

Distribution, déplacements et utilisation de l'habitat des différents clans du Grand Cachalot dans les Petites Antilles

Mémoire présenté pour l'obtention du diplôme d'Ingénieur de l'ENGEES



Ecole : Ecole Nationale du Génie de l'Eau et de l'Environnement de Strasbourg

Organisme d'accueil : Caribbean Cetacean Society

Encadrant : Jeffrey Bernus (Directeur de la CCS)

Louise SIMON
09/08/2023



Résumé

Cette étude complète les données existantes du Dominican Sperm Whale Project et de Felicia Vachon, élargissant la zone d'étude d'Anguilla à Grenade sur la période 2021-2022. Elle dévoile 62 individus dont 45 déjà connus et 17 nouveaux individus, potentiels nouveaux membres d'unités sociales déjà existantes ou membres d'unités non encore décrites. La majorité de ces observations se concentrent au niveau de la Guadeloupe, de la Dominique et de Saint-Vincent-et-les-Grenadines. Cette étude confirme le caractère sédentaire des cachalots des Petites Antilles en précisant les déplacements de plus grande ampleur des individus, des unités sociales et des clans vocaux le long de l'arc antillais. L'analyse acoustique met en lumière des zones de répartition distinctes entre les différents clans. En effet, le clan EC1 est majoritairement présent autour des îles de la zone Centre, alors que les clans EC2 et EC3 occupent plutôt la zone Sud, de Sainte-Lucie à Grenade. L'établissement de ces clans vocaux au niveau de zones précises semble également se retrouver à l'échelle des unités sociales, dont les domaines vitaux distincts sont relativement restreints. Enfin, un modèle additif généralisé a été déployé à partir des observations visuelles. Ce dernier indique que la distribution des cachalots semble être majoritairement influencée par quatre facteurs que sont la distance à la côte, la chlorophylle a, la bathymétrie et la température des eaux de surface. En outre, il révèle une zone à forte densité d'individus au niveau de la côte Atlantique de la Guadeloupe, encore non échantillonnée. Ces résultats soulignent l'importance de la collaboration entre les îles des Petites Antilles pour la conservation et la protection de cette population de cachalots unique et dont la survie est menacée.

Abstract

This study completes existing data of Dominican Sperm Whale Project and Felicia Vachon by enlarging the study area from Anguilla to Grenade over the period 2021-2022. It emphasizes 62 individuals whose 45 already known as well as 17 new individuals. These latter are potential new members of existing social units or part of unknown social units. Most of the observations focus on the Guadeloupe, Dominica and Saint-Vincent-and-the-Grenadines area. This project underscores the sedentary behavior of the sperm whales of the Lesser Antilles and precise the long-range displacements of individuals, social units and vocal clans. The acoustic analysis confirms the repartition area of the clans. As a matter of fact, EC1 clan is mainly present around islands of center area whereas EC2 and EC3 clans occupy south region from Saint-Lucia to Grenade. We retrieve this characteristic at social unit scale, whose core areas are particularly restricted. Finally, a Generalized Additive Model has been calibrated on visual observations. It indicates that the sperm whale's distribution seems to be mainly associated with four predictors such as the coast distance, the chlorophyl a, the bathymetry as well as the sea surface temperature. It highlights a region with high density of individuals not yet sampled on the Atlantic side of Guadeloupe. These results emphasize collaboration importance between the Lesser Antilles islands for conservation and protection of this unique and threatened sperm whale's population.

Remerciements

Merci à tous les partenaires de la CCS qui ont rendu cette aventure possible et plus particulièrement à WWF Pays-Bas, à la Collectivité Territoriale de Martinique, au Parc Naturel Régional de Martinique, à la Région Guadeloupe, à la Dutch Caribbean Nature Alliance, à Corail Caraïbes, à la Fondation Groupe EDF, à la Société Anonyme de Raffinerie des Antilles, à la Blue Marine Foundation, à l'Animal Welfare Institute, au ministère de l'Éducation nationale et de la jeunesse, au Museum national d'Histoire naturelle et au CAR.SPAW.RAC pour leurs soutiens financiers. Merci à Hal Whitehead et à l'équipe de Flukebook pour leur aide à l'utilisation de la plateforme. Un grand merci à Shane Gero (DSWP) et à Felicia Vachon pour leurs données, leur temps mais surtout, leurs précieuses expertises.

Je remercie toute l'équipe de la CCS, Jeff, Valentin et Rocío, pour leur aide et leur bonne humeur journalière qui m'a permis d'évoluer dans des conditions de travail plus qu'agréables. Un grand merci également pour l'expédition qui m'a permis de découvrir des personnes plus inspirantes les unes que les autres et de revenir en France avec des étoiles pleins les yeux et un bronzage de qualité ! Big up à Val, chef d'expédition en or, merci pour ton calme, ton écoute et tes supers conseils, merci d'avoir rendu cette expédition ce qu'elle a été, des moments de qualité avant tout, toujours dans la bonne humeur collective.

Gros merci à ma super coloc pour ces quelques mois de folie, pour le carnaval et toutes les soirées organisées, pour la playlist de la coloc des glingos, pour toutes nos balades, nos après-midi chill à la plage, nos parties de carte endiablées et surtout nos moments dessins animés et drames en mode pleureur en bande organisée !

Merci à Victor pour les highlights culturels, les conseils sur R et les escapades en forêt impétueuses mais infructueuses (toujours à la recherche de la Grande Sophie, snif). Merci à Salomé pour tes convictions, tes mots apaisants, ta patience au volant et tes fabuleux dessins (RIP Basil). Merci à Rebeca, d'avoir été une personne avec des intentions si pure et une si bonne amie, toujours à l'écoute. Merci pour tes rires sincères m'encourageant à continuer mes blagues les plus nulles et pour nos moments précieux de confession à deux. Merci pour tes mots rigolos et d'avoir toujours été à fond pour en apprendre plus sur les merveilles de la langue française. Merci à ma Soso pour nos moments sous l'eau de jour ou de nuit, merci d'avoir supporté mes bizarreries nocturnes, merci d'avoir veillé sur moi dans les moments plus mouvementés, merci pour nos aventures culinaires presque toujours réussies, merci pour tous les rires et tout simplement, merci d'être la personne que tu es, un cœur de pierre en surface mais un petit moelleux en profondeur. Merci à Lola d'avoir toujours été là pour m'accompagner dans nos délires à gogo, de jpp repose en paix, première soirée, aux inlassables répliques du tierquar, en passant par le merveilleux duo Gina et Joshua. T'es le sang, tu peux venir à Marseille sans problème maintenant. Et merci pour la vaisselle parce qu'on ne l'a pas assez dit. Merci à Lucas, pour nos aventures passées et à venir, pour nos moments sérieux que t'avais toutes déjà vues, pour nos longues discussions nocturnes sur tout et rien, pour ton écoute attentive et tes précieux conseils au boulot comme dans la vie, pour tes mèmes et tes blagues sans arrêt. J'ai trop hâte de nos prochaines aventures.

Je remercie toute ma petite famille qui s'élargit, pour leur rayonnement téléphonique international qui réussit toujours à réchauffer mon cœur. Merci à mes parents pour leur dévouement à la relecture de ce mémoire, merci à mon frère pour ses explications impeccables, merci à ma jums Lil pour son soutien incommensurable et sa visite à mi-chemin pour recharger mes batteries. Toujours là, peu importe la distance, peu importe le moment et pour toujours. Je vous aime.

Table des matières

I-	Contexte général de l'étude	9
1-	Présentation de l'organisation et de l'espèce étudiée.....	9
1.1	Présentation de la CCS	9
1.2	Présentation de l'espèce étudiée : le Grand cachalot.....	9
2-	Problématique.....	12
2.1	Etat de l'art	12
2.2	Objectifs de la présente étude	13
2.3	Pressions & Enjeux	14
II-	Présentation et mise en œuvre méthodologiques :.....	17
1-	Zone d'étude & Collecte des données	17
2-	Méthodologies associées à l'identification visuelle et la détection acoustique	21
2.1	Processus d'identification via Flukebook	21
2.2	Processus d'identification acoustique via Audacity	24
2.3	Détermination de l'appartenance de chaque individu à un clan vocal.....	26
2.4	Détermination des relations entre individus.....	26
2.5	Evaluation de l'abondance de la population.....	27
2.6	Détermination des déplacements des unités sociales	28
3-	Création d'un modèle de distribution d'espèce de type GAM (Generalized Additive Model) .	28
3.1	Description du modèle additif généralisé	28
3.2	Acquisition des données.....	29
3.3	Prétraitement des données.....	31
3.4	Calibration	34
3.5	Validation	35
III-	Résultats	36
1-	Identification à l'échelle de l'individu, détermination de l'abondance et de la distribution....	36
1.1	Photo-identification.....	36
1.2	Détermination de l'abondance	38
2-	Identification des liens sociaux	40
3-	Identification des clans vocaux & aires de répartition.....	43
4-	Détermination des déplacements.....	46
4.1	Détermination des déplacements à l'échelle individuelle	46
4.2	Détermination des déplacements à l'échelle des unités sociales & clans vocaux	47
5-	Modélisation de la distribution des cachalots	48
5.1	Modélisation & Quantification des erreurs.....	49
5.2	Validation	51

IV-	Discussion	52
1-	Une organisation sociale ordonnée	52
1.1	Le regroupement d'unités sociales ensemble : description de nouveaux groupes ?	52
1.2	Les nouveaux individus : nouvelles unités sociales ou nouveaux individus d'unités déjà décrites ?	53
2-	Des nomades de l'océan ?.....	55
3-	Des points chauds où se rassemblent les cachalots ?.....	55
3.1	Des facteurs environnementaux définissant des habitats préférentiels	55
3.2	Une densité plus importante du côté Caraïbe ?.....	56
3.3	Une probabilité d'observer des cachalots plus importante dans les régions Centre et Sud ..	57
4-	La non-mixité entre clans est-elle réelle ?	57
4.1	Une répartition des clans vocaux bien définie	57
4.2	Présence simultanée de deux clans : Rencontre entre clans vocaux ou apprentissage social ?	58
5-	L'appartenance à un clan vocal, un facteur déterminant dans la distribution ?	59
6-	Limitations et pistes d'améliorations	61
V-	Conclusion	66
VI-	Bibliographie.....	68
VII-	Annexes	73
	Annexe 1 : Description des enregistrements totaux, exploitables et de ceux où étaient présents des codas	73
	Annexe 2 : Liste des produits Copernicus desquelles les variables sont extraites (Annexe 2.A) et explicitation du choix des paramètres moyennes et variances pour chaque variable (Annexe 2.B)	74
	Annexe 3 : Explicitation du choix de chaque variable environnementale utilisée dans le modèle ..	75
	Annexe 4 : Implémentation d'une variable environnementale sur la grille hexagonale	76
	Annexe 5 : Tableau récapitulatif de tous les individus capturés par la CCS, de la date de la dernière recapture par la CCS et par DSWP (pour les individus de DSWP recapturés par la CCS), ainsi que le nombre de captures total pour chaque individu (dont entre parenthèses le nombre de recapturés par la CCS).....	77
	Annexe 6 : Individus capturés et recapturés.....	78
	Annexe 7 : Présentation des différentes unités sociales rencontrées par la CCS, du nombre de fois où elles ont été vues, ainsi que du nombre d'individus dans cette unité.....	79
	Annexe 8 : Planches photos des unités sociales/groupements les plus vus par la CCS.....	79
	Annexe 9 : Matrice des corrélations des variables environnementales utilisées dans le modèle GAM.....	82
	Annexe 10 : Courbe ROC représentant l'évolution de la sensibilité (taux de vrais positifs) en fonction de 1-spécificité (taux de faux positifs) lorsque le seuil t varie (au-dessus duquel la prédiction est positive et en-dessous duquel la prédiction est négative).	82

Annexe 11 : Arbre de régression multivarié en fonction des unités sociales et des 4 prédicteurs du meilleur modèle	83
--	----

Table des figures & tableaux

Figure 1: Figure issue de Andreas et al. (2022) avec a : sac frontal, b : « case » (enveloppe musculaire), c : l'organe du spermaceti, d : passage nasal gauche, e : passage nasal droit, f : compartiment « junk », g : l'évent, h : sac distal, i : museau du singe (ou lèvres phoniques). Les flèches représentent le processus d'émissions de sons.....	10
Figure 2: Détail de l'organisation sociale des cachalots.....	11
Figure 3: Aires marines protégées des sanctuaires Agoa (orange) et Yarari (vert) (Adaptée de https://www.ofb.gouv.fr).....	13
Figure 4: Présentation des objectifs de l'étude (DSWP : Dominican Sperm Whale Project ; CMR : Capture, Marquage, Recapture ; CCS : Caribbean Cetacean Society). Les ronds verts représentent les questionnements auxquels on cherchera à répondre. Les ronds rouges représentent les données utilisées pour chacun des trois axes.....	14
Figure 5: Echouage d'un Grand Cachalot à l'Anse Coulevre dans les années 60 (données Sepenmar)	16
Figure 6: Zone d'étude de "Ti Whale An Nou"	17
Figure 7: Présentation des caractéristiques géomorphologiques (A) et des courants océaniques (B) au niveau de la zone d'étude (Annexe 2 pour la source des données utilisées). Les flèches présentées en figure B sont purement schématisées pour montrer le sens des courants dans la zone. La légende représente les valeurs prises par la composante Est de ces courants.....	18
Figure 8: Effort d'échantillonnage en 2021 & 2022 (A, en rouge) et organisation à bord montrant la position des observateurs et de l'hydrophone câblé (B)	19
Figure 9: Exemple de résultats proposés par l'algorithme PIE de Flukebook, à l'issue de la phase d'identification	23
Figure 10: Exemple de fonctionnement de l'algorithme CurvRank. La ligne jaune représente la courbe tracée par l'algorithme à partir des irrégularités de la nageoire caudale.....	23
Figure 11: Points d'observations en 2021-2022 (A) et densité d'observations en fonction de la distance à laquelle l'animal a été détecté (B). La ligne orange en pointillé indique un maximum de cachalots détectés autour de 512 m. Les différentes zones colorées représentent les différents buffers et leurs distances associées	31
Figure 12: Effort d'observation du cachalot.....	32
Figure 13: Représentation schématisée des étapes nécessaires à la prédiction de la distribution suivant un modèle GAM.....	33
Figure 14: Planche photos présentant les 17 nouveaux individus découverts par la CCS.....	37
Figure 15: Carte représentant le nombre d'individus différents aperçus par île (ronds oranges)	38
Figure 16: Carte représentant les individus observés par Vachon (2022) et capturés ou non par la CCS (respectivement points bleus et oranges) (A) et carte zoomée sur les points d'observations des unités A,N, G12 et K en Dominique et Martinique (B).....	40
Figure 17: Dendrogramme présentant les différents regroupements possibles des individus et les unités sociales déjà connues	42
Figure 18: Extraits de spectrogrammes faisant apparaître les deux types de codas EC1 (1+1+3) & EC2 (5R)	44
Figure 19: Distribution des différents clans vocaux sur la zone d'étude avec données CCS (A) et données Vachon (2022) (B).....	45

Figure 20: Observations & recaptures des individus de la CCS (données CCS, Vachon, DSWP).....	46
Figure 21: Déplacements de plus de 100km de 8 individus appartenant à des unités sociales et des groupes vocaux différents (données CCS, Vachon, DSWP).....	47
Figure 22: Déplacements des différentes unités sociales en fonction de leur clan vocal (données CCS, Vachon, DSWP).....	48
Figure 23: Relation entre espérance de la variable réponse et chacune des variables explicatives pour les deux premiers modèles (à gauche, modèle 1 et à droite, modèle 2). L'histogramme en noir en bas des figures représente la densité d'observations dans l'espace des variables explicatives. Les tirets rouges représentent les valeurs observées du prédicteur.	50
Figure 24: Cartes représentant la prédiction de la distribution des cachalots selon les quatre prédicteurs choisis (à gauche). Les coefficients de variations reliées à ces prédictions sont présentés à droite.....	51
Figure 25: Planche photos des individus 14,15,17 & 18, possiblement nouvelle unité sociale.....	54
Tableau 1: Expéditions et dates d'observations des cachalots.....	21
Tableau 2: Tableau des individus de DSWP recapturés par la CCS.....	36
Tableau 3: Estimation de l'abondance via deux méthodes avec les données de la CCS uniquement, les données de Vachon (2022), et les deux réunions.....	39
Tableau 4: Comparaison des individus totaux identifiés et recapturés en fixant une durée minimale entre deux recaptures ou en laissant cette durée libre.....	39
Tableau 5: Regroupement des individus en cluster et attribution potentielle à une unité sociale déjà connue. « no » représente le nombre d'occurrences sur lequel se base le regroupement en cluster. Les groupements marqués en bleu dans le tableau ci-dessus sont ceux pour lesquels les individus regroupés n'ont été vus ensemble qu'une seule fois. Les groupements en vert représentent ceux pour lesquels les eigenvalues étaient égales ou inférieures à 0. Les individus marqués en orange sont les nouveaux individus appartenant potentiellement à une unité déjà décrite.....	41
Tableau 6: Comparaison du résultat de deux méthodes de clustering.....	42
Tableau 7: Détermination de l'appartenance potentielle des nouveaux individus à un groupement/unité sociale. Les lettres colorées en orange constituent les groupes potentiels auxquels seraient rattachés les nouveaux individus. « ni » représente le nombre de fois où l'individu a été vu en 2021, et en 2022.....	43
Tableau 8: Descriptif des enregistrements où deux clans ont été entendus : durée totale d'enregistrement & (nombre de codas associé à cette durée entre parenthèses).....	44
Tableau 9: Nombre d'occurrences de chaque clan vocal à partir des données ayant servi à faire la Figure 22.....	47
Tableau 10: Deux premiers modèles expliquant le mieux la distribution des cachalots.....	49
Tableau 11: Matrice de confusion réalisée à partir du jeu de validation.....	51

I- Contexte général de l'étude

1- Présentation de l'organisation et de l'espèce étudiée

1.1 Présentation de la CCS

La Caribbean Cetacean Society (CCS) est une association à but non-lucratif, qui a été cofondée par Laura Pittino et Jeffrey Bernus en 2020 et dont l'objectif principal est d'assurer la conservation des cétacés dans la Caraïbe, via la coopération inter-îles de différents types d'acteurs. Elle se base sur un constat simple : les cétacés ne connaissant pas de frontières, ils devraient donc pouvoir se déplacer d'une île à l'autre, sans risquer un changement de politique de conservation soudain, menaçant alors la prospérité des espèces en question. C'est dans ce contexte que la CCS a créé le programme Ti Whale An Nou en 2021 afin de réaliser des campagnes d'échantillonnages dans l'ensemble des îles des Petites Antilles, en suivant un protocole standardisé. Ces objectifs sont multiples : collaborer pour estimer la taille des populations, les déplacements, la distribution des différentes espèces en fonction des variables environnementales ; contribuer à l'alimentation des bases de données acoustiques pour permettre une meilleure identification des espèces ; renforcer la coopération du réseau Caribéen, accroître l'attractivité du territoire, favoriser le développement de l'économie bleue, définir l'importance des différentes menaces sur les cétacés localement pour pouvoir agir concrètement. A long terme, la CCS aspire à la mise en place de mesures adéquates de protection avec les gouvernements et à la création d'Aires Marines Protégées pour que les cétacés puissent être protégés efficacement contre l'ensemble des menaces anthropiques et sur l'ensemble des îles qu'ils traversent.

1.2 Présentation de l'espèce étudiée : le Grand cachalot

a- Généralités

Les cétacés sont usuellement décrits comme l'ensemble des dauphins, baleines et marsouins. On distingue d'ailleurs les cétacés à fanons (mysticètes) et les cétacés à dents (odontocètes). Le cachalot (*Physeter macrocephalus* ; Linnaeus 1758) représente le plus grand des odontocètes et bien que longtemps l'objet de débats, il est l'unique représentant de son genre. Il présente un dimorphisme sexuel très marqué aussi bien par sa taille que par son poids, avec des femelles pesant entre 10 et 20 tonnes et mesurant environ 11 m de long alors que les mâles peuvent atteindre 19.2 m et 70 tonnes (Carwardine et al., 2020). Pour les deux sexes, l'espèce est reconnaissable par son imposante tête carrée en forme de tronc (*Physeter macrocephalus* du grec *Physeter makros* et *kephale*, littéralement « le souffleur à grosse tête ») qui occupe jusqu'à un tiers de la longueur totale de son corps (Cranford, 1999). Cette tête abrite l'organe du spermaceti, qui lui permet notamment d'accéder à une grande variété de ressources, en dépensant le moins d'énergie pour cela. En effet, c'est sa tête massive et plus particulièrement, la solidification du spermaceti ; liquide gras composé majoritairement de cires et de triglycérides; qui lui permet de plonger rapidement, jusqu'à 1200 m pour des durées variant de 40 à 50 minutes (Cantor et al., 2019). Lors de ces épisodes de recherche de nourriture, qui occupent 75% de son temps (Whitehead & Weilgart, 1991), il chasse principalement des proies mésopélagiques telles que des céphalopodes (Waring, 2004), et affectionne particulièrement les calmars d'eaux profondes, qu'il parviendrait à attirer grâce à la coloration blanche de sa mâchoire inférieure (Cantor et al., 2019).

b- Un répertoire acoustique déterminant l'organisation sociale

Le cachalot utilise divers types de clics adaptés à différentes situations. Lorsqu'il chasse en profondeur, il repère ses proies par l'intermédiaire de trains de clics réguliers, dont l'intervalle se réduit à mesure

qu'il se rapproche de ces dernières, pour former des buzz, aussi appelés « creaks » (Miller et al., 2004). Le son est émis par le museau du singe (i) et est concentré jusqu'à l'avant de la tête en traversant le spermaceti (c). C'est la connexion entre l'organe du spermaceti et le compartiment junk (f) (constitué de nombreux corps aqueux agissant comme des lentilles acoustiques) qui permet au front d'onde d'être amplifié puis propagé dans l'eau (Figure 1, Andreas et al., 2022). Ce processus particulier d'émissions de sons fait du cachalot l'espèce ayant le plus gros bio-sonar sur Terre (Cantor et al., 2019).

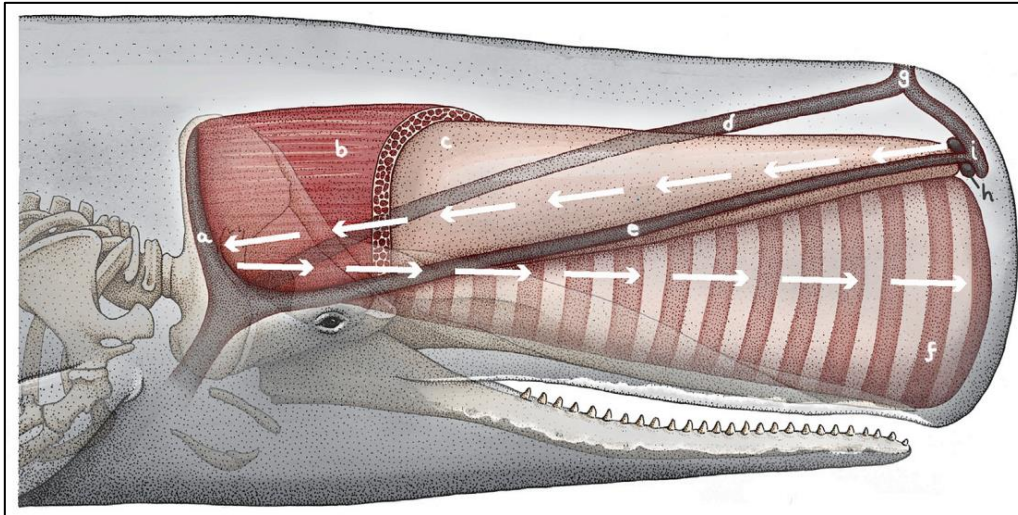


Figure 1: Figure issue de Andreas et al. (2022) avec a : sac frontal, b : « case » (enveloppe musculaire), c : l'organe du spermaceti, d : passage nasal gauche, e : passage nasal droit, f : compartiment « junk »¹, g : l'évent, h : sac distal, i : museau du singe (ou lèvres phoniques). Les flèches représentent le processus d'émissions de sons.

Il existe également des clics dont le rôle est davantage social comme les codas (Watkins, 1977), contenant entre 3 à 20 clics et durant entre 0.2 à 2 s, qui sont fortement utilisés dans les phases d'interactions sociales en surface (Whitehead & Weilgart, 1991); les clics lents (ou clangs), répétés toutes les 6 à 8 s, employés par les mâles pour assurer leur succès reproducteur notamment (Whitehead, 2003); ou encore, les « squeals », ces courtes rafales de clics répétés rapidement prenant ainsi la forme d'un miaulement, et qui sont majoritairement associés à des interactions sociales proches ou à un état d'excitation élevé (Weir et al., 2007). L'ensemble de ces sons est émis entre 2 et 40 kHz (Cantor et al., 2019).

La complexité du système de communication observé chez cette espèce révèle des capacités cognitives particulièrement développées, à l'origine d'une organisation sociale hiérarchisée, qui détermine l'appartenance de chaque individu à une unité familiale ou clan social. Ces clans sociaux, dont l'effectif atteint parfois une vingtaine d'individus (Lyrholm et al., 1999 ; Konrad et al., 2018), s'unissent souvent et temporairement avec d'autres clans pour former des groupes, afin de se protéger des prédateurs ou pour chasser par exemple, faisant alors partie d'un même « clan vocal » (L. E. Rendell & Whitehead, 2003). Ces dialectes bien particuliers sont associés à des comportements sociaux, de chasse et de mouvements spécifiques qui sont propres à chaque sous-population (Whitehead, 2018). Cela explique alors que, comme pour les humains, des populations géographiquement éloignées comme celles d'Atlantique et du Pacifique ne puissent pas se comprendre (Whitehead et al., 2012). Cette organisation sociale est détaillée en Figure 2.

¹ Apparenté au melon des autres cétacés

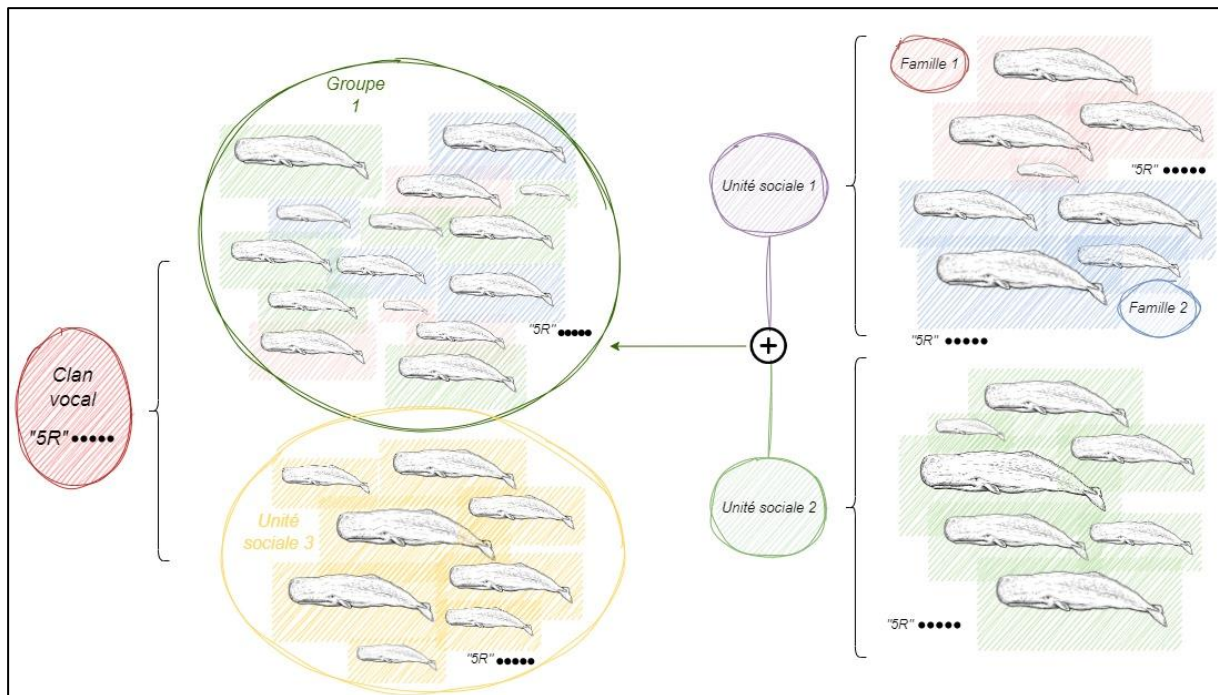


Figure 2: Détail de l'organisation sociale des cachalots

c- Entre répartition mondiale et résidence insulaire : une espèce vulnérable

L'ensemble de ces sous-populations et de leurs associations temporaires ont valu au cachalot son surnom de « nomade de l'océan » (Mizroch & Rice, 2013). En effet, ce dernier est globalement présent dans tous les océans du monde, des tropiques aux abords des hémisphères, parcourant ainsi des milliers de kilomètres par an. Il habite les eaux riches en production primaire, généralement associées aux événements d'upwelling océaniques mais également toutes les zones dont les caractéristiques bathymétriques et océanographiques concentrent les proies, telles que les canyons ou monts sous-marins (Hann et al., 2016; Moors-Murphy, 2014), les pentes escarpées des plateaux continentaux ou encore les zones de fronts thermiques et de tourbillons (Praca et al., 2009; Virgili et al., 2019).

Cependant, de pair avec le dimorphisme sexuel marqué, une forte ségrégation spatiale existe entre les deux sexes. En outre, les femelles occupent majoritairement les eaux tropicales et subtropicales situées à des latitudes inférieures à 40° (Carwardine et al., 2020) et seraient plus sédentaires que les mâles qui migrent des hautes latitudes vers les basses latitudes. Parallèlement à leur comportement migrateur, les mâles passent la majorité de leur vie seuls, sauf lorsqu'ils sont sexuellement immatures, période qui peut s'étendre jusqu'à 21 ans (Whitehead, 2018).

Pendant cette période, ils appartiennent aux unités familiales matrilineaires, principalement composées de femelles, de veaux (jeunes de l'année) et de juvéniles. Au sein de ces unités, un soin particulier est apporté à la descendance, notamment par le biais des soins allo-parentaux (Gero et al., 2009). En outre, certaines femelles sont « nourrices » et peuvent allaiter des nouveau-nés qui ne sont pas les leurs ; d'autres adoptent le statut de « baby-sitters », surveillant la progéniture restée à la surface lors des parties de chasse des individus plus âgés et d'autres encore présentent les deux types de comportements (Gero et al., 2009). Les « veaux » seraient donc des facteurs importants dans la construction d'une unité sociale, agissant comme ciment des relations entre les femelles (Gero et al., 2013). Les longues périodes de gestation et d'allaitement ainsi que la basse fréquence de mise-bas rendent la survie de chaque veau essentielle à celle plus globale de l'espèce (Whitehead 2003).

2- Problématique

2.1 Etat de l'art

Dans la mer des Caraïbes, Gero et al. (2007) ont prouvé l'existence d'une population de grands cachalots isolée, bien distincte par exemple de celle de la mer des Sargasses ou encore de celle du Golfe du Mexique (Engelhaupt et al. 2009). Cette étude a notamment permis de réaliser une première estimation de la population de cachalots dans la Caraïbe, grâce aux données collectées par plusieurs organisations sur 22 ans (1984 à 2006) dont l'association Evasion Tropicale (Guadeloupe, Caroline & Renato Rinaldi, 2000-2003) et la Dalhousie University (Dominique, Shane Gero, 2005-2006). Au total : 194 individus ont été identifiés au niveau de 5 îles des Petites Antilles, à savoir : la Dominique, la Guadeloupe, la Grenade, Sainte-Lucie et la Martinique.

Dans les années suivant cette estimation, plusieurs études se sont intéressées à la dynamique de la population de cachalots dans les Petites Antilles. Si, au premier abord, la population de cachalots peut sembler en bonne santé (taux de croissance de $+3.4\% \text{ yr}^{-1}$), le déséquilibre entre fécondité et mortalité fait des Petites Antilles un piège écologique (Battin, 2004) pour cette espèce. Les eaux caribéennes agissent donc comme un puits attractif pour les cachalots, grâce à l'abondance de leurs ressources, mais cet effet est contrecarré par les caractéristiques négatives du cycle vital (mortalité > fécondité ; taux de croissance projetée de $-2.7\% \text{ yr}^{-1}$) (Whitehead & Gero, 2015). Depuis, le déclin de cette population a été confirmé de nombreuses fois, en Dominique (2010-2015) et en Guadeloupe (2001-2013), avec des taux respectivement de 4.5 et 6.2%/an ((Gero & Whitehead, 2016) (Rinaldi et al. (2021)).

Au cours des deux dernières décennies, les efforts collectifs ont permis d'avoir une meilleure idée de la dynamique mais également des caractéristiques sociales de cette population (Gero et al., 2014). La continuité des efforts dans le temps a notamment permis d'augmenter le nombre d'individus photo-identifiés à 536. Cependant, la disparité des données collectées, principalement en Dominique et en Guadeloupe, pourrait introduire un biais dans ces premières observations, notamment en sous-estimant la taille et l'étendue de la population caribéenne. C'est pourquoi, Vachon et al. ((Vachon, Hersh, et al., 2022), (Vachon, Eguiguren, et al., 2022)) en reconnaissant le caractère multinational des eaux antillaises, ont tenté de repousser ces frontières préétablies et d'élargir l'étude de cette population à une plus grande partie des îles formant les petites Antilles, en l'étendant du Centre nord (îles de Saint Kitts & Nevis) jusqu'au Sud (Grenade) (2019-2020). Cela a rendu possible l'identification de 145 nouveaux individus, augmentant ainsi de plus de 20% l'estimation de la population totale de cachalots (Vachon, Hersh, et al., 2022), pour un total de 681 individus.

L'existence de trois clans vocaux a été mise en évidence : EC1 (type de codas : 1 + 1 + 3) et EC2 (type de codas : 1 + 1 + 1 + 1 + 1), déjà documentés ((Gero, Whitehead, et al., 2016), (Gero, Bøttcher, et al., 2016)) et le potentiel EC3, constitué d'une seule unité sociale (Unité 12, type de codas: 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1). Contrairement à ce qui était préalablement supposé, le clan vocal EC2 a été rencontré au moins autant de fois que l'EC1 par Vachon (2022), faisant de lui un clan vocal à part entière des eaux Caribéennes. Une nette séparation géographique a été observée entre les deux clans avec EC1 présent plus au Nord (jusqu'à la Dominique) et EC2 plus au Sud (entre la Martinique et Sainte-Lucie). La préférence des clans vocaux pour certaines îles plutôt que d'autres ne serait pas liée à des variables environnementales (territorialité, spécialisation en fonction du type de proies ou d'habitats) mais plutôt le fait d'une transmission culturelle, familiale, reliée à l'endroit où ils ont été élevés, ont appris à chasser, etc (Vachon, Eguiguren, et al., 2022).

