North Atlantic Right Whale upcall Localization with a Four-Element Acoustic Array on a Slocum Glider

Michael MacGillivray[1], Mae L. Seto[1], S. Bruce Martin[2], and Laird Bolt[2]

[1]Dalhousie University
[2]JASCO Applied Sciences

Michael D MacGillivray, Dalhousie University, 5269 Morris St, Halifax, NS B3J 1B6
Email: michael.macgillivray@dal.ca

Abstract: North Atlantic right whales (NARW) are critically endangered, with an estimated 350 individuals remaining. Passive acoustic monitoring offers a means to localize the whales, improving the accuracy of exclusion zones to mitigate vessel strikes and fishing entanglements. This study evaluates multiple time-difference-of-arrival azimuth estimation methods for a compact volumetric array (CVA) integrated onto a Slocum glider. The methods tested include minimum difference estimation, k-means clustering, kernel density estimation (KDE), TDoA maximum likelihood estimation, and azigram-KDE estimation. Simulations in a ROS/Gazebo environment, incorporating realistic underwater acoustic conditions, are used to test each of the methods. Verification was achieved with an in-water trial off the coast of Clam Harbour, Nova Scotia, Canada (2024). The NARW upcall is emulated using a moored source producing linear frequency modulated pulses. The trial revealed that the azigram-KDE estimator performed the best, yielding a mean absolute error of 25.6° ***(20.5 % improvement over next best method) and an interquartile range of*** 10.2° ***(40.7 % improvement). Successful localization of NARWs with a glider-mounted CVA advances autonomous marine mammal monitoring and supports ongoing conservation efforts.***

Keywords: *passive acoustic localization, underwater glider, compact volumetric arrays, array processing, signal processing*

1. INTRODUCTION

North Atlantic right whales (NARW) are a critically endangered large whale species with an estimated 350 individuals in existence. To decrease the probability of human-related fatalities, Canada enforces multiple mitigation measures including vessel speed regulations and fishing closures. For instance, upon detection of a NARW, a 2000 km^2 area centred around the point of detection is designated as an exclusion zone for fishing activities [1].

Passive acoustic localization of their vocalizations can more accurately define the exclusion zone around a detected animal compared to centreing around the detectors location. Based on the NARW underwater vocalizations, its relative azimuth and elevation angles with respect to an acoustic receiver can be estimated [7]. To fully define the location of the animal in 3-D space, a range estimate is required. Rough range estimations can be determined from the sonar equation. To calculate this, estimates of vocalization source level (SL), propagation loss (PL), and ambient noise levels (NL) are required. However, the scope of this work pertains only to azimuth estimations of NARW vocalizations.

An emerging and innovative approach uses underwater gliders (UG) as acoustic sensor platforms for passive localization. The Slocum UG (Teledyne Webb Research, Falmouth, MA, USA) is an autonomous underwater vehicle with a variable ballast system to propel itself through the water column in a sawtooth pattern. Integrating UGs with a compact volumetric hydrophone array (CVA) enables passive azimuth estimation of mobile underwater acoustic targets. Using the UG, compared to stationary sonar systems such as sonobouys, enhances localization duration, increases situational awareness for vessel traffic and fishing activity, and makes autonomous mission re-planning meaningful towards longer-duration localization. Prior studies demonstrated the effectiveness of UGs with CVAs in localizing surface vessels and low-frequency acoustic sources operating at depth within the 50–200 Hz band [5]. The application of UGs for marine mammal detection, monitoring, and localization is also documented in the literature [3]. However, there is less development on this for accurate azimuth estimations given irregular and short-duration acoustic sources, such as NARW upcalls.

This study evaluates five methods to estimate the azimuth of NARW upcalls relative to a CVA on an UG. The Slocum G3 UG is integrated with 4 hydrophones, to form a CVA. The azimuth estimation methods include minimum difference estimation, k-means clustering, kernel density estimation (KDE), TDoA maximum likelihood estimation (MLE), and azigram-KDE estimation [6]. Initial algorithmic verification is conducted within a ROS/Gazebo simulation environment [2][8]. Then, the five methods are verified in-water using a stationary underwater acoustic source that emits one-second linear frequency modulated (LFM) pulses to emulate a NARW upcall.

