Interim Report

Does Depth Matter? Examining Factors That Could Influence the Acoustic Identification of Odontocete Species on Bottom-moored Recorders

Submitted to:

Naval Facilities Engineering Command Atlantic under HDR Environmental, Operations and Construction, Inc.
Contract No. N62470-10-D-3011, Task Order 21, Issued to HDR, Inc.

Prepared by:

Bio-Waves, Inc.
364 2nd Street, Suite #3
Encinitas, CA 92024
(760) 452-2575

and

Oceanwide Science Institute
P.O. Box 61692
Honolulu, HI 96739

Submitted by:

HDR
Virginia Beach, VA

November 2016

Suggested Citation:

Oswald, J.N., M.O. Lammers, R. Walker, A. Kügler, C. Hom-Weaver, and S. Coates. 2016. Does depth matter? Examining factors that could influence the acoustic identification of odontocete species on bottom-moored recorders. Interim Report. Submitted to Naval Facilities Engineering Command Atlantic, Norfolk, Virginia, under Contract No. N62470-10-3011, Task Order 21, issued to HDR, Inc.

Cover Photo Credits:

Pantropical spotted dolphins *(Stenella attenuata)* taken by Julie Oswald, Bio-Waves, Inc.

Work conducted under following contract between Bio-Waves, Inc. and HDR
MSA #: CON-005-4394-009
Subproject #164744, Agreement #105067, Task CTO21

This project is co-funded by US Fleet Forces Command and ONR Marine Mammals and Biology Program managed by Naval Facilities Engineering Command Atlantic as part of the US Navy's marine species monitoring program

Table of Contents

Figures

Tables

Abbreviations and Acronyms

dB	decibel(s)
CTD	conductivity, temperature, depth
EAR2	second generation Ecological Acoustic Recorder
LMR	Living Marine Resources
LTSA	Long-term Spectral Average
kHz	kilohertz
NAVFAC LANT	Naval Facilities Engineering Command, Atlantic
ONR	Office of Naval Research
OSI	Oceanwide Science Institute
PAM	passive acoustic monitoring
ROCCA	Real-time Odontocete Call Classification Algorithm

1. Introduction

Over the past decade, passive acoustic monitoring (PAM) methods have been adopted for obtaining information about the occurrence, distribution and behavior of marine mammals. Seafloor recorders are increasingly being used to monitor marine mammals because they allow continuous data to be collected for long periods of time without requiring the presence of a human operator. However, recordings collected using instruments on the sea floor do not typically have visual observations associated with them, so species must be identified based on the characteristics of their calls. Acoustic-based species identification can be difficult because variability in the acoustic repertoire of most marine mammal species makes them challenging to classify. The sounds produced by delphinids exhibit a large amount of overlap in time-frequency characteristics among species, making them particularly difficult for use in identifying species. Most delphinids produce two general types of vocalizations: 1) tonal, frequency modulated whistles, and; 2) short, broadband clicks. Traditionally, whistles have been used for delphinid acoustic species identification (e.g., Sturtivant and Datta 1997, Rendell et al. 1999, Oswald et al. 2013), but it has also been proposed that echolocation clicks can be used to identify species such as Risso's dolphins (*Grampus griseus*), Pacific white-sided dolphins (*Lagenorhynchus obliquidens*), short-beaked common dolphins (*Delphinus delphis*), melon-headed whales (*Peponocephala electra*), false killer whales (*Pseudorca crassidens*), and short-finned pilot whales (*Globicephala macrorhynchus*; Soldevilla et al. 2008; Jarvis et al. 2008; Baumann-Pickering et al. 2010, 2011; Roch et al. 2011).

Considerable research has focused on marine mammal acoustic species identification over the past several years and, as a result, several delphinid whistle classifiers with relatively high correct classification scores have been developed. For example, the Real-time Odontocete Call Classification Algorithm (ROCCA), a whistle classifier that has been incorporated into the PAMGuard acoustic data processing software platform, contains classifiers for species in the tropical Pacific and northwest Atlantic oceans, with temperate Pacific and Hawaii classifiers currently in development. Overall correct classification scores are 86 percent and 60 percent for the northwest Atlantic and tropical Pacific classifiers, respectively (Oswald 2013, Oswald et al. 2013). These classifiers were trained and tested using data collected near the sea surface using towed hydrophone arrays. However, marine mammals are commonly monitored using moored acoustic recorders deployed at depths ranging from tens to thousands of meters. As sound travels through the water column, physical processes such as transmission loss and multi-path propagation can cause localized maxima and minima in acoustic intensity (Au and Hastings 2008) and change the characteristics of sounds arriving at the receiver. The characteristics of sounds that arrive at a receiver are also affected by the directionality of the sound producing mechanisms, the relative position and distance of the signaler to the receiver, the seafloor type, bathymetry, sea surface roughness, and variables such as temperature, salinity and pressure gradients in the water column. It has also been suggested that the whistles produced by dolphins may change with depth due to the effects of pressure on lung volume and the sound production structures (Ridgway et al. 2001). Because of these factors, sounds arriving at a seafloor recorder may have different characteristics from those observed at the sea surface. This in turn could have an effect on the performance (i.e., correct classification scores) of classifiers trained using data obtained at the sea surface.

The use of echolocation clicks for species identification in bottom-moored recorder data is rapidly evolving methodology and presents similar challenges that exist for whistles, but with the additional confounding effects of the high directionality of clicks, uncertain amounts of intra- and inter-specific variability in click characteristics and the potential for distortions of the frequency spectrum from multipath components. These factors have the potential for yielding ambiguous or incorrect classification results or results that are not easily generalized among populations or locations. Several efforts to use echolocation clicks for species-specific identification have proposed that relevant cues are found in the characteristics of off-axis clicks (e.g., Soldevilla et al. 2008, Baumann-Pickering et al. 2010, Roch et al. 2011). However, from a biosonar standpoint, this can be difficult to reconcile with the fact that delphinid clicks lose approximately 25 to 30 decibels (dB) of amplitude when only 30° to 45° off-axis (Au et al. 2012a), that click spectral properties vary greatly with increasing off-axis angle (Au et al. 2012a), and that multiple pulses with angle-dependent temporal separation propagate out of the animal's head off-axis (Lammers and Castellote 2009, Au et al. 2012b). In other words, off-axis clicks are lower in amplitude and more variable in their spectral properties relative to on-axis clicks, so it is not clear how consistent species-specific cues can emerge from off-axis clicks. However, numerous authors report being able to discriminate among the echolocation clicks of certain species (Jarvis et al. 2008, Soldevilla et al. 2008, Baumann-Pickering et al 2010, Roch et al. 2011) which raises the following questions: 1) What features drive the species-specificity of the clicks reported? and 2) Are the cues observed tied to anatomy and physiology or are they related to species-specific behavior and/or the effects of signal propagation? It is important to address these questions to validate the underlying assumptions needed for performing species identification using echolocation clicks, namely that species-specific cues are due to variations in the sound production morphology of different species (and can therefore be generalized across populations of the same species), and that these cues are robust to different recording scenarios.

This report describes the first year of a 2-year study to explore some of the factors that may contribute to ambiguity in species identification using whistles and echolocation clicks. We have obtained field recordings of wild and trained (captive) odontocetes in their natural environment under a variety of scenarios, including at the sea surface, at multiple depths in the water column, in different geographic locations, and in different behavioral states. The characteristics of whistles and echolocation clicks recorded at varying depths are being compared to answer the following questions:

1. Does the depth at which dolphin whistles are recorded affect the received signal characteristics?
2. If received whistle characteristics are different when recorded at different depths, does the performance of species classifiers developed using whistles recorded at the surface change when applied to data from different recording depths?
3. Do reported species-specific click characteristics remain consistent across recording depths? If they do, are the cues consistent across behaviors, such as diving, surface milling and travel? If cues do not remain consistent across recording depths or behaviors, what drives the observed variability?

This project is a collaboration between Bio-Waves, Inc. and Oceanwide Science Institute (OSI), with help from the National Marine Mammal Foundation and is being jointly funded by US Fleet Forces Command under the US Navy's Marine Species Monitoring Program, the Office of Naval Research (ONR) and the US Navy's Living Marine Resources Program (LMR). This report focused on the analysis of whistles, which is the portion of the project funded by US Fleet Forces Command and led by Bio-Waves, Inc.

2. Statement of Navy Relevance

PAM is used extensively to collect information regarding marine mammal occurrence, distribution and behavior in areas with high naval activity, and mitigation efforts rely heavily on data obtained by seafloor recorders. However, the suitability of using species classifiers trained with surface data for analyzing recordings obtained at depth is currently unknown. If classifiers perform differently on data recorded at depth, it may be necessary to re-train them or develop new classifiers to ensure accurate results. Similarly, if the behavior of animals or signal propagation affects the identification of species using classifiers developed for echolocation clicks, this must be understood and integrated into analysis methods. It is important to accurately identify odontocete species in acoustic recordings because different species are known react differently to naval activities, and understanding species-specific responses is important for implementing effective mitigation measures. This study tests previously unexamined assumptions associated with the acoustic identification of odontocete species. The results of this effort will ultimately provide a better understanding of the methods presently being employed for marine mammal monitoring and mitigation, and will lead to greater confidence in their application.

3. Methods

3.1 Data Collection

Two types of vertical arrays were used to obtain recordings of whistles and echolocation clicks at different depths: 1) a surface array of microMARS recorders and four broadband hydrophones used to collect data for localization analysis, and 2) a bottom array of second generation Ecological Acoustic Recorders (EAR2s).

3.1.1 Surface Array

The surface microMARS array was deployed from a small (approximately 24-foot) vessel and was composed of two vertical sub-arrays: a localization array with four broadband hydrophones (Cetacean Research Technology C75) spaced 10 meters apart, and a line array made up of five microMARS recorders spaced 50 meters apart (**Figure 1**). The four hydrophones in the localization array were sampled simultaneously on a high-resolution portable recorder. This allowed fine-scale localization of phonating animals so their depth and distance relative to the vertical array could be established. The array of microMARS recorders extended beyond the localization array with individual recorders separated by 50 meters to a maximum depth of 250 meters. The microMARS (http://www.desertstar.com/acoustic-recorders.html) is a new type of low-cost acoustic recorder with a relatively high maximum sampling rate of 250 kilohertz (kHz) and up to 512 gigabytes of storage space. It measures only 19.5 cm x 6.5 cm, so it can be configured into a hand- deployable/retrievable vertical line array (**Figure 2**). The microMARS array was used to record visually detected schools of odontocetes at multiple depths. When a group of odontocetes was sighted by the observers on the vessel, the observers spent 30 minutes to 1 hour observing the animals to determine their behavior, direction and rate of travel. When the observers determined the behavioral state and movement patterns of the school, the vessel moved approximately 1 kilometer ahead of the school and the surface array was deployed. Recordings were made as the school approached the boat until the animals moved out of acoustic detection range. If time and conditions allowed, the array was recovered and the process was repeated. While in range of the animals, acoustic data were recorded, as well as visual data such as animal location, behavior and species identification. Four of the five microMARS were paired with a Star-Oddi DST tilt sensor, which measured and recorded temperature, depth (pressure) and tilt (in three directions). The fifth (deepest) microMARS was paired with a Star-Oddi conductivity, temperature, depth (CTD) probe, which is a miniature salinity, temperature and depth data logger. This probe was fixed to the bottom of the array to obtain profiles upon deployment and recovery. These data will be used to calculate sound speed profiles, in order to investigate the effects of sound propagation on received signal properties.

