Institut Pythéas
Observatoire des Sciences de l'Univers
Aix-Marseille Université

UAlg FCT
UNIVERSIDADE DO ALGARVE
FACULDADE DE CIÊNCIAS E TECNOLOGIA

SiPLABoratory

MASTER SCIENCES DE LA MER

Parcours : Océanographie Physique et Biogéochimique

Benoît CHAMBON

Advancing Low Frequency Marine Mammal Calls Detection Algorithm

Internship report written in the laboratory: ***SiPLABoratory***

Supervised by: ***Sergio JESUS***

Academic year: ***2023–2024***

SiPLAB, FCT
University of Algarve

Gambelas Campus

PT-8005-139
Faro, Portugal

Acknowledgments

First of all, I would like to thank the SiPLABoratory team for giving me the opportunity to carry out an internship in the laboratory. This experience was both enriching and stimulating, enabling me to develop my skills and acquire new knowledge in the field of marine acoustic research. I would also like to express my sincere thanks to my internship supervisor, Sergio JESUS. The advice he gave me, his expertise and constant support, as well as his kindness and availability, have greatly helped me to progress professionally. I'm grateful for this experience and the valuable knowledge it has given me.

Contents

1. Introduction

As human activities increasingly impact on marine ecosystems, the marine acoustic environment is changing significantly, posing a threat to marine life. Marine mammals rely on sound for essential functions like navigation, communication, and prey detection (Edds-Walton, 1997). Commercial shipping, offshore exploration, seismic surveys, and military sonar operations have been identified as particularly harmful to marine fauna (Merchant et al., 2014), leading to adverse effects including physical injury, hearing loss, altered behaviour, and disrupted social and reproductive patterns (Weilgart, 2007; Merchant et al., 2014). Responses to these disturbances vary depending on factors such as species, age, and previous exposure to noise, complicating ecosystem management efforts (Weilgart, 2007; Gillespie et al., 2013). Despite growing research attention (Gillespie et al., 2013), understanding of marine ecosystems and mammal responses to disturbances remains limited (Zimmer, 2011; Correia et al., 2021).

Cetaceans emit a wide range of sounds spanning from low frequency (LF) to high frequency (HF), covering 10 Hz to 100 kHz (Herman, 1980; Richardson et al., 1995). These sounds vary in duration, modulation, and structure, ranging from brief calls to prolonged series lasting hours (Sayigh, 2013). Baleen whales, distributed globally, produce unique acoustic calls, with LF calls (< 500 Hz) being specific to this group (Harland and Armstrong, 2004; Nowacek et al., 2007). Call types within species share similarities but may differ in duration, frequency, and structure. For instance, fin whales often produce 20 Hz calls lasting 1-2 seconds, while blue whales produce amplitude- and frequency-modulated calls with longer durations (Watkins et al., 1987; McDonald et al., 2006). These calls are associated with social interactions or feeding behaviours (Širović et al., 2014; Lewis et al., 2018), and their variations provide insights into geographic and sub-population differences (Thompson et al., 1992; McDonald et al., 2006). Acoustic detection and monitoring of these calls aid in species identification, subpopulation assessment, and understanding migratory patterns (Stafford et al., 2001; Davis et al., 2020).

Several techniques have been developed for automatic detection and classification of cetacean signals, considering factors such as background noise, species sound variability, and relevant acoustic parameters (Marques et al., 2013). Performance of these techniques relies on species, recording environment, dataset size, and extracted descriptors (Mellinger and Bradbury, 2007). Evaluation metrics include precision, sensitivity, specificity, receiver operating characteristic (ROC), recall (true positive rate), F-measure, and positive/false positive/false negative rates (Sokolova and Lapalme, 2009; Knight et al., 2017). Various automatic detection algorithms have been employed for detecting the LF calls of baleen whales. These include matched filters (Stafford et al., 1998; Harris et al., 2013), energy detectors (Morano et al., 2012), and subspace projection detectors (Socheleau et al., 2015; Leroy et al., 2016). However, spectrogram correlation has emerged as the most widely used algorithm (Mellinger and Clark, 2000), implemented in various software (Figueroa and Robbins, 2008) and applied across diverse datasets (Balcazar et al., 2015; Aulich et al., 2019).

The spectrogram correlation, like matched filters, evaluates spectrogram output pixels by pixels rather than in time or frequency domains, by comparing image models or kernels directly with spectrogram results. However, its effectiveness, as highlighted by Miller et al. (2021), fluctuates according to factors such as signal-to-noise ratio (SNR) and other environmental variables such as acoustic propagation and call density. While this facilitates comparisons between sites and time periods, manual verification may be required to deal with false positives, which is problematic with large datasets. Then, deep neural networks offer a promising alternative to spectrogram correlation detectors, potentially yielding superior performance (Usman et al., 2020). These advanced algorithms not only resolve the uncertainties involved in estimating call density, but also reduce the human effort required to confirm true positives and deal with false positives. Despite their advantages such as pattern recognition capabilities, adaptability to the environment and flexibility, neural networks have limitations, including the need for large training datasets (Ibrahim et al., 2018; Rasmussen and Širović, 2021). Finally, the subspace algorithm for LF calls detection, developed by Scharf and Friedlander (1994); Kraut et al. (2001), offers accurate detection capabilities suitable for this study. Previously demonstrated by Socheleau et al. (2015) to significantly outperform XBAT (spectrogram correlation) by 15-20%, this method adapts well to transient signals, which are important for detecting the temporal variations of whale calls in dynamic acoustic environments. Although it requires prior knowledge of signal modes, unlike generally machine learning, this approach allows precise control of performance, making it promising for the automatic detection of LF calls. The approach has to date been applied exclusively to *B. musculus* and tested with real data (Socheleau et al., 2015; Leroy et al., 2016).

Expanding this method to include additional baleen whale species to perform a library, alongside *B. musculus*, is the focus of the study. We aim to develop an adaptable approach for LF vocalization detection using acoustic gliders in a simulated marine environment, in the context of the TRIDENT project (begins on June, 2024) at the Tropic Seamount (23.90°N, -20.70°E). Acoustic gliders, known for their silent operation and ability efficiently to cover vast distances, offer a valuable platform for passive acoustic monitoring (Baumgartner et al., 2020), overcoming limitations of other marine monitoring technologies (Goldstein and Bentley, 2010). We seek to assess its robustness and performance across various contexts, aiming to enhance our understanding of marine acoustic biodiversity and provide effective tools for cetacean conservation.

2. Materials and methods

2.1. Study areas

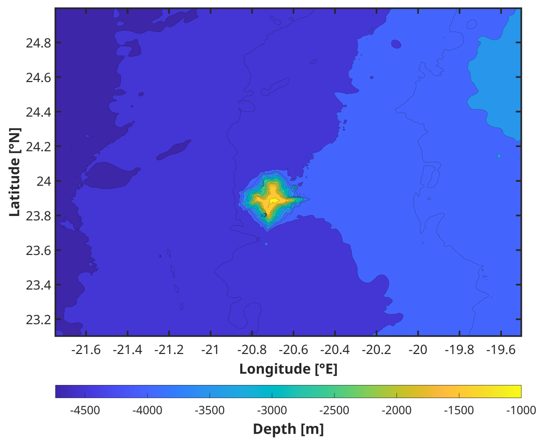

Fig. 1. Tropic Seamount bathymetry (23.90°N, -20.70°E).

