Detection of boats and fin whales using passive acoustic techniques

Alexandra Constaratas

ENGEES
ÉCOLE NATIONALE DU GÉNIE DE L'EAU
ET DE L'ENVIRONNEMENT DE STRASBOURG

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

NIWA
Taihoro Nukurangi

Graduation thesis

Detection of boats and fin whales using passive acoustic techniques

---

National Institute of Water and Atmospheric Research
Wellington, New Zealand

Alexandra CONSTARATAS
Graduating class: Freiburg, 3rd year specialization: hydrosystems
Thesis submitted to obtain a double degree: ENGEES engineer and Master of Marine Sciences, OSU Pytheas Institute

Internship carried out from January 14 to July 15, 2019

Photographs of the cover page (from left to right):

Aliens. *New Zealand Geographic* [online]. [accessed on 27/01/2019]. https://www.nzgeo.com/stories/aliens/

Why fin whales won't be targeted off Iceland this year. *Whale and Dolphin Conservation* [online]. [accessed on 15/02/2019]. https://uk.whales.org/blog/2017/03/why-fin-whales-wont-be-targeted-off-iceland-this-year

I really would like to thank my internship supervisor, Dr Giacomo Giorli, for his patience, availability and advice through my internship project

Détection des bateaux et des rorquals communs en utilisant des techniques d'acoustique passive

Résumé

Les océans représentent le plus grand écosystème de la planète. Surveiller des écosystèmes aussi vastes est un défi, car les chercheurs ne peuvent atteindre facilement (et à moindre coût) certaines régions éloignées. L'Aire Marine Protégée (AMP) de la mer de Ross, créée en décembre 2017, en est un exemple. Les techniques d'acoustique passive représentent une solution économique pour la détection dans les régions difficiles d'accès des océans. Cependant, les données acoustiques contiennent des signaux audios de différentes sources sonores. Les chercheurs doivent donc être en mesure de détecter les signaux ciblés dans les données acoustiques.

Nous développons ici des algorithmes de traitement du signal pour traiter trois tâches dans l'AMP de la mer de Ross et dans la région du détroit de Cook : la détection des bateaux, la mesure de leurs fréquences et la détection des appels de type 'doublet' des rorquals communs. Nous générons enfin des graphiques d'occurrence temporelle des bateaux et des rorquals communs dans les zones étudiées. Les bateaux ont été fortement détectés dans la mer de Ross, mais leur occurrence diffère en fonction de la station d'étude, ce qui semble lié à la localisation spatiale de la station. Les rorquals communs ont été détectés en juin 2017 dans la partie Est du détroit de Cook. Cependant, aucun rorqual commun n'a été détecté dans la mer de Ross.

Detection of boats and fin whales using passive acoustic techniques

Abstract

Oceans represent the largest ecosystem on Earth. Monitoring such large ecosystems is challenging because researchers cannot easily (and affordably) reach some remote areas. An example is the Ross Sea Marine Protected Area (MPA), established in December 2017. Passive acoustic techniques represent a cost-effective solution to sense hard-to-access regions of the Oceans. However, acoustic data contains audio signals from various sound sources, and researchers must be able to detect the target signals in the acoustic data.

Here we develop signal processing algorithms to address three tasks in the Ross Sea MPA and in the Cook Strait region: the detection of passing boats, the measurement of their tonal sounds and the detection of fin whale ‘doublet’ calls. We finally generate graphs of temporal occurrence of boats and fin whales in the study areas. Boats have been highly detected in the Ross Sea, but their occurrence differ depending on the station of study, which seems to be linked to the spatial location of the station. Fin whales have been detected in June 2017 in the East part of the Cook Strait. However, no fin whale has been detected in the Ross Sea.

Contents

Table of Figures

Table of Tables

Table of Abbreviations

AMAR: Autonomous Multichannel Acoustic Recorders

FN: False Negative

FP: False Positive

MPA: Marine Protected Area

PSD: Power Spectral Density

ROC: Receiver Operating Characteristics

TN: True Negative

TP: True Positive

1. Introduction

Passive acoustics is a cost-effective technique to monitor large marine ecosystems. By recording underwater signals, it is possible to detect passing ships (Sorensen *et al.*, 2010). The detection of boat tonal sounds is useful to identify and classify the type of transiting vessels which are passing (Mckenna *et al.*, 2012) and to discriminate illegal from legal fishing (Abileah and Lewis, 1996). The use of passive acoustics is also useful to make a species inventory and even to calculate biodiversity index (Parks *et al.*, 2014) by knowing population densities, for example for fin whales (McDonald and Fox, 1999).

Fin whale populations (*Balaenoptera physalus*, Linnaeus, 1758) are at very low levels in the Southern Ocean (Branch and Butterworth, 2001). This is due to commercial exploitation during the 20th century (Clapham and Baker, 2001). Therefore, visual and acoustic surveys have been conducted to assess the occurrence of marine mammals near Antarctica, including fin whales (Gedamke and Robinson, 2010). Fin whales are spotted in tropical/subtropical waters during winter for breeding and polar/subpolar waters during summer for feeding, with migrations that are often between 2,000 km to 5,000 km long. Indeed, cold latitudes offer them rich food, such as krill (Norris, 1977). Fin whales occur near Antarctica between February and June (Širović *et al.*, 2004). However, their presence in the Ross Sea is not well known (Širović *et al.*, 2009). Signals emitted by fin whales are contained in a low-frequency range (or bandwidth); the ‘doublet 20 Hz’, the ‘20- to 35- Hz irregular repetition interval’, and the ‘30- to 90- Hz shorter and more irregular repetition intervals’ (McDonald and Fox, 1999). Since these signals have low frequencies, they can travel long distances in sea water; a characteristic that makes them suitable for long-range passive acoustic monitoring (Širović *et al.*, 2007). The objectives of this thesis were to develop signal processing techniques to study the occurrence of boats and fin whales in the Ross Sea and Cook Strait. This study also aimed at understanding whether a potential fin whale migration between the Cook Strait and the Ross Sea could be studied acoustically.

2. Material and methods

2.1. Data collection

2.1.1. Instruments

The acoustic data were collected using Autonomous Multichannel Acoustic Recorders (AMAR; JASCO Applied Sciences) equipped with a M36-V35-100 hydrophone (Geospectrum Technologies Inc.). This instrument is used to measure underwater sound levels, either in shallow waters or deep waters (Figure 1).

Figure 1: Autonomous Multichannel Acoustic Recorder (AMAR). Left: for shallow waters, on a bottom base plate. Right: for deep waters, housed in a glass sphere. *Source:* Giorli and Goetz, 2018

AMARs can be deployed with small hydrophone arrays and can autonomously record signals for a period of a few months to a year. All recordings are saved in the .wav format. To convert the digital data collected by the recorders into acoustic pressure values, it is necessary to take into account the characteristics of the Analog-To-Digital Converter of the recorders (such as the bit depth and the gain) and the sensitivity of the hydrophones. An example of a sensitivity curve for the hydrophones used in this study is reported in Figure 2.

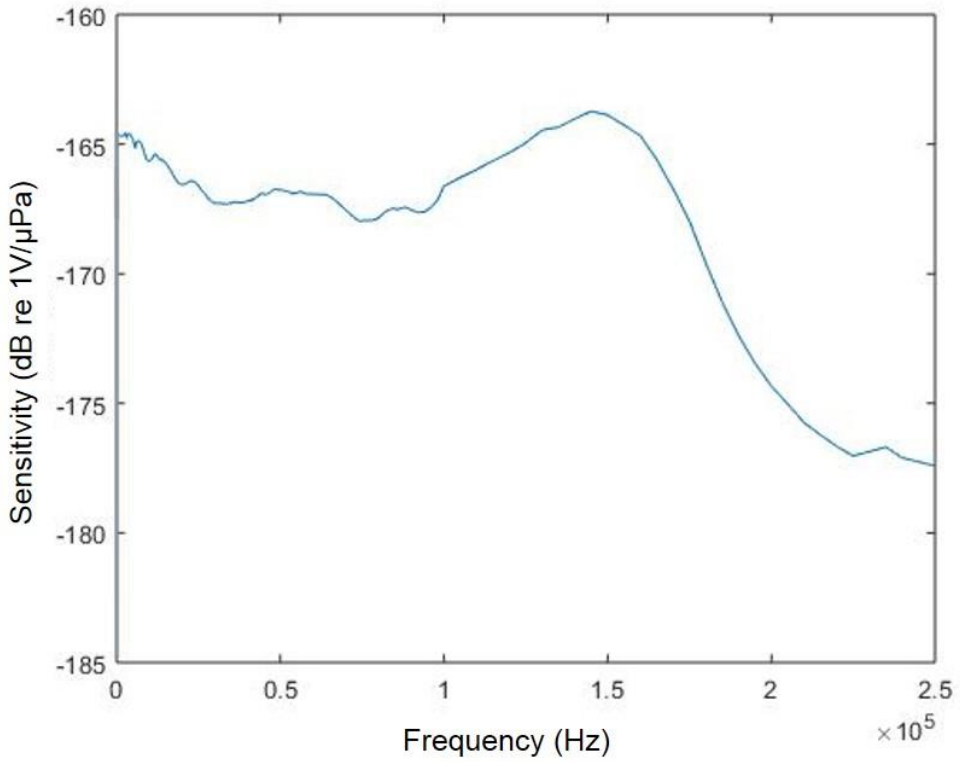

Figure 2: Example of a hydrophone calibration curve showing the sensitivity (dB re 1V/µPa) as a function of frequency (Hz)

Data from each recorder was converted to pressure values using the specific sensitivity of the hydrophone on the AMAR and accounting for the Analog-To-Digital Converter characteristics (Table 1).

