

ANALYSIS OF ACOUSTIC SIGNALS DETECTED ON AN OCEAN-BOTTOM  
SEISMOGRAPH ARRAY IN THE LAU BASIN, SOUTH PACIFIC OCEAN

A THESIS SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF  
HAWAII AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE  
DEGREE OF

MASTER OF SCIENCE

IN

GEOLOGY & GEOPHYSICS

MAY 2015

By

Dana C. Brodie

Thesis Committee:

Robert A. Dunn, *Chairperson*

L. Neil Frazer

Eva-Marie Nosal

Keywords: baleen whale, Lau basin, bioacoustics

## ACKNOWLEDGEMENTS

I would like to thank the captain, crew, and science parties of the R/V Revelle leg RR0915 and the R/V Kilo Moana leg KM1023. I also thank my committee (Robert Dunn, Neil Frazer, and Eva-Marie Nosal), Fred Duennebier, my family for their continued support, and my fellow graduate students for consistently reminding me of how the whale calls might sound. This project was partially funded by NSF grant OCE042642 to R. Dunn and Fred M. Bullard Endowed Graduate Fellowship to D. Brodie via the Department of Geology and Geophysics, University of Hawaii.

## ABSTRACT

Ocean-bottom seismic networks are often deployed in remote regions to study the structure of the oceanic crust and mantle, and are able to detect both seismic signals and ocean acoustic signals. This thesis examines records from a large ocean-bottom seismic network to decipher the various acoustic signals that were recorded over nearly one year, including baleen whale calls and volcanic landslides. Specifically, ten months of broadband seismic data, recorded on twenty-nine ocean-bottom seismographs located in and adjacent to the Lau Basin, were utilized to identify baleen whale species and to determine species' spatial, temporal, and diel patterns. Probable whale sounds that could not be matched to published spectrograms, as well as non-biologic sounds that are likely of volcanogenic origin, were also recorded. An automatic detection algorithm utilizing empirical orthogonal functions and multinomial logistical regression was developed to detect select sound types from the large data set.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENTS .....</b>	<b>i</b>
<b>ABSTRACT.....</b>	<b>ii</b>
<b>LIST OF TABLES .....</b>	<b>v</b>
<b>LIST OF FIGURES .....</b>	<b>vi</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>viii</b>
<b>CHAPTER 1. INTRODUCTION.....</b>	<b>1</b>
<b>CHAPTER 2. WHALE CALL IDENTIFICATION .....</b>	<b>3</b>
<b>I. Introduction.....</b>	<b>3</b>
<b>II. Methods.....</b>	<b>6</b>
<b>III. Sound Characterization .....</b>	<b>10</b>
<b>IV. Discussion .....</b>	<b>18</b>
<b>V. Conclusions.....</b>	<b>24</b>
<b>CHAPTER 3. AUTOMATIC DETECTION ALGORITHM.....</b>	<b>26</b>
<b>I. Introduction.....</b>	<b>26</b>
<b>II. Data and Methods .....</b>	<b>26</b>
<b>III. Algorithm Results.....</b>	<b>34</b>
<b>IV. Discussion .....</b>	<b>50</b>

V. Conclusions.....	55
<b>CHAPTER 4. CONCLUSIONS.....</b>	<b>56</b>
<b>Appendix A. Procedure for downloading and preparing data from IRIS .....</b>	<b>58</b>
<b>Appendix B. Instrument Responses .....</b>	<b>60</b>
<b>Appendix C. Background Noise .....</b>	<b>67</b>
<b>Appendix D. OBS name and number correlation.....</b>	<b>81</b>
<b>Appendix E. Algorithm Training Set .....</b>	<b>82</b>
<b>Appendix F. Detection and classification algorithm.....</b>	<b>85</b>
<b>REFERENCES.....</b>	<b>89</b>

## LIST OF TABLES

<b>Table 2-1.</b> Characteristics of baleen whale calls and unknown sounds.....	8
<b>Table B-1.</b> Instrument response Stages 1 and 2 for OBS instrument #14.....	63
<b>Table B-2.</b> Instrument response Stage 3 coefficients for OBS instrument #14.....	64
<b>Table C-1.</b> Instrument response gains .....	70
<b>Table C-2.</b> Normalized median power between 15-40 Hz without instrument response gains ...	71
<b>Table C-3.</b> Normalized p-wave amplitudes compared to background noise levels .....	72

## LIST OF FIGURES

<b>Figure 2-1.</b> Locations of the ocean-bottom seismographs .....	5
<b>Figure 2-2.</b> Seasonal occurrence of whale calls and other sounds .....	9
<b>Figure 2-3.</b> Spectrograms of whale calls attributed to fin, Bryde's, and blue whales .....	12
<b>Figure 2-4.</b> Spectrograms of unidentified sounds .....	15
<b>Figure 2-5.</b> Spectrograms of possibly biologic and non-biologic sounds .....	17
<b>Figure 3-1.</b> Location of the ocean-bottom seismographs .....	27
<b>Figure 3-2.</b> Spectrograms attributed to fin, Bryde's, and Antarctic blue whales .....	28
<b>Figure 3-3.</b> Detection error tradeoff graphs .....	32
<b>Figure 3-4.</b> Receiver-operator curves .....	33
<b>Figure 3-5.</b> Fin whale call detections for varying probability thresholds .....	35
<b>Figure 3-6.</b> Bryde's whale call detections for varying probability thresholds .....	36
<b>Figure 3-7.</b> Antarctic blue whale call detections for varying probability thresholds .....	37
<b>Figure 3-8.</b> Fin whale call detections throughout the recording period .....	39
<b>Figure 3-9.</b> Fin whale call diel pattern .....	40
<b>Figure 3-10.</b> Fin whale call detections by regions of the Lau basin .....	41
<b>Figure 3-11.</b> Bryde's whale call detections throughout the recording period .....	43

<b>Figure 3-12.</b> Bryde’s whale call diel pattern .....	44
<b>Figure 3-13.</b> Bryde’s whale call detections by regions of the Lau basin .....	45
<b>Figure 3-14.</b> Antarctic blue whale call detections throughout the recording period .....	47
<b>Figure 3-15.</b> Antarctic blue whale call diel pattern .....	48
<b>Figure 3-16.</b> Antarctic blue whale call detections by regions of the Lau basin .....	49
<b>Figure B-1.</b> Separate amplitude and phase response .....	65
<b>Figure B-2.</b> Combined amplitude and phase response .....	66
<b>Figure C-1.</b> Power versus frequency for instrument C17W .....	73
<b>Figure C-2.</b> Power versus frequency for instrument C11W .....	74
<b>Figure C-3.</b> Power versus frequency for instrument C17W showing two-week intervals .....	75
<b>Figure C-4.</b> Power versus frequency for instrument C11W showing two-week intervals .....	76
<b>Figure C-5.</b> $P_{\text{median}}$ of each two-week period for instrument C17W.....	77
<b>Figure C-6.</b> $P_{\text{median}}$ of each two-week period for instrument C11W .....	78
<b>Figure C-7.</b> $P_{\text{median}}$ of each two-week period for instrument B01W .....	79
<b>Figure C-8.</b> Background noise patchwork plot .....	80

## LIST OF ABBREVIATIONS

dB	decibel
DET	detection error threshold
EOF	empirical orthogonal function
FFT	Fast Fourier Transform
FNR	false negative rate
FP	false positive
FPR	false positive rate
Hz	hertz
kHz	kilohertz
km	kilometer
m	meter
NMP	normalized mean power
NPWA	normalized p-wave amplitudes
OBS	ocean-bottom seismograph
OBSIP	United States Ocean Bottom Seismic Instrumentation Pool
PC	principal component
ROC	receiver-operator curve

s	second
SEED	Standard for the Exchange of Earthquake Data
SOFAR	sound fixing and ranging
TPR	true positive rate
WHOI	Woods Hole Oceanographic Institution

## CHAPTER 1. INTRODUCTION

Ocean-bottom seismic networks are often deployed in remote regions to study the structure of the oceanic crust and mantle. As technology improves, these networks are able to record for increasing periods of time, up to a year or more, or for several years if they are able to be serviced. In addition to recording valuable seismic data, these seismic networks are able to record information useful for other scientists. For example, ocean-bottom seismographs (OBS) are able to record volcanic and iceberg tremors and biological calls from whales. By studying the other information recorded on the OBS, we can gain valuable information in scientific fields other than seismology and utilize the data sets more completely.

Baleen whales are present throughout the world's oceans, and each of the approximately 14 major species produces a variety of unique acoustic calls. While higher frequency calls ( $>1\text{kHz}$ ) are common among whales, only baleen species produce low frequency acoustic calls ( $<100\text{ Hz}$ ). Since duration, frequency, structure, and repeat intervals of calls are similar, hydroacoustic detection and monitoring of whale calls can aid researchers in identifying species, subpopulations, and their respective migratory patterns. Passive hydroacoustic monitoring using moored sensors, including OBS has become an established method for collecting whale call data and monitoring whale populations in sizable, remote regions. OBS networks are often deployed over broad areas for six months or more at a time and thus provide an opportunity for long-term, large-scale baleen whale acoustical studies.

This thesis presents two studies: first, a preliminary visual examination of ten months of broadband seismic data recorded on six ocean-bottom seismographs situated in the central Lau basin in the southwestern tropical Pacific Ocean between Fiji and Tonga; and second, the results of an automatic detection algorithm developed for this particular data set using multinomial logistic regression applied to all instruments in the seismic array. Although whales are known to

seasonally reside in the broader regions around the study area, no survey of baleen whale acoustics has been published for the Lau basin with the exception of high-frequency humpback whale studies near the islands of Fiji and Tonga. The purpose of this study is to characterize the different biological, as well as some non-biological, sounds recorded on a data collected by a large marine seismic experiment. Then, using the identified calls, an algorithm was developed to automatically select whale calls from the extensive data set to determine seasonality trends.

## CHAPTER 2. WHALE CALL IDENTIFICATION

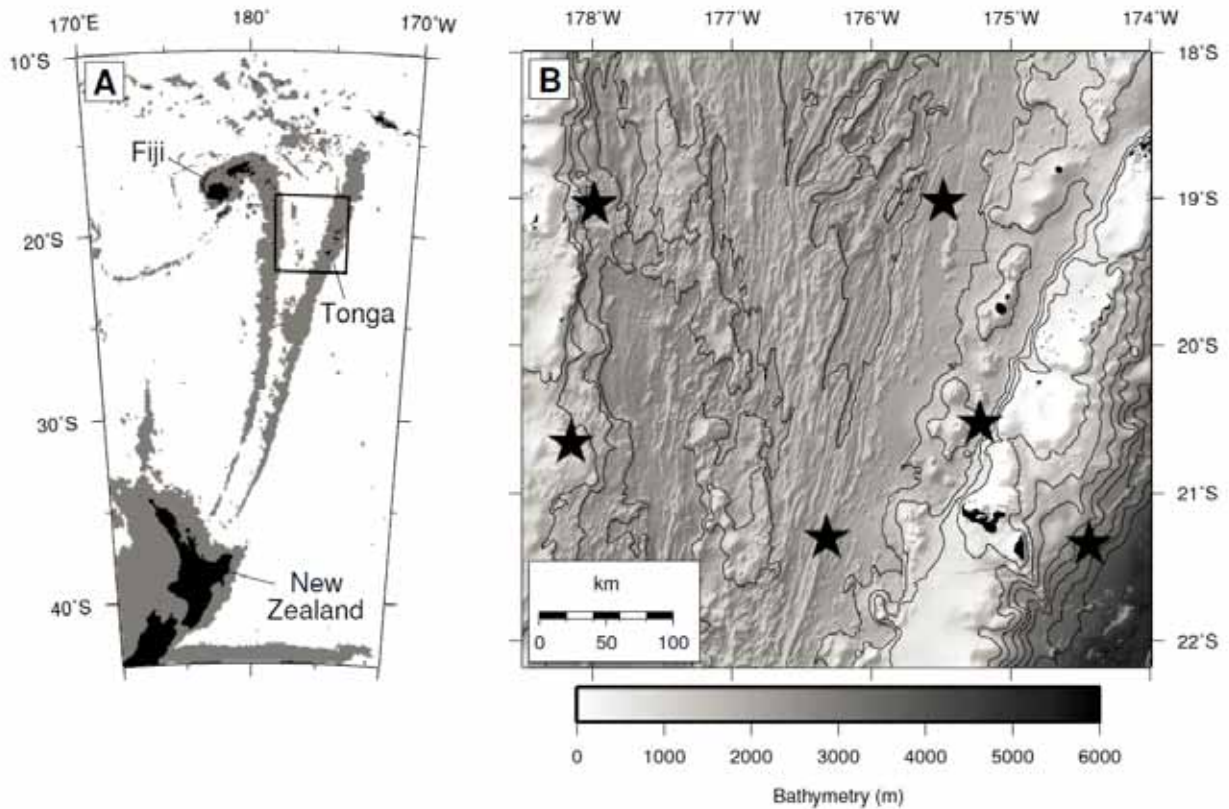
### I. Introduction

Baleen whales are present throughout the world's oceans, and each of the approximately 14 major species produces a variety of unique acoustic calls. While higher frequency calls (>1 kHz) are common among whales, only baleen species produce low frequency (<100 Hz) acoustic calls. Although call types vary in duration, frequency, structure, and repeat interval, they are similar within each species. For example, fin whales (*Balaenoptera physalus*) are widely observed to produce repeated sequences of ~1-2 second duration 20-Hz calls (e.g. Watkins, 1981), whereas blue whales (*B. musculus*) are known to produce repeated sequences of amplitude- and frequency-modulated calls with longer durations (e.g. McDonald et al., 2006). Some species' calls show geographic and subpopulation variations (e.g. Thompson et al., 1992; Oleson et al., 2003; McDonald et al., 2006).

Acoustic detection and monitoring of whale calls can aid researchers in identifying species, subpopulations, and their migratory patterns. To achieve these goals, complete call repertoires of individual species must be known; otherwise, identifying species by their calls alone is problematic. While past acoustic studies have furthered the understanding of baleen whale call patterns, they have focused on limited portions of the world's oceans, leaving many regions understudied. These gaps in coverage not only obscure population counts and migratory behavior, but also leave us with an incomplete catalog of whale calls and calling behavior.

Passive acoustic monitoring using moored sensors has become an established method for collecting whale call data and monitoring whale populations in sizable, remote areas (e.g. Širović et al., 2004; Heimlich et al., 2005; Stafford et al., 2007). One such methodology employs ocean-bottom seismographs (OBSs), which can record low frequency calls emitted by baleen whales at distances of up to tens of kilometers (e.g. Dunn and Hernandez, 2009). OBS networks are often deployed over broad areas for six months or more at a time and thus provide a unique opportunity for long-term, large-scale baleen whale acoustical studies.

This chapter presents a preliminary study of ten months of broadband seismic data recorded on OBSs situated within or adjacent to the central Lau basin. The Lau basin is located in the southwestern tropical Pacific Ocean between Fiji and Tonga, approximately 1,700 km northeast of New Zealand (**Figure 2-1a**). Five OBSs were located within the basin and one OBS was located east of Tonga, outside the basin (**Figure 2-1b**). Although whales are known to seasonally reside in the broader regions around the study area (Ohsumi, 1978; 1979), no survey of baleen whale acoustics has been previously published for the Lau basin, with the exception of studies of high-frequency humpback whale (*Megaptera novaeangliae*) near the islands of Fiji and Tonga (Dawbin, 1966; Helweg et al., 1998; Garland et al., 2013). The purpose of this preliminary study is to examine a representative subset of data from a large marine seismic experiment and characterize the different biological, as well as some non-biological, sounds recorded. This study may then act as a springboard for future investigations (using the full database) of individual whale species, their geographic distribution, and their seasonality within and around the basin. Spectrograms created using the Lau basin data reveal four known whale calls (fin, Bryde's, and two types of blue whale calls), some sounds that are similar to other published calls but lack sufficient information for unique species identification, a few possible whale calls that, to the authors' knowledge, have not yet appeared in the literature, and some non-biological sounds. The spectrograms' information supplements findings from previous studies to form a more complete picture of whale call characteristics and provides new information on baleen whale presence in this remote, effectively unstudied, location.



**Figure 2-1.** Location of the ocean-bottom seismographs used to detect baleen whale calls and other sounds. The instruments were located in the western tropical Pacific Ocean between Tonga and Fiji within the Lau basin (one instrument was located eastward of Tonga). In panel (A) water depths shallower than 2000 meters are shaded gray and land is shaded black; the experiment area, shown in panel (B), is outlined with a black box. (B) Shaded bathymetric map of the study area with land shaded black; the contour interval is 500 meters. The black star symbols show the locations of the six ocean-bottom seismographs used in the analysis. The instruments were deployed during the period Nov. 2009 to Dec. 2010.

## II. Methods

A yearlong deployment of broadband OBSs was carried out across the central Lau basin from late November 2009 through early December 2010 as part of a passive-recording marine seismic experiment (Wei et al., 2012). The instruments resided on the seafloor and provided continuous data records throughout the deployment, but many stopped recording after the first ~10 months due to battery exhaustion. The data examined in this study were recorded on OBSs designed and built by the Woods Hole Oceanographic Institution (WHOI) for the U.S. Ocean Bottom Seismic Instrumentation Pool (OBSIP; <http://www.obsip.org>), and were equipped with a Guralp CMG-3T (0.01 - 50 Hz response) 3-component seismometer and a digital pressure gauge. Only data from the vertical component of the seismometer were used for analysis. The horizontal channels of OBSs tend to have much higher noise levels than the vertical channel, and data from the digital pressure gauge were recorded at half the sampling rate and thus with a much more restricted frequency range. The vertical channel data were recorded at 100 samples per second, providing a theoretical maximum frequency of recording of 50 Hz. Internal to the instruments, an anti-aliasing filter was applied to the raw data stream before digital sampling, resulted in a useable frequency range between <1 and 45 Hz.

To provide broad coverage across the study area, six OBSs (**Figure 2-1b**) located throughout the study area were selected from 29 possible WHOI instruments (20%) and visually examined for whale calls. We chose to visually examine the data because automatic detection routines are best used in situations where species types are already known. On the basis of our visual assessment of Lau basin species, the observed calls can later be used to target individual species within the full dataset using automatic-detection algorithms.

The data are archived as Standard for the Exchange of Earthquake Data (SEED) format. For processing, SEED data were cut into day-length (24 hour) files and converted to Matlab® format (The Mathworks Inc., [www.mathworks.com](http://www.mathworks.com)). Then, the time series of each day file was demeaned, detrended, and high-pass filtered with a 4-pole, 6-Hz corner frequency Chebyshev filter. For each day and station, the data were divided into six-minute periods, and then the time-domain waveform and spectrogram of every fourth six-minute period (25% of the data) were examined for biological sounds. Additional six-minute periods were examined when faint or

unknown signals were observed. While examining the data, records were kept of the types of sounds observed each day. Frequency and temporal parameters of each sound type (**Table 2-1**) were measured from the spectrograms of each available instrument throughout the year; though not all sounds were identified on all stations. Spectrograms were created using a 256-point Hamming window with 90% overlap and a 256-point Fast Fourier Transform (FFT). Frequency ranges were measured for any upswept, down-swept, or broadband pulses, and the center frequency was measured for tones.

Sounds detected in the Lau basin data were identified by matching their acoustic characteristics and spectrograms to published accounts identified as whale calls or non-biological sounds. Several calls or sounds whose spectrograms could not be identified from the literature were also observed in the data. Of those sounds, we present those that were recorded on multiple instruments at different times and/or have “biological” characteristics including band-limited frequency range, repeating patterns, and call duration similar to known whale calls. The numerous earthquake recordings that were recorded in this tectonically active basin were ignored in this study. Although we did not examine all of the available data, we feel that the instrument distribution and analysis are sufficient to detect each vocalizing species traveling through the study area. No visual sightings of whales were available to correlate with the acoustic observations.

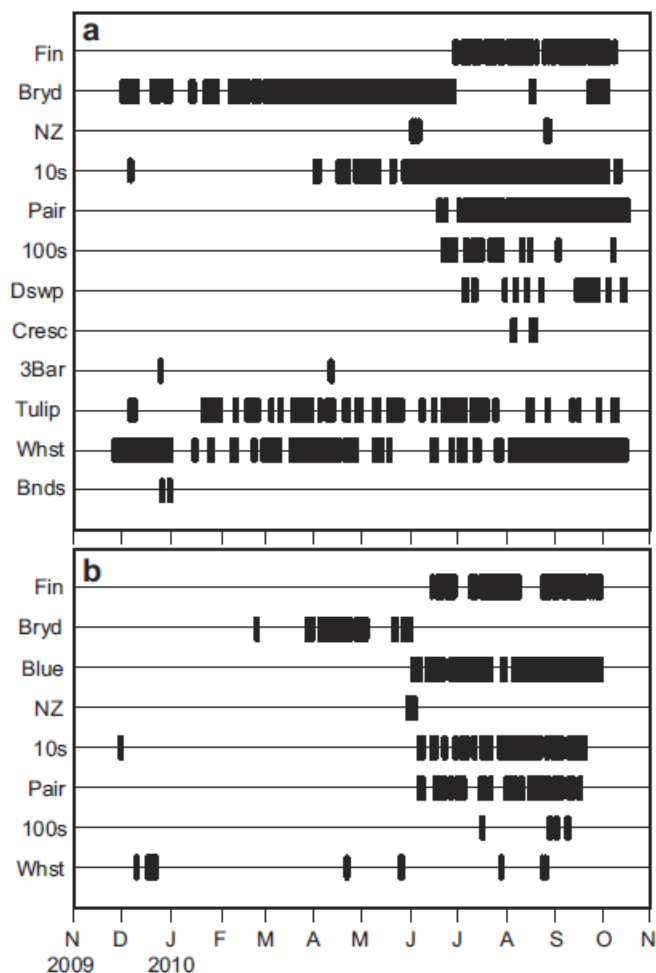
The temporal occurrence of the calls and other sound observations are summarized in **Figure 2-2**. We did not attempt to track or count individuals in this study, though it was clear that in some instances more than one vocalizing whale was present at a given time (overlapping call sequences being an obvious indicator). Nevertheless, most of the vocalizations could have come from just a few individuals, and one should not assume that large numbers of animals were present.

Table 2-1. Characteristics of baleen whale calls and unknown sounds

Call type	N measured	Component	Frequency (Hz)	Duration (s)	Intercall/pulse interval (s)
Fin	8	A	13.8 - 23.0* (11.3 - 27.9)†	1.5 (0.8 - 1.8)	A-B: 10.8 (8.0 - 13.5)
		B	14.2 - 34.4 (12.9 - 36.5)	1.6 (1.3 - 2.3)	B-A: 20.1 (18.4 - 22.9)
Bryde's	13	Burst	26.1 - 42.3 (23.8 - 45+)	2.5 (1.5 - 3.1)	296 (202 - 385)
		Tone	24.1 (21.7 - 25.2)	3.2 (1.8 - 4.1)	
Antarctic Blue	8	1st Tone	26.3 (26.1 - 26.5)	10.1 (9.6 - 11.4)	66.5 (64.4 - 68.6)
		2nd Tone	18.6 (18.5 - 18.7)	8.6 (6.5 - 10.9)	
New Zealand Blue	5	1	24.7 (24.6 - 24.8)	15.8 (12.2 - 17.1)	86.8 (83.7 - 90.4)
		2	25.3 (25.0 - 25.6)	4.3 (3.4 - 5.5)	
		3	18.1	8.2 (6.5 - 9.9)	
Upsweep	15	single	34.1 - 45+ (31.6 - 45+)	1.2 (0.8 - 1.3)	10.2 (9.1 - 11.4)
Paired Upsweep	7	double	34.4 - 45+ (32.4 - 45+)	3.4 (3.1 - 3.9)	9.9 (8.3 - 10.6)
100s Upsweep	7	single	33.0 - 45+ (31.2 - 45+)	1.1 (0.7 - 1.3)	97.3 (86.8 - 110.0)
Down-sweep	10	single	36.7 - 39.7 (33.7 - 43.4)	2.5 (1.0 - 5.0)	irregular
Crescent	7	Top Peak	33.0 (31.8 - 35.1)	7.9 (6.7 - 10.1)	none
	9	Mid Peak	28.0 (26.7 - 29.6)		
		Low Peak	22.4 (21.4 - 23.8)		
Three Bar	2	Top Center	30.7 (30.5 - 30.9)	9.1 (8.1 - 10.1)	none
		Mid Center	20.1 (19.7 - 20.5)		
		Low Center	10.2 (9.9 - 10.5)		
Tulip	16	1	16.9 - 19.9 (12.7 - 24.8)	5.3 (3.9 - 9.5)	none to irregular
		2	9.3 - 16.6 (7.0 - 18.5)		
		3	6.1 - 15.1 (5.1 - 20.3)		
Monochromatic Whistle	-	none	<5 - 45	10 - >1800	none
Harmonic Frequency Bands	2	all	(6.3 - 41.0)	(21.8 - 72.5)	none

\* Mean values and ranges are reported for each component of the sounds described in this study.

† Values in parentheses indicate the total range of the observations.



**Figure 2-2.** Seasonal occurrence of whale calls and other sounds recorded between November 2009 and October 2010. (a) Sounds recorded on the five instruments located within the Lau Basin. (b) Sounds recorded on the single station located just to the east of the basin. Each tick mark indicates that at least one observation of a particular sound was noted for the given day. Abbreviations: Bryd = Bryde’s whale; Blue = Antarctic blue whale; NZ = New Zealand blue whale; 10s = Single upsweep with a 10 s separation; Pair = Doublet upsweep; 100s = Single upsweep with a 100 s separation; Dswp = Down-sweep; Cresc = Crescent; 3 Bar = 3Bar; Whst = Monochromatic whistle; Bnds = Harmonic frequency bands.

### III. Sound Characterization

#### A. Fin whale call

A commonly observed sound pattern in the data is characterized by a two-part (A-B) configuration with a dominant frequency of  $\sim 20$  Hz (**Figure 2-3a**). Part A of the call has an average frequency range of 13.8 to 23.0 Hz, and part B has an average frequency range of  $\sim 14.2$  to 34.4 Hz, with each part of the call lasting an average of  $\sim 1.5$  s (**Table 2-1**). This pattern was detected in repeated sequences lasting 1 to 12 minutes with alternating 10.8 s (A-B) and 20.1 s (B-A) interpulse intervals (**Table 2-1**). Call sequences typically started with part A, but ended on either part A or B. Occasionally either part A or B would be skipped in the middle of a repeating set. Water column multiples of acoustic energy were often observed following both parts of the call. These “doublet” patterns are attributed to fin whales (e.g., Watkins et al., 1987; Thompson et al., 1992); however, other studies have suggested that these calls might be produced by Sei whales (*Balaenotera borealis*; Rankin and Barlow, 2007). These calls were detected on the five instruments within the basin between mid-June and mid-October 2010 (**Figure 2-2a**) and on the single instrument located east of the basin between mid-June and late September 2010 (**Figure 2-2b**).