The Tropic Seamount is set in the Canary Islands seamount region, west of the Western Sahara coast and southwest of the Canary Islands, north of Cape Verde (23.90°N, -20.70°E). It rises to a depth of 970 meters (Fig.1). Seamounts are recognized for their role in the functioning of various underwater ecosystems, affecting the currents and increasing the flow of prey, ultimately creating an ”oasis” of productivity (Pitcher et al., 2007), where trophic cascades are triggered by attracting a variety of marine megafauna (Morato et al., 2008). Studies have identified cetaceans and seabirds around seamounts, suggesting them as ecological ”hotspots” (Yen et al., 2004; Garrigue et al., 2015).

The Tropic Seamount is on the migration route of 3 major species of baleen whales in Atlantic waters (Romagosa et al., 2020; Valente et al., 2019): *B. musculus* (blue whale), *B. physalus* (fin whale), and *B. borealis* (sei whale). Baleen whales are known to utilize seamounts extensively for various activities such as feeding, breeding, resting, and navigation, highlighting the critical role of seamounts in their offshore habitat (Garrigue et al., 2015).

2.2. Simulated Environment

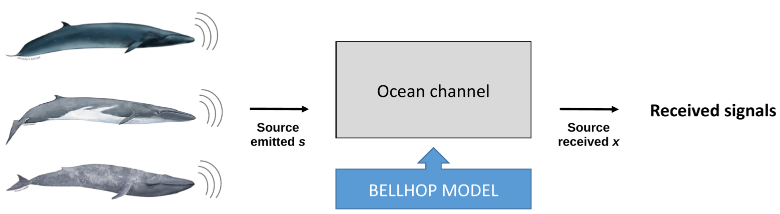

Fig. 2. Study problem diagram.

Here, our detection problem is applied in a simulated environment in which the different LF calls of the 3 baleen whale species pass through an oceanic acoustic channel modeled using BELLHOP (Fig.2). We aim to define a library of the LF calls and to test the performance of the detector in discriminating among the calls in the library. This configuration provides a close-to-reality scenario for evaluating the performance of our detection algorithm and its ability to accurately identify whale calls in the complexity of the underwater acoustic environment.

2.2.1. BELLHOP model

The BELLHOP model is a ray-tracing acoustic pressure calculation tool developed by Porter (2011). It uses beam tracing techniques and employs Gaussian and hat-shaped beams with geometric and physical spreading laws. In our case, we simulate our simulated environment, using BELLHOP, by calculating the acoustic channel where whale LF calls pass through and are received by a receiver. The model takes into account various environmental factors, such as sound speed profile (SSP), ocean surface and bottom properties, bathymetry, and reflection coefficients. The model outputs transmission loss, eigenrays, arrival times, and received time data, enabling thorough acoustic analysis (Fig.3). The model dynamically reads input files according to the options selected in the environment file.

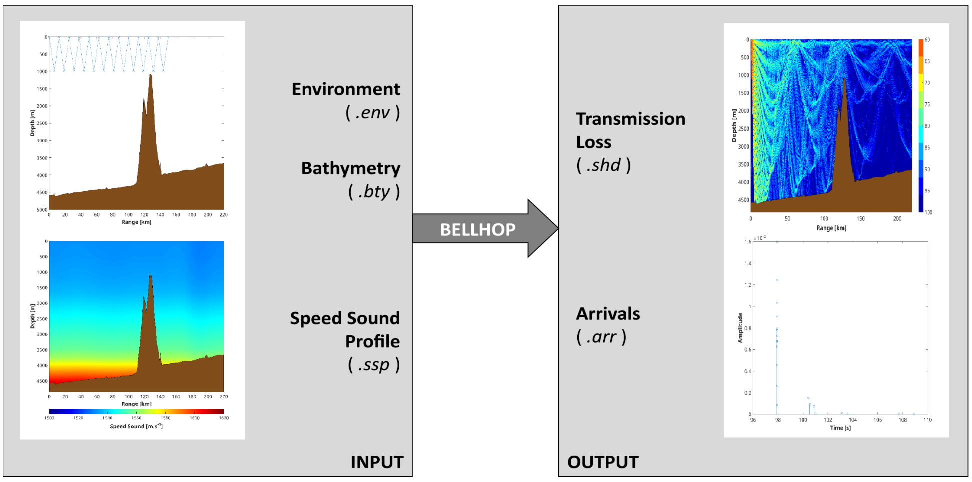

Fig. 3. BELLHOP model structure.

2.2.2. Channel design

The acoustic channel is an important factor when designing underwater acoustic detection systems, as it plays an important role in signal propagation from the source. The acoustic channel affects how signals travel through the water, influencing amplitude, phase and frequency due to interactions with environmental parameters such as bathymetry, SSP and acoustic layer distribution. For the simulation of the Tropic Seamount region, the incorporation of detailed bathymetric data from GEBCO (General Bathymetric Chart of the Ocean) and sound velocity profiles derived from temperature and salinity profiles from HYCOM (HYbrid Coordinate Ocean Model) enables accurate modeling of the environmental parameters of the channel. Considering the source, receiver, boundary properties and distribution of the acoustic layer, the challenge is to theoretically understand the potential impact of this channel between source s and x (Fig.2). This will enable us to predict how the acoustic signal will be modified as it passes through the underwater environment, and to optimize the detection parameters to improve the detector's reliability and accuracy under real-life conditions.

2.3. LF Calls

2.3.1. "AB"-Calls

Atlantic blue whale vocalizations, characterized by long, patterned sequences of very-low-frequency chirps ranging from 15 to 30 Hz (Mellinger and Clark, 2003), exhibit a hierarchical organization into phrases, each composed of one to four distinct unit types. Typically structured as two-part phrases "AB", these sequences feature a monotonic "A" segment around 19 Hz and lasting approximately 8 seconds, followed by a slight frequency-modulated 1 Hz downsweep "B" segment lasting approximately 11 seconds, recurring every 70 seconds (Mellinger and Clark, 2003). Variants may involve repetitions of only the "A" or "B" segment. Additionally, sequences may include a hybrid "AB" unit, called a "C" unit, where the "A" and "B" units are combined without a gap between them. Some phrases conclude with a final unit, "D", a monotonal 9-Hz unit lasting a few seconds. Silent periods between sequences average slightly over 4 minutes (Edds, 1982; Mellinger and Clark, 2003).

Monotonic units are the prevalent components found in every song, often accompanied by downsweeps, while hybrid units frequently initiate phrases (Berchok et al., 2006). The 9-Hz unit commonly follows hybrid units but may also occur after other unit types (Mellinger and Clark, 2003). In our case, we can approximate the "AB" signal as a logistic model for the time-frequency function $f(t)$, as Socheleau et al. (2015) defined by:

$$f(t) = f_c + \frac{1}{2\pi}\frac{d\phi(t)}{dt} = f_c + L + \frac{U-L}{1+e^{\alpha(t-M)}} \tag{1}$$

Here in $Eq.$ 1, f_c represents the central frequency of the signal, between f_{min} and f_{max}. The upper (U) and lower (L) asymptotes of the sigmoid are defined respectively as : $U = f_{max} - f_c$; $L = f_{min} - f_c$. Then, α defined the slope of the downsweep. Finally, parameter M represents the time shift (s) associated with f_c, when downsweep occurs. In the case of a more accurate "AB" signal description, α is set to 0.5 and M is given by $\frac{4}{5}$ of the signal duration.

