



## Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar

ELENA MCCARTHY

DAVID MORETTI

Naval Undersea Warfare Center, Code 74,

Newport, Rhode Island 02840, U.S.A.

E-mail: emccarthy@whoi.edu

LEN THOMAS

Centre for Research into

Ecological and Environmental Modeling,

The Observatory, Buchanan Gardens, University of St. Andrews,

St. Andrews, Fife KY16 9LZ, United Kingdom

NANCY DIMARZIO

RONALD MORRISSEY

Naval Undersea Warfare Center, Code 70,

Newport, Rhode Island 02840, U.S.A.

SUSAN JARVIS

JESSICA WARD

ANNAMARIA IZZI

ASHLEY DILLEY

Naval Undersea Warfare Center, Code 74,

Newport, Rhode Island 02840, U.S.A.

### ABSTRACT

The number and distribution of vocalizing groups of Blainville's beaked whales (*Mesoplodon densirostris*) were analyzed before, during, and after multiship mid-frequency active sonar operations at the US Navy's Atlantic Undersea Test and Evaluation Center (AUTECE) in the Bahamas. Groups of foraging animals were isolated by detecting their echolocation clicks using an array of bottom-mounted hydrophones. Two data sets were evaluated consisting of 115 and 240 h of acoustic data in May 2007 and 2008, respectively. Vocal activity was observed to decline during active sonar exercises and increase upon cessation of sonar transmissions in both data sets. Vocal activity did not recover to preexposure levels in the postexposure time period in 2007 nor in the initial postexposure period in the 2008 data set. Clicks detected during sonar operations were generally found to be on the periphery

of the hydrophone field and vocal durations declined for those groups that remained on the range in that time period. Receive levels were calculated for several vocal groups of whales and indicated that animals continued to forage when exposed to sonar at levels as high as 157 dB re:  $\mu\text{Pa}$ .

**Key words:** Blainville's beaked whale, mid-frequency, sonar, *Mesoplodon densirostris*, mass stranding, AUTEK.

Although of great concern, little is known about the effects of sonar and other types of anthropogenic noise on groups of foraging animals, in particular beaked whales, which have stranded during military activities. This paper quantifies behavior of foraging groups of Blainville's beaked whales (*Mesoplodon densirostris*), hereafter *Md*, by measuring the duration and location of vocalizations before, during, and after multiship navy exercises. Typically, these exercises involve several surface ships using mid-frequency active (MFA) sonars (nominally 2–10 kHz) in addition to other acoustic sources such as tracking pingers and countermeasures.

This opportunistic study is unique among beaked whale studies in that it uses existing data that are routinely recorded during military exercises. These exercises take place semi-annually at the U.S. Navy's Atlantic Undersea Test and Evaluation Center (AUTEK) in the Bahamas and provide an opportunity to study the effects of sonar operations on beaked whales exposed to actual sonar, without the significant funding, manpower, and permits required for controlled exposure experiments.

Since 2004 *Md* has been the focus of research at AUTEK as it is the species most often sighted by observers and is routinely detected acoustically. Data from passive acoustic monitoring combined with visual sightings have produced estimates of population density of *Md* on the AUTEK range between 25.3 and 30.0 per 1,000 km<sup>2</sup> (Moretti *et al.* 2006, Marques *et al.* 2009). The species has been identified in stranding events associated with MFA sonar in the past (Cox *et al.* 2006). Perhaps the most widely studied stranding event occurred in March 2000 in the Northwest Providence Channel, 110 km north of AUTEK. The use of multiple MFA sonars over an extended time in conjunction with canyon-like underwater bathymetry, excellent noise propagation conditions, and a constricted channel with limited egress was considered the most plausible cause of the stranding event (Evans and England 2001).

This study measures the response of *Md* to sonar through passive acoustic monitoring of their vocalizations combined with surface ship tracks from AUTEK. It tests the hypotheses that these animals exhibit avoidance behavior that causes them to move off the AUTEK range and return after cessation of operations and that their vocal durations decrease during sonar activity. It also determines the receive levels on foraging animals that were present on the range during active sonar operations.

#### MATERIALS AND METHODS

AUTEK is located south of the Northwest Providence Channel in a deep canyon known as the Tongue of the Ocean (TOTO) with water depths to 2,000 m. It is bounded on two sides by very shallow waters (Fig. 1). The TOTO's bathymetry, sound propagation characteristics, and animal populations are in many ways similar

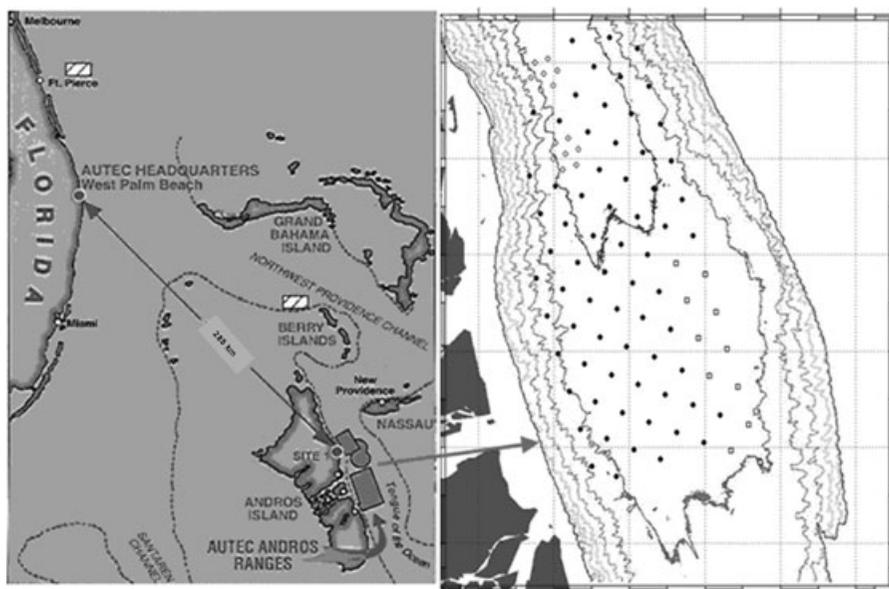


Figure 1. The undersea ranges of the AUTEC are used to collect passive acoustic recordings. This study site contains 82 hydrophones and covers  $\sim 1,500 \text{ km}^2$ . The range hydrophones are shown as dots in the image on the right.

to the Northwest Providence Channel. As such, it provides an ideal *in situ* laboratory for studying beaked whales in the presence of Navy sonar.

The AUTEC range covers an area of approximately  $1,500 \text{ km}^2$  which is instrumented with 82 bottom-mounted hydrophones that are used to provide three-dimensional tracking of undersea vehicles. Fourteen single-cabled hydrophones comprise two high-resolution arrays in the northwest corner of the range. These hydrophones, with a bandwidth of 8–50 kHz, are arranged in two hexagonal arrays with a center hydrophone, on approximately 1.8 km baselines. Sixty-eight additional hydrophones, with a bandwidth from 50 Hz to 48 kHz, comprise the remainder of the range sensors used in this study. These are arranged in offset rows on approximately 3.7 km baselines. The hydrophones are optimally spaced for direct-path tracking of underwater targets and vehicles equipped with 35 kHz pingers. This sensor layout guarantees a 5 dB signal-to-noise ratio (SNR) (as measured in a 3 kHz noise bandwidth) for tracking signals emitted by pingers with a nominal source level of 192 dB re  $\mu\text{Pa}$  at 1 m (hereafter referred to as dB). Because *Ms* produce echolocation clicks with source levels similar to those of *Ziphius cavirostris*, which are estimated at up to 214 dB, this layout assures their detection (Zimmer *et al.* 2005, Tyack *et al.* 2006b).