#### B. Bryde’s whale call

Another commonly observed sound in the data is characterized by a broadband “burst” from 26.1 Hz to at least 45 Hz (maximum recordable frequency) with an average 2.5 s duration, followed by either a flat or slightly down-swept pulse at 24.1 Hz with an average 3.2 s duration (**Figure 2-3b**). Within the burst, some frequencies have higher amplitudes than others, although this varied between calls. An 18 Hz beat, with  $\sim 2$  s duration, occasionally coincided with the initial broadband pulse. Water column multiples were often observed following the call components. These calls form repeated sequences with intercall intervals ranging from 200 to 385 s (296 s average). This sound is attributed to Bryde’s whales (Oleson et al., 2003; Heimlich et al., 2005; McDonald, 2006). Bryde’s whale calls were primarily detected on the five instruments located within the basin between early December 2009 and late June 2010, but were also observed on

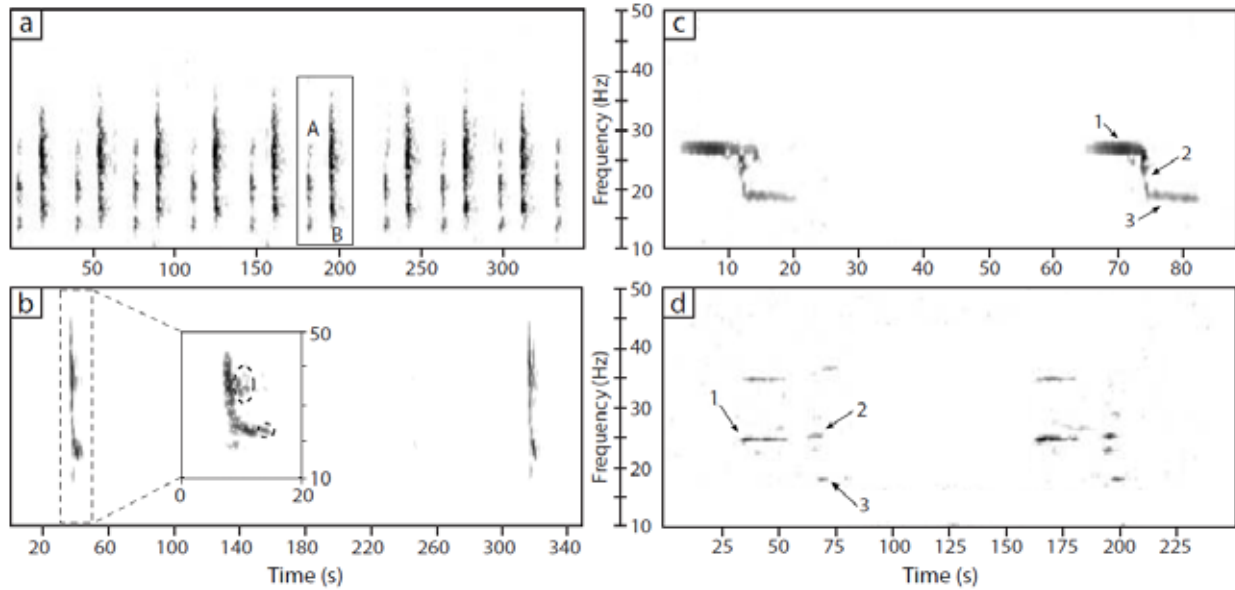
fewer instruments in mid-August 2010, and between late September and early October 2010 (**Figure 2-2a**). Bryde's whale calls were also detected on the instrument located east of the basin between February and May 2010 (**Figure 2-2b**).

### **C. Antarctic blue whale call**

A less commonly observed sound in the data consists of three simple components. A 26.3 Hz tone, lasting an average of 10.1 s, is followed by a 1 s duration, 26 to 18 Hz down-swept component, which is followed by an 8.6 s tone at 18.6 Hz. The entire call pattern lasts 20 to 21 s. Repeating call sequences of 2 to 17 calls exhibit an average intercall interval of 66.5 s (**Figure 2-3c**). A variation on this call consists of just the first component (without the following two components) with the same intercall spacing. This call type is attributed to Antarctic blue whales (e.g., Ljungblad et al., 1998, Rankin et al., 2005). Antarctic blue whale calls were only detected on the instrument located east of the basin between early June and late September 2010 (**Figure 2-2b**).

### **D. New Zealand blue whale call**

Another uncommonly observed, multiple-component sound consists of an initial 24.7 Hz tone of 15.8 s duration (labeled as "1" in **Figure 2-3d**), followed 2 s later by a 25.3 Hz tone lasting 4.3 s (labeled as "2" in the figure), followed by a final 18.1 Hz tone of 8.2 s duration (labeled as "3" in the figure). The second tone is not always present. The main tones in the spectrogram are often coincident with inharmonic overtones about 10 Hz higher, but with lower amplitude. For example, the initial ~25 Hz tone often occurs with a matching ~35 Hz tone of similar duration. The complete call lasts ~52 s and exhibits an average intercall interval of 86.7 s. The call is similar to the "New Zealand blue whale song" described by McDonald (2006) and Miller et al. (2014). In our data, the call was observed on the three southernmost instruments within the basin and on the one instrument located outside of the basin around the beginning of June. The call was observed within the basin again in late August 2010 (**Figure 2-2**).



**Figure 2-3.** Spectrograms of whale calls attributed to fin, Bryde's, and blue whales. The frequency axis is the same for all panels, but the time axes differ. (a) A sequence of fin whale doublet calls (256-point FFT). (b) A pair of Bryde's whale calls (128-point FFT). The inset shows an individual call on a magnified time axis with water column multiples circled. (c) A pair of Antarctic blue whale calls (128-point FFT). (d) A pair of New Zealand blue whale calls (256-point FFT). Spectrogram parameters for all panels: 90% overlap, Hamming window. In this and the following figures, some areas of random background noise were intentionally masked to avoid confusion with the indicated sounds.

### **E. The Upsweep**

The upsweep sound is characterized by an approximately 1 s duration upswept sound from 31 to 45 Hz (**Figure 2-4a**). The upswept sound typically appears in groups of 2 to 8 with an average spacing of 10.2 s between each upsweep. The direct arrival from a single upsweep is sometimes followed by one or two fainter arrivals at times consistent with water column multiples. In addition, two variations of this sound were observed. The first variation consists of a paired version of the single call, with pairs of calls spaced ~2 s apart (**Figure 2-4b**). The paired version typically appears in groups of 2 to 5 with an average spacing of 9.9 s between each upswept pair. The second part of each pair is separated from the first by too much time to be a water column multiple based on an examination of the water depths in the vicinity of the recording stations. The second variation is similar to the single upsweep sound already described, except that the average spacing between each upsweep is much longer at 97.3 s (**Figure 2-4c**). These three types of upswept sounds were detected on all six stations between late May and mid-October 2010. The single pulsed version with 10 s spacing was also observed on three instruments located within the basin between late November and early December 2009 and on one instrument located within the basin from early April to late May 2010 (**Figure 2-2a**).

### **F. The Down-sweep**

Down-swept sounds were observed with an average duration of 2.5 s and an irregular interpulse interval (ranging from 4 to 22 s when appearing in groups of 2 or more; **Figure 2-4d**). These sounds were observed most commonly between 33 and 43 Hz. Some of the down-sweep pulses may begin at a higher frequency outside the recording band, though many clearly do not. Sets of such sounds lasted between ~10 - 500 s. The down-sweeps were seldom observed in the data, but were sporadically detected within the basin between early June and early October 2010 (**Figure 2-2a**).

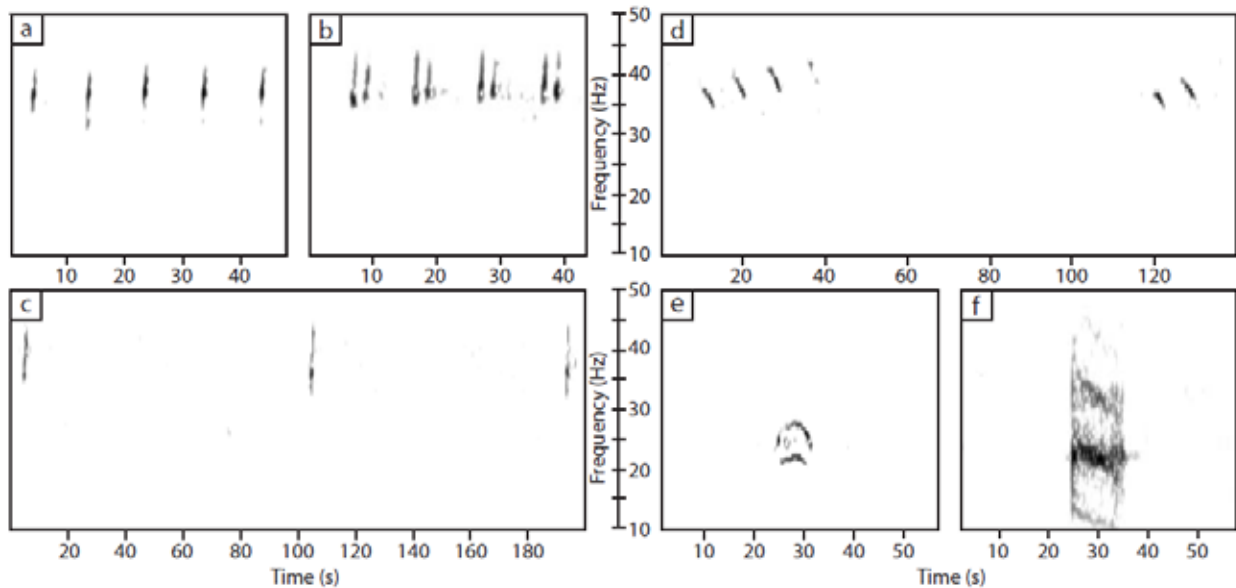
### **G. The Crescents**

The crescent sound (**Figure 2-4e**) consists of two or three crescent-shaped components occurring simultaneously in an inharmonic stack with an average duration of 7.9 s. The lowest frequency component has a peak at 22.4 Hz, and the overtones have peaks at 28.0 Hz and 33.0 Hz, on

average. The 33.0 Hz component was rarely observed. Although this sound did not occur in repeating patterns, multiple instances were typically detected within a few hours of each other. Crescents were recorded on the easternmost instrument located within the basin in early August 2010 (**Figure 2-2a**).

#### ***H. The Three-Bar***

The three-bar sound consists of three simultaneous and roughly parallel down-swept components lasting 9.1 s on average with center frequencies at roughly 10, 20, and 30 Hz (**Figure 2-4f**). The top and bottom components are likely sidebands around a central tone due to the pulsive nature of the amplitude of this sound over time. Unfortunately, the spectrograms of this sound are all low quality, suggesting a distant source, and precise characterization of the sound components was not possible. The three-bar sound did not occur in regular repeating patterns, and was recorded on the northernmost instrument within the basin in late December 2009 and again in mid-April 2010 (**Figure 2-2a**).



**Figure 2-4.** Spectrograms of unidentified sounds observed in the data. Time and frequency axes differ for some panels. (a) A sequence of up-swept sounds exhibiting ~10 s call separation. (b) A sequence of paired up-swept sounds exhibiting ~10 s call separation. (c) A sequence of up-swept sounds exhibiting ~100 s call separation. (d) Two sequences of down-swept sounds. (e) A crescent sound. (f) Three-bar sound. Spectrogram parameters for all panels: 128-point FFT, 90% overlap, Hamming window.

## **I. The Tulip**

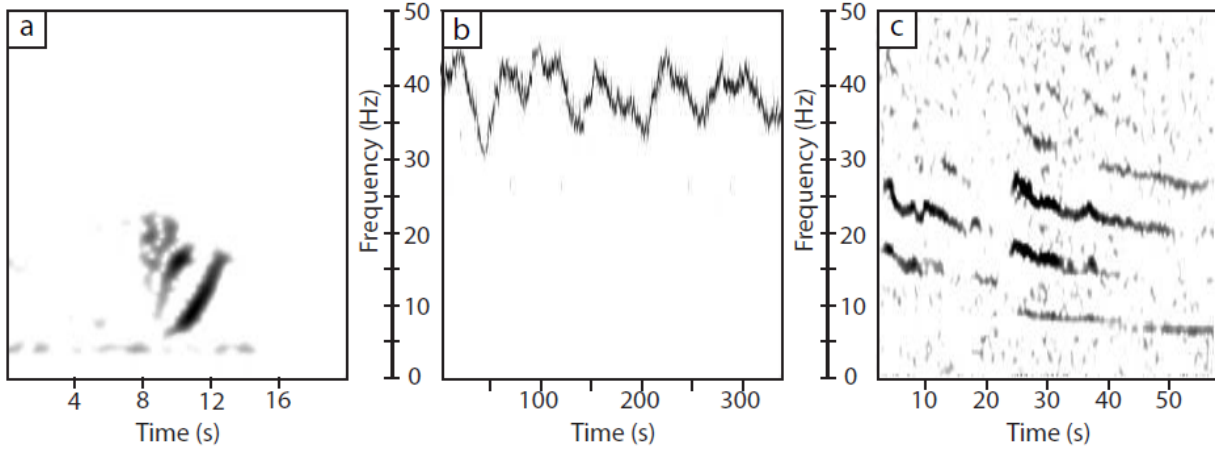
The tulip sound (**Figure 2-5a**) consists of two or more (typically two or three) up-swept components in a cluster with an average duration of 5.3 s. The earlier components tend to span higher frequencies, 9.3 to 19.9 Hz on average, than the final component, which spans 6.1 to 15.1 Hz on average and usually exhibits the largest amplitude. Although single “tulips” were typical, sets of 2 to 7 with irregularly spaced time intervals were documented. Tulip sounds were observed sporadically throughout the recording period on the two westernmost instruments located within the basin (**Figure 2-2a**).

## **J. Monochromatic whistle**

Long periods of a continuous narrow-band frequency-modulated sound (**Figure 2-5b**) were observed throughout the recording period in the <5-45 Hz range. This “whistle” passed in and out of the upper frequency limit of detection, and was most commonly observed above 30 Hz. Only one fluctuating tone was present at a time on a given station, yet higher frequency components outside of the detection band could exist. Durations of these sounds ranged from 10 s to over 30 minutes, and were observed on all six instruments throughout the year (**Figure 2-2**). The whistle appears to have a geologic origin rather than biological origin as discussed in Section IV.

## **K. Harmonic frequency bands**

Two instances of harmonic, frequency-swept tones were recorded (**Figure 2-5c**) that swept downward in frequency from 41 to 6 Hz. One instance lasted roughly 22 s and the other 73 s. These harmonic signals were only observed on the northernmost instrument in December 2009 (**Figure 2-2a**). Like the whistle describe above, this sound appears to have a geologic origin.



**Figure 2-5.** Spectrograms of possibly biologic and non-biologic sounds observed in the data. (a) Tulip sound; source unknown. (b) Monochromatic whistle exhibiting a single frequency modulated tone; probably volcanogenic. (c) Harmonic frequency bands; also likely to be volcanogenic. Spectrogram parameters for all panels: 128-point FFT, 90% overlap, Hamming window.

## **IV. Discussion**

### **A. Fin whale call**

Fin and Sei whales produce similar calls, and one of these whale species is likely source species of the doublet call observed in the Lau basin data. Since it seems more published accounts have suggested that these calls are produced by fin whales, we also suggest that these calls are produced by fin whales. Fin whales are known to produce 20-Hz doublet calls throughout the Pacific Ocean (e.g., McDonald et al., 1995; McDonald and Fox, 1999; McDonald, 2006; Stafford et al., 2007), but these and other types of fin calls have not been previously documented in the Lau basin. It is unclear whether fin whales are year-round residents of the Lau basin and surrounding areas, or if their seasonal migration paths pass through the basin. Previous studies suggest that groups of fin whales inhabit the Southern Ocean during the austral summer and move northward into the southern Pacific Ocean in the winter months (Macintosh, 1966; Širović et al., 2004; Širović et al., 2009). Fin whale calls have been detected near the Western Antarctic Peninsula between February and June; with peak call detection in May (Širović et al., 2004; Širović et al., 2009). Off the coast of New Zealand, fin calls have been detected between June and September (McDonald, 2006). In this study, fin whale doublet calls were detected in the Lau Basin between mid-June and mid-October 2010. In the past, fin whales have also been sighted by whaling vessels ~525 km south of the Lau basin between October and November 1977 (Ohsumi, 1979). Thus, it is possible, though far from being confirmed, that some fin whales reside in the Southern Ocean from roughly February to June, travel north past New Zealand from June to September, pass through the study area from mid-June to mid-October, and then return to the Southern Ocean perhaps via a similar route.

Alternatively, some populations of fin whales have been shown to remain in a specific area year-round (e.g. Tershy et al., 1993; Širović et al., 2013). Furthermore, differences between peak 20-Hz call detection and peak visual observation of fin whales (Clark and Fristrup, 1997) indicate that fin whales do not always produce 20-Hz calls while inhabiting a particular region. Fin whales also produce other call types or components at frequencies higher than our instruments can detect (e.g. Watkins, 1981; Širović et al., 2004; 2013). Therefore, we cannot rule out the

possibility that fin whales reside in the Lau basin outside of the call detection period, but could not be acoustically detected by this study.

## **B. Bryde's whale call**

Bryde's whale calls are relatively little studied, and few published descriptions are available at the frequency range of this study; furthermore, most published accounts of call detections were not visually confirmed. Nevertheless, on the basis of published calls, including the "Be3" call described by Oleson et al. (2003) and "Low Burst-tonal" call described by Heimlich et al. (2005), we suggest that the calls in our data (**Figure 2-3b**) are also Bryde's whale calls.

Although the observed calls have longer durations than these published accounts, it has been shown that call variations exist for geographically distinct populations (Oleson et al., 2003). Calls recorded off the coast of New Zealand that are attributed to Bryde's whales, but not visually confirmed, exhibit similar structure and duration to the calls recorded in the Lau Basin (McDonald, 2006). In published accounts, the burst portion of the call extends to frequencies above our recording limit, restricting the ability to match the spectrograms completely.

Bryde's whales are known to frequent the region between 40°N to 40°S (Cummings, 1985), and have been visually identified near Fiji and Tonga (Ohsumi, 1978; 1979) and off the coast of New Zealand (e.g. Gaskin, 1977; Wiseman et al., 2011). "Bryde's-like" whales have also been observed near Fiji, Nauru, New Caledonia, and Samoa (SPREP, 2007). The taxonomy of Bryde's whales is not clear, causing confusion in their identification. Recent genetic tests on specimens previously identified as Bryde's whales suggest three distinct species: *B. brydei*, *B. edeni*, and *B. omurai* (Wada et al., 2003; Sazaki et al., 2006).

Although Bryde's whale calls were primarily detected between February and June, calls were also observed at other times of the year. Bryde's whales have been shown to undertake shorter, more localized seasonal migrations than the long migrations of other baleen whales (Omura, 1962; Cummings, 1985), probably following food sources (Gaskin, 1977; Best, 2001; Kerosky et al., 2012). Other populations appear to reside in an area year-round where food sources remain consistent (Best, 2001; Wiseman et al., 2011). It is possible that Bryde's whales transit around and through the Lau basin year-round while following food sources in the general area.

### **C. Antarctic Blue whale call**

Three-component calls similar to those recorded east of the Lau basin (**Figure 2-3c**) have been recorded in the presence of blue whales near Antarctica (e.g. Ljungblad et al., 1998, Rankin et al., 2005). Such “Antarctic blue whale” calls have also been acoustically detected in the Pacific, Indian, and Southern Oceans (e.g. Clark and Fowler, 2001; Širović et al., 2004; Stafford et al., 2004; Samaran et al., 2010). In addition, the single-component version of the calls, also attributed to blue whales, has been detected in the Southern Ocean (Clark and Fowler, 2001) and eastern tropical Pacific Ocean (Stafford et al., 2004).

It is generally understood that many blue whales feed at high latitudes during the summer and migrate to lower latitudes to mate and birth during the winter (Mackintosh, 1966); however, the exact migration paths and extent of winter mating grounds are poorly understood. By determining when and where Antarctic-type blue whale calls are detected, the migration paths may be determined. Antarctic-type blue whale calls were detected in the Southern Ocean year-round with peaks in February-May and October-November (Širović et al., 2004; Širović et al., 2009), and adjacent to the Lau basin between June and September. The timing of Lau basin detections coincides with acoustic detections in the eastern tropical Pacific Ocean, where Antarctic-type blue whale calls were recorded between June and September, with a peak in July (Stafford et al., 2004). Additionally, Antarctic-type blue whale calls were detected off the coast of New Zealand between May and July (McDonald, 2006), and a July peak in calls was reported in the Indian Ocean (Samaran et al., 2010). These detection periods agree with the general understanding of blue whale migration, and suggest that at least some Antarctic blue whales migrate to a range of low latitude locations during austral winter.

### **D. New Zealand blue whale call**

This multiple-component call recorded in and around the Lau basin (**Figure 2-3d**) is similar to the “New Zealand blue whale song,” which has been definitively correlated with blue whales off the coast of New Zealand (Miller et al., 2014). Although the calls recorded in the Lau data do not have the first of the four components described in some accounts, it is similar to the version described by Kibblewhite et al. (1967). The last component of this study’s version has a shorter

duration than the published accounts, which may be due to the faintness of the calls caused by the whales being far from the instruments. Previous studies off the coast of New Zealand have recorded the New Zealand blue whale song in June, July, September, and December of 1997 (McDonald, 2006), and between January and March 2013 (Miller et al., 2014), while this call type was recorded on Lau basin instruments between late May and early June and in August. As a result of the limited observations of this call, it is unclear if the blue whales that produce these calls are migratory or if they reside in the area year-round.

### **E. The Upsweep**

The seasonal pattern of the three types of 30+ Hz upswept sounds recorded in the study area (**Figure 2-4a-c**) is similar to, but longer than those of humpback (e.g. Dawbin, 1966; Helweg et al., 1998), fin, and blue whales in or near the basin. Short pulses at 30 to 100 Hz have been attributed to fin whales (e.g. Watkins, 1981; Širović et al., 2013); yet these short-pulsed fin whale calls are described as typically down-swept in frequency, rather than upswept, with regular sequences generally not indicated. While blue whale D calls are often seen within the same frequency range as this upswept call (e.g. Oleson et al., 2007; Dunn and Hernandez, 2009), D calls are typically down-swept and do not exhibit the repeating pattern observed with this call. Multiple repetitions of short broadband units have been attributed to humpback whales, but the repeat interval is often significantly shorter and the frequency higher than the sounds reported here (Helweg et al., 1998). Furthermore, humpback songs are usually recorded and reported at higher frequencies than our observation band, thus discouraging our attempts to match the down-swept sound with previously identified humpback calls. Consequently, we classify these sounds as having unknown origin.

### **F. The Down-sweep**

The source of the down-swept sound (**Figure 2-4d**) is unclear, and is made more difficult by the limited number of examples detected in the data set. The down-sweeps were detected between June and October, which coincides with humpback, fin, and blue whale occurrence in or near the basin. As previously mentioned, both fin and blue whales produce low frequency, down-swept calls (e.g. Watkins, 1981; Oleson et al., 2007), but the frequency ranges do not match the

particular down-swept sound observed in the Lau data. On the other hand, calls recorded previously off the coast of New Zealand, are very similar to our down-swept sounds. Those calls also exhibit short, down-swept pulses occurring in groups with varying frequency from pulse to pulse. McDonald (2006) attributes those calls to humpback whales, but there was no visual confirmation.

### **G. The Crescents**

A literature search did not uncover spectrograms showing calls with the same structure as the crescents (**Figure 2-4e**), but the crescent sounds do share some similarities to some types of blue whale calls. The crescent sound is vaguely reminiscent of the “arch sound” attributed to North Atlantic blue whales by Mellinger and Clark (2003), but the Lau sounds occur at a lower frequency and without a repeating pattern. The crescent sound also shares some similarities to the continuous version of the North Atlantic blue whale AB calls (Mellinger and Clark, 2003; Berchok et al., 2006), except that the Lau sounds occur at a slightly higher frequency, not in a repeating pattern, and with more than one crescent stacked together. Calls similar to the continuous North Atlantic AB blue calls with stacked tones at 25 and 35 Hz were recorded in the Southern Ocean (Pangerc, 2010), suggesting it is possible to detect North Atlantic-like blue whale calls in the southern hemisphere. While the stacked calls described by Pangerc (2010) have similar frequencies to the crescent sound, their stacked calls are shorter in duration and were described with additional components not seen in our crescent sound. In summary, the source cannot be determined due to the lack of a good spectrogram match.

### **H. The Three-Bar**

Although our spectrograms of the three-bar sound (**Figure 2-4f**) are not of the best quality, the basic structure is fairly clear. More than one type of baleen whale is known to produce long duration (~10 s) frequency and amplitude modulated tonal sounds like the three-bar sound. Blue whales, for example, produce sidebanding in some calls like the AB call of northeast Pacific blue whales (e.g. Dunn and Hernandez, 2009). However, given the few rather poor examples of this sound, the origin of the sidebands is hard to determine. Sidebands could be produced by the

whale itself via amplitude modulation of the central tone; or the amplitude pulsing might be a sound propagation effect.

While the three-bar sound is somewhat similar to calls that have been attributed to “Madagascar-type” pygmy blue whales in the Indian Ocean (e.g. Samaran et al., 2010), the previously published calls exhibit different frequencies, are twice as long in duration, and contain additional call components that were not observed in the Lau data. Pygmy blue whales are hard to identify and their vocal patterns are not well established. They have not been documented in the Lau area, but there are reports of pygmy blue whales off the coast of northern New Zealand and New Caledonia (Borsa and Hoarau, 2004; Branch et al., 2007). Due to the lack of a good spectrogram match, this sound remains unclassified.

## **I. The Tulip**

No sound structure similar to the tulip sound (**Figure 2-5a**) was found during a literature search of both biological and non-biological sounds. Several observations suggest that the tulip sound has a biological source. For example, the sound has a consistent, complex structure with a short duration much like other identified whale calls, and it was observed on at least two widely spaced instruments, with possibly similar, but much fainter, candidates recorded on one other station. The consistency of its intricate structure from one recording to the next as seen on more than one station suggests a non-geological source. Other observations are more ambiguous. For example, on a wide time scale, many instances of the sound cluster in the time series, similar to sounds produced by vocalizing whales (**Figure 2-2**); yet there is no clear seasonal pattern to the tulip sounds like most whales in the basin. On shorter time scales the tulip sound does not occur in sequences with a regular repeat interval, but not all known whale calls occur in regular patterns either. The irregular time intervals of the tulip sound indicate that it does not have an anthropogenic source; pile drivers and airgun surveys are performed with long sequences (hours to days) of regularly spaced pulses. Finally, some observations seem to indicate a non-biological source. For example, the structure of the tulip sound is unlike published spectrograms of any whale call at  $<50$  Hz that we have observed, has the look of an interference pattern due to sound propagation effects, and the last component of the tulip sound has frequencies extending lower than the lowest known calls of baleen whales. The lowest frequency baleen whale call in the

literature appears to be a well-documented 9 Hz blue whale call (Mellinger and Clark, 2003). Given that some spectrogram smearing to lower frequencies possibly occurs in our analysis, we suggest that the tulip sound is not out of the realm of possibility for a blue whale. Given the lack of any corroborating evidence, the source of this sound remains unknown.

#### **J. Monochromatic whistle**

Long-duration, frequency-fluctuating whistles similar to those observed in our data (**Figure 2-5b**) were also recorded in the northern Lau basin on hydrophones during the same time frame as our study (Dziak et al., 2013). Those authors suggest a volcanogenic source, in particular the West Mata deep submarine volcano (coordinates 15°6.06'S, 173°45.54'W), which is active seismically, magmatically, and hydrothermally. They speculate that the source of the sound is focused degassing or hydrothermal flow through a restricted conduit or vent.

#### **K. Harmonic frequency bands**

Sweeping harmonic signals of the type seen in our data (**Figure 2-5c**) have also been recorded near volcanoes in general (Dziak et al., 2008) and on northern Lau basin hydrophones located near West Mata volcano (Dziak et al., 2013). Again, these signals have been associated with submarine volcanic eruptions, tremor, or submarine landslides (Dziak et al., 2013; Caplan-Auerbach et al., 2014). Harmonic sounds associated with volcanic activity are more common in the literature than the non-harmonic tone discussed above, but in our data the non-harmonic tone was significantly more common. Another possible source for this sound is Antarctic ice tremor. Ice tremor sounds have a structure similar to the harmonic tremor described here, and have been recorded over 7500 km from their source (e.g. Chapp et al., 2005). However, the close spatial and temporal association with the hydrophone recordings around West Mata volcano indicate that the volcano is the more likely source.

