



HAL
open science

Méthodes d'identification de sources acoustiques paramétriques par mesures d'antennerie

Valentin Baron

► **To cite this version:**

Valentin Baron. Méthodes d'identification de sources acoustiques paramétriques par mesures d'antennerie. Acoustique [physics.class-ph]. Université de Lyon, 2020. Français. NNT : 2020LY-SEI121 . tel-03199353

HAL Id: tel-03199353

<https://theses.hal.science/tel-03199353v1>

Submitted on 15 Apr 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.



N°d'ordre NNT : 2020LYSEI121

THESE de DOCTORAT **DE L'UNIVERSITE DE LYON**
opérée au sein de
INSA Lyon

Ecole Doctorale N° 160
Electronique, Electrotechnique et Automatique

Spécialité/discipline de doctorat :
Traitement du signal

Soutenue publiquement le 14/12/2020, par :
Valentin Baron

Méthodes d'identification de sources acoustiques paramétriques par mesures d'antennerie

Devant le jury composé de :

Larzabal, Pascal	Professeur des Universités	ENS Cachan	Président du jury
Marcos, Sylvie	Directrice de Recherche	CentraleSupélec	Rapporteuse
Thomas, Jean-Hugh	Professeur des Universités	ENSIM	Rapporteur
Bavu, Eric	Maître de Conférences HDR	CNAM	Examineur
Antoni, Jérôme	Professeur des Universités	INSA Lyon	Examineur
Nicolas, Barbara	Directrice de Recherche	INSA Lyon	Directrice de thèse
Bouley, Simon	Docteur	MicrodB	Encadrant industriel
Mars, Jérôme	Professeur des Universités	Grenoble INP	Co-directeur de thèse

Département FEDORA – INSA Lyon - Ecoles Doctorales – Quinquennal 2016-2020

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	<u>CHIMIE DE LYON</u> http://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr INSA : R. GOURDON	M. Stéphane DANIELE Institut de recherches sur la catalyse et l'environnement de Lyon IRCELYON-UMR 5256 Équipe CDFA 2 Avenue Albert EINSTEIN 69 626 Villeurbanne CEDEX directeur@edchimie-lyon.fr
E.E.A.	<u>ÉLECTRONIQUE,</u> <u>ÉLECTROTECHNIQUE,</u> <u>AUTOMATIQUE</u> http://edeea.ec-lyon.fr Sec. : M.C. HAVGOUDOUKIAN ecole-doctorale.eea@ec-lyon.fr	M. Gérard SCORLETTI École Centrale de Lyon 36 Avenue Guy DE COLLONGUE 69 134 Écully Tél : 04.72.18.60.97 Fax 04.78.43.37.17 gerard.scorletti@ec-lyon.fr
E2M2	<u>ÉVOLUTION, ÉCOSYSTÈME,</u> <u>MICROBIOLOGIE, MODÉLISATION</u> http://e2m2.universite-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : H. CHARLES secretariat.e2m2@univ-lyon1.fr	M. Philippe NORMAND UMR 5557 Lab. d'Ecologie Microbienne Université Claude Bernard Lyon 1 Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69 622 Villeurbanne CEDEX philippe.normand@univ-lyon1.fr
EDISS	<u>INTERDISCIPLINAIRE</u> <u>SCIENCES-SANTÉ</u> http://www.ediss-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 INSA : M. LAGARDE secretariat.ediss@univ-lyon1.fr	Mme Sylvie RICARD-BLUM Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS) - UMR 5246 CNRS - Université Lyon 1 Bâtiment Curien - 3ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tel : +33(0)4 72 44 82 32 sylvie.ricard-blum@univ-lyon1.fr
INFOMATHS	<u>INFORMATIQUE ET</u> <u>MATHÉMATIQUES</u> http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	M. Hamamache KHEDDOUCI Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tel : 04.72.44.83.69 hamamache.kheddouci@univ-lyon1.fr
Matériaux	<u>MATÉRIAUX DE LYON</u> http://ed34.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIÈRE INSA de Lyon MATEIS - Bât. Saint-Exupéry 7 Avenue Jean CAPELLE 69 621 Villeurbanne CEDEX Tél : 04.72.43.71.70 Fax : 04.72.43.85.28 jean-yves.buffiere@insa-lyon.fr
MEGA	<u>MÉCANIQUE, ÉNERGÉTIQUE,</u> <u>GÉNIE CIVIL, ACOUSTIQUE</u> http://edmega.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bât. Direction mega@insa-lyon.fr	M. Jocelyn BONJOUR INSA de Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69 621 Villeurbanne CEDEX jocelyn.bonjour@insa-lyon.fr
ScSo	<u>ScSo*</u> http://ed483.univ-lyon2.fr Sec. : Véronique GUICHARD INSA : J.Y. TOUSSAINT Tél : 04.78.69.72.76 veronique.cervantes@univ-lyon2.fr	M. Christian MONTES Université Lyon 2 86 Rue Pasteur 69 365 Lyon CEDEX 07 christian.montes@univ-lyon2.fr

Remerciements

Après trois ans de travail, voilà la fin de l'aventure, et quelle aventure ! Une thèse, c'est pour beaucoup des études, pour d'autres un diplôme qui ne sert à rien de plus que les autres. Pour moi, et c'est bien personnel, la thèse est une formation professionnelle exigeante pour apprendre à faire de la recherche. Et quand on parle de formation, il faut forcément des formateurs pour la mener à bien. Ce sont eux que je veux remercier en premier.

Merci à Barbara pour ces trois ans durant lesquels tu as su me montrer la voie sans m'étouffer, voire même parfois en me laissant me perdre un peu. Sur l'instant on se dit : "mais pourquoi elle ne m'explique pas plus?", et avec le recul je me dis que tu as su comprendre ma façon de travailler et savoir me ré-aiguiller aux moments où il fallait. Merci à Arthur de m'avoir fait confiance et m'avoir lancé en stage puis au démarrage de la thèse, avec une motivation de tous les instants sur n'importe quel sujet. Cette dynamique un peu folle m'a parfois manqué sur la fin, mais c'est peut-être pas plus mal comme ça parce que la rédaction c'est calme ! Merci à Simon de m'avoir apporté ce que j'ai moins : une rigueur scientifique de haut niveau, que ce soit sur la manière d'aborder les problèmes mais aussi de les présenter. J'aurais certainement fait une bonne thèse sans toi, mais il m'aurait manqué une facette de ce que doit être un bon chercheur. Et enfin merci à Jérôme, qui parmi toutes ses activités a réussi à trouver un peu de temps pour me distiller de précieux conseils, toujours francs et sans filtres mais qui font bien comprendre comment valoriser son travail (et surtout "vendre mieux!!").

La thèse c'est aussi un instant unique dans la vie de chercheur, celui où son travail effectué sur trois ans est évalué dans les moindres détails par des experts du domaine. Je remercie donc Sylvie Marcos et Jean-Hugh Thomas pour leur relecture très attentive de mon manuscrit et les améliorations qui en ressortent, mais aussi Pascal Larzabal, Eric Bavu et Jérôme Antoni pour l'intérêt qu'ils ont montré à examiner mon travail et me poser de nombreuses questions. L'évaluation est une constante du travail de recherche, rien que pour vérifier que ce qu'on raconte est basé sur des références solides, et sans un jury de qualité qui pose des questions compliquées et intéressées, il est difficile de savoir jusqu'à quel point on maîtrise son sujet.

L'aventure de thèse, c'est aussi des rencontres qui font avancer, que ce soit scientifiquement, professionnellement ou personnellement, et sur ce point il y en a des personnes à remercier !

Les plus proches d'abord, mes copains docteurs ! Merci Maxime pour ces interminables discussions où à la fin on ne sait pas trop si on est plus avancé qu'au début... C'est important de prendre (perdre ?) son temps quand on cherche, sinon on ne peut pas vraiment créer. Merci Aneline d'être ce que tu es, c'est-à-dire un peu comme moi, c'est vraiment un plaisir de se retrouver dans quelqu'un et de pouvoir en discuter, on se sent moins seul ! Merci Emeline pour ta générosité et ta gentillesse, mais on ne devrait pas être trop étonnés, ce qui touche au cœur ça te connaît. Thanks to the "Bureau des bouffons" members : Pei, Bingqing and Vincent (in English now you're abroad), I hope we will be able to celebrate our PhD on these days. Merci aussi à tous les autres au labo Matthieu, Tom, Pierre, l'équipe US (maintenant ULTIM!) etc... J'ai passé d'excellentes années en votre compagnie même s'il faut bien avouer que les ultrasons ne sont pas mon domaine d'expertise scientifique... A défaut on a fait du triathlon !

Une mention spéciale pour une rencontre imprévue mais qui m'a permis de me poser beaucoup de questions sur moi-même : merci Pierrick pour ces heures de discussion sur la science et le rôle d'un chercheur là-dedans. J'espère qu'on pourra continuer les projets ensemble, j'ai encore envie de parler ! Merci Suzanne d'avoir permis cette rencontre et ces séances de répétitions pour le concours MT180s (Maxime, à bientôt pour la finale!).

A tous les collègues de MicrodB et Vibratec, merci de m'avoir accueilli, de croire en moi et à ce que j'apporte, et surtout de m'avoir permis de mener ma thèse sans aucune contrainte, ce qui n'est a priori pas le cas de toutes les Cifre. Merci Christine pour toutes ces discussions absolument pas scientifiques mais absolument nécessaires dans la vie, à Thibaut pour sa motivation scientifique débordante, à Lucille de me faire confiance, à Allan pour son expertise technique (que je n'aurais jamais) et ses questionnements sur le monde (en télétravail ça arrive moins les discussions sur les auteurs de l'Usine Nouvelle non ? :)), à Martin, Bruno et Olivier pour avoir enfilé une trifonction, aux autres sportifs du midi qui m'ont parfois accompagné dans la

boue du bois de Serre, et à tous ceux qui ont participé à mes répétitions pour le concours MT180s, on y croit encore ça va avoir lieu !

Je suis extrêmement chanceux d'avoir rencontré aussi beaucoup de monde au cours de mes différents projets : merci Matthieu d'avoir fait voler des drones pour que je puisse les écouter, Rania et Florent pour la source tête en haut ou tête en bas, Marie-Edith, Lucie et Bernard pour cette virée en bateau à Saint-Tropez, et à toutes les personnes que j'ai croisées pour Abysound ou pour mon projet sur les drones. Participer à des projets, c'est aussi comprendre que faire comprendre ce qu'on a dans le crâne à d'autres personnes, c'est pas toujours évident !

Pour terminer dans le personnel il y a aussi les copains et la famille. Merci aux présents de l'AG 2020 du CRV de m'avoir permis une dernière répétition en public ! Merci les Faidherbards, notamment de bien vouloir relire des trucs comme un manuscrit de thèse ! Merci à mon Papa de s'être aussi prêté au jeu et de toujours être curieux de tout, et au reste de la famille, qui s'agrandit, pour leurs félicitations qui me vont droit au cœur (même Dory y a participé !).

Et peut-on parler de famille pour toi qui te reconnaîtra ? Officiellement pas encore mais ça ne saurait tarder... Merci de me supporter, de m'épauler, et surtout de m'aimer.

Résumé

La caractérisation de sources acoustiques a pour but de décrire des émetteurs sonores à travers certains paramètres comme leur localisation dans l'espace, le niveau sonore qu'ils produisent ou encore leur identification grâce à leur signature acoustique. Dans cette thèse, l'objectif est d'obtenir certains de ces trois paramètres dans deux cas d'application concrets, pour des sources situées en champ lointain et grâce à des antennes acoustiques.

La première application concerne l'impact acoustique de la prospection minière sous-marine par grands fonds dans le cadre du projet FUI Alysound. Au sein du projet, la thèse cherche à caractériser les machines excavatrices placées sur les fonds marins en obtenant leur localisation et leur niveau sonore. Une première phase de design a abouti à la construction d'une antenne conique de 3 m. En s'appuyant ensuite sur les données issues de deux campagnes expérimentales menées en mer Méditerranée avec cette antenne, la méthode haute-résolution MUSIC localise précisément les sources mises en œuvre, qu'elles soient mobiles et à plus de 600 m de l'antenne, ou immergées par 700 m de fond. Leur niveau sonore est ensuite estimé par formation de voies, et les niveaux attendus sont retrouvés pour des signaux monochromatiques ou à large bande.

Dans la seconde application, une procédure complète pour la localisation et l'identification de drones est proposée pour la protection de sites sensibles. Elle combine traitement d'antenne et apprentissage statistique en s'articulant autour de trois étapes clés : la localisation, la focalisation et l'identification. La méthode MUSIC localise à nouveau les sources acoustiques présentes autour de l'antenne industrielle utilisée, puis la focalisation reconstruit le signal temporel de chacune, et un modèle SVM les identifie comme drone ou non. Les validations expérimentales, en intérieur comme en extérieur, constituent une contribution importante de ce travail de thèse. Les données acquises montrent entre autres que la procédure localise des drones à 3° près en extérieur, les détecte à 99 %, ou encore les identifie en présence d'une source perturbatrice plus puissante.

Abstract

Acoustic sources characterization aims to describe sound emitters through some parameters like their localization in space, the sound level they produce or their identification thanks to their acoustic signature. In this thesis, the objective is to obtain some of these parameters in two industrial application cases, for sources located in far-field and by the use of acoustic arrays.

The first application concerns deep-sea mining acoustic impact in the context of Alysound FUI project. Within it, the thesis searches to characterize the excavation machine located on the seabed by assessing their localization and their sound level. First, a design phase has led to the construction of a 3 m size acoustic array. Then, using data from two experimental campaigns conducted in the Mediterranean Sea with this array, the high-resolution method MUSIC accurately localizes the used acoustic sources, either mobile and more than 600 m away from the array, or immersed by 700 m depth. Their sound level is then estimated by beamforming, and the expected levels are verified for monochromatic or wideband signals.

In the second application, a complete procedure for the localization and the identification of drones is proposed to protect sensitive areas. It combines array processing and machine learning through three key steps : localization, focalization, and identification. MUSIC localizes again nearby acoustic sources around the industrial array used, then focalization reconstructs each temporal signal, and a SVM model identifies them as drone or not. Experimental validations, inside and outside, establish an important contribution of this thesis work. Acquired data show for instance that the procedure localizes drones with 3° accuracy outside, detects them at 99 %, or identifies them despite the presence of a more powerful source.

- J'espère que vous allez trouver ce que vous cherchez.
- Oui moi aussi. Vous ne sauriez pas ce que c'était par hasard ?

- *Dory et un poisson, Le monde de Dory*



Crédits : [Agathe Girard.com](http://AgatheGirard.com)

Table des matières

Table des matières	vii
Introduction	2
1 Ondes et sources acoustiques	5
1.1 Cadre acoustique de la thèse	6
1.2 Modèles de propagation	7
1.3 Sources acoustiques	9
Modèles théoriques de sources	9
Modélisation de source dans la thèse	12
1.4 Conclusion	14
2 Outils de traitement du signal pour la caractérisation de sources acoustiques	16
2.1 Localisation, quantification et identification de sources acoustiques	17
2.2 Traitement d'antenne	18
Formation de voies conventionnelle	24
Méthode MUSIC	27
Champ lointain en antennerie	30
Conclusion pour le traitement d'antenne	30
2.3 Apprentissage statistique	31
Classification binaire	33
Classification à une classe	37
2.4 Conclusion	39
3 Localisation et quantification de source en prospection minière sous-marine	41
3.1 Contexte	42
3.2 Conception du système d'écoute	44
3.3 Essais expérimentaux en mer	53
Configuration 1 : deux sources en petit fond	54
Configuration 2 : antenne en grand fond et source référencée en petit fond	54
Configuration 3 : antenne et source à large bande en grand fond	55
Références pour la localisation et la quantification	55
3.4 Validation du système d'écoute	56
3.5 Apport de la localisation à haute résolution	60
3.6 Validation en situation réelle par grand fond	63
3.7 Conclusion	67

4 Localisation et identification couplées de drones par acoustique	69
4.1 Contexte	70
4.2 Description de la chaine de traitement	76
Localisation	76
Focalisation	77
Identification	78
4.3 Bases de données expérimentales produites	81
Environnement intérieur	81
Environnement extérieur	82
4.4 Validation de la localisation et de l'identification pour des sources seules en intérieur	84
4.5 Etude des caractéristiques pour l'identification	88
Liens entre les signaux et les caractéristiques	88
Modification du jeu de caractéristiques	93
4.6 Limites de fonctionnement en localisation	94
4.7 Intérêt du couplage traitement d'antenne / apprentissage statistique pour l'identification	96
4.8 Limites de fonctionnement en focalisation en présence de plusieurs sources	101
4.9 Performances en extérieur	103
Sources présentes une à une	103
Drone en présence d'une source perturbatrice	107
Performances finales en fonction de la distance	111
4.10 Conclusion	112
Conclusion	115
ANNEXES	118
A Description du projet Abysound	119
B Caractéristiques pour l'identification de drone	120
C Liste des publications	121
Bibliographie	122

Table des figures

1.1	Schéma d'une réflexion	8
1.2	Illustration du champ lointain	10
1.3	Illustration de la compacité	11
1.4	Illustration de sources monopolaires et dipolaires	12
1.5	Schéma d'une source à multiples composantes	13
1.6	Illustration d'une source composée de quatre couples de monopôles	14
2.1	Schéma de la grille de scrutation	20
2.2	Schéma illustrant le principe de la formation de voies	24
2.3	Comparaison des méthodes d'imagerie	29
2.4	Configuration pour la simulation de signaux temporels	30
2.5	Comparaison de l'aspect sphérique des ondes	31
2.6	Modèle SVM : cas séparable	34
2.7	Modèle SVM : cas non-séparable	36
2.8	Modèle SVM : cas non-linéaire	37
2.9	Modèle SVM : cas à une classe	38
3.1	Illustration d'un site minier sous-marin	42
3.2	Spectres de bruits de forage et de dragage	43
3.3	Gabarit du spectre d'une machine excavatrice	44
3.4	Géométries axisymétriques comparées lors de la conception de l'antenne	49
3.5	Schéma de la configuration de simulation	49
3.6	L_{-3dB} moyenne en fonction du NMLS moyen pour les différentes géométries	50
3.8	Schéma de l'antenne conique optimale	51
3.7	L_{-3dB} moyenne en fonction du NMLS moyen pour les antennes coniques	51
3.9	Validation des performances de l'antenne	52
3.10	Photographie de l'antenne construite	53
3.11	Schéma de la configuration 1	54
3.12	Schéma de la configuration 2	54
3.13	Schéma de la configuration 3	55
3.14	Exemple de décroissance des valeurs propres	57
3.15	Résultats de localisation : configuration 2	58
3.16	Carte de MUSIC obtenue pour une propagation avec un modèle avec une source image	59
3.17	Résultats de quantification : configuration 2	60
3.18	Apport de la haute-résolution : sources proches	61
3.19	Apport de la haute-résolution : sources à niveau équivalent	62
3.20	Apport de la haute-résolution : sources à niveau différent	63
3.21	Exemples de cartes de MUSIC en grand fond	65
3.22	Résultats de localisation : configuration 3	65
3.23	Résultats de quantification : configuration 3	66
3.24	DSP d'un hydrophone unique	66
4.1	Chaîne de traitement pour les drones	76
4.2	Illustration du découpage des signaux	76
4.3	Exemple de carte de MUSIC angulaire	77
4.4	Illustration de l'étape de focalisation	78

4.5	Illustration de l'étape d'identification	79
4.6	Photographie de l'antenne utilisée	81
4.7	Schéma de la trajectoire linéaire	81
4.8	Schéma de la trajectoire 5 cercles	81
4.9	Schéma de la trajectoire 3 lignes droites	82
4.10	Photographie des trois drones utilisés	82
4.11	Schéma de la configuration extérieure	83
4.12	Photographie du drone en extérieur	84
4.13	Localisation du drone avec les données intérieures	85
4.14	Erreur de localisation pour les vols de cinq cercles en intérieur	86
4.15	Résultats d'identification en intérieur	87
4.16	Précision du modèle SVM à quatre classes en fonction du nombre d'observations en apprentissage	89
4.17	Signaux d'entrée de l'étape d'identification	90
4.18	Exemples de caractéristiques séparant bien les sons	91
4.19	Exemples de caractéristiques séparant mal les sons	92
4.20	Schéma de la position du haut-parleur seul dans la salle	94
4.21	Comparaison des angles d'élévation pour le haut-parleur seul	95
4.22	Comparaison en simulation de l'erreur d'estimation de l'angle d'élévation pour différentes valeurs	96
4.23	Illustration de la procédure de labélisation	97
4.24	Résultats de localisation en présence de deux sources	98
4.25	Comparaison du signal de drone et du signal de sinus à 1 kHz en intérieur	99
4.26	Résultats d'identification à deux sources en intérieur	100
4.27	Exemples de caractéristiques modifiées en présence de bruit	102
4.28	Comparaison des signaux d'entrée pour les exemples de séquences de drone focalisées	103
4.29	Estimation des angles φ et θ pour les vols extérieurs	104
4.30	Résultats d'identification en extérieur pour des sources seules	105
4.31	Comparaison des densités spectrales de puissance du drone et des sons joués en extérieur	107
4.32	Estimation des angles φ et θ pour les vols extérieurs avec deux sources	108
4.33	Exemple de décroissance des valeurs propres de la matrice interspectrale en extérieur	109
4.34	Résultats d'identification à deux sources en extérieur	110
4.35	Performances en identification à deux sources en extérieur en fonction de la distance drone / antenne	111

Liste des tableaux

3.1	Variation du NMLS pour les antennes	50
4.1	Exemple de matrice de confusion	80
4.2	Données de drone intérieures	83
4.3	Données de sons intérieures	83
4.4	Données extérieures	84
4.5	Matrice de confusion pour le modèle SVM	87
4.6	Matrice de confusion pour le modèle SVM-1C	87
4.7	Matrice de confusion pour l'identification à quatre classes	89
4.8	Comparaison des résultats de prédiction du SVM et du SVM-1C sur des sons non utilisés en apprentissage	106
4.9	Résultats de localisation en extérieur avec deux sources	109
B.1	Liste des caractéristiques	120

Introduction

Que ce soit pour lutter contre une forme de pollution encore largement sous-étudiée, la pollution sonore, ou pour aller plus loin que le cadre classique de l'imagerie acoustique en apportant de l'information supplémentaire sur les sources sonores imagées, la description des paysages sonores se veut de plus en plus précise. En fonction de l'environnement acoustique étudié, cette description s'intéresse à des grandeurs différentes : l'intensité du bruit ambiant dans des zones protégées, le temps de réverbération en acoustique des salles, la détermination des sources acoustiques nuisibles pour le conducteur en soufflerie automobile, ou encore la directivité acoustique d'un moteur avec des mesures sur banc d'essai.

Parmi toutes les informations qui peuvent être recherchées, cette thèse s'intéresse spécifiquement à la caractérisation de sources acoustiques, selon des paramètres définis à chaque cas d'application. Une source acoustique peut être définie comme un modèle mathématique décrivant un objet émettant des signaux acoustiques. Toutes sortes de paramètres peuvent alors lui être attribués : sa taille, sa localisation, son niveau sonore, sa directivité ou encore son type selon des catégories prédéfinies. Pour des objets aux géométries complexes, mobiles dans l'espace, et relativement loin de l'endroit depuis lequel ils sont observés, il est possible d'établir un modèle de source acoustique adapté et mathématiquement pratique à manipuler : le monopôle acoustique. Une fois ce choix de modélisation effectué, la directivité, supposée omnidirectionnelle, et la taille de la source, supposée nulle car la source est réduite à un point, ne sont plus des paramètres d'intérêt.

Une fois l'objet d'étude défini, il faut néanmoins des moyens de mesures pour le caractériser. Dans ce travail, l'imagerie acoustique est utilisée, dont le but est de produire une image des signaux acoustiques émis par les sources présentes, grâce à un ensemble de capteurs acoustiques regroupés sous la forme d'une antenne.

Dans ce cadre, l'objectif de la thèse est alors de caractériser par antennerie acoustique les paramètres d'intérêt de sources supposées monopolaires dans deux champs d'application distincts : la prospection minière dans les grands fonds marins, et la détection de drones dans l'espace aérien.

Dans le premier cas, le but du projet Abysound, dans lequel ces travaux s'insèrent, est d'étudier le bruit généré par les machines excavatrices qui creusent le plancher océanique à la recherche de minerais. Afin d'être capable de prédire leur impact sur la vie des mammifères marins se trouvant aux alentours, les machines doivent être localisées et leur niveau sonore quantifié, pour pouvoir alimenter ensuite des modèles de propagation en acoustique sous-marine. Cette thèse répond alors au double enjeu de la conception d'un tel système alliant mesures, localisation et quantification, mais aussi de sa validation en eaux profondes pour démontrer ses capacités.

Bien éloignée de ces considérations environnementales, la seconde application cherche à protéger les infrastructures ou les personnes, qu'elles soient civiles ou militaires, contre des attaques de drones malveillants. Un enjeu est alors commun avec l'application précédente, celui de localisation, qui est crucial pour envisager d'éventuelles mesures de mitigation. Le second enjeu concerne l'identification de la menace selon sa signature acoustique afin d'affirmer que la cible localisée est dangereuse, et ainsi éviter la génération de fausses alarmes intempestives.

Dans la suite de cette thèse, des solutions méthodologiques pour caractériser les sources acoustiques à décrire dans ces deux applications sont proposées puis validées.

Le Chapitre 1 donne ainsi les définitions des grandeurs acoustiques nécessaires à la bonne compréhension du manuscrit, avant de justifier en détail le choix d'un modèle monopolaire pour décrire l'ensemble des sources acoustiques au cours des applications.

Une synthèse de la littérature scientifique de deux domaines du traitement du signal employés dans ce travail de thèse est menée dans le Chapitre 2 : le traitement d'antenne et l'apprentissage statistique. Les particularités des applications visées permettent alors de dégager de cette synthèse des méthodes qui leur sont adaptées. La méthode haute-résolution *Multiple Signal Classification* (MUSIC) est ainsi sélectionnée pour la localisation dans toutes les configurations. La formation de voies est utilisée lors de la conception de l'antenne dans le cadre de l'application en grand fond, et répond aussi à son besoin de quantification. Enfin l'identification des drones est quant à elle assurée par les modèles *Support Vector Machines* (SVM) et SVM à une classe (SVM-1C) issus de la littérature en apprentissage statistique.

A la suite de ces choix méthodologiques, les deux cas d'applications sont analysés en profondeur et représentent le cœur de ce travail de thèse. Grâce aux nombreuses données expérimentales acquises, la robustesse des méthodes proposées est ainsi démontrée.

Le Chapitre 3 démarre avec la conception du système d'écoute capable de répondre aux enjeux technologiques et sécuritaires liés à cet environnement de mesure très particulier. Après construction d'un prototype de l'antenne acoustique conçue, deux campagnes expérimentales ont été effectuées en avril 2018 et mai 2019 en mer Méditerranée pour valider ses performances. Diverses configurations ont ainsi pu être testées : deux sources à faible profondeur (104 m), une source mobile à faible immersion (80 m) avec l'antenne en grande profondeur (660 m), et enfin la source et l'antenne en grande profondeur (680 et 660 m). A travers ces trois configurations, la capacité qu'a l'antenne conçue à localiser précisément et quantifier fidèlement les sources présentes est démontrée, tout en illustrant l'apport de la méthode haute-résolution pour la localisation.

Le Chapitre 4 montre quant à lui l'intérêt de combiner différents domaines du traitement du signal pour établir une localisation, mais surtout une identification précise de drone. Après une contextualisation à la fois industrielle et académique, une chaîne de traitement en trois étapes est proposée : localisation, focalisation, et identification. Puis les données créées durant la totalité des trois ans de thèse sont analysées

afin d'établir les capacités et les limites de la procédure, qu'elles soient industrielles ou académiques. Les performances en intérieur sont ainsi établies, puis la précision en localisation et en identification, grâce à des configurations expérimentales spécialement dédiées, est étudiée. Une dernière partie détaille les performances atteintes en extérieur, pour un drone seul ou pour un drone en présence d'une source fortement perturbatrice.

Clés pour lire cette thèse : la volonté affichée de cette thèse, qui est régie par une CIFRE (Convention Industrielle de Formation par la Recherche), est d'essayer d'apporter au lecteur plusieurs niveaux de lecture, tout en gardant un ensemble concis. L'argumentation qui y est tenue est condensée, appuyée par des considérations pratiques pour justifier certains choix. Afin d'aider le lecteur à appréhender certains points à retenir, deux types de boîtes ont été mises en place : les boîtes bleues *informations importantes* (comme ci-dessous) pour mettre en valeur les informations importantes, et les boîtes *résultats* (comme ci-contre) pour aider à la navigation au sein des nombreux résultats expérimentaux.

Exemple de boîte informations importantes

Dans ces boîtes sont consignées les informations importantes comme par exemple :

- ▶ une comparaison de méthodes
- ▶ des définitions de termes spécifiques
- ▶ un rappel sur des hypothèses importantes

Bonne lecture.

Exemple de boîte résultats

Dans ces boîtes sont consignées les résultats quantitatifs pour par exemple :

- ▶ les mettre en avant
- ▶ y revenir rapidement

Ondes et sources acoustiques

1

Ce chapitre présente les notions d'acoustique nécessaires à la bonne compréhension du manuscrit. Il détaille les hypothèses qui établissent les modèles de source et de propagation acoustique qui doivent être mis en place en fonction des applications expérimentales visées. Pour ce faire les équations générales de l'acoustique nécessaires pour établir ces modèles sont présentées, avant d'être simplifiées grâce à des arguments liés à la nature des sources acoustiques à modéliser dans cette thèse.

Table des matières

1.1	Cadre acoustique de la thèse	6
1.2	Modèles de propagation	7
1.3	Sources acoustiques	9
	Modèles théoriques de sources	9
	Modélisation de source dans la thèse	12
1.4	Conclusion	14

Deux livres ont principalement été utilisés pour rédiger ce chapitre [Mor+87] (en anglais) et [Bru98] (en français).

1.1 Cadre acoustique de la thèse

Les équations de l'acoustique consistent à décrire les mouvements de particules d'un fluide autour de leur position d'équilibre. Une onde acoustique qui se propage dans une direction donnée fait ainsi osciller les particules de fluide qu'elle rencontre de proche en proche, sans déplacement final de la matière¹. Lorsque le mouvement des particules se fait dans la même direction que le mouvement de l'onde, comme c'est le cas pour les ondes acoustiques, celles-ci sont dites longitudinales². Tout l'enjeu des équations de l'acoustique est alors de décrire les grandeurs physiques mises en jeu dans ces mouvements d'onde selon des conditions imposées par le milieu fluide considéré : absence ou présence d'obstacle, température variable ou milieu fluide en mouvement par exemple.

Les grandeurs physiques qui caractérisent un milieu fluide sont³ : sa densité ρ , sa pression p , la vitesse de ses particules \mathbf{u} , sa température T_{emp} , son entropie s . Dans de nombreux cas traités par le traitement du signal, seule la pression est utilisée. C'est la seule grandeur physique mesurée par les microphones ou les hydrophones, et donc à disposition pour la suite des traitements lorsque ces capteurs sont utilisés. Cependant, une variation locale de la pression implique la variation des autres grandeurs physiques du milieu fluide. En exploitant les liens entre ces grandeurs, et des hypothèses sur les propriétés du fluide et sur l'amplitude des ondes considérées, l'équation des ondes peut être dérivée. Ses solutions décrivent la variation de pression dans le milieu fluide en fonction du temps et de l'espace, en accord avec les hypothèses simplificatrices décidées, dont la liste est donnée dans l'encart ci-dessous⁴.

Hypothèses simplificatrices dans le cadre de la thèse

- ▶ **Milieu uniforme** : implique qu'en l'absence d'onde, la densité, la pression et la température sont uniformes dans le milieu.
- ▶ **Milieu au repos** : implique que dans tout le milieu, et en l'absence d'onde, les particules sont immobiles.
- ▶ **Milieu non visqueux et de conductivité thermique nulle** : hypothèses qui permettent d'assimiler le mouvement acoustique comme dû uniquement à une énergie mécanique.
- ▶ **Cadre de l'acoustique linéaire** : cela signifie que seules les fluctuations de faibles amplitudes autour des positions d'équilibre des grandeurs physiques sont considérées. Techniquement, cette hypothèse permet de ne considérer que les perturbations du premier ordre dans les développements de Taylor des équations.

Sous ces hypothèses, les équations de conservation de la masse et de la quantité de mouvement produisent deux relations entre la densité, dont la valeur moyenne dans le milieu fluide est donnée par ρ_0 , la vitesse des particules et la pression⁵ :

$$\rho_0 \frac{\partial \mathbf{u}}{\partial t} = -\nabla \tilde{p} \quad \text{et} \quad \frac{\partial \tilde{p}}{\partial t} = -\rho_0 \operatorname{div} \mathbf{u}. \quad (1.1)$$

1: Un exemple très visuel d'une onde qui avance sans que la matière ne la fasse consiste à agiter une corde de bas en haut : bien qu'une onde la traverse, la corde, elle, n'avance pas.

2: Par opposition à notre exemple précédent, où si l'on considère la corde comme une succession de points, leur mouvement est perpendiculaire à l'onde qui les traverse : l'onde est transversale.

3: Les grandeurs sont données en un point du milieu, l'ensemble des valeurs prises en tous les points du milieu étant les **champs** de ces grandeurs.

4: Pour plus de détails sur ces hypothèses et sur la suite de cette partie, voir [Mor+87], Chapitre 6.2.

5: Le signal de pression dans le domaine temporel est indiqué avec un tilde \tilde{p} , alors que dans le domaine fréquentiel il est indiqué sans ce tilde p , car c'est celui-ci qui est utilisé dans toute la suite de la thèse.

La première relation traduit qu'un gradient de pression provoque une accélération du fluide⁶, tandis que la seconde montre qu'un gradient de vitesse produit une compression du fluide⁷, ce qui montre bien l'interconnexion entre ces différentes grandeurs physiques.

Une dernière relation entre la pression et la densité est tirée du caractère adiabatique de la propagation acoustique, et fait apparaître la célérité des ondes acoustiques c_0 : $\tilde{p} = c_0^2 \rho$. En combinant ces relations, l'équation homogène des ondes sur la pression est obtenue :

$$\Delta \tilde{p} - \frac{1}{c_0^2} \frac{\partial^2 \tilde{p}}{\partial t^2} = 0. \quad (1.2)$$

Les solutions de cette équation homogène ne décrivent que la propagation des ondes qui intervient dans le milieu fluide, à condition que les hypothèses mentionnées sur celui-ci soient respectées. Lorsqu'une source acoustique est présente, un second membre qui décrit la contribution de celle-ci doit être ajouté donnant lieu à une équation dite inhomogène.

Dans le cadre de la thèse, les signaux réels sont étudiés sur des durées au cours desquelles ils peuvent être considérés comme stationnaires et ergodiques. Sous cette hypothèse, la résolution de l'équation des ondes peut être entreprise dans le domaine fréquentiel par passage par transformée de Fourier, en considérant des solutions harmoniques : $\tilde{p}(t) = \tilde{A} \cos(\omega t + \varphi) = \text{Re}\{\tilde{A} e^{i\varphi} e^{i\omega t}\} = \text{Re}\{p e^{i\omega t}\}$, avec $\omega = 2\pi f$ la pulsation de l'harmonique et φ sa phase. L'équation des ondes homogène peut alors être simplifiée, avec cette forme particulière pour \tilde{p} pour une pulsation donnée, en l'équation de Helmholtz :

$$\Delta p + k^2 p = 0, \quad (1.3)$$

avec $k = \omega/c_0$ le nombre d'onde et p l'amplitude complexe des ondes harmoniques. Sa solution décrit la propagation d'une onde à une fréquence donnée dans le milieu. Pour gérer le cas des signaux à large bande, la résolution est alors faite pour toutes les fréquences à étudier de manière indépendante. La partie suivante détermine les solutions de cette équation en fonction du repère de coordonnées choisi pour être adapté à l'application visée, ce qui donne lieu à différents modèles de propagation.

1.2 Modèles de propagation

Les modèles de propagation acoustique ont pour but de décrire le champ de pression issu du déplacement des ondes dans l'ensemble du milieu fluide. Ils se réfèrent donc à un système de coordonnées défini, dans un environnement lui aussi défini. Une grande variété de modèles est alors possible du cas le plus simple, un milieu infini sans aucun obstacle, aux cas les plus complexes, une pièce réverbérante avec des obstacles et un écoulement d'air par exemple. Dans les applications expérimentales de cette thèse, seuls deux types de modèle sont étudiés : le champ libre et la réflexion sur une interface unique. Alors que le premier est adapté aux cas où il n'y a pas d'obstacle dans un milieu infini

6: La particule va accélérer dans la direction où la pression est plus faible pour aller y retrouver un état plus stable.

7: Si la vitesse d'une particule est plus forte que celle de sa voisine, cette dernière est plus compressée.

mais aussi aux cas où les réflexions sur une interface sont faibles en comparaison avec le trajet direct, le second permet de prendre en compte les cas où ces réflexions ne sont plus négligeables.

Champ libre : pour ce cas, un système de coordonnées sphérique est particulièrement adapté pour décrire la propagation tri-dimensionnelle d'un signal émis depuis, ou se propageant vers, un point de l'espace \mathbf{x}_0 donné. Dans le cadre harmonique étudié, la pression acoustique ne dépend alors que de la distance entre ce point \mathbf{x}_0 et un point de l'espace considéré \mathbf{x} ⁸ :

$$p(\mathbf{x}) = A_d \frac{e^{-ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|} + A_c \frac{e^{ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|}, \quad (1.4)$$

avec A_d qui représente l'amplitude des ondes dites divergentes⁹, et A_c celle des ondes convergentes¹⁰. Les ondes se déplacent alors sur une sphère de surface $4\pi\|\mathbf{x}-\mathbf{x}_0\|^2$, avec une amplitude qui évolue en $1/\|\mathbf{x}-\mathbf{x}_0\|$ pour respecter le principe de conservation de l'énergie : les ondes sphériques voient leur amplitude s'atténuer avec la distance¹¹.

Ondes planes : en s'écartant de \mathbf{x}_0 , les fronts d'ondes peuvent être assimilés localement à des plans, donnant lieu à une description de la propagation acoustique par des ondes planes. Cette approximation est valable localement à partir d'une distance grande devant la longueur d'onde considérée, mais en étudiant la propagation vers une zone étendue de l'espace, cette distance doit être agrandie comme discuté dans la Partie 2.2, spécifiquement pour le cas des antennes acoustiques. Le modèle en ondes planes est adapté aux problèmes qui requièrent une description angulaire du champ de pression acoustique autour d'un point de l'espace. La direction de propagation des ondes planes est décrite par le vecteur de norme unitaire orthogonal aux fronts d'ondes $\boldsymbol{\alpha}$. La résolution de l'équation de Helmholtz se fait dans un repère de coordonnées cartésiennes pour des vecteurs de positions \mathbf{x} , et ses solutions sont de la forme¹² :

$$p(\mathbf{x}) = A_d e^{-ik\boldsymbol{\alpha}\cdot\mathbf{x}} + A_c e^{ik\boldsymbol{\alpha}\cdot\mathbf{x}}. \quad (1.5)$$

L'opérateur \cdot représente le produit scalaire entre deux vecteurs de l'espace. Contrairement à la propagation sphérique, les amplitudes des ondes planes ne varient pas avec la distance entre \mathbf{x}_0 et \mathbf{x} .

Réflexion sur une interface : pour prendre en compte la présence d'une interface dans le cas d'une propagation dans un milieu fini¹³, une réflexion peut être prise en compte dans les modèles précédents. Le principe est schématisé dans la Figure 1.1, et fait intervenir, dans le cas d'une propagation sphérique, la distance entre le symétrique du point du milieu considéré par rapport à l'interface, appelé point image \mathbf{x}_i , et le point \mathbf{x}_0 :

$$p(\mathbf{x}) = A_d \left(\frac{e^{-ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|} + R \frac{e^{-ik\|\mathbf{x}_i-\mathbf{x}_0\|}}{\|\mathbf{x}_i-\mathbf{x}_0\|} \right) + A_c \left(\frac{e^{ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|} + R \frac{e^{ik\|\mathbf{x}_i-\mathbf{x}_0\|}}{\|\mathbf{x}_i-\mathbf{x}_0\|} \right). \quad (1.6)$$

8: Voir [Mor+87] chapitre 7.1 pour plus de détails.

9: Les ondes qui se propagent depuis le point \mathbf{x}_0 .

10: Inversement celles qui se propagent vers \mathbf{x}_0 .

11: Il faut bien noter que cette atténuation est purement géométrique, et n'est en aucun cas due à la viscosité du milieu.

12: Voir [Mor+87] chapitre 6.3 pour plus de détails.

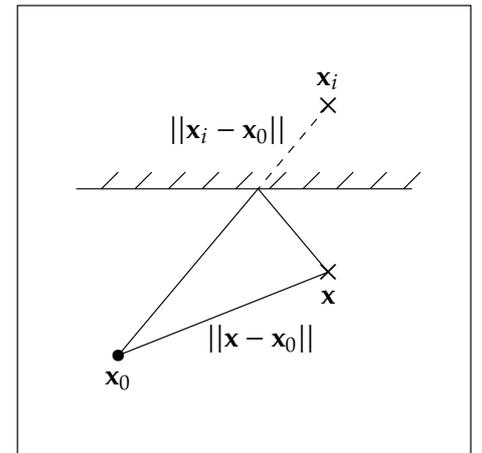


FIGURE 1.1 – Schéma illustrant une propagation entre un point \mathbf{x} de l'espace et un autre point \mathbf{x}_0 en présence d'une réflexion sur une interface. Le trajet réfléchi peut alors être modélisé par une source image \mathbf{x}_i .

13: Par exemple une propagation dans l'eau avec un fond solide et/ou de l'air au-dessus de la surface.

Le coefficient de réflexion R prend des valeurs dont le module varie entre 0 et 1 en fonction de l'intensité de la réflexion qui a lieu sur l'interface¹⁴ et plus sa valeur est proche de 1, plus la présence d'interférences constructives et destructives est importante dans le champ de pression solution.

En conclusion, le choix d'un modèle de propagation se base sur la géométrie du problème à résoudre : sans obstacle ou en présence de réflexions mineures le champ libre est à privilégier, alors qu'avec une interface marquée entre deux milieux de densités différentes un modèle qui prend en compte les réflexions est plus adapté. Pour des cas d'applications concrets, des conditions supplémentaires simplifient les solutions mathématiques déduites de l'équation de Helmholtz. Ainsi, dans un milieu infini et pour un signal acoustique créé en un point unique de l'espace x_0 , aucune onde convergente n'est à prendre en compte vers ce point, supprimant de fait les seconds termes des Equation 1.4, Equation 1.5 et Equation 1.6. Mais pour terminer la description d'un champ de pression acoustique, le mécanisme de création du signal acoustique mentionné doit être expliqué. Cette notion nécessite un point de départ, apporté par le concept de source acoustique.

14: Plus les milieux séparés par l'interface voient leur densité différer, plus le module du coefficient de réflexion augmente.

1.3 Sources acoustiques

Modèles théoriques de sources

Une source acoustique peut être définie comme un modèle mathématique pour décrire les effets produits par un objet ou une combinaison d'objets, un émetteur sonore, qui génère des ondes acoustiques dans le milieu dans lequel il est plongé. Cette génération provient de la transformation d'un phénomène oscillatoire depuis l'émetteur, comme les vibrations lors de l'impact d'une pelle sur le sol ou les turbulences créées en bordure d'une pale en rotation, en ondes acoustiques dans le milieu qui l'entoure selon le phénomène physique appelé rayonnement acoustique. C'est alors l'équation des ondes inhomogènes qui régit le problème acoustique à résoudre, avec la présence d'un second membre non nul.

Les ondes acoustiques peuvent provenir d'émetteurs très variés, du piston dans une cavité jusqu'au bruit de jet à la sortie d'un moteur d'avion. Une façon de décrire ces émetteurs consiste à considérer un ensemble de sources élémentaires disposées sur leur peau¹⁵, elles-mêmes modélisées par une combinaison de monopôles acoustiques. Un monopôle n'existe pas de manière physique : il résulte d'un passage à la limite des équations de rayonnement obtenues dans le cas d'une sphère pulsante. Une sphère de rayon a , dont la surface vibre avec une vitesse donnée selon la direction radiale à sa surface, rayonne autour d'elle un champ acoustique. L'équation de Helmholtz inhomogène associée à ce problème est alors de la forme¹⁶ :

15: Une surface englobante qui est souvent confondue avec la géométrie de l'émetteur.

16: Les détails mathématiques pour aboutir à l'équation sont donnés dans le chapitre 7.1 de [Mor+87].

$$\Delta p + k^2 p = -\rho_0 i \omega q, \quad (1.7)$$

avec q l'amplitude complexe d'un débit source défini par l'approche de sphère pulsante, qui représente la valeur instantanée du flux d'air ex-

pulsé par la sphère, et qui permet d'aboutir à la représentation théorique des monopôles acoustiques : à savoir un débit donné pour une sphère de rayon $a = 0$ ¹⁷. Le passage à la limite de la solution de l'équation de Helmholtz inhomogène pour a qui tend vers 0 permet de déterminer l'amplitude complexe du champ de pression généré par un monopôle seul placé en un point \mathbf{x}_0 de l'espace, en tout point du milieu par¹⁸ :

$$p(\mathbf{x}) = \frac{\rho i \omega q(\mathbf{x}_0)}{4\pi} \frac{e^{-ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|} = A(\mathbf{x}_0) \frac{e^{-ik\|\mathbf{x}-\mathbf{x}_0\|}}{\|\mathbf{x}-\mathbf{x}_0\|}. \quad (1.8)$$

L'amplitude complexe d'un monopôle situé en \mathbf{x}_0 est ainsi donnée par $A(\mathbf{x}_0) = \rho i \omega q(\mathbf{x}_0)/4\pi$, dont le module représente par construction l'amplitude du signal de pression généré par la source à 1 m. Ce niveau de pression à 1 m est couramment utilisé pour caractériser l'impact d'une source acoustique sur son environnement : sa quantification¹⁹. Il est exprimé en décibel, défini par dix fois la valeur logarithmique en base 10 du rapport entre la valeur efficace au carré de la pression acoustique générée à 1 m par la source, et une pression de référence dépendante du milieu fluide dans laquelle la source est placée²⁰ :

$$L_p = 10 \log_{10} \left(\frac{A_{eff}^2}{p_{ref}^2} \right). \quad (1.9)$$

Pour aller plus loin que le monopôle acoustique, c'est-à-dire un débit source d'entrée localisé dans l'espace en une position unique \mathbf{x}_0 qui génère des fronts d'ondes purement sphériques, des distributions de sources continues dans l'espace peuvent être modélisées en intégrant des distributions de monopôles dans une région donnée D . Ainsi le débit généré en un point de l'espace \mathbf{x} sera égal à l'intégrale, pour toutes les positions \mathbf{x}_d au sein de la région source D , des débits élémentaires relevés en cette position $q(\mathbf{x}_d)$:

$$q(\mathbf{x}) = \int_D q(\mathbf{x}_d) \delta(\mathbf{x} - \mathbf{x}_d) d\mathbf{x}_d. \quad (1.10)$$

Les solutions de l'équation d'Helmholtz inhomogène sont alors modifiées avec une amplitude de source $A(\mathbf{x}_d)$ qui devient fonction de la position dans la source distribuée :

$$p(\mathbf{x}) = \int_D A(\mathbf{x}_d) \frac{e^{-ik\|\mathbf{x}-\mathbf{x}_d\|}}{\|\mathbf{x}-\mathbf{x}_d\|} d\mathbf{x}_d. \quad (1.11)$$

Les distributions de monopôles apportent une flexibilité dans la modélisation des sources acoustiques qui peut alors être plus précise. Mais cette précision est acquise au prix d'un réglage complexe de la distribution de monopôles dans la région d'intégration D , pour faire correspondre le modèle au mieux à l'émetteur étudié. Sous certaines hypothèses, le gain en précision de cette distribution optimisée devient même négligeable et une forme plus générale de distribution de source peut être dérivée pour aboutir à une modélisation aussi précise de cet émetteur.

17: C'est dans cette approximation que réside le fait qu'un monopôle acoustique n'est qu'une représentation théorique : ce n'est pas possible de produire de la puissance acoustique à partir d'un point stricto sensu.

18: En supposant une propagation sphérique en champ libre : la pression ne dépend alors que de la distance à la source $\|\mathbf{x} - \mathbf{x}_0\|$.

19: Le terme quantification tel que défini ne réfère pas dans cette thèse à l'affectation d'une valeur dans un alphabet fini. Il décrit bien le niveau sonore d'une source acoustique donnée.

20: $p_{ref} = 20 \mu\text{Pa}$ en air contre $1 \mu\text{Pa}$ en eau.

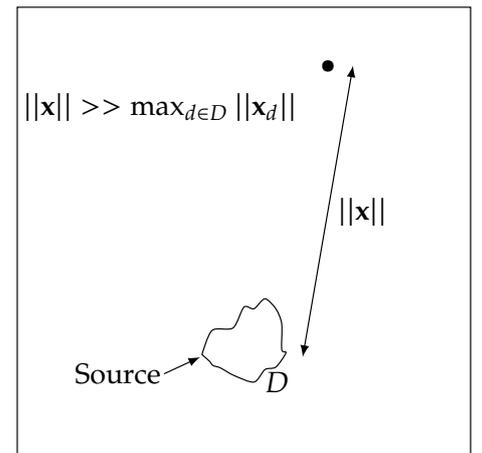


FIGURE 1.2 – Illustration de l'hypothèse de champ lointain en acoustique. Pour être en champ lointain d'une source acoustique, il faut que la distance source / point d'intérêt $\|\mathbf{x}\|$ soit grande devant la taille caractéristique de la source D .

En premier lieu, si le point de l'espace depuis lequel la source est observée \mathbf{x} est situé dans le **champ lointain géométrique**, représenté en Figure 1.2 et qui est traduit par une distance à la source ($\|\mathbf{x}\|$) grande devant la taille caractéristique de la source ($\max_{d \in D} \|\mathbf{x}_d\|$), la courbure des fronts d'ondes en ce point sont perçus comme identiques, quelle que soit la partie de la source qui les a générés. Mathématiquement, cette hypothèse permet une simplification du dénominateur de l'équation 1.11, qui peut être approximé par la distance $\|\mathbf{x}\|$ seule :

$$p(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|} \int_D A(\mathbf{x}_d) e^{-ik\|\mathbf{x}-\mathbf{x}_d\|} d\mathbf{x}_d. \quad (1.12)$$

En second lieu, et pour une source toujours placée en champ lointain, cette dernière est dite **compacte** si les retards spécifiques entre chaque monopôle de la distribution D et le point d'observation \mathbf{x} peuvent être assimilés à un retard global égal pour tous les points de la distribution. La compacité d'une source est vérifiée si elle émet un signal de longueur d'onde grande devant sa taille caractéristique comme représenté en Figure 1.3. Mathématiquement, cette condition permet la simplification de l'exposant de l'exponentielle complexe aboutissant à un champ de pression en \mathbf{x} donné par :

$$p(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|} e^{-ik\|\mathbf{x}\|} \int_D A(\mathbf{x}_d) d\mathbf{x}_d. \quad (1.13)$$

Cette solution est alors équivalente à une source monopolaire placée en un point unique de l'espace mais dont l'amplitude totale de la source distribuée est égale à :

$$A_D = \int_D A(\mathbf{x}_d) d\mathbf{x}_d. \quad (1.14)$$

Cette modélisation équivalente d'une source selon les deux hypothèses de champ lointain géométrique et de compacité est adaptée pour les cas où l'intégrale sur la région source D , qui donne l'amplitude totale équivalente, est différente de 0. La source modélisée rayonne alors de manière omnidirectionnelle, comme représenté en Figure 1.4a avec une source placée en $(0, 0)$ et un champ de pression calculé sur une région de $10 \text{ m} \times 10 \text{ m}$, générant des fronts d'ondes purement sphériques, montrés ici dans le plan (x, y) . La Figure 1.4b montre l'impact de l'ajout d'un second monopôle en phase avec le premier et séparé d'une distance faible ($\lambda/10$) devant la longueur d'onde émise : la propagation est légèrement modifiée, mais reste proche d'un comportement monopolaire. Néanmoins, pour les distributions dont l'intégration des amplitudes sur la région D est nulle, cette approche par amplitude équivalente n'est pas satisfaisante, elle soutiendrait que la source ne rayonne aucune énergie acoustique. Or de telles distributions sont envisageables, comme montré dans la Figure 1.4c, où est donné le champ de pression rayonné par deux monopôles en opposition de phase et séparés d'une distance égale à $\lambda/10$. En intégrant les amplitudes de ces sources sur une région D qui les englobe, l'amplitude totale est bien nulle par compensation de l'une par l'autre, mais un rayonnement acoustique est bien observé. Celui-ci est favorisé dans la direction qui passe par les deux monopôles, et au

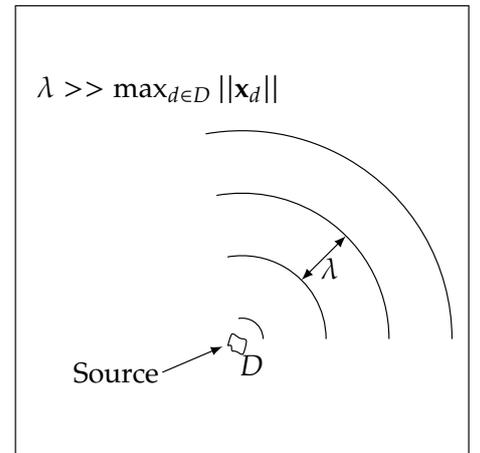


FIGURE 1.3 – Illustration de l'hypothèse de compacité en acoustique. Une source est dite compacte si elle génère des signaux acoustique de longueur d'onde λ grande devant sa dimension caractéristique D .

contraire nul dans la direction perpendiculaire : ce comportement est dit dipolaire, définissant ainsi cette distribution de source comme un dipôle. Une généralisation de ce processus est possible, donnant lieu aux sources multipolaires, qui permettent de décrire à chaque ajout de composante monopolaire un rayonnement de plus en plus complexe.

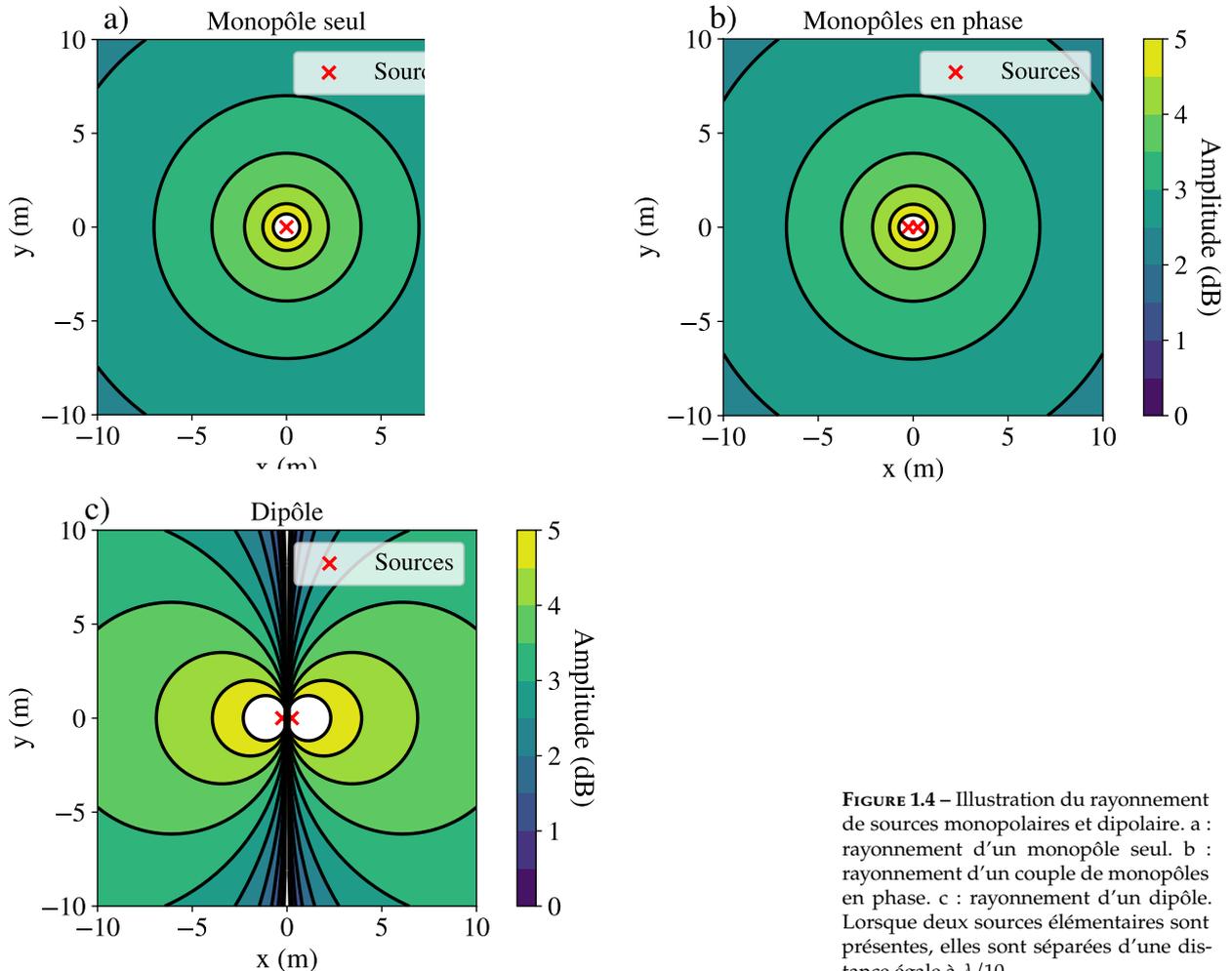


FIGURE 1.4 – Illustration du rayonnement de sources monopolaires et dipolaire. a : rayonnement d'un monopôle seul. b : rayonnement d'un couple de monopôles en phase. c : rayonnement d'un dipôle. Lorsque deux sources élémentaires sont présentes, elles sont séparées d'une distance égale à $\lambda/10$.

L'objectif de la modélisation de source est alors de combiner ces modèles théoriques de source, pour s'approcher au mieux du comportement du rayonnement de l'émetteur étudié. La partie suivante détaille ce processus de modélisation dans le cadre des émetteurs rencontrés dans cette thèse.

Modélisation de source dans la thèse

En combinant des distributions de monopôles, dipôles et sources multipolaires, une source acoustique peut être modélisée de manière très précise en décrivant des directivités non plus purement sphériques mais de forme arbitraire. Cependant, modéliser la directivité effectivement observée pour un émetteur donné requiert une étude poussée, qui doit être effectuée à chaque changement dans la configuration étudiée : changement de position ou d'orientation, masquage d'une partie de l'émetteur, changement de régime pour une machine en rotation, ou

autre. Lorsque l'objectif est de caractériser cette directivité pour modéliser une source de bruit complexe, cette étude est nécessaire et s'effectue dans des conditions expérimentales contrôlées : l'émetteur étudié est immobile, un grand nombre de capteurs est utilisé et un rapport signal à bruit important est requis²¹.

Dans cette thèse l'objectif est de caractériser des émetteurs lentement mobiles, situés en champ lointain géométrique, et dont le bruit rayonné résulte de la superposition de plusieurs phénomènes physiques différents²². Il est dès lors contre-productif d'optimiser le modèle de source pour une de ces configurations, car celui-ci provoquera des erreurs de modèles plus conséquentes qu'un modèle plus général lors des différents changements qui auront lieu pour l'émetteur ou son environnement. Il est préférable d'envisager un modèle *moyenné*, où l'émetteur est décrit de façon moins précise à un instant donné, mais pour lequel cette description est plus robuste.

Pour illustrer les effets de moyennage qui ont lieu pour un émetteur composé de multiples composantes bruyantes, une distribution plus complexe est proposée et schématisée en Figure 1.5. Elle est composée de quatre couples de monopôles, deux en phase dans les cadrans supérieur gauche et inférieur droit, et deux en opposition de phase (des dipôles) dans les cadrans supérieur droit et inférieur gauche. La position des monopôles dans les cadrans est prise aléatoirement, selon trois tirages consécutifs, pour générer différentes configurations afin de simuler les changements sur l'émetteur ou son environnement évoqués précédemment.

La Figure 1.6, qui représente le rayonnement de cette source simulée pour chacun des trois tirages, montre un comportement beaucoup plus régulier que pour les distributions théoriques de la partie précédente : la réponse globale des quatre couples tend vers un rayonnement omnidirectionnel par le moyennage des effets locaux de chacun d'entre eux. De plus, entre chaque tirage, le rayonnement obtenu varie, illustrant le caractère imprédictible du rayonnement d'un émetteur comprenant plusieurs composantes bruyantes.

En conclusion, les émetteurs rencontrés dans la thèse sont composés de multiples parties rayonnantes dont les contributions particulières se moyennent pour produire un rayonnement spécifique. De par son caractère imprédictible lors du mouvement de l'émetteur, du masquage de certaines de ses parties ou encore des fluctuations dans son régime d'utilisation, leur modélisation est faite par des sources monopolaires. Il s'ensuit une description simplifiée des émetteurs à un instant donné, mais robuste aux changements qui adviennent au cours du temps dans les conditions expérimentales visées.

21: La caractérisation de moteurs, que ce soit d'avion ou de voiture, se fait ainsi sur des bancs moteurs dont les paramètres de réglages sont très précis.

22: Pour des machines excavatrices le moteur, les bruits d'impacts sur le sol et les bruits de rotation des différents éléments se superposent. Pour les drones, le bruit des turbulences créées en bordure des pâles s'ajoute au bruit du moteur.

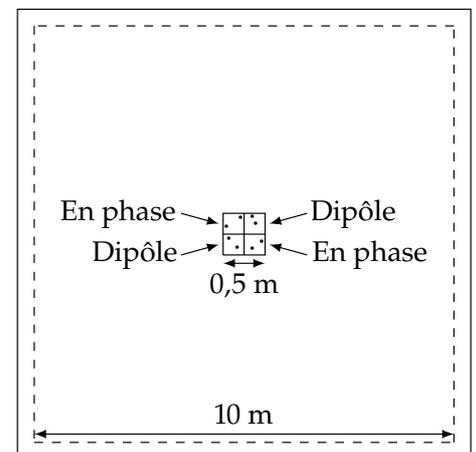


FIGURE 1.5 – Schéma de la distribution des couples de monopôles pour créer une source simulée le comportement d'un émetteur comprenant plusieurs composantes bruyantes. Chaque couple est disposé dans un cadran de 0,25 m de côté, et les positions des deux monopôles dans ce cadran sont tirées aléatoirement. Les trois tirages permettent de simuler des changements de configuration pour l'émetteur.

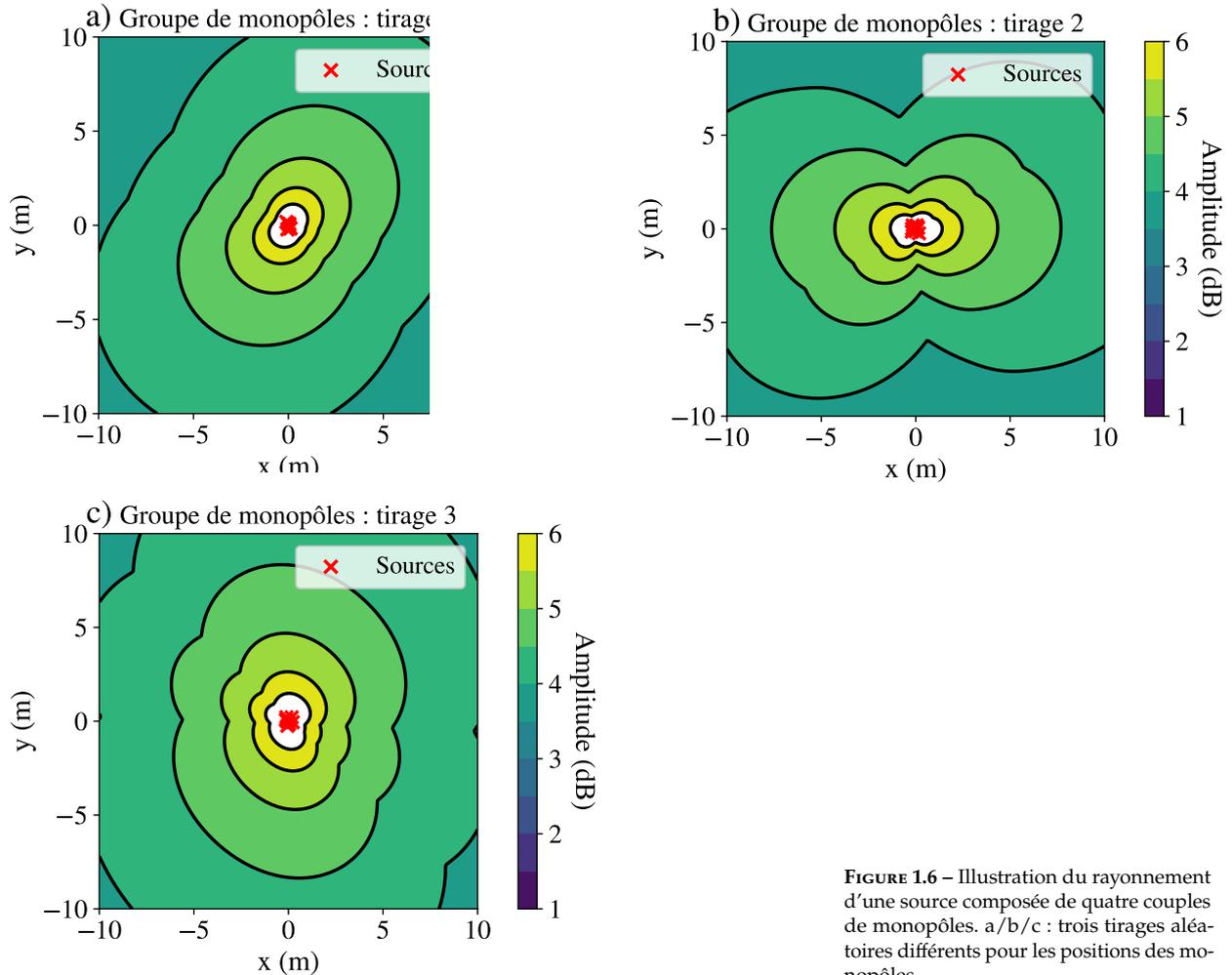


FIGURE 1.6 – Illustration du rayonnement d'une source composée de quatre couples de monopôles. a/b/c : trois tirages aléatoires différents pour les positions des monopôles.

1.4 Conclusion

Ce chapitre a permis d'établir le cadre acoustique de la thèse en commençant par établir les hypothèses considérées pour aboutir à l'équation des ondes : le milieu fluide est supposé uniforme, non visqueux, de conductivité thermique nulle et au repos, et les fluctuations d'amplitudes considérées sont faibles pour rester dans le cadre de l'acoustique linéaire. Sur les durées de traitement visées, les signaux acoustiques étudiés peuvent être considérés comme stationnaires et ergodiques, ce qui rend possible la résolution du problème dans le domaine fréquentiel, dont la formulation est donnée par l'équation de Helmholtz.