Table 1: Analog-To-Digital converter characteristics (up) and equation for converting digital values to dB re 1 µPa²/Hz (bottom). s is the Power Spectral Density (PSD) estimate of the input acoustic data and H the sensitivity of the hydrophone given by the calibration curve

Bit depth	Voltage (V)	Digital values Full Scale (FS)	Gain G (V/V)	Conversion factor M (V/bit)	Conversion factor N (dB re FS/V)
24	-2.5 to 2.5	-8 388 608 to 8 388 607 (= $\pm 2^{24}/2$)	2	2.98 x 10^{-7} (= 5 V / 2^{24} bits)	-7.96 (= $20\log_{10}(1/2.5\ V)$)

Data calibration to (dB re 1 µPa²/Hz)	
Power Spectral Density (dB re 1 µPa²/Hz)	PSD = $10\log_{10}(s) - 20\log_{10}(G) - N - H$

2.1.2. Cook Strait dataset

Figure 3 reports the location of the recorders in the Cook Strait region, New Zealand.

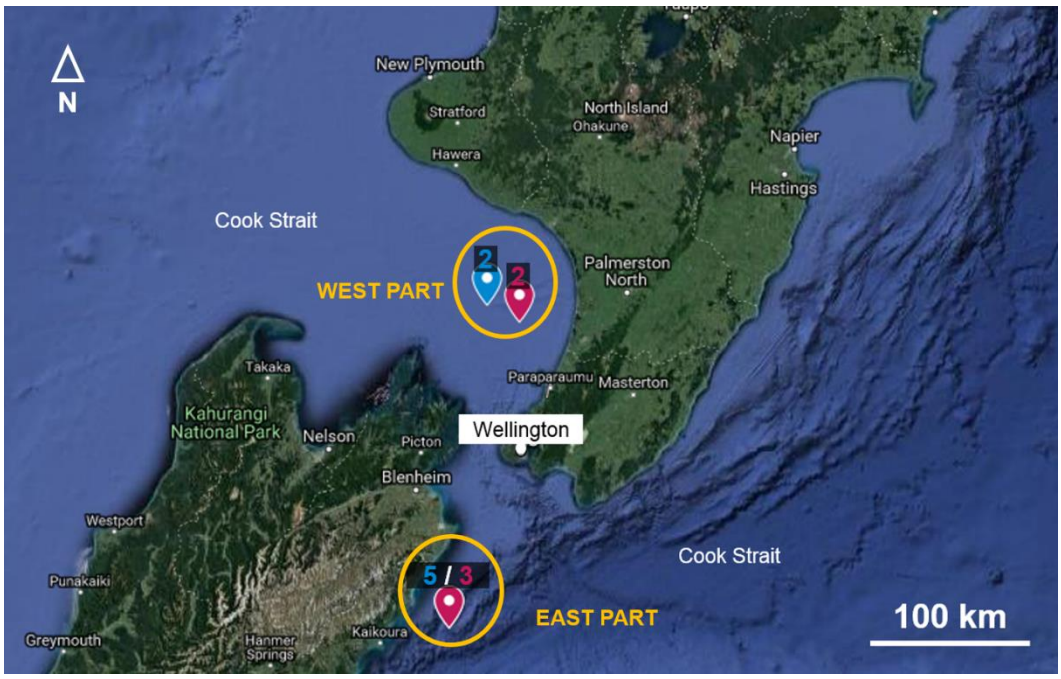

Figure 3: Location and number of the stations for the AMARs deployed in the Cook Strait, New Zealand. Blue: deployed in 2016. Pink: deployed in 2017. *Source: Google Maps*

The four recording stations were located on the two sides of the Cook Strait. The AMARs operated on a duty cycle recording 629 s of data at a sampling rate of 16 000 Hz and then turning off for 271 s (total duty cycle length: 900 s). Station 2 was moved between 2016 and 2017 due to weather conditions. Full deployment and recordings details are reported in Table 2.

Table 2: Deployments details for the AMARs in the Cook Strait

Part	Station	Hydrophone	Year of deployment	Latitude	Longitude	Depth (m)	Start date of recording (DD/MM/YYYY)	End date of recording (DD/MM/YYYY)	Number of files analyzed	Total duration recording (days)
East	5	AMAR 490	2016	-42.3087	174.2145	1 251	28/04/2016	21/12/2016	19 022	138
	3	AMAR 492	2017	-42.3071	174.2139	1 200	15/02/2017	08/09/2017	19 092	139
West	2	AMAR 498	2016	-40.419567	174.5074	110	04/06/2016	21/12/2016	19 040	139
	2	AMAR 498	2017	-40.5259	174.7571	100	15/02/2017	04/09/2017	17 982	131

The stations on the East part were deployed in deep waters (Table 2), so a sphere AMAR buoy was used (Figure 1, right), whereas the stations of the West part were deployed in shallow waters (Table 2), so a bottom base plate AMAR buoy was used (Figure 1, left).

2.1.3. Ross Sea dataset

Figure 4 reports the location of the recorders in the Ross Sea, Antarctica.

Figure 4: Location and number of the stations for the AMARs deployed in the Ross Sea, Antarctica, during 2018.
Source: Google Maps

Three AMARs were deployed in 2018 and were recovered in February 2019 by the research vessel *Tangaroa*. However, one AMAR (corresponding to Station 2) could not be recovered due to a technical issue. The AMARs operated on a duty cycle recording 341 s of data at a sampling rate of 48 000 Hz and then turning off for 799 s (total duty cycle length: 1 140 s). The details of the deployments are summarized in Table 3.

Table 3: Deployments details for the AMARs in the Ross Sea

Station	Hydrophone	Year of deployment	Latitude	Longitude	Depth (m)	Start date of recording (DD/MM/YYYY)	End date of recording (DD/MM/YYYY)	Number of files analyzed	Total duration recording (days)
1	AMAR 489	2018	-73.2869	177.0400	1 100	22/02/2018	27/01/2019	25 635	101
3	AMAR 491	2018	-63.6833	176.3286	1 500	13/03/2018	14/01/2019	23 285	92

The stations were deployed in deep waters (Table 3), so sphere AMARs were used (Figure 1, right).

2.2. Boat detection

A boat detector was developed with MATLAB® (version R2017b) using the dataset from the Cook Strait in New Zealand. Data from Station 5 (Figure 3) was used as training dataset to develop and test the detector algorithm. A 10th order low-pass filter with cut-off frequency of 1 kHz was applied on the signal. Most of sounds produced by boats are in fact below this cut-off frequency (Lourens, 1988).

Power Spectral Densities (PSD; Lourens, 1990) over consecutive one-minute long data snippets were calculated. A PSD gives the power (in dB re 1 µPa²/Hz) of a signal versus its frequency (in Hz). The equation used to calibrate the acoustic data is reported in Table 1. Each PSD (Figure 5) was calculated with the function `pwelch`.

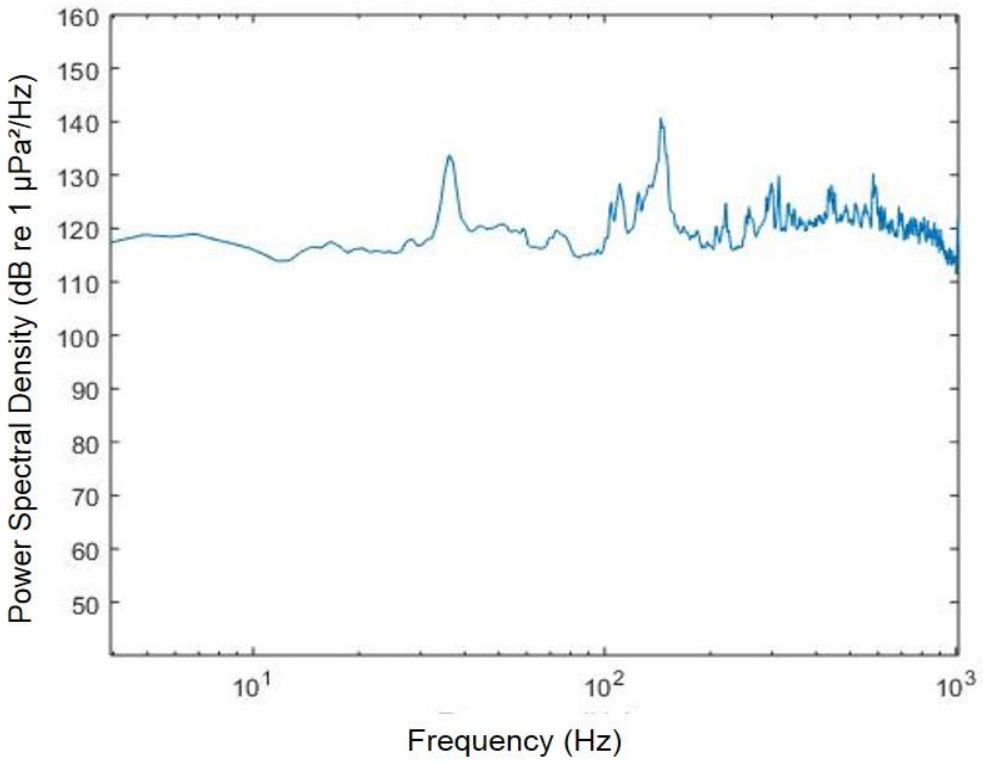

Figure 5: Power Spectral Density of one-minute long data snippet containing the sound signal of a boat (from Cook Strait dataset)

Consecutive one-minute PSDs were concatenated to obtain spectrograms of the acoustic data (Figure 6).