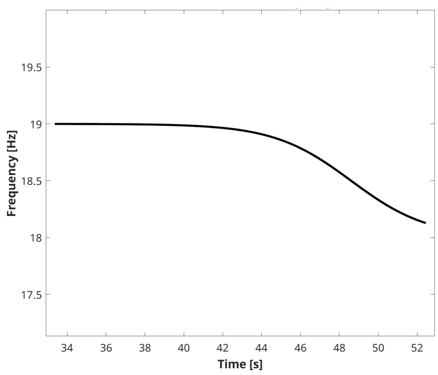

Fig. 4. "AB"-Call waveform in time frequency.

2.3.2. Downsweeps

Downsweeps (DS) chirps are a prominent feature in the vocalizations of fin, sei whales. Atlantic fin whale calls exhibit two typical downward sweeping: ”20-Hz” downsweep (DS-20), from 28 to 15 Hz and ”40-Hz” downsweep (DS-40), from 75 to 40 Hz, both lasting approximately 1 s and may occur as single events, short irregular series, or long regular series lasting over 24 hours (Watkins et al., 1987; Castellote et al., 2012; Romagosa et al., 2021). Similarly, Atlantic sei whale vocalizations feature downsweeps (DS-Sei) lasting around 1.4 seconds, spanning from 82 to 34 Hz on average, extending from an average maximum frequency of 90 to 105 Hz down to an average minimum of 35 to 42 Hz over 1.6 seconds, with a peak frequency occurring between 65 and 70 Hz. Recorded sei whale calls encompass a frequency range from 30 Hz to 129.4 Hz, often observed in pairs but also occurring as triplets or single calls (Baumgartner and Fratantoni, 2008; Español-Jiménez et al., 2019). Thus, downsweeps can be described as a hyperbolic decay, as indicated by $Eq.$ 2, in which f_{ini} represents the beginning frequency of the signal and f_{end} the ending frequency:

$$f(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} = \frac{f_{ini} f_{end}}{f_{end} + \left(f_{ini} - f_{end}\right) t} \tag{2}$$

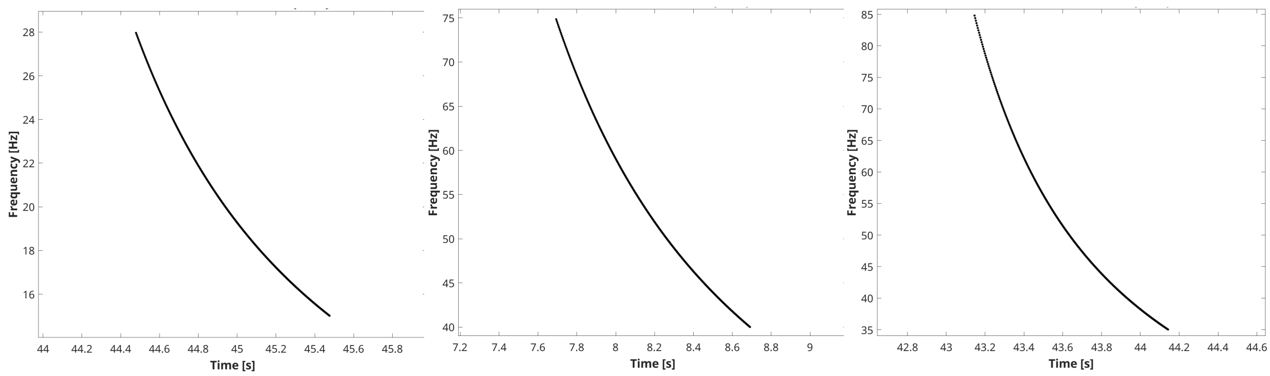

Fig. 5. Waveform in time-frequency of 3 downsweeps: 20-Hz (**left**), 40-Hz (**middle**) and sei whale (**right**).

2.3.3. ”D”-Call Downsweeps

Downsweeps chirps characterized as ”D”-Call is emitted by blue whales with a wide frequency range from 80 to 30 Hz and durations spanning between 0.6 to 3s (Schall et al., 2019; Romagosa et al., 2020). In our study, we approximate the ”D”-Call downsweep similarly as $Eq.$ 1, in which $\alpha = 2$, but the sequence is focused on the downsweep without considering the monotonic parts with signal duration. Monotonic segments are present as the duration increases, but in the case of the ”D” downsweep of a relatively short duration, these segments are not included.

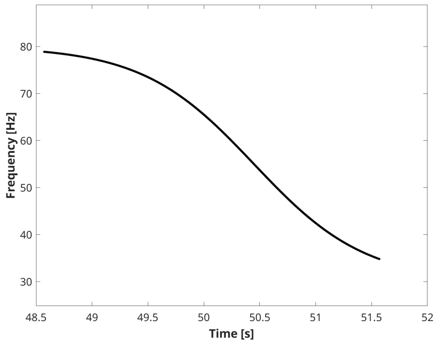

Fig. 6. "D"-Call waveform in time frequency.

In regions where blue whales coexist with fin or sei whales, their overlapping downsweep calls make acoustic identification challenging. This overlap can lead to misidentification, complicating passive acoustic monitoring and potentially affecting conservation efforts. Accurate species identification is essential for understanding population distributions, ecological interactions, and implementing effective conservation strategies.

2.4. Matched Subspace Detector

2.4.1. Detection problem

A fundamental approach to the analysis of baleen whale chirps in the Tropical Seamounts region is provided by the Matched Subspace Detector (MSD) method (Scharf and Friedlander, 1994). The algorithm consists of the design of a signal subspace using signal patterns. In particular, considering an observation window of N samples, the observation vector $y \in C^N$ is expressed as follows:

$$y = \mu x + w \tag{3}$$

Based on y the noisy observation model ($Eq.$ 3), our detection problem is to determine whether μ is 0 or 1, indicating respectively the absence or presence of signal x. The decision must be made in the presence of Gaussian white noise w.

2.4.2. Subspace design

MSD are generally used when the noisy observation (Eq. 3) includes a signal x which is a linear combination of p modes. In this case, x is represented by $\boldsymbol{x} = \boldsymbol{H\theta}$, where H is a known $N \times p$ matrix and θ is a $p \times 1$ vector containing the coordinates of signal. The p modes parameter represents the number of variations in amplitude and phase of the signal. For each signal type described in previous section, the function iterates over p subspaces, calculating the complex exponential component $e^{i \cdot \phi(\cdot)}$, resulting: $h_l = \left[e^{i\phi(l\lfloor N/p \rfloor)}\, e^{i\phi(l\lfloor N/p \rfloor+1)} \ldots e^{i\phi((l+1)\lfloor N/p \rfloor-1)} \right]^T$ the l-th basis vector of size N/p and $\phi(\cdot)$ the time-varying phase.