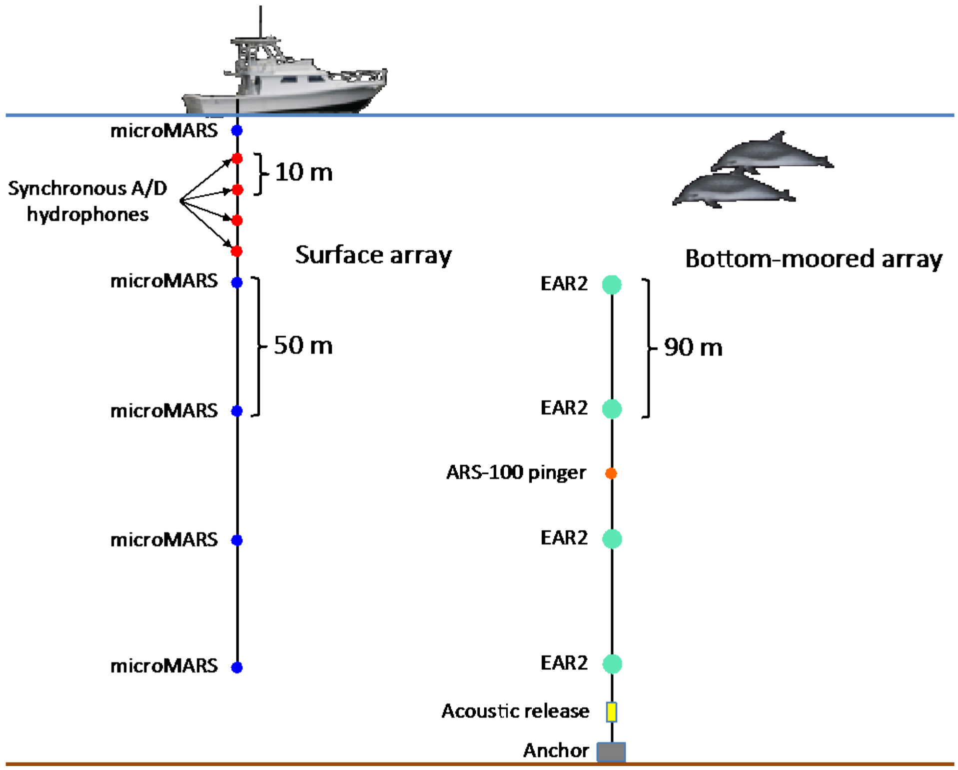

Figure 1. Schematic diagram of the surface microMARS array and bottom-moored EAR2 array.

Figure 2. MicroMARS recorder included in the surface array.

3.1.2 Bottom Array

The bottom-moored vertical array was made up of four EAR2s spaced 90 meters apart (**Figures 1** and **3**). The EAR2 is a redesigned version of the EAR and has a maximum sampling rate of 125 kHz, up to 1 terabyte of storage space and is able to sample continuously. The array also included a RJE International ARS-100 pinger, which provided a 4 to 7 kHz synchronization pulse every 30 minutes. This pulse was recorded on the four EAR2 recorders and was used to precisely time-align recordings during post-processing and analysis in order to localize signaling animals and determine their range and depth. A Star-Oddi DST centi miniature temperature and depth data logger was paired with the shallowest EAR2. The middle two EAR2s were paired with Star-Oddi DST-tilt sensor. A Star-Oddi miniature conductivity, temperature, depth probe was fixed to the bottom of the moored array to allow calculation of the sound speed profile at the deployment site at the time of deployment and the time of recovery. The EAR2 array was deployed on the sea floor at locations of known high odontocete activity at bottom depths between 500 and 950 meters for periods of 1 to 2 weeks during all field work efforts. The EAR2s were set to record on a 33 percent duty cycle (2 minutes every 6 minutes during the pilot work in Lanai, and 10 minutes every 30 minutes during the Kona and San Diego deployments.) This ensured that a sufficiently large sample size of odontocete encounters was obtained for performing statistically meaningful comparisons.

Figure 3. Bottom-moored EAR2 array showing four EAR2s, an acoustic release, 300 meters of line, and anchor (concrete blocks and sand bags).

3.1.3 Pilot Study

A pilot field effort was conducted in Hawaii from 2 to 12 August 2015. At the beginning of this effort, the bottom-moored array was deployed approximately 3 nautical miles south of the island of Lanai in waters 355 meters deep (**Figure 4**). Approximately 400 pounds of weight (sand bags and a concrete block) were used to keep the array moored in place until it was recovered on 12 August. The bottom array was set to record for 2 minutes every 6 minutes. After deployment of the bottom array, visual surveys were conducted off the islands of Maui and Lanai using a 26-foot research vessel (the *Aloha Kai*) and a 21-foot research vessel (the *Coho*; due to engine problems on the *Aloha Kai*). The surface microMARS array was used to record groups of odontocetes that were encountered during these surveys.

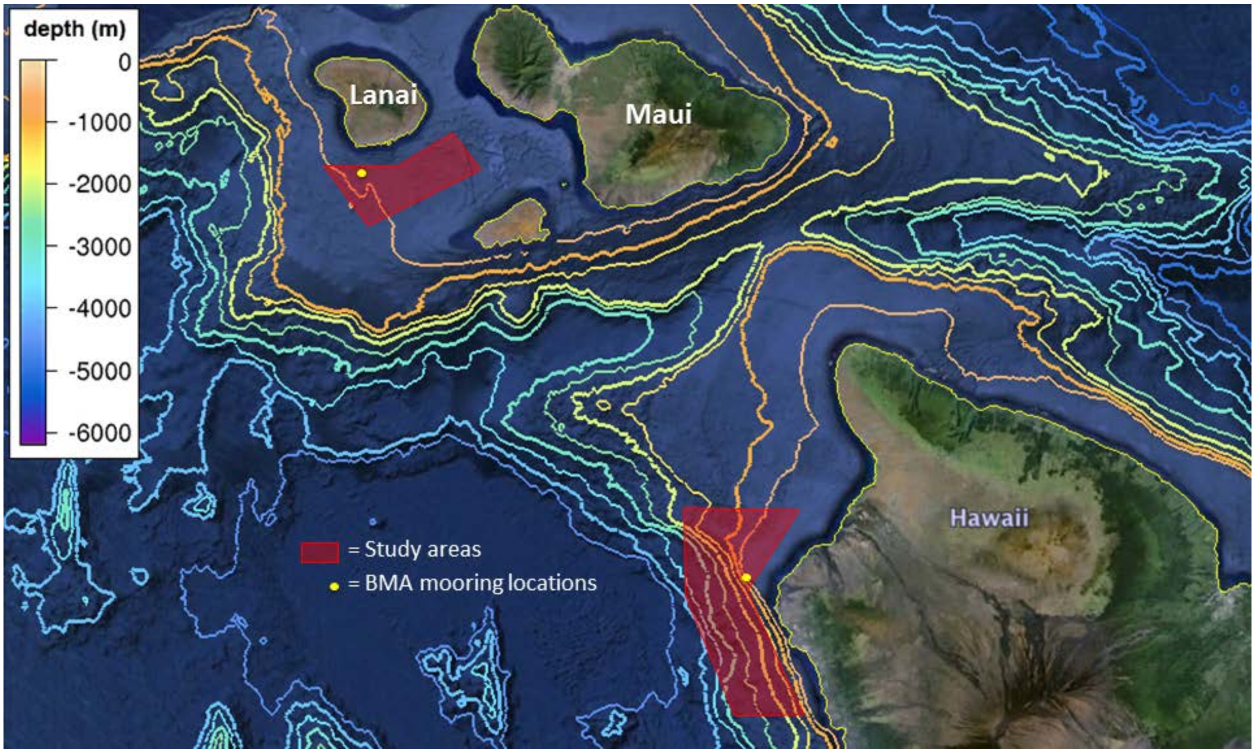

Figure 4. Lanai and Kona coast study areas, which bottom-moored array locations shown as yellow circles.

Three different microMARS hydrophones with different sensitivities and frequency ranges were compared to determine which would be best suited for this project. Two of the hydrophones had flat frequency responses up to 33 kHz and one had a flat frequency response up to 125 kHz (**Figures 5** and **6**). The configuration of microMARS hydrophones on the array was changed several times during the field testing period to compare their performance relative to one another. Various strategies for deployment and recovery of the surface array were also tested. Finally, the sub-array hydrophones and broadband recorder were tested using different gain levels.

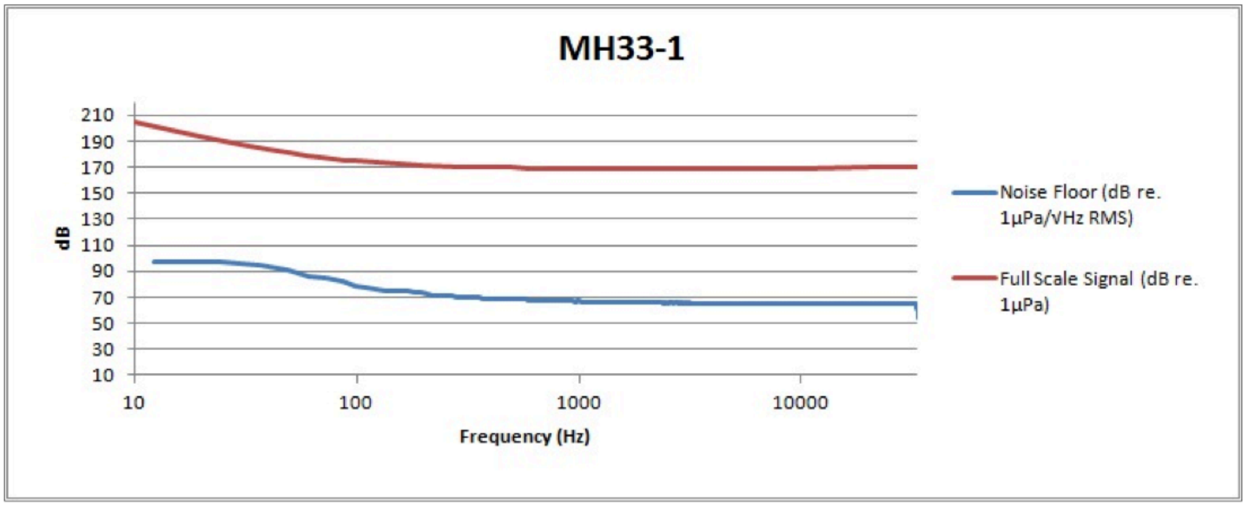

Figure 5. Generic sensitivity curve for MH33-1 microMARS hydrophone used in pilot work.