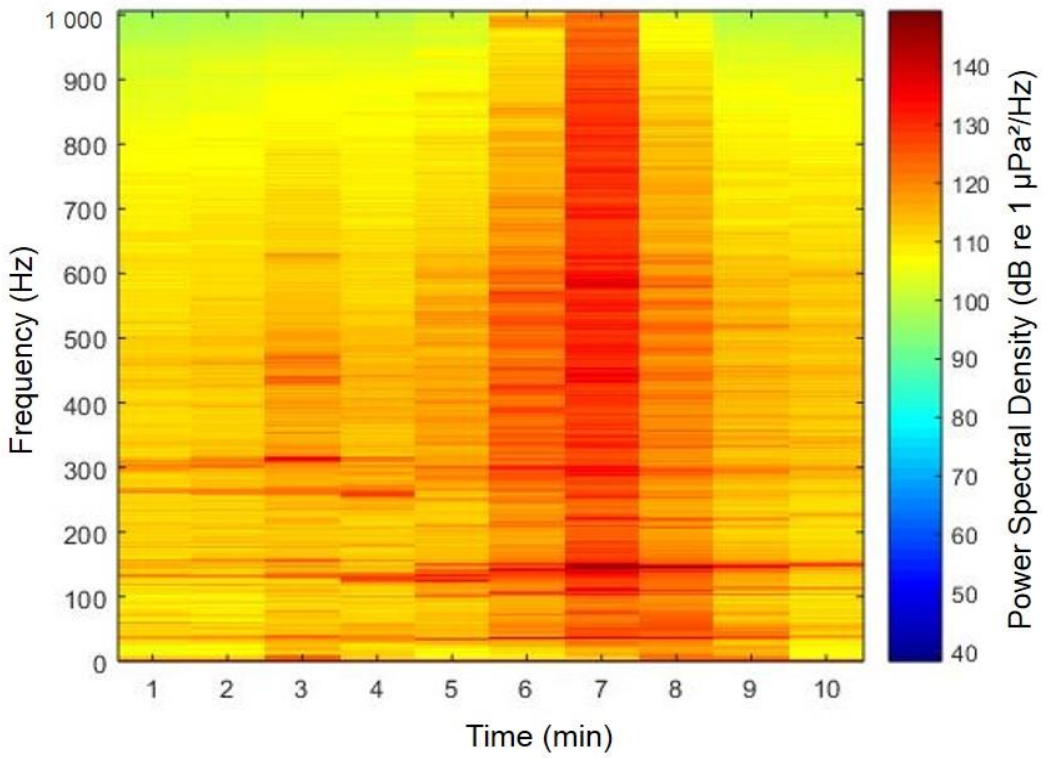

Figure 6: Spectrogram of the sound signal of a boat obtained with the concatenation of Power Spectral Densities (from Cook Strait dataset, frequency resolution = 1 Hz, time resolution = 1 min)

The detection of vessels consists in tracking tonal sounds, which are produced by the engines (Seto, 2011). Graphically, these tonals appear as horizontal lines in the spectrogram (see Figure 6) because their sound frequency remains almost constant over time (Lourens, 1990). These horizontal lines correspond to peaks for the PSDs (Figure 5). Therefore, local maxima in the PSDs were identified in order to track tonal sounds (Seto, 2011). The function `findchangepts` was used here to find these local maxima as abrupt changes in each PSD. This function gives the frequency where there is the most significant change from the mean signal. 100 changing points were found on each PSD, resulting in a residual error of about 100 Hz, which seems suitable compared to the frequency bandwidth Δf = 1 000 Hz (about 10 % of the frequency bandwidth). The frequency difference between consecutive changing points over time (i.e. over each PSD) was then calculated and it is assumed that a tonal sound is detected if a changing point does not move more than 10 Hz away. The next step consisted in measuring the duration of the tonal sounds.

A second step consisted in calculating the difference (in dB) between each PSD in a .wav file and their median PSD. The tonal sounds detection is done by choosing a threshold for this difference with the median.

A boat is therefore detected if the two criteria previously described are satisfied: a certain duration of the tonal sounds and a certain difference in dB for the PSD compared to the median PSD. These two parameters of the algorithm of the boat detector and their values are determined in Table 5.

The frequencies of detected tonal sounds and their power were extracted with the detection. After being validated with the Cook Strait dataset, the detector was launched on the two stations of the Ross Sea dataset (Figure 4).

2.3. Fin whale detection

2.3.1. Creation of a 'doublet' call template

The fin whale detector was developed with MATLAB® with the same dataset (Cook Strait) and a different station (Station 3) from the one that was used for the boat detector (see Figure 3). A 6^{th} order bandpass filter was applied to keep only frequencies above 10 Hz and below 120 Hz. This filter may allow a proper detection for fin whales, as all the different kind of calls emitted by the species are contained in this frequency bandwidth (McDonald and Fox, 1999).

Fin whales have three different calls (McDonald and Fox, 1999) that can be sorted in two categories. The 'doublet' call which is a stereotyped pattern around 20 Hz. Fin whales also emit some irregular calls that can be found in a large frequency bandwidth (20 Hz to 90 Hz; McDonald and Fox, 1999). Figure 7 shows 'doublet' calls recorded in Cook Strait.

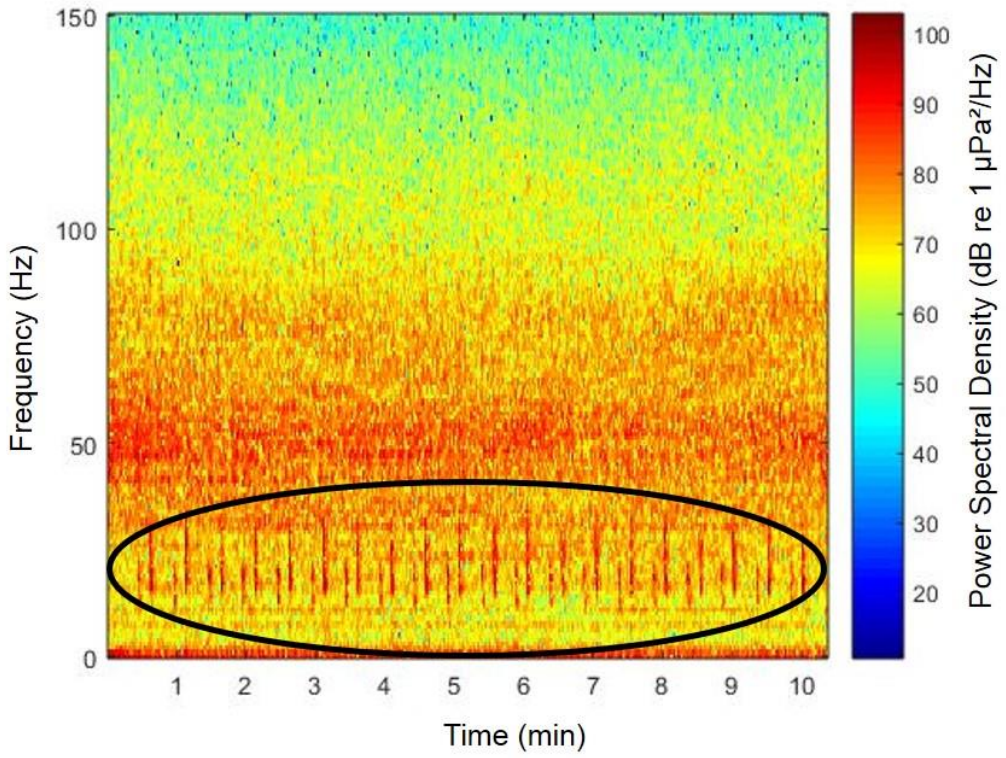

Figure 7: Spectrogram of one .wav file showing fin whales 'doublet' calls (from Cook Strait dataset, frequency resolution = 1 Hz, time resolution = 0.5 s)

A 'doublet' call consists of two calls repeated in a pattern. The first one has a narrower frequency bandwidth than the second one (Figure 7). As this type of call is the most known and dominant for the fin whale (Ou *et al.*, 2015), it was used here for the detection task. Indeed, this approach of detecting a pattern of two calls was chosen to reduce the possibility of falsely detecting other baleen whale's species. A fin whale was considered as detected only if at least one 'doublet' call was present in the signal.

The fin whale detection is here based on the cross-correlation between a 'doublet' call template and the spectrogram of the acoustic data (Mellinger and Clark, 2000). This approach performs well for the stereotyped calls of fin whales (Mouy, 2007). To create the 'doublet' call template, 10 .wav files containing fin whale's 'doublet' calls were selected. The spectrogram of each .wav file was obtained applying a Fast Fourier Transform (FFT) to the data and the PSD (Figure 7). The equation used to calibrate the acoustic data is in Table 1.

One 'doublet' call was manually selected from each spectrogram of the 10 .wav files using the function `ginput`. This created a dataset of 10 'doublet' calls. The first 'doublet' call selected was forced to have a frequency bandwidth ranging between 13 Hz and 32 Hz. This helped designing the frequency size of the other 'doublet' calls selected. All the 10

selected ‘doublet’ calls were horizontally cut with the frequency bandwidth of the first one. The first ‘doublet’ call was then cross-correlated with the other 9 so that the ‘doublet’ calls could be precisely synchronized in time. This allowed to cut them vertically to the same time duration. With this method, the selected ‘doublet’ calls were cut to the same time-frequency sizes and were averaged to compute one single ‘doublet’ call template, which was then binarized (Figure 8).

