

# RECOMMANDATIONS OPÉRATIONNELLES POUR LE SUIVI DE LA MÉGAFaUNE MARINE À L'ÉCHELLE D'UN PARC ÉOLIEN EN MER

---



FRANCE  
ÉNERGIES  
MARINES

Editions





# RECOMMANDATIONS OPÉRATIONNELLES POUR LE SUIVI DE LA MÉGAFAUNE MARINE À L'ÉCHELLE D'UN PARC ÉOLIEN EN MER

## AUTEURS :

Maud Quéroùé, France Energies Marines, Centre d'Écologie Fonctionnelle et Évolutive  
Sébastien Lefèvre, Université de Bretagne Sud  
Minh-Tan Pham, Université de Bretagne Sud  
Mathieu Authier, Université de La Rochelle  
Aurélien Besnard, Centre d'Écologie Fonctionnelle et Évolutive  
Karine Heerah, France Energies Marines

Tous droits réservés.

Les textes de ce rapport de synthèse sont la propriété des partenaires du projet OWFSOMM (France Energies Marines, Université de Bretagne Sud, Centre National de la Recherche Scientifique, Observatoire PELAGIS, WIPSEA, EDF Renouvelables, Ocean Winds, Shell, RWE, Ifremer, Eoliennes en mer Dieppe Le Tréport, CE-Sigma, ENSTA Bretagne). Ils ne peuvent être reproduits ou utilisés sans citer la source et sans autorisation préalable. Les photos, les schémas et les tableaux (sauf indication contraire) sont protégés par le droit d'auteur. Ils restent la propriété des partenaires du projet OWFSOMM ou de leurs auteurs (si reproduits) et ne peuvent être reproduits sous quelque forme ou par quelque moyen que ce soit, sans l'autorisation écrite préalable des partenaires du projet OWFSOMM ou de leurs auteurs.

Citer le document comme ci-dessous :

Quéroùé M., Lefèvre S., Pham M.-T., Authier M., Besnard A. et Heerah K.

Recommandations opérationnelles pour le suivi de la mégafaune marine à l'échelle d'un parc éolien en mer Plouzané : France Energies Marines Editions, 2025, 28 pages.

Edition : Février 2025

Photo de couverture : Groupe de dauphins immortalisé par un système de prise de vue automatisé

© Pelagis / Hytech-imaging - STORMM

Dépôt légal à parution

# Table des matières

	Page
<b>Introduction</b>	7
<b>1. Intercalibration des méthodes de suivi aérien de la mégafaune marine</b>	8
1.1. Contexte .....	8
1.2. Objectifs .....	8
1.3. Principales réalisations .....	9
<b>Recommandations</b> .....	<b>12</b>
<b>2. Développement d'outils d'IA pour le suivi de la mégafaune marine</b>	16
2.1. Contexte .....	16
2.2. Objectifs .....	17
2.3. Principales réalisations .....	17
<b>Recommandations</b> .....	<b>20</b>
<b>Conclusion</b>	24
<b>Liste des figures</b>	5
<b>Liste des tableaux</b>	5
<b>Acronymes</b>	5
<b>Remerciements</b>	6
<b>Bibliographie</b>	25

## Liste des figures

Page

<b>Fig. 1</b> • Représentation schématique des méthodes d'acquisition de données déployées pendant le suivi des parcs éoliens en mer .....	8
<b>Fig. 2</b> • Carte des sites étudiés pendant le projet OWFSOMM .....	9
<b>Fig. 3</b> • Représentation schématique d'une acquisition de données vidéo (A) et acoustique sous-marine (B) pour le suivi des oiseaux et mammifères marins autour d'une plateforme instrumentée. ....	16
<b>Fig. 4</b> • Déploiement d'un hydrophone large bande à proximité du mât de mesures situé au large de Fécamp. ....	17
<b>Fig. 5</b> • Résultats de détection et classification de mammifères marins par IA sur les spectrogrammes des données acoustiques. ....	18
<b>Fig. 6</b> • Résultats de détection d'oiseaux par IA sur des données vidéo HD 360°. Comparaison graphique des approches d'apprentissage supervisé et de détection d'anomalies faiblement supervisée. ....	19

## Liste des tableaux

<b>Tab. 1</b> • Descriptif des modes d'acquisition de données déployés lors des différentes campagnes réalisées. ....	10
---	----

## Acronymes

<b>IA</b>	Intelligence artificielle
<b>OWFSOMM</b>	Offshore Wind Farm Surveys Of Marine Megafauna

## Remerciements

Ce rapport a été rédigé dans le cadre du projet collaboratif OWFSOMM (Offshore Wind Farm Surveys Of Marine Megafauna), piloté par France Energies Marine, le Centre d'Ecologie Fonctionnelle et Evolutive et l'IRISA. Le projet OWFSOMM a bénéficié d'un financement de France Energies Marines, de ses membres et partenaires, de l'Office Français de la Biodiversité, de la Direction Générale de l'Energie et du Climat ainsi que d'une aide de l'Etat français gérée par l'Agence Nationale de la Recherche dans le cadre du Plan d'Investissement France 2030 (ANR-10-IEED-0006-34). Nous tenons à remercier les 13 partenaires du projet (France Energies Marines, Université de Bretagne Sud, Centre National de la Recherche Scientifique, Observatoire PELAGIS, WIPSEA, EDF Renouvelables, Ocean Winds, Shell, RWE, Ifremer, Eoliennes en mer Dieppe Le Tréport, CE-Sigma, ENSTA Bretagne) pour leurs contributions tout au long du projet OWFSOMM. Nous remercions toutes les personnes qui ont participé aux relevés aériens, collecté les données et analysé les images, notamment Hytech-imaging, Biotope, Setec In-vivo et Bretagne Vivante. Nous remercions également EDF Renouvelables, Ocean Winds et le gouvernement français (Direction Générale de l'Energie et du Climat) qui nous ont fourni des données provenant de relevés aériens effectués dans le cadre d'études d'évaluation environnementale réglementaires. Nous remercions également EDF Renouvelables et la société Eoliennes en Mer des Hautes Falaises (EOHF) pour avoir fourni le jeu de données acoustiques passives sous-marines utilisé dans cette étude, ainsi que l'équipe OSmoSE (ENSTA Bretagne, Lab-STICC, UMR CNRS 6285) pour ses conseils en bioacoustique.

## Introduction

Le nombre croissant de parcs éoliens en mer soulève des questions environnementales et sociétales majeures quant à leur impact sur la faune et la flore. Les espèces de la mégafaune marine, telles que les poissons, les tortues, les oiseaux et les mammifères marins, sont au cœur des préoccupations compte tenu des risques de mortalités directes (e.g. collision avec les turbines) et indirectes (e.g. perte d'habitats fonctionnels, perturbations sonores) induits par ces nouvelles structures (Bailey *et al.*, 2014; Croll *et al.*, 2022). Ces espèces, dont beaucoup sont protégées, en danger ou menacées, représentent le haut des chaînes trophiques et jouent un rôle clé dans le fonctionnement des écosystèmes. Leur écologie et leur distribution intègrent les variations spatio-temporelles de leurs sources de nourriture. Ces espèces peuvent donc être utilisées comme indicateurs pour surveiller la santé des écosystèmes marins (Hazen *et al.* 2019, Jelichich *et al.* 2022). En conséquence, pour chaque parc éolien, le suivi de la mégafaune marine est crucial à chaque étape, de l'autorisation à la construction et à l'exploitation. Pourtant, l'observation des espèces de la mégafaune marine reste difficile car elles passent la plupart de leur temps sous l'eau et au large.

Le suivi de la mégafaune marine peut être effectué par des observateurs embarqués sur un bateau, ou par des méthodes de suivi aérien. En France, les évaluations de base et les recensements préalables à l'installation sont principalement réalisés par observation visuelle aérienne à une altitude inférieure ou égale à 200 mètres afin de pouvoir détecter et identifier de manière satisfaisante les petites espèces telles que les oiseaux marins ou les marsouins. Pourtant, lors des phases opérationnelles, les techniques de suivi aérien nécessiteront une altitude minimale de 300 mètres afin de respecter les règles de sécurité. A cette altitude plus élevée, l'acquisition des données se fera par des moyens digitaux. Ce changement dans l'acquisition des données pendant le cycle de vie des parcs éoliens en mer peut avoir des conséquences sur le processus d'observation (c'est-à-dire la détection de la mégafaune marine) et potentiellement sur les estimations d'abondance de la mégafaune. Il est donc essentiel de pouvoir intercaler, au stade

de l'analyse, ces données collectées selon des méthodes différentes (visuelles vs digitales).