Les solutions de cette équation produisent différents modèles de propagation à sélectionner en fonction de la géométrie du problème à résoudre. Les deux modèles de propagation qui répondent aux enjeux expérimentaux de cette thèse sont alors détaillés : la propagation en champ libre est adaptée au cas où les ondes acoustiques se propagent dans un milieu considéré comme infini sans obstacle ou pour des interfaces sur lesquelles la réflexion est faible, tandis que la prise en compte d'un point image permet de prendre en compte les cas où les réflexions sur une interface ne sont plus négligeables.

Enfin, le modèle de source utilisé pour décrire les émetteurs de bruit

dans la thèse est déterminé. En partant des modèles théoriques de source que sont les distributions de sources monopolaires ou multipolaires, il est illustré qu'une distribution décrivant précisément le comportement du rayonnement des émetteurs n'est pas envisageable pour des configurations changeantes. Pour les émetteurs sonores visés dans cette thèse, composés de multiples composantes bruyantes dont la contribution globale tend vers un comportement monopolaire, cette description poussée serait même contre-productive. Chaque source acoustique est dès lors modélisée par un monopôle acoustique unique dans ce manuscrit, assurant une description robuste des émetteurs de signaux sonores au cours des essais expérimentaux. Pour les caractériser, c'est-à-dire estimer leur position, leur niveau sonore ou les classer dans une catégorie, les outils de traitement de signal que sont le traitement d'antenne et l'apprentissage statistique sont détaillés dans le chapitre suivant.

Outils de traitement du signal pour la caractérisation de sources acoustiques

2

Ce chapitre détaille les outils de traitement du signal utilisés pour caractériser les sources acoustiques rencontrées dans les applications expérimentales. La caractérisation de source est d'abord définie selon les besoins de ces applications, et s'oriente autour de trois axes qui ne sont pas systématiquement recherchés dans chacune d'entre elle : la localisation de la source, la quantification de son niveau de pression ou sa classification dans une catégorie de signal acoustique. Cette caractérisation est mise en œuvre à travers les outils que sont le traitement d'antenne pour la localisation et la quantification, et l'apprentissage statistique pour l'identification. Pour chacun d'entre eux, une synthèse des concepts qu'ils englobent est donnée avant de sélectionner plusieurs méthodes parmi la littérature pour répondre aux enjeux expérimentaux traités dans les chapitres applicatifs suivants.

Table des matières

2.1	Localisation, quantification et identification de sources acoustiques	17
2.2	Traitement d'antenne	18
	Formation de voies conventionnelle	24
	Méthode MUSIC	27
	Champ lointain en antennerie	30
	Conclusion pour le traitement d'antenne	30
2.3	Apprentissage statistique	31
	Classification binaire	33
	Classification à une classe	37
2.4	Conclusion	39

2.1 Localisation, quantification et identification de sources acoustiques

Comme vu dans le Chapitre 1, les émetteurs de signaux acoustiques sont modélisés dans cette thèse par des monopôles. Le but de cette partie est de montrer quelle caractérisation est faite de ces sources, en accord avec les enjeux expérimentaux des applications qui suivent. Une première information d'intérêt, commune aux deux applications, est l'estimation de la position spatiale des sources, par l'étape dite de **localisation**. Pour localiser une source acoustique, le traitement d'antenne, domaine qui se base sur la combinaison de signaux enregistrés par un réseau de N capteurs acoustiques, est particulièrement adapté. Quelques exemples utilisant un capteur unique existent mais il est alors nécessaire de connaître précisément le modèle de propagation et avoir connaissance des trajets multiples qui vont apporter de la diversité dans l'information [Jes+00; Tie+06]. Concernant l'application en prospection minière sous-marine, le niveau sonore émis par les machines sur le plancher océanique, c'est-à-dire la quantité de bruit qu'elles rayonnent dans leur environnement, constitue la seconde grandeur d'intérêt. Son estimation, appelée **quantification**, est possible grâce à diverses méthodes. De manière normée, c'est l'intensimétrie [ISO93; ISO96] ou la mesure de la puissance acoustique à partir de plusieurs capteurs indépendants à des distances préconisées, que ce soit en air [ISO01; ISO15] ou plus récemment en eau [ISO16; ISO19; ANS19], qui sont privilégiées. Mais les méthodes de traitement d'antenne permettant une quantification précise des sources se développent [Ant12; Le 16; Fle+11; Oer+07; Sij+04]¹, ce qui entrainera peut-être une évolution des normes en place. Leur intérêt est réel, car elles permettent de quantifier une source parmi d'autres grâce à l'obtention conjointe de la localisation et de la quantification². Dans le cadre de la protection de sites sensibles contre les drones, c'est l'**identification** de la source qui est faite en parallèle de sa localisation. Elle permet d'aller jusqu'à la classification de la source, selon des groupes, ou *classes*, qui décrivent le type de source étudié³. De nombreuses applications existent déjà en classification de signaux acoustiques comme par exemple l'identification de phénomènes sismiques [Hib+17], la détection d'évènements sonores anormaux [Mar+17], ou encore la reconnaissance de genres musicaux [Esm+04].

La caractérisation d'une source par sa **localisation**, sa **quantification** ou son **identification** est non-exhaustive, mais elle est suffisante pour répondre aux enjeux des différentes applications expérimentales visées. Dans d'autres cas, des grandeurs différentes peuvent avoir leur intérêt, comme par exemple la directivité⁴, qui nécessite un modèle de source plus avancé. Pour résumer, les sources acoustiques sont caractérisées dans cette thèse par deux des trois grandeurs suivantes, selon les cas d'applications :

Caractérisation possible des sources acoustiques

- ▶ **Localisation** de la position x dans l'espace tri-dimensionnel
- ▶ **Quantification** du niveau de pression sonore L_p à 1 m de la source
- ▶ **Identification** dans une classe C_i

1: Voir chapitre 6 de la thèse [Le 16] pour les applications expérimentales.

2: En prenant un exemple pour une voiture, il est possible de quantifier non pas le bruit de la voiture dans son ensemble, mais plutôt de séparer le bruit du moteur de celui du roulement.

3: Pour exemple, les types drone, jetski ou encore avion peuvent être envisagés.

4: i.e. la propension qu'a une source à rayonner dans certaines directions plutôt que d'autres.

Et pour établir cette caractérisation, des méthodes issues de deux domaines du traitement de signal sont utilisées :

- ▶ le traitement d'antenne pour assurer la **localisation** et la **quantification** ;
- ▶ l'apprentissage statistique pour assurer l'**identification**.

2.2 Traitement d'antenne

Le traitement d'antenne est un vaste domaine dont l'ensemble des applications dépasse le cadre acoustique de cette thèse. Il est utilisé pour localiser des sources électromagnétiques [Neh+94], pour concevoir des systèmes optiques capables de détecter des défauts [Kri+96], pour le traitement des signaux radar [Hay+93] ou encore pour la radio astronomie [Kra+05] ou le traitement de signaux sismiques [Gre+66; Car65]. Mais même en restant limité au traitement d'antenne pour l'acoustique, les applications restent très diverses comme par exemple :

- ▶ **en acoustique sous-marine** : l'application la plus connue étant le sonar développée en grande partie à des fins militaires [Bur91; Las77], mais d'autres enjeux existent comme la détection d'espèces marines [War+17; Tho04] ou la quantification de bruit de navires au passage [Oud15; Oud+18].
- ▶ **en acoustique aérienne** : avec la localisation d'avions en phase d'atterrissage [Mer+16], la caractérisation de moteurs sur banc d'essai [Le +20] ou la localisation de sources sur des émetteurs en mouvement comme les trains [Le 12] ou les avions [Sij+01].
- ▶ **dans l'imagerie médicale** : les ultrasons sont aussi régis par les équations de l'acoustique, généralement les sondes ultrasonores sont des systèmes actifs [Whi07], mais des exemples d'imagerie passive des ultrasons existent pour contrôler certains traitements [Pol19].
- ▶ **en traitement de la parole** : que ce soit pour localiser des orateurs [Bra+97], ou séparer des sources mélangées [Ehl+97] dans le cadre du problème de *cocktail party* [Che53].

Pour résoudre les problèmes soulevés par ces applications, de nombreuses méthodes ont été développées, avec comme point commun pour un grand nombre d'entre-elles de travailler avec un réseau de capteurs acoustiques⁵. Ce choix implique, pour une majorité des cas, un travail avec la grandeur physique de la pression, mais d'autres choix sont possibles comme exploiter la vitesse particulière acoustique [Wol+33; Ram19; Ram+18; Won+00; Win+09]⁶.

En se restreignant à l'utilisation des mesures de pression sur chaque capteur, il s'agit alors de les combiner pour caractériser les sources acoustiques environnantes. Les besoins en localisation et en quantification dans cette thèse permettent de limiter la littérature au domaine de l'imagerie acoustique qui répond à ces enjeux. Ses méthodes sont bien formalisées en acoustique aérienne, notamment grâce à des revues qui en couvrent un large spectre [Mer+19], mais qui omettent une partie d'entre elles que sont les méthodes haute résolution. Au contraire, l'acoustique sous-marine a permis l'essor de celles-ci [Red79; Bie+85; Mar98],⁷ mais celles issues

5: Des microphones dans l'air, des hydrophones dans l'eau.

6: Chapitre 2 de la thèse [Ram19].

7: Voir Chapitre 1 et 25 pour [Mar98].

du formalisme aérien commencent seulement à être envisagées [Sun+20; Yan18]. Cette synthèse tente ainsi de rassembler ces méthodes, pour le moment peu échangées d'un domaine à l'autre, mais dont l'utilisation peut s'avérer pertinente à la fois en aérien et en sous-marin.

L'objectif de l'imagerie acoustique est de produire une image d'une ou plusieurs sources de bruit présentes face à l'antenne⁸. Cette image peut représenter la localisation de ces sources, leur quantification, voire les deux. Une première possibilité pour obtenir la localisation des sources dans cette image consiste à déterminer les retards entre les mesures des différents capteurs dont les positions sont connues, avant de résoudre un problème inverse qui combine ces informations pour aboutir à une estimation de la position de la source [Car81; Cha+94; Ala+14; Bla+19b; Bla19]⁹. Néanmoins, cette solution ne donne pas accès à la quantification de la source. Pour l'estimer, les signaux temporels peuvent être utilisés directement, par exemple en les retardant et en les sommant comme dans le cadre d'une formation de voies conventionnelle [Ben+08]¹⁰. Ce choix assure un traitement pour l'ensemble des fréquences qu'ils contiennent, mais les opérations d'interpolation nécessaires à la remise en phase de signaux temporels sont généralement longues [Cou+19]. En exploitant la stationnarité des signaux étudiés sur les durées de traitement visées, il est possible de travailler dans le domaine fréquentiel grâce à la transformée de Fourier qui simplifie l'équation des ondes en celle de Helmholtz comme décrit dans le Chapitre 1. Les temps de calculs sont nettement réduits et l'analyse fréquentielle facilitée par la sélection et le traitement des signaux uniquement aux fréquences d'intérêt. Le passage en large bande se fait alors en rassemblant les résultats obtenus pour toutes les fréquences d'intérêt.

Dans ce cadre d'analyse, les signaux de pression sont concaténés dans un vecteur \mathbf{p} de taille $(M \times 1)$, représentant la transformée de Fourier des signaux mesurés par chacun des M capteurs à une fréquence donnée. L'imagerie acoustique cherche alors à décrire une distribution de K sources, dont les amplitudes sont données par $s_k, k \in \llbracket 1, K \rrbracket$, qui a produit ces mesures de pression \mathbf{p} en se propageant dans le milieu acoustique considéré selon les fonctions de transfert $\mathbf{h}_k, k \in \llbracket 1, K \rrbracket$. En ajoutant \mathbf{b} le vecteur qui décrit le bruit de mesure sur chacun des capteurs, le modèle suivant est obtenu :

$$\mathbf{p} = \sum_{k=1}^K s_k \mathbf{h}_k + \mathbf{b} \quad (2.1)$$

L'Equation 2.1 traduit le fait que les K sources se propagent vers les M capteurs selon les fonctions de transfert \mathbf{h}_k . Lorsque la position d'une source acoustique doit être estimée en pratique, cette équation est modifiée pour faire apparaître une grille de scrutation, comme illustré dans la Figure 2.1. Les sources sont alors recherchées sur cette grille de taille N_g ¹¹, en estimant les amplitudes de chaque point de cette grille, regroupées dans un vecteur \mathbf{q} ($N_g \times 1$). Les propagations entre ces positions potentielles de sources et les capteurs sont décrites par les fonctions de transfert \mathbf{G} ($M \times N_g$). Les colonnes de cette matrice \mathbf{g} , appelées vecteurs de pointage, correspondent aux vecteurs de propagation \mathbf{h}_k dans les cas

8: La généralisation en trois dimensions est possible, la reconstruction se faisant alors sur le volume qui entoure l'antenne.

9: Chapitre 2 pour la thèse [Bla19].

10: Voir chapitre 3.3.

11: Qui décrit des positions tridimensionnelles pour une propagation sphérique, et des angles d'arrivée pour une propagation en ondes planes.

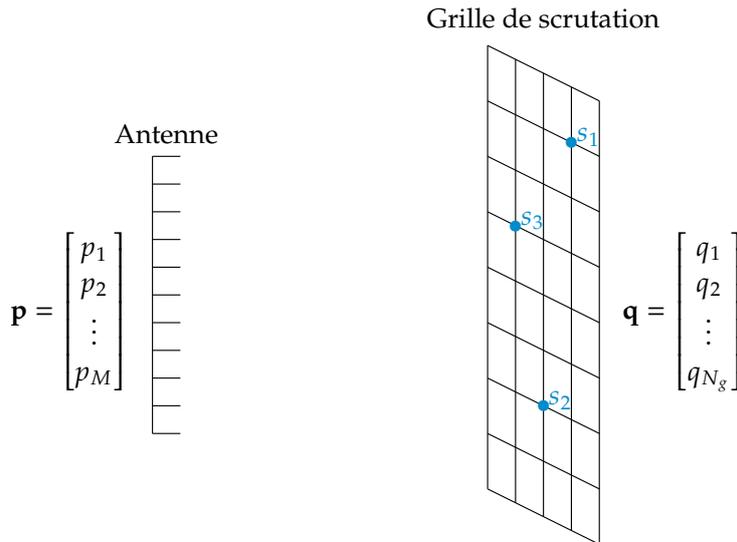


FIGURE 2.1 – Schéma des éléments qui composent un problème d’imagerie acoustique. Les K sources, représentées par les point bleus, propagent leur signaux jusqu’aux capteurs de l’antenne. Ces signaux sont passés dans le domaine fréquentiel dont la transformée de Fourier prise en une fréquence donne le vecteur \mathbf{p} . Les amplitudes des N_g positions de la grille de scrutation, regroupées dans le vecteur \mathbf{q} , sont ensuite recherchées pour établir si une source est effectivement positionnée au nœud testé.

où les positions des sources coïncident parfaitement avec un nœud de la grille. Le modèle à inverser prend ainsi la forme suivante :

$$\mathbf{p} = \mathbf{G}\mathbf{q} + \mathbf{b} \quad (2.2)$$

Les méthodes développées pour résoudre ce problème peuvent être regroupées dans deux familles [Lec+17; Sto+05; Mar98]¹² : les méthodes basées sur le filtrage spatial, et les méthodes inverses. Alors que la première famille teste les N_g positions une à une et de manière indépendante, en leur associant une probabilité de présence de source¹³, la seconde cherche à inverser l’équation Equation 2.1 dans sa globalité, et combine les N_g sources potentielles mises à disposition dans le vecteur \mathbf{q} pour décrire le champ acoustique effectivement reçu sur l’antenne.

Les méthodes basées sur le filtrage spatial peuvent être divisées en deux sous-catégories : les méthodes non-paramétriques et paramétriques.

Méthodes non-paramétriques basées sur le filtrage spatial : les méthodes non-paramétriques voient le problème d’imagerie acoustique comme un problème de conception d’un filtre spatial, avec un degré de complexité qui peut varier en fonction de la sélectivité spatiale souhaitée [Van+88]. Une grande proximité entre le traitement d’antenne et l’analyse spectrale découle de cette vision, avec même une équivalence entre les méthodes dans le cas de l’antenne linéaire uniforme, comme en témoignent les ouvrages qui traitent les deux domaines simultanément [Sto+05; Mar98]. Le point d’entrée de ces méthodes est la formation de voies, dont le principe est de compenser les retards qu’une source positionnée en \mathbf{x} aurait provoquée sur les capteurs de l’antenne, avant de sommer la sortie de chaque capteur. Si les retards ont été correctement corrigés, la somme sera alors constructive et augmentera le rapport signal à bruit d’une source effectivement placée à cette position. En pratique, la formation de voies est alors le résultat de N_g filtres spatiaux, qui couvrent toutes les N_g positions \mathbf{x} d’intérêt¹⁴. La position de la source est estimée en prenant le maximum parmi toutes celles testées, i.e. la position pour laquelle la somme a été la plus constructive. En étudiant toutes les positions de l’espace de manière indépendante, la

12: La synthèse de l’imagerie acoustique présentée ici fusionne les familles établies dans chacune de ces contributions.

13: Qui n’a souvent de signification physique que si une source est effectivement présente en la position testée.

14: \mathbf{x} pouvant représenter ici des positions angulaires ou spatiales.

formation de voies suppose que le bruit ambiant impactant l'antenne est réparti uniformément dans l'espace¹⁵, ce qui en représente sa principale limitation.

Les algorithmes dits adaptatifs cherchent alors à prendre en compte les données qui arrivent effectivement sur l'antenne en exploitant la matrice de corrélation des mesures faites par les capteurs [Cap69; Sto+02; Li+03; Vor+03]¹⁶. D'autres améliorations sont envisageables autour de l'optimisation des poids du filtre spatial pour obtenir des réponses d'antenne contraintes afin de réduire les artefacts provenant de positions spatiales prédéfinies [Van+88], ou en favorisant les positions de l'espace qui contiennent du signal en utilisant des filtres non-linéaires qui généralisent les approches adaptatives [Pis72; Dou14a; Dou14b].

Une dernière façon d'améliorer les performances brutes de la formation de voies est d'utiliser la connaissance de la réponse d'antenne, directement décrite par la disposition géométrique des capteurs. En prenant en compte cette connaissance supplémentaire, il faut néanmoins réaliser que ces méthodes sortent un peu du cadre de la sous-catégorie des méthodes non-paramétriques étudiées ici. Le principe suivi est alors celui de déconvolution, qui consiste à retirer la réponse d'antenne à la carte initiale de formation de voies, pour les sources qui auront alors été détectées. Deux manières d'envisager la déconvolution existent : soit en retirant à chaque itération la réponse d'antenne du maximum actuel de la carte disponible [Sij07], soit en résolvant un problème d'optimisation qui cherche à réduire l'erreur entre la carte initiale, et une distribution potentielle de source [Bro+06; Ehr+07].

Méthodes paramétriques basées sur le filtrage spatial : le problème intrinsèque des méthodes non-paramétriques est le fait que leurs performances en résolution sont limitées, même à durée d'observation infinie : elles n'arrivent pas à séparer deux sources spatialement proches [Mar98]. En imposant aux signaux d'observations le fait de suivre un modèle défini par un nombre donné de paramètres, d'autres approches qui contournent ce problème ont pu être envisagées. C'est l'objectif des méthodes dites haute-résolution¹⁷ qui voient leurs origines remonter aux problèmes d'estimation de sommes de signaux uniquement composées d'exponentielles par Prony [Pro95], soit un modèle aux paramètres très restrictifs. Une première amélioration par le modèle de Prony étendu développé par Hildebrand ne change pas le modèle mais cherche à rendre plus robuste l'estimation [Hil56]. En considérant le modèle de Wold qui représente le signal comme étant la sortie d'un filtre causal excité par un bruit blanc, ce sont les modèles auto-régressifs (AR) et auto-régressifs avec moyenne mobile (ARMA, *autoregressive moving average*) qui sont à utiliser en fonction du type de signal étudié [Mar98]¹⁸. Mais le réel essor des méthodes haute-résolution a lieu grâce aux méthodes de décomposition en sous-espaces de la matrice de corrélation [Mer76]. Elles s'appliquent pour un modèle paramétrique composé de K signaux de matrice de corrélation non singulière, auxquels s'ajoute un bruit de mesure spatialement blanc¹⁹. Ces méthodes utilisent la diagonalisation pour séparer les données d'entrée en deux sous-espaces appelés sous-espace signal et sous-espace bruit. L'ensemble ou une partie de ces espaces est utilisé pour déduire les positions de sources recherchées. Ainsi, il est possible de travailler avec le sous-espace bruit en utilisant sa propriété d'orthogonalité avec le signal, en prenant un vecteur de ce sous-espace

15: Ce qui correspond techniquement à minimiser la puissance en sortie d'un filtre spatial alimenté par un bruit spatialement blanc [Sto+05], chapitre 6.

16: En cherchant cette fois à minimiser la puissance en sortie d'un filtre alimenté par les données [Sto+05], chapitre 6.

17: Pour des synthèses bibliographiques plus poussées des méthodes haute-résolution uniquement, voir les thèses de R. Badeau [Bad05] et G. Bouleux [Bou07]

18: pp. 65-70

19: $E[\mathbf{b}\mathbf{b}^T] = \sigma_b^2 I_M$

estimé [Kum+83] ou en les combinant comme avec la méthode Multiple Signal Classification (MUSIC) [Bie+83; Sch86]. Mais le sous-espace signal peut aussi être utilisé, en exploitant une géométrie particulière d'antenne composée de deux sous-partie translattées l'une par rapport à l'autre [Roy+89]. Pour le cas d'une antenne linéaire uniforme, les équations qui décrivent le problème prennent la forme de polynômes dont les zéros correspondent aux positions des sources, donnant lieu aux versions *root* des algorithmes précédents qui produisent des estimations plus précises [Rao+89].

Un des inconvénients des méthodes haute-résolution réside dans leur non prise en compte de la corrélation entre les sources potentiellement présentes dans leur environnement acoustique. C'est pourquoi elles ont été améliorées au fil du temps, les faisant ainsi ressembler, dans leur philosophie, aux méthodes inverses : le but de ces extensions est de prendre en compte le problème dans sa globalité. C'est l'objectif de la méthode du maximum de vraisemblance par exemple, qui se décline sous une forme déterministe dans laquelle les amplitudes des signaux sont des paramètres déterministes à estimer, ou sous une forme stochastique où elles sont des variables aléatoires dont il faut déterminer les statistiques [Mar98]²⁰. L'algorithme de MUSIC a aussi été étendu dans une forme multidimensionnelle [Cad88], et exploite toujours le sous-espace bruit alors que les méthodes de pondération des sous-espaces [Vib+91] généralisent la façon de combiner les sous-espaces entre eux. L'intérêt de ces extensions réside dans leur possibilité de prendre en compte des sources corrélées, mais cela à un prix : la simplicité de mise en œuvre des méthodes haute-résolution précédemment décrites est perdue et les équations obtenues sont bien plus complexes à optimiser, nécessitant des méthodes numériques.

20: Chapitres 4.3 et 4.2, respectivement.

Méthodes inverses : l'objectif de l'ajout des extensions aux méthodes haute-résolution et de les faire fonctionner dans des cas de figures que leurs hypothèses de base ne prennent pas en compte. Elles tendent en ce sens à ressembler aux méthodes inverses qui elles, cherchent à inverser directement l'équation Equation 2.2, ce qui leur permet de prendre en compte ces cas complexes en recherchant simultanément les positions, les niveaux de pression acoustiques, et les corrélations qu'il existe entre les sources. Dans ce cadre, c'est la nature des sources et de leur propagation qui définit la méthode à utiliser [Lec+17], chaque méthode permettant de répondre à une modélisation spécifique des sources. Pour des sources décrites par des ondes planes en champ proche, les algorithmes d'holographie acoustique [Wil+80; Hal09] ont été développés. Pour les distributions monopolaires, des discussions sur la faisabilité de l'inversion sous certaines conditions [Nel+00; Lec09] ont été suivies par une vision plus globale du problème par des approches bayésiennes [Ant12]. Une représentation des sources sous forme d'harmoniques sphériques est aussi proposée [Wu98]. Enfin une dernière méthode résout les équations de l'acoustique par la méthode des éléments de frontière [Bai92], ce qui permet de s'affranchir de contraintes géométriques sur le problème à résoudre.

Comme toujours une synthèse de méthodes est perméable, et celle proposée dans cette thèse n'échappe pas à la règle. Les extensions de la formation de voies à travers la déconvolution, et celles des méthodes haute-résolution par la prise en compte de la corrélation entre les sources,

peuvent aussi être vues comme d'autres méthodes inverses. Elles ne proviennent pas directement du formalisme décrit dans la partie précédente mais s'en sont approchées au fur et à mesure, pour leur permettre de s'adapter à des cas pratiques que leurs hypothèses de départ n'appréhendaient pas.

Pour répondre aux enjeux des deux applications expérimentales visées, les méthodes de traitement d'antenne doivent être sélectionnées au sein de cette littérature selon leur adéquation aux attentes spécifiques de ces projets :

Attentes vis-à-vis des méthodes de traitement d'antenne

- ▶ Elles sont souhaitées **robustes** pour fonctionner dans des cadres expérimentaux variés, aérien et sous-marin, et potentiellement en présence de bruit ambiant important.
- ▶ Une **localisation précise** doit être obtenue, en présence de sources qui peuvent être proches spatialement ou angulairement, ou pour des sources dont les niveaux sonores peuvent être fortement différents.
- ▶ La **quantification** du niveau de pression acoustique d'une source doit être apportée.
- ▶ Le **filtrage spatial** d'un signal dans le domaine **temporel** doit être possible, pour permettre la focalisation de l'écoute dans une direction donnée sans distorsion de ce dernier.
- ▶ Un moyen de **comparaison des performances** d'une antenne par rapport à une autre est nécessaire, pour déterminer quelle antenne est à concevoir parmi un panel d'arrangements de capteurs.
- ▶ Leur **temps de calcul** doit être **court**, afin d'autoriser des extensions temps-réel pour la suite des projets abordés dans les applications expérimentales.

La première condition écarte les méthodes inverses, dont les paramètres qui permettent d'ajuster le modèle à la configuration pourraient être adaptés à une application mais pas à une autre. Seul le modèle bayésien permet plus de souplesse, mais au prix d'un temps de calcul plus important qui entre en contradiction avec le dernier point cité ci-dessus. Pour leur précision supérieure à celle des méthodes adaptatives, et leur capacité à distinguer des sources proches ou de niveaux sonores différents, les méthodes haute-résolution sont choisies pour effectuer l'étape de localisation. L'étude de leurs performances montre que MUSIC est la méthode haute-résolution sans contrainte géométrique sur l'antenne qui localise les sources avec le moins de variance [Li+91; Thu+85; Sha+84; Kri+92], mais elle peut parfois être surpassée en pouvoir de résolution [Sha+85; Thu+86]. Seule la méthode ESPRIT (Estimation of Signal Parameters via Rotational Invariance Techniques en anglais) parvient à être plus précise, mais en prenant en compte une antenne composée de deux sous-parties translattées l'une par rapport à l'autre [Ott+91]. Etant donné les géométries non régulières des antennes utilisées par la suite, cette famille de méthode ne peut être sélectionnée et le choix de la méthode pour localiser les sources se porte donc sur la méthode MUSIC.

Pour répondre au besoin de comparaison des performances d'an-

tenne, la formation de voies dans sa forme normalisée est sélectionnée afin d'obtenir les réponses d'antenne de chacun des arrangements de capteurs testés. Des critères interprétables physiquement, détaillés dans la sous-partie suivante, peuvent être calculés sur ces réponses d'antenne avant d'être comparés pour établir quel arrangement est à concevoir [Sar16].

Concernant le filtrage spatial sans distorsion, peu de méthodes vont à ce jour plus loin que la formation de voies, et celles-ci demandent un accroissement du temps de calcul non négligeable [Cou+19]. La formation de voies est par conséquent à nouveau choisie, mais dans le domaine temporel, pour effectuer la focalisation des signaux.

Enfin pour la quantification, une première approche est d'utiliser aussi la formation de voies, dont une version peut quantifier une source si elle est correctement séparée des autres sources présentes dans l'environnement acoustique. Bien que sous-optimal, c'est ce choix qui est fait dans cette thèse en ne quantifiant que sur une zone située autour de la localisation faite au préalable par la méthode MUSIC.

Les deux sous-parties suivantes détaillent les deux méthodes sélectionnées qui répondent aux besoins applicatifs : la formation de voies et la méthode MUSIC.

Formation de voies conventionnelle

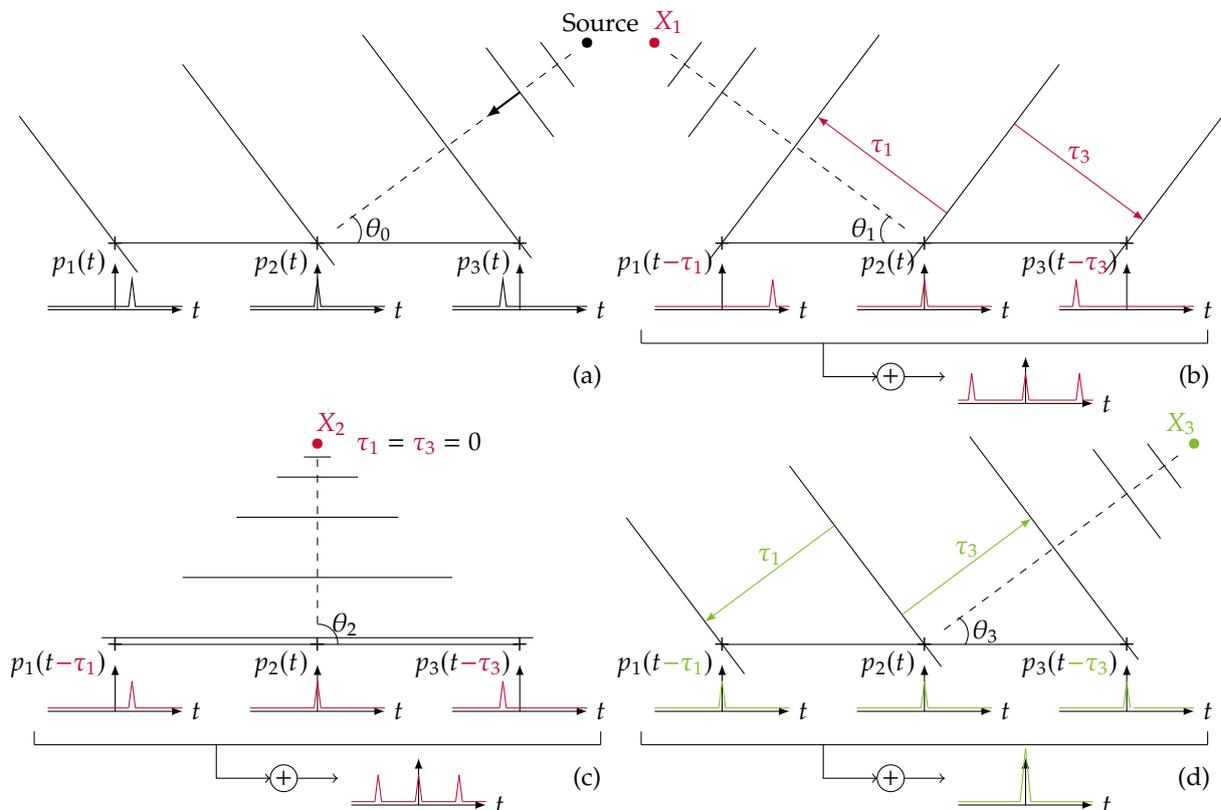


FIGURE 2.2 – Schéma illustrant le principe de la formation de voies. Trois positions X_1 , X_2 et X_3 sont successivement testées en b, c et d, pour la présence de source. Les signaux sont retardés en accord avec la position en cours de test, puis sommés pour obtenir des interférences constructives dans le cas de la présence d'une source. Pour la position X_3 , où est effectivement placée la source comme montré en a, la somme est bien constructive et le signal de cette source est retrouvé avec une plus grande amplitude.

Le principe de la formation de voies s'illustre particulièrement bien dans le domaine temporel et en ondes planes, comme montré en Figure 2.2 : il s'agit d'appliquer un retard aux signaux enregistrés par les différents capteurs puis de sommer ces signaux retardés. Le retard est déterminé par la position dans la grille qui est en cours de test (X_1 , X_2 et X_3 successivement dans la Figure 2.2, donc $N_g = 3$). Pour une source effectivement présente (ici en X_3), la somme des signaux retardés sera constructive, ce qui permet d'obtenir un signal de puissance plus élevée que pour les autres positions (X_1 et X_2). Ces valeurs de puissance pour chaque position de la grille de scrutation constituent la carte de formation de voies recherchée.

Le signal temporel obtenu en Figure 2.2d est ainsi filtré spatialement, et correspond à la somme des M signaux enregistrés sur les capteurs. Ce processus est appelé focalisation dans ce manuscrit, et il permet d'améliorer le rapport signal à bruit d'une source en comparaison avec son écoute par un seul capteur.

Ce principe illustré ici dans le domaine temporel est valable dans le domaine fréquentiel, où les signaux sont passés par transformée de Fourier, dont une fréquence unique est gardée pour tous les capteurs dans le vecteur \mathbf{p} comme détaillé dans la partie précédente. Les retards pour une position \mathbf{x} de la grille de scrutation sont alors appliqués sur les données en les multipliant par le vecteur de pointage $\mathbf{g}(\mathbf{x})$ ²¹. En fonction de l'objectif envisagé, localisation ou quantification, différents estimateurs sont dérivés et dépendent de la façon dont sont normalisés les produits entre les données et ces vecteurs de pointage [Sar12].

21: Pour la clarté des équations à venir, la position \mathbf{x} sera omise lorsque le vecteur de pointage \mathbf{g} sera utilisé.

Lorsque l'objectif est de localiser uniquement la position de la source, la formation de voies peut être vue comme une mesure de colinéarité entre le vecteur de pointage \mathbf{g} et le vecteur des pressions mesurées \mathbf{p} . Le cosinus entre \mathbf{g} et \mathbf{p} établit cet indicateur de colinéarité :

$$\cos(\mathbf{p}, \mathbf{g}) = \frac{\mathbf{p}^* \mathbf{g}}{\|\mathbf{p}\| \|\mathbf{g}\|}, \quad (2.3)$$

avec $*$ qui représente la transposée conjuguée.

Une estimation de \mathbf{p} est nécessaire pour obtenir cet indicateur. Il est alors commun à un grand nombre de méthodes d'estimer non pas le vecteur de pression \mathbf{p} ($M \times 1$) directement, mais plutôt la matrice interspectrale à une fréquence donnée $\Gamma(f) = \mathbb{E}[\mathbf{p}\mathbf{p}^*](f)$ ($M \times M$). En théorie, la matrice interspectrale correspond à la transformée de Fourier de la matrice de corrélation des signaux d'entrée. A une fréquence donnée, elle contient donc sur sa diagonale les auto-spectres des capteurs, et hors de sa diagonale les inter-spectres des capteurs entre eux. En pratique, elle est estimée en moyennant, pour L blocs d'observations successives de signal, le produit conjugué des transformées de Fourier des signaux à la fréquence souhaitée pour obtenir un périodogramme de Welch :

$$\hat{\Gamma}(f) = \frac{1}{L} \sum_{l=1}^L \mathbf{p}_l(f) \mathbf{p}_l^*(f), \text{ avec } \mathbf{p}_l(f) = TF[\tilde{\mathbf{p}}_l(t)](f). \quad (2.4)$$

Les signaux temporels reçus sur l'antenne pour le bloc l sont donnés par $\tilde{\mathbf{p}}_l$. La transformée de Fourier est donnée par $TF[\cdot]$. Par souci de

clarté, la variable f sera omise dans le reste du manuscrit. En prenant la forme quadratique de l'équation Equation 2.3, un estimateur classique de la formation de voies qui utilise $\hat{\mathbf{\Gamma}}$ est obtenu :

$$I_{loc}(\mathbf{x}) = \frac{\mathbf{g}^* \hat{\mathbf{\Gamma}} \mathbf{g}}{\|\mathbf{g}\|^2 \text{Tr}(\hat{\mathbf{\Gamma}})}, \quad (2.5)$$

avec $\text{Tr}(\cdot)$ l'opérateur trace d'une matrice.

Cet estimateur localise la source sans biais, et les cartes d'imagerie qu'il produit sont normalisées et représentées en décibel²².

$$22: I_{loc}^{dB}(\mathbf{x}) = 10 \log_{10}(I_{loc}(\mathbf{x}))$$

Lorsque l'objectif est de quantifier le niveau de pression acoustique des sources en présence, une autre dérivation de la formation de voies est nécessaire. De plus, l'estimation de ce niveau n'est valable que dans des conditions restreintes : pour une source seule ou pour des sources éloignées spatialement et décorréliées. Dans le cas contraire, la réponse d'antenne inhérente à la formation de voies perturbe les niveaux obtenus, en additionnant au niveau réel à une position spatiale donnée la contribution des lobes secondaires des autres sources en présence. Dans les conditions mentionnées, la version quantitative de la formation de voies a pour objectif d'estimer correctement l'amplitude d'une source donnée grâce aux mesures. Il est alors naturel de chercher à minimiser l'écart entre ces mesures, et la propagation d'une source d'amplitude A à la position \mathbf{x} pointée par le vecteur de pointage \mathbf{g} :

$$\epsilon(\mathbf{g}, A) = \|\mathbf{p} - A\mathbf{g}\|^2. \quad (2.6)$$

Minimiser cette erreur quadratique amène à estimer l'amplitude à la position \mathbf{x} par :

$$A(\mathbf{x}) = \frac{\mathbf{p}^* \mathbf{g}}{\|\mathbf{g}\|^2}, \quad (2.7)$$

et de la même manière que pour l'indicateur I_{loc} , l'estimation se fait directement en utilisant la forme quadratique de l'équation Equation 2.7, ce qui aboutit à une estimation de la moyenne quadratique de l'amplitude de la pression source par :

$$I_{quant}(\mathbf{x}) = \widehat{A}_{eff_g}^2(\mathbf{x}) = \frac{\mathbf{g}^* \hat{\mathbf{\Gamma}} \mathbf{g}}{\|\mathbf{g}\|^4}, \quad (2.8)$$

qui, passé en décibels par l'équation Equation 1.9, aboutit à l'estimation du niveau de pression à 1 m \hat{L}_p de la position \mathbf{x} de grille considérée par le vecteur de pointage \mathbf{g} . Cet estimateur quantifie le niveau de pression d'une source effectivement présente à la position de grille testée sans biais, mais pour tous les autres points, cette estimation n'est pas représentative d'un niveau sonore physique. Dans cette thèse il est donc fait le choix d'utiliser cet estimateur de manière locale, à la position d'une source effectivement localisée par une autre méthode au préalable.

Le critère I_{loc}^{dB} correspond à un passage en décibel avec une pression de référence égale à 1 Pa, impliquant ainsi un niveau maximum de la carte égal à 0 dB. Des exemples de cartes simulées pour une source située

en (0 m, 50 m), sur une grille de 200 m par 200 m, avec un pas de 0.5 m dans chaque direction, et de niveau de pression à 1 m égal à 150 dB sont données dans la Figure 2.3. L'antenne utilisée correspond à celle utilisée dans le projet Alysound, qui sera décrite dans le Chapitre 3. Elle est placée 50 m au-dessus du plan reconstruit²³. Pour la carte de localisation calculée à partir du critère I_{loc}^{dB} et donnée en Figure 2.3a, la normalisation est bien observée avec un maximum de la carte à 0 dB. Pour la carte de quantification qui estime le niveau de pression \hat{L}_p en Figure 2.3b, scruter des positions trop éloignées du dessous de l'antenne produit une surestimation qui empêche la localisation correcte de la source²⁴. En revanche, à la position de la source (0 m, 50 m), le niveau de la carte est exactement égal à 150 dB. Pour comparer les antennes l'une par rapport à l'autre, il est alors logique de travailler avec la carte I_{loc}^{dB} qui est indépendante du niveau de source émis en entrée. Deux critères sont calculés pour pouvoir établir les performances des antennes :

- ▶ le niveau maximum des lobes secondaires (NMLS), qui sont les lobes présents dans la carte de formation de voies obtenue, mais qui ne correspondent pas à des sources physiques. Leur niveau maximum traduit la capacité qu'a une antenne à détecter deux sources de niveaux différents. Plus le NMLS sera bas, plus l'antenne sera capable de détecter une source de niveau faible par rapport à la source la plus forte, qui serait sinon masquée par les lobes secondaires de la plus forte. Dans l'exemple donné en Figure 2.3a, les lobes secondaires les plus hauts sont marqués par les croix × et leur niveau est égal à $NMLS = -7,2$ dB.
- ▶ la résolution, évaluée grâce à la largeur à -3 dB du lobe principal, qui traduit la capacité de l'antenne à séparer deux sources spatialement proches. Plus la résolution de l'antenne est élevée, plus le lobe principal d'une source est petit, et ainsi deux sources proches pourront être distinguées l'une de l'autre. Dans l'exemple, la largeur à -3 dB est illustrée par la double-flèche blanche et égale à $L_{-3dB} = 25$ m.

Ces deux critères sont concurrents lorsque des limites sont introduites sur le nombre de capteurs à utiliser ou la taille maximum de l'antenne : une antenne avec des lobes secondaires bas ne sera pas forcément bien résolue. Pour pallier cette limitation, les méthodes haute-résolution se basent sur un modèle paramétrique des signaux et de bruit. La plus robuste d'entre elle, qui servira à la localisation des sources dans les applications expérimentales, est développée dans la partie suivante.

Méthode MUSIC

La méthode MUSIC est basée sur le concept de sous-espaces [Mer76]. L'idée est de séparer les données en deux, une partie dans le sous-espace signal et l'autre dans le sous-espace bruit, ces deux sous-espaces étant orthogonaux. Une fois cette séparation réalisée, MUSIC exploite cette orthogonalité en projetant le vecteur de pointage à tester sur l'espace bruit, pour obtenir une valeur très faible en présence de signal. La dernière étape est d'inverser cette projection afin d'obtenir des valeurs théoriquement infinies pour les vecteurs qui pointent effectivement vers une source physique, et des valeurs plus faibles ailleurs.

23: Qui correspond en pratique au plan du sol.

24: Le maximum de la carte étant situé en (0 m, -200 m).

En pratique, la séparation des sous-espaces se fait par une décomposition en éléments propres de la matrice interspectrale Γ qui permet d'obtenir une matrice diagonale \mathbf{D} , qui contient les valeurs propres λ_k de la décomposition, rangées par ordre décroissant, et une matrice unitaire \mathbf{U} :

$$\Gamma = \mathbf{U}\mathbf{D}\mathbf{U}^* = \sum_{k=1}^{K_s} \lambda_k \mathbf{u}_k \mathbf{u}_k^* + \sum_{l=K_s+1}^M \lambda_l \mathbf{u}_l \mathbf{u}_l^*. \quad (2.9)$$

Les K_s premières valeurs propres correspondent à la puissance des sources présentes, alors que les suivantes représentent la puissance du bruit de mesure. Les K_s premières colonnes \mathbf{u}_k de \mathbf{U} qui correspondent aux premières valeurs propres contiennent une combinaison des vecteurs de pointage normalisés vers les sources physiquement présentes. L'enjeu de la décomposition est de séparer les vecteurs propres \mathbf{u}_k qui correspondent aux sous-espaces signal et bruit, respectivement. Ce choix se fait en étudiant la décroissance des valeurs propres issues de la décomposition, données par \mathbf{D} . La frontière est déterminée par le point d'inflexion observé entre une partie décroissante liée aux K_s premières valeurs propres représentant le signal, et un plateau qui correspond au reste des valeurs propres représentant le bruit de mesure. Même si en théorie ce point d'inflexion doit refléter le nombre de sources en présence K , en pratique ce n'est pas toujours le cas, si bien qu'un seuillage est nécessaire pour un bon fonctionnement de la méthode. Bien qu'il existe de nombreux critères qui ont été développés afin d'estimer automatiquement le nombre de valeurs propres à associer aux sous-espaces [Aka98; Bie+83; Sch78; Wax+85; Qui+06; Fer+06], ils ne sont pas utilisés dans nos applications pour lesquelles ce nombre est déterminé manuellement. En acoustique sous-marine, ce réglage manuel s'est avéré nécessaire pour obtenir les cartes les plus lisibles possibles en le modifiant plusieurs fois si besoin par essai / erreur. En aérien, c'est le nombre de sources attendues dans chaque configuration qui définit le nombre de valeurs propres à garder. Ce choix, bien que sous-optimal, a donné des résultats satisfaisants et robustes, ce qui a évité de devoir passer plus de temps sur ce point technique qui peut être chronophage. Enfin, il est à noter qu'une sous-estimation du nombre de valeurs propres à inclure dans le sous-espace signal est beaucoup plus problématique qu'une surestimation [Mar98]²⁵. En cas de doute, il faudra ajouter des valeurs propres dans ce sous-espace.

25: Voir p. 94.

Une fois les sous-espaces déterminés, l'indicateur de MUSIC est alors donné, pour une position \mathbf{x} pointée par le vecteur de pointage \mathbf{g} , par l'inverse de la projection sur le sous-espace bruit de ce vecteur de pointage [Bie+83; Sch86]. En notant, $\hat{\mathbf{U}}_B$ la matrice dont les colonnes correspondent aux $M - K_s$ vecteurs propres associés au sous-espace bruit de la matrice interspectrale estimée à partir des données, il vient :

$$I_{MUSIC}(\mathbf{x}) = \frac{\|\mathbf{g}\|^2}{\mathbf{g}^* \hat{\mathbf{U}}_B \hat{\mathbf{U}}_B^* \mathbf{g}}. \quad (2.10)$$

Cet indicateur de MUSIC est généralement présenté pour des vecteurs de pointage de norme unitaire, rendant le numérateur égal à 1. Ceci est le cas lorsqu'ils décrivent des directions d'arrivées angulaires.

Mais dans le cas d'une propagation sphérique, la norme des vecteurs de pointage n'est plus égale à 1 et la présence de leur norme au carré au numérateur est requise pour obtenir une carte de localisation qui ne dépend pas de la distance.

Les valeurs calculées par cet indicateur ne reflètent pas le niveau de pression des sources. Une source d'amplitude plus faible qu'une autre peut même produire une valeur plus forte que celle-ci dans la carte MUSIC obtenue. Cette valeur traduit juste l'orthogonalité entre les données à disposition et le vecteur de pointage testé. Elle est généralement donnée en decibel comme pour le critère I_{loc}^{dB} ²⁶.

Pour MUSIC, les critères d'évaluation décrits pour la formation de voies ne sont théoriquement pas applicables : les lobes secondaires n'ont pas de réel sens physique, au même titre que le niveau du maximum, et la résolution est théoriquement infinie. Ceci s'observe dans la Figure 2.3c, où l'indicateur I_{MUSIC} est calculé sur le même cas simulé que pour les deux cartes de formation de voies. La carte ne montre qu'une valeur très forte obtenue exactement à la position de la source (0 m, 50 m), rendant le calcul des critères définis précédemment non pertinent. Il n'en reste pas moins que sur des données réelles, un lobe principal et des artéfacts sont observables dans les cartes MUSIC, ce qui rend possible la détermination d'une résolution et d'un niveau comparable au NMLS, donné par la différence entre le maximum de la carte et le niveau maximum des artéfacts de la carte.

$$26: I_{MUSIC}^{dB}(\mathbf{x}) = 10 \log_{10}(I_{MUSIC}(\mathbf{x}))$$

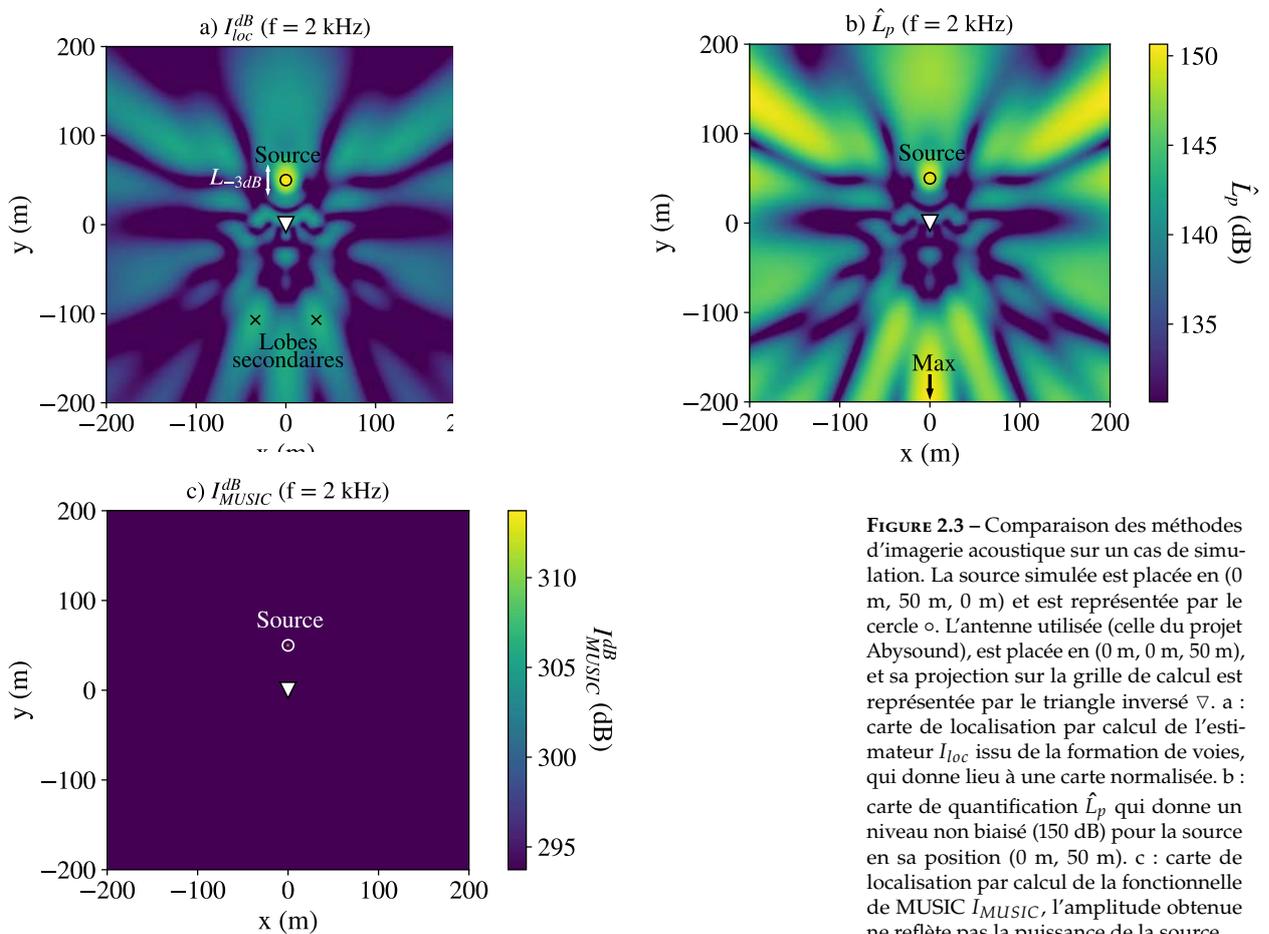


FIGURE 2.3 – Comparaison des méthodes d'imagerie acoustique sur un cas de simulation. La source simulée est placée en (0 m, 50 m, 0 m) et est représentée par le cercle \circ . L'antenne utilisée (celle du projet Alysound), est placée en (0 m, 0 m, 50 m), et sa projection sur la grille de calcul est représentée par le triangle inversé ∇ . a : carte de localisation par calcul de l'estimateur I_{loc} issu de la formation de voies, qui donne lieu à une carte normalisée. b : carte de quantification \hat{L}_p qui donne un niveau non biaisé (150 dB) pour la source en sa position (0 m, 50 m). c : carte de localisation par calcul de la fonctionnelle de MUSIC I_{MUSIC} , l'amplitude obtenue ne reflète pas la puissance de la source.

Les trois méthodes détaillées modélisent la propagation de manière commune à travers les vecteurs de pointage \mathbf{g} . Ceux-ci prennent une forme différente en fonction du modèle de propagation qui est en jeu : modèle en ondes sphériques ou en ondes planes. La partie suivante étend alors la notion de champ lointain défini localement dans le Chapitre 1 pour une antenne complète, afin de déterminer le modèle de propagation à sélectionner en fonction de la géométrie de chaque cas d'application.

Champ lointain en antennerie

Comme expliqué dans la partie Partie 1.3, pour une grande distance entre la source et un point d'observation par rapport à la taille caractéristique de cette source, les fronts d'ondes qu'elle génère peuvent être supposés localement comme plans. Cependant dans le cadre du traitement d'antenne, l'aspect local est à approfondir : une antenne, même placée extrêmement loin d'une source, pourra voir les fronts d'ondes comme non plans, pourvu qu'elle soit assez grande. Il est alors nécessaire de définir un critère, pouvant assurer que le front d'onde vu par tous les capteurs de l'antenne soit plan. Dans la littérature, ce critère provient d'études sur la stabilité des réponses des filtres spatiaux [Ken+98], mais son origine remonte à une règle empirique [Mai05]²⁷ et n'a pas été définie clairement par l'acoustique. Il est ainsi recommandé de considérer les fronts d'ondes comme plan pour des distances larges devant :

$$D_{lim} = 2L^2/\lambda, \quad (2.11)$$

avec L la distance entre les capteurs les plus éloignés de l'antenne, et λ la longueur d'onde étudiée. Pour vérifier la pertinence de ce critère, une simulation est menée selon la configuration décrite en Figure 2.4. Le signal temporel d'une source harmonique d'amplitude unitaire et de longueur d'onde fixée est propagé jusqu'à chacun des 30 capteurs d'une antenne linéaire de taille L , en y appliquant un retard équivalent au temps de parcours entre la source et le capteur considéré. Trois distances sont étudiées : $D_{lim}/2$, D_{lim} et $2D_{lim}$. La Figure 2.5 montre les signaux retardés pour chacun des capteurs en ordonnée en fonction du temps.

Les amplitudes sont données par l'intensité de couleur, qui illustre les oscillations de la source harmonique mais avec une amplitude inférieure à 1 : la divergence géométrique est prise en compte donc l'amplitude diminue lorsque la source s'éloigne. De la Figure 2.5a à la Figure 2.5c, les fronts d'ondes s'aplanissent, validant la représentation en ondes planes d'une source éloignée de l'antenne. A D_{lim} , la courbure des fronts est encore visible, ce qui valide le critère spécifié. De plus à $2D_{lim}$, la courbure commence à être négligeable, mais encore légèrement visible, ce qui montre bien que ce sont des distances grandes devant D_{lim} qui doivent être considérées pour valider la description d'une source monopolaire par un modèle en ondes planes.

Conclusion pour le traitement d'antenne

Cette première partie a permis de sélectionner les méthodes de traitement d'antenne adaptées aux applications expérimentales visées,

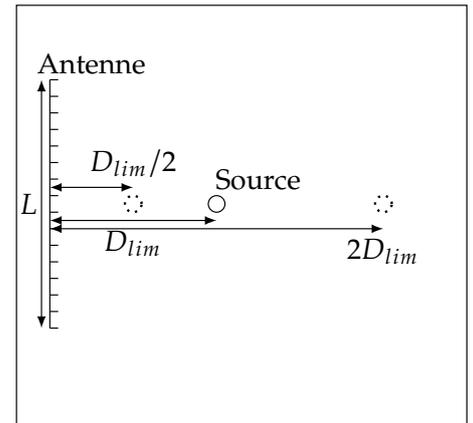


FIGURE 2.4 – Configuration utilisée pour la simulation de la réception de signaux temporels provenant d'une source monopolaire face à l'antenne. Trois distances différentes sont étudiées : $D_{lim}/2$, D_{lim} et $2D_{lim}$

27: Voir Chapitre 1.2

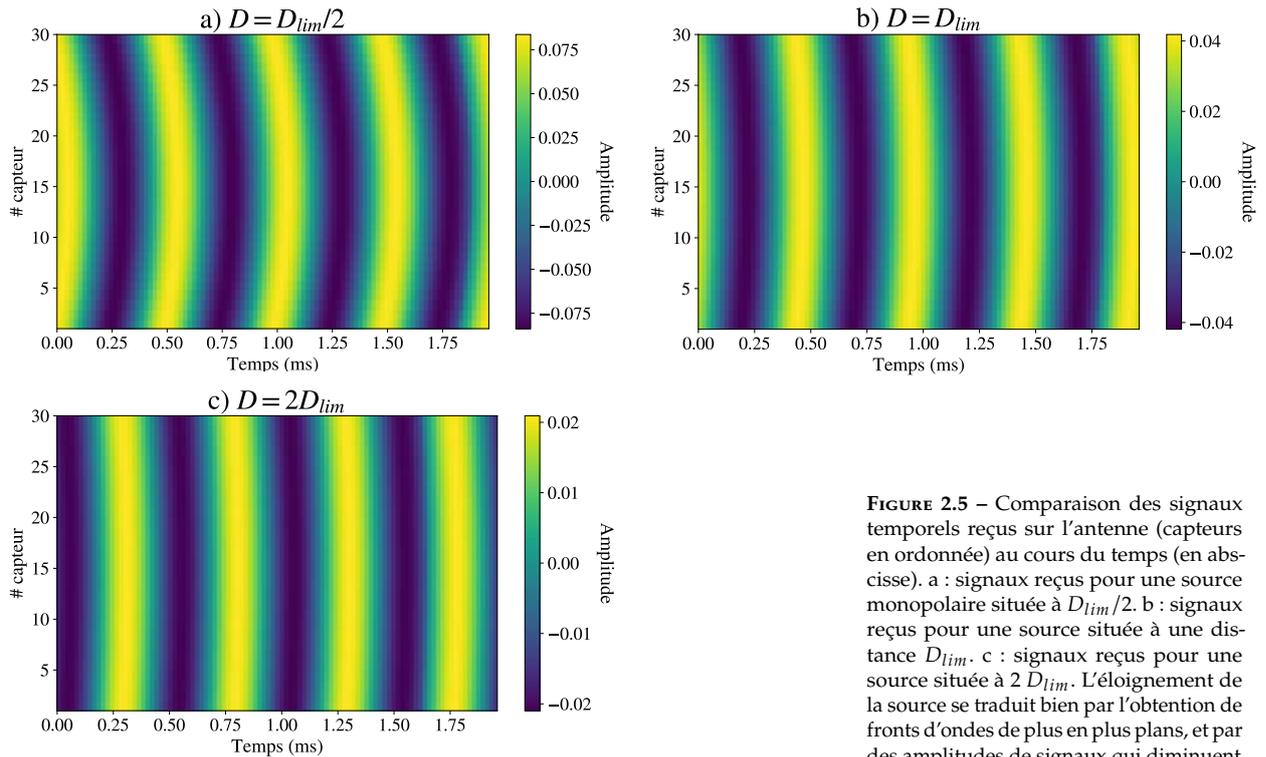


FIGURE 2.5 – Comparaison des signaux temporels reçus sur l’antenne (capteurs en ordonnée) au cours du temps (en abscisse). a : signaux reçus pour une source monopolaire située à $D_{lim}/2$. b : signaux reçus pour une source située à une distance D_{lim} . c : signaux reçus pour une source située à $2 D_{lim}$. L’éloignement de la source se traduit bien par l’obtention de fronts d’ondes de plus en plus plans, et par des amplitudes de signaux qui diminuent.

dont le but est de localiser et quantifier les sources acoustiques. La localisation est réalisée par la méthode MUSIC tandis que la formation de voies quantifie les sources trouvées.

En fonction de la distance entre l’antenne et la source, le modèle de propagation est à adapter : lorsque la distance entre la source et l’antenne est grande devant D_{lim} , un modèle en ondes planes est utilisé, sinon un modèle prenant en compte des ondes sphériques est privilégié.

Les méthodes de traitement d’antenne ne permettent pas de s’attaquer à l’identification de source acoustique. Pour pouvoir classer une source dans une catégorie donnée, les méthodes d’apprentissage statistique sont nécessaires. Une synthèse en est proposée dans la partie suivante pour sélectionner celles à utiliser dans les chapitres suivants.

2.3 Apprentissage statistique

L’apprentissage statistique, ou apprentissage machine (*machine learning*), est un ensemble de techniques dont l’objectif est de prédire le résultat d’un processus à partir de données d’entrée brutes sur lesquelles sont calculées certaines **caractéristiques** (*features*). Les **observations** sont alors définies comme la concaténation de ces caractéristiques pour un élément de la base de données d’entrée. La prédiction du résultat peut être quantitative ou catégorique, différenciant les méthodes de régression et de classification. Tout l’enjeu pour obtenir une bonne prédiction est alors de choisir les bonnes caractéristiques pour représenter le processus, mais aussi de sélectionner un modèle qui les associera fidèlement au résultat attendu en sortie. Lors de la première étape d’apprentissage,

Deux livres sont à la base de la rédaction de cette partie, [Dud+01] avec une introduction très pratique du problème de classification, et [Has+09] avec une approche plus tournée vers les statistiques.

les paramètres des modèles choisis sont déterminés grâce à des observations labélisées, i.e. dont le résultat est connu à l'avance. Les couples d'apprentissage (observation / label) servent ainsi à établir le modèle qui retrouve le plus précisément les quantités ou classes attendues. Puis lors de la phase de test, les nouvelles données d'entrée produisent des prédictions quantitatives (pour une régression) ou catégoriques (en classification), estimées à partir du modèle appris au préalable. Selon le degré de connaissance sur les couples d'apprentissage (observation / label) disponibles, trois grands types de méthodes peuvent être définis :

- ▶ **l'apprentissage supervisé** : dans ce cas ils sont connus, et il est donc possible de déterminer le modèle qui fonctionne le mieux à l'avance. L'enjeu de ces méthodes sera alors d'éviter le sur-apprentissage, c'est-à-dire un modèle extrêmement précis sur les données d'apprentissage, mais généralisant très mal les nouveaux résultats à prédire;
- ▶ **l'apprentissage non-supervisé** : ici les résultats des observations sont inconnus. Le but de ces méthodes est alors de trouver des groupes parmi les données d'entrée, la ressemblance entre les observations étant directement contrôlée par l'algorithme de regroupement sélectionné;
- ▶ **l'apprentissage par renforcement** : où les couples ne sont pas parfaitement connus, mais où le résultat obtenu par le modèle peut être déterminé comme bon ou mauvais. En donnant au modèle le retour de ce jugement, il peut être amélioré de façon itérative.

En acoustique, l'utilisation de l'apprentissage statistique se développe rapidement et s'attaque surtout aux cas supervisés et non-supervisés [Bia+19; Bar+15; Pot+94; Ace+09; Fis+17; Mic+95]. Dans cette thèse, l'objectif est d'identifier la classe d'un seul type de signal sonore, les signaux acoustiques de drone, de manière supervisée, i.e. les labels associés aux observations disponibles pour l'apprentissage sont connus. Deux manières d'envisager l'identification sont alors possibles :

Classification pour une seule classe d'intérêt

- ▶ par la **classification binaire**, pour laquelle deux classes sont à séparer : la classe cible, soit la classe drone dans l'application visée, et une autre classe appelée *autre bruit*, qui représente tout ce qui n'est pas la cible. Des données exhaustivement représentatives des deux classes sont alors requises pour la première phase d'apprentissage, afin de s'assurer que le modèle sera capable de généraliser lors de la phase de test, en présence de nouvelles données non présentes lors de la phase initiale.
- ▶ par la **classification à une classe**, pour laquelle seule une classe est apprise, la classe cible drone : en donnant uniquement des exemples de cette classe, la frontière qui les englobe au mieux est déterminé en rejetant les quelques observations considérées comme aberrantes. Tout l'enjeu de ce modèle est alors de comprendre comment sont organisées les données de la classe cible sans être trop concentré autour de celles-ci, pour être à la fois capable de rejeter des observations qui ne lui appartiennent pas sans en manquer des légèrement différentes qui devraient y

être classées.

Alors que la classification binaire permet de construire un modèle plus robuste si la variabilité des deux classes est bien prise en compte, la classification à une classe possède l'avantage de se concentrer sur les données de la classe drone, évitant ainsi des oublis dans la classe autre bruit composée d'une infinité de signaux acoustiques²⁸. Diverses méthodes existent pour chacune des façons d'aborder le problème d'identification avec une seule classe d'intérêt, et elles sont détaillées dans les parties suivantes.

28: Tous les sons qui peuvent être entendus sur terre, sauf ceux de drones.

Classification binaire

L'apprentissage supervisé requiert, dans le cadre de la classification, la connaissance de couples (observation / classe), notées (\mathbf{x}_i, y_i) , $i \in \llbracket 0, N_{app} \rrbracket$ avec N_{app} le nombre de couples sélectionnés pour l'apprentissage parmi les N_{obs} disponibles. Le vecteur \mathbf{x}_i , de taille $N_c \times 1$ avec N_c le nombre de caractéristiques utilisées, représente une observation. Le scalaire y_i représente lui la classe attendue pour cette observation, drone ou autre bruit par exemple, mais à qui une valeur de 0 au nombre de classes présentes est attribuée en pratique dans les algorithmes. Pour la suite, le cas binaire uniquement sera considéré, c'est-à-dire que le nombre de classes est égal à deux. L'objectif est alors de trouver un modèle f , basé sur un ensemble de paramètres donnés par le vecteur θ , qui va correctement prédire la classe y_i à partir de l'observation \mathbf{x}_i . La notion de prédiction correcte est rendue possible par l'introduction des fonctions de coût Fc_i ²⁹ qui vont mesurer la différence entre la prédiction faite par le modèle, et le résultat attendu. En sommant cette fonction de coût pour l'ensemble des N_{app} observations à disposition, on obtient une fonction de coût globale Fc ³⁰ sur la base de données d'apprentissage, qui dépend des observations données, et des paramètres du modèle θ qui intègrent notamment le nombre de couples d'apprentissage N_{app} .

$$29: Fc_i(f_{\theta}(\mathbf{x}_i), y_i) = \begin{cases} 0 & \text{si } y_i = f_{\theta}(\mathbf{x}_i) \\ 1 & \text{sinon} \end{cases}$$

par exemple pour de la classification

$$30: Fc(\theta) = \sum_{i=1}^{N_{app}} Fc_i(f_{\theta}(\mathbf{x}_i), y_i)$$

L'objectif des différentes méthodes disponibles est alors de minimiser cette fonction de coût par rapport aux paramètres du modèle, en cherchant à atteindre les performances optimales données par le classifieur de Bayes. Ce dernier n'est pas utilisable dans des cas pratiques, étant donné qu'une connaissance du modèle analytique entre les entrées et les sorties est nécessaire pour sa construction. Sans connaissance sur la distribution des données, des méthodes cherchent alors à les expliquer selon des lois gaussiennes, afin d'en déduire ensuite les meilleures frontières à tracer entre celles-ci, qu'elles soient linéaires ou quadratiques [Fis36]. La régression logistique, elle, cherche des séparations linéaires en se basant sur les distributions de probabilités postérieures [Cox58]. Pour s'abstraire du besoin de modéliser les distributions de données, des algorithmes dits *boîtes noires* ont vu le jour, comme par exemple les méthodes basées sur des arbres de décision, qu'ils soient utilisés seuls [Qui86] ou en groupes [Bre01]. Un autre exemple concerne les méthodes qui cherchent un hyperplan, c'est-à-dire une séparation linéaire en deux parties de l'espace des caractéristiques de dimension N_c , pour laquelle chacun des deux sous-espaces résultant contient les données d'une des deux classes. Le perceptron [Ros57] en est un exemple, mais il aboutit à différents hyperplans en fonction de l'initialisation. Pour contourner ce

problème la méthode SVM (*Support Vector Machine*) cherche quant à elle à trouver l'hyperplan optimal pour séparer ces deux classes, s'assurant d'en trouver un unique [Cor+95].

