'objectif de cette sous partie était de détecter automatiquement la présence d'Orques (*Orcinus orca*) présentes dans les eaux de la Colombie-Britannique à coté de la station de recherche Orcalab.

La partie méthode a exposé les performances du modèle utilisé. Une fois entraîné, ce modèle a été exécuté sur 3 ans d'enregistrements (2015, 2016 et 2017) et en 2 jours de calcul sur un GPU NVIDIA Titan X Pascal.

Les résultats en figure 4.5 montrent la proportion de fichiers avec des détections d'Orques parmi tous les enregistrements pour chaque mois et pour chaque hydrophone.

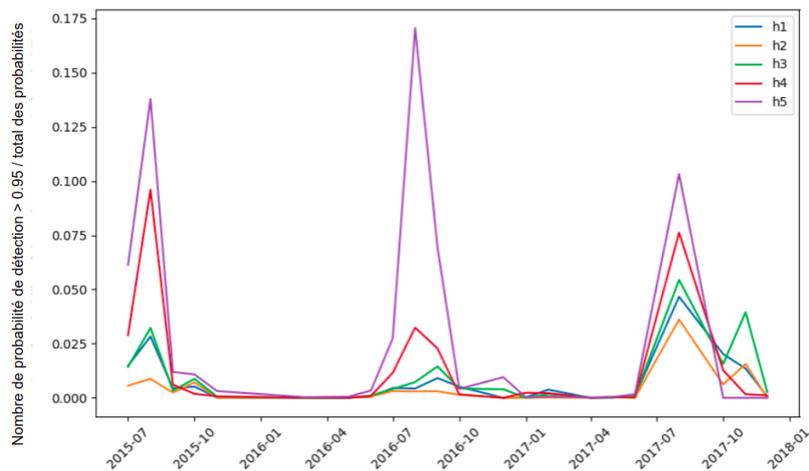


FIG. 4.5 – Proportion d'enregistrements de deux minutes contenant des vocalises d'Orques par mois et par hydrophone, de 2015 à 2017.

Les Orques seraient présentes (acoustiquement) principalement pendant l'été (juin, juillet, août et septembre). Cette migration est confirmée par Nichol and Shackleton (1996). Cette zone est très empruntée par les Orques durant cette période. En effet, elles sont attirées par la présence du Saumon chinook (*Oncorhynchus tshawytscha*), proie principale des Orques résidentes de la Colombie-Britannique.

Afin d'estimer les trajectoires des animaux, on peut mesurer l'activité acoustique des Orques dans la gamme de chaque hydrophone dans le temps.

Cette idée a été concrétisée le 24 août 2017 avec plusieurs techniques afin de démontrer la performance du modèle utilisé.

Nous avons d'abord mis en place l'extracteur automatique d'événements acoustiques, en prenant

le nombre d'événements détectés afin d'estimer l'activité acoustique des Orques. Le résultat est présenté figure 4.6. La trajectoire de H5 à H1 (voir carte 2.15) peut être déduite par les taux de détection élevés successifs des hydrophones, mais le signal est très bruyant car ce détecteur n'est pas spécifique aux Orques et détecte beaucoup de faux positifs.

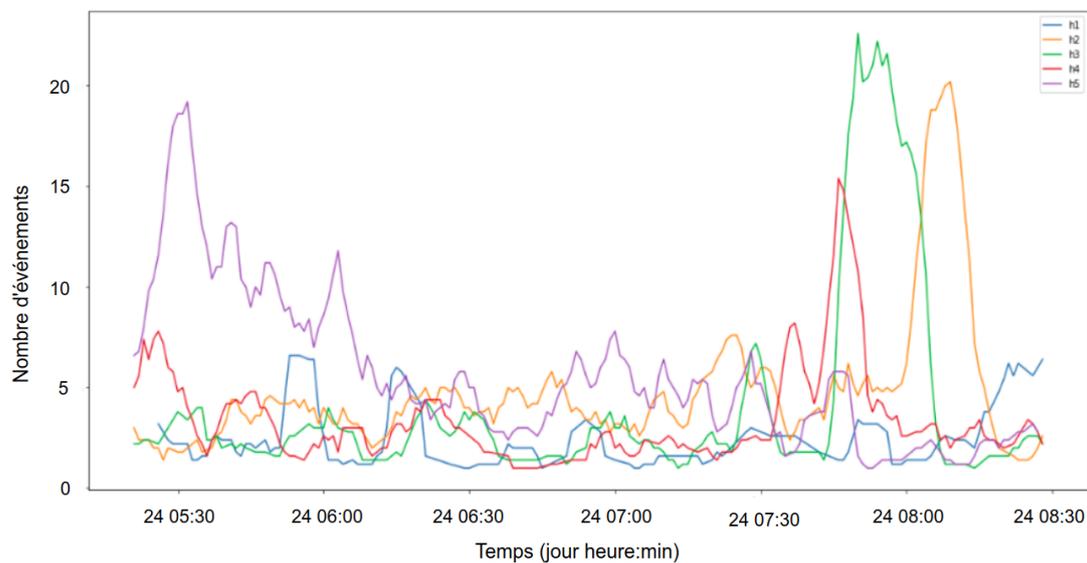


FIG. 4.6 – Évolution du nombre d'événements acoustiques d'Orques détectées de 5h30 à 8h30 le 24 août 2017. Résultats issus du détecteur "d'évènements acoustiques".

La deuxième méthode utilisée était notre modèle d'apprentissage entraîné pour la détection des vocalises d'Orques (CNN), donnant ainsi une bien meilleure estimation de l'activité des Orques dans le temps et sur chaque hydrophone (voir figure 4.7). Les résultats valident la déduction primaire du détecteur d'évènements acoustiques (Fig 4.7, trajet de H5 à H1 à 5h30), mais de manière plus précise.

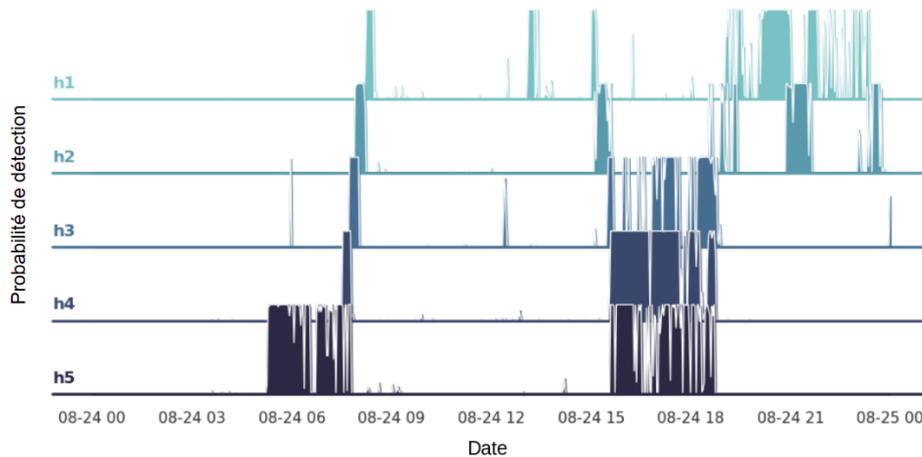


FIG. 4.7 – Exemple d'évolution de la probabilité de détection de vocalises pour chaque hydrophone sur une journée (24 août 2017). Le matin, un groupe d'Orques vient de l'est (voir carte 2.15) sur H5, et se déplace sur H4, H3, H2 puis sur H1. Différents voyages allers-retours sont effectués pendant la journée. Résultat issu du CCN.

Nous avons ensuite analysé plus en détail les activités acoustiques dans certaines plages horaires :

- La journée est de 11h à 17h (25% des enregistrements)
- La nuit est de 23h à 5h (25% des enregistrements)
- La pleine lune est de 4 jours avant à 4 jours après une pleine lune (31% des enregistrements)
- La nouvelle lune est de 4 jours avant à 4 jours après une nouvelle lune (24% des enregistrements)
- La marée montante est une période avec un différentiel positif de 40 cm du niveau de la mer (11% des enregistrements)
- La marée descendante est une période avec un différentiel négatif de 40 cm du niveau de la mer (11% des enregistrements).

Nous montrons (Fig.4.8) l'évolution globale des vocalises des Orques en fonction de ces conditions. Nous voyons que le nombre de vocalisations est plus élevé pendant la journée, la pleine lune et la marée montante par rapport à la nuit, la nouvelle lune et la marée descendante respectivement. Plus précisément, quatre fois plus d'enregistrements incluaient des Orques pendant la pleine lune (14%) par rapport à la nouvelle lune (3%) et 8 fois plus pendant les marées montantes (25%) par rapport aux marées descendantes (3%).

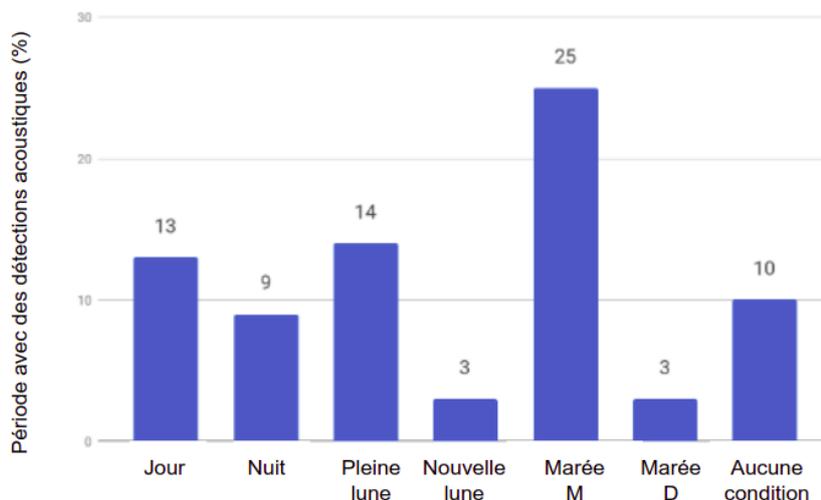


FIG. 4.8 – Pourcentage d'enregistrements avec des détections acoustiques d'Orques en fonction des conditions temporelles. Marée M= Marée Montante, Marée D= marée Descendante.

À notre connaissance, il s'agit de la première analyse de détections automatiques à grande échelle sur les Orques. Les résultats sont cohérents avec Benoit-Bird et al. (2009) ; Simonis et al. (2017) qui ont montré que la phase lunaire pourrait jouer un rôle important sur le comportement de deux espèces de cétacés, suggérant que leur présence serait corrélée avec des proies qui migrent verticalement (Roper and Scarnecchia, 1999). Nous pouvons supposer que, durant la pleine lune, les proies sont d'avantage présentes dans les couches d'eau supérieures, facilitant ainsi le travail des Orques pour les atteindre et donc pourrait modifier leur activité acoustique. Cette analyse nous a permis de comprendre le comportement de l'espèce sur de grandes échelles temporelles et spatiales grâce à l'automatisation de l'analyse des données, notamment sur l'extraction des vocalises.

Cette étude pourra être fondamentale pour la classification automatique des types d'appels émis dans le but d'identifier les différents clans présents dans la zone. L'ensemble de cette étude est présenté dans (Poupard et al., 2019a).

4.1.3 Analyse de la présence de Cachalots sur la bouée Bombxy

La deuxième application sous-marine pour la détection d'espèces se focalise sur les Cachalots de Méditerranée. L'ensemble de la méthode a été décrite dans la partie précédente. La figure 4.9 montre l'effort d'échantillonnage réalisé sur la bouée BOMBYX. Les enregistrements ont débuté en Juin 2015 et ont terminé en Novembre 2018. La période de Juillet à Septembre 2016 a été enregistrée et a déjà été traitée dans le rapport de l'étude commandé par le sanctuaire des cétacés PELAGOS (Glotin et al., 2017). Elle n'a pu être recalculée avec notre méthode améliorée dans cette thèse, car nous avons subi une perte de ces données brutes sur notre NAS de 500 To. Elle n'est donc pas reportée dans cette thèse. Cependant la méthode de traitement proposée dans (Glotin et al., 2017) a été appliquée sur une session de 2017, et sera donc comparée avec la méthode améliorée de cette thèse.

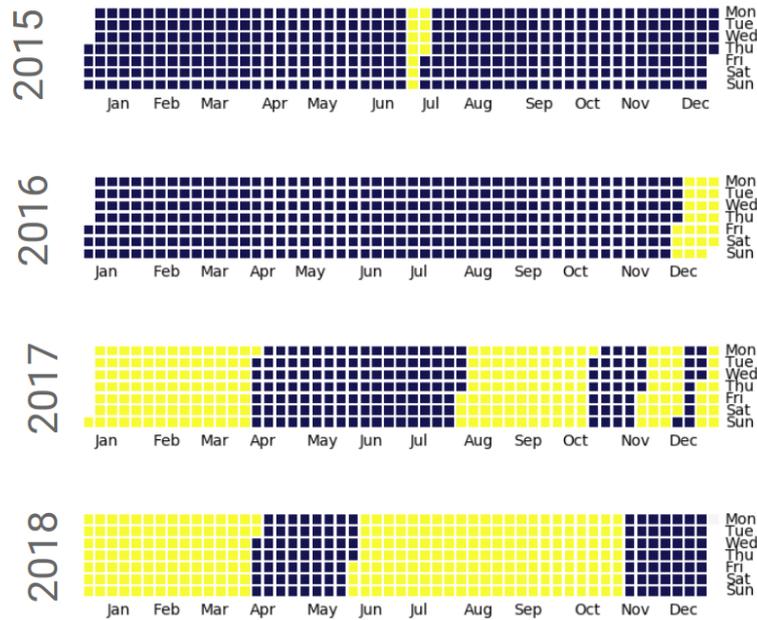


FIG. 4.9 – Effort des enregistrements sur la bouée BOMBYX, Jaune : enregistrements. Noir : sans enregistrement de 2015 à 2018.

Une fois les passages extraits, j’ai pu compter le nombre d’animaux sur l’ensemble des sessions. Ce résultat est présenté figure 4.10. Sur ce graphique, les quatre années sont représentées en terme de nombre d’individus par jour.



FIG. 4.10 – Calendrier représentant la présence de Cachalots sur l’ensemble des sessions. Les cases jaunes claires représentent les jours sans Cachalot, et les cases blanches les jours sans enregistrement.

Le mois de Décembre 2016 a été très fréquenté par cette espèce, tandis que décembre 2017 et 2018 ont été presque vides. Ce résultat global nous montre qu'il n'y aurait pas de véritable cycle saisonnier ou annuel pour cette espèce.

La différence de fréquentation de la zone entre décembre 2016, 2017 et 2018 pourrait être expliquée par les variations du courant Liguro-Provençal. Ce dernier s'écoule depuis le Golfe de Gêne, aux Baléares, en passant par les Îles d'Hyères et le Golfe du Lion. Ce courant venant de l'Est s'amplifie en hiver (>0.8 m/s), et est plus faible en été (<0.5 m/s) (Petrenko, 2003). Lorsque le courant est fort en hiver, il devient très instable et génère des méandres (Guihou et al., 2013) et donc potentiellement des zones ponctuelles d'accumulation de matière organique. Ces zones peuvent donc être, pendant quelques semaines fréquentées par les Cachalots chassant les céphalopodes. Ce qui expliquerait la forte présence des animaux durant Décembre 2016.

Une collaboration avec des scientifiques du laboratoire MIO (Institut Méditerranée d'Océanologie) de l'Université de Toulon débute, afin de mettre en commun leurs modèles courantologiques de la zone avec nos données de présence des Cachalots, et d'expliquer ces différentes fréquentations. On peut constater qu'il y a régulièrement plusieurs passages d'individus isolés (probablement des mâles) et plus rarement une concentration de plusieurs animaux (jusqu'à 9 individus en même temps). Ceci concorde avec Best (1979); Whitehead et al. (1991) qui ont montré que les grands groupes de Cachalots étaient composés de mères avec des jeunes, tandis que les individus isolés étaient des grands mâles.

Différentes études ont été réalisées sur la présence de Cachalots dans le bassin méditerranéen. Frantzis et al. (2011) ont travaillé en observation visuelle, et Borsani and Pavan (1994) ont travaillé sur les émissions acoustiques. Ils ont constaté que le nombre maximal d'individus dans un groupe n'a jamais dépassé 4. L'étude de Drouot et al. (2004) a notamment montré que la présence de Cachalots était assez rare (Mer Tyrhénienne) suggérant une tendance à la baisse de cette espèce.

Leur étude a été réalisée sur cinq ans et elle a relevé la présence de groupes de Cachalots atteignant plus de 7 individus. Cette étude a notamment travaillé sur les tailles des individus grâce à l'IPI (Inter Pulse Interval). Elle a mis également en évidence les variations des tailles des groupes en fonction de certaines régions méditerranéennes.

Additionnellement au nombre de passages, j'ai travaillé sur le nombre de sessions annotées positives à la présence de Cachalots. Afin d'homogénéiser le temps d'échantillonnages entre les sessions (temps d'enregistrements/temps de pause), les nombres de sessions annotées positives ont été multipliées par $Beta = \text{tempsderecord} + \text{tempsdepause}$ (voir Annexe E).

La figure 4.11 présente la moyenne du nombre de sessions positives pour le jour et pour la nuit.

La moyenne du nombre de sessions annotées positives à la présence de Cachalot est de 21 sessions pour la nuit, et de 11 sessions pour le jour. Afin de vérifier si les échantillons suivaient une loi normale, le test de Shapiro wilk a été réalisé.

La p-value était inférieure à 0.05 donc l'échantillon ne suivait pas une loi normale, et un test non paramétrique a été réalisé : le test de Wilcoxon-Mann-Whitney. Ce dernier permet de tester l'hypothèse selon laquelle les médianes de chacun des deux groupes de données sont proches (Wilcoxon, 1992). Ce test a été réalisé entre les distributions la nuit vs le jour. La p-value étant inférieure à 0.05, on peut donc rejeter H_0 et affirmer que les distributions sont

différentes. Donc, les nombres de détections la nuit vs le jour sont statistiquement différents : les Cachalots fréquentent d'avantage la bouée durant le jour.

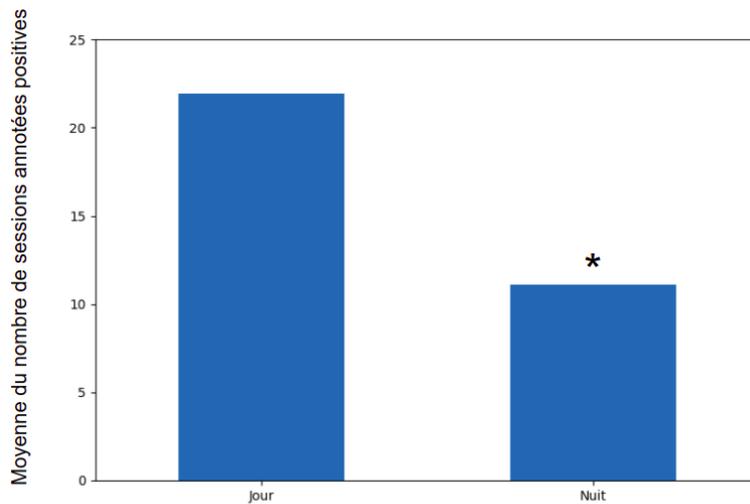


FIG. 4.11 – Moyenne du nombre de sessions (1min) annotées positives par jour (24h) pour la nuit et le jour avec sa différence significative.

Afin de visualiser la fréquentation de la bouée par les Cachalots durant le cycle journalier la figure 4.12 présente la somme des sessions annotées positives à la présence de Cachalots pour chaque heure de la journée. Sur cette figure 4.12, il est possible de voir le pic journalier de la fréquentation de la bouée autour de 13h.

Le rapport scientifique Glotin et al. (2017) présente les résultats d'un autre détecteur sur les enregistrements de BOMBYX en 2016 (Juillet à septembre). Il est probable que leur détecteur obtienne plus de faux négatifs car aucune vérification manuelle n'a été réalisée. Il n'y avait pas de tendance horaire dans les totaux dans leur travail alors notre étude démontre un effet net avec une fréquentation forte vers 13h.

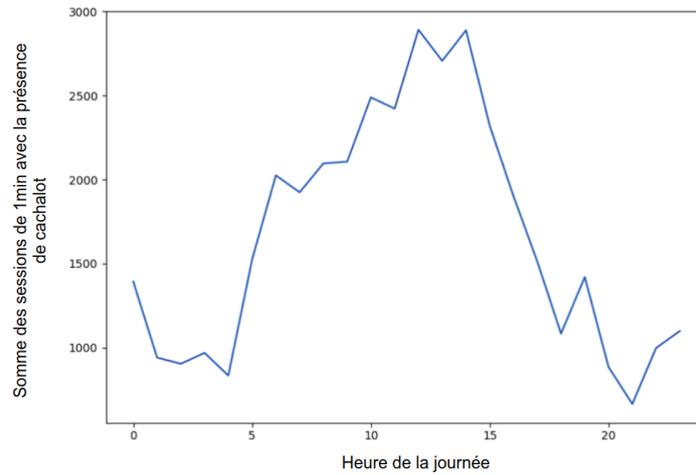


FIG. 4.12 – Évolution de la somme des sessions annotées positives à la présence de Cachalots au cours de la journée.

Puis, il est possible de dénombrer le nombre de pistes par jour (24h). On peut voir sur la figure 4.13 que pour 39 jours il n'y a eu qu'une piste (un seul individu), pour 26 jours il y a eu deux passages. Les jours où il y a plus de 5 pistes détectées sont relativement rares (7 jours en tout).

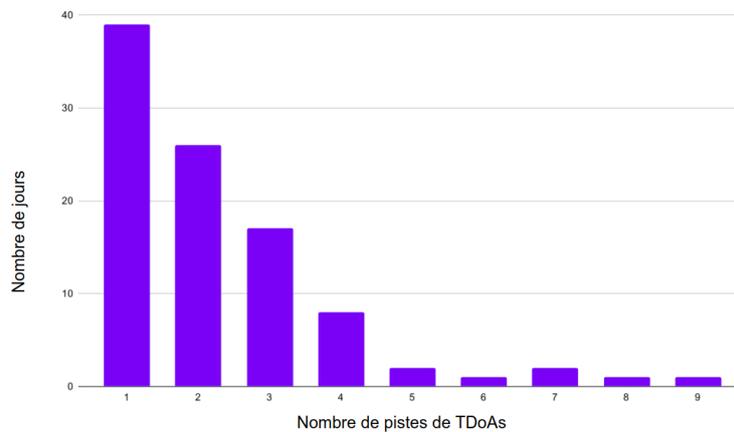


FIG. 4.13 – Histogramme du nombre de pistes par jour.

La figure 4.14 montre les pistes TDoAs théoriques que j'ai pu observer dans l'analyse des sessions. Cinq types de pistes ont été observées (a à e). Par exemple un Cachalot qui suit la piste "a" (en vert) va venir de l'ouest et aller à l'est.

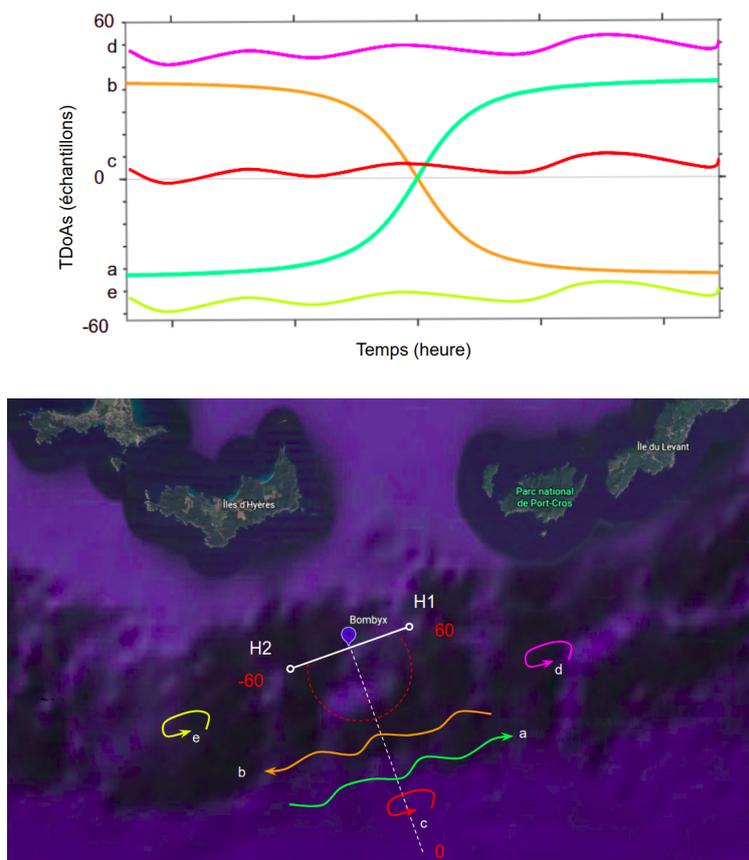


FIG. 4.14 – Haut : pistes possibles théoriques (a à e) en terme de temps de délais d'arrivée.
 Bas : pistes correspondantes projetées dans l'axe de la bouée.

Le récapitulatif du sens des animaux sur l'ensemble des sessions est présenté dans la figure 4.15. Lorsqu'il n'y a qu'une seule lettre (ex : a), un seul individu était présent. On peut voir sur ce graphique que la plupart du temps ce sont des individus seuls qui fréquentent la zone utilisant le sens a (ouest vers l'est), b (est vers l'ouest) et d (sens constant à l'est de la bouée). La lettre "n" signifie que le sens de déplacement n'a pas pu être déterminé.

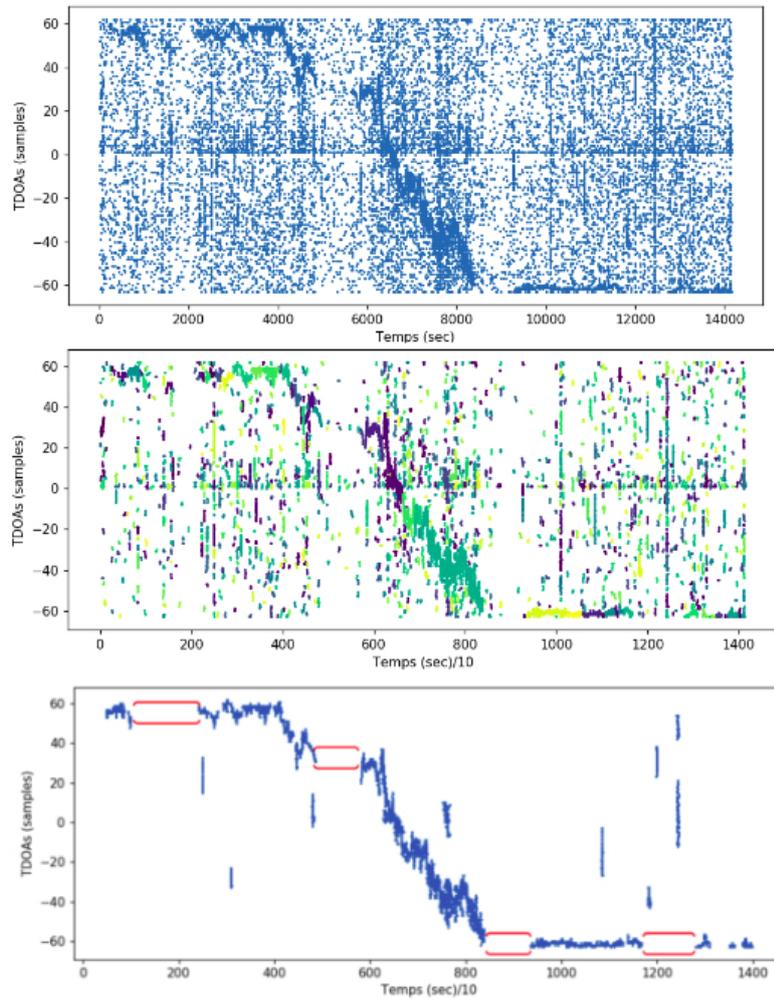


FIG. 4.16 – Haut : Affichage de la piste du Cachalot le 25 décembre 2018 en TDoAs. Milieu : Affichage de la piste du Cachalot, chaque couleur représente un cluster trouvé avec les paramètres suivants $eps=2$, échantillon minimum=3. Il a été nécessaire de dilater l'échelle temporelle afin que les clusters soient équilibrés sur les échelles des TDoAs et sur l'échelle temporelle. Bas : piste du Cachalot en TDoAs, en supprimant les points appartenant aux clusters avec une taille inférieure à 5 points. Chaque accolade rouge correspond au temps de surface ou l'animal ne produit pas de clic.

La figure 4.16 donne le résultat du clustering. La deuxième partie de cette figure représente pour chaque couleur un cluster trouvé par le HDBSCAN. La dernière partie montre la piste du Cachalot une fois que les clusters composés de moins de 5 points ont été retirés. La piste est beaucoup plus claire et il est donc possible d'estimer le nombre de plongées que l'animal a réalisé.

L'échantillonnage en continu nous permet de suivre l'animal pendant plus de 4h, et de voir que durant 4h il effectue 3 plongées et 4 temps de repos en surface ou l'animal n'émet pas de clics. Les 3 plongées durent en moyennes 40,5 minutes, et les temps de repos moyen est

de 19 minutes. Ce résultat est cohérent avec les études de Watkins et al. (1993) ; Jaquet et al. (2000). Ce clustering est donc très bien adapté pour l'isolation des clics, pour un enregistrement continu.

Le rapport Glotin et al. (2017) a présenté également des résultats intéressants concernant la présence de Cachalots sur cette bouée. Ils ont également construit un détecteur de clics (Ricard and Glotin, 2018), calculé les TDoAs, et relevé manuellement les pistes pour une petite période temporelle (du 21/08/2017 au 21/08/2017). Les étapes de leur détecteur sont, pour chaque fenêtre glissante de longueur 50ms (glissées de 2ms) les suivantes :

- filtrage passe bas (< 20 kHz) et passe haut (> 5 kHz).
- découpage du signal en 2 bandes de fréquence.
 - bande 1 = 5 kHz à 10 kHz
 - bande 2 = 10 kHz à 20 kHz
- dans chaque bande, mesure de la variation relative d'amplitude du signal temporel, (seuil à 0.2 en échelle log).
- regroupement des variations par bande, et décision finale par seuillage.

Il est donc possible, pour ces 10 jours de comparer leur méthode avec la notre. La figure 4.17 représente cette comparaison.

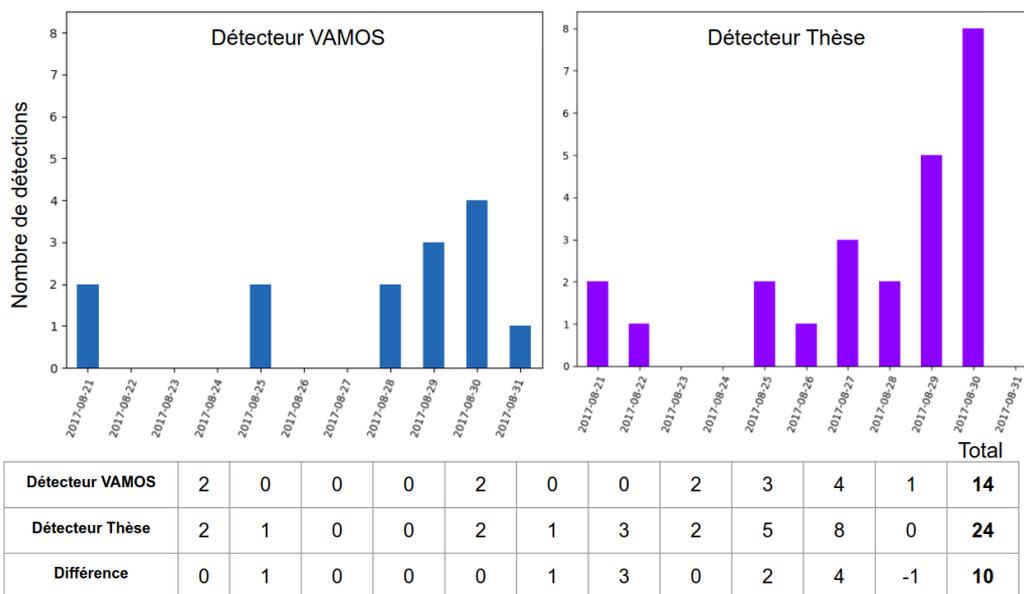


FIG. 4.17 – Comparaison des détecteurs : à gauche les résultats développés dans l'étude 'VAMOS' (Glotin et al., 2017), à droite ceux développés pendant la thèse, appliqués sur 10 jours de Juillet 2017.

Pour ces 10 jours, notre méthode aboutit à la détection de 24 pistes, tandis que le détecteur de VAMOS aboutit à 14. Nous pouvons alors calculer le taux de faux positifs et faux négatifs du détecteur de l'étude VAMOS en considérant la nouvelle expertise comme référence :

- $11/24 = 46\%$ de taux de faux négatifs
- $1/24 = 4,2\%$ de taux de faux positifs.

Le détecteur VAMOS aurait donc un faible rappel et une forte précision (taux de faux positifs très faible, taux de faux négatifs élevé).

Ceci illustre bien qu'afin de valider ou contre-valider la partie manuelle d'un détecteur, il faudrait la tester avec plusieurs annotateurs humains. Néanmoins nous démontrons ici un net progrès en rappel entre la première version du procédé d'analyse de Bombyx (étude VAMOS) et cette thèse.

De prochains travaux pourraient être également être réalisés, notamment dans l'entraînement de modèles d'apprentissages profonds (Deep learning) pour l'automatisation complète du processus de détection. Nous aimerions également extraire les IPI de chaque individu (Gordon, 1991) afin de pouvoir extrapoler les tailles de ces derniers automatiquement (Abeille et al., 2014).

Les résultats de la détection d'espèces pour la bioacoustique se sont articulés en trois différentes parties : la détection d'oiseaux, la détection d'Orques et la détection de Cachalots.

Ces trois cas d'étude nous ont montré que la détection d'espèces était un point clé pour des études en bioacoustique.

En effet, une fois l'espèce détectée il est possible de mettre en corrélation le nombre de détections avec des paramètres extérieurs comme la marée ou encore la saisonnalité.

4.2 Détection de population

4.2.1 Classification de clans d'Orques

La bioacoustique populationnelle se décline sur l'apprentissage d'un modèle d'intelligence artificielle pour la classification de clan d'Orques au Canada (méthodes décrites dans la partie 3.3). Pour rappel, nous disposons de 3 classes : le clan A, le clan G et le bruit. Après différents essais, le deuxième modèle avait un taux de bonne classification de 90% (tableau 3.4). Afin de valider ce dernier modèle, j'ai sélectionné 6 passages d'animaux devant Orcalab : 2 avec le clan A, 2 avec le G et 2 avec les deux clans voyageant ensemble.

Le modèle a produit une probabilité de présence des 3 classes (A, G et bruit) pour chaque segments de 5 secondes. Plusieurs stratégies étaient alors possibles : prendre le maximum, le 95^{ème} percentile ou la médiane des probabilités pour chaque enregistrement de 2 minutes.

La figure 4.18 représente l'évolution des prédictions de détections pour deux passages du **clan G**. On peut voir que pour le premier passage (gauche), les prédictions les plus fortes sont pour le clan A (courbe rouge), pour le maximum des prédictions (haut) et également pour le 95^{ème} percentiles (bas). Pour le deuxième exemple (droite), c'est bien le clan G qui a été détecté : la courbe bleue est bien supérieure que les deux autres courbes.

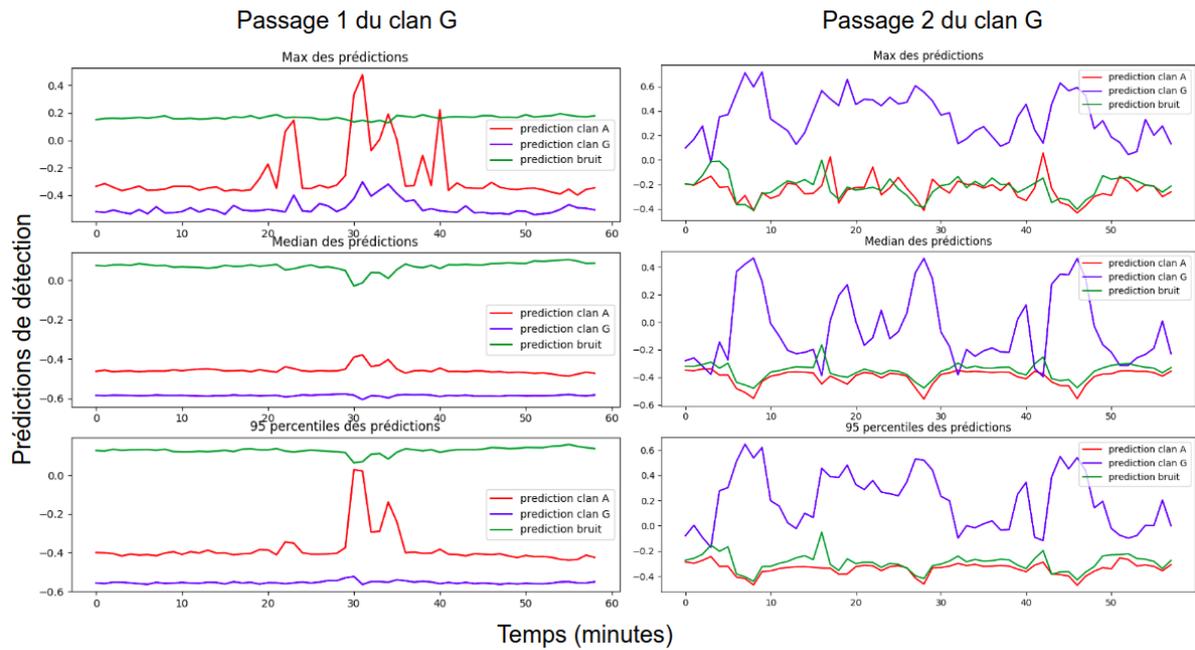


FIG. 4.18 – Évolution des prédictions de détections (maximum, médiane et 95 percentile) des 3 classes (A, G et bruit) pour deux passages du clan G.

Donc, pour deux passages différents du clan G, le modèle se trompe une fois (gauche sur la figure). Pour les deux autres exemples (A et A+G), le classifieur s'est trompé pour au moins un des deux passages.

En considérant uniquement le TBC du test, on peut supposer que le modèle est très performant (90%) mais lorsqu'on le teste sur des passages différents, le modèle n'arrive pas généraliser.

Les données d'entraînement et du test 1 proviennent d'enregistrements de 2017, et le test 2 était composé uniquement de données de 2019. Il est donc possible que durant l'entraînement, le modèle ait appris le bruit de fond, ou les caractéristiques de 2017, et donc s'est trompé sur 2019.

Un troisième entraînement a été réalisé, en ajoutant un troisième jeu de données afin que le modèle arrive à généraliser sur les caractéristiques des vocalises. Pour ce dernier, les enregistrements de la base de données Orhive ont été utilisés. Le TBC pour ce nouvel entraînement était de 68%.

Les trois derniers entraînements n'ont pas été concluants, ce qui nous a poussé à avoir une réflexion sur la qualité de notre jeu de données. Nous avons conclu qu'il existait probablement une trop grande variabilité au sein d'une même classe. En effet, le clan A et G peuvent produire environ 15 types d'appels différents. Il est possible que la variabilité au sein des clans soit plus importante que la variabilité entre les clans. C'est pour cette raison que le modèle aurait du mal à généraliser.

Afin de contrer cette grande variabilité acoustique au sein des clans, j'ai essayé de sélectionner un type d'appel par clan, représentatif de ce dernier, et d'entraîner le modèle à reconnaître uniquement ces types d'appels.

Pour le clan A, 1250 No4s ont été extraits, pour le clan G 160 N23s ont été extraits (jeu de donnée 4). Une troisième classe contenant tous les autres appels a été définie (384 exemples). La figure 4.19 représente la forme d'onde et le spectrogramme de ces deux types d'appels. Le choix de ces derniers s'est fait car ils sont très différents au niveau temps/fréquence, et qu'il sont produits en grande quantité pour chaque clan.

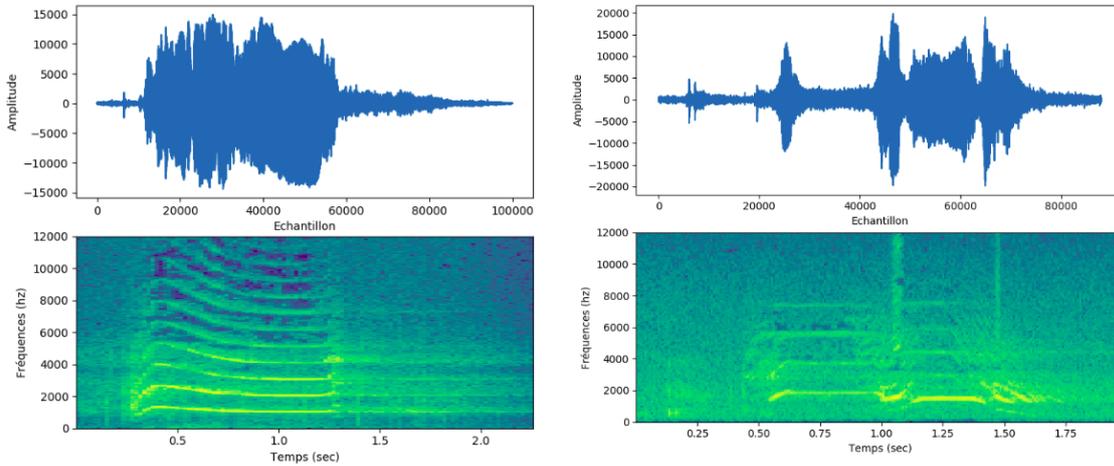


FIG. 4.19 – Forme d'onde et spectrogramme des deux types d'appels pour l'entraînement. Droite : No4 (clan A). Gauche N23 (clan G).

Le résultat de ce dernier entraînement est présenté dans le tableau 3.4. Le TBC de cet entraînement était de 51%. En analysant la matrice de confusion, la plupart des appels étaient classés dans le clan A. Cette erreur est probablement due au déséquilibre des classes durant l'entraînement (1250 No4s, 160 N23s).

Nous aurions pu équilibrer les classes en entraînant avec 160 exemples de chaque classe, mais ce type de modèle (CNN) nécessite un jeu de donnée d'entraînement assez conséquent et 160 éléments par classe n'est pas suffisant.

L'approche supervisée pour cette tâche de détection de population n'a donc pas été assez performante pour avoir des scores satisfaisants.

Pour cette échelle étude, aucun résultat en terme de bioacoustique ne sera présenté, mais les méthodes utilisées seront mises en comparaison dans la discussion.

4.3 Analyse de mouvements individuels

Comme décrites dans le chapitre Matériel, deux méthodes peuvent être mises en place pour l'identification individuelle sonore : l'identification grâce à une signature individuelle précise et reconnue et l'identification grâce à la localisation précise de l'individu. Cette dernière sera développée dans cette thèse.

L'identification individuelle est basée sur la localisation de la source grâce à plusieurs hydrophones. Je vais donc présenter deux cas d'étude sous-marins et un terrestre, relativement différents de par leurs émissions acoustiques (transitoires vs stationnaires) et leur comportement.

4.3.1 Localisation d'Orques pour l'identification individuelle sonore

Pour rappel, la localisation d'Orques pour l'identification des individus s'est faite grâce à une antenne composée de 4 hydrophones et des observations de surface depuis la station de recherche Orcalab (méthode présentée dans la partie 2.2.3.1).

Les sessions d'enregistrements ont eu lieu du 20 juillet au 1er août, du 4 août au 22 août et du 23 août au 15 septembre 2019. Au total, nous avons enregistré 42 passages vocaux qui comprenaient plusieurs Orques (Tableau. 4.1), et 29 passages sans vocalises.

TAB. 4.1 – Résumé des enregistrements pour toutes les sessions.

Session	Début	Fin	Cumul Temps h :min	Cumul jours	Passages vocaux	Passages non vocaux	Nbr Total des passages
1	20/07	01/08	17 :40	11	10	9	19
2	04/08	22/08	17 :30	17	12	6	18
3	23/08	15/08	17 :40	11	19	15	34
Total			52 :50	39	41	30	71

Premièrement, nous avons étudié l'activité acoustique globale des Orques en fonction de la marée. Le Blackney Pass (BP) en face d'Orcalab est étroit, il y a donc des courants forts lors des changements de marée.

La figure 4.20 montre l'évolution des passages d'Orques en fonction de la marée ; pour les groupes venant du sud (de l'île Flower à l'île Parson) par rapport au Nord (de l'île Parson à l'île Flower) (voir carte 2.15).

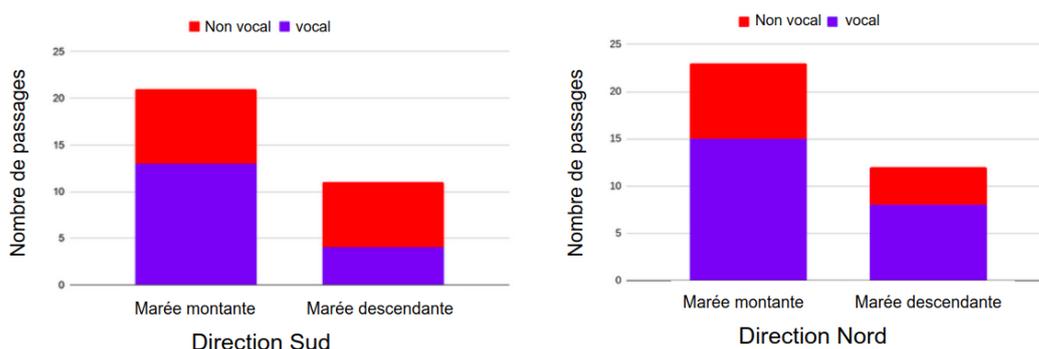


FIG. 4.20 – Résumé des passages en fonction des marées lorsque les animaux se dirigent vers le Sud et vers le Nord.

La figure 4.20 montre que les Orques viennent à travers BP lors de la marée montante (21 vs 11), quelle que soit la direction des groupes (23 vs 12). Lorsque les Orques vont vers le Sud

lors de la marée montante (nage avec le courant), il y a plus de passages avec vocalisations (13 vs 8), que lorsqu'elles vont à contre-courant.

Quand les Orques vont vers le Nord, il y a plus de passages vocaux à marée descendante, donc quand elles nagent avec le courant. En résumé les Orques sont plus vocales lorsqu'elles nagent avec le courant. Nous constatons également que les Orques se déplacent plus souvent pendant la marée montante (le double que pendant la marée descendante).

Dans les 71 passages, 57 étaient pendant la journée et 29 incluaient des vocalisations. Nous avons donc été en mesure d'isoler les vocalises de trois individus (trois mâles), car les mâles ont été à l'écart de leur groupe (minimum 50 m). Au total, 85 vocalisations ont été produites par ces trois mâles. L'ensemble des vocalises isolées est décrit dans le tableau 4.2.

