

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**Site specific passive acoustic detection and densities of humpback
whale calls off the coast of California**

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Oceanography

by

Tyler Adam Helble

Committee in charge:

Gerald L. D'Spain, Chair
Lisa T. Ballance
Peter J.S. Franks
Yoav Freund
John A. Hildebrand
Marie A. Roch

2013

UMI Number: 3558092

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3558092

Published by ProQuest LLC (2013). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against unauthorized copying under Title 17, United States Code



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Copyright
Tyler Adam Helble, 2013
All rights reserved.

The dissertation of Tyler Adam Helble is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

To Dr. Glenn Ierley: teacher, mentor, and lifelong friend.

EPIGRAPH

*If you want to sing out, sing out, and if you want to be free, be free, cause there's
a million ways to be, you know that there are.*

—Cat Stevens

TABLE OF CONTENTS

	Signature Page	iii
	Dedication	iv
	Epigraph	v
	Table of Contents	vi
	List of Figures	ix
	List of Tables	xv
	Acknowledgements	xvi
	Vita and Publications	xx
	Abstract of the Dissertation	xxi
Chapter 1	Introduction	1
	References	9
Chapter 2	A generalized power-law detection algorithm for humpback whale vocalizations	11
	2.1 Introduction	12
	2.2 Detector design considerations	14
	2.3 Theory	16
	2.3.1 Statistics of unit normalization for white noise	21
	2.3.2 Unnormalized statistics for white noise only, with mean removal	23
	2.3.3 Signal plus noise	29
	2.3.4 Summary	34
	2.4 Specific considerations for GPL algorithm used on HARP data for humpback detection	34
	2.5 Monte Carlo simulations	39
	2.5.1 Simulations comparing detector performance	41
	2.5.2 Simulations comparing power-law detectors to trained human analysts	45
	2.6 Parameter estimation	46
	2.7 Observational results	47
	2.8 Conclusions	50
	2.A Mathematical details	55
	References	57

Chapter 3	Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones	60
	3.1 Introduction	61
	3.2 Passive acoustic recording of transiting humpback whales off the California coast	65
	3.2.1 The humpback whale population off California	65
	3.2.2 HARP recording sites	67
	3.2.3 Probability of detection with the recorded data	74
	3.3 Probability of detection - modeling	75
	3.3.1 Approach - numerical modeling for environmental effects	78
	3.3.2 CRAM	82
	3.3.3 Results	84
	3.4 Model/Data Comparison	93
	3.5 Discussion	95
	3.6 Conclusions	98
	References	100
Chapter 4	Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics	104
	4.1 Introduction	105
	4.2 Methods	106
	4.3 Results	111
	4.4 Discussion	113
	References	115
Chapter 5	Humpback whale vocalization activity at Sur Ridge and in the Santa Barbara Channel from 2008-2009, using environmentally corrected call counts	117
	5.1 Introduction	118
	5.2 Methods	120
	5.2.1 Uncertainty Estimates	121
	5.3 Results	129
	5.3.1 Monthly and daily calling activity	129
	5.3.2 Call diel patterns	130
	5.3.3 Call density and lunar illumination	131
	5.3.4 Call density and ocean noise	131
	5.4 Discussion	132
	5.4.1 Seasonal comparison	132
	5.4.2 Diel comparison	135
	5.4.3 Calling behavior and ocean noise	137

	5.4.4 Population density estimates for humpback whales using single-fixed sensors	138
	References	140
Chapter 6	Conclusions and Future Work	146
	6.1 Improving animal density estimates from passive acoustics	147
	6.2 Improvements to studying migrating humpback whales in coastal California	150
	6.3 Improvements to the GPL detector	151
	6.4 Marine mammals as a source for geoacoustic inversions .	152

LIST OF FIGURES

Figure 2.1:	(Color online) Computed pdfs for the L^P norm in Eq. (2.18) for $p = 2, 6, \infty$ along with a Gaussian.	26
Figure 2.2:	(Color online) A comparison of numerical and analytic forms for the cdf of Eq. (2.17) for a) $p = 2$ and b) $p = 6$, emphasizing the tail of the distribution.	27
Figure 2.3:	(Color online) Comparison of the tails of the cdfs for local shipping (asterisk), distant shipping (open square), and wind driven (open circle) noise conditions versus ideal white noise (dashed).	29
Figure 2.4:	(Color online) Pdfs for a) $f_{GPL}^{(\infty)}$, b) f_E for signal amplitudes of 0 (dashed) and 2, 3, 4, 5 (solid) from left to right in each plot.	32
Figure 2.5:	Visual comparison of energy and GPL for six humpback call units in the presence of local shipping noise starting with a) conventional spectrogram ($ \mathbf{X} $) and b) resulting energy sum, c) energy with whitener ($ \hat{\mathbf{X}} $), d) resulting sum, and finally e) \mathbf{N} as defined in Sect. 2.3, and f) GPL detector output $T^g(\mathbf{X})$. Units are highlighted in e) with white boxes. GPL detector output in f) shows eight groupings of detector statistic values above threshold (horizontal line). The six whale call units (red) meet the minimum time requirements, but the four detections (green) resulting from shipping noise do not, and so are not considered detections. All grams in units of normalized magnitude (dB).	36
Figure 2.6:	(Color online) Six humpback units used in Monte Carlo Simulations.	40
Figure 2.7:	(Color online) DET results for Units 1-6 with SNR -3 dB in noise dominated by a) wind-driven noise, b) distant shipping, and c) local shipping, for GPL (closed circle), Nuttall (open triangle), entropy (asterisk), $E^{(1)}$ (open circle), and $E^{(2)}$ (open square).	44
Figure 2.8:	(Color online) DET results for HARP deployments at a) Site SurRidge, b) Site B, and c) Site N for GPL (closed circle), energy sums $E^{(1)}$ (open circle), and $E^{(2)}$ (open square).	51
Figure 2.9:	(Color online) Normalized histogram of detector outputs for signal and signal+noise for Site N deployment.	52

Figure 3.1:	Map of coastal California showing the three HARP locations: site SBC, site SR, and site Hoke (stars). The expanded region of the Santa Barbara Channel shows northbound (upper) and southbound (lower) shipping lanes in relation to site SBC. Ship traffic from the Automatic Identification System (AIS) is shown for region north of 32°N and east of 125°W. The color scale indicates shipping densities, which represent the number of minutes a vessel spent in each grid unit of 1 arc-min x 1 arc-min size in the month of May 2010. White perimeters represent marine sanctuaries. Shipping densities provided by Chris Miller (Naval Postgraduate School).	64
Figure 3.2:	(Color online) Six representative humpback whale units used in the modeling. Units labeled 1-6 from left to right.	67
Figure 3.3:	Bathymetry of site SBC, site SR, and site Hoke (left to right) with accompanying transmission loss (TL) plots. The TL plots are incoherently averaged over the 150 Hz to 1800 Hz band and plotted in dB (the color scale for these plots is given on the far right). The location of the HARP in the upper row of plots is marked with a black asterisk.	69
Figure 3.4:	Sound speed profiles for site SBC, site SR, and site Hoke (top to bottom), for winter (blue) and summer (red) months. These data span the years 1965 to 2008.	70
Figure 3.5:	Noise spectral density levels for site SBC, site SR, and site Hoke (top to bottom). The curves indicate the 90th percentile (upper blue), 50th percentile (black), and 10th percentile (lower blue) of frequency-integrated noise levels for one year at site SBC and site SR, nine months at site Hoke. The gray shaded area indicates 10th and 90th percentile levels for wind-driven noise used for modeling.	73

Figure 3.6: (Color online) (a) Measured humpback whale source signal rescaled to a source level of 160 dB re 1 μ Pa @ 1 m, (b) simulated received signal from a 20-m-deep source to a 540-m-deep receiver at 5 km range in the Santa Barbara Channel, with no background noise added, (c) simulated received signal as in (b) but with low-level background noise measured at site SBC added. The upper row of figures are spectrograms over the 0.20 to 1.8 kHz band and with 2.4 sec duration, and the lower row are the corresponding time series over the same time period as the spectrograms. The received signal and signal-plus-noise time series amplitudes in the 2nd and 3rd columns have been multiplied by a factor of 1000 (equal to adding 60 dB to the corresponding spectrograms) so that these received signals are on the same amplitude scale as the source signal in the first column. This example results in a detection with recorded $SNR_{est} = 2.54$ dB. 80

Figure 3.7: Probability of detecting a call based on the geographical position of a humpback whale in relation to the hydrophone during periods dominated by wind-driven noise at site SBC (upper left), site SR (upper center), and site Hoke (upper right), averaged over unit type. Assuming a maximum detection distance of $w = 20$ km, average $\hat{P} = 0.1080$ for site SBC, $\hat{P} = 0.0874$ for site SR, and $\hat{P} = 0.0551$ for site Hoke. The latitude and longitude axes in the uppermost row of plots is in decimal degrees. The detection probability functions for the three sites, resulting from averaging over azimuth, are shown in the middle row and the corresponding PDFs of detected distances are shown in the lower row. Solid (dashed) lines indicate functions with (without) the additional -1 dB SNR_{est} threshold applied at the output of GPL detector. 83

Figure 3.8: Geographical locations of detected calls (green dots mark the source locations where detections occur) and associated probability of detection (\hat{P} , listed in the upper right corner of each plot) for calls 1-6 (left to right, starting at the top row) in a 20 km radial distance from the hydrophone for a single realization of low wind-driven noise at site SBC. The latitude and longitude scales on each of the six plots are the same as in the upper lefthand plot of Fig. 3.7. 86

Figure 3.9:	Site SBC (upper) and site SR (lower) \hat{P} versus noise level for the sediment property and SSP pairing that maximizes \hat{P} (red), the sediment/SSP pairing that minimizes \hat{P} (green), and the best-estimate environmental parameters (blue). Vertical error bars indicate the standard deviation among call unit types, and horizontal error bars indicate the standard deviation of the noise measurement. The noise was estimated by integrating the spectral density over the 150 Hz to 1800 Hz frequency bands using twelve samples of noise within a 75 s period.	90
Figure 3.10:	Shaded gray indicates normalized histogram of received SNR estimates (SNR_{est}) for humpback units at site SBC, site SR, and site Hoke (top to bottom). Model best environmental estimates (black line), and model upper environmental estimates (green line). The cyan line indicates best estimate results with 4 km radial calling "exclusion zone" at site Hoke.	91
Figure 4.1:	Ocean noise levels in the 150-1800 Hz band over the 2008-2009 period at site SBC (upper) and SR (lower). The gray curves indicate the noise levels averaged over 75 sec increments, the green curves are the running mean with a 7 day window, and the black curve (site SR only) is a plot of the average noise levels in a 7-day window measured at the times adjacent to each detected humpback unit. White spaces indicate periods with no data. The blue vertical lines mark the start of enforcement of CARB law.	107
Figure 4.2:	Ocean noise levels at site SBC in May, 2008 (upper), probability of detecting a humpback unit (\hat{P}) within a 20 km radius of site SBC in May 2008 (middle), and the number of humpback units detected in uncorrected form (n_c) at site SBC for the same time period (lower). Shaded time periods indicates sunset to sunrise. The vertical grid lines indicate midnight local time.	108
Figure 4.3:	(color online) Uncorrected number of humpback units detected (n_c) in the 2008-2009 period at site SR (upper), estimated probability of detecting a humpback unit (\hat{P}) within a 20 km radius of site SR (middle), and the corrected estimated number of units occurring per unit area (\hat{N}_c) at site SR for the same time period (lower).	109

- Figure 5.1: Uncorrected call counts n_c , normalized for effort (recording duty cycle) and tallied in 1-month bins for site SR (green) and SBC (blue) (upper panel), corrected estimated call density, $\hat{\rho}_c$, for site SR (green) and site SBC (blue) (middle panels) tallied in 1-month bins. The same datasets are repeated in both panels to illustrate scale. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Values of $\hat{\rho}_c$, for site SR (green) and site SBC (blue) are also repeated in the lower plot on a log scale to illustrate detail. 122
- Figure 5.2: Average daily estimated call density, $\hat{\rho}_c$ shown in 1 hour time bins to illustrate diel cycle for site SR (upper panel) and site SBC (lower panel) for time period covering April 16, 2008 to Dec 31, 2009. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots. 123
- Figure 5.3: Average daily estimated call density, $\hat{\rho}_c$ at site SBC shown in 1 hour local time bins to illustrate diel cycle. The spring season (Apr 7-May 27, 2009) at site SBC (upper panel) shows stronger diel pattern and higher call densities than the fall season (Oct 15-Dec 4, 2009) at site SBC (lower panel). The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots. 124

- Figure 5.4: Average daily estimated call density, $\hat{\rho}_c$, shown in 10% lunar illumination bins, where units are aggregated over the entire deployment for site SR (upper panel) and site SBC (lower panel). Lunar illumination numbers do not account for cloud cover. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots. 125
- Figure 5.5: Estimated call density, $\hat{\rho}_c$ shown in 2 dB ocean noise bins for full 2-year deployment for site SR (upper panel), and site SBC (middle panel), adjusted for recording effort in each noise band. Numerically-estimated uncorrected call counts, \hat{n}_c , shown for site SBC (lower panel) for all detected calls (1,104,749), adjusted for recording effort in each noise band. 126

LIST OF TABLES

Table 2.1:	Distribution of Moments for Eq. (2.17).	50
Table 2.2:	Probability of missed detection and probability of false alarm (P_{MD}/P_{FA} , given as percentage) using η_{thresh} for Units 1-6, varying SNR and noise cases, 10,000 trials per statistic.	52
Table 2.3:	Probability of missed detection (P_{MD} , given as a percentage) for GPL versus baseline power-law detector (Nuttall) and human analysts for varying SNR. Detector threshold values were established such that Case 3 $P_{FA} < 6\%$ and applied to Cases 1 and 2.	53
Table 2.4:	Start-time bias Δt_s , end time bias Δt_e , start time standard deviation σ_s , and end time stand deviation σ_e in seconds for Unit 1 (duration 3.34 s) and Unit 3 (duration 1.3 s)	54
Table 3.1:	Best-estimate and extremal predictions for \hat{P} for wind-driven noise conditions, given the uncertainty in input parameters of SSP and sediment structure for each site, as outlined in Sec. 3.2.2. Each estimate of P assumes the remaining variables are fixed at best-estimate values. The \hat{P} values assume a detection radius of $w = 20$ km from the instrument center. . . .	89

ACKNOWLEDGEMENTS

Many people have contributed to the successful completion of my dissertation. First and foremost, I'd like to thank Dr. Glenn Ierley, whose unwavering support made this dissertation possible. While at Scripps, Glenn provided countless hours of support to all of his students, working endlessly to make them the best scientists possible. Personally, Glenn bestowed an enormity of Matlab skills upon me, without which the work in my thesis would not be possible. Glenn also showed me through his own ten-year pursuit of what he covertly referred to as "LT": that solving any scientific problem is possible with enough discipline and dedication.

My thesis advisor, Dr. Gerald D'Spain, went well above and beyond the call of duty in helping me develop my skills to become a successful scientist. Gerald allowed me the freedom to take full creative responsibility of my thesis, while insisting that I ground my research with a strong theoretical foundation. While writing—and rewriting—each chapter was painstaking, the final product is something of which I will always be proud. I will truly miss our multi-hour brainstorming sessions, his general good-nature, and late-night scientific email exchanges that always led me to wonder if, indeed, he required sleep. My unofficial co-advisor, Dr. John Hildebrand, was also instrumental to the success of my thesis. John welcomed me into the Whale Acoustics Lab with open arms, providing research feedback, resources, and personnel support that were crucial to my research. I will remember his acoustics classes fondly (despite the long haul to upper-campus). The rest of my committee deserves my gratitude for their support and guidance: Dr. Marie Roch, who was very helpful in teaching me about detection performance characterization, was always available to meet, and I'll miss our spontaneous office chats and lunches; Dr. Peter Franks not only dedicated an immense amount of time to his students, his first-year biological oceanography class was one of my favorites at Scripps and I was extremely impressed by the thorough review he gave to each manuscript I sent him; Dr. Lisa Ballance's marine tetrapod class inspired me to include marine mammals as part of my Ph.D. research, and her contagious enthusiasm always gave me a great sense of motivation; Dr. Yoav Freund provided

feedback on my research from a computational learning theory perspective, which was greatly appreciated.

In addition to my Ph.D. committee, a number of other mentors at Scripps deserve much thanks. Dr. Clint Winant worked with me after class to teach me partial differential equations while I was enrolled in his fluid mechanics class. He dedicated much of his time to my success, and I am truly appreciative. In addition to teaching four of the classes critical to my success at Scripps, Dr. Bill Hodgkiss also made time to meet with me outside of class, despite his busy schedule. His feedback at the early stages of my research were crucial in getting me on my feet. Special thanks to Heidi Batchelor and Dr. Stephen Lynch at MPL, who both allowed me to vent my frustrations while concurrently helping with Matlab coding and mapmaking.

Each member of the Scripps Whale Acoustics Lab (both past and present) contributed to the success of my research. Greg Campbell and Amanda Debich were instrumental in teaching me the ins-and-outs of human-aided analysis of marine mammal vocalizations. Without their feedback, the GPL detector described in Ch. 2 would have never gotten off the ground. Additionally, Greg and Amanda both spent considerable time pruning the datasets used in this thesis to remove false-alarms from the detection process. Thanks to Dr. Sean Wiggins for teaching me how to use the calibration files for HARP sensors. To Karli Merkins: in addition to being a great friend, thank you for reviewing my manuscripts and providing insightful feedback on density estimation. Liz Vu and Aly Fleming: your knowledge of humpback whales is incredible - thanks for passing some of it along to me. Megan Mckenna was extremely helpful for sharing her knowledge of ship noise in coastal California. She spent many hours chatting with me on the phone in her free time, sharing Matlab code, and brainstorming ideas for research. I would also like to thank Kait Fraiser, Bruce Thayre, Sara Kerosky, Ana Sirovic, and Simone Baumann-Pickering for offering their assistance.

To my other friends at Scripps (Tara Whitty, Todd Johnson, Jilian Maloney, Michelle Lande, Alexis Pasulka, and Guangming Zheng): thanks for making the graduate experience so memorable. To Tamara Beitzel: I am so glad we have

become such great friends. I could not think of a better companion to survive my first year with! I would like to thank Brianne Baxa for being my officemate, running and swimming buddy, improvised dance partner, and friend. Big thanks to Timothy Ray, whose smiling face always lit up the room – I will try my hardest to spread Tim’s passion and excitement for conservation and science throughout my career.

This thesis would not have been possible without the support of the Space and Naval Warfare (SPAWAR) Systems Command Center Pacific In-House Laboratory Independent Research program and the Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship program. Rich Arrieta, Greg Kwik, Dave Reese, Roger Boss, and Lynn Collins were all responsible for making this thesis possible.

I would also like to thank Richard Campbell and Kevin Heaney at Ocean Acoustical and Instrumentation Systems (OASIS) for allowing me to use the CRAM software package for my research, in addition to providing a great deal of technical support.

I would like to thank my professors at Duke University for providing me with the guidance and skills necessary for making my career at Scripps a reality, especially Dr. Emily Klein, Dr. Susan Lozier, and Dr. Michael Gustafson. Thanks to all of my teachers in the Okemos Public School system, especially John Olstad, who solidified my love for science.

To Katie Gerard, my 4th grade girlfriend and lifelong friend: thanks for being my "life coach".

And last but not least, I would like to thank my extraordinary family. My parents Ed and Charlene Helble have provided me with the means to explore my creativity since the moment I was born; none of this would be possible without their unwavering support and guidance. Thanks to my talented brothers, Nick and Mitch Helble, from whom I draw strength and inspiration on a daily basis. I would also like to thank my partner in life, Aaron Schroeder; the journey would not be the same without you.

This dissertation is a collection of papers that have been accepted,

submitted, or are in preparation for publication.

Chapter 2 is, in full, a reprint of material published in *The Journal of the Acoustical Society of America*: Tyler A. Helble, Glenn R. Ierley, Gerald L. D’Spain, Marie A. Roch, and John A Hildebrand, “A generalized power-law detection algorithm for humpback whale vocalizations”. The dissertation author was the primary investigator and author of this paper.

Chapter 3 is, in full, a reprint of material accepted for publication in *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D’Spain, John A. Hildebrand, Greg S. Campbell, Richard L. Campbell, and Kevin D. Heaney “Site specific probability of passive acoustic detection of humpback whale class from single fixed hydrophones”. The dissertation author was the primary investigator and author of this paper.

Chapter 4 is a manuscript in preparation for submission to *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D’Spain, Greg S. Campbell, and John A. Hildebrand, “Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics”. The dissertation author was the primary investigator and author of this paper.

Chapter 5 is a manuscript in preparation for submission to *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D’Spain, Greg S. Campbell, and John A. Hildebrand, “Humpback whale vocalization activity at Sur Ridge and in the Santa Barbara Channel from 2008-2009, using environmentally corrected call counts”. The dissertation author was the primary investigator and author of this paper.

VITA

2004	B.S.E., Electrical Engineering Duke University
2004	B.S., Environmental Science Duke University
2010	M.S., Oceanography - Applied Ocean Sciences Scripps Institution of Oceanography, University of California, San Diego
2013	Ph.D., Oceanography - Applied Ocean Sciences Scripps Institution of Oceanography, University of California, San Diego
2007-2013	Graduate Student Researcher Marine Physical Laboratory, University of California, San Diego

PUBLICATIONS

Journals

1. Tyler A. Helble, Gerald L. D'Spain, John A. Hildebrand, Greg S. Campbell, Richard L. Campbell, and Kevin D. Heaney, "Site specific probability of passive acoustic detection of humpback whale class from single fixed hydrophones", *J. Acoust. Soc. Am.*, accepted.
2. Tyler A. Helble, Glenn R. Ierley, Gerald L. D'Spain, Marie A. Roch, and John A Hildebrand, "A generalized power-law detection algorithm for humpback whale vocalizations", *J. Acoust. Soc. Am.*, Volume 131, Issue 4, pp. 2682-2699 (2012)

Conferences

1. Tyler A. Helble, Glenn R. Ierley, Gerald L. D'Spain, Marie A. Roch, and John A Hildebrand, "A generalized power-law detection algorithm for humpback whale vocalizations", Fifth International Workshop on Detection, Classification, Localization, and Density Estimation of Marine Mammals using Passive Acoustics. Mount Hood, Oregon. (2011)

ABSTRACT OF THE DISSERTATION

**Site specific passive acoustic detection and densities of humpback
whale calls off the coast of California**

by

Tyler Adam Helble

Doctor of Philosophy in Oceanography

University of California, San Diego, 2013

Gerald L. D'Spain, Chair

Passive acoustic monitoring of marine mammal calls is an increasingly important method for assessing population numbers, distribution, and behavior. Automated methods are needed to aid in the analyses of the recorded data. When a mammal vocalizes in the marine environment, the received signal is a filtered version of the original waveform emitted by the marine mammal. The waveform is reduced in amplitude and distorted due to propagation effects that are influenced by the bathymetry and environment. It is important to account for these effects to determine a site-specific probability of detection for marine mammal calls in a given study area. A knowledge of that probability function over a range of environmental and ocean noise conditions allows vocalization statistics from recordings of single,

fixed, omnidirectional sensors to be compared across sensors and at the same sensor over time with less bias and uncertainty in the results than direct comparison of the raw statistics.

This dissertation focuses on both the development of new tools needed to automatically detect humpback whale vocalizations from single-fixed omnidirectional sensors as well as the determination of the site-specific probability of detection for monitoring sites off the coast of California. Using these tools, detected humpback calls are "calibrated" for environmental properties using the site-specific probability of detection values, and presented as call densities (calls per square kilometer per time). A two-year monitoring effort using these calibrated call densities reveals important biological and ecological information on migrating humpback whales off the coast of California. Call density trends are compared between the monitoring sites and at the same monitoring site over time. Call densities also are compared to several natural and human-influenced variables including season, time of day, lunar illumination, and ocean noise. The results reveal substantial differences in call densities between the two sites which were not noticeable using uncorrected (raw) call counts. Additionally, a Lombard effect was observed for humpback whale vocalizations in response to increasing ocean noise. The results presented in this thesis develop techniques to accurately measure marine mammal abundances from passive acoustic sensors.

Chapter 1

Introduction

The use of passive acoustics to study marine life is an evolving field. Interest in underwater sound has been noted as early as 1490, when Leonardo Da Vinci wrote, "If you cause your ship to stop and place the head of a long tube in the water and place the outer extremity to your ear, you will hear ships at a great distance from you"[1]. Along with ships, whales also produce sound underwater, and this thesis addresses some of the earliest observations noted by Da Vinci. To what "great distance" is a whale heard? What is the probability you will hear that whale? How does this probability change under different environmental conditions? How has the sound been altered at the receiving end, after it has traveled this great distance? Does the sound produced by the ships Da Vinci noted, when heard by whales, affect the whales' behavior? These questions, simple in nature, prove to be complex and multidisciplinary to answer.