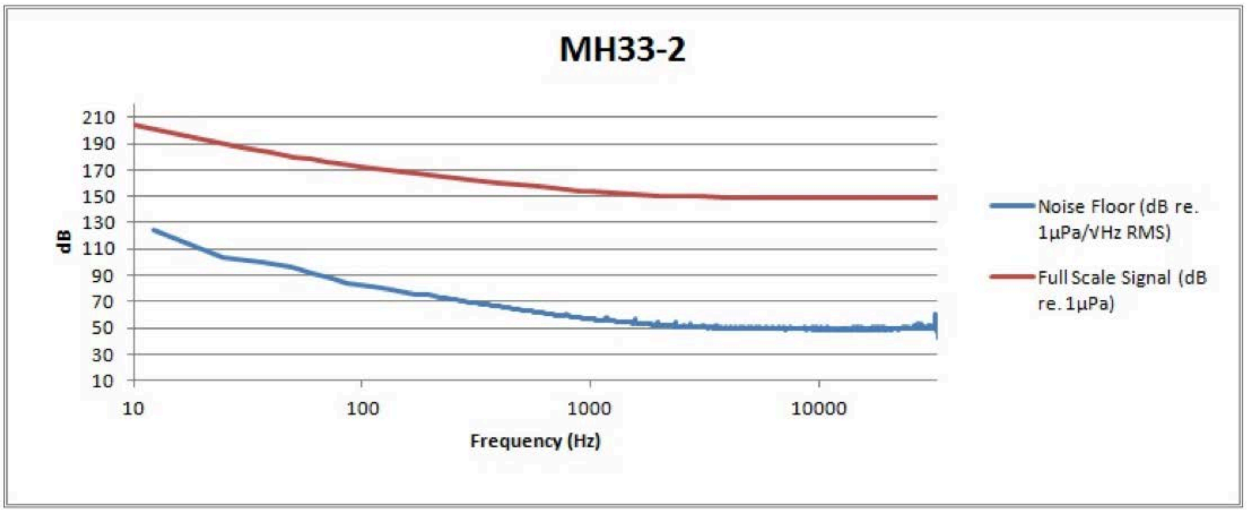

Figure 6. Generic sensitivity curve for MH33-2 microMARS hydrophone used in pilot work.

Because of the differences in microMARS sensitivities and configurations throughout the pilot study work, it was not possible to compare whistles and clicks among depths for this particular dataset.

3.1.4 Kona Data Collection

A field data collection effort was conducted off Kona, Hawaii from 2 to 14 November 2015. The bottom-moored EAR2 array was deployed north of Kona in a water depth of 400 meters for the duration of this field effort. The array was set to record for 10 minutes every 30 minutes. Visual surveys of marine mammals were conducted along the Kona coast (**Figure 4**) using a 26-foot research vessel (the *Hopena*). The surface array of microMARS was deployed near groups of odontocetes that were encountered.

3.1.5 San Diego Data Collection

A 2-week field data collection effort was conducted off San Diego, California, from 15 to 27 May 2016. During this effort, the EAR2 array was moored on the sea floor approximately 7 miles off La Jolla, California (**Figure 7**), in a water depth of 465 meters and recorded for 10 minutes every 30 minutes. Visual surveys were conducted between Point Loma and La Jolla, and on the leeward (east) side of Catalina Island using one of two 28-foot sport fishing vessels (the *Seasons* and the *Ugly Guy)*. The surface microMARS array was deployed near groups of odontocetes that were encountered during these surveys.

In addition to the vessel surveys, a controlled data collection experiment was conducted in collaboration with the Navy's Marine Mammal Program. During this effort, a trained Navy bottlenose dolphin was transported by boat to an open water location (approximately 1,000 meters in depth) off Point Loma. The animal was instructed to swim to and bite on a bite-plate positioned at a depth of 5 meters. The distance and orientation of the dolphin relative to the surface array were recorded at the beginning and ending of each trial. The dolphin was instructed to produce whistles while on station at the bite-plate. Whistles were recorded using the surface microMARS array at distances of approximately 50, 100, 250 and 400 meters away from the dolphin. The dolphin was oriented in two positions: 1) facing directly towards the array and, 2) facing directly away from the array. The magnetic bearing of the dolphin's orientation was measured and recorded using a compass app (SensorLog) on an iPad by orienting the iPad's edge along the posterior/anterior axis of the animal. The distance between the dolphin and the surface array was measured using a range finder.

3.2 Whistle Analysis

Initial whistle analyses focused on data recorded during the Lanai pilot study deployment and the Kona deployment of the bottom-mounted EAR2 array, as the San Diego data were just collected. For the Lanai data, Triton software (Wiggins 2007) was used to create long-term spectral averages (LTSAs) from recordings from all four EAR2s in the array. An analyst reviewed these LTSAs manually and created a log of all delphinid acoustic encounters in the data. Because there were no visual observations associated with EAR recordings, it was not possible to determine when one school left the area (or stopped vocalizing) and another school entered (or started vocalizing). As a proxy, for the Lanai data, an encounter was defined based on elapsed time between vocalizations. A new encounter was delineated when 30 or more minutes had elapsed between whistles or clicks. This was chosen based on our previous experience observing behavior and recording vocalizations of delphinids. A slightly different definition of encounter had to be used for the Kona data because of the duty cycle used during this deployment (10 minutes on, 20 minutes off). Since the EAR2s recorded every 30 minutes, which was also the time gap used to denote separate encounters in the Lanai data, each 10-minute sound file was treated as a separate encounter. Analysts examined each sound file from the shallowest EAR2 in Triton, and selected a subset of 19 events that contained sufficient whistles and clicks for analysis.

For each encounter, analysts used Raven software to examine spectrograms from each EAR2. Up to 25 whistles were randomly selected from the shallowest EAR2 (EAR2-1) and up to 25 different whistles were randomly selected from the deepest EAR2 (EAR2-4). These same whistles (25 from EAR2-1 and 25 from EAR2-4) were then selected from all other EAR2s, if they were present in those recordings. This resulted in a sample size of up to 50 whistles per EAR2

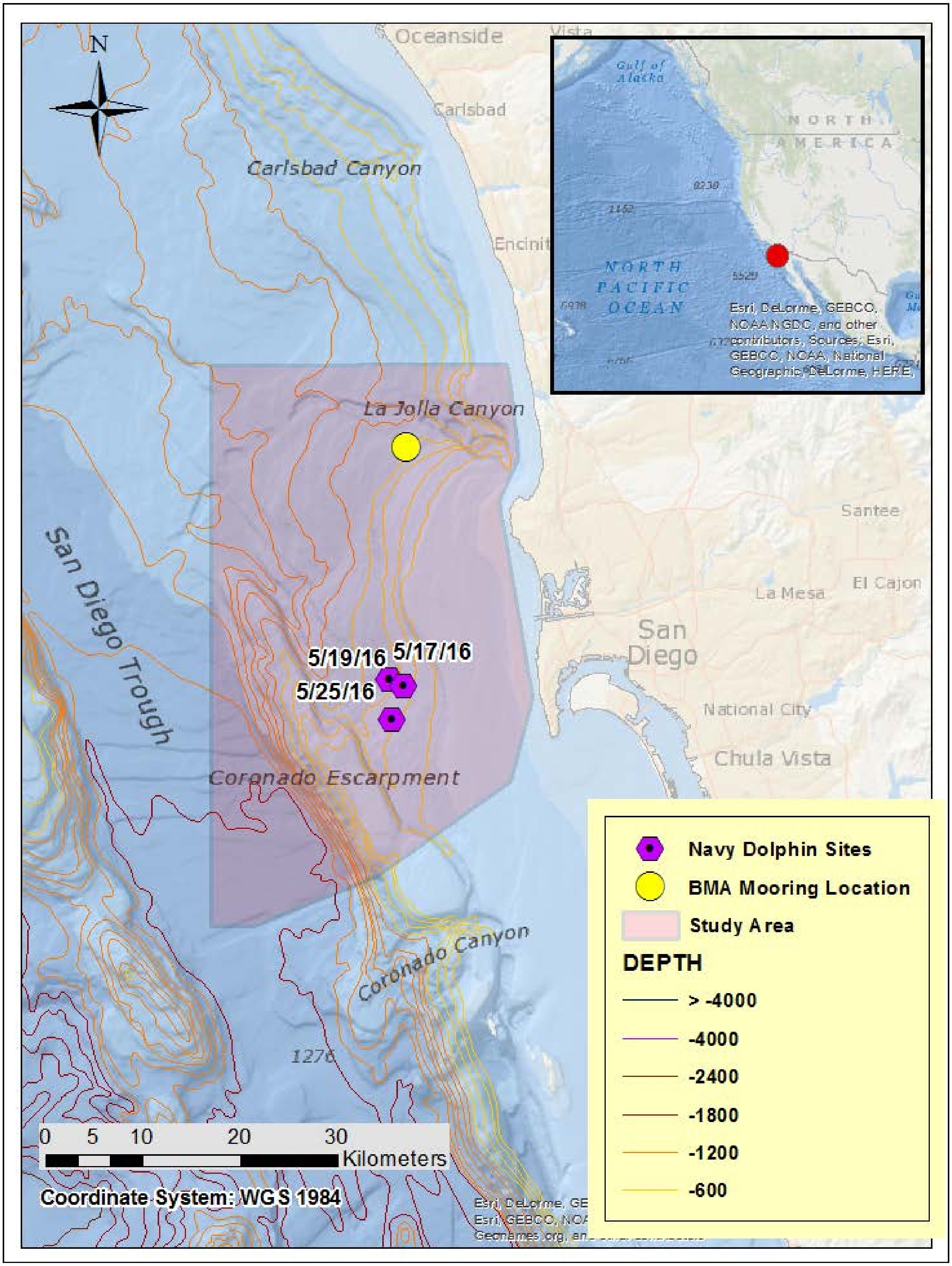

Figure 7. Map of San Diego study area with location of the bottom-mounted EAR2 array (yellow circle), and locations of Navy dolphin trials (purple hexagons).

per encounter. Selecting whistles from both EAR2-1 and EAR2-4 reduced the risk of missing whistles that might only be heard on the shallowest or deepest EAR2s due to the position of the animal relative to the array. Only encounters that contained at least 10 whistles with moderate to good signal-to-noise ratios (i.e., at least 3 dB) from EAR2-1 and EAR2-4 were included in the analyses.

Analysts then extracted time-frequency contours from the selected whistles in each encounter using the ROCCA (Oswald et al. 2013) module in the acoustic data processing software platform, PAMGuard (Gillespie et al. 2008). To extract time-frequency contours, the analyst traced contours on ROCCA's spectrographic display using a computer touch-pad (**Figure 8**). ROCCA then automatically measured 50 variables from each extracted contour, including duration, frequencies (e.g., minimum, maximum, beginning, ending, and at various points along the whistle), slopes, and variables describing shape of the whistles (e.g., number of inflection points and steps; see Barkley et al. 2011 for a complete list and description of variables measured). Since not all selected whistles appeared on all four EAR2s (e.g., **Figure 9**), two different analyses were conducted. Whistle variables were compared among EAR2s for each encounter using Kruskall-Wallis tests and post-hoc Dunn's tests with Bonferonni correction, to determine if there were significant differences among whistles recorded at different depths. In addition, the same statistical analyses were performed on only whistles that appeared on all four EARs in each encounter, to determine if the contour of specific whistles changed significantly between different depths.

Whistle variables were also used to classify whistles to species using a random forest classifier within ROCCA. A random forest is a collection of decision trees grown using binary partitioning of the data. Each binary partition is based on the value of one whistle variable (Breiman 2001). Randomness is introduced into the tree-growing process by examining a random subsample of all of the variables at each node. The variable that produces the most homogeneous split is chosen at each partition. When whistle variables are run through a random forest, each of the trees in the forest produces a species classification. Each tree can be considered one 'vote' for a given species classification. Votes are then tallied over all trees and the whistle classification is based on the species with the most 'votes.' In addition to classifying individual whistles, entire acoustic encounters are classified based on the number of tree classifications for each species, summed over all of the whistles that were analyzed for that encounter.