Blainville's beaked whales in the Bahamas associate in small groups of 1–11 animals with a mean observed group size of 4.1 (Claridge 2006). They vocalize only during deep foraging dives which can exceed 1,000 m in depth, may last for more than 60 min, and occur approximately every 2 h (Tyack *et al.* 2006a, Baird *et al.* 2008). Vocalizations generally occur below 200 m and consist of short 219–321  $\mu\text{sec}$  upsweeps from 26–51 kHz with an interclick interval (ICI) of 0.2 to 0.4 s (Johnson *et al.* 2004, 2006).

The beam pattern for Cuvier's beaked whale (*Ziphius cavirostris*), a species similar to *Md*, was estimated using acoustic data from two simultaneously tagged animals. Digital Tags (DTAG), noninvasive tags developed by the Woods Hole Oceanographic Institution (WHOI), were used to record the sound and orientation of individual whales (Johnson and Tyack 2003, Zimmer *et al.* 2005). Their echolocation clicks have a narrow beam with estimated on-axis rms source levels in excess of 200 dB. This suggests that at AUTECH, using bottom-mounted hydrophones, clicks are detected predominantly in front of the animal. Using a surface deployed sensor, Zimmer reported a maximum detection range of approximately 4,000 m (Zimmer *et al.* 2008). Recent measurements made at AUTECH using DTAGs and bottom-mounted hydrophone data from *Md* suggest a detection range of 6,500 m with few off-axis clicks detected beyond 1,200 m (Ward *et al.* 2008).

As a group forages at depth, its clicks can be detected on several hydrophones, with detections shifting back and forth between adjacent sensors, presumably as the orientation of each individual in the group changes. A digital signal processor monitors all hydrophones for transient signals including marine mammal vocalizations. A Fast Fourier Transform (FFT)-based detection algorithm is used, which implements a noise-variable adaptive threshold in each bin of the FFT. If signal energy in any of the 1,024 bins exceeds the threshold, a "detection report" is generated and archived. An FFT detection report identifies the hydrophone and precise time (<10 ms) of detection, and includes a binary frequency map indicating which bins of the FFT were above threshold. The amplitude level of the highest FFT bin is also included. If at least 10 out of 1,024 possible bins are above threshold, the detection is considered a "click" event. Clicks are classified as *Md* when the maximum signal energy is in the 24–48 kHz band and the ICI is between 0.2 and 0.4 s (Morrissey *et al.* 2006). In addition to the detection reports, the output from the hydrophones may be recorded on a bank of eight Alesis hard disk recorders, each of which records raw acoustic data from 11 range hydrophones along with IRIG-B encoded time. The performance of the FFT detector was quantified by comparing its output with *Md* data from the DTAGs (Ward *et al.* 2008).

The detection data from all hydrophones were recorded to hard disk. Transient signals including *Md* vocalizations and sonar transmissions before, during, and after active military operations on range were then viewed in the laboratory as hard-limited spectrograms (Fig. 2). Vocal groups of beaked whales were identified and the duration of their foraging activity or vocal period was determined based on the frequency, ICI, duration and location of vocalizations. Based on previous studies, it was assumed that if animals vocalized within a baseline of a hydrophone (~4 km), then at least some of their clicks would be detected during a deep foraging dive (Zimmer *et al.* 2005, DiMarzio *et al.* 2008). A "Group Vocal Period" (GVP) was considered to be a temporally and spatially unique set of vocalizations that represented a single group of beaked whales that vocalized during a deep foraging dive. The start of a GVP was considered to be the occurrence of five or more distinct *Md* clicks, known as a "click train". The end of the GVP was considered to occur upon cessation of click trains on all hydrophones in a group. These GVPs do not represent the number of distinct groups on range, but instead reflect the total number of foraging dives—some of which may be associated with the same group. No attempt was made to isolate individual animals. The start and stop times of the GVP and the location of the associated hydrophones were recorded and analyzed.

The number of GVPs before, during, and after the sonar operations were evaluated to test the following hypotheses using the Z-test:

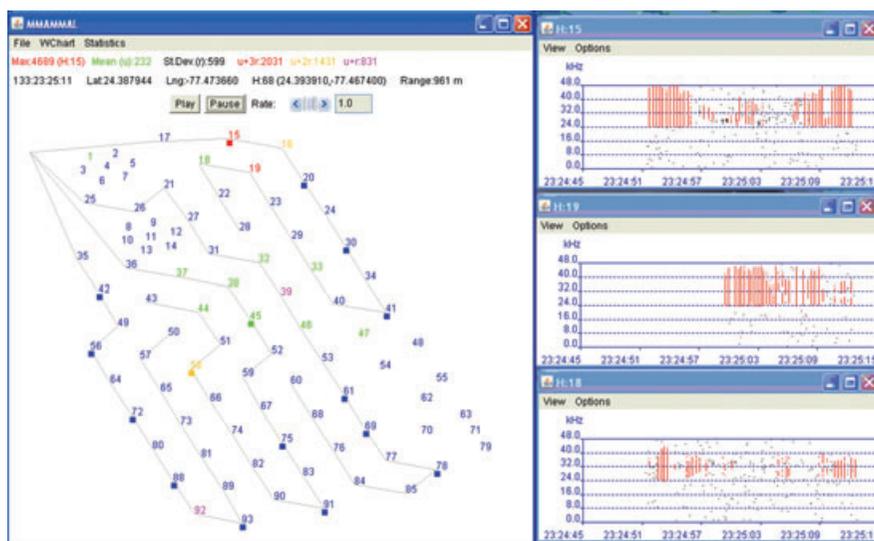


Figure 2. The AUTEc range monitor display is shown above left. The hydrophones showing the most acoustic activity above ambient noise levels are colored red. Note that hydrophones 15 and 19 show high levels of acoustic activity and that hydrophone 18 indicates a slightly lower level. On the right, three hard-limited spectrograms show characteristic beaked whale clicks from 24 to 48 kHz on hydrophones 15, 18, and 19. Thirty seconds of data are displayed on the  $x$ -axis and the frequency range from 0 to 48 kHz appears on the  $y$ -axis. This display is used for monitoring beaked whale activity at AUTEc in real-time and also in postprocessed archived data files.

H1:  $g_d/g_b < 1$ ; GVPs are lower during an exercise than before,

H2:  $g_d/g_a < 1$ ; GVPs are lower during an exercise than after,

where  $g_d$  represents the number of GVPs during the exercise,  $g_b$  represents the number before, and  $g_a$  represents the number after the exercises.

The data from both 2007 and 2008 were binned in 5 h increments and a Generalized Additive Model (GAM) was used to re-evaluate hypotheses 1 and 2 and to test hypothesis 3 as follows:

H3:  $g_a/g_b = 1$ ; GVPs recover after the exercise to their preexercise level,

where  $g_b$  represents the mean number of GVPs per 5 h period estimated from the GAM to have been on the range before the exercise and  $g_a$  represents the mean number of GVPs per 5 h period estimated from the GAM to have been on the range after the exercise.

The data were further divided into those GVPs detected within the range boundaries and those detected on edge hydrophones only and evaluated using the GAM. Edge hydrophones are those located on the edge of the range and not surrounded on all sides by hydrophones. If a vocalizing group is within the field of hydrophones, detections generally can be expected on up to seven hydrophones given the animals' beam pattern and known click detection range. If a group is outside the range boundaries, but within detection range of the edge hydrophones ( $\sim 6.5$  km), detections will

be limited to these edge hydrophones only. This led to the fourth hypothesis:

$$H4 : g_{d,i}/g_{b,i} - g_{d,e}/g_{b,e} < 0.$$

The decline in GVPs during an exercise compared with before is less on the edge hydrophones than it is on the inner hydrophones, where  $g_{b,i}$  represents the mean number of GVPs estimated by the GAM to have been on the inner hydrophones per 5 h period before the exercise,  $g_{b,e}$  represents the mean number on the edge, and  $d, a, i, e$  represent before, during, after, inner, and edge, respectively.

