COLLEGES BRETAGNE LOIRE — SCIENCES POUR L'INGENIEUR ET LE NUMERIQUE

ENSTA — INSTITUT POLYTECHNIQUE DE PARIS

THÈSE DE DOCTORAT DE

L'ÉCOLE NATIONALE SUPÉRIEURE
DE TECHNIQUES AVANCÉES

ÉCOLE DOCTORALE N° 648
Sciences pour l'Ingénieur et le Numérique
Spécialité : *Signal, Image, Vision*

Par
Mathilde MICHEL

Optimisation du suivi des cétacés par acoustique passive

Thèse présentée et soutenue à Brest, le 08/12/2025
Unité de recherche : Lab-STICC UMR CNRS 6285

Rapporteur·trice·s avant soutenance :

Sofie VAN PARIJS — Professeure, National Oceanic and Atmospheric Administration (NOAA)
Nigel Gilles YOCCOZ — Professeur, The Arctic University of Norway

Composition du Jury :

Président :	Mathieu PAQUIER	Professeur, UBO, Lab-STICC
Examinateur·trice·s :	Mathieu PAQUIER	Professeur, UBO, Lab-STICC
	Emmanuelle LEROY	Chargée de recherche, GLOBICE
	Gaëtan RICHARD	Ingénieur d'étude, SOMME
Dir. de thèse :	Angélique DREMEAU	Professeure, ENSTA, IPP, Lab-STICC
Encadrante :	Flore SAMARAN	Enseignante-chercheuse, ENSTA, IPP, Lab-STICC

Invitée :

Sara Bazin — Professeure, UBO, IUEM, Geo-Ocean

Remerciements

On m'avait dit que les remerciements de thèse seraient la chose la plus difficile à écrire. Et pourtant, il m'a semblé étrangement évident de savoir à qui je devais cette étape importante de ma vie.

Mes premiers remerciements vont tout naturellement à **Flore Samaran**, sans qui tout cela n'aurait pas été possible. **Flore**, je me rappelle encore avoir pris mon courage à deux mains pour venir te parler de moi, de mon profil et de mon envie de travailler avec toi. Je ne te remercierai jamais assez d'avoir eu confiance en moi, de m'avoir proposé ce stage, puis d'avoir accepté de m'accompagner dans cette thèse. Merci d'avoir cru en moi, de m'avoir encouragée, de m'avoir laissé la liberté et l'autonomie dont j'avais besoin, tout en gardant un regard bienveillant sur mon travail. Merci pour tes propositions, ton écoute, tes conseils, et de t'être battue pour que je puisse réaliser certaines envies. Merci aussi d'avoir tant œuvré pour que je puisse réaliser cette thèse aussi sereinement que possible, et d'y être parvenue avec brio. Je te serai toujours reconnaissante de m'avoir permis d'embarquer sur le Marion Dufresne (j'ai d'ailleurs tellement hâte de vivre ces aventures que tu m'as si souvent racontées). Merci enfin de m'avoir inspirée et de m'avoir transmis tes connaissances sur les baleines, mais surtout de m'avoir appris que la recherche, ce n'est pas seulement des chiffres, c'est aussi des histoires que l'on essaie de comprendre et de raconter.

Merci également à **Angélique Drémeau**, sans qui j'aurais eu bien du mal à concrétiser ce projet. Merci d'avoir accepté d'être ma directrice de thèse et de m'avoir fait confiance. Merci pour ta bonne humeur dans les couloirs, pour ta répartie sans faille et pour ta bienveillance avec tes « ça va ? » réguliers, qui voulaient toujours dire : « Mathilde, comment vas-tu réellement ? ».

Je souhaite aussi remercier mon jury : **Sara Bazin**, d'avoir accepté de présider ce jury ; **Sofie Van Parijs** et **Nigel Yoccoz**, d'avoir accepté d'être rapporteurs, de lire et de corriger attentivement mon manuscrit ; **Emmanuelle Leroy** et **Gaëtan Richard**, et aussi **Mathieu Paquier**. Merci à vous tous pour le temps et l'attention que vous avez

consacrés à mon travail, ainsi que pour vos remarques précieuses.

Baptiste Alglave, **Maxime Olmos**, **Auriane Virgili**, les boss de la modélisation et des statistiques. Merci à vous trois de m'avoir accompagnée, de m'avoir orientée et guidée avec autant de rigueur que de douceur. Ça a été un plaisir de travailler avec vous sur ces projets. J'ai énormément appris à vos côtés, alors merci pour ces discussions enrichissantes d'écologue et de statisticiens. Mention spéciale à toi, **Baptiste**, pour toutes ces petites réunions auxquelles tu as toujours répondu présent.

Un gigantesque merci à **Maëlle** et **Mathieu**. Tant de choses à vous dire, et bien que vous soyez un couple formidable, je pense que je vous dois des remerciements à chacun de vous.

Maëlle, merci d'avoir été cette machine de savoir et de connaissances à qui aucune question ne résiste. Merci pour ton optimisme sans faille tout au long de mon travail et pour avoir réussi à me déstresser en dédramatisant. Merci pour toutes tes expressions que j'emporterai avec moi, et pour le peu de mots bretons que je sais dire. Je n'ai aucun doute que le mot « pur » va finir par se propager dans toute la France...

Mathieu, merci d'avoir rendu ma vie à Brest plus joyeuse et solaire. Je me souviens encore de nos premiers moments, et il ne m'a pas fallu longtemps pour savoir qu'on s'entendrait si bien. Merci pour ces fous rires, ces discussions sans queue ni tête (et pour clore le débat une bonne fois pour toutes : oui, c'est un étendage, pas un tancarville). Merci aussi pour tes imitations (Eh. oui) et ton sens de l'autodérision qui m'ont toujours beaucoup fait rire. Merci aussi d'avoir été une oreille attentive, toujours prête à écouter toutes mes conneries sans jamais émettre le moindre jugement.

Avec vous deux, j'ai découvert tellement de choses : la Bretagne sous toutes ses formes, la passion du sarrasin, la couture express (merci Maëlle !) et mille petites touches de votre univers que vous m'avez partagées. J'espère que vous garderez un peu de mes petites références lyonnaises : les quenelles, la vogue, les rouleaux de printemps, et même Pouloulou ! Bref, on s'est échangé des morceaux de nos vies, et c'est sans doute ce que j'ai préféré. Merci à vous deux d'avoir été mes parents brestois : pour tous ces dimanches run, bouffe et Tuches, ces randos sur les îles, ces parties endiablées sur la switch, pour BBJuni que vous avez si bien adoptée et gardée, pour ces voyages de folie en Galice ou aux Açores (pas de merci pour SATA(N) et Monica par contre). J'emporte avec moi tous ces souvenirs (on a vu des baleines bleues...), ces photos, ces rires et ces bons repas. De collègues de bureau,

vous êtes devenus des amis, et vous allez énormément me manquer. Vivement qu'on se retrouve avec un masque et une bouteille de plongée dans le lagon de Tahiti !

Merci à toi **Gauthier**, parce que même si nous ne partageons pas toujours les mêmes goûts en matière de culture (toujours pas de Magic Eye, de films bizarres ou de musiques douteuses) et que nos références semblent parfois sur deux droites parallèles, ça a toujours été un plaisir d'être en ta compagnie. Je te dois mille mercis pour la relecture attentive de ma thèse, pour tous ces commentaires rigoureux et pertinents que j'ai d'abord détesté lire, puis adoré corriger. Je me rends bien compte de l'aide que tu m'as apportée, alors merci pour toutes ces figures que tu m'as aidée à refaire, pour le temps passé à gérer mes problèmes `LaTeX` ou `PDF`, et pour ta générosité. Si cette thèse est aussi belle, c'est en partie grâce à toi. Alors maintenant qu'elle est terminée, tu n'as plus d'excuse pour enfin écrire ton papier...

Merci aussi à toi **Julie**, pour avoir été si accueillante, chaleureuse et présente à mon arrivée à Brest. Merci de m'avoir accompagnée et de m'avoir permis de me sentir bien dans cet environnement. Merci pour ta bienveillance... et pour m'avoir embarquée sur les petites missions sur la mer d'Iroise. Bon, je ne sais pas si je devrais vraiment te remercier pour ça, parce que finalement, j'ai surtout découvert que j'avais le mal de mer !

Salomé, merci d'avoir été à l'origine une stagiaire de choc et très rapidement une amie. Merci pour ton aide sur ma thèse (pour ces calculs de front de glace, heureusement qu'une de nous maîtrisait `QGIS`...), pour ton stage que tu n'as pas lâché malgré les données bancales, et pour ces bons moments sur la mer Méditerranée à scruter l'horizon. Je te souhaite de t'épanouir dans ta vie de biologiste, que ce soit en thèse ou non, à Brest ou ailleurs. J'espère que l'on se reverra, et peut-être que tu me feras découvrir Madagascar...

Merci à **Anatole**, d'avoir été mon co-bureau pendant ces 3 années, d'avoir supporté ma déco, mes humeurs et mes histoires, et pour toutes nos discussions (parfois un peu utopiques) sur la recherche... sans oublier ce fameux soir où on a rêvé, un peu éméchés, de lancer notre voilier océanographique ! Merci aussi d'avoir transformé le bureau en mini-scène de karaoké et d'avoir animé nos parties de `Geoguessr`... même si c'était un peu agaçant de ne jamais réussir à te battre !

Merci aussi à tous mes collègues de l'ENSTA : **Naïs**, pour tous ces bons gâteaux qui ont

sauvé plus d'une journée ; **Élodie**, pour avoir cédé à mes caprices et modifié les raccourcis sur `APLOSE` juste pour m'éviter de râler ; **Gabi**, pour ta bonne humeur à chaque passage sur Brest et pour ton partage de l'amour de la montagne ; **Alex**, pour ton aide en propa ou `LaTeX` et surtout pour tes envois de mèmes qui redonnent le sourire ; **Élise**, pour tes rires stridents qui traversent tout le couloir et qu'on reconnaît entre mille ; **Pierre-Yves** (vivement le Marion ! !), **Camila** et toutes les personnes de ce couloir que j'ai pu croiser durant ces quatre années.

Enfin, merci à tous les Brestois pour avoir rendu ma vie à Brest plus joyeuse et ensoleillée que je ne l'aurais imaginée.

Merci aussi à **Mathieu Authier** et **Linda Gardelle** d'avoir accepté de faire partie de mon CSI. Ça a toujours été un plaisir de vous présenter mon travail et de bénéficier de vos retours. Merci pour vos idées, votre écoute et la considération de mes aspirations. Merci à toi, **Linda**, pour les cours que nous avons partagés : je pourrais presque animer une frise de la biodiversité maintenant !

Merci à la team MIRACETI : **Céline**, **Laurine**, **Laurène**, **Claire**... pour m'avoir embarquée sur votre beau Ti'blé et sur votre magnifique mer Méditerranée. Merci à **Jérôme**, skippeur de choc, car grâce à toi j'ai adoré naviguer. Merci de m'avoir transmis votre passion pour la mer, la voile et pour ces longues attentes à scruter l'horizon en attente de dauphins... Je suis maintenant incollable sur les cétacés de Méditerranée, alors merci !

Merci à l'équipe des JJBA : **Mathilde**, **Anatole**, **Gabriel**, **Salomé**, **Quentin** et **Virgil**. Organiser cet événement avec vous a été une expérience incroyable, alors merci d'avoir rendu cela possible dans une ambiance top !

Merci à **mes parents**. Je ne sais même pas par où commencer. Merci à vous d'avoir été aussi compréhensifs et conciliants quand je vous ai appelés pour vous dire que j'arrêtais mes études dans le commerce, puis que je voulais finalement faire une thèse en biologie. On n'y connaissait rien à ce milieu, moi la première, mais c'est grâce à votre ouverture et votre confiance que j'ai pu me lancer. Merci de m'avoir soutenue depuis toute petite dans tout ce que j'aspirais à faire, d'avoir été dans l'ombre pour que je puisse continuer à

poursuivre mes rêves et mes projets.

Merci à toi, **papa**. Je me souviens encore de la phrase que tu m'as dite quand je t'ai parlé de la biologie : « Mathilde, tu peux faire ce que tu veux dans la vie, tant que tu y mets tout ton cœur ». Merci d'avoir assuré pour tous ces changements de vie, de m'avoir permis de faire toutes ces années d'études sereinement, de m'avoir écoutée quand je voulais tout arrêter, de m'avoir poussée à continuer, et de m'encourager à voir plus grand, à croire que tout était possible.

Merci à toi, **maman**, d'avoir toujours été ce regard aimant, sans jugement et toujours disponible. Je ne saurais même pas comment décrire le sentiment que je ressens à chaque fois que je t'appelle, mais ta force et ta sérénité m'ont toujours permis de m'endormir en me disant : « tout va bien ». Merci aussi pour toutes ces fois où tu es venue en Bretagne, pour ces magnifiques week-ends spa et balades, qui m'ont apporté énormément de réconfort.

Vous m'avez tous les deux fait croire que je pouvais tout faire, même une thèse quand je pensais que je n'y arriverais jamais. Alors cette thèse, elle est aussi grâce à vous.

Merci à **mes frères**, mes meilleurs amis, mes pirates indisciplinés, qui m'ont tellement chouchoutée que j'ai parfois eu l'impression d'avoir grandi comme une princesse, qui ont toujours partagé leurs rires, leurs bêtises et leurs conseils, et m'ont accompagnée dans toutes mes aventures... Grandir avec vous a été la meilleure chose du monde, et je suis tellement fière du trio soudé que nous formons.

À **Arthur**, pour le grand frère protecteur et inspirant qu'il est, que ce soit dans sa vie pro, perso, ou sur les chemins de montagne... même si je dois avouer que ça m'agace un peu que tu sois si fort en trail ! Merci d'avoir toujours pris soin de moi dès que tu en avais l'occasion (mention spéciale aux montagnes de gorgonzola que tu m'achètes à chaque fois...).

À **Alex**, merci d'avoir toujours été là, même quand j'étais un peu perdue. Parfois je me demande qui de nous deux est le plus âgé et qui veille sur l'autre... Même si tu es le petit dernier, je crois que ces dernières années, c'est plutôt toi qui as veillé sur moi.

Et je ne peux pas ne pas mentionner notre petit côté visionnaire avec ce tatouage de baleine, choisi bien avant même que je ne pense travailler dessus... Je veux croire qu'il m'a porté chance !

Merci également à **Margot** et **Eve**, mes formidables belles-sœurs. À **Margot**, pour toutes ces années de rires partagés, que ce soit au poney, sur les skis, en soirées ou autour des moments de famille recomposée que tu as contribué à réunir avec Arthur. Merci pour ta bonne humeur constante, ton humour contagieux, tes potins sans fin, ton rire si... discret et ton accueil toujours chaleureux (mention spéciale aussi à toi pour les montagnes de gorgonzola !). À **Eve**, pour nos discussions passionnantes sur la vie et tes anecdotes d'hôtellerie parisienne, qui m'ont toujours fait énormément rire.

Merci à **mes grands-parents**, parce que j'ai toujours ressenti de la fierté dans votre regard et pour l'amour inconditionnel que vous m'avez transmis.

Merci à vous **Papy** et **Mamy** d'être un exemple de sagesse et d'amour. **Papy**, pour le meilleur couscous de France et **Mamy**, pour ta douceur infinie.

Merci aussi à **Papou et Mamité** pour tous ces bons repas du mercredi midi (je prends déjà commande pour le prochain : riz, tomates à la Provençale, frites et aïoli...) et vos appels réguliers. À toi **Papou**, pour toutes ces connaissances sur l'histoire et la France qui m'ont donné envie d'explorer le pays à travers tes récits. Tes anecdotes m'accompagnent partout, et j'espère un jour en avoir autant que toi à raconter. Merci à **Mamité**, pour la femme exceptionnelle que tu es. Merci de toujours partager avec moi tes histoires de vie et tous tes précieux conseils pour les randos. Si Papou a visité tous les Géant Casino de France, toi, tu as conquis toutes les montagnes françaises ; j'espère seulement être à la hauteur pour franchir tous les cols que tu as traversés ! Merci aussi pour ces parties de Rummikub... même si, depuis un moment, tu ne me laisses plus gagner !

Merci aussi à toute ma famille d'avoir toujours été là, pour votre générosité infinie et pour tous ces bons moments : **Sophie**, **Gilbert**, **Nathalie**, et mes cousin · es chéri · es, **Sébastien**, **Charlotte**, **Mei**... À mon parrain, **Pierre-Jean**, je pense qu'on aura le même souvenir d'une Bretagne pluvieuse...

Merci à **Louise** et **Marie**, mes meilleures amies. Merci à vous, les filles, de m'accompagner depuis si longtemps. Merci pour tout ce qu'on a traversé ensemble, pour cette jeunesse de folie, ces vacances mémorables, ces soirées infinies et tous ces moments qui ont façonné ce que nous sommes aujourd'hui. Merci d'avoir fait en sorte que la distance n'ait jamais eu raison de notre amitié, et que même à des kilomètres les unes des autres, nous

restions inséparables. Merci de m'avoir permis de grandir et de m'épanouir sereinement en sachant que je pourrai toujours vous appeler (même pour aller sauver mon chat quand je suis à l'autre bout du monde...). Merci d'être si fortes et justes, car ça me permet de l'être aussi.

À toi **Louise**, pour toutes nos aventures depuis petites : le cheval, le ski, nos aventures du collège, les karaokés dans la voiture, les vacances dans le Sud, le 4L Trophy, Marseille... Vingt ans à me nourrir de chaque instant passé avec toi, et j'ai toujours autant de plaisir à regarder en arrière pour mesurer le chemin incroyable que nous avons parcouru ensemble. Merci pour ta folie douce, ton énergie contagieuse, et tous ces souvenirs qui font que chaque moment avec toi devient unique. J'ai hâte de continuer à en écrire de nouveaux... et surtout, merci de rester ma partenaire d'aventures, même quand elles nous entraînent dans des situations totalement improbables.

À toi **Marie**, parce que tu es sans doute la meilleure pour comprendre ce que ces années ont représenté. Merci d'avoir toujours été à l'écoute, patiente, et d'avoir su m'accompagner avec tant de douceur et de bienveillance. J'ai toujours admiré ta détermination et ton ambition, qui me poussent à être moi aussi plus audacieuse et persévérante. De petites Lyonnaises fashion au lycée avec nos sacs de stars, nous sommes devenues des femmes accomplies, je pense qu'on peut être fières de nous. J'ai hâte de pouvoir à mon tour t'appeler « docteur », parce que oui, madame va devenir pneumologue !

Merci aussi à tous mes amis qui m'ont permis de m'évader quand j'en avais besoin. Un merci tout particulier à **William** et **Geoffrey** pour ces soirées de rires, de beaucoup de conneries, de danse jusqu'au bout de la nuit et... de beaucoup trop de bières !

Et enfin, mes derniers remerciements iront à **Christophe**, parce que si tu es initialement entré dans ma vie pour rendre celle de ma maman meilleure, c'est aussi la mienne que tu as égayée. Merci pour tous les conseils que tu m'as donnés, pour m'avoir mis le pied à l'étrier de la recherche... et pour avoir ouvert la première porte (sinon, qui sait si j'en serais là aujourd'hui !). Même si je t'ai parfois maudit avec tes : « Mathilde, fais des maths, tu verras, ça paiera », je suis convaincue que si tu n'avais pas été là, cette thèse n'existerait pas. Alors un grand merci de m'avoir prise sous ton aile et de faire partie de notre vie.

Et je ne peux pas terminer sans mentionner **June**, alias BBJuni, mon petit (gros) chat roux, qui a été en première ligne de toutes mes humeurs. Toujours là pour des câlins

interminables ou pour me faire sourire, elle a conquis mon cœur... et accessoirement celui de toute l'équipe, au point de devenir notre mascotte officielle !

Et puisqu'aucun de nous ne connaissait rien à la mer, ni même aux baleines ou à l'acoustique, j'espère, au travers de cette thèse, vous transporter dans un univers fascinant que j'aie appris à comprendre, contempler et aimer. Cette thèse, je l'ai aussi écrite pour vous, pour que vous compreniez tout ce que je vous raconte depuis tout ce temps...

Et comme je n'ai jamais osé placer cette phrase si mythique avant, il me semble que je peux enfin dire que... le travail est fait !

Kenavo.

VALORISATION DES RÉSULTATS DE THÈSE

Publications scientifiques

En tant que première autrice

— **Michel, M.**, Virgili, A., Dréo, R., Alglave, B., Torterotot, M., Bazin, S., Samaran, F. (en cours de révision dans Global Ecology and Conservation). *Ecological habitat characterization of baleen whales in the northern Mozambique Channel : linking seasonal acoustic detections with environmental drivers.*

— **Michel, M.**, Béesau, J., Torterotot, M., Todd, N., Samaran, F. (2025). *Effects of duty cycle on passive acoustic monitoring metrics : the case of odontocetes vocalizations. The Journal of the Acoustical Society of America*, 158, 5033-5046. doi:10.1121/10.0039925

— **Michel, M.**, Alglave, B., Olmos, M., Torterotot, M., Virgili, A., Martin-Marin, S., Royer, J.-Y., Samaran, F. (2025). *Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the Southern Indian Ocean. Scientific Reports*, 15, 23218. doi:10.1038/s41598-025-02941-9

— **Michel, M.**, Guichard, B., Béesau, J., Samaran, F. (2024). *Passive acoustic monitoring for assessing marine mammals population in European waters : Workshop conclusions and perspectives. Marine Policy*, 160, 105983. doi:10.1016/j.marpol.2023.105983

— **Michel, M.**, Torterotot, M., Royer, J.-Y., Samaran, F. (2024). *Effects of duty cycle on passive acoustic monitoring metrics : The case of blue whale songs. The Journal of the Acoustical Society of America*, 155, 2538–2548. doi:10.1121/10.0025545

Autres publications

— Stavenow Jerremalm, J., Béesau, J., Carlström, J., Eriksson, P., Jessopp, M., Kavanagh, A., Loisa, O., **Michel, M.**, Owen, K., Rasmussen, M. H., Samaran, F., Todd,

N., Rogan, E. (2025). *Large-scale passive acoustic monitoring data shows seasonal and diel diversity in foraging behaviour of harbour porpoises within their distributional range in the Northeast Atlantic. Ecology and Evolution.* doi:10.1002/ece3.72566

— Dubus, G., Torterotot, M., Béesau, J., Dupont, M., Gros-Martial, A., **Michel, M.**, Morin, E., Nguyen Hong Duc, P., Raumer, P.-Y., Adam, O., Samaran, F., Cazau, D. (2025). *APLOSE : A web-based annotation platform for underwater passive acoustic monitoring. SoftwareX*, Volume 29. doi:10.1016/j.softx.2025.102055

Communications scientifiques

Présentations par poster

— **36e conférence annuelle de la European Cetacean Society**, 14–16 mai 2025, Ponta Delgada, Açores, Portugal — *Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the Southern Indian Ocean.* doi:10.13140/RG.2.2.19236.80007

— **34e conférence annuelle de la European Cetacean Society**, 18–20 avril 2023, O'Grove, Galice, Espagne — *Less is more ? How the choice of a recording duty cycle could affect monitoring results of passive acoustic studies on cetaceans.* doi:10.13140/RG.2.2.19236.80007

Présentations orales

— **Kerguelen Plateau Symposium**, 31 mars–2 avril 2025, Concarneau, France — *Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the Southern Indian Ocean.*

— **Emerging Bioacousticians Days**, 28–30 juin 2023, St Etienne, France — *Less is more ? How the choice of a recording duty cycle could affect monitoring results of passive acoustic studies on cetaceans.*

— **SERENADE**, 27–30 juin 2022, Brest, France — *Effets du sous-échantillonnage temporel : le cas des chants de baleines dans l'océan Indien.*

Workshop

— **34e conférence annuelle de la European Cetacean Society**, 18–20 avril 2023, O'Grove, Galice, Espagne — *Organisation d'un atelier scientifique : Prospective Discussion on Passive Acoustic Monitoring of Marine Mammals for Descriptor D1 of the Marine Strategy Framework Directive (MSFD).*

Activités complémentaires

Missions en mer

Participation à deux missions en Méditerranée avec l'Association MIRACETI à bord du bateau Ti'blé, du 25 novembre au 2 décembre 2023, et du 23 mars au 2 mai 2024. Réalisation d'un inventaire des cétacés par le biais d'enregistrements acoustiques sous-marins et d'observations visuelles protocolées selon la méthode des transects en ligne.

Communication grand public

— Participation à la Nuit des Chercheurs 2022 et 2023 à Océanopolis, Brest.

— Organisation le 25 juin 2024 d'une journée dédiée à la bioacoustique sous-marine dans le cadre de la Journée des Jeunes Bioacousticien · nes, à Brest (musée de l'océan : 70.8). Evénement grand public avec des activités interactives pour découvrir les sons de l'océan, leur utilisation par les animaux marins, ainsi que leur importance pour la conservation des écosystèmes marins. Plus d'infos : Page de l'événement.

— Conférence pour l'association Astrolabe Expéditions lors d'un Café Océan à la PAM, à Brest, le 28 janvier 2025, sur le suivi par acoustique passive des cétacés.

Comité d'organisation

Trésorière de l'association ALBA (Association Ligérienne de Bioacoustique), organisatrice des Journées Jeunes Bioacousticiens (JJBA) qui se sont tenues à Brest du 25 au 28 juin 2024. Conférence internationale : 70 participant · es, 9 sessions scientifiques, 56 présentations, 3 conférences plénières, 4 ateliers, diffusion en ligne avec environ 100 spectateurs par session. Responsabilités : gestion financière, planification logistique, coordination des intervenants, communication, gestion des inscriptions et rédaction du bilan post-événement. Plus d'infos : Site de la conférence.

Encadrement de stages

Encadrement du stage de Master 2 de Salomé Martin-Marin (International Master of Bioacoustics, MoBi, Université de Saint-Étienne, promotion 2023-2024) de janvier à juin 2024 : Combinaison des approches acoustiques passives et actives pour étudier les relations proie-prédateur des baleines bleues dans l'océan Indien austral. Présentation orale des résultats au WGFAST, Brest, du 9 au 12 avril 2024, ainsi qu'un poster présenté aux Journées Jeunes Bioacousticien · nes (JJBA), Brest, du 25 au 28 juin 2024.

Enseignement

Activités d'enseignement dispensées à l'ENSTA, campus de Brest, dans le cadre de la formation d'ingénieur généraliste (sous statut étudiant, FISE et en alternance, FIPA), ainsi que dans le Mastère spécialisé « Expert en énergies marines renouvelables » (MSEEMR). Ensemble représentant un total d'environ 117 heures équivalentes de TD sur les années universitaires 2023–2024 et 2024–2025.

— **Analyse des données spatiales** (TP) — FISE 1A.

— **Probabilités et statistiques** (TD) — FISE 1A.

— **Vibrations et dynamique des structures** (CM et TP) — FIPAMECA 2A.

— **Projet tutoré : « découverte des systèmes »** — FISE 1A.

— **Sociétés humaines et environnement** (CM et TD) — FISE 1A et 2A.

— **Écologie marine et perturbation des écosystèmes** (CM) — MSEEMR 1A.

TABLE DES MATIÈRES

Introduction **25**

Préambule - Positionnement de la thèse . . . 25
Les cétacés, espèces clés des écosystèmes marins . . . 28
Des mammifères singuliers . . . 28
Importance écologique et fonctionnelle . . . 32
Vulnérabilité et pressions anthropiques . . . 34
Monitorer les cétacés : le suivi par acoustique passive . . . 38
L'acoustique passive : une méthode puissante . . . 38
Les cétacés : des candidats idéaux . . . 41
Méthodologie et défis scientifiques du suivi par acoustique passive . . . 55
L'acquisition de la donnée . . . 55
Traitement des enregistrements . . . 64
Choix des métriques et méthodes statistiques . . . 68

1 Optimisation des stratégies de sous-échantillonnage **75**

1.1 Introduction . . . 75
1.1.1 État de l'art . . . 77
1.1.2 Les jeux de données utilisés . . . 82
1.2 Méthodologie : explorer l'effet du *duty cycle* par sous-échantillonnage artificiel 92
1.2.1 Principe général . . . 92
1.2.2 Construction des jeux de données sous-échantillonnés . . . 92
1.2.3 Analyses comparatives et approche statistique . . . 93
1.3 Effects of duty cycle on passive acoustic monitoring metrics : The case of blue whale songs . . . 94
1.4 Effects of duty cycle on passive acoustic monitoring metrics : The case of odontocetes vocalizations . . . 106
1.5 Discussion . . . 123
1.5.1 Aspect méthodologique . . . 123
1.5.2 Aspect écologique . . . 124
1.5.3 Perspectives méthodologiques . . . 124

2 Modélisation spatio-temporelle de la présence acoustique 127
2.1 Introduction . . . 127
2.1.1 État de l'art : Les cétacés de l'océan Indien Austral . . . 128
2.1.2 Les jeux de données utilisés . . . 150
2.1.3 Modéliser la présence acoustique : les outils statistiques . . . 155
2.1.4 Les principales techniques de modélisation de l'habitat . . . 165
2.2 Méthodologie : modélisation de la présence acoustique des baleines . . . 168
2.2.1 Étude 1 : Océan Indien Austral - Réseau OHASISBIO . . . 169
2.2.2 Étude 2 : Nord du canal du Mozambique - Réseau REVOSIMA . . 172
2.3 Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern indian ocean . . . 173
2.4 Ecological habitat characterization of baleen whales east of Mayotte : linking seasonal acoustic detections with environmental drivers . . . 190
2.5 Stage de Master : Combining passive and active acoustic approaches to study predator-prey interactions of blue whales in the southern Indian Ocean 208
2.6 Discussion . . . 212
2.6.1 Aspects méthodologiques . . . 212
2.6.2 Aspect écologique . . . 215

3 Structurer les suivis par acoustique passive 217
3.1 Introduction . . . 217
3.1.1 La DCSMM comme cadre de gestion . . . 218
3.1.2 Le suivi par acoustique passive comme données complémentaires . . 223
3.1.3 Les défis liés à l'absence de standardisation . . . 226
3.2 *Workshop* 2023 : consensus et recommandations . . . 227
3.3 Conclusion . . . 232

Conclusion générale 233

Bibliography 239

Annexe A 261

Annexe B 267

Annexe C 284

Annexe D 288

TABLE DES FIGURES

1 Histoire phylogénétique des cétacés. 29
2 Caractéristiques évolutives des cétacés. 31
3 Illustration des mécanismes regroupés sous le terme *whale pump*. 33
4 Superposition spatiale entre baleines et trafic maritime. 35
5 Captures mondiales de baleines bleues. 36
6 Mortalité des dauphins par capture accidentelle. 38
7 Techniques de surveillance des cétacés. 40
8 Étapes d'un suivi par acoustique passive. 41
9 Gammes de fréquences des sons émis par les cétacés. 43
10 Spectrogramme de vocalisations de dauphins de Risso. 44
11 Principales caractéristiques physiques d'un son. 45
12 Mécanismes de production des clics d'écholocalisation chez les odontocètes. 47
13 Spectrogramme de vocalisations de dauphins. 48
14 Spectrogramme de codas de cachalot. 48
15 Migrations des grandes baleines. 50
16 Production sonore chez les baleines à fanons. 51
17 Spectrogramme de Z-calls de baleine bleue Antarctique (ANT BW). 52
18 Spectrogramme de chants de baleine bleue pygmée du sud-est de l'océan Indien (SEIO PBW). 52
19 Spectrogramme de chants de baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW). 53
20 Spectrogramme de D-calls. 54
21 Spectrogramme de chants de baleine à bosse. 55
22 Propagation du son sous-marin. 59
23 Panel non exhaustif d'enregistreurs acoustiques disponibles. 61
24 Exemple de *duty cycle* selon le pourcentage d'enregistrement et la période d'écoute 64
25 Spectrogrammes illustrant le masquage acoustique et le chevauchement des signaux. 66
26 Exemple d'un spectrogramme annoté sur la plateforme `APLOSE`. 67

27 Effet du choix la fenêtre temporelle sur une même métrique. 70
28 Effet du choix de la métrique sur la même fenêtre temporelle. 70
29 Effet de la moyenne sur la même métrique et à fenêtre temporelle identique. 71

1.1 Localisation des sites du réseau Observatoire HydroAcoustique de la SISmicité et de la BIOdiversité (OHASISBIO). 83
1.2 Périodes d'enregistrement des sites du réseau OHASISBIO. 84
1.3 Schéma d'une ligne de mouillage du réseau OHASISBIO. 85
1.4 Profil de la vitesse du son et illustration du canal SOFAR. 86
1.5 Spectrogrammes de chants et D-calls de baleines bleues. 87
1.6 Spectrogrammes de sifflements et clics de delphinidés et marsouins. 89
1.7 Localisation des sites F, G du projet CETIROISE et du site IR en Irlande. 90

2.1 Cartes des limites de l'Océan Indien. 129
2.2 Principaux courant de l'Océan Indien Austral. 130
2.3 Plateau de Kerguelen et sa concentration en chlorophylle a. 132
2.4 Bathymétrie de la région autour de Madagascar, de Mayotte et du canal du Mozambique avec les courants principaux. 133
2.5 Trajectoires de baleines à bosse et bleues dans l'océan Indien. 136
2.6 Aires de répartition connues des populations acoustiques de baleines bleues dans l'Océan Indien. 137
2.7 Capture d'écran d'une vidéo montrant une baleine bleue se nourrissant de krill. 138
2.8 Scénarios migratoires possibles des baleines bleues antarctiques. 142
2.9 Capture d'écran d'une vidéo montrant une baleine bleue avec son petit. . . 144
2.10 Trajectoires de déplacement de deux baleines bleues pygmées d'Australie équipées de balises satellites. 145
2.11 Présence acoustique mensuelle des SWIO PBW dans l'océan Indien. 146
2.12 Spectrogramme de vocalisations de rorqual commun. 148
2.13 Spectrogramme de vocalisations de petit rorqual antarctique. 149
2.14 Localisation des sites de mouillage du réseau REVOSIMA. 152
2.15 Périodes d'enregistrement des sites du réseau MAHY. 152
2.16 Spectrogrammes des chants de quatre espèces de baleines. 154
2.17 Le concept de niche écologique. 156
2.18 Diagramme BAM de la niche écologique. 157

2.19 Structure hiérarchique de l'environnement marin selon l'échelle spatiale et temporelle. 163
2.20 Points de mesure d'acoustique active du programme Toward Hydroacoustics and Ecology of Mid-trophic levels in Indian and Southern Ocean (THEMISTO). 209
2.21 Principe de l'acoustique passive vs acoustique active. 210

3.1 11 descripteurs utilisés pour évaluer le BEE dans la DCSMM. 219
3.2 Façades maritimes françaises de la DCSMM. 220
3.3 Carte des emplacements des F-POD le long de la côte de la mer Noire. . . 224
3.4 Carte des emplacements des sites du projet CETIROISE dans le Parc Naturel Marin d'Iroise. 225
3.5 Participants du *workshop* à l'ECS 2023. 227

Liste des tableaux

1.1 Exemples d'études internationales de suivi par acoustique passive (PAM) de cétacés utilisant un *duty cycle* dans l'acquisition des données. 79
1.2 Exemples de rapports nationaux de PAM de cétacés utilisant un duty cycle dans l'acquisition. 80

2.1 Types de données issues du PAM et leur interprétation biologique. 160

Acronymes

ACC courant circumpolaire antarctique. 129

AMW petit rorqual Antarctique. 147–150, 153–155, 172, 213

ANT BW baleine bleue Antarctique. 17, 52, 66, 86, 87, 137, 139–141, 153–155, 170, 172, 208, 210, 211, 213, 267

BEE Bon État Écologique. 218–223, 226, 228, 232

Chla chlorophylle a. 131, 161, 170–173

DCSMM Directive Cadre Stratégie pour le Milieu Marin. 25, 28, 217–221, 226–228, 232

DPH nombre d'heures positives à la détection par jour. 82, 91, 92, 155, 160, 168, 211, 213

DPM nombre de minutes positives à la détection par jour. 82, 91, 92, 151, 160, 168, 213

ENSO El Niño–Southern Oscillation. 133

EOF fonction orthogonale empirique. 171, 212, 213, 234

FW rorqual commun. 66, 147–149, 153–155, 172, 213

GAM modèle additif généralisé. 161, 167, 168, 170, 172, 212–214, 234

GLM modèle linéaire généralisé. 167

IOD Dipôle de l'océan Indien. 133, 134

IWC Commission Baleinière Internationale. 134

MAHY Réseau Acoustique pour la surveillance du nouveau volcan sous-marin de Mayotte. 151, 152, 154, 172

MAYOBS Réseau d'Ocean Bottom Seismometers (OBS) pour la surveillance sismologique du volcan sous-marin de Mayotte. 151, 152, 154, 172

MSI concentration moyenne de glace de mer. 170, 171

NASC Nautical Area Scattering Coefficient. 209, 211

OHASISBIO Observatoire HydroAcoustique de la SISmicité et de la BIOdiversité. 18, 82–85, 87, 150, 151, 169, 208, 210, 211, 213

OI Océan Indien. 128–131, 133–140, 143, 146–149, 168, 169, 208, 212–216

PAM Passive Acoustic Monitoring. 20, 25–28, 39, 40, 42, 44, 55–58, 62, 64, 68, 73–77, 79, 80, 82, 123–125, 127, 128, 135, 136, 139, 159–161, 212, 214, 217, 218, 223, 226–228, 232–238

REVOSIMA Réseau de Surveillance Volcanologique et Sismologique de Mayotte. 151, 172, 213

SDM modèle de distribution d'espèces. 73, 158, 159, 161, 164

SEIO PBW baleine bleue pygmée du sud-est de l'océan Indien. 17, 52, 66, 86, 87, 137, 139, 143, 170, 213, 267

SST température de surface de la mer. 170–173

SWH hauteur significative des vagues. 172, 173

SWIO PBW baleine bleue pygmée du sud-ouest de l'océan Indien. 17, 18, 53, 86, 87, 137, 139, 145, 146, 149, 153–155, 170, 172, 213, 267

THEMISTO Toward Hydroacoustics and Ecology of Mid-trophic levels in Indian and Southern Ocean. 19, 128, 161, 208–211, 215

Les sons des océans, complémentaires aux spectrogrammes

Les cétacés et leur environnement sous-marin produisent des sons fascinants qui ne peuvent être pleinement appréciés à travers des spectrogrammes seuls. Pour permettre au lecteur de mieux s'immerger, il est possible d'écouter les sons présentés dans le manuscrit en même temps que la lecture. Chaque fichier audio est nommé selon la figure à laquelle il est associé, l'espèce et l'accélération éventuelle (e.g. Fig_1.5a_SEIOPBW_x20). Certains sons ont été accélérés ou ralentis afin de rendre audibles à l'oreille humaine des fréquences trop graves ou trop aiguës, tout en conservant les caractéristiques du signal. Pour profiter pleinement de cette expérience, scannez le QR code ci-dessous ou cliquez dessus pour écouter directement les enregistrements au fil de la lecture.



The sounds of the oceans, complementing the spectrograms

Cetaceans and their underwater environment produce fascinating sounds that cannot be fully appreciated through spectrograms alone. To allow the reader to better immerse themselves, it is possible to listen to the sounds presented in the manuscript while reading. Each audio file is named according to the corresponding figure, species, and any temporal acceleration (e.g. Fig_1.5a_SEIOPBW_x20). Some sounds have been sped up or slowed down to make frequencies that are too low or too high audible to the human ear, while preserving the characteristics of the signal. To fully enjoy this experience, scan the QR code above or click on it to listen to the recordings as the read through the manuscript.

INTRODUCTION

Préambule - Positionnement de la thèse

Surveiller les populations de cétacés, souvent mobiles et discrètes, impose de concevoir les efforts de suivi à l'échelle de grandes zones géographiques, parfois partagées entre plusieurs pays. Cette dimension transfrontalière a favorisé, dès les années 1970, le développement de programmes d'observation visuelle fondés sur des protocoles standardisés, partagés entre institutions et entre nations. Qu'ils soient réalisés depuis des navires, la côte ou les airs, ces suivis ont permis de produire des séries temporelles cohérentes et comparables à grande échelle. Ces données sont indispensables pour comprendre les dynamiques de populations de cétacés et construire des politiques de conservation coordonnées. Ces politiques s'inscrivent notamment dans le cadre de la Directive Cadre Stratégie pour le Milieu Marin (DCSMM), qui engage les États membres de l'Union européenne à surveiller et préserver le bon état écologique des milieux marins, incluant les populations de cétacés et leurs habitats.

Depuis les années 2000, le développement de la bioacoustique et du suivi par acoustique passive (PAM, pour *Passive Acoustic Monitoring*) a apporté une approche complémentaire aux dispositifs de suivi existants. Cette technique repose sur l'enregistrement passif des sons sous-marins, c'est-à-dire la captation des sons ambiants sans émission de signal ou interaction directe avec l'environnement, à l'aide d'enregistreurs autonomes fixes ou mobiles. L'objectif est de détecter et d'analyser dans ce paysage sonore sous-marin, les émissions sonores produites par les cétacés. Cette méthode permet une détection continue, autonome et non intrusive de leur présence acoustique. Elle offre ainsi des perspectives particulièrement adaptées au suivi d'espèces discrètes, présentes dans des zones profondes, isolées ou mal couvertes par les méthodes de suivi visuelles traditionnelles.

Les activités humaines et le changement global entraînent des phénomènes tels que le déplacement des populations, les modifications des aires de répartition, l'accroissement du

bruit sous-marin ou l'expansion des infrastructures marines. Face à ces effets, le besoin de dispositifs de suivi robustes, autonomes et durables devient de plus en plus pressant. Le PAM s'inscrit donc dans cette dynamique et constitue aujourd'hui un levier prometteur pour améliorer la couverture spatio-temporelle des suivis tout en réduisant la dépendance à l'effort humain direct.

Cependant, l'adoption croissante de cette méthode s'est accompagnée d'une grande diversité de pratiques directement liées aux objectifs d'étude, aux choix des espèces ciblées, aux choix des enregistreurs, aux modalités de déploiement, à la configuration des enregistreurs, aux niveaux d'automatisation du traitement des signaux ou encore aux choix des indicateurs retenus pour quantifier la présence acoustique. Cette hétérogénéité, bien qu'enrichissante, rend difficile la comparaison entre projets et freine la structuration d'un cadre de suivi standardisé à grande échelle. À cela s'ajoutent des contraintes logistiques ou techniques, parfois incontournables, qui peuvent influencer fortement les résultats et l'interprétation écologique qui en découle. Enfin, malgré la richesse des données produites, leur valorisation reste encore limitée par un manque de consensus sur les bonnes pratiques et par certaines difficultés à interpréter ces indicateurs acoustiques dans leur dimension biologique, écologique ou fonctionnelle.

C'est dans ce contexte que s'inscrit cette thèse, qui vise à proposer des outils et des recommandations pour améliorer la conception, l'interprétation et la standardisation du PAM réalisés à partir d'enregistreurs autonomes fixes.

Cette thèse s'articule autour de trois chapitres, chacun poursuivant des objectifs spécifiques :

Chapitre 1 – Optimisation des stratégies de sous-échantillonnage : Ce chapitre évalue l'impact des choix de *duty cycle*, c'est-à-dire l'alternance programmée de phases d'enregistrement et de phases de pause, sur les indicateurs de présence acoustique. L'objectif est de formuler des recommandations sur les compromis entre l'effort d'enregistrement, l'autonomie des enregistreurs et la robustesse des indicateurs de présence acoustique produits.

— **Objectif méthodologique :** évaluer comment différentes configurations de *duty cycle* influencent la précision et la robustesse des indicateurs de présence acoustique, en tenant compte des contraintes logistiques liées à l'autonomie des enregistreurs et

à la gestion des données.

— **Objectif écologique :** examiner les conséquences de ces choix techniques sur l'interprétation des dynamiques de présence, la richesse des signaux détectés, et plus largement sur la compréhension des comportements et des habitats des cétacés étudiés.

Cette double approche vise à proposer des recommandations pratiques et adaptées, permettant d'optimiser les protocoles PAM selon les caractéristiques biologiques des espèces ciblées et les objectifs scientifiques ou de gestion associés.

Chapitre 2 – Modélisation spatio-temporelle de la présence acoustique : Ce chapitre explore comment la modélisation spatio-temporelle peut enrichir l'analyse des données de PAM en reliant la présence acoustique des cétacés à des variables environnementales, afin de mieux caractériser leurs habitats et d'identifier les zones écologiquement importantes. Il s'appuie sur des séries temporelles longues issues de différents jeux de données, souvent l'unique moyen de suivre des espèces discrètes, mobiles et peu accessibles. Bien que ce type d'approche soit généralement appliqué à des données d'observation visuelle, son utilisation avec des données acoustiques reste peu courante. Le chapitre inclut également les travaux réalisés dans le cadre d'un stage de Master 2, qui ont permis de tester l'intégration de mesures in situ de biomasse de proies.

— **Objectif méthodologique :** proposer des approches statistiques robustes, adaptées aux spécificités des données PAM, et démontrer leur potentiel pour caractériser la présence acoustique des espèces dans l'espace et le temps.

— **Objectif écologique :** utiliser ces outils pour identifier les zones de présence acoustique, susceptibles de jouer un rôle fonctionnel pour les espèces (zones de reproduction, d'alimentation, de repos, etc.), et analyser les liens entre cette présence et les conditions environnementales.

Cette double approche vise à renforcer l'interprétation écologique des données PAM à travers des modèles généralisables, et à appuyer le développement de stratégies de suivi et de conservation fondées sur une meilleure compréhension spatio-temporelle des habitats des cétacés.

Chapitre 3 – Structuration d'un cadre standardisé pour les suivis PAM : Ce chapitre propose une synthèse issue d'un atelier de travail scientifique européen. Il identifie les leviers et obstacles à la standardisation du PAM et discute de leur intégration dans les cadres d'évaluation réglementaires, notamment la DCSMM.

À travers ces trois volets, cette thèse s'inscrit dans une réflexion plus large sur le manque de standardisation des suivis en PAM, tant sur le plan des protocoles d'acquisition que des méthodes d'analyse et d'interprétation. Elle propose, à travers différentes composantes du suivi, des éléments de structuration méthodologique visant à renforcer la rigueur scientifique, la comparabilité des résultats et leur utilité pour la conservation des cétacés. La suite de cette introduction présente les bases scientifiques et techniques nécessaires à la compréhension des enjeux du PAM chez les cétacés.

Les cétacés, espèces clés des écosystèmes marins

Des mammifères singuliers

Les cétacés jouent un rôle essentiel dans l'équilibre des écosystèmes marins. De plus, la fascination qu'ils exercent sur l'être humain en fait un groupe emblématique de la faune marine. Ils sont classés dans l'ordre des *Cetacea*, qui regroupe environ 90 espèces connues à ce jour, réparties en deux sous-ordres distincts : les *Mysticeti*, ou baleines à fanons, et les *Odontoceti*, comprenant les cétacés à dents tels que les dauphins, orques ou cachalots (Jefferson, Leatherwood et Webber 1993) (Fig. 1). Les cétacés partagent un ancêtre commun avec les hippopotames modernes, au sein du groupe des *Cetartiodactyla*, qui regroupe les cétacés et les ongulés à doigts pairs [1]. Cet ancêtre, probablement proche du genre fossile *Indohyus*, était un petit mammifère terrestre semi-aquatique vivant il y a environ 55 millions d'années (Thewissen, Cooper, Clementz et al. 2007). Il présentait déjà des adaptations comportementales à l'eau, facilitant la transition évolutive vers une vie entièrement aquatique. Cette évolution a donné naissance à une lignée de cétacés archaïques, tels que *Pakicetus* ou *Ambulocetus*, illustrant les étapes intermédiaires vers l'occupation totale du milieu marin (Gatesy et al. 2013 ; Thewissen, Cooper, George et al. 2009).

1. Ordre des mammifères comprenant par exemple les cervidés, bovins, porcins et hippopotames.

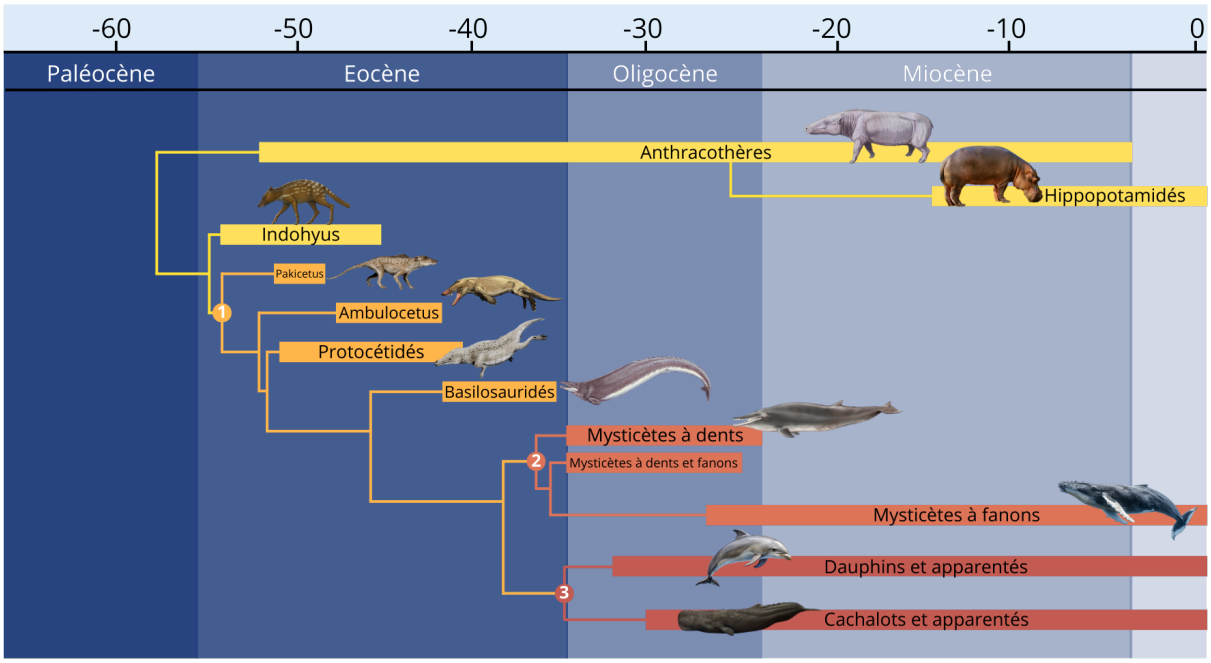


FIGURE 1 – Histoire phylogénétique des cétacés : 1-Cétacés, 2-Mysticètes, 3-Odontocètes. Figure adaptée et refaite de EVOLUTION BIOLOGIQUE 2022.

Ce retour à l'eau s'est accompagné d'adaptations morphologiques profondes, façonnant progressivement une silhouette parfaitement adaptée au milieu marin (Fig. 2). Certaines de ces transformations correspondent à des pertes évolutives, telles que la disparition des membres postérieurs visibles, des oreilles externes, des poils et, chez les mysticètes, des dents (GATESY et al. 2013). D'autres sont des innovations morphologiques uniques parmi les mammifères, comme l'apparition de la nageoire dorsale, du melon chez les odontocètes, des fanons chez les mysticètes, ou encore d'un crâne « télescopé » et asymétrique [2]. Cette modification facilite une respiration rapide et efficace en surface sans interrompre la nage. Par ailleurs, certaines espèces présentent aussi des sillons ou plis ventraux longeant la gorge pour faciliter l'évacuation de l'eau de la bouche après la prise de nourriture (GATESY et al. 2013).

L'ensemble du corps a ainsi été remodelé pour améliorer l'hydrodynamisme : les membres antérieurs se sont transformés en nageoires et la queue s'est dotée d'un puissant aileron horizontal assurant la propulsion (FISH et ROHR 1999).

2. Déplacement évolutif des os du crâne vers l'arrière et le haut, permettant notamment la migration des narines vers le sommet du crâne pour former l'évent.

Aujourd'hui, les cétacés occupent une grande diversité d'habitats marins, depuis les eaux froides des régions polaires jusqu'aux lagons tropicaux, et des plateaux continentaux aux abysses (Hooker et Gerber 2004). Leur écologie reflète une diversité comparable : certaines espèces effectuent de longues migrations saisonnières (Mate et al. 2011), tandis que d'autres restent fidèles à une région restreinte (Barco et al. 1999). Cette diversité morphologique, comportementale et écologique fait des cétacés un sujet d'étude à la fois riche et complexe, renforçant ainsi l'urgence de mieux comprendre leurs dynamiques dans un océan en perpétuelle transformation.

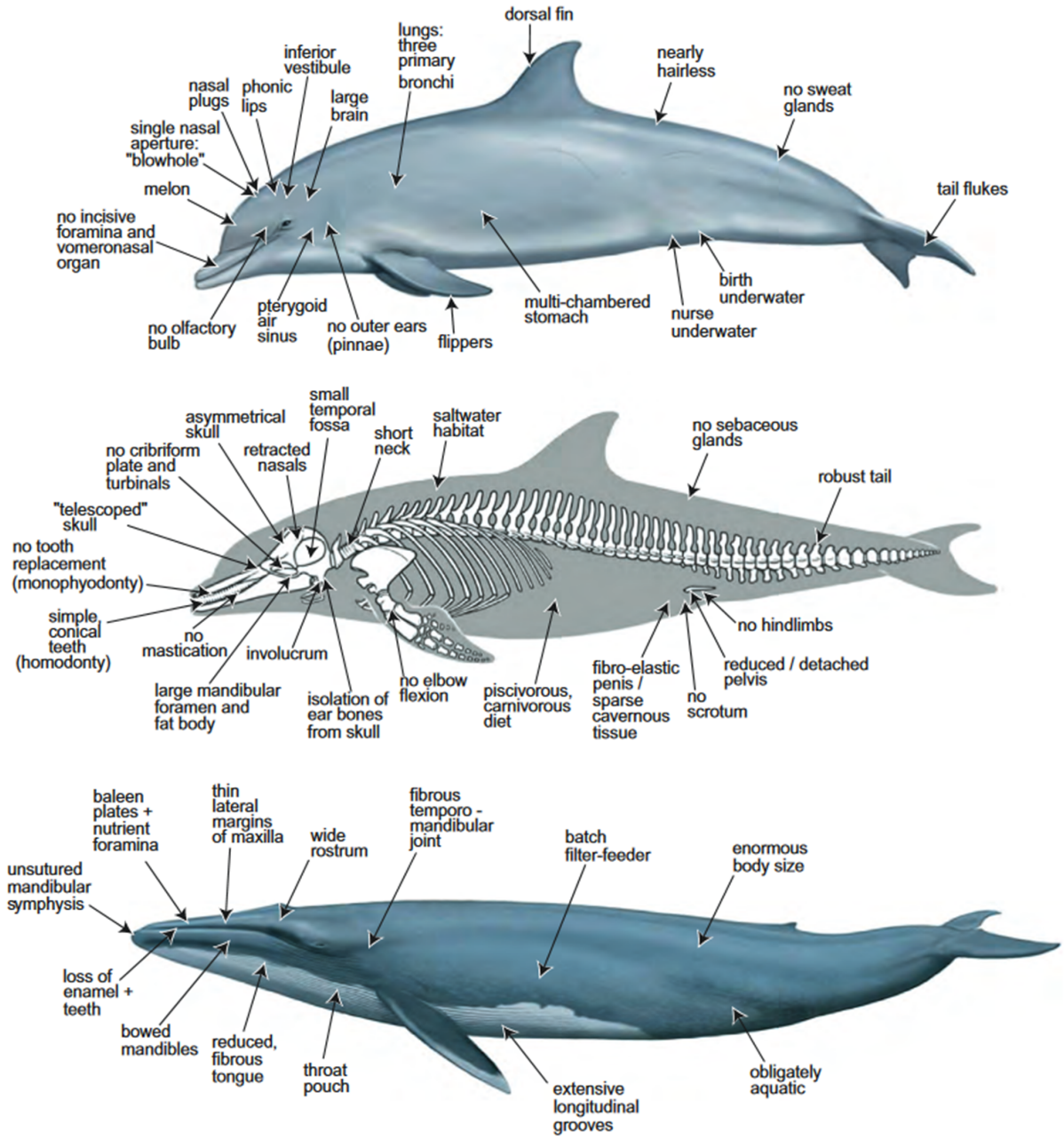


FIGURE 2 – Caractéristiques évolutives des cétacés (GATESY et al. 2013).

Importance écologique et fonctionnelle

Les cétacés jouent un rôle structurant dans les écosystèmes marins à travers divers mécanismes écologiques essentiels au maintien de la biodiversité, au fonctionnement biogéochimique des océans et à la résilience des réseaux trophiques (Roman et al. 2014). En tant que prédateurs supérieurs, les odontocètes, tels que les dauphins, les orques ou les cachalots, exercent une pression de prédation sur une large gamme de proies (e.g. céphalopodes, poissons, et parfois d'autres mammifères marins), modulant la distribution, l'abondance et le comportement de ces espèces. Ce contrôle du sommet vers la base peut engendrer des effets trophiques indirects, incluant une réduction de la pression sur les niveaux trophiques inférieurs (effets en cascade) (Estes et al. 2011). Concernant les mysticètes, leur consommation massive de zooplancton (e.g. krill, copépodes) influence directement la dynamique de ces populations et le transfert d'énergie entre niveaux trophiques. Une baleine bleue peut consommer jusqu'à 16 tonnes de krill par jour durant la saison d'alimentation (Hazen, Friedlaender et Goldbogen 2015), et il est estimé que les grandes baleines de l'océan Austral consomment au total environ 430 millions de tonnes de krill par an (Savoca et al. 2021).

Parmi les mécanismes fonctionnels majeurs, la *whale pump*[3] regroupe plusieurs processus écologiques clés comme illustré à la Fig. 3 :

(1) Fertilisation de surface : Les cétacés plongent dans la colonne d'eau pour se nourrir (e.g. krill, calmars, etc.) puis remontent régulièrement à la surface pour respirer. Ils y excrètent des matières riches en nutriments, notamment en azote, phosphore et fer, qui fertilisent la zone euphotique, favorisant ainsi la croissance du phytoplancton.

(2) Transport horizontal de nutriments : De nombreuses espèces de grandes baleines migrent entre des zones d'alimentation polaires riches en nutriments et des zones de reproduction tropicales pauvres en éléments nutritifs. Leurs rejets naturels dans ces zones pauvres enrichissent localement les eaux, stimulant la production primaire.

(3) Stockage de carbone biologique : Tout organisme vivant est constitué de carbone, stocké dans sa biomasse corporelle. En raison de leur grande taille et de leur longévité, les cétacés accumulent et retiennent ce carbone sur de longues périodes, ce qui en

3. Concept décrivant l'ensemble des processus par lesquels les baleines recyclent et redistribuent les nutriments, jouant ainsi un rôle dans la productivité primaire et le cycle du carbone.

fait des réservoirs temporaires importants. À la mort des cétacés, leur carbone est restitué à l'environnement, contribuant à la séquestration profonde et au recyclage des nutriments.

(4) **Séquestration profonde via les carcasses** : À leur mort, les carcasses de cétacés coulent vers les abysses, créant des habitats riches et temporaires (*whale falls*) qui peuvent soutenir des communautés benthiques pendant des décennies, tout en favorisant le stockage de carbone et l'enrichissement des écosystèmes profonds.

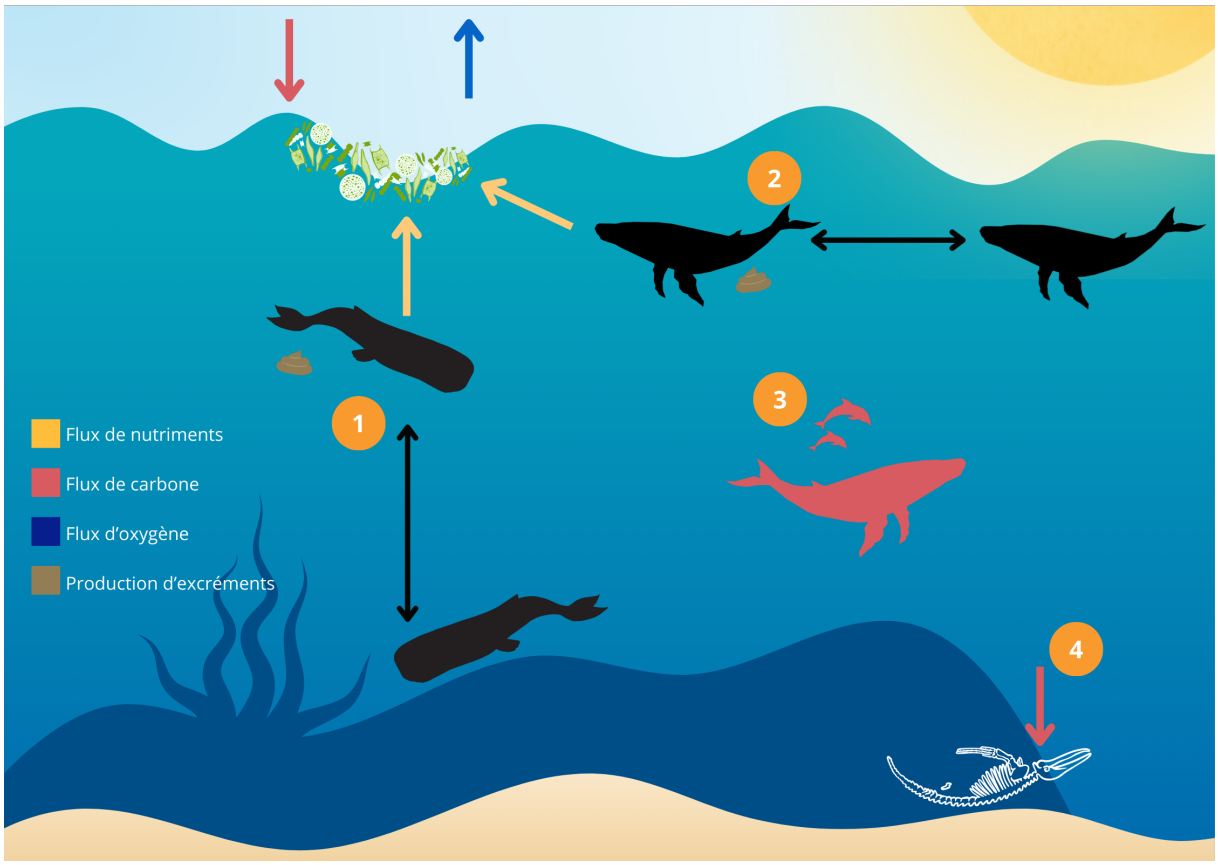


FIGURE 3 – Illustration des mécanismes regroupés sous le terme *whale pump* : (1) fertilisation verticale en surface par les excréments riches en nutriments ; (2) transport horizontal de nutriments via les migrations ; (3) stockage de carbone dans la biomasse des cétacés ; (4) séquestration profonde et apport nutritif par les carcasses. Figure adaptée et refaite de CHAMI 2019.

Ce mécanisme global rend les populations de cétacés indispensables à la dynamique des cycles biogéochimiques océaniques. À titre d'exemple, les cachalots, grands plongeurs, jouent un rôle crucial dans le recyclage du fer, un nutriment indispensable au phytoplancton

mais souvent en quantité limitée dans l'océan Austral. En excrétant du fer lors de leurs remontées en surface, ces cétacés fertilisent les eaux de surface, stimulant ainsi la croissance du phytoplancton, qui est constitué de micro-organismes photosynthétiques à la base de la chaîne alimentaire marine. Cette augmentation de la production primaire peut atteindre près de 40% localement. En favorisant la croissance du phytoplancton, le recyclage du fer par les cachalots accroît la capture de dioxyde de carbone (CO_2) atmosphérique, estimée à environ 240 000 tonnes de carbone supplémentaires par an dans la colonne d'eau (Lavery et al. 2010).

Vulnérabilité et pressions anthropiques

Les cétacés font face à une combinaison complexe de pressions anthropiques, dont les effets cumulés menacent gravement leur survie, altèrent leur comportement et compromettent les fonctions écologiques qu'ils remplissent dans les écosystèmes marins. Le changement climatique constitue l'une des principales menaces globales : l'augmentation des températures modifie la distribution verticale et horizontale des masses d'eau, affectant la disponibilité des proies et les routes migratoires saisonnières. Ces bouleversements impactent particulièrement les espèces dépendantes de la banquise, comme la baleine boréale (*Balaena mysticetus*), dont la distribution estivale est étroitement liée à la présence de glace de mer (Moore et Reeves 2018). De plus, la croissance exponentielle du trafic maritime mondial engendre une hausse des risques de collision avec les cétacés (Fig. 4). Au moins 20 000 grands cétacés sont tués chaque année par des navires (Vanderlaan et Taggart 2007 ; Schoeman, Patterson-Abrolat et Plön 2020 ; Nisi et al. 2024), une pression d'autant plus critique pour les espèces à faible effectif comme la baleine noire de l'Atlantique Nord (*Eubalaena glacialis*), dont la population totale est aujourd'hui inférieure à 360 individus (Knowlton et al. 2012). En Méditerranée, le rorqual commun (*Balaenoptera physalus*), seul mysticète régulièrement présent dans cette mer semi-fermée, est également exposé aux risques de collision. Le Sanctuaire Pelagos, dédié à la protection des mammifères marins [4], concentre une surveillance plus intensive : dans cette zone, la densité du trafic maritime, la pollution sonore chronique et la fragmentation de l'habitat compromettent la survie de l'espèce (Castellote, Clark et Lammers 2012 ; Panigada et al. 2006).

4. Zone maritime protégée créée en 1999 par la France, l'Italie et Monaco, couvrant environ 87 500 km² en Méditerranée, visant à préserver les mammifères marins.

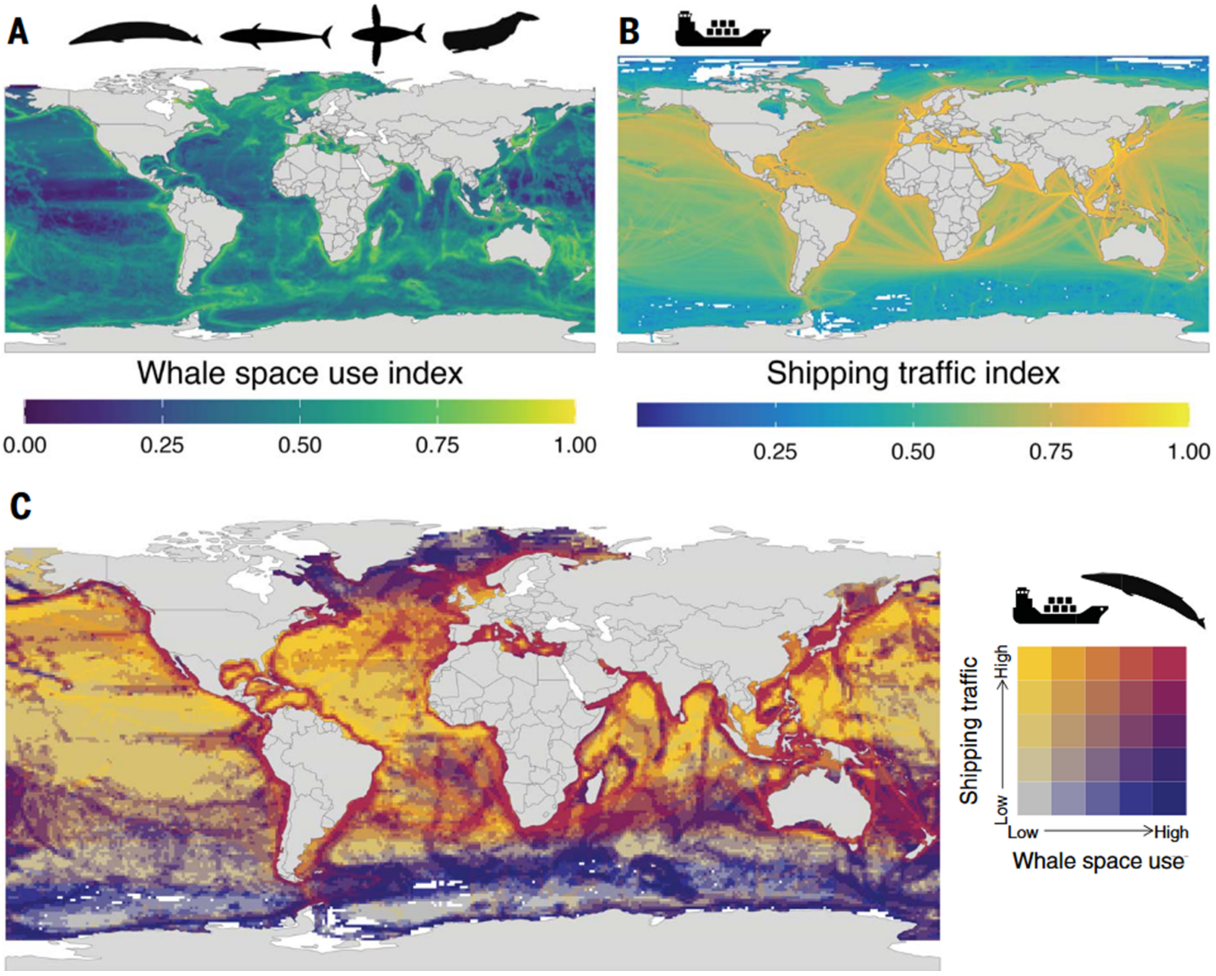


FIGURE 4 – Superposition spatiale entre les baleines et le trafic maritime. (A) Temps moyen occupé par les baleines dans la zone, regroupant les baleines bleues, les rorquals communs, les baleines à bosse et les cachalots. (B) Trafic maritime mondial pour les grands navires (> 300 tonnes brutes), basé sur les données AIS de 2017 à 2022. L'indice de trafic maritime pondère la densité des navires par leur vitesse, selon une échelle logarithmique normalisée entre 0 et 1. (C) Carte bivariée montrant l'intensité conjointe de l'utilisation de l'espace par les baleines et du trafic maritime pour chaque cellule de grille de 1° par 1° (NISI et al. 2024).

Les pollutions chimiques constituent une autre menace chronique. Les composés organochlorés (PCB, DDT) et les métaux lourds (mercure, plomb) s'accumulent dans les tissus adipeux des cétacés au fil du temps, notamment chez les espèces longévives situées en haut de la chaîne alimentaire comme les odontocètes. Ces polluants sont associés à des effets immunosuppresseurs, neurotoxiques et perturbateurs endocriniens susceptibles d'altérer la reproduction, la croissance et la réponse aux agents pathogènes (FOSSI et

al. 2004 ; Jepson et al. 2016). En parallèle, la pollution plastique a pris une ampleur considérable : une grande baleine filtrant plusieurs milliers de mètres cubes d'eau quotidiennement pourrait ingérer entre 5 et 10 millions de microplastiques par jour, en plus des menaces posées par les macroplastiques (e.g. ingestion, obstruction, contamination chimique indirecte) (Kahane-Rapport et al. 2022).

La chasse commerciale, bien que désormais limitée à quelques nations (Japon, Norvège, Islande), continue de causer des mortalités ciblées pouvant désorganiser les structures sociales et réduire la diversité génétique des populations locales (Parsons et Rose 2022). Ce phénomène a été particulièrement marqué chez les baleines bleues de l'océan Indien sud, qui ont été fortement décimées par la chasse industrielle au cours du XXe siècle (Fig. 5). Des centaines de milliers d'individus y ont été abattus, avec une prédominance pour les femelles gestantes en raison de leur grande taille et de leur densité en huile. Cette pression intense a laissé une empreinte durable : les populations de baleines bleues de cette région ne montrent aujourd'hui qu'un rétablissement très lent, et restent encore bien en deçà des niveaux pré-exploitation (Branch, Stafford et al. 2007 ; Branch, Monnahan et al. 2025).

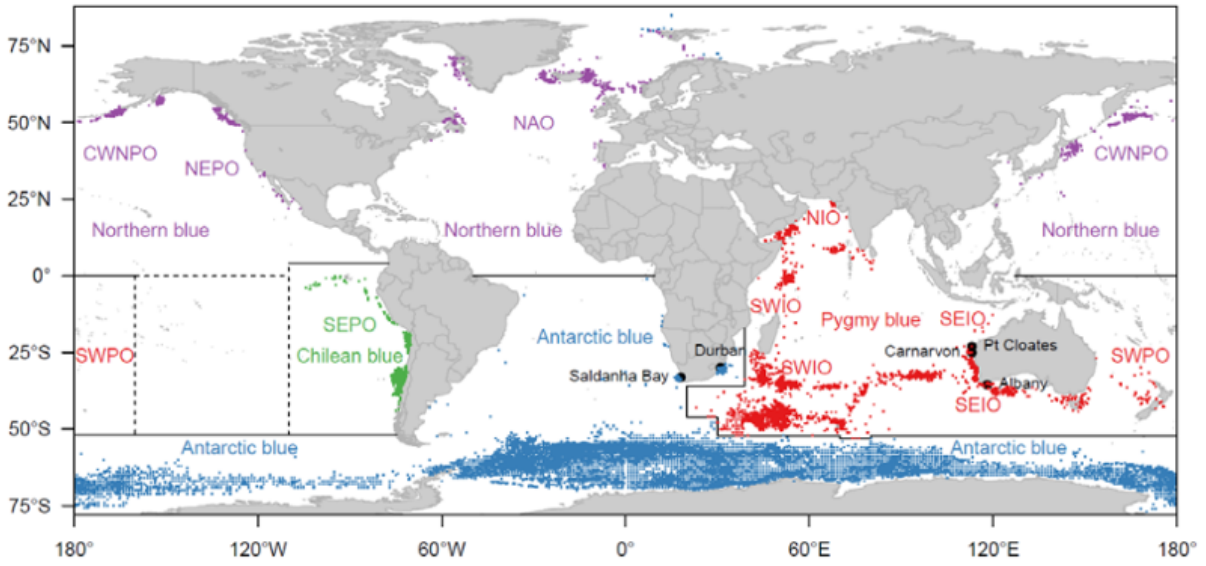


Figure 5 – Captures mondiales des différentes sous-espèces de baleines bleues (baleine bleue de l'hémisphère nord, baleine bleue du Chili, baleine bleue Antarctique et baleine bleue pygmée), avec les limites présumées des aires de capture de chaque sous-espèce indiquées en noir (Branch, Monnahan et al. 2025).

En parallèle, les captures accidentelles dans les engins de pêche, ou *bycatch*, représentent une pression anthropique majeure sur les populations de cétacés à l'échelle mondiale. Ces captures non ciblées surviennent principalement dans les filets maillants, les chaluts et autres dispositifs de pêche, piégeant involontairement des individus qui ne peuvent souvent pas se libérer, entraînant une mortalité significative. Les captures accidentelles affectent particulièrement les espèces côtières et celles évoluant dans des zones à forte activité de pêche, contribuant à une réduction directe de la taille des populations et perturbant leur structure démographique. Par exemple, des études ont documenté des taux élevés de mortalité par *bycatch* chez les dauphins communs (*Delphinus delphis*) dans le Golfe de Gascogne (Fig. 6) et chez les marsouins (*Phocoena phocoena*) en Atlantique Nord-Est (Peltier, Authier, Deaville et al. 2016 ; Peltier, Laran et al. 2024). Les mysticètes ne sont pas épargnés : la baleine franche de l'Atlantique Nord (*Eubalaena glacialis*) peut se retrouver accidentellement piégée dans les bouts des casiers à homard aux États-Unis, entraînant des blessures graves ou la mort (Bisack et Magnusson 2021). Malgré le développement de dispositifs de réduction des captures accidentelles (e.g. dispositifs acoustiques, modifications des engins), l'efficacité de ces mesures reste variable selon les contextes locaux et les types de pêcheries. Ainsi, les captures accidentelles constituent une menace persistante, qui s'ajoute aux autres pressions anthropiques, amplifiant les risques de déclin et compromettant la conservation des cétacés.

Enfin, le bruit anthropique sous-marin, généré par les moteurs de navires, les sonars militaires, ou la prospection sismique pour les hydrocarbures, perturbe profondément la communication acoustique des cétacés, qui repose sur des signaux sonores sophistiqués pour la navigation, la recherche alimentaire et les interactions sociales. Ces perturbations peuvent induire des modifications comportementales, un stress chronique, des abandons d'habitat, voire des échouages massifs lors d'expositions aiguës (Weilgart 2007 ; Williams et al. 2014). Le cumul de ces stress environnementaux se reflète dans la tendance observée à l'échelle globale : entre 1991 et 2021, la proportion d'espèces de cétacés classées comme menacées par l'UICN [5] est passée de 15% à 26%, traduisant un déclin généralisé de la viabilité des populations dans l'ensemble des bassins océaniques (O'Hara et al. 2019).

5. Union Internationale pour la Conservation de la Nature, organisation qui évalue le statut de conservation des espèces au niveau mondial à travers sa Liste rouge.

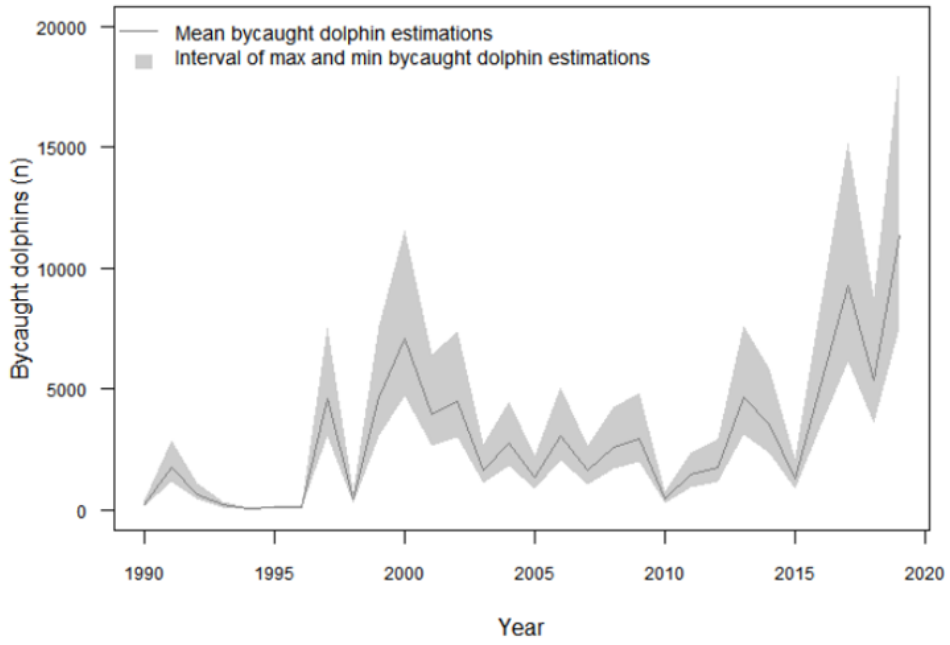


FIGURE 6 – Estimation de mortalité absolue de dauphin commun par capture accidentelle entre janvier 1990 et avril 2019 (PELTIER, AUTHIER, CAURANT et al. 2019).

Monitorer les cétacés : le suivi par acoustique passive

L'acoustique passive : une méthode puissante

Dans un contexte où les cétacés sont soumis à des pressions multiples, leur suivi devient indispensable pour documenter leur présence, leur comportement et l'évolution de leur distribution. Ces données sont essentielles pour évaluer leur état de conservation et mettre en place des mesures de gestion adaptées, notamment dans les zones sensibles ou peu accessibles. Historiquement, le suivi des cétacés reposait principalement sur des méthodes visuelles, telles que les observations depuis la côte, les campagnes en mer ou les relevés aériens (Fig. 7). Ces approches ont permis d'estimer l'abondance, d'identifier les aires de répartition ou encore de mieux comprendre les comportements des cétacés. Toutefois, elles présentent plusieurs limites : leur efficacité dépend des conditions météorologiques et de visibilité, elles nécessitent d'importants moyens logistiques et humains, et elles ne permettent généralement qu'un échantillonnage ponctuel dans le temps. Certaines espèces, discrètes ou au comportement cryptique, échappent par ailleurs facilement à la détection

visuelle. Ainsi, le PAM constitue une méthode particulièrement adaptée à l'étude des mammifères marins, dont la communication, la navigation, et les interactions sociales reposent largement sur l'utilisation du son. De plus, la propagation particulièrement efficace des ondes acoustiques dans l'eau permet la détection de leurs vocalisations sur de grandes distances. Le PAM exploite cette propriété en enregistrant le paysage sonore sous-marin de manière passive, c'est-à-dire sans émettre de signal ni interagir directement avec l'environnement ou les animaux. Cette méthode non intrusive permet de collecter les signaux sonores émis par les espèces présentes. L'analyse de ces signaux permet d'identifier les espèces produisant des sons, d'évaluer leur présence dans une zone, et de mieux comprendre leur distribution spatio-temporelle, sous réserve que leurs répertoires vocales [6] soient bien caractérisés et référencés.

Avantages et limites des suivis par acoustique passive

Avantages :

- Détecter les animaux même la nuit et par mauvais temps
- Détecter les animaux à une distance supérieure à celle de la vision
- Assurer une surveillance sur le long terme et une collecte continue de données
- Couvrir de larges zones géographiques
- Ne pas perturber les animaux (méthode non invasive)
- S'adapter aux espèces discrètes, rares ou difficiles à observer
- Être efficace dans les zones éloignées ou difficiles d'accès

Limites :

- Ne détecter que la présence acoustique, ce qui empêche de différencier l'absence d'animaux de la présence d'animaux silencieux
- Ne pas fournir d'estimation du nombre d'individus présents
- Ne pas pouvoir identifier précisément la source de certains sons
- Générer de grandes quantités de données nécessitant un traitement et un archivage lourds

6. Ensemble des types de sons émis par une espèce et aux contextes comportementaux associés, permettant l'identification et l'étude du comportement.

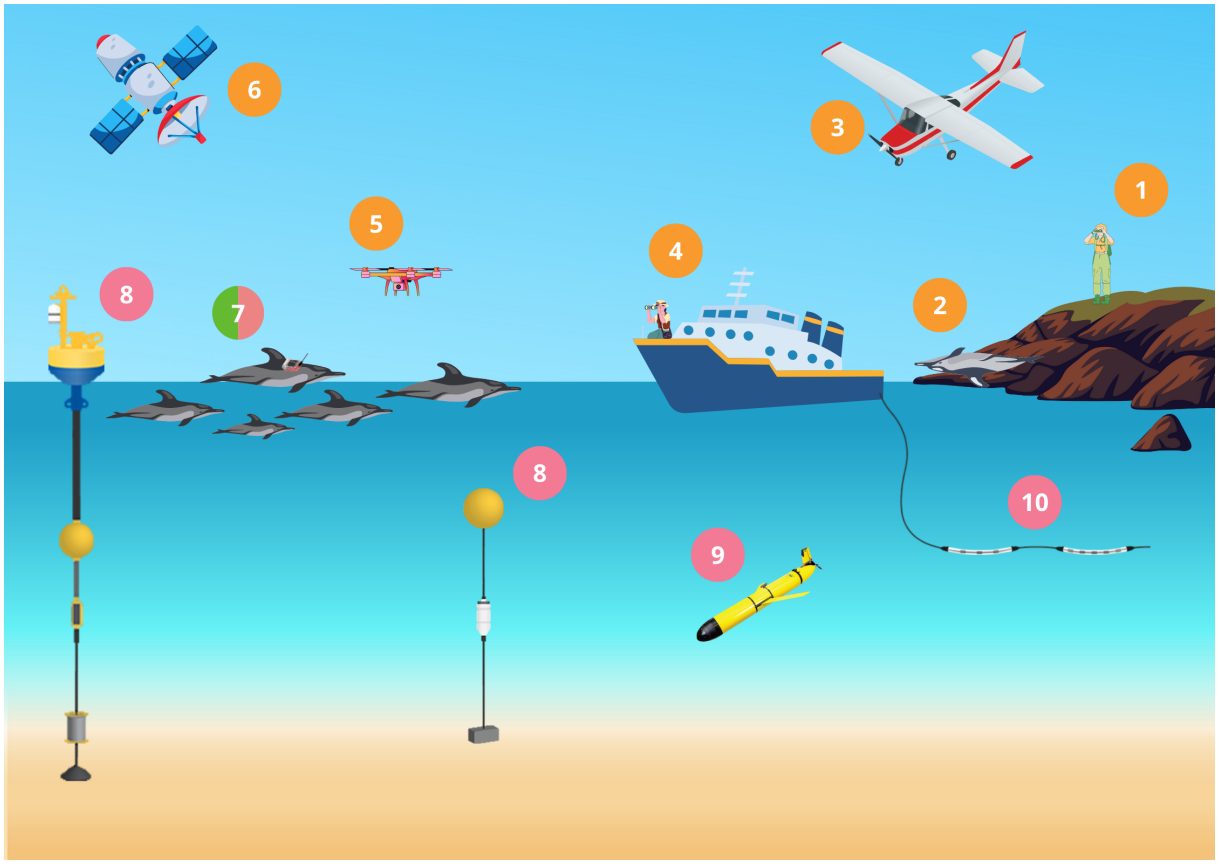


FIGURE 7 – Différentes techniques de surveillance des cétacés, codées par couleur : en orange, les suivis visuels (1) observation depuis la côte, (2) grâce aux échouages, (3) depuis un avion, (4) depuis un bateau, (5) par drone, (6) par image satellite ; en vert, les suivis par télémétrie (7) ; en rose, les suivis acoustiques (7) enregistrement avec des balises acoustiques (8) enregistrement avec enregistreurs fixe, (9) enregistrement avec enregistreurs mobile type sous-marin autonome, (10) enregistrement avec enregistreurs mobile type hydrophone tracté.

Le PAM se décompose classiquement en trois étapes : (1) la collecte des données acoustiques à l'aide d'enregistreurs autonomes équipés d'hydrophones, dont les caractéristiques sont choisies en fonction des espèces ciblées et de l'environnement ; (2) l'analyse de ces enregistrements, via de l'annotation manuelle et/ou l'utilisation d'algorithmes de détection automatique pour identifier les vocalisations et les espèces présentes ; et (3) l'interprétation biologique et écologique des résultats, pour mieux comprendre leur distribution spatiale et temporelle. Ces trois phases seront développées dans la section «Méthodologie et défis scientifiques du suivi par acoustique passive», en soulignant leurs objectifs, leurs limites et les défis méthodologiques associés.

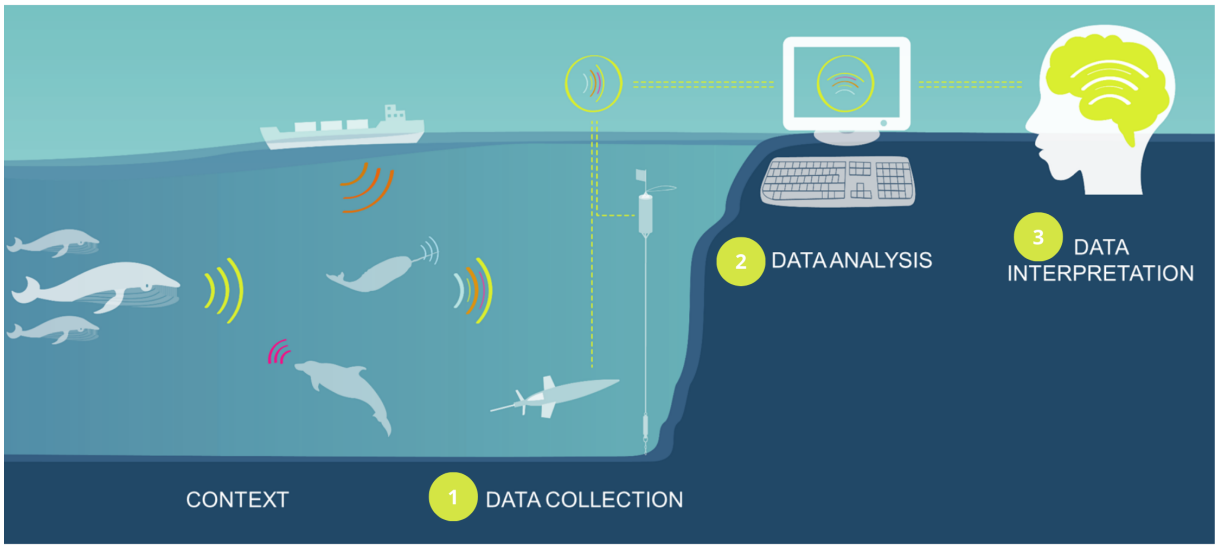


FIGURE 8 – Étapes d'un suivi par acoustique passive.

Les cétacés : des candidats idéaux

Les ondes acoustiques se propagent dans l'eau environ cinq fois plus vite que dans l'air, avec une atténuation relativement faible sur de grandes distances. Cette propriété physique fait du son le vecteur privilégié pour la communication et la perception sous-marine, notamment dans un milieu où la lumière est rapidement absorbée par la colonne d'eau, limitant fortement la vision au-delà de quelques dizaines de mètres. En profondeur, les capacités visuelles des cétacés sont donc réduites, même si leur vision reste utile en surface. Ils disposent d'une vision monochromatique, avec une capacité restreinte à percevoir les couleurs, adaptée aux conditions de faible luminosité grâce à une forte densité de bâtonnets (photorécepteurs sensibles à la lumière faible) ainsi qu'à la présence d'un *tapetum fibrosum*, une membrane réfléchissante située derrière la rétine qui amplifie la lumière reçue (GRIEBEL et PEICHL 2003 ; MASS et SUPIN 2007). De plus, leurs yeux sont situés sur les côtés de la tête, ce qui leur offre un large champ visuel mais limite leur vision binoculaire, réduisant leur perception de la profondeur (MOBLEY JR et HELWEG 1990). Par conséquent, les cétacés utilisent principalement leur audition pour percevoir leur environnement, mais aussi l'émission de sons pour communiquer, s'orienter ou chasser (TYACK et MILLER 2002). Les cétacés émettent des vocalisations, c'est-à-dire des signaux acoustiques intenses, réguliers et structurés, souvent spécifiques à chaque espèce et formant le répertoire vocal.

Ces vocalisations remplissent plusieurs fonctions essentielles : elles jouent un rôle dans la communication sociale et la reproduction et permettent, chez les odontocètes, la navigation et la localisation des proies via l'écholocation (Tyack 2000 ; Au et Hastings 2008).

Chez les cétacés, cette diversité acoustique est particulièrement marquée et reflète leur grande variabilité morphologique, écologique et sociale. La fréquence, la structure et la complexité de ces répertoires varient notamment en fonction de la taille corporelle, du régime alimentaire et de l'organisation sociale des espèces. Ainsi, les cétacés de petite taille émettent généralement des sons à haute fréquence, tandis que les plus grandes espèces produisent des vocalisations plus graves (Fig. 9) (Mellinger, Stafford et al. 2007). Ces répertoires vocaux permettent non seulement d'identifier les espèces à partir de leurs signatures sonores, mais aussi d'étudier leur comportement, leurs interactions sociales, ainsi que leur présence dans différentes zones. Ils offrent ainsi des informations essentielles sur la distribution spatiale et saisonnière des cétacés, leurs schémas migratoires, et plus largement sur leur écologie spatio-temporelle. Cette diversité acoustique traduit une adaptation fine entre fréquence d'émission et portée de propagation sous-marine, un aspect essentiel à comprendre pour interpréter correctement les données issues du PAM.

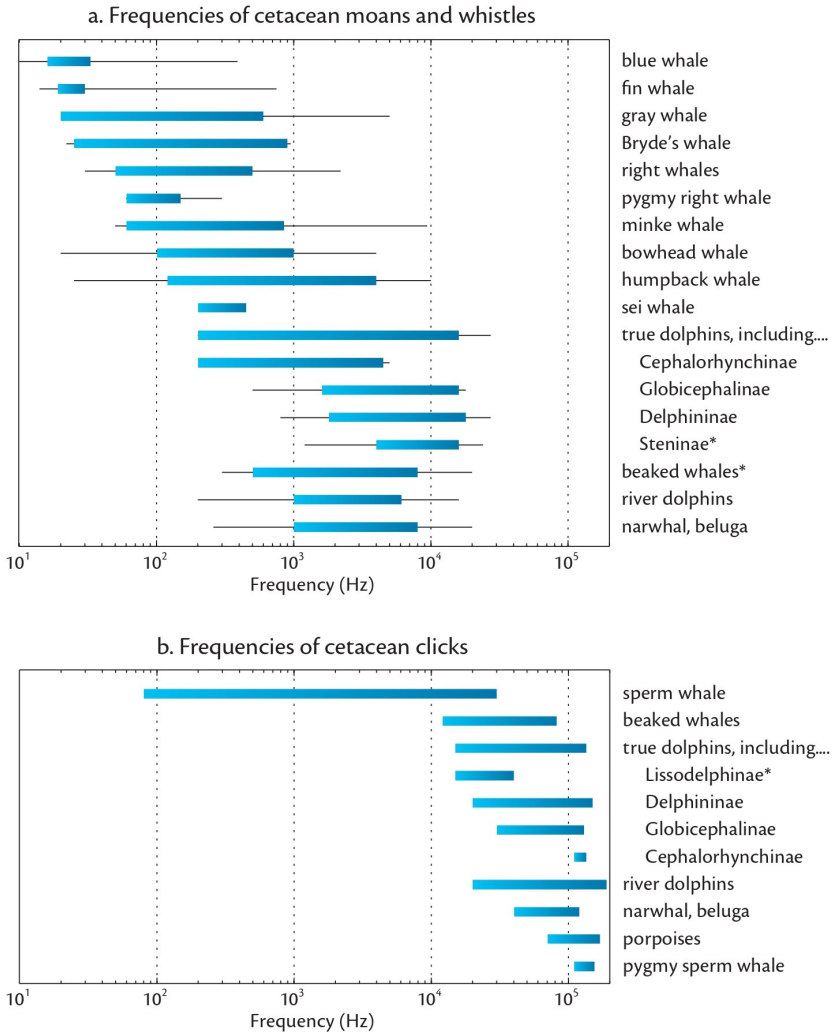


FIGURE 9 – Gammes de fréquences utilisées par les différentes espèces de cétacés. Les mysticètes sont listés par espèce, tandis que les odontocètes sont regroupés par famille. a) Sons tonaux («grondements» et «sifflements») émis par les mysticètes et odontocètes. b) Clics d'écholocation produits uniquement par les odontocètes. L'astérisque (*) indique que la fréquence maximale est inconnue en raison des limites du matériel d'enregistrement (MELLINGER, STAFFORD et al. 2007).

La portée de propagation d'une onde sonore dépend donc à la fois de ses propres caractéristiques physiques et des propriétés du milieu marin, qui influencent la manière dont ce son se propage. Ces aspects sont régis par des phénomènes physiques développés dans la suite de la thèse (voir encadré sur la propagation du son dans la section « Contexte environnemental et contraintes logistiques»).

De ce fait :

— Les **basses fréquences** (< 100 Hz), peu absorbées et peu diffusées, peuvent parcourir plusieurs centaines de kilomètres.

— Les **fréquences intermédiaires** (1–20 kHz) se propagent sur quelques kilomètres.

— Les **hautes fréquences** (> 100 kHz), fortement absorbées et sensibles à la diffusion, ne franchissent souvent que quelques centaines de mètres.

L'outil central de cette approche, permettant de représenter et d'interpréter visuellement les signaux enregistrés, est le spectrogramme, dont le principe est présenté dans l'encadré suivant.

Qu'est-ce qu'un spectrogramme ?

Un **spectrogramme** est une représentation visuelle d'un son qui affiche la répartition de l'énergie en fonction du temps et de la fréquence. L'énergie du signal à chaque valeur de temps (axe horizontal) et bande de fréquence (axe vertical) est représentée par une couleur selon une échelle (généralement un gradient), telle que visible Fig. 10.

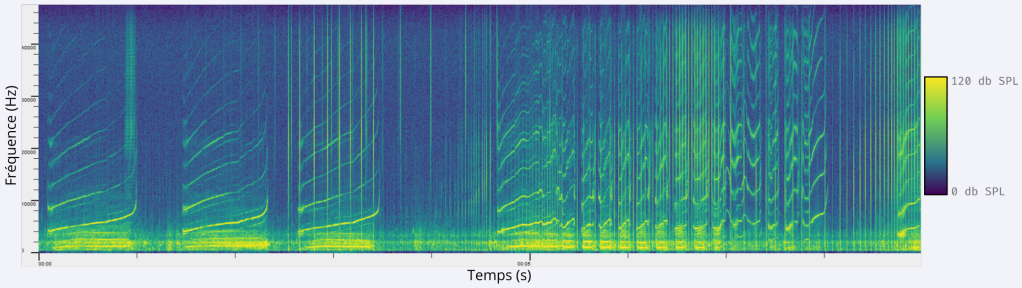


FIGURE 10 – Spectrogramme illustrant des vocalisations de dauphins de Risso sur 10 secondes. Enregistrement réalisé le 15 Juin 2022 en Mer Méditerranée par l'association MIRACETI.

Dans le PAM des cétacés, notamment en méthodes de détection manuelle, le spectrogramme est l'outil de référence. Il permet de visualiser facilement les vocalisations, de repérer leur forme, leur durée et leurs caractéristiques fréquentielles, ce qui est souvent beaucoup plus efficace que d'écouter les enregistrements bruts.

Caractéristiques physiques d'une onde sonore

Un son est une vibration mécanique longitudinale qui se propage à travers un milieu (comme l'air, l'eau ou un solide) sous forme d'ondes acoustiques. Il est caractérisé par plusieurs grandeurs :

- **Période** (T, en s) : intervalle de temps nécessaire pour qu'un point fixe du milieu dans lequel l'onde se propage (air, eau, solide) revienne à un état vibratoire identique au précédent.
- **Longueur d'onde** (λ, en m) : distance parcourue par l'onde pendant une période.
- **Fréquence** (f, en Hz) : nombre d'oscillations par seconde ; elle est l'inverse de la période T ($f = 1/T$) et détermine la tonalité (grave ou aiguë).
- **Amplitude** (A, en Pa) : correspond à l'ampleur de la variation de la pression du milieu, appelée *pression acoustique*.
- **Niveau de pression acoustique (SPL)** : le niveau sonore est souvent exprimé en décibels (dB) avec une échelle logarithmique définie par :

$$\mathrm{SPL} = 20 \log_{10} \left(\frac{p}{p_0} \right)$$

où p est la pression acoustique mesurée et p_0 la pression de référence (en acoustique sous-marine, elle est de $20\,\mu$Pa). L'utilisation d'une échelle logarithmique permet de représenter de façon lisible des variations de pression acoustique pouvant s'étendre sur une large échelle de valeurs, facilitant ainsi la comparaison des niveaux sonores.

- **Vitesse de propagation** (en m/s) : dépend du milieu (dans l'océan elle dépend de la température, de la salinité et de la profondeur et est d'environ 1500 m/s).
- **Durée** (en s) : temps pendant lequel le milieu est perturbé.

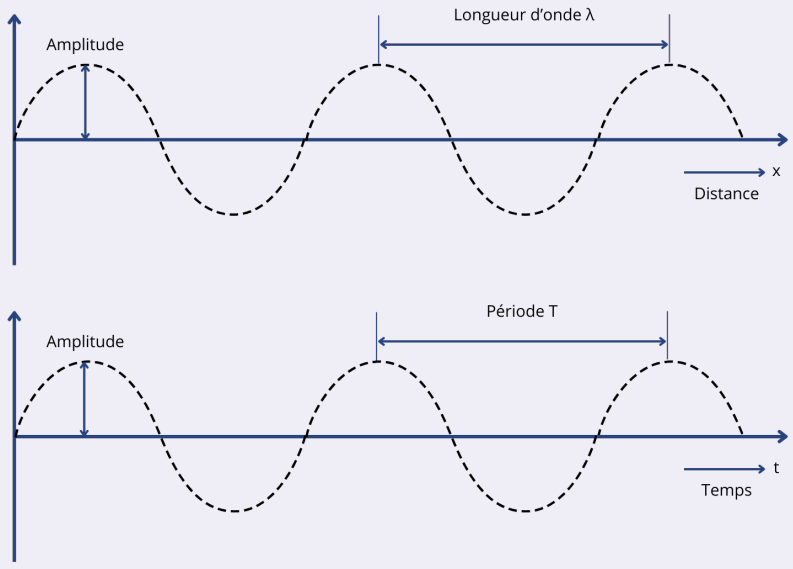


FIGURE 11 – Principales caractéristiques physiques d'un son.

Odontocètes : l'écholocalisation et la communication sociale

Les odontocètes, tels que les dauphins, les marsouins, les globicéphales, les cachalots ou les orques, possèdent un répertoire vocal riche et diversifié, influencé par deux besoins majeurs. D'une part, la vie sociale : ces cétacés vivent souvent en groupes dont la taille et la composition varient selon les espèces, nécessitant des mécanismes de communication sophistiqués pour maintenir les liens sociaux, coordonner les activités et reconnaître les individus à distance. D'autre part, la chasse et la navigation : en tant que prédateurs actifs, ils ont développé des systèmes d'écholocalisation comparables à un sonar, leur permettant de détecter et localiser leurs proies dans des environnements où la vision est limitée, comme les eaux turbides, les habitats profonds ou lors d'activités nocturnes. Cette combinaison de contraintes sociales et alimentaires explique la grande diversité et la complexité des vocalisations observées chez les odontocètes, de manière analogue à certaines chauves-souris qui utilisent l'écholocalisation pour chasser efficacement dans l'obscurité. Cette dépendance forte a conduit au développement de systèmes vocaux sophistiqués, souvent spécifiques à chaque espèce, groupe social ou individu (Janik, Sayigh et Wells 2006 ; Tyack 2000). Leur vocalisation comprend deux grandes classes de signaux :

(1) Des clics d'écholocation, impulsions brèves de large bande (10–200 µs, 20 à plus de 150 kHz), utilisés pour la navigation et la chasse. Ils sont produits dans la tête de l'animal par des structures spécialisées situées autour du passage nasal supérieur (Fig. 12). L'air expiré des poumons traverse le passage nasal, provoquant la vibration des *phonic lips*, générant ainsi des impulsions acoustiques. Ces sons sont ensuite focalisés par le melon, un organe graisseux qui fonctionne comme une lentille acoustique, projetant les clics sous forme de faisceaux directionnels vers l'environnement. Les clics permettent une détection et une localisation fines des proies ou obstacles à courte distance (quelques centaines de mètres à un kilomètre). Lorsqu'un objet ou une proie reflète ces ondes, les échos sont captés par la cavité graisseuse de la mâchoire inférieure, qui transmet les vibrations aux oreilles internes. Le cerveau analyse ensuite ces signaux pour créer une image acoustique précise de l'environnement, permettant à l'animal de localiser et d'identifier les obstacles et proies à courte distance (McKenna et al. 2012 ; Frankel 2009).

(2) Des signaux de communication sociale, notamment des sifflements (1 à 20 kHz), modulés en fréquence et porteurs d'informations individuelles ou contextuelles.

Ces sons sont produits par les mêmes structures nasales et *phonic lips* que pour les clics, mais sous forme de sons continus et modulés. Le melon focalise également ces sifflements, qui, grâce à leurs fréquences plus basses, peuvent voyager sur plusieurs kilomètres, facilitant la communication entre individus éloignés (Madsen et al. 2012 ; Au 1993 ; Au et Benoit-Bird 2003).

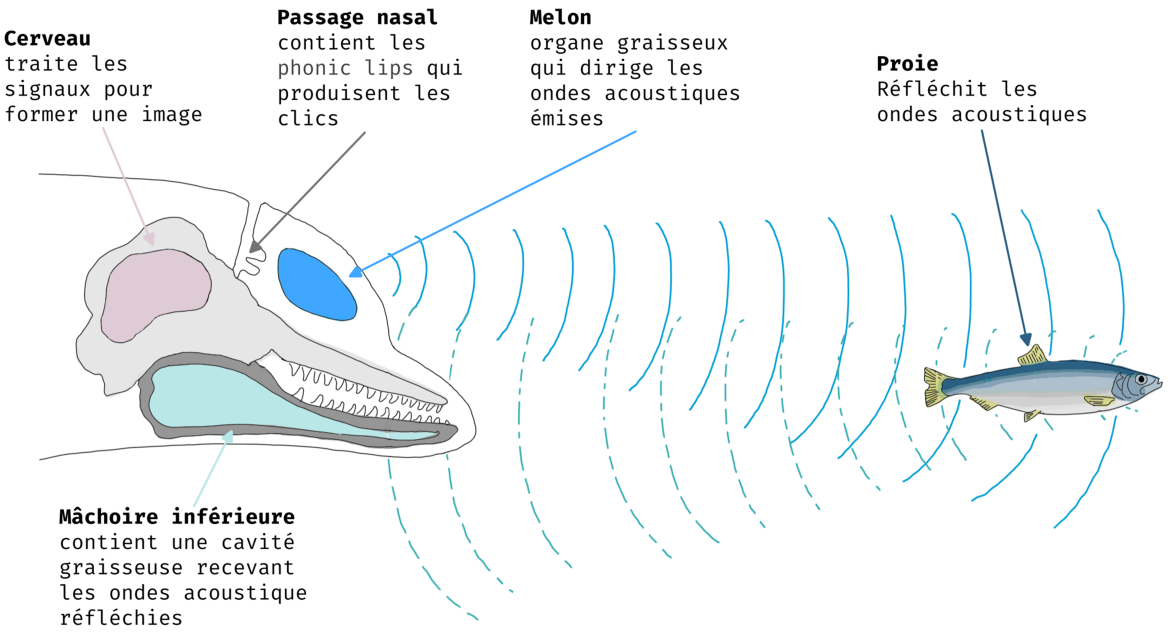


Figure 12 – Mécanismes de production des clics d'écholocalisation chez les odontocètes. Figure traduite de Minton et al. 2021.

La diversité de ces signaux reflète directement l'écologie sociale et comportementale des espèces. Par exemple, les grands dauphins *Tursiops truncatus*, espèce emblématique des systèmes dits « fission-fusion » [7]. Une telle organisation nécessite des mécanismes fins de reconnaissance individuelle et de cohésion à distance. Pour y répondre, les individus développent des sifflements signatures, vocalisations individualisées, jouant le rôle d'identifiants acoustiques uniques (Janik, Sayigh et Wells 2006). Ces vocalisations permettent aux individus de maintenir des liens sociaux durables malgré la séparation spatiale.

7. Organisation sociale dans laquelle les groupes se forment, se séparent et se recomposent fréquemment selon les activités, la disponibilité des ressources ou les interactions sociales.

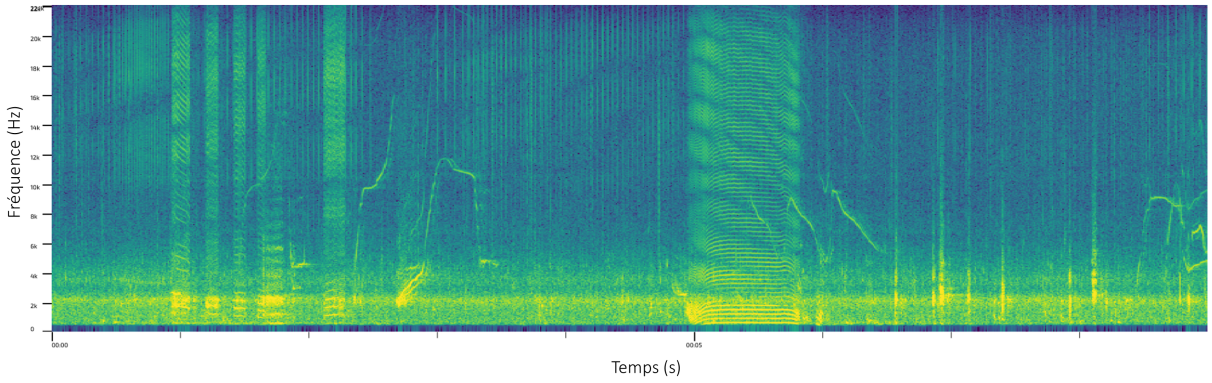


FIGURE 13 – Spectrogramme illustrant des vocalisations de grands dauphins (clics, buzzs et sifflements) sur 10 secondes. Enregistrement réalisé le 10 Septembre 2020 dans la mer Méditerranée par l'association MIRACETI.

À l'opposé, les marsouins communs (*Phocoena phocoena*) sont plus discrets et solitaires, formant rarement de grands groupes. Leur activité acoustique est dominée par des clics d'écholocation à très haute fréquence (> 130 kHz), bien adaptés à la détection fine de proies dans des habitats côtiers turbides, mais difficilement détectables à distance. En raison de leur comportement furtif et de leur taille réduite, ces animaux échappent fréquemment aux méthodes de suivi visuel, rendant les approches acoustiques essentielles pour leur étude.

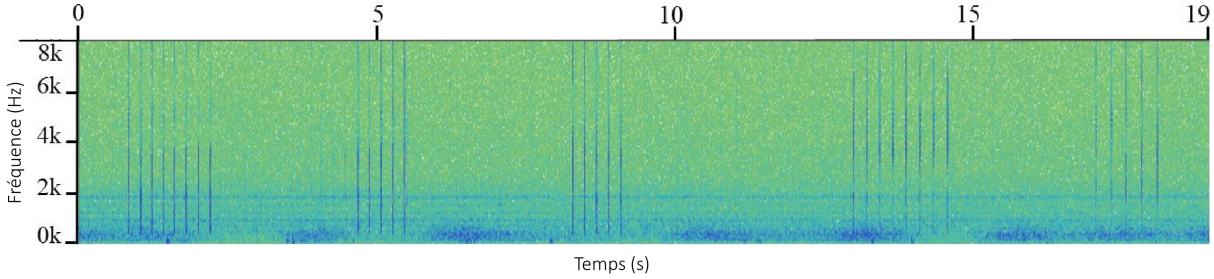


FIGURE 14 – Spectrogramme illustrant cinq codas produits par un cachalot d'un clan sur 19 secondes. De gauche à droite, les codas contiennent 8, 5, 5, 8, et 5 clics respectivement. Enregistrement réalisé le 30 mai 1995 au large des îles Galápagos (HERSH et al. 2022).

Certaines espèces comme les orques (*Orcinus orca*) ou les cachalots (*Physeter macrocephalus*) vivent dans des groupes sociaux stables, généralement fondés sur des liens familiaux, et développent des répertoires vocaux propres à chaque groupe, traduisant de véritables cultures acoustiques (RENDELL et WHITEHEAD 2003 ; DEECKE, FORD et SPONG 2000). Chez les cachalots, par exemple, la culture acoustique se manifeste principalement par la

production de codas, séquences rythmiques de clics utilisées lors des interactions sociales (Fig. 14). Ces vocalisations sont apprises et transmises culturellement entre individus d'un même groupe, indépendamment des liens génétiques ou de la proximité géographique (Hersh et al. 2022). Chaque clan social possède un répertoire distinct de codas, servant d'identifiant acoustique collectif, comparable à un dialecte.

Mysticètes : les basses fréquences pour la communication longue distance

Les mysticètes, tels que la baleine bleue (*Balaenoptera musculus*) ou la baleine franche (*Eubalaena glacialis*), adoptent une stratégie écologique fondée sur la filtration de petites proies. Exclusivement planctonophages, ils se nourrissent de krill, de zooplancton, de copépodes et parfois de petits poissons, à l'aide de techniques spécialisées : engouffrement chez les rorquals, nage continue chez les baleines franches, ou agitation du sédiment chez les baleines grises (Goldbogen, Cade et al. 2017 ; Mayo et Marx 1990 ; Nerini et Oliver 1983). Contrairement aux odontocètes, ils ne sont donc pas des chasseurs actifs et ne disposent pas de système d'écholocalisation ; ils ne produisent pas de clics.

En revanche, ces espèces émettent principalement des sons à basses fréquences ($<$ 10 kHz), dont la propagation peut s'étendre sur plusieurs dizaines, voire centaines de kilomètres dans l'océan (Širović, Hildebrand et Wiggins 2007 ; Samaran, Adam et Guinet 2010a).

Ce mode de communication s'inscrit dans un contexte écologique marqué par de vastes migrations saisonnières. La plupart des mysticètes alternent entre des zones d'alimentation estivales situées aux hautes latitudes, riches en ressources, et des zones de reproduction hivernales localisées dans les régions tropicales ou subtropicales (Fig. 15). Ces déplacements peuvent atteindre plusieurs milliers de kilomètres, constituant parmi les plus longues migrations du règne animal (Szesciorka et al. 2020 ; Horton et al. 2022).

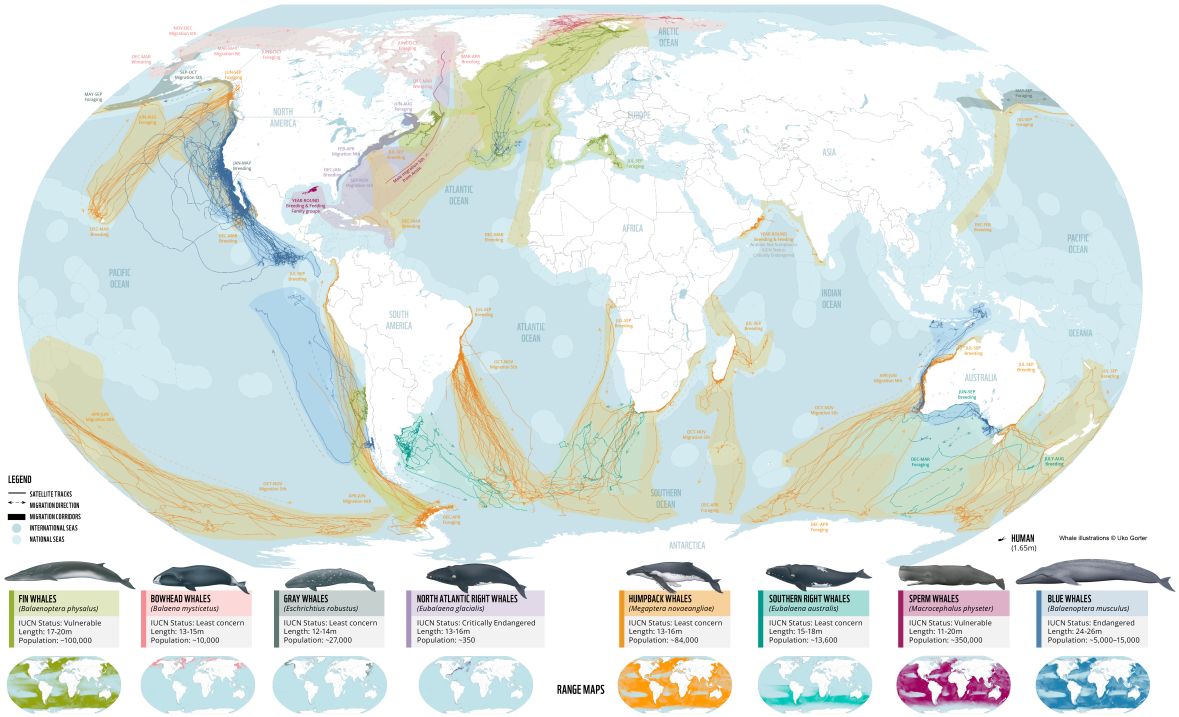


FIGURE 15 – Migrations saisonnières des principales espèces de grandes baleines entre zones de reproduction (latitudes basses) et zones d'alimentation (latitudes hautes) (JOHNSON et al. 2022).

Dans ce cadre, les vocalisations remplissent des fonctions essentielles. Elles permettent le maintien des contacts sociaux à grande distance en assurant une connectivité entre individus isolés, la coordination des déplacements collectifs au cours des migrations, ainsi que l'interaction entre individus dispersés dans des habitats vastes et ouverts. Selon l'espèce et le contexte, elles peuvent également contribuer à la synchronisation des comportements collectifs (e.g. lors de la chasse coopérative chez les baleines à bosse). Les chants sont des séquences sonores complexes et organisées, composées de motifs tonaux stéréotypés et répétés. Selon l'hypothèse retenue actuellement, ils sont principalement produits par les mâles et pourraient jouer un rôle central dans l'attraction des partenaires reproducteurs, la délimitation des territoires, ainsi que dans la structuration des échanges sociaux (CUMMINGS et THOMPSON 1971 ; MANN 2000 ; PERRIN, WÜRSIG et THEWISSEN 2009).

La production de ces chants repose sur des structures laryngées spécialisées (Fig. 16). Contrairement aux odontocètes qui utilisent les *phonic lips* dans le passage nasal, les baleines à fanons disposent de plis vocaux en forme de U situés dans le larynx, soutenus par les cartilages aryténoïdes. Ces plis vibrent sous l'effet du flux d'air expiré entre les

poumons et le sac laryngé (tissu orange), générant des sons à basses fréquences puissants et modulés (Reidenberg et Laitman 2007 ; Damien et al. 2019). Les vibrations se propagent ensuite à travers les plis vocaux (jaune) et le sac laryngé dans les plis ventraux de la gorge, pour se diffuser dans l'eau (ondes roses).

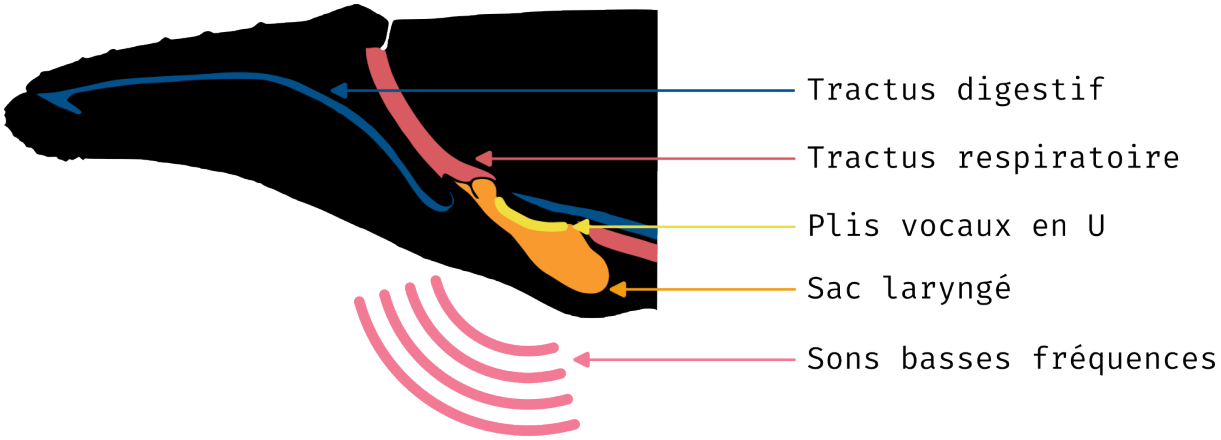


Figure 16 – Schéma du mécanisme potentiel de production des sons chez les baleines à fanons. Figure adaptée et refaite de Reidenberg et Laitman 2007.

Chez les baleines bleues, les mâles émettent des chants complexes, souvent longs (> 15 s), organisés en séries répétitives séparées par des périodes de silence. Ces signaux, dont l'intensité peut atteindre jusqu'à 189 dB re 1 µPa @1m, sont particulièrement bien adaptés à la transmission sur de longues distances (Širović, Hildebrand et Wiggins 2007 ; Samaran, Adam et Guinet 2010a). Les chants de cette espèce sont également de précieux marqueurs acoustiques pour distinguer les sous-espèces et les populations, en raison de leurs signatures spectrales stables dans le temps et propres à chaque groupe géographique (Leroy, Royer, Bonnel et al. 2018 ; McDonald, Mesnick et Hildebrand 2006 ; Clark et Garland 2022).

Par exemple, les baleines bleues antarctiques (*Balaenoptera musculus intermedia*) émettent toute l'année un chant caractéristique appelé Z-call, composé de trois unités successives (Fig.17) (Shannon et al. 2005). Ce chant, qui constitue une signature acoustique propre à cette population, peut être détecté sur plusieurs centaines de kilomètres (Širović, Hildebrand et Wiggins 2007). Les individus peuvent produire ce motif vocal en boucle pendant plusieurs jours, suggérant une fonction persistante, probablement liée à la reproduction ou au maintien du contact sur de longues distances (Leroy, Samaran, Bonnel et al. 2016).

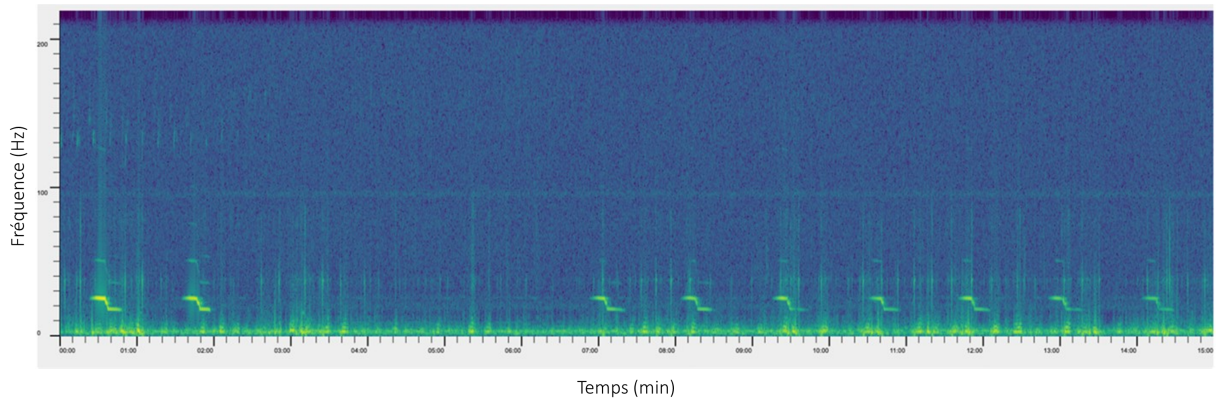


FIGURE 17 – Spectrogramme illustrant des Z-calls de ANT BW sur 15 minutes. Enregistrement réalisé en 2020 dans l'Océan Indien.

De manière similaire, les baleines bleues pygmées de l'océan Indien se répartissent en plusieurs populations acoustiquement distinctes, chacune produisant un chant spécifique, parfois modulé et comportant des harmoniques, ce qui permet leur identification acoustique (McDONALD, MESNICK et HILDEBRAND 2006). Par exemple, la population du sud-est de l'Océan Indien est caractérisée par un chant à trois unités tonales stéréotypées enregistrées de façon saisonnière sur les côtes ouest et sud de l'Australie (Fig.18) (McCAULEY, JENNER et al. 2001), tandis que la population du sud-ouest de l'Océan Indien est identifiable par un motif vocal spécifique enregistré autour de Madagascar (Fig.19) (LJUNGBLAD, CLARK et SHIMADA 1998).

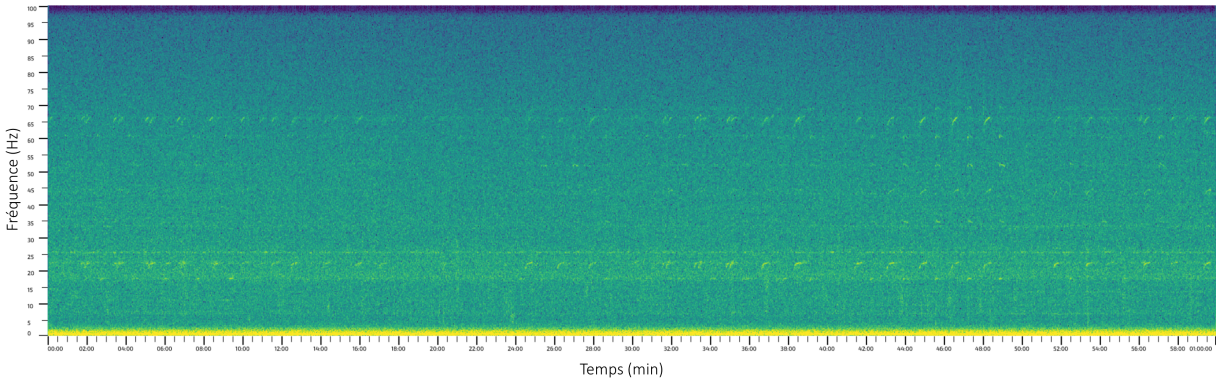


FIGURE 18 – Spectrogramme illustrant des chants de SEIO PBW sur 1 heure. Enregistrement réalisé en 2020 dans l'Océan Indien.

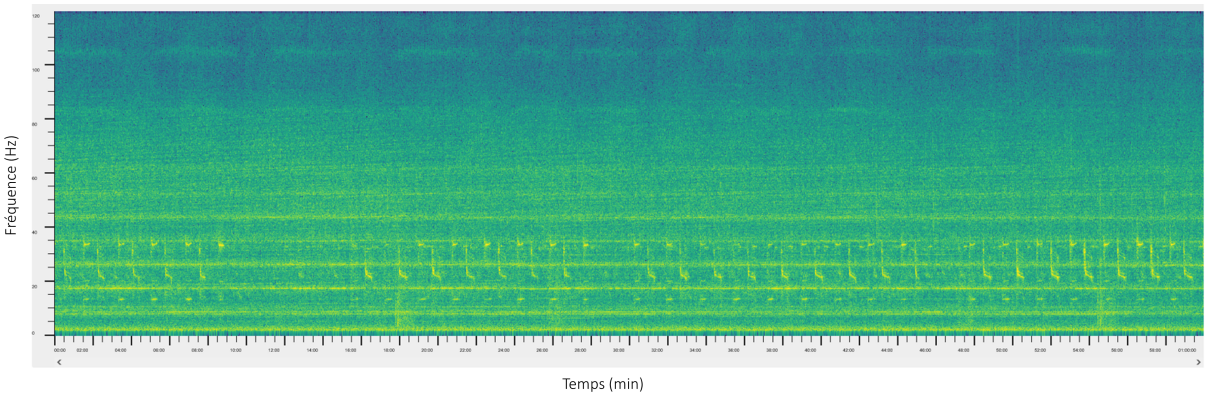


FIGURE 19 – Spectrogramme illustrant des chants de SWIO PBW sur 1 heure. Enregistrement réalisé en 2020 dans l'Océan Indien.

En plus de ces chants stéréotypés, les baleines bleues (toutes sous-espèces confondues) produisent également des signaux non structurés, souvent appelés D-calls, Frequency-Modulated (FM) calls, ou archsounds (THOMPSON et al. 1996 ; LJUNGBLAD, STAFFORD et al. 1997 ; THODE, D'SPAIN et KUPERMAN 1999). Ces signaux sont généralement plus courts et modulés que les chants, avec une fréquence très variable décroissant d'environ 100 Hz à 30 Hz (Fig. 20). Contrairement aux chants, ils ne présentent pas de variation géographique marquée dans leur forme et ont été enregistrés dans toutes les régions où les chants de baleines bleues ont été documentés (BARLOW et TORRES 2021 ; BUCHAN et al. 2021 ; MCDONALD, CALAMBOKIDIS et al. 2001 ; MCDONALD, MESNICK et HILDEBRAND 2006 ; MELLINGER et CLARK 2003 ; RANKIN et al. 2005 ; SAMARAN, ADAM et GUINET 2010a ; TORTEROTOT, SAMARAN et ROYER 2023).

Ces D-calls seraient produits par les mâles et les femelles (LEWIS et al. 2018 ; OLESON et al. 2007). Les études disponibles, principalement menées sur les zones d'alimentation de Californie du Sud, suggèrent que ces signaux sont liés à des comportements sociaux et à la coordination pendant l'alimentation (LEWIS et al. 2018 ; SZESCIORKA et al. 2020). Par exemple, les données issues de balises acoustiques ont montré que les D-calls sont souvent émis lors de plongées superficielles entre des plongées profondes de nourrissage, ce qui pourrait permettre d'attirer des congénères sur des zones riches en ressources ou de maintenir la communication entre individus dispersés (OLESON et al. 2007 ; LEWIS et al. 2018). Dans l'estuaire du Saint-Laurent, au Canada, et le long de la côte chilienne, ces signaux semblent également jouer un rôle dans le maintien de la cohésion entre couples

mâle-femelle, notamment lorsqu'un troisième individu rejoint le groupe (Schall et al. 2019).

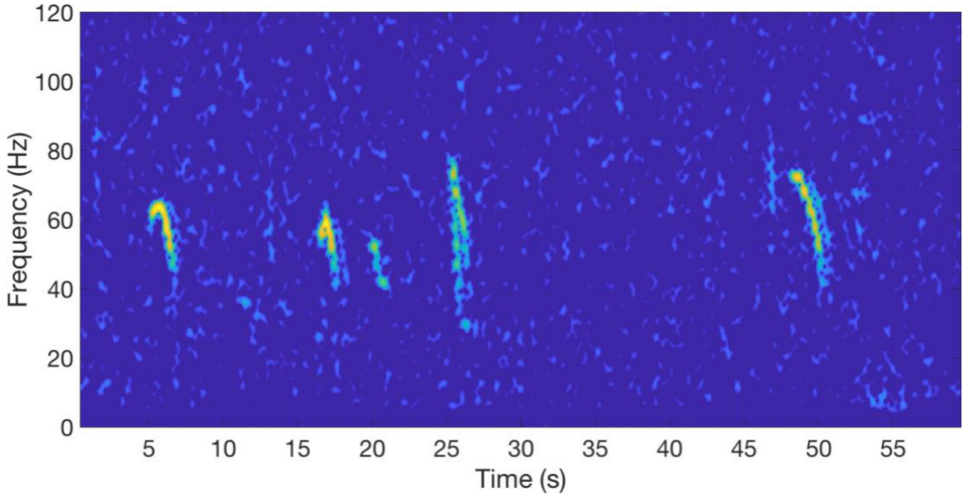


Figure 20 – Spectrogramme illustrant 5 D-calls sur 1 minute. Enregistrement réalisé dans l'Océan Indien.

Un autre exemple emblématique de communication acoustique chez les mysticètes est celui des baleines à bosse (*Megaptera novaeangliae*). Les mâles émettent des chants structurés et complexes, principalement pendant la saison de reproduction. Bien qu'ils soient le plus souvent enregistrés sur les zones de reproduction, ils sont également entendus sur les routes migratoires et, plus rarement, dans les zones d'alimentation. Ces chants jouent un rôle dans la sélection sexuelle, bien que leur fonction exacte reste débattue (Clapham 1996 ; Darling, Jones et Nicklin 2006). Contrairement à ceux des baleines bleues, les chants des baleines à bosse couvrent une gamme de fréquences plus large (de quelques dizaines de Hz à plusieurs kHz) et évoluent au cours du temps (Fig. 21).