TAB. 4.2 – Résumé des 3 individus où nous avons pu isoler des vocalises. Le mâle A66 est apparu plusieurs fois éloigné de son groupe.

Date	Pod Id.	Matriline Id.	Individus Id.	Nombre de vocalises (85 au Total)
04/09	A05	A42	A66	16
31/08	I11	I4	I76	24
22/07	A05	A23	A60	19
30/08	A05	A42	A66	26

Nous nous sommes concentrés sur trois événements particuliers : 30, 31 et 11 août. Nous listons sur la figure 4.21 toutes les vocalisations que nous avons détecté et annoté. La couleur de chaque appel correspond à la couleur du nuage de points des figures 4.23, 4.28, 4.34.

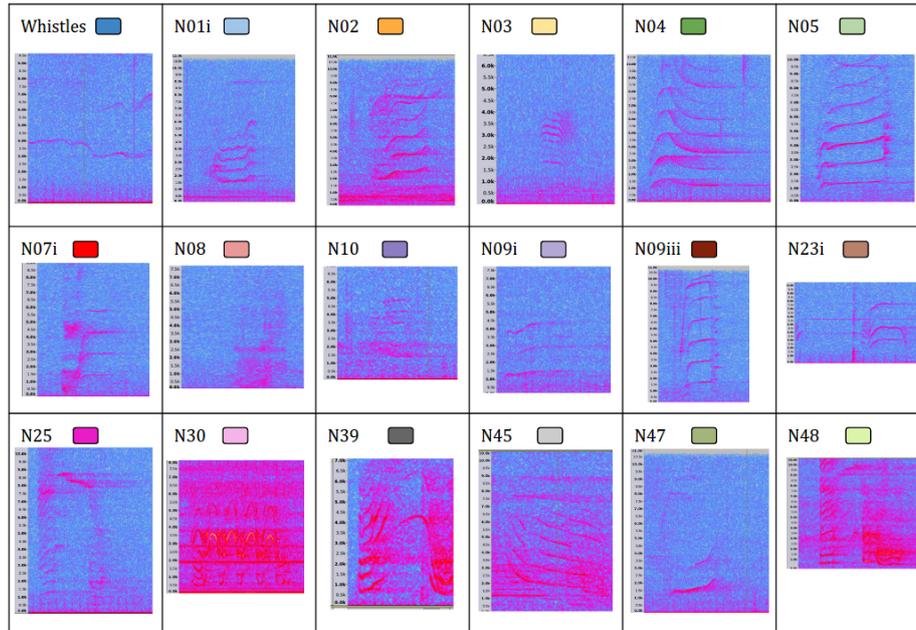


FIG. 4.21 – Liste des vocalises trouvées et annotées dans les enregistrements (Ford et al., 1987).

4.3.1.1 Premier cas d'étude (30 août)

Le premier cas d'étude est le 30 août, dans lequel différents groupes sont passés à côté de l'antenne. Le premier groupe n'était composé que de deux individus : A66 et un autre individu. Le deuxième groupe était le pod A05 (matriline A42 et A25) et le pod A01 (matriline A54). La figure 4.22 représente le passage du premier groupe suivi par le deuxième groupe. Chaque point représente un appel placé en azimut ; les courbes bleues et oranges sont la trace visuelle de chaque groupe. Chaque observation visuelle (courbes oranges et bleues) correspond à l'azimut calculé de chaque appel. Ainsi, nos relevés visuels valident nos estimations d'azimuts.

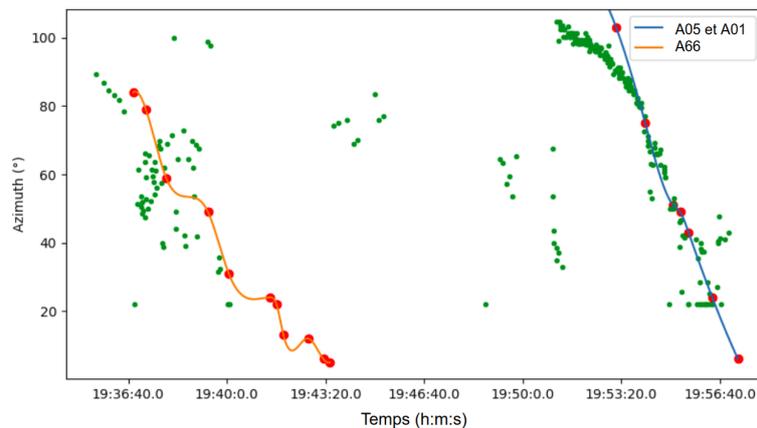


FIG. 4.22 – Graphique du passage du 30 août. Chaque point vert représente une vocalisation placée en azimut. La courbe orange est la position interpolée du mâle A66 détectée visuellement et la courbe bleue est le groupe 1.

Pour chaque appel, nous avons extrait le type d'appel et nous l'avons affiché en fonction de l'azimut des Orques (Fig.4.23).

La première piste est composée des appels N47, No4, No8i, No7i et No9i. Pendant ce premier passage, nous avons vu le mâle A66 et un autre individu. Le mâle A66 provient du pod A05, ce qui signifie qu'il pourrait émettre les appels No4, No8i et No7i (Ford, 1991). Nous savons que le deuxième individu provient du pod A01 parce qu'il a produit un appel N47. Cette première trace est notable car les Orques ne suivent pas une direction précise et semblent visuellement se nourrir. La deuxième piste est compacte, et le groupe 1 (A42, A54 et A25) a voyagé très rapidement vers le Nord. Ce groupe était composé de deux pods différents (A05 et A01).

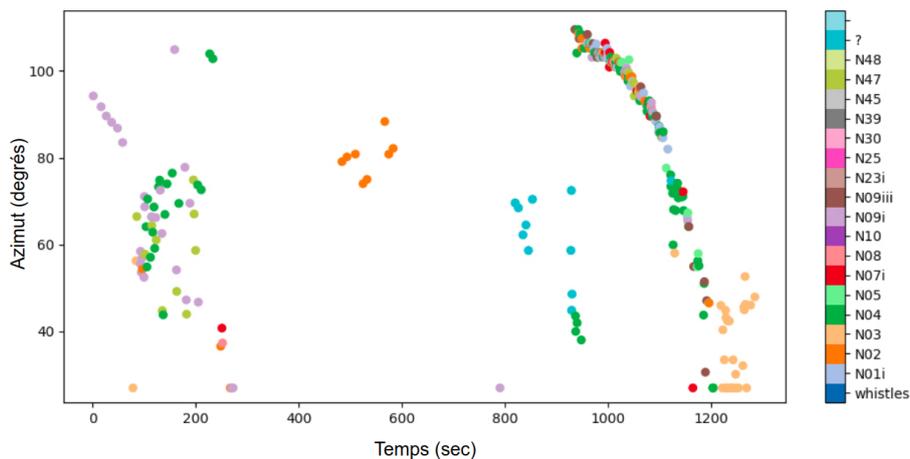


FIG. 4.23 – Graphique du passage du 30 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel (voir figure 4.21).

Dans la deuxième piste, il semblait y avoir beaucoup d'appels différents, mais la fin de la trajectoire se termine par des appels No3s (couleur beige). Ce dernier pourrait jouer un rôle dans la communication entre les matrilines, car à la fin de la piste, les individus n'ont émis que cet appel. Cela peut indiquer un changement de comportement particulier.

En se concentrant sur la dernière partie de la piste, nous voyons que les A01 et A05 ont voyagé ensemble et nous pouvons attribuer des appels à un pod précis. La figure 4.24 montre la trajectoire des A01, A05 et de leurs appels émis. Nous pouvons couper la piste en six parties. Tout d'abord, le pod A01 émet des types No2 et No4, tandis que le pod A05 ne produit pratiquement aucun appel. Lors de la deuxième partie, les deux pods émettent des vocalisations. Le pod A01 émet une succession de No9i et No4 tandis que le pod A05 émet No4, No9iii et No2.

Après cela, dans la troisième partie, les A01 n'ont émis que des N47 et N01i, et les A05 ont répondu avec des No4, No7i et No9iii. Au cours de la quatrième partie, il y a eu beaucoup d'appels produits par A01 et presque aucun par A05. C'est l'inverse pour la cinquième partie. Au cours de la sixième et dernière partie, il y a eu de nombreux appels A05. Nous pouvons suggérer que la succession de vocalisations n'est pas due au hasard et est probablement volontaire.

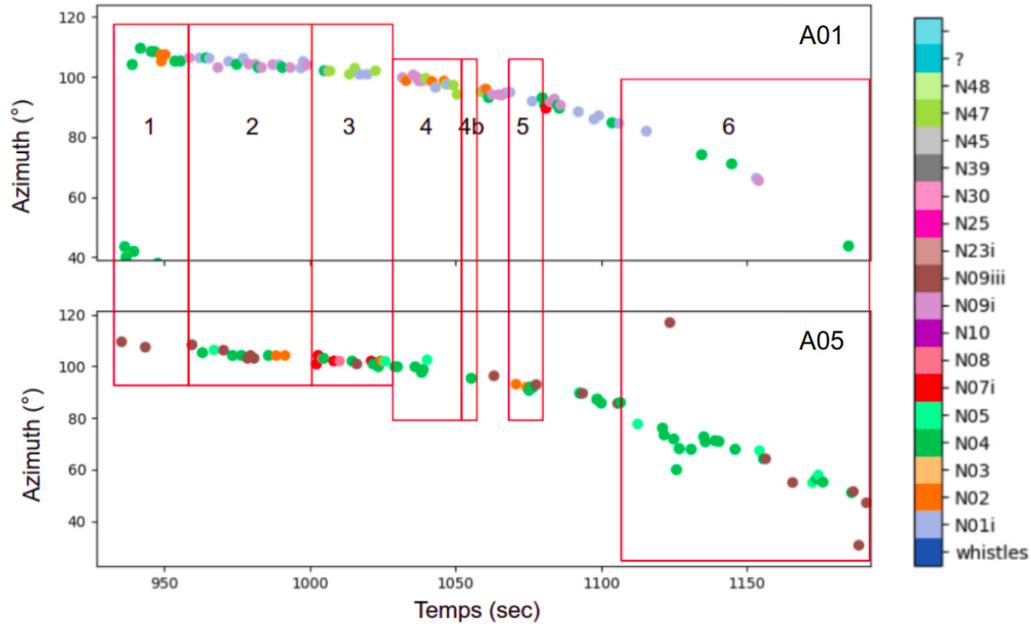


FIG. 4.24 – Graphique des passages de la dernière partie du 30 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel. Haut : Trajectoire du pod A01, Bas : Trajectoire du pod A05.

Nous avons mesuré le taux d'appel (call Rate= CR) pour les A01 et les A05, puis nous avons calculé le KDE (Estimation par noyau) du taux d'appel sur la position de chaque appel. Ces résultats sont montrés en figure 4.25. On peut voir que les CR pour chaque pod sont entrelacés pendant la période 'a' la densité des appels pour A01 augmente, tandis que celui des A05 diminue. C'est l'inverse en période 'b'. Ce résultat suggère que lorsque ces 2 pods voyagent ensemble, les A05 émettent tandis que les A01 sont silencieux (durant quelques secondes), et vice versa. Par conséquent, ils ne chevauchent pas leurs appels, ce qui permet de maximiser leur efficacité de communication.

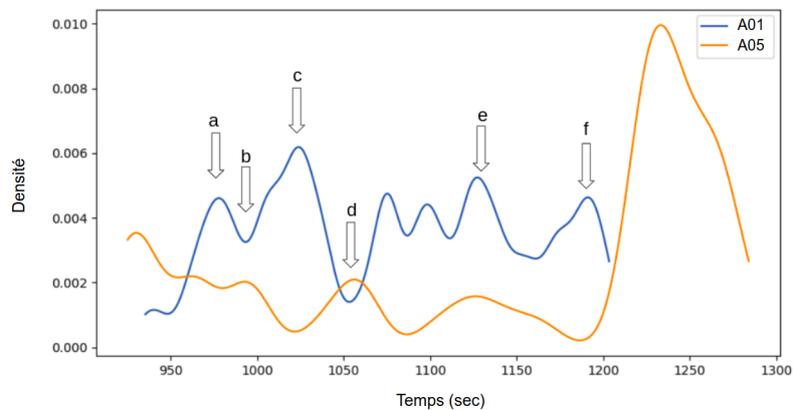


FIG. 4.25 – Densité du CR dans le temps, pour le 30 d'août.

Pour différents appels, nous avons extrait les 89 MFCC et les avons redimensionné en deux dimensions avec la t-SNE. Le résultat des appels de type No4 est illustré dans la figure 4.26, où chaque couleur représente un No4 provenant d'un pod identifié visuellement.

On voit que les appels No4 émis par les A05 sont différents des No4 émis par les A01. Suite à la t-SNE nous avons appliqué un clustering (BNP) afin d'attribuer des clusters à cette distribution (forme sur la figure). Le BNP clustering a trouvé 5 clusters.

L'information mutuelle normalisée (NMI) entre les pods (en couleur) et les clusters est égale à 0,20. Cela montre une corrélation non linéaire entre ces deux distributions. Ainsi, lorsque ces Orques sont regroupées, les appels de type No4 émis des pods A05 et A01 sont différents et représentent la moitié des appels produits. Ces caractéristiques du type d'appel No4 avaient déjà été mesurées, mais manuellement dans Ford et al. (1987).

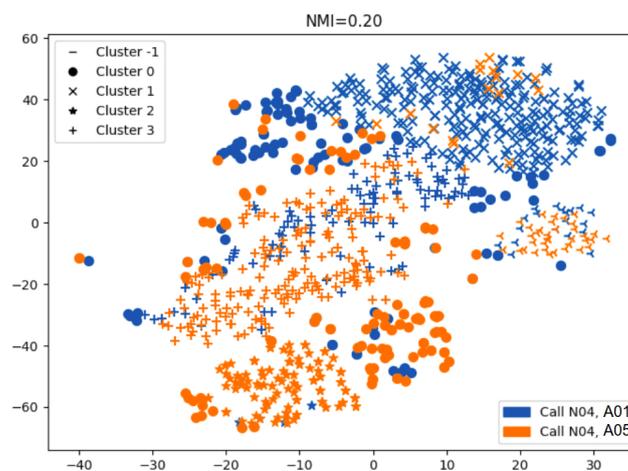


FIG. 4.26 – Représentation de la t-SNE des MFCC (29 + 30 + 30) pour les No4s pour le 30 août.

4.3.1.2 Deuxième cas d'étude (31 août)

Nous nous concentrons ici sur le 31 août, lorsque le groupe 1 composé des matriline I16, I65, I04 et I27 est passé par le BP. Pendant ce passage, le groupe 1 composé de ces quatre matriline voyageait ensemble ; le mâle I76, voyageait seul derrière le groupe, pour les rejoindre à la fin du passage.

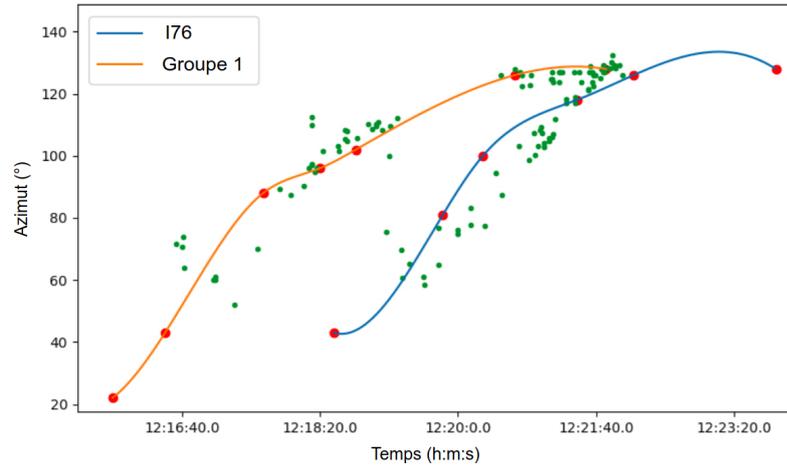


FIG. 4.27 – Graphique du passage du 31 août. Chaque point vert représente une vocalisation placée en azimut. La courbe bleue est la position d’interpolation du mâle I76 détectée visuellement, la courbe orange est le groupe 1.

La figure 4.27 représente ce passage. Chaque point représente une vocalisation placée par l’azimut estimé. La courbe bleu et orange est l’azimut visuel interpolé du mâle I76 et du groupe 1, respectivement. On voit qu’il est possible d’affecter des vocalisations soit au mâle soit au groupe 1 car chaque position acoustique correspond à une position visuelle. Ce résultat nous permet d’isoler 24 appels du mâle I76. A partir des annotations expertes manuelles, nous avons fait la corrélation entre l’azimut et le type d’appel pour chaque groupe (groupe 1 et I76).

Pour ces enregistrements, sept vocalises différentes ont été produites (voir figure 4.28). Chaque point représente un appel avec son type en couleur. Nous voyons que la piste de droite (I76) est composée de des types N48 (rose), N25 (vert) et N23i (marron).

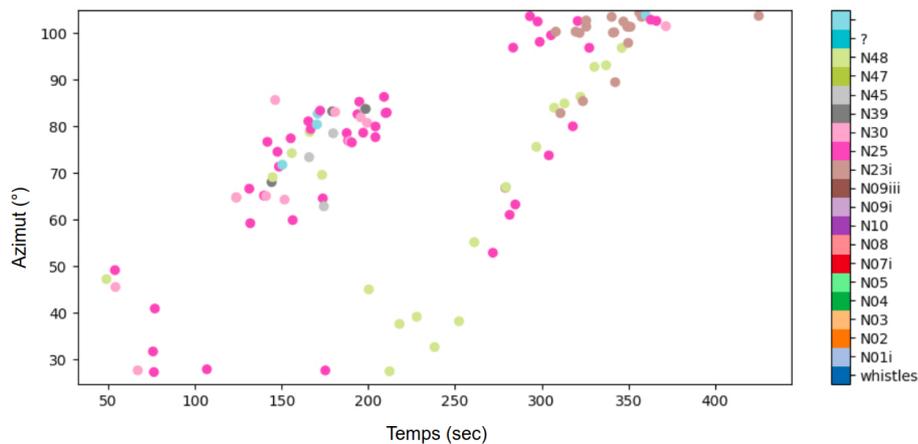


FIG. 4.28 – Graphique du passage du 31 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d’appel. Les types les plus courants sont N25, N48 et N23i.

Ainsi, nous pouvons supposer que lorsque le mâle était seul, il a fait une combinaison de deux appels : N48 et N25, et lorsqu'il était assez proche de son groupe, il produit des appels N23i auxquels son groupe lui a répondu avec le même type. Le groupe 1 a produit tous les types d'appels, y compris le type N23i (brun) à la fin de la piste. Nous pouvons voir que l'appel de type N23i est produit juste avant et après la rencontre des deux groupes.

Puis nous nous sommes concentrés sur les appels de type N48 et N25, car ils sont produits par le groupe 1 et le mâle I76. La figure 4.29 montre la projection t-SNE de tous les appels N48 et N25. Les N48 du groupe 1 (bleu) sont regroupés, de sorte que les N48 de ce groupe ont globalement la même forme. Les N48 des mâles sont différents. Il existe deux types d'appels N48 émis par le mâle I76.

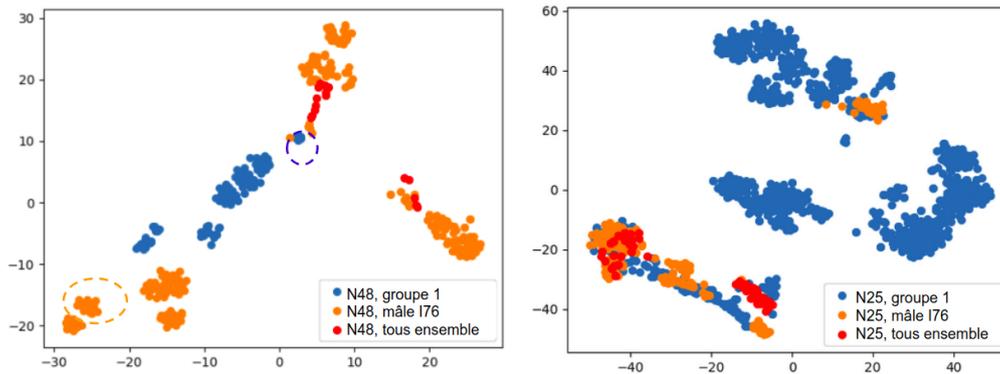


FIG. 4.29 – t-SNE des MFCC pour les appels N48 (à gauche) et N25 (à droite) le 31 août 2019. Les points oranges sont produits par le mâle I76 tandis que les points bleus sont produits par le groupe 1. Les points encadrés sont les vocalisations représentées Fig.4.30.

La figure 4.30 représente un exemple de deux N48 : un produit par le mâle I76 et un par le groupe 1. Ces deux appels sont très éloignés dans la t-SNE (cercles, Fig 4.29). Ce spectrogramme montre qu'il y a beaucoup d'informations intégrées à l'intérieur de chaque N48. Il est donc possible que les individus arrivent à s'identifier à l'aide d'informations fréquentielles et temporelles verrouillées à l'intérieur des N48. Cette information dans chaque type d'appels pourrait donc être utilisée pour la reconnaissance de chaque individu.

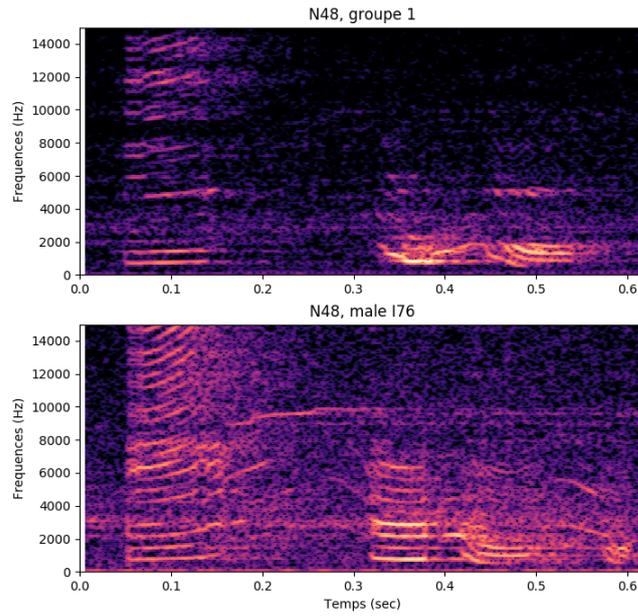


FIG. 4.30 – Exemple d'un spectrogramme de N48 du groupe 1 et du mâle I76.

Les N48 qui sont émis lorsque les deux groupes sont réunis (rouge) dans le cluster orange (appel du mâle), ce qui nous fait penser que ces vocalisations appartiennent au mâle I76.

La deuxième partie Fig.4.29 présente la projection t-SNE des appels N25. Les N25 émis par le mâle sont similaires à une partie des autres N25 émis par le groupe (cluster en bas à gauche). Les N25 du groupe 1 sont regroupés en quatre groupes différents (bleu). Nous pouvons conclure qu'il existe des variations dans les appels N25 du groupe 1 (4 grappes), mais pas dans les vocalisations du mâle I76 (1 cluster).

Nous avons mesuré le CR pour le groupe 1 et le mâle I76, afin de voir s'il y a différentes périodes d'échanges d'appels. Nous avons calculé le KDE du CR sur la position de chaque vocalisation (Fig.4.31). La courbe bleue représente la densité des appels pour le groupe 1, et la courbe orange représente la densité des appels pour le mâle I76. Nous pouvons constater quatre périodes.

Dans la première période (1), seul le groupe 1 est acoustiquement actif (bleu), dans la deuxième période (2), il n'y a que des appels du mâle. Dans la troisième période (juste avant le regroupement), il y a des appels des deux groupes, et de même dans la dernière période (4), lorsque les deux groupes sont ensemble.

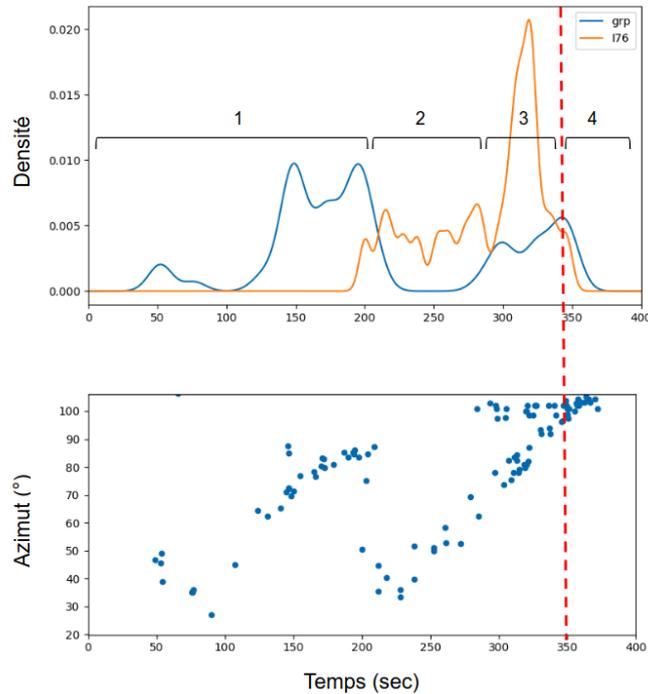


FIG. 4.31 – Haut : Fonction densité proportionnelle au CR dans le temps. Bleu : la densité des appels du groupe 1. Orange : la densité des appels du mâle I76. Bas : Graphique du passage du 31 août. La ligne rouge représente le regroupement des individus.

Nous avons observé que lorsque le mâle commence à appeler (période 2), le groupe 1 n'a pas émis de vocalise. Pendant la deuxième période, le mâle produit beaucoup de vocalisations, et pendant la troisième période, les deux groupes sont vocaux.

4.3.1.3 Troisième cas d'étude (11 août)

La dernière expérience naturelle a eu lieu le 11 août, lorsqu'un grand groupe composé des martilines A25, A50, A54 et I04 (mat.) est passé devant l'antenne à 16h36, représentant respectivement les pods A05, A01, I11. Les Orques étaient calmes et se déplaçaient à une vitesse de nage modérée, mais à 17h08 et 20sec, un bateau de croisière (figure 4.32) est arrivé (se déplaçant dans la même direction que les Orques).



FIG. 4.32 – Photographie prise pendant le passage du bateau de croisière le 11 août.

La présence du bateau a provoqué la dispersion du groupe et plusieurs Orques ont surfé dans son sillage, ce qui a rendu l'identification visuelle individuelle difficile.

La première partie (jusqu'à 17h15) nous a permis de reconnaître huit individus des pods A05 et A01 (Fig.4.33). Pendant le passage du bateau, les Orques se sont séparées donc nous avons perdu les identifications de quatre baleines. De 17h16 à 17h30, nous avons identifié A61 (matriline A25), A72 (mat. A50), A108 (mat. A50) et I76 (mat. I04). Les observations visuelles ont été complétées par des analyses acoustiques. De 16h56 à 17h10, les Orques voyageaient à une vitesse lente. Lorsque le bateau de croisière est arrivé, les baleines ont commencé à accélérer et à suivre le bateau jusqu'à l'azimut 100° .

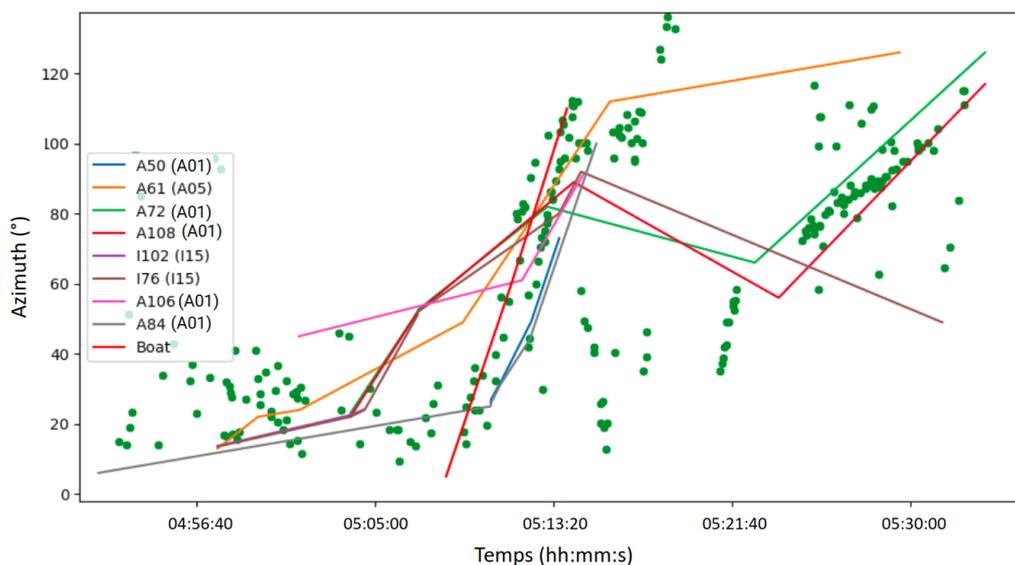


FIG. 4.33 – Graphique du passage du 11 août. Chaque point vert représente une vocalisation placée en azimut. Les différentes courbes sont les positions de chaque individu détecté visuellement.

Nous pouvons observer la trajectoire rapide visuellement et acoustiquement de 16h06 à 17h15. Au cours de la dernière partie de la piste, de 17h23 à 17h33, nous voyons que certaines Orques

ont poursuivi leur trajectoire, tandis que quatre individus (de trois matriline différentes) ont fait demi-tour : A61 (mat. A25s), A72 (mat. A50s), A108 (mat. A50s) et I76 (mat. I04s). Le mâle A61 (orange) est seul, mais les individus A72, A108 et I76 ont voyagé ensemble pendant dix minutes.

Nous avons pu extraire les types de vocalisations (Fig.4.34) que nous voyons dans cette piste en trois parties.

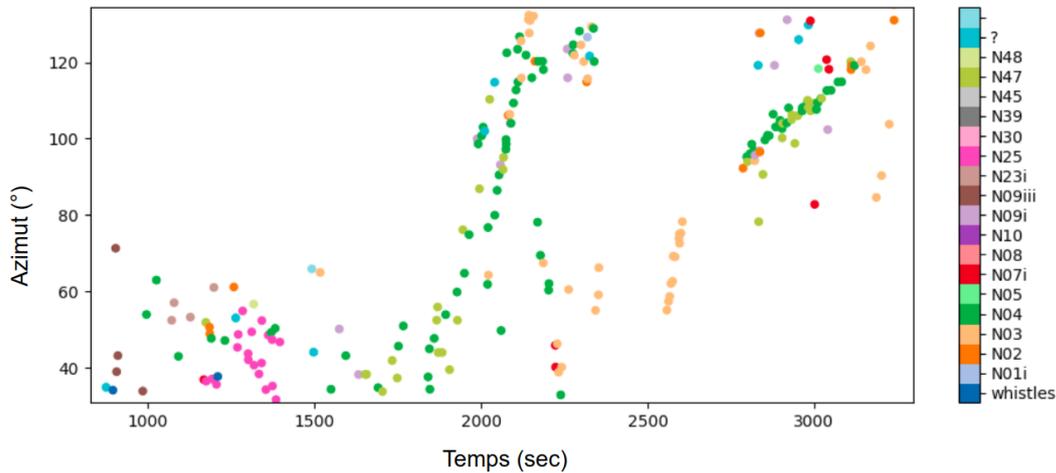


FIG. 4.34 – Graphique du passage du 11 août. Chaque point représente une vocalisation placée en azimuth. Chaque couleur représente un type d’appel. Les types les plus courants sont N04, N03, N48.

La première partie est avant l’arrivée du navire de croisière, lorsque les Orques voyagent lentement et émettent des types d’appel N25, N04, N02, N09iii. La deuxième partie est lorsque le bateau est arrivé et que les individus ont suivi le bateau (surf), en produisant uniquement des appels N04, N47 (de 1700 à 2300 secondes).

Après avoir surfé (troisième partie), elles produisent beaucoup d’appels N03, et certaines ont fait demi-tour. Pendant la troisième partie, de 2700 à 3200 secondes, la piste de 90° à 120° est composée d’appels N04, N47. D’après ces observations, nous pouvons supposer que le bateau a non seulement eu une influence sur la vitesse des Orques, mais aussi sur les types d’appels utilisés.

Dans le prolongement de ce raisonnement, nous avons compté le nombre d’appels (CR) et appliqué le KDE afin de voir si le bateau avait une influence sur le taux d’appel. La figure 4.35 montre le débit d’appel (KDE, densité d’appels) en fonction de la trajectoire des Orques. Pendant la première période (avant l’arrivée du bateau), la densité des appels est faible. Lorsque le bateau arrive (ligne rouge), la densité des appels a commencé à augmenter jusqu’à ce que le bateau s’éloigne. Puis le nombre de vocalise a beaucoup diminué juste après le passage du bateau.

Ensuite, la densité a de nouveau augmenté lorsque A108, I72 et I76 ont voyagé ensemble. Cela montre que le navire de croisière a eu une influence sur le type d’appels ainsi que sur le taux d’appel.

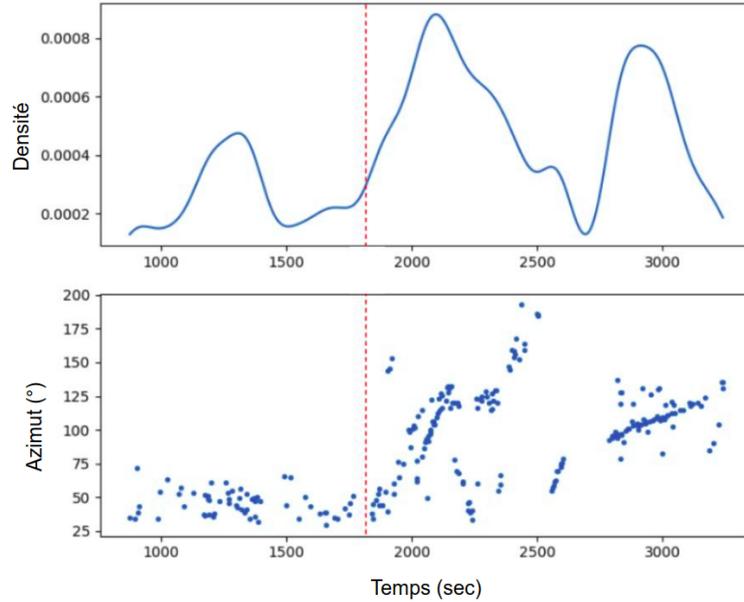


FIG. 4.35 – Haut : Fonction densité proportionnelle au CR dans le temps. Bas : Graphique du passage du 11 août. La ligne rouge représente l'arrivée du bateau.

Cette expérience nous a permis d'isoler plus de 85 vocalises appartenant à 3 mâles différents, mais elle nous a surtout permis de mettre en évidence que la communication chez ces Orques peut dépendre de :

- la distance entre les individus
- des émissions acoustiques des autres pods
- la présence de bateaux
- du comportement des animaux.

Le tableau 4.3 présente les différentes tendances bioacoustiques de ce cas d'étude et les compare avec la bibliographie de cette population d'Orques et des autres populations du monde. Il est important de signaler que ces résultats ne sont que des "tendances" car elles ont été constatées sur des événements rares ($N < 3$). Nous n'avons donc pas pu réaliser de tests statistiques robustes pour valider ces tendances.

TAB. 4.3 – Résumé de nos résultats préliminaires concernant la bioacoustique individuelle des Orques de Colombie-Britannique en fonction de la bibliographie actuelle (en excluant les études avec des animaux captifs).

Expérience naturelle	Résultats bioacoustiques	NRKW Bibliographie	Bibliographie des autres Orques
Tout	Isolation et identification de 85 vocalises provenant de 3 mâles	Cette étude	
30 août			
	Isolation de 26 vocalises du mâle A66	Cette étude	
	No3 utilisé pour la communication intra-groupe	Ford (1991),	
	Les successions des vocalisations sont dépendantes les unes des autres	Miller et al. (2004) Weiß et al. (2007), Cette étude	Filatova et al. (2013) Karenina et al. (2013); Vester (2017)
	No4 sont séparés en sous-types	Ford (1989); Deecke et al. (2010) Deeke (1998) Grebner (2009), Cette étude	Danishevskaya et al. (2020) Moore et al. (1988); Vester (2017)
31 août			
	Isolation de 24 appels du mâle I76	Cette étude	
	N48, N25 et N23 sont produits pour la cohésion de groupe	Cette étude	
	N48 et N25 des mâles sont différents (signatures individuelles ?)	Cette étude	
	Le taux d'appel change en fonction de la distance entre les groupes	Grebner (2009), Cette étude	Filatova et al. (2013)
11 août			
	La trajectoire des Orques change à cause du ferry , séparation des individus	Williams et al. (2002, 2014), Cette étude	Karenina et al. (2013)
	Variation du taux d'appel à cause du bateau	Talus (2000), Cette étude	
	N48, No4s produits pendant le "surf"	Cette étude	
	No3 produit après le passage du bateau	Cette étude	

L'ensemble de cette étude est en soumission pour la revue Scientific report (Poupard et al., 2020)

4.3.2 Localisation de Cachalot depuis le drone *Sphyrna*

La partie 3.4.2 dans le chapitre Méthode a développé les calculs réalisés afin d'extraire la trajectographie du Cachalot en trois dimensions.

À partir des détections de clics localisés restantes, nous calculons une trajectoire dans le temps, illustrée Fig.4.36.

Après sa première plongée, l'animal passe 15 minutes à la surface (silence). La deuxième plongée commence directement avec des clics à la surface et dure 32min, la profondeur maximale est de -1,2km.

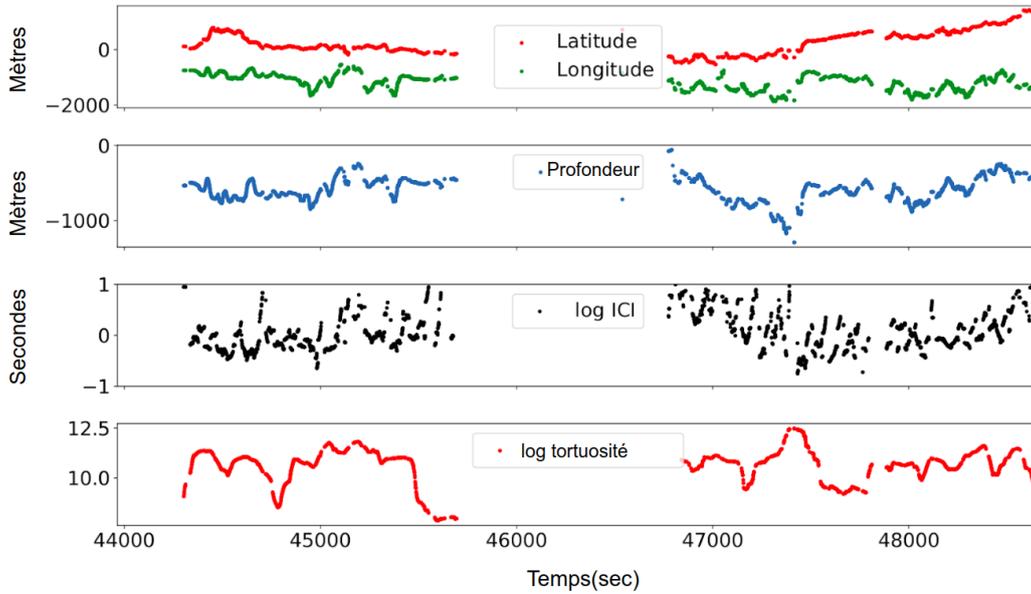


FIG. 4.36 – De haut en bas : position du Cachalot, latitude X (Nord), longitude Y (Est), profondeur Z (m), log (ICI) (s), log (tortuosité) dans le temps.

Enfin, afin de démontrer quels types d'informations peuvent être acquis à partir d'un suivi 3D détaillé, nous pouvons comparer l'ICI (t), la profondeur (m) avec la tortuosité T , qui est un indice du comportement du mouvement de l'animal (Wilson et al., 2007 ; Almeida et al., 2010). La tortuosité est la somme des variances (V) des trois dimensions (X, Y et Z) du déplacement dans une fenêtre de temps (ici 300s) :

$$T = V(X_1, \dots, X_n) + V(Y_1, \dots, Y_n) + V(Z_1, \dots, Z_n) \quad (4.1)$$

La figure 4.36 montre que lorsque la tortuosité (le déplacement) de l'animal est importante, la profondeur est maximale.

Cette étude a permis de localiser précisément un individu à une distance de plus de 2km de l'antenne avec un réseau d'hydrophones mobiles. Cette localisation a été possible grâce à la stabilité du drone et la haute fréquence d'échantillonnage de la carte JASON, permettant ainsi le calcul des TDoAS de façon très précise.

Nous prévoyons de réaliser cette mission avec plusieurs drones de surface afin de valider les position X,Y et Z de l'animale. Le placement des drones se ferait automatiquement à l'aide de systèmes d'évitement comme décrits dans Soriano et al. (2016). L'ensemble de cette étude est décrite dans l'article Poupard et al. (2019d).

Cette partie a présenté les différents résultats en bioacoustique sur l'échelle individuelle.

La localisation nous a permis d'isoler des signaux acoustiques, et donc des individus. Pour le Cachalot, ces signaux individuels ont été corrélés avec des paramètres extérieurs comme la tortuosité ou encore la profondeur. On a donc confirmé que les émissions acoustiques étaient bien liées à un comportement de chasse.

Puis, pour la détection individuelle d'Orques à Hanson isl., différents paramètres extérieurs ont influé les émissions acoustiques des individus comme la présence d'un ferry ou encore le

comportement des individus.

4.3.3 Identification individuelle aviaire

La dernière application en bioacoustique individuelle s'applique à un cas d'étude terrestre : la localisation d'oiseaux. L'ensemble des méthodes utilisées est présenté dans la partie 3.4.3. Une fois que nous avons sélectionné les meilleurs paramètres pour l'extracteur d'évènements acoustiques. Les 13 enregistrements annotés ont été analysés par le détecteur. Un exemple de 6 secondes est présenté en figure 4.37.

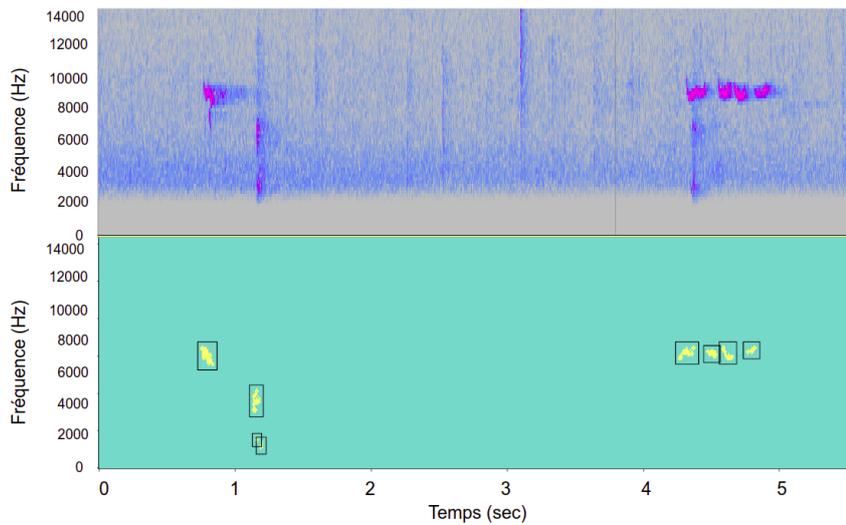


FIG. 4.37 – Spectrogramme de 6 secondes contenant deux individus (Haut) associé au détecteur automatique d'évènements acoustiques (Bas).

Les TDoAs ont été calculés pour chacune des détections sur les 4 voies et pour les 14 enregistrements annotés. Puis nous avons récupéré seulement les 3 TDoAs indépendants ($\tau_{21}, \tau_{31}, \tau_{41}$) et les afficher dans un graphique 3D. Puis nous avons essayé d'attribuer des clusters pour chaque point dans cette représentation à l'aide du clustering DBSCAN. Sur cette représentation, chaque point représente donc un chant positionné en TDoAs. La figure 4.38 représente les trois TDoAs indépendants pour deux enregistrements contenant deux individus. Sur la gauche l'algorithme a bien trouvé deux clusters (correspondant à deux oiseaux). Tandis que sur la droite l'algorithme a défini 3 clusters alors qu'il n'y avait que deux oiseaux dans l'enregistrement.

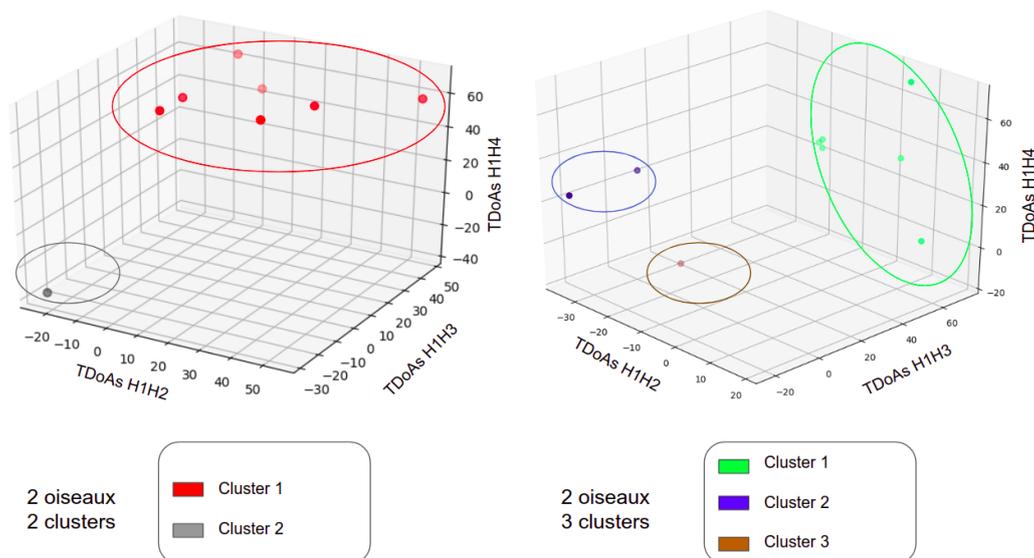


FIG. 4.38 – Représentation 3D des trois TDoAs indépendants pour deux exemples contenant deux individus. Chaque couleur correspond au clusters attribués par le DBSCAN.

Le clustering **DBSCAN** comporte différents paramètres dont *Epsilon* caractérisé par la distance euclidienne maximum entre 2 points. L'optimisation de ce paramètre a été réalisé afin de minimiser l'erreur de notre modèle. Le tableau 4.4 présente les meilleurs résultats obtenus. Sur 14 enregistrements, le modèle s'est trompé 7 fois (50 % d'erreurs). Nous pouvons également calculer le taux de faux positifs (rappel) et le taux de faux négatifs (précision) :

- $7/12 = 58\%$ de taux de faux positifs
- $3/14 = 21\%$ de taux de faux négatifs.