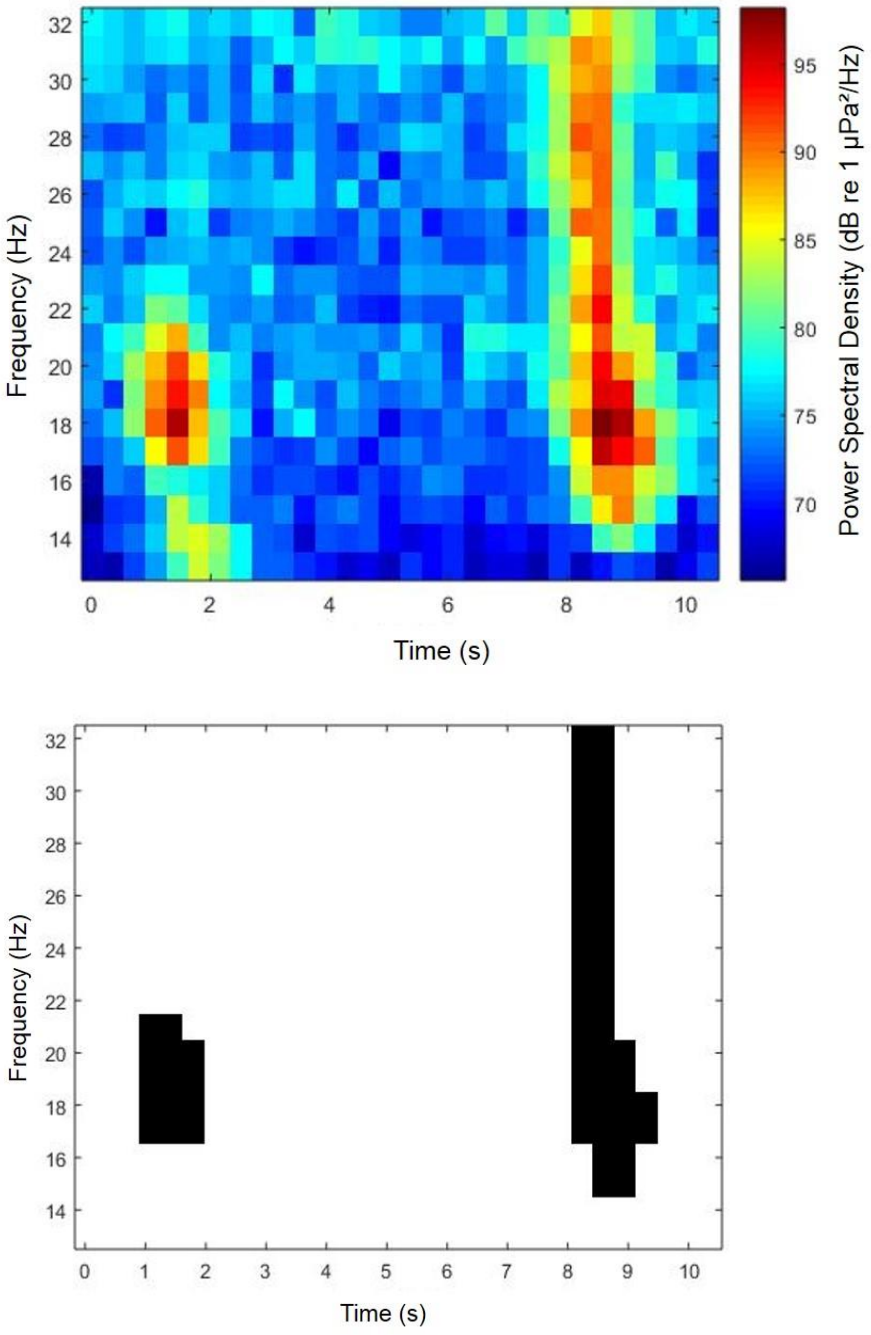

Figure 8: Binarization of the ‘doublet’ call template. Up: averaged ‘doublet’ call template of a fin whale ‘doublet’ call. Down: binarized ‘doublet’ call template (made with Cook Strait dataset, frequency resolution = 1 Hz, time resolution = 0.5 s, binarization threshold = 85 dB)

The binarized ‘doublet’ template call was saved with its frequency bandwidth and was used by the detector algorithm. The binarization approach is well-used for making templates of stereotyped calls (Mellinger and Clark, 2000). It was here made with a threshold of 85 dB, manually assessed to keep most of the pattern of the call. If the value of the PSD went beyond this threshold, the pixel value was assigned to 1 (it appears in black in Figure 8, down); if the value was smaller than the threshold, the pixel value was assigned to 0 (it appears in white in Figure 8, down).

2.3.2. Detection algorithm

To detect fin whales in the acoustic data, spectrograms of the .wav files were cut within the frequency bandwidth of the ‘doublet’ call template. The spectrogram was binarized with a threshold to highlight the peaks of the ‘doublet’ calls (Figure 9).

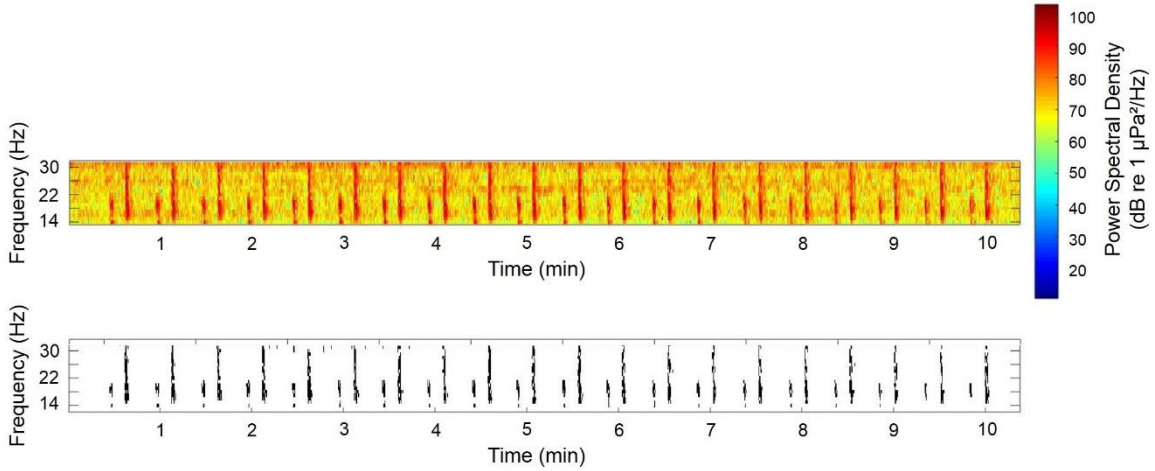

Figure 9: Spectrogram binarization. Up: spectrogram of one .wav file showing fin whales ‘doublet’ calls. Down: binarized spectrogram (from Cook Strait dataset, frequency resolution = 1 Hz, time resolution = 0.5 s, binarization threshold = determined with trials)

The binarization threshold is a parameter of the algorithm of the fin whale detector and its value is determined in Table 6.

A cross-correlation was calculated between the binarized ‘doublet’ call template and the binarized spectrogram. Then, a ‘doublet’ call was detected when there was a match between them, which means when the cross-correlation reached a peak (Figure 10).

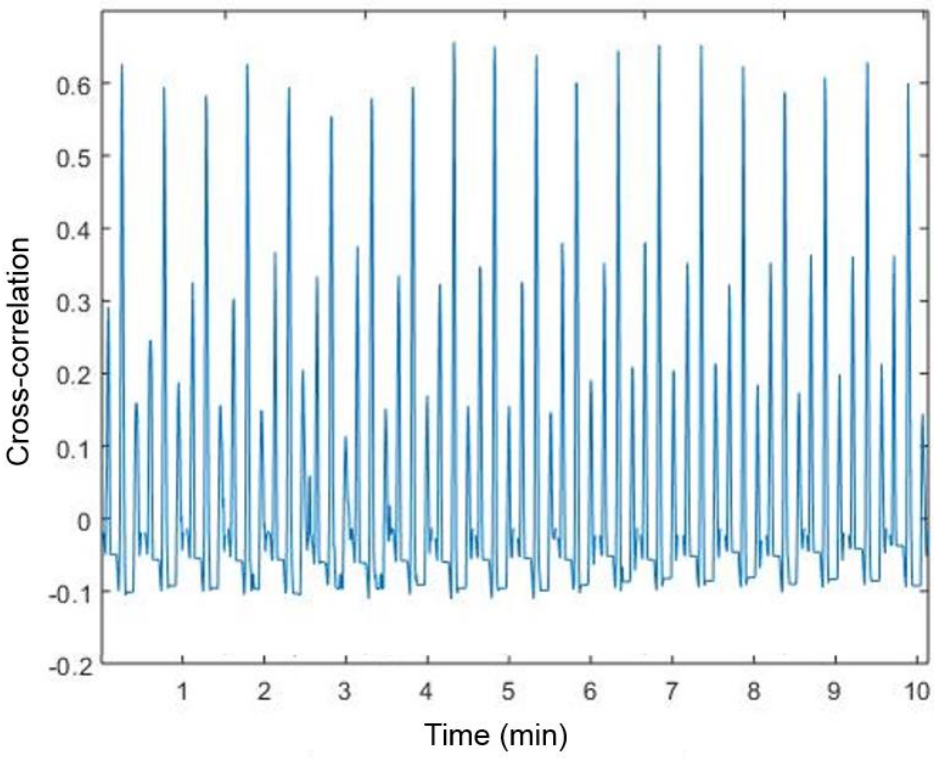

Figure 10: Cross-correlation between the binarized 'doublet' call template and the binarized spectrogram

All the cross-correlation coefficient peaks were detected with the function `findpeaks`. However, this function requires a correlation threshold. So, a 'doublet' call was detected when the cross-correlation coefficient exceeded the threshold. This threshold value was determined with trials to optimize the detection (Table 6).

Like the boat detector, the fin whale detector requires the choice of two thresholds: a binarization threshold for the spectrogram and a cross-correlation threshold to assess the match between the 'doublet' call template and the spectrogram (Table 6).

Finally, one 'doublet' call must be at least spotted by the detector to assume that a fin whale has appeared in the signal. The outputs extracted were the number of 'doublet' calls, their average PSD and the spectrograms of the detected 'doublet' calls. After being validated with the Cook Strait dataset, the detector was launched on the two stations of the Ross Sea dataset (Figure 4) and on the four stations of the Cook Strait dataset (Figure 3).

2.4. Validation of the detectors and outputs analysis

2.4.1. Receiver Operating Characteristics (ROC) curve

Receiver Operating Characteristics (ROC; Fawcett, 2006) curves were used to optimize the threshold and binarization parameters used by the detectors (boat and fin whale). Two training datasets (one for the boat case, and one for the fin whale case) consisting of 68 .wav files from the Cook Strait recordings were created. Both the training datasets contained 34 .wav files with the target signal (either boat tonal sounds or fin whale 'doublet' calls), and 34 .wav files with no target signal. After running the detectors on these training datasets, four outputs were possible (Table 4): True Positive (TP; detected and present), False Negative (FN; not detected but present), False Positive (FP; detected but absent) and True Negative (TN; not detected and absent).