2. PASSIVE ACOUSTIC AZIMUTH ESTIMATION

Passive acoustic azimuth estimation analyzes time, or phase, delays of signals received from multiple hydrophones of known locations (in this case, on the UG) to estimate its azimuth to an underwater source. Assuming the source is in the far-field, plane wave analysis can be applied. The acoustic signal's angle-of-arrival (AoA) to a pair of hydrophones is estimated using equation 1,

$$\theta = \arccos\left(\frac{c \cdot \tau_{12}}{d}\right) \tag{1}$$

where θ is the relative AoA to the hydrophone pair, c is the local speed of sound underwater, τ_{12} is the time-difference-of-arrival (TDoA) estimated from the cross-correlation of the signal arrivals at the two hydrophones, and d is the Euclidean distance between the hydrophones. There is directional ambiguity for a pair of hydrophones, addressed by adding additional hydrophones on a different axis to form a CVA. The next section discussed method to resolve the best azimuth estimate from the CVA.

2.1. AZIMUTH SOLVERS

The performance of five TDoA-based azimuth estimation methods are compared and contrasted for a CVA mounted on an UG. Each method offers a distinct strategy for resolving the directional ambiguity inherent in hydrophone-based measurements. The methods can be broken down into ambiguous azimuth estimation based (first three) and MLE based solvers (last two).

The first method is the minimum-difference solver. The four-element CVA yields 12 ambiguous azimuth estimates. All valid combinations of 3 estimates are generated, excluding pairs from the same hydrophone baseline. Each group's mean angular difference (spread) is computed, and the group with the smallest spread determines the final azimuth.

The second method uses K-Means to cluster the azimuths based on the angular difference. Once completed, the cluster with the minimum angular difference is selected. If the number of values in the cluster is greater than five, the values are sent through the minimum difference solver. If the size is less than 5, the angular mean is taken as the azimuth estimate.

The third method uses a KDE to solve for the optimal azimuth. The probability density function (PDF) of the group of 12 is determined using a von Mises kernel. The von Mises kernel is employed as it is specifically designed for circular data [6]. The peak value of that PDF is the azimuth estimation.

The fourth method implemented is a TDoA-MLE [7]. An array of expected time differences is generated for all possible azimuths, at a specific angular resolution. The optimal azimuth estimation is obtained by using least squares (LS) optimization.

The final method utilizes an azigram-KDE introduced in [6], but adapts the algorithm for CVA azimuth estimation compared to its original use with sonobuoys. This method first constructs an azigram. To do this, short-term Fourier transforms for each signal are computed. A threshold filter is applied to the spectrograms removing uncorrelated noise. Sub-band processing is used, therefore each time-frequency bin is analyzed individually. For each bin, the phase difference is calculated for all possible hydrophone pairs, creating a group of six phase differences. The group is sent to a MLE, which will yield an azimuth estimation. Phase differences are used to estimate the optimal angle because the expected TDoA can corresponds to phase shifts of multiple wavelengths at higher frequencies.

The remaining cells in the azigram are used to generate an azimuth density estimation. Similar to the third method, the von Mises KDE is applied to the azigram results, and the peak of the resulting PDF is the azimuth estimate.

3. NARW LOCALIZATION ROS/GAZEBO SIMULATIONS

The simulation's purpose was to test and validate the five TDoA-based azimuth estimation methods before using them with in-water trial data. The simulation environment was built using

ROS Noetic and Gazebo 11. The Gazebo simulation is built on top of the DAVE project [8]. The UG's mission consisted of pre-defined waypoint tracking. However, the UG is capable of autonomously updating its waypoints based on underwater acoustic localizations, paving the way for future work focused on real-time tracking of marine mammals.