The GAM modeled the number of GVPs per 5 h period as a Poisson random variable, with explanatory variables time (continuous) and inner/edge hydrophone (categorical) linked to the response via a log link function. The effect of time was modeled using a thin plate regression spline with the degree of smoothness selected using generalized cross validation (Wood 2006). Modeling was done using the *mgcv* package in the statistical software R (R Core Development Team 2010). GAMs assume the response variables are conditionally independent (given the model); however the data are a time series of possibly the same diving animals, making it important to check for residual correlation. This was done using plots of temporal autocorrelation in scaled Pearson residuals, and by correlating yearly pairs of annual residuals in edge and interior hydrophones to look for spatial correlation. In addition, generalized additive mixed models (GAMMs) were fit, with an autoregressive error term of order 1 to model the temporal autocorrelation.

Predicted GVPs from the fitted model were used to construct test statistics for each of the four hypotheses. For example, for H1, the mean predicted GVP over the 5 h periods during the exercise and summed across inside and edge hydrophones was divided by the mean predicted GVP over the 5 h periods before the exercise and summed across inside and edge hydrophones. A parametric bootstrap with 10,000 replicate simulations was used to generate two-sided 95% confidence intervals about each of the test statistics, and the hypothesis was judged to have been rejected if the hypothesized value lay outside the calculated confidence intervals.

The duration of GVPs before, during, and after active sonar operations was tested using the Z-test for the following hypotheses:

$$H5 : D_d/D_b < 1; \text{ average duration of GVPs is lower during an exercise than before,}$$

$$H6 : D_d/D_a < 1; \text{ average duration of GVPs is lower during an exercise than after,}$$

where  $D$  equals average duration of GVP in minutes and where  $b, d, a$ , represent before, during, and after, respectively.

In evaluating these hypotheses, only GVPs on inner hydrophones were considered for the following reasons. First, if animals are outside of the range boundaries and detected only the edge hydrophones, the duration of their vocal period is dependent on their physical orientation. This could lead to a lower average vocal duration. Second, if animals avoid sonar and move beyond the range boundaries, this increases the number of animals detected only on edge hydrophones, which in turn could lead to a lower average vocal duration.

In order to relate animal presence to anthropogenic activity, the distributions of animals and anthropogenic sound sources on the range were mapped. In MATLAB, a triangle-based linear interpolation was used to create a visual map of animal distribution based on the number of GVPs per hour for vocalizations on "center"

hydrophones. The “center” hydrophone was considered the hydrophone that recorded the most vocal activity during a GVP. To generate a map of the anthropogenic sound field during the exercises, the occurrences of anthropogenic noise were logged along with parameters, such as source type, frequency range, location, and duration. Both the 2007 and 2008 data sets consisted of multiship sonar exercises which included three surface ships using active sonar. Two of the surface ships used AN/SQS-56 or similar type sonars with a frequency range of 4.5–8.2 kHz and a published rms source level of 223 dB. One ship used an AN/SQS-53C sonar with a frequency range of 2.6–3.5 kHz and a published rms source level of 235 dB (D’Spain *et al.* 2006). The precise location of each ship was determined by AUTECH tracking systems to within 10 m and positions were logged every second. In general, ship movements were restricted to the area within the range boundary.

To illustrate the distribution of sound during these exercises, the cumulative energy levels over the range produced during the 68 h of the 2008 test were calculated using the Navy Exposure Model (NEMO), presently under development at the Naval Undersea Warfare Center. NEMO is used to estimate the sound exposure levels on marine mammals during Navy exercises. Inputs to the model include the ship tracks from surface vessels, active sonar systems characteristics (source level, beam pattern, depth), and the transmission times and duration of the active sonar. The cumulative energy over the entire exercise was calculated in 100 m intervals. Propagation loss was evaluated using the CASS/GRAB propagation model (Weinberg and Keenan 1996) at nine locations on the range over eight radials per location. NEMO predicted the sound exposure levels by updating the three ships’ positions every second and calculating the range and true bearing from the sonar to each range cell. The CASS/GRAB result at the closest location and radial was used to calculate the receive level at each cell.

In addition to looking at how GVPs were distributed, the pattern in which animals returned to the range was also considered. For the 2007 data set, hydrophone locations were plotted and color-coded according to the presence of *Md* detections in 6 h increments after the cessation of sonar, and the resultant pattern was examined.

The sonar receive levels for confirmed cases of foraging groups of *Md* were calculated. This required knowledge of the location of the animals and the sources, in addition to an understanding of the source characteristics and environmental parameters. The position of the center hydrophone was used as an approximation of the location of the vocalizing group. Positions of the surface ships and their hull-mounted sonars were then related to the location of vocalizing animals. The position, speed, and heading of each ship were precisely known from AUTECH track data and the durations of sonar transmissions were obtained from ships’ logs and hydrophone recordings. Finally, the relevant environmental parameters (bathymetry, sound velocity) were extracted from a US Navy database. Propagation loss was then calculated from these parameters using the CASS/GRAB algorithms. The receive level at the animal could then be determined based on the modeled propagation loss curves.

## RESULTS

The data sets from 2007 and 2008 consisted of archived data that contained detections of transient signals from all hydrophones on the range. The 2007 set consisted of detection reports from 115 h of data: 17 h before the start of the sonar transmissions, 75 h during the active sonar operations, and 23 h after the end of the exercise. The 2008 data set was considerably larger and consisted of detection reports

from 240 h of data: 65 h before the start of sonar, 68 h during the sonar exercises, and 108 h after the end of the exercise.

In addition to whale vocalizations, the 2007 data set contained 34.9 h of sonar transmissions during the 75 h of operations. The average sonar transmission was 11.78 min in duration with a standard deviation of 9.29 min. The 2008 data set contained 41.45 h of sonar during the 68 h of the operation with an average sonar transmission of 60.6 min and a standard deviation of 49.4 min.

Additional sources of sonar included the occasional use of dipping sonars deployed from hovering helicopters, which operate in a frequency range from 1.2 to 5.6 kHz, and active sonobuoys, which operate at selectable frequencies of 6.5, 7.5, 8.5, and 9.5 kHz at a source level less than 200 dB (Funnell 2009). Other sources of anthropogenic noise were also present during these exercises. These included countermeasures, which emit broadband noise designed to mask vessels' acoustic signatures and confuse sonars; pingers, which transmit a short (5–50 ms) 13 or 37 kHz signal used to track undersea vehicles; and acoustic communication signals in the 8–15 kHz bandwidth at source levels below 195 dB (Funnell 2009). Ships' propulsion and flow noise are other sources of anthropogenic sound with low-frequency components below 1.5 kHz and higher frequency (>10 kHz) broadband transients (Wenz 1962). Of all the anthropogenic sound sources present during the exercises, the surface ship sonars made up the majority of transmissions and had the highest source level. Countermeasures, the second most common source of anthropogenic noise throughout the range, were present less than half as often as the sonars.

#### *Number of Group Vocal Periods*

The 2007 data were divided into three periods, before (17 h), during (75 h), and after (23 h) sonar operations. The 2008 data were divided into four periods, before (65 h), during (68 h), immediately postexercise (65 h), and remainder (43 h) postsonar operations. The duration of the initial postexposure phase was chosen to match that of the preexposure phase (65 h). The data were initially tested using a simple *Z*-test for hypotheses 1 and 2. The basic statistics for GVPs detected during both exercises are presented in Table 1.

In both years a significantly greater number of GVPs occurred before the sonar than during the sonar (for 2007:  $|Z| = 39.1$ ,  $P < 0.00$ ; for 2008:  $|Z| = 69.6$ ,  $P < 0.00$ ). In 2007 the number of vocal periods during the sonar was significantly less than the number after the sonar ( $|Z| = 4.4$ ,  $P < 0.00$ ).

