Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence

Cédric Gervaise, Yvan Simard, Florian Aulanier, and Nathalie Roy

Fisheries and Oceans Canada
Science Branch
Maurice Lamontagne Institute
850 route de la Mer, P.O. Box 1000
Mont-Joli, Québec
Canada G5H 3Z4

2019

Canadian Technical Report of Fisheries and Aquatic Sciences 3345

Fisheries and Oceans Canada | Pêches et Océans Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. II n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences 3345

2019

Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence

by

Cédric Gervaise[1], Yvan Simard[2], Florian Aulanier[2], and Nathalie Roy[2]

[1]CHORUS research institute
PhelmaMinatec
3 Parvis Louis Neel
38000 Grenoble
France

[2]Fisheries and Oceans Canada
Sciences Branch
Maurice Lamontagne Institute
850 route de la Mer, P.O. Box 1000
Mont-Joli, Québec
Canada G5H 3Z4

Cat. No. Fs 97-6/3345E-PDF ISBN 978-0-660-33169-0 ISSN 1488-5379

Correct citation for this publication:

Gervaise, C., Simard, Y., Aulanier, F., and Roy, N. 2019. Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 3345: ix + 58 p.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS iii
LIST OF TABLES v
LIST OF FIGURES v
ACRONYMS vii
ABSTRACT viii
RÉSUMÉ ix
1. INTRODUCTION 1
1.1 NARW CURRENT CONTEXT AND THREATS IN GSL 1
1.2 OBJECTIVES 3
1.3 CONTENT OF THE REPORT 4
2. MATERIALS AND METHODS 5
2.1 SETTING THE ORDERS OF MAGNITUDE 5
2.2 ACOUSTIC AND ENVIRONMENTAL INPUTS 7
2.2.1 Objectives 7
2.2.2 NARW acoustics 7
2.2.3 Marine traffic within the dynamic management zones 9
2.2.4 Simulation of acoustic propagation and ambient noise within the study area 10
2.3 PERFORMANCES OF THREE NARW UPCALL DETECTORS FROM SINGLE-HYDROPHONE OR HYDROPHONE-ARRAY SETUPS 11
2.3.1 Objectives 11
2.3.2 The data and detection problem 11
2.3.3 Choice of three NARW upcall detectors 13
2.3.4 Description of the three NARW upcall detectors 14
2.3.5 Receiver Operational Characteristics (P_d, P_{fa}, SNR) of the three detectors 18
2.3.6 Accounting for the use of a hydrophone array 20
2.3.7 NARW localization with hydrophone-array PAM systems at 2 coastal positions and its precision 24
3. RESULTS 25
3.1 MAPPING OF THE AREA OF 'GOOD DETECTION' WITHIN THE NARW FEEDING GROUND AREA OFF GASPÉ 25
3.1.1 Scenario 25
3.1.2 Examples of instantaneous maps of detection probability 28
3.1.3 Maps of the proportion of times where $P_d > 0.5$ for 1 single hydrophone and a circular array of 20 hydrophones with TFBD 31
3.1.4 Maps of Good Detection Area and their geometrical characteristics 33
3.1.5 Optimal Good Detection Area and positioning of 2 coastal PAM systems for NARW upcall detection 36
3.1.6 Optimal positioning of 2 PAM systems for NARW upcall localization 38
4. DISCUSSION 40
4.1 CHOOSING THE RIGHT DETECTOR AND NEEDS FOR A PRECISE MAP OF THE DETECTION PERFORMANCE 41
4.2 MAPPING OF THE DETECTION PERFORMANCE THROUGH THE REALISTIC SIMULATION OF TRAFFIC OVER 3 MONTHS 42

4.3 Effective Range of Detection of a single hydrophone and TFBD 42
4.4 Effective Range of Detection with hydrophone array PAM systems with TFBD at 1 or 2 coastal positions 42
4.5 Area for good localization of NARW with 2 cooperating PAM systems using a hydrophone array and TFBD 43
4.6 Operational conclusions: toward an *ad hoc* PAM system to survey NARW in their feeding ground area off Gaspé 43
ACKNOWLEDGMENTS 44
REFERENCES 44
ANNEX 1: OPTIMAL FFT WINDOW LENGTH FOR NARW UPCALL 48
Processing Gain estimation from the spectrogram of monochromatic non-modulated sounds 48
Processing Gain estimation from the spectrogram of linear frequency-modulated sounds 49
ANNEX 2: DETAILS OF ROC CURVES OF THE UPCALL DETECTOR 53
ANNEX 3: EFFECT ON THE PROCESSING GAIN OF A MISFIT BETWEEN THE REFERENCE AND THE TRUE SIGNAL FOR CCBD 56
ANNEX 4: ACCURACY OF LOCALIZATION OF THE NARW WITH THE FORMALISM OF CRAMER RAO BOUNDS 57

LIST OF TABLES

Table 1. AIS traffic density (vessel per day) in the FGA off Gaspé in 2018. 10
Table 2. The four cases of the detection problem. 12
Table 3. Settings of the CCBD, TFBD, and EBD detectors. 17
Table 4. Formula of the ROCs of CCBD, TFBD, and EBD 19
Table 5. Geometrical characteristics of good detection areas. 35
Table 6. ESD and ERD for NARW upcall detection using hydrophone-array PAM systems with TFBD located at 2 coastal positions. 36
Table 7. ESD and ERD using hydrophone-array PAM systems at 2 coastal positions for NARW upcall localization. 38

LIST OF FIGURES

Figure 1. Context of the feeding ground area off Gaspé, with the monthly shipping traffic from June to December from Simard et al. (2014), and the three-dimensional underwater topography with the envelope (white line) of the NARW feeding ground basin, where sightings are concentrated from June to December (from https://whalemap.ocean.dal.ca/WhaleMap/). 2
Figure 2. Positions of the six coastal PAM systems compared for NARW detection and localization. 4
Figure 3. Map and geometry of the study area, showing the shipping routes (pink) and bathymetry of the feeding ground area where NARW sightings are concentrated. (from https://whalemap.ocean.dal.ca/WhaleMap/). 6
Figure 4. Spectrogram examples of NARW upcalls recorded in southern GSL in 2018. 8
Figure 5. Ships positions from AIS data from August 1st 2018 to October 30th. 9
Figure 6. Environmental configuration of the detection of NARW upcalls embedded in ship noise 12
Figure 7. The two-step logic of the detection task. 12
Figure 8. Schemes of the three detectors. 15
Figure 9. Step by step processing of the data for TFBD. 15
Figure 10. Step qualitative assessment of the detection performance of the three detectors. 16
Figure 11. ROCs at $P_{fa} = 2\times10^{-6}$ for the three detectors. 20
Figure 12. Geometry and conventions used for a hydrophone array. 21
Figure 13. Beamforming processing gain for a preferred direction $\alpha_1 = 0°$. 22
Figure 14. Beamforming processing gain for a preferred direction $\alpha_1 = 60°$. 23
Figure 15. Scenario for a realistic case where a whale at azimuth α_w emits an upcall of source level WSL, the transmission loss to reach the array is TL_w, and several ships at azimuth α_{si} with source levels SSL_i and transmission losses TL_i. 24
Figure 16. NARW localization with hydrophone-arrays at 2 coastal positions. 25
Figure 17. Good detection area (yellow) for a hydrophone array located at C1 with TFBD. 27
Figure 18. Map of TL (dB) @ 100 Hz between the whale at position (x, y, z = 15 m) and a PAM system receiving at the C_1 position. 28
Figure 19. Instantaneous maps of detection probability for 1 single hydrophone receiving at C1 position, 20-m depth, with a CCBD (A) or a TFBD detector (B), and for a circular array with a CCBD (C) or a TFBD (D), the bathymetry is shown in E. 29

Figure 20.Instantaneous maps of the detection probability or PAM system (August 5th, 23:00) at C1 position, 1 single hydrophone receiving with a CCBD(A) or a TFBD detector (B), a circular array with CCBD (C) and TFBD (D) with 5 ships (red triangles)..... 30
Figure 21. Residual masking areas (angular sectors θ_1 and θ_2) of NARW upcall by the shipping noise when a circular array is used at C1 and the ship and calling whale are co-located within the same beam. 31
Figure 22. Maps of the proportion of time from August 1st to October 30th 2018 where the NARW upcall detection probability is more than 0.5 for PAM systems located at the 6 studied coastal positions (C1 (left) to C6 (right)). 32
Figure 23. Maps of the good detection area (yellow) for single-hydrophone or hydrophone-array PAM systems with TFBD PAM systems located at the 6 studied coastal positions (C1 (left) to C6 (right)). 34
Figure 24. Superposition of the good detection area of an array of hydrophone at C1 with the approximate envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (from https://whalemap.ocean.dal.ca/WhaleMap/). 35
Figure 25. Good detection area for one PAM system at C1 (left) or at C4 (middle), and combination (right). North-South dimension of IRB is 100 km and East-West dimension is 96 km. 37
Figure 26. Superposition of the good detection area of hydrophone arrays at C1 and C4 with the envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (from https://whalemap.ocean.dal.ca/WhaleMap/).. 37
Figure 27. Good detection area for one PAM system at C1 (left), at C2 (middle) and for simultaneous detection at the 2 PAM systems (right). North-South dimension of IRB is 44 km and East-West dimension is 66 km. 39
Figure 28. Superposition of the area for possible localization of hydrophone-array PAM systems at positions C1 and C2 with the envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (while line). 39
Figure 29. Error ellipse for NARW upcall localization with 2 hydrophone-array PAM systems located at positions C1 and C2. 40
Figure 30.Theoretical SNR and processing gain (PG) of the STFT as a function of the window length L. 50
Figure 31. Application of the STFT real recordings containing two whistles of beluga whales (*Delphinapterus leucas*) with constant frequency. 51
Figure 32. Application of the STFT on real recordings containing three whistles of bottlenose dolphins (*Tursiops truncatus*) with a portion of the sound representing a linear frequency-modulated signal. 52
Figure 33. TFBD processing chain. 55
Figure 34. Loss in processing gain as a function of the level of misfit between the true and the used references, thick black curve ΔPG, thin black curves: $\Delta PG \pm \sigma \Delta PG$. 56

ACRONYMS

AIS:	Automatic Identification System
CCBD:	Cross Correlation Based Detector
DFO:	Department of Fisheries and Oceans
DoA:	Direction of Arrival
EBD:	Energy Based Detector
ESD:	Effective Surface of Detection
ERD:	Effective Range of Detection
FFT:	Fast Fourier Transform
FGA:	Feeding Ground Area
GSL:	Gulf of St. Lawrence
GDA:	Good Detection Area
IRB:	Inclusive Rectangular Box
LFM:	Linear Frequency Modulation
NARW:	North Atlantic Right Whale
OOS:	Ocean Observing System
PAM:	Passive Acoustic Monitoring
P_d:	Probability of detection
P_{fa}:	Probability of false alarm
PG:	Processing Gain
SAG:	Surface Active Group
SL:	Source Level
SNR:	Signal to Noise Ratio
STFT:	Short-Time Fourier Transform
TFBD:	Time Frequency Based Detector
TL:	Transmission Loss

ABSTRACT

Gervaise, C., Simard, Y., Aulanier, F., and Roy, N. 2019. Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 3345: ix + 58 p.

This report addresses the problem of detecting and localizing whales located in a feeding ground bordering a noisy seaway with passive acoustic monitoring (PAM) systems reporting the detections and the localization in real time. This general problem is methodologically addressed using the detection theory, and the relevant formulas for assessing the PAM performance are provided. The method is applied to the special case of North Atlantic right whales (NARW) in the feeding ground area (FGA) off Gaspé in the Gulf of St. Lawrence. The relevant scales at play in the area are first established before setting the parameters of realistic simulations fed with the actual shipping traffic, transmission loss (TL) from a regional acoustic propagation model accounting for the bathymetric and environmental structures, estimated NARW upcall source level (SL), and measured ship SLs of the local fleet. The tested scenarios include single hydrophone and hydrophone-array PAM systems cabled to the coast at six pre-defined positions where simple deployments with short cables are possible. The simultaneous use of two coastal stations is considered to achieve NARW localization or to maximize the area of good detection. Three families of NARW upcall detectors, providing different levels of minimal signal to noise ratio (SNR) for detection, are compared for each scenario. A low false-alarm rate of 1 per day is imposed and a probability of detection greater than 0.5 is retained as indicator of good performance. Time Frequency Based Detector (TFBD) prevailed as the best trade-off between performance of detection and robustness against natural variability of NARW upcalls. For PAM systems equipped with a single hydrophone, the Effective Surface of Detection (ESD, 256 ± 91 km^2, 1.6 % of the FGA surface) and the Effective Range of Detection (ERD, 12 ± 2 km) are small. A coastal single-hydrophone PAM system failed to cover the entire width of the FGA from the Gaspesian coast. A total of 62 single-hydrophone offshore PAM buoys or gliders are required to monitor the entire FGA. The use of a hydrophone array provides an ESD that is 14 times larger than the ESD of single-hydrophone PAM systems. A coastal PAM system equipped with a hydrophone array and TFBD detects NARW upcalls over an area of 2664 km^2 (15 % of the FGA surface). The combination of 2 such PAM systems, optimally positioned on the Gaspesian coast, detects NARW upcalls over an area of 4 240 km^2 (25 % of the FGA surface); the size of the corresponding Inclusive Rectangular Box (IRB) of the Good Detection Area (GDA) is then 100 km × 96 km. A combination of 5 coastal PAM systems equipped with an hydrophone array and TFBD, plus one autonomous offshore PAM buoy, well located, is able to monitor the whole surface of the FGA. For localizing and tracking NARWs at a hotspot of interaction with human activities within the FGA, 2 cooperating coastal PAM systems equipped with a hydrophone array and TFBD, 20-km apart between Gaspé and Percé, can localize and track NARW upcalls within an area of 30 km ×30 km, with a North-South accuracy of 1.5 km and an East-West accuracy of 5 km.

RÉSUMÉ

Gervaise, C., Simard, Y., Aulanier, F., and Roy, N. 2019. Optimal passive acoustic systems for real-time detection and localization of North Atlantic right whales in their feeding ground off Gaspé in the Gulf of St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci. 3345: ix + 58 p.

Ce rapport traite du problème de détection des baleines dans une aire d'alimentation à proximité d'une voie de navigation bruyante au moyen d'un système d'acoustique passive (PAM) transmettant ses détections et ses localisations en temps réel. Cette problématique est abordée méthodiquement à l'aide de la théorie de la détection et les formules pertinentes pour estimer la performance PAM sont énoncées. La méthode est appliquée au cas particulier de la détection de baleines noires du nord Atlantique (NARW) dans l'aire d'alimentation au large de Gaspé dans le Golfe du Saint-Laurent. Les dimensions pertinentes au problème dans cette région sont d'abord rappelées avant de déterminer les paramètres de simulations réalistes, nourries par le trafic observé, les pertes par transmission (TL) issues d'un modèle régional de propagation acoustique considérant la structure bathymétrique et environnementale, le niveau à la source (SL) estimé des sons de contact des NARWs ainsi que les niveaux SL mesurés des navires de la flotte locale. Les scénarios testés incluent des systèmes PAM à hydrophone unique et à réseaux d'hydrophones reliés à la côte par un câble. Ces systèmes sont pré-positionnés en 6 sites offrant des facilités de déploiement simples avec des câbles courts. Il est envisagé de faire fonctionner simultanément deux systèmes pour permettre la localisation ou optimiser la surface de bonne détection. Trois familles de détecteurs des sons de contact de NARWs, fournissant différents niveaux minimaux de rapports signal-sur-bruit (SNR) pour la détection, sont comparées pour chaque scénario. Un faible taux de fausses détections de 1 par jour est imposé et une probabilité de détection supérieure à 0.5 est utilisée comme indicateur de bonne performance. Le détecteur basé sur le patron temps-fréquence (TFBD) s'impose comme le meilleur compromis entre la performance de localisation et la robustesse à la variabilité naturelle des sons de contact de NARWs. Pour un système PAM utilisant un seul hydrophone, la surface efficace de détection (ESD, 256± 91 km^2, 1.6% de la surface de l'aire d'alimentation (FGA)) et le rayon efficace de détection (ERD, 12± 2 km) sont faibles et insuffisants pour couvrir toute la largeur de la FGA depuis la côte Gaspésienne. Un total de 62 systèmes PAM à 1 hydrophone (bouées ou planeurs) déployés au large dans le FGA sont nécessaires pour couvrir toute sa surface. L'utilisation d'un réseau d'hydrophones permet de multiplier par un facteur 14 la surface efficace de détection d'un système à 1 hydrophone. Un système PAM côtier avec un réseau d'hydrophones et un TFBD permet de bien détecter les sons de contact de NARW sur une surface de 2 664 km^2, soit 15% du FGA. L'usage de 2 systèmes PAM avec un réseau d'hydrophones, optimalement positionnés sur la côte gaspésienne, permet une bonne détection sur 4 240 km^2, soit 25% du FGA répartis au sein d'un secteur rectangulaire de 100 km × 96 km. Cinq systèmes PAM à réseaux d'hydrophones plus une bouée PAM autonome, bien positionnés, permettent de couvrir tout le FGA. Pour le positionnement et le suivi à la trace des NARWs dans des zones sensibles d'interactions avec les activités humaines à l'intérieur du FGA, un ensemble de 2 systèmes PAM à réseau d'hydrophones et TBFD coopérants, espacés de 20 km sur la côte Gaspésienne, permet de localiser les NARWs dans une région de 30 km × 30 km avec une précision de 5 km dans la direction est-ouest et 1.5 km dans la direction nord-sud.

1. INTRODUCTION

1.1 NARW current context and threats in GSL

North Atlantic right whale (*Eubalaena glacialis*) (NARW) is listed as "endangered" according to the Canadian Species at Risk Act (SARA) (DFO 2014). In 2017, the population was estimated by some models to about 411 individuals (Pettis et al. 2018). Historically, it was mainly distributed off eastern USA coast from Florida, in the south, up to the Canadian coast off Labrador, in the north (Cole et al. 2013, Davis et al. 2017, DFO 2014). Individuals have been sporadically sighted in Eastern Atlantic (*Ibid.*). In the last decades, its annual distribution reflected a seasonal migration pattern between the winter breeding grounds in the south and the summer feeding grounds in the north, notably the Gulf of Maine and the Bay of Fundy(Cole et al. 2013). This pattern changed in early 2010s, when the whales deserted these traditional feeding zones (Davis et al. 2017), where their main prey declined, possibly in response to hydro-climatic changes (Grieve et al. 2017, Meyer-Gutbrod and Greene 2018, Meyer-Gutbrod et al. 2018). Further north, in southern Gulf of St. Lawrence (GSL), NARW observations have considerably increased in last few years (DFO 2018, Simard et al. 2019) (see Figure 1).

This latter part of Northwest Atlantic is crossed by the main shipping traffic between the Atlantic and the Great Lakes (Simard et al. 2014). The area is also the site of an intensive fixed-gear fishing activity, notably for show crab. In 2017, the marked increase in NARW occurrence in GSL resulted in the death of 12 individuals, notably due to collision with ships and entanglement in fishing gears (Daoust et al. 2017). No mortalities in GSL were reported in 2018. In 2019, 9 mortalities were reported for June and July.

The Canadian Government has since taken action to minimize this NARW mortality, by limiting fixed-gear fishing and reducing ship speed in areas of high NARW occurrence (DFO 2019, Transport Canada 2019). This latter management measure includes a large static speed restriction zone during NARW presence season, and dynamic management sectors in the main seaways where the speed restriction is triggered by the presence of a whale. This information is presently obtained from visual observations from aircrafts or boats. Such observations require adequate meteorological conditions and are limited to the daylight period. To increase the surveillance effort, real-time acoustic detection techniques based on NARW specific calls, known as PAM (Passive Acoustic Monitoring), are considered. Such techniques are presently used from fixed Ocean Observing Systems (OOS) buoys (DFO OOS Viking-WOW buoys, https://ogsl.ca/viking/) and moving platforms (Slocum glider, https://whalemap.ocean.dal.ca/WhaleMap/). The present study examines the actual possibilities of different PAM systems for real-time detection of NARW specific calls in their feeding ground off Gaspé peninsula.

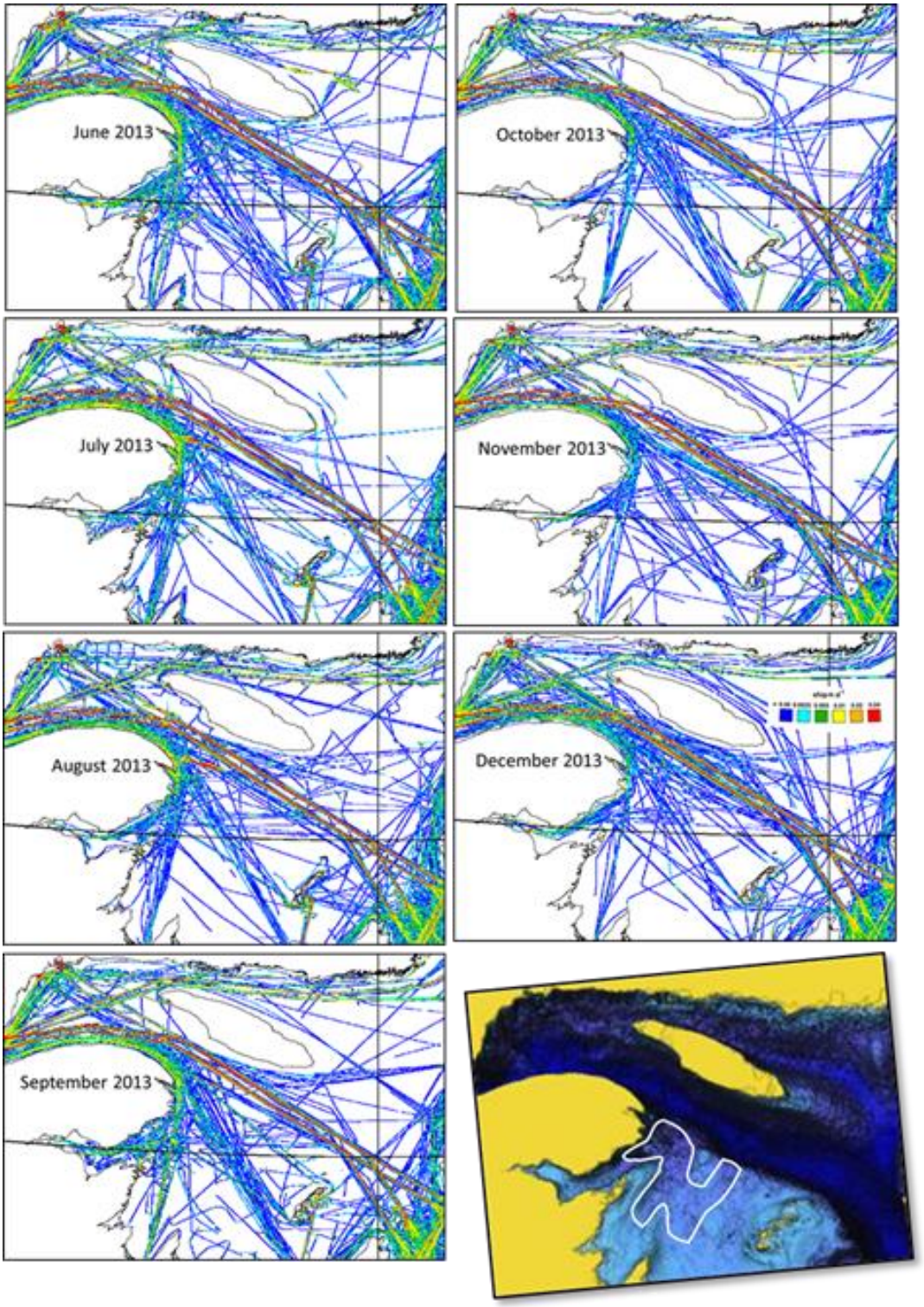

Figure 1. Context of the feeding ground area off Gaspé, with the monthly shipping traffic from June to December from Simard et al. (2014), and the three-dimensional underwater topography with the envelope (white line) of the NARW feeding ground basin, where sightings are concentrated from June to December (from https://whalemap.ocean.dal.ca/WhaleMap/).

1.2 Objectives

In November 2018, CHORUS Research Institute (www.chorusacoustics.com) was mandated by DFO to perform a desktop study of the abilities of several PAM systems to detect and localize NARW in real-time, from their specific calls, in their feeding ground off Gaspé. In present study, by real-time we refer to a maximum of 1-h interval after the occurrence of the event. The study was expected to rely upon a strong theoretical background and to be as realistic as possible. Its goals were to assess the detection areas and the abilities of the tested PAM systems to detect and localize the NARW calls, and to dimension the *ad hoc* PAM networks required to accomplish these tasks.

At six predefined positions on the Gaspesian coast, we considered two types of PAM systems:

- an isotropic single hydrophone,
- a hydrophone array offering directional hearing capability for separating ship noise and NARW emissions, thanks to beamforming.

Each PAM system, with or without beamforming, was examined using three types of detection algorithms, presented in descending order of theoretical performance:

- the Cross Correlation Based Detector (CCBD), based on the cross-correlation with a copy of the signal emitted by NARWs,
- the Time Frequency Based Detector (TFBD), based on the degree of match with a series of attributes or a stereotype of the signal emitted by NARWs,
- the Energy Based Detector (EBD), based on the amount of signal energy recorded in the bandwidth of NARW emissions.