L'ensemble des résultats est présenté dans notre rapport de recherche Poupard et al. (2018e).

TAB. 4.4 – Tableau récapitulatif des prédictions du modèle en fonction des annotations de référence ('vérité') faites par T. Lengagne.

Vérité	Prédiction	Différence
0 individu (A)	0 individu	0
0 individu (B)	0 individu	0
1 individu (C)	3 individus	2
1 individu (D)	1 individu	0
1 individu (E)	2 individus	1
2 individus (F)	2 individus	0
2 individus (G)	2 individus	0
2 individus (H)	2 individus	0
2 individus (I)	5 individus	3
2 individus (J)	3 individus	1
3 individus (K)	2 individus	-1
3 individus (L)	3 individus	0
4 individus (M)	2 individus	-2
4 individus (N)	3 individus	-1

Malgré différentes méthodes utilisées, je n'ai pas réussi à compter précisément le nombre d'oiseaux enregistrés de manière automatique. Les différents TDoAs des détections ne correspondaient pas aux nombres de sources dans l'enregistrement. Nous n'avons donc pas appliqué cette chaîne de méthodes (détection, TDoAs, clustering) sur d'avantages d'enregistrement pour réaliser d'autres études bioacoustiques à plus large échelle.

Nous discuterons dans le chapitre suivant les raisons pour laquelle nous n'avons pas réussi à localiser individuellement les oiseaux dans ces enregistrements.

Ces différentes études (Orques, Cachalots et oiseaux) nous ont donc montré que la bioacoustique serait applicable à l'échelle **individuelle**.

4.4 Résultats en bioacoustique comportementale

Cette partie présente les différents résultats de la bioacoustique dite comportementale. Premièrement l'impact du whale watching sur les dauphins tacheté pantropical est présenté, puis l'influence des stimuli chimiques sur les Baleines à bosse est exposée.

4.4.1 Impact du trafic maritime sur le Dauphin tacheté pantropical

Notre algorithme de détection automatique des sifflements a détecté 2502 sifflements avec une taille de fenêtre de 256, et 2851 avec une taille de fenêtre de 512. Nous avons ensuite extrait les caractéristiques de chaque sifflement (fréquences initiales, finales etc..)

Afin de comparer différentes méthodes de réduction de dimensionnalité, une ACP et une t-SNE ont été réalisées sur ce jeu de données extrait automatiquement. Les groupes théoriques ont également été calculés (clustering), ainsi que le NMI (Normalized Mutual Information).

Le plus grand NMI est atteint en utilisant 4 paramètres acoustiques (fréquences maximales, minimales, durée et vitesse des sifflements), en imposant 2 groupes dans le regroupement automatique et en utilisant une taille de fenêtre de 512 échantillons.

Afin de maximiser encore d'avantage le NMI, les 500 plus longs sifflements ont été sélectionnés. Cette nouvelle sélection a dû être réalisée car la durée des sifflements varie en fonction de la distance entre les animaux et l'hydrophone. La puissance du signal est positivement corrélée aux distances par rapport aux dauphins. Une partie du sifflement n'est pas détectée si les animaux sont trop éloignés. Donc, plus les animaux seront loin, plus leurs sifflements seront courts. En sélectionnant les sifflements les plus longs, nous choisissons indirectement des dauphins les plus proches de l'hydrophone.

Sur la figure 4.39, chaque point représente un sifflement. Il est difficile sur cette représentation de mettre en évidence différents groupes. La faible valeur du NMI (0.11) confirme que l'analyse des sifflements par l'ACP n'a pas permis de mettre en évidence des groupes acoustiques associés aux comportements.

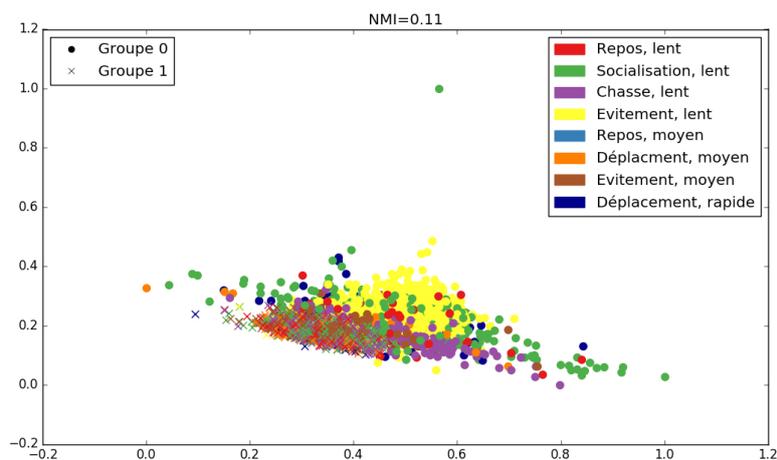


FIG. 4.39 – Visualisation d'une projection ACP de sifflements de *Stenella attenuata* en fonction de leur comportement et leurs vitesse de nage (avec 4 paramètres utilisés : Fréquence initiale, finale, durée et vitesse des sifflements mis en évidence par regroupement acoustique automatisé, NMI=0.11).

La figure 4.40 représente également tous les paramètres acoustiques mais avec la t-SNE. Le groupe 1 (croix) est composé de tous les comportements, tandis que le groupe 0 (cercle) est composé presque uniquement de sifflements émis lors du comportement d'évitement et en socialisation (tous les deux en vitesse lente). Le NMI est égale à 0.30 : il y a donc une corrélation entre les émissions acoustiques et les comportements de surface des dauphins.

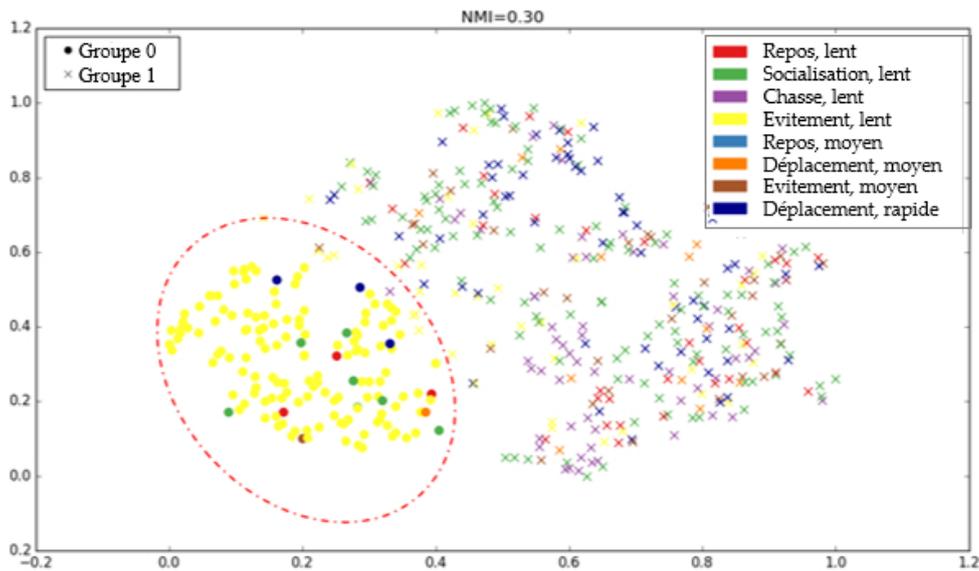


FIG. 4.40 – Visualisation d'une projection t-SNE des sifflements de *Stenella attenuata* en fonction de leurs comportements et leurs vitesses de nage (avec 4 paramètres utilisés : fréquence initiale, finale, durée et vitesse des sifflements) mis en évidence par regroupement acoustique automatisé, NMI=0.30.

L'objectif de cette étude était le développement d'une méthode de détection, d'analyse et de représentation de données en bioacoustique et son application sur une problématique précise : l'influence des comportements sur la communication des dauphins tachetés pantropicaux. Dans cette étude, deux méthodes de projections ont été testées : l'ACP et la t-SNE. En comparant visuellement la Figure 4.40 et la Figure 4.39, les groupes sont d'avantage mis en évidence par la t-SNE. En effet, l'ACP était composée uniquement d'objets répartis dans l'espace sans réellement appartenir à un groupe. De plus, le NMI est supérieur pour la t-SNE (=0.30) que pour l'ACP (=0.11). Ceci est explicable car l'ACP est une technique linéaire qui ne préserve pas les distances locales des données à haute dimension, donc dans notre contexte il est donc préférable d'utiliser une méthode non linéaire.

Notre méthode permet de mettre en évidence une modification de la structure des sifflements en fonction des comportements, particulièrement lorsque les individus ont un comportement d'évitement des bateaux et en socialisation (Poupard et al., 2019c, 2018a). Parallèlement à cette analyse, nous avons également effectué une analyse manuelle des sifflements afin de valider notre méthode automatique. Les résultats de cette analyse manuelle sont basés sur l'analyse des fréquences maximales et de la durée de chaque sifflement.

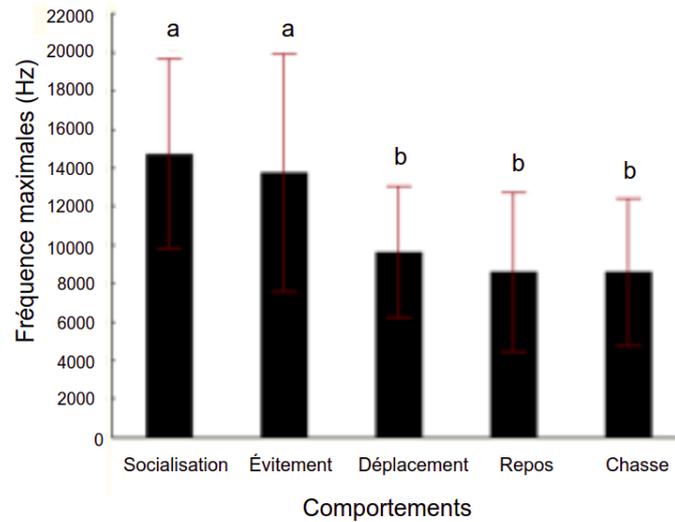


FIG. 4.41 – Variation des fréquences maximales des sifflements en fonction des différents comportements. Les lettres différentes indiquent des différences significatives entre les comportements.

La figure 4.41 représente la variation des fréquences maximales en fonction des cinq comportements. On peut voir que les fréquences maximales des sifflements en socialisation et en harcèlement sont statistiquement différentes des trois autres comportements.

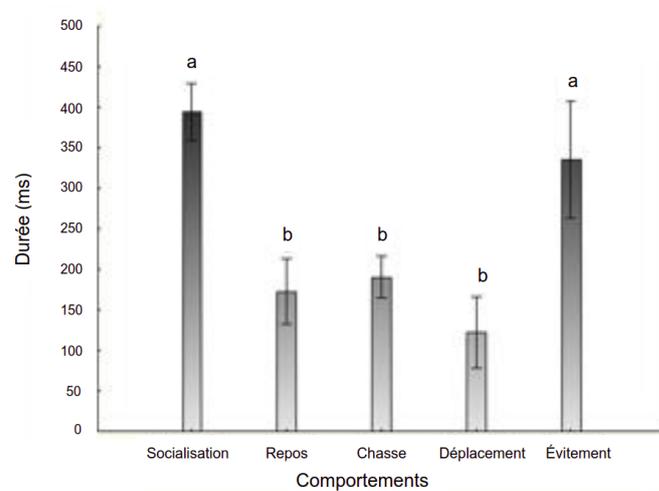


FIG. 4.42 – Variation de la durée des sifflements en fonction des différents comportements. Les lettres différentes indiquent des variations significatives entre les comportements.

On constate que la durée des sifflements est statistiquement identique entre la socialisation et le harcèlement (Fig.4.42). En effet, la moyenne des durées des sifflements pour la socialisation atteint 400ms, tandis que pour le comportement de chasse, la moyenne est égale à 195ms. Les durées de sifflements entre le repos, la chasse et le déplacement sont statistiquement identiques entre elles (170ms pour le repos, 195ms pour la chasse et 130ms pour le déplacement). L'analyse automatique nous a permis de montrer que durant le comportement d'évitement les sifflements étaient différents. Tandis que l'analyse manuelle (fréquence maximale et durée des

sifflements) nous a montré que les sifflements étaient statistiquement identiques en socialisation et en évitement. L'analyse automatique nous a permis de travailler sur beaucoup plus de paramètres acoustique (16 en tout), et donc devrait décrire d'avantage la forme des sifflements. Grâce à l'analyse manuelle, nous avons pu valider la méthode automatique, particulièrement pour le comportement d'évitement où les sifflements sont bien différents.

L'ensemble de cette étude est publiée dans (Poupard et al., 2019c)

4.4.2 Réponse acoustique des Baleines à bosse en fonction d'un stimulus chimique

Cette dernière étude concernait l'influence des stimuli chimiques sur l'émission acoustique des Baleines à bosse.

Notre algorithme a détecté un total de 2314 vocalisations à partir des enregistrements à Madagascar.

Dans notre analyse, nous avons considéré les 1000 appels les plus longs (c'est-à-dire les plus complets), car ceux-ci ont été émis par les baleines les plus proches du diffuseur de stimulus où l'hydrophone était attaché (la réception sonore et la puissance étant négativement corrélées à la distance à la source).

Le clustering avec un NMI maximum (NMI=0,27) est obtenu avec trois groupes (clusters) (Fig.4.43).

Le premier groupe (nommé 'o') est composé de vocalises émises dans toutes les conditions. Le cluster 2, comporte des sons émis avec la présence de CTL (contrôle) dans le milieu. Tandis que le cluster 1 comporte uniquement des sons émis lorsqu'il y avait de véritables stimuli olfactifs dans le milieu (DMS et Krill).

Ces résultats suggèrent que les vocalisations produites lors des expositions à des signaux chimiques ont des caractéristiques acoustiques différentes de celles produites lors des essais de contrôle.

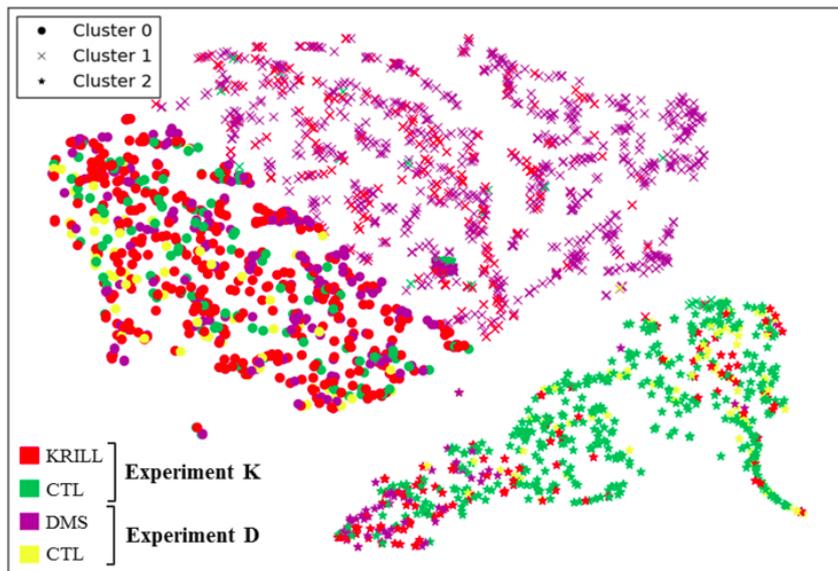


FIG. 4.43 – Analyse des vocalisations des baleines émises à Madagascar lors des expériences d'exposition. Après la réduction de la dimensionalité (t-SNE) de tous les paramètres acoustiques, un clustering non paramétrique bayésien (BNP) est appliqué aux données. Un score NMI maximal (0,27) est obtenu en utilisant 3 clusters.

Cette dernière étude peut être qualifiée de comportementale car elle a mis en évidence la corrélation entre des émissions acoustiques, et le comportement des animaux qui a évolué en fonction des stimuli chimiques placés dans la colonne d'eau. L'ensemble de cette étude est publié dans Bouchard et al. (2019).

Dans cette partie, les résultats en terme de bioacoustique comportementale ont été présentés. Nous avons mis au point une méthode (détecteur de signaux, réduction de dimensionnalité, clustering) applicable sur deux espèces différentes sans détecter acoustiquement les espèces.

Chapitre 5

Discussion générale

Comme développée dans l'introduction, la communication acoustique animale est utilisée pour diverses fonctions comme se nourrir (Elowson et al., 1991 ; Madsen et al., 2004), se reproduire (Boughman and Moss, 2003 ; Welbergen and Davies, 2008 ; Vignal et al., 2004), pour les interactions sociales (Ford, 1989 ; Sousa-Lima, 2005), les reconnaissances individuelles dans un groupe (Sousa-Lima et al., 2002) et pour beaucoup d'autres fonctions.

Ce dernier chapitre de thèse discute de l'ensemble des résultats obtenus dans les parties précédentes. Chaque cas d'étude a été accompagné de résultats préliminaires en bioacoustique. Cette discussion répond au questionnement initial qui était :

Comment la bioacoustique peut étudier la faune sur différentes échelles d'études : l'échelle spécifique, l'échelle populationnelle, l'échelle individuelle et l'échelle comportementale ?

Ce raisonnement sera basé sur l'ensemble des résultats présentés ainsi que les études réalisées dans la bibliographie. La première partie de ma discussion portera sur les avancées écologiques et méthodologiques des cas d'étude présentés.

La deuxième partie de cette discussion présentera différentes comparaisons afin de justifier le choix d'utilisation de certaines méthodes et des modèles d'études (espèces).

La sélection de différentes espèces, produisant différents types de signaux, évoluant dans différents milieux est volontaire afin de pouvoir explorer et mettre en comparaison les résultats et les méthodes utilisées.

Enfin, cette argumentation sera fondée sur le fait que la bioacoustique peut s'appliquer à toutes les échelles d'études : les espèces, les populations, les individus et comportementale.

La dernière partie de ce manuscrit présentera les conclusions de cette analyse, et s'appuiera sur de futures recherches.

5.1 Avancées méthodologiques et bioacoustiques

Comme explicitée dans l'introduction, la bioacoustique est une discipline récente développée et utilisée largement auprès de nombreux scientifiques.

En effet, les performances technologiques des enregistrements et des capacités de stockage n'ont fait qu'augmenter, impliquant l'anticipation de nouvelles méthodes automatiques pour extraire les signaux d'intérêt. L'objectif de cette thèse a été de contribuer au développement de ces nouvelles méthodes afin qu'elles soient utilisées facilement par la communauté scientifique. Afin de mettre en avant les différentes avancées de cette thèse le tableau 5.1 présente les contributions bioacoustiques et méthodologiques de nos résultats.

Au total, 9 cas d'étude ont été appliqués pour 4 échelles d'analyses : spécifique, populationnelle,

individuelle et comportementale. Chaque étude présente des avancées à la fois en bioacoustique mais également méthodologiques notamment sur l'association des méthodes supervisées ou non supervisées et automatiques ou manuelles.

Ces avancées seront décrites dans les parties suivantes. Il est important de rappeler que certains résultats en bioacoustique ne sont pour l'instant que des tendances, et n'ont pas encore été généralisés.

TAB. 5.1 – Résumé des différentes avancées pour tous les cas d'étude associés à chaque échelle d'analyse.

Protocole	Avancées/tendances bioacoustiques	Avancées méthodologiques	Perspectives de travail
Bioacoustique spécifique			
Détection d'espèces avifauniques		Efficacité des micros des mobiles CNN efficace pour la classification d'oiseaux	Construction d'une application mobile de reconnaissance avec la LPO Utilisation de cette application : Sciences citoyennes
Détection d'Orques	Présence des animaux en marée montante et à la pleine lune Présence des animaux durant les mois d'été	Première détection automatique d'Orques sur une grande échelle temporelle (3 ans) CNN efficace pour cette tâche	Actualiser la détection sur 2020 Mise en place de ce détecteur pour Orcalab
Détection de Cachalots	Présence d'individus seuls régulièrement. Présence de groupes (jusqu'à 9 individus) Présence des individus pendant la journée Sens Est/Ouest et Ouest/Est les plus courants	Protocole identique appliqué sur des centaines d'heures d'enregistrements. Méthode "semi-automatique" efficace pour les transitoires	Calcul des IPI pour l'identification individuelle Utilisation des annotations pour l'entraînement de nouveaux modèles d'IA
Bioacoustique populationnelle			
Classification de clans d'Orques		CNN pas suffisant avec 160 exemples Uniformisation des classes pas suffisant	Changer de modèle de classification Classification des pods, puis matriline Créer davantage de données d'entraînement
Bioacoustique Individuelle			
Détection d'individus (Orques)	Succession des vocalises dépendantes les unes des autres : N48, No4s produit pendant le "surf" Possibilité de relier certains appels à des perturbations (ferry) Taux d'appels changeant en fonction de la distance au groupe	Extractions de 85 appels appartenant à 3 individus CNN pas assez précis pour le contour des vocalises 4 hydrophones sont suffisants pour l'azimut	Améliorer le protocole visuel Intensifier les enregistrements pour généraliser
Détection d'individus (Cachalot)	Confirmation des temps de plongée/repos : Temps de surface : 15 min Temps de plongée : 42 minutes Profondeur : 1.2km Lien entre la tortuosité, profondeur et ICI	Localisation depuis un drone mobile de surface	Élargissement du protocole pour d'autres espèces (Globicéphales)
Détection d'individus (oiseaux)		FE et distance entre les micros trop faibles pour la localisation, micros trop différents	Améliorer le protocole acoustique
Bioacoustique comportementale			
Dauphin tacheté pantropical	Changement de communication quand ils sont harcelés par des bateaux	Chaîne des méthodes inédites : extraction, t-SNE et clustering.	Application à d'autres odontocètes
Baleines à bosse	Communication différente lorsqu'il y a des stimuli chimiques dans l'eau	Chaîne des méthodes inédites : extraction, t-SNE et clustering.	Application à d'autres mysticètes

Les différentes avancées seront donc séparées en deux grandes parties : les avancées méthodologiques et les avancées en bioacoustique.

5.1.1 Avancées méthodologiques

Au cours de la réalisation de cette thèse, différents jeux de données ont été analysés et ce, pour les quatre échelles d'analyses (spécifique, populationnelle, individuelle et comportementale). Ces jeux de données peuvent être qualifiés de "Big Data". En effet, pour l'échelle spécifique, 10,8 Teras d'enregistrements ont été analysés automatiquement (Orques et Cachalots), pour l'échelle individuelle nous avons analysé 500 Gigas de sons, et pour l'échelle comportementale 430 Gigas ont été analysés.

La première échelle d'analyse (spécifique) sous-marine a reposé sur la détection d'Orques et de Cachalots. A notre connaissance, il s'agit des premières analyses automatiques à grande échelle d'Orques et des Cachalots (période de 3 ans).

Certaines études ont réalisé des analyses sur autant de données, notamment chez la Baleine bleue avec plus de 7 ans d'enregistrements (Rice et al., 2015), ou encore des clics de cétacés analysés par Frasier et al. (2017) (52 millions de détections). Mais globalement des analyses sur plusieurs années d'enregistrements sont rares, qu'elles soient terrestres ou sous-marines.

Les Orques en Colombie-Britannique sont très étudiées (Ford et al., 1987 ; Deecke et al., 2000, 2005), mais la plupart de ces études ont utilisé des méthodes manuelles pour extraire les signaux.

Récemment, l'étude de Bergler et al. (2019) a mis au point un détecteur automatique d'Orques en utilisant 4 jeux de données différents. Leur score de classification était performant (93% de TBC), cependant, les exemples de tous les jeux de données se trouvaient dans chaque ensemble (train, validation, test). Cette répartition ne reflète donc pas les capacités de généralisation du modèle.

L'article de Best et al. (2020) a présenté différentes architectures de réseaux neurones pour la détection de vocalises d'Orques. Cette étude s'est particulièrement intéressée sur l'impact de l'hétérogénéité des données d'entraînement pour la généralisation des différents modèles. Ness et al. (2013) a mis en place un système aidant les chercheurs à écouter, voir et annoter des enregistrements d'Orques afin de pouvoir détecter automatiquement cette espèce dans les enregistrements. Cette base de données collaborative est composée de plus de 20 000 heures de vocalisations d'Orques collectées en 30 ans. Ces données annotées pourraient nous être utiles dans l'amélioration des performances de notre modèle pour la détection d'Orques, mais également pour la détection de clans acoustiques.

En effet, la tâche complexe de la classification automatique de clans (échelle populationnelle) chez les résidents de Colombie-Britannique est au coeur des recherches. La première étude proposant un modèle automatique est Deecke et al. (1999). Ils ont essayé d'apprendre un indice de similitude entre les différents types d'appels produits par les résidents. Les résultats ont montré une corrélation significative entre les indices proposés par le réseau neuronal et les similarités acoustiques entre les individus. Ce travail préliminaire a donc prouvé qu'il était possible de mettre en place des modèles basés sur ces différences.

Schröter et al. (2019) à utilisé les enregistrements de la base de données d'Orchive pour classer

12 appels différents. Ils ont obtenu 87 % de TBC, mais uniquement sur 70 appels différents, ce qui est un jeu de données très succinct. Leur modèle n'est pas disponible en ligne, et n'a pas été testé sur de longues périodes temporelles.

Concernant nos modèles pour la classification de clans, les performances de chaque détecteur testé étaient trop faibles.

Pour remédier à cela, nous aimerions tester d'autres modèles de classification. Il faudrait également obtenir davantage d'annotations par classe afin de pouvoir entraîner de manière efficace nos modèles.

La deuxième solution pour la détection de population a été d'utiliser des méthodes non supervisées comme le clustering suite à une extraction du pitch (Fo). Mais les résultats pour cette méthode n'étaient malheureusement pas suffisants pour classer les pods efficacement.

Aucune des méthodes (supervisées et non supervisées) n'a permis d'aboutir à de résultats satisfaisants pour la classification de pods d'Orques (échelle populationnelle).

Concernant le modèle de classification de chants d'oiseaux, nos performances étaient de 88% de bonne reconnaissance sur une cinquantaine d'espèces. Comme présentée dans l'Introduction, la tâche de classification de sons d'oiseaux est de plus en plus étudiée. La création de Challenges internationaux comme BirdClef (Goëau et al., 2016) motive la communauté scientifique à développer et comparer des modèles de classification de chants d'oiseaux, dont on constate que les meilleurs modèles sont de type deep learning.

Les meilleures scores dans les conditions d'usage identique à birdify (1 espèce vocalisant par fichier de test), avoisines les 72% TBC pour 1500 espèces d'oiseaux (Sevilla and Glotin, 2017).

Ces modèles sont relativement performants sur un nombre d'espèces important, mais l'objectif de notre classifieur était d'avoir des performances permettant un usage pour le citoyen et l'expert, avec un score supérieur à 85% TBC, et donc sur un nombre plus restreint d'espèces (notre modèle atteint plus de 88% TBC sur 50 espèces). Nous voulions que notre classifieur d'espèces soit très robuste pour les espèces françaises les plus communes. Pour cela, un long travail de sélection des données d'entraînement a été réalisé. Les bonnes performances de ce modèle sont dues à la bonne architecture de celui-ci mais également à la bonne qualité des données d'entraînement, la sélection en diversité des types de chants et en qualité RSB pour l'apprentissage de notre modèle.

Nous aimerions maintenant proposer nos différents modèles à la communauté scientifique.

Le modèle de détection d'Orques pourrait servir aux écovolontaires d'Orcalab. En effet, cette station de recherche met en place tous les ans, un système d'écovolontaires afin d'écouter H24 7/7j les 6 hydrophones de la station. Notre modèle installé au sein de la station pourrait alléger le travail de ces derniers.

Le modèle mis en place pour la classification de chants d'oiseaux (échelle spécifique) n'a pas encore été appliqué sur de gros jeux de données. Mais le modèle a été construit pour la création d'une application mobile qui sera mis à disposition des membres de la LPO prochainement. Une fois l'application utilisée par un certain nombre de personnes, nous pourrons faire des analyses statistiques plus approfondies afin d'évaluer la fréquentation de certaines espèces d'oiseaux dans l'espace et le temps.

Un prototype d'application a été construit à l'aide d'un stagiaire en développement mobile. Des fiches d'identification ont été élaborées afin de donner des informations sur l'espèce

reconnue dans l'enregistrement (Fig.5.1). Ce prototype est disponible à l'adresse suivante : <https://birdify.univ-tln.fr/labo>.

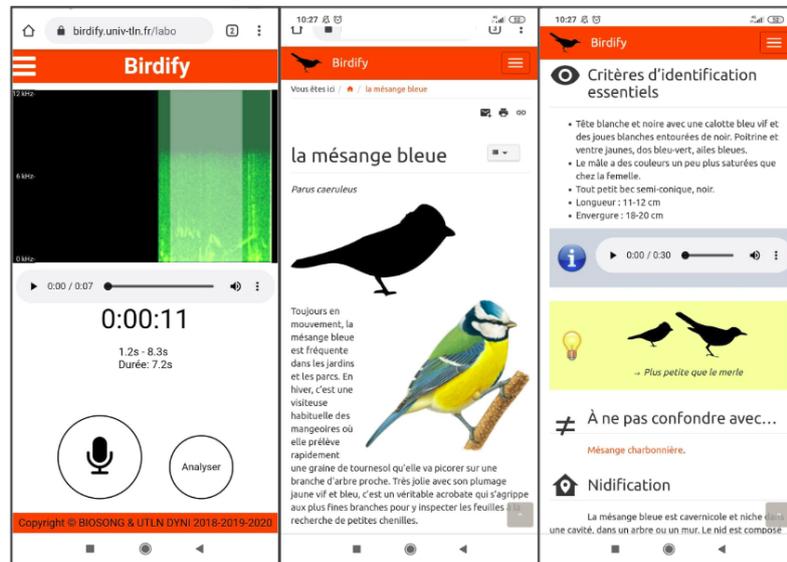


FIG. 5.1 – Image du prototype sur mobile. Gauche : Lancement de l'enregistrement, sélection du spectrogramme par l'utilisateur, et analyse de l'enregistrement. Milieu et droite : Réponse du modèle avec la fiche d'identification de l'espèce correspondante.

Ce prototype a été discuté et proposé à mes financeurs (BIOSONG), puis il pourrait être testé à une plus large échelle par des membres de la LPO. Une fois testé, ce prototype sera amélioré et mis en ligne sous forme d'application mobile gratuite.

Concernant la détection de Cachalots sur la bouée Bombyx, c'est la première étude sur le Cachalot en Méditerranée où un protocole unique a été appliqué sur un nombre d'heures d'enregistrements aussi important.

L'étude de Morrissey et al. (2006) a également mis au point un détecteur de clics de Cachalots afin de pouvoir les localiser par la suite. Leur détecteur est efficace pour des hydrophones placés sur le fond. Il n'y a donc pas beaucoup de bruit de surface et les clics possèdent un RSB très fort (en comparaison avec Bombyx), rendant la tâche de détection moins complexe. Leur détecteur ne serait donc pas efficient pour notre jeu de données.

L'étude de Caruso et al. (2015) a détecté les clics et analysé les IPI des Cachalots de 2100 enregistrements de 5 minutes (représentant env 175h d'enregistrement). Leur détection de clics est réalisée en deux étapes. La première est une investigation visuelle des enregistrements afin d'extraire des périodes temporelles contenant des clics de Cachalots. Puis, la deuxième étape s'effectue en appliquant un filtre passe-bande de 3 à 16 kHz. Un seuil adaptatif est calculé sur la valeur médiane de l'énergie sonore totale dans l'enregistrement de 5 min, et ensuite appliqué. Leur étude est donc qualifiée de "semi-automatique".

Notre détecteur peut être également qualifié de semi-automatique, car un premier détecteur de transitoires a été appliqué, puis une analyse manuelle des trajectoires issues des TDoAs a été nécessaire. Nous allons prochainement travailler sur la mise en ligne de ce processus (détection,

TDoAs et suivi des pistes) afin qu'il soit disponible et utilisable facilement par la communauté. Concernant l'échelle individuelle, différentes avancées méthodologiques sont à signaler. Trois protocoles ont été mis en place pour l'identification individuelle : la localisation d'Orques depuis la côte, la localisation de Cachalot depuis un drone de surface et la localisation d'oiseaux depuis une antenne fixe.

En terme méthodologique, c'est la première fois que des vocalises individuelles d'Orques ont été extraites depuis la côte.

Certaines études ont déjà identifié un émetteur dans un groupe d'Orques (Miller et al., 2004 ; Miller and Tyack, 1998) constatant que les types et le moment des appels sont influencés par les appels des autres membres du groupe. Mais ces études ont utilisé un réseau de 15 hydrophones tractés derrière un bateau scientifique. Il est probable que les résultats de ces études soient influencés par la présence du bateau. Le troisième cas d'étude pour l'identification individuelle des Orques a montré que les individus changeaient leur taux d'appels et les types d'appels lors du passage d'un ferry (tendance). Il paraît donc indispensable de mettre en place des protocoles passifs pour ces problématiques.

L'identification individuelle a également été réalisée pour le Cachalot depuis un drone de surface. Beaucoup d'études performantes utilisent les TDoAs pour localiser des signaux depuis une bouée sous-marine (Simard and Roy, 2008 ; Roy et al., 2010 ; Gillespie et al., 2020), mais elles sont fixes, et par conséquent ne peuvent donc pas suivre les animaux sur de longues distances. Dans Thode (2004), un réseau de deux hydrophones remorqué à grande ouverture était utilisé pour le suivi 2D des profils de plongée du Cachalot, mais ne permettant pas la localisation 3D de l'animal.

Notre étude a démontré la première piste de plongée 3D précise à plus de 2 km de distance, par surveillance acoustique passive à partir d'un drone, avec un réseau à petite ouverture de 5 hydrophones espacés de manière non uniforme. Le suivi de piste a été possible grâce à la dérive silencieuse du drone et la fréquence d'échantillonnage élevée de la carte JASON. Nous démontrons ainsi qu'un observatoire mobile haute définition pourrait élargir nos connaissances du comportement de plongée des cétacés.

La dernière échelle d'analyse explorée est la bioacoustique dite "comportementale". Pour cette échelle, la détection **acoustique** d'espèce, de population ou d'individu n'a pas été nécessaire.

Cette nouvelle chaîne des méthodes automatiques non-supervisées (détection, t-SNE, clustering et NMI) s'est révélée efficace pour des études en bioacoustique comportementale.

Les avancées méthodologiques de cette thèse sont validées par le traitement de jeux de données massifs, et seront mises à disposition de la communauté scientifique. Différentes comparaisons méthodologiques seront réalisées dans le chapitre suivant afin d'attribuer une chaîne de méthode spécifique en fonction de la problématique étudiée.

5.1.2 *Tendances bioacoustiques*

Cette thèse nous a permis de mettre en place des méthodes novatrices sur des jeux de données importants, mais elle a également mis en évidence des **tendances** bioacoustiques pour chaque cas d'étude.

La première échelle d'analyse est le niveau spécifique (4 premières lignes du tableau 5.1). Pour

cette dernière, trois cas d'étude ont été présentés : la classification d'oiseaux, la détection d'Orques et de Cachalots. En terme d'avancées bioacoustiques, cette échelle fut assez prometteuse.

Le premier cas d'étude pour la bioacoustique à l'échelle spécifique concernait la détection automatique d'Orques. Cette étude nous a permis de comprendre plus clairement les mouvements de cette espèce dans une échelle temporelle importante (3 ans). Les tendances bioacoustiques nous ont permis de mettre en évidence une plus grande fréquentation du Blackney Pass lorsque la marée montait, et que les cycles lunaires pouvaient avoir une influence considérable sur le mouvement des Orques. Nos résultats préliminaires valident ceux de Benoit-Bird et al. (2009) ; Lin et al. (2013).

Le deuxième cas d'étude pour la détection d'espèces concerne le Cachalot (*Physeter macrocephalus*) au large des côtes toulonnaises.

Cette détection a démontré que les individus pouvaient chasser seul (la plupart du temps) mais également en groupes constitués (jusqu'à 9 individus). Nos tendances confirment les résultats de Whitehead and Weilgart (1991) ; Pierantonio et al. (2008) ; Gannier et al. (2002) montrant que les groupes de Cachalots en Méditerranée peuvent aller jusqu'à 10 individus.

Après des tests statistiques, nous avons montré que les animaux étaient davantage présents acoustiquement durant la journée (de 8h à 18h) que durant la nuit.

Puis, la troisième échelle d'étude portait sur la détection individuelle (Cachalots, Orques et oiseaux). La détection individuelle d'Orques par localisation nous a permis d'obtenir des tendances très intéressantes en terme de bioacoustique. En effet, trois expériences naturelles ont été présentées. Une qui nous a permis de mettre en évidence quel type d'appel serait utilisé pour la réunification d'un pod. La deuxième expérience nous a montré le déroulement des échanges entre deux matriline nageant ensemble et les successions des vocalisations seraient dépendantes les unes des autres. Nos tendances sont en accord avec Miller et al. (2004) qui a démontré que la production acoustique au sein d'une matriline reposait sur la succession d'appels consécutifs qui n'était pas aléatoire.

Enfin, la dernière expérience a mis en évidence la réaction acoustique et comportementale d'un groupe d'Orques en présence d'un ferry. Ces tendances complètent les résultats de Holt et al. (2009) qui a montré que les individus ont augmenté l'amplitude de leurs appels de 1 dB pour chaque augmentation de 1 dB de bruit de moteurs. Notre étude n'a pas travaillé sur le niveau des signaux émis mais sur les types d'appels émis et leur taux d'appels. Les orques changeraient donc l'intensité, la quantité et le type d'appels face aux bruits anthropiques.

Le deuxième cas d'étude pour l'identification d'individus concernait la localisation de Cachalot depuis un drone de surface mobile. Cette étude nous a permis notamment de valider les conclusions de Watwood et al. (2006) confirmant que cette espèce pouvait plonger à plus de 1.2km sous la surface pendant plus de 40 minutes.

Nous avons également démontré que la tortuosité de l'individu serait liée avec son ICI et sa profondeur. Des calculs de tortuosité (en 2D) ont déjà été effectués pour différentes espèces sous-marines (Wilson et al., 2007). Ils ont notamment démontré que le calcul de la tortuosité pour les pinnipèdes et les tortues était très efficace dans l'étude du comportement spatial et temporel de ces espèces.

Notre détecteur de signaux stationnaires pour l'échelle comportementale nous a permis de

démontrer que les Dauphins tachetés pantropicaux changeraient leur communication en présence de bateaux et pendant la socialisation.

D'autres études ont déjà démontré ces résultats (Busnel and Dziedzic, 1966 ; Smolker et al., 1993). Par exemple, les dauphins communs en Méditerranée ont des fréquences initiales plus élevées en socialisation (Lopez and Shirai, 2009). Les auteurs ont en effet mis en évidence une variation des émissions entre les comportements : les résultats ont démontré que le nombre d'émissions a augmenté lors de la socialisation (Cook et al., 2004). Donc, en terme de quantité d'émissions acoustiques, les dauphins émettraient plus de sifflements lorsqu'ils se socialisent. Dreher and Evans (1964) a également révélé que, pour chaque espèce, un sifflement était utilisé dans un contexte comportemental particulier.

Il est possible d'affirmer, avec nos résultats, que les émissions acoustiques lors de pressions anthropiques sont différentes des autres comportements (Fig.4.40). En effet, lors du comportement d'évitement, les groupes de sifflements sont rassemblés dans le groupe o. Ce groupe est donc caractéristique de ce comportement particulier.

Différentes études ont été réalisées sur l'analyse comportementale des dauphins en présence de bateaux (Lesage et al., 1999 ; Nowacek et al., 2001 ; Lusseau et al., 2009). Les différents auteurs ont constaté une nage plus rapide, moins de temps en surface, et également une modification des profils de plongées (donc la recherche de nourriture) (Aguilar et al., 2006). Nous pouvons donc corrélérer cette modification de comportements aux variations acoustiques que nous avons observé.

La bioacoustique comportementale a également montré que les Baleines à bosse émettaient des vocalises différentes lorsqu'il y avait des stimuli chimiques dans le milieu.

Il est important de rappeler que l'ensemble de ces résultats n'est qu'une "tendance", et nous ne pouvons pas généraliser ces effets sur l'ensemble des individus/populations/espèces. Les événements acoustiques étaient relativement rares, de ce fait, nous n'avons pas pu réaliser de tests statistiques robustes pour l'ensemble de nos cas d'étude afin de valider nos tendances. En revanche, les méthodes utilisées peuvent être validées.

L'ensemble des avancées bioacoustiques et méthodologiques est présenté dans le tableau 5.1.

5.2 Comparaisons : signaux, milieux et méthodes

Différents paradigmes et méthodes ont été mis en comparaison dans cette thèse afin d'évaluer les problématiques des différents modèles d'études (espèces), au sein d'écosystèmes distincts (canal de transmission) et des différentes méthodes d'analyses utilisées.

Cette partie va donc comparer ces différentes notions en s'appuyant sur des cas d'étude concrets issus des résultats et de la bibliographie.

5.2.1 Les types de signaux étudiés

Au cours de cette thèse, différents modèles d'études ont été utilisés, produisant différents types de signaux.

En fonction de ces derniers, certaines méthodes peuvent être appliquées et d'autres non.

Les signaux analysés peuvent être définis en deux principales catégories : les signaux stationnaires et les signaux transitoires.

Les signaux dits stationnaires sont définis comme étant relativement longs dans le temps. La représentation temps fréquence via la transformation de Fourier est particulièrement efficace pour représenter des signaux stationnaires.

Afin de visualiser ces différents types de signaux, la figure 5.2 montre à gauche 7 secondes de clics enregistrés sur BOMBYX (transitoires) et à droite 7 secondes de vocalises enregistrées à Orcalab (stationnaires).

La première étape du traitement du signal en bioacoustique est de choisir une bonne représentation de ce dernier, afin d'optimiser la détection, la classification ou encore la localisation.

Ces différentes représentations (forme d'ondes et spectrogramme) ont été décrites dans la première partie du Chapitre Méthodes.

A droite de la figure 5.2, sont représentés quelques clics d'écholocation, on peut voir qu'il est facile de les identifier depuis la forme d'onde, car ces pulses sont très courts en temps (quelques ms) et très larges en fréquence (3-30 kHz).

Tandis que sur la gauche sont représentées des vocalises d'Orques, qui sont beaucoup plus longues en temps, mais moins larges en fréquences.

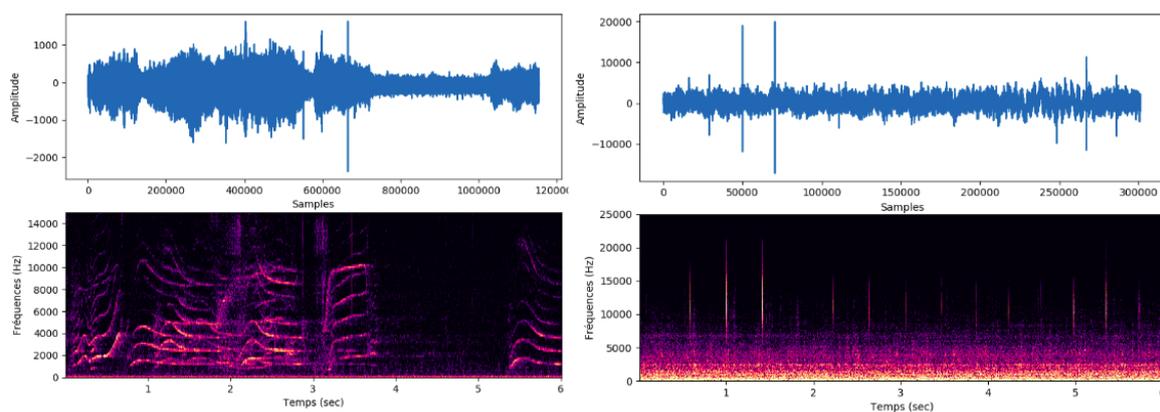


FIG. 5.2 – Représentation de la forme d'onde et du spectrogramme de 6 secondes de vocalises d'Orques (stationnaires à gauche) et de 6 secondes de clics d'un Cachalot (transitoires à droite).

La forme des signaux va présenter certains avantages et inconvénients pour leur analyse et certaines méthodes vont se révéler plus ou moins efficaces.

Pour la détection et l'extraction d'événements acoustiques comme des clics ou des vocalisations, différentes méthodes peuvent être utilisées et sont présentées dans la partie suivante.

5.2.1.1 Les signaux stationnaires

Les signaux stationnaires comme les vocalises d'Orques ou encore des chants d'oiseaux sont très riches en information temporelle. Donc la représentation en spectrogramme est très efficace pour des réseaux de type Deep Learning, ou le modèle apprend directement sur une image (Knight et al., 2020 ; Branson et al., 2014).

Le modèle apprend donc à discriminer un motif spectro-temporel au sein de l'enregistrement. Bien entraîné, il peut détecter ces motifs même en présence de bruit. Les bruits sous-marins ou terrestres ne présentent pas les mêmes caractéristiques temporelles et fréquentielles que les chants ou vocalises (voir Introduction).

Il est donc plus facile d'extraire ces signaux stationnaires en deep learning.

Nous avons profité d'un modèle de détection de vocalises d'oiseaux pour apprendre le détecteur de vocalises d'Orques (méthode dite de 'transfert learning'). L'apprentissage de stationnaires en bioacoustique est donc transférable d'une espèce à l'autre, même évoluant dans des milieux différents. Les premières couches du réseau transférées s'adaptent rapidement aux nouvelles conditions, s'il s'agit de signaux stationnaires, les deux modèles partagent de nombreux points communs.

Le chapitre "Résultats" a présenté différents réseaux neurones pour la reconnaissance automatique d'espèces (Orques et espèces d'oiseaux). Il s'est avéré que ces méthodes sont très efficaces pour la classification de ces signaux stationnaires.

La détection de signaux stationnaires peut également impliquer d'autres méthodes que l'intelligence artificielle. La première méthode, présentée en Introduction, est assez répandue et est la détection manuelle des vocalises/chants/sifflements par analyse du spectrogramme (Truskinger et al., 2013 ; Ford, 1989). Ces méthodes sont relativement efficaces mais très difficiles à reproduire sur de grandes quantités de données.

Différentes méthodes automatiques ont été développées pour la détection de signaux stationnaires basées sur l'analyse du spectrogramme (Gillespie et al., 2013 ; Halkias and Ellis, 2006 ; Shapiro, 2006 ; Halkias et al., 2013).