Dans la communauté acoustique, le modèle SVM fait partie des modèles les plus couramment utilisés [Bia+19], notamment parce qu'il ne fait pas d'hypothèse sur les distributions des données d'entrée, et grâce à l'interprétabilité physique que ses caractéristiques d'entrée offrent. De plus, dans l'application expérimentale visée dans cette thèse, l'identification de drones, sa robustesse face aux valeurs aberrantes est bénéfique. Dans la chaîne de traitement mise en place, l'identification se fait sur des durées de traitement courtes (200 ms) au cours du temps³¹. Or dans des conditions expérimentales variables, un certain nombre de ces observations ne sera pas représentatif du bruit d'un drone, comme par exemple lorsque celui-ci chute au sol, ou lorsqu'il passe trop près de l'antenne provoquant un bruit de souffle sur les capteurs. Un tri manuel de ces données n'étant pas envisageable au vu de leur nombre (150 observations chaque minute), la capacité qu'a le modèle SVM à produire un hyperplan robuste malgré ces quelques valeurs d'apprentissage erronées, comme illustré en Figure 2.6, conforte ce choix. Cette méthode est donc détaillée mathématiquement pour comprendre les différents paramètres qui agissent sur le modèle.

C'est en construisant un hyperplan à partir de valeurs particulières parmi tous les couples N_{app} mis à disposition pendant la phase d'apprentissage que la méthode SVM n'est pas affectée par les observations aberrantes. Un hyperplan \mathcal{H} ³² de paramètres θ est défini par l'équation suivante :

$$f_{\theta}(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}^T \mathbf{x} = 0, \quad (2.12)$$

avec \mathbf{x} un vecteur de données de dimension N_c , β_0 l'ordonnée à l'origine, et $\boldsymbol{\beta}$ le coefficient directeur de l'hyperplan. Le vecteur $\boldsymbol{\beta}$ est alors perpendiculaire à l'hyperplan et la distance signée d_s d'un point de l'espace à l'hyperplan peut être calculée³³.

En classification binaire, la valeur de sortie y_i est choisie dans l'ensemble $\{-1, 1\}$ pour chaque vecteur d'entrée \mathbf{x}_i selon sa classe d'appartenance. En combinant cette valeur de sortie avec la distance signée, et pour un hyperplan parfaitement séparateur comme schématisé en Figure 2.6, l'inégalité suivante est valable pour tout couple (\mathbf{x}_i, y_i) de l'ensemble d'apprentissage :

$$\frac{1}{\|\boldsymbol{\beta}\|} y_i (\boldsymbol{\beta}^T \mathbf{x}_i + \beta_0) \geq 0. \quad (2.13)$$

La marge est alors définie comme la distance entre les couples les plus proches de l'hyperplan et l'hyperplan. Elle est notée \mathcal{M} et est atteinte pour quelques vecteurs de l'ensemble d'apprentissage. Ces vecteurs sont appelés les vecteurs supports (indités v_s et situés sur les lignes pointillées en Figure 2.6), et ce sont eux seuls qui décident de la forme de l'hyperplan. Il est possible de normaliser $\boldsymbol{\beta}$ et β_0 dans l'équation Equation 2.13 afin de faire en sorte que $y_{v_s} (\boldsymbol{\beta}^T \mathbf{x}_{v_s} + \beta_0) = 1$ pour les vecteurs supports. Cette normalisation implique alors que $\mathcal{M} = 1/\|\boldsymbol{\beta}\|$, et le problème de

31: Plus de détails sont donnés dans le Chapitre 4

32: Plus rigoureusement, la définition donnée ici est celle d'un ensemble affine (car il ne passe pas par le vecteur nul de l'espace vectoriel). Pour simplifier la discussion, le terme hyperplan est utilisé sans distinction pour les ensembles passant par 0 ou non.

$$33: d_s(\mathbf{x}, \mathcal{H}) = \frac{\boldsymbol{\beta}^T \mathbf{x} + \beta_0}{\|\boldsymbol{\beta}\|}$$

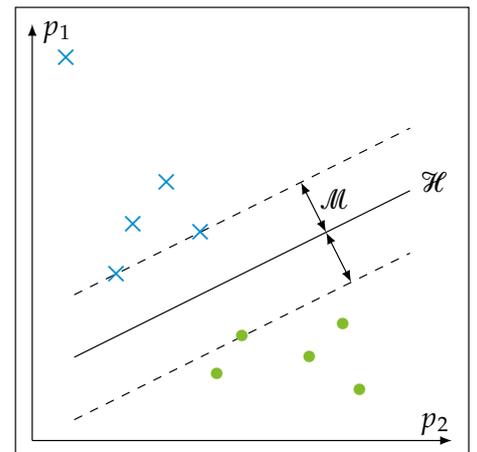


FIGURE 2.6 – Illustration d'un problème de classification à deux classes dans un espace de dimension $N_c = 2$. Le cas est ici séparable, c'est-à-dire qu'il est possible de trouver un hyperplan (en deux dimensions une droite) qui sépare parfaitement les deux classes.

maximisation de la marge est ainsi équivalent au problème d'optimisation quadratique sous-contraintes suivant :

$$\min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2, \quad (2.14)$$

sous contraintes $y_i(\beta^T \mathbf{x}_i + \beta_0) \geq 1$, pour $i \in \llbracket 1, N_{app} \rrbracket$.

Ce problème d'optimisation convexe se résout en introduisant les multiplicateurs de Lagrange α_i :

$$L(\beta, \beta_0, \alpha) = \frac{1}{2} \|\beta\|^2 - \sum_{i=1}^{N_{app}} \alpha_i [y_i(\beta^T \mathbf{x}_i + \beta_0) - 1]. \quad (2.15)$$

La première étape de la résolution est de satisfaire aux conditions de Kuhn-Tucker du premier ordre en égalisant les dérivées partielles à 0 pour obtenir :

$$\beta = \sum_{i=1}^{N_{app}} \alpha_i y_i \mathbf{x}_i, \quad (2.16)$$

$$\sum_{i=1}^{N_{app}} \alpha_i y_i = 0. \quad (2.17)$$

En substituant ces expressions, l'expression duale L_d à maximiser est obtenue :

$$L_d(\alpha) = \sum_{i=1}^{N_{app}} \alpha_i - \frac{1}{2} \sum_{i=1}^{N_{app}} \sum_{k=1}^{N_{app}} \alpha_i \alpha_k y_i y_k \mathbf{x}_i^T \mathbf{x}_k, \quad (2.18)$$

sous contraintes $\alpha_i \geq 0$, pour $i \in \llbracket 1, N_{app} \rrbracket$.

Sa maximisation représente un problème d'optimisation convexe simple, et des méthodes numériques standards le résolvent. Les solutions trouvées doivent satisfaire à une dernière condition de Kuhn-Tucker :

$$\alpha_i [y_i(\beta^T \mathbf{x}_i + \beta_0) - 1] = 0, \forall i \in \llbracket 1, N_{app} \rrbracket. \quad (2.19)$$

Deux cas sont alors possibles :

- si $\alpha_i > 0$, alors $y_i(\beta^T \mathbf{x}_i + \beta_0) = 1$, donc le vecteur \mathbf{x}_i considéré est un vecteur support \mathbf{x}_{vs} .
- si $y_i(\beta^T \mathbf{x}_i + \beta_0) > 1$, le vecteur considéré a une distance à l'hyperplan supérieure à la marge et $\alpha_i = 0$.

Cette condition démontre que la construction de l'hyperplan à marge maximale se base uniquement sur les vecteurs supports, et cet hyperplan sépare alors l'espace en deux parties. La capacité de la méthode à rejeter les valeurs aberrantes est ainsi vérifiée, avec comme exemple dans la Figure 2.6, la croix bleue en haut à gauche qui représente mal les valeurs

de la classe croix bleue qui n'est pas du tout prise en compte dans le choix de l'hyperplan et de la marge.

Mais la construction telle que décrite jusqu'à maintenant possède deux principaux défauts : elle ne permet pas de considérer le cas non-séparable, et elle ne peut pas séparer des données dont la frontière entre elles serait non linéaire.

Pour le cas non-séparable représenté en Figure 2.7, la solution est de relâcher les contraintes sur la marge en introduisant les variables de détente $\xi_i > 0$. Ainsi, le problème d'optimisation de départ donné à l'équation Equation 2.14 se transforme en :

$$\min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^{N_{app}} \xi_i, \quad (2.20)$$

sous contraintes $y_i(\beta^T x_i + \beta_0) \geq 1 - \xi_i$, pour $i \in \llbracket 1, N_{app} \rrbracket$,

où C est un paramètre à optimiser, qui contrôle la perméabilité des vecteurs de support par rapport à la marge maximale en construction. Le passage au problème dual en substituant les conditions de Kuhn-Tucker au premier ordre amène à la même équation que l'équation Equation 2.18, avec pour seul changement $0 \leq \alpha_i \leq C$: les multiplicateurs de Lagrange sont bornés par le coefficient C . Ce problème d'optimisation quadratique sous contraintes peut à nouveau être résolu par des techniques standards, et l'hyperplan obtenu est à nouveau calculé à partir des vecteurs supports, qui sont de deux types dans ce cas non-séparables :

- ▶ ceux qui sont exactement sur la marge (pour lesquels $\xi_i = 0$) et qui voient leur multiplicateur de Lagrange associé strictement borné $0 < \alpha_i < C$.
- ▶ ceux qui sont à l'intérieur de celle-ci ($\xi_i > 0$) pour lesquels $\alpha_i = C$.

Le coefficient C contrôle ainsi la confiance accordée aux données d'entrée. Un coefficient C qui tend vers 0 va permettre de prendre beaucoup de vecteurs de support du mauvais côté de la marge (car leur impact dans la minimisation de l'Equation 2.20 sera très faible) et donc va provoquer un risque de sur-apprentissage³⁴. Au contraire si C tend vers l'infini, aucun vecteur support ne sera autorisé du mauvais côté de la marge (sinon la minimisation de l'Equation 2.20 est impossible), et le problème d'optimisation tend vers le cas séparable.

Bien que ce relâchement permette d'assouplir les règles strictes de départ, la contrainte de linéarité n'est toujours pas levée avec le coefficient C : pour une frontière entre deux classes non linéaire comme illustré en Figure 2.8a, aucun hyperplan linéaire ne sera capable de séparer efficacement les données. Pour s'abstraire de ce problème, une dernière amélioration a été apportée à la méthode grâce à l'introduction d'un noyau.

Cette technique repose sur l'idée que des données non-séparables dans un espace de dimension donnée N_c le seront dans un espace élargi de dimension N_+ . Le but est alors de trouver des fonctions de base $\mathbf{h}(x_i) = (h_1(x_i), h_2(x_i), \dots, h_{N_+}(x_i))$ pour passer de l'espace de départ à ce nouvel espace élargi. Une représentation de cette séparation obtenue

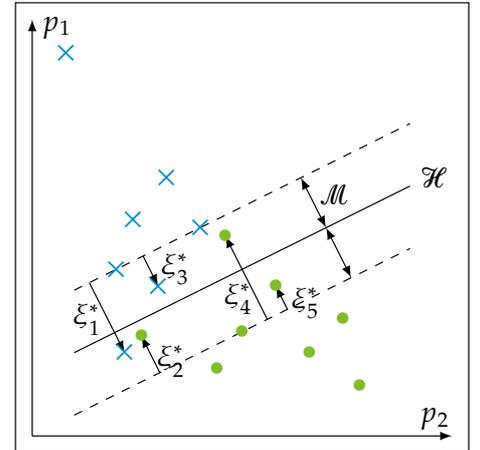


FIGURE 2.7 – Illustration d'un problème de classification à deux classes dans un espace de dimension $N_c = 2$. Le cas est ici non-séparable, c'est-à-dire qu'il est impossible de trouver un hyperplan (en deux dimensions une droite) qui sépare parfaitement les deux classes. Certaines observations se retrouvent du mauvais côté de leur marge, d'une quantité $\xi_i^* = M\xi_i$

34: Le modèle sera très bien adapté aux données d'apprentissage, mais tellement particulier qu'il ne généralisera pas.

dans l'espace élargi est donnée en Figure 2.8b : un hyperplan \mathcal{H}_+ peut désormais séparer linéairement les données des deux classes dans ce nouvel espace.

Le problème est qu'en pratique, ces bases peuvent être très grandes, voire potentiellement infinies, et le passage de l'espace de départ à celui-ci s'avère donc très coûteux. Cependant, lors du passage du problème primal au problème dual, les observations d'entrée \mathbf{x}_i ne sont plus utilisées que lors d'un produit scalaire $\mathbf{x}_i^T \mathbf{x}_k = \langle \mathbf{x}_i | \mathbf{x}_k \rangle_{N_c}$ dans le calcul du dual en Equation 2.18³⁵.

En notant $\langle \cdot | \cdot \rangle_{N_+}$ le produit scalaire dans l'espace élargi, le dual dans cet espace est alors obtenu directement en remplaçant le produit scalaire de départ par $\langle \mathbf{h}(\mathbf{x}_i) | \mathbf{h}(\mathbf{x}_k) \rangle_{N_+}$. La technique du noyau consiste alors à ne calculer que la valeur du noyau $N(\mathbf{x}_k, \mathbf{x}_l) = \langle \mathbf{h}(\mathbf{x}_k) | \mathbf{h}(\mathbf{x}_l) \rangle_{N_+}$ plutôt que de passer effectivement les données dans l'espace élargi avant d'en calculer le produit scalaire : la charge de calcul s'en retrouve fortement réduite. Un exemple de noyau largement utilisé est la fonction à base radiale³⁶ dont le paramètre γ doit être réglé. La résolution du problème d'optimisation convexe qui s'ensuit suit la même procédure que décrite précédemment, le changement de produit scalaire ne changeant pas la nature du problème à résoudre.

Suite à tous ces ajouts, la méthode SVM est armée pour répondre à des problèmes de classification non-linéaires complexes. Néanmoins, le choix de la fonction noyau reste largement fait empiriquement, sans explication mathématique claire sur la validité de ce choix même si la validation croisée permet de s'assurer de son bon fonctionnement sur les données en cours de traitement. C'est en ce sens que la méthode est vue encore comme une *boîte noire*, même si des travaux mathématiques tentent désormais d'expliquer leur fonctionnement [Mal16; Ben+00; Bir+17], ce qui fait écho au besoin de clarification qui se fait sentir dans le domaine [Mil19]. Cependant, en restant dans le cas linéaire, une grande interprétabilité des résultats de l'optimisation reste possible, et est d'un intérêt certain lorsqu'il s'agit de sélectionner des caractéristiques au profit d'autres afin d'améliorer les performances de classification.

La principale limite de la méthode dans l'objectif de n'identifier qu'un seul type de signal acoustique réside dans la construction de la base de données pour la classe concurrente, qui doit représenter toute sa variabilité. Pour s'en passer, la classification à une classe n'apprend sa frontière de décision qu'à partir d'exemples de la classe d'intérêt. Les méthodes qui le permettent sont analysées dans la partie suivante.

Classification à une classe

Avant d'avoir été formalisé par des travaux de thèse [Tax01], le problème de classification à une classe a été étudié sous différents noms selon les cas d'applications abordés : la détection d'anomalie [Rit+97], la détection d'innovation [Bis94], ou encore l'apprentissage de concept (*concept learning*) [Jap99]. Toutes ces études ont en commun de travailler avec un grand nombre d'observations d'une classe cible, et peu voire pas d'observations qui correspondent au reste des événements possibles.

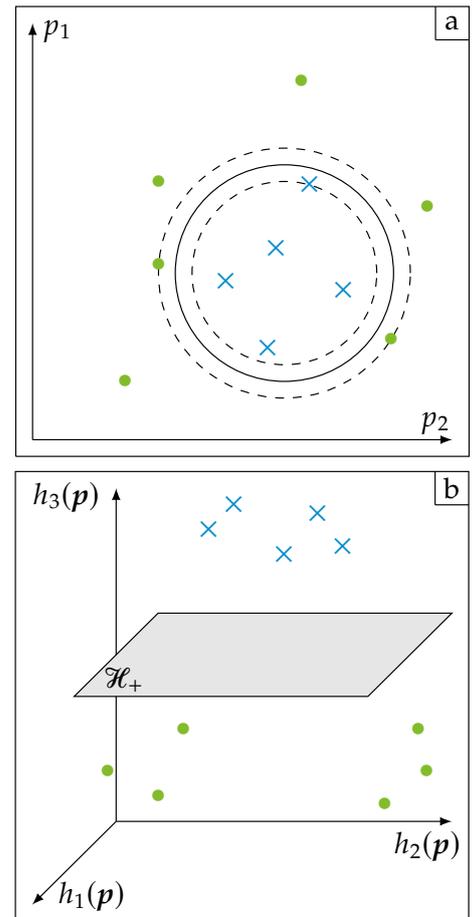


FIGURE 2.8 – Illustration d'un problème de classification à deux classes dans un espace de dimension $N_c = 2$. Le cas est ici non-linéaire, c'est-à-dire qu'il est possible de trouver un hyperplan qui sépare les deux classes, mais celui-ci est non linéaire. a : représentation des observations dans l'espace des caractéristiques (espace de départ). b : représentation des observations dans l'espace élargi. Une frontière linéaire est capable de séparer les observations des deux classes dans l'espace élargi.

35: Avec $\langle \cdot | \cdot \rangle_{N_c}$ le produit scalaire dans l'espace des caractéristiques.

36: $N(\mathbf{x}_k, \mathbf{x}_l) = \exp(-\gamma \|\mathbf{x}_k - \mathbf{x}_l\|^2)$

En synthétisant les taxonomies proposées dans plusieurs revues de la littérature [Kha+14; Tax01; Maz06; Bar11], trois grandes familles de méthodes peuvent être proposées pour traiter un tel problème de classification :

- ▶ celles qui cherchent à estimer la **densités de probabilités** de la classe cible, afin de déterminer ensuite si de nouvelles observations à classer en font partie ou non. C'est le cas par exemple des chaînes de Markov cachées [Cho+03; Yeu+03; Lan+03], des mixtures de gaussiennes [Ye+01; Yam+04], de l'estimation de densité de Parzen [Yeu+02] ou encore de l'algorithme des k plus proches voisins [Cab+09; Cab+07; Di +07; Di +08];
- ▶ celles qui cherchent à construire un **modèle simplifié** des données de la classe cible. Une fois ce modèle obtenu, ses paramètres pour de nouvelles observations sont déterminés et comparés à ceux qui décrivent fidèlement la classe cible pour établir son appartenance à celle-ci ou non. Les algorithmes k -moyennes (*k-means*) [Hol+99; Zan+04] et l'analyse en composantes principales [Shy+03] peuvent être cités ici;
- ▶ enfin les méthodes qui cherchent à séparer les données cibles du reste grâce à la construction d'une **frontière de décision**. C'est le placement par rapport à cette frontière qui décide ensuite de la détection d'une nouvelle observation comme cible ou non. Une première approche consiste à construire une hypersphère autour des données cibles [Tax+01], quand une autre basée sur la méthode SVM cherche à maximiser la distance entre l'origine de l'espace des caractéristiques et ces mêmes données [Sch+01]. Les réseaux de neurones [Ska03] et les méthodes basées sur des arbres de décision [Den+05; Dés+12] font eux aussi partie cette famille.

Afin de combiner les développements et de comparer deux algorithmes aux fonctionnements proches, la méthode basée sur l'algorithme SVM [Sch+01] est sélectionnée dans cette thèse. Elle est communément appelée SVM à une classe (SVM-1C), et sa formulation est très proche de celle développée dans la partie précédente pour le cas non séparable de l'Equation 2.20 :

$$\min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2 + \frac{1}{\nu N_{app}} \sum_{i=1}^{N_{app}} \xi_i - \rho \quad (2.21)$$

sous contraintes $(\beta^T x_i + \beta_0) \geq \rho - \xi_i$, pour $i \in \llbracket 1, N_{app} \rrbracket$

La différence réside dans l'introduction de la distance de l'hyperplan à l'origine ρ , qui est à maximiser.³⁷ Les sorties y_i ne sont plus considérées dans la contrainte, ce qui traduit le fait que seules les observations correspondant à la classe cible ont été conservées. La contrainte reflète le fait que la distance d'un point à l'hyperplan doit être plus grande que la distance de l'hyperplan à l'origine (moins la variable de détente pour assouplir la contrainte). La proportion d'observations à considérer comme *anormales* est alors contrôlée directement par le paramètre $\nu \in [0, 1]$.³⁸

Dans le cas linéaire, l'hyperplan recherché n'est pas très performant et dépend fortement de la position de départ des observations dans

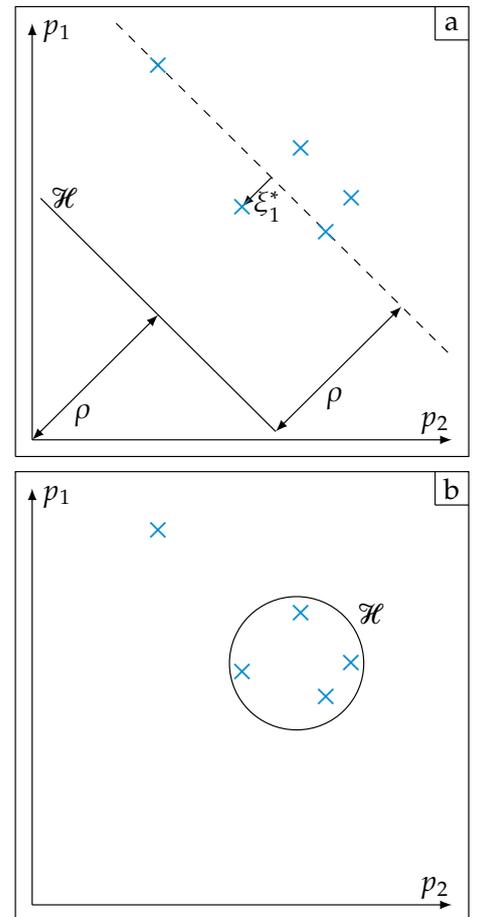


FIGURE 2.9 – Illustration d'un problème de détection d'anomalie dans un espace de dimension $N_c = 2$ pour $\nu = 0.2$ (soit une observation sur cinq à considérer comme anormale). a : cas linéaire, la distance entre l'hyperplan et l'origine est maximisée, sans réussir à isoler une observation. b : cas non-linéaire, grâce à l'introduction d'un noyau, une frontière non-linéaire peut être obtenue dans l'espace des caractéristiques qui permet d'isoler la valeur anormale. $\xi_i^* = \rho \xi_i$.

37: Ce qui est équivalent à minimiser $-\rho$ ce qui est fait en pratique.

38: Si $\nu = 0.1$, 10 % des observations seront classées en dehors de l'hyperplan.

l'espace des caractéristiques comme montré en Figure 2.9a. Cette méthode montre tout son potentiel grâce aux noyaux. C'est en cherchant la séparation entre les données et l'origine dans un espace élargi qu'il sera alors possible d'obtenir dans l'espace de départ une frontière non-linéaire comme illustré en Figure 2.9b. Pour des noyaux qui ne dépendent que de la différence $\mathbf{x}_i - \mathbf{x}_k$, des hypersphères sont retrouvées.

Dans cette thèse, une limite nette est tracée entre la classification binaire et la classification à une classe : dans le premier cas toutes les données sont labélisées alors que dans le second aucune. Mais dans la littérature, cette limite n'est pas aussi marquée. Des méthodes de classification à une classe permettant de prendre en compte une partie de données labélisées existent, afin d'espérer construire une frontière de décision moins encline au sur-apprentissage. Cette thèse n'étudie pas ces modèles et se concentre plutôt sur la comparaison des deux cas extrêmes, pour déterminer leurs avantages et inconvénients.

2.4 Conclusion

Ce chapitre a permis de détailler les outils de traitement du signal qui seront utilisés dans les chapitres applicatifs suivants, mais aussi de les placer au sein de la littérature afin de comprendre pourquoi ils étaient pertinents.

En premier lieu, le traitement d'antenne permet de générer des cartes de localisation de trois types différents : des cartes de formation de voies normalisées, d'autres quantitatives et des cartes de localisation à haute-résolution grâce à l'algorithme MUSIC. Les premières serviront à concevoir une antenne dans le chapitre suivant car le fait qu'elles soient normalisées offre la possibilité de comparer les performances de différents arrangements de capteurs entre eux. Le deuxième type servira à la quantification en un point donné de l'espace afin d'obtenir la quantification de sources acoustiques sous-marines. Enfin les cartes de MUSIC seront utilisées pour la localisation de toutes les sources acoustiques des deux chapitres suivants. Ses capacités de résolution pour deux sources spatialement proches ou de niveaux sonores fortement différents seront précieuses pour permettre des localisations correctes dans des environnements parfois très bruités.

En second lieu, les performances des deux modèles d'apprentissage statistiques sélectionnés, le SVM et le SVM-1C, seront comparés dans le cadre de l'identification de drone. Leur capacité de généralisation et leur robustesse face à des sources perturbatrices seront établies pour conclure sur les avantages et inconvénients de chacun.

Pour conclure, il est important de noter que dans les applications expérimentales, c'est l'enchaînement des outils présentés ici qui permettra d'établir des résultats innovants. En acoustique sous-marine dans le Chapitre 3, c'est une combinaison entre les différentes méthodes d'imagerie acoustique qui permettra d'établir la quantification par grand fond, jamais établie jusqu'alors, grâce à un système d'écoute spécifiquement conçu pour cette application.

Pour la localisation et l'identification de drone en Chapitre 4, la combinaison du traitement d'antenne et de l'apprentissage statistique permet une identification précise des classes en présence grâce à un filtrage spatial préalable. Des configurations avec des sources acoustiques perturbatrices sont même traitées, alors que l'utilisation de l'apprentissage statistique seul sur ce cas est très limitée.

Localisation et quantification de source en prospection minière sous-marine

3

Dans ce chapitre, une première caractérisation de sources acoustiques est effectuée dans un cadre spécifique lié à la prospection minière sous-marine. Cette caractérisation a pour objectif la **localisation** et la **quantification** du niveau sonore de machines excavatrices qui aplanissent et collectent les minerais situés sur le plancher océanique. Grâce à ces premières informations, l'objectif du projet de recherche Abysound dans lequel s'insèrent les travaux de cette thèse est de déterminer l'impact acoustique de ces machines sur leur environnement proche et lointain. Un système d'écoute est conçu en accord avec les contraintes de sécurité d'un site minier et les difficultés de déploiement. Puis les capacités en localisation et en quantification d'un prototype construit sont validées grâce à trois configurations expérimentales, issues de deux campagnes différentes effectuées en 2018 et en 2019 en mer Méditerranée. Une première configuration met en œuvre une source harmonique mobile référencée par des mesures en bassin. Une autre dispose de deux sources proches l'une de l'autre, une harmonique et une large bande, ce qui permet de mettre en évidence le bon fonctionnement du système d'écoute et de justifier le choix des méthodes dans des cas d'utilisation réels. Enfin, la dernière utilise une source large bande située à 680 m de fond et valide le fonctionnement du prototype dans des conditions similaires au grand fond.

Table des matières

3.1	Contexte	42
3.2	Conception du système d'écoute	44
3.3	Essais expérimentaux en mer	53
	Configuration 1 : deux sources en petit fond	54
	Configuration 2 : antenne en grand fond et source référencée en petit fond	54
	Configuration 3 : antenne et source à large bande en grand fond	55
	Références pour la localisation et la quantification	55
3.4	Validation du système d'écoute	56
3.5	Apport de la localisation à haute résolution	60
3.6	Validation en situation réelle par grand fond	63
3.7	Conclusion	67

3.1 Contexte

Après avoir connu plusieurs phases de développement plus ou moins marquées [Gla00], l'extraction de minerai par grand fond¹ devient depuis moins de cinq ans un moyen évoqué pour alimenter le marché des terres rares en expansion constante [Pit18]. Cette expansion répond à la hausse de production attendue des usines du monde entier, qui nécessite pour être satisfaite toujours plus de matières premières que les gisements terrestres ne parviennent plus à produire en quantité suffisante. Ceci explique l'intérêt croissant des industriels pour des gisements complexes et inexplorés à exploiter, situés sur les planchers océaniques, mais qui contiennent une grande quantité de nodules polymétalliques, riches de nombreux métaux nécessaires au développement des nouvelles technologies comme le cuivre, le cobalt ou encore les terres rares. Cet intérêt se traduit par le démarrage de différents projets à travers le globe [Chi+20; Min; New17], mais qui se heurtent à des résistances locales, nationales et internationales pour des considérations environnementales [Dee+19; The19] : ces activités sont situées dans des zones riches en biodiversité marine [Kai+17; Amo+16], et elles détruisent en grande partie ces habitats que ce soit par l'excavation elle-même, ou par les nombreux dépôts qui en résultent [Hal+07]. Une autre pollution moins étudiée par la littérature concerne l'impact acoustique de ces travaux. Bien qu'il existe quelques études menées par les entreprises minières elles-mêmes [Cof08]², celles-ci restent superficielles et cantonnées à l'impact des navires de surface qui collectent les minerais une fois récoltés grâce à des machines placées sur le sol, et remontés par l'intermédiaire d'une pompe placée dans l'eau, quelques dizaines de mètres au-dessus du plancher océanique. Malgré le principe de précaution qui devrait alors prévaloir, des projets sont lancés alors même que la littérature scientifique qui évalue les risques de ces travaux n'en est qu'à ses débuts [Chi+20].

Ces travaux de thèse s'inscrivent dans le cadre du projet Abyssound³ dont l'objectif est d'établir une procédure pour mesurer le bruit directement causé par ces machines excavatrices situées sur le plancher océanique, afin de produire des cartes de leur impact acoustique [Bas+19]. Les signaux acoustiques qu'elles génèrent sont d'abord mesurés par un système d'écoute conçu spécifiquement pour les localiser et les quantifier avant qu'un simulateur de propagation acoustique sous-marin [Cri+19] ne produise les cartes d'impact. Une fois générées, l'importance de la nuisance sur les espèces présentes aux alentours peut être étudiée pour en déduire si les niveaux sonores émis doivent être contrôlés.

En prenant comme exemple le projet Solwara 1 de Nautilus Minerals [Min] qui était le plus avancé dans le domaine, la Figure 3.1 illustre un site minier sous-marin. Le bateau (appelé *Production Support Vessel*) qui collecte les minerais est situé au-dessus de la zone exploitée. Sur le plancher océanique, trois types de machines excavatrices⁴ sont utilisées pour le mettre à niveau et collecter les minerais convoités. Ces machines sont reliées par des câbles à une pompe et une conduite, elle-même connectée au bateau pour y remonter les minerais. Toute l'activité est donc concentrée dans une colonne d'eau cylindrique qui part du plancher océanique pour aller jusqu'à la surface, dans laquelle les différents éléments sont mobiles.

1: Le terme grand fond est communément admis pour des profondeurs supérieures à 1000 m.

2: Ce rapport n'est plus disponible en ligne suite à la liquidation de la société Nautilus Minerals.

3: Une description du projet et de ses partenaires est donnée en Annexe A.

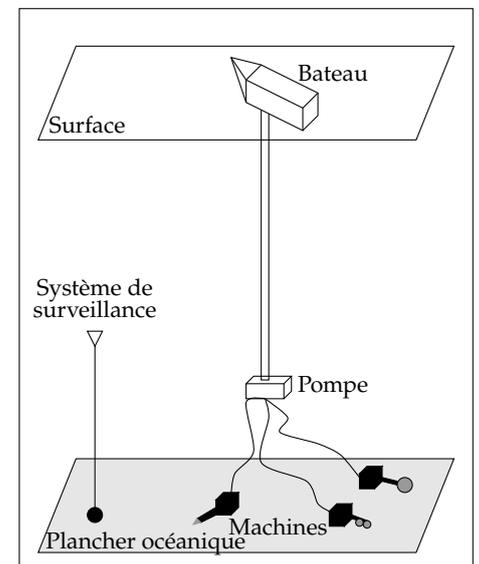


FIGURE 3.1 – Illustration de l'organisation d'un site minier sous-marin. Les machines situées sur le plancher océanique récupèrent les minerais, qui sont transférés au bateau collecteur grâce à la pompe et au système de conduite associé. Le système de surveillance acoustique est représenté par ∇ , et doit être placé au plus proche des machines à contrôler, tout en respectant les contraintes de sécurité.

4: Voir photos [ici](#).

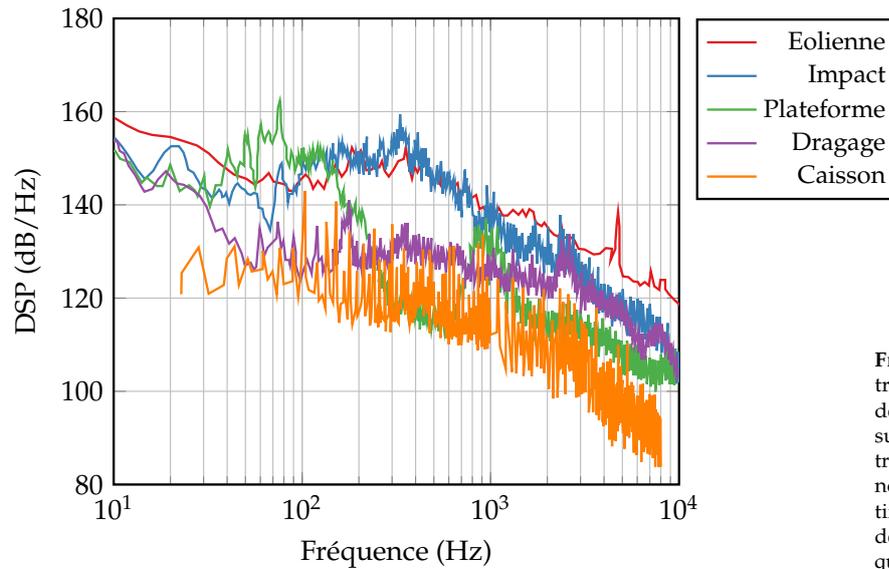


FIGURE 3.2 – Exemples de densités spectrales de puissance (DSP) de machines de forage et de dragage sous-marin issus de la littérature. Les machines excavatrices d'opérations minières sous-marines, non disponibles dans la littérature scientifique, devraient présenter le même type de spectre au vu des outils mécaniques qu'elles mettent en jeu.

Comme annoncé précédemment, l'acoustique des machines excavatrices à caractériser n'a pas encore été étudiée. Néanmoins, leur similitude mécanique avec des machines de forage ou de dragage permet d'établir une Densité Spectrale de Puissance (DSP) attendue, sans conclure précisément sur un niveau sonore final. Des mesures acoustiques sont disponibles dans le cadre de l'étude d'impact d'installations éoliennes en mer [Bro+14]. Le bruit de forage du plancher océanique peut y être caractérisé seul grâce au fait que la coque du bateau utilisé dans ces opérations est émergée grâce à des supports, l'empêchant de propager dans l'eau le bruit directement lié au bateau. D'autres mesures issues d'opérations de dragage dans le port de New-York ont donné lieu à une importante étude d'impact acoustique pour considérer les effets du bruit sur la faune qui y vit [Rei+14; Rei+12]. Les sources de bruit sont labélisées grâce à des enregistrements vidéos synchronisés avec les mesures acoustiques, et deux d'entre elles sont comparables au bruit de machines excavatrices : le bruit causé par la tête de la machine de dragage qui vient fracturer les roches, et le bruit des impacts au sol de la tractopelle lors de ramassage des roches. Des études d'impact acoustique ont aussi été menées autour de forages pétroliers en mer du Beaufort (Alaska) [Gre87; Bla+04]. Dans la première étude, des enregistrements sans bruit de bateau sont disponibles pour un caisson artificiel dédié au forage. Pour toutes les autres mesures, le forage est fait directement par des machines installées sur un bateau, ce qui empêche la caractérisation du bruit des machines seules. La deuxième étude est réalisée aux abords d'une plateforme pétrolière sans aucun bateau aux alentours.

Les densités spectrales des sources, récupérées à partir des contributions scientifiques citées, sont consignées en Figure 3.2. Elles montrent toutes un niveau constant à partir de 100 Hz, puis une partie décroissante qui ne démarre pas à la même fréquence pour toutes les sources, mais de pente commune de -20 dB/décade. Le bruit affecte toute la bande passante de l'audible jusqu'à 20 kHz, mais le spectre est principalement modifié jusqu'à environ 2 kHz. L'étude du forage pour les éoliennes en mer le montre en comparant des mesures avant et pendant le forage, pour lesquelles des tonales à 5 kHz sont déjà présentes alors que le

forage n'a pas débuté [Bro+14]. Le niveau constant n'est pas égal pour toutes les sources, et varie entre 130 et 150 dB/Hz, traduisant la diversité des environnements acoustiques présents⁵. Il est même montré qu'au sein du même environnement, les niveaux obtenus sont variables d'un enregistrement à l'autre [Rei+14; Bla+04]. En compilant ces études, le gabarit de source attendu pour une machine excavatrice est représenté en Figure 3.3 par un spectre large bande, entre 100 Hz et 20 kHz, démarré avec un niveau constant de 150 dB/Hz jusqu'à 500 Hz, suivi par une décroissance de pente -20 dB/décade. Des précisions sont à prévoir sur ce gabarit en fonction de l'avancée de la littérature dans le domaine, mais en pratique cette connaissance actuellement restreinte ne limite pas les méthodes mises en place qui sont indépendantes des niveaux observés.

Ce travail de thèse s'intéresse à la première étape nécessaire à l'étude d'impact de ces machines : leur **localisation** et la **quantification** de leur niveau de pression acoustique à 1 m. Afin de les caractériser au mieux en présence des autres sources de bruit que sont la pompe, la conduite de remontée et le bateau, un système d'écoute placé au plus près de la zone comme illustré en Figure 3.1, et respectant les contraintes de sécurité liées à cet environnement minier est un bon candidat. En partant des méthodes qui permettent de localiser et quantifier des sources acoustiques, la partie suivante aborde la conception d'un tel système d'écoute en montrant comment les contraintes en jeu réduisent les possibilités offertes, pour aboutir à une structure mécanique optimisée pour atteindre les objectifs fixés.

3.2 Conception du système d'écoute

Ayant pour objectif de localiser et quantifier des sources acoustiques, l'efficacité du système d'écoute à concevoir est jugée sur sa capacité à minimiser les erreurs de localisation et de quantification de ces sources. Pour assurer cette minimisation, le système doit être conçu pour permettre l'utilisation de méthodes qui répondent aux problématiques scientifiques de localisation et quantification, mais il doit aussi avoir une structure adaptée aux problématiques techniques que le contexte minier sous-marin induit. Cette partie répond à ces exigences scientifiques et techniques par le choix des méthodes d'imagerie acoustique à utiliser puis celui de la structure mécanique du système à construire.

Sélection des méthodes : en localisation acoustique sous-marine, les premières méthodes se développent durant la Seconde Guerre mondiale grâce aux progrès technologiques concrétisés par les mesures hydrophoniques [Las77; Vac98]. En combinant plusieurs hydrophones, des méthodes de traitement d'antenne ont été mises au point pour déterminer la localisation d'une source acoustique que ce soit par la formation de voies [Bur91]⁶, les méthodes adaptatives [Wid+67; Gri69; Cap69], ou haute-résolution [Pis73; Sch86; Red79; Bie+85]. Une autre possibilité consiste à estimer les retards entre les voies de mesures en exploitant leurs intercorrélations, pour les utiliser ensuite séparément pour la localisation [Car81; Kna+76; Kir+81; Fer89]. Enfin, en associant aux méthodes de traitement d'antenne des connaissances sur le canal de propagation comme son profil de célérité où les coefficients de réflexion de ses interfaces, la localisation est rendue plus précise grâce aux méthodes dites

5: Donnant même lieu pour certaines études [Rei+14; Gre87; Bla+04] à des coefficients de propagation trouvés par régression.

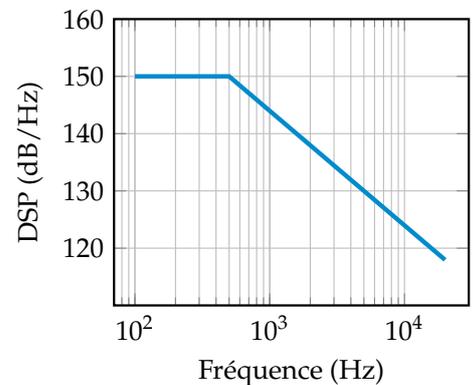


FIGURE 3.3 – Gabarit de la DSP du bruit supposé émis par une machine excavatrice. Il est composé d'une partie constante à 150 dB entre 100 et 500 Hz, suivie d'une décroissance de -20 dB/décade.

6: Voir chapitre 11.

de *Matched Field Processing* (MFP), sous condition que cette information supplémentaire décrive fidèlement le canal [Bag+88; Bag+93; Dos+13]. Toutes ces méthodes n'ont pas été systématiquement appliquées à des données réelles du fait de leur développement théorique antérieur à l'apparition de la puissance de calcul actuelle, mais des études comparatives sont tout de même disponibles pour les méthodes adaptatives et haute-résolution [Ger+95; Sch+85; Luc+85; Gao+94]. Dans le domaine du MFP les données réelles font partie intégrante du développement étant donné que les méthodes cherchent à s'adapter aux conditions réelles de propagation [Ham+89].

Historiquement, la quantification de sources acoustiques sous-marines est basée sur la mesure d'un hydrophone dont la distance à la source est connue, et dont le niveau est donné directement [Bro+14; Gre87]. Elle peut ensuite être rehaussée en compensant la perte de niveau en propagation par un modèle choisi selon la topographie de la zone, ou déduit des mesures [Rei+14; Gre87; Bla+04]. La quantification du bruit acoustique sous-marin n'est à ce jour pas systématique, même si l'intérêt qui lui est porté est grandissant au vu des études scientifiques qui lui sont consacrées [Ber+17a]. Des décisions politiques commencent d'ailleurs à être prises pour le réduire comme par exemple la Directive-Cadre Stratégie pour le Milieu Marin (DCSMM) [Par+08] dont une composante cherche à minimiser l'impact du bruit dans les mers au niveau européen. Le trafic maritime prend une part prépondérante de la globalité du bruit océanique. Par conséquent, les normes les plus récentes et les plus abouties liées à la quantification de sources acoustiques sous-marines concernent le bruit de passage des navires [ISO16; ISO19; ANS19]. Mais ces normes ne prennent en compte que la moyenne de signaux enregistrés par des hydrophones placés à des positions spécifiques [Gas+17], ce qui suppose que l'élément étudié est seul ou prépondérant dans l'environnement acoustique où sont placés les hydrophones. Le traitement d'antenne permet quant à lui une plus grande précision dans la quantification des bruits en séparant la contribution de différentes sources, comme par exemple les contributions de différentes parties d'un navire [Oud+18].

Aux abords de zones d'exploitation minière sous-marines, l'utilisation de certaines méthodes n'est pas envisageable à cause de deux conditions pratiques majeures :

- ▶ dans ces zones, plusieurs sources sont présentes de manière simultanée : les trois machines excavatrices et la pompe. Il est donc nécessaire de concevoir un système qui sera capable de séparer les contributions de chacune d'entre elles, que ce soit spatialement en les séparant grâce à une résolution suffisante, ou en niveau par le fait d'être capable de distinguer une source faible en présence d'une autre source plus forte;
- ▶ par grand fond, il est très difficile de garantir le placement correct des hydrophones pour respecter les distances demandées par les normes. De plus, les sources sont mobiles sur le plancher océanique, ce qui nécessite un nouveau placement des hydrophones à chaque mouvement des machines dans le cadre d'un contrôle continu de la zone d'exploitation en respectant les distances préconisées par les normes.

Ces particularités orientent le choix pour la localisation vers les méthodes à haute-résolution et en particulier la méthode MUSIC. Elle permet de localiser les sources qu'elles soient proches spatialement, ou de niveaux fortement distincts [Sch+85]. Elle est aussi plus précise que les méthodes adaptatives et s'abstrait des problèmes de rapport signal à bruit faible pouvant avoir lieu dans ce contexte industriel bruyant [Gao+94]. Sans connaissance précise sur le milieu de propagation et avec des distances entre les capteurs et les sources réduites, l'apport des méthodes MFP n'est pas significatif, et peut même nuire à la localisation en cas de mauvaise estimation des paramètres du milieu.

Ces deux conditions imposent ensuite l'utilisation du traitement d'antenne lors de la phase de quantification pour permettre de retrouver le niveau sonore de chaque source en présence séparément. Dans ce contexte expérimental inédit, le niveau de pression à 1 m est ainsi déterminé par calcul d'une formation de voie quantitative, faite pour la position du maximum de localisation estimé par la méthode haute-résolution MUSIC. Cette procédure est adaptée lorsque la source à quantifier est seule ou située loin d'autres sources acoustiques, ce qui est le cas lors des quantifications effectuées dans cette thèse. Il est néanmoins important de noter que pour traiter des cas sortant de ce cadre, les algorithmes de déconvolution qui prennent en compte la réponse de l'antenne, ou les méthodes inverses qui permettent la description de toutes les sources en présence de manière simultanée, devront être mis en place.

Choix de la procédure de conception : les méthodes sélectionnées impliquent un système d'écoute sous la forme d'une antenne d'hydrophones, dont la structure mécanique doit être conçue. Dans la littérature scientifique, la conception d'antenne est vue comme un processus d'optimisation qui cherche un compromis entre résolution spatiale faible et niveaux des lobes secondaires faibles. Ce problème d'optimisation multi-critères est résolu par des algorithmes heuristiques comme les algorithmes génétiques [Le 12; Hau94], ou le recuit simulé [Bai+10; Hol+01]. Les algorithmes de recherche Tabou simplifient ces méthodes en orientant l'optimisation vers les arrangements d'hydrophones les plus prometteurs [Mer+08], mais ils sont restreints à des géométries d'antennes linéaires. Or une antenne placée proche d'un site minier doit respecter plusieurs contraintes de sécurité. Compte tenu de la taille des câbles nécessaires à la bonne collecte des minerais [Spa+16; Yu+10], une distance de sécurité de 50 m doit être respectée entre l'antenne et la zone de travail. Elle doit aussi avoir une bonne tenue mécanique pour résister à des transports en bateau et des déploiements de plusieurs centaines de mètres de profondeur. Enfin, elle doit assurer des performances équivalentes autour de son axe vertical, car assurer un cap défini à l'avance lors du déploiement d'un système en mer entraîne un coup de déploiement prohibitif en grand fond. Pour résumer, la structure de l'antenne à concevoir doit respecter les contraintes suivantes :

Contraintes sur la structure de l'antenne à concevoir

- ▶ **Forme tri-dimensionnelle** : pour rendre possible la discrimination spatiale de sources acoustiques provenant de toutes les directions de l'espace, et ainsi déterminer les contributions des

différents éléments de la zone de travail.

- ▶ **Géométrie axisymétrique** autour de l'axe vertical : pour assurer des performances identiques quelle que soit la position finale de l'antenne une fois arrivée sur le plancher océanique.
- ▶ **Résolution de 20 m** : l'ordre de grandeur de la taille des machines excavatrices étant de 20 m, le système conçu doit être capable de séparer deux sources espacées de cette distance sur toute la bande de fréquence de travail, soit la bande de fréquence dans laquelle le son émis par ces machines est prépondérant [100 Hz ; 2 kHz]. La résolution étant liée directement à la taille de l'antenne pour la formation de voies [Eli97], cela implique une antenne de 10 m d'envergure pour atteindre 20 m de résolution moyenne sur la bande de fréquence [100 Hz ; 2 kHz], à 50 m de distance.
- ▶ **Des lobes secondaires inférieurs à -8 dB** : il est admis dans la communauté de l'imagerie acoustique que des niveaux de lobes secondaires inférieurs à -8 dB permettent une bonne interprétation physique des cartes obtenues. Ce niveau étant proportionnel au nombre d'hydrophones à disposition [Hol+01], un nombre de 48 hydrophones a été déterminé comme nécessaire afin de les atteindre en moyenne sur la gamme de fréquence d'intérêt [100 Hz ; 2 kHz].

Dans la littérature scientifique, il n'existe pas de système acoustique qui réponde à l'ensemble de ces contraintes à la fois. Un exemple d'antenne pour localiser des objets immergés à partir du bruit ambiant possède de nombreuses voies de mesures pour assurer des lobes secondaires bas mais il n'a pas été utilisé en grands fonds et il nécessite un placement précis à cause de sa géométrie plane [Pal+16]. Le système ALMA (*Acoustic Laboratory for Marine Applications*) de la Direction Générale de l'Armement (DGA) est plus modulable mais toujours limité à des utilisations en petit fond, et sans avoir à ce jour testé une géométrie tridimensionnelle d'antenne [Rea+18; Fat+19]. Un système complet d'observation dédié à la détection de neutrinos a été effectivement déposé en grand fond [Ric09], mais il ne possède que quatre hydrophones ce qui limite les performances atteignables en traitement d'antenne. Enfin un autre système déployé en grand fond exploite un système linéaire actif pour recenser la population benthopélagique à des profondeurs allant jusqu'à 5800 m [Ric+85; Smi+89], mais la géométrie linéaire empêche la localisation en cap d'une machine et travailler en actif ne permet pas d'en déterminer le niveau de bruit émis.

Il est donc nécessaire de créer une nouvelle antenne pour répondre aux contraintes particulières des sites miniers sous-marins. Le projet ne cherchant qu'à proposer un prototype d'un tel système, sa dimension sera réduite pour allier réduction de coût, et utilisation d'un bateau de recherche spécifique pour le déploiement de l'antenne lors de campagnes expérimentales. Grâce au consortium lié au projet Abysound, le navire océanographique *L'Europe* de l'IFREMER a été utilisé, mais il est limité en taille à un volume d'un cube de 3 m de côté pour le matériel sur la plage arrière.

Pour établir la forme de l'antenne et la position de ses hydrophones, une procédure d'optimisation doit être sélectionnée. Restreinte à des

géométries linéaires, la recherche Tabou est alors écartée. Puis l'ensemble des contraintes évoquées précédemment limitent les possibilités d'arrangements hydrophoniques, rendant les méthodes heuristiques trop coûteuses en temps de réglage, de mise en place et de calcul par rapport à un test exhaustif de tous les cas mécaniquement réalisables. Ce test exhaustif est alors privilégié pour concevoir l'antenne.

La méthode de conception proposée [Bar+20] compare alors la résolution et le niveau des lobes secondaires obtenus sur des cartes de formation de voies calculées avec le critère I_{loc} . Celles-ci sont obtenues pour des arrangements hydrophoniques techniquement réalisables, afin d'identifier ceux qui permettent d'atteindre le compromis entre résolution de 20 m et niveaux des lobes secondaires de -8 dB sur la gamme fréquentielle d'étude. Pour respecter dans un premier temps les contraintes techniques seules, puis établir la structure finale du prototype à construire, deux étapes sont suivies :

- ▶ une première étape pour déterminer la **géométrie de l'antenne** dans les conditions les plus proches possibles de l'application industrielle finale : antenne de 10 m avec 48 hydrophones pour localiser et quantifier sur la bande de fréquence [100 Hz ; 2 kHz].
- ▶ une seconde étape de **mise à l'échelle** de la forme obtenue puis d'optimisation du placement des hydrophones sur celle-ci, pour respecter les critères de coûts et d'encombrement physique inhérents au projet de recherche Abysound : antenne de 3 m avec 21 hydrophones. Pour s'approcher des résultats à l'échelle industrielle en résolution, la bande de fréquence est alors multipliée par 3,5 pour compenser la perte de taille et passe donc à [350 Hz, 7 kHz].

Sélection de la géométrie de l'antenne : pour la première étape, cinq antennes sont générées sur lesquelles les hydrophones sont placés de manière homogène et sans optimisation particulière, selon les géométries suivantes, car mécaniquement réalisables, et illustrées en Figure 3.4 :

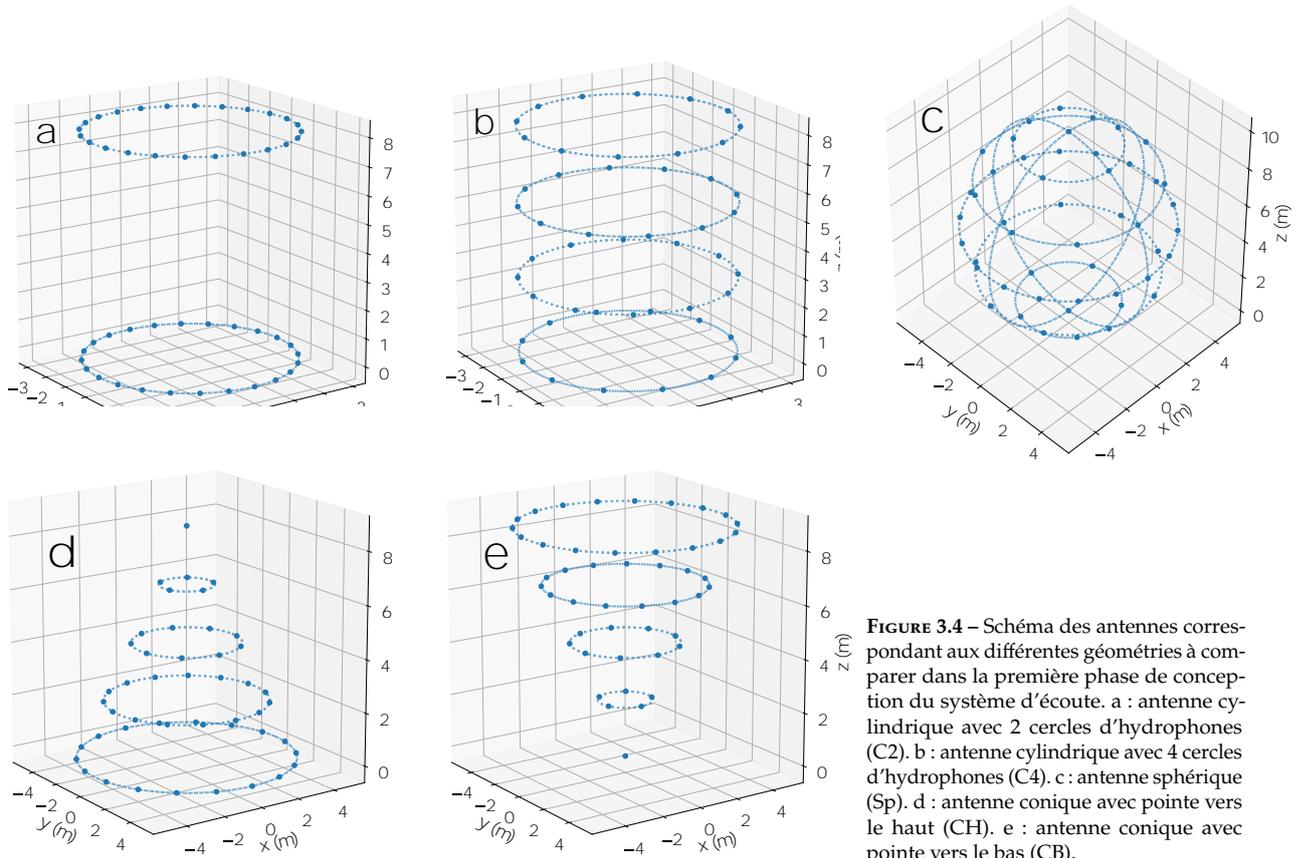


FIGURE 3.4 – Schéma des antennes correspondant aux différentes géométries à comparer dans la première phase de conception du système d'écoute. a : antenne cylindrique avec 2 cercles d'hydrophones (C2). b : antenne cylindrique avec 4 cercles d'hydrophones (C4). c : antenne sphérique (Sp). d : antenne conique avec pointe vers le haut (CH). e : antenne conique avec pointe vers le bas (CB).

- ▶ des cylindres constitués de deux et quatre cercles d'hydrophones ;
- ▶ des cônes avec la pointe vers le haut ou la pointe vers le bas ;
- ▶ une sphère.

Pour chaque antenne, une simulation est menée selon la configuration décrite en Figure 3.5, soit une antenne située à 50 m au-dessus du sol et une source placée sur le sol à 75 m au nord de l'antenne. Une carte de formation de voies est calculée, en utilisant le critère I_{loc} donné en Equation 2.5, pour chaque fréquence entre 100 Hz et 2 kHz par pas de 20 Hz. Le niveau maximum des lobes secondaires (NMLS) et la largeur du lobe principal à -3 dB (L_{-3dB}), décrits en Partie 2.2, sont déterminés pour chacune des cartes, puis moyennés sur les bandes de fréquence [160 Hz ; 2kHz] et [180 Hz ; 2 kHz], respectivement⁷. Finalement, la L_{-3dB} moyenne de chaque antenne est représentée en fonction du NMLS moyen [Sar16] pour aboutir à une comparaison synthétique des performances de chaque antenne, comme montré en Figure 3.6.

Les seuils de niveau pour les lobes secondaires et de résolution sont donnés par les traits pointillés rouges. Une antenne qui les respecte tous les deux se situe dans le cadran inférieur gauche de la figure, ayant en moyenne sur la bande de fréquence considérée des lobes secondaires inférieurs à -8 dB et une largeur de lobe principal inférieure à 20 m. Bien qu'aucune des géométries testées n'atteint ces contraintes en moyenne sur toute la bande de fréquence, les géométries cylindrique à deux cercles (C2) et conique avec pointe vers le haut (CH) en sont loin et elles sont donc écartées pour la suite. Entre les trois restantes, le choix doit se porter sur celle qui apporte une réponse d'antenne stable en fonction de la

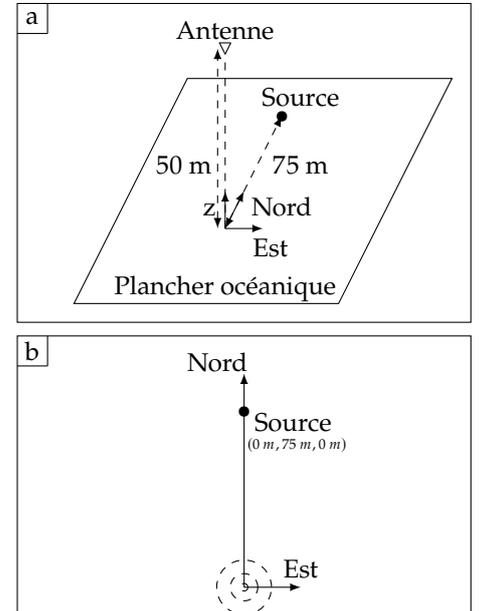


FIGURE 3.5 – Schéma de la configuration de simulation. La source simulée représentée par le ● est positionnée sur le sol, à 75 m au nord de l'antenne. a : vue isométrique, l'antenne est représentée par le triangle inversé ▽. b : vue de dessus, l'antenne est représentée par les cercles en pointillés.

⁷: En-dessous de ces fréquences, le lobe principal obtenu est trop étendu pour pouvoir calculer les valeurs.

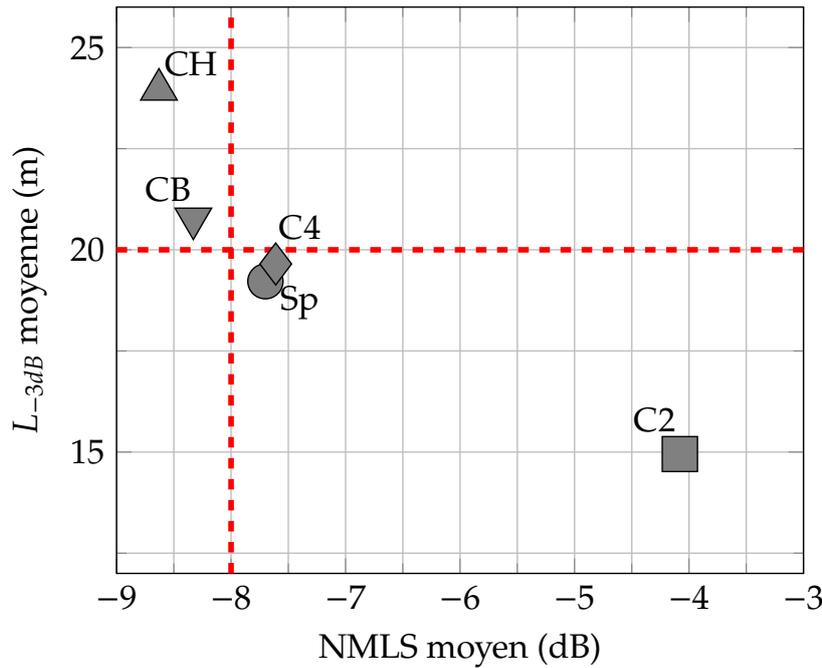


FIGURE 3.6 – L_{-3dB} moyenne en fonction du NMLS moyen pour les différentes géométries. Chaque marqueur représente une géométrie : cylindrique à deux cercles (C2), cylindrique à quatre cercles (C4), conique pointe vers le haut (CH), conique pointe vers le bas (CB) et sphérique (Sp). Les seuils de -8 dB pour le NMLS et de 20 m pour la résolution sont donnés par les traits pointillés rouges --.

fréquence, de sorte à éviter de mauvaises performances à des fréquences particulières qui ne se verraient pas dans la moyenne du NMLS.

La bande de fréquence est donc découpée en trois sous-bandes : basse fréquence de 180 Hz à 600 Hz, moyenne fréquence de 600 Hz à 1250 Hz, et haute fréquence de 1250 Hz à 2 kHz. La moyenne et l'écart-type du NMLS sont donnés pour ces trois bandes, et pour les trois géométries d'antenne, dans le Tableau 3.1.

Avec des moyennes supérieures à -8 dB pour chaque bande de fréquence, la géométrie cylindrique à quatre cercles (C4) est écartée. Entre les deux géométries restantes, les moyennes sont légèrement meilleures par bande de fréquence pour l'antenne conique pointe vers le bas (CB) par rapport à la sphérique (Sp), et ses écarts-types sont 1,4 et 0,4 dB plus faibles pour les bandes de fréquence basse et moyenne. Par conséquent, la géométrie retenue est conique avec la pointe vers le bas, représentée en Figure 3.4d.

Positionnement des hydrophones sur le cône : une fois la géométrie déterminée, le positionnement des hydrophones sur le prototype d'antenne à construire peut être optimisé. Pour être déployé par *L'Europe*, la hauteur de ce prototype est limitée à 2,7 m, et son diamètre maximum à 2,9 m. Le nombre d'hydrophone est fixé à 21 pour des raisons de coûts. Afin de concevoir une armature mécaniquement robuste sans modifier le champ acoustique à son voisinage, il n'est pas concevable de placer les hydrophones de manière aléatoire sur toute la surface conique de l'antenne, ce qui multiplierait les éléments de structure nécessaires à leur placement. Selon le même processus que pour la définition de la géométrie de l'antenne, 385 arrangements hydrophoniques sont alors créés, en respectant cette contrainte de faisabilité mécanique grâce aux critères suivants :

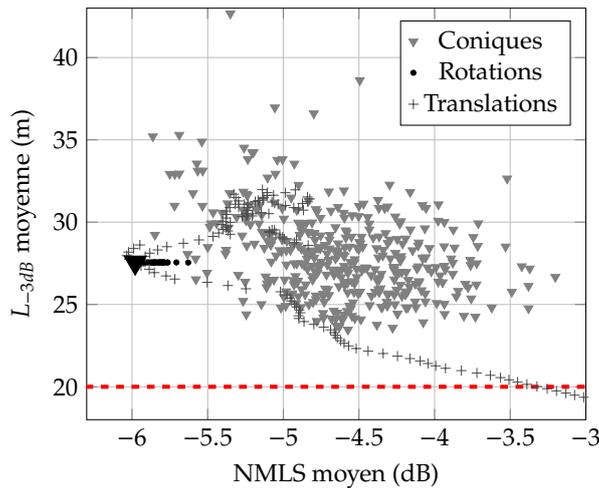
- l'antenne est formée d'un hydrophone à la pointe, et de 1, 2, 3, 4 ou 5 cercle(s) d'hydrophones ;

TABLEAU 3.1 – Moyennes et écarts-types du NMLS pour les trois antennes les plus proches des seuils, et pour les trois sous-bandes de fréquence : Basse [180 Hz ; 600 Hz], Moyenne [600 Hz ; 1250 Hz], et Haute [1250 Hz ; 2 kHz].

	Basse (dB)	Moyenne (dB)
Sp	-10.13 ± 1.70	-7.35 ± 1.22
C4	-7.91 ± 0.02	-7.61 ± 0.34
CB	-10.31 ± 0.26	-8.19 ± 0.80
	Haute (dB)	
Sp	-6.48 ± 0.74	
C4	-7.40 ± 0.38	
CB	-7.25 ± 0.88	

- ▶ chaque cercle est composé au minimum de trois hydrophones ;
- ▶ la distribution angulaire des hydrophones d'un cercle est uniforme ;
- ▶ la distribution des cercles est uniforme selon la hauteur de l'antenne.

Pour chaque arrangement, le processus décrit précédemment est reproduit, mais en calculant les moyennes du NMLS et de la L_{-3dB} sur la bande de fréquence [800 Hz ; 7 kHz]⁸. La L_{-3dB} moyenne en fonction du NMLS moyen est représentée en Figure 3.7, où chaque triangle inversé ▼ donne les performances d'un des 385 arrangements testés.



8: Elle correspond à nouveau à la bande de fréquence maximale sur laquelle les valeurs peuvent être calculées, mais pour la bande de fréquence mise à l'échelle pour le prototype donc multipliée par 3,5.

FIGURE 3.7 – L_{-3dB} moyenne en fonction du NMLS moyen pour les antennes coniques. Les 385 arrangements coniques de départ sont représentés par les marqueurs ▼. L'arrangement avec le plus faible NMLS moyen est donné par le triangle inversé noir ▼. Les points • correspondent aux arrangements obtenus par rotation des cercles de cet arrangement de départ, et les croix + correspondent à ceux obtenus par translation du cercle central de celui-ci. Le seuil de 20 m pour la résolution est donné par le trait pointillé rouge - - (le seuil pour le NMLS est trop faible pour être représenté).

