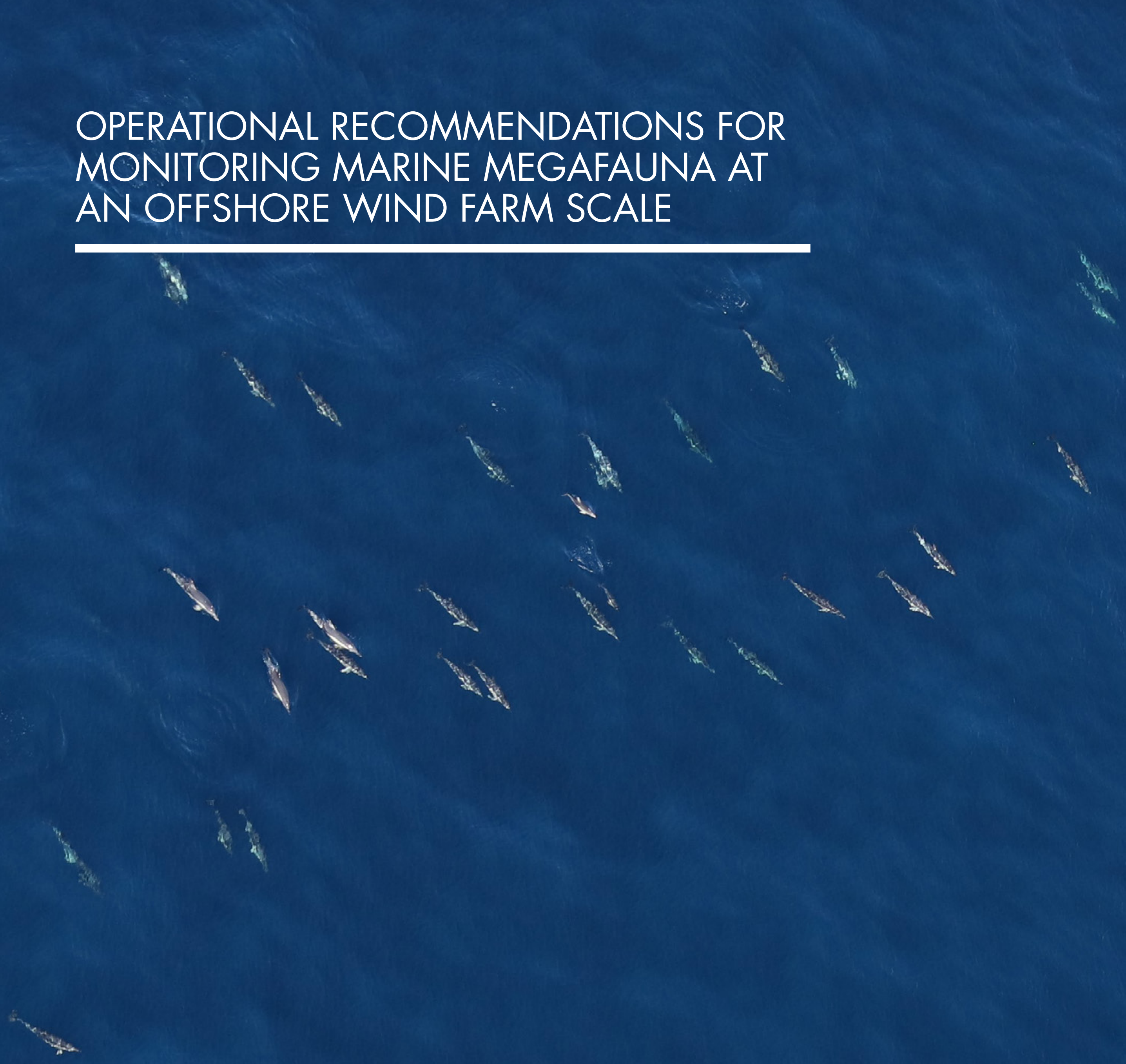
OPERATIONAL RECOMMENDATIONS FOR MONITORING MARINE MEGAFAUNA AT AN OFFSHORE WIND FARM SCALE



FRANCE ENERGIES MARINES
Editions

OPERATIONAL RECOMMENDATIONS FOR MONITORING MARINE MEGAFAUNA AT AN OFFSHORE WIND FARM SCALE

AUTHORS:
Maud Quéroué, France Energies Marines, Centre for Functional and Evolutionary Ecology
Sébastien Lefèvre, University of South Brittany
Minh-Tan Pham, University of South Brittany
Matthieu Authier, La Rochelle Université
Aurélien Besnard, Centre for Functional and Evolutionary Ecology
Karine Heerah, France Energies Marines

The document should be cited as follows :
Quéroué M., Lefèvre S., Pham M.-T., Authier M., Besnard A. et Heerah K.
Operational recommendations for monitoring marine megafauna at an offshore wind farm scale
Plouzané: France Energies Marines Editions, 2025, 28 pages.
Published: February 2025
Photo cover: Group of dolphins captured by an automated imaging system
© Pelagis / Hytech-imaging - STORMM

Legal deposit at publication

Table of contents

List of figures

Page

List of tables

Acronyms

AI	Artificial Intelligence
OWFSOMM	Offshore Wind Farm Surveys Of Marine Megafauna

Acknowledgements

This report was written as part of the collaborative OWFSOMM project (Offshore Wind Farm Surveys Of Marine Megafauna), led by France Energies Marines, the Centre for Functional and Evolutionary Ecology, and IRISA. The OWFSOMM project received funding from France Energies Marines, its members and partners, the French Office for Biodiversity, the Directorate General for Energy and Climate, as well as support from the French State managed by the National Research Agency under the France 2030 Investment Plan (ANR-10-IEED-0006-34). We would like to thank the 13 project partners (France Energies Marines, University of South Brittany, National Centre for Scientific Research, PELAGIS Observatory, WIPSEA, EDF Renouveblables, Ocean Winds, Shell, RWE, Ifremer, Eoliennes en mer Dieppe Le Tréport, CE-Sigma, ENSTA Bretagne) for their contributions throughout the OWFSOMM project. We thank all the individuals who participated in the aerial surveys, collected data, and analysed images, including Hytech-imaging, Biotope, Setec In-vivo, and Bretagne Vivante. We also thank EDF Renewables, Ocean Winds, and the French government (Directorate General for Energy and Climate) for providing data from aerial surveys conducted as part of regulatory environmental assessment studies. We also thank EDF Renouvelables and the company Eoliennes en Mer des Hautes Falaises (EOHF) for providing the passive underwater acoustic dataset used in this study, as well as the OSmOSE team (ENSTA Bretagne, Lab-STICC, UMR CNRS 6285) for their bioacoustic advice.