In 2008 the number of vocal periods during sonar operations was not significantly lower than in the 65 h immediately after the sonar ( $|Z| = 0.24$ ,  $P = 0.59$ ). However, the average number of vocal periods in the 65–108 h after the sonar ended was significantly greater than that before, during, and immediately after exposure (before:  $|Z| = 4.69$ ,  $P < 0.00$ ; during:  $|Z| = 22.2$ ,  $P < 0.00$ ; immediately after:  $|Z| = 9.48$ ,  $P < 0.00$ ).

The data from both 2007 and 2008 were also binned in 5 h increments (Fig. 3) and a Generalized Additive Model (GAM) was fitted to the data and used to test H3 and retest H1 and H2. Additionally, the data were further divided into those GVPs detected within the range boundaries and those detected on edge hydrophones only. This allowed for the evaluation of H4.

Plots of the GAM fit to the data are given in Figure 4. The data were separated into those GVPs detected on inner hydrophones and those detected only on edge hydrophones. For both years, the model provided a good fit to the observed data;

Table 1. Group vocalization statistics for 2007 and 2008 data.

	Group vocal periods	
	2007	2008
Before active sonar operations		
Measurement duration (h)	17	65
Number of vocal periods recorded	63	263
Number of vocal periods/h	3.7, SD = 1.82	4.04, SD = 1.87
Mean inner phone vocal duration (min)	33.92	29.65
Number of edge phone detections	22	94
% edge phone detections	35	36
Number edge phone detections/h	1.29	1.45
During active sonar operations		
Measurement duration (h)	75	68.13
Number of vocal periods recorded	82	93
Number of vocal periods/h	1.09, SD = 1.1	1.36, SD = 1.53
Mean inner phone vocal duration (min)	25.08	25.5
Number of edge phone detections	56	45
% edge phone detections	68	48
Number edge phone detections/h	0.75	0.66
After active sonar operations		
Measurement duration (h)	23	108.25
Number of vocal periods recorded	51	362
Number of vocal periods/h	2.22, SD = 1.6	3.34, SD = 2.91; entire postexposure period 1.49, SD = 1.73; 0–65 h after sonar 6.16, SD = 2.9; 65–108 h after sonar
Mean inner phone vocal duration (min)	32.92	30.63
Number of edge phone detections	13	88
% edge phone detections	25	24
Number edge phone detections/h	0.56	0.81

plots of residual autocorrelation (Fig. 5) showed no sign of temporal correlation in residuals. Estimated values from the GAMM (not shown) were almost identical, and the estimated temporal autocorrelation was low: 0.044 for 2007 and 0.179 for 2008. There was also little evidence of spatial correlation in the GAM residuals: correlation between yearly pairs of edge and inner residuals was 0.056 (95% CI: 0.374–0.467) in 2007 and 0.210 (95% CI : 0.085–0.472) in 2008. Given the lack of significant residual correlation, hypothesis test statistics are presented only for the GAM.

Test statistics addressing the four hypotheses and calculated from the GAM model are given in Table 2, together with the bootstrap confidence intervals. Results strongly support hypotheses 1 and 2 (that GVPs are lower during an exercise than before, and are lower during an exercise than after). As an example, the test statistic for H1 was the ratio of the number of GVPs during the exercise to the number before, and this was hypothesized to be less than 1. The estimated value for the 2007 experiment was 0.31, with 95% CI 0.29–0.52, indicating strong support for the hypothesis.

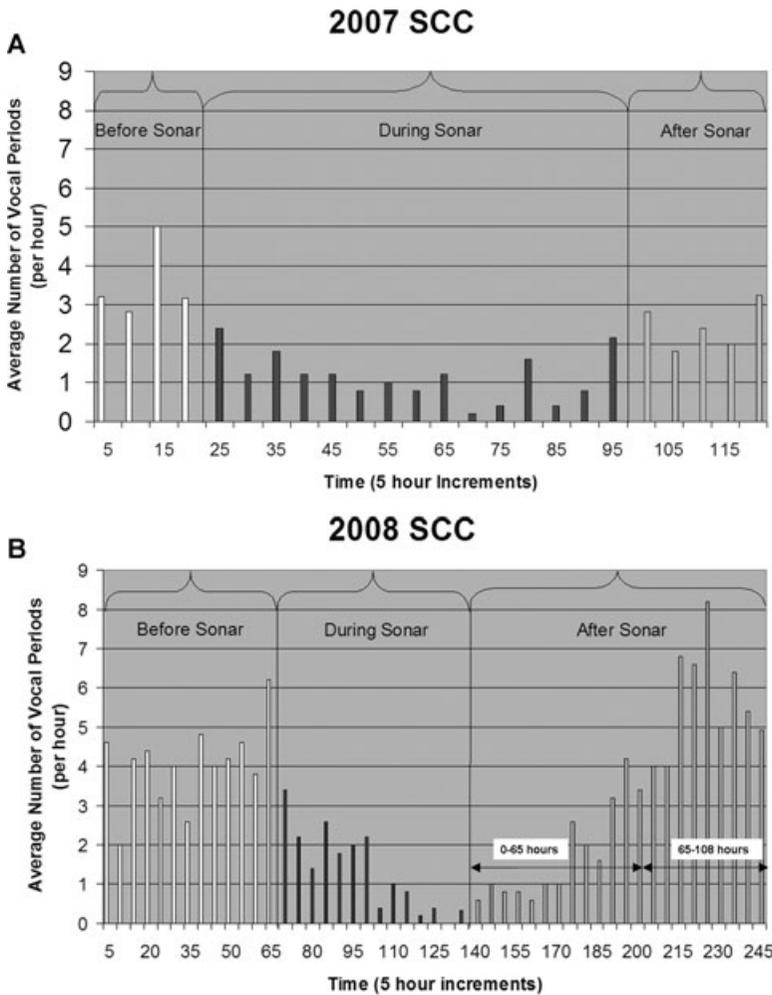


Figure 3. Average number of vocal periods per hour in 5 h intervals before, during, and after sonar operations for 2007(A) and 2008 (B) exercises.

Results did not support hypothesis 3, that the number of GVPs recovered after the exercise to its preexercise level. The ratio of the number of GVPs after sonar exercises to the number before was 0.61 (95% CI: 0.39–0.90) in 2007 and 0.78 (95% CI: 0.65–0.94) in 2008. Inspection of the GVP data and fitted trend (Fig. 4) show that in 2008 the GVPs increased slightly after the exercise, but remained low for approximately 35 h before climbing back to preexercise levels, even exceeding them on inner hydrophones at 75 h postexercise. If the test statistic is recalculated, using only data starting 35 h postexercise for the “after” component, then the ratio of the number of pre- to postexercise GVPs becomes 1.06 with 95% CI 0.88–1.28. The same calculations could not be repeated for the 2007 data as only 23 h of postexercise data were available.