Le NMLS minimum est de -6 dB pour l'ensemble des 385 arrangements. Cette valeur est liée au nombre d'hydrophones réduit à disposition pour le prototype qui entraîne une augmentation du niveau des lobes secondaires. La résolution augmente légèrement, avec une L_{-3dB} qui passe de 21 m pour l'antenne de 10 m à 27 m pour une antenne de 3 m. La mise à l'échelle de la bande de fréquence pour compenser la diminution de la taille du prototype permet de conserver les mêmes ordres de grandeur en résolution.

Pour réduire au maximum l'impact de l'augmentation des lobes secondaires dans la lisibilité des cartes d'imagerie acoustique, l'arrangement qui produit la moyenne de NMLS la plus basse est sélectionné pour la suite de l'optimisation. Celui-ci est composé d'un premier cercle de 11 hydrophones à 1,35 m de hauteur et 1,45 m de diamètre, et un deuxième de 9 hydrophones à 2,7 m de hauteur et de 2,9 m de diamètre. Il est illustré dans le schéma de la Figure 3.8.

Vérification des performances : afin de vérifier que des arrangements voisins de la structure sélectionnée ne présentent pas des NMLS moyens plus faibles, les cercles de l'antenne sont tournés autour de l'axe vertical pour aboutir à de nouveaux arrangements dits *rotations* et le cercle intermédiaire est déplacé de bas en haut par pas de 2 cm pour aboutir aux arrangements dits *translations*. Les performances des arrangements résultants de ces procédures sont représentées dans la Figure 3.7. Les *rotations* représentées par les points noirs ne voient que leur NMLS moyen augmenter à L_{-3dB} constante par rapport à l'arrangement de départ. Au contraire, pour les *translations*, le NMLS et la L_{-3dB} varient tous les deux, mais sans parvenir à un NMLS moyen et une L_{-3dB} moyenne inférieurs à ceux de l'arrangement de départ. C'est donc celui-ci qui est conservé, et

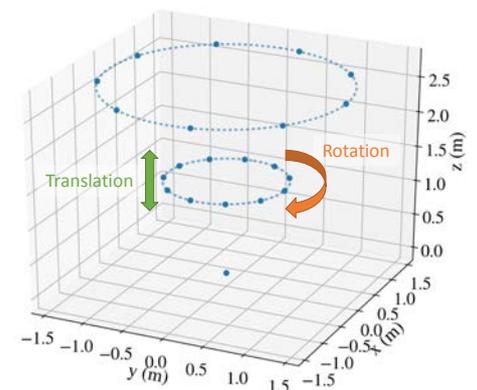


FIGURE 3.8 – Schéma de l'arrangement conique qui produit le NMLS moyen minimal parmi toutes les arrangements coniques. Des *rotations* et des *translations* de son cercle central sont effectuées afin de vérifier si des configurations proches ne produisent pas de meilleures performances.

pour s'assurer de la stabilité de ses performances pour des configurations différentes de celle qui a établi sa géométrie en Figure 3.5, de nouvelles cartes de formation de voies sont calculées en simulation pour une source toujours placée sur le sol mais positionnée de 25 à 65 m au Nord de l'antenne, par pas successifs de 5 m.

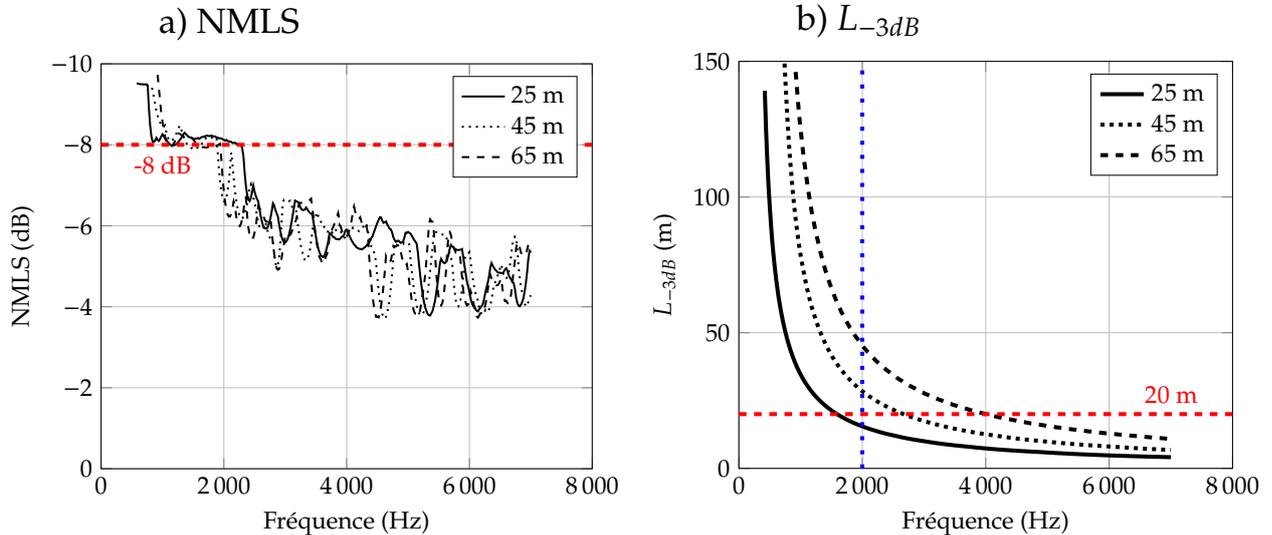


FIGURE 3.9 – Validation des performances de l'antenne en simulation pour des sources situées à 25 m, 45 m et 65 m au Nord de l'antenne. a : NMLS en fonction de la fréquence. b : L_{-3dB} en fonction de la fréquence.

Par souci de clarté, la Figure 3.9a et la Figure 3.9b donnent le NMLS et la L_{-3dB} en fonction de la fréquence, pour les sources à 25 m, 45 m et 65 m au nord de l'antenne uniquement, dont les résultats sont comparables à ceux des autres décalages. Pour la gamme de fréquence [350 Hz ; 2 kHz], l'arrangement sélectionné permet d'atteindre le seuil de -8 dB mais au-delà le NMLS oscille entre -4 et -6 dB, pour les trois décalages différents. En résolution, plus la source est écartée de l'antenne, plus celle-ci est dégradée, avec à 2 kHz une L_{-3dB} qui passe de 15 m à 45 m pour un décalage qui passe de 25 m à 65 m, respectivement. Ceci s'explique par la relation qui lie la résolution à la taille de l'antenne et à la distance à la source déjà vue précédemment [Eli97].

L'ensemble de ces résultats montre que l'antenne peut être utilisée pour des distances source / antenne variant entre 50 et 80 m, voire même au-delà, validant ainsi l'arrangement d'hydrophones trouvé. Une photographie de la version finale construite dans le cadre du projet Abyssound et validée mécaniquement par la société OSEAN SAS lors des déploiements [Jau+19] est montrée en Figure 3.10 sur le pont arrière du navire *L'Europe*. Les 21 voies de mesures enregistrent les signaux acoustiques avec une fréquence d'échantillonnage de 25 kHz.

La suite de ce chapitre est consacrée à la validation du système d'écoute et des méthodes associées, en localisation et en quantification, à travers trois configurations expérimentales.

Caractéristiques techniques de l'antenne

- ▶ Construite par OSEAN SAS
- ▶ 17 hydrophones HTI-96-MIN
- ▶ 4 hydrophones Neptune D/60
- ▶ Échantillonnage à 25 kHz
- ▶ Cercle 1 : 1,45 m / 11 hydro.
- ▶ Cercle 2 : 2.9 m / 9 hydro.
- ▶ Hauteur : 2,7 m



FIGURE 3.10 – Photographie de l'antenne finale construite sur le pont arrière du navire *L'Europe*. Elle est composée de 17 hydrophones HTI-96-MIN et de 4 hydrophones Neptune D/60. (Crédits : OSEAN SAS)

3.3 Essais expérimentaux en mer

L'objectif des essais est de valider le comportement du prototype de 2,7 m de hauteur avec une source référencée, de montrer expérimentalement les apports de la méthode haute-résolution MUSIC, et de démontrer les capacités du système à localiser et quantifier des sources acoustiques en grand fond. Dans le cadre des essais réalisés, le terme petit fond désigne des profondeurs d'une centaine de mètres alors que le grand fond est atteint pour la configuration où le plancher océanique était situé à 700 m de profondeur. Les 1000 m de profondeur ne sont pas atteints au cours de ces essais, mais aucun facteur n'est limitant à ce jour avec le matériel utilisé pour aller jusqu'à 2000 m.

Sans machine excavatrice disponible, les sources acoustiques mises en œuvre dans ces essais sont des transducteurs ITC 1001, ou des transducteurs équivalents. Ils émettent en fonction des configurations testées des signaux monochromatiques ou le signal acoustique type d'une machine excavatrice. Pour obtenir une réponse omnidirectionnelle, le constructeur préconise une utilisation dans la bande de fréquence [10 Hz ; 25 kHz], ce qui englobe les fréquences étudiées. Cependant, plus la limite de la bande de fréquence est basse, plus le besoin en puissance d'entrée est important pour générer les signaux. Or la puissance disponible durant les différents essais effectués, notamment lorsque la source a été mouillée à 680 m de profondeur, a contraint la bande de fréquence à être au moins supérieure à 1 kHz. Les spectres de bruit à large bande sont ainsi générés sur une bande de fréquence égale à [1 kHz ; 7 kHz] au lieu des [350 Hz ; 7 kHz] initialement prévus.

Les différentes configurations, mises en place lors d'essais en Mer Méditerranée qui ont eu lieu en avril 2018 et en mai 2019, sont décrites dans les sous-parties suivantes. Elles ont été rendues possibles par l'implication de tous les partenaires du projet Abyssound, dont la liste est donnée en Annexe A. En avril 2018, les essais ont eu lieu durant deux jours au large de la baie de Cavalaire. L'antenne a été mouillée et remontée chaque jour, et six configurations différentes ont été mises en place. En mai 2019, cinq jours d'essais ont été effectués, avec quatre jours en mer

au large de Saint-Tropez et un jour dans le bassin de l'IFREMER pour faire des mesures de référence. L'antenne a été mouillée une première fois par petit fond, puis une seconde fois par grand fond où elle est restée en place deux jours consécutifs.

Pour les trois configurations, le plan de localisation est défini comme le plan horizontal et parallèle au plancher océanique, à une hauteur égale à celle de la source par rapport au sol. L'estimation nécessaire de la position de ce plan est liée à l'expérimentation réalisée dans ce projet car dans les cas industriels visés, ce plan est directement confondu avec le plancher océanique sur lesquelles sont posées les machines excavatrices, et dont la profondeur est connue. Ici, sa position est donc déterminée pour chaque instant d'émission d'une source, même lorsque celle-ci bouge. Il est ensuite discrétisé avec un pas de 0,5 m dans les deux directions, pour créer la grille sur laquelle la méthode MUSIC est calculée, pour créer la carte de localisation pour chacune des trois configurations suivantes.

Configuration 1 : deux sources en petit fond

La configuration 1 est représentée en Figure 3.11 et voit tous ses éléments placés dans une zone dont le plancher océanique est profond de 104 m. L'antenne est située à 53 m au-dessus du fond, alors que les deux sources sont à 34 m de hauteur et espacées de 40 à 100 m en fonction des courants. Vu de dessus, les deux sources sont situées au nord-est du centre de l'antenne, dont la projection sur le plancher océanique est pris comme le centre du repère de coordonnées global. En fonction des instants de traitement choisis, la position des sources varie de quelques dizaines de mètres, à cause de la dérive des bateaux auxquelles elles sont accrochées par câble. Le plan de localisation est situé 34 m au-dessus du plancher océanique et contient les deux sources.

La source 1 émet des signaux monochromatiques de 10 s, avec des fréquences variant de 3 à 7 kHz par pas de 1 kHz. La source 2 émet un bruit à large bande qui décrit le bruit d'une pompe, soit une pente de -15 dB/décade démarrant à 1 kHz jusque 7 kHz, et dont le niveau d'environ 95 dB à 1 kHz a été réglé empiriquement durant les essais pour permettre au signal d'émerger par rapport au bruit ambiant.

Configuration 2 : antenne en grand fond et source référencée en petit fond

La configuration 2 correspond à une zone plus profonde, avec un plancher océanique situé à 700 m de profondeur. La source est accrochée au bateau utilisé pendant les essais pour être alimentée, ce qui la rend mobile au cours du temps à cause de la dérive de ce dernier. Son immersion varie au cours du temps entre 25 et 85 m, faisant varier la hauteur du plan de localisation utilisé entre 675 et 615 m. L'antenne est lestée par une ligne longue de 40 m, et est donc immergée à 660 m de profondeur. La Figure 3.12 illustre la position des différents éléments, avec un exemple de trajectoire suivie par la source mobile.

La source émet six signaux monochromatiques durant 10 s, de 2 à 7 kHz par pas de 1 kHz. Deux niveaux sont disponibles : un niveau

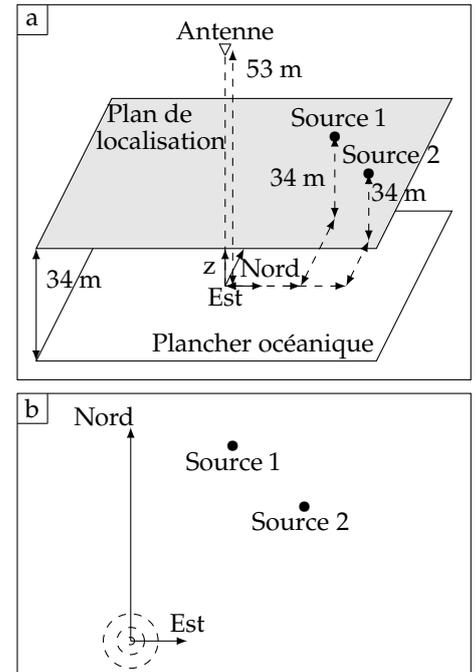


FIGURE 3.11 – Schéma de la configuration 1 avec l'antenne et les sources en petit fond. Les sources sont situées 34 m au-dessus du plancher océanique et dérivent au cours du temps tout en restant au nord-est de l'antenne. a : vue isométrique. b : vue de dessus.

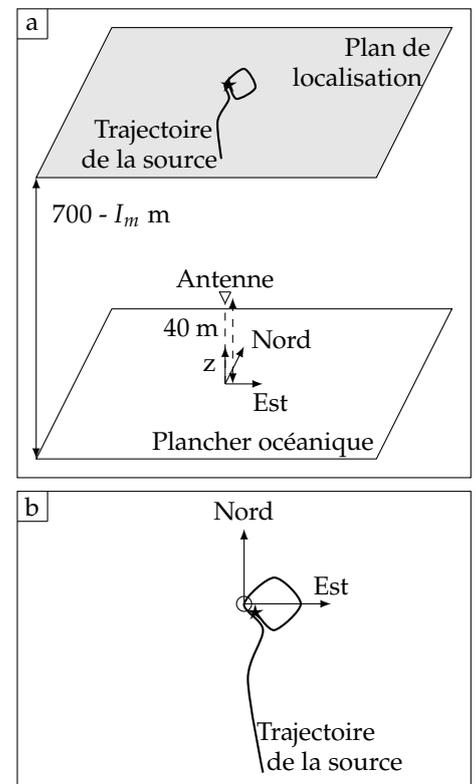


FIGURE 3.12 – Schéma de la configuration 2 avec l'antenne en grand fond et une source référencée qui est mobile en petit fond. a : vue isométrique. b : vue de dessus.

nominal joué dix fois où la source émerge de 30 à 40 dB par rapport au bruit ambiant en fonction des fréquences, et un niveau atténué de 17 dB par rapport au niveau nominal joué cinq fois.

Configuration 3 : antenne et source à large bande en grand fond

Cette dernière configuration schématisée en Figure 3.13 est une représentation fidèle de ce qui est attendu pour un site minier. Elle correspond à la même zone de 700 m de profondeur que dans la configuration précédente. L'antenne et la source sont toutes deux lestées à quelques mètres au-dessus du fond : 40 m pour l'antenne et 17 m pour la source. Elles ne sont soumises qu'aux courants sous-marins, et sont ainsi considérées comme immobiles durant les essais.

La source émet un bruit typique de machine excavatrice tel que décrit dans la partie 3.1, avec un niveau variable de 75 dB à 120 dB par pas de 5 dB pour la partie constante du spectre.

Références pour la localisation et la quantification

Pour valider les méthodes de localisation et de quantification proposées, des vérités terrain pour ces deux aspects doivent être établies.

La configuration 1 avait pour but initial de ne tester que le déploiement de l'antenne et des sources, sans chercher à obtenir des données localisées précisément, et encore moins quantifiées. C'est pourquoi la vérité terrain de la position de l'antenne est prise comme l'endroit où elle a été déployée, et celle des sources est prise comme la position des bateaux qui les tractent grâce à leur signal GPS. L'imprécision de ces mesures GPS et les courants sous-marins qui peuvent faire dériver les sources de l'aplomb des bateaux induisent alors une incertitude sur le placement des sources. Après calcul, cette incertitude est représentée par un cercle de rayon 11,5 m pour la source 1, et 23,5 m pour la source 2. La source 1 est portée par un zodiac, plus petit que le bateau de l'IFREMER l'Europe qui porte la source 2, ce qui explique la différence. Aucune de ces sources n'est quantifiée par une mesure de référence.

Pour les deux autres configurations, la vérité terrain des positions de tous les éléments est assurée par un système acoustique actif placé sur le bateau. Il communique avec des balises situées sur les lignes de mouillage de l'antenne et de la source, et les temps de parcours entre émission et réception des différents signaux permettent de déterminer les positions des éléments au cours du temps. Ce système étant acoustique, son utilisation doit se faire de manière non régulière au cours du temps, en dehors des périodes d'émission de signal par la source elle-même pour éviter une pollution de ses signaux. Un couloir d'incertitude sur la position de la source est alors obtenu en compilant les incertitudes sur tous les éléments dues à l'imprécision de l'estimation des temps de parcours acoustiques, et sa largeur dépend principalement de la distance entre le bateau et les différents éléments qu'il localise, mais aussi de la distance des éléments entre eux. Ce couloir définit la vérité terrain et la

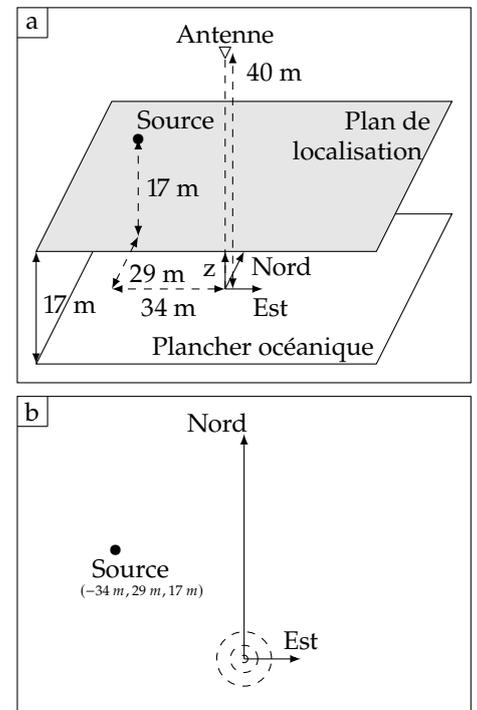


FIGURE 3.13 – Schéma de la configuration 3 avec l'antenne en grand fond et la source large bande en grand fond. a : vue isométrique. b : vue de dessus.

méthode de localisation a pour but d'estimer les positions de la source à l'intérieur de ce dernier.

C'est grâce à ce suivi de la position de la source au cours du temps par le système actif que la hauteur du plan de localisation peut être réglée correctement, en déterminant l'immersion de la source pour chaque instant à traiter. Pour rappel, cette connaissance n'est pas nécessaire dans les cas industriels visés, cette hauteur étant nulle car les machines excavatrices à étudier sont directement posées sur le sol.

Concernant les mesures de référence pour la quantification, seule la source monochromatique de la configuration 2 bénéficie d'une mesure de référence par un hydrophone indépendant réalisée en bassin. A partir de l'enregistrement temporel de cet hydrophone, l'amplitude crête-à-crête est déterminée avant l'arrivée des réflexions dues aux parois du bassin. Son niveau de pression à 1 m effectivement émis est ainsi connu à une incertitude de mesure près. Cette incertitude s'avère relativement élevée à cause du faible nombre d'arches temporelles qui peuvent être prises pour les signaux monochromatiques à référencer dans le cas des basses fréquences 2 et 3 kHz. L'estimation de ce niveau donnée par la formation de voies à partir des mesures faites durant les essais en mer pourra lui être comparée.

La source à large bande de la configuration 3 n'a quant à elle pas pu être référencée par des mesures en bassin, ce qui n'autorise qu'une validation relative de la quantification d'un niveau à l'autre.

La validation des méthodes passe alors par leur accord avec ces vérités terrains et leur incertitude respective, en commençant par la source référencée en localisation et en quantification.

3.4 Validation du système d'écoute

Pour valider le système, la configuration 2 est utilisée car sa source a l'avantage d'être référencée à la fois en localisation et en quantification. Néanmoins, cette configuration particulière où la source est accrochée au bateau induit la sélection d'un modèle de propagation spécifique en comparaison avec les autres configurations.

Modèle de propagation : pour spécifier le modèle de propagation à utiliser entre la source et l'antenne, la distance D_{lim} issue de l'Equation 2.11 doit être prise en compte, ainsi que l'objectif final de l'application. L'objectif des méthodes proposées est d'aboutir à une estimation du niveau de pression à 1 m de la source en présence. Dans ce contexte, c'est un modèle en ondes sphériques qui est le plus adapté car il permet d'estimer directement le niveau de pression à 1 m pour un monopôle situé en une position de la grille de scrutation.

De plus, pour les distances de sécurité envisagées, de l'ordre de 50 m, et pour la fréquence maximale atteinte dans ces essais, 7 kHz, la distance D_{lim} est égale à 84 m. Par conséquent, même si celle-ci est largement dépassée dans le cadre de la configuration 2, les distances entre la source et l'antenne atteintes dans les deux autres configurations sont plus petites, impliquant à nouveau la nécessité d'utiliser un modèle en ondes sphériques.

Pour un système d'écoute placé au plus près des sources acoustiques à caractériser sans danger, la propagation des signaux entre la source et l'antenne peut être considérée en champ libre sans créer d'erreur de modèle importante : la vitesse du son est quasiment identique en tout point du trajet, seul un faible nombre d'obstacles peut être présent, et l'atténuation des ondes acoustiques dans le milieu est négligeable. La seule source d'erreur importante concerne alors la réflexion sur l'interface eau / sol, voire eau / air dans le cas de la configuration 2 où la source est faiblement immergée.

Concernant la réflexion sur le sol, sa prise en compte n'est pas envisagée dans cette thèse car les caractéristiques bathymétriques et sédimentaires des zones d'essais sont inconnues. Sa prise en compte induirait un modèle plus précis certes, mais mal réglé et pouvant donc provoquer des erreurs de modèle plus importantes qu'en prenant un modèle champ libre.

En revanche dans le cas de l'interface eau / air de la configuration 2, le coefficient de réflexion est beaucoup plus important. Par conséquent, un modèle avec source image est utilisé, avec un coefficient de réflexion égal à 1, et une interface située à 700 m de hauteur, soit la hauteur de la colonne d'eau. Il est important de noter que ces paramètres ne sont pas précisément établis, et peuvent donc provoquer comme discuté plus haut des erreurs de modèle avec une variance plus importante que pour l'utilisation d'un modèle champ libre.

Une fois le modèle avec source image sélectionné dans cette configuration, les cartes de MUSIC peuvent alors être calculées sur le plan de localisation discrétisé dont la hauteur est donnée par le système de localisation actif servant à établir la vérité terrain.

Choix du nombre de valeurs propres : à partir de la matrice interspectrale, estimée dans ce cas à partir de 10 s de signal et avec des blocs de 2048 échantillons (résolution fréquentielle de 12 Hz) et un recouvrement de 50 %, le paramètre important de la méthode MUSIC est le nombre de valeurs propres à garder dans l'espace signal comme vu en Partie 2.2. Il est déterminé ici en sélectionnant manuellement le nombre de valeurs propres présentes avant le point d'inflexion qui apparaît dans la courbe de décroissance des valeurs propres, dont un exemple pour la source qui émet au niveau nominal est donné en Figure 3.14.

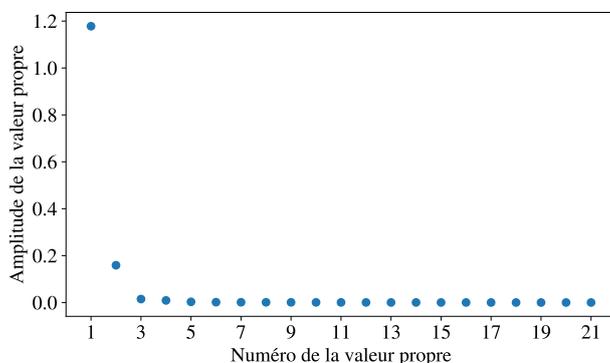


FIGURE 3.14 – Exemple de courbe de décroissance des valeurs propres. Une cassure est observée entre la deuxième et la troisième valeur propre : le nombre de valeurs propres sélectionnées pour le traitement par MUSIC est alors égal à deux.

Dans l'exemple donné, deux valeurs propres sont prépondérantes et sont donc sélectionnées pour représenter l'espace signal. Ces deux valeurs propres correspondent à la source, qui émerge de 30 à 40 dB du

bruit ambiant, et à sa réflexion sur l'interface eau / air. Pour l'ensemble des 60 localisations effectuées pour la source qui émet à un niveau nominal, le nombre de valeurs propres à garder pour composer l'espace signal est d'ailleurs toujours égal à deux, sauf pour une émission où trois valeurs propres étaient prépondérantes. Pour le niveau atténué, le nombre de valeurs propres à sélectionner varie de un à quatre, mais la sélection de deux valeurs propres reste le choix le plus récurrent avec 20 occurrences sur 30.

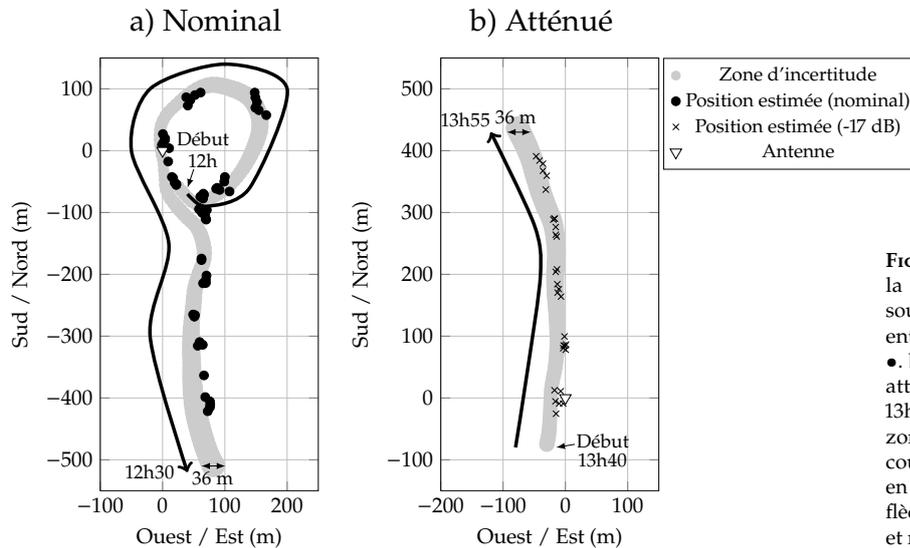


FIGURE 3.15 – Résultats de localisation de la configuration 2. a : localisation de la source au niveau nominal (10×6 signaux) entre 12h et 12h30, donnée par les points \bullet . b : localisation de la source au niveau atténué de 17 dB (5×6 signaux) entre 13h40 et 13h55, donnée par les croix \times . La zone d'incertitude est représentée par le couloir grisé \bullet . La position de l'antenne en (0 m, 0 m, 40 m) est illustrée par ∇ . Les flèches représentent le sens de la trajectoire et ne sont données qu'à titre indicatif.

Résultats de localisation : la Figure 3.15a donne la localisation de la source mobile pour l'ensemble des émissions de la source au niveau nominal. Chaque groupe de points correspond à une séquence de 6 signaux sinusoïdaux de 2 à 7 kHz. La zone d'incertitude produite grâce au système de localisation actif du bateau est représentée par la zone grisée. 50 % des estimations de position de la source sont localisées à l'intérieur de celle-ci. L'écart maximal entre les estimations et la zone d'incertitude reste inférieur à 17 m.

La localisation pour les 5 groupes de 6 émissions du cas atténué de 17 dB est donnée en Figure 3.15b, et 80 % des estimations de position de la source sont situées à l'intérieur de la zone d'incertitude. L'écart maximum entre les estimations et la zone d'incertitude est de 10 m. Avec 60 % de détection des positions de la source à l'intérieur de la zone d'incertitude au global, et des erreurs maximales bornées en-dessous de la taille caractéristique d'une machine excavatrice, ces résultats montrent que malgré le cadre expérimental complexe dû à la distance entre la source et l'antenne qui varie entre 570 m et 800 m et à la présence de l'interface eau / air, la méthode MUSIC associée à un modèle de propagation avec prise en compte de la source image est robuste pour la localisation de source. Des imprécisions sont observées, notamment par des sauts de localisation visibles d'une fréquence à l'autre vers la fin de la trajectoire de la Figure 3.15a de 200 m à 400 m au sud de l'antenne⁹, mais les estimations erronées restent proches de la zone d'incertitude.

Procédure pour la quantification : à chaque instant de traitement, une fois la localisation effectuée, la quantification du niveau de pression à 1 m L_p est calculée pour la position qui correspond au maximum trouvé dans la carte de MUSIC. Pour calculer ce niveau, il est nécessaire de

9: Les groupes de 6 signaux de fréquences différentes ne sont plus correctement rassemblés.

prendre une propagation en champ libre pour éviter une estimation du niveau L_p qui dépend des franges d'interférences que le modèle avec source image crée comme illustré en Figure 3.16 à 3 kHz, mais qui ne représentent pas la physique du problème étudié.

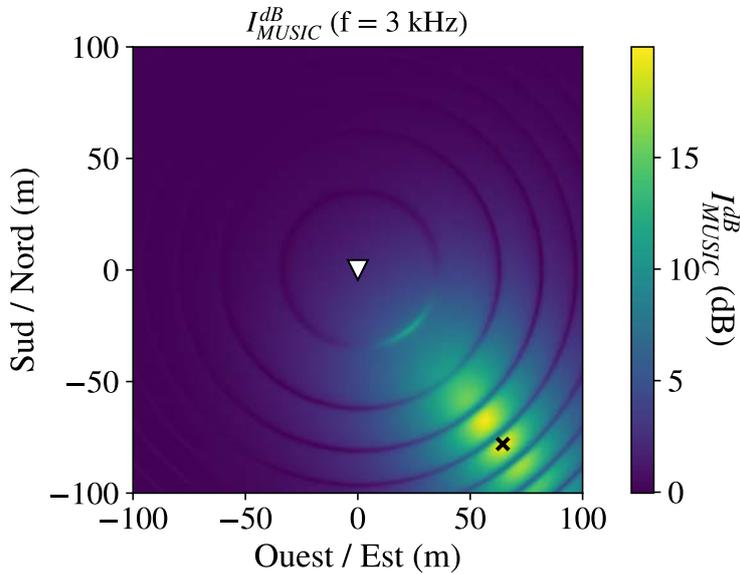


FIGURE 3.16 – Exemple d’une carte de MUSIC obtenue pour la configuration 2 avec un modèle avec source image. La projection de l’antenne sur le plan de localisation est donnée par le triangle inversé ∇ et constitue le centre du repère de coordonnées. La croix \times représente le maximum de la carte de MUSIC.

Résultats de quantification : les niveaux des sources estimés pour chaque fréquence d’émission et pour les deux niveaux d’émission de la source sont donnés en Figure 3.17. L’incertitude de mesure pour l’hydrophone de référence est donnée par les rectangles grisés, dans lesquels doivent être retrouvées les estimations des niveaux.

Pour les fréquences supérieures ou égales à 4 kHz, les niveaux estimés sont trouvés à l’intérieur des bornes d’incertitude. Seules les fréquences 2 et 3 kHz pour le niveau nominal surestiment le niveau de référence. Toutes les estimations à chaque fréquence et chaque niveau d’émission sont concentrées dans un intervalle de 3,5 dB, sauf pour le cas atténué à 2 kHz. Cette concentration des résultats démontre la robustesse de la méthode de quantification proposée au cours des différentes répétitions de la séquence, où la source se déplace sur des distances allant jusqu’à 500 m entre deux estimations différentes de son niveau de pression.

Pour les fréquences 2 et 3 kHz, la surestimation observée s’explique par le fait qu’à ces deux fréquences et à ce niveau élevé, la source sature. Cette saturation a été remarquée lors des essais en bassin, postérieurs aux essais en mer, et provoquent un écrêtage des signaux sur lesquels sont calculés le niveau de référence. Il s’ensuit que l’amplitude crête-à-crête de ces signaux est sous-estimée, ce qui explique les résultats observés.

Ces résultats valident la quantification par formation de voies quantitative sur le maximum de la carte de MUSIC, même pour une source qui se déplace jusqu’à 800 m de l’antenne. Il faut néanmoins chercher à améliorer l’incertitude de la mesure de référence pour éviter d’obtenir des rectangles qui s’étendent jusqu’à 12 dB, empêchant une évaluation précise de la quantification. Sans cette amélioration, il n’est pas possible d’aller plus loin dans l’analyse de la méthode de quantification, et notamment de regarder l’impact du choix du modèle de propagation.

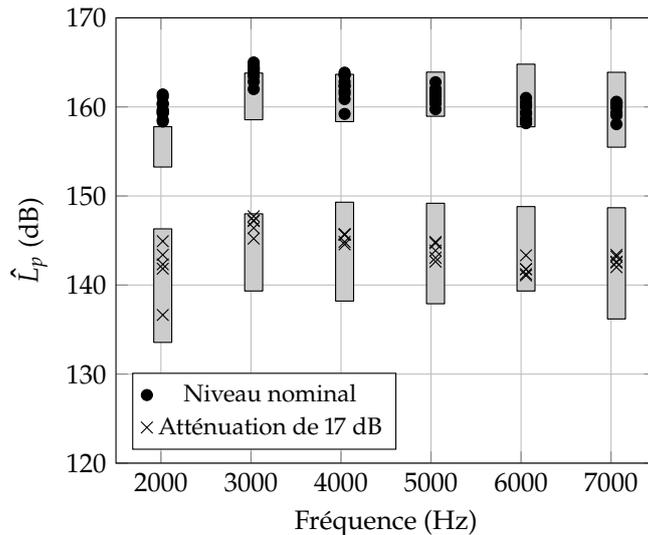


FIGURE 3.17 – Résultats de quantification de la configuration 2 en fonction des fréquences et du niveau d'émission de la source. Les niveaux pour la source émettant au niveau nominal (10×6 signaux) sont donnés par les points ●, ceux pour la source émettant au niveau atténué de 17 dB (5×6 signaux) sont donnés par les croix ×. Les rectangles grisés correspondent à l'incertitude de la mesure de référence.

Conclusion : le système d'écoute composé d'une antenne associée à des méthodes d'imagerie acoustique est ainsi capable de localiser et quantifier une source mobile émettant des signaux sinusoïdaux, à des distances de 570 m à 800 m de l'antenne. La méthode MUSIC combinée à un modèle de propagation avec prise en compte de la réflexion sur l'interface eau / air de la source se montre robuste aux erreurs de modèle, avec des erreurs de localisation inférieures à 17 m, soit moins que la taille d'une machine excavatrice. Les niveaux estimés par formation de voies quantitative sont conformes à ceux de référence, et ils sont répétables d'une émission de source à l'autre, même lorsque celle-ci s'écarte de l'antenne.

Pour aller plus loin dans l'analyse de ces résultats, la partie suivante démontre l'importance de l'algorithme MUSIC, et donc du passage à une localisation haute-résolution, pour s'assurer de trouver toutes les sources en présence dans le contexte minier sous-marin.

Résultats : source référencée

- ▶ 60 % de bonne localisation
- ▶ Erreurs maximum :
 - Localisation : 17 m
 - Quantification : 0,7 dB
- ▶ Estimations des niveaux répétable : 3,5 dB d'écart maximum entre deux estimations

3.5 Apport de la localisation à haute résolution

L'importance du passage à des algorithmes de localisation à haute-résolution dans le cadre minier se justifie par deux raisons :

- ▶ les sources peuvent être proches les unes des autres et ne pas être distinguables par la formation de voies ;
- ▶ les sources peuvent avoir des niveaux sonores très différents, ce qui entraîne des effets de masquage avec la formation de voies.

L'objectif des méthodes haute-résolution étant d'apporter une résolution théoriquement infinie, les sources spatialement proches seront alors distinguées. De plus, le fait que le critère ne prenne en compte que la présence ou l'absence de signal sans tenir compte du niveau relatif de chacune des sources permet de s'abstraire du problème lié au niveau des lobes secondaires de la formation de voies. Pour démontrer ces deux apports, la configuration 1 qui contient deux sources à la même profondeur, soit 34 m au-dessus du plancher océanique et dont le schéma est donné en Figure 3.11, est utilisée [Pha+19a ; Pha+19b].

Cas de deux sources proches : la Figure 3.18 compare les résultats de localisation par les méthodes de formation de voies et MUSIC pour les deux sources situées à 43 m l'une de l'autre, avec la source 2 émettant légèrement plus fort que la source 1. Les positions des bateaux qui tractent les sources sont données par les points de couleur, avec les cercles d'incertitude associés au positionnement de celles-ci par rapport aux bateaux. En Figure 3.18a, le résultat de formation de voies montre un lobe principal unique situé entre les deux sources, sans parvenir à les distinguer au contraire de la Figure 3.18b dans laquelle deux sources sont identifiées par la méthode MUSIC. Trois maxima sont représentés par des croix dans chaque carte, étant donné qu'une réflexion due à la source 2 est trouvée comme deuxième maximum dans la carte MUSIC. Les maxima trouvés par la méthode MUSIC permettent de séparer deux sources spatialement proches qui ne l'étaient pas par la formation de voies, en les retrouvant à l'intérieur des cercles d'incertitude liés à chaque bateau.

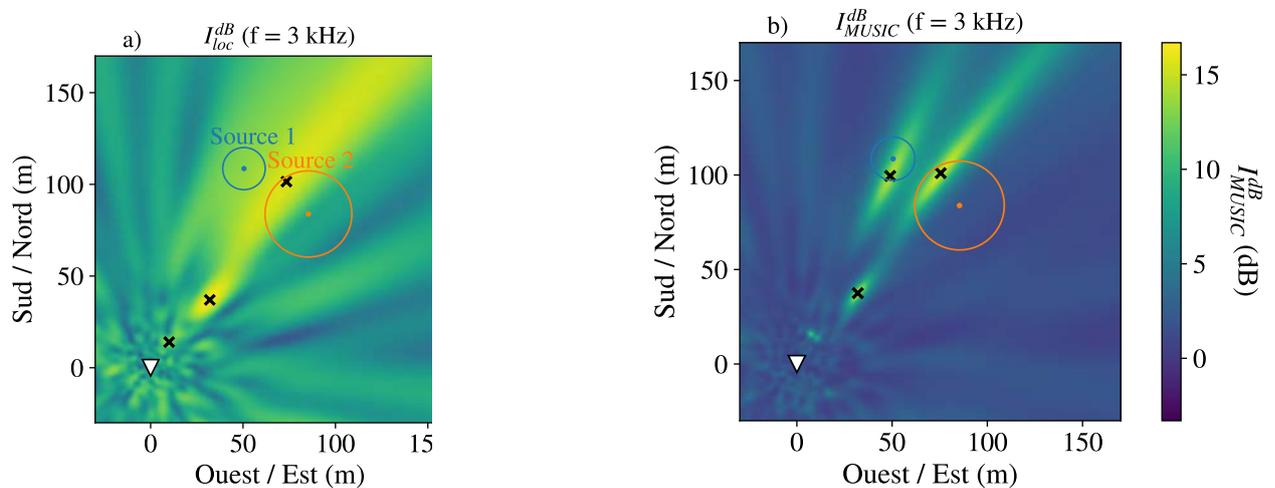


FIGURE 3.18 – Comparaison des cartes de formation de voies et de MUSIC pour le cas où les deux sources sont spatialement proches. a : carte de formation de voies. b : carte de MUSIC. Les bateaux sont situés à 43 m d'écart l'un de l'autre, le point bleu \bullet et son cercle d'incertitude associé représente le bateau qui porte la source 1, alors que le point orange \bullet représente celui qui porte la source 2. L'antenne est représentée par le triangle inversé ∇ et les trois premiers maxima de chaque carte sont donnés par les croix \times .

Pour étudier l'apport de la haute-résolution concernant deux sources de niveaux différents, deux instants temporellement proches sont sélectionnés dans les données de la configuration 1 : un premier où les sources émettent des signaux de niveau équivalent, et un second où la source 1 émet 6 dB plus fort.

Cas de deux sources de niveau équivalent : la Figure 3.19 les cartes de formation de voies (en a) et de MUSIC (en b) pour les sources avec un niveau équivalent. Quelle que soit la méthode d'imagerie utilisée, les deux sources sont distinguées, la source 2 étant correctement localisée alors que la source 1 est localisée 4 m en dehors du cercle d'incertitude associée à son bateau. Sur ce cas, la méthode haute-résolution n'apporte pas d'amélioration par rapport à la formation de voies.

Cas de deux sources de 6 dB d'écart en niveau : quatre minutes plus tôt, les bateaux étaient 13 m (celui de la source 1) et 22 m (celui de la source 2) plus au sud-ouest et la source 1 émettait 6 dB plus fort. Les Figure 3.20a et b donnent les cartes de formation de voies et de MUSIC pour ce cas. Pour la formation de voies, seule la source 1 qui émet plus

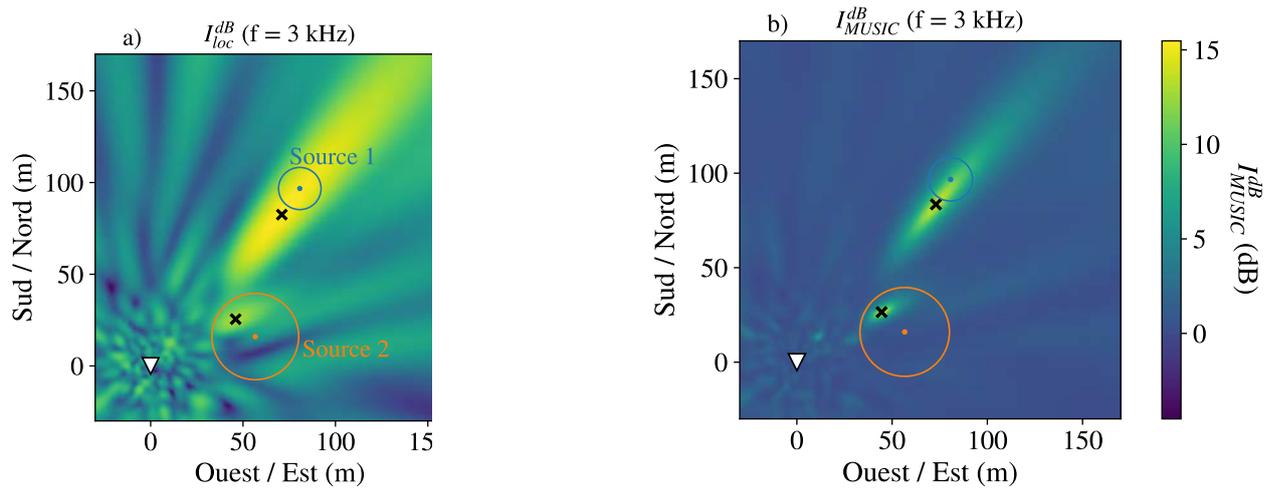


FIGURE 3.19 – Comparaison des cartes de formation de voies et de MUSIC pour le cas où les deux sources ont un niveau équivalent. a : carte de formation de voies. b : carte de MUSIC. Les bateaux sont situés à 84 m d'écart l'un de l'autre, le point bleu • et son cercle d'incertitude associé représente le bateau qui porte la source 1, alors que le point orange • représente celui qui porte la source 2. L'antenne est représentée par le triangle inversé ▽ et les deux premiers maxima de chaque carte sont donnés par les croix ✕.

fort est correctement localisée, et le second maximum trouvé dans la carte correspond à sa réflexion sur le sol. Aucun maximum local n'est présent à l'intérieur du cercle d'incertitude associé à la source 2. Dans la carte de MUSIC au contraire, les deux sources sont correctement localisées. Les réflexions des deux sources sont d'ailleurs aussi visibles dans la carte. Cet exemple illustre le fait que la méthode haute-résolution MUSIC est utile pour retrouver des sources cachées par les lobes secondaires obtenus en formation de voies, qui sont hauts dans ce cas à cause du faible nombre d'hydrophones de l'antenne.

Les capacités en localisation et en quantification du système ont été démontrées, en détaillant en quoi la méthode MUSIC apporte en localisation dans ce contexte minier sous-marin où les sources peuvent être proches ou de niveau différent. La dernière étape de validation du système d'écoute consiste à le tester dans des conditions proches de celles d'un site minier sous-marin réel : avec l'antenne et la source au fond, et un bruit émis sur une large bande similaire au spectre d'une machine excavatrice. C'est l'objet de la dernière partie.

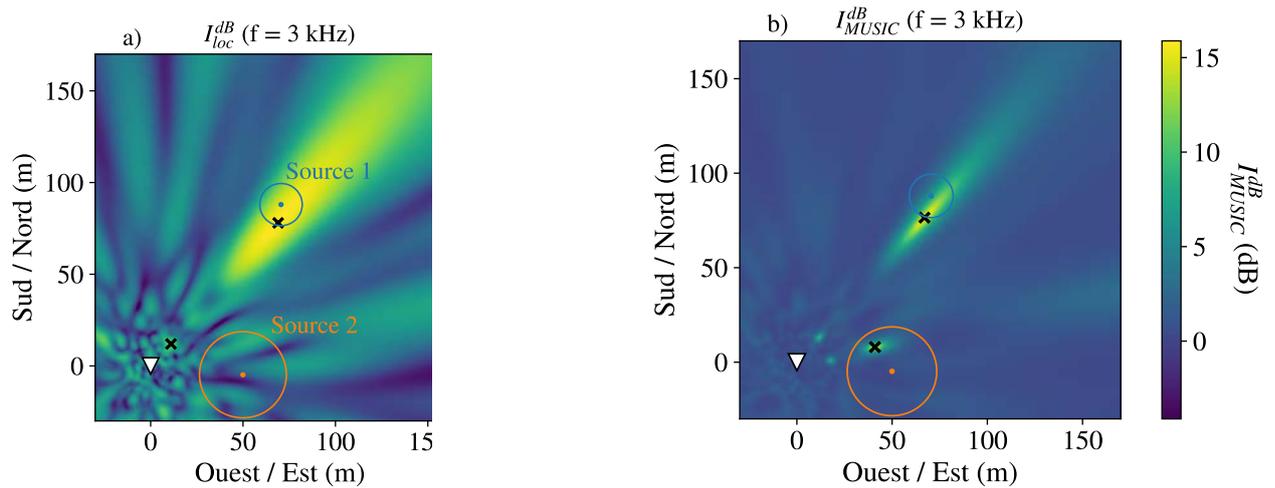


FIGURE 3.20 – Comparaison des cartes de formation de voies et de MUSIC pour le cas où la source 1 émet 6 dB plus fort que lors du cas traité précédemment. a : carte de formation de voies. b : carte de MUSIC. Les bateaux sont situés à 95 m d'écart l'un de l'autre, le point bleu • et son cercle d'incertitude associé représente le bateau qui porte la source 1, alors que le point orange • représente celui qui porte la source 2. L'antenne est représentée par le triangle inversé ∇ et les deux premiers maxima de chaque carte sont donnés par les croix \times .

3.6 Validation en situation réelle par grand fond

Pour étudier la validité du système proposé dans cette étude dans le cadre de la prospection minière sous-marine, la configuration 3 est utilisée. Cependant, quelques différences avec un site minier réel sont à noter :

- ▶ la profondeur de 700 m atteinte dans la configuration 3, qui est moindre par rapport aux profondeurs des mines sous-marines en cours de prospection, plutôt estimées à 1500 m ;
- ▶ la source utilisée qui est un transducteur au lieu d'une machine réelle ;
- ▶ la taille du prototype de l'antenne et la bande de fréquence adaptée à celui-ci ;
- ▶ le fait que la source ne soit pas posée directement sur le sol pour des raisons techniques, contrairement aux machines excavatrices. Ceci implique un traitement sur un plan de localisation parallèle au plan océanique et à 17 m au-dessus de ce dernier, hauteur déterminée grâce au système de localisation actif du bateau.

Ces différences n'empêchent pas la validation de la conception et du bon fonctionnement d'un prototype. Dans l'application industrielle, la mise à l'échelle résoudra le problème de taille de l'antenne et de la bande de fréquence à étudier, l'antenne telle que conçue peut déjà être déployée jusqu'à 2000 m de profondeur et les machines excavatrices seront effectivement présentes sur un site réel et sur le sol, réglant ainsi le problème de l'estimation de la hauteur du plan de localisation, alors égale à zéro. Cette position sur le sol aide par ailleurs au traitement en évitant les réflexions sur ce dernier, avec des sources qui peuvent être considérées comme bafflées.

Choix du nombre de valeurs propres : pour tous les niveaux d'émission de 75 à 120 dB, une carte de MUSIC est calculée à partir d'une matrice interspectrale estimée sur 30 s de signal, avec des blocs de 2048

échantillons et un recouvrement de 50 %. Cette carte est obtenue dans le plan de localisation, parallèle au plancher océanique et à 17 m au-dessus de ce dernier. De la même manière que pour la source harmonique, la sélection du nombre de valeurs propres se fait manuellement par rapport à leur courbe de décroissance, et donne lieu à une sélection de une à dix valeurs propres pour tous les cas grand fond testés. Dans la plupart des cas, plus le niveau d'émission de la source est faible, plus le nombre de valeurs propres à incorporer dans l'espace signal est grand, ce qui traduit la difficulté de la méthode à distinguer les deux sous-espaces lorsque le rapport signal à bruit diminue.

Cartes de localisation : la Figure 3.21a donne un exemple de carte de MUSIC obtenue sur une zone de 100 m par 100 m autour de l'antenne, pour la configuration 3 avec la source qui émet à 120 dB traitée à une fréquence de 2,8 kHz. Une seule valeur propre est nécessaire dans l'espace signal pour ce cas où le niveau d'émission est maximal. La localisation de la source donnée en exemple est estimée par le maximum de la carte en (-34 m, 32 m) à l'intérieur du cercle d'incertitude de diamètre 14 m, centré en (-35 m, 35 m). Bien que la source émette un signal puissant, aucune réflexion sur le sol n'est remarquée. Les sédiments présents sur la zone absorbent beaucoup plus d'énergie acoustique que l'interface eau / air. L'aspect haute-résolution de l'algorithme est particulièrement visible, avec un lobe principal concentré sur une zone spatiale de moins de 5 m, ce qui permet de localiser précisément la source recherchée, avec une "résolution" pratique¹⁰ quatre fois supérieure à celle requise de 20 m.

Il est à noter que pour les faibles niveaux d'émission de la source, la projection du bateau sur le plan de localisation a été retrouvée dans la carte de MUSIC, ayant un niveau pour le critère I_{MUSIC}^{dB} parfois supérieur à celui de la source, comme montré en Figure 3.21b pour une source émettant à 80 dB. Dans cette carte, zoomée sur une zone de 50 m par 50 m autour de l'antenne pour mettre en avant la source parasite, le maximum d'une valeur de 19,3 dB est retrouvé en (-3 m, -2,5 m), et il correspond à l'image du bateau dans le plan de localisation¹¹. La source n'est quant à elle retrouvée que pour le deuxième maximum de la carte, de niveau égal à 14,8 dB. Dans ces cas particuliers, les deux premiers maxima de la carte ont été déterminés, puis seul le maximum lié à la source a été quantifié.

Résultats de localisation : en calculant la distance de la source à l'antenne pour chaque niveau d'émission, les localisations de la source sont présentées de manière synthétique en Figure 3.22, en vérifiant visuellement que toutes sont estimées dans le quart nord-ouest du plan de localisation. La projection du cercle d'incertitude est représentée par rectangles grisés. Pour tous les niveaux d'émission¹², la distance de la source à l'antenne est estimée à l'intérieur des bornes d'incertitude qui s'étalent sur 14 m, validant ainsi la localisation de la source dans cette configuration, avec une précision atteinte inférieure à la taille de 20 m d'une machine excavatrice.

Résultats de quantification : pour la quantification, aucune mesure de référence n'est disponible. Cependant, il est possible de reconstruire les DSP à 1 m du signal source, en calculant comme en partie 3.4 une formation de voies quantitative pour le maximum estimé de la carte de MUSIC. La seule différence entre DSP et niveau de pression à 1 m tel que défini par L_p réside dans la division de la valeur efficace de

10: Théoriquement, la méthode MUSIC a une résolution infinie. Le lobe principal obtenu en pratique provient de la durée finie d'observation et des erreurs de modèle présentes.

11: Celui-ci étant situé très haut par rapport au plan de localisation, sa projection dans ce plan sera toujours située autour de l'origine.

12: Certains sont émis plusieurs fois.

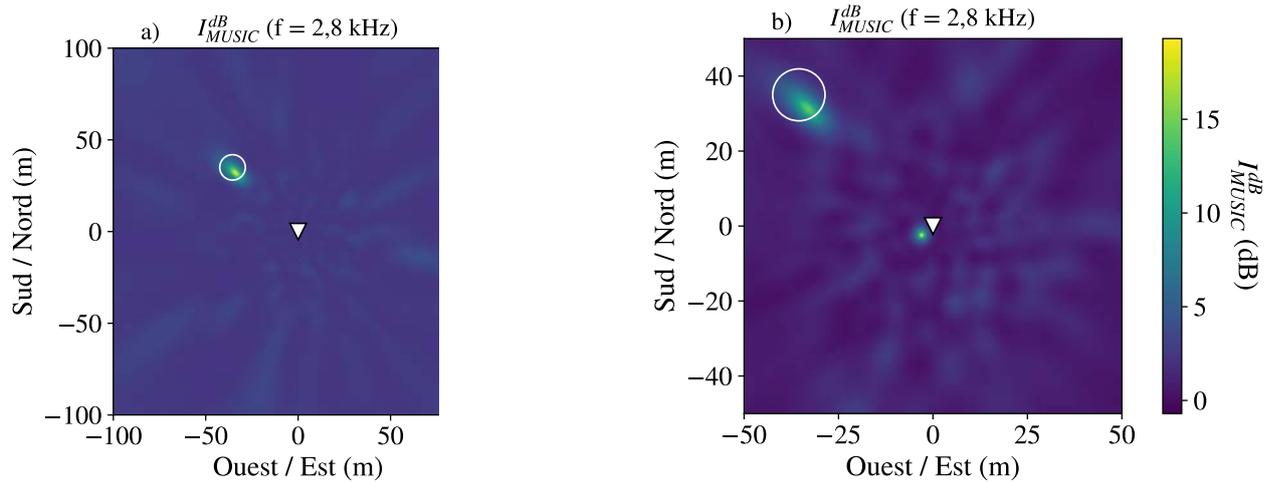


FIGURE 3.21 – Exemple de deux cartes de MUSIC en grand fond. La projection de l’antenne sur le plan de localisation est donnée par le triangle inversé ∇ et constitue le centre du repère. Les cercles d’incertitude ont un diamètre de 14 m et les maxima des cartes ne sont pas représentés pour éviter de masquer la carte. a : carte obtenue pour un niveau d’émission égal à 120 dB. Le cercle est centré en (-35 m, 35 m) et le maximum de la carte est localisé en (-34 m, 32 m). b : carte obtenue pour un niveau d’émission égal à 80 dB. Le cercle est centré en (-37 m, 34 m) et les deux premiers maxima de la carte sont situés en (-3 m, -2,5 m) et (-33 m, 31 m), avec le premier d’entre eux à un niveau de 19,3 dB qui correspond à la projection de la position du bateau dans le plan de localisation étudié, et le second à un niveau de 14,8 dB qui correspond effectivement à la source.

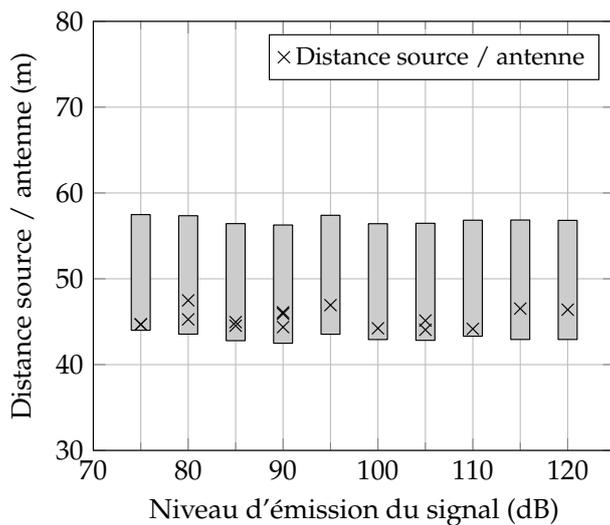


FIGURE 3.22 – Résultats de localisation de la configuration : distance source / antenne en fonction du niveau d’émission de la source. Les rectangles grisés correspondent à la projection du cercle d’incertitude. Les croix \times donnent les distances de la source à l’antenne pour chaque niveau d’émission.

l’amplitude au carré par la résolution fréquentielle pour la DSP. Une estimation du niveau est alors faite pour le maximum courant pour des fréquences allant de 500Hz à 10 kHz par pas de 50 Hz. La Figure 3.23 montre les DSP à 1 m reconstruites de la source pour tous les niveaux d’émission. La forme des spectres reconstruits montre bien une partie constante suivie d’une décroissance, mais celle-ci est plus forte que les -20 dB/décade attendus, et elle démarre à 3 kHz au lieu des 1750 Hz prévus. Pour deux fréquences particulières, 3,4 kHz et 4,15 kHz, la puissance de la source est beaucoup plus faible que pour les autres fréquences. En l’absence de mesure de référence, aucune conclusion ne peut être tirée sur ces observations particulières, étant donné qu’elles peuvent être dues directement à la source et non à la topographie du site des essais.

Néanmoins, grâce à la répétabilité des DSP reconstruites pour les différents niveaux d’émission, une validation par comparaison relative est possible. Chaque courbe de quantification est bien décalée de 5 dB par

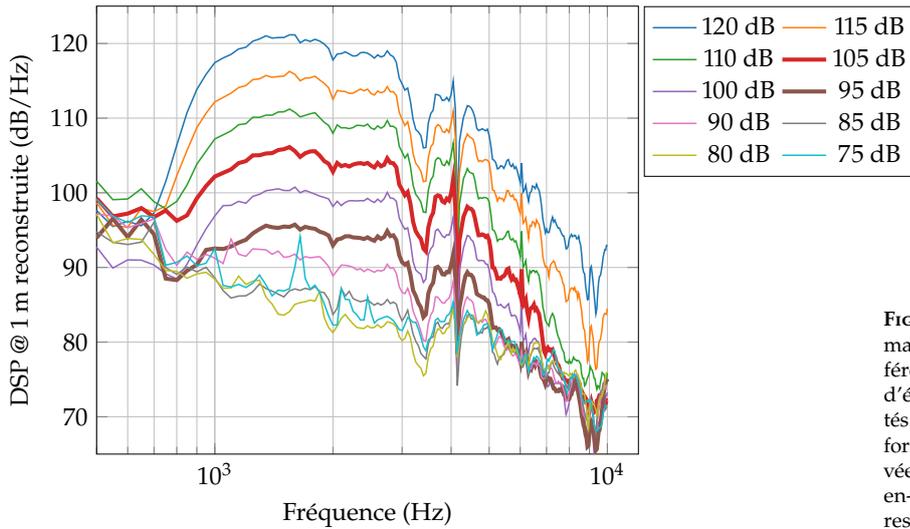


FIGURE 3.23 – DSP reconstruites après formation de voies quantitative pour les différents niveau d’émission. Les niveaux d’émission redondants ne sont représentés qu’une seule fois par souci de clarté. La forme du spectre est partiellement retrouvée jusqu’au niveau d’émission à 90 dB, en-dessous duquel les DSP reconstruites ressemblent à du bruit ambiant.

rapport à celle qui lui est supérieure. La quantification est assurée jusqu’à un niveau d’émission de 90 dB, en-dessous duquel les DSP obtenues ne sont plus représentatives du bruit de machine : la forme particulière du spectre retrouvée jusqu’alors n’est plus distinguable.

Utilité de l’antennerie pour la quantification : pour déterminer l’intérêt de l’antennerie en quantification, il est possible de comparer les DSP reconstruites après formation de voies aux DSP calculées sur un hydrophone unique de l’antenne. Pour les niveaux d’émission de 105 dB et 95 dB, dessinés en gras dans la Figure 3.23, ces DSP sont représentées en Figure 3.24a et Figure 3.24b, respectivement. Alors que pour 105 dB la partie constante du signal émis par la source émerge de 10 dB du bruit ambiant à 2 kHz, pour une source à 95 dB les deux spectres sont confondus. Etant donné que la quantification devient impossible pour une source à 85 dB avec l’antenne, le gain apporté par l’antennerie se situe entre 5 et 10 dB.

Résultats : grand fond

- ▶ Profondeur atteinte : 680 m
- ▶ Précision localisation < 14 m
- ▶ Gain d’antenne : 5 à 10 dB

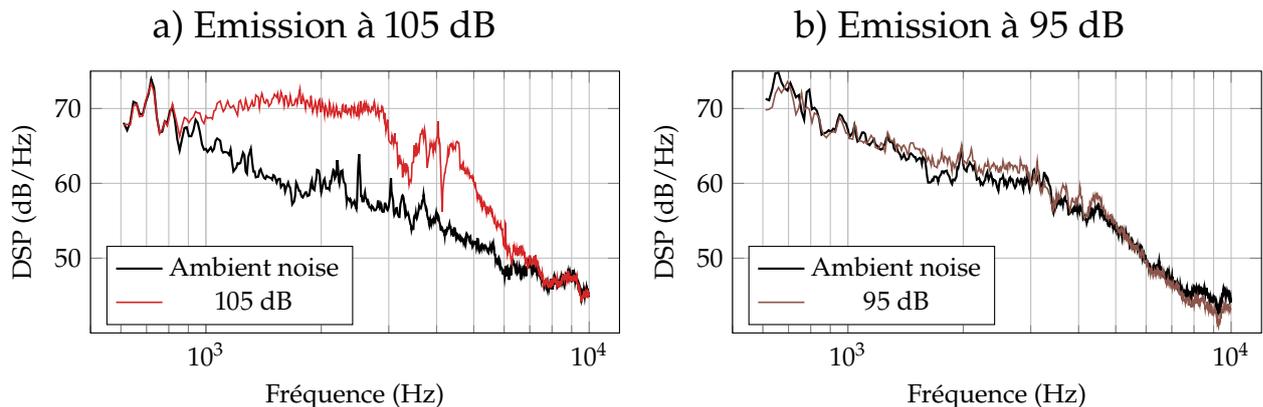


FIGURE 3.24 – DSP d’un hydrophone unique de l’antenne. a : la source émet à 105 dB. b : la source émet à 95 dB. La DSP calculée lorsque la source émet est représentée en couleur en fonction du niveau d’émission, alors que celle calculée juste avant que la source ne commence à émettre, pour décrire le bruit ambiant environnant, est représentée en noir.

Pour une antenne de M hydrophones, le gain d’antenne est défini par $10 \log_{10}(M) = 13$ dB pour 21 hydrophones. Même s’il est estimé 3 à 8 dB plus haut que ce qui est observé en pratique, il reste dans l’ordre de grandeur de ce qui est effectivement gagné en pratique lorsque la source

est quantifiée par antennerie.

En conclusion dans le cas grand fond, le prototype construit localise correctement et produit une quantification relative cohérente d'une source située à 17 m au-dessus du sol. La réflexion de la source sur les sédiments du sol, non présente pour des machines situées directement sur le plancher océanique, n'est pas problématique pour localiser la source dans les cartes MUSIC calculées. La seule source de perturbation provient du bateau qui servait au déploiement des différentes sources, qui lors de niveaux d'émission trop bas devient visible et doit être retiré de la carte. Ces résultats valident l'utilisation du prototype en grand fond et permettent d'aborder le passage à l'échelle sereinement, celui-ci devant permettre de résoudre les dernières contraintes techniques qu'un prototype de 3 m ne peut surmonter.

3.7 Conclusion

Ce chapitre a permis la caractérisation d'un premier type de source acoustique dans son environnement : les machines excavatrices pour la prospection minière par grand fond. L'objectif de cette caractérisation est de déterminer la **localisation et le niveau sonore** de ces machines grâce à un système d'écoute à concevoir. Le système envisagé consiste en une antenne d'hydrophones, dont la conception doit permettre le respect des contraintes sécuritaires liées aux éléments présents dans un site minier sous-marin, mais aussi assurer des performances optimales malgré l'impossibilité de prévoir son cap une fois posée au fond de l'eau. Une première étape de conception sélectionne ainsi la géométrie de l'antenne, pour aboutir au choix d'une antenne conique avec pointe vers le bas pour obtenir des niveaux de lobes secondaires faibles et stables dans toute la bande de fréquence du bruit de machine. La seconde étape optimise l'arrangement des hydrophones sur cette forme, pour créer le prototype final à échelle réduite, qui consiste en une **antenne de 2,7 m de haut et 2,9 m de diamètre maximum, avec 21 hydrophones** arrangés sur le cône en deux cercles et un hydrophone à la pointe. Ce prototype est utilisé dans trois configurations réelles d'application en mer Méditerranée.

En présence d'une source monochromatique référencée et mobile, la configuration 2 valide la capacité du prototype à la **localiser au cours du temps** avec une erreur de 17 m maximum, et la **quantifier** avec des estimations répétables à 3,5 dB près malgré des distances entre cette source et l'antenne qui varient de 570 à 800 m, soit des distances largement supérieures à celles visées sur des sites miniers en grand fond.

Pour les essais en petit fond de la configuration 1, la présence de deux sources permet de **démontrer l'apport de la méthode haute-résolution MUSIC** pour séparer deux sources spatialement proches ou être capable de retrouver une source de niveau faible par rapport à une autre. Dans le premier cas, la formation de voies est limitée en résolution spatiale et estime une position de source située entre les deux sources lorsqu'elles sont espacées d'environ 40 m, alors que MUSIC estime correctement la localisation des deux sources. Dans le second cas, deux sources de niveaux équivalents sont distinguées par la formation de voies, mais

lorsque l'une d'entre elles émet 6 dB plus fort, la seconde passe sous le niveau des lobes secondaires de celle-ci. Avec MUSIC, les deux sources sont correctement localisées peu importe les niveaux d'émission.