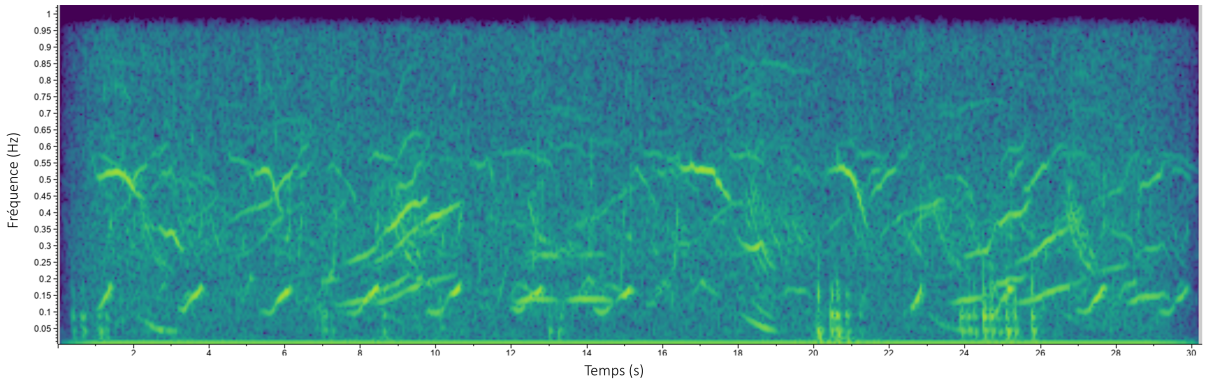


FIGURE 21 – Spectrogramme illustrant des chants d'une baleine à bosse sur 30 secondes. Enregistrement réalisé dans le sanctuaire Agoa (Antilles françaises).

Les mâles d'une même population partagent une structure de chant commune, qu'ils modifient collectivement au fil des saisons : certains motifs peuvent être progressivement transformés, simplifiés ou enrichis, tandis que, dans certains cas, le répertoire complet peut être remplacé par un chant importé d'une autre population (NOAD et al. 2000). Cette diffusion d'un répertoire partagé repose sur une transmission horizontale [8], c'est-à-dire un apprentissage entre individus d'une même génération, par imitation sociale. À l'instar des codas de cachalots cités dans la section précédente, ce phénomène constitue l'un des rares exemples documentés de culture acoustique chez les non-humains.

Méthodologie et défis scientifiques du suivi par acoustique passive

L'acquisition de la donnée

Dans le contexte du PAM, le choix du matériel acoustique et des paramètres techniques de déploiement repose sur une multiplicité de facteurs qui doivent être anticipés et soigneusement dimensionnés : espèces ciblées, objectifs écologiques, contraintes logistiques, durée souhaitée du suivi, conditions environnementales locales, capacité de traitement des

8. Apprentissage et acquisition d'un comportement ou d'un signal entre individus de la même génération, par imitation sociale, plutôt que par héritage génétique ou apprentissage parental.

données, etc. Il ne s'agit pas d'un simple choix technologique, mais d'un défi méthodologique fondamental. Les décisions relatives au matériel et aux protocoles d'enregistrement influencent directement la qualité et la portée des données recueillies et doivent être prises en amont pour garantir la robustesse et la pertinence des résultats à long terme. Il n'existe donc pas de modèle unique, mais une diversité de dispositifs et de stratégies qui reflètent à la fois la souplesse du PAM et la complexité de sa mise en œuvre à des fins de suivi standardisé (Mellinger, Stafford et al. 2007).

Identification des espèces ciblées dans la zone

La première étape d'une stratégie PAM consiste à identifier précisément les espèces de cétacés présentes et que l'on souhaite suivre dans la zone d'étude. Cette identification est essentielle car elle conditionne les questions scientifiques que l'on pourra aborder, ainsi que l'orientation générale du protocole de recherche. Comme présenté dans la section « Les cétacés : des candidats idéaux », les différentes espèces possèdent des caractéristiques acoustiques distinctes, liées à leur biologie et à leur écologie. Cette diversité se traduit notamment par des gammes de fréquences variées, influençant la distance sur laquelle leurs sons peuvent être détectés. Le choix des espèces dépend donc directement des objectifs scientifiques. Parmi les questions classiques auxquelles le PAM peut répondre, il est possible de citer :

- **Où et quand les espèces sont-elles présentes dans une zone donnée, y compris les espèces rares ou difficiles à observer ?** Cette question de distribution spatiale et temporelle est essentielle pour la conservation. Elle permet d'identifier des habitats critiques, des corridors de migration ou des périodes sensibles.
- **Quels sont les comportements vocaux, sociaux et alimentaires des espèces ciblées ?** L'analyse des vocalisations permet d'explorer les modes de communication, les structures sociales et les stratégies de reproduction. L'évolution des comportement vocaux peut aussi révéler des comportements d'apprentissage ou d'adaptation culturelle.
- **Comment les espèces réagissent-elles aux activités humaines, notamment aux sources de bruit sous-marin ?** Le PAM est un outil particulièrement adapté pour documenter les effets des pressions anthropiques sur les cétacés. Il permet par exemple de détecter des modifications dans les émissions sonores, des changements

de comportements vocaux, ou des variations dans la présence acoustique au sein d'une zone donnée. Dans le contexte du développement d'activités en mer, comme les parcs éoliens offshore, cette approche permet un suivi avant, pendant et après les travaux.

Chaque question oriente donc le choix des espèces à étudier et les modalités du protocole, ce qui souligne l'importance d'une connaissance préalable approfondie des espèces présentes dans la zone. Cette étape s'appuie généralement sur des inventaires précédents, des observations visuelles ou des campagnes pilotes d'enregistrement acoustique.

Contexte environnemental et contraintes logistiques

La définition d'une stratégie de PAM ne dépend pas uniquement des espèces ciblées ou des objectifs scientifiques. Le contexte environnemental et les contraintes logistiques jouent un rôle tout aussi déterminant, influençant les choix de matériel, de déploiement et de traitement des données. Intégrer ces éléments dès la phase de conception permet d'optimiser la qualité des enregistrements tout en assurant la faisabilité du dispositif.

Les conditions physiques du milieu, comme la bathymétrie, la température, la salinité, la turbidité ou les courants, affectent directement la propagation du son sous l'eau (Lutron 2002). Ces facteurs doivent être pris en compte pour estimer la portée des capteurs et adapter la stratégie de déploiement. Le niveau de bruit ambiant, qu'il soit d'origine naturelle (vagues, pluie, faune non ciblée) ou anthropique (trafic maritime, chantiers, moteurs), conditionne également la capacité à détecter les vocalisations : dans un environnement bruyant, les signaux peuvent être partiellement ou totalement masqués.

Le positionnement des enregistreurs est par ailleurs soumis à des contraintes réglementaires et pratiques. Il n'est pas toujours possible de les déployer librement : espaces protégés, zones militaires, trafic maritime ou activités de pêche peuvent limiter les emplacements accessibles. L'obtention des autorisations nécessaires, notamment pour les dispositifs laissés en mer sur le long terme, constitue souvent une étape incontournable.

Enfin, les réalités logistiques orientent les choix techniques. En zone côtière exposée, il est nécessaire d'opter pour des dispositifs robustes, faciles à manipuler et résistants aux conditions de mer. En zone hauturière, les contraintes d'accès sont plus importantes, mais

les enregistreurs peuvent être laissés plus longtemps sur site, car ils y sont moins exposés aux risques de dégradation liés aux activités humaines (pêche, navigation, vandalisme) ou aux perturbations acoustiques côtières. La stratégie globale dépend donc étroitement de la durée des missions, des moyens nautiques disponibles, du budget et de la capacité à traiter les données collectées.

Propagation du son en milieu marin – 1. Propriétés du milieu

Quelques notions fondamentales en propagation acoustique sont utiles pour mieux appréhender les enjeux liés au PAM des cétacés. Le son se propage bien plus rapidement dans l'eau que dans l'air : sa vitesse (ou célérité) y est environ cinq fois plus élevée (environ 1500 m/s), et les ondes, notamment à basse fréquence, peuvent parcourir de grandes distances avec peu d'atténuation (Lutron 2002 ; Hamilton et Bachman 1982).

Propriétés du milieu influençant la vitesse et le trajet des ondes :

- **Température** : la vitesse du son augmente avec la température (une variation de 1°C augmente la vitesse de 3m/s).
- **Salinité** : elle modifie modérément mais significativement la vitesse de propagation (une variation de 1g/l augmente la vitesse de 1m/s).
- **Pression** : la pression hydrostatique augmente avec la profondeur, ce qui accroît la célérité du son (environ +1.7 m/s tous les 100 m de profondeur).
- **Stratification** : des couches d'eau de températures ou salinités différentes peuvent dévier le son (réfraction), créant des zones d'ombre ou de concentration acoustique.
- **Bathymétrie** : le relief sous-marin, notamment en zone côtière, influence la réflexion et les pertes d'énergie selon la nature du fond (meuble, rocheux, canyon, etc.).

Propagation du son en milieu marin – 2. Phénomènes physiques

Les propriétés physiques du milieu (température, salinité, pression, stratification, bathymétrie) conditionnent directement la propagation du son. Elles expliquent en particulier divers phénomènes physiques qui modifient la transmission du signal acoustique entre l'émetteur (e.g. cétacé) et le récepteur (e.g. hydrophone). Ces mécanismes influencent la portée, la direction et l'intensité du son reçu.

Phénomènes physiques affectant la propagation acoustique :

- **Absorption** : une partie de l'énergie est absorbée par les molécules d'eau ou le sol ; ce phénomène s'intensifie avec la fréquence.
- **Diffusion** : des bulles, particules et organismes dispersent le signal, surtout en hautes fréquences.
- **Réflexion** : les ondes peuvent rebondir sur la surface ou le fond, créant des trajets multiples.
- **Réfraction** : les gradients de température, salinité et pression modifient la célérité de l'onde dans le milieu et modifie sa trajectoire.
- **Diffraction** : lorsqu'une onde acoustique rencontre un obstacle, les hautes fréquences sont diffusées par les aspérités de sa surface tandis que les basses fréquences le contournent en étant moins altérées. Il en résulte une perte d'énergie dépendant de la fréquence.

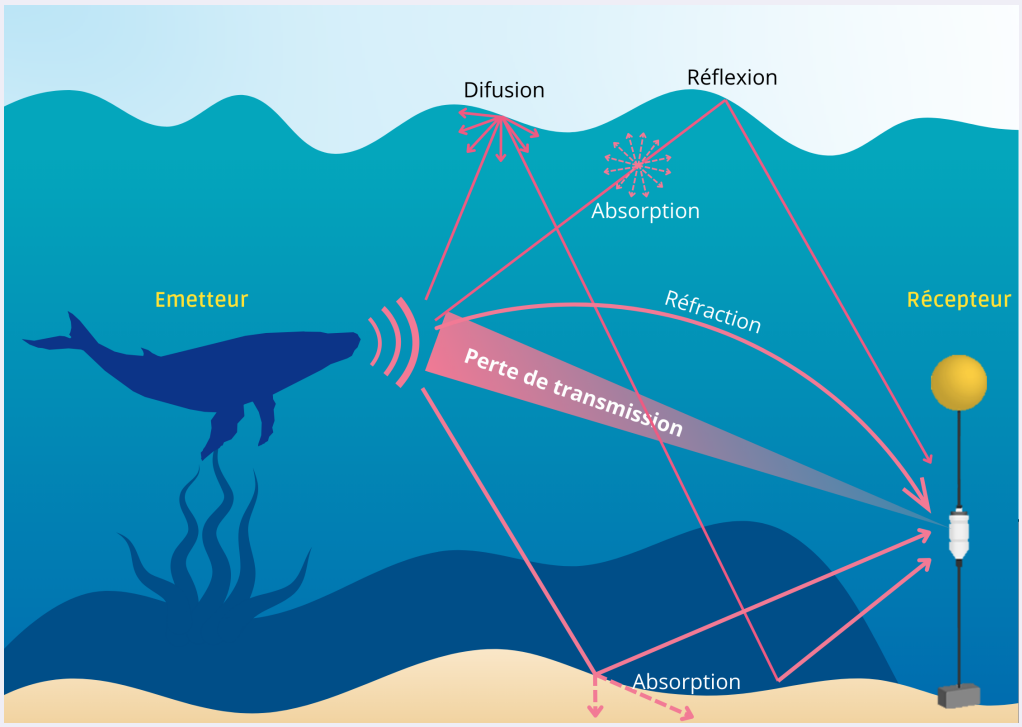


FIGURE 22 – Illustration des principaux phénomènes influençant la propagation d'un signal acoustique sous-marin, entre un émetteur (e.g. cétacé) et un récepteur (e.g. l'enregistreur et le capteur i.e. hydrophone).

Choix du matériel acoustique

Un enregistreur acoustique autonome est un dispositif électronique conçu pour enregistrer, numériser et stocker les sons. Quand les enregistrements ont lieu dans l'eau, il est associé à un hydrophone, un capteur acoustique spécialement conçu pour détecter les vibrations sonores dans cet environnement. Fonctionnant comme un microphone adapté au milieu aquatique, l'hydrophone convertit les variations de pression acoustique en signaux électriques, permettant ainsi l'enregistrement et l'analyse des sons. Le choix du matériel acoustique repose sur une prise en compte simultanée des objectifs scientifiques, des caractéristiques des espèces ciblées, du contexte environnemental et des contraintes logistiques. Il s'agit de sélectionner des enregistreurs et des capteurs capables de fournir des données de qualité, adaptées aux types de vocalisations attendues, tout en répondant aux impératifs opérationnels.

Les enregistreurs large bande sont des dispositifs capables de capturer un large spectre fréquentiel, incluant à la fois les sons basse fréquence (i.e. les chants et les sifflements) et les émissions à haute fréquence (i.e. les clics d'écholocalisation) (Fig. 23). Cette polyvalence leur permet d'enregistrer simultanément une grande variété de signaux biologiques, ainsi que le bruit ambiant naturel et les sources anthropiques. Utilisés non seulement pour l'étude des cétacés, ces enregistreurs servent également dans des contextes plus larges, comme la caractérisation des paysages sonores marins, la mesure du bruit ambiant ou la surveillance sismique.

En parallèle, certains dispositifs spécialisés, comme les détecteurs de clics «C-PODS et F-PODs» (Chelonia Limited), sont conçus pour détecter spécifiquement des sons impulsionnels comme les clics d'écholocalisation d'odontocètes. Ces enregistreurs autonomes effectuent un traitement embarqué et automatisé des signaux impulsionnels, enregistrant uniquement les instants auxquels ils ont été détectés et générant des indicateurs synthétiques (fréquences, durées). Cette absence d'enregistrement des données brutes constitue une limite importante, puisqu'elle restreint la possibilité d'analyses complémentaires ou d'évaluations rétrospectives. Une fois la donnée récupérée, le constructeur fournit un logiciel qui permet de classer ces impulsions en clic de marsouin, clic d'odontocète ou sonar de bateau. Ce logiciel fournit ainsi une quantité de clics par unité de temps directement exploitable. Ces instruments sont largement utilisés dans des programmes collaboratifs multipartenaires de suivis à grande échelle comme le projet SAMBAH à l'échelle de la mer Baltique



FIGURE 23 – Panels d'enregistreurs acoustiques disponibles (de gauche à droite : Loggerhead, Ocean Instrument NZ, Seiche Marine Technology, JASCO, RTsys, Osean Underwater technology, Wildlife, MTE, Cetacean Research Technology, SAEZ Electronica Submarina DBV Technology) (GERVAISE et al. 2018).

(CARLSTRÖM et al. 2011). Cette approche réduit considérablement la charge de post-traitement et facilite la mise en place de réseaux standardisés ainsi que la comparaison entre différents sites ou périodes. De plus, ces dispositifs bénéficient généralement d'une longue autonomie, ce qui les rend adaptés aux déploiements de longue durée. Toutefois, cette spécialisation restreint la portée acoustique aux clics, excluant les autres types de vocalisations et limitant la détection aux espèces produisant des sons dans la plage ciblée par l'enregistreur.

Il est important de noter que la plupart de ces enregistreurs acoustiques fonctionnent de manière autonome et stockent les données localement. Ainsi, il est nécessaire de retourner sur le site de déploiement pour récupérer physiquement les enregistreurs et en extraire les données, ce qui peut contraindre la fréquence des relevés et complexifier la logistique, notamment dans des zones éloignées ou difficiles d'accès. Contrairement aux systèmes actifs ou aux technologies de communication à distance, ces dispositifs ne permettent pas une transmission en temps réel des données. Le choix entre ces types d'enregistreurs dépend donc directement des questions scientifiques posées, de la durée et de la fréquence des enregistrements envisagées, des contraintes d'alimentation et de récupération, ainsi que des conditions physiques du site.

Réglages techniques des enregistreurs acoustiques

Les enregistreurs acoustiques présentés précédemment offrent une grande flexibilité de configuration, permettant d'adapter les paramètres techniques (e.g. fréquence d'échantillonnage, *duty cycle*, gain, etc.) aux objectifs spécifiques du PAM. Cette modularité constitue un atout majeur, mais elle représente également un enjeu important : ces réglages influencent directement la qualité des données recueillies, la durée de fonctionnement des dispositifs ainsi que la pertinence et la robustesse des analyses écologiques qui en découlent. Il est donc essentiel de choisir, de justifier et de documenter ces paramètres avec soin afin d'optimiser à la fois la performance des enregistreurs et la fiabilité des indicateurs acoustiques produits.

— **Fréquence d'échantillonnage** : Elle correspond au nombre de mesures acoustiques prises par seconde par l'enregistreur. Selon le théorème de Nyquist-Shannon, elle doit être au moins deux fois supérieure à la fréquence maximale du signal à enregistrer, afin d'éviter le repliement spectral et de conserver l'intégrité du signal (Shannon 2001). Par exemple, pour capter les clics d'odontocètes dont la fréquence peut dépasser 70 kHz, la fréquence d'échantillonnage doit être supérieure à 140 kHz, souvent réglée autour de 148 kHz. Une fréquence échantillonnage élevé garantit une meilleure résolution temporelle, essentielle pour l'identification fine des espèces et l'analyse détaillée des signaux. Cependant, elle génère également un volume de données plus important, ce qui augmente les besoins en capacité de stockage et en énergie, réduisant ainsi la durée d'enregistrement possible sans intervention. À titre d'exemple, un enregistreur large bande paramétré à 148 kHz pour être en capacité d'enregistrer les clics de delphinidés pourra rester actif environ trois mois (Mellinger, Stafford et al. 2007), tandis qu'un dispositif limité aux basses fréquences des chants de baleines (< 100 Hz, échantillonnage à 240 Hz) pourra rester déployé plus d'un an (Torterotot, Samaran, Stafford et al. 2020).

— ***Duty cycle*** : Il désigne la proportion de temps pendant laquelle l'enregistreur est activement en train d'enregistrer par rapport à une période totale donnée (voir encadré ci-dessous pour le détail de cette notion clé). Par exemple, un *duty cycle* avec une période d'écoute de 50% signifie que l'enregistreur enregistre pendant la moitié du temps et reste inactif l'autre moitié. Cette stratégie vise à optimiser l'autonomie

énergétique et la capacité de stockage en réduisant la durée effective d'enregistrement. Cette méthode est particulièrement utilisée dans les campagnes de longue durée ou lorsque les ressources logistiques sont limitées. Cependant, un *duty cycle* réduit peut introduire un biais d'échantillonnage : certains événements acoustiques risquent d'être manqués, en particulier lorsqu'il s'agit de vocalisations rares, intermittentes ou ponctuelles, ce qui affecte la représentativité des données. Cette notion de *duty cycle*, ses enjeux et impacts sur la représentativité de la mesure des données acoustiques font l'objet d'une analyse approfondie dans le chapitre 1 de cette thèse.

— **Gain** : C'est un amplificateur du signal reçu par l'hydrophone dont on peut régler sa valeur avant numérisation. Un gain trop faible peut empêcher la détection des sons faibles, réduisant la sensibilité du capteur. À l'inverse, un gain trop élevé peut saturer l'enregistreur face à des sons forts, introduisant des distorsions et des artefacts. Le réglage du gain doit être adapté à l'environnement acoustique et aux objectifs.

— **Filtres** : Les filtres (passe-bas, passe-haut, passe-bande) permettent d'atténuer certaines plages de fréquence, réduisant ainsi les bruits parasites hors de la bande d'intérêt. Leur réglage dépend des espèces ciblées et des types de vocalisations attendues.

— D'autres paramètres, comme la résolution en bits de la conversion analogique-numérique, influencent aussi la qualité du signal numérisé. Une résolution élevée (16 à 24 bits) garantit une meilleure fidélité sonore.

Une notion essentielle : le duty cycle

Le *duty cycle* désigne le régime d'enregistrement des enregistreurs acoustiques. Plutôt que d'enregistrer en continu, les enregistreurs peuvent fonctionner par intermittence, selon un cycle défini à l'avance. Deux paramètres le caractérisent :

— **le pourcentage d'enregistrement,** D, qui indique la proportion de temps pendant laquelle l'enregistreur est actif (par exemple 50%, 33%, etc.) ;

— **la période d'écoute,** L_p, qui précise la durée de chaque segment d'enregistrement (par exemple 1 minute, 10 minutes, 30 minutes, etc.).

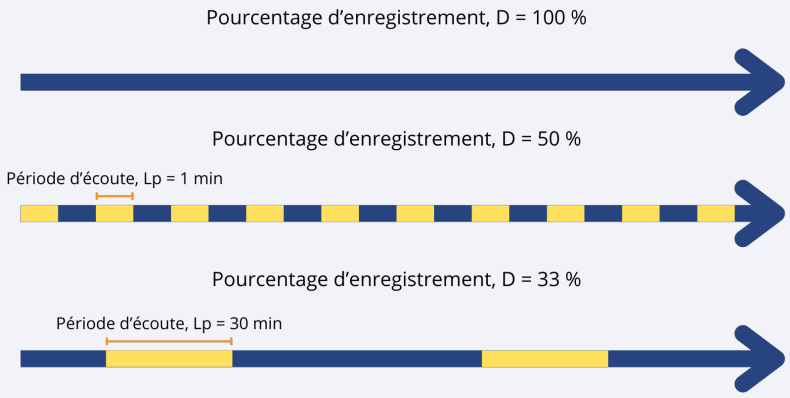


FIGURE 24 – Exemple de différents *duty cycle* selon le pourcentage d'enregistrement D et la période d'écoute L_p.

Le couple **pourcentage d'enregistrement – période d'écoute** détermine la fenêtre temporelle pendant laquelle l'enregistreur enregistre l'environnement sonore. Par exemple, un *duty cycle* avec un pourcentage d'écoute de 33% et une période d'écoute de 10 minutes signifie que l'enregistreur enregistre 10 minutes toutes les 30 minutes.
Son choix représente donc un compromis essentiel dans la conception d'un suivi PAM : un *duty cycle* faible permet d'économiser l'énergie des enregistreurs autonomes, prolongeant ainsi leur autonomie et la durée totale du suivi, mais il réduit également la quantité de données enregistrées, ce qui peut entraîner la perte de certains événements acoustiques. À l'inverse, un *duty cycle* élevé assure une meilleure couverture temporelle et une détection plus précise, au prix d'une consommation énergétique plus importante. Le terme *duty cycle* sera donc utilisé tout au long du manuscrit.

Traitement des enregistrements

L'analyse des enregistrements constitue une étape cruciale dans les suivis PAM. Elle assure le passage des données brutes collectées vers des informations interprétables sur le plan biologique et écologique. Similairement au choix de l'équipement, cette étape requiert

une prise de décision méthodologique essentielle, susceptible d'influencer la robustesse et la fiabilité des résultats obtenus. Le traitement peut être réalisé de manière **manuelle**, **semi-automatisée** ou **automatisée** :

— La méthode **manuelle** consiste à examiner visuellement et auditivement les spectrogrammes pour identifier les signaux d'intérêt.

— Les méthodes **automatisées** confient cette tâche à des algorithmes qui analysent les spectrogrammes ou les enregistrements bruts pour détecter et classer les signaux, éventuellement avec un apprentissage supervisé à partir de jeux de données validés en amont.

Les méthodes manuelles, bien que chronophages, restent largement utilisées et constituent parfois la seule approche possible dans certains suivis acoustiques. Elles sont également indispensables comme base de référence pour le développement et la validation des approches de détection automatique. Cette analyse permet une annotation fine, notamment dans des contextes difficiles tels que la présence de bruit ambiant intense (e.g. trafic maritime, activités industrielles), qui peut masquer partiellement ou totalement les signaux d'intérêt, ou encore dans des situations de chevauchement acoustique entre plusieurs espèces présentes au même moment, dans une même zone (sympatrie acoustique). Ces conditions, illustrées à la Fig. 25, compliquent l'attribution correcte des signaux et soulignent l'importance de disposer d'un jeu de données validé manuellement pour entraîner et tester les méthodes automatisées. Il convient cependant de noter que même l'annotation humaine peut être sujette à des biais ou des erreurs, notamment en raison des limitations des représentations de spectrogrammes et de la subjectivité de l'interprétation (Dubus, Cazau et al. 2024 ; Leroy, Thomisch et al. 2018).

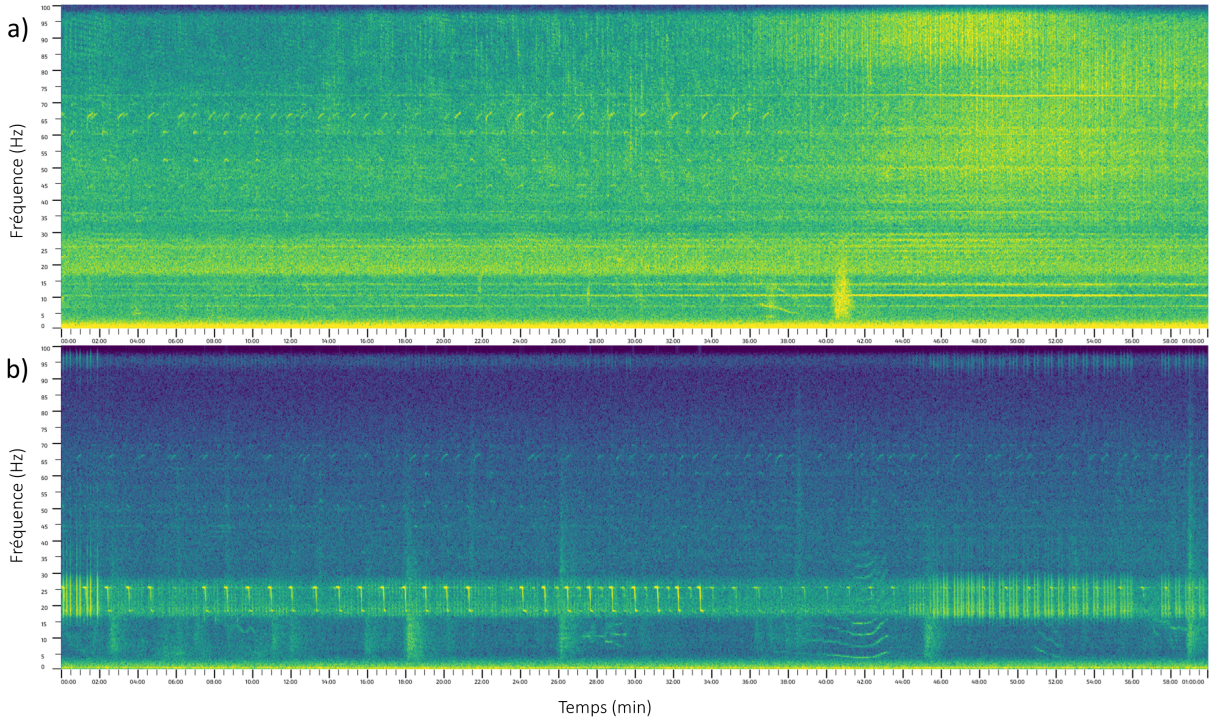


FIGURE 25 – Spectrogrammes illustrant a) des chants de SEIO PBW sur 1 heure masqué par des bruits de bateaux b) le chevauchement de signaux avec des chants de SEIO PBW, ANT BW, rorqual commun (FW) et bruit de cryogénie sur 1 heure. Enregistrements réalisé en 2020 dans l'Océan Indien.

Cependant, des plateformes d'annotation collaborative en ligne, telles qu'`APLOSE`[9] (DUBUS, TORTEROTOT et al. 2025 ; KERIBIN, MORIN et VOVARD 2024), simplifient cette tâche en proposant des outils ergonomiques pour la visualisation, l'annotation et le partage des données acoustiques (Fig. 26).

De leur côté, les approches automatisées, telles que `PAMGuard` (GILLESPIE, MELLINGER et al. 2008), `DeepSqueak` (COFFEY, MARX et NEUMAIER 2019) ou les réseaux de neurones appliqués aux spectrogrammes (THOMAS et al. 2019), permettent de traiter à grande échelle des séries temporelles longues ou massives et d'automatiser la détection et la classification d'une grande variété de signaux. Ces algorithmes d'apprentissage automatique reconnaissent les motifs acoustiques caractéristiques des différentes espèces à partir de jeux de données annotés manuellement. La qualité et la diversité de ces exemples, ainsi qu'une validation rigoureuse (par comparaison aux données annotées ou via des tests croisés), déterminent

9. Plateforme open-source et en ligne de l'équipe OSmOSE pour l'annotation de données acoustiques et pour faciliter la collaboration en acoustique marine.

directement la capacité de l'outil à détecter correctement les signaux, même dans des conditions de bruit élevé ou en présence de plusieurs espèces (Zhong et al. 2021 ; Dréo, Crawford et al. 2025 ; Dubus, Cazau et al. 2024 ; Thomas et al. 2019).

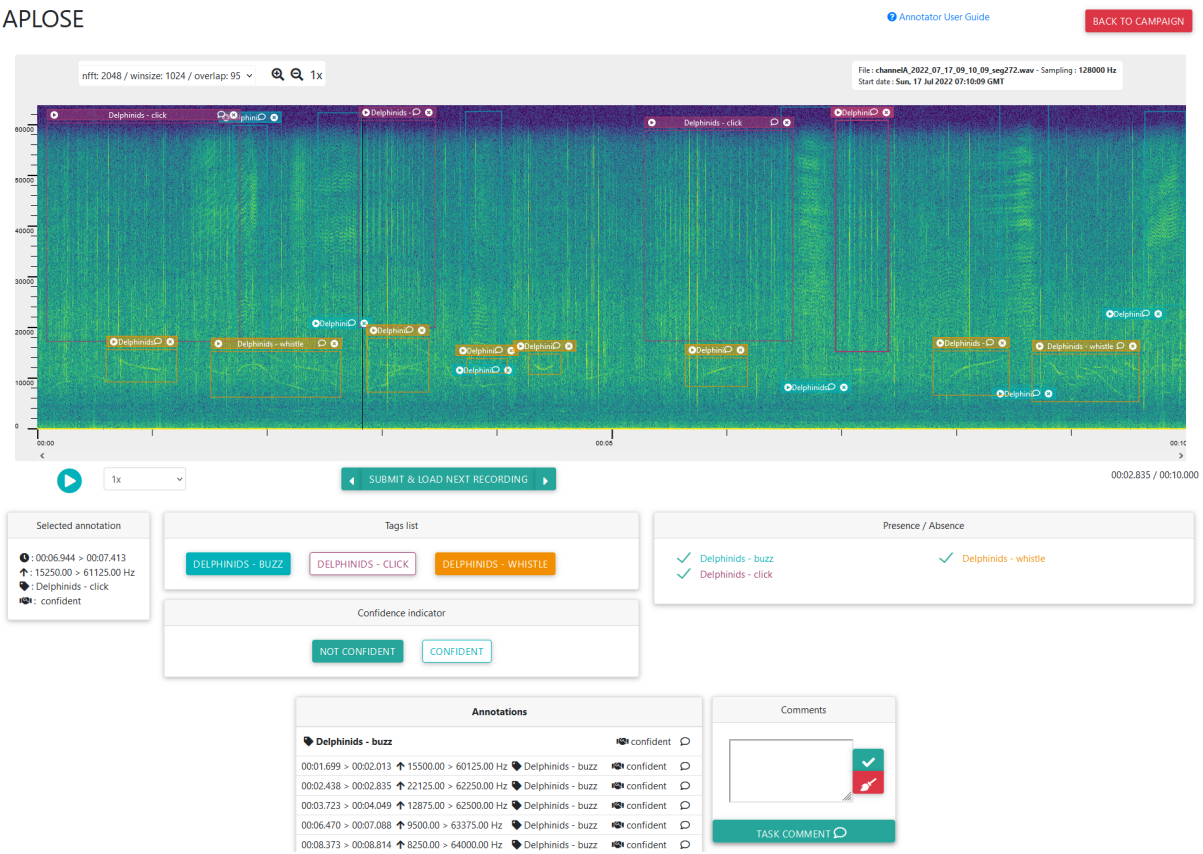


Figure 26 – Exemple d'un spectrogramme annoté sur la plateforme `APLOSE` (Keribin, Morin et Vovard 2024).

Une fois que les signaux sont identifiés et classifiés, ils sont agrégés sous forme de métriques acoustiques pour représenter la présence acoustique (ces notions sont détaillées dans la section « Choix des métriques et méthodes statistiques » ci-dessous). Ces choix méthodologiques ont une incidence directe sur la sensibilité des indicateurs à la détection des espèces et sur leur comparabilité entre études. Ils doivent donc être documentés de manière transparente et rigoureuse pour garantir la reproductibilité et l'interopérabilité des résultats.

Il est important de noter que cette thèse ne vise pas à détailler l'ensemble des méthodes de traitement des enregistrements acoustiques, ni à en comparer les performances ou

les objectifs spécifiques. Pour un exposé plus complet des stratégies de détection et de classification, notamment dans le contexte combinant annotation manuelle, science participative et apprentissage profond, le lecteur pourra se référer à la thèse de Gabriel Dubus 2025 (Sorbonne Université) qui aborde ces questions en détail : *From crowdsourcing to supervised deep learning : addressing new challenges in the Detection and Classification of Cetaceans using Passive Acoustic Monitoring.*

Choix des métriques et méthodes statistiques

Métriques de représentation de la présence acoustique

L'analyse des données PAM repose sur deux dimensions fondamentales qu'il est essentiel de définir en amont de toute étude :

1. **La fenêtre temporelle**, soit l'échelle à laquelle on résume les détections dans le temps ;
2. **Le type de métrique**, c'est-à-dire la manière dont on mesure l'activité acoustique à partir des détections.

Ces deux paramètres structurent profondément l'interprétation des résultats. Leur choix dépend de la biologie de l'espèce, de l'objectif scientifique, des caractéristiques du jeu de données, des conditions d'enregistrement et des méthodes de détection employées.

Exemples de métriques acoustiques courantes :

— **Présence/absence** : métrique binaire (0 ou 1), indiquant si au moins un signal d'intérêt a été détecté dans une unité temporelle donnée (e.g. minutes, heures, jours).

— **Nombre de fenêtres positives** : total d'unités temporelles (e.g. minutes, heures, jours) contenant au moins une détection.

— **Nombre brut de signaux d'intérêt** : total des signaux détectés dans une unité temporelle (e.g. minutes, heures, jours).

— **Taux de vocalisations** : nombre de vocalisations rapporté à une unité de temps (e.g. heure, jour, etc.).

La fenêtre temporelle contrôle la résolution de l'analyse. Des fenêtres plus fines (e.g. à l'échelle horaire) permettent de détecter des variations à court terme, comme des rythmes

circadiens, des changements de comportement en réponse à un stimulus ou des événements isolés. À l'inverse, des fenêtres plus larges (e.g. à l'échelle hebdomadaire ou mensuelle) offrent une vision synthétique utile pour mettre en évidence des tendances saisonnières, des schémas migratoires ou pour comparer différents sites d'enregistrement. L'adaptation de la fenêtre temporelle à l'échelle des processus biologiques observés est donc essentielle pour interpréter correctement les données.

Normalisation à l'effort d'enregistrement

Pour être comparables, les métriques doivent être **normalisées en fonction de l'effort d'enregistrement**, c'est-à-dire corrigées par la durée effective des données disponibles dans chaque fenêtre temporelle. Par exemple, une journée avec seulement 6 heures d'enregistrement ne peut pas être directement comparée à une journée complète de 24 heures. On peut alors exprimer le nombre de détections ou d'événements sous forme de **taux ou de pourcentage** par unité de temps. Cette normalisation est essentielle pour éviter des biais dus à des pertes de données ou à des configurations de déploiement différentes. Ainsi, chaque métrique décrite ci-dessus doit être systématiquement rapportée à l'effort d'enregistrement pour garantir des comparaisons valides entre périodes ou entre sites.

Pour illustrer l'impact concret de ces choix méthodologiques sur l'analyse et l'interprétation des résultats, trois figures comparatives sont présentées, basées sur une année complète de données (2019) issues d'un même site. Les détections représentent la présence acoustique de la baleine bleue antarctique, identifiée à l'aide d'un détecteur automatique (Torterotot, Royer et Samaran 2019).

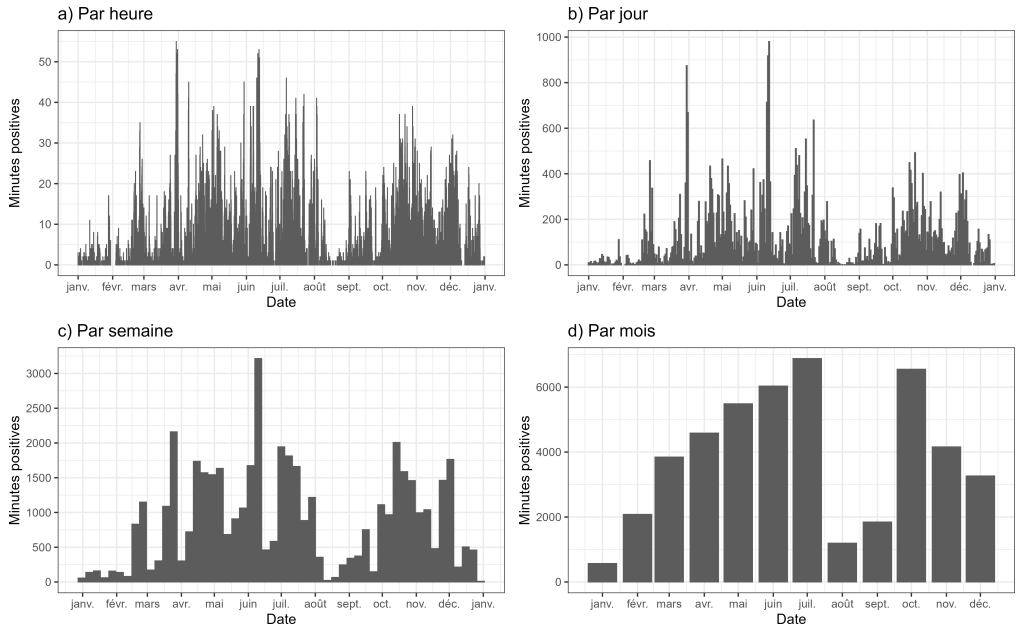


FIGURE 27 – Effet de la fenêtre temporelle sur la représentation du nombre de minutes positives : a) par heure ; b) par jour ; c) par semaine ; d) par mois.

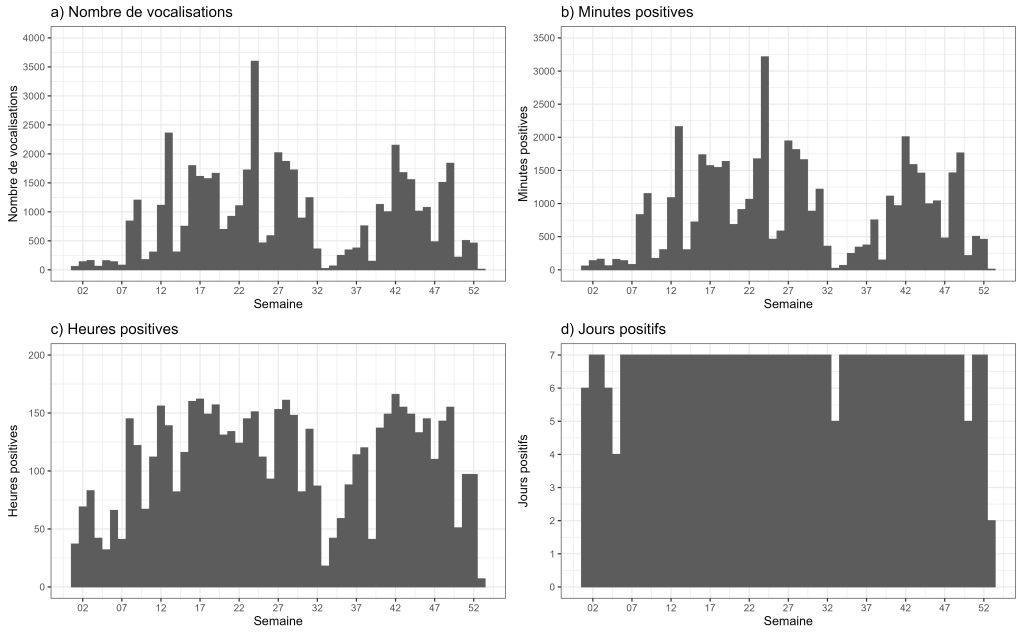


FIGURE 28 – Effet du choix de la métrique sur la fenêtre temporelle « semaine » : a) nombre de vocalisations ; b) nombre de minutes positives ; c) nombre d'heures positives ; d) nombre de jours positifs.

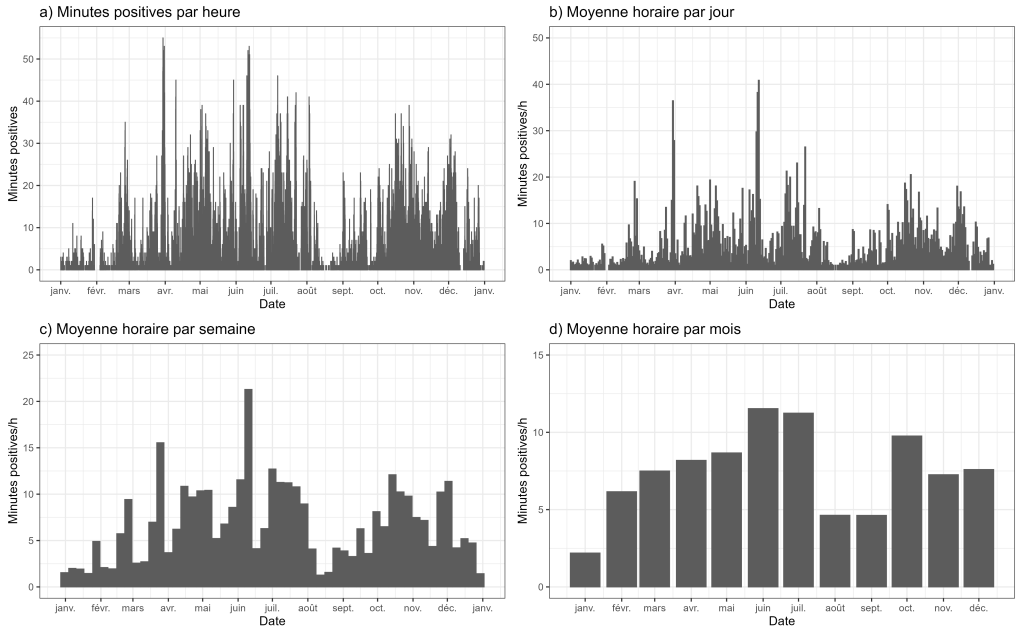


FIGURE 29 – Effet de la moyenne sur la métrique minutes positives avec une fenêtre temporelle par heure : a) données brutes (par heure) ; b) moyennes journalières ; c) moyennes hebdomadaires ; d) moyennes mensuelles.

Les figures 27, 28 et 29 mettent en évidence trois aspects méthodologiques majeurs :

— **Effet de la fenêtre temporelle :** Les agrégations fines (par heure) permettent de détecter des pics d'activité ponctuels ou des rythmes circadiens, tandis que les agrégations larges (mensuelles) synthétisent les tendances générales mais réduisent la résolution temporelle des signaux observés (Fig. 27).

— **Effet du type de métrique :** L'utilisation de métriques détaillées, telles que le nombre de minutes ou d'heures positives, révèle des variations fines de l'activité, alors que des métriques plus globales (jours positifs) facilitent les comparaisons inter-sites ou saisonnières (Fig. 28).

— **Effet du lissage par moyenne :** L'application de moyennes journalières, hebdomadaires ou mensuelles réduit la variabilité locale, permettant l'identification de tendances globales à long terme tout en masquant les événements ponctuels (Fig. 29).

Ces illustrations soulignent que le choix de la fenêtre temporelle, du type de métrique

et du niveau d'agrégation influence fortement l'interprétation écologique des données et doit être adapté aux objectifs analytiques poursuivis.

Types de représentations descriptives

Comme détaillé précédemment, le choix des métriques et des fenêtres temporelles influence directement le type de représentation descriptive produite. Ces représentations descriptives se déclinent en plusieurs types :

— **Patron journalier** : représentation de la variation de la présence acoustique au cours d'un cycle de 24 heures, généralement basée sur la proportion de minutes ou d'heures positives par tranche horaire. Ce format permet d'identifier des rythmes d'activité circadiens ou des comportements à périodicité quotidienne (e.g. émissions vocales nocturnes) (Shabangu, Munoz et al. 2024 ; Cascão et al. 2020 ; Sol et al. 2024).

— **Patron saisonnier** : représentation de la variation mensuelle ou hebdomadaire de la présence acoustique sur une année, en utilisant des métriques agrégées (e.g. jours positifs ou nombre moyen d'heures positives par mois). Cette représentation met en évidence les rythmes saisonniers ou les schémas migratoires (Torterotot, Samaran, Stafford et al. 2020 ; Leroy, Samaran, Bonnel et al. 2016 ; Dréo, Crawford et al. 2025 ; Samaran, Adam et Guinet 2010b ; Cascão et al. 2020).

— **Taux de co-occurrence** : proportion d'unités temporelles (e.g. heure, jour) durant lesquelles plusieurs espèces ou types de signaux sont détectés simultanément. Cette métrique est utile pour explorer les relations de co-présence, de chevauchement temporel ou de ségrégation acoustique entre espèces ou groupes fonctionnels (Calonge et al. 2024 ; Torterotot, Béesau et al. 2022).

— **Indices de diversité acoustique** : mesures dérivées (e.g. richesse, indice de Shannon, complexité spectrale) calculées à partir de la variété des signaux détectés dans une fenêtre donnée (e.g. heure, jour, site). Ces indices permettent de caractériser la diversité des communautés acoustiques, y compris en l'absence d'identification taxonomique précise (Lin et al. 2015 ; Alcocer et al. 2022 ; Harris, Shears et Radford 2016 ; Bertucci et al. 2016).

— **Pression acoustique moyenne** : moyenne des niveaux sonores sur une période donnée (e.g. par heure, par jour ou par mois). Cette métrique permet de quantifier

l'énergie acoustique à long terme et de détecter des gradients d'activité sonore, qu'ils soient biologiques, anthropiques ou abiotiques (Harris, Shears et Radford 2016 ; Bertucci et al. 2016).

Cette distinction claire entre métriques, fenêtres temporelles et représentations synthétiques permet de structurer l'analyse acoustique et de relier rigoureusement les données brutes aux interprétations écologiques. Elle conditionne également la comparabilité des résultats entre les jeux de données, les périodes ou les sites.

Les méthodes statistiques

Le point de départ de toute analyse PAM repose souvent sur des méthodes descriptives, permettant de visualiser les structures présentes dans les données : rythmes journaliers, saisonnalité, différences inter-sites ou effets d'un changement de protocole (Wiggins et Hildebrand 2007 ; Kent, Marques et Harris 2022). Ces méthodes incluent le calcul de taux de détection (e.g. par heure, jour ou mois), des représentations graphiques (e.g. courbes temporelles, spectrogrammes moyens, *heatmaps*) et, lorsque cela est pertinent, des analyses multivariées comme l'analyse en composantes principales (PCA).

Lorsque les données se diversifient et couvrent de larges séries temporelles et spatiales, il devient possible d'adopter des approches plus avancées, basées sur un cadre théorique solide : le concept de niche écologique. En reliant la présence acoustique aux variables environnementales pertinentes (e.g. température, profondeur, dynamique des proies, activité anthropique), ces approches permettent de modéliser la distribution spatiale et temporelle des espèces de manière prédictive, en identifiant les conditions favorables à leur activité ou à leur présence (Merchant et al. 2015 ; Fleishman et al. 2023).

Dans ce contexte, les modèles de distribution d'espèces (SDMs) offrent un outil statistique formalisé pour exploiter ces relations en reliant directement la distribution observée à la niche écologique potentielle de l'espèce (Elith et Leathwick 2009 ; Shabangu, Munoz et al. 2024). Ces modèles permettent de tester des hypothèses écologiques sur les facteurs influençant la présence ou l'activité des espèces, tels que les migrations saisonnières, les préférences d'habitat, ou les réponses comportementales face aux pressions anthropiques (bruit sous-marin, activités de pêche) (Shabangu, Yemane et al. 2017). Ils constituent ainsi un prolongement naturel des analyses descriptives, en traduisant les patrons de pré-

sence acoustique observés en mécanismes écologiques interprétables. Leur application aux données PAM est centrale dans cette thèse et fait l'objet d'un développement approfondi dans la section « Modéliser la présence acoustique : les outils statistiques» du chapitre 2.

Chapitre 1

Optimisation des stratégies de sous-échantillonnage

1.1 Introduction

Le suivi par acoustique passive (PAM) des cétacés repose sur la capacité à enregistrer les sons de l'environnement marin, permettant de détecter et caractériser les vocalisations produites par différentes espèces. L'enregistrement continu, c'est-à-dire la captation des signaux acoustiques 24 heures sur 24 et 7 jours sur 7, constitue une stratégie optimale pour capter l'ensemble des signaux présents, sans lacune temporelle. Toutefois, cette approche se heurte à des contraintes logistiques majeures : la consommation énergétique, les volumes de stockage particulièrement importants, le temps de traitement et les coûts opérationnels, notamment dans le cadre de suivis de longue durée, multi-sites ou réalisés en milieux isolés.

Pour pallier ces limites, de nombreux protocoles PAM intègrent un *duty cycle*, c'est-à-dire un schéma de sous-échantillonnage temporel dans lequel l'enregistrement acoustique est activé uniquement durant certaines plages horaires, alternant automatiquement les périodes d'enregistrement et les périodes non enregistrées. Par exemple, un enregistreur peut être programmé pour enregistrer 10 minutes toutes les 30 minutes, réduisant ainsi par un facteur trois l'effort de collecte, le volume de données produites et le besoin de stockage associé. La notion de *duty cycle*, ainsi que ses modalités de mise en œuvre, sont définies et détaillées dans la section « Réglages techniques des enregistreurs acoustiques » de l'introduction de cette thèse.

L'intégration d'un *duty cycle* poursuit plusieurs objectifs :

— **Augmenter la durée des campagnes**, en économisant la mémoire et les batteries disponibles sur chaque dispositif ;

— **Étendre la couverture spatiale**, en permettant le déploiement simultané de plusieurs enregistreurs sur différents sites avec les mêmes ressources ;
— **Réduire le volume de données** à stocker et à analyser. Les enregistrements acoustiques génèrent des fichiers très volumineux, ce qui complique leur gestion ;
— **Diminuer la fréquence des interventions sur site** permet de réduire les coûts logistiques et opérationnels liés aux manutentions et aux déplacements.

L'intégration d'un *duty cycle* dans un protocole PAM peut néanmoins impacter plusieurs aspects cruciaux de la collecte et de l'analyse des données :

— **Qualité des enregistrements** : chaque redémarrage de l'instrument demande un temps de stabilisation pendant lequel la qualité sonore peut être dégradée, réduisant ainsi la qualité globale des données ;
— **Probabilité de détection des événements acoustiques** : en enregistrant par intermittence, certains signaux peu fréquents et donc certaines espèces peuvent ne pas être enregistrées ;
— **Performance des détecteurs automatiques** : certains algorithmes de détection sont optimisés pour un enregistrement continu. Le sous-échantillonnage peut modifier la qualité et la continuité des signaux, réduisant l'efficacité et la fiabilité des détecteurs ;
— **Représentation temporelle des dynamiques biologiques** : la fenêtre d'observation réduite peut biaiser les estimations de présence, d'activité ou de diversité des espèces à différentes échelles temporelles ;
— **Comparabilité des données** : la modification de la probabilité de détection complique la comparaison entre sites, périodes ou espèces lorsque le *duty cycle* diffère.

Ainsi, bien que le *duty cycle* soit un outil précieux pour optimiser la collecte, il convient d'en mesurer précisément les impacts afin d'adapter au mieux la conception des campagnes de suivi.

Cette réflexion conduit à la question centrale de ce chapitre :

Dans quelle mesure la mise en œuvre d'un *duty cycle* influence-t-elle la fidélité des résultats PAM par rapport aux données obtenues en enregistrement continu ?

Pour répondre à cette question, ce chapitre propose une étude sur les effets de l'utilisation d'un *duty cycle* dans les protocoles PAM, ses impacts potentiels et les méthodes permettant d'évaluer ses conséquences. Il s'organise selon les étapes suivantes : (i) une mise en contexte basée sur un état de l'art, comprenant une section sur les pratiques actuelles ainsi qu'une présentation des études antérieures ayant évalué les effets du *duty cycle* ; (ii) une présentation des jeux de données utilisés pour cette évaluation ; (iii) une description du cadre théorique et méthodologique développé permettant d'analyser les effets du *duty cycle* sur les représentations descriptives ; (iv) la présentation des résultats à travers deux publications scientifiques :

— **Etude 1** : Effects of duty cycle on passive acoustic monitoring metrics : The case of blue whale songs, publiée en 2024 dans la revue *Journal of the Acoustical Society of America (JASA)* ;

— **Etude 2** : Effects of duty cycle on passive acoustic monitoring metrics : The case of odontocetes vocalizations, actuellement soumise et en cours d'évaluation dans la revue *JASA* également.

(v) une synthèse des résultats obtenus visant à guider l'intégration du *duty cycle* dans les protocoles PAM dans une perspective de suivi à grande échelle.

1.1.1 État de l'art

Pratiques actuelles des stratégies de *duty cycle* en suivi par acoustique passive

Dans le contexte du PAM, le recours aux stratégies de *duty cycle* s'est fortement développé pour répondre à diverses contraintes techniques et logistiques. Une revue croisée de la littérature scientifique internationale et des rapports techniques nationaux révèle une grande hétérogénéité dans les pratiques employées. Les paramètres de *duty cycle*, tels que la durée des périodes d'écoute ou le pourcentage d'écoute, varient considérablement selon les projets, allant de quelques minutes à une heure, avec des pourcentages d'écoute très divers. Dans certains cas, toutes les informations relatives au *duty cycle* ne sont pas communiquées :

certaines études ne précisent que la durée des périodes d'écoute ou seulement le pourcentage d'enregistrement, ce qui complique encore davantage la comparaison et l'évaluation des stratégies employées. Pour illustrer la diversité des approches en matière de *duty cycle*, cinq exemples représentatifs ont été sélectionnés pour chacun des deux types de sources étudiées : publications scientifiques internationales (Tableau 1.1) et rapports techniques nationaux (Tableau 1.2).

Les exemples de publications scientifiques ont été choisis en veillant à couvrir une grande diversité de contextes, incluant différents pays, espèces ciblées et types d'enregistreurs, afin d'être aussi exhaustifs que possible. Concernant les rapports techniques nationaux, les études présentées proviennent de plusieurs bureaux d'études français spécialisés, reflétant ainsi la variété des pratiques dans le cadre de projets opérationnels, notamment ceux liés à l'éolien offshore.

Cette variabilité est majoritairement le fruit de décisions empiriques, souvent sans justification scientifique explicite ni prise en compte des caractéristiques comportementales des espèces ciblées ou des objectifs écologiques du suivi. En effet, aucune des études consultées, qu'il s'agisse d'articles scientifiques ou de documents techniques, ne justifie explicitement ses choix de *duty cycle*, ni n'évalue rigoureusement l'impact de cette stratégie sur la qualité des données acoustiques, la représentativité temporelle ou la robustesse des indicateurs biologiques dérivés.

Cette revue met ainsi en lumière le besoin d'un encadrement méthodologique plus rigoureux afin de guider les décisions relatives au *duty cycle* et d'assurer la fiabilité des suivis acoustiques.

TABLE 1.1 – Exemples d'études internationales de PAM de cétacés utilisant un *duty cycle* dans l'acquisition des données.

Étude	**Espèce(s) ciblée(s)**	**Type d'enregistreurs**	**Duty cycle utilisé**	**Remarques / choix et impacts du duty cycle**
VAN GEEL et al. 2022	Mysticètes, odontocètes	Sylence-LP-440	25 min ON / 35 min OFF	Duty cycle choisi pour limiter le volume de données ; impact sur la qualité des détections non évalué.
STANISTREET, NOWACEK, BAUMANN-PICKERING et al. 2017	Baleines à bec (plusieurs espèces)	AMAR / HARP	9% à 100% selon site et période	Duty cycle adapté aux contraintes logistiques ; impact potentiel sur certaines espèces mentionné sans quantification.
RISCH, WILSON et al. 2019	Baleine de Minke	SM2M + HTI-94	10 min ON / 20 min OFF (33%)	Duty cycle hérité du protocole ECOMMAS (axé initialement sur les grands dauphins) ; aucune justification spécifique ni évaluation d'impact.
DAVIS et al. 2017	Baleine noire de l'Atlantique Nord	HARP, MARU, AMAR, HARU, Guardbuoy	12 à 95% (majorité en continu)	Duty cycle variable selon projets partenaires ; périodes d'écoute non précisées ; pas d'analyse d'impact.
STAFFORD, CITTA et al. 2021	Baleine boréale et béluga	Aural-M2	5 à 15 min ON / 60 min OFF	Duty cycle choisi pour autonomie ; détection du béluga possiblement affectée mais non évaluée.

TABLE 1.2 – Exemples de rapports nationaux de PAM de cétacés utilisant un duty cycle dans l'acquisition.

Étude	**Espèces ciblées**	**Type d'enregistreur / capteur**	**Duty cycle utilisé**	**Remarques / choix et impacts du duty cycle**
BELLANGER et al. 2023	Delphinidés et Phocoenidés	SoundTrap HT 600 HF	3 min ON / 7 min OFF (30%)	Duty cycle choisi pour concilier capacité de stockage et fréquence des relevages (2,5 mois) ; justification uniquement logistique, sans évaluation scientifique de l'impact.
ROCK et al. 2023	Delphinidés et Phocoenidés	SoundTrap HT 600 HF	30%	Duty cycle fixé pour assurer 90 jours d'autonomie ; durée des périodes d'écoute et impact non précisés.
DRIRA et al. 2021	Cétacés (espèces non précisées)	RTsys EA-SD14 / Sylence EA-SDALP + HDI-99-HF	5 min ON / 10 min OFF (50%)	Duty cycle choisi pour équilibre entre autonomie (15–26 jours) et représentativité des signaux ; impact non évalué.
CHOMPRET et al. 2021	Cétacés (divers)	Wildlife Acoustics SM3M / SM2M	20 min ON / 40 min OFF (30%)	Duty cycle de 30% Choix équilibré entre durée de suivi et contraintes logistiques.
GERVAISE et al. 2018	Diverses espèces marines	RTSYS SDA 14 / SM2M / Hydrophones HTI92	25–100% (15 à 60 min ON par heure selon l'enregistreur)	Duty cycle variable selon l'enregistreur et contexte ; choix dictés par autonomie et logistique.

État des connaissances scientifiques sur l'impact du *duty cycle*

En complément de cette diversité de pratiques opérationnelles, quelques travaux scientifiques ont cherché à évaluer de manière quantitative les conséquences de l'intégration d'un *duty cycle* sur la qualité des données acquises et la fiabilité des représentations descriptives associées. Ces études restent toutefois rares, ciblant un nombre limité d'espèces et de contextes, et s'attachent le plus souvent à des indicateurs ponctuels ou de court terme. Par exemple, Thomisch, Boebel, Zitterbart et al. 2015 ont étudié l'effet du *duty cycle* sur les chants de baleine bleue Antarctique en testant plusieurs pourcentages d'écoute (de 50% à 1,67%). Ils ont évalué notamment la précision de l'estimation de la présence acoustique quotidienne ainsi que le taux d'appels, démontrant une diminution progressive de la détection avec la diminution du pourcentage d'écoute, impactant négativement ces métriques. De leur côté, Stanistreet, Nowacek, Read et al. 2016 se sont focalisés sur la proportion moyenne de journées correctement évaluées pour la présence acoustique de différentes espèces de baleines à bec, sur plusieurs sites. Ils concluent que des périodes d'écoute courtes et fréquentes sont les plus efficaces pour estimer la présence quotidienne.

Deux études sur les orques ont apporté des informations complémentaires. Riera, Ford et Ross Chapman 2013 ont examiné les effets de deux pourcentages d'écoute : 33% et 67% sur des métriques comportementales, telles que le nombre de rencontres acoustiques, leur durée, ainsi que l'identification des groupes. Ils ont observé une baisse significative de la détection des vocalisations et de l'identification positive des groupes avec un pourcentage plus faible, ce qui conduit à une sous-estimation du temps de présence. Enfin, Rand, Wood et Oswald 2022 ont exploré une large gamme de pourcentages d'écoute (1% à 67%), évaluant l'impact sur la présence acoustique quotidienne des orques, les variations journalières dans la détection horaire des appels, ainsi que sur la détection et l'identification des séquences comportementales. Ils concluent que la détection de la présence quotidienne est peu affectée par un *duty cycle*, mais que des périodes d'écoute de 15 à 30 minutes et des pourcentages d'écoute $\geq 33\%$ sont nécessaires pour des calculs précis de la durée des séquences comportementales.

Cependant, ces recherches présentent plusieurs limites. D'une part, elles ne concernent qu'un nombre restreint d'espèces, aucune n'étant dédiée aux dauphins et marsouins, bien qu'il s'agisse de groupes très fréquemment suivis par acoustique passive. D'autre part, elles s'appuient principalement sur des données issues d'enregistreurs large bande, sans évaluer

les effets spécifiques du *duty cycle* sur des dispositifs fondés sur des algorithmes de détection automatisée embarquée, tels que les F-PODs. Par ailleurs, la plupart de ces travaux se sont concentrés sur des représentations descriptives de la donnée à court terme (présence quotidienne, nombre d'occurrences, etc.) ou comportementales (durée des séquences, identification de groupes), sans s'attacher à l'impact du *duty cycle* sur des tendances temporelles de plus long terme, comme les variations saisonnières ou interannuelles de la présence acoustique. Enfin, les tests menés dans ces études n'ont que rarement été réalisés sur des sites présentant des niveaux de présence contrastés, ce qui limite l'évaluation de la robustesse des métriques à travers différents contextes écologiques.

Afin de répondre à ces limitations et d'éclairer les choix de protocoles PAM, ce chapitre propose deux volets d'analyse complémentaires visant à évaluer l'effet de différents *duty cycles* sur la fiabilité des indicateurs de présence acoustique à différentes échelles temporelles. Le premier volet porte sur les chants spécifiques de trois populations acoustiques de baleines bleues de l'océan Indien et d'une vocalisation commune, à partir de données continues sous-échantillonnées artificiellement. Il examine l'impact du *duty cycle* sur 3 représentations descriptives des données : le patron saisonnier, l'indice de diversité acoustique et le taux d'émission quotidien. Le second volet se concentre sur les odontocètes, en particulier les delphinidés et les marsouins communs, et explore l'effet du *duty cycle* sur les clics et les sifflements détectés via des enregistreurs large bande et des F-PODs. Il s'appuie sur des métriques temporelles multi-échelles : nombre de minutes positives à la détection par jour (DPM) et nombre d'heures positives à la détection par jour (DPH) pour évaluer la présence acoustique quotidienne et saisonnière. Les deux analyses ont été menées sur des sites aux niveaux de détection contrastés, afin d'identifier les configurations de sous-échantillonnage les plus robustes selon les espèces, les signaux étudiés, les instruments utilisés et le contexte écologique.

1.1.2 Les jeux de données utilisés

Chants des mysticètes dans l'océan Indien : le réseau OHASISBIO

Le réseau Observatoire HydroAcoustique de la SISmicité et de la BIOdiversité (OHASISBIO) a été mis en place pour surveiller à la fois l'activité sismique et la biodiversité des grandes baleines dans le Sud-Ouest de l'océan Indien (Royer 2009). Il repose sur

le déploiement d'enregistreurs acoustiques autonomes, capables d'enregistrer les basses fréquences sur de longues périodes. Ces dispositifs permettent de détecter les signaux produits par les événements géophysiques, mais aussi par les mysticètes, comme les baleines bleues, dont les chants se situent dans des gammes de fréquences similaires (généralement inférieures à 100 Hz). Le réseau couvre une vaste zone géographique, s'étendant de 24°S à 56°S en latitude et de 52°E à 83°E en longitude, et comprend dix sites de mouillage (Fig. 1.1).

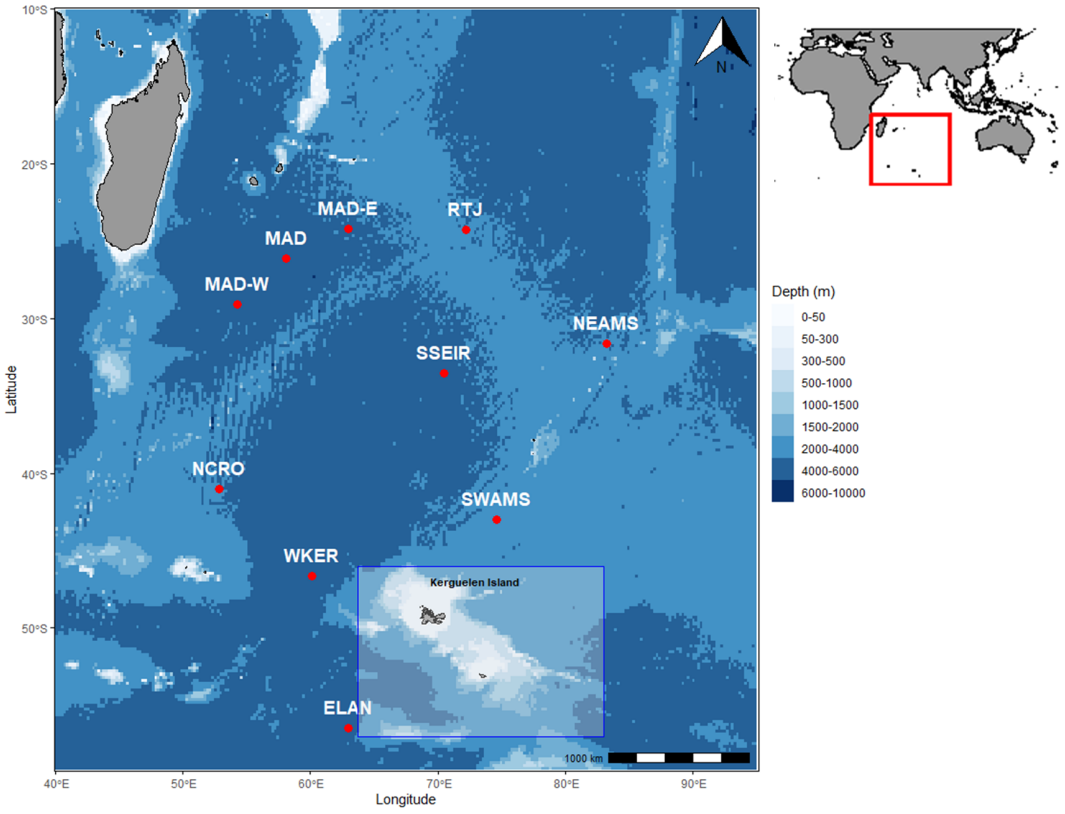


FIGURE 1.1 – Localisation des sites de mouillage du réseau OHASISBIO dans le sud ouest de l'océan Indien représenté par les points rouge. La zone ombrée en bleu avec contour bleu représente la zone de l'île de Kerguelen avec son plateau.

Il a fonctionné entre fin 2009 et début 2023, avec des périodes d'enregistrement spécifiques à chaque site selon les campagnes de déploiement et de récupération des instruments (Fig. 1.2). Chaque ligne de mouillage est constituée d'un lest, d'un largueur acoustique, d'une ligne de longueur ajustable, et d'une bouée immergée contenant l'enregistreur avec

son hydrophone (Fig. 1.3). L'hydrophone est positionné entre 1 000 et 1 300 mètres de profondeur, dans le canal SOFAR (voir encadré ci-dessous). Les enregistrements sont réalisés de manière continue à une fréquence d'échantillonnage de 240 Hz. La sensibilité des hydrophones est de -163.5 dB ref. 1V/μPa, sauf pour les mouillages doubles déployés à SWAMS en 2015 et 2016, pour lesquels elle était de -168.6 dB ref. 1V/μPa. L'autonomie des batteries permet une année complète d'enregistrement, avec une récupération planifiée chaque année en janvier ou février, lors de la rotation annuelle du R/V Marion Dufresne 2 dans les Terres australes et antarctiques françaises (TAAF) [1]. Cependant, des arrêts prématurés peuvent survenir en raison d'une décharge anticipée des batteries.

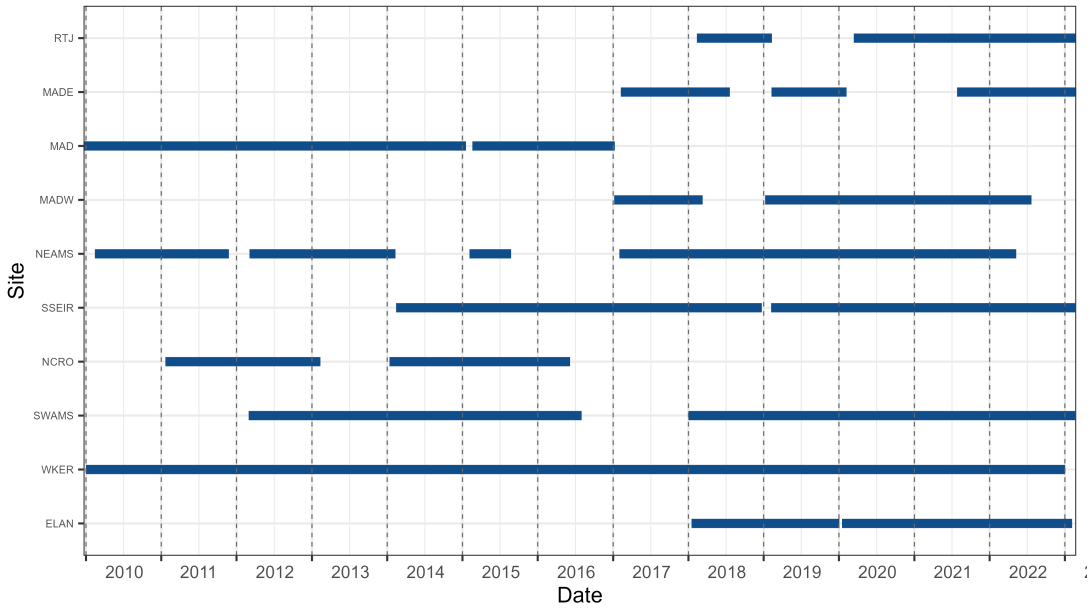


FIGURE 1.2 – Périodes d'enregistrement des sites du réseau OHASISBIO classé du site le plus au Nord au plus au Sud.

1. Territoires français situés dans l'hémisphère sud, comprenant les îles Kerguelen, Crozet, Saint-Paul et Amsterdam, ainsi que la Terre Adélie en Antarctique. Ces territoires sont administrés par un préfet et sont principalement dédiés à la recherche scientifique et à la conservation de la biodiversité.

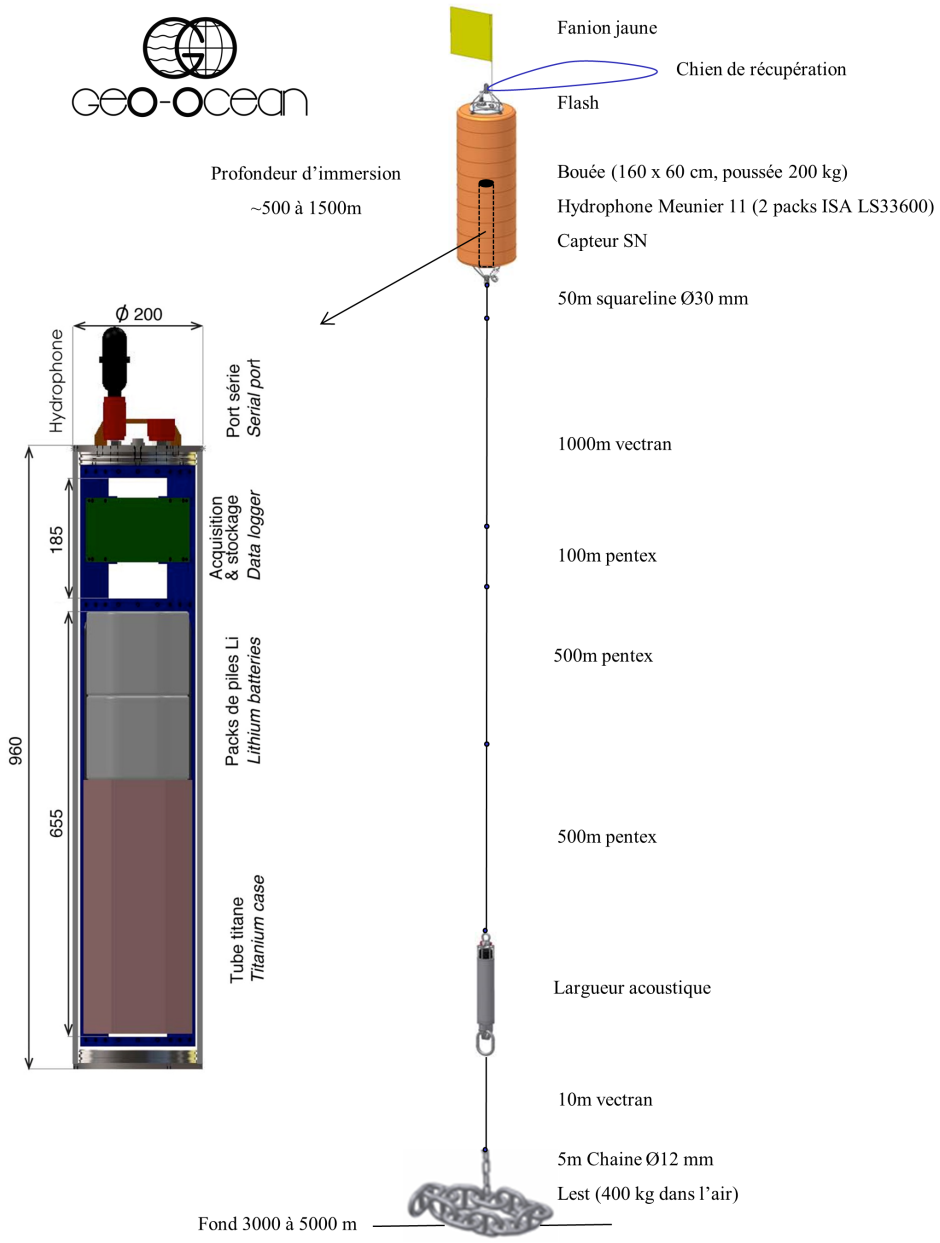


FIGURE 1.3 – Schéma d'une ligne de mouillage du réseau OHASISBIO.

Le canal SOFAR

Le **canal SOFAR** (pour *SOund Fixing And Ranging*, également appelé *deep sound channel*) est une couche océanique, généralement située entre 800 et 1 500 mètres de profondeur, où la vitesse du son atteint un minimum. Cette configuration résulte des effets combinés de la température et de la salinité (qui diminuent rapidement avec la profondeur) et de la pression (qui augmente linéairement), créant un profil en « U » de la vitesse du son (Fig. 1.4).

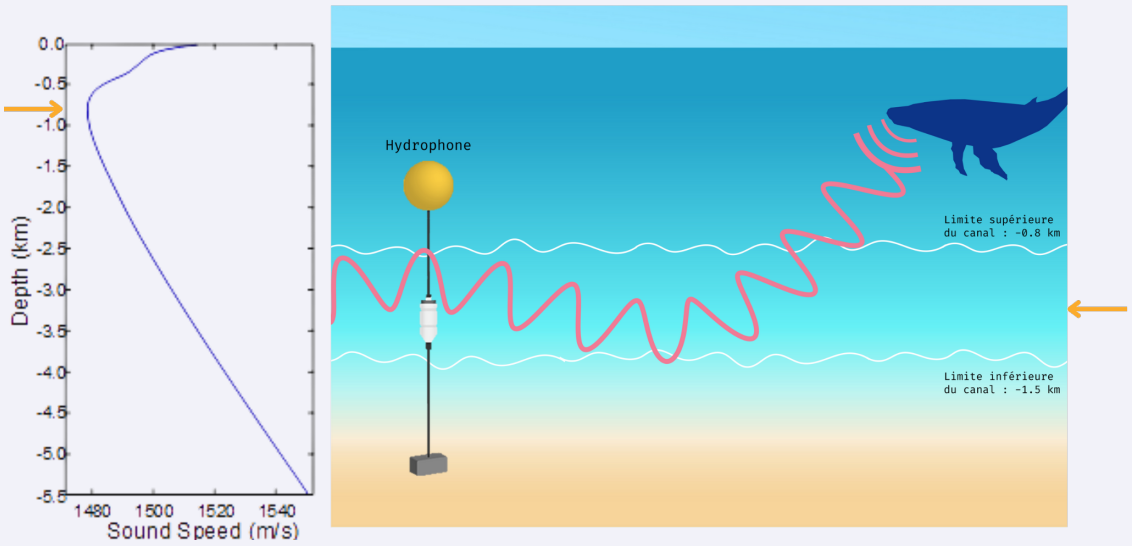


FIGURE 1.4 – A gauche : Vitesse du son en fonction de la profondeur. Données issues de mesures effectuées au nord de Hawaï dans l'océan Pacifique et tirées de l'édition 2005 de l'Atlas Mondial de l'Océan (World Ocean Atlas). A droite : Illustration du canal SOFAR.

Les ondes acoustiques basse fréquence s'y trouvent piégées par réfraction : lorsqu'elles s'éloignent de l'axe du canal (vers la surface ou le fond), elles sont déviées de nouveau vers celui-ci. Ainsi, elles se propagent sur plusieurs centaines de kilomètres avec peu d'atténuation. Les hydrophones déployés dans ce canal exploitent ce mécanisme pour maximiser la portée de détection acoustique (LUTRON 2002).

Deux sous-espèces de baleines bleues sont présentes dans cette région : la baleine bleue Antarctique (ANT BW) (*B. m. intermedia*), et la baleine bleue pygmée (*B. m. brevicauda*). Deux populations distinctes de baleines bleues pygmées ont été identifiées dans les enregistrements acoustiques : la baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW) et la baleine bleue pygmée du sud-est de l'océan Indien (SEIO PBW) (Fig. 1.5). Chaque sous-espèce émet un chant stéréotypé distinct, ainsi que des signaux non stéréotypés communs (*D-calls*) (TORTEROTOT, ROYER et SAMARAN 2019).

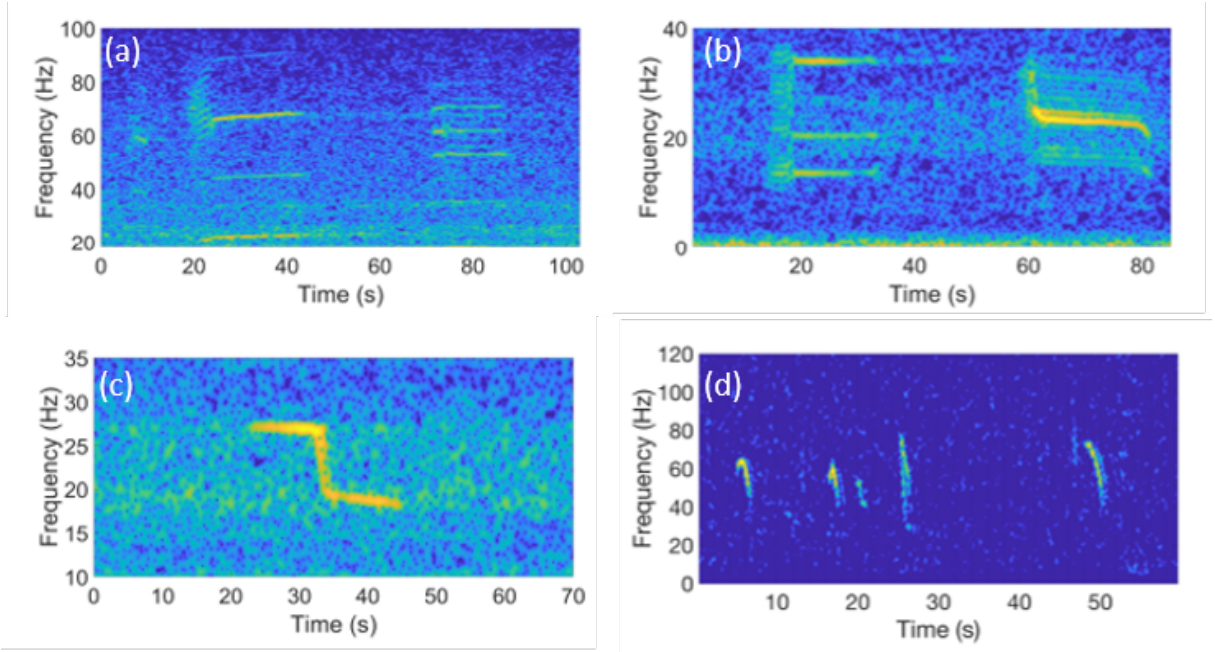


FIGURE 1.5 – Spectrogrammes illustrant (a) des chants SEIO PBW, (b) des chants SWIO PBW, (c) des chants ANT BW et (d) de D-calls. Enregistrements réalisés dans la zone d'étude du réseaux OHASISBIO (Océan Indien sud).

Les chants des baleines bleues ont été détectés automatiquement à l'aide du *Sparse Representation-based Detector* (SRD), un algorithme basé sur l'apprentissage par dictionnaire et la représentation parcimonieuse (SOCHELEAU et SAMARAN 2018). Ce détecteur a été conçu pour être à la fois fiable, en limitant les fausses alarmes, et rapide, afin de traiter de longues séries d'enregistrements en un temps raisonnable. Des dictionnaires spécifiques sont définis pour chaque population acoustique et pour les différents segments de chants ciblés, ce qui permet de prendre en compte la variabilité des vocalisations (TORTEROTOT, ROYER et SAMARAN 2019).

Dans le cadre de ce chapitre et des résultats présentés, les détections ont été exploitées sous forme de nombre de vocalisations par unité de temps (par jour), fournissant une mesure quantitative précise de l'activité acoustique. De plus, seuls trois sites du réseau ont été sélectionnés pour l'analyse en raison de la présence régulière de plusieurs populations acoustiques de baleines bleues : à l'ouest de l'île Kerguelen (WKER), au nord-est des îles Saint-Paul et Amsterdam (NEAMS, en service de 2010 à 2019), et au sud de la crête sud-est de l'océan Indien (SSEIR, en service de 2014 à 2018).

Vocalisations d'odontocètes : le projet CETIROISE et les données d'Irlande

Les données acoustiques utilisées pour l'étude des odontocètes proviennent majoritairement du projet CETIROISE (Samaran, Torterotot et Béesau 2023). Ce programme se concentre sur les espèces présentes en mer d'Iroise, en particulier le grand dauphin (*Tursiops truncatus*), le dauphin commun (*Delphinus delphis*) et le marsouin commun (*Phocoena phocoena*). L'étude se focalise sur les clics d'écholocation émis par les marsouins communs et sur les clics et sifflements émis par les dauphins (Fig. 1.6). Il est important de noter qu'il n'est pas possible de différencier les sifflements et les clics des différentes espèces de dauphins à partir des données acoustiques utilisées ici (Van Geel et al. 2022). Ainsi, les clics et sifflements produits par ces derniers seront qualifiés globalement sous le terme « delphinidés » dans la suite de l'étude.

Les enregistrements ont été collectés grâce à des dispositifs de suivi déployés sur sept sites dans la mer d'Iroise, au large de l'île d'Ouessant, dans la zone approximative délimitée par les coordonnées GPS suivantes : 48.3°N - 48.8°N et 4.5°W - 5.5°W. Ces sites ont été sélectionnés en fonction de leur profondeur, allant de 25 à 120 mètres. Selon la profondeur, deux types de systèmes de mouillage ont été utilisés : un dispositif adapté aux sites côtiers (sites E et F, entre 25 et 35 mètres), permettant une intervention par des plongeurs pour relever les instruments, et un autre système conçu pour les sites hauturiers (sites A, B, C, D et G, entre 60 et 120 mètres), pour lesquels un système de largage acoustique permet de remonter l'intégralité du dispositif sans intervention de plongeurs. Les enregistreurs étaient relevés environ tous les trois mois afin de récupérer les données et d'assurer leur bon fonctionnement. Dans le cadre de cette étude, seuls les sites F et G ont été retenus en raison de leur fréquentation acoustique contrastée (Fig. 1.7).

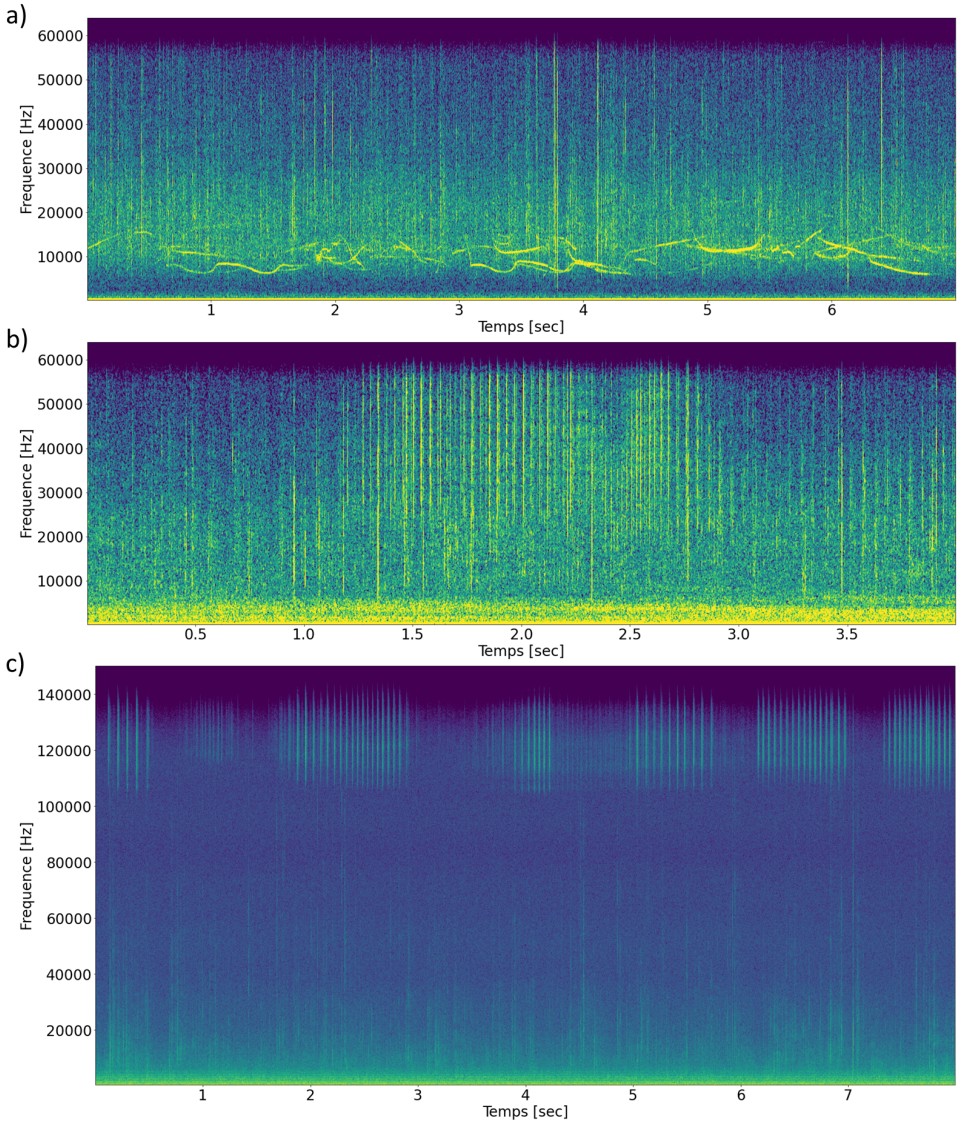


FIGURE 1.6 – Spectrogrammes illustrant (a) de sifflements de delphinidés, (b) de clics de delphinidés et (c) de clics de marsouins communs. Enregistrements réalisés en Mer d'Iroise.

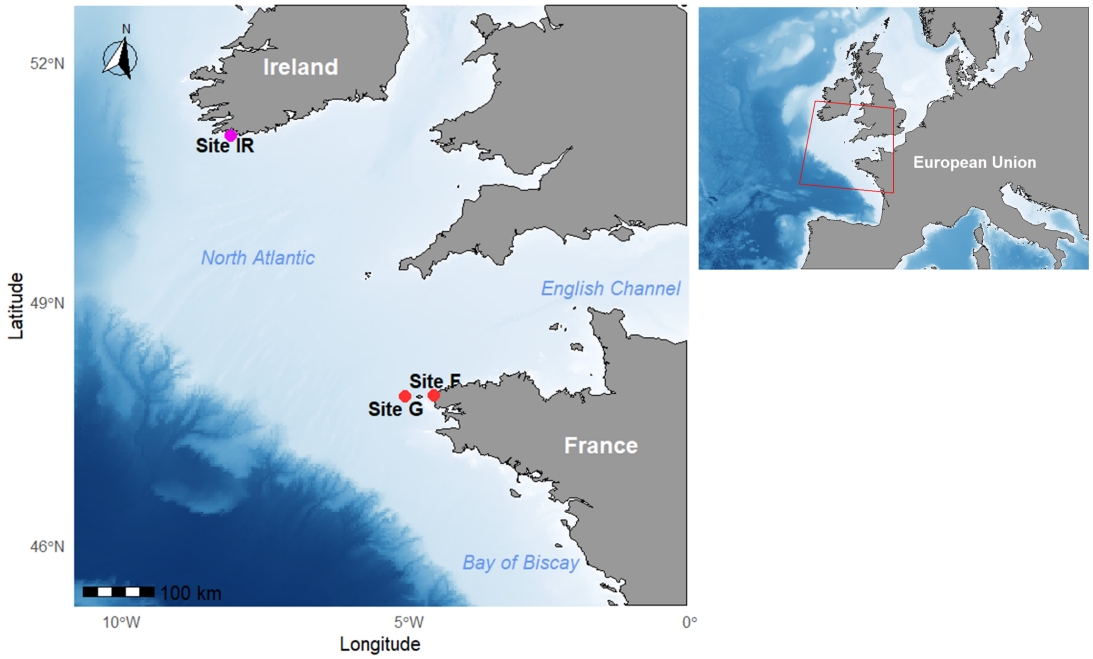


FIGURE 1.7 – Carte de la zone d'étude montrant les trois sites de mouillage. Le marqueur violet indique le site irlandais (Site IR), équipé d'un enregistreur F-POD. Les marqueurs rouges indiquent les sites français (Site F : côtier ; Site G : au large), chacun équipé d'un enregistreur large bande et d'un F-POD.

Deux instruments ont été utilisés dans le cadre de ce projet :

— **L'enregistreur acoustique large bande (Sylence – LP 440, RTSYS)** : cet enregistreur est équipé d'un hydrophone HTI – 99HF et permet d'enregistrer une large gamme de fréquences allant de 2 Hz à 64 kHz, avec une fréquence d'échantillonnage de 128 kHz. L'enregistrement est continu grâce à un *duty cycle* de 100%, et les données sont sauvegardées sur des cartes mémoire de 1 To. Ce dispositif couvre l'ensemble des vocalisations des cétacés, y compris celles des odontocètes. Les données collectées nécessitent un traitement post-acquisition pour extraire des informations sur la présence et l'activité acoustique des odontocètes.

— **Le détecteur de clics d'écholocation (F-POD D12, Chelonia Ltd)** : le F-POD est conçu pour détecter les clics d'écholocation des odontocètes (dauphins, marsouins). Il fonctionne dans une plage de fréquence de 17 à 220 kHz et enregistre les clics sous forme de détections sans conserver le signal brut. Ces détections sont ensuite classées par un algorithme de traitement développé par le constructeur (KERNO-F classifier).

À ces données françaises s'ajoutent celles issues d'un site en Irlande (site IR), où un F-POD a été déployé entre 2021 et 2022 dans le cadre d'une étude comparative entre C-POD et F-POD (Todd, Kavanagh et al. 2023). Ce dispositif était fixé à environ 5 mètres au-dessus du fond marin à 23 mètres de profondeur, sur un mouillage ancré et marqué par une bouée de surface. Le F-POD irlandais était récupéré et reprogrammé tous les 4 à 6 mois selon les conditions météorologiques et la facilité d'accès au site. Les enregistreurs utilisés ont été calibrés et standardisés par le fabricant (Fig. 1.7).

La détection automatique des signaux d'odontocètes est réalisée à l'aide d'algorithmes adaptés à la nature des vocalisations de ces espèces. Contrairement au F-POD, qui cible principalement les clics, les enregistrements large bande doivent être post-traités pour identifier les sifflements et autres vocalisations tonales. Pour cela, le logiciel `PAMGuard` (version 2.02.15) a été utilisé avec le module *Whistle and Moan Detector* (Gillespie, Caillat et al. 2013). Cet algorithme détecte les vocalisations tonales en se basant sur des critères de continuité spectrale et d'intensité, et il a été configuré à l'aide de paramètres spécifiques. Les clics de delphinidés et de marsouins communs enregistrés par les F-POD aux trois sites (G, F et IR) ont été traités avec le logiciel `F-POD.exe` (version 1.1) utilisant le classificateur KERNO-F (Chelonia Ltd. 2022). Ce classificateur distingue automatiquement les trains de clics en groupes d'espèces selon leurs caractéristiques spectrales : « NBHF » pour les clics étroits à haute fréquence du marsouin commun, et « Other cet » pour les delphinidés.

Enfin, l'activité acoustique des odontocètes a été quantifiée à l'aide des métriques DPM et DPH, qui correspondent respectivement au nombre de minutes ou d'heures par jour avec présence acoustique confirmée. Les détections sont agrégées sur des intervalles temporels fixes (par jour), ce qui fournit une mesure standardisée comparable entre sites et instruments (enregistreurs large bande et F-POD).

1.2 Méthodologie : explorer l'effet du *duty cycle* par sous-échantillonnage artificiel

1.2.1 Principe général

Une même approche a été appliquée aux deux études (étude 1 : mysticètes et étude 2 : odontocètes) pour évaluer l'impact du *duty cycle*. Elle consiste à créer artificiellement des jeux de données sous-échantillonnés à partir d'enregistrements initialement continus. Chaque schéma de sous-échantillonnage est défini par deux paramètres principaux : une période d'écoute L_p (en minutes), correspondant à la durée effective d'enregistrement à chaque cycle, et un pourcentage d'écoute D (en %), exprimant la proportion de temps enregistré au sein de chaque cycle temporel. Ces deux paramètres déterminent ensemble la durée et la répartition des périodes d'écoute dans le temps. Les représentations descriptives (détaillées dans la section suivante) ont été exprimées dans des métriques adaptées à chaque groupe taxonomique : pour les odontocètes, les taux de détection sont donnés en nombre de minutes positives à la détection par jour (DPM) et nombre d'heures positives à la détection par jour (DPH) ; pour les mysticètes, les taux sont exprimés en nombre de vocalisations par jour.

1.2.2 Construction des jeux de données sous-échantillonnés

Deux approches distinctes ont été mises en œuvre pour construire les jeux de données sous-échantillonnés, en fonction du type d'enregistreur utilisé : enregistreurs large bande ou F-POD.

Pour les enregistreurs large bande, les données enregistrées en continu ont été sous-échantillonnées selon 21 régimes distincts, combinant des périodes d'écoute L_p à durée variable (de 1 minute à 6 heures) avec des pourcentages d'écoute D de 50 %, 33 % et 25 %. Pour chaque combinaison, plusieurs phases d'écoute non superposées ont été simulées (deux pour 50 %, trois pour 33 %, quatre pour 25 %) afin d'évaluer la variabilité liée au positionnement temporel des fenêtres d'écoute. En effet, pour un schéma donné, par exemple, 5 minutes toutes les 15 minutes, la fenêtre d'enregistrement peut être positionnée au début, au milieu ou à la fin du cycle. L'ensemble des positions possibles a été testé, et

les résultats ont ensuite été moyennés pour chaque configuration, ce qui a permis d'estimer un effet moyen du sous-échantillonnage.

Des règles d'inclusion spécifiques ont été appliquées en fonction de la nature des signaux :

— **Odontocètes** : toute minute ou heure de détection tombant dans une période d'écoute était conservée.

— **Mysticètes** : seules les détections dont la vocalisation complète se trouvait à l'intérieur de la période d'écoute étaient retenues, afin de respecter les contraintes du détecteur automatique utilisé qui nécessite d'avoir l'entièreté du signal.

Pour les clics de delphinidés et de marsouins enregistrés via les F-PODs, le sous-échantillonnage repose sur quatre régimes prédéfinis par le constructeur, caractérisés par différents pourcentages d'écoute D de 50%, 33%, 20% ou 10%, avec une période d'écoute fixe $L_p = 1$ min. Contrairement aux configurations flexibles appliquées aux enregistreurs large bande, les F-PODs sont programmés pour enregistrer systématiquement la première minute de chaque cycle temporel. Aucun moyennage entre plusieurs réalisations possibles n'est alors effectué. Le détecteur KERNO-F exploite le contexte temporel autour des clics en analysant les extraits situés avant et après la minute évaluée. Afin de préserver la cohérence des résultats malgré le sous-échantillonnage, il ne serait pas pertinent de supprimer les périodes non enregistrées en post-traitement. Le détecteur KERNO-F a donc été réappliqué directement sur les extraits sous-échantillonnés, en conservant les mêmes paramètres que ceux utilisés pour la détection sur des données continues. Cette approche permet de respecter la logique du détecteur et de garantir la fiabilité des détections.

1.2.3 Analyses comparatives et approche statistique

Pour chaque schéma de sous-échantillonnage, les représentations descriptives sélectionnées ont été comparées à celles issues des données continues selon des méthodes développées :

— **Patron saisonnier** (mysticètes et odontocètes) : corrélation (coefficient R^2) entre séries temporelles sous échantillonnées et non sous échantillonnées.

— **Taux journalier de détection** (mysticètes et odontocètes) : calcul du ratio entre le taux moyen de présence acoustique sous échantillonné et le taux sur données non

sous échantillonnées.

— **Patron journalier** (odontocètes) : corrélation (coefficient R^2) entre séries temporelles sous échantillonnées et non sous échantillonnées.

— **Richesse en types de signaux** (mysticètes) : proportion de jours où la richesse estimée par sous-échantillonnage est égale à celle des données non sous échantillonnées.

Cette méthode permet d'évaluer les compromis entre l'effort d'enregistrement et la fiabilité des représentations descriptives, en fonction des objectifs spécifiques à chaque groupe taxonomique. Les résultats issus de ces analyses sont présentés dans les publications ci-dessous.

1.3 Effects of duty cycle on passive acoustic monitoring metrics : The case of blue whale songs

JASA ARTICLE

Effects of duty cycle on passive acoustic monitoring metrics: The case of blue whale songs

Mathilde Michel,[1,a] Maëlle Torterotot,[1] Jean-Yves Royer,[2] and Flore Samaran[1]
[1]Centre National de la Recherche Scientifique Laboratoire des Sciences et Techniques de l'information, de la Communication et de la Connaissance, Ecole Nationale Supérieur de Techniques Avancées Bretagne, Brest, France
[2]Lab Geo-Ocean, Centre National de la Recherche Scientifique, University of Brest, Ifremer, F-29280, Plouzané, France

ABSTRACT:
Long-term fixed passive acoustic monitoring of cetacean populations is a logistical and technological challenge, often limited by the battery capacity of the autonomous recorders. Depending on the research scope and target species, temporal subsampling of the data may become necessary to extend the deployment period. This study explores the effects of different duty cycles on metrics that describe patterns of seasonal presence, call type richness, and daily call rate of three blue whale acoustics populations in the Southern Indian Ocean. Detections of blue whale calls from continuous acoustic data were subsampled with three different duty cycles of 50%, 33%, and 25% within listening periods ranging from 1 min to 6 h. Results show that reducing the percentage of recording time reduces the accuracy of the observed seasonal patterns as well as the estimation of daily call rate and call type richness. For a specific duty cycle, short listening periods (5–30 min) are preferred to longer listening periods (1–6 h). The effects of subsampling are greater the lower the species' vocal activity or the shorter their periods of presence. These results emphasize the importance of selecting a subsampling scheme adapted to the target species.
© 2024 Acoustical Society of America. https://doi.org/10.1121/10.0025545
(Received 28 August 2023; revised 11 March 2024; accepted 22 March 2024; published online 9 April 2024)
[Editor: Klaus Lucke] Pages: 2538–2548

15 April 2024 10:12:38

I. INTRODUCTION

Passive acoustic monitoring (PAM) is commonly used to study the seasonal and spatial distribution and acoustic behaviour of cetaceans (e.g., Mellinger *et al.*, 2007; Van Parijs *et al.*, 2009). PAM is a tool of choice for collecting long-term data in remote areas such as the Southern Indian Ocean (e.g., Samaran *et al.*, 2010; Leroy *et al.*, 2016; Torterotot *et al.*, 2020). Unlike visual methods, it offers the possibility of collecting data during day and night, in bad-weather conditions, or when individuals are under the water's surface (Stanistreet *et al.*, 2016).

Blue whales are ideal candidates for such monitoring because of their long (longer than 15 s), frequently repeated loud (more than 180 dB re 1 μPa at 1 m), and low frequency (20–100 Hz) calls (Cummings and Thompson, 1971). The Southern Indian Ocean is inhabited by at least two subspecies of blue whales: the Antarctic blue whale (*Balaenoptera musculus intermedia*) and several populations of pygmy blue whales (*Balaenoptera musculus brevicauda*). Their distinct acoustic repertoire allows distinguishing different acoustic populations (Ichihara, 1966; LeDuc *et al.*, 2003; McDonald *et al.*, 2006; Stafford *et al.*, 2001). Due to the sparse distribution, low density, and wide extent of habitat of these acoustic populations, establishing their migration patterns in a remote area poses a real challenge that can be partly overcome by deploying wide arrays of autonomous recorders (Samaran *et al.*, 2010; Royer 2009).

Long-term PAM of cetacean's species is generally achieved by deploying autonomous recorders, which are anchored to the seabed, that archive data until their recovery after which data can be retrieved and processed on land. Thus, battery capacity generally limits the recording duration of instruments (see review by Sousa-Lima, 2013). Logistics and maintenance costs for redeploying recorders for a long monitoring period must also be optimized. Hence, the choices are either to reduce the frequential sampling rate or recording schedule through temporal subsampling (e.g., Au *et al.*, 2013). Decreasing the sampling rate will impact the recorded frequency band and may restrict the recording of some species. Temporal subsampling involves setting a duty cycle in which the recorder only collects data for a limited fixed time on a repeated basis (Rand *et al.*, 2022). For example, a duty cycle or listening proportion of 0.5 can be set by recording 30 min of data every hour. By adjusting the listening proportion and recording duration, different subsampling schedules can be designed.

However, a subsampling schedule can affect the ecological conclusions drawn from PAM studies (e.g., Thomisch *et al.*, 2015). For instance, inappropriate duty cycles may bias passive acoustic detections (Rand *et al.*, 2022; Riera *et al.*, 2013). Thus, the choice of an adequate subsampling scheme is paramount, must be made according to the research scope and prior knowledge of the acoustic

[a]Email: mathilde.michel@ensta-bretagne.org

 © 2024 Acoustical Society of America

https://doi.org/10.1121/10.0025545

behaviour and detectability of the target species, and then according to the available power and deployment plan of the recorders (Rand *et al.*, 2022; Sousa-Lima, 2013). Fortunately, the introduction of low-power, high-capacity storage has significantly reduced storage limitations in recorders.

Previous studies on the effects of PAM subsampling focused on species' daily presence, daily calling rates (Thomisch *et al.*, 2015; Stanistreet *et al.*, 2016), or the detectability of vocalizations (Rand *et al.*, 2022). Here, we analyse the effects of duty cycles on the detection of stereotyped and non-stereotyped vocalizations from three acoustic populations of blue whales at three locations in the Southern Indian Ocean with two main objectives: first, to assess the effects on long-term monitoring of seasonal patterns; and second, to evaluate the effects on daily call rates and the conclusions that can be inferred from them about the presence of several species of large blue whales at the same site.

II. MATERIAL AND METHODS

A. Data acquisition

Acoustic data used in this study were acquired by the OHASISBIO (Observatoire Hydro-Acoustique de la Sismicite et de la Biodiversité; Royer 2009) network of autonomous hydrophones, located in the Southwestern Indian Ocean, spanning from 24°S to 56°S in latitude and from 52°E to 83°E in longitude. Among the nine sites deployed, three sites were selected based on the consistent presence of three different blue whale acoustic populations. These instruments were located west of Kerguelen Island (WKER) and northeast of St. Paul and Amsterdam Islands (NEAMS) from 2010 to 2019 and south of the Southeast Indian Ridge (SSEIR) from 2014 to 2018 (Fig. 1).

At least two subspecies of blue whales live in this area: the Antarctic blue whale (*B. m. intermedia*), hereafter referred to as ANT BW, and the pygmy blue whale (*B. m. brevicauda*), of which two populations were clearly identified in acoustic recordings: Southwestern Indian Ocean (SWIO PBW) and Southeastern Indian Ocean (SEIO PBW) pygmy blue whales. Each of these subspecies emits a distinctive, stereotyped song phrase, as well as possibly common, non-stereotyped *D*-calls (Fig. 2; Torterotot *et al.*, 2019).

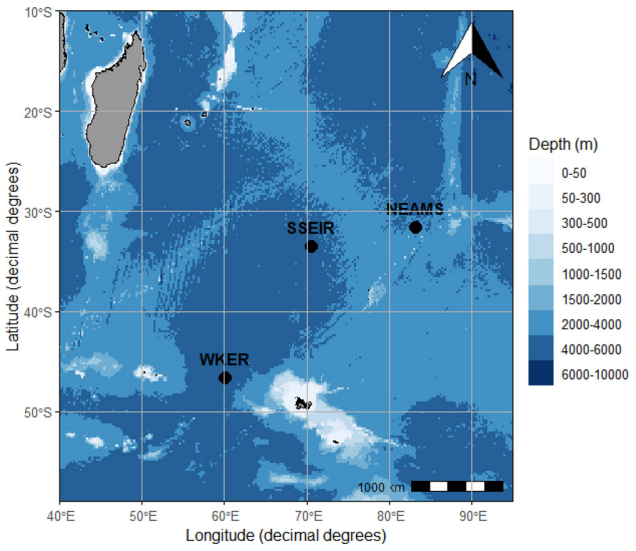


FIG. 1. (Color online) Mooring sites of the OHASISBIO hydrophone network used in this paper.

Each autonomous mooring consisted of an anchor, an acoustic release, an adjustable line, and a submerged buoy hosting the recording system. Instruments were moored between 1000 and 1300 m below sea-surface and recorded continuously at a sampling rate of 240 Hz (see Torterotot *et al.*, 2020, for more details).

B. Call automated detections

The study uses the detections from Torterotot *et al.* (2020), who applied an automatic detection algorithm based on dictionary learning and sparse representation of blue whale calls (Socheleau and Samaran, 2017; Torterotot *et al.*, 2020). The algorithm scanned the acoustic data with a sliding window of duration, t, and tried to reconstruct the observed signal with a combination of calls composing the dictionary: for SEIO PBW, only the second unit of the song phrase was targeted, $t = 25$ s; for SWIO PBW, only the first unit of the song phrase was targeted, $t = 20$ s; for ANT BW, the all song phrase was targeted, $t = 18$ s; and for D-calls, the whole call was targeted, $t = 8$ s (Torterotot *et al.*, 2019). Each part of the song phrase defined above will be named "call" in the rest of the paper. The data analysed herein consists of the call detections from the three acoustic populations described above, which have already been processed at the three sites over a period of 9 or 4 yr. The detection data included the recording site, the detected acoustic population, as well as the start date and time of each detected call.

C. Subsampling analysis

1. Subsampling schemes

Subsampling schemes are defined by a listening period, L (in minutes), and a cycle period, Tc (in minutes); the ratio L/Tc defines the duty cycle, D (in percent), and N is the number of cycles per day (Thomisch *et al.*, 2015). Note that a given duty cycle, D, can be designed with different subsampling strategies, e.g., a single long listening period per day or several short windows evenly distributed over the day but amounting to the same listening period in a day.

The continuous data containing call detections were subsampled according to 23 temporal subsampling schemes with listening periods from 1 min to 6 h and duty cycles, D, of 50%, 33%, and 25%. Potential effects of these subsampling schemes were explored by varying the listening period, L (hence, the number of cycles per day, N), and the duty cycle, D (Table I).

Because a listening period can start anytime within a cycle period, Tc (6 h every 12 h can be either the first 6 h or the last 6 h), different listening phases, r (without overlap), must be considered. For statistical analyses of the effects of subsampling data, each possible realization r ($r = 2$ for

15 April 2024 10:12:38

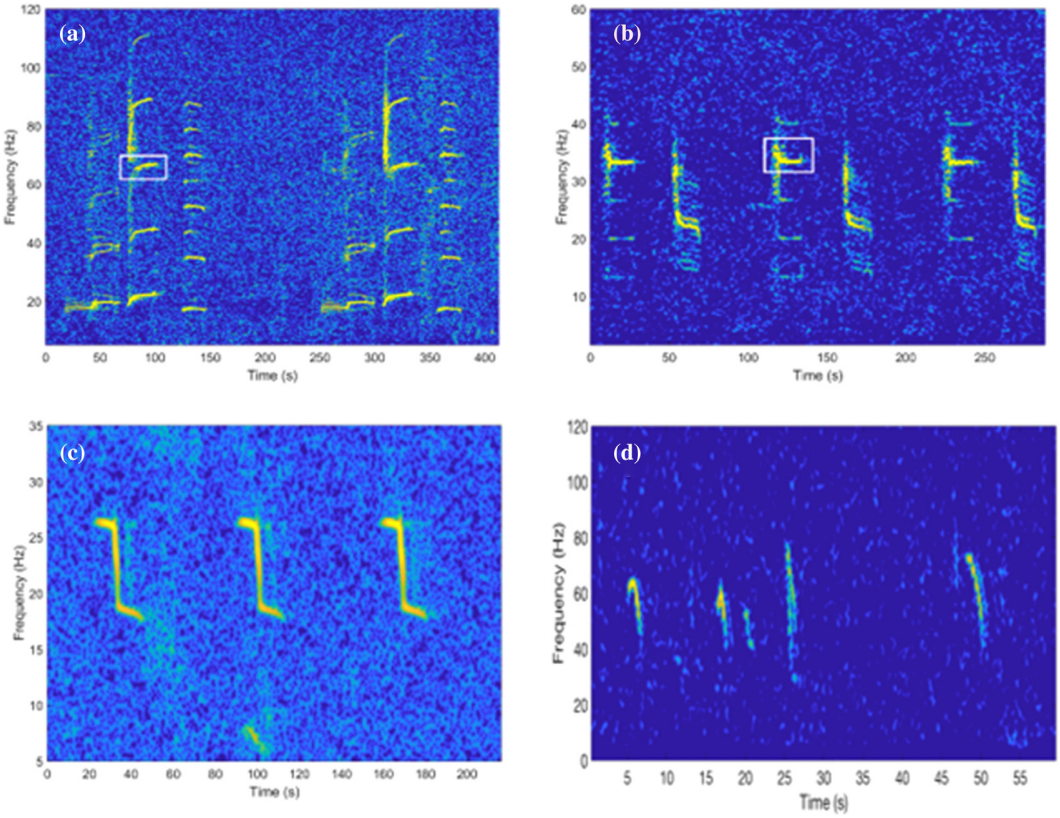


FIG. 2. (Color online) Typical spectrograms of three stereotyped song phrase of (a) a Southeastern Indian Ocean pygmy blue whale (SEIO PBW), (b) a Southwestern Indian Ocean pygmy blue whale (SWIO PBW), (c) an Antarctic blue whale (ANT BW), and (d) a series of five non-stereotyped *D*-calls, all recorded in the Southern Indian Ocean by the OHASISBIO network.

$D = 50\%$, $r = 3$ for $D = 33\%$, and $r = 4$ for $D = 25\%$) of each listening period for all cycles, N, and each duty cycle, D, were processed and averaged.

If the detection of a blue whale call in continuous data occurred during the listening period for a determined duty cycle, it was included in the subsampled data set for this duty cycle. To account for long calls that might only be partially captured by subsampled data, detections must have 100% of the call duration falling within the listening period to be included in the set. This choice was justified to agree with the automatic detector used by Torterotot *et al.* (2020), which operates from the calls described in Sec. II B. The average percentage of excluded calls for each subsampled dataset is presented in Table II.

Different subsampling schemes were applied on each day of each year of the recording period at three sites and over 9 yr. Prior to averaging data across years, an analysis of variance (ANOVA) was performed; this preliminary analysis was used to check that the results did not vary statistically significantly across years for the different metrics: seasonality, daily call rate estimates, and call type richness. These different metrics will be described in Secs. II C 2–II C 4. All statistical analyses and subsampling schemes were implemented in *R* (R Core Team, 2021).

TABLE I. Different subsampling designs used for blue whale detections. The duty cycle, D, (%) is the ratio between the listening period, L, and the cycle period, Tc, repeated throughout a day.

	$D = 50\%$		$D = 33\%$		$D = 25\%$	
	$r = 2$		$r = 3$		$r = 4$	
Listening period Lp	Cycle period Tc	Cycles per day N	Cycle period Tc	Cycles per day N	Cycle period Tc	Cycles per day N
1 min	2 min	720	3 min	480	4 min	360
5 min	10 min	144	15 min	96	20 min	72
15min	30 min	48	45 min	32	1 h	24
30 min	1 h	24	1 h30 min	16	2 h	12
1 h	2 h	12	3 h	8	4 h	6
3 h	6 h	4	9 h	2.66	12 h	2
6 h	12 h	2	18 h	1.33	24 h	1

15 April 2024 10:12:38

https://doi.org/10.1121/10.0025545

TABLE II. Average percentage and relative standard deviation of excluded calls from subsampled datasets for each type of call described in Fig. 2, based on each duty cycle, *D*, and listening period., *Lp*. This is an average across all possible realizations, *r*.

Duty cycle, D	Listening period, Lp	SWIO PBW Mean of excluded calls (%)	SEIO PBW Mean of excluded calls (%)	ANT BW Mean of excluded calls (%)	*D*-calls Mean of excluded calls (%)
50%	1 min	66.6 ± 0,4	70.8 ± 0.3	65.0 ± 0.4	56.6 ± 0.2
	5 min	53.3 ± 0.3	54.0 ± 0.2	53.0 ± 0.5	51.3 ± 1.1
	15 min	51.1 ± 0.4	51.4 ± 0.8	51.0 ± 0.1	50.5 ± 0.2
	30 min	50.6 ± 0.9	50.7 ± 0.1	50.5 ± 0.5	50.2 ± 0.8
	1 h	50.3 ± 0.9	50.3 ± 1	50.3 ± 0	50.1 ± 0.3
	3 h	50.3 ± 0.7	50.3 ± 1.7	50.3 ± 1.5	50.1 ± 2.3
	6 h	50.3 ± 0.6	50.3 ± 1.6	50.3 ± 2.1	50.1 ± 4
33%	1 min	77.7 ± 0.9	80.5 ± 0.7	76.7 ± 0.2	71.0 ± 0.9
	5 min	68.9 ± 0.3	69.3 ± 0.8	68.7 ± 0.1	67.5 ± 1.1
	15min	68.9 ± 0.5	69.3 ± 0.8	68.7 ± 0.2	67.5 ± 1.6
	30 min	67.1 ± 0.1	67.1 ± 1.5	67.0 ± 0.3	66.8 ± 3.2
	1 h	66.9 ± 0.4	66.9 ± 1.6	66.8 ± 0.6	66.7 ± 3.4
	3 h	66.9 ± 0.5	66.9 ± 1	66.8 ± 0.1	66.7 ± 2.7
	6 h	66.9 ± 2.3	66.9 ± 0.7	66.8 ± 0.8	66.7 ± 1.9
25%	1 min	83.3 ± 0.4	85.4 ± 0.3	82.5 ± 0.4	78.3 ± 0.5
	5 min	76.7 ± 0.9	77.0 ± 0.4	76.5 ± 0.5	75.7 ± 2.2
	15 min	76.7 ± 0.9	77.0 ± 0.8	76.5 ± 0.5	75.7 ± 1.2
	30 min	75.3 ± 1	75.3 ± 1.8	75.3 ± 0.5	75.1 ± 1.7
	1 h	75.2 ± 1.6	75.2 ± 1.4	75.1 ± 0.5	75.1 ± 1
	3 h	75.2 ± 1.7	75.2 ± 2.1	75.1 ± 3.6	75.1 ± 5.2
	6 h	75.2 ± 1.1	75.2 ± 4.7	75.1 ± 4.4	75.1 ± 11.2

2. Seasonal patterns

To explore the effect of duty cycles on the seasonality, we examined the seasonal patterns described in Torterotot *et al.* (2020). Continuous data were subsampled according to the different patterns presented in Table I. Seasonal patterns were here represented based on the absolute number of detections per day for the calls as described in Sec. II B of the three acoustic populations of interest, as well as for the *D*-calls. Four types of seasonal patterns were identified and used to compare the effect of subsampling (Fig. 3): (i) a strong seasonal pattern with a high number of detections (SEIO PBW at the NEAMS site), (ii) a strong seasonal pattern with a low number of detections (SWIO PBW at the NEAMS site), (iii) a no clear seasonal pattern with a high number of detections (ANT BW at the WKER site), and (iv) a no clear seasonal pattern with a low number of detections (SEIO PBW at the SSEIR site).

To assess the accuracy of the seasonality implied by the subsampled data, the ratio of sum of squares of residuals to total sum of squares was calculated between all different types of subsampling and continuous data. Given the objective of obtaining a more precise depiction of the seasonality in continuous data, the regression coefficient was determined as follows:

$$R^2 = \left(1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}\right), \quad (1)$$

where n is the number of measurements, y_i is the value of ith measurement, $\hat{y}_i$ is corresponding predicted value, and $\overline{y}$ is the mean of measurements.

A regression coefficient result close to one indicates a good fit of the model and, thus, a good representation of the seasonality from the subsampled data (three examples of linear regression for three listening periods with a duty cycle, $D = 50\%$; Fig. 4).

Furthermore, an ANOVA test was conducted for each distinct dataset (site and vocalization type) to validate differences in R^2 values between each listening period of each duty cycle. Subsequently, pairwise comparison tests were performed for datasets in which the ANOVA p-value was significant, between each listening period and the subsequent period, to compare the mean R^2 values pairwise using the Tukey test to adjust the p-values.

3. Daily call rate estimation

The daily call rate or average number of calls detected over a day was computed with the formula used in Thomisch *et al.* (2015):

$$\overline{\gamma_{r,J}} = \frac{1}{N}\sum_{j=1}^{N}\gamma_{r,j}, \quad (2)$$

where $\overline{\gamma_{r,J}}$ was the daily call rate estimated from call rates, $\gamma_{r,j}$, of the rth realization in all N cycles of the jth day. This estimation was accomplished for all possible independent realizations, r, of listening periods for all cycles, N, and each duty cycle, D.

15 April 2024 10:12:38

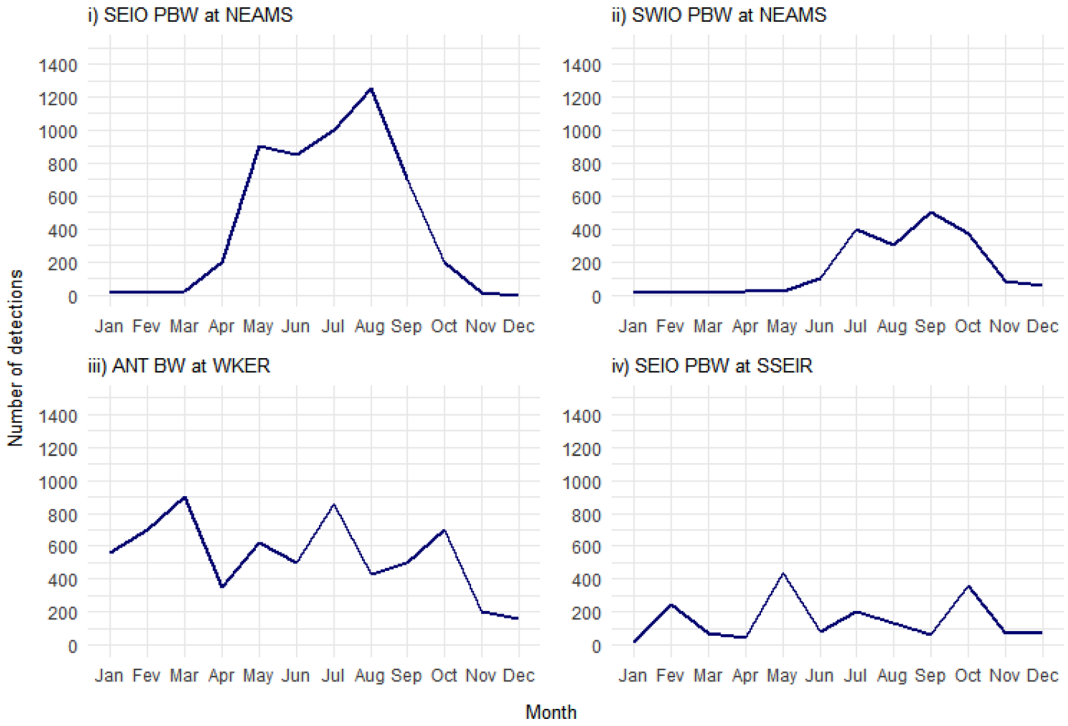


FIG. 3. (Color online) Schematic seasonal patterns, showing (i) and (ii) strong seasonal patterns with high and low numbers of detections by month, respectively; (iii) and (iv) no clear seasonal patterns with high and low numbers of detections by month, respectively. See Fig. 4 of Torterotot *et al.* (2020) for an actual example of such seasonal distributions.

To assess the variability of estimated call rates according to the type of subsampling, the ratio of call rate of subsampled data to call rate of continuous data was calculated for each $\overline{\gamma}_{r,j}$:

$$\text{ratio}_{r,j} = \frac{\overline{\gamma_{r,J}} - \gamma_{\text{true}j}}{\gamma_{\text{true}j}}. \tag{3}$$

The continuous data were subsampled according to different schemes in Table I. This procedure was repeated for all D percentages for each day to provide a mean and standard deviation of call rate estimates according to defined subsampling scheme.

4. Call type richness

Call type richness is a measure of the number of acoustic populations present in a habitat at a given time. In this study, only specific calls of three blue whale acoustic populations were used, as D-calls cannot be unambiguously

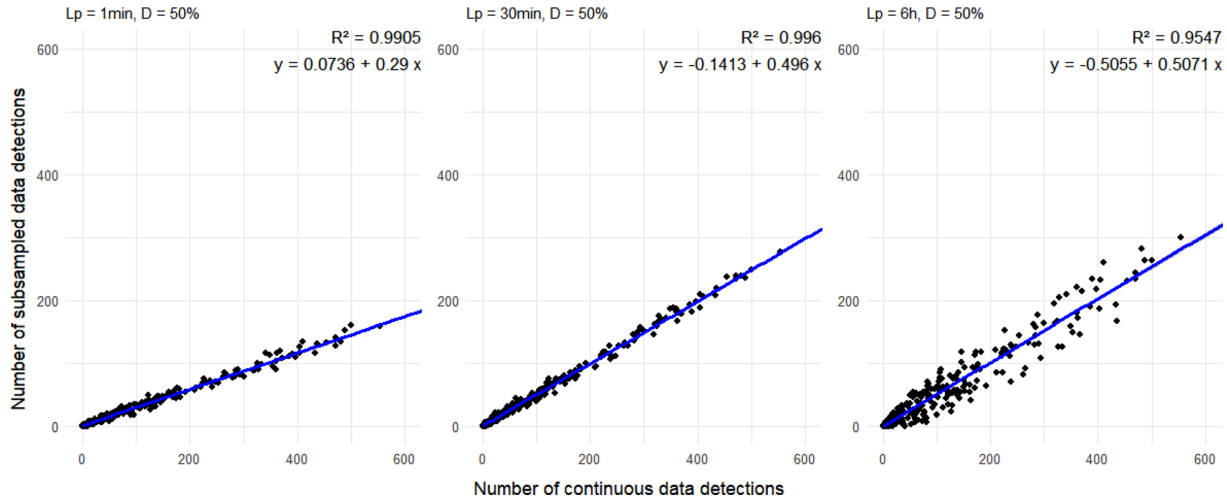


FIG. 4. (Color online) Relationship between the number of continuous data detections and the number of subsampled data detections for SWIO PBW call type at NEAMS site across three listening periods (1 min, 30 min, and 6 h) with a duty cycle, $D = 50\%$. The data points (black circles) are fitted to a linear regression model (blue line). X axis represents the number of detections per day for continuous data, and Y axis represents the number of detections per day for subsampled data.

15 April 2024 10:12:38

linked to any specific acoustic populations. For continuous data, an acoustic population was considered present if there was at least one call detected during the day. In the subsampled data, an acoustic population was considered present acoustically if there was a detection in any of the N cycles of the day. Call type richness at a site for continuous and subsampled data was, therefore, calculated by counting the number of acoustic populations present at the site per day (maximum call type richness = 3).

If the call types richness between subsampled and continuous data contained the same number of acoustic populations, then the decision was correct ($d = 1$) otherwise it was incorrect ($d = 0$). To compute an average proportion of days p_d for which call type richness was correctly assessed using each subsampling programme, d was determined for each day, j, and for each duty cycle, D, and then averaged over all, i, listening period phases. This allowed establishing an average proportion of days in which call type richness was well determined for each subsampling scheme such that

$$\overline{p_d} = \frac{1}{nr}\sum_{i=1}^{r}\sum_{j=1}^{n} d_{ij}, \tag{4}$$

where n is the number of days with detections, r is the number of listening phases, and d_{ij} is the value of d for phase i on day j.

To compare the results of different duty cycles, comparison of means was performed. As the data were not normally distributed, a nonparametric Wilcoxon test was applied for each site.

III. RESULTS

A. Seasonal patterns

Effects of subsampling strategies were first evaluated on seasonal patterns in the occurrence of three blue whale acoustic populations and D-calls at three selected sites by calculating R^2, which is the regression coefficient. Results for different duty cycles of 50%, 33%, and 25% are presented in Fig. 5.

For all acoustic populations and at all sites, the highest duty cycle (largest proportion of listening, 50%) gave the highest value of the regression coefficient, indicating a better accuracy of seasonal model. Unsurprisingly, a 33% duty cycle gave better estimates than a 25% duty cycle.

Within each duty cycle, short listening periods of 5–30 min improved the accuracy of the seasonal model compared to longer listening periods of 1–6 h. Indeed, in many cases, a significant difference was found between $Lp = 1$ h and $Lp = 3$ h (Fig. 5), indicating that from $Lp = 3$ h onward, significantly lower mean regression coefficient values were measured.

However, a listening period that was too short also had a negative impact on the seasonal occurrence pattern. Significant differences primarily occurred between 1-min ($Lp = 1$ min) and 5-min ($Lp = 5$ min) listening periods, resulting in lower mean regression coefficient values compared to other Lp durations, especially for long calls from the SEIO PBW acoustic population [Figs. 5(d)–5(f)].

The best fit between subsampled and continuous data was achieved for acoustic populations with a strong seasonal pattern and many detections. This was the case for the SEIO PBW at NEAMS site, where regression coefficient was approximately 0.95 for a cycle period 5 min and a duty cycle of 50% [Fig. 5(d)]. Additionally, this dataset is the only one in which there is no significant difference between successive Lp (p-value > 0.05). High regression coefficient was also measured when an acoustic population had a no clear seasonal pattern with many detections, such as the ANT BW at the WKER site, where the regression coefficient was always higher than 0.9 for a duty cycle of 50% and listening time of 5 min [Fig. 5(i)]. Moreover, a significant difference between successive Lp was observed only from $Lp = 1$ h onward.

In contrast, the regression coefficient was lower for datasets with few detections and a no clear seasonal pattern [e.g., SEIO PBW at SSEIR, regression coefficient for 30 min at 50% $\approx$ 0.7; Fig. 5(e)] or few detections and a strong seasonal model [SWIO PBW at NEAMS, regression coefficient for 5 min at 50% $\approx$ 0.75; Fig. 5(a)]. Notably, on the WKER site, differences in mean regression coefficient values emerged earlier, between 30 min and 1 h of listening period [Figs. 5(c), 5(f), 5(i), and 5(l)]. Conversely, for D-calls, significant p-values between consecutive listening periods varied across sites: for the SSEIR site [Fig. 5(k)], significant differences were observed between $Lp = 5$ min and $Lp = 15$min as well as between $Lp = 15$ min and $Lp = 30$ min. However, for the WKER site [Fig. 5(l)], the first significant differences between two listening periods occurred between $Lp = 30$ min and $Lp = 1$ h.

B. Daily call rate estimates

The call rate of the subsampled data were very sensitive to the duty cycle D and the listening period L (Fig. 6). The average value of the ratio of subsampled to continuous data were consistently below 0 for all listening durations, indicating a subsampled data underestimation of the daily call rate for different listening durations. This underestimation was most pronounced for the 1-min listening duration, which deviated significantly from the expected value of 0 for longer duration calls [Figs. 6(a)–6(c)]. This deviation from expected values was also observed for all listening periods for usage cycles of 33% and 25%.

In general, the ratio variability [as defined in Eq. (3)] increased as the listening period increased for a given duty cycle percentage and increased as the duty cycle percentage decreased. For a given duty cycle, its variability was greater for the longest listening times (1–6 h).