2.2 Objectifs de la présente étude

a- Objectifs généraux

Ces nouvelles découvertes remettent en cause les présupposés sur cette espèce, et notamment sur son caractère de « nomade de l'océan », puisque les cachalots des Petites Antilles affichent une forte résidence insulaire, associée à un faible déplacement annuel de moins de 200 km (Vachon, Hersh, et al., 2022). Il faut donc interpréter avec prudence les données antérieurement collectées autour de la Dominique et de la Guadeloupe, puisqu'il a été démontré qu'elles n'étaient pas représentatives de la situation globale des cachalots dans les Petites Antilles.

Le présent travail s'inscrit dans ce cadre, en constatant le manque de continuité dans le temps des actions ainsi que la pluralité des méthodes utilisées, qui rendent difficile l'interprétation de ces données à l'échelle du territoire global des Petites Antilles. Ainsi, depuis 2021, la CCS a élargi l'échantillonnage aux îles d'Anguilla, de St-Kitts-et-Nevis, d'Antigua & Barbuda ; à deux îles du sanctuaire Yarari (Saint-Eustache et Saba) ; et enfin, à deux îles du sanctuaire Agoa (statut d'aire marine protégée depuis 2010), celles de Saint-Martin et Saint-Barthélemy (Figure 3). En appliquant le protocole standardisé déjà développé par Vachon, Hersh, et al., (2022), la CCS s'inscrit dans la continuité des actions menées par un ensemble d'associations (DSWP, Sepenmar, Evasion Tropicale, etc) et a la volonté de collaborer avec l'ensemble des acteurs impliqués dans la conservation des cétacés dans les Petites Antilles.

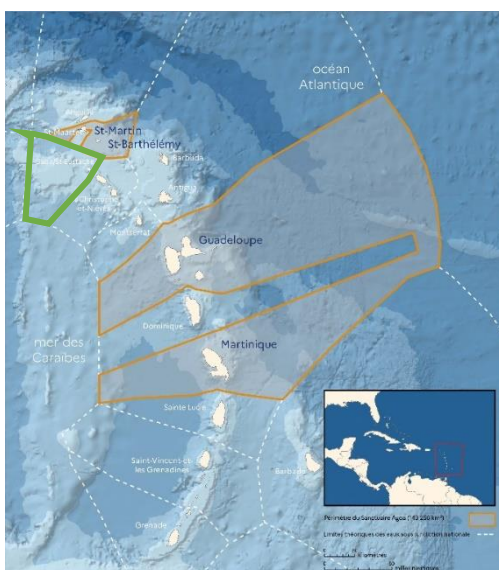


Figure 3: Aires marines protégées des sanctuaires Agoa (orange) et Yarari (vert) (Adaptée de <https://www.ofb.aouv.fr>).

Cela permet ainsi de considérer les aspects de conservation et de protection à une échelle multi-îles, considération essentielle dans le cas des cachalots, comme déjà suggéré par Gero et al. (2007). Cet aspect de collaboration inter-îles est particulièrement pertinent si l'on considère que la plupart de ces îles sont signataires de la convention de Carthagène, pour la Protection et la Mise en Valeur du Milieu Marin dans la Région des Caraïbes (1983) et ont ratifié le protocole SPAW (Specially Protected Areas and Wildlife, mis en place en 2000), relatif à la protection des espèces et espaces dans la région Caraïbe.

b- Objectifs propres à cette étude

Les objectifs globaux de cette étude sont :

- ✓ D'en apprendre plus sur les relations sociales existantes entre les cachalots vivant dans les Petites Antilles à différentes échelles (famille, groupe, clan) → Quels types ?
- ✓ De préciser l'estimation de l'abondance faite par Vachon (2022) → Combien ?
- ✓ D'analyser la répartition spatiale des différents clans du cachalot par le biais de l'étude de leurs déplacements et de leur possible distribution → Où ?

Ces objectifs seront étudiés selon trois grands axes (photo-identification, acoustique et modèle de distribution) et se déclinent en plusieurs questionnements selon le diagramme suivant (Figure 4).

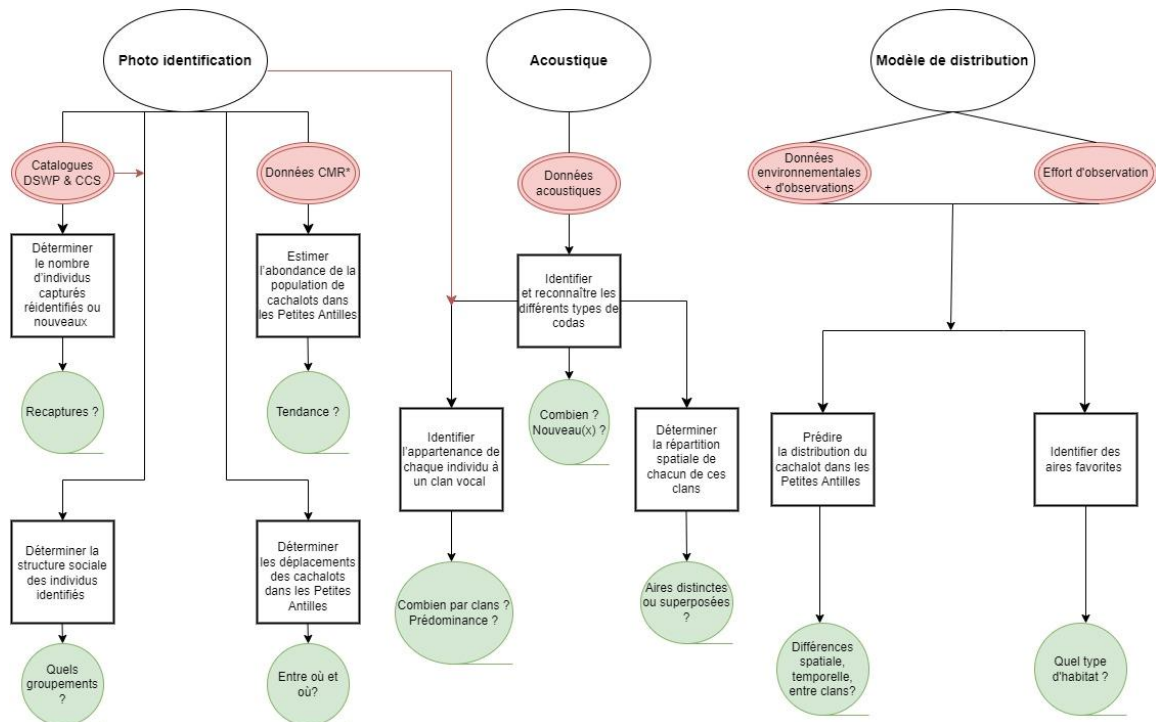


Figure 4: Présentation des objectifs de l'étude (DSWP : Dominican Sperm Whale Project ; CMR : Capture, Marquage, Recapture ; CCS : Caribbean Cetacean Society). Les ronds verts représentent les questionnements auxquels on cherchera à répondre. Les ronds rouges représentent les données utilisées pour chacun des trois axes.

2.3 Pressions & Enjeux

Les enjeux principaux liés à l'étude des cachalots sont ceux visant à prévenir leur exploitation excessive et à restaurer leurs populations, en prenant en considération les pressions passées et actuelles pesant sur ces géants, dans un but plus global de protection de l'espèce.

a- Enjeux passés : la chasse à la baleine

Le nom même de « sperm whale » (cachalot en anglais) est la preuve de la chasse abondante qui a été perpétrée sur les cachalots jusqu'à la fin du XX^{ème} siècle. En effet, son nom lui vient tout droit d'une mécompréhension, par les chasseurs, de la fonction du liquide huileux baignant le spermaceti, à cause de sa ressemblance avec le sperme des mammifères. Les cachalots ont été chassés en premier lieu pour leur lard², duquel on extrayait de l'huile pour s'en servir comme lubrificateur ou pour allumer les lampadaires (Ellis, 2018), puis pour leur ambre gris (concrétion intestinale) utilisé en parfumerie (Srinivasan, 2015) et plus tard encore pour son spermaceti, de par son caractère durcissant lorsqu'exposé à l'air libre, utilisé pour faire des bougies (Ellis, 2018).

Des initiatives internationales telles que le moratoire sur la chasse à la baleine de 1982 ou encore le classement du cachalot à l'annexe I de la Convention sur le commerce international des espèces de faune et de flore sauvages menacées d'extinction (CITES, 1975) ont été mises en place pour réglementer la chasse à la baleine et le commerce de produits dérivés des cachalots afin de prévenir notamment la surexploitation et la fragmentation de leurs populations. Dans les Petites Antilles, la

² « blubber » en anglais qui désigne l'épaisse couche de graisse permettant aux cétacés homéothermes d'être isolés thermiquement du milieu ambiant froid que représente l'eau

chasse au cachalot a cessé à des périodes différentes selon les territoires mais cette dernière a été définitivement abandonnée en 1986.

Depuis la cessation de cette chasse, de nombreuses institutions ont tenté d'évaluer le stock mondial des cachalots (Commission baleinière internationale (1991), Evans (1990)) mais les méthodes utilisées ont surestimé la population par un facteur 5. Selon Whitehead (2002), étude sur laquelle repose le statut IUCN « vulnérable » du cachalot, la chasse à la baleine aurait réduit de deux tiers la population initiale mondiale des cachalots que l'on estime aujourd'hui à environ 360 000 animaux.

Cette estimation pourrait présenter quelques biais puisqu'elle repose sur une extrapolation d'études, acoustiques, visuelles et aériennes, menées dans neuf zones principales dont les Petites Antilles sont exclues (Whitehead, 2002). De plus, elle ne tient pas compte des différences culturelles liées aux clans vocaux observées dans les sociétés matrilineales, qui pourraient même constituer des sous-populations génétiquement distinctes (L. Rendell & Whitehead, 2005).

b- Enjeux actuels

La compréhension des enjeux passés a permis de comprendre pourquoi le cachalot est aujourd'hui une espèce vulnérable et donc essentielle à protéger. Cela est d'autant plus important si l'on considère les services écosystémiques qu'ils procurent à l'Homme. En outre, en consommant environ 100 millions de tonnes de biomasse par an, la population mondiale de cachalots agit comme une pompe à carbone biologique non-négligeable (Cantor et al., 2019). De plus, en se nourrissant dans les profondeurs riches en nutriments et en venant ensuite déféquer à la surface, ils stimulent l'export de matière minérale riche en azote et en fer à la surface, éléments essentiels à la croissance du plancton. Au niveau de l'hémisphère Sud, où les eaux sont particulièrement appauvries en fer, leur présence contribue à éliminer 200 000 tonnes de carbone supplémentaires (Lavery et al., 2010).

Cependant, ces services écosystémiques sont remis en cause par les nouvelles menaces pesant sur ces géants, toutes liées de manière directe ou indirecte à l'activité humaine.

Concernant les menaces directes, on peut par exemple citer l'augmentation des potentiels enchevêtrements (Whitehead & Gero, 2015); la recrudescence du trafic maritime et donc des collisions (Fais et al., 2016; Rinaldi et al., 2021) ; l'intensification du bruit sous-marin ; la pollution marine (Gero & Whitehead, 2016); l'utilisation de sonars militaires (Sivle et al., 2012) ; la surpêche (N. & D., 2002) ou encore, la montée en puissance du whale-watching³ non-organisé et non-réglé (Rinaldi & Rinaldi, 2007). Le whale-watching, qui n'a pas arrêté de croître depuis 1992, dans le contexte de l'écotourisme, atteint aujourd'hui des sommets presque industriels (bénéfices estimés à 2 milliards de dollars en 2008 (SANDRON et al., 2015)). Dans les Caraïbes, où le tourisme joue un rôle économique majeur, le whale-watching a connu une croissance rapide et diversifiée, incluant aujourd'hui l'observation commerciale des baleines, les opérations de recherche et les activités de nage avec les cétacés (Hoyt & Hvenegaard, 2002). Lorsqu'elles sont pratiquées en excès ou qu'elles ne respectent pas les réglementations, ces pratiques peuvent avoir des conséquences négatives pour les animaux approchés, entraînant un stress accru, des ajustements énergétiques et une habitude à la présence humaine (Lundquist et al., 2013). Afin de limiter l'impact de ces activités sur les cétacés tout en promouvant l'écotourisme local, des formations ont été proposées aux whale-watchers à Sainte-Lucie et à Saint-Vincent par la CCS, dans les îles françaises par le sanctuaire Agoa et également dans d'autres îles. En Dominique, les actions du DSWP ont permis de réguler ces activités, dont l'intensité avait conduit à un harcèlement des individus.

³ Terme utilisé en science et faisant référence à l'observation des cétacés en mer à des fins touristiques



Figure 5: Echouage d'un Grand Cachalot à l'Anse Couleuvre dans les années 60 (données Sepenmar)

Les cachalots sont aussi la cible de menaces plus indirectes, telles que le dérèglement climatique. En effet, les changements climatiques déjà en cours et dont l'amplitude ne va pas cesser d'augmenter à l'avenir, menacent directement la survie de cette espèce. Une étude dans le Pacifique a montré que le succès alimentaire des clans vocaux varie au cours des cycles océanographiques El Niño et La Niña, ce qui signifie qu'avec l'augmentation prévue de ces cycles, les impacts au niveau de la population pourraient ne pas être uniformes (Brakes et al., 2019) et induire dans le pire des cas, une disparition d'un voire plusieurs de ces clans. Chez les cachalots, pour lesquels la transmission

culturelle est presque plus importante que l'environnement lui-même dans la survie, l'impact humain est alors décuplé et menace l'équilibre, voire même l'existence, des clans vocaux déjà fragiles (Gero, 2017).

Malgré l'ensemble des efforts de conservation mis en place depuis quelques décennies, consécutifs à leur forte exploitation par le passé et aux nouvelles menaces, les populations de *Pm* peinent aujourd'hui à se rétablir (Figure 5).

Les enjeux actuels relatifs à la protection des cachalots sont donc de limiter l'impact global de ces nouvelles pressions anthropiques afin de parvenir à restaurer leurs populations de manière durable. Pour l'ensemble des raisons évoquées précédemment, la considération des clans sociaux pour estimer la taille des populations et leurs vulnérabilités, plutôt que des relevés de populations classiques, apparaît alors évidente et pertinente pour la conservation des cachalots ((Whitehead & Gero, 2015) ; (Vachon, Eguiguren, et al., 2022)) et ce d'autant plus que leur compréhension reste encore largement méconnue.

II- Présentation et mise en œuvre méthodologiques :

1- Zone d'étude & Collecte des données

a- Présentation de la zone d'étude

Sur les 93 espèces de cétacés décrites mondialement, 26 d'entre elles sont présentes dans notre zone d'étude (AGOA, 2019). Cette zone comprend 15 territoires insulaires, d'Anguilla au Nord jusqu'à Grenade au Sud (entre 11 à 19°N et -65 à -57°E) pour un total de 76,896 km² parcourus. Etant donné la forte résidence observée des populations de cachalots dans les Petites Antilles (Vachon, Hersh, et al., 2022), il semblerait que ces dernières évoluent dans cette zone d'étude tout au long de l'année. Etudier cette zone est donc particulièrement intéressant pour comprendre la façon dont ces populations utilisent ce territoire.

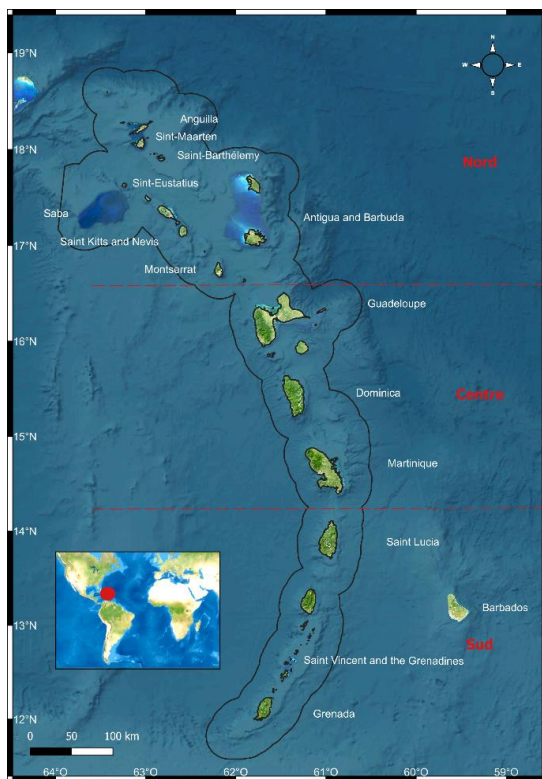


Figure 6: Zone d'étude de "Ti Whale An Nou"

Elle est divisée en trois grandes zones : la zone Nord qui s'étend d'Anguilla à Montserrat ; la zone Centre depuis la Guadeloupe jusqu'à la Martinique et enfin, la zone Sud, s'étalant de Sainte-Lucie à Grenade (Figure 6). Ces trois zones sont incluses à la fois dans la Mer des Caraïbes (côte Ouest de l'arc insulaire) et dans l'océan Atlantique (côte Est de l'arc insulaire). En effet, les îles des Petites Antilles sont formées par une série d'arcs volcaniques, résultant de la subduction de la plaque sud-américaine sous la plaque caraïbe, le long de la fosse des Caraïbes. Pour cette raison, les caractéristiques géomorphologiques et océaniques des mer et océan de part et d'autre de cette zone diffèrent grandement.

L'étude des caractéristiques géomorphologiques sur notre zone d'étude et notamment des isobathes 200, 1000 et 2000 indique que le plateau est plus large du côté atlantique et que de manière générale, les grandes profondeurs sont atteintes plus rapidement du côté Caraïbe (Figure 7, A). De nombreux canyons⁴ sont présents entre les îles et au niveau de la partie Centre de la zone d'étude (Guadeloupe, Dominique, Martinique). Ces canyons sont des lieux intéressants pour les cachalots en raison de leur productivité biologique élevée ainsi qu'à leur forte variation de profondeur, abritant une grande diversité de proies et d'habitats. De plus, leur forme crée des caractéristiques acoustiques favorables à la détection des proies.

Concernant la circulation océanique, qui peut déterminer les mouvements des proies à une méso-échelle, elle est influencée par des courants régionaux, tels que le courant des Caraïbes du côté Ouest

⁴ Source : <https://bluehabitats.org/>

de l'arc insulaire, et par des courants océaniques plus larges du côté Est, tels que le courant nord-atlantique (Figure 7, B).

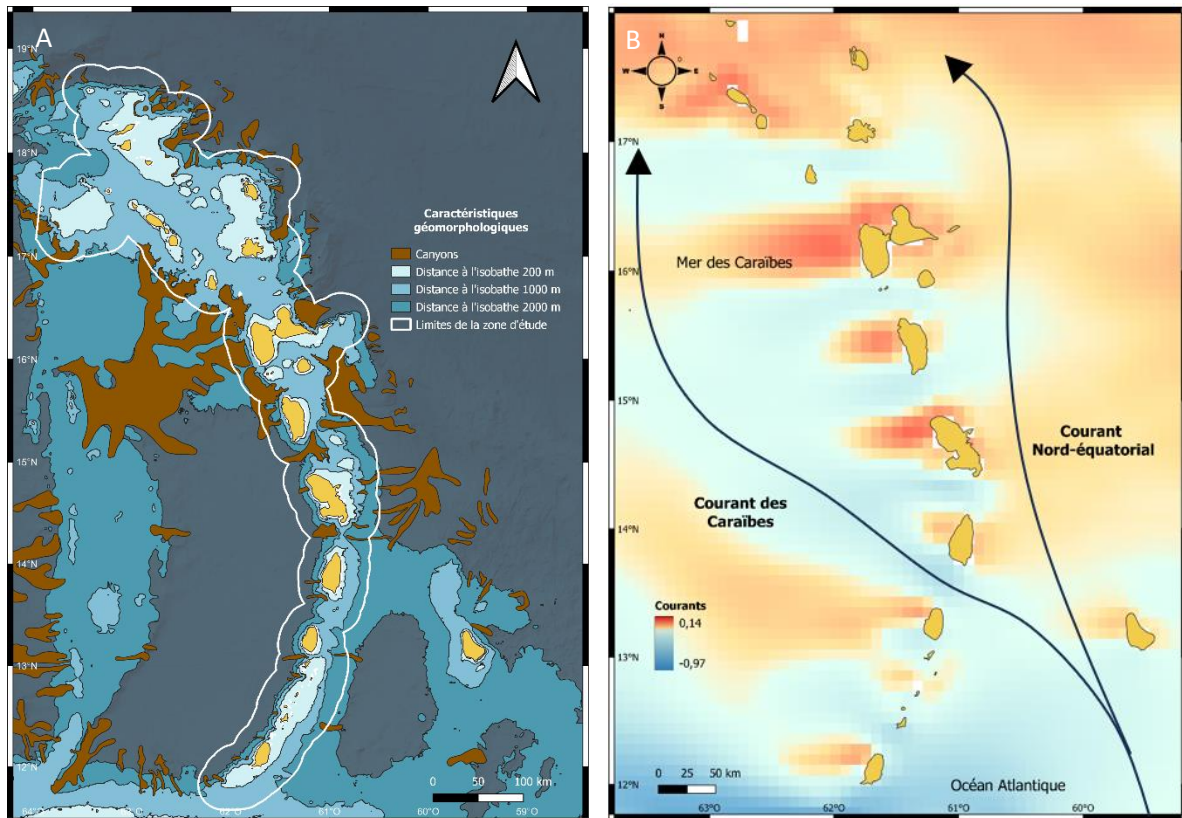


Figure 7: Présentation des caractéristiques géomorphologiques (A) et des courants océaniques (B) au niveau de la zone d'étude (Annexe 2 pour la source des données utilisées). Les flèches présentées en figure B sont purement schématique pour montrer le sens des courants dans la zone. La légende représente les valeurs prises par la composante Est de ces courants.

De la même façon, des différences existent également entre le Nord et le Sud des Petites Antilles, et ce, surtout du côté Caraïbe, ceci étant principalement dû au déplacement de l'arc insulaire au cours des temps géologiques. Ainsi, la profondeur moyenne augmente en se déplaçant vers le sud tandis que la largeur du plateau continental diminue à mesure qu'on se rapproche de Saint-Vincent-et-les-Grenadines (Figure 7, A).

Les différences observées font des Petites Antilles une région particulièrement intéressante à étudier de par la grande diversité des reliefs observés, abritant une grande pluralité d'espèces. Étudier cette zone permettra notamment de mieux comprendre les zones de résidence préférentielle des cachalots, en fonction de leurs caractéristiques géomorphologiques, océaniques et d'autres paramètres environnementaux variables au cours du temps.

b- Collecte des données

Pour permettre un suivi à long terme viable, dans le cadre du programme Ti Whale An Nou, la CCS a mis en place une méthode standardisée, en reprenant notamment celle développée par Vachon, Hersh, et al., (2022). Les expéditions, d'une durée de 15 jours chacune, ont été menées entre mars et août en 2021, 2022 et 2023 en parcourant l'ensemble des îles citées précédemment. Des campagnes supplémentaires ont également été réalisées au mois de septembre. Ce rapport se concentrera sur 2021 et 2022, la campagne 2023 n'étant pas encore terminée. Ainsi, notre période d'étude réelle s'étend du mois de mai à la mi-septembre 2021 et du mois de mars à la fin septembre 2022. Les raisons principales ayant menées au choix de cette période sont les conditions météorologiques favorables.

Afin de rendre compte de la présence des cétacés, le suivi a été réalisé sur la zone d'étude présentée en Figure 8, A, qui s'étend jusqu'à 25 km des côtes environ. Cet effort d'échantillonnage (Figure 8, A) n'est pas réalisé de façon systématique mais décidé par le chef d'expédition, en prenant en considération les conditions météorologiques (i.e arrêt de l'effort si vent supérieur à 21 nœuds, correspondant au niveau 4 sur l'échelle de Beaufort, ou pluie intense), le temps de navigation et le point de chute de la journée d'expédition. A bord, l'équipe est constituée de 6 à 10 personnes, réparties en groupes de deux dont les positions et rôles changent toutes les deux heures (Figure 8, B). La récolte des données est ainsi séparée en trois branches principales : données GPS, données photographiques et données acoustiques.

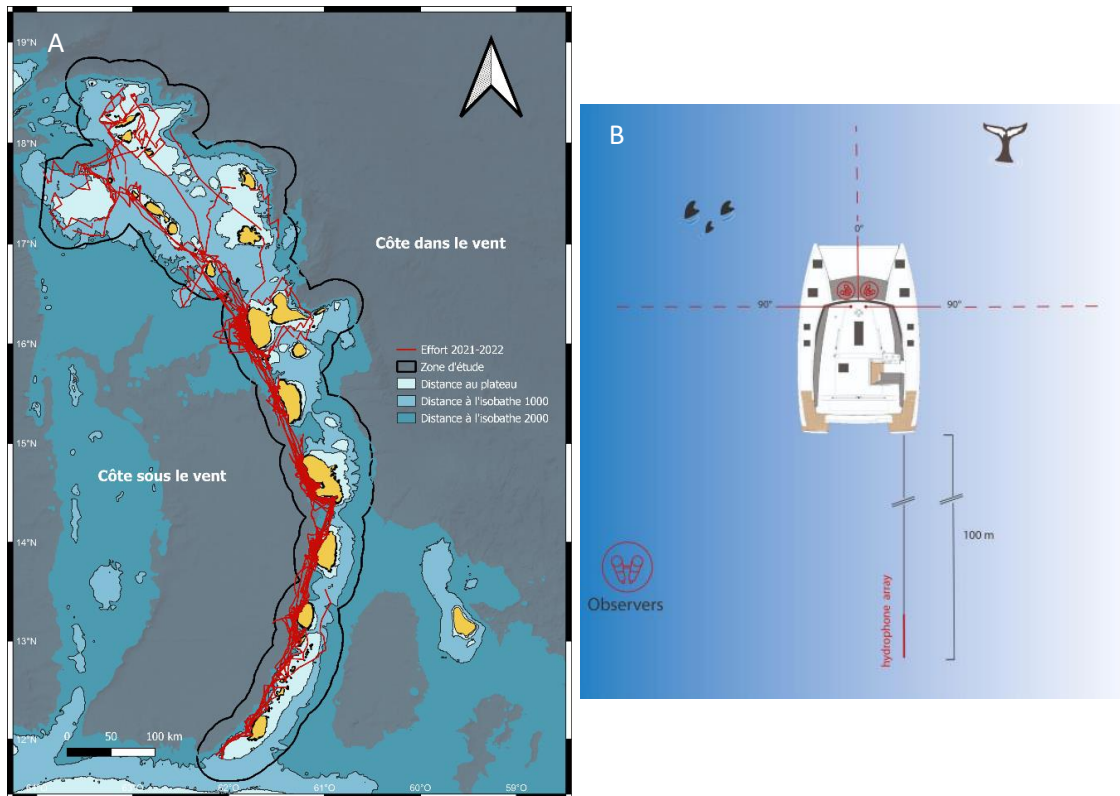


Figure 8: Effort d'échantillonnage en 2021 & 2022 (A, en rouge) et organisation à bord montrant la position des observateurs et de l'hydrophone câblé (B)

Données GPS : Afin d'enregistrer les paramètres climatologiques, le trafic maritime et les observations faites en mer, l'application ObsEnMER, développée par le GECC (Groupe d'Etude des Cétacés du Cotentin) est utilisée. De nombreuses informations sont ainsi collectées telles que :

- La position géographique du bateau à chaque instant (latitude, longitude et timestamp associé)
- Le nombre de bateaux et leurs types
- Les différents types de sons entendus (bruits de moteurs, bruits parasites, sons émis par les cétacés)
- Des paramètres climatologiques comme la houle, le vent et sa direction, la couverture nuageuse ou encore la visibilité
- Des paramètres logistiques : entrée/sortie de l'eau de l'hydrophone, début/fin de l'effort, début/fin suivi, etc.

Lors des observations de cétacés, des informations complémentaires sont rentrées dans l'application telles que l'espèce identifiée, la distance entre le bateau et l'observation, le nombre d'individus total, le nombre d'individus classés en fonction de leur âge (adulte, juvénile, nouveau-né).

Données acoustiques :

Un système d'hydrophones tractés à 100 m derrière le bateau, a été utilisé pour enregistrer les sons abiotiques, biotiques et anthropiques en continu (Figure 8, B). Ce dernier a été déployé seulement lorsque la profondeur était supérieure à 50 m et que le trafic maritime était évalué comme peu dense pour éviter les risques de coupure/perte de matériel. Dès lors que ces conditions sont atteintes, l'enregistrement débute jusqu'à approximativement la fin de l'effort journalier. Chaque enregistrement journalier est divisé en sous-enregistrements durant entre 5min35 et 9min 42, dépendamment de l'année où il a été enregistré (respectivement 2022 et 2021). L'hydrophone est connecté en temps réel à une unité pour l'acquisition des données et à un ordinateur avec le logiciel PAMGuard (Gillespie et al., 2008). Ce logiciel d'acquisition acoustique est spécialisé dans la détection, le traitement, la visualisation et l'analyse des sons sous-marins, pour classer les événements acoustiques d'intérêt.

L'utilisation de l'acoustique et de logiciel comme PAMGuard est particulièrement intéressante dans le cas des cachalots car il s'agit d'une espèce qualifiée de « grand plongeur », passant plus de 70% de son temps sous l'eau, et également d'une espèce très sociale, dont les vocalises sont fréquentes. En ne nécessitant pas d'observation visuelle et donc conséquemment que l'animal soit en surface, l'acoustique permet d'avoir des yeux sous l'eau et fournit ainsi des informations précieuses sur la taille des groupes, le type de comportement, la répartition géographique, ainsi que l'appartenance à un clan vocal par exemple. Le logiciel PAMGuard (Gillespie et al., 2008) permet ainsi de détecter les clics émis par les cachalots lorsque ceux-ci n'ont pas encore été visuellement détectés ou que la visibilité est réduite (nuit, brume, ...) et facilite la localisation de ces géants.

Cette étude acoustique est combinée à un effort visuel continu de 6h à 18h (en conditions météorologiques parfaites) et à la prise de photographies lors d'une rencontre avec des cétacés.

Données photographiques :

Les aspects acoustiques et visuels sont ici étroitement liés car dans la plupart des cas, les cachalots ont été d'abord repérés acoustiquement, ce qui a permis leur détection visuelle par la suite. En outre, le logiciel PamGuard permet, grâce à un procédé de triangulation, d'avoir une idée de la direction dans laquelle chercher les cachalots lorsque ceux-ci ont été détectés acoustiquement via leurs clics. En combinant ce procédé à l'intensité acoustique à laquelle sont émis les clics, la détection visuelle en est alors facilitée.

Les cachalots sont facilement reconnaissables de loin, par la forme et l'orientation de leur souffle (en buisson, émis vers l'avant et vers la gauche). Après avoir repris son souffle en surface, il tend à rester immobile ou à nager très doucement pour prendre sa respiration avant la prochaine plongée (phase qui dure environ 9min). Juste avant de sombrer à nouveau dans les profondeurs, le cachalot lève sa nageoire caudale (Whitehead, 2018). C'est à ce moment-là qu'intervient la photo-identification de l'espèce. En outre, chaque nageoire caudale, de par sa forme et ses marques, est unique, à l'image de l'empreinte digitale des hommes. Ainsi, lors de la prise de photographie, si possible, la partie ventrale de la nageoire caudale a été prise en photo pour permettre l'identification individuelle. La photo-identification permet ainsi d'identifier chaque individu et de savoir si certains d'entre eux ont été recapturés d'une année à l'autre.

La méthode de photo-identification a été choisie car elle est particulièrement efficace pour suivre les individus sur de longues périodes et notamment dans le cadre de la coopération scientifique puisqu'il s'agit de données facilement partageables. De plus, il s'agit d'une méthode peu invasive, ne nécessitant pas de réelle capture ou marquage et n'induisant pas/ ou peu de perturbation du comportement naturel de l'espèce. La phase de rapprochement des animaux, nécessaire à la collecte des données, a été conforme aux règles fixées par l'arrêté n° R-02-2017-03-15-003, c'est-à-dire que la

trajectoire du bateau était toujours parallèle à la trajectoire de l'animal et localisé derrière lui si possible et la vitesse régulière était calée sur l'animal le plus lent et limitée à 5 nœuds.

c- Equipements utilisés

Les relevés ont été effectués à bord de 4 types de voiliers différents (des catamarans Fontaine-Pajot dont la taille varie entre 12 à 14 m) dont les caractéristiques de navigation étaient similaires (stabilité, poids, taille, etc). Lors de chaque expédition, les équipements suivants seront utilisés :

- Appareils photos (Canon 5D, Canon 90D, Sony A7RIV avec des objectifs de 70-300 mm et 100-400 mm)
- Système d'hydrophones tractés réunissant deux hydrophones Magrec HPO3 hautes fréquences (entre 2 kHz et 250 kHz), chacun de ces hydrophones étant reliés à un préamplificateur de type Magrec HP02.
- Logiciel d'acquisition acoustique PamGuard (Gillespie et al., 2008) associé à un ordinateur et une carte son de type Pambox (Marine research ecology ltd).
- Tablette avec l'application ObsEnMer (Altitude creation company, release 3.08)

2- Méthodologies associées à l'identification visuelle et la détection acoustique

2.1 Processus d'identification via Flukebook

Afin de comprendre la structure sociale et donc la composition des différents clans vocaux, il est primordial d'identifier individuellement chaque membre de cette communauté. Le Tableau 1 regroupe par expédition, les journées où des cachalots ont été observés et capturés photographiquement ainsi que le nombre d'observations associé à chacune de ces journées.

Tableau 1: Expéditions et dates d'observations des cachalots

Nom de l'expédition	Dates d'observations (nombre d'observations associé)
Expe_2021_1_N	21/05 (1) ; 23/05 (2) ; 26/05 (1) ; 28/05 (1)
Expe_2021_2_N	03/06 (1) ; 12/06 (2)
Expe_2021_3_S	26/06 (1) ; 27/06 (2)
Expe_2021_4_C	07/07 (2); 12/07 (1)
Expe_2021_5_S	21/07 (1)
Expéditions supplémentaires_ formations_2021	14/09 (1)
Expe_2022_1_S	26/03 (1) ; 27/03 (2); 28/03 (2); 30/03 (1)
Expe_2022_2_C	22/04 (1) ; 24/04 (1)
Expe_2022_4_S	19/06 (1) ; 20/06 (2) ; 22/06 (2)
Expe_2022_5_C	16/07 (1); 23/07 (2)
Expe_2022_6_N	16/08 (1); 17/08 (1); 29/08 (1)
Expéditions supplémentaires_ formations_2022	23/09 (1) ; 24/09 (1) ; 27/09 (1) ; 28/09 (2) ; 29/09 (1)
TOTAL : 10 expéditions + campagnes supplémentaires	TOTAL : 41 observations avec photos

a- Présentation de Flukebook

Flukebook est une plateforme d'intelligence artificielle en ligne collaborative, utilisée dans le cadre de la photo-identification des cétacés (Levenson et al., 2015). A partir de photographies de la nageoire caudale des cachalots, elle parvient, grâce à différents algorithmes, à trouver des correspondances, au sein des photos soumises par un même utilisateur ou par plusieurs utilisateurs, entre deux photos présentant de fortes similarités et donc sur lesquelles apparaissent possiblement le même individu. Cela permet notamment d'identifier individuellement les baleines mais également leurs relations sociales, en dressant des tables de co-occurrences d'individus vus plusieurs fois ensemble, ce qui est particulièrement intéressant dans le cadre de l'organisation sociale complexe des cachalots. Via la coopération entre les différents utilisateurs, Flukebook facilite le partage et l'analyse des données et est précieux dans le cadre de la conservation des baleines, en identifiant leur possible migration et la dynamique de leurs populations (Levenson et al., 2015).