En parallèle, des plateformes multi-capteurs peuvent être installées pour surveiller les sites éoliens pendant les différentes phases de leur cycle de vie, de l'évaluation de base jusqu'au démantèlement. Ces plateformes sont d'une grande importance car elles permettent d'acquérir des données affinées sur l'écologie de la mégafaune à l'échelle de la ferme éolienne et d'envisager des suivis continus et à long terme. Par ailleurs, l'utilisation de plusieurs capteurs en simultané permet de pallier les limites de chacune des technologies quant au niveau d'information qu'elles peuvent apporter sur la surveillance de la mégafaune marine. Ces données peuvent ensuite alimenter différents types de modèles de prédiction, notamment pour évaluer les risques de collision pour les populations d'oiseaux marins et l'exposition au bruit pour les mammifères marins et les poissons. Dans ce contexte, OWFSOMM avait pour objectifs :

- de développer des méthodes et outils d'intercalibration pour le suivi de la mégafaune marine à partir de survol aérien afin d'assurer la comparabilité entre les observations visuelles et digitales,
- de développer des algorithmes de détection et d'identification automatique de la mégafaune marine à partir de sources de données multimodales,
- de fournir des recommandations méthodologiques afin d'optimiser le suivi de la mégafaune marine à l'échelle d'un parc éolien.

L'objectif de ce rapport est de proposer des recommandations aussi bien opérationnelles que techniques afin d'optimiser la collecte de données et leur traitement en vue d'assurer un suivi pertinent de la mégafaune marine. Ces recommandations sont fournies au regard des résultats obtenus dans le projet OWFSOMM et sont détaillées suivant deux parties. La première concerne l'intercalibration des méthodes pour le suivi de la mégafaune marine à partir de survols aériens. La deuxième traite du développement d'outils d'IA pour le suivi de la mégafaune marine.

## 1 - Intercalibration des méthodes de suivi aérien de la mégafaune marine

### 1.1 Contexte

Au cours de ces dernières années, les solutions digitales de surveillance de la mégafaune marine par survols aériens ont connu des développements techniques importants. Pouvant être déployées à haute altitude, elles permettent de respecter les conditions de survol des parcs éoliens (Figure 1) et seront prochainement utilisées dans le cadre du suivi environnemental des différents projets français d'éolien en mer.

Cependant, les données environnementales existantes sont issues de survols aériens embarquant des observateurs naviguant à basse altitude. Il est donc nécessaire de garantir la commensurabilité des résultats obtenus par différentes méthodes avec, d'une part, les méthodes traditionnelles de survols aériens visuels, et d'autre part, les techniques digitales plus récentes.

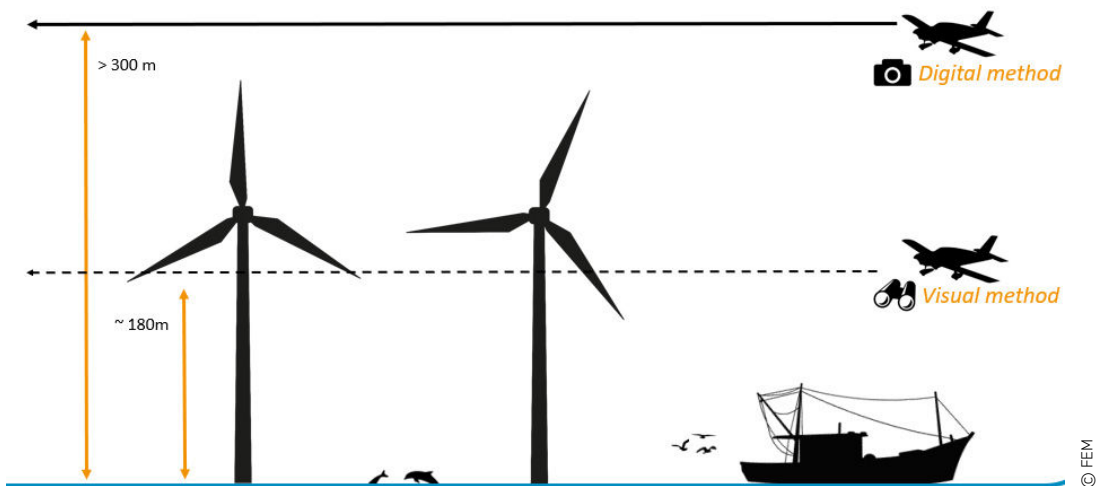


Fig. 1 : Représentation schématique des méthodes d'acquisition de données déployées pendant le suivi des parcs éoliens en mer.

### 1.2 Objectifs du projet

- Fournir une feuille de route opérationnelle pour une intercalibration fiable des suivis aériens de la mégafaune marine dans les parcs éoliens en mer en utilisant différentes technologies.
- Elaborer une méthodologie d'intercalibration entre les suivis aériens embarquant des observateurs humains et ceux réalisés avec des techniques digitales.
- Rédiger des recommandations pour la surveillance de la mégafaune dans le cadre des projets d'éolien en mer afin d'assurer l'interopérabilité de l'ensemble des données.



## 1.3 Principales réalisations

### Un protocole de format de données pour les méthodes digitales

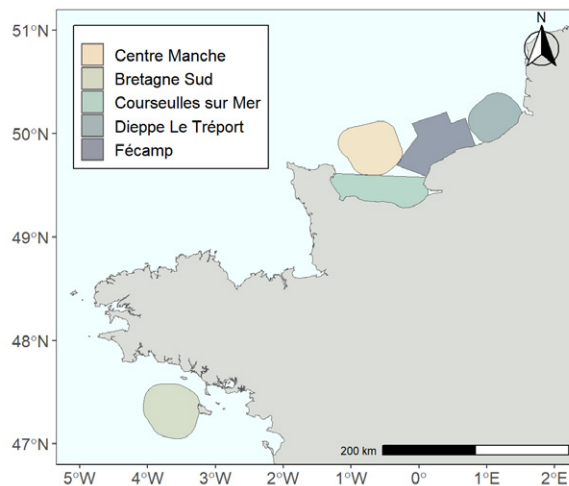
Un format de données standardisé adapté aux méthodes d'acquisition de données digitales a été mis au point en concertation avec différents partenaires impliqués dans les suivis digitaux. Ce protocole permet d'inclure les mêmes infor-

mations que celles collectées durant les survols aériens visuels tout en tenant compte des spécificités associées aux survols digitaux et au traitement des données.

### Des campagnes d'intercalibration des méthodes visuelles et digitales

14 campagnes d'intercalibration ont été réalisées sur 5 sites différents (Figure 2) et à différentes périodes de l'année (Tableau 1). Durant ces campagnes, deux avions circulant à un maximum de 15 minutes d'intervalle, effectuaient les mêmes transects sur la zone. Un avion naviguant

à basse altitude (~180 m) embarquait des observateurs ainsi qu'un système digital de manière à effectuer un suivi digital et un suivi visuel simultanément. Un deuxième avion naviguant à haute altitude, à plus de 300 m, embarquait un système d'acquisition de données digitales.



© Projet OWFSOMM

Fig. 2 : Carte des sites étudiés pendant le projet OWFSOMM

Campagne			Basse altitude Observation visuelle	Basse altitude Observation digitale		Haute altitude Observation digitale	
Site	Date	Etat de la mer (Moyenne, Beaufort)	Altitude	Altitude	Méthode digitale	Altitude	Méthode digitale
Fécamp	22/03/2021	1,4	~180 m	~180 m	A	-	-
Courseulles-sur-Mer	02/05/2021	1,0	~180 m	~180 m	A	~330 m	B
Dieppe Le Tréport (EMDT1)	27/02/2022	3,3	~180 m	~180 m	B	~550 m	C
Dieppe Le Tréport (EMDT2)	09/06/2022	1,9	~180 m	~180 m	B	~550 m	C
Dieppe Le Tréport (EMDT3)	04/09/2022	3,1	~180 m	~180 m	B	~550 m	C
Dieppe Le Tréport (EMDT4)	14/02/2023	1,9	~180 m	~180 m	B	~550 m	C
Centre Manche (A041)	25/03/2022	3,0	~180 m	~180 m	B	~550 m	C
Centre Manche (A042)	13/06/2022	1,0	~180 m	~180 m	B	~550 m	C
Centre Manche (A043)	12/09/2022	2,8	~180 m	~180 m	B	~550 m	C
Centre Manche (A044)	13/12/2022	3,1	~180 m	~180 m	B	~550 m	C
Sud Bretagne (A051)	23/07/2022	2,8	~180 m	~180 m	B	~550 m	B
Sud Bretagne (A052)	22/09/2022	1,1	~180 m	~180 m	B	~550 m	B
Sud Bretagne (A053)	21/01/2023	3,9	~180 m	~180 m	B	~550 m	B
Sud Bretagne (A054)	27/03/2023	2,3	~180 m	~180 m	B	~550 m	B

Tab. 1 : Descriptif des modes d'acquisition de données déployés lors des différentes campagnes réalisées. Méthodes digitales utilisées : A: *WIPSEA*, B: *Hytech-imaging*, C: *HiDef*

## Une analyse de comparabilité des différentes méthodes visuelles et digitales

Pour chaque jeu de données obtenu par les différents survols lors d'une campagne d'intercalibration, une estimation d'abondance et de distribution des individus a été réalisée. Etant donné que le nombre réel d'individus présents dans la zone d'étude était inconnu, il n'était pas

question d'évaluer la capacité des méthodes à estimer correctement l'abondance sur la zone, mais uniquement de comparer les résultats obtenus par l'une et l'autre des méthodes. Un facteur d'intercalibration est ainsi calculé comme étant le ratio entre l'abondance estimée par une

méthode et l'abondance estimée par une seconde méthode. De cette manière, les facteurs d'intercalibration ont été calculés entre les méthodes visuelles et les méthodes digitales, mais aussi entre les méthodes digitales pour l'ensemble des campagnes, pour les différentes altitudes de vol et pour les différents groupes d'espèces (fous, alcidés, laridés, plongeurs, cétacés, requins etc.).