Thus, the <H> subspace library of LF calls is defined according to the following nomenclature: $<H_{AB}>$, $<H_{DS-20}>$, $<H_{DS-40}>$, $<H_{DS-Sei}>$ and $<H_D>$. The projection matrices onto $<H>$ is defined as: $\boldsymbol{P_H} = \boldsymbol{H}(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H})^{-1}\boldsymbol{H}^{\mathrm{T}}$ enabling the isolation and analysis of each signal of interest within the subspace detector framework. According to Socheleau et al. (2015), the model described by $\boldsymbol{x} = \boldsymbol{H\theta}$ serves as an approximation of reality, yet it remains valid due to the significant portion of LF call energy concentrated within the subspace $<H>$. Essentially, when projecting signal x onto subspace <H>, denoted as $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}_H\boldsymbol{x}$, a considerable portion of its total energy $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$ is captured, particularly when p, the dimension of the subspace, is adequately chosen. In our simulated case, with no real whale call data, p is set to 1, simplifying our <H> subspace to a single dimension. This choice, based on simulated signals, avoids accounting for all signal variations, easing calculations for the projection matrices P_H. While real whale calls vary in amplitude, frequency, and phase, using a single mode here allows us to focus on the theoretical evaluation of our detector's performance without the complexity of real signal variations. It is important to note that while the energy ratio increases with larger values of p, the ability of the model to distinguish between different signals decreases. In the extreme case, where $p = N$, any random signal of size N will satisfy $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{P}_H\boldsymbol{x} = \boldsymbol{s}^{\mathrm{T}}\boldsymbol{x}$.

2.4.3. Detection statistic test

Detectors can be affected by background noise, interference and mismatched models, leading to errors or false alarms. The initial detection problem is defined in terms of a constant false alarm rate statistic T, compared with a threshold η for detecting LF call chirps. However, it is impossible to use signal x in the real case where there are only noisy observations y recorded by receivers, as defined by the model (Eq. 3). This statistic is given by the following relation:

$$T(y) = \frac{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}_H\boldsymbol{y}}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{P}_H^{\perp}\boldsymbol{y}} \geq \eta \iff \frac{\|\boldsymbol{P}_H^{\mathrm{T}}\boldsymbol{y}\|^2}{\|\boldsymbol{P}_H^{\perp}\boldsymbol{y}\|^2} \geq \eta \tag{4}$$

Where $\boldsymbol{P}_H^{\perp} = (\boldsymbol{I} - \boldsymbol{P}_H)$ is the orthogonal projection matrix. The T statistic measures the ratio of the energy of the noisy observation vector y in the whale-call subspace <H> to the energy in the orthogonal noise <$H^{\perp}$>. In our detection framework, the test statistic $T(y)$ follows a Fisher distribution $F_{2p+2,N-2p-2}$. To set the detection threshold, we consider the probability of false alarm P_{FA} and the noise variance σ^2. Thus, η is given by:

$$\eta(\sigma^2, P_{FA}) = \sigma^2 \frac{\chi^2_{\mathrm{inv}}(1 - P_{FA}, 2p+2)}{\chi^2_{\mathrm{inv}}(1 - P_{FA}, N-2p-2)} \tag{5}$$

If $T(y) \geq \eta$, the MSD concludes that $\mu = 1$ with presence of the interest signal; otherwise, it concludes that $\mu = 0$. In a simulated environment, σ^2 can be easily defined. In the real case, the true value cannot be known, so it is necessary to estimate it, in this case with P-percentile method. The method is defined by the statistic P, generally $10 \leq P \leq 20\%$, to ensure that we avoid including the signal component, which is expected to have higher energy than the noise. It is important to rank $|y|^2$ in order, as the lowest values are, by hypothesis, only characterized by noise from which the first P values are taken to estimate the noise variance. The P-percentile method is more accurate when the N samples of y are sufficiently large (Jesus, 2019).

3. Simulated data results

3.1. BELLHOP Acoustic channel

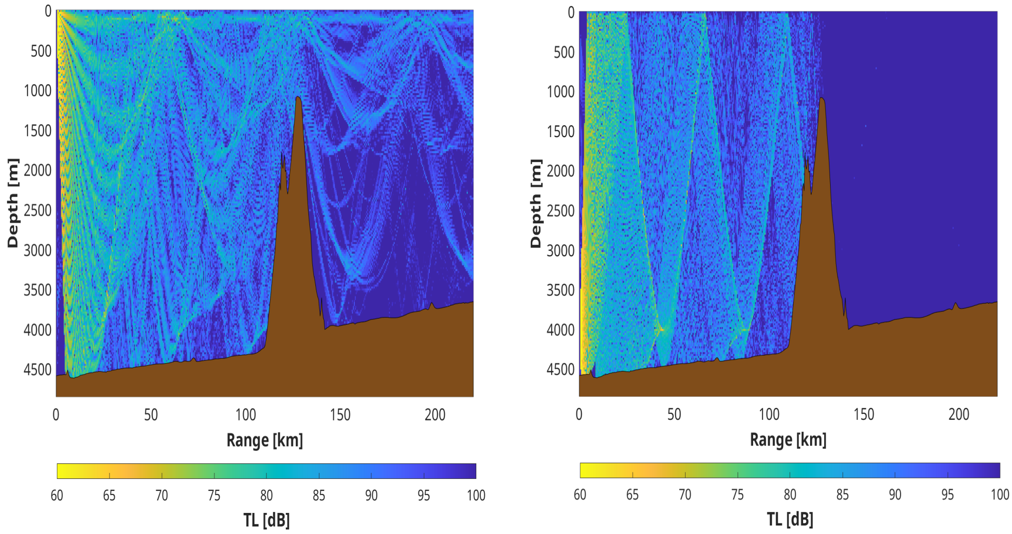

Fig. 7. Acoustic channel in the Tropic Seamount region modeled with BELLHOP (latitudinal section at 23.9°N): simulation of transmission loss due to environmental factors with a source at 100 m (**left**) and at 4000 m (**right**).

In an ocean acoustic channel, factors such as bathymetry, SSP, attenuation, source depth and propagation distance significantly influence signal transmission. The bathymetry affects the trajectories of acoustic rays through reflections and refractions, while the SSP causes depth-dependent variations in sound velocity, bending the rays. As a result of absorption and scattering, attenuation reduces signal intensity over distance. Source depth and transmission angle affect reflections on the surface and seabed, and longer distances increase signal loss. For deep sources, acoustic rays may be absorbed or reflected by underwater reliefs (Fig.7), limiting coverage. For shallow sources (0-100 m), rays can travel further, reducing attenuation. These factors affect signal properties such as amplitude, phase and frequency, which in turn affect detection and analysis, and also affects the time arrivals of the rays received, depending on their direct or non-direct path (Fig.8).

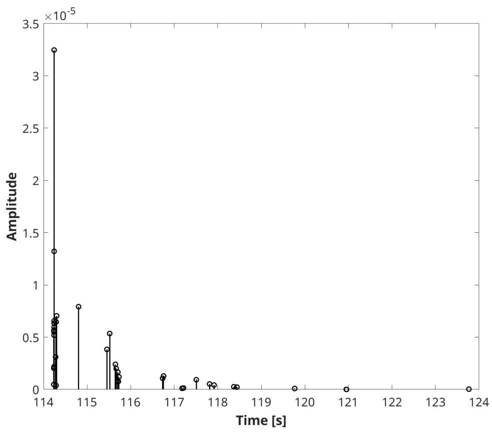

Fig. 8. Time arrivals of signal from a source (at 100m), calculated with BELLHOP, received by receiver (at 175 km from the source): the peaks correspond to an arrival, so overlapping is possible.