#### **V. Conclusions**

Our study presents the first long-term and systematic acoustic investigation of baleen whale sounds across and adjacent to the Lau basin. In this study, we detected known calls from fin, blue, and Bryde's whales. Calls produced by humpback whales may also have been recorded,

but exact spectrogram matches could not be made. Although fin, Bryde's, blue, and humpback whales were previously known to reside in the South Pacific Ocean in the general area of Fiji and Tonga, very little information has been available about their population dynamics, seasonality, and call repertoires in this region. The low frequency baleen whale calls detected on OBSs around the Lau basin between November 2009 and October 2010 suggest that some blue whales and possibly fin whales follow migration patterns that lead them from Antarctic regions to the study area in the austral winter. The incidence of Bryde's whales, however, appears to peak in the austral autumn, with some individuals residing in and near the study area almost year-round.

Several unknown signals were also detected in the data set. Although no visual observations were available to correlate the unknown signals with a particular source, these signals share characteristics of baleen whale calls, such as repetition patterns and frequency ranges, suggesting that they have biological sources. The difficulty in matching some of the observed sounds to published whale call accounts, as well as the possible discovery of new call types, emphasizes the limited knowledge of the full call repertoires of baleen whales. The Lau basin is just one of many areas that have received little attention in the past with regard to whale acoustics; therefore, it is possible that many other unpublished calls or versions of calls exist. By utilizing data from long-term OBS surveys in remote areas, whale calls and other sounds can be further documented, improving our understanding of marine bioacoustics.

## CHAPTER 3. AUTOMATIC DETECTION ALGORITHM

### I. Introduction

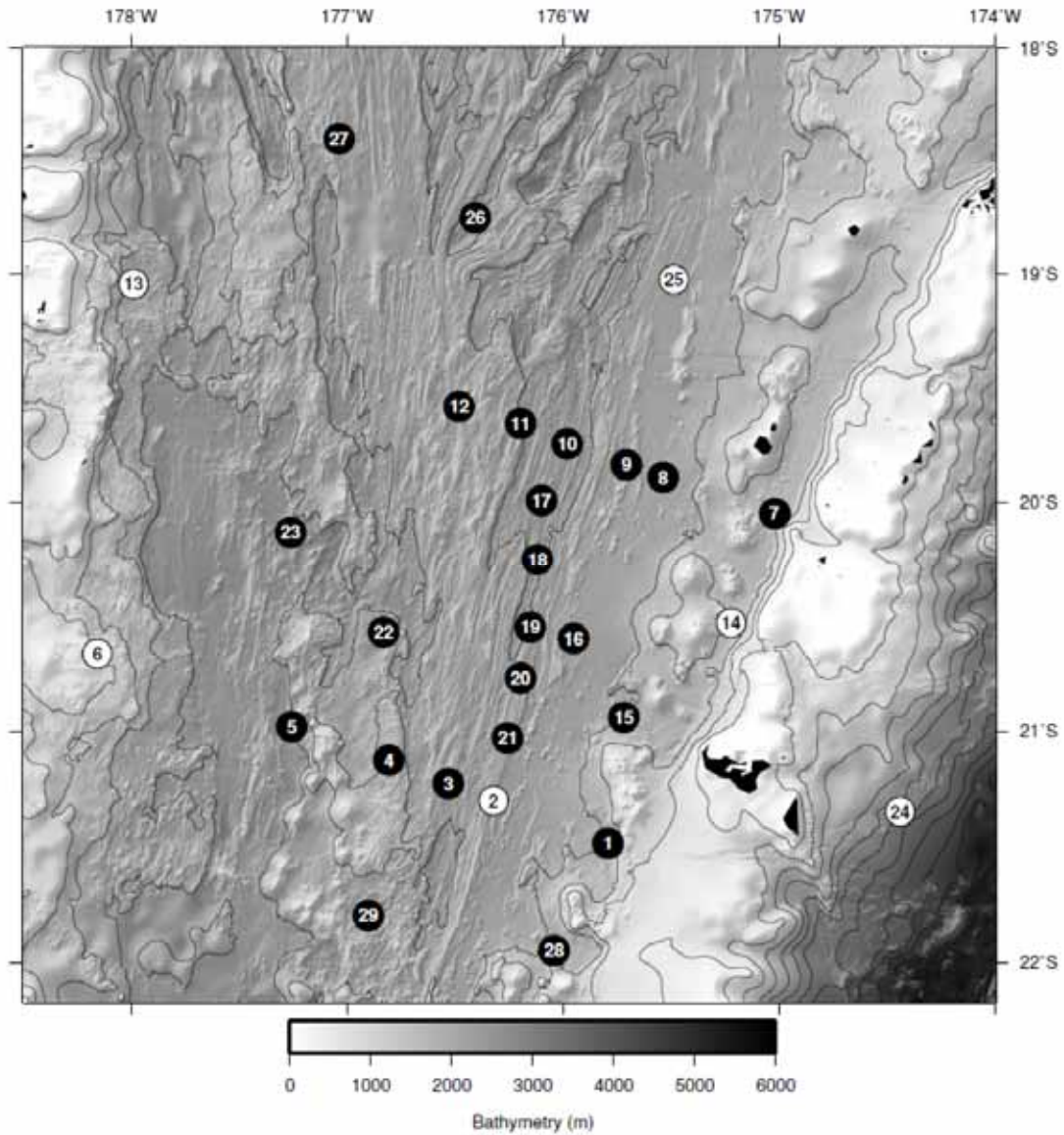
The full Lau basin data set consists of approximately 10 months of continuous recording on 29 broadband OBS instruments (**Figure 3-1**). Given its large size, the most practical method of identifying acoustic signals is to use an automatic detection algorithm, since sorting through the entire dataset visually would be overly time consuming. A variety of approaches for auto-detection routines have been implemented previously, some as simple as examining signal-to-noise ratios, while others using pattern matching or feature identification algorithms that require extensive training sets. Automatic detection algorithms reduce human bias associated with visual inspection of the acoustic records; although there can be other types of bias associated with the training aspects of pattern matching routines. This chapter presents the development and utilization of a detection algorithm built specifically for the Lau basin ocean-bottom seismograph (OBS) data set to sort through the acoustic spectrogram data.

### II. Data and Methods

To create a whale call identification algorithm, we chose to utilize principal component (PC) analysis and multinomial logistical regression methods. PC analysis allows each sound type to be characterized with the minimum number of variables required. PC analysis was chosen to project the 3 second time-frequency spectrogram window, which inhabits a space of very high dimension, onto a subspace of lower dimension in which the signals of the training set might be more easily distinguished from each other. It is thought to act as a signal to noise enhancing filter, but this aspect was not compared with other classification algorithms.

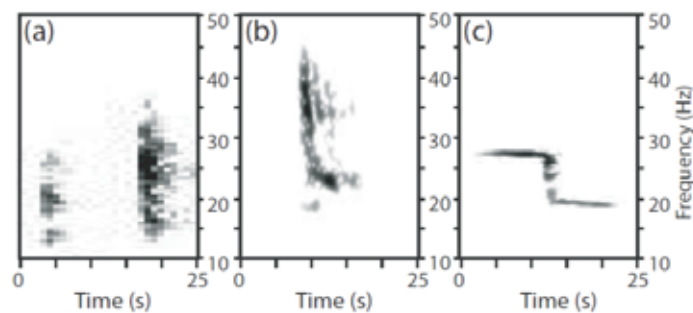
PC analysis transforms a number of possibly correlated variables into a smaller number of uncorrelated variables called principal components. This transformation is defined in such a way that the first PC accounts for the most variance in the data, the second PC accounts for most of what remains, and so forth. The data are efficiently compared to example sound types which have also been transformed into PC space. Multinomial logistical regression can then be used to

classify each segment of the data set (in PC space) by calculating the probability that the data segment includes a certain sound type.



**Figure 3-1.** Location of the OBS instruments used to detect baleen whale calls. The instruments were located in the western tropical Pacific Ocean between Tonga and Fiji within the Lau basin (with the exception of instrument 24 located east of Tonga). The shaded bathymetric map of the study area shows 500 m contour intervals. The white circle symbols show the locations of the six OBS used in the preliminary visual analysis and by the automatic detection algorithm. The black circle symbols show the locations of the other 23 ocean-bottom seismographs used by the automatic detection algorithm.

As discussed in Chapter 2, four types of low-frequency baleen whale calls were identified during the preliminary study including doublet fin whale calls, Bryde’s whale calls, Antarctic blue whale calls, and New Zealand blue whale calls. Considering observations and results of the preliminary study, we have excluded the New Zealand blue whale call from the detection algorithm development process due to its long duration, and we have instead focused on the remaining three whale calls (**Figure 3-2**).



**Figure 3-2.** Spectrograms of whale calls attributed to fin, Bryde’s, and Antarctic blue whales. The frequency and time axes are the same for all three panels. (a) A fin whale doublet call (256-point FFT). (b) A Bryde’s whale call (128-point FFT). (c) An Antarctic blue whale call (128-point FFT). Spectrogram parameters for all panels: 90% overlap, Hamming window. In this figure, some areas of random background noise were intentionally masked to avoid confusion with the indicated sounds.

The PC-multinomial logistical regression automatic detection algorithm was developed in several stages. First, a training set comprised of various examples of the possible output classes (acoustic sound types) was created. These output classes included fin, Bryde’s, and Antarctic blue whale calls in addition to earthquakes, random background noise, instrument noise, and other sounds encountered in the data set. Several different versions of each output class were included in the training set to account for variability in each sound. For whale call sounds, each variation was comprised of only one call rather than call trains consisting of multiple calls. Each sound example was created as a 25-second spectrogram using a 100-point Hamming window with 50% overlap and a 256-point Fast Fourier Transform (FFT). The power of each frequency

band was normalized to reduce horizontal bands in the spectrograms by dividing each frequency band by its median value.

After all of the example sounds were saved, the training set was converted to PC space. This was accomplished by first vectorizing each example sound as a column and then taking the natural log of each sound example to create positive and negative values. Then each of these vectorised sounds was put in a matrix A. Then, the empirical orthogonal functions (EOFs) of matrix A were determined using:

$$n^{-1}AA^T = U\Lambda U^T$$

where the columns of U are the EOFs, and n is the number of columns in A. Individual example sounds were then put into PC space by:

$$r = U^T R$$

where R is the original individual example sound vector, and r is the example sound in PC space. Each PC example sound was then normalized and saved for later use.

The training set was verified against itself for self-consistency, as well as verified against a test set comprised of hand-picked whale calls and other sounds from the data set to ensure that the algorithm correctly identified each sound type. This was accomplished by putting the entire training set in PC space, multiplying it by the PC matrix, and then normalizing. Then, the PC training set was multiplied by the transpose of each PC sound example, and the probabilities of each sound type were calculated by:

$$\frac{P_r(t)}{P_o(t)} = e^{r^T \Psi(t)} \text{ with } \frac{1}{P_o} = 1 + \sum_{r \neq 0} \frac{P_r}{P_o}$$

where r is a specific example sound,  $P_r(t)$  is the probability that the tested sound is that specific example sound,  $P_o(t)$  is the probability that the sound is in an unidentified category (the “other” class), and  $\Psi$  is the PC training set. The same protocol was followed for the verification test set.

To determine the best method for using the probabilities described above to accurately select valid sound detections, the probabilities were combined in two different ways: (1) “Top Two” – Multiple versions of each sound type were included in the training set. This probability combination method only considered algorithm detections valid if the sound categories with the two highest probabilities for a window were the same; and (2) “Combination” – For each window, the probabilities of the multiple versions of each example sound types were added together (e.g., all of the individual Bryde’s call example probabilities were combined into one Bryde’s call probability) so that the algorithm considered the combined probabilities of each sound type rather than individual examples of each sound type. The results of the two different probability-combination tests were compared on detection error threshold (DET) plots and receiver-operator curve (ROC) plots to determine the best method and to determine the appropriate probability threshold. To create the DET and ROC plots, the algorithm was applied to six days of data from the dataset using 25-second windows with a 3 second step size. Each window of data from the six days was transferred into PC space, and then the probabilities were calculated for each sound type in the same way the algorithm analyzed the verification test set described above. The six chosen days came from various stations and included two days with confirmed call examples for each of the three whale call types. The results were visually inspected to determine the false positive rate (FPR), true positive rate (TPR), and false negative rate (FNR) where

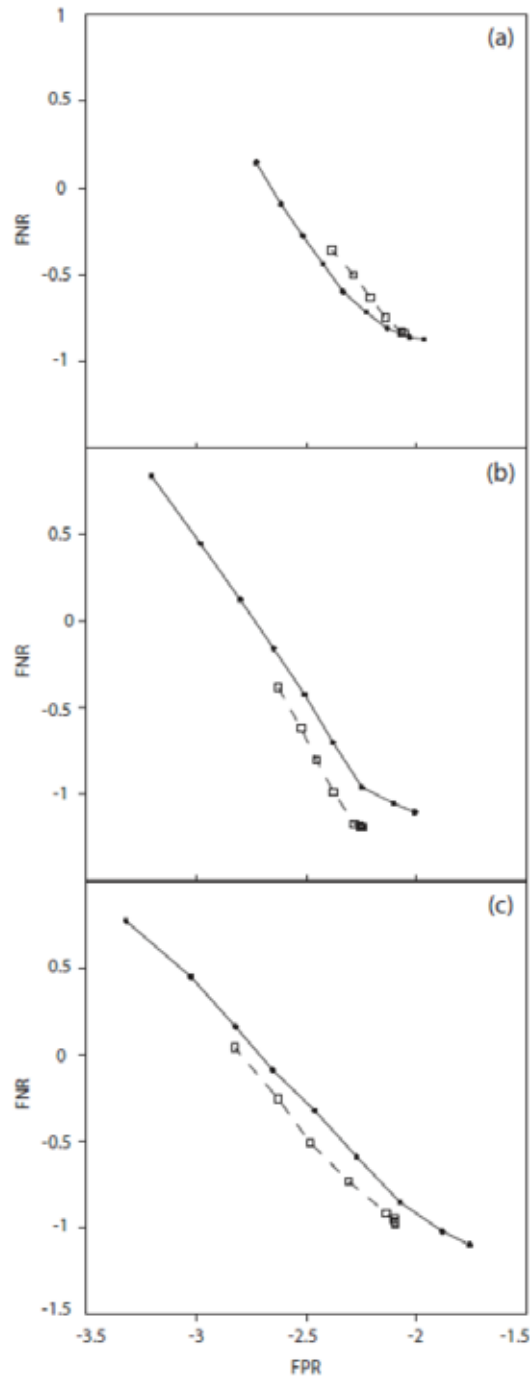
$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives}$$

$$TRP = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

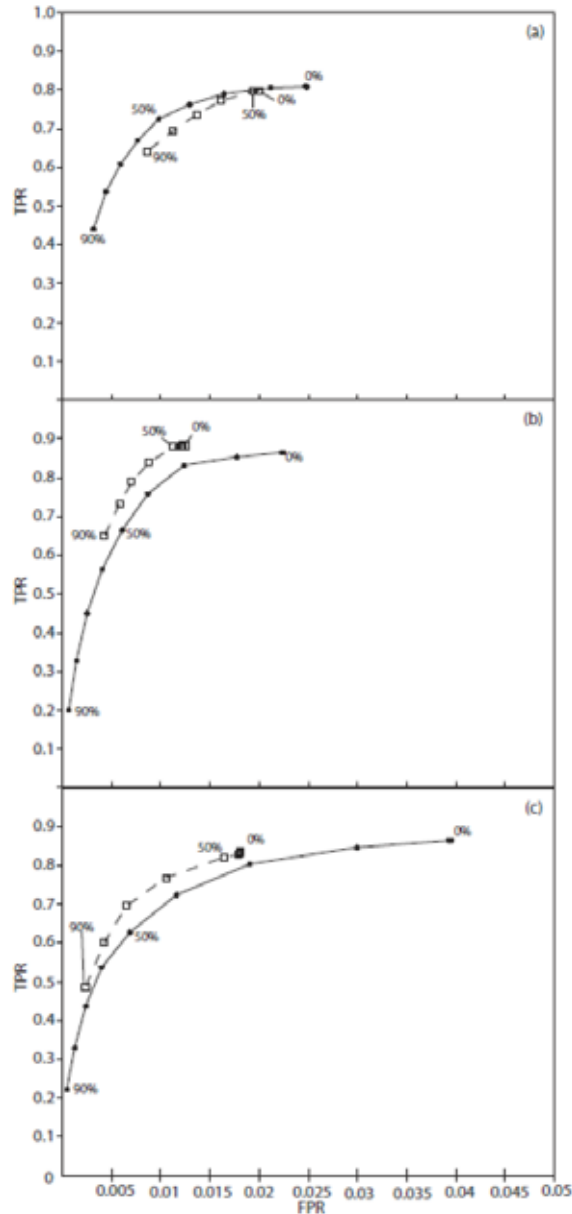
$$FNR = \frac{False\ Negatives}{False\ Negatives + True\ Positives}$$

by comparing the algorithm detections against spectrograms of the data. False positives included duplicate detections due to the 3 second step size.

Probability thresholds were examined in 10% increments for each of the three whale call types to analyze how varying the probability threshold impacted the number of detections. The DET plots indicated that the “Combination” method produced better results than the “Top Two” method for Bryde’s and Antarctic blue whale calls, but the “Top Two” method produced better results for fin whale calls (**Figure 3-3**). The ROC plots indicated that using only the 90-100% probability detections would be most appropriate for the purposes of this study because the 90% probability threshold minimizes false positive detections while still providing adequate true positive detections (**Figure 3-4**). Based on these results, the algorithm analyzed the data from each of the 29 instruments using the “Combination” method and a 90% probability cutoff. However, for comparison purposes, whale call detection information for other probabilities was saved as well.



**Figure 3-3.** Detection error tradeoff plots showing false positive rate (FPR) versus false negative rate (FNR) for the six days of data analyzed by the automatic detection algorithm. The FPR and FNR have been non-linearly scaled by their standard normal deviates. The axes are the same for each panel. Solid lines with dot points indicate “Top Two” probability method, and dashed lines with square points indicate the “Combination” probability method. Symbols (dots or squares) are shown for every 10% probability threshold between 0% and 90%. Each panel shows the DET curves for a separate call type: (a) Fin whale calls; (b) Bryde’s whale calls; and (c) Antarctic blue whale calls.



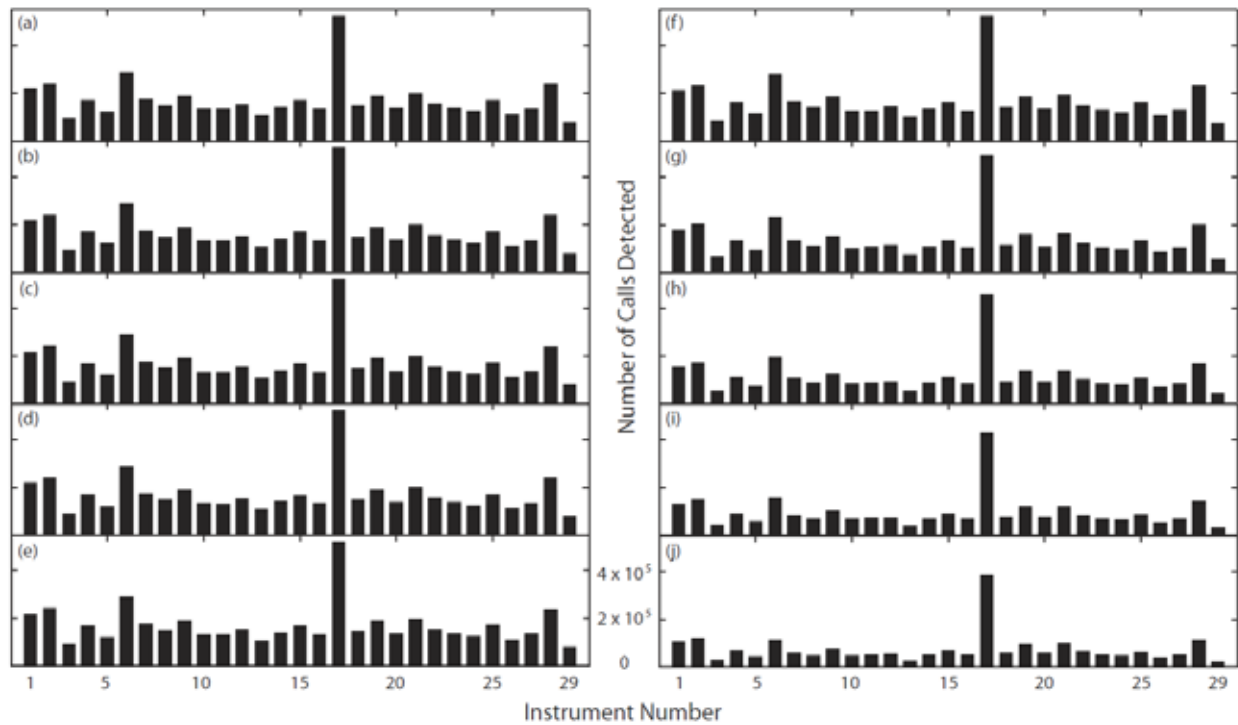
**Figure 3-4.** Receiver-operator curve plots showing false positive rate (FPR) versus true positive rate (TPR) for the six days of data analyzed by the automatic detection algorithm. The axes are not scaled the same. Solid lines with dot points indicate the “Top Two” probability method, and dashed lines with square points indicate the “Combination” probability method. Symbols (dots or squares) are shown for every 10% probability threshold between 0% and 90%. Each panel shows the ROC plot for a separate call type: (a) Fin whale calls; (b) Bryde’s whale calls; and (c) Antarctic blue whale calls.

### III. Algorithm Results

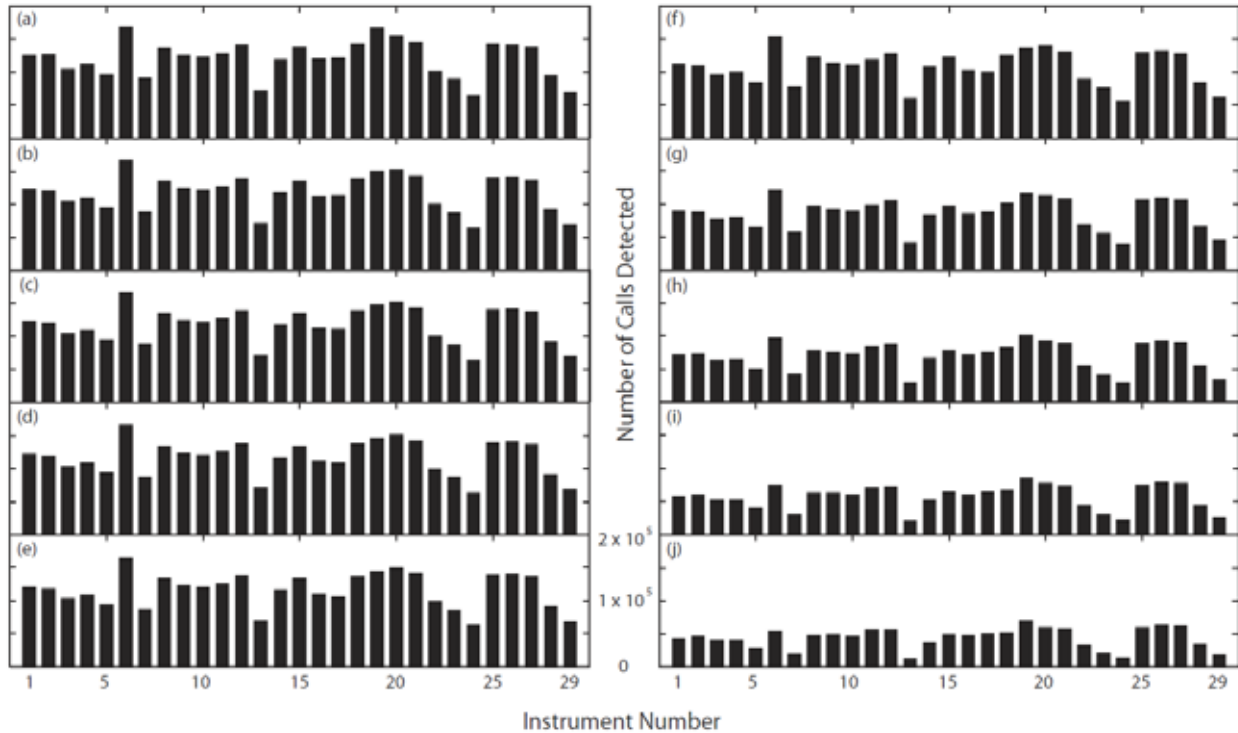
Algorithm detection results were compared for the range of probability thresholds. In general, the number of whale calls detected by the algorithm decreased as the probability threshold was increased to higher percentage ranges as shown on **Figure 3-5** for fin whale calls, **Figure 3-6** for Bryde's whale calls, and **Figure 3-7** for Antarctic blue whale calls.

Each of the three whale call types were detected on instruments throughout the entire recording period. Considering the results of the visual study (Chapter 2), it is more likely that the algorithm incorrectly identified background noise or other sounds as whale calls than it is that whales were actually calling throughout the entire recording period; therefore, when the number of whale calls is above the daily FPR calculated for the DET and ROC plots ("FP Cutoff"), it is more likely that actual whale calls were recorded in the data set.

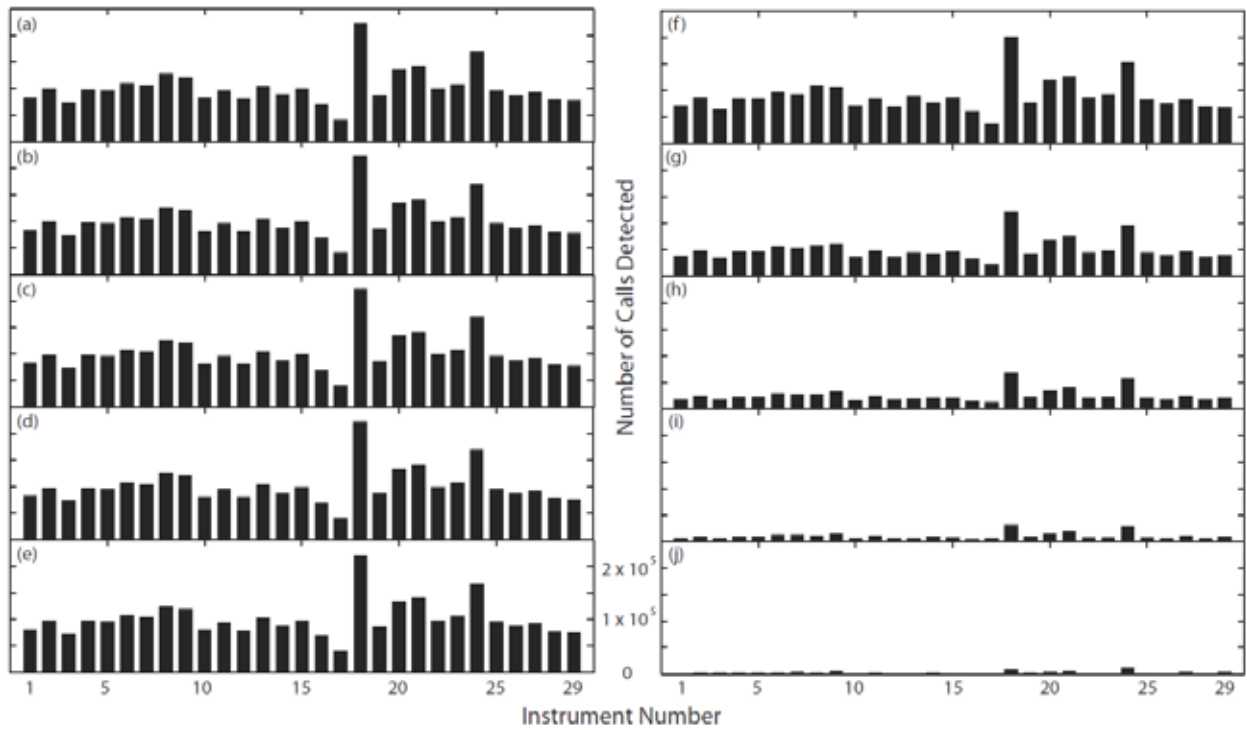
Results were also differentiated by the time of day that calls were detected to determine the diel patterns of the call detections. The diel pattern plots were divided into detections for each month of the recording period to see if the patterns varied over the recording period. Algorithm detection results were also compared for different regions of the study area to determine spatial patterns and distribution. These regions include the north basin, north central basin, central basin, west central basin, south central basin, south basin, northeast basin, northwest basin, southwest basin, and outside of the basin.