Les résultats ont montré qu'en moyenne les méthodes digitales estiment des abondances supérieures aux abondances estimées par les méthodes visuelles. Les facteurs d'intercalibration sont néanmoins très variables selon l'état de la mer et les espèces étudiées. Les méthodes produisent des estimations d'abondance assez proches lorsque les analyses concernent des espèces facilement détectables telles que les fous de Bassan (*Morus bassanus*). Au contraire, pour les espèces discrètes comme les alcidés, les plongeurs ou les grèbes, les méthodes visuelles fournissent en moyenne des estimations d'abondances plus faibles que les méthodes digitales. Par ailleurs, lorsque la mer est agitée la diffé-

rence d'estimation d'abondance entre les deux méthodes est plus forte qu'en condition de mer calme, avec une sous-estimation des méthodes visuelles. Une grosse partie de la variabilité (>60%) des facteurs d'intercalibration estimés n'est cependant pas expliquée par ces différents éléments. Ces résultats montrent la difficulté, voire l'impossibilité, de la définition générale et automatiquement transférable de facteurs d'intercalibration pour comparer des estimations d'abondance obtenues par les différentes méthodes visuelles et digitales sans tenir compte du contexte local dans lequel les données ont été récoltées (site, état de la mer, espèces cibles, etc.).

Enfin, les résultats montrent qu'en moyenne, les estimations d'abondance des méthodes digitales à haute et basse altitude étaient cohérentes avec des facteurs d'intercalibration proches de 1. Néanmoins, il existe une variabilité autour de ce facteur d'intercalibration moyen qui n'a pu être expliquée qu'à 18% par la méthode digitale employée et l'état de la mer.



© PIXSEA

### Une analyse des facteurs influençant la détection des algorithmes de détection d'anomalies

Une banque de 1054 images issues des campagnes comportant des cibles d'intérêt, à identifier ou non, a été constituée en vue de caractériser minutieusement la qualité des images ainsi que les conditions environnementales dans lesquelles elles avaient été prises.

Cette caractérisation a été réalisée par un opérateur humain. Les images ont ainsi été examinées afin de rechercher et d'identifier les « objets » présents, tels que les oiseaux ou mammifères marins, en se basant sur des guides d'identification de référence. Une comparaison des détec-

tions a ensuite été réalisée entre les détections de l'opérateur et celles des algorithmes utilisés pour détecter les objets d'intérêt sur les images. Cette comparaison permet de repérer d'éventuels faux-négatifs, c'est-à-dire des objets présents sur une image mais non détectés par l'algorithme. Les résultats de ce travail montrent

que la fréquence de faux-négatifs est fortement corrélée à la netteté de l'image et aux reflets du soleil et qu'elle diffère selon le groupe d'espèce étudié. Cette analyse souligne l'importance de disposer d'une mesure quantitative de la qualité des images afin de tenir compte de la détection imparfaite des objets d'intérêt.

### Un package R 'sismow' permettant la simulation d'efforts d'échantillonnage et la simulation d'individus dans une zone d'étude

Un package encapsulant une suite de fonctions dédiées pour réaliser des études de simulations numériques en vue de réaliser une étude d'intercalibration est disponible à l'adresse suivante : <https://github.com/maudqueroue/sismow>. Ce package permet de simuler des distributions 'réalistes' de la mégafaune marine dans une zone donnée, puis de simuler un design de campagne d'échantillonnage et enfin de simuler

une méthode d'acquisition de données. Le jeu de données ainsi simulé peut reproduire les principales caractéristiques d'un vrai jeu de données. Les simulations vont permettre ici de réaliser des analyses de puissance statistique afin de dimensionner (en termes d'effort) de manière adéquate et prospective des campagnes de survols aériens de la mégafaune marine.

## RECOMMANDATIONS

Les recommandations mentionnées ci-dessous concernent le suivi environnemental des parcs éoliens en mer. Elles sont uniquement valables dans le contexte de suivi de la mégafaune marine dans le but d'évaluer l'impact des parcs éoliens en mer sur l'ensemble de leur cycle de vie. Ces recommandations ne concernent donc pas le suivi à grande échelle d'espèces mobiles.

## PLANIFICATION ET RÉCOLTE

### 1 Favoriser l'utilisation des méthodes digitales pour les suivis écologiques de la mégafaune marine menés dans le cadre du développement de l'éolien en mer

Les résultats du projet ont montré que l'utilisation de facteurs d'intercalibration permettant de passer d'une méthode visuelle d'acquisition de données à une méthode digitale d'acquisition de données n'était pas aisée. Une utilisation correcte de facteurs d'intercalibration est soumise à de nombreux facteurs contextuels qui en limitent grandement la transférabilité et la validité. Considérant que les méthodes digitales seront obligatoirement utilisées pour des questions

de sécurité lors des phases de construction et d'utilisation des parcs éoliens, nous conseillons d'utiliser les méthodes digitales dès les premiers suivis environnementaux. L'objectif est d'utiliser le même type de méthode tout au long du cycle de vie des parcs de manière à pouvoir analyser des tendances temporelles d'abondance de la mégafaune marine sans qu'elles soient biaisées par des changements de méthodes d'acquisition des données.

### 2 Planifier les survols en favorisant autant que possible des conditions de mer calmes

Pour limiter autant que possible les sources de variabilité entre différents survols digitaux au cours du temps, des conditions similaires de

survol doivent être privilégiées. Il est notamment important de privilégier des conditions de mer les plus calmes possible. Une mer agitée peut

rendre indisponible à la détection des individus pourtant présents. C'est notamment le cas des espèces passant principalement leur temps sous l'eau. Le risque associé à la variabilité de l'état de la mer est de sous-estimer le nombre d'individus lorsque les conditions en mer sont détériorées. De ce fait, une non prise en compte de ces biais de perception et de disponibilité des espèces peut aboutir à des conclusions écologiques erronées. Par ailleurs, la détection des cibles est complexifiée lorsque la mer est agitée. Il y a notamment

le risque d'obtenir un grand nombre de faux positifs liés à la détection de vaguelettes par les algorithmes de détection d'anomalies, ce qui entraîne en conséquence un long travail d'identification des cibles. Inversement, les algorithmes entraînés à ne pas détecter de vaguelettes peuvent manquer des individus qui se confondraient dans les vagues, induisant alors des faux négatifs et une sous-estimation du nombre d'individus dans la zone d'étude.

### 3 Limiter la présence de reflets du soleil et optimiser la netteté des images

Les résultats de l'analyse des facteurs influençant la détection des méthodes digitales ont mis en évidence l'impact des reflets du soleil et de la netteté des images sur la quantité de faux négatifs. La présence de reflets du soleil sur les images augmente en effet le risque que les algorithmes manquent un individu pourtant bien présent sur l'image. De la même manière, il a été

observé que la capacité de détection des algorithmes était diminuée lorsque les images étaient floues. Lors des transects aériens, il est donc important de limiter au maximum la présence de reflets de soleil sur les images et d'assurer une bonne mise au point lors de la prise d'image pour avoir les cibles les plus nettes possible.

## ACQUISITION & TRAITEMENT DES DONNÉES

### 4 Se référer au protocole du format de données OWFSOMM pour les données collectées lors des survols digitaux

Ce projet a permis de mettre en place un format de données standardisé adapté aux méthodes digitales. Nous recommandons l'utilisation de ce format de données car il permet de récupérer

l'ensemble des informations nécessaires pour mettre en place les analyses d'abondance et de distribution à partir des méthodes d'acquisition de données digitales.