Enfin la configuration 3 valide en **grand fond (700 m) la localisation et la quantification relative** d'une source émettant un bruit large bande typique d'une machine excavatrice, avec des distances source / antenne estimées correctement dans des bornes d'incertitude de 14 m, et des densités spectrales de puissances reconstruites qui s'échelonnent par pas de 5 dB en accord avec les niveaux émis par la source. Grâce à l'antenne, la source peut être quantifiée jusqu'à un niveau d'émission de 90 dB, alors que pour une source à 95 dB ce n'est déjà plus possible à partir d'un hydrophone unique. Le gain apporté en pratique est alors de 5 à 10 dB, légèrement inférieur au gain d'antenne calculé à 13 dB.

La perspective prioritaire à la suite de cette validation serait d'obtenir des mesures de sources à large bande référencées afin de s'assurer du bon fonctionnement du système sur l'ensemble de la bande de fréquence. Des travaux sur les algorithmes de déconvolution pour fusionner les étapes de localisation et de quantification en une seule méthode ont été entrepris. Ils permettent de débruiter la contribution du bateau par la prise en compte de deux plans de localisation lors de la déconvolution, mais n'améliorent pas de manière significative les prédictions de localisation et de quantification. La suite concrète de ce projet n'est pas définie et dépend fortement des évolutions normatives qui accompagneront les projets de prospection minière sous-marine à travers le monde.

Localisation et identification couplées de drones par acoustique

4

Ce chapitre décrit un second exemple de caractérisation de source acoustique, cette fois-ci aérienne : les drones. L'objectif est à nouveau de donner leur **localisation**, mais associé à leur **identification** pour aborder des applications de surveillance de sites stratégiques en limitant le nombre de fausses alarmes générées. Le chapitre débute avec le contexte lié à cette problématique, avant de décrire la chaîne de traitement acoustique proposée dont l'innovation réside dans le couplage du traitement d'antenne et de l'apprentissage statistique. Cet enchaînement permet de localiser et filtrer spatialement les signaux provenant des sources pour améliorer leur rapport signal à bruit, puis de comparer deux approches pour l'identification : la classification binaire supervisée et la classification à une classe non supervisée. La chaîne est validée grâce à des données expérimentales créées durant la thèse qui comportent des sources seules dans un environnement intérieur. Les limites de la localisation acoustique et la pertinence des caractéristiques retenues pour l'apprentissage sont étudiées avant de démontrer l'intérêt du couplage proposé. Celui-ci permet la localisation et l'identification de deux sources présentes simultanément, jusqu'à une certaine limite déterminée par la différence de niveau sonore entre les sources. Enfin, un cas extérieur plus proche des applications finales visées est étudié grâce à des données également produites durant la thèse, avec la présence d'un drone et d'une source perturbatrice, simultanément ou non.

Table des matières

4.1	Contexte	70
4.2	Description de la chaîne de traitement	76
	Localisation	76
	Focalisation	77
	Identification	78
4.3	Bases de données expérimentales produites	81
	Environnement intérieur	81
	Environnement extérieur	82
4.4	Validation de la localisation et de l'identification pour des sources seules en intérieur	84
4.5	Etude des caractéristiques pour l'identification	88
	Liens entre les signaux et les caractéristiques	88
	Modification du jeu de caractéristiques	93
4.6	Limites de fonctionnement en localisation	94
4.7	Intérêt du couplage traitement d'antenne / ap- prentissage statistique pour l'identification	96
4.8	Limites de fonctionnement en focalisation en pré- sence de plusieurs sources	101
4.9	Performances en extérieur	103
	Sources présentes une à une	103
	Drone en présence d'une source perturbatrice	107
	Performances finales en fonction de la distance	111
4.10	Conclusion	112

4.1 Contexte

Grâce à un faible coût et une facilité d'utilisation apportés par les avancées technologiques, le marché du drone est passé de 2 à 22 milliards de dollars entre 2016 et 2020 et est estimé à 42 milliards en 2025 [Gio+17; Dro20]. Bien qu'en grande partie utilisés à des fins honnêtes, une partie de ces drones peuvent servir à attaquer des sites sensibles comme les centrales nucléaires [Sol+18; Age14] ou des personnes publiques comme la chancelière allemande Angela Merkel [Der13], voire même paralyser des zones entières comme l'aéroport de Gatwick perturbant dix mille voyageurs [BBC18]. En fonction de l'objectif affiché des malfaiteurs, le drone utilisé n'est pas le même. Le terme **drone** regroupe en effet une quantité importante de systèmes, englobant aussi bien les machines qui évoluent dans les airs, sur le sol, ou dans la mer. Dans cette thèse, ce sont les drones volants qui sont étudiés. Mais même au sein de cette catégorie, différentes familles existent, dont une taxonomie est proposée en fonction de leur taille dans un récent appel à projet européen H2020 [Eur20] :

- ▶ les drones de classes I a et b (micro drones) : ce sont les drones de moins de 2 kg ;
- ▶ les drones de classe I c (mini drones) : entre 2 et 20 kg ;
- ▶ les drones de classe I d (petits drones) : entre 20 et 150 kg ;
- ▶ les drones de classe II (drones tactiques) : entre 150 et 600 kg.

Un problème important qui se pose alors pour les acteurs responsables de la protection militaire ou civile est de se prémunir d'attaques pouvant provenir de menaces variant grandement en termes de tailles, de formes, et de technologies utilisées. Ainsi, si les technologies militaires classiques comme le radar sont efficaces contre les drones de classe II, elles le sont beaucoup moins sur ceux de classe I, composée en grande partie de drones **quadrirotors** qui constituent le type de drone étudié dans cette thèse, et dont la surface équivalente radar est faible [Mic19; DS+17].

Or ces drones sont impliqués dans un nombre croissant d'incidents que ce soit dans le domaine civil [BBC20; Ham20; Hub+19; New20] ou militaire [Joh20; Woo20]¹. Pour limiter leur utilisation malveillante, des systèmes industriels de détection sont développés, et leur nombre croit au fur et à mesure des années passant de 10 en 2015 [Bir+15] à 537 en 2019 [Mic18; Mic19]. Ils sont basés sur différentes technologies qui ont chacune leurs avantages et inconvénients :

1: Pour avoir un échantillon des problèmes causés par ces drones, suivre la [page LinkedIn](#) de la société CERBAIR.

Technologies utilisées pour la détection de drones

- ▶ **La radio-fréquence** : cette technologie cherche à intercepter les ondes radio émises entre le pilote du drone et son drone. La détection est possible jusqu'à des distances de 5 à 10 kilomètres et n'est pas limitée par l'environnement ou les conditions d'opération. Par contre, des drones peuvent être désormais pilotés grâce à des réseaux mobiles, ou d'autres voient leur trajectoire pré-programmée pour voler jusqu'à leur cible de manière autonome. Ces nouvelles solutions de pilotage empêchent la

détection par ce type de moyens.

- ▶ **Le radar** : les systèmes radar s’inspirent directement de ce qui est fait pour détecter des cibles ennemies dans le domaine militaire : elles envoient un signal radar dans une direction donnée, s’il se réfléchit sur une cible, le temps entre l’émission et la réception du signal par le système permet de déterminer la distance à la cible. Elles peuvent atteindre des distances de détection de 5 à 10 kilomètres, et fonctionnent dans des conditions météorologiques difficiles. Elles sont par contre peu robustes au masquage par des obstacles physiques, et peuvent émettre des fausses alarmes par exemple en présence d’oiseau. De plus, pour le cas des drones, une couverture à 360° de l’espace est nécessaire, ce qui n’est actuellement pas ou peu proposé par ce type de systèmes. Enfin, ce matériel est coûteux par rapport à l’ordre de prix des menaces auxquelles il a à faire face [Poi+17].
- ▶ **L’optronique** : les systèmes optroniques cherchent à identifier les drones grâce à un capteur optique associé à un système de traitement d’images. Ils fonctionnent sur des distances plus courtes allant jusqu’au kilomètre. En utilisant des algorithmes de reconnaissances de formes, ils peuvent réduire la probabilité de fausse alarme en s’assurant que la cible repérée est bien un drone. Ils servent aussi de moyen légal pour prouver l’intrusion sur un site d’un drone en particulier. Néanmoins, ils sont rapidement limités dans des environnements complexes, où les masquages sont possibles comme les environnements urbains, ou encore lorsque les conditions météorologiques sont difficiles. Un capteur infrarouge est souvent associé pour détecter aussi la signature thermique du drone ciblé, et pouvoir fonctionner de nuit par exemple.
- ▶ **L’acoustique** : cette technologie exploite le signal sonore émis par les drones pour la détection, grâce à un ou plusieurs microphones. Elle permet une détection jusqu’à des distances de l’ordre de 500 m, voire d’un kilomètre dans des environnements calmes. Elle est souvent utilisée pour identifier la signature acoustique de la cible comme étant bien un drone grâce à une comparaison avec une base de données de référence. Le niveau sonore du bruit ambiant de la zone à surveiller constitue son principal défaut, avec des distances de détection qui rétrécissent lorsque ce niveau augmente.

Les industriels qui travaillent sur ces enjeux sont actuellement en train de prendre conscience que les faiblesses de chaque technologie ne seront compensées que par l’intégration de plusieurs d’entre elles dans un système de détection global [Lyk+20; Koc16]. Dans cette thèse, l’objectif est de concevoir une chaîne de traitement exploitant la modalité acoustique et d’en comprendre les avantages et les limitations.

Le choix de l’acoustique pour la localisation et l’identification des quadrirotors est cohérent au vu du système de propulsion qu’ils utilisent. Celui-ci est basé sur un système de plusieurs pâles en rotation, dont le signal acoustique émis associe une composante harmonique et une composante à large bande comme montré expérimentalement [Sin+13]. Les harmoniques proviennent du caractère périodique du mouvement

de la pâle, qui est relié à son déplacement au sein du milieu fluide et à la charge aérodynamique à porter. Le bruit à large bande est dû aux différentes structures turbulentes qui apparaissent en bordures de pâle, mais aussi le long du profil [Ser+19; Can+19]. De plus, l'interaction avec la structure mécanique du drone accentue la composante harmonique du bruit rayonné [Zaw+20], donnant lieu aux spectres caractéristiques observés expérimentalement dans la littérature, et associés au signal acoustique émis par un drone composé de plusieurs couples moteur / pâle [Bla19]².

2: Voir Chapitre 3 de la thèse.

Dans une étude sur les **acteurs industriels** du marché des solutions anti-drones [Mic19], 34 systèmes sur 323 sont annoncés comme exploitant l'acoustique. Au sein de ces 34 systèmes, seuls 5 d'entre eux exploitent la possibilité de combiner plusieurs microphones pour utiliser du traitement d'antenne, sans détailler précisément les algorithmes utilisés, certainement pour protéger leurs avancées technologiques :

- ▶ le produit Discovair de la société *SquareHead* utilise les signaux acoustiques issus des drones pour les localiser et les identifier. Il est composé de 128 voies de mesures MEMS (*Microelectromechanical Systems*) auxquelles est ajoutée une caméra centrale qui sert à l'étape d'identification ;
- ▶ la société *Advanced Protection System* développe le produit CTRL+Sky. Ce système exploite toutes les technologies possibles, et sa partie acoustique est composée d'une antenne de 8 microphones qui est utilisée pour trouver une direction d'arrivée grâce à un algorithme de formation de voies. Ils annoncent une détection et une classification de drone à une distance de 200 m grâce à leur antenne acoustique ;
- ▶ le produit *DroneFence* est développé par la société du même nom. C'est à nouveau un système multi-capteurs, et il utilise une antenne acoustique linéaire de 8 microphones. Afin d'obtenir des positions de drone dans l'espace, ils proposent d'utiliser la triangulation grâce à deux systèmes synchronisés ;
- ▶ la société *DeDrone* a commercialisé un produit appelé DroneTracker, qui était composé de quelques microphones agencés autour d'une caméra centrale. Il est aujourd'hui difficile de retrouver les caractéristiques de ce produit, dont la production semble avoir été arrêtée pour être remplacé par un autre qui exploite les ondes radio ;
- ▶ *Panasonic* a développé une antenne acoustique de 32 microphones avec une caméra centrale pour la détection de drones. Le manque d'information récente sur ce produit amène à penser qu'il n'est plus maintenu.

Dans le **domaine académique**, les contributions scientifiques sur ce sujet sont à ce jour peu nombreuses mais l'intérêt qui y est porté est grandissant ces dernières années. Elles n'abordent souvent qu'un des aspects entre localisation et identification, et c'est pourquoi la synthèse présentée ci-après sépare les deux objectifs, alors que dans la chaîne de traitement proposée dans cette thèse ils sont intrinsèquement liés : la sortie de la localisation aide à l'identification et inversement l'identification permet de ne garder que la localisation qui correspond bien à un drone pour établir sa trajectoire. Bien que certains travaux, qui seront par la suite

signalés, indiquent travailler sur les deux objectifs, ceux-ci sont étudiés de manière indépendante en n'exploitant pas la possibilité de filtrage spatial qu'apporte le traitement d'antenne, et qui permet d'augmenter le rapport signal à bruit d'un signal acoustique dans une direction donnée pour aider à l'identification.

Méthodes pour la localisation : une première approche consiste à exploiter l'effet Doppler afin de reconstruire la trajectoire d'un drone en déterminant le point d'approche le plus proche, puis de déterminer la droite sur laquelle celui-ci se déplace en exploitant la méthode du maximum de vraisemblance pour obtenir les positions successives en azimut [Ton+16; Ton+13]. Cette approche est limitée à la reconstruction de trajectoires linéaires, qui ne sont que peu rencontrées en pratique.

Pour obtenir des angles d'arrivée de drones avec une antenne acoustique constituée de 4 branches de seulement 3,75 cm de long spécialement conçue pour le problème, une méthode d'estimation de la vitesse particulière acoustique au centre de l'antenne est utilisée dans le cadre d'une thèse [Ram19]. L'objectif de cette antenne, qui doit être combinée à terme à d'autres unités du même type pour former un réseau, est de déterminer la direction d'arrivée d'une cible et de l'identifier comme drone afin de pointer un système optronique perfectionné dans cette direction pour confirmer ou infirmer la présence du drone. Les deux aspects localisation et identification sont ainsi abordés, avec des capacités de localisation validées en intérieur [Ram+16], puis en extérieur avec aussi des résultats d'identification [Ram+18]. L'utilisation du filtrage spatial est évoquée en Chapitre 5 de la thèse [Ram19] sans être appliqué aux données expérimentales collectées.

Une approche partagée par un plus grand nombre d'équipes qui travaillent sur le sujet consiste à déterminer les retards entre les différentes voies de mesure par des mesures de corrélation entre les signaux mesurés [Sed+18; Cha+18b; Wu+19; Bla+19b]. Les propriétés harmoniques des signaux de drone sont ensuite utilisées par certains pour améliorer la localisation [Bla+19a; Bla+19c], ou pour identifier les sources trouvées comme étant bien des drones [Sed+19], mais toujours sans exploiter le filtrage spatial. L'identification est aussi menée de manière indépendante avant la localisation par ces méthodes [Shi+20].

La formation de voies est aussi utilisée pour localiser des drones, que ce soit avec une antenne spécialement conçue pour être facilement modulable [Zel+09], en utilisant deux antennes pour pouvoir faire de la triangulation et obtenir une position en trois dimensions dans l'espace [Bus+15], ou encore en effectuant directement une formation de voies fonctionnelle sur une grille en trois dimensions puis en utilisant un filtre de Kalman pour obtenir la trajectoire du drone dans une salle anéchoïque [Her+20].

Enfin, les performances en localisation de drones des méthodes haute-résolution ont aussi été étudiées, avec un premier exemple validant un algorithme de type MUSIC sur une large bande de fréquences³ par des essais expérimentaux menés par des laboratoires liés à l'armée américaine [Pha+04]. Une seconde équipe exploite comme vu plus haut les propriétés harmoniques des signaux de drones en ne les traitant par MUSIC qu'aux fréquences qui correspondent aux harmoniques du drone en présence [Cha+18a; Yan+18].

3: Avec l'algorithme IWM, *Incoherent Wideband MUSIC*.

Pour l'application visée dans cette thèse, la localisation est envisagée de manière angulaire car les sources à caractériser se situent dans le champ lointain de l'antenne. En prenant deux sources situées loin de l'antenne, il est alors probable que même avec une distance grande entre elles deux, leur position angulaire par rapport à l'antenne soit proche. En prenant ensuite une source proche et une source lointaine, les niveaux sonores arrivant sur l'antenne seront fortement distincts pour ces deux sources. Pour parer à une mauvaise localisation dans ces deux cas de figure, la méthode haute-résolution MUSIC est sélectionnée pour localiser précisément des sources potentiellement proches, ou de niveau sonore à l'antenne distincts.

Méthodes pour l'identification : la littérature concernant cet objectif est récente et non mature à ce jour. De nombreuses méthodes sont testées sur des bases de données comprenant peu d'observations, souvent créées pour l'étude seulement, et les résultats générés n'ont pas un format commun ce qui empêche une comparaison claire même si des revues bibliographiques commencent à faire leur apparition [Tah+19].

Il est ainsi montré que des mesures de corrélation [Mez+16], ou une méthode d'empreinte digitale audio [Mez+15], entre un signal acoustique de drone et des signaux à classer permettent de correctement décrire la ressemblance entre des signaux de même catégorie. Une détection de drone par cellules spatiales est aussi envisageable [Yan+19], avec la possibilité d'avoir la position de ce drone dans l'espace [Seo+20]. Mais les résultats de ces méthodes sont uniquement qualitatifs, sans donner de performances chiffrées.

Une autre approche compare la méthode PIL (*Plotted Image Learning*), qui mesure la ressemblance entre le spectre du signal à classer et celui d'un drone labélisé, et la méthode des k plus proches voisins (k -NN : *k-Nearest Neighbors*) qui construit une frontière entre deux classes en se basant sur le label des k voisins les plus proches de chaque observation d'apprentissage. Elles atteignent respectivement 83 % et 61 % de précision [Kim+17 ; Kim+18] sur des données contenant un drone volant seul générées par eux-mêmes, et du bruit environnant sans drone récupéré dans des vidéos Youtube.

La régression logistique a également été utilisée avec une précision de 99 % pour la classification d'un drone seul, posé sur un trépied, dans une salle reproduisant un hall de centre commercial [Ian+20]. Les caractéristiques utilisées dans cette étude sont les niveaux en tiers d'octave enregistrés par le microphone, ce qui limite le fonctionnement de cette méthode à des cas où le signal du drone est prédominant.

Les méthodes mettant en jeu des modèles pour représenter les données tels que les mélanges de gaussiennes [Bor+19], ou les modèles de Markov cachés [Shi+18 ; Nij+16], atteignent des performances fortement variables entre 40 % et 100 %, avec une sensibilité accrue aux données d'entrée : lorsque le jeu d'apprentissage est réduit, les performances passent de 100 % de bonne classification à 75 %.

Un des angles les plus utilisés pour identifier les drones grâce à un enregistrement acoustique est d'utiliser des réseaux de neurones. Il existe ainsi des travaux qui comparent des mélanges de gaussiennes à des réseaux convolutifs et récurrents [Jeo+17], des réseaux entre eux [Al+19],

ou qui exploitent les propriétés temporelles des signaux avec des réseaux qui intègrent une mémoire [Kol+20]. Une autre façon largement étudiée consiste à passer les signaux dans le plan temps-fréquence pour utiliser ensuite des réseaux convolutifs tirés du traitement d'image [Dum+20; Seo+18; Cab+19]. Les performances en classification varient entre 80 à 96 % en fonction des données utilisées, sachant que le passage à ce type de méthodes implique un plus grand nombre de données d'apprentissage nécessaire pour éviter le sur-apprentissage des modèles.

Un autre angle grandement étudié est l'utilisation du modèle SVM, que ce soit en comparaison avec des réseaux de neurones [Gan+20], ou en utilisant des caractéristiques basées sur le traitement de la parole, les MFCC (*Mel-Frequency Cepstral Coefficients*) [Wan+18], en y ajoutant des caractéristiques basées sur les statistiques du signal [Ber+17b], ou en les comparant à d'autres coefficients cepstraux, les LPCC (*Linear Predictive Cepstral Coefficients*) [Anw+19]. Des approches exploitant des caractéristiques liées aux propriétés harmoniques des signaux de drones ou leur transformée en ondelettes sont aussi étudiées [He+19]. Les performances de ces approches oscillent entre 60 et 98 %, avec un net avantage pour les méthodes qui n'utilisent pas qu'un seul type de caractéristiques comme les MFCC, mais en y ajoutant par exemple des mesures sur les statistiques des signaux [Ber+17b].

Enfin, la question de la détection simultanée de plusieurs sources est abordée grâce à des méthodes de séparation de source aveugle [Gha+20], avec des performances en classification après cette séparation qui s'établissent autour de 90 % [Udd+20].

Au vu des performances apportées par l'algorithme SVM, et sa capacité à passer à un cas non-supervisé par l'utilisation du SVM à une classe (SVM-1C), ce sont ces deux modèles qui sont sélectionnés pour l'identification dans cette thèse. Ils sont combinés à des caractéristiques issus de travaux sur la classification de signaux acoustiques [Mal18], qui représentent les statistiques, la quantité d'information et les propriétés énergétiques des signaux à classer. Ils sont détaillés plus en détail dans la partie consacrée à l'identification.

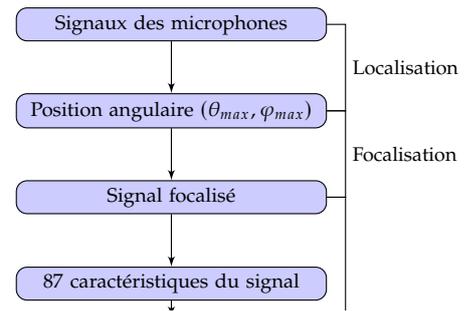
Couplage des deux objectifs : dans cette thèse, l'innovation majeure réside dans le fait de coupler le traitement d'antenne et l'apprentissage statistique. Ce couplage permet dans un premier temps de tirer profit du filtrage spatial qu'offre le traitement d'antenne avant de passer à l'étape d'identification du drone. Les signaux acoustiques à identifier possèdent ainsi un meilleur rapport signal à bruit, ce qui aide à leur identification. Dans un second temps, la localisation unique du drone est rendue possible en tenant compte du résultat d'identification lors de la reconstruction finale de la trajectoire de la menace. Les sources perturbatrices sont ainsi ignorées, évitant la génération de fausses alarmes.

La description de la chaîne de traitement proposée dans la partie suivante permet d'appréhender ce couplage lié au choix dans l'enchaînement des méthodes sélectionnées, et qui améliore les performances en identification par rapport aux méthodes vues dans la synthèse bibliographique qui voient les deux étapes de manière totalement indépendantes.

4.2 Description de la chaîne de traitement

La chaîne générale est décrite en Figure 4.1 et contient trois étapes successives :

- ▶ la **localisation** durant laquelle un algorithme de traitement d'antenne est utilisé pour localiser les sources acoustiques présentes autour de l'antenne utilisée ;
- ▶ la **focalisation** qui effectue un filtrage spatial dans les directions pour lesquelles des sources acoustiques ont été trouvées, afin d'améliorer le rapport signal à bruit de ces sources par rapport aux signaux provenant d'autres directions ;
- ▶ l'**identification** pendant laquelle des caractéristiques sont calculées sur le signal focalisé, pour ensuite utiliser un algorithme d'apprentissage statistique afin de classer ce signal comme drone ou autre bruit.



Afin d'obtenir une localisation et une identification au cours du temps, les signaux enregistrés à une fréquence d'échantillonnage de 51,2 kHz sont découpés en **séquences** successives comme illustré en Figure 4.2. Une durée de **200 ms** est choisie, soit une distance de 3 m parcourue pour un drone volant à une vitesse élevée de 15 m/s, ce qui assure une trajectoire bien échantillonnée spatialement au cours du temps.

La suite de cette partie détaille les trois étapes de localisation, focalisation et identification, qui sont appliquées sur chacune des séquences successives de 200 ms de signal enregistrées par l'antenne acoustique.

Localisation

L'objectif de l'étape de localisation est de fournir les positions des sources acoustiques présentes autour de l'antenne. Il est important de préciser qu'à cette étape, toutes les sources acoustiques sont considérées, qu'elles correspondent à la classe drone ou non.

Dans cette application en air, et pour la taille caractéristique de l'antenne utilisée de 60 cm, la distance D_{lim} donnée en Equation 2.11 est égale à 2,1 m pour une source émettant à une fréquence de 1 kHz. Pour des applications industrielles où l'objectif est d'émettre une alerte le plus tôt possible, la localisation et l'identification de la source est souhaitée bien au-delà de cette distance. Le modèle de propagation est donc choisi en **ondes planes**, impliquant une recherche des sources par leur direction d'arrivée dans le domaine angulaire (θ, φ) , avec θ l'angle d'élévation par rapport à l'horizontale, et φ l'angle d'azimut autour du centre du repère de coordonnées.

Concrètement, l'antenne acoustique utilisée, qui est plane, est positionnée parallèlement au sol. A partir de la séquence courante de 200 ms enregistrée par les M microphones de l'antenne, une **matrice interspectrale** est estimée grâce au périodogramme de Welch comme décrit en équation Equation 2.4, avec un nombre de blocs L égal à 10, soit une résolution fréquentielle de 50 Hz, et un recouvrement de 50 %. L'indicateur I_{MUSIC}^{dB} est ensuite calculé sur une **carte hémisphérique**,

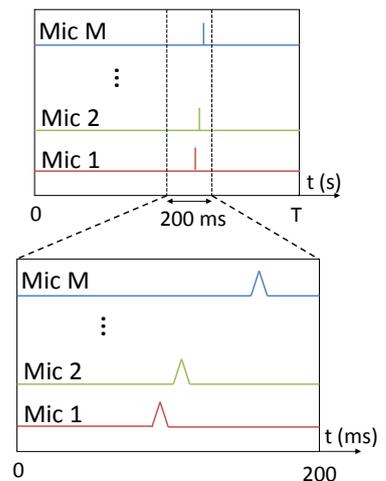


FIGURE 4.2 – Illustration du découpage des signaux enregistrés en séquences de 200 ms.

qui correspond à la partie de l'espace situé au-dessus du plan de l'antenne, dont un exemple est donné en Figure 4.3. Les angles d'élévation et d'azimut sont référencés par rapport au centre de l'antenne, et un système permettant de localiser la source de manière indépendante au cours du temps est utilisé pour établir la vérité terrain. En Figure 4.3, elle est donnée de manière absolue dans l'espace par le point noir ●. La carte de MUSIC étant angulaire, ce sont les directions d'arrivée données par les traits pleins qui doivent être comparées et non les positions absolues⁴.

Le **nombre de valeurs propres** sélectionnées pour l'espace signal correspond au nombre de sources acoustiques attendues dans la configuration étudiée : une ou deux en fonction des cas détaillés dans la Partie 4.3. Il faut noter que ce choix peut être optimisé en étudiant plus précisément la décroissance des valeurs propres, mais cette optimisation n'a pas été mise en place.

La **fréquence de traitement** est égale à 1 kHz en intérieur et 5 kHz en extérieur. Pour l'intérieur, ce choix découle du fait qu'un des exemples de sons destinés à caractériser la classe autre bruit est un sinus pur à une fréquence de 1 kHz. Afin de traiter tous les signaux sur un même pied d'égalité en localisation, sans donner d'avantage aux signaux de drone en se plaçant sur une de leurs harmoniques par exemple, c'est cette fréquence de 1 kHz qui est sélectionnée comme fréquence de localisation. Pour l'extérieur, le choix d'une localisation à 5 kHz est basé sur l'étude des erreurs de localisation, détaillée en Partie 4.6.

Là encore, ces choix pourraient être optimisés : localiser à plusieurs fréquences permettrait de moyenniser ensuite les résultats pour éviter les erreurs de localisation. Cette piste de travail n'a pas été étudiée dans la thèse, pour se concentrer sur la compréhension des différentes étapes de traitement et de leur enchaînement.

Pour terminer, la position de source est estimée en prenant le **maximum de la carte** obtenue dans le cas d'une source unique recherchée, et des deux premiers maxima pour deux sources recherchées. Aucun post-traitement supplémentaire n'est effectué sur les cartes de MUSIC.

Focalisation

L'objectif de l'étape de focalisation est de filtrer spatialement les signaux provenant de la (ou des) direction(s) déterminée(s) à l'étape de localisation, afin d'isoler le signal temporel de la source acoustique localisée en cette direction, et améliorer son rapport signal à bruit pour aider à son identification.

Elle est schématisée en Figure 4.4 et utilise le maximum, ou successivement sur les deux maxima, estimés à l'étape de localisation. Pour le **maximum courant** situé en $(\theta_{max}, \varphi_{max})$, la première phase de la focalisation consiste à **retarder les signaux** de l'antenne par rapport à un microphone de référence⁵ d'une durée égale pour le microphone i à :

$$\Delta t_i = \frac{\Delta D_i}{c_0} = \frac{\mathbf{v}_{max} \cdot \mathbf{u}_i}{c_0}, \quad (4.1)$$

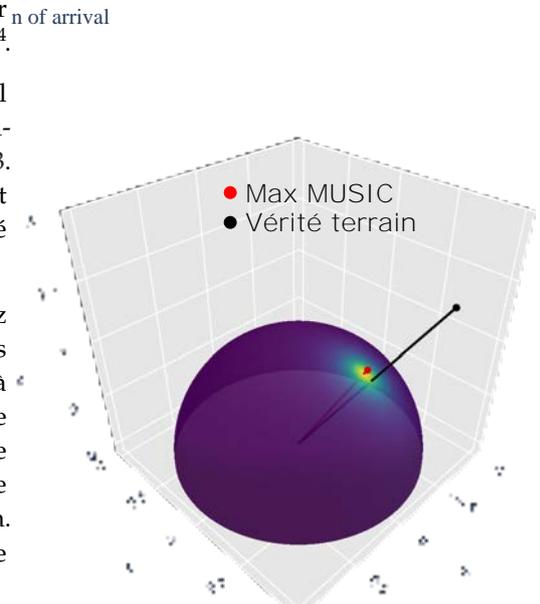


FIGURE 4.3 – Exemple de carte angulaire calculée pour l'indicateur I_{MUSIC}^{dB} . Le maximum de la carte $(\theta_{max}, \varphi_{max})$ est indiqué par le trait (-) et le point rouge (●). Le trait noir (-) et le point noir (●) représentent la vérité terrain.

5: Souvent pris par défaut comme le microphone numéroté 1 sur l'antenne.

avec $\mathbf{v}_{max} = [\cos(\theta_{max}) \cos(\varphi_{max}), \cos(\theta_{max}) \sin(\varphi_{max}), \sin(\theta_{max})]^T$ le vecteur d'onde plane dans la direction du maximum, et $\mathbf{u}_i = \overrightarrow{M_1 M_i}$ le vecteur entre le microphone 1 de référence et le microphone i .

Une fois les retards déterminés, plusieurs approches ont été envisagées pour les compenser. La première est d'arrondir le retard pour chaque voie au plus proche échantillon temporel et de décaler le signal de ce nombre d'échantillons, mais ceci empêche la remise en phase exacte des signaux. Le deuxième est d'interpoler les signaux temporels pour avoir une remise en phase exacte, mais elle est coûteuse en temps de calcul. Enfin la troisième, qui allie temps de calcul rapide et remise en phase exacte, exploite le fait qu'un retard temporel est équivalent à un **déphasage fréquentiel**. Ainsi, la voie i à retarder est passée dans le domaine fréquentiel par transformée de Fourier, puis déphasée d'une phase :

$$\Delta\phi_i = e^{-2i\pi\Delta t_i f}, \quad (4.2)$$

pour toutes les fréquences f de la bande $[0 \text{ Hz}; f_e / 2 \text{ Hz}]$ avec f_e la fréquence d'échantillonnage. Le signal résultant est repassé dans le domaine temporel par transformée de Fourier inverse, pour obtenir le signal retardé.

Pour assurer la remise en phase des signaux enregistrés par les deux microphones les plus éloignés sur l'antenne sans problème de cyclicité sur les phases, un délai supplémentaire est pris en compte lors de la sélection de la séquence courante. L'espacement maximum entre deux microphones étant de 60 cm, le retard maximum sur l'antenne est inférieur à 2 ms. Des morceaux de 5 ms de signal sont donc ajoutés de part et d'autre de la séquence avant sa localisation et sa focalisation, puis supprimés une fois la transformée de Fourier inverse effectuée.

Les M signaux remis en phase sont ensuite **sommés** pour obtenir le **signal temporel filtré spatialement dans la direction du maximum courant**, dont le rapport signal à bruit est augmenté grâce à la focalisation. Une séquence unique de signal de 200 ms au départ pour les M voies de mesures produit alors une ou deux séquence(s) focalisée(s) en fonction du nombre de sources acoustiques à localiser dans la configuration étudiée. L'ensemble de ces séquences focalisées constituent les signaux d'entrée de l'étape d'identification décrite ci-après.

Identification

L'objectif de l'étape d'identification est de classer la séquence focalisée courante comme faisant partie de la classe drone ou autre bruit.

Etant donné le choix des modèles SVM et SVM-1C qui fonctionnent avec des vecteurs d'observation non ordonnés, il est nécessaire de calculer des **caractéristiques** sur la séquence focalisée de 200 ms d'entrée comme représenté en Figure 4.5. Des travaux sur la classification automatique de signaux acoustiques [Mal18] ont permis de sélectionner 32 caractéristiques particulièrement performantes pour la classification de signaux volcaniques [Mal+18b; Mal+17a; Mal+18a; Mal+17b] et pour classer des poissons par famille en acoustique sous-marine [Mal+18c; Mal+18d].

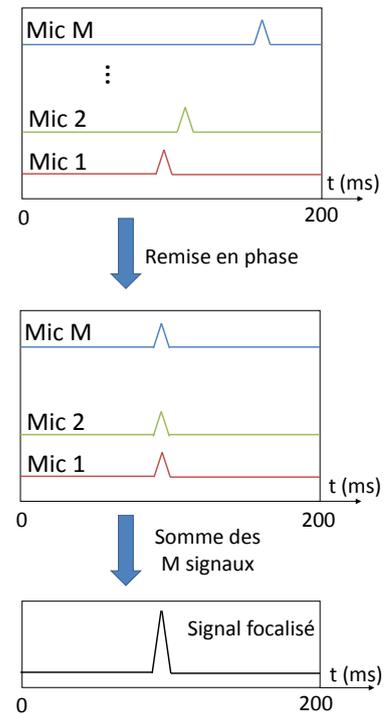


FIGURE 4.4 – Illustration de la focalisation des signaux durant laquelle les signaux de l'antenne sont remis en phase et sommés.

Dans le cadre de cette thèse, 29 d'entre elles sont conservées⁶ et peuvent être résumées en trois familles⁷ :

Caractéristiques utilisées pour l'identification

- ▶ Les caractéristiques liées à la **forme** du signal : le centroïde, la bande d'énergie autour du centroïde, le taux d'attaque, le taux d'affaiblissement, le taux de croisement (réglé avec trois seuils différents), le taux de silence (réglé avec trois seuils différents).
- ▶ Les caractéristiques liées aux **statistiques** du signal : le minimum, le maximum, la moyenne, le minimum sur la moyenne, le maximum sur la moyenne, la *skewness* et le *kurtosis* et leur version centrée sur le centroïde, l'écart-type.
- ▶ Les caractéristiques liées à la **quantité d'information** portée par le signal : l'entropie de Shannon (avec trois réglages différents) et l'entropie de Rényi (avec six réglages différents).

Ces caractéristiques sont calculées grâce au module Python AAA, disponible sur Github [Mal+], sur les séquences focalisées qui ont été normalisées en énergie pour s'assurer une indépendance par rapport à la distance entre la source à identifier et l'antenne. Ce calcul est effectué dans trois domaines de travail : temporel, spectral et cepstral. La description des signaux mesurés dans ces trois domaines permet d'obtenir des résultats de classification supérieurs à 90 % dans les deux cas expérimentaux évoqués précédemment [Mal+18b ; Mal+18c].

Pour les premiers cas d'application, une **observation** est alors un vecteur de dimension $N_c = 87$, composé des valeurs des 29 caractéristiques calculées dans les trois domaines temporel, spectral et cepstral à partir d'une séquence focalisée donnée. La pertinence de la description des signaux de drone par ces caractéristiques sera discutée en Partie 4.5.

En fonction des cas d'application, tout ou partie de ces observations sera utilisé pour apprendre ou tester les modèles d'apprentissage statistique sélectionnés : le **SVM avec un noyau linéaire** et le **SVM-1C avec un noyau gaussien**. Chacun des cas détaillera précisément les données utilisées que ce soit en phase d'apprentissage, en validation croisée ou en phase de test.

Note : il est important de noter que pour tous ces cas, une attention particulière a été prêtée au centrage et à la réduction des données : ce sont les données d'apprentissage uniquement qui permettent d'établir les paramètres de cette opération, soit leur moyenne et variance dans l'espace des caractéristiques. Les données de test, que ce soit en validation croisée ou en test réel, sont translatées et réduites en accord avec ces paramètres préalablement établis, dans une seconde phase qui est indépendante. Sans cette attention particulière, les résultats en phase de test sont biaisés car le centrage et la réduction pour l'ensemble des données à la fois leur assurent un certain placement dans l'espace des caractéristiques, qui peut aider à l'apprentissage sous certaines conditions.

Pour **régler les hyperparamètres** des deux modèles et **évaluer** leurs performances, le contexte de détection dans lequel cette application s'inscrit oriente les **métriques** à utiliser. En détection, chaque observation est classée comme appartenant ou non à une classe cible, qui correspond

6: Les trois caractéristiques relatives à l'énergie du signal n'étaient pas pertinentes pour des signaux normalisés en énergie comme c'est le cas dans cette thèse.

7: Voir la liste détaillée en Annexe B.

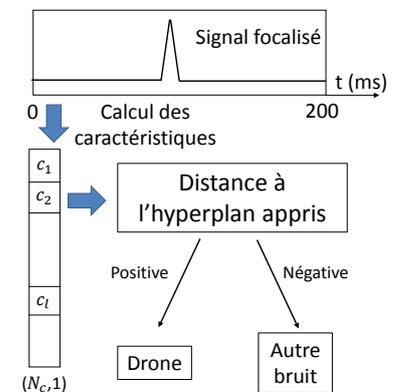


FIGURE 4.5 – Illustration de l'identification des séquences. Les caractéristiques sont calculées sur les séquences focalisées pour donner un vecteur d'observation dont la distance signée à l'hyperplan établi durant la phase d'apprentissage est calculée. En fonction de son signe, l'observation est identifiée comme appartenant à la classe drone (distance positive), ou à la classe autre bruit (distance négative).

ici à la classe drone. L'autre classe qui correspond à une observation ne comportant pas de cible est appelée autre bruit pour l'ensemble de ce chapitre. En utilisant les notations de la matrice de confusion (Tableau 4.1) dont les lignes représentent les classes réelles et les colonnes représentent les classes prédites, les trois métriques utilisées sont alors :

- la **précision** qui représente le nombre d'observations bien classées parmi toutes les observations disponibles :

$$P_r = \frac{VP + VN}{VP + FP + FN + VN} ; \quad (4.3)$$

- le **taux de détection** qui donne le nombre d'observations bien prédites parmi toutes les observations de la classe drone :

$$T_d = \frac{VP}{VP + FN} ; \quad (4.4)$$

- le **taux de fausse alarme** qui donne le nombre d'observations anormalement prédite parmi toutes les observations de la classe autre bruit :

$$T_{fa} = \frac{FP}{FP + VN} . \quad (4.5)$$

Le réglage des deux modèles consiste alors à regarder la variation de ces trois métriques pour plusieurs valeurs de leurs hyperparamètres et sélectionner celui qui produit le taux de détection le plus haut pour le taux de fausse alarme le plus bas, dont un bon compromis est associé à la valeur maximale de la précision.

Pour le modèle SVM, les trois métriques atteignent un plateau pour des valeurs pour C supérieures à 0,01, quels que soient les cas d'application. Le **paramètre C** est donc choisi **égal à 10** pour tous les cas d'applications qui suivent pour s'assurer d'être sur le plateau, et en même temps ne pas prendre des valeurs trop grandes pour C qui régulariseraient trop le problème à optimiser. Son noyau étant linéaire, aucun hyperparamètre ne lui est associé.

Pour le paramètre ν du SVM-1C, les trois métriques voient aussi leur performance stable pour des valeurs comprises entre 0,0001 et 0,1. Ce **paramètre ν est donc fixé à 0,001** pour tous les cas d'application. Pour le **paramètre γ** qui contrôle la largeur du noyau gaussien utilisé, les trois métriques varient plus fortement selon les cas d'application. Pour rendre compte de cette variabilité dans les résultats, **50 modèles SVM-1C** sont appris pour chaque cas d'application qui l'utilise, pour 50 valeurs de γ espacées logarithmiquement entre 10^{-5} et 10.

Les performances du modèle SVM et des 50 modèles SVM-1C générés à chaque cas sont ensuite comparées dans le plan taux de détection en fonction du taux de fausse alarme, pour rendre compte de manière synthétique de leur capacité de détection de la classe cible en même temps que celle de réjection des fausses alarmes.

La partie suivante détaille l'ensemble des données créées durant la thèse qui vont servir à valider les trois étapes proposées pour la localisation et l'identification de drone dans divers cas d'application.

TABLEAU 4.1 – Exemple de matrice de confusion. Les classes réelles sont données par les lignes et les classes prédites par les colonnes. Chaque case donne alors : les vrais positifs (VP), les faux positifs (FP), les faux négatifs (FN) et les vrais négatifs (VN).

		Classe prédite	
		Drone	Autre bruit
Classe réelle	Drone	VP	FN
	Autre bruit	FP	VN

4.3 Bases de données expérimentales produites

Durant les trois années de thèse, trois sessions de mesures ont été effectuées dont deux en intérieur et une en extérieur. Grâce à une répartition régulière tout au long de la thèse, ces différents essais ont permis d'améliorer les protocoles à chaque nouvelle session, pour obtenir des données spécifiquement dédiées à la résolution des problèmes rencontrés au fur et à mesure des développements.

L'antenne utilisée pour l'ensemble de ces essais est une Simcenter Sound Camera Digital Array [Sie], conçue et produite par MicrodB et distribuée par Siemens PLM Software. Elle est utilisée avec des extensions SMB (*Short Mother Board*), et possède ainsi 81 microphones MEMS répartis sur un diamètre de 60 cm. Une photographie de l'antenne en fonctionnement lors des essais extérieurs est montrée en Figure 4.6.

Les deux parties suivantes listent les données produites dans les environnements intérieur et extérieur, acquises selon des protocoles différents.

Environnement intérieur

Les données intérieures ont été acquises au sein de la *salle de capture de mouvement du Gipsa-Lab*. Elle est équipée d'un système de 12 caméras infrarouges Vicon qui permettent de localiser en temps réel des objets sur lesquels des réflecteurs auront été préalablement fixés [Mer+17]. Ce système permet de positionner les éléments souhaités en trois dimensions dans la salle, avec une précision de l'ordre du centimètre. Grâce à ce système, les drones testés et l'antenne sont précisément positionnés, et une vérité terrain fiable est disponible pour valider la méthode de localisation.

Un axe de recherche de l'équipex Robotex du Gipsa-Lab consiste à contrôler les trajectoires de drone grâce au système composé des caméras Vicon. Afin de produire des trajectoires précises et reproductibles, les vols de drone effectués dans cette thèse bénéficient de cette technologie. Les trajectoires sont ainsi décidées en amont et entrées dans le logiciel de commande qui pilote ensuite le drone une fois celui-ci démarré. En fonction des besoins, les trajectoires ont donc pu être adaptées pour mettre en difficulté les méthodes de localisation et d'identification sélectionnées, pour aboutir à la liste suivante :

- ▶ une trajectoire linéaire : cette trajectoire, représentée en Figure 4.7, est la première à avoir été implémentée. Elle alterne des vols stationnaires avec une montée et un décalage, effectués dans un plan vertical passant le centre de l'antenne. Elle possède l'avantage d'avoir été parcourue par plusieurs drones différents, et sera donc utilisée pour étudier la pertinence des caractéristiques retenues pour l'identification des drones ;
- ▶ cinq cercles successifs espacés de 20 cm en hauteur : la trajectoire linéaire ne permettait pas d'étudier la performance de l'estimation en azimut, cette trajectoire illustrée en Figure 4.8 le permet grâce à des cercles effectués autour de l'antenne. Les différentes hauteurs

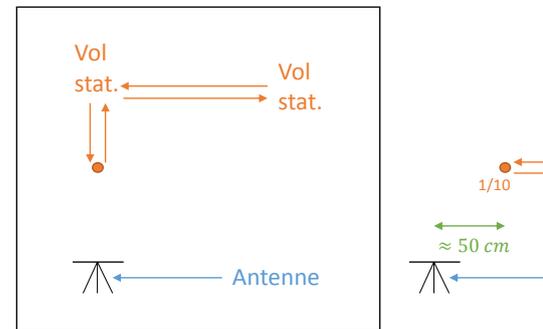


FIGURE 4.7 – Schéma de la trajectoire linéaire suivie par le drone.

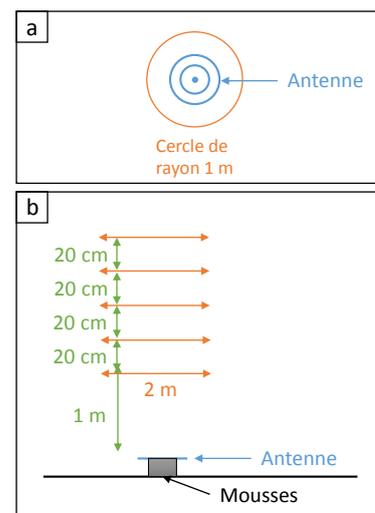


FIGURE 4.8 – Schéma de la trajectoire où le drone vole selon cinq cercles successifs espacés de 20 cm en hauteur. a : vue de dessus. b : vue de côté.

permettent d'étudier les performances d'estimation de l'angle d'élévation ;

- ▶ trois lignes droites à deux hauteurs successives : pour étudier les performances en identification, cette trajectoire schématisée en Figure 4.9 a été effectuée avec un haut-parleur fixe qui émettait des sons (ou non) en même temps que le drone volait. Les trois lignes droites sont faites à la droite, au-dessus, et à la gauche de l'antenne, ce qui permet d'obtenir un cas où une source perturbatrice est présente, avec une large gamme de distances entre le drone et l'antenne, et par conséquent de nombreuses différences de niveaux sonores entre le drone et le haut-parleur.

Le matériel utilisé pour créer les données est constitué de trois drones représentés en Figure 4.10 aux côtés de l'antenne. Ils sont tous propulsés par quatre moteurs, avec pour le plus petit (un drone de loisir) du courant continu et des quadripâles de 3 pouces (D3), et pour les deux autres des moteurs sans balais équipés de bipâles de 5 pouces pour celui de taille moyenne (D5) et de 6 pouces pour le plus grand (D6).

Tous les essais effectués sont listés dans le Tableau 4.2 pour les données comprenant des vols de drone, et dans le Tableau 4.3 pour celles contenant uniquement des sons joués par le haut-parleur. Pour chaque configuration, un identifiant est donné pour aider à leur référencement dans la suite du chapitre. Pour certaines configurations, le niveau sonore du haut-parleur a été réglé sur trois volumes différents, indiqués par ordre croissant par les valeurs 1 à 3 indiquées entre parenthèses.

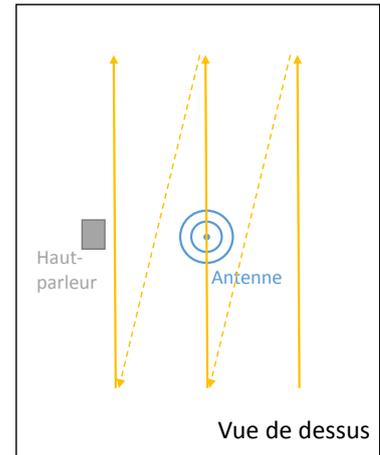


FIGURE 4.9 – Schéma de la trajectoire où le drone suit trois lignes droites à deux hauteurs successives. Cette trajectoire vue de dessus est répétée deux fois par vol.

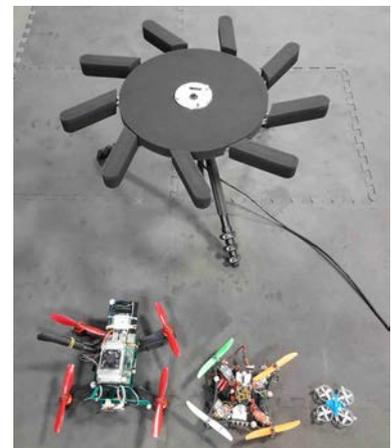


FIGURE 4.10 – Photographie des trois drones de différentes tailles utilisés en intérieur.

Environnement extérieur

En extérieur, les données ont été acquises dans une clairière près du lac de Freydières, dans la commune de Revel. Pour positionner les différents objets à l'extérieur, un système GPS RTK (*Real Time Kinematic*) est employé. Sa technologie est basée sur les signaux GPS, mais pour améliorer leur précision de l'ordre du mètre, les éléments ne sont pas positionnés de manière absolue mais de manière relative par rapport à une base pour atteindre des précisions de quelques dizaines de centimètres. En pratique durant les mesures, la précision pour la coordonnée z , moins bien gérée par les signaux GPS, est restée de l'ordre du mètre.

TABLEAU 4.2 – Liste des données intérieures créées durant la thèse avec la présence d'un drone.

Type de trajectoire	Drone	Son HP (volume)	Répétitions	Temps de mesure	Nombre de séquences	Identifiant
Linéaire	D3	/	15	12 min 45,2 s	3826	Int-Lin-D3
Linéaire	D5	/	12	9 min 51,2 s	2956	Int-Lin-D5
Linéaire	D6	/	7	6 min 7,8 s	1839	Int-Lin-D6
5 cercles	D6	/	5	6 min 21,8 s	1909	Int-5c-D6
3 lignes droites	D6	/	6	8 min 57,6 s	2688	Int-3ld-D6
		Bruit blanc (1)	1	1 min 36,4 s	482	Int-3ld-D6-bb1
3 lignes droites	D6	Bruit blanc (2)	1	1 min 34,8 s	474	Int-3ld-D6-bb2
		Bruit blanc (3)	1	1 min 36,2 s	481	Int-3ld-D6-bb3
		Sinus 1 kHz (1)	1	1 min 38,4 s	492	Int-3ld-D6-s1k1
3 lignes droites	D6	Sinus 1 kHz (2)	1	1 min 36,6 s	483	Int-3ld-D6-s1k2
		Sinus 1 kHz (3)	2	2 min 25,8 s	729	Int-3ld-D6-s1k3
		Voiture (1)	1	1 min 35,8 s	479	Int-3ld-D6-voit1
3 lignes droites	D6	Voiture (2)	1	1 min 36,4 s	482	Int-3ld-D6-voit2
		Voiture (3)	1	1 min 42,6 s	513	Int-3ld-D6-voit3

TABLEAU 4.3 – Liste des sons enregistrés durant la création les données intérieures durant la thèse.

Son HP (volume)	Répétitions	Temps de mesure	Nombre de séquences	Identifiant
Sinus 1 kHz	3	1 min 33,4 s	467	Int-ND-s1k
Bruit blanc	3	1 min 33,8 s	469	Int-ND-bb
Perceuse	3	1 min 34,6 s	473	Int-ND-perc
Hélicoptère 1	3	1 min 35,2 s	476	Int-ND-hel1
Hélicoptère 2	3	1 min 34,2 s	471	Int-ND-hel2
Moteur	3	1 min 34,6 s	473	Int-ND-mot
Bruit blanc (1/2/3)	1	3 × 30 s	3 × 150	Int-ND-bb1/2/3
Sinus 1 kHz (1/2/3)	1	3 × 30 s	3 × 150	Int-ND-s1k1/2/3
Voiture (1/2/3)	1	3 ×		

Un seul type de vol a été mis en place sur la journée de mesure. Il s'agit de vols en cercles autour de l'antenne comme représenté dans le schéma de la Figure 4.11, à différentes hauteurs. Pour ces vols, le drone était piloté manuellement, ce qui implique des trajectoires beaucoup moins reproductibles qu'en intérieur. Un haut-parleur était situé à 7,4 m au nord et 1,2 m à l'est de l'antenne, et à 2,7 m de hauteur, afin de jouer des sons à deux niveaux sonores différents, pour représenter la classe autre bruit. Il sert aussi de source perturbatrice pour évaluer les méthodes lorsque le drone n'est plus seul.

Le drone D6 est utilisé et montré en Figure 4.12, avec une adaptation du matériel qu'il transporte afin de pouvoir être reconnu par le système de positionnement GPS RTK. Son point de démarrage est toujours situé au même endroit, dans la zone appelée *home*, afin d'établir les positions de l'antenne et du haut-parleur de manière relative à ce point de départ. Ainsi les dérives des mesures GPS qui ont lieu durant la journée sont compensées, pour obtenir des positions fiables des éléments au cours du temps.

Les données enregistrées tout au long de la journée sont listées dans le Tableau 4.4, avec leur identifiant respectif.

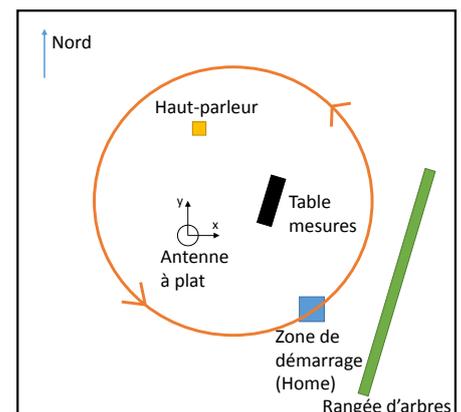


FIGURE 4.11 – Schéma du positionnement des différents éléments et de la trajectoire suivie par le drone durant les essais extérieurs.

TABLEAU 4.4 – Liste des données extérieures créées durant la thèse.

Type de trajectoire	Drone	Son HP	Répétitions	Temps de mesure	Nombre de séquences	Identifiant
Vols en cercles	D6	/	2	10 min	3000	Ext-D6
Vols en cercles	D6	Bruit blanc (vol 1)	1	1 min 35 s	475	Ext-D6-bb1
		Bruit blanc (vol 2)	1	1 min 46,6 s	533	Ext-D6-bb2
Vols en cercles	D6	Voiture (vol 1)	1	1 min 32,4 s	462	Ext-D6-voit1
		Voiture (vol 2)	1	1 min 26,8 s	434	Ext-D6-voit2
Pas de vol	/	Bruit blanc (vol 1)	1	16 s	80	Ext-ND-bb1
		Bruit blanc (vol 2)	1	20 s	100	Ext-ND-bb2
Pas de vol	/	Voiture (vol 1)	1	22 s	110	Ext-ND-voit1
		Voiture (vol 2)	1	20 s	100	Ext-ND-voit2
Pas de vol	/	Sinus 1 kHz	1	20 s	100	Ext-ND-s1k
Pas de vol	/	Sinus 3 kHz	1	20 s	100	Ext-ND-s3k
Pas de vol	/	Perceuse	1	20 s	100	Ext-ND-perc
Pas de vol	/	Hélicoptère 1	1	20 s	100	Ext-ND-hel1
Pas de vol	/	Hélicoptère 2	1	20 s	100	Ext-ND-hel2
Pas de vol	/	Moteur	1	20 s	100	Ext-ND-mot

A partir de cet ensemble de mesures, les parties suivantes valident puis discutent les performances des méthodes sélectionnées dans différents cas de figure applicatifs.

4.4 Validation de la localisation et de l'identification pour des sources seules en intérieur

L'objectif de cette première partie de résultats est de valider la chaîne de traitement proposée sur des données intérieures, où une seule source est présente autour de l'antenne à la fois et pour laquelle une vérité terrain très précise en localisation est disponible.

Validation de la localisation : la première étape du traitement à valider est la localisation. Celle-ci est réalisée grâce à l'ensemble des données Int-5c-D6, dont la trajectoire a été spécialement conçue pour faire varier les deux angles de localisation. Pour visualiser qualitativement les performances en localisation de la méthode MUSIC, la Figure 4.13 représente la trajectoire angulaire du drone au cours de son vol pour le système composé des caméras Vicon donnant la vérité terrain en vert, et celle reconstruite en prenant les maxima successifs des cartes de MUSIC calculées au cours du temps en rouge. Pour des raisons de clarté, seul le cercle le plus bas ($\theta = 36^\circ$) et le cercle le plus haut ($\theta = 57^\circ$) sont représentés⁸.

Visuellement, la correspondance entre l'estimation par MUSIC et la vérité terrain est bonne pour le cercle à 57° . La trajectoire est bien suivie avec une légère surestimation de l'angle d'élévation. Pour le cercle à 36° , la correspondance est nettement moins bonne, avec des variations fortes de l'angle d'élévation estimé par le maximum de la carte de MUSIC. Ceci amène à une première observation partagée par les travaux consacrés à



FIGURE 4.12 – Photographie du drone utilisé pour les essais extérieurs.

8: Une version 3D est disponible sur cette [page web](#).

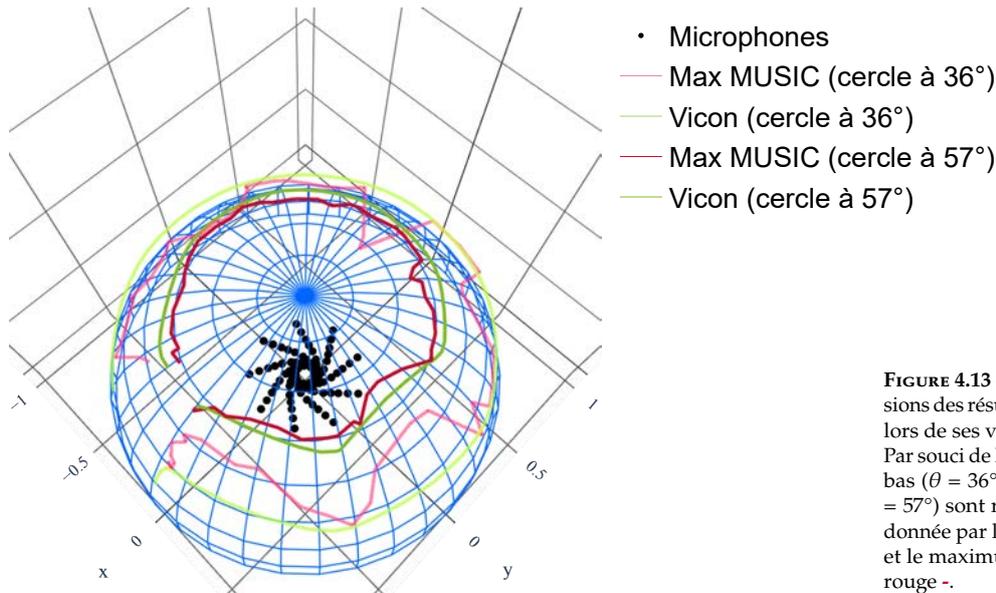


FIGURE 4.13 – Illustration en trois dimensions des résultats de localisation du drone lors de ses vols en cinq cercles successifs. Par souci de lisibilité, seuls le cercle le plus bas ($\theta = 36^\circ$), et le cercle le plus haut ($\theta = 57^\circ$) sont représentés. La vérité terrain donnée par les caméras Vicon est en vert et le maximum de la carte de MUSIC en rouge.

la localisation de drone [Ram19; Bla19] : la précision en localisation selon l'angle d'élévation varie en fonction de l'angle d'élévation lui-même.

Pour quantifier ces tendances observées qualitativement, la Figure 4.14 donne les erreurs en degrés entre l'estimation des angles d'élévation et d'azimut pour les maxima des cartes de MUSIC calculées en fonction du temps et la vérité terrain, pour chacun des cinq cercles effectués par le drone.

Grâce au système de contrôle de la trajectoire utilisé, l'ensemble des répétitions peuvent être synchronisées pour établir la moyenne et l'écart-type⁹ des erreurs obtenues pour les deux angles de localisation. Les erreurs sont données en valeur réelle, pour montrer le comportement différent entre l'angle d'azimut et l'angle d'élévation, mais les moyennes ensuite énoncées sont prises sur les valeurs absolues de ces erreurs.

L'erreur sur l'angle d'azimut φ varie autour de 0 et est en valeur absolue de $4 \pm 3^\circ$, sur l'ensemble des séquences localisées au cours des cinq répétitions du vol entier. De plus les performances selon cet angle sont globalement stables au cours de la trajectoire, même si une amélioration est visible pour le cercle le plus haut en moyenne, mais pas en écart-type.

Pour l'angle d'élévation θ , l'erreur moyenne sur l'ensemble des vols est de $8 \pm 6^\circ$ et elle varie selon la position du drone en élévation. Toutes ces erreurs sont systématiquement positives ce qui traduit une surestimation de l'angle. Au début de la trajectoire, i.e. pour l'angle d'élévation le plus bas de 36° , les erreurs sont de $13,1^\circ$ en moyenne contre seulement $3,2^\circ$ en fin de trajectoire pour l'angle d'élévation le plus élevé de 57° . L'écart-type de cette erreur décroît elle aussi avec l'angle d'élévation qui augmente, mais moins fortement en passant de $8,5^\circ$ à $2,4^\circ$ de $\theta = 36^\circ$ à $\theta = 57^\circ$.

Les réflexions sur le sol et le plafond sont à l'origine de ces erreurs plus importantes en élévation (étude sur ce sujet en Partie 4.6). La géométrie plane de l'antenne constitue un autre problème pour cet angle, avec des ondes acoustiques rasantes pouvant être diffractées sur la structure de l'antenne.

9: N'étant basé que sur cinq répétitions pour le tracé, le module Python Seaborn utilise une stratégie de *bootstrapping* pour établir ses intervalles de confiance.

Localisation intérieur : 1 source

Erreurs moyennes globales :

- ▶ En azimut φ : $4 \pm 3^\circ$
- ▶ En élévation θ : $8 \pm 6^\circ$

Erreurs moyennes vol rasant :

- ▶ En azimut φ : $4,7 \pm 4,2^\circ$
- ▶ En élévation θ : $13,1 \pm 8,5^\circ$

Erreurs moyennes vol en surplomb :

- ▶ En azimut φ : $0,2 \pm 3,4^\circ$
- ▶ En élévation θ : $3,2 \pm 2,4^\circ$

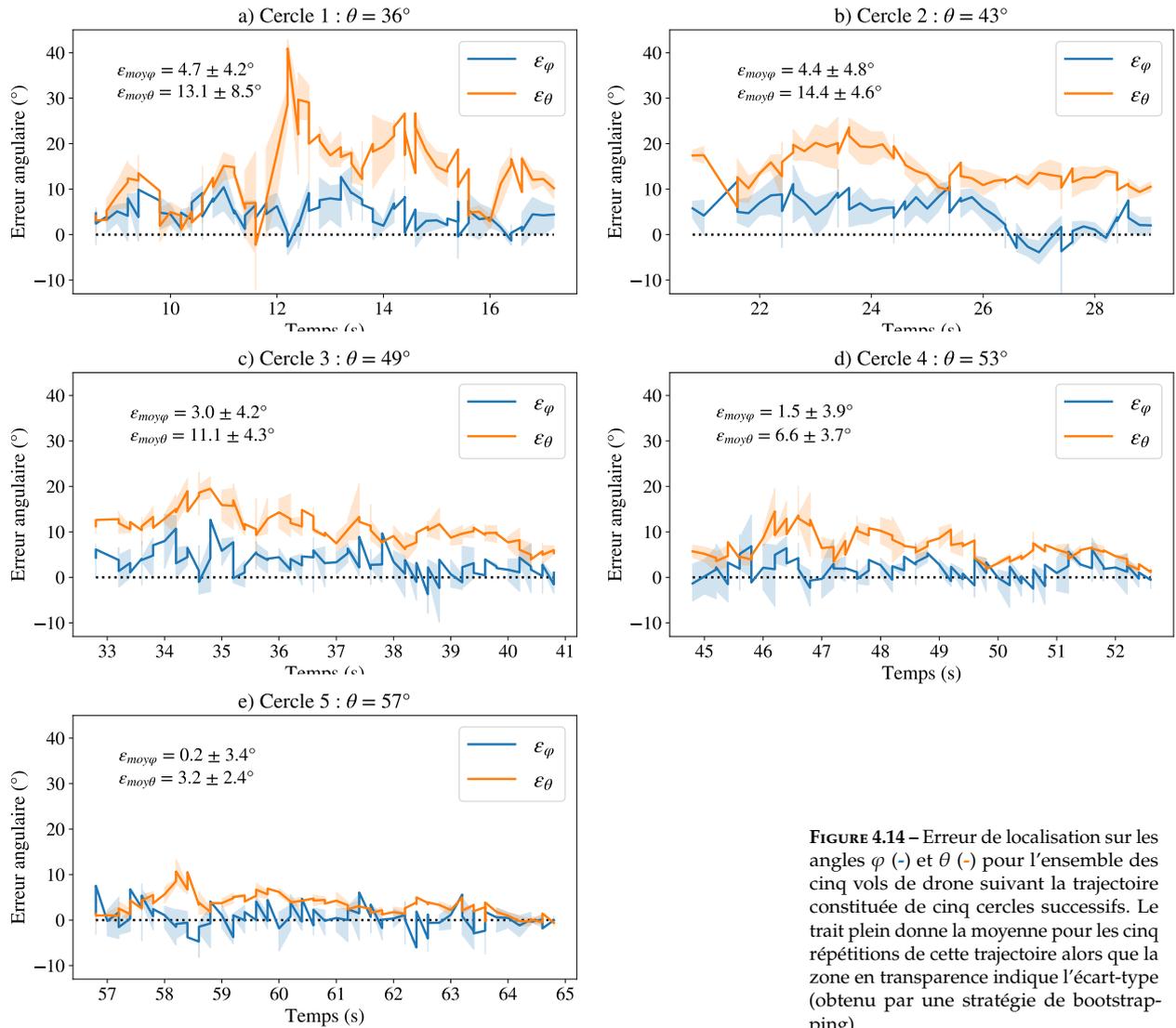


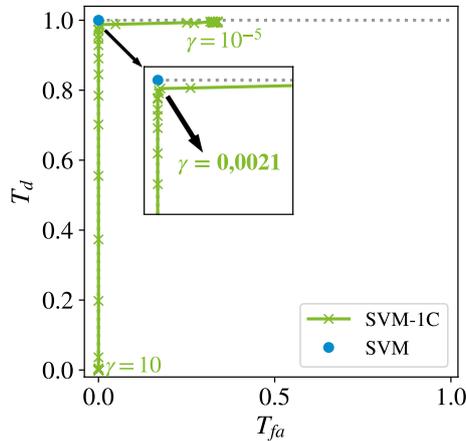
FIGURE 4.14 – Erreur de localisation sur les angles φ (-) et θ (-) pour l'ensemble des cinq vols de drone suivant la trajectoire constituée de cinq cercles successifs. Le trait plein donne la moyenne pour les cinq répétitions de cette trajectoire alors que la zone en transparence indique l'écart-type (obtenu par une stratégie de bootstrapping).