The six potential positions of PAM systems were chosen such that:

- there exist possibilities to connect the PAM system on shore, and
- a depth of at least 20 m can be reached at less than 1 km from the coast.

These different cases (position, single hydrophone or hydrophone array, type of detector) were compared to determine the *ad hoc* system suitable for PAM of NARW in their feeding ground off Gaspé with their respective detection and localization abilities; '*ad hoc*' here indicates that the PAM system maximizes the areas of good detection or good localization.

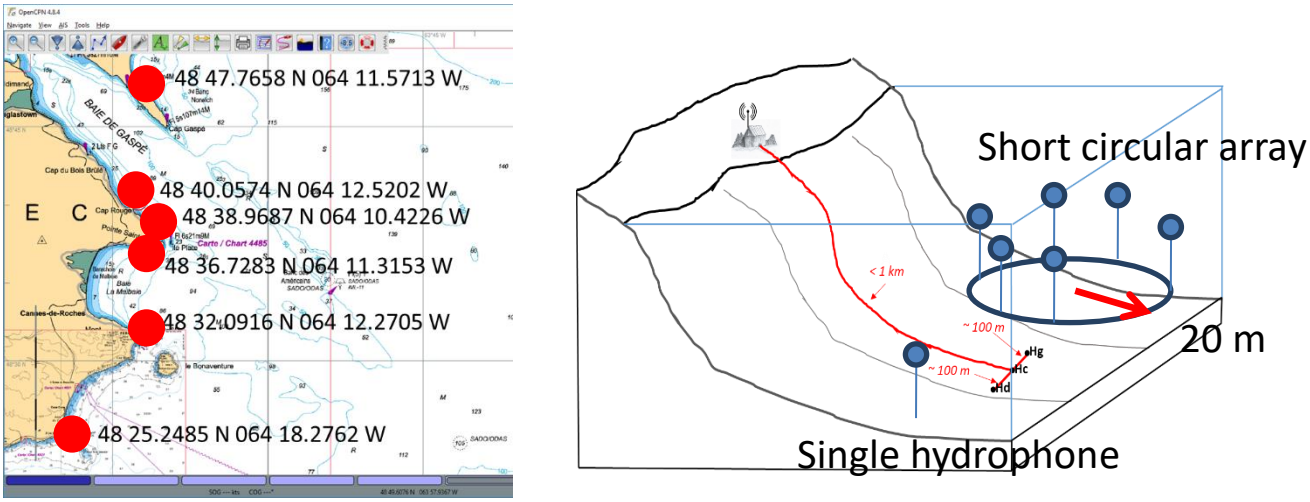

Figure 2. Positions of the six coastal PAM systems compared for NARW detection and localization.

1.3 Content of the report

The Material and Methods section is divided in the following 3 sub-sections.

Section 2.1 presents the approach used to estimate the propagation range of individual-ship noise in presence of ambient noise, and the detection range of a NARW upcall facing masking by ship noise using simplified but realistic assumption about the spatial-temporal distribution of the shipping. It clearly demonstrates that the sounds radiated by individual ships dominate the diffused ambient noise (*sensu* NRC 2003, Wenz 1962). This discrete space-time location of the dominant ocean noise (sensu NRC 2003) source offers the opportunity to exploit beamforming filtration from a hydrophone array to enhance the signal to noise ratio (SNR) and to increase the detection range.

Section 2.2 sets all the inputs and assumptions required to perform a more realistic evaluation. In section 2.2.2, we characterize NARW 'upcalls' or 'contact calls' (chirps, 1 s, ~[100 Hz-200 Hz], 165 dB re 1µPa_{rms} @ 1m) as the best candidate for PAM of NARW because of their recurrence rate, their relatively large propagation range due to their low frequency and relatively loud source level (SL) and their specific time-frequency modulation allowing their recognition among calls from other species. Section 2.2.3 presents the August-October 2018 AIS shipping traffic in the feeding ground area off Gaspé, and how we set the speed and SL of individual ships based on Simard et al. (2016). Section2.2.4 describes the regional model and environmental data used to compute the transmission loss (TL) in a 100-km radius around the receiving positions (Aulanier et al. 2016a) around the receiving positions, and the diffuse ambient noise from Wenz (1962)'s empirical model.

Section 2.3 presents the three detectors (i.e. detection algorithms) considered to detect the NARW-upcall, in descending order of performance (CCBD, TFBD, EBD). We derive the PAM system performance (Probability of Detection (P_d), Probability of False Alarm (P_{fa}), SNR) for each detector. To insure rigorous comparisons between the three detectors, the performances are presented considering the same SNR between the NARW upcall and the total noise. A functioning

point is defined ($P_d = 0.5$, $P_{fa} = 2\times10^{-6}$corresponding to 1 false alarm per day, validated by a human supervisor), and minimum SNRs to achieve this functioning point are given. The section ends with the case of a hydrophone array as PAM system (section 2.3.6). We derive the processing gain (PG) produced by the use of beamforming and the SNR expression for a NARW upcall embedded in diffuse ambient noise and discrete noise fields radiated by individual ships, with and without beamforming PG.

The Results (section 3) compares the detection performance estimates with maps of the proportion of time where P_d is higher than 0.5 and $P_{fa} = 2\times10^{-6}$, for the different scenarios (position, single hydrophone, hydrophone array, type of detector) within 100-km radius around the receiving PAM system. These maps are obtained by simulating the true shipping obtained from AIS data of the 2018 quarter of August 1st to October 30th (675 ships, 2490 transits) divided in 13 392 time steps of 10 min. The effective surfaces of good detection area (GDA) and the size of its inclusive rectangular box (IRB) are then computed for each (position, single hydrophone, hydrophone array, detector type) and recorded in dedicated tables.

The Discussion (section 4) compares the relative advantages and shortcomings of *ad hoc* solutions for PAM of NARWs upcalls in the FGA off Gaspé. PAM solutions allowing NARW localization are highlighted. Guidelines are given to choose the best PAM network to optimize the GDA or the precise localization area and to monitor the entire FGA.

2. MATERIALS AND METHODS

2.1 Setting the orders of magnitude

The objective of this part is to set the orders of magnitude that are used in this report. We develop a simplified approach to present the propagation range of the ship noise (i.e. the range over which the noise radiated by a given ship exceeds the diffuse ambient noise), and the NARW upcall detection range given masking by such ship noise.

The geography and geometry of the area are shown in Figure 3. The FGA off Gaspé forms a trapezoid oriented southwest to northeast with a 103-km southern base, a 144-km northern base and a height of 160 km. The FGA is ~20 km off the coastline and 30 km from the St. Lawrence Seaway. The FGA is also crossed by local shipping traffic.

The AIS data from August 1st 2018 to October 30th 2018 indicate (see section 2.2.3 for details) that there are ~ 6 to 7 ships present simultaneously inside the FGA or within 50-km from its contour.

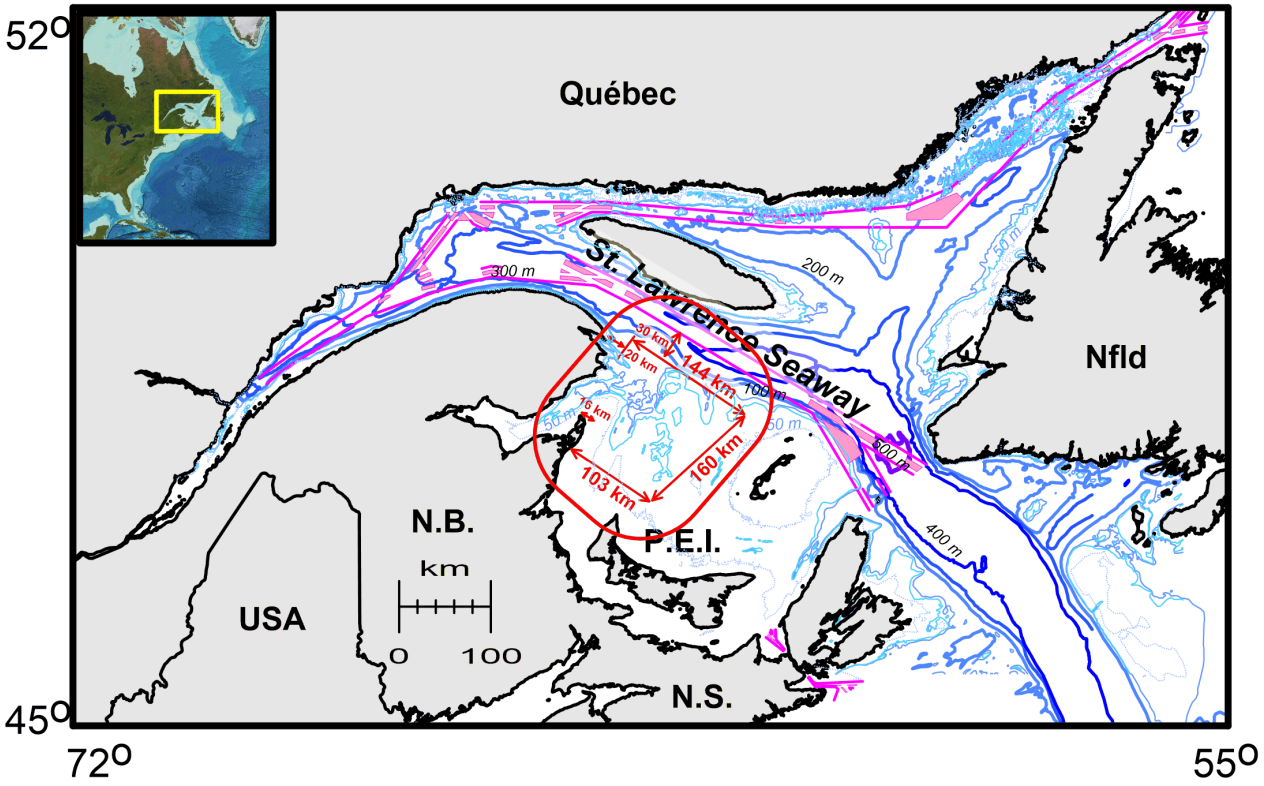

Figure 3. Map and geometry of the study area, showing the shipping routes (pink) and bathymetry of the feeding ground area where NARW sightings are concentrated. (from https://whalemap.ocean.dal.ca/WhaleMap/).

To assess the acoustic range of a ship in diffuse ambient noise (*sensu* Wenz, (1962)) and the range of NARW upcalls, we make the following assumptions:

- Main bandwidth of NARW upcalls (see Parks et al. 2011) [100 Hz, 200 Hz]
- Typical wideband SL of a ship (Simard et al. 2016) 185 dB re 1µPa @ 1m
- Typical narrow band ship SL at 150 Hz (Simard et al. 2016) 165 dB re 1µPa^2/Hz @1m
- Typical wideband SL of a NARW upcall (Clark et al. 2011) 165 dB re 1µPa @ 1m
- Typical diffuse wideband [100, 200]Hz low noise level (Wenz 1962) 68 dB re 1µPa
- Typical diffuse narrow band low noise level (Wenz 1962) 48 dB re 1µPa
- Transmission Loss model 20 $\log_{10}(r)$
- SNR (P_d = 0.5, 1 false alarm/day, TFBD) 16 dB

Under these assumptions,

- at 10 dB SNR condition, the noise radiated by a ship exceeds the diffuse wideband Wenz noise up to a range of 224 km ($10^{\frac{185-68-10}{20}}/1000$),
- at 10 dB SNR condition, the NARW upcall can exceed diffuse narrow band Wenz noise up- to a range of 224 km ($10^{\frac{165-48-10}{20}}/1000$).

When an individual ship is present at a distance R_s from the PAM system, it becomes the main noise contributor within a radius of 224 km and the NARW detection range, R_w, follows the relation $R_w = 1/20 \times R_s$ (see section2.3.5)[1].

These simplified computations demonstrate that, in the FGA off Gaspé,

- the range of NARW detection under pristine diffuse ambient noise is of the same order of magnitude than the size of the FGA,
- the noise limiting the detection of NARW upcalls should not be considered as diffuse ambient noise but coming from several punctual individual ships,
- the shipping noise drastically reduces the NARW upcalls detection range, which is highly dependent on the specific arrangement of ships in and around the FGA at a given time t,
- as the sources of noise are punctual ships, the use of a hydrophone array to perform a directional hearing should be considered.

2.2 Acoustic and environmental inputs

2.2.1 Objectives

In this section we set all the inputs and assumptions required for realistic NARW upcall detection performance estimates using numerical simulation modeling. First, we describe the NARW upcalls. Then we document the shipping traffic in the FGA off Gaspé with AIS data collected by DFO from August 1st2018 to October 30th2018. We indicate how we set the speed and SLs of individual ships using Simard et al. (2016) 's St. Lawrence fleet model. We end this part by presenting the PAM performance simulation method with the oceano-geo-acoustic parameters used to compute the transmission loss (TL) in a disc of 200 km in diameter around the position of the PAM system within the coast off Gaspé (Aulanier et al. 2016a, Aulanier et al. 2016b) and the diffuse ambient noise from the empirical model of Wenz (1962).

2.2.2 NARW acoustics

NARW sonic production has been described by comparison with other similar species (South Pacific right whales, North Pacific right whales) and from NARW previous and extensive studies since 2000 in their summering habitats of the north-eastern coasts of USA (Cape Code, Stellwagen Bank, Gulf of Maine) and Canada (Scotian shelf, Bay of Fundy) including recently the Gulf of St. Lawrence.

NARW calls fall into three categories presenting quite different features:

- The upcall, a stereotyped emission used to maintain contact between individuals (Clark 1982) over several miles. The upcall is a 1-s long increasing frequency modulation between 100 Hz and 200 Hz where these parameters are affected by natural variability (Clark 1982, Matthews et al. 2001, Mussoline et al. 2012, Parks et al. 2011, Simard et al. 2019) (Figure 4).

---

[1] $SL_w + 10log_{10}(L_{fft}) - 20log_{10}(R_w) - (SL_s - 10log_{10}(B_w) - 20log_{10}(R_s)) = 16$

- The gunshot, a stereotyped emission appearing as a short and loud impulse extending from very low frequency (50 Hz) to mid frequency (4 kHz). The gunshot has been documented to be emitted during Surface Active Group (SAG), when several individuals cluster for socializing, and for reproductive advertisement (Parks et al. 2012, Parks and Tyack 2005).
- Sonic emissions during SAG are more frequent and more diverse (tonal, moans, impulses…). As the communication range is short in a SAG, the SL of some sounds may be weaker than the gunshot and the upcall (Parks et al. 2011, Parks and Tyack 2005, Trygonis et al. 2013).

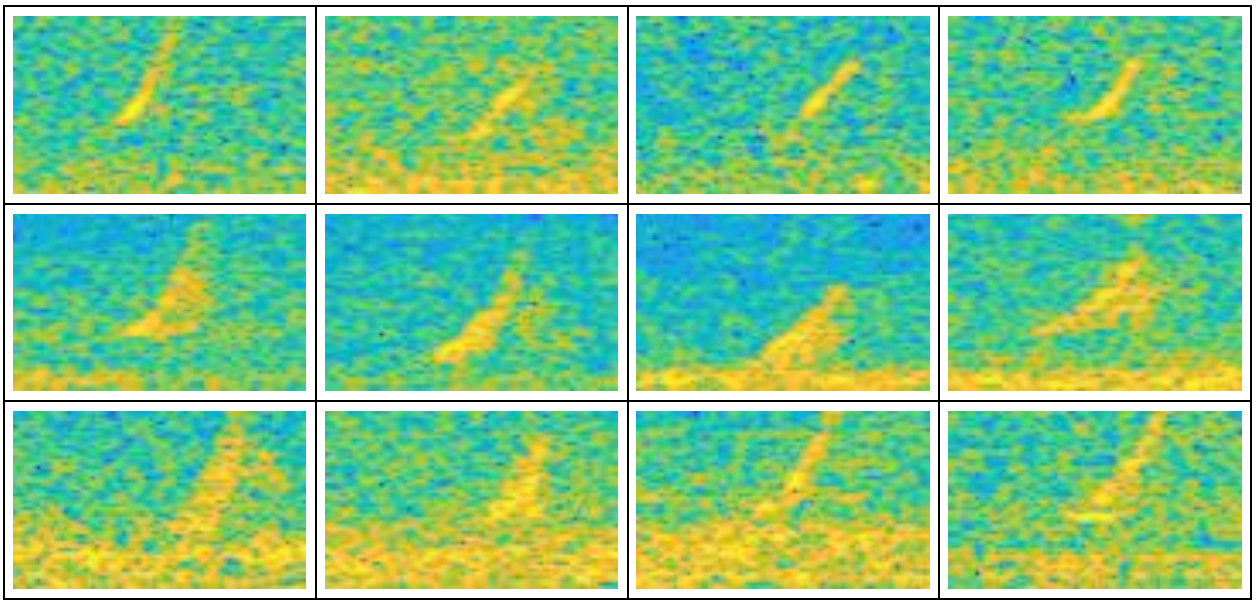

Figure 4. Spectrogram examples of NARW upcalls recorded in southern GSL in 2018. Y axis: frequency from 100 to 200 Hz; X axis: time 3 s. Resolution 3.9 Hz ×32 ms.

Although all NARW calls may be useful for monitoring purposes, the upcalls are the best candidates because they are stereotyped and can be used to identify the species. They are emitted regularly at a variable but significant rate, and their low frequency coupled with the frequency modulation ensures a long range propagation and detection (Clark et al. 2010). Upcalls were proposed and used by Mellinger et al. (2007), Clark et al. (2010), and Baumgartner et al. (2013) for NARW PAM (including from glider platforms) on the north-eastern coast of USA. Upcall detection ranges were reported to be around 10 km within an urbanized area (Clark et al. 2010) and more than 50 km for North Pacific right whales within the more pristine environment of Bering Sea (Munger et al. 2011, Wiggins et al. 2004).

The upcall features used in this present report are from (Clark et al. 2011, Parks and Tyack 2005):

- Upsweep frequency modulation,
- Duration: 1 s
- Main frequency support: [100 Hz, 200 Hz],
- SL: 165 ± 3 dB re 1µPa @ 1m.

The SL from Clark et al. (2011) is conservative[1] since Munger et al. (2011) report SL ranging within 177 ±1 dB re 1µPa @ 1m. for North Pacific right whales.

2.2.3 Marine traffic within the dynamic management zones

AIS messages from the ships transiting within the feeding ground area in 2018 (August 1st to October 30th) were collected by DFO (cf. Simard et al. 2014) . They are reported on a daily basis in Table 1. The traffic density was 27 ± 6 ships per day. During the period, 675 different ships performed 2490 transits in the FGA. The cumulated time of presence of ships in the area was 621 days, the average time of transit inside the area per ship was 6 hours and ~6 ships were present at the same time within the FGA. The traffic is mostly in the St. Laurence Seaway (60.68%) but there is also a more local and costal traffic (39.32%).

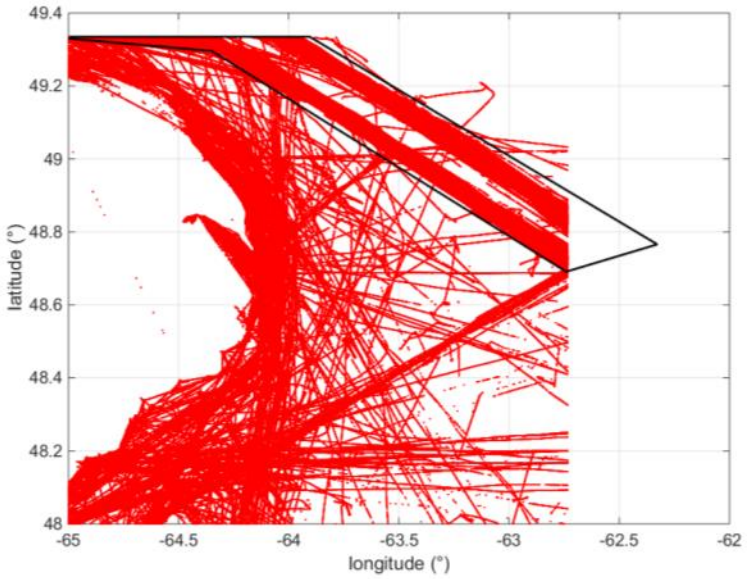

Figure 5. Ships positions from AIS data from August 1st 2018 to October 30th. Black line: St. Laurence Seaway.

[1] In the sense of minimizing the detection range.

Table 1. AIS traffic density (vessel per day) in the FGA off Gaspé in 2018.

Day	Aug.	Sept.	Oct.
1	xxx	23	29
2	37	26	23
3	33	23	35
4	37	22	33
5	35	22	23
6	21	33	32
7	30	21	42
8	34	20	25
9	28	22	20
10	26	27	27
11	26	27	31
12	33	24	20
13	38	21	30
14	33	26	29
15	31	xxx	25
16	40	xxx	23

Day	Aug.	Sept.	Oct.
17	29	xxx	29
18	34	xxx	21
19	31	xxx	17
20	30	xxx	25
21	24	xxx	31
22	16	xxx	26
23	16	xxx	23
24	32	xxx	28
25	38	xxx	39
26	26	12	25
27	28	32	28
28	34	33	29
29	21	31	30
30	23	27	27
31	30	xxx	xxx
Total	**894**	**472**	**825**
Mean	**29**	**25**	**27**

2.2.4 Simulation of acoustic propagation and ambient noise within the study area

TL were computed using a 2.5-D model (i.e. 3D TL field reconstruction from the concatenation of vertical slices between sources and receivers) implemented by DFO for the St. Lawrence and Canadian Arctic (Aulanier et al. 2016a, Aulanier et al. 2017). The propagation was simulated between the sources (ship location or hypothetical whales) and receivers (candidate hydrophones or hypothetical whales) with the classical RAM code (http://oalib.hlsresearch.com/PE/RAM/) (Collins 1993), a 2D parabolic equation (PE) model fed with the high-resolution bathymetry from the DFO Canadian Hydrographic Service, range-dependent environmental conditions from several DFO operational models and published data (Loring and Nota 1973, Senneville and Lefaivre 2015).

To get NARW upcall detection probability, TL maps were computed for a 1-km × 1-km grid of potential sources within a disc of 200 km in diameter around the PAM system at 3 frequencies (100 Hz, 150 Hz, and 200 Hz). The receivers were 1 m above the bottom depth at the positions of the PAM system (see Figure 2) and their depths ranged from 17 m to 42 m. The NARW call depth was set to 15 m based on the observations of the upper water column calling from NARWs tracked with D-tags (Parks et al. 2011).

Sound speed profiles over the water column and the geo-acoustic parameters of the seafloor were derived using the data and methodology described in Aulanier et al. (2016a, p. 3).

The diffuse ambient noise level was set to 69 dB re 1µPa^2/Hz at 150 Hz, corresponding to a mid-traffic condition of Wenz's empirical model (Wenz 1962).

2.3 Performances of three NARW upcall detectors from single-hydrophone or hydrophone-array setups

2.3.1 Objectives

This section presents the theoretical background for the detection of NARW upcall embedded in ship noise. The NARW upcall detection problem is presented and three detector types (CCBD, TFBD, and EBD) are examined. Performance analysis is made for each detector type using the three following criteria: Probability of Detection (P_d), Probability of False Alarm (P_{fa}), and Signal to Noise ratio (SNR). For rigorous comparisons between the detector types, the performances are presented with a unified and consistent theoretical framework, based on the same expression of SNR. The minimum SNRs to achieve functioning point defined as (P_d = 0.5, P_{fa} = 2×10^{-6}, corresponding to 1 false alarm per day, which can be eliminated by a human supervisor) are given for the three detectors. The section ends with the case of PAM systems using hydrophone arrays. We derive the processing gain (PG) produced by using beamforming techniques and the SNR required for detecting a NARW upcall embedded in diffuse ambient noise and individual ships, with or without using beamforming.

2.3.2 The data and detection problem

Let us consider an operational PAM sensor recording under noise-only conditions when a NARW upcall is absent (see Figure 6). In the specific context of the feeding ground area off Gaspé and according to the conclusions of section 2.1, the noise is the sum of the emissions of several individual ships and a diffuse ambient noise *sensu* Wenz (1962). This 'noise-only' case is referred to as hypothesis H_0. If a NARW upcall is present, the acoustic measurement (m(t)) is the sum of this upcall and the sum of the radiated noise from the ships and the diffuse ambient noise. This 'upcall present' case is referred to as hypothesis H_1:

- Hypothesis H_0: $m(t) = \sum_{N_s} s(i,t) + b(t) = b_{tot}(t)$
- Hypothesis H_1: $m(t) = uc(t) + \sum_{N_s} s(i,t) + b(t)$
-

where N_s stands for the number of ships present at the same time in the neighborhood, s(i,t) is the noise produced by ship of index i, b(t) is the diffuse ambient noise *sensu* Wenz(1962), b_{tot}(t) is the total noise and uc(t) is the upcall. The noise signals b(t) and s(i,t) are supposed to be random and normally distributed (0 mean, variance σ^2), whereas uc(t) is a deterministic signal:

$$uc(t) = 0 \; if \; t \notin [0,T], uc(t) = A\cos(2\pi f_{min} t + 2\pi \frac{f_{max} - f_{min}}{2T} t^2)$$

Noise level (σ^2) and upcall level ($A/\sqrt{2}$) are given in section 2.1 for a realistic configuration.