The random forest classifier used in this analysis was a two-stage classifier trained using whistles recorded from single-species schools in the tropical Pacific Ocean. Six species were included in the model: short-finned pilot whales (*Globicephala macrorhynchus*), false killer whales (*Pseudorca crassidens*), pantropical spotted dolphins (*Stenella attenuata*), bottlenose dolphins (*Tursiops truncatus*), rough-toothed dolphins (*Steno bredanensis*), and spinner dolphins (*Stenella longirostris*). The first stage consisted of classifying whistles to one of two categories: 'large delphinids-*Steno'* (including false killer whales, pilot whales and rough-toothed dolphins) and '*Stenella-Tursiops*' (including spinner, spotted, and bottlenose dolphins). In stage two of the model, whistles within each category were then classified to species (**Figure 10**). Four-fold cross-validation was used to test the performance of the classifier. To accomplish this, the training dataset was randomly divided into four subsets of data, with whistles from the same encounter kept together in the same dataset. One dataset was used to train the classifier and the other three were used to test the classifier. The datasets were then swapped so that each was used as both a training and a testing dataset. This procedure was repeated 50 times and the results were compiled (**Table 1**). Classification success was evaluated by examining the average percent of encounters that were correctly classified for each species and comparing that to the classification score that would be expected by chance (17 percent for six species). Overall, 58 percent of encounters were correctly classified in the training dataset. All correct classification scores were significantly greater than expected by chance (Fisher's exact test, α = 0.05).

The EAR2 encounters were analyzed with this classifier and encounters were classified based on classification results summed over all whistles in the encounter. Individual whistle classifications and overall encounter classifications were compared among EAR2s for each encounter to evaluate whether observed differences in whistle structure among EAR2s affected classifier performance. The classification analysis was performed on two different sets of whistles for each encounter. First, all whistles recorded on each EAR2 were included to examine whether the whistles available for analysis affect the results. Second, only whistles recorded on all four EAR2s were included to examine how differences in whistle structure affect classification results.

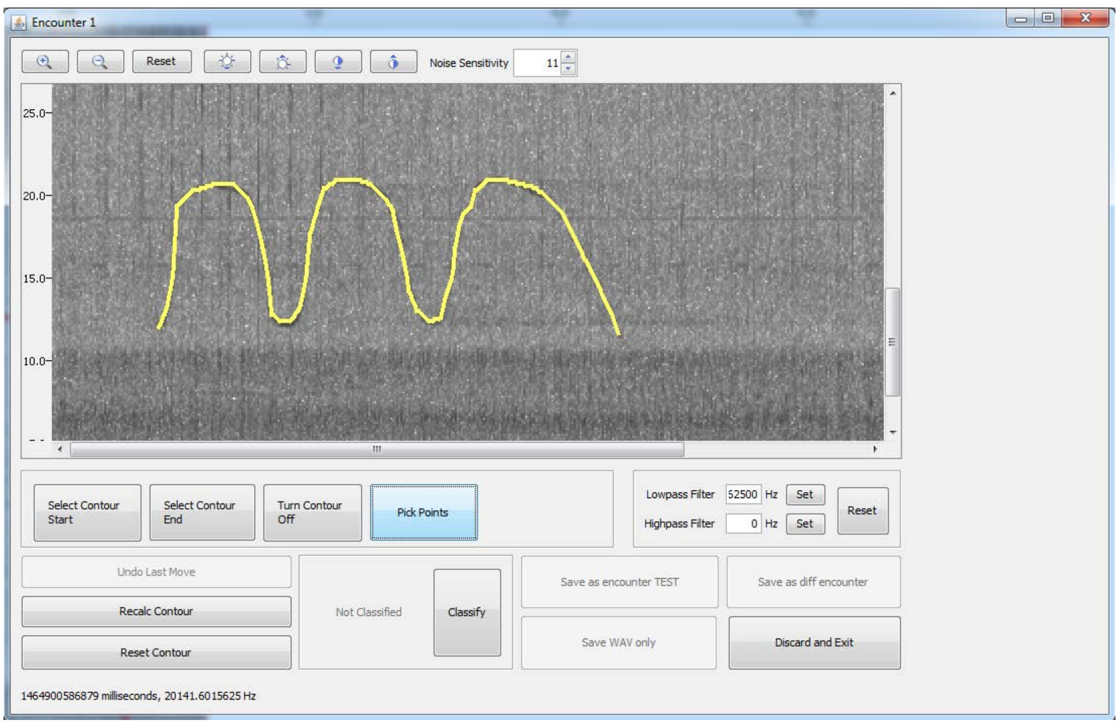

Figure 8. Example of whistle contour traced in ROCCA module in PAMGuard.

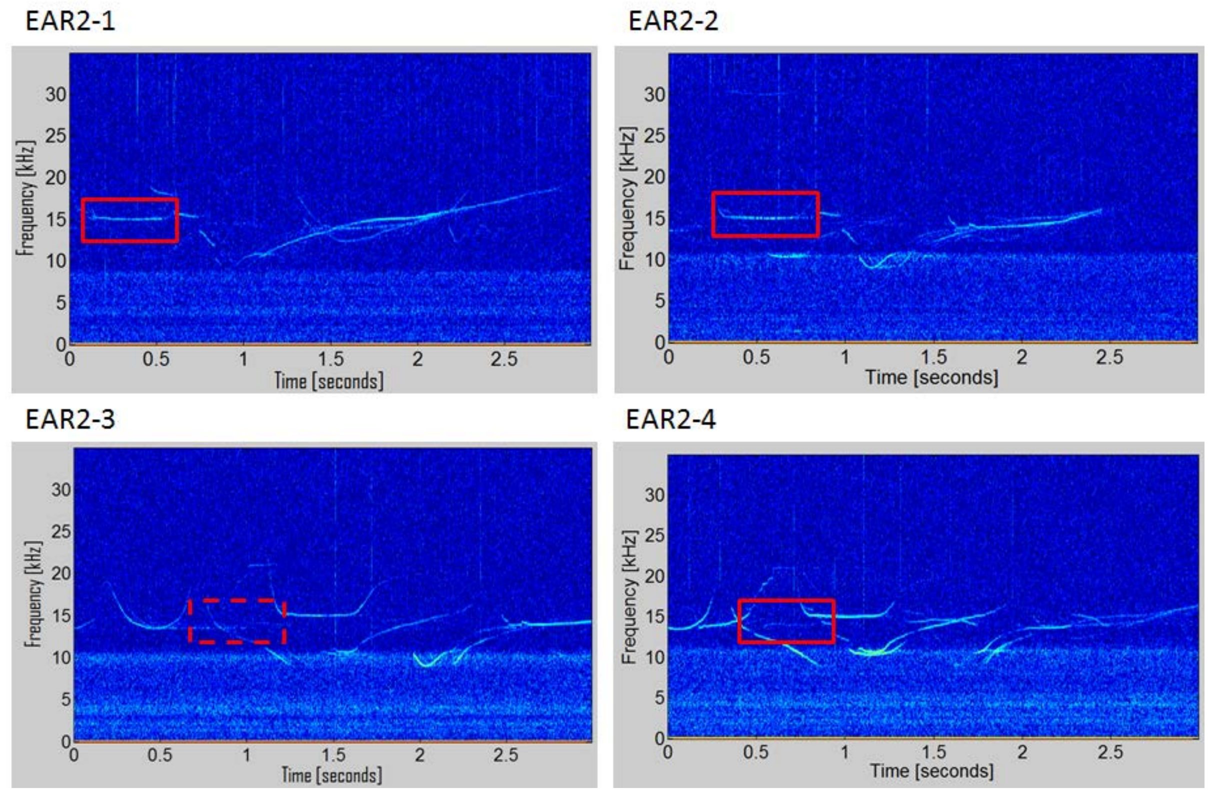

Figure 9. Example of whistle contour not appearing on all EAR2s. **Solid red lines show EAR2s where the whistle contour was visible, dashed red line shows where contour should appear but was not visible.**

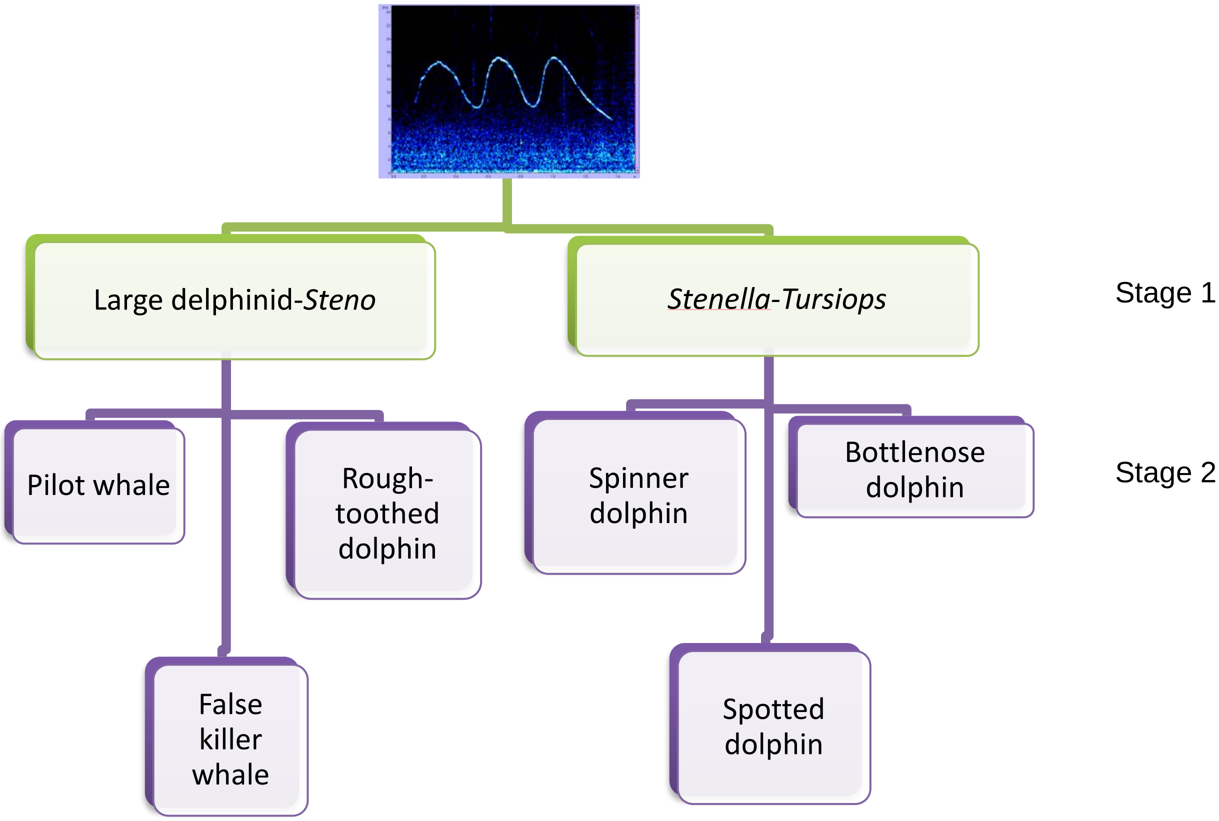

Figure 10. Schematic diagram of the two-stage random forest classifier. In stage one, whistles are classified to one of two broad categories ('large delphinid-steno' or 'Stenella-Tursiops'). In stage two, whistles within each category are classified to individual species or species-group.

Table 1. Confusion matrix for the two-stage classifier used to classify encounters. The percentage of encounters correctly classified for each species is in bold, with standard deviations in parentheses. The number of encounters in the training dataset are given for each species.