Table 4: Confusion matrix that shows the four possible outputs for the comparison between detection and reality (Fawcett, 2006)

		Reality	
		Presence	Absence
Detection	Detected	True Positive (TP)	False Positive (FP)
	Not detected	False Negative (FN)	True Negative (TN)

The ROC is based on the determination of TP and FP ratios (Fawcett, 2006). A good detector is characterized by high TP ratio (at least 85%) and small FP ratio (at most 10%). Both detectors were run 10 times on their training dataset to make 10 trials. These trials were performed by changing the parameters of the detectors: duration of tonal sounds and dB difference compared to the median PSD for the boat detector; binarization threshold and cross-correlation threshold for the fin whale detector. Of the 10 trials, the parameters that yielded the best detector performance were saved and used for the detection in the real datasets from the Ross Sea and Cook Strait.

2.4.2. Estimation of the detectors' performance

After launching the detectors on all the different stations, an amount of 250 files were taken randomly from the outputs of each station to determine the detectors' performance on the real dataset. The performance was assessed with four parameters: precision (P), recall

(R), F-score (F) and accuracy (A) (Fawcett, 2006). P, R, F and A were calculated using the following equations:

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN} \qquad F = \frac{(1 + \beta^2)P * R}{\beta^2 P + R} \qquad A = \frac{TP + TN}{N}$$

where TP, FN, TN and FP are the True Positive, False Negative, True Negative and False Positive ratios (as described in section 2.4.1).

P represents the proportion of detections that are TP. As an example, a P value of 0.9 means that 90% of the detections were correct without saying whether all tonal sounds or 'doublet' calls in the dataset were identified. R represents the proportion of boat tonal sounds and fin whale 'doublet' calls in the dataset that are detected by the detector. An R value of 0.8 means that 80% of all boat tonal sounds and fin whale 'doublet' calls in the dataset were detected without saying how many detections were incorrect. Thus, a perfect detector would have P and R values equal to 1. The F-score is a combined measure of P and R. An F-score of 1 indicates perfect performance of the detector. β is the relative weight between the recall and precision. For example, a β of 0.5 means the recall has half the weight of the precision. Finally, A represents how likely the detectors identify TP or TN. If A is higher than 0.5, then the detectors were better at identifying TP or TN than FP or FN. N is the number of .wav files analysed (250 in this case).

2.4.3. Sound frequencies of detected boats

The frequencies of all boat tonal sounds that were detected for each station in the Ross Sea were analysed with MATLAB®. All the boat tonal sounds values were sorted, and their occurrence was counted. This approach was chosen to understand which boat frequencies were the most dominant in the outputs. Therefore, it was possible to assess the kind of boat that was the most present in the acoustic data (Mckenna *et al.*, 2012).

2.4.4. Analysis of abundance over time

To have an idea of the abundance of boats over time, the number of files with a detection for each month was counted and compared with the total number of .wav files for each month to get a ratio, using a `barplot` with the software R (version 3.6.0). That kind of plot can show how abundant the boats were in each month.

The same calculation was performed with the detections from the fin whale detector. However, an additional approach was considered: the number of ‘doublet’ calls detected for each month. The number of ‘doublet’ calls detected for each month was compared with the total number of ‘doublet’ calls, showing how the ‘doublet’ calls are distributed during the year.

3. Results

3.1. ROC curves

3.1.1. Boat detector

After the 10 trials on the training dataset (68 .wav files described in section 2.4.1.), a ROC curve was obtained for the boat detector (Figure 11).

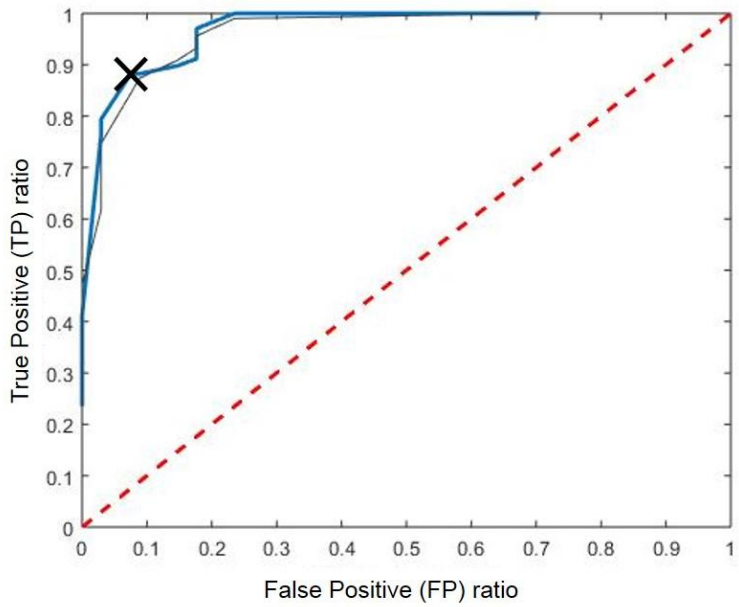

Figure 11: ROC curve for the boat detector (blue line: ROC curve, black line: smoothed ROC curve with moving average, red dotted line: random guess, black cross: trial whose parameters were selected)

The values of the parameters that returned the closest point to the upper left corner of the graph (i.e. the best classification; Figure 11) were selected to be used on the real datasets. Concerning the boat detector, the tonal sounds must last at least 2 minutes and the PSD must be at least 0.9 dB above the median PSD. Table 5 summarizes all the 10 trials with the values of parameters and TP, FP, TN, and FN ratios.

Table 5: Boat detector trials. Red: Best performance trial. Values of parameters of Trial 5: tonal sounds of 2 minutes and PSD 0.9 dB above the median PSD)

Trial	Tonal sounds duration (min)	Difference with the median PSD (dB)	True Positive (TP)	False Negative (FN)	True Negative (TN)	False Positive (FP)
1	5	1	76 %	24 %	97 %	3 %
2	3	1	79 %	21 %	97 %	3 %
3	3	0.8	88 %	12 %	85 %	15 %
4	2	0.8	91 %	9 %	82 %	18 %
5	2	0.9	88 %	12 %	91 %	9 %
6	5	3	24 %	76 %	100 %	0 %
7	1	0.2	100 %	0 %	29 %	71 %
8	2	0.5	100 %	0 %	76 %	24 %
9	2	0.65	97 %	3 %	82 %	18 %
10	5	1.5	41 %	59 %	100 %	0 %

These parameters from Trial 5 (Table 5) were used to run the detector on the Ross Sea dataset.

3.1.2. Fin whale detector

After the 10 trials on the training dataset (68 .wav files described in section 2.4.1.), a ROC curve was obtained for the fin whale detector (Figure 12). The detection approach was different from the boat detector. The TP, FP, FN and TN ratios do not depend on whether the fin whale is detected or not in the .wav file. The ratios are now given by the number of ‘doublet’ calls detected/not detected in all the 68 .wav files divided by the total number of ‘doublet’ calls of all the 68 .wav files.

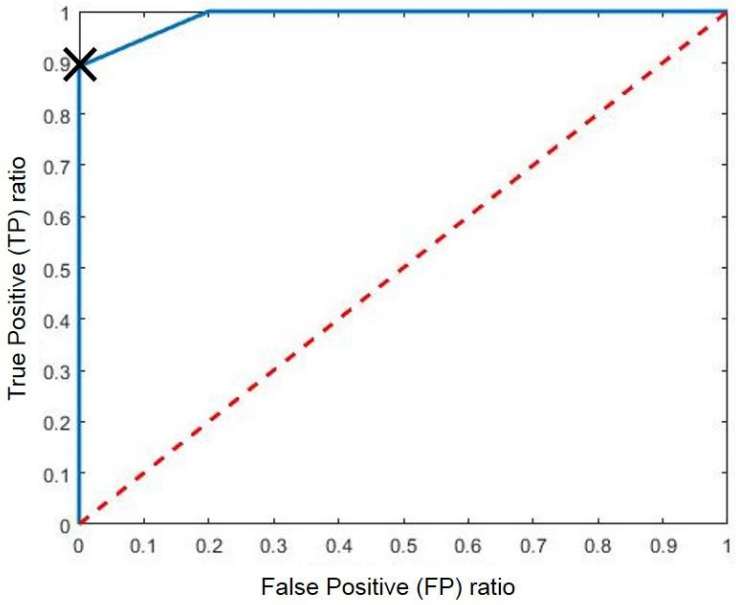

Figure 12: ROC curve for the fin whale detector (blue line: ROC curve, black line: smoothed ROC curve with moving average, red dotted line: random guess, black cross: trial whose parameters were selected)

The best detector performance was obtained using a binarization threshold of 88 dB and a cross-correlation threshold of 0.5 (Table 6).

Table 6: Fin whale detector trials. Red: Best performance trial. Values of parameters of Trial 4: binarization threshold = 88 dB and cross-correlation threshold = 0.5)

Trial	Binarization threshold (dB)	Cross-correlation threshold	True Positive (TP)	False Negative (FN)	True Negative (TN)	False Positive (FP)
1	83	0.6	30 %	70%	100 %	0 %
2	83	0.5	75 %	25 %	100 %	0 %
3	73	0.5	0 %	100 %	100 %	0 %
4	88	0.5	89 %	11 %	100 %	0 %
5	78	0.5	0 %	100 %	100 %	0 %
6	93	0.5	55 %	45 %	100 %	0 %
7	63	0.3	0 %	100 %	100 %	0 %
8	88	0.6	67 %	33 %	100 %	0 %
9	93	0.6	29 %	71 %	100 %	0 %
10	73	0.2	100 %	0 %	80 %	20 %

These thresholds from Trial 4 (Table 6) were used to run the detector on the Cook Strait (East and West) and Ross Sea dataset.