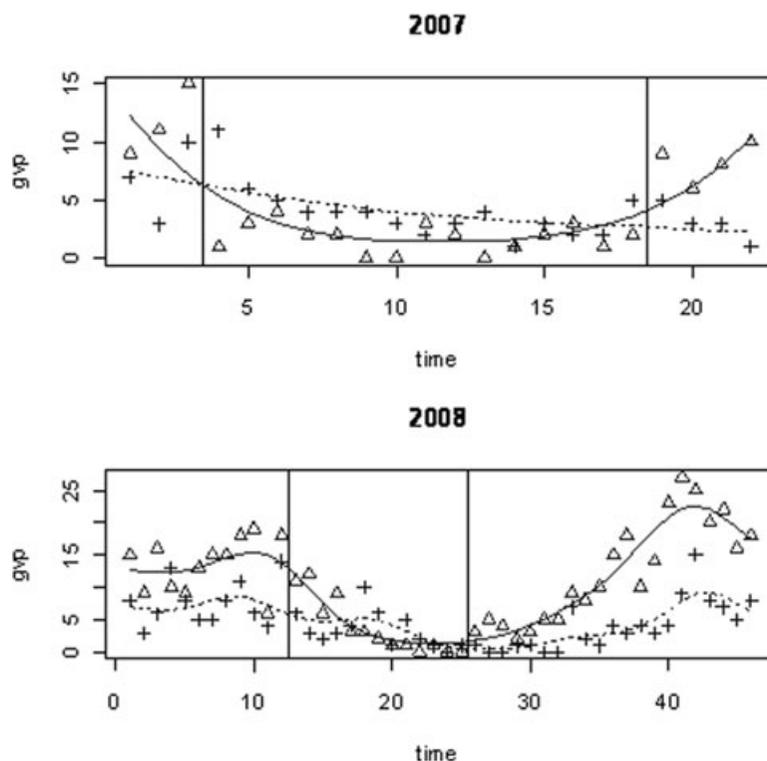


Figure 4. The results of the GAM (solid line) fit to the GVPs within the range boundaries (diamond), and the GAM (dashed line) fit to GVPs on edge hydrophones (plus). The x-axis represents the number of 5 h time increments. The plots are divided by vertical lines into the time periods before, during and after sonar operations.

Hypothesis 4, that the decrease in GVPs was stronger on interior hydrophones than edge hydrophones, was strongly supported in the 2007 data, but marginally nonsignificant in the 2008 data.

Results from the GAMM (not shown) were very similar to those from the GAM, and the estimated first-order temporal autocorrelation was very low: 0.04 in 2007 and 0.18 in 2008.

#### *Distribution of Group Vocal Periods*

The distribution of animals for the 2007 and 2008 data sets are illustrated by the surface plots given in Figures 6 and 7, respectively. The plot on the far left of each figure shows the average number of vocal periods per hour detected on each hydrophone before the sonar exercise. These detections represent foraging clicks from groups of animals and detections are found on almost every hydrophone across the range. The second plot presents the average number of vocal periods detected on each hydrophone during the active sonar exercises during which the majority of ship activity occurred in the center of the range. Beaked whale detections during these

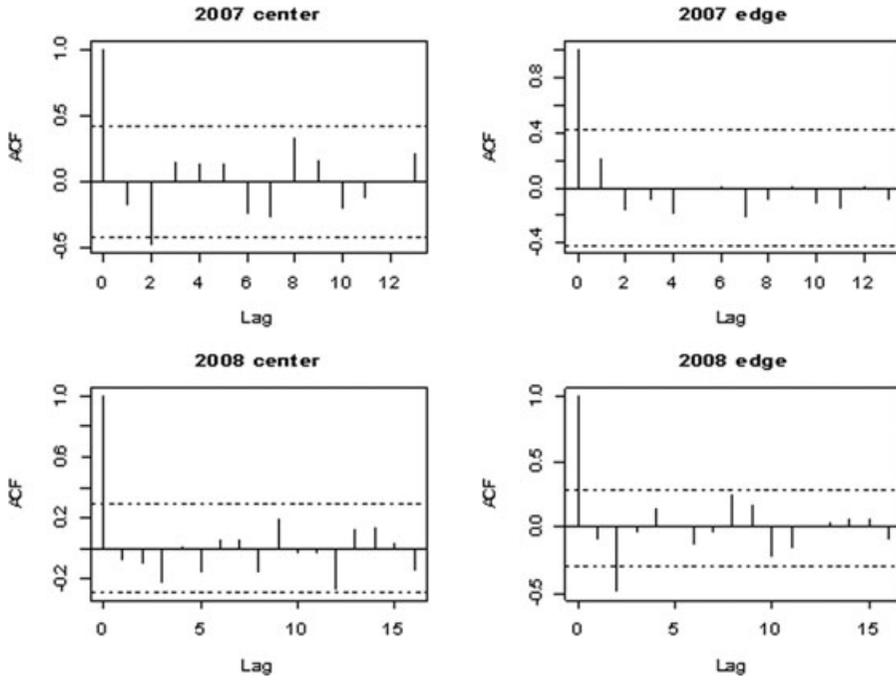


Figure 5. Temporal autocorrelation plots based on residuals from GAM model for 2007 and 2008 inner and edge GVPs. Blue dashed lines indicate critical values above which an autorrelation would be considered significantly different from zero assuming a type I error rate ( $\alpha$ -level) of 0.05.

exercises were distributed on the periphery of the range with no detections in the center where the military operations were concentrated (Fig. 8).

The pattern in which *Md* groups returned to the range was also considered. This is illustrated in Figure 9, which plots detections in 6 h intervals after the cessation of active exercises. In the first 2 h after the end of sonar operations, clicks were detected on hydrophones in the southwest corner of the range. In the next 3 h period, clicks were detected adjacent to these hydrophones but further north into the range and also on the far northwest edge hydrophones. As time increased, click detections were observed on hydrophones successively further into the center of the range.

Table 2. Test statistics, with 96%CI in brackets for each hypothesis. *g* denotes group vocal period and the subscripts *b, d, a, i, e* denote before, during, after, inner, and edge, respectively.

Hypothesis	2007	2008
H1: $g_d/g_b < 1$	0.31 (0.29–0.52)	0.34 (0.29–0.47)
H2: $g_d/g_a < 1$	0.50 (0.46–0.92)	0.43 (0.27–0.61)
H3: $g_a/g_b = 1$	0.61 (0.39–0.90)	0.78 (0.65–0.94)
H4: $g_{d,i}/g_{b,i} - g_{d,e}/g_{b,e} < 0$	-0.56 (-0.59–0.09)	-0.56 (-0.45–0.00)

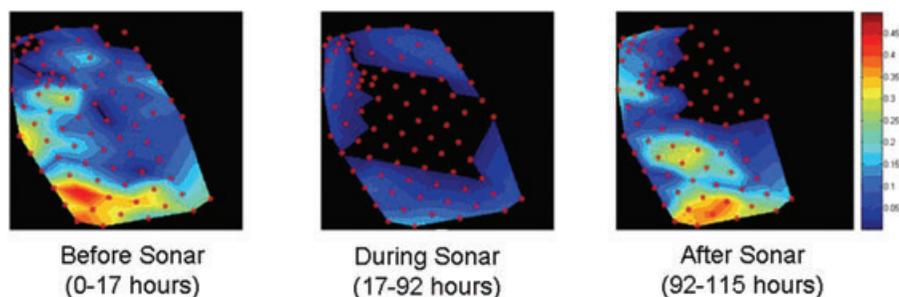


Figure 6. AUTECH hydrophone map showing average number of *Md* vocalizations per hour on each hydrophone during the 2007 exercise. Hydrophones are shown as red circles. The color bar indicates the average number of vocal periods per hour detected on each hydrophone. These images were created by plotting the number of times a hydrophone was at the center of a vocalizing group and using a triangle-based linear interpolation in MATLAB.

#### Group Vocal Duration

A decrease in average duration of vocalizations was also measured during sonar operations, supporting hypotheses 5 and 6. The average duration of vocalizations on inner hydrophones dropped from 33.92 min to 25.08 during the exercise in 2007 and then increased to 32.92 min after the exercise. In 2008 the average duration of vocalizations on inner hydrophones decreased from 29.65 min prior to the exercise to 25.5 min during the sonar transmissions and then rose to 30.63 after the exercise ended (for 2007:  $|Z| = 6.02$ ,  $P < 0.00$ ; for 2008:  $|Z| = 3.40$ ,  $P < 0.00$ ).