### 5 Solliciter des naturalistes qualifiés pour l'identification des cibles détectées par les méthodes digitales

Tout comme les personnes embarquant dans les avions pour les survols visuels, les personnes analysant les images acquises durant les survols digitaux doivent être compétentes pour l'identification des espèces. Il est essentiel que les personnes en charge de ce travail aient une

forte compétence naturaliste en ce qui concerne l'identification de la mégafaune marine et particulièrement pour les oiseaux marins qui sont les espèces les plus représentées dans les zones de parcs éoliens en mer.

### 6 Informer sur les performances des algorithmes de détection d'anomalies utilisés

Des erreurs de détection des méthodes digitales ont été observées en lien avec (1) les conditions environnementales comme la présence de reflets du soleil sur les images, (2) la qualité des

images comme la netteté et (3) selon les espèces observées. Pour être en capacité de maîtriser ces sources de biais et de pouvoir corriger les abondances estimées, il est nécessaire de connaître

les performances des algorithmes de détection d'anomalies utilisés. Ces performances peuvent être transmises sous forme d'une matrice de confusion par espèces, par conditions environnementales et selon la qualité des images, tout en informant sur le jeu de données sur lequel l'al-

gorithme a été testé. Par ailleurs, de telles informations permettront aussi de pouvoir suivre les évolutions des capacités des algorithmes utilisés au cours du temps et d'en tenir compte dans l'interprétation des résultats.

## 7 Instaurer des analyses de qualité pour les phases de détection et d'identification des cibles

Une manière de garantir la qualité des jeux de données fournis par les méthodes digitales est de mettre en place des procédures de vérification. Ces analyses ont pour objectif de valider la qualité du processus de traitement des données et concernent prioritairement les deux phases suivantes : la détection des cibles et l'identification des cibles. Concernant la détection des cibles, il est recommandé qu'une partie des images soit re-analysée par un opérateur humain et que les détections opérées par vision humaine soient comparées aux détections de l'algorithme.

Il est nécessaire de s'assurer que les capacités de détection de l'algorithme correspondent aux performances mentionnées (cf. recommandation 6). Concernant l'identification des cibles, il est recommandé qu'une personne n'ayant pas participé à l'identification des cibles détectées dans le jeu de données analyse indépendamment un sous échantillon des cibles détectées afin de confirmer ou d'infirmer l'identification au plus petit échelon taxonomique possible. L'objectif est d'assurer la justesse des identifications des espèces.

## UTILISATION DES RÉSULTATS POUR LES SUIVIS ÉCOLOGIQUES

### 8 Ne pas estimer de tendances temporelles en utilisant des résultats provenant à la fois de méthodes visuelles et de méthodes digitales

Les campagnes d'intercalibration ont montré la grande variabilité des facteurs d'intercalibration entre méthodes visuelles et digitales et leurs étroites dépendances à des facteurs contextuels élucidés partiellement. Utiliser des facteurs d'intercalibration sans tenir compte du contexte local dans lequel les données ont été récoltées (site, état de la mer, espèces cibles, méthodes utilisées, observateurs, etc.) ne permettrait pas

d'évaluer de manière fiable des fluctuations d'abondance au cours du temps. Le risque étant que, sur le cycle de vie d'un parc éolien en mer, les variations observées ne soient que la résultante d'un changement de méthode de suivi et non un réel changement d'abondance. En conséquence, l'utilisation de résultats provenant à la fois des méthodes visuelles et digitales pour établir des tendances temporelles est à éviter.

### 9 Ne pas interpréter de légères variations d'abondance au cours du temps

Les résultats des campagnes d'intercalibration ont montré que les estimations d'abondance des méthodes digitales à haute et basse altitude étaient, en moyenne, cohérentes et présentaient des facteurs d'intercalibration proches de 1. Cependant, il existe tout de même une variabilité autour du facteur d'intercalibration moyen qui n'est que très peu expliquée par différentes

variables comme la méthode digitale utilisée ou l'état de la mer. Ainsi, lors de l'estimation de tendances temporelles utilisant des résultats provenant de méthodes digitales, il faut veiller à ne pas interpréter de faibles changements d'abondance qui peuvent plutôt résulter de la variabilité entre les suivis plutôt qu'à une réalité écologique.

## 10 Prendre en compte l'amélioration des méthodes digitales

Les méthodes digitales sont encore en plein développement. Il est donc à anticiper que ces méthodes s'améliorent au cours du temps, et donc au cours du cycle de vie des parcs éoliens en mer. De même, il est raisonnable de s'attendre à ce que les algorithmes de détection d'anomalies deviennent plus performants ou que la qualité des images s'améliore. Ainsi, la probabilité de détection d'une cible (vrai positif) par les méthodes informatiques et algorithmiques

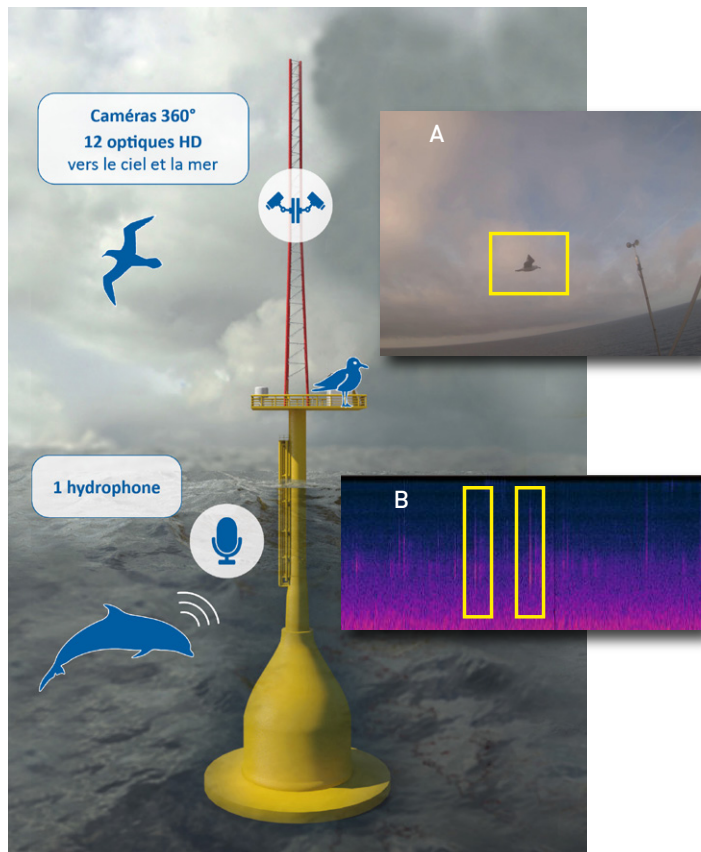
devrait constamment s'améliorer. En conséquence, il faudra être en mesure de prendre en compte l'amélioration des méthodes lors de l'interprétation de résultats provenant de méthodes en constante progression. Le risque étant, une fois de plus, d'observer des variations d'abondance non pas liées à une réalité écologique mais à une évolution des méthodes d'acquisition des données.

## 2 - Développement d'outils d'IA pour le suivi de la mégafaune marine

### 2.1 Contexte

Traditionnellement, les données sur la distribution et l'abondance des espèces de mégafaune marine reposent sur des observations directes à partir de suivis aériens et/ou maritimes (Hammond *et al.* 2021, Waggitt *et al.* 2020). Bien que ces méthodes offrent la possibilité d'échantillonner de vastes zones, elles ne fournissent qu'un aperçu à court terme de la réalité écologique. Les plateformes multi-instrumentées (Figure 3) émergent (mâts de mesures, tels que les mâts FINO et des bouées [akrocean] ou structures flottantes [données OCG] ; Wingett *et al.*, 2015) et sont prometteuses pour (i) étudier simultanément plusieurs compartiments de l'écosystème marin, (ii) pallier les limites propres de chaque instrument en collectant des informations écologiques complémentaires pour chaque compartiment (iii) mettre en place des suivis à long terme pour intégrer la variabilité temporelle de l'abondance et de la distribution des espèces, induite soit par la variabilité naturelle de l'en-

vironnement, soit par les activités humaines. Ce domaine de recherche en est encore à ses balbutiements car plusieurs barrières technologiques et techniques subsistent. Techniquement, la mise en place de suivis écologiques à haute fréquence, long terme, génèrent des ensembles de données volumineux et complexes, exigeants en termes de travail humain. Avec la percée significative de l'intelligence artificielle (IA), en particulier avec les réseaux de neurones profonds (par exemple récurrents et/ou convolutifs), l'automatisation de ces processus de surveillance est devenue un objectif réaliste (Goodwin *et al.* 2022). L'automatisation de la collecte et du traitement des données offre également l'opportunité unique de transmettre des informations écologiques en temps réel, ce qui est crucial pour optimiser la rentabilité des stratégies de suivi au sein des parcs éoliens qui sont appelés à être de plus en plus éloignés des côtes.