3.2. Theoretical detector performance

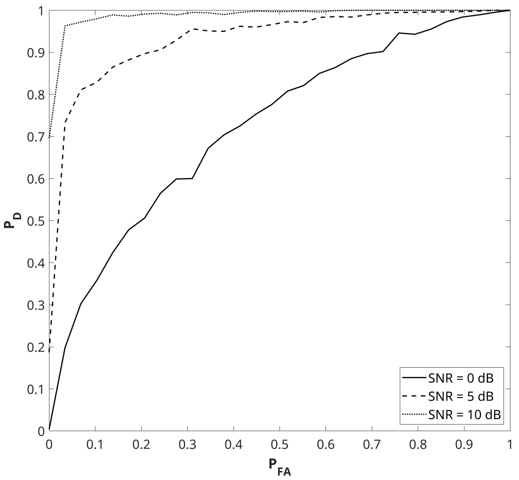

Fig. 9. Probability of detection (P_D) versus probability of false alarm (P_{FA}) for the MSD detection of the "D"-Call signal.

In a signal detection context, an experimental assessment of the performance of the detector is fundamental to optimize its ability to perform in signal analysis. The receiver operating characteristic (ROC) curve is used to measure the sensitivity of the detector to discriminate between signal and noise. Our ROC analysis for the MSD gives an overview of the performance of the detector under different signal-noise ratios (SNR), as following:

$$SNR = 10 \log \left(\frac{\bar{x}^2}{\sigma^2} \right) \tag{6}$$

Here, Eq. 6 describes signal quality relative to noise using a logarithmic scale in decibels (dB), comparing mean signal power $\bar{x}^2$ to σ^2. High SNR indicates better signal quality. At 0 dB, the ROC curve is the lowest (Fig.9), indicating difficulty in detecting weak signals (Fig.11). At 10 dB SNR, the ROC curve is the highest, showing effective detection of strong signals with minimal false alarms.

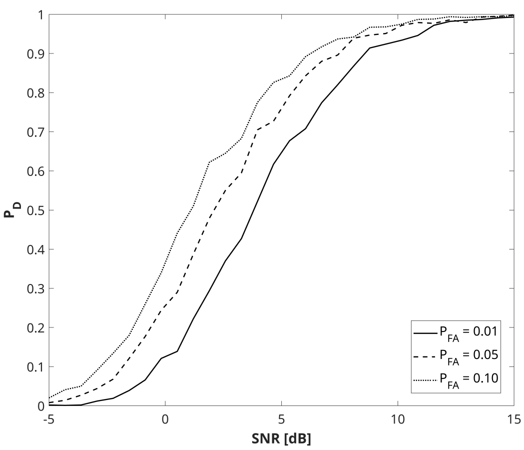

Fig. 10. Detection probability (P_D) as a function of SNR for the MSD detection of the "D"-Call signal.

Additionally, as P_{FA} increases, our MSD becomes less strict, detecting weaker signals at the cost of more false alarms (Fig.10). For a low P_{FA} of 0.01, the detector needs a higher SNR to maintain a high P_D, indicating high reliability but lower sensitivity to weak signals. For a moderate P_{FA} of 0.05, the detector balances sensitivity and reliability, requiring a moderately high SNR for reasonable P_D. For a high P_{FA} of 0.1, the detector is highly sensitive, detecting signals at lower SNR, but this increases the false alarm rate.

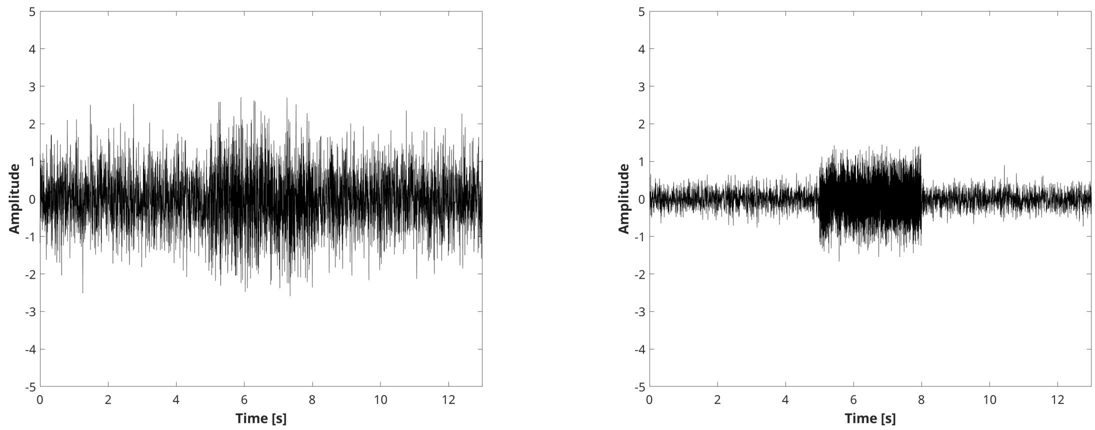

Fig. 11. "D"-Call signal with noise as function of time: SNR at 0 dB (**left**) and at 10 dB (**right**).

3.2. Discriminatory performance

Our subspace detector performs best in high SNR environments and can be adjusted for higher detection rates by accepting more false alarms. Ongoing tests evaluate the MSD's ability to discriminate LF calls from the library (Fig.12), using a fixed SNR and P_{FA}. However, the discrimination results, for each whale call with all P_H matrices, indicate that all signals are perfectly detected. A discrimination problem remains between each signal. The hypothesis is proposed to assess the issue: the projection matrices are not sufficiently different to be discriminating. Thus, to test the hypothesis, we use the Grassmann distance method to have the degree of similarity of the P_H matrices, as following:

$$\mathrm{dist}(P_{H_a}, P_{H_b}) = \sqrt{\sum_{n=1}^{k} \theta_i^2} \tag{7}$$

Where θ_i are the principal angles based on the singular values of the SVD (Singular Value decomposition) for the two projection matrices P_{H_a} and P_{H_b}, while k is the number of singular values.

Tab. 1. Discrimination performance of P_H matrices: Grassmann distance for pairwise comparison of the similarity distance of P_H matrices.

	$\boldsymbol{P}_{AB}$	$\boldsymbol{P}_{D}$	$\boldsymbol{P}_{DS-20}$	$\boldsymbol{P}_{DS-40}$	$\boldsymbol{P}_{DS-Sei}$
$\boldsymbol{P}_{AB}$	0	118.592	118.592	118.592	118.591
$\boldsymbol{P}_{D}$		0	47.124	47.124	47.123
$\boldsymbol{P}_{DS-20}$			0	27.162	27.176
$\boldsymbol{P}_{DS-40}$				0	27.175
$\boldsymbol{P}_{DS-Sei}$					0

Here, as shown Tab.1, P_{AB} has a significantly high value (118) compared with the other matrices, indicating high dissimilarity and therefore meaning that it is distinct from the others. Likewise, the P_D matrix indicates a more moderate dissimilarity (47) with P_{DS-20}, P_{DS-40} and P_{DS-Sei}. This score indicates a moderate distinction between the different downsweeps of the blue whale and those of the fin and sei whale. Finally, the distinction between each DS with P_{DS-20}, P_{DS-40} and P_{DS-Sei} is relatively poor (27), indicating a slight similarity between these different fin and sei whale calls. The result is consistent with the approximation of the mathematical formulation of DS ($Eq.$ 2), which is similar between these 3 calls.