Introduction

The increasing number of offshore wind farms raises major environmental and societal questions regarding their impact on wildlife. Marine megafauna species, such as fish, turtles, birds, and marine mammals, are at the heart of concerns due to the risks of direct mortality (e.g., collision with turbines) and indirect mortality (e.g., loss of functional habitats, noise disturbances) induced by these new structures (Bailey et al., 2014; Croll et al., 2022). These species, many of which are protected, endangered, or threatened, represent the top of the food chains and play a key role in ecosystem functioning. Their ecology and distribution integrate the spatio-temporal variations of their food sources. Therefore, these species can be used as indicators to monitor the health of marine ecosystems (Hazen et al., 2019; Jelicich et al., 2022). Consequently, for each wind farm, monitoring marine megafauna is crucial at every stage, from authorisation to construction and operation. However, observing marine megafauna species remains difficult as they spend most of their time underwater and/or offshore.

Marine megafauna monitoring can be carried out by observers on a boat or by aerial survey methods. In France, baseline assessments and pre-installation surveys are mainly conducted by aerial visual observation at an altitude of 200 metres or less to satisfactorily detect and identify small species such as seabirds or porpoises. However, during operational phases, aerial survey techniques will require a minimum altitude of 300 metres to comply with safety regulations. At this higher altitude, data acquisition will be done by digital means. This change in data acquisition during the lifecycle of offshore wind farms can affect the observation process (i.e., detection of marine megafauna) and potentially the abundance estimates of megafauna. Therefore, it is essential to intercalibrate, at the analysis stage, these data collected using different methods (visual vs. digital).

In parallel, multi-sensor platforms can be installed to monitor wind farm sites during the different phases of their lifecycle, from baseline assessment to decommissioning. These platforms are of great importance as they allow for refined data acquisition on megafauna ecology at the wind farm scale and enable continuous and long-term monitoring. Moreover, the use of multiple sensors simultaneously helps to overcome the limitations of each technology regarding the level of information they can provide on marine megafauna monitoring. These data can then feed different types of predictive models, particularly to assess collision risks for seabird populations and noise exposure for marine mammals and fish. In this context, OWFSOMM aimed to:

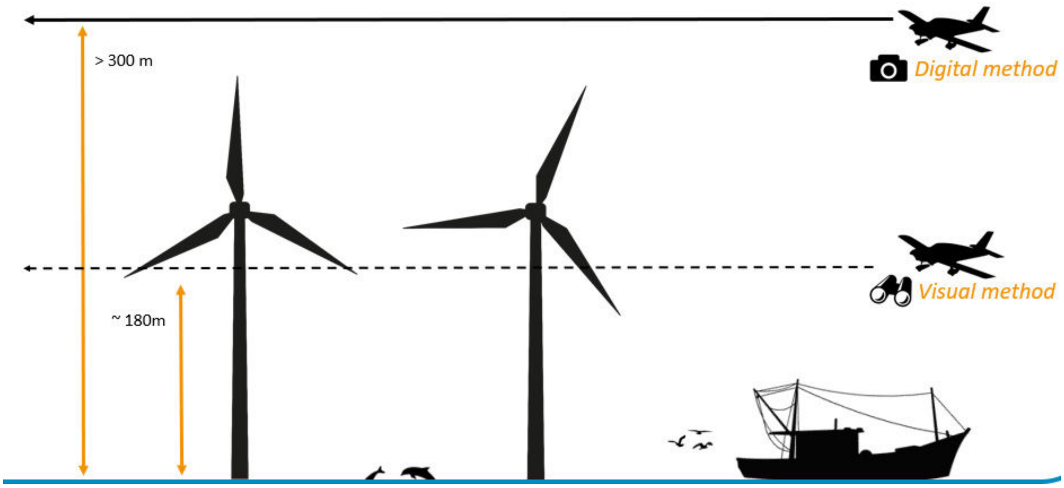
- Develop intercalibration methods and tools for marine megafauna monitoring from aerial surveys to ensure comparability between visual and digital observations,
- Develop automatic detection and identification algorithms for marine megafauna from multimodal data sources,
- Provide methodological recommendations to optimise marine megafauna monitoring at the wind farm scale.

The objective of this report is to propose both operational and technical recommendations to optimise data collection and processing to ensure relevant monitoring of marine megafauna. These recommendations are provided in light of the results obtained in the OWFSOMM project and are detailed in two parts. The first part concerns the intercalibration of methods for marine megafauna monitoring from aerial surveys. The second part deals with the development of AI tools for marine megafauna monitoring.

1 - Intercalibration of aerial survey methods for marine megafauna monitoring

1.1 Context

In recent years, digital solutions for monitoring marine megafauna through aerial surveys have seen significant technical developments. Deployable at high altitudes, they comply with wind farm overflight requirements **(Figure 1)** and will soon be used in the environmental monitoring of various French offshore wind projects. However, existing environmental data comes from aerial surveys with observers flying at low altitudes. Therefore, it is necessary to ensure the commensurability of results obtained by different methods, both with traditional visual aerial survey methods and with more recent digital techniques.



© FEM

Fig. 1: *Schematic representation of the data acquisition methods deployed during the aerial monitoring of offshore wind farms.*

1.2 Project objectives

- Provide an operational roadmap for reliable intercalibration of aerial surveys of marine megafauna in offshore wind farms using different technologies.
- Develop a methodology for intercalibration between aerial surveys with human observers and those conducted with digital techniques.
- Draft recommendations for monitoring megafauna in offshore wind projects to ensure the interoperability of all data.

© Pelagis / Hytech-imaging - STORMM

1.3 Main achievements

A data format protocol for digital methods

A standardised data format adapted to digital data acquisition methods has been developed in consultation with various partners involved in digital monitoring.

This protocol allows for the inclusion of the same information collected during visual aerial surveys while taking into account the specificities associated with digital surveys and data processing.