Actual species	Percent classified as						Number of encounters
	Pilot whale	False killer whale	Spotted dolphin	Rough-toothed dolphin	Spinner dolphin	Bottlenose dolphin	
Pilot whale	**59 (0.1)**	24 (0.1)	0 (0.04)	0.1 (0.02)	0 (0.04)	0 (0.04)	12
False killer whale	16 (0.05)	**72 (0.1)**	0 (0)	0.1 (0.1)	0 (0)	0 (0)	9
Spotted dolphin	0 (0)	0 (0.04)	**58 (0.1)**	0 (0.01)	20 (0.06)	19 (0.1)	17
Rough-toothed dolphin	0 (0.05)	13 (0.07)	0 (0.01)	**57 (0.09)**	13 (0.04)	0 (0.03)	12
Spinner dolphin	0 (0.05)	0 (0.03)	0 (0.07)	0 (0.02)	**56 (0.1)**	20 (0.07)	14
Bottlenose dolphin	0 (0.03)	0 (0.04)	16 (0.09)	15 (0.08)	19 (0.1)	**47 (0.1)**	8

4. Results

4.1 Pilot Field Effort

A total of 31 hours of field survey effort were spent searching for and/or recording odontocetes over 6 days. Odontocetes were encountered and recorded with the surface microMARS array on four occasions. These encounters included two groups of spotted dolphins, one group of spinner dolphins, and one group of short-finned pilot whales.

All EAR2s recorded successfully during the deployment period and the ARS-100 pinger transmitted signals every 30 minutes as expected. The average depth of each EAR2, determined from depth recorder measurements, is provided in **Table 2**. The EAR2 array recordings yielded 2,063 2-minute files, or approximately 68.7 hours of data per recorder. These recordings contained 28 delphinid acoustic encounters.

Table 2. Average depth (standard deviation in parentheses) of each EAR2. No depth sensors were used for the pilot work, so depths are estimated based on the depth of deployment and the spacing between EAR2s. Depths for the Kona field effort are based on depth recorder readings throughout the deployment. Depths for the San Diego field effort are based on the deployment depth of the array and the distance between EAR2s.

EAR2 number	Average Depth (m)		
	Pilot work	Kona	San Diego
EAR2-1	70	118 (1.2)	176.6
EAR2-2	160	209 (0.9)	266.6
EAR2-3	250	289 (0.6)	356.6
EAR2-4	340	389 (1.8)	446.6

4.2 Kona Field Effort

A total of 80 hours of survey effort were spent searching for and/or recording odontocetes over 8 days. Odontocetes were encountered and recorded with the surface microMARS array on 14 occasions. These encounters are summarized in **Table 3**. The EAR2 recordings yielded 538 10-minute files, or approximately 90 hours of data per recorder. Each 10-minute recording was treated as a separate encounter, resulting in 538 encounters. The average depth of each EAR2, based on depth recorder readings throughout the deployment is given in **Table 2**.

Table 3. Number of encounters per species recorded with the microMARS surface array during the Kona field effort.

Species	Number of encounters
Spotted dolphin	7
Short-finned pilot whale	3
Rough-toothed dolphin	1
False killer whale	1
Spinner dolphin	1
Pygmy killer whale	1

4.3 San Diego Field Effort

A total of approximately 75 hours of survey effort were spent searching for and/or recording odontocetes over 11 days. Odontocetes were encountered and recorded with the surface microMARS array on 16 occasions. These encounters are summarized in **Table 4**. The EAR2s yielded 521 10-minute files, or approximately 87 hours of data per recorder. The average depth of each EAR2, based on deployment depth and distance between EAR2s, is given in **Table 2**.

Table 4. Number of encounters per species recorded with the microMARS surface array during the San Diego field effort.

Species	Number of encounters
Common dolphin (sp.)	8
Short-beaked common dolphin	3
Long-beaked common dolphin	2
Bottlenose dolphin	2
Unidentified Odontocete	1

Approximately 8 hours of the 75 total survey hours (10.6 percent) were spent working with the Navy dolphin to collect controlled data. Results of these trials will be analyzed in the next few months and included in a later report.

4.4 Whistle Analysis

4.4.1 Pilot Study Analysis

The surface microMARS array that was used during the pilot field effort contained different hydrophones deployed at different depths. These hydrophones had different sensitivities and frequency responses and so these data cannot be used to compare signal characteristics among depths. These data did, however, provide valuable information regarding which microMARS hydrophone is most appropriate for this work. Based on examinations of the microMARS recordings, it was decided that the broadband hydrophones (125 kHz) with the highest sensitivity are necessary to capture the most whistles and echolocation clicks. These hydrophones were subsequently used in all microMARS during the Kona and San Diego field work.

Whistles were measured and classified from 20 of the 28 encounters recorded with the Lanai EAR2 array (**Table 5**). Eight of the 28 encounters did not contain enough whistles to be included in the analysis. In the 20 encounters that were included in the analysis, not all whistles were detected on all EAR2s. **Figure 11** shows the percentage of whistles detected on only one EAR2, on only two EAR2s, on only three EAR2s and on all four EAR2s for each encounter. For some encounters (e.g., encounters 25, 27, and 28) a high percentage of whistles were detected on all four EAR2s, but for others (e.g., encounters 9, 10, and 11), the majority of whistles were detected on only one EAR2. In most cases (16 out of 20 encounters), the greatest number of whistles was detected on the deepest EAR2 (EAR2-4; **Table 6**).

Table 5. Date, start time, end time for each acoustic encounter recorded on the EAR2 array during the pilot work off Lanai. **Detections with an asterisk were used to compare whistle contours that appeared on all four EAR2s.**

Encounter	Date	Start time	End time
1	8/4/2015	4:24:13	6:25:52
2*	8/4/2015	12:52:42	13:30:04
3*	8/4/2015-8/5/2015	18:42:51	3:30:03
5	8/5/2015	7:48:03	8:19:05
7	8/6/2015	3:54:55	4:13:03
9	8/6/2015	10:42:05	11:06:04
10	8/6/2015	12:43:04	13:07:04
11	8/6/2015	17:00:03	17:30:07
12*	8/6/2015-8/7/2015	18:42:34	1:49:39
13*	8/7/2015	3:48:04	6:07:44
14*	8/7/2015-8/8/2015	20:06:35	1:36:03
16*	8/8/2015	10:54:04	11:48:03
18*	8/8/2015-8/9/2015	20:12:34	7:54:04
19*	8/9/2015	8:36:04	12:12:04
21*	8/9/2015-8/10/2015	16:43:05	0:54:04
22*	8/10/2015	2:55:34	7:30:04
23	8/10/2015	11:48:03	13:31:12
25*	8/10/2015-8/11/2015	22:00:03	9:24:34
27*	8/11/2015	17:24:59	20:54:35
28*	8/11/2015-8/12/2015	21:55:43	6:36:04

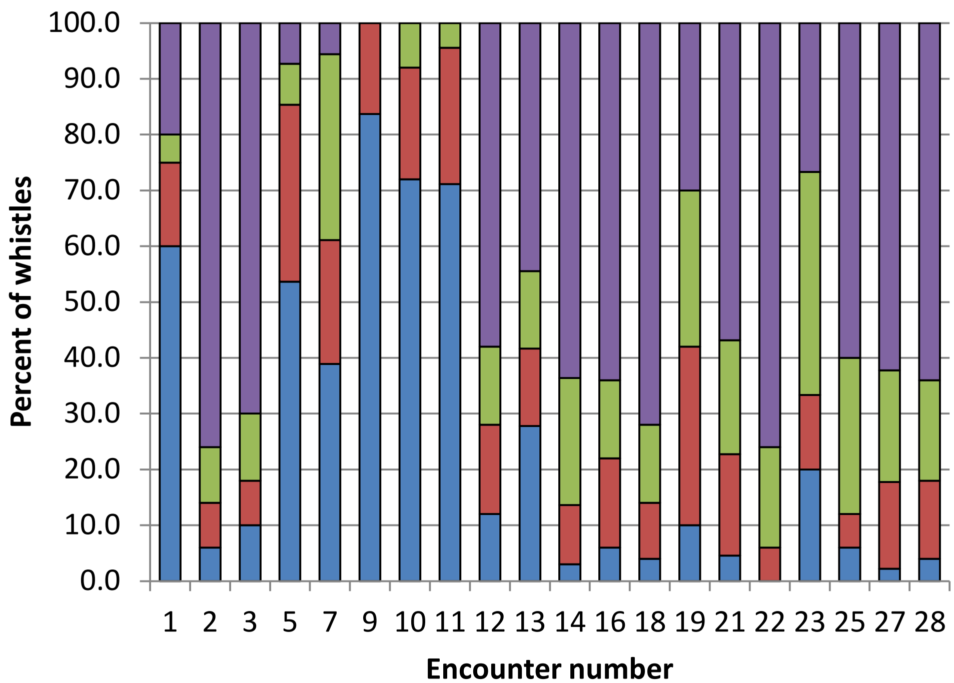

Figure 11. Percentage of whistles detected on only one EAR2 (purple), on two EAR2s (green), on three EAR2s (red) and on all four EAR2s (blue).

Table 6. Variables that were significantly different when whistles were measured from recordings made at different depths, including whistles that did not appear on all EAR2s. **Species that encounters were classified as based on those same whistles are given for each EAR2, with number of whistles included in the analyses in parentheses.**

Encounter	Significant Variable	EAR2s	p	Classified as			
				EAR2-1	EAR2-2	EAR2-3	EAR2-4
1	Mean frequency	EAR2-1-EAR2-4	0.008	Spinner (5)	Bottlenose (6)	Bottlenose (7)	Bottlenose (19)
		EAR2-2-EAR2-4	0.04				
		EAR2-3-EAR2-4	0.05				
	Median frequency	EAR2-3-EAR2-4	0.08				
	Maximum frequency	EAR2-1-EAR2-4	0.02				
		EAR2-2-EAR2-4	0.04				
	Minimum frequency	EAR2-1-EAR2-4	0.04				
	Center frequency	EAR2-1-EAR2-4	0.004				
		EAR2-3-EAR2-4	0.02				
2	None			Bottlenose (44)	Bottlenose (42)	Bottlenose (43)	Bottlenose (49)
3	None			Rough-toothed (40)	Rough-toothed (40)	Bottlenose (43)	Rough-toothed (48)
5	Mean frequency	EAR2-1-EAR2-4	0.01	Spinner (16)	Spinner (12)	Spinner (20)	Spinner (21)
	Median frequency	EAR2-1-EAR2-4	0.02				
	Beginning frequency	EAR2-1-EAR2-4	0.02				
	Minimum frequency	EAR2-1-EAR2-3	0.02				
		EAR2-1-EAR2-4	0.04				
	Center frequency	EAR2-1-EAR2-3	0.02				
		EAR2-1-EAR2-4	0.006				
	Mean slope	EAR2-1-EAR2-4	0.02				
7	Median frequency	EAR2-1-EAR2-4	0.02	Spinner (12)	Spinner (10)	Spotted (9)	Spotted (6)
	Center frequency	EAR2-1-EAR2-4	0.04				
	Mean negative slope	EAR2-1-EAR2-4	0.04				
9	none			Spinner (1)	Bottlenose (7)	Spinner (13)	Spinner (29)
10	Mean frequency	EAR2-2-EAR2-4	0.04	n/a (0)	Spotted (2)	Spinner (12)	Spinner (20)
		EAR2-3-EAR2-4	0.02				
	Median frequency	EAR2-2-EAR2-4	0.04				
		EAR2-3-EAR2-4	0.02				
	Maximum frequency	EAR2-2-EAR2-4	0.04				

