PASSIVE ACOUSTIC DETECTION AND LOCALIZATION OF MESOPLODON DENSIROSTRIS (BLAINVILLE'S BEAKED WHALE) VOCALIZATIONS USING DISTRIBUTED BOTTOM-MOUNTED HYDROPHONES IN CONJUNCTION WITH A DIGITAL TAG (DTAG) RECORDING

Jessica Ward[1], Ronald Morrissey[1], David Moretti[1], Nancy DiMarzio[1], Susan Jarvis[1], Mark Johnson[2], Peter Tyack[2], and Charles White[3]
1 - Naval Undersea Warfare Center Division Newport, 1176 Howell St., Newport, R.I., USA
2 – Woods Hole Oceanographic Institute, Woods Hole, MA., USA
3 – University of Rhode Island, Narragansett, R.I., USA

ABSTRACT

Click data from a tagged *Mesoplodon densirostris* was compared with broadband acoustic recordings from an 82 hydrophone wide-baseline array located in the Tongue of the Ocean, Bahamas. Two detectors, a Fast Fourier Transform (FFT) based detector and matched filter, were evaluated in white noise and with the acoustic recordings from the array for performance detecting *M. densirostris* clicks. The matched filter performed the best, allowing 92% of the tagged animal's clicks to be detected on at least one hydrophone. Time Difference of Arrivals (TDOAs) between the DTag and the surrounding hydrophones were computed. These TDOAs were used to compute a three-dimensional hyperbolic localization track of the tagged animal. A maximum detection range of 6500 m from the tagged animal to the recording hydrophone was observed. Offset aspect angles were determined from the DTag heading information and the bearing to the receiving hydrophone. Clicks within ±30 degrees were detected at the farthest ranges, while clicks were detected at all off-set angles at closer ranges.

SOMMAIRE

Les données de « clics », obtenues pour un Mesoplodon densirostris marqué, ont été comparées à des enregistrements acoustiques à large bande obtenus à l'aide d'un vaste réseau de référence à 82 hydrophones situés dans la Langue de l'Océan (Bahamas). Deux détecteurs, soit un détecteur à transformée de Fourier rapide (FFT, de l'anglais Fast Fourier Transform) et un filtre adapté, ont été évalués en bruit blanc et avec les enregistrements acoustiques obtenus grâce au réseau pour détecter les « clics » des M. densirostris. Le filtre adapté a réalisé la meilleure performance, ayant permis de détecter 92 % des clics de l'animal marqué au moins avec un des hydrophones. La différence entre les temps d'arrivée entre le DTag et les hydrophones avoisinants ont été calculées. Les différences ainsi calculées ont été utilisées pour effectuer un suivi tridimensionnel et hyperbolique des déplacements de l'animal marqué. Une plage de détection maximale de 6 500 m entre l'animal marqué et l'hydrophone enregistreur a été observée. Les angles correcteurs ont été déterminés à partir de l'information du DTag et le relèvement géographique a été effectué par rapport à l'hydrophone récepteur. Les clics se trouvant à l'intérieur des angles de ±30 degrés ont été détectés aux distances les plus éloignées, alors que des clics se trouvant à l'intérieur de tous les angles correcteurs ont été détectés à des distances plus courtes.

1. INTRODUCTION

On October 23, 2006, a Woods Hole Oceanographic Institute (WHOI) DTag was placed on one individual in a group believed to consist of four *Mesoplodon densirostris* in the Atlantic Undersea Test and Evaluation Center (AUTEC) underwater tracking range. The tag remained attached for approximately 19 hours over which time seven deep dives were recorded. Whale vocalizations were simultaneously monitored using 82 AUTEC bottom-mounted hydrophones. The hydrophones are at depths of ~2 km and are separated by ~4 km baselines. Two detectors are evaluated for use in *M. densirostris* click detection: a FFT based detector and matched filter. Detection performance is first evaluated in the presence of Gaussian white noise, and then with hydrophone recordings corresponding to the tagging event.

2. METHODS

2.1 DTag Data Description

The DTag was attached to a probable female *M. densirostris* in a group believed to consist of two mother-juvenile pairs at 11:37:38 a.m. (+/- 5 seconds) local time on October 23, 2006. The tagging GPS location was 24° 30.412' N, 77° 35.320' W. A male *M. densirostris* may have been in the vicinity. The DTag recorded stereo audio at a 192 kHz sampling rate with an audio sensitivity of –171 dB re

1.0/μPa. The pitch, roll, heading, and depth of the whale were determined from the accelerometer, magnetometer, and pressure sensors sampled at 50 Hz. The DTag measurements were processed using the methods described in Johnson and Tyack [1] resulting in orientation and depth data with 5 Hz resolution.