The subsampled data underestimation and variability were most significant for the SEIO PBW, which had the longest call duration [25 s; Fig. 6(b)]. In contrast, the subsampled data underestimation of the daily call rate was lower for D-calls, which were the shortest calls [8 s; Fig. 6(d)]. Note that for this particular type of call, variability

15 April 2024 10:12:38

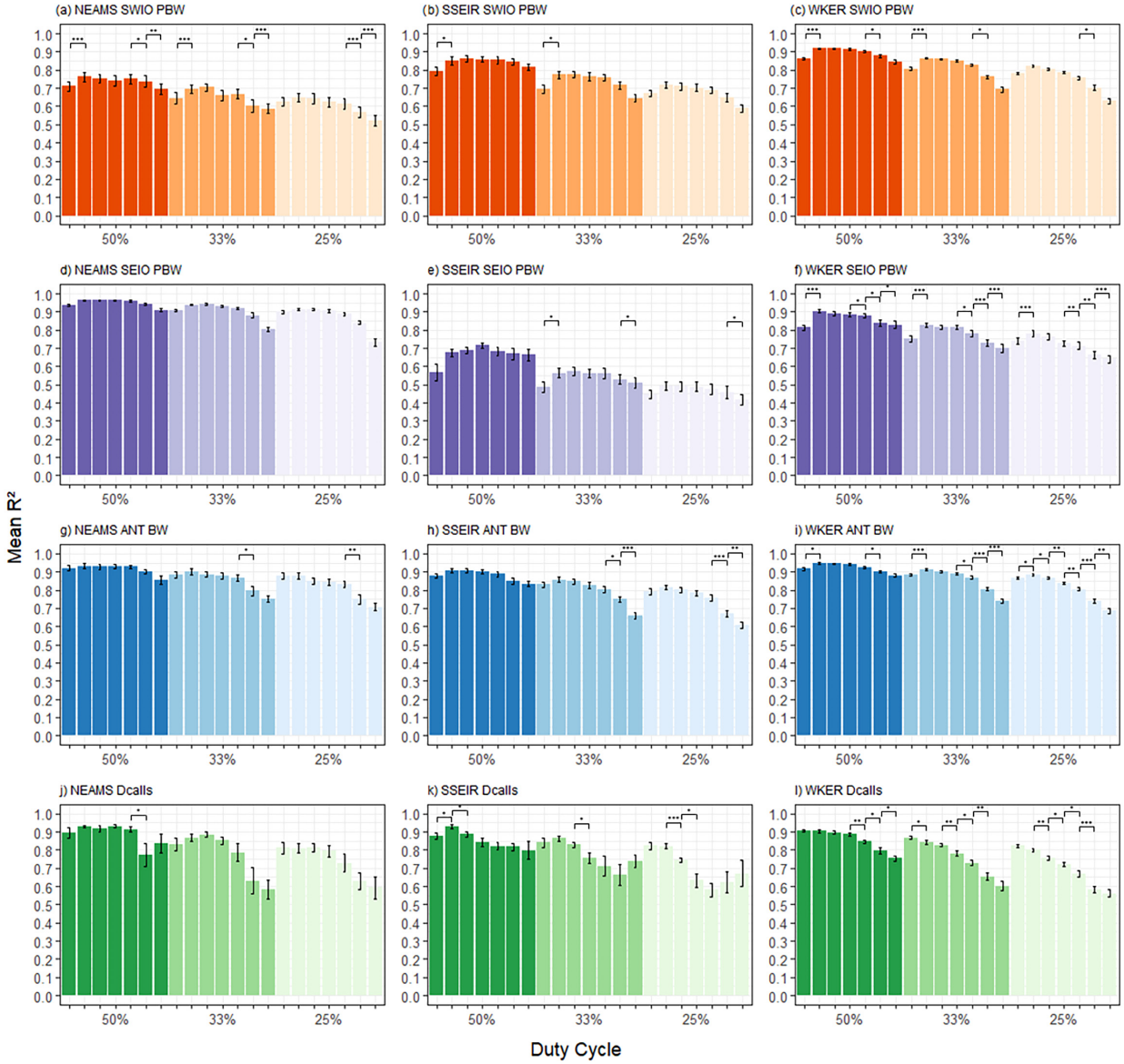


FIG. 5. (Color online) Comparison of different subsampling schemas in which each column represents a site (from left to right, NEAMS, SSEIR, and WKER) and each row represents a call (from top to bottom, SWIO PBW, SEIO PBW, ANT BW, and *D*-calls). Each shade of colour corresponds to a given duty cycle, *D* (from left to right, 50%, 33%, and 25%), and bar plots represent different listening periods, *L*, for a given duty cycle (from left to right, 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, and 6 h, respectively). Vertical black lines are the standard deviations of the regression coefficient. The four types of seasonal patterns described in Sec. II C 2 are represented here as (a) SWIO PBW to NEAMS with a strong seasonal pattern and few detections, (d) SEIO PBW to NEAMS with a strong seasonal pattern and many detections, (e) SEIO PBW to SSEIR with a no clear seasonal pattern and few detections, and (i) ANT BW to WKER with a no clear seasonal pattern and many detections. Horizontal bars marked with "*," "**," and "***" show domains where Tukey tests yielded $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

was most pronounced for each listening period, likely stemming from the irregularity in its occurrence.

C. Call type richness

The proportion of days in which call type richness was correctly assessed according to different subsampling schemes also depends on the selected duty cycle, *D*, and duration of listening period, *L* (Fig. 7). Highest proportions of days with correctly assessed call type richness was observed for duty cycles of 50%, then 33%, and finally 25% for each of the three sites. Results of Wilcoxon tests between different duty cycles for each site show significant *p*-values ($p < 0.0001$), indicating that the average number of days with a correct estimate of the call type richness differs according to the subsampling scheme. For each duty cycle, a 1-min listening period was less likely to correctly capture call type richness than longer listening periods. Short periods (between 5 and 30 min) gave a higher average proportion of days with correctly assessed call type richness than longer listening period (more than 1 h). WKER site had more days correctly evaluated compared to the other two

15 April 2024 10:12:38

https://doi.org/10.1121/10.0025545

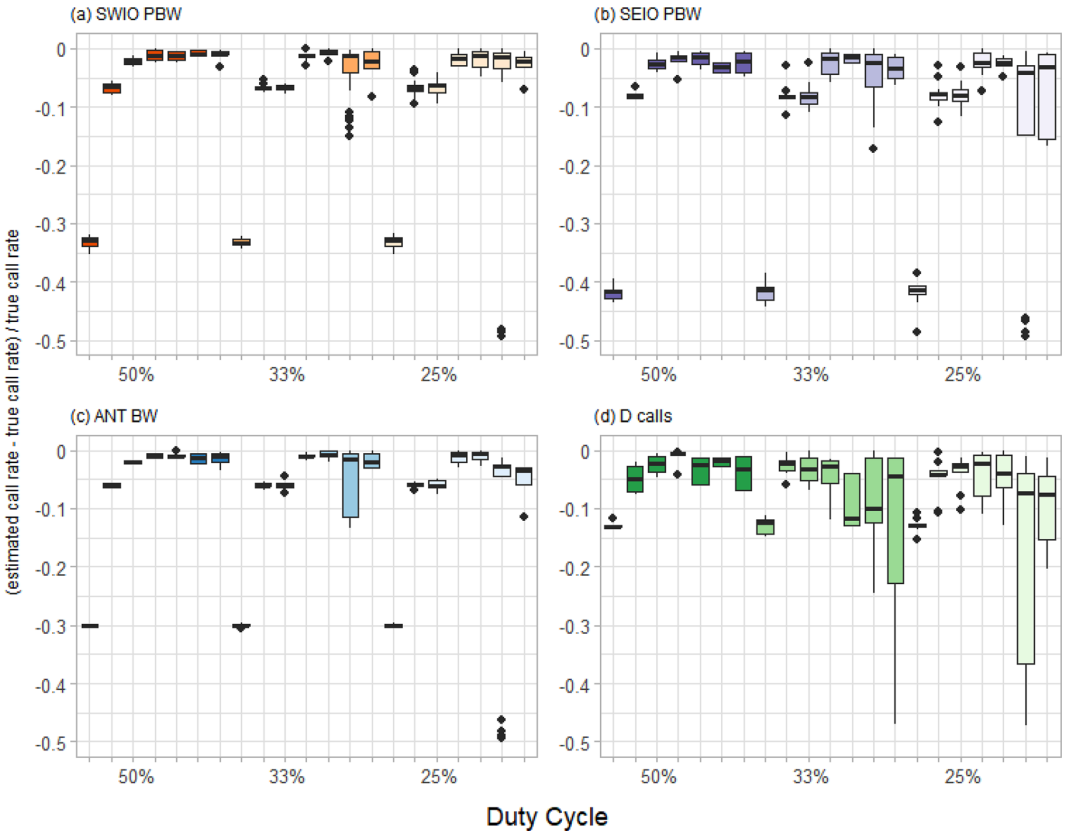


FIG. 6. (Color online) Comparison of different subsampling patterns on the ratios between estimated and actual daily call rates for different duty cycles and acoustic populations: a) SWIO PBW, b) SEIO PBW, c) ANT BW, d) D-calls. For each duty cycle, boxplots represent different listening periods L (from left to right: 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, 6 h respectively).

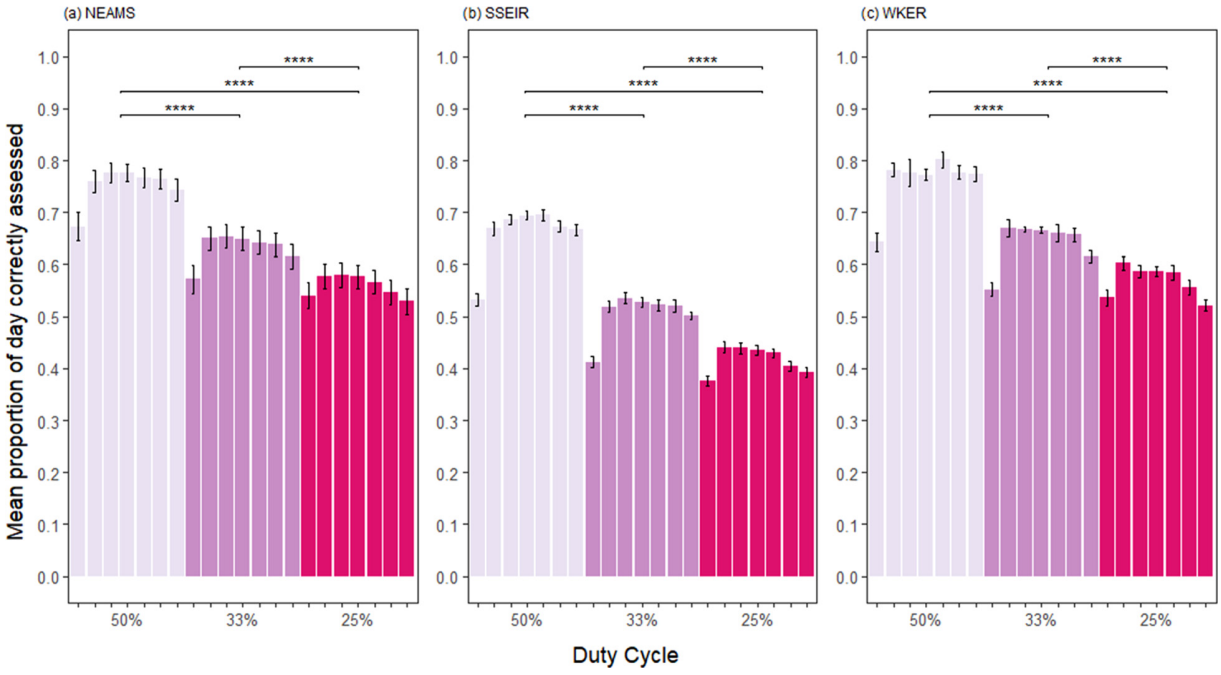


FIG. 7. (Color online) Comparison of proportion of days with a correct assessment of site call type richness with its standard deviation according to different subsampling schemes, showing (a) NEAMS, (b) SSEIR, and (c) WKER. Each shade of colour matches a different duty cycle, D (from left to right, 50%, 33%, and 25%). For each duty cycle, barplots represent different listening periods, L (from left to right, 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, and 6 h respectively). Vertical black lines are the standard deviations. Horizontal bars with "****" show domains where Wilcoxon tests yielded $p < 0.0001$.

15 April 2024 10:12:38

sites [Fig. 7(c)]. For example, call type richness was correctly assessed at WKER site, with an average probability of 75% for $D = 50\%$ with probabilities decreasing with smaller D (approximately 55% for $D = 25\%$).

IV. DISCUSSION

The aim of this study was to evaluate how temporal subsampling schemes can potentially bias metrics commonly used in passive acoustic studies to assess the occurrence of acoustic populations of blue whales in the Southern Indian Ocean based on call detections. Effects of subsampling are not the same, depending on the selected metric and chosen subsampling strategy.

It is important to note that the results of this study are specific to how calls were processed and detected. Here, the detection algorithm was based on dictionary learning and sparse representation of blue whale calls within specified time windows. Other algorithms, for instance, based on call rhythm or chorus intensity (e.g., Leroy *et al.*, 2018), will not be affected in the same way as in this study.

A. Effect of temporal subsampling on seasonal patterns

High duty cycles had higher regression coefficient, indicating that the seasonality was best represented with high duty cycles (or, conversely, poorly represented with low duty cycles). The best results were obtained when a species had a very marked seasonality with a high number of detections in few months. At NEAMS site, SEIO PBW has a very marked seasonality (April–July) with a minimum of 200 detections per week when the species was present and stable over the years (Torterotot *et al.*, 2020). In such a case, real and estimated seasonal patterns match even with a duty cycle of 25% [$R^2 > 0.9$; Fig. 5(d)].

High regression coefficient was also obtained for the ANT BW species. Number of detections of this species was high at the 3 selected sites (> 500 per week). However, unlike the SEIO PBW at NEAMS, its peaks of presence were less marked: significant emissions are observed over a more spread-out number of months. This result suggests that areas where vocalizations are numerous and constant over time provide a good representation of seasonal patterns.

In contrast, SWIO PBW calls at NEAMS site were only present occasionally between April and May. Its seasonal pattern was, therefore, more difficult to represent when data were subsampled, even for a high duty cycle ($R^2 < 0.75$ for 50%). Similarly, SEIO PBW calls at SSEIR site have a low coefficient of regression, suggesting a less accurate representation of seasonal patterns for this acoustic population ($R^2 < 0.6$ for 50%). These results are explained by its low acoustic activity (<90 calls per week) with calls, therefore, less likely to be captured in subsampled data.

Even though D-calls are shorter in duration and, thus, a smaller percentage of calls are removed by subsampling, the results show that they exhibit similar patterns to the signals emitted by the three studied acoustic populations. This suggests that irregularly emitted vocalizations, despite being shorter, should also be considered when choosing the duty cycle.

High acoustic activity with a high number of detections is the most important input for accurately representing the patterns of species presence from subsampled data. Conversely, continuous data collection seems unavoidable to monitor species that frequent an area irregularly or are present without acoustic activity or with a sparse acoustic activity.

For calls from the three acoustic populations and D-calls, long listening periods (≥ 1 h) result in a less accurate representation of continuous data compared to shorter listening periods. Similarly, a 1-min listening period leads to the less accurate results on the representativeness of seasonality in subsampled data. Indeed, many vocalizations will be cut off by an excessively short listening period and, thus, are less likely to be detected. This effect is even greater for acoustic populations with the longest window length used by the automatic detection algorithm (25 s for SEIO PBW). Although short listening periods seem to give better results, it is necessary to consider the duration of call (and/or the function and parameters of the detection algorithm) of the target species as a decisive factor in designing the subsampling scheme. In this regard, listening periods between 5 and 30 min provide similar representations of continuous data by subsampled data and, therefore, seem most suitable for our acoustic populations in the area.

B. Effects of temporal subsampling on daily call rate and call type richness

The accuracy of daily call rate estimates is highly influenced by subsampling designs, as depicted in Fig. 6. There is a notable variation between the estimated call rate and the actual call rate, particularly for extended listening periods ranging from 1 to 6 h. Furthermore, this discrepancy becomes more pronounced as the duty cycle decreases. These findings align with those observed by Thomisch *et al.* (2015). Specifically, a listening period of only 1 min consistently leads to a strong underestimation of the call rate, accompanied by low variability. This suggests that such a short duration is inadequate for our target species, which produces long calls between 18 and 25 s.

This is especially notable for the acoustic population with the longest call duration, SEIO PBW. As partial calls were excluded from the subsampled datasets to retain only complete calls, a substantial proportion was removed (70.8% for $Lp = 1$ min and $D = 50\%$; Table II), resulting in a persistent underestimation regardless of the listening period.

Additionally, the variability is amplified when calls are generated in irregular patterns (such as D-calls, for example) or there are few detections. These findings are consistent with those of Thomisch *et al.* (2015), who observed a greater disparity between call rate estimates and the actual rate for vocalization patterns characterized by low call rates.

15 April 2024 10:12:38

https://doi.org/10.1121/10.0025545

Regarding *D*-calls, their irregular emission and shorter duration align with the conclusions of the study by Rand *et al.* (2022) on killer whales (*Orcinus orca*), where the optimal duty cycle has shorter listening lengths than for the stereotyped vocalizations of the acoustic populations studied here. These cycles would facilitate capturing the variation in acoustic behaviour throughout the day.

Regardless of the selected duty cycle (50%, 33%, or 25%), the daily call rates across all listening periods are consistently underestimated. Nevertheless, the 30-min listening period provided the most accurate daily call rate (i.e., the closest to the daily call rate from continuous sampling) and the lowest variability for all four types of calls, as shown in Fig. 6. This suggests that a 30-min listening period may represent the optimal compromise for studying stereotyped vocalizations and irregular *D*-calls. The richness of call types was as underestimated as the percentage of work cycles decreased: a duty cycle of 50% yields more accurate results than a duty cycle of 25%. These findings match the trends observed by Thomisch *et al.* (2015), who showed that a duty cycle of 10% led to a 26% underestimation of the daily presence of North Atlantic right whales (*Eubalaena glacialis*), or those observed by Riera *et al.* (2013), who found that a duty cycle of 25% resulted in a 24% decrease in killer whale (*O. orca*) acoustic encounters compared to continuous recording.

Short listening periods (between 5 and 30 min) increase the likelihood of accurately assessing call type richness regardless of duty cycle. For listening periods of 1 min or longer than 1 h, call type richness is more likely to be underestimated (Fig. 7). These results are consistent with those found for the daily acoustic presence of other marine mammal species such as killer whales (Rand *et al.*, 2022) or beaked whales (Stanistreet *et al.*, 2016).

It is important to note that the studies cited above relate to the daily presence of a single species, whereas this study relates to the call type richness of a site, i.e., the total number of call types present per day. The effect of subsampling is greater on the call type richness metric than on the daily presence metric computed by Thomisch *et al.* (2015): the proportion of days with a correct assessment of call type richness is lower (maximum 77% for a duty cycle of 50%). In comparison, Thomisch *et al.* (2015) found a 100% probability of correctly assessing daily presence of ANT BWs at a site for a duty cycle of 50%.

In sites where multiple acoustic populations of blue whales coexist, with many detections for each (e.g., WKER), number of days with a correctly assessed call type richness is better at a site recording a majority of calls from a single population and few calls from the other acoustic populations (e.g., SSEIR; Fig. 7). Therefore, duty cycle has a significant impact on the estimation of call type richness in the study area. A good knowledge of the use and presence of species in the study area is, thus, critical to choose the most appropriate recording strategy so as not to miss detection of a rare species due to inadequate subsampling.

V. CONCLUSION

This study demonstrates that temporally subsampling acoustic data can have heterogeneous effects on results of long-term monitoring of blue whale acoustic populations as well as the assessment of call type richness daily call rate based on their presence in different geographic areas.

If subsampling is necessary because of technical or logistical constraints, it is preferable to collect data in several short listening periods compatible with the duration of the vocalizations sought. Such a subsampling scheme leads to many cycles per day and allows for a better representation of daily call rate over the course of a day as well as a better representation of seasonal patterns.

The occurrence of different types of seasonal patterns may affect the accuracy with which subsampling patterns represent continuous data. Number of detections is a decisive factor for using a low duty cycle (25%) to obtain good results. This study also showed that when acoustic populations are rare in an area, continuous sampling may be required to not lose too much information in representation of seasonal patterns, the daily call rate, and estimation of call type richness. Indeed, if the objective of a passive acoustics project is to characterize a soundscape and explore its acoustic biodiversity, continuous recordings are necessary to avoid missing rare or transient species over a very short period.

The subsampling strategy must, therefore, be chosen based on biological considerations and a good understanding of the vocal behaviour of the target species. In addition, for long-term and large-scale monitoring studies, selection of the temporal subsampling strategy should also be based on an understanding of the different migration and occurrence patterns of the species in the study area. For a multiyear project, it would be appropriate, for example, to monitor continuously an area during the first year to gain knowledge about the acoustic presence of species, and then adapt the subsampling plan accordingly for the remainder of the monitoring period.

Depending on the automatic detection algorithm, subsampling may affect study results differently. The implementation of subsampling must, therefore, also consider the detector that will be used to process the data. In this study, 100% of the vocalization had to be included in the listening period because the detector used to detect the vocalization worked this way (i.e., a detector cannot detect incomplete calls if they are not sufficiently similar to the calls used to build the learning dictionary; Socheleau and Samaran, 2017). As other detectors exist and may operate differently, it is important to define their operation upstream, which may influence the subsampling strategy.

Embedding acoustic signal processing algorithms in acquisition systems now offers new possibilities for acoustic monitoring. A random subsampling scheme could, for instance, limit the effect of single species dependence by generalizing it to multiple species. However, the effect of such schemes would first need to be evaluated on continuous

15 April 2024 10:12:38

data. The data in this study focused on low frequency emissions from blue whales. Over a wider frequency range, many other species with their different behaviours might require specific subsampling schemes for each. Optimizing passive acoustic data collection procedures, including the selection of appropriate sampling strategies, is among the issues for future management and conservation studies of underwater biodiversity.

ACKNOWLEDGMENTS

The authors wish to thank the captains and crews of RV Marion Dufresne for the successful deployments and recoveries of the instruments of the OHASISBIO array. Cruises were funded by the French Polar Institute and the French Oceanographic Fleet with additional support from Institut National des Sciences de l'Univers Centre National de la Recherche Scientifique. The authors also wish to thank Alain de Cheveigné for his help and advice on the statistical analyses. This work was supported by the professors of the International Master of Bioacoustics, Nicolas Mathevon and David Reby.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

No animals were harmed during the project.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Au, W. W. L., Giorli, G., Chen, J., Copeland, A., Lammers, M., Richlen, M., Jarvis, S., Morrissey, R., Moretti, D., and Klinck, H. (**2013**). "Nighttime foraging by deep diving echolocating odontocetes off the Hawaiian Islands of Kauai and Ni'ihau as determined by passive acoustic monitors,"J. Acoust. Soc. Am. **133**, 3119–3127.

Cummings, W. C., and Thompson, P. O. (**1971**). "Underwater sounds from the blue whale, *Balaenoptera musculus*," J. Acoust. Soc. Am. **50**, 1193–1198.

Ichihara, T. (**1966**). "The pygmy blue whale, *Balaenoptera musculus brevicauda*, a new subspecies from the Antarctic," in *Whales, Dolphins, and Porpoises*, edited by K. S. Norris (University of California Press, Berkeley, CA), p 79–111.

LeDuc, R. G., Dizon, A. E., Goto, M., Pastene, L. A., Kato, H., Nishiwaki, S., LeDuc, C. A., and Brownell, R. L. (2003). "Patterns of genetic variation in Southern Hemisphere blue whales and the use of assignment test to detect mixing on the feeding grounds," J. Cetacean Res. Manage. **9**, 73–80.

Leroy, E., Samaran, F., Stafford, K., Bonnel, J., and Royer, J. (**2018**). "Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the Southern Indian Ocean," Endang. Species Res. **37**, 289–300.

Leroy, E. C., Samaran, F., Bonnel, J., and Royer, J.-Y. (**2016**). "Seasonal and diel vocalization patterns of Antarctic blue whale (*Balaenoptera musculus intermedia*) in the Southern Indian Ocean: A multi-year and multi-site study," PLoS One **11**, e0163587.

McDonald, M., Hildebrand, J., and Mesnick, S. (**2006**). "Biogeographic characterisation of blue whale song worldwide: Using song to identify populations," J. Cetacean Res. Manage. **8**, 55–65.

Mellinger, D., Stafford, K., Moore, S., Dziak, R., and Matsumoto, H. (**2007**). "An overview of fixed passive acoustic observation methods for cetaceans," Oceanography **20**, 36–45.

Rand, Z. R., Wood, J. D., and Oswald, J. N. (**2022**). "Effects of duty cycles on passive acoustic monitoring of southern resident killer whale (*Orcinus orca*) occurrence and behavior," J. Acoust. Soc. Am. **151**, 1651–1660.

R Core Team (**2021**). "*R*: A language and environment for statistical computing" (R Foundation for Statistical Computing, Vienna, Austria), available at https://www.R-project.org/ (Last viewed 15 January 2024).

Riera, A., Ford, J. K., and Chapman, N. R. (**2013**). "Effects of different analysis techniques and recording duty cycles on passive acoustic monitoring of killer whales," J. Acoust. Soc. Am. **134**, 2393–2404.

Royer, J.-Y. (**2009**). "OHA-SIS-BIO: HydroAcousticity from SISmicity and Biodiversity in the Indian Ocean," available at https://campagnes.flotteoceanographique.fr/series/229/ (Last viewed 10 January 2022).

Samaran, F., Adam, O., and Guinet, C. (**2010**). "Discovery of a mid-latitude sympatric area for two Southern Hemisphere blue whale subspecies," Endang. Species Res. **12**, 157–165.

Socheleau, F.-X., and Samaran, F. (**2017**). Detection of mysticete calls: A sparse representation-based approach, [Research Report] RR-2017-04-SC, Dépt. Signal et Communications (Institut Mines-Télécom-IMT Atlantique-UBL); Laboratoire en sciences et technologies de l'information, de la communication et de la connaissance (UMR 6285 - CNRS - IMT Atlantique - Université de Bretagne Occidentale - Université de Bretagne Sud - ENSTA Bretagne - Ecole Nationale d'ingénieurs de Brest); École nationale supérieure de techniques avancées Bretagne. (Ministère de la Défense). 2017, hal-01736178v2.

Sousa-Lima, R. (**2013**). "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," Aquat. Mamm. **39**, 23–53.

Stafford, K. M., Nieukirk, S. L., and Fox, C. (**2001**). "Geographic and seasonal variation of blue whale calls in the North Pacific," J. Cetacean Res. Manage. **3**, 65–76.

Stanistreet, J. E., Nowacek, D. P., Read, A. J., Baumann-Pickering, S., Moors-Murphy, H. B., and Van Parijs, S. M. (**2016**). "Effects of duty-cycled passive acoustic recordings on detecting the presence of beaked whales in the northwest Atlantic," J. Acoust. Soc. Am. **140**, EL31–EL37.

Thomisch, K., Boebel, O., Zitterbart, D. P., Samaran, F., Van Parijs, S., and Van Opzeeland, I. (**2015**). "Effects of subsampling of passive acoustic recordings on acoustic metrics," J. Acoust. Soc. Am. **138**, 267–278.

Torterotot, M., Royer, J.-Y., and Samaran, F. (**2019**). "Detection strategy for long-term acoustic monitoring of blue whale stereotyped and non-stereotyped calls in the Southern Indian Ocean," *OCEANS 2019*, Marseille, France (17–20 June) (IEEE, New York), pp. 1–10.

Torterotot, M., Samaran, F., Stafford, K. M., and Royer, J.-Y. (**2020**). "Distribution of blue whale populations in the Southern Indian Ocean based on a decade of acoustic monitoring," Deep Sea Res. Part II **179**, 104874.

Van Parijs, S., Clark, C., Sousa-Lima, R., Parks, S., Rankin, S., Risch, D., and Van Opzeeland, I. (**2009**). "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," Mar. Ecol. Prog. Ser. **395**, 21–36.

15 April 2024 10:12:38

1.4 Effects of duty cycle on passive acoustic monitoring metrics : The case of odontocetes vocalizations

Effects of duty cycle on passive acoustic monitoring metrics: the case of odontocetes vocalizations

Mathilde Michel,[1] Julie Béesau,[1] Maëlle Torterotot,[1] Nicole Todd,[2] and Flore Samaran[1]
[1] *CNRS Lab-STICC, ENSTA, IPP, Brest, 29200, France*
[2] *MaREI Centre, Environmental Research Institute, University College Cork, Cork, P43 C573, Ireland*

(Dated: 11 September 2025)

Passive Acoustic Monitoring (PAM) is critical for long-term odontocete monitoring using autonomous recording devices. However, technical constraints such as storage capacity and data processing limitations often require temporal subsampling. This study investigates how varying duty cycles (50% to 10%) and listening periods (1 minute to 6 hours) affect the detection of delphinid whistles and clicks, and harbor porpoise clicks. Two types of instruments were used: broadband recorders for whistles and F-PODs for clicks. As each device offers different configuration options, subsampling schemes were tailored to each signal type. The impact of duty cycles on seasonal patterns was evaluated using daily positive detection minutes and hours, and diel patterns were assessed using hourly positive minutes and daily positive detection ratios. Results indicate that higher duty cycles (50%) better preserve temporal pattern representations, particularly in high-activity sites, across both instruments and signal types. Lower duty cycles reduce the quality of data representation, especially in low-activity areas. Short listening periods (5–30 minutes) most closely approximate metrics from continuous recordings. These findings highlight the importance of adapting subsampling strategies to instrument capabilities and the overall level of acoustic activity, which varies across taxa and sites, to obtain an accurate representation of odontocete acoustic presence.

[https://doi.org(DOI number)]

[XYZ] Pages: 1–18

I. INTRODUCTION

Odontocetes, such as dolphins (*Delphinidea*) and the harbor porpoise (*Phocoena phocoena*), are exemplary subjects for passive acoustic monitoring due to their frequent and distinguishable vocalizations. Dolphins produce a variety of acoustic signals, including whistles and clicks, which can be used to infer a range of behavioral contexts. Although species-level discrimination of delphinid vocalizations can be achieved in some regions, the high similarity of vocalizations among species and their frequent spatial and temporal overlap often prevent reliable separation of detections[1]. Therefore, delphinid signals are commonly analyzed as a single group. Whistles are often associated with social interactions and group cohesion, whereas clicks are mainly used for echolocation during navigation and foraging[2,3]. In contrast, harbor porpoises exclusively emit narrow-band high-frequency (NBHF) clicks; behavioral information for this species is primarily derived from parameters such as inter-click intervals and click rates, which can indicate foraging activity or general orientation[4]. Cetacean vocalizations are essential for ascertaining species presence, distribution, and acoustic behaviors in various habitats[5,6]. However, variable species distributions, calling behaviors, and acoustic characteristics (i.e., call types and frequency range) pose significant challenges to long-term monitoring, particularly across vast and/or remote areas.

Passive acoustic monitoring (PAM) is widely used to investigate seasonal and spatial distribution as well as the acoustic behavior of cetaceans[7–9]. It is favored for the collection of long-term data and for monitoring expansive areas[10,11]. In contrast to visual methodologies, PAM platforms can record audio continuously, regardless of adverse weather conditions or the submersion of individuals below the water surface[12]. Long-term monitoring of cetaceans is often performed using autonomous platforms anchored to the seabed, such as broadband recorders or those with integrated detection algorithms (e.g., the F-POD for odontocete clicks), which archive data until retrieval[7–11]. However, the battery capacity and onboard storage limitations of these devices often constrain deployment duration and continuous recording capacity[13]. Platform deployment and maintenance can be logistically challenging and expensive, particularly over long-term surveys with numerous monitoring periods. Common strategies to overcome PAM's limitations include reducing the sampling rate (which can lead to species being missed) or implementing temporal subsampling[14]. The latter strategy entails defining a duty cycle in which the recorder only captures data during a designated, repeated interval, such as 30 minutes every hour. By modulating the listening proportion and duration, various subsampling schedules can be designed. However, inappropriate subsampling can affect the ecological conclusions derived from PAM studies[15–18].

Hence, an effective subsampling scheme should consider the research objective, existing knowledge of the acoustic behavior of the target species, and constraints related to battery life, storage capacity, and recorder deployment plans[13].

Previous studies have primarily focused on the impact of duty cycles on species producing long, stereotyped, and continuous vocalizations, such as baleen whale songs[15,16], or on specific odontocetes like killer whales, whose vocal behavior has been studied using broadband recorders[17,18]. Some odontocetes producing clicks, such as beaked whales, have also been examined under duty-cycled recordings[19], but both their vocalizations and the behavioral contexts in which they are emitted differ from those of dolphins and harbor porpoises. Duty cycle strategies have received little attention for these latter species, which produce irregular and short-duration vocalizations including clicks and whistles. Dolphin clicks, for example, can vary considerably in detectability depending on context and recording configuration[20], highlighting the challenges of subsampling for these species. While individual clicks are brief and inter-click intervals are often regular, the overall sequences of clicks can vary in duration and occurrence, which may increase the risk of missed detections or misinterpretation under subsampling. Furthermore, temporal subsampling on F-PODs, recorders specifically designed for detecting clicks from dolphins and harbor porpoises, has not been previously investigated.

This study addresses the current lack of standardized guidance for selecting duty cycles in passive acoustic monitoring (PAM) of harbor porpoises and dolphins in general. Duty cycles are often chosen empirically, without evaluating their potential impact on ecological conclusions, particularly for species producing short, irregular vocalizations such as dolphin whistles and clicks, or harbor porpoise clicks. To address this gap, the effects of different duty cycles, varying both recording percentage (10–50%) and listening period (1 minute to 6 hours), were systematically evaluated by comparing subsampled datasets to continuous recordings of dolphin and harbor porpoise vocalizations. By combining broadband recorders for whistles and F-PODs for clicks, a recorder-specific evaluation of temporal subsampling schemes was conducted. This work contributes to the standardization of PAM methodology by providing evidence-based recommendations for duty cycle selection that optimize monitoring efficiency, improve data reliability, and support consistent ecological interpretation across studies. Ultimately, this approach guides the design of structured, methodologically robust PAM deployments that balance recorder constraints with the need for accurate assessment of species presence, behavior, and temporal patterns.

II. MATERIAL AND METHODS

A. Acoustic data acquisition

The acoustic data used in this study were collected from two types of acoustic recorders deployed along the French and Irish coasts (Fig. 1). Three sites were selected: Sites G and F in Western France and Site IR in Ireland. Sites G and F were equipped with co-deployed SYLENCE-LP (RTSYS) broadband recorders and F-POD D12 (Chelonia Limited[21]), which were recorded from May 2022 to June 2023. The SYLENCE recorders were used to detect delphinid whistles, while the F-POD captured clicks from both delphinids and harbor porpoises. The SYLENCE broadband recorders were continuously sampling (duty cycle 100%) at 128 kHz with 16-bit resolution and 0 dB gain. Each recorder was equipped with an HTI-99HF hydrophone and included a high-pass filter at 3 Hz. Similarly, the F-PODs were also continuously recording throughout the deployment period. At each French deployment sites, the two instruments were integrated into a mooring system anchored to the seafloor. Two mooring configurations were designed depending on site depth: a coastal system (site F, 35-meter depth) and an offshore system (site G, 120-meter depth). At the Irish site, only an F-POD D12 was deployed from April 2021 to July 2022. This deployment was part of a broader study on passive acoustic monitoring, but only F-POD data are analyzed here[22]. The device was mounted approximately 5 meters above the seabed on an anchored mooring line and marked with a surface buoy. The depth of the water at the site was 23 meters. The F-POD was retrieved and reprogrammed every 4 to 6 months, depending on the weather conditions and accessibility of the site. All F-PODs were calibrated and standardized by the manufacturer prior to deployment and devices were rotated between deployment periods to minimize potential systematic bias related to differences in device sensitivity.

B. Acoustic data analysis

1. Delphinids whistles

Delphinid whistles were detected using the "Whistle and Moan Detector" module implemented in PAMGuard (version 2.02.15)[23] (Fig. 2a). This algorithm identifies tonal vocalizations in broadband recordings based on spectral continuity and intensity criteria. The detector was configured using the parameter set presented in Appendix A, Table IV.

Detection performance was evaluated through a visual inspection of the spectrograms using the web-based annotation platform APLOSE[24]. For site G, performance was calculated based on detections of one data day in May 2022, resulting in precision 64% and recall 93%. For site F, due to the low number of whistle-positive minutes per day, all recordings were manually reviewed to validate automatic detections and limit false positives. Precision and recall metrics are defined in de-

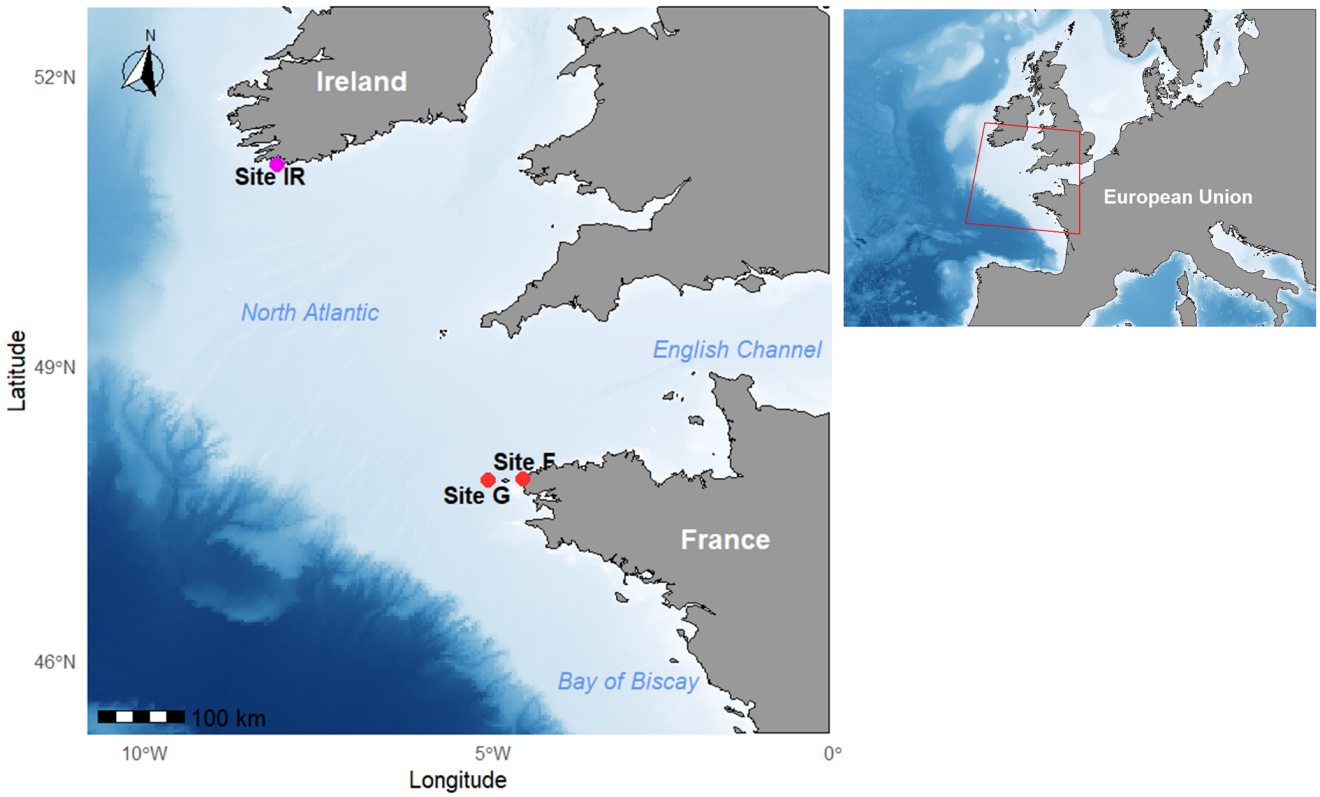


FIG. 1. Map of the study area showing the three mooring sites. The violet marker indicates the Irish site (Site IR), equipped with an F-POD recorder. The red markers indicate the French sites (Site F: coastal; Site G: offshore), each equipped with both a broadband recorder and an F-POD. Major geographic features are also labeled (France, Ireland, English Channel, Bay of Biscay, North Atlantic).

tail by Hildebrand et al. (2022)[25], who provide a comprehensive overview of performance metrics for the detection and classification of marine mammal signals.

2. Delphinids and harbor porpoise clicks

The clicks from delphinids and harbor porpoises recorded by F-POD at all sites (G, F and IR) were processed using dedicated F-POD.exe software (version 1.1) with the KERNO-F classifier[26] (Fig. 2b,c). This classifier automatically identifies and categorizes click trains into species groups based on signal characteristics. In this study, click trains were classified as "NBHF" for harbor porpoises, which emit narrow-band high-frequency clicks, or as "Other cet" for delphinid species.

To ensure data reliability at the Irish site (IR), the detection positive minutes (DPM) were visually verified following the manufacturer's recommendations[26]. This quality control step involved evaluating a random subset of DPMs, with a minimum of 100 sampling points suggested, to confirm the accuracy of the automatic classification and to filter out false positives or ambiguous detections. For the French sites (G,F), detection performance was calculated by visual inspection of spectrograms using the web-based annotation platform APLOSE[24]. The detection of Delphinid clicks was evaluated using one data day from May 2022 at site G, generating a precision of 75% and a recall of 30%. The performance of the detection of harbor porpoise clicks was evaluated on one day of data from August 2023 at site F. The precision reached 81%, while the recall was 46%.

C. Subsampling analysis

1. Subsampling schemes

a. Delphinids whistles.

Subsampling schemes were defined following Thomisch et al. (2015)[15], where a listening period L_p (in minutes) and a cycle period T_c (in minutes) determine the duty cycle D (in percent) using the ratio L_p/T_c, with N representing the number of cycles per day. The same cycle configurations as in Michel et al. (2024)[16] were applied, implementing 21 temporal subsampling schemes with L_p ranging from 1 min to 6 h and D values of 50%, 33%, and 25% (Table I). These schemes correspond to standard duty cycle configurations commonly offered by acoustic recorders and allowed us to explore different strategies by varying

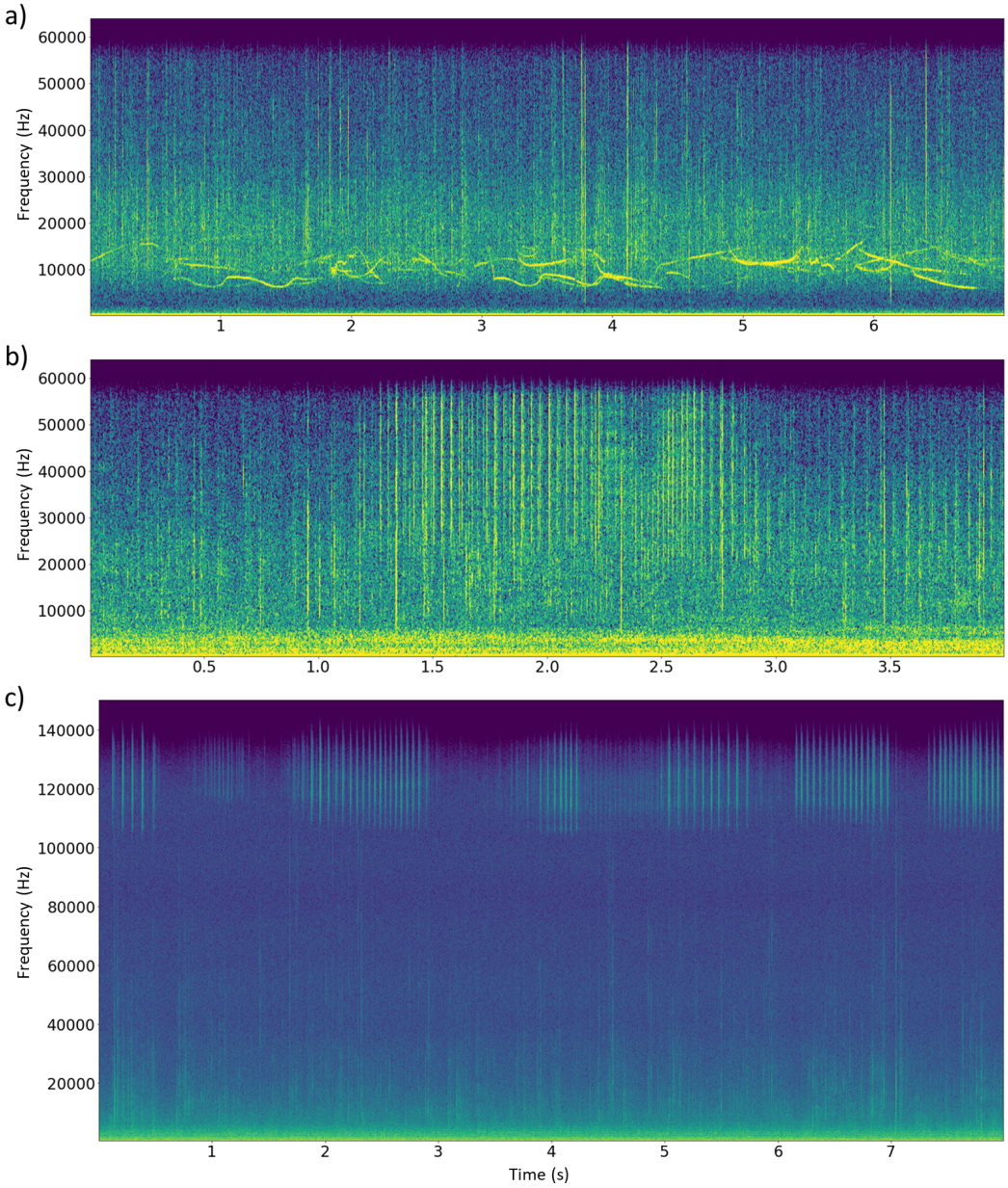


FIG. 2. Spectrograms of the three types of vocalisations used in this study: (a) delphinid whistles, (b) delphinid clicks, and (c) harbor porpoise clicks.

L_p (and consequently N) while maintaining constant D. For delphinid whistles, detection positive minutes (DPM) served as the metric: Any minute overlapping with a listening period was included in the subsampled dataset. Since listening periods could begin at any point within T_c (6 h could cover either the first or second half of a 12 h cycle), all possible non-overlapping listening phases r ($r = 2$ for $D = 50\%$, $r = 3$ for $D = 33\%$, and $r = 4$ for $D = 25\%$) were accounted for, and results were processed and averaged across all iterations to assess subsampling effects.

b. Delphinids and harbor porpoise clicks.

Unlike the broadband acoustic recorders used to detect whistles, where temporal subsampling schemes can typically be freely customized, the F-POD used to record delphinid and harbor porpoise clicks is limited to a set of four predefined subsampling modes. Each of these modes is characterized by a fixed duty cycle (D) of 50%, 33%, 20%, or 10%, and a standardized listening period (L_p) of 1 minute (Table II). In contrast to the flexible configurations available for delphinid whistles, F-POD instruments are programmed to always record only the

TABLE I. Different subsampling designs used for delphinid whistle detection. The duty cycle, D (%), is the ratio between the listening period, L_p, and the cycle period, T_c, repeated throughout a day.

D = **50%**			D = **33%**		D = **25%**	
r = 2			**r = 3**		**r = 4**	
Listening period L_p	Cycle period T_c	Cycles per day N	Cycle period T_c	Cycles per day N	Cycle period T_c	Cycles per day N
1min	2min	720	3min	480	4min	360
5min	10min	144	15min	96	20min	72
15min	30min	48	45min	32	1h	24
30min	1h	24	1h30	16	2h	12
1h	2h	12	3h	8	4h	6
3h	6h	4	9h	2.66	12h	2
6h	12h	2	18h	1.33	24h	1

first minute of each duty cycle. Consequently, unlike for whistle data, where results can be averaged across multiple iterations for each duty cycle and listening period combination, F-POD data are directly constrained by the fixed one-minute recording segments associated with each duty cycle mode. Importantly, the KERNO-F classifier used to detect clicks does not rely solely on the acoustic content within the isolated recorded minute but also considers click characteristics in adjacent time windows. To accurately simulate the effects of reduced duty cycles and maintain the integrity of the detection process, the KERNO-F classifier was applied to artificially subsampled datasets reflecting the specific recording schedules of the four duty cycle modes. This approach ensures that detection results remain consistent with the classifier's operation in real deployment scenarios.

2. Acoustic metrics

To assess the influence of duty cycle on temporal detection patterns, data sets were selected for each vocalization type from two monitoring sites showing contrasting levels of acoustic activity: one with high detection rates and one with low detection rates. The same site pairs were used for seasonal and diel pattern analyses. High detection sites included site G for delphinid whistles and clicks, and site IR for porpoise clicks (Fig.3a,c,e; 4a,c,e). Low detection sites included site F for delphinid whistles and clicks, and site G for porpoise clicks (Fig.3b,d,f; 4b,d,f).

a. Seasonal patterns.
The distribution of Positive Detection Minutes (DPM) and Daily Positive Hours (DPH) was analyzed throughout the year to evaluate the influence of duty cycling on seasonal trends. Although DPM offers finer temporal resolution, DPH reflects a broader pattern of detection throughout the day. Time series from subsampled datasets were compared with continuous data using R^2 values, providing a quantitative assessment of how well seasonal patterns were preserved across different duty cycles and detection contexts. Examples of these relationships for selected cases are provided in Appendix B (Fig. 11 and Fig. 12). For each duty cycle and vocalization/site combination, R^2 was calculated between subsampled and continuous time series of DPM and DPH. A value of R^2 close to 1 indicates that the seasonal trend of the subsampled data is closely aligned with that of the continuous data.

b. Diurnal patterns.
The influence of duty cycling on diurnal patterns was evaluated by calculating the mean number of positive detection minutes per hour (HPM). For each duty cycle, HPM time series were generated and compared with the continuous dataset using linear regression (R^2). A value of R^2 close to 1 indicates that the subsampled data closely reflect the diurnal trends of the continuous data.

c. Daily positive detection minute ratio estimation.

The daily positive detection minute ratio, γ_j, was estimated as the average of positive detection minutes in all independent subsampling configurations for each iteration. Specifically, for each realization r and cycle i, the positive detection minutes were averaged over the N cycles for the jth day:

$$\overline{\gamma}_{r,j} = \frac{1}{N}\sum_{i=1}^{N}\gamma_{r,j}, \tag{1}$$

where $\overline{\gamma}_{r,j}$ represents the estimate of the positive detection minute ratio of the rth realization in all N cycles of the jth day.

To evaluate the impact of subsampling on the representation of daily positive detection rates, the ratio between the estimated daily positive detection minute rate

TABLE II. Different subsampling designs used for delphinid and harbor porpoise click detection. The duty cycle, D (%), is the ratio between the listening period, L_p, and the cycle period, T_c, repeated throughout a day.

	D = **50%**		D = **33%**		D = **20%**		D = **10%**	
Listening period L_p	Cycle period T_c	Cycles per day N	Cycle period T_c	Cycles per day N	Cycle period T_c	Cycles per day N	Cycle period T_c	Cycles per day N
1min	2min	720	3min	480	5min	288	10min	144

$(\overline{\gamma}r, j)$ from subsampled data and the true continuous data (γtrue) was calculated:

$$\text{ratio}_{r,j} = \frac{\overline{\gamma}_{r,j} - \gamma_{\text{true}}}{\gamma_{\text{true}}}, \tag{2}$$

This calculation was performed for each subsampling scheme (with different duty cycles) and for each day, depending on the type of vocalization, to assess the variability of the estimated daily positive detection minute rate relative to continuous data. All analyses were performed using R (R Core Team, 2021)[27].

III. RESULTS

A. Data overview: seasonal and diurnal patterns

The acoustic data sets acquired from various locations showed different seasonal and diurnal patterns in the occurrence of delphinid and harbor porpoise vocalizations (Fig. 3, 4).

For delphinid whistles, the site with high detection (site G) demonstrated a clear seasonal peak in acoustic activity from July to October (mean = 283 DPM; range = 14–852 DPM, Table III) (Fig. 3a). The diurnal pattern indicated a consistent presence throughout the day, with a pronounced increase in activity during nighttime hours, particularly between 22:00 and 02:00 (HPM $\simeq$ 14), and a noticeable decrease in activity during the late morning hours (HPM $<$ 8) (Fig. 4a). In contrast, the site with low detection (site F) showed a bimodal seasonal pattern, with a primary peak in summer and a secondary peak in winter (December to April), coinciding with a marked reduction in activity at site G (mean = 16 DPM; range = 0–52 DPM, Table III). The diurnal activity was more diffuse, with minor peaks around 09:00 (HPM $\simeq$ 1) and 16:00–17:00 (HPM $\simeq$ 0.9) (Fig. 4b).

Regarding delphinid clicks, the high detection site (site G) exhibited a seasonal peak in August (mean = 68 DPM; range = 0–228 DPM, Table III) (Fig. 3c), with a strong diurnal pattern marked by elevated nighttime activity between 18:00 and 05:00 (HPM $\simeq$ 3–5), and a sharp decline in detections during the late morning hours (HPM $<$ 1) (Fig. 4c). The low detection site (site F) presented weaker seasonal trends (mean = 7 DPM; range = 0–37 DPM, Table III) (Fig. 3d), with a seasonal peak observed in January–February. A distinct diurnal increase in activity was observed from 18:00 to 23:00 (max. $>$ 0.75 HPM), extending until 06:00, with minimal detections between 08:00 and 15:00 (HPM $<$ 0.2) (Fig. 4d).

For harbor porpoise clicks, Site IR did not show a clear seasonal trend, maintaining a year-round presence (mean = 44 DPM; range = 1–149 DPM, Table III) (Fig. 3e). Diurnal activity peaked between 07:00 and 16:00 (HPM $>$ 2), decreasing considerably during nighttime hours (HPM $<$ 1.5) (Fig. 4e). In contrast, Site G exhibited a seasonal increase in acoustic detections between June and October (mean = 2 DPM; range = 0–9 DPM, Table III), while diurnal patterns were irregular with poorly defined peaks around 22:00 or 06:00 (Fig. 4f).

B. Effect of subsampling on seasonal patterns

The effects of subsampling strategies on seasonal patterns in the occurrence of delphinid and harbor porpoise vocalizations were examined by computing the mean R^2 values across all possible non-overlapping listening phases (r) for each combination of duty cycle and listening period, using two aggregation metrics: DPM and DPH. In general, the patterns observed with DPH mirrored those obtained with DPM, although the R^2 values were generally lower with DPH, particularly at shorter listening periods. This discrepancy likely stems from the reduced temporal resolution inherent in the DPH metric, which moderates the contrasts of seasonal variation. An exception to this was observed in delphinid whistles with low detection rates, where DPH produced slightly higher R^2 values than DPM for longer listening periods (Fig. 5b).

Focusing on the DPM metric, the results showed that higher duty cycles consistently improved the agreement between continuous and subsampled data, with optimal mean R^2 values attained at a 50% duty cycle across all datasets. This effect was particularly pronounced in datasets with low detection rates, where reduced duty cycles (33% and 25%) resulted in a substantial decline in subsampling scheme performance. In contrast, data sets with high detection rates maintained high R^2 values even under reduced duty cycles. For example, data sets of delphinid whistles with high detection rates and short listening periods (i.e., 1-15 minutes) resulted in values of R^2 that remained close to 1, even at a duty cycle of 25% (Fig.5a). However, lower duty cycles combined with longer listening periods (over 30 minutes) led to a noticeable drop in R^2, especially in low-detection-rate datasets

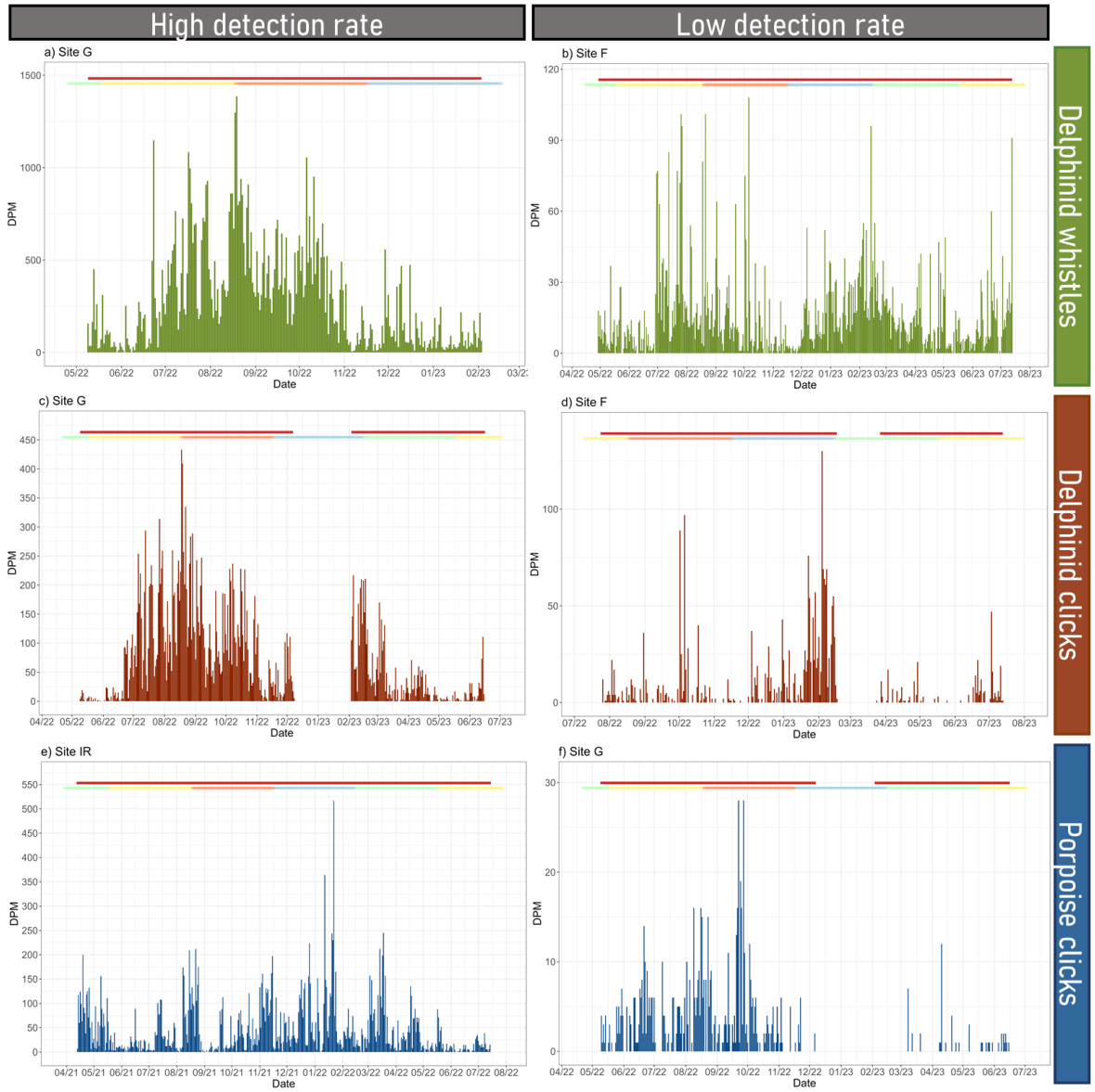


FIG. 3. Distribution of daily positive minutes (DPM) for three types of vocalizations: a) and b) for delphinid whistles, c) and d) for delphinid clicks, and e) and f) for porpoise clicks. The left column shows sites with higher detection rates, while the right column displays sites with fewer DPMs. Red horizontal bars indicate the recording periods for each site. Horizontal bars in yellow, orange, blue, and green represent the seasons: summer, autumn, winter, and spring, respectively. The scale of the bars differ between sites and vocalization types.

where the alignment between subsampled and continuous data weakened. In such cases, the mean values of R^2 dropped below 0.9 even for the duty cycle of 50% and dropped to 0.5 or lower for longer listening periods (3–6 hours) with duty cycles of 33% or 25% (Fig.5b).

A similar trend was observed for delphinid clicks and porpoise clicks. In high detection rate data sets, the R^2 values remained high (> 0.9) even at a 20% duty cycle (Fig.6a,c), with only modest decreases at the lowest duty cycle of 10%. These results indicate a higher resilience to subsampling in datasets with dense detections. In contrast, for sites with lower detection rates, the R^2 values dropped more rapidly. In these cases, R^2 decreased steadily at or below a duty cycle of 33%, and for porpoise clicks with low detection rates, decreased below 0.5 at 10% (Fig.6b,d). Note that for click datasets, each combination of listening period and duty cycle yields a single R^2 value; unlike delphinid whistles, no averaging across iterations was performed.

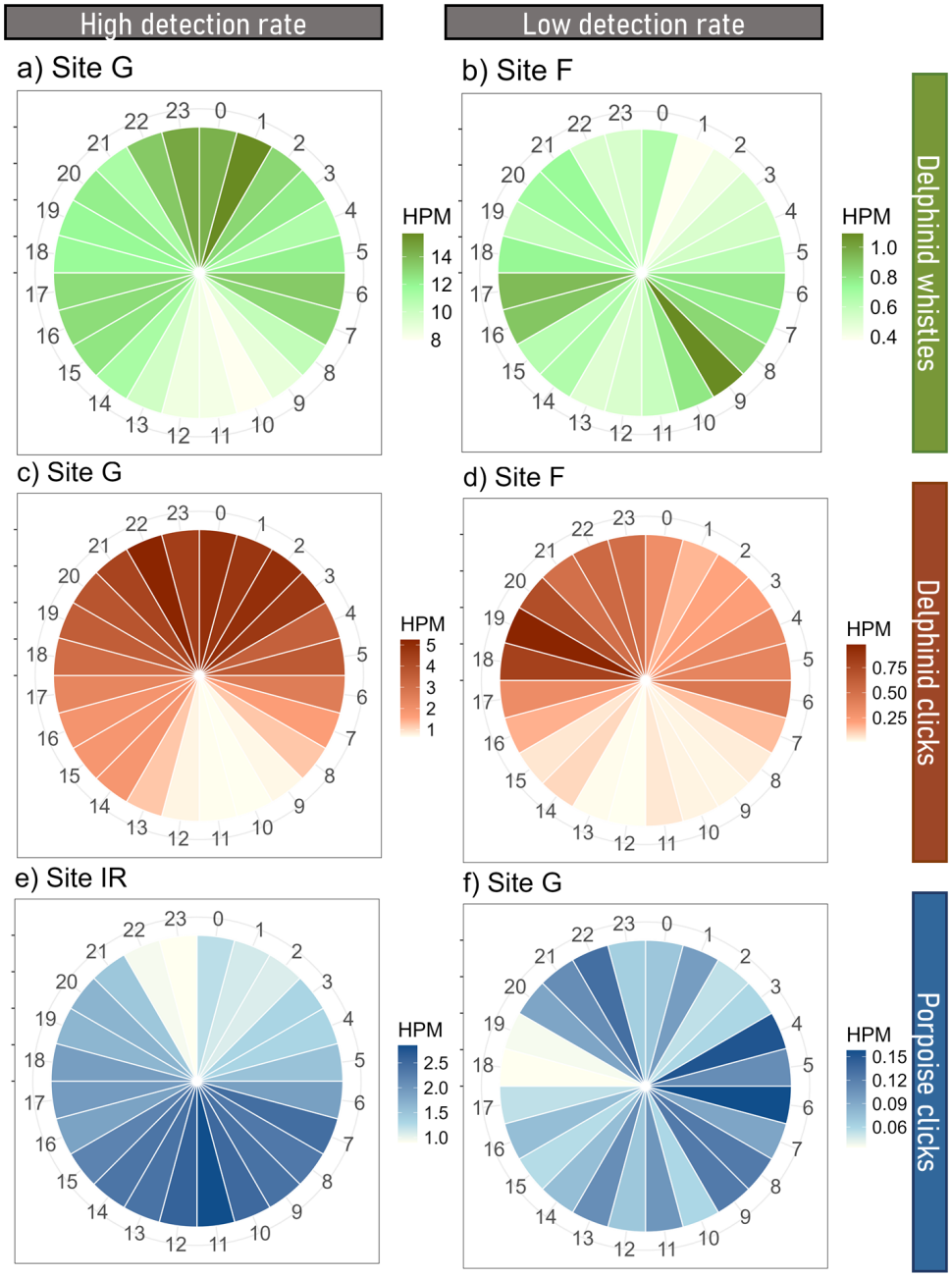


FIG. 4. Distribution of average detection positive minutes per hour (HPM) for three types of vocalizations: a) and b) for delphinid whistles, c) and d) for delphinid clicks, and e) and f) for porpoise clicks. The left column shows sites with high detection rates, while the right column displays sites with low detection rates. The intensity of the color represents the average value of positive minutes per hour, and the scales of the plots differ between sites and vocalization types.

Upon examination of the R^2 values acquired over a 1-minute listening period across varying duty cycles, it is evident that the type of vocalization, regardless of the detection rate, has little influence on delphinids, as both whistles and clicks demonstrate comparable performance. In contrast, harbor porpoise clicks consistently exhibit lower R^2 values at all duty cycles for sites with low detection rates. This reflects the generally lower frequency of porpoise detections at these sites, which makes the subsampling scheme less effective compared to delphinid clicks (Fig. 13).

TABLE III. Summary of detection positive minutes (DPM) for delphinid and harbor porpoise vocalizations across selected monitoring sites. Values include the mean, percentile range (p5–p95), and maximum observed values. The 5th and 95th percentiles indicate the range within which 90% of the observed DPM values fall.

Detection level	**Site**	**Vocalization**	**Mean**	**Percentile [p5–p95]**	**Max**
High detection	G	Delphinid whistles	283	[14–852]	1386
	G	Delphinid clicks	68	[0–228]	433
	IR	Porpoise clicks	44	[1–149]	517
Low detection	F	Delphinid whistles	16	[0–52]	108
	F	Delphinid clicks	7	[0–36.7]	130
	G	Porpoise clicks	2	[0–9.25]	28

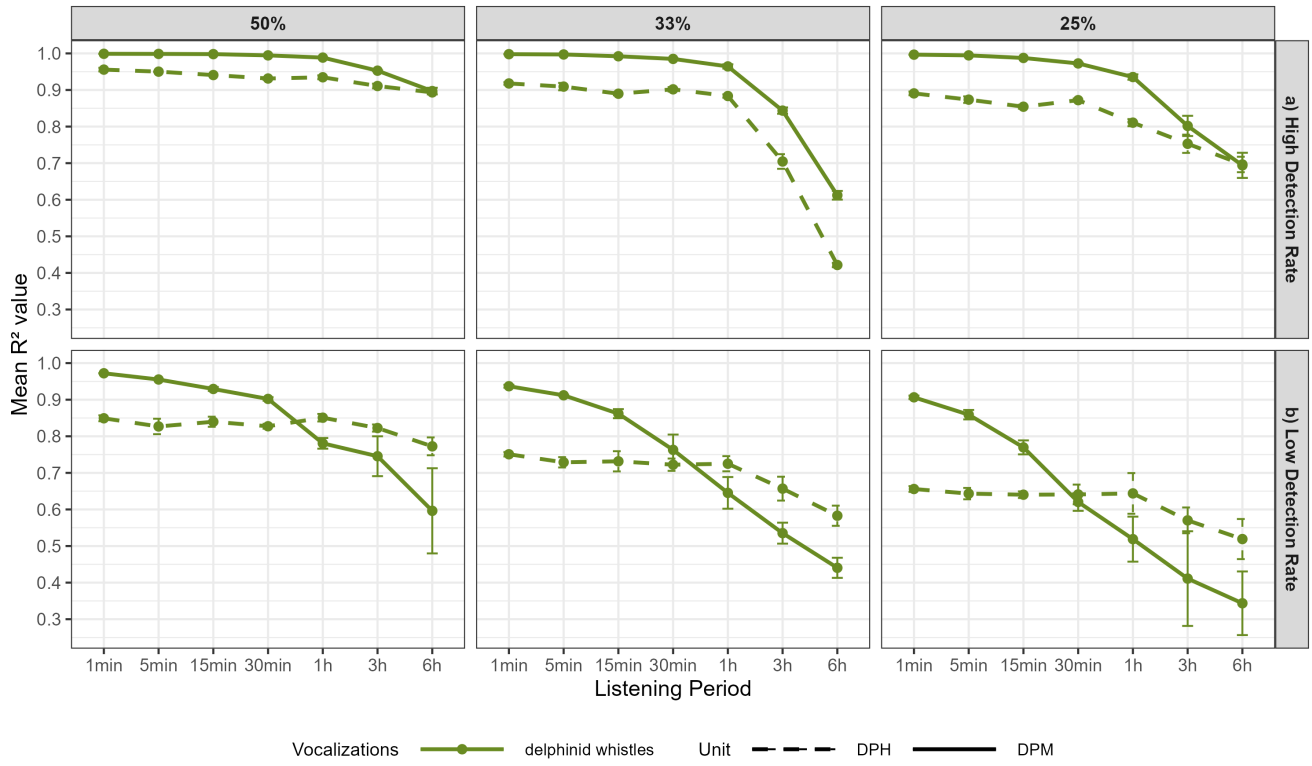


FIG. 5. Comparison of different subsampling schemes for delphinid whistles to capture seasonal patterns. Each row represents a detection rate category: a) High and b) Low. Each column corresponds to a duty cycle, D (from left to right, 50%, 33%, and 25%). Each point represents the mean R^2 value for different listening periods, L_p, within a given duty cycle (from left to right: 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, and 6 h, respectively). Vertical green lines indicate the standard deviations of the R^2 values. The dashed line represents the R^2 values for DPH, while the solid line represents the R^2 values for DPM.

C. Diurnal patterns

The results indicate that elevated duty cycles, particularly 50%, resulted in slightly higher R^2 values for high detection rate datasets at shorter listening periods; however, all datasets experienced substantial declines in R^2 at increasing listening periods, regardless of detection rate (Fig. 7, 8).

For delphinid whistles, a marked decline in the R^2 values was observed when the listening duration exceeded 5, 15, and 30 minutes for duty cycles of 33, 25, and 50%, respectively, with R^2 values dropping close to zero

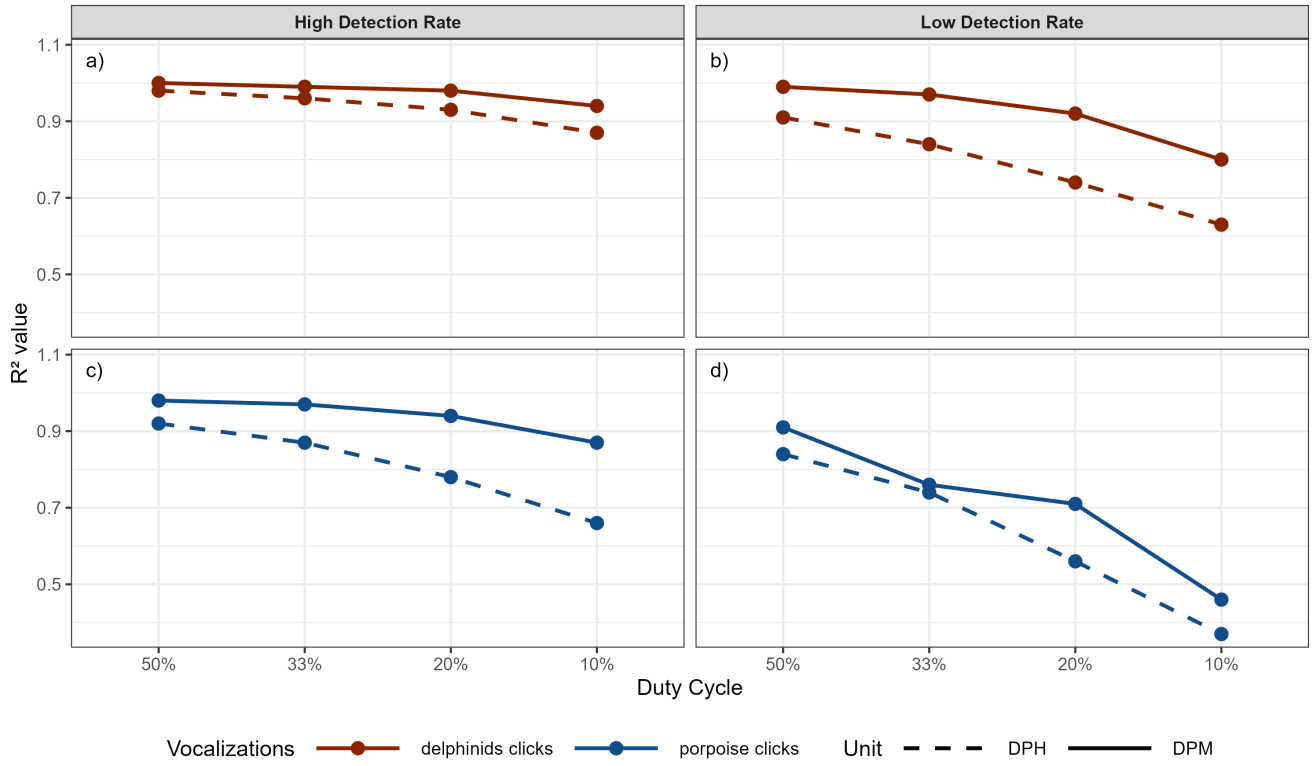


FIG. 6. Comparison of different subsampling schemes for high and low detection rate of delphinids (a and b) and and harbor porpoise (c and d) clicks, respectively, to capture seasonal patterns. The left column corresponds to sites with high detection rates, and the right column to sites with low detection rates. Each point represents the R^2 value for a given duty cycle, D (from left to right: 50%, 33%, 25%, and 10%). The dashed lines indicate R^2 values for the DPH metric, while the solid lines correspond to the DPM metric.

(Fig.7). This pattern was observed in both high and low detection rate datasets, with the decline more pronounced in low detection rate datasets due to the sparser presence of detections. For a duty cycle of 33%, the values of R^2 increased again for longer listening intervals (3 and 6 hours), reaching approximately 0.8 in high detection rate datasets and 0.5 in low detection rate datasets (Fig.7a,b). However, short listening intervals, particularly 1 and 5 minutes, provided the most reliable and representative results to capture diurnal trends.

For clicks, high R^2 values ($R^2 > 0.9$) were observed in data sets with a high detection rate across all duty cycles (Fig.8a,c). Even in low detection rate data sets for delphinid clicks, a relatively stable fit persisted across all duty cycles (Fig.8b). However, for harbor porpoise clicks in low detection rate datasets, a noticeable reduction in R^2 was evident as the duty cycle decreased. Although R^2 remained high (0.95) for duty cycle 50%, it dropped to 0.6 for duty cycle 10%, indicating a substantial loss in the subsampling scheme's ability to accurately capture the underlying diurnal patterns (Fig.8d).

When comparing the R^2 values at 1-minute listening durations across duty cycles, the type of vocalization had little influence on the subsampling scheme performance for delphinids: Both whistles and clicks yielded comparable results. In contrast, for harbor porpoise clicks, sites with low detection rates consistently showed lower R^2 values at all duty cycles. This is a direct consequence of the low frequency of porpoise detections at these sites, resulting in a higher likelihood of missing acoustic events under subsampling, relative to delphinid clicks (Fig.14).

D. Daily positive detection minute ratio estimation

The analysis shows that in all datasets, higher duty cycles (i.e. 50%) produced ratio values close to zero with low variability, indicating reliable estimates of daily positive minutes (Figs.9, 10).

Datasets characterized by high detection rates demonstrated greater robustness to subsampling, whereas those with low detection rates displayed more deviations and variability in the ratios. For delphinid whistles, a 50% duty cycle yielded ratios nearing zero, with a lower standard deviation than those observed with lower duty cycles (33% and 25%), particularly for shorter listening durations (1–30 minutes) (Fig. 9).

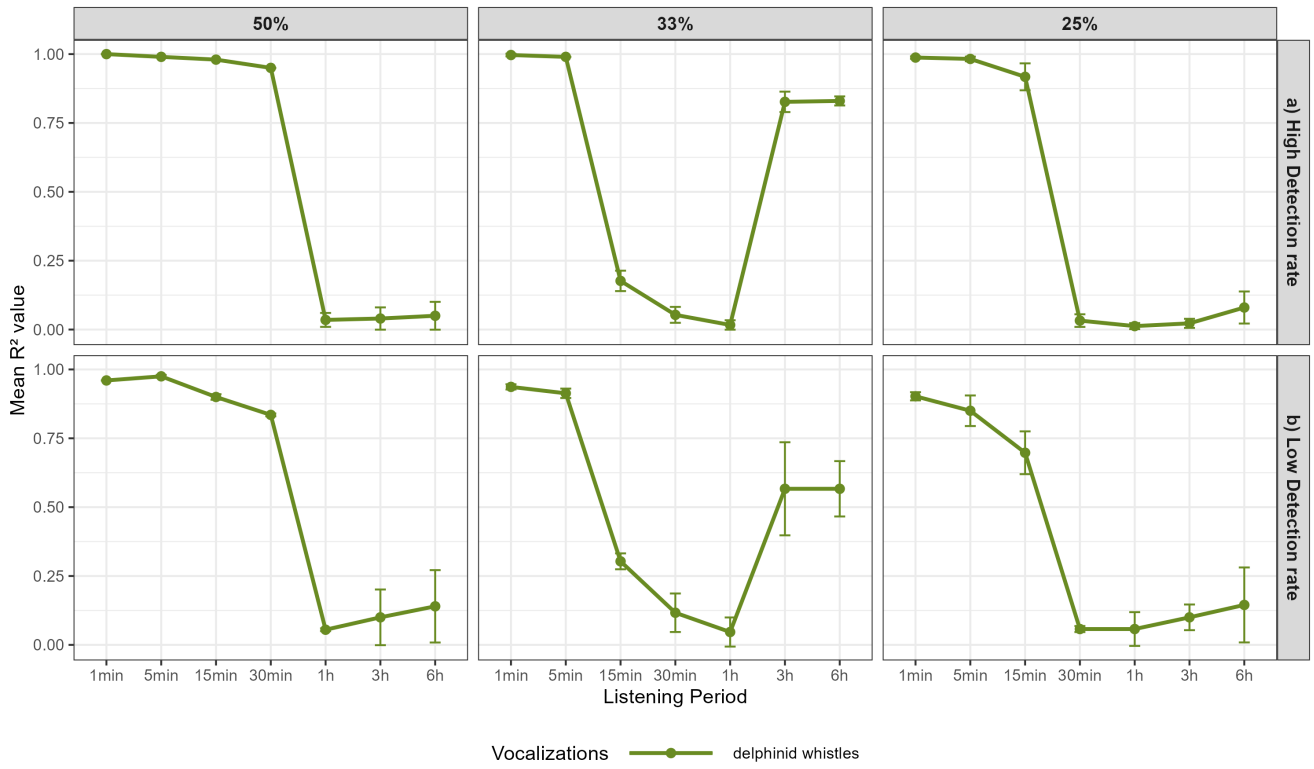


FIG. 7. Comparison of different subsampling schemas for delphinid whistles to capture diurnal patterns. Each row represents a detection rate category: a) High and b) Low. Each column corresponds to a duty cycle, D (from left to right, 50%, 33%, and 25%). Each point represents the mean R^2 value for different listening periods, L_p, within a given duty cycle (from left to right: 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, and 6 h, respectively). Vertical green lines indicate the standard deviations of the R^2 values.

In datasets with low detection rates, notably with a 25% duty cycle and extended listening periods (1–6 hours), median ratio values fell below zero, indicating a systematic underestimation of daily detections. For example, the median ratio approached –0.5 at L_p 6h, indicating a 50% reduction in detection positive minutes. Conversely, cases of extreme positive ratios (up to +3) were occasionally observed, suggesting that in rare instances, the number of detection positive minutes could be overestimated by 200% or more, resulting in a tripling of the actual value. These substantial deviations were particularly prominent in datasets with sparse detections.

For acoustic click analysis, analogous patterns were discerned. In data sets characterized by a high detection rate, the ratios remained predominantly close to zero regardless of the duty cycle employed (Fig.10). In contrast, in data sets with a low detection rate, the ratios exhibited increased variability with reduced duty cycle (i.e., 10-20%) (Fig.10). For example, with regard to harbor porpoise clicks, the median ratio for a 10% duty cycle decreased to nearly -1, reflecting a near-complete loss of daily detections. However, extreme positive outliers (ratios > 2.5, occasionally > 5) were observed, suggesting that on certain days, detections could be overestimated by more than 500%, resulting in up to six times more detection positive minutes than actually occurred. Although these extreme values were less frequently encountered, they underscore the pronounced variability and unpredictability associated with subsampling effects in low-detection rate data sets under minimal duty cycles.

IV. DISCUSSION

This study assessed the influence of temporal subsampling on quantitative metrics employed in passive acoustic monitoring of odontocetes, taking into account the types of vocalization (whistles and clicks), the detection densities, and the characteristics of the recording devices (broadband hydrophones versus F-POD). The impact of temporal subsampling on the metrics studied depended primarily on the period and duration of the recording cycles, as well as on the density of odontocete vocalizations. These factors strongly influenced the representativeness of the derived metrics. Although seasonal patterns remained relatively robust to subsam-

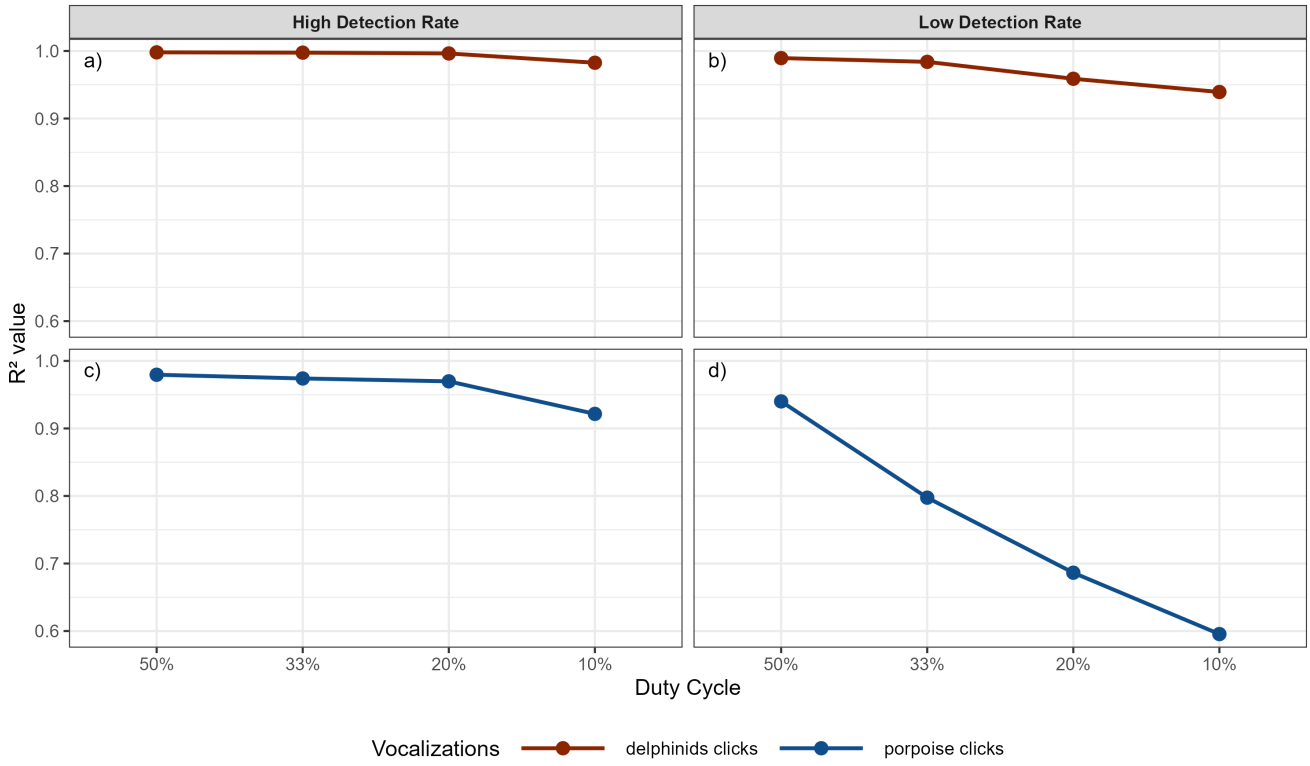


FIG. 8. Comparison of different subsampling schemes for high and low detection rates of delphinid (a and b) and harbor porpoise (c and d) clicks, respectively, to capture diurnal patterns. Each point represents the R^2 value for a given duty cycle, D (from left to right: 50%, 33%, 25%, and 10%).

pling, diurnal cycles and daily detection estimates were more vulnerable, especially in contexts where vocalizations were sparse or irregular. This was particularly evident for whistles, which tend to be produced in a highly variable and unpredictable manner compared to clicks.

A. Effects of temporal subsampling on seasonal and diurnal patterns

Seasonal patterns were still evident using a moderate subsampling scheme (25–33%) when acoustic activity was elevated. In these scenarios, the R^2 values remained in excess of 0.9, corroborating results of studies on blue whale songs that indicate consistent high calling rates that maintain integrity of seasonal trends even under conditions of reduced sampling[16]. In contrast, at sites characterized by low activity, the agreement between subsampled and continuous data notably declined below a duty cycle of 33%, with R^2 values often falling below 0.5. These observations are consistent with the findings of Thomisch et al. (2015)[15] and Stanistreet et al. (2016)[19], who reported that recordings with low duty cycles generally fail to capture seasonal variability in species exhibiting sparse or clustered vocal activity. In contrast, diurnal trends demonstrated a clear higher sensitivity to subsampling, particularly regarding whistles. Across all duty cycles tested, the performance of subsampling schemes showed considerable variability, with R^2 values that reached zero. At a duty cycle of 33%, listening periods longer than 15 minutes led to a nearly complete loss of diurnal structure ($R^2 \simeq 0$), even at sites with high odontocete activity. Similarly, mysticete D-calls, which are short, variable, and strongly influenced by behavioral context, are susceptible to loss under aggressive subsampling schemes. Unlike the long, repetitive songs, D-calls are brief and irregular, making it difficult for sparse recordings to capture representative diurnal or seasonal activity[16].

Although delphinid clicks generally showed greater robustness to subsampling, this is likely due to the longer duration of their click trains and moderate repetition rates (50–100 clicks/s with inter-click intervals of 10–20 ms[28]). Delphinid clicks, while directional, often occur over longer durations and wider spatial coverage, increasing redundancy in detections and making them more resilient to temporal subsampling. These patterns are consistent with previous findings on killer whale clicks, which also demonstrated greater resilience of click-based metrics when acoustic behavior is frequent[18]. In contrast, harbor porpoise clicks are extremely brief ($\simeq$ 44–113 μs),

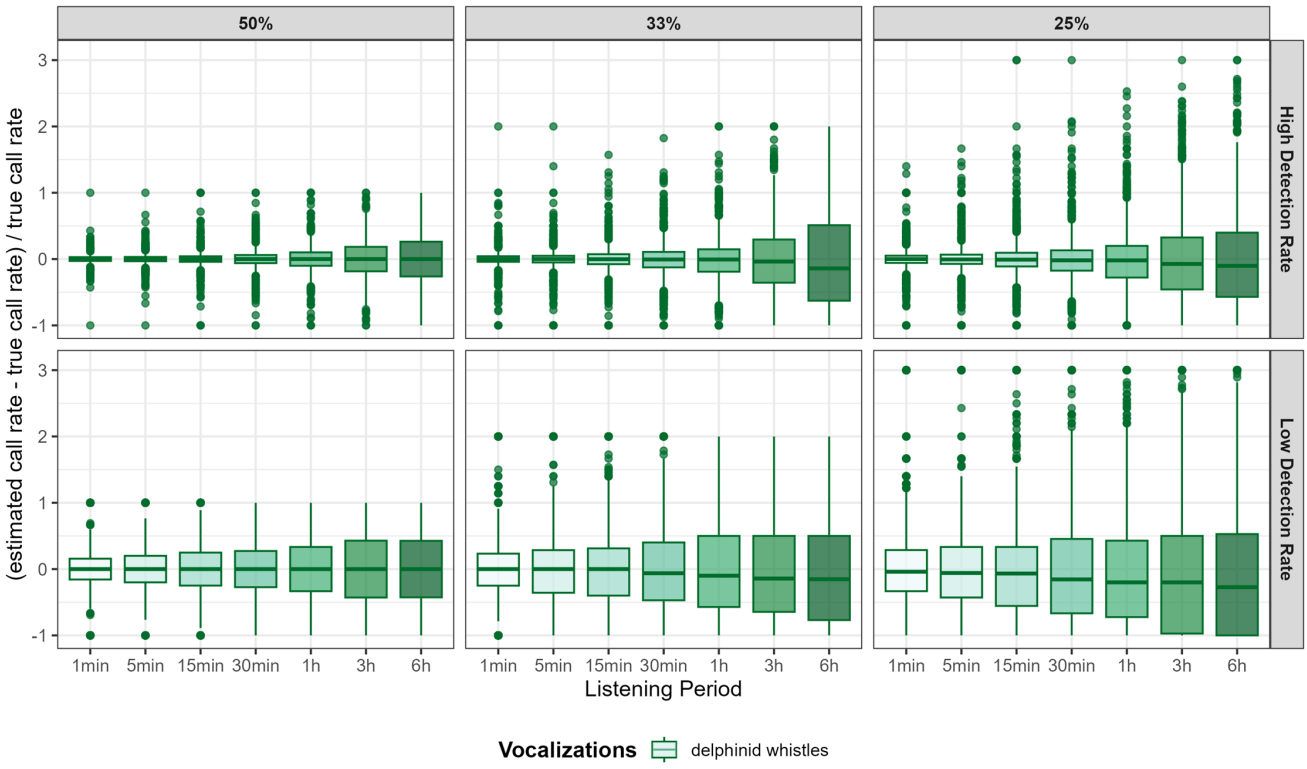


FIG. 9. Comparison of different subsampling patterns on the ratios between estimated and actual daily detection positive minutes ratio for delphinid whistles. The top row represents high detection rate conditions, and the bottom row low detection rate conditions. For each duty cycle, boxplots represent the distribution of ratios across different listening periods, L_p (from left to right: 1 min, 5 min, 15 min, 30 min, 1 h, 3 h, and 6 h, respectively). Boxplots display the median (center line), the interquartile range (25th–75th percentiles; box), whiskers extending to the 5th and 95th percentiles, and points representing outliers.

produced at very high frequencies ($\simeq$ 130 kHz), and organized in short, high-repetition rate trains (30–200 ms inter-click intervals[29,30]). Like all odontocete clicks, they are directional, but their high frequency leads to rapid absorption in seawater[31], sharply reducing detection range. Consequently, even if porpoises produce clicks nearly continuously, subsampled recordings are less likely to capture representative activity patterns, particularly in low-density sites. This combination of brevity, high frequency, and directionality explains the reduced resilience of porpoise clicks to subsampling.

The DPM metric, which aggregates at the minute level, preserves fine temporal resolution, allowing fluctuations in acoustic activity to be captured even when detections are sparse. This makes DPM particularly effective at reflecting seasonal trends under moderate to low sampling, as the risk of missing short or infrequent acoustic events is minimized. In contrast, DPH relies on hourly aggregation, which smooths out short-term variations. This can be advantageous when sampling effort is very low and detections are scarce, mitigating the impact of missed rare events, but DPH generally loses temporal resolution and may fail to capture short-term fluctuations that are essential for accurate representation of seasonal trends. Consequently, even with reduced subsampling, DPM provides a more consistent estimate of seasonal patterns than DPH, which may only be advantageous under conditions of extremely low detection density.

B. Effects of temporal subsampling on daily positive detection minute ratio estimation

Estimates of the positive detection minute ratio daily showed increased sensitivity to fluctuations in the duty cycle. Subsampling at a level of 25% for whistles introduced notable distortions, with some extreme outliers leading to overestimations of the call ratio exceeding 300% at sites with low acoustic activity. However, on average, estimated rates remained below zero, indicating an overall tendency toward underestimation. In the case of clicks, underestimation was frequently observed in scenarios characterized by low density, particularly in the F-POD data, where duty cycles of 10% resulted in median ratios approaching -1 for the detections of harbor porpoise at site G. This degree of distortion was considerably greater than the 20–30% underestimations docu-

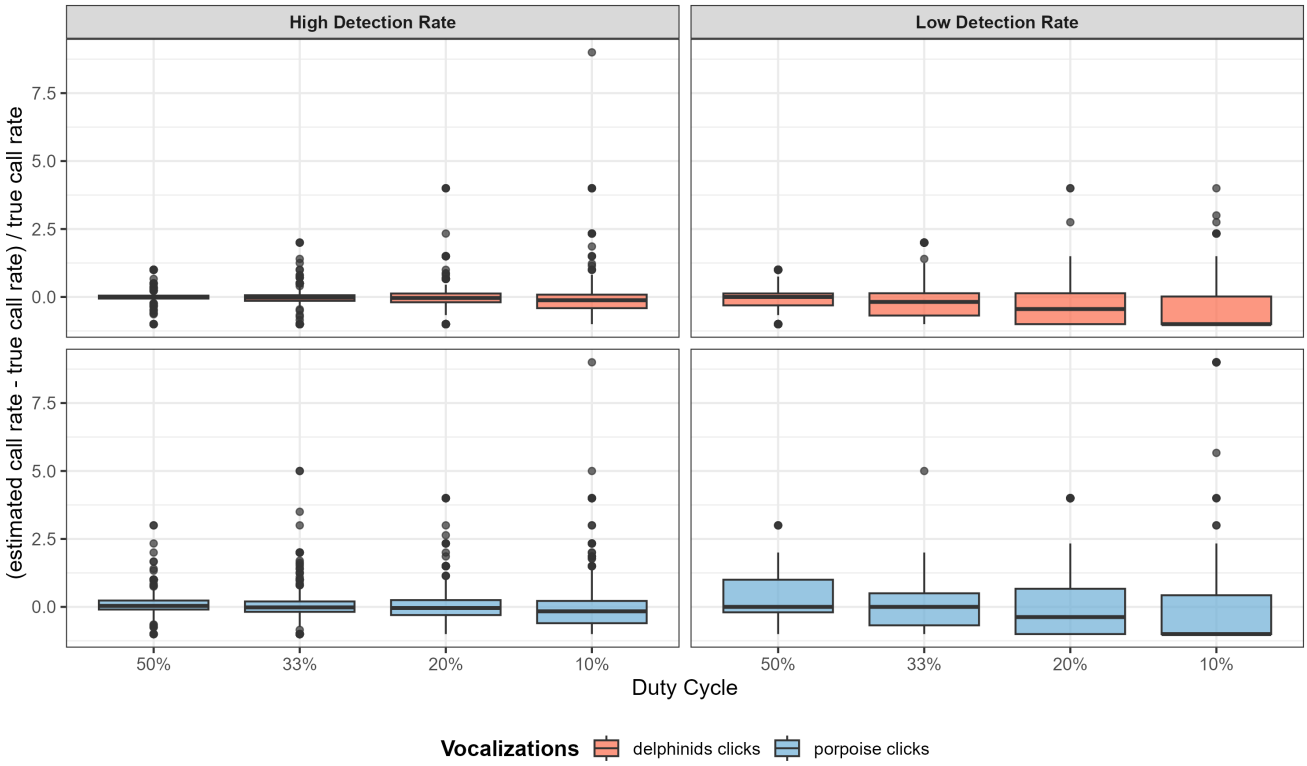


FIG. 10. Comparison of different subsampling patterns on the ratios between estimated and actual daily detection positive minutes ratio for delphinid and porpoise clicks. The left column corresponds to high detection rate conditions, and the right column to low detection rate conditions. Red boxplots represent delphinid clicks, and blue boxplots represent harbor porpoise clicks. Each boxplot represents the distribution of ratio values for a given duty cycle, D (from left to right: 50%, 33%, 25%, and 10%). Boxplots display the median (center line), the interquartile range (25th–75th percentiles; box), whiskers extending to the 5th and 95th percentiles, and points representing outliers.

mented in previous studies on killer whales[17] and beaked whales[19], indicating that greater vulnerability to detection of delphinids and porpoise clicks was biased by duty cycles. In the latter, detection rates were based on the number of positive days and mean detections per day, and subsampling effects varied between sites: those with high daily detection rates were more resilient, while low-activity sites showed stronger underestimation, a pattern similar to ours. However, their observed biases were generally less extreme than in our results, likely due to their coarser daily-scale analysis and differences in signal characteristics.

C. Implications for monitoring strategy optimization

From an operational perspective, our findings emphasize the need to customize subsampling strategies according to detection rates, signal typology, and scientific goals. The choice of temporal metrics, such as DPH or DPM, emerges as a crucial factor in this optimization process. For species exhibiting consistent acoustic presence and elevated detection rates, duty cycles as minimal as 25–33% may be adequate to derive robust seasonal and even diurnal tendencies, provided that short listening periods are used. In contrast, for signals with low detection rates, duty cycles below 50% considerably undermine the reliability of seasonal and diurnal metrics. These results are corroborated by previous studies[16,18] that demonstrate that signals regularly produced exhibit greater resistance to subsampling than short or unpredictable vocalizations.

The duration of listening periods constitutes another critical parameter. Our analysis corroborates the advisability of 5-30 minute listening intervals, aligned with Michel et al. (2024)[16], as a reasonable equilibrium that captures sufficient vocal activity without disproportionately increasing call rate variability or platform energy consumption. Importantly, very short intervals (1 minute), although potentially advantageous for maximizing temporal resolution, may conflict with technical specifications or requirements. Several instrument manufacturers have warned that such brief duty cycles may inhibit adequate system stabilization required for high-quality recordings and may increase power usage due to

frequent restart cycles. This paradox underlines the imperative to balance temporal resolution with hardware constraints and recording quality. For F-POD, which are designed for long-term deployments and efficient click detection, a duty cycle of 100% allows continuous operation for approximately five months. Reducing this to 50% can extend battery life. Although this reduction implies a theoretical loss of up to half of the potential detections, particularly in low-activity contexts, our results indicate that such losses are, on average, limited, with mean detection ratios near zero and R^2 values exceeding 80% between scenarios. This suggests that a duty cycle of 50% provides a reliable approximation of actual detection patterns, especially in the context of extended monitoring campaigns. In contrast, for short-term deployments that aim to capture detailed activity patterns over days or weeks, continuous recording remains preferable. In this context, moderate duty cycles (50%) appear to strike an optimal balance between operational autonomy and data reliability for most F-POD applications.

Based on the current study and the results presented in Michel et al. (2024)[16] on mysticete songs, it is recommended that, during simultaneous monitoring of multiple cetacean species, relatively short but meaningful listening intervals, preferably within the range of 5 to 15 minutes, be used, together with a duty cycle of approximately 50%. This methodology proves effective in capturing both the extended, episodic acoustic patterns typical of mysticete songs and the more sporadic, irregular vocalizations such as odontocete clicks and whistles. Mysticete songs and odontocete whistles are generally omnidirectional, increasing their likelihood of detection, whereas clicks are highly directional and therefore more easily missed unless the animal is oriented toward the hydrophone. Crucially, this approach retains its effectiveness in environments where species are rare or vocalize infrequently, thereby enhancing the ecological validity and precision of passive acoustic monitoring initiatives directed at diverse cetacean communities.

In conclusion, this study demonstrates that the impact of duty cycle on acoustic presence metrics depends primarily on technical parameters, such as the duration and repetition rate of listening periods. Longer listening intervals and lower duty cycle percentages reduce the reliability of both seasonal and diurnal metrics, while biological factors, including the level of acoustic activity at the recording site and the presence of diel vocalization patterns, also influence the robustness of subsampled data. The effects of subsampling are generally governed by the structure of the recording protocol; however, species-specific signal characteristics, such as the brevity, high frequency, and directionality of harbor porpoise clicks, can modulate sensitivity to reduced sampling, making some species more vulnerable than others. Currently, duty cycles are often chosen empirically, without systematic evaluation. Our results highlight that this choice can significantly affect ecological inferences and, importantly, provide a framework for decision-making when designing passive acoustic monitoring programs for odontocetes using duty cycles. By accounting for duty cycle effects, researchers can standardize methods, optimize sampling strategies, and improve the reliability of long-term monitoring for conservation purposes. Temporal subsampling, when properly planned, therefore represents an effective compromise between data volume, power consumption, and ecological representativeness.

ACKNOWLEDGMENTS

The research data from the French sites was collected as part of the CETIROISE project, funded with the support of the European Union, NextGenerationEU, France Relance and the Office Français de la Biodiversité (OFB). The authors thank the Division Technique de l'INSU – CNRS for the design of deep-sea moorings deployed in French waters. We are also grateful to the Céladon association for their logistical support during the deployment and recovery of offshore moorings, and to the Parc naturel marin d'Iroise for additional logistical assistance at coastal sites. The research data from the Irish site were facilitated by the financial support of Taighde Éireann, Research Ireland under grant number [GOIPG/2019/2173]. The authors also thank Michael Collins, the skipper of the Kestrel, and the various field assistants who supported the deployment and recovery of the devices.

AUTHOR DECLARATIONS

The authors have no conflicts to disclose.

AUTHOR CONTRIBUTIONS STATEMENT

Mathilde Michel: Conceptualization, Methodology, Formal Analysis, Original Draft Writing, Visualization. **Julie Béesau**: Formal Analysis, Writing-Review & Editing. **Maëlle Torterotot**: Formal Analysis, Writing-Review & Editing. **Nicole Todd**: Project administration, Formal Analysis, Resources, Writing-Review & Editing. **Flore Samaran**: Validation, Writing-Review & Editing, Project Administration, Funding Acquisition.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1]N. C. Van Geel, D. Risch, S. Benjamins, T. Brook, R. M. Culloch, E. W. Edwards, C. Stevens, and B. Wilson, "Monitoring cetacean occurrence and variability in ambient sound in scottish offshore waters," Frontiers in Remote Sensing **3**, 934681 (2022).

[2]B. Jones, M. Zapetis, M. M. Samuelson, and S. Ridgway, "Sounds produced by bottlenose dolphins (tursiops): A review of the defining characteristics and acoustic criteria of the dolphin vocal repertoire," Bioacoustics **29**(4), 399–440 (2020).

[3]W. Verboom and R. Kastelein, "Acoustic signals by harbour porpoises (phocoena phocoena)," Harbour porpoises: Laboratory studies to reduce bycatch 1–39 (1995).

[4]S. Kameyama, T. Akamatsu, A. Dede, A. Amaha Öztürk, and N. Arai, "Acoustic discrimination between harbor porpoises and delphinids by using a simple two-band comparison," The Journal of the Acoustical Society of America **136**(2), 922–929 (2014).

[5]J. Ivanchikova, N. Tregenza, D. Popov, G. Meshkova, R.-M. Paiu, C. Timofte, A. Amaha Öztürk, A. M. Tonay, A. Dede, U. Özsandıkçı, *et al.*, "Seasonal and diel patterns in black sea harbour porpoise acoustic activity in 2020–2022," Ecology and Evolution **14**(10), e70182 (2024).

[6]B. Zein, B. Woelfing, M. Dähne, T. Schaffeld, S. Ludwig, J. H. Rye, J. Baltzer, A. Ruser, and U. Siebert, "Time and tide: Seasonal, diel and tidal rhythms in wadden sea harbour porpoises (phocoena phocoena)," PLoS One **14**(3), e0213348 (2019).

[7]U. K. Verfuß, C. G. Honnef, A. Meding, M. Dähne, R. Mundry, and H. Benke, "Geographical and seasonal variation of harbour porpoise (phocoena phocoena) presence in the german baltic sea revealed by passive acoustic monitoring," Journal of the Marine Biological Association of the United Kingdom **87**(1), 165–176 (2007).

[8]F. Caruso, G. Alonge, G. Bellia, E. De Domenico, R. Grammauta, G. Larosa, S. Mazzola, G. Riccobene, G. Pavan, E. Papale, *et al.*, "Long-term monitoring of dolphin biosonar activity in deep pelagic waters of the mediterranean sea," Scientific Reports **7**(1), 4321 (2017).

[9]A. D. Marian, A. Monczak, B. C. Balmer, L. B. Hart, J. Soueidan, and E. W. Montie, "Long-term passive acoustics to assess spatial and temporal vocalization patterns of atlantic common bottlenose dolphins (tursiops truncatus) in the may river estuary, south carolina," Marine Mammal Science **37**(3), 1060–1084 (2021).

[10]M. S. Soldevilla, J. A. Hildebrand, S. M. Wiggins, and M. A. Roch, "Long-term passive acoustic monitoring of delphinids in the southern california bight," The Journal of the Acoustical Society of America **123**(5_Supplement), 3101–3101 (2008).

[11]N. R. Todd, M. Cronin, C. Luck, A. Bennison, M. Jessopp, and A. S. Kavanagh, "Using passive acoustic monitoring to investigate the occurrence of cetaceans in a protected marine area in northwest ireland," Estuarine, Coastal and Shelf Science **232**, 106509 (2020).

[12]S. M. Van Parijs, C. W. Clark, R. S. Sousa-Lima, S. E. Parks, S. Rankin, D. Risch, and I. C. Van Opzeeland, "Management and research applications of real-time and archival passive acoustic sensors over varying temporal and spatial scales," Marine Ecology Progress Series **395**, 21–36 (2009).

[13]R. S. Sousa-Lima, T. F. Norris, J. N. Oswald, and D. P. Fernandes, "A review and inventory of fixed autonomous recorders for passive acoustic monitoring of marine mammals," Aquatic Mammals **39**(1), 23–53 (2013).

[14]W. W. Au, G. Giorli, J. Chen, A. Copeland, M. Lammers, M. Richlen, S. Jarvis, R. Morrissey, D. Moretti, and H. Klinck, "Nighttime foraging by deep diving echolocating odontocetes off the hawaiian islands of kauai and ni'ihau as determined by passive acoustic monitors," The Journal of the Acoustical Society of America **133**(5), 3119–3127 (2013).

[15]K. Thomisch, O. Boebel, D. P. Zitterbart, F. Samaran, S. Van Parijs, and I. Van Opzeeland, "Effects of subsampling of passive acoustic recordings on acoustic metrics," The Journal of the Acoustical Society of America **138**(1), 267–278 (2015).

[16]M. Michel, M. Torterotot, J.-Y. Royer, and F. Samaran, "Effects of duty cycle on passive acoustic monitoring metrics: The case of blue whale songs," The Journal of the Acoustical Society of America **155**(4), 2538–2548 (2024).

[17]A. Riera, J. K. Ford, and N. Ross Chapman, "Effects of different analysis techniques and recording duty cycles on passive acoustic monitoring of killer whales," The Journal of the Acoustical Society of America **134**(3), 2393–2404 (2013).

[18]Z. R. Rand, J. D. Wood, and J. N. Oswald, "Effects of duty cycles on passive acoustic monitoring of southern resident killer whale (orcinus orca) occurrence and behavior," The Journal of the Acoustical Society of America **151**(3), 1651–1660 (2022).

[19]J. E. Stanistreet, D. P. Nowacek, A. J. Read, S. Baumann-Pickering, H. B. Moors-Murphy, and S. M. Van Parijs, "Effects of duty-cycled passive acoustic recordings on detecting the presence of beaked whales in the northwest atlantic," The Journal of the Acoustical Society of America **140**(1), EL31–EL37 (2016).

[20]J. Fearey, S. H Elwen, G. Distiller, and T. Gridley, "Improving detectability of dolphin signature whistles for capture-recapture analysis: an examination of array configuration using real-world data," Marine Mammal Science **38**(4), 1489–1507 (2022).

[21]J. Ivanchikova and N. Tregenza, "Validation of the f-pod—a fully automated cetacean monitoring system," PLoS One **18**(11), e0293402 (2023).

[22]N. R. E. Todd, A. S. Kavanagh, E. Rogan, and M. J. Jessopp, "What the f-pod? comparing the f-pod and c-pod for monitoring of harbor porpoise (phocoena phocoena)," Ecology and Evolution **13**(6), e10186 (2023).

[23]D. Gillespie, M. Caillat, J. Gordon, and P. White, "Automatic detection and classification of odontocete whistles," The Journal of the Acoustical Society of America **134**(3), 2427–2437 (2013).

[24]G. Dubus, M. Torterotot, J. Béesau, M. Dupont, A. Gros-Martial, M. Michel, E. Morin, P. N. H. Duc, P.-Y. Raumer, O. Adam, *et al.*, "Aplose: A web-based annotation platform for underwater passive acoustic monitoring," SoftwareX **29**, 102055 (2025).

[25]J. A. Hildebrand, K. E. Frasier, T. A. Helble, and M. A. Roch, "Performance metrics for marine mammal signal detection and classification," The Journal of the Acoustical Society of America **151**(1), 414–427 (2022).

[26]Chelonia Ltd., "Chelonia ltd." (2022) https://www.chelonia.co.uk/.

[27]R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2021), https://www.R-project.org/.

[28]W. W. Au, *The sonar of dolphins* (Springer Science & Business Media, 1993).

[29]A. Villadsgaard, M. Wahlberg, and J. Tougaard, "Echolocation signals of wild harbour porpoises, phocoena phocoena," Journal of Experimental Biology **210**(1), 56–64 (2007).

[30]K. T. Clausen, M. Wahlberg, K. Beedholm, S. Deruiter, and P. T. Madsen, "Click communication in harbour porpoises phocoena phocoena," Bioacoustics **20**(1), 1–28 (2011).

[31]W. W. Au and M. C. Hastings, *Principles of marine bioacoustics*, Vol. 510 (Springer, 2008).

1.5 Discussion

Dans un contexte où le PAM des cétacés repose de plus en plus sur des enregistreurs acoustiques fixes à large échelle temporelle et spatiale, ce chapitre avait pour objectif d'évaluer de manière rigoureuse les effets des stratégies de *duty cycle* sur la robustesse des représentations descriptives et de proposer des recommandations concrètes pour orienter les choix méthodologiques futurs. Cette démarche répond à l'un des enjeux centraux de la thèse : contribuer à la standardisation des pratiques du PAM, en tenant compte à la fois des contraintes techniques et des objectifs écologiques.

1.5.1 Aspect méthodologique

Les deux études de cas menées sur des espèces de mysticètes (baleines bleues) et d'odontocètes (delphinidés et marsouins) illustrent de manière complémentaire l'impact du *duty cycle* sur différents types de signaux, d'instruments et de métriques acoustiques. Les résultats montrent de façon convergente que :

— Les régimes de sous-échantillonnage réduisent la précision des indicateurs acoustiques, affectant à des degrés divers les patrons saisonniers, journaliers, les taux de détection et la richesse des types de signaux.
— Les configurations optimales varient selon l'espèce, le type de vocalisation et l'instrument utilisé, soulignant la nécessité d'adapter les choix de *duty cycle* aux spécificités biologiques et techniques du dispositif.
— Des périodes d'écoute courtes et fréquentes (5 à 30 minutes) donnent de meilleurs résultats que des fenêtres longues et espacées, à pourcentage d'écoute constant.
— Les effets négatifs du sous-échantillonnage sont amplifiés dans les contextes de faible activité vocale ou de présence acoustique brève, ce qui impose une vigilance accrue dans ces situations.

Un point important à souligner est que dans les zones et pour les espèces caractérisées par une présence acoustique quasi-continue, il est possible de descendre à des *duty cycle* plus faibles, autour de 25 à 33%, sans compromettre la qualité des indicateurs saisonniers et parfois même circadiens. Ceci offre une marge de manœuvre intéressante pour optimiser

l'autonomie des enregistreurs dans ces contextes spécifiques.

Enfin, il est nécessaire d'attirer l'attention sur la durée des fenêtres d'écoute : des périodes trop courtes, par exemple de l'ordre de 1 minute, bien que tentantes pour maximiser la résolution temporelle, sont généralement techniquement déconseillées. Plusieurs fabricants soulignent que ces très courtes durées peuvent compromettre la stabilité des systèmes, dégrader la qualité des enregistrements et augmenter la consommation énergétique liée aux cycles de redémarrage fréquents. Une durée d'écoute comprise entre 5 et 30 minutes représente un compromis satisfaisant entre résolution temporelle, qualité des données et autonomie des dispositifs.

1.5.2 Aspect écologique

Au-delà des enjeux techniques, ces résultats ont des implications directes sur l'interprétation écologique des données PAM. La réduction de la précision des indicateurs par le *duty cycle* peut limiter la capacité à détecter fidèlement les dynamiques spatio-temporelles des populations de cétacés, en particulier pour des espèces à activité vocale faible ou à présence rare ou irrégulière. Ceci peut affecter la détection de phénomènes clés comme les rythmes saisonniers ou journaliers, ou encore la diversité comportementale traduite par la richesse en types de signaux. Dès lors, la stratégie de *duty cycle* adoptée influence la qualité des conclusions écologiques et la pertinence des recommandations pour la conservation.

Il est aussi essentiel de rappeler que l'adéquation entre le protocole de *duty cycle* et les caractéristiques biologiques des espèces (niveau d'activité acoustique, nature des vocalisations, rythme circadien ou saisonnier) est primordiale pour assurer une bonne représentativité des indicateurs.

1.5.3 Perspectives méthodologiques

Pour poursuivre l'évaluation et l'optimisation des stratégies de collecte de données acoustiques via un *duty cycle*, plusieurs axes de travail pourraient s'avérer particulièrement prometteurs.

Il serait pertinent de mettre en œuvre une manipulation expérimentale directe : déployer

simultanément deux enregistreurs sur un même site, l'un en enregistrement continu et l'autre suivant des configurations de *duty cycle* spécifiques. Cette approche ne se limiterait pas à comparer des scénarios théoriques : elle permettrait d'obtenir la vérité de terrain sur les différences induites par le *duty cycle*. En intégrant les variations naturelles de l'activité vocale et des conditions environnementales, cette méthode offrirait une estimation empirique fiable de l'impact réel des choix de *duty cycle* sur les indicateurs acoustiques.

Cette stratégie pourrait ensuite être appliquée dans des contextes écologiques diversifiés pour tester la robustesse des recommandations : populations nombreuses ou rares, espèces à vocalisations fréquentes ou sporadiques, et différents types de cétacés (odontocètes et mysticètes). Il serait également pertinent de travailler sur des jeux de données contenant plusieurs types de vocalisations simultanément. Dans la réalité, différents comportements vocaux et espèces peuvent cohabiter dans un même environnement acoustique. Étudier ces situations permettrait de tester l'efficacité d'un *duty cycle* unique pour l'ensemble des signaux et d'évaluer son impact sur la détection et la quantification des activités vocales multi-espèces. Une piste serait d'explorer des *duty cycles* aléatoires, dont la distribution des périodes d'écoute pour un même pourcentage pourrait limiter les biais liés à la périodicité et s'adapter aux variations des comportements acoustiques, offrant ainsi une vision plus représentative de l'activité réelle.

Bien que les travaux présentés ici aient été réalisés avec des enregistreurs fixes, il serait intéressant d'étendre ces expérimentations à d'autres types de dispositifs, tels que des bouées dérivantes, des balises ou des plateformes embarquées sur des navires. Cela permettrait d'évaluer la performance des *duty cycles* dans des environnements dynamiques et variés, et d'identifier les compromis optimaux pour des suivis PAM multi-espèces et multi-contextes.

En combinant ces expérimentations directes avec des simulations complémentaires, il serait possible de générer une approche complète et systématique pour affiner les protocoles de *duty cycle*, renforcer leur robustesse, et garantir leur transférabilité dans différents contextes biologiques et environnementaux. Ces travaux permettraient également d'anticiper les situations où certaines stratégies pourraient sous-estimer des phénomènes rares ou épisodiques, et d'identifier des configurations optimales pour maximiser la détection et la fiabilité des indicateurs acoustiques.

Chapitre 2

Modélisation spatio-temporelle de la présence acoustique

2.1 Introduction

Le suivi par acoustique passive (PAM) offre un accès continu à la présence acoustique des cétacés dans leur habitat naturel, permettant de documenter leur distribution spatio-temporelle, leurs mouvements migratoires ou encore leurs comportements vocaux. Cependant, pour comprendre les mécanismes écologiques sous-jacents à ces observations, il est indispensable de relier cette présence à des variables environnementales dynamiques : qu'elles soient physiques (température, courants, bathymétrie) ou biologiques (distribution et abondance des proies).

La modélisation spatio-temporelle constitue un outil pertinent pour identifier les facteurs influençant la probabilité de présence acoustique et les préférences d'habitat des espèces. Elle transforme ainsi les données PAM descriptives en outils d'interprétation, essentiels non seulement pour anticiper les effets des changements environnementaux, mais aussi pour approfondir la compréhension de l'écologie des espèces, de leurs interactions avec les proies et de leurs relations avec l'environnement. Cette démarche propose donc un cadre méthodologique pour intégrer et combiner des sources de données complémentaires (e.g. acoustique passive et active, modèles océanographiques, bases de données satellitaires, mesures *in situ*) afin de mieux caractériser les dynamiques écosystémiques à différentes échelles spatiales et temporelles.

C'est dans ce cadre de recherche que s'inscrit la question directrice de ce chapitre :

Dans quelle mesure la modélisation spatio-temporelle des données PAM permet-elle de comprendre et prédire la distribution acoustique des cétacés en fonction des facteurs environnementaux ?

Pour y répondre, ce chapitre propose une méthodologie visant à optimiser la compréhension et l'exploitation des données acoustiques à travers la modélisation spatio-temporelle. Il s'articule autour de plusieurs sections : (i) un état de l'art présentant la zone d'étude de l'Océan Indien Austral ainsi que les espèces ciblées ; (ii) une description des jeux de données acoustiques utilisés ; (iii) un rappel du cadre théorique de la modélisation spatio-temporelle appliquée au milieu marin, accompagné d'une présentation de la méthodologie ; (iv) la présentation des résultats à travers deux études publiées ou en relecture :

— **Etude 1** : Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern Indian Ocean, publiée en 2025 dans la revue *Scientific Reports* ;

— **Etude 2** : Ecological habitat characterization of baleen whales east of Mayotte : linking seasonal acoustic detections with environmental drivers, actuellement soumise et en cours d'évaluation dans la revue *Global Ecology and Conservation*.

(v) la présentation du stage de Master 2 de Salomé Martin-Marin (International Master of Bioacoustics, MoBi, université de Saint-Étienne, promotion 2023-2024) explorant l'intégration de mesures *in situ* de biomasse de proies via le programme d'acoustique active THEMISTO pour enrichir les variables environnementales et mieux comprendre les interactions prédateur-proie ; (vi) une conclusion générale sur les apports de ces travaux, tant sur le plan des connaissances écologiques acquises que sur le choix des outils méthodologiques.

2.1.1 État de l'art : Les cétacés de l'océan Indien Austral

L'Océan Indien Austral

L'Océan Indien (OI), troisième plus vaste océan mondial, s'étend sur environ 75 millions de kilomètres carrés, ce qui représente près de 21% de la surface terrestre (Fig. 2.1). Contrairement aux océans Atlantique et Pacifique, il ne borde pas la région polaire

arctique. Ses limites géographiques sont définies à l'ouest par les côtes africaines et la péninsule Arabique, au nord par l'Asie, notamment l'Iran et l'Inde, et à l'est par la Birmanie et l'Australie. Par convention, il est séparé de l'océan Atlantique au niveau du cap des Aiguilles (environ 20° de longitude Est) et de l'océan Pacifique à la longitude de l'île de Tasmanie (vers 147° Est). Sa limite méridionale est fixée par l'océan Austral à partir du 60e parallèle sud. Toutefois, d'un point de vue océanographique, cette frontière est souvent placée plus au nord, autour de 45°S, correspondant à la zone frontale antarctique qui marque la transition entre les eaux froides et denses de l'Antarctique et les eaux plus tempérées subantarctiques (Llano 1978).



Figure 2.1 – Cartes des imites de l'Océan Indien (Wikipédia 2025).

La dynamique océanique de l'OI est complexe et stratifiée, particulièrement dans sa partie australe. La région sud-ouest de l'OI est caractérisée par la convergence de plusieurs courants majeurs et fronts océaniques, notamment les fronts subtropicaux, subantarctiques et polaires, ainsi que le courant circumpolaire antarctique (ACC) (Fig. 2.2). Ces courants forment une mosaïque dynamique de masses d'eau aux propriétés physiques contrastées, générant des gradients importants de température, salinité et nutriments. Ces variations jouent un rôle déterminant dans la structuration des écosystèmes marins et la productivité biologique régionale (Orsi, Whitworth III et Nowlin Jr 1995). En particulier, le ACC, courant froid et puissant circulant d'ouest en est autour de l'Antarctique, constitue

un véritable cloisonnement entre les eaux antarctiques, froides et pauvres en sels, et les eaux subtropicales plus chaudes et salées. La zone frontale subantarctique, située approximativement entre 45°S et 50°S, est une interface où le brassage vertical induit par les vents dominants d'ouest favorise une productivité primaire élevée (Allison, Stramski et Mitchell 2010 ; Constable et al. 2014).

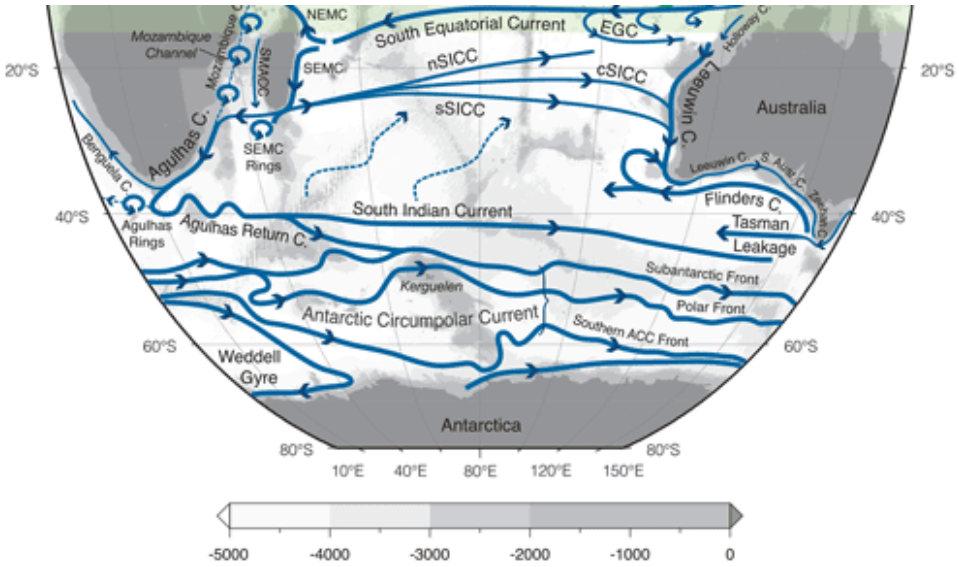


Figure 2.2 – Principaux courant de l'OI Austral. Les flèches indiquent les flux moyens tout au long de l'année. Les acronymes sont : SEMC, courant du sud-est de Madagascar ; SMACC, courant côtier du sud ouest de Madagascar ; SICC, contre-courant du sud de l'Inde (nord, central et sud) (Phillips et al. 2021).

Les saisons australes

Dans l'hémisphère sud, les saisons sont définies à l'inverse de l'hémisphère nord, c'est-à-dire :

- **Été austral** : décembre à février
- **Automne austral** : mars à mai
- **Hiver austral** : juin à août
- **Printemps austral** : septembre à novembre

Les variations saisonnières influencent fortement la dynamique physique et biologique de l'OI austral. En hiver austral, l'extension maximale de la banquise provoque un refroidissement intense de la surface et une augmentation de la salinité, ce qui rend l'eau plus dense et favorise un mélange vertical profond de la colonne d'eau. Ce processus entraîne la remontée de nutriments depuis les couches profondes vers la surface, préparant

la base de la production primaire. Au printemps et en été australs, la fonte de la glace et le réchauffement de la surface renforcent la stratification de la couche superficielle, limitant l'apport de nouveaux nutriments (Rintoul 2007). Néanmoins, les nutriments accumulés pendant l'hiver restent disponibles et, combinés à l'ensoleillement maximal, permettent une croissance importante du phytoplancton. Ces conditions provoquent des variations saisonnières marquées de la biomasse phytoplanctonique, avec des pics de chlorophylle a (Chla) généralement observés en fin d'été austral, notamment au niveau de la zone frontale subantarctique (Moore et Abbott 2000).

Dans l'OI, plusieurs zones se distinguent par leurs caractéristiques océanographiques et écologiques particulières. Parmi celles-ci, deux sites sont au cœur de ce chapitre : le plateau de Kerguelen, au niveau de l'OI subantarctique, et la région tropicale autour de Mayotte et au nord du Canal du Mozambique. Le plateau de Kerguelen, situé autour de 49°S et 70°E, se distingue par une topographie sous-marine complexe qui interagit avec la dérive circumpolaire antarctique pour générer un brassage vertical intense (Civel-Mazens et al. 2021) (Fig. 2.3a) . Ce mécanisme favorise la remontée d'eaux profondes riches en nutriments (*upwelling*), alimentant une productivité biologique importante à ces latitudes. Les concentrations de Chla mesurées par satellite (Fig. 2.3b) témoignent de cette haute productivité, qui fait de cette zone un *hotspot* écologique majeur, abritant une biodiversité marine diversifiée, incluant de nombreuses espèces pélagiques et migratrices ainsi que plusieurs prédateurs marins supérieurs (Bestley et al. 2020 ; Todd et Williamson 2022 ; AméziANe et al. 2011).

Plus au nord, la région tropicale autour de Mayotte (environ 12,75°S, 45,17°E), à l'entrée nord du canal du Mozambique, présente une dynamique océanographique fortement influencée par des variations saisonnières et méso-échelle (Fig. 2.4). Cette variabilité concerne la température de surface, les vents et l'activité tourbillonnaire, liée notamment à la propagation de vagues de Rossby et à la dynamique du courant Nord-Est de Madagascar (Schouten et al. 2003 ; Mawren et al. 2023). Toutefois, la topographie locale confère à Mayotte un abri relatif, atténuant les effets du refroidissement saisonnier et maintenant des températures de surface relativement élevées, même en hiver austral. Depuis 2018, l'activité hydrothermale du volcan sous-marin Fani Maoré constitue également une source ponctuelle de nutriments, renforçant la productivité biologique dans ce secteur (Feuillet et al. 2021 ; Lemoine et al. 2020 ; Saurel et al. 2021).

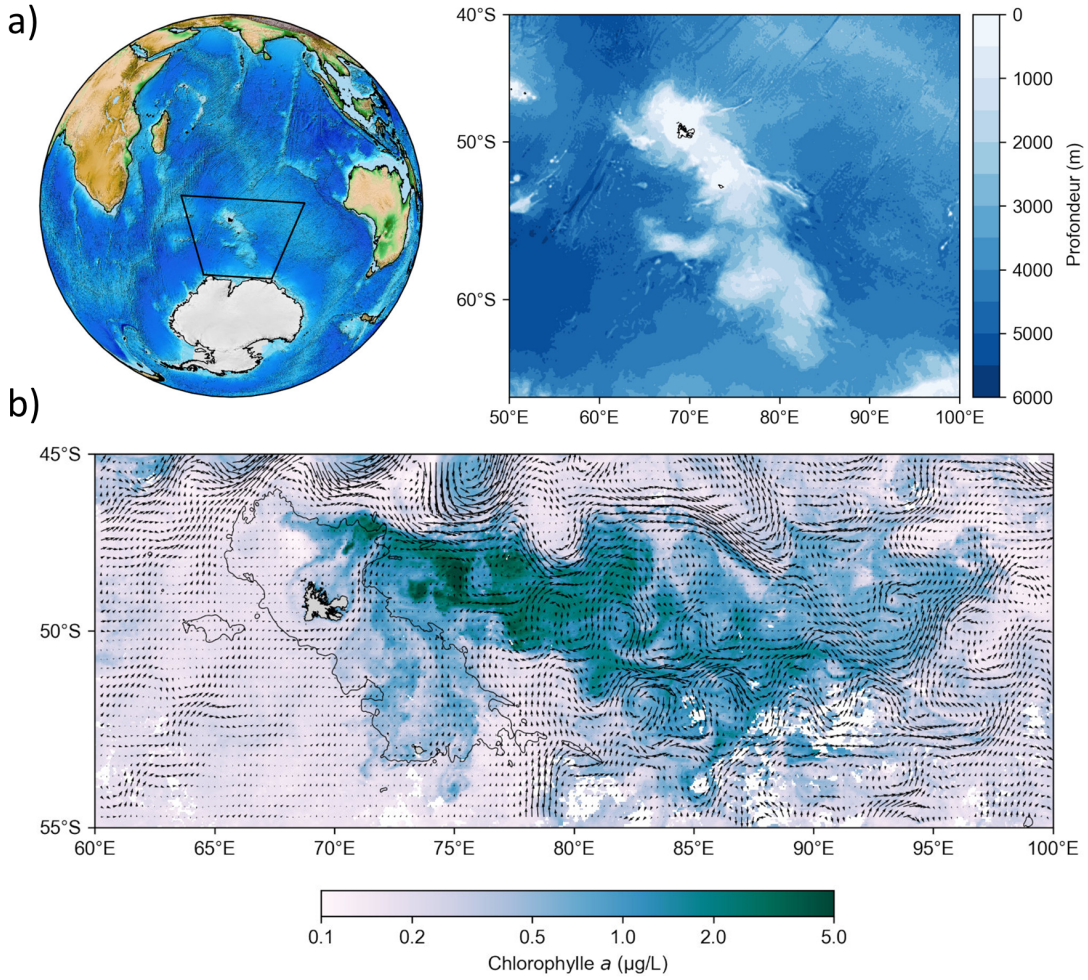


FIGURE 2.3 – a) Topographie en globe et carte bathymétrique centrées sur le plateau de Kerguelen et son île. b) Concentration en chlorophylle a (Satelitte MODIS) et courant géostrophique dérivés de la topographie dynamique (produit AVISO issu de plusieurs satellites altimétriques). Données moyennées sur le mois de décembre 2014, les pixels blancs correspondent à une absence de données à cause des nuages. L'isobathe 1000 m est représentée par un trait noir (REMBAUVILLE et DEQUINCEY 2024).