The use of underwater recording devices to study marine mammals began in 1949 when William E. Schevill and B. Lawrence deployed hydrophones (microphones that detect sound waves underwater) into the Saguenay River of Quebec, recording the Beluga (*Delphinapterus leucas*) whale for the first time in the wild [2]. Since then, passive acoustic monitoring has been used to study nearly all aspects of marine mammal ecology and biology. Initial passive acoustic studies often focused on deciphering marine mammal "language", in which scientists attempted to determine the purpose of different types of vocalizations by relating them to social, feeding, and mating behaviors[3, 4]. To this day, this field remains

an area of active research.

A more recent application of passive acoustic monitoring is to measure marine mammal abundance, which is critical for managing endangered or threatened species. Abundance studies in the past have primarily relied on visual sighting techniques. Some of the earliest visual sighting techniques for measuring marine mammal abundance employed methods of counting individuals from stationary locations. Scientists often focused on areas where marine mammals aggregated in colonies (during breeding for example), or along narrow corridors of migration routes[5, 6]. Mark-recapture methods, which use natural markings or man-made tags to mark a subset of the population, have also been employed. The total population size can then be derived using statistical methods after the population is resampled[7].

An alternative and often preferable tool for visual abundance estimates is the distance sampling method[8], which has become widely used by the marine mammal community. Two primary methods of distance sampling exist - line transect and point transect sampling. The line transect method is the most widely used, which employs a ship or aircraft to survey an area. The observers move in systematically-placed straight lines through the study area, counting the number and distance to individual animals, groups of animals, or visual cues from animals, such as blow hole spray. Because every individual in a population cannot be counted, each visual survey method requires observers to make a certain set of assumptions about the study animals. Errors in estimates occur when these assumptions are violated. For line transect methods, it is assumed that animals on, or very close to, the line are certain to be detected, animals are detected before responding to the presence of the observer, and that distances to the animals are accurately measured. If these assumptions are met, animal densities can be calculated. The detection function, which is the probability of detecting the species as a function of distance, is not needed a priori, and is in fact derived from the sampling data after the survey. Calculating the detection function is a crucial step for estimating animal densities, and so deriving this function directly from the dataset is advantageous. Additionally, the distribution of animals in the survey

area need not be random, making the survey technique fairly robust.

An alternative to visual sighting techniques for abundance estimates is the use of passive acoustic methods. Acoustic arrays in particular can be used in place of visual observers in a line transect survey[9]. Using passive acoustics is particularly advantageous for highly vocal species that may spend little time at the surface, which violates the visual assumption that animals along a transect are always detectable. Arrays contain multiple hydrophones and information can be coherently combined across the hydrophones, in a process known as beamforming, which allows bearings and/or locations of vocalizing animals to be estimated. If the probability of detecting an animal is less than 100% along the transect line, the probability along the line needs to be estimated using auxiliary information. An acoustic "cue" (vocalization) rate may also need to be estimated for the species, since it may not be possible to distinguish vocalizations from individuals traveling in groups.

Because both visual and acoustic line-transect methods are costly and cannot practically be conducted on a continuous, long-term basis, fixed passive acoustic sensors have been increasingly used throughout the marine mammal community. Fixed sensors are usually anchored to the seafloor, and often record continuously over several months or years. When hydrophone arrays or single hydrophone systems with overlapping coverage are deployed, it is still possible to localize marine mammals. If animal locations are known, the detection function and distribution of animals can be estimated, allowing for animal abundance to be calculated in the monitored area.

This thesis concerns the use of bottom-mounted passive acoustic monitoring systems composed of a single omnidirectional hydrophone, which are often deployed in place of hydrophone array systems because they are typically easier to deploy, require less bandwidth and electrical power, and are less expensive to construct. The main drawback to using single, fixed omnidirectional sensors is that the detection function is often unknown a priori and it is usually not possible to determine distances to vocalizing marine mammals using these sensors - a step required to establish the detection function from sensor data. Additionally,

the distribution of animals in the area cannot be determined from the sensor itself. For single, fixed omnidirectional sensors, the detection function, animal distribution, and cue rate are all needed in order to determine accurate density estimates. Scientists have generally avoided animal density estimate calculations from single, fixed omnidirectional sensors because of the difficulties in measuring these quantities, although successful instances of doing so have been published. [10, 11]. Despite not knowing the detection function in a study area, many scientists mark the presence/absence of detections or tabulate cue counts from these sensors, and use these numbers as a proxy to compare activity at the same sensor over varying time scales, or compare activity across widely separated sensors. The work in this thesis focuses on developing tools to both optimally detect acoustic cues and develop site-specific detection functions for single, fixed omnidirectional sensors in order to estimate the probability of detecting marine mammal calls in a given area with changing environmental and ocean noise conditions. In doing so, calling activity can be compared at the same sensor over time or across sensors with less bias and uncertainty. Rather than comparing detected call counts across sensors or at the same sensor over time, the calibration methods described in this thesis allow for the comparison of call densities, which is the number of calls produced per area per time. The hypothesis of this thesis is that using call densities from properly calibrated single, fixed omnidirectional sensors can reveal substantial biological and ecological information about transiting humpback whales off the coast of California. This information may not be available from detected call counts alone.

A key eventual goal of acoustic monitoring is estimating animal abundance, which in turn requires that one know the density of animals throughout a region versus time. But what a single hydrophone records is an acoustic cue. In general it is not possible to tell from the record of cues itself how many individuals are represented but, as an intermediate result, it is possible to determine the call density. Because the cues are masked to a varying degree by background noise and environmental properties that vary over space and time, inevitably not all calls are detected in the recording and so it is necessary to correct for this systematic

undercounting (using the detection function) to estimate the true value. If the cue rate of a species is known (and stable over some period of time), then animal densities can also be estimated using this method from single, fixed omnidirectional sensors. The situation under consideration is in some ways analogous to counting stars in the nighttime sky - depending on the cloud cover, light pollution, and phase of the moon, a human observer may count no stars or thousands of stars. In all situations, the number of stars observed is an underrepresentation of the true number. However, if the probability of detecting a star is known for each set of conditions, then the true number can be estimated.

Humpback whales have long captured the interest of scientists, producing perhaps the most diverse and complex vocalizations of all marine mammals. Humpback whales produce underwater 'song', a hierarchical structure of individual sounds termed 'units'. These units are grouped into 'phrases', and phrases are grouped into 'themes', which combine to make up the song[12]. Songs are produced by mature males and are thought to have important social and mating functions. Song has been observed on all humpback whale breeding grounds, and has been noted to occur on migration routes and even at high latitude feeding grounds. Other sounds are produced throughout the year by both male and female humpback whales, and some of these sounds have been linked to certain social and feeding behaviors[13]. Humpback whales are an endangered species. Prior to commercial whaling, worldwide population estimates suggests as many as 240,000 individuals[14]. An estimated 5-10% of the original population remained when an international ban on whaling was established in 1964. Since then, the humpback whale population has made an encouraging recovery with roughly 80,000 individuals estimated world wide[15, 16, 17, 18]. Nevertheless, certain sub-populations are particularly vulnerable and since humpbacks cover a wide range of coastal and island waters, increasing human activity in these regions may pose a risk.

The combination of a complex and evolving vocal structure, relatively unstudied migration routes, and an endangered population of animals makes the humpback whale both a challenging and rewarding candidate to study using

passive acoustic monitoring. Historically, humpback whale vocalizations have been monitored from passive acoustic recordings using trained human operators to note the presence and absence of song and social calls. However, in order to answer more complex questions about humpback whale ecology and biology from passive acoustics, a much greater sample size of detected calls was needed. The first half of this thesis focuses on developing the tools needed to detect humpback cues in an automated and optimal way, and to calibrate the single, fixed omnidirectional sensors to more accurately estimate humpback call densities. The second half of the thesis focuses on the importance of using calling densities over uncorrected acoustic cue counting, while revealing biological and ecological relevant information on humpback whales off the coast of California.

Following this introduction, Chapter 2 of this thesis details the generalized power-law (GPL) detector, which was developed to optimally detect and efficiently mark the start-time and end-time of nearly every human-identifiable humpback unit (each unit is considered an acoustic cue) in an acoustic record. Aside from being labor and time-prohibitive, using humans to mark vocalizations in an acoustic record is problematic because the performance of a human operator is highly variable and nearly impossible to characterize quantitatively. The development of the GPL detector is a unique contribution to marine mammal monitoring community for several reasons. Practically, its performance allows for the reliable detection of humpback units even in highly variable ocean-noise conditions, allowing scientist to monitor long acoustic records with higher fidelity than previously possible. Theoretically, analysis proves that the GPL detector, which is based on Nuttall's original power-law processor[19], is the near-optimal approach to detecting transient marine mammal vocalizations with unknown location, structure, extent, and arbitrary strength. The performance with these types of signals is a vast improvement over the energy detector, which is commonly used throughout the marine mammal community.

Chapter 3 focuses on the development of a second tool - a modeling suite that outputs probability of detection maps (analogous to the detection function described earlier) for humpback whale calls within each geographical

area containing a single, fixed omnidirectional sensor. The approach uses the Range-dependent Acoustic Model (RAM) that uses environmental inputs such as bathymetry, ocean bottom geoacoustic properties, and sound-speed profiles to predict the received sounds of simulated humpback whale vocalizations from locations surrounding each sensor. The simulated acoustic pressure time series of the whale calls are then summed with time series realizations of ocean noise and processed by the GPL detector, and the detection performance is recorded in order to estimate the probability of detection maps around each sensor. The locations of the three fixed sensors under consideration are shown in Fig. 3.1, and the study area is fully described in Ch. 3.2.2. The material in Ch. 3 is unique in that the probability of detection maps and the associated uncertainties are estimated over a wide range of likely environmental characteristics using full wave field acoustic models. Additionally, real instances of ocean noise that contain a wide range of spectral characteristics are used in the detection process. The full wave-field model allows the transmitted humpback signal to attenuate over frequency and accounts for phase distortions (due to dispersion and multipath), which can affect the detection process. Using real noise and a range of likely environmental properties results in the most accurate calculations of probability of detection maps and the associated uncertainties for fixed, omnidirectional sensors with non-overlapping coverage. Published related research employs the use of simple transmission-loss models and generally characterizes the transmission, noise, and detection processes separately, resulting in a much less realistic model. Additionally, most previous research has focused on high-frequency calling animals and the influence of environmental properties on the detection process has been minimized or ignored. Using the same published techniques in this thesis research would be an oversimplification for the propagation properties of mid and low-frequency humpback whale calls.

Chapter 4 establishes the importance of using both the GPL detector and acoustic modeling tools developed in the previous chapters by illustrating the differences between uncorrected call counts (acoustic cue counting) and corrected call densities at two hydrophone locations off the coast of California. Due to

changes in the world economy and the enforcement of new air pollution regulations, ocean noise decreased at both locations over a two-year period. The uncorrected call counts show a significant increase in detections in the second season at Sur Ridge, a site located off the coast of Monterey, CA. After the original call counts were corrected for the probability of detection, the resulting calling densities appeared roughly the same between the two years. A second example highlighting the variability of shipping noise on an hourly scale shows how uncorrected call counts vary inversely with shipping noise. A diel pattern in the number of uncorrected calls appears to show increased calling during nighttime hours, a pattern which disappears in certain months after correcting for the probability of detection. The analysis in Ch. 4 is perhaps the first study to ever systematically address the influence of changing ocean conditions on single, fixed omnidirectional passive acoustic monitoring results using datasets containing marine mammal calls.

Chapter 5 utilizes the tools and observations from the previous three chapters to address the hypothesis of this thesis - can passive acoustics, when calibrated for site specific probability of detection, reveal significant biological and ecological information on humpback whales off the coast of California? Humpback calling densities are presented for the Santa Barbara Channel (site SBC), and Sur Ridge (site SR) off the coast of Monterey covering a two-year study period from January 2008 through December 2009. Comparing call densities between the two sites reveal that call densities were roughly four times higher at site SR than site SBC. These results could indicate that only a portion of migrating whales choose to enter into the Santa Barbara Channel. Additionally, the call densities between years at site SBC are much more variable than at site SR, indicating the Santa Barbara Channel could be an opportunistic feeding source for migrating humpback whales. Call densities were also compared against a variety of environmental properties, including time of day, lunar illumination, and ocean noise. Results indicate that humpback whales have a tendency to call during nighttime hours, particularly in spring months, although the diel pattern varied noticeably between the two locations. Substantial evidence also exists that humpback whales have a vocal response to increasing ocean noise - either by increasing vocalization rates

and/or increasing the average source level of their calls. These results do reveal in an objective, quantitative way important biological and ecological information on transiting humpback whales and the potential impact human activity can have on their behavior. Additionally, the highly variable cue rate across seasons as shown in Ch. 5, combined with the potential for this cue rate to change with varying ocean noise and other environmental inputs calls the use of passive acoustics for accurate animal density estimates of this species into question.

Concluding remarks, including recommendations and directions for future research, are provided in the final chapter (Ch. 6).

References

- [1] R.J. Urick. *Principles of Underwater Sound*, volume 3, pages 19–22. McGraw-Hill, New York, NY, 1983.
- [2] W.E. Schevill and B. Lawrence. Underwater listening to the white porpoise (*Delphinapterus leucas*). *Science (New York, NY)*, 109(2824):143, 1949.
- [3] J. Wood. Underwater sound production and concurrent behavior of captive porpoises, *Tursiops truncatus* and *Stenella plagiodon*. *Bulletin of Marine Science*, 3(2):120–133, 1953.
- [4] W.E. Schevill. Underwater sounds of cetaceans. *Marine bio-acoustics*, 1:307–316, 1964.
- [5] P.M. Thompson and J. Harwood. Methods for estimating the population size of common seals, *Phoca vitulina*. *Journal of Applied Ecology*, pages 924–938, 1990.
- [6] W.H. Dawbin. The migrations of humpback whales which pass the New Zealand coast. *Transactions of the Royal Society of New Zealand*, 84(1):147–196, 1956.
- [7] L.L. Eberhardt, D.G. Chapman, and J.R. Gilbert. A review of marine mammal census methods. *Wildlife Monographs*, (63):3–46, 1979.
- [8] S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, and L. Thomas. *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, pages 1–448. Oxford University Press, New York, NY, 2001.

- [9] J. Barlow and B.L. Taylor. Estimates of sperm whale abundance in the northeastern temperate Pacific from a combined acoustic and visual survey. *Marine Mammal Science*, 21(3):429–445, 2005.
- [10] E.T. Küsel, D.K. Mellinger, L. Thomas, T.A. Marques, D. Moretti, and J. Ward. Cetacean population density estimation from single fixed sensors using passive acoustics. *J. Acoust. Soc. Am.*, 129(6):3610–3622, 2011.
- [11] T.A. Marques, L. Munger, L. Thomas, S. Wiggins, and J.A. Hildebrand. Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research*, 13:163–172, 2011.
- [12] R.S. Payne and S. McVay. Songs of humpback whales. *Science*, 173(3997):585–597, 1971.
- [13] S. Cerchio and M. Dahlheim. Variation in feeding vocalizations of humpback whales *Megaptera novaeangliae* from southeast Alaska. *Bioacoustics*, 11(4):277–295, 2001.
- [14] J. Roman and S.R. Palumbi. Whales before whaling in the North Atlantic. *Science*, 301(5632):508–510, 2003.
- [15] J. Calambokidis, E.A. Falcone, T.J. Quinn, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, R. LeDuc, D. Mattila, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J.R. Urban, D. Weller, B.H. Witteveen, M. Yamaguchi, A. Bendlin, D. Camacho, K. Flynn, A. Havron, J. Huggins, and N. Maloney. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Technical report, Cascadia Research Collective, Olympia, WA, 2008.
- [16] T.A. Branch. Humpback whale abundance south of 60 s from three complete circumpolar sets of surveys. *J. Cetacean Res. Manage*, 2010.
- [17] T.D. Smith, J. Allen, P.J. Clapham, P.S. Hammond, S. Katona, F. Larsen, J. Lien, D. Mattila, P.J. Palsbøll, J. Sigurjónsson, et al. An ocean-basin-wide mark-recapture study of the North Atlantic humpback whale (*Megaptera novaeangliae*). *Marine Mammal Science*, 15(1):1–32, 1999.
- [18] A. Fleming and J. Jackson. Global review of humpback whales (*Megaptera novaeangliae*). NOAA Technical Memorandum NMFS. Technical report, U.S. Department of Commerce, Washington, D.C., 2011.
- [19] A.H. Nuttall. Detection performance of power-law processors for random signals of unknown location, structure, extent, and strength. Technical report, NUWC-NPT, Newport, RI, 1994.

Chapter 2

A generalized power-law detection algorithm for humpback whale vocalizations

Abstract

Conventional detection of humpback vocalizations is often based on frequency summation of band-limited spectrograms, under the assumption that energy (square of the Fourier amplitude) is the appropriate metric. Power-law detectors allow for a higher power of the Fourier amplitude, appropriate when the signal occupies a limited but unknown subset of these frequencies. Shipping noise is non-stationary and colored, and problematic for many marine mammal detection algorithms. Modifications to the standard power-law form are introduced in order to minimize the effects of this noise. These same modifications also allow for a fixed detection threshold, applicable to broadly varying ocean acoustic environments. The detection algorithm is general enough to detect all types of humpback vocalizations. Tests presented in this paper show this algorithm matches human detection performance with an acceptably small probability of false alarms ($P_{FA} < 6\%$) for even the noisiest environments. The detector outperforms energy detection techniques, providing a probability of detection $P_D = 95\%$ for

$P_{FA} < 5\%$ for three acoustic deployments, compared to $P_{FA} > 40\%$ for two energy-based techniques. The generalized power-law detector also can be used for basic parameter estimation, and can be adapted for other types of transient sounds.

2.1 Introduction

Detecting humpback whale (*Megaptera novaeangliae*) vocalizations from acoustic records has proven to be difficult for automated detection algorithms. Humpback songs consist of a sequence of discrete sound elements, called units, that are separated by silence[1]. Both the units and their sequence evolve over time and cover a wide range of frequencies and durations[1, 2]. In addition, individual units may not repeat in a predictable manner, especially during non-song or broken song vocalizations, or in the presence of multiple singers with overlapping songs [1, 2]. Many types of marine mammal detection and classification techniques have been developed, using methods of spectrogram correlation[3], neural networks[4], Hidden Markov Models[5, 6], and frequency contour tracking[7], among others. Depending on the species of marine mammal, noise condition, and type of vocalization, many of these methods have been shown to be effective in producing high probabilities of detection (P_D) with low probabilities of false alarm (P_{FA}). However, for humpback vocalizations, these techniques often provide low P_D if the P_{FA} is to remain adequately low. Abbot *et al.* [8] used a kernel-based spectrogram correlation to identify the presence of humpback whales with extremely low P_{FA} . However, their approach requires 15 kernel matches within a three minute window in order to trigger a detection. Therefore, the goal is not to detect every humpback unit, but rather to predict the presence of song when enough predefined kernels are matched. Energy detection algorithms, readily available in acoustic analysis software such as Ishmael[9], XBAT[10], and PAMGuard[11] have proven effective for detecting all types of humpback call units. However, in order to avoid an exorbitant number of false detections, these methods generally require high signal-to-noise ratio (SNR): the hydrophones are in close proximity to the whales, and/or the shipping noise is low. Erbe and King[12] recently developed an entropy detector that can outperform

energy detection methods for a variety of marine mammal vocalizations. However, this method is inadequate for detecting humpback vocalizations for data sets that contain considerable shipping noise. Therefore, a need still exists for an automated detection capability in low SNR scenarios that is able to achieve low probability of false alarms, yet is general enough to achieve high probability of detection for all humpback units, including those with poorly defined spectral characteristics.

Nuttall introduced a general class of power-law detectors for a white noise environment[13, 14]. The energy method – based on the square of the Fourier amplitude – is a particular case, optimum when the signal occupies all the frequency bands over which energy summation occurs. However, in the case of narrowband transient signals that fall within a wide range of monitored frequencies (characteristic of humpback vocalizations), the optimal detector from Nuttall’s work has a markedly higher power than the square. This paper builds on this insight but with suitable adaptation for the highly colored and variable noise environment characteristic of the Southern California Bight, notably containing interfering sounds from large transiting vessels. Unlike most commonly used detectors, the generalized power-law detector (GPL) introduced here uses detection threshold parameters that are robust enough not to require operator adjustments while reviewing deployments with highly varying ocean noise conditions that can span months to years. Such a technique has the potential to significantly reduce operator analysis time for determining humpback presence/absence information, as well as the capacity to determine basic call unit parameters, such as unit duration, that are normally time-prohibitive to obtain using manual techniques. The goal for this detector is to detect nearly all humanly-audible humpback call units, allowing for occasional false detections in periods of heavy shipping. This detector is not designed to discriminate between transient biological signals that occur in overlapping spectral bands and of similar duration. However the method has a limited capacity for classification; namely the ability to separate shipping noise from narrowband, transient signals. Therefore, additional classification may be necessary if other acoustic sources meet the GPL detection criteria. Conversely, the GPL detector has proven to perform well for detecting other biological signals.

In unpublished experiments, suitable selection of spectral analysis parameters has provided good results for detecting blue whale (*Balaenoptera musculus*) "D" calls, minke (*Balaenoptera acutorostrata*) "boings", and killer whales (*Orcinus orca*) in the Southern California Bight (blue and minke whales) and in the coastal waters of Washington State (killer whales).

This paper is divided into six parts: Sect. 2.2 describes commonly-employed manual detection techniques, which guide the design constraints for an acceptable automated detector. Sect. 2.3 presents theoretical analysis for the GPL algorithm, highlighting the departures from the Nuttall form, which are motivated by these design constraints. Readers primarily interested in the application of the detector can move directly to Sect. 2.4, which discusses the particular application of the GPL algorithm to observational data, including the parameters chosen to best suit these data sets. Sect. 2.5 discusses the results of Monte Carlo simulations conducted to characterize the performance of the detector in comparison to: Nuttall's original power-law processor, the Erbe and King entropy method, and two energy-based detection algorithms. These simulations provide detection error trade-off (DET) curves for various humpback units, SNR, and noise conditions. In addition, results are given from simulations conducted to measure the performance of these algorithms against trained human analysts. Sect. 2.6 quantifies the ability of the GPL algorithm to measure call duration parameters. Finally, Sect. 2.7 presents the results from applying the GPL algorithm to 20 hours of recordings from three different deployments where humpback units were previously marked by trained human analysts. These 60 hours of acoustic data contain 21,037 individual humpback units occurring over a variety of ocean conditions and SNR. Although they perform poorly, the two energy detection algorithms are also included in this analysis because they are commonly used.

2.2 Detector design considerations

Detector design considerations were developed based on data sets collected by the Scripps Whale Acoustics Lab. However, similar detection requirements

are representative of the needs of the marine mammal acoustics community in general. The data sets for detecting humpback vocalizations were recorded by High-frequency Acoustic Recording Packages (HARP)[15]. These packages contain a hydrophone tethered above a seafloor-mounted instrument frame deployed in depths ranging from 200 m to 1500 m, covering a wide geographic area in the southern California Bight, and record more or less continuously over all seasons. HARP data are used to study the range and distribution of a wide variety of vocalizing marine mammals. The first step is to identify marine mammal vocalizations in the data. Depending on the type of marine mammal, this process can be labor intensive. Humpback recordings are particularly difficult. Humpback units can be described as transient signals, whose structure, strength, frequency, duration, and arrival time are unknown. Additionally, these vocalizations often occur in the same frequency bands that contain colored noise with additional contamination created by large transiting vessels. Depending on the distance of the passing ship, ship sounds can appear non-stationary over the same time scales as humpback units. The structure of the shipping noise is unknown but is often broadband. In practice, this complicated signal and noise environment often leads analysts to abandon automated detection entirely, relying on manual techniques for identifying vocalizations.

Various methodologies are used by the Whale Acoustics Lab to ensure consistent manual detection of marine mammal vocalizations. The Triton software package[16] was developed by the lab, providing the analyst with the ability to look at the time series and resulting spectrogram, with adjustable dynamic range, window lengths, filters, de-noising features, and audio playback. These manual detection techniques often find humpback units that are otherwise missed by standard automated detectors. While the ability to correctly mark the beginning and end time of each humpback unit is desirable, this step is time-prohibitive for longer data sets, and often only binary humpback presence/absence information is logged.

An acceptable automated humpback whale detector must be able to keep the probability of missed detections (P_{MD}) at or below the level of trained

human analysts, with a P_{FA} less than 6% in the noisiest environments. The amount of analyst review time required to separate humpback units from false detections depends upon both P_{FA} and the level of humpback vocalization activity. In practice, the 6% limit on P_{FA} necessitated 16 hours of review for a 365 day continuously recorded deployment in the southern California Bight, containing greater than one million humpback units. A reliable fixed detection threshold which fits within these constraints is desired for the entire deployment. Additionally, the algorithm must run significantly faster than real-time and provide accurate humpback unit start times and end times.