© Yohann Boutin, FEM

Fig. 3 : Représentation schématique d'une acquisition de données vidéo (A) et acoustique sous-marine (B) pour le suivi des oiseaux et mammifères marins autour d'une plateforme instrumentée. Les rectangles rouges indiquent des annotations de présence (A) d'oiseaux, (B) de sons de mammifères marins, qui servent à entraîner des modèles de détections automatiques.

## 2.2 Objectifs du projet

Le projet OWFSOMM visait à développer une suite de modèles d'IA pour optimiser l'utilisation de capteurs multiples afin d'améliorer leur efficacité dans la détection, l'identification et la caractérisation de la mégafaune marine. Plus spécifiquement, deux modalités ont été étudiées : l'acoustique sous-marine et l'imagerie vidéo 360°. Les objectifs initiaux incluaient également la fusion de ces modalités et l'extraction automatique d'indicateurs écologiques à partir de données multimodales. Cependant,

l'acquisition et la mise à disposition de données multimodales – c'est-à-dire plusieurs sources de données collectées de manière synchronisée pour observer le même événement – n'a pas pu être sécurisée sur la durée du projet. Bien que ces objectifs n'aient pas été atteints à l'échelle du projet OWFSOMM, l'utilisation de l'IA pour des données multimodales reste une perspective de recherche prometteuse pour détecter automatiquement et en temps réel la mégafaune marine.

## 3.3 Principales réalisations

### IA et données acoustiques

La surveillance des mammifères marins et la collecte d'informations sur leur environnement, susceptible d'affecter leur présence et l'utilisation de leur habitat (e.g. bruit sous-marin généré par les activités humaines) tout au long du cycle de vie des parcs éoliens en mer, sont cruciales. Chaque espèce de mammifère marin émet des sons qui lui sont propres mais possède également un répertoire de sons associés à différents comportements. Ainsi, l'acoustique passive sous-marine est utilisée dans le monde entier pour le suivi à long-terme des mammifères marins. Cependant, la grande quantité d'enregistrements audio générée soulève le besoin d'automatiser la détection des événements acoustiques. L'objectif était donc d'évaluer les performances des modèles d'apprentissage profond pour la détection et la classification des sons de mammifères marins.

Un hydrophone large bande (Figure 4), déployé pendant trois semaines sur le site éolien de Fécamp dans la Manche, a enregistré le paysage sonore sous-marin, notamment les sons des mammifères marins présents dans la zone. Pour visualiser ces sons, des images temps-fréquence (spectrogrammes) de 15 secondes ont été calculées. À partir de ces derniers, un total de 2 599 événements sonores de dauphins (D) et 1 689 marsouins (P) ont été annotés manuellement, y compris différents types de sons : trains de clicks (DCT : 2028, PCT : 1613), buzzes (DB : 254, PB : 76) et sifflements (DW : 317). Les spectrogrammes ont ensuite été divisés en cinq ensembles de

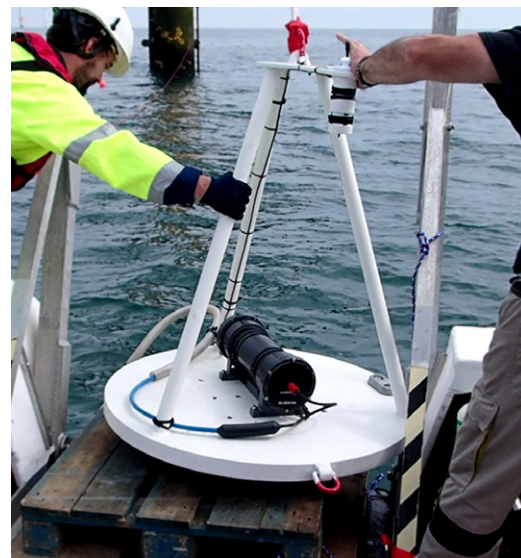


Fig. 4 : Déploiement d'un hydrophone large bande à proximité du mât de mesures situé au large de Fécamp.

données de validation croisée, chacun contenant la moitié des annotations manuelles et la moitié du bruit de fond uniquement. Un modèle Faster R-CNN+FPN a été entraîné pour détecter et localiser avec précision, dans le temps et en fréquence, les sons des mammifères marins dans les spectrogrammes ainsi que pour les classer par espèces et types de sons (Figure 5).

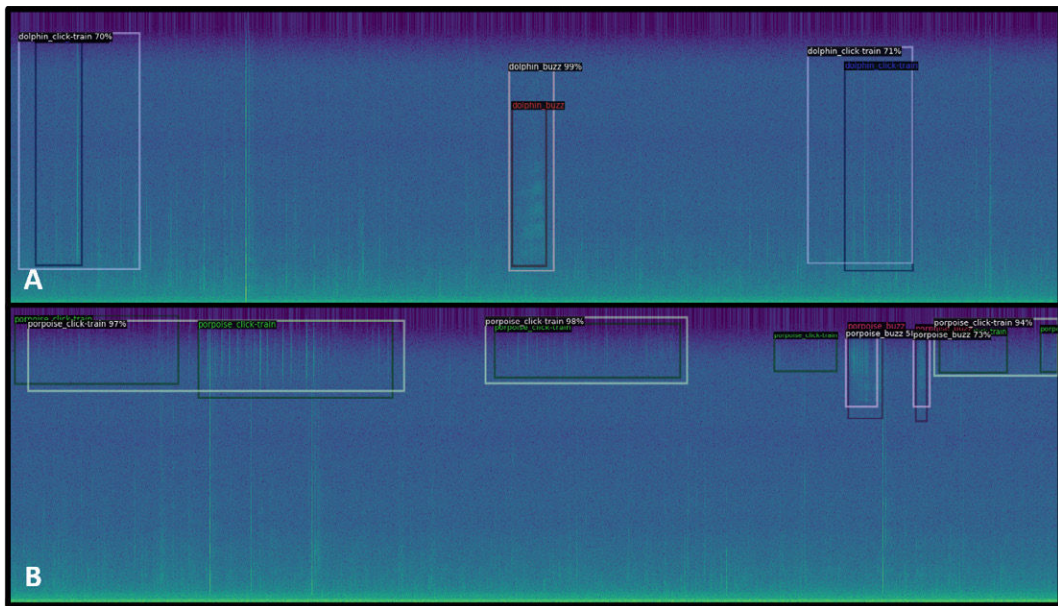
Trois configurations de sortie du modèle ont été utilisées : (1) détection globale des mammifères marins (présence ou absence), (2) détection et classification des espèces (deux classes : dauphin, marsouin) et (3) types de sons (cinq

classes : DCT, DB, DW, PCT, PB). Le modèle Faster R-CNN+FPN, pourtant conçu pour reconnaître des objets dans des images naturelles, a démontré qu'il restait une solution tout à fait pertinente dans le cas des spectrogrammes. En effet, nous avons pu atteindre des scores F1 allant de 0,88 pour le cas le plus simple (1) à 0,76 pour le cas le plus complexe (3).

Plusieurs paramètres du modèle peuvent être ajustés pour faire des compromis entre les détections manquées et les faux positifs. Ces choix et ajustements de paramètres doivent être soigneusement examinés et adaptés au problème. Par exemple, le modèle peut privilégier la sur-détection, laissant à l'expert la vali-

ation finale des différentes détections. Bien qu'un tel processus ne soit pas entièrement automatique, l'effort humain reste bien moindre qu'une analyse manuelle complète des données. Un tel scénario a été vérifié expérimentalement, permettant d'obtenir une précision de 96,3% en limitant la vérification de l'expert à 15,4% de l'ensemble des spectrogrammes.

En définitive, ces modèles sont prometteurs, que ce soit pour surveiller la présence de mammifères marins lors des constructions de parcs éoliens en mer ou pour des inférences écologiques tout au long du cycle de vie de ces parcs.



© Projet OWFSOMM

Fig. 5 : Résultats de détection et classification de mammifères marins par IA sur les spectrogrammes des données acoustiques. Détections correspondant aux annotations (A). Détections correspondant aux annotations, mais certaines annotations sont comprises dans une seule détection (B).

## IA et données optiques

Le développement de méthodes pour détecter les oiseaux traversant les parcs éoliens en mer et caractériser leur comportement à petite échelle autour de ces structures est crucial pour pouvoir mieux appréhender à terme les risques de collision des oiseaux avec les turbines. Le déploiement de caméras vidéo sur des structures offshore semble prometteur pour assurer une surveillance continue et à haute résolution de ces comportements à petite échelle. Pourtant, il

reste difficile d'analyser et de détecter automatiquement les oiseaux marins à partir de données vidéo. L'objectif était donc d'évaluer les performances de modèles d'apprentissage profond pour cette tâche de détection.