Intercalibration campaigns for visual and digital methods

14 intercalibration campaigns were conducted at 5 different sites (Figure 2) and at different times of the year (Table 1). During these campaigns, two planes flying no more than 15 minutes apart performed the same transects over the area. One plane flying at low altitude (~180 m) carried observers as well as a digital system to conduct simultaneous digital and visual monitoring. A second plane flying at high altitude, over 300 m, carried a digital data acquisition system.

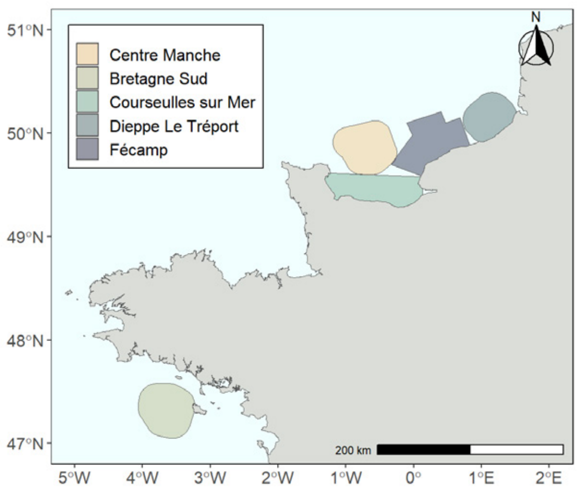


Fig. 2: *Map of the study sites of the OWFSOMM project*

Campaign			Low altitude Visual observation	Low altitude Digital obervation		High altitude Digital obervation	
Site	*Date*	*Sea state (Average, Beaufort)*	*Altitude*	*Altitude*	*Digital methods*	*Altitude*	*Digital methods*
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 Tré-port (EMDT1)	27/02/2022	3.3	~180 m	~180 m	B	~550 m	C
Dieppe Le Tré-port (EMDT2)	09/06/2022	1.9	~180 m	~180 m	B	~550 m	C
Dieppe Le Tré-port (EMDT3)	04/09/2022	3.1	~180 m	~180 m	B	~550 m	C
Dieppe Le Tré-port (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: *Description of the data acquisition methods deployed during the campaigns conducted.*
Digital methods used: A: WIPSEA, B: Hytech-imaging, C: HiDef

A comparability analysis of different visual and digital methods

For each dataset obtained from the various aerial surveys carried out during an intercalibration campaign, an estimate of the abundance and distribution of individuals was made (Quéroué *et al.*, 2024). Since the actual number of individuals present in the study area was unknown, the aim was not to assess the methods' ability to accurately estimate abundance in the area, but only to compare the results obtained by each method. An intercalibration factor was thus calculated as the ratio between the abundance estimated by one method and the abundance estimated by a second method. In this way, intercalibration factors were calculated between visual and digital methods, as well as between digital methods for all campaigns, for different flight altitudes, and for different species groups (gannets, alcids, gulls, divers, cetaceans, sharks, etc.).

The results showed that, on average, digital methods estimate higher abundances than visual methods. However, the intercalibration factors are highly variable depending on sea state and the species studied. The methods produce fairly similar abundance estimates when the analyses concern easily detectable species such as northern gannets (*Morus bassanus*). Conversely, for inconspicuous species such as alcids, divers, or grebes, visual methods generally provide lower abundance estimates than digital methods. Furthermore, when the sea is rough, the difference in abundance estimates between the two methods is greater than in calm sea conditions, with visual methods underestimating. A large part of the variability (>60%) of the estimated intercalibration factors is not explained by these different elements. These results highlight the difficulty, if not the impossibility, of defining general and automatically transferable intercalibration factors to compare abundance estimates obtained by different visual and digital methods without considering the local context in which the data were collected (site, sea state, target species, etc.).

Finally, the results show that, on average, the abundance estimates from high and low altitude digital methods were consistent, with intercalibration factors close to 1. However, there is variability around this average intercalibration factor, which could only be explained by 18% by the digital method used and the sea state.

An analysis of factors influencing anomaly detection algorithms

A database of 1054 images from the campaigns, containing targets of interest to be identified or not, was created to meticulously characterise the quality of the images and the environmental conditions in which they were taken. This characterisation was carried out by a human operator. The images were examined to search for and identify the «objects» present, such as birds or marine mammals, based on reference identification guides. A comparison of detections was then made between the operator's detections and those of the algorithms used to detect objects of interest in the images. This comparison helps to identify potential false negatives, i.e., objects present in an image but not detected by the algorithm. The results of this work show that the frequency of false negatives is strongly correlated with image sharpness and sun glare, and it varies according to the species group studied. This analysis highlights the importance of having a quantitative measure of image quality to account for the imperfect detection of objects of interest.

An R package 'sismow' for simulating sampling efforts and individuals in a study area

A package encapsulating a suite of dedicated functions for conducting numerical simulation studies for intercalibration studies is available at the following address: https://github.com/maudqueroue/sismow. This package allows for the simulation of 'realistic' distributions of marine megafauna in a given area, then simulates a sampling campaign design, and finally simulates a data acquisition method. The simulated dataset can reproduce the main characteristics of a real dataset. These simulations will enable statistical power analyses to adequately and prospectively dimension aerial survey campaigns of marine megafauna in terms of effort.

RECOMMENDATIONS

The recommendations mentioned below pertain to the environmental monitoring of offshore wind farms. They are only valid in the context of monitoring marine megafauna to assess the impact of offshore wind farms throughout their lifecycle. Therefore, these recommendations do not concern the large-scale monitoring of mobile species.

PLANNING AND DATA COLLECTION

1 Promote the use of digital methods for ecological monitoring of marine megafauna conducted as part of offshore wind farm development

The project results showed that using intercalibration factors to transition from a visual data acquisition method to a digital data acquisition method is not straightforward (Quéroué *et al.*, 2024). The correct use of intercalibration factors is subject to numerous contextual factors that greatly limit their transferability and validity. Considering that digital methods will necessarily be used for safety reasons during the construction and operational phases of wind farms, we recommend using digital methods from the initial environmental monitoring stages. The aim is to use the same type of method throughout the lifecycle of the wind farms to analyse temporal trends in marine megafauna abundance without being biased by changes in data acquisition methods.