2.3 Theory

One approach for detecting signals with unknown location, structure, extent, and arbitrary strength is the power-law processor. Using the likelihood ratio test, Nuttall derives the conditions for near-optimal performance of this processor in the presence of white noise, based on appropriate approximations[14]. Nuttall's signal absent hypothesis (H_0) is equivalent to assuming that the Short Time Fourier Transform (STFT) of the time series yields independent, identically distributed (iid) exponential random variables of unit norm. The signal present hypothesis (H_1) is that the STFT consists of two exponential populations. Wang and Willet[17] represent these exponential populations as:

$$\begin{aligned}
 H_0 : f(\mathbf{X}) &= \prod_{k=1}^K \frac{1}{\lambda_0} e^{-|X_k|^2/\lambda_0} \\
 H_1 : f(\mathbf{X}) &= \prod_{k \notin S} \frac{1}{\lambda_0} e^{-|X_k|^2/\lambda_0} \times \prod_{k \in S} \frac{1}{\lambda_1} e^{-|X_k|^2/\lambda_1}
 \end{aligned} \tag{2.1}$$

where

λ mean square amplitude;

K total number of frequency bins;

\mathbf{X} Fourier vector with components X_k ;

S subset of size M , the number of frequency bins occupied by signal.

(Notation here and in succeeding sections is standard for probability theory[18]: F is used to denote the cumulative distribution function (cdf) and f denotes the probability density function (pdf). In addition the upper case letters Y, Z denote general random variables and the lower case letters y, z are specific realizations of them. Owing to the particular needs of this paper, X is reserved for Fourier components. The upper case E indicates the expectation operator.) Application of the likelihood ratio test requires summing over all combinatorial possibilities in H_1 . For even moderate M , this step becomes infeasible. Hence, Nuttall develops various approximations to estimate a threshold for a power-law detection statistic of the form

$$T(\mathbf{X}) = \sum_{k=1}^K |X_k|^{2\nu}. \quad (2.2)$$

The variable ν is an adjustable exponent that can be optimized for a particular M . For the idealized case of white noise, Nuttall's work indicates a general purpose value of $\nu = 2.5$ when M is completely unknown. For a single snapshot in time one can assume that for a humpback unit the number of signal bins M is much less than the total number of bins K , which favors $\nu > 2.5$. A summation of energy over all STFT bins is equivalent to $\nu = 1$, which is only optimal for $M = K$, and hence inappropriate here. Nonetheless, it is used extensively in readily available marine mammal detection software, and so its performance is noted throughout this manuscript.

A complication in the determination of an optimal ν is that most data sets contain shipping sounds in addition to the colored noise typical of the marine environment. A trade-off is created between values of ν that favor humpback vocalizations and larger values that better discriminate against broadband shipping sounds. No single choice of ν can be ideal for both purposes, however, a generalized power-law detector can achieve a suitable compromise between these alternatives as well as a fixed threshold in all noise environments. The definition of this detection

problem is as follows:

$$\begin{aligned}
 H_0 : & \quad \begin{cases} n(t) & \text{or} \\ n(t) + s_1(t) \end{cases} \\
 H_1 : & \quad \begin{cases} n(t) + s_2(t) & \text{or} \\ n(t) + s_1(t) + s_2(t) \end{cases}
 \end{aligned} \tag{2.3}$$

where $n(t)$ is a time series generated from distant shipping and wind, which can be modeled as a Gaussian distributed stochastic process. Local shipping sounds created by a single nearby ship are represented by $s_1(t)$, which can be both non-stationary and contain intermittent coherent broadband structure in frequency. The quantity $s_2(t)$ is the humpback vocalization signal. Although not a contributing factor in the datasets used in this work, any additional acoustic sources determined not to be humpback whales are also considered noise, and categorized as H_0 . Associated with these hypotheses is a formal optimization problem subject to nonlinear inequality constraints:

$$\min_{\Theta} P_{FA}(T^g(\mathbf{X}; \Theta)) \tag{2.4}$$

subject to:

$$\begin{aligned}
 P(T^g(\mathbf{X}; \Theta) < \eta_{\text{thresh}} | H_1) &= P_{MD} \leq P_{MD}^H \\
 P(T^g(\mathbf{X}; \Theta) > \eta_{\text{thresh}} | H_0) &= P_{FA} \leq P_{FA}^{max}
 \end{aligned} \tag{2.5}$$

where

$T^g(\mathbf{X}; \Theta)$	generalized power-law detection statistic;
η_{thresh}	detector threshold value;
P_{FA}	detector probability of false alarms;
P_{FA}^{max}	upper bound on false alarms (6%);
P_{MD}	detector probability of missed detection;
P_{MD}^H	human probability of missed detection;
Θ	model parameters.

Hereafter, the argument Θ will be dropped, its dependence implicit. Note that the superscript g distinguishes the GPL power-law detector from the Nuttall form.

To be considered an acceptable solution, a constant set of values for Θ , including η_{thresh} , is necessary. As in many other constrained optimization problems, the optimal solution is likely to be attained by an end-point minimum. A more traditional approach would be to permit detection on both $s_1(t)$ and $s_2(t)$, deferring discrimination to subsequent classification. While further classification is always possible, it turns out that this discrimination can be done largely at the detection stage if the power-law processor is suitably adapted. This goal is in the spirit of Wang and Willet[17], who developed a plug-in transient detector suitably adapted for a colored noise environment.

The characteristics described for $s_1(t)$ require examination of whitening, normalization, and broadband noise suppression. The non-stationary nature of $s_1(t)$ and the time clustered nature of $s_2(t)$ together motivate the choice of a conditional whitener insensitive to outliers. Similarly, while stationary noise motivates a simple estimator to produce the desired unit mean noise level, this normalization is less appropriate for the varying noise environments of H_0 , where it is more important to bound the largest values generated by the test statistic. Lastly, broadband suppression requires unit normalization across frequency in addition to normalization within frequency.

Another consideration is discrimination based on temporal persistence of the test statistic. Provided ν is appropriately chosen, local shipping characteristically generates highly intermittent values of the test statistic while humpback vocalizations exhibit continuity in the test statistic over the typically longer duration of the call unit. An event is defined as a continuous sequence of test statistic values at least one of which exceeds a prescribed value η_{thresh} and which is delimited on each side by the first point for which the test statistic is at or below η_{noise} , a noise baseline. The expectation with this definition is that an event corresponds to a humpback call unit, and as such a minimum unit duration, τ_c , is a reasonable additional model parameter to incorporate into the detector (discussed in Sect. 2.4). Because the statistical distributions $H_{0,1}$ cannot be solved

for analytically, η_{thresh} and η_{noise} are determined empirically with guidance from theory.

The proposed modification of the power-law statistic that incorporates these adaptations and also reflects the time dependence, j , can be written in its most general form as

$$T^g(\mathbf{X})_j = \sum_{k=1}^K a_{k,j}^{2\nu_1} b_{k,j}^{2\nu_2} \equiv \sum_{k=1}^K n_{k,j}, \quad (2.6)$$

$$a_{k,j} = \frac{||X_{k,j}|^\gamma - \mu_k|}{\sqrt{\sum_{n=1}^K (|X_{n,j}|^\gamma - \mu_n)^2}}, \quad (2.7)$$

$$b_{k,j} = \frac{||X_{k,j}|^\gamma - \mu_k|}{\sqrt{\sum_{m=1}^J (|X_{k,m}|^\gamma - \mu_k)^2}} \quad (2.8)$$

where

\mathbf{X}	now represents a Fourier matrix with J STFTs;
j	snapshot index ranging from 1 to J ;
k	frequency index ranging from 1 to K ;
$\{a, b, n\}_{k,j}$	elements in the matrices \mathbf{A} , \mathbf{B} , \mathbf{N} respectively;
ν_1, ν_2, γ	adjustable exponents;
μ_k	conditional whitener, defined below.

It is helpful to note that \mathbf{A} is a matrix whose columns are of unit length. The normalization across frequency (Eq. (2.7)) enforces the desired broadband suppression. \mathbf{B} is a matrix whose rows are of unit length, resulting from a normalization across time (Eq. (2.8)). The average μ_k is defined by

$$\mu_k = \int_0^\infty z f_k(z) dz. \quad (2.9)$$

For the purpose of whitening, this is approximated by

$$\mu_k \approx \int_{F_k^{-1}(y_c)}^{F_k^{-1}(y_c+1/2)} z f_k(z) dz, \quad (2.10)$$

$$y_c = \min_{y \in [0, 1/2]} [F_k^{-1}(y + 1/2) - F_k^{-1}(y)]. \quad (2.11)$$

Eq. (2.10) includes fifty percent of the distribution centered about the steepest part of the cdf, corresponding to the peak of the pdf. This form is termed

“conditional” to reflect that the limits of integration are dynamically determined from the data rather than fixed, as in Eq. (2.9). This formula is one of several possible implementations of a whitener whose goal is to suppress one or more strong signals, such as the order-truncate-average[19]. Equation (2.10) is unbiased for f_k a symmetric pdf, but is biased to the low side for the skewed distributions of interest here. The bias is not large however hence a more elaborate estimator of μ_k has not been explored. The integrals are cast in discrete form as follows. Let s_j denote the sorted values (from small to large) of $|X_{k,j}|$ over $j = 1..J$ for a fixed k . Next find $j^* = \min_j (s_{j+J/2-1} - s_j)$. And finally

$$\mu_k = \frac{2}{J} \sum_{j=j^*}^{j^*+J/2-1} s_j.$$

The conditional restriction of the average to those points deemed in the noise level means that the numerators in Eqs. (2.7) and (2.8) using the μ_k above are not exactly zero mean, though small.

Obtaining analytical expressions in the analysis of Eqs. (2.6)–(2.11) for $H_{0,1}$ is a difficult task. However, the case of white noise permits reasonable progress in characterizing the normalization and the whitener, which are explored in the following subsections. For white noise, only the sum $\nu_1 + \nu_2$ matters and hence can be replaced by a single exponent ν . For conditions other than white noise, the choices of γ , ν_1 , and ν_2 must be set individually, deviating from Nuttall’s one parameter form. For the optimization problem stated in Eqs. (2.4) and (2.5), values of $\gamma = 1$, $\nu_1 = 1$, and $\nu_2 = 2$ yielded about the minimal P_{FA} . These values were obtained with the guidance of theory presented in the following subsections, and verified with Monte Carlo simulations and observational results. In the remainder of the paper, these are the values employed.

2.3.1 Statistics of unit normalization for white noise

To understand the importance of the normalized variables that enter into Eq. (2.6), consider the case of white noise. In this section, the focus is on normalization and hence μ_k is set to zero in Eq. (2.6). To represent the associated

Fourier coefficients X_k let

$$X_k = \frac{1}{\sqrt{2}} (\Re(X_k) + i \Im(X_k)) \quad (2.12)$$

where real and imaginary parts are each independent and identically distributed normal random variables of zero mean and unit variance. With this normalization, $|X_k|$ has a Rayleigh distribution, $E(|X_k|) = \sqrt{\pi}/2$, and $E(|X_k|^2) = 1$, independent of frequency.

First consider the statistics of $a_{k,j}^2$ alone, hence define the random variable Y by

$$Y = \frac{|X_k|^2}{\sum_{n=1}^K |X_n|^2}, \quad (2.13)$$

where K is the number of Fourier frequency bins in the retained band. The matrix column index is omitted for the moment. The pdf for Y , $f_Y(y)$, is now sought. Because the sum in the denominator includes the index k , it is not independent of the numerator. Accordingly it is useful to look instead at the reciprocal, which is denoted as $1 + Z$ where Z is then given by

$$Z = \frac{\sum_{n=1}^{K'} |X_n|^2}{|X_k|^2}. \quad (2.14)$$

and the prime on the sum denotes the restriction $n \neq k$. From this starting point, standard statistical arguments lead to the conclusion that Y has the exact pdf

$$f_Y(y) = (K - 1) (1 - y)^{K-2}. \quad (2.15)$$

(See the appendix for details. In practice a Hamming window is used with the STFT and so this result does not strictly apply. The practical differences in the distributions obtained with a window compared to those above are slight however.) From Eq. (2.15), it follows that $E(y) = 1/K$. Note that, also as expected from the normalized form, y is necessarily limited in range to $[0, 1]$. This reflects the stated preference of bounding the test statistic in lieu of enforcing a unit norm of the noise, as found in most implementations of the power-law processor. In the present case of white noise the distinction is trivial, but such a bound remains in force even for the complex environments of $H_{0,1}$.

Equation (2.15) is well approximated by the exponential form $(K - 1) \exp(-(K - 2)y)$ provided $\log(1 - y) \approx -y$. The result is not, however, exactly normalized. To form a suitable pdf it is appropriate to modify this expression to

$$f_Y(y) \sim (K - 2) e^{-(K-2)y}, \quad (2.16)$$

which has the proper unit area. A measure of the approximation error is seen in the modified mean, $E(y) = 1/(K - 2)$, which agrees with the exact result to only leading order in K . While Eq. (2.15) correctly incorporates the fact that y can never exceed unity, a consequence of the expansion is that Eq. (2.16) has an exponentially small tail extending to infinity.

As shown in the Appendix, for even the simplest product of \mathbf{A} and \mathbf{B} the statistics cannot be found in closed form. However, observe that if the denominator in Eq. (2.13) is replaced by its mean value of K , then the pdf for Y becomes simply a rescaled version of the numerator, namely $K \exp(-K y)$. This last result, while not formally asymptotic to Eq. (2.16), is nonetheless a useful approximation for large K , and hence in subsequent sections when values are referred back to Eqs. (2.6)–(2.8), all normalizations are replaced by their mean values.

2.3.2 Unnormalized statistics for white noise only, with mean removal

It is important to characterize the role of nonzero μ_k . The particular frequency is irrelevant hence the subscript k is dropped in this subsection and subsection C. For this purpose it is simplest to consider the unnormalized sum

$$Y = \sum_{n=1}^N ||X_n| - \mu|^p \quad (2.17)$$

where, with reference to Eq. (2.6), $p = 2\nu_1 + 2\nu_2$, leaving the summation index N general. In later plots $p = [2, 6, \infty]$ are considered. The first of these, $p = 2$, addresses statistics of the denominators in Eqs. (2.7) and (2.8), the last two cover the numerators of interest. The value of p can be regarded in visual terms as a contrast setting; small p corresponds to low contrast, large p corresponds

to high contrast, where ν_1 controls vertical contrast and ν_2 controls horizontal contrast through the relative weighting of the normalization (denominator) terms in Eqs. (2.7) and (2.8).

At certain points in this and the succeeding subsection, it is useful to form the related quantity

$$\left(\sum_{n=1}^N ||X_n| - \mu|^p \right)^{1/p}, \quad (2.18)$$

the classical L^p norm in \mathbb{R}^N to facilitate comparison of differing values of p . The limit of large p in this latter form yields the minimax, or infinity, norm which singles out the largest single entry in the k -th column. Using a measure with all its support concentrated at one point is probably not a good idea since humpback units commonly include very sharp upsweeps and downsweeps, as well as units with a number of harmonics of similar amplitudes. Additionally, if p is too large, temporal persistence of the test statistic is lost and discrimination between shipping and transients such as humpback units is compromised. As previously indicated, the optimal constrained solution of Eqs. (2.4) and (2.5) is achieved in the neighborhood of $(\nu_1 = 1, \nu_2 = 2)$ or equivalently $p = 6$.

Now $|X_n|$ is Rayleigh distributed with, as noted before, a mean of $\sqrt{\pi}/2$. Defining the random variable

$$Z = ||X_n| - \mu|^p, \quad (2.19)$$

the associated pdf follows by a change of independent variable (see Appendix). The mean, $\mu_Z^{(p)}$, and standard deviation, $\sigma_Z^{(p)}$, of Z can be calculated but the expressions become unwieldy so the exact result is given only for $p = 2$ in Table 2.1. The superscript (p) denotes the dependence on the exponent in Eq. (2.17). The salient features are: the value of moments grows exponentially with p and rate of exponential growth itself increases rapidly with the order of the moment. Hence the numerator and denominator in Eq. (2.6) do not approach the prediction of the central limit theorem at the same rate.

Evaluation of the N -fold convolution integral that represents the pdf for the sums in numerator and denominator leads to approximation in terms of the moment expansion of the characteristic function, of which the leading contribution

is given exactly by the central limit theorem. On this basis it is expected that Eq. (2.17) is well approximated as

$$Y \approx \mu_Z^{(p)} N + \sigma_Z^{(p)} N^{1/2} z_d \quad (2.20)$$

for sufficiently large N , where z_d is a normally distributed random variable of zero mean and unit variance. However, it remains to be shown whether or not the asymptotic normal form is in fact an accurate approximation of the actual distribution for parameter values that are typical in application.

The first correction to the Gaussian pdf is the skewness, given by

$$c_3 = \int_{-\infty}^{\infty} Z_d^3 f_{Z_d} dZ_d = \frac{\rho_Z^{(p)}}{6 \sqrt{2} N \pi (\sigma_Z^{(p)})^3},$$

and $\rho_Z^{(p)} = E(|Z|^3)$. Scaling the random variable by $\sqrt{2} N \sigma_Z^{(p)}$ to express it in terms of z_d , the corrected pdf assumes the form

$$f_Y \sim e^{-z_d^2/2} (1 + c_3 z_d (z_d^2 - 3)) .$$

This is a good approximation provided

$$|z_d| \ll \sqrt[3]{6/\rho_Z^{(p)}} N^{1/6} \sigma_Z^{(p)} .$$

For $p = 2$, i.e. the denominator in Eq. (2.6), this results in $c_3 = 0.0150$ valid for $|z_d| \ll 3$ while for the numerator with $p = 6$, the skewness is nearly twenty times larger at $c_3 = 0.2644$ and consequently the expansion holds for $|z_d| \ll 1$, i.e., only the immediate vicinity of the peak of the pdf. Characterization of the tail of the distribution is given below.

Figure 2.1 shows computed pdfs for the L^P norm in Eq. (2.18) for $p = 2, 6, \infty$ along with the Gaussian pdf for comparison. It is seen that $p = 2$ lies close to the normal distribution while $p = 6$ is reasonably close to the infinity norm pdf. This bears directly on the analysis in the final theory subsection.

Turning briefly to the tails of these distributions, see Fig. 2.2 where $\log(1 - F_Y)$ is plotted. The parabolic curves in each panel reflect the quadratic controlling factor in the asymptotic expansion of the error function. This factor deviates significantly from the curve for $p = 6$; the controlling factor in the correct

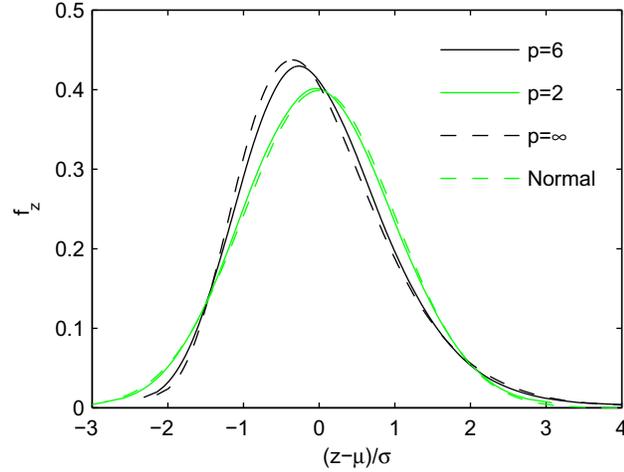


Figure 2.1: (Color online) Computed pdfs for the L^p norm in Eq. (2.18) for $p = 2, 6, \infty$ along with a Gaussian.

cdf is weaker than linear. How much weaker is made clear by switching from a global representation to a local approximation, namely

$$\log(1 - F_Y) \sim -\sqrt[3]{N} \left(\sqrt{\pi}/2 + y^{1/6} \right)^2 + \mathcal{O}(\log y). \quad (2.21)$$

Coefficients of the log and higher order corrections would derive from asymptotic matching. In lieu of that, here only the first term is used along with a numerically determined constant offset.

The results above individually characterize the numerator and denominator of Eq. (2.6). Because the terms in the denominator have large mean with small relative variance, as previously noted in Sect. 2.3.1, little error is incurred by replacing them with their mean value. It is really the numerator alone that controls the distribution of $T^g(\mathbf{X})$. For a normalized detector based strictly on *energy* ($p = 2$), no such partition is possible; the numerator and denominator scale comparably. This similarity of scaling is the basic cause of poor discrimination between shipping and humpback vocalizations for energy detectors.

The zeroth moment of the distribution is accurately estimated from the entries in Table 2.1 even though there is a long tail to the right, hence the average

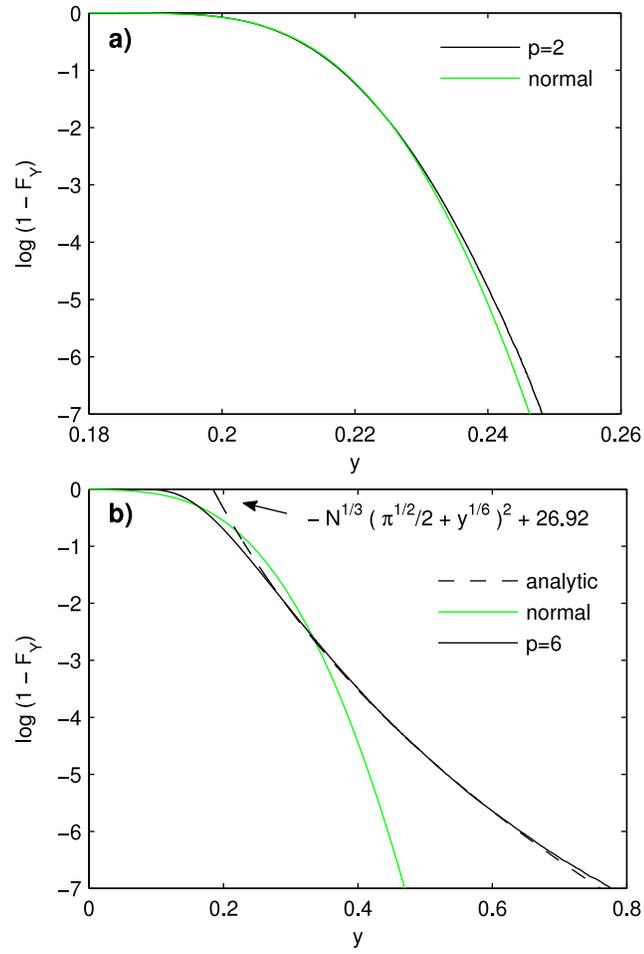


Figure 2.2: (Color online) A comparison of numerical and analytic forms for the cdf of Eq. (2.17) for a) $p = 2$ and b) $p = 6$, emphasizing the tail of the distribution.

test statistic for H_0 is

$$\bar{T}^g(\mathbf{X}) \approx \frac{\mu_Z^{(p)}}{J^{p/2-1} (\mu_Z^{(2)})^{p/2}}, \quad (2.22)$$

independent of K . For $J = 1460$, and $p = 6$, this works out to a prediction of $\bar{T}^g(\mathbf{X}) = 1.0223 \times 10^{-5}$. Simulations using Eq. (2.6) and the conditional whitener given in Eqs. (2.10) and (2.11) gives an average of 1.29×10^{-5} . In spite of real data leading to additional complications such as: 1) overlap of successive spectra, 2) dependence of the μ_k on frequency, 3) nonstationarity of shipping noise, and 4) sensor self-noise (discussed in Sect. 2.4), it is notable that the operational noise threshold for use with HARP data is set at $\eta_{\text{noise}} = 2.07 \times 10^{-5}$, just a factor of two larger than the value from Eq. (2.22). Recall the purpose of η_{noise} is to delimit the beginning time and end time of a particular humpback unit. Therefore, the final value was chosen in order to optimize the accuracy of this process, as described further in Sect. 2.6.