Avant de soumettre les données à Flukebook, une étape de nettoyage et d'organisation des données a été nécessaire. L'ensemble des photos a été renommé afin que le même nom ne soit pas présent plusieurs fois. Un fichier contenant toute l'information liée aux métadonnées des photos (heure et date de prise de vue, localisation, numéro de l'observation, photos en rafale, etc) a également été créé.

b- Fonctionnement de Flukebook

Le fonctionnement global de Flukebook peut se diviser en trois grandes étapes :

1ère étape : Chargement des photos & lien avec les métadonnées associées :

Cette étape peut être longue car elle nécessite le chargement de l'ensemble des photos sur les serveurs de Flukebook, processus lors duquel une connexion internet est nécessaire.

2^{ème} étape : Lancement de la détection :

Dans un premier temps, Flukebook va passer en revue toutes les photos et détecter la nageoire caudale du cachalot. Chaque individu ainsi repéré se verra attribuer un numéro de rencontre. Si plusieurs individus étaient présents sur la même photo et que Flukebook les a repérés, alors il attribue un numéro de rencontre différent à chaque nouvel individu. A l'issue de cette étape, une colonne dans Flukebook montre le résultat du processus de détection, en affichant un code constitué de la partie détectée (dans notre cas, la nageoire caudale) et de l'espèce (dans notre cas, le cachalot). Les rencontres pour lesquelles aucun code n'est affiché sont celles pour lesquelles Flukebook n'a pas été capable de repérer d'individus (qualité de la photo insuffisante, photo à contre-jour, etc). Il faut alors réaliser l'étape de détection manuellement en se rendant sur la page de la rencontre associée. Cette étape a également dû être réalisée lorsque plusieurs individus étaient présents sur la photo mais que seulement certains d'entre eux avaient été détectés.

3^{ème} étape : Lancement de l'identification :

Le processus de l'identification est celui pendant lequel les différents algorithmes de l'intelligence artificielle cherchent à trouver des correspondances entre les photos, afin de mettre en évidence un individu nouveau ou un individu déjà identifié sur d'autres photos déjà présentes sur Flukebook et/ou rentrées par d'autres utilisateurs. Avant de lancer ce processus, il faut préciser la région du monde sur laquelle les algorithmes doivent chercher des possibles ressemblances avec des photos déjà rentrées sur la plateforme. Il est ainsi possible de faire tourner les algorithmes à différentes échelles (régionale, internationale, etc).

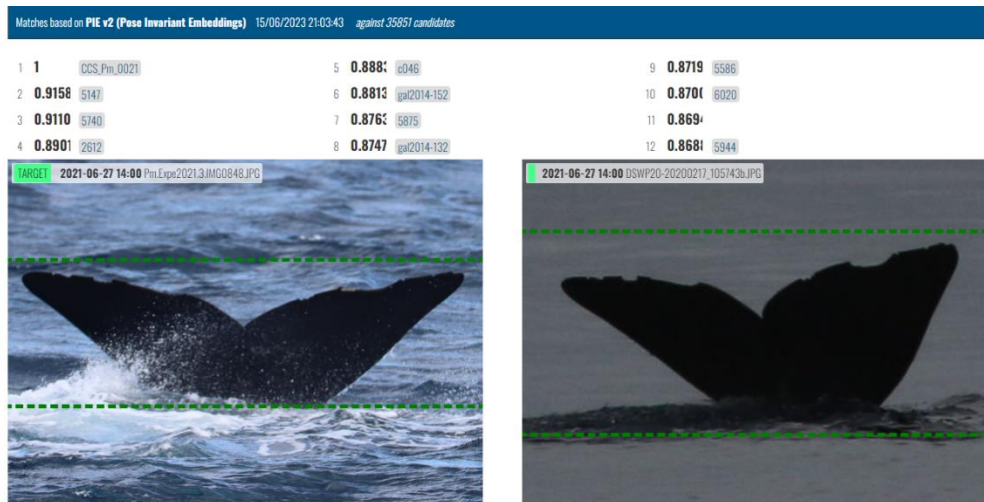


Figure 9: Exemple de résultats proposés par l'algorithme PIE de Flukebook, à l'issue de la phase d'identification

Dans le cadre des cachalots, deux algorithmes principaux sont utilisés : l'algorithme CurvRank et l'algorithme PIE (Pose Invariant Embeddings) (Blount et al., 2022). Bien qu'utilisant des méthodes différentes, ils se basent tous les deux sur l'analyse du bord de fuite de la nageoire caudale des cachalots, qui représente la partie la plus marquée et où sont présentes les entailles et encoches (Figure 9).

L'algorithme CurvRank :

L'objectif principal de l'algorithme CurvRank est de quantifier la similarité entre différents contours de nageoires caudales (flukes) en se basant sur leurs motifs caractéristiques, afin d'identifier les individus de manière précise.

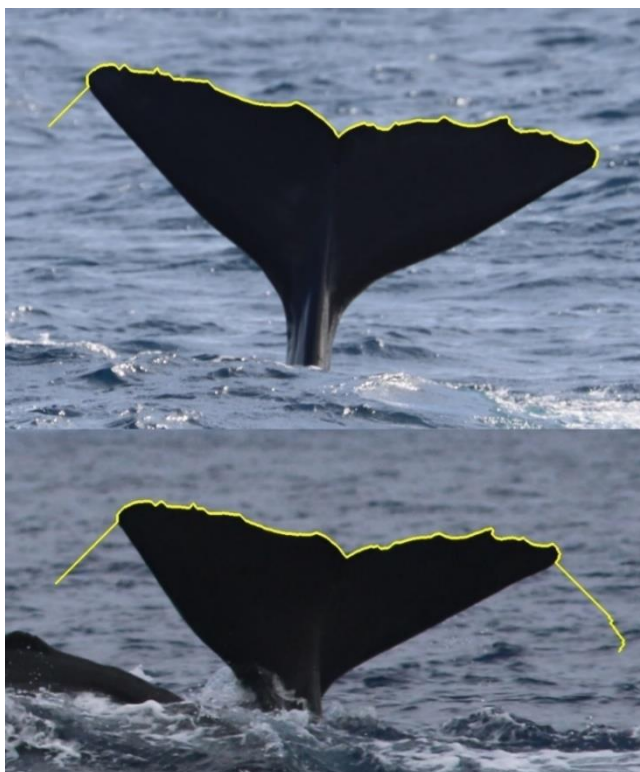


Figure 10: Exemple de fonctionnement de l'algorithme CurvRank. La ligne jaune représente la courbe tracée par l'algorithme à partir des irrégularités de la nageoire caudale

Pour cela, il utilise une mesure de similarité appelée "dynamic time warping" (DTW), qui mesure la distance entre deux courbes de contours de nageoires caudales et trouve le meilleur alignement possible, en minimisant la distance globale entre elles, même si elles ont des différences de longueur, de forme ou d'orientation (Weideman et al., 2017). Il applique une pondération pour mettre l'accent sur les caractéristiques les plus discriminantes. Cela permet de comparer les contours des nageoires caudales de différents cachalots et de quantifier leur similitude malgré les variations individuelles (Figure 10).

L'algorithme PIE :

L'objectif principal de l'algorithme PIE est de capturer les caractéristiques distinctives des baleines, indépendamment de leur pose ou de leur orientation dans les images. Pour cela, il s'appuie sur l'apprentissage de représentations dites "embeddings" (encodages) qui permettent de projeter les caractéristiques des nageoires caudales des baleines dans un espace vectoriel. Pour entraîner l'algorithme PIE, on utilise généralement des méthodes basées sur les "triplet loss", des fonctions de perte, utilisées pour comparer les similarités entre des paires d'échantillons (ancres et positifs) par rapport à des échantillons négatifs (Blount et al., 2022). Dans le cas des cachalots par exemple, une paire ancre-positif pourrait correspondre à deux images du même individu dans des poses différentes, tandis qu'un échantillon négatif serait une image d'un individu différent. L'objectif est ainsi de minimiser la distance entre les embeddings des paires ancre-positif, tout en maximisant la distance entre les embeddings des paires ancre-négatif, afin de regrouper les individus similaires et les séparer des individus différents.

c- Analyse des résultats issus de Flukebook

Les résultats du processus d'identification de Flukebook ont été utiles à différents niveaux.

Dans un premier temps et tout naturellement, ils ont permis d'identifier le nombre d'individus total photo-identifiables vus lors des expéditions 2021 et 2022 et également le nombre d'individus déjà identifiés par d'autres associations. Cela a notamment été utile afin de constituer les catalogues de photo-identification (meilleure photo de chaque individu identifié) et de capture-marquage-recapture⁵(meilleure photo à chaque occasion de capture des individus identifiés).

Ces informations ont contribué à réaliser des cartes de déplacements des différents individus afin d'étudier leur distribution globale. Ces cartes ont été réalisées en utilisant le logiciel QGIS (version 3.24.3). Ces données ont également été nécessaires pour déterminer l'abondance de la population et l'organisation de cette population en unités sociales ou groupements, notamment grâce au logiciel SOCPROG (Whitehead, 2009). L'identification individuelle et les correspondances avec des individus déjà identifiés par d'autres organisations (DSWP notamment) ont donné l'occasion d'obtenir des informations sur le clan vocal de ces individus, avant même d'avoir effectué l'analyse acoustique.

2.2 Processus d'identification acoustique via Audacity

Comme énoncé précédemment, les cachalots émettent différents types de clics dont les codas, qui sont plus rythmiques que les clics d'écholocation. L'objectif de l'analyse acoustique a été de repérer les codas au sein des enregistrements des cachalots, dans un premier temps, seulement les enregistrements qui concordaient avec des observations visuelles, puis, dans un second temps, également ceux qui révélaient la présence de cachalots hors effort ou hors observation.

a- Analyse préliminaire par écoute simple

Dans un premier temps, lorsqu'un cachalot avait été vu ou entendu au cours d'une journée d'expédition, l'ensemble des enregistrements de cette journée ont été récupérés. Ensuite, grâce à

⁵ Les termes de « capture » ou « recapture » utilisés dans ce mémoire font référence à la photo-identification d'un individu pour la première fois (« capture ») et des fois suivantes (les différentes « recaptures »). L'étape « marquage » de la méthode CMR ici correspond juste au fait de reconnaître un individu par les marques sur sa nageoire.

l'information temporelle du début de l'observation ou du premier enregistrement où des clics ont été repérés par PAMGuard (Gillespie et al., 2008), les différents enregistrements où des clics sont distingués ont été sélectionnés. Parfois, et particulièrement pour les enregistrements nocturnes réalisés en 2021, les cachalots étaient présents et avaient commencé à émettre des clics bien avant qu'ils soient détectés par les écoutes réalisées toutes les heures par les membres de l'équipage. Ainsi, pour l'ensemble des journées d'enregistrements, les supposées heures de début et de fin des clics ont été considérées avec précaution, et les enregistrements précédents et suivants le début et la fin potentielles des clics entendus ont tout de même été partiellement écoutés (10s par tranches de 30s pour les enregistrements dans lesquels des clics ne devaient pas être entendus).

Suite à un problème sur le canal d'acquisition de l'hydrophone, les enregistrements de l'expédition 1 et 5 de 2022 ont été inexploitable. D'autres problèmes tels que la présence d'enregistrements vides ou d'enregistrements où aucun cachalot n'a été entendu ont contribué à diminuer le nombre d'enregistrements finalement exploitables. Ainsi, sur les 52 journées d'enregistrements, répartis sur 9 expéditions, seulement 33 étaient exploitables (Annexe 1) et ont été écoutés avec un logiciel d'écoute simple (UniConverter Player).

Il se peut que sur une même journée d'enregistrements, plusieurs détections acoustiques aient été décelées ; par exemple, une associée à une observation et une autre entendue hors de cette observation. Ces enregistrements ont donc été considérés comme deux rencontres différentes. Dans l'Annexe 1 et dans le cas décrit précédemment, la journée n'a été comptée qu'une seule fois.

b- Étude plus approfondie sur Audacity et confirmation des codas entendus

L'ensemble des enregistrements a été écouté et visionné à l'aide du logiciel Audacity (Audacity® Cross-Platform Sound Editor) afin de dissocier les enregistrements ne présentant que des clics de ceux présentant à la fois clics et codas ou codas uniquement (rares). Lorsque des codas étaient détectés, les informations relatives à leur position temporelle dans l'enregistrement en question (ex : de 1min35s à 1min56s), à leur nombre (ex : 5) et à leur type (EC1, EC2, EC3, inconnu) ont été notées.

c- Analyse des résultats acoustiques

La distribution des différents clans (Figure 19.A) a été estimée via des cartes de chaleurs calculées à partir du modèle de densité de Kernel. Ce modèle utilise plusieurs paramètres pour créer la carte de chaleur tels que :

1. **Dimension des cellules** ou pixels utilisés pour représenter la carte de chaleur : La densité des rencontres acoustiques a été calculée à partir des individus de chaque clan à 0.001° de résolution, soit 0.1km. Cette valeur a été choisie afin de ne pas alourdir le calcul avec une valeur trop petite mais pour avoir une résolution tout de même assez fine.
2. **Le rayon du noyau de Kernel** : Il représente la portée d'influence de chaque point. Il a été fixé à 6 km dans notre cas, afin de représenter la distribution des clans vocaux sur le territoire de manière plus douce, sans pourtant être trop haut, pour éviter de surévaluer l'ampleur de la distribution de ces clans par rapport à leurs positions GPS initiales.
3. Le **paramètre de forme du noyau** déterminant la fonction de noyau utilisée pour lisser les points et estimer la densité : Celle utilisée dans notre cas est le noyau « Quartique » qui attribue une pondération plus grande aux points proches du centre du noyau, et une pondération décroissante à mesure que l'on s'éloigne.

2.3 Détermination de l'appartenance de chaque individu à un clan vocal

La confrontation des données visuelles et acoustiques a donné lieu à l'identification de l'appartenance de chaque individu ou groupes d'individus à un clan vocal. Pour l'attribution du clan vocal à une observation, cette dernière s'est faite naturellement lorsque la détection acoustique et l'observation avait eu lieu en même temps. Dans les autres cas de figures, l'attribution d'enregistrements acoustiques à une observation a été réalisée lorsqu'il y avait une continuité acoustique jusqu'à l'heure de l'observation permettant de considérer que c'était le même groupe d'individus. Si deux observations étaient présentes pour une continuité acoustique, tous les enregistrements qui étaient antérieurs à l'heure de début de la deuxième observation ont été attribués à la première observation.

Les enregistrements avec lesquels aucune observation n'a été associée, ou une observation a été associée mais aucune photo n'a été prise, ont été classés dans la catégorie des « autres enregistrements » qui n'apportent pas d'information sur l'appartenance de ces individus (non-identifiés) à un clan vocal. L'information manquante concernant les clans vocaux des individus déjà connus et pour lesquels aucun coda n'avait été entendu a pu être complétée par les données du DSWP et de la thèse de Vachon (2022).

2.4 Détermination des relations entre individus

L'analyse des relations sociales entre les individus a été réalisée sur SOCPROG (Whitehead, 2009). Ce logiciel a été choisi par rapport à Mark car il est plus adapté au traitement de seulement deux années de suivi. En effet, l'utilisation de Mark requiert d'avoir plus de deux occasions de capture.

Préparation des données nécessaire aux analyses sur SOCPROG

L'analyse exclut les individus mâles et les veaux et sera par conséquent menée seulement sur les femelles adultes et les individus juvéniles faisant partie intégrante des unités matrilineales pendant plusieurs années. Le regroupement des individus ensemble utilise la méthode des vecteurs propres de Newman (2006) qui représente le réseau sous forme d'une matrice d'adjacence, dans laquelle les poids de liens entre les individus du réseau sont représentés. Les individus sont ensuite partitionnés dans des clusters en fonction de leur degré de connexion les uns aux autres. Enfin, la dernière étape de cette méthode consiste en l'évaluation du partitionnement par le biais du calcul d'un indice de modularité quantifiant la densité des liens à l'intérieur des clusters par rapport aux liens entre les clusters.

Les associations entre les individus ont été étudiées en choisissant un indice d'association à poids égal (ou « half weight »). L'utilisation de cet indice est généralement recommandée lorsqu'on veut définir l'association entre deux individus en considérant à la fois les similitudes positives (c'est-à-dire les fois où ils ont effectivement été vus ensemble) et les similitudes négatives (c'est-à-dire les fois où ils ont été vus seuls) (Whitehead, 2008: 97-104). Cet indice attribue alors un poids égal aux similitudes positives et aux similitudes négatives et permet ainsi de fournir une mesure plus complète de l'association, en prenant en compte les caractéristiques communes et distinctives des individus. Il varie entre 0 et 1, un indice proche de 1 signifiant une similarité plus élevée entre les individus.

Analyse des associations entre individus selon la modularité

Afin d'étudier les relations entre les individus, l'option « Community division by modularity » de SOCPROG a été utilisée. Cette dernière permet de diviser la communauté étudiée en cluster, de telle sorte que les indices d'association soient généralement élevés parmi les individus d'un même cluster et généralement faibles parmi les individus de clusters différents. Cette division de la communauté en clusters est possible via le calcul de l'indice de modularité (Newman, 2004). Il correspond à la

différence entre la proportion de l'association totale au sein des clusters et la proportion attendue, compte tenu de la somme des associations des différents individus. Le type de modularité utilisé ici est la modularité de type 1, c'est-à-dire celle tenant compte du caractère grégaire⁶ de l'espèce (Godde et al., 2013). Une modularité supérieure à 0,3 est considérée comme indiquant une division utile de la population (Newman, 2004).

Analyse des associations entre individus selon le regroupement en cluster hiérarchique

L'indice d'association à poids égal est souvent utilisé avec des méthodes de clustering telles que la méthode de Ward. Cette méthode ainsi que celle dite de « lien moyen » (Average linkage) sont celles qui sont souvent préférées parmi les autres méthodes (Milligan & Cooper, 1987). La méthode de Ward favorise les clusters compacts et homogènes, où les individus au sein d'un même cluster sont similaires les uns aux autres. En revanche, la méthode "Average linkage" se focalise sur les similarités moyennes entre tous les individus, privilégiant ainsi les regroupements basés sur des similarités globales.

Il convient alors de tester les deux méthodes pour voir celle qui est la plus appropriée à nos données. Cela peut être fait en calculant les coefficients de corrélations cophénétique associées à chacune des méthodes. Ce coefficient indique dans quelle mesure le dendrogramme correspond à la matrice des indices d'association. Lorsque ce coefficient indique une valeur supérieure à 0.8 (Bridge, 1993) alors, le dendrogramme proposé est estimé comme être une bonne représentation des relations réelles existant entre les individus.

Définition utilisée pour la distinction entre unités sociales et groupes

Dans l'analyse ultérieure, un groupe est défini comme tous les individus vus au cours d'une même observation (Gero et al., 2014). Au cours du temps, plusieurs critères ont été utilisés pour distinguer les unités sociales des groupes. Le plus flexible considère qu'on peut considérer que deux individus font partie d'une même unité sociale s'ils ont été revus ensemble à 30 jours d'intervalle minimum (Christal et al., 1998). Le second, plus restrictif mais plus fiable, considère que les individus doivent avoir été recapturés ensemble sur au moins deux années différentes (Gero et al., 2014). L'utilisation de ces deux définitions sera développée dans la partie discussion, notamment pour comprendre l'attribution des nouveaux individus aux unités sociales déjà connues.

2.5 Evaluation de l'abondance de la population

De nombreuses politiques de conservation se basent sur la connaissance de l'abondance d'une population afin de mettre en place des mesures de protection pour cette espèce. C'est dans ce contexte que l'abondance de la population a souhaité être évaluée avec les données de la CCS. Cette volonté fait suite aux travaux réalisés par Vachon (2022) et a l'espoir de venir affiner les estimations préalablement réalisées.

Dans cette étude, une population sera définie comme tous les individus de la même espèce co-existant au sein de la même zone en même temps (Van Dyke 2008). L'estimation a été faite en utilisant les données de photo-identification se basant uniquement sur les individus adultes des sociétés matrilineales (femelles et mâles juvéniles) dont la population a été considérée comme fermée. Utiliser ce genre de modèles suggère de ne pas considérer les phénomènes d'émigration, d'immigration, de naissance et de décès. Or, dans notre cas, sur deux années d'étude, les hypothèses de non-décès et non-naissance sont très probablement vérifiées, et ce d'autant plus si l'on considère la longévité du

⁶ La tendance des individus à s'associer les uns avec les autres

cachalot (mortalité estimée à 0.0505/an dans les Petites Antilles, (Whitehead & Gero, 2015)). Concernant les phénomènes d'émigration et d'immigration, ils peuvent être considérés comme presque-nuls dans le cas des populations fortement résidentes de cachalots dans les Petites Antilles (Vachon, Hersh, et al., 2022).

Deux types de modèles fermés ont été utilisés dans le cadre de cette étude : le modèle de Peterson fermé avec modification de Chapman et celui de Schnabel utilisant la vraisemblance maximale.

- Le modèle de Peterson avec modification de Chapman utilise à la fois les données de recapture (Peterson « basique ») et les données de non-recapture (Chapman) pour estimer plus précisément l'abondance de la population. En tenant compte des individus marqués mais non recapturés lors des observations ultérieures, cette approche améliore l'estimation de l'abondance.
- Le modèle fermé de Schnabel avec maximum de vraisemblance : Ce test utilise des données de capture-recapture qui incluent à la fois les individus marqués et les individus non marqués observés lors de différentes occasions de capture ainsi que d'autres paramètres associés à la dynamique de la population. Il utilise des modèles statistiques spécifiques basés sur la vraisemblance maximale pour estimer la taille de la population.

Deux hypothèses importantes de ces tests sont donc la persistance des marques dans le temps, ainsi que l'homogénéité de la capture de tous les individus (probabilités égales de capturer tous les individus, peu importe l'importance de leurs marques). Dans le cadre des rencontres avec les cachalots, qui sont souvent présents en petits groupes (entre 7 à 9 individus, (Gero et al., 2014)), il est peu fréquent que l'ensemble des individus ne soient pas capturés.

2.6 Détermination des déplacements des unités sociales

Afin de mieux rendre compte du déplacement des cachalots à l'échelle de l'ensemble des Petites Antilles, une partie des données collectées par le DSWP depuis 2005 et par Vachon en 2019-2020 ont été prises en considération lors de la création des cartes (Figure 20, Figure 21, Figure 22). Seules les données des individus capturés au moins une fois par la CCS ont été utilisées dans ce cadre. En fonction de l'information à représenter sur la carte, les données ont été adaptées pour garder une carte compréhensible et non surchargée. Ainsi, dans la Figure 22 par exemple, uniquement les individus recapturés par la CCS et pour lesquels le clan vocal est connu ont été pris en considération. Pour chaque rencontre avec une unité sociale ou un groupement, une seule position GPS a été gardée par rencontre représentant la position de l'ensemble des individus de l'unité sociale en question. Même si tous les membres d'une unité sociale n'étaient pas présents lors d'une rencontre, il a été considéré que l'unité sociale était présente à cette date-là.

3- Création d'un modèle de distribution d'espèce de type GAM (Generalized Additive Model)

3.1 Description du modèle additif généralisé

Le modèle de régression linéaire fait l'hypothèse que la distribution conditionnelle de la réponse Y sachant les variables explicatives est une loi gaussienne. Le modèle linéaire généralisé (aussi appelé « GLM ») permet de prendre en compte n'importe quelle autre loi pour cette distribution conditionnelle. Ce type de modèle est couramment utilisé pour modéliser la présence-absence des espèces (Redfern et al., 2006). Dans le cas où la loi choisie fait partie de la famille des distributions

exponentielles, l'équation liant l'espérance de la loi retenue et les différentes variables explicatives s'écrit :

$$g(\mathbb{E}_Y(Y|X)) = \sum_{j=1}^p x_j \beta_j = X_j^T \beta_j$$

Les trois composantes du GLM sont la distribution conditionnelle retenue, la somme pondérée des variables explicatives (par les coefficients β_j) et la fonction de lien g . On observe que par cette formulation, le GLM ne peut modéliser que des relations linéaires entre les variables explicatives et la variable réponse. Pour modéliser des relations non-linéaires, le modèle additif généralisé (aussi appelé « GAM ») a été introduit par Hastie & Tibshirani (1986) et s'est avéré extrêmement efficace pour prédire la distribution des cétacés grands plongeurs, comme l'a démontré une étude réalisée par Virgili et al. (2019). Il permet d'intégrer d'une manière additive des fonctions f_j des variables explicatives tel que :

$$g(\mathbb{E}_Y(Y|X)) = \beta_0 + \sum_{j=1}^p f_j(x_j)$$

Les fonctions f_j des variables explicatives x_j sont modélisées à partir de différents types de splines de lissage. Pour modéliser la relation entre l'espérance de la loi conditionnelle et une variable explicative en particulier, une spline n'est souvent pas suffisante, on doit alors sommer plusieurs splines pour former la fonction finale $f_j(x_j)$. La jointure de ces différentes splines se fait par l'intermédiaire de points appelés nœuds qui influence la forme de la fonction $f_j(x_j)$.

3.2 Acquisition des données

a- Données environnementales

Le choix des variables environnementales est une étape clé dans le bon fonctionnement d'un modèle de prédiction de la distribution. Ces variables environnementales peuvent se diviser en deux grands types : les variables dites physiographiques statiques, c'est-à-dire que leurs caractéristiques ne changent pas dans le temps, et les variables dites océanographiques (ou hydrologiques) dynamiques dont la valeur évolue au fil du temps.

Les variables océanographiques dynamiques

D'un point de vue biologique, la meilleure variable pouvant influencer de manière directe la distribution des cachalots est la distribution de sa proie principale : le calmar d'eaux profondes. Cependant, comme aucune donnée sur la distribution spatiale des calmars n'était disponible pour notre zone d'étude, d'autres données environnementales telles que la chlorophylle A ont été utilisées comme proxy. Cependant, les cachalots ne sont pas directement influencés par les concentrations en chlorophylle a, car il existe un décalage dans les réseaux trophiques entre la production primaire et les céphalopodes (Jaquet, 1996). Néanmoins, comme considéré par Whitehead (2002), la distribution des calmars pourrait être influencée par la production primaire pendant les épisodes de blooms phytoplanctoniques.

Une étude menée par Virgili et al. (2021) suggère une préférence des cachalots pour les habitats associés à des particularités topographiques et des fronts thermiques, attisant l'intérêt de considérer

la température des eaux de surface (SST), la profondeur de la couche de mélange ainsi que les courants océaniques et leurs directions, comme des paramètres qui, liés, peuvent fortement influencer la croissance du phytoplancton. De surcroît, Mannocci et al. (2014) ont montré qu'il existait des relations fortes entre la distribution des cétacés et les variables environnementales dynamiques telles que la SST, la hauteur des eaux de surface (SSH), les courants océaniques (composantes U et V) ainsi que la vitesse moyenne de ces courants (CV). De même, il semblerait que les cachalots soient concentrés majoritairement au niveau des front de tourbillons (Virgili et al., 2021). L'énergie cinétique des tourbillons associée (EKE, « Eddy Kinetic Energy ») peut en fait être calculée à partir de l'équation suivante (1) dans laquelle u et v représentent les composantes respectivement Est et Nord de la vitesse totale (2) :

$$EKE = \frac{1}{2} * (u^2 + v^2) \quad (1)$$

$$Vitesse\ du\ courant\ (CV) = \sqrt{u^2 + v^2} \quad (2)$$

Les cachalots ayant déjà montré une faible dépendance aux variables environnementales de surface, d'autres facteurs environnementaux tels que la température du fond et la profondeur de la couche de mélange ont été pris en compte et permettent d'analyser l'influence d'une variabilité verticale sur la distribution de cette espèce. La profondeur de la couche de mélange est étroitement liée à la productivité primaire car elle détermine le mélange des eaux de surface et des eaux profondes, favorisant ainsi le recyclage des nutriments. Étant de grands plongeurs, les cachalots peuvent également être influencés par la température du fond, car c'est là qu'ils chassent principalement. (Chambault et al., 2021).

Ces variables océanographiques temporelles présentées ci-dessus ont donc été choisies comme potentiels facteurs influençant la distribution des cachalots. Elles ont toutes été moyennées sur la période d'étude, sur une base de données journalières provenant du site Copernicus (<http://marine.copernicus.eu/>). La liste exacte des produits Copernicus et des jeux de données dont elles sont issues sont présentées en Annexe 2.A.

Les variables physiographiques statiques

Concernant les variables statiques, la profondeur (ou bathymétrie), la pente et les distances aux isobathes 200, 1000 et 2000 ont été choisies puisqu'elles se sont révélées pertinentes pour les espèces teutophages (se nourrissant exclusivement de calmars) telles que les cachalots (Mangion and Gannier, 2002). Si la distance à la côte ne fait pas toujours l'unanimité dans le cas de l'étude des cachalots, il a été tout de même décidé de la conserver puisque des observations de cachalots ont eu lieu près des côtes, notamment lorsque le plateau continental est peu étendu, comme c'est le cas en Dominique par exemple. Enfin, la distance aux canyons a été choisie puisqu'elle est représentative des zones fréquemment visitées par les cachalots, selon une étude menée par Moors-Murphy (2014).

Plusieurs de ces variables, entre autres les distances aux isobathes, à la côte et aux canyons, ont été calculées grâce à la fonction « terrain » du package « raster » (Hijmans, 2020) à partir des données de bathymétrie. La bathymétrie, quant à elle, a été extraite du fichier netCDF X, provenant du site GEBCO (General bathymetric Chart of the Ocean; <https://www.gebco.net/>).

Le choix de chaque variable environnementale est explicité dans l'Annexe 3. Dans un premier temps, le modèle tourne avec l'ensemble des variables préalablement décrites, puis ensuite, le choix s'affinera afin de supprimer les variables fortement corrélées entre elles, expliquant ainsi de la même manière la distribution des cachalots.

b- Données de suivi des espèces

Les données de suivi ont été collectées selon le protocole décrit en section II-1-b-Collecte des données. Ces données regroupent les données rentrées sur ObsEnMer permettant d'accéder aux points GPS lors du suivi actif des espèces, donc à l'effort total actif sur les deux années d'étude ; ainsi qu'à la position GPS associée au début d'un suivi d'observation de cachalots, assumée comme représentant la position de chaque individu tout au long de cette observation.

3.3 Prétraitement des données

Calcul de l'effort d'échantillonnage propre au cachalot

Une grille constituée d'hexagones de 1,76 km de côté a été créée sur la zone d'étude (aire d'un hexagone = 7,8 km²). La forme hexagonale, grâce à ces frontières régulières ainsi qu'à sa structure de voisinage équilibrée, permet de minimiser les effets de bords⁷ indésirables qui peuvent être forts dans le cas de carrés (bords anguleux). Afin de ne prendre en compte que l'effort actif spécifique au cachalot (effort d'observation), plusieurs étapes ont été nécessaires :

1^{ère} étape : Etablissement des distances de détectabilité

A partir des 69 observations initiales de cachalot sur les années 2021 et 2022 (Figure 11,A) et de la distance à laquelle chaque groupe/individu a été détecté, une courbe de densité des observations en fonction de la distance de détection a été établie (Figure 11,B).

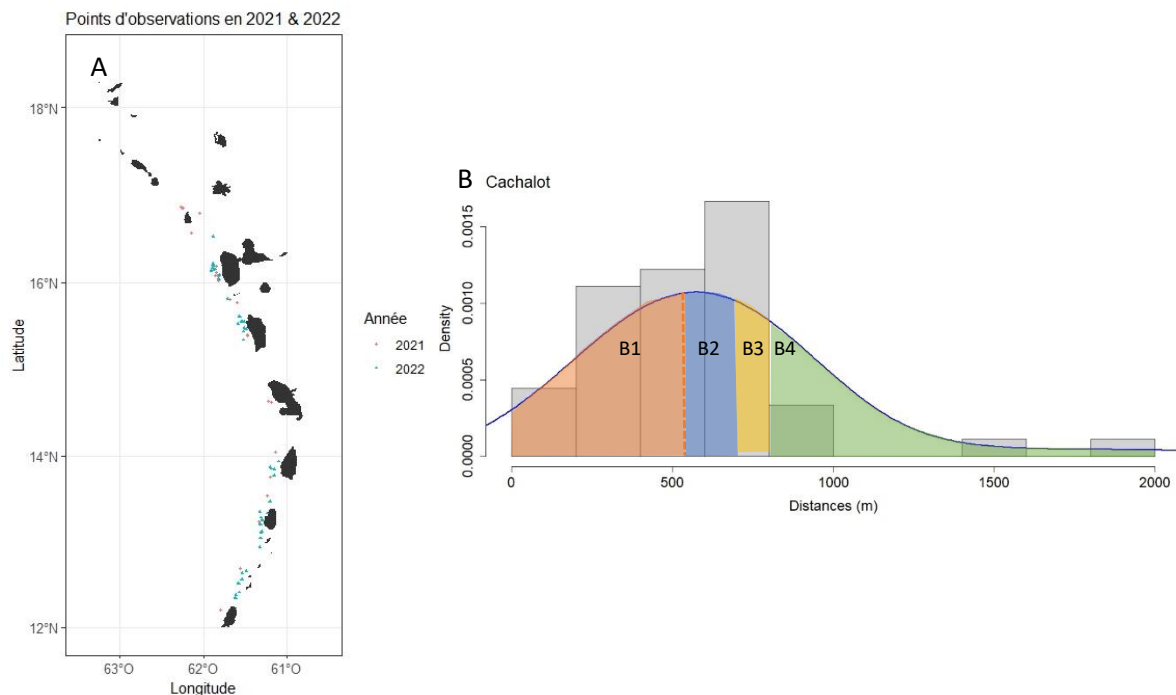


Figure 11: Points d'observations en 2021-2022 (A) et densité d'observations en fonction de la distance à laquelle l'animal a été détecté (B). La ligne orange en pointillé indique un maximum de cachalots détectés autour de 512 m. Les différentes zones colorées représentent les différents buffers et leurs distances associées

⁷ les caractéristiques ou les résultats d'une zone ou d'une entité sont influencés ou modifiés par la proximité de ses frontières ou limites

Cette figure indique un maximum de cachalots détectés autour de 512 m environ, ce qui correspond environ au premier buffer où la probabilité de détection est de 1 (c'est-à-dire qu'on considère qu'on détecte tous les individus entre 0 et 512 m). L'aire sous la courbe associée est d'environ 40%. On cherche ensuite à diviser le reste de l'aire sous la courbe de détection en parts égales de sorte à considérer les probabilités de détection de 0.75, 0.5 et 0.25. Les distances de détectabilité associées à ces probabilités de détection sont respectivement de 700, 800 et 2000 m. Au-delà de 2000m, on considère que les individus ne sont plus détectés. A mesure que la probabilité de détection diminue, l'aire de la zone dans laquelle les cachalots sont détectés augmente.