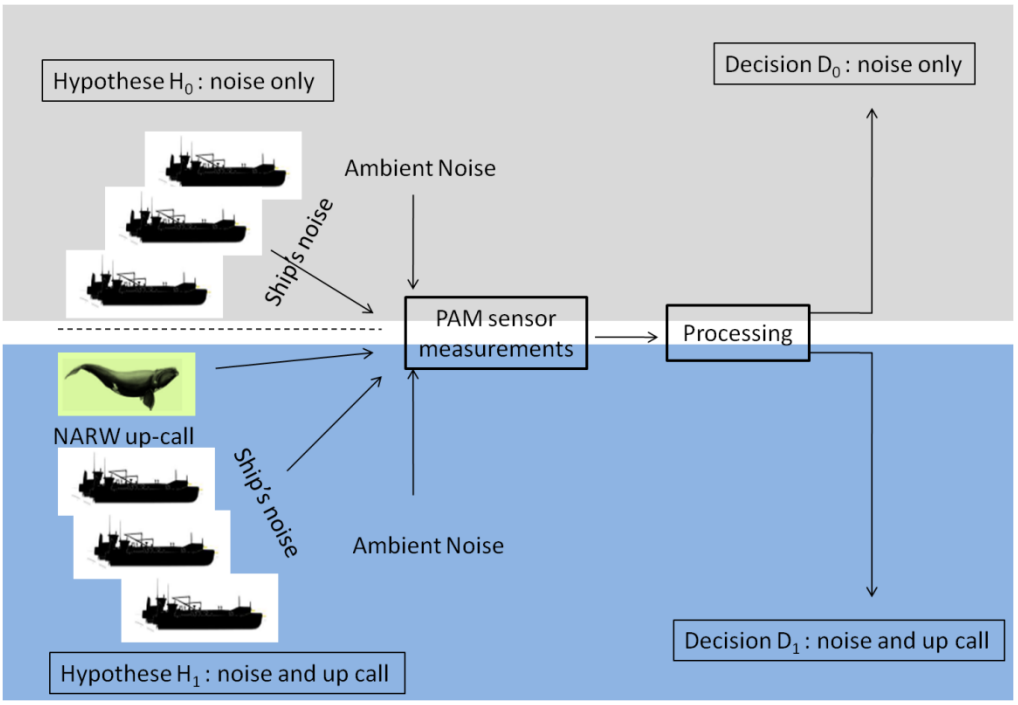

Figure 6. Environmental configuration of the detection of NARW upcalls embedded in ship noise

The detection problem is to decide if the upcall is present or absent based on the acoustic measurements m(t). Any detector follows the same 2-step logic: i) the measurements are processed to compute a statistical test T, and ii) the statistical test T is compared to a threshold λ which fixes the probability of false alarm (see Figure 7). If T is higher than λ, we decide that the upcall is present in the measurements and this decision is referred to as D_1; if T is lower than λ, we decide that the upcall is absent from the measurements and this decision is referred to as D_0. Depending on the initial hypotheses H_0 and H_1, the decisions D_0 or D_1 may be right or false. Table 2 sets the vocabulary.

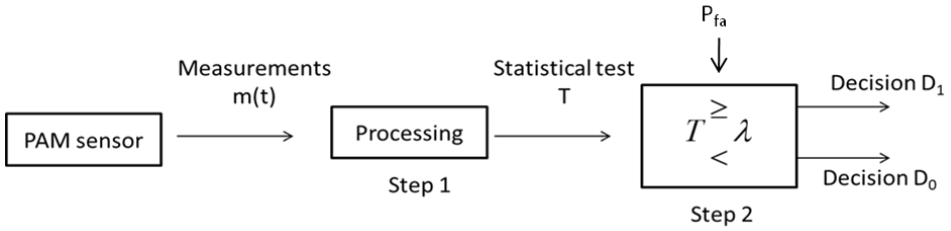

Figure 7. The two-step logic of the detection task.

Table 2. The four cases of the detection problem.

	Decision		
Hypothesis		D_0	D_1
	H_0	right decision of absence Probability 1-P_{fa}	wrong decision, false alarm Probability P_{fa}
	H_1	wrong decision, missed detection Probability 1-P_d	right decision of presence Probability P_d

We will promote detectors that ensure a constant false alarm rate and, for this family of detectors, we will look for the ones maximizing the detection probability to optimize the protection of NARWs (Kay 1998).

2.3.3 Choice of three NARW upcall detectors

A large range of solutions exists to detect short transient sound embedded in noise. The choice of the ideal solution is a huge task and results from a tradeoff between:

- the raw performances of detection (the best probability of detection for a given false alarm probability),
- the robustness versus the variability of the real data and the assumption used to derive the best detector (in the sense of optimizing the probability of detection)
- the computational load and the ability to run in real-time.

The scope of the present report is not to cover the whole range of solutions and to study in detail each solution but to identify the family of solutions (e.g. spectrogram based detectors, wavelet based detector, correlation with a template based detectors, etc.) that is suitable to the context of the detection of NARW upcalls in the eastern Canadian waters.

First, we chose to bound the performances of detection by an upper and a lower limit. To do so, we rely upon Kay (1998):

- When we do not assume any *a priori* knowledge about the waveform of the upcall except its bandwidth, the detector is based on the energy of the measurements contained within the bandwidth of the upcall (Kay 1998, p. 250), this Energy Based Detector (EBD) is simple, compatible with real-time processing, and robust, but it sets the lowest bound of the detection performances,
- When we assume to know the exact waveform of the received upcall, the detector is based on the cross-correlation of the measurements within the exact copy of the received upcall (Kay 1998, p. 101). This Cross Correlation Based Detector (CCBD) sets the highest bound of the detection performances but lakes of robusteness versus a deviation between the template used to compute the cross correlation and the true upcall emitted by the NARW and requires a large computational payload that can impair real-time abilities.

A bibliographic survey of solutions related by the scientific communities indicates efficient real-time solutions for NARW upcall within Western North Atlantic waters based on the processing of the spectrogram of the measurement (Baumgartner and Mussoline 2011, Gillespie 2004, Simard et al. 2019, Urazghildiiev and Clark 2007). These four solutions fall within the family of Time Frequency Based Detectors (TFBD). Lampert and O'Keefe (2010)rewieved more than 32 declinations of TFBD applied to the passive detection of cetaceans'sounds. Among these numerous algorithms, Urazghildiiev et al. (2009) demonstrates that the performances of TFBD using the expected time-frequency support of the upcall as *a priori* information tend towards that of CCBDs for low SNR conditions. For these reasons, we select as a third detector, a TFBD using the expected time-frequency support of the upcall. We expect (and we will demonstrate) that this TFBD sets an optimal tradeoff between performance, robustness and computational payload.

2.3.4 Description of the three NARW upcall detectors

Here, we describe the three NARW upcall detectors in decreasing order of performance. Figure 8 illustrates the detection steps for each of them.

CCBD, the first detector, is referred to as CCBD, for Cross Correlation Based Detector. Under certain conditions (see Kay (1998), p.101), CCBD is optimal in the sense that it maximizes the detection probability for a given probability of false alarm. CCBD requires to know the emitted waveform of the upcall and the eventual distortion due to acoustic propagation effects. Let ref(t) be this reference; then,

- step 1 calculates the cross-correlation of the measurement with the reference, and
- step 2 tests if the value of this cross-correlation exceeds the detection threshold for the presence of the upcall.

TFBD, the second detector, is referred to as TFBD for Time Frequency Based Detector. At this step in the report, nothing is known about its optimality. What is known is that: 1) according to Urazghildiiev et al. (2009), performances of TFBDs tend towards those of CCBDs for low SNR conditions, and 2) various TFBDs are operationally used (Baumgartner et al. 2013, Simard et al. 2019, Urazghildiiev et al. 2009). TFBD requires a time-frequency template or the attributes of the emitted upcall. Figure 8 and Figure 9 illustrate the TFBD process and its application on real data using the comparison with a time-frequency stereotype:

- step 1.1: the spectrogram of the data is computed with an adequate FFT window length (see annex 1 for the optimal FFT window length for a LFM signal such as NARW upcall).
- step 1.2: the spectrogram of the data is binarized for each time-frequency bin following the method proposed in (Dadouchi et al. 2013).
- step 1.3: the time-frequency template of the upcall (see arrow a1, plot C, Figure 9) is run over the binarized spectrogram and, for each time step, the number of time-frequency bins overlapping with the time-frequency template are counted.
- step 2: this matching sum or the proportion of match is tested against a threshold to detect or not detect the presence of an upcall.

EBD, the third detector, is referred as EBD for Energy Based Detector. EBD only requires knowing the bandwidth of the upcall. With only this limited a priori information, the EBD sets a worst bound of detection performance. Figure 8 illustrates the EBD process:

- step 1.1: the acoustic data are band-pass filtered for the upcall bandwidth,
- step 1.2: the energy of the data is computed for a period equal to the upcall duration,
- step 2: the energy is tested against a threshold to detect the presence of an upcall.

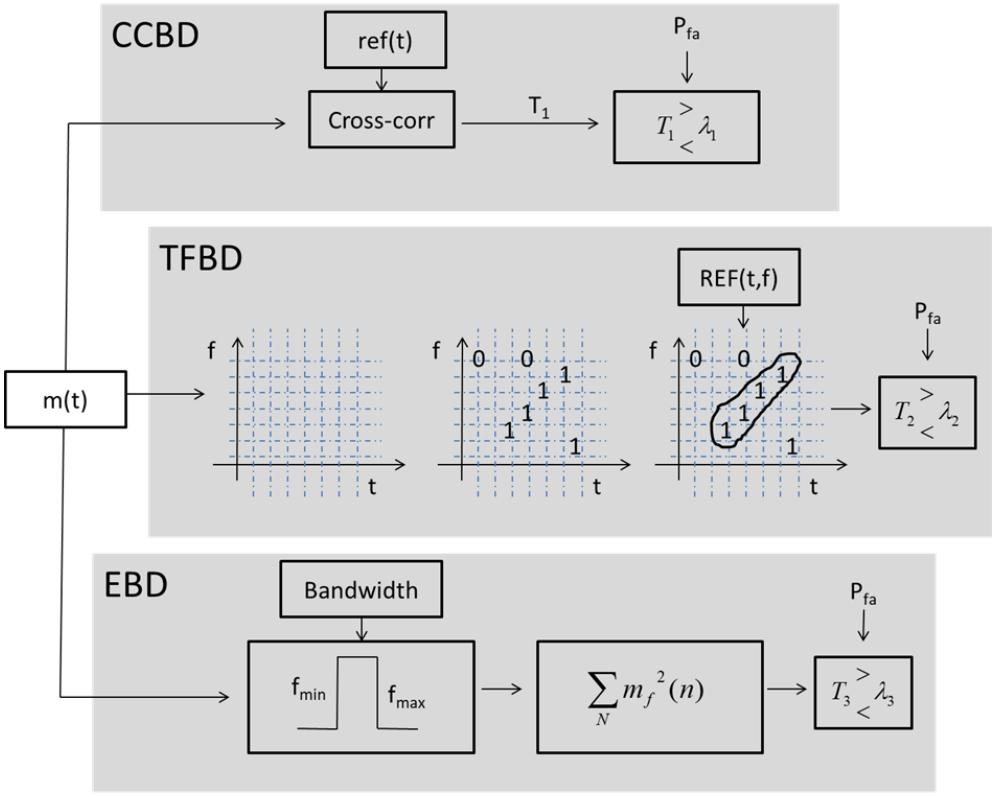

Figure 8. Schemes of the three detectors.

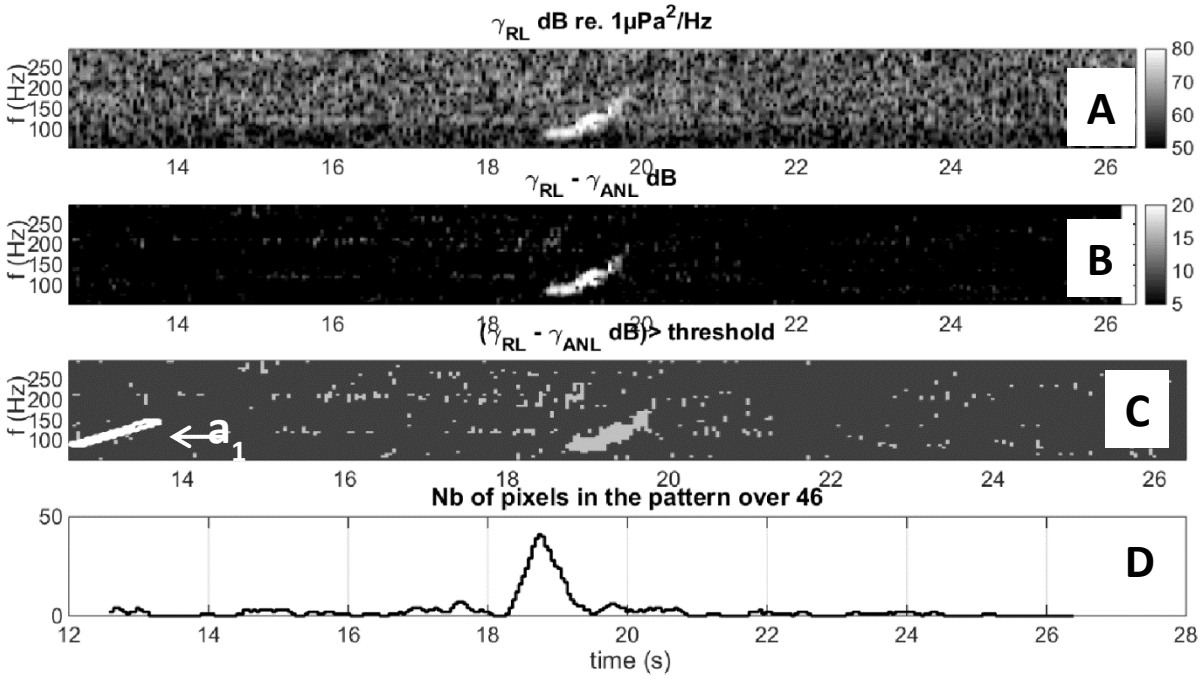

Figure 9. Step by step processing of the data for TFBD.
A: Computation of the spectrogram, B: Binarization of the spectrogram, C & D: Estimation of the proportion of match with the time-frequency template of the upcall. Application to real data collected in the Gaspé area (Gaspé B8A80989, Simard et al. 2019).

A qualitative assessment of the functioning and performances of the three detectors is presented in Figure 10, showing how each detector improves the SNR.. A negative SNR (-5 dB) synthetic data snapshot was created and used for CCBD, TFBD, and EBD tests. The outputs of step 2 of each detector are plotted and we quantify their differences under noise-only and upcall-present conditions. The higher is this difference, the better is the detection performance. The EBD shows a very low change (1.7/1.4 = 1.2) whereas the CCBD change is large (5500/500 = 11). The TFBD shows an intermediate change of 48/8 = 6. This level of change is mainly created by the computation of the spectrogram magnifying the presence of the upcall (see arrow c_2, Figure 10).

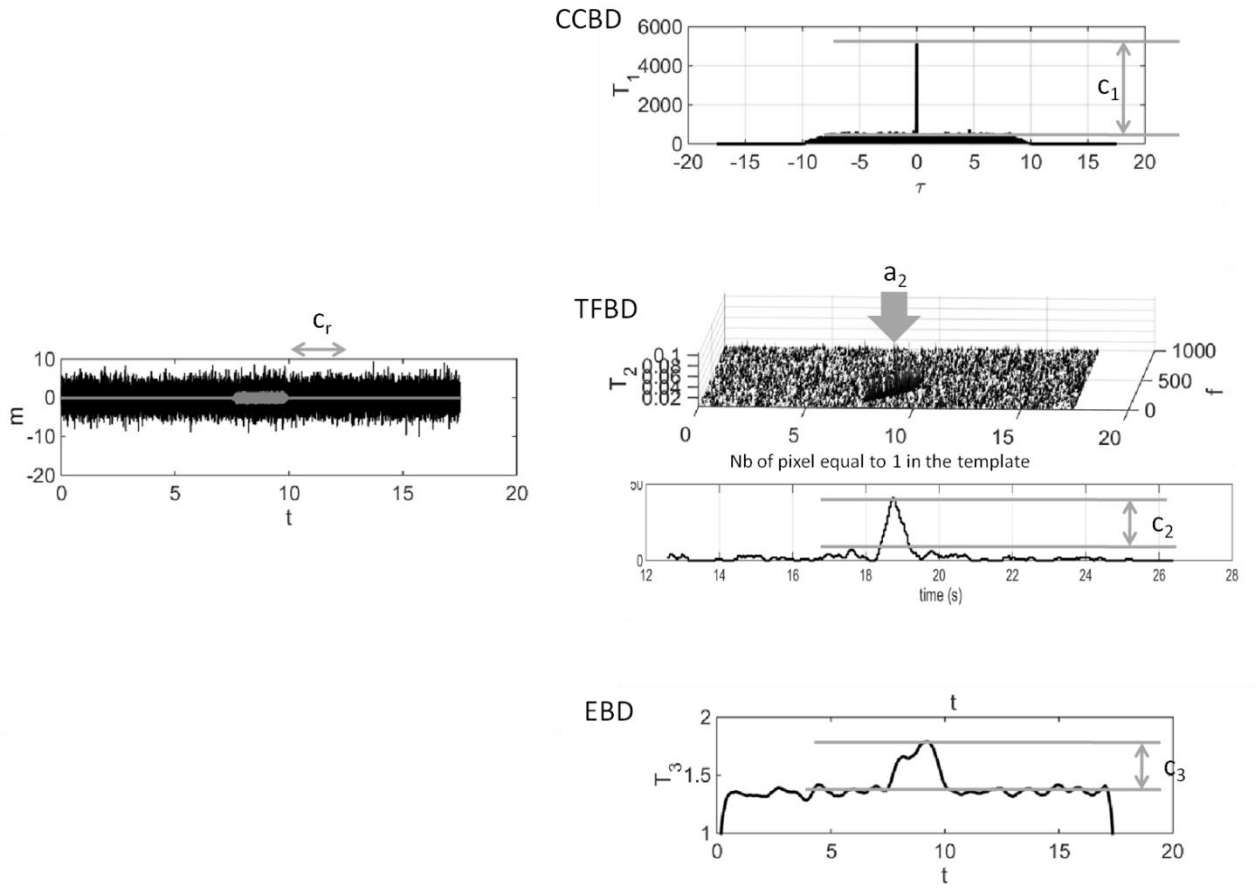

Figure 10. Step qualitative assessment of the detection performance of the three detectors. Left: negative SNR acoustic data, black noise and upcall; gray: upcall only. Top right: output of CCBD step 2. Middle right: spectrogram of the data and output of TFBD step 2. Bottom right: output of the EBD step 2. The arrows c_1, c_2, c_3 quantify the difference between the noise-only and upcall-present conditions.

Finally, the settings and the degrees of freedom for each detector are presented in Table 3.

Table 3. Settings of the CCBD, TFBD, and EBD detectors.

Setting	Usual values
CCBD	
Step 1 - Reference r(t)	$r(t) = \cos(2\pi f_{min} t + 2\pi \frac{f_{max-fmin}}{2T} t^2)$ T = 1 s, f_{min} = 100 Hz, f_{max} = 200 Hz
Step 2 - Frequency of evaluation of test T1	1 test every 0.1 s
Step 2 - Probability of false alarm	1 per day, 2×10^{-6} per test
TFBD	
Step 1.1: Length of the FFT windows	0.1 s
Step 1: Spectrogram weighting windows	Kaiser 180 dB
Step 1.1: Overlap	50%
Step 1.2: Ambient Noise estimation	Method in Kinda et al. (2013, p. 85-86), percentile = 20
Step 1.2: Probability of false alarm at the level of the time-frequency bin of the spectrogram	10^{-3}
Step 1.3: Time-frequency template of an upcall	Linear frequency modulation LFM T = 1 s, f_{min} = 100 Hz, f_{max} = 200 Hz Width of the template: ±2 spectrogram time steps around the upsweep time-frequency bin center
Step 2: Frequency of evaluation of test T2	1 test every 0.1 s
Step 2: Probability of false alarm when the template of the upcall is matched with the tested acoustic data	1 per day, 2×10^{-6} per test
EBD	
Step 1.1: bandwidth of the passband filter	[f_{min}, f_{max}] = [100 Hz, 200 Hz]
Step 1.2: Duration of energy integration	T = 1 s
Step 2: Frequency of evaluation of test T3	1 test every 0.1 s
Step 2: Probability of false alarm	1 per day, 2×10^{-6} per test

2.3.5 Receiver Operational Characteristics (P_d, P_{fa}, SNR) of the three detectors

To complete the qualitative approach of CCBD, TFBD, and EBD performances presented in section 2.3.3, we derive here the Receiver Operational Curves (ROC). The ROCs provide a rapid overview of the detection performance. For each detector, the ROC takes the form of an analytically or numerically derived function which links the detection probability, the probability of false alarm and the SNR:

$$ROC <=> P_d = function(P_{fa}, SNR).$$

If we chose to fix the probability of false alarm, the ROC expresses the relation between the detection probability and the SNR. For a given ambient noise (diffuse Wenz's noise and noise from individual ships), source levels and transmission losses, the ROCs are used to:

- quantify the detection probability as a function of whale range,
- quantify the expected detection range corresponding to a given detection probability.

On the first hand, as a detection of an upcall (right or false) will first elicit a validation by an human operator and second a decision about the regulation of the shipping or fishing activity, we want to avoid numerous false alarms and we target a low false alarm probability: 1 false alarm per day or 2×10^{-6} per test.

On the second hand, the PAM systems are used to protect the "endangered" species. So, we target a large probability of detection equal to 50% per test. Choosing a larger probability and at the same time expecting to have only one false alarm per day would cause a drastic increase of the required minimum SNR and a large drop in the range of detection. This level conforms with the state of the art of the PAM throughout the world as demonstrated during the bench-test of PAM DCLDE international workshop, in 2013 (https://soi.st-andrews.ac.uk/dclde2013/). As the NARWs regularly emit upcalls, the natural and intrinsic redundancy helps improving the hourly or daily probability of detection.

This functioning point is challenging if we expect it still holds at large ranges. Therefore, we propose here a precise and detailed bench-testing of several detectors and PAM systems (single-hydrophone vs hydrophone arrays).

The ROCs of CCBD and EBD under the data assumptions given in section 2.3.2 are classical and we used their expression as given by Kay (1998):

- CCBD see (Kay 1998, p. 101),
- EBD see (Kay 1998, p. 250).

The original derivation of ROC for TFBD is fully outlined in annex 2.

Table 4 presents the ROCs of CCBD, TFBD, and EBD. One advantage of ROCs of Table 4 is to give a clear and common expression of the SNR for the three types of detectors. These SNRs take into account two components:

- The SNR for raw data: $\frac{A^2}{2\gamma_0}$

- A processing gain (multiplication by T or T_{fft}).
- $A/\sqrt{2}$: rms Sound Pressure Level (SPL) of the upcall measured by the receiver
- γ_0: power spectral density of the total noise b_{tot} (diffuse noise and sum of individual ship noises, see section 2.3.2) at the mid frequency of the upcall (150 Hz)
- T: duration of the upcall
- T_{fft}: duration of the FFT windows used to compute the spectrogram
- cdfN: cumulative distribution function of a normal law with mean equal to 0 and variance equal to 1.
- P: number of time-frequency bins (i.e. pixels) of the spectrogram within the upcall time-frequency template
- $\binom{P}{j} = \frac{factorial(P)}{factorial(P-j)\times factorial(j)}$
- ncX2cdf(x,p,q): cumulative distribution function computed at x of a non-central chi^2 law with p degrees of freedom and a parameter of non-centrality equal to q
- $chi^2inv(x,p)$: reciprocal function computed at x of the cumulative distribution function of the centralized chi^2 law with p degrees of freedom.

Table 4. Formula of the ROCs of CCBD, TFBD, and EBD

Detector	**ROC**
CCBD	$P_d = 1 - cdfN(cdfN^{-1}(1-P_{fa}) - \sqrt{\frac{A^2T}{2\gamma_0}})$
TFBD	$P_{fa} = 1 - F_Q(i, H_0)$ $P_d = 1 - F_Q(i, H_1)$ $F_Q(i, H_0) = \sum_{j=0}^{i} \binom{P}{j} (P_{fa\ pix})^j (1-P_{fa\ pix})^{P-j}$ $F_Q(i, H_1) = \sum_{j=0}^{i} \binom{P}{j} (P_{d\ pix})^j (1-P_{d\ pix})^{P-j}$ $P_{d\ pix} = 1 - ncX2cdf(-2\log(P_{fa\ pix}), 2, 0.8\frac{A^2T_{fft}}{2\gamma_0})$
EBD	$P_d = 1 - ncX2cdf(chi2inv(1-P_{fa}, N), N, \frac{A^2T}{2\gamma_0})$

Figure 11 presents the ROCs of the three detectors for an expected P_{fa} equal to 2×10^{-6}. For each detector, the shift from bad detection ($P_d < 0.1$) to very good detection ($P_d > 0.9$) occurs within a 4-dB change in SNR. If spherical spreading loss ($20\log_{10}(r)$) applies, 4 dB corresponds to a change of range by a factor of 1.58. A 58% decrease of the whale range (for example 100 km to 42

km) changes an inefficient detector to an efficient one. To perform as well as TFBD, EBD needs a 12- dB higher SNR, which means that the EBD detection range is 4-times shorter than TFBD range. Similarly, TFBD requires a 3-dB higher SNR to perform as well as CCBD, which corresponds to a 1.4-time lower detection range (for example 100 km for CCBD compared to 60 km for TFBD).