Plus particulièrement des détecteurs de "pitch" (fréquences fondamentales) ont été réalisés (Shapiro and Wang, 2009 ; Poupard et al., 2019b). La fréquence fondamentale d'une émission peut fournir beaucoup d'informations concernant l'action de communication. Mais, pour certaines espèces, comme les Orques, l'extraction du pitch peut être parfois difficile à cause du chevauchement des vocalises en temps (voir chapitre Résultats, partie Détection de clans acoustiques). Également les bruits continus, comme les moteurs de bateaux, peuvent être confondus avec des vocalises

Une fois les vocalises sélectionnées, pour certaines problématiques (détection individuelle) il a fallu calculer le temps de délais d'arrivée entre ces dernières (TDoAs). Mais notre détecteur automatique n'était malheureusement pas assez précis pour faire le contour entier de la vocalise, et calculer avec précision leur TDoAs. Il a donc été nécessaire d'isoler manuellement l'ensemble des vocalises pour cette étude. La localisation est une tâche plus complexe pour des signaux stationnaires en milieu naturel.

5.2.1.2 Les signaux transitoires

A l'inverse, les signaux transitoires comme les clics de Cachalots sont très courts en temps, et ces motifs peuvent ressembler à du bruit, ou d'autres sons biologiques (cliquetis de crevettes) ou transitoires anthropiques (moteurs, chantiers). Il paraissait donc difficile de construire un réseau CNN sur ce type de signaux.

Différentes études ont déjà tenté de développer des réseaux neurones pour la détection ou

classification de clics d'odontocètes (Yang et al., 2020 ; Luo et al., 2019). Ces différentes méthodes peuvent utiliser le signal pur en entrée du réseau ou encore certains paramètres préalablement extraits.

Mais la classification de transitoires paraît plus difficile car ces travaux de classification sont basés sur la représentation spectrogramme, mais les transitoires sont trop courts pour que le modèle soit efficace dans l'apprentissage de motifs bien définis (Ferrari et al., 2020).

La détection de signaux transitoires comme les clics d'odontocètes peut également se faire grâce à des méthodes classiques d'analyse du signal. Des caractéristiques particulières des clics peuvent être extraites manuellement comme les valeurs de fréquence des pics, des enveloppes spectrales ou encore des intervalles entre les clics (Gerard et al., 2008 ; Jarvis et al., 2008 ; Roch et al., 2011).

La méthode utilisée dans nos différentes études (Sphyrna et BOMBYX) consiste en une succession de filtres appliqués sur le signal (voir chapitre Méthodes). Ces derniers ont été relativement efficaces pour le protocole réalisé depuis le drone Sphyrna, car sa forme particulière (deux coques asymétriques) impliquait peu de bruit de surface. L'antenne principale était placée à 2m sous la surface et peu de bruit de clapot et de vagues de surface ont été enregistrés sur les hydrophones. Ces différents bruits peuvent avoir la forme de transitoires des clics de Cachalots et donc peuvent avoir un impact sur les performances de nos détecteurs.

Concernant la méthode de détection de clics sur la bouée BOMBYX, une étape manuelle de vérification fut nécessaire. En effet, la bouée était placée à -25 m sous la surface, au sud de Port-Cros. Cette zone étant très fréquentée par les bateaux, beaucoup de bruits anthropiques étaient présents (transitoires). Il était donc impossible pour notre détecteur actuel de différencier ces bruits, des clics de Cachalots.

La stéréophonie de la bouée m'a donc permis de différencier les clics issus des Cachalots, des transitoires issus du bruit. En effet, il était possible de suivre les différents TDoAs des clics de Cachalots évoluant dans le temps, tandis que les clics de bruit possédaient toujours des TDoAs constants. A l'avenir, nous aimerions mettre en place un détecteur de clics automatiques, puis un classifieur automatique afin de classer les différentes espèces d'odontocètes présentes (Ferrari et al., 2020).

Pour la localisation de sources, les signaux transitoires sont très précis en temps, il est donc possible de calculer leur TDoAs précisément.

Nous avons volontairement fait le choix de sélectionner différents modèles d'études produisant des types de signaux bien différents (stationnaires vs transitoires).

Les deux types de signaux peuvent être utilisés en bioacoustique, mais nous avons vu que leurs caractéristiques en temps-fréquences les rend plus ou moins malléables pour certains travaux.

5.2.2 La nature du canal de transmission : terrestre et marin

Comme développé dans l'introduction, le canal de transmission a son importance dans la communication animale (Forrest, 1994). Les différents cas d'étude de cette thèse ont permis de montrer que la nature du canal de transmission (marin ou terrestre) a une réelle influence sur les méthodes d'analyses des signaux.

La présence de bruit dans les deux milieux est variable et peut nuire à la bonne détec-

tion/classification ou encore localisation. Le niveau de nuisance sonore peut varier à cause de son amplitude, sa fréquence, sa distribution spatiale mais également de sa répétabilité.

La première échelle d'analyse concerne la détection d'espèces. Différentes méthodes ont été utilisées notamment les modèles d'intelligence artificielle. Les deux réseaux utilisés (détection d'Orques et d'espèces avifauniques) ont été efficaces. Mais ces deux modèles ont nécessité un travail considérable de sélection pour les données d'entraînements.

Deux canaux de transmission ont été utilisés dans cette thèse : le milieu marin et terrestre. En fonction de nos différentes échelles d'analyse, la nature du canal de transmission a un impact sur nos différentes méthodes d'analyse.

5.2.2.1 Le milieu terrestre

Dans mon travail de thèse, la bioacoustique en milieu terrestre repose sur la mise au point d'un modèle de classification automatique d'espèces d'oiseaux. Les performances de ce dernier dépendent de la qualité des données d'entraînement et notamment si elles sont bruitées ou non (vent, pluie etc.).

La variabilité acoustique du canal de transmission peut être un frein à la bioacoustique mais dans certains cas, elle peut être utilisée comme support de communication, c'est le cas de certaines espèces d'oiseaux dans les forêts tropicales qui utilisent les différentes réverbérations du milieu pour communiquer (Slabbekoorn et al., 2002).

Dans le chapitre "Matériels", différentes manipulations ont été présentées afin de tester la performance des microphones des différentes marques de téléphones portables et également l'influence du milieu sur leurs réceptions.

Lors de mes expériences, plusieurs résultats ont été présentés concernant la propagation du son dans différents milieux terrestres. Les niveaux rapport signal sur bruit (RSB) ont été étudiés en fonction de la distance d'émission mais également de l'angle d'émission (Poupard et al., 2018c). Le RSB était plus important en milieu ouvert, et lorsqu'il n'y avait pas d'obstacle entre l'émetteur et le récepteur.

Le RSB est un facteur déterminant pour la problématique de détection ou classification. Notamment, en milieu marin, Wiley and Richards (1978) ; Dadouchi et al. (2015) confirme qu'une perte de 6 dB en RSB correspond à une portée de détection divisée par deux. Bradbury et al. (1998) a regroupé les différents bruits ambiants qui pourraient exister dans les milieux terrestres, en fonction des habitats. Les basses forêts tropicales produiraient du bruit homogène ambiant entre 4 et 10 kHz, tandis qu'en milieu ouvert, les fréquences du bruit ambiant sont beaucoup plus faibles.

Les turbulences acoustiques générées par le vent sont comprises entre 100 et 1000 Hz et jusqu'à 70 dB pour les vents forts (8 m/sec). Les micros des téléphones peuvent être très sensibles à ces perturbations, il faut donc préconiser des enregistrements sans vent pour les utilisateurs.

La figure 5.3 représente un exemple de 15 secondes de signal de Troglodyte mignon *Troglodytes troglodytes* avec et sans bonnette sur l'enregistreur, qui sont des accessoires en mousse pour protéger du bruit du vent les microphones.

Sans la bonnette (spectrogramme du bas) les signaux entre 50 et 200 Hz sont masqués par le

vent. En milieu terrestre, il est donc possible d'ajouter ce genre de système de protection afin d'augmenter le RSB, et donc favoriser une bonne classification.

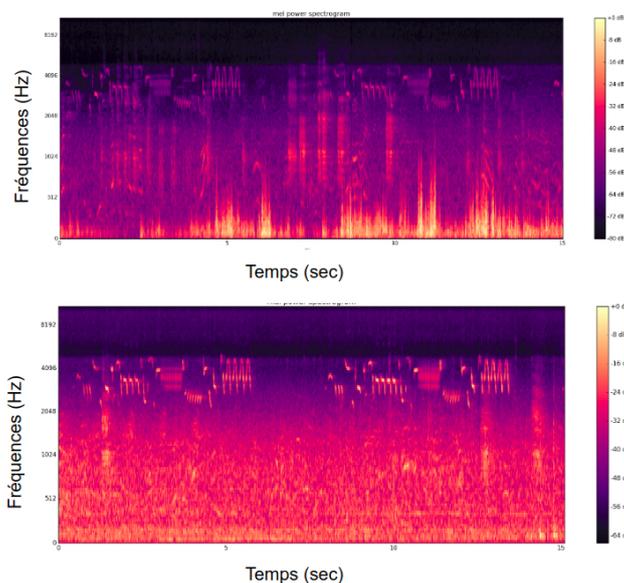


FIG. 5.3 – Spectrogrammes représentant les différences de dB entre les sons émis par l'enceinte (en haut avec bonnette et en bas sans bonnette à 12,5 m) de 15 secondes de Troglodyte mignon. Enregistrements réalisés Par T. Lengagne et moi même à Lyon, 2018.

Cependant les bruits non anthropiques terrestres ne proviennent pas uniquement du vent ou de la pluie, mais peuvent provenir d'autres espèces. Les sources majeures de bruit dans le milieu terrestre peuvent être produits par les anoues, les oiseaux, les cigales et des orthoptères (Bradbury et al., 1998). Ces types de bruit peuvent commencer à 500 Hz, et peuvent parfois se prolonger jusqu'aux ultrasons.

La localisation de source en milieu terrestre est par contre plus complexe dès que cette dernière est à plus d'une quinzaine de mètres des microphones.

Le calcul du temps de délai d'arrivée entre les microphones peut être réalisé de différentes manières (Liang et al., 2012 ; Blumstein et al., 2011). La localisation aviaire présentée dans la partie 4.3.3 n'a pas abouti à des résultats convaincants.

Quatre raisons en sont probablement la cause :

- La distance entre les microphones était trop faible (21cm maximum)
- Les microphones étaient différents : 2 microphones internes du Tascam DR40 et 2 microphones externes (Tbone LC 97 TWS)
- La fréquence d'échantillonnage était trop faible (96 kHz)
- Les oiseaux se déplaçaient très vite devant les microphones
- Il y avait trop de réverbérations (obstacles, arbres)
- Les microphones étaient trop directifs pour calculer les TDoAs

En effet, en moins de 2 secondes, une cible pouvait être du côté du premier microphone

(M₁), et se déplacer du côté de M₂.

Différents protocoles ont tenté de dénombrer le nombre de sources dans un milieu terrestre (Rhinehart et al., 2020).

Les premières études de Alexandridis and Mouchtaris (2015) ; Blumstein et al. (2011) ; Mennill et al. (2006) ont utilisé plusieurs micros espacés de 4 à 8m. Blumstein et al. (2011) ont montré que les milieux ouverts entraînaient de meilleurs performances dans la localisation car il y avait peu de réverbération. Il a notamment établi que son système permettait de localiser régulièrement les sons avec une précision de quelques dizaines de centimètres. Mais la distance entre ces 4 microphones était de 40 m soit une distance 200 fois plus grande que la notre. Le choix d'une distance aussi faible entre nos microphones était volontaire afin de faciliter au maximum le déploiement de l'antenne pour de futures utilisations par des ornithologues.

Blumstein et al. (2011) a également démontré que les performances de localisation dépendaient donc de l'habitat, de la distance de la source mais également du type de signal émis (modulation en fréquence).

Certaines équipes ont notamment travaillé sur la localisation grâce aux différences de niveaux d'énergies arrivant sur chaque microphone (Aarabi, 2003). La localisation de source dans le milieu terrestre nécessite donc un habitat ouvert (sans obstacle) mais surtout une distance entre les capteurs beaucoup plus importante que dans le milieu marin (40 cm pour le protocole Orcalab 4D).

La localisation de source peut se faire grâce à deux principales méthodes : le calcul du temps de délai d'arrivée entre les microphones, ou le calcul de la différence d'intensité de ces derniers. Pour le calcul des TDoAs, il faut que la forme d'onde soit identique sur tous les microphones, donc cela nécessite des microphones omnidirectionnels posés dans la même direction.

Tandis que pour le calcul de la variation d'intensité, il faudrait des microphones très directifs pointant vers différentes directions. (Glotin et al., 2015). La détection et la classification de sons dans le milieu terrestre est donc réalisable, par contre la localisation nécessite du matériel particulier et des méthodes très performantes.

5.2.2.2 Le milieu marin

L'écosystème marin est également composé de différents bruits, biologiques ou non, pouvant nuire ou altérer la détection d'espèces. Pour la détection d'Orques également, les données d'entraînement doivent être bien sélectionnées afin d'éviter des erreurs de classification. Généralement, les bruits sous-marins sont classés en 3 principales catégories en fonction de leur fréquences

(Hildebrand, 2009) :

- les bruits basses fréquences (10 à 500 Hz) produits par le trafic maritime principalement
- les bruits moyennes fréquences (500 Hz à 25 000 kHz) produits par les vagues, le vent, les déferlantes.
- les bruits hautes fréquences produits par le résultat du mouvement brownien des molécules d'eau sur l'hydrophone et des cliquetis des crustacés.

La communication est très utilisée en milieu marin car la propagation des sons est 4 fois plus rapide que dans l'air, mais les bruits se propagent également à une vitesse importante

(voir Introduction).

Une des échelles d'analyse de cette thèse était l'identification individuelle sonore obtenue à l'aide de la localisation de la source. Dans ce cadre, le milieu marin était illustré pour la localisation des Cachalots et des Orques.

Deux types de signaux ont été analysés et localisés (clics et vocalises). Le milieu sous-marin est donc favorable à la localisation des sources car il est possible de capter des clics à plus de 2km pour le Cachalot, et plusieurs centaines de mètres pour les vocalises d'Orques. Les Orques de Colombie-Britannique se déplacent relativement doucement (1.7m/sec) (Williams and Noren, 2009), le Cachalot, lui peut aller jusqu'à 4m/s en chasse (Amano and Yoshioka, 2003). Ces vitesses relativement faibles les rendent ainsi plus facilement localisables.

La détection et la localisation des sources en milieu sous-marin va donc dépendre de leur propagation et également de la configuration de l'antenne acoustique (Simard et al., 2008).

Différentes études ont mis au point des réseaux d'hydrophones (non tractés derrière un engin) très performants pour la localisation des mammifères marins (Simard et al., 2006 ; Grebner et al., 2007 ; Simard and Roy, 2008). Pour ces travaux, la distance entre les hydrophones était de plusieurs dizaines de mètres jusqu'à plusieurs kilomètres. Les TDoAs étaient donc moins complexes à calculer.

Plus la distance (D) entre les hydrophones sera faible, plus la fréquence d'échantillonnage (F_e) doit être importante. Ces deux paramètres donnent un critère sur le nombre d'échantillons ("bin") maximum (TDoAs maximum). En effet, chaque paire d'hydrophone doit évaluer un gisement entre 0 et 180 degrés, le TDoA maximum doit être supérieur à 180 bins pour avoir une bonne précision (1 bin pour 1 degrés).

Il est donc tout à fait possible ou réalisable de localiser des sources en milieu marin.

Les milieux terrestres et marins sont donc très différents pour la mise en oeuvre de la bioacoustique et selon les objectifs de l'étude, différentes méthodes et protocoles devront être adaptés pour chacun des milieux.

Il est important de noter que certaines espèces utilisent la communication dans les deux environnements. C'est le cas des pinnipèdes, capables de communiquer dans le milieu terrestre et marin. Southall et al. (2003) a démontré que ces espèces étaient capables de détecter des signaux terrestres et marins sur une large gamme de fréquences à des rapports signal / bruit relativement faibles. Donc pour cette espèce, il faudrait envisager deux protocoles : un terrestre et un marin simultanés afin d'étudier la communication dans les deux milieux.

5.2.3 Comparaisons des apports pour les méthodes supervisées et non supervisées en bioacoustique

La dernière comparaison se concentre sur les différentes méthodes présentées dans la thèse. La figure 5.4 (première version présentée dans le chapitre "Méthode") résume les différentes méthodes utilisées.

Les carrés gris clairs représentent les tâches de localisation via le calcul des TDoAs. On voit que le calcul des TDoAs peut servir à la fois pour la détection d'espèces (Cachalots) mais également pour le suivi individuel.

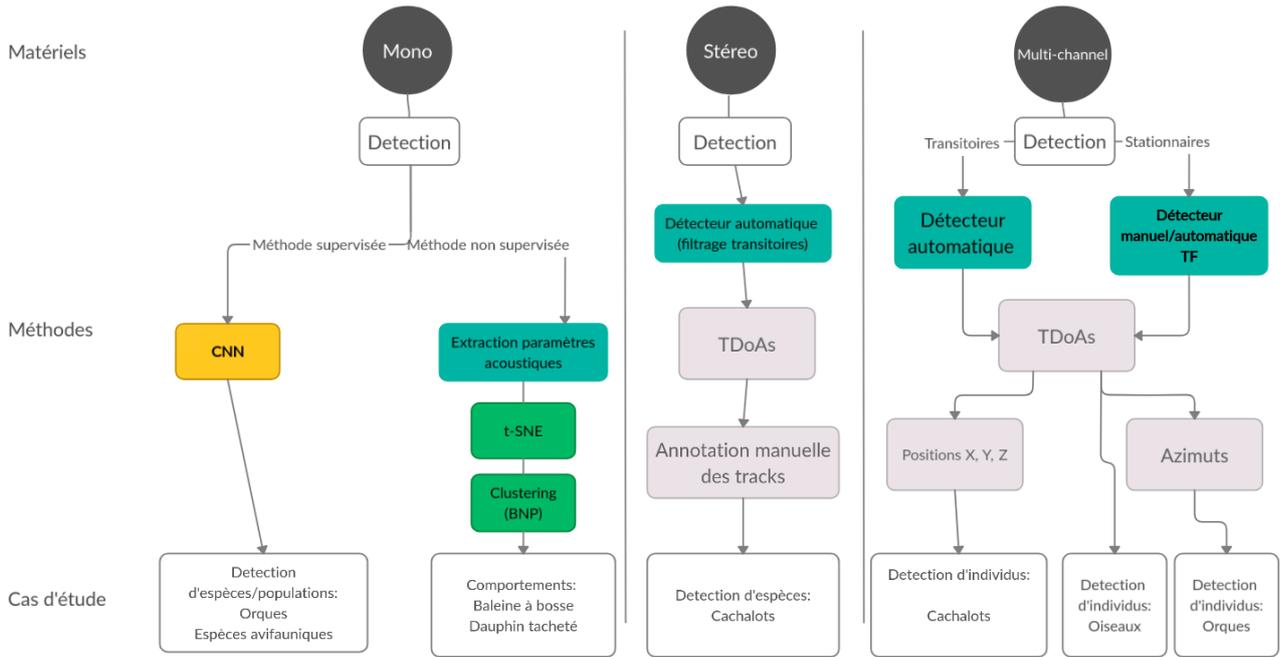


FIG. 5.4 – Résumé des méthodes utilisées associées à chaque cas d'étude. Orange : méthodes supervisées. Bleu : détections issues de méthode semi-manuelles (filtre, suivis TF). Gris : localisation.

A ce stade de la thèse, il paraît fondamental d'opposer les méthodes dites supervisées (orange dans la Fig.5.4) ou non supervisées (vert dans la Fig.5.4). Un état de l'art de ces méthodes a été présenté en première partie du chapitre Méthodes.

Le tableau 5.2 compare les différents avantages et inconvénients des méthodes supervisées et non supervisées dans le cadre de la bioacoustique.

En fonction des différents objectifs et des différentes échelles d'études (espèces, populations, individuelles et comportementale) certaines méthodes seront plus efficaces. Les méthodes supervisées comme le Deep learning sont très utiles pour la détection d'espèces. En effet, les deux modèles de Deep learning utilisés pour la détection d'espèces d'oiseaux et d'Orques se sont révélés efficaces.

Une fois la détection d'espèces effectuée, il est possible d'associer la présence de cette espèce à des paramètres extérieurs (marée, périodes saisonnières). Ces types de modèles sont donc très utiles pour la bioacoustique des espèces, et peuvent servir pour étudier leur fréquentation de certaines zones. Le principal avantage de ce type d'approche (supervisée) c'est qu'une fois entraînés les modèles sont relativement stables et peuvent être utilisés pour des tâches de détections sur des quantités d'enregistrements importantes. C'est le cas des Orques de Colombie-Britannique ou la détection d'Orques sur 3 ans d'enregistrement n'a pris que 48h. Mais les méthodes supervisées nécessitent malheureusement du matériel particulier comme des GPU (Graphical Processing Unit) et des calculateurs bien spécifiques. Ce matériel est généralement disponible dans des laboratoires d'informatique et non dans des laboratoires d'écologie.

Puis les méthodes supervisées nécessitent des compétences particulières en Machine Learning. Il

faut maîtriser des outils comme [Tensorflow](#) ou [Keras](#) indispensables pour la création de certains modèles.

Le principal inconvénient des méthodes supervisées est qu'elles nécessitent un large volume de données labélisées. Pour entraîner les modèles à reconnaître une espèce, ou une population d'Orques par exemple, il faut des centaines voire des milliers d'exemples de vocalises appartenant à chaque classe. Il est particulièrement difficile d'obtenir ce type de données particulièrement pour le milieu marin, ou mis à part les scientifiques, personne ne réalise d'enregistrement de la faune.

Au contraire, pour le milieu terrestre, et particulièrement pour les oiseaux, les enregistrements de la faune sont très courants. Les ornithologues sont très actifs en bioacoustique et ils mettent à disposition leurs enregistrements labélisés au service de la communauté scientifique.

Les données d'entraînement pour la classification d'espèces avifauniques ont été obtenues assez facilement pour la tâche de classification grâce à [Xeno-Canto](#). Ce site Web est dédié au partage d'enregistrements d'oiseaux du monde entier, il est donc facile d'obtenir des données d'entraînement pour les espèces les plus courantes.

Par ailleurs, quand les données annotées sont inexistantes, et que l'on a aucune idée des caractéristiques pouvant créer certaines classes, les méthodes non supervisées sont très utiles. C'est le cas par exemple de la réduction de dimensionnalité (t-SNE) des MFCC des différentes vocalises appartenant à deux groupes d'Orques différents. La t-SNE nous a permis de montrer que les appels produits par deux mâles étaient bien différents.

Les méthodes non supervisées ne nécessitent pas un bagage en informatique très important (comparé aux méthodes supervisées), et les bibliothèques notamment en Python sont très faciles d'utilisation (clustering et réduction de dimensionnalité).

Mais les méthodes non supervisées présentent également des inconvénients, le premier est la difficulté d'évaluer la robustesse des nouvelles classes créées. Les mesures de performances sont assez difficiles à mettre en place. Dans notre cas, les performances du clustering ont été mesurées avec le NMI (Normalized Mutual Information) qui a mesuré l'information mutuelle entre les clusters acoustiques et les clusters de comportements (pour les dauphins tachetés pantropicaux). Un NMI proche de 1 signifiait que les clusters acoustiques correspondaient aux clusters comportementaux et un score de 0 signifiait qu'il n'y avait aucune correspondance entre les deux distributions. Le NMI peut donc servir de mesure de performance, mais cette chaîne de méthodes est totalement nouvelle et n'a jamais été encore utilisée en bioacoustique. L'ensemble de ces avantages/inconvénients est décrit dans le tableau [5.2](#).

TAB. 5.2 – Récapitulatif des avantages/inconvénients des méthodes supervisées et non supervisées dans le cadre de la bioacoustique.

	Méthodes supervisées		Méthodes non supervisées	
	Avantages	Inconvénients	Avantages	Inconvénients
Matériel		Nécessité de GPU Matériel particulier	Machines classiques	
Compétences nécessaires	Solution sous Python pour le traitement des sons	Compétences nécessaires en informatique	Moins de compétences en informatique	
Complexité temporelle		Important temps de calcul	Calculs rapides	
Données d'entraînement		Besoin de beaucoup de d'enregistrements labélisés	Pas besoin de d'enregistrements labélisés	
Temps d'interprétation	Classes déjà définies Interprétation facile			Temps à l'interprétation des nouvelles étiquettes créées
Traitement Big data	Très efficace pour les masses de données sonores		Très efficace pour les masses de données sonores	
Mesures des performances	Facile d'évaluer la qualité des modèles TBC, matrice de confusion			Difficultés mesures des performances
Transfer learning	Transfer learning possible (ex oiseaux vers Orques)			Pas de Transfer learning
Stabilité des modèles	Relativement stable une fois entraîné		haute dépendance des paramètres (clustering et t-SNE)	
Informations complémentaires		Pas d'information sur la structure des données ("boite noire")	Possibilité d'extraire des informations sur la structure des signaux	

Une fois les informations acoustiques extraites, soit par des méthodes supervisées, soit par des méthodes non supervisées, il faut les mettre en lien avec des données "non acoustiques" ou environnementales. Ces données peuvent être extraites par les scientifiques (comportement) ou être récupérées sur des plateformes en ligne (exemple coefficients des marées).

L'acquisition de ces données "comportementales" peut être réalisée visuellement pour l'échelle de l'acoustique comportementale. En effet, pour cette échelle d'étude, l'acquisition de données visuelles possède deux principales fonctions :

- La détection de l'espèce présente
- La détection de comportement de surface

Ces deux fonctions sont totalement indispensables pour cette échelle d'étude car, ni l'espèce, ni la population et ni les individus ne sont identifiés acoustiquement.

Pour les autres échelles (espèces, populations, individus) un protocole visuel n'est pas forcément nécessaire. Les détections acoustiques peuvent être corrélées avec des paramètres extérieurs comme la marée ou la tortuosité.

Le cas d'étude pour l'identification individuelle des Orques de Colombie-Britannique a nécessité un

protocole visuel de "vérification" afin que les positions visuelles des individus servent de contrôle pour les positions acoustiques estimées. La détection d'individus s'est faite acoustiquement et non visuellement.

Cette partie "Comparaisons : signaux, milieux et méthodes" a mis en contradiction pour la bioacoustique :

- Les signaux transitoires vs les signaux stationnaires
- Les écosystèmes terrestres vs marins
- Les méthodes supervisées vs non supervisées.

L'ensemble de ces comparaisons est résumé en figure 5.5. Le premier niveau d'analyse pour la bioacoustique est donc les différentes échelles d'études : comportementales, espèces, populations et individus. Au sein de ces quatre échelles, il est possible de s'intéresser à l'écosystème terrestre ou marin. Puis, l'analyse repose sur deux principaux types de signaux : les transitoires (Tr) et stationnaires (St). En fonction de ces types de signaux, différentes méthodes sont utilisées (supervisées, non supervisées ou encore localisation).

Les triangles bleus représentent les travaux correspondants réalisés dans cette thèse tandis que les chiffres (1, 2 et 3) correspondent à des exemples de la bibliographie qui ont répondu à ces problématiques en bioacoustique.

Les croix vertes sont des branches où aucune étude en bioacoustique n'a encore été effectuée. Ce schéma permet de visualiser les différents champs d'études possibles en bioacoustique et les champs qu'il reste à explorer en terme méthodologique.

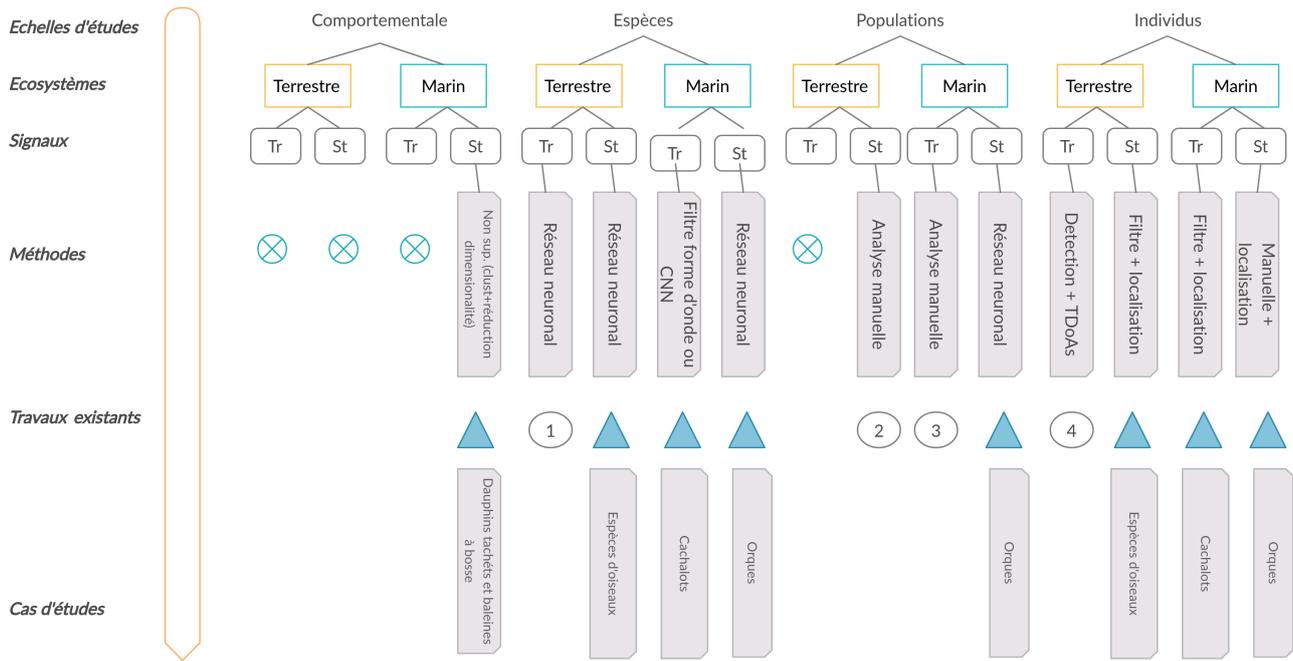


FIG. 5.5 – Résumé des différentes comparaisons discutées dans cette partie. Tr= Transitoires. St= stationnaires. Triangles bleus : travaux issus de ce travail de thèse. Croix bleues : travaux pas encore réalisés. Nombre (1, 2, 3, 4) : **exemples** de travaux existants. 1 : Mac Aodha et al. (2018), 2 : Joachim and Lauga (1996), 3 : Weilgart and Whitehead (1997), 4 : Glotin et al. (2015).

5.3 La bioacoustique appliquée sur l'ensemble des échelles

La bioacoustique permet d'aborder les systèmes de communication à différentes échelles : **spécifique, populationnelle, individuelle** et permet de travailler également sur le **comportement**. Ces différentes échelles ont été abordées au cours de mon travail. Les différentes expériences et cas d'étude ont montré que la bioacoustique peut s'appliquer également sur des signaux sans détecter (acoustiquement) les espèces, populations ou individus. En effet, la dernière échelle (partie 4.4) porte sur l'influence de certains paramètres (comportements ou alimentation) sur les Dauphins tachetés pantropicaux et Baleines à bosse. Pour ces deux cas d'étude, aucune détection acoustique d'espèce n'a été réalisée. Le détecteur automatique détecte des signaux stationnaires biologiques (vocalises ou sifflements) et non des espèces.

Avant de détecter les espèces, les populations ou les individus, l'échelle d'étude la plus vaste, est donc l'échelle "comportementale", où le détecteur extrait les signaux stationnaires biologiques (vocalises et sifflements).

Pour cette échelle d'étude, la détection d'espèces se fait donc visuellement depuis un bateau. La figure 5.6 illustre les différents paramètres pouvant influencer la bioacoustique sur l'ensemble des échelles d'études. Ces différents paramètres peuvent être les nuisances anthropiques, les relations sociales entre les individus, les facteurs environnementaux etc... Ces paramètres ont

une influence sur le comportement des individus, et donc sur leurs émissions acoustiques. Sur ce schéma, l'ensemble des échelles est représenté : l'échelle la plus large est la bioacoustique spécifique, puis populationnelle et enfin individuelle. La quatrième échelle : "bioacoustique comportementale" est transversale aux quatre autres échelles, car nous pouvons faire le lien entre des comportements et des émissions pour des individus, des populations et des espèces.

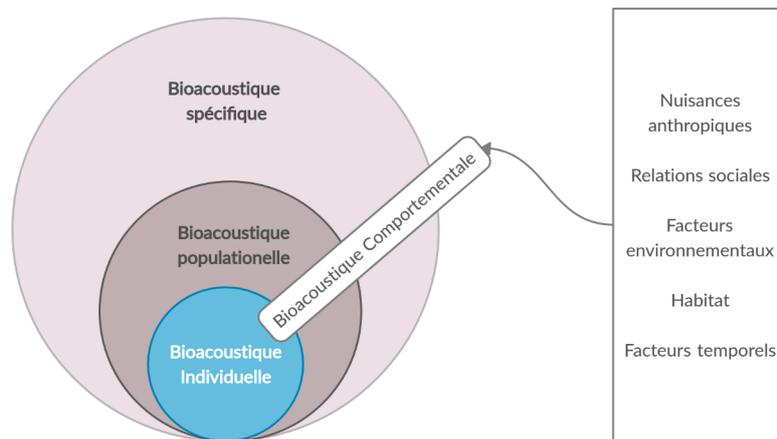


FIG. 5.6 – Schéma récapitulatif des facteurs influençant la bioacoustique pour les différentes échelles d'études. La bioacoustique comportementale est transversale aux autres échelles.

La cause d'un comportement représente les mécanismes qui ont déclenché un comportement particulier. Donc l'ensemble des facteurs cités précédemment peuvent être catégorisés comme la "Cause" en bioacoustique (Fig.5.6). En effet, les nuisances anthropiques, les relations sociales et les facteurs environnementaux sont les causes d'un comportement particulier qui va avoir un impact sur l'acoustique du sujet (espèce, population, individu).

Il est donc possible d'étudier l'influence de ces causes directement par les émissions acoustiques (sans étudier le comportement en lui-même).

Par exemple, pour la détection d'espèce, notamment chez les Orques de Colombie-Britannique, nous avons corrélé la présence des animaux avec l'évolution de la marée. Cette étude bioacoustique ne s'est pas intéressée directement au comportement des animaux, mais à la cause (marée) de ce comportement.

Les différentes études peuvent se faire directement en impliquant les comportements des animaux, mais également en s'intéressant à la cause de ces comportements.

Le tableau 5.3 présente donc le lien qui peut être fait entre des événements et des émissions acoustiques pour les différents cas d'étude.

TAB. 5.3 – Tableau représentant les différents cas d'études avec l'association des variabilités acoustiques et des variabilités environnementales.

	Méthodes	Variabilité acoustique	Variabilité environnementale/ comportementale
(1) Espèces d'oiseaux	CNN	Nombre de détections	Habitat particulier
(2) Orques à Orcalab	CNN	Nombre de détections	Quantité de saumons
(3) Cachalots sur Bombyx	Pistage(TDoAs)	Nombre de détections	Vitesse du courant
(4) Pods d'Orques à Orcalab	CNN	Diversité des types d'appels	Complexité sociale
(5) Cachalots depuis Sphyrna (individu)	Solveur Non Linéaire(TDoAs)	Intervalle inter-clic (ICI)	Tortuosité
(6) Orques à Orcalab (individu)	Solveur Linéaire(TDoAs)	Taux d'appel	Eloignement d'un individu avec son groupe
(7) Oiseaux (individu)	Clustering(TDoAs)	Taux d'appel	Mouvement d'un individu
(8)Dauphins tacheté pantropical	Clustering(t-SNE(pistage))	Fréquence initiale des sifflements	Comportements
(9) Baleines à bosse	Clustering(t-SNE(pistage))	Taux d'appels	Présence de krill

Nous pouvons chercher à élucider les **mécanismes environnementaux** à l'origine d'une émission acoustique. Nous ne sommes pas concentrés sur les mécanismes cognitifs qui ont déclenché l'émission.

5.4 L'importance de l'échelle temporelle d'analyse

Les différents cas d'étude de cette thèse nous ont démontré que l'échelle temporelle d'étude des signaux était primordiale pour des études en bioacoustique.

Afin de comparer des signaux avec des évènements précis, il faut considérer la juxtaposition des signaux (vocalises, sifflements, clics) dans leur ensemble. L'analyse bioacoustique d'une scène nécessite l'étude de la séquence entière et ne pas se focaliser sur une seule émission.

Par exemple, pour le cas d'étude de l'analyse individuelle des Orques de Colombie-Britannique, l'ensemble des appels a été pris en compte en fonction de la distance des animaux. La notion de séquence est l'échelle d'analyse la plus appropriée.

L'émission des transitoires par les Cachalots pour la communication sous forme de coda est l'exemple parfait de l'étude d'une séquence. L'étude d'un seul clic ne permettra pas de mettre en corrélation les différents comportements (ex : distance entre les individus) avec les successions des signaux (Pavan et al., 2000). Il faut étudier les successions des différents codas dans leur ensemble.

Un deuxième exemple intéressant est l'émission acoustique des Baleines à bosse pendant la reproduction. Les scientifiques décrivent habituellement ces émissions en "chants", ceci dû à leur répétabilité (Sousa-Lima, 2005). L'objectif principal de cette émission acoustique serait la sélection de partenaire sexuel. Certains scientifiques suggèrent que les qualités des chants

servent principalement à informer les autres baleines (en particulier les femelles) de la condition physique d'un chanteur (Winn and Winn, 1978), tandis que d'autres supposent que les chansons servent principalement à informer les chanteurs de ce qui se passe au-delà de leur portée visuelle (sonar) (Mercado, 2018). Mais dans les deux hypothèses l'analyse de la séquence entière a été nécessaire.

La mesure des variabilités dans les signaux peut donc dépendre de l'échelle temporelle à laquelle nous abordons les différentes problématiques (Lengagne et al., 2015). En effet, la discontinuité dans les périodes d'enregistrements peut nuire à certaines analyses. Prenons l'exemple de l'analyse de la présence de Cachalots sur la bouée Bombyx. L'enregistrement n'était pas réalisé en continu, cependant, nous avons pu dénombrer le nombre de pistes. Par contre l'échantillonnage ne nous permettait pas d'estimer les temps de plongées/repos des différentes pistes.

Le croisement de ces différents concepts pourraient être appliqué pour comparer les espèces entre elles, mais aussi les populations et les individus en bioacoustique. En effet, il est envisageable en bioacoustique de mettre en place des méthodes d'analyse et de traitement des signaux adéquats sur différentes espèces. C'est le cas de (Cholewiak et al., 2018) qui a examiné les séquences des différents chants de Baleines à bosse en utilisant des mesures utilisées pour les oiseaux (2 paramètres acoustiques). Il ont donc montré que les mâles modifiaient leur chant en présence d'autres mâles dans la zone (distance variable entre les deux mâles). Les méthodes utilisées sont applicables à plusieurs espèces (oiseaux et Baleines à bosse). En effet, ces deux espèces (Combassou du Sénégal *Vidua chalybeata*, et la Baleine à bosse *Megaptera novaeangliae*) partagent de nombreuses caractéristiques vocales et comportementales rendant les méthodes d'analyses génériques. Il faut donc mettre en place des échantillonnages temporels adaptés et mettre en place des méthodes génériques transposables sur différentes espèces.

5.5 Représentation synthétique des questionnements en bioacoustique

En conclusion, un tableau a été élaboré afin de conceptualiser les différentes questions auxquelles un bioacousticien pourrait se poser, en fonction du signal émis.

Les concepts de ce tableau ont été développés en introduction (partie 1.3). Il est composé de quatre principales questions : "quoi ?", "qui ?", "où ?" et "pourquoi ?". Ce tableau serait donc applicable pour l'ensemble des cas d'étude en bioacoustique. Pour illustrer cette nouvelle représentation, voici deux exemples :

- L'émission acoustique d'une Orque rejoignant son pod.
- L'émission acoustique des Cachalots pour la chasse.

Le premier exemple est donc l'émission sonore d'une Orque rejoignant son groupe.

L'ensemble des questions (quoi ?, qui ?, où ?, pourquoi ?) ont trouvé une réponse pour cette séquence.

La succession du même type d'appel (N48) répond à la question "quoi ?" qui peut être enregistré avec un seul hydrophone, et traité avec des méthodes supervisées (CNN pour classer les appels) ou non supervisées (clustering).

Maintenant que la question "quoi ?" a obtenu une réponse, il faut passer à la suivante : "qui ?", afin de savoir si l'individu volontairement ou non, envoie de l'information à ses congénères pour se faire reconnaître.

Pour cela, les méthodes non supervisées seront utilisées, considérant que l'on ne connaît pas l'information qui va le distinguer des autres individus à priori (exemple des clusters issus de la t-SNE). Puis, la troisième question "où ?" va consister en la localisation de l'émetteur, et essayer de savoir s'il émet ce type d'appel pour localiser les membres de son pod et/ou signaler sa présence. Pour cela, plus de matériel est nécessaire, notamment plusieurs hydrophones pour calculer les TDoAs, et donc localiser l'émetteur.

La dernière question relève de l'ordre de l'écologie comportementale : pourquoi ces types d'appels ont été émis (cause) ? L'individu aurait volontairement émis ces types d'appels pour le maintien des cohésions sociales de son groupe. Afin de répondre à cette question, il faut analyser le comportement de surface des deux groupes, et notamment travailler sur la distance entre les individus. Puis, il est possible que ces signaux soient émis involontairement.

Le tableau 5.4 résume ces quatre questions.

TAB. 5.4 – Tableau bioacoustique de l'émission acoustique d'une Orque rejoignant son pod.
 1H= 1 Hydrophone, BF = Basse Fréquence 1 à 100Hz ; MF = 100Hz à 10 kHz ; HF = 10 kHz à 500 kHz. En blanc les questions qui relèvent de la bioacoustique et en gris celles qui relèvent de l'écologie comportementale.

Vocalises d'une Orque ralliant son pod	Signal volontaire	Signal involontaire	Matériels	Méthodes
Quoi ?	Une succession de vocalises type N48		1H, MF	Supervisée (CNN) Non supervisée Pitch tracking
Qui ?	Signature individuelle volontaire verrouillée à l'intérieur de la vocalise ?	Signature involontaire (Fo)= timbre, Séquence...	1H, MF	Pitch tracking Non supervisée
Où ?	Il émet cette vocalise pour localiser ses congénères, ou pour que ses congénères le localisent ?	Il communique sur sa position sans réellement le vouloir ?	4H, HF	Détection, TDoAs, localisation. Besoin d'un support visuel pour vérifier les positions si individus regroupés
Pourquoi ?	Maintien des relations sociales volontaires Accélération de la vitesse de nage Cause : Besoin de maintenir les relations sociales Fonction : se rapprocher de son groupe	Besoin "involontaire" génétique et phénotypique de rejoindre son groupe	Données visuelles	Analyse de la distance (AZ, 2D,...) entre les individus

La bioacoustique comportementale repose donc sur l'interaction de la question "pourquoi" (comportements) avec les autres questions (qui ?, où ? quoi ?).

Le deuxième exemple illustre une séquence de clics d'écholocation pendant le comportement de chasse d'un Cachalot (Tableau 5.5).

Ce type de cas d'étude peut être appliqué à l'échelle spécifique (étude des Cachalots en

Méditerranée sur la bouée Bombyx). La première question ("quoi?") est d'identifier les clics d'écholocation, notamment à l'aide des intervalles inter-clics qui ne sont pas dépendants de l'animal en lui même, mais de la distance entre lui et sa proie. Pour cela, il faut deux hydrophones afin de pouvoir calculer les TDoAs et de confirmer la trajectoire d'un Cachalot. Pour ce cas d'étude, il est donc indispensable de répondre à la question "où?" en premier, car la détection de clics d'écholocation n'a pu se faire sans la réalisation des TDoAs. Puis, la question "qui?" dans un clic peut être définie par la signature individuelle volontaire à l'intérieur du clic, ou encore par l'interval-interpulse (involontaire) de ce dernier. La question en écologie comportementale "pourquoi?" peut être résolue par le comportement de chasse (volontaire) et des facultés cognitives et involontaires que l'animal a développé au cours de sa vie pour déclencher un tel comportement.

TAB. 5.5 – Tableau bioacoustique de l'émission acoustique d'un Cachalot en chasse. 1H= 1 hydrophone, BF = basse fréquence 1 à 100Hz ; MF = 100 Hz à 10 kHz ; HF = 10 kHz à 500 kHz. En gris les questions qui relèvent de l'acoustique et en orange celles qui relèvent de la bioacoustique comportementale.

Clics de chasse d'un Cachalot	Signal volontaire	Signal involontaire	Matériels	Méthodes
Quoi ?	Clic d'écholocation	Intervalle inter-clic (ICI)	2H, HF	Analyse temporelle Filtre TDoAs
Qui ?	Signature individuelle volontaire verrouillée à l'intérieur du clic ?	Intervalle Inter-pulse (IPI)	1H, HF	Analyse temporelle et spectrale Filtre
Où ?	Il émet ce clic pour localiser ses congénères (ou proie), mais également pour que ses congère le localisent (chasse en groupe)	Il communique sur sa position sans réellement le vouloir lors des activités de chasse solitaire	4H+ 1H, HF	Détection, TDoAs,
Pourquoi ?	Comportement de chasse : volontaire Cause : Besoin de se nourrir Fonction : Apport d'énergie	Besoin "involontaire" génétique et phénotypique de chasser	Données temporelles	Étude des cycles saisonniers, nyctéméraux, journaliers...

Afin de valider ces concepts, le tableau 5.5 présente les quatre questions ("quoi?", "qui?", "où?" et "pourquoi?") pour un cas générique applicable à la plupart des émissions acoustiques associées à un comportement.

Ce tableau bilan est applicable à l'ensemble des cas d'étude présentés dans cette thèse. Les méthodes de traitements proposées sont réalisables pour la majorité des signaux.

TAB. 5.6 – Tableau générique représentant les questionnements en bioacoustique multiéchelle.

1H= 1 Hydrophone

	Signal volontaire	Signal involontaire	Matériels	Méthodes
Quoi ?	Types d'émissions volontaires	Types d'émissions involontaires	1 H	Supervisée majoritairement
Qui ?	Information individuelle volontaire (ex : signature)	Information individuelle involontaire (ex : timbre)	1 H	Non supervisée en cas de nouvelles questions, et nouvelles connaissances
Où ?	Echolocation (volontaire) Signalisation (volontaire)	Signalisation involontaire	4 H	TDoAs, estimation de vitesse, Profondeur, localisation 3D Non supervisée
Pourquoi ?	Cause Fonction	Besoins involontaires Besoins phénotypique Capacités physiologiques		Apport de données écosystémiques (extérieur à l'acoustique)

Une étude bioacoustique peut donc reposer sur la mise en place de matériel et de méthodes afin d'associer des questions en écologie comportementale (pourquoi ?) avec des questions acoustiques (quoi ? qui ? où ?).

L'ensemble de ces questionnements peut être appliqué à toutes les échelles d'études appliquées dans cette thèse.