2 Plan surveys favouring calm sea conditions as much as possible

To minimise variability between different digital surveys over time, similar survey conditions should be prioritised. It is particularly important to favour the calmest possible sea conditions. Rough seas can make it difficult to detect individuals that are present, especially for species that spend most of their time underwater. The risk associated with sea state variability is underestimating the number of individuals when sea conditions are rough. Failure to account for these perception and availability biases can lead to erroneous ecological conclusions. Furthermore, target detection becomes more complex in rough seas. There is a risk of obtaining a large number of false positives due to wavelets being detected by automated detection algorithms, resulting in a lengthy target identification process. Conversely, algorithms trained not to detect wavelets may miss individuals that blend into the waves, leading to false negatives and an underestimation of the number of individuals in the study area.

3 Minimise sun glare and optimise image sharpness

The results of the analysis of factors influencing the detection of digital methods highlighted the impact of sun glare and image sharpness on the number of false negatives. The presence of sun glare on images increases the risk that algorithms will miss an individual that is actually present in the image. Similarly, it was observed that the detection capability of the algorithms was reduced when the images were blurry. During aerial transects, it is therefore important to minimise sun glare on the images as much as possible and to ensure proper focus during image capture to achieve the sharpest possible targets.

DATA ACQUISITION & PROCESSING

4 Refer to the OWFSOMM data format protocol for data collected during digital surveys

This project has established a standardised data format adapted to digital methods. We recommend using this data format as it allows for the collection of all necessary information to conduct abundance and distribution analyses from data obtained using digital methods.

5 Engage qualified naturalists for the identification of targets detected by digital methods

Just like the individuals boarding planes for visual surveys, those analysing images acquired during digital surveys must be skilled in species identification. It is essential that the people responsible for this work have strong naturalist expertise in identifying marine megafauna, particularly seabirds, which are the most represented species in offshore wind farm areas.

6 Provide information on the performance of the anomaly detection algorithms used

Detection errors in digital methods have been observed in relation to (1) environmental conditions such as the presence of sun glare on images, (2) image quality such as sharpness, and (3) the species observed. To manage these sources of bias and correct the estimated abundances, it is necessary to know the performance of the anomaly detection algorithms used. These performances can be presented in the form of a confusion matrix by species, environmental conditions, and image quality, while also providing information on the dataset on which the algorithm was tested. Such information will allow for tracking the evolution of the algorithms' capabilities over time and taking their progress into account when interpreting the results.

7 Implement quality analyses for target detection and identification stages

One way to ensure the quality of datasets provided by digital methods is to implement verification procedures. These analyses aim to validate the quality of the data processing and primarily concern the following two phases: target detection and target identification. Regarding target detection, it is recommended that a portion of the images be re-analysed by a human operator and that the detections made by human vision be compared to those of the algorithm. It is necessary to ensure that the detection capabilities of the algorithm match the mentioned performances (see recommendation 6). Regarding target identification, it is recommended that a person who did not participate in the identification of the detected targets independently analyse a subsample of the detected targets to confirm or refute the identification at the smallest possible taxonomic level. The objective is to ensure the accuracy of species identifications.

USE OF RESULTS FOR ECOLOGICAL MONITORING

8 Do not estimate temporal trends using results from both visual and digital methods

Intercalibration campaigns have shown the great variability of intercalibration factors between visual and digital methods and their close dependence on partially elucidated contextual factors. Using intercalibration factors without considering the local context in which the data were collected (site, sea state, target species, methods used, observers, etc.) would not allow reliable evaluation of abundance fluctuations over time. The risk is that, over the lifecycle of an offshore wind farm, the observed variations would only be the result of a change in monitoring method and not a real change in abundance. Consequently, the use of results from both visual and digital methods to establish temporal trends should be avoided.

9 Do not interpret slight variations in abundance over time

The results of intercalibration campaigns have shown that abundance estimates from digital methods at high and low altitudes were, on average, consistent and had intercalibration factors close to 1. However, there is still variability around the average intercalibration factor, which is only minimally explained by different variables such as the digital method used or sea state. Therefore, when estimating temporal trends using results from digital methods, care must be taken not to interpret slight changes in abundance, which may result from variability between surveys rather than an ecological reality.

10 Take into account the improvement of digital methods

Digital methods are still under development. Therefore, it is anticipated that these approaches will improve over time, and thus over the lifecycle of offshore wind farms. Similarly, it is reasonable to expect that anomaly detection algorithms will become more efficient or that image quality will improve. Consequently, the probability of detecting a target (true positive) by computational and algorithmic techniques should constantly improve. Therefore, it will be necessary to take into account these advancements when interpreting results from evolving technologies. The risk, once again, is to observe variations in abundance not linked to an ecological reality but to an evolution in data acquisition methods.

2 - Development of AI tools for monitoring marine megafauna

2.1 Context

Traditionally, data on the distribution and abundance of marine megafauna species rely on direct observations from aerial and/or maritime surveys (Hammond *et al.*, 2021, Waggitt et al. 2020). Although these methods offer the possibility of sampling large areas, they only provide a short-term snapshot of ecological reality. Multi-instrument platforms (Figure 3) are emerging (measurement masts, such as FINO masts and buoys [akrocean] or floating structures [OCG data]; Wingett *et al.,* 2015) and are promising for (i) simultaneously studying several compartments of the marine ecosystem, (ii) overcoming the limitations of each instrument by collecting complementary ecological information for each compartment, and (iii) establishing long-term monitoring to integrate the temporal variability of species abundance and distribution, induced either by natural environmental variability or by human activities. This research field is still in its infancy as several technological and technical barriers remain. Technically, setting up high-frequency, long-term ecological monitoring generates large and complex datasets, demanding in terms of human labour. With the significant breakthrough of artificial intelligence (AI), particularly with deep neural networks (e.g., recurrent and/or convolutional), automating these monitoring processes has become a realistic goal (Goodwin *et al.*, 2022). Automating data collection and processing also offers the unique opportunity to transmit ecological information in real-time, which is crucial for optimising the cost-effectiveness of monitoring strategies within wind farms that are increasingly located further offshore.