Encounter	Significant Variable	EAR2s	p	Classified as			
				EAR2-1	EAR2-2	EAR2-3	EAR2-4
	Center frequency	EAR2-2-EAR2-4	0.04				
		EAR2-3-EAR2-4	0.04				
11	Mean frequency	EAR2-2-EAR2-3	0.008	Spinner (6)	Rough-toothed (14)	Spinner (27)	Spinner (13)
	Median frequency	EAR2-1-EAR2-2	0.04				
		EAR2-2-EAR2-3	0.04				
	Ending frequency	EAR2-1-EAR2-2	0.004				
	Maximum frequency	EAR2-2-EAR2-3	0.01				
	Center frequency	EAR2-2-EAR2-3	0.01				
	Duration	EAR2-2-EAR2-3	0.04				
		EAR2-2-EAR2-4	0.006				
	Percent flat	EAR2-1-EAR2-2	0.001				
		EAR2-2-EAR2-4	0.01				
	Mean absolute slope	EAR2-1-EAR2-2	0.02				
12	None			Spinner (38)	Bottlenose (37)	Bottlenose (41)	Spinner (43)
13	Maximum frequency	EAR2-1-EAR2-4	0.01	Spinner (25)	Spinner (20)	Spinner (25)	Spinner (29)
14	None			Pilot whale (55)	Pilot whale (54)	Rough-toothed (57)	Rough-toothed (63)
16	None			Pilot whale (39)	Pilot whale (38)	Pilot whale (44)	Pilot whale (47)
18	None			Bottlenose (46)	Bottlenose (42)	Bottlenose (44)	Bottlenose (45)
19	Mean frequency	EAR2-1-EAR2-2	0.02	Spinner (35)	Spinner (23)	Spinner (38)	Spinner (43)
		EAR2-1-EAR2-4	0.01				
	Median frequency	EAR2-1-EAR2-2	0.02				
		EAR2-1-EAR2-4	0.01				
	Maximum frequency	EAR2-1-EAR2-2	0.04				
21	None			Bottlenose (34)	Bottlenose (30)	Bottlenose (40)	Bottlenose (41)
22	Minimum frequency	EAR2-1-EAR2-3	0.04	Bottlenose (43)	Bottlenose (45)	Bottlenose (48)	Bottlenose (49)
23	None			Spinner (5)	Spotted (10)	Bottlenose (12)	Bottlenose (14)
25	None			Bottlenose (46)	Bottlenose (38)	Bottlenose (44)	Bottlenose (43)
27	None			Bottlenose (34)	Spinner (40)	Spinner (40)	Bottlenose (40)
28	None			Bottlenose (42)	Bottlenose (40)	Bottlenose (42)	Bottlenose (47)

Results of statistical comparisons among EAR2 depths and the classification results are provided in **Table 6** for all whistles measured for each encounter, and in **Table 7** for only those whistles that appeared on all four EAR2s. For all whistles measured for each encounter, one or more variables were significantly different (Kruskall-Wallis test and post-hoc Dunn's tests with Bonferonni correction, α=0.05) when compared among EAR2s for 8 of the 20 encounters (40 percent). Most of the significant differences were for variables measuring frequency characteristics; mean, median and center frequency were significantly different among EAR2s for six of the encounters that had significant differences (**Table 5**). Slope variables, duration, and other frequency variables were also significantly different for some of the encounters. **Figure 12** shows an example of a whistle that was detected on more than one EAR2, but with only a portion of the contour evident in some of the recordings. For only those whistles that appeared on all four EAR2s, only 1 of the 13 encounters (8 percent) showed a significant difference (Kruskall-Wallis test and post-hoc Dunn's tests with Bonferonni correction, α=0.05).

Table 7. Percentage of trees that voted for each species, calculated using all whistles analyzed for each EAR2.

Encounter Number	EAR2 Number	Number of Whistles	Percent of Tree Votes					
			Pilot whale	False killer whale	Spotted dolphin	Rough-toothed dolphin	Spinner dolphin	Bottlenose dolphin
1	EAR2-1	5	0.0%	0.0%	18.7%	0.0%	41.7%	39.7%
	EAR2-2	6	0.0%	0.0%	22.6%	0.0%	36.7%	40.6%
	EAR2-3	7	0.0%	0.0%	17.4%	0.0%	40.7%	41.9%
	EAR2-4	19	7.2%	6.7%	11.6%	17.6%	21.5%	35.3%
2	EAR2-1	44	0.4%	1.1%	27.5%	3.1%	32.8%	35.1%
	EAR2-2	42	0.6%	1.0%	27.6%	3.1%	31.0%	36.6%
	EAR2-3	43	0.3%	0.5%	27.6%	1.5%	30.3%	39.7%
	EAR2-4	49	0.9%	1.6%	27.3%	5.6%	28.4%	36.2%
3	EAR2-1	40	14.9%	17.7%	9.7%	24.9%	11.2%	21.5%
	EAR2-2	40	13.0%	17.0%	9.4%	30.0%	9.5%	21.2%
	EAR2-3	43	14.8%	16.4%	10.9%	22.2%	11.7%	23.9%
	EAR2-4	48	13.8%	17.5%	9.4%	27.0%	10.1%	22.2%
5	EAR2-1	16	0.0%	0.0%	24.6%	0.0%	60.5%	14.9%
	EAR2-2	12	0.0%	0.0%	23.4%	0.0%	55.3%	21.4%
	EAR2-3	20	1.9%	6.3%	19.0%	11.8%	38.5%	22.5%
	EAR2-4	21	1.3%	3.0%	20.8%	10.0%	41.2%	23.7%
7	EAR2-1	12	0.0%	0.0%	13.5%	0.0%	59.6%	26.9%
	EAR2-2	10	0.0%	0.0%	27.5%	0.0%	42.3%	30.2%
	EAR2-3	9	3.2%	2.7%	28.5%	16.3%	28.0%	21.3%
	EAR2-4	6	2.1%	2.0%	34.9%	12.5%	24.1%	24.3%
9	EAR2-1	1	0.0%	0.0%	12.8%	0.0%	53.8%	33.5%
	EAR2-2	7	5.8%	12.1%	16.9%	25.1%	14.0%	26.3%
	EAR2-3	13	0.7%	1.5%	19.5%	13.2%	38.1%	27.1%
	EAR2-4	29	0.7%	1.9%	20.7%	4.3%	37.6%	34.8%
10	EAR2-2	2	0.0%	0.0%	35.2%	0.0%	30.1%	34.7%
	EAR2-3	12	0.0%	0.0%	33.1%	0.0%	39.8%	27.1%
	EAR2-4	20	0.7%	1.1%	22.9%	3.1%	38.6%	33.5%
11	EAR2-1	6	0.0%	0.0%	20.0%	0.0%	66.9%	13.1%
	EAR2-2	14	2.8%	5.8%	15.7%	34.3%	31.4%	10.0%
	EAR2-3	27	0.5%	1.3%	18.0%	9.3%	49.8%	21.1%

Encounter Number	EAR2 Number	Number of Whistles	Percent of Tree Votes					
			Pilot whale	False killer whale	Spotted dolphin	Rough-toothed dolphin	Spinner dolphin	Bottlenose dolphin
	EAR2-4	13	0.6%	0.6%	17.1%	6.5%	39.6%	35.6%
12	EAR2-1	38	0.0%	0.0%	20.0%	0.0%	40.0%	40.0%
	EAR2-2	37	0.0%	0.0%	20.3%	0.0%	39.8%	39.9%
	EAR2-3	41	0.0%	0.0%	20.1%	0.0%	39.2%	40.7%
	EAR2-4	43	0.0%	0.0%	20.7%	0.0%	40.6%	38.7%
13	EAR2-1	25	0.0%	0.0%	16.3%	0.0%	54.1%	29.5%
	EAR2-2	20	0.0%	0.0%	16.4%	0.0%	52.5%	31.1%
	EAR2-3	25	0.3%	0.5%	18.2%	3.3%	47.0%	30.8%
	EAR2-4	29	0.3%	0.3%	17.1%	2.9%	47.6%	31.9%
14	EAR2-1	55	24.6%	17.4%	7.0%	23.5%	18.1%	9.4%
	EAR2-2	54	28.0%	17.2%	6.2%	27.0%	11.6%	10.0%
	EAR2-3	57	23.9%	16.6%	6.4%	26.1%	16.5%	10.4%
	EAR2-4	63	22.3%	15.3%	6.6%	25.8%	19.4%	10.5%
16	EAR2-1	39	48.7%	35.0%	1.7%	11.1%	1.0%	2.5%
	EAR2-2	38	43.9%	37.8%	2.1%	10.4%	2.5%	3.3%
	EAR2-3	44	43.5%	35.0%	2.0%	14.7%	1.5%	3.4%
	EAR2-4	47	39.4%	34.2%	2.9%	13.6%	4.2%	5.7%
18	EAR2-1	46	1.6%	1.2%	24.8%	1.5%	28.6%	42.2%
	EAR2-2	42	0.2%	0.9%	24.9%	1.3%	28.3%	44.4%
	EAR2-3	44	0.8%	1.4%	24.3%	4.6%	26.2%	42.7%
	EAR2-4	45	0.6%	1.9%	24.0%	2.0%	29.9%	41.7%
19	EAR2-1	35	0.0%	0.0%	23.4%	0.0%	45.5%	31.1%
	EAR2-2	23	2.4%	4.1%	21.3%	6.6%	36.1%	29.6%
	EAR2-3	38	0.9%	0.7%	21.9%	3.7%	39.7%	33.1%
	EAR2-4	43	0.4%	0.4%	20.7%	3.9%	38.3%	36.3%
21	EAR2-1	34	0.7%	0.5%	24.1%	1.8%	34.4%	38.5%
	EAR2-2	30	0.0%	0.0%	32.1%	0.0%	26.8%	41.1%
	EAR2-3	40	0.0%	0.0%	27.6%	0.0%	31.2%	41.2%
	EAR2-4	41	0.0%	0.0%	27.8%	0.0%	29.4%	42.8%
22	EAR2-1	43	1.3%	2.0%	24.4%	3.8%	28.9%	39.7%
	EAR2-2	45	1.0%	2.1%	27.3%	3.6%	26.0%	40.0%
	EAR2-3	48	1.5%	2.8%	26.0%	4.0%	24.8%	40.9%
	EAR2-4	49	2.1%	3.3%	26.5%	4.8%	24.5%	38.8%
23	EAR2-1	5	0.0%	0.0%	26.1%	0.0%	45.1%	28.8%
	EAR2-2	10	0.0%	0.0%	39.2%	0.0%	28.2%	32.6%
	EAR2-3	12	0.4%	0.7%	30.9%	7.2%	28.1%	32.7%
	EAR2-4	14	0.0%	0.0%	29.4%	0.0%	31.6%	39.0%
25	EAR2-1	46	0.7%	1.2%	28.1%	2.4%	29.7%	37.9%
	EAR2-2	38	1.0%	1.9%	28.0%	5.1%	26.2%	38.0%
	EAR2-3	44	0.5%	1.4%	30.1%	2.6%	27.1%	38.2%
	EAR2-4	43	0.8%	1.7%	30.0%	4.4%	27.1%	35.9%
27	EAR2-1	34	6.1%	2.7%	13.7%	2.9%	36.7%	37.8%
	EAR2-2	40	4.6%	2.6%	15.9%	0.2%	38.7%	37.9%
	EAR2-3	40	4.0%	2.1%	16.7%	3.9%	37.5%	35.9%
	EAR2-4	40	0.5%	1.7%	17.5%	2.8%	36.8%	40.8%
28	EAR2-1	42	2.4%	5.0%	21.9%	6.9%	26.7%	37.1%
	EAR2-2	40	3.7%	8.3%	19.8%	8.0%	23.4%	36.8%
	EAR2-3	42	2.8%	6.4%	21.4%	9.8%	23.8%	35.8%
	EAR2-4	47	2.5%	6.3%	22.0%	8.2%	24.1%	36.8%