In lieu of a more elaborate model to incorporate the frequency dependence of μ_k , representative distributions are shown of $T^g(\mathbf{X})$ from recorded wind-driven noise, distant shipping, and local shipping data (discussed at greater length as Cases 1,2,3 respectively in Sect. 2.5) in comparison with the white noise result. In Fig. 2.3, a slightly different format for the tail of the distribution is used to bypass issues relating to a varying mean, μ_k , so the abscissa is now $\log(T^g(\mathbf{X}))$. Note how the tail of the wind-driven noise environment matches the ideal white noise result up to within a translation of about 0.5, which corresponds to a simple multiplicative rescaling of $T^g(\mathbf{X})$. The distributions of distant and local shipping, by contrast, decay more slowly although even for the latter on average a fraction of only about $\exp(-5)$ sample points per 75 s interval will exceed the indicated threshold. Whether these sample points produce an event detection is subject to the event duration requirement. Such persistent events come about not by a chance confluence of independent random spikes, which is quite rare, but from a spectral feature that does not fall to η_{noise} quickly enough to either side of the peak. How often that happens requires a more detailed model of shipping noise than is suitable to pursue here. A principal cause for excessively slow decay of the tail in Fig. 2.3 is failure of the whitener. During intervals of high level shipping, a prominent

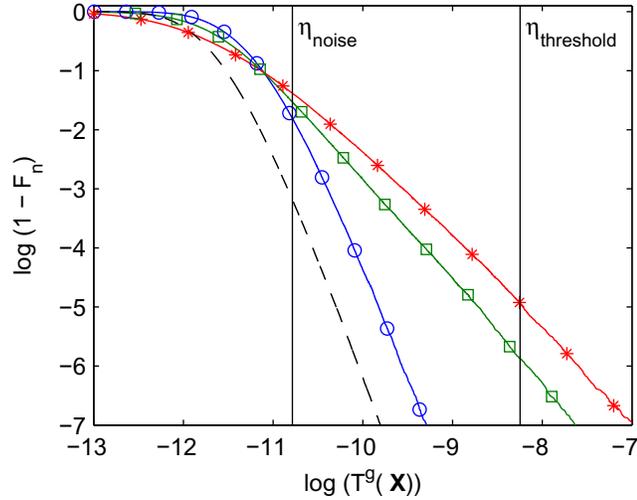


Figure 2.3: (Color online) Comparison of the tails of the cdfs for local shipping (asterisk), distant shipping (open square), and wind driven (open circle) noise conditions versus ideal white noise (dashed).

modulation of the spectrogram from ship propellor noise of 10 to 20 second period typically occurs. In this case, the use of a constant μ_k at each frequency over a time window of 75 s leaves a significant residual sinusoidal modulation.

2.3.3 Signal plus noise

To understand the response of GPL in the simplest setting the normalization can be omitted. Recall that its purpose is to allow fixed values for η_{noise} and η_{thresh} in $H_{0,1}$. With white noise of fixed variance this normalization is unnecessary. It is helpful here also to use the standard L^p form

$$\tilde{T}^g(\mathbf{X})_j^{(p)} = \left[\sum_{k=1}^K ||X_{k,j}| - \mu|^p \right]^{1/p}. \quad (2.23)$$

The tilde denotes the absence of normalization in the remainder of this subsection. The main issue is the statistics of an isolated snapshot. The correlation of $\tilde{T}^g(\mathbf{X})_j^{(p)}$ with adjacent values $\tilde{T}^g(\mathbf{X})_{j\pm 1}^{(p)}$ arising from overlap of successive STFT windows is hence neglected here. While characterizing the pdf for $\tilde{T}^g(\mathbf{X})^{(p)}$ in analytic form is not easy for intermediate p , the limiting case of the infinity norm is relatively

accessible. Moreover in Fig. 2.1, which shows the noise pdf for Eq. (2.23), the earlier noted similarity of results for $p = \infty$ and $p = 6$ suggests that qualitative aspects of the analysis below can be also expected to apply to the latter value of p .

For $p \rightarrow \infty$, Eq. (2.23) simplifies to

$$\tilde{T}^g(\mathbf{X})_j^{(\infty)} = \max_k ||X_{k,j}| - \mu|, \quad (2.24)$$

that is, the value assigned to \tilde{T} for time interval j is the single largest value in the k -th column of the whitened amplitude matrix. As an idealized model of this process, the signal is assumed to be a sine wave of amplitude s that lasts exactly one snapshot, superimposed on white noise. Denote the index of its frequency as k' . (The actual value is irrelevant in what follows.) What matters is that the maximum in Eq. (2.24) is taken over K values in the frequency domain. One of these values contains the signal plus noise; the remaining $K - 1$ contain only noise. For this detection scheme to be reliable, the signal must be large enough that the corresponding value of $||X_{k',j}| - \mu|$ exceeds the likely extremal value over the remaining $K - 1$ realizations of pure noise.

The cdf for the case of pure noise is given by

$$\tilde{F}_n(z; K - 1) = (1 - \exp(-(z + \mu)^2))^{K-1} \quad z > \mu. \quad (2.25)$$

For large K , the contribution in the range $z < \mu$ is exponentially small and may be neglected. The pdf for $||X_{k',j}| - \mu|$ is

$$\tilde{f}_s(z) = 2(z + \mu) \exp(-s^2 - (z + \mu)^2) \mathcal{I}_0(2s(z + \mu)) \quad z > \mu, \quad (2.26)$$

where \mathcal{I}_0 is the modified Bessel function of zeroth order. (For $0 \leq z \leq \mu$, the pdf is $\tilde{f}_s(z) + \tilde{f}_s(-z)$.) The accompanying cdf, $\tilde{F}_s(z)$, cannot be expressed in terms of known functions, however, its asymptotic and series expansions for large and small s respectively can both be found.

In terms of these quantities, the pdf for the random variable $z = \tilde{T}^g(\mathbf{X})$ summed over all frequencies including k' is given by

$$\tilde{f}_{GPL}^{(\infty)}(z) \sim \tilde{f}_s(z) \tilde{F}_n(z; K - 1) + \tilde{f}_n(z; K - 1) \tilde{F}_s(z), \quad (2.27)$$

with $K - 1$ equal to the total number of frequencies *not* counting that of the signal. From this construction, it follows automatically that $\int_0^\infty \tilde{f}_{GPL}^{(\infty)} dz = 1$. For large s and K Eq. (2.27) has the simple leading order asymptotic expansion

$$\tilde{f}_{GPL}^{(\infty)}(z) \sim \sqrt{\frac{z + \mu}{\pi s}} e^{-(z + \mu - s)^2}, \quad (2.28)$$

which is an excellent approximation for $s \geq 4$.

From the derivative of Eq. (2.25), the pdf of noise for $\tilde{f}_{GPL}^{(\infty)}$ reaches a maximum at $z \sim \sqrt{\log(K - 1)} - \mu$. The predicted separation of the peaks of signal plus noise and noise only pdfs is thus $s - \sqrt{\log(K - 1)}$. Pressing Eq. (2.28) somewhat beyond its formal range of applicability in this last result suggests for $K = 339$ that $s > 2.4$ is required for a signal to begin to emerge from the background. This predicted separation is qualitatively corroborated in Fig. 2.4a.

The case for the energy sum is given by Eq. (2.2) with $\nu = 1$. The sum of K noise terms has a cdf of $\Gamma(K, z)$. The pdf is well approximated by a normal distribution for the values of K considered here. The pdf for the signal follows from substituting $\mu = 0$ in Eq. (2.26) above and then making a variable change to reflect the choice of energy rather than amplitude as the independent variable. Hence

$$f_s(z) = \exp(-s^2 - z) \mathcal{I}_0(2s\sqrt{z}). \quad (2.29)$$

The equivalent of Eq. (2.27) is then given by the convolution

$$f_E(z) = \frac{1}{\Gamma(K)} \int_0^z (z - x)^{K-1} e^{x-z} f_s(x) dx. \quad (2.30)$$

This integral also cannot be found in closed form, but only approximated in various limits.

The displacement of the peak of f_E relative to the peak of the noise pdf at K is found to satisfy the approximate relation

$$4s^4 + (K - 1)(s^2 + z) = 2s^2 \frac{(2s^2 + K - 1)^{3/2}}{\sqrt{K - 1 + 2z}}, \quad (2.31)$$

which is equivalent to a cubic polynomial and has a K -independent exact root of $z = s^2$, as can be seen by inspection.

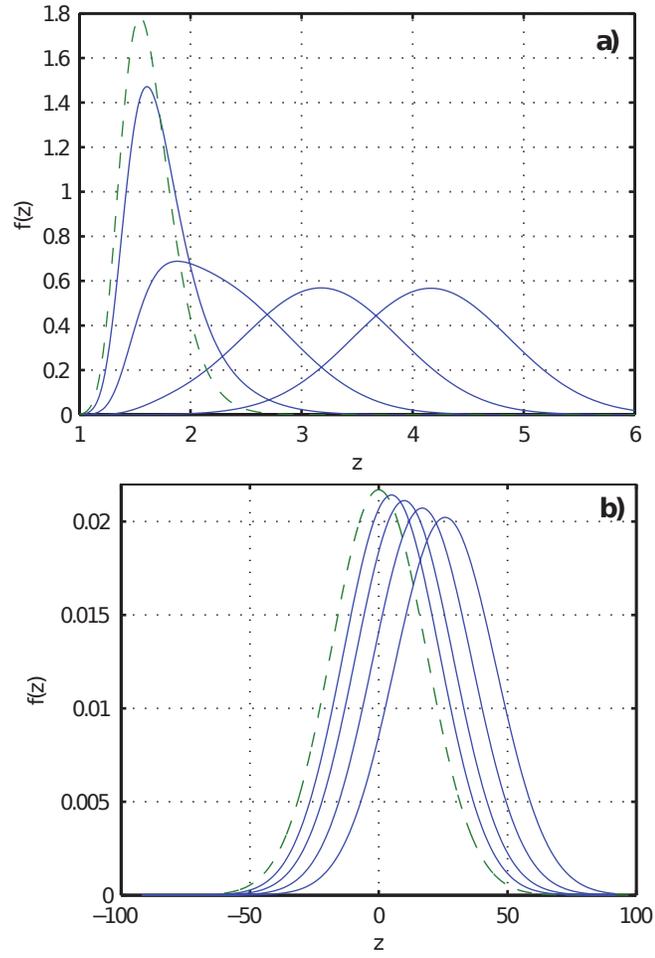


Figure 2.4: (Color online) Pdfs for a) $\tilde{f}_{GPL}^{(\infty)}$, b) f_E for signal amplitudes of 0 (dashed) and 2, 3, 4, 5 (solid) from left to right in each plot.

The plots in Fig. 2.4 show $\tilde{f}_{GPL}^{(\infty)}$ and f_E for signal amplitudes of $s = [0, 2, 3, 4, 5]$ (for, again, an rms noise amplitude of $\mu = \sqrt{\pi}/2$ per frequency and $K = 339$). Fig. 2.4 suggests that it takes about a 5 dB dynamic range for GPL to go from essentially no detection to nearly perfect detection. Taking $s = 4$ to define a suitable threshold for detection, it is useful for orientation to convert this choice of s into an associated (normalized) value of η_{thresh} for $p = 6$. The denominator of $T^g(\mathbf{X})$ is estimated as previously in Eq. (2.22). For the numerator it suffices to compute $\int_0^\infty z^6 \tilde{f}_s(z) dz$ with \tilde{f}_s as given in Eq. (2.26). The result is $\eta_{\text{thresh}} = 2.66 \times 10^{-4}$, virtually the exact value used in practice.

No algorithm based on $\nu = 1$ can compete with this performance; the linear separation of signal and noise with GPL is complete before the quadratic separation of the energy method begins to be effective. A formal measure of signal-to-noise statistics is the deflection ratio, defined as

$$d = \frac{|\mu_{s+n} - \mu_n|}{\sqrt{\sigma_{s+n}^2 + \sigma_n^2}}. \quad (2.32)$$

Asymptotic expansions for the means are tedious but, for large K , the distinction between the mean values and the peaks of the corresponding pdfs is slight. Accordingly the latter are used instead, yielding

$$d_{GPL} \approx \frac{\sqrt{2}(s - \sqrt{\log(K-1)})}{1 + 1/(2 \log(K-1))} \quad \text{and} \quad d_E \approx \frac{s^2}{\sqrt{2}K}. \quad (2.33)$$

The first of these reaches unit deflection ratio at $s = 3.2$, the second not until $s = 21.9$. Computed values of deflection ratio as defined in Eq. (2.32) based on statistics from simulations were compared against the analytical simplification for d_{GPL} in Eq. (2.33). Close agreement was found for $s > 4$, consistent with the approximation in Eq. (2.28) used to obtain d_{GPL} above. The computed values from simulation also corroborated a precise evaluation of Eq. (2.32) based on Gaussian quadrature with the exact pdf given in Eq. (2.27). Lastly, simulation confirms that $d_{GPL}(s)$ for $p = 6$ differs minimally from that for $p = \infty$, with an asymptotic slope reduced by only about 8%, thus discrimination for the ideal signal considered here is only slightly degraded by fixing $p = 6$ in place of the infinity norm, as anticipated.

Needless to say, real signals are not confined to a single frequency and the noise is neither white nor stationary. For these reasons, a more robust detector

is required but one that nonetheless approximates this sifting property of the L^∞ norm. The choice of $p = 6$ ($\nu_1 = 1$, $\nu_2 = 2$) is a good compromise.

2.3.4 Summary

It is not hard to see why GPL (or any other optimized power-law processor) is good at practical noise rejection: an overwhelming fraction of the final sample points $\{T^g(\mathbf{X})\}$ is tightly clustered near $\bar{T}^g(\mathbf{X})$. These points, which lie below η_{noise} , automatically define the snapshots at which events begin and end. Their ubiquity ensures that, although common noise sources (and ships particularly) do generate occasional spikes above threshold, the majority of the latter are subsequently discarded because their duration is nearly always less than the minimum unit duration subsequently imposed. More broadly, defining event duration is problematic for energy detection schemes both because no clean separation of signal and noise exists (equivalently the pdfs have excessive overlap) and because of the need to define an empirical *adaptive* threshold in contrast with the fixed value used in GPL.

What has been shown in the preceding subsections is that the modifications of normalization and whitening achieve white noise results comparable to those of Eq. (2.2). Analytical evaluation of these modifications in application to $H_{0,1}$ is not feasible. Rather, the evaluation is carried out in succeeding sections by means of both simulation and application to real data sets. It is shown that these modifications are necessary for an acceptable solution to the constrained optimization problem in Eqs. (2.4) and (2.5) using real ocean acoustic data and cannot be achieved with the power-law processor in Eq. (2.2).

2.4 Specific considerations for GPL algorithm used on HARP data for humpback detection

HARP data are recorded in either continuous or duty cycled format with a sampling frequency of 200 kHz. For the results presented in this paper, data

were processed in 75 s blocks, a time segment that was convenient for the duty cycle used in the HARP deployments. The time series is then lowpass filtered and decimated to a 10 kHz sampling rate. An STFT of length 2048 points is used with a 75% overlap and a Hamming window function, which corresponds to 4.9 Hz per frequency bin, 0.05 s per snapshot, and a total number of snapshots, J , equal to 1460. These parameters were found most effective for the majority of humpback vocalizations. The shortest call units could benefit from a shorter STFT length at the expense of a decrease in spectral resolution. No improvements in detection are realized for overlaps greater than 75%, therefore the overlap is fixed at 75% to avoid additional processing time. The output from the STFT is band-limited to a frequency range of 150 - 1800 Hz, and the number of frequency bins, K , is then 339. While humpback vocalizations can be recorded well above 1800 Hz and slightly below 150 Hz, sufficient energy for such units exists between these frequencies for good humpback detection performance.

The HARP data contain self-noise from the disk recording process. Therefore, a pattern matching algorithm based on singular value decomposition is used to remove short duration, broadband spectral features that coincide with the beginning and end of write-to-disk events. Additionally, the disk-write process produces narrowband, long duration (on the order of 10 s) noise contamination. While this narrowband noise is not problematic for higher order power-law processors, it does pose a problem for the energy-based detection methods (discussed in the following sections). Therefore, for energy detection only, a second algorithm is deployed that searches for the five strongest frequencies containing these narrowband features and removes these bands in the spectrogram. For both the energy methods and GPL, $|\mathbf{X}|$ as defined in Eqs. (2.7) and (2.8) is whitened following the discretized version of Eqs. (2.10) and (2.11), defining $|\widehat{\mathbf{X}}_k| = ||\mathbf{X}_k| - \mu_k|$.

Threshold values were guided by both the theoretical calculations and the nonlinear inequality constraints discussed in Sect. 2.3. Initially η_{thresh} was adjusted to match the performance of a trained human analyst. The theory in Sect. 2.3 provides an ex post facto analytical basis for this as a formal problem in separation

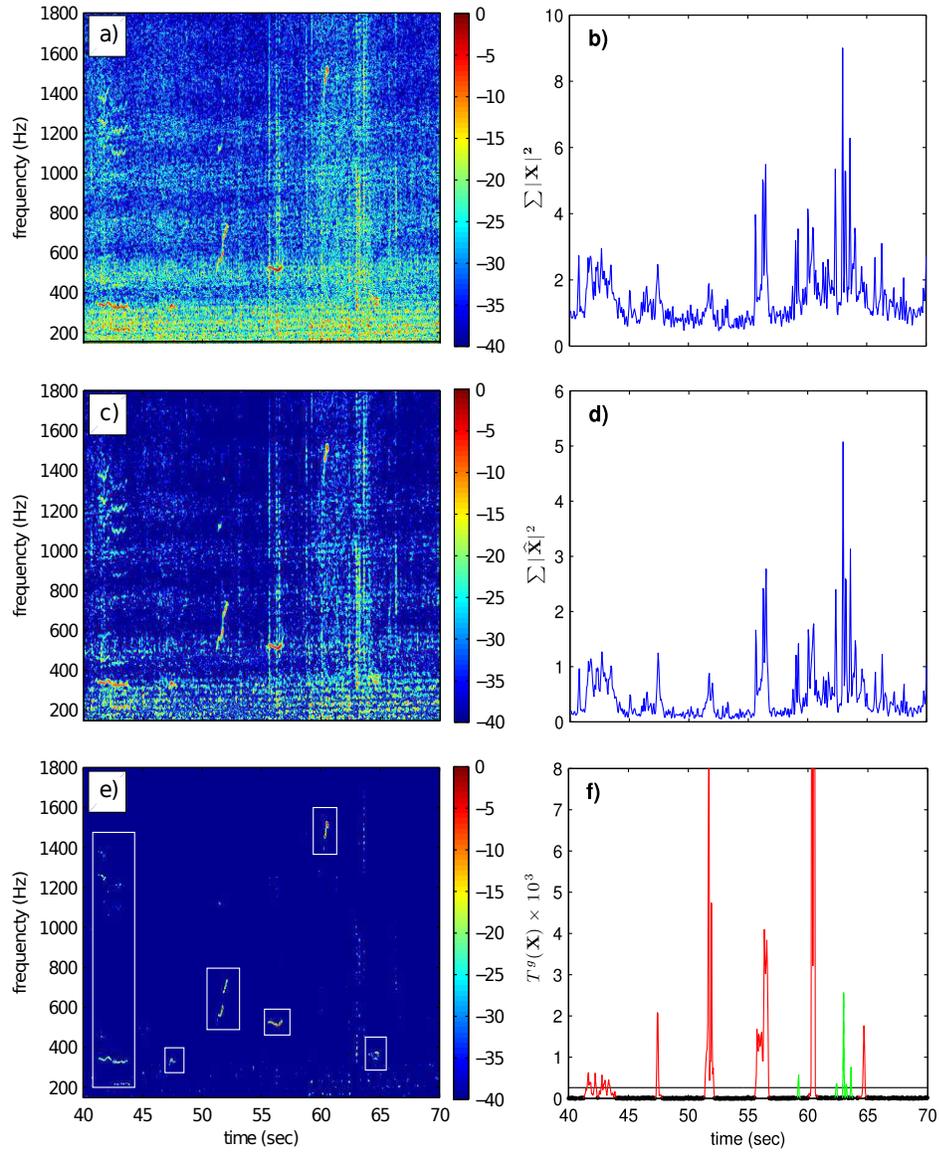


Figure 2.5: Visual comparison of energy and GPL for six humpback call units in the presence of local shipping noise starting with a) conventional spectrogram ($|\mathbf{X}|$) and b) resulting energy sum, c) energy with whitener ($|\hat{\mathbf{X}}|$), d) resulting sum, and finally e) \mathbf{N} as defined in Sect. 2.3, and f) GPL detector output $T^g(\mathbf{X})$. Units are highlighted in e) with white boxes. GPL detector output in f) shows eight groupings of detector statistic values above threshold (horizontal line). The six whale call units (red) meet the minimum time requirements, but the four detections (green) resulting from shipping noise do not, and so are not considered detections. All grams in units of normalized magnitude (dB).

of signal and noise. The simple choice of $s = 4$ gives a predicted η_{thresh} that lies fortuitously close to the chosen value but the factor of two discrepancy between the empirical and theoretical values for η_{noise} is more representative of the predictive accuracy one should expect. It was found that values of $\eta_{\text{noise}} = 2.07 \times 10^{-5}$ and $\eta_{\text{thresh}} = 2.62 \times 10^{-4}$ satisfied these constraints while keeping $P_{FA} < P_{FA}^{\text{max}}$ in the heaviest shipping environments. The detection test statistics for each time step j are evaluated according to Eqs. (2.6)-(2.8) as earlier noted using $\gamma = 1$, $\nu_1 = 1$, and $\nu_2 = 2$. Other values of γ , ν_1 , and ν_2 may be appropriate for other marine mammal vocalizations and/or noise conditions.

Using a normalized detection approach allows the user to set a fixed detection threshold, η_{thresh} , that works well over varying ocean conditions. However, during periods when the intercall interval between humpback units is short, the normalization approach reduces values of $T^g(\mathbf{X})$ for repeated units with shallow spectral slope, at times to values below η_{thresh} . Therefore, an iterative method is used in an attempt to adjust $|\hat{\mathbf{X}}|$ so that $T^g(\mathbf{X})$ gives similar values for a particular call unit, regardless of call activity. First a preprocessing step is done: T^g is computed from $|\hat{\mathbf{X}}|$. A submatrix $|\hat{\mathbf{X}}|^s$ is formed containing all columns of $|\hat{\mathbf{X}}|$ for which the corresponding $T^g < \eta_{\text{noise}}$. Next T^g is recomputed from $|\hat{\mathbf{X}}|^s$ with J adjusted to the size of the submatrix. All columns of $|\hat{\mathbf{X}}|^s$ for which $T^g > \eta_{\text{thresh}}$ are removed. Iteration then proceeds as follows:

T^g is computed from $|\hat{\mathbf{X}}|$. The detection with the highest value of T^g that exceeds threshold is recorded, its duration n fixed by the nearest neighbor to either side for which $T^g < \eta_{\text{noise}}$. Next the n columns in $|\hat{\mathbf{X}}|$ corresponding to this event are replaced by n columns of $|\hat{\mathbf{X}}|^s$ chosen at random. The process is repeated until no values of T^g exceed η_{thresh} .

In rare cases where the unit is repeated heavily, the normalization that reduces shipping noise also reduces the contribution of the calls to the test statistic. In such cases, the statistic may be below the detection threshold. Alternative techniques for normalization have shown promise.

It is possible to further reduce the effects of shipping noise in the data using a minimum unit duration requirement as described in the following. After

all events in the 75 second section of data have been determined, those events with a common terminus are merged into a single event. After qualifying events are merged, each event must exceed the minimum call duration requirement, τ_c , of 0.35 s. The modified detector output $T^{g*}(\mathbf{X})$ contains the values of $T^g(\mathbf{X})$ with detector values replaced by zero for events that do not meet these duration requirements. The formal optimization problems in Eqs. (2.4) and (2.5) should thus be changed so that $T^g(\mathbf{X})$ is replaced with $T^{g*}(\mathbf{X})$, and the model parameters contained in Θ are augmented to include $[\eta_{\text{thresh}}, \eta_{\text{noise}}, \tau_c]$. For an overlap of 75% a minimum call unit duration of 0.35 s corresponds to seven snapshots. The event duration, τ , is recorded for each detection. Shipping noise can sometimes produce high values of $T^g(\mathbf{X})$ albeit short in duration. Most of these events are shorter than τ_c . Using energy techniques, detections from shipping events and humpback units occur on similar time scales, and so this method of discrimination cannot be utilized. For comparison purposes, the performance of $T^g(\mathbf{X})$ and $T^{g*}(\mathbf{X})$ are discussed in the following sections.

Because the event duration is computed from Fourier components rather than the original time series, STFT length and window overlap define the terminal points of the event[20, 21]. For example, due to the 75% overlap, energy occurring entirely within the snapshot j can influence the test statistic from $X_{k,j-3}$ to $X_{k,j+3}$. This overlap can hence permit detection of events slightly shorter than τ_c which is useful in the case of detecting shorter humpback units, but can also increase false detection from shipping noise.