5.6 Conclusions

Les différents cas d'étude associés à des échelles d'analyse nous ont permis de définir plus précisément les différents outils nécessaires aux études bioacoustiques.

Cette thèse s'est appuyée sur des modèles d'études différents, produisant des signaux éloignés (transitoires vs stationnaires) évoluant dans des milieux différents (terrestres vs marins). L'objectif était d'homogénéiser les méthodes d'analyses afin de faciliter le développement de nouvelles études en bioacoustique. Les méthodes proposées ont été testées et validées sur certains protocoles de données massives : Orcalab (espèces) 8 To ; Orcalab (individus) 0.5 To ; Bombyx 3 To pionnières dans la communauté bioacoustique.

Un volet méthodologique a été développé et mis en place pour des échelles d'études précises. Les différents cas d'étude ont permis de mettre en évidence des corrélations entre des événements et des émissions acoustiques. Ce premier travail pourrait donc servir à de futures études en bioacoustique. La figure 5.7 regroupe les différentes étapes pour la mise en place d'une problématique en bioacoustique. La première étape est de connaître le **milieu** dans laquelle l'espèce évolue. Puis nous devons déterminer l'**échelle d'analyse** qui sera utilisée (spécifique, populationnelle, individuelle et comportementale) afin d'anticiper le **matériel** nécessaire à l'analyse. Une fois les enregistrements effectués, nous devons identifier les **types de signaux** produits par la faune (transitoires et stationnaires). Puis l'acquisition de données environnementales est nécessaire. Elles peuvent être illustrées par des comportements, des habitats ou encore des nuisances anthropiques. Enfin, la dernière étape est de mettre en place des méthodes d'analyses qui seront bien spécifiques en fonction des étapes précédentes (milieux, échelles, signaux).

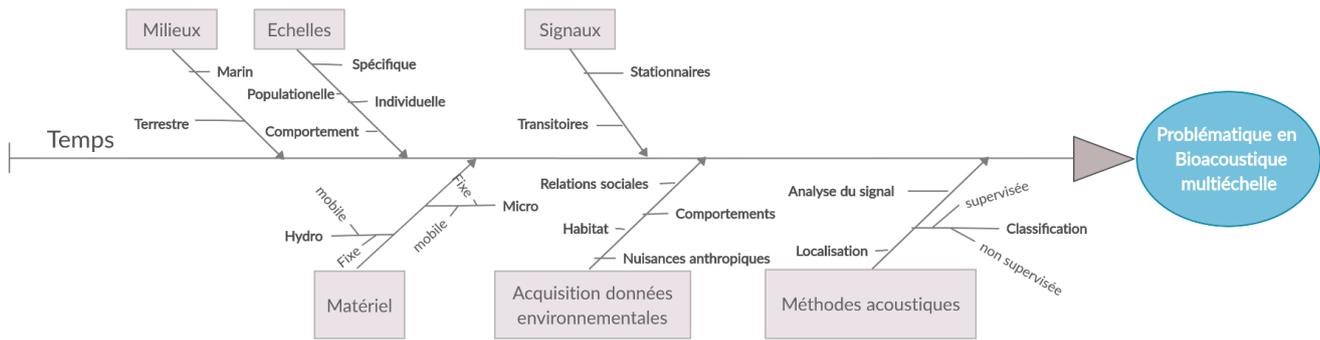


FIG. 5.7 – Diagramme d’Ishikawa représentant les différentes causes et effets pour la mise en place d’une étude bioacoustique au cours du temps.

Différents verrous peuvent encore limiter l’évolution de la bioacoustique. Le premier obstacle qui pourrait être franchi est celui de favoriser les échanges entre les informaticiens, bioacousticiens et les écologues.

La bioacoustique est une science multidisciplinaire nécessitant les compétences de scientifiques provenant de nombreux domaines. Prenons l’exemple de l’extraction individuelle d’Orques à Orcalab. Pour cette étude, nous avons besoin des connaissances comportementales des individus et des annotations des différents enregistrements, réalisées par les experts d’Orcalab. Mais également des connaissances en acoustique et en traitement des signaux pour construire une antenne capable de calculer les TDoAS ont été nécessaires. Enfin, des connaissances en informatique et en statistiques ont été indispensables pour extraire les différentes trajectoires des animaux. Donc le cloisonnement entre ces différentes sciences ne peut que ralentir l’évolution de cette discipline.

Les compétences du ”Data scientist” peuvent être mises à profit des gestionnaires de sites remarquables pour le suivi des espèces, populations et individus. Les compétences des écologues/biologistes peuvent être employées à une expertise sur les espèces et leur évolution. Les interactions entre les différents laboratoires sont donc fondamentales pour la bioacoustique. Le développement de nouvelles équipes au sein des unités de recherches devrait être multi-disciplinaire, afin que les compétences de chacun soient utilisées de manière optimale.

5.7 Perspectives

Différents types de perspectives sont à envisager suite à ce travail de thèse. Les premières concernent l’évolution des différentes méthodes proposées. Puis, des perspectives en bioacoustique sont prévues afin de valider les différentes tendances présentées dans le chapitre Résultats.

5.7.1 Perspectives méthodologiques

Différentes améliorations sont envisagées concernant les performances de localisations individuelles en milieu terrestre et marin.

Le premier protocole marin pour la bioacoustique individuelle concernait la localisation d’Orques devant la station de recherche Orcalab. Les résultats présentés dans ce manuscrit peuvent en

effet être améliorés. La localisation d'individus a pu être réalisée lorsque les individus étaient éloignés de plus de 50m. Mais la plupart du temps, les individus au sein d'une même matriliné se déplaçaient de manière très groupée (voir annexe C).

En général, quelques mètres (voire centimètres) séparent les différents individus, ce qui rend la tâche de localisation individuelle particulièrement difficile.

Nous aimerions donc durant l'été 2021 mettre en place deux antennes multi-hydrophones devant Orcalab afin d'obtenir une localisation beaucoup plus précise. Une carte prévisionnelle est présentée en figure 5.8.



FIG. 5.8 – Carte prévisionnelle de la position des deux antennes pour l'été 2021 à Orcalab ainsi que leur zone d'écoute envisagée

La deuxième antenne serait posée en face de la station d'Orcalab à la même profondeur. La distance entre les deux antennes serait de 1.5 km. Grâce à elle, il serait possible d'avoir l'azimut de chaque individu mais également la distance avec chaque antenne. Nous pourrions synchroniser les deux antennes à l'aide d'émission de PPS (Pulse Par Secondes) émis depuis la surface comme ce qui a été fait dans l'étude de Glotin et al. (2020). Chaque antenne pourrait donc être reliée à une petite bouée de surface équipée d'un PPS afin de les synchroniser.

Puis la bioacoustique individuelle s'applique également à l'écosystème terrestre. La tentative de localisation individuelle présentée dans le chapitre Discussion sera également améliorée afin de comptabiliser précisément le nombre de sources dans un enregistrement. Nous aimerions augmenter la fréquence d'échantillonnage, la distance entre les microphones et utiliser des microphones beaucoup moins directifs pour calculer les TDoAs. Ce nouveau protocole serait installé dans le jardin du Rayol afin d'estimer le nombre d'individus fréquentant le jardin en fonction des cycles journaliers et saisonniers.

5.7.2 Perspectives en bioacoustique

L'ensemble de cette thèse a fourni différentes tendances issues des protocoles mis en place. Afin de consolider nos résultats, il faudrait envisager de reproduire ces différentes manipulations afin de réaliser des tests statistiques sur ces tendances. Pour cela, nous allons continuer les

acquisitions pour l'ensemble des cas d'études.

En effet, la plupart de nos cas d'étude étaient accompagnés d'un seul évènement, et ont donc été testés sur un seul exemple. Deux principales perspectives sont envisageables : la première est d'intensifier les efforts d'enregistrements pour les cas d'étude où nous n'avons pas eu suffisamment d'exemples pour réaliser des tests statistiques. La deuxième perspective serait de réaliser des tests statistiques sur les cas d'études où nous disposons suffisamment d'enregistrements. L'ensemble de ces perspectives est présenté dans le tableau 5.7.2.

TAB. 5.7 – Tableau des différentes perspectives associées à chaque cas d'étude. En rouge les cas d'étude où l'acquisition de nouveaux enregistrements est nécessaire

	Variabilité acoustique	Variabilité environnementale/ comportementale	Perspectives
(1) Espèces d'oiseaux (échelle espèce)	Nombre de détections	Habitat particulier	Intensifier les enregistrements + Tests statistiques
(2) Orques à Orcalab (échelle espèce)	Nombre de détections	Quantité de saumons	Réaliser des tests statistiques
(3) Cachalots sur Bombyx (échelle espèce)	Nombre de détections	Vitesse du courant	Significatif (jour/nuit) Corrélation avec le courant
(4) Pods d'Orques à Orcalab (échelle population)	Diversité des types d'appels	Complexité sociale	Intensifier les enregistrements + Tests statistiques
(5) Cachalot depuis Sphyrna (échelle individus)	Intervalle inter-clic (ICI)	Tortuosité	Intensifier les enregistrements + Tests statistiques
(6) Orques à Orcalab (échelle individus)	Taux d'appel	Éloignement d'un individu avec son groupe	Intensifier les enregistrements + Tests statistiques
(7) Dauphin tacheté pantropical (échelle comportementale)	Fréquence initiale des sifflements	Comportements	Significatif
(8) Baleines à bosse (échelle comportementale)	Taux d'appels	Présence de krill	Réaliser uniquement des tests statistiques

Pour les cas d'études (3), et (7), des tests statistiques ont été réalisés, pour les cas (1), (4), (5), et (6), il faudrait intensifier les enregistrements pour réaliser des tests plus solides, puis pour les cas (2) et (8), il faudra réaliser des tests statistiques.

5.7.2.1 Le projet CARI'MAM

Une des perspectives de cette thèse sera d'appliquer les différentes méthodes développées au projet CARI'MAM pour la détection d'espèces et de populations dans le cadre d'études bioacoustiques.

Le projet CARI'MAM, pour Caribbean Marine Mammals preservation network est un projet qui a vu le jour en 2018. L'objectif principal de ce projet est de renforcer le réseau d'aires marines protégées dans les caraïbes. Pour cela CARI'MAM a mis en place des mesures de protection et d'études d'animaux à grande échelle temporelle et spatiale. Le porteur du projet est le sanctuaire

AGOA sous tutelle de l'Office National de la Biodiversité (OFB). Le sanctuaire AGOA, créé en 2010, est un véritable hotspot de biodiversité en terme de mammifères marins, en effet, plus de 20 espèces de cétacés fréquentent ces eaux (Fig. 5.9). Sa superficie recouvre plus de 138 000 km².

Mais différentes pressions anthropiques peuvent altérer le bien être des animaux, comme le tourisme (Mayol et al., 2015), la pêche, le trafic maritime, la pollution diffuse, les manoeuvres militaires ou encore les déchets présents. C'est pour toutes ces pressions anthropiques qu'il est nécessaire de mettre en place des protocoles de suivis d'espèces, de populations et d'individus afin de protéger cet écosystème unique.

Il y a donc des enjeux de conservation, des enjeux scientifiques, mais également des enjeux de développement économiques, pédagogiques et culturels pour la protection de ces espèces.

Mais au sein de la Grande Région Caraïbe (GWC), les mesures d'études et de protection présentent une forte hétérogénéité. En effet, certains territoires essaient de mettre en place des protocoles respectueux d'approches des mammifères marins (Martinique) alors que d'autres pratiquent encore la chasse à la baleine (St. Vincent et les Grenadines). Depuis plusieurs années, l'interaction entre les humains et les cétacés ne cesse d'accroître, ceci dû aux whale watching (activité commerciale d'observation de cétacés) (Heenehan et al., 2019 ; Gridley et al., 2016 ; Mayol et al., 2015).

Le projet Carimam a donc différents objectifs. Le premier est de fédérer, c'est à dire de créer un réseau d'acteurs impliqués dans la conservation des mammifères marins, puis de renforcer les compétences des gestionnaires sur l'ensemble des îles. Également, la mise en place d'outils de gestion pour permettre une approche respectueuse des animaux. Pour cela, les structures impliquées ont défini des "workpackages" comme les actions de communication et de sensibilisation à la protection des espèces. Le projet vise également à développer l'acquisition de connaissances scientifiques à l'échelle de la Caraïbe, via des transects d'observation ou encore la mise en place d'études acoustiques.

Ce dernier programme aurait pour objectif l'étude de la biodiversité, la saisonnalité des espèces de cétacés sur une grande échelle temporelle et spatiale. Pour cela, la première étape est de mettre en place une harmonisation des protocoles acoustiques sur l'échelle de la Caraïbe.

L'équipe DYNI a donc répondu à l'appel à projet de CARI'MAM et proposé les équipements nécessaires pour effectuer les enregistrements. Le matériel d'acquisition est la carte son Jason développée par SMioT de l'université de Toulon avec l'équipe DYNI. La carte est placée à l'intérieur d'un tube avec une carte SD de 512GB et des batteries. L'hydrophone utilisé est un C75 omnidirectionnel permettant une réponse en haute fréquence de 200 kHz. Afin de former les représentants de chaque île, je me suis rendue avec l'équipe DYNI en République Dominicaine en Octobre 2019. Lors de ce workshop, l'ensemble du matériel a été présenté et une introduction à la bioacoustique fut réalisée. Une concertation a permis de décider communément de l'emplacement des 20 stations acoustiques réparties stratégiquement sur l'ensemble des Caraïbes. Ce réseau sera déployé à partir de fin 2020, et les hydrophones seront placés simultanément autour des sites concernés afin d'enregistrer le paysage acoustique pendant 40 jours consécutifs, tous les 2 mois, durant 1 an. Une fois les deux mois terminés chaque responsable récupère les enregistrements et envoie la carte uSD à notre équipe à l'Université de Toulon.

La figure 5.9 donne les positions GPS des hydrophones posés par chacun des partenaires avec

un exemple de la Guadeloupe.

Nous avons donc défini la position stratégique de chacun des hydrophones en fonction de la disponibilité des partenaires, mais également en fonction de l'intérêt scientifique de la zone. Grâce à ces trois hydrophones en Guadeloupe, il sera possible d'évaluer le sens de déplacement des espèces, populations ou individus (comme à la station de recherche Orcalab).

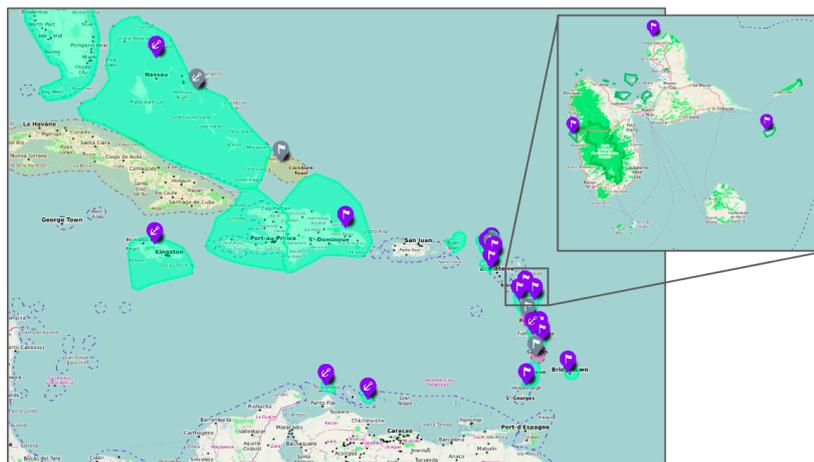


FIG. 5.9 – Carte des positionnements des hydrophones pour le projet CARI'MAM. Les ancres indiquent les points précis définis avec le partenaire, les drapeaux la localisation approximative.

Une fois les enregistrements archivés, nous aimerions nous inspirer des méthodes développées dans cette thèse pour l'analyse des données. Notamment sur l'analyse du sens de déplacement des différentes populations.

Différents objectifs sont à envisager :

- mettre en place des détecteurs/classifieurs automatiques de clics afin d'estimer les mouvements locaux d'odontocètes dans la zone (Ferrari et al., 2020)
- mettre en place des détecteurs/classifieurs automatiques de vocalises (Orques et baleines à bosse) afin d'estimer les périodes de fréquentation des îles par ces espèces passagères
- une fois les espèces détectées, essayer de cartographier leur déplacement sur l'ensemble de la zone Caraïbes
- mettre en corrélation le bruit ambiant (trafic maritime) avec la présence de certaines espèces.
- reproduire ces différentes analyses sur plusieurs années afin d'étudier les cycles saisonniers/annuels des différentes espèces.

Ce projet actuel est donc une future illustration d'une étude bioacoustique à grande échelle temporelle et spatiale qui serait basée sur les méthodes. L'interdisciplinarité de l'équipe DYNI pourra donc mettre en place des méthodes développées dans la thèse pour l'étude des mammifères marins dans la zone Caraïbe, afin d'améliorer leur protection.

Ces différentes perspectives de recherches témoignent de l'intérêt à fournir de nouveaux outils exploitables en bioacoustique. Nous espérons que la communauté scientifique pourra s'appuyer sur nos recherches et ces cas d'étude pour mieux étudier le monde sonore qui nous entoure.

Annexes

A. Liste des espèces aviaires étudiées

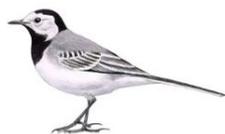
Sources : BIOSONG



Accenteur mouchet



Alouette des champs



Bergeronnette grise



Bruant jaune



Bruant zizi



Chardonneret élégant



Chevêche d'Athéna



Chouette hulotte



Choucas des tours



Corbeau



Cocou gris



Corneille noire



Effraie des clochers



Engoulevent d'Europe



Etourneau sansonnet



Faisan de colchide



Fauvette grisette



Fauvette à tête noire



Geai des chênes



Grive draine



Grimpereau des jardins



Grive musicienne



Hirondelle de fenêtre



Hirondelle rustique



Loriot d'Europe



Martinet noir



Merle noir



Mésange à longue queue



Mésange bleu



Mésange charbonnière



Mésange huppée



Mésange noire



Mésange nonnette



Moineau domestique



Pic épeiche



Pic vert



pie bavarde



Pigeon ramier



Pinson des arbres



Pipit des arbres



Pouillot véloce



Rouge-gorge



Rougequeue noir



Rougequeue à front blanc



Rossignol philomèle



Serin cini



Sittelle torche-pot



Tourterelle des bois



Tourterelle turque



Troglodyte mignon



Verdier d'europe

B. Réponse fréquentielle des hydrophones utilisés

<https://www.cetaceanresearch.com/hydrophones/c75-hydrophone/index.html>

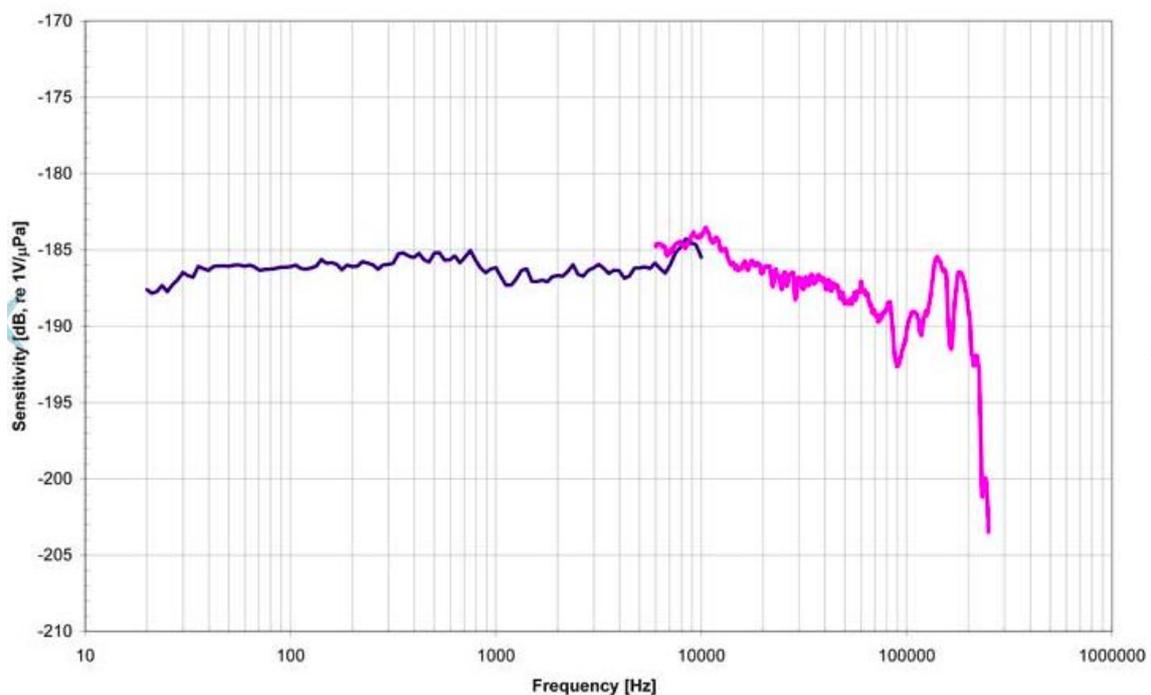
Sensibilité: 3 Hz à 250 kHz

Les hydrophones de la série Cetacean Research™ C75 sont des hydrophones omnidirectionnels préamplifiés.

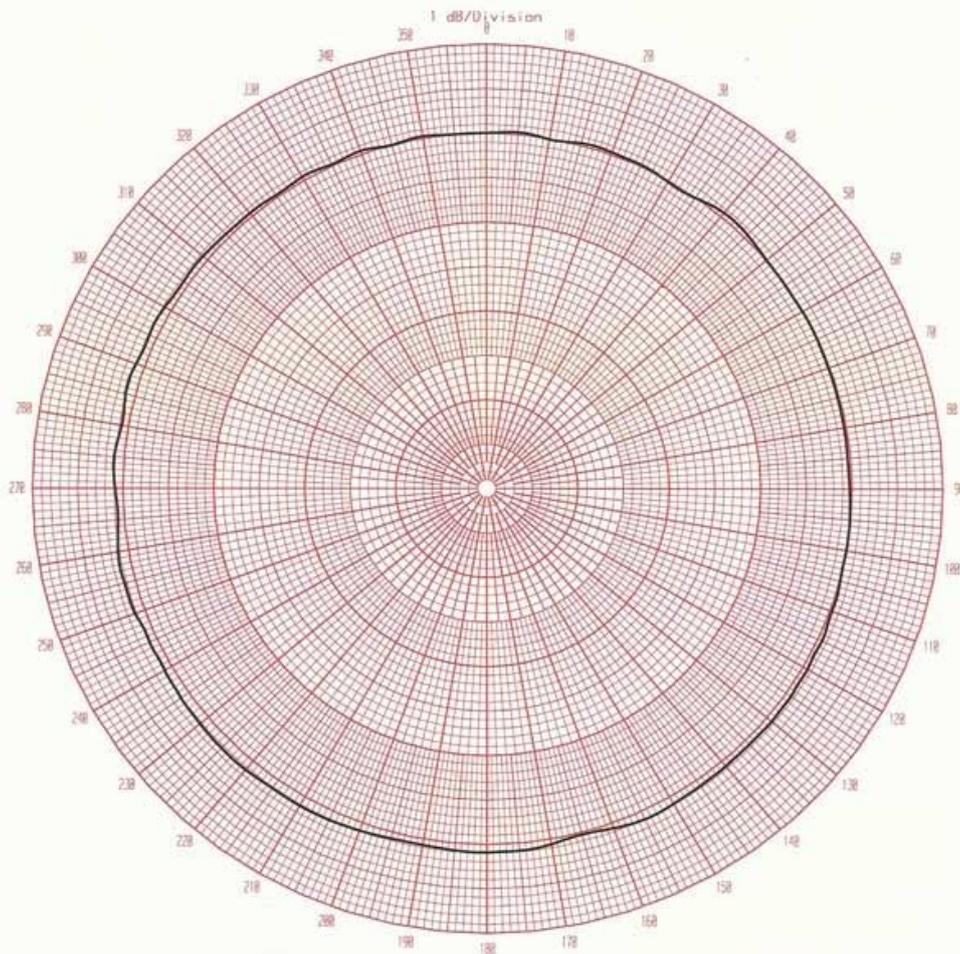
C75 hydrophone specifications	
	C75 / C75X
Linear Frequency Response Range (± 3 dB) [kHz]	0.010 to 170
Usable Frequency Range (+3/-12dB) [kHz]	0.003 to 250
Transducer Sensitivity* [dB, re 1V/ μ Pa]	-200
Preamplifier Gain [db]	20 / 33
Effective Sensitivity* [dB, re 1V/ μ Pa]	-180 / -167
SPL Equiv. Self Noise at 1kHz [dB, re 1 μ Pa/ \sqrt Hz]	51
Power Requirement [Vdc]	5 to 32
RMS Overload Acoustic Pressure [dB, re 1 μ Pa]	184 to 201 / 171 to 188
Maximum Operating Depth ** [m]	920
Operating Temperature Range [°C]	-40 to 85
Output Impedance [Ω]	10
Dimensions [mm]	92L x (25 to 18)dia.
Integral Connector†	Subconn MCBH3MSS
Directionality	omnidirectional

* Transducer Sensitivity + Preamplifier Gain = Effective Sensitivity.
 ** 1,000 meter versions are available. Append "-1k" to the model number to order a hydrophone with a depth rating >1,000 meters.
 *** Connectors are rated to a depth of 7,000 meters; see below (or to the side) for mating connector and cables.

C75/001 Frequency Response



CETACEAN RESEARCH™ C75 hydrophone Frequency Response graph - log scale



Water Temp	11.3 C	RECEIVING DIRECTIVITY RESPONSE		Acoustic Test Facility	
Frequency	50 kHz			NUWC Keyport	
Spacing	3.16 M	C75	SER. # 001	Time	12:48:16
Depth	3.81 M			Date	22 Apr 2010
Plane	XY			Technician	MF
Console	1	High Amp:-8.7 High angle: 228.2 Low Amp:-10.3 Low Angle: 88.6 Total Plot Variation is: 1.6 dB			

C57

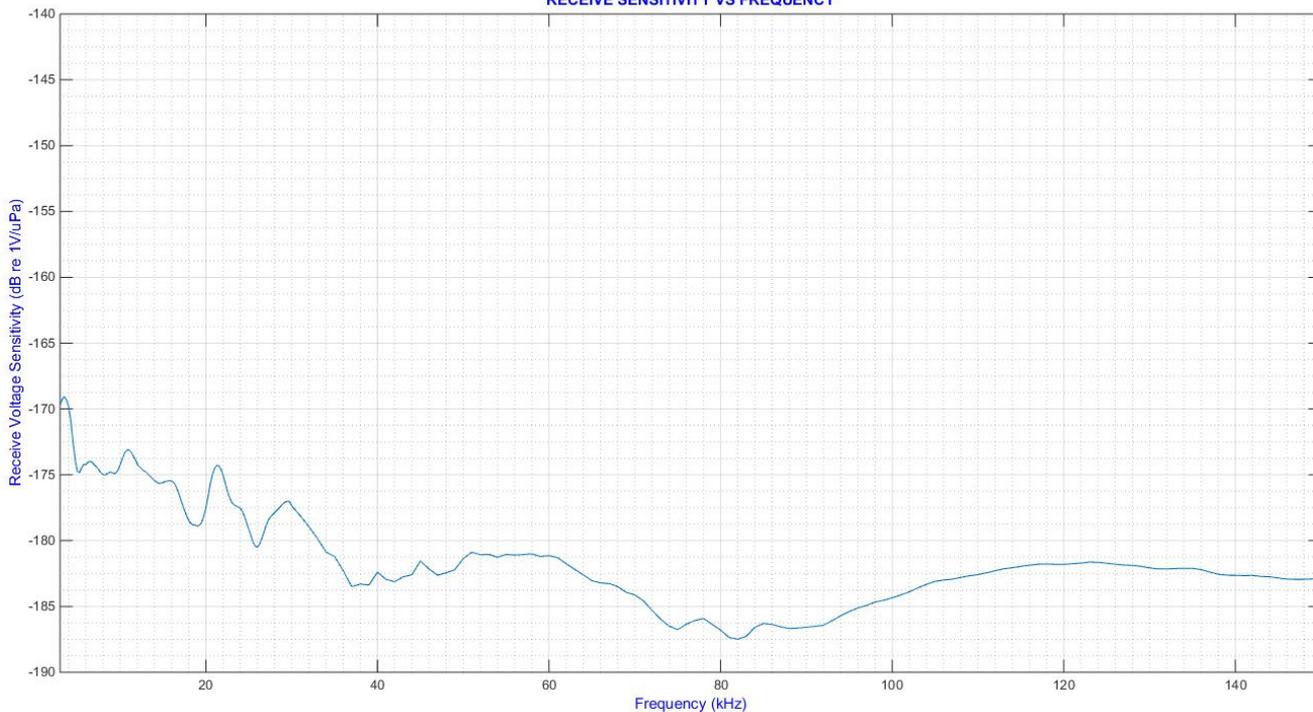
<https://www.cetaceanresearch.com/hydrophones/c57-hydrophone/index.html>

Sensibilité: 8 Hz à 100 kHz

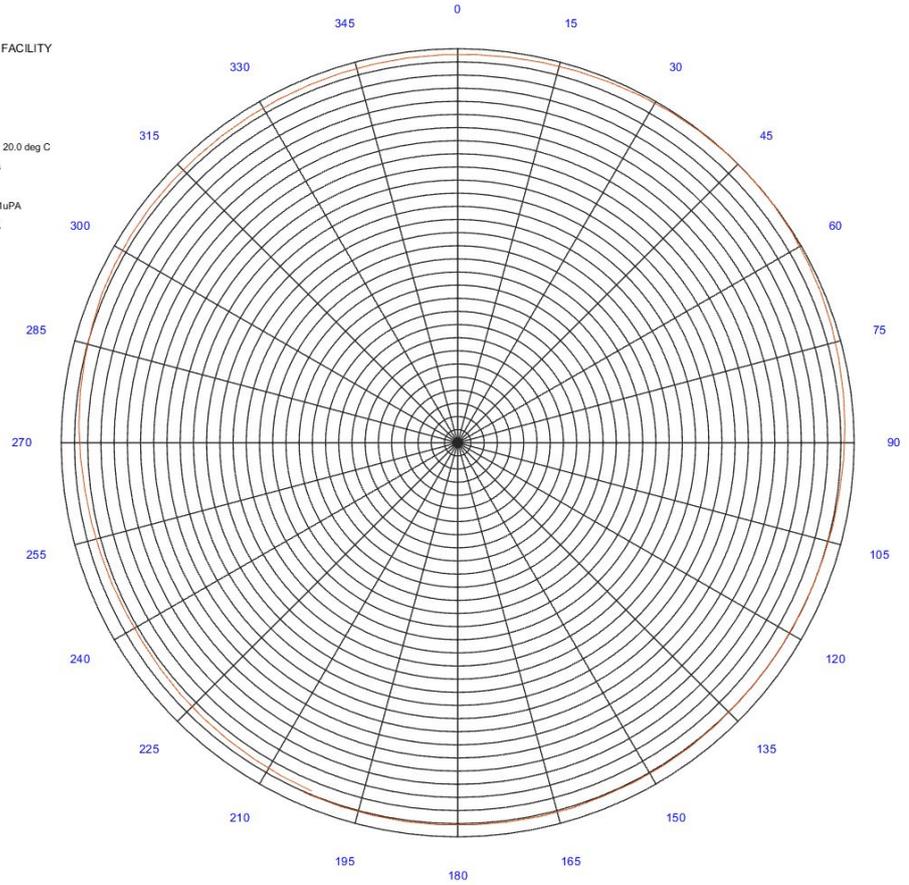
C57 hydrophone specifications		
	C57 / C57X	C57RS / C57XRS
Linear Frequency Range (± 3dB) [kHz]	0.015 to 45	0.015 to 50 & 124 to 250+
Usable Frequency Range (+3/-12dB) [kHz]	0.008 to 100	0.008 to 77 & 96 to 250+
Transducer Sensitivity* [dB, re 1V/μPa]	-187	-200
Preamplifier Gain [dB]	20 / 33	20 / 33
Effective Sensitivity* [dB, re 1V/μPa]	-167 / -154	-180 / -167
SPL Equiv. Self Noise at 1kHz [dB, re 1μPa/\sqrtHz]	46 (<i>Sea State Zero</i>)	61
Power Requirement [Vdc]	5 to 32	5 to 32
RMS Overload Acoustic Pressure [dB, re 1μPa]	171 to 188 / 158 to 175	184 to 201 / 171 to 188
Maximum Operating Depth [m]	370	920
Operating Temperature Range [°C]	-40 to 60	-40 to 60
Output Impedance [Ω]	10	10
Dimensions [mm]	116L x 25dia.	116L x 25dia.
Integral Connector ***	Subconn MCBH3MSS	Subconn MCBH3MSS

* Transducer Sensitivity + Preamplifier Gain = Effective Sensitivity.
 *** Connectors are rated to a depth of 7,000 meters; see below (or to the side) for mating connector and cables.

RECEIVE SENSITIVITY VS FREQUENCY



UNIVERSITY OF WASHINGTON
APPLIED PHYSICS LABORATORY
R/V HENDERSON ACOUSTIC TEST FACILITY
DATE: 5/14/2018
TIME: 13:39:03
REF TRANSDUCER = F37 Transducer
REF TRANSDUCER S/N = A51
REF TRANSDUCER CAL TEMPERATURE = 20.0 deg C
WAVEFORM DRIVE VOLTAGE = 1.00 Vrms
FREQUENCY = 10.0 kHz
PROJ SOURCE LEVEL AT 0 = 160.0 dB re 1uPA
RECEIVER LEVEL AT 0 = 150.0 dB re 1uPA
RANGE = 3.2 meters
DEPTH = 2.1 meters
PING AVG COUNT = 2
WATER TEMPERATURE = 17.8 deg C
ANGLE STEP = 5.0 deg
SCALE = 0 to -30 dB

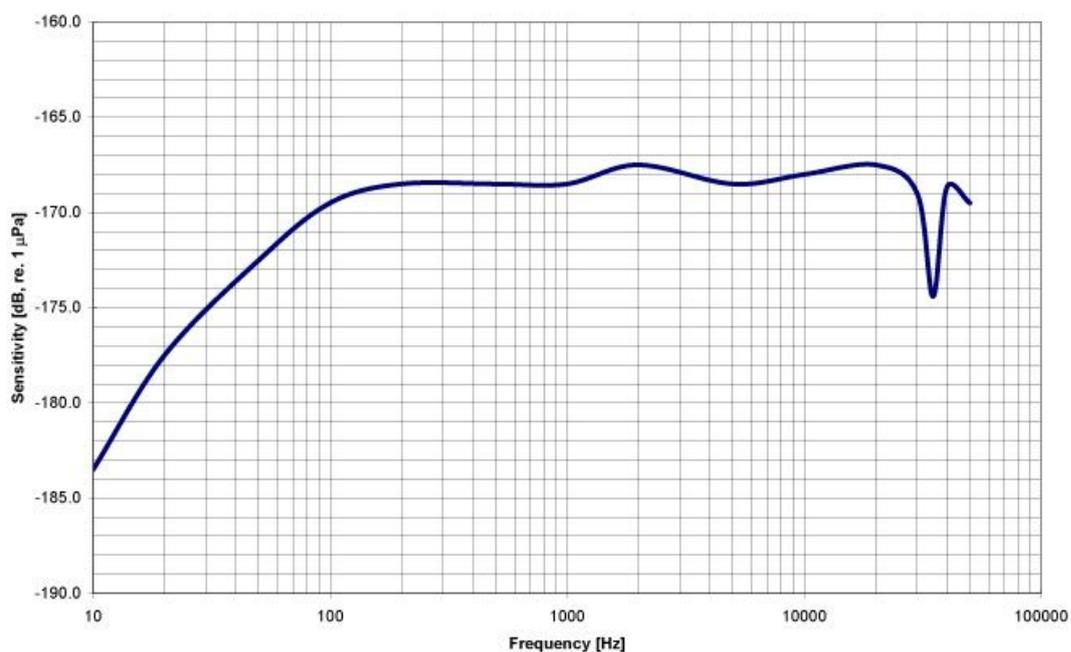


C57 10kHz Beam Pattern SN806
DATA FILE NAME: 10kHz Beam.PAT

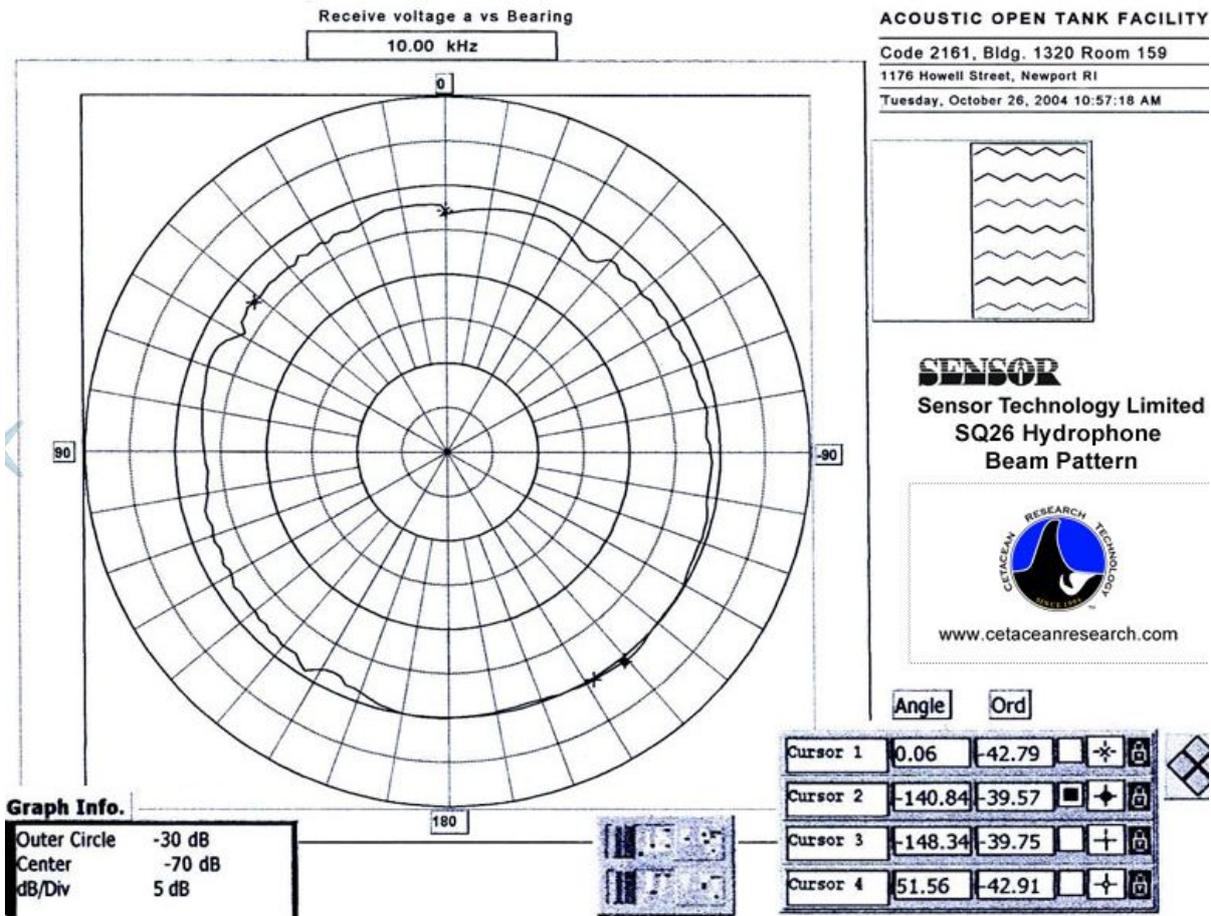
Sensibilité: 20 Hz à 45 kHz

SQ26-01 hydrophone specifications	
Linear Frequency Range (± 1 dB) [kHz]	.001† to 28
Usable Frequency Range (+3/-12dB) [kHz]	0.001† to > 50
Transducer Voltage Sensitivity @ 20°C [dB, re 1V/ μ PA]	-193.5
Transducer Voltage Sensitivity @ 20°C [V/bar]	20
Transducer Charge Sensitivity @ 20°C [nC/bar]	24
Acceleration Sensitivity [mbar/g]	< 0.2 when properly mounted
Maximum Operating Depth [m]	down to 2,000
Operating Temperature Range [°C]	-30 to 60
Capacitance (± 10 %) @ 20 °C [nF]	1.4
Sensitivity Variation with Temperature	less than 1dB loss from 0 to 35 °C
Capacitance Variation with Temperature	0.33% increase per °C
Capacitance Variation with Depth	7% loss per 1,000 m
Dimensions [mm]	25.4 L x 25.4 dia.
Mass [grams]	16
Electrical Leads	two-wire shielded, 28AWG, 30cm (12") long
Shielding	integral Faraday cage type
Electrical Insulation [Ω]	> 500 M
Water Blocked Leads	No

SQ26-07 Frequency Response



Beam: 10kHz



Plot of Beam Pattern for the SQ26-01 Hydrophone: 10kHz.