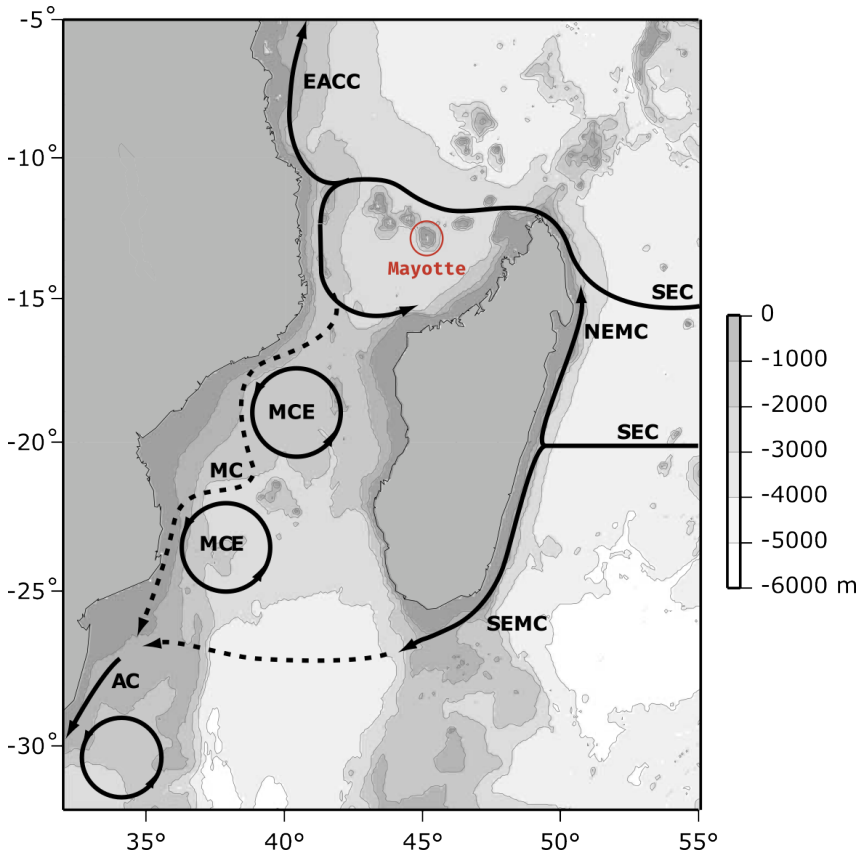


FIGURE 2.4 – La bathymétrie de la région autour de Madagascar, de Mayotte et du canal du Mozambique, ainsi que les principaux courants et caractéristiques de circulation, sont représentées schématiquement. Le cercle rouge entoure l'île de Mayotte. Les éléments illustrés comprennent le Courant Équatorial Sud (SEC), les Courants Madagascar Nord-Est et Sud-Est (NEMC et SEMC), le Courant Côtier Est-Africain (EACC), le Courant des Aghulas (AC), les tourbillons du Canal du Mozambique (MCE) et le Courant du Mozambique (MC). Ce dernier est représenté par une ligne pointillée, car son existence et sa nature demeurent incertaines. Il en va de même pour la connexion entre le SEMC et l'AC, qui pourrait être formée par des tourbillons générés près de la pointe sud de Madagascar. Figure adaptée et refaite de SCHOUTEN et al. 2003.

Ces deux zones, bien que très contrastées par leur latitude et leur dynamique océanographique, illustrent la diversité des habitats présents dans l'OI. Elles présentent cependant une vulnérabilité commune aux variations climatiques interannuelles, notamment le Dipôle de l'océan Indien (IOD) et, de manière indirecte, El Niño–Southern Oscillation (ENSO). L' IOD est un phénomène océan-atmosphère propre à l'OI, caractérisé par des anomalies de température de surface opposées entre les parties occidentale et orientale du bassin, modulant les régimes de mousson, la circulation océanique régionale, l'*upwelling* côtier et la productivité biologique (SAJI et al. 1999). L'ENSO, bien qu'étant un phénomène

du Pacifique tropical alternant entre une phase chaude (El Niño) et une phase froide (La Niña), influence indirectement l'OI via des téléconnexions atmosphériques. Il joue un rôle clé dans le déclenchement des événements IOD, avec un réchauffement de l'OI occidental et un refroidissement de l'OI sud-oriental (Zhang et Li 2023). Ces phénomènes conjugués influencent directement la température, la circulation océanique, la distribution des nutriments et la productivité biologique à l'échelle du bassin (Schott, Xie et McCreary Jr 2009 ; Iwakiri et Watanabe 2019).

Enfin, dans un contexte de préservation, l'OI bénéficie d'un cadre de protection international avec la création en 1979 du Sanctuaire des Baleines de l'OI par la Commission Baleinière Internationale (IWC) [1]. Ce sanctuaire interdit la chasse commerciale à l'échelle du bassin austral et joue un rôle majeur dans la conservation des grandes espèces de cétacés, notamment la baleine bleue (*Balaenoptera musculus*) (IWC 1980). À une échelle plus régionale, des mesures de protection viennent renforcer ce dispositif. Autour de Mayotte, le Parc naturel marin, créé en 2010, couvre l'ensemble des eaux territoriales de l'île et vise à concilier préservation de la biodiversité, activités traditionnelles et usages modernes (*Parc naturel marin de Mayotte* 2010). Plus au sud, la zone des Kerguelen est intégrée depuis 2016 à la Réserve naturelle nationale des Terres australes françaises, l'une des plus vastes aires marines protégées au monde, contribuant à la conservation des écosystèmes subantarctiques et de nombreuses espèces emblématiques, dont plusieurs cétacés (*Réserve naturelle nationale des Terres australes françaises* 2006).

La coexistence de ces différents niveaux de protection, international, régional et local, illustre la volonté de préserver la biodiversité dans un océan soumis à de fortes pressions environnementales et anthropiques. Néanmoins, la gestion de ces vastes espaces reste un défi majeur en raison de leur étendue géographique, de la diversité des écosystèmes concernés et des enjeux socio-économiques associés. Cela souligne la nécessité d'approches intégrées combinant données océanographiques, observations biologiques et suivis acoustiques pour mieux comprendre l'utilisation de l'habitat par les espèces et orienter les stratégies de conservation.

1. Organisation intergouvernementale créée en 1946 pour réglementer la chasse à la baleine et promouvoir la conservation des cétacés.

État des connaissances sur les baleines bleues

La majorité des connaissances sur les grandes baleines dans l'océan Indien (OI) repose sur les suivis par acoustique passive, qui constituent aujourd'hui la principale source d'information dans cette région. Ces données viennent compléter les archives historiques issues des carnets de bord de la chasse à la baleine [2], fournissant des informations précieuses sur la distribution, la présence et les déplacements de ces espèces (Branch, Stafford et al. 2007 ; Branch, Monnahan et al. 2025 ; Torterotot, Samaran, Stafford et al. 2020 ; Samaran, Adam et Guinet 2010b ; Leroy, Samaran, Stafford et al. 2018 ; Dréo, Bouffaut et al. 2019 ; Stafford, Bohnenstiehl et al. 2004 ; Stafford, Chapp et al. 2011).

Dans l'OI, les observations visuelles directes sont particulièrement limitées. Les grandes baleines fréquentent majoritairement les zones du large, loin des côtes, et aucune flotte scientifique n'est dédiée à la réalisation de suivis systématiques. Les campagnes ponctuelles existantes sont extrêmement coûteuses et chronophages, rendant ce type d'observation difficilement applicable à l'échelle régionale (Andrews-Goff, Smith et al. 2024 ; Sutton, Jenner et Jenner 2019 ; Laran et al. 2017).

Face à ces contraintes, le suivi satellitaire s'impose comme un outil prometteur pour documenter les migrations à grande échelle et suivre les déplacements individuels sur de longues distances. Cependant, il reste encore très limité pour les baleines bleues de l'OI, en raison de contraintes logistiques et biologiques importantes. Le projet Blue Corridors (Johnson et al. 2022 ; Palacios et al. 2025) illustre ce déséquilibre : seules quelques baleines bleues ont été équipées au large de l'Australie, alors que de nombreux individus de baleines à bosse ont déjà fait l'objet de suivis satellitaires. Or, les baleines bleues étant des migratrices de longue distance, ce type de suivi à grande échelle serait particulièrement pertinent pour comprendre leurs déplacements et leurs corridors migratoires.

Malgré ces développements récents, la rareté des suivis satellitaires et l'impraticabilité des observations visuelles directes ne compensent pas la richesse et la continuité des données fournies par le PAM, qui demeure un outil central et incontournable pour l'étude des cétacés dans l'OI.

2. La chasse commerciale à la baleine a été pratiquée intensivement dans l'océan Indien entre le XIXe et le XXe siècle, notamment par la Norvège, le Japon, la France et le Royaume-Uni.

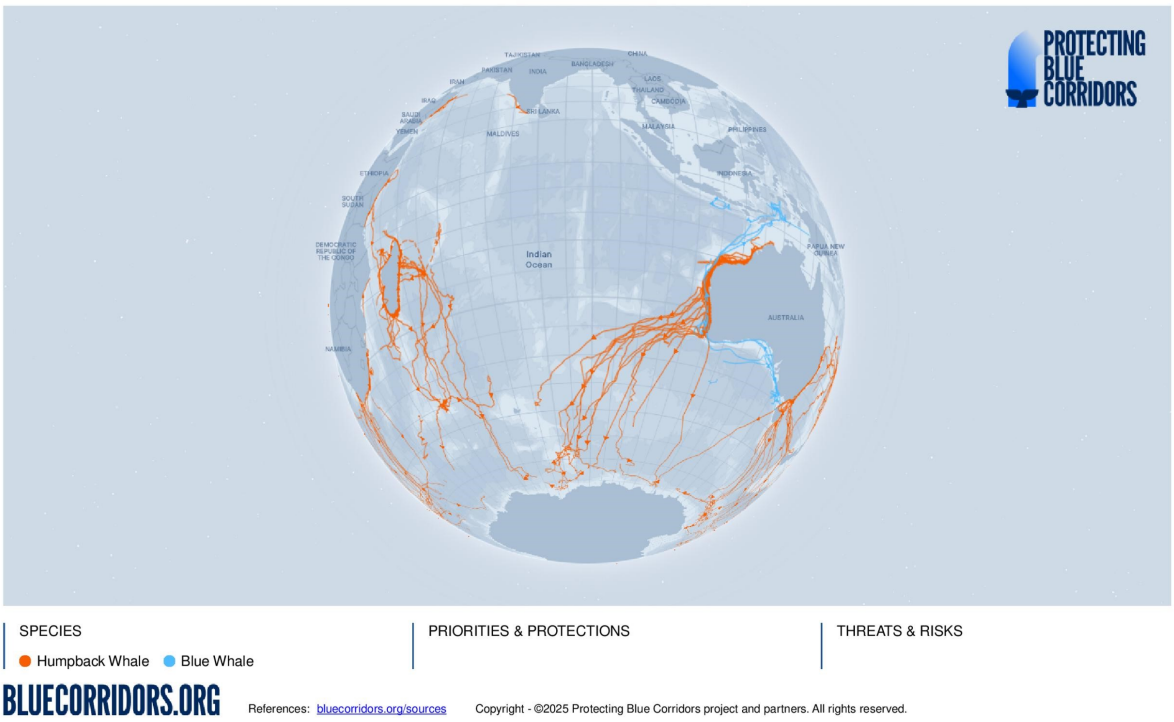


FIGURE 2.5 – Trajectoires des baleines à bosse (en orange) et des baleines bleues (en bleu) équipées de balises satellitaires dans l'OI, d'après les données du projet Blue Corridors (PALACIOS et al. 2025).

Parmi les espèces de mysticètes fréquentant l'OI austral, la baleine bleue (*Balaenoptera musculus*) occupe une place centrale, tant par sa taille exceptionnelle que par les enjeux de conservation qu'elle soulève. Deux sous-espèces sont actuellement reconnues dans cette région par le Comité de Taxonomie de la Society for Marine Mammalogy (2020) (COMMITTEE ON TAXONOMY, SOCIETY FOR MARINE MAMMALOGY 2020) : la baleine bleue antarctique (*Balaenoptera musculus intermedia*), largement répandue dans les hautes latitudes australes, et la baleine bleue pygmée (*Balaenoptera musculus brevicauda*), plus petite, présente dans les zones tempérées et tropicales de l'OI (Fig.2.6). Ces deux sous-espèces se distinguent notamment par leur morphologie, leur distribution et leur répertoire vocal (ŠIROVIĆ et OLESON 2022).

Grâce au PAM, plusieurs populations acoustiquement distinctes de baleines bleues pygmées ont été identifiées dans l'OI (Fig.2.6). Elles ne présentent pas de différences morphologiques (LEDUC et al. 2007) mais se distinguent par leurs chants caractéristiques, stéréotypés et répétitifs, ainsi que par leur distribution géographique et leur phénologie

migratoire (Samaran, Adam et Guinet 2010b ; Branch, Stafford et al. 2007 ; Leroy, Royer, Alling et al. 2021 ; Torterotot, Samaran, Stafford et al. 2020). Quatre populations acoustiques sont actuellement décrites dans cette région (Širović et Oleson 2022) :

— **La population du sud-est de l'océan Indien (SEIO PBW)**, présente entre l'Indonésie, l'Australie occidentale et le plateau de Kerguelen (McCauley, Jenner et al. 2001 ; Gavrilov et McCauley 2013) ;

— **La population du sud-ouest de l'océan Indien (SWIO PBW)**, présente entre le sud de Madagascar et les îles subantarctiques telles qu'Amsterdam, Marion et Prince Edward (Samaran, Stafford et al. 2013 ; Gavrilov et McCauley 2013 ; Torterotot, Samaran, Stafford et al. 2020 ; Shabangu, Munoz et al. 2024) ;

— **La population de Sri Lanka (SL PBW)**, enregistrée dans le nord-est de l'OI, autour du Sri Lanka et du sud de l'Inde (Miksis-Olds, Nieukirk et Harris 2018) ;

— **La population du nord-ouest de l'océan Indien (NWIO PBW)**, récemment suggérée à partir des caractéristiques des vocalisations enregistrées entre le canal du Mozambique, les Comores, Mayotte et jusqu'au golfe d'Aden, bien que cette entité reste encore peu caractérisée (Leroy, Royer, Alling et al. 2021 ; Cerchio, Willson et al. 2020).

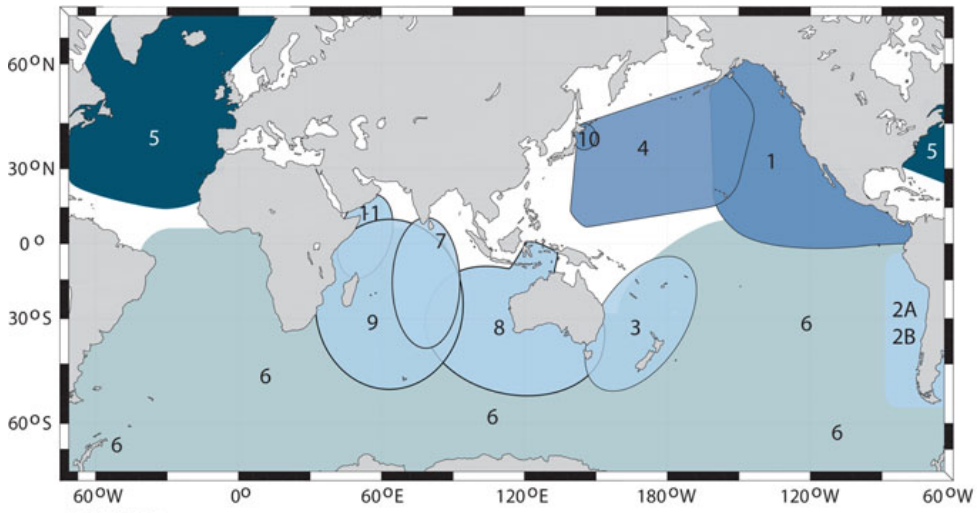


Figure 2.6 – Aires de répartition connues des populations acoustiques de baleines bleues dans le monde, les chiffres de 6 à 11 (hors 10) concernent l'OI : (6) baleine bleue antarctique (ANT BW), (7) baleine bleue pygmée du Sri Lanka (SL PBW), (8) baleine bleue pygmée du sud-est de l'océan Indien (SEIO PBW), (9) baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW), (11) population du nord-ouest de l'océan Indien (NWIO PBW). Les autres numéros représentent des populations situées dans d'autres océans (Širović et Oleson 2022).

Les baleines bleues de l'OI présentent une forte spécialisation alimentaire, se nourrissant presque exclusivement de krill (famille des *Euphausiidae*), dont l'espèce varie selon la région (Kawamura 1980 ; Kawamura 1994 ; Gill et al. 2002). Elles ciblent principalement les agrégations denses, souvent composées d'une seule espèce (Fig.2.7) (Miller et al. 2019). Ce choix influence directement leur distribution et leurs migrations, ainsi que la formation de petits groupes ou de rassemblements temporaires sur les zones d'alimentation. Sur ces zones, les baleines bleues sont généralement observées seules ou en petits groupes (Calambokidis et al. 2007), mais peuvent parfois former des couples mâle-femelle semi-stables durant l'alimentation estivale (Schall et al. 2019 ; Sears et Larsen 2002). Le comportement reproducteur reste mal connu, car ces associations n'ont pas été suivies au-delà des zones d'alimentation ; cependant, leur nature suggère un lien possible avec la reproduction (Širović et Oleson 2022). Les baleines pratiquent souvent le *lunge-feeding* ,[3] permettant une alimentation efficace en ciblant les agrégations les plus denses (Oleson et al. 2007 ; Gill et al. 2002 ; Calambokidis et al. 2007).



Figure 2.7 – Capture d'écran d'une vidéo de National Geographic montrant une baleine bleue se nourrissant de krill au large de la Nouvelle-Zélande dans l'océan Austral (Gibbens 2020).

3. Technique consistant à foncer sur une dense concentration de krill, à ouvrir largement la bouche pour engloutir un grand volume d'eau et de proies, puis à filtrer le krill grâce aux fanons,

Ce chapitre de thèse se concentre plus particulièrement sur trois populations acoustiquement différenciées : les baleine bleue Antarctique (ANT BW), les baleine bleue pygmée du sud-est de l'océan Indien (SEIO PBW) et les baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW).

La baleine bleue Antarctique - ANT BW Les baleines bleues Antarctique (ANT BWs) sont présentes dans tous les océans de l'hémisphère sud (Fig. 2.6). Elles ne traversent généralement pas l'équateur et se déplacent principalement autour de l'Antarctique (Attard et al. 2024). Depuis une vingtaine d'années, elles font l'objet d'un suivi PAM important grâce à des réseaux d'hydrophones déployés dans l'ensemble de l'OI, depuis les eaux subtropicales jusqu'aux régions antarctiques (Torterotot, Samaran, Stafford et al. 2020 ; Leroy, Samaran, Bonnel et al. 2016 ; Samaran, Stafford et al. 2013). Le chant typique de la ANT BW dure environ 26 secondes et se compose de trois unités tonales qui s'enchaînent sans silence (Ljungblad, Clark et Shimada 1998). La première unité, appelée unité A, est un son de fréquence constante autour de 28 Hz durant 8 à 12 secondes. Elle est suivie d'une diminution de fréquence de 28 à 20 Hz pendant 1 à 2 secondes, puis de l'unité B, légèrement modulée entre 20 et 18 Hz. Ces vocalisations ont une forme caractéristique en 'Z' sur le plan temps-fréquence, d'où leur appellation de vocalisations en 'Z' ou Z-calls (Fig. 1.5c). L'analyse de ce chant caractéristique de cette sous-espèce a permis de mettre en évidence un schéma de migration saisonnier entre les hautes et les basses latitudes (Leroy, Samaran, Bonnel et al. 2016).

Durant l'été austral, les ANT BW se concentrent dans les eaux froides de l'Antarctique, où elles s'alimentent principalement de krill. Cette présence est confirmée par un nombre conséquent de détections acoustiques dans la zone, suggérant qu'elle constitue une aire d'alimentation privilégiée (Stafford, Bohnenstiehl et al. 2004 ; Samaran, Adam et Guinet 2010b ; Leroy, Samaran, Bonnel et al. 2016). Lors de leur migration vers les basses latitudes en automne, les ANT BW empruntent des corridors migratoires jusqu'à Madagascar et même plus au nord. En hiver austral, des enregistrements acoustiques réalisés à plus basse latitude, notamment à l'est de Madagascar et même au nord du canal du Mozambique, révèlent leur présence dans des zones d'hivernage pouvant potentiellement servir de zone de reproduction ou de repos (Torterotot, Samaran, Stafford et al. 2020 ; Leroy, Samaran, Bonnel et al. 2016). Cependant, cet aspect de leur cycle de vie est encore peu documenté et mal connu. Au printemps, lors de leur retour vers les hautes

latitudes, un grand nombre de détections sont enregistrées autour des îles subantarctiques, telles que le plateau de Kerguelen et les îles Crozet et Amsterdam. Bien que leur rôle précis dans le cycle de vie des baleines reste encore incertain, ces zones apparaissent comme des habitats d'importance écologique majeure à cette période (Samaran, Adam et Guinet 2010b ; Samaran, Stafford et al. 2013 ; Stafford, Bohnenstiehl et al. 2004 ; Shabangu, Munoz et al. 2024 ; Leroy, Samaran, Bonnel et al. 2016).

Enfin, les données acoustiques révèlent que les schémas migratoires des ANT BW de l'OI pourraient être plus complexes que le modèle saisonnier classique, qui décrit un cycle estival de nourrissage dans les eaux productives de haute latitude et un hivernage dans les eaux chaudes de basse latitude (Mackintosh 1966 ; Kellogg 1929). En effet, des chants sont détectés de manière quasi permanente à des latitudes intermédiaires, notamment autour du plateau de Kerguelen, y compris durant l'été austral, alors que les individus devraient se nourrir exclusivement dans les eaux antarctiques (Samaran, Adam et Guinet 2010b ; Torterotot, Samaran, Stafford et al. 2020 ; Leroy, Samaran, Bonnel et al. 2016). De même, des enregistrements en hiver austral dans des zones de haute latitude, telles que la mer de Weddell, suggèrent la présence de certains individus dans des zones supposément hivernales subtropicales (Širović, Hildebrand, Wiggins et al. 2004 ; Thomisch, Boebel, Clark et al. 2016).

Ces observations indiquent l'existence de différentes stratégies migratoires chez les ANT BWs. Certains individus semblent effectuer une migration partielle, en restant parfois dans des zones de nourrissage ou d'hivernage au lieu de compléter leur migration annuelle, tandis que d'autres suivent une migration différenciée, avec des variations temporelles selon l'âge, le sexe, l'état de santé ou le statut reproductif (Geijer, Notarbartolo di Sciara et Panigada 2016 ; Branch, Stafford et al. 2007 ; Letsheleha et al. 2022 ; Shabangu, Findlay et al. 2019). Des analyses historiques viennent appuyer cette variabilité : l'exploitation des données de marquage-récupération du programme *Discovery*[4] (1926–1963) a mis en évidence des probabilités élevées de mouvement interannuel entre les trois grands bassins de l'océan Austral (Fig. 2.8). Par exemple, environ 15% des individus quittant l'Atlantique rejoignaient l'océan Indien et près de 30% migraient vers le Pacifique d'une année sur l'autre. Ces résultats suggèrent une faible fidélité aux aires de nourrissage

4. Expéditions océanographiques britanniques qui consistaient à tirer dans le dos des baleines des flèches métalliques numérotées ; en cas de capture ultérieure, le marquage permettait de retracer leurs déplacements. Ce fut l'un des premiers dispositifs à grande échelle pour étudier la migration et la structure des populations de grands cétacés.

propres à chaque bassin et fournissent peu d'indices en faveur d'une structuration de population liée à ces zones (Rand, Branch et Jackson 2024). Ainsi, les ANT BWs semblent disposer d'un répertoire de stratégies migratoires allant de la migration complète à une certaine forme de résidence, reflétant une flexibilité adaptative face aux conditions environnementales et aux contraintes individuelles (Chapman et al. 2011 ; Corkeron et Connor 1999 ; Clapham 2001).

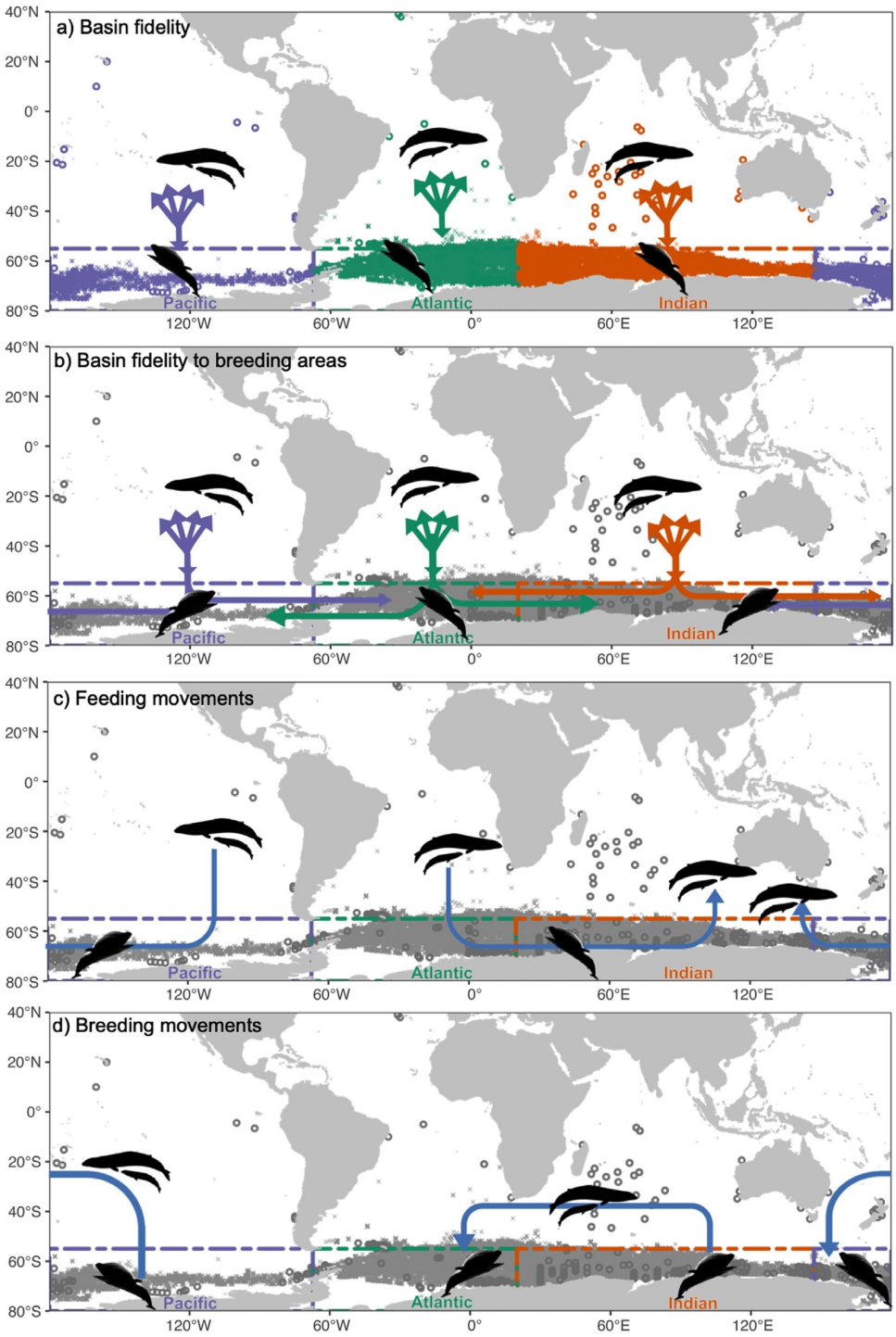


FIGURE 2.8 – Scénarios migratoires possibles des baleines bleues antarctiques dans l'océan Austral : (a) routes migratoires fidèles à un seul bassin, (b) déplacements inter-bassins sur les zones de nourrissage mais fidélité aux aires de reproduction, (c) mouvements inter-bassins aux hautes latitudes avant migration vers le nord, (d) déplacements inter-bassins aux basses latitudes mais peu de mouvement aux hautes latitudes. Les cercles indiquent des détections acoustiques et les croix des captures de baleines (RAND, BRANCH et JACKSON 2024).

La baleine bleue pygmée d'Australie - SEIO PBW Les baleines bleues pygmées du sud-est de l'océan Indien (SEIO PBWs), aussi qualifiées d'« Australie », présentent une distribution centrée sur les latitudes tropicales et tempérées de l'est de l'OI (Fig. 2.6). Cette population émet un chant composé de trois unités distinctes (McCauley, Jenner et al. 2001). La première est un son tonal à 19 Hz qui évolue à 21 Hz après 21 secondes et dure encore 22 secondes. La deuxième unité, apparaissant 5 à 10 secondes plus tard, est modulée en fréquence, débutant à 20 Hz et montant jusqu'à 26 Hz en 23 secondes. Enfin, la dernière unité arrive 23 secondes plus tard, consistant en une note tonale à 18-19 Hz d'une durée de 26 à 48 secondes. Toutes ces unités présentent également des harmoniques à plus haute fréquence (Fig. 1.5a).

La présence des SEIO PBW est principalement documentée grâce aux enregistrements acoustiques et aux suivis satellitaires, qui révèlent un schéma migratoire saisonnier entre les eaux tropicales de l'Indonésie (mers de Banda, Timor et Savu) et les eaux tempérées du sud-ouest de l'Australie, s'étendant parfois jusqu'aux îles subantarctiques comme le plateau de Kerguelen. Leurs vocalisations sont détectées tout au long de l'année le long de ce corridor, confirmant la continuité de leurs déplacements sur de larges échelles spatiales (Double et al. 2014 ; McCauley, Gavrilov et al. 2018 ; Gavrilov et McCauley 2013 ; Mustika et al. 2024). Les zones tropicales de l'archipel indonésien sont considérées comme des habitats hivernaux potentiels. Des vidéos filmées en 2016 et 2018 dans la mer de Banda suggèrent qu'elles pourraient également servir de zone de mise bas pour cette population (Fig. 2.9) (Pindito 2016 ; Pindito 2018).



FIGURE 2.9 – Capture d'écran d'une vidéo montrant une baleine bleue avec son petit dans la mer de Banda en 2016 (PINDITO 2016).

Des campagnes de marquage satellitaire menées entre 2009 et 2021 ont permis de suivre 37 individus équipés de balises satellitaires implantables [5] au large de l'Australie occidentale ou méridionale. Tous ont montré une migration nord-sud cohérente, reliant les eaux du sud-ouest de l'Australie aux eaux tropicales de l'archipel indonésien, notamment la mer de Banda (DOUBLE et al. 2014 ; MÖLLER et al. 2020 ; OWEN et al. 2016 ; THUMS et al. 2022). Deux nouvelles baleines taguées en 2024 ont permis de compléter les données en documentant un retour complet vers l'Australie pour un individu, ainsi qu'un déplacement vers les eaux subantarctiques pour un autre, suggérant une possible extension de l'aire de répartition estivale jusqu'au plateau de Kerguelen (Fig. 2.10) (MUSTIKA et al. 2024). Cette observation vient renforcer l'hypothèse déjà avancée par des enregistrements acoustiques plus anciens, qui détectaient une activité notable dans cette région pendant l'été et l'automne, période correspondant à la phase d'alimentation (SAMARAN, STAFFORD et al. 2013 ; TORTEROTOT, SAMARAN, STAFFORD et al. 2020).

5. Dispositifs composés d'un émetteur Spot 5 (Wildlife Computers) fixé à l'animal via un système de projectile, implanté dans la peau, le blubber et le muscle superficiel, permettant un suivi satellitaire détaillé des déplacements et comportements.

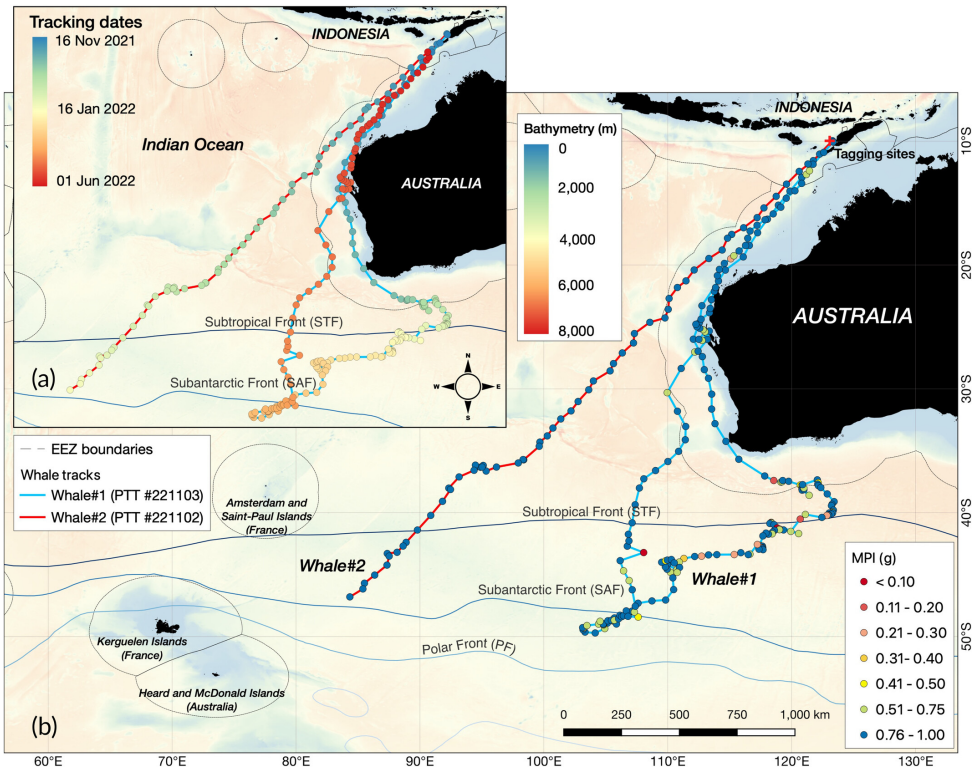


FIGURE 2.10 – Trajectoires de déplacement de deux baleines bleues pygmées d'Australie équipées de balises satellites. (a) Trajectoires des baleines selon les dates ; les points colorés représentent les dates de suivi ; (b) trajectoires des baleines selon l'indice de persistance du mouvement (MUSTIKA et al. 2024).

Bien que les données révèlent une forte saisonnalité des déplacements, la variabilité interannuelle et régionale observée dans les détections acoustiques suggère également une diversité de stratégies migratoires au sein de cette population. Le rôle précis de certaines zones comme celle du plateau de Kerguelen reste encore à clarifier, mais leur fréquentation saisonnière laisse supposer une importance écologique pour l'alimentation. Enfin, l'absence de détection au sud de 42°S confirme une préférence marquée pour des eaux plus chaudes que celles fréquentées par les baleines bleues antarctiques (DOUBLE et al. 2014).

La baleine bleue pygmée de Madagascar - SWIO PBW Les baleines bleues pygmées du sud-ouest de l'océan Indien (SWIO PBWs), également appelées baleines bleues pygmées de « Madagascar », émettent un chant comprenant deux unités (LJUNGBLAD, CLARK et SHIMADA 1998). La première est un son tonal autour de 38 Hz durant 10-12 secondes, suivi d'environ 25 secondes plus tard d'un son modulé en fréquence allant de 28

à 15 Hz, qui devient tonal sur les 15-20 secondes finales. Ces unités sont accompagnées d'harmoniques (Fig. 1.5b). Cette population acoustique est présente dans la partie sud-ouest de l'OI, mais ses habitats préférentiels restent mal connus et son schéma migratoire demeure encore flou (Fig. 2.11).

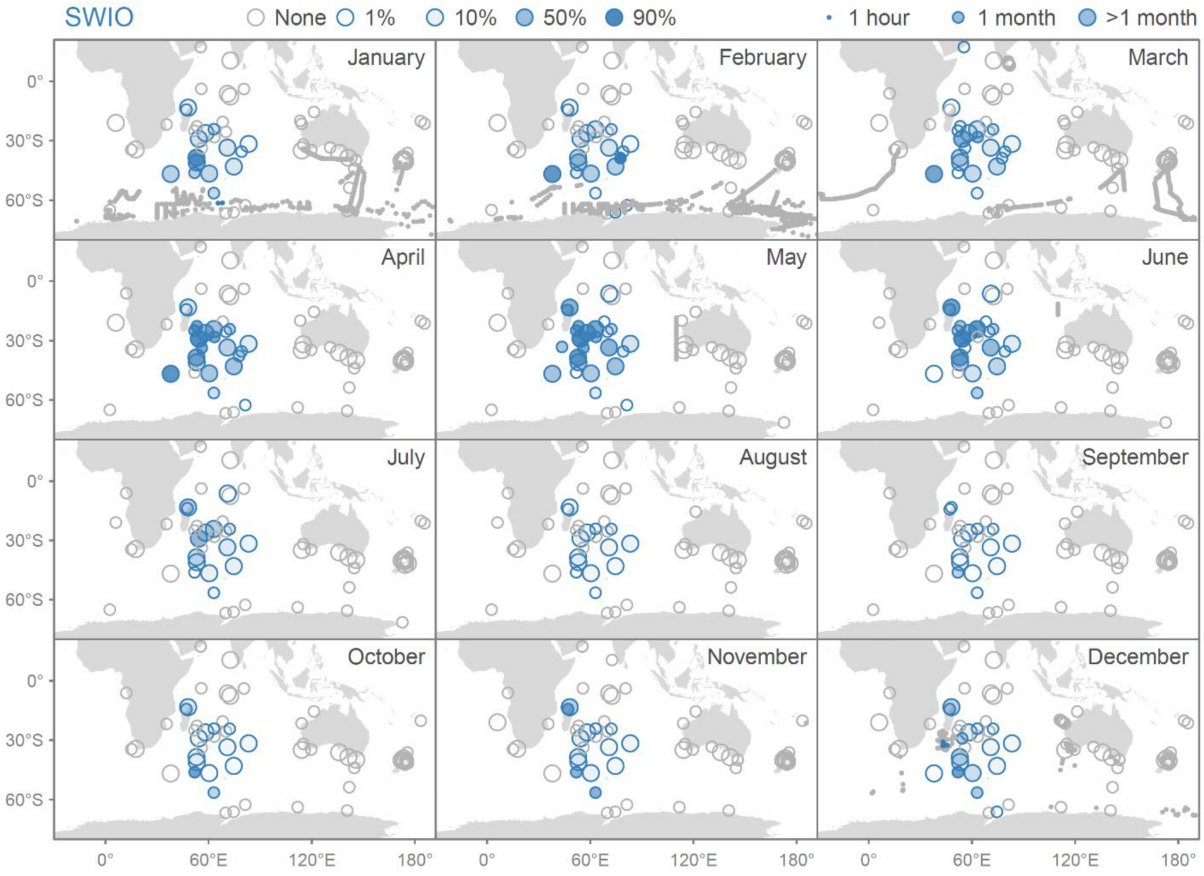


FIGURE 2.11 – Données mensuelles de chants disponibles pour les SWIO PBW dans l'océan Indien à chaque site. La couleur des cercles indique la proportion d'heures durant lesquelles des chants ont été détectés (gris clair : absence de chants). La taille des cercles reflète la quantité de données à chaque site : bouée acoustique (« 1 heure »), enregistreur fixe (« 1 mois ») ou enregistreur fixe sur plusieurs années (« >1 mois ») (BRANCH, MONNAHAN et al. 2025).

Néanmoins, les données acoustiques récentes permettent d'étoffer les connaissances sur les grandes tendances de sa présence saisonnière. Ses vocalisations sont détectées de manière saisonnière entre décembre et juillet dans les eaux entourant les îles subantarctiques, notamment Marion, Prince Edward, Amsterdam et Kerguelen, qui semblent jouer un rôle important comme zones d'alimentation (TORTEROTOT, SAMARAN, STAFFORD et al.

2020 ; Shabangu, Munoz et al. 2024). À partir d'avril, les individus amorcent une migration vers le nord en empruntant possiblement deux corridors : l'un à travers le canal du Mozambique, et l'autre le long de la côte est de Madagascar, incluant les régions de La Réunion et de l'île Maurice (Cerchio, Rasoloarijao et al. 2024 ; Torterotot, Samaran, Stafford et al. 2020). Des détections récentes dans l'est de Mayotte entre mai et juillet confirment la poursuite de ce mouvement migratoire vers des latitudes tropicales plus au nord (Dréo, Crawford et al. 2025).

En revanche, une absence quasi totale de détections est observée entre hiver austral (août à octobre) dans l'ensemble des stations d'écoute du sud-ouest de l'OI, y compris dans des zones théoriquement propices telles que Diego Garcia ou les Seychelles (Stafford, Boussarie et al. 2023 ; Leroy, Royer, Alling et al. 2021). Cette disparition saisonnière pourrait résulter soit d'une migration vers des habitats encore non couverts par les dispositifs d'écoute actuels, soit d'une réduction marquée de la production vocale durant cette période, comme cela a été suggéré pour d'autres espèces de baleines (Dréo, Bouffaut et al. 2019 ; Cerchio, Rasoloarijao et al. 2024). Les premières détections post-hivernales réapparaissent généralement en octobre autour de Madagascar, avant de s'étendre progressivement vers le sud avec l'approche de l'été austral (Cerchio, Rasoloarijao et al. 2024 ; Dréo, Crawford et al. 2025 ; Torterotot, Samaran, Stafford et al. 2020). Ces éléments suggèrent que Madagascar et les régions adjacentes pourraient constituer une zone de transition importante dans la migration des individus entre leurs habitats tropicaux et subantarctiques. Cependant, la complexité de leur cycle annuel, couplée à l'hétérogénéité spatiale des efforts de surveillance, appelle à une intensification des dispositifs d'écoute dans certaines régions encore peu couvertes afin de mieux cerner l'étendue et la fonction des habitats fréquentés.

Etat des connaissances sur le rorqual commun et le petit rorqual Antarctique

Outre les baleines bleues, l'OI austral est également fréquenté par deux autres espèces majeures de mysticètes : le rorqual commun (FW) (*Balaenoptera physalus*) et le petit rorqual Antarctique (AMW) (*Balaenoptera bonaerensis*).

Le rorqual commun - FW Le rorqual commun (FW) est une espèce cosmopolite, présente dans tous les océans et également dans le sud-ouest de l'OI où les données

acoustiques ont permis de mieux documenter sa phénologie et ses déplacements saisonniers. Il émet principalement deux types de vocalisations de basse fréquence (Fig. 2.12) : un chant stéréotypé à 20 Hz, composé d'impulsions puissantes et régulières émises majoritairement par les mâles en période de reproduction (Watkins et al. 1987 ; Simon et al. 2010), et des impulsions plus courtes autour de 40 Hz, associées à l'alimentation (Romagosa et al. 2021 ; Shabangu, Munoz et al. 2024).

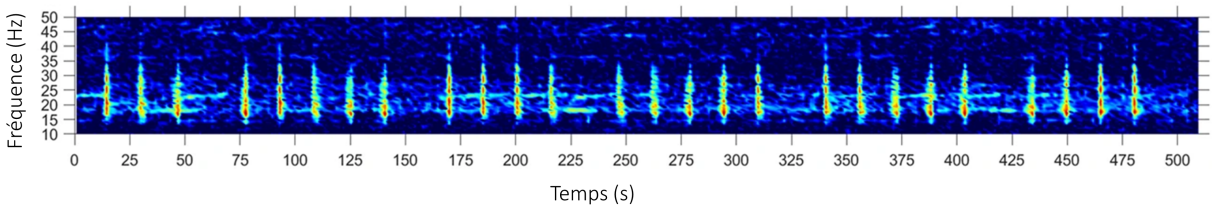


Figure 2.12 – Spectrogramme illustrant des vocalisations 20 Hz de FW sur 500 secondes. Enregistrement réalisé en Australie (Aulich et al. 2019).

Dans cette région, les chants à 20 Hz sont détectés de façon saisonnière, avec une arrivée progressive des individus en eaux subantarctiques dès l'automne austral (mars-avril), un mouvement migratoire vers le nord entre mai et août, puis un pic de détection au printemps austral dans les zones méridionales, avant un déclin à partir de l'été austral (octobre-novembre) (Leroy, Samaran, Stafford et al. 2018 ; Dréo, Bouffaut et al. 2019 ; Dréo, Crawford et al. 2025). En revanche, les impulsions à 40 Hz, associées à une activité d'alimentation, sont présentes toute l'année, notamment autour des îles du Prince-Édouard (Shabangu, Munoz et al. 2024).

Ces variations reflètent une alternance saisonnière entre comportements reproductifs, culminant en hiver austral, et alimentaires, plus fréquents en été austral. Ainsi, les eaux subantarctiques et subtropicales du sud-ouest de l'OI jouent un rôle clé dans le cycle de vie du FW, même si les sites exacts de reproduction restent encore mal connus, soulignant la nécessité d'une surveillance accrue.

Le petit rorqual Antarctique - AMW Le petit rorqual Antarctique (AMW) est une espèce de mysticète relativement peu connue, mais qui occupe une place importante dans les écosystèmes marins du sud-ouest de l'OI. Sa vocalisation principale consiste en une série de sons courts, pulsés et modulés en fréquence, généralement centrés entre 60 et 120

Hz, et pouvant se répéter en séquences rapides (Fig.2.13) (Risch, Gales et al. 2014). Ces vocalisations sont souvent associées à des comportements sociaux et à la communication entre individus au sein des groupes (Allen et al. 2024), et pourraient également jouer un rôle dans la recherche de nourriture collective.

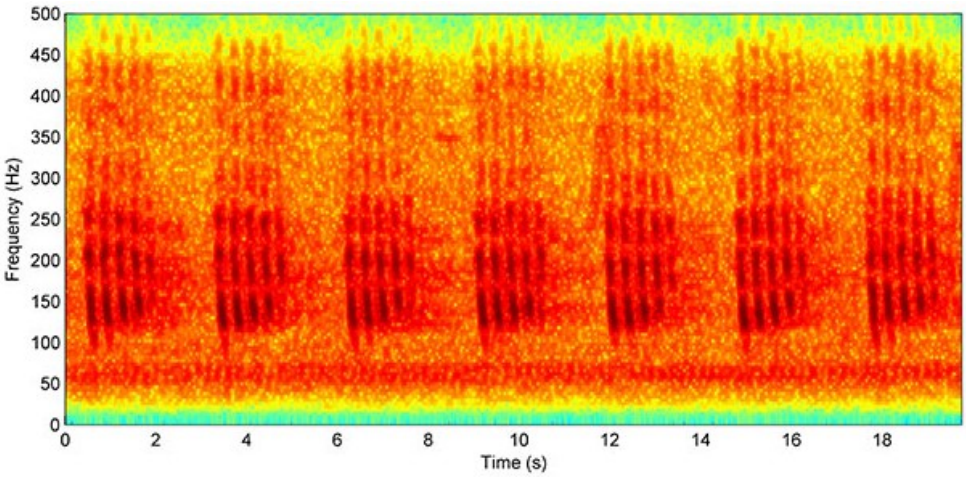


Figure 2.13 – Spectrogramme illustrant des vocalisations d'AMW sur 20 secondes. Enregistrement réalisé en 2003 dans le Perth Canyon (Australie-Occidentale) (Erbe et al. 2017).

Cette espèce fréquente principalement les eaux tempérées et subantarctiques, avec des détections acoustiques enregistrées surtout durant l'hiver austral, correspondant vraisemblablement à leur période d'alimentation (Filun et al. 2020). Les zones autour des îles subantarctiques du sud-ouest de l'OI, notamment les îles Kerguelen, Amsterdam et Prince Edward, apparaissent comme des habitats clés, en raison de la richesse trophique favorable à cette espèce (Leroy, Royer, Alling et al. 2021). Des enregistrements récents ont également confirmé sa présence acoustique à des latitudes plus tropicales, notamment dans la région de Mayotte, durant l'hiver et le printemps australs (Dréo, Crawford et al. 2025), suggérant une migration partielle vers le nord durant la seconde moitié de l'année. Contrairement au FW et à la SWIO PBW, cette migration tropicale n'avait pas encore été clairement mise en évidence et les données restent fragmentaires, témoignant de la nécessité d'intensifier la surveillance pour mieux comprendre les mouvements migratoires de l'espèce (Shabangu, Munoz et al. 2024).

Du point de vue social, l'AMW présente une structure de type fission-fusion, avec des associations à court terme et peu de liens durables (Casey et al. 2022). Les individus se rencontrent souvent en petits groupes de 2 à 6, mais de grandes agrégations allant jusqu'à

40 individus ont été observées, généralement en lien avec une forte disponibilité de proies (Risch, Gales et al. 2014). Les vocalisations et les regroupements pourraient faciliter la communication sociale et l'alimentation collective, indiquant que la socialité joue un rôle dans l'écologie alimentaire de l'espèce (Allen et al. 2024).

En raison de leur discrétion, de leurs intervalles de surface courts et de l'inaccessibilité de leur habitat polaire, l'AMW reste difficile à étudier. Néanmoins, les suivis satellitaires et les enregistrements acoustiques permettent d'obtenir des informations sur leur comportement, la dynamique de groupe et les stratégies de forage (Allen et al. 2024). Le renforcement des campagnes de suivi, tant acoustiques que visuelles, est nécessaire pour identifier les habitats critiques et mieux comprendre les réponses de l'espèce aux variations environnementales.

2.1.2 Les jeux de données utilisés

Le réseau OHASISBIO

Le réseau OHASISBIO, présenté en détail au Chapitre 1 (Chants des mysticètes dans l'océan Indien : le réseau OHASISBIO), constitue le jeu de données acoustiques utilisé dans l'étude 1 de ce chapitre. L'analyse et le traitement des enregistrements acoustiques ont déjà été abordés dans deux thèses doctorales antérieures : celle d'Emmanuelle Leroy 2017 (Université de Bretagne Occidentale), Surveillance acoustique des baleines bleues en Antarctique et dans l'océan Indien austral : traitement, analyse et interprétation, et celle de Maëlle Torterotot 2020 (Université de Bretagne Occidentale), Traitement et analyse de données bioacoustiques dans l'océan Indien austral : application aux baleines bleues. Ces travaux ont permis de détecter et de caractériser la présence acoustique de plusieurs espèces, en particulier différentes populations acoustiques de baleines bleues. La présente thèse s'appuie sur ces analyses, notamment sur les détections saisonnières de populations acoustiques décrites précédemment, et vise à approfondir leur interprétation écologique en examinant les relations avec des facteurs environnementaux associés.

La dernière thèse en date portait sur les données collectées jusqu'à fin 2018 (Torterotot 2020). Pour les travaux présentés ici, les données issues des enregistrements entre 2019 et 2023 ont été traitées par la même personne dans le cadre de l'ANR MUSIC (Royer et Samaran, 2024-2027) et avec les mêmes outils d'analyse automatique, permettant ainsi

de prolonger la série temporelle de manière homogène. Pour cette étude, l'ensemble des enregistrements disponibles dans le réseau OHASISBIO, couvrant tous les sites, a été utilisé. Le fonctionnement du détecteur automatique permet d'obtenir un nombre de vocalisations détectées par unité de temps. Dans le cadre de ce chapitre et des résultats présentés à travers l'étude 1, ce nombre de vocalisations a été agrégé en DPM, correspondant au nombre de minutes par jour avec présence acoustique confirmée. De plus, seules les vocalisations correspondant aux chants stéréotypés propres à chaque population acoustique de baleines bleues ont été incluses. Les appels non stéréotypés, tels que les D-calls, ont été exclus car ils ne peuvent pas être affiliés de manière fiable à une population acoustique spécifique (Torterotot, Samaran et Royer 2023).

Le réseau REVOSIMA

La zone étudiée se situe dans le Parc naturel marin de Mayotte, une aire protégée reconnue pour sa biodiversité marine exceptionnelle, mais où les connaissances sur les grandes baleines restent limitées. Le volcan sous-marin actif Fani Maoré, situé à l'est de Mayotte, a motivé le déploiement du programme de surveillance sismo-acoustique Réseau de Surveillance Volcanologique et Sismologique de Mayotte (REVOSIMA), visant à suivre l'activité volcanique tout en collectant des données acoustiques fines sur la faune marine (Volcanological et al. 2021). Ce dispositif combine deux réseaux complémentaires : le réseau hydroacoustique MAHY et les réseaux sismologiques MAYOBS.

Le réseau MAHY comprend quatre hydrophones autonomes disposés en quadrilatère d'environ 70 km de côté, centré à environ 50 km autour du volcan (Fig. 2.14). Ces stations sont déployées à des profondeurs comprises entre 1 300 et 1 366 mètres, dans le canal SOFAR, et enregistrent en continu à une fréquence d'échantillonnage de 240 Hz, permettant une captation fine des sons sous-marins. Le réseau MAHY a été initialement déployé en octobre 2020, puis les instruments ont été régulièrement entretenus et redéployés afin d'assurer une continuité des enregistrements jusqu'en octobre 2024 (Fig. 2.15). La configuration, la profondeur de mouillage et la fréquence d'échantillonnage sont similaires à celles utilisées sur le réseau OHASISBIO, situé plus au sud dans l'océan Indien (Fig. 1.3), ce qui facilite la comparaison et la complémentarité des données.

Parallèlement, les réseaux MAYOBS sont constitués de 3 à 16 sismomètres océaniques

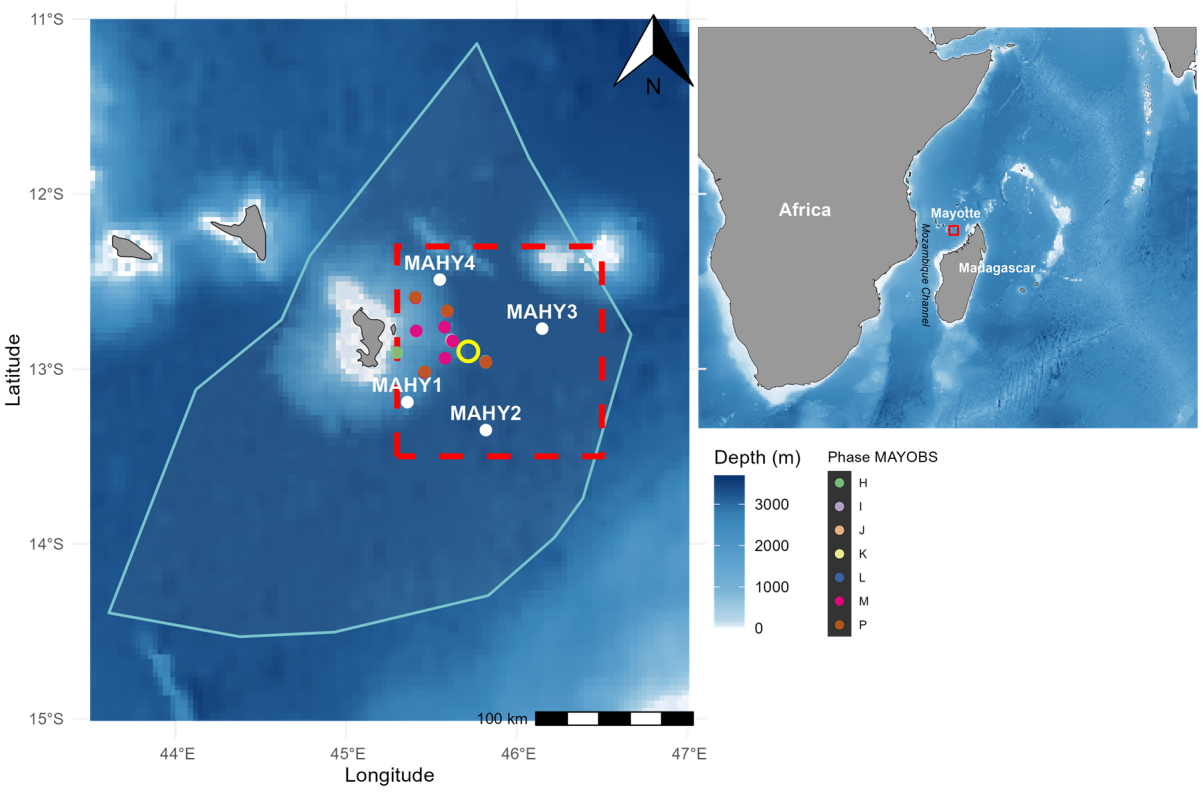


FIGURE 2.14 – Localisation des sites de mouillage du réseau REVOSIMA dans le sud-ouest de l'océan Indien. Les points blancs correspondent au réseau hydroacoustique MAHY, tandis que les points colorés représentent les déploiements du réseau sismologique MAYOBS durant différentes phases. Certains enregistreurs MAYOBS se superposent spatialement, donc tous les points ne sont pas forcément visibles. Le cercle jaune indique la position du volcan sous-marin Fani Maoré. Le rectangle rouge en pointillés délimite la zone sélectionnée pour l'extraction des variables environnementales autour des instruments. La zone ombrée en bleu avec contour bleu représente les limites du Parc naturel marin de Mayotte.

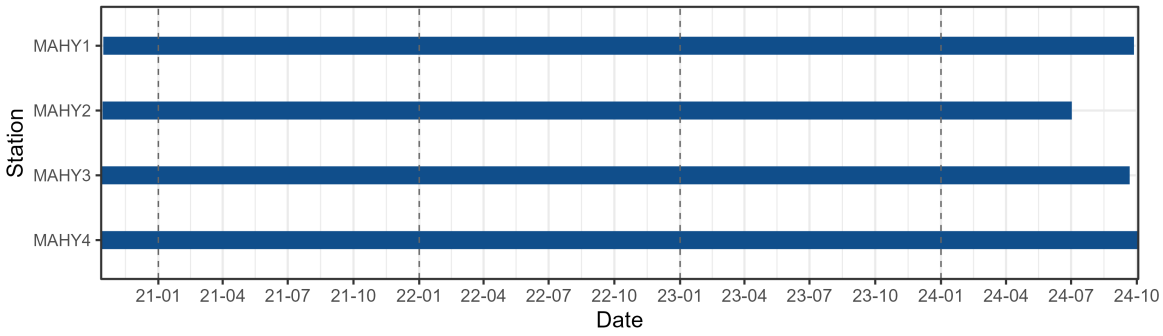


FIGURE 2.15 – Périodes d'enregistrement des sites du réseau MAHY.

(OBS) équipés de sismomètres à courtes périodes et d'hydrophones, déployés sur le fond marin à des profondeurs moyennes d'environ 2 000 mètres (Fig. 2.14). Leur localisation précise, les périodes d'enregistrement et la profondeur sont détaillées dans l'Annexe C. Ces instruments sont installés pour des durées variables de 1 à 8 mois, avec des taux d'échantillonnage plus faibles (62,5 à 125 Hz).

Quatre espèces de mysticètes ont été ciblées pour cette étude, leurs chants stéréotypés étant enregistrés dans la bande de fréquence programmée des hydrophones. Ces espèces sont la baleine bleue Antarctique (ANT BW), la baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW), le petit rorqual Antarctique (AMW) et la rorqual commun (FW), conformément aux descriptions issues de la littérature récente (Fig.2.16) (Leroy, Royer, Bonnel et al. 2018 ; Torterotot, Samaran, Stafford et al. 2020 ; Shabangu, Munoz et al. 2024). Le chant de ces espèces est constitué de séries de vocalisations répétées présentant un intervalle régulier, appelé ICI (inter-call interval), qui correspond au temps séparant deux unités sonores consécutives au sein d'une même série. Ces séries sont séparées par des périodes plus longues, appelées ISI (inter-series interval), durant lesquelles aucune vocalisation stéréotypée n'est produite. Cette structuration rythmique des chants facilite l'identification automatique à partir des enregistrements, en servant de signature acoustique stable pour chaque espèce ou population.

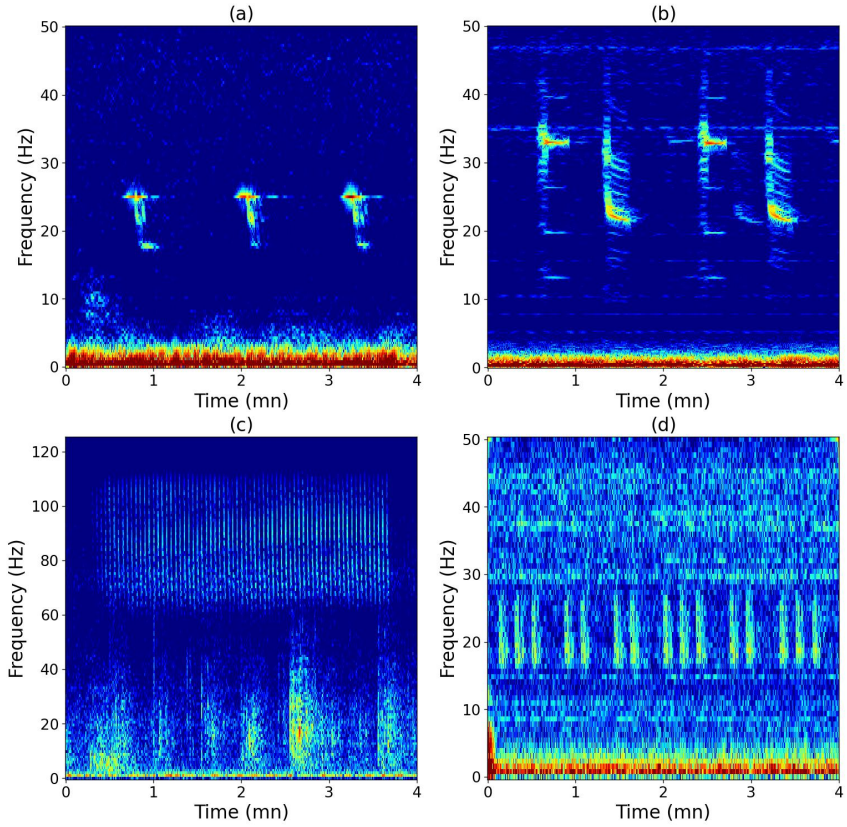


FIGURE 2.16 – Spectrogrammes illustrant des chants de a) ANT BW, b) SWIO PBW, c) AMW, et d) FW. Enregistrements réalisés dans la zone d'étude du réseaux MAHY (DRÉO, CRAWFORD et al. 2025).

La détection automatique de ce jeu de données repose sur la méthode développée par DRÉO, CRAWFORD et al. (2025), qui s'appuie sur l'analyse rythmique des chants via un traitement cepstral. Le cepstrogramme est une représentation du signal dans laquelle les périodicités spectrales sont mises en évidence, ce qui permet de détecter des motifs répétitifs dans le chant des baleines même en présence de bruit. Cette méthode robuste extrait des motifs rythmiques spécifiques et calcule un ratio pic-à-creux (p2vr) sur le cepstrogramme pour quantifier la force de répétition, produisant un score de détection horaire par espèce. Cette approche s'avère peu sensible aux variations de propagation acoustique et au bruit ambiant, assurant ainsi une détection fiable.

Pour cette étude, toutes les stations des réseaux MAHY et MAYOBS ont été combinées en une série temporelle quotidienne unique, c'est-à-dire que tous les points ont été rassemblés pour n'en former qu'un seul. Une espèce est considérée présente durant une

heure si au moins un enregistreur détecte son chant. Les détections horaires sont ensuite agrégées en une métrique quotidienne (DPH), correspondant au nombre d'heures par jour avec présence acoustique confirmée. L'analyse se concentre sur la période débutant en octobre 2020, lorsque les données des deux réseaux sont disponibles de manière continue et simultanée, garantissant un jeu de données homogène sur plusieurs années et couvrant l'ensemble de la zone étudiée.

Il est important de noter que certaines périodes de prospection sismique, notamment du 27 décembre 2020 au 4 février 2021 et du 20 février au 7 mars 2021, ont généré un bruit de fond important, masquant les détections des espèces ANT BW, SWIO PBW et FW. Les données de ces périodes ont été exclues pour ces espèces, tandis que les détections de AMW, non affectées par ce bruit, ont été conservées.

2.1.3 Modéliser la présence acoustique : les outils statistiques

Pour modéliser la présence acoustique des cétacés, il est nécessaire de s'appuyer sur des concepts écologiques robustes et des outils statistiques adaptés. Cette section introduit ces concepts avant de détailler les méthodes employées.

Le concept de niche écologique

Avant d'aborder la modélisation spatio-temporelle des distributions d'espèces, il est essentiel de comprendre le concept fondamental de la niche écologique.

> « La niche écologique est définie comme un hypervolume multidimensionnel regroupant l'ensemble des conditions abiotiques et biotiques nécessaires à une espèce pour maintenir un taux de croissance positif dans un écosystème. Elle inclut non seulement les facteurs abiotiques (e.g. température, lumière, etc.) et les interactions biotiques (e.g. compétition, prédation, mutualisme, etc.), mais aussi les interactions au sein de l'espèce elle-même (intra-spécifiques) ainsi que les impacts que l'espèce exerce sur son environnement. » (Hutchinson 1957)

La qualité d'un habitat dépend ainsi de sa capacité à fournir les conditions nécessaires à la survie, à la reproduction et à la persistance des populations (Block et Brennan 1993). On distingue donc deux types de variables : les conditions, sur lesquelles l'espèce

n'a pas d'effet direct, et les ressources avec lesquelles elle interagit (Fig. 2.17). La niche écologique correspond donc à l'ensemble des conditions environnementales dans lesquelles une espèce est capable de survivre et de se reproduire, en fonction de ses traits biologiques et de ses adaptations évolutives.

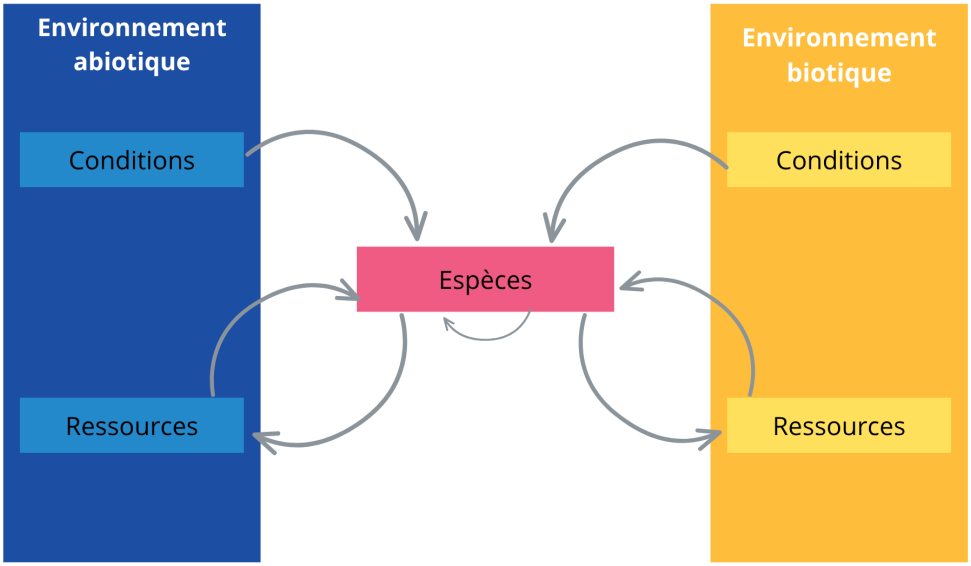


FIGURE 2.17 – Schéma illustrant le concept de niche écologique. Figure adaptée et refaite de MCINERNY et ETIENNE 2012b.

Toutefois, la distribution spatiale réellement observée de cette espèce résulte également de contraintes externes, telles que les barrières géographiques, les limitations de dispersion ou les interactions biotiques. Ainsi, la niche représente le potentiel écologique de l'espèce, tandis que sa distribution effective reflète l'influence combinée de ce potentiel et des contraintes du milieu. Le diagramme BAM (Fig. 2.18) illustre graphiquement ces relations entre niche, facteurs biotiques/abiotiques et accessibilité spatiale.

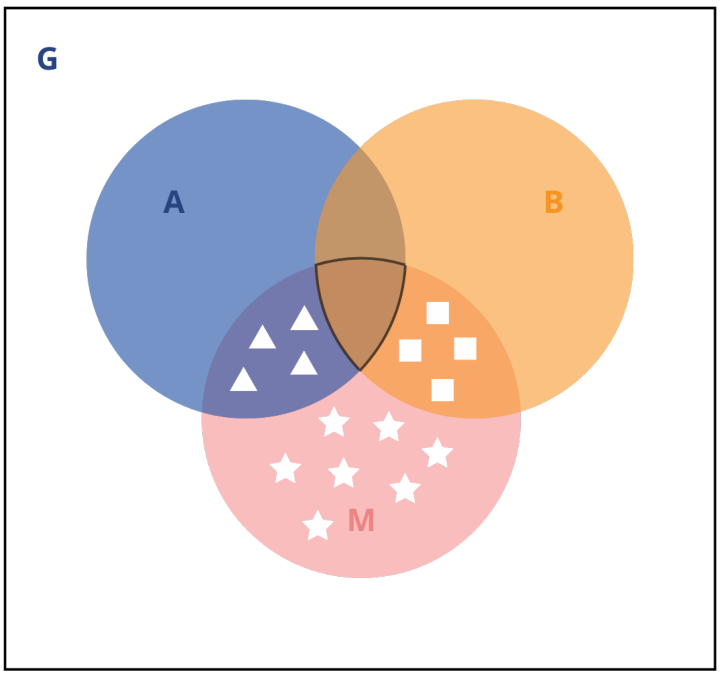


FIGURE 2.18 – Diagramme BAM : représentation heuristique de la niche écologique d'une espèce. G représente la superficie totale de l'étude. A prend en compte les facteurs abiotiques (bleu) c'est à dire l'espace géographique où les conditions environnementales sont suffisamment favorables pour que les populations aient un taux de croissance intrinsèque positif. B prend en compte les facteurs biotiques (orange) c'est à dire l'espace géographique où les interactions biotiques sont favorables (facilitation, symbiose) ou tolérés (compétition). M prend en compte les mouvements (rose) et représente l'aire géographique accessible à l'espèce en fonction de ses capacités de dispersion. L'interaction de A et de B correspond à la niche réalisée de l'espèce. L'interaction de A, B et M (trait noir) correspond à l'aire géographique potentielle de l'espèce résultant de l'interaction des 3 facteurs où sont présentes des populations sources. En dehors de cette zone, il est possible de trouver des populations puits dans M grâce à des événements de dispersion depuis la zone centrale, même si les fonctions sont défavorables d'un point de vue abiotique (triangles), biotique (carrés) ou les deux (étoiles). Figure adaptée et refaite de SOBERÓN 2007.

La niche écologique comporte également une dimension temporelle importante, reflétant la variabilité des activités vitales au cours du cycle de vie (alimentation, reproduction, migration, repos). Ainsi, la niche globale peut se décomposer en sous-niches fonctionnelles spécifiques à chaque activité, chacune associée à des habitats, des conditions environnementales et des interactions biotiques distincts. Ces sous-niches fonctionnelles correspondent à différentes phases du cycle biologique, chacune caractérisée par des exigences environnementales et comportementales spécifiques. Par exemple, chez les baleines, les zones de reproduction tropicales diffèrent nettement des aires d'alimentation polaires, illustrant cette segmentation spatio-temporelle (CLAPHAM 1996 ; HORTON et al. 2022).

Cette niche, dynamique et flexible, l'est particulièrement dans l'environnement océanique caractérisé par une forte hétérogénéité et variabilité spatio-temporelle. Les cétacés adaptent leur utilisation de l'espace et des ressources selon les saisons, les variations interannuelles, et les anomalies climatiques qui modifient la disponibilité des proies et les conditions environnementales (Santos et al. 2013 ; Baumgartner et Mate 2005). Cette plasticité niche-habitat est cruciale pour leur survie dans un milieu aussi changeant.

De plus, cette flexibilité peut varier à l'échelle individuelle, influencée par l'expérience, l'âge ou le statut reproducteur. Certains individus plus expérimentés ou dominants peuvent ainsi occuper des habitats préférentiels ou adopter des stratégies migratoires spécifiques, affectant la dynamique globale de la niche écologique au sein des populations (Austin, Bowen et McMillan 2004 ; Quinn et Brodeur 1991).

Les modèles de distribution d'espèces (SDM) en écologie marine

Les modèle de distribution d'espèces (SDM) exploitent ce cadre théorique pour prédire la répartition spatiale des espèces en reliant des données de présence, d'absence ou d'abondance à des variables environnementales des sites d'observation étudiés (Elith et Leathwick 2009 ; Sánchez-Mercado, Ferrer-Paris et Franklin 2010). Ces modèles reposent sur deux hypothèses principales : d'une part, la distribution des espèces est principalement déterminée par les variables environnementales ; d'autre part, les espèces sont en équilibre avec ces variables, de sorte que leur répartition reflète leurs préférences écologiques et leurs limites physiologiques (Guisan et Zimmermann 2000).

Il est essentiel de distinguer les SDM, qui visent principalement à prédire la distribution spatiale des espèces à partir de variables environnementales abiotiques (comme la température, la bathymétrie ou la productivité), des *Ecological Niche Models* (ENM), qui cherchent à modéliser la niche fondamentale d'une espèce en intégrant, de manière plus explicite, les interactions biotiques telles que la prédation, la compétition ou les relations trophiques (McInerny et Etienne 2012a). Ainsi, bien que les deux approches soient complémentaires, les SDM tendent à se focaliser sur l'habitat disponible, sans toujours refléter la complexité écologique des niches.

Appliqué aux prédateurs marins supérieurs (e.g. cétacés, oiseaux marins), ce cadre permet d'étudier comment la variabilité océanique spatio-temporelle façonne leur distri-

bution et les habitats occupés par ces espèces (Virgili, Racine et al. 2017 ; Lambert et al. 2014). Certaines espèces changent de niche ou d'habitat selon la saison, tandis que d'autres modulent leur distribution géographique tout en conservant une niche stable. Dans un environnement océanique dynamique, la plasticité niche-habitat est un élément clé, notamment pour les cétacés confrontés à des déplacements massifs et à une forte variabilité des ressources trophiques.

Nature des données et spécificités des données acoustiques passives

Les données utilisées dans les SDM se distinguent selon leur nature : données de présence seule (observations sans information sur l'absence), données de présence-absence (données systématiques indiquant la détection ou non d'une espèce), ou données d'abondance/comptage (nombre d'individus détectés). Comme résumé dans le tableau 2.1, les données issues du PAM peuvent prendre diverses formes, depuis des indicateurs binaires jusqu'à des mesures semi-quantitatives ou continues, telles que le taux d'activité acoustique ou le nombre d'heures positives à la détection par jour. Ces différents formats nécessitent des approches statistiques adaptées (Redfern et al. 2006 ; Mellinger, Stafford et al. 2007).

TABLE 2.1 – Types de données issues du PAM et leur interprétation biologique.

Type de donnée	**Format**	**Plage des valeurs**	**Interprétation biologique**
Présence acoustique (présence/absence)	Binaire (0/1)	$\{0, 1\}$	Indique la détection ou non de présence acoustique durant une période donnée (ex. jour, session).
Nombre brut de présence acoustique	Entier positif	$[0, +\infty[$	Quantification de l'activité acoustique (nombre de vocalisations, de rencontre acoustique etc.).
Taux d'activité acoustique (e.g. clics/min)	Continu positif	$[0, +\infty[$	Mesure normalisée permettant de comparer l'activité acoustique sur différentes périodes.
Heures ou minutes positives par jour	Continu borné	$[0, 24\,h]$ ou $[0, 1440\,min]$	Durée cumulée des fenêtres temporelles comportant au moins une détection acoustique.
Probabilité de détection	Probabilité	$[0, 1]$	Estimation statistique issue de modèles reflétant la confiance dans la présence d'une présence acoustique.

Dans cette thèse, les analyses portent sur des données exprimées en « heures ou minutes positives par jour », où la métrique correspond au nombre cumulé d'heures ou de minutes contenant au moins une détection acoustique, et la fenêtre temporelle d'agrégation retenue est la journée entière (24 heures), comme défini dans l'Introduction générale, à la section «Choix des métriques et méthodes statistiques». Ces mesures (DPM ou DPH) sont des données continues bornées (de 0 à 24 heures ou 1440 minutes) présentant une forte proportion de zéros (jours sans détection) et une variabilité continue positive (jours avec détection acoustique).

Ainsi, le terme présence acoustique sera employé pour qualifier la variable réponse modélisée et est défini comme suit :

> « La présence acoustique est un indice basé sur la détection de vocalisations, reflétant la présence acoustique des cétacés et, par extension, la présence relative des espèces. »

Il convient de noter que l'absence de détection acoustique ne garantit pas l'absence réelle des individus, notamment en raison de comportements silencieux ou de conditions acoustiques défavorables limitant la propagation du son. L'interprétation doit donc rester

prudente et privilégier l'analyse des tendances relatives spatio-temporelles plutôt qu'une estimation absolue de densité ou d'abondance (Marques et al. 2013).

Les données issues du PAM constituent un indicateur robuste et pertinent pour décrire la distribution relative des espèces à travers le temps et l'espace à condition d'avoir des protocoles d'enregistrement standardisés et des durées d'échantillonnage suffisantes (Redfern et al. 2006). La modélisation de données semi-quantitatives, telles que les heures positives, est particulièrement adaptée aux distributions *zero-inflated* (c'est-à-dire comportant un excès de valeurs nulles par rapport à une distribution de Poisson ou binomiale), et des méthodes comme les modèles modèle additif généralisé (GAM) avec distribution Tweedie sont couramment utilisées à cet effet (Wood, Pya et Säfken 2016).

Variables environnementales intégrées aux SDM

On distingue généralement deux catégories de variables environnementales dans les SDM : les variables proximales, qui influencent directement les espèces (e.g. la disponibilité en proies), et les variables distales, qui sont des indicateurs indirects de l'environnement (e.g. la bathymétrie) (Guisan et Zimmermann 2000). Ces dernières comprennent des descripteurs physionomiques statiques, comme la distance à la côte ou la pente. Elles incluent également des variables physico-chimiques dynamiques, telles que la température de surface, la hauteur de la mer ou la vitesse du vent. L'accessibilité de ces données et leur disponibilité via des bases de données publiques, à différentes résolutions spatiales et temporelles, facilitent leur intégration dans les modèles. En milieu marin, la concentration de surface en chlorophylle a (Chla) est fréquemment utilisée comme proxy de la biomasse primaire, tandis que les sorties de modèles écosystémiques simulant la biomasse et la production du zooplancton ou micronekton sont de plus en plus mobilisées pour représenter les niveaux trophiques intermédiaires (Lehodey, Senina et Murtugudde 2008 ; Virgili, Authier et al. 2019). Des programmes de mesure *in situ*, comme le programme THEMISTO, permettent de collecter des données sur la biomasse de proies pouvant être intégrées aux modèles lorsque leur résolution spatiale et temporelle le permet. Cependant, ces mesures restent rares et ponctuelles, notamment dans les environnements océaniques reculés et difficiles d'accès, ce qui limite fortement la représentativité et la précision des modèles. Une étude exploratoire sur cette intégration a été menée et est présentée en section «Stage de Master : Combining passive and active acoustic approaches to study

predator-prey interactions of blue whales in the southern Indian Ocean» de ce chapitre.

Échelles spatiales et temporelles en milieu marin

La distribution des ressources en milieu marin est hiérarchisée dans l'espace et le temps, avec des concentrations irrégulières influencées par des phénomènes physiques tels que les remontées d'eau (*upwellings*), les tourbillons, les fronts et les courants. Ces structures créent, à différentes échelles, des « oasis » de ressources favorables à l'agrégation des proies et de leurs prédateurs (Hunt et Schneider 1987 ; Longhurst 2010 ; Godø et al. 2012 ; Worm, Lotze et Myers 2003).

La prise en compte des échelles spatiales et temporelles est donc essentielle en écologie, car les comportements des prédateurs et la disponibilité des ressources varient selon ces dimensions (Levin 1992). Deux composantes fondamentales sont distinguées : l'étendue, correspondant à la surface ou au domaine spatial étudié, et le grain, c'est-à-dire la résolution spatiale minimale des données (Wu et Li 2006).

En milieu océanique, Hunt et Schneider (1987) identifie quatre échelles spatiales majeures (Fig.2.19) :

- **Fine (1 m – 1 km)** : phénomènes rapides et localisés, tels que les tourbillons de surface (circulation de Langmuir) ou les *upwellings* côtiers locaux. Ces structures influencent directement la distribution des nutriments et des proies à petite échelle, modulant le comportement alimentaire des prédateurs.
- **Coarse (1 – 100 km)** : remontées d'eau étendues induites par le vent, favorisant la productivité primaire et créant des zones d'alimentation localisées mais persistantes.
- **Méso (100 – 1000 km)** : courants océaniques et tourbillons associés, comme le Gulf Stream, redistribuant la chaleur, la salinité et les nutriments sur des centaines de kilomètres et structurant les migrations et l'agrégation des espèces.
- **Macro (>1000 km)** : circulation thermohaline globale, déterminant la redistribution mondiale des masses d'eau, des nutriments et de la chaleur sur des décennies et des milliers de kilomètres, servant de toile de fond aux processus à plus petite échelle.

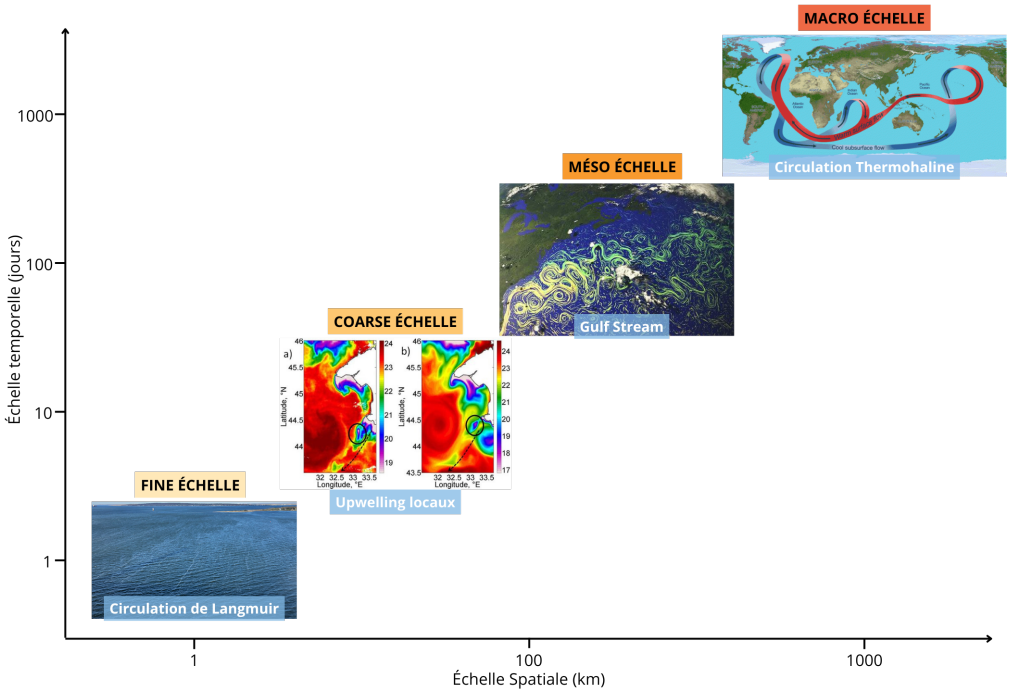


FIGURE 2.19 – Structure hiérarchique de l'environnement marin. Chaque échelle est définie par une dimension spatiale et temporelle : plus l'événement est vaste, plus sa durée est longue. Pour chaque échelle, un exemple de phénomène océanographique est fourni. À fine échelle, la circulation de Langmuir consiste en une série de tourbillons de surface induits par le vent (SCIENCE PRIMER 2020). À échelle coarse, les remontées d'eau (*upwellings*) sont provoquées par l'action du vent qui fait remonter les eaux froides et riches en nutriments (KUBRYAKOV, ALESKEROVA et MIZYUK 2024). À méso-échelle, le Gulf Stream est un courant chaud accélérant vers le nord au large de la côte est de l'Amérique du Nord et générant de multiples tourbillons (CNRS NEWS 2021). À macro-échelle, la circulation thermohaline est entraînée par les gradients de densité globaux créés par les flux de chaleur et d'eau douce en surface (CENTER FOR ENVIRONMENTAL STUDIES, FAU 2023). Figure adaptée et refaite de VIRGILI 2018

Les prédateurs marins adaptent donc leurs déplacements et leur comportement alimentaire à cette complexité multi-échelle. Ils combinent l'exploitation de caractéristiques environnementales persistantes à grande échelle (coarse à méso), telles que les fronts océaniques ou les bordures de plateau continental, pour optimiser la recherche de nourriture, avec des stratégies localisées de type Area-Restricted Search (ARS) lorsqu'ils rencontrent des zones à forte densité de proies. Dans ce cas, ils ralentissent leur vitesse et augmentent la fréquence des changements de direction afin de rester dans les zones riches en ressources (BALLANCE, PITMAN et FIEDLER 2006 ; FAUCHALD 1999 ; WEIMERSKIRCH 2007 ; HAYS et al. 2016).

Chez les cétacés, cette intégration multi-échelle se traduit par l'exploitation d'habitats stables (e.g. monts sous-marins, plateaux continentaux, fronts) et de phénomènes éphémères à méso-échelle (e.g. tourbillons, remontées d'eau côtières) (Kaschner 2007 ; Baumgartner, Mullin et al. 2001 ; Cotté et al. 2009 ; Johnston et al. 2008 ; Friedlaender et al. 2006). Les espèces à forte demande énergétique, comme certains delphinidés, sont particulièrement sensibles aux variations locales de production primaire et de densité de proies, tandis que les grands mysticètes adoptent des distributions plus homogènes (Mannocci et al. 2014).

La complexité multi-échelle des habitats et des comportements alimentaires souligne l'importance de sélectionner des variables environnementales adaptées aux différentes échelles spatiales et temporelles. Les comportements rapides nécessitent des variables instantanées à fine échelle, tandis que les processus saisonniers ou annuels se modélisent mieux avec des variables moyennées ou climatologiques (Virgili, Authier et al. 2019 ; Lambert et al. 2014). Dans le cadre de la modélisation à long terme de la distribution des prédateurs marins, il est recommandé d'utiliser des variables persistantes à échelle coarse ou méso, plutôt que des variables éphémères à fine échelle, afin de capturer la structure générale de la disponibilité des ressources et de mieux prédire la niche écologique des espèces.

Limites et contraintes des SDM en milieu marin

Les SDM permettent de comprendre et de prédire les relations entre les espèces et l'environnement, mais leur validité reste restreinte aux conditions observées lors des relevés. Lorsqu'ils sont appliqués à d'autres zones géographiques ou à des périodes différentes, leur capacité prédictive diminue sensiblement, en particulier dans des régions écologiquement divergentes (Elith et Leathwick 2009 ; Randin et al. 2006).

Cette faible transférabilité s'explique en grande partie par la non-stationnarité des relations espèce-environnement : en milieu marin, ces relations peuvent varier selon l'espace et le temps, sous l'effet de facteurs locaux, saisonniers ou comportementaux, tels que les dynamiques trophiques, la phénologie ou les interactions interspécifiques. Les modèles supposant des relations fixes risquent alors de mal représenter la réalité écologique, ce qui limite fortement leur usage pour des prédictions spatiales ou temporelles en dehors du

cadre d'entraînement (Redfern et al. 2006).

Cette contrainte méthodologique constitue un enjeu majeur pour la conservation et la gestion, qui nécessitent des outils capables d'anticiper des distributions futures ou dans des zones non échantillonnées. Cela motive le développement d'approches plus complexes, intégrant la variabilité spatio-temporelle et les interactions écologiques, bien que ces méthodes restent souvent limitées par la disponibilité des données.

2.1.4 Les principales techniques de modélisation de l'habitat

Un large éventail de techniques existe pour modéliser la distribution des espèces en écologie marine (Guisan et Zimmermann 2000 ; Redfern et al. 2006 ; Franklin 2010a ; Peterson et al. 2011).

Les méthodes d'ordination

Les méthodes d'ordination sont des techniques descriptives appartenant à la famille des analyses multivariées. Leur objectif principal est de résumer les gradients environnementaux en un nombre réduit de dimensions, plus facilement interprétables que les données brutes (Jongman, Ter Braak et Van Tongeren 1995). Ces méthodes permettent d'explorer la répartition des espèces le long de ces gradients, en partitionnant la variance des données sur plusieurs axes orthogonaux et indépendants, qui sont des combinaisons linéaires des variables environnementales. Selon la méthode, les relations entre espèces et environnement peuvent être supposées linéaires (analyse en composantes principales, analyse de redondance) ou unimodales (analyses de correspondance et de correspondance canonique).

Ces méthodes sont particulièrement utilisées en écologie communautaire pour analyser l'habitat utilisé par plusieurs espèces simultanément, comme dans le cas des dauphins tachetés de l'Atlantique et des grands dauphins au Mexique (Griffin et Griffin 2003), ou des oiseaux marins dans le Pacifique tropical (Ballance, Pitman et Fiedler 2006).

Les arbres de classification et de régression

Les arbres de classification et de régression (CART) constituent une méthode non paramétrique fondée sur un partitionnement récursif et binaire des données, permettant de diviser un ensemble en sous-groupes homogènes en fonction des variables prédictives (Elith, Leathwick et Hastie 2008 ; Guisan et Zimmermann 2000). Selon la nature de la variable réponse, on distingue les arbres de classification (pour une variable catégorielle) et les arbres de régression (pour une variable continue). Ces techniques permettent de révéler des interactions complexes entre variables et de prédire des patrons, mais peuvent générer des erreurs de classification, souvent réduites par l'assemblage de multiples arbres simples dans des modèles de type *boosted regression trees* ou *random forests* (Elith, Leathwick et Hastie 2008).

Bien que peu répandus dans l'étude des relations entre cétacés ou oiseaux marins et leur habitat, les modèles fondés sur les arbres ont montré leur intérêt pour analyser les préférences d'habitat des cétacés (MacLeod, Brereton et Martin 2009) ou modéliser la distribution passée de diverses espèces de baleines (Monsarrat et al. 2015). Plus récemment, des approches par *random forests* ont été utilisées pour modéliser la présence acoustique de grandes baleines autour d'îles subantarctiques à partir de données d'acoustique passive, illustrant la pertinence de ces méthodes pour analyser la variabilité spatiale et temporelle de l'occurrence des cétacés (Shabangu, Munoz et al. 2024). Toutefois, en raison de leur sensibilité aux données d'échantillonnage et de leur capacité limitée à modéliser des fonctions continues et lisses, les arbres présentent souvent une performance prédictive inférieure à celle des modèles de régression traditionnels (Hastie, Tibshirani et Friedman 2009 ; Franklin 2010b).

Les réseaux de neurones artificiels

Les réseaux de neurones artificiels (ANN) représentent une approche non paramétrique inspirée du fonctionnement du cerveau animal, composée de couches de neurones artificiels interconnectés (couches d'entrée, cachée et sortie) (Lek et Guégan 1999 ; Liu et al. 2010). Cette architecture permet d'apprendre des relations complexes et non linéaires entre les variables explicatives et la variable réponse, ce qui en fait un outil puissant pour la modélisation prédictive lorsque les relations sous-jacentes sont inconnues ou difficiles à

modéliser par des méthodes classiques.

Cependant, les ANN sont souvent considérés comme des «boîtes noires» en raison du manque d'interprétabilité de leurs décisions, et requièrent de grandes quantités de données pour un apprentissage efficace (Liu et al. 2010). Ils ont été utilisés avec succès, par exemple, pour prédire la probabilité d'occurrence du cachalot en Méditerranée centrale (Ai"ssi et al. 2014).

Les modèles de régression

Les modèles de régression constituent la méthode la plus couramment employée en écologie marine pour modéliser la distribution des espèces (Redfern et al. 2006). Ils établissent une relation entre une variable réponse, telle que la présence acoustique ou l'abondance d'une espèce, et un ou plusieurs prédicteurs environnementaux. Ces derniers peuvent être des variables brutes ou des gradients issus des méthodes d'ordination afin de limiter la colinéarité.

Les modèles linéaires simples supposent une relation linéaire entre la variable réponse et les prédicteurs, ainsi qu'une distribution normale et une variance constante, des hypothèses rarement vérifiées en écologie (Guisan et Zimmermann 2000). Les modèle linéaire généralisé (GLM) étendent cette approche en permettant d'utiliser différentes distributions (e.g. binomiale, Poisson, Gamma, etc.) et en liant la moyenne de la variable réponse à une fonction linéaire via une fonction de lien appropriée, ce qui permet de mieux ajuster des données non normales (McCullagh 2019).

Les modèles additifs généralisés (GAM)

Les modèle additif généralisé (GAM) constituent une extension souple et puissante des modèles linéaires généralisés (GLM), largement utilisés en écologie pour explorer les relations complexes et potentiellement non linéaires entre une variable réponse et plusieurs prédicteurs (Hastie 2017 ; Guisan et Zimmermann 2000 ; Redfern et al. 2006 ; Wood 2017b).

Contrairement aux GLM, qui imposent une forme fonctionnelle prédéfinie (souvent

linéaire), les GAM intègrent des fonctions de lissage non paramétriques, telles que les splines, pour modéliser des réponses flexibles aux variables explicatives. Cette propriété est particulièrement utile dans les systèmes écologiques, où les effets environnementaux peuvent être complexes, non linéaires et spécifiques à des seuils ou à des interactions (Virgili, Racine et al. 2017).

Cette flexibilité a rendu les GAM très populaires en écologie marine, notamment pour modéliser la répartition des espèces en fonction des gradients environnementaux, qu'ils soient physico-chimiques (température, salinité), biologiques (productivité, biomasse) ou spatio-temporels (Wood 2017b ; Zuur et al. 2009). En revanche, les GAM standards ne tiennent pas compte des effets aléatoires ni des structures spatiales ou temporelles complexes. Cela peut limiter leur capacité prédictive dans des contextes écologiques hétérogènes. Des extensions comme les GAMM (modèles additifs généralisés mixtes) ou les modèles hiérarchiques spatio-temporels peuvent pallier ces limites, à condition de disposer de données structurées (Wood 2017b).

2.2 Méthodologie : modélisation de la présence acoustique des baleines

Ce travail repose sur une approche de modélisation statistique fondée sur les GAM, afin d'étudier la présence acoustique des cétacés dans l'OI austral. Deux volets complémentaires ont été développés à partir de mesures semi-quantitatives issues de l'acoustique passive : d'une part, les DPM moyennés à l'échelle hebdomadaire (étude 1) et, d'autre part, les DPH (étude 2).

Les variables de réponse sont continues mais présentent une distribution très asymétrique, avec une proportion élevée de valeurs nulles. Pour tenir compte de cette structure particulière, tous les modèles ont été ajustés à l'aide d'une distribution Tweedie. Cette famille de distributions, appartenant à la loi exponentielle, est particulièrement adaptée aux données dites *zero-inflated*, car elle combine une masse de probabilité en zéro et une composante continue strictement positive (Wood 2017b). Elle permet ainsi de représenter simultanément la probabilité de détection et l'intensité de l'activité vocale enregistrée (Wood, Pya et Säfken 2016).

Une fonction lien logarithmique a été utilisée sur la variable réponse pour réduire la variance, linéariser les relations avec les prédicteurs et conserver la présence des zéros dans les données :

$$\mu_i = \log(\mathrm{DPM}/\mathrm{DPH}_i + 1) \tag{2.1}$$

où :

— μ_i est la variable réponse transformée pour l'observation i ;
— DPM_i représente le nombre de minutes positives à la détection par jour pour l'observation i ;
— DPH_i représente le nombre d'heure positive à la détection par jour pour l'observation i.

La sélection des variables explicatives s'est appuyée à la fois sur des hypothèses écologiques établies et sur un contrôle statistique de la colinéarité, évaluée par les coefficients de corrélation de Pearson. Dans la majorité des cas, les variables retenues présentaient des corrélations faibles à modérées ($|r| < 0.5$). Toutefois, une colinéarité plus forte a pu être acceptée lorsqu'elle était écologiquement pertinente, par exemple entre la température de surface et la chlorophylle dans l'étude 1.

L'ajustement des modèles a été évalué à l'aide de critères statistiques classiques. Le critère d'information d'Akaike (AIC) permet de comparer différents modèles en tenant compte à la fois de la qualité d'ajustement et du nombre de paramètres (Akaike 1974). Des valeurs d'AIC plus faibles indiquent, toutes choses égales par ailleurs (i.e. lorsque la vraisemblance est calculée sur le même jeu de données et la même fonction de vraisemblance), des modèles plus parcimonieux. La déviance expliquée (*deviance explained*) mesure la proportion de la déviance totale des données expliquée par le modèle, fournissant une indication de la capacité du modèle à reproduire la variabilité observée, de manière analogue à un R^2 pour les modèles linéaires. Les analyses ont été conduites dans `R` à l'aide de la bibliothèque `mgcv` (Wood 2017a ; R Core Team 2021).

2.2.1 Étude 1 : Océan Indien Austral - Réseau OHASISBIO

La première étude s'appuie sur les enregistrements du réseau OHASISBIO, qui couvre une vaste portion de l'OI subantarctique. Ce dispositif offre une perspective à la fois

spatiale et temporelle, avec une dizaine de sites suivis pendant plus d'une décennie.

Pour chaque point d'écoute, les variables environnementales ont été extraites puis moyennées dans une zone de 50×50 km centrée sur l'enregistreur. Cette approche permet de représenter de manière plus réaliste les conditions locales influençant la propagation du son et d'estimer les caractéristiques de l'habitat potentiellement utilisé par les individus détectés autour de l'enregistreur. Le choix d'une distance de 50 km correspond à une estimation conservatrice de la portée de détection. Ce seuil a été retenu afin de limiter les biais liés à la variabilité importante des distances de propagation rapportées dans la littérature, qui varient de 10 km à plus de 700 km selon les régions, les saisons et les populations acoustiques (Samaran, Adam et Guinet 2010a ; Širović, Hildebrand et Wiggins 2007). Les analyses portent sur trois populations acoustiques distinctes : ANT BW, SEIO PBW et SWIO PBW.

Et les variables environnementales retenues sont :

— **température de surface de la mer (SST)** : indicateur central des conditions thermiques régionales et de la distribution des proies. Les données proviennent de la réanalyse globale Copernicus Marine (DOI : 10.48670/moi-00021), à une résolution spatiale d'environ 8 km et une résolution temporelle quotidienne, agrégée ensuite à l'échelle hebdomadaire. La SST est exprimée en °C.

— **chlorophylle a (Chla)** : proxy de la biomasse phytoplanctonique et de la productivité primaire (Godet et al. 2020). Les données proviennent de Mercator Océan (DOI : 10.48670/moi-00019), avec une résolution spatiale d'environ 25 km (1/4°) et une résolution quotidienne, moyennée de façon hebdomadaire. La Chl-a est exprimée en $\mathrm{mg} \cdot \mathrm{m}^{-3}$.

— **concentration moyenne de glace de mer (MSI)** : paramètre influençant les migrations de baleines et la disponibilité de proies en zone subantarctique (Širović, Hildebrand, Wiggins et al. 2004 ; Nicol, Worby et Leaper 2008). Les données proviennent du NOAA/NSIDC (version 4), à résolution de 25 km et suivi journalier (Meier et al. 2014). Le pourcentage de glace de mer a été estimé sur la zone 50°E–85°E et -70°S—10°S, en comptabilisant les pixels dont la concentration dépasse 15 %, seuil couramment utilisé pour distinguer glace et eau libre (Orsi, Whitworth III et Nowlin Jr 1995).

Modélisation : Les GAM intègrent des termes lissés pour chacune de ces variables, ainsi

qu'un terme spatial bivarié (longitude, latitude) pour décrire la variabilité géographique. Des effets linéaires supplémentaires pour le mois et l'année capturent respectivement la saisonnalité et les tendances temporelles. Pour chaque combinaison mois–année, les prédictions ont été réalisées sur une grille régulière de 50×50 km couvrant l'ensemble de la zone d'étude. Le modèle général est le suivant :

$$\begin{aligned} g(\mu_i) &= \mu_0 + \beta_1 \cdot \mathrm{Month}_i + \beta_2 \cdot \mathrm{Year}_i \\ &\quad + s_1(\log(\mathrm{Chla}_i)) + s_2(\mathrm{SST}_i) + s_3(\mathrm{MSI}_i) + s_4(\mathrm{lon}_i, \mathrm{lat}_i) + \varepsilon_i \end{aligned} \tag{2.2}$$

où :

— μ_i est la valeur attendue (moyenne) de la variable réponse pour l'observation i ;
— μ_0 est l'ordonnée à l'origine (intercept) ;
— β_1 et β_2 représentent respectivement les effets linéaires du mois et de l'année ;
— $s_j(\cdot)$, pour $j \in \{1, \ldots, 4\}$, désignent les fonctions lissées associées aux covariables environnementales : chlorophylle-a (Chla), température de surface de la mer (SST), concentration moyenne en glace de mer (MSI) et effet spatial (longitude, latitude) ;
— ε_i est le terme résiduel aléatoire.

Analyse complémentaire par EOF : Afin d'extraire les principaux modes spatio-temporels présents dans les prédictions, une analyse en fonctions orthogonales empiriques (EOFs) a été réalisée (Thorson et al. 2020 ; Alglave et al. 2024). Cette méthode, proche d'une analyse en composantes principales, associe un motif spatial à une évolution temporelle et permet d'identifier des variations écologiques dominantes.

L'analyse EOF a été complétée par un *clustering* hiérarchique, regroupant à la fois les zones spatiales présentant des trajectoires temporelles similaires (habitats fonctionnels) et les périodes temporelles partageant des structures spatiales proches (saisons écologiques). Afin d'éviter que des valeurs extrêmes, notamment autour de Kerguelen, ne biaisent les résultats, cette zone a été exclue de l'analyse. L'ensemble a été implémenté dans `R` avec la bibliothèque `FactoMineR` (Lê, Josse et Husson 2008 ; R Core Team 2021).

2.2.2 Étude 2 : Nord du canal du Mozambique - Réseau REVOSIMA

La deuxième étude se concentre sur une zone tropicale située au nord du canal du Mozambique, à proximité de Mayotte et du volcan sous-marin Fani Maoré. Cette zone bénéficie d'une couverture instrumentale dense grâce au réseau REVOSIMA. Quatre espèces de baleines à fanons sont considérées : ANT BW, SWIO PBW, AMW et FW.

Les données de SST, d'anomalie de SST et de Chla proviennent des mêmes jeux que pour l'étude 1 (Mercator Océan et Copernicus Marine). Les variables environnementales retenues sont donc :

- **SST**
- **Chla**
- **SST anomaly** : différence entre la SST instantanée et la moyenne saisonnière climatologique (1995–2010). Cette variable renseigne sur la variabilité interannuelle et les tendances climatiques de long terme.
- **hauteur significative des vagues (SWH)** : indicateur de l'état de la mer, influençant la distribution des proies et le comportement des prédateurs. Données issues de la réanalyse ERA5 du Copernicus Climate Change Service (résolution 0.25°, suivi journalier) (DOI : 10.24381/cds.adbb2d47). La SWH est exprimée en mètres.

Les variables environnementales ont été moyennées sur une zone fixe de 130×130 km englobant tous les enregistreurs des réseaux MAHY et MAYOBS. Ce choix garantit que les individus détectés se situent bien dans la zone d'étude. La taille retenue (70 km de portée maximale) correspond à une valeur conservatrice de la détection acoustique, en cohérence avec les estimations disponibles dans la littérature, allant d'une dizaine à plusieurs centaines de kilomètres selon les conditions (Samaran, Guinet et al. 2010 ; Shabangu, Munoz et al. 2024 ; Širović, Hildebrand et Wiggins 2007).

Un prétraitement supplémentaire a été appliqué : exclusion des valeurs extrêmes (0,5 % les plus basses et hautes) et lissage médian sur 15 jours de la SWH, afin de réduire la variabilité bruitée non écologique.

Modélisation : Le cadre de modélisation appliqué à chaque espèce repose sur des GAM. Les variables environnementales sont introduites sous forme lissée, certaines en interaction avec l'année pour capturer la variabilité interannuelle. Le terme `Week` est

également modélisé de façon lissée, représentant la dynamique saisonnière intra-annuelle.

$$\begin{aligned} g(\mu_i) &= \mu_0 + s_1(\text{Week}_i, \text{Year}_i) \\ &\quad + s_2(\log(\text{Chla}_i), \text{Year}_i) + s_3(\text{SST}_i, \text{Year}_i) \\ &\quad + s_4(\text{SST anomaly}_i, \text{Year}_i) + s_5(\text{SWH}_i, \text{Year}_i) + \varepsilon_i \end{aligned} \tag{2.3}$$

où :

— μ_i est la valeur attendue (moyenne) de la variable réponse pour l'observation i ;

— μ_0 est l'ordonnée à l'origine (intercept) ;

— $s_j(\cdot)$, pour $j \in \{1, \ldots, 5\}$, désignent les fonctions lissées capturant les effets non linéaires des covariables environnementales : semaine de l'année (*Week*), chlorophylle-*a* transformée en logarithme (Chla), température de surface de la mer (SST), anomalie de température de surface (SST anomaly) et hauteur significative des vagues (SWH) ;

— ε_i est le terme résiduel aléatoire.

Lorsque l'ajout d'interactions avec l'année ne conduisait pas à une amélioration significative (selon l'AIC et la déviance expliquée), des effets lissés simples ont été retenus.

2.3 Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern indian ocean

scientific reports

Check for updates

OPEN

Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern Indian Ocean

Mathilde Michel[1]✉, Baptiste Alglave[2], Maxime Olmos[3], Maëlle Torterotot[1], Auriane Virgili[4], Salomé Martin-Marin[1], Jean-Yves Royer[5] & Flore Samaran[1]

Blue whales in the Indian Ocean have been severely depleted by previous extensive commercial whaling. A good understanding of their spatio-temporal distribution is crucial for conservation. The songs of three blue whale acoustic populations - Antarctic blue whales (*Balaenoptera musculus intermedia*, ANT BW) and pygmy blue whales (*B. musculus brevicauda*) from the Southeast (SEIO PBW) and Southwest Indian Ocean (SWIO PBW) - were analyzed using 13 years of passive acoustic recordings from 10 sites in the southwest Indian Ocean. Although blue whale vocalizations comprise both songs and non-song calls (e.g., D-calls), the present study concentrates on the examination of songs. Generalized additive models (GAMs) were used to relate acoustic presence, measured by the number of positive minutes per day (averaged weekly), to environmental drivers such as sea surface temperature (SST), chlorophyll-a concentrations, and sea ice extent. These models allowed predictions of blue whale acoustic presence across the region. Empirical orthogonal functions (EOFs) were applied for dimensionality reduction to identify key habitats, including the Kerguelen Plateau and Madagascar Basin, which may serve as important feeding and resting zones based on acoustic presence and environmental data. Antarctic blue whales were predominantly detected in austral winter and spring, associated with lower SST and higher chlorophyll-a. In contrast, SEIO and SWIO pygmy blue whales were more frequent in summer and autumn, with some overlap suggesting ecological interactions. These findings lay the groundwork for targeted conservation efforts to protect critical blue whale habitats in a rapidly changing ocean.

Keywords Spatio-temporal modelisation, Blue whale, Acoustic recording

Blue whale (*Balaenoptera musculus*) populations in the southern hemisphere have been severely depleted by previous extensive commercial whaling. As a consequence, the International Whaling Commission (IWC) established the Indian Ocean Whale Sanctuary (IOWS) in 1979, prohibiting commercial whaling in the region[1]. The rare nature of blue whales, combined with the vast and remote expanses of their habitat, complicates efforts to observe them directly. Given these challenges, passive acoustic monitoring has emerged as the most effective method to study these cetaceans, allowing researchers to track their presence and behavior through sounds they produced[2–7]. This approach is particularly appropriate for blue whales, which are highly vocal animals characterized by long low-frequency sounds[8].

Blue whale sounds are broadly categorized into two types: songs and non-song calls. Songs are complex long-duration calls and are made up of repeated units organized in series, separated by regular intervals (intercall interval, ICI)[5]. They are recorded in all seasons in both feeding and breeding areas and are characterized by a stereotype time-frequency structure, allowing differentiation between populations[3,4]. In contrast, non-song calls, such as D-calls, are shorter, frequency-modulated vocalizations typically associated with foraging

[1]ENSTA, IPP, CNRS Lab-STICC, 29200 Brest, France. [2]Université Bretagne Sud, Lab-STICC, 56000 Vannes, France. [3]DECOD (Ecosystem Dynamics and Sustainability), L'Institut Agro, Ifremer, INRAE, 29280 Plouzané, France. [4]Share the Ocean, 56870 Larmor-Baden, France. [5]Geo-Ocean, CNRS, Université de Brest, Ifremer, 29280 Plouzané, France. ✉email: mathilde.michel@ensta.fr

behavior[9]. D-calls exhibit highly variable frequencies and are produced by both males and females, without a clear geographic variation in their time-frequency shape[10–13].

Temporally stable differences in songs provide characteristics that can be used to distinguish the two subspecies and to define several acoustic populations of pygmy blue whales in the Indian Ocean[14]. Although specific to each subspecies and acoustic population, blue whale songs share similar acoustic properties: the duration of song units exceeds 15 seconds, their amplitude is greater than 180 dB re 1 μPa at 1 m, and their pitch ranges from 20 to 100 Hz. Songs can be composed of multiple units, depending on the subspecies or acoustic population, and intercall intervals are periodic and do not exceed a few minutes. The year-round recording of blue whale vocal activity in various locations suggests that individuals produce acoustic emissions continuously throughout the year at the population level[3,4,7,15]. Furthermore, long-term analyzes have revealed a gradual decline in the frequency of blue whale calls in recent decades, although the underlying causes remain uncertain[16]. The differences in vocal repertoire between the blue whale subspecies and the acoustic populations of pygmy blue whales provide a unique means of tracking their migratory movements throughout an ocean basin[7,15,17,18].

The southern region of the Indian Ocean hosts seasonally several acoustic populations of blue whales, including different subspecies and acoustic populations. Two subspecies are regularly present in this area: the Antarctic blue whale (*B. m. intermedia*) and the pygmy blue whale (*B. m. brevicauda*)[19]. There are notable differences in the way populations use their habitats, each of which exhibits distinct migration patterns. Antarctic blue whales (ANT BW) follow a seasonal migration from subantarctic regions in austral autumn to subtropical areas in austral winter, before returning to subantarctic waters in austral spring (September to November)[2,3,5,20].

The acoustic populations of pygmy blue whales in the Southeast Indian Ocean (SEIO PBW) and the Southwest Indian Ocean (SWIO PBW) show distinct distribution patterns within the Indian Ocean and are predominantly observed during the austral autumn and winter seasons[3,4,7,15]. The SEIO PBW acoustic population is recorded mainly along the southwest coast of Australia, predominantly between March and May[21,22], while the SWIO PBW acoustic population is found in regions further west, within the southwestern Indian Ocean[23–25]. However, recent findings reveal that these acoustic populations have been identified beyond their previously recognized distributional limits. SEIO PBW have been observed farther west than expected, in proximity to Kerguelen Island[3,5,7], while SWIO PBW songs have been detected further south and west, near the Crozet Archipelago and southwest of Amsterdam Island[3,7].

Oceanographic parameters, including sea surface temperature, chlorophyll-a concentration, and sea ice extent, are critical in determining prey availability and subsequently influence whale migratory movements and foraging behavior. Previous research has concentrated primarily on modelling the associations between these environmental variables and the distribution of cetaceans on extensive spatial scales, predominantly using visual or aerial observation data[26–28]. Other studies have used passive acoustic data, but have been limited to smaller spatial and temporal scales, often focusing only on correlations between environmental variables and acoustic presence[29,30]. Building on this foundation, this study seeks to extend these analyses to a larger geographic scale and a longer time frame. By integrating passive acoustic monitoring with oceanographic data, we provide comprehensive predictions for the entire study region, thus enhancing our understanding of key blue whale habitats and supporting conservation efforts.

Specifically, this study aims to improve the characterization of blue whale habitat based on 13 years of acoustic recordings collected in the southern Indian Ocean. Using a generalized additive model (GAM) framework, we evaluated the influence of oceanographic covariates on the acoustic presence, defined as an acoustic intensity metric based on the number of positive detection minutes per day, averaged weekly. This modelling approach allows us to (i) identify key environmental drivers of whale distribution and assess their non-linear effects, and (ii) predict spatio-temporal variations in blue whale acoustic presence, including inter-annual and seasonal patterns. In a subsequent phase, we will delineate functional areas using Empirical Orthogonal Functions (EOFs), a spatio-temporal dimension reduction technique. The identified spatial and temporal patterns are then compared with the existing literature to ensure alignment with established knowledge. These results provide valuable information for the conservation of this iconic species.

Methods

Area Description

The study area in the southern Indian Ocean extends from 20°S to 60°S latitude and from 45°E to 90°E longitude. In this region, the SST varies significantly, ranging from about 28 °C at lower latitudes to nearly 0 °C near the South Pole. This gradient is influenced by the separation of the water masses at the subantarctic and polar fronts. The subantarctic front, at about 45°S, separates warm mid-latitude waters from cooler Antarctic waters, improving biological productivity[31]. Further south, between 50°S and 60°S, the polar front separates the subantarctic from the polar waters, promoting the formation of sea ice[32]. Antarctic sea ice expands seasonally, peaking in the austral winter and covering much of the southern Indian Ocean. These variations correspond to the natural cycles of sea ice formation and melting dictated by the regional climate. The chlorophyll-a distribution shows spatial variability, with high concentrations in nutrient-rich colder waters. On the Kerguelen Plateau, concentrations peak at 3.96 mg/m^3 due to nutrient accumulation and upwelling[33]. In contrast, lower latitudes have lower concentrations, around 0.02 mg/m^3, due to subsurface features that affect currents and nutrient transport[34]. Phytoplankton blooms on the polar front are highly variable in space and time, especially in December, and are associated with increased solar radiation and an increase in SST in the southern ocean[35].

Acoustic data collection

The acoustic data used in this study were collected by a wide network of autonomous hydrophones called the Hydro-Acoustic Observatory of Seismicity and Biodiversity (OHA-SIS-BIO)[36]. This network was deployed

in the southern Indian Ocean from December 2009 to February 2023 (see Fig. 1). Initially and until 2016, it included five sites : south of la Réunion Island (MAD), north of the Crozet Archipelago (NCRO), west of Kerguelen Island (WKER), southwest (SWAMS) and northeast (NEAMS) of the Saint-Paul and Amsterdam Islands. An additional site south of the Southeast Indian Ridge (SSEIR) was installed in 2014. In 2017, the geometry of the network was slightly modified to improve coverage in the northern regions and better pinpoint the location of southern blue whale feeding grounds. Five new sites were added: near the junction of the three Indian Ocean seafloor spreading ridges (RTJ), west (MAD-W) and east (MAD-E) of the Madagascar Basin, south of the southwest Indian ridge (SSWIR) and south of the Kerguelen Plateau (ELAN), while the MAD and NCRO sites were no longer deployed.

Each mooring consisted of an anchor, an acoustic release, an adjustable line, and a submerged buoy that houses the recording system. The instruments were deployed at depths ranging from 1000 to 1300 m on the axis of the Sound Fixing and Ranging (SOFAR) channel, which facilitates the transmission of low-frequency sounds over long distances[37]. The recording was continuous at 240 Hz sampling with a 24-bit analog-to-digital conversion. The hydrophones used had a sensitivity of -163.5 dB re 1V/µPa, except for the moorings at SWAMS in 2015 and 2016, where the sensitivity was -168.6 dB re 1V/µPa. The data collected were stored locally on each device[38]. The moorings were maintained during the annual scientific cruise of the R/V Marion Dufresne in the French Southern and Antarctic Territories in January/February each year. The database used in this study contains almost continuous acoustic recordings over 13 years, from 2010 to 2022.

Acoustic data analysis

In this study, blue whale song detections from 2010 to 2018 were taken from Torterotot et al.[7], and from 2019 to 2022 were processed using the same detection method, as follows.

A detection algorithm using dictionary learning and sparse representation was applied to automatically detect blue whale songs[39]. This technique uses large dictionaries to capture song variability and uses linear combinations of dictionary entries to represent song complexity. Without relying on a fixed template, the algorithm builds dictionaries directly from temporal call signals extracted from the data and reconstructs

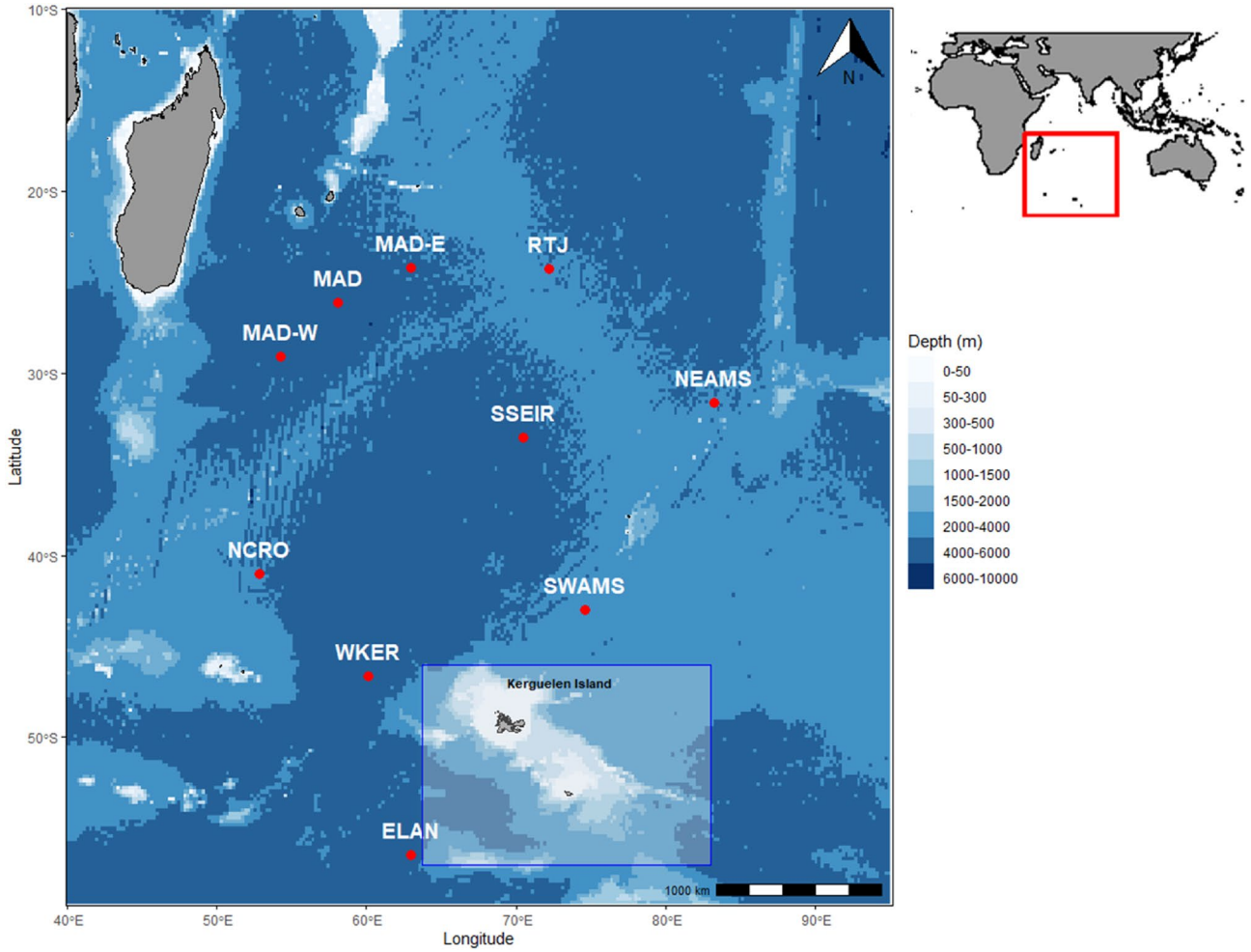


Fig. 1. Locations of the mooring sites of the OHA-SIS-BIO hydrophone network. The blue square represents the Kerguelen Plateau region. The map was created using R (version 4.1.2; https://www.r-project.org/)[51].

observed signals with sparse combinations of waveform entries. Higher-quality reconstructions correspond to higher similarity scores[39,40]. The signal used for recognition was the entire call for the ANT BW and a harmonic of a single unit of the call for both pygmy blue whales (see the white frames in Fig. 2). These harmonics were chosen for their high amplitude and frequency range that did not overlap with other blue whale songs.

The performance of this detection algorithm was evaluated on a substantial subset of manually annotated data spanning different seasons and locations within the OHA-SIS-BIO array. By setting the detection threshold to allow for one false positive per hour, we achieved a 90% recall for all calls with an SNR greater than 0 dB. A detailed review of all detections at the WKER location throughout the 2015 recordings confirmed that the number of false positives was very low, most of them caused by calls masked within the chorus[41]. The year-to-year frequency decrease in blue whale calls is very small (from -0.12 to .0.32 Hz/yr depending on the species)[15]. As a result, the dictionary from one year can still be applied effectively to the adjacent year. Although some calls may be missed due to a slight drop in frequency, most will still be detected successfully. A new dictionary was then built for each year of data, based on the calls detected by the adjacent year's detector. This approach effectively accounts for the frequency shift in the calls. A detailed evaluation of the algorithm performance in the dataset used is provided by Torterotot et al. (2019)[41].

Environmental variables

In this study, three dynamic environmental variables were analyzed to explain the acoustic presence of blue whales derived from acoustic detections. These variables were chosen because they effectively characterize the environment in the study area, as discussed in the "Area description" section.

- Sea Surface Temperature (SST) Changes in sea surface temperature can signal shifts in regional biological and chemical processes that affect productivity. The SST data in this study come from the Copernicus Marine global ocean reanalysis (DOI: 10.48670/moi-00021), which covers the period from 1992 to 2024, with a high spatial resolution of 1/12°(approximately 8 km). The reanalysis includes observations from satellites such as TOPEX POSEIDON and ERS-1 and models of 3D thermodynamic and dynamic variables. SST is measured in degrees Celsius (°C) with daily temporal resolution. Daily data were extracted and averaged over a 50 km area around each hydrophone, assuming the songs originated in that region. In addition, these averages were calculated weekly to align with the response variable under study. Figure 3 summarizes the SST data for the study area for each season.
- Chlorophyll-a concentration Chlorophyll-a concentration, a common indicator of primary productivity, is considered as a distal rather than proximal variable for prey such as zooplankton[42]. Although proximal variables typically better explain species distributions, the environmental variables chosen reflect prey distribution patterns[42]. Chlorophyll-a data were obtained from Mercator-Ocean, from January 1, 1993, to February 29, 2024, with a resolution of 1/4° and concentrations reported in mg/m^3(DOI: 10.48670/moi-00019). Data

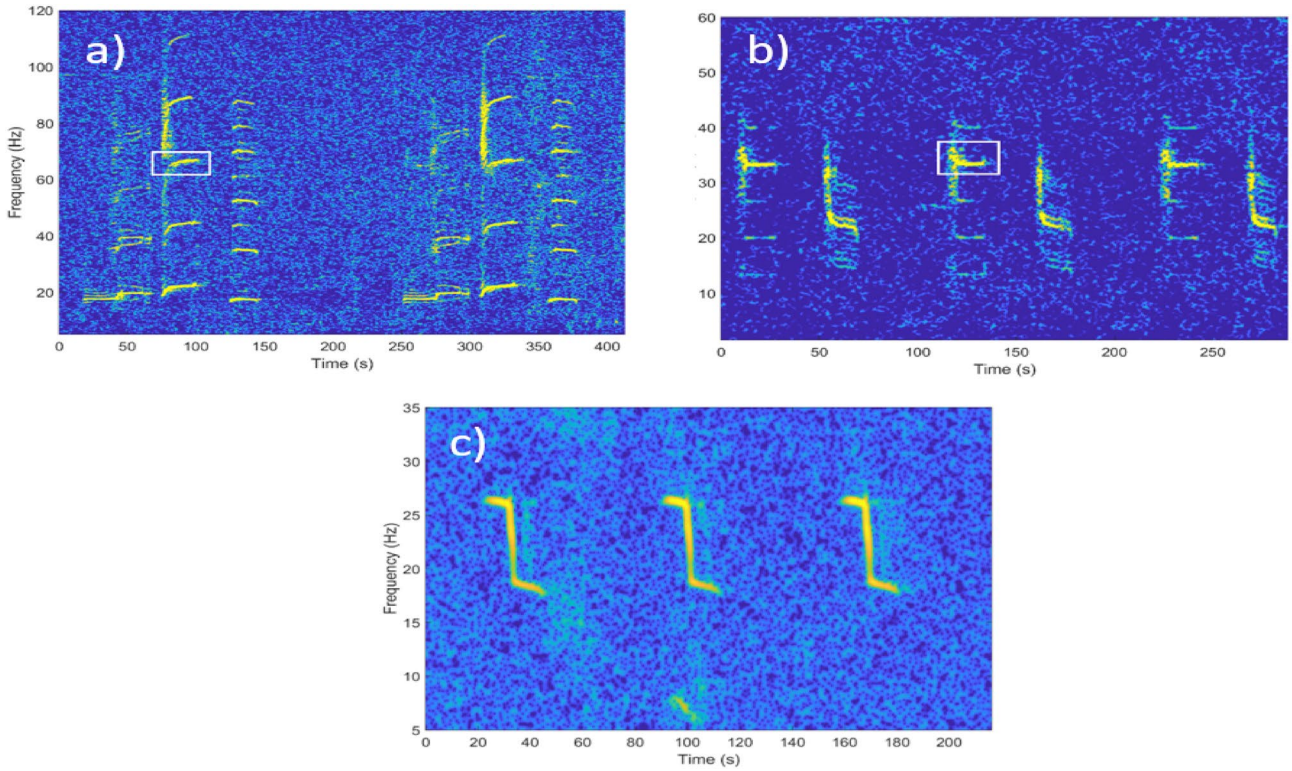


Fig. 2. Spectrograms of (**a**) SEIO PBW songs (**b**) SWIO PBW songs, and (**c**) ANT BW songs from the southern Indian Ocean. The units selected for the detector are framed in white.

were collected daily and averaged over a 50 km area per site to account for spatial variability. An overview of these data for the study area is presented in Fig. 3, which summarizes the chlorophyll-a data for each season.

- Sea Ice Concentration Sea ice significantly affects marine habitats, the availability of whale prey, and the migration of whales[28,43]. The sea ice concentration data come from the NOAA/NSIDC Climate Data Record, Version 4[44], based on passive microwave measurements following NOAA criteria. It combines estimates from NASA Team and Bootstrap algorithms, using the highest precision[45]. Data are available at 25 km x 25 km resolution, daily and monthly. We analyze sea ice in a specific area (50°E to 85°E, -70°S to -10°S) by calculating pixels that exceed a concentration of 15% ice, a standard threshold to distinguish sea ice from open water[32]. An overview of these data for the study area is presented in Fig. 3, which summarizes the sea ice concentration data for each season.
- Geographic coordinates Geographic coordinates of the study sites were included in the models to capture spatial variability not captured by the other covariates and to account for large-scale distributions resulting from the geographical extent of the subspecies[5].
- Temporal Parameters: Months and years Months and years were used as temporal predictors in the models to account for seasonal and annual variations in the acoustic presence of an acoustic population. These variations reflect the migratory patterns of the species and the inter-annual variability[7].

In this study, the austral seasons are defined as follows: summer includes December, January and February; autumn includes March, April, and May; winter includes June, July, and August; and spring includes September, October, and November. To simplify the visual representation of the seasonal patterns in Figs. 3 and 5, a single representative month was selected for each season. Specifically, January was designated for summer, April for autumn, July for winter, and November for spring.

Habitat modelling

The environmental variables (described above) may not have a linear relationship with the spatio-temporal distribution of the whales. Consequently, we modeled the acoustic presence of the whales with a spatio-temporal GAM[46]. This approach accounts for non-linear relationships with covariates, can deal with zero inflation through non-standard observation models (e.g. tweedie distribution), and models spatio-temporal correlations through spatial and temporal splines[47,48].

To fit the model, we used observations from the average number of minutes with positive detections per day (DPM), averaged weekly. Observations were zero-inflated positive with heavy tails. Consequently, we tested two different likelihood distributions that can be considered for count data that exhibit overdispersion and excess zeros (Tweedie or Negative Binomial, Table 1) to model the observations. The averaged number of minutes with positive DPM μ_i is modeled through the equation:

$$g(\mu_i) = \mu_0 + \beta_{\text{Month}} \cdot \text{Month}_i + \beta_{\text{Year}} \cdot \text{Year}_i + f_1(\text{Chla}_i) + f_2(\text{SST}_i) + f_3(\text{MSI}_i) + f_4(\text{lon}_i, \text{lat}_i)$$

where g is a link function (here $\log$), μ_0 is the intercept, the linear terms related to months and years are β_{Month} and β_{Year}, $f_j, j \in \{1, \cdots, 4\}$ are the smoothed functions related to the covariates chlorophyll-a (Chla), sea surface temperature (SST), mean sea ice (MSI), and the spatial effect related to longitude (lon) and latitude (lat). Collinearity among environmental covariates (Chla, SST and MSI) was assessed using pairwise correlation coefficients. A strong negative correlation was found between Chla and SST (r = -0.81), but both variables were retained in the model due to their ecological relevance and because AIC-based model selection supported their inclusion.

From this model, we provide spatio-temporal maps of the acoustic presence of the whales by predicting $\hat{\mu}$ over a 50 km^2 spatial grid that covers the spatial domain at each time step. Different models were tested to investigate the spatio-temporal variations in the acoustic data (Table 1). In the M2 reference model, we modeled acoustic data as a function of months, years, environmental covariates, and spatial effects related to longitude and latitude. Two models of lower complexity were also considered (M0, M1, Table 1). We used Akaike's information criterion (AIC) for model selection as a measure of the goodness of fit penalized by the number of parameters to account for model parsimony[49]. To avoid overfitting, we split the data into 80% training and 20% validation sets, and cross-validation was performed using Spearman's rank correlation (R_{Spearman}) to assess model performance. The results showed improved performance for more complex models that incorporated temporal and spatial effects. Furthermore, the contribution of each covariate to the explained variance in the GAM was assessed by comparing the deviation of the full model with reduced models, where one covariate was removed at a time. This approach enabled calculation of the proportion of variance explained by each covariate as the relative change in deviation, normalized by the null deviation. The asterisks (*) indicate significant contributions ($p < 0.001$). These significance values are provided to indicate the importance of each covariate in explaining the variance in the model (Fig. Supplementary 4). Finally, the coefficients of variation of the model predictions were calculated as the standard deviation divided by the mean of the predictions (Fig. Supplementary 5). The model was implemented using the 'mgcv' package[50] in R[51].

EOFs to identify functional areas

To identify functional areas based on GAM spatio-temporal predictions, we relied on standard spatio-temporal dimension reduction methods known as Empirical Orthogonal Functions (EOFs)[52,53].