An example of the GPL process can be seen in Fig. 2.5, whose corresponding time series was created by adding a HARP recording containing strong shipping noise to a filtered HARP recording of humpback units (details discussed in Sect. 2.5 and shown in Fig. 2.6). Visual representations of \mathbf{X} , $|\widehat{\mathbf{X}}|$, and \mathbf{N} for 30 seconds of data are shown in Fig. 2.5(a,c,e). The incoherent sum over frequency for these matrices as a function of time are shown in Fig. 2.5(b,d,f), where Fig. 2.5(b) represents the energy sum, Fig. 2.5(d) represents the whitened energy sum, and Fig. 2.5(f) shows the values of $T^g(\mathbf{X})$. In Fig. 2.5(f) the detection threshold η_{thresh} is represented by a black horizontal line, while $T^g(\mathbf{X})$ values below the

noise level η_{noise} are illustrated with black dots. Events where $T^g(\mathbf{X}) > \eta_{\text{thresh}}$ are highlighted in red, while green represents events that fail to meet the event duration requirement in $T^{g^*}(\mathbf{X})$. The evolution from Fig. 2.5(b) to 2.5(f) shows significant improvement in humpback unit detectability: choosing a threshold value that would include all six humpback units in Fig. 2.5(b) would include a significant amount of shipping noise, while a threshold in Fig. 2.5(f) can be chosen to include all six humpback units with no inclusion of shipping noise.

The start time, end time, and duration for all events that meet detection requirements are recorded in a log file. A human analyst then prunes false detections from the log file. To aid operator review of the detections in a efficient manner, a graphical user interface (GUI) was designed. The GUI provides a tool for the operator to review time-condensed spectrograms containing the detections, to listen to the detections with adjustable band-passed audio, and to accept or reject each detection. The resulting subset of operator-selected detections can later be used for additional classification.

2.5 Monte Carlo simulations

In order to quantify the performance of GPL with known signals over a range of SNR, Monte Carlo simulations were conducted and the GPL algorithm performance was compared with Nuttall’s original power-law processor, two types of energy detection methods, Erbe and King’s entropy method, and trained human analysts.

Simulations were considered for three types of noise environments: wind dominated (Case 1), distant shipping (Case 2), and local shipping (Case 3). Case 1 approximates the circumstance of $H_0 = n(t)$, while Cases 2 and 3 reflect $H_0 = n(t) + s_1(t)$ with variation in relative contribution of single ship noise, $s_1(t)$, to the total noise field. It is worth noting that Case 3 is composed of shipping events recorded in the Santa Barbara channel when one or more large freight vessels were within 5 km of the HARP recording package (depth = 580 m). Six humpback units were selected that spanned varying frequency and temporal ranges in an attempt

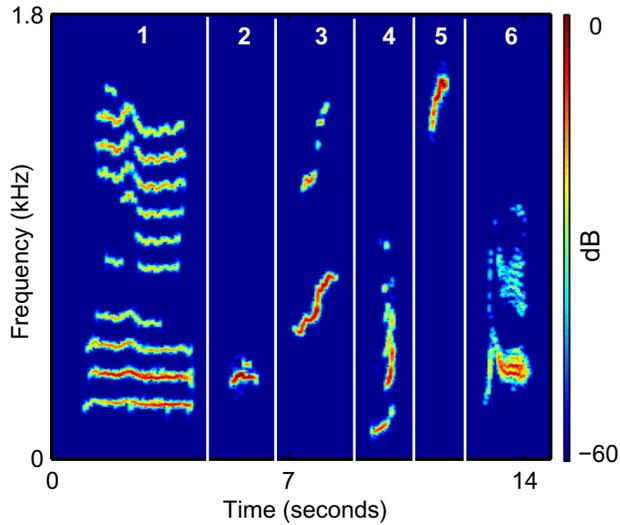


Figure 2.6: (Color online) Six humpback units used in Monte Carlo Simulations.

to characterize detector performance for the wide variety of humpback call units typically seen in acoustic recordings. Ninety-minute segments for each type of noise environment were selected from HARP data free of detectable humpback vocalizations and HARP self-noise. The six characteristic call units (shown in Fig. 2.6) were selected from a different HARP dataset that contained humpback vocalizations with high SNR. Noise in these recordings was further reduced using a masking filter in the Fourier domain, and then converted back to the time domain, to ensure that broadband background noise was not included in the signals of interest. Scalloping (spectral modulation) was avoided by using windows with 93.75% overlap, dividing out the window amplitude in each filtered STFT segment, and overlapping successive central segments by 50% [22]. Call units were added in the time domain to a random section of noise for each noise condition. Detection results were recorded for each detection method as described in Kay [23], using the binary hypothesis test in Eq. (2.3). Following Kay’s example, the observation interval is defined as the duration of the humpback unit of interest. When appropriate, detection error tradeoff (DET) curves[24] were created to compare

the performance of each detector with varying SNR, where SNR is defined as:

$$SNR = 10 \log_{10} \frac{\langle p_s^2 \rangle}{\langle p_n^2 \rangle}$$

where

$$\langle p_s^2 \rangle \equiv \frac{1}{T} \int_0^T p_s^2(t) dt$$

and where p represents the recorded pressure of the time series, bandpass filtered between 150 Hz and 1800 Hz, and T is the duration of the signal. Note that negative SNR in the time domain does not imply negative SNR for individual frequencies following a transformation into the Fourier domain. Detection Error Tradeoff curves are plots of the two error types from the binary hypothesis test: missed detections (P_{MD}) versus false alarms (P_{FA}). These error types are plotted as a function of detection threshold. DET curves are preferred over traditional receiver operator characteristic (ROC) curves[23] because the missed detection and false alarm axes are scaled to normal distribution fits of the scores of segments with and without signal. DET curves make use of the entire plotting space and are more capable of showing detail when comparing well-performing systems. Best detector performance in the DET space is represented by the point in the lower left corner of DET plots, where the P_{MD} is 0.05% and the P_{FA} is also 0.05%. The point in upper right corner of the plot represents no skill in the detector.

2.5.1 Simulations comparing detector performance

In addition to the entropy method described by Erbe and King, two types of energy detectors were included in the analysis. Detector $E^{(1)}$ is defined as a simple energy sum over the frequency range of 150 Hz to 1800 Hz, which is the equivalent to Nuttall's power-law processor described in Eq. (2.2) with $\nu = 1$. Assuming an approximate duration of the signal is known, $E^{(1)}$ can be enhanced by using a split window approach [25]. Detector $E^{(2)}$ represents this modified approach, as indicated in Eq. (2.34). For most units, $E^{(2)}$ performs optimally when the number of signal snapshots m_0 corresponds to one-third the signal duration and the number

of background snapshots M spans 20 s.

$$E_j^{(2)} = \frac{\sum_{m=-m_0}^{m_0} E_{j+m}^{(1)}}{\sum_{m=-M}^M E_{j+m}^{(1)} - \sum_{m=-m_0}^{m_0} E_{j+m}^{(1)}}. \quad (2.34)$$

The value of m_0 was adjusted for each unit type during the Monte Carlo simulations but in practice a single m_0 value would likely be chosen. Additionally, closely spaced call units were not in the simulations, allowing $E^{(2)}$ to perform at its best. Nuttall's power-law processor $T(\mathbf{X})$ was included in the analysis with an exponent $\nu = 3$, which was found to be the optimal exponent for the simulations. Simulations for GPL were conducted with and without the parameter metric enhancements $T^{g*}(\mathbf{X})$.

In order to minimize the influence of the whitener, both energy methods and the entropy method used the conditional whitener prescribed in Eqs. (2.10) and (2.11), as it increased performance for all three methods. The conditional whitener was not used with Nuttall's original power-law processor, as it decreased performance.

For each of the detectors, Monte Carlo simulations were conducted for all six unit types in Fig. 2.6, with SNR ranging from -10 dB to 10 dB, and noise Cases 1-3. Based on examination of trained human analysts' picks, a SNR of -3 dB corresponds to a human P_{MD} of approximately 15% in Case 1, 18% in Case 2, and over 20% for Case 3. The detector DET statistics for Units 1-6 were combined and are shown for each detector in Fig. 2.7 with 10,000 trials for each unit, noise condition and SNR. The GPL test statistic $T^g(\mathbf{X})$ is shown in preference to $T^{g*}(\mathbf{X})$ to put all the detection algorithms on an equal footing. In noise Case 1, all detection methods meet the inequality constraints in Eq. (2.5). In noise Case 2, both $T(\mathbf{X})$ and $T^g(\mathbf{X})$ meet the constraints. In noise Case 3, only $T^g(\mathbf{X})$ satisfies the constraints. The DET statistics do not address the stability of η_{thresh} among noise conditions, which is discussed further in succeeding sections. It is worth noting that the performance of $E^{(2)}$ is susceptible to considerable performance degradation when the short-term averaging duration is not selected carefully. In wind-driven noise conditions, it is found that a simple energy sum often has better detector performance than $E^{(2)}$. However, in the presence of shipping noise, detection method $E^{(2)}$ consistently

outperformed $E^{(1)}$.

Table 2.2 summarizes the GPL threshold DET statistics using the parameter enhancement $T^{g^*}(\mathbf{X})$ for all call units and noise conditions, over a range of SNR using the defined value for η_{thresh} . Threshold DET statistics are not provided for the other detection techniques since they do not satisfy the inequality constraints, and also establishing appropriate threshold values is somewhat arbitrary. GPL had nearly perfect detection scores for all six unit types in all three noise cases for SNR of 0 dB and higher. For SNR -2 dB, GPL had P_{MD} below 2% for all unit types and noise cases, except Unit 4. The majority of energy in Unit 4 is contained within a very narrow time interval of 0.3 s. Therefore, Unit 4 required slightly higher SNR than the rest of the unit types in order to consistently meet the minimum event duration requirement. It is also worth noting that the DET statistics are better in Cases 2 and 3 than Case 1 in very low SNR conditions. Since SNR is defined as the ratio of time-integrated squared pressure band-limited between 150 Hz to 1800 Hz, the low frequency distribution of noise in Case 2 and Case 3 can allow for locally higher SNR in the frequency bands in which the unit occurs, and results in an increase in detectability for very low SNR units. In general, units with the shortest durations, lowest frequencies, and units lacking frequency sweeps prove hardest to detect using the GPL algorithm. This result is expected, since units at low SNR with very short duration may be rejected for failing to meet τ_c . Low frequency units tend to be more susceptible to masking by shipping, and monotone units are more liable to be suppressed during normalization. The first two weaknesses in detection are also shared by human analysts, the third applies to GPL alone.

Humpback call analysts would like the ability to categorize humpback song into types of units. To this end, Table 2.2 will help provide guidelines for minimum SNR conditions that should be met before the detector can reliably detect all humpback units. The augmented model parameters $[\Theta, \eta_{\text{thresh}}, \eta_{\text{noise}}, \tau_c]$ were found to be robust for two years of data analyzed at multiple locations throughout the southern California Bight, the coast of Washington state, and Hawaii. However, these values may need to be adjusted slightly if ocean noise conditions change

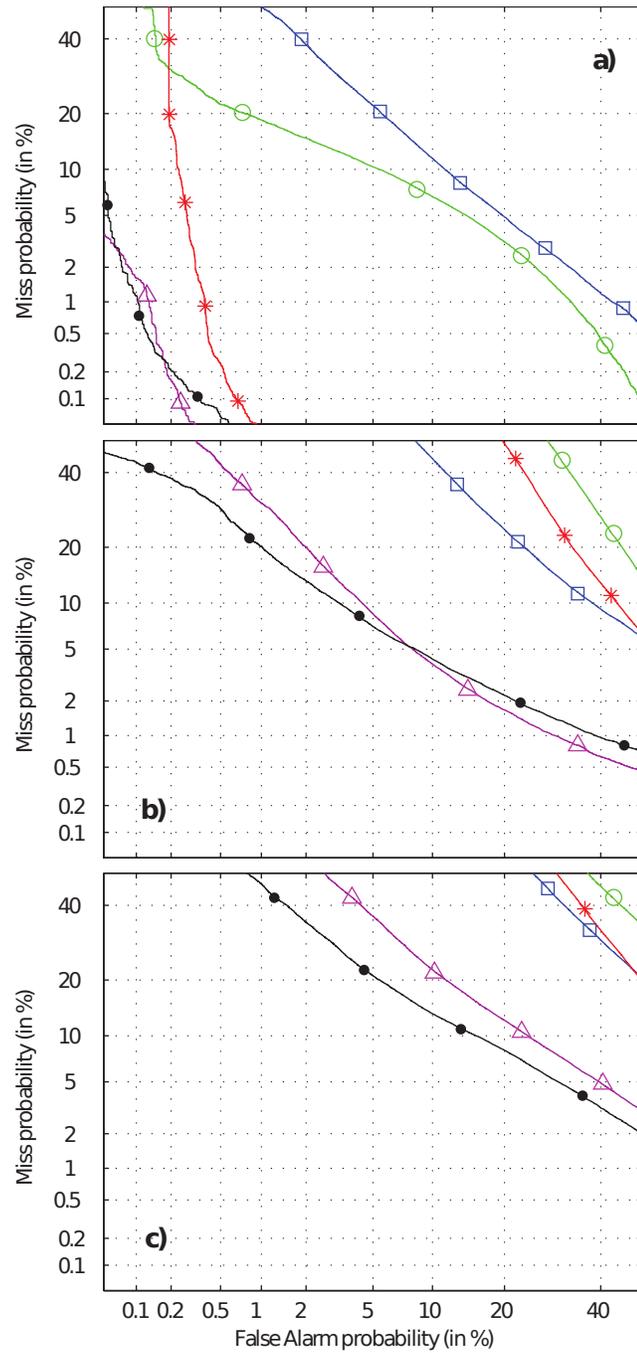


Figure 2.7: (Color online) DET results for Units 1-6 with SNR -3 dB in noise dominated by a) wind-driven noise, b) distant shipping, and c) local shipping, for GPL (closed circle), Nuttall (open triangle), entropy (asterisk), $E^{(1)}$ (open circle), and $E^{(2)}$ (open square).

appreciably from the noise recorded at these locations. Hydrophones located at shallower depths, sea ice noise, and the presence of noise generated from oil exploration are some circumstances that may warrant adjustments.

2.5.2 Simulations comparing power-law detectors to trained human analysts

A second set of simulations was conducted in order to compare the performance of $T^{g^*}(\mathbf{X})$ and Nuttall's test statistic $T(\mathbf{X})$ with trained human analysts. Here, five additional humpback units were included with the original six units shown in Fig. 2.6 in order to prevent the operators from recognizing repeated units. These eleven units were inserted into the ninety-minute recordings of Cases 1-3 with varying SNR, totaling 220 units for each of the three noise conditions. Each human analyst was asked to identify all humpback units and was not told the number, locations, or SNR of the signals present. The GPL P_{MD} values were calculated using the standard value of η_{thresh} , which was chosen so that $P_{FA} < P_{FA}^{max}$ for the strongest shipping conditions. The results using this threshold, shown in Table 2.3, illustrate that the GPL algorithm was able to detect lower SNR signals slightly better than the human analysts, and performed roughly on a par with the human analysts for higher SNR. Each operator was able to improve their performance by reviewing the output of the GPL detector.

For comparison purposes Eq. (2.2) with $\nu = 3$ was included in Table 2.3 to show the performance of a constant threshold using Nuttall's original power-law processor. A threshold was chosen using the same construction as for GPL, shown in Fig. 2.3, limiting the relative proportion of false detections in Case 3 to the same level. In doing so, the P_{MD} for Cases 1 and 2 violate the constraints stated in Eqs. (2.4) and (2.5), as humans were able to identify a significantly higher number of units at low SNR. For this reason Eq. (2.2) is not further considered.

2.6 Parameter estimation

In addition to detecting the presence or absence of a humpback unit, it is often desired to mark the beginning and end times of the humpback unit in the time series. If this can be done automatically and accurately, then that unit can be selected from the time series and passed to a classification scheme that can measure additional metrics about the unit. Even without further classification, unit timing parameters are provided by GPL itself, providing useful statistics on call rate, repetition, and both short-term and long-term calling trends. Parameter estimation algorithms and human analysts may provide different start and end time estimates for the same call unit depending on the noise condition and SNR. As SNR decreases, the edges of the unit may often be indistinguishable from the noise, and so a human analyst or automated algorithm tends to mark a shorter unit duration at lower SNR, even when the vocalizing source is producing a unit with the same duration in both cases. Additionally, all three detectors and human analysts are subject to the limitations imposed by the STFT length and window overlap as previously discussed. The bias and standard deviation in estimating unit duration are documented in this section for the GPL algorithm over a range of SNR, noise conditions, and unit types. Using the same six unit types from the Monte Carlo simulations, the units were inserted into the three noise conditions with SNR varying from -4 dB to 10 dB, with 500 trials per condition. For comparison, the two energy detectors were also included in this analysis, where the unit duration was marked by the time that passed in which the energy of the unit was above threshold. This method is similar to that used in Ishmael[9], in which the user is able to extract time series segments for calls that pass the user-defined threshold. For consistency in comparison with GPL, a threshold value for the energy techniques was chosen in which on average the P_{MD} was 10% for call Units 1-6 for noise Case 1, with SNR of -2 dB. For noise Case 1, an SNR of -2 dB was sufficiently high for a human to consistently and accurately detect nearly all call units in the record. The threshold and baseline values for marking call units with the GPL algorithm remained consistent with those described in Sect. 2.4.

Table 2.4 shows call duration parameters for Units 1 and 3, with Unit

1 representing the most error in parameter estimation for GPL, while Unit 3 represents typical performance. The quantity Δt_s represents the bias of the estimated unit start time in seconds from the true unit start time ($\hat{t}_s - t_s^{true}$), σ_s represents the standard deviation of \hat{t}_s . Likewise, the quantity Δt_e represents the bias in seconds of the unit end time estimate ($\hat{t}_e - t_e^{true}$), and σ_e represents the standard deviation for \hat{t}_e .

For units greater than 2 dB SNR in noise Cases 1 and 2, GPL is able to accurately measure start and end times, with Δt_s and Δt_e at 0.09 s or smaller and both σ_s and σ_e at 0.10 s or smaller. The two energy methods are also fairly effective at measuring these parameters at 2 dB or higher in noise Case 1. $E^{(1)}$ is not useful in either noise Case 2 or 3, because the threshold chosen for $E^{(1)}$ to work well in noise Case 1 creates large overestimates when ship noise is present. While at first glance $E^{(2)}$ appears to also work well in noise cases 2 and 3, using the threshold optimized for noise Case 1 results in many false alarms. Raising the threshold reduces P_{FA} , but unit durations are then drastically underestimated and the standard deviation is large.

2.7 Observational results

The performance of GPL using $T^{g*}(\mathbf{X})$ was established for three HARP deployments with varying humpback unit structure, SNR, depth, and noise conditions. Although the entropy detector, Nuttall's original power-law processor, and the energy methods violate the constraints in Eq. (2.5), $E^{(1)}$ and $E^{(2)}$ were included in the observational results because of their prevalence in marine mammal detection software. Twenty hours of acoustic recordings were first examined by trained human analysts, and humpback call units were identified for each of the three locations off the California coast. Additionally, operators reviewed the detections produced by GPL and energy-based methods in order to include any units first missed by the operators but captured by the detectors. Unlike the Monte Carlo simulations where the humpback unit locations are known regardless of signal strength, in the observational data the locations of humpback units are

only known within the detection ability of a trained operator. This operator-derived information was used as ground truth. As in the Monte Carlo simulations, binary hypothesis test metrics are used to evaluate the detector performances. An observation interval of 3 s is used for determining the detector output. Specifically, the maximum value of each detector output is recorded in a 3 s window surrounding each known humpback unit. The portions of the acoustic record that contained only noise are also broken into 3 s observation windows. The maximum detector output is recorded for each noise observation window using the same method as the signal-present windows. DET curves were produced for each of the three HARP deployments for GPL, $E^{(1)}$, and $E^{(2)}$.

Site SurRidge is 50 km southwest of Monterey, and the recording package is at a depth of 1386 m. Site B, located inside the Santa Barbara shipping channel, is 25 km north of Santa Rosa Island and the recording package is at a depth of 580 m. Site N is located 50 km southwest of San Clemente Island, and contains a recording package at a depth of 750 m.

Fig. 2.8(a) shows the DET curves for twenty hours of duty cycled acoustic recordings at site SurRidge spanning January 26-28, 2008. The analysis period contains 1,041 humpback call units, with most units categorized as low SNR with few identifiable harmonics. Local shipping noise is dominant during 14% of the record, distant shipping is dominant during 62% of the record, and wind-dominated noise is dominant during 24% of the record. Both $E^{(1)}$ and $E^{(2)}$ perform poorly during this period, with $E^{(1)}$ performing worse than $E^{(2)}$. The GPL algorithm performs reasonably well, and is able to detect all the units marked by the operator with a 4% P_{FA} .

Fig. 2.8(b) shows the DET curves for twenty hours of duty cycled recordings at site B spanning April 16-18, 2008. The analysis period contains 4,546 humpback call units, with most units categorized as moderate SNR with occasional calling bouts with high SNR. Local shipping noise is dominant during 36% of the record, distant shipping is dominant during 59% of the record, and wind-dominated noise is dominant during 5% of the record. Both $E^{(1)}$ and $E^{(2)}$ perform poorly during this period, with $E^{(1)}$ performing worse than $E^{(2)}$. The GPL algorithm performs

well, and is able to detect all the units marked by the operator with just over 2% P_{FA} .

Fig. 2.8(c) shows the DET curves for twenty hours of continuous recordings at site N spanning December 6-7, 2009. The analysis period contains 15,450 humpback call units, with most units categorized as high SNR containing many harmonics, with occasional calling at low SNR. Local shipping noise is dominant during 15% of the record, distant shipping is dominant during 23% of the record, and wind-dominated noise is dominant during 62% of the record. The detector $E^{(1)}$ performs better than $E^{(2)}$ in this scenario, which can be attributed to the extremely high call rate for this recording. Because $E^{(2)}$ uses a short-term average compared with a long-term average, units in close proximity often decrease the detector output. Because the GPL algorithm uses an iterative strategy in determining units, it is less affected by high calling rates. Therefore, the GPL algorithm outperforms $E^{(1)}$ and $E^{(2)}$ by a wide margin in this environment, detecting every unit marked by the operator with just over 0.5% P_{FA} .

Each deployment contains a handful of questionable humpback signals. When the questionable signals are included as units, the P_{MD} becomes nonzero, but remains 2% or less for each deployment.

At first glance, the steep vertical slope of the DET curve for GPL performance in Fig. 2.8 can lead to the conclusion of an unstable detection threshold, because a seemingly small change in P_{FA} appears to have a large effect on P_{MD} . The reason for this steep slope is twofold: Using the statistic $T^{g*}(\mathbf{X})$ instead of $T^g(\mathbf{X})$ enhances the non-Gaussian distribution of the test statistic, as shown in the histogram in Fig. 2.9. Here, one can see that a vast majority of events have detector output values of zero, because detections that do not meet the τ_c duration requirement are forced to zero. This binary decision within the GPL logic creates a sharp, but stable elbow in the DET curve. Additionally, low SNR units that would have received low values of $T^{g*}(\mathbf{X})$ were not identified by human analysts, which also alters the shape of the DET curves as compared to Fig. 2.7.

In order to evaluate the stability in the GPL threshold value among the

Table 2.1: Distribution of Moments for Eq. (2.17).

p	$\mu_Z^{(p)}$	$(\sigma_Z^{(p)})^2$	$\rho_Z^{(p)}$
2	$1 - \pi/4$	$1 + \pi/2 - \pi^2/4$	$2 + 15\pi/8 - \pi^3/4$
4	0.1494	0.4842	0.6481×10^1
5	0.1663	0.1613×10^1	0.7703×10^2
6	0.2154	0.6654×10^1	0.1257×10^4
22	0.7885×10^5	0.1922×10^{18}	0.2279×10^{33}

three HARP deployments, the P_{FA} and P_{MD} are calculated using the standard threshold of $\eta_{\text{thresh}} = 2.62 \times 10^{-4}$. Site SurRidge had $P_{FA} = 3.7\%$ and $P_{MD} = 0\%$, site N had $P_{FA} = 1.1\%$ and $P_{MD} = 0\%$, and site B had $P_{FA} = 3.2\%$ and $P_{MD} = 0\%$. These results suggest that the chosen value of η_{thresh} is both a stable and a sensible choice for all three HARP deployments, despite varying signal and noise conditions. Undoubtedly, the GPL algorithm misses some humpback units that occurred in these records. However, since human analysts are used to establish a ground truth of humpback unit occurrences, the low P_{MD} values verify that the GPL algorithm is able to find nearly all units that could be verified by human analysts.