**Figure 3-5.** Fin whale call detections for varying probability thresholds. The axes are the same for each panel where the x-axis is the instrument number, and the y-axis is the number of fin call detections. Each panel shows a different probability threshold range: (a) 0 – 100%; (b) 10 – 100%; (c) 20 – 100%; (d) 30 – 100%; (e) 40 – 100%; (f) 50 – 100%; (g) 60 – 100%; (h) 70 – 100%; (i) 80 – 100%; and (j) 90 – 100%.



**Figure 3-6.** Bryde’s whale call detections for varying probability thresholds. The axes are the same for each panel where the x-axis is the instrument number, and the y-axis is the number of Bryde’s call detections. Each panel shows a different probability threshold range: (a) 0 – 100%; (b) 10 – 100%; (c) 20 – 100%; (d) 30 – 100%; (e) 40 – 100%; (f) 50 – 100%; (g) 60 – 100%; (h) 70 – 100%; (i) 80 – 100%; and (j) 90 – 100%.



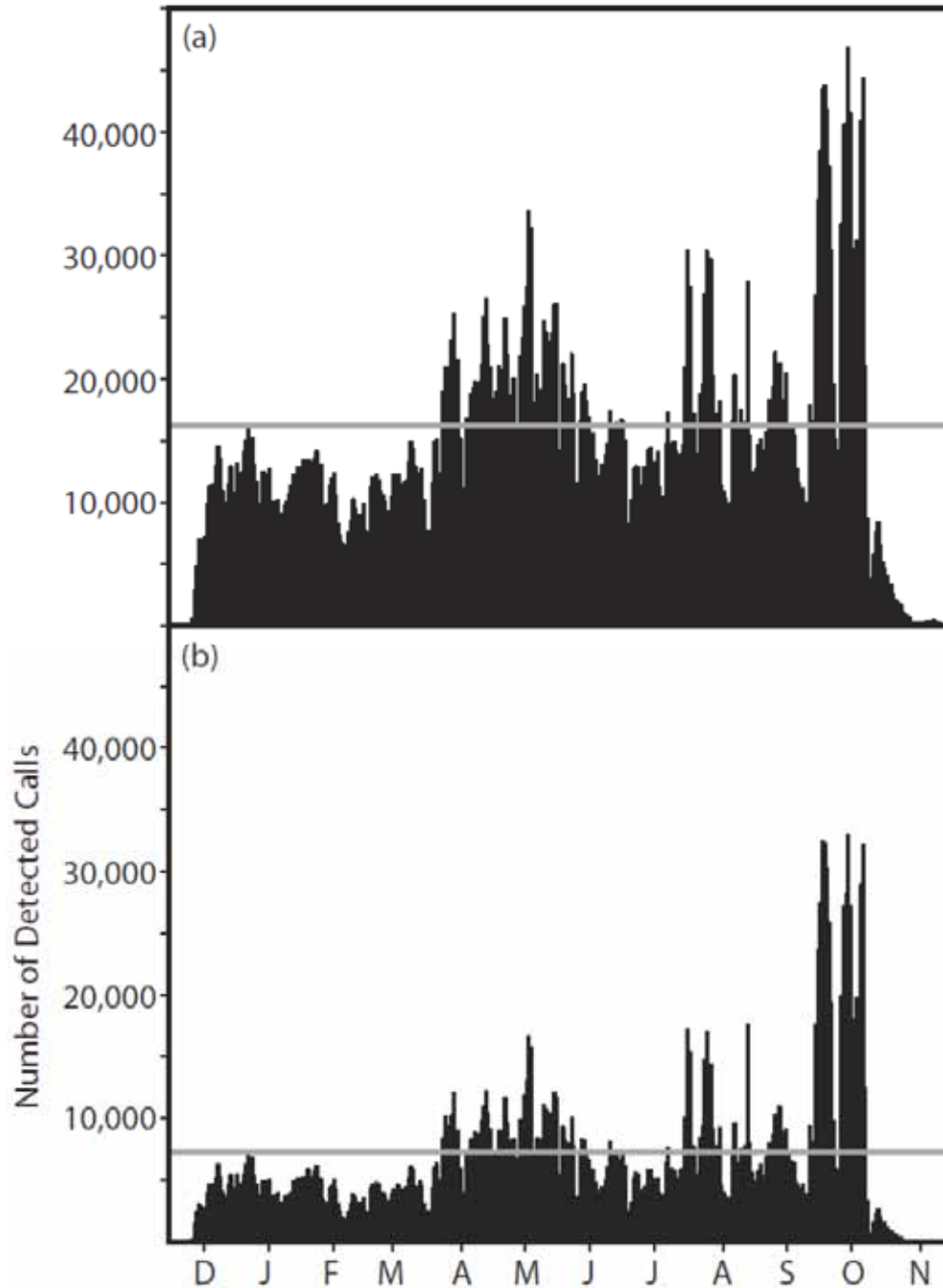
**Figure 3-7.** Antarctic blue whale call detections for varying probability thresholds. The axes are the same for each panel where the x-axis is the instrument number, and the y-axis is the number of blue call detections. Each panel shows a different probability threshold range: (a) 0 – 100%; (b) 10 – 100%; (c) 20 – 100%; (d) 30 – 100%; (e) 40 – 100%; (f) 50 – 100%; (g) 60 – 100%; (h) 70 – 100%; (i) 80 – 100%; and (j) 90 – 100%.

## A. Fin whale calls

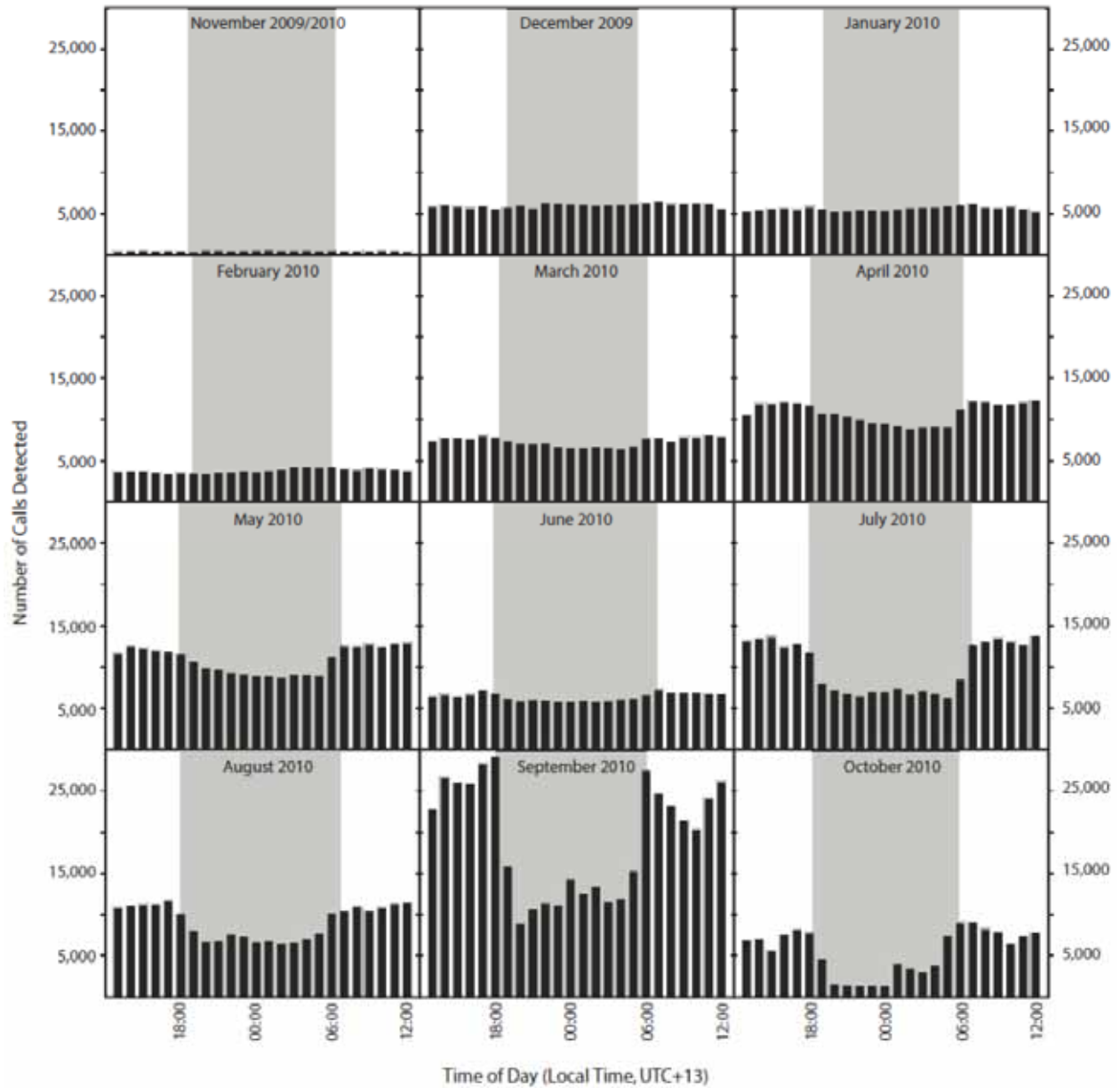
Fin whale calls were identified by the algorithm throughout the recording period (**Figure 3-8**), but were primarily observed above the FP Cutoff between mid-March and late May 2010 and between mid-July and early October 2010. The number of call detections was greatest between mid-September and early October 2010.

Fin calls were detected throughout the day and night during Bryde's season, with more calls detected during daylight hours than during the night (**Figure 3-9**). The diel pattern plots also show seasonal trends of call detections, specifically that the number of calls detected increased significantly in September 2010. Similar, but less extreme, diel patterns were also observed between March and May 2010 and in June, August, and October 2010. The time of day that calls were detected in September 2010 show that significantly more calls were detected during daylight hours than during the night. It should be noted that only a few days of data were available for November 2009, and the low detections are not indicative of the absence of whale vocalizations.

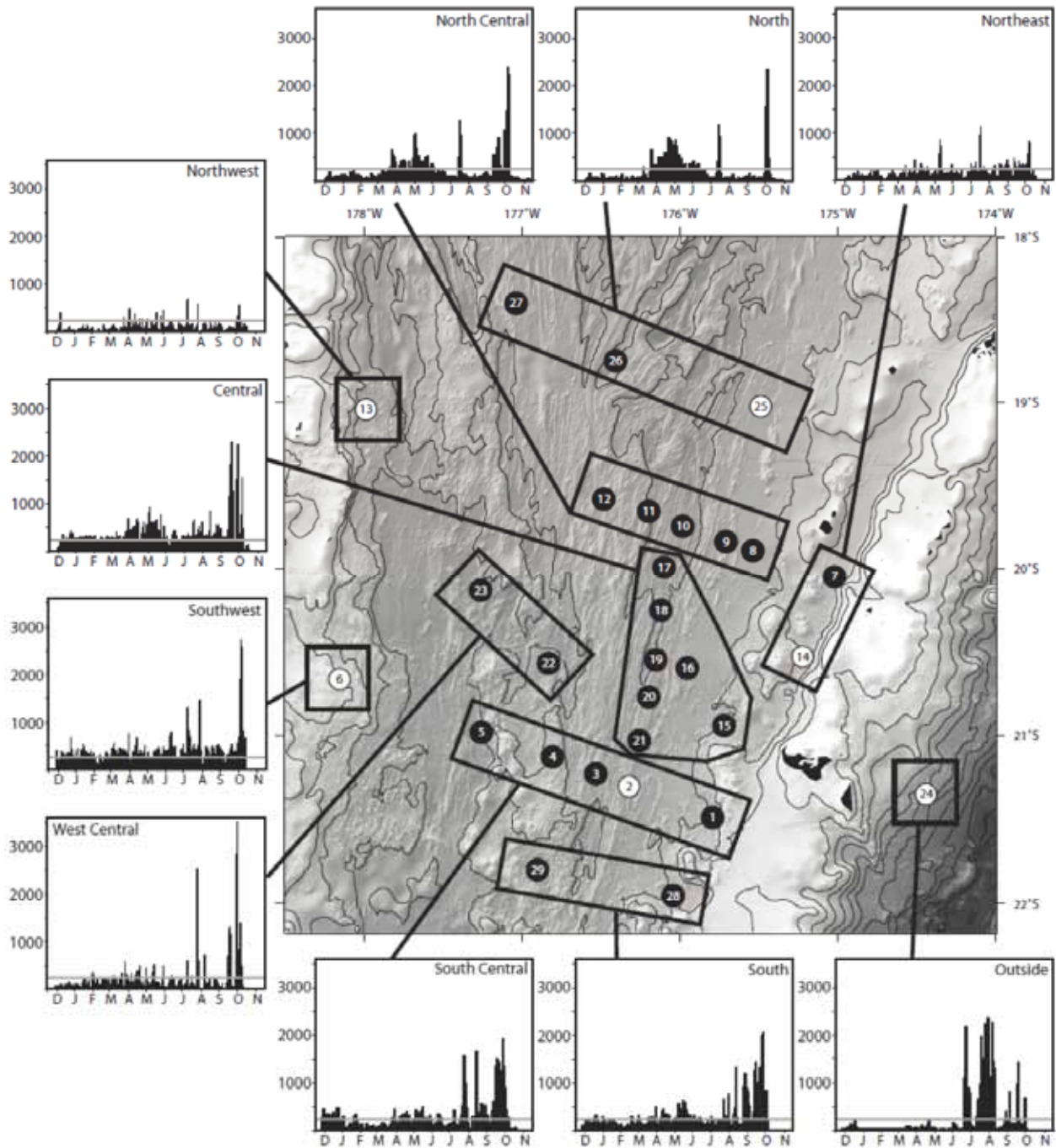
A comparison of fin whale call detections for different regions of the study area indicates that more fin whale calls were detected on instruments in the north, north central, and central regions between April and June 2010 (**Figure 3-10**). High peaks of fin whale calls were detected on instruments in six regions (north, north central, west central, southwest, south central, and outside) in July 2010, in three regions (south central, south, and outside) in August 2010, and in eight regions in October 2010 (north, north central, central, west central, southwest, south central, south, and outside). Relatively few detections were identified on the instruments in the northeast and northwest regions of the basin. An individual comparison of each instrument for 90 – 100% probability detections indicates that calls detected on 23 of the instruments were relatively similar (**Figure 3-5j**). Of the remaining OBS, three instruments (instruments 3, 13, and 29) had fewer call detections than the other instruments, and one instrument (instrument 17) had significantly more detections than the other OBS.



**Figure 3-8.** Fin whale call detections throughout the recording period for all 29 OBS instruments. The axes are the same for both panels where the x-axis indicates the months during the recording period from December 2009 (D) through November 2010 (N) and the y-axis indicates the number of calls detected. The panels show different probability threshold ranges: (a) Detected calls with 0 – 100% probabilities; and (b) Detected calls with 90 – 100% probabilities. The gray horizontal lines indicate the FP Cutoff.



**Figure 3-9.** Fin whale call diel pattern. Call detections include those from all 29 instruments throughout the recording period using the 90% probability threshold. Each panel shows one month of the recording period as indicated. The axes are the same for each panel where the x-axis is the time of day in local Tongan time, and the y-axis is the number of calls detected. Gray shading indicates the time between sunset and sunrise (the time between sunrise and sunset is day; the time between sunset and sunrise is night).



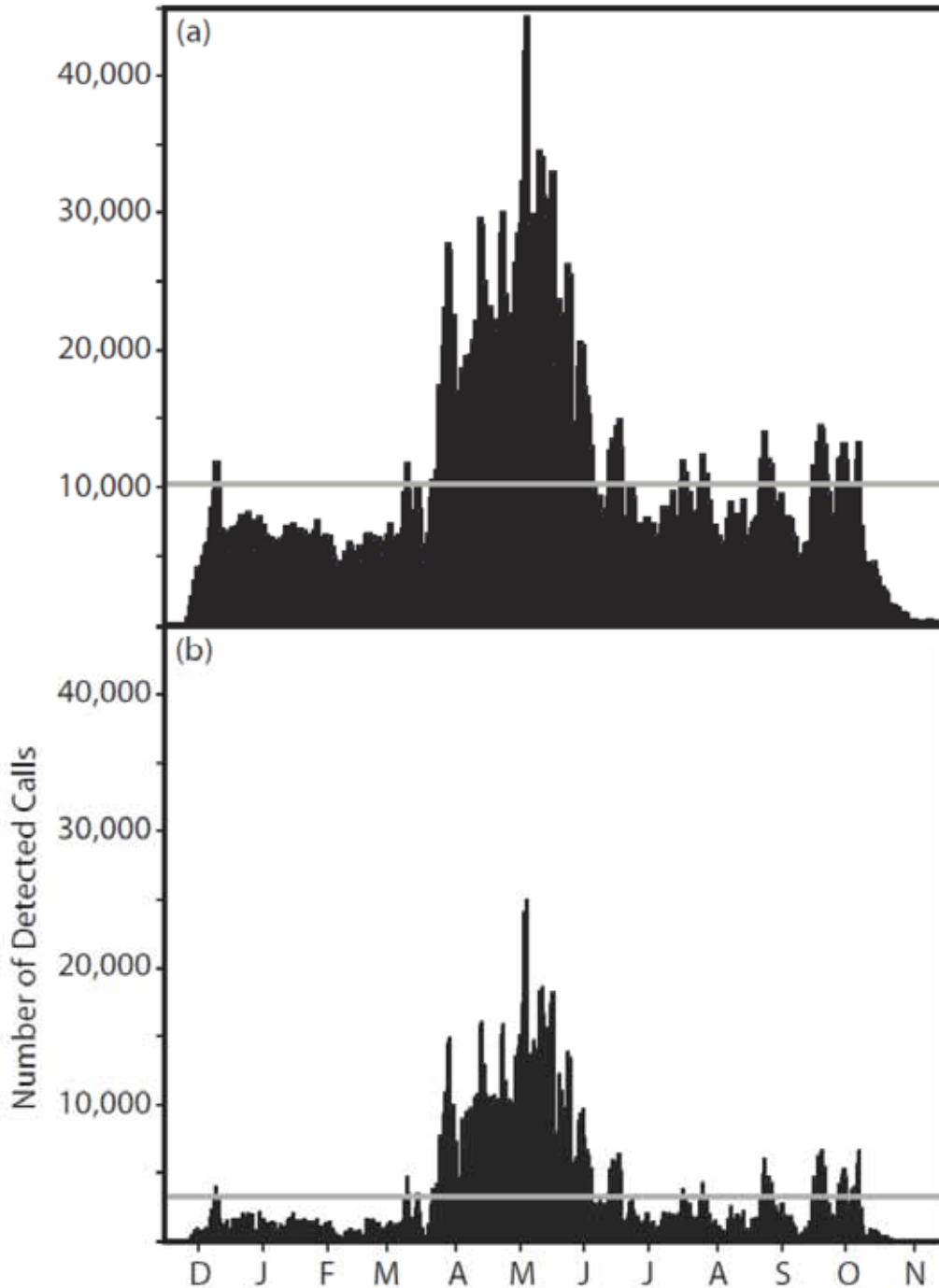
**Figure 3-10.** Fin whale call detections by region of the Lau basin. Call detections were separated into the ten regions shown. The center map shows locations of the 29 OBS instruments deployed within and adjacent to the Lau basin. The surrounding histograms depict the mean number of calls detected on the regions' instruments for each day using the 90% probability threshold. The axes are the same for each plot where the x-axis displays tick marks indicating the month beginning with December 2009 and ending in November 2010, and the y-axis measures the mean number of detected fin whale calls. The FP Cutoff is shown as horizontal gray lines on each histogram.

## **B. Bryde's whale calls**

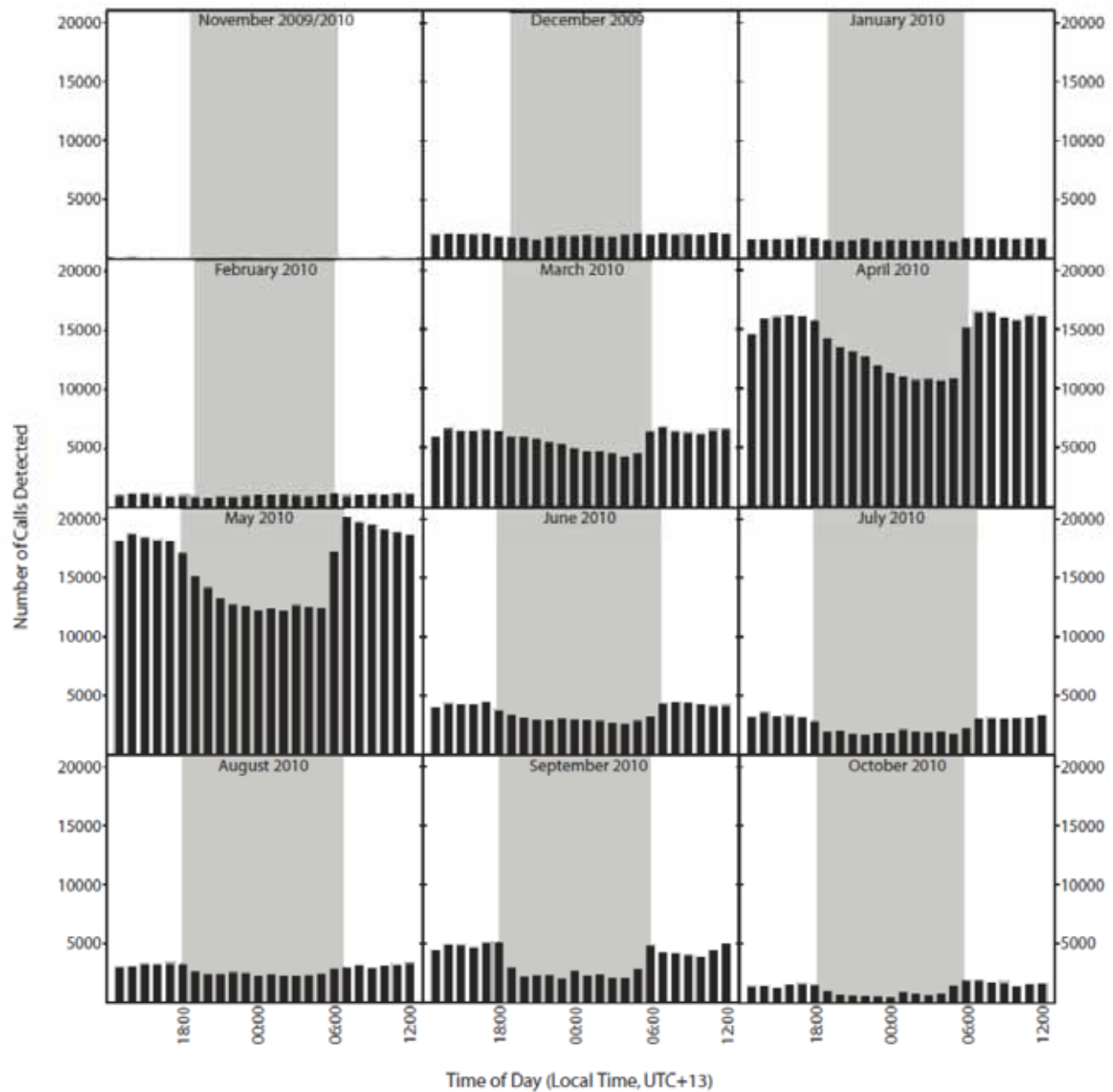
Bryde's whale calls were also detected throughout the recording period (**Figure 3-11**), but they were primarily detected above the FP Cutoff between mid-March and mid-June 2010, in mid- to late August 2010, and between mid-September and early October 2010. The number of call detections was greatest between mid-March and late May 2010.

Bryde's calls were detected throughout the day and night during Bryde's season, with more calls detected during daylight hours than during the night (**Figure 3-12**). The diel pattern plots also show seasonal trends of call detections, specifically that the number of calls detected increased significantly in April and May 2010. Less extreme examples of the Bryde's whale call diel pattern were also observed in March 2010 and between June and September 2010. As noted above, only a few days of data were available for November 2009, and the low detections are not indicative of the absence of whale vocalizations.

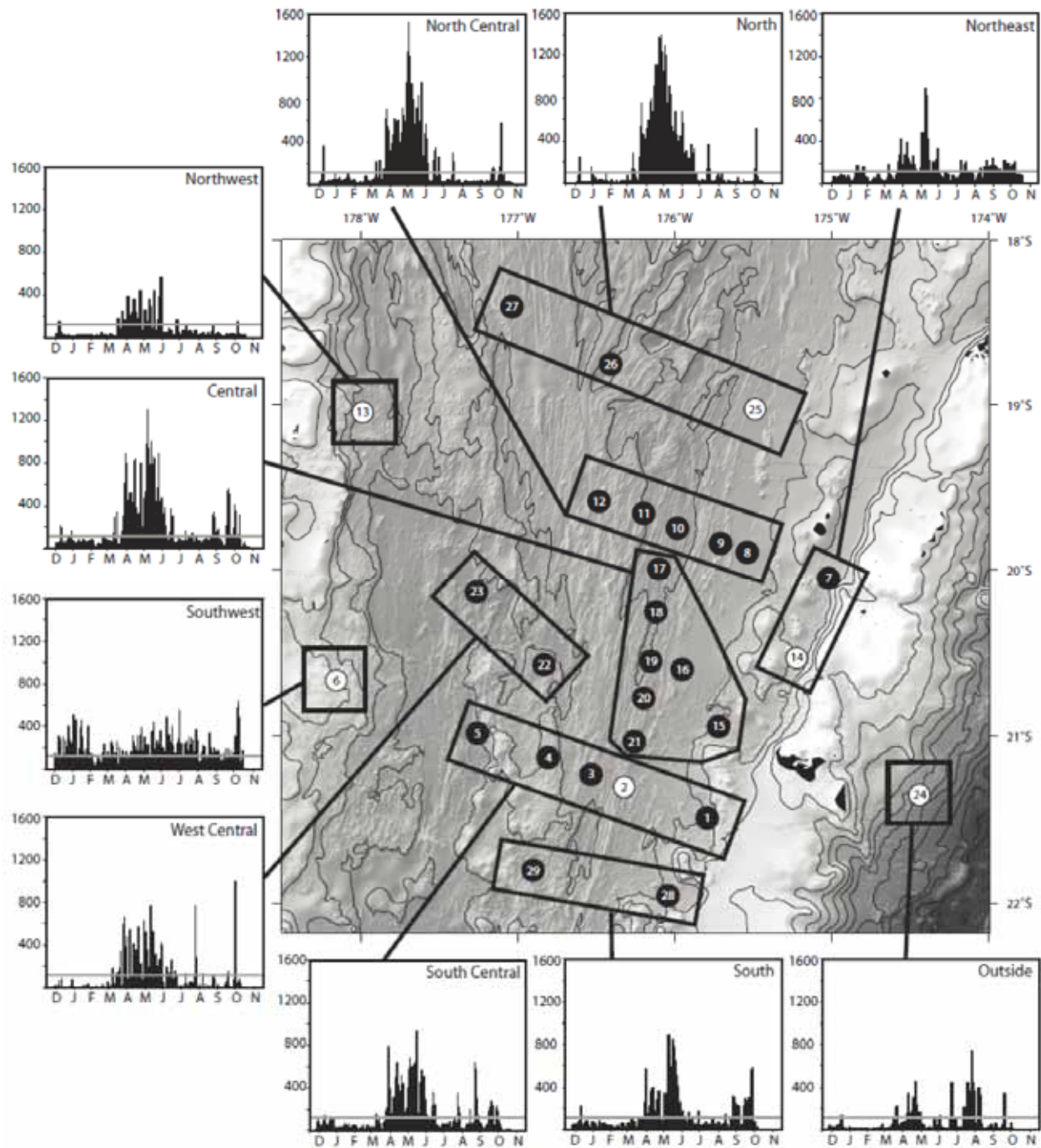
A comparison of Bryde's whale call detections for different regions of the study area indicates that more Bryde's whale calls were detected on instruments in the north, north central, and central regions of the basin than on instruments in the other regions (**Figure 3-13**). The fewest detections were recorded on the instruments in the northwest and southwest regions of the basin and outside the basin. The northern regions histograms show heavier concentrations of calls between April and June 2010, but calls were detected more frequently in the central and southern instruments between August and October 2010, and more calls were detected on the southwest instrument in December 2009 and January 2010. A individual comparison of each instrument for 90 – 100% probability detections indicates that calls detected on 23 of the instruments were relatively similar, and the remaining six instruments (instruments 5, 7, 13, 23, 24, and 29) had fewer call detections (**Figure 3-6j**).