HN1 Neptune D140

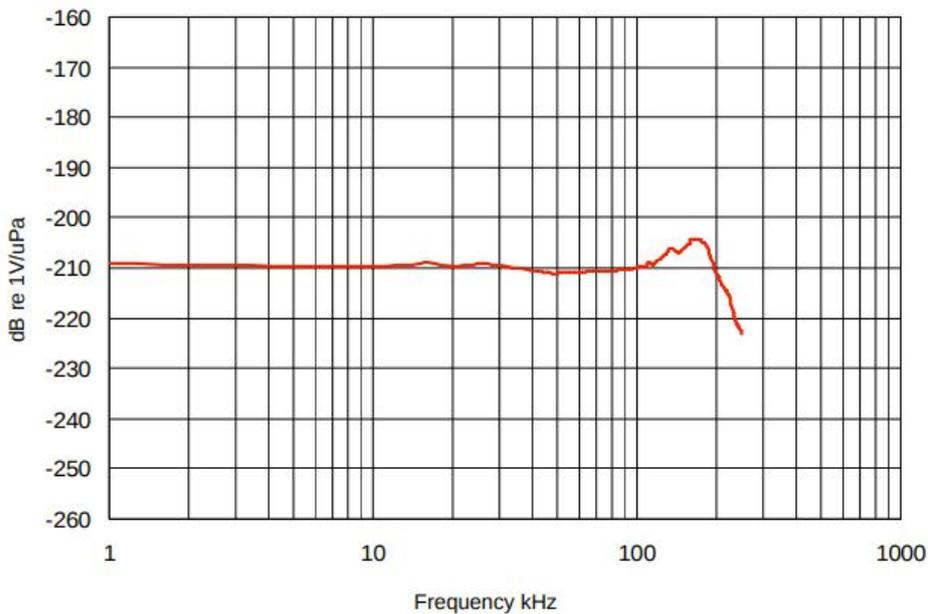
Sensibilité: 100 Hz 40 kHz

<http://www.neptune-sonar.co.uk/product-category/standard-transducer-products/hydrophone/s/>

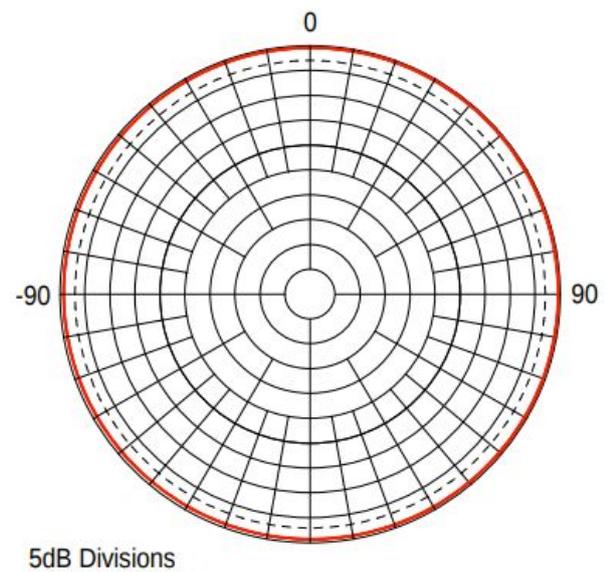
Technical Specification

Resonant Frequency	150 kHz (Nominal)
Beam Pattern	Omni \pm 2dB up to 160kHz
Receive Sensitivity	See Graph
Transmit Sensitivity	See Graph
Capacitance at 1 kHz	3500pF
Input Power	100 Watts around resonance
Operating Depth	1500 Metres
Operating Temperature	-5 to +40°C
Storage Temperature	-40 to +80°C
Cable Type	Polyurethane \varnothing 6mm Low Noise Coaxial
Cable Length	10 metres standard Additional lengths supplied to order

Receive Graph



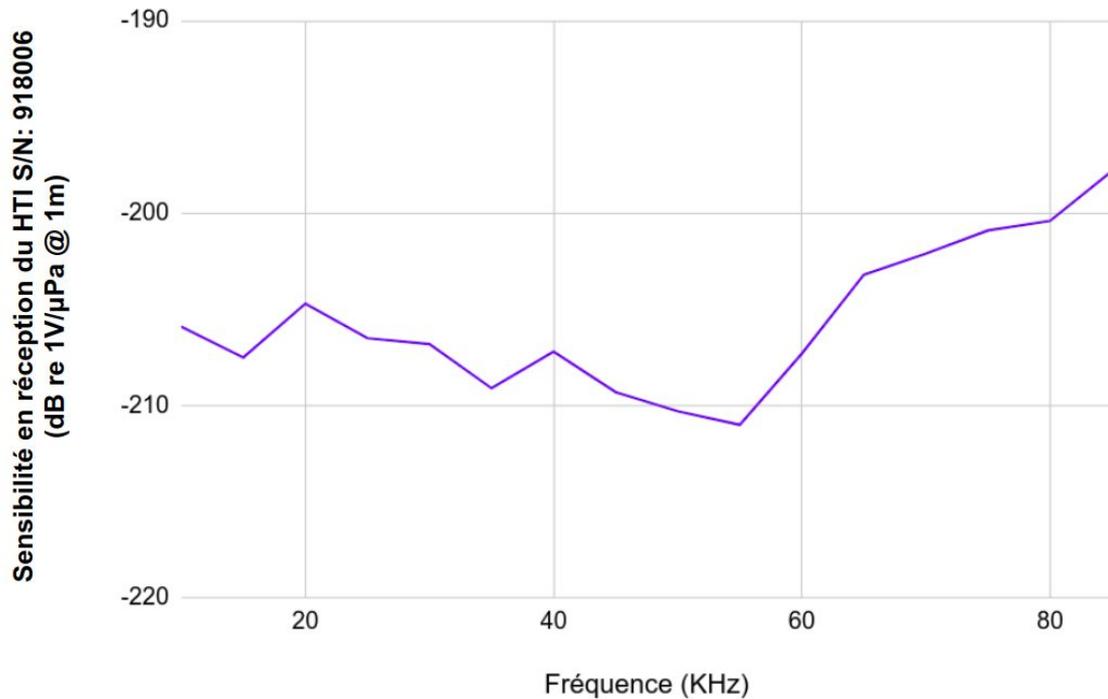
Beam Pattern at 150 kHz



hydrophone HTI

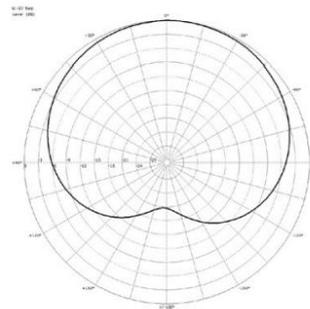
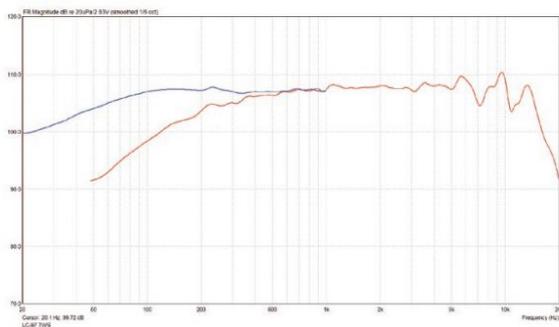
Sensibilité: 100 Hz à 80 kHz

<http://www.hightechincusa.com/products/hydrophones/hti96min.html>



Microphone Tbone LC 97 TWS

https://images.static-thomann.de/pics/atg/atgdata/document/specs/183388_c_183388_datasheet_v4_en_online.pdf



Polar pattern @ 1 kHz

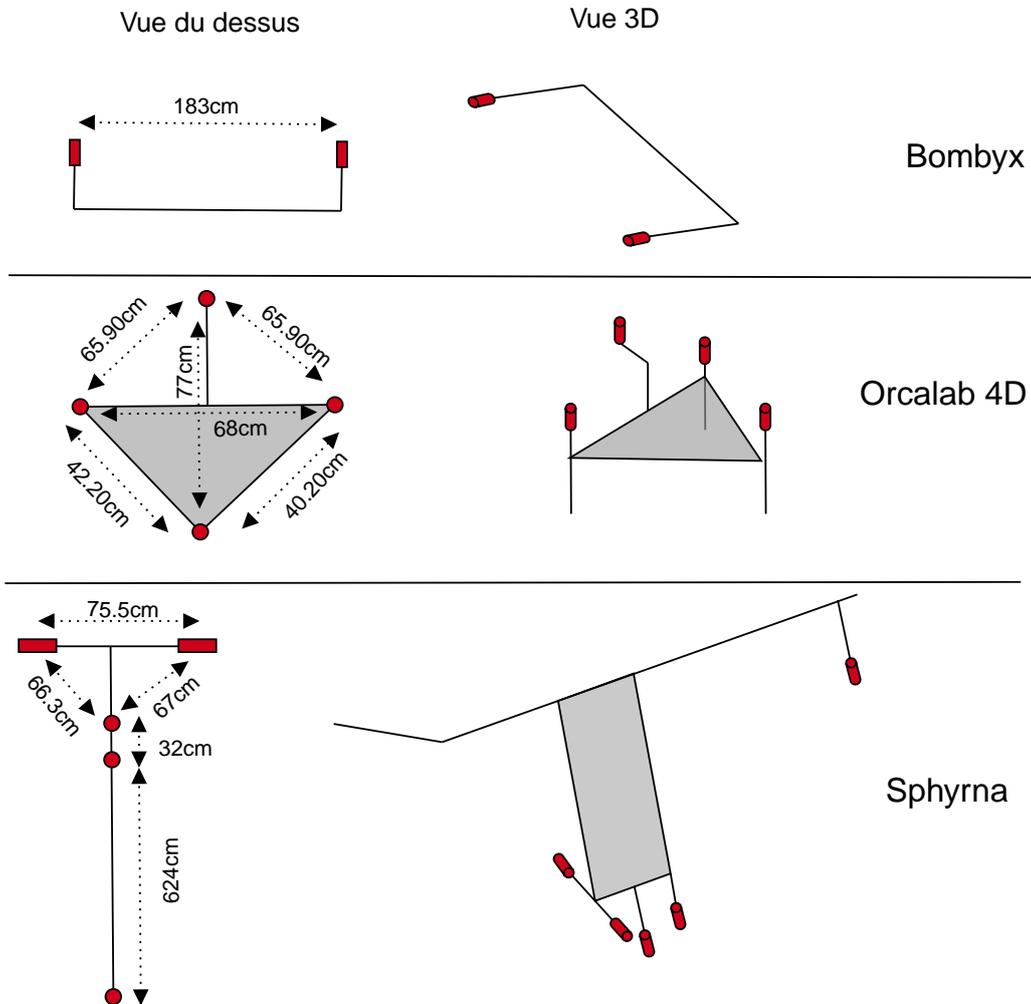
Connection / Supply	Mini-XLR / 1.5 V – 10.0 V
Transducer concept	Uni-directional electret condenser
Directivity	cardioid
Field sensitivity [mV / Pa]	4.46
Field sensitivity [dB re 1V / Pa]	-47 ± 4
Equivalent noise level [dB(A)]	< 34
Signal-to-noise ratio [dB(A)]	> 60
Maximum SPL [dB SPL @ 1% THD]	110
Electric impedance [kOhm @ 1 kHz]	2.2
Current draw [mA]	0.5
Frequency response [Hz, 0.1 m / -10 dB]	100 – 20.000
Dimensions [mm]	13 × 37
Weight [g]	19

C. Exemple d'un passage de deux matriline (A54 et A50) devant l'antenne acoustique (environ 500 m)



D. Schéma représentant les dimensions des différentes antennes mises en place.

- Haut : Bouée Bombyx (espèces)
- Milieu : Antenne à Orcalab (individus)
- Bas : Réseau d'hydrophones sous le drone Sphyrna (individus)

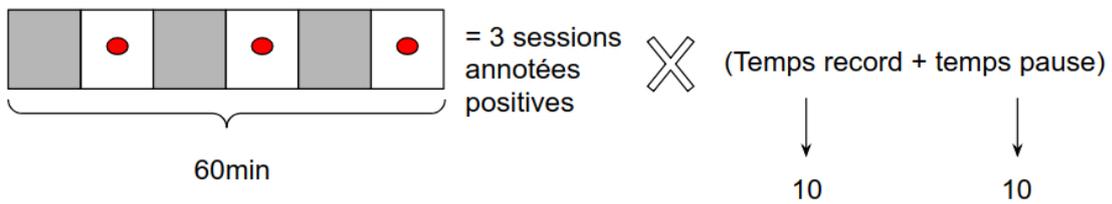


E. Schéma représentant la normalisation des enregistrements réalisés sur Bombyx pour la détection de Cachalots

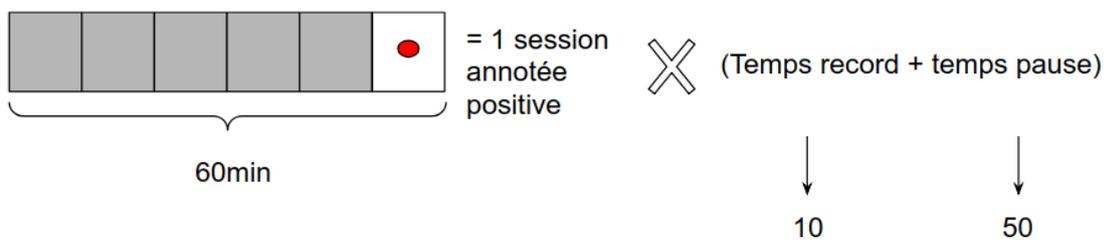
Lors de nos résultats, nous avons multiplié l'ensemble de nos sessions annotées positives par *Beta* qui est égale au Temps d'enregistrements + le Temps de pause. Le schéma suivant récapitule comment ce *Beta* a été calculé. Prenons l'exemple ou un Cachalot était présent pendant 1h sur Bombyx. Pour le premier cas, nous avons 10min d'enregistrement, 10min de pause, ce qui donne 3 sessions annotées positives à la présence de Cachalot. Multiplié par *Beta*, cela donne 60 détections.

Puis, durant le deuxième cas, nous enregistrons 10min et faisons 50min de pause, ce qui fait 1 session annotée positive. Multipliée par *Beta*, on obtient également 60 détections.

Ce procédé nous permet donc de pouvoir comparer les sessions entre elles, même si elles n'ont pas les mêmes efforts d'enregistrements.



$$3 \times 20 = 60 \text{ détections}$$



$$1 \times 60 = 60 \text{ détections}$$

F. Métriques des performances en détection et classification

Dans cette thèse, différents classifieurs et détecteurs ont été mis en place. Afin d'estimer leurs performances, différentes mesures peuvent être utilisées. Les deux premières sont la **précision** et le **rappel** pour caractériser un détecteur. Ces dernières sont expliquées dans la figure suivante. Prenons l'exemple d'un détecteur de clics de Cachalots. L'ensemble des clics détectés est rassemblé dans le cercle. Parmi eux, il y a des vrais positifs (VP) et des faux positifs (FP). A l'inverse, les transitoires non détectés peuvent être des clics de bruit (Vrais négatifs : VN) et des clics de Cachalots faux négatifs (FN). La précision égale le rapport des vrais positifs sur la somme des vrais positifs et des faux positifs. Tandis que le rappel est le rapport des vrais positifs sur la somme des vrais positifs et des faux négatifs. Il faut donc trouver un bon équilibre entre une bonne précision (qui signifie peu de faux positifs) et un bon rappel (peu de faux négatifs).

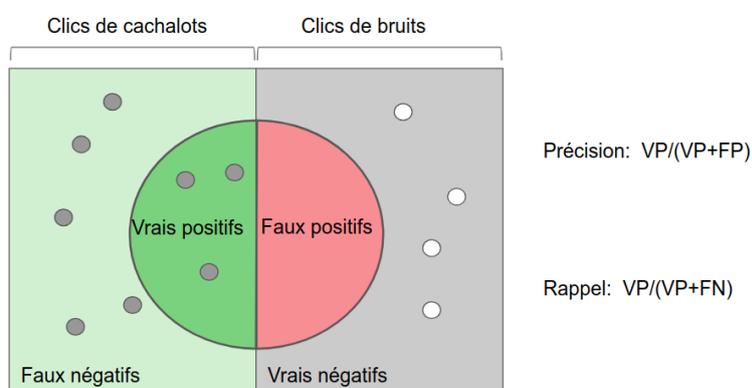


FIG. 5.10 – Résumé des notions de précision et rappel dans le cadre d'une détection de clics de Cachalots. Les transitoires détectés par notre modèle sont à l'intérieur du cercle (vrais positifs + faux positifs).

La troisième mesure de performance est le Taux de Bonne Classification (TBC). Généralement en pourcentage, il calcule la somme des vrais positifs et des faux positifs sur l'ensemble des échantillons :

$$TBC = \frac{(VP + VN)}{(VP + VN + FP + FN)} \quad (5.1)$$

La moyenne des performances possibles d'un détecteur est donnée également par l'aire (AUC) sous la courbe ROC (Receiver Operating Characteristic), qui est la courbe des performances d'un détecteur : son taux de VP en fonction du taux FP.

Bibliographie

- Aarabi, P. (2003). The fusion of distributed microphone arrays for sound localization. *EURASIP Journal on Advances in Signal Processing*, 2003(4) :860465.
- Abeille, R., Doh, Y., Dufour, O., Giraudet, P., Halkias, X., Glotin, H., Prevot, J., Rabouy, C., and Razik, J. (2012). Détection et classification sur transect audiovisuel de populations de cétacés du nord pelagos aux îles d'or. Technical report, PELAGOS International Sanctuary. http://sabiiod.org/pub/PELAGOS_DECAV.pdf.
- Abeille, R., Doh, Y., Giraudet, P., Glotin, H., Prevot, J.-M., and Rabouy, C. (2014). Estimation robuste par acoustique passive de l'intervalle-inter-pulse des clics de *Physeter macrocephalus* : méthode et application sur le parc national de Port-Cros. *Journal of the Scientific Reports of Port-Cros National Park*, 28.
- Acevedo, M. A. and Villanueva-Rivera, L. J. (2006). From the field : Using automated digital recording systems as effective tools for the monitoring of birds and amphibians. *Wildlife Society Bulletin*, 34(1) :211-214.
- Adam, O., Cazau, D., Gandilhon, N., Fabre, B., Laitman, J. T., and Reidenberg, J. S. (2013). New acoustic model for humpback whale sound production. *Applied Acoustics*, 74(10) :1182-1190.
- Aguilar, S., Johnson, M., Madsen, P., Tyack, P., Bocconcelli, A., and Fabrizio, B. (2006). Does intense ship noise disrupt foraging in deep-diving cuvier's beaked whales ? *Marine Mammal Science*, 22(3) :690-699.
- Alberts, A. C. (1992). Constraints on the design of chemical communication systems in terrestrial vertebrates. *The American Naturalist*, 139 :S62-S89.
- Alexandridis, A. and Mouchtaris, A. (2015). Multiple sound source location estimation and counting in a wireless acoustic sensor network. In *2015 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, pages 1-5. IEEE.
- Almeida, P. J., Vieira, M. V., Kajin, M., Forero-Medina, G., and Cerqueira, R. (2010). Indices of movement behaviour : conceptual background, effects of scale and location errors. *Zoologia (Curitiba)*, 27(5) :674-680.
- Amano, M. and Yoshioka, M. (2003). Sperm whale diving behavior monitored using a suction-cup-attached tdr tag. *Marine Ecology Progress Series*, 258 :291-295.
- American Cetacean Society (2017). American cetacean society. https://www.acsonline.org/index.php?option=com_content&view=article&id=72:spotted-dolphin&catid=20:site-content. Last checked on Sept 01, 2019.
- Anderson, R. C., Searcy, W. A., Peters, S., and Nowicki, S. (2008). Soft song in song sparrows : acoustic structure and implications for signal function. *Ethology*, 114(7) :662-676.
- Au, W. W. (1993). *The sonar of dolphins*. Springer Science & Business Media.
- Au, W. W., Carder, D. A., Penner, R. H., and Scronce, B. L. (1985). Demonstration of adaptation in beluga whale echolocation signals. *The Journal of the Acoustical Society of America*, 77(2) :726-730.
- Au, W. W. and Lammers, M. O. (2016). *Listening in the Ocean*. Springer.
- Audacity-Team (2018). Audacity : free audio editor and recorder.
- Azar, J. F. and Bell, B. D. (2016). Acoustic features within a forest bird community of native and introduced species in new zealand. *Emu-Austral Ornithology*, 116(1) :22-31.
- Bain, D. (1986). Acoustic behavior of orcinus : Sequences, periodicity, behavioral correlates and an automated technique for call classification. *Behavioral biology of Killer whales*, pages 335-371.

- Barca, B., Lindon, A., and Root-Bernstein, M. (2016). Environmentalism in the crosshairs : Perspectives on migratory bird hunting and poaching conflicts in Italy. *Global ecology and conservation*, 6 :189–207.
- Bardeli, R., Wolff, D., Kurth, F., Koch, M., Tauchert, K.-H., and Frommolt, K.-H. (2010). Detecting bird sounds in a complex acoustic environment and application to bioacoustic monitoring. *Pattern Recognition Letters*, 31(12) :1524–1534.
- Barnard, C. and Burk, T. (1979). Dominance hierarchies and the evolution of “individual recognition”. *Journal of theoretical Biology*, 81(1) :65–73.
- Barragán-Barrera, D. C., Alejandra, C.-C. P., Farias-Curtidor, N., Lancheros-Neva, R., Botero-Acosta, N., Buenos, P., Moreno, I., Bolanos-Jimenez, J., Bouveret, L., Nataly, C.-M. D., et al. (2019). Ecological niche modeling of three species of *Stenella* dolphins in the Caribbean basin, with application to the seaflower biosphere reserve. *Marine Conservation and Sustainability*.
- Bee, M. A. and Gerhardt, H. C. (2002). Individual voice recognition in a territorial frog (*Rana catesbeiana*). *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 269(1499) :1443–1448.
- Beecher, M. D. (1988). Spectrographic analysis of animal vocalizations : implications of the “uncertainty principle”. *Bioacoustics*, 1(2-3) :187–208.
- Beer, C. G. (1971). Individual recognition of voice in the social behavior of birds. In *Advances in the Study of Behavior*, volume 3, pages 27–74. Elsevier.
- Bénard-Caudal, F., Giraudet, P., and Glotin, H. (2010). Whale 3d monitoring using astrophysic nemo onde two meters wide platform with state optimal filtering by Rao-Blackwell Monte Carlo data association. *Applied Acoustics*, 71(11) :994–999.
- Benoit-Bird, K. J., Dahood, A. D., and Würsig, B. (2009). Using active acoustics to compare lunar effects on predator–prey behavior in two marine mammal species. *Marine Ecology Progress Series*, 395 :119–135.
- Benton, T. G., Bryant, D. M., Cole, L., and Crick, H. Q. (2002). Linking agricultural practice to insect and bird populations : a historical study over three decades. *Journal of Applied Ecology*, 39(4) :673–687.
- Bergler, C., Schröter, H., Cheng, R. X., Barth, V., Weber, M., Nöth, E., Hofer, H., and Maier, A. (2019). Orca-spot : An automatic killer whale sound detection toolkit using deep learning. *Scientific reports*, 9(1) :1–17.
- Bermant, P. C., Bronstein, M. M., Wood, R. J., Gero, S., and Gruber, D. F. (2019). Deep machine learning techniques for the detection and classification of sperm whale bioacoustics. *Scientific reports*, 9(1) :1–10.
- Best, P., Ferrari, M., Poupard, M., Paris, S., Marxer, R., Symonds, H., Spong, P., and Glotin, H. (2020). Deep learning and domain transfer for orca vocalization detection. In *International joint conference on neural networks*.
- Best, P., Findlay, K., Sekiguchi, K., Peddemors, V., Rakotonirina, B., Rossouw, A., and Gove, D. (1998). Winter distribution and possible migration routes of humpback whales *Megaptera novaeangliae* in the southwest Indian ocean. *Marine Ecology Progress Series*, 162 :287–299.
- Best, P. B. (1979). Social organization in sperm whales, *Physeter macrocephalus*. In *Behavior of marine animals*, pages 227–289. Springer.
- Bianco, M. J., Gerstoft, P., Traer, J., Ozanich, E., Roch, M. A., Gannot, S., and Deledalle, C.-A. (2019). Machine learning in acoustics : Theory and applications. *The Journal of the Acoustical Society of America*, 146(5) :3590–3628.
- Bigg, M., Olesiuk, P., Ellis, G. M., Ford, J., and Balcomb, K. C. (1990). Social organization and genealogy of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington state. *Report of the International Whaling Commission*, 12 :383–405.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J. L., Krakauer, A. H., Clark, C., Cortopassi, K. A., et al. (2011). Acoustic monitoring in terrestrial environments using microphone arrays : applications, technological considerations and prospectus. *Journal of Applied Ecology*, 48(3) :758–767.

Bibliographie

- Borker, A. L., Halbert, P., Mckown, M. W., Tershy, B. R., and Croll, D. A. (2015). A comparison of automated and traditional monitoring techniques for marbled murrelets using passive acoustic sensors. *Wildlife Society Bulletin*, 39(4) :813–818.
- Borker, A. L., McKown, M. W., Ackerman, J. T., EAGLES-SMITH, C. A., Tershy, B. R., and Croll, D. A. (2014). Vocal activity as a low cost and scalable index of seabird colony size. *Conservation biology*, 28(4) :1100–1108.
- Borsani, J. F. and Pavan, G. (1994). Oral communication : Acoustics and aspects of Sperm whale (*Physeter macrocephalus*) behaviour and ecology in the mediterranean sea. *Italian journal of Zoology*.
- Bouchard, B., Barnagaud, J.-Y., Poupard, M., Glotin, H., Gauffier, P., Ortiz, S. T., Lisney, T. J., Campagna, S., Rasmussen, M., and Célérier, A. (2019). Behavioural responses of humpback whales to food-related chemical stimuli. *PLoS one*, 14(2) :e0212515.
- Boughman, J. W. and Moss, C. F. (2003). Social sounds : vocal learning and development of mammal and bird calls. In *Acoustic communication*, pages 138–224. Springer.
- Bradbury, J. W., Vehrencamp, S. L., et al. (1998). *Principles of animal communication*. Sinauer, 2011.
- Branson, S., Van Horn, G., Belongie, S., and Perona, P. (2014). Bird species categorization using pose normalized deep convolutional nets. *arXiv preprint arXiv :1406.2952*.
- Brill, R. L., Sevenich, M. L., Sullivan, T. J., Sustman, J. D., and Witt, R. E. (1988). Behavioral evidence for hearing through the lower jaw by an echolocating dolphin (*tursiops truncatus*). *Marine Mammal Science*, 4(3) :223–230.
- Brown, J. L., Li, S.-H., and Bhagabati, N. (1999). Long-term trend toward earlier breeding in an american bird : A response to global warming ? *Proceedings of the National Academy of Sciences*, 96(10) :5565–5569.
- Brumm, H. and Slabbekoorn, H. (2005). Acoustic communication in noise. *Advances in the Study of Behavior*, 35 :151–209.
- Brumm, H., Voss, K., Köllmer, I., and Todt, D. (2004). Acoustic communication in noise : regulation of call characteristics in a new world monkey. *Journal of Experimental Biology*, 207(3) :443–448.
- Burham, R., Palm, R., Duffus, D., Mouy, X., and Riera, A. (2016). The combined use of visual and acoustic data collection techniques for winter Killer whale (*Orcinus orca*) observations. *Global ecology and conservation*, 8 :24–30.
- Busnel, R. and Dziedzic, A. (1966). Acoustic signals of the pilot whale *Globicephala melaen* and of the porpoises *Delphinus delphis* and *Phocoena phocoena*. *Whales, dolphins and porpoises*, pages 607–646.
- Buxton, R. T. and Jones, I. L. (2012). Measuring nocturnal seabird activity and status using acoustic recording devices : applications for island restoration. *Journal of Field Ornithology*, 83(1) :47–60.
- Calupca, T., Fristrup, K., and Clark, C. (2000). A compact digital recording system for autonomous bioacoustic monitoring. *The Journal of the Acoustical Society of America*, 108(5) :2582–2582.
- Camacho, A., García-Rodríguez, A., and Bolaños, F. (2011). Automatic detection of vocalizations of the frog diasporus hylaeformis in audio recordings. In *Proceedings of Meetings on Acoustics 162ASA*, volume 14, page 010003. Acoustical Society of America.
- Campbell, G. S. (2004). Quantitative comparison of bottlenose dolphin (*Tursiops ssp*) whistles from three geographic regions. *MS thesis*.
- Caruso, F., Sciacca, V., Bellia, G., De Domenico, E., Larosa, G., Papale, E., Pellegrino, C., Pulvirenti, S., Riccobene, G., Simeone, F., et al. (2015). Size distribution of sperm whales acoustically identified during long term deep-sea monitoring in the ionian sea. *PLoS One*, 10(12) :e0144503.
- Catchpole, C. K. and Slater, P. J. (2003). *Bird song : biological themes and variations*. Cambridge university press.
- Cavalli-Sforza, L. L., Piazza, A., Menozzi, P., and Mountain, J. (1988). Reconstruction of human evolution : bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences*, 85(16) :6002–6006.

- Célérier, A., Bon, C., Malapert, A., Palmas, P., and Bonadonna, F. (2011). Chemical kin label in seabirds. *Biology letters*, 7(6) :807–810.
- Cerchio, S. and Dahlheim, M. (2001). Variation in feeding vocalizations of humpback whales *Megaptera novaeangliae* from southeast alaska. *Bioacoustics*, 11(4) :277–295.
- Chapuis, C. (1971). Un exemple de l'influence du milieu sur les émissions vocales des oiseaux : l'évolution des chants en forêt équatoriale. *La terre et la vie*.
- Charrier, I. and Harcourt, R. G. (2006). Individual vocal identity in mother and pup australian sea lions (*Neophoca cinerea*). *Journal of Mammalogy*, 87(5) :929–938.
- Chesmore, D. (2004). Automated bioacoustic identification of species. *Anais da Academia Brasileira de Ciências*, 76(2) :436–440.
- Cholewiak, D. M., Cerchio, S., Jacobsen, J. K., Urbán-R, J., and Clark, C. W. (2018). Songbird dynamics under the sea : acoustic interactions between humpback whales suggest song mediates male interactions. *Royal Society open science*, 5(2) :171298.
- Cholewiak, D. M., Sousa-Lima, R. S., and Cerchio, S. (2013). Humpback whale song hierarchical structure : Historical context and discussion of current classification issues. *Marine Mammal Science*, 29(3) :E312–E332.
- Clayton, N. (1988). Song learning and sexual preference in estrildids raised by two species. *Animal Behaviour*, 36 :1589–1600.
- Clemins, P. J. and Johnson, M. T. (2002). Automatic type classification and speaker identification of african elephant vocalizations. *Animal behavior*.
- Conner, D. A. (1982). Dialects versus geographic variation in mammalian vocalizations. *Animal Behaviour*.
- Cook, M., Sayigh, L., Blum, J., and Wells, R. (2004). Signature-whistle production in undisturbed free-ranging bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society of London-B*, 271(1543) :1043–1050.
- Council, N. R. et al. (2003). *Ocean noise and marine mammals*. National Academies Press.
- Dadouchi, F., Villa, A., Gervaise, C., Ioana, C., Huillery, J., and Mars, J. (2015). Détection statistique temps-fréquence dans un contexte acoustique à faible rsb. In *GRETSI 2015*.
- Dalheim, M., Schulman-Janiger, A., Black, N., Ternullo, R., Ellifrit, D., and Balcomb III, K. (2008). Eastern temperate north pacific offshore killer whales (orcinus orca) : Occurrence, movements, and insights into feeding ecology. *Publications, Agencies and Staff of the U.S. Department of Commerce*. 43.
- Danishevskaya, A. Y., Filatova, O. A., Samarra, F. I. P., Miller, P. J. O., Ford, J. K., Yurk, H., Matkin, C. O., and Hoyt, E. (2020). Crowd intelligence can discern between repertoires of Killer whale ecotypes. *Bioacoustics*, 29(1) :15–27.
- Das, A., Kumar, A., and Bahl, R. (2013). Marine vessel classification based on passive sonar data : the cepstrum-based approach. *IET Radar, Sonar & Navigation*, 7(1) :87–93.
- Dawson, D. K. and Efford, M. G. (2009). Bird population density estimated from acoustic signals. *Journal of Applied Ecology*, 46(6) :1201–1209.
- de Jong, K., Amorim, M., Fonseca, P. J., and Heubel, K. U. (2018). Noise affects multimodal communication during courtship in a marine fish. *Frontiers in Ecology and Evolution*, 6 :113.
- De Mauro, A., Greco, M., and Grimaldi, M. (2016). A formal definition of big data based on its essential features. *Library Review*.
- Deecke, V. B. (2006). Studying marine mammal cognition in the wild : a review of four decades of playback experiments. *Aquatic mammals*, 32(4) :461–482.
- Deecke, V. B., Barrett-Lennard, L. G., Spong, P., and Ford, J. K. (2010). The structure of stereotyped calls reflects kinship and social affiliation in resident Killer whales (*Orcinus orca*). *Naturwissenschaften*, 97(5) :513–518.

Bibliographie

- Deecke, V. B., Ford, J. K., and Slater, P. J. (2005). The vocal behaviour of mammal-eating Killer whales : communicating with costly calls. *Animal Behaviour*, 69(2) :395–405.
- Deecke, V. B., Ford, J. K., and Spong, P. (1999). Quantifying complex patterns of bioacoustic variation : Use of a neural network to compare killer whale (*Orcinus orca*) dialects. *The Journal of the Acoustical Society of America*, 105(4) :2499–2507.
- Deecke, V. B., Ford, J. K., and Spong, P. (2000). Dialect change in resident killer whales : implications for vocal learning and cultural transmission. *Animal behaviour*, 60(5) :629–638.
- Deecke, V. B. and Janik, V. M. (2006). Automated categorization of bioacoustic signals : avoiding perceptual pitfalls. *The Journal of the Acoustical Society of America*, 119(1) :645–653.
- Deeke, V. B. (1998). *Stability and change of Killer whale (Orcinus orca) dialects*. PhD thesis, University of British Columbia.
- Desprat, J. L., Lengagne, T., Dumet, A., Desouhant, E., and Mondy, N. (2015). Immunocompetence handicap hypothesis in tree frog : trade-off between sexual signals and immunity? *Behavioral Ecology*, 26(4) :1138–1146.
- Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3) :297–302.
- Dooling, R. J., Lohr, B., and Dent, M. L. (2000). Hearing in birds and reptiles. In *Comparative hearing : birds and reptiles*, pages 308–359. Springer.
- Dreher, J. and Evans, W. (1964). Cetacean communication. *Marine bioacoustics*, 1 :373–393.
- Drouot, V., Gannier, A., and Goold, J. C. (2004). Summer social distribution of Sperm whales (*Physeter macrocephalus*) in the mediterranean sea. *Journal of the Marine Biological Association of the United Kingdom*, 84(3) :675–680.
- Duan, S., Zhang, J., Roe, P., Wimmer, J., Dong, X., Truskinger, A., and Towsey, M. (2013). Timed probabilistic automaton : a bridge between raven and song scope for automatic species recognition. In *Twenty-Fifth IAAI Conference*.
- Duarte de Figueiredo, L. and Simão, S. M. (2009). Possible occurrence of signature whistles in a population of sotalia guianensis (cetacea, delphinidae) living in sepetiba bay, brazil. *The Journal of the Acoustical Society of America*, 126(3) :1563–1569.
- Dufour, O., Artières, T., Glotin, H., and Giraudet, P. (2013). Clusterized mel filter cepstral coefficients and support vector machines for bird song identification. In *Soundscape Semiotics—Localization and Categorization*, number 2013 in 1, pages 89–93. InTech.
- Dugan, P., Zollweg, J., Popescu, M., Risch, D., Glotin, H., LeCun, Y., et al. (2015). High performance computer acoustic data accelerator : A new system for exploring marine mammal acoustics for big data applications. *arXiv preprint arXiv :1509.03591*.
- Dunlop, R. A., Cato, D. H., and Noad, M. J. (2008). Non-song acoustic communication in migrating humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science*, 24(3) :613–629.
- Duquet, M. and Thiollay, J.-M. (1993). *Glossaire d'écologie fondamentale*. Nathan.
- Elemans, C., Rasmussen, J. H., Herbst, C. T., Düring, D. N., Zollinger, S. A., Brumm, H., Srivastava, K., Svane, N., Ding, M., Larsen, O. N., et al. (2015). Universal mechanisms of sound production and control in birds and mammals. *Nature communications*, 6(1) :1–13.
- Elowson, A. M., Tannenbaum, P. L., and Snowdon, C. T. (1991). Food-associated calls correlate with food preferences in cotton-top tamarins. *Animal Behaviour*, 42(6) :931–937.
- Engel, M. H., Fagundes, N. J., Rosenbaum, H. C., Leslie, M. S., Ott, P. H., Schmitt, R., Secchi, E., Dalla Rosa, L., and Bonatto, S. L. (2008). Mitochondrial dna diversity of the southwestern atlantic humpback whale (*Megaptera novaeangliae*) breeding area off brazil, and the potential connections to antarctic feeding areas. *Conservation Genetics*, 9(5) :1253–1262.
- Farina, A. (2013). *Soundscape ecology : principles, patterns, methods and applications*. Springer.

- Fearey, J., Elwen, S. H., James, B., and Gridley, T. (2019). Identification of potential signature whistles from free-ranging common dolphins (*Delphinus delphis*) in south africa. *Animal cognition*, 22(5) :777–789.
- Ferrari, M., Glotin, H., Marxer, R., and Asch, M. (2020). Docc10 : Open access dataset of marine mammal transient studies and end-to-end cnn classification. In *Int. Joint Conf. on Neural Net.*
- Filatova, O., Burdin, A., and Hoyt, E. (2010). Horizontal transmission of vocal traditions in Killer whale *Orcinus orca* dialects. *Biology bulletin*, 37(9) :965–971.
- Filatova, O., Fedutin, I., Burdin, A., and Hoyt, E. (2006). Using a mobile hydrophone stereo system for real-time acoustic localization of Killer whales (*Orcinus orca*). *Applied acoustics*, 67(11-12) :1243–1248.
- Filatova, O., Guzeev, M., Fedutin, I., Burdin, A., and Hoyt, E. (2013). Dependence of Killer whale *Orcinus orca* acoustic signals on the type of activity and social context. *Biology Bulletin*, 40(9) :790–796.
- Fitzpatrick, M. C., Preisser, E. L., Ellison, A. M., and Elkinton, J. S. (2009). Observer bias and the detection of low-density populations. *Ecological applications*, 19(7) :1673–1679.
- Fleischer, G. (1973). On structure and function of the middle ear in the bottlenosed dolphin (*Tursiops truncatus*). In *Ninth Annual Conference on Biological Sonar and Diving Mammals. Menlo Park, CA : Stanford Research Institute Press*, volume 137179.
- Fletcher, N. H. and Tarnopolsky, A. (1999). Acoustics of the avian vocal tract. *The Journal of the Acoustical Society of America*, 105(1) :35–49.
- Ford, J. K. (1989). Acoustic behaviour of resident Killer whales (*Orcinus orca*) off vancouver island, british columbia. *Canadian Journal of Zoology*, 67(3) :727–745.
- Ford, J. K. (1991). Vocal traditions among resident Killer whales (*Orcinus orca*) in coastal waters of british columbia. *Canadian journal of zoology*, 69(6) :1454–1483.
- Ford, J. K. (2009). Killer whale : *Orcinus orca*. In *Encyclopedia of marine mammals*, pages 650–657. Elsevier.
- Ford, J. K., Ellis, G. M., Barrett-Lennard, L. G., Morton, A. B., Palm, R. S., and Balcomb III, K. C. (1998). Dietary specialization in two sympatric populations of Killer whales (*Orcinus orca*) in coastal british columbia and adjacent waters. *Canadian Journal of Zoology*, 76(8) :1456–1471.
- Ford, J. K. et al. (1987). *A catalogue of underwater calls produced by Killer whales (Orcinus orca) in British Columbia*. Department of Fisheries and Oceans, Fisheries Research Branch, Canada.
- Forrest, T. (1994). From sender to receiver : propagation and environmental effects on acoustic signals. *American zoologist*, 34(6) :644–654.
- Fourniol, M., Gies, V., Barchasz, V., Kussener, E., Barthelemy, H., Vauche, R., and Glotin, H. (2018). Low-power wake-up system based on frequency analysis for environmental internet of things. In *14th Int. Conference on Mechatronic and Embedded Systems and Applications (MESA)*, pages 1–6. IEEE.
- Fowlkes, E. and Mallows, C. (1983). A method for comparing two hierarchical clusterings. *Journal of the American statistical association*, 78(383) :553–569.
- Frankel, A., Clark, C., Herman, L., and Gabriele, C. (1995). Spatial distribution, habitat utilization, and social interactions of humpback whales off hawai'i, determined using acoustic and visual techniques. *Canadian Journal of Zoology*, 73(6) :1134–1146.
- Frankel, A. S. (2009). Sound production. In *Encyclopedia of marine mammals*, pages 1056–1071. Elsevier.
- Frantzis, A., Airoidi, S., Notarbartolo-di Sciara, G., Johnson, C., and Mazzariol, S. (2011). Inter-basin movements of mediterranean Sperm whales provide insight into their population structure and conservation. *Deep Sea Research Part I : Oceanographic Research Papers*, 58(4) :454–459.
- Frasier, K., Elizabeth Henderson, E., Bassett, H., and Roch, M. (2016). Automated identification and clustering of subunits within *delphinid* vocalizations. *Marine Mammal Science*.

Bibliographie

- Frasier, K. E., Roch, M. A., Soldevilla, M. S., Wiggins, S. M., Garrison, L. P., and Hildebrand, J. A. (2017). Automated classification of dolphin echolocation click types from the gulf of mexico. *PLoS computational biology*, 13(12) :e1005823.
- Gannier, A., Drouot, V., and Goold, J. C. (2002). Distribution and relative abundance of sperm whales in the mediterranean sea. *Marine Ecology Progress Series*, 243 :281–293.
- Garcia, M., Herbst, C. T., Bowling, D. L., Dunn, J. C., and Fitch, W. T. (2017). Acoustic allometry revisited : morphological determinants of fundamental frequency in primate vocal production. *Scientific reports*, 7(1) :1–11.
- Gedamke, J., Costa, D. P., and Dunstan, A. (2001). Localization and visual verification of a complex minke whale vocalization. *The Journal of the Acoustical Society of America*, 109(6) :3038–3047.
- Gerard, O., Carthel, C., Coraluppi, S., and Willett, P. (2008). Feature-aided tracking for marine mammal detection and classification. *Canadian Acoustics*, 36(1) :13–19.
- Ghahramani, Z. (2015). Probabilistic machine learning and artificial intelligence. *Nature*, 521(7553) :452–459.
- Gillespie, D., Caillat, M., Gordon, J., and White, P. (2013). Automatic detection and classification of odontocete whistles. *The Journal of the Acoustical Society of America*, 134(3) :2427–2437.
- Gillespie, D., Palmer, L., Macaulay, J., Sparling, C., and Hastie, G. (2020). Passive acoustic methods for tracking the 3d movements of small cetaceans around marine structures. *Plos one*, 15(5) :e0229058.
- Gingras, B., Mohandesan, E., Boko, D., and Fitch, W. T. (2013). Phylogenetic signal in the acoustic parameters of the advertisement calls of four clades of anurans. *BMC evolutionary biology*, 13(1) :1–12.
- Giraudet, P. and Glotin, H. (2006). Real-time 3d tracking of whales by echo-robust precise tdoa estimates with a widely-spaced hydrophone array. *Applied Acoustics*, 67(11-12) :1106–1117.
- Glotin, H., Caudal, F., and Giraudet, P. (2008). Whale cocktail party : real-time multiple tracking and signal analyses. *Canadian acoustics*, 36(1) :139–145.
- Glotin, H., Clark, C., LeCun, Y., Dugan, P., Halkias, X., and Sueur, J. (2013a). *The Bioacoustic Challenges in ICML4B*. ICML, SABIOD.
- Glotin, H., Clark, C., LeCun, Y., Dugan, P., Halkias, X., and Sueur, J. (2013b). *Proc. 1st workshop on Machine Learning for Bioacoustics - ICML4B*. ICML int. Conf., USA. http://sabiiod.univ-tln.fr/ICML4B2013_book.pdf.
- Glotin, H., Giraudet, P., Ricard, J., Malige, F., Patris, P., Roger, V., Prévot, J.-M., Poupard, M., Philippe, O., and Cosentino, P. (2017). Projet VAMOS : Visées aériennes de mammifères marins jointes aux observations acoustiques sous-marines de la bouée BOMBYX et antares : nouveaux modèles en suivis et lois allométriques du *Physeter macrocephalus*, *Ziphius Cavirostris* et autres cétacés. Technical report, Univ. Toulon. http://sabiiod.org/pub/PELAGOS_VAMOS.pdf.
- Glotin, H., LeCun, Y., Artieres, T., Mallat, S., Tchernichovski, O., and Halkias, X. (2013c). *Neural information processing scaled for bioacoustics, from neurons to big data*. NeurIPS (NIPS) Workshop. http://sabiiod.org/NIPS4B2013_book.pdf.
- Glotin, H., Thellier, N., Best, P., Poupard, M., Ferrari, M., Viera, S., Giés, V., Oger, M., Giraudet, P., Mercier, M., Donzé, G., Campana, M., Chevallier, J., Malige, F., Patris, J., Prévot, J., Cosentino, P., Prévot d’Alvise, N., Ourmières, Y., Barchasz, V., Lehir, A., Marzetti, S., Sarano, F., Benveniste, J., Gaillard, S., and de Varenne, F. (2020). *Sphyrna-Odyssey 2019-2020, Rapport I : Découvertes Ethoacoustiques de Chasses Collaboratives de Cachalots en Abysses et Impacts en Mer du Confinement COVID19*. 197p, CNRS LIS, Université de Toulon. <http://sabiiod.org/pub/S01.pdf>.
- Glotin, H. G. P., Giraudet, P. V., and Caudal, F. (2014). Real-time robust method for determining the trajectory of one or more cetaceans by means of passive acoustics, using a laptop computer. US Patent 8,638,641.
- Glotin, H. G. P., Mischenko, A., and Giraudet, P. V. (2015). Joint constraints imposed on multiband time transitivity and doppler-effect differences, for separating, characterizing, and locating sound sources via passive acoustics. Patent WO2015177172A1 <https://patentimages.storage.googleapis.com/50/2d/de/a0db00f3ca53fb/WO2015177172A1.pdf>.

- Goëau, H., Glotin, H., Vellinga, W.-P., Planqué, R., and Joly, A. (2016). Lifeclef bird identification task 2016 : The arrival of deep learning. In *CLEF : Conference and Labs of the Evaluation Forum*.
- Gomez, D., Théry, M., Gauthier, A.-L., and Lengagne, T. (2011). Costly help of audiovisual bimodality for female mate choice in a nocturnal anuran (hyla arborea). *Behavioral Ecology*, 22(4) :889–898.
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). Machine learning basics. In *Deep learning*, volume 1, pages 98–164. MIT press.
- Gordon, J. C. (1991). Evaluation of a method for determining the length of Sperm whales (*Physeter catodon*) from their vocalizations. *Journal of Zoology*, 224(2) :301–314.
- Grebner, D. M. (2009). Individual Killer whale vocal variation during intra-group behavioral dynamics. *Phd Thesis, Pennsylvania State Univ.*
- Grebner, D. M., Bradley, D. L., and Capone, D. E. (2007). Localization of Killer whale (*Orcinus orca*) vocalizations using a triangular hydrophone array in johnstone strait, british columbia, canada. *The Journal of the Acoustical Society of America*, 121(5) :3105–3105.
- Green, S. and Marler, P. (1979). The analysis of animal communication. In *Social behavior and communication*, pages 73–158. Springer.
- Gridley, T., Elwen, S. H., Rashley, G., Badenas Krakauer, A., and Heiler, J. (2016). Bottlenose dolphins change their whistling characteristics in relation to vessel presence, surface behavior and group composition. In *Proceedings of Meetings on Acoustics*. Acoustical Society of America.
- Grigg, G., Taylor, A., Mc Callum, H., and Watson, G. (1996). Monitoring frog communities : an application of machine learning. In *Proceedings of eighth innovative applications of artificial intelligence conference, Portland Oregon*, pages 1564–1569.
- Grill, T. and Schlüter, J. (2017). Two convolutional neural networks for bird detection in audio signals. In *2017 25th European Signal Processing Conference (EUSIPCO)*, pages 1764–1768. IEEE.
- Growcott, A., Miller, B., Sirguey, P., Slooten, E., and Dawson, S. (2011). Measuring body length of male Sperm whales from their clicks : the relationship between inter-pulse intervals and photogrammetrically measured lengths. *The Journal of the Acoustical Society of America*, 130(1) :568–573.
- Guihou, K., Marmain, J., Ourmieres, Y., Molcard, A., Zakardjian, B., and Forget, P. (2013). A case study of the mesoscale dynamics in the north-western mediterranean sea : a combined data–model approach. *Ocean Dynamics*, 63(7) :793–808.
- Halkias, X. C. and Ellis, D. P. (2006). Call detection and extraction using bayesian inference. *Applied Acoustics*, 67(11-12) :1164–1174.
- Halkias, X. C., Paris, S., and Glotin, H. (2013). Classification of mysticete sounds using machine learning techniques. *The Journal of the Acoustical Society of America*, 134(5) :3496–3505.
- Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of intelligent information systems*, 17(2-3) :107–145.
- Haren, A. M. (2007). Reducing noise pollution from commercial shipping in the channel islands national marine sanctuary : a case study in marine protected area management of underwater noise. *Journal of International Wildlife Law and Policy*, 10(2) :153–173.
- Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The elements of statistical learning : data mining, inference, and prediction*. Springer Science & Business Media.
- Hawkins, E. R. and Gartside, D. F. (2010). Whistle emissions of indo-pacific bottlenose dolphins (*Tursiops aduncus*) differ with group composition and surface behaviors. *The Journal of the Acoustical Society of America*, 127(4) :2652–2663.
- Hedwig, D., Hammerschmidt, K., Mundry, R., Robbins, M. M., and Boesch, C. (2014). Acoustic structure and variation in mountain and western gorilla close calls : a syntactic approach. *Behaviour*, 151(8) :1091–1120.

Bibliographie

- Heenehan, H., Stanistreet, J. E., Corkeron, P. J., Bouveret, L., Chalifour, J., Davis, G. E., Henriquez, A., Kiszka, J., Kline, L., Reed, C., et al. (2019). Caribbean sea soundscapes : monitoring humpback whales, biological sounds, geological events, and anthropogenic impacts of vessel noise. *Frontiers in Marine Science*, 6 :347.
- Herzing, D. (2000). Acoustics and social behavior of wild dolphins : implications for a sound society. In *Hearing by whales and dolphins*, pages 225–272. Springer.
- Hildebrand, J. A. (2009). Anthropogenic and natural sources of ambient noise in the ocean. *Marine Ecology Progress Series*, 395 :5–20.
- Hill, R. and Hüppop, O. (2008). Birds and bats : automatic recording of flight calls and their value for the study of migration. *Computational bioacoustics for assessing biodiversity. BfN-Skripten*, 234 :135–141.
- Hinton, G. and Roweis, S. (2002). Stochastic neighbor embedding. In *NIPS*, volume 15, pages 833–840.
- Hobson, K. A., Rempel, R. S., Greenwood, H., Turnbull, B., and Van Wilgenburg, S. L. (2002). Acoustic surveys of birds using electronic recordings : new potential from an omnidirectional microphone system. *Wildlife Society Bulletin*, pages 709–720.
- Hoelzel, A. R. and Dover, G. A. (1991). Genetic differentiation between sympatric Killer whale populations. *Heredity*, 66(2) :191–195.
- Holt, M. M., Noren, D. P., Veirs, V., Emmons, C. K., and Veirs, S. (2009). Speaking up : Killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. *The Journal of the Acoustical Society of America*, 125(1) :EL27–EL32.
- Horton, T. W., Holdaway, R. N., Zerbini, A. N., Hauser, N., Garrigue, C., Andriolo, A., and Clapham, P. J. (2011). Straight as an arrow : humpback whales swim constant course tracks during long-distance migration. *Biology letters*, 7(5) :674–679.
- Houtsma, A. J. (1997). Pitch and timbre : Definition, meaning and use. *Journal of New Music Research*, 26(2) :104–115.
- Irwin, D., Thimgan, M., and Irwin, J. (2008). Call divergence is correlated with geographic and genetic distance in greenish warblers (*phylloscopus trochiloides*) : a strong role for stochasticity in signal evolution? *Journal of evolutionary biology*, 21(2) :435–448.
- Jackson, J., Ross-Gillespie, A., Butterworth, D., Findlay, K., Holloway, S., Robbins, J., and Zerbini, A. (2015). Southern hemisphere humpback whale comprehensive assessment—a synthesis and summary : 2005–2015. *IWC SC/66a/SH3*, 38 :1–38.
- Janik, V., Van Parijs, S., and Thompson, P. (2000). A two-dimensional acoustic localization system for marine mammals. *Marine Mammal Science*, 16(2) :437–447.
- Janik, V. M., Sayigh, L. S., and Wells, R. S. (2006). Signature whistle shape conveys identity information to bottlenose dolphins. *Proceedings of the National Academy of Sciences*, 103(21) :8293–8297.
- Janik, V. M. and Slater, P. J. (1998). Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls. *Animal behaviour*, 56(4) :829–838.
- Janik, V. M., Todt, D., and Dehnhardt, G. (1994). Signature whistle variations in a bottlenosed dolphin, *Tursiops truncatus*. *Behavioral Ecology and Sociobiology*, 35(4) :243–248.
- Jaquet, N., Dawson, S., and Douglas, L. (2001). Vocal behavior of male Sperm whales : Why do they click? *The Journal of the Acoustical Society of America*, 109(5) :2254–2259.
- Jaquet, N., Dawson, S., and Slooten, E. (2000). Seasonal distribution and diving behaviour of male Sperm whales off kaikoura : foraging implications. *Canadian Journal of Zoology*, 78(3) :407–419.
- Jaramillo-Legorreta, A., Cardenas-Hinojosa, G., Nieto-Garcia, E., Rojas-Bracho, L., Ver Hoef, J., Moore, J., Tregenza, N., Barlow, J., Gerrodette, T., Thomas, L., et al. (2017). Passive acoustic monitoring of the decline of mexico's critically endangered vaquita. *Conservation Biology*, 31(1) :183–191.
- Jarvis, S., DiMarzio, N., Morrissey, R., and Moretti, D. (2008). A novel multi-class support vector machine classifier for automated classification of beaked whales and other small odontocetes. *Canadian Acoustics*, 36(1) :34–40.