2.8 Conclusions

The generalized power-law processor outperforms energy detection techniques for finding humpback vocalizations in the presence of shipping noise and wind-generated noise in the southern California Bight. The normalization over both frequency and time permits fixed thresholds that can be used throughout long deployments having varying ocean noise conditions. The algorithm capitalizes on basic parameters of the signal and noise environments, yet remains general enough to capture all types of humpback units, without the need for predefined templates. The detector is designed to capture all humpback units that are detectable by trained human analysts, while maintaining a low probability of false alarms. The

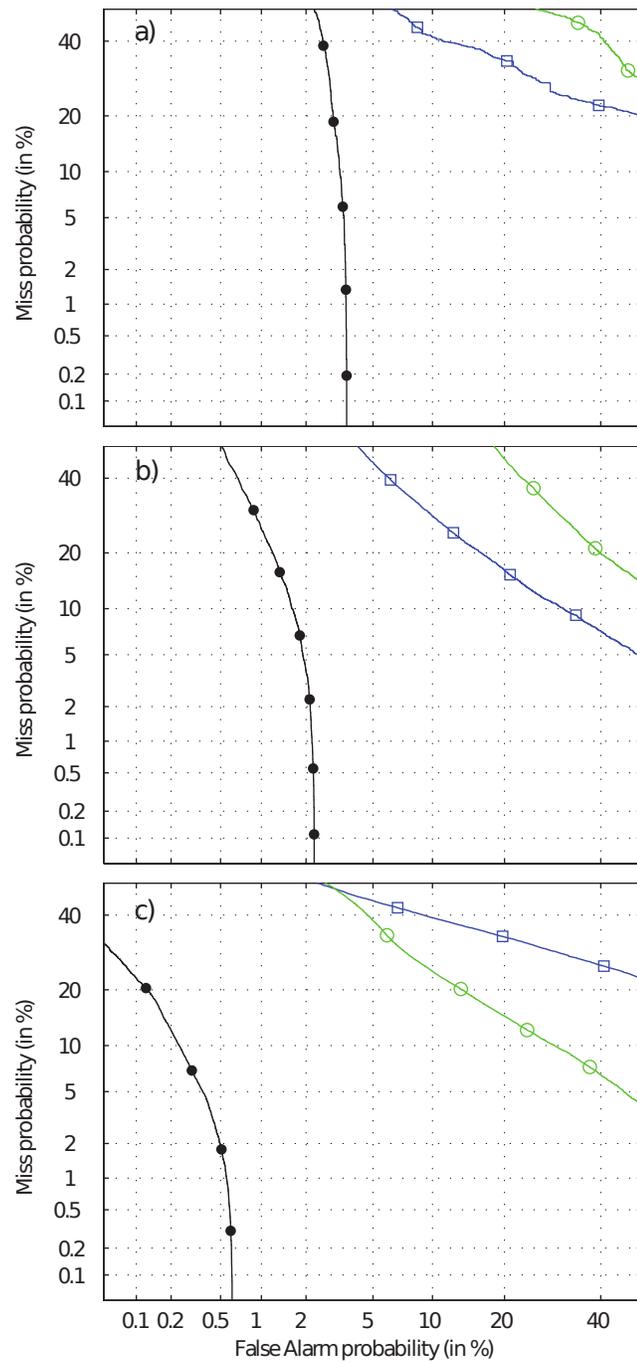


Figure 2.8: (Color online) DET results for HARP deployments at a) Site SurRidge, b) Site B, and c) Site N for GPL (closed circle), energy sums $E^{(1)}$ (open circle), and $E^{(2)}$ (open square).

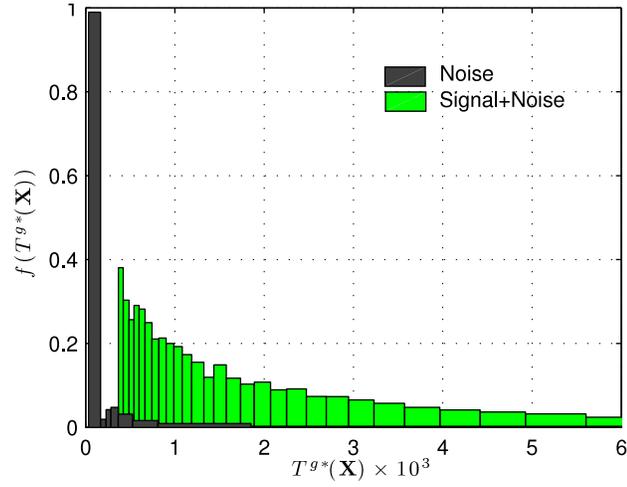


Figure 2.9: (Color online) Normalized histogram of detector outputs for signal and signal+noise for Site N deployment.

Table 2.2: Probability of missed detection and probability of false alarm (P_{MD}/P_{FA} , given as percentage) using η_{thresh} for Units 1-6, varying SNR and noise cases, 10,000 trials per statistic.

SNR	Noise	Unit 1	Unit 2	Unit 3	Unit 4	Unit 5	Unit 6
-6 dB	Case 1	98.5/1.0	87.2/0.0	98.2/0.0	100/0.0	98.9/0.0	95.4/0.0
	Case 2	87.9/4.8	77.7/4.7	84.0/4.9	94.7/4.5	78.8/4.1	89.6/4.5
	Case 3	78.5/6.0	81.6/5.7	73.1/6.5	92.1/5.7	31.6/5.0	83.2/4.7
-4 dB	Case 1	18.7/0.0	14.8/0.0	8.0/0.0	98.8/0.0	10.2/0.0	0.7/0.0
	Case 2	21.5/5.2	10.6/4.5	1.9/4.7	92.7/3.8	0.4/4.2	16.7/4.6
	Case 3	32.3/6.3	26.2/5.7	4.0/6.1	89.3/5.3	0.0/4.8	39.3/6.8
-2 dB	Case 1	0.0/0.0	0.0/0.0	0.0/0.0	23.8/0.0	0.0/0.0	0.0/0.0
	Case 2	0.1/5.0	0.1/4.3	0.0/4.9	47.0/4.1	0.0/4.2	0.2/4.8
	Case 3	0.0/6.9	0.6/5.6	0.0/6.6	62.2/5.3	0.0/5.2	1.6/6.5
0 dB	Case 1	0.0/0.0	0.0/0.0	0.0/0.0	0.0/0.0	0.0/0.0	0.0/0.0
	Case 2	0.0/5.1	0.0/4.4	0.0/4.8	3.4/4.4	0.0/4.5	0.0/5.1
	Case 3	0.0/6.3	0.0/5.3	0.0/6.7	0.0/5.5	0.0/5.0	0.0/6.4

Table 2.3: Probability of missed detection (P_{MD} , given as a percentage) for GPL versus baseline power-law detector (Nuttall) and human analysts for varying SNR. Detector threshold values were established such that Case 3 $P_{FA} < 6\%$ and applied to Cases 1 and 2.

	SNR	-6 dB	-4 dB	-2 dB	0 dB
Case 1	GPL	74.6	10.9	10.9	0.0
	Nuttall	94.6	32.7	10.9	0.0
	Analyst 1	74.6	21.8	12.7	3.6
	Analyst 2	76.4	18.2	9.1	5.4
Case 2	GPL	60.0	14.6	12.7	7.3
	Nuttall	81.8	41.8	14.6	7.3
	Analyst 1	78.0	24.0	12.0	6.0
	Analyst 2	81.9	27.3	10.9	7.3
Case 3	GPL	61.8	27.3	9.1	5.5
	Nuttall	61.8	29.1	7.3	3.6
	Analyst 1	84.0	48.0	14.0	14.0
	Analyst 2	56.4	23.7	7.3	3.7

Table 2.4: Start-time bias Δt_s , end time bias Δt_e , start time standard deviation σ_s , and end time stand deviation σ_e in seconds for Unit 1 (duration 3.34 s) and Unit 3 (duration 1.3 s)

		Noise Case 1				Noise Case 2				Noise Case 3			
Unit	type	Δt_s	σ_s	Δt_e	σ_e	Δt_s	σ_s	Δt_e	σ_e	Δt_s	σ_s	Δt_e	σ_e
Unit 1 -2 dB	E1	-1.38	0.63	-0.62	0.50	-0.78	2.27	-0.66	3.65	22.22	21.41	23.83	22.33
	E2	-1.00	0.41	-0.71	0.27	-0.96	0.55	-0.84	0.54	-1.00	0.71	-0.85	0.69
	GPL	-0.34	0.08	-0.02	0.06	-0.35	0.17	-0.16	0.33	-0.34	0.20	-0.19	0.28
Unit 1 0 dB	E1	-0.49	0.21	-0.23	0.10	-0.48	3.48	0.14	3.29	22.71	21.92	23.43	22.26
	E2	-0.43	0.06	-0.39	0.06	-0.46	0.22	-0.43	0.23	-0.50	0.35	-0.44	0.32
	GPL	-0.21	0.10	0.01	0.03	-0.21	0.14	-0.02	0.11	-0.22	0.14	-0.02	0.11
Unit 1 2 dB	E1	-0.31	0.10	-0.15	0.03	0.29	3.54	0.63	3.84	20.63	20.64	25.36	23.06
	E2	-0.28	0.04	-0.23	0.04	-0.29	0.10	-0.25	0.09	-0.29	0.15	-0.25	0.15
	GPL	-0.09	0.05	0.03	0.03	-0.09	0.10	0.02	0.08	-0.09	0.09	0.03	0.10
Unit 3	type	Δt_s	σ_s	Δt_e	σ_e	Δt_s	σ_s	Δt_e	σ_e	Δt_s	σ_s	Δt_e	σ_e
Unit 3 -2 dB	E1	-0.46	0.21	-0.36	0.16	0.26	3.63	0.34	4.28	23.04	22.36	23.95	23.18
	E2	-0.39	0.15	-0.47	0.19	-0.36	0.22	-0.41	0.18	-0.33	0.32	-0.36	0.33
	GPL	-0.01	0.05	0.01	0.04	0.02	0.16	0.04	0.15	0.00	0.12	0.05	0.13
Unit 3 0 dB	E1	-0.20	0.09	-0.20	0.04	0.43	4.31	0.59	4.49	22.46	22.62	22.58	22.58
	E2	-0.22	0.09	-0.29	0.06	-0.21	0.19	-0.29	0.14	-0.21	0.24	-0.29	0.23
	GPL	0.03	0.04	0.05	0.04	0.06	0.31	0.09	0.41	0.06	0.11	0.07	0.12
Unit 3 2 dB	E1	-0.11	0.03	-0.15	0.03	0.52	3.64	0.28	2.51	24.15	22.25	23.70	22.14
	E2	-0.07	0.05	-0.21	0.03	-0.08	0.10	-0.21	0.06	-0.09	0.18	-0.20	0.18
	GPL	0.06	0.04	0.07	0.03	0.07	0.08	0.08	0.06	0.08	0.11	0.10	0.12

detector performance was verified by inserting humpback units with varying SNR into three noise conditions and comparing the detector output to that of two trained operators. Additionally, the GPL algorithm is able to detect nearly all humpback units previously identified by human analysts in three different deployments off the coast of California, with a result of $P_{FA} = 3.7\%$ or better. This performance allows a human analyst to review a much smaller subset of data when looking for humpback units.

Once the periods of data containing humpback units have been identified, basic call parameters such as unit duration, center frequency, number of units, and inter-call interval can be automatically tabulated. The GPL process provides considerably more detail than basic presence/absence metrics to which human analysis is typically restricted, owing to the labor intensive nature of manually selecting individual units. Parameter estimation performance obtained from simulations show that GPL commonly yields precision of 0.1 s or less for estimating the beginning and end of a unit for reasonable SNR under all but heavy shipping noise. By contrast, measuring unit duration parameters using energy detection techniques proved unfeasible except in high SNR situations. Although the analysis here has focused on algorithm settings tuned to the specific characteristics of humpback vocalizations, the GPL algorithm has in fact the potential to be modified for many types of marine mammal vocalizations, and is likely to prove useful as a precursor to classification techniques.

2.A Mathematical details

The numerator in Eq. (2.14) has a pdf of $\chi_{K-1}^2(z)$ and the denominator $\chi_2^2(z)$ so the quantity $X/(K-1)$ is thus an F-distribution of the form

$$f_X(x) = \left(\frac{(K-1)x}{1+(K-1)x} \right)^{K-2} \left(\frac{K-1}{1+(K-1)x} \right)^2. \quad (2.35)$$

Observe that

$$\begin{aligned} P(Y < y) &= P(X > (K-1)^{-1}(1/y-1)) \\ &= 1 - F_X((K-1)^{-1}(1/y-1)), \end{aligned}$$

accordingly

$$\begin{aligned} f_Y(y) &= \frac{1}{y^2} f_X((K-1)^{-1}(1/y-1)) \\ &= (K-1)(1-y)^{K-2} \end{aligned} \quad (2.36)$$

and therefore

$$F_Y(y) = 1 - (1-y)^{K-1}.$$

With the statistics of entries in \mathbf{A} thus characterized, it is logical to try to extend this line of reasoning to the product form of Eq. (2.6) by attempting first to reproduce the equivalent of Eq. (2.15). For simplicity, consider $J = K$ and $\gamma = 1$. Then the reciprocal leads to a homogeneous form $1 + Z_1 + Z_2$ where

$$\begin{aligned} Z_1 &= \frac{\sum_{n=1}^{K'} |X_{n,j}|^2 + \sum_{m=1}^{K'} |X_{k,m}|^2}{|X_{k,j}|^2}, \\ Z_2 &= \frac{\sum_{n=1}^{K'} |X_{n,j}|^2 \sum_{m=1}^{K'} |X_{k,m}|^2}{|X_{k,j}|^4}. \end{aligned} \quad (2.37)$$

The first term in Eq. (2.38) is another F -distribution as in Eq. (2.35) but with K replaced by $2K$. The difficulty comes from the second term. For the second term the pdfs for its numerator and denominator are

$$\frac{(2K-3)z^{K-2}}{\Gamma(K-1/2)^2} \mathcal{K}_1(2\sqrt{z}) \quad \text{and} \quad \frac{1}{2} z^{-1/2} e^{-z^{1/2}}$$

respectively, where \mathcal{K} is the modified Bessel function of the second kind. This ratio is not an F -distribution and appears not to be characterized. Thus even for this first extension of normalization beyond Eq. (2.13), immediate recourse to asymptotic approximation is necessary.

Lastly, for the pdf governing Eq. (2.19) it is immediate on a change of variable that

$$\begin{aligned} f_Z^{(p)}(z) &= \frac{2}{pz^{(p-1)/p}} (\sqrt{\pi}/2 + \sqrt[2]{z}) e^{-(\sqrt{\pi}/2 + \sqrt[2]{z})^2} \\ &\quad z > \pi^{p/2}/2^p, \end{aligned} \quad (2.38)$$

and the symmetric combination $f_Z^{(p)}(z) + f_Z^{(p)}(-z)$ applies for $0 \leq z \leq \pi^{p/2}/2^p$ to account for both roots in that interval.

Acknowledgements

The authors are extremely grateful to Greg Campbell, Amanda Cummins, and Sara Kerosky, who provided operator-identified humpback whale unit locations and trained human analyst expertise. Special thanks to Sean Wiggins and the entire Scripps Whale Acoustics lab for providing thousands of hours of high quality acoustic recordings. Bill Hodgkiss was extremely helpful in providing feedback in areas of signal processing, Monte Carlo simulations, and detection theory. The authors are grateful to Peter Rickwood, who at the early stages in this work provided time, expertise, and software in our initial evaluation of schemes for classification. The first author would like to thank the Department of Defense Science, Mathematics and Research for Transformation Scholarship program, the Space and Naval Warfare Systems Command Center (SPAWAR) Pacific In-House Laboratory Independent Research program, and Rich Arrieta from the SPAWAR Unmanned Maritime Vehicles Lab for continued financial and technical support. Work was also supported by the Office of Naval Research, Code 32, CNO N45, and the Naval Postgraduate School.

Chapter 2 is, in full, a reprint of material published in *The Journal of the Acoustical Society of America*: Tyler A. Helble, Glenn R. Ierley, Gerald L. D'Spain, Marie A. Roch, and John A Hildebrand, "A generalized power-law detection algorithm for humpback whale vocalizations". The dissertation author was the primary investigator and author of this paper.

References

- [1] R.S. Payne and S. McVay. Songs of humpback whales. *Science*, 173(3997):585–597, 1971.
- [2] S. Cerchio, J.K. Jacobsen, and T.F. Norris. Temporal and geographical variation in songs of humpback whales, Megaptera novaeangliae: synchronous change in Hawaiian and Mexican breeding assemblages. *Animal Behaviour*, 62(2):313–329, 2001.
- [3] D.K. Mellinger and C.W. Clark. Recognizing transient low-frequency whale sounds by spectrogram correlation. *J. Acoust. Soc. Am.*, 107:3518–3529, 2000.

- [4] J.R. Potter, D.K. Mellinger, and C.W. Clark. Marine mammal call discrimination using artificial neural networks. *J. Acoust. Soc. Am.*, 96:1255–1262, 1994.
- [5] J.C. Brown and P. Smaragdis. Hidden Markov and Gaussian mixture models for automatic call classification. *J. Acoust. Soc. Am.*, 125(6):EL221–EL224, 2009.
- [6] P. Rickwood and A. Taylor. Methods for automatically analyzing humpback song units. *J. Acoust. Soc. Am.*, 123(3):1763–1772, 2008.
- [7] X. Mouy, M. Bahoura, and Y. Simard. Automatic recognition of fin and blue whale calls for real-time monitoring in the St. Lawrence. *J. Acoust. Soc. Am.*, 126:2918–2928, 2009.
- [8] T.A. Abbot, V.E. Premus, and P.A. Abbot. A real-time method for autonomous passive acoustic detection-classification of humpback whales. *J. Acoust. Soc. Am.*, 127:2894–2903, 2010.
- [9] D.K. Mellinger. Ishmael 1.0 users guide. *NOAA Technical Memorandum OAR PMEL-120, available from NOAA/PMEL*, 7600:98115–6349, 2001.
- [10] H. Figueroa. XBAT. v5. Cornell University Bioacoustics Research Program, 2007.
- [11] D. Gillespie, D.K. Mellinger, J. Gordon, D. McLaren, P. Redmond, R. McHugh, P. Trinder, X.Y. Deng, and A. Thode. PAMGUARD: Semiautomated, open source software for real-time acoustic detection and localization of cetaceans. *J. Acoust. Soc. Am.*, 125:2547–2547, 2009.
- [12] C. Erbe and A.R. King. Automatic detection of marine mammals using information entropy. *J. Acoust. Soc. Am.*, 124:2833–2840, 2008.
- [13] A.H. Nuttall. Detection performance of power-law processors for random signals of unknown location, structure, extent, and strength. *NUWC-NPT Tech. Rep*, 1994.
- [14] A.H. Nuttall. Near-optimum detection performance of power-law processors for random signals of unknown locations, structure, extent, and arbitrary strengths. *NUWC-NPT Tech. Rep*, 1996.
- [15] S. Wiggins. Autonomous Acoustic Recording Packages (ARPs) for long-term monitoring of whale sounds. *Marine Tech. Soc. J.*, 37(2):13–22, 2003.
- [16] S.M. Wiggins, M.A. Roch, and J.A. Hildebrand. Triton software package: Analyzing large passive acoustic monitoring data sets using matlab. *J. Acoust. Soc. Am.*, 128:2299–2299, 2010.

- [17] Z. Wang and P.K. Willett. All-purpose and plug-in power-law detectors for transient signals. *Signal Processing, IEEE Transactions on*, 49(11):2454–2466, 2001.
- [18] A. Stuart and K. Ord. *Kendall's advanced theory of statistics, Vol. 1: Distribution Theory*, chapter 1-5, 8-11. J. Wiley, New York, NY, 2009.
- [19] W.A. Struzinski and E.D. Lowe. A performance comparison of four noise background normalization schemes proposed for signal detection systems. *J. Acoust. Soc. Am.*, 76:1738–1742, 1984.
- [20] R.A. Charif, C.W. Clark, and K.M. Fristrup. Raven 1.2 users manual, Appendix B: A Biologists Introduction to Spectrum Analysis. *Cornell Laboratory of Ornithology, Ithaca, New York*, 2004.
- [21] M.D. Beecher. Spectrographic analysis of animal vocalizations: Implications of the uncertainty principle. *Bioacoustics*, 1:187–208, 1988.
- [22] R.W. Lowdermilk and F. Harris. Using the FFT as an arbitrary function generator. In *Proc. AUTOTESTCON 2005*, pages 408–412. IEEE, 2005.
- [23] S.M. Kay. *Fundamentals of Statistical Signal Processing: Detection Theory*, pages 7, 41, 238. Prentice-Hall, Englewood Cliffs, NJ, 1998.
- [24] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The DET curve in assessment of detection task performance. In *Proc. Eurospeech*, volume 97, pages 1895–1898, 1997.
- [25] R.O. Nielsen. *Sonar signal processing*, pages 145–147. Artech House, Inc., Norwood, MA, 1991.

Chapter 3

Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones

Abstract

Passive acoustic monitoring of marine mammal calls is an increasingly important method for assessing population numbers, distribution, and behavior. A common mistake in the analysis of marine mammal acoustic data is formulating conclusions about these animals without first understanding how environmental properties such as bathymetry, sediment properties, water column sound speed, and ocean acoustic noise influence the detection and character of vocalizations in the acoustic data. The approach in this paper is to use Monte Carlo simulations with a full wave field acoustic propagation model to characterize the site specific probability of detection of six types of humpback whale calls at three passive acoustic monitoring locations off the California coast. Results show that the probability of detection can vary by factors greater than ten when comparing detections across locations, or comparing detections at the same location over

time, due to environmental effects. Effects of uncertainties in the inputs to the propagation model are also quantified, and the model accuracy is assessed by comparing calling statistics amassed from 24,690 humpback units recorded in the month of October 2008. Under certain conditions, the probability of detection can be estimated with uncertainties sufficiently small to allow for accurate density estimates.

3.1 Introduction

A common mistake in passive acoustic monitoring of marine mammal vocalizations and other biological sounds is to assume many of the features in the recorded data are associated with properties of the marine animals themselves, without accounting for other important aspects. Once a sound is emitted by a marine animal, its propagation through the ocean environment can cause significant distortion and loss in energy[1]. These environmental effects can be readily seen in the ocean-bottom-mounted acoustic data recorded in California waters that are presented in this paper. Spatial variability in bathymetry at shallow-to-mid-depth monitoring sites can be significant over propagation distances typical of those for low (10-500 Hz) and mid (500-20 kHz) frequency calling animals. Bathymetric effects can break the azimuthal symmetry so that detection range becomes a function of bearing from the data recording package. In addition to this spatial variability, the site-specific propagation characteristics change over time due to changes in water column properties, leading to changes in the sound speed profile[1]. Solar heating during summertime increases both the sound speed and the vertical gradient in sound speed in the shallow waters where many marine mammal species vocalize. Larger near-surface gradients in sound speed refract the sound more strongly towards the ocean bottom. In contrast, surface ducts that often form and deepen during wintertime can trap sound near the surface[2]. Depending on the location and depth of the receivers, these changes in sound speed profiles can increase or decrease the detectability of calls.

Detection is a function not only of the properties of the received signal, but

also of the noise. Differences in overall level of the noise (defined in this paper as all recorded sounds excluding calls from marine mammal species) can vary by more than two orders of magnitude in energy (i.e., by more than 20 dB). In addition, the spectral character of the noise at each site can differ. For example, the variability as a function of frequency in the noise levels is significantly greater at sites with nearby shipping due to the frequency variability of radiated noise from commercial ships[3]. For a given average noise level, signal detection is more difficult in noise with frequency-varying levels than in noise that is flat (i.e., white noise).

All of these site-specific and time-varying environmental effects must be taken into account when evaluating the passive acoustic monitoring capabilities of a recording system deployed in a given location over a given period of time. They also should be taken into account when comparing the passive acoustic monitoring results collected at one location to those from another location. Therefore, it is important to estimate the site specific probability of detection (P is the true underlying detection, and \hat{P} is its estimate) for species-specific acoustic cues within a dataset. As part of this calculation, it is necessary to estimate the azimuth-dependent range over which the detections can occur for each deployed sensor. These estimates must be frequently updated as environmental properties change. One application where these site-specific and time-varying environmental effects are particularly important to take into account is in estimating the areal density of various marine mammal species using passive acoustic data.

Significant progress has been made recently in estimating marine mammal population densities using passive acoustic monitoring techniques, most notably in the Density Estimation for Cetaceans from passive Acoustic Fixed sensors (DECAF) project [4]. In addition to being of basic scientific interest, information on population densities is important in regions of human activities, or potential activities, to properly evaluate the potential impact of these activities on the environment. In the DECAF project and in other efforts, a variety of methods are used to calculate \hat{P} . It is often derived from estimating the detection function - the probability of detecting an acoustic cue as a function of distance from the receiving sensor[5]. Using distance sampling methods, it is necessary to calculate distances

to the vocalizing marine mammal, often a time-consuming task in which multiple sensors for localization are usually needed. Additionally, the detection function may need to be recalculated as environmental parameters change, particularly for low-and mid-frequency vocalizations.

When single fixed sensors are used for density estimation, the probability of detection must be estimated in part from acoustic propagation models. For marine mammals vocalizing at high frequencies (greater than 20 kHz), simple spherical spreading models are sufficient. Küsel et al.[6] demonstrated the feasibility of using spherical spreading propagation models in estimating the density of Blainville’s beaked whales (*Mesoplodon densirostris*) from passive acoustic recordings, calculating \hat{P} with acceptable uncertainty. For whales vocalizing at lower frequencies, full wave field acoustic models are necessary, and the uncertainties in the input parameters in these models can lead to large uncertainties in \hat{P} .