2.2 AUTEC Hydrophone Array

Prior to and for the duration the tag was attached to the whale, audio data from the 82 bottom mounted hydrophones of the AUTEC tracking range were simultaneously recorded digitally on multiple Alesis HD24 hard drive recorders at a 96 kHz sampling rate. Each recorder can accommodate 12-channels of data with the last channel recording an IRIG-B modulated time signal.

The hydrophones are mounted 4-5 meters off the sea floor with an upward, roughly hemispherical, beam pattern. There are 68 wideband hydrophones with a usable bandwidth from 50 Hz to approximately 45 kHz. There are an additional 14 hydrophones with a bandwidth from roughly 8 kHz to over 50 kHz installed in two 7 hydrophone arrays. Hydrophone data is digitized at a sampling rate of 96 kHz. This is a standard audio rate that allows for Nyquist sampling of the wideband hydrophones. The upper 2 kHz of the 14 wider bandwidth hydrophones is aliased. This folding has not been found to have a significant effect on the determination of click arrival times.

2.3 Detection

M. densirostris produce echolocation clicks with a frequency modulated upsweep. The peak source level of similiarly sized delphinids has been estimated at 220 dB re. 1 μPa [2]. Tag data from another species of beaked whale (*Ziphius cavirostris*) indicate a pronounced beam pattern with a 3 dB beam width of 6° [3]. Due to the narrow beam width, determination of the arrival time of a specific click on multiple hydrophones in a widely spaced array such as at AUTEC is a challenge. The hyperbolic localization technique used requires a minimum of three hydrophones for a two dimensional position to be determined. Improved detector performance is critical in order to maximize the probability that a given click will be detected on enough hydrophones to produce a position. Accordingly, two detection methods, a FFT based detector and a matched filter, were compared with respect to detection performance. Detection performance was first compared in Gaussian white noise, and then on recorded data in the vicinity of the tagged animal.

DTag Click Detector

Clicks recorded on the DTag were classified as belonging to the tagged whale based on two features. The attachment of the tag to the whale results in a low-frequency energy component that is not present in clicks from conspecific whales [3]. Second, the angle of arrival for clicks from the tagged whale is close to zero between the two hydrophones on the tag, while it varies as the whale moves for clicks from conspecifics [4].

FFT based detector

A multi-stage FFT based energy detector has been successfully used for detection of clicks from a variety of echo-locating odontocetes, including sperm whales [5] and beaked whales [6]. A 2048 point FFT with a 50% overlap is used for this analysis. At the 96 kHz sampling rate this provides a frequency resolution (per bin) of 46.875 Hz and a time resolution (per FFT) of 10.67 ms. Each bin of the FFT is independently thresholded against an exponentially decaying time average of the data in that bin as given in Eq. 1:

n>0: $NVT[n] = (1-a)bin[n] + aNVT[n-1]$ *Eq. 1a*

n=0: $NVT[n] = 0$ *Eq. 1b*

where, the parameter *a* has been chosen empirically to provide a time constant of 0.2 seconds.

The binary output of the thresholding process is combined into a single detection report. If any of the 1024 bins have passed threshold, the first stage declares a detection and passes the detection report on to the next stage.

The output of the first stage of the detector is then examined to determine whether the event was triggered by a beaked whale. Since clicks are broadband events, the detection report may be broadly classified as a click by counting the number of FFT bins which triggered. Assuming a click event is declared, the frequency content of the thresholded detection report is examined. A set of five frequency bands roughly conforming to species of interest (cut off by the hydrophone response) have been selected, where beaked whales comprise band 2 (Table 1).

Table 1: Frequency bands

Band	*Low Frequency (kHz)*	*High Frequency (kHz)*
1	45	48
2	24	48
3	12	48
4	1.5	18
5	0	1.5

A ratio of bins above threshold to the total number of bins in each band is computed. If band two is selected, then the detection is tentatively classified as a beaked whale. Due to the fact that many of the bands overlap, a second check is performed by examining the number of bins set out-of-band. If this exceeds 10%, then the detection is reclassified as a dolphin as they are more likely to have significant spectral energy below 24 kHz.