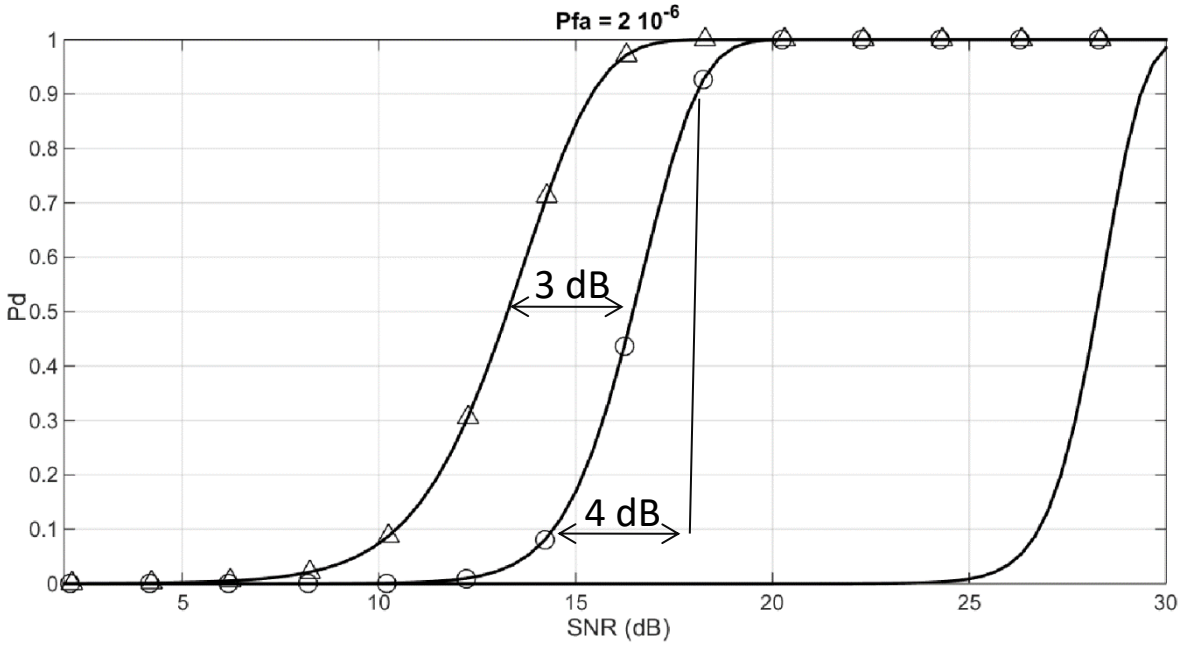

Figure 11. ROCs at $P_{fa} = 2\times10^{-6}$ for the three detectors.
Triangles: CCBD; circles: TFBD; continuous line without symbols: EBD.

2.3.6 Accounting for the use of a hydrophone array

Here, we incorporate the use of a hydrophone array into the computation of ROCs. We assume a planar 2D hydrophone array and a horizontal 2D configuration between the hydrophone array and the different sources of noise, since the range between the sources and the array is much larger than the depth of the acoustic propagation channel. Figure 12 sets the geometry and the conventions. A 2D set of coordinates is used, the x-axis is oriented towards the east and the y-axis towards the north. The azimuth is the angle between the east and the line of sight of the source. A 2D network of N_c hydrophones is placed at location (x_i, y_i) for hydrophone number i. The angle α_w is the azimuth of the whale and α_s is the azimuth of a ship.

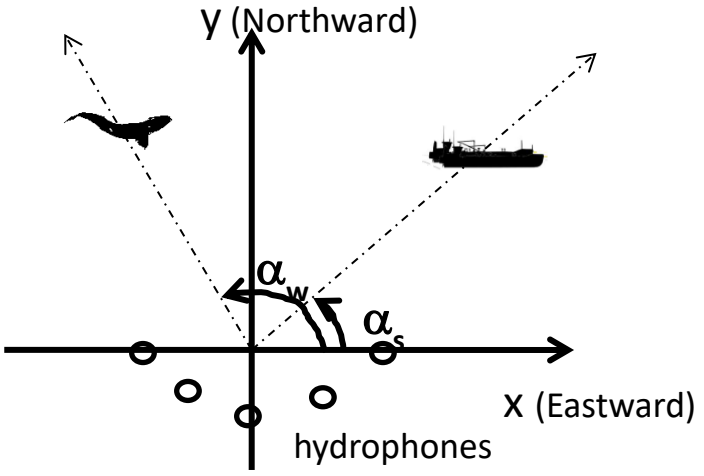

Figure 12. Geometry and conventions used for a hydrophone array.

The measurements of each hydrophone can be linearly combined, the weights applied to each measurement prior to summation allow for tuning the directivity of the hearing towards a preferred direction while drastically decreasing the hearing in the other directions. This succession of weighting and summation is named beamforming (Haykin 1985, For individual items see A85-43961 to A85-43963.). Let be α_2 the preferred azimuth and α_1 the azimuth of a generic source at frequency f, the processing gain of using beamforming is:

$$G(\alpha_2, \alpha_1, f) = \sum_{i=1:N_c} \frac{1}{\sqrt{N_c}} \exp(2\pi j \frac{c}{f}[-(c_i|u_2) + (c_i|u_1)]),$$
$$\text{with } c_i = \begin{pmatrix} x_i \\ yi \end{pmatrix}, u_1 = \begin{pmatrix} -\cos(\alpha_1) \\ -\sin(\alpha_1) \end{pmatrix}, u_2 = \begin{pmatrix} -\cos(\alpha_2) \\ -\sin(\alpha_2) \end{pmatrix}.$$

The processing gain G is such as:$|G(\alpha_1, \alpha_1, f|^2 = 1,\ and\ |G(\alpha_2, \alpha_1, f|^2 \leq 1$.

The wideband beamforming processing gain is computed by:

$$\overline{G^2}(\alpha_2, \alpha_1) = \frac{1}{N_f} \sum_{f \in [f_{min}, f_{max}]} G(\alpha_2, \alpha_1, f).$$

Figure 13 and Figure 14 present the processing gain of the beamforming for a 125-m long, 20-hydrophone linear array and that of a 20-hydrophone circular array with a 20-m radius. For comparison, the radius of the circular array is chosen such as its perimeter is equal to the length of the linear array. The processing gain presents a maximum toward the preferred azimuth, where a main lobe is formed with a width equal to the angular resolution flanked with side lobes. The difference between the gain at the preferred azimuth and other azimuths sets the abilities of the array to focus on the preferred azimuth (the direction of the whale) while filtering out the other azimuths (the direction of the ship). When the preferred azimuth is orthogonal to the linear array, the linear array shows a better angular resolution and weaker side lobes than the circular array. When the preferred azimuth tends to be parallel to the linear array, the resolution of the linear array drops and an ambiguity appears. This does not happen with a circular array, whose response does not depend on the preferred azimuth. With such hydrophone arrays the unwanted azimuth signal can drop by at least 15 dB.

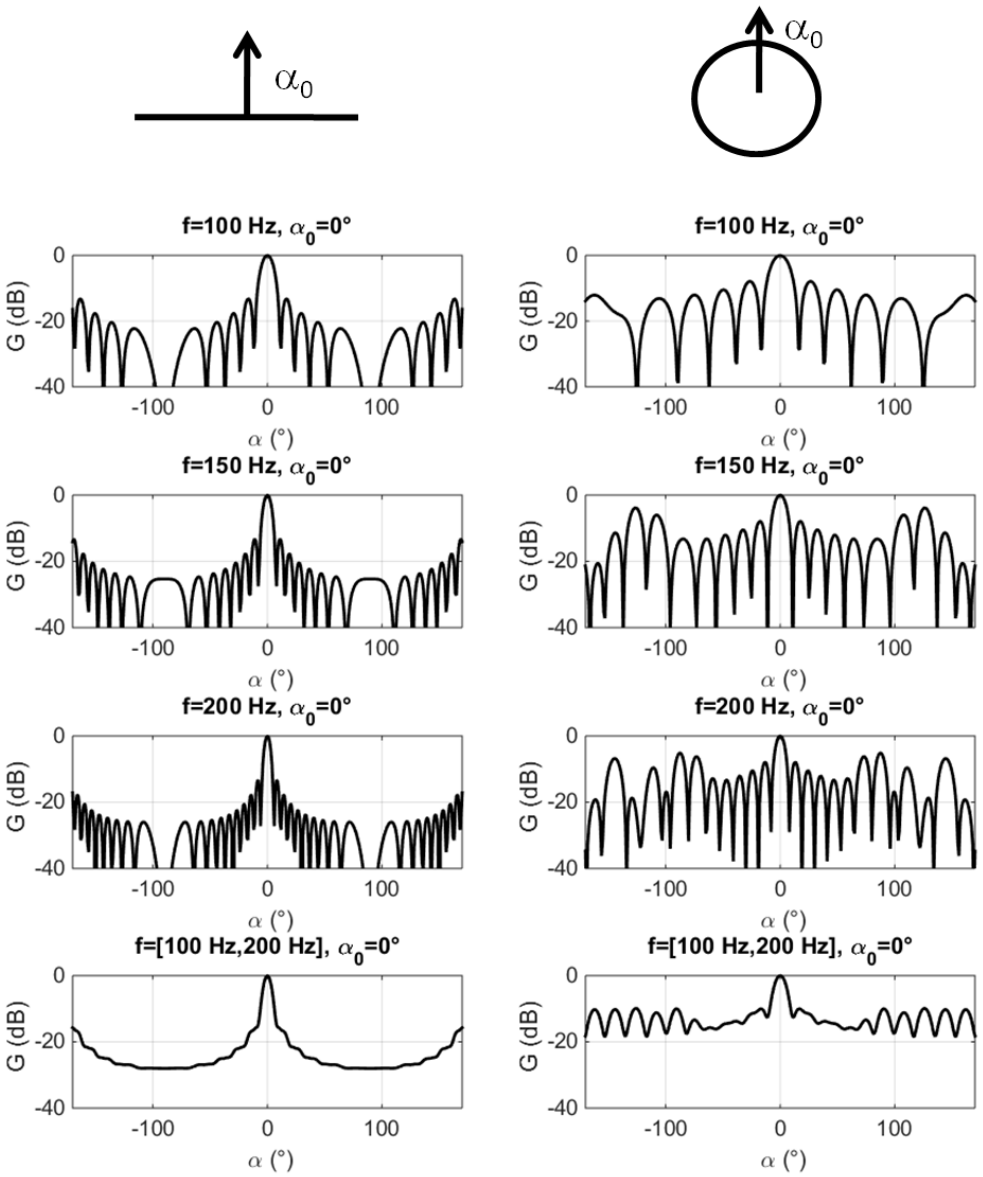

Figure 13. Beamforming processing gain for a preferred direction $\alpha_1 = 0°$.
Left: for a linear array (20 hydrophones, length = 125 m);
right: for a circular array (20 hydrophones, radius = 20 m).

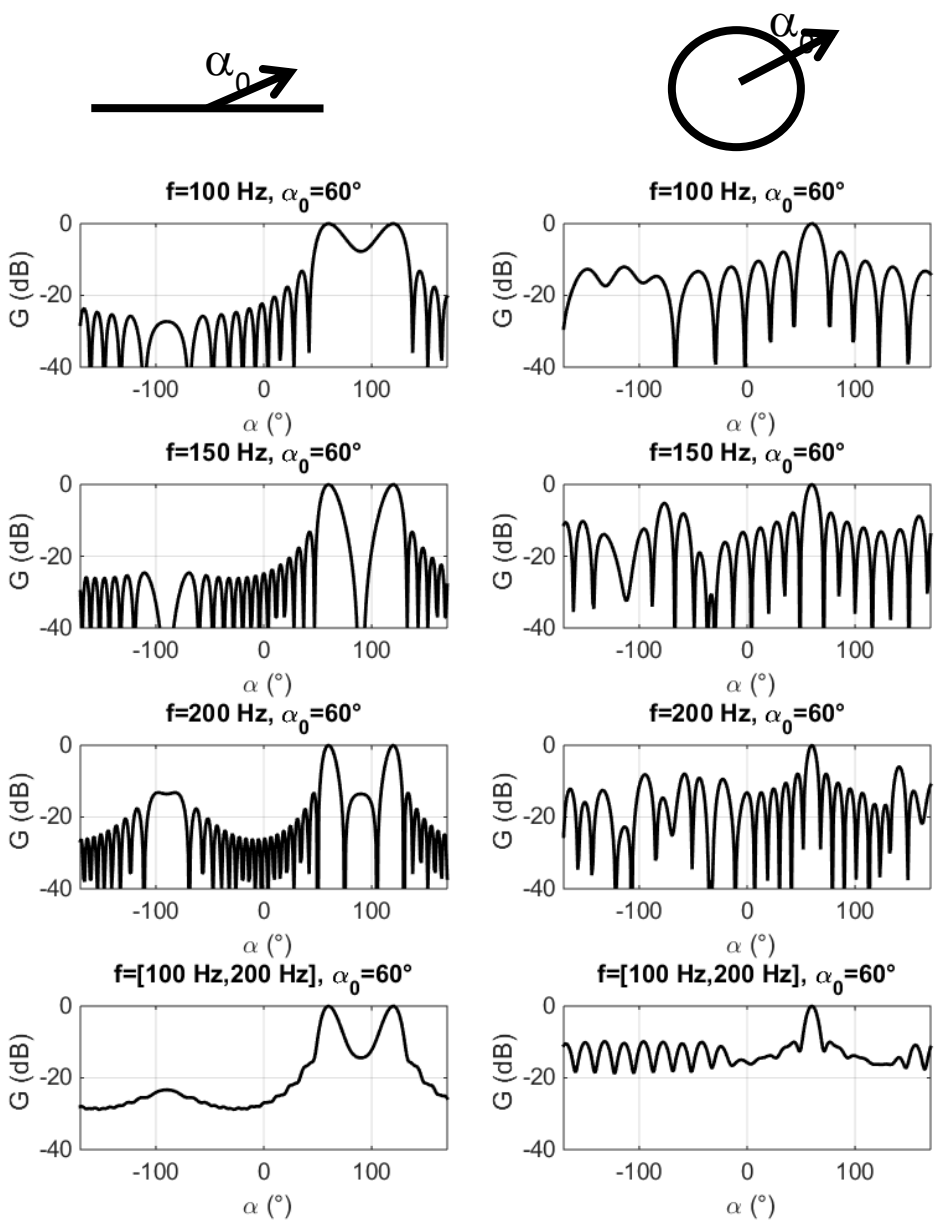

Figure 14. Beamforming processing gain for a preferred direction $\alpha_1 = 60°$.
Left: for a linear array (20 hydrophones, length = 125 m);
right: for a circular array (20 hydrophones, radius = 20 m).

Let us consider a complex realistic situation with a whale emitting an upcall and several ships radiating their own noises (see Figure 15), the processing gain of the array can be incorporated into the ROC formula by weighting each term of the SNR by its beamforming gain:

$$A_{dB} = SL_w - TL_W - \overline{G_{dB}}(\alpha_w, \alpha_w)$$

$$\gamma_{0_{dB}} = 10log_{10}(\sum_{i=1:N} 10^{\frac{\gamma_{SSL}(i) - TL_i - \mathrm{G_{dB}}(\alpha_{si}, \alpha_w)}{10}} + 10^{\frac{1}{Nc}\frac{\gamma_{wenz}}{10}}),$$

with $\mathrm{G_{dB}}(\theta_1, \theta_2) = 10\log_{10}\left(\overline{G_{dB}(\theta_1, \theta_2)^2}\right)$.

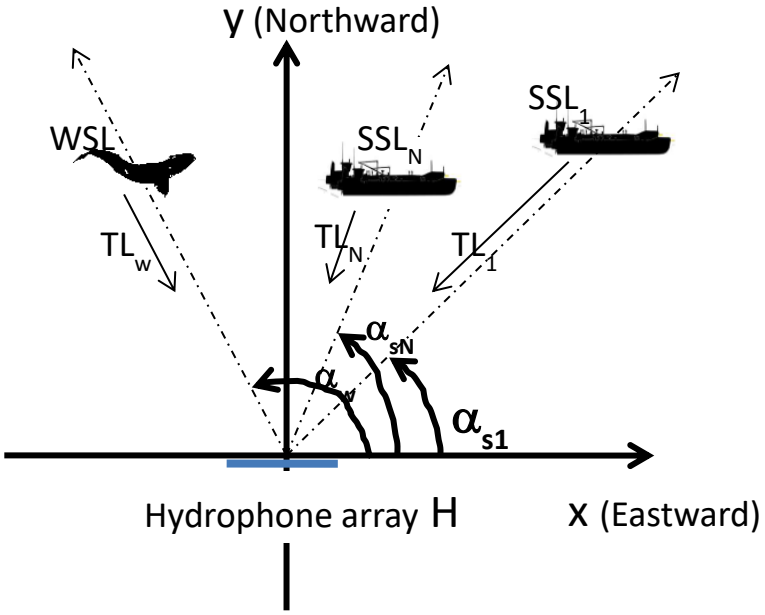

Figure 15. Scenario for a realistic case where a whale at azimuth α_w emits an upcall of source level WSL, the transmission loss to reach the array is TL_w, and several ships at azimuth α_{si} with source levels SSL_i and transmission losses TL_i.

2.3.7 NARW localization with hydrophone-array PAM systems at 2 coastal positions and its precision

We examine the localization of the NARW when hydrophone-array PAM systems positioned at 2 different locations are used simultaneously (Figure 16). As each PAM system must detect the upcall prior to localization, the largest area of possible localization is the intersection of the 2 detection areas. Within this intersection area, we assume that each array estimates the azimuth of the NARW upcall and the NARW position is the intersection of the NARW upcall DoA from each array.

The accuracy of the NARW localization depends on the accuracy of the azimuth estimation, which in turn depends on the SNR and the array shape (see section 2.3.6). To set the order of magnitude without complicating the study, we postulate that:

- the accuracy (standard deviation) is equal to the 3-dB half-angle of the main lobe of the directivity function of the hydrophone array, here equal to 3° (see Fig. 13, 14),
- the azimuth estimation from each hydrophone array is the sum of the true azimuth and a random error (Gaussian, zero mean, standard deviation 3°),
- the error at array 1 is independent from the one at array 2.

Under these hypotheses, we compute the NARW location error (standard deviation of x coordinate, standard deviation of y coordinate, ellipse of error), thanks to the formalism of the Cramèr-Rao Bounds (Kay 1993, p. 29). The derivation of the Cramèr-Rao Bounds of the position of the NARW is detailed in annex 4.

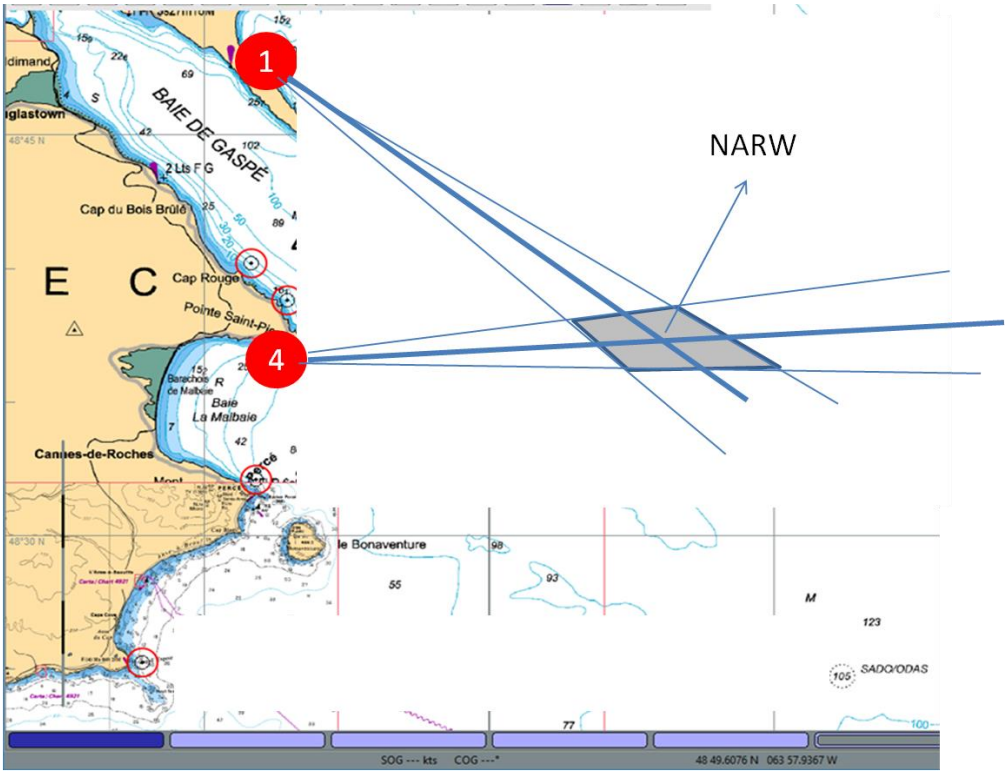

Figure 16. NARW localization with hydrophone-arrays at 2 coastal positions.
The directions of arrival (DoA, thick blue lines) given by the 2 coastal hydrophone arrays (red points) are crossed to find the position of the NARW upcall; the error of measurement of the azimuth (thin blue lines) creates an error of measurement of the NARW (blue quadrilateral).

3. RESULTS

3.1 Mapping of the area of 'good detection' within the NARW feeding ground area off Gaspé

3.1.1 Scenario

Here, we explain how we mapped the NARW upcall detection performances of a given PAM system based upon a realistic simulation of a quarter of year 2018 (August 1st to October 30th) by steps of 10 minutes, which provides 13 392 acoustical situations.

The TL maps (see section 2.2.4) from the upcall and from a ship to the PAM system were computed for frequencies 100 Hz, 150 Hz, and 200 Hz on a 1 km × 1 km grid over a 100-km radius and stored in memory. The wideband processing gain of the beamforming from the hydrophone array, $\overline{G^2}(\alpha_2, \alpha_1)$, was computed from -180° to 180° with a step of 1° for the two azimuths α_2 and α_1 and 101 frequencies between 100 Hz and 200 Hz. Results were stored in memory. For a single hydrophone, $\overline{G^2}(\alpha_2, \alpha_1)$ was set to 1.

A realistic shipping traffic within the feeding ground area was simulated for the following conditions:

- 2490 ship transits for the quarter were at positions transmitted by AIS,
- the narrow-band SL (γ_0) of each ship at 150 Hz was randomly chosen following the probability distribution measured by Simard et al. (2016).

For each sample of the 2018 time series, an instantaneous detection performance is taken by:

- looking for the number, the position, and the SL of ships present in the FGA at this time,
- computing the SNR at this time for each point of the TL grid, thanks to formulas at the end of section 2.3.6,
- computing the detection probability for CCBD, TFBD, and EBD and for each point of the TL grid, thanks to the ROC formula (Table 4), fed with $P_{fa} = 2\times10^{-6}$ and the SNR computed at the previous step,
- storing the map of the detection probability of CCBD, TFBD, and EBD.

Then the quarter time series of the detection maps is summarized in 2 maps:

- map of "wrong functioning", which expresses, for each point of the grid, the proportion of time where P_d is less than 0.1,
- map of "right functioning", which expresses, for each point of the grid, the proportion of time where P_d is more than 0.5.

We build a binary map of "Good Detection Area" (GDA) by looking for the positions where the proportion of time where $P_d > 0.5$ exceeds 50%.

From this map we estimate 3 descriptors (see Figure 17):

- ESD: the Effective Surface of Detection as the surface around the sensor where the proportion of time where P_d is more than 0.5 exceeds 50%,
- ERD: the Effective Range of Detection as the radius of a disc that would have the same surface than ESD noting that the coast reduces by half the effective surface so that:

$$ERD = \sqrt{\frac{2ESD}{\pi}}$$

- IRB: the inclusive rectangular box surrounding the GDA and its East-West and North-South dimensions.