2^{ème} étape : Création des tampons sur QGIS

A partir de ces distances de détectabilité, les zones tampons autour de la ligne d'effort active ont pu être tracées. Pour chacune des 4 zones tampons, 209 sous-tampons ont été différenciés (relatifs aux 209 journées d'expéditions distinctes). Ces tampons se superposent seulement si le bateau est effectivement passé plusieurs fois au même endroit ; autrement, les 4 zones tampons ont des surfaces distinctes. Chacune d'entre elles se voit associer une probabilité de détection de 1 pour le premier buffer, de 0.75 pour le deuxième, etc.

3^{ème} étape : Union de l'effort d'observation lié à P_m avec notre grille

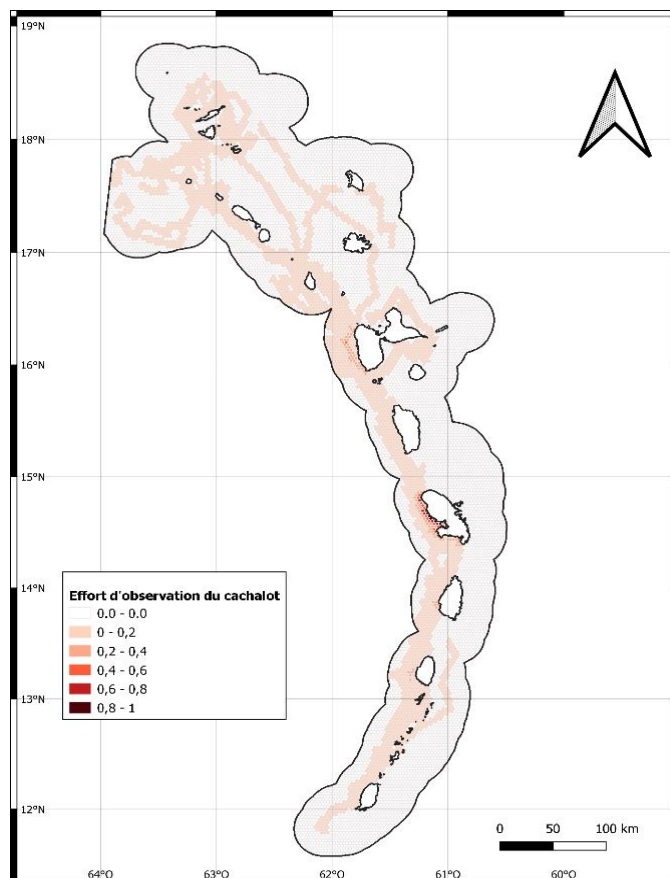


Figure 12: Effort d'observation du cachalot

Ensuite, les tampons sont fusionnés pour former une couche contenant tous les buffers liés aux probabilités de détection, représentant l'effort d'observation total du cachalot. Ensuite, la grille hexagonale est combinée avec cette couche, les buffers hors de la zone d'étude sont supprimés, et l'aire totale⁸ pondérée par la probabilité de détection est calculée pour chaque hexagone afin d'obtenir l'effort d'observation relatif. Le résultat est présenté dans la Figure 12 après normalisation des valeurs par la valeur maximale de l'effort.

L'obtention de cet effort d'échantillonnage est essentielle afin de pondérer les observations dans le modèle de distribution par cet effort, c'est-à-dire de prendre en compte le fait qu'une zone ait été plus échantillonnée qu'une autre, ce qui peut impacter la prédiction finale.

⁸ somme de toutes les entités présentes dans un hexagone, représentée par toutes les fois où le bateau est passé dans cette zone ou qu'une partie d'un tampon se retrouve dans cette zone

Variables environnementales

Le choix a été fait de calculer la valeur moyenne sur les deux années d'études par rapport à celle de la médiane de sorte à ce que les valeurs extrêmes de chaque variable environnementale soient bien représentées. L'écart-type a été sélectionné quand il apportait une information supplémentaire par rapport à la moyenne (Annexe 2.B). Quand il était uniforme sur l'ensemble de la zone d'étude ou alors que sa variation suivait la même tendance que la moyenne, il n'a donc pas été sélectionné.

Lors de l'union de chaque variable environnementale à la grille hexagonale (Annexe 4), la valeur moyenne au sein de l'hexagone a été choisie par rapport à la valeur au centre. Cela permet de prendre en considération le cas où aucune valeur d'une variable environnementale donnée n'est présente au centre de l'hexagone mais que pour autant, une valeur de cette variable environnementale est présente à un autre endroit de l'hexagone. Le choix de prendre la valeur moyenne plutôt que la médiane est là aussi important car il permet de prendre en compte l'ensemble des valeurs (même les plus aberrantes) présentes au sein de l'hexagone.

Observations de cachalots

Chaque point d'observation correspond à la position GPS du moment où le suivi du cachalot a débuté. Ainsi, un seul point GPS par observation est considéré dans un modèle de type GAM. Le jeu de données qui va être utilisé dans le modèle contient les données d'observations de cachalots, l'effort d'observation ainsi que chaque variable environnementale associée à chaque hexagone.

Plusieurs étapes ont été nécessaires pour traiter les données d'observation, ceci étant notamment dû à leur confrontation avec le quadrillage de la zone d'étude ou les variables environnementales. La recherche de valeurs extrêmes⁹ dans les données environnementales a révélé que 5 de nos observations étaient dans des zones de fortes valeurs de courants. En investiguant de manière plus approfondie, il s'avère que ces valeurs ne sont pas inquiétantes puisqu'il s'agit de la vitesse réelle des courants dans cette zone. Ne pas prendre en compte ces valeurs aberrantes peut amener à ne pas considérer une zone aux caractéristiques intéressantes pour les cachalots. Les observations en question ont donc été conservées dans le modèle. Les différentes étapes nécessaires à la prédiction de la distribution suivant un modèle GAM sont reprises dans la Figure 13.

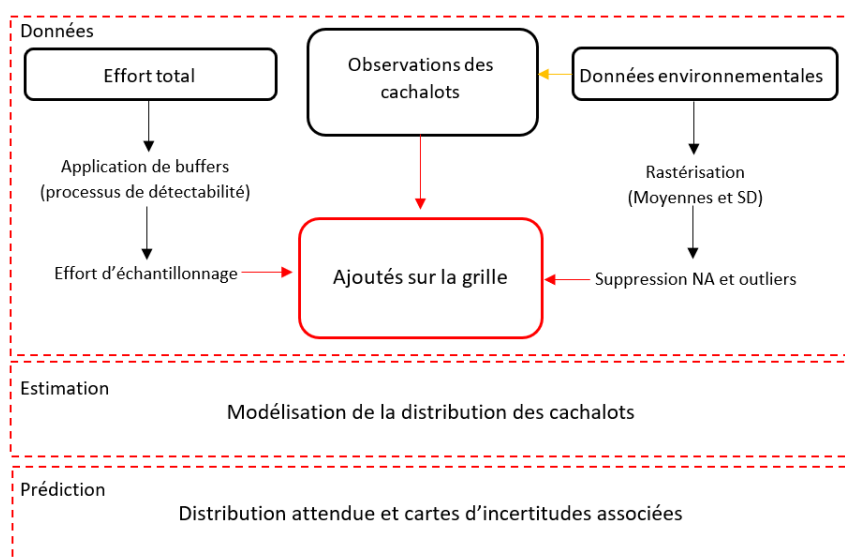


Figure 13: Représentation schématique des étapes nécessaires à la prédiction de la distribution suivant un modèle GAM

⁹ Dont la valeur normalisée excède 3

3.4 Calibration

Afin de calibrer le modèle, plusieurs paramètres sont à prendre en compte, tels que la corrélation maximum entre les variables, la complexité de lissage, le nombre maximum de prédicteurs.

Choix de la distribution

Dans notre étude, notre variable réponse représente le nombre de groupes observés. Nos variables explicatives sont des variables environnementales détaillées dans la section précédente. Étant donné que notre variable réponse est une variable de comptage, on suppose que la distribution conditionnelle de la variable réponse sachant les variables explicatives est une loi de Poisson de paramètre μ , avec une fonction de lien logarithmique. De plus, pour pondérer chacune des observations, il est usuel d'utiliser un terme d'exposition T qui dans notre cas représente l'effort d'observation (Figure 12). La formulation de notre GAM s'écrit donc :

$$\log(\mathbb{E}_Y(Y|X)) = \beta_0 + \sum_{j=1}^p f_j(x_j) + \log(T)$$

Choix de la complexité de lissage

Les splines retenues sont des splines de lissage à plaques minces. La complexité de lissage (paramètre k) pour chaque fonction f_j représente le nombre de nœuds utilisés (endroits où la courbe se coupe en deux sous-portions de courbe). Par défaut, lorsque ce paramètre est libre, sa valeur est généralement entre 8 et 10. Plus sa valeur augmente, plus la courbe de lissage va être ajustée à la distribution initiale de nos données d'observations. Or, l'objectif de cette étude est d'établir un modèle transférable à l'échelle des Petites Antilles et non pas de capturer la particularité de nos données. Afin de résoudre ce problème de sur-adaptation de la courbe de régression à nos données, la complexité de lissage a été limitée à 4 (S. N. Wood, 2006).

Choix de la méthode de sélection des lissages

Au départ, le degré de lissage était ajusté à l'aide d'une méthode de validation croisée qui exclut une donnée à chaque itération, puis estime la capacité moyenne des modèles, construits sur les données restantes, à prédire la donnée exclue. Néanmoins, Simon Wood, créateur de la bibliothèque `mgcv` dédiée à l'utilisation des modèles GAM sur R, a observé que l'utilisation de la méthode REML (Restricted Maximum Likelihood) était nettement plus efficace pour l'estimation des différents paramètres de lissage (les coefficients β_j) et de la loi conditionnelle liant la variable réponse aux variables explicatives. Le choix a donc été fait d'utiliser une méthode de type REML (S. Wood, 2012). Pour plus d'information sur cette méthode, nous référons le lecteur à la section 6.2.5 de ce livre.

Choix du maximum de corrélation

Le maximum de corrélation entre les variables a été fixé à 0.7 (Virgili et al., 2019) de sorte à ce que deux variables qui seraient corrélées à plus de 0.7 ne puissent pas être présentes en même temps dans un même modèle de distribution. En effet, prendre en compte la combinaison de ces deux variables fortement corrélées pourrait altérer la qualité d'ajustement du modèle car aucune interaction n'est prise en compte dans le modèle. Le test de corrélation utilisé ici est celui de Pearson.

Choix du nombre de prédicteurs maximal

Le nombre maximal de prédicteurs choisis ici pour prédire la distribution des cachalots est de 4. Limiter le nombre de prédicteurs à 4 permet de réduire la complexité du modèle, ce qui facilite l'interprétation des résultats. Cela permet également de réduire le risque de surajustement donc une meilleure généralisation du modèle aux nouvelles données. Enfin, choisir un nombre réduit de prédicteurs

permet d'éviter les effets d'interaction entre les variables explicatives et donc de mieux isoler les effets individuels des variables explicatives.

Choix du meilleur modèle

Afin d'établir quel était le modèle qui décrivait le mieux nos données, le critère d'information d'Akaike (AIC) a été utilisé (Akaike, 1974). Il évalue à la fois le maximum de vraisemblance du modèle avec les données (c'est-à-dire l'écart minimal de chaque donnée d'observation aux courbes splines sélectionnées pour chaque variable) et sa parcimonie, c'est-à-dire qu'il pénalise les modèles faisant intervenir un nombre trop important de paramètres. En outre, les modèles qui prennent en compte un plus grand nombre de variables environnementales pour améliorer leur ajustement aux données seront pénalisés. Cela permet notamment de s'assurer que le meilleur modèle ne soit pas un modèle sur-adapté aux données initiales mais bien celui qui prédise le mieux la distribution possible des cachalots par rapport aux données manquantes. On considérera comme meilleur modèle celui dont l'AIC est le plus faible et également tous les modèles dont l'écart avec l'AIC le plus faible est inférieur à 2.

3.5 Validation

La précision statistique des prédictions du modèle a été évaluée en comparant les probabilités de présence prédites par la Figure 24 et les données réellement observées du jeu de validation. Ainsi, les données de prédiction sur la zone des 20% sont extraites depuis les prédictions faites à partir des données réellement observées et des quatre meilleurs prédicteurs du jeu de données de calibration. Ces données prédites sur la zone du jeu de données de validation sont ensuite comparées aux données réellement observées sur cette zone, ce qui permet de tester la robustesse du modèle et sa transférabilité à d'autres jeux de données.

La performance du modèle a été évaluée selon une méthode de validation communément utilisée : la matrice de confusion. Cette dernière compare les données prédites avec les données réelles d'absence et de présence (Fielding & Bell, 1997). Ces données de présence et d'absence ne sont pas ici des données binaires mais bien des probabilités dont la valeur varie entre 0 et 1. Ainsi, le seuil généralement utilisé pour différencier les présences des absences de 0.5 doit être recalculé dans notre cas. Pour cela, la probabilité prédite du point pour lequel la distance, entre la Receiving Operating Characteristic (ROC) et la première bissectrice, est maximale a été utilisée. Cette distance a été calculée à partir du package ROCR de R (Sing et al., 2005).

A partir de cette matrice de confusion, deux métriques complémentaires résumant la performance du modèle peuvent être calculées : la sensibilité (taux d'individus positifs correctement prédits) et la spécificité (taux d'individus négatifs correctement prédits). Elles traduisent les capacités du modèle à détecter l'ensemble des individus réellement présents et l'ensemble des individus réellement absents. Ainsi, si la sensibilité et la spécificité sont élevées, alors on peut avoir confiance dans le fait que le modèle prédira bien les présences et absences de cachalots respectivement.

La qualité de l'ajustement des données prédites par rapport aux données réelles a également été évaluée par le biais du calcul de l'aire sous la ROC courbe (AUC), qui donne une idée générale de la performance du modèle (Fielding & Bell, 1997). Un modèle parfait a un AUC de 1 tandis qu'un modèle aléatoire a un AUC de 0.5. Le plus proche l'AUC est de 1, meilleure est la prédiction faite (Praca et al., 2009).

III- Résultats

1- Identification à l'échelle de l'individu, détermination de l'abondance et de la distribution

1.1 Photo-identification

Sur un nombre total de 1988 photos, 1078 contenaient des nageoires caudales de cachalots et ont donc été sélectionnées pour être soumises à Flukebook. Sur les 69 observations (27 en 2021 et 42 en 2022), 41 rencontres (soit 31 journées distinctes, Annexe 5) ont pu être utilisées. Cela correspond à deux types de situations :

-celle où aucune photo n'a été prise lors de l'observation due à une durée trop courte d'observation, à de mauvaises conditions météorologiques, ou à la rencontre d'une autre espèce en même temps,

-celle où des photographies ont été prises mais aucune n'était de qualité suffisante pour permettre l'identification des individus (individus trop loin, mauvaise visibilité, photos floues, etc).

Sur l'ensemble des rencontres restantes, on estime qu'en moyenne 2 individus ont été loupés (avec au minimum 0 individu loupé et au maximum 8 individus loupés par rapport à la meilleure estimation). Sur ces 41 rencontres, 100 cachalots ont été observés. Leur photo-identification via Flukebook a permis de révéler l'existence de 62 individus dont 1 juvénile et 1 mâle. Sur les 62 individus, 20 individus ont été recapturés sur les deux années d'étude (soit la même année, soit sur deux années différentes). 45 de ces individus avaient déjà été capturés par DSWP et Felicia Vachon. L'ensemble des individus recapturés sont présentés dans le Tableau 2 ci-dessous, qui présente la correspondance entre le nom de l'individu dans le catalogue CCS et son nom dans les catalogues déjà existants (catalogue de DSWP, catalogue Vachon (2022)).

Tableau 2: Tableau des individus de DSWP recapturés par la CCS

Nom CCS	Nom DSWP/Felicia	Nom CCS	Nom DSWP/Felicia	Nom CCS	Nom DSWP/Felicia
CCS_Pm_0001	6058 (Canopener)	CCS_Pm_0025	5562 (Knife)	CCS_Pm_0042	6016
CCS_Pm_0002	5560 (Pinchy)	CCS_Pm_0026	6465	CCS_Pm_0045	6196
CCS_Pm_0003	5722 (Fingers)	CCS_Pm_0027	6426	CCS_Pm_0046	6205
CCS_Pm_0004	6219 (Digit)	CCS_Pm_0030	6313	CCS_Pm_0047	6197
CCS_Pm_0005	6348	CCS_Pm_0032	6335	CCS_Pm_0048	6408
CCS_Pm_0007	5163 (Tooth)	CCS_Pm_0033	6331	CCS_Pm_0049	6411
CCS_Pm_0009	5126 (1674)	CCS_Pm_0034	6332	CCS_Pm_0050	6466
CCS_Pm_0010	5946 (Tabitha)	CCS_Pm_0035	6473	CCS_Pm_0052	6413
CCS_Pm_0011	6093 (Top)	CCS_Pm_0036	5144 (Ziggy)	CCS_Pm_0053	6409
CCS_Pm_0012	5585 (Tip)	CCS_Pm_0037	5162	CCS_Pm_0054	6247
CCS_Pm_0013	6242	CCS_Pm_0038	5588	CCS_Pm_0055	6237
CCS_Pm_0020	5740 (Nova)	CCS_Pm_0039	6088	CCS_Pm_0058	6022
CCS_Pm_0023	5127 (Zandoli)	CCS_Pm_0040	5586	CCS_Pm_0059	6230
CCS_Pm_0024	5151 (Fork)	CCS_Pm_0041	6017	CCS_Pm_0062	6377
CCS_Pm_0021	5147 (Niva)	CCS_Pm_0043	5575	CCS_Pm_0061	6422

Sur les 45 individus déjà capturés par DSWP, 28 avaient été capturés moins de 10 fois (Annexe 5). Sur ces 28, la CCS a participé à plus de la moitié des recaptures dans 5 cas (individus en vert dans le tableau) et a permis d'augmenter le taux de recapture d'en moyenne 46% pour les 23 individus restants.

Toutes données confondues, le taux de recapture moyen a été évalué à 1,51/an par individu.

Les 17 individus restants, ont été jugés comme nouveaux au regard de leur absence dans les catalogues déjà existants. Ces derniers sont présentés dans la planche photo ci-après (Figure 14).



Figure 14: Planche photos présentant les 17 nouveaux individus découverts par la CCS

Ces nouveaux individus présentent un taux de recapture moyen sur les deux années d'étude de 0.64 recaptures/an, avec 1 individu recapturé 3 fois, 3 individus recapturés 2 fois et le reste des individus capturé une fois.

Sur l'ensemble de la zone d'étude, en prenant en compte seulement les données de captures et recaptures de la CCS, certaines zones regroupent un plus grand nombre d'individus différents. Sur la Figure 15, le nombre total d'individus aperçus autour de chaque île a été représenté à l'aide de ronds de taille différente. Si un individu avait été vu plusieurs fois autour de l'île, il n'a donc été compté qu'une fois.

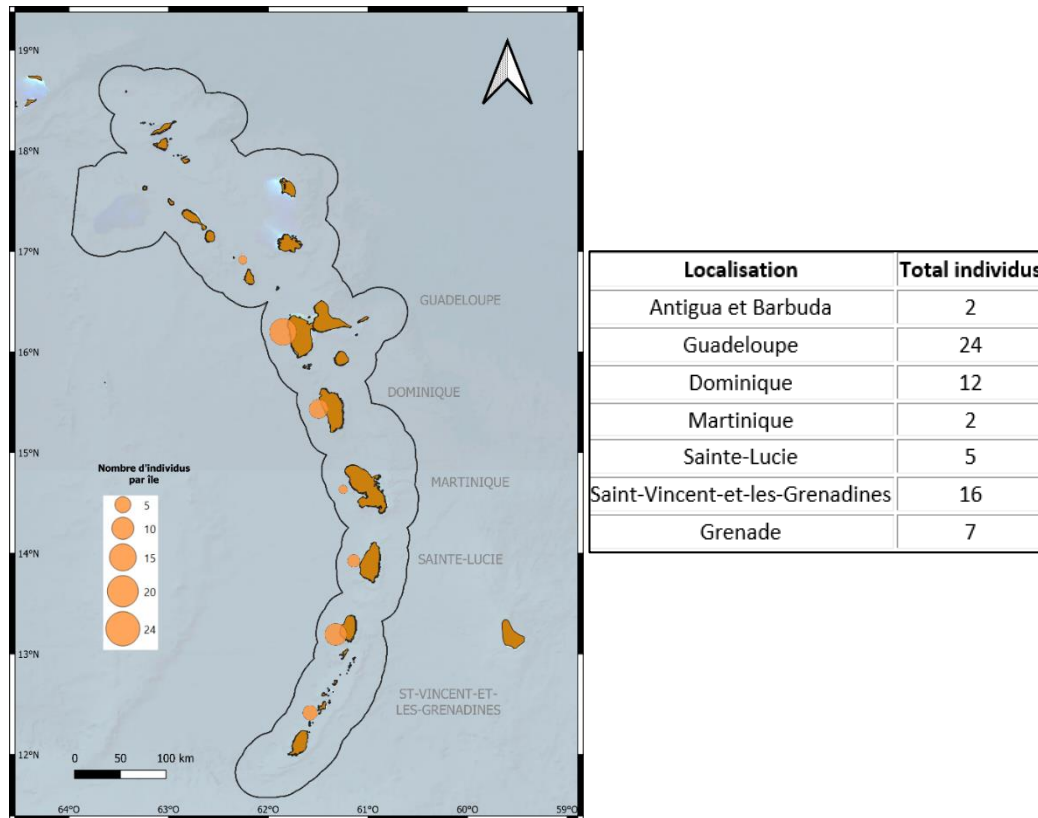


Figure 15: Carte représentant le nombre d'individus différents aperçus par île (ronds oranges)

Les 68 individus différents observés sur toutes les îles sont répartis surtout entre la région Sud (41%) et la région Centre (56%). Peu d'individus ont été observés dans la région Nord pour un total de seulement 2 individus distincts à Antigua et Barbuda. Sans surprise, les îles de la Guadeloupe et de la Dominique, qui sont le terrain de jeu de la plupart des cachalots observés par DSWP depuis 2005, présentent à elles deux, plus de la moitié du nombre d'individus aperçus. Les observations dans la région Sud sont majoritairement au niveau de Saint-Vincent, qui est la deuxième île de l'arc avec le plus d'individus différents observés selon nos données.

1.2 Détermination de l'abondance

Dans un premier temps, l'estimation de la population a été réalisée à l'aide des deux modèles présentés en II.2.4, avec uniquement les données collectées par la CCS en 2021 et 2022 sur les 60 individus adultes. Ces résultats ont ensuite été comparés avec ceux obtenus par Vachon (2022, p77) sur 213 individus adultes. Enfin, les données de Vachon (2022) et de la CCS ont été mises bout à bout afin d'obtenir une estimation plus robuste. Cette jointure entre deux jeux de données différents a été possible dans la mesure où les méthodes d'échantillonnage utilisées étaient similaires. Les résultats des différents tests sont présentés dans le tableau ci-dessous (Tableau 3).

Tableau 3: Estimation de l'abondance via deux méthodes avec les données de la CCS uniquement, les données de Vachon (2022), et les deux réunies

Données	Echelle	Nombre d'individus considérés	Méthode	Estimation population	SE	IC à 95%
CCS (2021-2022)	Années, individus	60	Peterson fermée	111.0	24.0	79.6 - 158.4
			Schnabel fermé	116.8	34.8	88.6 - 231.5
Felicia (2019-2020)	Années, individus	213	Peterson fermé	413.6	46.1	337.7 - 489.5
			Schnabel fermé	419.03	42.3	348.4 - 562.1
CCS + Felicia (2019-2022)	Années, individus	238	Schnabel fermé	384.81	24.9	342.9 - 434.3

L'estimation de l'abondance en utilisant les données de la CCS uniquement sous-estime la taille de la population estimée par Vachon (2022) en étant de 116.8 (contre 419.03 (Vachon, 2022)).

Or, la première définition d'une recapture fixait comme recapture, toute nouvelle observation de cet individu, peu importe le temps écoulé entre les deux observations. Ainsi, cette définition amenait à considérer qu'un individu avait été recapturé s'il avait été vu deux fois au sein de la même journée ou au sein de la même expédition. Afin d'évaluer un nombre de recaptures plus représentatif, il a été décidé de fixer la durée minimale entre deux recaptures à au moins deux semaines d'intervalle, de sorte à ce qu'un individu déjà vu lors d'une expédition et revu quelques jours plus tard dans ces mêmes eaux ne soit pas considéré deux fois (Tableau 4). Cette méthode a uniquement été appliquée afin d'étudier les effets sur les estimations d'abondance et non pas dans le reste de l'analyse.

Tableau 4: Comparaison des individus totaux identifiés et recapturés en fixant une durée minimale entre deux recaptures ou en laissant cette durée libre

Organisation	Nombre d'individus totaux identifiés	Durée minimale de recapture libre	Durée minimale de recapture fixe (15j)
CCS	60	21	17
Vachon	213	148	63

En ne fixant aucune durée minimale entre deux recaptures, en moyenne, 70% des individus sont recapturés par Vachon (2022) tandis que cette moyenne chute à seulement 30% en prenant en compte les 15 jours minimum entre deux recaptures. Concernant la CCS, la variation est minime entre les deux méthodes utilisées, et rejoint le taux de recapture de Vachon (2022) puisqu'environ 1 individu sur 3 est recapturé. Les résultats de l'estimation recalculée en prenant en compte un intervalle de recapture minimal de 15 jours confirment cela, puisqu'aucune différence significative n'est obtenue avec les résultats précédents. En outre, les résultats par la méthode de Schnabel donnent toujours une estimation moyenne de 419 individus pour Vachon et modifient légèrement à la hausse l'estimation précédente de la CCS en passant de 117 à 126 individus. Les taux de recapture entre Vachon (2022) et la CCS sont donc similaires. Ainsi, de manière exploratoire, une carte a été réalisée sur laquelle plusieurs données ont été représentées :

-d'une part, les individus de Vachon (2022) non recapturés par la CCS,

-d'autre part, les individus de Vachon (2022) recapturés par la CCS et ayant été observés par Vachon (2022) en 2019 ou 2020

-l'effort de la CCS permettant de comprendre une potentielle différence de la zone totale échantillonnée entre la CCS et Vachon (2022).

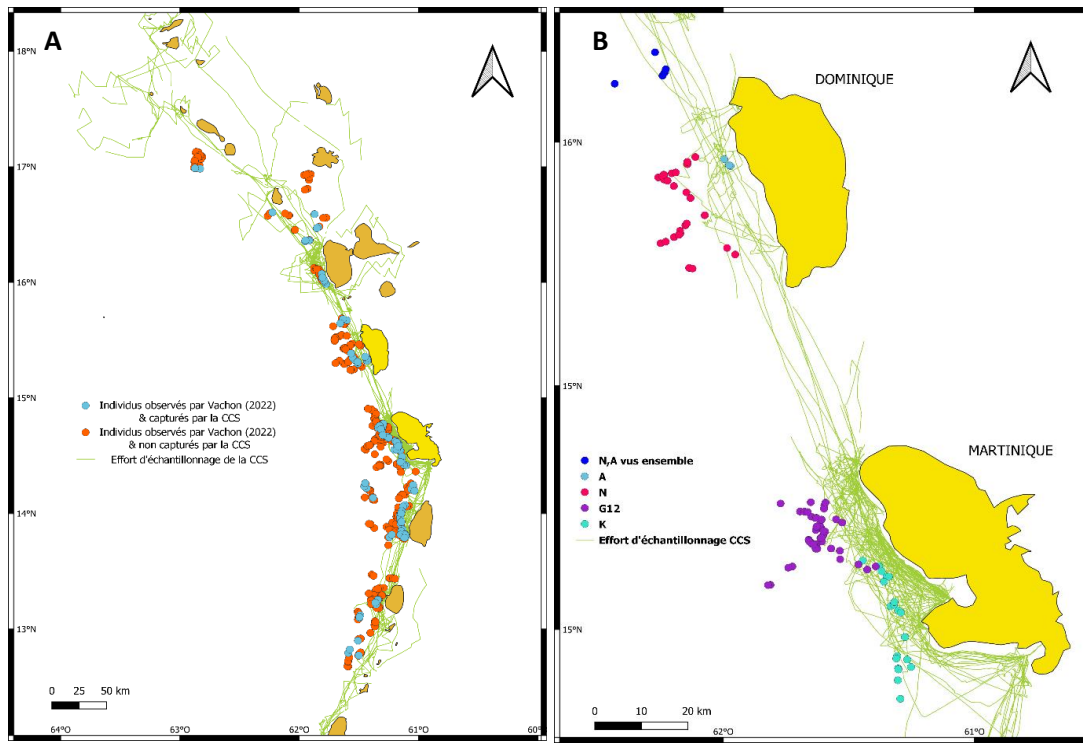


Figure 16: Carte représentant les individus observés par Vachon (2022) et capturés ou non par la CCS (respectivement points bleus et oranges) (A) et carte zoomé sur les points d'observations des unités A,N, G12 et K en Dominique et Martinique (B)

Cette carte (Figure 16.A) laisse apparaître deux zones distinctes : une près des côtes (zone 1), échantillonnée par la CCS, où on retrouve la majorité des observations de Vachon (2022) correspondant à ses individus recapturés par la CCS ; l'autre, plus au large (zone 2), qui comprend majoritairement toutes les autres observations des individus de Vachon (2022) non recapturés par la CCS. La Figure 16.B présente plusieurs unités sociales et leurs aires de répartition lorsqu'elles ont été vues seules ou en groupe.

2- Identification des liens sociaux

Dans un premier temps, le jeu de données a été divisé en groupes, en utilisant comme variable de regroupement le numéro de l'observation (« Grp-code » unique à chaque observation). Tous les individus qui ont été vus en même temps sur la période d'échantillonnage (heures ou jours en fonction de la durée de l'observation) seront alors considérés dans le même groupe sur cette période. L'indice d'association a ensuite été calculé à partir du regroupement des individus sur une même période d'échantillonnage. Les individus sont considérés associés (association=1) dans la période d'échantillonnage s'ils ont été trouvés au moins une fois au sein d'une observation. Ils sont considérés comme non-associés (association=0) s'ils n'ont jamais été vus ensemble dans une même observation sur les deux années d'échantillonnage.

Les associations entre individus ont d'abord été examinées sous l'angle de la modularité. Cela a notamment permis d'affiner le choix des individus qui apparaîtront par la suite dans la représentation

hiérarchique des liens entre les individus par un arbre d'associations. Les résultats sont présentés dans le Tableau 5 où « no » représente le nombre d'occurrences sur lequel se base le regroupement en cluster (c'est-à-dire le nombre de fois où au moins deux individus appartenant au groupe ont été vus ensemble). La modularité de cet arrangement est de 0.823, donc cet arrangement est considéré comme indiquant une division utile de la population (Newman, 2004).

Tableau 5: Regroupement des individus en cluster et attribution potentielle à une unité sociale déjà connue. « no » représente le nombre d'occurrences sur lequel se base le regroupement en cluster. Les groupements marqués en bleu dans le tableau ci-dessus sont ceux pour lesquels les individus regroupés n'ont été vus ensemble qu'une seule fois. Les groupements en vert représentent ceux pour lesquels les valeurs propres étaient égales ou inférieures à 0. Les individus marqués en orange sont les nouveaux individus appartenant potentiellement à une unité déjà décrite.

Individus	Valeur propre moyenne	Cluster (no)	Groupe/unité sociale déjà connus	Clan vocal
28, 6408, 6409, 6411, 6413, 6426, 6465, 6466	0.26	1 (2)	G15, G17	EC2
5946, 6093, 5126, 5163, 5585, 6242	0.29	2 (8)	T, V	EC1
5740, 5147, 22, 5127, 5144, 6196, 6197, 6205	0.02	3 (2)	Z, A, N	EC1
14, 15, 17, 18	0.41	4 (1)	Aucun	EC1
5151, 5560, 5562, 5722, 6058, 6219, 8	0.32	5 (3)	U, F	EC1
29, 57, 6331, 6332, 6335	0.32	6 (2)	G7	EC1
60, 6022, 6230	0.53	7 (2)	L	EC1
19, 5575	0.50	8 (1)	Aucun	EC2/EC3
6, 6348	0.50	9 (1)	G10	EC1
6016, 6017	0.52	10 (1)	K	EC2
5586, 6088	0.50	11 (1)	ATW	EC1
6422, 6377	0.58	12 (1)	G12	EC3
16, 44, 5162, 5588	-0.29	13 (3)	V	EC1
6237, 6247	-0.50	14 (1)	X	EC2
31, 6313	-0.50	15 (1)	G18	EC1

Les groupements marqués en bleu dans le tableau ci-dessus sont ceux pour lesquels le nombre d'occurrences sur lequel se base le regroupement en cluster est égal à 1, c'est-à-dire que les individus regroupés n'ont été vus ensemble qu'une seule fois. Cela correspond notamment à des petits regroupements (2 à 4 individus maximum). Ces groupements ne seront pas représentés dans le dendrogramme.

Les groupements en vert représentent ceux pour lesquels les valeurs propres étaient égales ou inférieures à 0, rendant l'association des individus en groupement incertaine. En effet, les valeurs propres correspondent au vecteur propre correspondant à la dernière bifurcation impliquant cet individu. Les valeurs proches de zéro indiquent une incertitude dans l'affectation de l'individu. Les individus marqués en vert seront donc également exclus du dendrogramme.

Les individus 6473 et 56 ont été vus seuls, une seule fois uniquement et n’ont donc été regroupés avec aucun autre.

Au total, 17 unités sociales/groupements déjà connus ont été identifiés (Annexe 7). En considérant la taille d’un groupe moyen entre 7 et 9 individus (Gero et al., 2014), seuls 2.5 individus de chaque groupe ont été capturés par la CCS. Ce nombre varie beaucoup d’une unité sociale à l’autre avec par exemple, 1 seul individu capturé sur les 9 de l’unité sociale ATW ou encore, tous les individus capturés des unités sociales U et F. Les unités ou groupements les plus vus sont présentés en Annexe 8.

En prenant en considération les différents ajustements décrits précédemment, la représentation en arbre hiérarchique de l’organisation sociale des cachalots a donc pu être établie. Les deux méthodes de clustering décrites en II.2.4 ont été testées préalablement afin de déterminer celle qui décrivait le mieux les données (Tableau 6).