Matched filter detector

A linear matched filter can be shown to be the optimal detector for known signals in white gaussian noise [7]. A high signal to noise ratio *M. densirostris* click extracted from the data set was used as the match template. The instantaneous output of the filter is then compared to an exponentially decaying time average of the filter output with a time constant of 0.1 seconds. If the instantaneous output exceeds the time average by a specified threshold, a detection is declared.

False alarm statistics

Both the FFT detector and the matched filter have been implemented as constant false alarm rate (CFAR) detectors. A direct comparison of detection performance between the two requires normalizing the false alarm rates. False alarm statistics have been computed in the presence of white Gaussian noise using the Box-Mueller pseudorandom noise generation algorithm from the GNU Scientific Library (GSL). The false alarm rate was then computed by dividing the number of false detections by the total run time for each threshold.

Two sets of results were compiled for the FFT detector. The first set indicates the performance of the first stage of the detector. As can be seen, this stage runs with a high false alarm rate.

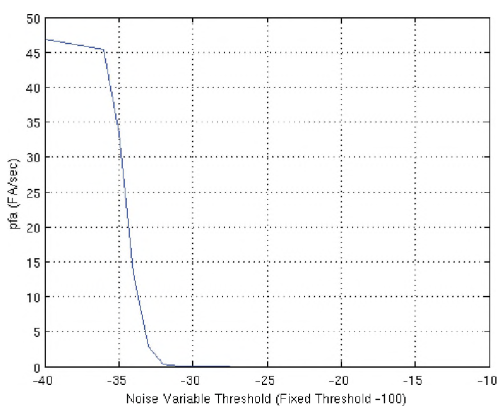

Figure 1: FFT Detector first stage false alarm curve

The false alarm rate drops dramatically at the output of the second stage. The main parameter determining performance is the click threshold used to determine when a sufficient number of bins have been detected to declare a click event.

The matched filter curve is typical and indicates a false alarm rate dropping exponentially with an increase in the threshold.

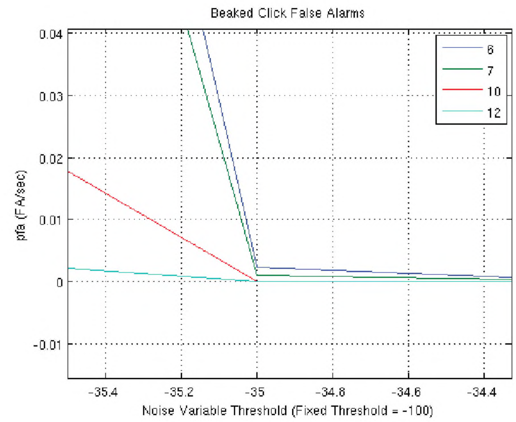

Figure 2: FFT Detector False Alarm Curve after filtering for beaked whales

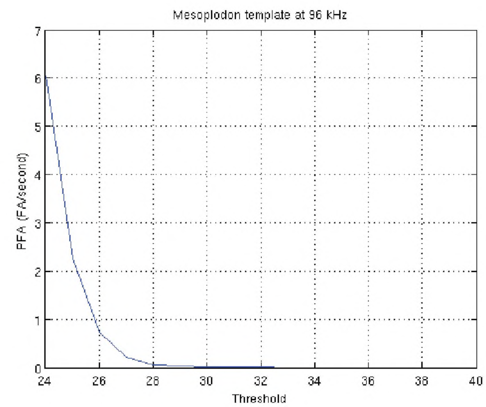

Figure 3: Matched filter false alarm curve

A false alarm rate of $1x10^{-3}$ was chosen as the test criterion. To achieve this rate, the following thresholds were chosen based on the false alarm curves:

FFT Detector: -35.1432
Matched Filter: 28.7009

Probability of Detection

Probability of detection statistics were compiled by creating a series of test data sets consisting of a high SNR click identified as *M. densirostris*. The click was scaled by a specified constant to achieve a desired signal level and repeated 827 times at a regular interval of 4/second. White Gaussian Noise was added to the signal to obtain the desired SNR.