A growing number of single fixed acoustic sensor packages have been located in the southern California Bight since 2001. Each High-frequency Acoustic Recording Package (HARP)[7], contains a hydrophone tethered above a seafloor-mounted instrument frame, and is deployed in water depths ranging from 200 m up to about 1000 m. Analysts monitor records from these packages for a variety of marine mammal species, including humpback whales (*Megaptera novaeangliae*). Humpback songs consist of a sequence of discrete sound elements, called units, that are separated by silence[8]. Traditionally, analysts mark the presence of humpback whales within a region by indicating each hour in which a vocalization occurred. The recent development of a generalized power-law (GPL) detector for humpback vocalizations[9] has provided the ability to count nearly all human-detectable humpback units within the acoustic record. However, comparing statistics from calling activity between HARP sensors, between seasons, and across years is still constrained by the ability to estimate the spatial and temporal-varying P for these vocalizations, and the areal coverage in which these vocalizations are detected. Comparing activity between geographical locations or at the same location over time without accounting for the acoustic propagation properties of the environment

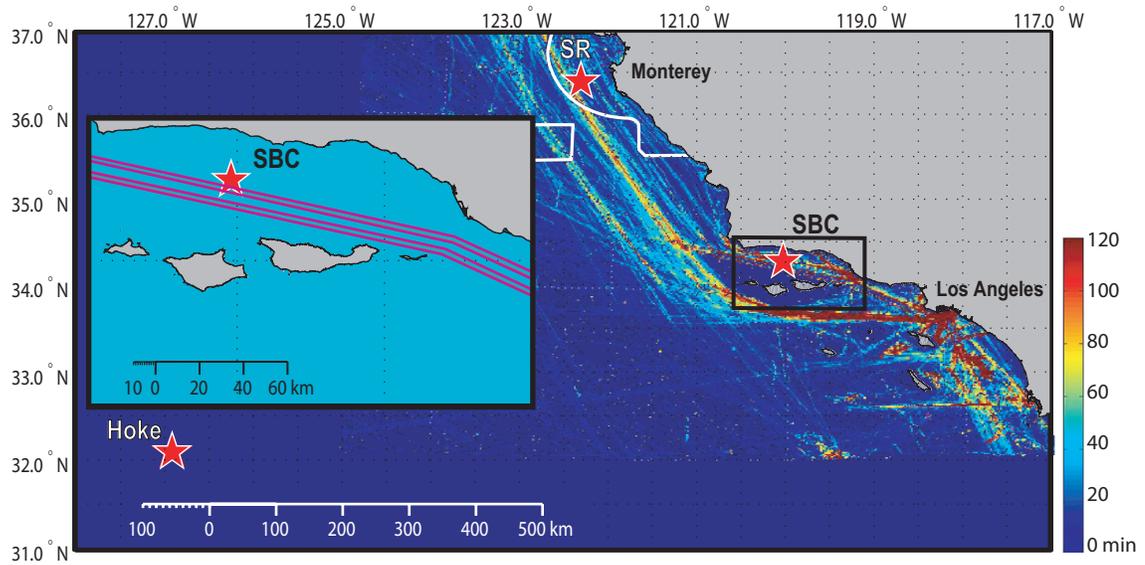


Figure 3.1: Map of coastal California showing the three HARP locations: site SBC, site SR, and site Hoke (stars). The expanded region of the Santa Barbara Channel shows northbound (upper) and southbound (lower) shipping lanes in relation to site SBC. Ship traffic from the Automatic Identification System (AIS) is shown for region north of 32°N and east of 125°W. The color scale indicates shipping densities, which represent the number of minutes a vessel spent in each grid unit of 1 arc-min x 1 arc-min size in the month of May 2010. White perimeters represent marine sanctuaries. Shipping densities provided by Chris Miller (Naval Postgraduate School).

can be extremely misleading, as the probability of detection can vary by factors of ten or more as shown in Sec. 3.3.3.

This paper focuses on three geographical areas off the coast of California, each with distinct bathymetry, ocean bottom sediment structure, sound speed profiles, and ocean noise conditions. This study highlights the variability that bathymetric and other environmental properties create when calculating \hat{P} for humpback whales. Section 3.2 gives a brief description of humpback whale activity in the north Pacific, followed by a description of bathymetric and environmental conditions at the three HARP locations off the California coast. This section also highlights the data collection and analysis effort to date for these three HARP locations. Section 3.3 outlines the acoustic modeling used to determine \hat{P} for each of the three HARP locations, with the environmental and bathymetric information described in Section 3.2.2 as inputs to the model. Estimates of P are presented for each of the three sites as well as uncertainties for these estimates. Section 3.4 explores the accuracy of the model by comparing detection statistics of 24,690 humpback units from the data collection effort to statistics generated from the model. Section 3.5 discusses the importance of various input parameters to the model, giving insight into ways to minimize uncertainty in \hat{P} . Additionally, a discussion on the potential for accurate density estimation at the three locations is given. The final section summarizes the conclusions from this work.

3.2 Passive acoustic recording of transiting humpback whales off the California coast

3.2.1 The humpback whale population off California

Humpback whales in the north Pacific Ocean exhibit a dynamic population distribution driven by seasonal migration and maternally directed site fidelity [10, 11, 12]. They typically feed during spring, summer, and fall in temperate to near polar waters along the northern rim of the Pacific, extending from southern California in the east northward to the Gulf of Alaska, and then westward to

the Kamchatka peninsula. During winter months, the majority of the population migrates to warm temperate and tropical sites for mating and birthing.

Although the International Whaling Commission only recognizes a single stock of humpback whales in the north Pacific[13], good evidence now exists for multiple populations[14, 15, 10, 12, 16, 17, 11]. Based on both DNA analysis[12] and sightings of distinctively-marked individuals[11], four relatively separate migratory populations have been identified: 1) the eastern north Pacific stock which extends from feeding grounds in coastal California, Oregon, and Washington to breeding grounds along the coast of Mexico and Central America; 2) the Mexico offshore island stock which ranges from as yet undetermined feeding grounds to offshore islands of Mexico; 3) the central north Pacific stock which ranges from feeding grounds off Alaska to breeding grounds around the Hawaiian Islands; and 4) the western north Pacific stock which extends from probable feeding grounds in the Aleutian Islands to breeding areas off Japan[18, 17, 19, 11, 20].

Within the northeastern Pacific region, where the data presented in this paper were collected, photo-ID data indicate migratory movements of humpback whales are complex; however, a high degree of structure exists. Long-term individual site fidelity to both breeding and feeding habitats for the two populations that migrate off the U.S. west coast (populations 1 and 2 in the previous paragraph) has been described[11]. The mark-recapture population estimate from 2007/2008 for California and Oregon is 2,043 and with a coefficient of variation (CV) of 0.10, this estimate has the greatest level of precision[21]. Mark-recapture data also indicate a long-term increase in the eastern north Pacific stock of 7.5% per year[21], although short-term declines have occurred during this period, perhaps due to changes in whale distribution relative to the areas sampled. Intriguing variations in seasonal calling patterns between the three data recording sites reported on in this paper have been observed[22], suggesting that the animals' behavior may differ among these three habitats.

Based on the humpback song recorded at many locations off the coast of California, six representative units were selected as inputs to the acoustic propagation model, and are shown in Fig. 3.2. These commonly recorded units

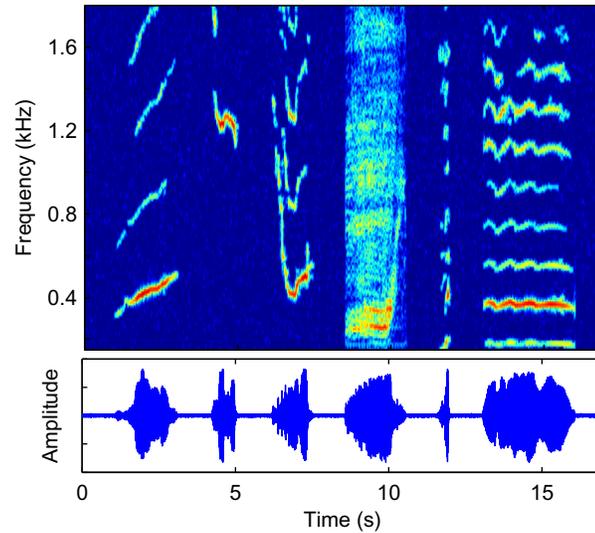


Figure 3.2: (Color online) Six representative humpback whale units used in the modeling. Units labeled 1-6 from left to right.

of humpback song represent diversity in length, frequency content, and number of harmonics - all which influence the probability of detecting the units. Vocalizations were selected from a different data source than the HARP recordings so as to capture high SNR vocalizations near to the source, minimizing attenuation and multipath effects[23].

3.2.2 HARP recording sites

Three HARP locations were selected for this study. Site SBC (34.2754° , -120.0238°) is located in the center of the Santa Barbara Channel, site SR (36.3127° , -122.3926°) is on Sur Ridge, a feature 45 km southwest of Monterey, and site Hoke (32.1036° , -126.9082°) is located on the Hoke seamount, 800 km west of Los Angeles. A map of coastal California showing the HARP locations and the Santa Barbara Channel commercial shipping lanes can be seen in Fig. 3.1. Acoustic data collected at each of these sites indicates the occurrence of humpback song over much of the fall, winter, and spring.

Bathymetry

The bathymetry for each of the three sites can be seen in the upper row of Fig. 3.3. Bathymetry information for site SR and site SBC was collected from the National Oceanographic and Atmosphere Administration (NOAA) National Geophysical Data Center U.S. Coastal Relief Model[24]. Bathymetry information for site Hoke was collected by combining data from the Monterey Bay Aquarium and Research Institute (MBARI) Atlantis cruise ID AT15L24 with data from the ETOPO1 1 Arc-minute Global Relief Model[25] for depths greater than 2000 m. At site SBC the bathymetry forms a basin with the HARP located near the center of the basin at a depth of 540 m. The walls of the basin slope up to meet the channel islands to the south and the California coastline to the north. The HARP at site SR is located at a depth of 833 m on a narrow steep ridge approximately 15 km long with a width of 3 km trending east-west. To the east the ridge slopes upwards to the continental shelf, and to the west is downward sloping to the deep ocean floor. Site Hoke is located near the shallowest point of the Hoke seamount, at a depth of 770 m. The seamount walls slope downward nearly uniformly in all directions to a depth of 4000 m.

Ocean sound speed

Sound speed profiles (SSP) were calculated from conductivity, temperature, and depth (CTD) casts in the NOAA World Ocean Database[26] that were recorded in near proximity to each of the three sites. Several hundred CTD casts were used in the analysis, covering all seasons and for years ranging from 1965 - 2008. When available, additional CTD casts were taken during the same time period as the HARP deployments[3]. Figure 3.4 shows a representative sample of the sound speed profiles collected near each of the three sites, with red indicating summer profiles (Jul-Sept.) and blue indicating winter profiles (Jan - Mar). The plots illustrate effects of warm surface waters in the summer on the sound speed profiles, especially at site SBC and site Hoke, with a deeper mixed layer occurring at site Hoke. The variation between summer and winter profiles is not as prominent at site SR, which is exposed to cooler mixed waters during the summer months than

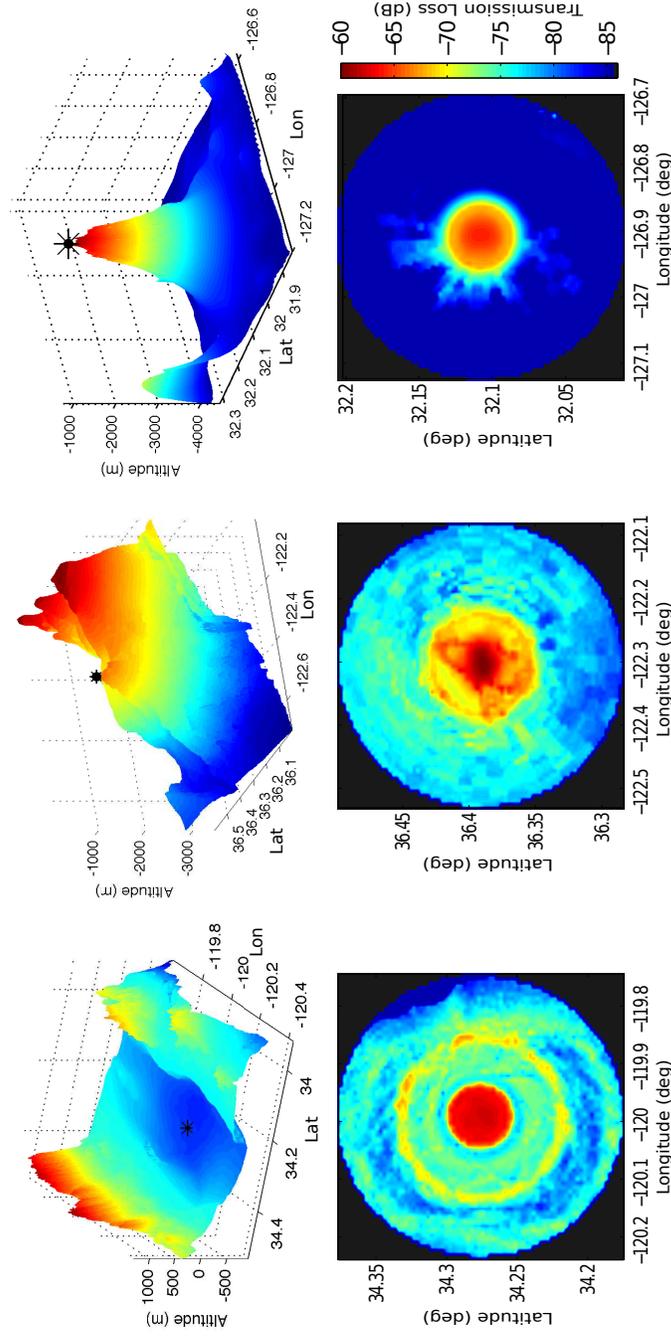


Figure 3.3: Bathymetry of site SBC, site SR, and site Hoke (left to right) with accompanying transmission loss (TL) plots. The TL plots are incoherently averaged over the 150 Hz to 1800 Hz band and plotted in dB (the color scale for these plots is given on the far right). The location of the HARP in the upper row of plots is marked with a black asterisk.

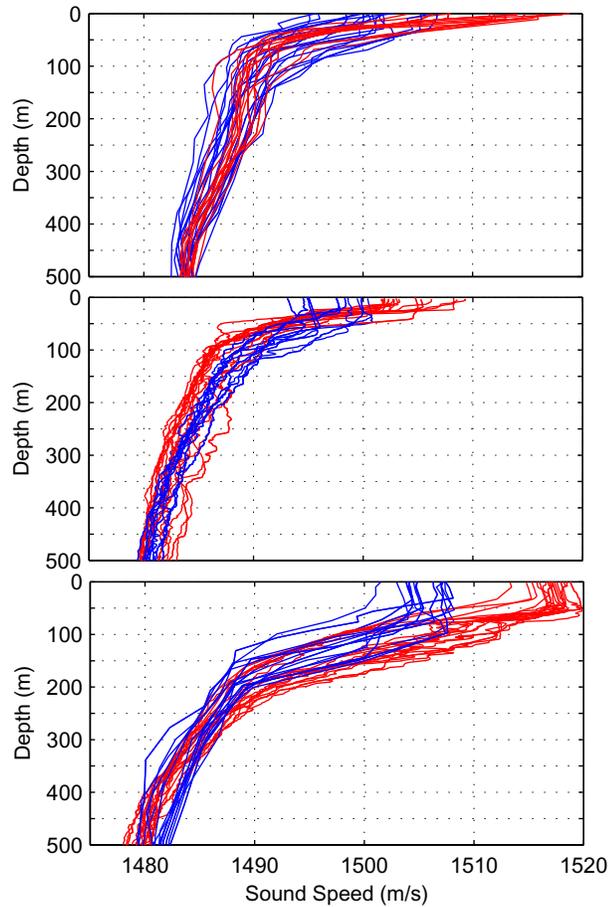


Figure 3.4: Sound speed profiles for site SBC, site SR, and site Hoke (top to bottom), for winter (blue) and summer (red) months. These data span the years 1965 to 2008.

the other two sites.

Solar heating during summertime increases both the sound speed and the vertical gradient in sound speed in the shallow waters where humpbacks vocalize. Larger near-surface gradients in sound speed refract the sound more strongly towards the ocean bottom, influencing the surface area over which sound propagates directly to the hydrophone. Additionally, surface ducts that often form and deepen during wintertime (most clearly seen in the profiles at site Hoke) can trap sound near the surface, influencing the intensity and spectral characteristics of sound propagating to the bottom-mounted hydrophone.

Ocean bottom properties

Ocean bottom characteristics are important input parameters to the acoustic propagation model. A combination of methods was used to characterize the bottom at site SBC. Bottom sound speed profile information was obtained from an experiment conducted in the area in which geoacoustic inversion methods were used to calculate the sound speed[27]. The results of this experiment combined with relationships from Hamilton[28, 29] suggest that the bottom is comprised of a sediment layer extending beyond 100-m in thickness, containing fine sand material (grain size of $\phi = 2.85$ on the Krumbein phi (ϕ) scale[30, 31]). A separate study was conducted in which sediment core samples were taken very near the location of the HARP. Information from the core suggests a sediment layer extending at least the full length of the 100 m core. The material contained within the core varied from clayey silt to silty clay, with intermediate layers of fine sand[32]. An estimated grain size of $\phi = 7.75$ was used to characterize the core. Most of the transects from the sonar study were nearer to the coastline rather than over the center of the basin, which may partly explain the variability in bottom type between the two studies. It was assumed that these two studies represent the endpoints of uncertainty of the sediment layer in the Santa Barbara channel. Therefore, in addition to these endpoint parameters, a best-estimate value of $\phi = 5.4$ extending to 100 m depth was used for the modeling effort, corresponding to a silty bottom. Below this layer was assumed to be sedimentary rock, (sound speed = 2374 m/s, density = 1.97 g/cm³, attenuation = 0.04 dB/m/kHz).

Submersible dives conducted by MBARI along with sediment cores were used to characterize the bottom at site SR. Correspondence with Gary Greene (Moss Landing Marine Laboratories) suggests the ridge itself is thought to be mostly deprived of sediment and composed of sedimentary rock. Surrounding the ridge is sediment covered seafloor - the region east of the ridge contains sediments mostly consisting of fine sand ($\phi = 3$). To the west, the sediment is characterized by clayey silt ($\phi = 7$)[33, 34]. Eleven sediment cores are available in this region to a depth of only 1 m below the ocean-sediment interface, and so the thickness of the sediment layer is unknown. The best estimate at this site assumes sedimentary rock

(sound speed = 2374 m/s, density = 1.97 g/cm³, attenuation = 0.04 dB/m/kHz), devoid of sediment out to a range of 4 km from the HARP's location. Beyond the ridge, the sedimentary rock is assumed to have a 10-m sediment cover. Ideally, the modeling would incorporate range and azimuth dependent sediment type - fine sand to the east and clayey silt to the west. However, to increase the speed of the computations, the "best" estimate used in the model assumes the sediment layer is uniform with an average grain size of $\phi = 5$. Since the exact sediment type and layer thickness are unknown, the endpoints for the bottom parameters allow the sediment structure to range from the thickest and most acoustically absorptive (sediment thickness of 50 m and clayey silt, $\phi = 7$), to least absorptive (sediment thickness of 1 m consisting of fine sand, $\phi = 3$).

For site Hoke, sediment samples were collected from the Alvin submarine in 2007 during the deployment of the HARP. Correspondence with David Clague (MBARI) suggests that the rock samples contain common alkalic basalt samples with minimal vesicles. Pictures of the HARP at its resting location on the seamount confirm that the hydrophone is surrounded by this type of rock. No sediments were observed at this site, and sediment deposit is not expected on the slopes of the seamount due to steep bathymetry and strong ocean currents. Detailed studies on the composition of nearby seamounts[35] in combination with Hamilton's[28, 29] study suggest that the density of this rock can range from just over 2.0 g/cm³ to 3.0 g/cm³, with corresponding compressional wave speeds ranging from 3.5 km/s to 6.5 km/s. A best estimate was chosen using a density of 2.58 g/cm³, compressional speed of 4.5 km/s and attenuation of 0.03 dB/m/kHz. It was assumed that the uncertainties in the bottom properties on the seamount could span the documented range of values for basalts.

Ocean noise levels

The ocean noise was characterized at each site using 75 s samples taken every hour of the HARP recordings over the 2008-2009 calendar year. No data were available from Hoke during June - August, so the noise was characterized using the remaining nine months of data. Figure 3.5 shows the noise spectrum

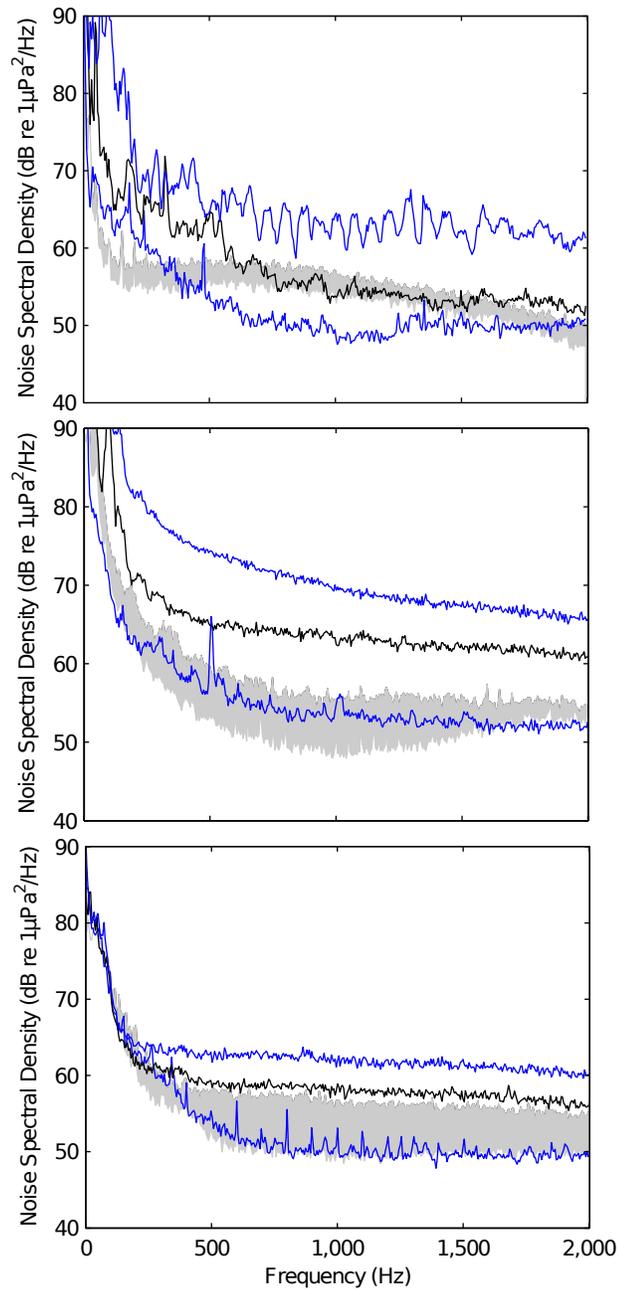


Figure 3.5: Noise spectral density levels for site SBC, site SR, and site Hoke (top to bottom). The curves indicate the 90th percentile (upper blue), 50th percentile (black), and 10th percentile (lower blue) of frequency-integrated noise levels for one year at site SBC and site SR, nine months at site Hoke. The gray shaded area indicates 10th and 90th percentile levels for wind-driven noise used for modeling.

levels at each of the three sites, with the 90th percentile, 50th percentile, and the 10th percentile noise levels illustrated. The percentile bands were determined from the integrated spectral density levels over the 150 - 1800 Hz band. The gray shaded area in each plot represents the 10th and 90th percentile range from 30 min of HARP recordings used to represent wind-driven conditions over which \hat{P} will be characterized during model simulations.