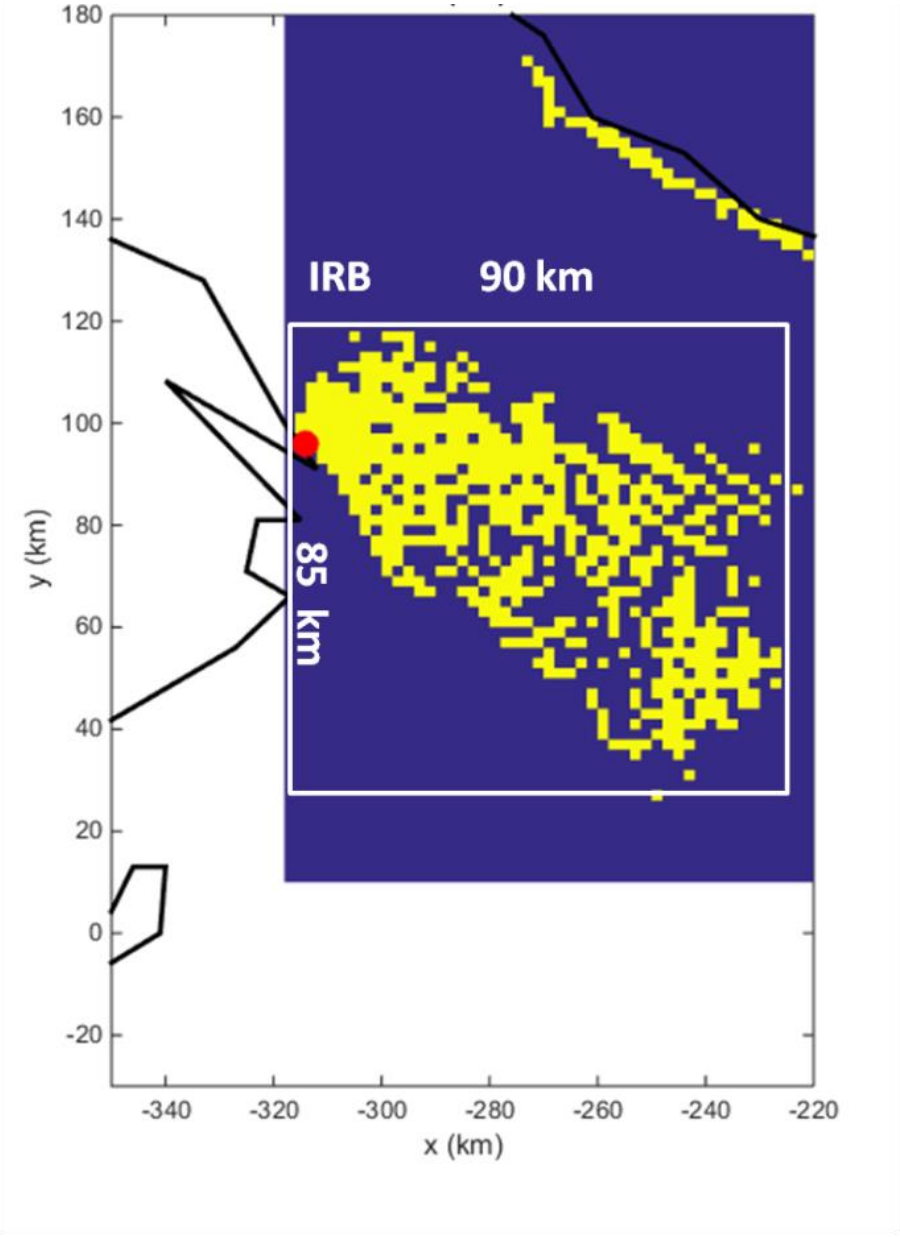

Figure 17. Good detection area (yellow) for a hydrophone array located at C1 with TFBD. The ESD is 2664 km^2, its ERD is 41 km, the IRB is the white rectangle, its North-South dimension is 85 km and the East-West dimension is 90 km.

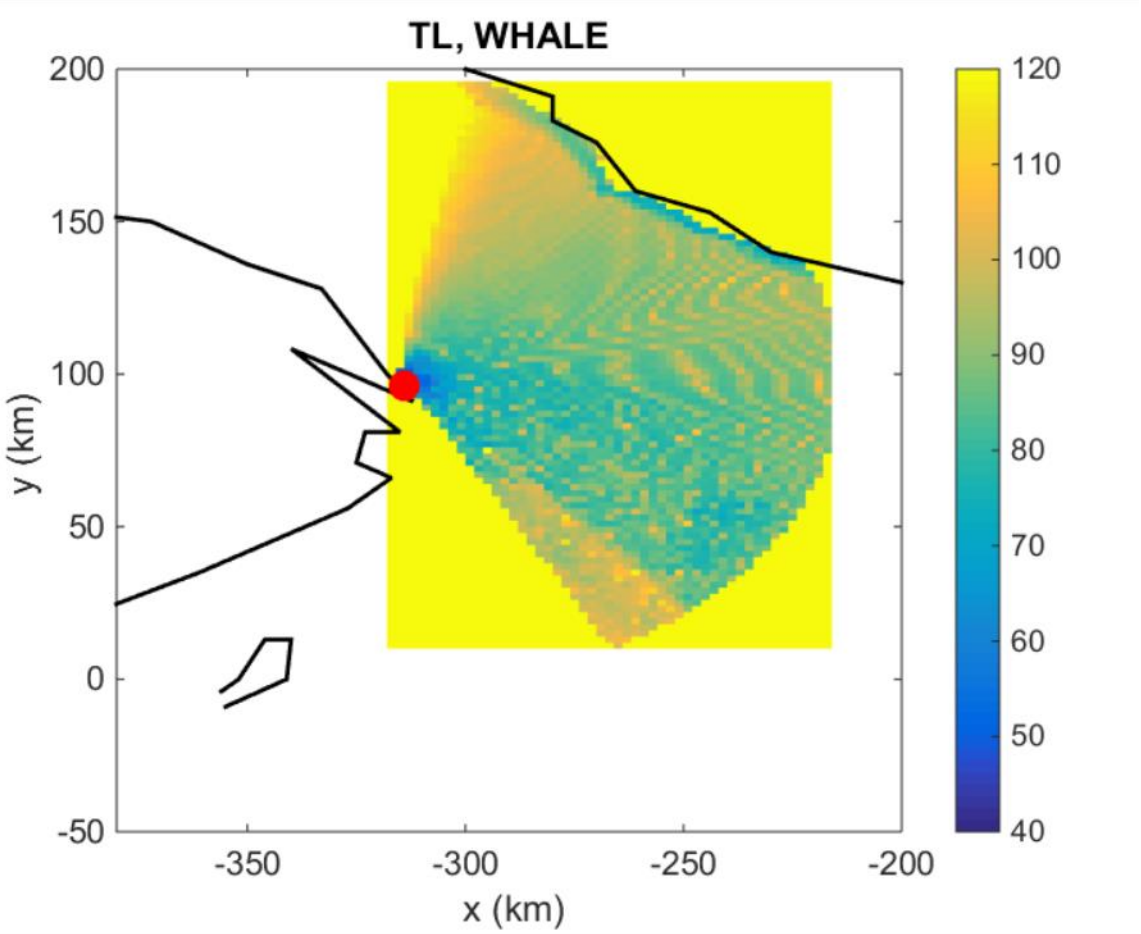

Figure 18. Map of TL (dB) @ 100 Hz between the whale at position (x, y, z = 15 m) and a PAM system receiving at the C_1 position.

3.1.2 Examples of instantaneous maps of detection probability

Figure 19 sets the detection performance with a PAM system at C_1 position under pristine ambient noise conditions, i.e. no ship in the FGA off Gaspé and only diffuse Wenz noise. The area of high detection probability when CCBD is used is larger than when TFBD is used. With a single-hydrophone PAM system, the area of high detection probability partially covers the southeastern part of the FGA (Figure 19A, B). Detection range is driven by the local bathymetry with a larger range for shallow water than for the deeper waters of the Laurentian channel. When the circular array of 20 hydrophones is used as PAM system (Figure 19C, D), the high processing gain of the beamforming compensates for the larger TL in deep waters, and the area of high detection probability then covers the whole FGA and even extends northwards into the Honguedo strait, both for CCBD and TFBD.

Figure 20 presents the case where ships transited within the area (August 5th, 23:00): 3 ships were cruising in the Honguedo seaway and 1 ship was cruising close to Gaspé. For a single-hydrophone PAM system at C1 position (Figure 20A, B), the surface of high detection probability is drastically reduced compared with the pristine case with no traffic (Figure 19A, B). The range

of high detection probability at this C1 shallow coastal position is less than 7 km. When a circular array is used the area of high detection probability enlarges (Figure 20C, D). The map of detection probability is bi-lobed as P_d is a function of the azimuth around the array. When the 6°-aperture beam towards the azimuth does not cross a ship, the range of detection is close to that of the pristine case with no ship. On the contrary, when a ship is inside the beam, masking may appear and the detection range may decrease; the magnitude of this decrease is a function of the SL of the ship (see Figure 21).

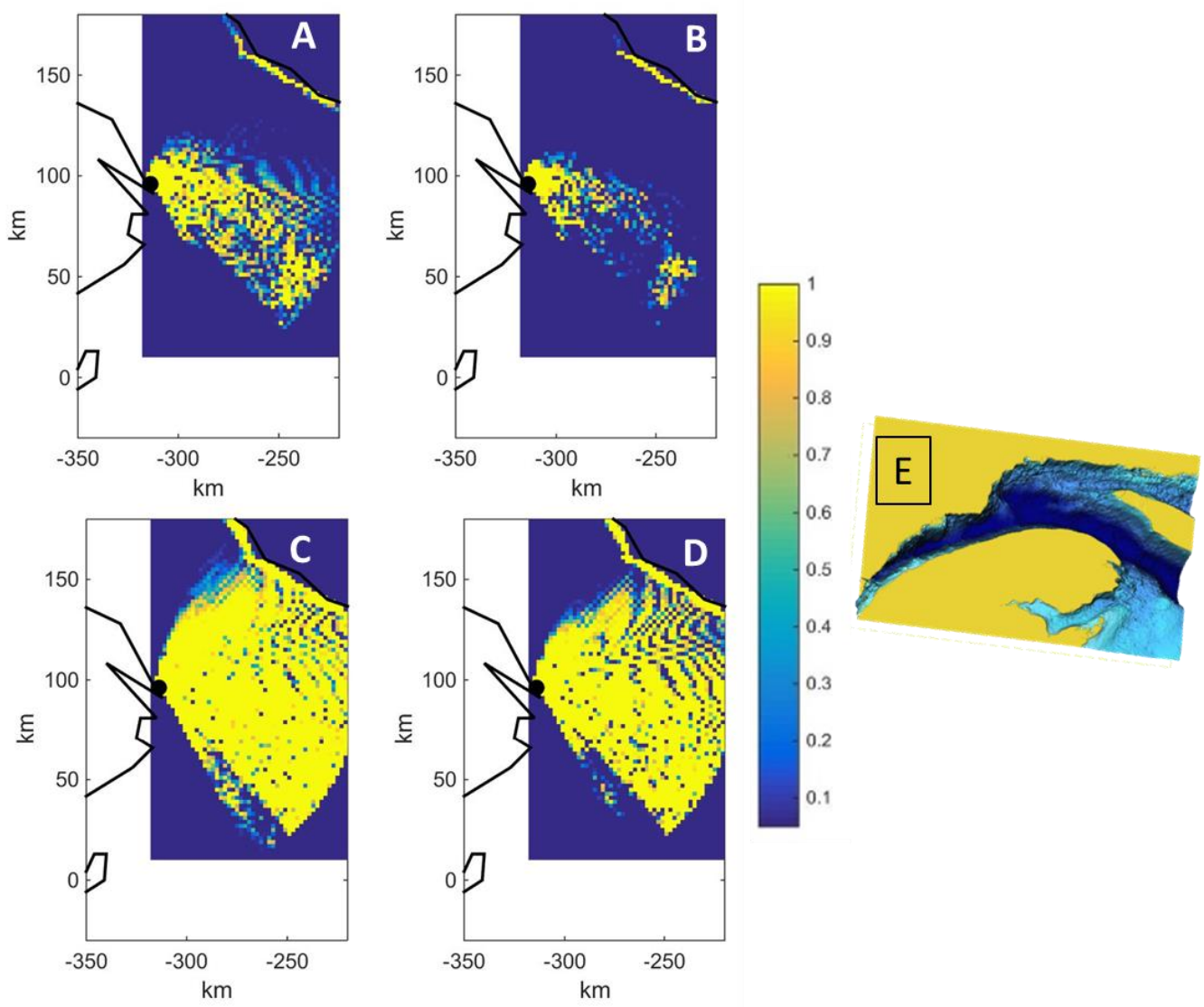

Figure 19. Instantaneous maps of detection probability for 1 single hydrophone receiving at C1 position, 20-m depth, with a CCBD (A) or a TFBD detector (B), and for a circular array with a CCBD (C) or a TFBD (D), the bathymetry is shown in E.

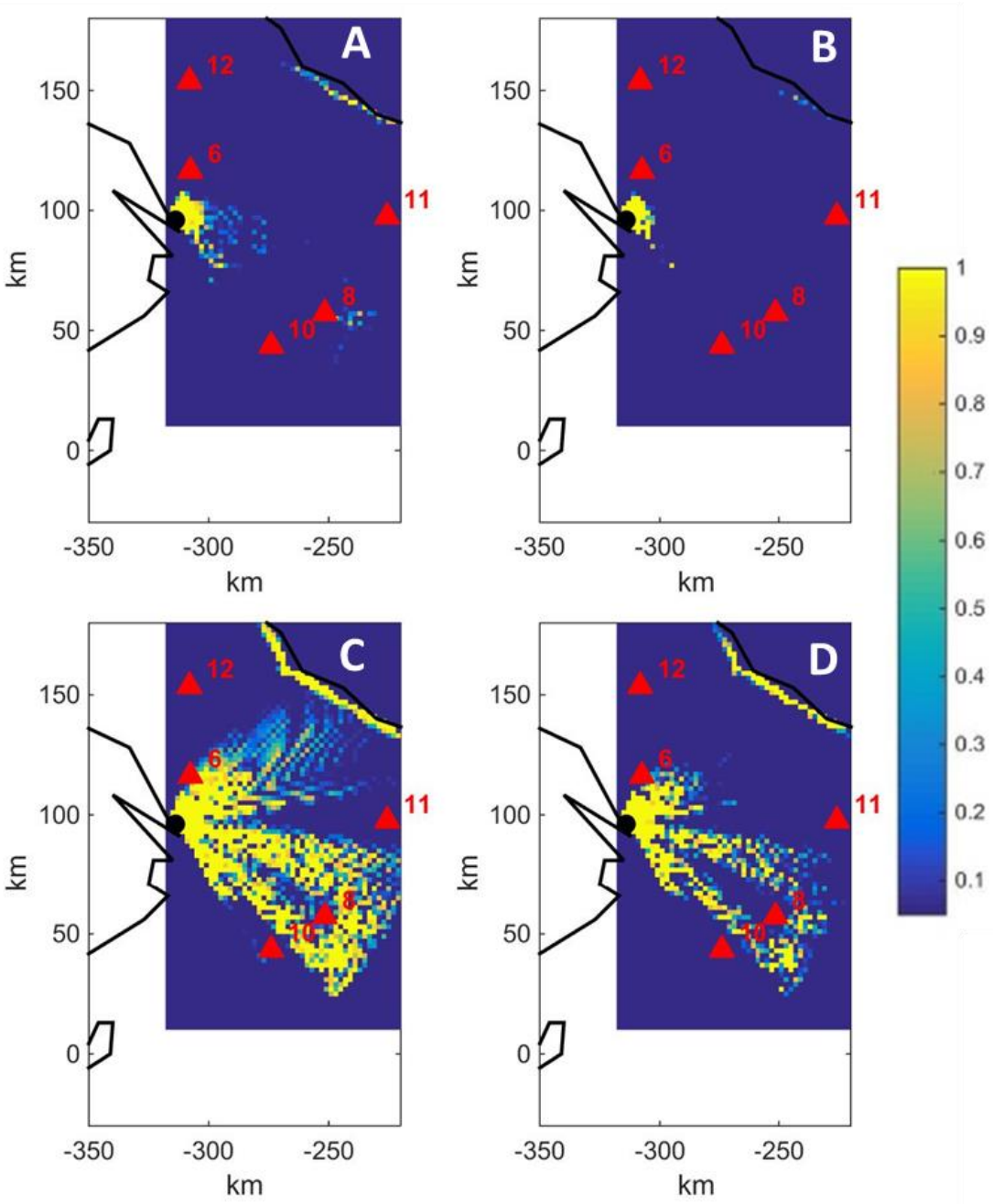

Figure 20.Instantaneous maps of the detection probability or PAM system (August 5th, 23:00) at C1 position, 1 single hydrophone receiving with a CCBD(A) or a TFBD detector (B), a circular array with CCBD (C) and TFBD (D) with 5 ships (red triangles)

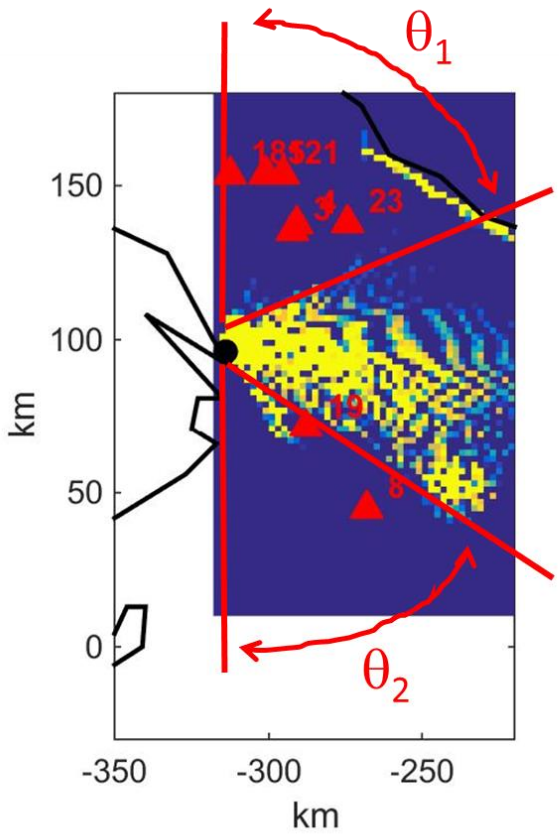

Figure 21. Residual masking areas (angular sectors θ_1 and θ_2) of NARW upcall by the shipping noise when a circular array is used at C1 and the ship and calling whale are co-located within the same beam.

3.1.3 Maps of the proportion of times where $P_d > 0.5$ for 1 single hydrophone and a circular array of 20 hydrophones with TFBD

Figure 22 maps the proportion of times from August 1st to October 30th 2018 where $P_d > 0.5$ for each of the 6 coastal positions for single-hydrophone and hydrophone-array PAM systems with TFBD. When single-hydrophone and TFBD PAM systems are used, the area with a high proportion of time of good detection is small and restricted to the immediate neighborhood of PAM system position (Figure 22 upper panels). The area with a high proportion of time where $P_d > 0.5$ increases when hydrophone arrays are used and can cover a large part of the FGA (Figure 22 lower panels). The surface and the shape of the area changes with the position of the PAM systems and is a function of both the coastal configuration, which provokes masking, and bathymetry that may favor or block the acoustic propagation.

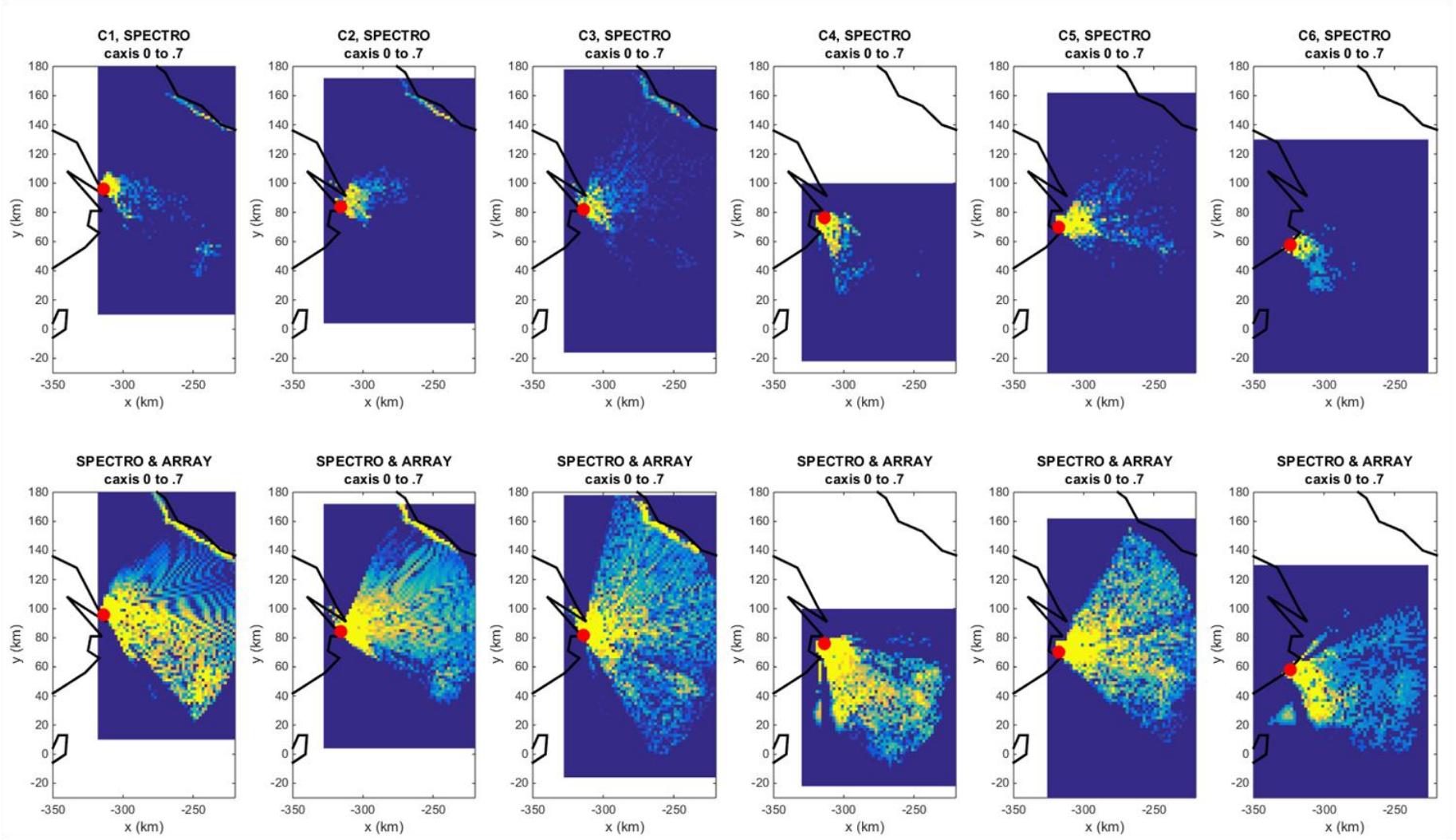

Figure 22. Maps of the proportion of time from August 1st to October 30th 2018 where the NARW upcall detection probability is more than 0.5 for PAM systems located at the 6 studied coastal positions (C1 (left) to C6 (right)).
Upper row: with single-hydrophone and TFBD PAM systems. Bottom row: with a circular array of 20 hydrophones and TFBD PAM systems. Color scale: blue (0), yellow (0.7).

3.1.4 Maps of Good Detection Area and their geometrical characteristics

Figure 23 maps of the GDAs for single-hydrophone or hydrophone-array PAM systems with TFBD at the 6 coastal positions, and Table 5 summarizes their respective geometrical characteristics (position × (single hydrophone or hydrophone array)).

The mean ± SD detection characteristics with single-hydrophone PAM systems at the 6 coastal positions are for: ESD, 256 ± 91 km^2; ERD, 12 ± 2 km; North-South IRB dimension, 25± 3 km; and East-West IRB dimension, 22 ±3 km. The maximum ESD is reached at position C5 (Table 5).

The same characteristics with hydrophone-array PAM systems are for: ESD, 1694 ± 600 km^2; for ERD, 32± 6 km; for North-South IRB dimension 53± 19 km; and East-West IRB dimension, 64 ± 18 km. The maximum ESD is reached at position C1 (Table 5).

The use of a hydrophone array increases the average GDA over all 6 PAM system positions by a factor of 6. For the C1 position alone, the GDA increases by a factor of 14. Figure 24 superimposes the GDA for a hydrophone-array PAM system with TFBD at coastal position C1 and the approximate envelope of the 2018-2019 NARW sightings (from https://whalemap.ocean.dal.ca/WhaleMap/) in the FGA off Gaspé.