- Jeliazkov, A., Bas, Y., Kerbirou, C., Julien, J.-F., Penone, C., and Le Viol, I. (2016). Large-scale semi-automated acoustic monitoring allows to detect temporal decline of bush-crickets. *Global Ecology and Conservation*, 6 :208–218.
- Jenkins, P. F. (1978). Cultural transmission of song patterns and dialect development in a free-living bird population. *Animal Behaviour*, 26 :50–78.
- Joachim, J. and Lauga, J. (1996). Populations et dialectes chez le pinson. *Comptes rendus de l'Académie des sciences. Série 3, Sciences de la vie*, 319(3) :231–239.
- Johnson, C. S. (1967). Sound detection thresholds in marine mammals. *Marine bioacoustics*, pages 247–260.
- Johnson, M., de Soto, N. A., and Madsen, P. T. (2009). Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags : a review. *Marine Ecology Progress Series*, 395 :55–73.
- Johnstone, R. A. (1997). Recognition and the evolution of distinctive signatures : when does it pay to reveal identity ? *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 264(1387) :1547–1553.
- Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.-P., Lombardo, J.-C., Planque, R., Palazzo, S., and Müller, H. (2017). Lifeclef 2017 lab overview : multimedia species identification challenges. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 255–274. Springer.
- Joly, A., Goëau, H., Glotin, H., Spampinato, C., Bonnet, P., Vellinga, W.-P., Planqué, R., Rauber, A., Palazzo, S., Fisher, B., et al. (2015). Lifeclef 2015 : multimedia life species identification challenges. In *International Conference of the Cross-Language Evaluation Forum for European Languages*, pages 462–483. Springer.
- Jouventin, P. and Aubin, T. (2002). Acoustic systems are adapted to breeding ecologies : individual recognition in nesting penguins. *Animal Behaviour*, 64(5) :747–757.
- Jouventin, P., Aubin, T., and Lengagne, T. (1999). Finding a parent in a king penguin colony : the acoustic system of individual recognition. *Animal Behaviour*, 57(6) :1175–1183.
- Kahl, S., Clapp, M., Hopping, A., Goëau, H., Glotin, H., Planqué, R., Vellinga, W., and Joly, A. (2020). Overview of birdclef 2020 : bird sound recognition in complex acoustic environments. *CLEF Task Overview*.
- Kahl, S., Wilhelm-Stein, T., Hussein, H., Klinck, H., Kowerko, D., Ritter, M., and Eibl, M. (2017). Large-scale bird sound classification using convolutional neural networks. In *CLEF (Working Notes)*.
- Kajiura, S. M. and Holland, K. N. (2002). Electroreception in juvenile scalloped hammerhead and sandbar sharks. *Journal of Experimental Biology*, 205(23) :3609–3621.
- Kandia, V. and Stylianou, Y. (2006). Detection of Sperm whale clicks based on the teager–kaiser energy operator. *Applied Acoustics*, 67(11-12) :1144–1163.
- Kaplan, M. B., Aran Mooney, T., Sayigh, L. S., and Baird, R. W. (2014). Repeated call types in hawaiian melon-headed whales (*Peponocephala electra*). *The Journal of the Acoustical Society of America*, 136(3) :1394–1401.
- Karenina, K., Giljov, A., Ivkovich, T., Burdin, A., and Malashichev, Y. (2013). Lateralization of spatial relationships between wild mother and infant orcas, *Orcinus orca*. *Animal behaviour*, 86(6) :1225–1231.
- Kays, R., McShea, W. J., and Wikelski, M. (2020). Born-digital biodiversity data : Millions and billions. *Diversity and Distributions*, 26(5) :644–648.
- Kimura, S., Akamatsu, T., Li, S., Dong, S., Dong, L., Wang, K., Wang, D., and Arai, N. (2010). Density estimation of yangtze finless porpoises using passive acoustic sensors and automated click train detection. *The Journal of the Acoustical Society of America*, 128(3) :1435–1445.
- Kishida, T., Thewissen, J., Hayakawa, T., Imai, H., and Agata, K. (2015). Aquatic adaptation and the evolution of smell and taste in whales. *Zoological letters*, 1(1) :9.
- Klingbeil, B. T. and Willig, M. R. (2015). Bird biodiversity assessments in temperate forest : the value of point count versus acoustic monitoring protocols. *PeerJ*, 3 :e973.

Bibliographie

- Klump, G. and Shalter, M. (1984). Acoustic behaviour of birds and mammals in the predator context ; i. factors affecting the structure of alarm signals. ii. the functional significance and evolution of alarm signals. *Zeitschrift für Tierpsychologie*, 66(3) :189–226.
- Knight, E. C., Poo Hernandez, S., Bayne, E. M., Bulitko, V., and Tucker, B. V. (2020). Pre-processing spectrogram parameters improve the accuracy of bioacoustic classification using convolutional neural networks. *Bioacoustics*, 29(3) :337–355.
- Koga, S. (2015). Major challenges and solutions for utilizing big data in the maritime industry. *World Maritime University Dissertations*.
- Krause, B. L. (1993). The niche hypothesis : a virtual symphony of animal sounds, the origins of musical expression and the health of habitats. *The Soundscape Newsletter*, 6 :6–10.
- Krebs, J. R. and Kroodsma, D. E. (1980). Repertoires and geographical variation in bird song. In *Advances in the Study of Behavior*, volume 11, pages 143–177. Elsevier.
- Kroodsma, D. E. (1974). Song learning, dialects, and dispersal in the bewick's wren. *Zeitschrift für Tierpsychologie*, 35(4) :352–380.
- Kumar, A. (2003). Acoustic communication in birds. *Resonance*, 8(6) :44–55.
- La Manna, G., Manghi, M., Perretti, F., and Sara, G. (2016). Behavioral response of brown meagre (*Sciaena umbra*) to boat noise. *Marine pollution bulletin*, 110(1) :324–334.
- Laiolo, P. (2010). The emerging significance of bioacoustics in animal species conservation. *Biological Conservation*, 143(7) :1635–1645.
- Laist, D. W., Knowlton, A. R., Mead, J. G., Collet, A. S., and Podesta, M. (2001). Collisions between ships and whales. *Marine Mammal Science*, 17(1) :35–75.
- Lammers, M. O., Au, W. W., and Herzog, D. L. (2003). The broadband social acoustic signaling behavior of spinner and spotted dolphins. *The Journal of the Acoustical Society of America*, 114(3) :1629–1639.
- Lammers, M. O., Brainard, R. E., Au, W. W., Mooney, T. A., and Wong, K. B. (2008). An ecological acoustic recorder (ear) for long-term monitoring of biological and anthropogenic sounds on coral reefs and other marine habitats. *The Journal of the Acoustical Society of America*, 123(3) :1720–1728.
- Lasseck, M. (2013). Bird song classification in field recordings : winning solution for nips4b 2013 competition. In *Proc. of int. symp. Neural Information Scaled for Bioacoustics, sabiod. org/nips4b, joint to NIPS, Nevada*, pages 176–181.
- Leavens, D. A., Russell, J. L., and Hopkins, W. D. (2010). Multimodal communication by captive chimpanzees (pan troglodytes). *Animal cognition*, 13(1) :33–40.
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *nature*, 521(7553) :436–444.
- Lengagne, T., Aubin, T., Lauga, J., and Jouventin, P. (1999). How do king penguins (*aptenodytes patagonicus*) apply the mathematical theory of information to communicate in windy conditions ? *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 266(1429) :1623–1628.
- Lengagne, T., Ferrandiz-Rovira, M., Superbie, C., Figueroa, I., Bichet, C., Claramunt-Lopez, B., and Cohas, A. (2020). Geographic variation in marmots' alarm calls causes different responses. *Behavioral Ecology and Sociobiology*, 74(8) :1–13.
- Lengagne, T., Gomez, D., Josserand, R., and Voituren, Y. (2015). Long recording sequences : how to track the intra-individual variability of acoustic signals. *PloS one*, 10(5) :e0123828.
- Lengagne, T., Lauga, J., and Aubin, T. (2001). Intra-syllabic acoustic signatures used by the king penguin in parent-chick recognition : an experimental approach. *Journal of Experimental Biology*, 204(4) :663–672.
- Lengagne, T. and Slater, P. J. (2002). The effects of rain on acoustic communication : tawny owls have good reason for calling less in wet weather. *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 269(1505) :2121–2125.

- Lesage, V., Barrette, C., Kingsley, M., and Sjare, B. (1999). The effect of vessel noise on the vocal behavior of belugas in the St. Lawrence River estuary, Canada. *Marine Mammal Science*, 15(1) :65–84.
- Levréro, F., Durand, L., Vignal, C., Blanc, A., and Mathevon, N. (2009). Begging calls support offspring individual identity and recognition by zebra finch parents. *Comptes rendus biologies*, 332(6) :579–589.
- Liang, Q., Zhang, B., Zhao, C., and Pi, Y. (2012). Tdoa for passive localization : Underwater versus terrestrial environment. *Transactions on Parallel and Distributed Systems*, 24(10) :2100–2108.
- Lin, T.-H., Akamatsu, T., and Chou, L.-S. (2013). Tidal influences on the habitat use of Indo-Pacific humpback dolphins in an estuary. *Marine Biology*, 160(6) :1353–1363.
- Liu, S., Liu, M., Wang, M., Ma, T., and Qing, X. (2018). Classification of cetacean whistles based on convolutional neural network. In *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5. IEEE.
- Logan, B. et al. (2000). Mel frequency cepstral coefficients for music modeling. In *ISMIR*, volume 270, pages 1–11.
- Lopez, B. and Shirai, J. (2009). Mediterranean common bottlenose dolphins repertoire and communication use. *Dolphins : Anatomy, behavior, and threats*, pages 129–148.
- LPO (1912). Ligue pour la protection des oiseaux. <https://www.lpo.fr/>. Last checked on Aug 01, 2019.
- Luo, W., Yang, W., and Zhang, Y. (2019). Convolutional neural network for detecting odontocete echolocation clicks. *The Journal of the Acoustical Society of America*, 145(1) :EL7–EL12.
- Lusseau, D., Bain, D. E., Williams, R., and Smith, J. C. (2009). Vessel traffic disrupts the foraging behavior of southern resident killer whales *Orcinus orca*. *Endangered Species Research*, 6(3) :211–221.
- Lutz, M. and Biernat, E. (2015). *Data Science : fondamentaux et études de cas : Machine Learning avec Python et R*. Editions Eyrolles.
- Mac Aodha, O., Gibb, R., Barlow, K. E., Browning, E., Firman, M., Freeman, R., Harder, B., Kinsey, L., Mead, G. R., Newson, S. E., et al. (2018). Bat detective—deep learning tools for bat acoustic signal detection. *PLoS computational biology*, 14(3) :e1005995.
- MacQueen, J. et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA.
- Madsen, P., Kerr, I., and Payne, R. (2004). Echolocation clicks of two free-ranging, oceanic delphinids with different food preferences : false killer whales *Pseudorca crassidens* and Risso's dolphins *Grampus griseus*. *Journal of Experimental Biology*, 207(11) :1811–1823.
- Madsen, P., Payne, R., Kristiansen, N., Wahlberg, M., Kerr, I., and Møhl, B. (2002). Sperm whale sound production studied with ultrasound time/depth-recording tags. *Journal of Experimental Biology*, 205(13) :1899–1906.
- Magrath, R. D., Haff, T. M., Horn, A. G., and Leonard, M. L. (2010). Calling in the face of danger : predation risk and acoustic communication by parent birds and their offspring. In *Advances in the Study of Behavior*, volume 41, pages 187–253. Elsevier.
- Malfante, M., Mars, J. I., Dalla Mura, M., and Gervaise, C. (2018). Automatic fish sounds classification. *The Journal of the Acoustical Society of America*, 143(5) :2834–2846.
- Manabe, K., Sadr, E. I., and Dooling, R. J. (1998). Control of vocal intensity in budgerigars (*Melopsittacus undulatus*) : differential reinforcement of vocal intensity and the Lombard effect. *The Journal of the Acoustical Society of America*, 103(2) :1190–1198.
- Marler, P. (1987). Sensitive periods and the roles of specific and general sensory stimulation in birdsong learning. *Imprinting and cortical plasticity*, pages 99–135.
- Marler, P. and Peters, S. (1982). Long-term storage of learned birdsongs prior to production. *Animal Behaviour*, 30(2) :479–482.

Bibliographie

- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2) :431–441.
- Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). Estimating animal population density using passive acoustics. *Biological Reviews*, 88(2) :287–309.
- Marshall, C. D., Amin, H., Kovacs, K. M., and Lydersen, C. (2006). Microstructure and innervation of the mystacial vibrissal follicle-sinus complex in bearded seals, *erignathus barbatus* (pinnipedia : Phocidae). *The Anatomical Record Part A : Discoveries in Molecular, Cellular, and Evolutionary Biology : An Official Publication of the American Association of Anatomists*, 288(1) :13–25.
- Marten, K. and Marler, P. (1977). Sound transmission and its significance for animal vocalization. *Behavioral ecology and sociobiology*, 2(3) :271–290.
- Mathevon, N., Charrier, I., and Jouventin, P. (2003). Potential for individual recognition in acoustic signals : a comparative study of two gulls with different nesting patterns. *Comptes Rendus Biologies*, 326(3) :329–337.
- Mayol, P., de Mongolfier, B., Ratel, M., Bordes, R., Costales, L., Latropoulos, D., Ortolé, C., and Belhadjer, A. (2015). Caractérisation des activités d'observation commerciale des cétacés à l'échelle du sanctuaire agoa. Technical report, Aquasearch.
- McCowan, B. and Reiss, D. (2001). The fallacy of 'signature whistles' in bottlenose dolphins : a comparative perspective of 'signature information' in animal vocalizations. *Animal behaviour*, 62(6) :1151–1162.
- McCracken, K. G. and Sheldon, F. H. (1997). Avian vocalizations and phylogenetic signal. *Proceedings of the National Academy of Sciences*, 94(8) :3833–3836.
- McGowen, M. R., Clark, C., and Gatesy, J. (2008). The vestigial olfactory receptor subgenome of odontocete whales : phylogenetic congruence between gene-tree reconciliation and supermatrix methods. *Systematic biology*, 57(4) :574–590.
- McGregor, P. (2000). Conservation applications of behaviour. *Behaviour and Conservation* (ed. Gosling, L., Morris, L. & Sutherland, WJ).—Cambridge University Press, Cambridge.
- McGregor, P. K. and Krebs, J. R. (1982). Song types in a population of great tits (*Parus major*) : their distribution, abundance and acquisition by individuals. *Behaviour*, 79(2-4) :126–152.
- McIlraith, A. L. and Card, H. (1995). Birdsong recognition with dsp and neural networks. In *IEEE WESCANEX 95. Communications, Power, and Computing. Conference Proceedings*, volume 2, pages 409–414. IEEE.
- McInnes, L., Healy, J., and Melville, J. (2018). Umap : Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv :1802.03426*.
- Mellinger, D. K., Stafford, K. M., Moore, S. E., Dziak, R. P., and Matsumoto, H. (2007). An overview of fixed passive acoustic observation methods for cetaceans. *Oceanography*, 20(4) :36–45.
- Mennill, D. J., Burt, J. M., Fristrup, K. M., and Vehrencamp, S. L. (2006). Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *The Journal of the Acoustical Society of America*, 119(5) :2832–2839.
- Mercado, E. (2018). The sonar model for humpback whale song revised. *Frontiers in psychology*, 9 :1156.
- Miller, P. J., Shapiro, A., Tyack, P., and Solow, A. (2004). Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Animal Behaviour*, 67(6) :1099–1107.
- Miller, P. J. and Tyack, P. L. (1998). A small towed beamforming array to identify vocalizing resident Killer whales *Orcinus orca* concurrent with focal behavioral observations. *Deep Sea Research Part II : Topical Studies in Oceanography*, 45(7) :1389–1405.
- Mitani, J. C., Hasegawa, T., Gros-Louis, J., Marler, P., and Byrne, R. (1992). Dialects in wild chimpanzees ? *American Journal of Primatology*, 27(4) :233–243.
- Mobley Jr, J. R. and Herman, L. M. (1985). Transience of social affiliations among humpback whales *Megaptera novaeangliae* on the hawaiian wintering grounds. *Canadian Journal of Zoology*, 63(4) :762–772.

- Moore, S. E., Francine, J. K., Bowles, A. E., and Ford, J. K. (1988). Analysis of calls of Killer whales, *Orcinus orca*, from iceland and norway. *Rit Fiskideildar*, 11 :225–250.
- Morfi, V., Bas, Y., Pamula, H., Glotin, H., and Stowell, D. (2019). Nips4bplus : a richly annotated birdsong audio dataset. *PeerJ Computer Science*, 5 :e223.
- Morrissey, R., Ward, J., DiMarzio, N., Jarvis, S., and Moretti, D. (2006). Passive acoustic detection and localization of Sperm whales (*Physeter macrocephalus*) in the tongue of the ocean. *Applied acoustics*, 67(11-12) :1091–1105.
- Mortega, K. G., Flinks, H., and Helm, B. (2014). Behavioural response of a migratory songbird to geographic variation in song and morphology. *Frontiers in zoology*, 11(1) :85.
- Mossbridge, J. A. and Thomas, J. A. (1999). An “acoustic niche” for antarctic Killer whale and leopard seal sounds 1. *Marine Mammal Science*, 15(4) :1351–1357.
- Mullarney, K., Grant, P. J., and Christie, D. S. (2009). *Birds of Europe*. Princeton University Press.
- Mundinger, P. (1975). Song dialects and colonization in the house finch, *Carpodacus mexicanus*, on the east coast. *The Condor*, 77(4) :407–422.
- Murray, S. O., Mercado, E., and Roitblat, H. L. (1998). The neural network classification of false Killer whale (*Pseudorca crassidens*) vocalizations. *The Journal of the Acoustical Society of America*, 104(6) :3626–3633.
- Naguib, M. and Wiley, R. H. (2001). Estimating the distance to a source of sound : mechanisms and adaptations for long-range communication. *Animal behaviour*, 62(5) :825–837.
- Ness, S. R., Symonds, H., Spong, P., and Tzanetakis, G. (2013). The orchive : Data mining a massive bioacoustic archive. In *Proc. of wkp Bioacoutics, in Int. Conf on Machine Learning*. Glotin, LeCun et al., Univ. Toulon.
- Nichol, L. M. and Shackleton, D. M. (1996). Seasonal movements and foraging behaviour of northern resident Killer whales (*Orcinus orca*) in relation to the inshore distribution of salmon (*oncorhynchus* spp.) in british columbia. *Canadian Journal of Zoology*, 74(6) :983–991.
- Nicholls, J. A. (2008). Site specificity in advertisement calls and responses to playbacks of local and foreign call variants in satin bowerbirds, *Ptilonorhynchus violaceus*. *Behavioral Ecology and Sociobiology*, 62(5) :831–841.
- Nicholls, J. A. and Goldizen, A. W. (2006). Habitat type and density influence vocal signal design in satin bowerbirds. *Journal of Animal Ecology*, 75(2) :549–558.
- NOAA, C. (1998). Impact, underwater acoustic, NOAA technical memorandum. Technical report, NOAA-TM-N M FS-SW FSC-256.
- NOAA Fisheries (2011a). Nationail oceanic and atmospheric administration. <https://www.fisheries.noaa.gov/species/sperm-whale>. Last checked on Nov 01, 2019.
- NOAA Fisheries (2011b). Nationail oceanic and atmospheric administration. <https://www.fisheries.noaa.gov/species/killer-whale>. Last checked on Nov 01, 2019.
- Nottebohm, F. (1972). The origins of vocal learning. *The American Naturalist*, 106(947) :116–140.
- Nottebohm, F., Stokes, T. M., and Leonard, C. M. (1976). Central control of song in the canary, *serinus canarius*. *Journal of Comparative Neurology*, 165(4) :457–486.
- Nowacek, S., Wells, S., and Solow, A. (2001). Short-term effects of boat traffic on bottlenose dolphins in sarasota bay florida. *Marine Mammal Science*, 17(4) :673–688.
- Nummela, S. (2009). Hearing. In *Encyclopedia of marine mammals*, pages 553–562. Elsevier.
- Oswald, J. N., Barlow, J., and Norris, T. F. (2003). Acoustic identification of nine delphinid species in the eastern tropical pacific ocean. *Marine mammal science*, 19(1) :20–037.
- Oswald, J. N., Rankin, S., and Barlow, J. (2004). The effect of recording and analysis bandwidth on acoustic identification of delphinid species. *The Journal of the Acoustical Society of America*, 116(5) :3178–3185.

Bibliographie

- Owens, I. P. and Bennett, P. M. (2000). Ecological basis of extinction risk in birds : habitat loss versus human persecution and introduced predators. *Proceedings of the National Academy of Sciences*, 97(22) :12144–12148.
- O’Connell-Rodwell, C., Arnason, B., and Hart, L. A. (2000). Seismic properties of asian elephant (*Elephas maximus*) vocalizations and locomotion. *The Journal of the Acoustical Society of America*, 108(6) :3066–3072.
- Pan, S. J. and Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10) :1345–1359.
- Parker III, T. A. (1991). On the use of tape recorders in avifaunal surveys. *The Auk*, 108(2) :443–444.
- Partan, S. R. and Marler, P. (2005). Issues in the classification of multimodal communication signals. *The American Naturalist*, 166(2) :231–245.
- Pavan, G., Hayward, T. v., Borsani, J., Priano, M., Manghi, M., Fossati, C., and Gordon, J. (2000). Time patterns of Sperm whale codas recorded in the mediterranean sea 1985–1996. *The Journal of the Acoustical Society of America*, 107(6) :3487–3495.
- Payne, K. and Payne, K. (1998). *Silent thunder : in the presence of elephants*. Simon and Schuster.
- Payne, K. and Payne, R. (1985). Large scale changes over 19 years in songs of humpback whales in bermuda. *Zeitschrift für Tierpsychologie*, 68(2) :89–114.
- Payne, R. and Webb, D. (1971). Orientation by means of long range acoustic signaling in baleen whales. *Annals of the New York Academy of Sciences*, 188(1) :110–141.
- Payne, R. B. (1978). *Population structure and social behavior : models for testing the ecological significance of song dialects in birds*. Museum of Zoology, University of Michigan.
- Payne, R. S. and McVay, S. (1971). Songs of humpback whales. *Science*, 173(3997) :585–597.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn : Machine learning in Python. *Journal of Machine Learning Research*, 12 :2825–2830.
- Perrin, W. F. (2009). Pantropical spotted dolphin : *Stenella attenuata*. In *Encyclopedia of marine mammals*, pages 819–821. Elsevier.
- Perrin, W. F., Würsig, B., and Thewissen, J. (2009). *Encyclopedia of marine mammals*. Academic Press.
- Perry, E. A. and Terhune, J. M. (1999). Variation of harp seal (*pagophilus groenlandicus*) underwater vocalizations among three breeding locations. *Journal of Zoology*, 249(2) :181–186.
- Petrenko, A. A. (2003). Variability of circulation features in the gulf of lion nw mediterranean sea. importance of inertial currents. *Oceanologica acta*, 26(4) :323–338.
- Pettorelli, N., Baillie, J. E., and Durant, S. M. (2013). Indicator bats program : a system for the global acoustic monitoring of bats. *Biodiversity monitoring and conservation*, page 213.
- Pierantonio, N., Pavan, G., Airoidi, S., and Panigada, S. (2008). Presence of socializing sperm whales in the ligurian sea assessed through acoustic, behavioral and photo-identification data. In *22nd Annual Conference of the European Cetacean Society. Egmond aan Zee, The Netherlands*, pages 10–12.
- Pijanowski, B. C., Villanueva-Rivera, L. J., Dumyahn, S. L., Farina, A., Krause, B. L., Napoletano, B. M., Gage, S. H., and Pieretti, N. (2011). Soundscape ecology : the science of sound in the landscape. *BioScience*, 61(3) :203–216.
- Podos, J. and Warren, P. S. (2007). The evolution of geographic variation in birdsong. *Advances in the Study of Behavior*, 37 :403–458.
- Potamitis, I. (2014). Automatic classification of a taxon-rich community recorded in the wild. *PLoS one*, 9(5).
- Poupard, M., Best, P., Schlüter, J., Prévot, J.-M., Symonds, H., Spong, P., and Glotin, H. (2019a). Deep learning for ethoacoustics of orcas on three years pentaphonic continuous recording at Orcalab revealing tide, moon and diel effects. In *OCEANS Conf.*, pages 1–7. IEEE.

- Poupard, M., Best, P., Schlüter, J., Symonds, H., Spong, P., and Glotin, H. (2019b). Large-scale unsupervised clustering of orca vocalizations : a model for describing orca communication systems. Technical report, PeerJ Preprints.
- Poupard, M., de Montgolfier, B., and Glotin, H. (2019c). Ethoacoustic by bayesian non parametric and stochastic neighbor embedding to forecast anthropic pressure on dolphins. In *OCEANS Conf.*, pages 1–5. IEEE.
- Poupard, M., De Montgolfier, B., Roger, V., Lohani, D., and Glotin, H. (2018a). Ethoacoustics : a model based on t-SNE & clustering, applied on pantropical spotted dolphin during whale watching. In *8th Int. DCLDE (Detection, Classification, Localization, and Density Estimation) Workshop*. <https://hal.archives-ouvertes.fr/hal-01873336/document>.
- Poupard, M., Ferrari, M., Schluter, J., Marxer, R., Giraudet, P., Barchasz, V., Gies, V., Pavan, G., and Glotin, H. (2019d). Real-time passive acoustic 3d tracking of deep diving cetacean by small non-uniform mobile surface antenna. In *Int. Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8251–8255. IEEE.
- Poupard, M., Glotin, H., Bedu, A.-L., Dubourg, A. B., and Lengagne, T. (2018b). Rapport d’avancement sur le modèle de reconnaissance précise sur une cinquantaine d’espèces avec estimation des conditions de terrain. Technical report, Université de Toulon, Equipe DYNI.
- Poupard, M., Glotin, H., Lengagne, T., Bougrain-Dubourg, A., and Bedu, A.-L. (2018c). Rapport de recherche : Précision sur système, programme des expériences de terrain. Technical report, Université de Toulon.
- Poupard, M., Helena, S., Paul, S., and Hervé, G. (2020). Individual separation and identification of orcas calls in the wild : study of inter and intra-group communication. *Scientific report (in submission)*.
- Poupard, M., Lengagne, T., and Glotin, H. (2018d). Première étude comparée des applications sur mobile (iphone et android) d’identification acoustique des espèces cibles d’oiseaux de france métropolitaine. Technical report, Université de Toulon, CNRS LIS, Equipe DYNI.
- Poupard, M., Lengagne, T., and Glotin, H. (2018e). Préparation du challenge birdclef 2019 “bird4d”. Technical report, DYNI LIS CNRS Toulon.
- Pourhomayoun, M., Dugan, P., Popescu, M., Risch, D., Lewis, H., and Clark, C. (2013). Classification for big dataset of bioacoustic signals based on human scoring system and artificial neural network. *arXiv preprint arXiv :1305.3633*.
- Priyadarshani, N., Castro, I., and Marsland, S. (2018a). The impact of environmental factors in birdsong acquisition using automated recorders. *Ecology and evolution*, 8(10) :5016–5033.
- Priyadarshani, N., Marsland, S., and Castro, I. (2018b). Automated birdsong recognition in complex acoustic environments : a review. *Journal of Avian Biology*, 49(5) :jav-01447.
- Pukelsheim, F. (1994). The three sigma rule. *The American Statistician*, 48(2) :88–91.
- Purwins, H., Li, B., Virtanen, T., Schlüter, J., Chang, S.-Y., and Sainath, T. (2019). Deep learning for audio signal processing. *Journal of Selected Topics in Signal Processing*, 13(2) :206–219.
- Reby, D., Lek, S., Dimopoulos, I., Joachim, J., Lauga, J., and Aulagnier, S. (1997). Artificial neural networks as a classification method in the behavioural sciences. *Behavioural processes*, 40(1) :35–43.
- Reeve, H. K. (1989). The evolution of conspecific acceptance thresholds. *The American Naturalist*, 133(3) :407–435.
- Rendall, D., Kollias, S., Ney, C., and Lloyd, P. (2005). Pitch (fo) and formant profiles of human vowels and vowel-like baboon grunts : The role of vocalizer body size and voice-acoustic allometry. *The Journal of the Acoustical Society of America*, 117(2) :944–955.
- Rendell, L. E. and Whitehead, H. (2003). Vocal clans in Sperm whales (*Hyseter macrocephalus*). *Proceedings of the Royal Society of London. Series B : Biological Sciences*, 270(1512) :225–231.
- Rhinehart, T. A., Chronister, L. M., Devlin, T., and Kitzes, J. (2020). Acoustic localization of terrestrial wildlife : Current practices and future opportunities. *Ecology and Evolution*.
- Ricard, J. and Glotin, H. (2018). Dynclick : open-source toolbox for stereo click detection, analysis and tracking. DCLDE 2018 Paris Sorbonne. Poster.

Bibliographie

- Rice, A., Chou, E., Hildebrand, J. A., Wiggins, S. M., Roch, M. A., et al. (2015). Seven years of blue and fin whale call abundance in the southern california bight. *Endangered Species Research*, 28(1) :61–76.
- Richardson, W. J., Greene Jr, C. R., Malme, C. I., and Thomson, D. H. (2013). *Marine mammals and noise*. Academic press.
- Rickwood, P. and Taylor, A. (2008). Methods for automatically analyzing humpback song units. *The Journal of the Acoustical Society of America*, 123(3) :1763–1772.
- Ridgway, S. H. (2000). The auditory central nervous system of dolphins. In *Hearing by whales and dolphins*, pages 273–293. Springer.
- Riede, K. (1993). Monitoring biodiversity : analysis of amazonian rainforest sounds. *Ambio*, pages 546–548.
- Riede, K. (1998). Acoustic monitoring of orthoptera and its potential for conservation. *Journal of Insect Conservation*, 2(3-4) :217–223.
- Risch, D., Clark, C. W., Corkeron, P. J., Elepfandt, A., Kovacs, K. M., Lydersen, C., Stirling, I., and Van Parijs, S. M. (2007). Vocalizations of male bearded seals, *erignathus barbatus* : classification and geographical variation. *Animal Behaviour*, 73(5) :747–762.
- Rittschof, D., Kratt, C., and Clare, A. (1990). Gastropod predation sites : the role of predator and prey in chemical attraction of the hermit crab *clibanarius vittatus*. *Journal of the Marine Biological Association of the United Kingdom*, 70(03) :583–596.
- Robert, A., Lengagne, T., Melo, M., Gardette, V., Julien, S., Covas, R., Gomez, D., and Doutrelant, C. (2019). The theory of island biogeography and soundscapes : Species diversity and the organization of acoustic communities. *Journal of Biogeography*, 46(9) :1901–1911.
- Roch, M., Klinck, H., Baumann-Pickering, S., Mellinger, D., Qui, S., Soldevilla, M., and Hildebrand, J. (2011). Classification of echolocation clicks from odontocetes in the southern california bight. *The Journal of the Acoustical Society of America*, 129(1) :467–475.
- Romer, H. and Lewald, J. (1992). High-frequency sound transmission in natural habitats : implications for the evolution of insect acoustic communication. *Behavioral Ecology and Sociobiology*, 29(6) :437–444.
- Roper, B. B. and Scarnecchia, D. L. (1999). Emigration of age-0 chinook salmon (*oncorhynchus tshawytscha*) smolts from the upper south umpqua river basin, oregon, usa. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(6) :939–946.
- Rosenbaum, H. C., Pomilla, C., Mendez, M., Leslie, M. S., Best, P. B., Findlay, K. P., Minton, G., Ersts, P. J., Collins, T., Engel, M. H., et al. (2009). Population structure of humpback whales from their breeding grounds in the south atlantic and indian oceans. *PLoS One*, 4(10).
- Rosenstock, S. S., Anderson, D. R., Giesen, K. M., Leukering, T., and Carter, M. F. (2002). Landbird counting techniques : current practices and an alternative. *The Auk*, 119(1) :46–53.
- Rossing, T. (2007). *Springer handbook of acoustics*. Springer Science & Business Media.
- Rougier, G., Rey, V., and Molcard, A. (2015). Wave-current interactions in deep water conditions : field measurements and analyses. *EGUGA*, page 4719.
- Rountree, R. A., Gilmore, R. G., Goudey, C. A., Hawkins, A. D., Luczkovich, J. J., and Mann, D. A. (2006). Listening to fish : applications of passive acoustics to fisheries science. *Fisheries*, 31(9) :433–446.
- Roy, N., Simard, Y., and Gervaise, C. (2010). 3d tracking of foraging belugas from their clicks : Experiment from a coastal hydrophone array. *Applied Acoustics*, 71(11) :1050–1056.
- Russo, A. P. (2001). Acoustic signature recognition and identification. US Patent 6,173,074.
- Ryan, M. J. and Brenowitz, E. A. (1985). The role of body size, phylogeny, and ambient noise in the evolution of bird song. *The American Naturalist*, 126(1) :87–100.

- Sayigh, L. S., Tyack, P. L., and Wells, R. S. (1993). Recording underwater sounds of free-ranging dolphins while underway in a small boat. *Marine mammal science*, 9(2) :209–213.
- Sayigh, L. S., Tyack, P. L., Wells, R. S., and Scott, M. D. (1990). Signature whistles of free-ranging bottlenose dolphins *Tursiops truncatus* : stability and mother-offspring comparisons. *Behavioral Ecology and Sociobiology*, 26(4) :247–260.
- Schröter, H., Nöth, E., Maier, A., Cheng, R., Barth, V., and Bergler, C. (2019). Segmentation, classification, and visualization of orca calls using deep learning. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8231–8235. IEEE.
- Selouani, S.-A., Kardouchi, M., Hervet, E., and Roy, D. (2005). Automatic birdsong recognition based on autoregressive time-delay neural networks. In *2005 ICSC Congress on Computational Intelligence Methods and Applications*, pages 6–pp. IEEE.
- Sevilla, A. and Glotin, H. (2017). Audio bird classification with inception-v4 extended with time and time-frequency attention mechanisms. In *CLEF Cross-language image retrieval evaluations Working Notes*.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3) :379–423.
- Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1) :10–21.
- Shapiro, A. D. (2006). Preliminary evidence for signature vocalizations among free-ranging narwhals (*Monodon monoceros*). *The Journal of the Acoustical Society of America*, 120(3) :1695–1705.
- Shapiro, A. D. and Wang, C. (2009). A versatile pitch tracking algorithm : From human speech to Killer whale vocalizations. *The Journal of the Acoustical Society of America*, 126(1) :451–459.
- Sharp, S. and Hatchwell, B. (2005). Individuality in the contact calls of cooperatively breeding long-tailed tits (*aegithalos caudatus*). *Behaviour*, 142(11-12) :1559–1575.
- Sheng, K., Dong, W., Li, W., Razik, J., Huang, F., and Hu, B. (2017). Centroid-aware local discriminative metric learning in speaker verification. *Pattern Recognition*, 72 :176–185.
- Shin, S., Yang, H., Kim, J., Choresca Jr, C. H., Han, J., Jun, J., Han, S., Park, S., et al. (2012). Rapid detection and isolation of salmonella sp. from amphibians and reptiles. *African Journal of Biotechnology*, 11(3) :682–686.
- Shirihai, H. (2007). *Guide des mammifères marins du monde : toutes les espèces décrites et illustrées*. Delachaux et Niestle.
- Shonfield, J. and Bayne, E. (2017). Autonomous recording units in avian ecological research : current use and future applications. *Avian Conservation and Ecology*, 12(1).
- Simard, Y., Bahoura, M., Park, C., Rouat, J., Sirois, M., Mouy, X., Seebarruth, D., Roy, N., and Lepage, R. (2006). Development and experimentation of a satellite buoy network for real-time acoustic localization of whales in the st. lawrence. In *OCEANS Conf.*, pages 1–6. IEEE.
- Simard, Y. and Roy, N. (2008). Detection and localization of blue and fin whales from large-aperture autonomous hydrophone arrays : A case study from the st. lawrence estuary. *Canadian Acoustics*, 36(1) :104–110.
- Simard, Y., Roy, N., and Gervaise, C. (2008). Passive acoustic detection and localization of whales : Effects of shipping noise in saguenay–st. lawrence marine park. *The Journal of the Acoustical Society of America*, 123(6) :4109–4117.
- Simmonds, M., Dolman, S., and Weilgart, L. (2004). *Oceans of noise*. Whale and dolphin conservation Society.
- Simonis, A. E., Roch, M. A., Bailey, B., Barlow, J., Clemesha, R. E., Iacobellis, S., Hildebrand, J. A., and Baumann-Pickering, S. (2017). Lunar cycles affect common dolphin *Delphinus delphis* foraging in the southern california bight. *Marine Ecology Progress Series*, 577 :221–235.
- Sinsch, U., Lümekemann, K., Rosar, K., Schwarz, C., and Dehling, M. (2012). Acoustic niche partitioning in an anuran community inhabiting an afromontane wetland (butare, rwanda). *African Zoology*, 47(1) :60–73.
- Skowronski, M. D. and Harris, J. G. (2006). Acoustic detection and classification of microchiroptera using machine learning : lessons learned from automatic speech recognition. *The Journal of the Acoustical Society of America*, 119(3) :1817–1833.

Bibliographie

- Slabbekoorn, H., Ellers, J., and Smith, T. B. (2002). Birdsong and sound transmission : the benefits of reverberations. *The Condor*, 104(3) :564–573.
- Slater, P. (1986). The cultural transmission of bird song. *Trends in Ecology & Evolution*, 1(4) :94–97.
- Slater, P. J. and Mann, N. I. (2004). Why do the females of many bird species sing in the tropics? *Journal of Avian Biology*, 35(4) :289–294.
- Smolker, R., Mann, J., and Smuts, B. (1993). Use of signature whistles during separations and reunions by wild bottlenose dolphin mothers and infants. *Behavioral Ecology and Sociobiology*, 33(6) :393–402.
- Snowdon, C. T. and Cleveland, J. (1980). Individual recognition of contact calls by pygmy marmosets. *Animal Behaviour*, 28(3) :717–727.
- Soriano, T., Hien, N., Tuan, K., and Anh, T. (2016). An object-unified approach to develop controllers for autonomous underwater vehicles. *Mechatronics*, 35 :54–70.
- Sousa-Lima, R. (2005). Songs indicate interaction between humpback whale (megaptera novaeangliae) populations in the western and eastern south atlantic ocean. *Marine Mammal Science*, 21(3) :557–566.
- Sousa-Lima, R. S., Engel, M. H., Sábato, V., Lima, B. R., Queiróz, T. S., Brito, M. R., Fernandes, D. P., Martins, C. A., Hatum, P. S., Casagrande, T., et al. (2018). Acoustic ecology of humpback whales in brazilian waters investigated with basic and sophisticated passive acoustic technologies over 17 years. *Western Indian Ocean Journal of Marine Science*, pages 23–40.
- Sousa-Lima, R. S., Norris, T. F., Oswald, J. N., and Fernandes, D. P. (2013). A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals. *Aquatic Mammals*, 39(1).
- Sousa-Lima, R. S., Paglia, A. P., and Da Fonseca, G. A. (2002). Signature information and individual recognition in the isolation calls of amazonian manatees, *Trichechus inunguis* (Mammalia : Sirenia). *Animal Behaviour*, 63(2) :301–310.
- Southall, B. L., Schusterman, R. J., and Kastak, D. (2003). Auditory masking in three pinnipeds : Aerial critical ratios and direct critical bandwidth measurements. *The Journal of the Acoustical Society of America*, 114(3) :1660–1666.
- Sprengel, E., Jaggi, M., Kilcher, Y., and Hofmann, T. (2016). Audio based bird species identification using deep learning techniques. Technical report, Eidgenössische Technische Hochschule (ETH).
- Stone, E. (2000). Separating the noise from the noise : a finding in support of the “niche hypothesis,” that birds are influenced by human-induced noise in natural habitats. *Anthrozoös*, 13(4) :225–231.
- Stowell, D., Wood, M. D., Pamuła, H., Stylianou, Y., and Glotin, H. (2019). Automatic acoustic detection of birds through deep learning : the first bird audio detection challenge. *Methods in Ecology and Evolution*, 10(3) :368–380.
- Strehl, A. and Ghosh, J. (2002). Cluster ensembles, a knowledge reuse framework for combining multiple partitions. *Journal of machine learning research*, 3(Dec) :583–617.
- Sueur, J. (2002). Cicada acoustic communication : potential sound partitioning in a multispecies community from mexico (hemiptera : Cicadomorpha icadidae). *Biological Journal of the Linnean Society*, 75(3) :379–394.
- Sueur, J. and Farina, A. (2015). Ecoacoustics : the ecological investigation and interpretation of environmental sound. *Biosemiotics*, 8(3) :493–502.
- Sueur, J., Farina, A., Gasc, A., Pieretti, N., and Pavoine, S. (2014). Acoustic indices for biodiversity assessment and landscape investigation. *Acta Acustica united with Acustica*, 100(4) :772–781.
- Sueur, J., Pavoine, S., Hamerlynck, O., and Duvail, S. (2008). Rapid acoustic survey for biodiversity appraisal. *PloS one*, 3(12).
- Suter, S. M., Giordano, M., Nietlispach, S., Apollonio, M., and Passilongo, D. (2017). Non-invasive acoustic detection of wolves. *Bioacoustics*, 26(3) :237–248.
- Suthers, R. A. (2004). How birds sing and why it matters. *Nature’s music : the science of birdsong*. Elsevier Academic Press, San Diego, pages 272–295.

- Taga, M. E. and Bassler, B. L. (2003). Chemical communication among bacteria. *Proceedings of the National Academy of Sciences*, 100(suppl 2) :14549–14554.
- Talus, C. E. (2000). Analysis of the vocalizations of (*Orcinus orca*) in response to anthropogenic noise. Master's thesis, University of Alaska Fairbanks.
- Terrell, G. R. and Scott, D. W. (1992). Variable kernel density estimation. *The Annals of Statistics*, pages 1236–1265.
- Thode, A. (2004). Tracking sperm whale (*physeter macrocephalus*) dive profiles using a towed passive acoustic array. *The Journal of the Acoustical Society of America*, 116(1) :245–253.
- Thomas, J. A. and Stirling, I. (1983). Geographic variation in the underwater vocalizations of weddell seals (*Leptonychotes weddelli*) from palmer peninsula and mcmurdo sound, antarctica. *Canadian Journal of Zoology*, 61(10) :2203–2212.
- Thompson, P., Cummings, W., and Kennison, S. (1977). Sound production of humpback whales, *Megaptera novaeangliae*, in alaskan waters. *The Journal of the Acoustical Society of America*, 62(S1) :S89–S89.
- Tinbergen, N. (1948). Social releasers and the experimental method required for their study. *The Wilson Bulletin*, pages 6–51.
- Todt, D. and Naguib, M. (2000). Vocal interactions in birds : the use of song as a model in communication. In *Advances in the Study of Behavior*, volume 29, pages 247–296. Elsevier.
- Towers, J. (2015). *Photo-identification catalogue and status of the northern resident Killer whale population in 2014*. Fisheries and Oceans Canada.
- Trudelle, L., Cerchio, S., Zerbini, A. N., Geyer, Y., Mayer, F.-X., Jung, J.-L., Hervé, M. R., Pous, S., Sallée, J.-B., Rosenbaum, H. C., et al. (2016). Influence of environmental parameters on movements and habitat utilization of humpback whales *Megaptera novaeangliae* in the madagascar breeding ground. *Royal Society open science*, 3(12) :160616.
- Truskinger, A., Cottman-Fields, M., Johnson, D., and Roe, P. (2013). Rapid scanning of spectrograms for efficient identification of bioacoustic events in big data. In *2013 IEEE 9th International Conference on e-Science*, pages 270–277. IEEE.
- Van der Maaten, L. and Hinton, G. (2012). Visualizing non-metric similarities in multiple maps. *Machine learning*, 87(1) :33–55.
- Van Wilgenburg, S., Sólymos, P., Kardynal, K., and Frey, M. (2017). Paired sampling standardizes point count data from humans and acoustic recorders. *Avian Conservation and Ecology*, 12(1).
- Vellinga, W.-P. and Planqué, R. (2015). The xeno-canto collection and its relation to sound recognition and classification. In *CLEF (Working Notes)*.
- Vester, H. (2017). *Vocal repertoires of two matrilineal social whale species Long-finned Pilot whales (*Globicephala melas*) & Killer whales (*Orcinus orca*) in northern Norway*. PhD thesis, Phd Thesis, Georg-August-Universität Göttingen.
- Vignal, C. and Kelley, D. (2007). Significance of temporal and spectral acoustic cues for sexual recognition in *xenopus laevis*. *Proceedings of the Royal Society B : Biological Sciences*, 274(1609) :479–488.
- Vignal, C., Mathevon, N., and Mottin, S. (2004). Audience drives male songbird response to partner's voice. *Nature*, 430(6998) :448–451.
- Voituron, Y., Brepson, L., Richardson, C., Joly, P., and Lengagne, T. (2012). Energetics of calling in the male treefrog *hyla arborea* : when being large means being sexy at low cost, 149, 775–793.
- Wahlberg, M. (2002). The acoustic behaviour of diving Sperm whales observed with a hydrophone array. *Journal of Experimental Marine Biology and Ecology*, 281(1-2) :53–62.
- Watkins, W. A., Daher, M. A., Fristrup, K. M., Howald, T. J., and Di Sciara, G. N. (1993). Sperm whales tagged with transponders and tracked underwater by sonar. *Marine mammal science*, 9(1) :55–67.