Une caméra HD 360°, conçue pour détecter les oiseaux volant dans un rayon de 100 à 500 mètres, a été déployée sur le mât de mesures de Fécamp dans la Manche. Sur la période étudiée

(une semaine continue d'acquisition de données), nous avons examiné manuellement environ 10% des données brutes, ce qui a conduit à l'identification de 101 passages d'oiseaux (un passage correspondant à une séquence d'images avec le même oiseau) et aux annotations de plus de 15 000 images. Cet ensemble de données a ensuite été utilisé pour entraîner, valider et évaluer deux approches distinctes de réseau profond pour la détection des oiseaux marins (Figure 6). La première approche implique un réseau populaire de détection d'objets connu sous le nom de Faster R-CNN, qui fonctionne dans un cadre supervisé. La deuxième approche, VAE-GRF, relève de la détection d'anomalies faiblement supervisée. Même si l'on s'attend à ce que cette dernière ait une capacité prédictive moindre, une telle approche offre l'avantage de ne pas s'appuyer sur une base de données annotées.

Le modèle Faster R-CNN a démontré des performances prédictives impressionnantes, atteignant un score F1 supérieur à 0,85 sur un ensemble de données de test. Cependant, son application dans les opérations en temps réel est actuellement limitée en raison (1) du temps de calcul important, (2) des biais de détection associés aux

conditions météorologiques et (3) de la nécessité de disposer d'ensembles de données annotées volumineux et de haute qualité. Malgré ces limites, le modèle Faster R-CNN pourrait être extrêmement utile pour détecter efficacement les passages d'oiseaux dans les données archivées de caméras vidéo HD à 360°, fournissant ainsi des informations précieuses sur le comportement de vol des oiseaux à petite échelle. En revanche, le modèle faiblement supervisé, VAE-GRF présentait un score F1 nettement inférieur, d'environ 0,08, en grande partie à cause d'un nombre excessif de fausses détections. Malgré cela, l'approche s'avère prometteuse car (1) elle ne dépend pas de données annotées manuellement, (2) elle est 70 fois plus rapide que le Faster R-CNN et (3) elle devrait présenter des propriétés de généralisation supérieures.

Des recherches supplémentaires sont nécessaires pour une détection précise et rapide des oiseaux marins à partir de données vidéo. Nous proposons d'explorer de nouvelles architectures de réseaux profonds et de développer conjointement des techniques supervisées et faiblement supervisées pour exploiter les avantages des deux approches.

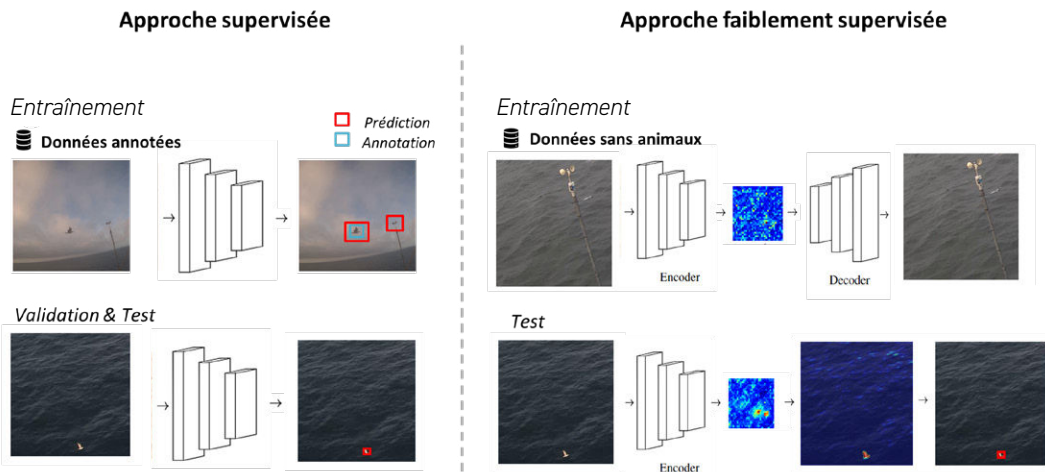


Fig. 6 : Résultats de détection d'oiseaux par IA sur des données vidéo HD 360°. Comparaison graphique des approches d'apprentissage supervisé et de détection d'anomalies faiblement supervisée. Les deux approches sont illustrées par les architectures Faster-R-CNN et VAE-GRF respectivement. La principale différence entre les deux approches est l'exigence en termes d'ensembles de données : une approche supervisée est basée sur des données annotées (c'est-à-dire avec des exemples de détection réels), tandis que l'approche faiblement supervisée ne nécessite que des exemples d'images sans oiseaux marins et est utilisée pour détecter des anomalies (c'est-à-dire des caractéristiques non rencontrées lors de l'apprentissage).

## RECOMMANDATIONS

Le projet OWFSOMM a permis de démontrer la pertinence de l'IA comme solution automatisée au suivi à fine échelle de la mégafaune marine au sein des parcs éoliens en mer, en considérant deux modalités distinctes : l'acoustique sous-marine pour la détection et la classification des mammifères marins, et l'imagerie vidéo 360° pour la détection des oiseaux. Au-delà des résultats prometteurs obtenus, les travaux ont également permis de mieux apprécier les conditions dans lesquelles les réseaux de neurones profonds peuvent être utilisés, et les verrous qu'il reste à lever pour un déploiement de ces solutions dans un contexte opérationnel à large échelle.

### LE RÔLE CENTRAL DE LA DONNÉE (ANNOTÉE)

#### 1 Entraîner les modèles d'IA avec de grandes quantités de données préalablement annotées

Les réseaux de neurones profonds sont devenus en quelques années une méthode de référence dans de nombreuses applications de la science des données, y compris en vision par ordinateur. Nous avons montré qu'ils pouvaient être appliqués avec succès aux données acoustiques ou aux vidéos HD 360° afin de détecter et reconnaître différentes espèces de mégafaune marine. Il n'a pourtant pas été nécessaire de concevoir des architectures spécifiques aux espèces étudiées ou aux données traitées. En effet, la capacité des réseaux de neurones à s'adapter à un contexte particulier (caractéristiques des sources de données ou des éléments d'intérêt à identifier) dépend surtout de la qualité du jeu de données avec lequel ils ont été entraînés. Ainsi, pour chaque source de données (acoustique sous-marine ou imagerie vidéo 360°), il a été nécessaire de préparer un jeu de données de qualité et de quantité suffisante. L'étape de collecte de données doit être suivie par un nettoyage ou filtrage permettant de limiter l'impact des données bruitées ou manquantes, voire un rééchantillonnage pour s'assurer que les classes (individus, espèces) minoritaires soient suffisamment représentées pour pouvoir être apprises (et ultérieurement reconnues) par les modèles d'IA. Mais le plus souvent, l'effort le plus important n'est pas tant dans la collecte des données que dans leur annotation. En effet,

dans le cadre de l'apprentissage supervisé (le paradigme le plus répandu en IA et celui aussi suivi dans la plupart de nos travaux), l'entraînement d'un modèle d'IA nécessite de fournir de très nombreux exemples des objets à détecter et à reconnaître. Ainsi, il ne suffit pas de collecter des données acoustiques ou des images 360° dans lesquelles peuvent être observés des indicateurs de mégafaune marine, mais il est aussi nécessaire d'annoter ces données. Pour la tâche de détection et classification d'objet qui a été étudiée dans ce projet, la phase d'annotation consiste à délimiter les régions d'intérêt et à les associer avec des classes d'intérêt. Cette étape est le plus souvent réalisée de façon manuelle, amenant un travail extrêmement fastidieux d'analyse visuelle de dizaines voire de centaines de milliers d'images. Bien que des outils automatiques d'annotation existent, leur performance reste insuffisante pour remplacer le travail d'un expert. **Pour atteindre un niveau de performance satisfaisant, il est nécessaire d'entraîner les modèles d'IA avec de grandes quantités de données qui auront au préalable été annotées : ce travail de collecte et d'annotation est le principal obstacle à la conception et au déploiement de ces modèles pour le suivi de la mégafaune marine dans les parcs éoliens en mer.**