Validation de l'identification : pour étudier ensuite les performances en identification, les données Int-5c-D6 précédemment localisées représentent la classe drone. Pour la classe autre bruit, ce sont les sons Int-ND-bb1/2/3, Int-ND-s1k1/2/3 et Int-ND-voit1/2/3 qui sont sélectionnés. Afin d'avoir en apprentissage un jeu de données équilibré, seules 1260 observations de chacune des deux classes sont conservées, avec une égale répartition pour chaque son et chacun de ses volumes de 140 observations.

Afin d'établir les performances en identification sur l'ensemble du jeu de données, qui sert à la fois pour l'apprentissage et pour le test, une validation croisée est composée de 10 plis est mise en place, avec pour chacun d'entre eux 70 % des données sélectionnées aléatoirement pour l'apprentissage et les 30 % restantes sont utilisées pour la phase de test.

Pour chaque pli, le modèle SVM linéaire avec le paramètre C égal à 10 et les 50 modèles SVM-1C avec $\nu = 0,001$ et γ variant entre 10^{-5} et 10 sont générés comme expliqué dans la Partie 4.2. L'hyperplan pour chacun d'entre eux est appris sur les données d'apprentissage du pli courant.

Les différents modèles obtenus sont comparés en Figure 4.15 dans le plan taux de détection en fonction du taux de fausse alarme. Pour le SVM, un seul point est tracé (●) et représente la moyenne des 10 plis, alors que pour le modèle SVM-1C, 50 croix sont tracées (×) en fonction des performances obtenues pour les 50 valeurs du paramètre γ , formant une courbe. Chaque croix représente les performances moyennes des 10 plis générés pour une valeur de γ donnée.



Le modèle SVM possède en moyenne sur les 10 plis un taux de fausse alarme de $7 \cdot 10^{-4}$ et un taux de détection égal à 1. La matrice de confusion associée au résultat le moins bon est donnée en Tableau 4.5, en nombre d'observations à classer. Les différents modèles SVM-1C voient leur performance varier entre un taux de détection et de fausse alarme nuls pour γ égal à 10, et un taux de fausse alarme égal à 32,3 % pour un taux de détection de 99,7 % pour γ égal à 10^{-5} . Etant donné que γ représente l'inverse de la taille du noyau gaussien, ces résultats illustrent bien le fait qu'un noyau gaussien grand va prendre une grande portion de l'espace en compte, et donc bien détecter les observations de la classe drone au prix d'une détection de celles de la classe autre bruit du mauvais côté de l'hyperplan, donc un taux de fausse alarme élevé. Pour $\gamma = 0,0021$ qui se situe dans le coude de la courbe donnée en Figure 4.15, la matrice de confusion pour le pli générant les moins bons résultats est donnée dans le Tableau 4.6. Les observations mal classées sont plus nombreuses, mais la précision reste élevée, égale à 96,8 %.

Performances de la chaîne entière : en conclusion, pour des sources uniques en environnement intérieur, la localisation est assurée en azimut avec 4° d'erreur en moyenne, et 8° en élévation. En présence d'une seule source à la fois autour de l'antenne, toutes les séquences de signal donnent lieu à une séquence focalisée qui décrit soit la classe drone lorsque c'est celui-ci qui est présent, soit la classe autre bruit lorsque c'est le haut-parleur qui joue. Les 1260 observations de la classe drone issues des séquences focalisées sont alors identifiées à 100 % avec le modèle SVM, contre 98,5 % de détection en moyenne et 1,4 % de fausse alarme pour le modèle SVM-1C situé au coude de la courbe, soit pour un γ égal à 0,0021.

Pour comprendre comment une telle précision est possible en identification, la partie suivante démontre la pertinence des caractéristiques sélectionnées pour décrire les séquences acoustiques sous forme d'observations bien séparées selon les classes en présence.

FIGURE 4.15 – Taux de détection en fonction du taux de fausse alarme pour les différents modèles testés. La performance du SVM est représentée par le point ● tandis que celles des différents modèles générés pour les coefficients γ sont données par les croix ×.

TABLEAU 4.5 – Matrice de confusion obtenue pour le pli générant les moins bons résultats durant la validation croisée pour le modèle SVM.

		Classe prédite	
		Drone	Autre bruit
Classe réelle	Drone	377	1
	Autre bruit	2	376

TABLEAU 4.6 – Matrice de confusion pour le pli générant les moins bons résultats pour le modèle SVM-1C avec $\gamma = 0,0021$, dans le coude de la courbe générée en Figure 4.15.

		Classe prédite	
		Drone	Autre bruit
Classe réelle	Drone	374	4
	Autre bruit	20	358

Identification intérieur : 1 source

- ▶ T_d du SVM : 100 %
- ▶ T_{fa} du SVM : 0 %
- ▶ T_d du SVM-1C : 98,5 %
- ▶ T_{fa} du SVM-1C : 1,4 %

4.5 Etude des caractéristiques pour l'identification

Cette partie a pour but de démontrer la pertinence des caractéristiques sélectionnées pour décrire les séquences acoustiques. Les figures qui y sont créées sont illustratives et arbitrairement choisies pour discuter sur des points spécifiques, sans être exhaustif et faire toutes les remarques possibles sur chacune des caractéristiques. Dans la première sous-partie, les liens entre les signaux sur lesquels sont calculées les caractéristiques et les valeurs contenues dans leur vecteur d'observation associé est discuté. Cette partie utilise les 29 caractéristiques dans les trois domaines, avec les mêmes réglages pour l'ensemble d'entre elles comme proposé dans le travail de [Mal18]. Suite aux remarques faites, la seconde sous-partie repère les caractéristiques produisant des valeurs aberrantes pour les modifier voire les supprimer si elles ne sont pas pertinentes.

Liens entre les signaux et les caractéristiques

Pour montrer la pertinence des caractéristiques pour décrire des signaux acoustiques, les données Int-Lin-D3/D5/D6 et Int-ND-s1k/bb/perc/hel1/hel2/mot sont utilisées. Les trois drones ne sont pas mis ensemble dans une seule classe mais composent trois classes D3, D5 et D6 afin de montrer que les caractéristiques parviennent non seulement à différencier des drones d'autres sons, mais aussi différents types de drones entre eux. Dans chacune des classes, 1800 observations sont gardées, ce qui correspond pour la classe autre bruit à 300 observations par son joué.

Seul le modèle SVM est utilisé dans cette partie car le SVM-1C ne peut prendre en compte un modèle à quatre classes. De plus, l'intérêt de cette partie est de se concentrer sur la description des séquences de signal focalisées par les caractéristiques, sans chercher le meilleur modèle. Le SVM est appris sur l'ensemble du jeu de données, selon une validation croisée composée de 10 plis, en prenant 70 % des observations pour l'apprentissage et 30 % pour le test. Cela signifie que 540 observations sont à classer en phase de test, soit 108 secondes de signal.

En première approche, la vitesse de convergence du modèle en fonction du nombre d'observations données en apprentissage est étudiée. La Figure 4.16 donne la précision du modèle sur les 10 plis de la validation croisée en fonction du nombre d'observations prises pour établir le modèle SVM à quatre classes.

La précision est croissante en fonction du nombre d'observations, avec une précision qui dépasse les 98 % à partir de la prise en compte de 930 observations en apprentissage, puis qui augmente plus lentement ensuite pour atteindre 98,7 % pour 1800 observations. L'intervalle de confiance à 99 % est représenté en transparence autour de la moyenne pour les 10 plis et il vaut 6 % pour un apprentissage avec 10 observations, pour se réduire à 0,2 % avec l'ensemble des 1800 observations disponibles au maximum pour apprendre. L'apprentissage est ainsi répétable d'un pli à l'autre sans être favorisé par un choix particulier des données d'apprentissage. Grâce à un modèle volontairement gardé linéaire pour

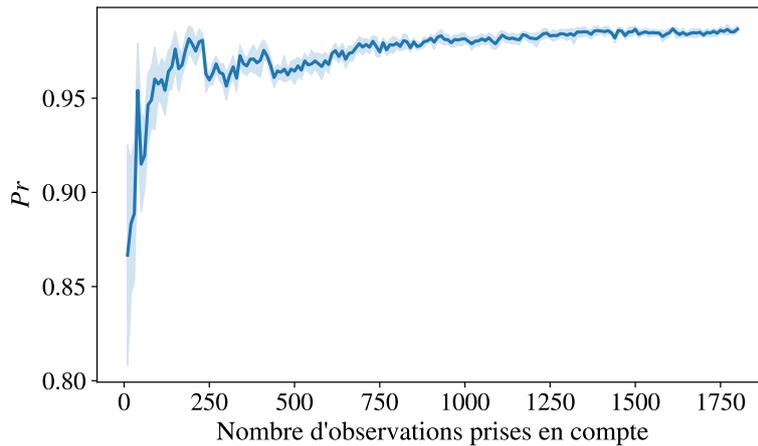


FIGURE 4.16 – Précision du modèle SVM à quatre classes en fonction du nombre d'observations prises en compte lors de la phase d'apprentissage. La courbe représente la moyenne sur les 10 plis de la validation croisée et son intervalle de confiance à 99 % est représenté en transparence (toujours selon une stratégie de bootstrapping).

ne pas le complexifier, l'apprentissage est établi pour un nombre de données d'apprentissage faible, même pour un problème multi-classes : à partir de 1000 observations pour apprendre, soit 3 min 20 s de signal pour chaque classe, la précision ne varie plus que de 0,7 %.

Une fois ce constat établi, l'ensemble des 1800 observations est utilisé pour générer un modèle dont le Tableau 4.7 donne la moyenne et l'erreur maximale des 10 matrices de confusion de la validation croisée. Les résultats dans cette matrice sont exprimés en pourcentage de données par rapport au 108 s de signal à classer.

		Classe prédite			
		D3	D5	D6	Autre bruit
Classe réelle	D3	98,3 ± 2,4	1,2 ± 1,5	0,6 ± 1,3	0 ± 0
	D5	0,8 ± 1,9	97,9 ± 3,3	1,1 ± 1,5	0,3 ± 0,4
	D6	0,3 ± 0,6	0,9 ± 1,9	98,5 ± 2,4	0,2 ± 0,6
	Autre bruit	0 ± 0	0 ± 0	0,1 ± 0,2	99,9 ± 0,2

TABLEAU 4.7 – Moyenne et valeur maximale de l'erreur des 10 matrices de confusion établies pendant la validation croisée sur le problème d'identification à quatre classes.

Le modèle SVM avec quatre classes sépare les données issues de trois types de drone différents et les autres bruits joués par le haut-parleur avec une précision de 98,6 % et se trompe au maximum de 2,1 %. Les caractéristiques se montrent ainsi pertinentes pour le problème de classification de signaux acoustiques. Diverses remarques peuvent être faites sur cette matrice de confusion :

- ▶ les données de la classe autre bruit ne sont pas confondues avec les données de drone, avec au maximum 0,1 % qui sont classées dans la classe D6 soit 0,1 s ;
- ▶ pour les drones, les erreurs sont le plus souvent faites sur le drone de taille voisine. Ainsi c'est la classe D5 qui est la moins bien classée puisqu'elle représente le drone de taille intermédiaire, mais 97,9 % de ces observations sont tout de même bien classées.

Pour illustrer ces résultats, les signaux sur lesquels sont calculés les caractéristiques sont montrés en Figure 4.17 dans les trois domaines temporel, spectral et cepstral. Pour raison de clarté une seule séquence focalisée, qui représente correctement sa classe, est sélectionnée pour tracer cette figure, pour le drone D6, le son d'hélicoptère 1, le sinus à 1 kHz et le son de perceuse¹⁰.

Résultats de classification à quatre classes

- ▶ Précision : 98,6 ± 2,1 %
- ▶ 98 % de précision pour 930 observations en apprentissage

Pour 108 secondes à classer :

- ▶ 106,1 s de D3 bien classées
- ▶ 105,7 s de D5 bien classées
- ▶ 106,4 s de D6 bien classées
- ▶ 107,9 s de sons bien classées

10: Dix séquences de tous les signaux sont disponibles dans une animation disponible sur la [page web suivante](#) pour des résultats plus exhaustifs.

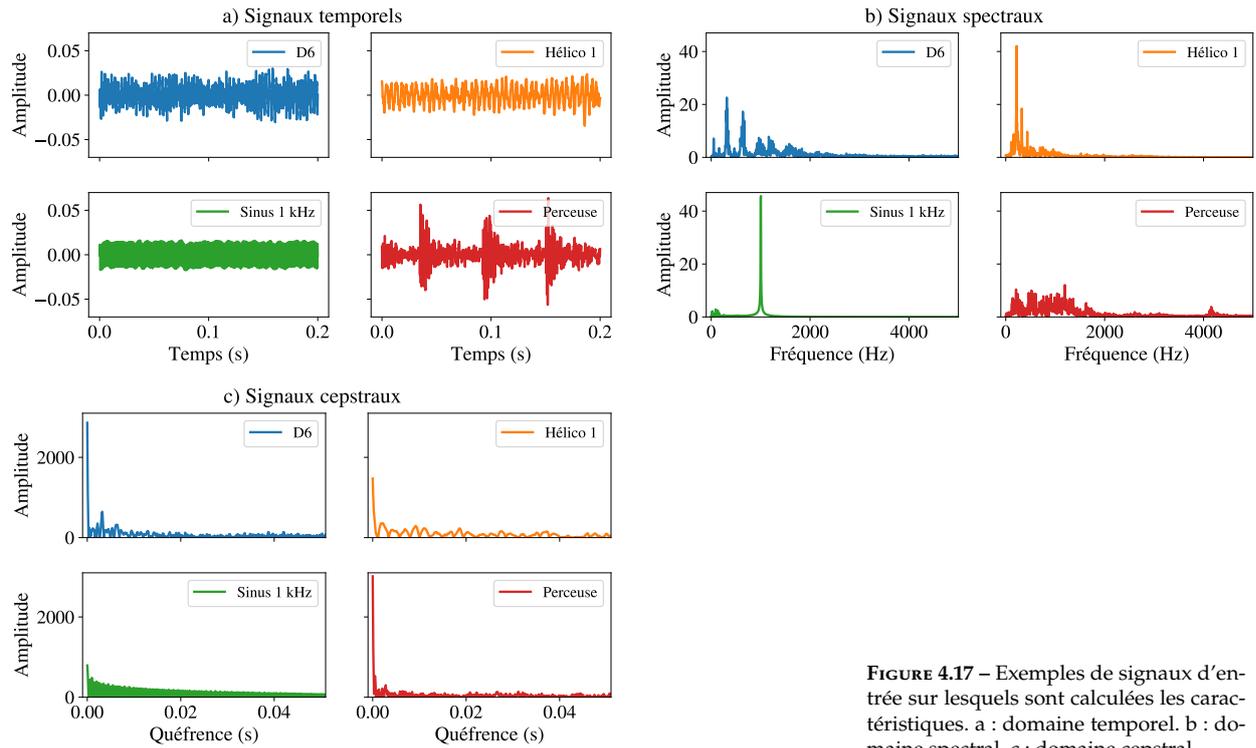


FIGURE 4.17 – Exemples de signaux d’entrée sur lesquels sont calculées les caractéristiques. a : domaine temporel. b : domaine spectral. c : domaine cepstral.

Les signaux dans le domaine temporel sont normalisés en énergie, afin de ne pas favoriser des sources proches de l’antenne au détriment de sources lointaines. Les signaux du domaine spectral sont les modules de la transformée de Fourier des signaux temporels. Ils sont alors symétriques, et ne sont représentés que jusqu’à 5 kHz car au-delà ils sont quasi nuls. De la même façon les signaux du domaine cepstral sont les modules de la transformée de Fourier des signaux du domaine spectral¹¹. Ils sont à nouveau symétriques et sont donnés en quéfrencé jusqu’à 0,05 car au-delà le cepstre reste similaire. L’amplitude est volontairement donnée sans unité dans la Figure 4.17 afin d’insister sur le fait qu’une fois en phase d’identification, les valeurs prises par les signaux ne sont pas primordiales, au contraire de leurs variations. Les taux de croisement et les taux de silence fonctionnent d’ailleurs sur les signaux normalisés, non représentés ici, plutôt que sur les signaux bruts, ce qui renforce cette idée.

Six caractéristiques sont sélectionnées pour donner une idée de la représentation des signaux qu’elles offrent. Trois séparent bien les différents sons en présence alors que trois autres ne sont pas efficaces.

La Figure 4.18 montre les histogrammes pour l’ensemble de 300 observations de chaque son sélectionné pour le taux d’attaque en temporel (Figure 4.18a), pour la largeur de la bande d’énergie en spectral (Figure 4.18b), et pour l’entropie de Rényi avec des paramètres 500 et ∞ en cepstral (Figure 4.18c)¹².

Pour le taux d’attaque, l’hélicoptère est séparé assez nettement des autres sons et le sinus à 1 kHz génère des valeurs resserrées autour de 0,1. En regardant les signaux temporels correspondant en Figure 4.17a, ces deux sons génèrent les signaux temporels les moins chahutés, possédant

11: Attention, le log est volontairement omis dans la définition du domaine cepstral dans les travaux issus de [Mal18].

12: Pour tous les sons utilisés dans la classification, l’ensemble des histogrammes est disponible sur la [page web suivante](#).

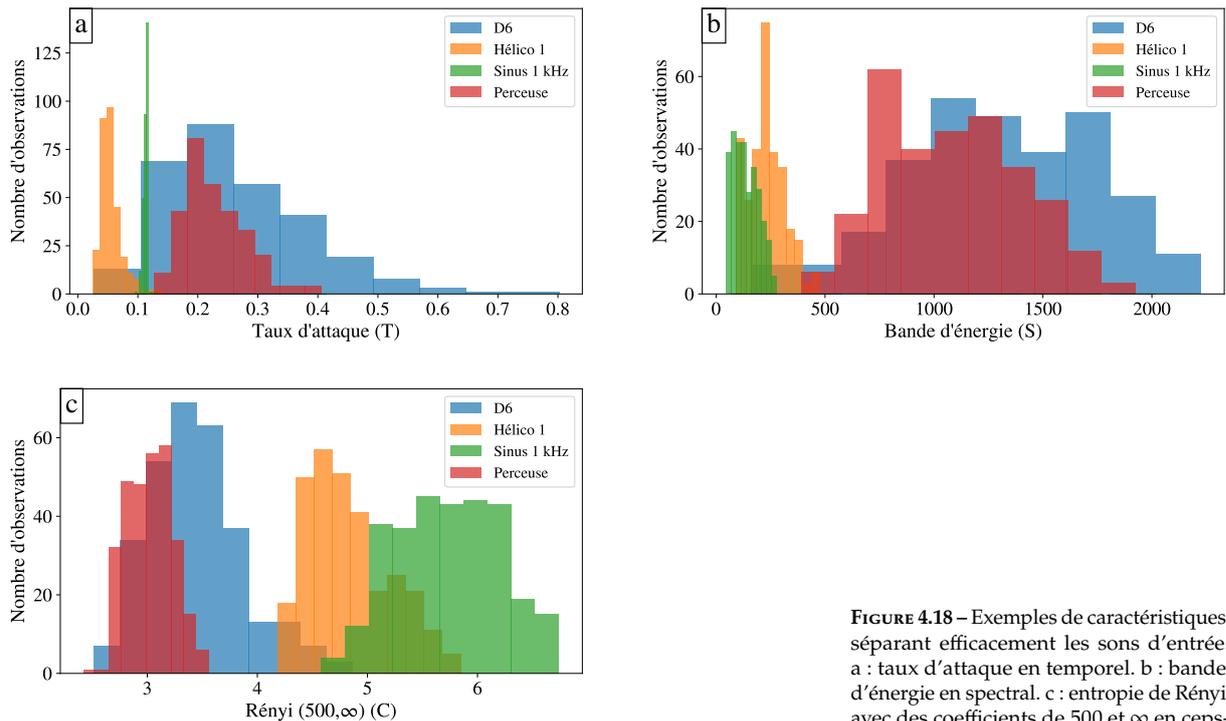


FIGURE 4.18 – Exemples de caractéristiques séparant efficacement les sons d’entrée. a : taux d’attaque en temporel. b : bande d’énergie en spectral. c : entropie de Rényi avec des coefficients de 500 et ∞ en cepstral.

le plus de basses fréquences. Pour le drone et la perceuse, il n’y a pas de séparation.

La largeur de la bande d’énergie spectrale attribuée logiquement des faibles valeurs pour le sinus à 1 kHz, puis des valeurs croissantes pour l’hélicoptère, la perceuse et dans une moindre mesure le drone. Visuellement en Figure 4.17b cette variation est retrouvée en regardant la largeur sur laquelle le spectre s’étale.

Enfin, l’entropie de Rényi, qui donne une mesure de la quantité d’information que porte un signal, sépare bien les quatre sons dans le domaine cepstral. Pour interpréter les valeurs prises, il faut comprendre que la quantité d’information portée par un signal représente son degré de prédictibilité : plus celui-ci est prédictible, plus sa quantité d’information est faible. En étudiant la Figure 4.17c, ce sont bien les signaux cepstraux générant le plus de valeurs en amplitude différentes (le sinus et l’hélicoptère), qui portent le plus d’information et donc voient leur valeur d’entropie plus grande que ceux dont les valeurs d’amplitudes sont presque toutes les mêmes (drone D6 et perceuse).

La Figure 4.19 s’intéresse ensuite à des caractéristiques moins efficaces pour séparer les différents sons avec la moyenne temporelle (Figure 4.19a), le taux de silence avec un seuil de 0,75 en spectral (Figure 4.19b), et le seuil de croisement avec un seuil de 0,5 en cepstral (Figure 4.19c).

Le point commun de ces trois caractéristiques est de produire des valeurs identiques pour tous les sons à identifier. Mais la cause n’est pas identique pour toutes. Pour la moyenne temporelle, le problème est que ce choix n’est pas pertinent, la moyenne d’un signal acoustique temporel, représentatif de la variation de pression autour de la pression statique, est effectivement toujours nulle. Pour les deux autres, le problème est

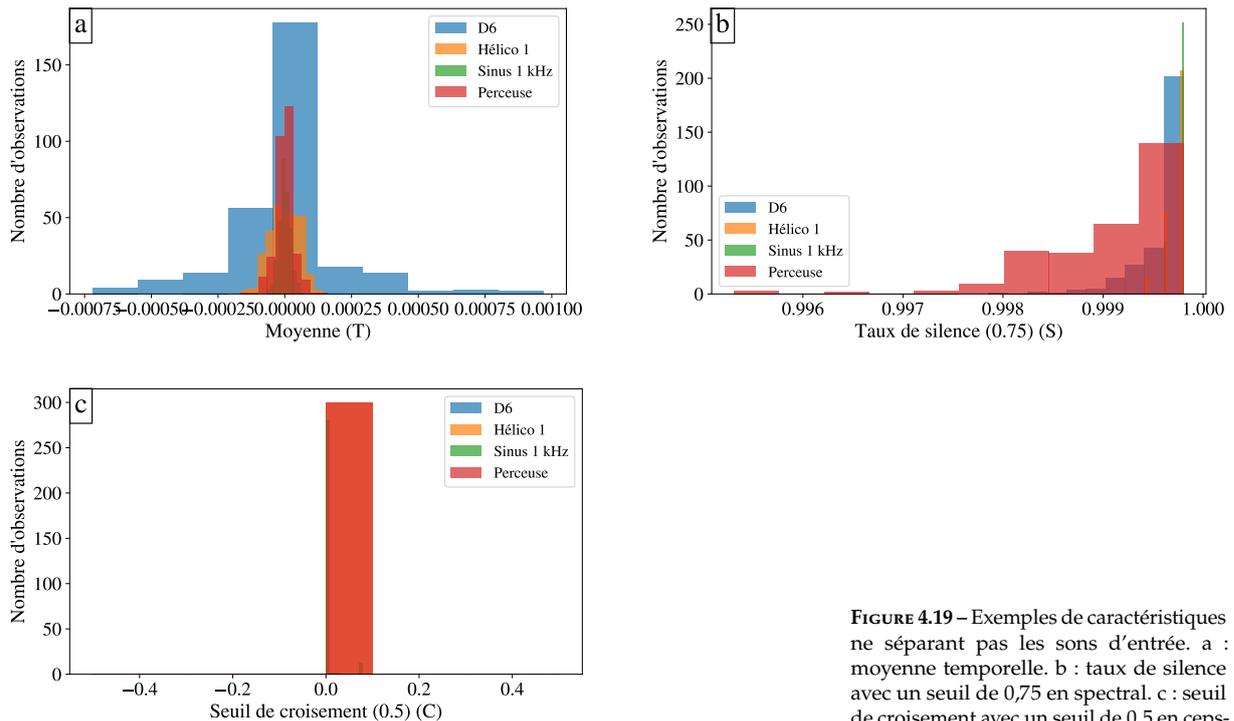


FIGURE 4.19 – Exemples de caractéristiques ne séparant pas les sons d’entrée. a : moyenne temporelle. b : taux de silence avec un seuil de 0,75 en spectral. c : seuil de croisement avec un seuil de 0,5 en cepstral.

que les seuils utilisés pour les établir, fixés de manière unique dans le code pour les trois domaines, n’est pas adapté à des signaux spectraux et cepstraux normalisés, car ils possèdent une valeur très importante à basse fréquence ou basse quérfrence, qui provoque lors de la normalisation un affaiblissement fort de l’ensemble des autres valeurs, qui passent toutes sous les seuils choisis.

A ce stade, il faut bien réaliser que c’est une combinaison de l’ensemble des 87 caractéristiques qui permet de construire l’hyperplan entre les classes en présence. Ainsi, une caractéristique qui ne sépare pas bien deux classes peut être épaulée par une autre pour laquelle la séparation est évidente. C’est bien grâce à un passage dans un espace de dimension élevée par la multiplicité et la pertinence des caractéristiques choisies que les performances d’identification affichées en début de partie peuvent être atteintes, même avec un hyperplan linéaire.

Cependant, il est peu souhaitable d’utiliser les caractéristiques non pertinentes pour la classification car l’hyperplan doit alors chercher des séparations qui n’existent pas. Ces caractéristiques peuvent être rangées dans deux familles définies comme :

- ▶ celles qui donnent des valeurs physiquement interprétables mais qui ne sont pas adaptées pour séparer les sons. C’est le cas de la moyenne temporelle illustrée en Figure 4.19a par exemple ;
- ▶ celles qui donnent des valeurs erronés, par un mauvais choix des paramètres qui les contrôlent, comme par exemple le taux de silence de la Figure 4.19b et le taux de croisement de seuil de la Figure 4.19c.

Pour éviter de se reposer sur ces caractéristiques erronées, la partie suivante détaille leur modification ou suppression pour aboutir à un jeu

de caractéristiques mieux adapté au problème de classification acoustique de drones.

Modification du jeu de caractéristiques

Pour établir les caractéristiques erronées afin de les modifier ou les supprimer du vecteur d'observation, les trois domaines sont évalués successivement.

Dans le domaine **temporel**, outre la moyenne dont il a déjà été question, le maximum divisé par la moyenne et le minimum divisé par la moyenne sont eux aussi mal adaptés à cause de la valeur de la moyenne proche de zéro. Le centroïde, la largeur de bande d'énergie et l'écart-type sont eux aussi toujours les mêmes pour toutes les observations, égaux à 0,1 soit la moitié de la séquence de signal, 0,05 soit le quart de cette séquence, et 0,00988. Enfin les taux de silence, peu importe le seuil utilisé, sont eux aussi de valeurs proches sauf pour le sinus à 1 kHz, étant donné que tous les autres sons sont à large bande et possèdent des niveaux comparables du fait de la normalisation en énergie. Par conséquent, neuf caractéristiques sont supprimées en temporel.

Dans le domaine **spectral**, comme ce sont des valeurs absolues qui sont utilisées, le minimum et le maximum sur la moyenne sont systématiquement nuls. Comme les valeurs sont issues d'une transformée de Fourier brute, elles possèdent une forte variance qui empêchent une utilisation efficace des taux d'attaque et d'affaiblissement. Enfin, les seuils des taux de croisement et de silence sont adaptés pour mieux correspondre aux valeurs obtenues après normalisation, pour donner des seuils de 0,05, 0,1 et 0,2. Ce qui implique la suppression de quatre caractéristiques et la modification de six autres.

Enfin dans le domaine **cepstral**, la même remarque est valable pour le minimum et le maximum sur la moyenne, et seul le taux d'affaiblissement est erroné : les valeurs absolues des cepstres possèdent une variance beaucoup moins forte, mais la différence systématiquement très importante entre la première valeur et la seconde implique un taux d'affaiblissement similaire pour toutes les observations. Les seuils des taux de croisement et de silence doivent à nouveau être adaptés, en étant réglés aux valeurs de 0,025, 0,05 et 0,075. Trois caractéristiques sont à nouveau supprimées et six autres modifiées.

Par conséquent, pour la suite de cette partie, les observations calculées à partir des séquences de signal de 200 ms sont constituées des 71 caractéristiques non supprimées, pour lesquelles des seuils adaptés à chaque domaine sont proposés pour les taux de silence et les taux de croisement.

Mais avant de valider ce choix sur un cas en extérieur, les parties suivantes esquissent les limitations des méthodes utilisées en localisation et en focalisation, et démontrent l'intérêt du couplage entre traitement d'antenne et apprentissage statistique, surtout en présence d'une source perturbatrice aux côtés du drone.

4.6 Limites de fonctionnement en localisation

Lors de la validation de la chaîne de traitement en intérieur, le point le plus problématique était l'erreur d'estimation de l'angle d'élévation, avec un maximum d'erreur de 40° pour des angles d'élévation faibles. Pour expliquer ces erreurs, le haut-parleur, qui est statique lors des essais, est utilisé afin de comparer les localisations en élévation qu'il induit par rapport à des cas simulés.

Afin de s'assurer une localisation possible pour un grand nombre de fréquences, ce sont les données Int-ND-bb3 qui sont utilisées pour localiser le haut-parleur avec des données réelles. Dans la configuration des vols en trois lignes droites vue de côté en Figure 4.20, il est situé dans la salle en $(-1,41 \text{ m}; 0,08 \text{ m}; 0,65 \text{ m})$. L'antenne est quant à elle posée sur

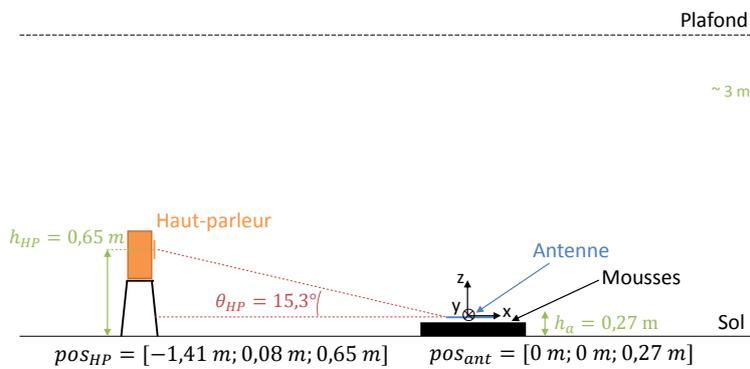


FIGURE 4.20 – Schéma donnant la position du haut-parleur dans la salle lors des essais avec les vols en trois lignes droites. L'antenne est située en $(0 \text{ m}, 0 \text{ m}, 0,27 \text{ m})$ tandis que le haut-parleur est en $(-1,41 \text{ m}, 0,08 \text{ m}, 0,65 \text{ m})$. L'angle d'élévation θ_{HP} entre le haut-parleur et l'antenne est de $15,3^\circ$.

En intérieur, les réflexions sur les différentes parois de la pièce, qui n'étaient pas du tout traitées acoustiquement, provoquent des erreurs en localisation. De plus, les distances sont limitées à 2 m maximum à cause de la taille de la zone de vol. Ainsi, le modèle de propagation en ondes planes utilisé pour la localisation n'est pas valide théoriquement à 1 kHz , fréquence à laquelle D_{lim} est égal à $2,1 \text{ m}$. Pour étudier l'impact possible des réflexions et du modèle en ondes planes, la Figure 4.21 donne l'angle d'élévation du haut-parleur θ_{HP} en fonction de la fréquence, en comparaison avec son estimation dans la salle par le maximum de la carte de MUSIC θ_{max} , avec un cas simulé $\theta_{simu-OP}$ où le modèle de propagation entre la source et l'antenne utilisé lors du traitement pas MUSIC est en ondes planes, et avec le même cas simulé mais une propagation en ondes sphériques $\theta_{simu-OS}$.

Pour la simulation, les réflexions sur le sol et le plafond sont prises en compte, et la propagation directe entre le haut-parleur et l'antenne est effectuée avec un modèle sphérique. Les coefficients de réflexion ont été réglés manuellement pour faire correspondre la courbe obtenue en simulation à la courbe réelle, pour aboutir à un coefficient de $0,9$ pour le plafond et $0,2$ pour le sol. Étant donné que des tapis de gymnastique étaient posés sur le sol de la salle pour éviter les chutes lors des chutes de drone, ces coefficients de réflexions sont cohérents.

A cause des réflexions, une oscillation des estimations de l'angle d'élévation est visible, provoquant une erreur maximale de $44,4^\circ$ à 550

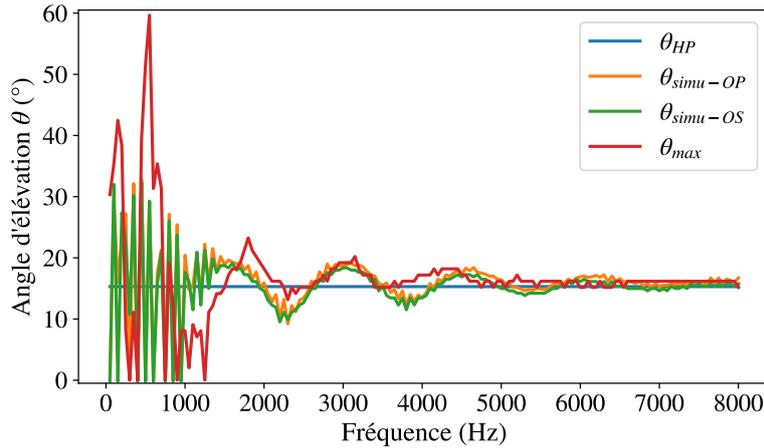


FIGURE 4.21 – Comparaison des angles d'élevation réel (θ_{HP}), estimé par le maximum de la carte de MUSIC avec les données réelles (θ_{max}) et estimé lors de la simulation avec des ondes planes ($\theta_{simu-OP}$) et des ondes sphériques ($\theta_{simu-OS}$) en fonction de la fréquence.

Hz pour les données réelles (en rouge -) et de $17,5^\circ$ à 450 Hz en simulation. Les oscillations basses fréquences, dues aux réflexions sur le sol, ont lieu aux mêmes fréquences pour les cas réel et de simulation jusqu'à 4 kHz, avec un léger décalage ensuite. Les oscillations hautes fréquences, causées par les réflexions sur le plafond, provoquent des erreurs qui diminuent avec la fréquence, passant de près de 45° à moins de 1° .

L'impact du modèle de propagation est quant à lui limité. Les deux courbes de simulation obtenues ont les mêmes variations, avec une surestimation légèrement plus importante pour le modèle en ondes planes. Malgré le fait que la distance entre le haut-parleur et le drone soit de l'ordre de grandeur de D_{lim} , le modèle en ondes planes peut être utilisé sans causer d'erreur importante sur l'angle d'élevation estimé par MUSIC.

Suite à ces résultats, il est recommandé pour la suite de travailler à des fréquences plus élevées, pour éviter la bande de fréquence [0 Hz ; 1 kHz]. À partir de 5 kHz, les oscillations de l'angle d'élevation estimé sont bornées sous les 2° , ce qui permet d'envisager des résultats de localisation précis au-delà de cette fréquence.

Pour valider ensuite l'observation faite en Partie 4.4 selon laquelle l'erreur sur l'angle d'élevation se réduit lorsque ce dernier augmente, la Figure 4.22 compare les résultats en simulation avec la même configuration que celle établie ci-dessus, mais pour trois angles d'élevation croissants de $15,3^\circ$, 45° et 80° pour la source haut-parleur. Pour assurer des performances comparables d'un point de vue du modèle, la distance entre la source simulée et le centre de l'antenne est gardée fixe pour les différentes configurations, ce qui implique que la source s'élève selon une trajectoire circulaire à rayon constant autour de l'antenne.

Plus l'angle d'élevation à estimer est élevé, moins les oscillations basses fréquences sont visibles. De plus, les oscillations marquées sur la bande de fréquence [0 Hz ; 1 kHz] s'atténuent pour passer de 33° d'amplitude à $15,3^\circ$ à 5° seulement à 80° . Inversement, à haute fréquence l'erreur augmente, passant de 2° pour l'angle de $15,3^\circ$ à 10° pour l'angle de 80° . Ce phénomène peut s'expliquer par les réflexions sur le plafond dont l'amplitude augmente avec l'angle d'élevation considéré étant donné que la source se rapproche alors du plafond (1,69 m de hauteur pour $\theta = 80^\circ$). Ces résultats confirment bien ce qui a été vu lors de la localisation à 1 kHz pour différents angles d'élevation en Partie 4.4.

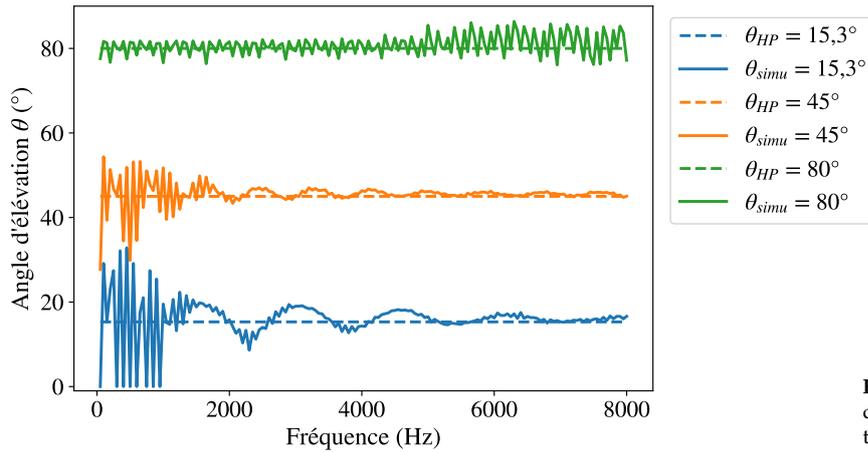


FIGURE 4.22 – Comparaison de l'erreur d'estimation de l'angle d'élevation pour trois simulation à $\theta_{HP} = 15,3^\circ$ (-), $\theta_{HP} = 45^\circ$ (-), et $\theta_{HP} = 80^\circ$ (-).

Pour conclure, la **principale limitation** en localisation est due aux **réflexions**, fortement présentes dans une salle intérieure. La **distance entre la source et l'antenne**, de l'ordre de grandeur de D_{lim} **ne provoque pas d'erreur** de localisation importante même avec l'utilisation d'un modèle de propagation en ondes planes.

En fonction du cas extérieur visé, les réflexions peuvent être moins présentes et atténuées. Dans la configuration de cette thèse en prairie, seule la réflexion du sol est présente, et fortement atténuée grâce au terrain herbeux. Un modèle en champ libre est alors sélectionné, et pour encore diminuer l'impact des réflexions résiduelles, la fréquence de traitement est choisie à 5 kHz. Pour des cas urbains plus réverbérants, la question de la prise en compte de ces réflexions reste ouverte et à étudier.

Avec ces limitations en localisation maîtrisées, la partie suivante démontre l'intérêt qu'apporte le couplage des deux outils méthodologiques que sont le traitement d'antenne et l'apprentissage statistique pour améliorer les performances en identification par rapport à un microphone unique, ce qui reste actuellement largement privilégié dans la littérature.

4.7 Intérêt du couplage traitement d'antenne / apprentissage statistique pour l'identification

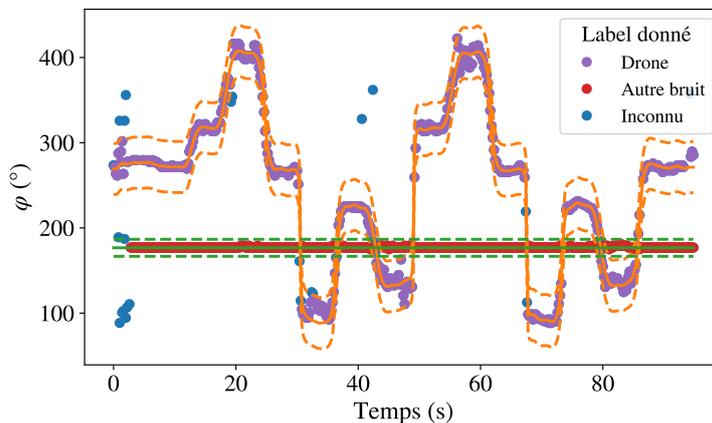
Au-delà de la localisation qu'il produit au contraire d'un microphone unique, le couplage entre le traitement d'antenne à l'apprentissage statistique permet de filtrer spatialement les signaux grâce à l'étape dite de focalisation. En présence d'une source unique, le gain apporté par la focalisation permet d'identifier de manière plus efficace lorsque le bruit ambiant augmente. Mais son intérêt principal réside dans le fait de pouvoir identifier plusieurs sources présentes simultanément autour de l'antenne.

Pour démontrer cet aspect, les données Int-3ld-D6 avec le drone seul, Int-ND-s1k1/2/3 avec le signal de sinus à 1 kHz seul et Int-3ld-D6-s1k1/2/3 avec les deux en même temps sont utilisées. Seul le son de

sinus à 1 kHz, facilement séparable du son émis par le drone comme vu en Partie 4.5, est pris ici pour s'assurer un apprentissage facile, et se concentrer sur les aspects liés à la différence de niveau sonore entre les deux sources. Les données comprenant chaque source séparément servent alors à établir l'hyperplan entre les classes drone et autre bruit lors de l'étape d'apprentissage, tandis que les données où les deux sources sont présentes simultanément sont utilisées indépendamment en phase de test. La validation de l'apprentissage étant étudiée dans les parties précédentes, les résultats qui suivent ne traitent que de la précision en localisation et en identification des enregistrements pour lesquels les deux sources sont présentes en même temps, ce qui nécessite une procédure spécifique pour labéliser les signaux provenant de chaque source.

Procédure de labélisation

En passant à une configuration mettant en jeu deux sources, la labélisation des données localisées, nécessaire pour évaluer les résultats d'identification, n'est plus possible automatiquement comme c'était le cas jusqu'alors¹³. Dans ce cas à deux sources, pour chaque séquence de 200 ms de signal en entrée, deux séquences focalisées sont obtenues après estimation des deux premiers maxima de la carte de MUSIC. Ainsi toutes les 200 ms, deux observations associées à chacune des deux séquences focalisées sont à labéliser comme drone ou autre bruit afin d'obtenir leur vérité terrain. Grâce à la précision de localisation en azimuth, cette labélisation est entreprise de manière semi-automatique comme illustré en Figure 4.23. Chaque point y représente une observation, dont l'angle d'azimut est donné au cours du temps.



13: Le même label était attribué à toutes les observations issues des séquences focalisées, en fonction de la source qui était en présence.

FIGURE 4.23 – Illustration de la procédure de labélisation suivie pour déterminer la classe de chaque observation à classer. Des seuils sur les angles en azimuth donnés par la vérité terrain sont établis manuellement. Pour chaque observation issue des deux maxima estimés sur la carte de MUSIC, si elle se situe entre les seuils entourant la vérité terrain du drone elle est labélisée comme drone (●), entre les seuils de la vérité terrain du haut-parleur elle est labélisée comme autre bruit (●), sinon elle est labélisée comme inconnu (●).

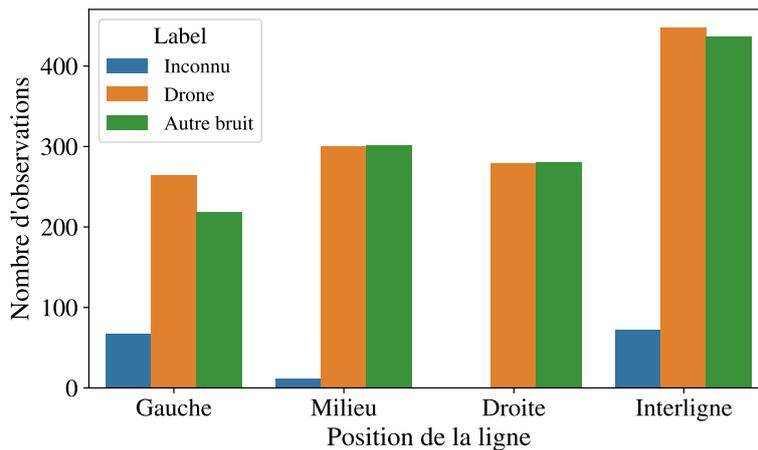
Pour labéliser une observation, le principe suivi est de comparer son angle d'azimut à celui donné par les deux vérités terrains pour le drone et pour le haut-parleur. Si cet angle se situe entre les seuils supérieurs et inférieurs autour de la vérité terrain du drone (en orange - dans la Figure 4.23), l'observation est labélisée comme drone (●). Pour une estimation située entre les seuils liés à la position du haut-parleur (en vert -) elle est labélisée comme autre bruit (●). Dans les autres cas, un label inconnu (●) lui est attribué, qui traduit le fait que la classe de cette observation ne peut être déterminée à partir de son angle d'azimut. Les observations

fonctionnant toujours par paire toutes les 200 ms, la génération d'un label inconnu provoque un label manquant au cours du temps pour l'une ou l'autre des classes. Pour la classe *autre bruit* la conséquence n'est pas importante car ses observations ne sont pas recherchées, mais pour la classe *drone* ce manque entraîne directement une baisse de son taux de détection.

Identification de deux sources

Analyse de la localisation : sur l'ensemble des données issues des trois vols pour les trois volumes de son du haut-parleur différents, le drone est localisé 92,9 % du temps, et le haut-parleur 88,9 % du temps. Sur le nombre total d'observations à classer, 5,4 % ont donné lieu à un label inconnu.

L'histogramme de toutes les séquences focalisées est présenté en Figure 4.24, en fonction de la ligne sur laquelle le drone était situé pour chaque séquence focalisée entre la ligne de gauche, du milieu ou de droite. Le groupe interligne correspond aux moments où le drone vole entre deux lignes successives pour entamer une nouvelle ligne droite.



Localisation intérieur : 2 sources

Proportions de séquences localisées :

- ▶ Drone : 92,9 %
- ▶ Haut-parleur : 88,9 %

Sur les 2×1391 séquences à classer :

- ▶ 5,4 % d'inconnu générés

FIGURE 4.24 – Histogramme des séquences focalisées par label en fonction de la position durant la trajectoire du vol en trois lignes droites du drone. La ligne de gauche correspond à celle où le drone passe au-dessus du haut-parleur et en est donc le plus proche.

Lorsque le drone vole sur la ligne de droite, soit au plus loin du haut-parleur, toutes les séquences focalisées donnent lieu à une labélisation selon les classes *drone* ou *autre bruit*. Pour la ligne du milieu, seules 1,8 % des séquences donnent lieu à des labélisations dans la classe *inconnu*, contre 12,2 % et 7,5 % pour la ligne de gauche et les interlignes, respectivement. Plus les deux sources sont proches, plus les résultats en localisation sont perturbés.

Définition de la différence de niveau ΔL_{moy} : pour la suite des résultats en identification, la différence de niveau entre les deux sources au cours de la trajectoire a été déterminée. Pour l'obtenir, le haut-parleur émet d'abord seul, puis grâce au système de contrôle de trajectoire, le drone parcourt exactement la même trajectoire avec et sans haut-parleur qui émet. La différence de niveau à un instant donné de la trajectoire correspond alors à la différence entre les densités spectrales de puissance (DSP) lorsque les deux sources sont seules. Par exemple lorsque le drone se situe au milieu de la ligne de droite, la DSP du signal focalisé obtenu

lorsque le drone est seul est comparée à celle du son joué par le haut-parleur réglé au volume 1 en Figure 4.25. Elles sont toutes deux calculées grâce à un périodogramme de Welch.

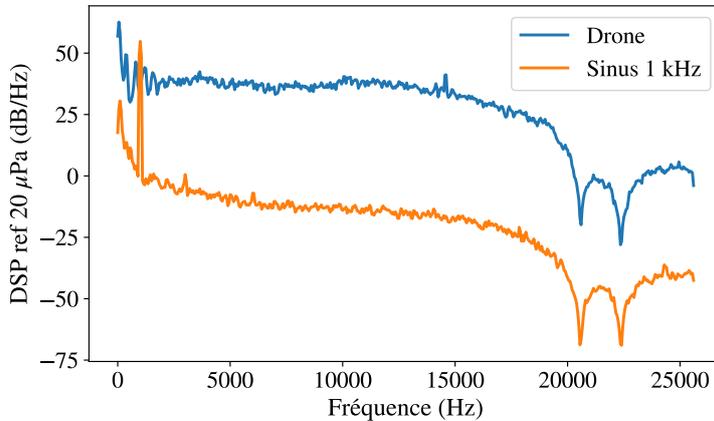


FIGURE 4.25 – Comparaison entre le signal de drone et le signal de sinus à 1 kHz au volume 1 joué par le haut-parleur lors des essais intérieurs avec le vol de drone selon trois lignes droites. Les deux densités spectrales ont été calculées pour le drone et le son joués seuls, et pour un drone situé au milieu de la première ligne droite.

La différence de niveau ΔL_{moy} est alors définie par le rapport entre le niveau moyenné sur toute la bande de fréquence de ces deux densités spectrales de puissance :

$$\Delta L_{moy} = 10 \log_{10} \left(\frac{\overline{DSP}_{drone}}{\overline{DSP}_{son}} \right), \quad (4.6)$$

avec \overline{DSP}_{drone} et \overline{DSP}_{son} les moyennes des valeurs linéaires prises par les densités spectrales de puissance calculées sur le signal de drone seul, et sur le signal du haut-parleur seul. Ce calcul est répété tout au long de la trajectoire pour obtenir une différence de niveau qui dépend de la position du drone dans la salle.

Analyse de l'identification : pour analyser les résultats d'identification par rapport à la différence de niveau ΔL_{moy} , les distributions des prédictions faites sur les observations relatives à chaque label sont tracées en Figure 4.26. Pour chaque classe labélisée, les prédictions correctes sont tracées en vert (●), et les prédictions incorrectes en rouge (●). Pour les observations labélisées comme inconnu, leur prédiction est considérée comme correcte lorsqu'elles sont classées dans la classe autre bruit, mais elle est donnée à titre indicatif car il est évident que des erreurs seront commises : ces dernières ne décrivent correctement aucune des deux classes. Les boîtes à moustaches indiquent les statistiques des distributions obtenues, avec le trait central pour sa moyenne, les limites de la boîte pour ses quartiles et les tirets aux extrémités les valeurs minimale et maximale. Deux cas sont comparés pour démontrer l'apport du traitement d'antenne :

- ▶ L'apprentissage statistique est effectué sur le signal focalisé grâce aux 81 microphones de l'antenne ;
- ▶ l'apprentissage statistique est effectué sur un microphone unique, choisi ici comme le microphone 1 de l'antenne.

Pour le signal focalisé en Figure 4.26a, 86,6 % des observations labélisées comme drone sont correctement prédites et pour la classe

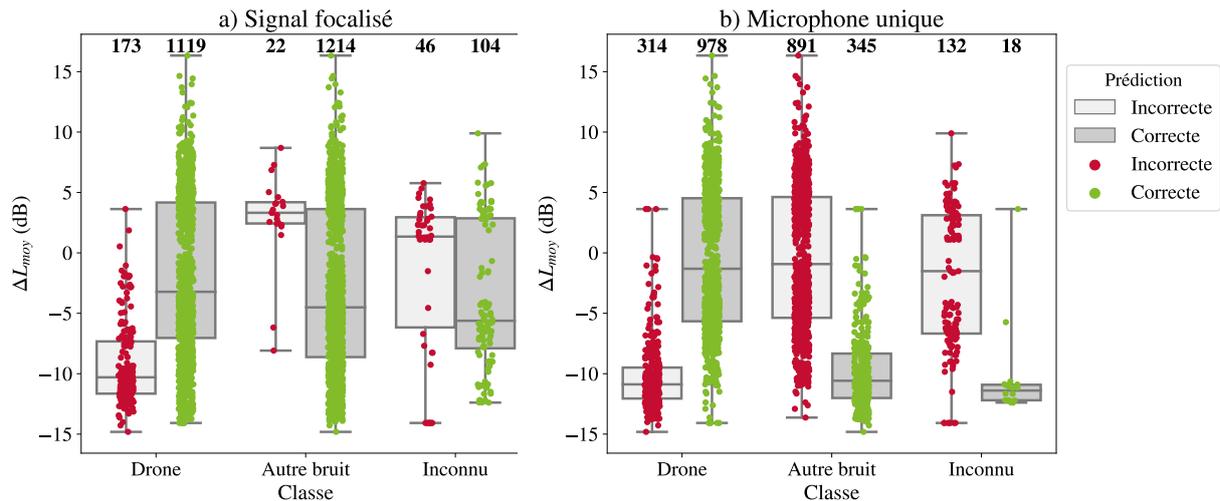


FIGURE 4.26 – Résultats d'identification à deux sources en intérieur. Les distributions correspondent, pour chaque classe en abscisse, aux observations correctement prédites en vert (●), et incorrectement prédite en rouge (●), en fonction de la différence de niveau entre le drone et le son joué par le haut-parleur ΔL_{moy} . Le nombre d'observations contenues dans les distributions est indiqué au-dessus de chacune. a : résultats pour les signaux focalisés. b : résultats pour les signaux issus d'un microphone unique.

autre bruit, le taux de prédiction correcte est de 98,2 %. Pour les observations correctement localisées, l'identification est précise, avec un avantage pour les observations de la classe autre bruit. Pour des différences de niveau supérieures à 4 dB, plus aucune erreur n'est commise pour l'identification des drones. Dans la classe autre bruit, très peu d'erreurs sont commises et leur distribution ne permet pas de conclure sur un comportement particulier des prédictions incorrectes. Enfin, les observations labélisées inconnu voient les distributions de leurs prédictions correctes et incorrectes similaires, traduisant l'aspect imprédictible de ces observations.

Lorsque la classification est entreprise avec les signaux issus d'un microphone unique, l'analyse est totalement différente. Quelle que soit la classe de départ, plus aucune prédiction n'est faite dans la classe autre bruit au-dessus de 4 dB de différence de niveau entre le signal émis par le drone et celui du haut-parleur. Pour la classe drone, cela n'est pas trop problématique puisque 75,7 % des observations restent tout de même bien classées. Par contre pour la classe autre bruit, seules 27,9 % des observations sont correctement classées, ce qui est nettement moins important qu'avec le signal focalisé.

Ces résultats montrent que pour ces essais intérieurs, le signal du drone prend le pas sur les signaux émis par le haut-parleur lors de la classification par le modèle SVM. Lorsqu'il est le plus puissant, toutes les observations issues des séquences focalisées du microphone unique lui correspondent, mais la réciproque n'est pas vraie : pour des différences de niveaux ΔL_{moy} négatives, certaines observations de drone sont toujours correctement identifiées. Il faut aller au-delà de -10 dB pour commencer à voir disparaître presque toutes les identifications correctes de drone.

Pour conclure, du point de vue de la chaîne de traitement dans son ensemble, sur les 1391 séquences de drone qui étaient à identifier au total, avant l'étape de localisation, l'utilisation d'un signal focalisé permet un taux de détection de 80,4 % avec 4,9 % de taux de fausse alarme, en regroupant les observations des classes autre bruit et inconnu pour

Identification intérieur : 2 sources

Différence de niveau entre les sources :

- ▶ $-15 \text{ dB} < \Delta L_{moy} < 15 \text{ dB}$

T_d avec le signal focalisé :

- ▶ Drone : 86,6 %
- ▶ Autre bruit : 98,2 %

T_d avec le microphone unique :

- ▶ Drone : 75,7 %
- ▶ Autre bruit : 27,9 %

Chaîne entière intérieur : 2 sources

Avec le signal focalisé :

- ▶ $T_d = 80,4 \%$
- ▶ $T_{fa} = 4,9 \%$

Avec un microphone unique :

- ▶ $T_d = 70,3 \%$
- ▶ $T_{fa} = 73,5 \%$

calculer ce taux. En comparaison, le microphone unique détecte 70,3 % des drones, avec taux de fausse alarme important de 73,5 %.

Cependant, même en ayant amélioré fortement le taux de fausse alarme grâce au filtrage spatial amené par le traitement d'antenne, ces résultats montrent que la focalisation ne retire pas complètement la contribution d'une source perturbatrice. La partie suivante illustre cette limitation.

4.8 Limites de fonctionnement en focalisation en présence de plusieurs sources

Idéalement, la focalisation permet d'éliminer entièrement le signal qui ne provient pas de la direction spatiale visée. Cependant, avec une focalisation basée sur une formation de voies, des résidus provenant des signaux présents autour de l'antenne restent présents dans le signal focalisé final.

Pour illustrer cet effet, les données utilisées dans la partie précédente pour lesquelles le drone et le haut-parleur sont simultanément en train d'émettre, avec le haut-parleur jouant un sinus à 1 kHz, sont à nouveau exploitées (les données Int-3ld-D6-s1k1/2/3).

L'impact de l'imprécision de la focalisation sur les caractéristiques est alors montrée en Figure 4.27, en représentant les distributions de trois d'entre elles pour des observations calculées sur différents groupes de séquences focalisées vers la source acoustique qui correspond au drone :

- ▶ des séquences où le drone est seul (Drone seul);
- ▶ des séquences où le drone vole et le haut-parleur joue le sinus au volume 1 (Son (vol 1));
- ▶ des séquences où le drone vole et le haut-parleur joue le sinus au volume 2 (Son (vol 2));
- ▶ et des séquences où le drone vole et le haut-parleur joue le sinus au volume 3 (Son (vol 3)).

Alors que toutes les séquences doivent être prédites comme drone, les groupes sélectionnés ne produisent pas des distributions qui se superposent pour ces trois caractéristiques. Pour la première en temporel en Figure 4.27a, seul le volume 3 perturbe la distribution des séquences sélectionnées, ce qui montre que cette caractéristique est relativement robuste au son perturbateur. Par contre pour les caractéristiques spectrale et cepstrale en Figure 4.27b et Figure 4.27c, pour chaque niveau sonore une distribution différente est obtenue, ce qui montre qu'à partir de ces caractéristiques, il va être difficile de classer correctement une séquence de drone avec précision dès lors qu'un sinus perturbateur sera présent en même temps que le drone aux alentours de l'antenne.

Pour illustrer ces différences observées dans les distributions, les signaux d'entrée de l'étape d'identification pour chaque domaine sont donnés en Figure 4.28, pour un exemple de séquence pour chacun des quatre volumes sonores étudiés pour le haut-parleur.

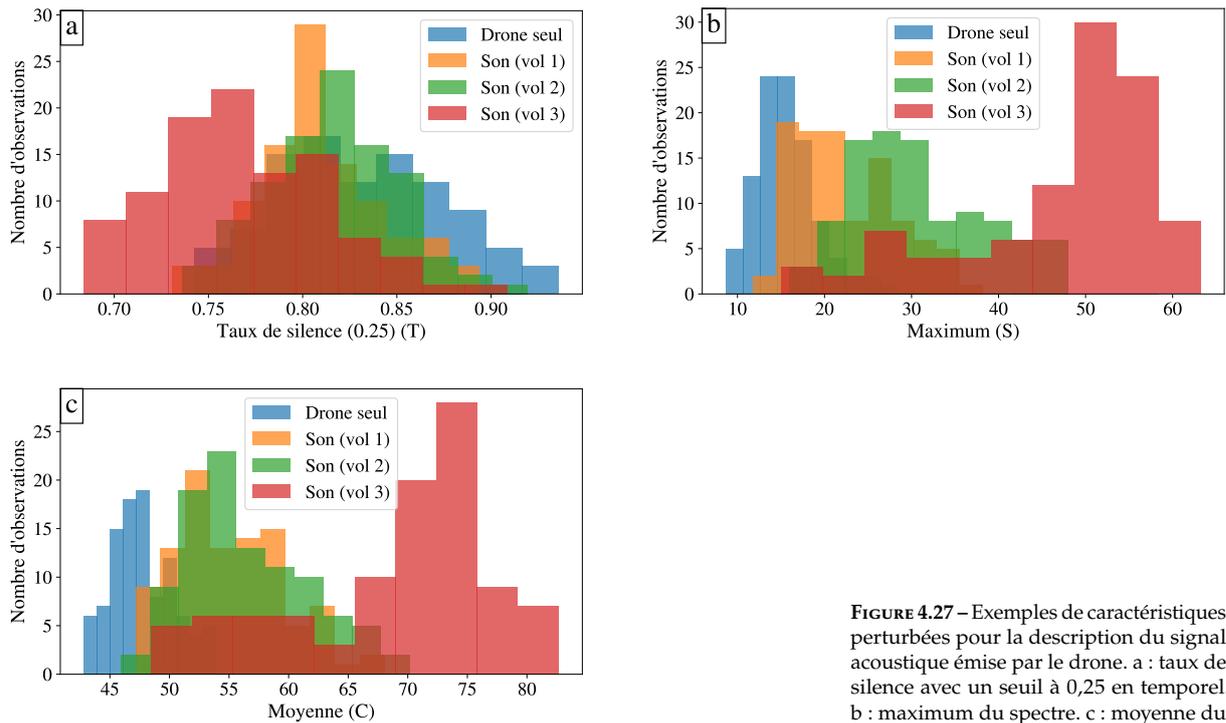


FIGURE 4.27 – Exemples de caractéristiques perturbées pour la description du signal acoustique émise par le drone. a : taux de silence avec un seuil à 0,25 en temporel. b : maximum du spectre. c : moyenne du cepstre.

Bien que pour les signaux temporels le lien avec les valeurs obtenues dans l’histogramme de la Figure 4.27a soit difficile à établir, pour le domaine spectral le maximum de chaque séquence focalisée, qui correspond à la fréquence de 1 kHz jouée par le haut-parleur, augmente bien avec le volume du son, pour dépasser 45 dans le cas du son au volume 3. Pour le domaine cepstral la tendance observée dans l’histogramme de la Figure 4.27c est aussi retrouvée avec une moyenne des valeurs cepstrales de plus en plus forte avec l’augmentation du volume du son joué, qui s’aperçoit à travers la disparition progressive des creux dans les valeurs prises par les signaux cepstraux aux alentours de 0,025 et 0,075 s.

Cette étude est effectuée pour le sinus ici pour conserver une cohérence avec la partie précédente. En étudiant la perturbation des caractéristiques pour d’autres sons joués par le haut-parleur, les sons à larges bandes posent plus de problème encore, avec un plus grand nombre de distributions de caractéristiques modifiées, et de manière plus importante. Tous ces cas ne peuvent être présentés ici par souci de concision¹⁴.

Pour terminer la validation de la chaîne de traitement, ses performances en extérieur sont étudiées dans la partie suivante, qui met en œuvre toutes les remarques faites au cours du chapitre dans ce cas plus réaliste. Elle mettra en jeu des sources à larges bandes, pour faire écho à la remarque précédente.

14: Les résultats pour l’ensemble des caractéristiques avant modification et pour les différents sons sont néanmoins disponibles sur une [page web dédiée](#).

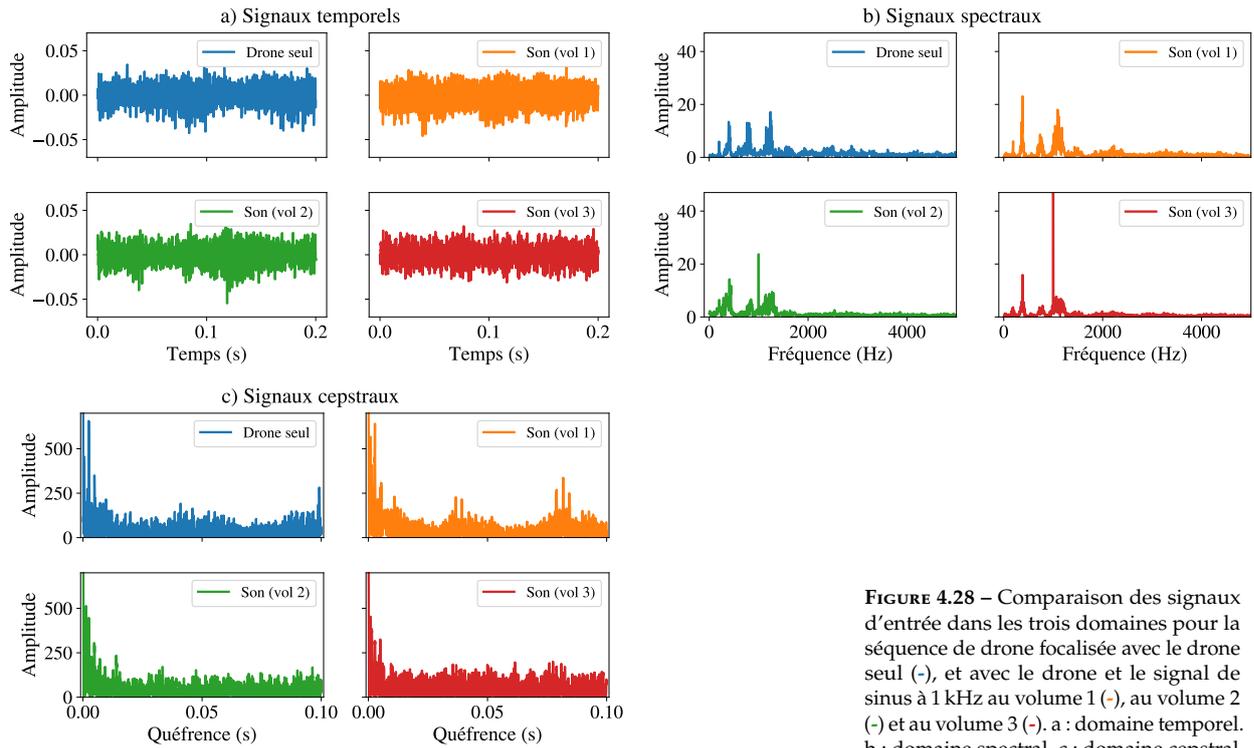


FIGURE 4.28 – Comparaison des signaux d’entrée dans les trois domaines pour la séquence de drone focalisée avec le drone seul (-), et avec le drone et le signal de sinus à 1 kHz au volume 1 (-), au volume 2 (-) et au volume 3 (-). a : domaine temporel. b : domaine spectral. c : domaine cepstral.

4.9 Performances en extérieur

Dans cette dernière partie, l’objectif est de montrer les performances atteintes par la chaîne de traitement en localisation et en identification en extérieur. Etant donné les cas d’application visés, l’extérieur est en toute logique le cadre dans lequel le système final sera le plus amené à évoluer.