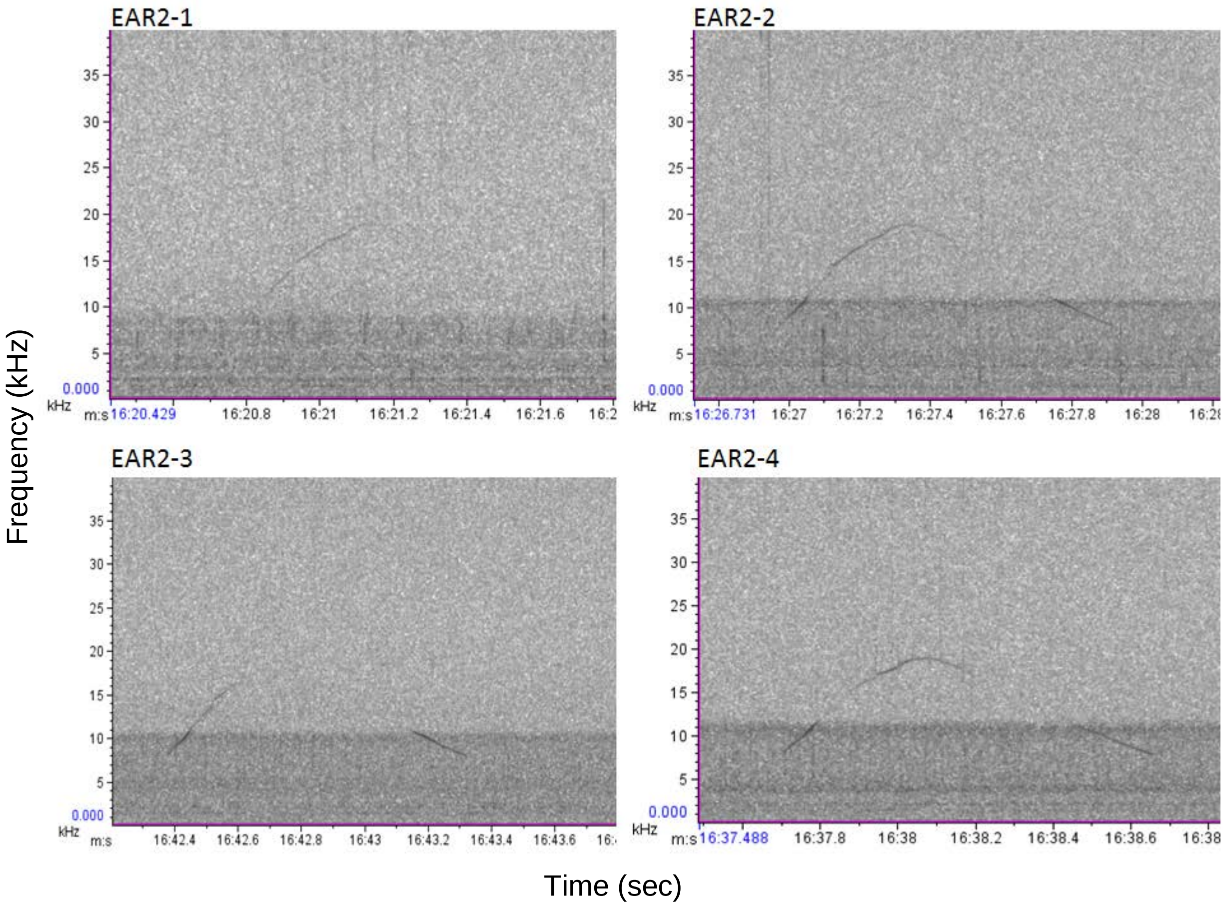

Figure 12. Example of a whistle detected on more than one EAR2, but with only a portion of the contour evident in some of the recordings.

ROCCA classification results for all whistles measured on all EARs are provided in **Table 6**, along with the number of whistles analyzed from each EAR2 recording. For 10 of the 20 encounters (50 percent), ROCCA classified the encounter as a different species based on whistles recorded at different depths. Usually, one out of the four EAR2s had a different classification result. Four of the 10 encounters (40 percent) that were classified differently on different EAR2s had significant differences in whistle variables and six (60 percent) did not. Of the 10 encounters that were classified as the same species on all four EAR2s, four (40 percent) had significant differences in whistle variables among EAR2s. When classification results were different on one or more of the EAR2s, the percent of tree votes was often similar for the two species that the encounter was classified as on the different EAR2s (**Table 7**). For example, encounter 12 was classified as spinner dolphin on EAR2-1, and EAR2-4 and as bottlenose dolphin on EAR2-2 and EAR2-3. The percentage of trees that classified the encounter as spinner dolphin was less than 1 percent greater than the percentage of trees that classified the encounter as bottlenose dolphin for EAR2-1 and EAR2-4, and the same was true for bottlenose dolphin vs. spinner dolphin on EAR2-2 and EAR2-3.

Thirteen of the 20 (65 percent) encounters included a sufficient number of whistles that were detected on all four EAR2s (at least 10 whistles). Variables from only the whistles that occurred on all four EAR2s were compared statistically to determine if there were any significant differences in the characteristics of the contours. Significant differences were found in only one encounter (8 percent) (encounter 22) and for only one variable. Minimum frequency was significantly different between EAR2-1 and EAR2-3 (p=0.04). However, this difference did not affect classification success as all four EARs classified the encounter as the same species. Classification results for this dataset are provided in **Table 8**. Four of the encounters (31 percent) were classified as one species for three of the EAR2s and as a different species on the fourth EAR2. The EAR2 that differed was not consistent among encounters. For all four of these encounters, the percent of trees votes was similar for the two species in question (**Table 9**). The remaining nine encounters (69 percent) were classified as the same species on all four EAR2s.

Table 8. Species classification results by encounter for whistles appearing on all four EAR2s. Number of whistles included in the analysis is in parentheses.

Encounter	Classified as				Number of whistles per EAR2
	EAR2-1	EAR2-2	EAR2-3	EAR2-4	
2	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	38
3	Rough-toothed dolphin	Rough-toothed dolphin	Rough-toothed dolphin	Rough-toothed dolphin	35
12	Spinner dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	29
13	Spinner dolphin	Spinner dolphin	Spinner dolphin	Spinner dolphin	16
14	Pilot whale	Pilot whale	Rough-toothed dolphin	Pilot whale	42
16	Pilot whale	Pilot whale	Pilot whale	Pilot whale	32
18	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	36
19	Spinner dolphin	Spinner dolphin	Spinner dolphin	Bottlenose dolphin	15
21	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	25
22	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	38
25	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	30
27	Spinner dolphin	Spinner dolphin	Spinner dolphin	Bottlenose dolphin	28
28	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	Bottlenose dolphin	32

Table 9. Percentage of trees that voted for each species, based on whistles that were detected on all four EAR2s.

Encounter Number	EAR2 Number	Number of Whistles	Percent of tree votes					
			Pilot whale	False killer whale	Spotted dolphin	Rough-toothed dolphin	Spinner dolphin	Bottlenose dolphin
2	EAR2-1	38	0.3%	0.5%	25.5%	1.8%	34.9%	36.9%
	EAR2-2	38	0.2%	0.4%	25.5%	2.0%	33.7%	38.2%
	EAR2-3	38	0.3%	0.3%	25.1%	2.0%	32.1%	40.2%
	EAR2-4	38	0.3%	0.6%	27.6%	1.8%	32.3%	37.5%
3	EAR2-1	35	14.8%	16.0%	9.0%	26.4%	12.8%	21.1%
	EAR2-2	35	14.6%	15.6%	7.5%	29.8%	11.1%	21.4%
	EAR2-3	35	15.9%	14.6%	8.4%	26.6%	12.6%	21.9%
	EAR2-4	35	13.8%	17.4%	8.4%	28.7%	10.8%	20.8%
12	EAR2-1	29	0.0%	0.0%	18.5%	0.0%	40.9%	40.6%
	EAR2-2	29	0.0%	0.0%	19.4%	0.0%	39.2%	41.4%
	EAR2-3	29	0.0%	0.0%	18.0%	0.0%	40.1%	41.9%
	EAR2-4	29	0.0%	0.0%	19.1%	0.0%	39.1%	41.8%
13	EAR2-1	16	0.9%	0.8%	11.6%	4.6%	58.0%	24.2%
	EAR2-2	16	0.0%	0.0%	12.6%	0.0%	57.9%	29.5%
	EAR2-3	16	0.5%	0.7%	11.7%	5.0%	54.2%	27.8%
	EAR2-4	16	0.9%	0.6%	11.1%	4.8%	54.6%	28.1%
14	EAR2-1	42	29.2%	16.0%	5.5%	28.6%	12.2%	8.5%
	EAR2-2	42	30.9%	14.1%	6.0%	28.8%	10.4%	9.8%
	EAR2-3	42	30.2%	13.4%	6.3%	30.3%	10.7%	9.1%
	EAR2-4	42	30.2%	14.4%	5.7%	29.3%	10.8%	9.7%
16	EAR2-1	32	49.4%	36.0%	1.8%	8.4%	1.5%	3.0%
	EAR2-2	32	46.6%	37.2%	1.6%	10.0%	1.4%	3.2%
	EAR2-3	32	47.3%	36.9%	1.6%	9.5%	1.4%	3.3%
	EAR2-4	32	45.9%	37.4%	1.7%	10.4%	1.4%	3.2%
18	EAR2-1	36	0.4%	0.7%	22.6%	1.7%	30.0%	44.7%
	EAR2-2	36	0.5%	0.9%	21.4%	1.4%	29.2%	46.7%
	EAR2-3	36	0.3%	0.6%	22.9%	1.9%	28.2%	46.1%
	EAR2-4	36	0.4%	0.9%	21.4%	1.5%	31.4%	44.4%
19	EAR2-1	15	0.0%	0.0%	24.4%	0.0%	40.9%	34.7%
	EAR2-2	15	0.0%	0.0%	24.0%	0.0%	41.7%	34.3%
	EAR2-3	15	0.0%	0.0%	24.0%	0.0%	40.8%	35.2%
	EAR2-4	15	0.0%	0.0%	25.3%	0.0%	36.9%	37.9%
21	EAR2-1	25	0.0%	0.0%	28.1%	0.0%	32.1%	39.8%
	EAR2-2	25	0.0%	0.0%	32.5%	0.0%	26.0%	41.5%
	EAR2-3	25	0.0%	0.0%	32.1%	0.0%	27.1%	40.8%
	EAR2-4	25	0.0%	0.0%	31.9%	0.0%	27.8%	40.4%
22	EAR2-1	38	1.9%	1.6%	20.9%	4.4%	31.6%	39.6%
	EAR2-2	38	1.4%	1.7%	23.5%	4.7%	26.9%	41.7%
	EAR2-3	38	1.6%	1.7%	23.4%	4.7%	27.4%	41.3%