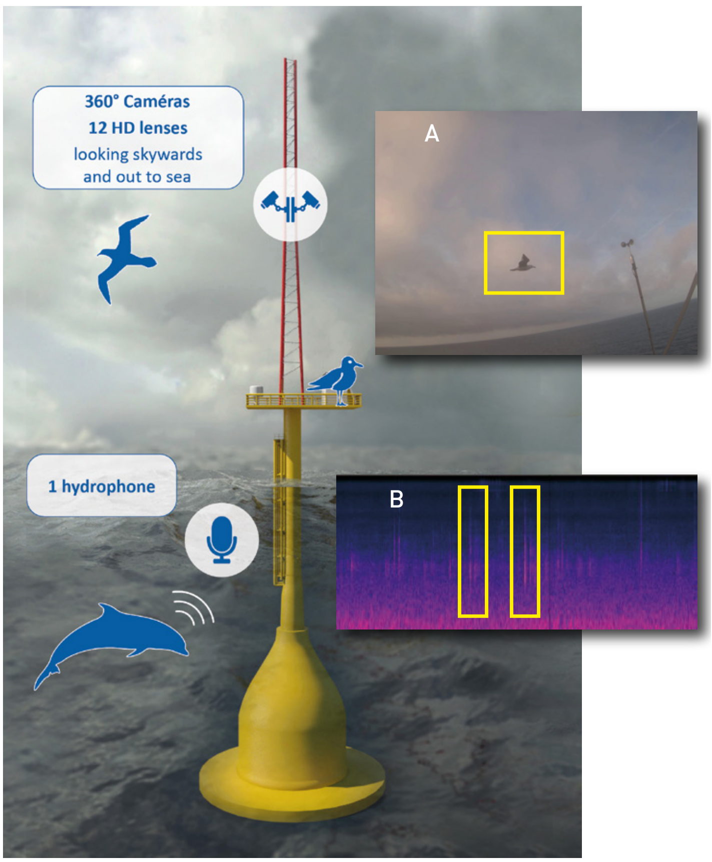


Fig. 3: *Schematic representation of video (A) and underwater acoustic (B) data acquisition for monitoring birds and marine mammals around an instrumented platform. The yellow rectangles indicate presence annotations of (A) birds and (B) marine mammal sounds, which are used to train automatic detection models.*

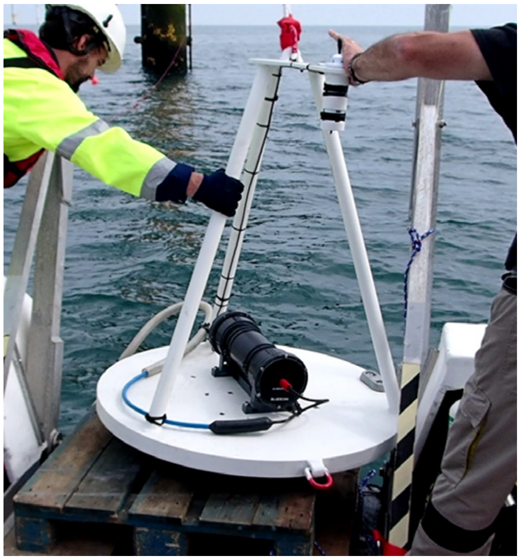
2.2 Project objectives

The OWFSOMM project aimed to develop a suite of AI models to optimise the use of multiple sensors and improve their efficiency in detecting, identifying, and characterising marine megafauna. Specifically, two modalities were studied: underwater acoustics and 360° video imaging. The initial objectives also included the fusion of these modalities and the automatic extraction of ecological indicators from multimodal data. However, the acquisition and availability of multimodal data – that is, multiple data sources collected synchronously to observe the same event – could not be secured over the project's duration. Although these objectives were not achieved within the OWFSOMM project, the use of AI for multimodal data remains a promising research perspective for automatically and in real-time detecting marine megafauna.

3.3 Main achievements

AI and acoustic data

Monitoring marine mammals and collecting information about their environment, which may affect their presence and habitat use (e.g., underwater noise generated by human activities) throughout the lifecycle of offshore wind farms, is crucial. Each marine mammal species emits unique sounds but also has a repertoire of sounds associated with different behaviours. Thus, passive underwater acoustics is used worldwide for the long-term monitoring of marine mammals. However, the large amount of audio recordings generated raises the need to automate the detection of acoustic events. Here, the objective was to evaluate the performance of deep learning models for detecting and classifying marine mammal sounds.



© FEM

Fig. 4: *Deployment of a broadband hydrophone near the measurement mast located off the coast of Fécamp.*

A broadband hydrophone (Figure 4), deployed for three weeks at the Fécamp wind farm site in the English Channel, recorded the underwater soundscape, including the sounds of marine mammals present in the area. To visualise these sounds, 15-second time-frequency images (spectrograms) were calculated. From these, a total of 2,599 dolphin (D) and 1,689 porpoise (P) sound events were manually annotated, including different types of sounds: click trains (DCT: 2028, PCT: 1613), buzzes (DB: 254, PB: 76), and whistles (DW: 317). The spectrograms were then divided into five cross-validation datasets, each containing half of the manual annotations and half of the background noise only. A Faster R-CNN+FPN model was trained to accurately detect and locate, in time and frequency, marine mammal sounds in the spectrograms and classify them by species and sound types (Figure 5).

Three model output configurations were used: (1) global detection of marine mammals (presence or absence), (2) detection and classification of species (two classes: dolphin, porpoise), and (3) sound types (five classes: DCT, DB, DW, PCT, PB). The Faster R-CNN+FPN model, although designed to recognise objects in natural images, proved to be a highly relevant solution for spectrograms. Indeed, we achieved F1 scores ranging from 0.88 for the simplest case (1) to 0.76 for the most complex case (3).

Several model parameters can be adjusted to balance missed detections and false positives. These choices and parameter adjustments must be carefully examined and adapted to the problem. For example, the model can favour over-detection, leaving the final validation of the detections to the expert. Although such a process is not entirely automatic, the human effort remains much less than a complete manual analysis of the data.

This scenario was experimentally verified, achieving an accuracy of 96.3% while limiting expert verification to 15.4% of the total spectrograms.

Ultimately, these models are promising, whether for monitoring the presence of marine mammals during the construction of offshore wind farms or for ecological inferences throughout the lifecycle of these farms.