## ACCROÎTRE LA TRANSFÉRABILITÉ

### 2 Mutualiser des données acquises dans des contextes différents

Les bonnes performances obtenues en détection et classification, que ce soit sur les données acoustiques ou sur les données vidéo 360°, pourraient laisser croire à l'obtention de solutions facilement déployables vers de nouveaux sites d'étude. Mais la réalité est tout autre. En effet, nos tests ont montré qu'un modèle d'IA entraîné sur un jeu de données (correspondant le plus souvent à un capteur spécifique installé dans un environnement donné) obtenait des performances bien plus faibles qu'escompté lorsqu'il était appliqué à des données provenant d'un autre capteur ou d'un autre site. Cette problématique, connue sous le nom de transférabilité ou d'adaptation de domaine, est centrale en science des données, et de nombreuses solutions ont été proposées dans la littérature scientifique. Nous n'avons cependant pas pu évaluer la capacité de

ces méthodes existantes à permettre cette transférabilité dans le contexte du projet OWFSOMM. **En l'absence de solution de transférabilité dont la pertinence aurait été démontrée dans le contexte du suivi de la mégafaune marine dans les parcs éoliens en mer, il reste nécessaire d'entraîner un modèle d'IA pour chaque nouveau capteur et/ou site d'étude. Toutefois, un modèle plus générique pourrait être entraîné en agrégeant des jeux de données acquis depuis plusieurs sites d'étude, permettant ainsi d'accroître la variété des éléments observés (et appris) par le réseau de neurones. Il est donc nécessaire d'envisager des projets collaboratifs permettant de mutualiser des données acquises dans des contextes différents (i.e. sites, paramètres des instruments).**

## L'EFFORT D'ANNOTATION

### 3 Limiter l'effort d'annotation en explorant d'autres stratégies d'apprentissage que la méthode supervisée consommatrice de ressources humaines

L'effort d'annotation a été identifié comme un des freins majeurs à l'adoption de l'IA comme solution automatisée de suivi de la mégafaune marine dans les parcs éoliens en mer. Cette difficulté étant directement liée au paradigme de l'apprentissage supervisé, il semble primordial d'explorer d'autres stratégies d'apprentissage. Nous avons montré que l'apprentissage non ou faiblement supervisé permettait de réduire drastiquement l'effort demandé à l'expert, en se limitant par exemple à une sélection manuelle des images ne contenant pas d'animaux. Toutefois, les performances obtenues restent encore bien inférieures à celles des modèles supervisés, illustrant le manque de maturité scientifique de l'apprentissage faiblement supervisé. Un autre paradigme aujourd'hui largement plébiscité par la communauté scientifique est celui de l'apprentissage auto-supervisé, qui consiste à générer automatiquement des données annotées à partir

des données brutes, permettant ainsi de mettre en œuvre un entraînement supervisé à la manière des réseaux utilisés dans le projet OWFSOMM. Une fois entraînés, ces réseaux auto-supervisés nécessitent cependant une phase de réglage (dite de fine-tuning) sur un jeu de données annotées, mais de taille réduite. **L'apprentissage supervisé, bien que représentant le paradigme le plus répandu dans les applications opérationnelles de l'IA, est fortement consommateur de ressources humaines pour la phase d'annotation préalable des données : les promesses des paradigmes alternatifs (apprentissage non-, faiblement-, semi-, ou auto-supervisé) appellent à engager des travaux pour apprécier leur pertinence dans le cadre du suivi de la mégafaune marine au sein des parcs éoliens en mer.**

## UNE NÉCESSAIRE TRANSPARENCE SUR LES MODÈLES ET LES PROTOCOLES

### 4 Favoriser le partage des codes sources, des modèles entraînés et des données utilisées

L'engouement pour les méthodes d'IA dans des domaines variés, y compris celui du suivi de la mégafaune marine, se traduit par un foisonnement d'études scientifiques démontrant l'intérêt de ces techniques. Cependant, il reste très difficile de reproduire les résultats des études publiées par autrui, ce qui nuit à une appréciation fiable du niveau de performance pouvant être atteint. Très souvent, le protocole d'évaluation (par exemple le découpage entre données d'entraînement et de test) et le paramétrage du modèle ne sont pas indiqués. Le partage des codes sources, des modèles entraînés (fichiers contenant les paramètres optimaux obtenus à

l'issue de la phase d'entraînement), mais aussi des données utilisées (et des annotations) est un prérequis à une transparence nécessaire pour pouvoir comparer les résultats des différentes études, des différents modèles, et pouvoir ainsi qualifier leur pertinence dans le contexte des parcs éoliens en mer. **Cette démarche de transparence ne doit pas seulement être suivie par les acteurs académiques mais également, dans la mesure du possible, par les opérateurs commerciaux, afin d'offrir à tous les acteurs de la filière une vision homogène et non-biaisée des possibilités de l'IA.**

## LA SYNCHRONISATION AU CŒUR DU CADRE MULTIMODAL

### 5 Obtenir des données réellement synchronisées

Parmi les objectifs initiaux du projet OWFSOMM, une attention particulière avait été portée à la conception de réseaux de neurones profonds aptes à opérer dans un cadre multimodal, c'est-à-dire à combiner différentes sources de données (sonar, lidar, optique, etc.) pour mieux caractériser les animaux pouvant être observés simultanément par différents capteurs. Ce cadre multimodal, largement répandu dans d'autres domaines (comme celui du véhicule autonome), reste cependant très difficile à mettre en œuvre en pratique dans le contexte des parcs éoliens en mer. En effet, les données expérimentales qui ont pu être acquises depuis le mât instrumenté de

Fécamp n'ont pas permis de combiner différentes observations des mêmes individus. Alors que les données acoustiques sous-marines ont été utilisées pour identifier les mammifères marins, les données vidéo 360° ont permis la détection des oiseaux. **Le potentiel du cadre multimodal reste donc à démontrer, mais suppose l'obtention de données réellement synchronisées qu'il sera ensuite possible de fusionner à l'aide de modèles de réseaux de neurones spécifiques. Ce faisant, la complémentarité des observations devrait permettre d'accroître la précision et la fiabilité des résultats obtenus.**

## LES CONTRAINTES DE DÉPLOIEMENT

### 6 Bien dimensionner la complexité des réseaux de neurones en vue de leur déploiement opérationnel et de leur exécution au plus près des capteurs

Les campagnes d'acquisition de données expérimentales conduites dans le projet ont permis de constituer des jeux de données particulièrement utiles pour évaluer le potentiel de l'IA pour le suivi de la mégafaune marine dans les parcs éoliens

en mer. Pour autant, de nombreux obstacles restent à franchir avant de permettre un déploiement opérationnel de ces solutions, afin d'aller vers de la détection de la mégafaune marine en temps réel. Un tel déploiement nécessitera d'exé-

cuter les algorithmes dans un contexte en ligne (sur site), et non pas hors ligne (sur le serveur d'un laboratoire). **Les ressources informatiques accessibles depuis un parc éolien en mer étant limitées, une attention particulière devra être portée en amont à la complexité des réseaux**

**de neurones, pour qu'ils puissent s'exécuter au plus près des capteurs. Cette problématique, connue sous le nom de *edge AI*, nécessite par exemple de compresser ou élaguer les réseaux préalablement entraînés en laboratoire.**

## LA SUITE ? LES MODÈLES FONDATIONS ET L'IA GÉNÉRATIVE

---

L'IA est un domaine en très forte évolution. Bien que les modèles les plus fréquemment déployés dans des contextes opérationnels ont été introduits dans la littérature il y a moins de 10 ans, la communauté scientifique est confrontée à une évolution majeure, qui a déjà trouvé un public plus large : les modèles fondations d'IA générative, dont ChatGPT et DALL-E sont deux des exemples les plus connus. Ces très grands modèles,

entraînés sur d'immenses volumes de données dans un cadre auto-supervisé, opèrent sur une ou plusieurs modalités (texte et image notamment), et peuvent répondre à des tâches variées. Nul doute que dans un futur plus ou moins proche, il sera possible de les mobiliser autour de tâches de suivi de la mégafaune marine dans les parcs éoliens en mer.

## Conclusion

Les solutions digitales de surveillance de la mégafaune marine par survols aériens ont connu des développements techniques importants et sont désormais utilisées dans le cadre du suivi environnemental de parcs éoliens en mer français. Garantir la commensurabilité des données ainsi acquises avec celles issues des survols aériens embarquant des observateurs est essentiel. En outre, l'utilisation de plateformes instrumentées en mer pour le suivi de la mégafaune marine est de plus en plus envisagée et prometteuse. Des développements méthodologiques sont toutefois nécessaires pour automatiser les futurs suivis et extraire les informations écologiques pertinentes des données multimodales obtenues.

Dans le cadre du projet OWFSOMM, 14 campagnes de suivis aériens de la mégafaune marine, impliquant une surveillance visuelle à basse altitude et une surveillance digitale à la fois à basse et à haute altitude, ont été menées sur cinq sites différents et à différentes périodes de l'année au sein d'un même site. Ces données ont permis le calcul d'une centaine de facteurs d'intercalibration sur lesquels l'influence de variables, telles que l'espèce étudiée ou les conditions spécifiques au site, a été explorée. Par ailleurs, le projet OWFSOMM a permis le développement d'algorithmes de détection automatisée d'oiseaux et de mammifères marins à partir de

données vidéos 360° et d'acoustique sous-marine, respectivement.