3.2. Estimation of the detectors' performance

3.2.1. Boat detector

The detector was run on the two stations of the Ross Sea (Figure 4). Then, 250 .wav files were picked up randomly from each station and the TP, FP, TN, and FN ratios were calculated to assess the performance of the detector. The validation of the detection was done manually: each .wav file was visually inspected to check whether the classification was correct or not, to calculate the TP, FP, TN, and FN ratios as well as the performance parameters (Table 7 and 8).

Table 7: Confusion matrix of the boat detector on the Ross Sea dataset, Antarctica, during 2018

Station 1					
Presence			Absence		
Detected	True Positive (TP)	64 %	Detected	False Positive (FP)	0 %
Not detected	False Negative (FN)	36 %	Not detected	True Negative (TN)	100 %
Station 3					
Presence			Absence		
Detected	True Positive (TP)	86 %	Detected	False Positive (FP)	4 %
Not detected	False Negative (FN)	14 %	Not detected	True Negative (TN)	96 %

Table 8: Performance of the boat detector on the Ross Sea dataset, Antarctica, during 2018

	Station 1	Station 3
	N = 250	N = 250
Precision $P = \frac{TP}{TP + FP}$	100 %	96 %
Recall $R = \frac{TP}{TP + FN}$	64 %	86 %
Accuracy $A = \frac{TP + TN}{N}$	66 %	73 %

	Station 1	Station 3
β	F-score $F = \frac{(1 + \beta^2)P * R}{\beta^2 P + R}$	
2	0.69	0.87
1.5	0.72	0.88
1	0.78	0.90
0.5	0.90	0.93

The TP ratio is higher for the Station 3 compared to Station 1 (Table 7), as well as the recall, accuracy and F-score (Table 8).

3.2.2. Fin whale detector

The detector was run on four stations of the Cook Strait (Figure 3) and on the two stations of the Ross Sea (Figure 4). Then, 250 files were picked up randomly from each station and the TP, FP, TN, and FN ratios were calculated to assess the performance of the detector.

No fin whale was detected in the West part of the Cook Strait, either in 2016 or 2017: the detector gave a very few FP detections. Moreover, when there was no detection, there was no fin whale (yielding a TN of 100 %). The same outputs were also obtained for the East part of the Cook Strait in 2016. Furthermore, no fin whale was detected in the Ross Sea, at each station. In other words, the species was only detected in the East part of the Cook Strait in 2017 (i.e. Station 3; Figure 3). As in the case of the boat detector, the validation of the fin whale detector performance (Table 9 and 10) was performed with a visual inspection of the 250 randomly selected .wav files.

Table 9: Confusion matrix of the fin whale detector on the Cook Strait dataset, East part, during 2017

Station 3					
Presence			Absence		
Detected	True Positive (TP)	86 %	Detected	False Positive (FP)	0 %
Not detected	False Negative (FN)	14 %	Not detected	True Negative (TN)	100 %

Table 10: Performance of the fin whale detector on the Cook Strait dataset, East part, during 2017

	Station 3
	N = 250
Precision $P = \frac{TP}{TP + FP}$	100 %
Recall $R = \frac{TP}{TP + FN}$	86 %
Accuracy $A = \frac{TP + TN}{N}$	74 %

β	Station 3 F-score $F = \frac{(1 + \beta^2)P * R}{\beta^2 P + R}$
2	0.88
1.5	0.90
1	0.92
0.5	0.97

3.3. Sound frequencies of detected boats

Here are the results of the occurrences of boat tonal sounds for the Ross Sea dataset:

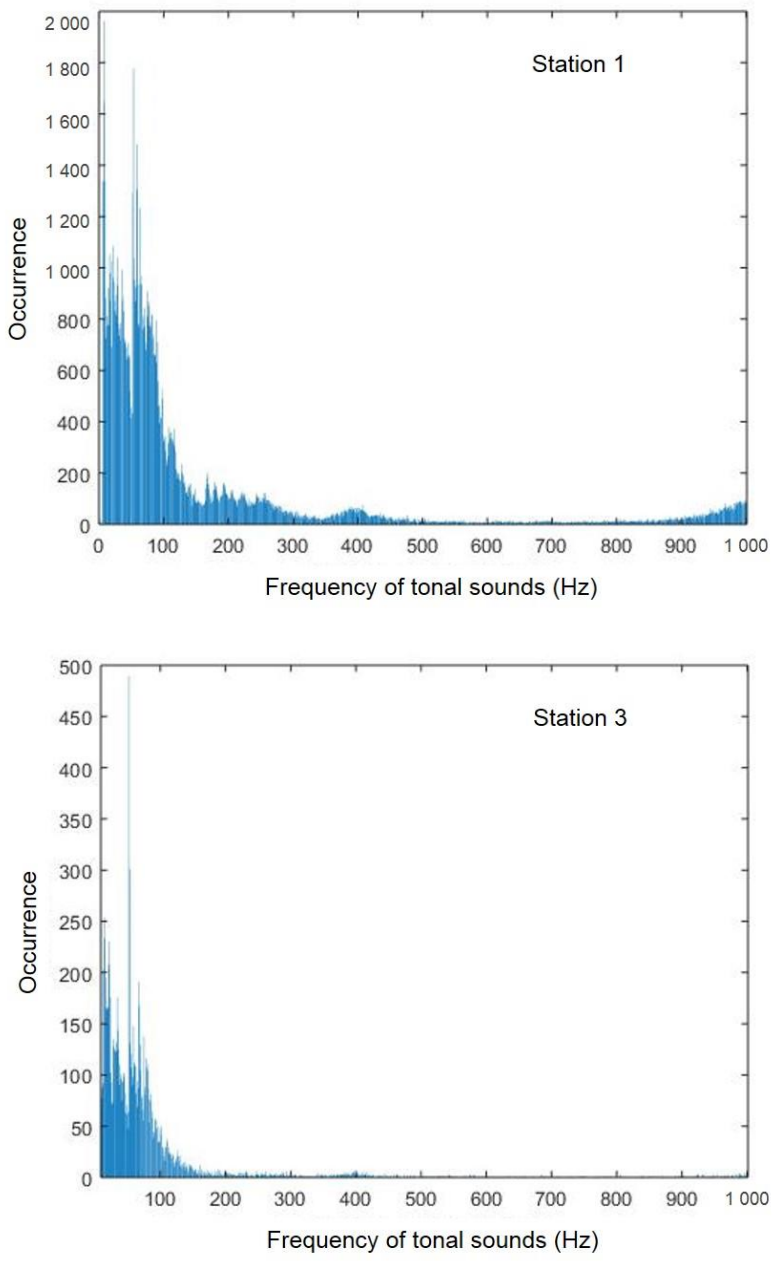

Figure 13: Occurrence of boat tonal sounds for boat detector in the Ross Sea. Up: Station 1. Down: Station 3

The highest amount of boat tonal sounds occurred in the first 100 Hz for both stations (and up to 200 Hz for Station 1; Figure 13, up). A higher number of boat tonal sounds occurred at Station 1 (about 1 000; Figure 13, up) than at Station 3 (about 200; Figure 13, down).

3.4. Analysis of abundance over time

3.4.1. Boats

The abundance of boats in the Ross Sea (Figure 14) was assessed by calculating a ratio between the number of .wav files with a detection in each month divided by the total number of .wav files in each month.

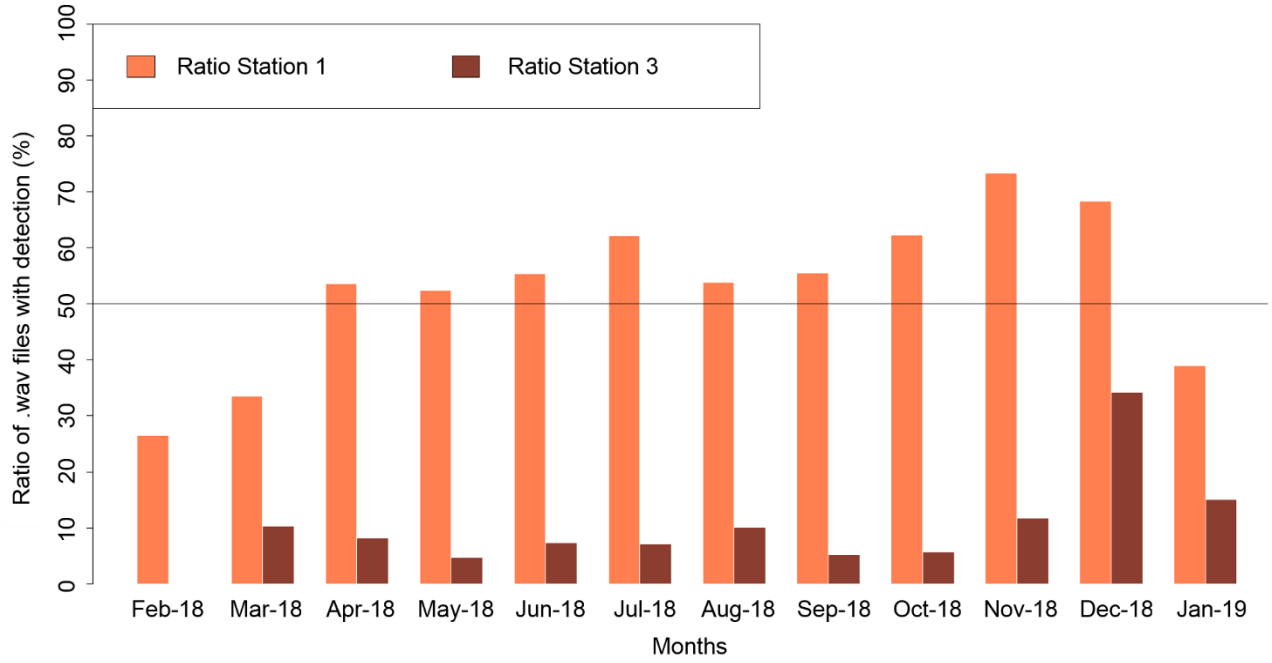

Figure 14: Abundance ratio of boats for each month in the Ross Sea, Antarctica, during 2018 (orange: Station 1, brown: Station 3)

There was a significant difference between Station 1 and Station 3. Indeed, from April 2018 to December 2018, more than one out of two .wav files contained a boat detection for Station 1 (Figure 14). The maximum of abundance is reached at the end of spring: in November 2018 for Station 1 (about 70 % of the .wav files of November) and in December 2018 for Station 3 (about 40 % of the .wav files of December).