**Figure 3-11.** Bryde’s whale call detections throughout the recording period for all 29 instruments. The axes are the same for both panels where the x-axis indicates the months during the recording period from December 2009 (D) through November 2010 (N), and the y-axis indicates the number of calls detected. The panels show different probability threshold ranges: (a) Detected calls with 0 – 100% probabilities; and (b) Detected calls with 90 – 100% probabilities. The gray horizontal lines indicate the FP Cutoff.



**Figure 3-12.** Bryde’s whale call diel pattern. Call detections include those from all 29 instruments throughout the recording period using the 90% probability threshold. Each panel shows one month of the recording period as indicated. The axes are the same for each panel where the x-axis is the time of day in local Tongan time, and the y-axis is the number of calls detected. Gray shading indicates the time between sunset and sunrise.



**Figure 3-13.** Bryde’s whale call detections by region of the Lau basin. Call detections were separated into the ten regions shown. The center map shows locations of the 29 OBS instruments deployed within and adjacent to the Lau basin. The surrounding histograms depict the mean number of calls detected on the region’s instruments for each day using the 90% probability threshold. The axes are the same for each plot where the x-axis displays tick marks indicating the month beginning with December 2009 and ending in November 2010, and the y-axis is the mean number of detected Bryde’s calls. The FP Cutoff is shown as horizontal gray lines on each histogram.

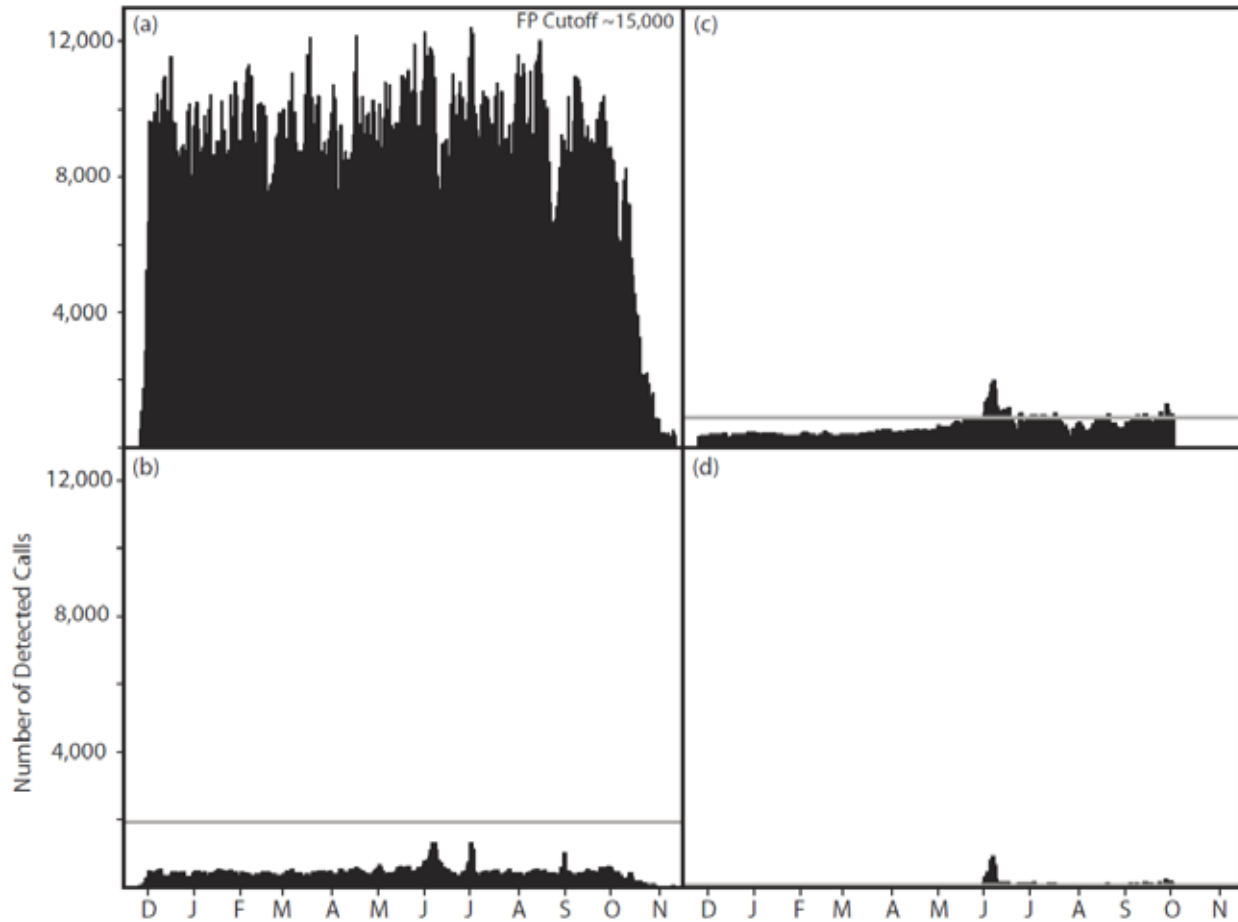
### C. Antarctic blue whale calls

Antarctic blue whale calls were detected throughout the recording period, although fewer Antarctic calls were detected than fin or Bryde's calls. Blue whale call detections were fairly evenly distributed over the recording period when analyzing a combination of all 29 instruments, and no obvious seasonality when analyzing all probability thresholds (**Figure 3-14a**). Although small detection peaks in June, July, and September 2010 are shown when analyzing 90 – 100% probability thresholds, these detections are below the FP Cutoff (**Figure 3-14b**). The preliminary study (Chapter 2) indicated that blue whales were only detected on the OBS instrument located outside of the basin (instrument 24); therefore, this instrument was also analyzed separately. Detections of blue whale calls on this instrument alone indicated that a detection peak in June 2010 and sporadic detection above the FP Cutoff between July and October 2010 (**Figure 3-14c,d**).

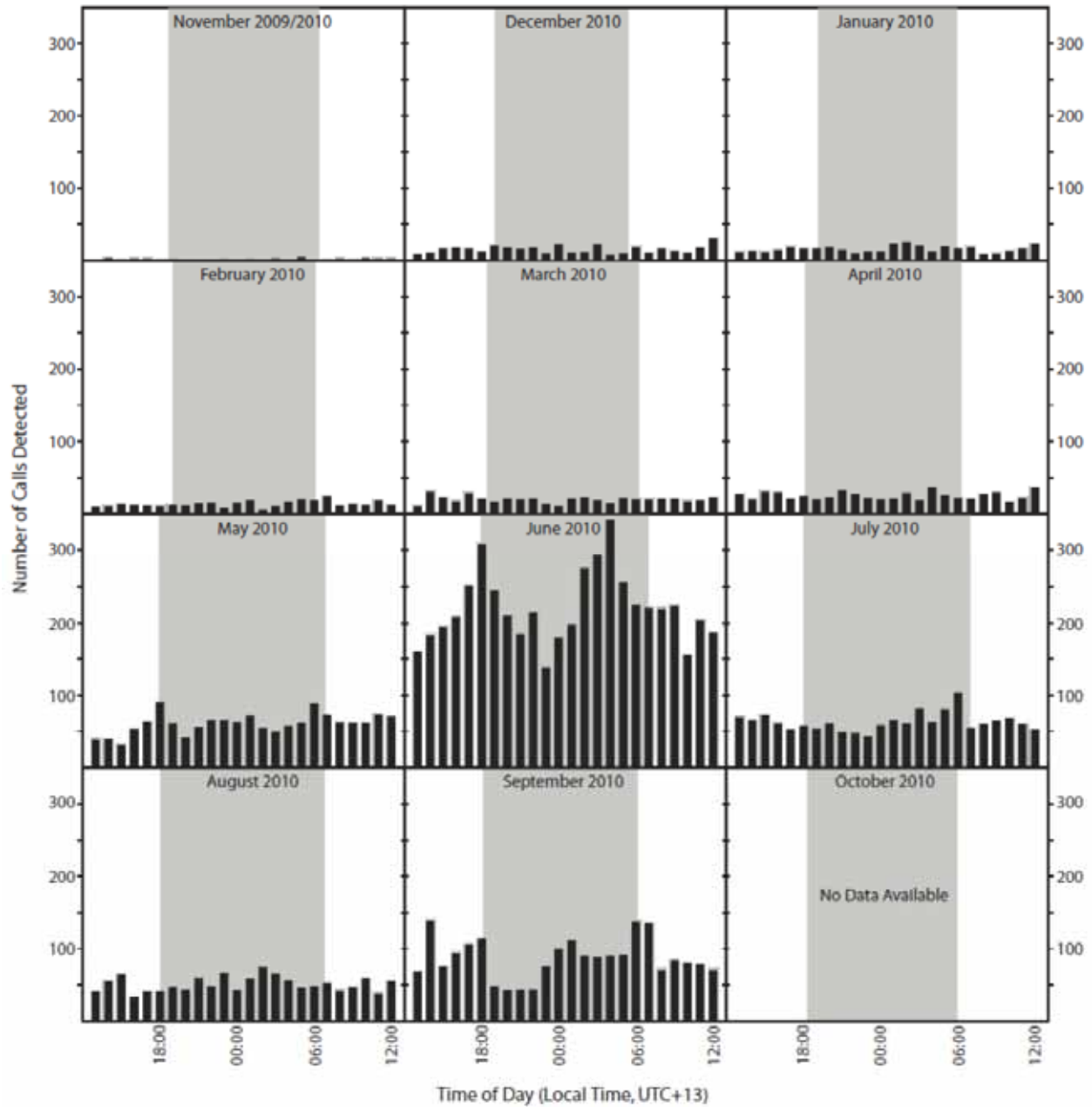
Blue call detection diel pattern plots were only analyzed for calls detected on the instrument outside of the basin, since including other instruments would likely bias the results. Antarctic blue whale calls were detected throughout the day and night during blue whale season on the OBS instrument located outside of the basin (**Figure 3-15**). In May, June, and September 2010, more calls were detected before and after sunset and in the hours before sunrise than other times. In July and August 2010, calls fluctuated over the course of the day with peaks in the morning before sunrise and in mid-afternoon. The time of day plots also show seasonal trends of call detections, specifically that the number of detected calls increased significantly in June 2010, but were also detected above background levels in May and between July and September 2010. No data was available from the OBS instrument outside the basin for the month of October 2010, and few days were available for November 2009. The low number of detections in November 2009 is not indicative of the absence of whale vocalizations.

A comparison of blue whale call detections for different regions of the study area indicated that the vast majority of calls were detected on the instrument outside of the basin (**Figure 3-16**). A few detection peaks above the FP Cutoff were identified on instruments in the northeast region in December 2009, in the south central regions in July 2010, and on the outside region in October 2010. A comparison of each instrument individual for 90 – 100% probability detections indicates

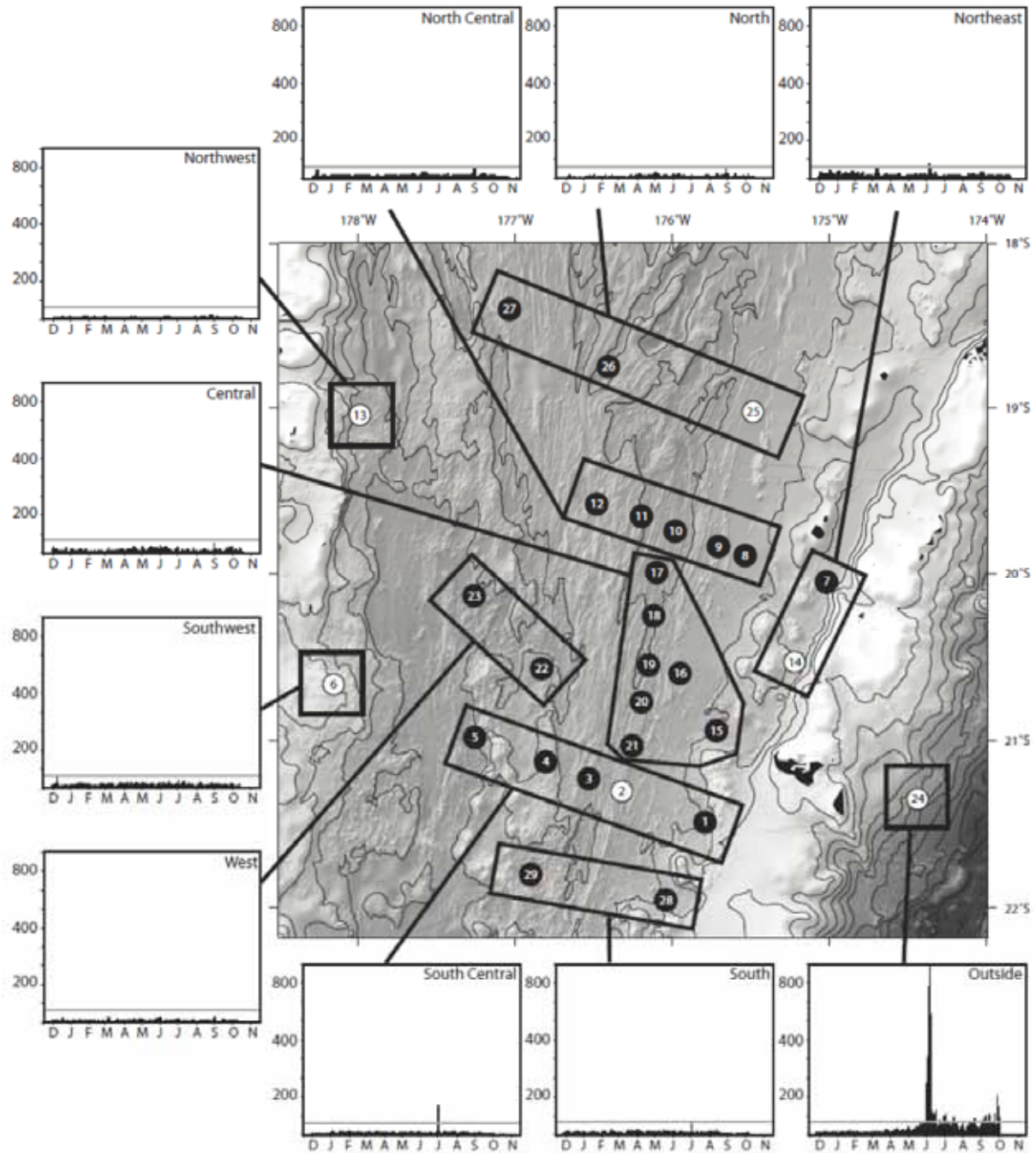
that calls detected on 27 of the instruments were had few detections, and the remaining two instruments (instruments 18 and 24) had the most call detections (**Figure 3-7j**).



**Figure 3-14.** Bryde's whale call detections throughout the recording period. The axes are the same for both panels where the x-axis indicates the months during the recording period from December 2009 (D) to November 2010 (N) and the y-axis indicates the number of detected calls. The panels show different probability threshold ranges for different instrument configurations: (a) Probability thresholds 0 – 100% for all 29 instruments; (b) Probability thresholds 90 – 100% for all 29 instruments; (c) Probability thresholds 0 – 100% for only the instrument outside of the basin; (b) Probability thresholds 90 – 100% for only the instrument outside of the basin. The gray horizontal lines indicate FP Cutoff of the detections; although panel (a) does not show the FP Cutoff because it is above the y-axis range as indicated.



**Figure 3-15.** Antarctic blue whale call diel pattern. Call detections include those from all 29 instruments throughout the recording period using the 90% probability threshold. Each panel shows one month of the recording period as indicated. Axes are the same for each panel where the x-axis is the time of day in local Tongan time and the y-axis is the number of calls detected. Gray shading indicates the time between sunrise and sunset.



**Figure 3-16.** Antarctic blue whale call detections by region of the Lau basin. Call detections were separated into the ten regions shown. The center map shows locations of the 29 OBS instruments deployed within and adjacent to the Lau basin. The surrounding histograms depict the mean number of calls detected on the region’s instruments for each day using the 90% probability threshold. The axes are the same for each plot where the x-axis displays tick marks indicating the month beginning with December 2009 and ending in November 2010 and the y-axis is the mean number of detected blue calls. The FP Cutoff is shown as horizontal gray lines on each histogram.

#### **IV. Discussion**

As discussed in the Results section above, fin, Bryde's, and Antarctic blue whale calls were recorded on all 29 instruments throughout the recording period, which is inconsistent with the results of the preliminary investigation (Chapter 2). Observations made during the algorithm testing process, in addition to the ROC curves, indicate that the algorithm identified false positive signals. Although the FPR is very low, the quantity of data results in many false positive detections. Along the same line, although the false negative rate is low, many whale calls may have been incorrectly classified and, therefore, missed. Overall, the algorithm used to analyze the data set worked best for Bryde's whale calls (TPR = 0.88, FPR = 0.01 for 0 – 100% probability thresholds; TPR = 0.65; FPR = 0.004 for 90 – 100% probability threshold), but still worked reasonably well for fin whale calls (TPR = 0.79, FPR = 0.012 for 0 – 100% probability thresholds; TPR = 0.64; FPR = 0.003 for 90 – 100% probability threshold) and Antarctic blue whale calls (TPR = 0.84, FPR = 0.014 for 0 – 100% probability thresholds; TPR = 0.49; FPR = 0.001 for 90 – 100% probability threshold).

False positive and false negative detections may result from several different causes. As discussed in further detail in the sections below, some of the calls have similar structure to random background noises or electronic data spikes resulting in false positives. Additionally, some whale calls may also have been obscured by the numerous earthquakes or other strong signals (e.g., strong electronic spikes) recorded on the instruments resulting in false negatives.

During the development of the algorithm, a 3 second step size was chosen because it helped reduce the missed calls associated with a larger step size and helped reduce the duplicate call detection associated with a smaller step size. However, some calls were likely missed and some calls were likely duplicated for this reason. To account for duplications, duplicate calls were evaluated during the FPR calculation and are included in the FP Cutoff.

Detections of whale calls on the instruments may be influenced by propagation effects in the ocean which can distort the call or cause the call to lose power (Helble et al., 2013). For example, propagation effects may be influenced by seasonal variations in water temperature

which change sound speeds. These effects may reduce the detections of calls, especially during winter months when colder water at the surface may trap sound, preventing it from reaching the instruments on the ocean floor. Therefore, a lack of call detections does not necessarily mean that whales are not present in the study area, and may be due to propagation effects prohibiting from being recorded on instruments or the whales emitting different, higher frequency calls that cannot be detected in our frequency range as discussed in Chapter 2.

### **A. Fin whale call**

Fin whale calls were primarily detected between mid-March and late May and between mid-July and early October. The preliminary visual study indicated fin whales calls were recorded in the basin between mid-June and mid-October (Chapter 2), which corresponds with the second period of algorithm-based fin whale detections. However, the preliminary visual study did not identify fin whales between March and May. The calls detected by the algorithm between March and late May coincide with the Bryde's whale call seasonality identified during the preliminary visual study; therefore, it is possible that these fin whale calls are incorrectly identified Bryde's calls. Regional plots show numerous detections of both Bryde's and fin calls in the same regions (north, north central, central, south central, and south) between April and June 2010, which might be further evidence that the fin whale calls detected during this time period are in fact misidentified Bryde's calls. Although a similar pattern was observed sporadically between July and October 2010, both fin and Bryde's were observed during that time of year in the visual investigation (Chapter 2). Additionally, it was noted during development and testing that the algorithm often misidentified Bryde's calls as fin calls in addition to correctly identifying them. Attempts were made during algorithm development to reduce this incorrect selection, which improved the selections, but did not fix the problem entirely.

In addition to incorrectly selecting Bryde's whale calls as fin calls, the algorithm also incorrectly identified data spikes and other strong vertical signals as fin calls, likely due to the relatively plain structure of the fin whale call, especially in PC space. Future algorithms might be more successful if trains of fin whale doublet calls are converted into PC space rather than an individual doublet set. By increasing the length of the example call window, the call structure

would be more unique, and the algorithm would be less likely to mistake single spikes, random background noise, or Bryde's whale calls.

Diel patterns of the detected calls show that fin whales are more vocal during daylight hours, although they also appear to make vocalizations during the night. Previous studies have shown that the 20 Hz fin whale call diel patterns vary by region and possibly time of year (Širović et al., 2013). Data from the Bering Sea indicated 20 Hz calls were more prevalent during the day and night than during dusk or dawn, whereas data collected off the coast of California and in the Gulf of California indicated that calls were more prevalent during the night and at dawn (Širović et al., 2013). The diel patterns likely stem from changing activity patterns throughout the day; however, it remains unclear what purpose the calls serve or activity they may indicate.

Fin whale calls were detected above the FP Cutoff in all analyzed regions of the study area. Detections were most frequent in the central and southern regions (including north central, central, south central, south, and outside regions) between July and October. Relatively few calls were detected in the northwest and northeast regions during the same time period, indicating regional preferences; however, it is unclear why the fin whales would prefer the south and central portions of the basin. One possibility is food source prevalence in this area, but there is no data to support or refute that possibility.

As discussed in Chapter 2, it is unclear whether fin whales are year-round residents of the Lau basin and surrounding areas, or if their seasonal migration paths pass through the basin. Previous studies suggest that groups of fin whales inhabit the Southern Ocean during the austral summer and migrate northward into the southern Pacific Ocean in the winter months (Macintosh, 1966; Širović et al., 2004; Širović et al., 2009). Other studies have shown that some populations of fin whales remain in a specific area year-round (e.g. Tershy et al., 1993; Širović et al., 2013). Since fin whales produce other call types or components at frequencies higher than our instruments can detect (e.g. Watkins, 1981; Širović et al., 2004; 2013), we cannot rule out the possibility that fin whales reside in the Lau basin outside of the call detection period, but could not be detected acoustically by algorithm.

## **B. Bryde's whale call**

Bryde's whale calls were primarily detected in the basin between mid-March and mid-June 2010, but also during other times of the recording period. These results coincide well with the results of the preliminary visual study, which indicated Bryde's whale calls were recorded between February and June 2010 and in August and October 2010 (Chapter 2).

Diel patterns of Bryde's whale call detections indicate that Bryde's whales produce low frequency calls during the day, and slightly decrease the frequency of vocalization during the night. Vocalizations tended to taper off slowly in the hours after sunset, but quickly increased in numbers around sunrise. Bryde's whale calls in the Gulf of Mexico were shown to be more prevalent during dusk and night (Širović et al., 2014); however, the calls detected in the Gulf of Mexico study were a different type of call than those detected in the Lau basin data set, and, therefore, may have a different purpose for Bryde's whales leading to the difference in diel pattern.

Call detections between March and June were densest in the north, north central, and central regions of the Lau basin, whereas fewer calls were detected in the northwest, southwest, and outside regions. However, when Bryde's whale calls were detected during other times of year, more calls were detected on instruments in the south central, south, and outside regions of the basin. This spatial distribution may indicate that the Bryde's whales preferred the northern and middle portions of the study area between March and June 2010, and preferred the southern portions of the study area between July and October 2010. Spatial preference may be due to food sources migrating through different portions of the study area during different times of year.

Overall, the algorithm was most successful at identifying Bryde's whale calls likely due to their unique structure. Although the spacing between Bryde's whale calls was shown to vary in the preliminary study (Chapter 2), future algorithms might benefit from including a pair or train of Bryde's whale calls in the detection window, rather than just an individual call to match against.

Bryde's whales prefer the warmer waters between 40°N to 40°S, and are not known to undertake the long migrations that blue and fin whales are known for (Omura, 1962; Cummings, 1985). Some populations of Bryde's whales have been shown to reside in a region year-round if food

sources remain consistent (Best, 2001; Wiseman et al., 2011). As discussed in Chapter 2, it is possible that the Bryde's whales stay in the areas around the Lau basin and follow food sources as they migrate around the general area.

### **C. Antarctic Blue whale call**

Seasonality of the Antarctic blue whale within the Lau basin was difficult to determine from the results of the algorithm, because few, if any, calls were detected on instruments within the basin. Calls were detected on the instrument located outside of the basin between June and October 2010, which coincides with the results of the visual investigation (Chapter 2). The majority of calls were detected in June 2010. Calls detected between June and October coincide with general migration trends of blue whales from Antarctica and are similar to acoustic detections of Antarctic-type blue whale calls in the eastern tropical Pacific Ocean between June and September (Stafford et al., 2004).

Diel patterns of Antarctic blue whale call detections indicate that vocalizations increase during the afternoon until dusk. A second vocalization increase occurs in the early hours of the morning before sunrise, which then tapers off after sunrise. This diel pattern is similar to those of studies conducted in the eastern tropical Pacific and off the coast of southern California, which showed significantly more vocalizations during the night and twilight hours (Stafford et al., 2003, Wiggins et al., 2005).

Although the vast majority of blue whale calls were detected on the OBS instrument located outside of the basin, small call detection spikes were detected on some instruments in December 2009 and in June, July, September 2010. Call detection spikes in July were recorded on instruments located in the south central and south regions of the basin, and call detection spikes in September were detected on instruments located in the north and north central regions of the basin. This may indicate that blue whales were near or within the southern portion of the study area in July, and in the northern portion of the study area in September. Overall, the detection results indicate the Antarctic blue whales prefer the deeper waters outside of the basin.

Although blue whale calls are able to travel long distances due to their low frequency and power, relatively few calls were detected in the data set. This might be due to the depth of the

instruments, which were positioned deeper than the SOFAR channel. Background noise in the blue whale call frequency bands from distant calling whales might also have reduced the number of detected calls since increased noise in the frequency bands would make it harder for the algorithm to pick out the blue whale calls from the background.

## **V. Conclusions**

This chapter presents the results of an automatic detection algorithm developed to identify low frequency fin, Bryde's, and Antarctic blue whale calls in a data set collected in the Lau basin between November 2009 and October 2010. The algorithm detection results were compared to the results of the preliminary study (Chapter 2), which indicated that the detections above the FP Cutoff were similar to the visual study results. The algorithm appears to work well for Bryde's whale calls, but encountered some issues when selecting fin and blue whale calls. The algorithm results helped to determine seasonal, regional, and diel patterns for all three whale call types for the entire study area.

In general, the number of whale calls detected by the algorithm decreased as the probability threshold was increased to higher percentage ranges. The 90 - 100% probability threshold was typically used when generating figures, which biased number of detected calls to be low. However, using lower probability thresholds increases the likelihood that false positive detections would be included.