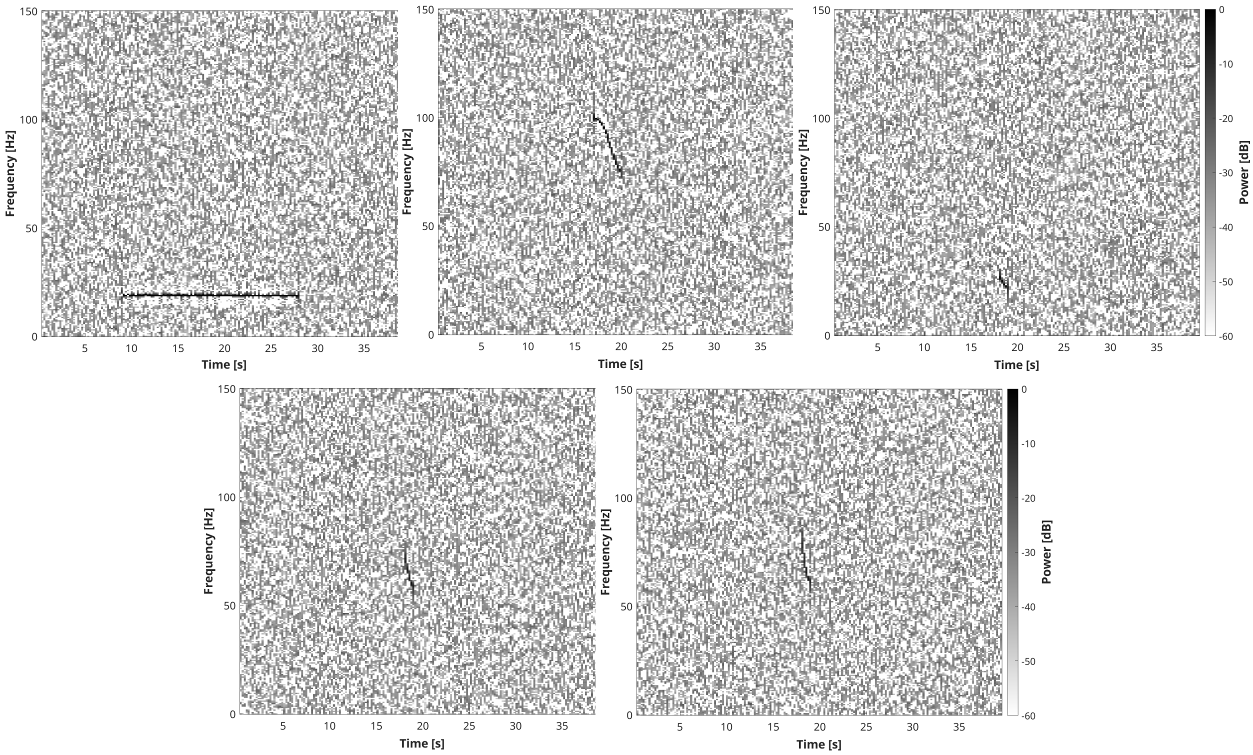

Fig. 12. Spectrogram of 5 LF calls SNR at 10 dB, Kaiser window (128, 18), FFT length (256 points), window overlap (75%): "AB"-Call (**top left**), "D"-Call (**top middle**), "DS-20" Call (**top right**), "DS-40" Call (**bottom left**) and "DS-Sei" Call (**bottom right**).

4. Discussion

The MSD detector algorithm, based on the subspace method, is implemented to detect and identify different LF calls of baleen whales from a library of Mid-Atlantic species, described in literature (Mellinger and Clark, 2003; Schall et al., 2019; Romagosa et al., 2021). By working in a fully simulated environment, we can control all parameters influencing the detector, enabling us to optimize it realistically for experimental detection scenarios. The theoretical performance of the MSD highlights its efficiency under optimal noise conditions, becoming more effective with high SNR ratios (Fig.9-10). Distinguishing signals in a noisy environment is challenging. Theoretically, by increasing the SNR, the signal from the whale is more distinct from the noise, and therefore the discrimination is significantly greater. In our case, the results indicate an absence of discrimination for each subspace, independent of the value of SNR and P_{FA}. According to the degrees of similarity between each pair of P_H matrices (Tab.1), it is possible to discriminate the signals on the basis of this factor alone. This poor discrimination can be explained hypothetically by the fact that the η detection threshold is ill-suited to the detection of several different signals. For each noisy signal containing a different whale sound, to which the variance of the noise is estimated, this threshold is approximately identical between each signal, and therefore doesn't allow the whale sounds to be dissociated. Thus, it is necessary to further estimate an adaptive detection threshold for each whale signal, in order to complement the discriminant qualities of the P_H matrices. Meanwhile, the degree of similarity of P_H shows that the approximation of the downsweeps "DS-20", "DS-40" and "DS-Sei" with the same equation ($Eq.$ 2) on relatively close frequency ranges, will also be a problem to be managed *a posteriori* of the adaptation of detection thresholds, to avoid the confusion of these signals produced by fin and sei whales. In addition, Socheleau et al. (2015) addressed interferences in their work, recognizing that signals similar to whale calls can disrupt detection performance in subspaces. In our study, we opted to simplify the problem by focusing solely on an observation y that includes the LF calls and noise. This simplification is crucial for evaluating performance in basic scenarios. Future work will build on this foundation by considering more complex cases, including external sources that interfere with the signals of interest.

The effectiveness of the MSD detector depends on a number of factors, including noise and the complex nature of the acoustic environment. By modeling the acoustic channel using BELLHOP, a comprehensive examination of environmental variables such as bathymetry, SSP, attenuation, source depth and propagation distance is performed, enabling us to understand their collective influence on signal transmission (Fig.7-8). These variables modulate signal properties. The amplitude is reduced due to attenuation, particularly at HF. The signal may also show a frequency shift due to the Doppler effect, and a dispersion that extends the spectrum. Multiple reflections cause phase shifts and reverberations, altering the original temporal structure. In addition, ambient background noise is added to the signal, reducing the *SNR* and making certain parts of the signal less distinct. These combined impacts modify the initial characteristics of the signal, significantly affecting detection and analysis performance (Stojanovic and Preisig, 2009). However, it is imperative to note that performance in real acoustic environments has yet to be tested, which represents a critical step in the final evaluation. In our study, the absence of testing for the $x = s$ signal, implies that its representation may not fully capture post-acoustic channel signal behavior. Furthermore, the inherent complexity of realistic marine environments suggests potential challenges for the MSD detector compared to fully simulated conditions. In the real case of an ocean acoustic channel, environmental parameters significantly control signal transmission dynamics. The simulations performed by BELLHOP are an approximation, to which the inputs and outputs of the model only reflect a vision of the reality of the simulated region and therefore of the acoustic channel. This holistic understanding underlines the importance of evaluating the performance of the MSD method in real acoustic channel scenarios with at-sea trials, recognizing the complex interplay between environmental factors and signal features.

5. Conclusions

The MSD detector algorithm, based on the subspace method, demonstrated theoretical efficiency in detecting and identifying various LF calls of baleen whales from a Mid-Atlantic species library. The controlled, fully simulated environment allowed for realistic optimization, highlighting the detector's effectiveness under optimal noise conditions, particularly with high SNR ratios. However, in noisy environments, the current detection threshold proved inadequate for multiple signals, necessitating adaptive thresholds for better discrimination. Similar downsweeps "DS-20", "DS-40", and "DS-Sei" require refined detection strategies to avoid confusion. External interferences can disrupt detection, emphasizing the need for more complex scenario testing. Acoustic channel modeling with BELLHOP provided insights into how environmental variables like bathymetry, SSP, attenuation, source depth, and propagation distance influence signal transmission, affecting amplitude, frequency, dispersion, phase shifts, and background noise. While BELLHOP simulations offer a close approximation, real-world trials are essential to evaluate MSD performance in actual marine environments. This approach will enhance understanding of baleen whale behavior in the Tropic Seamount region, aiding conservation and sustainable ocean management.