Tableau 6: Comparaison du résultat de deux méthodes de clustering

Méthode utilisée	Wards	Average Linkage
Coefficient de corrélation cophénétique	0.72	0.93

La méthode de « Average Linkage » étant la seule donnant un résultat supérieur à 0.8 (Bridge, 1993), seul le dendrogramme issu de cette méthode est interprétable. La modularité associée est de 0.74.

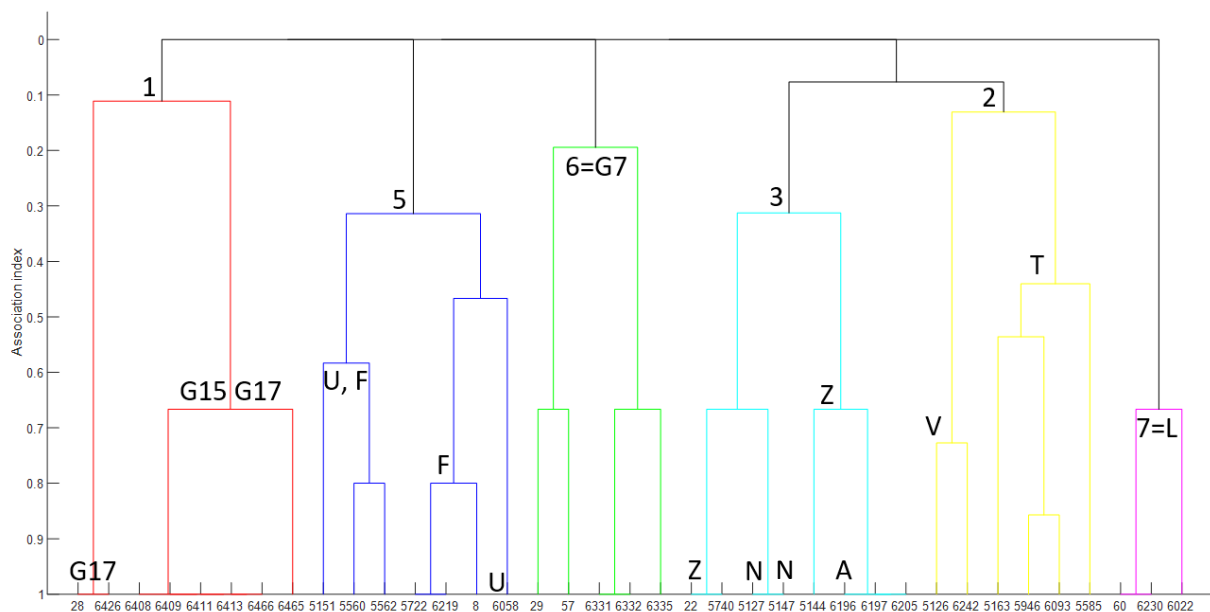


Figure 17: Dendrogramme des différents regroupements possibles des individus et les unités sociales déjà connues

Au niveau des premiers nœuds du dendrogramme (Figure 17), six groupes majoritaires apparaissent, qui se subdivisent en 11 unités sociales déjà décrites. Parmi les 6 regroupements réalisés, deux d’entre eux correspondent à des groupements/unités sociales déjà établis (groupe 6= groupement G7 découvert par Vachon (2022) ; groupe 7 = unité sociale L mise en évidence par le DSWP). Le reste des groupes correspond au rassemblement des unités sociales suivantes : les unités F & U ; les unités T & V ; les unités Z, N & A ; les groupements 15 & 17.

Ce dendrogramme ainsi que le regroupement en cluster via l’indice de modularité permettent d’étudier de façon plus appliquée la relation entre les individus et notamment la reconnaissance d’unités sociales déjà connues. Certains individus nouvellement découverts ont été regroupés au sein

de ces unités ou groupements et pourraient donc constituer de nouveaux individus de ces unités. Pour vérifier cela, le Tableau 7 fait figurer chaque nouvel individu (« ni » représentant le nombre de fois où l'individu a été vu en 2021, et en 2022), en compagnie de quel individu il a été vu, le nombre de fois où ces individus ont été vus ensemble, et les unités sociales des individus connus. Seuls les nouveaux individus qui ont été vus une seule fois mais en compagnie d'au moins deux individus, ou plusieurs fois en compagnie d'au moins un individu sont étudiés ici Tableau 7: Détermination de l'appartenance potentielle des nouveaux individus à un groupement/unité sociale (Tableau 7).

Tableau 7: Détermination de l'appartenance potentielle des nouveaux individus à un groupement/unité sociale. Les lettres colorées en orange constituent les groupes potentiels auxquels seraient rattachés les nouveaux individus. « ni » représente le nombre de fois où l'individu a été vu en 2021, et en 2022.

Individu (ni21, ni22)	Nombre de jours maximal entre les recaptures	Identifié en compagnie de	Nombre de fois où les individus ont été vus ensemble	Type de l'association	Unité(s) sociale(s) des individus déjà connus	Groupe supposé
22 (1,0)		5946, 5740, 5147, 5127	(1,1, 1,1)	Incertaine	T, N, N, Z	
28 (0,1)		6426, 6465	(1,1)	Incertaine	G17	
8 (0,3)	114	5562, 5560, 5151, 5722, 6219, 6058	(2,2, 1,2, 2,1)	Certaine	U, F, U, F, F, F	U, F
16 (1,1)	406	44, 5588	(1,1)	Incertaine	V	
44 (1,0)		16	(1)	Incertaine		
29 (0,2)	84	57, 6335	(1,1)	Incertaine	G7	
57 (0,1)		29, 6335	(1,1)	Incertaine	G7	
60 (0,2)	<1	6022, 6230	(1,2)	Certaine	L	L

L'association a été évaluée comme incertaine si l'individu a été vu une seule fois sur les deux ans (cas des individus 22, 28, 44 et 57), ou une seule fois en compagnie de chaque individu (cas des individus 16 et 29). Pour les deux individus 8 et 60, l'association a été évaluée comme certaine puisque ces individus ont été vus au moins deux fois en compagnie d'au moins un même individu. Les unités sociales déjà connues auxquelles ils pourraient potentiellement appartenir sont présentées en orange dans le Tableau 7.

3- Identification des clans vocaux & aires de répartition

Sur un nombre total de 603.8h d'effort acoustique où les cachalots avaient potentiellement été entendus (sans compter les expéditions où aucun enregistrement n'était disponible), des codas ou clics étaient présents sur 145h dont 98.6h en 2021 et 46.4h en 2022. Grâce au travail d'écoute et de visionnage sur Audacity, 1097 codas ont été repérés acoustiquement. Sur les 33 journées d'enregistrements, 24 présentaient des codas pour un total d'environ 12,3 heures d'enregistrements (Annexe 1). Leur détection se base sur les clans vocaux précédemment documentés EC1, EC2 et EC3. Les codas identitaires majoritairement identifiés pour ces clans sont le coda 1+1+3 pour le clan EC1, 5R (pour 5 Regular clics) pour le clan EC2 et 9R pour le clan EC3. Les spectrogrammes relatifs aux codas EC1 et EC2 sont présentés en Figure 18. 13 groupements/unités sociales ont participé à l'émission de ces codas identitaires qui ont rendu possible leur identification.

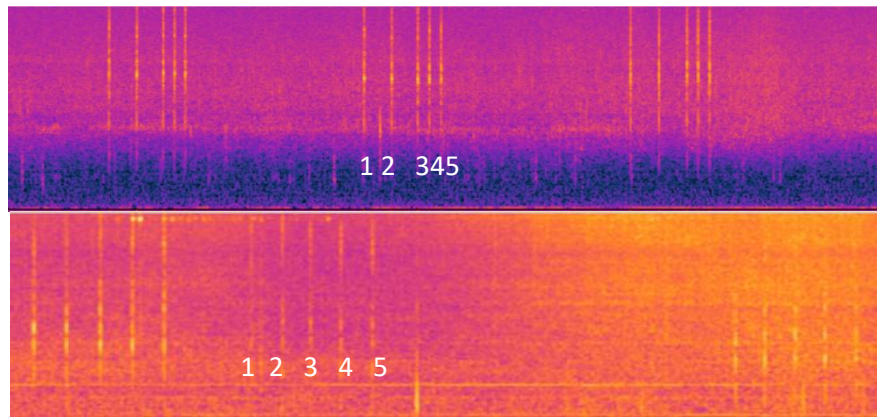


Figure 18: Extraits de spectrogrammes faisant apparaître les deux types de codas EC1 (1+1+3) & EC2 (5R)

Grâce aux 18 journées où des enregistrements étaient présents en même temps qu’une identification visuelle, l’appartenance certaine de 32 individus (dont 8 nouveaux individus) à un clan vocal a pu être déterminée. Certains de ces individus pour lesquels l’appartenance à un clan vocal est certaine ont été vus à d’autres moments en compagnie d’individus pour lesquels l’attribution d’un clan vocal n’avait pas pu être faite en raison d’enregistrements manquants ou de la non-présence de codas pour ces journées d’observation. Lorsque deux individus ont été vus ensemble dans un même groupe d’individus, il a été considéré qu’ils appartenaient à un même clan vocal. Cette méthode a permis d’attribuer à 15 individus supplémentaires un clan vocal (dont 5 nouveaux individus). Ainsi, grâce aux enregistrements de la CCS en 2021-2022, 47 individus se sont vus attribués un clan vocal. Ensuite, les données de DSWP et Vachon (2022) ont permis de venir vérifier et compléter l’attribution à un clan vocal réalisée grâce aux données de la CCS. Cela a notamment permis d’attribuer à 13 autres individus un clan vocal. Finalement, 2 individus seulement n’ont pas de clan vocal certain : l’individu 63, vu une unique fois, seul le 26/03/2022 et l’individu 19, vu une unique fois et en même temps que l’individu 5575 dont le clan vocal est toujours inconnu. Pour ce dernier individu, l’appartenance à un clan vocal reste incertaine puisque les enregistrements liés à son observation font apparaître les deux clans EC2 et EC3. L’appartenance de chaque individu à un clan vocal est résumée dans le Tableau 5. Au total, sur les 62 individus identifiés par la CCS, 45 appartiennent au clan vocal EC1, 12 à EC2 et 2 à EC3.

Finalement, en considérant les 32 rencontres avec les cachalots (associées ou non à une observation), 22 d’entre elles font apparaître le clan vocal EC1, 1 le clan vocal EC2, et 1 contient un mix de codas inconnus (Annexe 1). Plusieurs journées présentent plusieurs clans en même temps telles que le 25/05, 23/06 et 07/07/21 où EC1 et EC2 ont été entendus ensemble (Tableau 8). Une autre journée contient à la fois EC2 et EC3 (23/09/22).

Tableau 8: Descriptif des enregistrements où deux clans ont été entendus : durée totale d’enregistrement & (nombre de codas associé à cette durée entre parenthèses)

Journée	Observation	Durée totale	Durée EC1	Durée EC2	Durée EC3
26/05/21	OUI	12min et 27s (218)	12min et 16s (204)	11s (14)	
23/06/21	NON	5min et 10s (58)	3min et 57s (46)	1min et 13s (12)	
07/07/21	OUI	14min et 48s (272)	14min et 35s (269)	13s (3)	
23/09/22	OUI	34 s (10)		22s (6)	12s (4)

La répartition de ces différents clans vocaux sur l’arc insulaire des Petites Antilles est présentée en Figure 19 (A). Cette répartition prend en considération tous les enregistrements (associés ou non à une observation) pour lesquels des codas identitaires ont été détectés. Afin d’être le plus conservatif possible, pour chaque journée où des codas identitaires ont été repérés, les positions GPS de tous les

points entre la première heure et la dernière heure de détection de clics ou codas associés à un même groupe ont été considérées pour l'établissement de cette carte. Par exemple, lorsque des codas ont été détectés entre 15h32 et 16h05 mais que les cachalots émettant ces codas avaient été repérés dès 9h54 par le biais de clics et ont continué d'émettre des clics jusqu'à 17h01, l'ensemble des positions GPS relatives à la présence de ces cachalots a été pris en compte dans l'établissement de la carte.

Le modèle de densité de Kernel attribue une valeur de densité à chaque cellule de la grille sur laquelle la carte est basée. Plus la valeur est élevée, plus la couleur est foncée et plus la densité est grande dans cette zone.

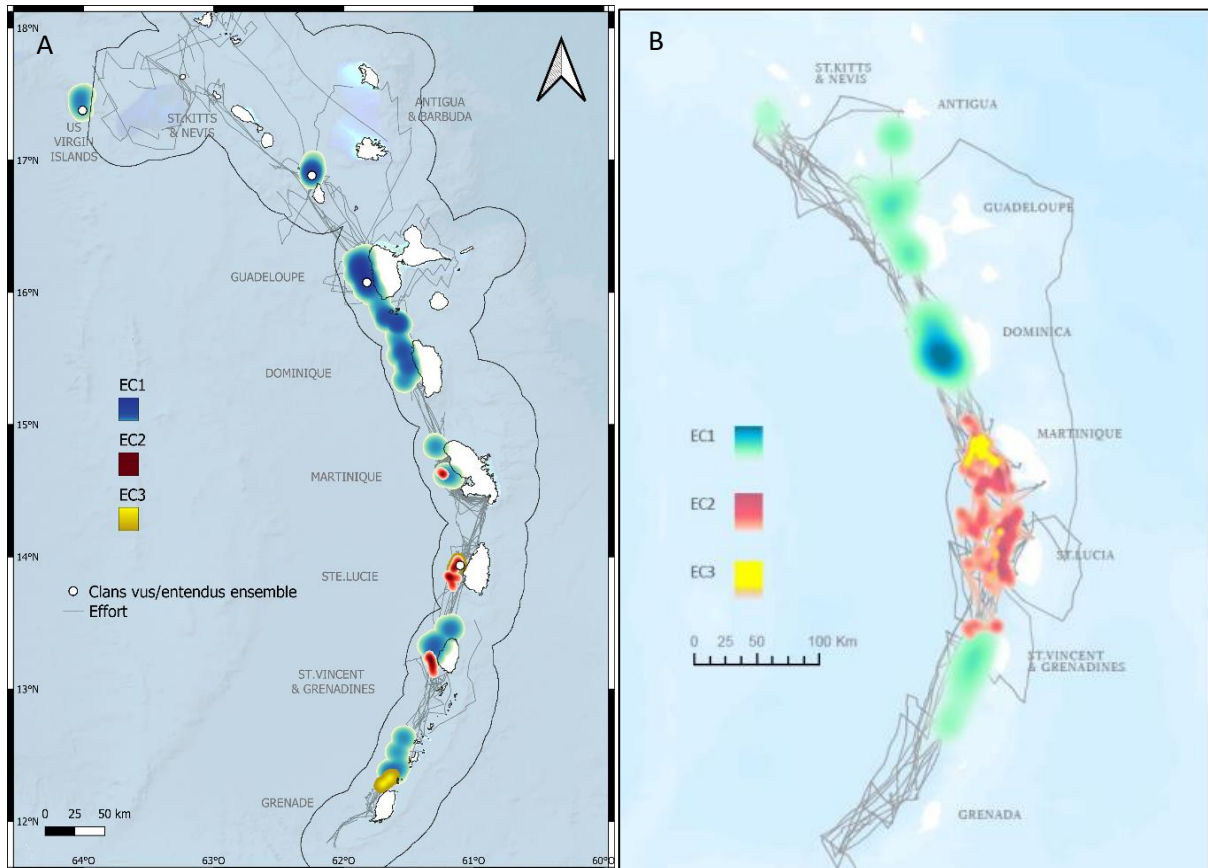


Figure 19: Distribution des différents clans vocaux sur la zone d'étude avec données CCS (A) et données Vachon (2022) (B)

Le clan vocal le plus présent (EC1) se répartit sur l'ensemble de la zone d'étude, bien que plus abondant dans la zone Centre. La distribution des autres clans est beaucoup plus localisée autour de la zone Sud majoritairement, avec la présence du clan vocal EC2 au niveau de la Martinique, Sainte-Lucie et Saint-Vincent-et-les-Grenadines et celle du clan vocal EC3 autour de Grenade uniquement (Figure 19.A). Les points blancs sur la carte représentent la localisation des jours où plusieurs clans ont été entendus ensemble, soit EC1 et EC2 au niveau de la Guadeloupe, de Montserrat et des îles vierges des Etats-Unis et EC2 & EC3 en Dominique.

4- Détermination des déplacements

4.1 Détermination des déplacements à l'échelle individuelle

Une carte a été établie, reprenant les déplacements de l'ensemble des individus identifiés par la CCS (Figure 20). Pour les individus recapturés, les positions où ils avaient été observés préalablement par DSWP ou Vachon (2022), ont été prises en compte afin de mieux rendre compte du réel déplacement de ces individus sur notre zone d'étude. Le déplacement des individus a pu être établi seulement pour les individus ayant été capturés plus d'une fois (soit 47 individus). Sur cette carte, seuls les points GPS les plus éloignés pour chaque individu ont été conservés afin de ne pas surcharger la carte avec toutes les positions GPS et pour montrer la plus grande distance parcourue par chaque individu. Les lignes qui représentent les déplacements sont donc à considérer avec précaution puisqu'il ne s'agit pas du trajet réel de l'animal en ligne droite mais juste d'une représentation de la distance maximale parcourue par cet individu.

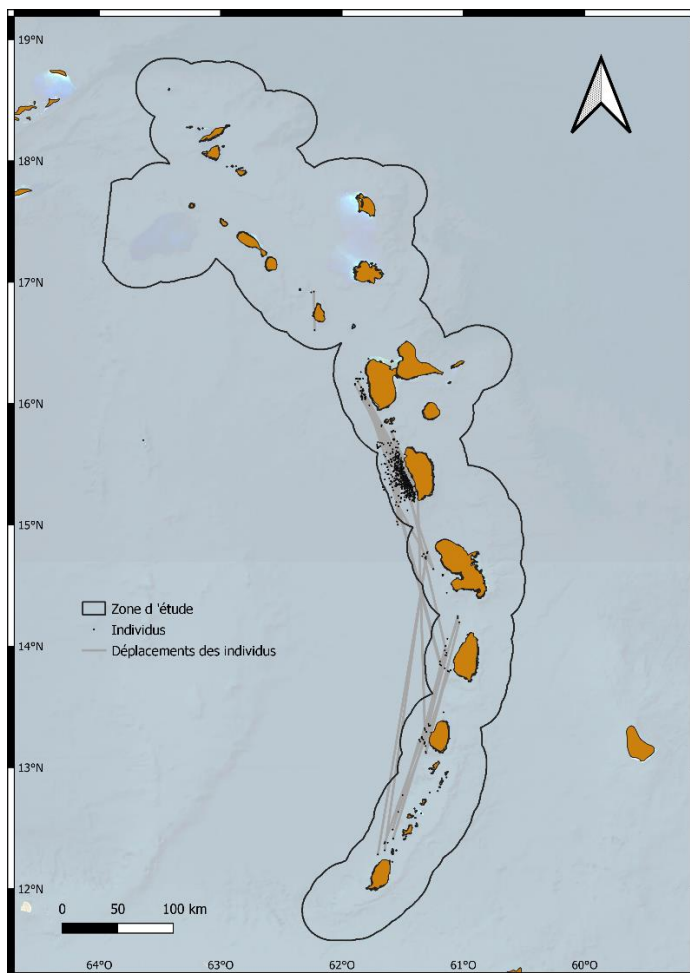


Figure 20: Observations & recaptures des individus de la CCS (données CCS, Vachon, DSWP)

Sur les 47 individus recapturés, 20 ont parcouru une distance supérieure à 100km. La durée entre les deux événements de recapture responsables de la distance maximale varie entre 85j (individu 6313 recapturé les deux fois par la CCS, une fois à Saint-Vincent-et-les-Grenadines le 27/03/2022 et l'autre fois à Grenade le 20/06/2022) et 10 ans (individu 5575 capturé pour la dernière fois par DSWP en 2011 et recapturé le 23/09/2022 à Sainte-Lucie par la CCS) (Annexe 6.B). La médiane est de 4,17 ans pour parcourir 146 km environ.

Parmi les 20 ayant parcouru plus de 100km, 4 ont parcouru plus de 200 km en une durée moyenne de 2 ans et 53 jours pour les individus 6377 et 6422 entre la Martinique et Grenade et environ 5 ans pour les individus 6022 et 6230 (Dominique → Saint-Vincent-et-les-Grenadines) (Annexe 6.B).

Sur les 20 individus ayant parcourus plus de 100 km, la plupart sont globalement restés dans la même zone (zone centre pour 7 individus ; zone Sud pour 6 individus) tandis que 7 autres se sont déplacés entre la zone Centre et la zone Sud en parcourant plus de 150 km (dont 3 entre la Dominique et Sainte-Lucie ; 2 entre la Dominique et Saint-Vincent-et-les-Grenadines ; 2 entre la Martinique et la Grenade).

4.2 Détermination des déplacements à l'échelle des unités sociales & clans vocaux

Les déplacements à plus grande échelle ne sont pas individuels mais s'effectuent bien à l'échelle de l'unité sociale, voire même du groupe (groupe T,V par exemple). La Figure 21 représente le déplacement d'individus appartenant à trois unités sociales distinctes (T,K et G12). Les individus appartenant à une même unité sociale voyagent ensemble vers les mêmes destinations, qui sont d'ailleurs différentes en fonction de l'unité sociale en question.

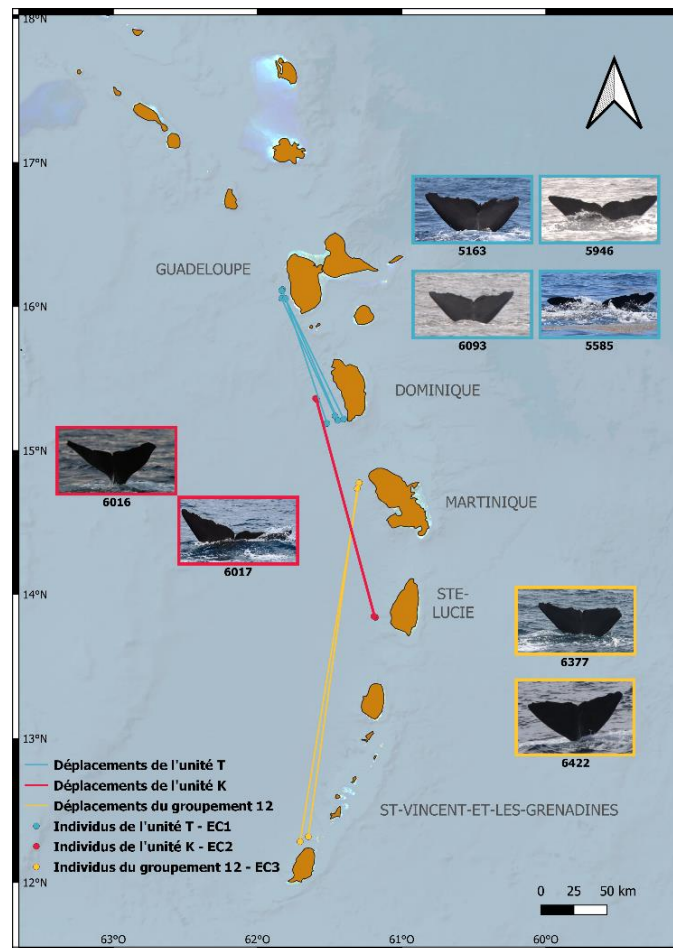


Figure 21: Déplacements de plus de 100km de 8 individus appartenant à des unités sociales et des groupes vocaux différents (données CCS, Vachon, DSWP)

Plus largement, le clan vocal auquel chacun de ces unités sociales/groupes appartient influence les destinations vers lesquelles voyagent les individus (Figure 22).

La Figure 22 laisse apparaître des zones relativement distinctes où se déplacent les différentes unités sociales en fonction de leur clan vocal. Sur les 17 unités sociales préalablement identifiées, 12 appartiennent au clan vocal EC1, 5 au clan vocal EC2 et 1 seule au clan vocal EC3. Le nombre de fois où toutes les unités sociales appartenant à un même clan vocal ont été vues est résumé dans le Tableau 9.

Tableau 9: Nombre d'occurrences de chaque clan vocal à partir des données ayant servi à faire la Figure 22

Clan vocal	EC1	EC2	EC3
Nombre de fois où chaque clan a été vu	247	18	6

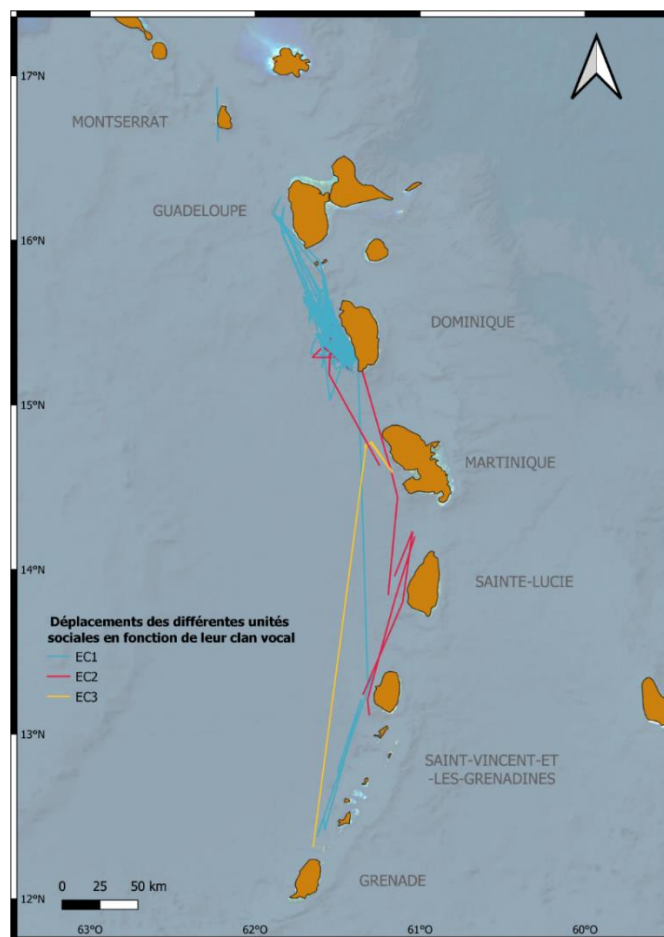


Figure 22: Déplacements des différentes unités sociales en fonction de leur clan vocal (données CCS, Vachon, DSWP)

Les unités sociales du clan vocal EC1 effectuent de courts déplacements entre la Guadeloupe et la Dominique, bien que certaines unités aient été observées plus au Sud (Figure 20). Les unités sociales/groupements des clans vocaux EC2 et EC3 se déplacent majoritairement vers le Sud, bien que l’ampleur de ces déplacements et la destination exacte diffèrent entre les deux groupes. En outre, les unités du clan EC2 dépassent rarement Saint-Vincent alors que les unités du clan EC3 ont été aperçues bien plus au Sud, au niveau de Grenade, ayant parcouru plus de 270km pour atteindre cette destination.

Sur les 20 individus préalablement identifiés comme parcourant plus de 100km, 12 appartiennent au clan vocal EC1, 5 au clan vocal EC2 et 2 au clan vocal EC3. En pondérant par le nombre d’individus identifiés total appartenant à chaque clan, 27 % de EC1 et 42% de EC2 parcourent plus de 100km, tandis que 100% des individus EC3 capturés parcourent plus de 270km. Ces pourcentages viennent confirmer l’ampleur visuelle des déplacements des différents clans affichés par la Figure 22.

5- Modélisation de la distribution des cachalots

Le jeu de données a été aléatoirement divisé en deux jeux de données : un jeu de données de calibration et un jeu de données de validation, représentant respectivement 80% et 20% du jeu de données initial. Ainsi, le calage du modèle a été réalisé sur le jeu de données de calibration tandis que la validation de ce dernier a été réalisée sur le jeu de données de validation (c’est-à-dire les 20% restants n’ayant pas servi à la modélisation). L’intérêt d’une telle manœuvre est de tester la robustesse des prédictions faites. En outre, le modèle est entraîné sur une partie des données pour qu’il trouve

les meilleurs prédicteurs et fasse des prédictions à partir de ces derniers puis sa robustesse est testée en conservant les prédicteurs choisis par la première simulation et en les appliquant aux 20% de données restantes n'ayant pas servi jusqu'ici, et que l'on peut donc considérer comme un nouveau jeu de données.

5.1 Modélisation & Quantification des erreurs

A partir des données environnementales choisies, 2731 combinaisons variant entre 1 et 4 variables explicatives au maximum ont été testées. Les variables environnementales qui étaient fortement corrélées n'ont pas pu être sélectionnées toutes les deux comme variables prédictives dans un même modèle (Annexe 9). Le meilleur modèle est représenté par celui ayant la plus faible valeur d'AIC.

Les deux premiers modèles ainsi que leurs combinaisons des variables explicatives, leurs AIC, et leurs déviations expliquées sont représentés dans le Tableau 10.

Tableau 10: Deux premiers modèles expliquant le mieux la distribution des cachalots

Modèle	Co-variables environnementales	AIC	Déviance expliquée (%)
1	Moy_bathy + Moy_CHL + distcoast + Moy_SST	255,49	24,1
2	Moy_bathy + Moy_CHL + distcoast + MLD	258,31	23,4

Les deux premiers modèles présentent à chaque fois quatre co-variables pouvant prédire la distribution des cachalots. Parmi ces quatre, trois sont identiques entre les modèles, à savoir la bathymétrie (c'est-à-dire la profondeur), la chlorophylle A et la distance à la côte. Ces trois variables sont donc d'une importance capitale pour expliquer la distribution des cachalots.

Le premier modèle a une valeur d'AIC de 255,49 et explique 24.1 % de la distribution prédite avec comme quatrième variable la température de l'eau de surface. Il s'agit là du meilleur modèle si l'on considère l'AIC du second modèle (258,31) dont l'écart avec le premier est bien supérieur à 2 (environ 2,8) ainsi que la déviance expliquée qui est plus faible que celle du premier modèle (23,4 contre 24,1). Le premier modèle est celui qui fait le meilleur compromis entre la précision du modèle et la complexité (nombre de paramètres).

Les courbes de lissage des quatre variables environnementales du meilleur modèle sont représentées en Figure 23. Ces courbes permettent de visualiser la contribution de chaque prédicteur, en représentant la relation estimée entre la variable réponse et le prédicteur en question. L'axe X correspond aux valeurs de la variable environnementale décrite tandis que l'axe Y représente la prédiction de la fréquence d'observation dans cette zone de l'espace du prédicteur considéré. Sur cette figure, les courbes en noir représentent les courbes de lissage non-linéaires choisies comme meilleur ajustement de chaque variable par rapport au nombre d'observations relatif par hexagone. La partie en bleu autour de cette courbe montre l'intervalle de confiance à 95% associé à cet ajustement. Les tirets noirs présents sur l'axe X représentent la distribution de la variable environnementale sur l'ensemble de la zone d'étude (c'est-à-dire la densité de points de cette variable qui se situe entre telle ou telle valeur de la variable). Les tirets rouges mettent en évidence les valeurs de la variable environnementale où les cachalots ont été réellement observés.

En ce qui concerne la profondeur, plus cette dernière augmente et plus la fréquence estimée de groupes de cachalots est importante (Figure 23) avec un intervalle de confiance à 95% qui augmente à mesure qu'on se rapproche des plus grandes profondeurs. La tendance s'inverse pour la distance à la côte où la prédiction du nombre de groupes de cachalots diminue à mesure que l'on s'en éloigne.

La courbe de lissage relative à la température indique elle-aussi une tendance à la hausse de la densité d'observations quand ses valeurs augmentent.

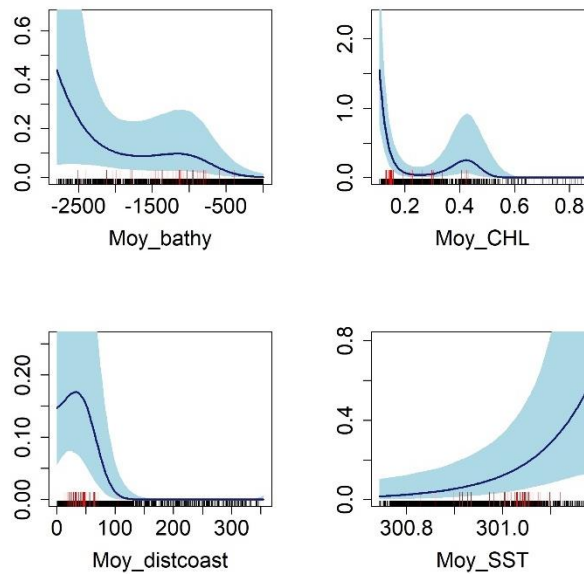


Figure 23: Relation entre espérance de la variable réponse et chacune des variables explicatives pour les deux premiers modèles (à gauche, modèle 1 et à droite, modèle 2). L'histogramme en noir en bas des figures représente la densité d'observations dans l'espace des variables explicatives. Les tirets rouges représentent les valeurs observées du prédicteur.

Les prédictions associées au meilleur modèle sont présentées en Figure 24. Dans chaque hexagone, la valeur retenue pour la prédiction est la moyenne des prédictions à l'intérieur de l'hexagone :

$$\widehat{\mu}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} \widehat{\mu}_i$$

Où h représente l'hexagone considéré et n_h le nombre de prédictions associées à l'intérieur de cet hexagone. Ainsi, la probabilité d'observer aucun groupe sur l'hexagone s'écrit :

$$\mathbb{P}(\hat{Y} = 0) = e^{-\widehat{\mu}_h}$$

Et la probabilité au moins un groupe sur l'hexagone est donc égale à :

$$\mathbb{P}(\hat{Y} \geq 1) = 1 - e^{-\widehat{\mu}_h}$$

Ces valeurs sont associées à des coefficients de variations changeant entre 0.2 et 62.5%.

La Figure 24 indique que la distribution potentielle des cachalots est notamment plus importante du côté Ouest de l'arc Antillais, et ce, sans exception du Nord au Sud. La probabilité de rencontrer un groupe de cachalots au niveau de la partie Mer des Caraïbes de notre zone d'étude est globalement comprise entre 0.05 et 0.25 (cas de la Martinique, de Sainte-Lucie et des îles de la zone Nord) contre 0 à 0.1 pour la partie Atlantique. Une distribution intermédiaire a été prédite aux abords de la Guadeloupe et de la Dominique, dont les probabilités varient entre 0.4 et 0.79.

Au niveau de la partie Sud de la zone d'étude, notamment entre les îles de Grenade et Saint-Vincent-et-les-Grenadines, la probabilité d'observer un groupe de cachalots est majoritairement comprise entre 0.79 et 1, traduisant une probabilité de rencontre quasiment certaine. Ces fortes densités prédites se retrouvent également au niveau de la partie Atlantique de la Guadeloupe, et plus particulièrement au niveau d'une zone de forte profondeur (probabilité quasiment égale à 1).