EOFs are derived from climate science to extract the main spatial and temporal patterns in spatio-temporal data. They are principal component analysis (PCA) in a matrix where locations are individuals and time steps are variables. They have been widely used in marine ecology to analyze ecosystem variability, species distribution, and phenology[53,54]. Here, the spatio-temporal data are the model predictions μ at each location of the spatial

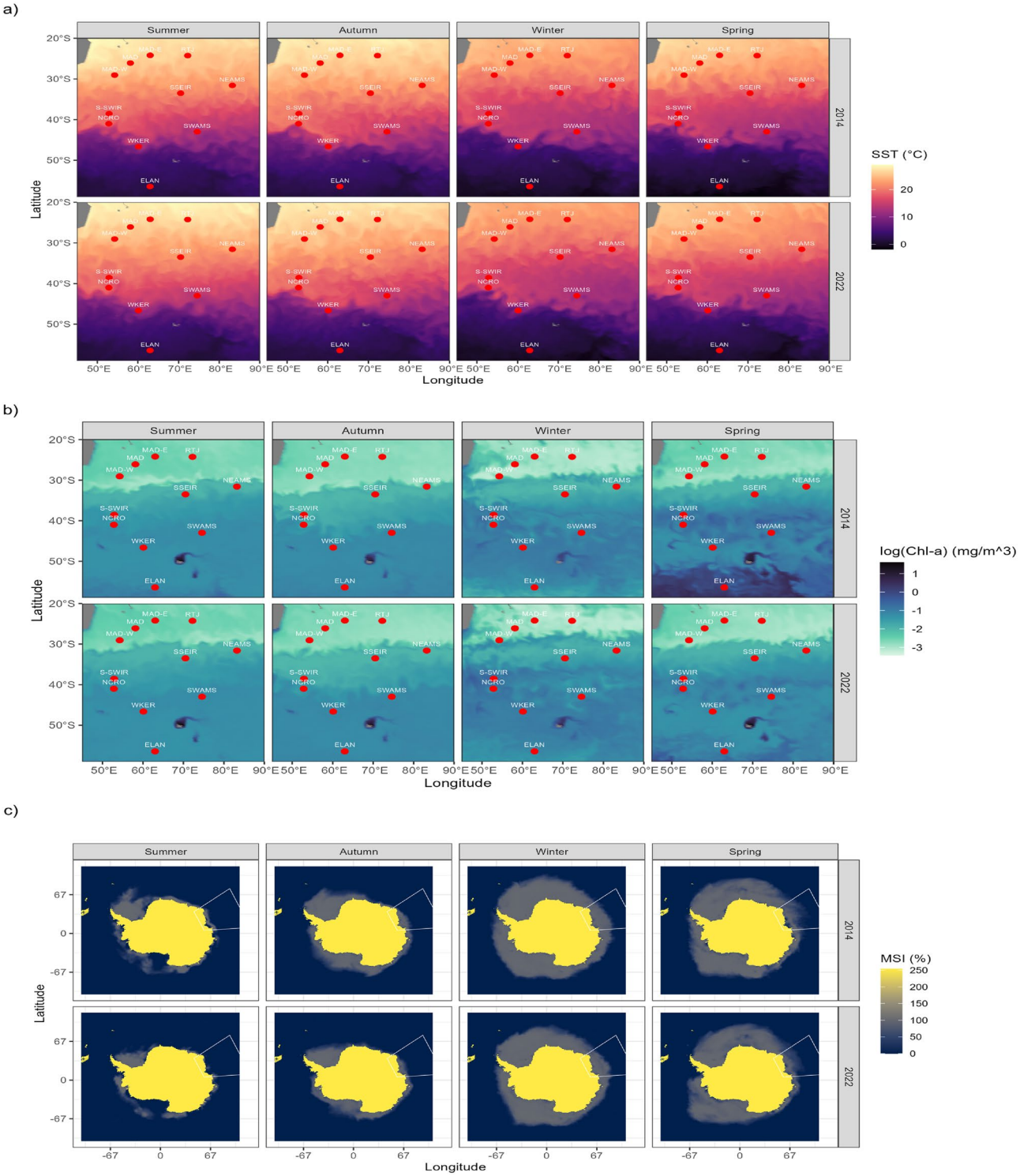


Fig. 3. Monthly averages of covariates over the entire study area in the Indian Ocean for the four seasons, represented by January (summer), April (autumn), July (winter), and November (spring), for the years 2014 and 2022. The variables represented are: (**a**) SST (°C), (**b**) log (Chl-a) (mg/m^3), and (**c**) MSI (%). The red dots indicate the locations of the hydrophones used in this study.

grid for each time step (month of each year). They are stacked in a matrix $n \times m$ where n is the number of spatial locations (here X) and m is the number of time steps (here Y).

EOFs summarize information from a sequence of maps into a smaller set of r maps, denoted $\mathrm{p}_j = (p_j(x_1), \cdots, p_j(x_n))$ where n is the number of spatial locations. These terms (called spatial factors) best represent the spatio-temporal variability patterns of the spatio-temporal variable. The first spatial factor of the

Models	Spatio-temporal variations in DPM are explained by	Ecological hypothesis	Acoustic population	Distribution	Deviance explained (%)	AIC	ΔAIC	Cross Validation (R_{Spearman})
M0	Chlorophyll-a + sea surface temperature + mean sea ice	Only depends on environmental variables, does not take into account monthly and annual variations in acoustic presence	ANT BW	Tweedie	11.6	92019	3234	0.38
			SEIO PBW	Tweedie	43.5	34165	3748	0.49
			SWIO PBW	Tweedie	20.9	76628	6755	0.47
M1	M0 + Month and Year effect	Account for inter-annual and monthly changes in acoustic whale distribution	ANT BW	Tweedie	16.4	90471	1686	0.45
			SEIO PBW	Tweedie	52	31952	1535	0.52
			SWIO PBW	Tweedie	32.2	72155	2282	0.58
M2	M1 + longitude and latitude effects	Take into account the areas of distribution of species and migration patterns	ANT BW	Tweedie	21.2	88785	0*	0.51
			ANT BW	Negative binomial	20.9	88809	24	0.51
			SEIO PBW	Tweedie	57	30417	0*	0.54
			SEIO PBW	Negative binomial	56.9	31569	1152	0.54
			SWIO PBW	Tweedie	37.6	69873	0*	0.63
			SWIO PBW	Negative binomial	37.1	70617	744	0.63

Table 1. Summary of the hypotheses tested, the associated model configurations, and AIC values attributed to each model for each acoustic population: ANT BW, SEIO PBW, and SWIO PBW. Explained deviance (%) quantifies the proportion of variation explained by the model. ΔAIC is the difference in the AIC score between the best model (*) and the model being compared.

EOF p_1 captures the highest amount of variability in the spatio-temporal data; the second p_2 is orthogonal to the first spatial factor p_1 and captures the second highest amount of variability. The same principle applies to the next dimensions.

For each dimension *j*, the spatial factor is related to a set of temporal indices called (temporal) loadings defined as $\boldsymbol{\alpha}_j = (\alpha_j(t_1), \cdots, \alpha_j(t_m))$, where *m* is the number of time steps. If the loading of a time step is positive ($\alpha_j(t) > 0$), then the spatio-temporal variable is distributed according to the related spatial factor p_j. If the loading of a time step is negative ($\alpha_j(t) < 0$), the spatio-temporal variable is distributed according to the opposite of the related spatial factor $-\mathrm{p}_j$. More details on the interpretation of EOFs can be found in Alglave et al. (2024)[55].

As with standard PCA, EOFs can be combined with clustering analysis (here Hierarchical Clustering on Principal Components) to regroup locations that have similar temporal trends (clustering on spatial factors) or time steps that have similar spatial patterns (clustering on temporal loadings)[53]. Clusters of locations can be interpreted as distinct functional habitats, while clusters of time steps can be interpreted as ecological seasons. Both can be plotted on a single graph, and locations can be related to seasons if they are in the same direction on the graph (similar to the case in PCA).

Note that both EOF dimensions and clustering can be strongly affected by extreme values. Given the high predictions surrounding Kerguelen Island, which distort the EOF representation of broad-scale spatial patterns, we exclusively retained predictions outside the Kerguelen area to perform the EOFs and clustering for the SEIO PBW and SWIO PBW acoustic population. EOF analysis and clustering were carried out using the 'FactoMineR' library[56] in R[51].

Results

Acoustic effort

Acoustic effort and detection data for each site and acoustic population are summarized in Table Supplementary 1. The total recording effort exhibited variation between sites, ranging from 1,630 days (NCRO) to 4,750 days (WKER), with a range of recording durations between sites, totaling approximately 2.3 million minutes (NCRO) to 6.8 million minutes (WKER). The proportion of days with acoustic presence exhibited variation across sites and acoustic populations, ranging from 0.0% at MADW for the SEIO PBW acoustic population up to 95.6% at NEAMS for the ANT BW acoustic population. Acoustic effort data reflect continuous and intermittent recording throughout the study period, with the frequency of detections varying over time. Although some sites show nearly constant acoustic presence (e.g., MAD, NEAMS), others exhibit more intermittent detections (e.g., RTJ, ELAN). The acoustic presence of each acoustic population and the recording periods are detailed in the supplementary materials (Fig. Supplementary 1, Supplementary 2, Supplementary 3).

Spatio-temporal variations in environmental covariates

The thermal gradient in the study area (Fig. 3) ranges from -4°C to more than 25°C, exhibiting a discernible latitudinal pattern, particularly evident during austral summer and autumn, with a SST peak in the northwest. These temperatures remain consistent from year to year and demonstrate minimal variability (Fig. 3a). Conversely, chlorophyll-a concentration exhibits an inverse latitudinal gradient with concentrations ranging from 7.38 mg/m^3 in nutrient-rich zones to 0.04 mg/m^3 in nutrient-poor zones, reaching its peak near the Kerguelen Plateau in austral summer. In 2014, inter-annual variability was observed, with a notable peak in primary productivity in austral spring, compared to the other years (Fig. 3b). The sea ice cover (Fig. 3 c) shows stable seasonal patterns, with slight variability, peaks in austral winter, with ice cover ranging from 0% to almost 50% of the region.

Model selection

Table 1 delineates the hypotheses, model configurations, and AIC values pertinent to each acoustic population: ANT BW, SEIO PBW and SWIO PBW. For all species examined, model M2, characterized by its incorporation of spatial, monthly and annual effects in conjunction with all environmental covariates such as chlorophyll-a concentration, sea surface temperature, and sea ice cover, exhibited the lowest AIC and the highest explanatory power regarding deviation, although anticipated. Among the two tested distributions, the Tweedie distribution demonstrated superior performance in the context of model M2. In addition to AIC-based model selection, cross-validation using Spearman's rank correlation (R_{Spearman}) further supported the superiority of model M2, revealing improved predictive performance for more complex models incorporating temporal and spatial effects. Ultimately, the chosen model represented 21. 2% of the deviation for ANT BW, 57% for SEIO PBW, and 37.6% for SWIO PBW, and was validated through both AIC and cross-validation metrics.

Monthly habitat models and predicted distributions

ANT BW

ANT BW songs were detected at all recording stations, demonstrating a large representation within the study area (Fig. Supplementary 1). Detection rates are significantly higher at the NEAMS site and comparatively lower at the SSEIR and RTJ sites (Fig. Supplementary 1). The covariates that contribute the most significantly to the variance in our model include sea surface temperature, chlorophyll-a concentration, sea ice extent, GPS coordinates, temporal year parameters, and month. These covariates were identified as statistically significant ($p < 0.001$)) (Fig. Supplementary 4). There is a stable positive correlation between chlorophyll-a concentration and acoustic presence, except at very high chlorophyll-a concentrations, where further increases in acoustic presence are observed (Fig. 4a). In particular, this correlation starts with a logarithmic value of log(mean Chla) = -1.8 mg/m^3(Fig. 4a). Seasonal peaks in acoustic events were notably discernible, particularly during austral winter and spring (Fig. 4a). In contrast, there is a negative correlation between temperature and acoustic presence, markedly above 20°C, with a plateau and even a second peak in the range of 10 to 20°C (Fig. 4a). The relationship between sea ice extent and acoustic presence reveals variability at lower percentages of sea ice coverage, with an increase in daily positive minutes (DPM) concurrent with initial ice formation, followed by a secondary positive peak at higher percentages of ice coverage (Fig. 4a). The annual variability in the acoustic presence exhibited significant peaks in the years 2014, 2017, and 2019, with predicted contributions of approximately 0.3, 0.35, and 0.4, respectively, followed by an observable downward trend post-2019 (Fig. 4a). The inter-annual variability in the predicted contributions over the years ranged from around 0.1 in 2010 to a maximum of 0.4 in 2019, subsequently decreasing to approximately 0.2 in 2022.

The predicted areas of acoustic presence extend throughout the study region, with a higher acoustic presence predicted around Kerguelen Island and its plateau, and adjacent to the offshore areas of Madagascar (Fig. 5a). The intra-annual variability in acoustic presence is evident in all four seasons. The highest acoustic presence is observed during austral winter and spring, with predicted contributions ranging between 0.3 and 0.4. In contrast, summer and autumn show significantly lower contributions, ranging between 0 and 0.1 (Fig. 4a). Despite these seasonal fluctuations, the Kerguelen area remains consistently occupied throughout the year, with vocalizations detected in all seasons (Fig. 4a). All predicted DPM maps are available in the supplementary material, from Fig. Supplementary 8 to Fig. Supplementary 20.

SEIO PBW

SEIO PBW are mainly detected acoustically at locations located east of the observation network, specifically at SWAMS and NEAMS, whereas they are nearly absent from northwestern sites such as MAD and southern locations such as ELAN (Fig. Supplementary 2). The covariates that account for the most significant variance in the model include chlorophyll-a concentration, sea surface temperature, sea ice coverage, GPS coordinates, month, and year. All of these covariates are statistically significant ($p < 0.001$) (Fig. Supplementary 4).

Chlorophyll-a concentration exhibits a notable relationship with acoustic presence. Very low chlorophyll-a concentrations (log (mean Chla) = -3 mg/m^3) are associated with a very low acoustic presence, while higher chlorophyll-a concentrations (log (mean Chla) = -0.4 mg/m^3) are associated with a stronger acoustic presence. Between these values, the relationship remains stable (Fig. 4b). Regarding sea surface temperature, a negative relationship with acoustic presence is observed at low temperatures, specifically 0 to 5°C, followed by a rebound and stabilization between 10 and 20°C, and a return to a negative relationship at temperatures exceeding 20°C (Fig. 4b). There is no significant correlation with the percentage of sea ice coverage, except in higher percentages (Fig. 4b). GPS coordinates contribute significantly to spatial variability, with a higher acoustic presence observed in the eastern part of the study area, particularly around the Kerguelen Islands (Fig. 5b).

In terms of temporal variability, there is marked seasonal variation in the acoustic presence. Activity peaks during the austral summer and extends into the autumn (between 0 and 1). The acoustic presence decreases significantly during the winter months, with values dropping to near -2 and remaining low during early spring (-1; Fig. 4b). The inter-annual variability in acoustic presence is evident, with a steady increase observed until 2019 (from approximately -0.7 to 0), followed by a subsequent decline to below -1 in 2022 (Fig. 4b). All predicted DPM maps are available in the supplementary material, from Fig. Supplementary 21 to Fig. Supplementary 33.

SWIO PBW

SWIO PBW are broadly present within the network, but are quasi-absent from extreme northeastern sites like RTJ and extreme southern sites like ELAN (Fig. Supplementary 3). The highest detection rates occur in the western part of the network, particularly at offshore sites near Madagascar, such as MAD, MADE, and MADW (Fig. Supplementary 3).

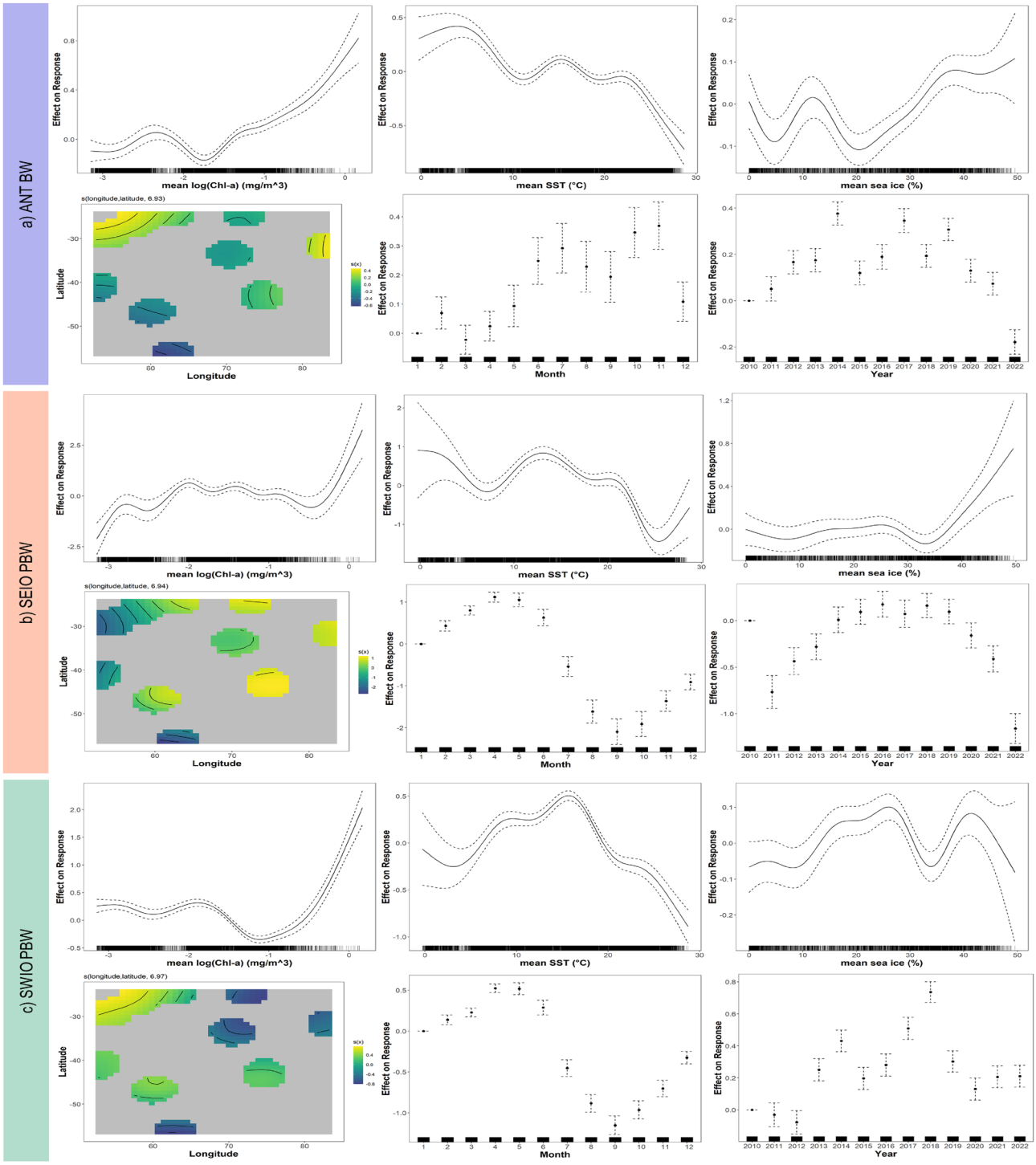


Fig. 4. Functional relationships obtained with generalized additive models between DPM and environmental variables for the three acoustic populations: (**a**) ANT BW (explained deviance = 21.2%), (**b**) SEIO PBW (explained deviance = 57%), and c) SWIO PBW (explained deviance = 37.6%). Solid lines represent the estimated smooth functions and the dashed lines represent the approximate 95% confidence intervals. The x-axis corresponds to the linear predictor and the y-axis to the covariate. Tick marks on the x-axis show the distribution of data points.

The covariates that have the most significant impact on the variance of our model include chlorophyll-a concentration, sea surface temperature, GPS coordinates, month, and year. All of these covariates are statistically significant ($p < 0.001$)) (Fig. Supplementary 4). Chlorophyll-a concentration shows a positive correlation with acoustic presence only at higher concentrations (log (mean Chla) -1.2 mg/m^3) (Fig. 4c). For SST, there is a generally negative relationship with acoustic presence at temperatures near 0°C and a stronger negative

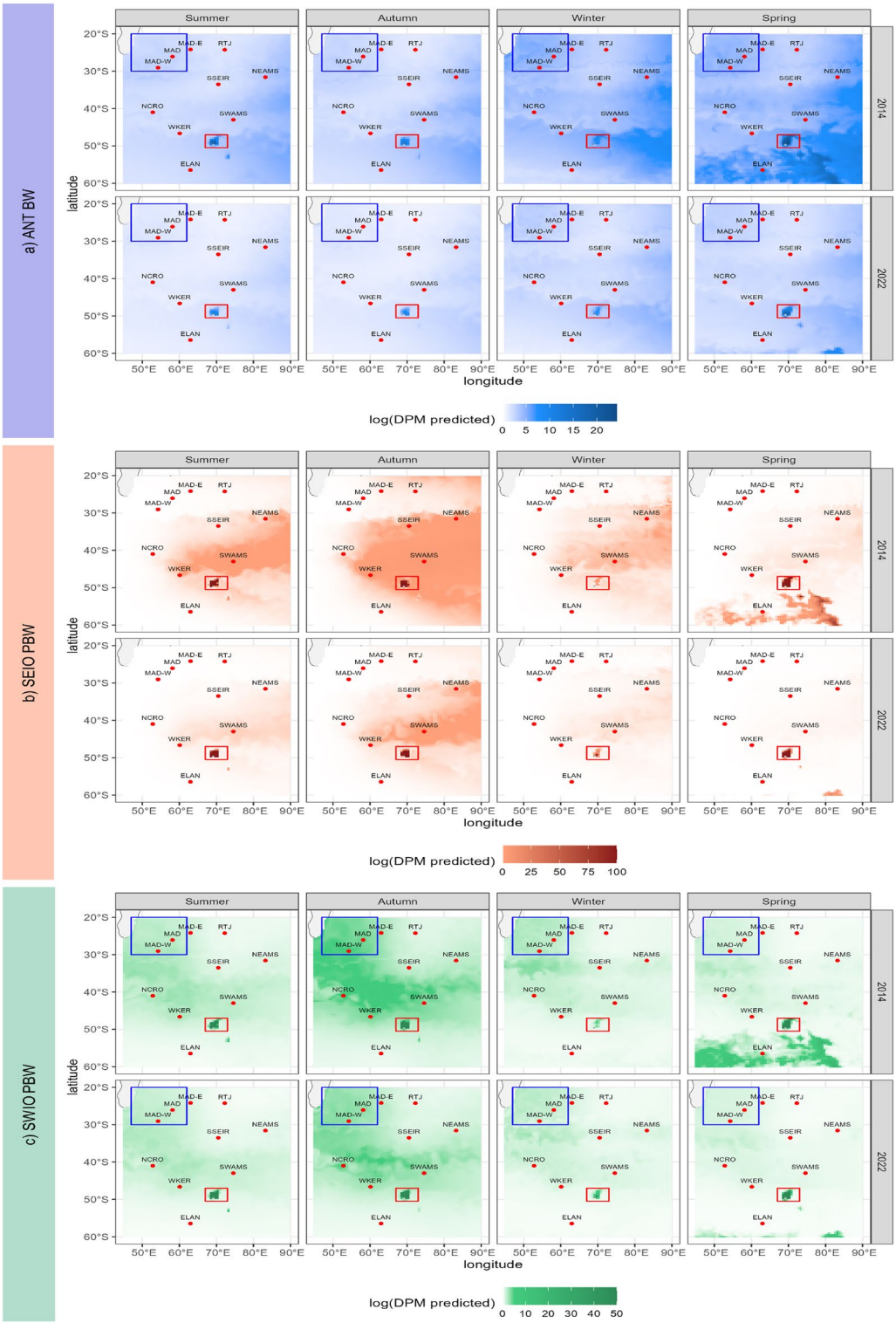


Fig. 5. Predicted log(DPM) distribution for the three acoustic populations: (**a**) ANT BW, (**b**) SEIO PBW, and (**c**) SWIO PBW for the four seasons, represented by January (summer), April (autumn), July (winter), and November (spring), for the years 2014 and 2022. The red square indicates the Kerguelen Island region, and the blue square corresponds to the area east of Madagascar. The color scale is not the same for all acoustic populations. Predictions for each month and year are provided in supplementary materials.

correlation at temperatures exceeding approximately 15°C (Fig. 4c). No significant correlation is observed with the increasing percentage of sea ice coverage (Fig. 4c). The predicted areas of acoustic presence are spread throughout the region, with notably high predictions off Madagascar and around Kerguelen, while the acoustic population is not predicted to be present south of 50°S or in the northeast region (Fig. 5 c). Seasonally, the acoustic population exhibits strong peaks in activity during austral summer and autumn (between 0 and 0.5), with a notable peak in 2018 (approximately 0.8). Since 2019, the acoustic population has shown a more stable pattern (around 0.2) (Fig. 4c). All predicted DPM maps are available in the supplementary material, from Fig. Supplementary 34 to Fig. Supplementary 46.

EOF to identify functional areas

For the three acoustic populations, the first two dimensions of the EOF were selected, which explain more than 60% of the variance for the ANT BW and SWIO PBW acoustic populations, and more than 75% for the SEIO PBW acoustic population. Additional dimensions were excluded from the analysis due to their tendency to represent noise (Fig. Supplementary 6). Each acoustic population exhibits strong spatio-seasonal patterns with succession between a summer/autumn season and a spring/winter season. The cluster analysis highlights the different functional areas predicted for whales during the different seasons for each species (Fig. Supplementary 7).

For the ANT BW subspecies, three temporal clusters and four spatial clusters were predicted (Fig. 6a). Spatial clusters 1 and 4 are predicted for summer and autumn, mainly in the Kerguelen Plateau and higher latitudes. Spatial cluster 2 is predicted for spring and shows limited acoustic activity, concentrated mainly in the southwestern region. Spatial cluster 3 is predicted for winter, predominantly near the coasts or islands within the area.

For SEIO PBW, two temporal clusters and three spatial clusters were predicted (Fig. 6b). Spatial clusters 1 and 2 are predicted for summer and autumn, mainly in the eastern areas surrounding the Kerguelen Plateau. Spatial cluster 3, predicted for winter and spring, suggests that this acoustic population is likely outside the study area.

For SWIO PBW, three temporal clusters and three spatial clusters were predicted (Fig. 6c). In winter, the acoustic population is expected to be mainly located in spatial cluster 3, off Madagascar. In spring, acoustic activity is predicted in clusters 1 and 3, either off Madagascar or beyond the zone. In summer and autumn, the acoustic presence is predicted in clusters 1 and 2, extending from off Madagascar to the Kerguelen Plateau.

Discussion

Using generalized additive models (GAMs) and empirical orthogonal functions (EOFs), we identified key relationships between acoustic presence and environmental variables. ANT BW acoustic activity was strongly associated with low sea surface temperatures, high chlorophyll-a concentrations, and extensive sea ice cover, primarily during austral winter and spring. In contrast, the acoustic populations of SEIO and SWIO PBW showed peak activity in temperate waters (12-15 ° C) with intermediate chlorophyll-a concentrations, revealing different ecological niches. Three critical habitats were identified in the southern Indian Ocean: (1) Madagascar, which serves as a key potential resting area during the austral winter; (2) the Kerguelen Plateau, an essential summer and fall habitat characterized by high primary productivity, which likely supports both feeding and resting activities; and (3) an overlap zone for the SEIO and SWIO PBW acoustic populations in summer, suggesting possible interactions such as hybridization or competition. Our approach combines the flexibility of GAMs with the dimensionality reduction power of EOFs. These methods, coupled with the extensive spatial and temporal scales of the dataset, provide novel insights into the migratory patterns, habitat use, and ecological drivers of blue whale distributions[57,58].

A principal advantage of our approach is its ability to predict the acoustic presence of whales in unsampled regions, offering a comprehensive understanding of their distribution over vast marine areas. We used the number of positive minutes per day, averaged weekly, metric to effectively capture inter-annual and seasonal variations of blue whale acoustic presence over a dataset with extensive spatial and temporal coverage. The incorporation of EOFs further improved spatio-temporal dimensionality reduction, providing novel insights into critical habitats and seasonal dynamics of blue whale acoustic populations.

The results of our model align with the established knowledge on the ANT BW acoustic population. These cetaceans are predominantly detected during austral winter and spring, with peak acoustic activity correlated with low temperatures, high chlorophyll-a concentrations, and extensive sea ice cover. The spatial predictions of the model, reflecting the temporal variability, corroborate known migratory behaviors, where the ANT BW feed in the southern ocean during the summer and migrate to lower latitudes during the winter months[2,3,5,20]. Furthermore, inter-annual variability in detection patterns is evident, with a notable decline in acoustic presence after 2019. This fluctuation could be linked to changes in environmental conditions, such as variations in sea ice extent and primary productivity, or anthropogenic factors, suggesting dynamic habitat use and potential shifts in blue whale distribution over time.

These findings are consistent with previous research that emphasizes the importance of sea ice and the availability of prey in the distribution of ANT BW. The positive correlation between acoustic detections and sea ice extent supports the results of several studies[29,30]. Furthermore, the prominence of chlorophyll-a as a proxy of prey availability is consistent with observations in similarly productive areas such as the Northeast Pacific and Northern Patagonia[59,60].

In particular, our results suggest that the eastern waters of Madagascar may serve as a significant resting area for ANT BW during austral winter. These regions, characterized by moderate temperatures and productivity, are mainly visited during this season, indicating their potential role as temporary refuges or resting areas. The Kerguelen Plateau appears to play a prominent role during the austral summer and autumn, where a significant acoustic presence of blue whales is observed. This area may function both as a resting zone before migration

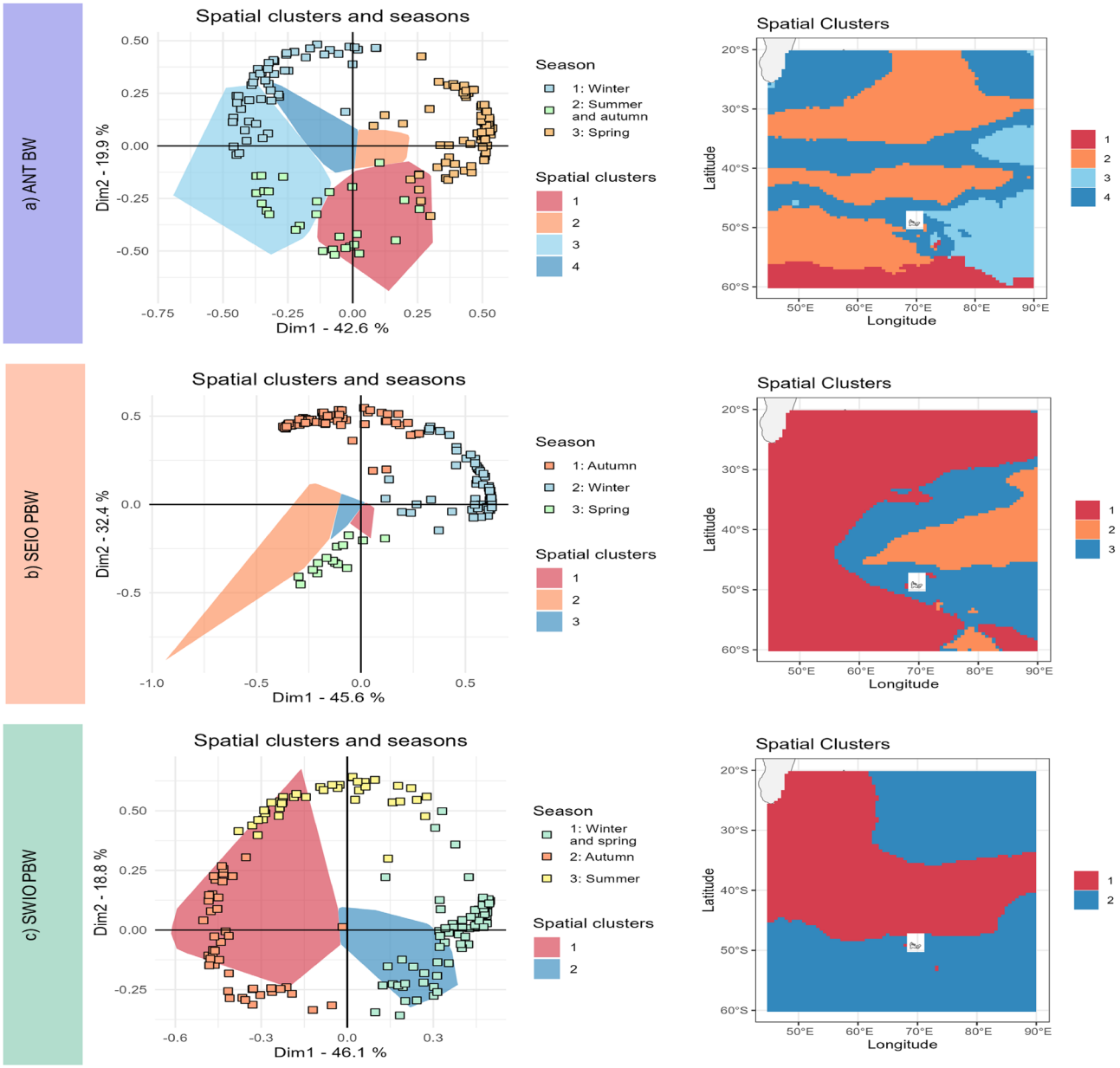


Fig. 6. Functional areas and associated seasons for the three acoustic populations: (**a**) ANT BW, (**b**) SEIO PBW, and (**c**) SWIO PBW. (Left column) Projection of the standardized temporal loadings (seasons) and the standardized spatial factors (spatial clusters) onto the first two dimensions of the EOF. (Right column) Mapping of spatial clusters using standardized spatial factors to identify functional areas. The same colors are used to represent spatial clusters on the maps (right column) and on the first two dimensions of the EOF (polygons in the left column). In the left column, each season is represented by a square with a specific color. If the loadings of a season are in the same direction as a spatial cluster, then they are correlated and correspond to a couple season x functional area.

to higher latitudes to feed in summer and as a feeding site for non-migrating individuals[5,61]. Its substantial primary productivity, particularly in krill abundance[62], supports this hypothesis, highlighting the importance of Kerguelen as a critical habitat for blue whale acoustic populations.

SEIO and SWIO PBW exhibit different spatial distribution patterns throughout the year. However, these acoustic populations were recorded simultaneously during the austral summer and autumn, with maximum activity observed in regions characterized by intermediate temperatures (12–15 °C) and high chlorophyll-a concentrations. This spatial and temporal overlap supports hypotheses from previous studies, which suggest potential interactions between these two acoustic populations, such as hybridization or competition for food resources[63]. EOF analysis further improves understanding of habitat use by identifying spatio-temporal patterns consistent with those described in other dynamic models[64,65].

In contrast to ANT BW, pygmy blue whales show a minimal association with sea ice, which has limited influence on the predictions in our model. This observation is consistent with studies of SEIO PBW acoustic populations, which do not migrate as far south to feed. Satellite telemetry studies have shown that these whales are not detected south of 42°S, which confirms their tendency to avoid regions with extensive sea ice cover[66,67]. The hypothesis of geographical separation within the SEIO PBW acoustic population is supported by recent studies documenting the presence of SEIO PBW in the Kerguelen Plateau and southern Australian waters during the same austral summer and autumn season[21,22]. Acoustic data from Kerguelen and southern Australia overlap seasonally, reinforcing the idea that the Kerguelen Plateau is an important feeding ground for these whales, especially given its high primary productivity. Satellite tracking further supports this, showing that SEIO PBW individuals migrate south from Indonesia to Kerguelen for feeding but also travel to southern Australia, indicating a potential division within the population into two distinct migratory routes[66,67]. These observations suggest that the SEIO PBW population may follow different migratory routes, with migrations to Kerguelen driven by foraging rather than reproduction, as these whales are known to reproduce in Indonesian waters. Although Kerguelen's abundant krill[62] makes it a critical food source, separation into distinct migratory routes and ecological zones may not be fixed.

Similarly, SWIO PBWs disperse as they migrate across the southwestern Indian Ocean. Their range extends from the south of Marion and Prince Edward Islands, where acoustic detections have been reported between December and April[68] to the south of Amsterdam Island in the east, with detections occurring from December to July[7]. The region surrounding Prince Edward Island has been identified as a significant feeding ground during this period[68]. The concomitant recording of SWIO PBW songs near Kerguelen in our study posits the potential for Kerguelen to also function as a feeding site. This suggests that there may be a spatial dispersion within the SWIO PBW population during summer - mid-fall, with individuals following different migratory pathways while exploiting similar ecological resources. Northward migration begins in April, with individuals traveling toward Madagascar, where the migratory route splits: one branch passes through the Mozambique Channel[24], while the other skirts offshore around Réunion and Mauritius[7]. The northernmost acoustic detections have been recorded near Mayotte between May and July[25].However, all acoustic monitoring stations across the Indian Ocean report an absence of calls from August to October[6,7,24,68], which may indicate a shift in vocal behavior or a seasonal movement of the population to a more northerly, equatorial region that remains unsurveyed. Migration to the central and eastern tropical Indian Ocean is unlikely, as no calls have been detected around the Seychelles or Diego Garcia[69,70]. During their return south (October to January), SWIO PBW calls are not detected by offshore acoustic stations in the western Indian Ocean[7]. However, they are consistently detected along the western and eastern coasts of Madagascar[6,24,25], suggesting that coastal waters provide a more favorable habitat during this part of the migratory cycle.

Our results further advance current knowledge by emphasizing the role of the Kerguelen Plateau for pygmy blue whale acoustic populations as a shared habitat during austral summer and autumn. This area, with its high primary productivity and abundance of krills[62], is likely to support feeding and resting activities for these acoustic populations. Identifying key habitats underscores the ecological importance of these areas for the conservation of blue whales. The region serves as a biodiversity hotspot, supporting not only blue whales but also other predators such as elephant seals and penguins[71,72].

This study has certain limitations to consider. First, ambient noise, which can affect call detection, was not explicitly taken into account in our model. High levels of noise from natural or anthropogenic sources could mask whale calls and lead to underestimations of acoustic activity[59,64]. Including ambient noise as a covariate in future studies would improve the accuracy of the model. To address these limitations, we estimate a detection range of 50 km around the hydrophones to account for sound propagation and minimize bias. An estimated detection range of 50 km was determined as a conservative figure, reflecting the substantial variability in the detection ranges documented in the literature. These ranges exhibit substantial variation, ranging from 10 km to more than 700 km, depending on factors such as geographical region, seasonal variations, and source levels between acoustic populations[68,73,74].

Secondly, this analysis intentionally excluded D-calls, which represent short acoustic signals typically associated with foraging activities and are believed to aid in prey detection or coordination during feeding[9]. These vocalizations were omitted because of their inability to be reliably attributed to a distinct acoustic population. However, the inclusion of D-calls in future research could provide deeper insights into the various behavioral contexts of blue whale vocalization, especially in highly productive areas such as the Kerguelen Plateau. This proposition is corroborated by Torterotot et al. (2023)[75], who reported high occurrence of D-calls near the Kerguelen Plateau. In contrast, the blue whale songs examined in this study are predominantly associated with reproductive activities, similar to those observed in humpback whales[13,76]. However, their presence in feeding regions[77,78] implies that they may perform additional functions.

Third, we used chlorophyll-a concentration as a proxy for prey availability, which may not directly reflect krill biomass, the primary food source for blue whales. This limitation stems from the difficulty of obtaining direct measurements of the biomass of krill in such vast regions and time frames. Future work should incorporate direct measurements of krill biomass, such as acoustic surveys or satellite data, to better capture prey dynamics and refine habitat models.

In addition, we focus on broad-scale variables such as SST, chlorophyll-a, and sea ice extent. Although this approach is ecologically relevant on large scales, finer-scale factors such as nutrient upwelling or mesoscale eddies could improve local habitat models and provide deeper ecological insights.

In conclusion, this study sheds light on the complex relationships between the acoustic populations of blue whales and their environment. By identifying functional zones shaped by environmental factors, it establishes a foundation for targeted conservation efforts to preserve these critical marine ecosystems.

Data availability

The environmental data used in this study are publicly available from dedicated databases: Sea Surface Temperature (SST): Data were obtained from the Copernicus Marine global ocean reanalysis (DOI: 10.48670/moi-00021), covering the period 1992 to 2024 with a spatial resolution of 1/12° and daily temporal resolution. Chlorophyll-a concentration: Data were sourced from Mercator-Ocean (DOI: 10.48670/moi-00019), available from January 1993 to February 2024 at a 1/4° resolution. Sea Ice Concentration: Data were obtained from the NOAA/NSIDC Climate Data Record, Version 4 (DOI: 10.7265/efmz-2t65), based on passive microwave measurements and available at a 25 km × 25 km resolution with daily and monthly temporal scales. These environmental data were used to generate Fig. 3 in the manuscript and were included as explanatory variables in the models used throughout the study. The acoustic data used in this study (DOI: 10.18142/229) were used to generate Fig. 2 and served as the response variable in our model. The raw acoustic data is not publicly available, but the detection data can be obtained from Jean-Yves Royer (jean-yves.royer@univ-brest.fr) upon reasonable request. Additional data sets were not generated or analyzed beyond those specified.

Received: 3 February 2025; Accepted: 16 May 2025
Published online: 02 July 2025

References

1. IWC. Chairman's report of the thirty-first annual meeting. *Rep. int. Whal. Commn.* vol. 30, 25–41 (1980).
2. Samaran, F., Adam, O. & Guinet, C. Discovery of a mid-latitude sympatric area for two southern hemisphere blue whale subspecies. *Endanger. Species Res.* **12**, 157–165 (2010).
3. Samaran, F. et al. Seasonal and geographic variation of southern blue whale subspecies in the Indian ocean. *PLoS ONE* **8**, e71561 (2013).
4. Stafford, K. M., Chapp, E., Bohnenstiel, D. R. & Tolstoy, M. Seasonal detection of three types of "pygmy" blue whale calls in the Indian ocean. *Mar. Mamm. Sci.* **27**, 828–840 (2011).
5. Leroy, E. C., Samaran, F., Bonnel, J. & Royer, J.-Y. Seasonal and diel vocalization patterns of Antarctic blue whale (*Balaenoptera musculus intermedia*) in the southern Indian ocean: A multi-year and multi-site study. *PLoS ONE* **11**, e0163587 (2016).
6. Dréo, R., Bouffaut, L., Leroy, E., Barruol, G. & Samaran, F. Baleen whale distribution and seasonal occurrence revealed by an ocean bottom seismometer network in the western Indian ocean. *Deep Sea Res. Part II* **161**, 132–144 (2019).
7. Torterotot, M., Samaran, F., Stafford, K. M. & Royer, J.-Y. Distribution of blue whale populations in the southern Indian ocean based on a decade of acoustic monitoring. *Deep Sea Res. Part II* **179**, 104874 (2020).
8. Cummings, W. C. & Thompson, P. O. Underwater sounds from the blue whale, *Balaenoptera musculus*. *J. Acoust. Soc. Am.* **50**, 1193–1198 (1971).
9. Mellinger, D. K. & Clark, C. W. Blue whale (*Balaenoptera musculus*) sounds from the north Atlantic. *J. Acoust. Soc. Am.* **114**, 1108–1119 (2003).
10. Thompson, P. O., Findley, L. T., Vidal, O. & Cummings, W. C. Underwater sounds of blue whales, *Balaenoptera musculus*, in the gulf of California, Mexico. *Mar. Mamm. Sci.* **12**, 288–293 (1996).
11. Thode, A. M., D'Spain, G. L. & Kuperman, W. A. Physical constraints on blue whale sound production mechanisms derived from matched-field processing. *J. Acoust. Soc. Am.* **105**, 993–993 (1999).
12. Lewis, L. A. et al. Context-dependent variability in blue whale acoustic behaviour. *R. Soc. Open Sci.* **5**, 180241 (2018).
13. Oleson, E. M., Calambokidis, J., Barlow, J. & Hildebrand, J. A. Blue whale visual and acoustic encounter rates in the southern California bight. *Mar. Mamm. Sci.* **23**, 574–597 (2007).
14. McDonald, M. A., Mesnick, S. L. & Hildebrand, J. A. Biogeographic characterization of blue whale song worldwide: Using song to identify populations. *J. Cetac. Res. Manag.* **8**, 55–65 (2006).
15. Leroy, E. C., Samaran, F., Stafford, K. M., Bonnel, J. & Royer, J.-Y. Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the southern Indian ocean. *Endanger. Species Res.* **37**, 289–300 (2018).
16. Leroy, E. C., Royer, J.-Y., Bonnel, J. & Samaran, F. Long-term and seasonal changes of large whale call frequency in the southern Indian ocean. *J. Geophys. Res. Oceans* **123**, 8568–8580 (2018).
17. Miller, B. S. et al. Antarctic sonobuoy surveys for blue whales from 2006–2021 reveal contemporary distribution, changes over time, and paths to further our understanding of their distribution and biology. *Front. Mar. Sci.* **11**, 1324816 (2024).
18. Davis, G. E. et al. Exploring movement patterns and changing distributions of baleen whales in the western north Atlantic using a decade of passive acoustic data. *Glob. Change Biol.* **26**, 4812–4840 (2020).
19. Ichihara, T. The pygmy blue whale, *Balaenoptera musculus brevicauda*, a new subspecies from the Antarctic. *Whales, dolphins, and porpoises* 79–111 (1966).
20. Stafford, K. M. et al. Antarctic-type blue whale calls recorded at low latitudes in the Indian and eastern Pacific oceans. *Deep Sea Res. Part I* **51**, 1337–1346 (2004).
21. Gavrilov, A. N. & McCauley, R. D. Acoustic detection and long-term monitoring of pygmy blue whales over the continental slope in southwest Australia. *J. Acoust. Soc. Am.* **134**, 2505–2513 (2013).
22. McCauley, R. D., Gavrilov, A. N., Jolliffe, C. D., Ward, R. & Gill, P. C. Pygmy blue and Antarctic blue whale presence, distribution and population parameters in southern Australia based on passive acoustics. *Deep Sea Res. Part II* **157**, 154–168 (2018).
23. Ljungblad, D., Clark, C. W. & Shimada, H. Sc/49/sh17 a comparison of sounds attributed to pygmy blue whales (*Balaenoptera musculus brevicauda*) recorded south of the Madagascar plateau and those attributed to 'true' blue whales (*Balaenoptera musculus*) recorded off Antarctica. *Rep.-Int. Whaling Commiss.* **48**, 439–442 (1998).
24. Cerchio, S., Rasoloarijao, T., Mueller-Brennan, B. & Cholewiak, D. Migratory movements and acoustic activity of blue whales (*Balaenoptera musculus*) and fin whales (*Balaenoptera physalus*) in the eastern Mozambique channel. *Technical report* (2024).
25. Dréo, R. *et al.* Singing around the volcano: Detecting baleen whales in the Mozambique channel based on their song rhythms, from seismic and hydroacoustic data. *J. Acoust. Soc. Am.* (2025).
26. Virgili, A. et al. Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: Large-scale modelling of deep-diving cetacean habitats. *Glob. Ecol. Biogeogr.* **28**, 300–314 (2019).
27. Lambert, C., Mannocci, L., Lehodey, P. & Ridoux, V. Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions. *PLoS ONE* **9**, e105958 (2014).
28. Nicol, S., Worby, A. & Leaper, R. Changes in the Antarctic sea ice ecosystem: Potential effects on krill and baleen whales. *Mar. Freshw. Res.* **59**, 361–382 (2008).
29. Shabangu, F. W., Yemane, D., Stafford, K. M., Ensor, P. & Findlay, K. P. Modelling the effects of environmental conditions on the acoustic occurrence and behaviour of Antarctic blue whales. *PLoS ONE* **12**, e0172705 (2017).
30. Shabangu, F. W. et al. Seasonal occurrence and diel calling behaviour of Antarctic blue whales and fin whales in relation to environmental conditions off the west coast of South Africa. *J. Mar. Syst.* **190**, 25–39 (2019).

31. Belkin, I. M. & Gordon, A. L. Southern ocean fronts from the Greenwich meridian to Tasmania. *J. Geophys. Res. Oceans* **101**, 3675–3696 (1996).
32. Orsi, A. H., Whitworth, T. III. & Nowlin, W. D. Jr. On the meridional extent and fronts of the Antarctic circumpolar current. *Deep Sea Res. Part I* **42**, 641–673 (1995).
33. Behrenfeld, M. J. et al. Climate-driven trends in contemporary ocean productivity. *Nature* **444**, 752–755 (2006).
34. Smith, W. H. & Sandwell, D. T. Global sea floor topography from satellite altimetry and ship depth soundings. *Science* **277**, 1956–1962 (1997).
35. Moore, J. K. & Abbott, M. R. Surface chlorophyll concentrations in relation to the Antarctic polar front: seasonal and spatial patterns from satellite observations. *J. Mar. Syst.* **37**, 69–86 (2002).
36. Royer, J.-Y. Oha-sis-bio-observatoire hydroacoustique, https://doi.org/10.18142/229 (2009).
37. Lutron, X. An introduction to underwater acoustics (2002).
38. d'Eu, J.-F., Royer, J.-Y. & Perrot, J. Long-term autonomous hydrophones for large-scale hydroacoustic monitoring of the oceans. In *2012 Oceans-Yeosu*, 1–6 (IEEE, 2012).
39. Socheleau, F.-X. & Samaran, F. Detection of mysticete calls: a sparse representation-based approach. *Technical report* (2018).
40. Guilment, T., Socheleau, F.-X., Pastor, D. & Vallez, S. Sparse representation-based classification of mysticete calls. *J. Acoust. Soc. Am.* **144**, 1550–1563 (2018).
41. Torterotot, M., Royer, J.-Y. & Samaran, F. Detection strategy for long-term acoustic monitoring of blue whale stereotyped and non-stereotyped calls in the southern indian ocean. In *OCEANS 2019-Marseille*, 1–10 (IEEE, 2019).
42. Godet, C. et al. Matching zooplankton abundance and environment in the south Indian ocean and southern ocean. *Deep Sea Res. Part I* **163**, 103347 (2020).
43. Širović, A. et al. Seasonality of blue and fin whale calls and the influence of sea ice in the western Antarctic peninsula. *Deep Sea Res. Part II* **51**, 2327–2344 (2004).
44. Meier, W. N. et al. Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity. *Rev. Geophys.* **52**, 185–217 (2014).
45. Meier, W. N., Fetterer, F., Windnagel, A. K. & Stewart, J. S. Noaa/nsidc climate data record of passive microwave sea ice concentration. (g02202, version 4). (National Snow and Ice Data Center, 2021).
46. Wood, S. N. *Generalized Additive Models: An Introduction with R* (Chapman and Hall/CRC, 2017).
47. Arcuti, S. et al. Spatio-temporal modelling of zero-inflated deep-sea shrimp data by tweedie generalized additive. *Statistica (Bologna)* **73**, 87–101 (2013).
48. Cressie, N., Sainsbury-Dale, M. & Zammit-Mangion, A. Basis-function models in spatial statistics. *Annu. Rev. Stat. Appl.* **9**, 373–400 (2022).
49. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723 (1974).
50. Wood, S. *Generalized Additive Models: An Introduction with R* 2 edn (Chapman and Hall/CRC, 2017).
51. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, 2021).
52. Thorson, J. T. et al. Empirical orthogonal function regression: Linking population biology to spatial varying environmental conditions using climate projections. *Glob. Change Biol.* **26**, 4638–4649 (2020).
53. Alglave, B. et al. Investigating fish reproduction phenology and essential habitats by identifying the main spatio-temporal patterns of fish distribution. *ICES J. Mar. Sci.* **81**, 1563–1574 (2024).
54. Thorson, J. T., Ciannelli, L. & Litzow, M. A. Defining indices of ecosystem variability using biological samples of fish communities: a generalization of empirical orthogonal functions. *Prog. Oceanogr.* **181**, 102244 (2020).
55. Alglave, B., Dufée, B., Obakrim, S. & Thorson, J. T. Empirical Orthogonal Functions for ecology (2024). Preprint.
56. Lê, S., Josse, J. & Husson, F. FactoMineR: A package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18, https://doi.org/10.18637/jss.v025.i01 (2008).
57. Monnahan, C. C., Branch, T. A., Stafford, K. M., Ivashchenko, Y. V. & Oleson, E. M. Estimating historical eastern north pacific blue whale catches using spatial calling patterns. *PLoS ONE* **9**, e98974 (2014).
58. Anderson, S. C., Ward, E. J., English, P. A. & Barnett, L. A. sdmtmb: an r package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *BioRxiv* 2022–03 (2022).
59. Burtenshaw, J. C. et al. Acoustic and satellite remote sensing of blue whale seasonality and habitat in the northeast pacific. *Deep Sea Res. Part II* **51**, 967–986 (2004).
60. Buchan, S. J. & Quiñones, R. A. First insights into the oceanographic characteristics of a blue whale feeding ground in northern Patagonia, Chile. *Mar. Ecol. Prog. Ser.* **554**, 183–199 (2016).
61. Branch, T. A. et al. Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the southern hemisphere and northern Indian ocean. *Mammal Rev.* **37**, 116–175 (2007).
62. Hunt, B. & Swadling, K. Macrozooplankton and micronekton community structure and diel vertical migration in the heard island region, central Kerguelen plateau. *J. Mar. Syst.* **221**, 103575 (2021).
63. Attard, C. R. et al. Hybridization of southern hemisphere blue whale subspecies and a sympatric area off Antarctica: Impacts of whaling or climate change?. *Mol. Ecol.* **21**, 5715–5727 (2012).
64. Barlow, D. R. & Torres, L. G. Planning ahead: dynamic models forecast blue whale distribution with applications for spatial management. *J. Appl. Ecol.* **58**, 2493–2504 (2021).
65. Abrahms, B. et al. Memory and resource tracking drive blue whale migrations. *Proc. Natl. Acad. Sci.* **116**, 5582–5587 (2019).
66. Double, M. C. et al. Migratory movements of pygmy blue whales (*Balaenoptera musculus brevicauda*) between Australia and Indonesia as revealed by satellite telemetry. *PLoS ONE* **9**, e93578 (2014).
67. Mustika, P. L. K. *et al.* The first record of the southbound movements of satellite-tagged pygmy blue whales (*B. m. brevicauda*) from Savu Sea (Indonesia) to the subantarctic waters. *Mar. Mamm. Sci.* e13167 (2024).
68. Shabangu, F. W. et al. Diverse baleen whale acoustic occurrence around two sub-antarctic islands: A tale of residents and visitors. *Sci. Rep.* **14**, 21663 (2024).
69. Stafford, K. M. et al. Acoustic detections and sightings of blue whales *Balaenoptera musculus* in the Seychelles, western tropical Indian ocean (2020–2022). *Endanger. Species Res.* **52**, 203–208 (2023).
70. Leroy, E. C., Royer, J.-Y., Alling, A., Maslen, B. & Rogers, T. L. Multiple pygmy blue whale acoustic populations in the Indian ocean: Whale song identifies a possible new population. *Sci. Rep.* **11**, 8762 (2021).
71. Toole, M. O., Guinet, C., Lea, M.-A. & Hindell, M. A. Marine predators and phytoplankton: how elephant seals use the recurrent Kerguelen plume. *Mar. Ecol. Prog. Ser.* **581**, 215–227 (2017).
72. Hindell, M. A. *et al.* Foraging habitats of top predators, and areas of ecological significance, on the Kerguelen plateau. *The Kerguelen Plateau Mar. Ecosyst. Fish.* 203–215 (2011).
73. Samaran, F., Guinet, C., Adam, O., Motsch, J.-F. & Cansi, Y. Source level estimation of two blue whale subspecies in southwestern Indian ocean. *J. Acoust. Soc. Am.* **127**, 3800–3808 (2010).
74. Širović, A., Hildebrand, J. A. & Wiggins, S. M. Blue and fin whale call source levels and propagation range in the southern ocean. *J. Acoust. Soc. Am.* **122**, 1208–1215 (2007).
75. Torterotot, M., Samaran, F. & Royer, J.-Y. Long-term acoustic monitoring of nonstereotyped blue whale calls in the southern indian ocean. *Mar. Mamm. Sci.* **39**, 594–610 (2023).
76. McDonald, M. A., Calambokidis, J., Teranishi, A. M. & Hildebrand, J. A. The acoustic calls of blue whales off California with gender data. *J. Acoust. Soc. Am.* **109**, 1728–1735 (2001).

77. Širović, A., Hildebrand, J. A., Wiggins, S. M. & Thiele, D. Blue and fin whale acoustic presence around Antarctica during 2003 and 2004. *Mar. Mamm. Sci.* **25**, 125–136 (2009).
78. Thomisch, K. et al. Spatio-temporal patterns in acoustic presence and distribution of Antarctic blue whales *Balaenoptera musculus* intermedia in the Weddell sea. *Endanger. Species Res.* **30**, 239–253 (2016).

Acknowledgements
This project was funded by the ANR (grant ESDIR2023). The authors wish to thank the captains and crews of the Marion Dufresne R/V for the successful deployments and recoveries of the instruments of the OHA-SIS-BIO array. The cruises were funded by the French Polar Institute and the French Oceanographic Fleet with additional support from Institut National des Sciences de l'Univers Centre National de la Recherche Scientifique.

Author contributions
Mathilde Michel: Conceptualization, Methodology, Formal Analysis, Original Draft Writing, Visualization. Baptiste Algave: Conceptualization, Methodology, Validation, Formal Analysis, Writing-Review & Editing. Maxime Olmos: Conceptualization, Methodology, Validation, Formal Analysis, Writing-Review & Editing. Maëlle Torterotot: Validation, Formal Analysis, Resources, Data Curation, Writing-Review & Editing. Auriane Virgili: Validation, Writing-Review & Editing. Salomé Martin-Marin: Formal Analysis, Resources. Jean-Yves Royer: Writing-Review & Editing, Project Administration, Funding Acquisition. Flore Samaran: Conceptualization, Validation, Writing-Review & Editing, Project Administration, Funding Acquisition.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-02941-9.

Correspondence and requests for materials should be addressed to M.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2025

2.4 Ecological habitat characterization of baleen whales east of Mayotte : linking seasonal acoustic detections with environmental drivers

Ecological habitat characterization of baleen whales in the northern Mozambique Channel: linking seasonal acoustic detections with environmental drivers

Mathilde Michel,[1] Auriane Virgili,[2] Richard Dréo,[3,4] Baptiste Alglave,[5] Maëlle Torterotot,[1] Sara Bazin,[6] and Flore Samaran[1]

[1] *CNRS Lab-STICC, ENSTA, IP Paris, Brest, 29200, France*

[2] *Share the Ocean, Larmor-Baden, 56870, France*

[3] *Institut de physique du globe de Paris, Université de Paris, CNRS, Paris, 75005, France*

[4] *SAS Boksound, Kernevez, Rosnoën, 20590, France*

[5] *CNRS Lab-STICC, Université Bretagne Sud, Vannes, 56000, France*

[6] *Geo-Ocean, Univ Brest, CNRS, Ifremer, IUEM, UMR6538, Plouzané, 29280, France*

(Dated: 7 October 2025)

Baleen whales (*Mysticeti*) play vital ecological roles in marine ecosystems, but their habitat use in the southwest Indian Ocean remains poorly understood due to logistical constraints and low population densities. This study analyzes five years of passive acoustic monitoring data collected by the REVOSIMA (volcanological and seismological observatory of Mayotte) in the northern Mozambique Channel near Mayotte Island. It focuses on three species of baleen whales, including two subspecies of blue whales: the Antarctic blue whale (ANT BW) (*Balaenoptera musculus intermedia*), the Southwestern Indian Ocean pygmy blue whale (SWIO PBW) (*B. m. brevicauda*), the Fin whale (FW) (*B. physalus*) and the Antarctic minke whale (AMW) (*B. bonaerensis*). The acoustic presence was quantified using Daily positive hour (DPH), and Generalized additive models (GAM) assessed the relationships between DPH and environmental variables, Sea surface temperature (SST), Sea surface temperature anomaly (SST anomaly), Chlorophyll-a (Chl) concentration, and Significant wave height (SWH).
The results reveal species-specific habitat use strategies. ANT BW and AMW use the area as a stable migratory resting habitat during austral winter and spring; SWIO PBW exhibit bimodal seasonal peaks, consistent with migratory transit; FW shows sporadic occurrence linked to thermal preferences. SST, SST anomaly, and SWH emerged as key environmental drivers, while Chl showed no consistent effect, indicating limited local feeding activity.
These findings position the northern Mozambique Channel as a multipurpose habitat that supports various ecological strategies among baleen whales. Using long-term acoustic data from a seismological observatory, this study highlights the value of interdisciplinary monitoring in data-limited tropical environments. The results provide essential baseline knowledge to guide conservation and management within the Indian Ocean Whale Sanctuary, especially under climate change and increasing anthropogenic pressures.

[https://doi.org(DOI number)]

[XYZ] Pages: 1–21

I. INTRODUCTION

Baleen whales (*Mysticeti*) fulfill essential ecological functions within marine ecosystems, such as the recycling of nutrients through the "whale pump" and the regulation of zooplankton populations[1]. However, critical knowledge gaps remain regarding the use of baleen whale habitat in the southwest Indian Ocean, due to the vast and difficult-to-access marine environment[2,3], low population densities following extensive whaling[4,5], and the logistical challenges and costs of sustained visual surveys over the open ocean. Although French aerial surveys have significantly improved the understanding of main predator habitats[6,7], these efforts provide only seasonal snapshots, mainly during austral summer, when the presence of whales is low. Consequently, our knowledge of baleen whale ecology has traditionally relied on fragmented observations from whaling records and sporadic ship-based surveys[8,9].

The advent of passive acoustic monitoring (PAM) has significantly advanced cetacean research by leveraging two important advantages: the efficient propagation of low-frequency sounds in seawater and the species-specific vocal signatures of baleen whales[10,11]. This methodological approach is particularly valuable for monitoring rare and widely distributed baleen whale species, whose powerful low-frequency calls (below 100

Hz) can travel up to around 100 kilometers underwater, making them detectable via seismic acoustic monitoring recorders[12,13].

Only two studies have specifically addressed the distribution of baleen whales in waters east of Mayotte and north of the Mozambique Channel, in the southwest Indian Ocean[14,15]. Among them, Dréo et al. (2025)[15] notably detected and characterized the seasonal acoustic presence of three species of baleen whales, including two subspecies of blue whales, in this region. These are the Antarctic blue whale (ANT BW) (*Balaenoptera musculus intermedia*), the Southwestern Indian Ocean pygmy blue whale (SWIO PBW) (*B. m. brevicauda*), the Fin whale (FW) (*B. physalus*), and the Antarctic minke whale (AMW) (*B. bonaerensis*). Each species exhibits specific patterns of habitat use and migration. The ANT BW appear to reach the northern limit of their migratory range in this region. Seasonal detections of their characteristic 18-26 Hz Z calls during the austral winter are indicative of their movements from the feeding grounds of the Southern Ocean to tropical or subtropical areas[11,16,17]. The SWIO PBW is recognized as a transiting population, identified by unique call sequences of 15 to 35 Hz that show bimodal seasonal peaks (early austral summer and early austral winter), likely reflecting migrations between subantarctic feeding grounds and poorly defined tropical breeding areas[14,18,19]. The FW are intermittently detected in the northern Mozambique Channel, mainly during late austral winter, and are also recorded year-round around the Prince Edward Islands, with a peak in detections during austral winter[14,20], indicating their use of a broad area throughout the southwest Indian Ocean. The AMW, identified by their "bioduck" vocalizations (50 to 300 Hz), shows sparse acoustic detections and limited visual observations in these subtropical waters. Consequently, current data are insufficient for a comprehensive understanding of their habitat use and migratory behavior in the region[21]. In general, these observations highlight significant knowledge gaps about the spatial ecology and seasonal dynamics of baleen whales in the southwest Indian Ocean[12,22].

Building on the work of Dréo et al. (2025)[15], who detected and characterized the seasonal acoustic presence of these baleen whale species in the region, this study aims to better understand the environmental drivers of their acoustic presence. Specifically, we examined the relationships between the acoustic presence of whales and four environmental variables: Sea surface temperature (SST) as an indicator of thermal preference[23]; Sea surface temperature anomaly (SST anomaly), reflecting long-term environmental variability; Chlorophyll-a (Chl), a proxy of primary productivity[24,25]; and Significant wave height (SWH), a proxy for ocean surface dynamics, including wind-driven turbulence and variability of the sea state[26]. These covariates are known to influence the distribution of baleen whales in other ocean regions[20,22,27,28].

A quantitative habitat model for baleen whales is developed for the north of the Mozambique Channel, based on five years of continuous passive acoustic data collected from 2020 to 2024 by the REVOSIMA network. Originally established to monitor potentially dangerous submarine volcanic activity in the region, REVOSIMA involved successive short-term deployments (3-12 months) of ocean-bottom seismometers and moored hydrophones in the water column (MAHY network) during the MAYOBS cruises[29–31]. Although not originally intended for biological monitoring, the acoustic dataset provides a valuable opportunity to investigate the presence and use of baleen whale habitat in this key area of the southwest Indian Ocean. The acoustic presence of whales was quantified using Daily positive hour (DPH), defined as the number of hours per day with confirmed acoustic detections. This metric provides a robust way to characterize the temporal variations in the vocal activity of whales. Generalized additive models (GAM)[32] were used to analyze the relationships between DPH and environmental predictors. This approach allows for (i) the identification of the main environmental drivers of whale acoustic presence, (ii) the description of non-linear responses, and (iii) the assessment of the functional role of the area. Ultimately, while previous studies have demonstrated the potential of using seismological or hydroacoustic networks for the detection of baleen whales[13,33–36], this study extends the approach by integrating long-term acoustic monitoring with quantitative habitat modeling. It establishes a robust framework for leveraging opportunistic acoustic data to support marine spatial planning and conservation, and provides critical baseline data for the management of the Indian Ocean Whale Sanctuary[2].

II. MATERIAL AND METHODS

A. Study area

The study region is located East of Mayotte Island in the southwest Indian Ocean, spanning coordinates from 45.3°E to 46.5°E longitude and 13.5°S to 12.3°S latitude (Fig. 1). It corresponds to the northern entrance of the Mozambique Channel, within the boundaries of the Mayotte Marine Natural Park. The area is characterized by a tropical marine environment influenced by the westward-flowing South Equatorial Current and mesoscale eddy activity. Although the main Agulhas current flows south along the African margin, its retroflections and eddies affect circulation patterns in this area. Bathymetry varies from about 1,300 m near the island to more than 3,500 m offshore. Since 2018, the emergence of the Fani Maoré submarine volcano approximately 50 km offshore has introduced notable geological activity, with hydrothermal emissions that could improve local nutrient availability and biological productivity[37–39]. Hydrographic data indicate a stratified water column with warm and well-mixed surface waters (around 27–29°C) above seasonal and permanent thermoclines and relatively stable salinity values (around 35 PSU), reflecting typical tropical oceanographic conditions. The region experiences two main seasons: a hot and rainy austral summer (roughly November to April) with increased freshwater input and

wind-driven mixing, and a cooler and drier austral winter (May to October). Together, these seasons modulate surface water properties and local oceanographic dynamics.

B. Acoustic data collection

The acoustic data for this study were collected thanks to MAYOBS cruises and the MAHY hydroacoustic network, both components of the REVOSIMA established to monitor the seismo-acoustic environment linked to the volcanotectonic activity of the submarine near Mayotte[29]. For the analysis, data recorded by all sensors from both networks were aggregated into a single composite daily time series.

The Ocean Bottom Seismometers (OBS) network consists of between 3 and 16 seafloor instruments deployed and recovered periodically, with durations ranging from 1 to 8 months (Fig. 1). These instruments are installed directly on the seafloor at average depths around 2,000 meters, with precise locations, recording periods, and seafloor elevations detailed in the appendix A. Most OBS units are equipped with a short-period, 3-component seismometer and a hydrophone, providing continuous recordings of four channels at sampling rates varying between 62.5 and 125 Hz, depending on the deployment phase and station.

The MAHY hydroacoustic network consists of four hydrophone stations arranged in a near-square configuration, covering a geographic area approximately between 45.36°E and 46.15°E longitude, and 12.49°S to 13.35°S latitude, at depths ranging from 1,300 to 1,370 meters (Fig. 1). This quadrilateral spans roughly 70 km on each side and is centered about 50 km around the new volcano. The hydroacoustic array was first deployed in October 2020. Since then, autonomous instruments have been regularly serviced and redeployed four times to ensure data continuity and equipment maintenance. Each instrument continuously records underwater sounds at a higher sampling frequency of 240 Hz.

Because the MAHY network has a higher sampling frequency compared to the OBS network, it is capable of capturing a wider range of biological sounds, potentially enabling the detection of a greater diversity of species. Therefore, the analysis focuses on detections from both networks starting in mid-October 2020, when continuous and overlapping data became available. The recording periods of the stations, extending from October 2020 to October 2024, are represented by the red lines in Fig. 2.

C. Acoustic data analysis

Acoustic detections were produced using the result of Dréo et al. (2025)[15], which identifies baleen whale songs based on their rhythmic structure, specifically the regular intercall intervals (ICI), rather than their spectral content. The species studied include Antarctic blue whale (ANT BW), Southwestern Indian Ocean pygmy blue whale (SWIO PBW), Fin whale (FW), Antarctic minke whale (AMW). Their calls are illustrated in the spectrograms shown in Figure 3. This method is robust to propagation effects and background noise. For each species, the acoustic data were processed using cepstrogram-based analysis to highlight these rhythmic patterns. A species-specific peak-to-valley ratio (p2vr) was calculated on the cepstrogram to quantify repetition strength and produce hourly detection scores. The detector achieved good performance for ANT BW, SWIO PBW, and FW, with true positive rates close to 80% at a fixed false positive rate of 5%. Performance was lower for AMW, likely due to ship noise that affected detections.

To improve detection reliability, data from the four hydrophone stations of the MAHY network and the multiple Ocean Bottom Seismometers of the MAYOBS network were combined into a single detection dataset. A species was considered present during a given hour if at least one sensor in either network recorded a detection. The hourly detections on all sensors were then summed to produce a daily metric, Daily positive hour (DPH), representing the number of hours per day with confirmed acoustic presence anywhere within the monitored area.

Exploration seismic surveys conducted from 2020/12/27 to 2021/02/04 and from 2021/02/20 to 2021/03/07 introduced low-frequency noise that masked detections for ANT BW, SWIO PBW and FW. Detections for these species were excluded during those periods. AMW detections were not affected and were included in the analyses (Fig. 2).

D. Environmental Factors

Environmental variables were obtained from the sea surface within a study area defined by the coordinates 45.3°E to 46.5°E longitude and 13.5°S to 12.3°S latitude. This area covers approximately 17,250 km^2 (see the red rectangle in Fig. 1). The daily values for each variable were spatially averaged in this region to provide representative estimates of local environmental conditions. This averaging approach also aligns with the treatment of acoustic detection results, which were aggregated from all sensors in the REVOSIMA network, including the four MAHY hydrophones and the OBS network, into a single combined DPH metric, representing the study zone as a unified acoustic monitoring unit. It also ensured consistency despite differences in the native spatial resolution of the datasets.

Importantly, the study area fully encompasses all REVOSIMA sensors, none located further than approximately 70 km from the boundaries of the defined zone. An estimated detection range of around 70 km was used as a conservative figure, reflecting the substantial variability in detection ranges reported in the literature, which can vary from about 10 km to several hundred km depending on species, environmental conditions, and acoustic source levels[20,40,41]. This configuration helps to ensure that the acoustic detections correspond primarily to the animals present within the study area.

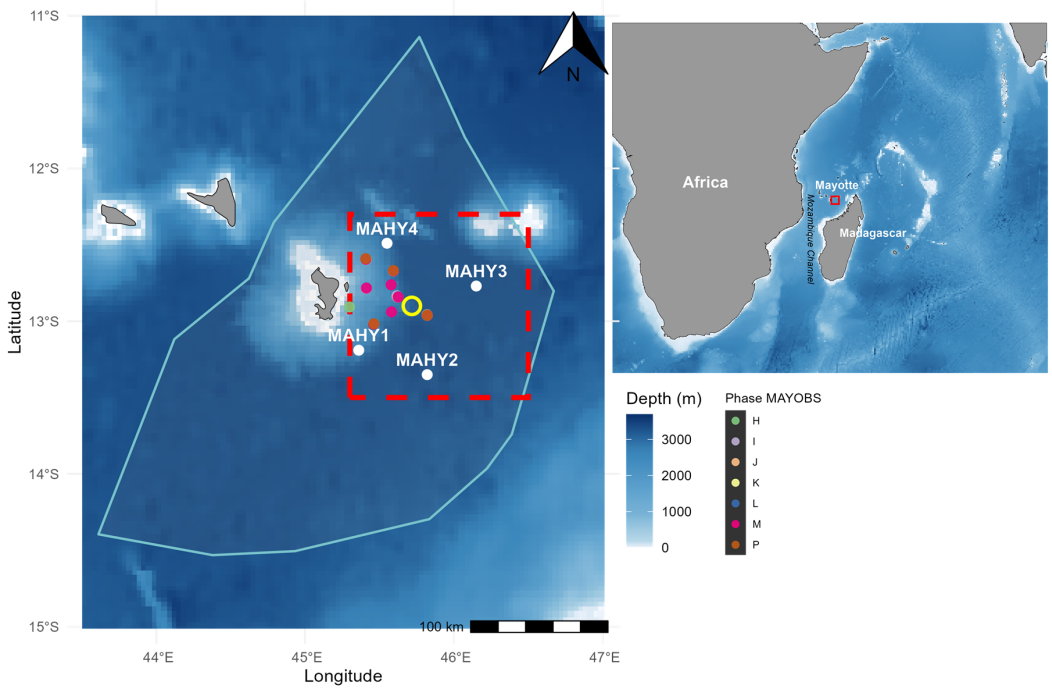


FIG. 1: Locations of the mooring sites of the REVOSIMA network in the southwest Indian Ocean. White points correspond to the MAHY hydroacoustic network, while colored points represent the OBS network deployments during different phases. Some MOBS sensors overlap spatially, so not all points may be visible. The yellow circle marks the position of the Fani Maoré submarine volcano. The red dashed rectangle delineates the area where environmental variables were extracted around the hydrophones. The blue shaded area with a blue border represents the boundaries of the Mayotte Marine Park. Detailed deployment information for MAYOBS is provided in the Appendix A.

Four primary variables were considered: Sea surface temperature (SST), Sea surface temperature anomaly (SST anomaly) Chlorophyll-a (Chl) (log transformed) and Significant wave height (SWH) (Table I). Specifically, SST serves as an indicator of thermal habitat preference, SST anomaly reflects large-scale environmental variability, Chl concentration is a proxy of primary productivity, and SWH is related to ocean surface dynamics and sea state conditions that can affect acoustic detectability and whale behavior (Table I).

The SST, SST anomaly and Chl data were obtained from Copernicus Marine Service products (DOI: 10.48670/moi-00016 for SST and DOI: 10.48670/moi-00015 for Chl), while the SWH data were sourced from the ERA5 hourly reanalysis dataset, aggregated to daily resolution, provided by the Copernicus Climate Change Service (C3S) (DOI: 10.24381/cds.adbb2d47). The SST anomaly was calculated by subtracting the daily climatological mean over a baseline period (1995–2010) from the observed values. This methodology allows for the assessment of thermal variability relative to long-term average conditions, which is essential to understanding the environmental influences on marine ecosystems. The Chl concentration ($mg.m^{-3}$) was logarithmically transformed to reduce the skewness of its distribution and diminish the impact of extreme values, thus enhancing the ecological interpretation of the primary productivity patterns.

To improve the robustness of the dataset, 1% of extreme outliers were removed from the environmental variables using a quantile-based saturation method (values below the 0.5th percentile and above the 99.5th percentile were trimmed). This preprocessing step was implemented via a custom function to ensure that extreme values would not influence the analysis disproportionately. A 15-day median filter was applied to the SWH data to smooth short-term variability and reduce noise, thereby enhancing the robustness of this variable as a proxy for sea-state conditions influencing marine species' behavior.

For this study, the austral seasons are defined as summer (December, January, February), autumn (March, April, May), winter (June, July, August) and spring (September, October, November). In Appendix B, a detailed description and temporal analysis of the environmental covariates used in this study are provided, including interannual variability, seasonal patterns, and notable anomalies observed during the study period.

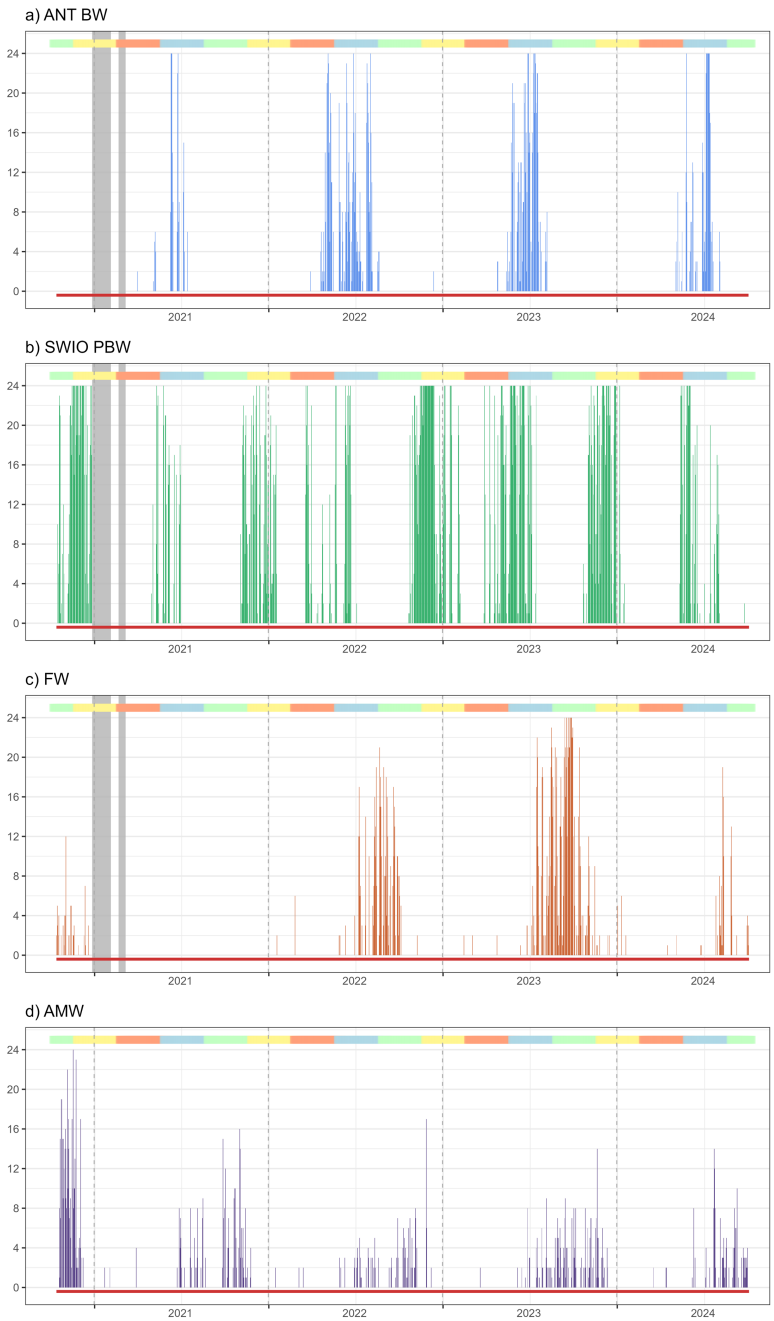


FIG. 2: Distribution of daily positive hours (DPH) for a) ANTW BW, b) SWIO PBW, c) FW, d) AMW throughout the deployment period. The DPH values represent detections aggregated across all REVOSIMA sites, including both the MAHY hydrophone network and the MAYOBS seismometer network. The red horizontal bars indicate the recording periods for each species. The horizontal bars in yellow, orange, blue, and green represent the austral seasons: summer, autumn, winter, and spring, respectively. For panels a) to c), the shaded grey square areas indicate seismic experiment periods during which acoustic detections are considered unreliable or impossible. This does not apply to AMW (d).

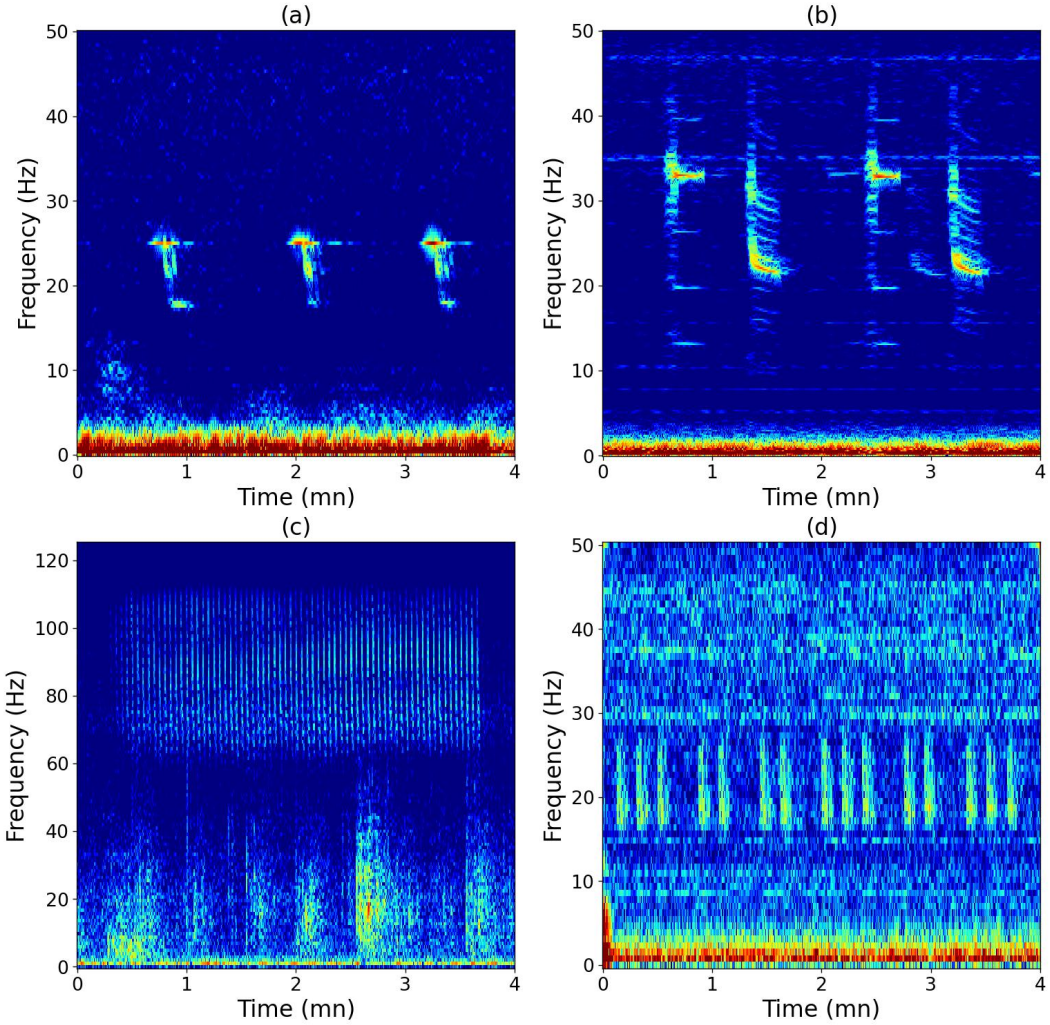


FIG. 3: Spectrograms of sounds from a) ANT BW, b) SWIO PBW, c) AMW, and d) FW recorded in the study area east of Mayotte, from Dréo et al., 2025[15].

TABLE I: Summary of environmental variables used in the study

Environmental Variable (Units)	Spatial Resolution	Temporal Resolution	Source	Expected Effects on Baleen Whale Ecosystems
Chlorophyll-a concentration (mg m^{-3})	0.25° × 0.25°	Daily	Global Ocean Biogeochemistry Analysis and Forecast (Mercator Océan International); DOI: https://doi.org/10.48670/moi-00015	Proxy for primary productivity; indicates phytoplankton biomass.
Sea Surface Temperature (°C)	0.083° × 0.083°	Daily	Global Ocean Physics Analysis and Forecast (Mercator Océan International); DOI: https://doi.org/10.48670/moi-00016	Influences prey distribution and meets the physiological needs of whales.
Sea Surface Temperature Anomaly (°C)	0.083° × 0.083°	Daily	Derived from the same dataset as SST	Reflects thermal variability relative to long-term average conditions, impacting ecosystem dynamics and prey availability.
Significant Wave Height (m)	0.25° × 0.25°	Daily	ERA5 hourly dataset aggregated daily; Copernicus Climate Change Service (C3S); DOI: https://doi.org/10.24381/cds.adbb2d47	Affects local sea conditions, influencing prey and predator behavior.

E. Modeling predictors of acoustic presence

Considering the improbability that the relationships between the aforementioned environmental variables and the spatio-temporal distribution of whales follow a strictly linear pattern, Generalized additive models (GAM) incorporating spatio-temporal components were used[32]. This flexible modeling framework captures non-linear effects of covariates, handles zero-inflated and overdispersed data using appropriate distributions, and accounts for spatio-temporal dependencies through smooth spatial and temporal splines[42,43].

For each species (ANT BW, SWIO PBW, FW, AMW), the DPH of acoustic detections was modeled independently using a GAM with a logarithmic response transformation:

$$\mu_i = \log(\mathrm{DPH}_i + 1) \tag{1}$$

The same set of environmental covariates was included in all models. This modeling choice is based on the hypothesis that the relationships between environmental conditions and acoustic presence may vary over years, and they are not necessarily stationary over time. Interannual variability in oceanographic and ecological processes could alter how cetaceans respond to environmental drivers. Including interactions between year and covariates thus enables the model to capture these potential temporal shifts in ecological relationships.

Initially, a complete reference structure was fitted for each species, including all environmental covariates with year-specific smooth interactions. This setup allowed for interannual variation in functional relationships. These year effects were retained when they improved model performance, as assessed by explained deviance and AIC[44]. However, when these year-specific interactions did not significantly improve performance, based on the significance of the smooth terms or increases in explained deviance, simpler smoothers without year interactions were preferred, following the principle of parsimony.

Collinearity among covariates was assessed using pairwise Pearson's correlation coefficients. All values were below 0.5, indicating low multicollinearity (Fig. S2).

The models were fitted using a Tweedie distribution, which is well-suited for zero-inflated and overdispersed data, with a log link function. The general structure of the model is defined as follows:

$$\begin{aligned} g(\mu_i) = \mu_0 &+ s_1(\mathrm{Week}_i, \mathrm{Year}_i) + s_2(\log(\mathrm{Chl}_i), \mathrm{Year}_i) \\ &+ s_3(\mathrm{SST}_i, \mathrm{Year}_i) + s_4(\text{SST Anomaly}_i, \mathrm{Year}_i) \\ &+ s_5(\mathrm{SWH}_i, \mathrm{Year}_i) \end{aligned} \tag{2}$$

where μ_i is the transformed response variable and μ_0 the intercept. The s_j terms ($j \in \{1, \ldots, 5\}$) represent smooth functions capturing non-linear effects of the environmental covariates: week of the year (Week), log-transformed Chl (log(Chl)), SST, SST anomaly, and SWH. Year-specific interactions were retained only when they improved model fit.

The contribution of each covariate to the explained variance was evaluated by comparing the deviance of the full model to reduced models excluding one covariate at a time. The relative change in deviance, normalized by the null deviance, was used to quantify the explanatory power of each variable (Fig. 4).

All models were implemented using the `mgcv` package in R[45,46].

III. RESULTS

A. Model Selection

Species-specific GAM were fitted to identify environmental drivers of the acoustic presence of baleen whales.

The final structure of the model varied by species, particularly in the inclusion of year-specific interactions. For ANT BW and FW year-specific interactions were initially included but subsequently removed when associated variables were not significant. The reduced models without these interactions were then tested and compared to the full models. In both cases, the simplified models showed equal or improved explanatory power with lower AIC values and were therefore selected as final models according to the principle of parsimony. Table II summarizes the model structures, AIC scores, and explained deviance for each species.

Each final model substantially explained the deviance in the acoustic presence. For ANT BW, the selected model explained 71.6% of the deviance, with all covariates significantly contributing except SWH; week was the most influential predictor (Figure 4a). The SWIO PBW model explained 65.2% of the deviance, with all predictors significant, and this week again showed the highest contribution (Figure 4b). The FW model accounted for 62.9% of the deviance, with only the week, SST anomaly, and SST being significant, and the week being the most important variable (Figure 4c). For AMW, the selected model explained 44.6% of the deviance, with all predictors significant except SWH; in this case, the SST anomaly was the dominant contributor (Figure 4d).

B. Habitat Characterization and Distribution Modeling

1. Antarctic blue whales – ANT BW

ANT BW were detected each year throughout the study period (from 2021 to 2024; Fig. 2a). **Regarding seasonal variation**, the detections were mainly concentrated between late austral autumn and early winter, the acoustic presence typically reaching its maximum during austral winter (around week 25). This intra-annual structure remained consistent across most years. However, in 2023 and 2024, the seasonal peak occurred slightly earlier, around week 20, and both years exhibited the highest intensities of acoustic presence (Fig. 5, week). **In terms of interannual variation**, a notable increase in acoustic presence was observed between 2021 and 2024. These changes were particularly pronounced in 2023 and 2024, which not only had a higher overall acoustic presence. **With respect to environmental covariates**, the relationships varied between years and variables. Log(Chl) did not show a consistent trend, although in 2024 higher concentrations were positively associated with the acoustic presence, while in 2021 the relationship was inverse (Fig. 5, log(Chl). SWH had a weak overall effect, although a slight negative relationship appeared at lower values (below 1 m), although it was not a statistically significant predictor in the final model (Figure 4a). SST generally showed a negative relationship with acoustic presence, with detections starting to decline above 26°C and a stronger negative trend observed beyond 29°C, especially in 2023 and 2024. However, this pattern was not consistent in all years; for example, 2022 showed a slight positive trend (Fig. 5, SST). In contrast, the effects of the SST anomaly were more consistent, with a negative relationship observed in 2022, 2023, and 2024, particularly at anomalies above +0.5°C (Fig. 5, SST anomaly).

2. SWIO pygmy blue whales – SWIO PBW

SWIO PBW were detected each year from 2020 to 2024. **Regarding seasonal variation**, the acoustic activity occurred during two distinct intra-annual periods: one in late autumn to early winter (weeks 20–30), and another in late spring to early summer (weeks 45–50) (Fig. 2b). This bimodal seasonal pattern remained relatively consistent over time, with two periods of low acoustic presence typically observed around weeks 10 (late winter) and 35 (early spring). Although this pattern remained stable from 2021 to 2023, greater variability was observed in 2020 and 2024, although these years lacked full temporal coverage (Fig. 6, Week). **In terms of interannual variation**, the acoustic presence fluctuated moderately between years. The seasonal structure remained detectable each year, but the overall intensity and timing of the peaks were more variable in 2020 and 2024 compared to the more stable years of 2021–2023. **With respect to environmental covariates**, the relationships were variable between years and covariates. Log(Chl) showed no clear effect until the values fell below -1.8. Below this threshold, interannual differences emerged: weak positive associations appeared in 2021 and 2022, while weak negative effects were observed in 2020 and 2023 (Fig. 6, log(Chl)). SWH showed a consistent negative relationship with the acoustic presence, especially at values greater than 1.0 m, and this effect remained stable throughout all years (Fig. 6, SWH). The SST did not appear to have an apparent influence on the acoustic presence, the response remaining flat throughout the range (Fig. 6, SST). In contrast, the effects of the SST anomaly were more variable: a positive relationship emerged in 2021 and 2022, with higher acoustic presence at anomalies above +0.5 °C, whereas a negative effect was observed in 2023 and 2024 (Fig. 6, SST anomaly).

3. Fin whales – FW

FW were recorded in all years from 2020 to 2024, except in 2021, when no acoustic presence was detected. **Regarding seasonal variation**, the acoustic presence occurred predominantly during austral winter and spring, with a marked decrease during mid-autumn (weeks 15–25), followed by a peak in late winter to early spring (around weeks 35–40) (Fig. 7, week). **In terms of interannual variation**, detections peaked in 2023, followed by a notable decline in 2024. The year 2021 was marked by a complete absence of detections, contrasting sharply with other years (Fig.2c). **With respect to environmental covariates**, log(Chl) generally showed a flat response, except in 2021 and particularly in 2022, where a positive relationship was observed for values above –2.3 (Fig.7, log(Chl)). For SWH, no consistent relationship was found for wave heights below

TABLE II: Summary of model selection for each species. All full models are based on the generalized additive model structure described in Eq. 2. For ANT BW and AMW, alternative reduced models omitting some year-specific interactions are shown to illustrate improvements in model parsimony and performance.

Species	**Model description**	**AIC**	**Deviance explained (%)**
ANT BW	Full model: all covariates with year interaction	1 206	70.8
	Final model: SWH without year interaction	**1 127**	**71.6**
SWIO PBW	**Final model: all covariates with year interaction**	**2 326**	**65.2**
FW	**Final model: all covariates with year interaction**	**1 188**	**62.9**
AMW	Full model: all covariates with year interaction	1 527	44.3
	Final model: SST and SST anomaly without year interaction	**1 524**	**44.6**

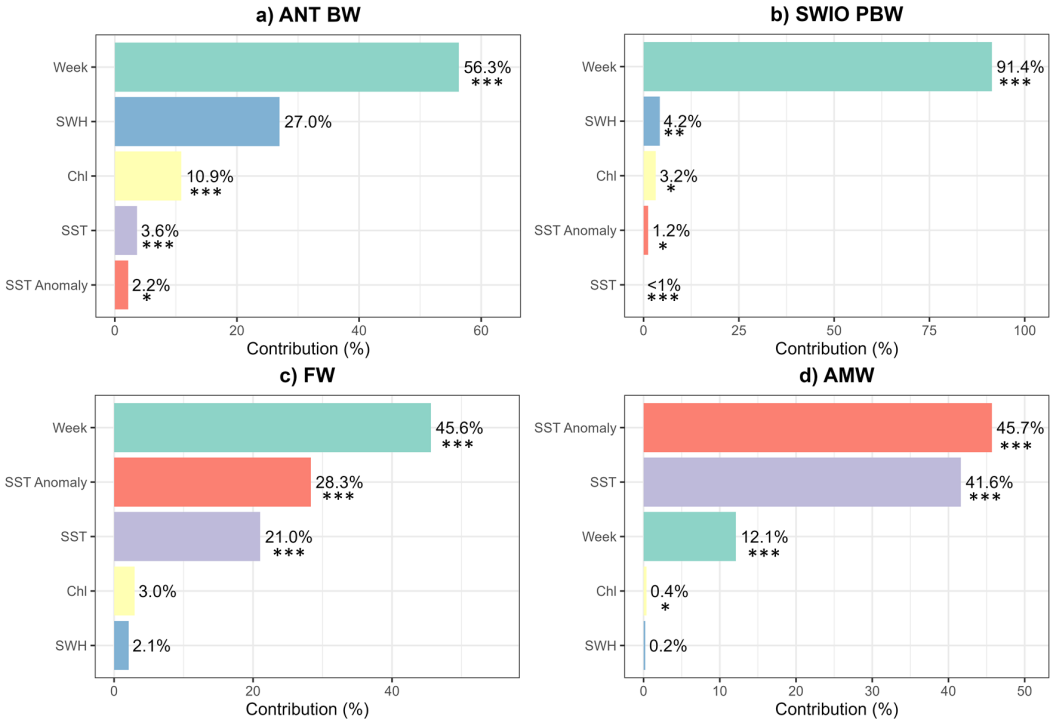


FIG. 4: Contribution (%) of environmental variables used in the GAMs for the four species: a) ANT BW, b) SWIO PBW, c) FW and d) AMW. Asterisks indicate statistically significant contributions: * $p < 0.05$, **$p < 0.01$, *** $p < 0.001$.

1.5 m, while a negative effect emerged for higher values, especially in 2023 and 2024 (Fig.7, SWH). SST showed no clear relationship in 2020 and 2021, but a strong negative effect was evident in 2022, 2023, and 2024, with a lower acoustic presence at temperatures above approximately 28.5 °C (Fig.7, SST). The effects of the SST anomaly varied between years: positive relationships were observed in 2021, 2023, and 2024, with increased acoustic presence at anomalies above +0.5 °C (Fig. 7, SST anomaly).

4. Antarctic minke whales – AMW

AMW were detected each year between 2020 and 2024, although their acoustic presence remained relatively low compared to other baleen whale species (Fig. 2d). **Regarding seasonal variation**, a gradual and consistent increase in detections was observed throughout the year, usually peaking in the late austral spring. **In terms of interannual variation**, no substantial year-to-year changes in acoustic intensity were detected, and the overall signal remained relatively constant over the 5-year period. **With respect to environ-**

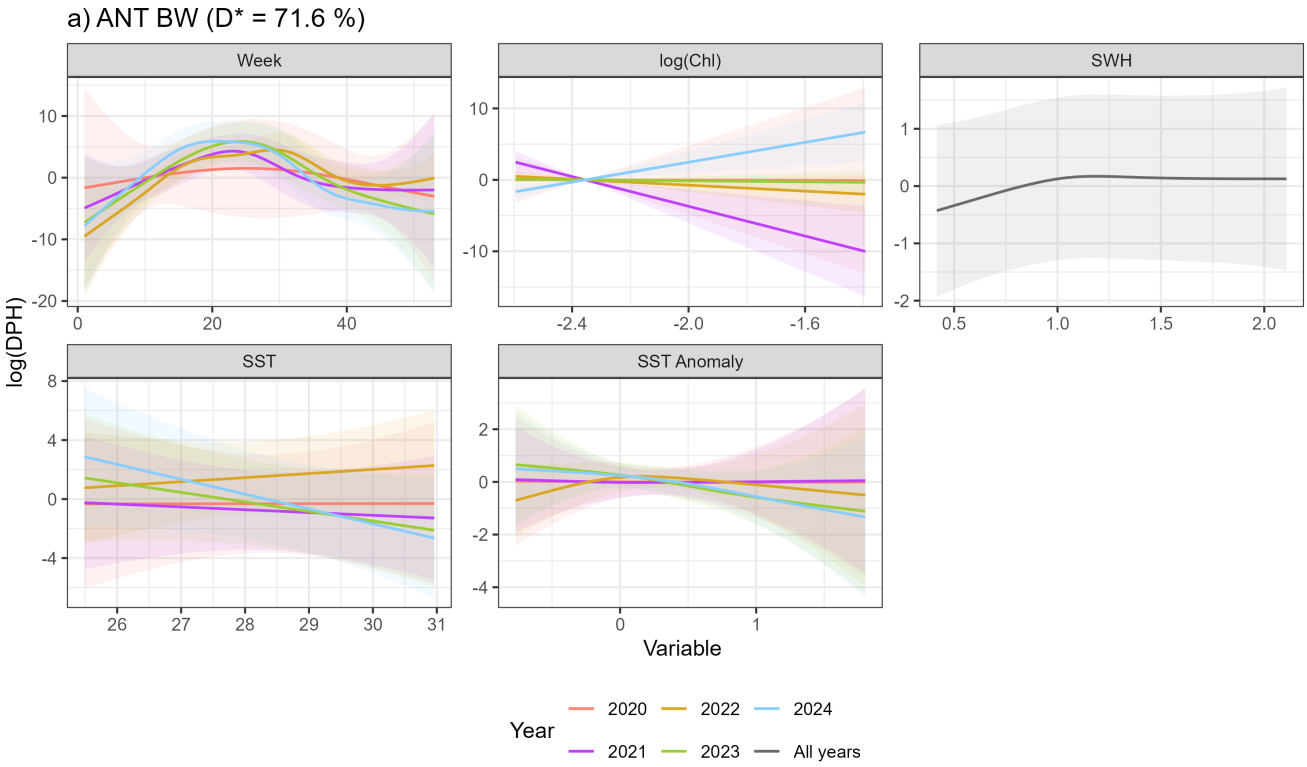


FIG. 5: Functional relationships obtained with the GAM between the log-transformed daily acoustic presence of ANT BW and five environmental covariates: week of the year, log(Chl), SWH, SST, and SST anomaly. The model explained 71.6% of the deviance. Solid lines represent the estimated smooth functions for each year (2020–2024) and the overall multi-year fit ("All years" in black), while shaded areas indicate the approximate 95% confidence intervals. Note that for SWH, the interaction with year was removed in the model due to better performance. The x-axis corresponds to the covariate value, and the y-axis to the smooth term.

mental covariates, log(Chl) showed a positive association with the acoustic presence in 2023 for values greater than –2.2, while other years showed flat or weakly negative trends, most notably in 2021, which showed a slight negative effect beyond –2.2 (Fig. 8, log(Chl)). SWH generally exhibited a negative effect, although slightly positive or neutral responses were observed in 2022 and 2023 (Fig. 8, SWH). SST consistently demonstrated a strong negative relationship with acoustic presence throughout all years, the effect becoming negative at temperatures above 27 °C (Fig. 8, SST). In contrast, SST anomaly was positively associated with the acoustic presence, especially for values exceeding +0.5 °C. These effects remained consistent throughout the study period and did not show clear interannual variability (Fig. 8, SST anomaly).

IV. DISCUSSION

Conducted in the heart of the Mayotte Natural Marine Park, this study uses a unique high-resolution time series from the REVOSIMA observatory, which integrates both the MAHY hydroacoustic and MAYOBS seismological networks, to provide novel insights into the seasonal and interannual habitat use of multiple baleen whale species at tropical latitudes. REVOSIMA represents a long-term acoustic monitoring initiative in the northern Mozambique Channel and around Mayotte[29]. This recording effort complements and extends earlier efforts near Nosy Be, northwest Madagascar (13°S, 48°E), where passive acoustic data were collected between 2016 and 2021[14]. More broadly, REVOSIMA expands the geographical coverage of passive acoustic observations in the southwestern Indian Ocean by addressing a major data gap left by the OHASISBIO observatory[47], which operated from 2009 to 2023 in mid- to high-latitude waters (20°S–60°S, 45°E–90°E) but did not include the northern Mozambique Channel[12,22]. As such, this study contributes to filling a critical gap in long-term monitoring of baleen whale occurrence in tropical regions, an area where habitat use remains insufficiently characterized despite growing interest and conservation relevance.

Based on the seasonal detection patterns identified by Dréo et al. (2025)[15], this study further examines the ecological importance of the region by exploring the relationship between the acoustic presence of three species

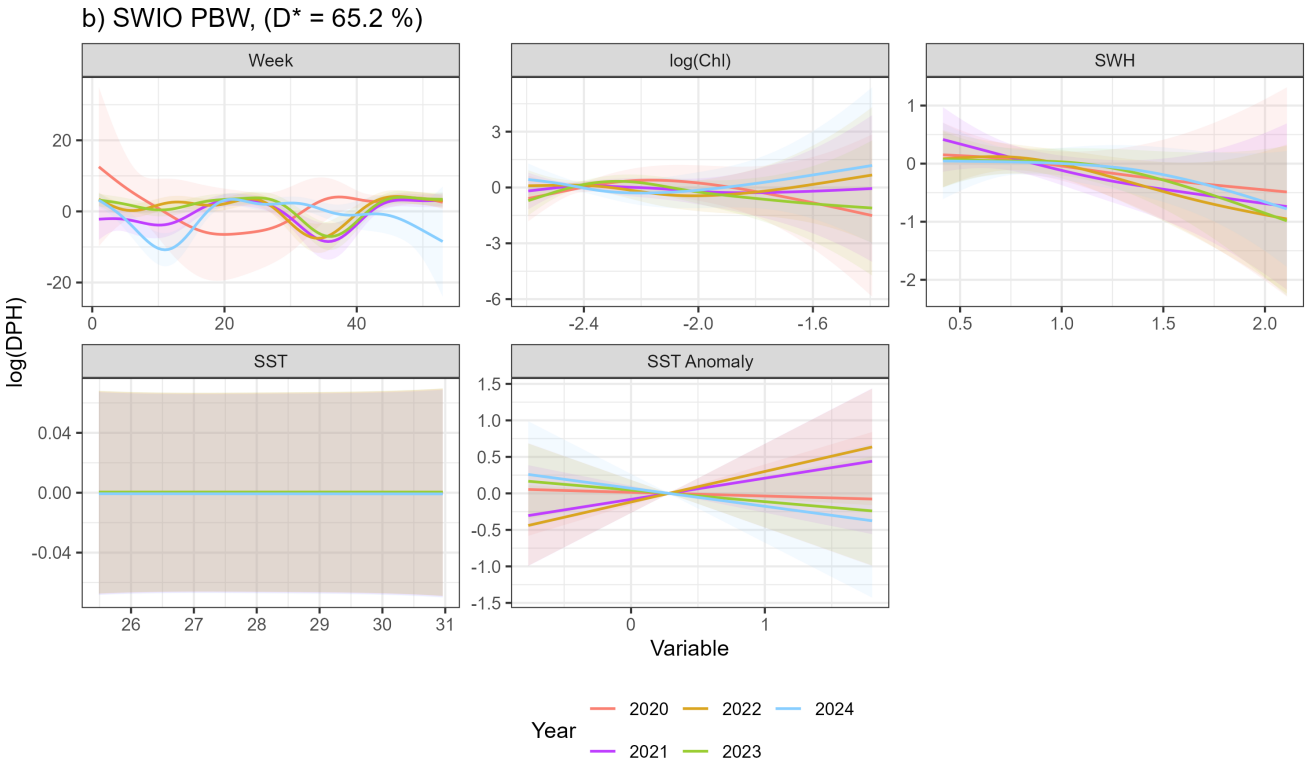


FIG. 6: Functional relationships obtained with the GAM between the log-transformed daily acoustic presence of SWIO PBW and five environmental covariates: week of the year, log(Chl), SWH, SST, and SST anomaly. The model explained 65.2% of the deviance. Solid lines represent the estimated smooth functions for each year (2020–2024), while shaded areas indicate the approximate 95% confidence intervals. The x-axis corresponds to the covariate value, and the y-axis to the smooth term.

of baleen whales, including two different subspecies of blue whales, in relation to key environmental drivers. One of the primary objectives was to characterize the functional role of this area for multiple species of baleen whales. Our analyses reveal that these species do not occur simultaneously, exhibit distinct interannual variability, and show differing responses to environmental covariates, with the exception that chlorophyll-a concentrations appear to have little effect across all species. This suggests that each species employs specific habitat use strategies shaped by its environmental preferences and behavioral ecology. The long-term dataset enables the evaluation of both seasonal dynamics and interannual variability in habitat usage, allowing the identification of contrasting strategies among species. The findings suggest that the northern Mozambique Channel serves as a regular habitat for several species, whose acoustic presence appears to be influenced by local environmental conditions. GAM revealed that SST, SST anomaly, and SWH are critical covariates, although their impacts varied between species, highlighting different habitat utilization patterns and ecological strategies. Consequently, the discussion that follows is structured as a species-by-species comparison to better elucidate these contrasting habitat use patterns and ecological roles within the region.

Consistent multi-year detections underline the importance of this subtropical region as a habitat supporting critical life stages of ANT BW. Acoustic detections peak during the austral winter, mainly between May and August[15]. Their wintering range covers a large portion of the Indian Ocean, including the central and western sectors[12,13,20]. The Mozambique Channel, monitored here, probably marks the northern limit of their acoustic presence in the region, as detections become scarce farther north[11]. This species undertakes annual migrations between Antarctic summer feeding grounds[48] and subtropical wintering areas for breeding or resting[11,13,14,16]. Consequently, the Mozambique Channel offers environmental conditions, such as relatively stable SST and moderate wave heights, that likely create a suitable resting habitat during migration. Although the seasonal timing of detections is generally consistent, with annual returns around the same period, interannual variability is evident. In particular, there was a marked decline in 2021 and a temporal shift between 2023 and 2024. This variability could be related to environmental or behavioral factors that are not captured by surface parameters

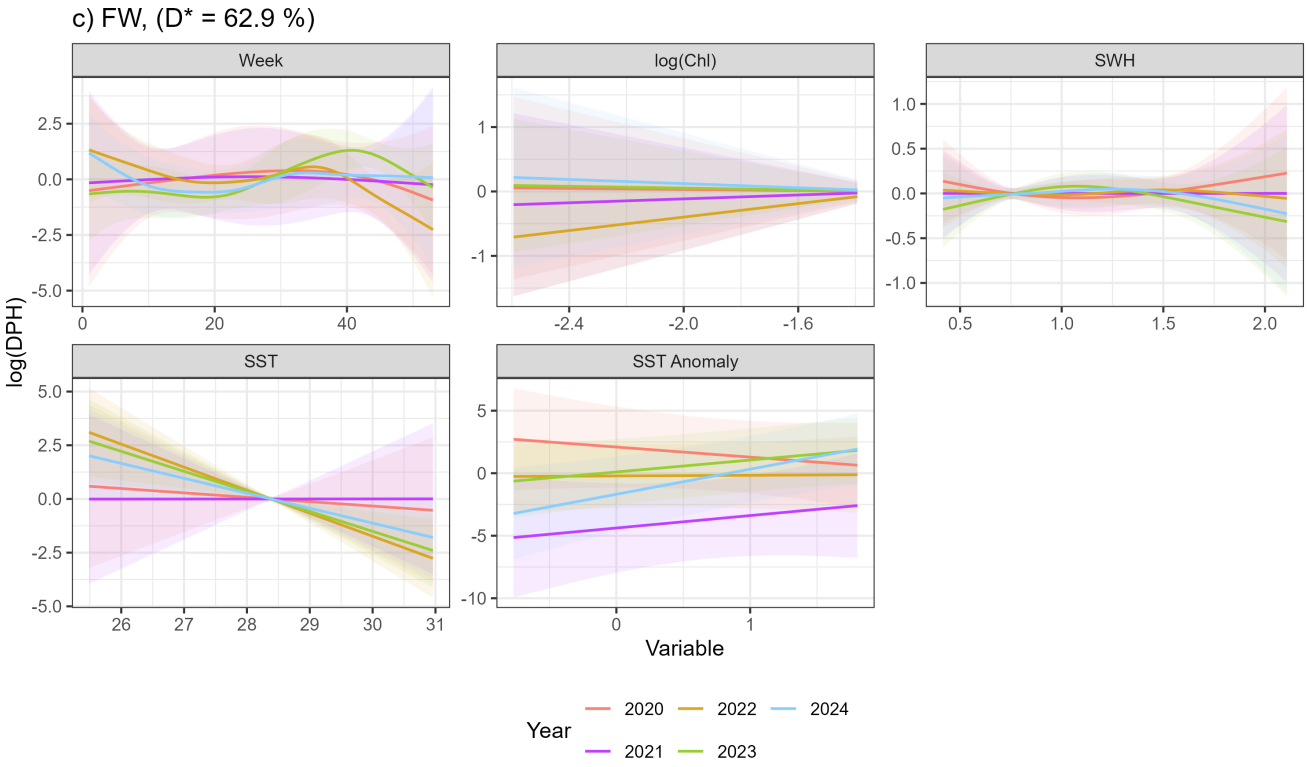


FIG. 7: Functional relationships obtained with the GAM between the log-transformed daily acoustic presence of FW and five environmental covariates: week of the year, log(Chl), SWH, SST, and SST anomaly. The model explained 60.1% of the deviance. Solid lines represent the estimated smooth functions for each year (2020–2024), while shaded areas indicate the approximate 95% confidence intervals. The x-axis corresponds to the covariate value, and the y-axis to the smooth term.

alone. In particular, it may reflect changes in oceanographic conditions farther south, either in their Antarctic feeding grounds or along their migratory routes. Although SST and SWH remained stable locally during these years, other variables, such as subsurface conditions or large-scale climatic anomalies, may play a more significant role. The absence of correlation with Chl further supports the idea that this region is not a feeding area, in agreement with the known foraging zones in Antarctic waters[48]. Although SST itself does not appear to drive the local acoustic presence, the slight negative correlation with SST anomaly suggests a limited tolerance for anomalously high temperatures, which could influence future habitat suitability.

SWIO PBW acoustic presence reflects their use of an extensive migratory corridor spanning the southwestern Indian Ocean. Acoustic detections cover a broad geographic range, from Diego Garcia in the north[16] to the Kerguelen Plateau in the south[22], and from Marion Island in the west[20] to waters southwest of Amsterdam Island in the east[12]. This wide spatial distribution indicates long-distance movements that rely on multiple key regions throughout the migratory cycle. Seasonal detection patterns consistently show two annual peaks: one in early austral winter and another in late spring to early summer[15]. These bimodal peaks likely correspond to biannual migrations[14], with the first movement in summer probably targeting subantarctic feeding grounds. This is supported by frequent detections and documented feeding behavior near Kerguelen, Marion, and the Prince Edward Islands during late summer and autumn[20]. The second phase may involve migration to northern resting or reproductive areas. However, exact locations remain unknown as no detections have been recorded at expected tropical sites such as the Seychelles[49]. The environmental conditions along this corridor, especially seasonally calm seas and a relatively stable SST of around 29 °C, appear to support both long-distance transit and temporary resting. The two annual acoustic peaks, in late austral autumn and early summer, coincide with these similar environmental windows, suggesting that SWIO PBW transit through the area when conditions are predictably favorable, rather than being directly driven by interannual variation. The strong negative correlation with high SWH confirms a preference for calmer seas, which is likely to enhance the efficiency of migration. Elevated SWH could also contribute to acoustic masking by increasing ambient noise levels, potentially reducing detection

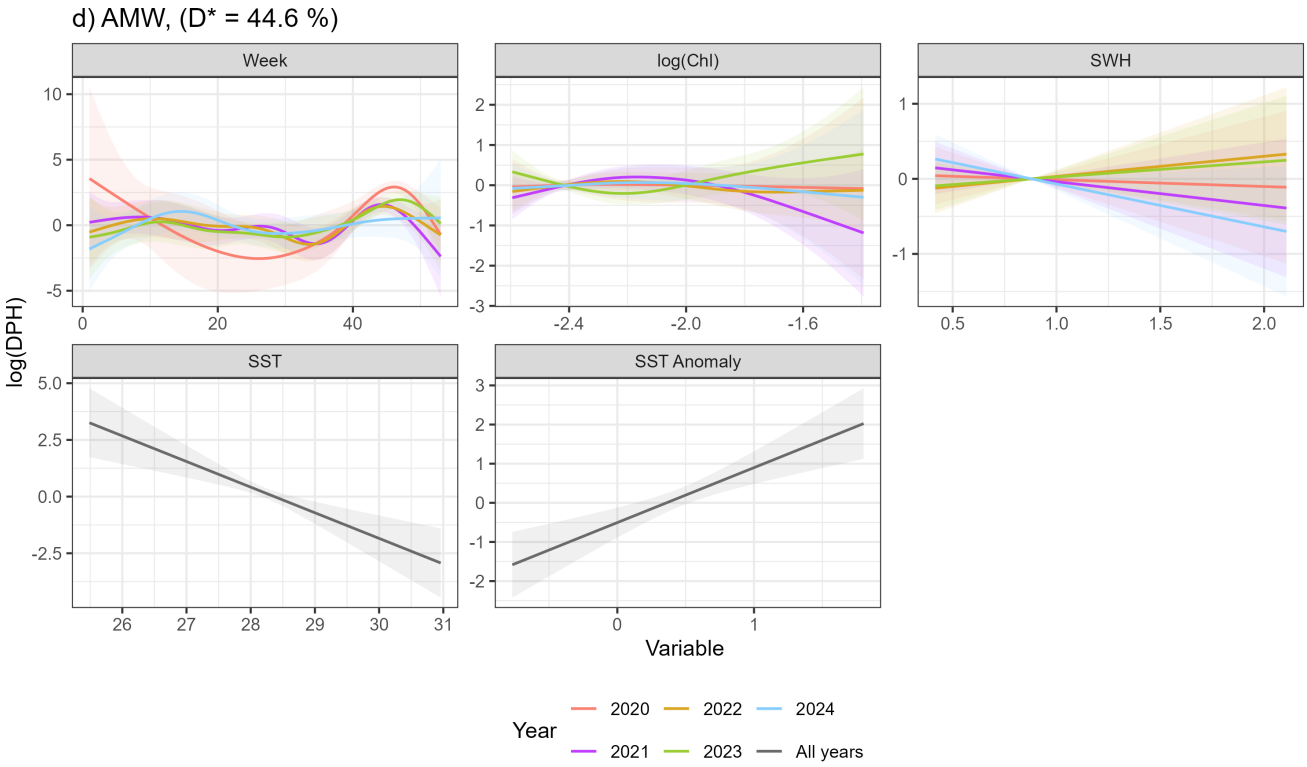


FIG. 8: Functional relationships obtained with the GAM between the log-transformed daily acoustic presence of ANT BW and five environmental covariates: week of the year, log(Chl), SWH, SST, and SST anomaly. The model explained 44.6% of the deviance. Solid lines represent the estimated smooth functions for each year (2020–2024) and the overall multi-year fit ("All years" in black), while shaded areas indicate the approximate 95% confidence intervals. Note that for SST and SST anomaly, the interaction with year was removed from the model due to better performance. The x-axis corresponds to the covariate value, and the y-axis to the smooth term.

rates. However, this explanation is less likely, as similar masking effects would be expected across all baleen whale species monitored here, given that their vocalizations occupy comparable low-frequency bands, which is not supported by our observations. Meanwhile, the absence of consistent SST effects over the years probably reflects the acoustic presence of whales in this region during thermally stable periods, rather than any strict thermal preference. The lack of correlation with Chl further supports the idea that this area is not a feeding ground. Although the seasonal timing of the detections remains consistent across years, variations in the amplitude of the detection can reflect fluctuations in migration intensity, social composition, or ecological drivers not captured by the measured parameters. The absence of a bimodal seasonal pattern in 2020 is due to incomplete recording coverage that year, rather than an ecological deviation. Despite such data limitations, the regular presence of the species over multiple years highlights the importance of this corridor in its migratory cycle and suggests resilience to local environmental variability.

The absence of FW detections in 2021, in contrast to their acoustic presence in all other years from 2020 to 2024, highlights the notable interannual variability in their occurrence within the Mayotte region[15]. This variability cannot be fully explained by key environmental parameters such as SST, SST anomaly, or Chl concentrations, which remained relatively stable over the years. The acoustic presence predominantly occurred during the austral winter and spring in the Mayotte region, aligning with known migration and reproductive periods in the western Indian Ocean. In contrast, feeding occurs primarily in Antarctic waters during austral autumn, indicating a clear seasonal separation between feeding and breeding grounds[13,19,48]. The analyzes reveal a negative correlation between the acoustic presence of FW and SST, with detections decreasing as the SST increases, suggesting thermal constraints or an upper temperature preference. Interestingly, a positive relationship was observed with increased positive SST anomaly in some years, suggesting that short-term warmer conditions within a certain range may enhance habitat suitability or trigger changes in whale distribution, potentially related to prey availability or migratory signals. Regarding the state of the sea, higher SWH values (¿1.5 m) negatively impacted the acoustic presence,

reflecting a preference for calmer seas, which probably facilitates acoustic communication and efficient migration. The lack of significant correlations with Chl supports the idea that Mayotte is a transition area that is used primarily for resting or social interactions, not for feeding. Although local environmental conditions in Mayotte remain relatively stable, the observed interannual variability in FW acoustic presence is likely influenced by non-environmental factors such as group dynamics or breeding cycles, which likely drive the interannual variability. In addition, environmental changes elsewhere along their migratory route, such as shifts in prey availability or conditions at feeding grounds, may also influence their timing and presence in Mayotte. This underscores the importance of studying both local social and demographic factors and broader environmental drivers throughout their migration to better understand FW habitat use in this tropical region.

The migratory behavior of AMW in the Indian Ocean is poorly understood. Acoustic detections confirm the presence of AMW in Antarctic and subantarctic waters during the late austral winter and early spring, particularly near the western Antarctic Peninsula and the Marion and Prince Edward Islands[50,51]. However, acoustic recordings near Mayotte also reveal a consistent presence over the same period, from late austral winter to spring, with notable interannual variability and a distinctive peak in 2020[15,21,52]. These simultaneous detections in widely separated locations suggest the existence of distinct subpopulations or multiple migratory strategies within the Indian Ocean[20]. The environmental conditions in the Mayotte region, characterized by relatively low SST (around 28 °C) and limited primary productivity, support the interpretation of this area as a migratory or resting habitat rather than a feeding ground. This is reinforced by a weak correlation with Chl and the absence of a clear relationship with the SWH. The acoustic presence is negatively related to SST, with detections decreasing as temperatures rise above 28 °C, suggesting a preference for the lowest seasonal values. A positive correlation with moderate SST anomaly suggests some tolerance to short-term warming, although detections tend to peak before SST anomaly increase, indicating a likely avoidance of warmer conditions than average. These observations point to a selective use of the area during cooler periods, while broader spatial patterns raise questions about connectivity and movement across distant sites. More research is needed to clarify the ecological drivers and migratory pathways of AMW in the Indian Ocean, especially to understand whether the observed patterns reflect the structure of the population or the flexibility of behavior.

These patterns reveal different habitat use strategies between species. ANT BW and SWIO PBW show contrasting patterns: the former appears to use the region as a stable resting area with potential social functions during migration, while the latter shows transient movements through a broad corridor with low site fidelity. In contrast, both FW and AMW exhibit sporadic and thermally restricted acoustic presence, suggesting peripheral or opportunistic use of the area at the northern edge of their distribution. Additionally, both species show seasonal patterns peaking between late austral winter and spring, and a preference for the cooler temperature range available in this tropical environment, further supporting the idea that their presence is governed by thermal limits and possibly other external factors influencing migratory timing.

Although studies on cetacean presence and distribution have traditionally relied on visual observations, there is a growing body of research using acoustic data as a valuable and complementary approach, particularly in remote or logistically challenging regions. It is important to emphasize that the acoustic presence reported in this study reflects the detection of vocalizing individuals and does not correspond to the absolute abundance or density of cetacean populations. Due to the nature of acoustic monitoring, only a subset of the animals present are captured, depending on their vocal behavior, species-specific call rates, and environmental conditions affecting sound propagation. Consequently, these data represent the relative presence and temporal patterns at the monitored locations, rather than the complete distribution or population estimates throughout the region. Therefore, the interpretation of acoustic presence must be contextualized within these limitations, acknowledging that silent animals or those outside the detection range are not taken into account. Despite these limitations, the continuous and non-invasive nature of acoustic monitoring provides key information on species temporal presence and migratory dynamics, especially in regions where long-term visual surveys are logistically infeasible.

Beyond ecological patterns and habitat use strategies, it is essential to consider the methodological context and monitoring approach underlying our results. Although this study provides valuable information, several limitations must be acknowledged. First, we focus on a limited set of broad-scale environmental variables. Although ecologically significant at large scales, these metrics may overlook finer-scale oceanographic processes such as upwelling zones, frontal systems, or mesoscale eddies that can significantly influence prey aggregation and cetacean habitat use. In addition, chlorophyll-a was used as a proxy for primary productivity and potential prey availability, although it does not directly reflect the abundance or distribution of specific prey species such as krill or small pelagic fish. Incorporating direct prey indicators, such as acoustic biomass surveys or satellite-derived trophic indices, would improve our understanding of ecological drivers and improve habitat models. Furthermore, intra- and interannual variations in cetacean presence may be influenced by environmental changes occurring elsewhere along their migratory routes, including shifts in prey availability or habitat conditions in feeding and breeding grounds outside the study area. This highlights the need to adopt a broader spatial perspective when investigating the drivers of habitat use and population dynamics in highly migratory species.

Initially designed for seismological applications, marine acoustic observatories have significant prospects for collecting long-term biological data. Their dual functionality enables the acquisition of continuous and extensive time series, which are crucial for enhancing our understanding of marine species' utilization of oceanic environments, particularly cetaceans. Such multidisciplinary systems markedly advance ecological knowledge and set the stage for comprehensive and sustainable monitoring in difficult-to-access regions. In the northern Mozambique Channel and the eastern region of Mayotte, acoustic analyses indicate that multiple mysticete species frequent this area, exhibiting distinct seasonal behaviors associated with resting, reproducing, or migrating transit phases. Mayotte appears to be a pivotal region acting as a stable ecological corridor for these species, where habitat use is influenced by local environmental conditions. Together, these findings suggest that the northern Mozambique Channel constitutes a multipurpose habitat used by baleen whales with different strategies. In light of global ocean warming and increasing anthropogenic pressures in the region, it is essential to maintain high-resolution acoustic monitoring and integrate these data with oceanographic and behavioral studies. These efforts are key in forming effective conservation strategies across international boundaries in this biodiversity hotspot.

V. CONCLUSION

This study demonstrates that the northern Mozambique Channel and the Mayotte region serve as critical habitats for multiple baleen whale species, each exhibiting distinct habitat use strategies. Long-term acoustic monitoring reveals seasonal and interannual variability in species presence, highlighting the area's role as a resting, breeding, or migratory corridor. Local environmental factors, such as sea surface temperature and wave height, influence acoustic presence, while additional factors, including prey availability and social dynamics, likely play a role. These findings underscore the value of multidisciplinary acoustic observatories for monitoring marine megafauna in tropical regions. By providing insights into whale ecology and migratory patterns, this work fills a major gap in long-term monitoring in the southwestern Indian Ocean. In the context of global warming and increasing anthropogenic pressures, maintaining and expanding high-resolution acoustic monitoring is essential for developing effective conservation strategies in this biodiversity hotspot.

DATA AVAILABILITY

The environmental data used in this study are publicly available from dedicated databases:

- Sea Surface Temperature: Data were obtained from the Copernicus Marine global ocean reanalysis (DOI: 10.48670/moi-00016), covering the period 1992 to 2024 with a spatial resolution of 1/12° (approximately 0.083°) and daily temporal resolution.
- Sea Surface Temperature Anomaly: The data was derived from the same dataset as SST, reflecting thermal variability relative to long-term averages, with identical spatial and temporal resolution.
- Chlorophyll-a concentration: Data were sourced from the Mercator Ocean Global Ocean Biogeochemistry Analysis and Forecast system (DOI: 10.48670/moi-00015), available from January 1993 to February 2024 at a spatial resolution of 0.25 ° and a daily temporal resolution.
- Significant Wave Height: Data were obtained from the ERA5 hourly dataset aggregated to daily values; Copernicus Climate Change Service (C3S) (DOI: 10.24381/cds.adbb2d47), with a spatial resolution of 0.25°.
- The MAHY acoustic data used in this study is available as FDSN network 1T (DOI 10.15778/RESIF.1T2018).

ACKNOWLEDGMENTS

This study benefited from the continuous seismic and hydroacoustic data provided by REVOSIMA, supported by CNRS/INSU, IPGP, IFREMER, and BRGM. All marine operations are performed as part of the MAYOBS set of cruises (DOI 10.18142/291) and contribute to the French Service National d'Observation en Volcanologie (SNOV). We thank all the technical and scientific teams involved in the MAYOBS cruises, whose efforts were essential for data collection and instrument deployment. Many thanks also go to the French Biodiversity Agency (OFB) and Boksound for their essential support. This work was supported by the MUSIC-ANR-23-EDIR-002 project.

AUTHOR CONTRIBUTIONS STATEMENT

Mathilde Michel: Conceptualization, Methodology, Formal Analysis, Original Draft Writing, Visualization. **Auriane Virgili**: Methodology, Validation, Formal Analysis, Writing-Review & Editing. **Richard Dréo**: Resources, Formal Analysis, Writing-Review & Editing. **Baptiste Alglave**: Methodology, Validation, Formal Analysis, Writing-Review & Editing. **Maëlle Torterotot**: Validation, Writing-Review & Editing. **Sara Bazin**: Project administration, Resources, Writing-Review & Editing. **Flore Samaran**: Validation, Writing-Review & Editing, Project Administration, Funding Acquisition.

COMPETING INTERESTS

The author(s) declare no competing interests.

[1]J. Roman, J. A. Estes, L. Morissette, C. Smith, D. Costa, J. McCarthy, J. B. Nation, S. Nicol, A. Pershing, and V. Smetacek, "Whales as marine ecosystem engineers," Frontiers in Ecology and the Environment **12**(7), 377–385 (2014).

[2]IWC, "Chairman's report of the thirty-first annual meeting.," Rep. int. Whal. Commn. 30:25–41 (1980).

[3]K. M. Stafford, E. Chapp, D. R. Bohnenstiel, and M. Tolstoy, "Seasonal detection of three types of "pygmy" blue whale calls in the indian ocean," Marine Mammal Science **27**(4), 828–840 (2011).

[4]T. A. Branch, K. Stafford, D. Palacios, C. Allison, J. Bannister, C. Burton, E. Cabrera, C. Carlson, B. Galletti Vernazzani, P. Gill, *et al.*, "Past and present distribution, densities and movements of blue whales balaenoptera musculus in the southern hemisphere and northern indian ocean," Mammal Review **37**(2), 116–175 (2007).

[5]T. A. Branch, C. C. Monnahan, E. C. Leroy, F. W. Shabangu, A. Širović, S. Cerchio, S. Al Harthi, C. Allison, N. B. Cabrera, D. R. Barlow, *et al.*, "Separating historical catches among pygmy blue whale populations using recent song detections," Marine Mammal Science e70003 (2025).

[6]L. Mannocci, S. Laran, P. Monestiez, G. Dorémus, O. Van Canneyt, P. Watremez, and V. Ridoux, "Predicting top predator habitats in the southwest indian ocean," Ecography **37**(3), 261–278 (2014).

[7]S. Laran, M. Authier, O. Van Canneyt, G. Dorémus, P. Watremez, and V. Ridoux, "A comprehensive survey of pelagic megafauna: their distribution, densities, and taxonomic richness in the tropical southwest indian ocean," Frontiers in Marine Science **4**, 139 (2017).

[8]T. Ichihara, "The pygmy blue whale, balaenoptera musculus brevicauda, a new subspecies from the antarctic," Whales, dolphins, and porpoises 79–111 (1966).

[9]D. Ljungblad, C. W. Clark, and H. Shimada, "Sc/49/sh17 a comparison of sounds attributed to pygmy blue whales (balaenoptera musculus brevicauda) recorded south of the madagascar plateau and those attributed to'true'blue whales (balaenoptera musculus) recorded off antarctica," Report-International Whaling Commission **48**, 439–442 (1998).

[10]F. Samaran, O. Adam, and C. Guinet, "Discovery of a mid-latitude sympatric area for two southern hemisphere blue whale subspecies," Endangered Species Research **12**(2), 157–165 (2010).

[11]E. C. Leroy, F. Samaran, J. Bonnel, and J.-Y. Royer, "Seasonal and diel vocalization patterns of antarctic blue whale (balaenoptera musculus intermedia) in the southern indian ocean: a multi-year and multi-site study," PloS one **11**(11), e0163587 (2016).

[12]M. Torterotot, F. Samaran, K. M. Stafford, and J.-Y. Royer, "Distribution of blue whale populations in the southern indian ocean based on a decade of acoustic monitoring," Deep Sea Research Part II: Topical Studies in Oceanography **179**, 104874 (2020).

[13]R. Dréo, L. Bouffaut, E. Leroy, G. Barruol, and F. Samaran, "Baleen whale distribution and seasonal occurrence revealed by an ocean bottom seismometer network in the western indian ocean," Deep Sea Research Part II: Topical Studies in Oceanography **161**, 132–144 (2019).

[14]S. Cerchio, T. Rasoloarijao, B. Mueller-Brennan, and D. Cholewiak, "Migratory movements and acoustic activity of blue whales (balaenoptera musculus) and fin whales (balaenoptera physalus) in the eastern mozambique channel," Technical report (2024).

[15]R. Dréo, W. Crawford, G. Barruol, S. Bazin, J.-Y. Royer, and F. Samaran, "Singing around the volcano: Detecting baleen whales in the mozambique channel based on their song rhythms, from seismic and hydroacoustic data," The Journal of the Acoustical Society of America (2025).

[16]K. M. Stafford, D. R. Bohnenstiehl, M. Tolstoy, E. Chapp, D. K. Mellinger, and S. E. Moore, "Antarctic-type blue whale calls recorded at low latitudes in the indian and eastern pacific oceans," Deep Sea Research Part I: Oceanographic Research Papers **51**(10), 1337–1346 (2004).