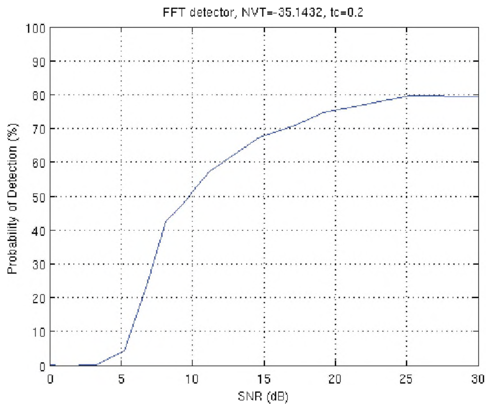

Figure 4: FFT detector probability of detection

The FFT detector tops out at approximately 80% probability of detection. Visual examination of the data indicates that some clicks are not present at the output of the first stage of the detector. The performance deficit at high SNR is therefore most likely linked to the choice of time constant for the exponential noise filter. A lower time constant may improve performance for regular clicks, however this remains to be investigated.

The matched filter provides the expected behavior when the match template exactly matches the signal present in the data set. This is never the case in practice. To estimate the effect of using an arbitrarily chosen high SNR template, a click from a completely separate data set (also collected at AUTEC) was used as a second match template. This is plotted in the rightmost curve in Figure 5. Using an arbitrarily chosen click degrades detector performance by approximately 2-3 dB. In either case probability of detection is at least 95% by 0 dB SNR.

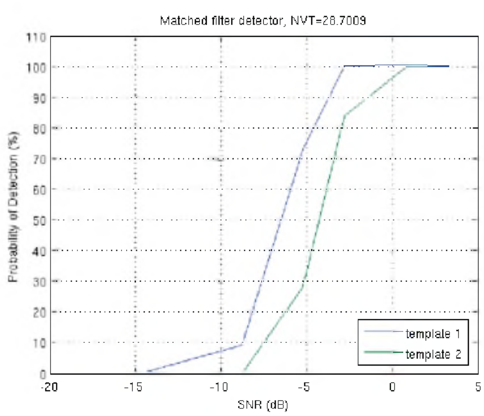

Figure 5: Matched filter probability of detection. Two match templates are plotted. First, template matching the click used to generate the dataset. Second template for high SNR click from a separate dataset.

Conclusion: 'Optimal' Detector

The matched filter significantly outperformed the FFT detector on the test data sets. This was true even when the click used as the match template was not the same as the click used to generate the data sets. The probability of detection data shows a performance gain of at least 25 dB for the test cases studied. This suggests that the click structure is relatively constant with reasonably low variance between clicks from different individuals.

High SNR clicks are typically chosen as match templates. Due to the narrow beam width emitted by the animal, it is expected that most high SNR clicks received will be received when the animal has the ensonified receiver directly in the beam. The structure of these clicks may not be representative of off-axis clicks. In this case the matched filter will be sub-optimal at any aspect angle other than the one at which the match template was tuned for. However, in a widely spaced array such as AUTEC, it is possible to enhance detection and association significantly by improving detection performance at hydrophones which are farther away from the animal, but still in the beam. In this instance the matched filter may be employed to significant advantage.

2.5 Data Association

Clicks originating from the tagged animal have been identified on the surrounding bottom mounted hydrophones by matching inter-click interval patterns [5,8]. These patterns have been found to be an effective means of associating patterns of detections among hydrophones for sperm whales (*Physeter macrocephalus*). A fundamental assumption is that each animal exhibits its own unique pattern of clicks. The unique pattern is used as a template for a comb sieve that is correlated against the beaked whale clicks detected on the surrounding hydrophones [5]. The window with greatest number of correlations between the template and the hydrophone is assigned as the TDOA between the DTag and the hydrophone. After the comb sieve is complete, the probability density function of the TDOAs for each hydrophone is calculated in one minute windows. TDOAs that are significantly above the noise level in each window are passed on for use in localization and considered valid. The remaining TDOAs are considered invalid and not used further.

2.6 3D Hyperbolic Localization

Positions are computed from the valid TDOA sets using a hyperbolic multilateration positioning algorithm developed by Vincent [9]. Two 2500-ft depth XBT profiles were collected on 23 October 2006. These profiles were combined with a standard deep water profile and converted to sound speed by AUTEC. The sound speed profile nearest to the tagging location was used for calculating the direct path effective sound velocity [9]. TDOAs are required between at least four hydrophones and the DTag to compute a 3-D position. Due to the directional nature of the clicks, there were very few instances when an individual click was correlated on 4 hydrophones. Therefore, the TDOA for each hydrophone was interpolated using a piecewise cubic

Hermite interpolating polynomial function. The four hydrophones with the greatest number of valid time-difference of arrivals, 37, 43, 44, and 50, were used to create a Time of Arrival (TOA) matrix. The time of emission, x, y and z position of the sound source were estimated using the TOA matrix in the hyperbolic multilateration algorithm.