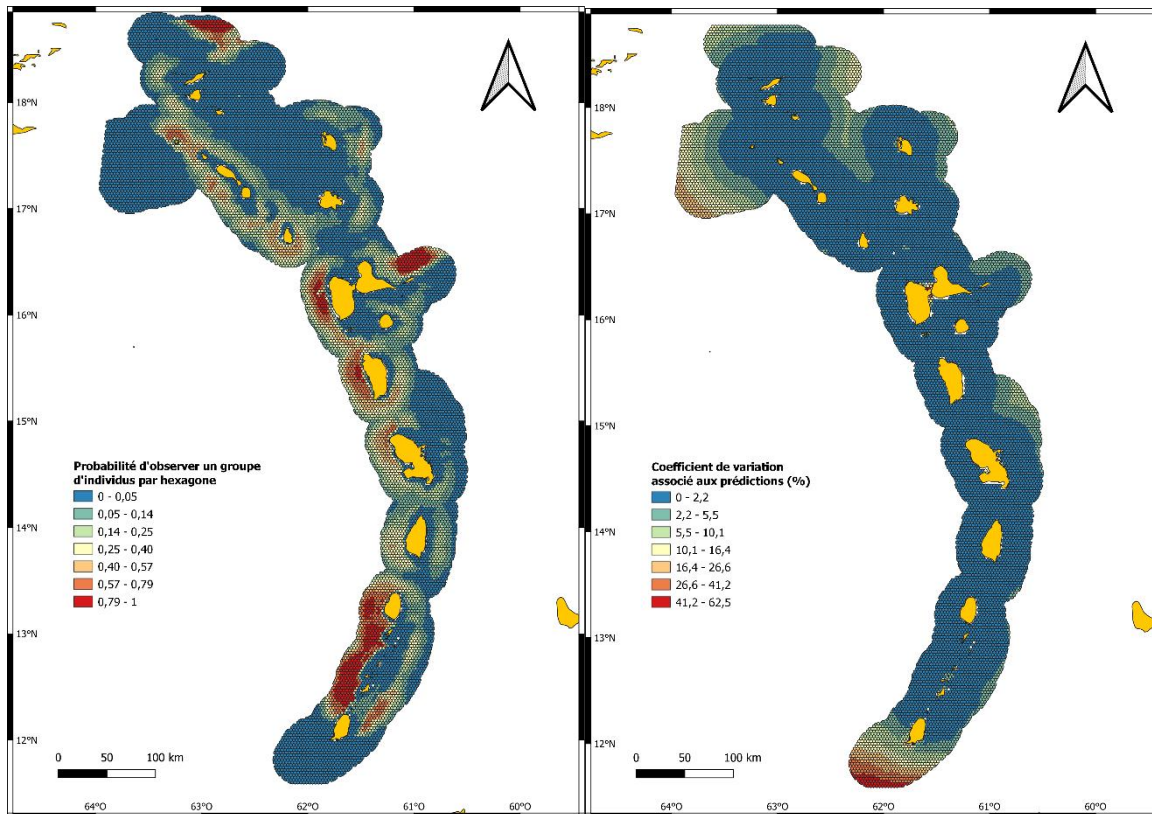


Figure 24: Cartes représentant la prédiction de la distribution des cachalots selon les quatre prédicteurs choisis (à gauche). Les coefficients de variations reliées à ces prédictions sont présentés à droite.

Les cartes d'incertitude (Figure 24.B) associées révèlent de fortes variations de la fréquence de groupes prédite au niveau des limites Sud et Nord de notre zone d'étude, ainsi qu'au niveau de la côte Atlantique. Ces coefficients sont faibles (inférieurs à 2.2 %) sur le reste de la zone d'étude et notamment au niveau des zones où les plus fortes fréquences de groupes de cachalots sont prédites.

5.2 Validation

Les données du jeu de validation ont été utilisées ici afin de rendre compte de la robustesse et de la transposabilité du modèle. La probabilité associée à la distance maximale entre la courbe ROC et la première bissectrice est de 0.047. La matrice de confusion obtenue est présentée dans le Tableau 11:

Tableau 11: Matrice de confusion réalisée à partir du jeu de validation

		Réalité	
		Absence : 0	Présence : 1
Prédiction	Matrice de confusion		
	Absence : 0	58% (429 vrais négatifs)	12% (1 faux négatif)
Présence : 1		42% (309 faux positifs)	88% (8 vrais positifs)

Les deux métriques résumant la performance du modèle sont représentées en vert dans le tableau. Ainsi, la sensibilité du modèle est de 88% puisque 8 des 9 présences ont été correctement prédites par le modèle. Concernant la spécificité, cette dernière vaut 58% car sur les 708 négatifs prédits, seuls 429 sont des vrais négatifs, le reste étant des données ayant été prédites comme présences alors qu'il s'agissait d'absences. L'AUC a également été évaluée à 0.79. Cela indique une performance du modèle relativement bonne et non aléatoire. La courbe ROC, toujours au-dessus de la première bissectrice (voir Annexe 11), indique que le taux de vrais positifs est toujours supérieur au taux de faux positifs donc que la prédiction n'est pas aléatoire.

IV- Discussion

1- Une organisation sociale ordonnée

1.1 Le regroupement d'unités sociales ensemble : description de nouveaux groupes ?

Au total 62 individus ont été identifiés dont 45 qui avaient déjà été capturés par le DSWP ou Vachon (2022). Ces individus recapturés se répartissent en 17 unités sociales distinctes déjà connues. Le dendrogramme (Figure 17), basé sur le nombre de fois où les individus ont été vus les uns avec les autres, a mis en évidence six groupements dont deux déjà connus et quatre dont les compositions correspondent plutôt à la définition d'un groupe établi par Gero et al. (2014) et soutiennent les observations de DSWP (2019) concernant le regroupement temporaire de plusieurs unités sociales en groupe pour chasser ou socialiser. Ces regroupements représentent :

- **Des groupes déjà documentés :**
- ✓ **Le Groupe 5 : Un mélange des unités sociales F & U:**

L'unité U (Utensils) et l'unité F (The Group of Seven) sont deux unités sociales qui ont l'habitude de se fréquenter et qui ont été vues ensemble très fréquemment depuis le début des travaux de DSWP (2005). Les femelles de ces deux unités se mélangeraient même pour baby-sitter les veaux des femelles de l'autre unité sociale. Ces deux unités ont été aperçues ensemble à 4 reprises sur la période 2021-2022.

- ✓ **Le Groupe 2 : Un mélange des unités sociales T & V :**

L'unité T (Tooth's Team) et l'unité V (Vive la France) passent beaucoup de temps ensemble (DSWP (2019)). Ces deux unités socialisent souvent sur la côte Nord-ouest de la Dominique. C'est d'ailleurs exactement là qu'elles ont été vues ensemble par la CCS, les 28/05 et 26/06/2021.

- **Des groupes encore jamais renseignés :**
- ✓ **Le Groupe 3 : Un mélange des unités Z, N & A:**

Les membres de l'unité N sont connus pour utiliser les eaux de la Dominique depuis 1995 mais ils ont été vus pour la première fois en tant qu'unité familiale en 2005. Cette unité est plutôt grande et évolue rapidement. Le grand nombre de veaux présents dans cette unité peut la rendre complexe à analyser puisque ces derniers obtiennent tellement de marques d'une année l'autre qu'ils peuvent alors devenir difficilement reconnaissables (DSWP, 2019).

Aucune observation de ces unités sociales auparavant ne montre qu'elles passeraient beaucoup de temps ensemble, justifiant leur regroupement au sein d'un même groupe. Concernant les données d'observation de la CCS, elles justifient ce regroupement des unités sociales Z & N & A puisqu'observées ensemble le 07/07/21. Les unités Z&N ont été observées une fois de plus ensemble le 27/06/21.

- ✓ **Le Groupe 1 : Un mélange des groupements 15 et 17 :**

De la même façon ici, l'association de ces deux groupes ensemble n'a pas vraiment été décrite auparavant. Le groupement 15 a été vu une seule fois par la CCS et c'était en compagnie du groupement 17 le 21/07/2021. Les données de Vachon (2022) montrent qu'ils avaient déjà été observés une fois ensemble le 12/03/2020. Il est donc fortement possible que ces deux groupes s'associent temporairement, voire qu'il s'agisse en fait d'une unique unité sociale. En effet, ces groupements ont été découverts par Vachon (2020) et il reste encore beaucoup à apprendre sur eux.

1.2 Les nouveaux individus : nouvelles unités sociales ou nouveaux individus d'unités déjà décrites ?

Parmi les 17 nouveaux individus restants, jamais capturés auparavant, certains d'entre eux ont été aperçus avec des unités sociales déjà connues (Figure 17), tandis que d'autres ont été capturés en compagnie d'aucun autre individu déjà connu. Parmi eux (Figure 14), 5 sont peu marqués et sont donc très probablement des individus juvéniles appartenant à des unités sociales déjà connues ou non encore décrites. Les différents cas sont présentés ci-après :

- D'après le Tableau 7, le nouvel individu CCS_Pm_0022 (Figure 14) pourrait appartenir à une des trois unités T, N ou Z. En effet, cet individu a été observé une seule fois dans un grand groupe réunissant ces trois unités sociales simultanément, ce qui rend impossible de savoir à laquelle de ces trois unités il appartient. Cependant, certaines hypothèses peuvent être émises sur son appartenance plus probable à l'une ou l'autre des unités. En outre, il s'agit ici d'un individu que l'on peut qualifier de juvénile, au vu du peu de marques présentes sur sa nageoire caudale. Or, le nombre de membres au sein de l'unité N est connu pour connaître de nombreuses évolutions, au regard du grand nombre de veaux et juvéniles qui évoluent en son sein. Il serait donc plus probable que ce nouvel individu appartienne à l'unité N, plutôt qu'aux unités T et Z dont la structure sociale est plutôt stable ces dernières années (Tableau 7).
- De même, les individus CCS_Pm_0016 et CCS_Pm_0044 (Figure 14) ont été aperçus ensemble et sont également des juvéniles. L'individu 16 a été aperçu une fois en compagnie de l'individu 5588 qui appartient à l'unité V (Tableau 7). Deux cas plausibles se présentent alors : soit les individus 16 & 44 sont deux nouveaux jeunes, nés après 2020, de l'unité V, soit il s'agit de nouveaux jeunes d'une autre unité sociale.
- L'individu 8 (juvénile, Figure 14) a été vu trois fois la même année, ce qui, selon la définition la plus restrictive d'une unité sociale (Gero et al., 2014) ne permet pas de le relier à un quelconque groupement. En considérant la définition la plus souple, les recaptures ont effectivement eu lieu à plus de 30 jours d'intervalle. Cependant, à chacune de ces trois captures, les deux unités sociales F et U étaient présentes (Tableau 7). Il est donc impossible de savoir exactement à laquelle il appartient.
- Quant à l'individu CCS_Pm_0019, il n'a été vu qu'une seule fois et en compagnie de l'individu 5575 dont l'appartenance à une unité sociale ou même à un clan vocal reste encore inconnu (Tableau 7). En effet, l'individu 5575 n'a été capturé que quatre fois au total donc une fois en 2022 par la CCS. Au vu du faible nombre de recaptures sur une période temporelle étendue (13 ans), il pourrait donc s'agir ici d'un mâle. L'individu CCS_Pm_0019 pourrait alors être également un jeune mâle qui a rejoint ce mâle plus âgé afin d'apprendre les codes sociaux nécessaires à l'accouplement. Une autre hypothèse serait également que l'individu 5575,

supposément mâle, aurait tenu compagnie à une unité matrilineale, dont seul l'individu CCS_Pm_0019 était visible, en quête d'un partenaire sexuel.

Les 12 nouveaux individus restants sont des individus adultes ou en tout cas des juvéniles moins jeunes au regard de leurs nageoires caudales bien marquées. Ces derniers peuvent se diviser en deux catégories : les individus aperçus en compagnie d'individus déjà capturés (Figure 17, Tableau 7) et ceux observés pour la première fois par la CCS et seuls (Tableau 5).

Les nouveaux individus CCS_Pm_0006, CCS_Pm_0028, CCS_Pm_0029, CCS_Pm_0057 et CCS_Pm_0060 font plutôt partie de la première catégorie. Etant donné que ces individus n'avaient jamais été décrits auparavant, il pourrait s'agir de mâles venus rendre visite à différentes unités sociales, en quête d'une femelle avec laquelle s'accoupler. Il est également possible que ces individus soient des femelles appartenant à un nouveau groupement qui n'a pas encore été décrit, et qui viennent socialiser avec ces unités sociales déjà décrites. Sur ces 5 individus, 3 n'ont été vus qu'une seule fois et les deux autres ont été, soit vus deux fois sur deux années distinctes mais en compagnie de deux individus différents, soit deux fois en compagnie des mêmes individus mais à moins de 30 jours d'intervalles. Aucune des situations précédentes ne permet d'être certain du rattachement de ces individus à l'unité sociale avec laquelle ils ont été aperçus, ni de connaître leur sexe.

Quant aux nouveaux individus 14, 15, 17 et 18, ils rentrent dans la deuxième catégorie puisqu'ils n'ont jamais été observés auparavant et ont été vus une fois ensemble le 27/06/2021. Cette unique capture n'est pas suffisante pour émettre quelconque conclusion sur leur possible appartenance à une même unité sociale (Figure 25).



Figure 25: Planche photos des individus 14,15,17 & 18, possiblement nouvelle unité sociale

Ainsi, pour l'ensemble des points discutés précédemment, seule la prise en compte de nouvelles données et la potentielle recapture de ces nouveaux individus pourront nous en apprendre plus sur l'interaction de ces nouveaux individus avec les unités sociales déjà décrites ou mettre en évidence leur appartenance à une nouvelle unité sociale.

2- Des nomades de l'océan ?

Les cachalots ont souvent été décrits comme des nomades de l'océan, parcourant de longues distances, pouvant atteindre un millier de kilomètres (Mizroch & Rice, 2013; L. Rendell & Frantzi, 2016). Cependant, d'après les résultats présentés dans cette étude, les unités sociales matrilineales des cachalots des Petites Antilles restent dans cette zone l'année durant et la distance maximale qu'elles parcourent en une durée minimale de 2 ans est inférieure à 300 km (Annexe 6.B). Les observations de Vachon (2022) corroborent particulièrement l'organisation de ces unités sociales à petite échelle spatiale.

La Figure 16.A vient elle aussi appuyer cette théorie en mettant en lumière le caractère sédentaire des unités sociales. La différence très distincte des individus observés entre la zone proche des côtes et la zone plus au large de Vachon (2022) suggère que ces individus ont un domaine vital assez restreint et occupent des habitats favorisés aux caractéristiques propres dont ils s'éloignent à de rares occasions. De surcroît, la Figure 16.B présente différentes familles observées plus ou moins proches de la côte. Au niveau de la Dominique, l'unité N vivrait dans la zone 2, plus au large, tandis que l'unité A aurait apprivoisé les eaux plus près de la côte dans la zone 1. Ces deux unités sociales/groupements se rassembleraient en un seul groupe pour socialiser ou chasser au large de la Dominique, quittant ainsi leurs domaines « vitaux » respectifs. D'un autre côté, les unités sociales G12 et K, semblent aussi avoir des aires de répartition bien distinctes. Cela est d'autant plus intéressant si on prend en compte que ces deux unités appartiennent à des clans vocaux différents. Cela viendrait étayer l'hypothèse selon laquelle les clans vocaux auraient des aires de répartition bien différentes (Vachon, Eguiguren, et al., 2022), en l'étendant même aux groupes et unités sociales.

Cette forte ségrégation spatiale observée entre les différentes unités sociales pourrait donc expliquer les différences entre les estimations d'abondance de Vachon (2022) et la nôtre. En outre, la tendance à la sous-estimation de l'abondance réelle des cachalots avec les données de la CCS viendrait d'une zone échantillonnée plus proche des côtes, associée à une forte sédentarité des cachalots.

3- Des points chauds où se rassemblent les cachalots ?

3.1 Des facteurs environnementaux définissant des habitats préférentiels

Afin de mieux comprendre la manière dont les cachalots sont distribués et leurs types d'habitats préférés, les résultats des modèles GAM ont été rigoureusement étudiés.

Dans un premier temps, l'exploration des courbes de lissage utilisées pour décrire la relation entre les variables prédictives et les données d'observations est déjà intéressante afin de mieux comprendre leur choix dans les meilleurs modèles. En outre, le choix de la profondeur et de la chlorophylle A comme meilleurs prédicteurs de la distribution laisse peu de place au doute et étaient également ressortis dans les modèles de prédiction suite aux campagnes REMMOA II (Laran et al., 2019).

En effet, l'augmentation de la fréquence d'observations prédite avec la profondeur est sans surprise si l'on considère le caractère de grand plongeur de *Pm* qui chasse et passe la majorité de son temps en eaux profondes (Figure 23). L'intervalle de confiance à 95% associée à cette courbe de lissage augmente à mesure qu'on se rapproche des plus grandes profondeurs. Cela peut s'expliquer par le faible nombre d'observations en zones plus profondes, étant donné que l'effort de la CCS ne s'éloigne que rarement à plus de 12 km de la côte. Ces grands plongeurs restent pour autant à proximité de la côte, comme l'indique la relation décroissante existante entre la distance à la côte et le nombre d'observations. De surcroît, la Figure 16 avait déjà montré la préférence des cachalots observés par la

CCS pour les zones calmes proches des côtes. Ces observations soutiennent celles faites par GANDILHON et al. (2012) montrant l'évolution des cachalots en zone côtière des îles de l'arc.

La présence de la chlorophylle A reflète la productivité primaire qui influence donc l'abondance du reste de la chaîne trophique, et est donc un bon proxy pour refléter la distribution des calmars (Figure 23). Les faibles variations de la température de surface rendent difficiles une quelconque interprétation de sa relation avec la densité d'observations des cachalots. Pour autant, la courbe de lissage montre une relation non-linéaire croissante qui indiquerait une préférence des cachalots pour les eaux plus chaudes en surface (échelle de moins de 1°C sur le graphe).

3.2 Une densité plus importante du côté Caraïbe ?

La Figure 24, relative aux prédictions de la distribution provenant du meilleur modèle, indique une densité plus importante du côté Mer des Caraïbes, où la Figure 7.A a montré que la bathymétrie atteignait rapidement de fortes profondeurs supérieures à 2000m. Ces gradients de profondeur élevés sont connus pour être des zones d'intérêt particulier pour les cachalots, comme cela avait déjà été montré dans le Golfe du Mexique (Davis et al., 1998). De même, cette densité diminue lorsqu'on s'éloigne de la côte. Les cachalots auraient donc une préférence pour les zones qui rassemblent une grande diversité de profondeurs et donc d'habitats, tout en restant près des côtes où les conditions sont plus favorables. La température plus chaude que du côté Atlantique (moyenne de 27.8°C environ) favorise également l'activité métabolique des organismes marins, en permettant notamment la prolifération du phytoplancton et en attirant les espèces thermophiles. Cette distribution préférentielle à proximité de l'arc Antillais avait déjà été mise en évidence par Laran et al. (2019).

Des corrélations similaires décrivant la répartition des cachalots ont été régulièrement observées dans différentes régions du monde (Pirodda et al., 2011; Rice et al., 2021). En outre, ces caractéristiques pourraient agir comme des proxies reflétant la disponibilité en proies, puisque les grandes variations de profondeur sont souvent associées à un mélange au sein de la colonne d'eau, duquel résulte une productivité primaire et secondaire décuplée.

Cela pourrait expliquer en partie pourquoi les cachalots affectionneraient plus le côté Caraïbe par rapport au côté Atlantique par exemple. De plus, des travaux de longues dates montraient déjà la préférence des cachalots pour la côte sous le vent (Watkins & Moore, 1982), en tant qu'habitat préférentiel de leurs proies potentielles. La partie Ouest des îles, à proximité des côtes, a été confirmée comme être un habitat favorable à la présence des cachalots dans des travaux beaucoup plus récents (HUGON & MAALOUF, 2023; Vachon, Eguiguren, et al., 2022) où leur probabilité de présence prédite autour des îles françaises était quasiment égale à 1 (HUGON & MAALOUF, 2023).

Cependant, cette possible préférence pour la zone Caraïbe n'est pas si évidente si l'on considère la forte probabilité prédite sur la côte Atlantique de la Guadeloupe. Cela ne semble pas improbable que des cachalots puissent réellement apprécier ces zones, qui présentent des caractéristiques bathymétriques favorables à leur présence. Cela rejoindrait notamment les prédictions de distribution issues du programme REMMOA II (Laran et al., 2019) qui révélaient une distribution des cachalots impactée par la saisonnalité, avec une préférence des cachalots pour l'Atlantique en saison humide.

Il est également important de souligner que la forte densité prédite du côté Caraïbes pourrait être le fait du suréchantillonnage de cette partie de la zone d'étude par rapport à l'océan Atlantique par exemple.

Les prédictors évalués ici comme les meilleurs pour rendre compte de la distribution potentielle des cachalots dans les Petites Antilles sont en partie similaires à ceux obtenus par Vachon (2022), notamment en ce qui concerne la chlorophylle et la distance à la côte. Néanmoins, malgré les découvertes de Vachon (2022) suggérant une préférence des cachalots pour les eaux des îles du Centre telles que la Dominique et la Martinique, nos prédictions révèlent une histoire divergente, comme nous le démontrerons par la suite.

3.3 Une probabilité d'observer des cachalots plus importante dans les régions Centre et Sud

Au vu de la Figure 15, les régions Centre et Sud sont celles où l'on retrouve les plus fortes densités d'observations de cachalots et ce, surtout du côté Mer des Caraïbes. Ces observations corroborent celles de Vachon (2022) dont respectivement 51% et 38% des individus avaient été observés dans les régions Centre et Sud pour seulement 11% dans la région Nord (voir Annexe 6.A). Les îles autour desquelles les cachalots sont majoritairement présents sont la Guadeloupe, Saint-Vincent-et-les-Grenadines ainsi que la Dominique. A elles-seules, ces trois îles regroupent 84% des observations.

En outre, les résultats des prédictions issues du meilleur modèle (Figure 24) s'accordent pour indiquer une forte distribution de cachalots au niveau de la partie Sud de la zone d'étude, notamment entre les îles de Grenade et Saint-Vincent-et-les-Grenadines. Entre ces deux îles existe une grande diversité de fonds dont un large plateau (Figure 7.A), servant de refuge à une large diversité de proies. Cela peut donc influencer la distribution des calmars et donc conséquemment celle des cachalots. Cette zone présente également une productivité primaire plus importante qu'au Nord avec en moyenne, une concentration en chlorophylle de 0.3mg/m³, qui pourrait également participer à expliquer la forte densité des cachalots dans cette aire. C'est également entre ces deux îles que les courants de surface sont les plus forts, notamment car c'est une des portes d'entrée du courant des Caraïbes. Ces deux caractéristiques pourraient donc également expliquer la forte probabilité d'observer des groupes au niveau de cette zone.

La Figure 24 rend également compte d'une zone de densité moyenne autour de la Guadeloupe et de la Dominique, zones au sein desquelles les captures/recaptures des cachalots sont nombreuses. La présence d'habitats favorables à la distribution des cachalots au niveau de ces îles a d'ailleurs également été mis en évidence par HUGON & MAALOUF (2023). La présence d'unités sociales de façon permanente dans cette zone pourrait être due à une production primaire décuplée, liée à l'existence d'un canyon massif dans cette région, favorisant la remontée d'eau riche en nutriments.

4- La non-mixité entre clans est-elle réelle ?

4.1 Une répartition des clans vocaux bien définie

Des aires de répartition distinctes

Les résultats de la distribution acoustique des différents clans vocaux (Figure 19.A) corroborent les résultats de Vachon (2022) concernant la distribution des clans EC1 et EC2 et confirment notamment la présence certaine du clan vocal EC2 dans les Petites Antilles. Bien que globalement présents sur l'ensemble de la zone d'étude, le clan EC1 est surtout présent au niveau de la zone Centre de la zone d'étude tandis que les clans EC2 et EC3 se répartissent majoritairement au niveau de la zone Sud.

La Figure 19.A apporte des informations supplémentaires par rapport à la Figure 19.B en montrant la présence de EC1 dans les eaux Martiniquaises et en mettant en lumière la présence d'individus EC3

très au Sud de l'arc insulaire. En effet, ce clan serait plutôt distribué au niveau de la région Sud des Petites Antilles, spécifiquement au niveau de Grenade. De surcroît, ce clan aurait également été entendu dans des enregistrements brésiliens ce qui viendrait appuyer sa présence dans le Sud de l'arc Antillais (S. Gero et al., communication personnelle, 5 juin 2023). Cependant, le peu d'enregistrements dans lesquels nous avons trouvé EC3 (2) ne permet pas de valider l'hypothèse de Vachon (2022) relatif à l'existence certaine d'un troisième clan vocal dans les Petites Antilles. La collecte de données supplémentaires lors des prochaines expéditions permettra de confirmer ou infirmer l'appartenance des individus 5575 et 19 au clan vocal EC3 (23/09/22), de collecter davantage de données acoustiques et donc d'affiner les hypothèses émises sur l'existence réelle de ce clan.

De même, le clan vocal EC2 aurait été entendu plus au Nord de la zone d'étude (Figure 19), au niveau de la Dominique, de la Guadeloupe, de Montserrat et des îles vierges des Etats-Unis mais toujours en compagnie d'un autre clan vocal (EC1 ou EC2) et faiblement présent acoustiquement ce qui rend sa présence incertaine.

Une différence observée dans l'ampleur et la durée des déplacements entrepris

Ce travail a également permis de mettre en évidence que des différences existent entre les clans vocaux, notamment en ce qui concerne le comportement migratoire de leurs unités sociales (Figure 21). En outre, le clan vocal EC3, bien que peu représenté, semble parcourir de plus grandes distances que le clan vocal EC1 par exemple, qui reste majoritairement autour des îles de Guadeloupe et de Dominique. De surcroît, en considérant la présence du clan vocal EC3 au Brésil, il est même possible que les individus de ce clan parcourent en réalité des distances bien plus grandes que celles que l'on imaginait jusqu'à présent.

Cependant, ces analyses des déplacements par clans restent à considérer avec précaution au vu du déséquilibre des données entre les clans.

4.2 Présence simultanée de deux clans : Rencontre entre clans vocaux ou apprentissage social ?

Sur les 3 journées d'enregistrements acoustiques où on a entendu EC1 et EC2, 2 d'entre elles étaient liées à des observations du clan vocal EC1 uniquement. En outre, lors de ces deux journées (26/05 & 07/07/21), sur les 490 des codas entendus, 473 étaient des codas du clan vocal EC1 et seulement 17 appartenaient au clan vocal EC2 (Tableau 8). Il est donc possible que la présence de ces codas identitaires de type 5R soit par exemple dû à une erreur des individus juvéniles du groupe. En outre, on peut bien imaginer que lors de l'apprentissage de ce coda identitaire propre à chaque clan (qui peut correspondre à notre définition de la langue), les jeunes individus fassent des erreurs qui ressemblent malencontreusement à un autre coda identitaire.

Le cas de CCS_Pm_0006

Cette théorie pourrait bel et bien s'appliquer dans le cadre de la journée du 26/05/21 où les deux clans EC1 et EC2 ont été entendus. En effet, l'individu CCS_Pm_0006 a été aperçu une fois et ce, en compagnie de l'individu 6348, appartenant au clan EC1 et que Vachon (2022) avait vu accompagné d'un veau en 2020. L'individu CCS_Pm_0006 pourrait donc être ce veau ayant gagné quelques marques lors de sa recapture en 2021. Au vu de la faible évolution des marques sur la caudale des cachalots d'une année à l'autre, cette théorie paraît moins probable puisque l'individu 6 était très marqué lors de sa recapture en 2021.

Une autre hypothèse pourrait être que l'individu CCS_Pm_0006 ait été classé à défaut comme appartenant au clan vocal EC1. En effet, à l'issue d'une seule journée d'enregistrement (26/05/21) où les deux clans EC1 et EC2 étaient présents et parce qu'il se trouvait en compagnie de l'individu 6348 dont le clan vocal est EC1, il n'a pas été envisagé que cet individu puisse appartenir au clan vocal EC2. En considérant que le nombre d'individus loupés est en moyenne de 2 individus par observation, il devient alors probable que plus d'individus que ceux ayant été aperçus à la surface aient contribué à l'émission de ces codas de type 5R et que l'individu CCS_Pm_0006 ainsi que les individus non-vus fassent en réalité bel et bien partie du clan vocal EC2.

Ainsi, même si sa caudale est fortement marquée, si l'on suit la logique de la première hypothèse, tout semble montrer que l'individu CCS_Pm_0006 est le veau avec lequel la femelle 6348 avait été aperçu en 2020. De plus, les marques sont acquises lors des premières années de vie et c'est à cette période que la caudale d'un cachalot change le plus. D'un autre côté, la présence simultanée des clans vocaux EC1 et EC2 sur les enregistrements du 21/05 vient semer le trouble quant au ralliement de ce nouvel individu non pas au clan vocal EC1, mais plutôt à EC2.

Présence simultanée des clans EC2 & EC3

Concernant la présence simultanée des groupes EC2 et EC3 le 23/09/22, elle s'explique par la présence de deux individus dont le clan vocal n'a encore jamais été décrit auparavant (Individu 5575 de DSWP et individu CCS_Pm_0019) et possiblement 3 individus loupés pour cette journée spécifiquement. De plus, les deux types de codas identifiés se basent sur un faible nombre de codas (respectivement 6 pour EC2 et 4 pour EC3). Considérant ce faible nombre, aucune hypothèse ne peut être établie sur l'appartenance certaine de chaque individu à un de deux clans vocaux ou encore sur la présence de ces deux clans de façon simultanée.

Un apprentissage social ?

Enfin, il a été mis en évidence (S. Gero et al., communication personnelle, 5 juin 2023) qu'il pouvait y avoir un apprentissage social entre les clans vocaux, si bien que des individus appartenant au clan vocal EC1 s'entraîneraient à imiter les codas 5R ou 9R des clans EC2 et EC3 et vice-versa. Un tel apprentissage social, combinée à la ségrégation spatiale connue entre chaque clan vocal, pourraient donc expliquer la « présence » simultanée de deux clans vocaux au sein d'un même enregistrement.

Enfin, il convient de rappeler que les limites des clans sont encore peu connues et que la compréhension de leurs distributions et de leurs « territoires » respectifs fait, par conséquent, toujours l'objet de recherches.

5- L'appartenance à un clan vocal, un facteur déterminant dans la distribution ?

La prédiction de la distribution n'a pas été faite à l'échelle des clans en raison d'un manque de données pour les clans EC2 et EC3. Cependant, l'analyse antérieure de la répartition des cachalots le long de la façade caraïbe met en évidence des zones de densité élevée clairement distinctes les unes des autres. Ces zones pourraient correspondre aux zones préalablement identifiées comme étant les habitats préférentiels distincts des différents clans vocaux (Vachon, Eguiguren, et al., 2022).

Des zones de densité intermédiaire ont été révélées par les prédictions autour des îles de la Guadeloupe et de la Dominique avec une probabilité élevée de rencontre (en moyenne autour de 0.6).

Cela vient appuyer les observations faites par le DSWP depuis 20 ans autour de ces deux îles (DSWP, 2019). La majorité des individus identifiés aux alentours de ces îles sont des individus du clan vocal EC1 (Figure 21, Figure 22), restant assez proches de la côte, où les conditions sont plus calmes. Or, ces deux zones sont caractérisées par des faibles vitesses de courants (Figure 7.B) et des canyons sous-marins importants. Cela pourrait alors constituer des habitats préférentiels pour les individus de ce clan.

Les cachalots du clan EC2, usuellement vus autour de Sainte-Lucie (Figure 19), pourraient être représentés ici par la zone de densité faible à intermédiaire autour de la Martinique et de Sainte-Lucie. Bien qu'éparses, les points d'observations de ces cachalots sont davantage situés au niveau des chenaux entre les îles (Figure 20, Figure 19.B) où les courants océaniques sont plus forts (Figure 7.B). Enfin, les cachalots du clan EC3, identifiés majoritairement dans la zone Sud entre Saint-Vincent et Grenade (Figure 21 & Figure 19), seraient plus attirés par des conditions de courants plus ardues, et des zones où la bathymétrie propose un large panel de profondeurs (Figure 7.A).

De surcroît, les fréquences prédites (Figure 24) semblent s'accorder sur un nombre de groupes de cachalots observés plus importants au Sud avec une probabilité d'observation quasiment égale à 1. Or, la plupart des groupements des clans vocaux EC2 et EC3 mis en évidence par Vachon (2022) présente un grand nombre d'individus. Etant donné que ces groupements/unités sociales sont nouveaux, de nombreux doutes subsistent encore sur la possible déclinaison d'un groupement en plusieurs sous-unités sociales (cas du groupement 20 et de l'unité sociale Y par exemple). Ainsi, le grand nombre de groupes prédit par le modèle pourrait faire référence au cas décrit ci-dessus et donc bel et bien représenter la distribution des clans vocaux EC2 et EC3 plus au Sud de l'île.

Enfin, les Figure 21 & Figure 22 ont permis de révéler des différences dans la longueur des déplacements entrepris en fonction de l'appartenance des unités sociales à un clan vocal. Cette différence dans leurs aires de répartition pourrait indiquer une flexibilité plus grande des clans qui se déplacent plus loin (EC3 par exemple), étant alors confrontés à une plus grande diversité d'habitats.

En définitive, comme démontré par Vachon, Eguiguren, et al. (2022), la ségrégation spatiale observée entre ces clans serait davantage due à la fidélité des clans envers les îles où ils sont nés et ont évolué, plutôt qu'à de réelles différences fondamentales dans les variables environnementales. Finalement, la culture agit ici comme une barrière empêchant complètement le mixage entre ces clans.

Cette forte territorialité à l'échelle des clans se retrouverait même à l'échelle de l'unité sociale, où la Figure 16.B a révélé des aires de répartition bien distinctes entre les unités sociales A et N, qui, lorsqu'elles se rencontrent, utilisent même un tiers endroit qui ne se chevauche ni avec l'une, ni avec l'autre des deux aires. Cette sédentarité accrue explique d'ailleurs les différences d'abondances estimées par la CCS et par Vachon (2022) (Tableau 3). En outre, ces différences ne sont ni liées à des différences dans la fréquence d'échantillonnage (8 campagnes d'échantillonnage de Vachon (2022) contre 12 pour la CCS) ou dans le taux de recapture (Tableau 4), mais bien à la zone échantillonnée plus ou moins proche des côtes, qui définit la présence ou l'absence des unités sociales.

Cependant, bien que rares, des enregistrements où on été entendus deux clans ensemble existent bien et pourraient alors remettre en question cette non-mixité entre clans. Il pourrait également s'agir d'individus mâles d'un clan vocal qui viendraient se reproduire avec les femelles d'un autre clan vocal, favorisant alors un plus grand brassage génétique.