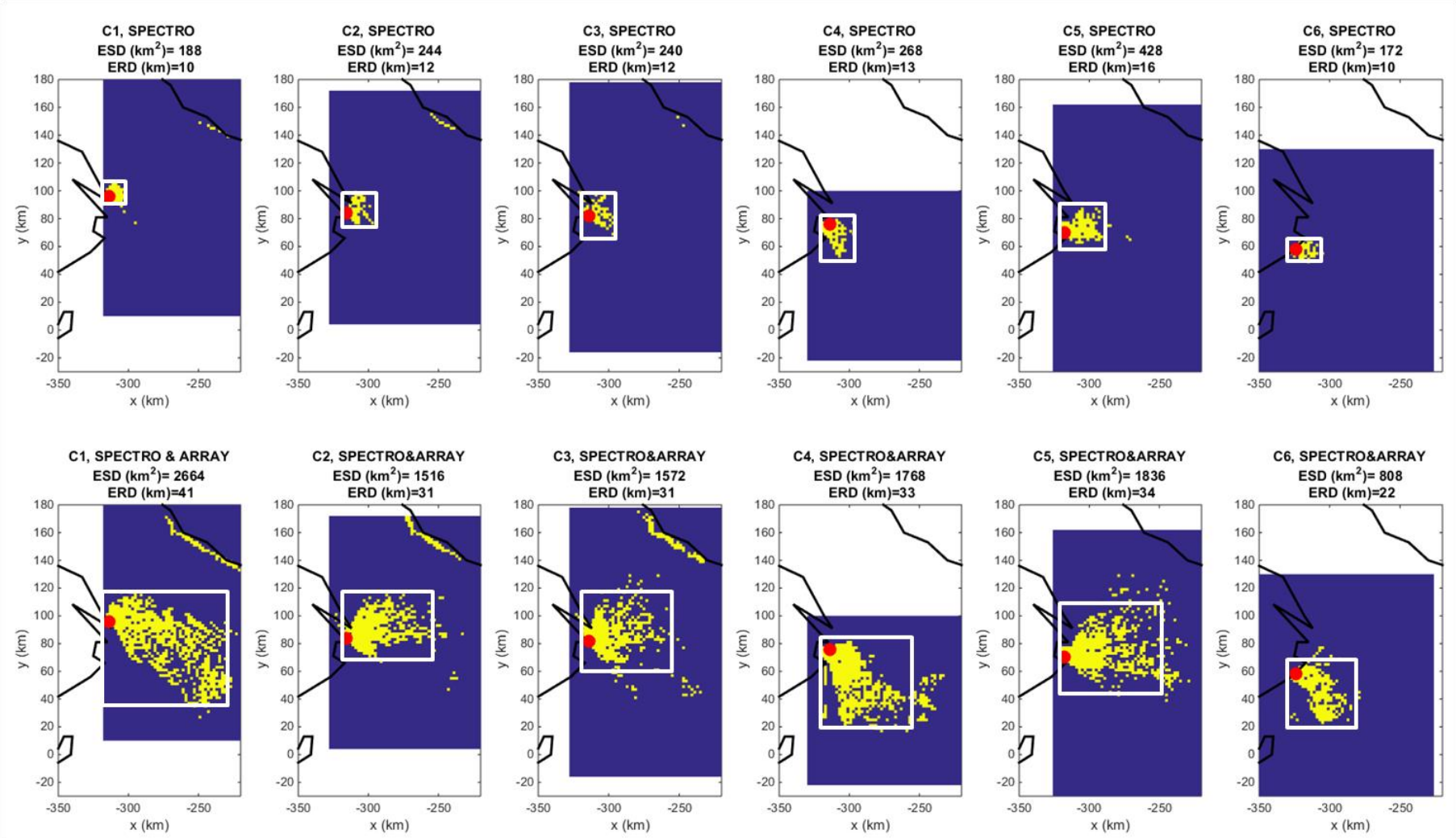

Figure 23. Maps of the good detection area (yellow) for single-hydrophone or hydrophone-array PAM systems with TFBD PAM systems located at the 6 studied coastal positions (C1 (left) to C6 (right)).
Top row: with single-hydrophone and TFBD PAM systems. Bottom row: with a circular array of 20 hydrophones and TFBD PAM systems. White box: IRB, ESD and ERD are reported up to their maps.

Table 5. Geometrical characteristics of good detection areas.

Site × Sensor × Detector	ESD (km^2)	ERD (km)	IRB North-South dimension (km)	IRB East-West dimension (km)
C1 × single-hyd.× TFBD	188	10	20	18
C2 × single-hyd.× TFBD	244	12	24	22
C3 × single-hyd.× TFBD	240	12	26	24
C4 × single-hyd.× TFBD	268	13	28	22
C5 × single-hyd.× TFBD	428	16	28	26
C6 × single-hyd× TFBD	172	10	28	24
C1 ×hyd. array× TFBD	2664	41	85	90
C2 ×hyd. array× TFBD	1516	31	50	62
C3 ×hyd. array× TFBD	1572	31	50	48
C4 ×hyd. array× TFBD	1768	33	58	72
C5 ×hyd. array× TFBD	1836	34	54	76
C6 ×hyd. array× TFBD	808	22	24	40

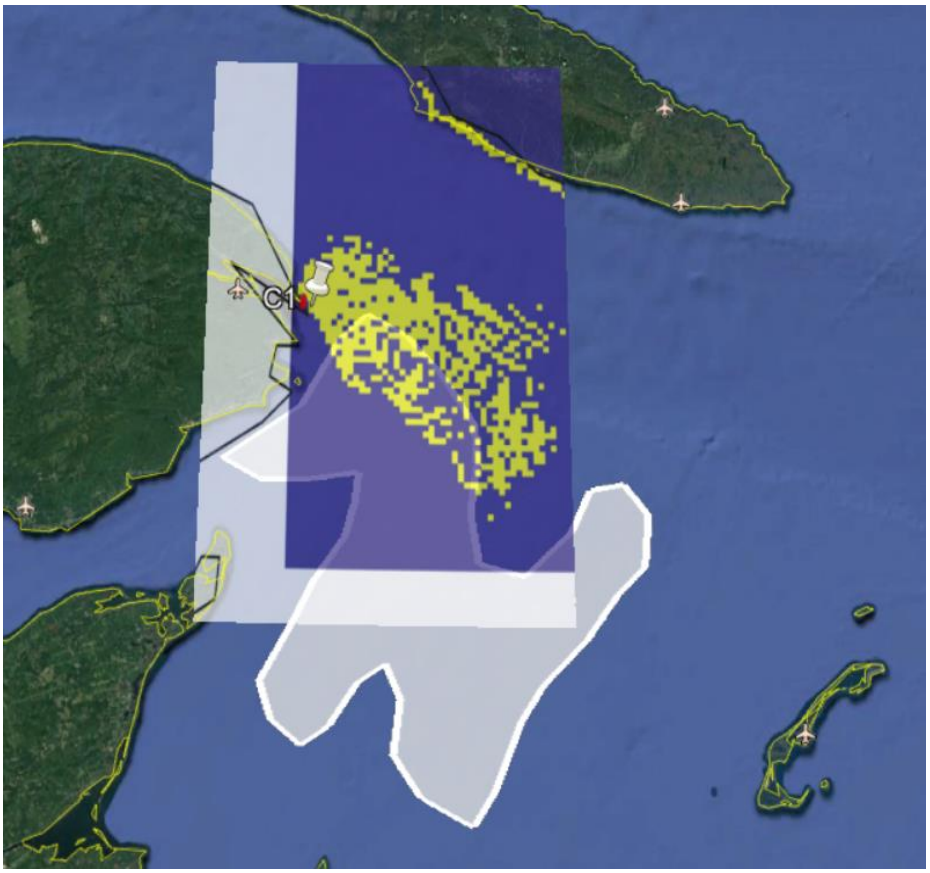

Figure 24. Superposition of the good detection area of an array of hydrophone at C1 with the approximate envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (from https://whalemap.ocean.dal.ca/WhaleMap/).

3.1.5 Optimal Good Detection Area and positioning of 2 coastal PAM systems for NARW upcall detection

We address the issue of increasing the GDA by using hydrophone-array PAM systems cabled to the coast from 2 positions. A NARW detection exists if at least one system detects the NARW upcall. Then the global GDA is the union of the individual GDAs from each PAM system. Table 6 reports the ESD and ERD when 2 PAM systems are used simultaneously to improve the GDA (union of individual GDAs, detection by at least 1 system). The mean ESD at the 6 coastal stations when a single hydrophone-array PAM system is used (1694± 600 km^2, Table 5) increases to 2956± 640 km^2 when two stations are combined (Table 6). Using 2 PAM systems may therefore multiply the mean ESD by a factor of 1.74. The best combination is one PAM system at C1 and one PAM system at C4 (ESD = 4 240 km^2, ERD = 52 km, Table 6). Figure 25 presents the shape of the corresponding individual and combined GDAs. The dimensions of its IRB are 100 km in North-South direction and 96 km in East-West direction. Figure 26 superimposes the GDA with PAM systems at C1 and C4.

Table 6. ESD and ERD for NARW upcall detection using hydrophone-array PAM systems with TFBD located at 2 coastal positions.

	PAM System 1 at position C1 [ESD km^2] [ERD km]	**PAM System 1 at position C2 [ESD km^2] [ERD km]**	**PAM System 1 at position C3 [ESD km^2] [ERD km]**	**PAM System 1 at position C4 [ESD km^2] [ERD km]**	**PAM System 1 at position C5 [ESD km^2] [ERD km]**
PAM System 2 at position C2	[3 220] [45]	[1516] [31]			
PAM System 2 at position C3	[3 392] [46]	[2 136] [36]	[1 572] [31]		
PAM System 2 at position C4	[4 240] [52]	[3 148] [44]	[3 072] [44]	[1 768] [33]	
PAM System 2 at position C5	[3 778] [48]	[2 820] [42]	[2 776] [42]	[3 036] [43]	[1 836] [34]
PAM System 2 at position C6	[3460] [47]	[2 300] [38]	[2 352] [38]	[2 096] [36]	[2 568] [40]

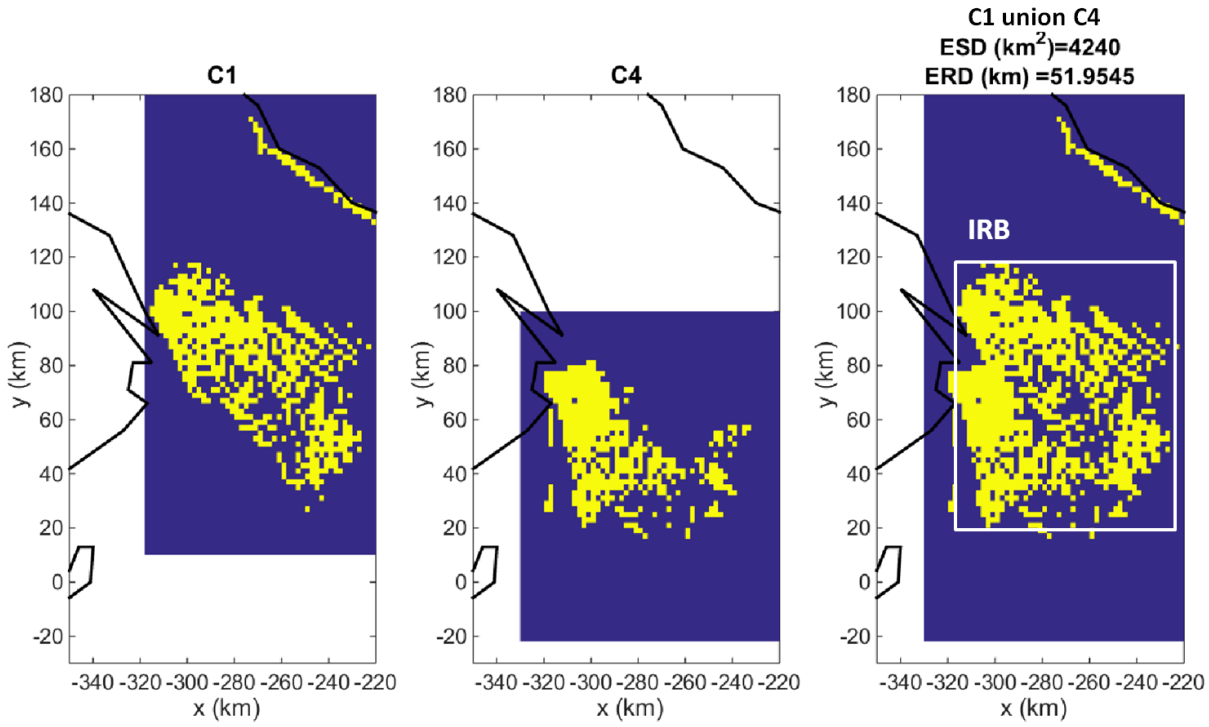

Figure 25. Good detection area for one PAM system at C1 (left) or at C4 (middle), and combination (right). North-South dimension of IRB is 100 km and East-West dimension is 96 km.

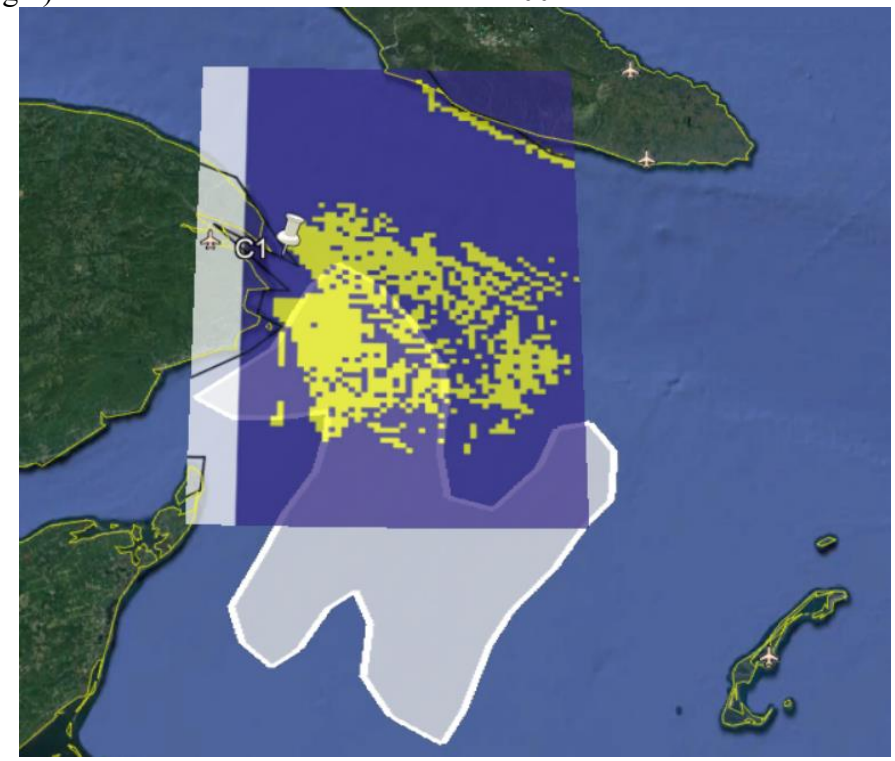

Figure 26. Superposition of the good detection area of hydrophone arrays at C1 and C4 with the envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (from https://whalemap.ocean.dal.ca/WhaleMap/)..

3.1.6 Optimal positioning of 2 PAM systems for NARW upcall localization

We examine the issue of NARW localization with 2 collaborating PAM systems equipped with a hydrophone array and TFBD. In this case, the upcall should be detected simultaneously by the two PAM systems; the NARW is localized by crossing the DoA from each PAM system (cf. section 2.3.7 and Figure 16).

Then the area for possible localization is the intersection of GDAs from each PAM system. Table 7 reports the ESD and ERD for the combination of hydrophone-array PAM systems at two coastal positions for NARW upcall localization. The mean ESD at the combination pairs of the 6 stations is 426 ± 307 km^2 while it is 1694± 600 km^2 when a single hydrophone-array PAM system is considered for the detection alone. Imposing the simultaneous detection at 2 PAM systems therefore divides the ESD by a factor of 3.7.

The largest area of possible localization of the NARW upcall is obtained from the combination of C1 and C2. Figure 27 presents this area. ESD is then 960 km^2 and the ERD is 24 km. Figure 28 superimposes this area of possible localization with hydrophone-array PAM systems at C1 and C2.

Within the area of possible localization, we have computed the accuracy of localization following section 2.3.7. Figure 29 presents the error ellipses for the NARW localization. The ellipses are oriented East-West; a standard deviation of the localization error smaller than 5 km in East-West direction and 1.5 km in North-South direction is achievable within a 30-km North-South × 35-km East-West rectangular box off the Gaspé coast.

Table 7. ESD and ERD using hydrophone-array PAM systems at 2 coastal positions for NARW upcall localization.

	PAM System 1 at position C1 [ESD km^2] [ERD km]	**PAM System 1 at position C2 [ESD km^2] [ERD km]**	**PAM System 1 at position C3 [ESD km^2] [ERD km]**	**PAM System 1 at position C4 [ESD km^2] [ERD km]**	**PAM System 1 at position C5 [ESD km^2] [ERD km]**
PAM System 2 at position C2	[960] [24]				
PAM System 2 at position C3	[844] [23]	[468] [17]			
PAM System 2 at position C4	[192] [11]	[592] [19]	[388] [15]		
PAM System 2 at position C5	[712] [22]	[444] [16]	[632] [20]	[568] [19]	
PAM System 2 at position C6	[8] [2]	[4] [1.5]	[28] [4]	[480] [18]	[76] [7]

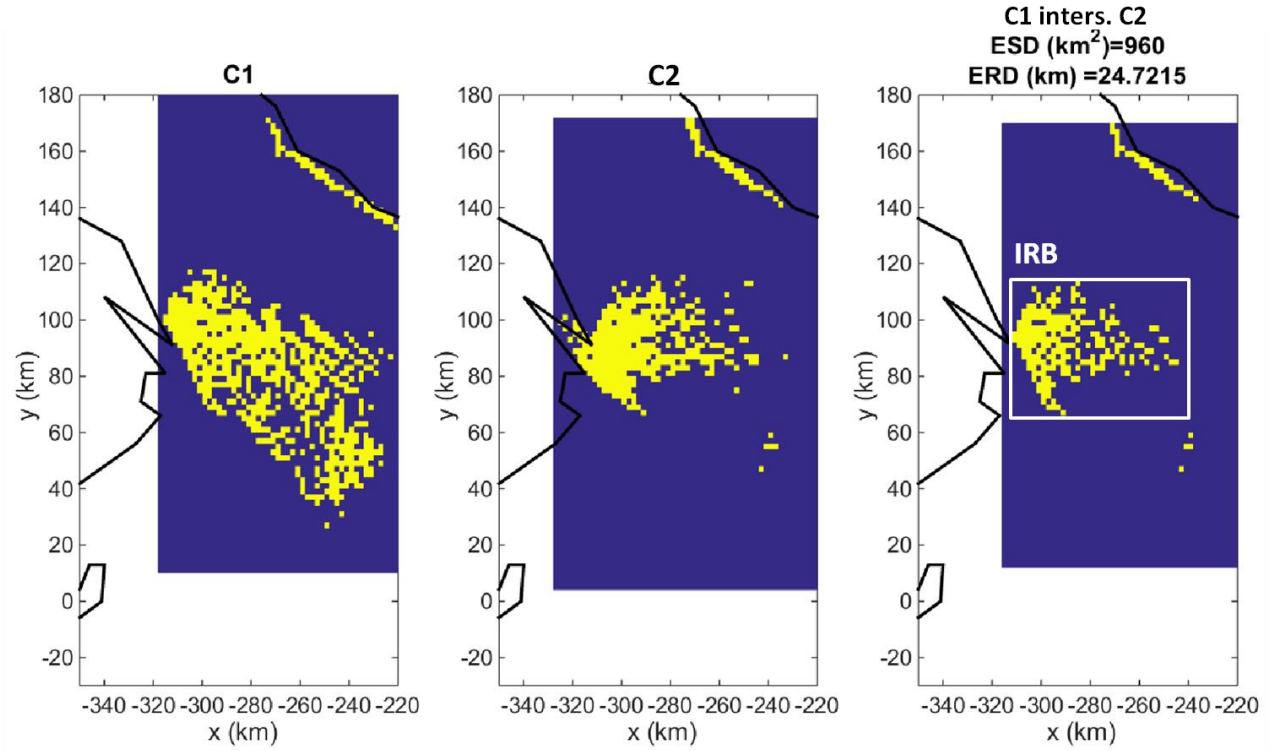

Figure 27. Good detection area for one PAM system at C1 (left), at C2 (middle) and for simultaneous detection at the 2 PAM systems (right). North-South dimension of IRB is 44 km and East-West dimension is 66 km.

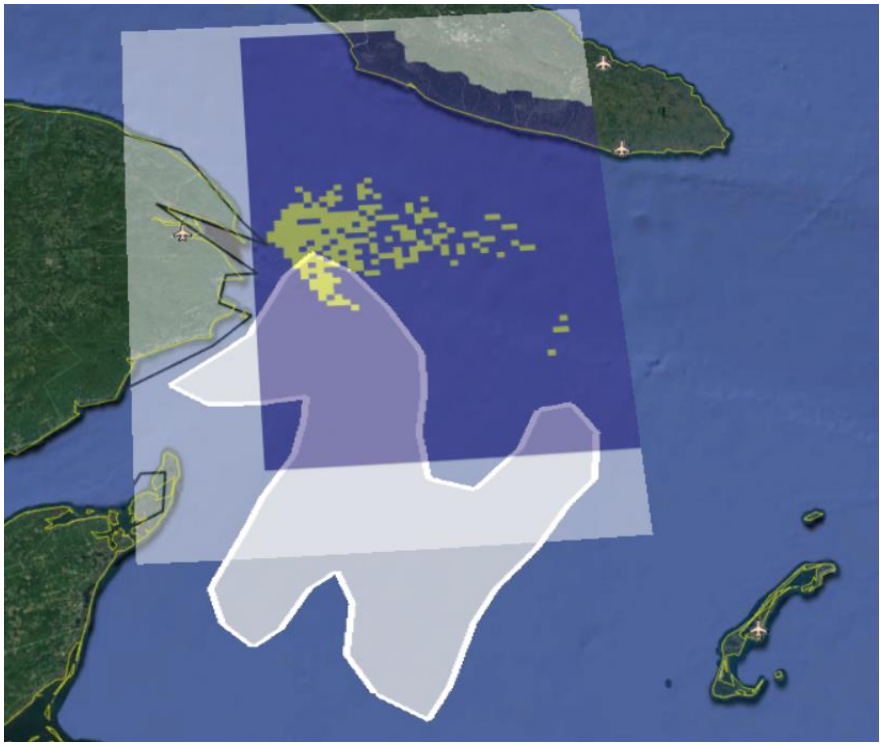

Figure 28. Superposition of the area for possible localization of hydrophone-array PAM systems at positions C1 and C2 with the envelope of 2018-2019 NARW sightings in the feeding ground area off Gaspé (while line).

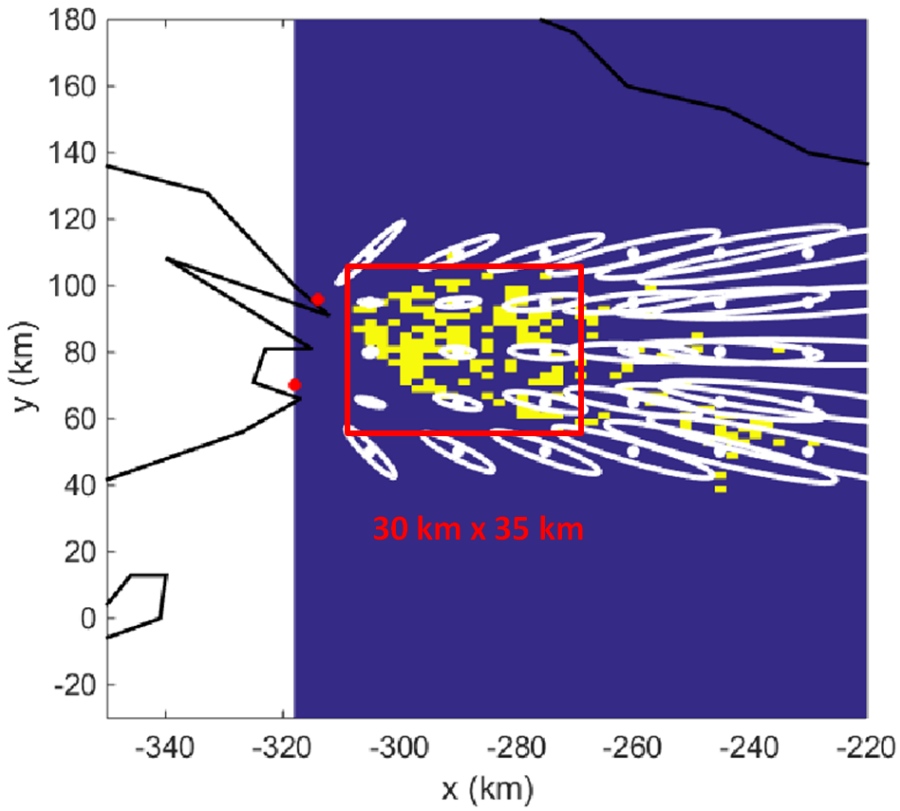

Figure 29. Error ellipse for NARW upcall localization with 2 hydrophone-array PAM systems located at positions C1 and C2.
Standard deviation of the localization error is less than 5 km East-West and 1.5 km North-South within the 30 x 35 km red rectangle.

4. DISCUSSION

In the present study, we addressed the issue of the detection of a NARW upcall within the FGA off Gaspé. With its size and mid-traffic conditions (144 km × 160 km, 27 ship transits per day, ~6 ships at the same time in the neighborhood), the FGA is noisy. The noise is created by discrete localized sources (i.e. the ships transiting in the area) and a hydrophone array can be used to filter out these punctual sources by beamforming to enhance the upcall SNR, its detection probability and range. We chose 6 positions for the installation of shore-cabled PAM systems on the Gaspesian coast, presenting logistical facilities (possibility to cable the PAM system onshore, sufficient depth at less than 1 km from the coast) for real-time processing and broadcasting of NARW detections and localizations. For detection purposes, the PAM system locations were considered individually and their detection performances with a single-hydrophone and a hydrophone-array were compared. For localization purposes, we considered two cooperating PAM systems with a hydrophone array where each PAM system estimates the azimuth of the NARW upcall, before estimating its position by crossing the DoAs from each PAM system.