3.4.2. Fin whales

Abundance of fin whales in the Cook Strait and Ross Sea was assessed by calculating a ratio between the number of .wav files with a detection in each month divided by the total number of .wav files in each month (Figure 15). Another ratio was also calculated between the number of 'doublet' calls in each month divided by the total number of 'doublet' calls in the whole year. Here only the results for the East part of the Cook Strait in 2017 (Station 3; Figure 3) are reported, since this was the only location were fin whales occurred.

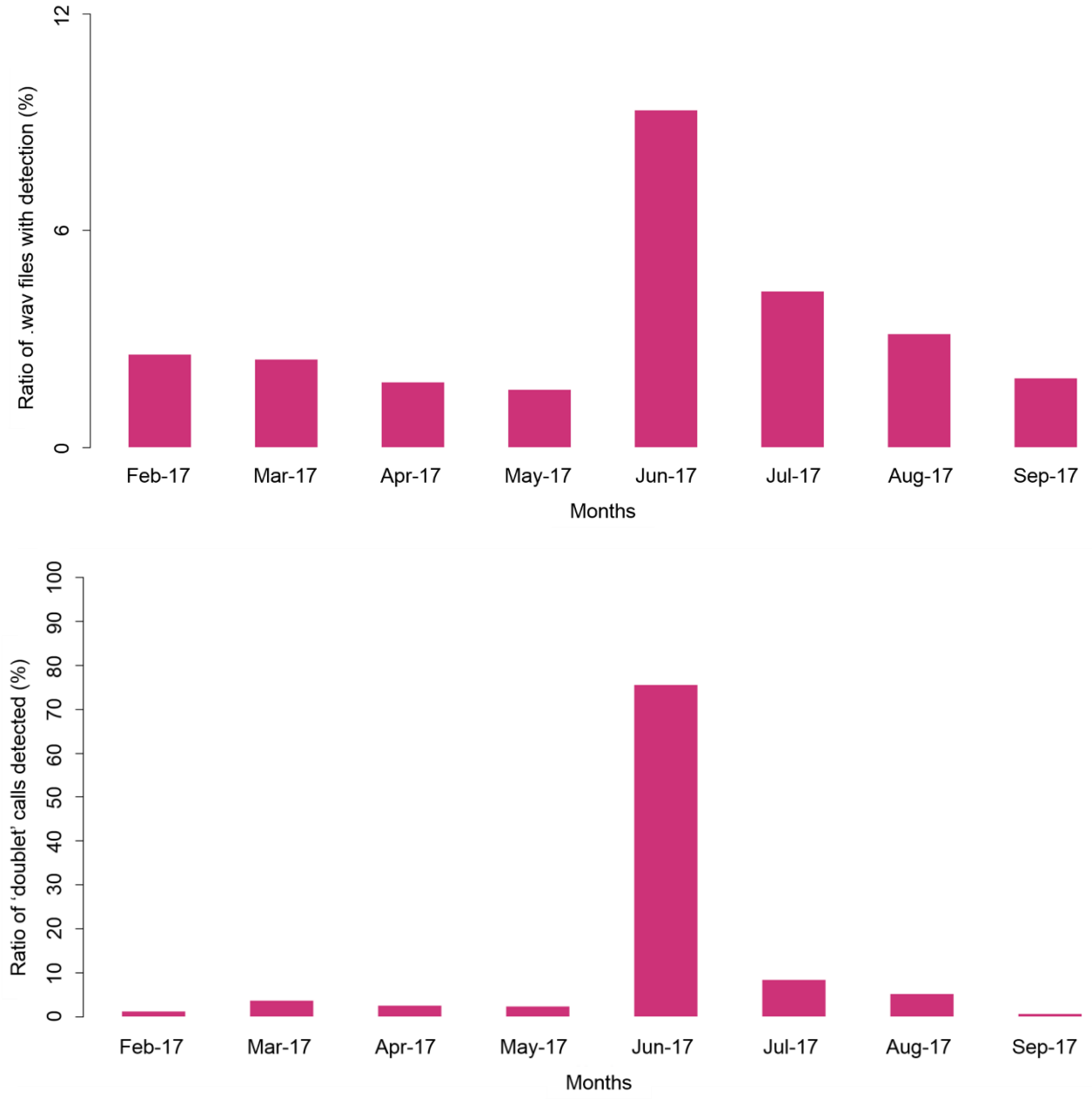

Figure 15: Abundance ratio of fin whales in the Cook Strait, East part, during 2017 (Station 3). Up: abundance ratio of detected .wav files for each month. Down: abundance ratio of 'doublet' calls for each month distributed in the whole year

Fin whales' 'doublet' calls were more abundant in June 2017. In particular, the ratio of .wav files with detection in this month might seems low (about 10 % of the .wav files in June; Figure 15, upper panel) but the ratio of 'doublet' calls is important (more than 50 % of the whole amount of 'doublet' calls in the dataset occurs in June 2017; Figure 15, lower panel). In other words, the fin whales were calling more frequently in a short time period.

4. Discussion

A significant difference between the two stations of the Ross Sea for the boat detection was apparent. Station 1 had more detections than the Station 3. However, this may be due to the location of the recorders. Indeed, Station 1 is close to the Antarctica continent, and most of the noise detected by the algorithm may have been generated by the ice cover and tidal flow. Indeed, the recorder at Station 1 was located on a place that was covered by sea ice from April to December, mostly during autumn (April to June; Holland, 2014). The seasonality of sea ice expansion in the Ross Sea coincides with the high amount of detection at Station 1. Ambient noise under the sea ice of the Antarctica was characterized by low-frequency sounds between 10 Hz and 1 000 Hz and it was correlated with wind and current stresses (Dahl *et al.*, 2007). Because the sound produced by the ice is in the same frequency bandwidth as boats, it can be assumed that the detector misclassified this ice generated sounds as boat tonal sounds. Although there were tonal sounds in the .wav files (a manual validation confirmed their presence) and the detector got good F-scores for both stations at the Ross Sea, these tonal sounds were not probably generated by boats at Station 1.

The frequencies of the tonal sounds that are detected are mainly below 100 Hz. This would be the signature of large boats, such as container ships, cargos or tankers (Mckenna *et al.*, 2012). Further analysis would be required to identify the exact type of boat which is transiting at the time of its detection. Therefore, it would be possible to assess if the boat had the right to cruise in the Ross Sea MPA at that time and whether it is fishing legally or not (Abileah and Lewis, 1996).

The highest abundance of boats has occurred in December for Station 3 at the Ross Sea. It is the end of spring and beginning of summer, so this period is conducive to fishing activities as the cold latitudes gets rich in nutrients (Norris, 1977).

Fin whales were detected only during a short time period: in the winter (mainly in June 2017) and in the East part of the Cook Strait (no detection for the West part). In fact, more than 50 % of the whole amount of ‘doublet’ calls was concentrated in June 2017. The occurrence of fin whales does not seem regular as they were not detected in 2016 but detected in 2017. They also seem to prefer the East part of the Cook Strait than the West part, as there was no detection in the West part. They have not been detected in the Ross Sea, so the hypothesis of a migration of the species between the Cook Strait and the Ross Sea remains unanswered. Fin whales are known to be rare in the Southern Ocean (Branch and Butterworth, 2001). Furthermore, no acoustic study has detected them in the Ross Sea

compared with the rest of Southern Ocean (Širović *et al.*, 2009). Hence, fin whales may have not occurred in 2018 in the Ross Sea, or they may have not sung. As they are known to stay in the Southern Ocean for feeding during summer (Norris, 1977), they might not sing while feeding. Besides, there is a negative correlation between the presence of fin whales near the Antarctica with sea ice cover (Širović *et al.*, 2004), so fin whales may not be detected at Station 1.

The false detections of the fin whale detector may be caused by other ambient sounds: ice cover in the Ross Sea (Dahl *et al.*, 2007; Holland, 2014) or the weather in the East part of the Cook Strait (in 2016) or West part. Since the recorder was deployed in shallow waters in the West part of the Cook Strait, the weather may have influenced the detection results. Indeed, a stormy day with a lot of wind produces low-frequency sounds (Dahl *et al.*, 2007), in the same bandwidth as fin whale's 'doublet' calls.

5. Conclusion

This project consisted in a first step to develop signal processing tools to conduct of an acoustics surveys of fin whales and monitoring boats noise. It has successfully been demonstrated that fin whales seem to prefer the East part than the West part of the Cook Strait. The fact that they were not detected in the Ross Sea is in line with previous published studies. There are many potential applications of the detection algorithms that have been developed through this project. The boat detection may be used with the intent to monitor illegal fishing activities by comparing the time of detected boats with. The detection can also be helpful to understand how busy is an MPA in terms of boat traffic and may be relevant to assess the level of noise pollution on marine life. The detection of animals using template of their stereotyped calls seems fair to identify the presence of the species in a suitable way. This kind of acoustic studies could help to understand the taxa diversity of 'noisy species', such as cetaceans, in some places that are hard to access and monitor for a long time. The knowledge of the acoustic level of detected target (either boats or animals) may give some information on the spatial location of the target. All these applications make passive acoustics a tool that improves the understanding of a marine ecosystem.