#### Receive Levels at Animals

Precise knowledge of ship position and the characteristics and transmission times of sonar allowed for a calculation of the propagation loss and a determination of receive level on vocalizing animals. These receive levels are presented in Table 3. For animals that continued to forage during exposure, the estimated received levels

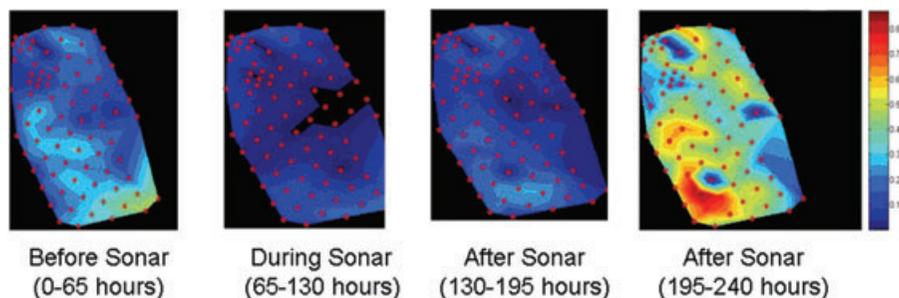


Figure 7. AUTECH hydrophone map showing average number of *Md* vocalizations per hour during the 2008 exercise. Note that the “before” and “after” portions of the 2008 data set (consisting of 65 and 110 h, respectively) were much longer than in the 2007 data set. This accounts for the fourth image above. Also note that the range of the color bar for the 2008 surface plots differs from the color bar used for the 2007 data set. This allows for a better visualization of the high vocal activity displayed in the 65–108 h after exposure.

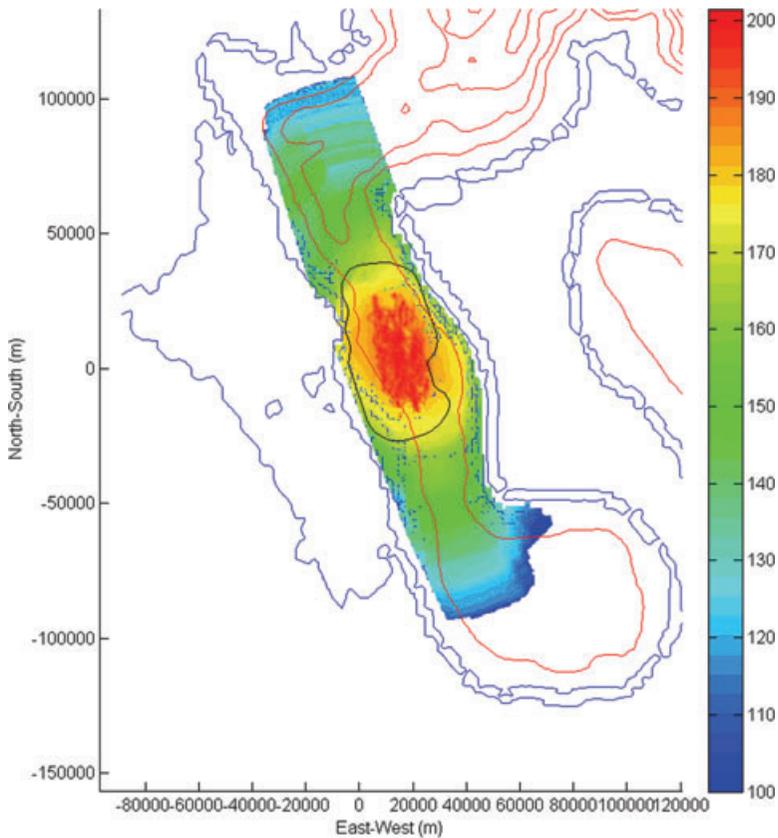
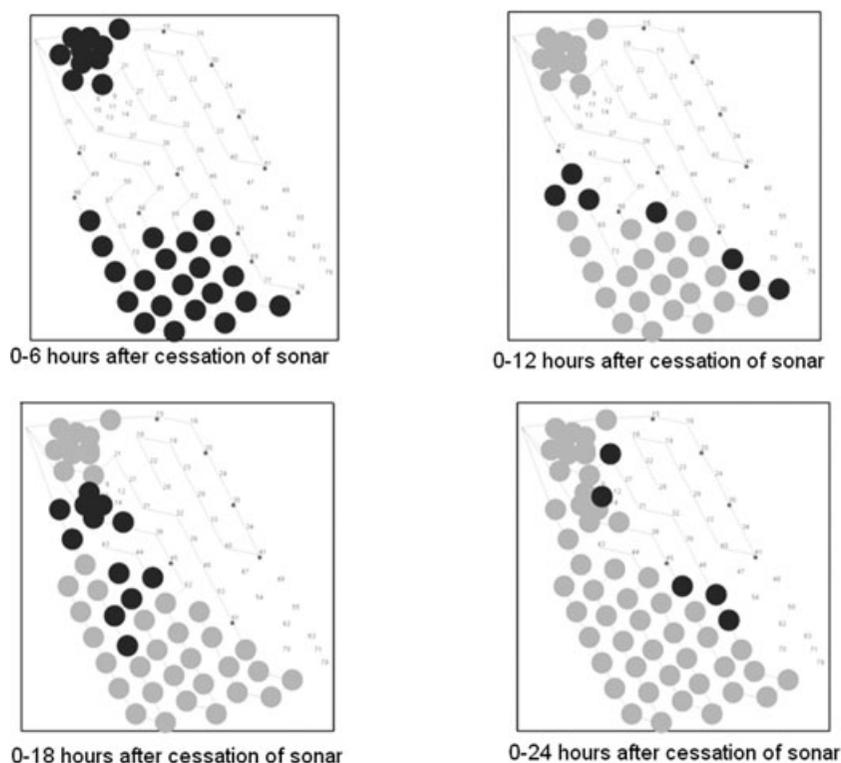


Figure 8. A bird's eye view of the cumulative sound exposure level at 100 m depth over the AUTECH range during sonar operations for the 2008 exercise. The color bar indicates rms level in dB re  $1 \mu\text{Pa}^2 \text{ s}$ . The boundary of the AUTECH range is outlined in black in the middle of the image.

ranged from 101 to 157 dB for the 17 vocal periods analyzed, with an average receive level of 128 dB and a standard deviation of 15 dB. In one instance, a group completed an average vocal period of 37 min while exposed to transmissions from two sonars (53C and 56) at receive levels of 128 db and 118 db, respectively.

## DISCUSSION

The decline in vocalization activity associated with foraging groups of *Md* during military exercises has three possible explanations: (1) the animals moved off the range but continued to vocalize, (2) the animals did not vocalize during the military operations, or (3) the system failed to detect whale vocalizations in the midst of noise associated with military operations (*i.e.*, masking occurred). The results of the analysis strongly suggest that the animals avoided ships using active sonar and moved



*Figure 9.* Distribution of vocalizations in the 24 h immediately after sonar transmissions during the 2007 exercise. The darker circles indicate hydrophones on which a vocalization was detected for the first time in a 6 h period after the sonar exercises stopped. The lighter-colored circles indicate those hydrophones where detections had occurred in previous 6 h periods. The initial vocalizations were not distributed randomly throughout the range, but rather were detected on the periphery of the range and gradually moved to the center.

off range during such exercises. Further, the data suggest animals return to the range after the cessation of sonar activity.

#### *Number of Group Vocal Periods*

Analysis of the data using the Z-test indicates a significant decline in the number of GVPs per hour during sonar operations. This decline was also observed when the data were binned in 5 h increments and analyzed using the GAM. In the 2007 data set the number of GVPs did not recover to preexposure levels within the 23 h observation period after sonar ended; in 2008 it did not recover until around 35 h postexercise, but then increased to a level even higher than that measured during the period before sonar operations, particularly on the inner hydrophones. This suggests that the whales may have taken 1–3 d to recover from exposure. It is unclear whether this can be attributed to an identical or lower number of animals exhibiting a higher foraging rate (suggesting hunger) or an increased number of animals exhibiting a

Table 3. Summary of receive levels on *Md* during military operations (\* indicates same vocal group).