[17]F. Samaran, K. M. Stafford, T. A. Branch, J. Gedamke, J.-Y. Royer, R. P. Dziak, and C. Guinet, "Seasonal and geographic variation of southern blue whale subspecies in the indian ocean," PloS one **8**(8), e71561 (2013).

[18]M. A. McDonald, S. L. Mesnick, and J. A. Hildebrand, "Biogeographic characterization of blue whale song worldwide: using song to identify populations," Journal of cetacean research and management **8**(1), 55–65 (2006).

[19]E. C. Leroy, F. Samaran, K. M. Stafford, J. Bonnel, and J.-Y. Royer, "Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the southern indian ocean," Endangered Species Research **37**, 289–300 (2018).

[20]F. W. Shabangu, T. Munoz, L. Van Uffelen, B. J. Estabrook, D. Yemane, K. M. Stafford, T. A. Branch, E. Vermeulen, M. A. van den Berg, and T. Lamont, "Diverse baleen whale acoustic occurrence around two sub-antarctic islands: A tale of residents and visitors," Scientific Reports **14**(1), 21663 (2024).

[21]S. Cerchio, T. Rasoloarijao, B. Mueller-Brennan, and D. Cholewiak, "Acoustic monitoring of blue whales (balaenoptera musculus) and fin whales (balaenoptera physalus) in the mozambique channel off the northwest coast of madagascar," Technical report (2018).

[22]M. Michel, B. Alglave, M. Olmos, M. Torterotot, A. Virgili, S. Martin-Marin, J.-Y. Royer, and F. Samaran, "Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern indian ocean," Scientific Reports **15**(1), 23218 (2025).

[23]A. Virgili, M. Authier, O. Boisseau, A. Cañadas, D. Claridge, T. Cole, P. Corkeron, G. Dorémus, L. David, N. Di-Méglio, *et al.*, "Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale: Large-scale modelling of deep-diving cetacean habitats," Global Ecology and Biogeography **28**(3), 300–314 (2019).

[24]C. Lambert, L. Mannocci, P. Lehodey, and V. Ridoux, "Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions," PLoS One **9**(8), e105958 (2014).

[25]S. Nicol, A. Worby, and R. Leaper, "Changes in the antarctic sea ice ecosystem: potential effects on krill and baleen whales," Marine and Freshwater Research **59**(5), 361–382 (2008).

[26]F. W. Shabangu, K. P. Findlay, D. Yemane, K. M. Stafford, M. Van den Berg, B. Blows, and R. K. Andrew, "Seasonal occurrence and diel calling behaviour of antarctic blue whales and fin whales in relation to environmental conditions off the west coast of south africa," Journal of Marine Systems **190**, 25–39 (2019).

[27]S. Laran and A. Gannier, "Spatial and temporal prediction of fin whale distribution in the northwestern mediterranean sea," ICES Journal of Marine Science **65**(7), 1260–1269 (2008).

[28]L. Bedriñana-Romano, R. Hucke-Gaete, F. A. Viddi, J. Morales, R. Williams, E. Ashe, J. Garcés-Vargas, J. P. Torres-Florez, and J. Ruiz, "Integrating multiple data sources for assessing blue whale abundance and distribution in chilean northern patagonia," Diversity and Distributions **24**(7), 991–1004 (2018).

[29]M. Volcanological, S. M. N. (REVOSIMA), I. de physique du globe de Paris (IPGP), B. de recherches géologiques et minières (BRGM), I. français de recherche pour l'exploitation de la mer (IFREMER), and C. national de la recherche scientifique (CNRS), "Data collection of the mayotte volcanological and seismological monitoring network (revosima) (version 1.0)" (2021), https://doi.org/10.18715/MAYOTTE.REVOSIMA, doi: 10.18715/MAYOTTE.REVOSIMA version 1.0.

[30]N. Feuillet, J. Van der Woerd, and RESIF, "Seismic sequence monitoring on land and at sea in mayotte, sismayotte (resif-sismob)" , https://doi.org/10.15778/RESIF.1T2018 (2022), doi: 10.15778/RESIF.1T2018 data available through RESIF.

[31] A. Lavayssière, S. Bazin, and J.-Y. Royer, "Hydroacoustic monitoring of mayotte submarine volcano during its eruptive phase," Geosciences **14**(6), 170 (2024).

[32] S. N. Wood, *Generalized additive models: an introduction with R* (chapman and hall/CRC, 2017).

[33] L. Bouffaut, R. Dréo, V. Labat, A.-O. Boudraa, and G. Barruol, "Passive stochastic matched filter for antarctic blue whale call detection," The Journal of the Acoustical Society of America **144**(2), 955–965 (2018).

[34] A. Pereira, D. Harris, P. Tyack, and L. Matias, "Fin whale acoustic presence and song characteristics in seas to the southwest of portugal," The Journal of the Acoustical Society of America **147**(4), 2235–2249 (2020).

[35] L. Matias and D. Harris, "A single-station method for the detection, classification and location of fin whale calls using ocean-bottom seismic stations," The Journal of the Acoustical Society of America **138**(1), 504–520 (2015).

[36] D. Harris, L. Matias, L. Thomas, J. Harwood, and W. H. Geissler, "Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast atlantic," The Journal of the Acoustical Society of America **134**(5), 3522–3535 (2013).

[37] N. Feuillet, S. Jorry, W. C. Crawford, C. Deplus, I. Thinon, E. Jacques, J. M. Saurel, A. Lemoine, F. Paquet, C. Satriano, *et al.*, "Birth of a large volcanic edifice offshore mayotte via lithosphere-scale dyke intrusion," Nature Geoscience **14**(10), 787–795 (2021).

[38] A. Lemoine, P. Briole, D. Bertil, A. Roullé, M. Foumelis, I. Thinon, D. Raucoules, M. de Michele, P. Valty, and R. Hoste Colomer, "The 2018–2019 seismo-volcanic crisis east of mayotte, comoros islands: seismicity and ground deformation markers of an exceptional submarine eruption," Geophysical Journal International **223**(1), 22–44 (2020).

[39] J.-M. Saurel, E. Jacques, C. Aiken, A. Lemoine, L. Retailleau, A. Lavayssière, O. Foix, A. Dofal, A. Laurent, N. Mercury, *et al.*, "Mayotte seismic crisis: building knowledge in near real-time by combining land and ocean-bottom seismometers, first results," Geophysical Journal International **228**(2), 1281–1293 (2021).

[40] F. Samaran, C. Guinet, O. Adam, J.-F. Motsch, and Y. Cansi, "Source level estimation of two blue whale subspecies in southwestern indian ocean," The Journal of the Acoustical Society of America **127**(6), 3800–3808 (2010).

[41] A. Širović, J. A. Hildebrand, and S. M. Wiggins, "Blue and fin whale call source levels and propagation range in the southern ocean," The Journal of the Acoustical Society of America **122**(2), 1208–1215 (2007).

[42] S. Arcuti, C. Calculli, A. Pollice, G. D'Onghia, P. Maiorano, and A. Tursi, "Spatio-temporal modelling of zero-inflated deep-sea shrimp data by tweedie generalized additive," Statistica **73**(1), 87–101 (2013).

[43] N. Cressie, M. Sainsbury-Dale, and A. Zammit-Mangion, "Basis-function models in spatial statistics," Annual Review of Statistics and Its Application **9**(1), 373–400 (2022).

[44] H. Akaike, "A new look at the statistical model identification," IEEE transactions on automatic control **19**(6), 716–723 (1974).

[45] S. Wood, *Generalized Additive Models: An Introduction with R*, 2 ed. (Chapman and Hall/CRC, 2017).

[46] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2021), https://www.R-project.org/.

[47] J.-Y. Royer, "Oha-sis-bio-observatoire hydroacoustique" (2009), doi: 10.18142/229.

[48] A. Širović, J. A. Hildebrand, S. M. Wiggins, M. A. McDonald, S. E. Moore, and D. Thiele, "Seasonality of blue and fin whale calls and the influence of sea ice in the western antarctic peninsula," Deep Sea Research Part II: Topical Studies in Oceanography **51**(17-19), 2327–2344 (2004).

[49] K. M. Stafford, G. Boussarie, M. Caputo, L. Irvine, S. Laing, E. Nancy, H. Pearson, and J. J. Kiszka, "Acoustic detections and sightings of blue whales balaenoptera musculus in the seychelles, western tropical indian ocean (2020-2022)," Endangered Species Research **52**, 203–208 (2023).

[50] T. Dominello and A. Širović, "Seasonality of antarctic minke whale (balaenoptera bonaerensis) calls off the western antarctic peninsula," Marine Mammal Science **32**(3), 826–838 (2016).

[51] F. W. Shabangu, K. Findlay, and K. M. Stafford, "Seasonal acoustic occurrence, diel-vocalizing patterns and bioduck call-type composition of antarctic minke whales off the west coast of south africa and the maud rise, antarctica," Marine Mammal Science **36**(2), 658–675 (2020).

[52] S. Cerchio, S. Laran, N. Andrianarivelo, A. Saloma, B. Andrianantenaina, O. Van Canneyt, and T. Rasoloarijao, "Cetacean species diversity in malagasy waters," The New Natural History of Madagascar **15**, 411–23 (2022).

[53] R. Yan, J. Wang, W. Ju, D. S. Goll, A. K. Jain, S. Sitch, H. Tian, P. Benjamin, F. Jiang, and H. Wang, "Interactive effects of the el niño-southern oscillation and indian ocean dipole on the tropical net ecosystem productivity," Agricultural and Forest meteorology **336**, 109472 (2023).

2.5 Stage de Master : Combining passive and active acoustic approaches to study predator-prey interactions of blue whales in the southern Indian Ocean

Dans le cadre de ce chapitre, un stage de Master 2 a été réalisé afin d'explorer la possibilité d'intégrer des mesures directes de biomasse de proies dans les modèles de distribution des baleines bleues. Ce stage a été motivé par le fait que les deux projets (THEMISTO et OHASISBIO) se déroulent dans la même zone géographique (le sud-ouest de l'OI) et se recoupent temporellement. L'objectif principal était de tester si les données acoustiques actives pouvaient être intégrées comme variables environnementales expliquant la présence acoustique des baleine bleue Antarctique (ANT BW) pour mieux comprendre les interactions prédateur-proie de la zone.

Le programme THEMISTO (*Toward Hydroacoustics and Ecology of Mid-trophic levels in Indian and Southern Ocean*, Cotte 2016) a réalisé des mesures de zooplancton et micronekton par acoustiques actives depuis le Marion Dufresne 2. Les campagnes de relevés ont eu lieu durant l'été austral, de janvier à mars, entre 2021 et 2023. Des échosondeurs multifréquences (SIMRAD, Norway) fonctionnant à 18, 38, 70, 120 et 200 kHz ont été utilisés pour quantifier, à des stations ponctuelles le long des transects, la distribution et la densité du zooplancton et du micronekton (Fig. 2.20).

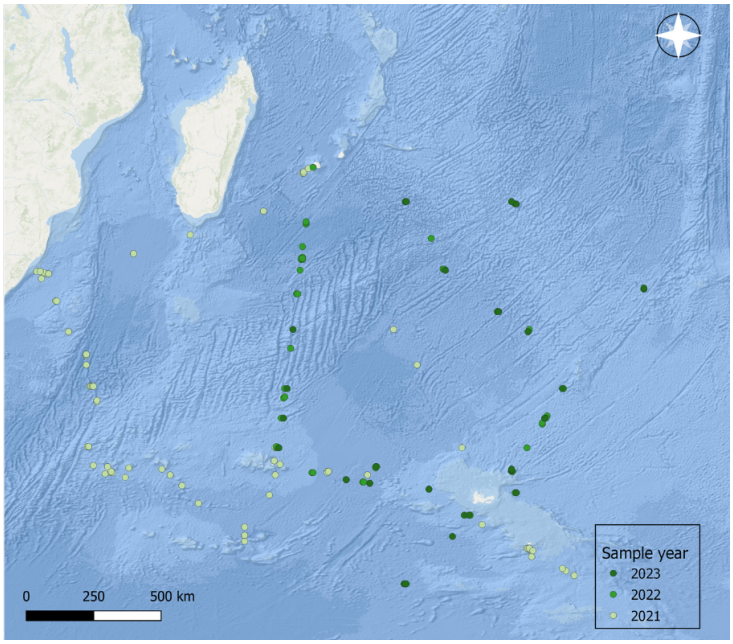


FIGURE 2.20 – Points de mesure d'acoustique active du programme THEMISTO le long des transects lors des campagnes 2021–2023. Les points sont codés par année : vert clair pour 2021, vert moyen pour 2022 et vert foncé pour 2023.

Les mesures couvraient jusqu'à 225 m de profondeur, correspondant aux plongées typiques des baleines bleues (GOLDBOGEN, CALAMBOKIDIS et al. 2011). Pour cette étude, signal à 120 kHz a été utilisé puisqu'il reflète le mieux la biomasse de la famille des euphausiidés, principale proie des baleines bleues (BENOIT-BIRD et LAWSON 2016). Ces mesures *in situ* ont ensuite été transformées en profils de rétro-diffusion acoustique et agrégées sous forme de *Nautical Area Scattering Coefficient (NASC)*[6]. Pour produire des cartes à l'échelle régionale, ces observations ponctuelles ont été étendues spatialement via des modèles *Random Forest*, utilisant comme variables explicatives des paramètres environnementaux (température de surface, salinité, chlorophylle a, oxygène dissous). Les modèles ont été calibrés séparément pour le jour et la nuit afin de prendre en compte la migration verticale des organismes (BRIERLEY 2014). La méthodologie employée pour le traitement et la modélisation des données suit l'approche développée par Lloyd IZARD 2023 (Sorbonne Université) dans sa thèse : Structuration spatiale et variabilité des écosystèmes mésopélagiques dans l'Océan Indien Sud, combinant analyse fonctionnelle des profils

6. Somme de la rétro-diffusion acoustique par unité de surface nautique et est utilisé comme indicateur de la densité de zooplancton et micronekton (proxy pour la biomasse des proies).

acoustiques et modélisation statistique pour produire des cartes de distribution de la biomasse mésopélagique.

Acoustique passive vs acoustique active

Acoustique passive : Cette méthode enregistre l'ensemble du paysage sonore sous-marin, incluant les sons biologiques (comme ceux des cétacés), mais aussi les sons anthropiques ou abiotiques, sans émettre de signal.
Acoustique active : Cette approche consiste à émettre un signal sonore et à mesurer l'écho réfléchi par les objets ou organismes rencontrés à l'aide d'un échosondeur mono ou multifréquences. Elle permet d'estimer la densité et la distribution des proies, comme le zooplancton et le micronekton. Elle est également utilisée pour la bathymétrie et la cartographie des fonds marins.

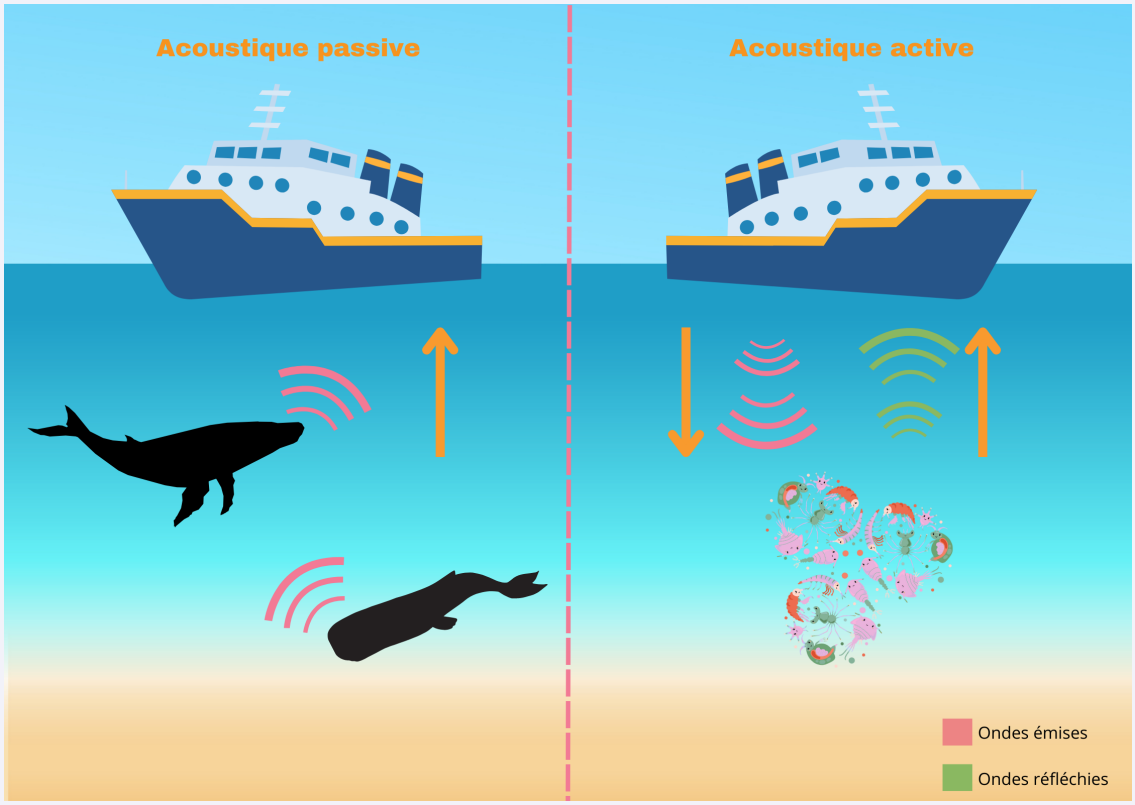


FIGURE 2.21 – Principe de l'acoustique passive vs acoustique active.

Le stage avait pour objectifs d'évaluer la variabilité spatio-temporelle de la biomasse de zooplancton et micronekton mesurée par THEMISTO, d'examiner la corrélation entre cette biomasse et la présence acoustique des ANT BW enregistrée par le réseau OHASISBIO, et de tester la faisabilité d'intégrer ces données dans des modèles de distribution spatio-

temporelle, en lien avec la disponibilité en proies. Pour ce faire, les anomalies hebdomadaires de NASC ont été calculées pour chaque site et comparées aux anomalies hebdomadaires de présence acoustique de ANT BW mesurées en DPH.

Les principaux résultats montrent que la biomasse de zooplancton et micronekton présente un gradient latitudinal, avec des valeurs plus élevées dans les zones sub-tropicales que dans les zones sub-Antarctiques, cohérentes avec la littérature précédente (Béhagle et al. 2016 ; Boersch-Supan, Rogers et Brierley 2017). Les détections acoustiques des ANT BW présentent une grande variabilité inter-annuelle et inter-sites, avec un nombre d'heures de détection plus élevé dans les zones sub-Antarctiques et certains sites spécifiques. Une corrélation négative significative a été observée entre anomalies de biomasse et anomalies de DPH pour l'année 2021, suggérant que certaines baleines peuvent se trouver en zones moins productives, mais aucune tendance globale claire liant la présence des baleines et des proies n'émerge sur l'ensemble des trois années. Ces résultats mettent en évidence la complexité des stratégies migratoires, montrant à la fois une migration partielle et différenciée, pouvant être influencée par l'âge, le sexe, l'état physiologique et la disponibilité des proies.

Plusieurs limites méthodologiques ont été identifiées dans cette étude. Les NASC utilisés incluent plusieurs groupes de zooplancton et ne se limitent pas aux euphausiidés. De plus, les délais trophiques entre la production primaire et la biomasse de zooplancton ne sont pas pris en compte. La présence de Z-calls semble principalement associée à des mâles et pourrait ne pas refléter directement l'activité de nourrissage ; l'analyse de la présence acoustique des baleines via les vocalisations de type D-calls permettrait d'explorer plus la présence acoustique des baleines ayant un comportement d'alimentation. Enfin, les variations interannuelles et les incertitudes liées aux modèles *Random Forest* peuvent limiter la précision des prédictions. Il convient également de souligner que les campagnes THEMISTO ne couvrent que l'été austral (janvier à mars), ce qui restreint la disponibilité des données à cette période. Pour mieux comprendre les stratégies migratoires et les interactions prédateur-proie sur l'ensemble de l'année, il serait souhaitable de disposer de mesures acoustiques actives couvrant toutes les saisons.

Malgré ces limites, ce stage a permis de réaliser la première intégration entre les deux principaux jeux de données de l'océan Indien : THEMISTO et OHASISBIO. La combinaison d'acoustique passive et active s'avère ainsi être un outil prometteur pour

étudier les interactions prédateur-proie et les comportements migratoires des baleines bleues. Le rapport complet du stage est disponible en Annexe D de cette thèse.

2.6 Discussion

Dans un contexte où la compréhension des dynamiques spatio-temporelles des grands cétacés repose de plus en plus sur l'exploitation de données issues du PAM, ce chapitre avait pour objectif de caractériser les patrons de distribution des mysticètes dans l'OI à partir d'analyses statistiques robustes et flexibles, tout en intégrant des informations environnementales complémentaires. Cette approche s'inscrit dans l'enjeu central de la thèse visant à mieux documenter les interactions entre les cétacés et leur environnement, afin d'éclairer les processus écologiques sous-jacents et de renforcer la portée des analyses en écologie acoustique.

2.6.1 Aspects méthodologiques

L'approche statistique mobilisée a permis de capter des relations complexes entre la présence acoustique des baleines et leur environnement, et de révéler des dynamiques spatio-temporelles difficilement accessibles par d'autres méthodes. L'utilisation des GAM a fourni une représentation souple et robuste, tandis que la combinaison avec l'analyse en EOF et le *clustering* hiérarchique a offert une lecture synthétique des habitats fonctionnels et des saisons écologiques.

Les EOF ont été déjà utilisées dans le domaine halieutique pour identifier les principaux modes spatio-temporels dans les séries de distribution de poissons. Par exemple, Alglave et al. 2024 ont appliqué cette méthode à des données de captures commerciales pour trois espèces de poissons dans le golfe de Gascogne, afin de synthétiser les signaux spatio-temporels liés à la phénologie de reproduction et aux habitats essentiels. Dans notre contexte, l'application des EOF à des données acoustiques de cétacés permet de transposer cet outil éprouvé à un système écologique différent, en capturant de manière synthétique les dynamiques de présence acoustique sur plusieurs sites et années, ce qui n'avait, à notre connaissance, jamais été réalisé auparavant dans ce domaine.

Les deux principales études conduites montrent la complémentarité des approches méthodologiques selon les contextes écologiques et les objectifs analytiques. Toutes deux reposent sur des suivis pluriannuels, mais leur protocole diffère : l'étude 1 (OHASISBIO), menée sur dix sites, vise à cartographier à grande échelle la présence acoustique des populations de baleines bleues (ANT BW, SEIO PBW, SWIO PBW) dans l'OI subantarctique. Des DPM hebdomadaires et une grille de 50×50 km y sont employés, ce qui permet de capturer finement la variabilité spatiale et saisonnière et d'extraire des modes spatio-temporels via l'EOF. À l'inverse, l'étude 2 (REVOSIMA), centrée sur un seul site au nord du canal du Mozambique, cherche à caractériser les habitats écologiques d'une zone tropicale spécifique. Des DPH, une résolution spatiale plus large (130×130 km) ainsi que l'intégration de l'interaction entre variables environnementales et année y sont incluses, afin de saisir la variabilité interannuelle et la dynamique intra-annuelle (`Week`) de différentes espèces de baleines à fanons (ANT BW, SWIO PBW, AMW, FW).

Ces différences méthodologiques ont plusieurs implications :

— La granularité temporelle et le type de données (DPM vs DPH) influencent la sensibilité des modèles aux signaux rares ou sporadiques.

— L'inclusion d'interactions avec l'année dans l'étude 2 permet de révéler des variations interannuelles spécifiques à chaque espèce, tandis que l'étude 1 privilégie une approche plus globale pour capter les tendances générales et étendre des prédictions.

— La résolution spatiale et la portée de détection choisies reflètent l'échelle écologique des populations étudiées et la densité des enregistreurs, affectant la finesse et les possibilités de prédictions.

— L'approche multi-espèces permet de comparer des stratégies écologiques contrastées, enrichissant l'interprétation des habitats critiques et des corridors migratoires.

— Les deux dispositifs reposent sur des suivis pluriannuels, mais l'étude 1, avec dix sites, permet des extrapolations spatiales et une cartographie à large échelle, tandis que l'étude 2, centrée sur un seul site, se prête davantage à l'analyse fine des dynamiques locales.

L'utilisation des GAM avec distribution Tweedie s'est révélée particulièrement adaptée

aux grandes baleines et, plus largement, à l'étude d'espèces rares. Elle permet de gérer la dispersion excessive des données (*overdispersion*[7]) et d'offrir un compromis robuste face au risque de surajustement (*overfitting*[8]). Ces propriétés sont cruciales pour modéliser efficacement la distribution malgré un faible nombre d'observations, notamment pour des populations peu détectées ou vocalisant rarement dans l'OI. Ces conclusions sont en accord avec Virgili 2018 et d'autres auteurs (Redfern et al. 2006 ; Mannocci et al. 2014 ; Shono 2008), qui soulignent la pertinence des GAM/Tweedie pour les espèces rares, avec des déviances expliquées souvent supérieures à 25%.

Bien que d'autres approches existent (e.g. modèles zéro-inflated (Welsh et al. 1996 ; Cunningham et Lindenmayer 2005), MaxEnt (Elith, Phillips et al. 2011), modèles hiérarchiques bayésiens (Latimer et al. 2006) ou modèles mécanistes (Morin et Thuiller 2009)), les GAM représentent un compromis robuste pour les données avec effort connu, conciliant flexibilité, interprétabilité et performance prédictive. Elles ouvrent également des perspectives pour la standardisation du PAM à grande échelle, en proposant un cadre analytique réplicable dans d'autres zones géographiques et adaptable selon le type de populations et le contexte écologique.

Cependant, plusieurs limites méthodologiques doivent être soulignées. D'une part, la grande flexibilité des GAM peut rendre les modèles difficiles à interpréter. En effet, lorsque les relations entre variables sont très complexes et non linéaires, elles deviennent moins lisibles. De plus, si les contraintes de lissage sont trop faibles, un risque de surajustement apparaît. D'autre part, ces modèles reposent sur l'hypothèse que les relations environnement–présence acoustique sont généralisables, alors que les réponses écologiques peuvent varier fortement selon les zones et les saisons. Cette absence de validation de la transférabilité spatiale constitue une limite importante pour l'usage prédictif comme souligné par Auriane Virgili 2018 dans sa thèse : *Modelling distributions of rare marine species : The deep-diving cetaceans*, des modèles calibrés dans une sous-région peuvent ne pas être applicables à des zones écologiquement distinctes.

Un autre frein majeur concerne l'absence d'informations directes sur la distribution des proies. Les travaux réalisés lors du stage de Master 2 ont montré que l'intégration de mesures

7. Variance des données beaucoup plus élevée que celle prédite par le modèle statistique, ce qui biaise l'ajustement et les tests de significativité.

8. Situation où le modèle s'adapte trop étroitement aux données observées, y compris au bruit, perdant ainsi sa capacité à généraliser et à prédire dans de nouveaux contextes.

in situ de biomasse de zooplancton et micronekton, issues du programme THEMISTO, combinées à la présence acoustique des baleines, permettait de tester des corrélations prédateur-proie. Même si les résultats n'ont pas révélé de lien clair sur l'ensemble des années, cette expérience démontre la valeur ajoutée de coupler des données passives et actives pour mieux appréhender les interactions écologiques. L'intégration future de sorties issues de modèles numériques comme SEAPODYM (Lehodey, Senina et Sibert 2023), simulant la biomasse et la distribution des proies, constituerait un progrès notable pour rapprocher les conditions environnementales distales des réponses écologiques observées.

2.6.2 Aspect écologique

Les résultats présentés renforcent la compréhension des stratégies spatio-temporelles des mysticètes dans l'OI, mettant en évidence la complexité et la diversité de leurs adaptations écologiques face à un environnement océanique hautement dynamique. L'identification de zones écologiquement critiques telles que le plateau de Kerguelen, le bassin de Madagascar ou le nord du canal du Mozambique corrobore leur rôle essentiel pour différentes fonctions biologiques, notamment comme zones d'alimentation, corridors migratoires ou aires de reproduction. Ces observations confirment le rôle central des structures océanographiques de méso- et de fine échelle (bathymétrie, upwelling, concentrations locales de krill) dans l'organisation spatiale des cétacés. Des exemples issus d'autres populations, comme les rorquals communs en mer Méditerranée ou les baleines à bosse en Antarctique, montrent que la disponibilité locale de ressources et les conditions physiques guident également leur distribution et leur activité (Notarbartolo-Di-Sciara et al. 2003 ; Andrews-Goff, Bestley et al. 2018). Ces comparaisons illustrent que la sensibilité aux conditions environnementales est un trait commun aux grands cétacés, renforçant la pertinence de nos observations dans l'OI.

Néanmoins, certaines limites doivent être prises en compte. Comme dans d'autres études acoustiques, les résultats reflètent une présence acoustique plutôt qu'une présence absolue d'individus. Par exemple, sur 121 déploiements de tags de baleines bleues au large de la Californie, seuls 22% contenaient des sons attribués à l'individu identifié et 65% aucun son (Lewis et al. 2018). Ces chiffres illustrent que l'absence de sons ne signifie pas nécessairement l'absence de l'animal, en raison de la variabilité comportementale dans l'émission des sons.

La variabilité saisonnière et interannuelle constitue un autre défi, car elle est difficile à caractériser sans séries temporelles étendues. Dans le Pacifique Est, l'analyse des mouvements migratoires des baleines bleues a montré que les corrélations robustes avec la productivité (chlorophylle-a) n'émergent qu'en comparant les déplacements aux moyennes climatologiques sur dix ans, plutôt qu'aux conditions contemporaines (Abrahms et al. 2019). Ces limites permettent de contextualiser la portée et l'interprétation des résultats acoustiques.

Enfin, la forte hétérogénéité écologique entre les différentes zones de l'OI suggère que les relations entre espèces et environnement peuvent varier localement. L'océan Austral, par exemple, comprend plusieurs bassins aux caractéristiques écologiques distinctes, avec des réseaux alimentaires régionaux différenciés par la dynamique des flux d'énergie, la composition et l'abondance relative des espèces, ainsi que par leur résilience face aux perturbations climatiques ou anthropiques (McCormack et al. 2021). Tester la transférabilité des modèles entre sous-régions permettrait de mieux identifier les mécanismes écologiques sous-jacents. À plus grande échelle, ce type d'approche comparative pourrait ouvrir la voie à des prédictions plus robustes sur la distribution et le comportement des grands cétacés dans l'océan Indien et au-delà, en intégrant la variabilité régionale et saisonnière des ressources.

Chapitre 3

Structurer les suivis par acoustique passive

3.1 Introduction

Ces dernières années, le suivi par acoustique passive (PAM) s'est fortement développé à l'échelle européenne, en réponse à une demande croissante d'outils capables de documenter la présence et l'usage de l'habitat des cétacés sur le long terme et sur de vastes zones. Cette dynamique s'inscrit dans un contexte de structuration progressive des politiques publiques de conservation marine, notamment à travers la Directive Cadre Stratégie pour le Milieu Marin (DCSMM), qui fixe des exigences précises en matière de suivi du milieu marin en Europe.

À l'échelle locale, de nombreux projets ont démontré la capacité du PAM à documenter les schémas d'activité, la saisonnalité ou les migrations des cétacés. Cependant, face aux enjeux de conservation marine à grande échelle, une question cruciale émerge :

Comment structurer et standardiser les suivis PAM pour en faire un outil robuste, comparable et exploitable dans le cadre des politiques de conservation internationales ?

Ce chapitre propose une réflexion intégrée sur les conditions de mise en œuvre d'un suivi PAM standardisé, fiable et reproductible à l'échelle européenne. Il s'appuie sur plusieurs niveaux d'analyse : (i) un état des lieux du cadre réglementaire et de la place actuelle du PAM dans la DCSMM, (ii) une discussion des verrous liés à l'hétérogénéité des pratiques, et (iii) une restitution des échanges émanant d'un atelier de travail européen organisé en

2023 à la Conférence de l'*European Cetacean Society*[1], dont les conclusions ont été publiées dans la revue *Marine Policy* en 2024 : « *Passive acoustic monitoring for assessing marine mammal populations in European waters : Workshop conclusions and perspectives* ».

Cette synthèse constitue une forme de conclusion appliquée de la thèse. Elle vise à identifier des leviers concrets pour améliorer la cohérence des approches PAM, à partir des enseignements tirés des travaux menés et des besoins identifiés par la communauté scientifique et les gestionnaires. Sans prétendre répondre à l'ensemble des défis, les outils développés au fil de cette thèse contribuent à éclairer certains choix méthodologiques et à poser les bases d'une standardisation à venir.

3.1.1 La DCSMM comme cadre de gestion

La **Directive Cadre Stratégie pour le Milieu Marin** (DCSMM, Directive 2008/56/CE) constitue l'outil central de l'Union européenne pour la protection, la préservation et la restauration du milieu marin. Elle vise à atteindre, ou à maintenir, le Bon État Écologique (BEE) des eaux marines européennes, conformément aux principes du développement durable.

Le BEE est défini comme suit :

> « L'état écologique des eaux marines tel que celles-ci conservent la diversité écologique, le fonctionnement et le dynamisme des océans et des mers sans impact des pressions anthropiques. L'utilisation du milieu marin doit être durable, sauvegardant ainsi son potentiel aux fins des utilisations et activités des générations actuelles et à venir. » (Directive 2008/56/CE)

Reposant sur une approche dite écosystémique, la DCSMM prend en compte l'ensemble des interactions entre les espèces, leurs habitats et les pressions d'origine humaine. Sa mise en œuvre se déploie à travers des cycles de planification de six ans. Trois cycles ont déjà eu lieu : 2012–2018, 2018–2024, et le cycle actuel couvre 2024–2030. Durant chaque cycle, les États membres sont tenus de :

— **Évaluer l'état initial** de leurs domaines maritimes, en identifiant les pressions existantes ainsi que les caractéristiques écologiques majeures ;

1. L'*European Cetacean Society* (ECS), créée en janvier 1987, a pour objectif de promouvoir et de développer les études scientifiques et les efforts de conservation des mammifères marins, ainsi que de collecter et diffuser l'information à ses membres et au grand public. L'ECS est coordonnée par un conseil de 11 membres et organise chaque année une conférence annuelle dans un pays européen.

— **Définir le BEE** à atteindre pour les différentes composantes du milieu marin ;
— **Fixer des objectifs environnementaux** et mettre en place des programmes de surveillance pour suivre les évolutions de l'état du milieu ;
— **Mettre en œuvre des programmes de mesures**, destinés à réduire les pressions identifiées et à restaurer les écosystèmes si nécessaire.

Pour structurer cette évaluation, la directive s'appuie sur 11 descripteurs correspondant chacun à un enjeu environnemental majeur (e.g. biodiversité, bruit sous-marin, déchets, contaminants, etc.) (Fig.3.1). Ces descripteurs fournissent une grille d'analyse qualitative et quantitative de la santé des écosystèmes marins.



FIGURE 3.1 – 11 descripteurs utilisés pour évaluer le BEE dans la DCSMM.

Parmi eux, le **Descripteur 1**, consacré à la « diversité biologique », occupe une place centrale. Il vise à garantir que les espèces marines et leurs habitats soient conservés dans un état compatible avec les conditions naturelles du milieu. La directive précise ainsi que :

> « La diversité biologique est conservée. La qualité des habitats et leur nombre, ainsi que la distribution et l'abondance des espèces, sont adaptées aux conditions physiographiques, géographiques et climatiques existantes. » (Directive 2008/56/CE)

Dans ce contexte, les mammifères marins jouent un rôle d'indicateurs clés de l'état écologique des écosystèmes. En effet, ces espèces sentinelles du milieu marin sont particulièrement vulnérables aux pressions anthropiques telles que la pollution, les collisions avec les navires, le bruit sous-marin, et les captures accidentelles dans les engins de pêches.

Leur suivi permet non seulement de mesurer l'impact des activités humaines, mais aussi d'orienter les politiques publiques de gestion environnementale.

Les façades maritimes évaluées

Dans le cadre de la DCSMM, les eaux marines européennes sont subdivisées en grandes zones géographiques appelées façades maritimes (Fig. 3.2). En France, on distingue quatre sous-régions, chacune disposant de son propre plan d'action :

— Manche - Mer du Nord
— Mers Celtiques
— Golfe de Gascogne
— Méditerranée occidentale

L'évaluation du BEE est donc conduite façade par façade, afin de mieux refléter les spécificités écologiques et les pressions propres à chaque région sous-marine.

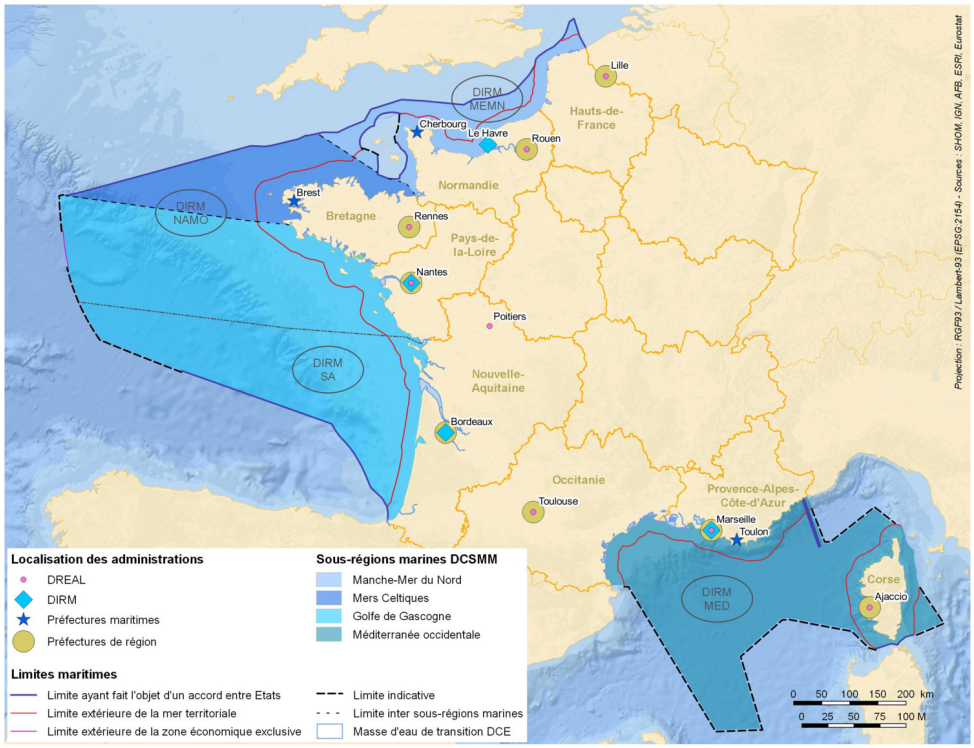


FIGURE 3.2 – Façades maritimes françaises de la DCSMM (IFREMER 2020).

Définition du BEE pour la composante « Mammifères marins »

L'évaluation du BEE pour les mammifères marins est encadrée par la décision 2017/848/UE. Celle-ci décline le Descripteur 1 en plusieurs critères, répartis selon cinq grands axes :

- **D1C1 :** Le taux de mortalité causée par les activités humaines (notamment les captures accidentelles) doit rester suffisamment bas pour ne pas compromettre la viabilité des populations.
- **D1C2 :** Les populations doivent présenter une abondance suffisante pour garantir leur bon état de conservation.
- **D1C3 :** Les événements de mortalité inhabituels ou extrêmes (comme les échouages massifs) doivent rester rares et limités.
- **D1C4 :** La distribution spatiale des individus ne doit pas être réduite ni fragmentée de manière significative par rapport à la situation naturelle.
- **D1C5 :** Les habitats essentiels doivent offrir une qualité suffisante pour répondre aux besoins biologiques des espèces.

Pour répondre à ces exigences, les États membres doivent identifier les espèces les plus représentatives de leurs façades maritimes, fixer des seuils écologiques pertinents et coordonner leurs efforts à l'échelle régionale. Cette coordination implique également une cohérence avec d'autres politiques européennes, telles que la Directive Habitats (92/43/CEE), et une articulation avec d'autres descripteurs de la DCSMM (par exemple, le D8 pour les contaminants, le D10 pour les déchets ou le D11 pour le bruit sous-marin).

Groupes d'espèces représentatives

Pour le suivi du Descripteur 1, quatre grands groupes de mammifères marins sont considérés comme représentatifs de la biodiversité marine européenne :

- **Les phoques** (e.g. phoque gris *Halichoerus grypus*, phoque veau-marin *Phoca vitulina*) ;
- **Les petits odontocètes** (e.g. marsouin commun *Phocoena phocoena*, grand dauphin *Tursiops truncatus*) ;

— **Les Odontocètes grands plongeurs** (e.g. dauphin de Risso *Grampus griseus*) ;
— **Les mysticètes** (e.g. petit rorqual *Balaenoptera acutorostrata*).

Ces groupes sont évalués selon leur présence, leur abondance, leur répartition spatiale et leur état de santé dans les différentes façades, fournissant ainsi les données nécessaires au renseignement des critères du descripteur.

Méthodes d'évaluation et sources de données

L'évaluation repose sur un ensemble de méthodes scientifiques basées principalement sur des campagnes d'observations visuelles réalisées depuis des avions ou des navires. Ces campagnes suivent des protocoles standardisés et permettent d'estimer la densité, la distribution et les tendances temporelles des populations. Cependant, l'évaluation du BEE demande une connaissance précise de la distribution et de l'abondance, critères difficiles à documenter pour ces espèces très mobiles et à faible densité. De plus, la définition d'un état de référence fiable demeure complexe, souvent basée sur la première estimation disponible qui peut déjà être affectée par des perturbations anthropiques. Les évaluations doivent aussi être cohérentes spatialement et temporellement avec les comportements des populations, ce qui nécessite des suivis à grande échelle, complexes et coûteux à mettre en œuvre. Ainsi, les données disponibles restent souvent incomplètes et parcellaire.

Les données et collaborations proviennent de sources nationales, régionales et internationales, et contribuent à la fois à la collecte d'informations et à l'harmonisation des méthodes :

— Programmes nationaux, comme le Réseau National Échouage (RNE) pour le suivi des mortalités (Office français de la biodiversité (OFB) 1991) ;
— Campagnes aériennes, comme SAMM (*Surveillance Aérienne de la Mégafaune Marine*, Observatoire PELAGIS – CNRS 2013) ;
— OSPAR (*Oslo-Paris Convention*, OSPAR Commission 1992), qui définit des indicateurs communs et coordonne le suivi à l'échelle de l'Atlantique Nord-Est ;
— ASCOBANS (*Agreement on the Conservation of Small Cetaceans of the Baltic, North East Atlantic, Irish and North Seas*, ASCOBANS 1994) et ACCOBAMS (*Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea*

and Contiguous Atlantic Area, ACCOBAMS 2001), pour la conservation régionale des cétacés ;

— ICES (*International Council for the Exploration of the Sea*, Ifremer 1902), à travers les groupes WGMME (écologie des mammifères marins) et WGBYC (prises accidentelles) ;

— Groupes d'experts européens dédiés aux phoques et aux cétacés, qui assurent la coordination scientifique et la définition des seuils d'évaluation.

Les résultats sont ensuite comparés à des seuils de référence. L'intégration des données suit la logique du « *One Out All Out* » (OOAO) : si un seul indicateur relatif à une espèce ne respecte pas le seuil fixé pour le BEE, alors l'ensemble du descripteur est considéré comme non atteint pour la sous-région concernée. Cette collaboration étroite permet d'harmoniser les approches, de mutualiser les connaissances et de garantir une gestion cohérente à l'échelle des régions marines.

3.1.2 Le suivi par acoustique passive comme données complémentaires

À ce jour, les données issues du PAM ne sont pas encore intégrées dans les indicateurs du Descripteur 1, malgré un potentiel désormais largement reconnu. Plusieurs projets pilotes à l'échelle européenne et régionale illustrent concrètement les apports du PAM dans une optique de suivi écologique. Le projet SAMBAH (*Static Acoustic Monitoring of the Baltic Sea Harbour Porpoise*) a, par exemple, mobilisé plus de 300 C-PODs entre 2011 et 2013, déployés par l'ensemble des pays riverains de la mer Baltique (Suède, Finlande, Estonie, Danemark, etc.), afin de suivre les populations de marsouins communs dans cette région (SAMBAH Project 2023). Ce projet a permis d'estimer l'abondance de la population (environ 497 individus), mais aussi d'identifier des zones critiques pour la reproduction et l'alimentation (Carlström et al. 2011). Ces résultats ont été utilisés pour appuyer la désignation de sites Natura 2000 [2], démontrant ainsi l'utilité directe du PAM dans la mise en œuvre des politiques européennes de conservation.

Un autre exemple majeur est le projet BlackCeTrends, une initiative internationale de

2. Réseau européen de sites protégés visant à préserver la biodiversité et les habitats naturels tout en conciliant activités humaines et conservation.

grande envergure dédiée à la surveillance acoustique des cétacés en mer Noire, réunissant six pays riverains (Turkish Marine Research Foundation 2021). Ce projet utilise en permanence des instruments F-PODs, spécialement conçus pour détecter les sons à haute fréquence, comme ceux émis par les marsouins communs de la mer Noire (*Phocoena phocoena relicta*). Entre septembre 2020 et octobre 2022, 19 F-PODs ont été déployés par plusieurs pays riverains de la mer Noire (e.g. Roumanie, Bulgarie, Turquie) (Fig. 3.3). Les données recueillies ont révélé des variations saisonnières et nycthémérales dans l'activité acoustique des marsouins, en lien avec les migrations de leurs proies comme l'anchois et le sprat (Ivanchikova et al. 2024). Ces données acoustiques complètent les observations visuelles régulières et contribuent à la protection des populations locales dans cet écosystème sensible. Lancé début 2021, BlackCeTrends illustre parfaitement le potentiel d'un suivi acoustique coordonné, standardisé et multi-pays à grande échelle.

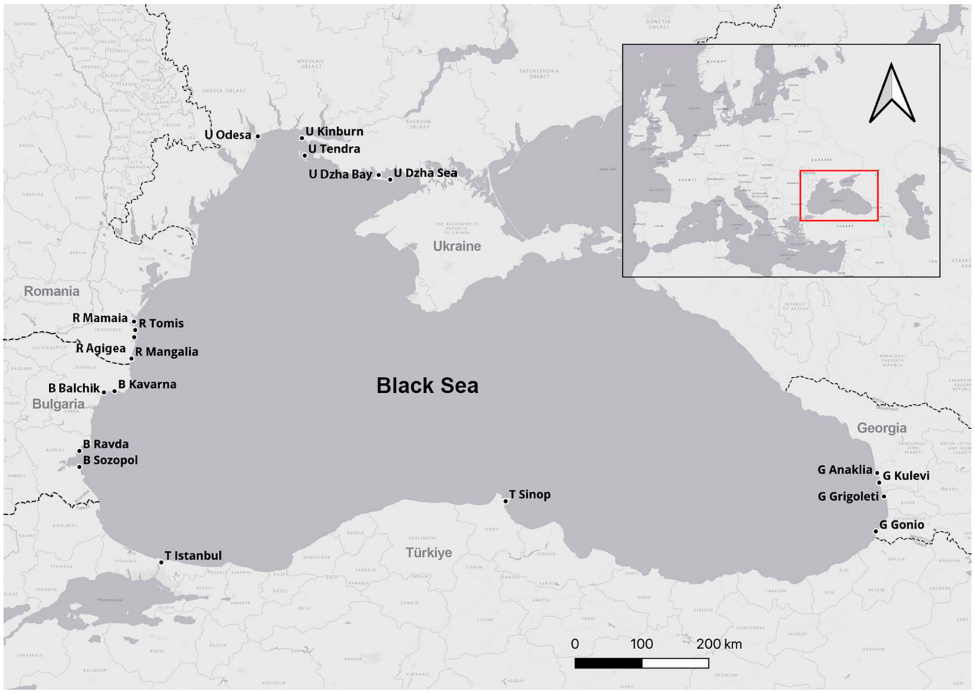


Figure 3.3 – Carte des emplacements des F-POD le long de la côte de la mer Noire (Ivanchikova et al. 2024).

De par les espèces ciblées, les projets mentionnés utilisent principalement des F-PODs, ce qui facilite la comparaison des données grâce à l'homogénéité des équipements et des protocoles. Cependant, l'acoustique passive peut aussi s'appuyer sur des enregistreurs large bande comme l'a démontré le programme CETIROISE en mer d'Iroise (ENSTA et Parc

Naturel Marin d'Iroise 2022). Ce projet, mené conjointement par le parc naturel marin d'Iroise [3] et l'ENSTA [4] en 2021-2022, a déployé plusieurs types d'enregistreurs acoustiques sur sept sites répartis entre zones côtières et hauturières (Fig. 3.4). Il vise à inventorier les différentes espèces de cétacés présentes, à évaluer leur fréquence d'utilisation et leur saisonnalité dans cette région. Le protocole combine des enregistreurs dédiés aux petits cétacés et des enregistreurs large bande enregistrant l'ensemble du paysage sonore sous-marin, permettant ainsi de détecter aussi bien les odontocètes que les mysticètes. Les analyses des données CETIROISE ont révélé des modèles de fréquentation distincts selon les saisons et les moments de la journée, confirmant notamment une présence continue des delphinidés au large d'Ouessant tout au long de l'année, ainsi que des pics saisonniers pour les marsouins et les petits rorquals (Samaran, Torterotot et Béesau 2023). Ces résultats illustrent la richesse et la complexité des communautés de cétacés dans le parc naturel marin, soulignant l'intérêt d'une surveillance acoustique continue et multi-échelles.

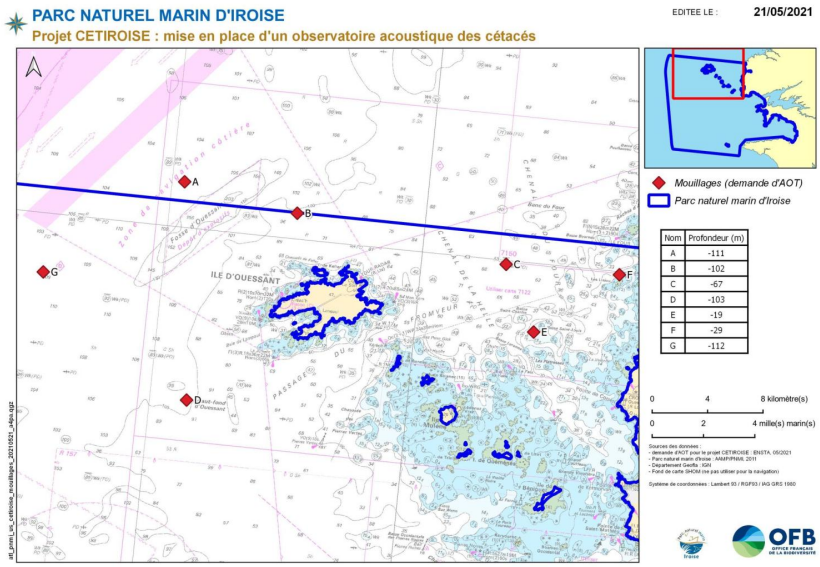


Figure 3.4 – Carte des emplacements des sites équipés d'enregistreurs large bande et de FPODs (point rouge) du projet CETIROISE dans le Parc Naturel Marin d'Iroise (délimité par le trait bleu) (ENSTA et Parc Naturel Marin d'Iroise 2022).

3. Créé en 2007, le Parc naturel marin d'Iroise protège les milieux marins autour de la Bretagne occidentale.

4. L'École Nationale Supérieure de Techniques Avancées (ENSTA) est une école d'ingénieurs et un pôle de recherche dans les domaines du maritime, du numérique, de la défense et sécurité, des transports et mobilités, de l'énergie et de la santé.

Ces exemples témoignent de la maturité croissante des méthodes PAM et de leur potentiel pour répondre aux exigences de la DCSMM. Ils montrent qu'un usage standardisé et coopératif à l'échelle régionale est non seulement possible, mais également porteur de résultats concrets et exploitables. Dans cette perspective, intégrer les données PAM au Descripteur 1 représenterait une avancée majeure, permettant de pallier plusieurs limites des suivis visuels actuels, notamment dans les zones reculées ou en offrant une couverture temporelle longue et continue.

3.1.3 Les défis liés à l'absence de standardisation

L'essor du PAM comme méthode de surveillance des mammifères marins s'est accompagné d'une grande diversité de pratiques, à toutes les étapes du processus, ce qui complique la comparaison des résultats et l'intégration harmonisée des données dans les cadres réglementaires comme la DCSMM.

Les principales sources de variabilité concernent la profondeur et le positionnement des hydrophones, la durée des campagnes, le traitement des données (détection, classification, seuils de validation), les indicateurs de présence (présence acoustique, taux de clics, indices d'abondance) et la qualité des métadonnées. Ces disparités s'expliquent par le manque de retours d'expérience consolidés, l'évolution rapide des technologies et le fait que le PAM est souvent conduit par des biologistes sans toujours disposer des compétences acoustiques nécessaires. Cette absence de standards empêche une agrégation fiable des données et peut introduire des biais dans l'évaluation du BEE.

Si le potentiel du PAM est désormais reconnu, sa mise en œuvre reste morcelée. L'absence d'une structure institutionnelle pérenne dédiée au PAM des mammifères marins constitue un frein majeur à l'émergence de standards partagés et à leur diffusion à l'échelle européenne. Contrairement à d'autres thématiques bénéficiant de groupes de travail structurés, comme le *Technical Group on Underwater Noise (TG Noise)*, qui travaillent spécifiquement sur les pressions acoustiques en lien avec le Descripteur 11 de la DCSMM relatif aux bruits sous-marins (Dekeling et al. 2014), les initiatives PAM reposent encore largement sur des dynamiques de projet, souvent temporaires, et portées par des consortiums hétérogènes. Cette fragmentation limite la coordination inter-états, la capitalisation des retours d'expérience, ainsi que la construction d'une vision stratégique

commune pour l'intégration du PAM dans les cadres réglementaires.

3.2 *Workshop* 2023 : consensus et recommandations

Afin de compléter la compréhension des enjeux liés à l'implication du PAM dans un cadre tel que la DCSMM, un atelier a été organisé à l'occasion de la 34ème conférence annuelle de l'*European Cetacean Society* qui s'est tenue à O'Grove en Galice, du 16 au 20 avril 2023. J'ai co-organisé cet événement avec Benjamin Guichard de l'Office Français de la Biodiversité (OFB) [5], au nom de l'ENSTA et de l'OFB.

Intitulé « *Prospective Discussion on Passive Acoustic Monitoring of Marine Mammals for Descriptor D1 of the Marine Strategy Framework Directive (MSFD)* », ce *workshop* a réuni 27 participants, chercheurs et gestionnaires, engagés dans des projets PAM dans différentes zones marines européennes (Fig. 3.5).



FIGURE 3.5 – Participants du *workshop* à l'ECS 2023.

5. Établissement public français chargé de la protection, de la connaissance et de la gestion de la biodiversité, incluant les espaces terrestres et marins, ainsi que de la mise en œuvre des politiques publiques de conservation.

L'objectif principal était de créer un espace de dialogue avec les États membres de l'Union européenne sur l'intégration de la PAM dans le suivi des cétacés, en lien avec l'évaluation du BEE au titre du Descripteur 1 (biodiversité) de la DCSMM. Les participants ont partagé leurs retours d'expérience à travers la présentation de projets existants, en cours de développement ou en phase exploratoire.

La seconde partie du *workshop* s'est concentrée sur les aspects méthodologiques, en particulier sur l'exploitation des données PAM dans l'évaluation du critère D1C5 (utilisation des habitats). Les discussions ont mis en évidence les enjeux d'harmonisation des protocoles, de standardisation des indicateurs et de consolidation d'un cadre commun facilitant l'intégration des données acoustiques dans les outils d'évaluation européens.

Le contenu détaillé, les échanges et les perspectives issues de cet atelier sont synthétisés dans un rapport publié sous forme d'article dans *Marine Policy : Passive acoustic monitoring for assessing marine mammal populations in European waters : Workshop conclusions and perspectives.*

Marine Policy xxx (xxxx) xxx

Contents lists available at ScienceDirect

Marine Policy

journal homepage: www.elsevier.com/locate/marpol

Conference reports

Passive acoustic monitoring for assessing marine mammals population in European waters: Workshop conclusions and perspectives

ARTICLE INFO

Keywords
Marine Strategy Framework Directive (MFSD)
Passive acoustic monitoring
Marine mammal
European working group

ABSTRACT

The European Marine Strategy Framework Directive (MSFD, 2008/56/CE) aims to achieve and maintain the 'good environmental status' (GES) of marine ecosystems, including the assessment of marine mammal populations. However, assessing their abundance and distribution is challenging due to their mobility, elusiveness, and difficult data acquisition. This article highlights the role of passive acoustic monitoring as a complementary approach to traditional methods (mainly plane and boat surveys). It also presents the conclusions and perspectives of a workshop organized at the 34th conference of the European Cetacean Society in O'Grove, Spain (18th to 20th April 2023). Building on the rich foundation established by the European Cetacean Society (ECS) in 1987, this society plays a pivotal role in advancing marine mammal research and conservation efforts across European waters. The workshop emphasized the need for harmonization of protocols, indicators, and guidelines to ensure comparability of results. The establishment of a dedicated European working group was proposed to coordinate passive acoustic monitoring efforts, foster collaboration, and share best practices. The workshop concluded that the consolidation of scientific knowledge, the development of common guidelines, and the establishment of a transnational working group would improve the understanding and conservation of these vulnerable species and their habitats.

1. European Cetacean Society context

Established in January 1987, the European Cetacean Society (ECS) is a leading organization dedicated to the scientific study and conservation of cetaceans, including whales, dolphins, and porpoises, in European waters. With a membership ranging from 350 to 550 individuals across over 25 European countries and beyond, ECS serves as a dynamic platform, uniting researchers, scientists, conservationists, and enthusiasts. The Society organizes an annual conference, a focal point for experts and stakeholders, providing a robust forum for cutting-edge research, discussions on cetacean sciences, and exchange of ideas.

The annual conference, a cornerstone of ECS's initiatives, serves as a focal point for experts and stakeholders in the field. This conference provides a robust forum for presenting cutting-edge research, discussing the latest developments in cetacean science, and exchanging ideas among professionals working towards the shared goal of cetacean conservation. Themes explored at ECS conferences cover a diverse range of topics such as cetacean behavior, ecology, population dynamics, conservation strategies, and the impact of human activities on marine mammal species. Alongside the annual conference, the ECS organizes and offers workshops on various specialized topics, providing a platform for in-depth discussions and hands-on experiences. It is within this framework that our study was conducted.

Given the influential role of ECS in the field of cetacean research and conservation, the conference not only serves as a platform for disseminating scientific knowledge but also as a catalyst for shaping policies in marine mammal conservation. Moreover, the ECS has demonstrated its commitment to proactive conservation efforts by actively engaging with policymakers. For instance, in 2017, ECS members composed a letter addressed to the Minister of Agriculture, Livestock, Rural Development, Fisheries, and Food in Mexico, highlighting the alarming situation of the Vaquita, a critically endangered species of porpoise. Additionally, the ECS has played a role in endorsing resolutions, as seen in the adoption of a significant resolution in 2008 concerning the accidental captures of porpoises in the Baltic Sea. These examples underscore the ECS's role not only as a scientific organization but also as an influential advocate for the protection and conservation of marine mammal species.

2. MSFD context

The Marine Strategy Framework Directive (MSFD, 2008/56/EC) aims to achieve or maintain 'Good Ecological Status' (GES) of the marine environment in Europe. GES is defined as a state in which marine ecosystems maintain their ecological diversity, functioning and dynamics without being adversely affected by anthropogenic pressures. Marine mammals are considered essential components of marine biodiversity and their long-term survival depends on the ecological status of the ecosystems on which they depend. These mammals, such as whales, dolphins and seals, are vulnerable to anthropogenic pressures, including by-catch, ship strikes, noise pollution and contamination by pollutants. They are therefore considered to be indicator species for the overall health of marine ecosystems. However, assessing their abundance and distribution is a scientific challenge due to their mobility, elusiveness and low densities. For some species available data often limited or incomplete, making it difficult to accurately measure their status [1]. To address this challenge, the MSFD uses a quantitative

https://doi.org/10.1016/j.marpol.2023.105983
Received 2 October 2023; Received in revised form 4 December 2023; Accepted 12 December 2023

indicator-based approach, where measurable parameters are monitored over time to assess the achievement of the GES. However, poor data availability and the lack of definition of consistent quantitative thresholds remain obstacles to build robust indicators. MSFD encourages the use of common indicators developed by regional conventions such as OSPAR or the Barcelona Convention. Where common indicators do not exist, national indicators can be developed to specifically assess uncovered criteria and species.

3. Contribution of passive acoustic data in the MSFD

Acoustic monitoring can play a crucial role in supplementing data for the assessment of good ecological status (GES) of cetaceans. Marine ecosystems are characterised by rich and complex acoustic environments in which sounds play a crucial role in the communication, foraging, reproduction and navigation of marine species [2]. Acoustic recordings allow the detection and study of specific sounds emitted by marine mammals [3], providing valuable information on their presence [3,4], abundance [5,6] and behaviour [7]. Passive acoustic techniques, such as hydrophones and autonomous acoustic recorders, are used to monitor the sound emissions of marine mammals over long periods of time and over large marine areas. These acoustic data complement information from other methods, such as distance sampling from plane and boat surveys, which may be limited in spatial and temporal coverage [8]. By integrating acoustic data into GES assessments, a more comprehensive picture of marine mammal occurrence and distribution activity could be obtained, contributing to a better understanding of the ecological status of marine ecosystems. This would also help to identify critical areas where specific conservation and management measures could be implemented to protect these species and maintain their habitat in a healthy state.

4. European Cetacean Society workshop: conclusions and perspectives

During the 34th Annual Conference of the European Cetacean Society in O'Grove, Galicia, ENSTA Bretagne and the French Biodiversity Agency organised a workshop entitled "Prospective Discussion on Passive Acoustic Monitoring of Marine Mammals for Descriptor D1 of the Marine Strategy Framework Directive (MSFD)". This workshop brought together 27 participants: scientist and managers in connection with passive acoustic monitoring projects in various areas of European waters.

The aim of the workshop was to discuss with other EU Member States the use of acoustics as a monitoring method for cetaceans under the MSFD and its integration into the assessment of Good Ecological Status (GES). Participants involved in or responsible for these monitoring programmes were invited to present ongoing, developing or future passive acoustic monitoring projects for marine mammals in line with MSFD biodiversity descriptor D1 in EU Member States.

In the second part of the workshop, the use and integration of acoustic data from these monitoring programmes into the assessment of Good Ecological Status for Descriptor D1 (e.g. for Criterion D1C5 on habitat use) was discussed. The aim was to attempt to synchronize studies by homogenizing data processing and descriptive measures in order to integrate these studies into indicators for assessing Good Ecological Status.

The workshop focused on the harmonization of protocols, acoustic indicators and guidelines for marine mammal monitoring within the MSFD. The various project presentations revealed a lack of standardization of methods, making it difficult to compare results between participating countries. It is therefore essential to establish a consistent and standardised approach to the collection and analysis of acoustic data in order to facilitate the comparability of results on a European scale. This convergence highlights the importance of close cooperation between Member States to develop common guidelines and share best practices for acoustic monitoring of marine mammals. In this sense, the BlackCeTrends [9] project bringing together several Black Sea coastal countries has demonstrated the importance of this cooperation in order to be able to carry out large-scale monitoring. By imposing a strategy for collecting and analyzing results common to several countries, the conclusions of this study allow an overall view and a comparison of the situation throughout the area.

Marine Renewable energy project such as offshore wind farm offer opportunities to collect data that could be exploited and made available. However, it was emphasized that such data is often private, limiting its use for research or monitoring and resulting in the loss of valuable information on marine species. Access to this data would be a significant opportunity for collaboration and would contribute to a better assessment of the ecological status of marine ecosystems.

It is noteworthy that some countries have already started to collect acoustic data to assess Descriptor D11 related to underwater noise and are now looking for ways to integrate them into the assessment of Descriptor D1. Such an integrated approach would allow for a more efficient use of resources and a better understanding of the overall impact of underwater noise on marine mammals, thus contributing to the conservation of these vulnerable species.

The workshop highlighted the need for large-scale coordination of passive acoustic monitoring of marine mammals under the MSFD. The establishment of a dedicated European working group, along the lines of the TG Noise or WGBYKE, coordinated by ICES, OSPAR, ASCOBANS/ACCOBAMS or the European Commission's Joint Research Center, would allow Member States to exchange knowledge, share experiences and work together to address common challenges. Such an initiative would promote transnational cooperation, harmonize practices and improve the assessment of Good Ecological Status under the MSFD.

5. Future efforts and conclusion

In conclusion, this workshop highlighted the critical importance of harmonising protocols, acoustic indicators and guidelines for the acoustic monitoring of marine mammals. The integration of acoustic data into assessments of good ecological status offers a promising perspective for improving our understanding of the presence, abundance and behaviour of these marine species. However, it became evident that to realize this potential, there is a pressing need for large-scale coordination and close cooperation among EU Member States.

The establishment of a dedicated European working group, similar to the TG Noise, would facilitate this coordination by promoting the exchange of information, sharing best practices and harmonising approaches to acoustic monitoring of marine mammals. Such an initiative would strengthen the protection of these vulnerable species and their habitats by providing a solid basis for assessing and managing the impact of underwater noise under the Marine Strategy Framework Directive.

It is crucial to acknowledge that while acoustic monitoring emerges as a valuable tool for tracking and evaluation marine mammals, the workshop brought to light a significant lack of coordination in methods at European level. By working together, Member States could improve the comparability of results and enhance the assessment of 'Good Ecological Status' in marine ecosystems. By consolidating scientific knowledge and developing common guidelines, we could better understand and protect marine mammals, ensuring the conservation of these valuable species for future generations.

Furthermore, conducting workshops of this nature within the framework of the European Cetacean Society (ECS) proves to be a genuine opportunity. The ECS Annual Conference brings together key stakeholders in the conservation of marine mammals in Europe. This convergence of experts and enthusiasts provides an ideal setting for addressing the challenges and opportunities associated with acoustic monitoring. Leveraging the collective expertise and commitment of ECS members enhances the workshop's impact, fostering collaborative

efforts that are crucial for the effective conservation of marine mammals in European waters.

CRediT authorship contribution statement

Mathilde Michel: Conceptualization, Methodology, Writing – Original Draft, Project administration. Benjamin Guichard: Writing – Review & Editing, Project administration, Funding acquisition. Julie Béesau: Resources, Writing – Review & Editing, Project Administration. Flore Samaran: Writing – Review & Editing, Funding acquisition.

Acknowledgments

ENSTA Bretagne and the French Biodiversity Agency would like to thank the European Cetacean Society for providing the opportunity to organise this workshop and for hosting it. We would also like to thank all the participants who registered and contributed to the success of this workshop and the enriching exchanges. We are particularly grateful to those who generously agreed to present their work during this session. Mathilde Michel is funded by French Biodiversity Agency.

References

[1] Spitz J., Peltier H., Authier M. (2017). Evaluation de l'état écologique des mammifères marins en France métropolitaine. Rapport scientifique pour l'évaluation 2018 au titre de la DCSMM. Observatoire PELAGIS – UMS 3462, Université de La Rochelle / CNRS, 170 PP.

[2] I. Sousa-Pinto, A.R. Amaral, A. Fais, The importance of sound in marine ecology: affirming the role of acoustic communication for marine species, Mar. Ecol. 40 (2) (2019), e12527. DOI: 10.1111/maec.12527.

[3] K.M. Stafford, S.L. Nieukirk, C.G. Fox, Low frequency whale sounds recorded on hydrophones moored in the eastern tropical Pacific, J. Acoust. Soc. Am. 106 (1999) 3687–3698. ⟨https://doi.org/10.1121/1.428220⟩.

[4] C.W. Clark, R.A. Charif, Acoustic Monitoring of Large Whales to the West of Britain and Ireland Using Bottom-mounted Hydrophone Arrays, October 1996-September 1997 (Report No. 281), Joint Nature Conservation Committee, Aberdeen, Scotland, 1998.

[5] D.K. Mellinger, K.M. Stafford, C.G. Fox, Seasonal occurrence of sperm whale (Physeter macrocephalus) sounds in the Gulf of Alaska, 19992001, Mar. Mammal. Sci. 20 (2004) 48–62, https://doi.org/10.1111/j.1748-7692.2004.tb01140.x.

[6] D.K. Mellinger, K.M. Stafford, S.E. Moore, L. Munger, C.G. Fox, Detection of North Pacific right whale (Eubalaena japonica) calls in the Gulf of Alaska, Mar. Mammal. Sci. 20 (2004) 872–879, https://doi.org/10.1111/j.1748-7692.2004.tb01198.x.

[7] C.W. Clark, Acoustic communication and behavior of the southern right whale (Eubalaena australis), in: R. Payne (Ed.), Communication and Behavior of Whales, Avalon Publishing, New York, NY, 1983, pp. 163–198.

[8] J.N. Oswald, J. Barlow, T.F. Norris, Integration of visual and acoustic approaches for studying marine mammals. Field Techniques for Estimating Abundance and Species Richness, Springer,, 2003, pp. 245–265, https://doi.org/10.1007/978-1-4757-5347-6_11.

[9] BlackCeTrends – Large coverage initiative for acoustic survey of cetaceans in the Black Sea. *(s. d.)*. TÜDAV | Türk Deniz Araştırmaları Vakfı. ⟨https://tudav.org/en/active-projects/blackcetrends/⟩.

Mathilde Michel[a,*], Benjamin Guichard[b], Julie Béesau[a], Flore Samaran[a]

[a] *CNRS Lab-STICC, ENSTA Bretagne, Brest, France*

[b] *Office français de la biodiversité, pôle de Brest, Service Evaluation, Connaissance et Usages du Milieu Marin, France*

[*] Correspondence to: 2, Rue François Verny, 29200 Brest, France.
E-mail address: mathilde.michel@ensta-bretagne.org (M. Michel).

3.3 Conclusion

La DCSMM fournit un cadre réglementaire exigeant pour l'évaluation du BEE des mammifères marins. L'atelier européen auquel ce chapitre est consacré a mis en évidence un besoin urgent de standardisation des protocoles d'acquisition et d'analyse en PAM, condition indispensable à la comparabilité et à la robustesse des données. Il ressort que la structuration de la communauté PAM reste limitée, en l'absence d'un groupe de travail européen pérenne dédié. Dans ce contexte, le travail mené par Camille Ollier 2024 (Université de La Rochelle) dans sa thèse, *Cetacean population monitoring, the merits and challenges of combining acoustic and visual data*, a montré l'intérêt de combiner les données acoustiques et visuelles pour améliorer la robustesse des suivis et faciliter la standardisation méthodologique. Cette approche intégrée illustre comment le couplage de différentes sources de données peut renforcer la comparabilité et la fiabilité des évaluations écologiques, un point particulièrement pertinent pour l'évaluation du Descripteur 1 de la DCSMM. Cette lacune freine l'émergence de standards communs et complique la coordination régionale, mais également nationale, comme l'illustrent les difficultés rencontrées dans les suivis menés autour des parcs éoliens offshore. L'établissement d'un cadre collaboratif, coordonné par des instances telles que l'ICES, OSPAR, ACCOBAMS ou la Commission européenne, constituerait une avancée décisive pour mutualiser les bonnes pratiques et assurer l'intégration cohérente des données acoustiques dans les politiques publiques.

CONCLUSION GÉNÉRALE

Cette thèse s'inscrit dans une démarche globale visant à optimiser les suivis par acoustique passive des cétacés (PAM), en combinant les dimensions technique, scientifique et institutionnelle. Dans un contexte de pressions anthropiques croissantes et de changements environnementaux, le PAM représente un outil unique pour documenter la présence et les dynamiques spatio-temporelles des populations, tout en éclairant la gestion adaptative et la conservation.

Contributions sur l'impact du *duty cycle*

Un premier apport concerne l'analyse du *duty cycle*, paramètre d'acquisition souvent perçu comme une contrainte logistique. Cette thèse montre au contraire qu'il constitue un levier stratégique pour garantir la robustesse et la comparabilité des indicateurs acoustiques. Le choix de la stratégie de sous-échantillonnage influence directement la capacité à détecter des dynamiques écologiques clés, en particulier pour les espèces à activité vocale faible ou irrégulière.

Une phase initiale d'enregistrement continu apparaît essentielle pour caractériser la fréquentation et l'activité vocale des espèces et pour calibrer les protocoles de *duty cycle*. Dans ce cadre, il est préférable de privilégier des pourcentages d'écoute intermédiaires à élevés, d'opter pour des périodes d'écoute courtes et répétées plutôt que longues et espacées, et d'adapter systématiquement les choix méthodologiques aux caractéristiques biologiques et techniques des suivis.

Le *duty cycle* ne se limite toutefois pas à la phase d'acquisition. Il intervient également lors de la constitution des jeux de données d'entraînement utilisés par les algorithmes de détection automatique, maillon critique de la chaîne du PAM. Ces algorithmes reposent sur des extraits annotés manuellement, dont le mode de sélection correspond à une forme de *duty cycle* appliqué *a posteriori*. Selon qu'elle soit aléatoire, regroupée en petits ensembles

ou régulièrement espacée dans le temps, cette sélection pourrait influencer la performance des détecteurs et la fiabilité des indicateurs acoustiques.

Enfin, le *duty cycle* peut être envisagé comme un outil de gestion et d'archivage des données. Les enregistrements acoustiques génèrent des volumes considérables, difficiles à stocker et à partager dans la durée. Une stratégie de *duty cycle*, visant à conserver une part représentative des suivis plutôt que l'intégralité des enregistrements, pourrait contribuer à réduire ces contraintes logistiques tout en préservant la valeur scientifique des bases de données.

La thèse de Perrine Bauchot (ENSTA, IPP) illustre cette problématique dans un autre domaine, celui de l'océanographie. En s'appuyant sur un modèle de la circulation de l'Atlantique Nord, elle a testé différentes stratégies de sous-échantillonnage pour améliorer la reconstruction de changements abrupts dans la circulation océanique (Bauchot et al. 2024). Son travail montre que, même lorsque la fréquence d'échantillonnage est faible, l'organisation en *clusters* d'observations consécutives régulièrement espacées dans le temps permet une meilleure détection des changements de régime que des stratégies purement aléatoires ou uniformes. Transposée au PAM, cette approche ouvre une perspective nouvelle : concevoir la sélection d'extraits acoustiques (acquisition, annotation, archivage) afin de maximiser la capacité à détecter des signaux rares ou intermittents et de renforcer la robustesse des suivis acoustiques.

Contributions en modélisation spatio-temporelle

L'utilisation conjointe de modèles additifs généralisés (GAMs) et d'analyses de type fonction orthogonale empirique (EOF) a permis de caractériser des dynamiques spatio-temporelles complexes, d'identifier des zones écologiquement critiques et d'évaluer les liens entre la présence acoustique des grandes baleines et les variables environnementales. Ces approches confirment l'intérêt du PAM longue durée et multi-sites pour suivre des espèces discrètes et mobiles dans des environnements océaniques vastes et hétérogènes.

Toutefois, la pertinence écologique des modèles dépend de la variabilité locale et régionale des habitats et des ressources. Toute application en gestion ou en conservation doit donc tenir compte de cette variabilité, afin d'éviter des décisions fondées uniquement sur des distributions moyennes (Authier, Saraux et Péron 2017 ; Elith, Kearney et

Phillips 2010).

Au-delà d'un seul jeu de données, l'assemblage d'informations issues de plusieurs campagnes, protocoles et méthodes constitue un levier majeur pour surmonter le faible nombre de détections et modéliser efficacement les espèces rares de cétacés (Virgili, Authier et al. 2019 ; Roberts et al. 2016 ; Mannocci et al. 2014). Cela renforce la nécessité d'une standardisation et d'une harmonisation des suivis PAM, afin de rendre comparables et interopérables les données collectées par différentes institutions et pays.

Intégration multi-sources et pluridisciplinarité

Au-delà des seules données acoustiques, l'intégration de sources complémentaires de suivi constitue un levier majeur pour accroître la robustesse des suivis et mieux comprendre les dynamiques de population. La thèse de Camille Ollier 2024 illustre concrètement les bénéfices de la combinaison de suivis visuels et acoustiques, permettant de caractériser plus finement la distribution des populations et d'associer comportement et disponibilité des ressources.

Dans cette même idée, le *Bio-Telemetry and Behavioral Ecology Laboratory* développe des approches intégrées mobilisant des capteurs innovants, la vidéo, l'acoustique active, ainsi que le suivi biologique et physiologique à long terme des cétacés (Friedlaender Lab, UCSC s. d.). Ces outils permettent d'étudier le comportement et l'écologie à différentes échelles spatiales et temporelles, d'évaluer les interactions avec les proies et les perturbations anthropiques, et de suivre la dynamique des populations dans des environnements changeants. Ces initiatives soulignent l'intérêt d'une coordination méthodologique et d'une harmonisation des pratiques pour maximiser la valeur scientifique des suivis et leur applicabilité à la conservation et à la gestion adaptative.

Ces résultats mettent également en évidence plusieurs limites : la dépendance aux signaux vocaux pour les suivis, l'absence d'informations directes sur la distribution des proies, et la difficulté à valider la transférabilité spatiale et temporelle des modèles. Ces manques traduisent la nécessité d'un renforcement des connaissances fondamentales, notamment sur l'écologie alimentaire et les interactions prédateur-proie. Les dispositifs intégrés offrent ainsi une vision plus complète de ces interactions et renforcent la portée opérationnelle des suivis acoustiques pour la conservation et la gestion adaptative.

Recommandations opérationnelles

Sur la base de ces analyses, il est possible de formuler plusieurs recommandations pour guider la conception, la mise en œuvre et la coordination des suivis acoustiques des cétacés :

- **Optimisation du *duty cycle* et des périodes d'écoute :** privilégier des *duty cycles* intermédiaires à élevés ($\geq$ 33 %) lorsque les contraintes techniques le permettent, et favoriser des régimes à périodes d'écoute courtes répétées fréquemment.
- **Phase initiale d'observation continue :** déployer des enregistrements continus pour caractériser la fréquentation et l'activité vocale des espèces, permettant d'ajuster les choix méthodologiques et de décider des indicateurs acoustiques.
- **Adaptation aux caractéristiques biologiques et techniques :** tenir compte des espèces, de la variabilité de leurs signaux et des capacités des détecteurs automatiques pour déterminer la stratégie de sous-échantillonnage et les paramètres de suivi.
- **Standardisation et coordination des suivis :** développer des réseaux multi-sites et long terme, harmoniser les protocoles d'acquisition et de traitement, favoriser le partage d'expériences et la formation pour renforcer la comparabilité et la rigueur scientifique.
- **Prise en compte de la variabilité écologique locale et régionale :** contextualiser les conclusions, éviter des décisions fondées uniquement sur des distributions moyennes, et utiliser des indicateurs de robustesse comme le coefficient de variation des densités prédites.
- **Intégration des informations sur les proies et données environnementales multi-sources :** inclure progressivement les informations issues de modèles numériques ou d'observations directes, ainsi que les données satellites et mesures *in situ*, pour réduire les incertitudes et anticiper les impacts du changement climatique.
- **Intégration multi-sources de données de collecte :** combiner PAM passive et active, balises, observations visuelles et modèles numériques pour enrichir l'interprétation écologique, relier présence, comportement et disponibilité des ressources, et tester la transférabilité spatiale et temporelle des modèles.
- **Approche pluridisciplinaire et intégrée :** allier acoustique, écologie, modélisation statistique et gestion pour maximiser la pertinence scientifique et opérationnelle des suivis et des mesures de conservation.

Structuration institutionnelle et initiatives

Au-delà des aspects méthodologiques et analytiques, cette thèse met en lumière la nécessité d'une structuration et d'une coordination des suivis PAM à toutes les échelles. La standardisation doit commencer au sein des équipes et projets, se développer entre laboratoires et disciplines (e.g. entre acoustique sous-marine et terrestre), et inclure la mise en place de termes communs, de protocoles harmonisés et de formats de métadonnées partagés. Ces efforts facilitent la comparabilité des données et la reproductibilité des analyses, avant même d'envisager une coordination nationale ou européenne.

De nombreuses initiatives illustrent cette dynamique. À l'échelle de projets ou d'équipes, des *workshops* comme celui organisé par Nicole Todd en Irlande, *Porpoises, PAM, and Perspectives for Conservation and Management*, permettent de partager les résultats de suivis visuels et acoustiques, de discuter des stratégies d'amélioration des suivis locaux et de diffuser des perspectives nationales et internationales. Des projets comme METADAX (ENSTA-SHOM) visent à structurer les métadonnées associées à l'acquisition de données acoustiques sous-marines, favorisant l'harmonisation des jeux de données et la comparabilité des suivis. Le projet ECoSS (*European Catalogue of Sound Signatures*, EMODnet 2025) développe une bibliothèque ouverte de signatures sonores sous-marines accompagnée d'un classificateur basé sur l'intelligence artificielle, permettant d'identifier et de suivre différentes sources sonores et de quantifier l'impact acoustique des activités humaines, tout en offrant une approche collaborative pour enrichir et améliorer la classification des sons.

Parallèlement, des initiatives terrestres comme SOL (*Consortium-HN Sounds of Life*, Consortium Huma-Num 2023) proposent des infrastructures et des outils pour explorer des données sonores dans le domaine de l'éco-acoustique, tandis que DataPhonia (BioPhonia, BioPhonia 2025) et *Safe and Sound* (WILDLABS 2025) offrent des cadres standardisés pour le stockage, l'annotation et la diffusion des données bioacoustiques. Ces plateformes incluent également des outils de visualisation du contexte spatial et sonore, des solutions pour l'archivage et la gestion des données, ainsi que des programmes de formation et de communication destinés à partager les bonnes pratiques. Enfin, un projet ISO/CD 23990 ISO/TC 43/SC 3 2025 est en cours pour définir la terminologie de la bioacoustique sous-marine, témoignant d'un effort global pour rendre les données comparables et réutilisables.

Ces projets montrent l'intérêt de centraliser, standardiser et structurer l'information à différentes échelles. Cependant, malgré la multiplication de ces initiatives, elles restent pour l'instant fragmentaires et souvent limitées à des équipes ou projets isolés, en l'absence de coordination nationale ou internationale. Cette fragmentation illustre l'importance des conclusions du *workshop* présenté dans le Chapitre 3, qui souligne la nécessité d'un PAM standardisé, capable de fédérer les pratiques, d'harmoniser les protocoles et de construire des bases de données interopérables pour faciliter la recherche collaborative et la gestion adaptative à grande échelle.

Bibliographie

Abrahms, Briana et al. (2019), « Memory and resource tracking drive blue whale migrations », in : *Proceedings of the National Academy of Sciences* 116.*12*, p. 5582-5587.

ACCOBAMS (2001), *ACCOBAMS | Accord sur la conservation des cétacés de la mer Noire, de la Méditerranée et de la zone Atlantique adjacente*, Consulté le 9 septembre 2025, url : https://accobams.org/fr/.

Ai"ssi, Mehdi et al. (2014), « Modelling predicted sperm whale habitat in the central Mediterranean Sea : requirement for protection beyond the Pelagos Sanctuary boundaries », in : *Aquatic Conservation : Marine and Freshwater Ecosystems* 24.*S1*, p. 50-58.

Akaike, Hirotugu (1974), « A new look at the statistical model identification », in : *IEEE transactions on automatic control* 19.*6*, p. 716-723.

Alcocer, Irene et al. (2022), « Acoustic indices as proxies for biodiversity : a meta-analysis », in : *Biological Reviews* 97.*6*, p. 2209-2236.

Alglave, Baptiste et al. (2024), « Investigating fish reproduction phenology and essential habitats by identifying the main spatio-temporal patterns of fish distribution », in : *ICES Journal of Marine Science* 81.*8*, p. 1563-1574.

Allen, Jenny A et al. (2024), « Evidence of sociality and group foraging in Antarctic minke whales (Balaenoptera bonaerensis) », in : *Behavioral Ecology and Sociobiology* 78.*5*, p. 61.

Allison, David B, Dariusz Stramski et B Greg Mitchell (2010), « Seasonal and interannual variability of particulate organic carbon within the Southern Ocean from satellite ocean color observations », in : *Journal of Geophysical Research : Oceans* 115.*C6*.

AméziANe, Nadia et al. (2011), « Biodiversity of the benthos off Kerguelen Islands : overview and perspectives », in : *The Kerguelen Plateau : marine ecosystem and fisheries. Société Française d'Ichtyologie, Paris*, p. 157-167.

Andrews-Goff, V, Sophie Bestley et al. (2018), « Humpback whale migrations to Antarctic summer foraging grounds through the southwest Pacific Ocean », in : *Scientific reports* 8.*1*, p. 12333.

Andrews-Goff, Virginia, Joshua N Smith et al. (2024), « Ship-based RPA operations for cetacean research in Antarctica : progress, opportunities and challenges », in : *Frontiers in Marine Science* 11, p. 1473471.

ASCOBANS (1994), *ASCOBANS | Agreement on the Conservation of Small Cetaceans of the Baltic, North East Atlantic, Irish and North Seas*, Consulté le 9 septembre 2025, url : https://www.ascobans.org/.

Attard, CRM et al. (2024), « Global conservation genomics of blue whales calls into question subspecies taxonomy and refines knowledge of population structure », in : *Animal Conservation* 27.*5*, p. 626-638.

Au, Whitlow WL (1993), *The sonar of dolphins*, Springer Science & Business Media.

Au, Whitlow WL et Kelly J Benoit-Bird (2003), « Automatic gain control in the echolocation system of dolphins », in : *Nature* 423.*6942*, p. 861-863.

Au, Whitlow WL et Mardi C Hastings (2008), *Principles of marine bioacoustics*, t. 510, Springer.

Aulich, Meghan G et al. (2019), « Fin whale (Balaenoptera physalus) migration in Australian waters using passive acoustic monitoring », in : *Scientific reports* 9.*1*, p. 8840.

Austin, Deborah, WD Bowen et JI McMillan (2004), « Intraspecific variation in movement patterns : modeling individual behaviour in a large marine predator », in : *Oikos* 105.*1*, p. 15-30.

Authier, Matthieu, Claire Saraux et Clara Péron (2017), « Variable selection and accurate predictions in habitat modelling : a shrinkage approach », in : *Ecography* 40.*4*, p. 549-560.

Ballance, L. T., R. L. Pitman et P. C. Fiedler (2006), « Oceanographic influences on seabirds and cetaceans of the eastern tropical Pacific : a review », in : *Progress in Oceanography* 69.*2-4*, p. 360-390.

Barco, Susan G et al. (1999), « Local abundance and distribution of bottlenose dolphins (Tursiops truncatus) in the nearshore waters of Virginia Beach, Virginia », in : *Marine Mammal Science* 15.*2*, p. 394-408.

Barlow, Dawn R et Leigh G Torres (2021), « Planning ahead : dynamic models forecast blue whale distribution with applications for spatial management », in : *Journal of Applied Ecology* 58.*11*, p. 2493-2504.

Bauchot, Perrine et al. (2024), « Impact of sampling strategies on the monitoring of climate regime shifts with a learning data assimilation method », in : *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, p. 7555-7559.

Baumgartner, Mark F et Bruce R Mate (2005), « Summer and fall habitat of North Atlantic right whales (Eubalaena glacialis) inferred from satellite telemetry », in : *Canadian Journal of Fisheries and Aquatic Sciences* 62.*3*, p. 527-543.

Baumgartner, Mark F, Keith D Mullin et al. (2001), « Cetacean habitats in the northern Gulf of Mexico. », in : *Fishery Bulletin* 99.*2*, p. 219-219.

Béhagle, Nolwenn et al. (2016), « Acoustic micronektonic distribution is structured by macroscale oceanographic processes across 20–50 S latitudes in the South-Western Indian Ocean », in : *Deep Sea Research Part I : Oceanographic Research Papers* 110, p. 20-32.

Bellanger, J. et al. (2023), *Suivi par acoustique passive du bruit ambiant et de la fréquentation des mammifères marins dans le cadre du parc éolien en mer de Normandie,*

État de référence, rapp. tech. QO.20210225.10.RAP.001.05A, Brest, France : Quiet-Oceans.

Benoit-Bird, Kelly J et Gareth L Lawson (2016), « Ecological insights from pelagic habitats acquired using active acoustic techniques », in : *Annual review of marine science* 8.*1*, p. 463-490.

Bertucci, Frédéric et al. (2016), « Acoustic indices provide information on the status of coral reefs : an example from Moorea Island in the South Pacific », in : *Scientific Reports* 6.*1*, p. 33326.

Bestley, Sophie et al. (2020), « Ocean circulation and frontal structure near the southern Kerguelen Plateau : the physical context for the Kerguelen Axis ecosystem study », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 174.

BioPhonia (2025), *DataPhonia*, Accès le 15 septembre 2025, url : `https://dataphonia.fr/login`.

Bisack, Kathryn D et Gisele M Magnusson (2021), « Spatial management to reduce entanglement risk to North Atlantic right whales in fishing gear : a case study of US Northeast lobster fishery 2002–2009 », in : *Frontiers in Marine Science* 8, p. 540966.

Block, William M et Leonard A Brennan (1993), « The habitat concept in ornithology : theory and applications », in : *Current ornithology : volume 11*, p. 35-91.

Boersch-Supan, Philipp H, Alex D Rogers et Andrew S Brierley (2017), « The distribution of pelagic sound scattering layers across the southwest Indian Ocean », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 136, p. 108-121.

Branch, Trevor A, Cole C Monnahan et al. (2025), « Separating Historical Catches Among Pygmy Blue Whale Populations Using Recent Song Detections », in : *Marine Mammal Science*, e70003.

Branch, Trevor A, KM Stafford et al. (2007), « Past and present distribution, densities and movements of blue whales Balaenoptera musculus in the Southern Hemisphere and northern Indian Ocean », in : *Mammal Review* 37.*2*, p. 116-175.

Brierley, Andrew S (2014), « Diel vertical migration », in : *Current biology* 24.*22*, R1074-R1076.

Buchan, Susannah J et al. (2021), « Intraseasonal variation in southeast Pacific blue whale acoustic presence, zooplankton backscatter, and oceanographic variables on a feeding ground in Northern Chilean Patagonia », in : *Progress in Oceanography* 199, p. 102709.

Calambokidis, John et al. (2007), « Insights into the underwater diving, feeding, and calling behavior of blue whales from a suction-cup-attached video-imaging tag (CRITTERCAM) », in : *Marine Technology Society Journal* 41.*4*, p. 19-29.

Calonge, Arienne et al. (2024), « Importance of multi-sensor observations to advance species co-occurrence knowledge : A demonstration of two acoustic technologies », in : *Marine Ecology Progress Series* 727, p. 49-65.

Carlström, J et al. (2011), « SAMBAH–Static Acoustic Monitoring of the Baltic Sea Harbour Porpoise », in : *Sweden.*

Cascão, Irma et al. (2020), « Temporal patterns in acoustic presence and foraging activity of oceanic dolphins at seamounts in the Azores », in : *Scientific reports* 10.*1*, p. 3610.
Casey, CB et al. (2022), « Acoustic signalling and behaviour of Antarctic minke whales (Balaenoptera bonaerensis) », in : *Royal Society Open Science* 9.*7*, p. 211557.
Castellote, Manuel, Christopher W Clark et Marc O Lammers (2012), « Acoustic and behavioural changes by fin whales (Balaenoptera physalus) in response to shipping and airgun noise », in : *Biological Conservation* 147.*1*, p. 115-122.
Center for Environmental Studies, FAU (2023), *Global Ocean Conveyor : Thermohaline Circulation*, https://www.ces.fau.edu/nasa/resources/global-ocean-conveyor.php, Consulté le 27 août 2025.
Cerchio, Salvatore, Tahina Rasoloarijao et al. (2024), « Migratory movements and acoustic activity of blue whales (Balaenoptera musculus) and fin whales (Balaenoptera physalus) in the eastern Mozambique Channel », in : *Technical report*.
Cerchio, Salvatore, Andrew Willson et al. (2020), « A new blue whale song-type described for the Arabian Sea and Western Indian Ocean », in : *Endangered Species Research* 43, p. 495-515.
Chami, Ralph (2019), *Nature's solution to climate change*, IMF Finance & Development, December 2019, Consulté le 21 août 2025, url : https://www.imf.org/en/Publications/fandd/issues/2019/12/natures-solution-to-climate-change-chami.
Chapman, Ben B et al. (2011), « The ecology and evolution of partial migration », in : *Oikos* 120.*12*, p. 1764-1775.
Chelonia Ltd. (2022), *Chelonia Ltd.* https://www.chelonia.co.uk/.
Chompret, J. et al. (août 2021), *Suivi en acoustique passive des cétacés – rapport final pour l'état de référence, Parc éolien en mer de la baie de Saint-Brieuc*, Rapport final V2.4, 37 pages, fichier : STB_BIOTOPE_QUIET_OCEANS_ACOUSTIQUE_Etat de référence_V4.pdf, France : Ailes Marines SAS.
Civel-Mazens, M et al. (2021), « Impact of the Agulhas Return Current on the oceanography of the Kerguelen Plateau region, Southern Ocean, over the last 40 kyrs », in : *Quaternary Science Reviews* 251, p. 106711.
Clapham, Phil (2001), « Why do baleen whales migrate ? A response to Corkeron and Connor », in : *Marine Mammal Science* 17.*2*, p. 432-436.
Clapham, Phillip J (1996), « The social and reproductive biology of humpback whales : an ecological perspective », in : *Mammal Review* 26.*1*, p. 27-49.
Clark, Christopher W et Ellen C Garland (2022), *Ethology and behavioral ecology of mysticetes*, Springer, chap. 9, p. 123-145.
CNRS News (2021), *Why the Gulf Stream is a misnomer*, https://news.cnrs.fr/articles/why-the-gulf-stream-is-a-misnomer, Consulté le 27 août 2025.
Coffey, Kevin R, Ruby E Marx et John F Neumaier (2019), « DeepSqueak : a deep learning-based system for detection and analysis of ultrasonic vocalizations », in : *Neuropsychopharmacology* 44.*5*, p. 859-868.

Committee on Taxonomy, Society for Marine Mammalogy (2020), *List of marine mammal species and subspecies*, Version mise à jour en mai 2020, site de la Society for Marine Mammalogy, URL : https://www.marinemammalscience.org/science-and-publications/list-marine-mammal-species-subspecies/.

Consortium Huma-Num (2023), *Consortium Sound of Life*, Accès le 15 septembre 2025, URL : https://consortium-sol.huma-num.fr/.

Constable, Andrew J et al. (2014), « Climate change and Southern Ocean ecosystems I : how changes in physical habitats directly affect marine biota », in : *Global change biology* 20.*10*, p. 3004-3025.

Corkeron, Peter J et Richard C Connor (1999), « Why do baleen whales migrate ? 1 », in : *Marine Mammal Science* 15.*4*, p. 1228-1245.

Cotte, Cédric (2016), *THEMISTO (Toward Hydroacoustics and Ecology of Mid-trophic levels in Indian and Southern Ocean)*, DOI : 10.18142/288.

Cotté, Cédric et al. (2009), « Scale-dependent habitat use by a large free-ranging predator, the Mediterranean fin whale », in : *Deep Sea Research Part I : Oceanographic Research Papers* 56.*5*, p. 801-811.

Cummings, William C et Paul O Thompson (1971), « Underwater sounds from the blue whale, Balaenoptera musculus », in : *The journal of the Acoustical Society of America* 50.*4B*, p. 1193-1198.

Cunningham, Ross B et David B Lindenmayer (2005), « Modeling count data of rare species : some statistical issues », in : *Ecology* 86.*5*, p. 1135-1142.

Damien, Juliette et al. (2019), « Anatomy and functional morphology of the mysticete rorqual whale larynx : phonation positions of the U-fold », in : *The Anatomical Record* 302.*5*, p. 703-717.

Darling, James D, Meagan E Jones et Charles P Nicklin (2006), « Humpback whale songs : Do they organize males during the breeding season ? », in : *Behaviour*, p. 1051-1101.

Davis, Genevieve E et al. (2017), « Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (Eubalaena glacialis) from 2004 to 2014 », in : *Scientific reports* 7.*1*, p. 13460.

Deecke, Volker B, John KB Ford et Paul Spong (2000), « Dialect change in resident killer whales : implications for vocal learning and cultural transmission », in : *Animal behaviour* 60.*5*, p. 629-638.

Dekeling, R. P. A. et al. (nov. 2014), *Progress Report on Monitoring of Underwater Noise : 3rd Report of the Technical Group on Underwater Noise (TG Noise)*, rapp. tech., Supported by the Coastal & Marine Union (EUCC) and Arcadis under framework contract ENV.D.2/FRA/2012/0025. Final editing by R. Dekeling, M. Tasker, and M. Ferreira, Technical Group on Underwater Noise (TG Noise), URL : https://mcc.jrc.ec.europa.eu/documents/201605183553.pdf.

Double, Michael C et al. (2014), « Migratory movements of pygmy blue whales (Balaenoptera musculus brevicauda) between Australia and Indonesia as revealed by satellite telemetry », in : *PLoS One* 9.*4*, e93578.

Dréo, Richard, Léa Bouffaut et al. (2019), « Baleen whale distribution and seasonal occurrence revealed by an ocean bottom seismometer network in the Western Indian Ocean », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 161, p. 132-144.

Dréo, Richard, Wayne Crawford et al. (2025), « Singing around the volcano : Detecting baleen whales in the Mozambique channel based on their song rhythms, from seismic and hydroacoustic data », in : *The Journal of the Acoustical Society of America*.

Drira, Achraf et al. (août 2021), *Suivi acoustique passif du bruit ambiant et des mammifères marins – État de référence dans le cadre du projet de parc éolien en mer de Fécamp*, rapp. tech. N°A1000160, Rapport d'étude acoustique dans le cadre du projet de parc éolien en mer de Fécamp, EDF Renouvelables – Projet Fécamp.

Dubus, Gabriel (2025), « From crowd-sourcing to supervised deep learning : addressing new challenges in the Detection and Classification of Cetaceans using Passive Acoustic Monitoring », English, NNT : 2025SORUS037, Thèse de doctorat, France : Sorbonne Université, url : https://theses.hal.science/tel-05109487.

Dubus, Gabriel, Dorian Cazau et al. (2024), « From citizen science to AI models : Advancing cetacean vocalization automatic detection through multi-annotator campaigns », in : *Ecological Informatics* 81, p. 102642.

Dubus, Gabriel, Maëlle Torterotot et al. (2025), « APLOSE : A web-based annotation platform for underwater passive acoustic monitoring », in : *SoftwareX* 29, p. 102055.

Elith, Jane, Michael Kearney et Steven Phillips (2010), « The art of modelling range-shifting species », in : *Methods in ecology and evolution* 1.*4*, p. 330-342.

Elith, Jane et John R Leathwick (2009), « Species distribution models : ecological explanation and prediction across space and time », in : *Annual review of ecology, evolution, and systematics* 40.*1*, p. 677-697.

Elith, Jane, John R Leathwick et Trevor Hastie (2008), « A working guide to boosted regression trees », in : *Journal of animal ecology* 77.*4*, p. 802-813.

Elith, Jane, Steven J Phillips et al. (2011), « A statistical explanation of MaxEnt for ecologists », in : *Diversity and distributions* 17.*1*, p. 43-57.

EMODnet (2025), *ECoSS - European Catalogue of Sound Signatures*, Accès le 15 septembre 2025, url : https://emodnet.ec.europa.eu/en/ecoss-european-catalogue-sound-signatures.

ENSTA et Parc Naturel Marin d'Iroise (2022), *CETIROISE : Un observatoire acoustique des cétacés en mer d'Iroise*, Consulté le 24 juin 2025, url : https://parc-marin-iroise.fr/editorial/cetiroise.

Erbe, Christine et al. (2017), « Review of underwater and in-air sounds emitted by Australian and Antarctic marine mammals », in : *Acoustics Australia* 45.*2*, p. 179-241.

Estes, James A et al. (2011), « Trophic downgrading of planet Earth », in : *science* 333.*6040*, p. 301-306.

Evolution Biologique (2022), *Histoire des cétacés*, http://www.evolution-biologique.org/echelle-du-temps/echelle-des-temps-geologiques/paleogene/histoire_cetaces.html, Consulté le 21 août 2025.

Fauchald, Per (1999), « Foraging in a hierarchical patch system », in : *The american naturalist* 153.*6*, p. 603-613.
Feuillet, Nathalie et al. (2021), « Birth of a large volcanic edifice offshore Mayotte via lithosphere-scale dyke intrusion », in : *Nature Geoscience* 14.*10*, p. 787-795.
Filun, Diego et al. (2020), « Frozen verses : Antarctic minke whales (Balaenoptera bonaerensis) call predominantly during austral winter », in : *Royal Society open science* 7.*10*, p. 192112.
Fish, FE et J Rohr (1999), « Review of dolphin hydrodynamics and swimming performance SPAWARS System Center Technical Report 1801 », in : *San Diego : SPAWAR Systems Center*.
Fleishman, Erica et al. (2023), « Ecological inferences about marine mammals from passive acoustic data », in : *Biological Reviews* 98.*5*, p. 1633-1647.
Fossi, M Cristina et al. (2004), « Assessment of toxicological status of a SW Mediterranean segment population of striped dolphin (Stenella coeruleoalba) using skin biopsy », in : *Marine Environmental Research* 58.*2-5*, p. 269-274.
Frankel, Adam S (2009), « Sound production », in : *Encyclopedia of marine mammals*, Elsevier, p. 1056-1071.
Franklin, Janet (2010a), *Mapping Species Distributions : Spatial Inference and Prediction*, Ecology, Biodiversity and Conservation, Cambridge : Cambridge University Press.
— (2010b), *Mapping Species Distributions : Spatial Inference and Prediction*, Ecology, Biodiversity and Conservation, Cambridge University Press.
Friedlaender, Ari S et al. (2006), « Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the Western Antarctic Peninsula », in : *Marine Ecology Progress Series* 317, p. 297-310.
Friedlaender Lab, UCSC (s. d.), *Bio-Telemetry & Behavioral Ecology Lab*, `https://btbel.pbsci.ucsc.edu/`, Accédé le 3 septembre 2025.
Gatesy, John et al. (2013), « A phylogenetic blueprint for a modern whale », in : *Molecular phylogenetics and evolution* 66.*2*, p. 479-506.
Gavrilov, Alexander N et Robert D McCauley (2013), « Acoustic detection and long-term monitoring of pygmy blue whales over the continental slope in southwest Australia », in : *The Journal of the Acoustical Society of America* 134.*3*, p. 2505-2513.
Geijer, Christina KA, Giuseppe Notarbartolo di Sciara et Simone Panigada (2016), « Mysticete migration revisited : are Mediterranean fin whales an anomaly ? », in : *Mammal Review* 46.*4*, p. 284-296.
Gervaise, Cédric et al. (nov. 2018), *Réseau CALME Caractérisation Acoustique du Littoral Méditerranéen et de ses Ecosystèmes : Synthèse des travaux réalisés pour la période [01/01/2015 – 01/08/2018]*, rapp. tech., CHORUS Research Institute, doi : `10.13140/RG.2.2.22870.50246`, url : `https://www.researchgate.net/publication/328838424`.
Gibbens, Sarah (jan. 2020), *Rare : des scientifiques ont filmé le repas d'une baleine bleue*, Consulté le 27 août 2025, url : `https://www.nationalgeographic.fr/animaux/rare-des-scientifiques-ont-filme-le-repas-dune-baleine-bleue`.

Gill, Peter C et al. (2002), « A blue whale (Balaenoptera musculus) feeding ground in a southern Australian coastal upwelling zone », in : *J. Cetacean Res. Manage.* 4.*2*, p. 179-184.

Gillespie, Douglas, Marjolaine Caillat et al. (2013), « Automatic detection and classification of odontocete whistles », in : *The Journal of the Acoustical Society of America* 134.*3*, p. 2427-2437.

Gillespie, Douglas, DK Mellinger et al. (2008), « PAMGUARD : Semiautomated, open source software for real-time acoustic detection and localisation of cetaceans », in : *Journal of the Acoustical Society of America* 30.*5*, p. 54-62.

Godet, Claire et al. (2020), « Matching zooplankton abundance and environment in the South Indian Ocean and Southern Ocean », in : *Deep Sea Research Part I : Oceanographic Research Papers* 163, p. 103347.

Godø, Olav R et al. (2012), « Mesoscale eddies are oases for higher trophic marine life », in : *PloS one* 7.*1*, e30161.

Goldbogen, Jeremy A, DE Cade et al. (2017), « How baleen whales feed : the biomechanics of engulfment and filtration », in : *Annual review of marine science* 9.*1*, p. 367-386.

Goldbogen, Jeremy A, John Calambokidis et al. (2011), « Mechanics, hydrodynamics and energetics of blue whale lunge feeding : efficiency dependence on krill density », in : *Journal of Experimental Biology* 214.*1*, p. 131-146.

Griebel, Ulrike et Leo Peichl (2003), « Colour vision in aquatic mammals—facts and open questions », in : *Aquatic mammals* 29.*1*, p. 18-30.

Griffin, R. B. et N. J. Griffin (2003), « Distribution, habitat partitioning, and abundance of Atlantic spotted dolphins, bottlenose dolphins, and loggerhead sea turtles on the eastern Gulf of Mexico continental shelf », in : *Gulf of Mexico Science* 21.*1*, p. 3.

Guisan, Antoine et Niklaus E Zimmermann (2000), « Predictive habitat distribution models in ecology », in : *Ecological modelling* 135.*2-3*, p. 147-186.

Hamilton, Edwin L et Richard T Bachman (1982), « Sound velocity and related properties of marine sediments », in : *The Journal of the Acoustical Society of America* 72.*6*, p. 1891-1904.

Harris, Sydney A, Nick T Shears et Craig A Radford (2016), « Ecoacoustic indices as proxies for biodiversity on temperate reefs », in : *Methods in Ecology and Evolution* 7.*6*, p. 713-724.

Hastie, T., R. Tibshirani et J. H. Friedman (2009), *The elements of statistical learning : data mining, inference, and prediction*, t. 2, New York : Springer, p. 1-758.

Hastie, Trevor J (2017), « Generalized additive models », in : *Statistical models in S*, p. 249-307.

Hays, Graeme C et al. (2016), « Key questions in marine megafauna movement ecology », in : *Trends in ecology & evolution* 31.*6*, p. 463-475.

Hazen, Elliott Lee, Ari Seth Friedlaender et Jeremy Arthur Goldbogen (2015), « Blue whales (Balaenoptera musculus) optimize foraging efficiency by balancing oxygen use and energy gain as a function of prey density », in : *Science Advances* 1.*9*, e1500469.

Hersh, Taylor A et al. (2022), « Evidence from sperm whale clans of symbolic marking in non-human cultures », in : *Proceedings of the National Academy of Sciences* 119.*37*, e2201692119.

Hooker, Sascha K et Leah R Gerber (2004), « Marine reserves as a tool for ecosystem-based management : the potential importance of megafauna », in : *Bioscience* 54.*1*, p. 27-39.

Horton, Travis W et al. (2022), « Baleen whale migration », in : *Ethology and behavioral ecology of mysticetes*, Springer, p. 71-104.

Hunt, GL et DC Schneider (1987), « Scale-dependent processes in the physical and biological environment of marine birds », in : *Seabirds : feeding biology and role in marine ecosystems. Cambridge University Press, Cambridge*, p. 7-41.

Hutchinson, G.E. (1957), « Concluding remarks », in : *Cold Spring Harbor Symposia on Quantitative Biology* 22, p. 415-427.

Ifremer (1902), *Ifremer in Europe and around the world*, Consulté le 9 septembre 2025, url : `https://www.ifremer.fr/en/ifremer-europe-and-around-world`.

— (2020), *DCSMM - Directive Cadre Stratégique sur le Milieu Marin*, Consulté le 27 août 2025, url : `https://sextant.ifremer.fr/Ressources/Liste-des-catalogues-thematiques/DCSMM-Directive-Cadre-Strategique-sur-le-Milieu-Marin`.

ISO/TC 43/SC 3 (2025), *ISO/CD 23990 : Acoustique sous-marine — terminologie bioacoustique*, Projet en cours, standardisation des termes et définitions pour la bioacoustique aquatique, y compris la biotremologie, Accès le 15 septembre 2025, examen du projet par le comité en cours.

Ivanchikova, Julia et al. (2024), « Seasonal and diel patterns in Black Sea harbour porpoise acoustic activity in 2020–2022 », in : *Ecology and Evolution* 14.*10*, e70182.

Iwakiri, T et M Watanabe (2019), « Strengthening of the Indian Ocean Dipole with increasing seasonal cycle in the mid-Holocene », in : *Geophysical Research Letters* 46.*14*, p. 8320-8328.

IWC (1980), « Chairman's report of the thirty-first annual meeting. », in : *Rep. int. Whal. Commn.*, 30 :25-41.

Izard, Lloyd (déc. 2023), « Structuration spatiale et variabilité des écosystèmes mésopélagiques dans l'Océan Indien Sud », Theses, Sorbonne Université, url : `https://theses.hal.science/tel-04524647`.

Janik, Vincent M, Laela S Sayigh et Randall S Wells (2006), « Signature whistle shape conveys identity information to bottlenose dolphins », in : *Proceedings of the National Academy of Sciences* 103.*21*, p. 8293-8297.

Jefferson, Thomas A., Stephen Leatherwood et Marc A. Webber (1993), « Marine Mammals of the world », in : *United Nations Environment Programme (UNEP), Food and Agriculture Organization of the United Nations (FAO), United Nations (UN), Rome, Italy.*

Jepson, Paul D et al. (2016), « PCB pollution continues to impact populations of orcas and other dolphins in European waters », in : *Scientific reports* 6.*1*, p. 18573.

Johnson, C. M. et al. (2022), *Protecting Blue Corridors – Challenges and solutions for migratory whales navigating national and international seas*, DOI : 10.5281/zenodo.6196131, URL : https://doi.org/10.5281/zenodo.6196131.

Johnston, DW et al. (2008), « Temporal patterns in the acoustic signals of beaked whales at Cross Seamount », in : *Biology Letters* 4.*2*, p. 208-211.

Jongman, R. H. G., C. J. Ter Braak et O. F. Van Tongeren (1995), *Data analysis in community and landscape ecology*, Cambridge university press.

Kahane-Rapport, SR et al. (2022), « Field measurements reveal exposure risk to microplastic ingestion by filter-feeding megafauna », in : *Nature communications* 13.*1*, p. 6327.

Kaschner, Kristin (2007), « Air-Breathing Visitors to Seamounts : Marine Mammals », in : *Seamounts : ecology, fisheries & conservation*, p. 230-238.

Kawamura, AKITO (1980), « A review of food of balaenopterid whales », in : *Sci. Rep. Whales Res. Inst* 32, p. 155-197.

Kawamura, Akito (1994), « A review of baleen whale feeding in the Southern Ocean. », in : *Rep. int whal. Commn* 44, p. 261-271.

Kellogg, Remington (1929), *What is known of the migrations of some of the whalebone whales*, US Government Printing Office.

Kent, Chandra Salgado, Tiago A Marques et Danielle Harris (2022), « Fundamental data analysis tools and concepts for bioacoustical research », in : *Exploring Animal Behavior Through Sound : Volume*, p. 319.

Keribin, Erwan, Élodie Morin et Romain Vovard (jan. 2024), *APLOSE : a scalable web-based annotation tool for marine bioacoustics - public repository*, version 1.6.4, Consulté le 27 juin 2025, DOI : 10.5281/zenodo.10468000, URL : https://github.com/Project-OSmOSE/osmose-app.

Knowlton, Amy R et al. (2012), « Monitoring North Atlantic right whale Eubalaena glacialis entanglement rates : a 30 yr retrospective », in : *Marine Ecology Progress Series* 466, p. 293-302.

Kubryakov, Arseny, Anna Aleskerova et Artem Mizyuk (2024), « Submesoscale features of coastal upwellings in the Black Sea : Observations and modeling », in : *Continental Shelf Research* 279, p. 105291.

Lambert, Christophe et al. (2014), « Predicting cetacean habitats from their energetic needs and the distribution of their prey in two contrasted tropical regions », in : *PLoS One* 9.*8*, e105958.

Laran, Sophie et al. (2017), « A comprehensive survey of pelagic megafauna : their distribution, densities, and taxonomic richness in the tropical Southwest Indian Ocean », in : *Frontiers in Marine Science* 4, p. 139.

Latimer, Andrew M et al. (2006), « Building statistical models to analyze species distributions », in : *Ecological applications* 16.*1*, p. 33-50.

Lavery, Trish J et al. (2010), « Iron defecation by sperm whales stimulates carbon export in the Southern Ocean », in : *Proceedings of the Royal Society B : Biological Sciences* 277.*1699*, p. 3527-3531.

Lê, Sébastien, Julie Josse et François Husson (2008), « FactoMineR : A Package for Multivariate Analysis », in : *Journal of Statistical Software* 25.*1*, p. 1-18, doi : 10.18637/jss.v025.i01.

LeDuc, RG et al. (2007), « Patterns of genetic variation in Southern Hemisphere blue whales and the use of assignment test to detect mixing on the feeding grounds », in : *J. Cetacean Res. Manage.* 9.*1*, p. 73-80.

Lehodey, P., I. Senina et J. Sibert (2023), *SEAPODYM : Spatial Ecosystem and Population Dynamics Model*, https://seapodym.groupcls.com/about-seapodym/, Consulté le 9 septembre 2025, url : https://seapodym.groupcls.com/about-seapodym/.

Lehodey, Patrick, Inna Senina et Raghu Murtugudde (2008), « A spatial ecosystem and populations dynamics model (SEAPODYM)–Modeling of tuna and tuna-like populations », in : *Progress in Oceanography* 78.*4*, p. 304-318.

Lek, Sovan et Jean-François Guégan (1999), « Artificial neural networks as a tool in ecological modelling, an introduction », in : *Ecological modelling* 120.*2-3*, p. 65-73.

Lemoine, Anne et al. (2020), « The 2018–2019 seismo-volcanic crisis east of Mayotte, Comoros islands : seismicity and ground deformation markers of an exceptional submarine eruption », in : *Geophysical Journal International* 223.*1*, p. 22-44.

Leroy, Emmanuelle (2017), « Surveillance acoustique des baleines bleues Antarctique dans l'océan Indien austral : traitement, analyse et interprétation », français, NNT : 2017BRES0069, Thèse de doctorat, Université de Bretagne occidentale - Brest, url : https://theses.hal.science/tel-01686443.

Leroy, Emmanuelle C, Jean-Yves Royer, Abigail Alling et al. (2021), « Multiple pygmy blue whale acoustic populations in the Indian Ocean : whale song identifies a possible new population », in : *Scientific Reports* 11.*1*, p. 8762.

Leroy, Emmanuelle C, Jean-Yves Royer, Julien Bonnel et al. (2018), « Long-term and seasonal changes of large whale call frequency in the southern Indian Ocean », in : *Journal of Geophysical Research : Oceans* 123.*11*, p. 8568-8580.

Leroy, Emmanuelle C, Flore Samaran, Julien Bonnel et al. (2016), « Seasonal and diel vocalization patterns of Antarctic blue whale (Balaenoptera musculus intermedia) in the Southern Indian Ocean : a multi-year and multi-site study », in : *PloS one* 11.*11*, e0163587.

Leroy, Emmanuelle C, Flore Samaran, Kathleen M Stafford et al. (2018), « Broad-scale study of the seasonal and geographic occurrence of blue and fin whales in the Southern Indian Ocean », in : *Endangered Species Research* 37, p. 289-300.

Leroy, Emmanuelle C, Karolin Thomisch et al. (2018), « On the reliability of acoustic annotations and automatic detections of Antarctic blue whale calls under different acoustic conditions », in : *The Journal of the Acoustical Society of America* 144.*2*, p. 740-754.

Letsheleha, Ishmail S et al. (2022), « Year-round acoustic monitoring of Antarctic blue and fin whales in relation to environmental conditions off the west coast of South Africa », in : *Marine Biology* 169.*3*, p. 41.

Levin, Simon A (1992), « The problem of pattern and scale in ecology : the Robert H. MacArthur award lecture », in : *Ecology* 73.*6*, p. 1943-1967.
Lewis, Leah A et al. (2018), « Context-dependent variability in blue whale acoustic behaviour », in : *Royal Society open science* 5.*8*, p. 180241.
Lin, Tzu-Hao et al. (2015), « Passive acoustic monitoring of the temporal variability of odontocete tonal sounds from a long-term marine observatory », in : *PLoS One* 10.*4*, e0123943.
Liu, ZeLin et al. (2010), « Application of artificial neural networks in global climate change and ecological research : An overview », in : *Chinese science bulletin* 55, p. 3853-3863.
Ljungblad, DK, Christopher W Clark et Hiroyuki Shimada (1998), « SC/49/SH17 A comparison of sounds attributed to pygmy blue whales (Balaenoptera musculus brevicauda) recorded south of the Madagascar Plateau and those attributed to'true'blue whales (Balaenoptera musculus) recorded off Antarctica », in : *Report-International Whaling Commission* 48, p. 439-442.
Ljungblad, DK, Kathleen M Stafford et al. (1997), « Sounds attributed to blue whales recorded off the southwest coast of Australia in December 1995 », in : *Report of the International Whaling Commission* 47, p. 435-439.
Llano, George A (1978), « Ecology of the Southern Ocean Region », in : *U. Miami L. Rev.* 33, p. 357.
Longhurst, Alan R (2010), *Ecological geography of the sea*, Elsevier.
Lutron, X (2002), *An introduction to underwater acoustics*.
Mackintosh, NA (1966), « The distribution of southern blue and fin whales », in : *Whales, dolphins, and porpoises*, p. 125-144.
MacLeod, C. D., T. Brereton et C. Martin (2009), « Changes in the occurrence of common dolphins, striped dolphins and harbour porpoises in the English Channel and Bay of Biscay », in : *Journal of the Marine Biological Association of the United Kingdom* 89.*5*, p. 1059-1065.
Madsen, Peter Teglberg et al. (2012), « Dolphin whistles : a functional misnomer revealed by heliox breathing », in : *Biology letters* 8.*2*, p. 211-213.
Mann, Janet (2000), *Cetacean societies : field studies of dolphins and whales*, University of Chicago Press.
Mannocci, Laura et al. (2014), « Predicting top predator habitats in the Southwest Indian Ocean », in : *Ecography* 37.*3*, p. 261-278.
Marques, Tiago A et al. (2013), « Estimating animal population density using passive acoustics », in : *Biological reviews* 88.*2*, p. 287-309.
Mass, Alla M et Alexander YA Supin (2007), « Adaptive features of aquatic mammals' eye », in : *The Anatomical Record : Advances in Integrative Anatomy and Evolutionary Biology : Advances in Integrative Anatomy and Evolutionary Biology* 290.*6*, p. 701-715.
Mate, BR et al. (2011), « Coastal, offshore, and migratory movements of South African right whales revealed by satellite telemetry », in : *Marine Mammal Science* 27.*3*, p. 455-476.

Mawren, D et al. (2023), « On the importance of the Mozambique Channel for the climate of southeastern Africa », in : *Climate Dynamics* 60.*1*, p. 279-299.
Mayo, Charles A et Marilyn K Marx (1990), « Surface foraging behaviour of the North Atlantic right whale, Eubalaena glacialis, and associated zooplankton characteristics », in : *Canadian Journal of Zoology* 68.*10*, p. 2214-2220.
McCauley, Robert D, Alexander N Gavrilov et al. (2018), « Pygmy blue and Antarctic blue whale presence, distribution and population parameters in southern Australia based on passive acoustics », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 157, p. 154-168.
McCauley, Robert D, Curt Jenner et al. (2001), « Blue whale calling in the Rottnest trench—2000, Western Australia », in : *Prepared for Environment Australia, from Centre for Marine Science and Technology, Curtin University, R2001–6*.
McCormack, Stacey A et al. (2021), « Southern Ocean food web modelling : progress, prognoses, and future priorities for research and policy makers », in : *Frontiers in Ecology and Evolution* 9, p. 624763.
McCullagh, Peter (2019), *Generalized linear models*, Routledge.
McDonald, Mark A, John Calambokidis et al. (2001), « The acoustic calls of blue whales off California with gender data », in : *The Journal of the Acoustical Society of America* 109.*4*, p. 1728-1735.
McDonald, Mark A, Sarah L Mesnick et John A Hildebrand (2006), « Biogeographic characterization of blue whale song worldwide : using song to identify populations », in : *Journal of cetacean research and management* 8.*1*, p. 55-65.
McInerny, Greg J et Rampal S Etienne (2012a), « Ditch the niche–is the niche a useful concept in ecology or species distribution modelling ? », in : *Journal of Biogeography* 39.*12*, p. 2096-2102.
— (2012b), « Stitch the niche–a practical philosophy and visual schematic for the niche concept », in : *Journal of Biogeography* 39.*12*, p. 2103-2111.
McKenna, Megan F et al. (2012), « Morphology of the odontocete melon and its implications for acoustic function », in : *Marine Mammal Science* 28.*4*, p. 690-713.
Meier, Walter N et al. (2014), « Arctic sea ice in transformation : A review of recent observed changes and impacts on biology and human activity », in : *Reviews of Geophysics* 52.*3*, p. 185-217.
Mellinger, David K et Christopher W Clark (2003), « Blue whale (Balaenoptera musculus) sounds from the North Atlantic », in : *The Journal of the Acoustical Society of America* 114.*2*, p. 1108-1119.
Mellinger, David K, Kathleen M Stafford et al. (2007), « An overview of fixed passive acoustic observation methods for cetaceans », in : *Oceanography* 20.*4*, p. 36-45.
Merchant, Nathan D et al. (2015), « Measuring acoustic habitats », in : *Methods in Ecology and Evolution* 6.*3*, p. 257-265.
Miksis-Olds, Jennifer L, Sharon L Nieukirk et Danielle V Harris (2018), « Two unit analysis of Sri Lankan pygmy blue whale song over a decade », in : *The Journal of the Acoustical Society of America* 144.*6*, p. 3618-3626.

Miller, EJ et al. (2019), « The characteristics of krill swarms in relation to aggregating Antarctic blue whales », in : *Scientific Reports* 9.*1*, p. 16487.
Minton, G. et al. (2021), *Shipping and cetaceans : A review of impacts and mitigation options for policy makers and other stakeholders*, Report 72 pages, WWF.
Mobley Jr, Joseph R et David A Helweg (1990), « Visual ecology and cognition in cetaceans », in : *Sensory abilities of cetaceans : Laboratory and field evidence*, p. 519-536.
Monsarrat, S. et al. (2015), « Historical summer distribution of the endangered North Atlantic right whale (Eubalaena glacialis) : a hypothesis based on environmental preferences of a congeneric species », in : *Diversity and Distributions* 21.*8*, p. 925-937.
Moore, J Keith et Mark R Abbott (2000), « Phytoplankton chlorophyll distributions and primary production in the Southern Ocean », in : *Journal of Geophysical Research : Oceans* 105.*C12*, p. 28709-28722.
Moore, Sue E et Randall R Reeves (2018), « Tracking arctic marine mammal resilience in an era of rapid ecosystem alteration », in : *PLoS Biology* 16.*10*, e2006708.
Morin, Xavier et Wilfried Thuiller (2009), « Comparing niche-and process-based models to reduce prediction uncertainty in species range shifts under climate change », in : *Ecology* 90.*5*, p. 1301-1313.
Mustika, Putu Liza Kusuma et al. (2024), « The first record of the southbound movements of satellite-tagged pygmy blue whales (B. m. brevicauda) from Savu Sea (Indonesia) to the subantarctic waters », in : *Marine Mammal Science*, e13167.
Möller, Luciana M et al. (2020), « Movements and behaviour of blue whales satellite tagged in an Australian upwelling system », in : *Scientific Reports* 10.*1*, p. 21165.
Nerini, MK et JS Oliver (1983), « Gray whales and the structure of the Bering Sea benthos », in : *Oecologia* 59.*2*, p. 224-225.
Nicol, Stephen, Anthony Worby et Rebecca Leaper (2008), « Changes in the Antarctic sea ice ecosystem : potential effects on krill and baleen whales », in : *Marine and Freshwater Research* 59.*5*, p. 361-382.
Nisi, Anna C et al. (2024), « Ship collision risk threatens whales across the world's oceans », in : *Science* 386.*6724*, p. 870-875.
Noad, Michael J et al. (2000), « Cultural revolution in whale songs », in : *Nature* 408.*6812*, p. 537-537.
Notarbartolo-Di-Sciara, Giuseppe et al. (2003), « The fin whale Balaenoptera physalus (L. 1758) in the Mediterranean Sea », in : *Mammal Review* 33.*2*, p. 105-150.
O'Hara, Casey C et al. (2019), « Mapping status and conservation of global at-risk marine biodiversity », in : *Conservation Letters* 12.*4*, e12651.
Observatoire PELAGIS – CNRS (2013), *SAMM – Suivi Aérien de la Mégafaune Marine*, Consulté le 9 septembre 2025, url : `https://www.observatoire-pelagis.cnrs.fr/pelagis-2/les-programmes/samm/`.
Office français de la biodiversité (OFB) (1991), *Les réseaux d'échouages*, Consulté le 9 septembre 2025, url : `https://www.ofb.gouv.fr/les-reseaux-dechouages`.
Oleson, Erin M et al. (2007), « Blue whale visual and acoustic encounter rates in the Southern California Bight », in : *Marine Mammal Science* 23.*3*, p. 574-597.

Ollier, Camille (sept. 2024), « Cetacean population monitoring : the merits and challenges of combining acoustic and visual data », Theses, Université de La Rochelle, url : https://theses.hal.science/tel-05106040.

Orsi, Alejandro H, Thomas Whitworth III et Worth D Nowlin Jr (1995), « On the meridional extent and fronts of the Antarctic Circumpolar Current », in : *Deep Sea Research Part I : Oceanographic Research Papers* 42.5, p. 641-673.

OSPAR Commission (1992), *Marine Mammals Thematic Assessment*, https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/thematic-assessments/marine-mammals/, Accès le 9 septembre 2025, url : https://oap.ospar.org/en/ospar-assessments/quality-status-reports/qsr-2023/thematic-assessments/marine-mammals/.

Owen, Kylie et al. (2016), « A week in the life of a pygmy blue whale : migratory dive depth overlaps with large vessel drafts », in : *Animal Biotelemetry* 4, p. 1-11.

Palacios, D. et al. (2025), *Protecting Blue Corridors Dataset*, https://bluecorridors.org, Accessed 2025-06-26.

Panigada, Simone et al. (2006), « Mediterranean fin whales at risk from fatal ship strikes », in : *Marine Pollution Bulletin* 52.10, p. 1287-1298.

Parc naturel marin de Mayotte (jan. 2010), Créé par décret n° 2010-71 du 18 janvier 2010, url : https://parc-marin-mayotte.fr/.

Parsons, ECM et NA Rose (2022), « The history of cetacean hunting and changing attitudes to whales and dolphins », in : *Marine Mammals : the Evolving Human Factor*, Springer, p. 219-254.

Peltier, H., M. Authier, F. Caurant et al. (2019), *État des connaissances sur les captures accidentelles de dauphins communs dans le golfe de Gascogne – Synthèse 2019*, Rapport scientifique Convention avec le MTES, La Rochelle : Observatoire PELAGIS – UMS 3462, La Rochelle Université / CNRS, p. 23.

Peltier, Hélène, Matthieu Authier, Rob Deaville et al. (2016), « Small cetacean bycatch as estimated from stranding schemes : the common dolphin case in the northeast Atlantic », in : *Environmental Science & Policy* 63, p. 7-18.

Peltier, Hélène, Sophie Laran et al. (2024), « From the sky and on the beaches : complementary tools to evaluate common dolphin bycatch in the Bay of Biscay », in : *Endangered Species Research* 53, p. 509-522.

Perrin, William F, Bernd Würsig et Johannes Gerardus Marie Thewissen (2009), *Encyclopedia of marine mammals*, Academic Press.

Peterson, A Townsend et al. (2011), « Ecological niches and geographic distributions », in : *Ecological niches and geographic distributions*, Princeton University Press.

Phillips, Helen E et al. (2021), « Progress in understanding of Indian Ocean circulation, variability, air–sea exchange, and impacts on biogeochemistry », in : *Ocean Science* 17.6, p. 1677-1751.

Pindito, E. (2016), *Blue whale and calf*, Video, YouTube, https://www.youtube.com/watch?v=1T-jyPVW1IA.

Pindito, E. (2018), *Blue whale 3 trips Pindito*, Video, YouTube, https://www.youtube.com/watch?v=p716WHPxtVE.

Quinn, Thomas P et Richard D Brodeur (1991), « Intra-specific variations in the movement patterns of marine animals », in : *American Zoologist* 31.*1*, p. 231-241.

R Core Team (2021), *R : A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, url : https://www.R-project.org/.

Rand, Zoe R, Trevor A Branch et Jennifer A Jackson (2024), « High historical movement rates of Antarctic blue whales on Southern Ocean feeding grounds estimated from Discovery mark data », in : *Endangered Species Research* 55, p. 109-128.

Rand, Zoe R, Jason D Wood et Julie N Oswald (2022), « Effects of duty cycles on passive acoustic monitoring of southern resident killer whale (Orcinus orca) occurrence and behavior », in : *The Journal of the Acoustical Society of America* 151.*3*, p. 1651-1660.

Randin, Christophe F et al. (2006), « Are niche-based species distribution models transferable in space ? », in : *Journal of biogeography* 33.*10*, p. 1689-1703.

Rankin, Shannon et al. (2005), « Vocalisations of Antarctic blue whales, Balaenoptera musculus intermedia, recorded during the 2001/2002 and 2002/2003 IWC/SOWER circumpolar cruises, Area V, Antarctica », in : *Journal of Cetacean Research and Management* 7.*1*, p. 13-20.

Redfern, JV et al. (2006), « Techniques for cetacean–habitat modeling », in : *Marine Ecology Progress Series* 310, p. 271-295.

Reidenberg, Joy S et Jeffrey T Laitman (2007), « Discovery of a low frequency sound source in Mysticeti (baleen whales) : anatomical establishment of a vocal fold homolog », in : *The Anatomical Record : Advances in Integrative Anatomy and Evolutionary Biology : Advances in Integrative Anatomy and Evolutionary Biology* 290.*6*, p. 745-759.

Rembauville, Mathieu et Olivier Dequincey (2024), *Le plateau des Kerguelen, du volcanisme à la pompe biologique de carbone*, Consulté le 4 septembre 2025, url : https://planet-terre.ens-lyon.fr/ressource/Kerguelen-volcan-carbone.xml.