The simulates NARW produces upcalls modelled as LFM pulses from 100 - 200 Hz, with a SL of 165 dB re $1\mu Pa$. According to [7], the Cramér-Rao lower bound azimuth variance does not depend on the frequency modulation of the signal and is determined by the SNR and the minimum and maximum frequencies of the signal spectrum, making the LFM signal an appropriate substitute for simulations. The vocalizations are simulated as time-series data sampled at 16 kHz. Gaussian noise is added to the signals based on SNR calculated using PL estimates from the BELLHOP Acoustic Toolbox [4] and ambient levels obtained from typical sound level distributions measured on a Slocum glider, provided by JASCO Applied Sciences. The localization simulator also includes Gaussian error terms for hydrophone synchronization ($X[\mathrm{s}] \sim \mathscr{N}(0, (\frac{2}{f_s})^2)$ and UG heading ($X[^\circ] \sim \mathscr{N}(0, 5^2)$). The simulation trials consisted of 5,000 localization measurements spanning source-receiver ranges of 100 m to 10 km and azimuths from -180° to 180° relative to the UG. The performance metrics are median error, mean absolute error (MAE), and interquartile range (IQR).

3.1. SIMULATION RESULTS

The results for each azimuth estimation method considered are shown in Table 1. All methods performed well in simulation. Performance-wise, the azigram KDE and MLE methods slightly edged out the others. These results are expected as both methods use the LS-MLE which is optimal for Gaussian error estimations – the case in simulation [7]. Note, real underwater acoustic environments introduce a range of complexities not captured in simulation, which may influence the effectiveness of the algorithms.

azimuth estimation method	**median error** (°)	**MAE** (°)	**IQR** (°)
minimum difference estimator	-0.81	13.1	18.7
k-means clustering estimator	-0.17	14.6	19.6
KDE estimator	-0.65	11.7	17.3
maximum likelihood estimator	-0.20	10.0	15.3
azigram KDE estimator	-0.65	10.2	15.7

Table 1: Simulated azimuth solver results (100–200 Hz). All methods have relatively high accuracy with the MLE and azigram-KDE performing slightly better, as expected.

4. MOORED SOURCE LOCALIZATION IN-WATER TRIALS

In March 2024 JASCO Applied Sciences completed a trial using an UG integrated with a four-element CVA and a JASCO OceanObserver that was performing real-time detection of the signals for telemetry to shore. The hydrophones were mounted on the UG's nose, tail, port side wing, and starboard side wings. The array geometry was defined by the following 2D coordinates (in meters): $[0.0, 0.0]$, $[0.04064, -2.3]$, $[-0.5, -1.6]$, $[0.5, -1.6]$.

The UG travelled both towards, and away from, a moored, calibrated omnidirectional source projecting acoustic LFM pulses from 700 - 850 Hz and 1250 - 1500 Hz. Higher frequencies were used compared to the simulation due to limitations of the source. The range between the UG and source varied from 125 m to 12 km. The source was at an altitude of 6 metres above the sea floor and at a depth of $\approx$ 70 metres. A repeating sequence of three source signals was projected during the trials: signal A (low-frequency LFM at 165 dB re 1 μPa) followed by a 1-min pause; signal B (low-frequency LFM at 170 dB re 1 μPa) followed by a 1-min pause, and signal C (high-frequency LFM at 170 dB re 1 μPa) followed by a 3-min pause.

4.1. IN-WATER TRIAL RESULTS

The experimental results are presented in Table 2. The first clear trend seen is an increase in azimuth MAE compared to the simulation results. This due to unmodeled uncertainties in the underwater environment, causing a number of outlier measurements. In future work with the simulator, these errors will be included to better reflect real-world conditions.

Among the five methods evaluated, the azigram-KDE estimator demonstrated the best overall performance, achieving a 20.5% reduction in MAE and a 40.7% reduction in IQR compared to the next best method (KDE estimator). Sub-band processing proved effective for accurate azimuth estimation in the presence of in-band noise. In contrast, the performance of the other broadband cross-correlation methods degraded as both correlated and uncorrelated noise levels increased.

azimuth estimation method	**median error** (°)	**MAE** (°)	**IQR** (°)
minimum difference estimator	2.70	34.8	20.7
k-means clustering estimator	-1.32	35.0	21.5
KDE estimator	-2.28	32.2	17.2
maximum likelihood estimator	-1.32	39.8	49.2
azigram KDE estimator	-1.06	25.6	10.2

Table 2: In-water azimuth estimation results. The azigram-KDE estimator outperformed all other methods, while the MLE estimator underperformed due to the Gaussian error assumption being unsuitable.