4.1 Choosing the right detector and needs for a precise map of the detection performance

We considered 3 detector types for the upcall and we carefully derived the ROCs. The CCBD sets the best detection performance bounds but lacks robustness and requires large computational capacities. EBD is basic and sets the worst detection performance bounds. TFBD represents the state-of-the-art used in present operational PAM systems. A common expression for SNR is shared by the three ROCs, permitting a precise comparison between the three detectors.

If the probability of false alarm is set to one per 24 h (i.e. probability of false alarm equal to 2×10^{-6} per test, one test every 0.1 second), a SNR variation of 4.23 ± 1.06 dB can shift from bad detection performance (P_d equal to 0.1) to very good performance (P_d equal to 0.1). Such SNR variation can be observed when the distance between the hydrophone and the calling whale increases. For instance, assuming spherical TL in first approximation, if bad detection performances are observed at a range d = 100 km, the detection performance becomes very good at 61 km (= 0.61d). The range shift from bad to very good detection performances occurs over a distance of only 39 km. So drastic changes of detection performances are obtained across a ~120-km wide area which is the scale of the FGA off Gaspé. Precise mapping of the detection probability is thus necessary to properly assess acoustic system detection performance.

A detection probability of 50% and one false alarm per 24 h are chosen as realistic functioning point of the acoustic detection system. To obtain this functioning point, minimum SNRs are 13.27 dB for CCBD, 16.43 dB for TFBD and 28.15 dB for EBD. Therefore, using spherical TL in first approximation, the maximum detection ranges, r_{CCBD}, r_{TFBD}, $r_{EBD,}$ correspond to the minimum necessary SNR for CCBD, TFBD, and EBD, and we can express r_{TFBD} and r_{EBD} with respect to r_{CCBD} as:

$$r_{EBD} = 10^{-(28.15-13.27)/20} r_{CCBD} = 0.18 r_{CCBD}$$
$$r_{TFBD} = 10^{-(16.43-13.27)/20} r_{CCBD} = 0.72 r_{CCBD} .$$

Using the basic EBD would decrease by 82 % the detection range obtained with the optimal CCBD. Considering this poor detection performance, EBD is definitely rejected.

The level of performance is proportional to the level of the *a priori* knowledge used by the detectors. EBD requires only knowing the bandwidth of the upcall, TFBD requires knowing a time-frequency template of the upcall, whereas CCBD needs the precise full waveform template.

Using the operational, robust and real-time-compatible TFBD decreases the detection range by only 28 % compared to the optimal, but not robust and not real-time compatible, CCBD range. Indeed, the optimality of CCBD is guaranteed by a perfect knowledge of the upcall at the receiver. However, in practice, it is quite impossible to possess this level of *a priori* knowledge, because the source itself is subject to some variability in its structuring features (e.g. minimum frequency, maximum frequency, duration,…). This intrinsic variability is magnified by the propagation within shallow waters, which introduces several distortions of the emitted signal. The propagated signal may include several echoes. Each echo corresponds to one propagation mode and its time-frequency support is modified by modal dispersion (Gervaise et al. 2008, Ioana et al. 2010). In annex 3, we studied the loss in processing gain of using the CCBD with a wrong reference. We showed that a small deviation of 10% in signal duration, minimum or maximum frequencies

degraded the SNR by 3 dB, (which is very close to the SNR variation necessary to shift from very good to bad detection performance). In addition, CCBD is more computational demanding that TFBD, which may create difficulties for integrated real-time applications. For these reasons, we conclude that TFBD is a very good trade-off between optimality and operability. If 100 km is the order of the range of good functioning of CCBD, then the range of good functioning of TFBD is reduced by 20 km, which is within the scale of the FGA off Gaspé. This again confirms the need for precise mapping of the detection probability.

4.2 Mapping of the detection performance through the realistic simulation of traffic over 3 months

We divided the period from August 1st 2018 – October 30th 2018 into 13 392 steps of 10 min. For each step, we mapped the detection probability of CCBD, EBD with or without beamforming, thanks to an acoustics simulation fed with the realistic TL, ship and upcall SLs, and the true AIS traffic. Then we summed up the 3-month time series by mapping the area of right functioning (location within the FGA where the proportion of time $P_d > 0.5$ if more than 50%) and infer the Effective Range and Surface of Detection. Using the traffic in the area during this quarter is conservative since larger than the average annual traffic in the area (Figure 1, Simard et al. 2014).

4.3 Effective Range of Detection of a single hydrophone and TFBD

The ERD for a single-hydrophone and TFBD PAM system is 12 ± 2 km, maximal (16 km) at position C5 and minimal at positions C1 and C6 (10 km) (see Figure 2 for positions and Table 5 for ERD). Whatever the PAM position is, the ERD with single-hydrophone PAM systems fails to reach the eastern part of the FGA (see Figure 3). Single-hydrophone PAM systems on the Gaspesian coast perform too poorly to detect the NARW within their FGA off Gaspé.

4.4 Effective Range of Detection with hydrophone array PAM systems with TFBD at 1 or 2 coastal positions

As the noise that affects NARW upcall detection is created by punctual ships at their respective positions, NARW upcalls and ship noise do not usually come from the same direction. A hydrophone array with beamforming allows for listening in one direction and not in the other directions. More precisely, a circular hydrophone array (20-m radius, 20 hydrophones) is able to focalize in one direction with a resolution of 7° and filter out the other directions by 15 dB within the bandwidth of the NARW upcall (see Figure 13, Figure 14).

Optimal detection performances were achieved for a PAM system at position C1. In the feeding ground area context (acoustical propagation, shipping), this additional 15-dB SNR offered by beamforming increases the ESD by a factor of 14 and the ERD by a factor of 4 for a PAM system at position C1. When superimposing the GDA of a PAM system at C1 and the envelope of NARW sightings in their FGA(see Figure 24), the GDA did not entirely cover the FGA but monitored its northern part and allows for detecting NARW displacements from their FGA in southern GSL toward the high-traffic St. Lawrence Seaway of Honguedo strait.

We addressed the issue of maximizing the ESD by using hydrophone-array PAM systems at 2 coastal positions (upcall detected at least on one PAM system). The optimal combination of positions was C1 and C4, where C1 system was covering the deep-water northern part of the FGA and C4 the shallow-water southern part of the FGA (see Table 6 and Figure 26). Then the IRB dimensions were 100 km x 96 km. Such IRB allowed for surveying the northeastern quarter of the FGA.

4.5 Area for good localization of NARW with 2 cooperating PAM systems using a hydrophone array and TFBD

Good detection was achieved by using a hydrophone array filtering out the noise from individuals ships. The beamforming at each array also produced an estimate of the azimuth of the upcall. Then two cooperating PAM systems allowed for localizing the NARW by crossing the directions of arrival given by each PAM system. A necessary condition for allowing localization was upcall detection by the two PAM systems simultaneously. This condition decreased the size of the area of possible localization compared with the GDA at one PAM systems alone. Imposing simultaneous detection at 2 PAM systems divides the ESD by a factor of 3.7. The combination maximizing the area of possible localization was one array at C1 and the other one at C2 (see Figure 27, Table 7). The IRB dimensions were then 44 km by 66 km. Within this IRB of possible localization, a sub-region of 30 km by 35 km allowed for a more precise localization (5 km East-West standard deviation, 1.5 km North-South standard deviation). This area of precise localization (1050 km^2) was quite limited (6.5 %) compared to the FGA size (16 000 km^2), but allowed for surveying possible NARW transits from their feeding ground areas towards the busy Honguedo seaway.

4.6 Operational conclusions: toward an *ad hoc* PAM system to survey NARW in their feeding ground area off Gaspé

The ESD of a single hydrophone PAM system was 256 ± 91 km^2 and ERD was 12 ± 2 km. Although the computation was made for 6 nearshore positions off the Gaspesian coast, we believe these values are representative of the whole FGA. Therefore, these ESD and ERD are comparable to those of single-hydrophone PAM systems mounted on autonomous buoys or gliders. As the surface of the feeding ground area off Gaspé is 16 000 km^2:

- 62 autonomous PAM buoys/gliders would be required to continuously monitor the whole FGA.
- A PAM glider can survey 1.5% of the FGA at any given time and only 3.2% of the FGA per day (15 km per day and a range of ±16 km around the position of the glider).

The use of hydrophone arrays increases the ESD and ERD and decreases the number of PAM systems required to monitor the FGA.

Through the 6 nearshore positions off Gaspesian coast, we demonstrated that position C1 was optimal for detection, with IRB dimensions of 85 km × 90 km.

The combination of 2 PAM systems at C1 and C4 led to an ESD of 4 240 km^2, equal to a quarter of the FGA.

Under the assumption that the size and shape of the good detection area at C1 was also valid for other positions in southern Gulf of St. Lawrence, we suggest that a network of nearshore cabled hydrophone-array PAM systems at 5 positions around the FGA and one autonomous buoy in the center of the FGA would be necessary to monitor the entire FGA:

- 2 PAM systems (nearshore, cabled, hydrophone array, TFBD) at positions C1 and C4,
- 1 PAM system (nearshore, cabled, hydrophone array, TFBD) on Lamèque Island near Shippagan,
- 1 PAM system (nearshore, cabled, hydrophone array, TFBD) north of Prince Edouard Island near Tignish,
- 1 PAM system (nearshore, cabled, hydrophone array, TFBD) on the southwest coast of Magdalen Island,
- 1 autonomous PAM buoy system (offshore, autonomous, single hydrophone, TFBD) in the center of the feeding ground area.

If localization and tracking of NARW at a hotspot within the FGA is desired, additional arrays may be used and located 20 to 30 km away from the above coastal positions. Then the combination of these two cooperating systems would allow NARW localization within an area of 30 km by 30 km.

ACKNOWLEDGMENTS

This work was supported by Fisheries and Oceans Canada. We are grateful to Bazile Kinda and Jérôme Mars for reviewing the manuscript and to Denis Chabot for his editorial work.

REFERENCES

Abramowitz, M., and Stegun, I. 1974. Handbook of mathematical functions with formulas, graphs and mathematical tables. Dover Publications, Inc., New York, NY, USA.

Aulanier, F., Simard, Y., Roy, N., Bandet, M., and Gervaise, C. 2016a. Groundtruthed probabilistic shipping noise modeling and mapping: Application to blue whale habitat in the Gulf of St. Lawrence, Acoust. Soc. Am. Proc. Meetings Acoust., 27, 070006, pp. 1-14.

Aulanier, F., Simard, Y., Roy, N., Gervaise, C., and Bandet, M. 2016b. Spatio-temporal exposure of blue whale habitats to shipping noise in St. Lawrence system. DFO Can. Sc. Advis. Sec. Res. Doc. 2016/090, p. vi + 26 p.

Aulanier, F., Simard, Y., Roy, N., Gervaise, C., and Bandet, M. 2017. Effects of shipping on marine acoustic habitats in Canadian Arctic estimated via probabilistic modeling and mapping. Mar. Poll. Bull. **125**(1): 115-131.

Baumgartner, M.F., Fratantoni, D.M., Hurst, T.P., Brown, M.W., Cole, T.V.N., Van Parijs, S.M., and Johnson, M. 2013. Real-time reporting of baleen whale passive acoustic detections from ocean gliders. J. Acoust. Soc. Am. **134**(3): 1814-1823.

Baumgartner, M.F., and Mussoline, S.E. 2011. A generalized baleen whale call detection and classification system. J. Acoust. Soc. Am. **129**(5): 2889-2902.

Clark, C., Ellison, W., Hatch, L., Merrick, R., Van Parijs, S., and Wiley, D. 2011. An ocean observing system for large-scale monitoring and mapping of noise throughout the Stellwagen Bank National Marine Sanctuary. Reports to the National Oceanographic Partnership Program, Stellwagen project, Award number N00014-07-1-1029, p. 11 pp.

Clark, C.W. 1982. The acoustic repertoire of the Southern right whale, a quantitative analysis. Animal Behaviour **30**(4): 1060-1071.

Clark, C.W., Brown, M.W., and Corkeron, P. 2010. Visual and acoustic surveys for North Atlantic right whales, *Eubalaena glacialis*, in Cape Cod Bay, Massachusetts, 2001-2005: Management implications. Mar. Mam. Sci. **26**(4): 837-854.

Cole, T.V.N., Hamilton, P., Henry, A.G., Duley, P., Pace III, R.M., White, B.M., and Frasier, T. 2013. Evidence of a North Atlantic right whale *Eubalaena glacialis* mating ground. End. Sp. Res. **21**: 55-64.

Collins, M.D. 1993. A split-step Padé solution for the parabolic equation method. J. Acoust. Soc. Am. **94**(4): 1736-1742.

Dadouchi, F., Gervaise, C., Ioana, C., Huillery, J., and Mars, J.I. 2013. Automated segmentation of linear time-frequency representations of marine-mammal sounds. J. Acoust. Soc. Am. **134(3)**(3): 2546–2555.

Daoust, P.-Y., Couture, E.L., Wimmer, T., and Bourque, L. 2017. Incident Report: North Atlantic right whale mortality event in the Gulf of St. Lawrence, 2017. Collaborative Report produced by: Canadian Wildlife Health Cooperative, Marine Animal Response Society, and Fisheries and Oceans Canada, p. 224 pp.

Davis, G.E., Baumgartner, M.F., Bonnell, J.M., Bell, J., Berchok, C., Bort Thornton, J., Brault, S., Buchanan, G., Charif, R.A., Cholewiak, D., Clark, C.W., Corkeron, P., Delarue, J., Dudzinski, K., Hatch, L., Hildebrand, J., Hodge, L., Klinck, H., Kraus, S., Martin, B., Mellinger, D.K., Moors-Murphy, H., Nieukirk, S., Nowacek, D.P., Parks, S., Read, A.J., Rice, A.N., Risch, D., Širović, A., Soldevilla, M., Stafford, K., Stanistreet, J.E., Summers, E., Todd, S., Warde, A., and Van Parijs, S.M. 2017. Long-term passive acoustic recordings track the changing distribution of North Atlantic right whales (*Eubalaena glacialis*) from 2004 to 2014. Sci. Rep. **7**(1): 13460.

DFO. 2014. Recovery strategy for the North Atlantic right whale (*Eubalaena glacialis*) in Atlantic Canadian waters [Final]. Species at Risk Act Recovery Strategy Series, p. vii +68 pp.

DFO. 2018. Science advice on timing of the mandatory slow-down zone for shipping traffic in the Gulf of St. Lawrence to protect the North Atlantic right whale. DFO Can. Sci. Advis. Sec. Sci. Resp. 2017/042, p. 16 p.

DFO. 2019. Fishing closures for North Atlantic right whale protection: 2019 North Atlantic right whale management measures. Available from http://www.dfo-mpo.gc.ca/fisheries-peches/commercial-commerciale/atl-arc/narw-bnan/index-eng.html [accessed 2019-05-23 2019].

Gervaise, C., Vallez, S., Stephan, Y., and Simard, Y. 2008. Robust 2D localization of low-frequency calls in shallow waters using modal propagation modelling. Can. Acoust. **36**(1): 153-159.

Gillespie, D. 2004. Detection and classification of right whale calls using an "edge" detector operating on a smoothed spectrogram. Can. Acoust. **32**(2): 39-47.

Grieve, B.D., Hare, J.A., and Saba, V.S. 2017. Projecting the effects of climate change on Calanus finmarchicus distribution within the U.S. Northeast Continental Shelf. Sci Rep **7**(1): 6264.

Haykin, S. 1985. Array signal processing. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Ioana, C., Jarrot, A., Gervaise, C., Stéphan, Y., and Quinquis, A. 2010. Localization in underwater dispersive channels using the time-frequency-phase continuity of signals. . IEEE Transactions on Signal Processing **58**(8): 4093–4107.

Kay, S.M. 1993. Fundamentals of statistical signal processing: estimation theory Prentice Hall PTR, Englewood Cliffs, N.J.

Kay, S.M. 1998. Fundamentals of statistical signal processing, Vol. II: Detection Theory Prentice Hall, Upper Saddle River, NJ.

Kinda, G.B., Simard, Y., Gervaise, C., Mars, J.I., and Fortier, L. 2013. Under-ice ambient noise in Eastern Beaufort Sea, Canadian Arctic, and its relation to environmental forcing. J. Acoust. Soc. Am. **134**: 77-87.

Lampert, T.A., and O'Keefe, S.E.M. 2010. A survey of spectrogram track detection algorithms. Appl. Acoust. **71**(2): 87-100.

Loring, D.H., and Nota, D.J.G. 1973. Morphology and sediments of the Gulf of St. Lawrence. Bull. Fish. Res. Bd. Can. **182**: 147 p. + 147 charts.

Matthews, J.N., Brown, S., Gillespie, D., Johnson, M., McLanaghan, R., Moscrop, A., Nowacek, D., Leaper, R., Lewis, T., and Tyack, P. 2001. Vocalisation rates of the North Atlantic right whale (*Eubalaena glacialis*). J. Cetacean Res. Manag. **3**(3): 271-282.

Mellinger, D.K., Nieukirk, S.L., Matsumoto, H., Heimlich, S.L., Dziak, R.P., Haxel, J., Fowler, M., Meinig, C., and Miller, H.V. 2007. Seasonal occurrence of North Atlantic right whale (*Eubalaena glacialis*) vocalizations at two sites on the Scotian Shelf. Mar. Mamm. Sci. **23**(4): 856-867.

Meyer-Gutbrod, E.L., and Greene, C.H. 2018. Uncertain recovery of the North Atlantic right whale in a changing ocean. Global Change Biol. **24**(1): 455-464.

Meyer-Gutbrod, E.L., Greene, C.H., and Davies, K.T.A. 2018. Marine species range shifts necessitate advanced policy planning: The case of the North Atlantic right whale. Oceanography **31**(2): 19-33.

Munger, L.M., Wiggins, S.M., and Hildebrand, J.A. 2011. North Pacific right whale up-call source levels and propagation distance on the southeastern Bering Sea shelf. The Journal of the Acoustical Society of America **129**(6): 4047-4054.

Mussoline, S.E., Risch, D., Hatch, L.T., Weinrich, M.T., Wiley, D.N., Thompson, M.A., Corkeron, P.J., and Van Parijs, S.M. 2012. Seasonal and diel variation in North Atlantic right whale up-calls: implications for management and conservation in the northwestern Atlantic Ocean. Endang. Species Res. **17**(1): 17-26.

NRC. 2003. Ocean noise and marine mammals. National Academy Press, Washington D.C.

Parks, S.E., Hotchkin, C.F., Cortopassi, K.A., and Clark, C.W. 2012. Characteristics of gunshot sound displays by North Atlantic right whales in the Bay of Fundy. J. Acoust. Soc. Am. **131**(4): 3173-3179.

Parks, S.E., Searby, A., Célérier, A., Johnson, M.P., Nowacek, D.P., and Tyack, P.L. 2011. Sound production behavior of individual North Atlantic right whales: implications for passive acoustic monitoring. Endang. Species Res. **15**: 63-76.

Parks, S.E., and Tyack, P.L. 2005. Sound production by North Atlantic right whales (*Eubalaena glacialis*) in surface active groups. J. Acoust. Soc. Am. **117**(5): 3297-3306.

Pettis, H.M., Pace, R.M.I., and Hamilton, P.K. 2018. North Atlantic Right Whale Consortium 2018 annual report card. Report to the North Atlantic Right Whale Consortium, p. 17 pp.

Senneville, S., and Lefaivre, D. 2015. Reproduction horaire et à trois dimensions des conditions hydrographiques et hydrodynamiques du Golfe du Saint-Laurent avec le modèle MoGSL

pour la période de 1997 à 2014. *In* Rapport 2014-2015 de l'entente de contribution entre le MPO et l'ISMER/UQAR pour la simulation numérique des océans. Fisheries and Oceans Canada, Mont-Joli, QC.

Simard, Y., Roy, N., Gervaise, C., and Giard, S. 2016. Analysis and modeling of 255 source levels of merchant ships from an acoustic observatory along St. Lawrence Seaway. J. Acoust. Soc. Am. **140**(3): 2002-2018.

Simard, Y., Roy, N., Giard, S., and Aulanier, F. 2019. North Atlantic right whale shift to the Gulf of St. Lawrence in 2015 as monitored by long-term passive acoustics. End. Sp. Res. **xx**: xxx-xxx.

Simard, Y., Roy, N., Giard, S., and Yayla, M. 2014. Canadian year-round shipping traffic atlas for 2013: Volume 1, East Coast marine waters. Can. Tech. Rep. Fish. Aquat. Sci. 3091(Vol.1)E, p. xviii + 327 pp.

Transport Canada. 2019. Protecting North Atlantic right whales from collisions with ships in the Gulf of St. Lawrence. Available from http://www.tc.gc.ca/en/services/marine/navigation-marine-conditions/protecting-north-atlantic-right-whales-collisions-ships-gulf-st-lawrence.html [accessed 2019-05-23 2019].

Trygonis, V., Gerstein, E., Moir, J., and McCulloch, S. 2013. Vocalization characteristics of North Atlantic right whale surface active groups in the calving habitat, southeastern United States. J. Acoust. Soc. Am. **134**(6): 4518-4531.

Urazghildiiev, I.R., and Clark, C.W. 2007. Acoustic detection of North Atlantic right whale contact calls using spectrogram-based statistics. J. Acoust. Soc. Am. **122**(2): 769-776.

Urazghildiiev, I.R., Clark, C.W., and Krein, T.P. 2009. Detection and recognition of North Atlantic right whale contact calls in the presence of ambient noise. IEEE J Ocean Eng **34**(3): 358-368.

Wenz, G.M. 1962. Acoustic ambient noise in the ocean: Spectra and sources. J. Acoust. Soc. Am. **34**(12): 1936-1956.

Wiggins, S.M., McDonald, M.A., Munger, L.M., Moore, S.E., and Hildebrand, J.A. 2004. Waveguide propagation allows range estimates for north pacific right whales in the bering sea. Can. Acoust. **32**(2): 146-154.

ANNEX 1: OPTIMAL FFT WINDOW LENGTH FOR NARW UPCALL

Here, we demonstrate that an optimal length for the window used to compute the spectrogram exists in order to maximize the processing gain (ratio of the SNR after the computation of the STFT (short-time Fourier transform) with the SNR of the raw data). The demonstration is run for 2 types of signal. The first one is a cosine of duration T and the optimal length is T. The second one is a LFM (NARW upcall like) and the optimal length depends on the modulation rate.

Processing Gain estimation from the spectrogram of monochromatic non-modulated sounds

Our measure is the sum of 1) a monochromatic, non-modulated signal of amplitude A and a duration of M samples centered around time t(o) and a frequency f_0 (with $f_0 = f(p) = pF_e/L$), and 2) of a Gaussian stationary white noise with mean $\mu = 0$ and variance σ^2.
The measure is defined by:

$$n \in [0, N-1]$$
$$m(nTe) = A\exp(2\pi\mathrm{j}f_0 nTe) + b(n) \quad \forall n \in [o - M/2 + 1, o + M/2]$$
$$m(nTe) = 0 + b(n) \quad elsewhere$$

where T_e is the sampling period, and b(n) is a Gaussian stationary white noise.

The SNR of the raw measurements, SNR_{be}, is defined as follows:

$$SNR_{be} = \frac{A^2}{\sigma^2}$$

We compute the STFT of the signal in [t(o),f(p)] with a rectangular window of length L. The analytical computations of the STFT and the spectrogram in the case of target signal-only data, as well as the analytical computations of the STFT and the mean value of the spectrogram of noise-only data are not referred here. The estimation of the SNR_{af} after calculating the STFT at the segment [t(o), f(p)] can be obtained through:

$$\frac{SNR_{af}}{SNR_{be}} = GT = M \times \frac{L}{M} = L; \ \forall L < M$$
$$\frac{SNR_{af}}{SNR_{be}} = GT = M \times \frac{M}{L} = \frac{M^2}{L}; \ \forall L \geq M$$

Figure 30 (box 1) reveals that the processing gain (PG) reaches a maximum when the length of the window L equals the length of the monochromatic signal. This phenomenon is further outlined in figure 5 using real data. The recordings were made in presence of belugas (*Delphinapterus leucas*) in the St. Lawrence Estuary, Canada at a depth of 5 m and with a sampling

frequency of 16384 sa/s. The temporal representation of the measure (Figure 31, A) does not enable the detection of the vocalizations. This is not the case of the spectrogram representation, in which the sounds are clearly visible (Figure 31, panel B) as monochromatic signals (constant frequency) of 1-s duration. One second corresponds to an optimal length L of 16384 samples. This is emphasized when comparing the spectrograms of panels C, computed using L = 2048 and D calculated using L = 16384 (Figure 31). The whistles are better evidenced in panel D than in panel C and have therefore a better SNR.

Processing Gain estimation from the spectrogram of linear frequency-modulated sounds

We address the case of linear frequency-modulated sounds (NARW upcall like, Figure 30, box 2).

In this case, the data are defined by:

$$n \in [0, N-1]$$
$$m(nT_e) = s_u(nT_e) + b(n) \quad \forall n \in [o - M/2 + 1, o + M/2]$$
$$m(nT_e) = 0 + b(n) \; elsewhere$$
$$s_u(nT_e) = A\exp(2\pi\mathrm{j} f_0 nT_e + 2\pi\frac{\alpha}{2}(nT_e)^2)$$

The artificial signal is the sum of a linear frequency-modulation of amplitude A, with a duration of M samples centered around time t(o) and a center frequency f_0 (with $f_0 = f(p) = pF_e/L$) as well as a modulation rate (slope) α (Hz/s), and of a Gaussian stationary white noise with $\mu = 0$ and a variance σ^2.