Rendell, Luke E et Hal Whitehead (2003), « Vocal clans in sperm whales (Physeter macrocephalus) », in : *Proceedings of the Royal Society of London. Series B : Biological Sciences* 270.*1512*, p. 225-231.

Réserve naturelle nationale des Terres australes françaises (oct. 2006), Créée par décret n° 2006-1211 du 3 octobre 2006, étendue en 2016, url : https://reserves-naturelles.org/reserves/terres-australes-francaises/.

Riera, Amalis, John K Ford et N Ross Chapman (2013), « Effects of different analysis techniques and recording duty cycles on passive acoustic monitoring of killer whales », in : *The Journal of the Acoustical Society of America* 134.*3*, p. 2393-2404.

Rintoul, Stephen R (2007), « Rapid freshening of Antarctic Bottom Water formed in the Indian and Pacific oceans », in : *Geophysical Research Letters* 34.*6*.

Risch, Denise, Nicholas J Gales et al. (2014), « Mysterious bio-duck sound attributed to the Antarctic minke whale (Balaenoptera bonaerensis) », in : *Biology letters* 10.*4*, p. 20140175.

Risch, Denise, Samuel C Wilson et al. (2019), « Seasonal and diel acoustic presence of North Atlantic minke whales in the North Sea », in : *Scientific Reports* 9.*1*, p. 3571.
Roberts, Jason J et al. (2016), « Habitat-based cetacean density models for the US Atlantic and Gulf of Mexico », in : *Scientific reports* 6.*1*, p. 22615.
Rock, Pauline De et al. (1er mars 2023), *État de référence du parc éolien en mer de Dieppe Le Tréport. Résultats des mesures de suivi SE1, SE2, SE2bis, SE2ter et E5*, rapp. tech. Contrat n°2020542, version Version 1, 311 pages. Rapport rédigé par Biotope., Le Colisée La Défense, Bât. A - 10, av. de l'Arche, 92400 Courbevoie, France : Biotope.
Romagosa, Miriam et al. (2021), « Food talk : 40-Hz fin whale calls are associated with prey biomass », in : *Proceedings of the Royal Society B* 288.*1954*, p. 20211156.
Roman, Joe et al. (2014), « Whales as marine ecosystem engineers », in : *Frontiers in Ecology and the Environment* 12.*7*, p. 377-385.
Royer, Jean-Yves (2009), *OHA-SIS-BIO-OBSERVATOIRE HYDROACOUSTIQUE*, doi : 10.18142/229.
Saji, NH et al. (1999), « A dipole mode in the tropical Indian Ocean », in : *Nature* 401.*6751*, p. 360-363.
Samaran, F., M. Torterotot et J. Béesau (2023), *Synthèse des données du programme CETIROISE (Mise en place d'un observatoire acoustique des cétacés dans le Parc naturel marin d'Iroise - 2022-2023)*, Rapport final, Parc naturel marin d'Iroise et Plan de Relance.
Samaran, Flore, Olivier Adam et Christophe Guinet (2010a), « Detection range modeling of blue whale calls in Southwestern Indian Ocean », in : *Applied Acoustics* 71.*11*, p. 1099-1106.
— (2010b), « Discovery of a mid-latitude sympatric area for two Southern Hemisphere blue whale subspecies », in : *Endangered Species Research* 12.*2*, p. 157-165.
Samaran, Flore, Christophe Guinet et al. (2010), « Source level estimation of two blue whale subspecies in southwestern Indian Ocean », in : *The Journal of the Acoustical Society of America* 127.*6*, p. 3800-3808.
Samaran, Flore, Kathleen M Stafford et al. (2013), « Seasonal and geographic variation of southern blue whale subspecies in the Indian Ocean », in : *PloS one* 8.*8*, e71561.
SAMBAH Project (2023), *Static Acoustic Monitoring of the Baltic Sea Harbour Porpoise*, Consulté le 24 juin 2025, url : https://www.sambah.org/.
Sánchez-Mercado, AY, JR Ferrer-Paris et Janet Franklin (2010), « Mapping species distributions : spatial inference and prediction », in : *Oryx* 44.*4*, p. 615.
Santos, M Begoña et al. (2013), « Long-term variation in common dolphin diet in relation to prey abundance », in : *Marine Ecology Progress Series* 481, p. 249-268.
Saurel, Jean-Marie et al. (2021), « Mayotte seismic crisis : building knowledge in near real-time by combining land and ocean-bottom seismometers, first results », in : *Geophysical Journal International* 228.*2*, p. 1281-1293.
Savoca, Matthew S et al. (2021), « Baleen whale prey consumption based on high-resolution foraging measurements », in : *Nature* 599.*7883*, p. 85-90.

Schall, Elena et al. (2019), « Visual and passive acoustic observations of blue whale trios from two distinct populations », in : *Marine Mammal Science* 36.*1*, p. 365-374.
Schoeman, Renée P, Claire Patterson-Abrolat et Stephanie Plön (2020), « A global review of vessel collisions with marine animals », in : *Frontiers in Marine Science* 7, p. 292.
Schott, Friedrich A, Shang-Ping Xie et Julian P McCreary Jr (2009), « Indian Ocean circulation and climate variability », in : *Reviews of Geophysics* 47.*1*.
Schouten, Mathijs W et al. (2003), « Eddies and variability in the Mozambique Channel », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 50.*12-13*, p. 1987-2003.
Science Primer (2020), *Langmuir Circulation*, `https://scienceprimer.com/langmuir-circulation`, Consulté le 27 août 2025.
Sears, Richard et Finn Larsen (2002), « Long range movements of a blue whale (Balaenoptera musculus) between the Gulf of St. Lawrence and West Greenland », in : *Marine mammal science* 18.*1*, p. 281-285.
Shabangu, Fannie W, Ken P Findlay et al. (2019), « Seasonal occurrence and diel calling behaviour of Antarctic blue whales and fin whales in relation to environmental conditions off the west coast of South Africa », in : *Journal of Marine Systems* 190, p. 25-39.
Shabangu, Fannie W, Tessa Munoz et al. (2024), « Diverse baleen whale acoustic occurrence around two sub-Antarctic islands : A tale of residents and visitors », in : *Scientific Reports* 14.*1*, p. 21663.
Shabangu, Fannie W, Dawit Yemane et al. (2017), « Modelling the effects of environmental conditions on the acoustic occurrence and behaviour of Antarctic blue whales », in : *PLoS One* 12.*2*, e0172705.
Shannon, Claude Elwood (2001), « A mathematical theory of communication », in : *ACM SIGMOBILE mobile computing and communications review* 5.*1*, p. 3-55.
Shannon, RANKIN' et al. (2005), « Vocalisations of Antarctic blue whales, Balaenoptera musculus intermedia, recorded during the 2001/2002 and 2002/2003 IWC/SOWER circumpolar cruises, Area V, Antarctica », in : *J. Cetacean Res. Manag* 7, p. 13-20.
Shono, Hiroshi (2008), « Application of the Tweedie distribution to zero-catch data in CPUE analysis », in : *Fisheries Research* 93.*1-2*, p. 154-162.
Simon, Malene et al. (2010), « Singing behavior of fin whales in the Davis Strait with implications for mating, migration and foraging », in : *The Journal of the Acoustical Society of America* 128.*5*, p. 3200-3210.
Širović, Ana, John A Hildebrand et Sean M Wiggins (2007), « Blue and fin whale call source levels and propagation range in the Southern Ocean », in : *The Journal of the Acoustical Society of America* 122.*2*, p. 1208-1215.
Širović, Ana, John A Hildebrand, Sean M Wiggins et al. (2004), « Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 51.*17-19*, p. 2327-2344.
Širović, Ana et Erin M. Oleson (2022), « The Bioacoustics of Blue Whales—Global Diversity and Behavioral Variability in a Foraging Specialist », in : *Ethology and*

Behavioral Ecology of Mysticetes, sous la dir. de Christopher W. Clark et Ellen C. Garland, Cham : Springer International Publishing, p. 195-221, isbn : 978-3-030-98449-6, doi : 10.1007/978-3-030-98449-6_9, url : https://doi.org/10.1007/978-3-030-98449-6_9.

Soberón, Jorge (2007), « Grinnellian and Eltonian niches and geographic distributions of species », in : *Ecology letters* 10.*12*, p. 1115-1123.

Socheleau, François-Xavier et Flore Samaran (2018), « Detection of Mysticete Calls : a Sparse Representation-Based Approach », in : *Technical report*.

Sol, Morgane et al. (2024), « Temporal patterns in dolphin foraging activity in the Mediterranean Sea : insights from vocalisations recorded during the ACCOBAMS Survey Initiative », in : *Frontiers in Marine Science* 11, p. 1378524.

Stafford, Kathleen M, DelWayne R Bohnenstiehl et al. (2004), « Antarctic-type blue whale calls recorded at low latitudes in the Indian and eastern Pacific Oceans », in : *Deep Sea Research Part I : Oceanographic Research Papers* 51.*10*, p. 1337-1346.

Stafford, Kathleen M, Germain Boussarie et al. (2023), « Acoustic detections and sightings of blue whales Balaenoptera musculus in the Seychelles, western tropical Indian Ocean (2020-2022) », in : *Endangered Species Research* 52, p. 203-208.

Stafford, Kathleen M, Emily Chapp et al. (2011), « Seasonal detection of three types of "pygmy" blue whale calls in the Indian Ocean », in : *Marine Mammal Science* 27.*4*, p. 828-840.

Stafford, Kathleen M, John J Citta et al. (2021), « Bowhead and beluga whale acoustic detections in the western Beaufort Sea 2008–2018 », in : *Plos one* 16.*6*, e0253929.

Stanistreet, Joy E, Douglas P Nowacek, Simone Baumann-Pickering et al. (2017), « Using passive acoustic monitoring to document the distribution of beaked whale species in the western North Atlantic Ocean », in : *Canadian Journal of Fisheries and Aquatic Sciences* 74.*12*, p. 2098-2109.

Stanistreet, Joy E, Douglas P Nowacek, Andrew J Read et al. (2016), « Effects of duty-cycled passive acoustic recordings on detecting the presence of beaked whales in the northwest Atlantic », in : *The Journal of the Acoustical Society of America* 140.*1*, EL31-EL37.

Sutton, Alicia L, K Curt S Jenner et Micheline-Nicole M Jenner (2019), « Habitat associations of cetaceans and seabirds in the tropical eastern Indian Ocean », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 166, p. 171-186.

Szesciorka, Angela R et al. (2020), « Timing is everything : Drivers of interannual variability in blue whale migration », in : *Scientific reports* 10.*1*, p. 7710.

Thewissen, JGM, Lisa Noelle Cooper, John C George et al. (2009), « From land to water : the origin of whales, dolphins, and porpoises », in : *Evolution : Education and Outreach* 2, p. 272-288.

Thewissen, Johannes GM, Lisa Noelle Cooper, Mark T Clementz et al. (2007), « Whales originated from aquatic artiodactyls in the Eocene epoch of India », in : *Nature* 450.*7173*, p. 1190-1194.

Thode, Aaron M, Gerald L D'Spain et William A Kuperman (1999), « Physical constraints on blue whale sound production mechanisms derived from matched-field processing », in : *The Journal of the Acoustical Society of America* 105.*2_Supplement*, p. 993-993.

Thomas, Mark et al. (2019), « Marine mammal species classification using convolutional neural networks and a novel acoustic representation », in : *Joint European conference on machine learning and knowledge discovery in databases*, Springer, p. 290-305.

Thomisch, Karolin, Olaf Boebel, Christopher W Clark et al. (2016), « Spatio-temporal patterns in acoustic presence and distribution of Antarctic blue whales Balaenoptera musculus intermedia in the Weddell Sea », in : *Endangered Species Research* 30, p. 239-253.

Thomisch, Karolin, Olaf Boebel, Daniel P Zitterbart et al. (2015), « Effects of subsampling of passive acoustic recordings on acoustic metrics », in : *The Journal of the Acoustical Society of America* 138.*1*, p. 267-278.

Thompson, Paul O et al. (1996), « Underwater sounds of blue whales, Balaenoptera musculus, in the Gulf of California, Mexico », in : *Marine Mammal Science* 12.*2*, p. 288-293.

Thorson, James T et al. (2020), « Empirical orthogonal function regression : Linking population biology to spatial varying environmental conditions using climate projections », in : *Global Change Biology* 26.*8*, p. 4638-4649.

Thums, Michele et al. (2022), « Pygmy blue whale movement, distribution and important areas in the Eastern Indian Ocean », in : *Global Ecology and Conservation* 35, e02054.

Todd, Nicole Rose Eileen, Ailbhe Sarah Kavanagh et al. (2023), « What the F-POD ? Comparing the F-POD and C-POD for monitoring of harbor porpoise (Phocoena phocoena) », in : *Ecology and Evolution* 13.*6*, e10186.

Todd, Victoria LG et Laura D Williamson (2022), « Cetacean distribution in relation to oceanographic features at the Kerguelen Plateau », in : *Polar Biology* 45.*1*, p. 113-126.

Torterotot, Maëlle (2020), « Traitement et analyse de données bioacoustiques dans l'océan Indien austral : application aux baleines bleues », français, NNT : 2020BRES0058, Thèse de doctorat, Université de Bretagne occidentale - Brest, url : `https://theses.hal.science/tel-03372735`.

Torterotot, Maëlle, Julie Béesau et al. (2022), « Assessing marine mammal diversity in remote Indian Ocean regions, using an acoustic glider », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 206, p. 105204.

Torterotot, Maëlle, Jean-Yves Royer et Flore Samaran (2019), « Detection strategy for long-term acoustic monitoring of blue whale stereotyped and non-stereotyped calls in the Southern Indian Ocean », in : *OCEANS 2019-Marseille*, IEEE, p. 1-10.

Torterotot, Maëlle, Flore Samaran et Jean-Yves Royer (2023), « Long-term acoustic monitoring of nonstereotyped blue whale calls in the southern Indian Ocean », in : *Marine Mammal Science* 39.*2*, p. 594-610.

Torterotot, Maëlle, Flore Samaran, Kathleen M Stafford et al. (2020), « Distribution of blue whale populations in the Southern Indian Ocean based on a decade of

acoustic monitoring », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 179, p. 104874.

Turkish Marine Research Foundation (2021), *BlackCeTrends – Large coverage initiative for acoustic survey of cetaceans in the Black Sea*, Consulté le 24 juin 2025, url : https://tudav.org/en/active-projects/blackcetrends/.

Tyack, Peter L (2000), *Functional aspects of cetacean communication. Pages 270–307inJ. MANN, R. CONNOR, PL TYACK AND H. WHITEHEAD, eds. Cetacean societies : Field studies of dolphins and whales*.

Tyack, Peter L et Edward H Miller (2002), « Vocal anatomy, acoustic communication and echolocation », in : *Marine mammal biology : an evolutionary approach* 59, p. 142-84.

Van Geel, Nienke CF et al. (2022), « Monitoring cetacean occurrence and variability in ambient sound in Scottish offshore waters », in : *Frontiers in Remote Sensing* 3, p. 934681.

Vanderlaan, Angelia SM et Christopher T Taggart (2007), « Vessel collisions with whales : the probability of lethal injury based on vessel speed », in : *Marine mammal science* 23.*1*, p. 144-156.

Virgili, Auriane (jan. 2018), « Modelling distributions of rare marine species : the deep-diving cetaceans », Theses, Université de La Rochelle, url : https://theses.hal.science/tel-02009798.

Virgili, Auriane, Matthieu Authier et al. (2019), « Combining multiple visual surveys to model the habitat of deep-diving cetaceans at the basin scale : Large-scale modelling of deep-diving cetacean habitats », in : *Global Ecology and Biogeography* 28.*3*, p. 300-314.

Virgili, Auriane, Mélanie Racine et al. (2017), « Comparison of habitat models for scarcely detected species », in : *Ecological Modelling* 346, p. 88-98.

Volcanological, Mayotte et al. (2021), *Data collection of the Mayotte volcanological and seismological monitoring network (REVOSIMA) (Version 1.0)*, Version 1.0, doi : 10.18715/MAYOTTE.REVOSIMA, url : https://doi.org/10.18715/MAYOTTE.REVOSIMA.

Watkins, William A et al. (1987), « The 20-Hz signals of finback whales (B alaenoptera physalus) », in : *The Journal of the Acoustical Society of America* 82.*6*, p. 1901-1912.

Weilgart, Lindy S (2007), « The impacts of anthropogenic ocean noise on cetaceans and implications for management », in : *Canadian journal of zoology* 85.*11*, p. 1091-1116.

Weimerskirch, Henri (2007), « Are seabirds foraging for unpredictable resources ? », in : *Deep Sea Research Part II : Topical Studies in Oceanography* 54.*3-4*, p. 211-223.

Welsh, Alan H et al. (1996), « Modelling the abundance of rare species : statistical models for counts with extra zeros », in : *Ecological Modelling* 88.*1-3*, p. 297-308.

Wiggins, Sean M et John A Hildebrand (2007), « High-frequency Acoustic Recording Package (HARP) for broad-band, long-term marine mammal monitoring », in : *2007 symposium on underwater technology and workshop on scientific use of submarine cables and related technologies*, IEEE, p. 551-557.

Wikipédia (2025), *Carte de l'océan Indien*, Consulté le 27 août 2025, url : https://fr.wikipedia.org/wiki/Oc%C3%A9an_Indien.

WILDLABS (2025), *A standard for bioacoustic data - Safe and Sound*, Accès le 15 septembre 2025, URL : https://wildlabs.net/event/standard-bioacoustic-data-safe-and-sound.

WILLIAMS, Rob et al. (2014), « Severity of killer whale behavioral responses to ship noise : A dose–response study », in : *Marine pollution bulletin* 79.*1-2*, p. 254-260.

WOOD, S.N (2017a), *Generalized Additive Models : An Introduction with R*, 2e éd., Chapman et Hall/CRC.

WOOD, Simon N (2017b), *Generalized additive models : an introduction with R*, chapman et hall/CRC.

WOOD, Simon N, Natalya PYA et Benjamin SÄFKEN (2016), « Smoothing parameter and model selection for general smooth models », in : *Journal of the American Statistical Association* 111.*516*, p. 1548-1563.

WORM, Boris, Heike K LOTZE et Ransom A MYERS (2003), « Predator diversity hotspots in the blue ocean », in : *Proceedings of the National Academy of Sciences* 100.*17*, p. 9884-9888.

WU, Jianguo et Harbin LI (2006), « Concepts of scale and scaling », in : *Scaling and uncertainty analysis in ecology*, Springer, p. 3-15.

ZHANG, Zhiyuan et Gen LI (2023), « Multi-decadal enhancement in the influence of El Niño on the Indian Ocean dipole mode », in : *Climate Dynamics* 61.*11*, p. 5331-5347.

ZHONG, Ming et al. (2021), « Detecting, classifying, and counting blue whale calls with Siamese neural networks », in : *The Journal of the Acoustical Society of America* 149.*5*, p. 3086-3094.

ZUUR, Alain F et al. (2009), *Mixed effects models and extensions in ecology with R*, t. 574, Springer.

Annexe A

Cette annexe correspond à l'**Étude 2** du Chapitre 1 : *Effects of duty cycle on passive acoustic monitoring metrics : The case of odontocetes vocalizations*, actuellement soumise et en cours d'évaluation dans *JASA*.

APPENDIX A

TABLE IV. Parameters used for delphinid whistle detection with the Whistle and Moan Detector module in PAMGuard.

FFT Parameter	**Value**
FFT length	2048
FFT Hop	2048
Window	Hann
Whistle and Moan Detector Parameter	**Value**
Minimum frequency	5000 Hz
Maximum frequency	30000 Hz
Connection type	Connect 8 (slides and diagonals)
Minimum length	10 time slices
Minimum total size	20 pixels
Crossing and joining	Re-link across joins
Maximum cross length	5 time slices
Median filter length	61
Average subtraction constant	0.02
Gaussian kernel smoothing	ON
Threshold	7 dB

APPENDIX B

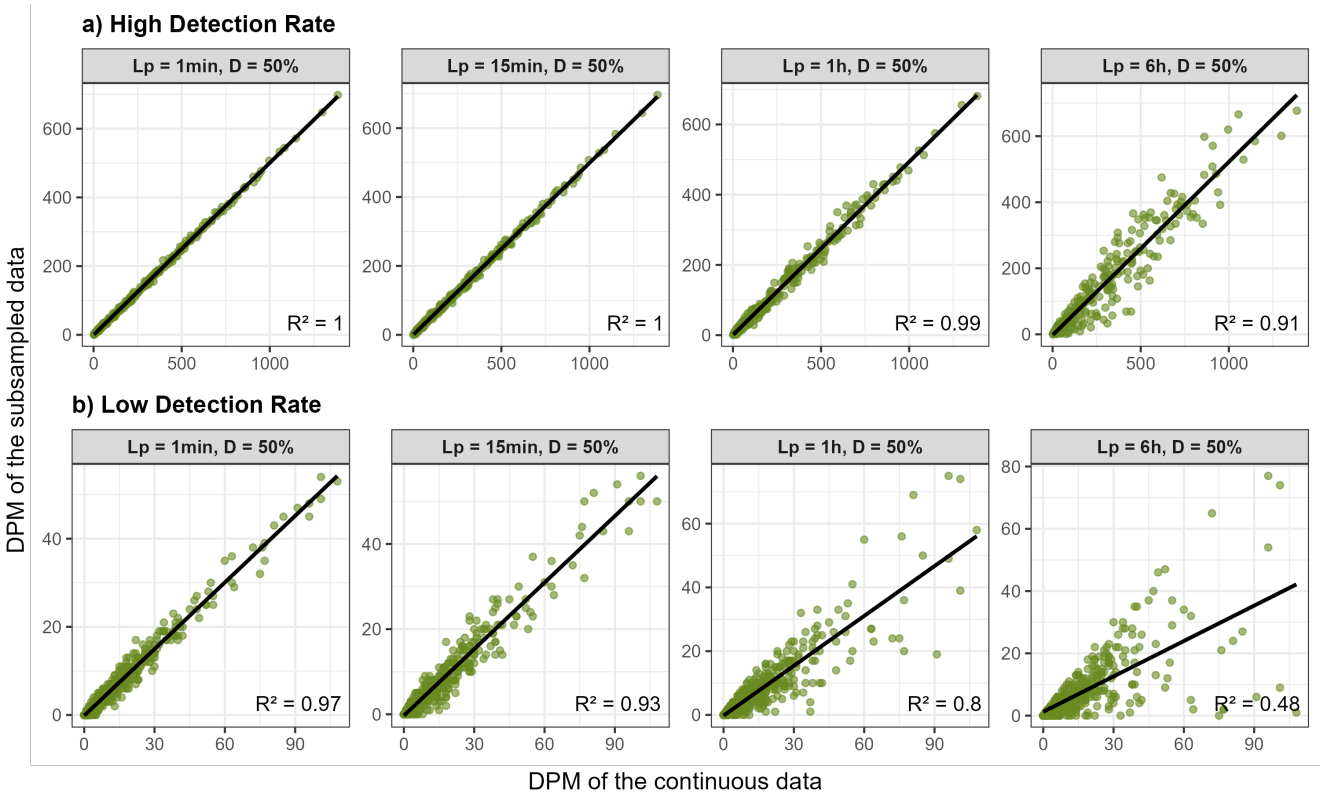


FIG. 11. Examples of the relationship between DPM of continuous data and DPM of subsampled data for delphinid whistles. Panel a) represents high detection rate conditions, and panel b) low detection rate conditions. Each panel shows results across four listening periods (L_p = 1 min, 15 min, 1 h, and 6 h). Data points (green circles) are fitted with a linear regression model (black line). The R^2 value of the linear fit is indicated on each panel.

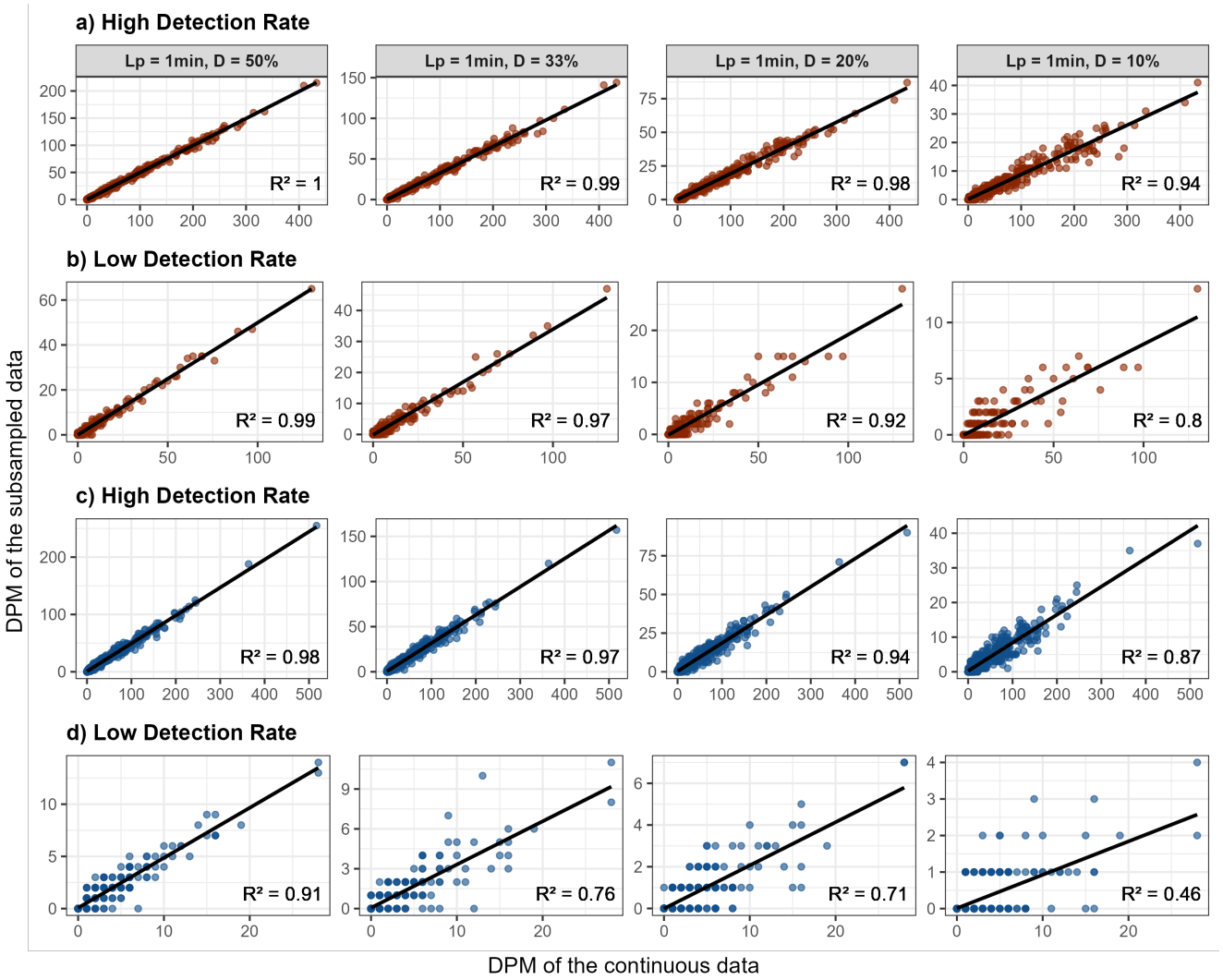


FIG. 12. Examples of the relationship between DPM of continuous data and DPM of subsampled data for delphinid and harbor porpoise clicks. Panels a) and c) correspond to high detection rate conditions, while panels b) and d) correspond to low detection rate conditions. Results are shown across four duty cycles ($D = 50\%$, 33%, 20%, and 10%) with a fixed listening period of $L_p = 1$ min. Red circles represent delphinid clicks, and blue circles represent harbor porpoise clicks. Data points are fitted with a linear regression model (black line). The R^2 value of the linear fit is indicated on each panel.

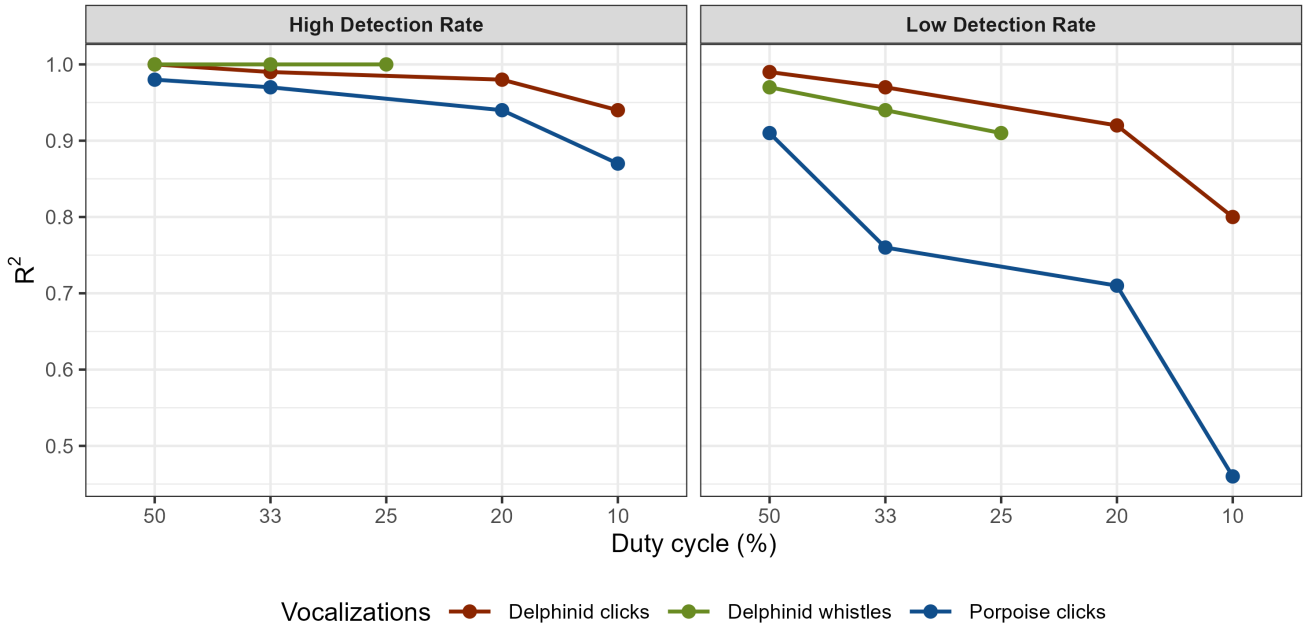


FIG. 13. Comparison of different duty cycles for a 1-minute listening period across three vocalization types for the seasonal patterns in daily detection positive minutes: delphinid whistles (green), delphinid clicks (red), and harbor porpoise clicks (blue). The left column shows sites with high detection rates, while the right column shows sites with low detection rates. Each point represents the R^2 value for a given duty cycle (D), ordered from left to right as 50%, 33%, 25%, 20%, and 10%.

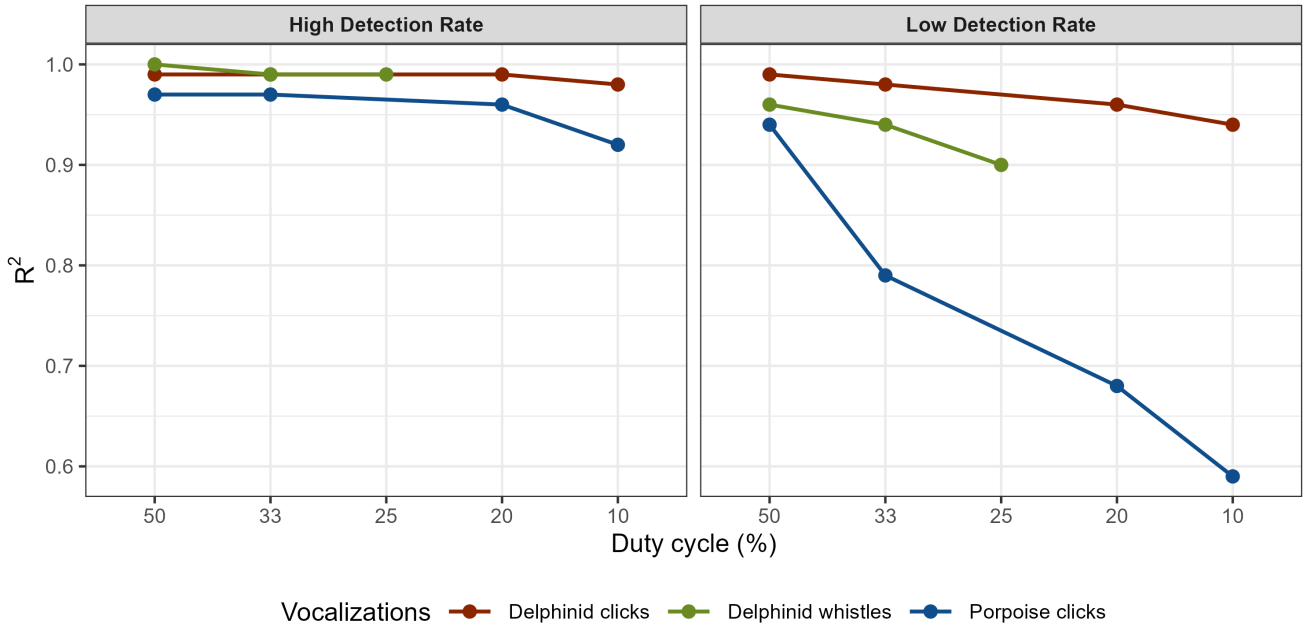


FIG. 14. Comparison of different duty cycles for a 1-minute listening period across three vocalization types for the diel patterns in daily detection positive hours: delphinid whistles (green), delphinid clicks (red), and harbor porpoise clicks (blue). The left column shows sites with high detection rates, while the right column shows sites with low detection rates. Each point represents the R^2 value for a given duty cycle (D), ordered from left to right as 50%, 33%, 25%, 20%, and 10%.

Annexe B

Cette annexe correspond à l'**Étude 1** du Chapitre 2 : *Modelling the influence of environmental factors on the acoustic presence of blue whale populations in the southern Indian Ocean*, publiée en 2025 dans la revue *Scientific Reports*.

General legend for the maps : The maps show the predicted monthly distribution of *log(DPM)* for the three acoustic populations of blue whales :

- Blue : baleine bleue Antarctique (ANT BW)
- Orange : baleine bleue pygmée du sud-est de l'océan Indien (SEIO PBW)
- Green : baleine bleue pygmée du sud-ouest de l'océan Indien (SWIO PBW)

Each map corresponds to one year. The color scale may vary from year to year.

Supplemental materials

Table Supplementary 1. Summary of acoustic effort for each site and acoustic population, including the total number of days recorded, minutes recorded, total positive minutes, and the percentage of days with acoustic presence.

Site	Acoustic Population	Total of Days Recorded	Total of Minutes Recorded	Total Positive Minutes	% of Days with Acoustic Presence
RTJ	ANT BW SEIO PBW SWIO PBW	1442	2,076,480	1182 88 406	82 6.1 28.2
MADE	ANT BW SEIO PBW SWIO PBW	1474	2,122,560	1333 117 1255	90.4 7.9 85.1
MAD	ANT BW SEIO PBW SWIO PBW	2548	3,669,120	2426 33 2004	95.2 1.3 78.6
MADW	ANT BW SEIO PBW SWIO PBW	1723	2,481,120	1624 0 1461	94.3 0 84.8
NEAMS	ANT BW SEIO PBW SWIO PBW	3489	5,024,160	3337 1139 2617	95.6 32.6 75
SSEIR	ANT BW SEIO PBW SWIO PBW	3251	4,681,440	2738 709 1716	84.2 21.8 52.8
NCRO	ANT BW SEIO PBW SWIO PBW	1630	2,347,200	1429 125 1262	87.7 7.6 77.4
SWAMS	ANT BW SEIO PBW SWIO PBW	3493	5,029,920	3296 1835 2713	94.4 52.5 77.7
WKER	ANT BW SEIO PBW SWIO PBW	4750	6,840,000	4059 1175 3247	85.5 24.7 68.4
ELAN	ANT BW SEIO PBW SWIO PBW	1835	2,642,400	1548 62 403	84.4 3.38 22

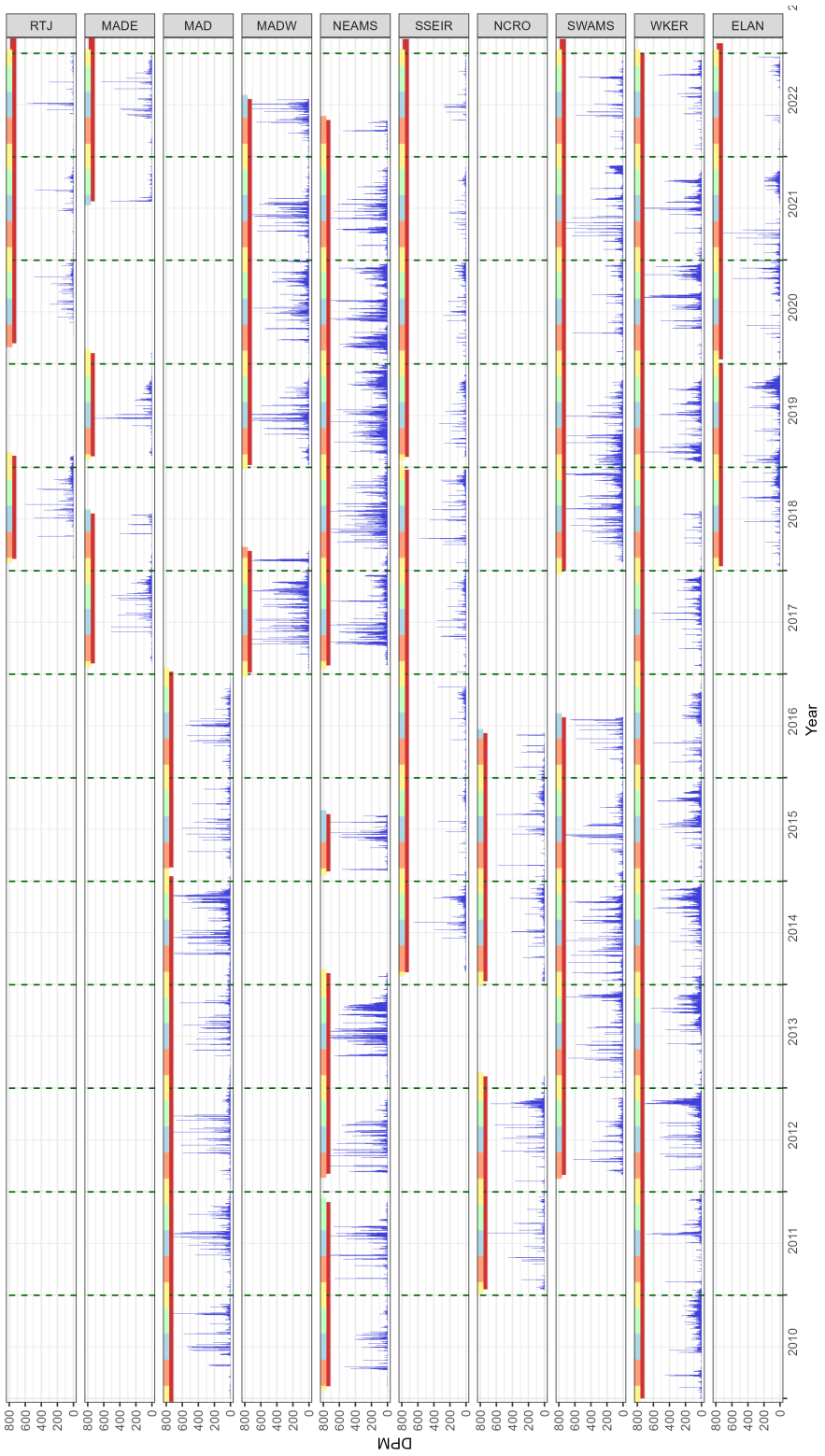


Figure Supplementary 1. Daily positive minutes (DPM) averaged by week for ANT BW acoustic population in the OHA-SISBIO hydrophone network throughout the deployment period. The red horizontal bars indicate the recording periods for each site. The horizontal bars in yellow, orange, blue, and green represent the seasons: summer, autumn, winter, and spring, respectively. The sites are ordered from north to south. The scale of the bars is not the same for each site.

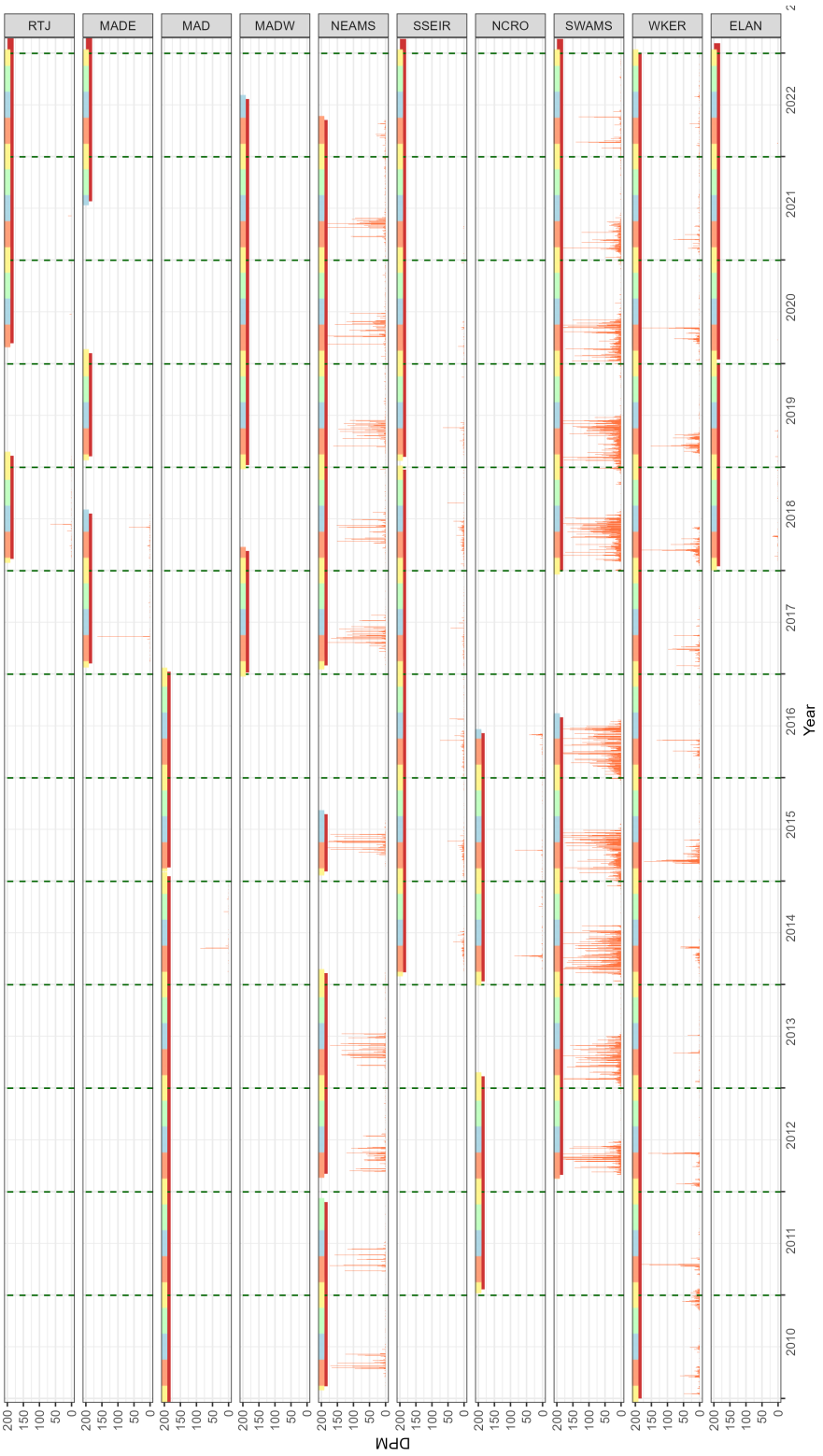


Figure Supplementary 2. Daily positive minutes (DPM) averaged by week for SEIO PBW acoustic population in the OHA-SISBIO hydrophone network throughout the deployment period. The red horizontal bars indicate the recording periods for each site. The horizontal bars in yellow, orange, blue, and green represent the seasons: summer, autumn, winter, and spring, respectively. The sites are ordered from north to south. The scale of the bars is not the same for each site.

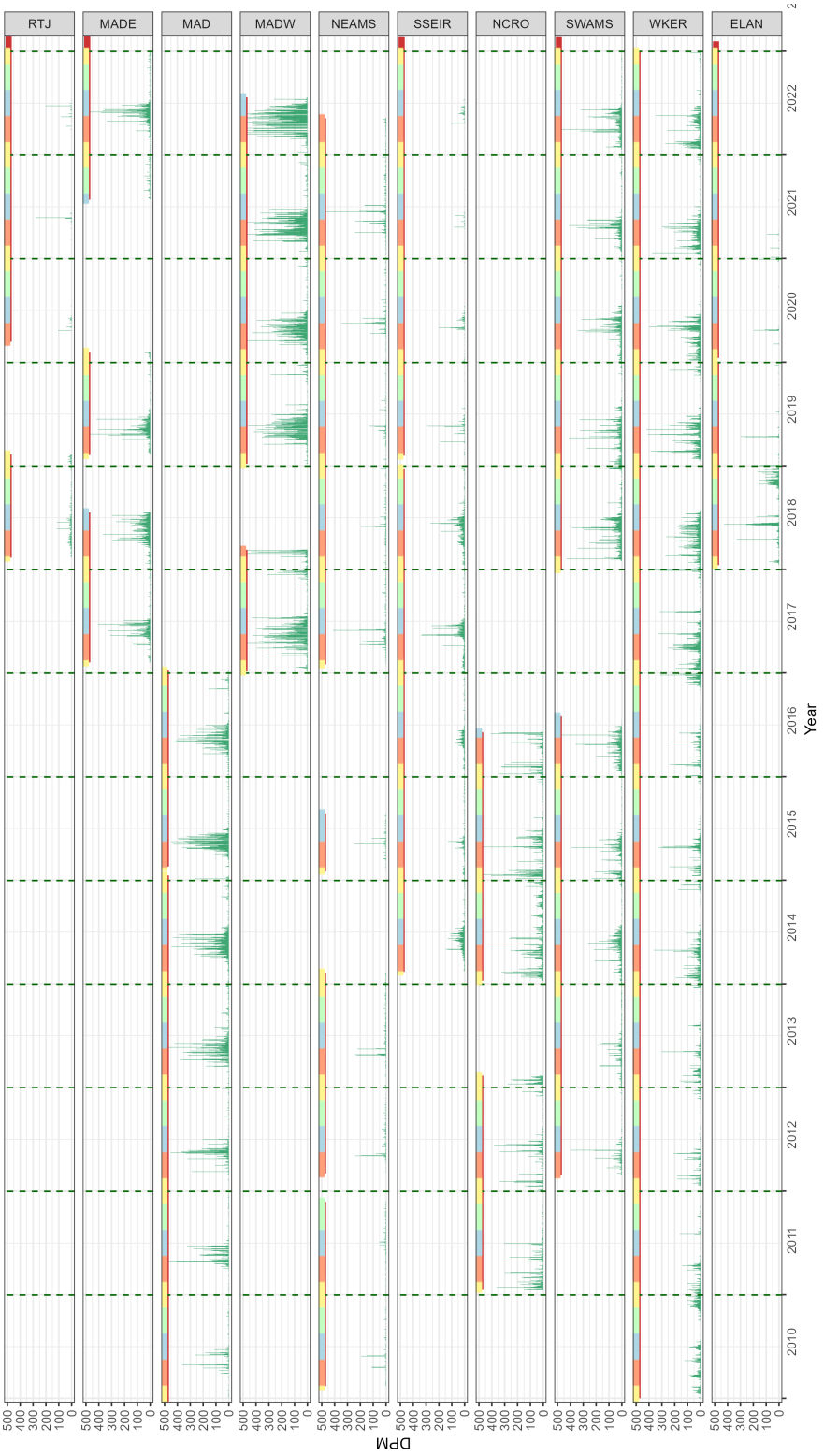


Figure Supplementary 3. Daily positive minutes (DPM) averaged by week for SWIO PBW acoustic population in the OHA-SISBIO hydrophone network throughout the deployment period. The red horizontal bars indicate the recording periods for each site. The horizontal bars in yellow, orange, blue, and green represent the seasons: summer, autumn, winter, and spring, respectively. The sites are ordered from north to south. The scale of the bars is not the same for each site.

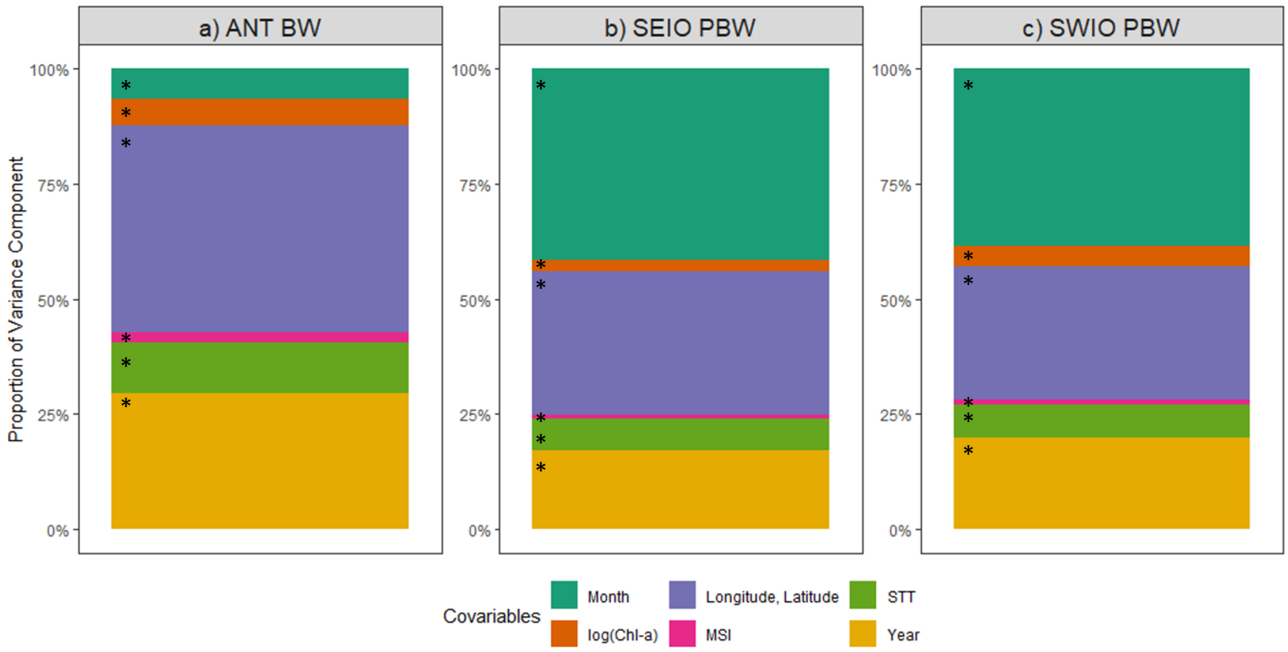


Figure Supplementary 4. Contribution (%) of environmental variables used in the GAM models for the three acoustic populations: (a) ANT BW, (b) SEIO PBW, and (c) SWIO PBW. Asterisks (*) indicate significant contribution ($p < 0.05$).

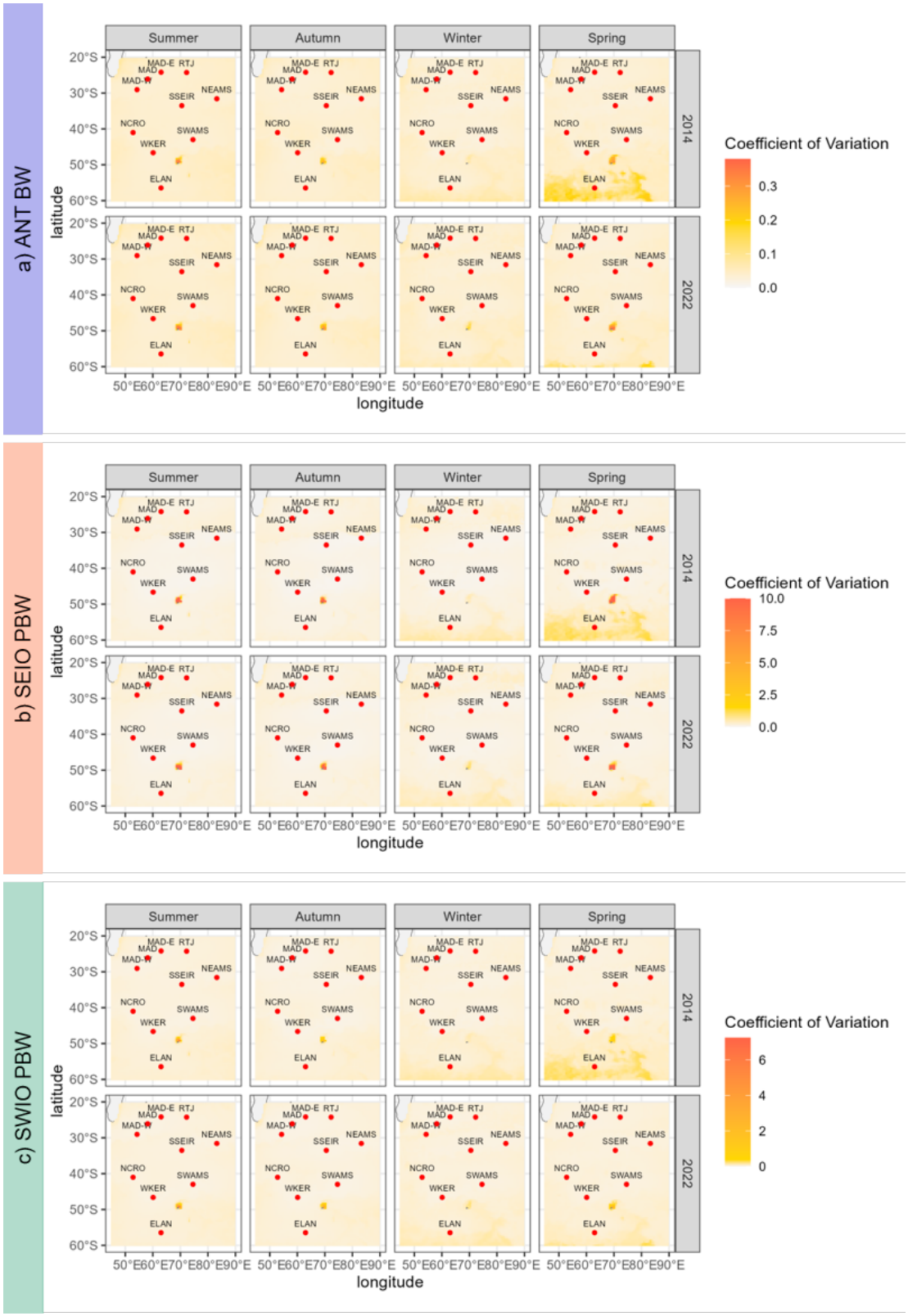


Figure Supplementary 5. Coefficient of variation of the predicted log(DPM) for the three acoustic populations: (a) ANT BW, (b) SEIO PBW, and (c) SWIO PBW across the four seasons, represented by January (summer), April (autumn), July (winter), and November (spring), for the years 2014 and 2022. The intensity of the shading indicates the magnitude of the CV (darker = higher variation), as shown by the color bar on the right. Red dots represent key observation locations.

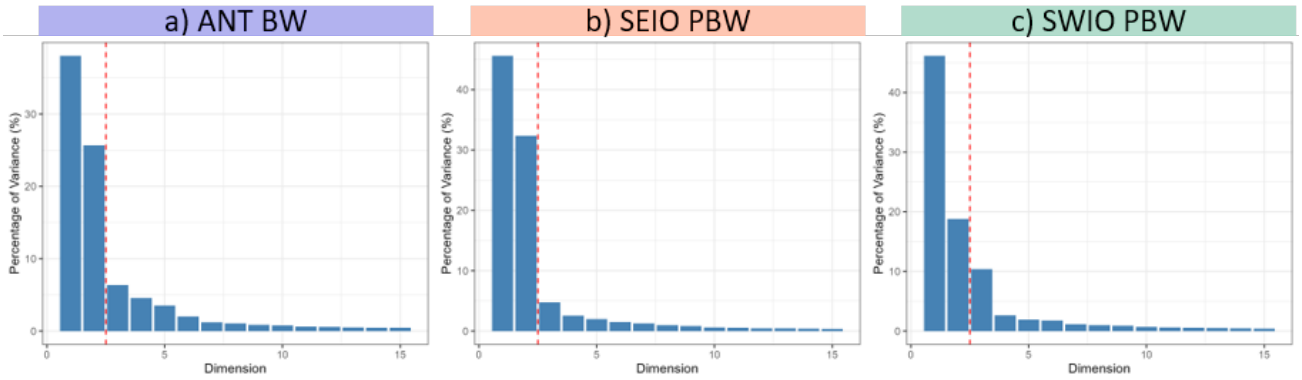


Figure Supplementary 6. Proportion of the total variance explained by each dimension of the EOF for each acoustic population. Dashed line: threshold used to filter the interpretable dimensions of the EOF.

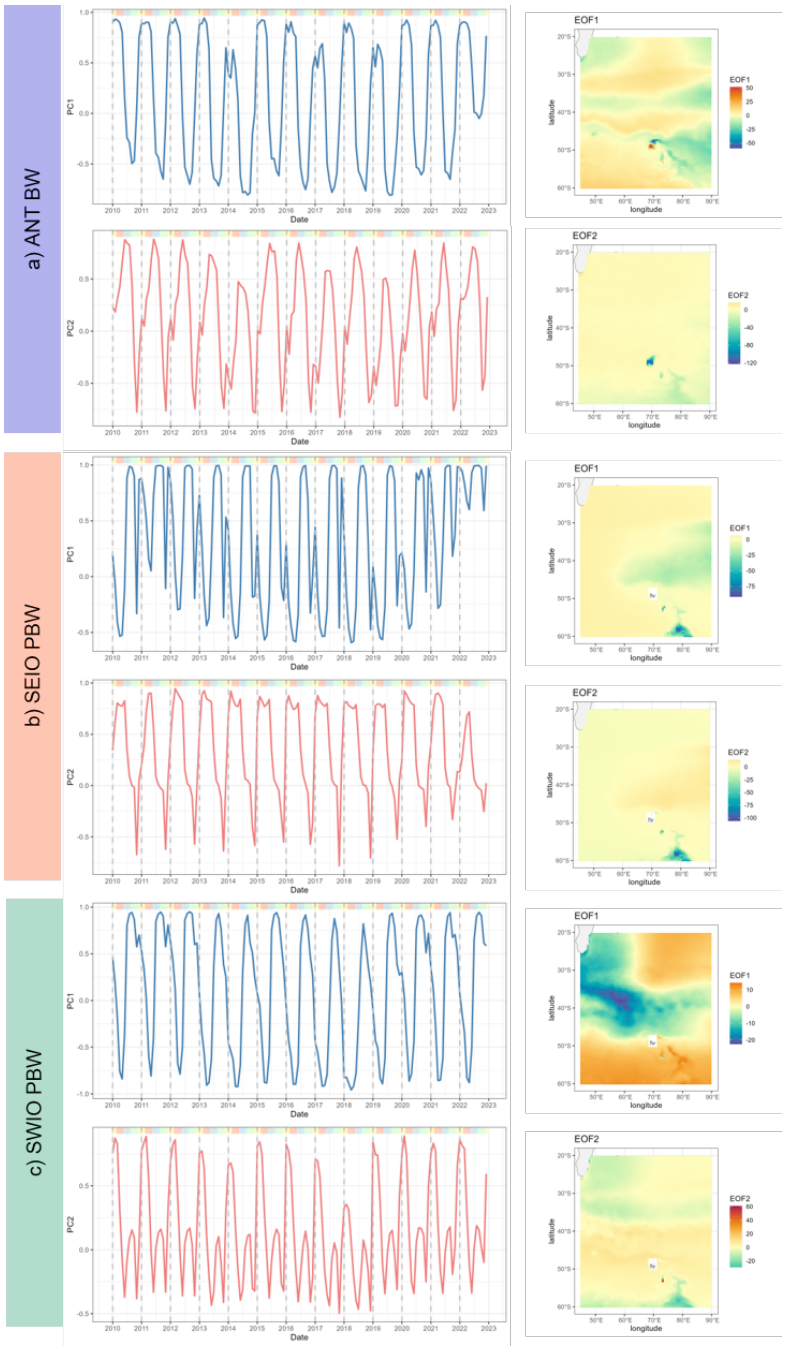
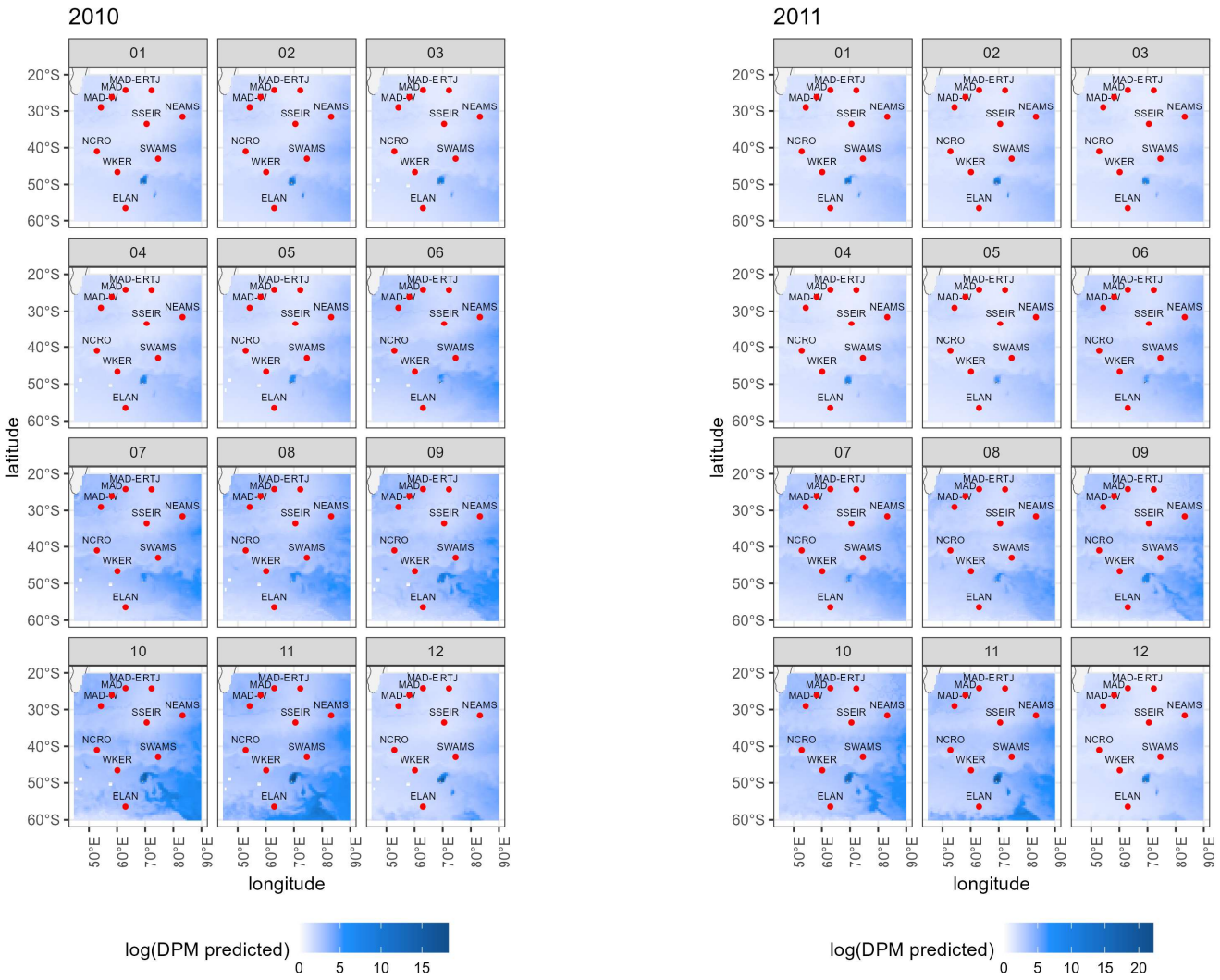
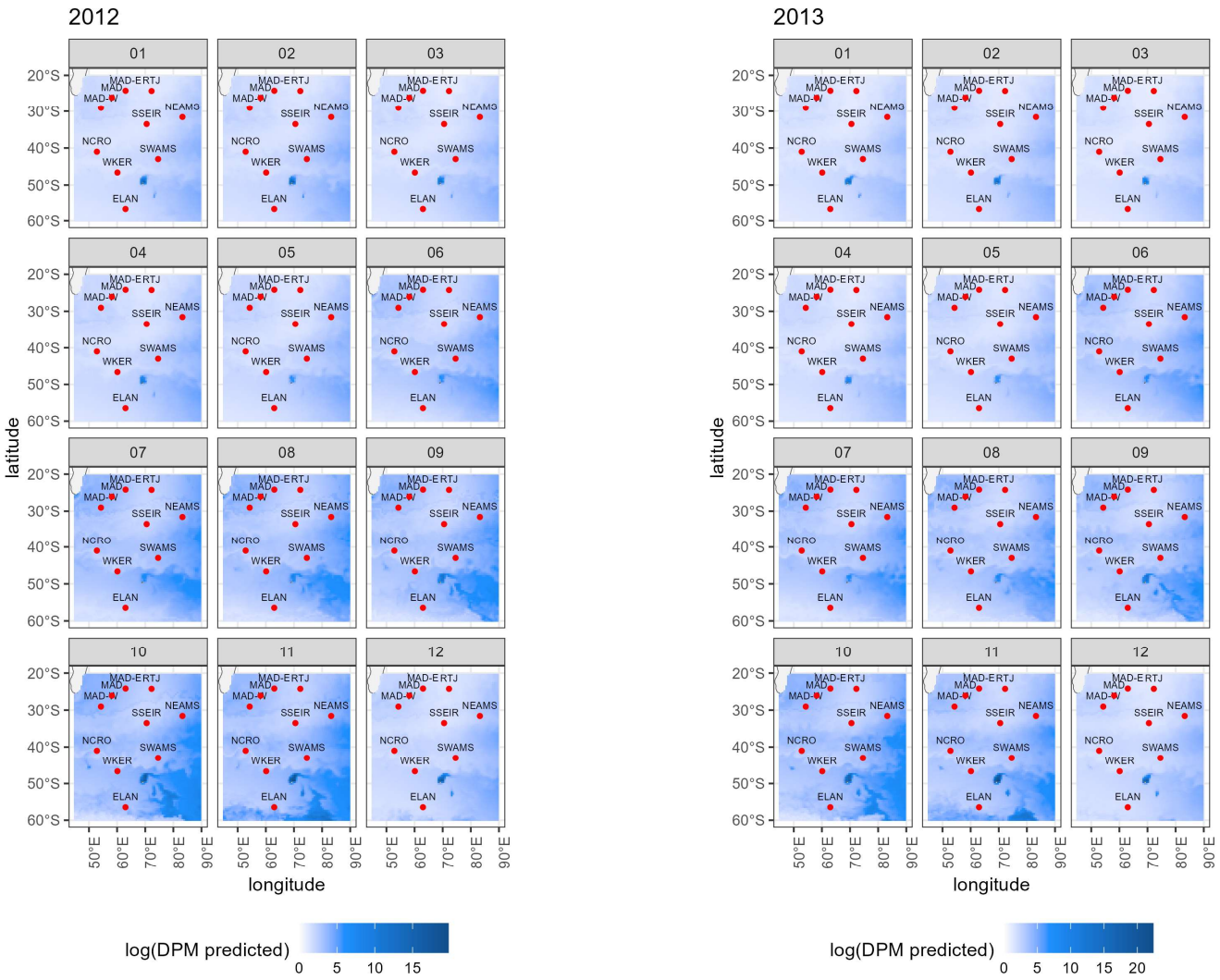
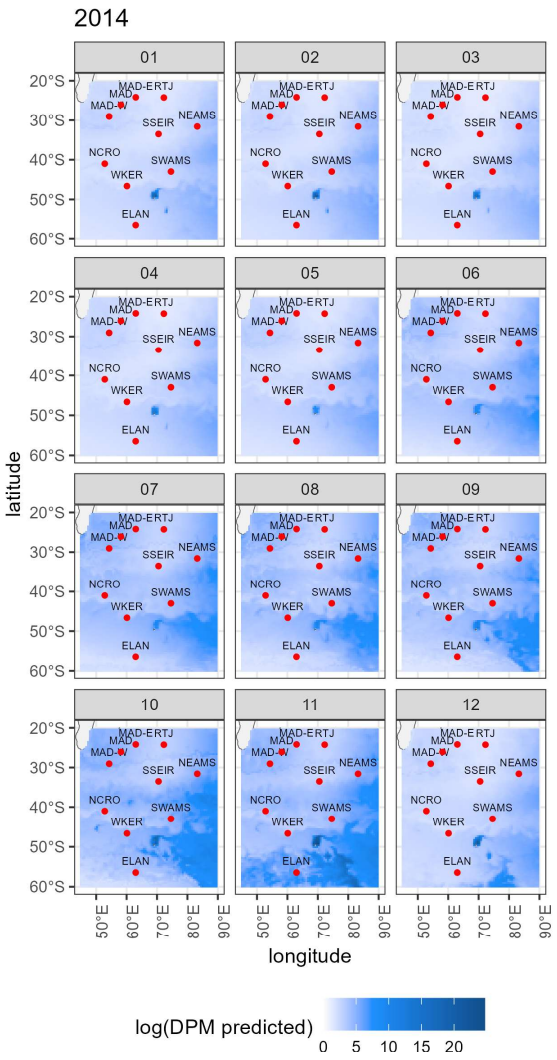
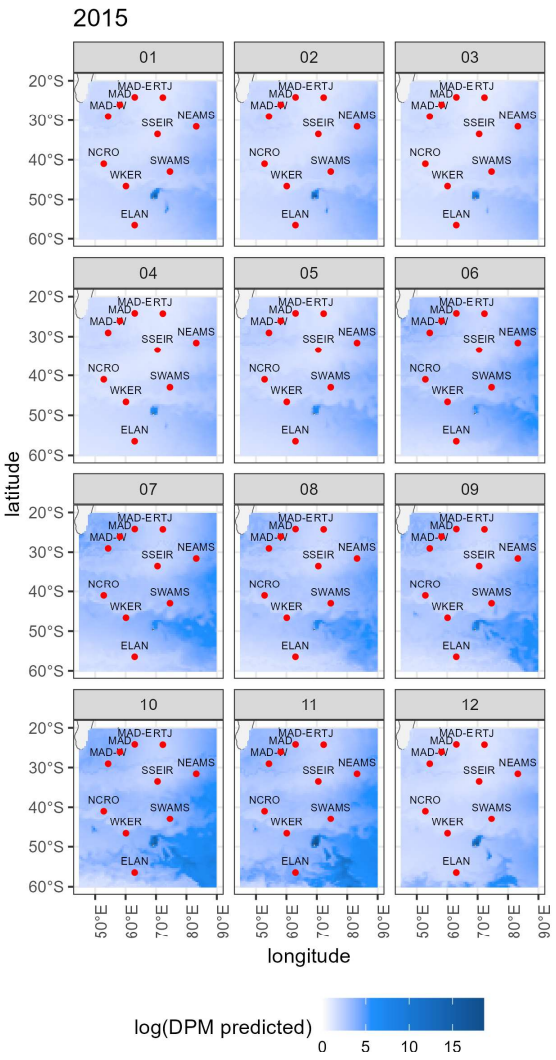


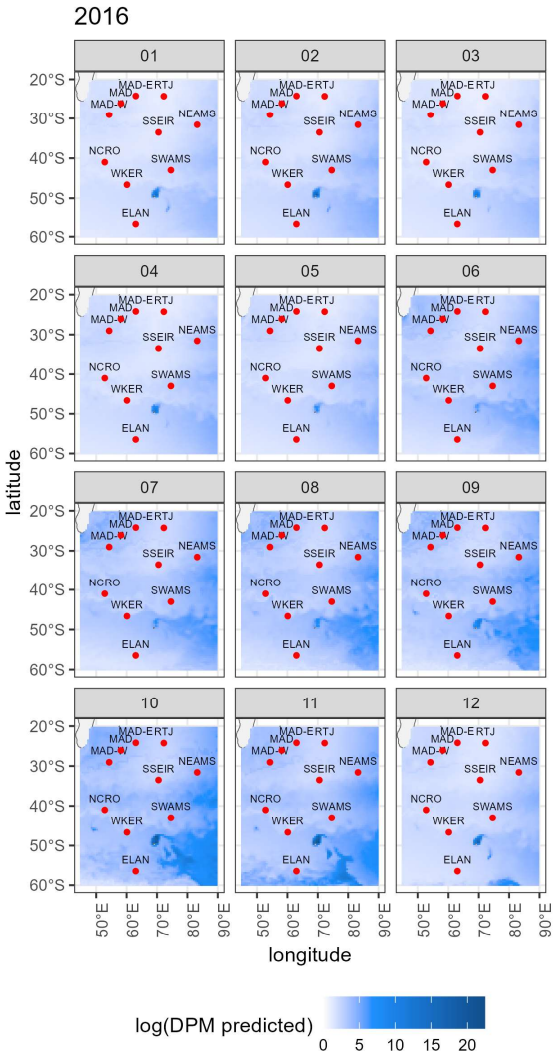
Figure Supplementary 7. Spatial factors (left) and EOF maps (right) for the three acoustic populations: a) ANT BW, b) SEIO PBW, and c) SWIO PBW. The dashed gray bars indicate the months of January for each year. The horizontal bars in yellow, orange, blue, and green represent the seasons: summer, autumn, winter, and spring, respectively.

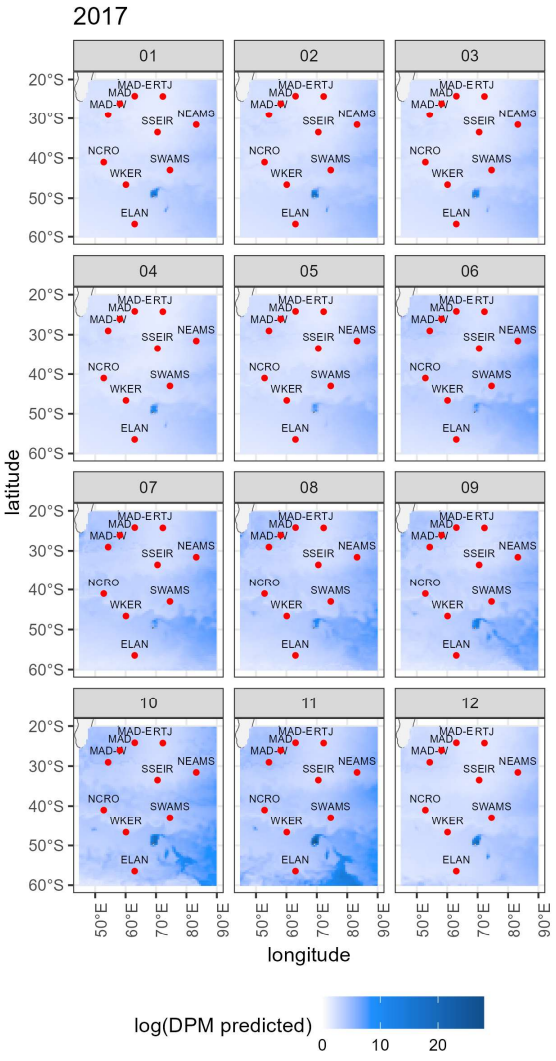

2010
2011
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20

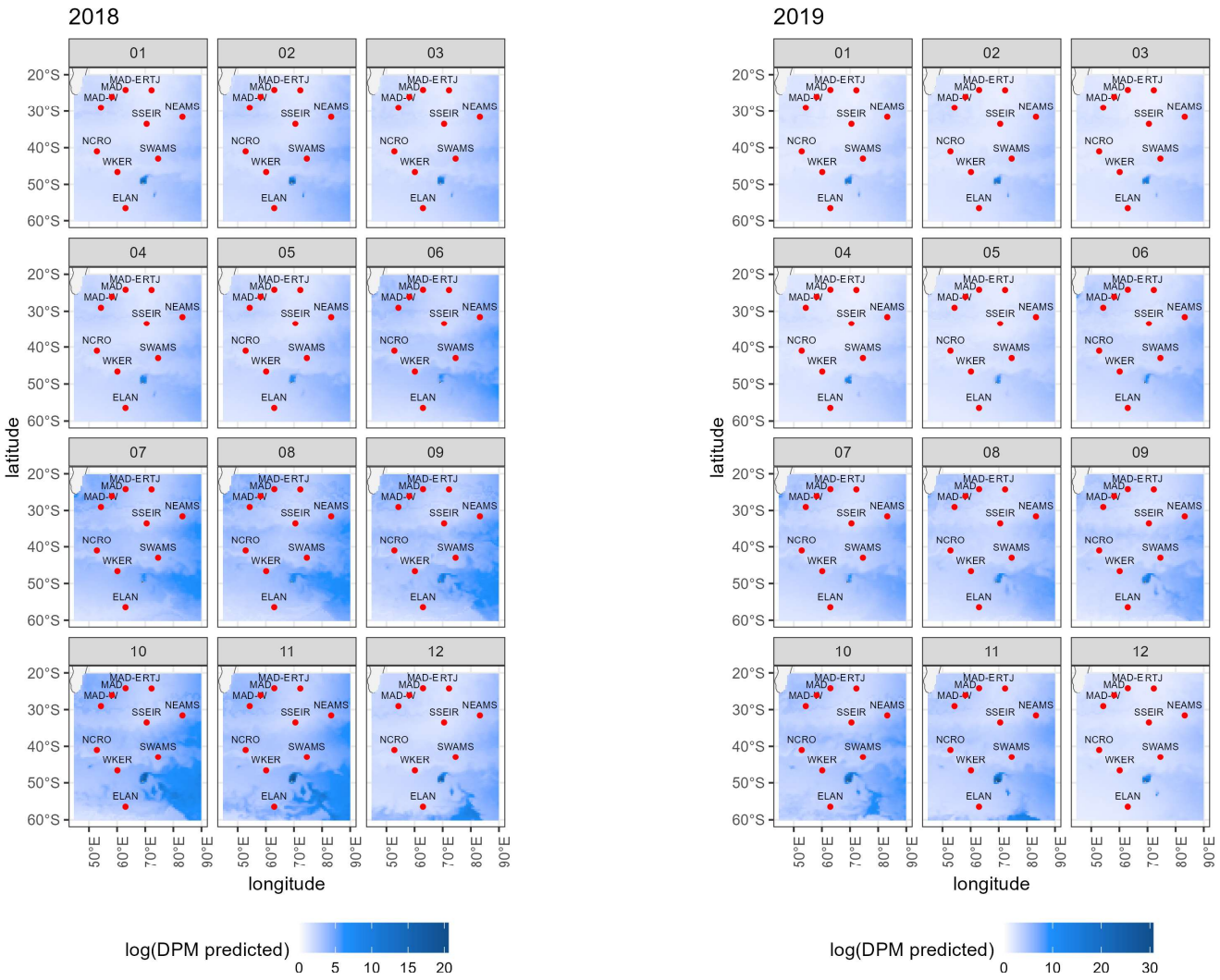

2012
2013
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20

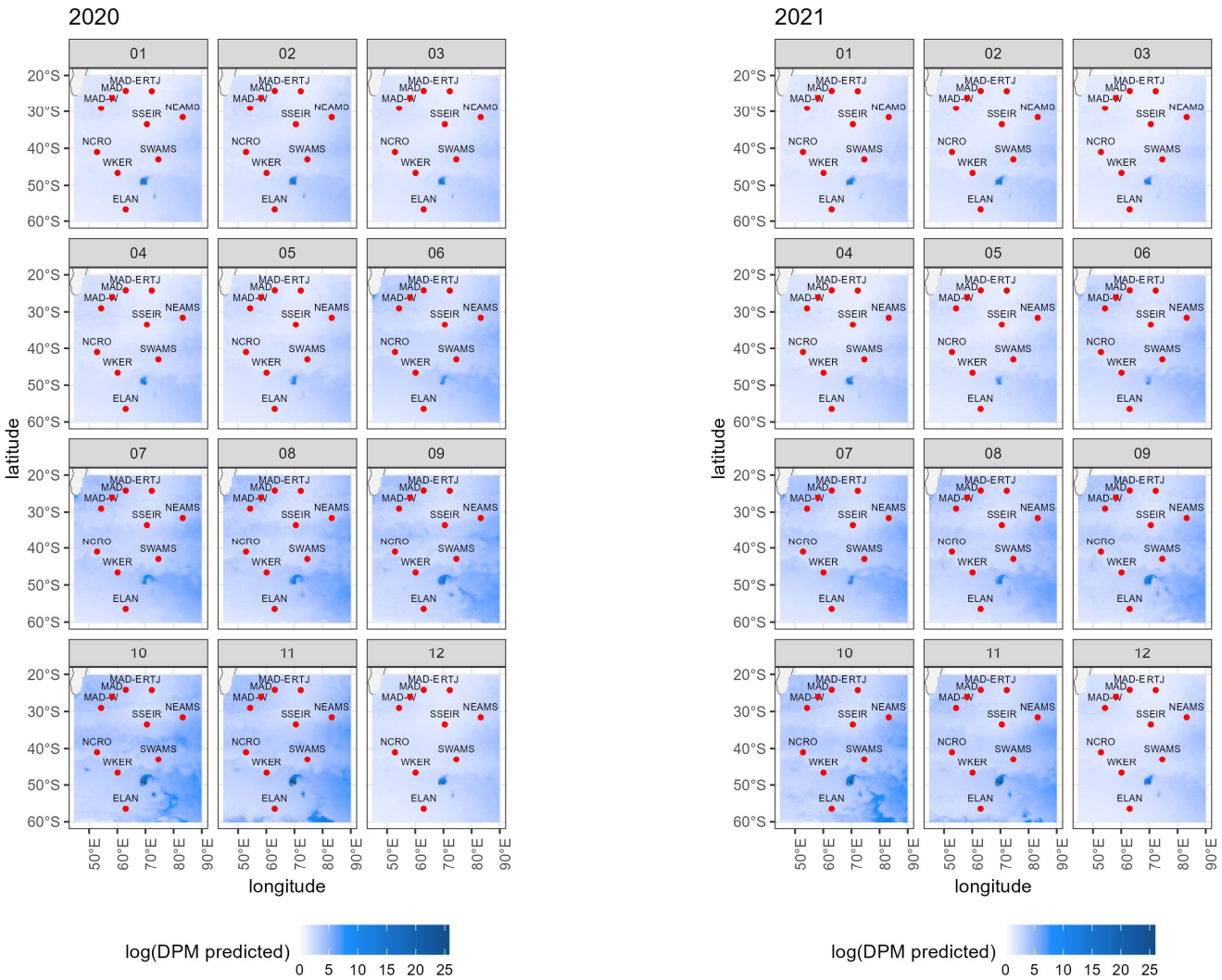

2014
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20

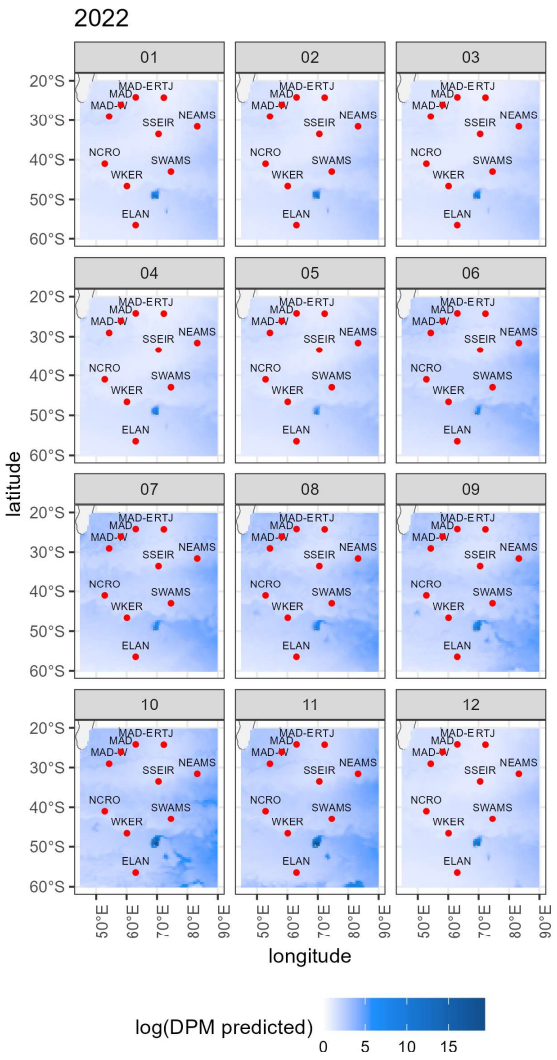

2015
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15

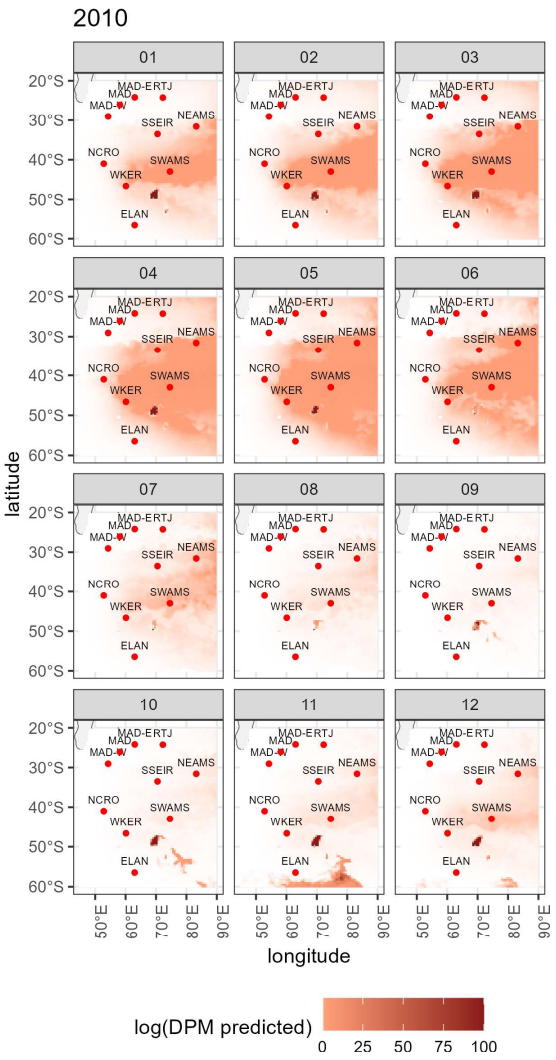

2016
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20

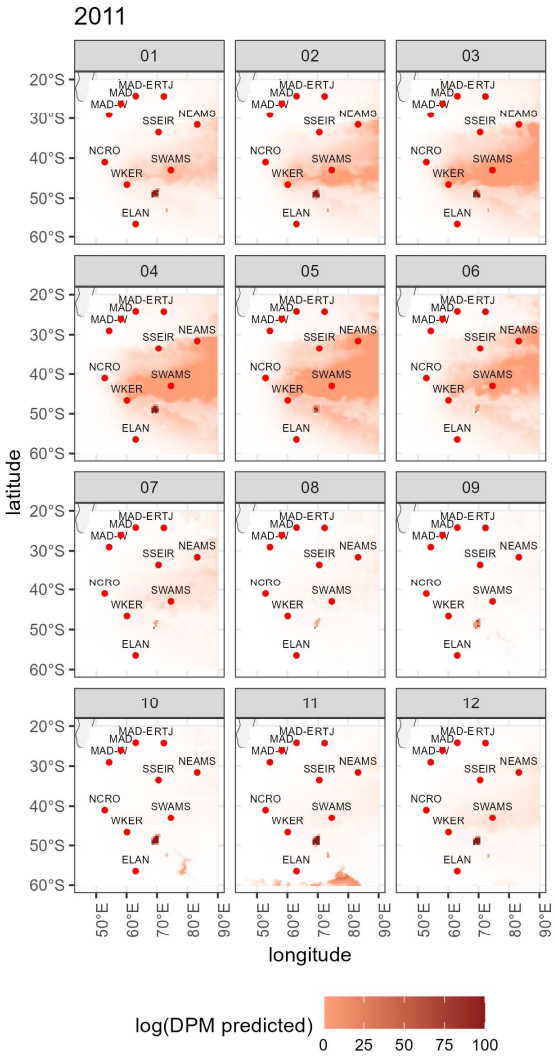

2017
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
10
20

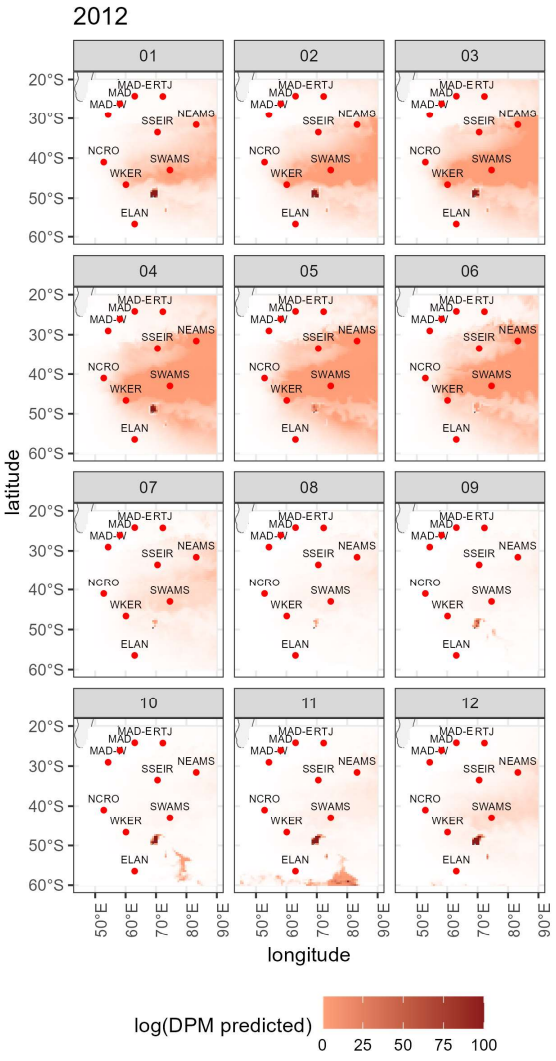

2018
2019
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20
30

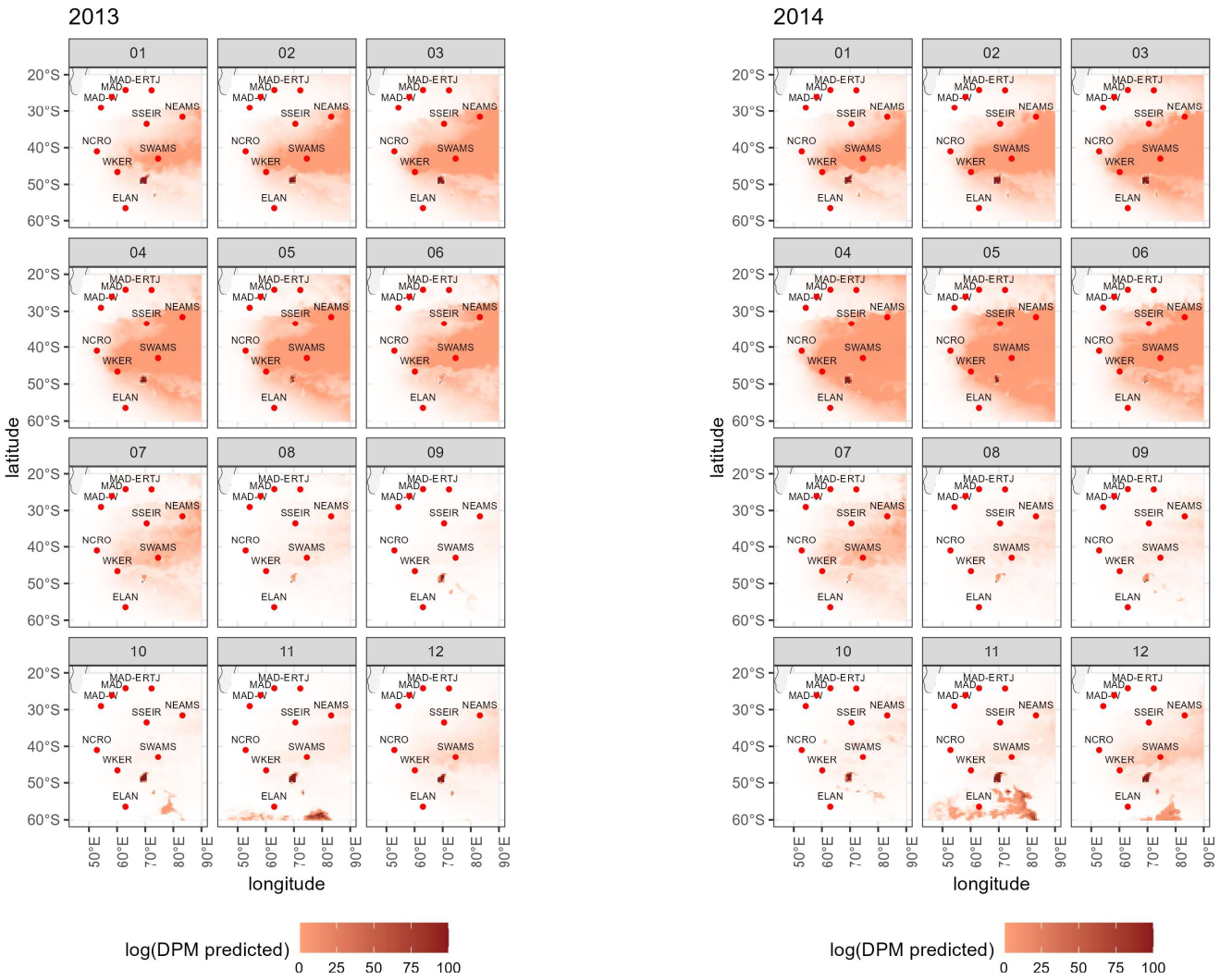

2020
2021
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15
20
25

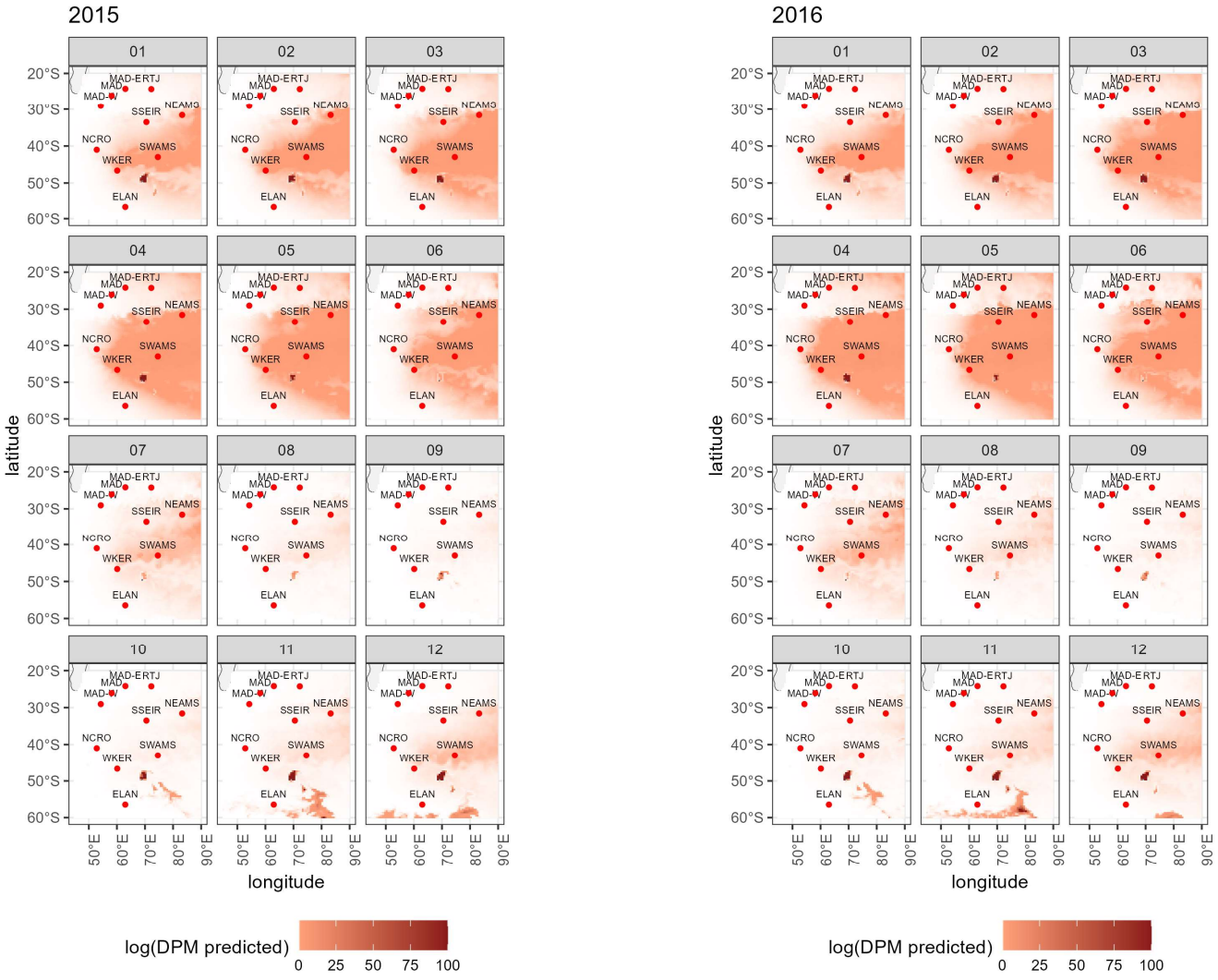

2022
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
5
10
15

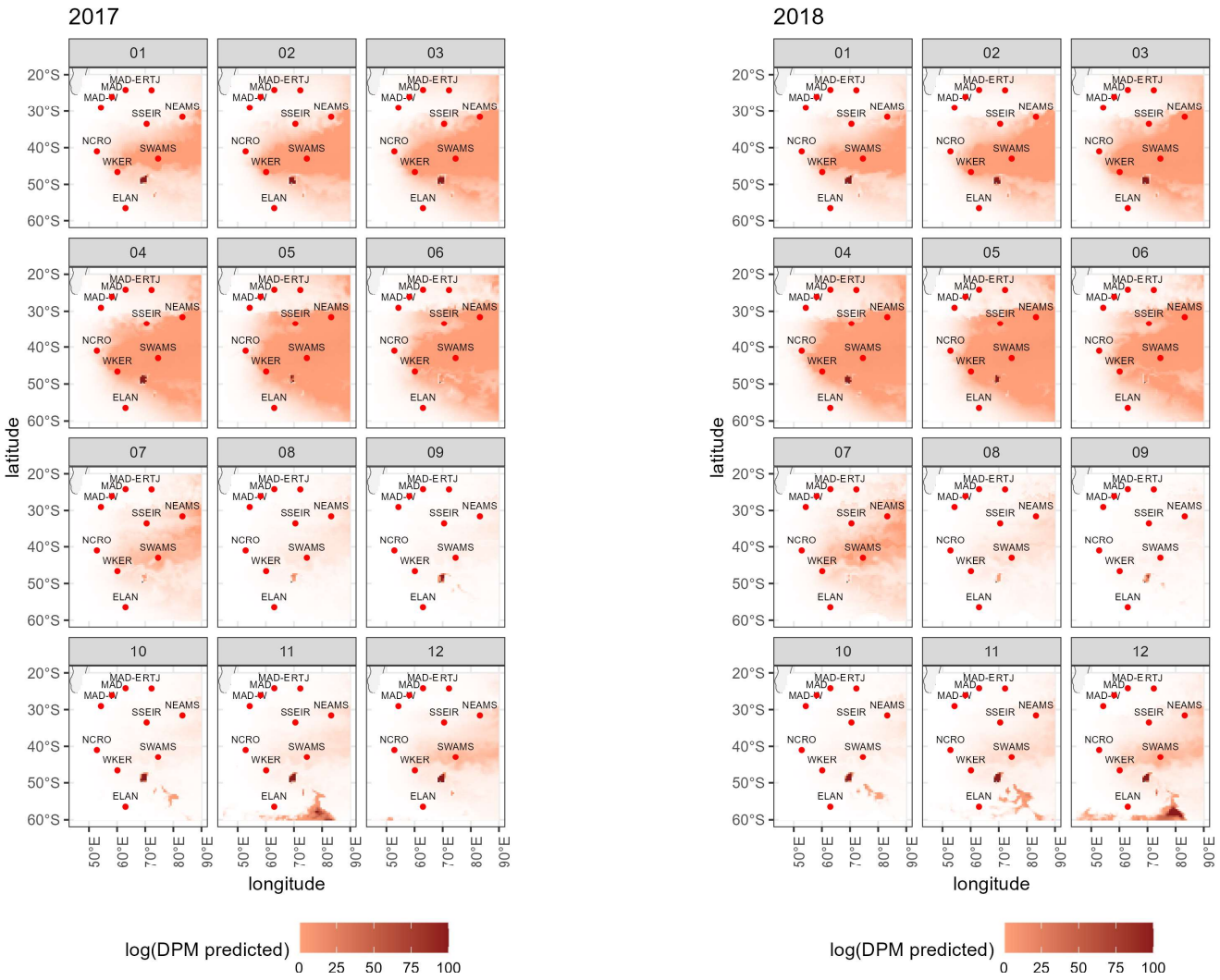

2010
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100

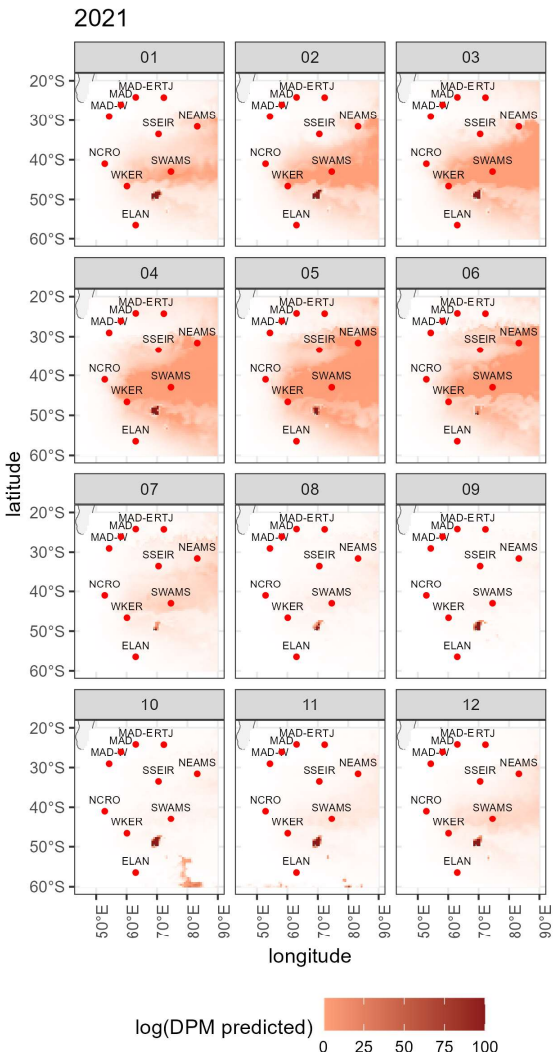

2011
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100

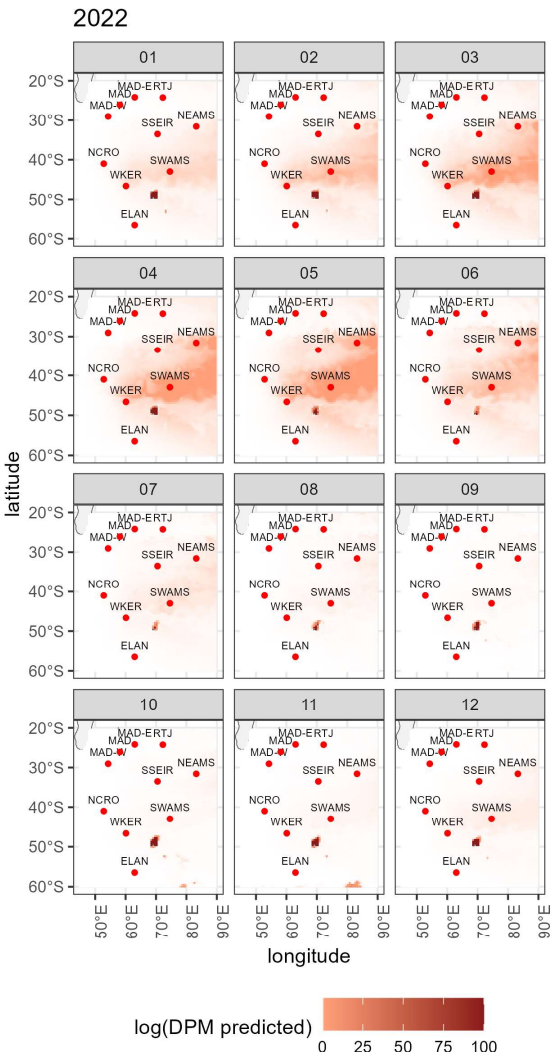

2012
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100

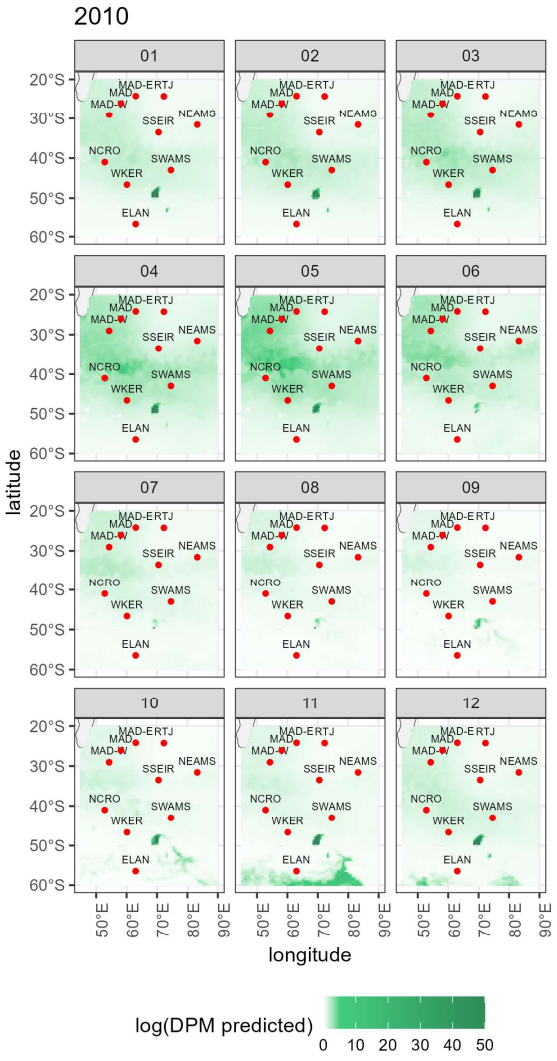

2013
2014
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100

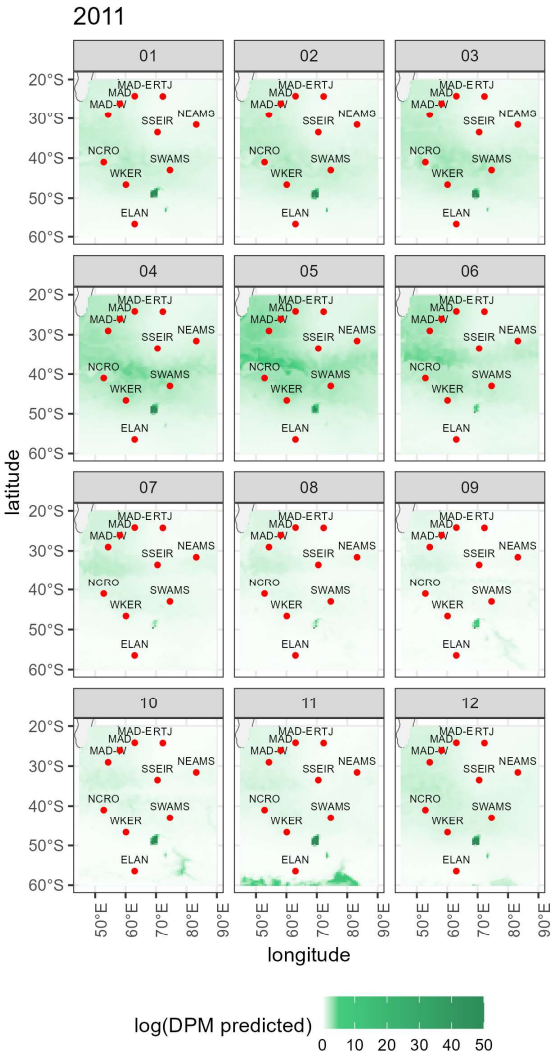

2015
2016
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100

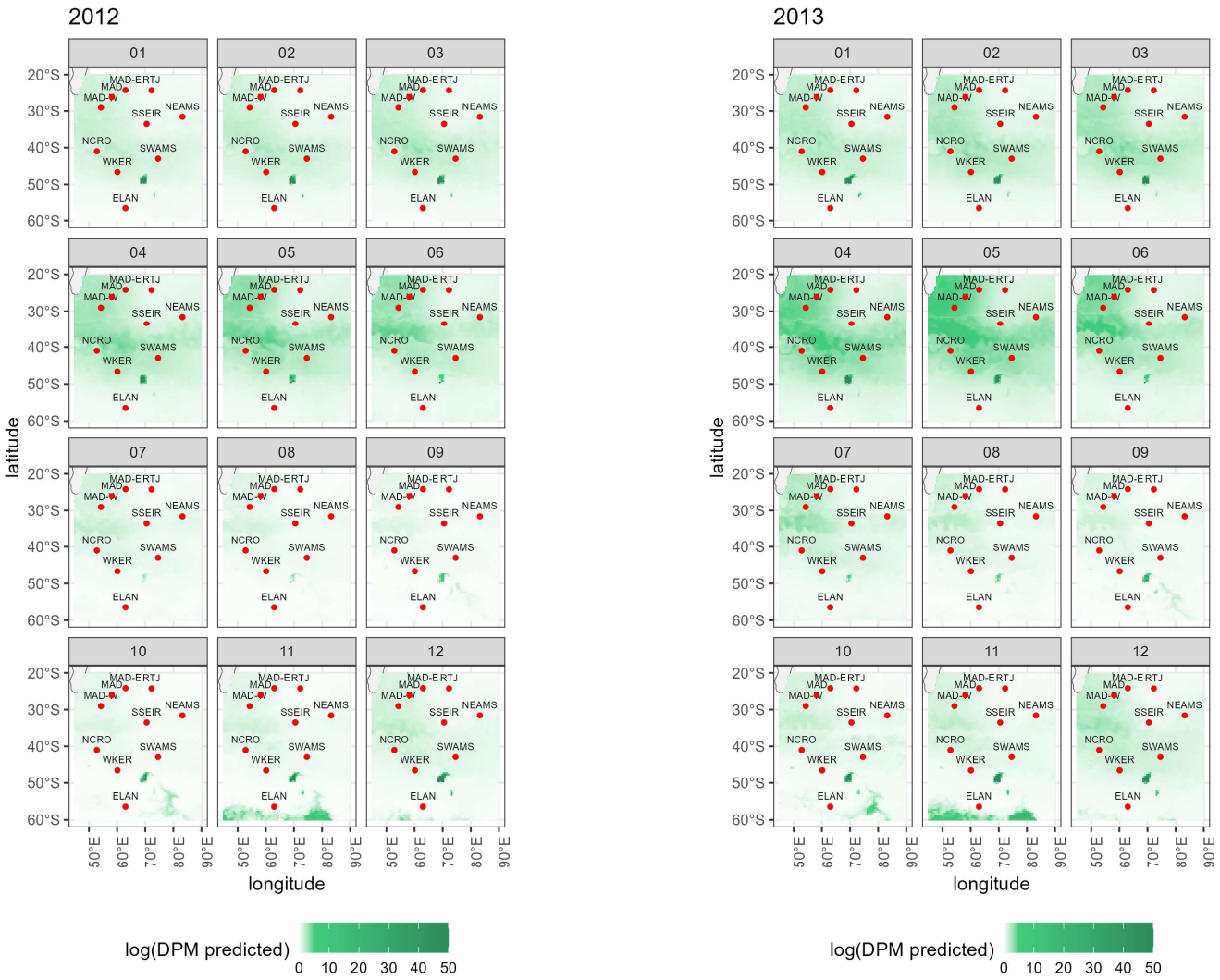

2017
2018
2019
2020
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100


2021
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100


2022
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
25
50
75
100


2010
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
10
20
30
40
50


2011
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
10
20
30
40
50


2012
2013
2014
2015
01
02
03
04
05
06
07
08
09
10
11
12
MAD-ERTJ
MAD
MAD-W
SSEIR
NEAMS
NCRO
SWAMS
WKER
ELAN
20°S
30°S
40°S
50°S
60°S
50°E
60°E
70°E
80°E
90°E
latitude
longitude
log(DPM predicted)
0
10
20
30
40
50

Annexe C

Cette annexe correspond à l'**Étude 2** du Chapitre 2 : *Ecological habitat characterization of baleen whales east of Mayotte : linking seasonal acoustic detections with environmental drivers*, actuellement soumise et en cours d'évaluation dans *Global Ecology and Conservation*.

APPENDIX A: SUMMARY OF MAYOBS DEPLOYMENTS

TABLE S1: Summary of MAYOBS seismological station deployments from 2020 to 2023. Each phase (H–P) includes the station name, deployment period (Start–End), station position (Latitude and Longitude), and seafloor depth (Elevation, in meters). All stations recorded at a sampling frequency (FS) of 125 Hz except phase I (62 Hz).

Phase	Station	Start	End	Latitude	Longitude	Elevation (m)
H	MOA	2020-10-15	2021-01-26	-12.9567	45.8192	-3514
H	MOB	2020-10-14	2021-01-25	-12.8415	45.6230	-3085
H	MOC	2020-10-12	2021-01-25	-12.7842	45.4110	-1722
H	MOD	2020-10-14	2021-01-25	-12.6707	45.5928	-3250
H	MON	2020-10-14	2021-01-25	-12.5958	45.4090	-2017
H	MOO	2020-10-07	2020-10-20	-12.9071	45.2981	-1260
H	MOS	2020-10-16	2021-01-26	-13.0248	45.4560	-2519
I	MOA	2021-02-05	2021-04-08	-12.9622	45.8196	-3524
I	MOB	2021-02-06	2021-04-07	-12.8333	45.6167	-3012
I	MOC	2021-02-05	2021-04-09	-12.7844	45.4114	-1720
I	MOD	2021-02-05	2021-04-11	-12.6709	45.5918	-3255
I	MON	2021-02-05	2021-04-11	-12.5957	45.4091	-2019
I	MOS	2021-02-05	2021-04-10	-13.0204	45.4583	-2531
J	MOA	2021-04-11	2021-09-17	-12.9621	45.8201	-3522
J	MOB	2021-04-12	2021-09-18	-12.8419	45.6217	-3108
J	MOC	2021-04-12	2021-09-19	-12.7839	45.4118	-1667
J	MOD	2021-04-12	2021-09-20	-12.6710	45.5929	-3202
J	MON	2021-04-13	2021-09-21	-12.5961	45.4094	-1962
J	MOS	2021-04-12	2021-09-22	-13.0209	45.4592	-2531
K	MOA	2021-09-29	2022-03-08	-12.9619	45.8206	-3525
K	MOB	2021-09-27	2022-03-08	-12.8414	45.6244	-3061
K	MOC	2021-09-26	2022-03-07	-12.7838	45.4112	-1723
K	MOD	2021-09-21	2022-03-09	-12.6709	45.5928	-3266
K	MON	2021-09-21	2022-03-07	-12.5954	45.4093	-2020
K	MOS	2021-09-22	2022-03-08	-13.0199	45.4599	-2534
L	MOA	2022-03-09	2022-07-15	-12.9626	45.8199	-3384
L	MOB	2022-03-10	2022-07-17	-12.8395	45.6254	-3115
L	MOC	2022-03-09	2022-07-11	-12.7822	45.4117	-1678
L	MOD	2022-03-09	2022-07-10	-12.6694	45.5924	-3190
L	MON	2022-03-10	2022-07-09	-12.5950	45.4082	-2000
L	MOS	2022-03-09	2022-07-13	-13.0239	45.4569	-2503
M	MOA	2022-07-16	2023-03-30	-12.9608	45.8210	-3524
M	MOB	2022-07-20	2023-03-30	-12.8421	45.6249	-3124
M	MOC	2022-07-14	2023-03-29	-12.7835	45.4114	-1732
M	MOD	2022-07-21	2023-03-30	-12.6709	45.5931	-3264
M	MON	2022-07-22	2023-03-29	-12.5953	45.4080	-2017
M	MOS	2022-07-14	2023-03-20	-13.0196	45.4601	-2549
M	MOT	2022-07-22	2023-09-15	-12.7620	45.5777	-2873
M	MOU	2022-07-20	2023-09-15	-12.9392	45.5802	-3038
P	MOA	2023-03-31	2023-09-14	-12.9624	45.8209	-3487
P	MOD	2023-03-30	2023-09-19	-12.6681	45.5922	-3463
P	MON	2023-04-01	2023-09-18	-12.5905	45.4066	-2010
P	MOS	2023-03-31	2023-09-16	-13.0167	45.4614	-2594

APPENDIX B: VARIATIONS IN ENVIRONMENTAL FACTORS

The temporal series of the main environmental variables throughout the period (2020–2024) is depicted in Fig.S1. The SST demonstrates a persistent seasonal cycle, with minimum values around 25 °C occurring in July-August and maximum values near 30 °C in February-March annually. Overlaying this seasonal pattern, interannual anomalies were identified, which are characterized by positive deviations reaching +1 °C in the early months of 2021 and 2023, and negative anomalies of approximately -0.5 °C in mid-2022 (Fig.S1d). In particular, the positive anomaly in early 2023 was stronger and occurred earlier than in previous years, probably influenced by the onset of the global El Niño event detected around February-March 2023[53]. This event may have driven an accelerated warming in the region, shifting sea surface conditions beyond typical interannual variability. Surface chlorophyll-a concentrations generally remain low throughout the year (< 0.1 mg $\cdot$ m^{-3}), indicative of oligotrophic conditions. However, distinct ephemeral peaks exceeding 0.3 mg $\cdot$ m^{-3} were documented in March-April 2021 and August 2022 (Fig. S1a). These phenomena frequently coincide with decreases in SST, implying transient enrichment processes. The SWH exhibits a pronounced seasonal cycle, with elevated values during the austral winter (June to September), often exceeding 2.5 m in 2022 and 2023, and reduced values around 1.5 m during the austral summer (December to April) (Fig. S1b). No long-term trend is evident, although variability in interannual amplitude is observed.

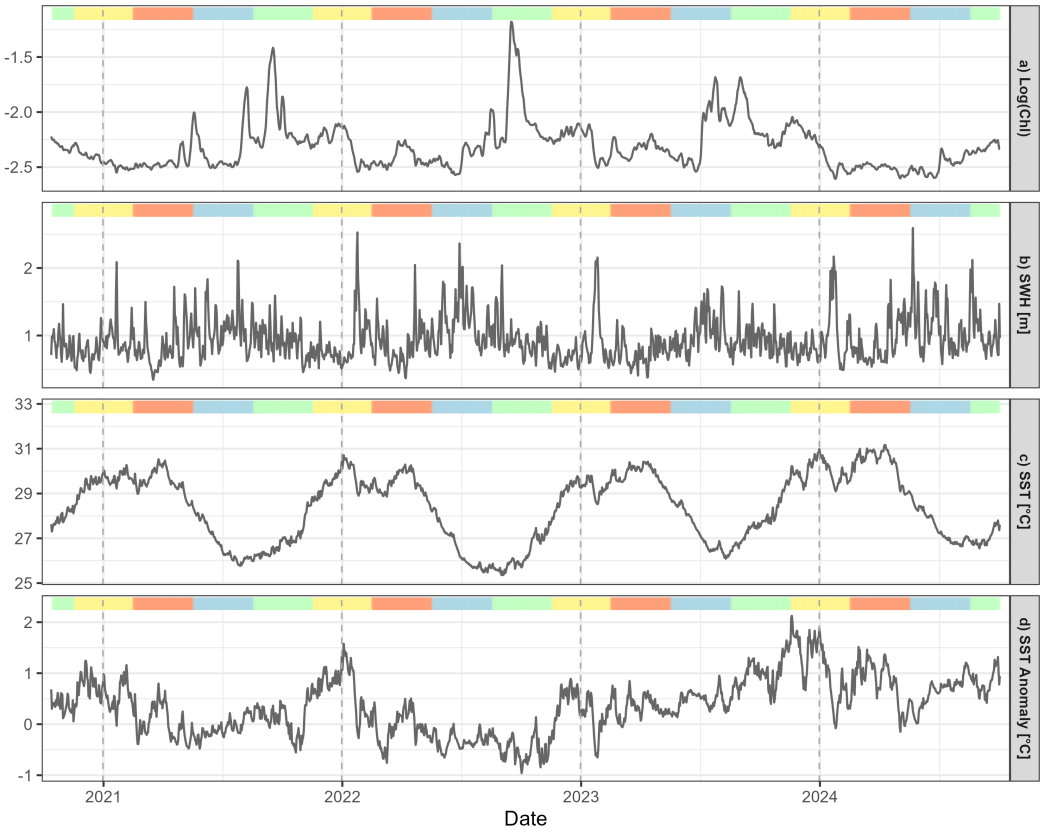


FIG. S1: Temporal series of daily values of four environmental covariates used in the GAMs. The variables shown are a) Log(Chla), b) SWH [m], c) SST [°C], and d) SST Anomaly [°C]. Grey lines represent daily measurements; colored horizontal segments indicate the seasons defined by calendar months: summer (December–February, yellow), autumn (March–May, orange), winter (June–August, blue), and spring (September–November, green). Vertical dashed lines mark January 1st of each year.

APPENDIX C: CORRELATION BETWEEN ENVIRONMENTAL COVARIATES

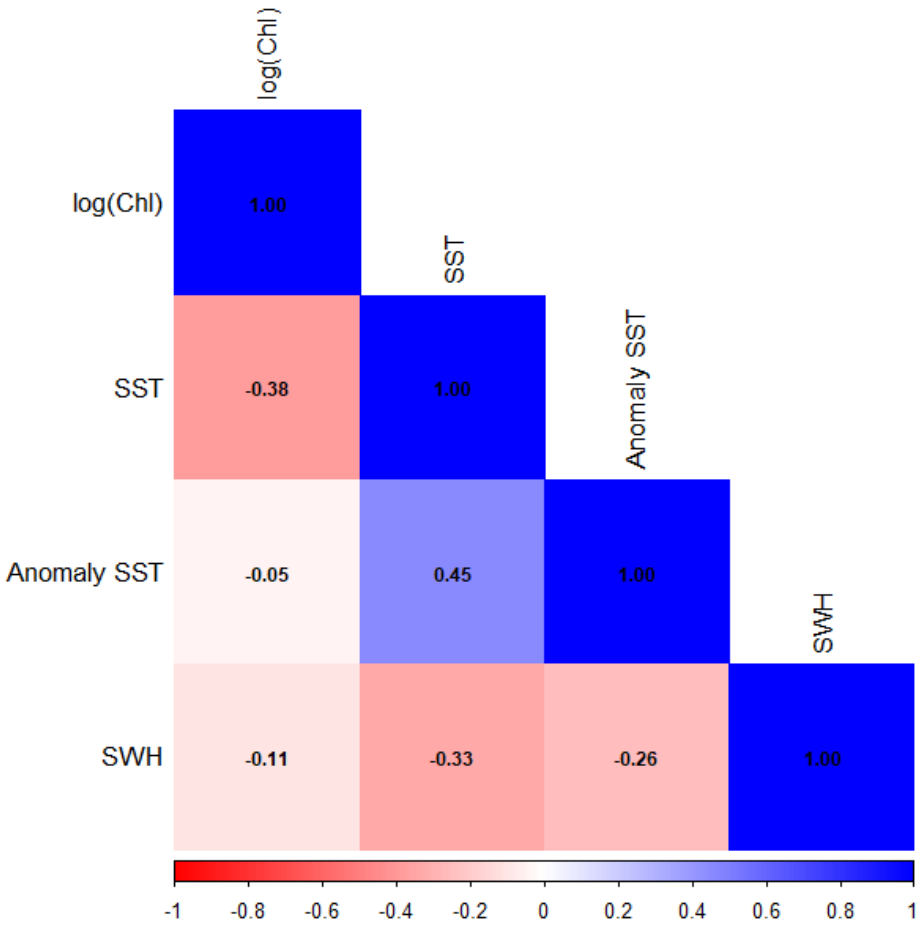


FIG. S2: Correlation matrix of explanatory variables used in the GAM models, calculated using pairwise Pearson correlation coefficients. The color scale indicates the direction and magnitude of the correlations, ranging from strong negative (blue) to strong positive (red).

Annexe D

Cette annexe reprend le rapport complet du stage de Master 2 de Salomé Martin-Marin (International Master of Bioacoustics, MoBi, université de Saint-Étienne, promotion 2023-2024) présenté dans le Chapitre 2.

Empirical project of The International Master of Bioacoustics

https://www.masterofbioacoustics.com/

Research:

Cite this article:

Received: 15 August 2024
Accepted:

Subject Area:

Ecology, Biological Modelling, Acoustics, Cetaceans

Keywords:
Monitoring - Occurrence - Modeling – Zooplankton
Passive acoustic – Active acoustic - Migration

Author for correspondence:

Salomé Martin-Marin: salome_MM@hotmail.fr

Animal behaviour

Combining passive and active acoustic approaches to study predator-prey interactions of blue whales in the southern Indian Ocean

Martin-Marin Salomé[1,2], Mathilde Michel[2], Lloyd Izard[3], Cédric Cotté[3], Flore Samaran[2]

[1]International Master of Bioacoustics, Université Jean-Monnet, Saint-Etienne, France, [2]CNRS Lab-STICC, ENSTA Bretagne, Brest, France, [3]LOCEAN-IPSL, Sorbonne Université, CNRS, IRD, MNHN, Paris, France

Abstract:

This study investigates the relationship between the seasonal acoustic presence of Antarctic blue whales (*Balaenoptera musculus intermedia*) and zooplankton biomass in the southern Indian Ocean. Using passive acoustic data from the OHA-SIS-BIO network and active acoustic data from the THEMISTO program, this research explores the hypothesis that Antarctic blue whales' presence outside their traditional summer feeding grounds is driven by prey availability in subtropical and sub-Antarctic regions.

Acoustic detections of Antarctic blue whale Z-calls were analyzed in relation to zooplankton acoustic densities modeled using echosounders data at 120 kHz, the frequency characteristic of their prey (*Euphausiid*). Results indicate significant spatial and temporal variability in blue whale presence in the low and mid latitudes, with a higher detection frequency in sub-Antarctic regions. However, the study found only a slight negative correlation between macrozooplankton and micronekton biomass and whale acoustic activity, suggesting that factors other than prey abundance may influence whale occurrence or the acoustic dataset might not reflect with enough precision the prey distribution for this species. This work contributes to understanding the complex migratory patterns of Antarctic blue whales and their foraging strategies in response to environmental variability, highlighting the nuanced nature of baleen whale migrations and their potential partial or differential migration behavior.

1. Introduction

Blue whales are good candidates for passive acoustic monitoring due to their calls that are repetitive, long (over 15 seconds), loud (over 180 dB ref 1 µPa at 1m), low-frequency (20-100 Hz) and well stereotyped (Cummings & Thompson, 1971). In the Indian Ocean, most knowledge of blue whales distribution comes from this survey method (Dréo et al., 2019; Leroy et al., 2016, 2021; Samaran et al., 2010, 2013; Stafford et al., 2010; Torterotot et al., 2020, 2023). Their calls are known to vary between regions and have been used to distinguish different blue whale subspecies and acoustic populations (Torterotot et al., 2020). The southern part of the Indian Ocean is known for having a very high acoustic diversity of blue whales (McDonald et al., 2006). In particular, one subspecies is seasonally present in this region: the Antarctic blue whale (*Balaenoptera musculus intermedia* (Linnaeus, 1758) displaying calls with clearly distinct characteristics (Samaran et al., 2010; Stafford et al., 2010). The Antarctic blue whale stands out as the most well-known among various acoustic populations. The duration of their vocalization is approximately 26 s and comprises three tonal units that follow one another without silence (Ljungblad et al., 1998). These vocalizations are known as Z-calls because of their 'Z' shape on a spectrogram. Z-calls are produced repeatedly in long periods of several tens of minutes, making it easy to distinguish the presence of the species in the underwater soundscape. Blue whale songs are thought to be produced only by males for reproductive purposes (based on a small sample of known callers). The reproductive purpose is not certain, however, as these calls are produced year-round (Lewis et al., 2018; McDonald et al., 2001; Oleson et al., 2007). Blue whale populations feed predominantly on dense patches of krill from the *Euphausiidae* family but switch foraging strategies in response to variation in prey density and depth to maximize energetic efficiency (Hazen et al., 2015); for instance they can also locate and feed on dense swarms of other types of prey, like sergestid shrimp (de Vos et al., 2018). To feed, blue whales can dive as deep as 315 m, with average dives of 201±52 m (Goldbogen et al., 2011). McQuinn et al. (2016) assuming that, if blue whales forage by targeting the best krill aggregations within their territory, the abundance, distribution, and local density of krill are likely among the main factors determining the distribution of foraging blue whales. Antarctic blue whales are mostly found in the polar zone close to the Antarctic during the austral summer, where they feed, and then spread out over a wider range in the mid and low latitudes of the southern hemisphere during the austral winter to breed (Branch et al., 2007). Acoustic studies reveal, however, that the paradigm of classic seasonal patterns of large migratory whales appears to be more complex for this species. Indeed, these studies have shown that individuals are present all year round in the Southern Ocean (Širović et al., 2009), but also all year round at mid (Samaran et al., 2010) and low latitudes (Branch et al., 2007). It's worth pointing out, however, that in each case, few acoustic detections are made, but enough nonetheless to attest to the presence of a few individuals skipping migration by remaining in either feeding or breeding areas all year round. However, given the high energy requirements of baleen whales, a migration jump in favor of overwintering in an area of low krill productivity is surprising, and we can therefore assume that individuals manage to find nutritional resources even at low latitudes. In this study we hypothesize that the presence of Antarctic blue whales outside their austral-summer (January-February-March) feeding grounds in the Antarctic polar zone, is motivated by the availability of prey in sub-tropical and sub-Antarctic areas. To verify this hypothesis, we use densities of intermediate trophic level organisms like macrozooplankton and micronekton, crucial prey for whales, measured with echosounders. This study establishes the first link between two

significant datasets in the Indian Ocean: OHA-SIS-BIO, a long-term passive acoustic program for monitoring the seasonal presence patterns of blue whales using a large hydrophone network, and THEMISTO, an active acoustic program for estimating mid-trophic levels distribution, abundance and biomass using echosounders data. This study combines these two datasets of passive and active acoustic data to better understand the relationships between the dynamics and densities of prey, and the seasonal acoustic activity patterns of Antarctic blue whales in their migration path.