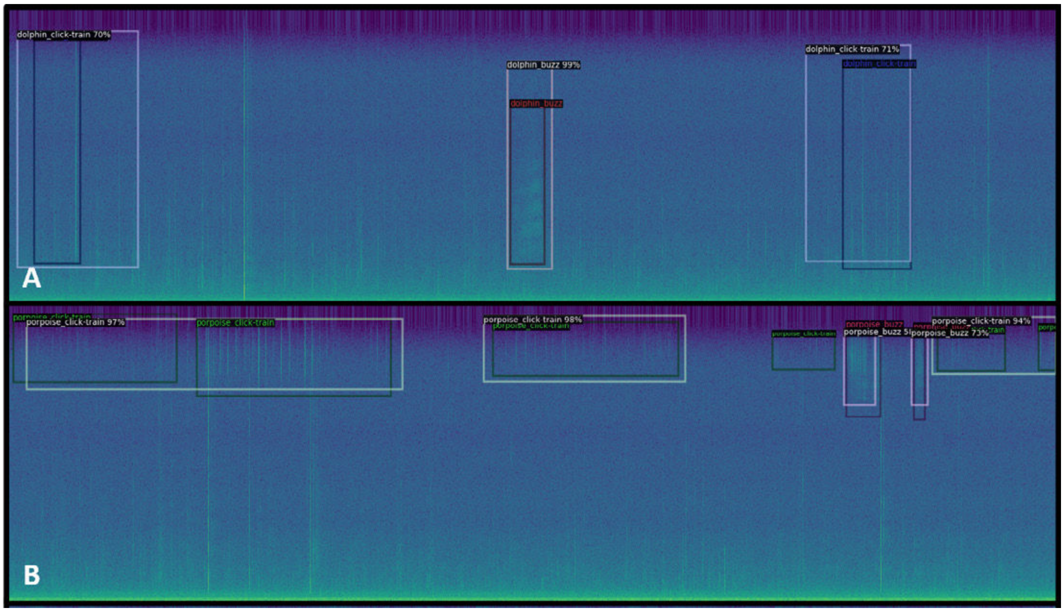


Fig. 5: *Results of marine mammal detection and classification by AI on the spectrograms of acoustic data. Detections corresponding to annotations (A). Detections corresponding to annotations, but some annotations are included in a single detection (B).*

AI and optical data

The development of methods to detect birds crossing offshore wind farms and characterise their behaviour on a small scale around these structures is crucial to better understand the risks of bird collisions with turbines. Deploying video cameras on offshore structures seems promising for ensuring continuous and high-resolution monitoring of these small-scale behaviours. However, it remains challenging to analyse and automatically detect seabirds from video data. The objective was therefore to evaluate the performance of deep learning models for this detection task.

A 360° HD camera, designed to detect birds flying within a radius of 100 to 500 metres, was deployed on the measurement mast located off the coast of Fécamp in the English Channel. We manually examined around 10% of the raw data collected over the study period (one week of continuous data acquisition), leading to the identification of 101 bird passages (a passage corresponding to a sequence of images with the same bird) and the annotation of over 15,000 images. This dataset was then used to train, validate, and evaluate two distinct deep network approaches for seabird detection (Figure 6). The first approach involves a popular object detection network known as Faster R-CNN, which operates in a supervised framework. The second approach, VAE-GRF, pertains to weakly supervised anomaly detection. Although the latter is expected to have lower predictive capacity, it offers the advantage of not relying on an annotated database.

The Faster R-CNN model has demonstrated impressive predictive performance, achieving an F1 score above 0.85 on a test dataset. However, its application in real-time operations is currently limited due to (1) significant computation time, (2) detection biases associated with weather conditions, and (3) the need for large, high-quality annotated datasets. Despite these limitations, the Faster R-CNN model could be extremely useful for effectively detecting bird passages in archived 360° HD video camera data, thus providing valuable insights into small-scale bird flight behaviour. In contrast, the weakly supervised model, VAE-GRF, exhibited a significantly lower F1 score of around 0.08, largely due to an excessive number of false detections. Nevertheless, the approach is promising because (1) it does not rely on manually annotated data, (2) it is 70 times faster than Faster R-CNN, and (3) it is expected to have superior generalisation properties.

Further research is needed for accurate and rapid detection of seabirds from video data. We propose exploring new deep network architectures and jointly developing supervised and weakly supervised techniques to leverage the advantages of both approaches.

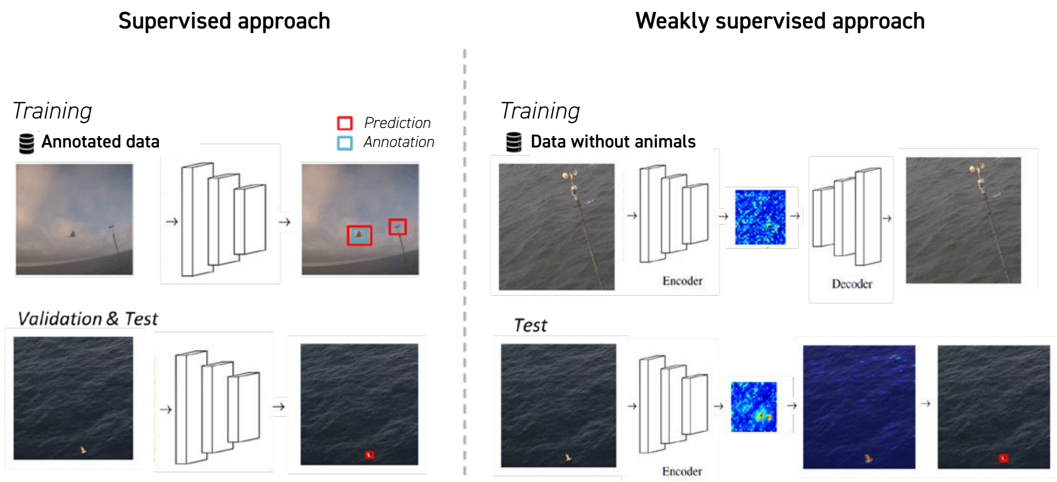


Fig. 6: *AI Bird Detection Results on 360° HD Video Data. Graphical comparison of supervised learning and weakly supervised anomaly detection approaches. The two approaches are illustrated by the Faster R-CNN and VAE-GRF architectures, respectively. The main difference between the two approaches is the dataset requirement: a supervised approach relies on annotated data (i.e., with actual detection examples), while the weakly supervised approach only requires examples of images without seabirds and is used to detect anomalies (i.e., features not encountered during training).*

RECOMMENDATIONS

The OWFSOMM project has demonstrated the relevance of AI as an automated solution for fine-scale monitoring of marine megafauna within offshore wind farms, considering two distinct modalities: underwater acoustics for the detection and classification of marine mammals, and 360° video imaging for bird detection. Beyond the promising results obtained, the work has also provided a better understanding of the conditions under which deep neural networks can be used, and the challenges that remain to be addressed for the deployment of these solutions in a large-scale operational context.