Source type	Frequency (kHz)	Range (m)	Bearing (degrees)	Known exposure duration (min)	Receive level (dB re: 1 $\mu$ Pa) (rms)	Duration of vocal period (min)
53C	3.5	18,967	327	4	127	26
56	8	21,391	7	8	114	8
53C	3.5	18,193	263	19	133	19
56	8	25,216	352	28	112	28
53C	3.5	14,727	215	29	133	29
56	8	21,101	177	13	118	31
56	8	12,212	228	10	117	26
Foreign	4.5	24,311	338	36	117	36
Foreign	4.5	3,416	259	26	154	26
Foreign	4.5	2,161	44	17	157	20
56	8	7,011	135	22	143	30
56	8	12,586	229	20	126	20
56	8	11,354	159	8	129	8
56	8	7,612	291	7	141	35
56	8	28,853	348	12	101	21
53C*	3.5	18,117	335	37	128	37
56*	8	25,789	327	37	118	37

lower foraging rate (suggesting an increased prey abundance after the sonar exercises due to a lack of predators on the range during sonar exercises).

#### *Distribution of Group Vocal Periods*

The application of the GAM considered both the number of GVP's and their distribution as related to edge/inner range hydrophones. This analysis found that there was a significant increase in the ratio of GVPs detected on edge hydrophones as compared to inner hydrophones during sonar operations in 2007, and an almost significant increase in 2008. Note that, although the ratio of detections on edge to inner hydrophones increased during sonar, the rate of GVPs per hour detected on edge hydrophones actually declined. This suggests that if animals continue to vocalize but move off-range, then a proportion of the groups is more than 6.5 km beyond the range boundaries, based on the maximum click detection range.

The distribution of animals shown in the surface plots of Figures 6 and 7 suggest that either the animals were distributed throughout the range but did not vocalize during the sonar operations, or that the animals moved off the range. If whales were present on the range but not vocalizing, the pattern of vocalizations upon cessation of sonar activity would reflect a distribution of animals throughout the entire range. Instead as observed in Figure 8, vocalizations were initially detected from the sides of the range, and gradually spread to the center and north. This "return" rate is consistent with rates of horizontal movement from tagged animals and measured shallow water vertical ascent and descent rates (Tyack *et al.* 2006a, Schorr *et al.* 2009). A similar pattern of vocalization distribution throughout the range was observed in the 2008 data. This suggests that the animals that were not detected during sonar exercises had moved off the range and returned upon completion of the exercises.

### *Group Vocal Duration*

The measured decrease in average GVP duration could be attributed to an increased number of detections on edge hydrophones as the animals moved off range. However, when the duration of vocalizations *not* on edge hydrophones during sonar is compared to average vocal durations before and after sonar, a decrease in duration is still observed. This could indicate that the animals that remained on range decreased the duration of their foraging dives.

### *Receive Level at Animals*

Although generally animals appear to move off range during active sonar operations, 17 cases of *Md* groups vocalizing coincident with active sonar were isolated. This suggests that these animals continued to forage despite the presence of sonar. The maximum estimated receive level was 157 dB with an average level of 128 dB. This provides insight into the levels at which *Md* react to sonar but did not consider the context, such as the ship movement relative to the vocalizing group.

### *Masking*

The possibility that the increased noise background created by sonar operations masked the detection of *Md* echolocation clicks was considered. Given the known source level of *Md*, their focused beam pattern, their average vocal period, and the high number of clicks produced by a foraging group of animals, it is reasonable to assume that at least *some* of the clicks during foraging dives would be detected on at least one of the range hydrophones. If the group is foraging within the range, the maximum distance from the closest hydrophone is approximately 2.0 km. If the on-axis source level is conservatively estimated as 200 dB, the SNR at the hydrophone will be in excess of 30 dB at sea state 4. In addition, the center frequencies of the 53C and the 56 sonars are below 10 kHz, while ship noise is dominant below 2 kHz (Wenz 1962). The majority of the energy in *Md* clicks is concentrated in the band from 25 to 60 kHz which provides significant separation from such sources. Moreover, clicks were detected even in the presence of operational noise (*i.e.*, low SNR). In these cases, animal vocalizations were observed coincident with sonar transmissions, as shown in Figure 10. Finally, if masking had occurred, the number of GVPs detected immediately after the cessation of operations would increase. In fact, in the first three postexercise days of the 2008 data set, the number of vocal periods per hour remained significantly lower than in the 3 d before the operations even though no sources were active. These factors strongly suggest that vocalizations are not masked by sounds associated with these sonar operations.

### *Future Research*

These data strongly suggest an avoidance response by *Md* during active sonar exercises. However, our current measurements are restricted to vocalizing animals within detection range of the hydrophones. Once outside AUTECH range boundaries, the nature and extent of their behavior is unknown. The use of long-term tags (>30 d) will provide additional information as to the range of beaked whale movement. Such tags, which are attached with a dorsal fin dart, have been used successfully on beaked whales in the Hawaiian Islands (Baird *et al.* 2006). Tagging efforts in collaboration

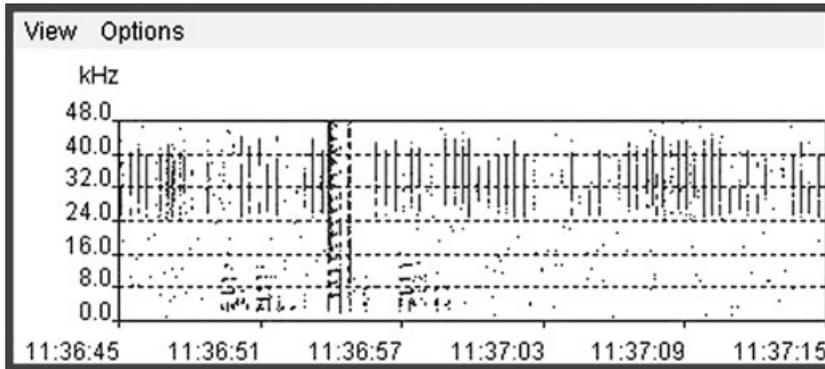


Figure 10. Output from hydrophone 49 showing *Md* clicks at 24–40 kHz concurrent with mid-frequency sonar transmissions shown at time 1136:54.

with the Bahamas Marine Mammal Research Organisation (BMMRO) and NOAA Southwest Fisheries are underway at AUTEK where researchers are also collecting photo-ID data that can be used to document calving rates and animal presence in the TOTO and Northwest Providence Channel. These data will provide information on large-scale movement of populations and could help determine whether *Md* move between the TOTO and the Northwest Providence Channel. These ongoing efforts will also help characterize whales' behavior on the AUTEK range and whether it is relevant to previous stranding events. This might aid in determining if individual *Md* move between AUTEK and the Northwest Providence Channel and whether animals near the AUTEK range are habituated to sonars and as such less likely to exhibit behaviors that lead to stranding.

Even if behavior that leads to stranding is not found, it is important to determine if military operations may impair the health of *Md* in other ways, *i.e.*, if military operations are affecting their ability to survive and reproduce at a rate to maintain their existing population. To this end, the ongoing Behavioral Response Study (BRS) will provide detailed data concerning the behavior of individual animals in response to sonar (Southall *et al.* 2008). Another study, by Duke University, will create predator–prey maps that will help define the extent of prey fields (Nowacek *et al.* 2009). Data from these two studies can then be combined with passive acoustic data from AUTEK to link foraging animals directly to their prey. This may help determine if anthropogenic noise is affecting the prey or the predator—that is, whether the sonar is driving the food source, not the whales, off the range. Moreover, the use of predictive habitat models will begin to link environmental data to animals in the TOTO. By combining these data in a population model, it may be possible to assess *Md* population health.