To improve the algorithm, using trains of calls for sound examples rather than individual calls should increase the uniqueness of the call and reduce false positives. New Zealand blue whale calls should be included in future algorithm attempts since they have a relatively unique structure and were observed on instruments within the basin during the preliminary visual study (Chapter 2). Increasing the window length to include more than one call might improve results by increasing the uniqueness of the call. Alternatively, the algorithm could use intercall intervals to increase or decrease the probability that a sound is a call by determining if another potential call was detected within a set intercall time frame.

## CHAPTER 4. CONCLUSIONS

The purpose of this thesis was to classify non-earthquake related acoustic sounds recorded on OBS instruments within and around the Lau basin. This included geologic sounds as well as marine mammal sounds. For marine mammals, large baleen whales are detected by the instruments, and a major aim of the project was to determine the spatial, diel, and temporal patterns exhibited by their calls. Few whale studies have included the Lau basin, which lies along the suspected migration route of blue and fin whales. Additionally, the basin is within territory of the relatively little studied Bryde's whale. Although the data set was originally obtained for a large-scale seismic survey, studying signals other than earthquakes allows the data to be utilized more fully.

Four low frequency baleen whale call types (doublet fin, Bryde's, Antarctic blue, and New Zealand blue whale calls) were identified in the data set by comparing spectrograms of the calls observed in the Lau basin data set during a preliminary visual inspection to similar calls associated with specific whales. These whale call types include one Bryde's whale call, the Antarctic blue whale call, the New Zealand blue whale call, and likely a doublet fin whale call (although it is also similar to a sei whale call). Identifying these calls confirms that these whale species reside in and around the basin for at least part of the year. In addition to these whale calls, many unknown or geological sounds were identified in the data set during the visual inspection, including some sounds that might be new to the whale call repertoire.

An automatic detection algorithm utilizing PC and multinomial logistical regression was developed to select the three most commonly observed calls from the data set. The algorithm detection results were used to determine temporal, diel, and spatial patterns for the three whale call types. The preliminary visual study and algorithm detection results indicate that fin whales most likely reside in the basin at least between June and October. Bryde's whales likely reside in the basin at least between March and June, but also appear to pass through the basin during other parts of the year. Antarctic blue whales seem to prefer to remain outside of the basin, and were identified on the only instrument located outside of the basin between June and October.

Although the presence of whale calls indicates that the respective whale is located in or near the basin, the absence of calls cannot be interpreted to mean that whales are not present. Calls may not have been recorded on the instrument due to propagation effects in the basin, or the whales located in and around the basin may have been emitting higher frequency calls than our instruments could detect.

Diel patterns indicate that fin and Bryde's whales vocalize more during daylight hours, but Antarctic blue whales vocalize more during the afternoon into dusk and at night. Spatial patterns were identified for each whale call type, although it is unclear what drives these patterns.

## Appendix A. Procedure for downloading and preparing data from IRIS

The data used in this experiment are archived in SEED format at the IRIS data management center (<http://www.iris.edu/dms/dmc/>). The following steps were used to download the data.

1. Request files from IRIS: [http://www.iris.edu/SeismiQuery/breq\\_fast.htm](http://www.iris.edu/SeismiQuery/breq_fast.htm)

Network = YL

Stations = A02W, A04W, A06W, A08W, A10W, A12W, A14W, B01W, B04W, B05W,  
B07W, B09W, B11W, B14W, C01W, C03W, C05W, C07W, C08W, C09W,  
C10W, C11W, C16W, C17W, F02W, N01W, N02W, N03W, S01W, S02W

Channels = HH1, HH2, HHZ, BDH

Start = 11-13-2009

End = 12-10-2010

Name = Dana Brodie

Institution = University of Hawaii at Manoa

Address = 1680 East-West Road, Honolulu, HI 96822

Email = dbrodie@hawaii.edu

Phone = 808-956-7640

Label = A02W\_HH1 (example for station A02W, channel HH1)

Media = electronic

**\*\*Start Query\*\***

Request “fullSEED”

**\*\*Submit\*\***

The user is sent to a page with a message similar to: “Your breq\_fast file ‘A02W\_HH1’ has been emailed to breq\_fast@iris.washington.edu”

2. Receive email with download instructions from IRIS.

3. Download SEED files to directory: /dbrodie/Research/KM1023/OBS\_DATA/[channel: HH1, HH2, HHZ, or BDH]

4. Decrypt SEED files:

Type the following into terminal:

```
“openssl enc -d -des-cbc -salt -in[FileName].openssl  
-out[NewFileName].seed -pass pass:<enter password>
```

5. Once decrypted, remove [FileName].openssl

6. Run msd2mat\_year2day.m

- change filename out to correct channel (line 111)
- change basepath (lines 19 and 20)
- change stat (line 25)

7. Move [FileName].mat into station directory, delete [FileName].msd files

8. Remove bad [FileName].mat files based on small files or non-continuous days.

## Appendix B. Instrument Responses

In this study, all ocean bottom seismometer instruments were of the same type, consisting of a Guralp CMG-3T three-component sensor and a Quanterra Q330 data logger. The instrument package was designed and constructed by the Woods Hole Oceanographic Institution for the U.S. Ocean Bottom Seismometer Instrument Pool. Instrument responses for these seismometers include the effects of the broadband sensor itself as well as anti-alias filters and gains applied by the data logger. The instrument responses are effectively the same across all instruments used in this study, with the exception of small differences in the applied gains (by  $\pm 3\%$ ). Instrument responses are provided by the instrument manufacturer and are stored with the data at the IRIS Data Management Center (<http://www.iris.edu/dms/dmc/>).

Instrument responses are composed of three stages. Stage 1 includes the sensor response composed of poles and zeros (**Table B-1**). The poles and zeros are polynomial roots that make up a transfer function (T) of the form:

$$T(s) = kg \frac{(s - z_1)(s - z_2)}{(s - p_1)(s - p_{21})(s - p_3)(s - p_4)(s - p_5)}$$

where zero values are the roots ( $z_1, z_2$ ) of the numerator, and the pole values are roots ( $p_1 - p_5$ ) of the denominator,  $k$  is the normalization factor,  $g$  is the gain applied in Stage 1 (**Table B-1**) and  $s$  is given by:

$$s = wj$$

where  $w$  is the angular frequency range and  $j$  is the imaginary unit ( $\sqrt{-1}$ ). Stage 2 consists of a constant gain applied across all frequencies (**Table B-1**). Stage 3 consists of an asymmetric finite impulse response (FIR) filter with 65 coefficients (**Table B-2**). For FIR filters, the output signal is a weighted sum of current and previous values of the input signal described by the equation:

$$y[n] = b_0x[n] + b_1x[n-1] + \dots + b_nx[n-N] = \sum_{i=0}^N b_i x[n-i]$$

where  $x[n]$  is the input signal,  $y[n]$  is the output signal, and  $b_i$  are the coefficients. To apply the Stage 3 coefficients, the transfer function of the FIR filter ( $H$ ) was determined using the summation equation:

$$H(j\phi) = \sum_{n=0}^N b_n (e^{j\phi})^{-n}$$

where  $b_n$  is the  $n^{\text{th}}$  coefficient,  $j$  is the imaginary unit ( $\sqrt{-1}$ ), and  $\phi$  is a complex argument given in radians and calculated from the equation:

$$\phi = 2\pi f / R$$

where  $f$  is frequency and  $R$  is the sampling rate of the instrument. To get a zero phase output from the FIR filter, the transfer function ( $H$ ) is multiplied by its conjugate and normalized:

$$H_{corrected} = N(H^* \text{conj}(H))$$

where  $N$  is the normalization factor calculated from:

$$N = \frac{|H|}{|H^* \text{conj}(H)|}$$

A gain of 1 counts per counts was applied in Stage 3. **Figure B-1** shows a breakdown of each stage of the amplitude and phase portions of the instrument response. Amplitude information in the plot is found by taking the magnitude of the complex response, and phase information is found by taking the argument of the complex response, unwrapping the angles (for example, MATLAB's® unwrap function), and converting from radians to degrees.

The total instrument response was determined by combining the three stages: the complex transfer function from Stage 1 ( $T$ ) was multiplied with the gain from Stage 2 and the complex transfer function from Stage 3 ( $H_{corrected}$ ). When multiplying the two frequency-dependent functions, it is important to make sure that they are both defined on the same discrete frequency

axis. **Figure B-2** shows the combined, total amplitude and phase responses. Only the high-frequency end of the response spectrum (1-50 Hz) is shown in **Figure B-2**, since that is the relevant range for marine mammal observations.

To remove the instrument response from the data, first the frequency axis of the total instrument response was unfolded so the zero-frequency component lies in the center of the spectrum to be compatible with Matlab's "fft" function (see MATLAB® function 'fftshift'). After detrending the data and applying a Tukey window with a 20% taper, the fast Fourier transform of the data was divided by the instrument response. When dividing the Fourier transform of the data by the instrument response, small values of the response (usually at the extreme ends of the response function) can cause the data to blow up at those frequencies – an unintended consequence of spectral division. Therefore, a constant value (sometimes referred to as a water level value) of 5% of the peak response amplitude was added to the total amplitude response across all frequencies to shift the response away from very small values. After spectral division, the data was converted back to a time series by taking the inverse fast Fourier transform of the result and clipping the result to the number of samples in the original times series (the fast Fourier transform pads the data with zeros to increase the vector length to the next power of 2 in size). When using MATLAB's® "ifft" function, the real part of the result must be taken, since numerical error adds a very small imaginary component.

A comparison of the uncorrected and corrected data in waveform and spectrograms did not show an appreciable difference. Frequencies between 45 and 50 Hz were more apparent; however, the frequency band focus of this study is between 15 and 40 Hz. Although the instrument response phase shift is ~70 degrees between 1 and 50 Hz (**Figure B-2**), the phase shift is ~27 degrees at the most common whale frequency (20 Hz), which equates to a time shift of only ~0.004 s. Therefore, correcting for instrument response is not necessary for this study.

Table B-1. Instrument response Stages 1 and 2 for OBS instrument #14

Stage	Poles	Zeros	Normalization	
			Factor	Gains*
1	-3.700800E-02+3.700800E-02i	0.000000+0.000000i	5.71521E+08	1.493970E+03
	-3.700800E-02-3.700800E-02i	0.000000+0.000000i		
	-5.026500E+02+0.000000E+00i			
	-1.005300E+03+0.000000E+00i			
	-1.131000E+03+0.000000E+00i			
2	NA	NA	NA	4.194300E+05

Notes:

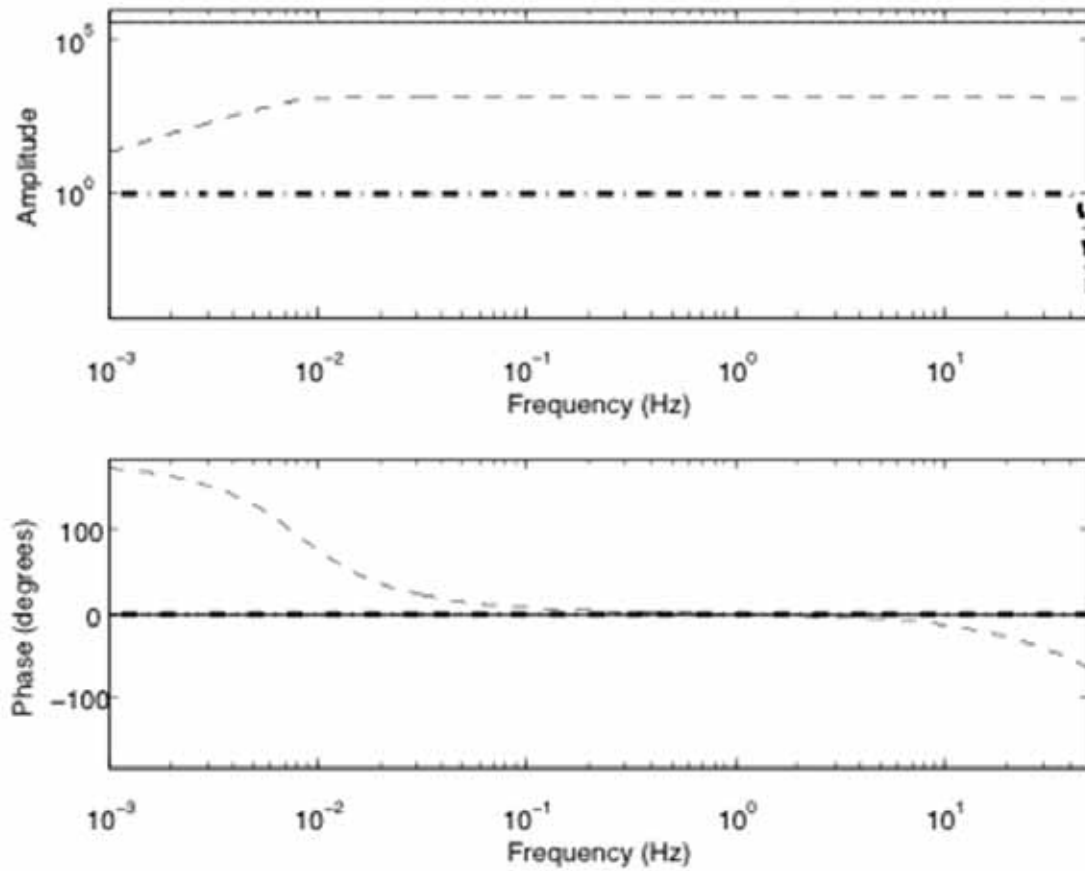
\* Stage 1 gain is given in (m/s)/volts and Stage 2 gain is given in volts/counts

NA = not applicable for that stage.

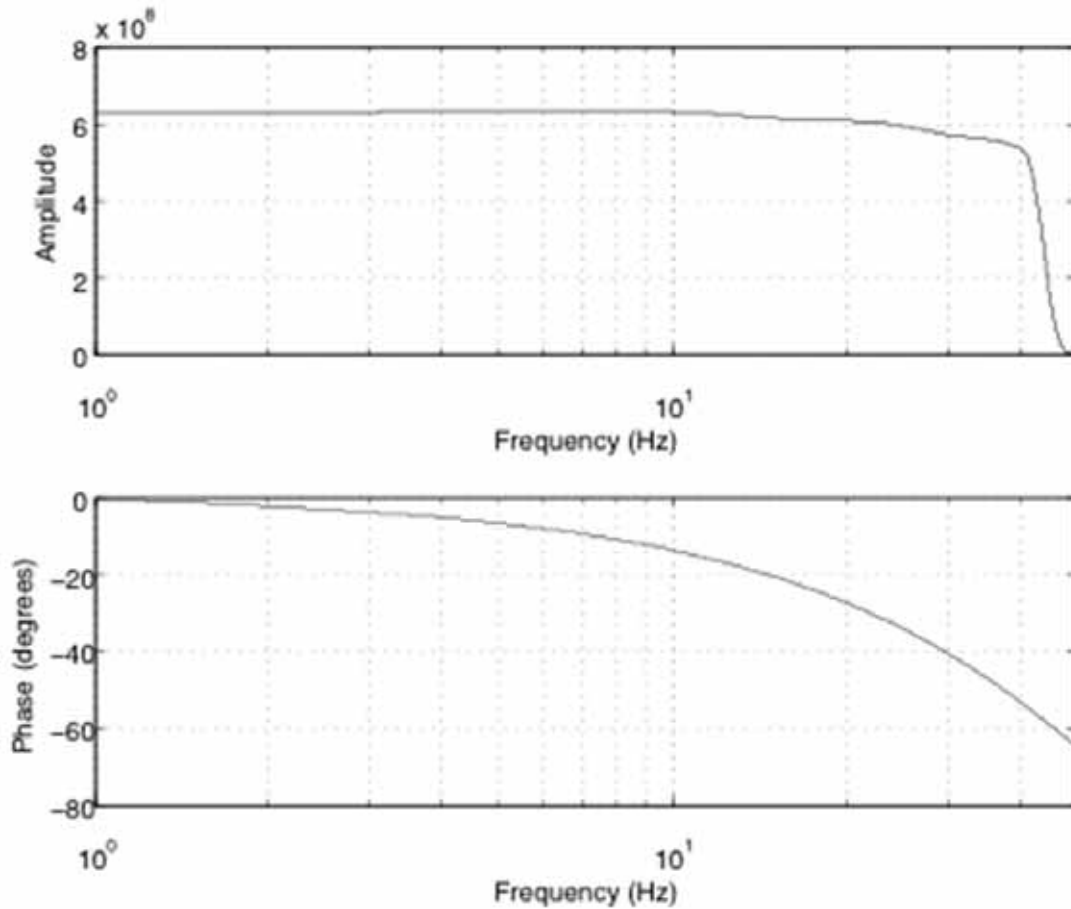
Table B-2. Instrument response Stage 3 coefficients for OBS instrument #14\*

n	Coefficient	n	Coefficient	n	Coefficient
0	6.437110E-14	22	-1.103990E-02	44	-1.214880E-02
1	6.582580E-08	23	8.813080E-03	45	1.102760E-02
2	-7.531180E-07	24	-3.077780E-03	46	-8.456630E-03
3	6.267430E-07	25	-4.648080E-03	47	5.394370E-03
4	2.470690E-06	26	1.601700E-02	58	-2.570720E-03
5	-3.152560E-06	27	-3.720880E-02	59	4.267280E-04
6	5.482400E-06	28	5.360130E-02	50	8.810930E-04
7	-7.781130E-06	29	-7.261630E-02	51	-1.413910E-03
8	5.105590E-06	30	8.785510E-02	52	1.379280E-03
9	3.315680E-05	31	-9.792760E-02	53	-1.029930E-03
10	-2.682380E-04	32	9.467820E-02	54	5.962990E-04
11	-3.530560E-04	33	8.888910E-01	55	-3.037780E-04
12	6.224890E-04	34	1.384220E-01	56	-2.987730E-04
13	-1.010060E-03	35	-1.153310E-01	57	3.410940E-05
14	1.272760E-03	36	9.511580E-02	58	5.269810E-06
15	-1.184310E-03	37	-7.431980E-02	59	-8.092690E-06
16	5.126970E-04	38	5.212640E-02	60	5.564590E-06
17	9.100910E-04	39	-3.396840E-02	61	-2.953830E-06
18	-3.090450E-03	40	1.228310E-02	62	2.261840E-06
19	5.811070E-03	41	-1.402390E-03	63	7.690850E-07
20	-8.577870E-03	42	-5.858370E-03	64	-7.497590E-07
21	1.063440E-02	43	1.076180E-02		

\* Stage 3 gain is 1 count per count



**Figure B-1.** The separate amplitude and phase components of the three stages that make up the total response of the instrument. Stage 1 (dashed) is the sensor response, which is effectively flat across the frequency band of the plot. A normalization factor and gain were applied in this stage. Stage 2 (solid) consists of a constant gain applied across all frequencies. Stage 3 (dash-dot) is an anti-alias filter that consists of an asymmetric FIR filter with 65 coefficients.



**Figure B-2.** The combined amplitude and phase response for the frequency band of interest. The amplitude begins to roll off slightly at  $\sim 10$  Hz due to the Stage 3 filter and decreases by  $\sim 10\%$  between 10 and 40 Hz. The amplitude then decreases by  $\sim 60\%$  between 40 and 45 Hz and  $\sim 99\%$  between 45 and 50 Hz. There is a phase shift of  $\sim 70$  degrees between 1 and 50 Hz.

## Appendix C. Background Noise

Background noise levels were compared for each station to determine if there were certain instruments or periods of time where background levels might bias call detection statistics. To compare background noise levels, the data were first corrected for the gains of the instrument responses to establish uniform recording levels (**Table C-1**). A few instruments exhibited large, widely separated in time, short-duration noise bursts of apparent electronic/instrument origin. While these spikes might not obstruct whale call detection to any significant degree, their large magnitudes (many orders of magnitude greater than the background noise) adversely affect the average power spectra of the data. Noise spikes that exceeded five standard deviations of each day of the data were zeroed using a 25-s cosine taper with 25% of the taper before the spike and 75% of the taper after the spike.

For each instrument, the average power spectrum for the entire recording period was computed. Then, to quantitatively compare these power spectra, the median of the power spectra ( $P_{\text{median}}$ ) was calculated between 15 and 40 Hz. Then the  $P_{\text{median}}$  values were all normalized by the minimum  $P_{\text{median}}$  value (station C17W) for comparison (**Table C-2**). Three stations (C05W, F02W, and S01W) had normalized  $P_{\text{median}}$  values greater than one order of magnitude higher than the minimum  $P_{\text{median}}$  value, and one station (C07W) had a normalized  $P_{\text{median}}$  value greater than two orders of magnitude higher than the minimum  $P_{\text{median}}$  value. Examples of median power are shown in **Figure C-1** and **Figure C-2** for cases with and without the noise bursts.

To investigate changes throughout the year, we computed and compared average power spectra for two-week intervals for each instrument. **Figure C-3** and **Figure C-4** show example power spectra for two instruments. For comparison, the median power in the 15-40 Hz band was computed for each two-week interval, and then all values were normalized by the lowest two-week median power value. Examples are shown in **Figure C-5** through **Figure C-7**. In these figures, the x-axis is one year long, January to December, and the 24<sup>th</sup> time period corresponds to the start of the data set towards the end of 2009 and 1<sup>st</sup> time period corresponds with the

beginning of 2010. **Figure C-8** shows the normalized power for each instrument for each two week period.

The two-week periods on most individual instruments were within  $<1$  to 2.5 dB of each other. In a few isolated cases, instruments had two-week periods that were  $\sim 4 - 5$  orders of magnitude apart. Three instruments (B11W, B14W, and C08W) initially had high noise levels in period 24. The difference between the average background noise level and the elevated noise level in period 24 ranged from 1 to 4 orders of magnitude, and is likely due to instrument noise created by the settling of the instrument after deployment and fewer days of data available during period 24. Additionally, four instruments (C05W, C07W, F02W, and S01W) exhibited consistently higher background noise over the entire recording period.

Several stations exhibited trends in the levels of background noise in the data set. One instrument (C07W) showed a decrease in noise at the end of the recording period (periods 19-21) of  $\sim 2$  orders of magnitude. Another instrument (F02W) showed a slight, gradual increase in noise of  $<1$  order of magnitude between periods 1 and 9, after which it remained relatively consistent. Two instruments (B01W and C05W) appeared to have slightly decreasing noise levels towards the middle of the year. Instrument B01W showed background noise decreasing by  $\sim 1$  order of magnitude between periods 1 and 12 later followed by a possible slight increasing trend in periods 20 and 21 (**Figure C-7**). Instrument C05W showed a decrease in background noise between periods 3 and 15 with  $<1$  order of magnitude difference. It is not clear why these stations show background noise trends. The patterns are not strictly seasonal, but may be related to environmental noise level changes nonetheless.

To determine if stations with relatively low background noise is a result of incorrect gains applied to those stations, we examined how each station recorded a large earthquake with a hypocenter at a distance of roughly 8020 km, along a great circle path. The selected 6.7 magnitude earthquake occurred in the Aleutian Islands ( $52.88^{\circ}\text{N}$ ,  $169.85^{\circ}\text{W}$ ) at a depth of 14 km on July 18, 2010 (day 199) at 5:56 UTC. The earthquake information was obtained from a NEIC earthquake search ([http://earthquake.usgs.gov/earthquakes/eqarchives/epic/epic\\_circ.php](http://earthquake.usgs.gov/earthquakes/eqarchives/epic/epic_circ.php)). The data for each station at the time of the earthquake was band pass filtered using a 0.0005 to 0.001 Hz filter. Amplitudes of the p-wave were measured from top peak to bottom peak of the initial

onset. These values were normalized against the mean value and were compared against the background noise level of the two-week period in which the earthquake occurred and against the total median background noise level for each station (**Table C-3**). No correlation was observed between background noise levels and p-wave amplitude, indicating that stations with low background noise do not necessarily have low detection abilities. Additionally, no spatial correlation in the study area was observed based on the p-wave amplitudes.

Table C-1. Instrument response gains.

Station	Normalization Factor	Stage 1 Gain ((m/s)/volts)	Stage 2 Gain (volts/count)	Stage 3 Gain (counts/count)	Total Gain ((m/s)/count)
A02W	5.71521E+08	1.47797E+03	4.19430E+05	1.00	3.54289E+17
A06W	5.71521E+08	1.49197E+03	4.19430E+05	1.00	3.57645E+17
A08W	5.71521E+08	1.48197E+03	4.19430E+05	1.00	3.55248E+17
A10W	5.71521E+08	1.49997E+03	4.19430E+05	1.00	3.59562E+17
A12W	5.71521E+08	1.48597E+03	4.19430E+05	1.00	3.56206E+17
A14W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
B01W	5.71521E+08	1.49197E+03	4.19430E+05	1.00	3.57645E+17
B04W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
B05W	5.71521E+08	1.48797E+03	4.19430E+05	1.00	3.56686E+17
B07W	5.71521E+08	1.49797E+03	4.19430E+05	1.00	3.59083E+17
B09W	5.71521E+08	1.48797E+03	4.19430E+05	1.00	3.56686E+17
B11W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
B14W	5.71521E+08	1.49797E+03	4.19430E+05	1.00	3.59083E+17
C01W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
C03W	5.71521E+08	1.49197E+03	4.19430E+05	1.00	3.57645E+17
C05W	5.71521E+08	1.49197E+03	4.19430E+05	1.00	3.57645E+17
C07W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
C08W	5.71521E+08	1.48197E+03	4.19430E+05	1.00	3.55248E+17
C09W	5.71521E+08	1.48997E+03	4.19430E+05	1.00	3.57165E+17
C10W	5.71521E+08	1.48197E+03	4.19430E+05	1.00	3.55248E+17
C11W	5.71521E+08	1.49397E+03	4.19430E+05	1.00	3.58124E+17
C16W	5.71521E+08	1.48797E+03	4.19430E+05	1.00	3.56686E+17
C17W	5.71521E+08	1.48197E+03	4.19430E+05	1.00	3.55248E+17
F02W	5.71521E+08	1.48597E+03	4.19430E+05	1.00	3.56206E+17
N01W	5.71521E+08	1.52397E+03	4.19430E+05	1.00	3.65316E+17
N02W	5.71521E+08	1.48597E+03	4.19430E+05	1.00	3.56206E+17
N03W	5.71521E+08	1.48997E+03	4.19430E+05	1.00	3.57165E+17
S01W	5.71521E+08	1.48597E+03	4.19430E+05	1.00	3.56206E+17
S02W	5.71521E+08	1.48197E+03	4.19430E+05	1.00	3.55248E+17

Table C-2. Normalized median power (NMP) between 15-40 Hz for each instrument with instrument response gains removed.

Station	NMP	Station	NMP
A02W	6.2	C05W	20.5
A06W	2.2	C07W	112.6
A08W	5.3	C08W	4.3
A10W	8.2	C09W	5.1
A12W	1.5	C10W	2.1
A14W	5.8	C11W	1.6
B01W	7.3	C16W	1.8
B04W	4.7	C17W	1.0
B05W	5.2	F02W	41.0
B07W	3.8	N01W	9.3
B09W	4.8	N02W	2.4
B11W	4.9	N03W	3.5
B14W	1.5	S01W	13.7
C01W	9.6	S02W	2.3
C03W	1.9		

Instrument C17W, with a median power value of  $7.25 \times 10^{-23}$  (nm/s)<sup>2</sup>/Hz was used for normalization.

Table C-3. Normalized p-wave amplitudes (NPWA) compared to background noise levels.