6- Limitations et pistes d'améliorations

Limites liées aux données

D'une part, la présente étude se base sur deux années de suivi uniquement et 69 observations de cachalots ce qui peut être considéré comme peu, voire pas assez de données pour prédire la distribution d'une espèce. Cela peut donc rendre la prédiction du modèle moins fiable pour extrapoler aux zones non échantillonnées. D'ailleurs, l'évaluation de la performance du modèle sur le jeu de données de validation indique un nombre assez important de faux positifs. Cela est dû au fait que trop peu de données ont été utilisées ici pour valider le modèle (jeu de validation = neuf données d'observations réelles). Les cartes d'incertitude associées aux cartes de prédiction confirment la difficulté du modèle à extrapoler à des données non connues, avec des incertitudes de 10% au niveau de certains endroits de la côte Atlantique par exemple (Figure 24.B). Cependant, les très faibles valeurs d'incertitude observées sur le reste de la zone sont rassurantes concernant la fiabilité des prédictions obtenues.

D'autre part, l'objectif de la CCS est d'étudier la répartition de l'ensemble des cétacés dans les Petites Antilles. Cela amène donc à devoir faire des choix dans la collecte des données, c'est-à-dire de décider d'arrêter le suivi d'une espèce même lorsque l'ensemble des données n'a pas été totalement collecté si une autre espèce, peut-être considérée plus rare, est aperçue au même moment mais pas au même endroit. Dans le cas du cachalot, cela peut se traduire par des photos de caudale manquantes ou une non-estimation de la distance de la première détection. Cela peut impliquer une diminution de la qualité des données collectées et donc influencer l'estimation de la distance à laquelle se trouvait le cachalot. En effet, sur les 69 observations initiales, 46 uniquement ont contribué au calcul des distances de détectabilité à cause de données non collectées pour le reste des observations. De la même façon, sur ces 69 observations initiales, 18 ne présentent pas de photos et en moyenne 2 individus par observation ont été loupés.

Limites & piste d'amélioration liées à l'estimation de l'abondance

Notre estimation d'abondance ne tient pas compte des clans vocaux, en raison de données insuffisantes pour les clans EC2 et EC3. Or, il a été montré que les estimations d'abondance des populations peuvent être tronquées par plusieurs facteurs tels que la méthode d'échantillonnage ou le biais géographique si les individus d'une même espèce présentent une distribution spatiale inégale par exemple. Or, les cachalots se sont révélés faire partie de sociétés organisées de manière complexe mais stable, de telle sorte que les clans vocaux soient rarement, voire jamais, amenés à se rencontrer et occuperaient d'ailleurs des niches différentes (Vachon, Eguiguren, et al., 2022). Ceci est important du point de vue de la conservation, car une telle structure de population a le potentiel de fausser les estimations de la population.

Cependant, Vachon (2022) a montré que l'estimation de la population prenant en compte les clans vocaux EC1 et EC2 séparément donnait des résultats similaires à l'estimation considérant les deux clans ensemble. Toutefois, son analyse ne tenait pas compte du clan vocal EC3, sur lequel peu d'informations sont disponibles. L'estimation de l'abondance pourrait donc avoir été largement sous-estimé par la non-prise en compte de ce clan et d'autant plus, si l'on considère la faible temporalité sur laquelle l'ensemble des îles des Petites Antilles ont été échantillonnées.

L'apport de nouvelles données pourra nous permettre d'affiner notre estimation afin de permettre de discriminer au moins deux, voire trois clans distincts (EC1 et EC2) dans l'estimation de l'abondance.

Limites liées à l'établissement des relations sociales & de l'acoustique

L'étude de ces deux thématiques a été intéressante afin d'obtenir une première description des relations existantes entre les individus, entre les unités sociales et plus largement entre les clans vocaux. Nos deux années de données n'ont cependant pas suffi à affirmer de manière certaine l'appartenance des nouveaux individus à une unité sociale déjà décrite ou nouvelle, ou encore, à confirmer la présence de EC3 dans les Petites Antilles. De plus, l'appartenance d'un individu à un clan vocal a été déterminée par la présence de données d'observations en même temps que des données acoustiques dans lesquelles des codas étaient présents. Cependant, il est possible que les individus vus ne soient pas ceux qui ont émis les codas et vice-versa. La validation du clan vocal grâce aux données du DSWP a toutefois permis de pallier ce problème de « fausse » attribution potentielle.

Les différences observées sur les deux cartes de distribution (Figure 19.A, Figure 19.B) peuvent s'expliquer par le plus grand nombre de codas analysées par Vachon, Hersh, et al. (2022) ainsi que les données choisies pour calculer ces densités. En effet, dans la Figure 19.B, la densité des rencontres acoustiques a été estimée à partir des unités sociales de chaque clan et a la même résolution que celle utilisée par la CCS. Ainsi, les individus dont l'unité sociale ou le groupement n'était pas connu n'ont pas été pris en compte dans l'estimation de la distribution des clans vocaux de Vachon, Hersh, et al. (2022) contrairement à la Figure 19.A.

L'apport de données futures devrait nous permettre de préciser ces questions laissées en suspens. Pour l'acoustique, l'utilisation des logiciels de reconnaissance automatique des clics ou codas (script sur Matlab ; utilisation de RavenPro) pourrait permettre la détection des signaux acoustiques de faible intensité ou encore, de rendre plus rigoureuse l'analyse acoustique en elle-même en calculant des métriques associées à la détection des clics et codas (intervalles inter-clics & inter-pulse, durée de chaque coda, nombre précis de codas, etc).

Limites & pistes d'améliorations liées au modèle de distribution

Un effort non-uniforme & non-systématique

Bien qu'un gradient de concentration de chlorophylle existe entre le Nord et le Sud de la zone d'étude, aucune différence significative entre la côte Ouest et Est de l'arc impliquerait une productivité primaire plus importante sur sa partie Ouest qui expliquerait que la plupart des individus ait été vue sur la côte Ouest de l'arc insulaire. Il est donc possible que les différences observées entre les côtes sous le vent et dans le vent des Petites Antilles ne soient nullement le fait d'une différence marquée de leurs variables océanographiques ou physiographiques mais plutôt d'un effort de recherche plus marqué du côté Ouest que du côté Est de l'arc insulaire.

D'autre part, l'échantillonnage n'étant pas systématisé, il ne permet pas d'utiliser une méthode robuste telle que le « distance sampling » pour estimer la probabilité de détection. En effet, ne suivant pas de transect prédéfini, l'effort est influencé par le repérage des animaux puisque dès qu'un animal est détecté, la trajectoire du bateau est modifiée afin d'aller à sa rencontre. De plus, l'effort n'est pas uniforme sur l'ensemble de la zone d'étude puisqu'il s'intensifie au niveau de la Martinique (Figure 12), qui est le point de départ commun à toutes les expéditions, ce qui peut entraîner une surconsidération de cette aire dans le modèle prédiction.

A l'heure actuelle, la compréhension des facteurs environnementaux influençant réellement la distribution des cachalots reste encore incomplète et ne permet pas d'expliquer de manière certaine les prédictions du modèle de distribution. Amplifier l'effort sur la partie Est des Petites Antilles pourrait permettre de comprendre un peu plus le sens des probabilités d'observations prédites.

Un effort surestimé

Un premier biais, lié à la détection visuelle de cétacés de manière générale, est que cette détection est hautement dépendante des observateurs en question. Cette différence est d'autant plus vraie dans le cadre des expéditions de la CCS puisqu'elles regroupent des personnes venant d'horizons différents, des professionnels du domaine aux passionnés de la mer qui n'ont peut-être jamais vus de cétacés auparavant.

De surcroît, dans le cas spécifique du cachalot, son observation visuelle est fortement influencée par sa détection acoustique préalable.

Dans un premier temps, on pourrait penser qu'en prenant en compte en compte la distance de détectabilité associée à un effort d'observation largement influencé par la détection acoustique, on sous-estime la distance de détectabilité réelle qui est environ deux fois plus grande si on considère la position du premier enregistrement où on a détecté acoustiquement les cachalots (distance maximale de détectabilité visuelle = 2 km et distance maximale de détectabilité acoustique = 9 km).

Cependant, le processus de détection en lui-même vise à tracer acoustiquement les cachalots afin de s'en rapprocher le plus possible, ce qui augmente donc la probabilité de voir les cachalots en question. Ensuite, lorsque les cachalots sont repérés acoustiquement, l'effort visuel est alors décuplé, si bien que le nombre d'observateurs pendant cette période est plus grand que le nombre d'observateurs pendant le reste de l'effort. L'effort relatif à la détection des cachalots s'en retrouve alors dédoublé. Ainsi, l'effort associé aux moments où aucun cachalot n'a été aperçu n'est alors pas identique à celui pendant lequel des cachalots ont été vus. La distance à laquelle les cachalots sont repérés peut alors s'avérer plus grande lorsque tous les observateurs sont en observation que lorsque seulement deux le sont. Cela met en lumière une possible surestimation des distances de détectabilité considérées dans le modèle, plutôt qu'une sous-estimation de ces distances.

Les biais décrits peuvent donc influencer le processus de détection des cachalots et par conséquent, la distance à laquelle ils sont repérés, ce qui peut introduire un biais dans l'établissement de l'effort d'observation des cachalots. Ainsi, la prédiction pourrait avoir été influencée par cet effort surestimé.

Un moyen de pallier le problème de la forte dépendance des données visuelles par rapport aux données acoustiques serait de considérer directement les données acoustiques plutôt que les données d'observations dans le modèle. Pour cela, des modèles, tels que le GLM_GEE ou le GAM_GEE, existent et ont déjà été utilisés en modélisation de l'habitat pour les cachalots (Pirota et al., 2011; Vachon, Eguiguren, et al., 2022). La construction de ce genre de modèles est quelque peu différente de celle du GAM. En effet, le GAM utilise un effort d'échantillonnage propre à l'espèce, des données d'observations liées à cette espèce ainsi que des variables environnementales et proposent les prédicteurs les plus adaptés pour décrire les données.

Une des hypothèses fortes de ce modèle est de supposer l'indépendance des points GPS des observations les uns par rapport aux autres et c'est d'ailleurs pourquoi, on utilise en général une seule position GPS pour représenter la position de l'ensemble des cachalots, sur la totalité de la durée de l'observation. Bien que communément utilisée, cette considération peut sembler peu adaptée dans le cadre des cachalots, pour lesquels les suivis durent souvent plusieurs heures et peuvent s'étendre sur

des dizaines de kilomètres (23.6 km au maximum dans notre cas) et cela encore plus, si on considère les suivis acoustiques qui ont pu durer jusqu'à plus de 12h en 2021.

Ainsi, un moyen de considérer l'ensemble des positions GPS sans violer l'hypothèse d'indépendance entre les résidus du modèle GAM est d'utiliser des « Generalized Estimating Equations » ou GEE. Ainsi, l'ensemble des points GPS correspondant à une même observation ou un même suivi acoustique, forment un bloc dont l'identifiant est unique. Ainsi, bien que les points GPS au sein d'une même observation soient fortement corrélés les uns aux autres, les points GPS entre les blocs sont indépendants, ce qui permet de vérifier l'hypothèse d'indépendance des résidus. En prenant en considération l'ensemble des points GPS relatifs à l'observation des cachalots, ils s'affranchissent du calcul de l'effort d'observation. Ils sont donc très adaptés dans le cadre de données acoustiques, pour lesquelles il peut être très difficile d'établir une distance de détectabilité acoustique.

Il serait donc intéressant d'utiliser un de ces modèles pour réétudier la distribution des cachalots dans les Petites Antilles et voir si les résultats du modèle « visuel » et du modèle « acoustique » diffèrent beaucoup.

Limitations liées aux choix de la zone d'étude & des variables choisies

La prédiction de la distribution des cachalots peut être affectée négativement par plusieurs facteurs liés aux données collectées :

- Zone échantillonnée :

La zone échantillonnée est large ce qui est très favorable pour avoir un panel de l'ensemble des habitats rencontrés dans la Caraïbe mais la grande étendue qu'il faut parcourir à chaque expédition a pour conséquence que moins d'informations sont collectées sur une même zone. Cet effet de sous-échantillonnage peut induire une généralisation des données ponctuelles particulières à un espace plus vaste, ce qui peut fausser la prédiction de la distribution.

- Choix des variables environnementales :

Des proxys environnementaux ont été utilisés ici afin de prédire la présence potentielle des cachalots. Cependant, des variables sociales prenant en compte les relations entre les espèces (compétition (Heithaus & Dill, 2002), prédation (Reeves et al., 2007), les besoins reproductifs (Ersts & Rosenbaum, 2003) ou encore, la présence de veaux (Cañadas & Hammond, 2008), peuvent également fortement influencer la distribution d'une espèce (Pirota et al., 2011). De la même façon, l'utilisation d'outils permettant de simuler la distribution des calmars d'eaux profondes (SEAPODYM) aurait pu être plus pertinente pour prédire la distribution des cachalots (Virgili et al., 2021).

La sur-représentation du clan EC1

L'étude de la distribution à l'échelle du clan vocal n'a pas pu être conduite en raison du peu de données collectées pour les clans EC2 et EC3. Cependant, afin d'observer de possibles différences d'habitats entre les clans, un arbre de régression multivarié a été construit. Ce dernier se base sur les quatre prédicteurs qui sont ressortis dans le meilleur modèle. Il indique notamment des valeurs seuils à partir desquelles les divisions se font à l'échelle des unités sociales. Cependant, en raison du peu de données représentant le clan EC2 notamment, l'arbre donne des résultats aberrants qui révèlent des préférences environnementales presque complètement différentes pour des unités qui ont pourtant l'habitude de fréquenter les mêmes eaux (cas de V et T par exemple). Les résultats de cet arbre sont présentés en Annexe 10.

Le problème de fausse-absence

Dans le contexte de l'étude des cétacés, il est souvent difficile de collecter des données d'absence réelles. En effet, il est compliqué de déterminer si un groupe d'individus était réellement absent ou simplement non détecté lors de l'échantillonnage (Royle & Link, 2006), ce qui est particulièrement vrai pour les cachalots qui passent la majeure partie de leur temps sous l'eau et qui sont difficiles à repérer en surface. Cela est davantage accentué par le caractère non-systématique de l'échantillonnage (comportement opportuniste du bateau lors d'une observation) qui peut également générer des données de fausses absences. Dans notre cas, si la plupart des détections acoustiques mènent à une détection visuelle, il existe vingt enregistrements où des cachalots ont été entendus mais aucune observation n'a été faite (Annexe 1). En outre, la capacité d'un individu à repérer un animal dans l'eau peut être impactée par des conditions météorologiques adverses (pluie, brouillard, vents forts, etc) pouvant rendre plus difficile la détection ou l'observation des animaux marins et donc entraîner des données de fausses absences. L'utilisation de données d'absence incorrectes peut introduire un biais considérable dans l'analyse et fausser les résultats des prédictions de la distribution (Praca et al., 2009).

Même si les GAM permettent de gérer plus ou moins bien ces données, l'étude du paramètre de spécificité sur le jeu de données de validation indique une part importante des données de faux positifs, c'est-à-dire des prédictions de présence aux endroits où des absences ont été observées. Deux raisons peuvent expliquer cela :

- Les données d'absence initiales sont des données de fausses absences et le modèle prédit alors des données de présence à juste titre ;
- Le peu de données de présence du jeu de validation (dû au nombre limité d'observations des données initiales) implique une part importante de données d'absence qui pourrait venir tronquer les prédictions du modèle.

Une façon de pallier ce problème de fausses absences sans changer la méthode d'échantillonnage serait de ne pas prendre en compte l'effort actif total pour décrire la distribution mais plutôt de sélectionner aléatoirement des portions de cet effort, ce qui permettrait alors réellement de considérer que chaque zone échantillonnée l'a été de manière uniforme.

Une faible déviance expliquée

Toute l'analyse liée aux prédictions de la distribution est basée sur le choix du meilleur modèle, étant considéré comme celui dont l'AIC est le plus faible et la déviance expliquée la plus forte. Cependant, dans notre cas précisément, la déviance expliquée du meilleur modèle est de 24.1% seulement. Cela signifie que le modèle qui génère nos prédictions finales n'explique que 24.1% des données réellement observées. Plusieurs améliorations sont possibles sur notre modèle GAM telles que la prise en compte de termes d'interactions qui permettraient de considérer l'influence de deux variables dont les variations sont liées sans pour autant être corrélées ; ou encore, la prise en compte du caractère linéaire de certaines des courbes de lissage (cas de la relation entre les données d'observations et la distance à la côte). Afin d'obtenir des prédictions prenant mieux en compte la diversité de nos données d'observations initiales, il pourrait être possible de tester plusieurs types de modèles (GLM, GEE, modèles de machine Learning) et de combiner les résultats des prédictions obtenus sur notre zone d'étude (HUGON & MAALOUF, 2023) .

Pour conclure, la collecte de nouvelles données pourrait permettre l'amélioration des prédictions du modèle et également de pouvoir dissocier les différents clans pour observer des différences de distribution possibles. Cela pourra également permettre de discriminer les deux saisons, sèche et humide, afin de montrer une possible différence d'habitat en fonction de la saisonnalité.

V- Conclusion

La présente étude s'inscrit dans la continuité des recherches sur les cachalots entreprises par le DSWP depuis 1995. Elle permet notamment de combler l'absence de données collectées par le DSWP sur les années 2021 & 2022. Elle reprend le cadre d'étude établi par Vachon (2022), qui visait à élargir l'étude des cachalots de la Guadeloupe et de la Dominique aux autres îles des Petites Antilles, d'Anguilla à Grenade. En recapturant 45 individus déjà identifiés et en capturant 17 nouveaux individus, la CCS contribue à l'estimation de la taille globale de la population. Ce rapport a également mis en évidence la possible découverte de nouvelles unités sociales et de nouveaux individus appartenant à des unités sociales déjà connues. La collecte de données sur l'ensemble de l'arc insulaire a également permis de préciser les déplacements à plus grande échelle des individus, puisque sur les 20 individus, déjà connus et recapturés par la CCS, ayant parcouru plus de 100 km, la recapture par la CCS de ces individus a toujours contribué à l'établissement de la distance maximale.

L'analyse acoustique a fourni des informations clés sur la répartition des trois principaux clans vocaux dans les Caraïbes, confirmant notamment la présence certaine du clan EC2 dans l'arc insulaire, en particulier aux îles de Sainte-Lucie et de Saint-Vincent. Bien que les données sur le clan EC3 soient encore limitées, sa possible présence dans le sud des Petites Antilles, associée à son comportement migratoire plus étendu par rapport aux autres clans, laisse présager de futures découvertes intéressantes concernant la répartition de ces clans et leurs interactions.

Enfin, un modèle de type GAM a été utilisé pour prédire la distribution des cachalots dans les Petites Antilles, en mettant en évidence trois variables environnementales essentielles : la bathymétrie, la chlorophylle A et la distance à la côte. La carte prédictive a souligné l'importance de la zone sud (de Saint-Vincent à Grenade) et de la Guadeloupe dans cette distribution. La carte de prédiction a également révélé des zones potentielles de fortes densités dans la partie atlantique, notamment sur la côte est de la Guadeloupe, ce qui pourrait s'expliquer par une chute prononcée de la profondeur dans cette zone.

Il convient de prendre en compte certains biais tels que la méthode d'échantillonnage et la durée limitée de l'échantillonnage, qui ont tendance à sous-estimer la taille de la population de cachalots dans les Petites Antilles, en particulier sur leur façade atlantique. Par conséquent, toutes les hypothèses émises doivent être considérées avec prudence, en attendant de pouvoir les vérifier ou les infirmer avec le temps.

Ces résultats soulignent l'importance de la collaboration entre les îles des Petites Antilles pour la conservation et la protection de cette population de cachalots unique. Alors que la chasse à la baleine a été pratiquée pendant des siècles, l'impact humain sur les cachalots n'a jamais été aussi important qu'à notre siècle, où nous le faisons plus par ignorance que par intention. Cela est d'autant plus préoccupant si l'on considère la non-mixité des clans vocaux préalablement décrits, puisque la prospérité de l'espèce dépend alors de la survie de chacun de ces clans. Ceci, combiné au déclin de 6.2% par an observé, rend la population de cachalots dans les Petites Antilles encore plus vulnérable que ce que l'on imaginait auparavant.

C'est pourquoi la CCS poursuivra ses actions de conservation au cours des prochaines années, en étendant la zone d'étude jusqu'aux îles ABC, proches du Venezuela et en intensifiant ses efforts de

sensibilisation afin d'assurer la survie à long terme de ces animaux et de maintenir l'équilibre des écosystèmes marins.

VI- Bibliographie

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723. <https://doi.org/10.1109/TAC.1974.1100705>
- Andreas, J., Beguš, G., Bronstein, M. M., Diamant, R., Delaney, D., Gero, S., Goldwasser, S., Gruber, D. F., de Haas, S., Malkin, P., Pavlov, N., Payne, R., Petri, G., Rus, D., Sharma, P., Tchernov, D., Tønnesen, P., Torralba, A., Vogt, D., & Wood, R. J. (2022). Toward understanding the communication in sperm whales. *iScience*, 25(6), 104393. <https://doi.org/10.1016/j.isci.2022.104393>
- Battin, J. (2004). When Good Animals Love Bad Habitats : Ecological Traps and the Conservation of Animal Populations: *Ecological Traps. Conservation Biology*, 18(6), 1482-1491. <https://doi.org/10.1111/j.1523-1739.2004.00417.x>
- Blount, D., Gero, S., Van Oast, J., Parham, J., Kingen, C., Scheiner, B., Stere, T., Fisher, M., Minton, G., Khan, C., Dulau, V., Thompson, J., Moskvayak, O., Berger-Wolf, T., Stewart, C. V., Holmberg, J., & Levenson, J. J. (2022). Flukebook : An open-source AI platform for cetacean photo identification. *Mammalian Biology*, 102(3), 1005-1023. <https://doi.org/10.1007/s42991-021-00221-3>
- Brakes, P., Dall, S. R. X., Aplin, L. M., Bearhop, S., Carroll, E. L., Ciucci, P., Fishlock, V., Ford, J. K. B., Garland, E. C., Keith, S. A., McGregor, P. K., Mesnick, S. L., Noad, M. J., Di Sciara, G. N., Robbins, M. M., Simmonds, M. P., Spina, F., Thornton, A., Wade, P. R., ... Rutz, C. (2019). Animal cultures matter for conservation. *Science*, 363(6431), 1032-1034. <https://doi.org/10.1126/science.aaw3557>
- Bridge, P. D. (1993). *Classification. Pages 219-242 in J. C. Fry, ed. Biological data analysis. Oxford University Press, Oxford, UK.*
- Cañadas, A., & Hammond, P. (2008). Abundance and habitat preferences of the short-beaked common dolphin *Delphinus delphis* in the southwestern Mediterranean : Implications for conservation. *Endangered Species Research*, 4, 309-331. <https://doi.org/10.3354/esr00073>
- Cantor, M., Gero, S., Whitehead, H., & Rendell, L. (2019). Sperm Whale : The Largest Toothed Creature on Earth. In B. Würsig (Éd.), *Ethology and Behavioral Ecology of Odontocetes* (p. 261-280). Springer International Publishing. https://doi.org/10.1007/978-3-030-16663-2_12
- Carwardine, M., Camm, M., Robinson, R., & Llobet, T. (2020). *Handbook of whales, dolphins and porpoises*. Bloomsbury.
- Chambault, P., Fossette, S., Heide-Jørgensen, M. P., Jouannet, D., & Vély, M. (2021). Predicting seasonal movements and distribution of the sperm whale using machine learning algorithms. *Ecology and Evolution*, 11(3), 1432-1445. <https://doi.org/10.1002/ece3.7154>
- Christal, J., Whitehead, H., & Lettevall, E. (1998). Sperm whale social units : Variation and change. *Canadian Journal of Zoology*, 76(8), 1431-1440. <https://doi.org/10.1139/z98-087>
- Cranford, T. W. (1999). THE SPERM WHALE'S NOSE : SEXUAL SELECTION ON A GRAND SCALE?1. *Marine Mammal Science*, 15(4), 1133-1157. <https://doi.org/10.1111/j.1748-7692.1999.tb00882.x>
- Davis, R. W., Fargion, G. S., May, N., Leming, T. D., Baumgartner, M., Evans, W. E., Hansen, L. J., & Mullin, K. (1998). PHYSICAL HABITAT OF CETACEANS ALONG THE CONTINENTAL SLOPE IN THE NORTHCENTRAL AND WESTERN GULF OF MEXICO. *Marine Mammal Science*, 14(3), 490-507. <https://doi.org/10.1111/j.1748-7692.1998.tb00738.x>
- Ellis, R. (2018). Whaling, Aboriginal and Western Traditional. In *Encyclopedia of Marine Mammals* (p. 1054-1063). Elsevier. <https://doi.org/10.1016/B978-0-12-804327-1.00269-7>
- Ersts, P. J., & Rosenbaum, H. C. (2003). Habitat preference reflects social organization of humpback whales (*Megaptera novaeangliae*) on a wintering ground. *Journal of Zoology*, 260(4), 337-345. <https://doi.org/10.1017/S0952836903003807>
- Evans, P. G. H. (1990). *The natural history of whales & dolphins* (Reprinted). Helm.

- Fais, A., Lewis, T. P., Zitterbart, D. P., Álvarez, O., Tejedor, A., & Aguilar Soto, N. (2016). Abundance and Distribution of Sperm Whales in the Canary Islands : Can Sperm Whales in the Archipelago Sustain the Current Level of Ship-Strike Mortalities? *PLOS ONE*, *11*(3), e0150660. <https://doi.org/10.1371/journal.pone.0150660>
- Fielding, A. H., & Bell, J. F. (1997). A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, *24*(1), 38-49. <https://doi.org/10.1017/S0376892997000088>
- GANDILHON, N., Robin, J.-P., Clapham, P., Nolibé, G., Adam, O., Louis, M., & Guinet, C. (2012). *CONTRIBUTION AU RECENSEMENT DES CÉTACÉS DANS L'ARCHIPEL DE GUADELOUPE*.
- Gero, S. (Réalisateur). (2017, septembre 28). *Les cultures perdues des baleines*. TED. https://www.ted.com/talks/shane_gero_the_lost_cultures_of_whales/transcript?language=fr
- Gero, S., Bøttcher, A., Whitehead, H., & Madsen, P. T. (2016). Socially segregated, sympatric sperm whale clans in the Atlantic Ocean. *Royal Society Open Science*, *3*(6), 160061. <https://doi.org/10.1098/rsos.160061>
- Gero, S., Engelhaupt, D., Rendell, L., & Whitehead, H. (2009). Who Cares? Between-group variation in alloparental caregiving in sperm whales. *Behavioral Ecology*, *20*(4), 838-843. <https://doi.org/10.1093/beheco/arp068>
- Gero, S., Gordon, J., Carlson, C., Evans, P., & Whitehead, H. (2007). *Population estimate and inter-island movement of sperm whales in the eastern caribbean*. 143-150.
- Gero, S., Gordon, J., & Whitehead, H. (2013). Calves as social hubs : Dynamics of the social network within sperm whale units. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1763), 20131113. <https://doi.org/10.1098/rspb.2013.1113>
- Gero, S., Milligan, M., Rinaldi, C., Francis, P., Gordon, J., Carlson, C., Steffen, A., Tyack, P., Evans, P., & Whitehead, H. (2014). Behavior and social structure of the sperm whales of Dominica, West Indies. *Marine Mammal Science*, *30*(3), 905-922. <https://doi.org/10.1111/mms.12086>
- Gero, S., Simon, L., & Bernus, J. (2023, juin 5). *Présentation des résultats partiels de la CCS (photo-identification & acoustique)*. [Communication personnelle].
- Gero, S., & Whitehead, H. (2016). Critical Decline of the Eastern Caribbean Sperm Whale Population. *PLOS ONE*, *11*(10), e0162019. <https://doi.org/10.1371/journal.pone.0162019>
- Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, *3*(1), 150372. <https://doi.org/10.1098/rsos.150372>
- Gillespie, D., Thode, A., Mellinger, D. K., Deng, X. Y., Gordon, J., Trinder, P. W., McLaren, D., McHugh, R., & Redmond, P. (2008). PAMGUARD: semiautomated, open source software for real-time acoustic detection and localisation of Cetaceans. In *Conference on underwater noise measurement, impact and mitigation, Eli Research D*. Curran Associates Inc.
- Godde, S., Humbert, L., Côté, S. D., Réale, D., & Whitehead, H. (2013). Correcting for the impact of gregariousness in social network analyses. *Animal Behaviour*, *85*(3), 553-558. <https://doi.org/10.1016/j.anbehav.2012.12.010>
- Hann, C. H., Smith, T. D., & Torres, L. G. (2016). A sperm whale's perspective : The importance of seasonality and seamount depth. *Marine Mammal Science*, *32*(4), 1470-1481. <https://doi.org/10.1111/mms.12320>
- Hastie, T., & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Sciences*, *Vol.1*(No.3), 297-318.
- Heithaus, M. R., & Dill, L. M. (2002). FOOD AVAILABILITY AND TIGER SHARK PREDATION RISK INFLUENCE BOTTLENOSE DOLPHIN HABITAT USE. *Ecology*, *83*(2), 480-491. [https://doi.org/10.1890/0012-9658\(2002\)083\[0480:FAATSP\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2002)083[0480:FAATSP]2.0.CO;2)
- Hijmans, R. J. (2020). *Package raster de R: Geographic Data Analysis and Modeling; R Package* (Version 3.0-12) [Logiciel].
- Hoyt, E., & Hvenegaard, G. T. (2002). A Review of Whale-Watching and Whaling with Applications for the Caribbean. *Coastal Management*, *30*(4), 381-399. <https://doi.org/10.1080/089207502900273>

- HUGON, F., & MAALOUF, J. P. (2023). *Analyse des habitats des mammifères marins dans le Sanctuaire Agoa*. Office Français de la Biodiversité.
- Jaquet, N. (1996). How spatial and temporal scales influence understanding of Sperm Whale distribution : A review. *Mammal Review*, 26(1), 51-65. <https://doi.org/10.1111/j.1365-2907.1996.tb00146.x>
- Laran, S., Bassols, N., Dorémus, G., Authier, M., Ridoux, V., & Van Canneyt, O. (2019). *Distribution et abondance de la mégafaune marine aux Petites Antilles et en Guyane. Campagne REMMOA - II. Rapport final pour l'Agence Française pour la Biodiversité*. (p. 80+ XXXII pp.).
- Lavery, T. J., Roudnew, B., Gill, P., Seymour, J., Seuront, L., Johnson, G., Mitchell, J. G., & Smetacek, V. (2010). Iron defecation by sperm whales stimulates carbon export in the Southern Ocean. *Proceedings of the Royal Society B: Biological Sciences*, 277(1699), 3527-3531. <https://doi.org/10.1098/rspb.2010.0863>
- Levenson, J., Gero, S., Van Oast, J., & Holmberg, J. (2015). *Flukebook : A cloud-based photo-identification analysis tools for marine mammal research* [Logiciel]. <https://www.flukebook.org>
- Lundquist, D., Sironi, M., Würsig, B., Rowntree, V., Martino, J., & Lundquist, L. (2013). Response of southern right whales to simulated swim-with-whale tourism at Península Valdés, Argentina. *Marine Mammal Science*, 29(2), E24-E45. <https://doi.org/10.1111/j.1748-7692.2012.00583.x>
- Mannocci, L., Catalogna, M., Dorémus, G., Laran, S., Lehodey, P., Massart, W., Monestiez, P., Van Canneyt, O., Watremez, P., & Ridoux, V. (2014). Predicting cetacean and seabird habitats across a productivity gradient in the South Pacific gyre. *Progress in Oceanography*, 120, 383-398. <https://doi.org/10.1016/j.pocean.2013.11.005>
- Miller, P. J. O., Johnson, M. P., & Tyack, P. L. (2004). Sperm whale behaviour indicates the use of echolocation click buzzes 'creaks' in prey capture. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271(1554), 2239-2247. <https://doi.org/10.1098/rspb.2004.2863>
- Milligan, G. W., & Cooper, M. C. (1987). Methodology Review : Clustering Methods. *Applied Psychological Measurement*, 11(4), 329-354. <https://doi.org/10.1177/014662168701100401>
- Mizroch, S. A., & Rice, D. W. (2013). Ocean nomads : Distribution and movements of sperm whales in the North Pacific shown by whaling data and Discovery marks. *Marine Mammal Science*, 29(2), E136-E165. <https://doi.org/10.1111/j.1748-7692.2012.00601.x>
- Moors-Murphy, H. B. (2014). Submarine canyons as important habitat for cetaceans, with special reference to the Gully : A review. *Deep Sea Research Part II: Topical Studies in Oceanography*, 104, 6-19. <https://doi.org/10.1016/j.dsr2.2013.12.016>
- N., J., & D., G. (2002). Distribution and relative abundance of sperm whales in relation to key environmental features, squid landings and the distribution of other cetacean species in the Gulf of California, Mexico. *Marine Biology*, 141(3), 591-601. <https://doi.org/10.1007/s00227-002-0839-0>
- Newman, M. E. J. (2004). Analysis of weighted networks. *Physical Review E*, 70(5), 056131. <https://doi.org/10.1103/PhysRevE.70.056131>
- Newman, M. E. J. (2006). Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23), 8577-8582. <https://doi.org/10.1073/pnas.0601602103>
- Pirotta, E., Matthiopoulos, J., MacKenzie, M., Scott-Hayward, L., & Rendell, L. (2011). Modelling sperm whale habitat preference : A novel approach combining transect and follow data. *Marine Ecology Progress Series*, 436, 257-272. <https://doi.org/10.3354/meps09236>
- Praca, E., Gannier, A., Das, K., & Laran, S. (2009). Modelling the habitat suitability of cetaceans : Example of the sperm whale in the northwestern Mediterranean Sea. *Deep Sea Research Part I: Oceanographic Research Papers*, 56(4), 648-657. <https://doi.org/10.1016/j.dsr.2008.11.001>
- Redfern, J., Ferguson, M., Becker, E., Hyrenbach, K., Good, C., Barlow, J., Kaschner, K., Baumgartner, M., Forney, K., Ballance, L., Fauchald, P., Halpin, P., Hamazaki, T., Pershing, A., Qian, S., Read, A., Reilly, S., Torres, L., & Werne, F. (2006). Techniques for cetacean habitat modeling. *Marine Ecology Progress Series*, 310, 271-295. <https://doi.org/10.3354/meps310271>