It should be emphasized that during these operations, sources of anthropogenic noise other than sonar were present. Although these sources operated at lower levels and at shorter durations, one future goal is to differentiate among various anthropogenic sources of noise and their potential for effects. A more complex time series analysis could focus on the significance of amount of time before the whales begin to vocalize after transmissions of various sources. Another goal is to determine what effects various source levels, frequencies, and signal type may have. It is also essential to determine what variations exist in beaked whale foraging patterns on a seasonal,

monthly, and daily basis. These baseline data will also aid in determining what “recovery” times can be expected upon the cessation of sonar operations. Without these baseline data, it will be impossible to determine what normal variations in beaked whale behavior might be.

The current focus of marine mammal mitigation policy has been driven by concern over the effects of anthropogenic sound on individual animals. The use of passive acoustics combined with surface observations has documented *Md* in significant numbers in an area known for repeated use of active sonar. For the first time, these opportunistic data show a reaction on a group level. This gives rise to complex questions. For example, what does the absence of vocalizations mean for *Md* in terms of their health and survival? How do these active sonar exercises differ from those that resulted in mass strandings of animals? Do these animals become habituated to the anthropogenic sounds to which they are exposed? Does displacement actually affect feeding? Most importantly, is this effect *biologically significant* and what are the effects of sonar on a *population* level? By expanding the analysis of these opportunistic data, we hope to address these questions in the future.

#### ACKNOWLEDGMENTS

We thank Peter Tyack and Mark Johnson of the Woods Hole Oceanographic Institution for their support and expertise with DTAGs, Diane Claridge and her staff at the Bahamas Marine Mammal Research Organization, Ian Boyd at the Sea Mammal Research Unit at the University of St Andrews, and Chris Clark at Cornell University's Laboratory of Ornithology. Thanks also to Bert Neales, Jr. at the Naval Undersea Warfare Center for his expertise with NEMO, Bill Sutphin for his knowledge of AUTECS's ship tracking system, and the interns there who aided in the creation of this manuscript, Margaret Prior and Carroll-Anne Ciminello. Special thanks are extended to Holly Turton who edited our manuscript. We would especially like to acknowledge the very helpful and thorough comments by three anonymous reviewers. The Chief of Naval Operations' Environmental Readiness Division and the Office of Naval Research's Marine Mammals and Biological Oceanography Program provided funding and guidance for this work.

#### LITERATURE CITED

- Baird, R., D. Webster, D. McSweeney, A. Ligon, G. Schorr and J. Barlow. 2006. Diving behaviour of Cuvier's (*Ziphius cavirostris*) and Blainville's (*Mesoplodon densirostris*) beaked whales in Hawai'i. *Canadian Journal of Zoology* 84:1120–1128.
- Baird, R., D. Webster, G. Schorr, D. McSweeney and J. Barlow. 2008. Diel variation in beaked whale diving behavior. *Marine Mammal Science* 24:630–642.
- Claridge, D. E. 2006. Fine scale distribution and habitat selection of beaked whales. M.Sc. thesis, University of Aberdeen, King's College, Scotland, U.K. 119 pp.
- Cox, T. M., T. J. Ragen, A. J. Read, *et al.* 2006. Understanding the impacts of anthropogenic sound on beaked whales. *Journal of Cetacean Research and Management* 7:177–187.
- D'Spain, G. L., A. D'Amico and D. M. Fromm. 2006. Properties of the underwater sound fields during some well documented beaked whale mass stranding events. *Journal of Cetacean Research and Management* 7:223–238.
- DiMarzio, N., D. Moretti, J. Ward, R. Morrissey, S. Jarvis and A. Izzi. 2008. Passive acoustic measurement of dive vocal behavior and group size of Blainville's beaked whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO). *Canadian Acoustics* 36:166–173.

- Evans, D. L. and G. R. England. 2001. Joint interim report Bahamas marine mammal stranding event of 15-16 March 2000. (unpublished). 61 pp. Available at [http://www.nmfs.noaa.gov/pr/pdfs/health/stranding\\_bahamas2000.pdf](http://www.nmfs.noaa.gov/pr/pdfs/health/stranding_bahamas2000.pdf) (accessed 17 November 2010).
- Funnell, C., ed. 2009. Jane's underwater warfare systems. 21st edition. Jane's Information Group, Coulsdon, Surrey, U.K.
- Johnson, M. P., and T. L. Tyack. 2003. A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Ocean Engineering* 28:3–12.
- Johnson, M., P. Madsen, W. M. X. Zimmer, N. Aguilar De Soto and P. L. Tyack. 2004. Beaked whales echolocate on prey. *Proceedings of the Royal Society of London, B* 271 (Suppl 6):S383–S386.
- Johnson, M., P. T. Madsen, W. M. X. Zimmer, N. Aguilar De Soto and P. L. Tyack. 2006. Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation. *Journal of Experimental Biology* 209:5038–5050.
- Marques, T. A., L. Thomas, J. Ward, N. DiMarzio and P. L. Tyack. 2009. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville's beaked whales. *Journal of the Acoustical Society of America* 125:1982–1994.
- Moretti, D., N. DiMarzio, R. Morrissey, J. Ward and S. Jarvis. 2006. Estimating the density of Blainville's beaked whale (*Mesoplodon densirostris*) in the Tongue of the Ocean (TOTO) using passive acoustics. *Proceedings of the MTS/IEEE Oceans 2006 Conference and Exhibition*. 5 pp.
- Morrissey, R., J. Ward, N. DiMarzio, S. Jarvis and D. J. Moretti. 2006. Passive acoustic detection of sperm whales (*Physeter macrocephalus*) in the tongue of the ocean. *Applied Acoustics* 67:1091–1105.
- Nowacek, D. P., L. St. Laurent, D. J. Moretti and P. N. Halpin. 2009. Prey fields and habitats of deep diving odontocetes: 3d characterization and modeling of beaked and sperm whale foraging areas in the Tongue of the Ocean. N00014-08-1-1162, 2008/09-2010/08, Office of Naval Research, Arlington, VA.
- R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Schorr, G., R. Baird, M. B. Hanson, D. Webster, D. McSweeney and R. Andrews. 2009. Movements of satellite-tagged Blainville's beaked whales off the island of Hawai'i. *Endangered Species Research* 10:203–213.
- Southall, B., I. Boyd, P. Tyack and D. Wartzok. 2008. Deep-diving odontocetes behavioral-response study (BRS). *Bioacoustics* 17(1–3):186–188.
- Tyack, P., M. Johnson, N. Aguilar Soto, A. Sturlese and P. T. Madsen. 2006a. Extreme diving of beaked whales. *Journal of Experimental Biology* 209:4238–4253.
- Tyack, P., M. Johnson, W. Zimmer and P. Madsen. 2006b. Acoustic behavior of beaked whales with implications for acoustic monitoring. *Oceans 2006*:1–6.
- Ward, J., R. Morrissey, D. J. Moretti, et al. 2008. Passive acoustics detection and localization of *Mesoplodon densirostris* (Blainville's beaked whale) vocalizations using distributed bottom-mounted hydrophones in conjunction with a digital tag (DTAG) recording. *Canadian Journal of Acoustics* 36:157–166.
- Weinberg, H., and R. Keenan. 1996. Gaussian ray bundles for modeling high-frequency propagation loss under shallow-water conditions. *Journal of Acoustical Society of America* 100:1421–1431.
- Wenz, G. M. 1962. Acoustic ambient noise in the ocean: Spectra and sources. *Journal of Acoustical Society of America* 34:1936–1956.
- Wood, S. N. 2006. Generalized additive models: An introduction with R. CRC Press, Boca Raton, FL.

- Zimmer, W. M. X., M. P. Johnson, P. T. Madsen and P. L. Tyack. 2005. Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). *Journal of the Acoustical Society of America* 117:3919–3927.
- Zimmer, W. M. X., J. Hardwood, M. P. Johnson, P. T. Madsen and P. L. Tyack. 2008. Passive acoustic detection of deep-diving beaked whales. *Journal of the Acoustical Society of America* 125:2823–2832.

Received: 2 August 2010

Accepted: 31 October 2010