The MLE performed worse out of all the methods, due to its Gaussian error assumption based on LS optimization. A Gaussian TDoA error is invalid, due to uncertainties in the underwater channel and the underwater acoustic ambient, which introduce outliers that LS optimization is particularly sensitive to. The solver was adapted to use least absolute error (LAE), assuming a Laplacian distribution. Changing the loss function improved the MAE and IQR to 34.8° and 19.9°, respectively. Additionally, the azigram-KDE estimator increase performance when using LAE as its loss function, with a MAE of 22.4° and a IQR of 9.89°.

From Table 3, the higher frequency (HF) signal resulted in a higher azimuth MAE. This is an interesting result, as it is expected the signal with higher frequency and higher bandwidth would provide a more accurate azimuth estimation. The higher MAE can be attributed to increased risk of spatial aliasing and attenuation/signal distortion for the higher frequency signal. However, this is not expected to pose a significant issue for localizing NARW upcalls, which predominantly occur within the 100–200 Hz frequency range. Additionally, it is seen that signal B performs better than the other two signals. This shows clearly that increasing SNR will improve the azimuth estimation accuracy.

signal type	median error (°)	MAE (°)	IQR (°)
Signal A (low-frequency LFM at 165 dB re 1 μPa)	1.21	23.3	9.27
Signal B (low-frequency LFM at 170 dB re 1 μPa)	0.407	17.7	8.03
Signal C (high-frequency LFM at 170 dB re 1 μPa)	-0.797	30.8	9.95

Table 3: Azigram-KDE results for each signal. The LF signals performed better than HF, indicating spatial aliasing or higher attenuation/signal distortion at higher frequencies. Signal B (LF and higher SL) performed best, indicating that higher SNR will improve localization accuracy.

5. CONCLUSION

This work examined five methods to solve for the azimuth angle between a UG and a vocalizing NARW producing upcalls. The azimuth estimation methods were first verified and tested in a simulation environment, where all methods demonstrated reliable and consistent performance. In future work the simulation environment will be used for tracking simulation.

In-water trial data was used to verify the azimuth estimation algorithms. The results show the azigram-KDE estimator outperforms the other broadband TDoA algorithms. While this method can exhibited spatial aliasing at higher frequencies, it is not expected to pose a significant issue when localizing NARW upcalls, which primarily occupy lower frequency bands.

REFERENCES

[1] Government of Canada. Protecting the north atlantic right whale: Speed restriction measures in the gulf of st. lawrence - ssb no.: 10/2024, 2024. Accessed: Apr. 2024.

[2] N. Koenig and A. Howard. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *IEEE/RSJ IROS 2004*, volume 3, pages 2149–2154 vol.3, 2004.

[3] Elizabeth T. Küsel and et al. Marine mammal tracks from two-hydrophone acoustic recordings made with a glider. *Ocean Science*, 13(2):273–288, 2017.

[4] Acoustical Society of America. Ocean acoustics library - oalib, February 2024. Accessed: Dec. 2024.

[5] Dajun Sun and et al. Low frequency three-dimensional doa estimation for underwater gliders using an arbitrary tetrahedral array. *Applied Acoustics*, 214:109707, 2023.

[6] Kamden Thebeau. Localization of north atlantic right whales using a deformable sonobuoy array. Master's thesis, Dalhousie University, December 2024.

[7] Ildar R. Urazghildiiev, Bruce Martin, and David E. Hannay. The accuracy of bearing estimates of wideband signals produced by marine animals. *IEEE Journal of Oceanic Engineering*, 46(3):1057–1067, 2021.

[8] Mabel M. et al. Zhang. Dave aquatic virtual environment: Toward a general underwater robotics simulator. In *IEEE/OES AUV Symposium*, 2022.