Pour étudier les performances avec et sans présence d’une source acoustique perturbatrice, les données utilisées dans cette partie sont les données Ext-D6, Ext-ND-bb1/bb2/voit1/voit2 et Ext-D6-bb1/bb2/voit1/voit2. Afin d’être plus équitable entre les sons joués par le haut-parleur, qui atténue les hautes fréquences, et le signal acoustique émis par le drone, toutes les séquences focalisées dans cette partie sont filtrées par un passe-bas à 8 kHz avant que leurs caractéristiques ne soient calculées. Cela permet de s’assurer que les modèles appris se concentrent sur des caractéristiques qui n’apprennent pas le fait qu’un signal soit juste joué par un haut-parleur ou non.

Sources présentes une à une

Localisation du drone : en premier lieu, la localisation du drone lorsqu’il vole seul est évaluée. Suite à l’étude des performances en localisation faite sur les données intérieures, la fréquence de traitement pour établir les cartes de MUSIC à chaque séquence est choisie à 5 kHz dans cette partie. Le drone et les sons joués par le haut-parleur, qui sont à large bande ici, émettent tous jusqu’à cette fréquence. Les angles d’azimut et d’élévation φ_{max} et θ_{max} estimés par le maximum des cartes de MUSIC sont comparés à ceux obtenus grâce au système de localisation

RTK en Figure 4.29, pour les deux vols disponibles de drone où il est seul et durant lesquels il a été piloté manuellement par un opérateur.

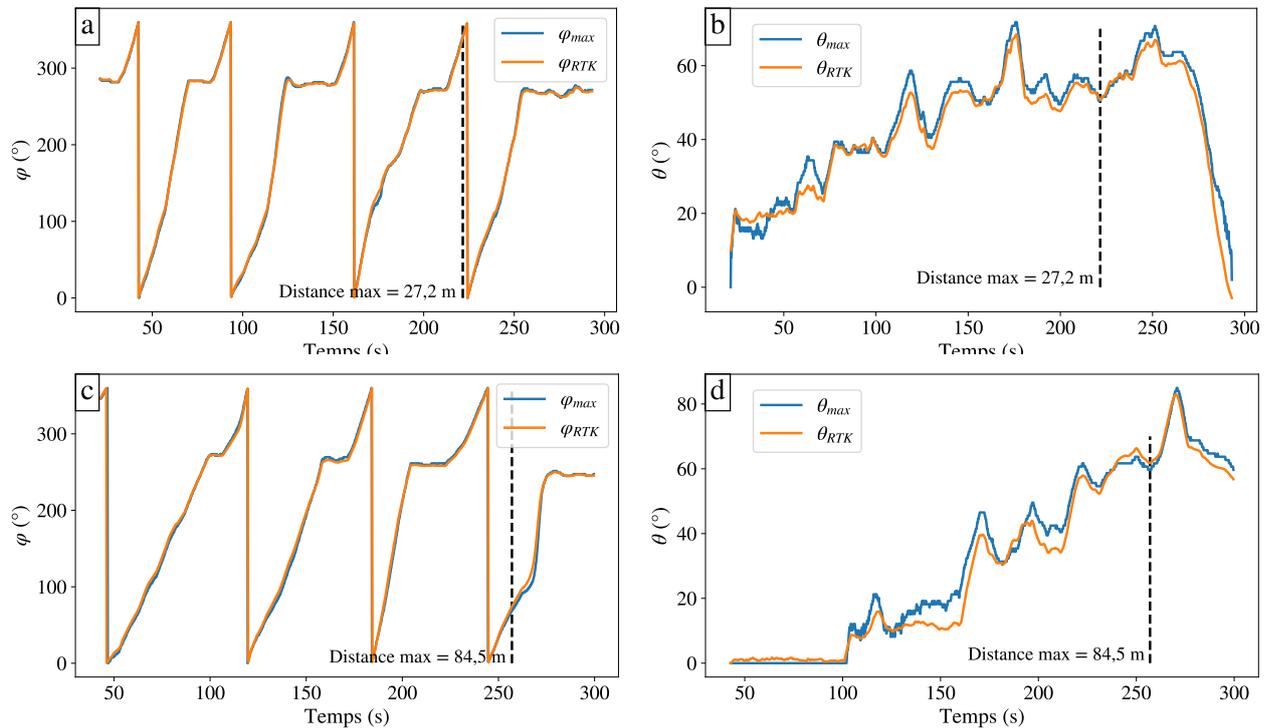


FIGURE 4.29 – Comparaison entre les angles estimés par le maximum des cartes de MUSIC (-) pour chaque séquence et les angles donnés par le système de localisation RTK (-). a : angle d'azimut du premier vol. b : angle d'élévation du premier vol. c : angle d'azimut du second vol. d : angle d'élévation du second vol.

L'erreur moyenne entre la localisation estimée et la vérité terrain donnée par le système RTK est de $2,9^\circ$ en azimut et de $3,0^\circ$ en élévation, pour l'ensemble des deux trajectoires suivies. L'erreur maximale angulaire est de $26,4^\circ$ pour l'azimut en un point particulier du vol visible à 265 s du second vol. En élévation cette erreur maximale est de $15,3^\circ$. Pour les angles d'élévation négatifs, qui correspondent aux moments où le drone était en train de décoller, ce dernier n'est pas correctement localisé car il faut qu'il passe au-dessus de l'antenne. C'est pourquoi la localisation du second vol ne démarre qu'à partir de 42 s.

Durant le premier essai le drone vole jusqu'à une distance maximale à l'antenne de 27,2 m, atteint à 220 s et représenté par les lignes verticales dans la Figure 4.29a et la Figure 4.29b. Durant le second vol, le drone s'écarte jusqu'à 84,5 m, sans impact sur la précision en localisation.

Performances en identification : l'analyse des résultats d'identification sur ce cas où les sources sont présentes individuellement autour de l'antenne permettent de situer les performances par rapport à ceux de la littérature. Grâce à la précision atteinte en localisation, 2998 séquences focalisées de drone sont disponibles pour l'apprentissage. Cependant, seules 348 séquences de sons joués par le haut-parleur seul le sont. Le modèle SVM est donc appris avec 348 observations pour chaque classe drone et autre bruit pour conserver un jeu de données équilibré, alors que le SVM-1C exploite l'ensemble des 2998 observations de drone à disposition.

Une validation croisée composée de 10 plis est menée pour valider

Localisation extérieur : 1 source

- ▶ Erreur moyenne φ : $2,9^\circ$
- ▶ Erreur moyenne θ : $3,0^\circ$
- ▶ Erreur maximale φ : $26,4^\circ$
- ▶ Erreur maximale θ : $15,3^\circ$

les performances du SVM avec ses paramètres fixes, et celles du SVM-1C avec ses 50 valeurs pour le paramètre γ . Dans le cas du SVM, les deux classes sont équilibrées avec 348 observations dans chacune, et 70 % sont prises à chaque pli pour apprendre, les 30 % restantes étant utilisées en test. Pour le SVM-1C, ce sont 70 % des 2298 observations de drone uniquement qui sont prises pour l'apprentissage à chaque pli, et le test est effectué sur 30 % de ces données auxquelles sont ajoutées 30 % des 348 observations de autre bruit. Les moyennes sur les 10 plis des taux de fausse alarme et de détection sont tracés pour chacun des modèles en Figure 4.30.

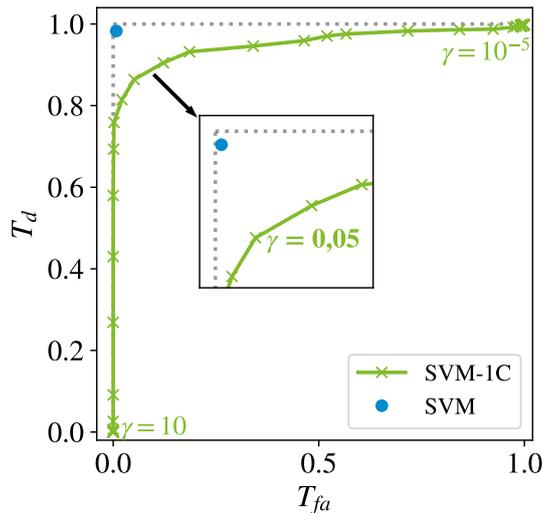


FIGURE 4.30 – Taux de détection en fonction du taux de fausse alarme pour les différents modèles testés en extérieur et avec des sources seules. La performance du SVM est représentée par le point \bullet tandis que celles des différents modèles générés pour les coefficients γ sont données par les croix \times .

Pour le modèle SVM le taux de détection moyen sur les 10 plis est de 99,1 % avec un taux de fausse alarme de 0,5 %. Pour le SVM-1C les performances atteintes sont diverses en fonction de γ , démarrant à un taux de détection et de fausse alarme à 1 pour $\gamma = 10^{-5}$ pour aller jusqu'à des taux de détection et de fausse alarme nuls pour $\gamma = 10$. Aucune valeur de γ ne permet d'atteindre les performances du SVM, avec comme compromis établi pour la suite une valeur de γ égale à 0,05, pour laquelle le taux de détection est égal à 86,9 % et celui de fausse alarme à 4,5 %. Cette valeur correspond au coude de la courbe obtenue pour l'ensemble des 50 modèles testés.

En comparant ces résultats à ceux obtenus en intérieur en Figure 4.15, le passage à l'extérieur a fait baisser les performances, ce qui est dû à la présence d'un bruit ambiant plus important lors des mesures extérieures, rendant les observations de drone moins ressemblantes de l'une à l'autre. Des discussions ont notamment été remarquées lors de l'écoute des enregistrements. Malgré cette baisse par rapport à l'intérieur, le modèle SVM surpasse les résultats de la littérature proches de ce cas d'application, c'est-à-dire en extérieur avec des signaux réellement produits [Jeo+17; Seo+18; Gan+20; Anw+19; Wan+18]. Le SVM-1C atteint seulement certains d'entre eux en fonction du paramètre γ qui est choisi. Il faut cependant nuancer cette comparaison car les configurations expérimentales sont toutes différentes d'une étude à l'autre : une base de donnée de référence est nécessaire pour comparer de manière fiable les chaînes de traitement proposées.

Capacité de généralisation des modèles : pour avoir une idée de la capacité de généralisation des deux modèles appris, les six autres

Identification extérieur : 1 source

- ▶ T_d du SVM : 99,1 %
- ▶ T_{fa} du SVM : 0,5 %
- ▶ T_d du SVM-1C : 86,9 %
- ▶ T_{fa} du SVM-1C : 4,5 %

sons Ext-ND-s1k/s3k/perc/hel1/hel2/mot¹⁵ joués par le haut-parleur en extérieur sont utilisés uniquement en phase de test, sans avoir été utilisés lors de la phase d'apprentissage où seules les données Ext-D6 et Ext-ND-bb1/bb2/voit1/voit2 sont conservées. Les résultats de la prédiction des deux modèles (à gauche pour le SVM / à droite pour le SVM-1C) sur 100 observations de chaque son sont donnés dans le Tableau 4.8. Elles sont correctement prédites si elles sont classées dans la classe *autre bruit*, la colonne de gauche étant censée rester nulle.

SVM / SVM-1C		Classe prédite	
		Drone	Autre bruit
Son joué	s1k	100 % / 0 %	0 % / 100 %
	s3k	99 % / 0 %	1 % / 100 %
	perc	95 % / 2 %	5 % / 98 %
	hel1	13 % / 14 %	87 % / 86 %
	hel2	13 % / 4 %	87 % / 96 %
	mot	18 % / 2 %	82 % / 98 %
Total		56,3 % / 3,7 %	43,7 % / 96,3 %

15: Ceux-ci sont aussi filtrés avant identification par un filtre passe-bas à 8 kHz.

TABLEAU 4.8 – Comparaison des résultats de prédiction pour les modèles SVM et SVM-1C sur six sons joués par le haut-parleur en extérieur qui ne sont pas utilisés lors de leur phase d'apprentissage.

La grande différence entre les résultats des deux modèles a lieu sur les sons de sinus à 1 et 3 kHz et celui de perceuse, avec un minimum sur ces trois-là de 98 % de bonne classification pour le modèle SVM-1C contre un maximum de 5 % pour le SVM. Pour ces sons qui sont dans leur structure fortement différents des sons utilisés pour l'apprentissage courant, un bruit blanc et un bruit de voiture, les caractéristiques calculées n'ont pas les mêmes valeurs et les observations obtenues se retrouvent donc situées du mauvais côté de l'hyperplan linéaire qui a été construit. En les ajoutant lors de l'apprentissage, comme cela a déjà été fait dans les parties précédentes pour le sinus à 1 kHz, l'hyperplan n'a aucun mal à s'adapter et les sépare à nouveau correctement des observations de drone. Pour le SVM-1C qui ne se soucie pas des observations de la classe *autre bruit* lors de la phase d'apprentissage, ces nouvelles observations ne sont pas assez ressemblantes à ce qu'il considère être un drone pour le perturber, et il les classe donc correctement.

Il est intéressant de noter que pour ce modèle, ce sont les observations d'hélicoptère qui engendrent le plus de mauvaise classification avec un taux de 14 % notamment pour le premier d'entre eux. Etant donné que ces sons avaient été choisis pour mettre en défaut si possible les modèles car ils sont acoustiquement similaires au son d'un drone, leur taux de mauvaise classification plus élevé que pour les autres confirme cette similarité.

Pour conclure, lorsque les données d'apprentissage n'englobent pas l'entière variabilité des sons qui peuvent advenir, le modèle SVM possède une capacité de généralisation moins importante que le SVM-1C même si celui-ci atteint des taux de détection et de fausse alarme plus faibles pour un apprentissage bien mené. Cette robustesse est liée au fait que l'hyperplan du SVM-1C n'est construit qu'à partir des observations de la classe drone, ce qui lui permet d'en être proche et de ne pas y classer des sons, même inédits. La partie suivante s'intéresse alors au problème inverse : le cas où les observations de drone sont perturbées lors de la phase de test, jusqu'à ne plus être reconnues comme telles.

Drone en présence d'une source perturbatrice

La suite de cette partie traite le problème de localisation et d'identification de drone en présence d'une seconde source acoustique, qui est ici perturbatrice car elle correspond au son joué par le haut-parleur. Ce problème n'est actuellement pas traité par la littérature, dont seules quelques contributions se limitent à l'identification de drone parmi plusieurs sources en présence grâce aux méthodes de séparation de sources aveugle [Gha+20; Udd+20].

Sources en présence : pour illustrer la difficulté de la configuration mise en place, la Figure 4.31 représente les densités spectrales de puissance de l'enregistrement d'un microphone unique de l'antenne lorsque le drone vole seul et lorsque le haut-parleur joue seul les deux sons proposés aux deux volumes différents.

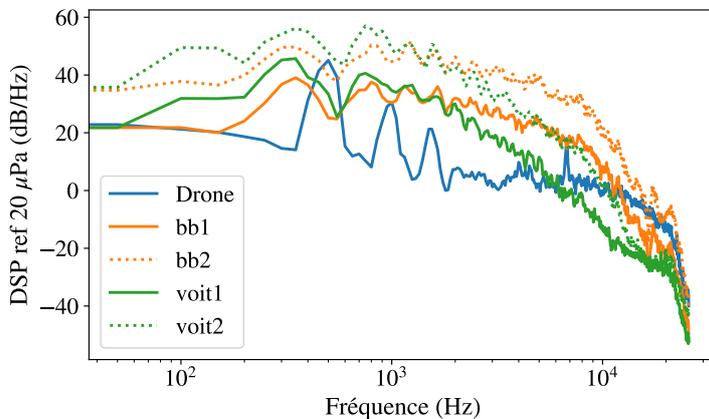


FIGURE 4.31 – Comparaison des densités spectrales de puissance obtenues en extérieur pour le drone (-), le son bb aux volumes 1 et au volume 2 (-, trait plein et trait pointillé), et le son voit aux volumes 1 et 2 (-, trait plein et trait pointillé).

En calculant les différences de niveaux ΔL_{moy} selon la définition donnée en Equation 4.6, on obtient -2,9 dB, -16,4 dB, -5,2 dB et -19,1 dB pour les bruits bb1, bb2, voit1 et voit2, respectivement. Ces cas correspondent ainsi à des sources fortement perturbatrices, pour lesquelles les prédictions de la classe drone de la Figure 4.26 n'étaient pas toutes correctes, voire en grande partie incorrectes.

Localisation simultanée de deux sources : un exemple de localisation pour le drone qui vole en même temps que le haut-parleur joue le son bb (en orange - dans la Figure 4.31) est donné avec en Figure 4.32a l'angle d'azimut et en Figure 4.32b l'angle d'élévation pour le volume 1 du son (bb1). La Figure 4.32c et la Figure 4.32d donnent ces mêmes angles pour le volume 2. L'angle d'azimut a été déroulé en partie pour obtenir une représentation plus claire de ses résultats. La vérité terrain donnée par le système RTK est représentée en orange (-) alors que la position du haut-parleur est donnée en vert (-). Les maxima des cartes de MUSIC successives sont donnés par les points de couleur ●, avec deux maxima estimés pour chaque séquence de 200 ms.

En comparant les deux lignes de la Figure 4.32, soit les deux volumes du son, il vient que la localisation des sources est fortement impactée par l'augmentation du volume sonore de la source perturbatrice avec l'apparition de nombreuses estimations ne correspondant pas au drone ou au haut-parleur. Pour l'angle d'élévation le drone est mieux localisé lorsqu'il s'élève alors que la localisation du haut-parleur n'est pas facile

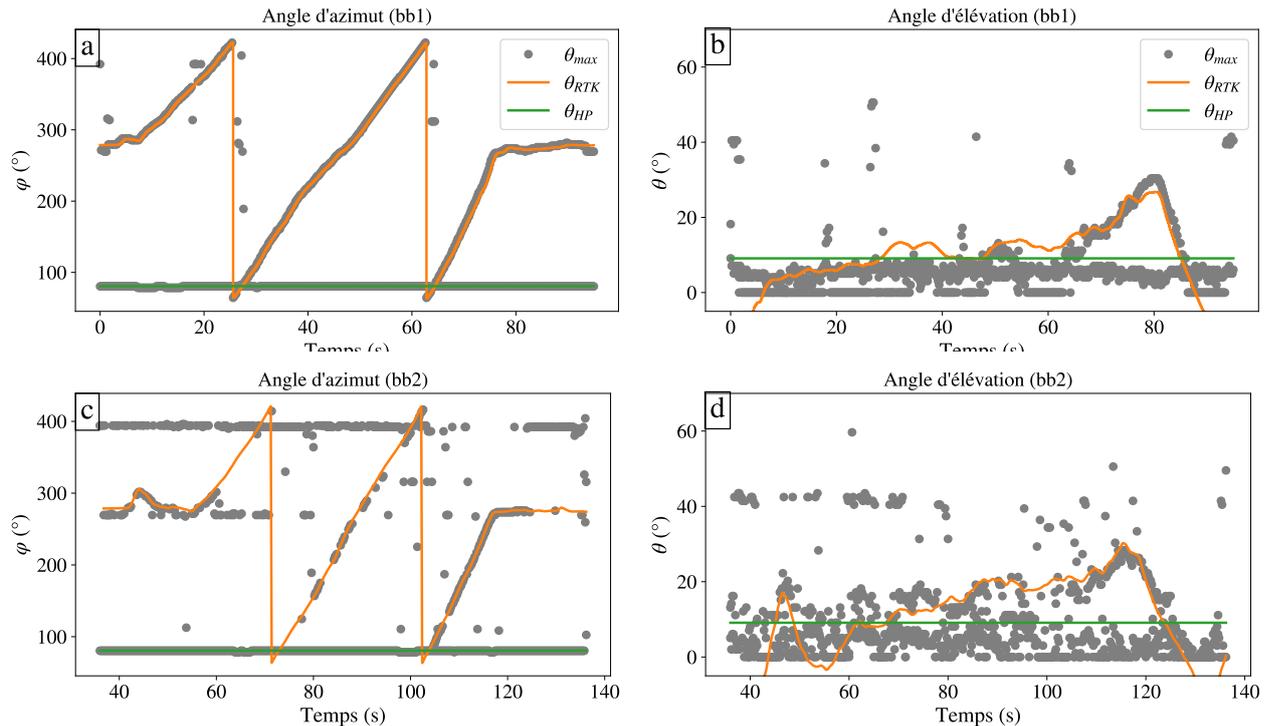


FIGURE 4.32 – Comparaison entre les angles estimés par les maxima des cartes de MUSIC (*) pour chaque séquence et les angles donnés par le système de localisation RTK pour le drone (-) et pour le haut-parleur (-). a : angle d'azimut pour le son bb1. b : angle d'élévation pour le son bb1. c : angle d'azimut pour le son bb2. d : angle d'élévation pour le son bb2.

à interpréter tant les estimations sont variées entre 0 et 15°. Lorsque le son émet plus fort, des estimations erronées apparaissent à $\theta = 40^\circ$. Pour l'angle d'azimut, très peu de mauvaises estimations sont faites au volume 1 alors qu'un grand nombre d'entre elles apparaissent au volume 2 juste sous les 400°. Aucune explication n'a été recherchée pour comprendre pourquoi les erreurs apparaissent pour cette valeur angulaire particulière.

Le fait de ne prendre que deux valeurs propres en présence de deux sources limite forcément les résultats de localisation présentés, étant donné qu'une sous-estimation du nombre de valeurs propres à prendre pour l'espace signal dans la méthode MUSIC entraîne des erreurs importantes d'estimation ensuite. Un axe d'amélioration de ces résultats consiste à passer d'un choix contraint du nombre de valeurs propres au nombre de sources attendues, à un choix automatique qui prendra mieux en compte les éléments de la décomposition de la matrice interspectrale.

Pour établir un bilan chiffré des résultats de localisation, la procédure de labélisation est à nouveau mise en place à partir des résultats de l'angle d'azimut. Pour ce cas, il est clair qu'un plus grand nombre de séquences focalisées vont générer des observations de la séquence inconnu étant donné le nombre important d'estimations situées loin des vérités terrains. Une fois la labélisation effectuée, le Tableau 4.9 donne pour chacun des quatre vols la proportion de séquences où le drone et le haut-parleur ont été localisés dans les seuils liés aux vérités terrain, et sur l'ensemble de la trajectoire suivie.

Alors que les deux première lignes donnent la proportion de séquences où le drone et le haut-parleur ont été correctement localisés pour

chaque cas, la troisième ligne donne la proportion de séquences inconnu générées durant la trajectoire étudiée.

Le drone est localisé à plus de 98 % lorsque le son joué est la voiture, peu importe le volume, ce qui est satisfaisant. Lorsque c'est le bruit blanc qui est joué, sa localisation est effective dans 94,7 % des cas au volume 1, et dans seulement 40,5 % des cas au volume 2 : le bruit blanc impacte fortement les performances en localisation du drone ce qui n'est pas satisfaisant d'un point de vue pratique car il est difficile de prévoir de quel type serait un son perturbateur à l'avance.

Pour le haut-parleur les performances ne varient pas de la même façon pour les deux sons. Alors qu'il est détecté dans plus de séquences lorsque le son qu'il joue augmente en volume pour le son de voiture, c'est l'inverse pour le bruit blanc. Pour expliquer pourquoi le bruit blanc perturbe autant la localisation, les 20 premières valeurs propres issues de la décomposition de la matrice interspectrale calculée lorsque le drone vole en présence de chacun des quatre sons sont tracées en Figure 4.33.

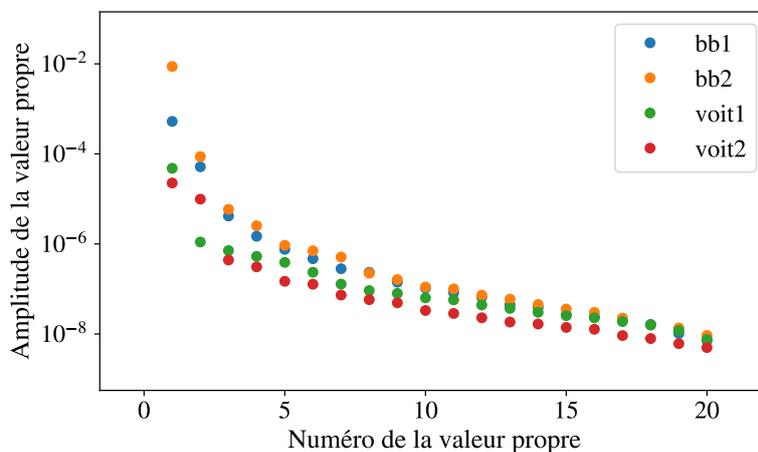


TABLEAU 4.9 – Résultats de localisation des deux sources en extérieur pour chacun des quatre vols avec les sons bb1/2 et voit1/2.

(%)	bb1	bb2	voit1	voit2
Drone	94,7	40,5	98,9	99,8
Haut-parleur	92,6	85,6	80,2	99,3
Inconnu	6,5	37,1	10,7	0,7

FIGURE 4.33 – Décroissance des valeurs propres observées pour une matrice interspectrale calculée sur une séquence de signal lorsque le drone vole alors que le haut-parleur émet un son.

Les amplitudes des différentes valeurs propres sont tracées en échelle logarithmique. Au-delà de la valeur propre numéro 20 pour tous les sons, un saut de neuf ordres de grandeur dû aux paramètres utilisés pour l'estimation de la matrice interspectrale est observé, d'où un tracé seulement pour ces premières valeurs. Les valeurs propres pour les sons bb1 et bb2 sont supérieures en amplitude à celles des sons voit1 et voit2, et décroissent plus lentement.

Dans la thèse, seules une ou deux valeurs propres sont conservées pour constituer le sous-espace signal, suivant le nombre de sources attendues. Ce choix est censé être adapté aux configurations étudiées dans lesquelles le nombre de sources acoustiques est connu, mais il apparaît clairement ici qu'un choix lié à la décroissance réelle des valeurs propres pourrait permettre d'améliorer les résultats de localisation.

Concernant les séquences de la classe inconnu, celles-ci sont bien générées en plus grande proportion lorsque l'une ou l'autre des sources n'a pas été correctement localisée. Ainsi pour le son bb2 elles représentent 37,1 % du nombre de séquences focalisées total à classer contre seulement 0,7 % dans le cas du son voit2 où les deux sources sont bien localisées.

A la suite de l'étape de localisation pour l'ensemble des quatre vols, 82,4 % des séquences de drone sont localisées, contre 89,4 % des

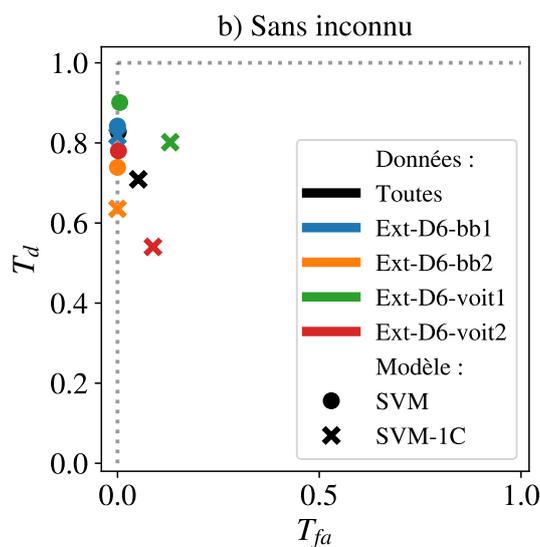
Localisation extérieur : 2 sources

- ▶ Drone localisé : 82,4 %
- ▶ Haut-parleur localisé : 89,4 %
- ▶ Inconnu générés : 14,3 %

séquences de son du haut-parleur. 536 séquences inconnues sont générées à la suite de l'étape de localisation, ce qui représente 14,3 % du nombre de séquences totales à classer en prenant les deux sources. La localisation avec deux sources en présence de manière simultanée estime des positions de sources correctes pour 85,9 % des séquences au global¹⁶. La précision en fonction des angles n'est pas étudiée ici mais elle est similaire à celle observée pour le drone seul au vu des résultats de la Figure 4.32.

Identification simultanée de deux sources : pour toutes ces séquences, la focalisation est entreprise puis les caractéristiques sélectionnées sont calculées pour produire les observations à apprendre et tester. Grâce à la procédure de labélisation menée pour évaluer les performances de localisation, les labels de ces différentes observations à classer sont connus. Le modèle SVM servant ensuite à les classer est appris sur l'ensemble des données qui ont servi à la validation croisée, soit 348 observations correspondant à des séquences de drone et de autre bruit seules. Pour le modèle SVM-1C, les 2298 observations de drone sont utilisées avec le coefficient γ établi à 0,05.

La Figure 4.34 donne alors les taux de fausse alarme et de détection pour ces deux modèles, pour l'ensemble des observations testées et pour chaque son séparément. Le but étant d'évaluer la performance de l'identification seule ici, les observations labélisées inconnu ne sont pas



16: La somme des séquences correctement et incorrectement localisées n'est pas égale à 100 % étant donné que pour quelques-unes d'entre elles, seul un maximum a été trouvé sur la carte de MUSIC.

FIGURE 4.34 – Résultats d'identification à deux sources en extérieur. Chaque point correspond au taux de détection en fonction du taux de fausse alarme pour un cas testé. Le résultat pour l'ensemble des données de test est donné en noir (•, ×), celui pour le son bb1 en bleu (•, ×), pour le son bb2 en orange (•, ×), pour le son voit1 en vert (•, ×), et pour le son voit2 en rouge (•, ×). Les points • représentent les performances du modèle SVM et les croix × celles du SVM-1C.

Pour l'ensemble des données de test, le SVM atteint un taux de détection de 83,9 % pour un taux de fausse alarme de 0,2 % (le point qui correspond à ses performances est masqué par la croix et le point bleus). Le SVM-1C détecte quant à lui 70,9 % des séquences drone pour une fausse alarme de 5,1 %. Les observations de la classe autre bruit sont ainsi mieux identifiées que celles de drone lors de l'identification des deux sources.

Quel que soit le son perturbateur, le SVM-1C produit des taux de détection variant entre 54 et 82 %, contre 74 à 90 % pour le SVM. Les performances observées pour les sources présentes individuellement se confirment avec un modèle SVM qui détecte mieux les observations de drone.

Le taux de fausse alarme varie entre 0 et 13 %, avec pour le modèle SVM un taux de fausse alarme maximum de 0,6 % atteint pour le son voit1. Par conséquent, pour des séquences correctement localisées, le modèle SVM prédit les observations avec un taux de fausse alarme quasi nul.

En analysant son par son, il est intéressant de noter que lorsque le volume d'un son augmente, le taux de détection obtenu diminue quel que soit le son. De plus, le taux de fausse alarme est nul pour les deux sons bb1 et bb2 alors qu'ils sont à 13,1 et 8,7 % pour voit1 et voit2. Le taux de fausse alarme observé est fortement dépendant du son qui est émis.

Pour conclure, le modèle SVM est plus robuste à la source perturbatrice que le SVM-1C. Il détecte plus d'observations de drone sans générer de fausse alarme, et ce quel que soit le son perturbateur. La partie suivante vérifie ensuite si cette identification dépend de la distance entre le drone et l'antenne, concluant sur la robustesse complète de la chaîne de traitement.

Identification extérieur : 2 sources

T_d avec le modèle SVM :

- ▶ Drone : 83,9 %
- ▶ Autre bruit : 99,8 %

T_d avec le modèle SVM-1C :

- ▶ Drone : 70,9 %
- ▶ Autre bruit : 94,9 %

Performances finales en fonction de la distance

Pour établir les performances de la chaîne de traitement en fonction de la distance entre le drone et l'antenne, la Figure 4.35 représente, pour chaque label donné lors de la labélisation, les distributions des observations correctement et incorrectement classées en fonction de la distance entre le drone et l'antenne. Les distributions brutes sont données par les marqueurs de couleur, en vert (●) pour des observations bien classées et en rouge (●) pour des observations mal classées. Les statistiques de ces distributions sont données par les boîtes à moustaches avec leur moyenne représentée par le trait horizontal à l'intérieur de la boîte, leur quartile par la boîte et les valeurs minimales et maximales avec les moustaches. Etant donné que le système RTK n'a pas fonctionné jusqu'à la fin du dernier vol pour le son voit2, seules les observations pour lesquelles la distance est disponible sont tracées, dont le nombre est donné au-dessus de chaque distribution.

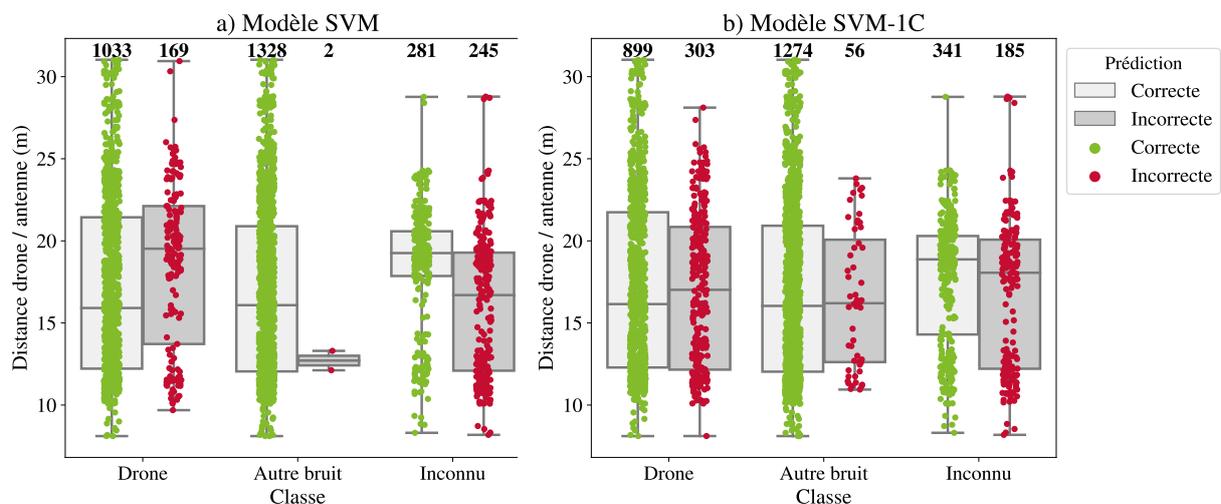


FIGURE 4.35 – Performances en identification pour le cas deux sources en extérieur en fonction de la distance entre le drone et l'antenne. Les données sont séparées en abscisse selon les trois classes, puis pour chacune les données correctement identifiées sont données en vert (●), alors que celles incorrectement identifiées sont données en rouge (●). a : résultats pour le modèle SVM. b : résultats pour le modèle SVM-1C.

Les résultats d'identification obtenus avec le modèle SVM, en Figure 4.35a, montrent que seulement deux observations sont mal identifiées pour les observations de la classe *autre bruit*, et qu'elles le sont pour un drone situé à environ 13 m de l'antenne. Un nombre trop faible d'erreur est commis pour conclure grâce à cette classe. Par conséquent, ce sont les résultats d'identification des deux autres classes qui permettent d'affirmer que l'identification faite est indépendante de la distance entre le drone et l'antenne. Pour ces deux classes, les distributions d'observations classées de manière correcte ou incorrecte sont similaires, quelle que soit la distance entre le drone et l'antenne.

Pour le modèle SVM-1C, où un plus grand nombre d'erreurs est commis pour la classe *autre bruit*, les statistiques des distributions des trois classes sont similaires. Ceci conforte le fait que les performances en identification ne dépendent pas de la distance entre le drone et l'antenne.

Avec ces derniers résultats qui permettent d'analyser plus finement la phase d'identification, celle-ci est validée dans ce cas difficile avec un drone en présence d'une source perturbatrice.

Performances de la chaîne entière : en conclusion pour la chaîne entière sur ce cas avec source perturbatrice, le modèle SVM détecte 70,0 % des drones sur l'ensemble des séquences avec un taux de fausse alarme de 13,5 %. Le modèle SVM-1C possède un taux de détection de seulement 58,4 % avec un taux de fausse alarme de 14,7%. L'amélioration principale de ces performances passe par une meilleure localisation des sources acoustiques en présence, afin de bénéficier des bonnes performances de l'étape d'identification qui ont été démontrées avec un taux de fausse alarme quasi nul. Pour améliorer encore la détection, la robustesse des modèles d'apprentissage doit être travaillée pour mieux identifier les observations de drone, surtout pour le SVM-1C lorsqu'une source perturbatrice est présente.

Chaîne entière extérieur : 2 sources

Avec le modèle SVM :

- ▶ $T_d = 70,0 \%$
- ▶ $T_{fa} = 13,5 \%$

Avec le modèle SVM-1C :

- ▶ $T_d = 58,4 \%$
- ▶ $T_{fa} = 14,7 \%$

4.10 Conclusion

Pour traiter le problème de localisation et d'identification de drone, ce chapitre propose une chaîne de traitement en **trois étapes** qui couple traitement d'antenne et apprentissage statistique. Une localisation est d'abord effectuée sur des **séquences** successives de **200 ms** pour déterminer la direction d'arrivée angulaire d'une ou de deux sources acoustiques en présence. Chacune est filtrée spatialement pour focaliser l'écoute de l'antenne et améliorer le rapport signal à bruit de la source visée. Enfin, 87 caractéristiques spécifiquement dédiées à la bonne description de signaux acoustiques sont calculées sur ces séquences temporelles focalisées pour obtenir des observations qui sont comparées à l'hyperplan d'un modèle SVM ou SVM-1C appris au préalable.

L'aspect expérimental est au cœur de ce travail. C'est pourquoi 1 heure et 15 minutes de mesures en intérieur et 20 minutes en extérieur ont été produites durant les trois ans de thèse. Associées à des systèmes établissant une vérité terrain précise, elles permettent de montrer que les drones ne sont jamais localisés en intérieur à plus de 8° d'erreur en moyenne selon l'angle d'élévation, et 4° en azimut. Pour l'extérieur, la

précision moyenne sur les deux angles est de 3° . Pour un drone situé à 30 m de l'antenne, cela signifie une erreur de position de 1,6 m.

Il a été montré que ce sont les **réflexions qui causent les principales erreurs** sur l'angle d'élévation, et qu'elles sont prépondérantes à basse fréquence. Le modèle en ondes planes s'avère quant à lui robuste, même si les sources acoustiques ne se situent pas à une distance de l'antenne grande D_{lim} . Pour améliorer la précision de la localisation en extérieur, un traitement à plus haute fréquence est alors entrepris, réduisant bien l'erreur moyenne faite sur l'angle d'élévation.

Une fois les drones localisés, une identification est entreprise. Pour justifier de la pertinence des caractéristiques qui décrivent les signaux acoustiques, une classification à quatre classes est mise en place. Ses résultats montrent qu'avec un modèle SVM linéaire, les **caractéristiques classent avec une précision globale de 98,6 % trois types de drones différents**. Finalement, 71 caractéristiques spécifiquement optimisées pour le problème d'identification à résoudre sont gardées pour traiter les configurations suivantes.

Que ce soit en intérieur ou en extérieur, **l'identification est très précise pour des sources acoustiques présentes tour à tour** aux environs de l'antenne. Le SVM détecte en extérieur 99,1 % des observations de drones, et le SVM-1C en détecte 86,9 %. Le taux de fausse alarme est lui faible avec 0,5 % pour le SVM et 4,5 % pour le SVM-1C. Puis lorsque deux sources sont présentes simultanément avec des niveaux sonores différents, leur localisation et identification est acquise grâce au couplage entre le traitement d'antenne et l'apprentissage statistique, alors qu'un microphone unique identifie en priorité la source qui émet le plus fort. Des données en intérieur valident alors la chaîne entière avec un taux de détection de drone de 80,4 % pour un taux de fausse alarme de 4,9 %.

Pour terminer la description des capacités de la chaîne de traitement, les données des vols du **drone en extérieur en présence de sources perturbatrices émettant jusqu'à 19,1 dB plus puissamment** que le drone sont traitées. Dès l'étape de localisation, des erreurs sont commises si bien que seules 82,4 % des séquences où le drone vole sont correctement localisées et 89,4 % de celles liées au haut-parleur. Pour ces observations correctement localisées, le **drone est identifié à 83,9 % avec le modèle SVM**, alors que les sons émis par le haut-parleur ne génèrent que 0,2 % de fausse alarme. Le **modèle SVM-1C lui atteint des performances de 70,9 % en détection**, et 5,1 % en fausse alarme, sur ces deux classes. Le modèle SVM est donc plus robuste aux sources perturbatrices que le SVM-1C. Tous ces résultats d'identification sont indépendants de la distance entre le drone et l'antenne, ce qui valide la chaîne de traitement pour un drone qui vole jusqu'à 31 m de l'antenne.

Pour conclure, la chaîne de traitement entière avec le modèle SVM détecte à ce jour en extérieur, et pour des sources présentes tour à tour, 7 min 36 s de signal de drone en générant 0,3 s de fausse alarme sur un enregistrement de 7 min 40 s. Avec une source perturbatrice, ce sont 4 min 15 s qui sont détectées en générant 44 s de fausse alarme, le tout pour 6 min 14 de signal enregistré. La **chaîne de traitement est donc particulièrement efficace pour des cas calmes** en présence de peu de perturbations, et elle est d'ores et déjà capable de détecter une proportion importante de drones dans des environnements pour lesquels des sources

acoustiques dont le niveau sonore est 19 dB supérieur à celui du drone au niveau de l'antenne sont présentes.

Conclusion

Dans cette thèse, des méthodes de la littérature du traitement d'antenne et de l'apprentissage supervisé ont été combinées afin de mettre en place des prototypes industrialisables pour répondre à deux besoins applicatifs concrets : la localisation et quantification de machines excavatrices pour les sites miniers marins par grand fond, et la localisation et l'identification de drones aériens aux abords de sites sensibles.

Dans les deux premiers chapitres, les grandeurs acoustiques nécessaires à la modélisation des sources acoustiques ont été présentées, pour aboutir à une description de celles-ci sous forme de monopôles acoustiques. Puis pour les caractériser, des méthodes de traitement de signal ont été sélectionnées grâce à une analyse de la littérature du traitement d'antenne et de l'apprentissage supervisé. La méthode *Multiple Signal Classification* (MUSIC) assure alors la localisation des sources, la formation de voies est utilisée pour la conception d'antenne et la quantification de niveau sonore, et enfin les modèles *Support Vector Machines* (SVM) et SVM à une classe (SVM-1C) identifient les signaux acoustiques dans des classes prédéfinies.

Les deux chapitres suivants démontrent que ces méthodes sont capables de localiser précisément des sources acoustiques, que ce soit en air ou en eau, mais aussi de les quantifier dans les grands fonds marins et de les identifier en aérien. La localisation en acoustique sous-marine est ainsi précise à 14 m près en grand fond, et la quantification est retrouvée par rapport à un niveau de référence pour une source monochromatique mobile, alors qu'elle est retrouvée de manière relative pour une source à large bande. Pour les drones, l'erreur de localisation en extérieur est inférieure à 3°, et la chaîne de traitement proposée détecte plus de 85 % du temps le drone lorsqu'il est seul en présence, quel que soit le modèle d'apprentissage utilisé.

De tous les résultats obtenus dans cette thèse, trois grands enseignements peuvent être tirés :

La méthode haute-résolution MUSIC permet de localiser précisément des sources peu nombreuses dans un environnement acoustique réel.

La capacité de MUSIC en résolution est démontrée à travers la séparation de sources spatialement proches mais aussi de niveaux sonores différents, dans toutes les configurations rencontrées. Par rapport à la formation de voies, la source la plus faible n'est pas masquée par les lobes secondaires de la source la plus puissante. La localisation obtenue est précise dans un environnement acoustique réel, que ce soit en intérieur ou en extérieur, et en eau ou en air. En présence d'une source perturbatrice, la précision de la localisation reste la même lorsque la source est trouvée, mais la méthode peut donner lieu à des localisations erronées au cours du temps.

Etant donné le nombre assez faible de références scientifiques employant cette méthode pour des applications expérimentales en acoustique, ce travail permet de montrer son utilité dans des environnements calmes. Des essais supplémentaires sont requis pour valider son fonctionnement dans des cas de figure plus bruités, avec par exemple un fort bruit ambiant sous-marin dans des cas de mer agitée, ou encore la présence de vent fort pour des applications aériennes.

Dans cette thèse, la méthode a été appliquée avec un choix fixe du nombre de valeurs propres pour composer le sous-espace signal, ce qui n'est pas optimal. Il est nécessaire de travailler sur la détection automatique du nombre de valeurs propres à sélectionner dans les deux cas d'application afin d'amener les systèmes proposés vers des produits industriels complets. Pour le cas des drones, ce travail est d'autant plus essentiel que pour le moment ce nombre est réglé en fonction du nombre de sources attendues dans la configuration traitée, qui sera inconnu dans une application réelle. Trouver une solution à ce problème permettra non seulement de proposer un système industriel finalisé, mais aussi de régler les problèmes de localisation observés en présence de plusieurs de sources simultanément, responsables de l'augmentation du taux de fausse alarme. En réglant ce problème, seul le taux de détection reste à optimiser.

Un système d'écoute spécifiquement conçu pour être déployé proche de sources acoustiques sous-marines à caractériser permet l'obtention de résultats précis en localisation et en quantification.

La formation de voies a permis la quantification de sources acoustiques sous-marine par grand fond. Ce succès obtenu avec une méthode optimale pour la quantification d'une source seule dans l'environnement acoustique total est certainement basé sur les choix de conception qui ont été menés. Ils permettent d'avoir une antenne avec une structure robuste, localisant précisément autour de son axe vertical et en trois dimensions, avec des distances entre la source et l'antenne relativement faibles dans le cas visé, i.e. le cas grand fond. En proposant un système d'écoute qui vient se placer au plus proche de la source à caractériser sans risque de choc avec les éléments de la zone, une propagation acoustique en champ libre peut être utilisée avec des résultats de localisation et de quantification se situant tous dans les bornes d'incertitudes des mesures de référence, validant l'ensemble de la solution industrielle proposée grâce à un prototype.

Le grand défi de cette application est désormais de passer à l'échelle, en cherchant à caractériser de vraies machines excavatrices dans des environnements miniers réels, apportant ainsi des défis de taille que sont des éléments encore plus imposants à déployer en sécurité, et des bruits ambiants beaucoup plus élevés pour des sites dans lesquels beaucoup de machines travaillent de concert. Dans ce contexte, la formation de voies dans sa version quantitative risque d'être limitée, et le passage à des méthodes de déconvolution en exploitant plusieurs plans de localisation parallèles semble être une piste à privilégier.

Le couplage entre traitement d'antenne et apprentissage statistique permet une identification précise de drone, en offrant la possibilité de filtrer la présence d'une source perturbatrice forte.

L'identification de drone par les modèles SVM et SVM-1C est très précise en présence d'une source à la fois autour de l'antenne acoustique, et prometteuse avec la présence d'une source perturbatrice. Le modèle SVM a d'ailleurs prouvé sa capacité de séparation des sources acoustiques en classant avec succès trois types de drones séparément.

L'innovation majeure de ce travail, qui réside dans le couplage entre traitement d'antenne et apprentissage statistique, permet d'envisager une chaîne de traitement pour la localisation et l'identification de drone robuste aux sources acoustiques perturbatrices, mais aussi aux bruits ambiants trop élevés sur les zones potentielles de déploiement.

Pour passer à une solution industrielle finalisée, le traitement en temps réel doit être mis en place pour aboutir à un système utilisable dans des conditions où la menace doit être détectée instantanément. En parallèle, des bases de données réalistes de signatures acoustiques doivent être construites pour décrire fidèlement les signaux potentiellement présents lors de campagnes réelles.

Enfin, pour atteindre de meilleures performances en détection dans le cas le plus complexe étudié dans cette thèse, trois pistes de travail peuvent être envisagées par ordre de priorité :

- ▶ l'amélioration de la localisation déjà évoquée plus haut : en absence de séquences mal localisées, le taux de fausse alarme est déjà quasi nul avec le modèle SVM, mais le taux de détection lui n'est pas fortement modifié ;
- ▶ l'amélioration du filtrage spatial : en utilisant des algorithmes de traitement d'antenne adaptatifs, l'importance des résidus de la source perturbatrice dans le signal focalisé peut être grandement réduite, ce qui permettrait de ne plus modifier les distributions des caractéristiques pour une classe donnée ;
- ▶ une étude plus poussée sur les caractéristiques : en identifiant les caractéristiques les plus robustes, les résultats d'identification peuvent être améliorés car leur distribution sera insensible à la présence d'une source perturbatrice. Cette sélection possède aussi l'avantage de simplifier les modèles d'apprentissage à apprendre grâce à une recherche d'hyperplan dans un espace de dimension réduit, et d'améliorer les temps de calcul.

ANNEXES



Description du projet Abysound

Le projet **Abysound** est un projet FUI (Fonds Unique Interministériel) 22 financé par par BPI France (Banque Publique d'Investissement) et par le Pôle Mer Méditerranée. Il est porté par Naval Group et regroupe des partenaires industriels que sont OSEAN SAS, MicrodB et Semantic TS, et des organismes de recherche, le LMA (Marseille), le Gipsa-Lab (Grenoble), l'IFREMER (Toulon) et le LSIS (Toulon).

L'objectif du projet Abysound est d'étudier la possibilité d'estimer le bruit rayonné par des machines excavatrices en grand fond (1500 m), grâce à un système d'écoute spécifiquement conçu. Le système doit permettre de localiser et quantifier les sources à caractériser, puis des méthodes de propagation permettent de prédire l'impact des niveaux sonores dans le milieu environnant. Une dernière étape consiste à regarder si ces niveaux dépassent les seuils de danger concernant la surdité temporaire ou permanente de la faune marine pouvant être présente aux abords des site concernés.

Une tâche supplémentaire dans le projet consiste à regarder l'impact du bruit potentiellement élevé aux abords des sites industriels sous-marins sur le fonctionnement acoustique de véhicules sous-marins autonomes. En cas de perturbations acoustiques trop importantes, des stratégies d'évitement sont mises en place pour éviter les accidents dans la zone de travail.

Le projet a donné lieu à un travail méthodologique pour concevoir le système et les méthodes de localisation et de quantification des sources acoustiques. C'est le travail détaillé dans cette thèse. Mais d'autres avancées méthodologiques ont été faites :

- ▶ un interfaçage dédié a été développé entre la méthode SPEC FM et les méthodes de tracé de rayons et parabolique. Il permet de prendre en compte des modèles de sources beaucoup plus précis en entrée des méthodes de propagation, tout en n'étant pas trop consommateur de ressources calculatoires ;
- ▶ un système de détection de niveau sonore trop élevé basé sur un rapport de vraisemblance généralisé a été mis en place afin de permettre à des véhicules sous-marins autonomes de rejoindre des zones acoustiquement calmes lorsque le bruit ambiant qui les entoure est trop important.

Puis une composante importante du projet concerne l'application concrète en mer des méthodes sélectionnées, à travers deux campagnes de mesures effectuées en avril 2018 au large de la baie de Cavalaire, et en mai 2019 au large de Saint-Tropez. La première a duré deux jours, et a permis de valider le déploiement et la bonne construction de l'antenne, grâce à des essais effectués dans une zone de 100 m de profondeur. L'année suivante, des essais ont été menés pendant 5 jours, permettant de valider la localisation et la quantification de source, y compris dans un cadre grand fond avec une antenne et une source immergées à 700 m de profondeur.

Pour conclure, ce projet a donné lieu à diverses contributions scientifiques en conférences [Bas+19; Jau+19; Pha+19a; Cri+19; Pha+19b] et dans le journal *IEEE Journal of Oceanic Engineering* [Bar+20].

Caractéristiques pour l'identification de drone

L'ensemble des 29 caractéristiques utilisées pour l'identification de drone et leur formule associée est donné dans le tableau ci-dessous. Elles sont calculées à partir d'un signal générique x_i composé de n échantillons, qui représente le signal dans l'un des trois domaines d'étude : temporel, spectral ou cepstral.

Remarques :

- $E = \sum_{i=1}^n x_i^2$ est l'énergie du signal

TABLEAU B.1 – Liste des caractéristiques pour l'identification de drone.

Caractéristique	Description	Calcul
Centroïde	Point autour duquel l'énergie est centrée	$\bar{i} = \frac{1}{E} \sum_{i=1}^n i \cdot x_i^2$
Bande d'énergie	Largeur de la bande d'énergie autour du centroïde	$B = \sqrt{\sum_{i=1}^n \frac{i^2 x_i^2}{E} - \bar{i}^2}$
Taux d'attaque	Pente croissante maximum entre deux échantillons du signal	$\max_i \left(\frac{x_i - x_{i-1}}{n} \right)$
Taux d'affaiblissement	Pente décroissante minimum entre deux échantillons du signal	$\min_i \left(\frac{x_i - x_{i+1}}{n} \right)$
Taux de croisement (3 seuils)	Nombre de fois où le signal passe par 0	$\frac{\#(\text{seuil croisé})}{n}$
Taux de silence (3 seuils)	Nombre d'échantillons situés sous un seuil donné	$\frac{\#(\text{échantillons} < \text{seuil})}{n}$
Minimum	Amplitude minimum du signal	$\min_i(x_i)$
Maximum	Amplitude maximum du signal	$\max_i(x_i)$
Moyenne	Moyenne du signal	$\mu = \frac{\sum_{i=1}^n x_i}{n}$
Minimum sur moyenne	Minimum du signal divisé par la moyenne	$\frac{\min_i(x_i)}{\mu}$
Maximum sur moyenne	Maximum du signal divisé par la moyenne	$\frac{\max_i(x_i)}{\mu}$
Ecart-type	Dispersion des valeurs d'amplitude du signal	$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2}$
Skewness	Mesure de l'asymétrie du signal	$\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^3$
Kurtosis	Mesure de l'aspect piqué du signal	$\frac{1}{n} \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right)^4$
Skewness centré	Skewness centré sur le centroïde	$\sqrt{\frac{\sum_{i=1}^n (i - \bar{i})^3 x_i^2}{EB^3}}$
Kurtosis centré	Kurtosis centré sur le centroïde	$\sqrt{\frac{\sum_{i=1}^n (i - \bar{i})^4 x_i^2}{EB^4}}$
Entropie de Shannon (3 réglages)	Mesure de la quantité d'information portée par le signal	$-\sum_{i=1}^n p(x_i) \log_2(p(x_i))$
Entropie de Rényi (6 réglages)	Mesure plus générale de l'entropie	$\frac{1}{1-\alpha} \log_2(\sum_{i=1}^n p(x_i)^\alpha)$



Liste des publications

Liste des publications liées à mon travail de thèse

- ▶ V. BARON, A. FINEZ, S. BOULEY, F. FAYET, J. MARS et B. NICOLAS. "Hydrophone array optimization, conception, and validation for localization of acoustic sources in deep-sea mining". In : *IEEE Journal of Oceanic Engineering*, 2020, p. 1-9.
- ▶ V. BARON, S. BOULEY, J. MARS, GT. PHAM et B. NICOLAS. "Experimental validation of an acoustic monitoring system for deep-sea mining". In : *IEEE Journal of Oceanic Engineering*. Soumis en décembre 2020.
- ▶ R. BASSILA, T. BERTET, L. SOMAGLINO, MÈ. BOUHIER, H. GLOTIN, P. BEST, V. BARON, C. NOËL, P. CRISTINI, F. FAYET, B. NICOLAS et J. MARS. "ABYSOUND, an end to end system for noise impact measurement of deep sea mining production tools". In : *MTS/IEEE Oceans 2019*. Marseille, 2019, p. 1-9.
- ▶ J. JAUMONT, V. BARON, B. NICOLAS, F. FAYET, A. FINEZ et J. MARS. "Conception of an underwater prototype array : from simulation design to mechanical assessment". In : *MTS/IEEE Oceans 2019*. Marseille, 2019, p. 1-9.
- ▶ GT. PHAM, V. BARON, A. FINEZ, J. MARS et B. NICOLAS. "High resolution source localization in underwater acoustics for deep sea mining monitoring". In : *MTS/IEEE Oceans 2019*. Marseille, 2019, p. 1-7.
- ▶ GT. PHAM, J. MARS, V. BARON, A. FINEZ et B. NICOLAS. "Localisation de sources de bruit en prospection minière sous-marine : une preuve de concept". In : *XXVIIème colloque GRETSI*. Lille, 2019.
- ▶ V. BARON, S. BOULEY, M. MUSCHINOWSKI, J. MARS et B. NICOLAS. "Localisation et identification acoustique de drones par mesures d'antennerie et apprentissage supervisé". In : *XXVIIème colloque GRETSI*. Lille, 2019.
- ▶ V. BARON, S. BOULEY, M. MUSCHINOWSKI, J. MARS et B. NICOLAS. "Drone localization and identification using an acoustic array and supervised learning". In : *SPIE Artificial Intelligence and Machine Learning in Defense Applications*. Strasbourg, 2019.
- ▶ V. BARON, A. FINEZ et B. NICOLAS. "Numerical and experimental assessment of functional beamforming for source quantification". In : *7th Berlin Beamforming Conference*. Berlin, 2018, p. 1-15.
- ▶ S. BARCET, V. BARON et A. CARBONELLI. "Gears and bearings defaults : from classification to diagnosis using machine learning". In : *Surveillance, Vishno and AVE Conferences*. Lyon, 2019.

Bibliographie

Bibliographie du chapitre 1

- [Bru98] M. BRUNEAU. *Manuel d'acoustique fondamentale*. Français. Paris : Hermes, 1998 (cf. p. 6).
- [Mor+87] P. MORSE et K. INGARD. *Theoretical Acoustics*. en. Princeton University Press, jan. 1987 (cf. p. 6, 8, 9).

Bibliographie du chapitre 2

- [Ace+09] M. ACEVEDO, C. CORRADA-BRAVO, H. CORRADA-BRAVO, L. VILLANUEVA-RIVERA et T. Mitchell AIDE. "Automated classification of bird and amphibian calls using machine learning : a comparison of methods". In : *Ecological Informatics* 4.4 (sept. 2009), p. 206-214 (cf. p. 32).
- [Aka98] H. AKAIKE. "Information theory and an extension of the maximum likelihood principle". In : *Parzen E., Tanabe K., Kitagawa G., Selected Papers of Hirotugu Akaike*. Springer Series in Statistics. New York : Springer, 1998 (cf. p. 28).
- [Ala+14] X. ALAMEDA-PINEDA et R. HORAUD. "A Geometric Approach to Sound Source Localization from Time-Delay Estimates". In : *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 22.6 (juin 2014), p. 1082-1095 (cf. p. 19).
- [ANS19] ANSI/ASA S12.64 PART 1. "American National Standard Quantities and Procedures for Description and Measurement of Underwater Sound from Ships – Part 1 : General Requirements". In : *American National Standards Institute* (sept. 2019) (cf. p. 17, 45).
- [Ant12] J ANTONI. "A Bayesian approach to sound source reconstruction : Optimal basis, regularization, and focusing". In : *The Journal of the Acoustical Society of America* 131.4 (avr. 2012), p. 2873-2890 (cf. p. 17, 22).
- [Bad05] R. BADEAU. "Méthodes à haute résolution pour l'estimation et le suivi de sinusoides modulées. Application aux signaux de musique". Français. Thèse de doct. Ecole Nationale Supérieure des Télécommunications, avr. 2005, p. 230 (cf. p. 21).
- [Bai92] M. BAI. "Application of BEM (boundary element method)-based acoustic holography to radiation analysis of sound sources with arbitrarily shaped geometries". In : *The Journal of the Acoustical Society of America* 92.1 (juill. 1992), p. 533-549 (cf. p. 22).
- [Bar+15] D. BARCHIESI, D. GIANNOULIS, D. STOWELL et M. PLUMBLY. "Acoustic scene classification : classifying environments from the sounds they produce". In : *IEEE Signal Processing Magazine* 32.3 (mai 2015), p. 16-34 (cf. p. 32).
- [Bar11] A. BARTKOWIAK. "Anomaly, novelty, one-class classification : a comprehensive introduction". In : *International Journal of Computer Information Systems and Industrial Management Applications* 3 (2011), p. 61-71 (cf. p. 38).
- [Ben+08] J. BENESTY, J. CHEN et Y. HUANG. *Microphone Array Signal Processing*. Springer Science & Business Media, mars 2008 (cf. p. 19).
- [Ben+00] K. BENNETT et E. BREDENSTEINER. "Duality and geometry in SVM classifiers". In : *17th International Conference on Machine Learning*. Stanford, USA, 2000, p. 57-64 (cf. p. 37).
- [Bia+19] M. BIANCO, P. GERSTOFT, J. TRAER, E. OZANICH, M. ROCH, S. GANNOT et C. DELEDALLE. "Machine learning in acoustics : Theory and applications". In : *The Journal of the Acoustical Society of America* 146.5 (nov. 2019), p. 3590-3628 (cf. p. 32, 34).
- [Bie+83] G. BIENVENU et L. KOPP. "Optimality of high resolution array processing using the eigensystem approach". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 31.5 (oct. 1983), p. 1235-1248 (cf. p. 22, 28).

- [Bie+85] G. BIENVENU et L. KOPP. "New Approaches in the Adaptive Array Processing Field". In : *Adaptive Methods in Underwater Acoustics*. Sous la dir. d'Heinz G. URBAN. NATO ASI Series. Dordrecht : Springer Netherlands, 1985, p. 375-384 (cf. p. 18, 44).
- [Bir+17] O. BIRAN et C. COTTON. "Explanation and justification in machine learning : a survey". In : *26th International Joint Conference on Artificial Intelligence, workshop on explainable AI*. Melbourne, août 2017 (cf. p. 37).
- [Bis94] C. BISHOP. "Novelty detection and neural network validation". In : *IEEE Proceedings on Vision, Image and Signal Processing* 141.4 (août 1994), p. 217-222 (cf. p. 37).
- [Bla19] T. BLANCHARD. "Caractérisation de drones en vue de leur localisation et de leur suivi à partir d'une antenne de microphones". Français. Thèse de doct. Université du Mans, déc. 2019, p. 117 (cf. p. 19, 72, 85).
- [Bla+19b] T. BLANCHARD, J. THOMAS et K. RAOOF. "Acoustic signature analysis for localization estimation of unmanned aerial vehicles using few number of microphones". In : *MATEC Web of Conferences* 283 (2019), p. 1-4 (cf. p. 19, 73).
- [Bou07] G. BOULEUX. "Traitement d'antenne adapté aux modèles linéaires intégrant une interférence structurée. Application aux signaux mécaniques." Français. Thèse de doct. Saint-Etienne : Université Jean Monnet, déc. 2007, p. 136 (cf. p. 21).
- [Bra+97] M. BRANDSTEIN et H. SILVERMAN. "A practical methodology for speech source localization with microphone arrays". In : *Computer Speech & Language* 11.2 (avr. 1997), p. 91-126 (cf. p. 18).
- [Bre01] L. BREIMAN. "Random Forests". In : *Machine Learning* 45.1 (oct. 2001), p. 5-32 (cf. p. 33).
- [Bro+06] T. BROOKS et W. HUMPHREYS. "A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays". In : *Journal of Sound and Vibration* 294.4 (juill. 2006), p. 856-879 (cf. p. 21).
- [Bur91] W. BURDIC. *Underwater acoustic system analysis*. Prentice Hall signal processing series. Prentice Hall, 1991 (cf. p. 18, 44).
- [Cab+07] G. CABRAL, A. OLIVEIRA et C. CAHU. "A novel method for one-class classification based on the nearest neighbor data description and structural risk minimization". In : *2007 International Joint Conference on Neural Networks*. Août 2007, p. 1976-1981 (cf. p. 38).
- [Cab+09] G. CABRAL, A. OLIVEIRA et C. CAHÚ. "Combining nearest neighbor data description and structural risk minimization for one-class classification". In : *Neural Computing and Applications* 18.2 (fév. 2009), p. 175-183 (cf. p. 38).
- [Cad88] J. CADZOW. "A high resolution direction-of-arrival algorithm for narrow-band coherent and incoherent sources". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.7 (juill. 1988), p. 965-979 (cf. p. 22).
- [Cap69] J. CAPON. "High-resolution frequency-wavenumber spectrum analysis". In : *Proceedings of the IEEE* 57.8 (août 1969), p. 1408-1418 (cf. p. 21, 44).
- [Car65] E. CARPENTER. "An historical review of seismometer array development". In : *Proceedings of the IEEE* 53.12 (déc. 1965), p. 1816-1821 (cf. p. 18).
- [Car81] G. CARTER. "Time delay estimation for passive sonar signal processing". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3 (juin 1981), p. 463-470 (cf. p. 19, 44).
- [Cha+94] Y. CHAN et K. HO. "A simple and efficient estimator for hyperbolic location". In : *IEEE Transactions on Signal Processing* 42.8 (août 1994), p. 1905-1915 (cf. p. 19).
- [Che53] E. COLIN CHERRY. "Some experiments on the recognition of speech, with one and with two Ears". In : *The Journal of the Acoustical Society of America* 25.5 (sept. 1953), p. 975-979 (cf. p. 18).
- [Cho+03] S. CHO et H. PARK. "Efficient anomaly detection by modeling privilege flows using hidden Markov model". In : *Computers & Security* 22.1 (jan. 2003), p. 45-55 (cf. p. 38).
- [Cor+95] C. CORTES et V. VAPNIK. "Support-vector networks". In : *Machine Learning* 20.3 (sept. 1995), p. 273-297 (cf. p. 34).