6. Bibliography

Abileah, R. and Lewis, D. (1996). Monitoring high-seas fisheries with long-range passive acoustic sensors. *OCEANS 96 MTS/IEEE Conference Proceedings*, pp. 378-382.

Branch, T. A. and Butterworth, D. S. (2001). Estimates of abundance south of 60 ◦S for cetacean species sighted frequently on the 1978/79 to 1997/98 IWC/IDCR-SOWER sighting surveys. *Journal of Cetacean Research and Management*, 3(3), pp. 251–270.

Clapham, P. J. and Baker, C. S. (2001). How many whales were killed in the Southern Hemisphere in the 20th century? Paper SC/53/O14. *IWC Scientific Committee*, pp. 1-3.

Dahl, P. H., Miller, J. H., Cato, D. H. and Andrew, R. K. (2007). Underwater Ambient Noise. *Acoustics Today*, 3(1), pp. 23.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), pp. 861-874.

Giorli G., Goetz, K. T. (2018). The Soundscape of Cook Strait, Results from one year of passive acoustic monitoring. *NIWA report*, 2018186WN, project CHE18301.

Gedamke, J. and Robinson, S. M. (2010). Acoustic survey for marine mammal occurrence and distribution off East Antarctica (30-80°E) in January-February 2006. *Deep Sea Research Part II: Topical Studies in Oceanography*, 57(9), pp. 968-981.

Holland, P. R. (2014). The seasonality of Antarctic sea ice trends. *Geophysical Research Letters*, 41(12), pp. 4230-4237.

Lourens, J. G. (1988). Classification of ships using underwater radiated noise. *IEEE South African Symposium on Communications and Signal Processing*, pp. 130-134.

Lourens, J. G. (1990). Passive Sonar Detection of Ships with Spectrograms. *IEEE South African Symposium on Communications and Signal Processing*, pp. 147-151.

McDonald, M. A. and Fox, C. G. (1999). Passive acoustic methods applied to fin whale population density estimation. *The Journal of the Acoustical Society of America*, 105(5), pp. 2643-2651.

Mckenna, M. F., Ross, D, Wiggins, S. M. and Hildebrand, J. A. (2012). Underwater radiated noise from modern commercial ships. *The Journal of the Acoustical Society of America*, 131(1), pp. 92-103.

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), pp. 3518-3529.

Mouy, X. (2007). Détection et identification automatique en temps-réel des vocalises de rorqual bleu (*Balaenoptera musculus*) et de rorqual commun (*Balaenoptera physalus*) dans l'estuaire du Saint-Laurent. *Institut des sciences de la mer de Rimouski*, Masters, Rimouski, Québec : Université du Québec à Rimouski, 86 p.

Norris, K. S. (1966). Whales, Dolphins, and Porpoises. *International Symposium on Cetacean Research*, *University of California Press*, 1, 789 p.

Ou, H., Au, W. W. L., Van Parijs, S., Oleson, E. M. and Rankin, S. (2015). Discrimination of frequency-modulated Baleen whale downsweep calls with overlapping frequencies. *The Journal of the Acoustical Society of America*, 137(6), pp. 3024-3032.

Parks, S. E., Miksis-Olds, J. L. and Denes, S. L. (2014). Assessing marine ecosystem acoustic diversity across ocean basins. *Ecological Informatics*, 21, pp. 81-88.

Seto, M. L. (2011). Application of Tonal Tracking to Ship Acoustic Signature Feature Identification. *Journal of Vibration and Acoustics*, 133(6), pp. 064501-064501-3.

Širović, A., Hildebrand, J. A. and Wiggins, S. M. (2007). Blue and fin whale call source levels and propagation range in the Southern Ocean. *The Journal of the Acoustical Society of America*, 122(2), pp. 1208-1215.

Širović, A., Hildebrand, J. A., Wiggins, S. M., Mcdonald, M. A., Moore, S. E. and Thiele, D. (2004). Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula. *Deep Sea Research Part II: Topical Studies in Oceanography*, 51(17), pp. 2327-2344.

Širović, A., Hildebrand, J. A., Wiggins, S. M. and Thiele, D. (2009). Blue and fin whale acoustic presence around Antarctica during 2003 and 2004. *Marine Mammal Science*, 25(1), pp. 125-136.

Sorensen, E., Ou, H. H., Zurk, L. M. and Siderius, M. (2010). Passive acoustic sensing for detection of small vessels. *OCEANS 2010 MTS/IEEE SEATTLE*, pp. 1-8.

7. Annexes

Annex 1: Introduction translated in French

L'acoustique passive est une technique rentable pour surveiller de grands écosystèmes marins. En enregistrant les signaux sous-marins, il est possible de détecter les navires qui passent (Sorensen *et al.*, 2010). La détection des sons des bateaux est utile pour identifier et classer les types de navires en transit (Mckenna *et al.*, 2012) et pour discriminer la pêche illégale de la pêche légale (Abileah et Lewis, 1996). L'utilisation de l'acoustique passive est également utile pour faire un inventaire des espèces et pour calculer des indices de biodiversité (Parks *et al.*, 2014) en connaissant les densités de population, par exemple pour les rorquals communs (McDonald et Fox, 1999).

Les populations de rorquals communs (*Balaenoptera physalus*, Linnaeus, 1758) sont très faibles dans l'océan Austral (Branch et Butterworth, 2001). Cela est dû à une exploitation commerciale au 20ème siècle (Clapham et Baker, 2001). Par conséquent, des études visuelles et acoustiques ont été menées pour évaluer la présence des mammifères marins près de l'Antarctique, y compris les rorquals communs (Gedamke et Robinson, 2010). Les rorquals communs sont observés dans les eaux tropicales / subtropicales en hiver pour se reproduire et dans les eaux polaires / subpolaires en été pour se nourrir, avec des migrations souvent longues : de 2 000 km à 5 000 km. En effet, les latitudes froides leur offrent une nourriture riche, telle que le krill (Norris, 1977). Les rorquals communs sont présents près de l'Antarctique entre février et juin (Širović *et al.*, 2004). Cependant, leur présence dans la mer de Ross n'est pas bien connue (Širović *et al.*, 2009). Les signaux émis par les rorquals communs sont contenus dans une plage (ou bande passante) de basses fréquences ; 'doublet 20 Hz', 'intervalle de répétition irrégulier de 20 à 35 Hz' et 'intervalles de répétition plus courts et plus irréguliers de 30 à 90 Hz' (McDonald et Fox, 1999). Comme ces signaux ont des fréquences basses, ils peuvent parcourir de longues distances dans l'eau de mer ; une caractéristique qui les rend aptes à la surveillance acoustique passive à longue portée (Širović *et al.*, 2007). Les objectifs de ce rapport étaient de développer des techniques de traitement du signal pour étudier la présence des bateaux et des rorquals communs dans la mer de Ross et le détroit de Cook. Cette étude visait également à comprendre si une migration potentielle des rorquals communs entre le détroit de Cook et la mer de Ross pouvait être étudiée de manière acoustique.

Annex 2: Conclusion translated in French

Ce projet consistait en la première étape du développement d'outils de traitement du signal permettant de réaliser des enquêtes acoustiques sur les rorquals communs et de surveiller le bruit des bateaux. Il a été démontré avec succès que les rorquals communs semblaient préférer la partie Est à la partie Ouest du détroit de Cook. Le fait qu'ils n'aient pas été détectés dans la mer de Ross est conforme aux études publiées précédemment. Il existe de nombreuses applications potentielles des algorithmes de détection développés dans le cadre de ce projet. La détection de bateaux peut être utilisée dans le but de surveiller les activités de pêche illégale en accédant à l'heure de détection. La détection peut également être utile pour comprendre à quel point une AMP est fréquentée en termes de trafic maritime et peut être utile pour évaluer le niveau de pollution sonore sur la vie marine. La détection des animaux à l'aide d'un modèle d'appels stéréotypés semble convenable pour identifier la présence d'une espèce de manière appropriée. Ce type d'étude acoustique pourrait aider à comprendre la diversité des taxons 'd'espèces bruyantes', telles que les Cétacés, dans certains endroits difficiles d'accès et pour de longues surveillances. La connaissance du niveau acoustique de la cible détectée (bateau ou animal) peut donner des informations sur la localisation spatiale de la cible. Toutes ces applications font de l'acoustique passive un outil qui améliore la compréhension d'un écosystème marin.

ENGEES
ÉCOLE NATIONALE DU GÉNIE DE L'EAU ET DE L'ENVIRONNEMENT DE STRASBOURG

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

MEMOIRE DE FIN D'ETUDES

Diplôme(s) : Ingénieure ENGEES – Master Sciences de la Mer

Spécialités : Hydrosystèmes – Océanographie Biologique et Écologie Marine

Auteur	**Année**
Alexandra CONSTARATAS	**2019**

Titre

Détection des bateaux et des rorquals communs en utilisant des techniques d'acoustique passive

Nombre de pages : texte : 38 annexes : 2

Nombre de références bibliographiques : 22

Structure d'accueil : National Institute of Water and Atmospheric Research (NIWA), Wellington, Nouvelle-Zélande

Maître de stage : Dr Giacomo GIORLI

Résumé

Les océans représentent le plus grand écosystème de la planète. Les surveiller est un défi : les chercheurs ne peuvent atteindre facilement (et à moindre coût) certaines régions éloignées. L'Aire Marine Protégée de la mer de Ross en est un exemple. Les techniques d'acoustique passive représentent une solution de surveillance. Cependant, les données acoustiques contiennent des signaux audios provenant de différentes sources sonores. Il faut pouvoir détecter les entités ciblées. Nous développons ici des algorithmes de traitement du signal pour trois tâches : la détection des bateaux, la mesure de leurs fréquences et la détection des appels de type 'doublet' des rorquals communs.

Mots-clés

Aire Marine Protégée – Acoustique passive – Traitement du signal – Détection – Bateaux – Rorquals communs