3. DISCUSSION

All results discussed in this paper are for the first deep dive recorded on the DTag, from approximately 34 to 91 minutes after tagging. The whale began vocalizing 6 minutes into the dive at 567 m depth, and continued to vocalize for approximately 36 minutes between 567 and 1049 m depth (Figure 6).

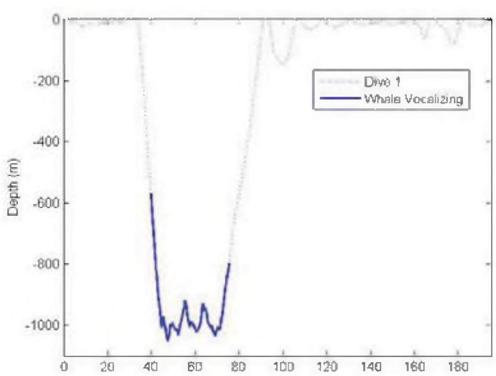

Figure 6: DTag Dive Profile Depth (m) vs. Time (minutes)

3.1 Detection Efficacy

Detection efficacy was assessed by evaluating how many of the clicks emitted by the tagged whale were successfully detected and associated on the nearby hydrophones. For each method, hydrophones 36, 37, 38, 42, 43, 44, 49, and 50 were processed through each detector and the beaked whale filter. These hydrophones were chosen by visually evaluating thresholded spectrogram data over the entire range for the presence of beaked whale clicks for the duration of the tagging event. Whale click times from the DTag were used as the template to search for correlation with the resulting TOAs produced using the FFT and matched filter detectors. Using the association algorithm detailed in Section 2.5, TDOAs were calculated between each hydrophone and the DTag.

During the 36 minute first dive, 5797 clicks were produced by the whale. Approximately 97% of these clicks were foraging clicks with an Inter-Click Interval (ICI) between 0.15 and 1 sec [3]. The mean foraging click ICI was 0.31 sec (std=0.05). This is in agreement with Johnson [4], who also observed a regular click ICI 0.37 seconds for a *M. densirostris* in the Canary Islands.

The FFT detector, implemented with a noise variable threshold of 34, was able to detect 49% of the clicks on at least one hydrophone. The matched filter detector, implemented using a threshold of 28.7 was able to detect 92% of the clicks on at least one hydrophone. The filter template was a *M. densirostris* click recorded on an AUTEC hydrophone from a previous year. On each hydrophone, the matched filter detector performed significantly better than the FFT detector (Figure 7).

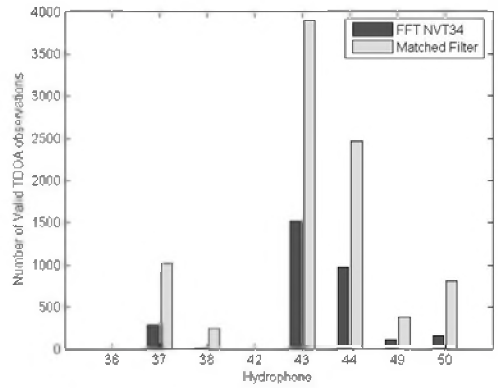

Figure 7: Detection efficacy per hydrophone

3.2 Localization

While the majority of the clicks were detected on at least one hydrophone using the matched filter, three dimensional localization was still difficult due to the need for at least 4 TDOAs between the DTag and hydrophone as input. Of the 5767 clicks associated with the hydrophones, only 1% were detected on four or more hydrophones. More commonly, the clicks were detected on only one hydrophone (44%), two hydrophones (36%), or three hydrophones (11%). Only 8% of the clicks were not detected at all. As a result, the TDOA trends were interpolated as discussed in section 2.6 prior to input into the hyperbolic multilateration algorithm. In addition, depth from the DTag was also used as an initialization parameter to provide better convergence of the solution.