- Reeves, R. R., Berger, J., & Clapham, P. J. (2007). Killer Whales as Predators of Large Baleen Whales and Sperm Whales. In J. Estes (Éd.), *Whales, Whaling, and Ocean Ecosystems* (p. 174-187). University of California Press. <https://doi.org/10.1525/california/9780520248847.003.0014>
- Rendell, L. E., & Whitehead, H. (2003). Vocal clans in sperm whales (*Physeter macrocephalus*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1512), 225-231. <https://doi.org/10.1098/rspb.2002.2239>
- Rendell, L., & Frantzis, A. (2016). Mediterranean Sperm Whales, *Physeter macrocephalus*. In *Advances in Marine Biology* (Vol. 75, p. 37-74). Elsevier. <https://doi.org/10.1016/bs.amb.2016.08.001>
- Rendell, L., & Whitehead, H. (2005). Spatial and temporal variation in sperm whale coda vocalizations : Stable usage and local dialects. *Animal Behaviour*, 70(1), 191-198. <https://doi.org/10.1016/j.anbehav.2005.03.001>
- Rice, A., Širović, A., Trickey, J. S., Debich, A. J., Gottlieb, R. S., Wiggins, S. M., Hildebrand, J. A., & Baumann-Pickering, S. (2021). Cetacean occurrence in the Gulf of Alaska from long-term passive acoustic monitoring. *Marine Biology*, 168(5), 72. <https://doi.org/10.1007/s00227-021-03884-1>
- Rinaldi, C., & Rinaldi, R. (2007). *Les cétacés dans l'archipel guadeloupéen et dans la caraïbe. Etat des connaissances et perspectives 1998 – 2006. Rapport AET mars 2007.*
- Rinaldi, C., Rinaldi, R., Laine, J., & Barbraud, C. (2021). Population dynamics of sperm whales (*PHYSETER MACROCEPHALUS*) in Guadeloupe, French Caribbean : A mark-recapture study from 2001 to 2013. *Marine Mammal Science*, 37(4), 1391-1405. <https://doi.org/10.1111/mms.12837>
- Royle, J. A., & Link, W. A. (2006). GENERALIZED SITE OCCUPANCY MODELS ALLOWING FOR FALSE POSITIVE AND FALSE NEGATIVE ERRORS. *Ecology*, 87(4), 835-841. [https://doi.org/10.1890/0012-9658\(2006\)87\[835:GSOMAF\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2006)87[835:GSOMAF]2.0.CO;2)
- SANDRON, F., SAISHO, M., GANNON, F., MOUYSSSET, L., RAKOTOHARIMALALA, S., MUÑOZ FORERO, P. A., ANDRIANANTOANDRO, T., FAULET, A., ANDRIAMARO, F., & RASOLOFOMANANA, R. (2015). *Résultats détaillés du programme AS2W; Analyse socioéconomique du whale watching à Madagascar et La Réunion.*
- Sing, T., Sander, O., Beerenwinkel, N., & Lengauer, T. (2005). ROC R : Visualizing classifier performance in R. *Bioinformatics*, 21(20), 3940-3941. <https://doi.org/10.1093/bioinformatics/bti623>
- Sivle, L. D., Kvadsheim, P. H., Fahlman, A., Lam, F. P. A., Tyack, P. L., & Miller, P. J. O. (2012). Changes in dive behavior during naval sonar exposure in killer whales, long-finned pilot whales, and sperm whales. *Frontiers in Physiology*, 3. <https://doi.org/10.3389/fphys.2012.00400>
- Srinivasan, T. M. (2015). Historical Note : Ambergris in Perfumery in the Past and Present Indian Context and the Western World. *Indian Journal of History of Science*, 50(2). <https://doi.org/10.16943/ijhs/2015/v50i2/48241>
- The Dominica Sperm Whale Project (DSWP). (2019). *The Fluke Book : The Natural History and Family Units of Sperm Whales off Dominica.* <http://www.thespermwhaleproject.org>
- Vachon, F. (2022). *ON CULTURAL INHERITANCE: EVOLUTION, BEHAVIOUR AND SOCIAL STRUCTURE OF EASTERN CARIBBEAN SPERM WHALES.*
- Vachon, F., Eguiguren, A., Rendell, L., Gero, S., & Whitehead, H. (2022). Distinctive, fine-scale distribution of Eastern Caribbean sperm whale vocal clans reflects island fidelity rather than environmental variables. *Ecology and Evolution*, 12(11). <https://doi.org/10.1002/ece3.9449>
- Vachon, F., Hersh, T. A., Rendell, L., Gero, S., & Whitehead, H. (2022). Ocean nomads or island specialists? Culturally driven habitat partitioning contrasts in scale between geographically isolated sperm whale populations. *Royal Society Open Science*, 9(5), 211737. <https://doi.org/10.1098/rsos.211737>
- Virgili, A., Authier, M., Boisseau, O., Cañadas, A., Claridge, D., Cole, T., Corkeron, P., Dorémus, G., David, L., Di-Méglio, N., Dunn, C., Dunn, T. E., García-Barón, I., Laran, S., Lauriano, G., Lewis, M., Louzao, M., Mannocci, L., Martínez-Cedeira, J., ... Ridoux, V. (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. *Global Ecology and Biogeography*, 28(3), 300-314. <https://doi.org/10.1111/geb.12850>
- Virgili, A., Hedon, L., Authier, M., Calmettes, B., Claridge, D., Cole, T., Corkeron, P., Dorémus, G., Dunn, C., Dunn, T. E., Laran, S., Lehodey, P., Lewis, M., Louzao, M., Mannocci, L., Martínez-Cedeira, J., Monestiez, P., Palka, D., Pettex, E., ... Ridoux, V. (2021). Towards a better characterisation of deep-diving whales' distributions by using prey distribution model outputs? *PLOS ONE*, 16(8), e0255667. <https://doi.org/10.1371/journal.pone.0255667>
- Waring, G. T. (2004). *Sperm Whales : Social Evolution in the Ocean* . By Hal Whitehead. Chicago (Illinois): University of Chicago Press . \$80.00 (hardcover); \$30.00 (paper). xxiii + 431 p; ill.; index. ISBN: 0-226-89517-3 (hc); 0-226-89518-1 (pb). 2003. *The Quarterly Review of Biology*, 79(2), 217-218. <https://doi.org/10.1086/423094>
- Watkins, W. A. (1977). Sperm whale codas. *The Journal of the Acoustical Society of America*, 62(6), 1485. <https://doi.org/10.1121/1.381678>
- Watkins, W. A., & Moore, K. E. (1982). An underwater acoustic survey for sperm whales (*Physeter catodon*) and other cetaceans in the southeast Caribbean. *Cetology*, 1-7.
- Weideman, H. J., Jablons, Z. M., Holmberg, J., Flynn, K., Calambokidis, J., Tyson, R. B., Allen, J. B., Wells, R. S., Hupman, K., Urian, K., & Stewart, C. V. (2017). *Integral Curvature Representation and Matching Algorithms for Identification of Dolphins and Whales* (arXiv:1708.07785). arXiv. <http://arxiv.org/abs/1708.07785>
- Weir, C. R., Frantzis, A., Alexiadou, P., & Goold, J. C. (2007). The burst-pulse nature of 'squeal' sounds emitted by sperm whales (*Physeter macrocephalus*). *Journal of the Marine Biological Association of the United Kingdom*, 87(1), 39-46. <https://doi.org/10.1017/S0025315407054549>
- Whitehead, H. (2002). Estimates of the current global population size and historical trajectory for sperm whales. *Marine Ecology Progress Series*, 242, 295-304. <https://doi.org/10.3354/meps242295>
- Whitehead, H. (2003). *Sperm whales : Social evolution in the ocean*. University of Chicago Press.
- Whitehead, H. (2008). Analyzing animal societies : Quantitative methods for vertebrate social analysis. *Choice Reviews Online*, 46(10), 46-5619. <https://doi.org/10.5860/CHOICE.46-5619>
- Whitehead, H. (2009). *SOCPROG programs : Analyzing animal social structures* [MATLAB]. Behavioral Ecology and Sociobiology 63: 765-778.
- Whitehead, H. (2018). Sperm Whale. In *Encyclopedia of Marine Mammals* (p. 919-925). Elsevier. <https://doi.org/10.1016/B978-0-12-804327-1.00242-9>
- Whitehead, H., Antunes, R., Gero, S., Wong, S. N. P., Engelhaupt, D., & Rendell, L. (2012). Multilevel Societies of Female Sperm Whales (*Physeter macrocephalus*) in the Atlantic and Pacific : Why Are They So Different? *International Journal of Primatology*, 33(5), 1142-1164. <https://doi.org/10.1007/s10764-012-9598-z>
- Whitehead, H., & Gero, S. (2015). Conflicting rates of increase in the sperm whale population of the eastern Caribbean : Positive observed rates do not reflect a healthy population. *Endangered Species Research*, 27(3), 207-218. <https://doi.org/10.3354/esr00657>
- Whitehead, H., & Weilgart, L. (1991). Patterns of Visually Observable Behaviour and Vocalizations in Groups of Female Sperm Whales. *Behaviour*, 118(3-4), 275-296. <https://doi.org/10.1163/156853991X00328>
- Wood, S. (2012). *mgcv : Mixed GAM Computation Vehicle with GCV/AIC/REML smoothness estimation* [Logiciel].
- Wood, S. N. (2006). Low-Rank Scale-Invariant Tensor Product Smooths for Generalized Additive Mixed Models. *Biometrics*, 62(4), 1025-1036. <https://doi.org/10.1111/j.1541-0420.2006.00574.x>
- Yee, T. W., & Wild, C. J. (1996). Vector Generalized Additive Models. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(3), 481-493. <https://doi.org/10.1111/j.2517-6161.1996.tb02095.x>

VII- Annexes

Annexe 1 : Description des enregistrements totaux, exploitables et de ceux où étaient présents des codas

	Année(s)	Enregistrements + observations	Enregistrements seuls	Total journées d'enregistrements
Enregistrements audibles	2022	26/03 ; 27/03 ; 28/03 ; 30/03 ; 22/04 ; 24/04 ; 19/06 ; 20/06 ; 22/06 ; 16/07 ; 23/07 ; 16/08 ; 17/08 ; 29/08 ; 23/09 ; 24/09 ; 27/09 ; 28/09 ; 29/09	29/03 ; 02/04 ; 03/04 ; 04/04 ; 26/04 ; 18/05 ; 19/05 ; 21/05 ; 30/05 ; 28/06 ; 15/07 ; 17/07 ; 19/07 ; 20/07 ; 22/07 ; 24/07 ; 15/08	36
	2021	21/05 ; 23/05 ; 26/05 ; 28/05 ; 03/06 ; 12/06 ; 26/06 ; 27/06 ; 07/07 ; 12/07 ; 21/07 ; 14/09 ;	27/05 ; 23/06 ; 06/07 ; 05/08	16
Enregistrements inaudibles ou autres problèmes*	2022	26/03 ; 27/03 ; 28/03 ; 30/03 ; 19/06 ; 16/07 ; 23/07	03/04 ; 18/05 ; 21/05 ; 30/05 ; 28/06 ; 29/03 15/07 ; 20/07 ; 22/07 ; 24/07 ; 15/08	18
	2021	12/07		1
Total de journées d'enregistrements exploitables	2021/2022	23	10	33
Dont codas	2021	21/05 ; 23/05 ; 26/05 ; 28/05 ; 03/06 ; 12/06 ; 26/06 ; 27/06 ; 07/07 ; 21/07	23/06 ; 06/07	24
	2022	24/04 ; 19/06 ; 20/06 ; 22/06 ; 16/08 ; 29/08 ; 23/09 ; 24/09 ; 27/09 ; 28/09 ; 29/09		

Annexe 2 : Liste des produits Copernicus desquelles les variables sont extraites (Annexe 2.A) et explicitation du choix des paramètres moyennes et variances pour chaque variable (Annexe 2.B)

Annexe 2.A :

Variable	Produit Copernicus	Dataset
Sea Surface Temperature-SST(°C)	Global Ocean OSTIA (“Operational Sea Surface Temperature and Ice Analysis”)	METOFFICE-GLO-SST-L4-REP-OBS-SST
Concentration en chlorophylle A - CHL (mg.m ⁻³)	The Operational Mercator Ocean biogeochemical global ocean analysis and forecast	global-analysis-forecast-bio-001-028-daily
Température du fond (°C)	The Operational Mercator global ocean analysis and forecast system	cmems_mod_glo_phy_anfc_0.083deg_P1D-m
Hauteur des eaux de surface par rapport au géoïde-SSH (m)		cmems_mod_glo_phy_anfc_0.083deg_P1D-m
Profondeur de la couche de mélange-MLD (m)		cmems_mod_glo_phy_anfc_0.083deg_P1D-m
Composante Est de la vitesse- Uo (m/s)		cmems_mod_glo_phy-cur_anfc_0.083deg_P1D-m
Composante Nord de la vitesse- Vo (m/s)		cmems_mod_glo_phy-cur_anfc_0.083deg_P1D-m

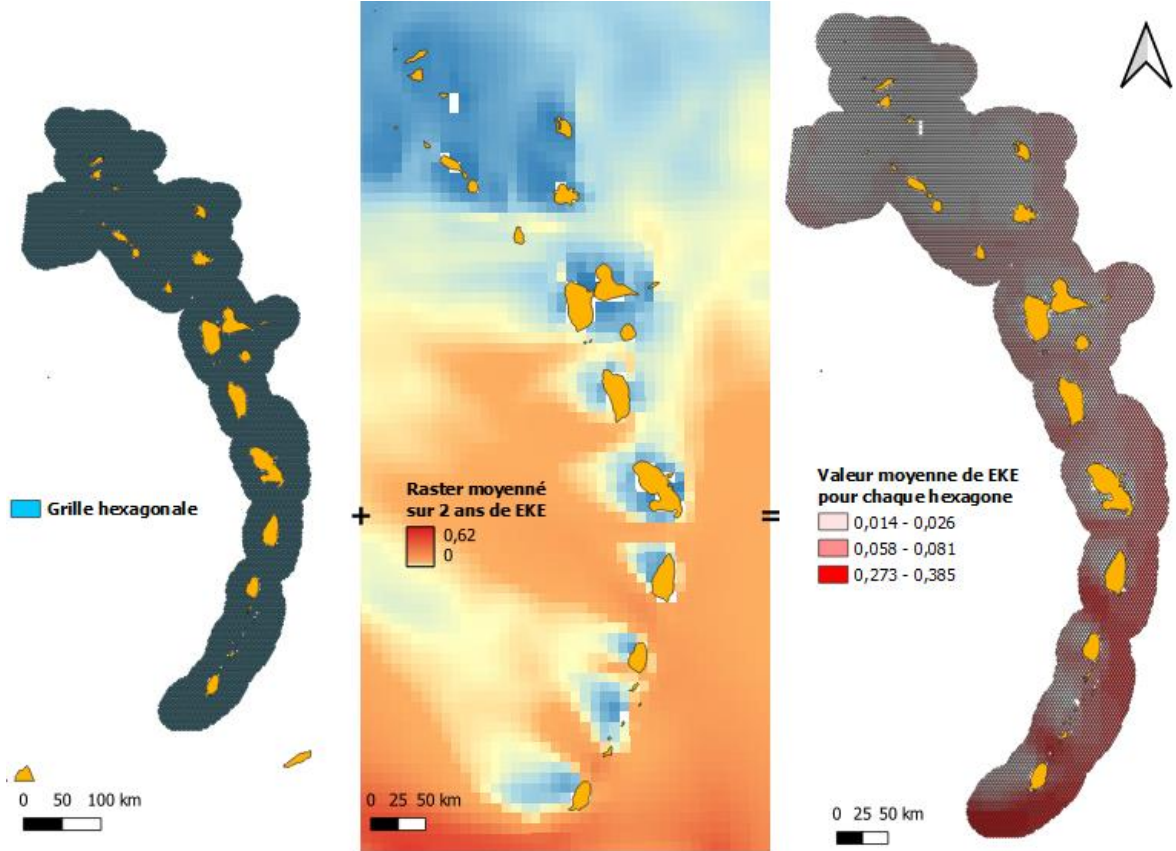
Annexe 2.B :

Variable environnementale	2021-2022
SST Moy	Oui
SST SD	Non
ChIA Moy	Oui
ChIA SD	Non
UO Moy	Oui
UO SD	Non
VO Moy	Oui
VO SD	Non
CV Moy	Oui
CV SD	Oui
EKE Moy	Oui
EKE SD	Oui
Bottom Temp Moy	Oui
Bottom Temp SD	Non
SSH Moy	Oui
SSH SD	Non
MLD Moy	Oui
MLD SD	Non

Annexe 3 : Explication du choix de chaque variable environnementale utilisée dans le modèle

Variables utilisées dans l'étude avec abréviations et unités	Résolution originale, fréquence	Effets sur les écosystèmes pélagiques d'intérêt potentiel pour les grands plongeurs
Physiographique		
Profondeur (m)	15 arc seconde	Les cachalots se nourrissent de calmars qui vivent dans les eaux profondes
Pente (°)	15 arc seconde	Associée aux courants, les fortes pentes induisent l'agrégation des proies ou une production primaire décuplée
Distance aux canyons (km)	15 arc seconde	Les grands plongeurs sont souvent associés aux canyons et monts sous-marins
Distance à la côte (km)	15 arc seconde	Permet d'identifier les préférences d'habitat ; les zones topographiques favorables à leur distribution
Distances à l'isobathe 200, 1000, 2000 (km)	15 arc seconde	Peut avoir une influence sur la distribution verticale des proies dans la colonne d'eau
Océanographique		
Température moyenne des eaux de surface-SST (°C)	0.05°x0.05°, journalière	Sa variabilité dans le temps peut être associée avec une productivité primaire augmentée
Température du fond (°C)	0.083° × 0.083°, journalière	Sa variation peut influencer directement la distribution des proies vivants dans les profondeurs comme les calmars
Hauteur moyenne des eaux de surface par rapport au géoïde-SSH (m)	0.083 x 0.083, journalière	Une haute SSH indique une activité à méso-échelle importante pouvant impliquer une agglomération des proies
Moyenne et variance de l'énergie cinétique des tourbillons-EKE (m ² .s ⁻²)	0.083 x 0.083, journalière	Un fort EKE indique la formation de tourbillons et donc la re-suspension de particules sédimentaires pouvant influencer une accréation de proies
Concentration en chlorophylle A moyenne-CHL (mg.m ⁻³)	0.25° x 0.25°, journalière	Utilisée comme un proxy pour rendre compte de la disponibilité en proies
Epaisseur de la couche de mélange moyenne-MLD (m)	0.083° × 0.083°, journalière	Une faible MLD favorise en général la productivité primaire et donc de manière indirecte les proies
Vitesse de courant moyenne et variance- CV (m/s)	0.083 x 0.083, journalière	Les courants océaniques peuvent influencer la dispersion et le déplacement des proies des cachalots
Moyenne de la composante Est de la vitesse- Uo (m/s)	0.083 x 0.083, journalière	Peut influencer la dérive des proies et/ou déterminer les schémas de migration saisonnière des cachalots le long des côtes
Moyenne de la composante Nord de la vitesse- Vo (m/s)	0.083 x 0.083, journalière	Peut affecter le mouvement vertical des cachalots et influencer la circulation océanique le long des isobathes et des fronts impactant ainsi les corridors de migration des cachalots

Annexe 4 : Implémentation d'une variable environnementale sur la grille hexagonale



Annexe 5 : Tableau récapitulatif de tous les individus capturés par la CCS, de la date de la dernière recapture par la CCS et par DSWP (pour les individus de DSWP recapturés par la CCS), ainsi que le nombre de captures total pour chaque individu (dont entre parenthèses le nombre de recapturés par la CCS).

Nom	Dernière date de capture DSWP, Vachon	Dernière date de capture CCS	Nombre de captures total (dont CCS)	Nom	Dernière date de capture Shane	Dernière date de capture CCS	Nombre de captures total (dont CCS)
5126	09/03/2020	29/08/2022	14(6)	6331	23/01/2020	22/06/2022	2(1)
5127	26/05/2017	27/06/2021	3(2)	6332	23/01/2020	22/06/2022	2(1)
5144	10/03/2020	16/07/2022	5(2)	6335	23/01/2020	22/06/2022	3(2)
5147	05/06/2012	07/07/2021	8(2)	6348	17/01/2020	26/05/2021	2(1)
5151	20/05/2018	24/04/2021	35(1)	6377	06/03/2020	30/03/2022	6(1)
5162	01/05/2017	23/07/2022	12(1)	6408	12/03/2020	21/07/2021	4(1)
5163	08/05/2017	07/07/2021	26(4)	6409	01/02/2020	21/07/2021	3(1)
5560	20/05/2018	24/04/2022	93(3)	6411	22/01/2020	21/07/2021	2(1)
5562	20/05/2018	16/08/2022	46(2)	6413	22/01/2020	21/07/2021	2(1)
5575	05/03/2011	23/09/2022	4(1)	6422	06/03/2020	30/03/2022	5(1)
5585	09/03/2020	22/04/2022	25(4)	6426	12/03/2020	19/06/2022	3(1)
5586	23/05/2018	23/07/2022	44(1)	6465	12/03/2020	19/06/2022	3(2)
5588	01/05/2017	17/08/2022	11(2)	6466	12/03/2020	21/07/2021	2(1)
5722	20/05/2018	29/08/2022	93(2)	6473	13/03/2020	22/06/2022	2(1)
5740	10/03/2020	27/06/2021	12(1)	Pm_0006		26/05/2021	1
5946	08/05/2017	22/04/2022	20(4)	Pm_0008		29/08/2022	3
6016	21/01/2020	24/09/2022	6(1)	Pm_0014		27/06/2021	1
6017	13/02/2020	24/09/2022	7(1)	Pm_0015		27/06/2021	1
6022	21/05/2017	28/03/2022	3(1)	Pm_0016		17/08/2022	2
6058	20/05/2018	29/08/2022	52(2)	Pm_0017		27/06/2021	1
6088	23/05/2018	23/07/2022	43(1)	Pm_0018		27/06/2021	1
6093	09/03/2020	22/04/2022	24(3)	Pm_0019		23/09/2022	1
6196	29/04/2017	07/07/2021	14(1)	Pm_0022		27/06/2021	1
6197	10/03/2020	07/07/2021	4(1)	Pm_0028		19/06/2022	1
6205	10/03/2020	07/07/2021	3(1)	Pm_0029		20/06/2022	2
6219	19/05/2018	29/08/2022	13(2)	Pm_0031		20/06/2022	1
6230	21/05/2017	28/03/2022	3(2)	Pm_0044		07/07/2021	1
6237	30/03/2018	14/09/2021	3(1)	Pm_0056		27/03/2022	1
6242	09/03/2020	29/09/2022	8(5)	Pm_0057		27/03/2022	1
6247	30/03/2018	14/09/2021	2(1)	Pm_0060		28/03/2022	2
6313	13/03/2020	20/06/2022	3(1)	Pm_0063		26/03/2022	1

Annexe 6 : Individus capturés et recapturés

Annexe 6.A : Nombre d'individus capturés autour de chaque île en 2019 & 2020 (Vachon, 2022).

Localisation	Total individus
Antigua-et-Barbuda	4
Saint-Kitts-et-Nevis	13
Montserrat	8
Guadeloupe	20
Dominique	42
Martinique	57
Saint-Vincent-et-les-Grenadines	42
Sainte-Lucie	48

Annexe 6.B : Individus recapturés ayant parcourus une distance supérieure à 100km

La durée entre les deux événements de recapture responsables de la distance maximale a été calculée en prenant en compte l'écart entre les deux dates où la distance maximale a été observée. Se pose alors la question de savoir si la deuxième distance maximale ne serait pas proche de la première et donnerait une durée plus courte. Cela a été vérifié en regardant toutes les dates intermédiaires entre la 1^{ère} date (c'est-à-dire la date la plus ancienne à partir de laquelle les positions GPS ont été considérées pour calculer la distance maximale) et la date la plus récente de recapture par DSWP. Dans les cas où cette date était supérieure à la 1^{ère} date, le point a été localisé soit au niveau de la localisation 2, soit en un lieu intermédiaire entre la localisation 1 et la localisation 2. Cela indique que l'évaluation de la durée mise pour parcourir cette distance maximale a été évaluée de manière correcte.

Dans le tableau ci-dessous, le G entre parenthèses signifie « Guadeloupe ». SVLG est l'abréviation pour Saint-Vincent-et-les-Grenadines. Les individus en rouge sont ceux pour lesquels la 1^{ère} date n'est pas la dernière date de recapture et pour lesquels l'évaluation de la durée a donc été vérifiée. Les lignes colorées en vert correspondent aux individus ayant parcouru une distance supérieure à 200km.

Nom	Distance max (m)	1 ^{ère} Date	Localisation 1	2 ^{ème} Date	Localisation 2	Durée (années)	Date la plus récente de recapture par DSWP
5126	109.72	01/05/2017	Dominique	28/05/2021	Guadeloupe	4.1	09/03/2020 (G)
5163	113.83	17/03/2011	Dominique	28/05/2021	Guadeloupe	10.2	08/05/2017 (G)
5575	154.15	05/03/2011	Dominique	23/09/2022	Sainte-Lucie	11.5	05/03/2011
5585	107.01	13/03/2010	Dominique	03/06/2021	Guadeloupe	11.2	09/03/2020 (G)
5946	108.07	13/03/2010	Dominique	28/05/2021	Guadeloupe	11.2	08/05/2017 (G)
6016	172.85	02/06/2012	Dominique	24/09/2022	Sainte-Lucie	10.3	21/01/2020 (Martinique)
6017	172.86	02/06/2012	Dominique	24/09/2022	Sainte-Lucie	10.3	13/02/2020 (Martinique)
6022	217.19	21/05/2017	Dominique	28/03/2022	SVLG	4.8	21/05/2017
6093	108.91	14/03/2011	Dominique	28/05/2021	Guadeloupe	10.2	09/03/2020 (G)
6196	100.49	29/04/2017	Dominique	07/07/2021	Guadeloupe	4.2	29/04/2017
6230	218.79	21/05/2017	Dominique	28/03/2022	SVLG	4.8	21/05/2017

6242	114.15	01/05/2017	Dominique	26/06/2021	Guadeloupe	4.2	09/03/2020 (G)
6313	114.24	27/03/2022	SVLG	20/06/2022	Grenade	0.2	13/03/2020
6331	101.66	23/01/2020	SVLG	22/06/2022	Grenade	2.4	23/01/2020
6332	102.92	23/01/2020	SVLG	22/06/2022	Grenade	2.4	23/01/2020
6377	279.37	17/02/2020	Martinique	30/03/2022	Grenade	2.1	06/03/2020
6408	113.91	01/02/2020	Sainte-Lucie	21/07/2021	SVLG	1.4	12/03/2020
6409	116.75	01/02/2020	Sainte-Lucie	21/07/2021	SVLG	1.4	01/02/2020
6422	270.10	26/01/2020	Martinique	30/03/2022	Grenade	2.2	06/03/2020
6426	121.72	06/02/2020	Sainte-Lucie	19/06/2022	SVLG	2.4	12/03/2020

Annexe 7 : Présentation des différentes unités sociales rencontrées par la CCS, du nombre de fois où elles ont été vues, ainsi que du nombre d'individus dans cette unité

Unité sociale/groupement	Nombre d'observations (nombre de jours si différent)	Nombre d'individus vus par la CCS appartenant à ce groupe	Nombre d'individus totaux dans le groupe (et nombre potentiel)
A	1	1	5
ATW	1	2	9
F	4	3	3 (4)
G10	1	1	6
G12	1	2	10
G15	1	4	7
G17	2	2	4 (5)
G18	2	2	14
G7	2	3	10 (12)
K	1	2	6
L	2 (1)	2	5 (6)
N	2	2	7 (8)
T	8 (7)	4	5 (6)
U	4	3	3 (4)
V	9 (8)	4	9 (10)
X	1	2	13
Z	3	4	

Annexe 8 : Planches photos des unités sociales/groupements les plus vus par la CCS

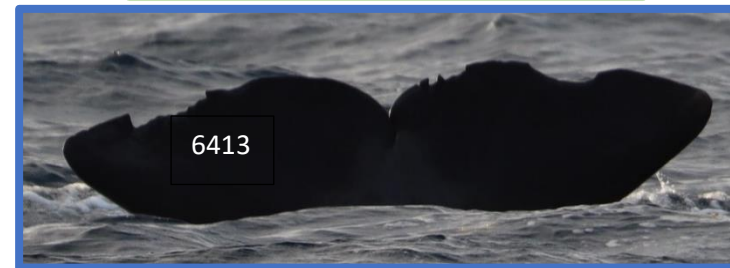
Dans les planches photos, les individus dont le cadre est de la même couleur au sein d'une planche ont été vus le même jour. Les individus dont le nom est en rouge n'ont pas été vus par la CCS mais par Vachon (2022) ou DSWP et sont les autres individus qui font partie de l'unité sociale ou du groupement en question. Les informations relatives à l'endroit où, ou la date à laquelle ils ont été vus sont également en rouge. Lorsqu'une seule localisation est présente et non colorée en rouge, cela signifie que les individus ont été observés aux mêmes endroits par la CCS que par Vachon (2022) ou DSWP.

Groupement 15

Localisation : SVLG ; Sainte-Lucie

Date : 21/07/2021 ; 2020

Clan : EC2

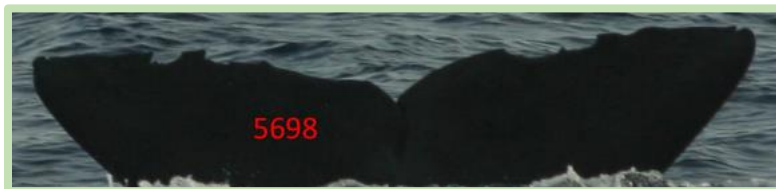


Unité T, Tooth's Team

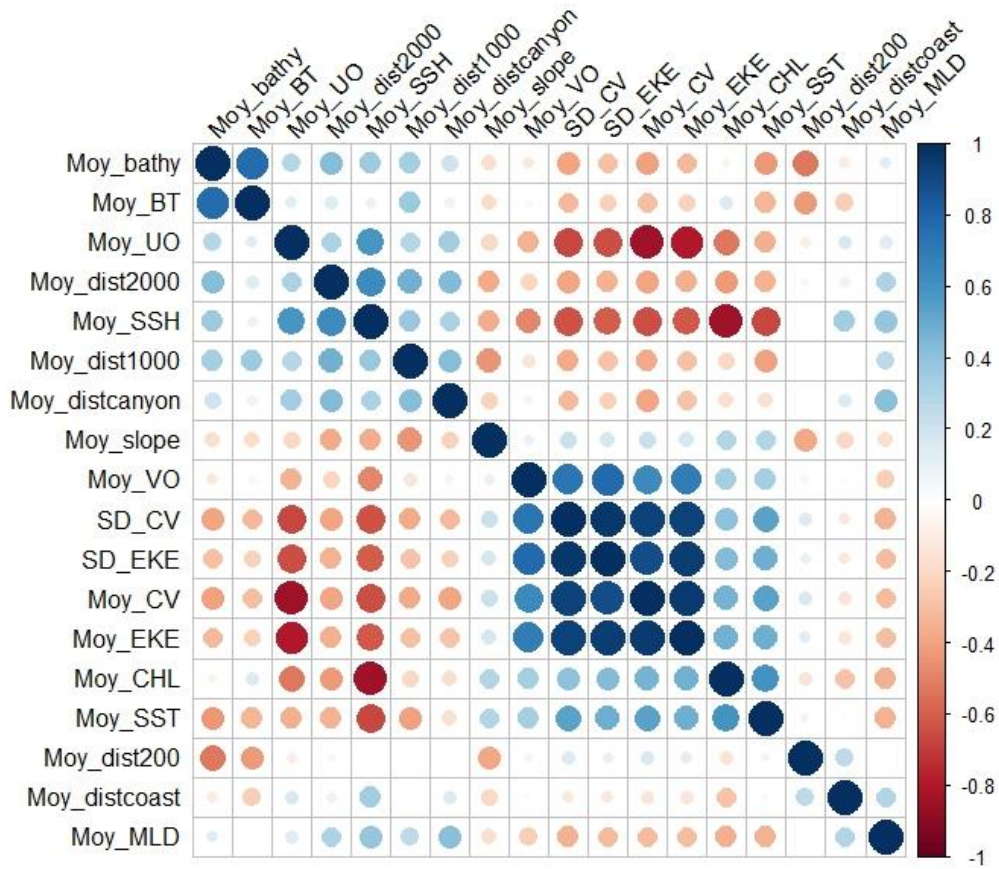
Localisation : Guadeloupe

Date : 28/05 ; 03/06 ; 12/06 ; 26/06 ; 27/06 ; 07/07/21 ; 22/04/22 ; 2020

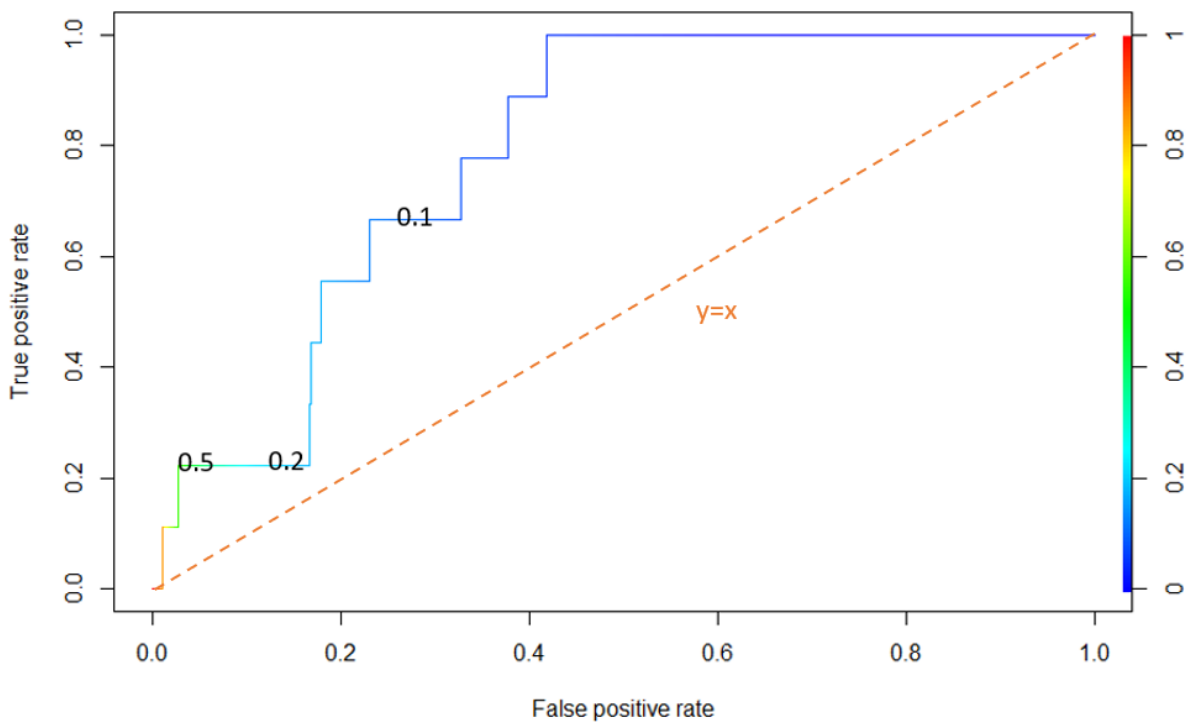
Clan : EC1



Annexe 9 : Matrice des corrélations des variables environnementales utilisées dans le modèle GAM



Annexe 10 : Courbe ROC représentant l'évolution de la sensibilité (taux de vrais positifs) en fonction de 1-spécificité (taux de faux positifs) lorsque le seuil t varie (au-dessus duquel la prédiction est positive et en-dessous duquel la prédiction est négative).



Annexe 11 : Arbre de régression multivarié en fonction des unités sociales et des 4 prédicteurs du meilleur modèle