Bibliographie

- Watwood, S. L., Miller, P. J., Johnson, M., Madsen, P. T., and Tyack, P. L. (2006). Deep-diving foraging behaviour of Sperm whales (*Physeter macrocephalus*). *Journal of Animal Ecology*, 75(3) :814–825.
- Weilgart, L. and Whitehead, H. (1997). Group-specific dialects and geographical variation in coda repertoire in south pacific Sperm whales. *Behavioral Ecology and Sociobiology*, 40(5) :277–285.
- Weiß, B. M., Symonds, H., Spong, P., and Ladich, F. (2007). Intra-and intergroup vocal behavior in resident Killer whales, (*Orcinus orca*). *The Journal of the Acoustical Society of America*, 122(6) :3710–3716.
- Welbergen, J. A. and Davies, N. B. (2008). Reed warblers discriminate cuckoos from sparrowhawks with graded alarm signals that attract mates and neighbours. *Animal Behaviour*, 76(3) :811–822.
- Whitehead, H. (2003). *Sperm whales : social evolution in the ocean*. University of Chicago press.
- Whitehead, H., Waters, S., and Lyrholm, T. (1991). Social organization of female Sperm whales and their offspring : constant companions and casual acquaintances. *Behavioral Ecology and Sociobiology*, 29(5) :385–389.
- Whitehead, H. and Weilgart, L. (1991). Patterns of visually observable behaviour and vocalizations in groups of female sperm whales. *Behaviour*, 118(3-4) :275–296.
- Wiggins, S. (2003). Autonomous acoustic recording packages (arps) for long-term monitoring of whale sounds. *Marine Technology Society Journal*, 37(2) :13–22.
- Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in statistics*, pages 196–202. Springer.
- Wiley, R. H. (1982). Adaptations for acoustic communication in birds : sound transmission and signal detection. *Acoustic communication in birds*, 1 :131–181.
- Wiley, R. H. and Richards, D. G. (1978). Physical constraints on acoustic communication in the atmosphere : implications for the evolution of animal vocalizations. *Behavioral ecology and sociobiology*, 3(1) :69–94.
- Williams, R., Erbe, C., Ashe, E., Beerman, A., and Smith, J. (2014). Severity of Killer whale behavioral responses to ship noise : a dose–response study. *Marine pollution bulletin*, 79(1-2) :254–260.
- Williams, R. and Noren, D. P. (2009). Swimming speed, respiration rate, and estimated cost of transport in adult Killer whales. *Marine Mammal Science*, 25(2) :327–350.
- Williams, R., Trites, A. W., and Bain, D. E. (2002). Behavioural responses of Killer whales (*Orcinus orca*) to whale-watching boats : opportunistic observations and experimental approaches. *Journal of Zoology*, 256(2) :255–270.
- Wilson, R. P., Liebsch, N., Davies, I. M., Quintana, F., Weimerskirch, H., Storch, S., Lucke, K., Siebert, U., Zankl, S., Müller, G., et al. (2007). All at sea with animal tracks ; methodological and analytical solutions for the resolution of movement. *Deep Sea Research Part II : Topical Studies in Oceanography*, 54(3-4) :193–210.
- Winn, H., Thompson, T., Cummings, W., Hain, J., Hudnall, J., Hays, H., and Steiner, W. (1981). Song of the humpback whale—population comparisons. *Behavioral ecology and sociobiology*, 8(1) :41–46.
- Winn, H. and Winn, L. (1978). The song of the humpback whale *Megaptera novaeangliae* in the west indies. *Marine Biology*, 47(2) :97–114.
- Witteveen, B. H., Foy, R. J., Wynne, K. M., and Tremblay, Y. (2008). Investigation of foraging habits and prey selection by humpback whales *Megaptera novaeangliae* using acoustic tags and concurrent fish surveys. *Marine Mammal Science*, 24(3) :516–534.
- Yang, W., Luo, W., and Zhang, Y. (2020). Classification of odontocete echolocation clicks using convolutional neural network. *The Journal of the Acoustical Society of America*, 147(1) :49–55.
- Zahavi, A. (1975). Mate selection—a selection for a handicap. *Journal of theoretical Biology*, 53(1) :205–214.
- Zimmer, W. M., Madsen, P. T., Teloni, V., Johnson, M. P., and Tyack, P. L. (2005). Off-axis effects on the multipulse structure of Sperm whale usual clicks with implications for sound production. *The Journal of the Acoustical Society of America*, 118(5) :3337–3345.
- Zollinger, S. A., Riede, T., and Suthers, R. A. (2008). Two-voice complexity from a single side of the syrinx in northern mockingbird mimus polyglottos vocalizations. *Journal of Experimental Biology*, 211(12) :1978–1991.

Liste des publications de l'auteur

Articles de conférence

- **Poupard**, M., Best, P., Schlüter, J., Prévot, J. M., Symonds, H., Spong, P., & Glotin, H. (2019, June). Deep Learning for Ethoacoustics of Orcas on three years pentaphonic continuous recording at Orcalab revealing tide, moon and diel effects. In OCEANS (7 p). IEEE Ed.
- **Poupard**, M., Ferrari, M., Schluter, J., Marxer, R., Giraudet, P., Barchasz, V., ... & Glotin, H. (2019, May). Real-time passive acoustic 3d tracking of deep diving cetacean by small non-uniform mobile surface antenna. In ICASSP International Conference on Acoustics, Speech and Signal Processing (pp. 8251-8255). IEEE Ed.
- **Poupard**, M., Ferrari, M., Schluter, J., Astruch, P., Schohn, B., Rouanet, B., ... & Gies, V. (2019, April). Passive acoustics to monitor flagship species near boat traffic in the unesco world heritage natural reserve of scandola. In Input Academy : International Conference on Innovation in Urban and regional planning.
- **Poupard**, M., Best, P., Schlüter, J., Symonds, H., Spong, P., & Glotin, H. (2019). Large-scale unsupervised clustering of Orca vocalizations : a model for describing orca communication systems (No. e27979v1). VIHAR. PeerJ Preprints Ed.
- **Poupard**, M., de Montgolfier, B., & Glotin, H. (2019, June). Ethoacoustics by bayesian non parametric and stochastic neighbor embedding to forecast anthropic pressure on dolphins. In OCEANS (5 p). IEEE Ed.
- Joly, A., Goëau, H., Botella, C., Kahl, S., **Poupard**, M., Servajean, M., ..., & Schlüter, J. (2019, April). Lifeclef 2019 : Biodiversity identification and prediction challenges. In European Conference on Information Retrieval (pp. 275-282). Springer, Cham Ed.
- Ferrari, M., **Poupard**, M., Giraudet, P., Marxer, R., Prévot, J. M., Soriano, T., & Glotin, H. (2019, June). Efficient artifacts filter by density-based clustering in long term 3D whale passive acoustic monitoring with five hydrophones fixed under an Autonomous Surface Vehicle. In OCEANS (7 p). IEEE Ed.
- Best, P., Ferrari, M., **Poupard**, M., Paris, S., Marxer, R., Symonds, H., ..., & Glotin, H. (2020, July). Deep Learning and Domain Transfer for Orca Vocalization Detection. In International joint conference on neural networks, IJCNN Ed.
- **Poupard**, M., De Montgolfier, B., Roger, V., Lohani, D., & Glotin, H. (2018). EthoAcoustics : a model based on t-SNE & Clustering, applied on Pantropical spotted dolphin during Whale Watching, Detection, Classification, Localization, and Density Estimation of marine mammals using passive acoustics, DCLDE, Paris Sorbonne Ed.

Articles de revue

- **Poupard**, M., Symonds, H., Spong, P., & Glotin, H. (2020). Individual separation and identification of orcas calls in the wild : study of inter and intra-group communication (16 p). To be submitted to Scientific Report, Nature Communication Ed
- Bouchard, B., Barnagaud, J.-Y., **Poupard**, M., Glotin, H., Gauffier, P., Torres Ortiz, S., Lisney, T. J., Campagna, S., Rasmussen, M., & Célérier, A. (2019). Behavioural responses of humpback whales to food-related chemical stimuli. PloS one, 14(2) :e0212515

Chapitre dans ouvrage

- Kobayashi, H. H., Kudo, H., Glotin, H., Roger, V., **Poupard**, M., Shimotoku, D., ... & Sezaki, K. (2018). A Real-Time Streaming and Detection System for Bio-acoustic Ecological Studies after the Fukushima Accident. In Multimedia Tools and Applications for Environmental and Biodiversity Informatics (pp. 53-66). Springer Book Ed.

Rapports scientifiques

- Glotin, H., Giraudet, P., Ricard, J., Malige, F., Patris, J., Roger, V., Prévot, J.-M., **Poupard**, M., Philippe, O., and Cosentino, P. (2016). Visées aériennes de mammifères marins jointes aux observations acoustiques sous-marines de bombyx et antares : nouveaux modèles en suivis et lois allométriques du *Physeter macrocephalus*, *Ziphius cavirostris* et autres cétacés, Rapport PELAGOS, www.sanctuaire-pelagos.org Ed., http://sabiiod.org/pub/PELAGOS_VAMOS.pdf
- Glotin, H., Thellier, N., Best, P., **Poupard**, M., Ferrari, M., Sara Viera and, V. G., Oger, M., Giraudet, P., Donzé, M. M. G., Campana, M., Chevallier, J., Malige, F., Patris, J., Prévot, J.-M., Cosentino, P., d'Alvise, N. P., Ourmières, Y., Barchasz, V., Lehir, A., Marzetti, S., Sarano, F., Gaillard, S., and de Varenne, F. (2020). Missions Sphyrna Odyssey : Découvertes éthoacoustiques de chasses collaboratives de Cachalots en abysse et impacts en mer du confinement COVID19. CNRS, Université de Toulon, Explorations de Monaco Ed., <http://sabiiod.org/pub/S01.pdf>

Table des figures

1.1	Schéma global de la structure de la communication animale selon Shannon (1948).	10
1.2	Exemple de production acoustique de 0.1Hz à 1GHz en fonction des capacités de l'oreille humaine (Bradbury et al., 1998).	12
1.3	Forme d'onde et spectrogramme de 8 secondes de chants d'oiseaux enregistrés au jardin du Rayol (Var) en 2019 par l'équipe DYNI (micro et carte JASON de SMIoT).	13
1.4	Exemple d'audiogramme de différents mammifères marins (odontocètes, mysticètes, pinnipèdes) et des humains (NOAA, 1998).	17
1.5	Représentation d'un clic de Cachalot enregistré à l'Ile Maurice en 2019. Haut : forme d'onde. Bas : spectrogramme (fenêtre de Hanning, taille de la FFT 256) . Protocole réalisé par l'équipe DYNI (Université de Toulon), carte son JASON SMIoT et Hydrophone C75 Cetacean Research.	20
1.6	Représentation des processus indispensables à la reconnaissance d'appartenance à une population (Bradbury et al., 1998).	24
1.7	Schéma tiré de Johnstone (1997) résumant les mécanismes de reconnaissance et la distribution des fréquences des deux types d'individus reconnus ou non.	26
1.8	Représentation de 6 spectrogrammes appartenants à deux individus d'un même groupe (<i>Turisops truncatus</i>). a) : un mâle, b) : une femelle (les harmoniques ont été supprimées). Figure tirée de Janik and Slater (1998).	27
1.9	Schéma représentant les différentes échelles d'analyses ainsi que leurs cas d'étude associés.	31
1.10	Schéma représentant les différentes étapes réalisées dans cette thèse.	33
1.11	Schéma représentant les chapitres Matériels, Méthodes et Résultats ainsi que les différentes options de lecture du manuscrit (couleurs).	34
2.1	Représentation du Cachalot (<i>Physeter macrocephalus</i>) associée à sa répartition géographique. Source : NOAA Fisheries (2011a).	36
2.2	Représentation de l'Orque (<i>Orcinus orca</i>) associée à sa répartition géographique. Source : NOAA Fisheries (2011b).	37
2.3	Organisation des Orques résidentes de la Colombie-Britannique. Les résidents sont regroupés en 2 communautés (Sud et Nord). Ces communautés sont composées de différents pods. Et chacun des pods est constitué de matriline qui sont composées d'individus liés de manière générationnelle (Ford et al., 1987).	38
2.4	Représentation du Dauphin tacheté pantropical (<i>Stenella attenuata</i>) associée à sa répartition géographique. Source : American Cetacean Society (2017).	38

2.5	Forme d'onde (haut) et spectrogramme (bas) de sifflements de Dauphins tachetés pantropiaux (<i>Stenella attenuata</i>) en Martinique. Spectrogramme : fenêtre de Hanning de taille 256 échantillons.	39
2.6	Représentation de la Baleine à bosse (<i>Megaptera novaeangliae</i>) associée à sa répartition géographique (Source : American Cetacean Society, National Oceanic and atmospheric administration)	40
2.7	Spectrogramme et forme d'onde de signaux de Baleine à bosse (<i>Megaptera novaeangliae</i>).	41
2.8	Schéma représentant une acquisition globale en bioacoustique. 1 : Émission acoustique de la faune, provoquant une onde sonore. 2 : Cette onde sonore est captée par un microphone/hydrophone. 3 : Transformation de l'onde sonore en un signal électrique transmis à un enregistreur (ou carte son) (4). 5 : Conversion de la tension électrique en valeurs numériques et enregistrement. 6 : Archivage sur des disques durs (ex : NAS). La dernière étape est la validation de la qualité des enregistrements.	43
2.9	Téléphones utilisés ainsi que l'enregistreur type "Zoom". Le cinquième téléphone est utilisé pour prendre la photo. Photographie tirée du rapport de recherche Poupard et al. (2018c).	46
2.10	Schéma représentant les émissions sonores, des espèces seules, puis des couples à différents niveaux d'énergie ainsi que la photographie de l'enceinte émettrice.	47
2.11	Gauche : Schéma de la station d'acquisition BOMBYX au large de Port-Cros. Droite : installation des 2 hydrophones.	48
2.12	Profil vertical de la direction moyenne des vitesses en 2015, les barres horizontales représentent l'écart-type. Figure tirée de Rougier et al. (2015).	49
2.13	Enregistreurs les plus couramment utilisés pour l'acquisition de sons d'oiseaux.	51
2.14	Représentation de l'acquisition des données, des enregistrements à Orcalab jusqu'à l'archivage sur les serveurs DYNi à Toulon.	52
2.15	Carte de la zone et des plages d'écoutes des 5 hydrophones. Les repères de H1 à H5 en gras indiquent l'emplacement des hydrophones. Les zones de détections acoustiques estimées sont indiquées en couleurs (estimées par les experts d'Orcalab).	53
2.16	L'antenne à quatre hydrophones, avec le câble 12 V (vert) connecté à la station de recherche Orcalab.	54
2.17	Carte d'acquisition Jason (SMIoT), 5 x 2 MHz. Le luxmètre n'a pas été utilisé pour ce protocole.	55
2.18	Protocole d'acquisition visuel à Orcalab. Crédit Y. Rose.	56
2.19	Représentation de la vue du Blackney Pass depuis la plateforme d'Orcalab.	56
2.20	Carte des plages visuelles et d'écoutes, couvrant une zone d'environ 5km ² avec une possible estimation azimutale audio et visuelle des passages des Orques.	57
2.21	Drone Sphyrna en mer (Crédit P. Cottaz).	58
2.22	Disposition du réseau d'hydrophones 4 + 1. Les distances interhydrophones sous la quille sont de 35, 59, 59, 63, 63, 70cm. L'hydrophone H5 est placé à l'arrière, à 7 m de distance de H4	59

2.23	Protocole réalisé pour la détection de source en milieu terrestre.	60
2.24	Présentation de la zone d'étude et du matériel. 1 : Hydrophone (dans sa protection), 2 : Enregistreur, 3 : Casque audio.	62
2.25	Gauche : Carte de la zone d'étude. Droite : flotteur contenant une batterie, un enregistreur (1), le câble des hydrophones (2) préparés par l'équipe DYNI (Bouchard et al., 2019). . .	64
2.26	Gauche : Photographie des deux stimuli : Clay (argile) et krill. Droite : Photographie de la zone de stimuli.	64
2.27	Résumé des différentes problématiques en bioacoustique associées à chaque échelle d'analyse.	66
3.1	Résumé des étapes globales d'un CNN.	69
3.2	Résumé des étapes globales de la classification supervisée.	69
3.3	Comparaisons des principaux algorithmes de clustering proposés par la librairie Scikit-learn (Halkidi et al., 2001). DBSCAN, en 7ème colonne, partitionne correctement pratiquement tous les ensembles.	71
3.4	Représentation théorique des distances utilisées en ACP (tiret) et t-SNE (continu).	72
3.5	Résumé des méthodes non supervisées.	72
3.6	Architecture du modèle utilisé. Image tirée de Kahl et al. (2017).	75
3.7	Matrice de confusion du test set pour la tâche de classification de 52 espèces d'oiseaux, avec un TBC de 88%.	78
3.8	Principales étapes de l'extraction des événements acoustiques : Binarisation et détection des composants connectés. Le spectrogramme montre des fréquences de 0 à 6,5 kHz pendant 2,5 sec.	79
3.9	Image tirée de Grill and Schlüter (2017) représentant l'architecture du modèle utilisé pour la détection d'Orques.	80
3.10	Représentation de la courbe "ROC" (Receiver Operating Characteristic) du détecteur acoustique d'Orques.	81
3.11	Représentation des différents filtrages réalisés pour détecter les clics de Cachalot pour un enregistrement de 7 secondes.	82
3.12	Exemple d'un clic visualisé sous Audacity, sur les deux voies de BOMBYX. La double flèche rouge représente le TDoA entre les deux voies.	83
3.13	Représentation de l'outil d'annotation réalisé par l'équipe DYNI. Haut : affichage de plusieurs heures de TDoAs. Bas : Affichage d'un clic en spectrogramme (point rouge) représentant un clic qui vient d'être écouté.	83
3.14	Affichage de la piste en TDoAs du Cachalot le 25 Décembre 2018.	84
3.15	Exemple de la piste TDoAs d'un Cachalot, avec en haut la piste sur les 4h (continu). Puis la piste décimée et concaténée à 25%, puis à 16%.	85

Table des figures

3.16	Diagramme représentant les étapes réalisées pour la détection de Cachalots sur la bouée BOMBYX. L'expertise manuelle pour l'extraction des pistes de Cachalots s'est faite via l'écoute du son, l'analyse de la forme d'onde et de la visualisation du spectrogramme du clic.	86
3.17	Diagramme représentant l'organisation hiérarchique de la communauté des résidents du Nord. Figure inspirée de Ford (1991).	87
3.18	Exemple d'extraction du pitch tirée de Poupard et al. (2019b).	90
3.19	Exemple d'une annotation de 6 secondes réalisée par Helena Symonds faite sous Audacity.	91
3.20	Architecture de la méthode utilisée pour l'extraction des vocalises pour l'identification individuelle.	92
3.21	Représentation des différents filtrages pour la détection de clics de Cachalots enregistrés depuis le drone Sphyrna.	94
3.22	En haut : filtrage des valeurs de corrélation croisée (H_4 et H_5). En bas : TDoAs entre H_1 et H_2 au fil du temps. Les points jaunes ne sont pas utilisés pour l'analyse.	95
3.23	Spectrogramme de 13 secondes d'un enregistrement contenant des signaux de <i>Stenella attenuata</i> et la représentation finale suite à la détection automatique pour les différentes fenêtres.	98
3.24	Architecture de l'analyse pour la bioacoustique comportementale du Dauphin tacheté pantropical.	99
3.25	Représentation de l'ensemble des méthodes utilisées dans cette thèse. Les carrés verts : méthodes automatiques, carrés gris : méthodes manuelles, carrés blancs : cas d'étude.	101
4.1	Spectrogrammes (échelle Mel) représentant les différences de dB entre les sons reçus. Haut Gauche le spectrogramme initial du son émis. Bas gauche : le sons reçu à 10m. Haut droit : son reçu à 50m, Bas droit : son perçu à 100m. Le son émis était 15 secondes de Troglodyte mignon enregistrées avec un iphone 4.	103
4.2	Probabilité de reconnaissance du Troglodyte mignon (<i>Troglodytes troglodytes</i>) en fonction du téléphone, de la distance d'émission et de l'angle d'écoute.	104
4.3	Évolution des probabilités maximales du modèle en fonction de la distance d'émission en émettant le chant du Troglodyte mignon.	104
4.4	Graphique récapitulatif des scores de TBC en fonction des différents modèles disponibles sur le web comparé à notre modèle DynetBird. Il y a une différence importante entre elles (toutes ont un TBC entre 50 et 33%) et la notre à 86%. Voir rapport de recherche Poupard et al. (2018d).	105
4.5	Proportion d'enregistrements de deux minutes contenant des vocalises d'Orques par mois et par hydrophone, de 2015 à 2017.	106
4.6	Évolution du nombre d'événements acoustiques d'Orques détectées de 5h30 à 8h30 le 24 août 2017. Résultats issus du détecteur "d'évènements acoustiques".	107

4.7	Exemple d'évolution de la probabilité de détection de vocalises pour chaque hydrophone sur une journée (24 août 2017). Le matin, un groupe d'Orques vient de l'est (voir carte 2.15) sur H ₅ , et se déplace sur H ₄ , H ₃ , H ₂ puis sur H ₁ . Différents voyages allers-retours sont effectués pendant la journée. Résultat issu du CCN.	108
4.8	Pourcentage d'enregistrements avec des détections acoustiques d'Orques en fonction des conditions temporelles. Marée M= Marée Montante, Marée D= marée Descendante.	109
4.9	Effort des enregistrements sur la bouée BOMBYX, Jaune : enregistrements. Noir : sans enregistrement de 2015 à 2018.	110
4.10	Calendrier représentant la présence de Cachalots sur l'ensemble des sessions. Les cases jaunes claires représentent les jours sans Cachalot, et les cases blanches les jours sans enregistrement.	110
4.11	Moyenne du nombre de sessions (1min) annotées positives par jour (24h) pour la nuit et le jour avec sa différence significative.	112
4.12	Évolution de la somme des sessions annotées positives à la présence de Cachalots au cours de la journée.	113
4.13	Histogramme du nombre de pistes par jour.	113
4.14	Haut : pistes possibles théoriques (a à e) en terme de temps de délais d'arrivée. Bas : pistes correspondantes projetées dans l'axe de la bouée.	114
4.15	Nombre des passages de 1 à 9 individus avec leur sens de déplacement (a à e)	115
4.16	Haut : Affichage de la piste du Cachalot le 25 décembre 2018 en TDoAs. Milieu : Affichage de la piste du Cachalot, chaque couleur représente un cluster trouvé avec les paramètres suivants $eps=2$, échantillon minimum=3. Il a été nécessaire de dilater l'échelle temporelle afin que les clusters soient équilibrés sur les échelles des TDoAs et sur l'échelle temporelle. Bas : piste du Cachalot en TDoAs, en supprimant les points appartenant aux clusters avec une taille inférieure à 5 points. Chaque accolade rouge correspond au temps de surface ou l'animal ne produit pas de clic.	116
4.17	Comparaison des détecteurs : à gauche les résultats développés dans l'étude 'VAMOS' (Glotin et al., 2017), à droite ceux développés pendant la thèse, appliqués sur 10 jours de Juillet 2017.	117
4.18	Évolution des prédictions de détections (maximum, médiane et 95 percentile) des 3 classes (A, G et bruit) pour deux passages du clan G.	119
4.19	Forme d'onde et spectrogramme des deux types d'appels pour l'entraînement. Droite : No ₄ (clan A). Gauche N ₂₃ (clan G).	120
4.20	Résumé des passages en fonction des marées lorsque les animaux se dirigent vers le Sud et vers le Nord.	121
4.21	Liste des vocalises trouvées et annotées dans les enregistrements (Ford et al., 1987).	123
4.22	Graphique du passage du 30 août. Chaque point vert représente une vocalisation placée en azimuth. La courbe orange est la position interpolée du mâle A66 détectée visuellement et la courbe bleue est le groupe 1.	123

Table des figures

4.23	Graphique du passage du 30 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel (voir figure 4.21).	124
4.24	Graphique des passages de la dernière partie du 30 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel. Haut : Trajectoire du pod A01, Bas : Trajectoire du pod A05.	125
4.25	Densité du CR dans le temps, pour le 30 d'août.	125
4.26	Représentation de la t-SNE des MFCC (29 + 30 + 30) pour les No4s pour le 30 août.	126
4.27	Graphique du passage du 31 août. Chaque point vert représente une vocalisation placée en azimut. La courbe bleue est la position d'interpolation du mâle I76 détectée visuellement, la courbe orange est le groupe 1.	127
4.28	Graphique du passage du 31 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel. Les types les plus courants sont N25, N48 et N23i.	127
4.29	t-SNE des MFCC pour les appels N48 (à gauche) et N25 (à droite) le 31 août 2019. Les points oranges sont produits par le mâle I76 tandis que les points bleus sont produits par le groupe 1. Les points encadrés sont les vocalisations représentées Fig.4.30.	128
4.30	Exemple d'un spectrogramme de N48 du groupe 1 et du mâle I76.	129
4.31	Haut : Fonction densité proportionnelle au CR dans le temps. Bleu : la densité des appels du groupe 1. Orange : la densité des appels du mâle I76. Bas : Graphique du passage du 31 août. La ligne rouge représente le regroupement des individus.	130
4.32	Photographie prise pendant le passage du bateau de croisière le 11 août.	131
4.33	Graphique du passage du 11 août. Chaque point vert représente une vocalisation placée en azimut. Les différentes courbes sont les positions de chaque individu détecté visuellement.	131
4.34	Graphique du passage du 11 août. Chaque point représente une vocalisation placée en azimut. Chaque couleur représente un type d'appel. Les types les plus courants sont No4, No3, N48.	132
4.35	Haut : Fonction densité proportionnelle au CR dans le temps. Bas : Graphique du passage du 11 août. La ligne rouge représente l'arrivée du bateau.	133
4.36	De haut en bas : position du Cachalot, latitude X (Nord), longitude Y (Est), profondeur Z (m), log (ICI) (s), log (tortuosité) dans le temps.	135
4.37	Spectrogramme de 6 secondes contenant deux individus (Haut) associé au détecteur automatique d'évènements acoustiques (Bas).	136
4.38	Représentation 3D des trois TDoAs indépendants pour deux exemples contenant deux individus. Chaque couleur correspond au clusters attribués par le DBSCAN.	137
4.39	Visualisation d'une projection ACP de sifflements de <i>Stenella attenuata</i> en fonction de leur comportement et leurs vitesse de nage (avec 4 paramètres utilisés : Fréquence initiale, finale, durée et vitesse des sifflements mis en évidence par regroupement acoustique automatisé, NMI=0.11).	139

4.40	Visualisation d'une projection t-SNE des sifflements de <i>Stenella attenuata</i> en fonction de leurs comportement et leurs vitesse de nage (avec 4 paramètres utilisés : fréquence initiale, finale, durée et vitesse des sifflements) mis en évidence par regroupement acoustique automatisé, NMI=0.30.	140
4.41	Variation des fréquences maximales des sifflements en fonction des différents comportements. Les lettres différentes indiquent des différences significatives entre les comportements.	141
4.42	Variation de la durée des sifflements en fonction des différents comportements. Les lettres différentes indiquent des variations significatives entre les comportements.	141
4.43	Analyse des vocalisations des baleines émises à Madagascar lors des expériences d'exposition. Après la réduction de la dimensionalité (t-SNE) de tous les paramètres acoustiques, un clustering non paramétrique bayésien (BNP) est appliqué aux données. Un score NMI maximal (0,27) est obtenu en utilisant 3 clusters.	143
5.1	Image du prototype sur mobile. Gauche : Lancement de l'enregistrement, sélection du spectrogramme par l'utilisateur, et analyse de l'enregistrement. Milieu et droite : Réponse du modèle avec la fiche d'identification de l'espèce correspondante.	149
5.2	Représentation de la forme d'onde et du spectrogramme de 6 secondes de vocalises d'Orques (stationnaires à gauche) et de 6 secondes de clics d'un Cachalot (transitoires à droite).	153
5.3	Spectrogrammes représentant les différences de dB entre les sons émis par l'enceinte (en haut avec bonnette et en bas sans bonnette à 12,5 m) de 15 secondes de Troglodyte mignon. Enregistrements réalisés Par T. Lengagne et moi même à Lyon, 2018.	157
5.4	Résumé des méthodes utilisées associées à chaque cas d'étude. Orange : méthodes supervisées. Bleu : détections issues de méthode semi-manuelles (filtre, suivis TF). Gris : localisation.	160
5.5	Résumé des différentes comparaisons discutées dans cette partie. Tr= Transitoires. St= stationnaires. Triangles bleus : travaux issus de ce travail de thèse. Croix bleues : travaux pas encore réalisés. Nombre (1, 2, 3, 4) : exemples de travaux existants. 1 : Mac Aodha et al. (2018), 2 : Joachim and Lauga (1996), 3 : Weilgart and Whitehead (1997), 4 : Glotin et al. (2015).	164
5.6	Schéma récapitulatif des facteurs influençant la bioacoustique pour les différentes échelles d'études. La bioacoustique comportementale est transversale aux autres échelles.	165
5.7	Diagramme d'Ishikawa représentant les différentes causes et effets pour la mise en place d'une étude bioacoustique au cours du temps.	171
5.8	Carte prévisionnelle de la position des deux antennes pour l'été 2021 à Orcalab ainsi que leur zone d'écoute envisagée	172
5.9	Carte des positionnements des hydrophones pour le projet CARI'MAM. Les ancres indiquent les points précis définis avec le partenaire, les drapeaux la localisation approximative.	175

5.10 Résumé des notions de précision et rappel dans le cadre d'une détection de clics de Cachalots. Les transitoires détectés par notre modèle sont à l'intérieur du cercle (vrais positifs + faux positifs).	191
---	-----

Liste des tableaux

1.1	Exemples d'études où l'espèce a été identifiée automatiquement par l'acoustique pour différents taxons.	22
2.1	Récapitulatif des enregistrements réalisés avec BOMBYX. La session B ne sera pas utilisée dans ce manuscrit (pas d'accès aux enregistrements mais déjà traitée dans le rapport Glotin et al. (2017)).	49
2.2	Présentation des différents comportements et de leurs descriptions associées.	62
2.3	Tableau récapitulatif du matériel utilisé et des paramètres des enregistrements pour chaque cas d'étude.	65
3.1	Tableau représentant les différents jeux de données utilisés pour la classification aviaire. .	76
3.2	Tableau représentant les différents runs effectués pour la classification aviaire avec leur Taux de Bonne Classification (TBC).	77
3.3	Performances du modèle d'apprentissage pour la détection d'Orques sur le test.	80
3.4	Récapitulatif des différents entraînements réalisés pour la classification des clans.	89
4.1	Résumé des enregistrements pour toutes les sessions.	121
4.2	Résumé des 3 individus où nous avons pu isoler des vocalises. Le mâle A66 est apparu plusieurs fois éloigné de son groupe.	122
4.3	Résumé de nos résultats préliminaires concernant la bioacoustique individuelle des Orques de Colombie-Britannique en fonction de la bibliographie actuelle (en excluant les études avec des animaux captifs).	134
4.4	Tableau récapitulatif des prédictions du modèle en fonction des annotations de référence ('vérité') faites par T. Lengagne.	138
5.1	Résumé des différentes avancées pour tous les cas d'étude associés à chaque échelle d'analyse.	146
5.2	Récapitulatif des avantages/inconvénients des méthodes supervisées et non supervisées dans le cadre de la bioacoustique.	162
5.3	Tableau représentant les différents cas d'études avec l'association des variabilités acoustiques et des variabilités environnementales.	166

Liste des tableaux

5.4	Tableau bioacoustique de l'émission acoustique d'une Orque rejoignant son pod. 1H= 1 Hydrophone, BF = Basse Fréquence 1 à 100Hz ; MF = 100Hz à 10 kHz ; HF = 10 kHz à 500 kHz. En blanc les questions qui relèvent de la bioacoustique et en gris celles qui relèvent de l'écologie comportementale.	168
5.5	Tableau bioacoustique de l'émission acoustique d'un Cachalot en chasse. 1H= 1 hydrophone, BF = basse fréquence 1 à 100Hz ; MF = 100 Hz à 10 kHz ; HF = 10 kHz à 500 kHz. En gris les questions qui relèvent de l'acoustique et en orange celles qui relèvent de la bioacoustique comportementale.	169
5.6	Tableau générique représentant les questionnements en bioacoustique multiéchelle. 1H= 1 Hydrophone	170
5.7	Tableau des différentes perspectives associées à chaque cas d'étude. En rouge les cas d'étude où l'acquisition de nouveaux enregistrements est nécessaire	173

Résumé

L'objectif de cette thèse est d'apporter différentes contributions méthodologiques en bioacoustique pour l'étude de la faune. En effet, la bioacoustique est une science récente, pluridisciplinaire et très efficace pour étudier et classer un écosystème. Beaucoup d'études ont mis au point des procédés acoustiques pour étudier la faune à des échelles spécifiques, populationnelles, individuelles et comportementales.

Ce travail de thèse propose d'étudier différents cas d'études présents dans ces quatre échelles d'analyses. L'objectif de cette thèse est de mettre en place des outils depuis la pose du matériel d'acquisition jusqu'à l'analyse des données pour l'ensemble des échelles présentées, de les discuter et de les mettre en perspective. La bioacoustique spécifique est illustrée ici par la classification automatique d'Orques, de Cachalots et d'oiseaux. Pour la bioacoustique populationnelle, la classification acoustique de clans d'Orques est étudiée. Puis l'échelle d'analyse s'affine et étudie les émissions sonores individuelles. Pour cela 3 cas d'études sont utilisés : la localisation individuelle d'Orques, de Cachalots et d'oiseaux. La dernière échelle est appelée bioacoustique comportementale, elle a pour but de mettre en corrélation des comportements avec des émissions acoustiques. Pour cela, l'influence du trafic maritime sur les Dauphins tachetés pantropicaux et l'impact de stimuli chimiques chez la Baleine à bosse est étudié.

Nous avons volontairement fait le choix de sélectionner différentes espèces produisant des types de signaux bien différents (stationnaires vs transitoires) évoluant dans des milieux différents (marins vs terrestres) afin d'homogénéiser les méthodes d'analyses pour faciliter le développement de nouvelles études en bioacoustique. Chaque cas d'étude présente des résultats intéressants en terme de bioacoustique et d'écologie comportementale. Ces résultats sont comparés avec la bibliographie. Puis, les résultats de chaque cas d'étude permettent de valider les méthodes proposées dans cette thèse. Les apports méthodologiques de cette thèse sont synthétisés, comparés et discutés, notamment l'impact des signaux stationnaires et transitoires, des milieux (marin et terrestre) sur la mise en place des méthodes. Les méthodes supervisées et non supervisées sont mises en comparaison. Les méthodes proposées ont été testées et validées sur certains protocoles de données massives (plusieurs dizaines de Tera).

En conclusion, cette thèse montre que les méthodes supervisées (notamment le Deep Learning) étaient très bien adaptées pour la classification de signaux stationnaires en bioacoustique spécifique et populationnelle pour le milieu terrestre et marin. Puis les méthodes non supervisées (clustering et réduction de dimensionnalité) peuvent être utilisées dans le cadre des études en bioacoustique comportementale pour identifier les signaux d'intérêt. Enfin, la bioacoustique individuelle peut se traduire par des méthodes de localisation comme l'estimation du temps de délais d'arrivée inter-capteur, réalisable pour les signaux transitoires, et plus complexe pour les signaux stationnaires.

Mots-clés : Bioacoustique, Masse de données, Espèce, Population, Individu, Localisation, Apprentissage supervisé et non supervisé.

Abstract

The objective of this thesis is to make different methodological contributions in bioacoustics for the study of fauna. Bioacoustics is a recent multidisciplinary science and is very effective for studying and classifying an ecosystem. Many past studies have developed acoustical methods to analyze wildlife across (1) specific, (2) populational, (3) individual and (4) behavioral scales. The research presented in this thesis aims to study different case methods in the four scales of analysis listed above while also setting up tools from the setup of the acquisition material to the analysis of the data for all the aforementioned scales, and finally the discussion of the studies and putting them into perspective.

In this study, (1) specific bioacoustics were illustrated by the automatic classification of orcas, sperm whales, and birds. The acoustic classification of orca clans were studied for (2) population analysis. Then the scale was refined and (3) individual acoustic emissions were studied through three different case studies : the individual locations of orcas, sperm whales, and birds. The last scale evaluated was (4) behavioral bioacoustics which aimed to correlate behaviors with acoustic emissions. In order to correlate certain behaviors with acoustic emissions, the influence of maritime traffic on pantropical spotted dolphins and the impact of chemical stimuli in humpbacks were evaluated and recorded. We deliberately chose to select a diverse pool of species that would produce a variety of different signals (stationary vs. transient) and had evolved in different environments (marine vs. terrestrial). This allows us to standardize analysis methods in order to facilitate the development of new studies in bioacoustics.

Each case study showed interesting results in terms of bioacoustics and behavioral ecology. These results were compared with past studies which can be found in the bibliography. The results of each case study validated the methods proposed in this thesis. In particular, our study yielded excellent results in the evaluation of bird songs and is now a sound-recognition application available on any type of mobile phone, making it easy to identify bird species. The methodological contributions of this thesis, specifically the difference between stationary and transient signals and those of marine or terrestrial evolution, were synthesized, compared, and discussed. Supervised and unsupervised methods were also compared. These proposed methods have been tested and validated using massive data (several tens of Tera), which are unique.

In conclusion, this thesis shows that supervised methods, in particular Deep Learning, are very well suited for the classification of stationary signals in specific and population-based bioacoustics for the terrestrial and marine environment. We also derived that unsupervised methods such as clustering and reduction of dimensionality, can be used within the framework of behavioral bioacoustics to identify signals of interest. Finally, individual bioacoustics can be translated into localization methods such as estimating the inter-sensor delay time which is feasible for transient signals and more complex for stationary signals.

Keywords : Bioacoustics, Big data, Species, Populations, Individuals, localization, Unsupervised and Supervised Learning.

Publications et conférences

Articles scientifiques

- [1] M POUPARD, H SYMONDS, P SPONG et H GLOTIN. “Evidences of Intra-Group Orca Call Rate Modulation Using A Small-Aperture Four Hydrophone Array”. In : *in submission to Scientific report* (2020). URL : https://assets.researchsquare.com/files/rs-116685/v1_stamped.pdf.
- [2] B. BOUCHARD, JY. BARNAGAUD, M. POUPARD, H. GLOTIN, P. GAUFFIER, L. ORTIZ S. Torres, Thomas J., S. CAMPAGNA, M. RASMUSSEN et A. CÉLÉRIER. “Behavioural responses of humpback whales to food-related chemical stimuli”. In : *PloS one* 14.2 (2019), e0212515. URL : <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0212515>.
- [3] H. KOBAYASHI, HIROKI, Hiromi K, H GLOTIN, V ROGER, M POUPARD, D SHIMOTOKU, A FUJIWARA, K NAKAMURA, K SAITO et K SEZAKI. “A Real-Time Streaming and Detection System for Bio-Acoustic Ecological Studies After the Fukushima Accident”. In : *Multimedia Tools and Applications for Environmental & Biodiversity Informatics*. 2018, p. 53–66. URL : https://link.springer.com/chapter/10.1007/978-3-319-76445-0_4.

Communication dans des actes de conférence et Workshop

- [1] P BEST, M FERRARI, M POUPARD, S PARIS, R MARXER, H SYMONDS, P SPONG et H GLOTIN. “Deep Learning and Domain Transfer for Orca Vocalization Detection”. In : *International joint conference on neural networks*. 2020. URL : <https://hal.archives-ouvertes.fr/hal-02865300/document>.
- [2] M POUPARD, P BEST, J SCHLÜTER, H SYMONDS, P SPONG et H GLOTIN. *Large-scale unsupervised clustering of Orca vocalizations : a model for describing orca communication systems*. Rapp. tech. PeerJ Preprints, 2019. URL : <https://peerj.com/preprints/27979.pdf>.
- [3] M POUPARD, P BEST, J SCHLÜTER, JM PRÉVOT, Helena SYMONDS, Paul SPONG et Hervé GLOTIN. “Deep Learning for Ethoacoustics of Oreas on three years pentaphonie continuous recording at Orealab revealing tide, moon and diel effects”. In : *OCEANS 2019-Marseille*. IEEE. 2019, p. 1–7. URL : <https://hal.archives-ouvertes.fr/hal-02445426/document>.
- [4] M POUPARD, B de MONTGOLFIER et H GLOTIN. “Ethoacoustic by bayesian non parametric and stochastic neighbor embedding to forecast anthropic pressure on dolphins”. In : *OCEANS 2019-Marseille*. IEEE. 2019, p. 1–5. URL : <https://ieeexplore.ieee.org/abstract/document/8867126>.
- [5] M FERRARI, M POUPARD, P GIRAUDET, R MARXER, J-M PRÉVOT, T SORIANO et H GLOTIN. “Efficient artifacts filter by density-based clustering in long term 3D whale passive acoustic monitoring with five hydrophones fixed under an Autonomous Surface Vehicle”. In : *OCEANS 2019-Marseille*. IEEE. 2019, p. 1–7. URL : <https://ieeexplore.ieee.org/abstract/document/8867416>.

- [6] M. POUPARD, M. FERRARI, J. SCHLUTER, R. MARXER, P. GIRAUDET, V. BARCHAZS, V. GIÈS, G. PAVAN et H. GLOTIN. “Real-time passive acoustic 3D tracking of deep diving cetacean by small non-uniform mobile surface antenna”. In : *Accepted to IEEE/ICASSP*. 2019. URL : http://sabiiod.univ-tln.fr/publications/documents/SphyrnaI_Poupardetal_ICASSP2019.pdf.
- [7] A JOLY, H GOËAU, C BOTELLA, S KAHL, M POUPARD, M SERVAJEAN, H GLOTIN, P BONNET, W VELLINGA, R PLANQUÉ et al. “Lifeclef 2019 : Biodiversity identification and prediction challenges”. In : *European Conference on Information Retrieval*. Springer. 2019, p. 275–282. URL : https://link.springer.com/chapter/10.1007/978-3-030-15719-7_37.
- [8] M. POUPARD, M. FERRARI, J. SCHLUTER, P. ASTRUCH, B. SCHOHN, B. ROUANET, A. GOUJARD, A. LYONNET, P. GIRAUDET, V. BARCHASZ, V. GIES, P. BEST, D. DOMINICI, T. LENGAGNE, T. SORIANO et H. GLOTIN. “Passive acoustics to monitor flagship species near boat traffic in the Unesco world heritage natural reserve of Scandola”. In : *Input Academy : International Conference on Innovation in Urban and regional planning*. April 2019.
- [9] M. POUPARD, B. DE MONTGOLFIER, V. ROGER, D. LOHANI et H. GLOTIN. “EthoAcoustics : a model based on t-SNE & Clustering, applied on Pantropical spotted dolphin during Whale Watching”. In : *8th International DCLDE (Detection, Classification, Localization, and Density Estimation) Workshop*. 2018. URL : <https://hal.archives-ouvertes.fr/hal-01873336/document>.
- [10] H. GLOTIN et M. POUPARD. “Big data passive acoustic for Baikal Lake (ОЗЕРО БАЙКАЛ) soundscape and ecosystem observatory”. In : *Proceedings of the Russian-French Workshop in Big Data and Applications, Executive editor V. Kornilov; National Research University Higher School of Economics - Moscow, ISBN 978-5-7598-1808-3, 2018*. 2018. URL : https://www.hse.ru/data/2018/10/04/1157264869/Proceedings_Russian_French_Workshop.pdf.

Rapports de recherche

- [1] M. POUPARD, T. LENGAGNE et H. GLOTIN. *Première étude comparée des applications sur mobile (Iphone et Android) d'identification acoustique des espèces cibles d'oiseaux de France métropolitaine*. Rapport de recherche Université de Toulon, 2018.
- [2] M. POUPARD, T. LENGAGNE et H. GLOTIN. *Rapport d'avancement sur le modèle de reconnaissance précise sur une cinquantaine d'espèces avec estimation des conditions de terrain, CNRS UMR 7020*. Rapport de recherche Université de Toulon et Biosong, CNRS UMR 7020, 2018.
- [3] H. GLOTIN, G. PASCALE, J. RICARD, F. MALIGE, J. PATRIS, M. POUPARD, O. PHILIPPE et P. COSENTINO. *VAMOS - Visees Aériennes de Mammifères marins jointes aux Observations acoustiques Sous-marines de BOMBYX et ANTARES : nouveaux modèles en suivis et lois allométriques du *Physeter macrocephalus*, *Ziphius Cavirostris* et autres cétacés*. Université de Toulon, 2017.

Diffusion scientifique

- Documentaire ARTE “Les orques résidentes de Colombie Britannique”, diffusion Septembre 2020 [Lien de la bande annonce](#)
- Reportage Thalassa “De Marseille à Toulon” pour la mission Sphyrna Odyssey (diffusion Novembre 2020) [Lien de la mission](#), [Lien de la vidéo](#)
- Support de Allain Bougrain-Dubourg pour la diffusion prochaine de l'application Birdify pour la classification de chant d'oiseaux [Lien de la vidéo](#)
- Journal télévisé France 2 pour la mission Sphyrna Odyssey, diffusion février 2020 [Lien de la vidéo](#)
- Journal télévisé TF1 pour la mission Sphyrna Odyssey, diffusion printemps 2020
- Reportage Thalassa avec l'association Corseacare “Du Cap Corse à Bonifacio, les gardiens d'une nature sauvage” du 15/10/2018 : [Lien de la vidéo](#)

- Interview pour un Article de "Corsematin" avec l'association Corsecare : [Lien de l'article](#)
- Participation au tournage du documentaire decouverte "Méditerranée" de Jérôme Espla : [Lien de la bande annonce](#)
- Animation réalisée avec le CNAM de Laval pour promouvoir la mission Sphyrna 2018 : [Lien Youtube](#)
- Retour sur la mission Sphyrna 2018 (vidéo réalisée par l'Université de Toulon) : [Vidéo Youtube](#)