We suppose that the STFT of the signal in [t(o),f(p)] is calculated with a rectangular window of length L. The analytical computations of the STFT and the spectrogram in the case of a target signal-only data as well as the analytical computations of the STFT and the mean value of the spectrogram of noise-only data are not referred here. The estimation of the SNR_{af} after STFT computing the FFT of the segment [t(o),f(p)] can be assessed from:

$$\frac{SNR_{af}}{SNR_{be}} = GT = \frac{L_0^2}{L}\left|C(\frac{L}{L_0}) + jS(\frac{L}{L_0})\right|^2$$
$$L_0\,(sample) = \sqrt{\frac{2}{\alpha}}\,f_e$$
$$T_0\,(second) = \sqrt{\frac{2}{\alpha}}$$
$$PG_{\max} = PG(L_0) = 0.8L_0$$

where C et S are the Bessel integrals in sine and cosine (Abramowitz and Stegun 1974)

Figure 30 (box 2) reveals that the processing gain (PG) reaches a maximum when the length of the window L equals L_0. This phenomenon is further outlined in (Figure 32) using real data. The recordings were made in presence of bottlenose dolphins (*Tursiop struncatus*) in the Iroise Sea, Brest, France at a depth 5 m and with a sampling frequency of 44000 sa/s. Plot A (Figure 32) shows the temporal representation of the recordings using a bandpass filter starting at 2000 Hz. In the spectrogram representation (Figure 32, plot B), three whistles are clearly visible, all with an almost linear frequency-modulated portion of 0.5 s duration and a frequency band of almost 10 kHz. The modulation rate therefore corresponds to 20000 Hz/s with an optimal STFT length of 512 points. This is emphasized when comparing the spectrograms computed using different window sizes L. Figure 32 illustrates three examples: spectrogram C, computed with a window length L = 128, spectrogram D with L = 512, and spectrogram E with L = 16384.The whistles are better evidenced in panel D compared to the two other panels. The window length that maximizes the SNR of the modulated signal corresponds to the one resulting in a better time-frequency resolution of the STFT. Optimizing the SNR therefore means to concentrate the energy along the modulated portion of the signal (e.g. narrower and higher peaks in panel D than in panel E).

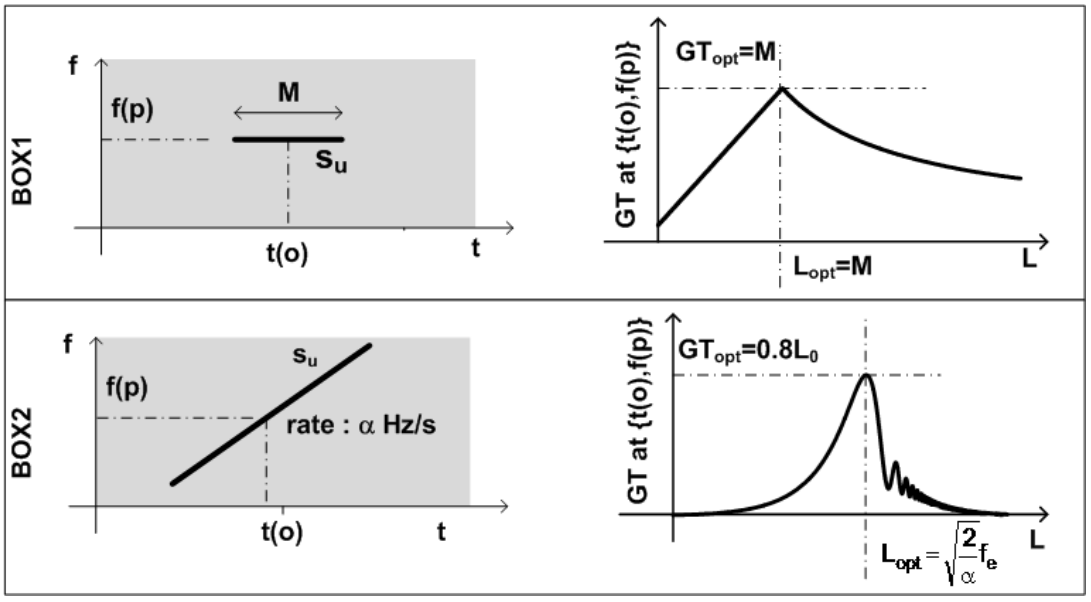

Figure 30.Theoretical SNR and processing gain (PG) of the STFT as a function of the window length L.
BOX1: non-modulated monochromatic pulses; BOX2: signals with linear frequency modulation.

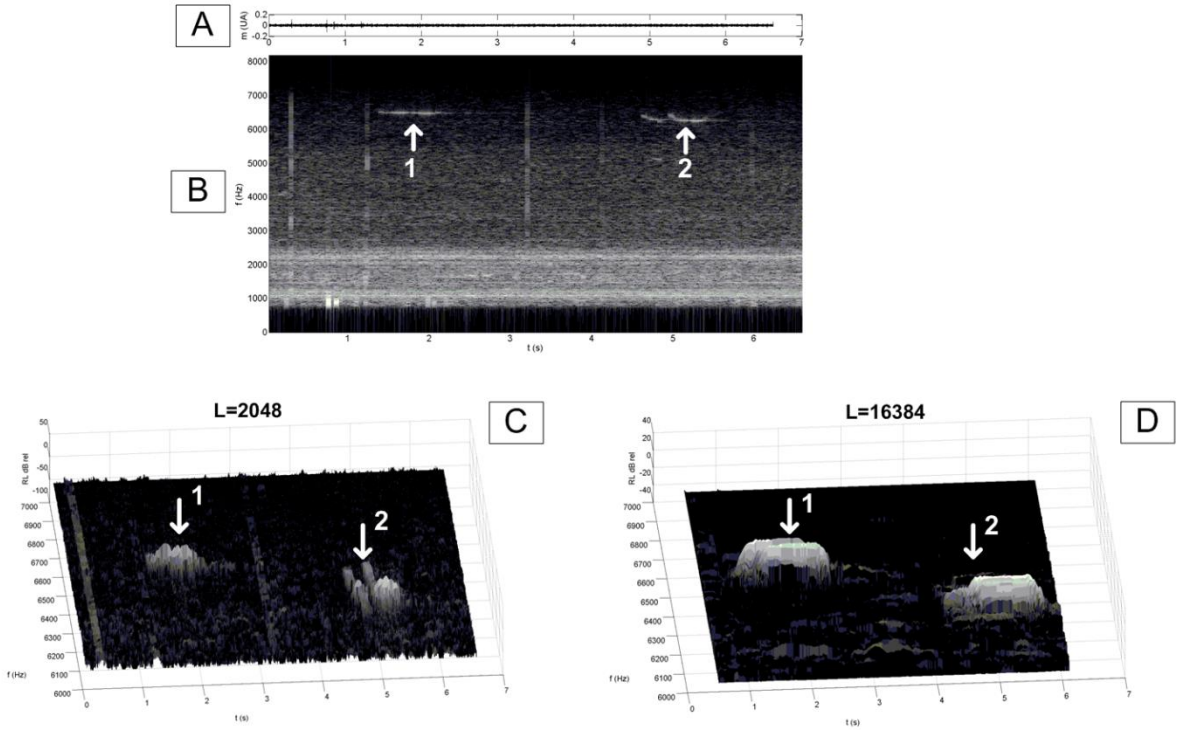

Figure 31. Application of the STFT real recordings containing two whistles of beluga whales (*Delphinapterus leucas*) with constant frequency.
A: raw data, B: 2D spectrogram with optimal length of the FFT window, C: 3D spectrogram with a smaller length of FFT window lower than the optimal one shown at panel D.

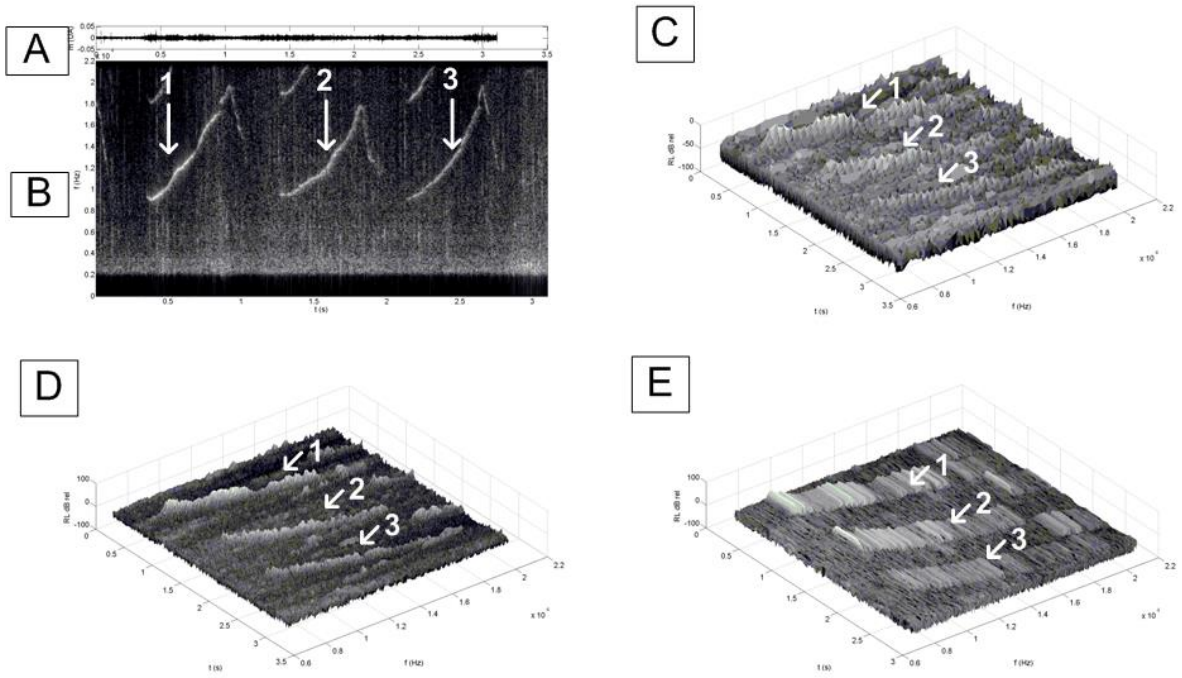

Figure 32. Application of the STFT on real recordings containing three whistles of bottlenose dolphins (*Tursiops truncatus*) with a portion of the sound representing a linear frequency-modulated signal.
A: raw data, B: 2D spectrogram with optimal length of FFT window, C) 3D spectrogram with a smaller window than the optimal one, D) 3D spectrogram with optimal length of FFT window, E) 3D spectrogram with larger window than the optimal one.

ANNEX 2: DETAILS OF ROC CURVES OF THE UPCALL DETECTOR

Here we detail the computation of the ROC of TFBD shown in Table 4. Let the measurements according to the model be:
Hypothesis H_0: $m(t) = b_{tot}(t)$
Hypothesis H_1: $m(t) = uc(t) + b_{tot}(t)$

$b_{tot}(t)$ is the noise and uc(t) is the upcall. $b_{tot}(t)$ is assumed to be random and normally distributed (0 mean, variance σ^2), whereas uc(t) is a deterministic signal:
$uc(t) = 0\ if\ t \notin [0,T], uc(t) = A\cos(2\pi f_{min} t + 2\pi \frac{f_{max}-f_{min}}{2T} t^2)$.

TFBD implies three stages of processing shown in Figure 33:

- Step 1: Computation of the spectrogram,
- Step 2: Binarization at the scale of the pixel of the spectrogram,
- Step 3: Evaluation of the number of time frequency bins coinciding with those inside the time-frequency template of the upcall
- Final decision: test this number against a detection threshold for upcall presence.

Steps 1 and 2 follow the development of Dadouchi et al. (2013). Let L_{fft} be the length (in samples) of the FFT used to compute the spectrogram (chosen to be optimal as in annex 1) and $P_{fa\ pix}$ and $P_{d\ pix}$ the probability of false alarm and detection at the scale of the pixel after stage 2.

Under the hypothesis H0 (noise only), the spectrogram is a random variable and follows a centralized chi^2 law with 2 degrees of freedom (mean $L_{fft}\sigma^2$, standard deviation $L_{fft}\sigma^2$).

Under the hypothesis H1 (noise and upcall), the spectrogram is a random variable and follows a shifted non-centered chi^2 law with 2 degrees of freedom (mean $A^2/2L_{fft}^2+L_{fft}\sigma^2$, standard deviation $L_{fft}\sigma^2$).

Without changing the values of $P_{fa\ pix}$ and $P_{d\ pix}$, we can scale the argument of the law by $L_{fft}\sigma^2$, then we have:

- Under the hypothesis H0 (noise only), the spectrogram is a random variable and follows a centralized chi^2 law with 2 degrees of freedom (mean 1, standard deviation 1).
- Under the hypothesis H1 (noise and upcall), the spectrogram is a random variable and follows a non-centered chi^2 law with 2 degrees of freedom (mean $1 + 0.8\frac{A^2}{2}\frac{L_{fft}}{\sigma^2}$, standard deviation 1).

The factor $\frac{A^2}{2}\frac{L_{fft}}{\sigma^2}$ is equal to $\frac{A^2}{2}\frac{L_{fft}}{\gamma_0 \times f_s} = \frac{A^2}{2}\frac{T_{fft}}{\gamma_0}$, where T_{fft} is the length of the FFT windows in second and γ_0 is the power spectrum of the ambient noise ($\mu Pa^2/Hz$).

Then we have:

$$P_{d\,pix} = 1 - ncX2cdf(-2\log(P_{fa\,pix}), 2, 0.8\frac{A^2 T_{fft}}{2\gamma_0})$$

with:

ncX2cdf(x,p,q): cumulative distribution function computed at x of a non-central chi^2 law with p degrees of freedom and a parameter of non-centrality equal to q

chi^{2}inv(x,p): reciprocal function computed at x of the cumulative distribution function of the centralized chi^2 law with p degrees of freedom.

At the end of step 2, we have a discrete spectrogram and step 3 processes it by looking for the instants where an upcall is present by counting the number of pixels equal to 1 in a time frequency template of the upcall; this number is tested to decide if an upcall is present. Let P be the number of time-frequency pixels needed to cover the time-frequency template of the upcall, we can calculate the probability to have j pixels equal to 1 into the P possible pixels. This is the probability to obtain j draws equal to 1 among P draws knowing that the probability to obtain 1 for 1 draw is $P_{d\,pix}$ under hypotheses H_1 and $P_{fa\,pix}$ under hypotheses H_0.

Considering that time-frequency pixels are independent in first approximation, the number i follows a binomial law:

$$f_Q(j, H_0) = \binom{P}{j} (P_{fa\,pix})^j (1 - P_{fa\,pix})^{P-j}$$

$$f_Q(j, H_1) = \binom{P}{j} (P_{d\,pix})^j (1 - P_{d\,pix})^{P-j}$$

The presence of an upcall (decision D1) is decided if the number of draws among P is more than a threshold λ_2, while the absence of an upcall (decision D0) is decided if the number of draws among P is less than a threshold λ_2, then we have:

$$P_{fa} = 1 - F_Q(\lambda_2, H_0)$$

$$P_d = 1 - F_Q(\lambda_2, H_1)$$

$$F_Q(\lambda_2, H_0) = \sum_{j=0}^{\lambda_2} \binom{P}{j} (P_{fa\,pix})^j (1 - P_{fa\,pix})^{P-j}$$

$$F_Q(\lambda_2, H_1) = \sum_{j=0}^{\lambda_2} \binom{P}{j} (P_{d\,pix})^j (1 - P_{d\,pix})^{P-j}.$$

To evaluate the ROC of TFBD, one must follow the steps:

- set $P_{fa\,pix} = 10^{-3}$, $P_{fa} = 2 \times 10^{-6}$
- from $P_{fa} = 1 - F_Q(\lambda_2, H_0)$ find λ_2
- from $P_{d\,pix} = 1 - ncX2cdf(-2\log(P_{fa\,pix}), 2, 0.8\frac{A^2 T_{fft}}{2\gamma_0})$ find $P_{d\,pix}$
- from $P_d = 1 - F_Q(\lambda_2, H_1)$ find Pd.

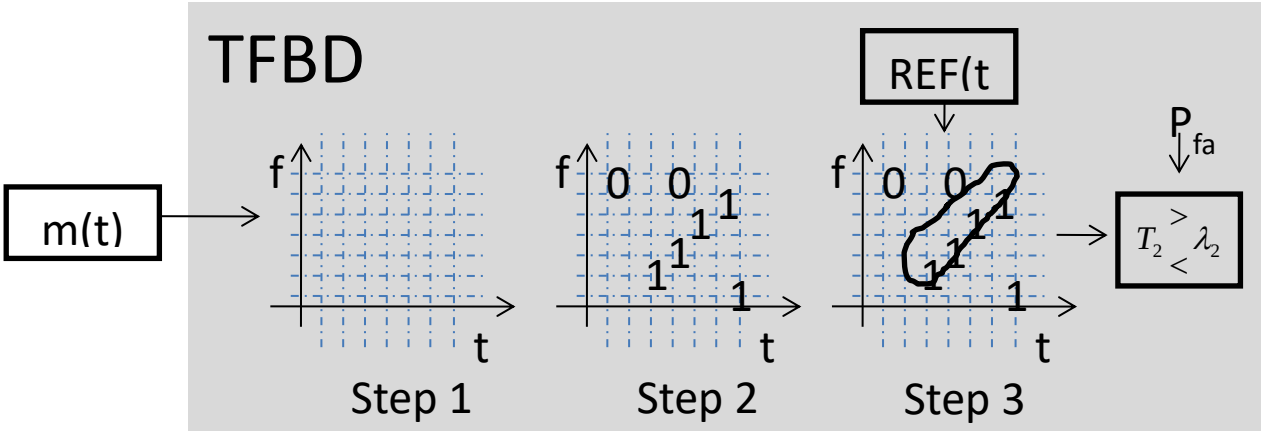

Figure 33. TFBD processing chain.

Step 1 computes the spectrogram of the data, step 2 binarizes the spectrogram at the scale of 1 pixel, step 3 works at the scale of 1 time-frequency region and looks for areas where upcall is present by counting the number of pixels equal to 1 in a time- frequency template of the upcall; this number is tested to decide if an upcall is present.

ANNEX 3: EFFECT ON THE PROCESSING GAIN OF A MISFIT BETWEEN THE REFERENCE AND THE TRUE SIGNAL FOR CCBD

Here we examine how the processing gain of CCBD decreases with the level of misfit between the measurement and the upcall reference used for the cross-correlation. We simulate a measurement made with noise plus a LFM upcall with features (T_0, $f_{min,0}$, $f_{max,0}$) called LFM_0. As optimal processing the measurement is cross-correlated with LFM_0, the maximum of the correlation Γ_0 is noted. Then we simulate a sub-optimal reference as a LFM with features (T, f_{min}, f_{max}) and the measurement is cross-correlated with LFM, then the maximum of the correlation Γ is noted. As LFM is different from LFM_0, $\Gamma < \Gamma_0$, we define the loss of processing gain as:

$$\Delta PG = 10log_{10}(\frac{\Gamma}{\Gamma_0}).$$

To assess the level processing gain loss as a function of the level of misfit between the true and the used references, we define a percentage of deviation of the features (T, f_{min}, f_{max}) vs (T_0, $f_{min,0}$, $f_{max,0}$) and we choose a number N_{MC} of Monte Carlo simulations. For each Monte Carlo simulation, we randomly draw the value of (T, f_{min}, f_{max}), we cross-correlate the measurement with LFM (T, f_{min}, f_{max}), and we evaluate Γ and ΔPG. We keep in memory each ΔPG for each MC simulation. At the end of the N_{MC} Monte Carlo simulations we evaluate the mean ($\overline{\Delta PG}$) and the standard deviation $\sigma\Delta PG$ of ΔPG.

Figure 34 presents $\overline{\Delta PG}$ and $\sigma\Delta PG$ for N_{MC} = 1000 and the percentage of misfit from 0 % to 50%. From 10% of misfit between the features of the LFM, the average loss of processing gain is at least 4 dB.

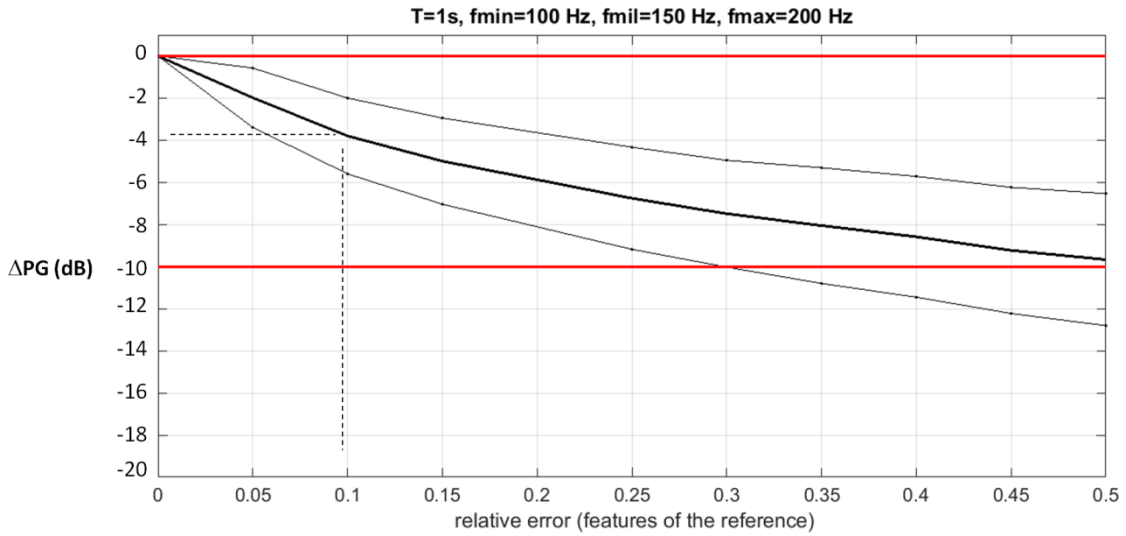

Figure 34. Loss in processing gain as a function of the level of misfit between the true and the used references, thick black curve $\overline{\Delta PG}$, thin black curves: $\overline{\Delta PG} \pm \sigma\Delta PG$.

ANNEX 4: ACCURACY OF LOCALIZATION OF THE NARW WITH THE FORMALISM OF CRAMER RAO BOUNDS

We detail the calculus of the accuracy of upcall localization and the error ellipses, thanks to the formalism of Cramèr Rao Bounds (Kay 1993, p. 29).

We assume:

- a planar 2D horizontal localization with 2 coordinates axis "x" and "y",
- a NARW located at the position (x_w, y_w),
- two arrays of hydrophones located at (x_1,y_1) and (x_2,y_2),
- each array measures the azimuth (a_1, a_2) of the NARW (direction of arrival of the NARW referenced to the East-West line passing through the center of the array),
- the measured azimuth is the true geometrical azimuth plus a random error,
- the random error at array 1 is independent of the error at array 2,
- the random error at array 1 or 2 is normally distributed with 0° mean and 3° standard deviation.

Under these assumptions, the data model is:

$$\alpha 1_m = \alpha 1_t(x_w, y_w, x_1, y_1) + d\alpha 1$$
$$\alpha 2_m = \alpha 2_t(x_w, y_w, x_2, y_2) + d\alpha 2$$

with:

$$\alpha i_t(x_w, y_w, x_i, y_i) = \text{atan}(\tfrac{y_w - y_i}{x_w - x_i}),$$
$$d\alpha 1, d\alpha 2 \ \# \ N(0, 3°),$$
$$pr(\alpha 1, \alpha 2) = pr(\alpha 1) \times pr(\alpha 2).$$

Under these assumptions, the Cramer Rao Bounds set a lower bound on the error of estimation of any unbiased (xe_w,ye_w) estimator of (x_w,y_w) from ($\alpha 1_m$, $\alpha 2_m$) such that:

$$COV(\text{xe}_\text{w}, \text{ye}_\text{w}) - \text{CRB} \geq 0 \ .$$

and

$$CRB = \sigma_\alpha^2 [FI]^{-1}$$

$$Fisher\ Information\ Matrix = [FI] = \begin{bmatrix} \frac{d\alpha 1}{dx_w} & \frac{d\alpha 2}{dx_w} \\ \frac{d\alpha 1}{dy_w} & \frac{d\alpha 2}{dy_w} \end{bmatrix}$$

$$\text{atan}(x)' = \frac{1}{1 + x^2}$$

We know that the Maximum Likelihood Estimator of (x_w,y_w) asymptotically ($SNR \to \infty$) reaches the CRB, so we assume that the CRB is a good proxy for the error of estimation of (x_w,y_w) from ($\alpha 1_m$, $\alpha 2_m$).

CRB is a 2 × 2 matrix where:

- $\sqrt{CRB(1,1)}$ is the standard deviation of measurement error of xw,
- $\sqrt{CRB(2,2)}$ is the standard deviation of the measurement error of yw,
- The eigen vectors of CRB are the axis of the error ellipse,
- The eigen values of CRB are the dimension of the error ellipse along the eigen vectors.