2. Material & methods

2.a) Passive acoustic monitoring

The passive acoustic data used in this study come from the hydrophone network of the Hydro-Acoustic Observatory of Seismicity and Biodiversity (OHA-SIS-BIO), as detailed by Royer Jean-Yves (2009). This is a long-term hydroacoustic program to monitor the seismic events and vocal activity of large marine mammals in the southern Indian Ocean. A network of autonomous hydrophones (9 moorings, 10 hydrophones (Fig.1) has been deployed and maintained annually between 2010 and 2023 between La Réunion, Crozet, Kerguelen and Amsterdam, during ObsAutral cruises of the oceanographic vessel Marion Dufresne II (Nicolas & Lo Monaco Claire, 2017).

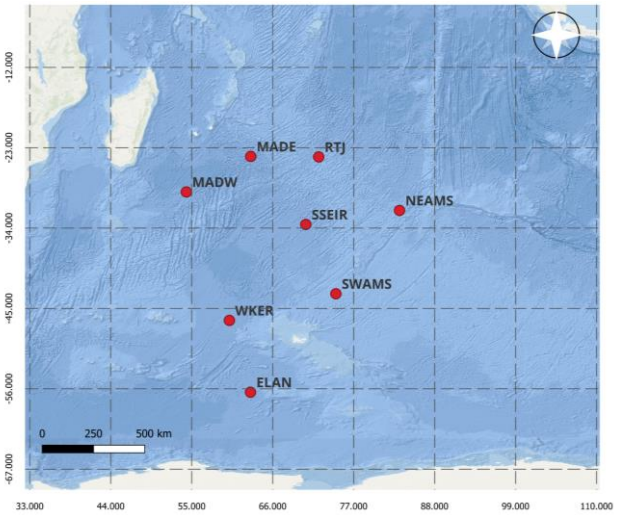


Figure 1. *The OHASISBIO hydrophone network in the Southern Indian Ocean. Each point represents a listening station equipped with a recorder.*

For this study, only acoustic data recorded between January and March 2020 to 2022 on each site were used to evaluate the presence/absence of the species from Z-calls detections. The hydrophones are moored between 1 000 and 1 300 m depth. The continuous recordings are sampled at 240 Hz using a 24-bit analog-to-digital conversion. An automated detection algorithm based on dictionary learning and sparse representation of Antarctic blue whale Z-calls was used to detect each calls (Fig.2, Socheleau & Samaran, 2017).

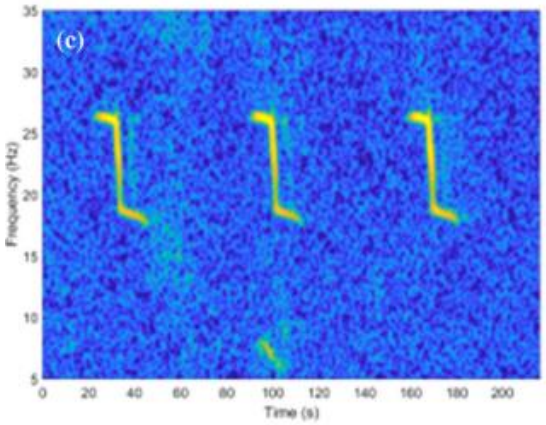


Figure 2. *Typical spectrograms of stereotyped Antarctic blue whale sounds, Z-calls.*

This technique employs extensive dictionaries to capture the variability in calls and uses linear combinations of dictionary entries to represent the complexity of the calls. Without relying on a fixed template, the algorithm builds dictionaries directly from temporal call signals extracted from the data and reconstructs observed signals with sparse combinations of waveform entries. However, the detector can be sensitive to interference and generate false detections that can distort results. Its detection performance is therefore calculated by a human annotator who double-check automatic detections using APLOSE tools[1]. The distance at which blue whale calls can be detected depends on many factors, including ambient noise levels, instrument type, receiver depth and automated detection method. In the region around Crozet Island, the estimated maximum detection range of Antarctic blue whales is 180 km, using a single hydrophone located in line with the sound channel and an automated detector (Samaran et al., 2010). A presence/absence of the species on each site between January and March 2020 to 2022, could be identified from the detections. Each hour of a day with at least one detection was defined as "positive". The weekly positive hours (WPH) metric, corresponding to the sum of all positive hours in a week, was then used for analysis.

2.b) Active acoustic analysis

The THEMISTO program (Toward Hydroacoustics and Ecology of Mid-trophic levels in Indian and SouThern Ocean) aims to study the 3-D distribution of zooplankton and micronekton and its variability in relation to physical and climatic phenomena. This study uses calibrated water column sonar data recorded onboard the oceanographic vessel Marion Dufresne II, during the austral summers of 2021, 2022 and 2023. The vessel is equipped with an EK80 echosounder (SIMRAD, Norway) operating at 18, 38, 70, 120, and 200 kHz frequencies. Our analyses are based on data obtained during stations and focus on 18 to 120 kHz frequencies to explore the distribution of pelagic fauna down to 225 m depth (i.e. approximate depth of blue whale dives (Goldbogen et al., 2011)). This intricate methodology (Izard et al., under review) unfolds in two steps: analyzing in situ acoustic observations and subsequently extending these observations to model backscatter across the entire region using environmental predictors. The complete method is detailed in the supplementary materials (Appendix 1). Firstly, each profile is smoothed using 10 B-spline basis functions. The smoothed profiles are then transformed into a matrix, which undergoes Functional Principal Component Analysis (fPCA). This analysis decomposes the matrix into principal components (PCs) that reveal the primary modes of vertical variability within the dataset. Daytime and nighttime data have been treated independently to take into account the effect of vertical migration of organisms, which can affect density in the 225m depth range (Brierley, 2014). The first four modes capture around 74 % of the variability for daytime data and 92 % for nighttime data. The second stage extends these in situ observations to a broader regional scale using Random Forest models. Each principal component identified in the first stage is modeled independently, with environmental fPCA scores (surface temperature, salinity, chlorophyll a concentration, and dissolved oxygen) serving as predictors of the acoustic fPCA scores. This study has been conducted using environmental parameters available from E.U. Copernicus Marine Service Information (Global Ocean Biogeochemistry Analysis and Forecast: https://doi.org/10.48670/moi-00015 and Global Ocean Physics Analysis and Forecast: https://doi.org/10.48670/moi-00016). This modelling process involves a robust hyperparameter tuning strategy to prevent

[1] https://osmose.ifremer.fr/app/

overfitting and ensure model accuracy. Once the best models are identified, they are used to predict principal components across the entire study area, resulting in a series of maps that illustrate the predicted vertical modes of variability. These predicted components are then used to reconstruct the acoustic profiles by combining the vertical modes weighted by the predicted PCs (e.g Pauthenet et al., 2018). The reconstructed profiles are transformed into the Nautical Area Scattering Coefficient (NASC), which serves as a proxy for integrated biomass. According to the blue whale detection zone, the NASC values obtained were then averaged within a 180 km radius around hydrophones sites. Although predictions were made for each frequency, only 120 kHz, which is characteristic of the euphausiid group as a fluid-like target (Benoit-Bird & Lawson, 2015), was used in this study to relate to acoustic detections of Antarctic blue whales Z-calls. Focusing on predictions at this frequency allows us to better approximate real blue whale prey distribution. For further analysis, in particular for the study of biomass anomalies, day and night predicted data were grouped.

2.c) Statistical analysis

An anomaly study was carried out to observe the inter-site differences of the two density metrics (WPH and NASC). For the first one, the total number of positive hours of detected Z-calls per week for each site was subtracted from the average number of positive hours for all sites to obtain the WPH anomaly. Thus, the WPH anomaly informs about how Antarctic blue whales relatively occurred at each site at the scale of the western Indian Ocean. If the anomaly is positive, then the site seems more frequented by Antarctic blue whales than the rest of the ocean; if it is negative, then the site is less frequently visited. Similarly, the biomass anomaly was obtained by subtracting the average of one site (NASC within 180 km) from the average biomass of all sites for a given week. The biomass anomaly was obtained from averaged day and night data. The biomass anomaly allows us to see how the NASC value within a 180 km radius of a site differs from the average biomass of other sites in the Indian Ocean. If the anomaly is positive, then the site has a higher zooplankton biomass than the rest of the sites; if it is negative, then the zooplankton biomass is lower at this site than at the other sites. Kruskal-Wallis tests and Dunn's post-hoc tests were performed to study the variability of blue whale detection data between the different sites and the three years. Wilcoxon tests were also performed to investigate the difference in WPH between day and night and between subtropical sites (between 23.5° and 40° latitude, i.e., MADE, MADW, RTJ, NEAMS, and SSEIR) or sub-Antarctic sites (above 40° latitude, i.e., ELAN, SWAMS, and WKER). Finally, Kendall tests were performed to study the linear relationship between zooplankton biomass anomaly and weakly positive hours anomaly. All statistical analyses were performed with R version 4.3.0. (2023-04-21 ucrt).

3. Results

3.a) Zooplankton biomass modeling

As specified in the method, for a more specific approach, only predictions for the 120 kHz frequency have been interpreted in this study. The night-time acoustic zooplankton density for the entire water column of interest in the study (25-225m) was greater than in daytime. Predicted biomass at 120 kHz ranged from 20 to 609m^2/nm^2 (NASC values) at night and from 6 to 135 m^2/nm^2 during the day for the entire area (30°W-85°W; 20°S-60°S). Biomass values at the sites during the day were a maximum of 84 m^2/nm^2 at the RTJ site and a minimum of 10 m^2/nm^2 at the ELAN site. At night, the maximum and minimum biomass values were 504 m^2/nm^2 for RTJ and 28 m^2/nm^2 for ELAN (Fig. 3 and Fig.4). We can also observe a latitudinal gradient in biomass, with higher

values in tropical zones and lower values in sub-Antarctic zones. Examples of a map of average monthly biomass during daytime and nighttime for the years 2020, 2021 and 2022 are presented in the appendices 2, 3 and 4 respectively.



Figure 3. *Daytime NASC biomass estimation (y-axis, m²/nm²) per weeks (x-axis) for each year (2020-2022) and sites (from north to south).*



Figure 4. *At night NASC biomass estimation (m²/nm²) per week (x-axis) for each year (2020-2022) and sites (from north to south).*

3.b) Acoustic detections of Antarctic blue whales

There was no significant difference in the number of detection-positive hours per week between day and night (Fig.5, Wilcoxon test, p-value > 0.05). Thus, for the subsequent analyses, day/night data were grouped.

.

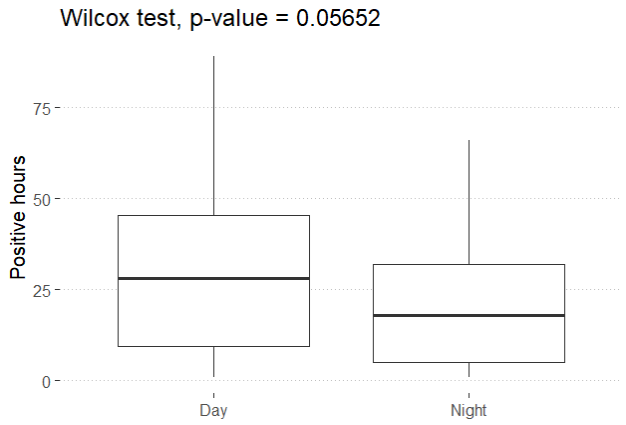


Figure 5. *Weekly positive hours of Antarctic blue whales Z-calls by day and night, Wilcoxon p-value indicates no significant difference.*

The distribution of WPH appears heterogeneous between years and sites, with more WPH in 2020 for the NEAMS site and for the SWAMS site in 2021 (Fig.6). Indeed, the WPH was significantly higher in 2021 (mean = 11.4 ± 17.5 h per week) than in the other two years (10.3 ± 29.7 h per week and 1.89 ± 4.29 h per week for 2020 and 2022, respectively) (Fig.7, Dunn test, p-values <0.05).

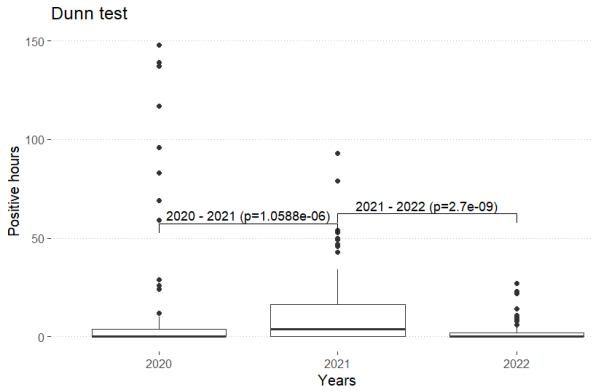


Figure 7. *Weekly positive hours of Antarctic blue whales Z-calls by year, brackets indicate significant differences according to the Dunn test.*

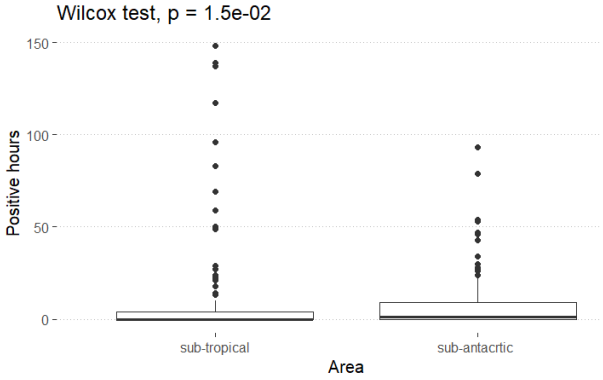


Figure 8. *Weekly positive hours of Antarctic blue whales by geographical area Z-calls, Wilcoxon p-value indicates a significant difference.*

The number of positive hours also varied between sites, with significantly more detections at NEAMS, ELAN, and SWAMS (Dunn test, p-values <0.05). The number of positive hours was further analyzed based on the geographical area in which the site was located (subtropical or sub-Antarctic). The number of detection-positive hours in the sub-Antarctic area (mean = 8.51 ± 16.2 h per week) was significantly higher (Fig.9, Wilcoxon test, p-value < 0.05) than in the subtropical area (mean = 7.47 ± 22.70 h per week). The quantification of this difference in positive hours between sites was studied using weekly anomaly calculations.

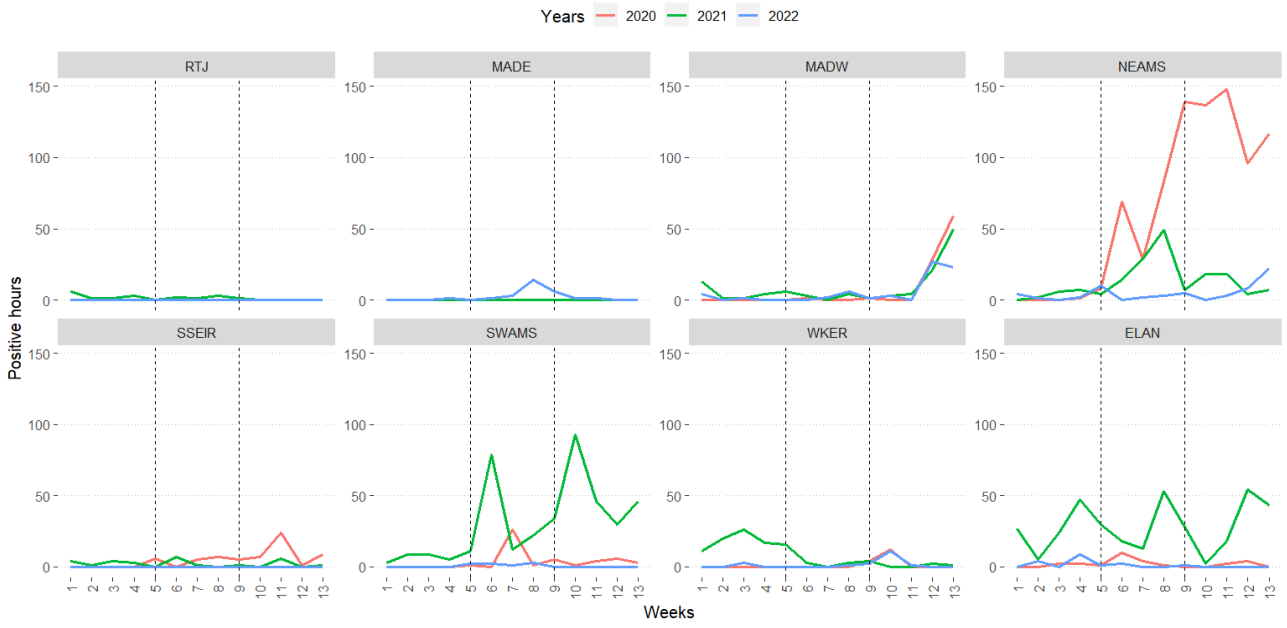


Figure 6. *Number of detection-positive hours for Antarctic blue whales Z-calls per week (x-axis) for each year and site (from north to south.*

3.c) Spatial anomaly

The results show a stronger positive WPH anomaly for NEAMS in 2020, indicating that Antarctic blue whale occurrence is more important in this site than the rest of the sites during weeks 5 to 13 (Fig.9). Positive anomaly peaks are also observed to a lesser extent in 2021 for the sites SWAMS (weeks 6 and 10) and ELAN (weeks 4, 8, and 12). The biomass study reveals a latitudinal gradient with stronger negative anomalies towards the south, indicating that the southernmost sites have lower zooplankton biomasses than other sites (Fig.9). Linear regression and Kendall's test reveal a significant linear correlation between the two variables (Appendix 5, p value < 0.05), indicating that there is a statistically significant correlation between the biomass and the WPH anomalies. As the value of Kendall's ϱ statistic is less than 0, the correlation between these two variables is negative. Meaning that as the biomass hours anomaly increases, the positive hours anomaly tends to decrease. However, as there is significant inter-annual variability in the WPH data, an analysis by year was also carried out (Appendix 5) and showed that this significant negative correlation (p-value < 0.05, ϱ < 0) is only observed for the year 2021. Tests carried out to distinguish between subtropical and subantarctic zones (Fig.10) showed no statistically significant correlation between the two anomalies as a function of geographical zone (Spearman test, p-value > 0.05). NASC anomalies were then grouped into qualitative categories to study the proportions of detection-positive hours as a function of biomass anomaly intensity (Fig.11). Most positive detection hours were obtained during neutral biomass anomalies, i.e. when the biomass at the site was little or not at all different from other sites in the Indian Ocean. Thus, although a negative correlation is observed for some sites in 2021, no clear trend seems to emerge on the influence of the biomass anomaly on positive detection hours of Antarctic blue whale Z-calls on a more global scale.

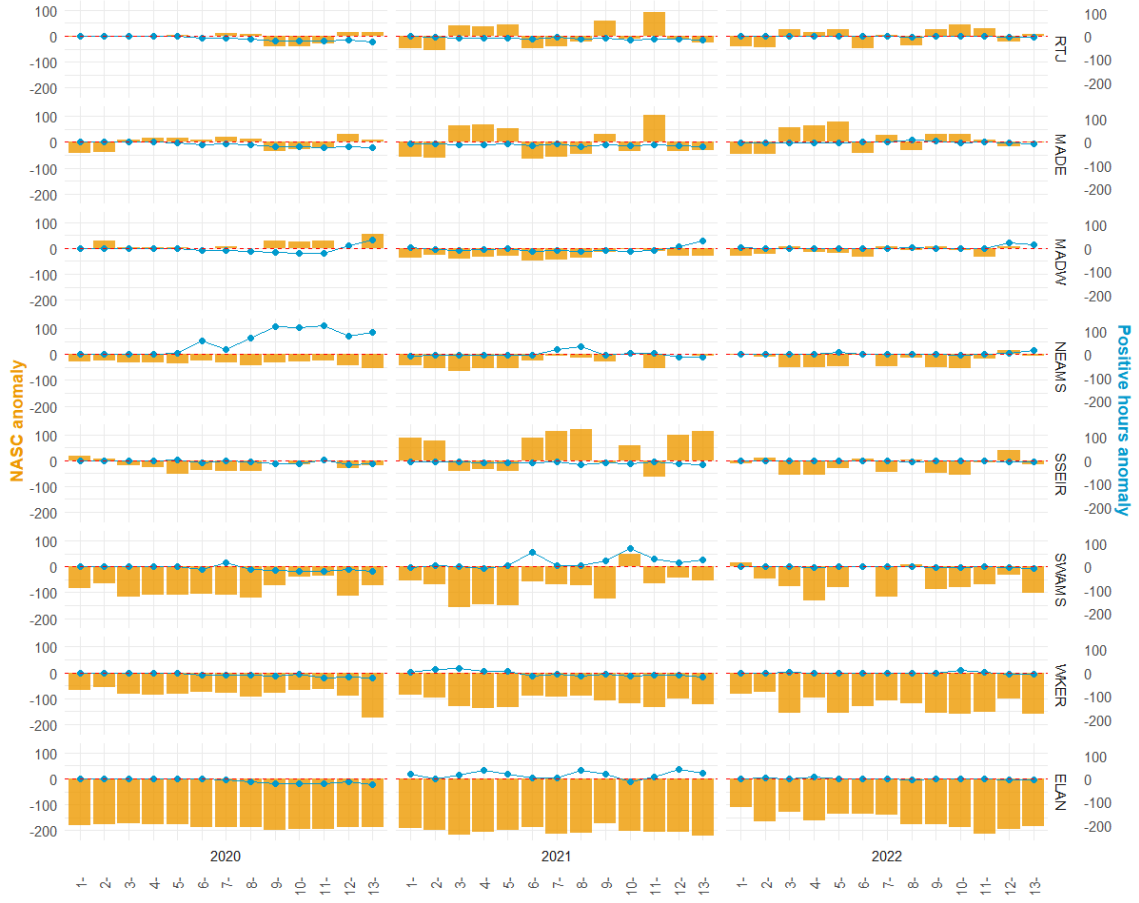


Figure 9. *NASC biomass anomaly (m^2/nm^2, left axis) and positive hours anomaly (right axis) by week (bottom axis), year and site (from north to south).*

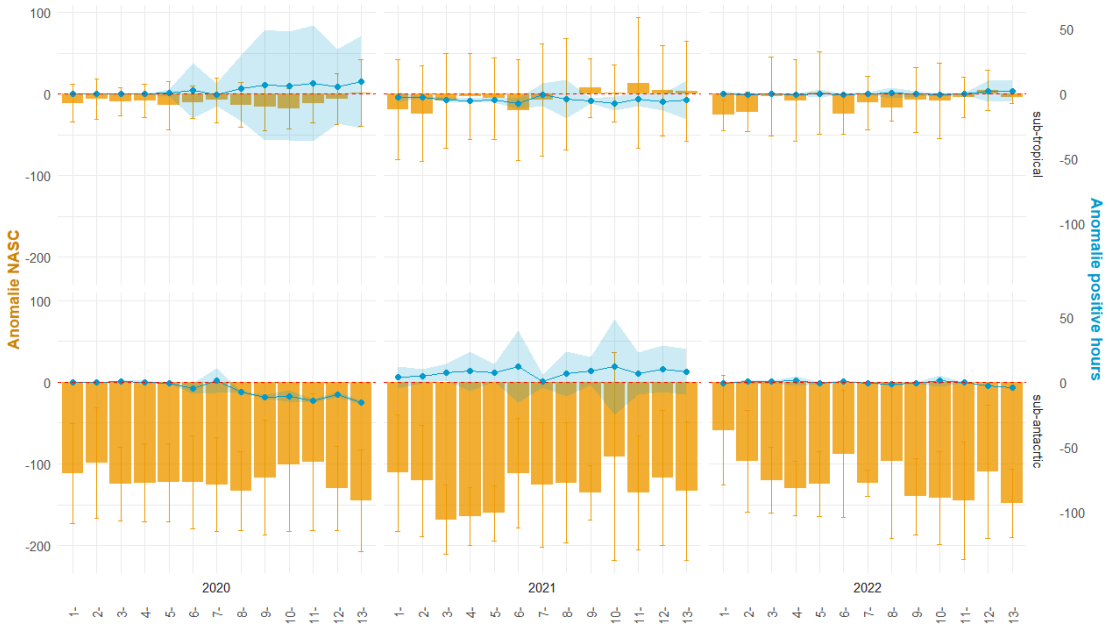
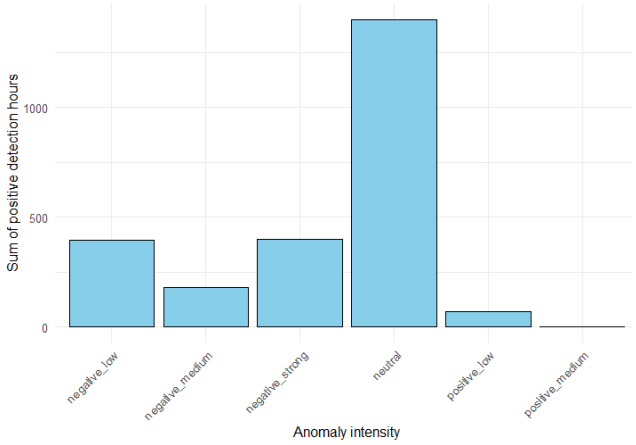


Figure 10. *NASC biomass anomaly (left axis) and positive hours anomaly (right axis) by week (bottom axis) per year and area (from north to south).*



NASC value	Category
-222 to -162	Negative strong
-162 to -108	Negative medium
-108 to -54	Negative low
-54 to 54	Neutral
54 to 108	Positive low
108 to 162	Positive medium
162 to 222	Positive strong

Figure 11 *Sum of positive detection hours as a function of biomass anomaly intensity (top graph) and biomass anomaly intensity categories (bottom table).*

4.. Discussion

4.a) Zooplankton biomass modeling

The NASC (Nautical Area Scattering Coefficient) values observed in our study on macrozooplankton in the South-Western Indian Ocean appear consistent with those reported in other studies, despite differences in geographic areas, acquisition frequencies, and biological targets. Indeed, although our study used a frequency of 120 kHz and focused on a depth range of 25 to 225 m, the biomass measurements (6 to 135 m^2/nm^2 during the day and 20 to 609 m^2/nm^2 at night) are comparable to those observed by Béhagle et al. (2016), who reported NASC values for micronekton (at a frequency of 38 kHz, from 10 to 200 m) in the same region, ranging from 83 ± 118 to 249 ± 116 m^2/nm^2 during the day and 170 ± 88 to 827 ± 324 m^2/nm^2 at night. Although the frequencies and biological targets differ, these similar magnitudes reinforce the validity of our observations, particularly regarding the higher nighttime NASC values, a phenomenon frequently observed in the daily vertical migration of macrozooplankton (Brierley, 2014). These results suggest consistent patterns of distribution and biomass in the region, regardless of the specific parameters of the study. Furthermore, the results of Kang et al. (2024), who explored macrozooplankton at latitudes between 0° and 15°S in the same area, using a method of backscatter intensity difference between frequencies of 38 and 120 kHz, showed NASC values ranging from 0.6 to 300 m^2/nm^2 for similar depths (0 to 150 m). This range of values, although recorded in a more restricted geographic area, also corroborates our observations. Thus, despite variations in methodologies, our results seem to align with previous observations. At the same time, our results show a latitudinal gradient, with higher NASC values in the sub-tropics than in the sub-Antarctic, indicating a heterogeneous distribution of pelagic backscatter in the southwest Indian Ocean, concordant with the literature (Boersch-Supan et al., 2017; Izard, 2023). Many physical and biological factors can explain this heterogeneity, including temperature, mixing, currents, predator density and food availability (Boersch-Supan et al., 2017).

4.b) Influence of zooplankton biomass on the summer presence of blue whales

Antarctic blue whales, known for their seasonal migrations between low-latitude breeding areas and high-latitude feeding areas, sometimes exhibit atypical migratory behaviors. Indeed, some of these whales do not migrate to the

Southern Ocean close to Antarctica during the austral summer but remain year-round in subtropical regions (Branch et al., 2007; Leroy et al., 2016; Samaran et al., 2010; Torterotot et al., 2020). Based on the hypothesis that, even though food is more abundant in the Southern Ocean, some northern areas may offer sufficient food resources for these whales, this study attempted to identify the cause of the absence of migration in certain individuals by studying the relationship between acoustic detections of Antarctic blue whales and prey abundance in subtropical and sub-Antarctic zones. This hypothesis was already suggested by Thomisch et al. (2019), who proposed that the year-round presence of this species off Namibia (Sub-tropical zone of the Atlantic Ocean) would be strongly influenced by prey availability and the associated high turbulence eddy regime.

In this study, a slight negative trend seems to emerge for the year 2021. Thus, based on the detection of Z-calls for this year, Antarctic blue whale occurrence seems more important at sites with lower plankton biomass. To date, the vocal feeding behavior of Antarctic blue whales is unknown, only the vocal behavior of the Northeast Pacific blue whale has been studied in the context of feeding behavior. In these studies, blue whales' songs are not produced during deep feeding dives (Lewis et al., 2018; Oleson et al., 2007). We can therefore assume that a higher number of calls could indicate a lower feeding dive rate, perhaps due to a lower plankton biomass. It should be noted, however, that in our case, most of the positive detection hours were recorded at sites with neutral anomalies, i.e. neither more nor less productive than the average of the other sites. Therefore, the results of our study do not allow us to visualize a clear correlation between prey abundance and the acoustic presence of Antarctic blue whales.

The sex of Antarctic blue whale's individuals producing Z-calls has not, to date, been investigated. If we assume, by analogy with other blue whales subspecies, that only males produce these sounds (Lewis et al., 2018; McDonald et al., 2001; Oleson et al., 2007), then we can assume that the presence of Z-calls most likely only indicates the presence of reproductively active male. However, blue whales can also produce non-stereotyped vocalizations, D-calls, emitted by both males and females (Lewis et al., 2018; Oleson et al., 2007) and documented across all blue whale populations around the world (Ljungblad et al., 1998; McDonald et al., 2001, 2006; Mellinger & Clark, 2003; Rankin et al., 2006; Samaran et al., 2010). This call type is believed to be associated with social (Oleson et al., 2007) or feeding behaviors. Acoustic tag data from North Pacific blue whales revealed that D-calls were produced during shallow dives between deeper feeding dives, suggesting they may be used to either attract conspecifics to productive feeding areas or maintain communication between individuals (Lewis et al., 2018; Oleson et al., 2007). Indeed, Shabangu et al. (2019) detected D-calls off the coast of Benguela (Angola), which suggests that some blue whales may feed in tropical ecosystems. These data particularly align with Best (1967), who observed that the stomachs of blue whales from the west coast of South Africa contained crab larvae, indicating that these whales were feeding in these waters. Thus, a similar study assessing the detection of D-calls in addition to Z-calls would be of interest in the future, to study blue whale feeding behavior more specifically. However, as these D-calls are produced by all acoustic populations, this would not allow the study of the acoustic occurence of Antarctic blue whales specifically.

4.c) Investigating the causes of the summer presence of Antarctic blue whales

Although the migration of mysticetes is generally described as a seasonal cycle of summer feeding in the productive high-latitude areas and winter breeding in the warm low-latitude waters

(Kellogg, 1929), a recent study shows that this migratory model is too simplistic to truly describe the wide range of migratory behaviors of baleen whales, which should rather be considered as a continuum of different strategies (Geijer et al., 2016). In the case of the Antarctic blue whale, the data available today tend more towards partial and differential migration. The observation of individuals year-round in the waters of mid and low latitudes could thus be linked either to individuals skipping migration to the Antarctic polar zone (partial migration) or to a time-shifted migration (differential migration) of the whales to and from these areas.

Partial Migration:

Partial migration occurs when only some individuals migrate during a given season while others remain in the same location. In a partial migration, either the individuals share a breeding area and winter separately, or they share a wintering area and breed separately (Shaw & Levin, 2011). For Antarctic blue whales, partial migration is evidenced by their over-wintering in the Southern Ocean, as demonstrated by year-round sightings at high latitudes, including the waters around South Georgia (Branch et al., 2007), the Western Antarctic Peninsula (Širović et al., 2009), the Weddell Sea (Thomisch, 2016), and Eastern Antarctica (Širović et al., 2004). These year-round detections could reflect the presence of facultative residents, who occasionally skip migration to low latitudes to remain in highly productive areas. Conversely, year-round acoustic detections in low latitudes, as in the present study, in the Indian Ocean (Branch et al., 2007; Leroy et al., 2016; Samaran et al., 2010), the equatorial Atlantic (Haver et al., 2017), and off Namibia (Thomisch et al., 2019) demonstrate that some individuals sometimes skip their migration to high latitudes to stay in warmer waters. While the causes of partial migration in favor of a summer in high latitudes are fairly understandable—such as the energy demands of migration and reproduction, along with an extended feeding period (Chapman et al., 2011). However, the factors that compel some individuals to remain in the breeding area during the feeding season are less clear. Overwintering is energetically demanding due to the scarcity of food at low latitude in the winter months and the considerable energy expenditure associated with migration and reproduction (Lockyer, 1984; Lockyer & Brown, 1981; Silva et al., 2013). The reasons for this migration skipping could be imagined based on the origins of baleen whale migration (see Thomisch, 2016).

Durban & Pitman (2012) suggest, based on observations of orcas and not baleen whales, that migration to low latitudes could have a physiological maintenance function, allowing skin regeneration at lower metabolic costs than at high latitudes and thus improving the energy budget of individuals. If this hypothesis applies to baleen whales, we might suppose that partial migration could fulfill a temporary energetic need, forcing some individuals to stay in warm waters to recover from an illness, wound, etc. Similarly, considering the winter catches off the west coast of South Africa, mostly composed of juveniles in poor body condition, Mackintosh & Wheeler (1929) proposed that non-migrating whales might be juveniles. These young individuals in poor health may not have migrated south the previous summer, preferring to remain in the wintering areas. Corkeron & Connor (1999) hypothesized that the most likely selective advantage of baleen whale migration to low-latitude areas is the reduction of predation pressure from orcas (*Orcinus orca* (Linnaeus, 1758) on calves. In the context of partial migration in Antarctic blue whales, we could assume that some individuals or groups of individuals (especially mother-calf pairs) make the compromise of not returning to feed in the polar zone and stay all year at low latitudes to increase the survival rate of calves against predation. However, this evolutionary strategy hypothesis was questioned by Clapham (2001),

who argued that lethal orca attacks on baleen whales are occasional events and that a more favorable energy balance for calves born in warm waters, leading to increased growth and future reproductive success, would rather represent the main selective force for the evolution of migratory behavior in baleen whales. Thus, the vulnerability of juveniles and smaller individuals to cold (Brodie, 1975; Thomisch, 2016) and predators (Corkeron & Connor, 1999) could be factors influencing either a migration skip to remain in the warmer, calmer waters of low latitudes, or a shortening of wintering at high latitudes. If we adhere to the hypothesis proposed earlier that Z calls are, like the other subspecies, only produced by adult males for reproductive purposes (Lewis et al., 2018; McDonald et al., 2001; Oleson et al., 2007). Females and sexually immature individuals were, therefore, probably not detected in our acoustic study. However, the summer presence of Z-calls in subtropical and sub-Antarctic zones, observed in this study (Fig.6), indicates that some sexually mature males also sometimes skip their migration to the Southern Ocean. The results obtained here (and in previous studies) could thus indicate that reproductive acoustic behaviors are not limited to the winter period and tropical zones (Clark & Clapham, 2004; Oleson et al., 2014; Vu et al., 2012). This also provides another clue to explain the migration skip of sexually mature males; some individuals (especially larger individuals, who are more resistant to fasting (Chapman et al., 2011) might remain in the tropical zone year-round to maximize their reproductive chances. Whaling data, which shows evidence of aseasonal conceptions and out-of-season births in baleen whales, could support this theory (Clark & Clapham, 2004; Mackintosh & Wheeler, 1929). Even if, in the case of the Antarctic blue whale, its breeding ground is not precisely known, and no reproduction or calving behavior has yet been documented.

While the year-round acoustic presence of Antarctic blue whales could also suggest the existence of resident individuals, no resident population has been recently described (Thomisch, 2016). Although such populations are likely to exist in the polar zone (Attard et al., 2016). Here the high inter- and intra-annual variation in detections (Fig.6 and 7) suggests that individuals are not particularly committed to one site throughout the year, but also from one year to the next. Year-round observations could however potentially represent remnants of former resident populations before the era of commercial whaling (Thomisch, 2016). However, the year-round presence of Antarctic blue whales in the subtropical zone might not be entirely due to partial migration but could also stem from variations in migratory behavior among individuals. The varying needs, particularly energetic, of individuals can lead to a time-shifted migration depending on the sex, age, health condition, and reproductive status of the individuals (Brown et al., 1995; Craig et al., 2003; Dawbin, 1997).

Differential Migration:

This staggered or differential migration has been particularly observed by studying the composition of the blue whale population in southern hemisphere whaling zones. Indeed, during the austral summer, adult whales dominate the population at the beginning of the season, while juvenile individuals and lactating females with dependent calves only arrive in the feeding areas in February (Mackintosh & Wheeler, 1929, see also Thomisch, 2016). Thus, individuals of different ages and sexes may differ in their schedules, as well as in their migratory routes and destinations (Mackintosh & Wheeler, 1929). In humpback whales (*Megaptera novaeangliae* (Borowski, 1781)), for example, it has been reported that lactating females spend less time in feeding areas than males, potentially to reduce their calves' exposure to the cold temperatures of the Southern Ocean waters

(Dawbin, 1966). Similarly, sexually immature juveniles are likely to leave their breeding grounds earlier (Craig et al., 2003). Širović et al., (2004, 2009) suggest that the intermittent acoustic occurrence of blue whales is consistent with a gradual and temporally separated migration of animals from Southern Ocean waters, as opposed to a mass movement. This could explain the overall increase in detections we see here over the summer (Fig.6). However, these migratory behaviors are likely to change from year to year, as shown by our results, with different weekly positive hours depending on the years (Fig.6). This change in migratory behavior has notably been observed in two studies (Letsheleha et al., 2022; Shabangu et al., 2019) conducted off the west coast of South Africa, showing different results in terms of seasonality. While Shabangu et al.'s study detected Z-calls only seasonally, Letsheleha et al.'s findings indicate that blue whale Z-calls were present year-round. This suggests that individuals within a population may change their migration patterns across different years or seasons.

Limitations:

In addition to all these hypotheses, prey availability remains relevant to explain the acoustic presence of Antarctic blue whales. Nonetheless, blue whales may not be drawn to areas with the highest total biomass or densest krill aggregations, but rather to regions where krill biomass is more accessible, particularly where swarms are denser and at shallower depths (McQuinn et al. 2016). Yet, we predict here sound scattering layers of pelagic fauna, consisting of various animals that may not be the primary targets of blue whales. A more focused study on krill patches, such as using the Python application RAPIDKRILL[2], rather than on overall zooplankton biomass, might be more relevant for understanding the feeding behavior of this species. Finally, the robustness of the predictions made in this study can also be discussed. Indeed, the R^2 of linear relationships between PCs predicted by Random Forest and values observed in situ are sometimes low, particularly for the third principal component (Appendix 1). In addition, there may be a time lag between environmental variables and their impact on zooplankton biomass. For example, a high chlorophyll a concentration is not instantly followed by a high zooplankton biomass, due to a trophic effect associated with the growth of populations in the trophic web. This time lag is not considered in this study and may have an impact on the results. A multitude of other methods can also be envisaged, to study the prey-predator relationships of blue whales, the interest of the present study stemmed above all from the first linkage of two abundant data sets in the Indian Ocean, THEMISTO and OHASISBIO.

Conclusion

PAM (Passive Acoustic Monitoring) allowed us to collect valuable information on the acoustic presence and distribution patterns of Antarctic blue whales. PAM serves as an alternative to visual methods, which come with high logistical and financial costs, especially in remote and seasonally inaccessible areas. Active acoustics, on the other hand, have enabled the modeling of the 3-D distribution of zooplankton and its variability in relation to physical and climatic phenomena. Although this study did not succeed in demonstrating a clear trophic picture between zooplankton biomass and the acoustic presence of blue whales at low and mid-latitudes during the summer period, the hypotheses posed here suggest a complexity and flexibility in the migratory patterns of Antarctic blue whales, influenced by environmental, dietary, social, and potentially anthropogenic factors. A deeper understanding of these behaviors would require further studies, particularly the use of acoustic

[2] https://www.bas.ac.uk/project/rapidkril

tags to more specifically analyze the feeding and social behaviors of this species in different regions.

Acknowledgements.

This study would not have been possible without the generous support of ISBLUE who financed this research internship, thanks to them. First of all, I'd like to thank the OSMOSE team: Julie Béesau, Mathieu Dupont, Maëlle Torterotot, Elodie Morin, Dorian Cazau and Anatole Gros-Martial for their warm welcome. I hope to have the chance to work alongside you again one day! Many thanks to Flore Samaran for her guidance and for giving me the chance to work alongside her on this fascinating subject. Thanks also to Lloyd Izard for holding my hand through the twists and turns of this complicated method, which was a big challenge for me, so thank you for your patience! And last but not least, a special thank you to Mathilde Michel for her support, kindness and advice. I was delighted to be your second brain for these few months! See you all soon, I hope.

References

Attard, C. R. M., Beheregaray, L. B., & Möller, L. M. (**2016**). Towards population-level conservation in the critically endangered Antarctic blue whale: The number and distribution of their populations. *Scientific Reports*, *6*(1), 22291. https://doi.org/10.1038/srep22291

Béhagle, N., Cotté, C., Ryan, T. E., Gauthier, O., Roudaut, G., Brehmer, P., Josse, E., & Cherel, Y. (**2016**). Acoustic micronektonic distribution is structured by macroscale oceanographic processes across 20–50°S latitudes in the South-Western Indian Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *110*, 20–32. https://doi.org/10.1016/j.dsr.2015.12.007

Benoit-Bird, K., & Lawson, G. (**2015**). Ecological Insights from Pelagic Habitats Acquired Using Active Acoustic Techniques. *Annual Review of Marine Science*, *8*. https://doi.org/10.1146/annurev-marine-122414-034001

Best, P. B. (**1967**). *Distribution and feeding habits of baleen whales off the Cape Province*. Republic of South Africa, Dept. of Commerce and Industries, Division of Sea Fisheries.

Boersch-Supan, P. H., Rogers, A. D., & Brierley, A. S. (**2017**). The distribution of pelagic sound scattering layers across the southwest Indian Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *136*, 108–121. https://doi.org/10.1016/j.dsr2.2015.06.023

Branch, T. A., Stafford, K. M., Palacios, D. M., Allison, C., Bannister, J. L., Burton, C. L. K., Cabrera, E., Carlson, C. A., Galletti Vernazzani, B., Gill, P. C., Hucke-Gaete, R., Jenner, K. C. S., Jenner, M. -N. M., Matsuoka, K., Mikhalev, Y. A., Miyashita, T., Morrice, M. G., Nishiwaki, S., Sturrock, V. J., … Warneke, R. M. (**2007**). Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the Southern Hemisphere and northern Indian Ocean. *Mammal Review*, *37*(2), 116–175. https://doi.org/10.1111/j.1365-2907.2007.00106.x

Brierley, A. S. (**2014**). Diel vertical migration. *Current Biology*, *24*(22), R1074–R1076. https://doi.org/10.1016/j.cub.2014.08.054

Brodie, P. F. (**1975**). Cetacean Energetics, an Overview of Intraspecific Size Variation. *Ecology*, *56*(1), 152–161. https://doi.org/10.2307/1935307

Brown, M. R., Corkeron, P. J., Hale, P. T., Schultz, K. W., & Bryden, M. M. (**1995**). Evidence for a sex-segregated migration in the humpback whale (Megaptera novaeangliae). *Proceedings. Biological Sciences*, *259*(1355), 229–234. https://doi.org/10.1098/rspb.1995.0034

Chapman, B. B., Brönmark, C., Nilsson, J.-Å., & Hansson, L.-A. (**2011**). The ecology and evolution of partial migration. *Oikos*, *120*(12), 1764–1775. https://doi.org/10.1111/j.1600-0706.2011.20131.x

Clapham, P. (**2001**). Why do Baleen Whales Migrate? *Marine Mammal Science*, *17*(2), 432–436. https://doi.org/10.1111/j.1748-7692.2001.tb01289.x

Clark, C. W., & Clapham, P. J. (**2004**). Acoustic monitoring on a humpback whale (Megaptera novaeangliae) feeding ground shows continual singing into late spring.

Proceedings of the Royal Society of London. Series B: Biological Sciences, *271*(1543), 1051–1057. https://doi.org/10.1098/rspb.2004.2699

Corkeron, P. J., & Connor, R. C. (**1999**). Why Do Baleen Whales Migrate? *Marine Mammal Science*, *15*(4), 1228–1245. https://doi.org/10.1111/j.1748-7692.1999.tb00887.x

Craig, A. S., Herman, L. M., Gabriele, C. M., & Pack, A. A. (**2003**). Migratory Timing of Humpback Whales (Megaptera novaeangliae) in the Central North Pacific Varies with Age, Sex and Reproductive Status. *Behaviour*, *140*(8/9), 981–1001.

Cummings, W. C., & Thompson, P. O. (**1971**). Underwater Sounds from the Blue Whale, Balaenoptera musculus. *The Journal of the Acoustical Society of America*, *50*(4B), 1193–1198. https://doi.org/10.1121/1.1912752

Dawbin, W. H. (William H. (**1966**, January 1). *The seasonal migratory cycle of humpback whales / William H. Dawbin.* The Seasonal Migratory Cycle of Humpb... | Items | National Library of New Zealand | National Library of New Zealand; [Calif. : University of California Press, 1966]. https://natlib.govt.nz/records/20907691

de Vos, A., Faux, C. E., Marthick, J., Dickinson, J., & Jarman, S. N. (**2018**). New Determination of Prey and Parasite Species for Northern Indian Ocean Blue Whales. *Frontiers in Marine Science*, *5*. https://www.frontiersin.org/articles/10.3389/fmars.2018.00104

Dréo, R., Bouffaut, L., Leroy, E., Barruol, G., & Samaran, F. (**2019**). Baleen Whale distribution and seasonal occurrence Revealed by an ocean bottom seismometer network in The Western Indian Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *161*, 132–144. https://doi.org/10.1016/j.dsr2.2018.04.005

Durban, J. W., & Pitman, R. L. (**2012**). Antarctic killer whales make rapid, round-trip movements to subtropical waters: Evidence for physiological maintenance migrations? *Biology Letters*, *8*(2), 274–277. https://doi.org/10.1098/rsbl.2011.0875

Geijer, C., Notarbartolo di Sciara, G., & Panigada, S. (**2016**). Mysticete migration revisited: Are Mediterranean fin whales an anomaly? *Mammal Review*, *46*. https://doi.org/10.1111/mam.12069

Goldbogen, J. A., Calambokidis, J., Oleson, E., Potvin, J., Pyenson, N. D., Schorr, G., & Shadwick, R. E. (**2011**). Mechanics, hydrodynamics and energetics of blue whale lunge feeding: Efficiency dependence on krill density. *Journal of Experimental Biology*, *214*(1), 131–146. https://doi.org/10.1242/jeb.048157

Haver, S. M., Klinck, H., Nieukirk, S. L., Matsumoto, H., Dziak, R. P., & Miksis-Olds, J. L. (**2017**). The not-so-silent world: Measuring Arctic, Equatorial, and Antarctic soundscapes in the Atlantic Ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, *122*, 95–104. https://doi.org/10.1016/j.dsr.2017.03.002

Hazen, E. L., Friedlaender, A. S., & Goldbogen, J. A. (**2015**). Blue whales (Balaenoptera musculus) optimize foraging efficiency by balancing oxygen use and energy gain as a function of prey density. *Science Advances*, *1*(9), e1500469. https://doi.org/10.1126/sciadv.1500469

Izard, L. (**2023**). *Structuration spatiale et variabilité des écosystèmes mésopélagiques dans l'Océan Indien Sud* [Phdthesis, Sorbonne Université]. https://theses.hal.science/tel-04524647

Izard,L, Ariza, A., Fonvieille, N., Goberville, E., Djian, V., Kristiansen, A., Charrassin,J-B., Cotté, C. (**2024**). Large differences in the distribution of pelagic biomass as a result of sonar frequency choice. *Proc. R. Soc. B. under review*

Kang, M., Oh, S., Oh, W., Kang, D.-J., Nam, S., & Lee, K. (**2024**). Acoustic characterization of fish and macroplankton communities in the Seychelles-Chagos Thermocline Ridge of the southwest Indian ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *213*, 105356. https://doi.org/10.1016/j.dsr2.2023.105356

Kellogg, R. (**1929**). *What is known of the migrations of some of the whalebone whales*. US Government Printing Office.

Leroy, E. C., Royer, J.-Y., Alling, A., Maslen, B., & Rogers, T. L. (**2021**). Multiple pygmy blue whale acoustic populations in the Indian Ocean: Whale song identifies a possible new population. *Scientific Reports*, *11*(1), 8762. https://doi.org/10.1038/s41598-021-

88062-5

Leroy, E. C., Royer, J.-Y., Bonnel, J., & Samaran, F. (**2016**). *Antarctic blue whales in the southern Indian Ocean: A multi-year and multi-site database of Z-call detections* [Dataset]. SEANOE. https://doi.org/10.17882/46124

Letsheleha, I. S., Shabangu, F. W., Farrell, D., Andrew, R. K., la Grange, P. L., & Findlay, K. P. (**2022**). Year-round acoustic monitoring of Antarctic blue and fin whales in relation to environmental conditions off the west coast of South Africa. *Marine Biology*, *169*(3), 41. https://doi.org/10.1007/s00227-022-04026-x

Lewis, L. A., Calambokidis, J., Stimpert, A. K., Fahlbusch, J., Friedlaender, A. S., McKenna, M. F., Mesnick, S. L., Oleson, E. M., Southall, B. L., Szesciorka, A. R., & Širović, A. (**2018**). Context-dependent variability in blue whale acoustic behaviour. *Royal Society Open Science*, *5*(8), 180241. https://doi.org/10.1098/rsos.180241

Ljungblad, D., Clark, C., & Shimada, H. (**1998**). A comparison of sounds attributed to pygmy blue whales (Balaenoptera musculus brevicauda) recorded south of the Madagascar Plateau and those attributed to "true" blue whales (Balaenoptera musculus) recorded off Antarctica. *Report of the International Whaling Commission*, *48*, 439–442.

Lockyer, C. (**1984**). Review of Baleen Whale (Mysticeti) Reproduction and Implications for Management. *Reproduction in Whales, Dolphins and Porpoises. Proc. Conference, La Jolla, CA, 1981*, *6*, 27–50.

Lockyer, C. H., & Brown, S. G. (**1981**). The migration of whales. *Animal Migration*, *13*, 105–137.

McDonald, M. A., Calambokidis, J., Teranishi, A. M., & Hildebrand, J. A. (**2001**). The acoustic calls of blue whales off California with gender data. *The Journal of the Acoustical Society of America*, *109*(4), 1728–1735. https://doi.org/10.1121/1.1353593

McDonald, M. A., Mesnick, S. L., & Hildebrand, J. A. (**2006**). Biogeographic characterization of blue whale song worldwide: Using song to identify populations. *Journal of Cetacean Research and Management*, *8*(1), 55–65.

McQuinn, I., Gosselin, J.-F., Bourassa, M.-N., Mosnier, A., Pierre, J.-F., Plourde, S., Lesage, V., & Raymond, A. (2016). The spatial association of blue whales (Balaenoptera musculus) with krill patches (Thysanoessa spp. And Meganyctiphanes norvegica) in the estuary and northwestern Gulf of St. Lawrence.

Mellinger, D. K., & Clark, C. W. (**2003**). Blue whale (Balaenoptera musculus) sounds from the North Atlantic. *The Journal of the Acoustical Society of America*, *114*(2), 1108–1119. https://doi.org/10.1121/1.1593066

Mackintosh N A & Wheeler J F G. (**1929**). *Southern Blue and Fin Whales*. http://archive.org/details/biostor-224615

Nicolas, M. & LO MONACO Claire. (**2017**). *MD 206 / OBSAUSTRAL cruise,Marion Dufresne R/V*. Sismer. https://doi.org/10.17600/17002400

Oleson, E. M., Calambokidis, J., Burgess, W. C., McDonald, M. A., LeDuc, C. A., & Hildebrand, J. A. (**2007**). Behavioral context of call production by eastern North Pacific blue whales. *Marine Ecology Progress Series*, *330*, 269–284. https://doi.org/10.3354/meps330269

Oleson, E. M., Širović, A., Bayless, A. R., & Hildebrand, J. A. (**2014**). Synchronous Seasonal Change in Fin Whale Song in the North Pacific. *PLOS ONE*, *9*(12), e115678. https://doi.org/10.1371/journal.pone.0115678

Pauthenet, E., Roquet, F., Madec, G., Guinet, C., Hindell, M., McMahon, C. R., Harcourt, R., & Nerini, D. (**2018**). Seasonal Meandering of the Polar Front Upstream of the Kerguelen Plateau. *Geophysical Research Letters*, *45*(18), 9774–9781. https://doi.org/10.1029/2018GL079614

R Core Team. (**2023**). *_R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.* [Computer software].

Rankin, S., Barlow, J., & Stafford, K. (**2006**). Blue whale (*balaenoptera musculus*) sightings and recordings south of the Aleutian islands. *United States Department of Commerce: Staff Publications*. https://digitalcommons.unl.edu/usdeptcommercepub/309

Royer Jean-Yves. (**2009**). *OHA-SIS-BIO - OBSERVATOIRE HYDROACOUSTIQUE*. Sismer. https://doi.org/10.18142/229

Samaran, F., Adam, O., & Guinet, C. (**2010**). Discovery of a mid-latitude sympatric area for two Southern Hemisphere

blue whale subspecies. *Endangered Species Research*, *12*(2), 157–165. https://doi.org/10.3354/esr00302

Samaran, F., Stafford, K. M., Branch, T. A., Gedamke, J., Royer, J.-Y., Dziak, R. P., & Guinet, C. (**2013**). Seasonal and Geographic Variation of Southern Blue Whale Subspecies in the Indian Ocean. *PLOS ONE*, *8*(8), e71561. https://doi.org/10.1371/journal.pone.0071561

Shabangu, F. W., Findlay, K. P., Yemane, D., Stafford, K. M., van den Berg, M., Blows, B., & Andrew, R. K. (**2019**). Seasonal occurrence and diel calling behaviour of Antarctic blue whales and fin whales in relation to environmental conditions off the west coast of South Africa. *Journal of Marine Systems*, *190*, 25–39. https://doi.org/10.1016/j.jmarsys.2018.11.002

Shaw, A. K., & Levin, S. A. (**2011**). To breed or not to breed: A model of partial migration. *Oikos*, *120*(12), 1871–1879. https://doi.org/10.1111/j.1600-0706.2011.19443.x

Silva, M. A., Prieto, R., Jonsen, I., Baumgartner, M. F., & Santos, R. S. (**2013**). North Atlantic blue and fin whales suspend their spring migration to forage in middle latitudes: Building up energy reserves for the journey? *PloS One*, *8*(10), e76507. https://doi.org/10.1371/journal.pone.0076507

Širović, A., Hildebrand, J. A., Wiggins, S. M., McDonald, M. A., Moore, S. E., & Thiele, D. (**2004**). Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula. *Deep Sea Research Part II: Topical Studies in Oceanography*, *51*(17), 2327–2344. https://doi.org/10.1016/j.dsr2.2004.08.005

Širović, A., Hildebrand, J. A., Wiggins, S. M., & Thiele, D. (**2009**). Blue and fin whale acoustic presence around Antarctica during 2003 and 2004. *Marine Mammal Science*, *25*(1), 125–136. https://doi.org/10.1111/j.1748-7692.2008.00239.x

Socheleau, F.-X., & Samaran, F. (**2017**). *Detection of Mysticete Calls: A Sparse Representation-Based Approach* (Research Report RR-2017-04-SC). Dépt. Signal et Communications (Institut Mines-Télécom-IMT Atlantique-UBL) ; Laboratoire en sciences et technologies de l'information, de la communication et de la connaissance (UMR 6285 - CNRS - IMT Atlantique - Université de Bretagne Occidentale - Université de Bretagne Sud - ENSTA Bretagne - Ecole Nationale d'ingénieurs de Brest) ; École nationale supérieure de techniques avancées Bretagne. (Ministère de la Défense). https://hal.science/hal-01736178

Stafford, K., Chapp, E., Bohnenstiel, D., & Tolstoy, M. (**2010**). Seasonal detection of three types of "pygmy" blue whale calls in the Indian Ocean. *Marine Mammal Science*, *27*, 828–840. https://doi.org/10.1111/j.1748-7692.2010.00437.x

Thomisch, K. (**2016**). *Distribution patterns and migratory behavior of Antarctic blue whales*.

Thomisch, K., Boebel, O., Bachmann, J., Filun, D., Neumann, S., Spiesecke, S., & van Opzeeland, I. (**2019**). Temporal patterns in the acoustic presence of baleen whale species in a presumed breeding area off Namibia. *MARINE ECOLOGY PROGRESS SERIES*, *620*, 201–214. https://doi.org/10.3354/meps12952

Torterotot, M., Samaran, F., & Royer, J. (**2023**). Long-term acoustic monitoring of nonstereotyped blue whale calls in the southern Indian Ocean. *Marine Mammal Science*, *39*(2), 594–610. https://doi.org/10.1111/mms.12998

Torterotot, M., Samaran, F., Stafford, K. M., & Royer, J.-Y. (**2020**). Distribution of blue whale populations in the Southern Indian Ocean based on a decade of acoustic monitoring. *Deep Sea Research Part II: Topical Studies in Oceanography*, *179*, 104874. https://doi.org/10.1016/j.dsr2.2020.104874

Vu, E. T., Risch, D., Clark, C. W., Gaylord, S., Hatch, L. T., Thompson, M. A., Wiley, D. N., & Van Parijs, S. M. (**2012**). Humpback whale song occurs extensively on feeding grounds in the western North Atlantic Ocean. *Aquatic Biology*, *14*(2), 175–183. https://doi.org/10.3354/ab00390

W H Dawbin. (**1997**). *Temporal segregation of humpback whales during migration in Southern Hemisphere waters*. http://archive.org/details/biostor-245684

Appendix 1: Supplementary materials (Based on Izard et al., under review).

The following supplementary material provides detailed information on the methodology and intermediate results used to describe and model the vertical distribution of backscatter across the Indian sector of the Southern Ocean. Firstly, we detail the analysis of *in situ* acoustic observations, focusing on the application of functional data analysis and clustering methods. Secondly, the backscatter modelling process using Random Forest with environmental variables as predictors is explained.

1. **Describing *in situ* acoustic observations**

The statistical workflow for the first part of our analyses is explained hereafter and visualized in Figure 1. Example is provided for simplicity for frequencies 18 and 38 kHz, but analyses have also been carried out for frequencies 70 and 120 kHz.

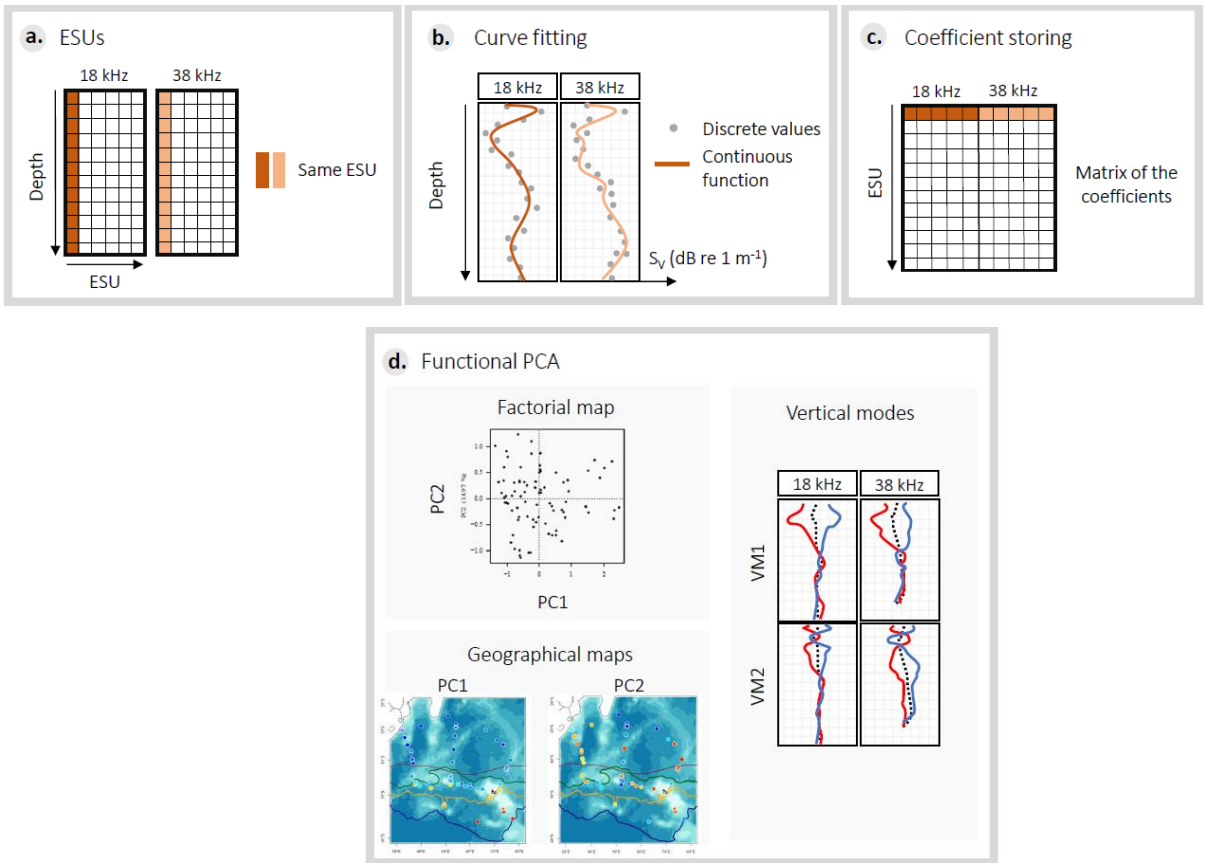


Figure 1. Workflow detailing the analysis of acoustic observations to identify recurring patterns in the vertical distribution of backscatter at 18, 38, 70 and 120 kHz. (a, b) **Smoothing of acoustic profiles**: an acoustic observation is defined as the combination of simultaneously emitted at 18, 38, 70 and 120 kHz profiles. The profiles are smoothed using 10 B-spline basis functions. (c) **Storing of smoothing coefficients**: the smoothing coefficients from both frequencies are stored in a matrix **X**. (d) **Functional Principal Component Analysis (fPCA)**: a multivariate functional Principal Component Analysis is

applied to matrix **X**. The resulting principal components (PCs), representing new coordinates of the observations, are visualized on a biplot combining both the principal component space and a geographical map. The associated vertical modes (eigenfunctions) indicate the depth maximizing variability.

Step a, b – Smoothing of acoustic profiles. The acoustic observations consist of pairs of S_v profiles (18, 38, 70 and 120 kHz) for both daytime and nighttime datasets. Here, we express the 18, 38, 70 and 120 kHz profiles as a linear combination of 10 B-splines functions to analyze their shape and how they vary together (Figure 2).

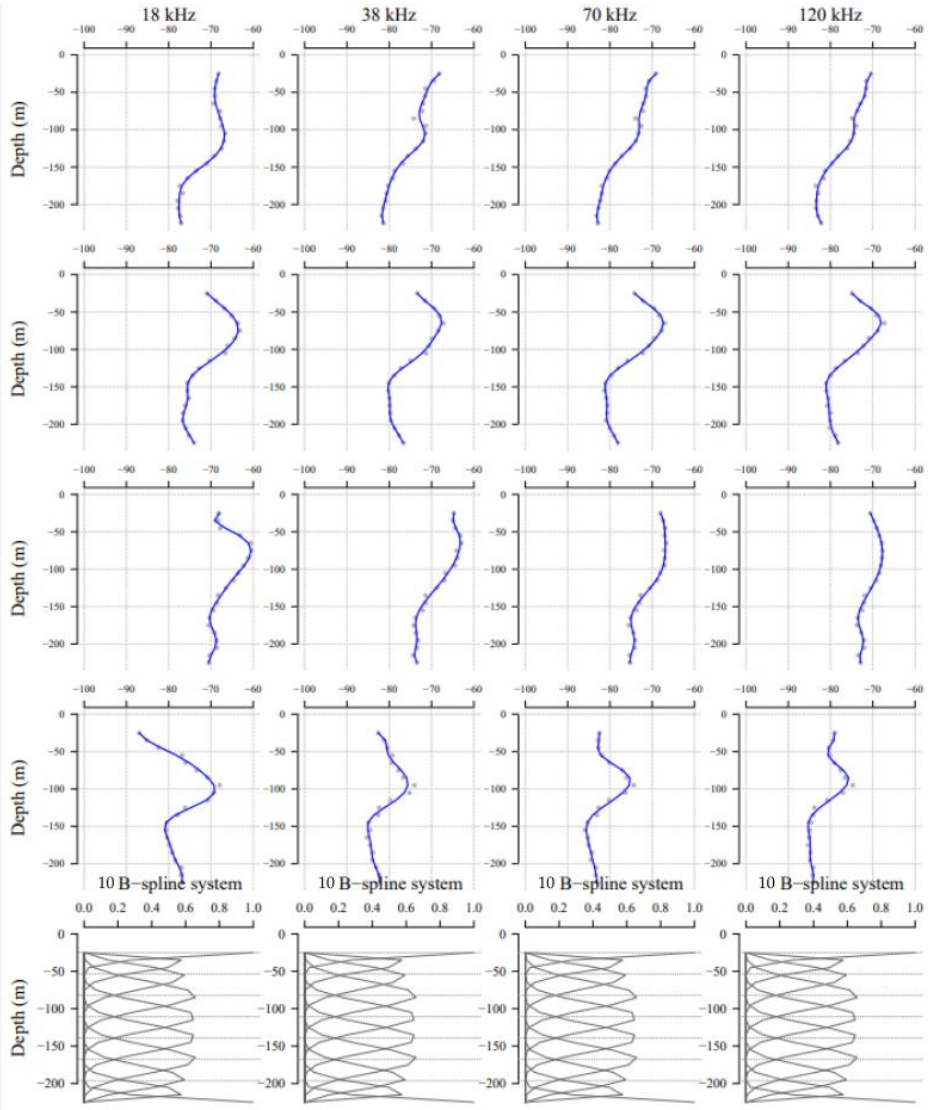


Figure 2. Examples of smoothing applied to four acoustic observations from the daytime dataset. The 18, 38, 70 and 120 kHz profiles are smoothed using a 10 B-spline basis system. Grey dots represent the original data, and the blue lines represent the smoothed functional data.

Step c, d - Storing of smoothing coefficients and fPCA. We apply a functional Principal Component Analysis (fPCA) to capture the key patterns in the variability of the smoothed acoustic profiles. The

smoothing coefficients from all frequencies (18, 38, 70 and 120 kHz) are concatenated into a single matrix X with dimensions N x 2K, where N represents the number of acoustic observations (varying between 82 for nighttime and 98 for daytime data) and K denotes the number of spline coefficients (which is the same for all frequencies). The fPCA decomposes this matrix **X** to identify the main modes of variability. The first four principal components (PCs) explain a substantial portion of the total variance, accountingfor 74% and 92% for the daytime and nighttime datasets, respectively. The contribution of each frequency (18, 38, 70 and 120 kHz) to these principal components differs between day and night but seem equivalent for the first 3 modes, explaining between 20 and 30% of the variance each (Figure 3).

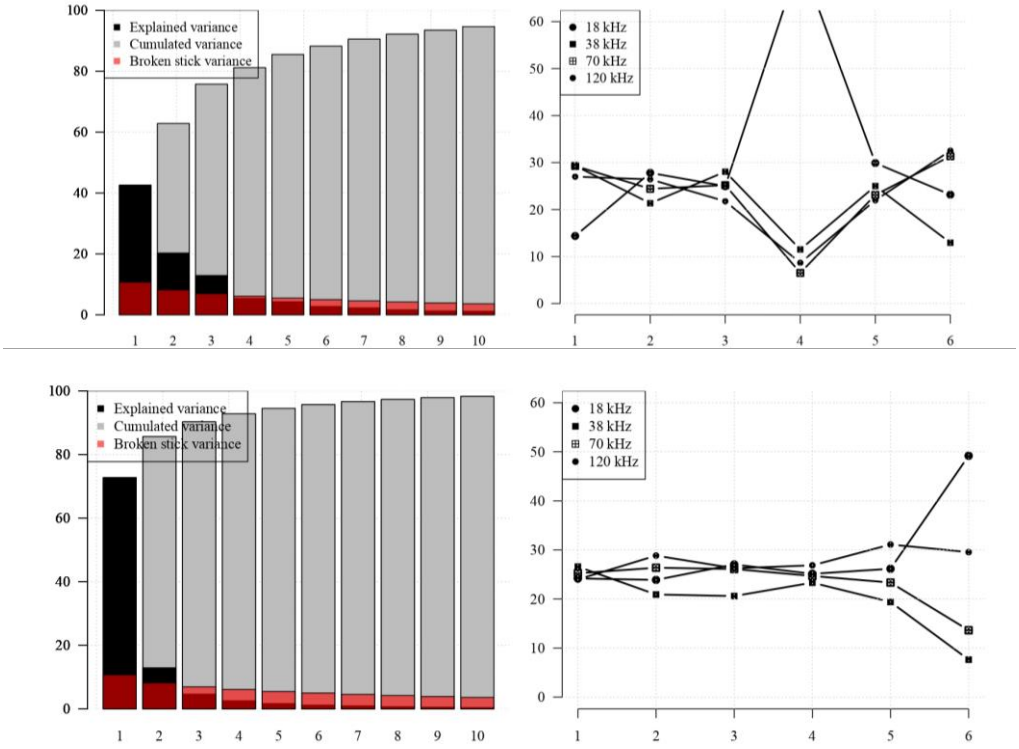


Figure 3. Contribution of each frequency (18, 38, 70 and 120 kHz) in percent (y-axis) to principal components (x-axis) for both daytime (top panel) and nigttime (bottome panel).

The principal components analysis results in linearly uncorrelated combinations of the original profiles (18, 38, 70 and 120 kHz), called vertical modes (VMs). These VMs capture the maximum variance in the data and essentially form a new basis for representing any combination of 18, 38, 70 and 120 kHz profiles. This property is used in Section 2 for further analyses. The PC scores for each observation represent the new coordinates in this space and indicate the degree to which a specific acoustic profile aligns with a given VM (Figure 4; Figure 5).

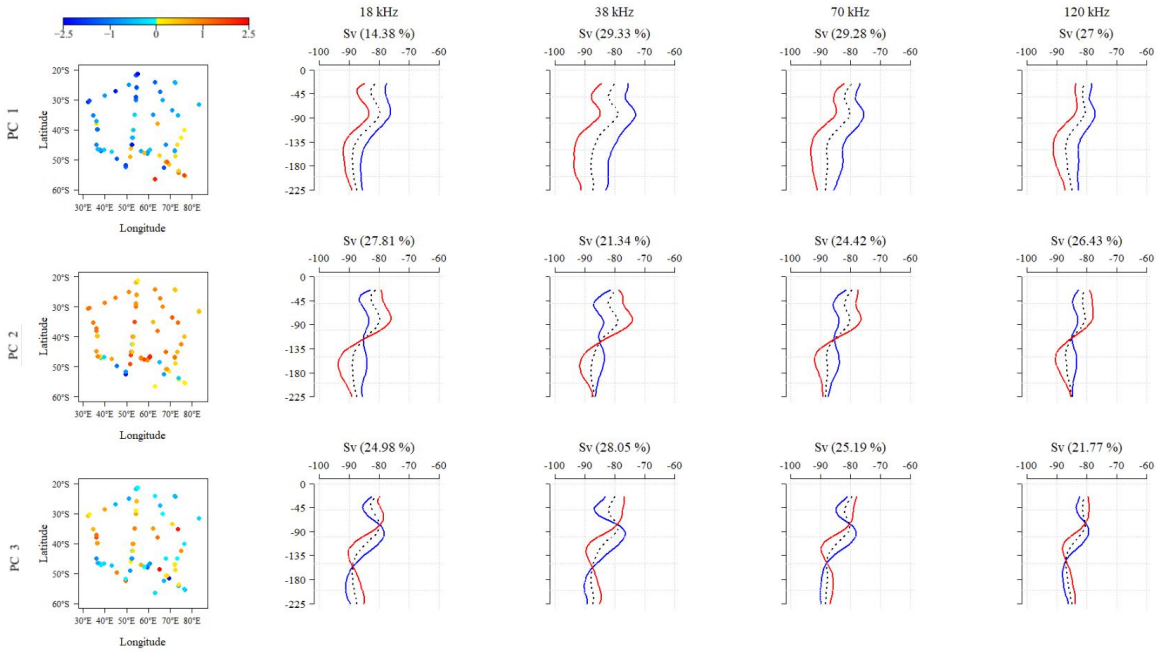


Figure 4. The three principal components (map) associated with the first three vertical modes of variability (eigenfunctions) by frequency (18, 38, 70 and 120 kHz) for daytime data.

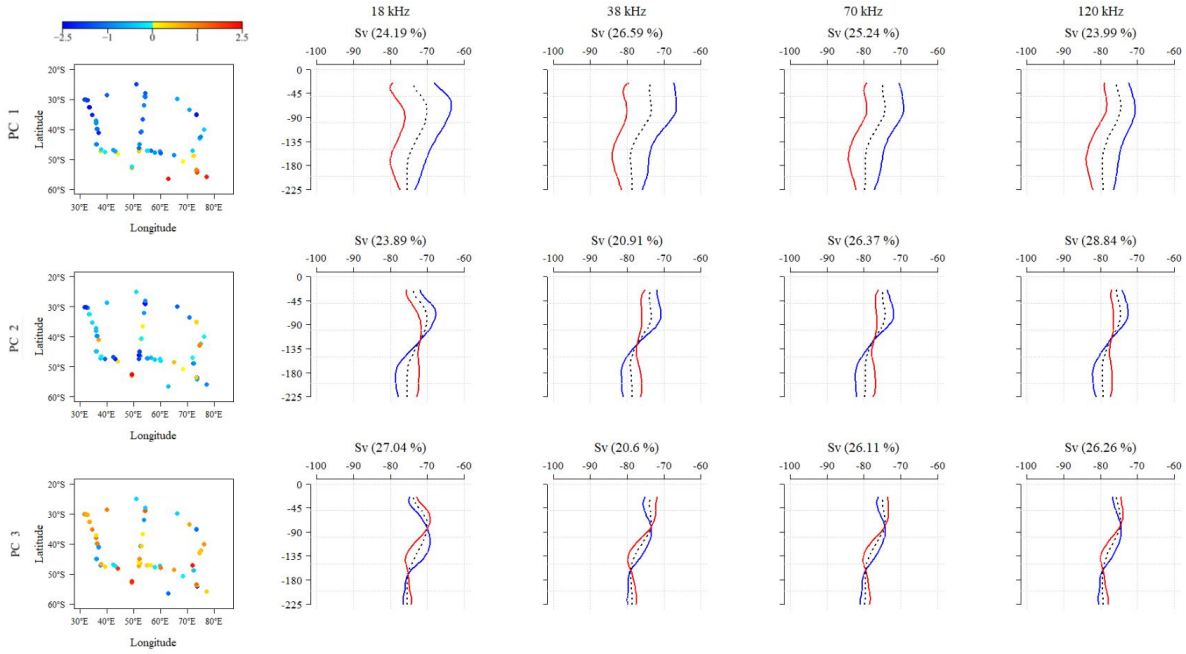


Figure 5. The three principal components (map) associated with the first three vertical modes of variability (eigenfunctions) by frequency (18, 38, 70 and 120 kHz) for night-time data.

2. Extending *in situ* observation to regional acoustic seascapes

The statistical workflow for the second part of our analyses is detailed hereafter and visualized in Figure 6. Step 2a. follows Step 1d.

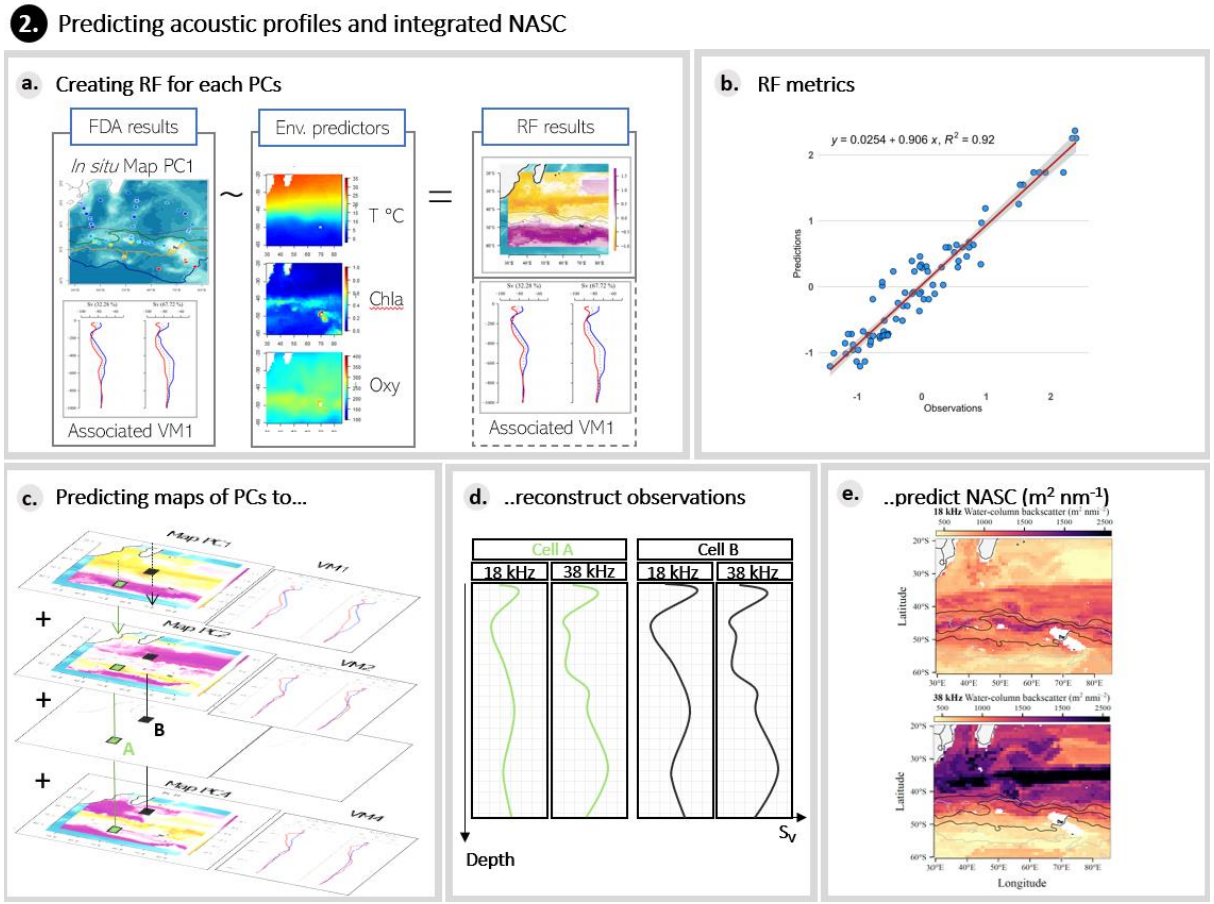


Figure 6. Workflow associated with extending clusters to acoustic seascapes. (a) **Modelling Principal Components** (PCs): each PC, representing a mode of variability in the backscatter profiles, is modelled independently using environmental variables as predictors. (b) **Model evaluation**: the quality of each model predicting a PC is assessed. (c) **Predicting vertical structure**: the modelling process yields a series of predicted PCs associated with vertical modes of variability. (d) **Reconstructing acoustic observations**: the predicted PCs are weighted by their respective predicted values and summed to reconstruct acoustic observation in S_v. (e) **Estimating biomass**: each predicted S_v profile can be transformed into NASC to investigate spatial patterns of integrated biomass across the seascape.

Step a - Modelling Principal Components. To extend the analysis beyond the identified clusters and create maps representing the key backscatter patterns across the entire region, we employ Random Forest models. This approach involves modelling each principal component as a function of environmental variables. We build a separate model for each PC, using environmental data (surface temperature (°C), chlorophyll a (mg m^{-3}), and dissolved oxygen (mmol kg^{-1})) to predict the spatial distribution of the PC

across the study area. This allows us to map the predicted values for each PC, revealing large-scale spatial patterns associated with the original backscatter signals captured by the PCs. It is important to note that this modelling process focuses solely on the PCs, leaving the underlying vertical modes (eigenfunctions) unchanged. These eigenfunctions provide a basis for reconstructing the acoustic observations in later steps.

Step b - Model evaluation. Results of the Random Forest predictions are generated for each principal component for both daytime and nighttime datasets, resulting in a total of eight models. To ensure model robustness and prevent overfitting, we implement a hyperparameter tuning strategy. Specifically, a hyperparameter grid search is conducted, selecting five key hyperparameters for optimization:

1. **The number of trees** balances the need for error stability with computational efficiency.
2. **The number of variables** randomly sampled at each split.
3. **The number of samples** used for training. By default, this corresponds to 63.25% of the training dataset to reflect the expected number of unique observations in a bootstrap sample. Deviating from this value can influence training time and introduce various levels of bias.
4. **The minimum number of samples within terminal nodes** influences tree complexity, depth, and intricacy. Smaller node sizes yield deeper and more complex trees, while larger nodes lead to shallower trees. This choice involves a trade-off between bias and variance, with deeper trees potentially introducing more variance (risk of overfitting) and shallower trees potentially increasing bias.
5. **The maximum number of terminal nodes** offers another way to manage tree complexity. More nodes result in deeper, more complex trees, while fewer nodes lead to shallower trees.

To ensure model robustness and prevent overfitting, a rigorous hyperparameter tuning strategy is implemented using the "RandomDiscrete" method within the R package h2o. This method explores a large number of hyperparameter combinations (20*3*5*6*4 = 7200 in our case) by randomly sampling different settings for: Number of trees in the forest (=20); number of variables considered at each split (=3); proportion of samples used for training (=5); minimum number of samples in terminal nodes (=6); maximum number of terminal nodes (=4).

The search process stops when a pre-defined stopping criterion is met, such as a significant improvement in model performance, reaching a time limit, or evaluating a specified number of models. Here, we consider all three criteria to identify the best models for each principal component. Model performance is evaluated using Out-of-Bag Root Mean Squared Error (OOB RMSE). While the random search approach may not guarantee the absolutely optimal model, it is known for effectively finding models with minimal

OOB RMSE. Additionally, we calculate mean residual deviance, mean absolute error (MAE), and R-squared to assess the overall quality of the models (Table 1, Figure 7). Once the best models are identified for each PC in both day and night data, we use them to predict the spatial distribution of PCs across the entire study area based on environmental data. Finally, we evaluate the agreement between the original PC values and the model-predicted values using the Pearson correlation coefficient.

Table 1. Performance evaluation of Random Forest models for each principal component (PC) for both daytime and nighttime datasets. RMSE: Root Mean Squared Error. MAE: Mean Absolute value of Errors. R-squared: Coefficient of determination. p: pvalue.

		Day			**Night**		
		PC1	**PC2**	**PC3**	**PC1**	**PC2**	**PC3**
Hyperparameter values	Number of trees	30	50	45	63	39	93
	Number of variables sampled at each split	2	2	3	2	2	3
	Number of samples used for training	0.632	0.700	0.800	0.700	0.550	0.632
	Minimum number of samples within terminal nodes	1	1	1	1	1	1
	Maximum number of terminal nodes	5	10	5	20	30	10
(Mean) Metric of evaluation	RMSE	0.738	0.647	0.428	0.258	0.368	0.235
	Mean residual deviance	0.544	0.418	0.183	0.073	0.161	0.058
	MAE	0.550	0.486	0.379	0.647	0.409	0.267
	R-squared	0.600	0.490	0.250	0.700	0.380	0.110
	Pearson correlation between observed and modelled PC	0.857 p <0.001	0.697 p<0.001	0.505 p<0.001	0.838 p<0.001	0.617 p<0.001	0.333 p<0.005

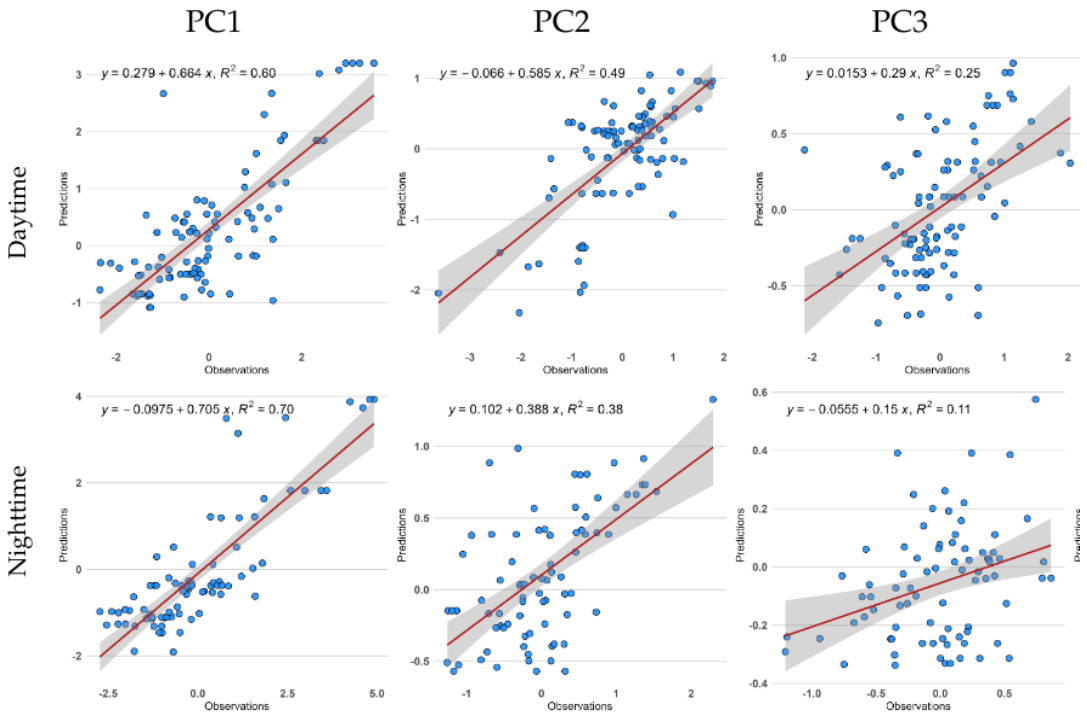


Figure 7. Random forest linear regression between predicted and observed principal component scores for both daytime (top panel) and nighttime (bottom panel).

The relative influence of each environmental variable on the predicted PCs is assessed for both daytime and nighttime data. We consider (i) the frequency of selection, indicating how often a variable is chosen to split the data during the construction of trees in the Random Forest models; and (ii) the error reduction to assess the extent to which a variable contributed to reducing the squared error across all the trees in the forest. Temperature emerges as the most influential variable for predicting PC1 in both daytime and nighttime data, although its impact is slightly weaker for PC2 at night. Chlorophyll a play an important role in explaining PC2 and PC3 during both daytime and nighttime (Figure 8).

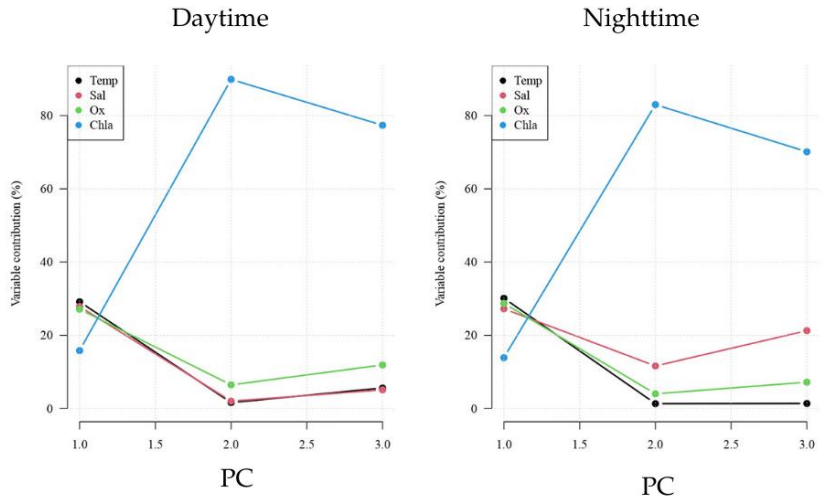


Figure 8. Percentage of importance of the three environmental variables used to model the first three principal components for both daytime (left panel) and nighttime (right panel).

Step c- Predicting vertical structure. Steps 2a and 2b culminate in a series of eight maps, one for each PC in both daytime and nighttime (Figure 8). Each map corresponds to a specific vertical backscatter pattern depicted in Figure 5. For instance, the first PC map (both in daytime and nighttime) reflects the latitudinal variation observed in the vertical distribution of organisms south of the Subantarctic Front. Other PCs, such as PC4 for daytime and PC3 for nighttime, capture more localized patterns, particularly around the subantarctic islands.

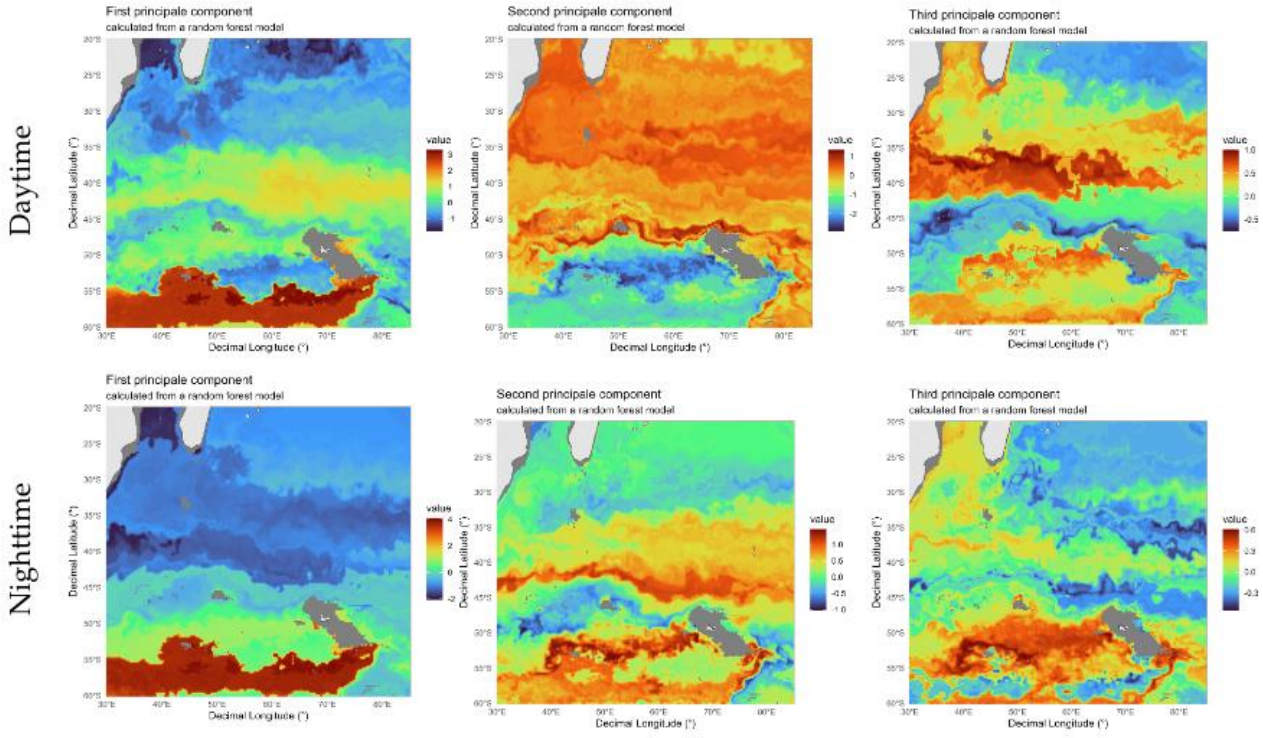


Figure 9. Predicted principal component scores from the Random Forest models. Top panels: daytime; Bottom panels: nighttime.

Step d - Reconstructing acoustic observations. Leveraging the predicted PC maps and the previously identified vertical modes (VMs; Figure 2a), we can reconstruct acoustic observations at any grid cell in the study area. This reconstruction follows the notation by Pauthenet et al. (2018):

$$\begin{cases} P18_n^{(q)}(z) = \overline{P18}(z) + \sum_{k=1}^{q} y_{n,k}\xi_k^{P18}(z) \\ P38_n^{(q)}(z) = \overline{P38}(z) + \sum_{k=1}^{q} y_{n,k}\xi_k^{P38}(z) \end{cases}$$

where q=4; n = 1, …, N; k = 1, …, 3 the mode number; z the depth in m (z ∈ 25-225) and $\underline{P_{38}}(z)$ and $\underline{P_{18}}(z)$ are the mean profiles.

Step e - Estimating biomass. The volume backscattering coefficient (s_v) is obtained by converting the predicted smoothed profiles from volume backscattering strength (S_v, dB re 1 m^{-1}). The integrated Nautical Area Scattering Coefficient (NASC) is calculated from the S_v profiles and displayed at the bottom of each profile. This NASC value serves as a proxy for the total biomass integrated within the water column.

$$s_v = 10^{(S_v/10)} \qquad (\mathrm{m}^{-1})$$

$$s_a = \int_{-1000}^{-25} s_v \, dz \qquad (\mathrm{m}^2\mathrm{m}^{-2})$$

$$NASC = 4\pi(1852)^2 s_a \qquad (\mathrm{m}^2\mathrm{nmi}^{-2})$$

This method was then applied to obtain prediction maps for each week between January and March 2020, 2021, and 2022.

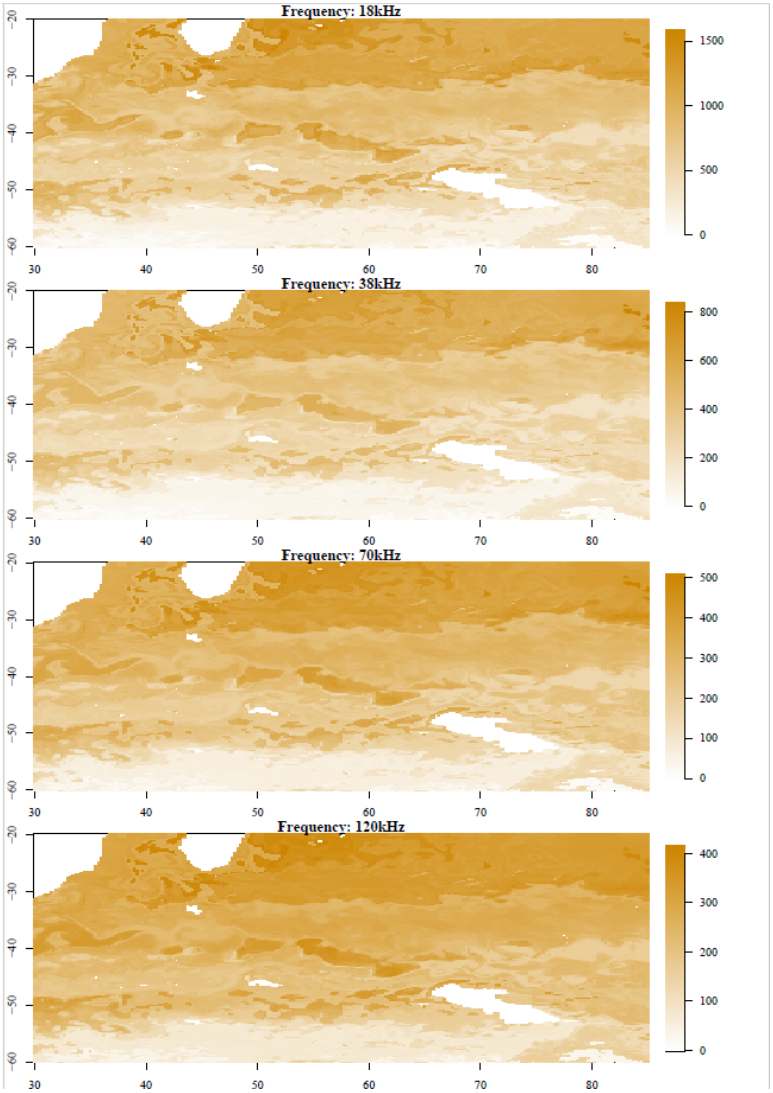


Figure 11. Nighttime water-column backscatter (Exemple for week 6 in 2020), represented as a proxy of animal biomass, represented in Nautical Area Scattering Coefficient (NASC, m^2 nmi^{-2}), integrated from 25 to 225 m depth, and measured at 38, 18, 70, and 120 kHz.

References

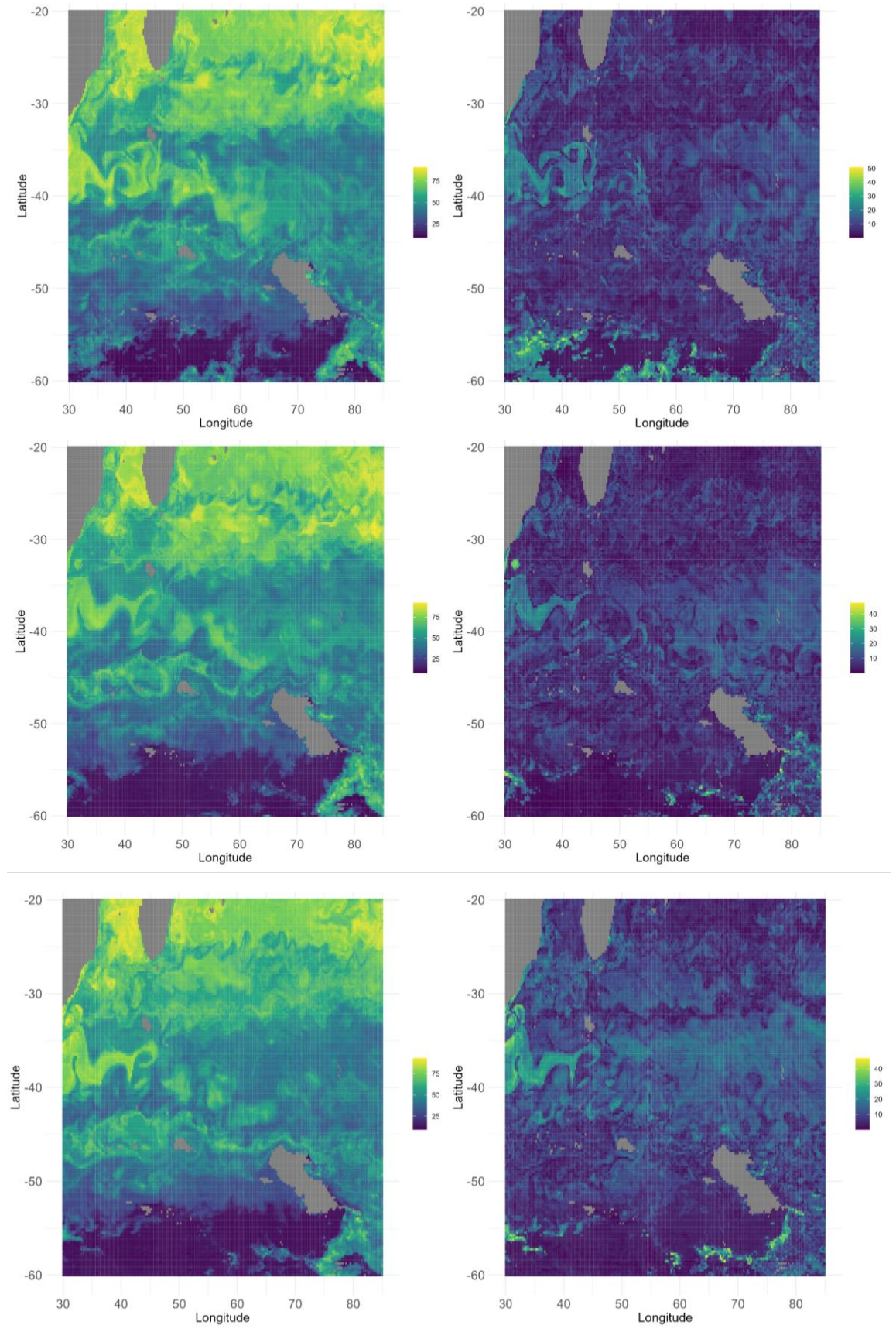
Godard, M., Manté, C., Guinet, C., Picard, B. and Nerini, D., 2020. Diving behavior of mirounga leonina: a functional data analysis approach. Frontiers in Marine Science, 7, p.595.

MacLennan, D.N., Fernandes, P.G. and Dalen, J., 2002. A consistent approach to definitions and symbols in fisheries acoustics. ICES Journal of Marine Science, 59(2), pp.365-369.

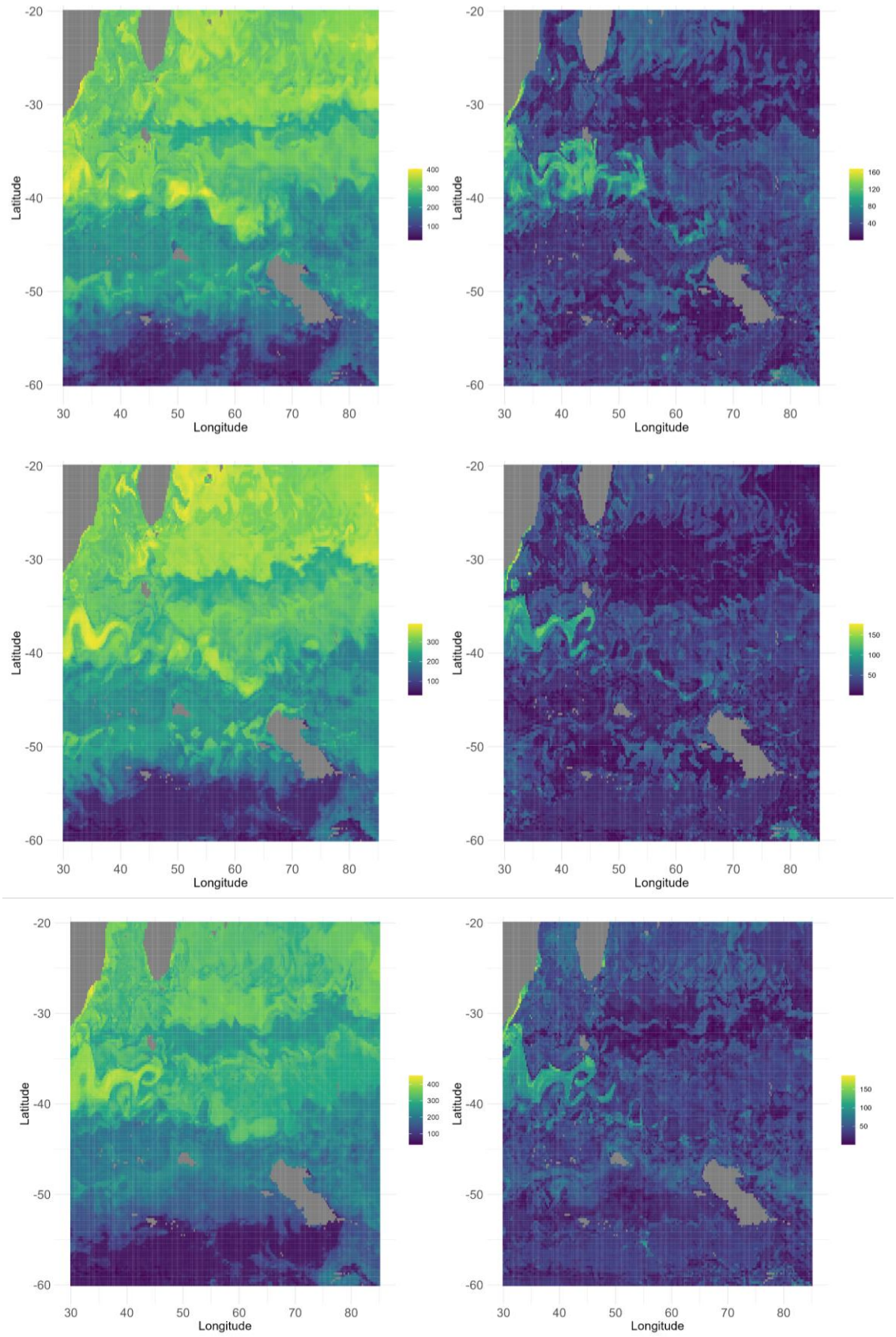
Park, Y.-H., Park, T., Kim, T.-W., Lee, S.-H., Hong, C.-S., Lee, J.-H., Rio, M.-H., Pujol, M.I., Ballarotta, M., Durand, I., Provost, C., 2019. Observations of the Antarctic Circumpolar Current over the Udintsev Fracture Zone, the narrowest choke point in the Southern Ocean. J. Geophys. Res. Oceans 124, 4511–4528. https://doi.org/ 10.1029/2019JC015024.

Pauthenet, E., Roquet, F., Madec, G., Guinet, C., Hindell, M., Mcmahon, C.R., Harcourt, R. and Nerini, D., 2018. Seasonal meandering of the Polar Front upstream of the Kerguelen Plateau. *Geophysical Research Letters*, *45*(18), pp.9774-9781.

Appendix 2 : Summer biomass prediction map for 2021

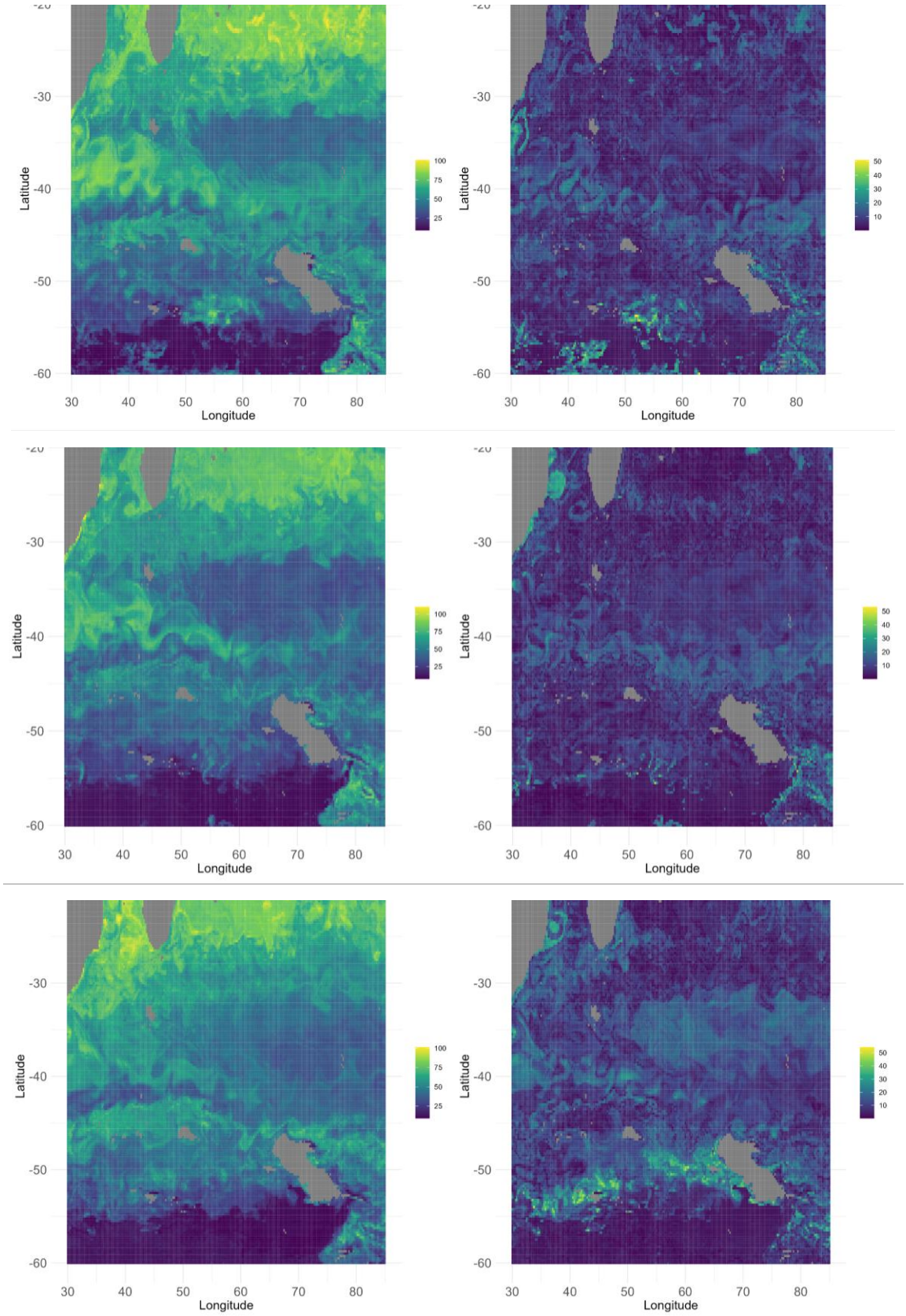


Day average NACS biomass prediction maps (left column) and standard deviation (right column) in January, February and March 2020 (first, second and third rows respectively) for the 120kHz frequency.

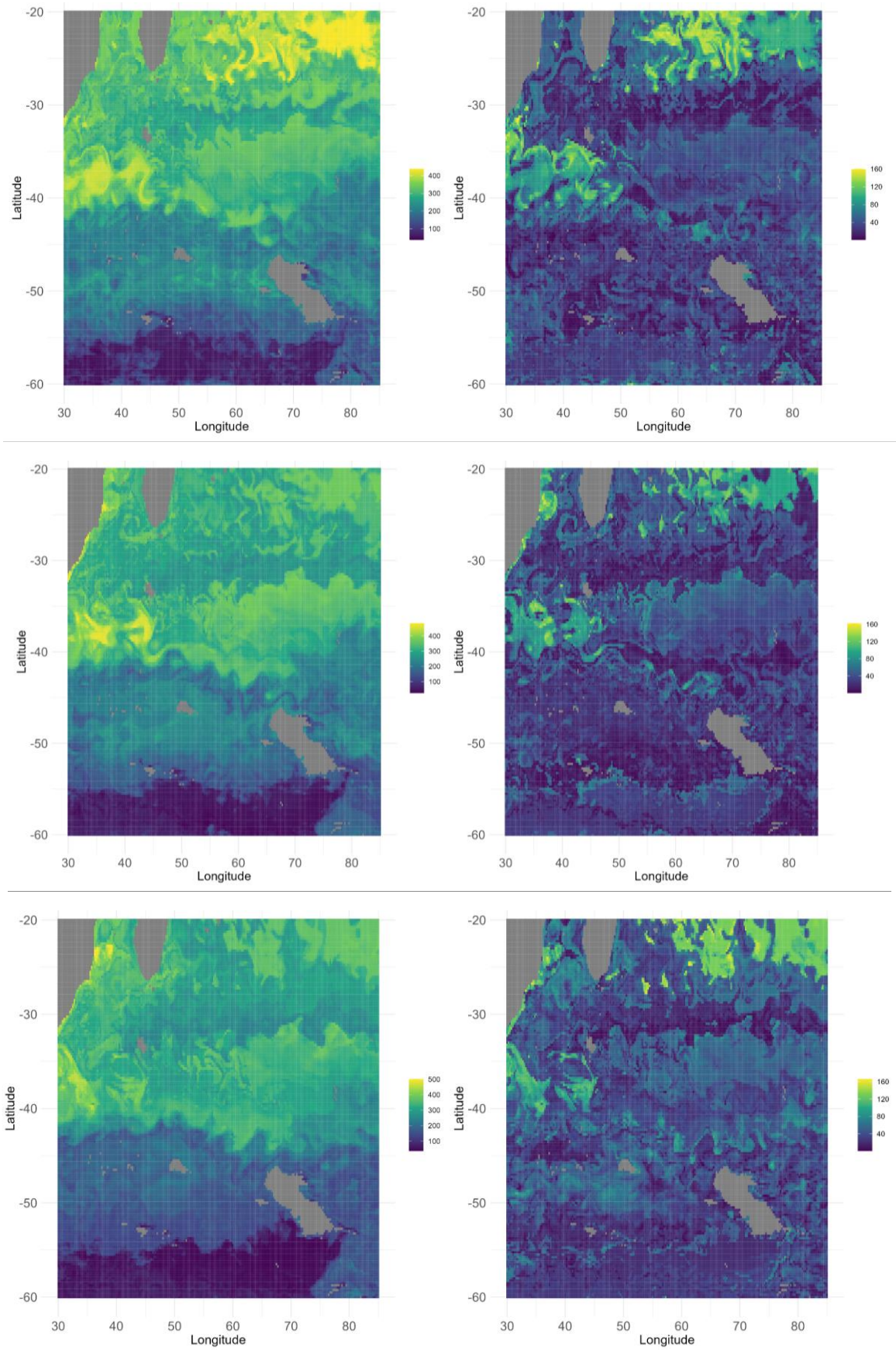


Night average NACS biomass prediction integrated from 0 to 225 m depth (a proxy of the overall crustacean abundance distribution) (left column) and standard deviation (right column) in January, February and March 2020 (first, second and third rows respectively) for the 120kHz frequency.

Appendix 3 : Summer biomass prediction map for 2021

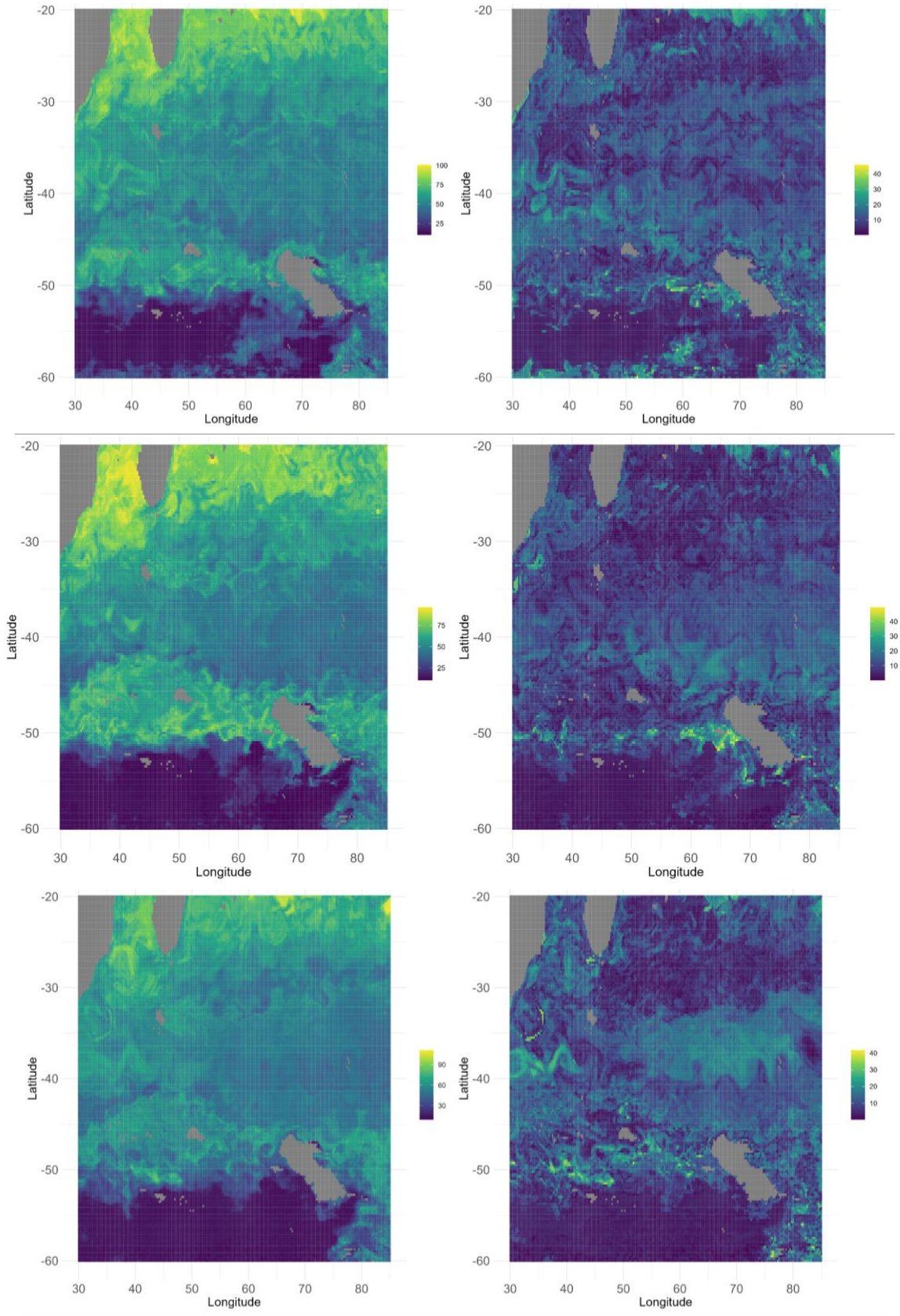


Day average NACS biomass prediction integrated from 0 to 225 m depth (a proxy of the overall crustacean abundance distribution) (left column) and standard deviation (right column) in January, February and March 2021 (first, second and third rows respectively) for the 120kHz frequency.

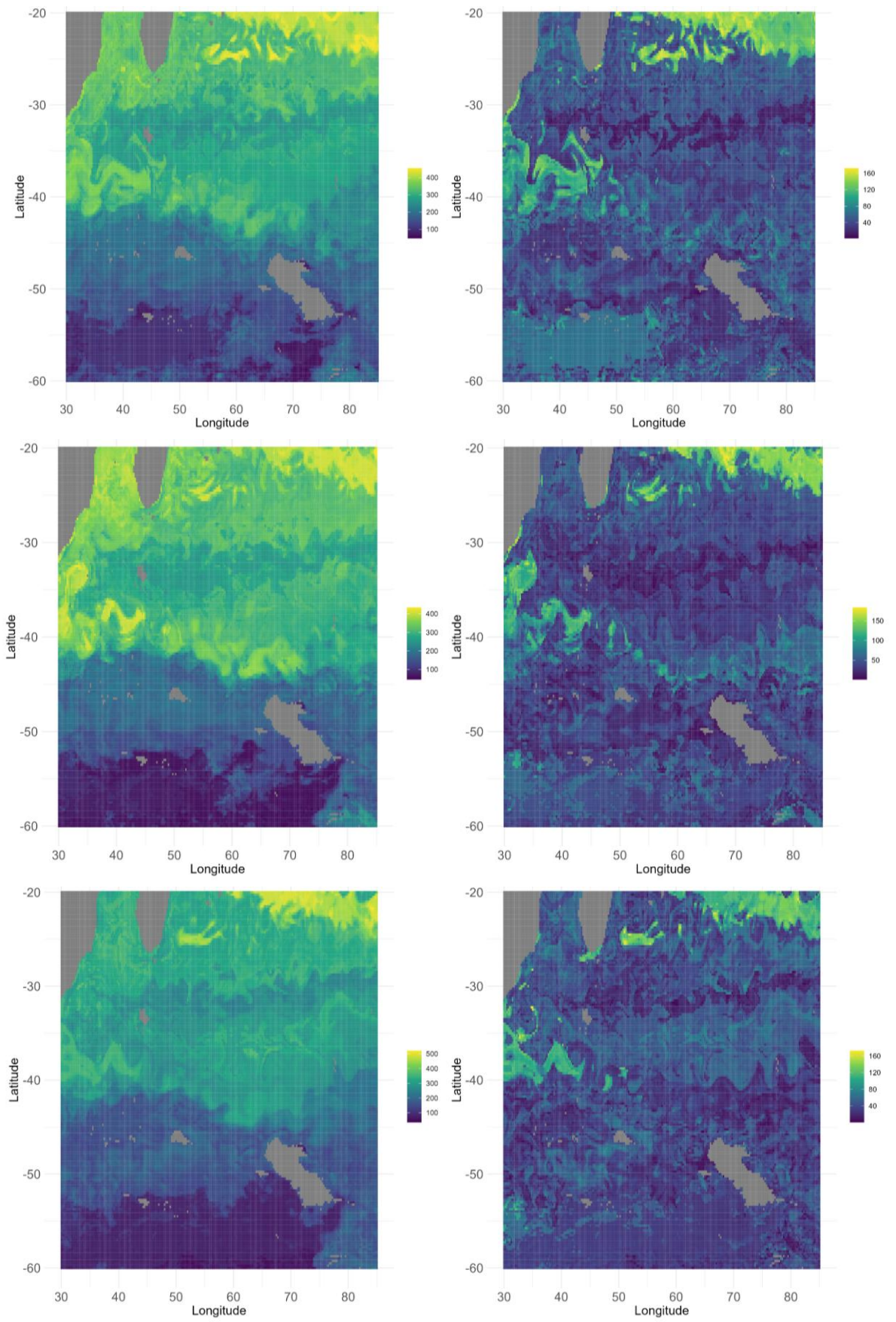


Night average NACS biomass prediction integrated from 0 to 225 m depth (a proxy of the overall crustacean abundance distribution) (left column) and standard deviation (right column) in January, February and March 2021 (first, second and third rows respectively) for the 120kHz frequency.

Appendix 4 : Summer biomass prediction map for 2022

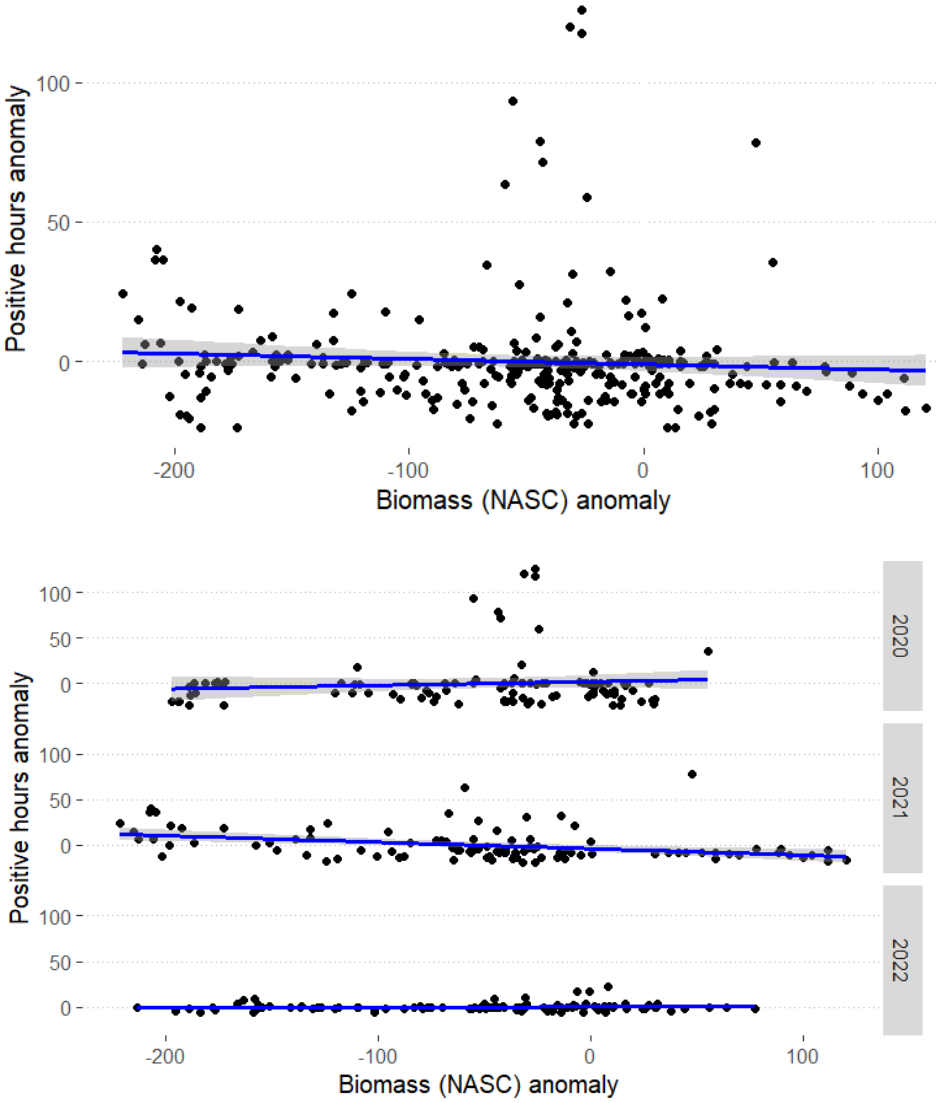


Day average NACS biomass prediction integrated from 0 to 225 m depth (a proxy of the overall crustacean abundance distribution) (left column) and standard deviation (right column) in January, February and March 2022 (first, second and third rows respectively) for the 120kHz frequency.



Night average NACS biomass prediction integrated from 0 to 225 m depth (a proxy of the overall crustacean abundance distribution)(left column) and standard deviation (right column) in January, February and March 2021 (first, second and third rows respectively) for the 120kHz frequency.

Appendix 5: Linear regression of NASC biomass anomaly and positive hours anomaly



Linear regression of NASC biomass anomaly (x axis) and positive hours anomaly (y axis), for all years (top graph) and by year (bottom graph).

COLLEGES BRETAGNE LOIRE | SCIENCES POUR L'INGENIEUR ET LE NUMERIQUE

ENSTA | INSTITUT POLYTECHNIQUE DE PARIS

Titre : Optimisation du suivi des cétacés par acoustique passive

Mot clés : cétacés, suivi par acoustique passive, standardisation, modélisation spatio-temporelle, duty cycle, océan indien

Résumé : Les cétacés sont exposés à une multitude de pressions anthropiques et environnementales. Dans ce contexte, leur suivi constitue un enjeu majeur pour mieux comprendre leur écologie, évaluer leur état de conservation et orienter les actions de gestion. Les méthodes visuelles, bien que structurées et comparables à grande échelle, restent limitées dans le temps et l'espace, en particulier pour les espèces discrètes ou évoluant dans des zones profondes ou isolées. Depuis les années 2000, le suivi par acoustique passive (PAM) s'est imposé comme une approche complémentaire, capable de détecter de manière autonome et non intrusive la présence acoustique des cétacés. Toutefois, le développement du PAM s'est accompagné d'une forte hétérogénéité des protocoles, ce qui limite la comparabilité des résultats et freine la mise en place d'un suivi standardisé à grande échelle. C'est dans ce cadre que cette thèse vise à proposer des outils et recommandations pour optimiser les protocoles PAM, renforcer la robustesse des indicateurs et améliorer l'interprétation écologique des données.

L'impact du sous-échantillonnage temporel sur la robustesse des indicateurs acoustiques a été évalué sur plusieurs espèces de cétacés. Les analyses montrent que la précision des résultats dépend du pourcentage de temps enregistré et de la durée des périodes d'écoute, en fonction du type de vocalisation, de la présence acoustique et de l'espèce étudiée. Ces travaux permettent de formuler des recommandations pratiques pour ajuster les protocoles selon les contraintes techniques et les caractéristiques biologiques des espèces ciblées.

Des données acoustiques pluriannuelles et multi-sites dans l'océan Indien ont été analysées pour caractériser les dynamiques spatio-temporelles des grandes baleines. L'utilisation de modèles additifs généralisés (GAMs) a permis d'identifier des zones écologiquement critiques, de relier la présence acoustique aux variables environnementales et de révéler des rythmes saisonniers et interannuels. Un suivi PAM long terme et standardisé permet ainsi d'obtenir des séries temporelles étendues, offrant la possibilité de passer de la description saisonnière de la présence acoustique à une meilleure compréhension des mécanismes écologiques sous-jacents à la distribution de ces espèces.

Ces travaux soulignent donc l'importance de définir des standards pour l'ensemble des choix méthodologiques d'un suivi par acoustique passive des cétacés. L'élaboration de protocoles harmonisés et documentés est nécessaire pour garantir des résultats comparables et exploitables pour la gestion et la conservation, tout en facilitant leur intégration future dans des cadres institutionnels et des coordinations européennes.

Title: Optimizing passive acoustic monitoring of cetaceans

Keywords: cetaceans, passive acoustic monitoring, duty cycle, standardization, spatio-temporal modeling, Indian Ocean

Abstract: Cetaceans are exposed to a wide range of anthropogenic and environmental pressures. In this context, their monitoring is a key challenge for improving our understanding of their ecology, assessing their conservation status, and guiding management actions. Visual survey methods, although structured and comparable at large scales, remain limited in temporal and spatial coverage, particularly for elusive species or those inhabiting deep or remote areas. Since the 2000s, passive acoustic monitoring (PAM) has emerged as a complementary approach, capable of autonomously and non-invasively detecting cetacean acoustic presence. However, the expansion of PAM has been accompanied by substantial heterogeneity in protocols, which constrains the comparability of results and hinders the establishment of standardized large-scale monitoring. Against this background, this thesis aims to provide tools and recommendations to optimize PAM protocols, enhance the robustness of indicators, and improve the ecological interpretation of acoustic data.

The influence of temporal sub-sampling on the robustness of acoustic indicators was assessed across several cetacean species. Analyses revealed that the accuracy of results depends on the proportion of recording effort and the duration of listening periods, with outcomes varying according to vocalization type, acoustic presence, and species studied. These findings enable the formulation of practical recommendations for tailoring protocols to technical constraints and the biological characteristics of the focal species.

Multi-year, multi-site acoustic datasets from the Indian Ocean were analyzed to characterize the spatio-temporal dynamics of large whales. The application of generalized additive models (GAMs) identified ecologically critical areas, linked acoustic presence to environmental variables, and revealed both seasonal and interannual patterns. Standardized long-term PAM programs therefore provide extended time series, enabling a shift from simple seasonal descriptions of acoustic presence to a deeper understanding of the ecological mechanisms underlying species distributions.

So these findings highlight the importance of defining standards for all methodological choices in passive acoustic monitoring of cetaceans. The development of harmonized and well-documented protocols is necessary to ensure results that are both comparable and usable for management and conservation, while also facilitating their future integration into institutional frameworks and European coordination.