OWFSOMM a mis en évidence la difficulté de comparer des estimations d'abondance obtenues par les méthodes visuelles et digitales, montrant ainsi qu'il est préférable d'opter pour une même méthode de suivi aérien tout au long du cycle de vie d'un parc éolien en mer. Compte tenu des altitudes de vol en vigueur dans les parcs, il convient de favoriser les survols digitaux dès l'état de référence. Le développement d'algorithmes automatisés est prometteur sous réserve de suivre un cadre de bonnes pratiques incluant une transparence quant aux performances avérées des modèles et des jeux de données utilisés pour leur entraînement et leur validation. Par ailleurs, il est important de confronter les modèles développés à une multitude de sources de données, collectées dans des environnements différents, afin d'en assurer leur transférabilité. Dans le cadre du projet, les développements d'algorithmes et la mise en œuvre d'une approche d'acquisition de données multimodales ont permis de poser la première brique opérationnelle vers des suivis intégrés, continus et à long terme des différents compartiments de la mégafaune marine, objectifs actuellement poursuivis dans le cadre du projet DRACCAR-MMERMAID.

## Bibliographie

- Bailey, H., Senior, B., Simmons, D., Rusin, J., Picken, G., Thompson, H.M., 2010. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Marine Pollution Bulletin* 60, 888–897. <https://doi.org/10.1016/j.marpolbul.2010.01.003>
- Berg, P., Santana Maia, D., Pham, M.-T., Lefèvre, S., 2022. Weakly Supervised Detection of Marine Animals in High Resolution Aerial Images. *Remote Sensing* 14, 339. <https://doi.org/10.3390/rs14020339>
- Beyan, C., Browman, H.I., 2020. Setting the stage for the machine intelligence era in marine science. *ICES Journal of Marine Science* 77, 1267–1273. <https://doi.org/10.1093/icesjms/fsaa084>
- Brautaset, O., Waldeland, A.U., Johnsen, E., Malde, K., Eikvil, L., Salberg, A.-B., Handegard, N.O., 2020. Acoustic classification in multifrequency echosounder data using deep convolutional neural networks. *ICES Journal of Marine Science* 77, 1391–1400. <https://doi.org/10.1093/icesjms/fsz235>
- Croll, D.A., Ellis, A.A., Adams, J., Cook, A.S.C.P., Garthe, S., Goodale, M.W., Hall, C.S., Hazen, E., Keitt, B.S., Kelsey, E.C., Leirness, J.B., Lyons, D.E., McKown, M.W., Potiek, A., Searle, K.R., Soudijn, F.H., Rockwood, R.C., Tershy, B.R., Tinker, M., VanderWerf, E.A., Williams, K.A., Young, L., Zilliacus, K., 2022. Framework for assessing and mitigating the impacts of offshore wind energy development on marine birds. *Biological Conservation* 276, 109795. <https://doi.org/10.1016/j.biocon.2022.109795>
- Desholm, M., Fox, A.D., Beasley, P.D.L., Kahlert, J., 2006. Remote techniques for counting and estimating the number of bird–wind turbine collisions at sea: a review. *Ibis* 148, 76–89. <https://doi.org/10.1111/j.1474-919X.2006.00509.x>
- Goodfellow, I., Bengio, Y., Courville, A., 2016. *Deep Learning*. MIT Press.
- Goodwin, M., Halvorsen, K.T., Jiao, L., Knausgård, K.M., Martin, A.H., Moyano, M., Oomen, R.A., Rasmussen, J.H., Sørtdalen, T.K., Thorbjørnsen, S.H., 2022. Unlocking the potential of deep learning for marine ecology: overview, applications, and outlook†. *ICES Journal of Marine Science* 79, 319–336. <https://doi.org/10.1093/icesjms/fsab255>
- Hammond, P.S., Francis, T.B., Heinemann, D., Long, K.J., Moore, J.E., Punt, A.E., Reeves, R.R., Sepúlveda, M., Sigurðsson, G.M., Siple, M.C., Víkingsson, G., Wade, P.R., Williams, R., Zerbini, A.N., 2021. Estimating the Abundance of Marine Mammal Populations. *Frontiers in Marine Science* 8
- Hazen, E.L., Abrahms, B., Brodie, S., Carroll, G., Jacox, M.G., Savoca, M.S., Scales, K.L., Sydeman, W.J., Bograd, S.J., 2019. Marine top predators as climate and ecosystem sentinels. *Frontiers in Ecology and the Environment* 17, 565–574. <https://doi.org/10.1002/fee.2125>
- Jelicich, R.M., Berón, P., Copello, S., Dellabianca, N.A., García, G., Labrada-Martagón, V., Viola, M.N.P., Paz, J.A., Riccialdelli, L., Martin, A.S., Pon, J.P.S., Torres, M.A., Favero, M., 2022. Marine Megafauna Sea Turtles, Seabirds and Marine Mammals, in: *Marine Biology*. CRC Press.
- Malde, K., Handegard, N.O., Eikvil, L., Salberg, A.-B., 2020. Machine intelligence and the data-driven future of marine science. *ICES Journal of Marine Science* 77, 1274–1285. <https://doi.org/10.1093/icesjms/fsz057>
- Mutanu, L., Gohil, J., Gupta, K., Wagio, P., Kotonya, G., 2022. A Review of Automated Bioacoustics and General Acoustics Classification Research. *Sensors* 22, 8361. <https://doi.org/10.3390/s22218361>
- Shiu, Y., Palmer, K.J., Roch, M.A., Fleishman, E., Liu, X., Nosal, E.-M., Helble, T., Cholewiak, D., Gillespie, D., Klinck, H., 2020. Deep neural networks for automated detection of marine mammal species. *Sci Rep* 10, 607. <https://doi.org/10.1038/s41598-020-57549-y>
- Waggitt, J., Evans, P., Andrade, J., Banks, A., Boisseau, O., Bolton, M., Bradbury, G., Brereton, T., Camphuysen, C., Durinck, J., Felce, T., Fijn, R., García-Barón, I., Garthe, S., Geelhoed, S., Gilles, A., Goodall, M., Haelters, J., Hamilton, S., Hiddink, J., 2019. Distribution maps of cetacean and seabird populations in the North East Atlantic. *Journal of Applied Ecology* 57, 253–269. <https://doi.org/10.1111/1365-2664.13525>

Wingett, S.W., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., Andrews, S., 2015. HiCUP: pipeline for mapping and processing Hi-C data. <https://doi.org/10.12688/f1000research.7334.1>



L'éolien en mer s'est imposé comme l'une des technologies les plus prometteuses du nouveau mix énergétique et connaît une expansion rapide. D'ici fin 2024, quatre parcs éoliens en mer seront en service en France et une cinquantaine de parcs sont projetés d'ici à 2050 sur le littoral français. Ces infrastructures peuvent avoir des impacts sur la mégafaune marine, impacts qui doivent être estimés et réduits dans le cadre des politiques environnementales. Le suivi de la mégafaune marine est donc crucial à chaque étape du cycle de vie d'un parc éolien en mer.

Dans ce contexte, OWFSOMM avait pour objectifs :

- de développer des méthodes et outils d'intercalibration pour le suivi de la mégafaune marine à partir de survol aérien afin de d'assurer la comparabilité entre les observations visuelles et digitales.
- de développer des algorithmes de détection et d'identification automatique de la mégafaune marine à partir de source de données multimodales.

A la lumière des résultats obtenus, ce rapport propose des recommandations aussi bien opérationnelles que techniques afin d'optimiser la collecte de données et leur traitement afin d'assurer un suivi pertinent de la mégafaune marine à l'échelle d'un parc éolien. Il est organisé en deux parties. La première concerne l'intercalibration des méthodes pour le suivi de la mégafaune marine à partir de survols aériens, la deuxième traite du développement d'outils d'IA pour le suivi de la mégafaune marine.

Ce rapport s'adresse à l'ensemble des acteurs de la filière de l'éolien en mer susceptibles de collecter, analyser, interpréter des données destinées à assurer les suivis de la mégafaune marine, impliquant aussi bien des campagnes de survols aériens que des suivis dont les résultats reposent sur l'utilisation d'algorithmes automatisés.



Bâtiment Cap Océan  
Technopôle Brest Iroise  
525, avenue Alexis de Rochon  
29280 PLOUZANÉ

02 98 49 98 69

[www.france-energies-marines.org](http://www.france-energies-marines.org)

ISBN 978-2-493115-40-9



9 782493 115409