THE CENTRAL ROLE OF (ANNOTATED) DATA

1 Training AI models with large amounts of pre-annotated data

Deep neural networks have become a reference method in many data science applications, including computer vision, in just a few years. We have shown that they can be successfully applied to acoustic data or 360° HD videos to detect and recognise different species of marine megafauna. It was not necessary to design architectures specific to the species studied or the data processed. The ability of neural networks to adapt to a particular context (characteristics of data sources or elements of interest to be identified) mainly depends on the quality of the dataset with which they have been trained. Thus, for each data source (underwater acoustics or 360° video imaging), it was necessary to prepare a dataset of sufficient quality and quantity. The data collection step must be followed by cleaning or filtering to limit the impact of noisy or missing data, or even resampling to ensure that minority classes (individuals, species) are sufficiently represented to be learned (and subsequently recognised) by AI models. However, the most significant effort is often not in data collection but in their annotation. In the context of supervised learning (the most widespread approach in AI and also followed in most of our work), training an AI model requires providing numerous examples of objects to be detected and recognised. Thus, it is not enough to collect acoustic data or 360° images in which indicators of marine megafauna can be observed, but it is also necessary to annotate these data. For the object detection and classification task studied in this project, the annotation phase consists of delimiting regions of interest and associating them with classes of interest. This step is most often carried out manually, leading to an extremely tedious visual analysis of tens or even hundreds of thousands of images. Although automatic annotation tools exist, their performance is still insufficient to replace the work of an expert. **To achieve a satisfactory level of performance, it is necessary to train AI models with large amounts of data that have been previously annotated: this data collection and annotation work is the main obstacle to the design and deployment of these models for monitoring marine megafauna in offshore wind farms.**

ENHANCING TRANSFERABILITY

2 Pooling Data Acquired in Different Contexts

The good performance achieved in detection and classification, whether on acoustic data or 360° video data, might suggest the possibility of easily deployable solutions to new study sites. However, the reality is quite different. Indeed, our tests have shown that an AI model trained on a dataset (usually corresponding to a specific sensor installed in a given environment) achieved much lower performance than expected when applied to data from another sensor or site. This issue, known as transferability or domain adaptation, is central in data science, and many solutions have been proposed in the scientific literature. However, we were unable to evaluate the ability of these existing methods to enable this transferability in the context of the OWFSOMM project. **In the absence of a transferability solution whose relevance has been demonstrated in the context of monitoring marine megafauna in offshore wind farms, it remains necessary to train an AI model for each new sensor and/or study site. However, a more generic model could be trained by aggregating datasets acquired from multiple study sites, thereby increasing the variety of elements observed (and learned) by the neural network. It is therefore necessary to consider collaborative projects that allow pooling data acquired in different contexts (i.e., sites, instrument settings).**

THE ANNOTATION EFFORT

3 Limiting the annotation effort by exploring learning strategies other than the resource-intensive supervised method

The annotation effort has been identified as one of the major obstacles to the adoption of AI as an automated solution for monitoring marine megafauna in offshore wind farms. This difficulty is directly related to the supervised learning framework, making it essential to explore other learning strategies. We have shown that unsupervised or weakly supervised learning can drastically reduce the effort required from experts, for example, by limiting manual selection to images without animals. However, the performance achieved is still far inferior to that of supervised models, illustrating the scientific immaturity of weakly supervised learning. Another framework widely endorsed by the scientific community today is self-supervised learning, which involves automatically generating annotated data from raw data, thus enabling supervised training similar to the networks used in the OWFSOMM project. Once trained, these self-supervised networks still require a fine-tuning phase on a smaller annotated dataset. **Supervised learning, although the most widespread framework in operational AI applications, is highly resource-intensive for the preliminary data annotation phase: the promises of alternative frameworks (unsupervised, weakly supervised, semi-supervised, or self-supervised learning) call for further research to assess their relevance in the context of monitoring marine megafauna in offshore wind farms.**

A NECESSARY TRANSPARENCY ON MODELS AND PROTOCOLS

4 Encouraging the sharing of source codes, trained models, and used data

The enthusiasm for AI methods in various fields, including the monitoring of marine megafauna, has led to a proliferation of scientific studies demonstrating the interest of these techniques. However, it remains very difficult to reproduce the results of studies published by others, which hinders a reliable assessment of the level of performance that can be achieved. Very often, the evaluation protocol (e.g., the split between training and test data) and the model parameters are not indicated. Sharing source codes, trained models (files containing the optimal parameters obtained at the end of the training phase), as well as the data used (and annotations), is a prerequisite for the necessary transparency to compare the results of different studies and models, and thus to assess their relevance in the context of offshore wind farms. **This transparency approach should not only be followed by academic actors but also, as far as possible, by commercial operators, in order to offer all stakeholders in the sector a homogeneous and unbiased view of the possibilities of AI.**

SYNCHRONISATION AT THE HEART OF THE MULTIMODAL FRAMEWORK

5 Obtaining truly synchronised data

Among the initial objectives of the OWFSOMM project, particular attention was paid to the design of deep neural networks capable of operating in a multimodal framework, i.e., combining different data sources (sonar, lidar, optical, etc.) to better characterise animals that can be observed simultaneously by different sensors. This multimodal framework, widely used in other fields (such as autonomous vehicles), remains very difficult to implement in practice in the context of offshore wind farms. The experimental data acquired from the instrumented mast at Fécamp did not allow for combining different observations of the same individuals. While underwater acoustic data were used to identify marine mammals, 360° video data enabled the detection of birds. **The potential of the multimodal framework remains to be demonstrated, but it requires obtaining truly synchronised data that can then be fused using specific neural network models. In doing so, the complementarity of observations should increase the accuracy and reliability of the results obtained.**