Station	NPWA	2-week background noise level <sup>1</sup>	Overall background noise levels <sup>2</sup>	Station	NPWA	2-week background noise level <sup>1</sup>	Overall background noise levels <sup>2</sup>
A02W	0.88	LOW-MED	LOW-MED	C05W	1.06	MED	MED-HIGH
A06W	0.98	LOW	LOW	C07W	0.50	HIGH	HIGH
A08W	0.61	LOW	LOW-MED	C08W	0.91	LOW-MED	LOW-MED
A10W	0.91	MED	LOW-MED	C09W	1.12	LOW-MED	LOW-MED
A12W	0.91	LOW	LOW	C10W	1.02	LOW	LOW
A14W	1.77	MED-HIGH	LOW-MED	C11W	1.00	LOW	LOW
B01W	0.98	LOW-MED	LOW-MED	C16W	1.04	LOW	LOW
B04W	1.06	LOW-MED	LOW-MED	C17W	0.99	LOW	LOW
B05W	1.33	LOW-MED	LOW-MED	F02W	0.88	HIGH	MED-HIGH
B07W	1.02	MED	LOW	N01W	1.01	MED-HIGH	MED
B09W	1.08	MED	LOW-MED	N02W	1.02	LOW	LOW
B11W	0.86	MED-HIGH	LOW-MED	N03W	1.03	LOW-MED	LOW
B14W	1.03	LOW	LOW	S01W	0.84	MED	MED
C01W	1.15	MED	MED	S02W	0.87	LOW	LOW
C03W	0.99	LOW	LOW				

The median p-wave amplitude value of 1607 was used for normalization.

<sup>1</sup> Two-week background noise levels for the period in which the earthquake occurred. Levels determined based on normalized median power values

LOW = <10

LOW-MED = 10-20

MED = 20-40

MED-HIGH = 40-100

HIGH = >100

<sup>2</sup> Overall median background noise levels for each station based on Table D2.

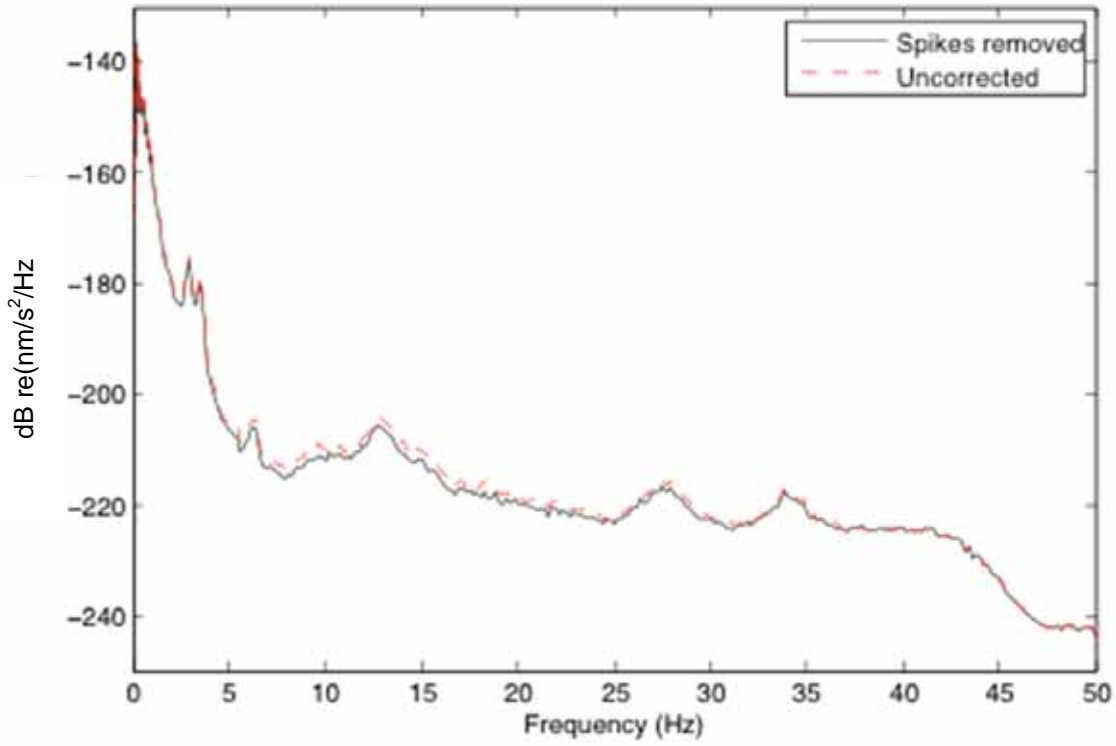
LOW = <4

LOW-MED = 4-9

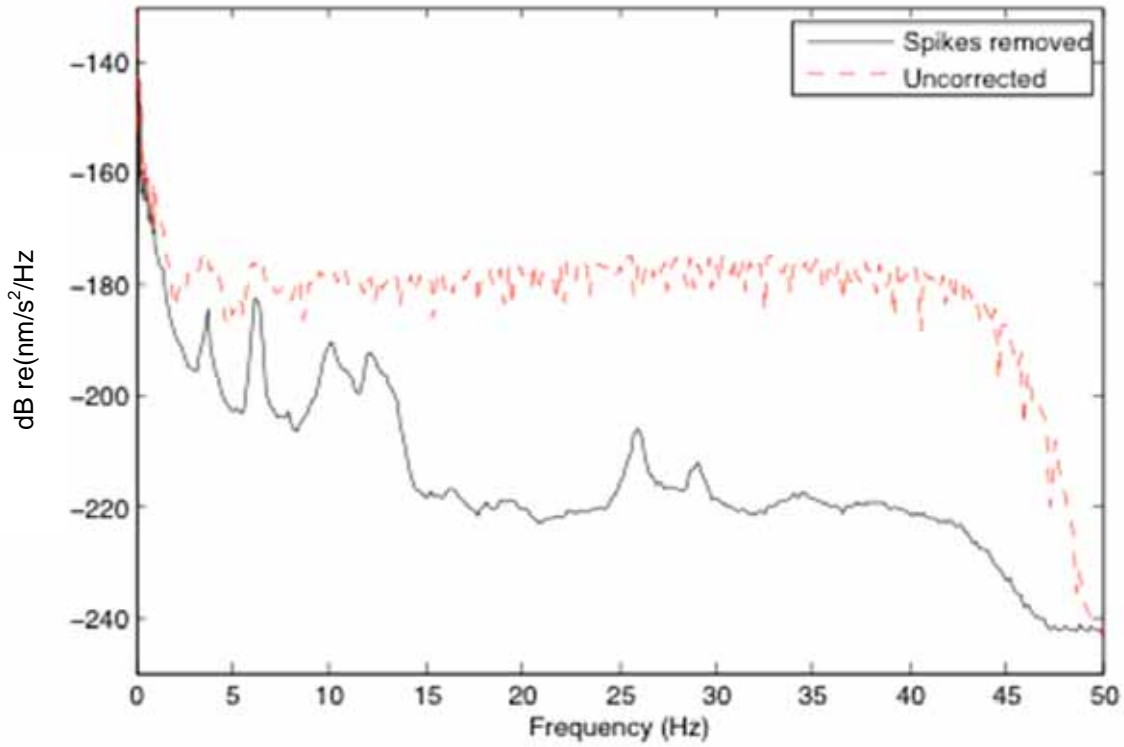
MED = 9-17

MED-HIGH = 17-50

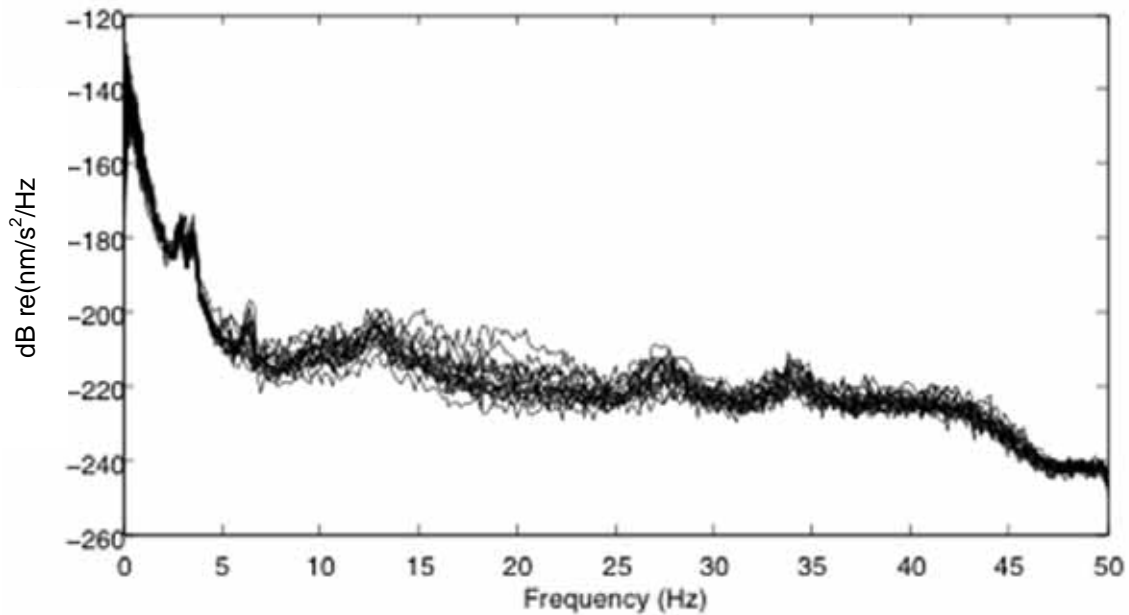
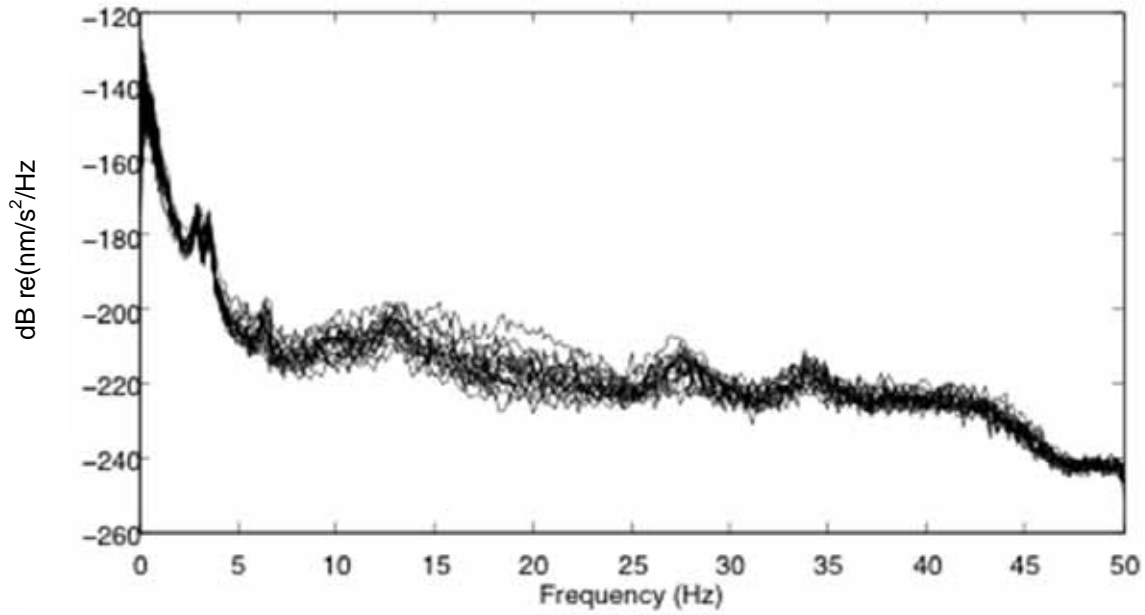
HIGH = >50



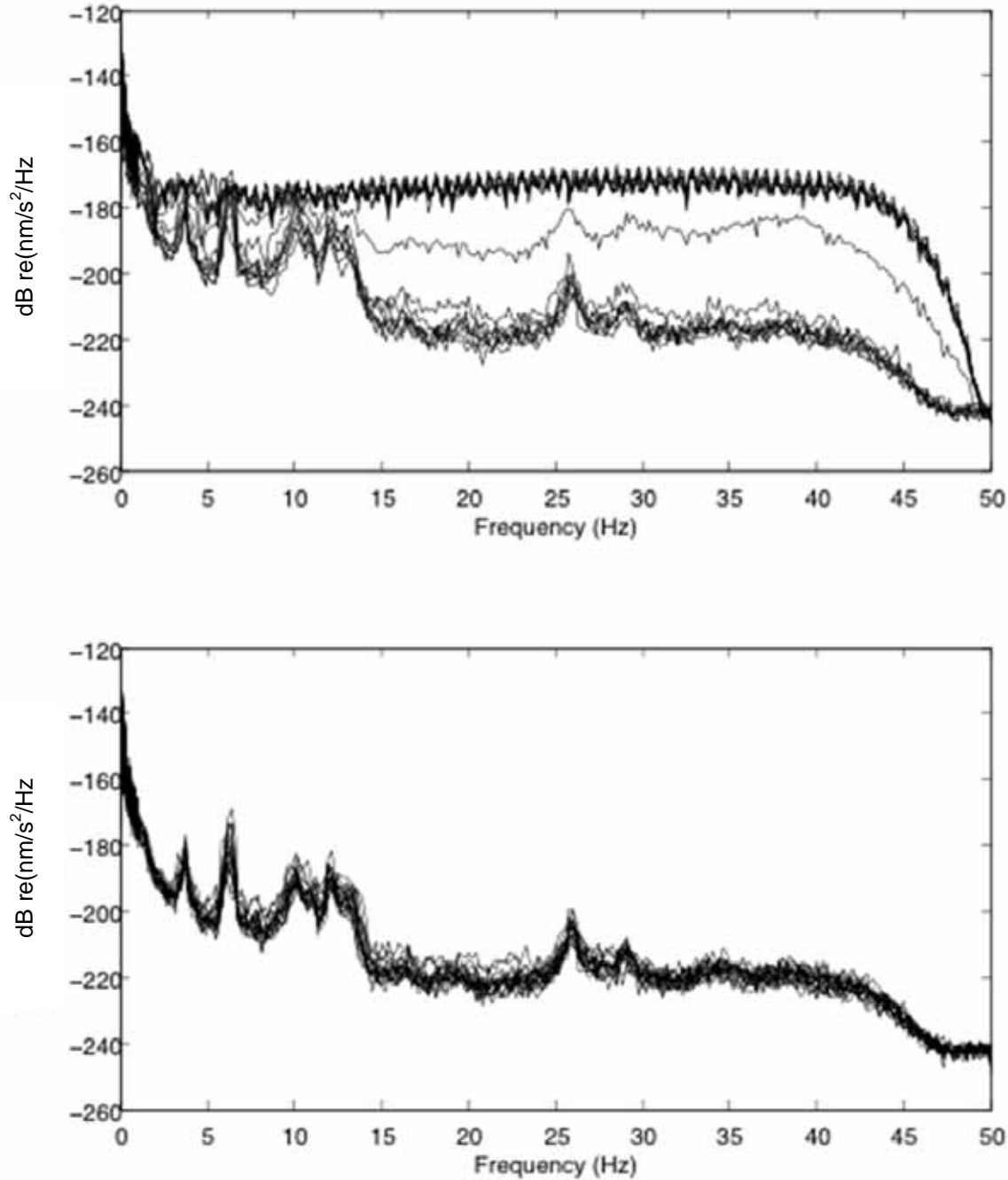
**Figure C-1.** Power versus frequency for instrument C17W showing relatively low background noise of both uncorrected data and the data with spikes removed using a cosine taper.



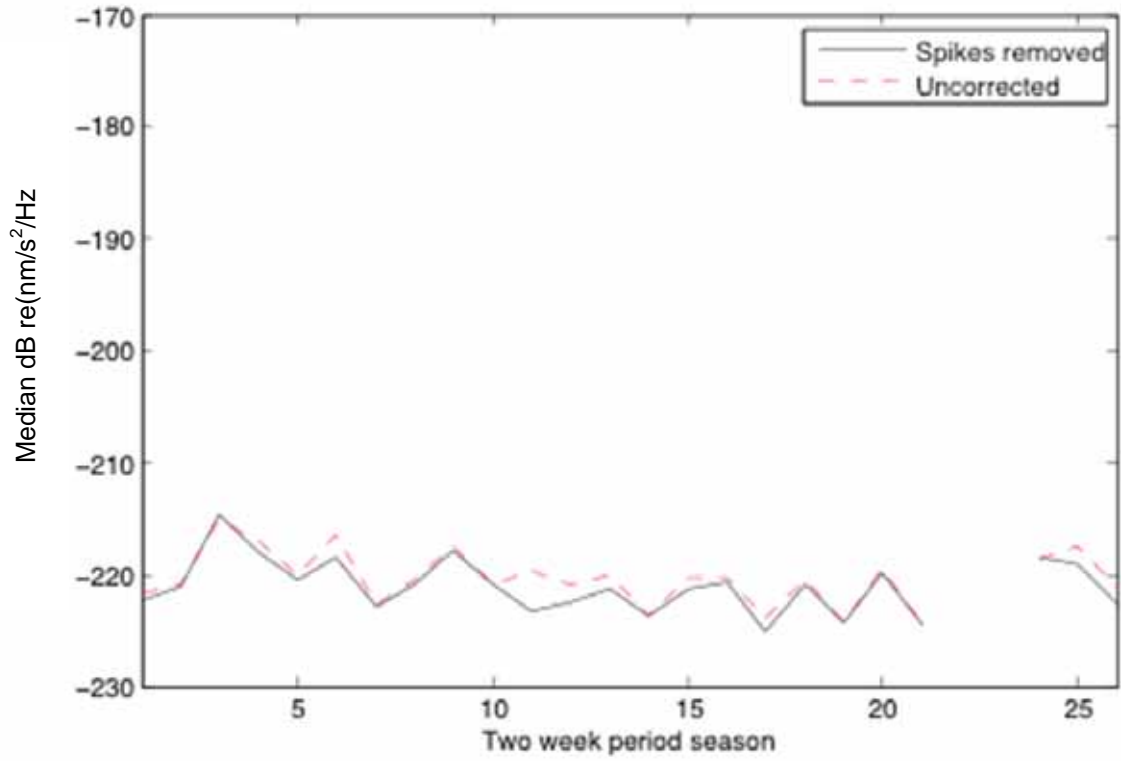
**Figure C-2.** Power versus frequency for instrument C11W showing high background noise in the uncorrected data and lower, more normal background noise in the data with spikes removed using a cosine taper.



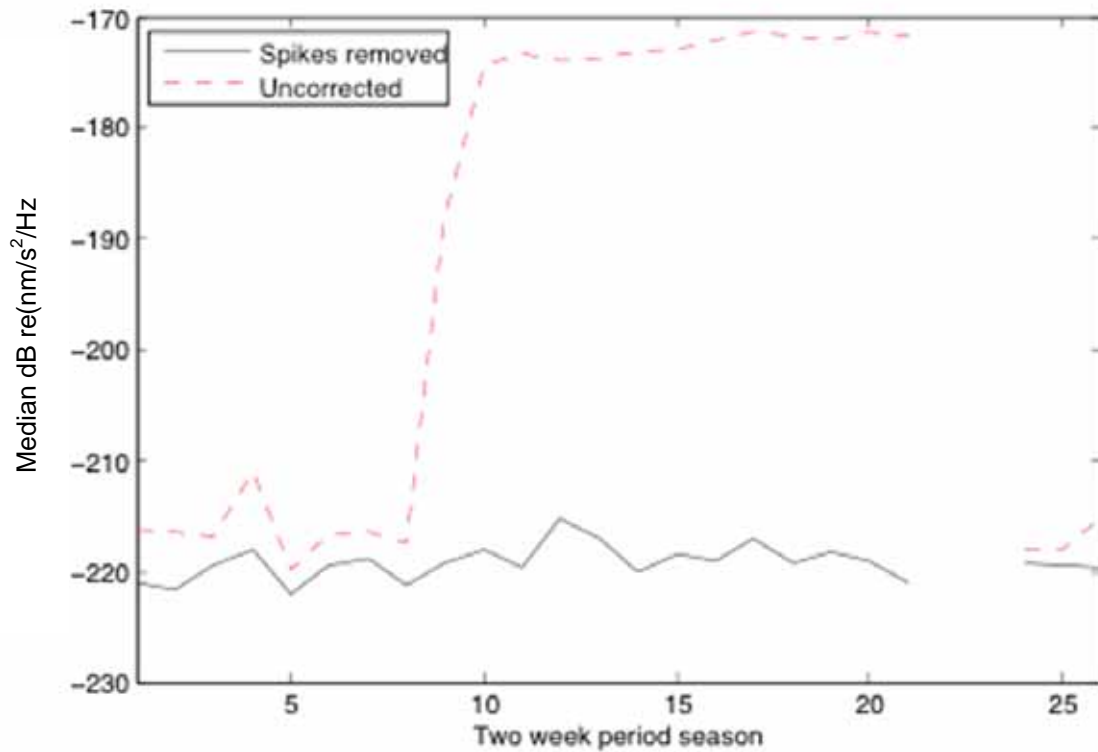
**Figure C-3.** Power versus frequency for instrument C17W showing all two-week intervals for the uncorrected data set (top) and the data set with spikes removed using a cosine taper (bottom). The differences in power range between  $\sim 1 - 3$  orders of magnitude.



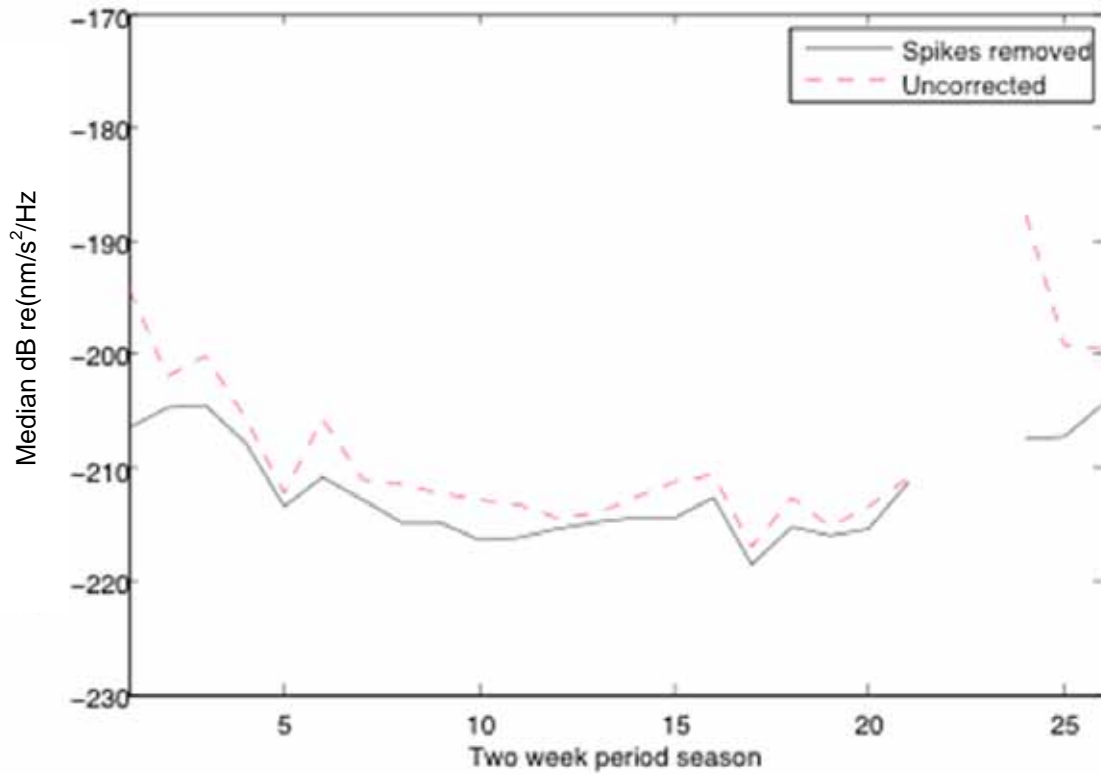
**Figure C-4.** Power versus frequency for instrument C11W showing all two-week period intervals for the uncorrected data set (top) and the data set with spikes removed using a cosine taper (bottom). The lines plotted at higher powers for the uncorrected data are for time periods 9 – 21 (days 114 to 295 of year 2010). This higher background noise was due to instrument-dependent electronic noise and the background noise level was reduced when the spikes were tapered to zero.



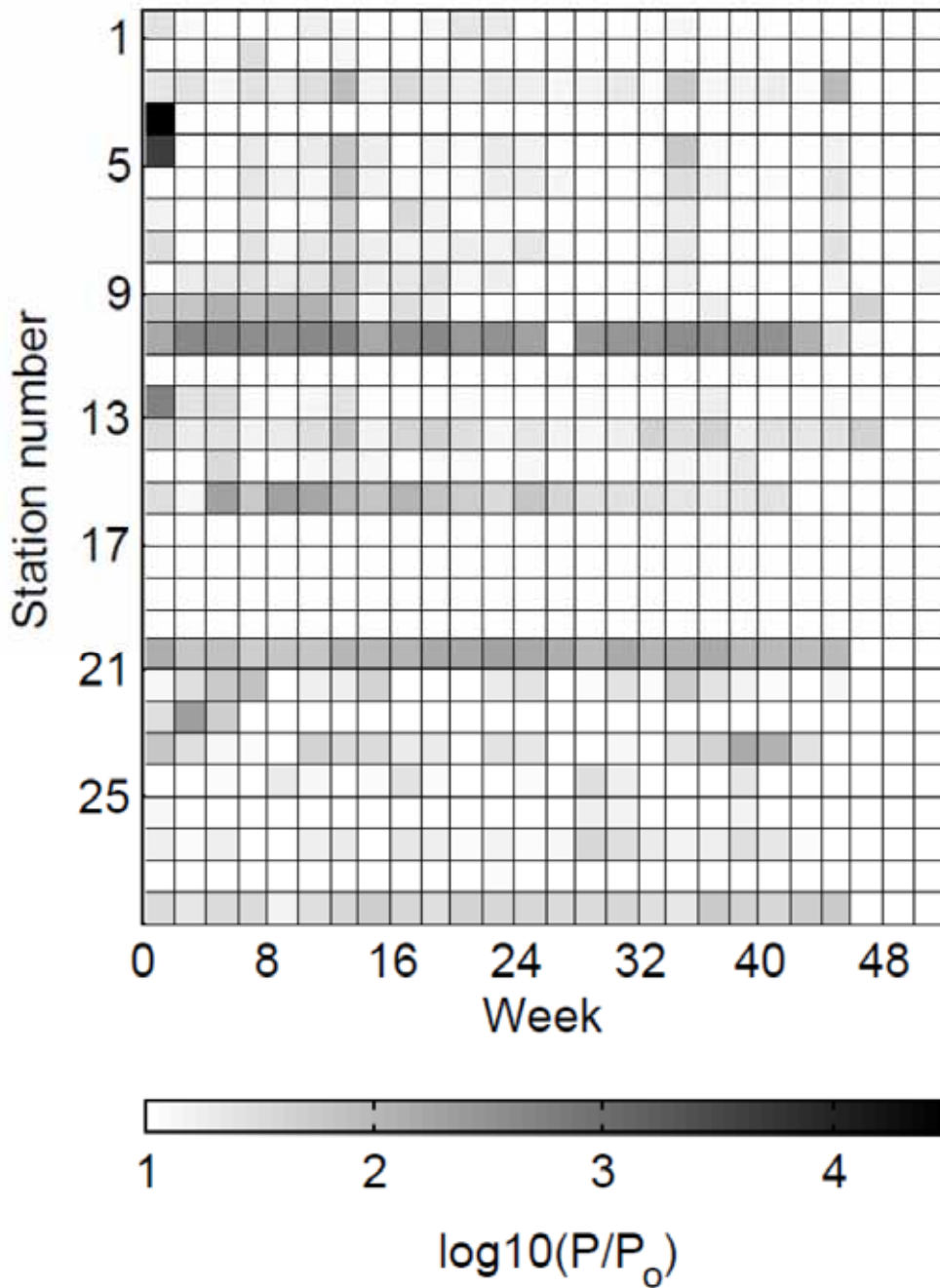
**Figure C-5.**  $P_{\text{median}}$  of each two-week period for instrument C17W showing little change in the background noise in both data sets. The plotted lines are not present during periods when data was unavailable.