References

Aulich, M. G., McCauley, R. D., Saunders, B. J., and Parsons, M. J. (2019). Fin whale (balaenoptera physalus) migration in australian waters using passive acoustic monitoring. *Scientific reports*, 9(1):8840.

Balcazar, N. E., Tripovich, J. S., Klinck, H., Nieukirk, S. L., Mellinger, D. K., Dziak, R. P., and Rogers, T. L. (2015). Calls reveal population structure of blue whales across the southeast indian ocean and the southwest pacific ocean. *Journal of Mammalogy*, 96(6):1184–1193.

Baumgartner, M. F., Bonnell, J., Corkeron, P. J., Van Parijs, S. M., Hotchkin, C., Hodges, B. A., Bort Thornton, J., Mensi, B. L., and Bruner, S. M. (2020). Slocum gliders provide accurate near real-time estimates of baleen whale presence from human-reviewed passive acoustic detection information. *Frontiers in Marine Science*, 7:100.

Baumgartner, M. F. and Fratantoni, D. M. (2008). Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders. *Limnology and Oceanography*, 53(5part2):2197–2209.

Berchok, C. L., Bradley, D. L., and Gabrielson, T. B. (2006). St. lawrence blue whale vocalizations revisited: Characterization of calls detected from 1998 to 2001. *The Journal of the Acoustical Society of America*, 120(4):2340–2354.

Castellote, M., Clark, C. W., and Lammers, M. O. (2012). Acoustic and behavioural changes by fin whales (balaenoptera physalus) in response to shipping and airgun noise. *Biological Conservation*, 147(1):115–122.

Correia, A. M., Sousa-Guedes, D., Gil, Á., Valente, R., Rosso, M., Sousa-Pinto, I., Sillero, N., and Pierce, G. J. (2021). Predicting cetacean distributions in the eastern north atlantic to support marine management. *Frontiers in Marine Science*, 8:643569.

Davis, G. E., Baumgartner, M. F., Corkeron, P. J., Bell, J., Berchok, C., Bonnell, J. M., Bort Thornton, J., Brault, S., Buchanan, G. A., Cholewiak, D. M., et al. (2020). Exploring movement patterns and changing distributions of baleen whales in the western north atlantic using a decade of passive acoustic data. *Global Change Biology*, 26(9):4812–4840.

Edds, P. L. (1982). Vocalizations of the blue whale, balaenoptera musculus, in the st. lawrence river. *Journal of Mammalogy*, 63(2):345–347.

Edds-Walton, P. L. (1997). Acoustic communication signals of mysticete whales. *Bioacoustics*, 8(1-2):47–60.

Español-Jiménez, S., Bahamonde, P. A., Chiang, G., and Häussermann, V. (2019). Discovering sounds in patagonia: Characterizing sei whale (balaenoptera borealis) downsweeps in the south-eastern pacific ocean. *Ocean Science*, 15(1):75–82.

Figueroa, H. and Robbins, M. (2008). Xbat: an open-source extensible platform for bioacoustic research and monitoring. *Computational bioacoustics for assessing biodiversity*, pages 143–155.

Garrigue, C., Clapham, P. J., Geyer, Y., Kennedy, A. S., and Zerbini, A. N. (2015). Satellite tracking reveals novel migratory patterns and the importance of seamounts for endangered south pacific humpback whales. *Royal Society open science*, 2(11):150489.

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

Goldstein, A. and Bentley, S. (2010). Use of highly portable micro-sized remotely operated vehicles for environmental monitoring and mapping. *OCEANS 2010 MTS/IEEE SEATTLE*, pages 1–6.

Harland, E. and Armstrong, M. (2004). The real-time detection of the calls of cetacean species. *Canadian Acoustics*, 32(2):76–82.

Harris, D., Matias, L., Thomas, L., Harwood, J., and Geissler, W. H. (2013). Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast atlantic. *The Journal of the Acoustical Society of America*, 134(5):3522–3535.

Herman, L. M. (1980). The communication systems of cetaceans. *Cetacean Behavior.*, pages 149–209.

Ibrahim, A. K., Zhuang, H., Chérubin, L. M., Schärer-Umpierre, M. T., and Erdol, N. (2018). Automatic classification of grouper species by their sounds using deep neural networks. *The Journal of the Acoustical Society of America*, 144(3):EL196–EL202.

Jesus, S. (2019). A simple detector for passive acoustic monitoring of ocean.

Knight, E. C., Hannah, K. C., Foley, G. J., Scott, C. D., Brigham, R. M., and Bayne, E. (2017). Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. *ACE*, 12(2):14.

Kraut, S., Scharf, L. L., and McWhorter, L. T. (2001). Adaptive subspace detectors. *IEEE Transactions on signal processing*, 49(1):1–16.

Leroy, E. C., Samaran, F., Bonnel, J., and Royer, J.-Y. (2016). Seasonal and diel vocalization patterns of antarctic blue whale (balaenoptera musculus intermedia) in the southern indian ocean: a multi-year and multi-site study. *PloS one*, 11(11):e0163587.

Lewis, L. A., Calambokidis, J., Stimpert, A. K., Fahlbusch, J., Friedlaender, A. S., McKenna, M. F., Mesnick, S. L., Oleson, E. M., Southall, B. L., Szesciorka, A. R., et al. (2018). Context-dependent variability in blue whale acoustic behaviour. *Royal Society open science*, 5(8):180241.

Marques, T. A., Thomas, L., Martin, S. W., Mellinger, D. K., Ward, J. A., Moretti, D. J., Harris, D., and Tyack, P. L. (2013). Estimating animal population density using passive acoustics. *Biological reviews*, 88(2):287–309.

McDonald, M. A., Mesnick, S. L., and Hildebrand, J. A. (2006). Biogeographic characterization of blue whale song worldwide: using song to identify populations. *Journal of cetacean research and management*, 8(1):55–65.

Mellinger, D. and Bradbury, J. (2007). Acoustic measurement of marine mammal sounds in noisy environments. In *Proceedings of the International Conference on underwater acoustical measurements: Technologies and results*, pages 273–280.

Mellinger, D. K. and Clark, C. W. (2000). Recognizing transient low-frequency whale sounds by spectrogram correlation. *The Journal of the Acoustical Society of America*, 107(6):3518–3529.

Mellinger, D. K. and Clark, C. W. (2003). Blue whale (balaenoptera musculus) sounds from the north atlantic. *The Journal of the Acoustical Society of America*, 114(2):1108–1119.

Merchant, N. D., Pirotta, E., Barton, T. R., and Thompson, P. M. (2014). Monitoring ship noise to assess the impact of coastal developments on marine mammals. *Marine Pollution Bulletin*, 78(1-2):85–95.

Miller, B. S., 15, I.-S. A. T. W. G. M. B. S. . S. K. M. . V. O. I. . H. D. . S. F. . Š. A. . B. S. . F. K. ., Balcazar, N., Nieukirk, S., Leroy, E. C., Aulich, M., Shabangu, F. W., Dziak, R. P., Lee, W. S., and Hong, J. K. (2021). An open access dataset for developing automated detectors of antarctic baleen whale sounds and performance evaluation of two commonly used detectors. *Scientific Reports*, 11(1):806.