The 577 localizations estimated using at least 3 measured TDOAs and only one interpolated value, were used to "ground-truth" the track estimated by the DTag alone [1] (Figure 8). Usually, the DTag "tag on" and "tag off" positions are known and can be used as absolute start and end positions. However, in this case the tagging vessel had to leave the range due to the presence of range operations and the "tag off" position is unknown. To "ground truth" the DTag track, small user-chosen sections of the track were individually fitted to time-synchronized 3D localizations by adjusting the swim speed to a least-squares match (Figure 9).

3.3 Detection Range

The 3D localizations created using the matched filter detection data were used to determine the range from the whale to the hydrophone for each detection method. Clicks from each hydrophone determined to be valid on the basis of their TDOA with the DTag were used to estimate range and bearing to the hydrophone. The maximum detection range for both methods was approximately 6500 m, significantly greater than previously estimated [10]. The whale was traveling generally in a north-east direction, but was observed to turn at various times in all directions (Figure 10). The off-axis aspect angle between the caudal-rostral axis of the tagged whale and the receiving hydrophone was determined by subtracting the bearing angle from the whale to the hydrophone from the heading measured by the DTag. The detection range as a function of off-axis aspect angle is depicted in Figure 11 and Figure 12. The majority of the clicks detected at far ranges were within ±30 degrees. With decreasing range, a greater number of clicks were detected further off-axis. While a -3 dB beam width of 6° has been suggested for *Z. cavirostris* by Zimmer [3], *M. densirostris* may be less directional due to their smaller body size and potentially smaller source aperature [10]. From these figures, it is evident that the matched filter detector significantly outperforms the FFT detector at longer ranges.

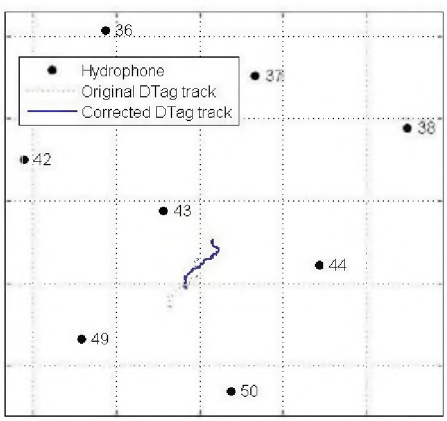

Figure 8: Original DTag Kalman-filtered track and corrected DTag track, each grid square is 2 km x 2 km.

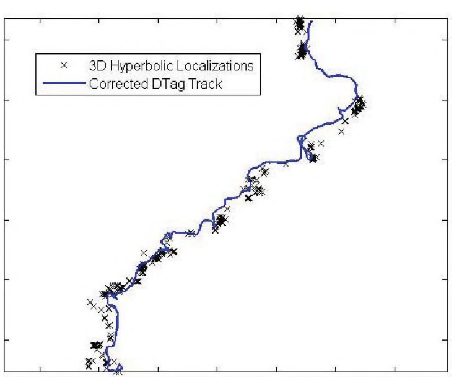

Figure 9:DTag track corrected based on 3D hyperbolic localizations, tick marks at 200-m increments

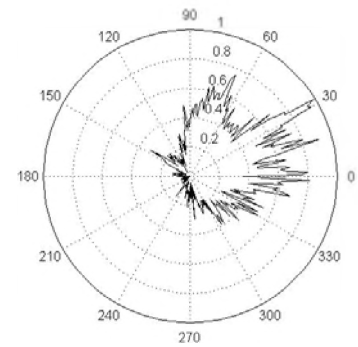

Figure 10:DTag heading for dive 1, Due North = 0, the radius axis is probability density (%)

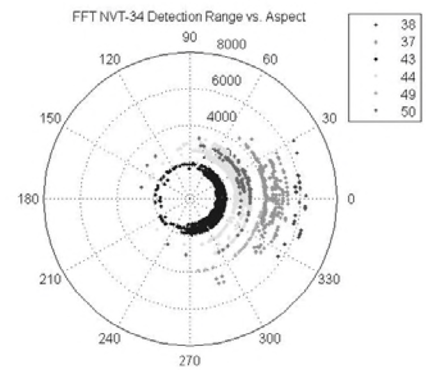

Figure 11: Detection range vs. aspect from the whale's head: FFT (NVT34) detector