DEPLOYMENT CONSTRAINTS

6 Properly sizing the complexity of neural networks for their operational deployment and execution close to the sensors

The experimental data acquisition campaigns conducted in the project have made it possible to create datasets that are particularly useful for evaluating the potential of AI for monitoring marine megafauna in offshore wind farms. However, many obstacles remain before these solutions can be operationally deployed to enable real-time detection of marine megafauna. Such deployment will require running the algorithms in an online context (on-site), rather than offline (on a laboratory server). **Given the limited computing resources available from an offshore wind farm, particular attention must be paid in advance to the complexity of the neural networks, so they can run close to the sensors. This issue, known as edge AI, requires, for example, compressing or pruning the networks previously trained in the laboratory.**

WHAT'S NEXT? FOUNDATION MODELS AND GENERATIVE AI

AI is a rapidly evolving field. Although the models most frequently deployed in operational contexts were introduced in the literature less than 10 years ago, the scientific community is facing a major evolution that has already found a broader audience: foundation models of generative AI, with ChatGPT and DALL-E being two of the most well-known examples. These very large models, trained on immense volumes of data in a self-supervised framework, operate on one or more modalities (notably text and image) and can perform various tasks. There is no doubt that in the near or distant future, it will be possible to leverage them for tasks related to monitoring marine megafauna in offshore wind farms.

Conclusion

Digital solutions for monitoring marine megafauna through aerial surveys have seen significant technical developments and are now used in the environmental monitoring of French offshore wind farms. Ensuring the commensurability of data acquired with new digital methods with those from aerial surveys involving observers is essential. Moreover, the use of instrumented platforms at sea for monitoring marine megafauna is increasingly considered and promising. However, methodological developments are necessary to automate future monitoring and extract relevant ecological information from the obtained multimodal data.

As part of the OWFSOMM project, 14 aerial survey campaigns of marine megafauna, involving low-altitude visual monitoring and both low and high-altitude digital surveillance, were conducted at five different sites and at different times of the year within the same site. These data allowed the calculation of about a hundred intercalibration factors, on which the influence of variables such as the studied species or site-specific conditions was explored. Furthermore, the OWFSOMM project enabled the development of automated detection algorithms for birds and marine mammals from 360° video data and underwater acoustics, respectively.

OWFSOMM highlighted the difficulty of comparing abundance estimates obtained by visual and digital methods, showing that it is preferable to opt for the same aerial survey method throughout the lifecycle of an offshore wind farm. Given the higher flight altitudes required in the farms, it is advisable to favour digital surveys from the reference state. The development of automated algorithms is promising, provided that a framework of best practices is followed, including transparency regarding the proven performance of the models and the datasets used for their training and validation. Moreover, it is important to confront the developed models with a multitude of data sources collected in different environments to ensure their transferability. Within the project, the development of algorithms and the implementation of a multimodal data acquisition approach have laid the first operational foundation for integrated, continuous, and long-term monitoring of the different compartments of marine megafauna, objectives currently pursued within the DRACCAR-MMERMAID project.

References

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., Sørdalen, 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

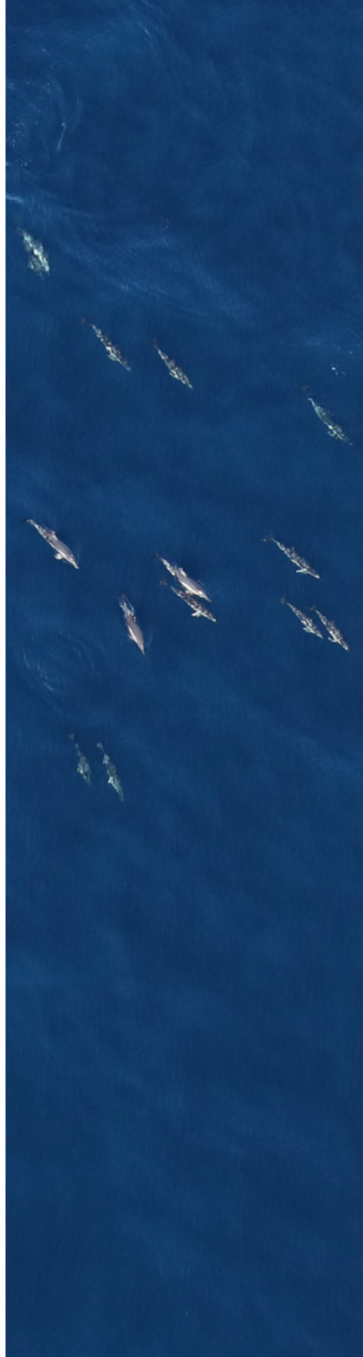
Quéroué, M., Authier, M., Besnard, A., Heerah, K., 2024 Going digital: challenges in monitoring marine megafauna when comparing results from visual and digital aerial surveys. Frontiers in Marine Science, 11:1432798. https://doi.org/10.3389/fmars.2024.1432798

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



Offshore wind power has emerged as one of the most promising technologies in the new energy mix and is experiencing rapid expansion. By the end of 2024, four offshore wind farms will be operational in France, and around fifty wind farms are projected by 2050 along the French coast. These infrastructures can impact marine megafauna, which must be estimated and mitigated within the framework of environmental policies. Monitoring marine megafauna is therefore crucial at every stage of the lifecycle of an offshore wind farm.

In this context, OWFSOMM aimed to:

- Develop methods and tools for intercalibration to ensure comparability between visual and digital observations for monitoring marine megafauna from aerial surveys.
- Develop algorithms for the automatic detection and identification of marine megafauna from multimodal data sources.

In light of the results obtained, this report proposes both operational and technical recommendations to optimise data collection and processing to ensure relevant monitoring of marine megafauna at the scale of a wind farm. It is organised into two parts. The first part concerns the intercalibration of methods for monitoring marine megafauna from aerial surveys, and the second part deals with the development of AI tools for monitoring marine megafauna.

This report is addressed to all stakeholders in the offshore wind sector who may collect, analyse, and interpret data to ensure the monitoring of marine megafauna, involving both aerial survey campaigns and monitoring whose results rely on the use of automated algorithms.

FRANCE ENERGIES MARINES Editions

Bâtiment Cap Océan
Technopôle Brest Iroise
525, avenue Alexis de Rochon
29280 PLOUZANÉ
02 98 49 98 69
www.france-energies-marines.org

ISBN 978-2-493115-41-6