- [Cou+19] R. COUSSON, Q. LECLÈRE, M. PALLAS et M. BÉRENGIER. "A time domain CLEAN approach for the identification of acoustic moving sources". In : *Journal of Sound and Vibration* 443 (mars 2019), p. 47-62 (cf. p. 19, 24).
- [Cox58] D. COX. "The Regression Analysis of Binary Sequences". In : *Journal of the Royal Statistical Society : Series B (Methodological)* 20.2 (1958), p. 215-232 (cf. p. 33).
- [Den+05] F. DENIS, R. GILLERON et F. LETOUZEY. "Learning from positive and unlabeled examples". In : *Theoretical Computer Science. Algorithmic Learning Theory (ALT 2000)* 348.1 (déc. 2005), p. 70-83 (cf. p. 38).
- [Dés+12] C. DÉSIR, S. BERNARD, C. PETITJEAN et L. HEUTTE. "A random forest based approach for one class classification in medical imaging". In : *Machine Learning in Medical Imaging. Lecture Notes in Computer Science. Berlin, Heidelberg, 2012*, p. 250-257 (cf. p. 38).
- [Di +07] V. DI GESÙ et G. LO BOSCO. "Combining one Class fuzzy KNN's". In : *International Workshop on Fuzzy Logic and Applications. Lecture Notes in Computer Science. Berlin, Heidelberg, 2007*, p. 152-160 (cf. p. 38).
- [Di +08] V. DI GESÙ, G. LO BOSCO et L. PINELLO. "A One Class Classifier for Signal Identification : A Biological Case Study". In : *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems. Lecture Notes in Computer Science. Berlin, Heidelberg, 2008*, p. 747-754 (cf. p. 38).
- [Dou14a] R. DOUGHERTY. "Functional beamforming". In : *5th Berlin Beamforming Conference. Berlin, fév. 2014*, p. 1-25 (cf. p. 21).
- [Dou14b] R. DOUGHERTY. "Functional beamforming for aeroacoustic source distributions". In : *20th AIAA/CEAS Aeroacoustics Conference. Atlanta, USA, juin 2014*, p. 1-18 (cf. p. 21).
- [Dud+01] R. DUDA, P. HART et D. STORK. *Pattern classification*. 2nd ed. New York : Wiley, 2001 (cf. p. 31).
- [Ehl+97] F. EHLERS et H. SCHUSTER. "Blind separation of convolutive mixtures and an application in automatic speech recognition in a noisy environment". In : *IEEE Transactions on Signal Processing* 45.10 (oct. 1997), p. 2608-2612 (cf. p. 18).
- [Ehr+07] K. EHRENFRIED et L. KOOP. "Comparison of iterative deconvolution algorithms for the mapping of acoustic sources". In : *AIAA Journal* 45.7 (2007), p. 1584-1595 (cf. p. 21).
- [Esm+04] S. ESMAILI, S. KRISHNAN et K. RAAHEMIFAR. "Content based audio classification and retrieval using joint time-frequency analysis". In : *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. T. 5. Mai 2004, p. 665-668 (cf. p. 17).
- [Fer+06] A. FERREOL, P. LARZABAL et M. VIBERG. "On the asymptotic performance analysis of subspace DOA estimation in the presence of modeling errors : case of MUSIC". In : *IEEE Transactions on Signal Processing* 54.3 (mars 2006), p. 907-920 (cf. p. 28).
- [Fis+17] E. FISCHELL et H. SCHMIDT. "Supervised machine learning for estimation of target aspect angle from bistatic acoustic scattering". In : *IEEE Journal of Oceanic Engineering* 42.4 (oct. 2017), p. 759-769 (cf. p. 32).
- [Fis36] R. FISHER. "The Use of Multiple Measurements in Taxonomic Problems". In : *Annals of Eugenics* 7.2 (1936), p. 179-188 (cf. p. 33).
- [Fle+11] V. FLEURY et J. BULTÉ. "Extension of deconvolution algorithms for the mapping of moving acoustic sources". In : *The Journal of the Acoustical Society of America* 129.3 (mars 2011), p. 1417-1428 (cf. p. 17).
- [Gre+66] P. GREEN, E. KELLY et M. LEVIN. "A Comparison of Seismic Array Processing Methods". In : *Geophysical Journal International* 11.1 (sept. 1966), p. 67-84 (cf. p. 18).
- [Hal09] J. HALD. "Basic theory and properties of statistically optimized near-field acoustical holography". In : *The Journal of the Acoustical Society of America* 125.4 (avr. 2009), p. 2105-2120 (cf. p. 22).
- [Has+09] T. HASTIE, R. TIBSHIRANI et J. FRIEDMAN. *The elements of statistical learning : data mining, inference, and prediction*. 2nd ed. Springer series in statistics. New York, NY : Springer, 2009 (cf. p. 31).

- [Hay+93] S. HAYKIN, J. LITVA et T. SHEPHERD. *Radar Array Processing*. Berlin, Heidelberg : Springer Berlin Heidelberg, 1993 (cf. p. 18).
- [Hib+17] C. HIBERT, F. PROVOST, J. MALET, A. MAGGI, A. STUMPF et V. FERRAZZINI. "Automatic identification of rockfalls and volcano-tectonic earthquakes at the Piton de la Fournaise volcano using a Random Forest algorithm". In : *Journal of Volcanology and Geothermal Research* 340 (juin 2017), p. 130-142 (cf. p. 17).
- [Hil56] F. HILDEBRAND. *Introduction to numerical analysis*. New York : McGraw-Hill, 1956 (cf. p. 21).
- [Hol+99] J. HOLLMEN, V. TRESP et O. SIMULA. "A self-organizing map for clustering probabilistic models". In : *9th International Conference on Artificial Neural Networks*. Edinburgh, Scotland, sept. 1999, p. 946-951 (cf. p. 38).
- [ISO16] ISO 17208-1. "Underwater acoustics — Quantities and procedures for description and measurement of underwater sound from ships — Part 1 : Requirements for precision measurements in deep water used for comparison purposes". In : *International Organization for Standardization* (mars 2016) (cf. p. 17, 45).
- [ISO19] ISO 17208-2. "Underwater acoustics — Quantities and procedures for description and measurement of underwater sound from ships — Part 2 : Determination of source levels from deep water measurements". In : *International Organization for Standardization* (juill. 2019), p. 1-13 (cf. p. 17, 45).
- [ISO15] ISO 362-1. "Measurement of noise emitted by accelerating road vehicles — Engineering method — Part 1 : M and N categories". In : *International Organization for Standardization* (sept. 2015), p. 1-77 (cf. p. 17).
- [ISO01] ISO 3740. "Acoustics — Determination of sound power levels of noise sources — Guidelines for the use of basic standards". In : *International Organization for Standardization* (juill. 2001), p. 1-35 (cf. p. 17).
- [ISO93] ISO 9614-1. "Acoustics — Determination of sound power levels of noise sources using sound intensity — Part 1 : Measurement at discrete points". In : *International Organization for Standardization* (juin 1993), p. 1-19 (cf. p. 17).
- [ISO96] ISO 9614-2. "Acoustics — Determination of sound power levels of noise sources using sound intensity — Part 2 : Measurement by scanning". In : *International Organization for Standardization* (août 1996), p. 1-21 (cf. p. 17).
- [Jap99] N JAPKOWICZ. "Concept-Learning in the Absence of Counter-Examples :An Autoassociation-Based Approach to Classification". Anglais. Thèse de doct. Oct. 1999 (cf. p. 37).
- [Jes+00] S. JESUS, M. PORTER, Y. STEPHAN, X. DEMOULIN, O. RODRIGUEZ et E. COELHO. "Single hydrophone source localization". In : *IEEE Journal of Oceanic Engineering* 25.3 (juill. 2000), p. 337-346 (cf. p. 17).
- [Ken+98] R. KENNEDY, T. ABHAYAPALA et D. WARD. "Broadband nearfield beamforming using a radial beampattern transformation". In : *IEEE Transactions on Signal Processing* 46.8 (août 1998), p. 2147-2156 (cf. p. 30).
- [Kha+14] S. KHAN et M. MADDEN. "One-class classification : taxonomy of study and review of techniques". In : *The Knowledge Engineering Review* 29.3 (juin 2014), p. 345-374 (cf. p. 38).
- [Kra+05] J. KRAUS, M. TIURI, A. RÄISÄNEN et T. CARR. *Radio astronomy*. Durham : Cygnus-Quasar, 2005 (cf. p. 18).
- [Kri+92] H. KRIM, P. FORSTER et J. PROAKIS. "Operator approach to performance analysis of root-MUSIC and root-min-norm". In : *IEEE Transactions on Signal Processing* 40.7 (juill. 1992), p. 1687-1696 (cf. p. 23).
- [Kri+96] H. KRIM et M. VIBERG. "Two decades of array signal processing research : the parametric approach". In : *IEEE Signal Processing Magazine* 13.4 (juill. 1996), p. 67-94 (cf. p. 18).
- [Kum+83] R. KUMARESAN et D. TUFTS. "Estimating the Angles of Arrival of Multiple Plane Waves". In : *IEEE Transactions on Aerospace and Electronic Systems* AES-19.1 (jan. 1983), p. 134-139 (cf. p. 22).
- [Lan+03] T. LANE et C. BRODLEY. "An empirical study of two approaches to sequence learning for anomaly detection". In : *Machine Learning* 51.1 (avr. 2003), p. 73-107 (cf. p. 38).

- [Las77] M. LASKY. "Review of undersea acoustics to 1950". In : *The Journal of the Acoustical Society of America* 61.2 (fév. 1977), p. 283-297 (cf. p. 18, 44).
- [Le 12] F. LE COURTOIS. "Caractérisation des sources acoustiques sur le matériel ferroviaire par méthode d'antennerie". Français. Thèse de doct. Université du Maine, sept. 2012, p. 114 (cf. p. 18, 46).
- [Le 16] T. LE MAGUERESSE. "Approche multidimensionnelle du problème d'identification acoustique inverse". Français. Thèse de doct. INSA Lyon, fév. 2016, p. 170 (cf. p. 17).
- [Le +20] T. LE MAGUERESSE, A. OUTREQUIN, M. THIVANT, J. ANTONI et E. ROBERT. "3D acoustical characterization of an electrical motor by bayesian focusing". In : *8th Berlin Beamforming Conference*. Berlin, mars 2020, p. 1-14 (cf. p. 18).
- [Lec09] Q. LECLÈRE. "Acoustic imaging using under-determined inverse approaches : Frequency limitations and optimal regularization". In : *Journal of Sound and Vibration* 321.3 (avr. 2009), p. 605-619 (cf. p. 22).
- [Lec+17] Q. LECLÈRE, A. PEREIRA, C. BAILLY, J. ANTONI et C. PICARD. "A unified formalism for acoustic imaging based on microphone array measurements". In : *International Journal of Aeroacoustics* 16.4-5 (juill. 2017), p. 431-456 (cf. p. 20, 22).
- [Li+91] F. LI et R. VACCARO. "Unified analysis for DOA estimation algorithms in array signal processing". In : *Signal Processing* 25.2 (nov. 1991), p. 147-169 (cf. p. 23).
- [Li+03] J. LI, P. STOICA et Z. WANG. "On robust Capon beamforming and diagonal loading". In : *IEEE Transactions on Signal Processing* 51.7 (juill. 2003), p. 1702-1715 (cf. p. 21).
- [Mai05] R. MAILLOUX. *Phased array antenna handbook*. 2. ed. Boston : Artech House, 2005 (cf. p. 30).
- [Mal16] S. MALLAT. "Understanding deep convolutional networks". In : *Philosophical Transactions of the Royal Society A : Mathematical, Physical and Engineering Sciences* 374.2065 (avr. 2016), p. 1-16 (cf. p. 37).
- [Mar+17] E. MARCHI, F. VESPERINI, S. SQUARTINI et B. SCHULLER. "Deep Recurrent Neural Network-Based Autoencoders for Acoustic Novelty Detection". In : *Computational Intelligence and Neuroscience* (2017) (cf. p. 17).
- [Mar98] S. MARCOS. *Les méthodes à haute résolution : traitement d'antenne et analyse spectrale*. Français. Paris : Hermes, 1998 (cf. p. 18, 20-22, 28).
- [Maz06] O. MAZHELIS. "One-class classifiers : a review and analysis of suitability in the context of mobile-masquerader detection". In : *South African Computer Journal* 2006.36 (juin 2006), p. 29-48 (cf. p. 38).
- [Mer+19] R. MERINO-MARTÍNEZ, P. SIJTSMA, M. SNELLEN, T. AHLEFELDT, J. ANTONI, C. BAHR, D. BLACODON, D. ERNST, A. FINEZ, S. FUNKE, T. GEYER, S. HAXTER, G. HEROLD, X. HUANG, W. HUMPHREYS, Q. LECLÈRE, A. MALGOEZAR, U. MICHEL, T. PADOIS, A. PEREIRA, C. PICARD, E. SARRADJ, H. SILLER, D. SIMONS et C. SPEHR. "A review of acoustic imaging methods using phased microphone arrays". In : *CEAS Aeronautical Journal* 10.1 (mars 2019), p. 197-230 (cf. p. 18).
- [Mer+16] R. MERINO-MARTÍNEZ, M. SNELLEN et D. SIMONS. "Functional beamforming applied to imaging of flyover noise on landing aircraft". In : *Journal of Aircraft* 53.6 (juin 2016), p. 1830-1843 (cf. p. 18).
- [Mer76] H. MERMOZ. "Imagerie, corrélation et modèles". fr. In : *Annales Des Télécommunications* 31.1 (jan. 1976), p. 17-36 (cf. p. 21, 27).
- [Mic+95] Z. MICHALOPOULOU, D. ALEXANDROU et C. de MOUSTIER. "Application of neural and statistical classifiers to the problem of seafloor characterization". In : *IEEE Journal of Oceanic Engineering* 20.3 (juill. 1995), p. 190-197 (cf. p. 32).
- [Mil19] T. MILLER. "Explanation in artificial intelligence : Insights from the social sciences". In : *Artificial Intelligence* 267 (fév. 2019), p. 1-38 (cf. p. 37).
- [Neh+94] A. NEHORAI et E. PALDI. "Vector-sensor array processing for electromagnetic source localization". In : *IEEE Transactions on Signal Processing* 42.2 (fév. 1994), p. 376-398 (cf. p. 18).

- [Nel+00] P. NELSON et S. YOON. "Estimation of acoustic source strength by inverse methods : part I, conditioning of the inverse problem". In : *Journal of Sound and Vibration* 233.4 (juin 2000), p. 639-664 (cf. p. 22).
- [Oer+07] S. OERLEMANS, L. BROERSMA et P. SIJTSMA. "Quantification of airframe noise using microphone arrays in open and closed wind tunnels". In : *International Journal of Aeroacoustics* 6.4 (déc. 2007), p. 309-333 (cf. p. 17).
- [Ott+91] B. OTTERSTEN, M. VIBERG et T. KAILATH. "Performance analysis of the total least squares ESPRIT algorithm". In : *IEEE Transactions on Signal Processing* 39.5 (mai 1991), p. 1122-1135 (cf. p. 23).
- [Oud15] B. OUDOMPHENG. "Localisation et contribution de sources acoustiques de navire au passage par traitement d'antenne réduite". Français. Thèse de doct. Université Grenoble Alpes, nov. 2015, p. 209 (cf. p. 18).
- [Oud+18] B. OUDOMPHENG, B. NICOLAS et L. LAMOTTE. "Localization and contribution of underwater acoustical sources of a moving surface ship". In : *IEEE Journal of Oceanic Engineering* 43.2 (avr. 2018), p. 536-546 (cf. p. 18, 45).
- [Pis72] V. PISARENKO. "On the estimation of spectra by means of non-linear functions of the covariance matrix". In : *Geophysical Journal of the Royal Astronomical Society* 28.5 (juin 1972), p. 511-531 (cf. p. 21).
- [Pol19] M. POLICHETTI. "Traitement d'antenne adaptatif pour l'imagerie ultrasonore passive de la cavitation". Français. Thèse de doct. Lyon : Université de Lyon, oct. 2019, p. 163 (cf. p. 18).
- [Pot+94] J. POTTER, D. MELLINGER et C. CLARK. "Marine mammal call discrimination using artificial neural networks". In : *The Journal of the Acoustical Society of America* 96.3 (sept. 1994), p. 1255-1262 (cf. p. 32).
- [Pro95] Baron de PRONY. "Essai expérimental et analytique : sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur d'eau et de la vapeur de l'alcool à différentes températures". In : *Journal de l'école polytechnique Paris* 1 :2 (1795), p. 24-76 (cf. p. 21).
- [Qui+06] A. QUINLAN, JP. BARBOT et P. LARZABAL. "Automatic determination of the number of targets present when using the time reversal operator". In : *The Journal of the Acoustical Society of America* 119.4 (mars 2006), p. 2220-2225 (cf. p. 28).
- [Qui86] J. QUINLAN. "Induction of decision trees". In : *Machine Learning* 1.1 (mars 1986), p. 81-106. (Visité le 21/08/2020) (cf. p. 33).
- [Ram19] A. RAMAMONJY. "Développement de nouvelles méthodes de classification/localisation de signaux acoustiques appliquées aux véhicules aériens". Français. Thèse de doct. Paris, CNAM, mai 2019, p. 238 (cf. p. 18, 73, 85).
- [Ram+18] A. RAMAMONJY, E. BAVU, A. GARCIA et S. HENGY. "Source localization and identification with a compact array of digital MEMS microphones". In : *ICSV25*. Hiroshima, juill. 2018, p. 1-8 (cf. p. 18, 73).
- [Rao+89] B. RAO et K. HARI. "Performance analysis of Root-Music". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.12 (déc. 1989), p. 1939-1949 (cf. p. 22).
- [Red79] S. REDDI. "Multiple source location-A digital approach". In : *IEEE Transactions on Aerospace and Electronic Systems* AES-15.1 (jan. 1979), p. 95-105 (cf. p. 18, 44).
- [Rit+97] G. RITTER et M. GALLEGOS. "Outliers in statistical pattern recognition and an application to automatic chromosome classification". In : *Pattern Recognition Letters* 18.6 (juin 1997), p. 525-539 (cf. p. 37).
- [Ros57] F. ROSENBLATT. *The perceptron, a perceiving and recognizing automaton*. 1957 (cf. p. 33).
- [Roy+89] R. ROY et T. KAILATH. "ESPRIT-estimation of signal parameters via rotational invariance techniques". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 37.7 (juill. 1989), p. 984-995 (cf. p. 22).
- [Sar12] E. SARRADJ. "Three-Dimensional Acoustic Source Mapping with Different Beamforming Steering Vector Formulations". In : *Advances in Acoustics and Vibration* 2012 (2012), p. 1-12 (cf. p. 25).

- [Sar16] E. SARRADJ. "A generic approach to synthesize optimal microphone arrangements". In : *6th Berlin Beamforming Conference*. Mars 2016, p. 1-12 (cf. p. 24, 49).
- [Sch86] R. SCHMIDT. "Multiple emitter location and signal parameter estimation". In : *IEEE Transactions on Antennas and Propagation* AP-34.3 (mars 1986), p. 276-280 (cf. p. 22, 28, 44).
- [Sch+01] B. SCHÖLKOPF, J. PLATT, J. SHAWE-TAYLOR, A. SMOLA et R. WILLIAMSON. "Estimating the Support of a High-Dimensional Distribution". In : *Neural Computation* 13.7 (juill. 2001), p. 1443-1471 (cf. p. 38).
- [Sch78] G. SCHWARZ. "Estimating the Dimension of a Model". In : *Annals of Statistics* 6.2 (mars 1978), p. 461-464 (cf. p. 28).
- [Sha+85] K. SHARMAN et T. DURRANI. "Resolving power of signal subspace methods for finite data lengths". In : *ICASSP '85. IEEE International Conference on Acoustics, Speech, and Signal Processing*. T. 10. Avr. 1985, p. 1501-1504 (cf. p. 23).
- [Sha+84] K. SHARMAN, T. DURRANI, M. WAX et T. KAILATH. "Asymptotic performance of eigenstructure spectral analysis methods". In : *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*. T. 9. Mars 1984, p. 440-443 (cf. p. 23).
- [Shy+03] L. SHYU, C. CHEN, K. SARINNAKORN et L. CHANG. "A novel anomaly detection scheme based on principal component classifier". In : *Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop*. 2003, p. 172-179 (cf. p. 38).
- [Sij07] P. SIJTSMA. "CLEAN based on spatial source coherence". In : *International Journal of Aeroacoustics* 6.4 (déc. 2007), p. 357-374 (cf. p. 21).
- [Sij+01] P. SIJTSMA, S. OERLEMANS et H. HOLTHUSEN. "Location of rotating sources by phased array measurements". In : *7th AIAA/CEAS Aeroacoustics Conference*. Maastricht, mai 2001 (cf. p. 18).
- [Sij+04] P. SIJTSMA et R. STOKER. "Determination of absolute contributions of aircraft noise components using fly-over array measurements". In : *10th AIAA/CEAS Aeroacoustics Conference*. Manchester, Great Britain, mai 2004 (cf. p. 17).
- [Ska03] A. SKABAR. "Single-class classifier learning using neural networks : an application to the prediction of mineral deposits". In : *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics*. T. 4. Nov. 2003, p. 2127-2132 (cf. p. 38).
- [Sto+05] P. STOICA et R. MOSES. *Spectral analysis of signals*. Upper Saddle River, N.J : Pearson/Prentice Hall, 2005 (cf. p. 20, 21).
- [Sto+02] P. STOICA, Z. WANG et J. LI. "Robust Capon beamforming". In : *Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*. T. 1. Nov. 2002, 876-880 vol.1 (cf. p. 21).
- [Sun+20] D. SUN, C. MA, J. MEI et W. SHI. "Improving the resolution of underwater acoustic image measurement by deconvolution". In : *Applied Acoustics* 165 (août 2020), p. 1-11 (cf. p. 19).
- [Tax01] D. TAX. "One-class classification : concept-learning in the absence of counter-examples". Anglais. Thèse de doct. Technische Universiteit Delft, juin 2001, p. 190 (cf. p. 37, 38).
- [Tax+01] D. TAX et R. DUIN. "Uniform object generation for optimizing one-class classifiers". In : *Journal of Machine Learning Research* 2 (déc. 2001), p. 155-173 (cf. p. 38).
- [Tho04] A. THODE. "Tracking sperm whale (*Physeter macrocephalus*) dive profiles using a towed passive acoustic array". In : *The Journal of the Acoustical Society of America* 116.1 (juill. 2004), p. 245-253 (cf. p. 18).
- [Thu+86] D. THUBER et L. KOPP. "Measurement accuracy and resolving power of high resolution passive methods". In : *Proceedings of EUSIPCO*. T. 2. 1986, p. 1037-1040 (cf. p. 23).
- [Thu+85] D. THUBERT, L. KOPP et J. LECADRE. "Précision des méthodes hautes résolution et bornes de cramer rao". Français. In : *Xème colloque GRETSI*. Nice, France, mai 1985 (cf. p. 23).
- [Tie+06] C. TIEMANN, A. THODE, J. STRALEY, V. O'CONNELL et K. FOLKERT. "Three-dimensional localization of sperm whales using a single hydrophone". In : *The Journal of the Acoustical Society of America* 120.4 (oct. 2006), p. 2355-2365 (cf. p. 17).

- [Van+88] B. VAN VEEN et K. BUCKLEY. "Beamforming : a versatile approach to spatial filtering". In : *IEEE ASSP Magazine* 5.2 (avr. 1988), p. 4-24 (cf. p. 20, 21).
- [Vib+91] M. VIBERG et B. OTTERSTEN. "Sensor array processing based on subspace fitting". In : *IEEE Transactions on Signal Processing* 39.5 (mai 1991), p. 1110-1121 (cf. p. 22).
- [Vor+03] S. VOROBYOV, A. GERSHMAN et Z. LUO. "Robust adaptive beamforming using worst-case performance optimization : a solution to the signal mismatch problem". In : *IEEE Transactions on Signal Processing* 51.2 (fév. 2003), p. 313-324 (cf. p. 21).
- [War+17] G. WARNER, S. DOSSO et D. HANNAY. "Bowhead whale localization using time-difference-of-arrival data from asynchronous recorders". In : *The Journal of the Acoustical Society of America* 141.3 (mars 2017), p. 1921-1935 (cf. p. 18).
- [Wax+85] M. WAX et T. KAILATH. "Detection of signals by information theoretic criteria". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 33.2 (avr. 1985), p. 387-392 (cf. p. 28).
- [Whi07] T. WHITTINGHAM. "Medical diagnostic applications and sources". In : *Progress in Biophysics and Molecular Biology. Effects of ultrasound and infrasound relevant to human health* 93.1 (jan. 2007), p. 84-110 (cf. p. 18).
- [Wil+80] E. WILLIAMS, J. MAYNARD et E. SKUDRZYK. "Sound source reconstructions using a microphone array". In : *The Journal of the Acoustical Society of America* 68.1 (juill. 1980), p. 340-344 (cf. p. 22).
- [Win+09] J. WIND, E. TIJS et H. de BREE. "Source localization using acoustic vector sensors : a music approach". In : *Noise and Vibration : Emerging Methods* (avr. 2009), p. 1-10 (cf. p. 18).
- [Wol+33] I. WOLFF et F. MASSA. "Use of pressure gradient microphones for acoustical measurements". In : *The Journal of the Acoustical Society of America* 4.3 (jan. 1933), p. 217-234 (cf. p. 18).
- [Won+00] K. WONG et M. ZOLTOWSKI. "Self-initiating MUSIC-based direction finding in underwater acoustic particle velocity-field beamspace". In : *IEEE Journal of Oceanic Engineering* 25.2 (avr. 2000), p. 262-273 (cf. p. 18).
- [Wu98] S. WU. "On reconstruction of acoustic pressure fields using the HELS method". In : *The Journal of the Acoustical Society of America* 104.3 (sept. 1998), p. 1775-1775 (cf. p. 22).
- [Yam+04] K. YAMANISHI, J. TAKEUCHI, G. WILLIAMS et P. MILNE. "On-line unsupervised outlier detection using finite mixtures with discounting learning algorithms". In : *Data Mining and Knowledge Discovery* 8.3 (mai 2004), p. 275-300 (cf. p. 38).
- [Yan18] T. YANG. "Deconvolved Conventional Beamforming for a Horizontal Line Array". In : *IEEE Journal of Oceanic Engineering* 43.1 (jan. 2018), p. 160-172 (cf. p. 19).
- [Ye+01] N. YE et Q. CHEN. "An anomaly detection technique based on a chi-square statistic for detecting intrusions into information systems". In : *Quality and Reliability Engineering International* 17.2 (2001), p. 105-112 (cf. p. 38).
- [Yeu+02] D. YEUNG et C. CHOW. "Parzen-window network intrusion detectors". In : *Object recognition supported by user interaction for service robots. T. 4. Août 2002*, p. 385-388 (cf. p. 38).
- [Yeu+03] D. YEUNG et Y. DING. "Host-based intrusion detection using dynamic and static behavioral models". In : *Pattern Recognition* 36.1 (jan. 2003), p. 229-243 (cf. p. 38).
- [Zan+04] S. ZANERO et S. SAVARESI. "Unsupervised learning techniques for an intrusion detection system". In : *Proceedings of the 2004 ACM symposium on Applied computing. New York, NY, USA, mars 2004*, p. 412-419 (cf. p. 38).

Bibliographie du chapitre 3

- [Amo+16] D. AMON, A. ZIEGLER, T. DAHLGREN, A. GLOVER, A. GOINEAU, A. GOODAY, H. WIKLUND et C. SMITH. "Insights into the abundance and diversity of abyssal megafauna in a polymetallic-nodule region in the eastern Clarion-Clipperton Zone". In : *Scientific Reports* 6.30492 (juill. 2016), p. 1-12 (cf. p. 42).

- [ANS19] ANSI/ASA S12.64 PART 1. "American National Standard Quantities and Procedures for Description and Measurement of Underwater Sound from Ships – Part 1 : General Requirements". In : *American National Standards Institute* (sept. 2019) (cf. p. 17, 45).
- [Ant12] J ANTONI. "A Bayesian approach to sound source reconstruction : Optimal basis, regularization, and focusing". In : *The Journal of the Acoustical Society of America* 131.4 (avr. 2012), p. 2873-2890 (cf. p. 17, 22).
- [Bag+93] A. BAGGEROER, W. KUPERMAN et P. MIKHALEVSKY. "An overview of matched field methods in ocean acoustics". In : *IEEE Journal of Oceanic Engineering* 18.4 (oct. 1993), p. 401-424 (cf. p. 45).
- [Bag+88] A. BAGGEROER, W. KUPERMAN et H. SCHMIDT. "Matched field processing : Source localization in correlated noise as an optimum parameter estimation problem". In : *The Journal of the Acoustical Society of America* 83.2 (fév. 1988), p. 571-587 (cf. p. 45).
- [Bai+10] M. BAI, J. LIN et K. LIU. "Optimized microphone deployment for near-field acoustic holography : To be, or not to be random, that is the question". In : *Journal of Sound and Vibration* 329.14 (juill. 2010), p. 2809-2824 (cf. p. 46).
- [Bar+20] V. BARON, A. FINEZ, S. BOULEY, F. FAYET, J. MARS et B. NICOLAS. "Hydrophone Array Optimization, Conception, and Validation for Localization of Acoustic Sources in Deep-Sea Mining". In : *IEEE Journal of Oceanic Engineering* (2020). Early access, p. 1-9 (cf. p. 48, 119).
- [Bas+19] R. BASSILA, T. BERTET, L. SOMAGLINO, ME. BOUHIER, H. GLOTIN, P. BEST, V. BARON, C. NOËL, P. CRISTINI, F. FAYET, B. NICOLAS et J. MARS. "ABYSOUND, an end to end system for noise impact measurement of deep sea mining production tools". In : *MTS/IEEE Oceans 2019*. Marseille, juin 2019, p. 1-9 (cf. p. 42, 119).
- [Ber+17a] H. BERKOWITZ et H. DUMEZ. "Racket in the oceans : why underwater noise matters, how to measure it, and how to manage it". In : *Paris : Observatory for Responsible Innovation / Palaiseau (France) : i3-CRG (CNRS – École polytechnique)* (2017), p. 1-123 (cf. p. 45).
- [Bie+85] G. BIENVENU et L. KOPP. "New Approaches in the Adaptive Array Processing Field". In : *Adaptive Methods in Underwater Acoustics*. Sous la dir. d'Heinz G. URBAN. NATO ASI Series. Dordrecht : Springer Netherlands, 1985, p. 375-384 (cf. p. 18, 44).
- [Bla+04] S. BLACKWELL, C. GREENE et W. RICHARDSON. "Drilling and operational sounds from an oil production island in the ice-covered Beaufort Sea". In : *The Journal of the Acoustical Society of America* 116.5 (nov. 2004), p. 3199-3211 (cf. p. 43-45).
- [Bro+14] M. BROUDIC, I. VOELLMY, P. DOBBINS, S. ROBINSON, P. BERGGREN, S. LAING, L. BLAKE, A. RADFORD, P. LEPPER, F. PACE, S. NEVES, S. SIMPSON et R. BRUINJES. "Underwater noise emission from the NOAH's drilling operation at the NaREC site, Blyth, UK". In : *2nd International Conference on Environmental Interactions of Marine Renewable Energy Technologies*. Scotland, mai 2014 (cf. p. 43-45).
- [Bur91] W. BURDIC. *Underwater acoustic system analysis*. Prentice Hall signal processing series. Prentice Hall, 1991 (cf. p. 18, 44).
- [Cap69] J. CAPON. "High-resolution frequency-wavenumber spectrum analysis". In : *Proceedings of the IEEE* 57.8 (août 1969), p. 1408-1418 (cf. p. 21, 44).
- [Car81] G. CARTER. "Time delay estimation for passive sonar signal processing". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3 (juin 1981), p. 463-470 (cf. p. 19, 44).
- [Chi+20] A. CHIN et K. HARI. *Predicting the impacts of mining of deep sea polymetallic nodules in the Pacific Ocean : A review of Scientific literature*. 52 pages. Mai 2020 (cf. p. 42).
- [Cof08] COFFEY NATURAL SYSTEMS PTY LTD. *Environmental impact statement : Nautilus Minerals Niugini Limited Solwara 1 project*. Volume B Appendix 13. Sept. 2008 (cf. p. 42).
- [Cri+19] P. CRISTINI, R. SABATINI, C. NOËL, S. MARCHETTI et JM. TEMMOS. "Design and implementation of an acoustic simulator of noise level distribution in the ocean". In : *MTS/IEEE Oceans 2019*. Marseille, juin 2019, p. 1-7 (cf. p. 42, 119).

- [Dee+19] DEEP SEA MINING CAMPAIGN, PNG COUNCIL OF CHURCHES, VOICE OF MILNE BAY, ALLIANCE OF SOLWARA WARRIORS, BISMARCK RAMU GROUP et CENTRE FOR ENVIRONMENTAL LAW AND COMMUNITY RIGHTS. *Joint Letter calling for the Papua New Guinea Government to cancel all Nautilus Minerals deep sea mining licences and to ban seabed mining in PNG*. Juin 2019. URL : <http://www.deepseaminingoutofourdepth.org/joint-letter-calling-for-the-papua-new-guinea-government-to-cancel-all-nautilus-minerals-deep-sea-mining-licences-and-to-ban-seabed-mining-in-png/> (visité le 06/08/2020) (cf. p. 42).
- [Dos+13] S. DOSSO et M. WILMUT. "Bayesian tracking of multiple acoustic sources in an uncertain ocean environment". In : *The Journal of the Acoustical Society of America* 133.4 (mars 2013), EL274-EL280 (cf. p. 45).
- [Eli97] G. ELIAS. *Aeroacoustic and active noise control - Experimental techniques for source location*. Sept. 1997 (cf. p. 47, 52).
- [Fat+19] D. FATTACCIOLI et G. REAL. "A brief five-year progress report on the DGA "ALMA" project : scientific program and projects, at-sea campaigns and data collected, system evolutions and improvements". In : *Underwater Acoustics Conference and Exhibition*. Crete, Greece, juill. 2019, p. 1-10 (cf. p. 47).
- [Fer89] B. FERGUSON. "Improved time-delay estimates of underwater acoustic signals using beamforming and prefiltering techniques". In : *IEEE Journal of Oceanic Engineering* 14.3 (juill. 1989), p. 238-244 (cf. p. 44).
- [Gao+94] SW. GAO et J. GRIFFITHS. "Experimental performance of high-resolution array processing algorithms in a towed sonar array environment". In : *The Journal of the Acoustical Society of America* 95.4 (avr. 1994), p. 2068-2080 (cf. p. 45, 46).
- [Gas+17] M. GASSMANN, S. WIGGINS et J. HILDEBRAND. "Deep-water measurements of container ship radiated noise signatures and directionality". In : *The Journal of the Acoustical Society of America* 142.3 (sept. 2017), p. 1563-1574 (cf. p. 45).
- [Ger+95] A. GERSHMAN, V. TURCHIN et V. ZVEREV. "Experimental results of localization of moving underwater signal by adaptive beamforming". In : *IEEE Transactions on Signal Processing* 43.10 (oct. 1995), p. 2249-2257 (cf. p. 45).
- [Gla00] G. GLASBY. "Lessons learned from deep-sea mining". In : *Science* 289.5479 (juill. 2000), p. 551-553 (cf. p. 42).
- [Gre87] C. GREENE. "Characteristics of oil industry dredge and drilling sounds in the Beaufort Sea". In : *The Journal of the Acoustical Society of America* 82.4 (oct. 1987), p. 1315-1324 (cf. p. 43-45).
- [Gri69] L. GRIFFITHS. "A simple adaptive algorithm for real-time processing in antenna arrays". In : *Proceedings of the IEEE* 57.10 (oct. 1969), p. 1696-1704 (cf. p. 44).
- [Hal+07] J. HALFAR et R. FUJITA. "Danger of deep-sea mining". In : *Science* 316.5827 (mai 2007), p. 987-987 (cf. p. 42).
- [Ham+89] R. HAMSON et R. HEITMEYER. "Environmental and system effects on source localization in shallow water by the matched-field processing of a vertical array". In : *The Journal of the Acoustical Society of America* 86.5 (nov. 1989), p. 1950-1959 (cf. p. 45).
- [Hau94] HAUPT, R. "Thinned arrays using genetic algorithms". In : *IEEE Transactions on Antennas and Propagation* 42.7 (juill. 1994), p. 993-999 (cf. p. 46).
- [Hol+01] S. HOLM, A. AUSTENG, K. IRANPOUR et JF. HOPPERSTAD. "Sparse Sampling in Array Processing". In : *Nonuniform Sampling : Theory and Practice*. Sous la dir. de FAROKH MARVASTI. Information Technology : Transmission, Processing, and Storage. Boston, MA : Springer US, 2001, p. 787-833 (cf. p. 46, 47).
- [ISO16] ISO 17208-1. "Underwater acoustics — Quantities and procedures for description and measurement of underwater sound from ships — Part 1 : Requirements for precision measurements in deep water used for comparison purposes". In : *International Organization for Standardization* (mars 2016) (cf. p. 17, 45).

- [ISO19] ISO 17208-2. "Underwater acoustics — Quantities and procedures for description and measurement of underwater sound from ships — Part 2 : Determination of source levels from deep water measurements". In : *International Organization for Standardization* (juill. 2019), p. 1-13 (cf. p. 17, 45).
- [Jau+19] J. JAUMONT, V. BARON, B. NICOLAS, F. FAYET, A. FINEZ et J. MARS. "Conception of an underwater prototype array : from simulation design to mechanical assessment". In : *MTS/IEEE Oceans 2019*. Marseille, juin 2019, p. 1-9 (cf. p. 52, 119).
- [Kai+17] S. KAISER, C. SMITH et P. ARBIZU. "Editorial : Biodiversity of the Clarion Clipperton Fracture Zone". In : *Marine Biodiversity* 47.2 (juin 2017), p. 259-264 (cf. p. 42).
- [Kir+81] R. KIRLIN, D. MOORE et R. KUBICHEK. "Improvement of delay measurements from sonar arrays via sequential state estimation". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 29.3 (juin 1981), p. 514-519 (cf. p. 44).
- [Kna+76] C. KNAPP et G. CARTER. "The generalized correlation method for estimation of time delay". In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 24.4 (août 1976), p. 320-327 (cf. p. 44).
- [Las77] M. LASKY. "Review of undersea acoustics to 1950". In : *The Journal of the Acoustical Society of America* 61.2 (fév. 1977), p. 283-297 (cf. p. 18, 44).
- [Le 12] F. LE COURTOIS. "Caractérisation des sources acoustiques sur le matériel ferroviaire par méthode d'antennerie". Français. Thèse de doct. Université du Maine, sept. 2012, p. 114 (cf. p. 18, 46).
- [Luc+85] B. LUCAS et JP. LE CADRE. "Experimentation of spatial processing methods". In : *Adaptive Methods in Underwater Acoustics*. Sous la dir. de H. URBAN. NATO ASI Series. Dordrecht : Springer Netherlands, 1985, p. 317-324 (cf. p. 45).
- [Mer+08] L. MERAD, F. BENDIMERAD et S. MERIAH. "Design of linear antenna arrays for side lobe reduction using the Tabu search method". In : *The Internatioanl Arab Journal of Information Technology* 5.3 (juill. 2008), p. 219-222 (cf. p. 46).
- [Min] MINING TECHNOLOGY. *Solwara Project*. URL : <https://www.mining-technology.com/projects/solwara-project/> (visité le 24/07/2020) (cf. p. 42).
- [New17] Kyodo NEWS. "Japan successfully undertakes large-scale deep-sea mineral extraction". In : *The Japan Times Online* (sept. 2017). (Visité le 20/02/2020) (cf. p. 42).
- [Oud+18] B. OUDOMPHENG, B. NICOLAS et L. LAMOTTE. "Localization and contribution of underwater acoustical sources of a moving surface ship". In : *IEEE Journal of Oceanic Engineering* 43.2 (avr. 2018), p. 536-546 (cf. p. 18, 45).
- [Pal+16] V. PALLAYIL, M. CHITRE, S. KUSELAN, A. RAICHUR, M. IGNATIUS et J. POTTER. "Development of a Second-Generation Underwater Acoustic Ambient Noise Imaging Camera". In : *IEEE Journal of Oceanic Engineering* 41.1 (jan. 2016), p. 175-189 (cf. p. 47).
- [Par+08] PARLEMENT EUROPÉEN et CONSEIL DE L'UNION EUROPÉENNE. *Directive-cadre Stratégie pour le milieu marin*. Directive 2008/56/CE. Juin 2008 (cf. p. 45).
- [Pha+19a] GT. PHAM, V. BARON, A. FINEZ, J. MARS et B. NICOLAS. "High resolution source localization in underwater acoustics for deep sea mining monitoring". In : *MTS/IEEE Oceans 2019*. Marseille, juin 2019, p. 1-7 (cf. p. 60, 119).
- [Pha+19b] GT. PHAM, J. MARS, V. BARON, A. FINEZ et B. NICOLAS. "Localisation de sources de bruit en prospection minière sous-marine : une preuve de concept". Français. In : *Gretsi 2019*. Lille, août 2019 (cf. p. 60, 119).
- [Pis73] V. PISARENKO. "The retrieval of harmonics from a covariance function". In : *Geophysical Journal International* 33.3 (sept. 1973), p. 347-366 (cf. p. 44).
- [Pit18] G. PITRON. *La guerre des métaux rares*. Les liens qui libèrent. Jan. 2018 (cf. p. 42).
- [Rea+18] G. REAL et D. FATTACCIOLI. "Acoustic Laboratory for Marine Applications : Overview of the ALMA system and data analysis." In : *MTS/IEEE Oceans 2018*. Charleston, oct. 2018, p. 1-7 (cf. p. 47).

- [Red79] S. REDDI. "Multiple source location-A digital approach". In : *IEEE Transactions on Aerospace and Electronic Systems* AES-15.1 (jan. 1979), p. 95-105 (cf. p. 18, 44).
- [Rei+12] K. REINE, D. CLARKE et C. DICKERSON. *Characterization of underwater sounds produced by a backhoe dredge excavating rock and gravel*. Report ERDC TN-DOER-E36. Déc. 2012 (cf. p. 43).
- [Rei+14] K. REINE, D. CLARKE et C. DICKERSON. "Characterization of underwater sounds produced by hydraulic and mechanical dredging operations". In : *The Journal of the Acoustical Society of America* 135.6 (juin 2014), p. 3280-3294 (cf. p. 43-45).
- [Ric09] G. RICCOBENE. "Long-term measurements of acoustic background noise in very deep sea". In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment* 604.1, Supplement (juin 2009), S149-S157 (cf. p. 47).
- [Ric+85] K. RICHTER, J. BENNETT et K. SMITH. "Bottom-moored acoustic array to monitor density and vertical movement of deep-sea benthopelagic animals". In : *IEEE Journal of Oceanic Engineering* 10.1 (jan. 1985), p. 32-37 (cf. p. 47).
- [Sar16] E. SARRADJ. "A generic approach to synthesize optimal microphone arrangements". In : *6th Berlin Beamforming Conference*. Mars 2016, p. 1-12 (cf. p. 24, 49).
- [Sch86] R. SCHMIDT. "Multiple emitter location and signal parameter estimation". In : *IEEE Transactions on Antennas and Propagation* AP-34.3 (mars 1986), p. 276-280 (cf. p. 22, 28, 44).
- [Sch+85] B. SCHOLZ et W. KROLL. "Application of high resolution spatial processing methods to real data of a fixed array in shallow water". In : *Adaptive Methods in Underwater Acoustics*. Sous la dir. de H. URBAN. NATO ASI Series. Dordrecht : Springer Netherlands, 1985, p. 309-315 (cf. p. 45, 46).
- [Smi+89] K. SMITH, D. ALEXANDROU et J. EDELMAN. "Acoustic detection and tracking of abyssopelagic animals : description of an autonomous split-beam acoustic array". In : *Deep Sea Research Part A. Oceanographic Research Papers* 36.9 (sept. 1989), p. 1427-1441 (cf. p. 47).
- [Spa+16] G. SPAGNOLI, S. MIEDEMA, C. HERRMANN, J. RONGAU, L. WEIXLER et J. DENEGRE. "Preliminary design of a trench cutter system for deep-sea mining applications under hyperbaric conditions". In : *IEEE Journal of Oceanic Engineering* 41.4 (oct. 2016), p. 930-943 (cf. p. 46).
- [The19] THE GUARDIAN. *Collapse of PNG deep-sea mining venture sparks calls for moratorium*. Sept. 2019. URL : <http://www.theguardian.com/world/2019/sep/16/collapse-of-png-deep-sea-mining-venture-sparks-calls-for-moratorium> (visité le 24/07/2020) (cf. p. 42).
- [Vac98] R. VACCARO. "The past, present, and the future of underwater acoustic signal processing". In : *IEEE Signal Processing Magazine* 15.4 (juill. 1998), p. 21-51 (cf. p. 44).
- [Wid+67] B. WIDROW, P. MANTEY, L. GRIFFITHS et B. GOODE. "Adaptive antenna systems". In : *Proceedings of the IEEE* 55.12 (déc. 1967), p. 2143-2159 (cf. p. 44).
- [Yu+10] C. YU et Y. CHENG. "System and method of utilizing monitoring data to enhance seafloor sulfide production for deep water mining system". US 7,784,201 B2 (Houston). Août 2010 (cf. p. 46).

Bibliographie du chapitre 4

- [Age14] AGENCE FRANCE-PRESSE. "More drones spotted over French nuclear power stations". In : *The Guardian* (oct. 2014). (Visité le 02/09/2020) (cf. p. 70).
- [Al +19] S. AL EMADI, A. AL ALI, A. MOHAMMAD et A. AL ALI. "Audio based drone detection and identification using deep learning". In : *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*. Juin 2019, p. 459-464 (cf. p. 74).
- [Anw+19] M. ANWAR, Z. KALEEM et A. JAMALIPOUR. "Machine learning inspired sound-based amateur drone detection for public safety applications". In : *IEEE Transactions on Vehicular Technology* 68.3 (mars 2019), p. 2526-2534 (cf. p. 75, 105).
- [BBC18] BBC. "Drones ground flights at Gatwick". In : *BBC News* (déc. 2018). (Visité le 02/09/2020) (cf. p. 70).

- [BBC20] BBC. "Drone and light aircraft in dangerous 'near miss'". In : *BBC News* (juill. 2020). (Visité le 02/09/2020) (cf. p. 70).
- [Ber+17b] A. BERNARDINI, F. MANGIATORDI, E. PALLOTTI et L. CAPODIFERRO. "Drone detection by acoustic signature identification". In : *IS&T International Symposium on Electronic Imaging*. T. 2017. Jan. 2017, p. 60-64 (cf. p. 75).
- [Bir+15] G. BIRCH, J. GRIFFIN et M. ERDMAN. *UAS detection, classification, and neutralization : market survey 2015*. 2015 (cf. p. 70).
- [Bla19] T. BLANCHARD. "Caractérisation de drones en vue de leur localisation et de leur suivi à partir d'une antenne de microphones". Français. Thèse de doct. Université du Mans, déc. 2019, p. 117 (cf. p. 19, 72, 85).
- [Bla+19a] T. BLANCHARD, J. THOMAS et K. RAOOF. "Acoustic localization estimation of an Unmanned Aerial Vehicle using microphone array". In : *Inter-Noise*. Madrid, sept. 2019 (cf. p. 73).
- [Bla+19b] T. BLANCHARD, J. THOMAS et K. RAOOF. "Acoustic signature analysis for localization estimation of unmanned aerial vehicles using few number of microphones". In : *MATEC Web of Conferences* 283 (2019), p. 1-4 (cf. p. 19, 73).
- [Bla+19c] T. BLANCHARD, J. THOMAS et K. RAOOF. "Caractérisation acoustique d'un drone pour la localisation à partir d'une antenne microphonique par formation de voies et goniométrie". Français. In : *Colloque Gretsi*. Lille, août 2019 (cf. p. 73).
- [Bor+19] A. BORGHGRAEF et M. VANDEWAL. "Evaluation of acoustic detection of UAVs using machine learning methods". In : *Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies III*. T. 11166. Oct. 2019, p. 1-10 (cf. p. 74).
- [Bus+15] J. BUSSET, F. PERRODIN, P. WELLIG, B. OTT, K. HEUTSCHI, T. RÜHL et T. NUSSBAUMER. "Detection and tracking of drones using advanced acoustic cameras". In : *Unmanned/Unattended Sensors and Sensor Networks XI; and Advanced Free-Space Optical Communication Techniques and Applications*. T. 9647. Oct. 2015, p. 1-9 (cf. p. 73).
- [Cab+19] A. CABRERA-PONCE, J. MARTINEZ-CARRANZA et C. RASCON. "Detection of nearby UAVs using CNN and spectrograms". In : *11th International Micro Air Vehicle Competition and Conference*. Madrid, Spain, oct. 2019, p. 138-143 (cf. p. 75).
- [Can+19] P. CANDELORO, T. PAGLIAROLI, D. RAGNI et S. DI FRANCESCO. "Small-scale rotor aeroacoustics for drone propulsion : a review of noise sources and control strategies". In : (oct. 2019). DOI : [10.20944/preprints201910.0078.v1](https://doi.org/10.20944/preprints201910.0078.v1) (cf. p. 72).
- [Cha+18a] X. CHANG, C. YANG, X. SHI, P. LI, Z. SHI et J. CHEN. "Feature extracted DOA estimation algorithm using acoustic array for drone surveillance". In : *2018 IEEE 87th Vehicular Technology Conference (VTC Spring)*. Juin 2018, p. 1-5 (cf. p. 73).
- [Cha+18b] X. CHANG, C. YANG, J. WU, X. SHI et Z. SHI. "A Surveillance System for Drone Localization and Tracking Using Acoustic Arrays". In : *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. Juill. 2018, p. 573-577 (cf. p. 73).
- [D S+17] D. SHIN, D. JUNG, D. KIM, J. HAM et S. PARK. "A Distributed FMCW Radar System Based on Fiber-Optic Links for Small Drone Detection". In : *IEEE Transactions on Instrumentation and Measurement* 66.2 (fév. 2017), p. 340-347 (cf. p. 70).
- [Der13] DER SPIEGEL. "Federal Police Report Explains Inaction Despite Mini Drone Attack". In : (sept. 2013). (Visité le 02/09/2020) (cf. p. 70).
- [Dro20] DRONE INDUSTRY INSIGHT. *The drone market report 2020-2025*. Août 2020 (cf. p. 70).
- [Dum+20] C. DUMITRESCU, M. MINEA, I. MADALINA COSTEA, I. COSMIN CHIVA et A. SEMENESCU. "Development of an Acoustic System for UAV Detection". In : *Sensors* 20.17 (août 2020), p. 1-27 (cf. p. 75).

- [Eur20] EUROPEAN DEFENCE INDUSTRIAL DEVELOPMENT PROGRAMME. *Capabilities to detect, classify, track, identify and/or counter UASs in defence scenarios*. Call ID : EDIDP-CUAS-2020. Mars 2020. URL : <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/edidp-cuas-2020;freeTextSearchKeyword=;typeCodes=1;statusCodes=31094501,31094502,31094503;programCode=EDIDP;programDivisionCode=null;focusAreaCode=null;crossCuttingPriorityCode=null;callCode=EDIDP-CUAS-2020;sortQuery=submissionStatus;orderBy=asc;onlyTenders=false;topicListKey=callTopicSearchTableState> (visité le 02/09/2020) (cf. p. 70).
- [Gan+20] U. GANAPATHI et M. SABARIMALAI MANIKANDAN. "Convolutional neural network based sound recognition methods for detecting presence of amateur drones in unauthorized zones". In : *Machine Learning, Image Processing, Network Security and Data Sciences*. Springer, juin 2020, p. 229-244 (cf. p. 75, 105).
- [Gha+20] S. GHANI et W. KHAN. "Extraction of UAV sound from a mixture of different sounds". In : *Acoustics Australia* (août 2020) (cf. p. 75, 107).
- [Gio+17] F. GIONES et A. BREM. "From toys to tools : the co-evolution of technological and entrepreneurial developments in the drone industry". In : *Business Horizons* 60.6 (nov. 2017), p. 875-884 (cf. p. 70).
- [Ham20] D. HAMBLING. "'Drone Swarm' Invaded Palo Verde Nuclear Power Plant Last September — Twice". In : *Forbes* (juill. 2020). (Visité le 02/09/2020) (cf. p. 70).
- [He+19] Y. HE, I. AHMAD, L. SHI et K. CHANG. "SVM-based drone sound recognition using the combination of HLA and WPT techniques in practical noisy environment". In : *KSII Transactions on Internet and Information Systems* 13.10 (oct. 2019) (cf. p. 75).
- [Her+20] G. HEROLD, A. KUJAWSKI, C. STRUMPFEL, S. HUSCHBECK, M. de HAAG et E. SARRADJ. "Detection and separate tracking of swarm quadcopter drones using microphone array measurements". In : *8th Berlin Beamforming Conference*. Berlin, mars 2020, p. 1-19 (cf. p. 73).
- [Hub+19] B. HUBBARD, P. KARASZ et S. REED. "Two Major Saudi Oil Installations Hit by Drone Strike, and U.S. Blames Iran". In : *The New York Times* (sept. 2019). (Visité le 02/09/2020) (cf. p. 70).
- [Ian+20] G. IANNACE, G. CIABURRO et A. TREMATERRA. "Acoustical unmanned aerial vehicle detection in indoor scenarios using logistic regression model". In : *Building Acoustics* (mai 2020), p. 1-20 (cf. p. 74).
- [Jeo+17] S. JEON, JW. SHIN, YJ. LEE, WH. KIM, YH. KWON et HY. YANG. "Empirical study of drone sound detection in real-life environment with deep neural networks". In : *2017 25th European Signal Processing Conference (EUSIPCO)*. Août 2017, p. 1858-1862 (cf. p. 74, 105).
- [Joh20] B. JOHNSON. "ISIS Shows Seizure, Study of U.S.-Made Drone in Sinai". In : *Homeland Security Today* (août 2020). (Visité le 02/09/2020) (cf. p. 70).
- [Kim+18] J. KIM et D. KIM. "Neural network based real-time UAV detection and analysis by sound". In : *Journal of JAITS* 8.1 (juill. 2018), p. 43-52 (cf. p. 74).
- [Kim+17] J. KIM, C. PARK, J. AHN, Y. KO, J. PARK et J. GALLAGHER. "Real-time UAV sound detection and analysis system". In : *2017 IEEE Sensors Applications Symposium (SAS)*. Mars 2017, p. 1-5 (cf. p. 74).
- [Koc16] W. KOCH. "On multiple sensor UAS security research aspects and first experimental results". In : *Security Research Conference : 11th Future Security*. Berlin, sept. 2016, p. 1-8 (cf. p. 71).
- [Kol+20] H. KOLAMUNNA, J. LI, T. DAHANAYAKA, S. SENEVIRATNE, K. THILAKARATNE, A. ZOMAYA et A. SENEVIRATNE. "AcousticPrint : acoustic signature based open set drone identification". In : *Proceedings of the 13th ACM Conference on Security and Privacy in Wireless and Mobile Networks*. New York, NY, USA : Association for Computing Machinery, juill. 2020, p. 349-350 (cf. p. 75).
- [Lyk+20] Georgia LYKOU, Dimitrios MOUSTAKAS et Dimitris GRITZALIS. "Defending Airports from UAS : A Survey on Cyber-Attacks and Counter-Drone Sensing Technologies". In : *Sensors* 20.12 (juin 2020), p. 1-40 (cf. p. 71).
- [Mal18] M. MALFANTE. "Automatic classification of natural signals for environmental monitoring". English. PhD thesis. Grenoble Alpes, oct. 2018 (cf. p. 75, 78, 88, 90).

- [Mal+17a] M. MALFANTE, M. DALLA MURA, B. BOULLAY, J.P. MÉTAXIAN et J. MARS. "Apprentissage statistique : classification automatique de signaux volcano-sismiques". In : *XXVIème colloque GRETSI (GRETSI 2017)*. Proceeding Actes de la conférence GretsI 2017. Juan-Les-Pins, France, sept. 2017 (cf. p. 78).
- [Mal+17b] M. MALFANTE, M. DALLA MURA, J. MARS et JP. MÉTAXIAN. "Machine learning for automatic classification of volcano-seismic signatures". In : *2017 25th European Signal Processing Conference (EUSIPCO)*. Août 2017, p. 2393-2397 (cf. p. 78).
- [Mal+18a] M. MALFANTE, M. DALLA MURA, J. MARS, JP. MÉTAXIAN, O. MACEDO et A. INZA. "Automatic Classification of Volcano Seismic Signatures". In : *Journal of Geophysical Research : Solid Earth* 123.12 (2018), p. 10, 645-10, 658 (cf. p. 78).
- [Mal+18b] M. MALFANTE, M. DALLA MURA, J.P. METAXIAN, J. MARS, O. MACEDO et A. INZA. "Machine learning for volcano-seismic signals : challenges and perspectives". In : *IEEE Signal Processing Magazine* 35.2 (mars 2018), p. 20-30 (cf. p. 78, 79).
- [Mal+] M. MALFANTE, J. MARS et M. DALLA MURA. *Automatic Analysis Architecture*. <https://zenodo.org/record/1216028>. Accédé le : 14/09/2020 (cf. p. 79).
- [Mal+18c] M. MALFANTE, J. MARS, M. DALLA MURA et C. GERVAISE. "Automatic fish sounds classification". In : *The Journal of the Acoustical Society of America* 143.5 (mai 2018), p. 2834-2846 (cf. p. 78, 79).
- [Mal+18d] M. MALFANTE, O. MOHAMMED, C. GERVAISE, M. DALLA MURA et J. MARS. "Use of deep features for the automatic classification of fish sounds". In : *MTS/IEEE Oceans 2018*. Kobe, Japan, mai 2018, p. 1-5 (cf. p. 78).
- [Mer+17] P. MERRIAUX, Y. DUPUIS, R. BOUTTEAU, P. VASSEUR et X. SAVATIER. "A study of vicon system positioning performance". In : *Sensors* 17.7 (juill. 2017), p. 1591 (cf. p. 81).
- [Mez+15] J. MEZEI, V. FIASKA et A. MOLNÁR. "Drone sound detection". In : *2015 16th IEEE International Symposium on Computational Intelligence and Informatics (CINTI)*. Nov. 2015, p. 333-338 (cf. p. 74).
- [Mez+16] J. MEZEI et A. MOLNÁR. "Drone sound detection by correlation". In : *2016 IEEE 11th International Symposium on Applied Computational Intelligence and Informatics (SACI)*. Mai 2016, p. 509-518 (cf. p. 74).
- [Mic18] A. MICHEL. *Counter-drone systems*. Fév. 2018 (cf. p. 70).
- [Mic19] A. MICHEL. *Counter-drone systems 2nd Edition*. Déc. 2019 (cf. p. 70, 72).
- [New20] Arab News. "Daesh supporter tried to modify drone for use in UK terror attack". In : *Arab News* (juill. 2020). (Visité le 02/09/2020) (cf. p. 70).
- [Nij+16] M. NIJIM et N. MANTRAWADI. "Drone classification and identification system by phenome analysis using data mining techniques". In : *2016 IEEE Symposium on Technologies for Homeland Security (HST)*. Mai 2016, p. 1-5 (cf. p. 74).
- [Pha+04] T. PHAM et N. SROUR. "TTCP AG-6 : acoustic detection and tracking of UAVs". In : *Unattended/Unmanned Ground, Ocean, and Air Sensor Technologies and Applications VI*. T. 5417. Sept. 2004, p. 24-31 (cf. p. 73).
- [Poi+17] P. POITEVIN, M. PELLETIER et P. LAMONTAGNE. "Challenges in detecting UAS with radar". In : *2017 International Carnahan Conference on Security Technology (ICCST)*. Oct. 2017, p. 1-6 (cf. p. 71).
- [Ram19] A. RAMAMONJY. "Développement de nouvelles méthodes de classification/localisation de signaux acoustiques appliquées aux véhicules aériens". Français. Thèse de doct. Paris, CNAM, mai 2019, p. 238 (cf. p. 18, 73, 85).
- [Ram+16] A. RAMAMONJY, E. BAVU et A. GARCIA. "Détection, classification et suivi de trajectoire de sources acoustiques par captation pression-vitesse sur capteurs MEMS numériques". Français. In : *CFA / Vishno*. Le Mans, avr. 2016, p. 7 (cf. p. 73).
- [Ram+18] A. RAMAMONJY, E. BAVU, A. GARCIA et S. HENGY. "Source localization and identification with a compact array of digital MEMS microphones". In : *ICSV25*. Hiroshima, juill. 2018, p. 1-8 (cf. p. 18, 73).

- [Sed+19] A. SEDUNOV, D. HADDAD, H. SALLOUM, A. SUTIN, N. SEDUNOV et A. YAKUBOVSKIY. "Stevens Drone Detection Acoustic System and Experiments in Acoustics UAV Tracking". In : *2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. Nov. 2019, p. 1-7 (cf. p. 73).
- [Sed+18] A. SEDUNOV, H. SALLOUM, A. SUTIN, N. SEDUNOV et S. TSYURYUPA. "UAV passive acoustic detection". In : *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*. Oct. 2018, p. 1-6 (cf. p. 73).
- [Seo+20] S. SEO, S. YEO, H. HAN, Y. KO, K. HO et E. MATSON. "Single node detection on direction of approach". In : *2020 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. Mai 2020, p. 1-6 (cf. p. 74).
- [Seo+18] Y. SEO, B. JANG et S. IM. "Drone detection using convolutional neural networks with acoustic STFT features". In : *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. Nov. 2018, p. 1-6 (cf. p. 75, 105).
- [Ser+19] R. SERRÉ, H. FOURNIER et JM. MOSCHETTA. "A design methodology for quiet and long endurance MAV rotors". In : *International Journal of Micro Air Vehicles* 11 (mai 2019), p. 1-14 (cf. p. 72).
- [Shi+18] L. SHI, I. AHMAD, Y. HE et K. CHANG. "Hidden markov model based drone sound recognition using MFCC technique in practical noisy environments". In : *Journal of Communications and Networks* 20.5 (oct. 2018), p. 509-518 (cf. p. 74).
- [Shi+20] Z. SHI, X. CHANG, C. YANG, Z. WU et J. WU. "An Acoustic-Based Surveillance System for Amateur Drones Detection and Localization". In : *IEEE Transactions on Vehicular Technology* 69.3 (mars 2020), p. 2731-2739 (cf. p. 73).
- [Sie] SIEMENS. *Simcenter Sound Camera*. URL : https://www.plm.automation.siemens.com/media/global/en/Siemens-PLM-Simcenter-Sound-Camera-eb-63902-A13_tcm27-56488.pdf (visité le 10/09/2020) (cf. p. 81).
- [Sin+13] G. SINIBALDI et L. MARINO. "Experimental analysis on the noise of propellers for small UAV". In : *Applied Acoustics* 74.1 (jan. 2013), p. 79-88 (cf. p. 71).
- [Sol+18] A. SOLODOV, A. WILLIAMS, S. AL HANAËI et B. GODDARD. "Analyzing the threat of unmanned aerial vehicles (UAV) to nuclear facilities". In : *Security Journal* 31.1 (fév. 2018), p. 305-324 (cf. p. 70).
- [Tah+19] B. TAHA et A. SHOUFAN. "Machine learning-based drone detection and classification : state-of-the-art in research". In : *IEEE Access* 7 (2019), p. 138669-138682 (cf. p. 74).
- [Ton+13] J. TONG, Y. HU, M. BAO et W. XIE. "Target tracking using acoustic signatures of light-weight aircraft propeller noise". In : *2013 IEEE China Summit and International Conference on Signal and Information Processing*. Juill. 2013, p. 20-24 (cf. p. 73).
- [Ton+16] J. TONG, W. XIE, Y. HU, M. BAO, X. LI et W. HE. "Estimation of low-altitude moving target trajectory using single acoustic array". In : *The Journal of the Acoustical Society of America* 139.4 (avr. 2016), p. 1848-1858 (cf. p. 73).
- [Udd+20] Z. UDDIN, M. ALTAË, M. BILAL, L. NKENYEREYE et A. BASHIR. "Amateur drones detection : a machine learning approach utilizing the acoustic signals in the presence of strong interference". In : *Computer Communications* 154 (mars 2020), p. 236-245 (cf. p. 75, 107).
- [Wan+18] Y. WANG, H. MA, S. WEI, S. ZHANG, Z. FENG et Z. WEI. "Sound Detection and Alarm System of Unmanned Aerial Vehicle". In : *Recent Developments in Intelligent Computing, Communication and Devices*. Springer, août 2018, p. 885-898 (cf. p. 75, 105).
- [Woo20] C. WOODY. "Drones are dropping bombs on US troops in Syria, and it's not clear who's doing it". In : *Business Insider France* (mars 2020). (Visité le 02/09/2020) (cf. p. 70).
- [Wu+19] Z. WU, C. YANG, X. CHANG, J. WU et Z. SHI. "A Drone Localization System Based on Multiple Time Delays Fusion". In : *2019 IEEE Radar Conference (RadarConf)*. Avr. 2019, p. 1-6 (cf. p. 73).
- [Yan+19] B. YANG, E. MATSON, A. SMITH, J. DIETZ et J. GALLAGHER. "UAV detection system with multiple acoustic nodes using machine learning models". In : *2019 Third IEEE International Conference on Robotic Computing (IRC)*. Fév. 2019, p. 493-498 (cf. p. 74).

- [Yan+18] C. YANG, Z. WU, X. CHANG, X. SHI, J. WO et Z. SHI. "DOA estimation using amateur drones harmonic acoustic signals". In : *2018 IEEE 10th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. Juill. 2018, p. 587-591 (cf. p. 73).
- [Zaw+20] N. ZAWODNY et D. BOYD. "Investigation of Rotor–Airframe Interaction Noise Associated with Small-Scale Rotary-Wing Unmanned Aircraft Systems". In : *Journal of the American Helicopter Society* 65.1 (jan. 2020), p. 1-17 (cf. p. 72).
- [Zel+09] A. ZELNIO, E. CASE et B. RIGLING. "A low-cost acoustic array for detecting and tracking small RC aircraft". In : jan. 2009, p. 121-125 (cf. p. 73).



FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : BARON
(avec précision du nom de jeune fille, le cas échéant)

DATE de SOUTENANCE : 14/12/2020

Prénoms : Valentin

TITRE : Méthodes d'identification de sources acoustiques paramétriques par mesures d'antennerie

NATURE : Doctorat

Numéro d'ordre : 2020LYSEI121

Ecole doctorale : EEA

Spécialité : Traitement du Signal

RESUME : La caractérisation de sources acoustiques a pour but de décrire des émetteurs sonores à travers certains paramètres comme leur localisation dans l'espace, le niveau sonore qu'ils produisent ou encore leur identification grâce à leur signature acoustique. Dans cette thèse, l'objectif est d'obtenir certains de ces trois paramètres dans deux cas d'application concrets, pour des sources situées en champ lointain et grâce à des antennes acoustiques.

La première application concerne l'impact acoustique de la prospection minière sous-marine par grands fonds dans le cadre du projet FUI Abyssound. Au sein du projet, la thèse cherche à caractériser les machines excavatrices placées sur les fonds marins en obtenant leur localisation et leur niveau sonore. Une première phase de design a abouti à la construction d'une antenne conique de 3 m. En s'appuyant ensuite sur les données issues de deux campagnes expérimentales menées en mer Méditerranée avec cette antenne, la méthode haute-résolution MUSIC localise précisément les sources mises en œuvre, qu'elles soient mobiles et à plus de 600 m de l'antenne, ou immergées par 700 m de fond. Leur niveau sonore est ensuite estimé par formation de voies, et les niveaux attendus sont retrouvés pour des signaux monochromatiques ou à large bande.

Dans la seconde application, une procédure complète pour la localisation et l'identification de drones est proposée pour la protection de sites sensibles. Elle combine traitement d'antenne et apprentissage statistique en s'articulant autour de trois étapes clés : la localisation, la focalisation et l'identification. La méthode MUSIC localise à nouveau les sources acoustiques présentes autour de l'antenne industrielle utilisée, puis la focalisation reconstruit le signal temporel de chacune, et un modèle SVM les identifie comme drone ou non. Les validations expérimentales, en intérieur comme en extérieur, constituent une contribution importante de ce travail de thèse. Les données acquises montrent entre autres que la procédure localise des drones à 3° près en extérieur, les détecte à 99 %, ou encore les identifie en présence d'une source perturbatrice plus puissante.

MOTS-CLÉS : traitement d'antenne, machine learning, détection de drones, estimation de niveau sonore.

Laboratoire (s) de recherche : CREATIS

Directeur de thèse: Barbara NICOLAS

Président de jury : Pascal LARZABAL

Composition du jury : Sylvie MARCOS, Jean-Hugh THOMAS, Eric BAVU, Jérôme ANTONI, Barbara NICOLAS, Simon BOULEY, Jérôme MARS