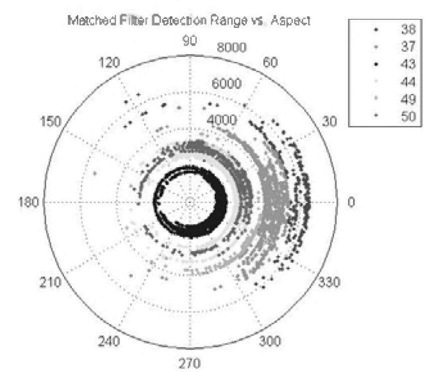

Figure 12: Detection range vs. aspect from the tagged whale's head: Matched filter detector

4. CONCLUSIONS

The matched filter detector performed significantly better than the FFT detector for *M. densirostris* foraging clicks. Using the matched filter detector, the wide-baseline AUTEC hydrophones were able to detect the tagged whale at up to 6500 m range. The off-axis aspect angle from the tagged whale to the hydrophone indicates the ability to detect signals significantly off-axis at lesser ranges and out to far ranges when close to the axis. For Dive 1, 92% of the clicks produced by the tagged whale were detected on at least one hydrophone within the array. This combination of long detection ranges and increased probability of detection with the matched filter indicates that wide-baseline, broadband arrays, such as at AUTEC, provide an excellent opportunity for long-term monitoring of beaked whale populations and successful passive acoustic based mitigation.

5. ACKNOWLEDGEMENTS

We would like to acknowledge our sponsor, Dr. Frank Stone at N45, as well as the Office of Naval Research. We would also like to thank the Bahamas Marine Mammal Research Organization who provided observer support on the water. Finally we would also like to acknowledge the NUWC Division Newport Independent Laboratory Innovative Research program manager, Richard Philips, for providing the initial funding to study the data association methods presented here on sperm whales, and the AUTEC range staff for access to their considerable infrastructure.

6. REFERENCES

1. M. P. Johnson and P. L. Tyack, "A digital acoustic recording tag for measuring the response of wild marine mammals to sound." IEEE Journal of Oceanic Engineering, 28(1), pp 3-12, 2003

2. P. T. Madsen, M. Johnson, N. Aguilar de Soto, W. M. X. Zimmer and P. Tyack, "Biosonar performance of foraging beaked whales (Mesoplodon densirostris)", Journal of Experimental Biology 208, 181-194, 2005

3. Walter M. X. Zimmer, Mark P. Johnson, Peter T. Madsen, and Peter L. Tyack, "Echolocation clicks of free-ranging Cuvier's beaked whales (Ziphius cavirostris)", J. Acoust. Soc. Am., Vol. 117, No. 6, June 2005, pp.3919-3927

4. M. Johnson, P. T. Madsen, W. M. X. Zimmer, N. Aguilar de Soto, and P. L. Tyack, "Foraging Blainville's beaked whales (Mesoplodon densitrostris) produce distinct click types matched to different phases of echolocation," The Journal of Experimental Biology (209), pp.5038-5050, 2006.

5. J. A. Ward,. Sperm whale bioacoustic characterization in the Tongue of the Ocean, Bahamas (U). NUWC-NPT TR 11,398. Naval Undersea Warfare Center, Division Newport, RI, 20 September 2002 (Unclassified).

6. D. Moretti, N. DiMarzio, R. Morrissey, J. Ward, and S. Jarvis, "Estimating the density of Blainville's beaked whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO) using passive acoustics," Oceans 2006, pp 1-5, September 2006.

7. Smith, " The Scientist and Engineer's Guide to Digital Signal Processing ", California Technical Publishing, San Diego, CA, 1997, pp. 307-310

8. R. P. Morrissey, J. Ward, N. DiMarzio, S. Jarvis, D. J. Moretti, "Passive acoustic detection and localization of sperm whales (Physeter macrocephalus) in the tongue of the ocean", Applied Acoustics 67, pp 1091–1105, 2006.

9. Vincent, H. "Models, Algorithms, and Measurements for Underwater Acoustic Positioning," Ph.D. Dissertation. University of Rhode Island, Kingston, R.I., 2001.

10. P. L. Tyack, M. P. Johnson, W. M. X. Zimmer, P. T. Madsen, M. A. de Soto. "Acoustic behavior of beaked whales, with implications for acoustic monitoring," Oceans 2006, pp 1-6, September 2006.

Photo Credit:Bahamas Marine Mammal Research Organisation