Encounter Number	EAR2 Number	Number of Whistles	Percent of tree votes					
			Pilot whale	False killer whale	Spotted dolphin	Rough-toothed dolphin	Spinner dolphin	Bottlenose dolphin
	EAR2-4	38	2.5%	2.0%	23.0%	6.1%	24.9%	41.5%
25	EAR2-1	30	2.1%	2.1%	27.4%	5.7%	27.7%	35.0%
	EAR2-2	30	1.3%	1.7%	27.5%	3.7%	29.3%	36.6%
	EAR2-3	30	0.7%	1.3%	30.4%	4.6%	27.5%	35.4%
	EAR2-4	30	0.8%	1.2%	31.0%	4.7%	29.6%	32.8%
27	EAR2-1	28	1.1%	2.8%	14.0%	3.2%	40.3%	38.6%
	EAR2-2	28	1.0%	2.2%	14.7%	0.3%	41.2%	40.5%
	EAR2-3	28	1.8%	2.6%	14.7%	2.7%	39.3%	38.9%
	EAR2-4	28	1.2%	0.9%	15.3%	5.0%	37.5%	40.0%
28	EAR2-1	32	2.8%	4.0%	20.6%	5.7%	29.1%	37.7%
	EAR2-2	32	2.7%	3.5%	19.3%	6.3%	27.3%	40.9%
	EAR2-3	32	1.8%	2.3%	22.8%	5.3%	26.7%	41.1%
	EAR2-4	32	1.1%	2.6%	22.5%	2.6%	28.4%	42.8%

4.4.2 Upcoming analyses

The analyses presented for the Lanai EAR2 dataset are currently being performed on the Kona and San Diego EAR2 and microMARS datasets. Differences in whistle variables and classification results at different depths will likely be affected by the position of the animals relative to the hydrophone array. Upcoming analyses include localizing whistles in order to examine these relationships, as well as performing sound propagation analyses to investigate further the causes of the observed differences and similarities in relation to recording depths. The analyses will be performed for the Lanai, Kona and San Diego datasets.

5. Discussion

The two sets of analyses performed on the Lanai EAR2 data address two facets of the effect of recording depth on classifier performance. The first analysis, which includes all whistles recorded on each EAR2, provides information on the set of whistles available to be analyzed at different depths, and whether the number and composition of whistles available affects classification results. The second analysis, which includes only whistles recorded on all four EAR2s, provides insight to the effect of differences in whistle structure recorded at different depths on classification results.

Whistles were chosen for analysis by randomly selecting 25 whistles from the shallowest EAR2 and 25 whistles from the deepest EAR2. This approach was taken in order to avoid biasing the sample to whistles that were only recorded on either the shallower or the deeper EAR2s. Not all of the selected whistles were detected on all four EAR2s (**Figure 11**). For over half of the encounters, 50 percent or more of the whistles appeared on only one EAR2. Because of this, a different set of whistles was available for classification analysis on each EAR2. In addition, even when the same whistle was detected on multiple EAR2s, the entire contour was not always present on every EAR2 (**Figure 12**). The different whistles, coupled with the fact that entire whistle contours were not always detected, led to significant differences in whistle variables among EAR2s for 8 out of the 20 encounters (40 percent) and different classification results among EAR2s for 10 out of the 20 encounters (50 percent).

When the same set of whistles was used to classify encounters on all four EAR2s, the majority of encounters were classified as the same species on all four EAR2s. This was not surprising, given that significant differences in whistle variables were found for only one encounter and only one variable (minimum frequency). The lack of significant differences in the variables included in the analysis suggests that, although entire whistle contours were not always detected on every EAR2, these incomplete whistles were less common than complete whistles and that whistle structure generally remained constant with recording depth. Therefore, the significant differences observed when including all whistles in the comparison were likely due mainly to different whistles being available on each EAR2 and not due to the same whistles appearing differently on EAR2 recordings from different depths. However, for four of the encounters (31 percent), the classification result was different for one of the four EAR2s, despite the lack of significant differences in whistle variables. For example, encounter 12 was classified as bottlenose dolphin on EAR2-2, EAR2-3 and EAR2-4, but was classified as spinner dolphin on EAR2-1. In each of these encounters, the percentage of trees that voted for the two species that the encounters were classified as were very similar (**Table 9**). In cases like this, where the percentage of tree votes are similar for two species, even minor differences in whistle variables can lead to different species classifications on different EAR2s.

The fact that many whistles were not detected on all four EAR2s was likely due to the position of the vocalizing animals relative to the EAR2 array. In most cases, the greatest number of whistles was detected on the deepest EAR2, which was somewhat surprising, as it was assumed that vocalizing dolphins would be closer to the surface and therefore closest to the shallowest EAR2. Upcoming analyses will include localizing vocalizing animals in order to examine these relationships, as well as performing sound propagation analysis to investigate further causes of the observed differences and similarities among depths. The analyses described here will also be performed for the Kona and San Diego EAR2 and microMARS array data, as well as for the microMARS data collected from the Navy dolphin. These data will provide increased sample sizes to further explore the trends observed in the Lanai EAR2 data and deeper insight to species-specific differences in the effects of depth on whistle structure, as the microMARS data have concurrent visual observations. Results may be different for the microMARS data, as those hydrophones were closer to the surface and likely will cross the thermocline, which could have a significant impact on whistle propagation and differential sound reception by each recorder in the array.

6. Acknowledgements

We would like to thank Gabriela Alongi, Mark Cotter, Liz Ferguson, Megan McElligott, Tom Norris, Tina Yack, and Eden Zang for their hard work preparing equipment and for their help in the field. We are grateful to our boat captains, Jamie Thinnes, Pete McCormick and Jeff Stock, for their dedication, patience and skill in getting us to where we needed to be. We thank Dorian Houser, Elaine Allen, Jefferson Filamor, Catlin Rixon and Nicole Simon at the National Marine Mammal Foundation for their efforts in training the Navy dolphin and for all of their help in the field trials. We would also like to thank Elizabeth Henderson and Jay Barlow for logistical assistance. We thank US Fleet Forces Command under the US Navy's Marine Species Monitoring Program, ONR and LMR for funding this project. Project management was provided by NAVFAC Atlantic (NAVFAC LANT). We are especially grateful to Dan Engelhaupt and Michael Richlen (HDR), Joel Bell (NAVFAC LANT), Mike Weise (ONR), Anu Kumar (LMR) and Mandy Shoemaker (LMR) for their assistance.

7. References

Au, W.W.L., and M. Hastings. 2008. Principles of marine bioacoustics. Springer, NY. 679pp.

Au, W.W.L., B. Branstetter, P.W. Moore, and J.J. Finneran. 2012a. The biosonar field around an Atlantic bottlenose dolphin (*Tursiops truncatus*). *Journal of the Acoustical Society of America* 131: 569–576.

Au, W.W.L., B. Branstetter, P.W. Moore and J.J. Finneran. 2012b. Dolphin biosonar signals measured at extreme off-axis angles: Insights to sound propagation in the head. *Journal of the Acoustical Society of America* 132: 1199–1206.

Barkley, Y., J.N. Oswald, J.V. Carretta, S. Rankin, A. Rudd, and M.O. Lammers. 2011. Comparison of real-time and post-cruise acoustic species identification of dolphin whistles using ROCCA (Real-time Odontocete Call Classification Algorithm). NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-473.National Marine Fisheries Service, La Jolla, CA. 29 pp.

Baumann-Pickering, S, S. Wiggins, J.A. Hildebrand, M.A. Roch and H.U. Schnitzer. 2010. Discriminating features of echolocation clicks of melon-headed whales (*Peponocephala electra*), bottlenose dolphins (*Tursiops truncatus*), and Gray's spinner dolphins (*Stenella longirostris longirostris*). *Journal of the Acoustical Society of America* 128: 2212-2224.

Bauman-Pickering, S., Simonis, A.E., Oleson, E.M., Rankin, S., Baird, R., Roch, M.A., Wiggins, S.M., and J.A. Hildebrand. 2011. False killer whale and short-finned pilot whale acoustic occurrences around the Hawaiian Islands. Proceedings of 19th Biennial Conference on the Biology of Marine Mammals, November, Tampa, FL.

Breiman, L. (2001). Random forests. *Machine Learning* 45: 5-32.

Gillespie, D., J. Gordon, R. McHugh, D. McLaren, D.K. Mellinger, P. Redmond, A. Thode, P. Trinder, and D. Xiao. 2008. PAMGUARD: Semiautomated, open-source software for real-time acoustic detection and localization of cetaceans. *Proceedings of the Institute of Acoustics* 30 (Part 5): 9 pp.

Jarvis, S., N. DiMarzio, R. Morrissey and D. Moretti. 2008. A novel multi-class support vector machine classifier for automated classification of beaked whales and other small odonotocetes. *Canadian Acoustics* 36: 34–40.

Lammers, M.O. and M. Castellote. 2009. The beluga whale produces two pulses to form its sonar signal. *Biology Letters* 5: 297-301.

Oswald, J.N. 2013. *Development of a Classifier for the Acoustic Identification of Delphinid Species in the Northwest Atlantic Ocean.* Final Report. Submitted to HDR Environmental, Operations and Construction, Inc. Norfolk, Virginia under Contract No. CON005-4394-009, Subproject 164744, Task Order 003, Agreement # 105067. Prepared by Bio-Waves, Inc., Encinitas, California.

Oswald, J.N., S. Rankin, J. Barlow, M. Oswald and M.O. Lammers. 2013. Real-time Call Classification Algorithm (ROCCA): software for species identification of delphinid whistles. In: Detection, Classification and Localization of Marine Mammals using Passive Acoustics, 2003-2013: 10 years of International Research, DIRAC NGO, Paris, France, pp. 245-266.

Rendell, L.E., J.N. Matthews, A. Gill, J.C.D. Gordon, and D.W. Macdonald. 1999. Quantitative analysis of tonal calls from five odontocete species, examining inter-specific and intra-specific variation. *Journal of Zoology* 249:403-410.

Ridgway, S.A., D.A. Carder, T. Kamolnick, R.R. Smith, C.E. Schlundt, and W.R. Elsberry. 2001. Hearing and whistling in the deep sea: depth influences whistle spectra but does not attenuate hearing by white whales (*Delphinapterus leucas*) (Odontoceti, Cetacea). *Journal of Experimental Biology* 204:3829-3841.

Roch M.A., H. Klinck, S. Baumann-Pickering, D.K. Mellinger, S. Qui, M.S. Soldevilla and J.A. Hildebrand. 2011. Classification of echolocation clicks from odontocetes in the Southern California Bight. *Journal of the Acoustical Society of America* 129: 467–475.

Soldevilla, M.S., E.E. Henderson, G.S. Campbell, S.M. Wiggins, J.A. Hildebrand and M.A. Roch. 2008. Classification of Risso's and Pacific white-sided dolphins using spectral properties of echolocation clicks. *Journal of the Acoustical Society of America* 124: 609–624.

Sturtivant, C., and S. Datta. 1997. Automatic dolphin whistle detection, extraction, encoding, and classification. *Proceedings of the Institute of Acoustics* 19: 259-266.

Wiggins, S. 2007. Triton (Version 1.80) [Acoustic Processing Software]. Scripps Institution of Oceanography, UC San Diego, La Jolla, California. Retrieved August 1, 2011. Available from www.cetus.ucsd.edu.