Morano, J. L., Salisbury, D. P., Rice, A. N., Conklin, K. L., Falk, K. L., and Clark, C. W. (2012). Seasonal and geographical patterns of fin whale song in the western north atlantic ocean. *The Journal of the Acoustical Society of America*, 132(2):1207–1212.

Morato, T., Varkey, D. A., Damaso, C., Machete, M., Santos, M., Prieto, R., Santos, R. S., and Pitcher, T. J. (2008). Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series*, 357:23–32.

Nowacek, D. P., Thorne, L. H., Johnston, D. W., and Tyack, P. L. (2007). Responses of cetaceans to anthropogenic noise. *Mammal Review*, 37(2):81–115.

Pitcher, T. J., Morato, T., Hart, P. J., Clark, M. R., Haggan, N., Santos, R. S., et al. (2007). *Seamounts: ecology, fisheries & conservation*, volume 12. Wiley Online Library.

Porter, M. B. (2011). The bellhop manual and user's guide: Preliminary draft. *Heat, Light, and Sound Research, Inc., La Jolla, CA, USA, Tech. Rep*, 260.

Rasmussen, J. H. and Širović, A. (2021). Automatic detection and classification of baleen whale social calls using convolutional neural networks. *The Journal of the Acoustical Society of America*, 149(5):3635–3644.

Richardson, W. J., Greene, C. R., Malme, C. I., and Thomson, D. H. (1995). Marine mammal hearing. *Marine mammals and noise*, pages 205–240.

Romagosa, M., Baumgartner, M., Cascão, I., Lammers, M. O., Marques, T. A., Santos, R. S., and Silva, M. A. (2020). Baleen whale acoustic presence and behaviour at a mid-atlantic migratory habitat, the azores archipelago. *Scientific Reports*, 10(1):4766.

Romagosa, M., Pérez-Jorge, S., Cascão, I., Mouriño, H., Lehodey, P., Pereira, A., Marques, T. A., Matias, L., and Silva, M. A. (2021). Food talk: 40-hz fin whale calls are associated with prey biomass. *Proceedings of the Royal Society B*, 288(1954):20211156.

Sayigh, L. S. (2013). Cetacean acoustic communication. *Biocommunication of animals*, pages 275–297.

Schall, E., Di Iorio, L., Berchok, C., Filún, D., Bedriñana-Romano, L., Buchan, S. J., Van Opzeeland, I., Sears, R., and Hucke-Gaete, R. (2019). Visual and passive acoustic observations of blue whale trios from two distinct populations. *Marine Mammal Science*.

Scharf, L. L. and Friedlander, B. (1994). Matched subspace detectors. *IEEE Transactions on signal processing*, 42(8):2146–2157.

Širović, A., Bassett, H. R., Johnson, S. C., Wiggins, S. M., and Hildebrand, J. A. (2014). Bryde's whale calls recorded in the gulf of mexico. *Marine Mammal Science*, 30(1):399–409.

Socheleau, F.-X., Leroy, E., Carvallo Pecci, A., Samaran, F., Bonnel, J., and Royer, J.-Y. (2015). Automated detection of antarctic blue whale calls. *The Journal of the Acoustical Society of America*, 138(5):3105–3117.

Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information processing & management*, 45(4):427–437.

Stafford, K. M., Fox, C. G., and Clark, D. S. (1998). Long-range acoustic detection and localization of blue whale calls in the northeast pacific ocean. *The Journal of the Acoustical Society of America*, 104(6):3616–3625.

Stafford, K. M., Nieukirk, S. L., Cox, C. G., et al. (2001). Geographic and seasonal variation of blue whale calls in the north pacific. *J. Cetacean Res. Manage.*, 3(1):65–76.

Stojanovic, M. and Preisig, J. (2009). Underwater acoustic communication channels: Propagation models and statistical characterization. *IEEE communications magazine*, 47(1):84–89.

Thompson, P. O., Findley, L. T., and Vidal, O. (1992). 20-hz pulses and other vocalizations of fin whales, b alaenoptera physalus, in the gulf of california, mexico. *The Journal of the Acoustical Society of America*, 92(6):3051–3057.

Usman, A. M., Ogundile, O. O., and Versfeld, D. J. (2020). Review of automatic detection and classification techniques for cetacean vocalization. *Ieee Access*, 8:105181–105206.

Valente, R., Correia, A. M., Gil, Á., Gonzalez Garcia, L., and Sousa-Pinto, I. (2019). Baleen whales in macaronesia: occurrence patterns revealed through a bibliographic review. *Mammal review*, 49(2):129–151.

Watkins, W. A., Tyack, P., Moore, K. E., and Bird, J. E. (1987). The 20-hz signals of finback whales (b alaenoptera physalus). *The Journal of the Acoustical Society of America*, 82(6):1901–1912.

Weilgart, L. S. (2007). A brief review of known effects of noise on marine mammals. *International Journal of Comparative Psychology*, 20(2).

Yen, P. P., Sydeman, W. J., and Hyrenbach, K. D. (2004). Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. *Journal of Marine systems*, 50(1-2):79–99.

Zimmer, W. M. (2011). *Passive acoustic monitoring of cetaceans*. Cambridge University Press.

Résumé

Les activités humaines affectent de plus en plus les écosystèmes marins. Il est essentiel de comprendre leurs effets sur la vie marine, en particulier sur les cétacés, pour en assurer la conservation. Les baleines à fanons s'appuient fortement sur des appels à basse fréquence pour communiquer et naviguer, chacun ayant des caractéristiques distinctes. En utilisant des techniques avancées de modélisation acoustique telles que BELLHOP, cette étude représente avec précision l'environnement acoustique sous-marin de la région du mont sous-marin Tropic. En utilisant le Matched Subspace Detector (MSD), l'étude vise à isoler les appels de baleines du bruit ambiant, mais rencontre des difficultés pour discriminer efficacement les catégories d'appels (« AB », « D », et « DS »). Si les similitudes entre les formes d'onde contribuent à ces difficultés, un seuil de détection insuffisant pour chaque signal peut également jouer un rôle. Les recherches futures devraient se concentrer sur l'adaptation des seuils de détection, en tenant compte des effets de la modélisation du canal acoustique par BELLHOP. Malgré ses limites, cette étude jette les bases d'une recherche plus approfondie sur le comportement acoustique des baleines à fanons pendant la migration et ses implications pour les efforts de conservation marine.

Mots clés : *Appels de Baleines, Basses-Fréquences, Environnement simulé, Détecteur par Sous-espaces*

Abstract

As human activities increasingly impact marine ecosystems, understanding their effects on marine life, particularly cetaceans, is vital for conservation. Baleen whales rely heavily on low-frequency calls for communication and navigation, each with distinct characteristics. Using advanced acoustic modeling techniques such as BELLHOP, this study accurately represents the underwater acoustic environment of the Tropic Seamount region. Employing the Matched Subspace Detector (MSD), the study aims to isolate whale calls from ambient noise, but encounters challenges in effectively discriminating between call categories ("AB", "D", and "DS"). While waveform similarities contribute to these challenges, an insufficient detection threshold for each signal may also play a role. Future research should focus on adapting detection thresholds, taking into account the effects of acoustic channel modeling by BELLHOP. Despite limitations, this study lays the groundwork for further investigation into the acoustic behavior of baleen whales during migration and its implications for marine conservation efforts.

Keywords: *Baleen Calls, Low-Frequency, Simulated Environment, Matched Subspace Detector*