**Figure C-6.**  $P_{\text{median}}$  of each two-week period for instrument C11W showing high levels of background noise beginning in the 9<sup>th</sup> two-week period and lasting until the end of the uncorrected data set and relatively little change in the data with spikes removed using a cosine taper (zeroed). The plotted lines are not present during periods when data was unavailable.



**Figure C-7.**  $P_{\text{median}}$  of each two-week period for instrument B01W for the uncorrected data set and the data set with spikes removed using a cosine taper. The uncorrected data set shows an initially high level of background noise at period 24 and a decreasing noise trend between periods 1 and 12 followed by a possible slight increasing trend between periods 20 and 21. The spike-zeroed data set shows a similar decreasing noise trend between periods 1 and 12 and small increase between periods 20 and 21 on a slightly smaller scale than the uncorrected data. The plotted lines are not present during periods when data was unavailable.



**Figure C-8.** Background noise patchwork plot showing normalized power for each OBS instrument and two week period. Power was normalized by the lowest two week period power. Darker shading indicates noisier periods. Although noise level at the ocean floor is known to correlate with wind speeds above the instruments, wind speed data is not available to determine if noise levels are primarily due to the wind.

Appendix D. OBS name and number correlation

<b>OBS name</b>	<b>OBS number</b>
A02W	1
A06W	2
A08W	3
A10W	4
A12W	5
A14W	6
B01W	7
B04W	8
B05W	9
B07W	10
B09W	11
B11W	12
B14W	13
C01W	14
C03W	15
C05W	16
C07W	17
C08W	18
C09W	19
C10W	20
C11W	21
C16W	22
C17W	23
F02W	24
N01W	25
N02W	26
N03W	27
S01W	28
S02W	29

## Appendix E. Algorithm Training Set

#	Sound Type	Instrument	Date	Time (UTC)
1	Bryde's	A14W	5/11/2010	22:42.8
2	Bryde's	C17W	4/25/2010	8:41.935
3	Fin	A14W	10/5/2010	2:05
4	Fin	A06W	7/23/2010	9:12.95
5	Fin	F02W	7/24/2010	17:49.48
6	Fin	N01W	7/14/2010	22:47.94
7	Antarctic Blue	F02W	6/4/2010	22:51.75
8	Antarctic Blue	F02W	6/5/2010	0:59.84
9	Antarctic Blue	F02W	6/4/2010	13:30
10	NZ Blue	F02W	6/1/2010	4:08.55
11	NZ Blue	A06W	8/29/2010	18:22.15
12	NZ Blue	C01W	6/3/2010	19:18.4
13	Tulip	B14W	5/25/2010	15:12.71
14	Tulip	B14W	1/29/2010	17:41.635
15	10s Upsweep	B14W	7/12/2010	8:32.075
16	10s Upsweep	C01W	6/17/2010	23:18.97
17	Doublet Upsweep	A06W	7/9/2010	4:26.54
18	100s Upsweep	A14W	6/23/2010	5:18.64
19	Crescent	C01W	8/6/2010	13:16.05
20	Crescent	C01W	8/6/2010	18:10.3
21	Background Noise	C08W	1/19/2010	19:19
22	Background Noise	C08W	1/10/2010	0:00
23	Background Noise	S02W	2/24/2010	23:59.5
24	Earthquake	C11W	8/29/2010	6:17.3
25	Earthquake	B01W	4/29/2010	1:14
26	Earthquake	C11W	8/29/2010	7:08.2
27	Earthquake	B01W	4/29/2010	1:14.15
28	Spike	C07W	5/11/2010	10:00.47
29	Spike	A08W	4/19/2010	9:10.56
30	Spike	A06W	11/30/2009	10:18.26
31	Instrument Noise	C01W	12/2/2009	15:28.25
32	Instrument Noise	C08W	11/30/2009	5:29.25
33	Instrument Noise	C11W	11/29/2009	6:25.95
34	Instrument Noise	C11W	11/29/2009	11:04
35	Instrument Noise	C11W	12/1/2009	15:21
36	Instrument Noise	B01W	12/3/2009	10:49
37	Instrument Noise	C11W	11/29/2009	11:00.28

#	Sound Type	Instrument	Date	Time (UTC)
38	Chatter	C01W	8/29/2010	6:50.4
39	Chatter	C01W	8/29/2010	22:06
40	Chatter	C01W	8/29/2010	16:56.5
41	Downsweep	F02W	7/5/2010	19:01.7
42	Three bar	N01W	4/12/2010	22:51.9
43	Bryde's	C01W	3/7/2010	11:57.932
44	Fin	A14W	10/5/2010	2:03.17
45	Fin	A14W	10/5/2010	3:06.28
46	Antarctic Blue	C01W	4/13/2010	0:54.65
47	Antarctic Blue	C01W	4/13/2010	0:06.95
48	Instrument Noise	C01W	4/13/2010	0:19.5
49	Instrument Noise	A14W	10/5/2010	0:23.4
50	Background Noise	A14W	10/5/2010	0:40.22
51	Background Noise	A14W	10/5/2010	3:04.9
52	Spike	F02W	6/5/2010	9:44.35
53	Spike	F02W	6/4/2010	9:25.585
54	Spike	A14W	5/11/2010	0:52.098
55	Spike	C17W	4/8/2010	5:16.915
56	Tulip	B14W	5/25/2010	15:11.435
57	Tulip	B14W	12/9/2009	8:50.9
58	10s Upsweep	A14W	6/23/2010	4:39.85
59	Doublet Upsweep	A14W	9/20/2010	22:34.6
60	100s Upsweep	A14W	6/23/2010	5:18.65
61	Crescent	C01W	8/6/2010	15:09.5
62	Crescent	C01W	8/6/2010	19:22.4
63	Background Noise	N01W	3/29/2010	0:54.7
64	Background Noise	S02W	7/19/2010	7:07.1
65	Spike	S02W	7/19/2010	7:41.85
66	Spike	A14W	10/5/2010	15:09.65
67	Spike	A14W	10/5/2010	0:40.4
68	Instrument Noise	A14W	10/5/2010	9:53.4
69	Instrument Noise	S02W	7/19/2010	1:07.2
70	Instrument Noise	S02W	7/19/2010	3:31.25
71	Instrument Noise	N01W	3/29/2010	1:26
72	Instrument Noise	N01W	3/29/2010	1:03
73	Antarctic Blue	F02W	6/4/2010	5:37.2
74	Spike	A08W	4/4/2010	11:56.9
75	Spike	A08W	4/4/2010	16:58.63
76	Bryde's	A08W	4/4/2010	1:26.2

#	Sound Type	Instrument	Date	Time (UTC)
77	Spike	B14W	7/8/2010	2:05.1
78	Spike	B14W	7/8/2010	2:35.8
79	Background Noise	B14W	7/8/2010	2:46.1
80	Spike	B14W	7/8/2010	5:41.48
81	Antarctic Blue	F02W	6/5/2010	2:52.4
82	Antarctic Blue	F02W	6/5/2010	3:36.75
83	CH	C01W	8/18/2010	1:02.75
84	CH	C01W	8/18/2010	2:07.75
85	CH	C01W	8/18/2010	2:48.6
86	Spike	C01W	8/18/2010	21:25.55
87	Spike	C01W	8/18/2010	16:33.45
88	Spike	C01W	6/3/2010	10:38.88
89	Background Noise	C01W	6/3/2010	8:03.55
90	Background Noise	C01W	6/3/2010	12:15.5
91	Earthquake	C01W	6/3/2010	13:23.85
92	Background Noise	C01W	6/3/2010	16:15.6
93	Bryde's	F02W	4/8/2010	11:15.08
94	Bryde's	A06W	3/27/2010	14:32.39
95	Antarctic Blue	F02W	6/5/2010	6:38.45
96	Antarctic Blue	F02W	6/5/2010	8:54.85
97	Background Noise	A02W	2/23/2010	1:16.35
98	Background Noise	B14W	6/8/2010	2:38.4
99	Background Noise	B14W	6/8/2010	4:13.65
100	Background Noise	A02W	2/23/2010	1:02.5
101	Background Noise	A02W	2/23/2010	4:55.15
102	Background Noise	A02W	2/23/2010	0:18.5
103	Background Noise	A02W	2/23/2010	1:07.98
104	Instrument Noise	B14W	6/8/2010	6:08.75
105	Instrument Noise	B14W	6/8/2010	11:49.5
106	Spike	B14W	6/8/2010	9:43.35
107	Spike	A02W	2/23/2010	3:38.3
108	Bryde's	C16W	3/30/2010	6:44.2
109	Bryde's	C16W	3/30/2010	7:00.88
110	Spike	C03W	4/22/2010	23:22.2
111	Instrument Noise	A06W	9/7/2010	4:33.3
112	Instrument Noise	A06W	9/7/2010	4:38.38
113	Instrument Noise	A06W	9/7/2010	4:38.24

## Appendix F. Detection and classification algorithm

This appendix provides the details of the principal component-multinomial logistic regression algorithm used in this study.

### 1. Preprocessing

As explained in the main text, a training set consisting of acoustic sound examples containing particular whale calls, earthquakes, and other acoustic phenomena notable in the data was assembled based on the visual observations discussed in Chapter 2. Each example sound was created to be 25 seconds long, from data that had been normalized to reduce background noise as explained in Chapter 2. Then, the logarithm of the spectrogram was taken and vectorized by rearranging its columns into a single long column. Each example sound in the training set was then a column vector of signal-to-noise ratios at different times and frequencies. Finally, each vector was shifted to have zero mean.

### 2. A basis of EOFs

The  $N$  preprocessed columns in the training set were assembled side by side in a large matrix  $A$ . The covariance matrix  $C = N^{-1}AA^T$  was then written as  $U\Lambda U^T$  where  $U = [u_{(1)}, u_{(2)}, \dots, u_{(N)}]$  is a matrix of orthonormal eigenvectors of  $C$  and  $\Lambda = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_N]$  is a diagonal matrix of eigenvalues. These eigenvectors are the empirical orthogonal functions (EOFs) of the training set. The EOFs with the  $p$  largest eigenvalues explain most of the covariance, and thus it is conventional to think of the partition  $U = [U_p, U_0]$ ,  $\Lambda = \text{diag}[\Lambda_p, \Lambda_0]$  and

$C = U_p \Lambda_p U_p^T + U_0 \Lambda_0 U_0^T$ . The  $p$  EOFs in  $U_p$  are the EOFs that we will use for detection and classification. In our procedure  $p$  was found by trial and error, and our results are insensitive to its value provided that it is large enough. A rule of thumb is that  $p$  should be large enough to explain 90% of the data variance, i.e.,  $\sum_{i=1}^p \lambda_i \geq 0.9 \sum_{i=1}^N \lambda_i$ .

### 3. Acoustic signatures

Each sound in the training matrix A was individually processed given the system of linear equations were:

$$\ln\left(\frac{P_{a1}}{P_0}\right) = m^T u_p^T a_{(1)}$$

$$\ln\left(\frac{P_{a2}}{P_0}\right) = m^T u_p^T a_{(2)}$$

...

$$\ln\left(\frac{P_{an}}{P_0}\right) = m^T u_p^T a_{(n)}$$

In this system of equations, which can also be written as  $q = B^T U_p m$ , the quantity  $P_{a1}$  is the estimated probability that  $a_{(1)}$  is the actual example sound type, and  $P_0$  is the estimated probability that  $a_{(1)}$  is an unidentified sound (something for which there is no category). The vector  $m$  is the acoustic signature of select training set sound, and can be estimated by solving the system of linear equations in the least squares sense:

$$\hat{m} = (U_p^T B B^T U_p)^{-1} U_p^T B q.$$

This procedure was repeated to get a signature for example sound in the training set.

#### 4. Validation

After the signatures of each sound was created, the detection/classification procedure was tested on a validation set of sounds consisting of examples from each sound category, but not the same exact sounds that were included in the training set. As the validation procedure is similar to the detection and classification procedure, only the latter is discussed here.

#### 5. Detection and classification

In order to apply the algorithm, the data stream was transformed into EOF space just as was done with the training set. Each day of the data set was divided into 25 second windows with a 3 second step size, so that 25 second windows were created and overlapped each other. For each noise window, the logarithm of the spectrogram was taken. Then it was vectorised and normalized by remove the mean. The result was a column vector  $d$  for each time  $t$ , referred to as  $d(t)$ . This vector  $d(t)$  was multiplied by  $U_p^T$  to transform it into a vector  $\psi(t)$  in EOF space. Then the relative probability of detection in categories  $\{1, 2, \dots, c\}$  at time  $t$  was estimated by substituting the signatures found above into the following equations.

$$P_1 / P_0 = e^{m_{(1)}^T \psi}; P_2 / P_0 = e^{m_{(2)}^T \psi}; \dots; P_c / P_0 = e^{m_{(c)}^T \psi}.$$

However, the probability of the ‘none of the above’ category,  $P_0$ , is still unknown. To obtain  $P_0$  we observe that any time  $t$  it is always true that  $P_1 + \dots + P_c + P_0 = 1$ , from which it follows that

$$\frac{P_1}{P_0} + \frac{P_2}{P_0} + \dots + \frac{P_c}{P_0} + \frac{P_0}{P_0} = \frac{1}{P_0}.$$

This last equation may be rewritten as  $e^{m_{(1)}^T \psi} + e^{m_{(2)}^T \psi} + \dots + e^{m_{(c)}^T \psi} + 1 = P_0^{-1}$ , and combining the above results gives:

$$\begin{aligned} P_0 &= \frac{1}{1 + \sum_{i=1}^c \exp(m_{(i)}^T \psi)} \\ P_1 &= \frac{\exp(m_{(1)}^T \psi)}{1 + \sum_{i=1}^c \exp(m_{(i)}^T \psi)} \\ &\dots \\ P_c &= \frac{\exp(m_{(c)}^T \psi)}{1 + \sum_{i=1}^c \exp(m_{(i)}^T \psi)}. \end{aligned}$$

Recalling that  $\psi = \psi(t)$  we have a continuous stream of probability channels  $P_j(t)$  that always sum to unity. Whenever  $P_j(1 - P_j)^{-1}$  exceeds a certain threshold we declared a detection in category  $j$ .

## **6. Multiple sound types from a single species**

In multinomial logistic regression, there is an assumption known as the independence of irrelevant alternatives, the practical meaning of which is that it is important not to overwork the none-of-the-above category  $P_0$ . In particular, if a whale species has, say, two types of calls, then a different category should be used for each type of call. If the resulting two categories are numbered 1 and 2, for example, then the probability of that particular whale is  $P_1 + P_2$ , and so we declare a detection when  $P_1 + P_2$  is the top probability and exceeds the set probability threshold.

## REFERENCES

- Berchok, C.L., Bradley, D.L., and Gabrielson, T.B. (2006). "St. Lawrence blue whale vocalizations revisited: characterization of calls detected from 1998 to 2001," *J. Acoust. Soc. Am.* **120**(4), 2340-2354.
- Best, P.B. (2001). "Distribution and population separation of Bryde's whales *Balaenoptera edeni* off southern Africa," *Mar. Ecol. Prog. Ser.* **220**, 277-289.
- Borsa P., and Hoarau, G. (2004). "A pygmy blue whale (Cetacea: Balaenopteridae) in the Inshore Waters of New Caledonia," *Pacific Science* **58**, 579-584.
- Branch, T.A., Stafford, K.M., Palacios, D.M., Allison, C., Bannister, J.L., et al. (2007). "Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the Southern Hemisphere and northern Indian Ocean," Publications, Agencies and Staff of the U.S. Department of Commerce Paper 103.
- Caplan-Auerbach, J., Dziak, R.P., Bohnenstiehl, D.R., Chadwick, W.W., and Lau, T.-K. (2014). "Hydroacoustic investigation of submarine landslides at West Mata volcano, Lau Basin," *Geophys. Res. Lett.*, **41**, 5927-5934.
- Chapp, E., Bohnenstiehl, D.R., and Tolstoy, M. (2005). "Sound-channel observations of ice-generated tremor in the Indian Ocean," *Geochem. Geophys. Geosys.* **6**, Q06993.
- Clark, C.W. and Fowler, M. (2001). "Status of archival and analysis effort of acoustic recordings during SOWER and IWC cruises 1996-2000," *Rep. Int. Whal. Comm. SC/53/IA28*.
- Clark, C.W. and Fristrup, K.M. (1997). "Whales '95: A combined visual and acoustic survey of blue and fin whales off southern California," *Rep. Int. Whal. Comm.* **47**, 583-600.
- Cummings, W.C. (1985). "Bryde's whale, *Balaenoptera edeni*," In *Handbook of Marine Mammals*, edited by S.H. Ridgway and R. Harrison (Academic Press, London), pp. 137-154.

- Dawbin, W.H. (1966). "The seasonal migratory cycle of humpback whales," In *Whales, Dolphins, and Porpoises*, edited by K.S. Norris (University of California Press, Berkeley and Los Angeles), pp. 145-170.
- Dunn, R.A. and Hernandez, O. (2009). "Tracking blue whales in the eastern tropical Pacific with an ocean-bottom seismometer and hydrophone array," *J. Acoust. Soc. Am.* **126**, 1084-1094.
- Dziak, R.P., Haxel, J.H., Matsumoto, H., Lau, T.K., Merle, S.G., de Ronde, C.E.J., Embley, R.W., and Mellinger, D.K. (2008). "Observations of regional seismicity and local harmonic tremor at Brothers volcano, south Kermadec arc, using an ocean bottom hydrophone array," *J. Geophys. Res.* **113**, B08S04.
- Dziak, R., Bohnenstiehl D., Baker, E., Matsumoto, H., Caplan-Auerbach, J., Mack, C., Embley, R., Merle, S., Walker, S., and Lau, T-K. (2013) "Acoustic and tephra records of explosive eruptions at West Mata submarine volcano, NE Lau Basin," Abstract V21C-2733 presented at the Fall Meeting, AGU, San Francisco CA.
- Garland, E.C., Noad, M.J., Goldizen, A.W., Lilley, M.S., Rekdahl, M.L., Garrigue, C., Constantine, R., Hauser, N.D., Poole, M.M., and Robbins, J. (2013). "Quantifying humpback whale song sequences to understand the dynamics of song exchange at the ocean basin scale," *J. Acoust. Soc. Am.* **133**(1), 560-569.
- Gaskin, D.E. (1977). "Sei and Bryde's whales in waters around New Zealand," Report of the International Whaling Commission (Special Issue 1), 50-52.
- Heimlich, S.L., Mellinger, D.K., Nieukirk, S.L., and Fox, C.G. (2005). "Types, distribution, and seasonal occurrence of sounds attributed to Bryde's whales (*Balaenoptera edeni*) recorded in the eastern tropical Pacific, 1999-2001," *J. Acoust. Soc. Am.* **118**(3), 1830-1837.
- Helble, T.A., D'Spain, G.L., Hildebrand, J.A., and Campbell, G.S. (2013). "Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones," *J. Acoust. Soc. Am.* **134**(3), 2556-2570.

- Helweg, D.A., Cato, D.H., Jenkins, P.F., Garrigue, C., and McCauley, R.D. (1998). "Geographic variation in South Pacific humpback whale songs," *Behaviour* **135**, 1-27.
- Kerosky, S.M., Širović, A., Roche, L.K., Baumann-Pickering, S., Wiggins, S.M., and Hildebrand, J.A. (2012). "Bryde's whale seasonal range expansion and increasing presence in the Southern California Bight from 2000 to 2010," *Deep-Sea Research I* **65**, 125-132.
- Kibblewhite, A.C., Denham, R.N., and Barnes, D.J. (1967). "Unusual low-frequency signals observed in New Zealand waters," *J. Acoust. Soc. Am.* **41**(3), 644-655.
- Ljungblad, D.K., Clark, C.W., and Shimada, H. (1998). "A comparison of sounds attributed to pygmy blue whales (*Balaenoptera musculus brevicauda*) recorded south of the Madagascar Plateau and those attributed to 'true' blue whales (*Balaenoptera musculus*) recorded off Antarctica," *Rep. Int. Whal. Comm. SC/49/SH17*.
- Mackintosh, N.A. (1966). "The distribution of southern blue and fin whales." In: *Whales, Dolphins, and Porpoises*, edited by K.S. Norris (University of California Press, Berkeley), pp. 125-144.
- McDonald, M.A., Hildebrand, J.A., and Webb, S.C. (1995). "Blue and fin whales observed on a seafloor array in the Northeast Pacific," *J. Acoust. Soc. Am.* **98**(2), 712 – 721.
- McDonald, M.A. and Fox, C.G. (1999). "Passive acoustic methods applied to fin whale population density estimation," *J. Acoust. Soc. Am.* **105**, 2643-2651.
- McDonald, M.A. (2006). "An acoustic survey of baleen whales off Great Barrier Island, New Zealand," *New Zealand Journal of Marine and Freshwater Research* **40**, 519-529.
- McDonald, M.A., Mesnick, S.L., and Hildebrand, J.A. (2006). "Biogeographic characterisation of blue whale songs worldwide: using song to identify populations," *J. Cetacean Res. Manage.* **8**(1), 55-65.
- Mellinger, D.K., and Clark, C.W. (2003). "Blue whale (*Balaenoptera musculus*) sounds from the North Atlantic," *J. Acoust. Soc. Am.* **114**(2), 1108-1119.

- Miller, B.S., Collins, K., Barlow, J., Calderan, S., Leaper, R., McDonald, M., Ensor, P., Olson, P., Olavarria, C., and Double, M.C. (2014). "Blue whale vocalizations recorded around New Zealand 1964 - 2013," J. Acoust. Soc. Am **135**(3), 1616-1623.
- Ohsumi, S. (1978). "Provisional report on the Bryde's whales caught under special permit in the Southern Hemisphere," Report of the International Whaling Commission, SC/29/Doc 38.
- Ohsumi, S. (1979). "Provisional report of the Bryde's whales caught under special permit in the Southern Hemisphere in 1977/78 and a research programme for 1978/79," Report of the International Whaling Commission, SC/30/Doc30.
- Oleson, E.M., Barlow, J., Gordon, J., Rankin, S., and Hildebrand, J.A. (2003). "Low frequency calls of Bryde's whales," Mar. Mamm. Sci. **19**(2), 407-419.
- Oleson, E.M., Calambokidis, J., Burgess, W.C., McDonald, M.A., LeDuc, C.A., and Hildebrand, J.A. (2007). "Behavioral context of call production by eastern North Pacific blue whales," Mar. Ecol. Prog. Ser. **330**, 269-284.
- Omura, H. (1962). "Further information on Bryde's whale from the coast of Japan," Sci. Rep. Whales Res. Inst., Tokyo **16**, 7-18.
- Pangerc, T. (2010). "Baleen whale acoustic presence around South Georgia," PhD thesis, University of East Anglia. [<http://ueaeprints.uea.ac.uk/19109>] accessed on 6 June 2013.
- Rankin, S., Ljungblad, D., Clark, C., and Kato, H. (2005). "Vocalisations of Antarctic blue whales, *Balaenoptera musculus intermedia*, recorded during the 2001/2002 and 2002/2003 IWC/SOWER circumpolar cruises, Area V, Antarctica," J. Cetacean Res. Manage. **7**(1), 13-20.
- Rankin, S. and Barlow, J. (2007). "Vocalizations of the Sei whale *Balaenoptera Borealis* off the Hawaiian Islands," Bioacoustics **16**, 137-145.

- Samaran, F., Guinet, C., Adam, O., Motsch, J-F., and Cansi, Y. (2010). "Source level estimations of two blue whale subspecies in southwestern Indian Ocean," *J. Acoust. Soc. Am.* **127**(6), 3800-3808.
- Sasaki, T., Nikaido, M., Wada, S., Yamada, T.K., Cao, Y., Hasegawa, M., and Okada, N. (2006). "Balaenoptera omurai is a newly discovered baleen whale that represents an ancient evolutionary lineage," *Molecular Phylogenetics and Evolution* **41**, 40-52.
- Širović, A., Hildebrand, J.A., Wiggins, S.M., McDonald, M.A., Moore, S.E., and Thiele, D. (2004). "Seasonality of blue and fin whale calls and the influence of sea ice in the Western Antarctic Peninsula," *Deep-Sea Research II* **51**, 2327-2344.
- Širović, A., Hildebrand, J.A., Wiggins, S.M., and Thiele, D. (2009). "Blue and fin whale acoustic presence around Antarctica during 2003 and 2004," *Mar. Mamm. Sci.* **25**(1), 125-136.
- Širović, A., Williams, L.N., Kerosky, S.M., Wiggins, S.M., and Hildebrand, J.A. (2013). "Temporal separation of two fin whale call types across the eastern North Pacific," *Mar. Bio.* **160**(1), 47-57.
- Širović, A., Bassett, H.R., Johnson, S.C., Wiggins, S.M., and Hildebrand, J.A. (2014). "Bryde's whale calls recorded in the Gulf of Mexico," *Mar. Mamm. Sci.* **30**(1), 399-409.
- Secretariat of the Pacific Regional Environmental Programme (SPREP) (2007). "Pacific Islands Regional Marine Species Programme 2008-2012." Apia, Samoa.
- Stafford, K.M., Bohnenstiehl, D.R., Tolstoy, M., Chapp, E., Mellinger, D.K., and Moore, S.E. (2004). "Antarctic-type blue whale calls recorded at low latitudes in the Indian and eastern Pacific Oceans," *Deep-Sea Research I* **51**, 1337-1346.
- Stafford, K.M., Moore, S.E., and Fox, C.G. (2005). "Diel variation in blue whale calls recorded in the eastern tropical Pacific," *Animal Behaviour* **69**, 951-958.

- Stafford, K.M., Melinger, D.K., Moore, S.E., and Fox, C.G. (2007). "Seasonal variability and detection range modeling of baleen whale calls in the Gulf of Alaska, 1999-2002," J. Acoust. Soc. Am **122**(6), 3378-3390.
- Tershy, B.R., Urbán-Ramírez, J., Breese, D., Rojas-Bracho, L., and Findley, L.T. (1993). "Are fin whales resident to the Gulf of California?" Rev. Inv. Cient. **1**, 69-72.
- Thompson, P.O., Findely, L.T., and Vidal, O. (1992). "20-Hz pulses and other vocalizations of fin whales, *Balaenoptera physalus*, in the Gulf of California, Mexico," J. Acoust. Soc. Am. **92**, 3051-3057.
- Wada, S., Oishi, M., and Yamada, T. (2003). "A newly discovered species of baleen whale," Nature **426**, 278-281.
- Watkins, W.A. (1981). "Activities and Underwater Sounds of Fin Whales," Sci. Rep. Whales Res. Inst. **93**, 83-117.
- Watkins, W.A., Tyack, P., Moore, K.E., and Bird, J.E. (1987). "The 20-Hz signals of finback whales (*Balaenoptera physalus*)," J. Acoust. Soc. Am. **82**(6), 1901-1912.
- Wei, S.S., Wiens, D.A., Webb, S.C., Blackman, D.K., Dunn, R.A., and Conder, J.A. (2012). "Shear velocity structure of the Tonga arc and Lau backarc basin from Rayleigh wave tomography," Abstract T53E-03 presented at 2012 Fall Meeting, AGU, San Francisco, Calif., 3-7 Dec.
- Wessel, P., and W. H. Smith (1998). "New, improved version of the Generic Mapping Tools released," Eos Trans. AGU **79**, 579.
- Wiggins, S.M., Oleson, E.M., McDonald, M.A., and Hildebrand, J.A. (2005). "Blue Whale (*Balaenoptera musculus*) Diel Call Patterns Offshore of Southern California," Aquatic Mammals, **31**(2), 161-168.

Wiseman, N., Parsons, S., Stockin, K.A., and Baker, C.S. (2011). "Seasonal occurrence and distribution of Bryde's whales in the Hauraki Gulf, New Zealand," *Mar. Mamm. Sci.* **27**(4), 253-267.