Noise levels at site SBC can change drastically over short time scales, sometimes varying between extremal values within an hour. The shallow bathymetry shields the basin from sound carried by the deep sound channel, creating at times an extremely low-noise-level environment. However, the channel is also one of the busiest shipping lanes worldwide[3], and so local shipping noise makes a significant contribution at this site. The upper plot in Fig. 3.5 illustrates the variation in the noise spectrum level with frequency, especially at high noise levels, indicating the presence of a large transiting vessel. Noise at site SR is characterized by wind-driven ocean surface processes, distant shipping, and local shipping. Sur Ridge is exposed to noise from the west traveling in the deep sound channel. Therefore, the lowest noise level times at this site are higher in level than the lowest-level times recorded at site SBC. Although not as prominent as site SBC, large ships do occasionally pass near to site SR, creating more variation across frequency than site Hoke, but less variation across frequency than site SBC. Ocean noise at the site Hoke is the least variable both spectrally and temporally among the three sites studied. The seamount is exposed to noise from all directions, and the HARP is exposed to noise traveling in the deep sound channel. However, nearby shipping noise is rare for this area of the ocean, and so the noise levels are much less variable than those found at the other two sites. HARP instrument noise can be seen in the lowest percentile curves for all three sites, where hard drive disk read/write events create narrowband contamination.

3.2.3 Probability of detection with the recorded data

Acoustic data were recorded at site SBC from Apr. 2008 to Jan. 2010, at site SR from Feb. 2008 to Jan. 2010, and at site Hoke from Sept. 2008 to

June 2009. The GPL detector was used to mark the start-time and end-time of nearly every human identifiable unit in the records, resulting in approximately 2,300,000 marked units. The GPL detector is a transient signal detector based on Nutall’s power-law processor[36], which is a near-optimal detector for identifying signals with unknown location, structure, extent, and arbitrary strength. The GPL detector is built on the theory of the power-law processor with modifications necessary to account for drastically changing ocean noise environments, including non-stationary and colored noise generated from shipping. The GPL detector has an average false alarm rate of approximately 5% at the detector threshold used in this research and for the datasets at hand. Therefore, trained human analysts eliminated the false detections manually, using a graphical user interface (GUI), which is part of the GPL software. The GUI allows the analysts to accept or reject large batches of detections at a time, allowing for much quicker data analysis time when compared to reviewing each detection individually. This pruning effort required approximately two weeks (112 hours) of trained human analyst time for the total 54 months of recorded data. Statistics obtained from the data analysis effort were used to verify the accuracy of the probability of detection modeling effort, discussed in Sec. 3.3.

3.3 Probability of detection - modeling

The accuracy of estimating P relies on characterizing the range, azimuth, and depth dependent detection function in accordance with the detector used. In this paper, the variation in depth of calling animals is not fully accounted for in the modeling, so that the detection function, $g(r, \theta)$, is taken as a function of range, r , and azimuth, θ , only. The detection function measures the probability of detection from the hydrophone out to the maximum radial distance (w) in which a detection is still possible, over all azimuths. The azimuthal dependence is added to the standard equation to emphasize the complexity caused by bathymetry. The probability of detection within a given area is then calculated by

$$\hat{P} = \int_0^w \int_0^{2\pi} g(r, \theta) \rho(r, \theta) r dr d\theta \quad (3.1)$$

where $\rho(r, \theta)$ represents the probability density function (PDF) of whale calling locations in the horizontal plane[5]. Throughout this study, a homogeneous random distribution of animals over the whole area of detection, πw^2 , is assumed, and so $\rho(r, \theta) = (1/\pi w^2)$. One way of calculating the detection function is to use a localization method to tabulate distances to whale vocalizations within an acoustic record. An appropriate parametric model for $g(r, \theta)$ is assumed, and $g(r, \theta)$ is estimated based on a PDF of detected distances[37]. This method is often preferred because variables that influence the detection function, such as source level and acoustic propagation properties, can remain unknown. From the single hydrophone data used in this analysis, tabulating distances to vocalizing animals using localization methods is not possible. Instead, a 2D acoustic propagation model is used to estimate P within a geographic area. This method requires knowledge about the acoustic environment and the source, and in general is more demanding and perhaps less accurate than methods in which distances to animals can be estimated. However, this method does have some advantages over distance estimation methods. Mainly, a parametric model is not assumed for $g(r, \theta)$, meaning the detection function can both increase and decrease with range. This variation in range is often overlooked using distance methods because a high localization accuracy is necessary, and many distances need to be calculated to make these variations statistically significant. Additionally, the use of single fixed sensors for acoustic monitoring can reduce the complexity and cost of the monitoring data acquisition system when compared to localizing systems.

Recent research results have been published on the successful characterization of \hat{P} for detecting marine mammals from single fixed omnidirectional sensors, some of which use acoustic models for calculating the detection function[6, 37, 38]. Most of these studies have involved higher frequency odontocete calls, such as those from beaked whales (family *Ziphiidae*), although some studies have included baleen whales. For higher frequency calls typical of odontocetes, the high absorption of sound with range limits uncertainties associated with

environmental parameters, and transmission loss (TL) is usually confined to spherical spreading plus absorption. Therefore, the variables that influence \hat{P} the most tend to be associated with the source, such as whale source level (SL), grouping, location, depth, and orientation due to the directionality of high frequency calls. These types of variations often can be modeled as independent random variables with an assumed distribution, characterized by Monte Carlo simulation. Apart from source level, these variables play a minimal role for acoustic censusing of humpback whales. Au et al. show that humpback whales tend to produce omni-directional sound over a very limited range in depth[39]. However, due to the lower frequency nature of the humpback vocalizations, variations in sound propagation due to environmental properties become large. Uncertainties in these variations, such as bottom type, sediment depth, water column sound speed, and bathymetry can lead to uncertainties in \hat{P} that overwhelm uncertainties attributed to other processes. To complicate the issue, the pressure field received at the hydrophone depends on these environmental parameters non-linearly.

To understand the influence of individual variables on \hat{P} , these variables are grouped into environmental variables and source variables, and an analysis is conducted on each group separately. The main focus is to characterize the influence of the environment. To do so, the source variable properties remain unchanged, assuming a random homogeneous, horizontal distribution of animals, a fixed source depth of 20 m, and a fixed omnidirectional source level of 160 dB rms re 1 μ Pa @ 1 m for each humpback unit. The dependence of \hat{P} on environmental variables is explored in two stages. In the first stage, variation is limited to a single input parameter, while holding others fixed at best-estimate values. In the second stage, combinations of variables that lead to extremal values of \hat{P} are characterized. After characterizing the influence of environmental variables, a limited analysis of uncertainties associated with variation originating from the source properties is carried out by holding environmental variables fixed at best-estimate values.

3.3.1 Approach - numerical modeling for environmental effects

This section describes the method for estimating the probability of detecting humpback units using a single fixed omni-directional sensor. This method is in many ways similar to that described by Küsel et al[6] for Blainville’s beaked whales, but with important differences needed to account for the propagation properties of lower frequency vocalizations. To accommodate the complex transmission of lower frequency calls, a full wave field acoustic propagation model is used. Additionally, unlike beaked whale clicks which have distinct and mostly uniform characteristics, humpback units cover a wide range of frequencies and time scales. As such, the probability of detecting individual units varies significantly - this variation comes about both from bias in the GPL detector, as well as the frequency dependent propagation characteristics of the acoustic environment. Since one important application of estimating \hat{P} is density estimation, establishing an average vocalization rate, or cue rate is required. Because humpback song can be highly variable, selecting a particular type of unit, or even a subset of units to use as acoustic cues would lead to inaccurate density estimates as the song changes. Additionally, a classification system would be needed to single out these units from an acoustic record. Counting all units over a wide frequency range overcomes some of the challenges associated with the variation in humpback song, but adds additional challenges to characterizing \hat{P} for all unit types.

The humpback units shown in Fig. 3.2 were used to simulate calls originating at various locations within a 20-km radius centered on the hydrophone. For this purpose, the Range-dependent Acoustic Model (RAM)[40] was used to simulate the call propagation from source to receiver, in amplitude and phase as a function of frequency. In previous studies[6], the passive sonar equation[41] was used to estimate the acoustic pressure squared level at the receiver. However, this method does not account for phase distortion of the signal, necessary for including propagation effects such as frequency-dependent dispersion. In addition, modeling both the acoustic field amplitude and phase as a function of frequency, which then can be inverse-FFT’d and added to a realization of noise taken from the measured

data, allows the synthesized calls to be processed in an identical way to that of the recorded data.

The RAM model is used to calculate the complex pressure field at 0.2 Hz spacing from 150 Hz to 1800 Hz. An inverse FFT of this complex pressure field results in a simulated time series with duration 5 s for data sampled at 10 kHz. This window encompasses the longest-duration humpback unit used in this study, with multipath distortion. The convolution of this pressure time series with the original unit yields the simulated unit as received by the sensor. A sample result is shown in Fig. 3.6. Once the waveform of a unit transmitted from a particular point on the grid is computed, a randomly-chosen HARP-specific noise sample (discussed in Sec. 3.2.2) is added and the resulting waveform is passed to the GPL detector. The output of the GPL detector determines whether this unit is detected, and updates the probability of detection for that location on the grid. Calls are simulated over each location on the geographic grid with 20 arc-second spacing. Based on these results, the truncation distance (w) can be chosen, allowing for the calculation of \hat{P} for the area defined by πw^2 . This process is repeated with a range of noise samples to produce a curve that links \hat{P} to the monitored noise level as shown in Fig. 3.9, and discussed further in Sec. 3.3. As previously outlined, these Monte Carlo simulations are also repeated allowing environmental and source inputs to vary so as to characterize uncertainty in \hat{P} .

For purposes of cetacean density estimation, it is sometimes necessary to further restrict the process of detection with an added received SNR constraint. The purpose of this constraint is threefold: a) to truncate detections to distances that result in stable determination of \hat{P} , b) minimize bias in the detector for varying unit types as outlined in Table II in Helble et al[9], and c) limit detections to SNRs easily detectable by human analysts used to verify the output of the detector. Additionally, comparing the estimated SNR in both the simulations and the real datasets allows the accuracy of the model to be assessed. The SNR is defined as:

$$SNR = 10 \log_{10} \frac{\langle p_s^2 \rangle}{\langle p_n^2 \rangle} \quad (3.2)$$

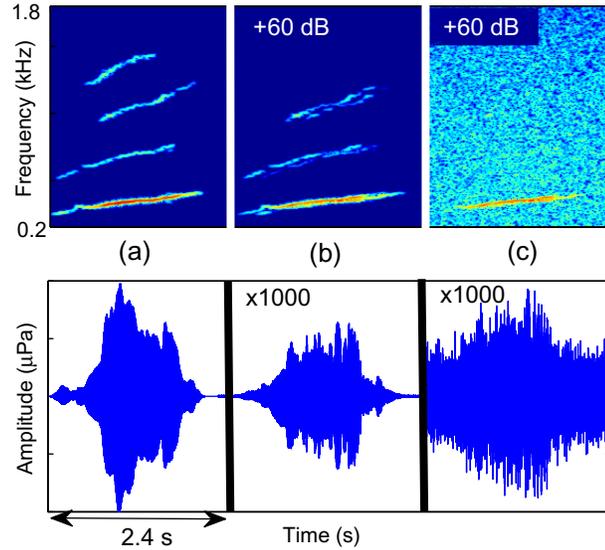


Figure 3.6: (Color online) (a) Measured humpback whale source signal rescaled to a source level of 160 dB re 1 μPa @ 1 m, (b) simulated received signal from a 20-m-deep source to a 540-m-deep receiver at 5 km range in the Santa Barbara Channel, with no background noise added, (c) simulated received signal as in (b) but with low-level background noise measured at site SBC added. The upper row of figures are spectrograms over the 0.20 to 1.8 kHz band and with 2.4 sec duration, and the lower row are the corresponding time series over the same time period as the spectrograms. The received signal and signal-plus-noise time series amplitudes in the 2nd and 3rd columns have been multiplied by a factor of 1000 (equal to adding 60 dB to the corresponding spectrograms) so that these received signals are on the same amplitude scale as the source signal in the first column. This example results in a detection with recorded $SNR_{est} = 2.54$ dB.

where

$$\langle p_{s,n}^2 \rangle \equiv \frac{1}{T} \int_0^T p_{s,n}^2(t) dt$$

and where p represents the recorded pressure of the time series, bandpass filtered between 150 Hz and 1800 Hz, and T is the duration of the time series under consideration.

The GPL detection software automatically estimates the SNR of each detected unit in the recorded data. With real data, the SNR defined in Eq. (3.2) must be estimated because the recorded pressure of the signal and noise can never be separated completely. This automated estimate of SNR, SNR_{est} , is assisted by the GPL detector, which is designed to identify narrowband features in the presence of broadband noise. Individual frequencies in the spectrogram are identified that correspond to the narrowband humpback signal. These frequency bins also contain noise, and the energy contributed by noise is estimated, by measuring the energy levels in the corresponding bands over a 1-s time period before and after the occurrence of the unit, and then subtracted. The resulting estimates of energy from the signal frequencies are averaged over the duration of the detected unit, and compared to energy in the spectrogram adjacent to the unit within the 150 to 1800 Hz band, resulting in SNR_{est} . Although the exact SNR of simulated data as defined in Eq. (3.2) could be calculated, SNR is estimated in the same way for both real and simulated data, so that calculations of \hat{P} from simulated data that use an SNR constraint will apply for the analysis of real data.

Choosing an $SNR_{est} = -1$ dB cutoff helps to minimize the bias in the detector over unit type in addition to limiting incoming detections to levels easily verifiable by human operators. The criteria for selecting detections corresponding to those propagation distances that result in a stable determination of \hat{P} are site specific. For simplicity the same threshold value of -1 dB SNR_{est} is employed throughout, although adjusting this value based on a number of factors is appropriate, as discussed in Sec. 3.5.

The modeling method outlined in this section is different than most published acoustic-based methods used to derive \hat{P} , in which the transmission loss, noise level, and SNR performance of the detector are characterized separately.

Using the method proposed in this paper, these quantities are interlinked owing to the site-specific environmental characteristics. Characterizing the detection process jointly gives a more realistic solution, at the cost of substantially greater computational effort.

3.3.2 CRAM

The C-program version of the Range-dependent Acoustic Model (CRAM) was developed as a general-purpose Nx2D, full wave field acoustic propagation model. At its core are the self-starter and range-marching algorithm of the RAM 2D parabolic equation model, originally developed and implemented in Fortran by Collins[40]. The parabolic equation (PE) model is an approximate solution to the full elliptic wave equation, in which the solution is reduced in computational complexity by assuming the outgoing acoustic energy dominates the backscattered energy. In CRAM, setup of the Nx2D propagation problem is handled automatically for desired receiver output grids in geographic coordinates. The assumptions inherent in the Nx2D approximation, versus full 3D propagation modeling, are that horizontal refraction and out-of-plane bathymetric scattering can be neglected in the environment of interest, so that adjacent radials can be computed independently without coupling. The set of independent radials, and the range-marching within each radial, are selected such that the complex pressure for each source-receiver pair is phase-exact in the along-range direction, and approximated in the much less sensitive cross-range direction by a controllable amount. This preservation of spatial coherence allows for beamforming and other post-processing operations which require high fidelity of the complex pressure output.

The RAM Fortran code was ported to the C programming language and refactored for efficiency on modern processor architectures, which have very different relative costs of computation and memory access than older processors. As much of the 2D PE grid setup as possible is reused over multiple frequencies, allowing for more rapid computation of broadband and time-domain pressure responses. To leverage the multiprocessor capability of modern computers, the

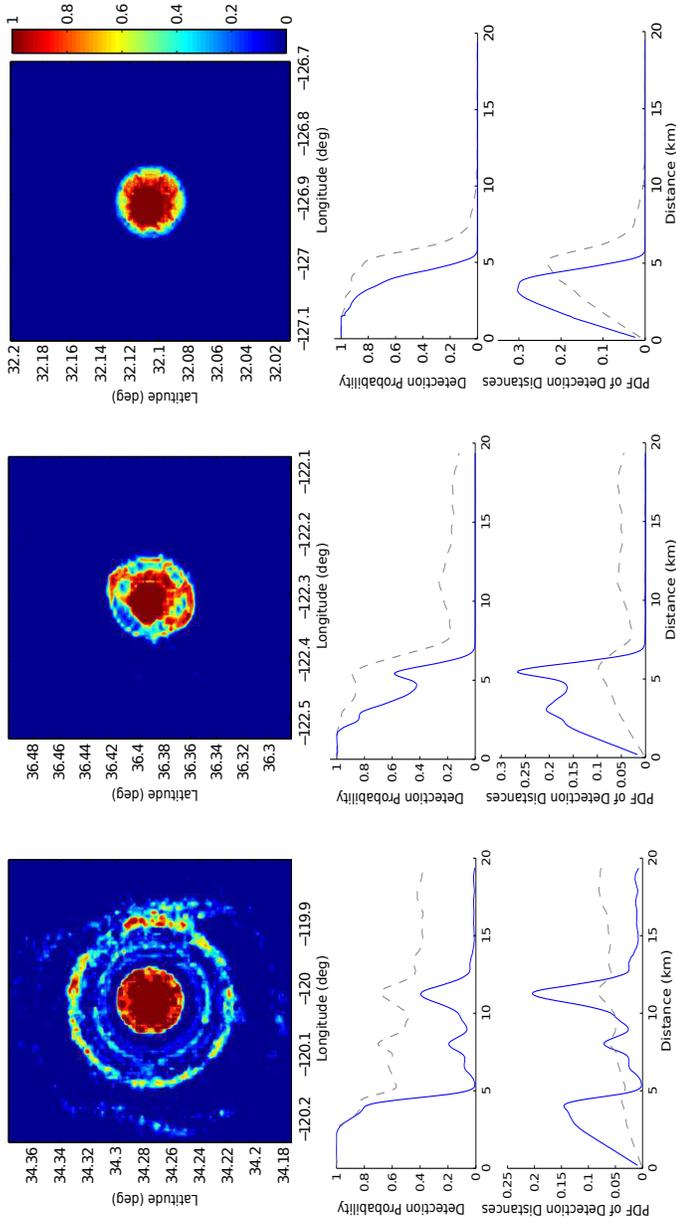


Figure 3.7: Probability of detecting a call based on the geographical position of a humpback whale in relation to the hydrophone during periods dominated by wind-driven noise at site SBC (upper left), site SR (upper center), and site Hoke (upper right), averaged over unit type. Assuming a maximum detection distance of $w = 20$ km, average $\hat{P} = 0.1080$ for site SBC, $\hat{P} = 0.0874$ for site SR, and $\hat{P} = 0.0551$ for site Hoke. The latitude and longitude axes in the uppermost row of plots is in decimal degrees. The detection probability functions for the three sites, resulting from averaging over azimuth, are shown in the middle row and the corresponding PDFs of detected distances are shown in the lower row. Solid (dashed) lines indicate functions with (without) the additional -1 dB SNR_{est} threshold applied at the output of GPL detector.

program is parallelized over the N independent radials as well as more limited parallelization over frequency and Pade coefficient index, without causing changes to the output.

Environmental inputs are interpolated from a variety of 4D (3D space plus time) ocean models and bathymetry databases as they are needed in the calculations. The model can use standard geoacoustic profiles that are range as well as depth dependent, but its ability to take a scalar mean grain size (ϕ), available from sediment cores or even from the sediment type read off a navigation chart, and convert this information into geoacoustic profiles using Hamilton's relations[28, 29] greatly facilitates the problem setup. Additionally, the model can output a variety of file formats including Keyhole Markup Language (KML) format that can be imported directly into popular viewers.

3.3.3 Results

The resulting transmission loss from the modeling effort as a function of range and azimuth for each site is shown in the lower row of plots in Fig. 3.3, using the best-estimate environmental parameters as outlined in Sec. 3.2.2. These plots were created by placing a horizontal grid of virtual humpback sources at 20-m water depth covering the area out to a 20-km radius from the HARP. The TL is calculated as a function of frequency from the sources to the receiver (HARP) at ranges from zero (source directly over the HARP) out to 20 km, at all azimuths. To reduce computation time, the principle of reciprocity is used - a single source is placed at the HARP sensor position and the acoustic field is propagated out to each of the grid points (receivers) at 20 m depth. The plotted TL in dB is the result of incoherently averaging over frequency from 150 Hz to 1800 Hz, covering the humpback whale call frequency band. The HARP latitude/longitude position is located in the center of each plot. As these TL plots illustrate, the propagation characteristics at each site are strikingly different. Whereas the TL is comparatively low only in a small-radius circle about the HARP location at site Hoke (the small red circle in the lower right-most plot in Fig. 3.3), the sound field at site SBC refocuses at greater range due to interaction with the bathymetry (the

outer yellow circular ring surrounding the red circle in the lower left-most plot). This yellow ring indicates that sources at this range can be detected more easily by the HARP than sources at somewhat shorter range. The bathymetry at each site also breaks the azimuthal symmetry so that detection range is a function of bearing from the HARP package.

Values of \hat{P} in wind-driven noise

The simulated probability of detecting units 1-6 averaged over unit type and in 30 min of wind-driven noise, randomly selected from the HARP data, for sites SBC, SR, and Hoke are shown in Fig. 3.7. These results use a sound speed profile taken in the month of October with the remaining environmental variables set to best-estimate values as described in Sec. 3.2.2. The plots in the uppermost row show $\hat{P}(r, \theta)$, the plots in the middle row show the detection function $g(r)$, averaged over azimuth, and the plots in the lower row show the area-weighted PDF that results. The values of \hat{P} are computed directly from the plots in the upper row; the remaining rows are provided for comparison with other distance sampling methods. The solid lines in the plots from the middle and lower rows indicate values obtained using the -1 dB SNR threshold applied to the GPL output, while the dashed lines illustrate the results in the absence of the -1 dB SNR threshold. The dashed lines clearly show that a substantial fraction of the low-SNR detections occur at distances greater than 20 km for site SBC. Using the SNR threshold, detections for all three sites are limited to $w = 20$ km, resulting in $\hat{P} = 0.1080$ for site SBC, $\hat{P} = 0.0874$ for site SR, and $\hat{P} = 0.0551$ for site Hoke. (For comparison purposes, w is set to the same range for all three sites, but in practice w should be calculated as outlined in Sec. 3.3.1.) Without the SNR constraint, the probability of detecting humpback units at site SBC can be greater than ten times the probability at site Hoke. The highly structured form of $\hat{P}(r, \theta)$ for both sites SBC and SR, due to the influence of bathymetric features, indicates the necessity of a fully 2-D simulation of detection. The detailed structure at site SBC also suggests that estimation of the detection function based on localized distances to vocalizing animals as in Marques et al[37] would require an enormous sample size and accurate distance

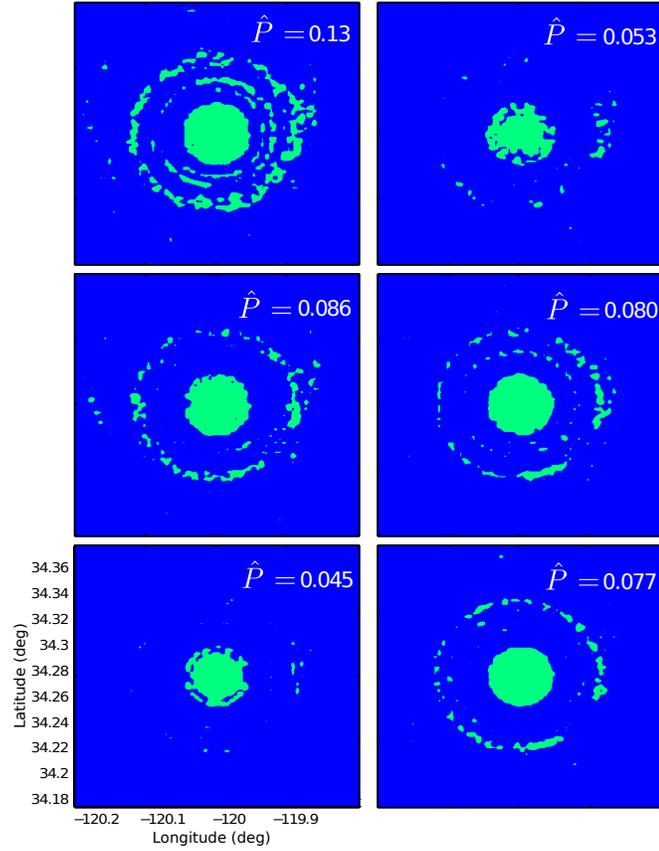


Figure 3.8: Geographical locations of detected calls (green dots mark the source locations where detections occur) and associated probability of detection (\hat{P} , listed in the upper right corner of each plot) for calls 1-6 (left to right, starting at the top row) in a 20 km radial distance from the hydrophone for a single realization of low wind-driven noise at site SBC. The latitude and longitude scales on each of the six plots are the same as in the upper lefthand plot of Fig. 3.7.

determination, particularly when an SNR threshold is not applied. Note that during a high noise period, such as when a ship was located within the Santa Barbara channel, detections at site SBC are confined to the inner red circular patch (4 km radial distance from HARP). This example emphasizes the necessity of continuous monitoring of noise to calculate \hat{P} as indicated by Fig. 3.9 and discussed in greater detail in this paper. Figure 3.8 illustrates an example of the variability in the detection across unit type during a sample of wind-driven noise conditions at site SBC. Units 2 and 5 from Fig. 3.2 are the ones most difficult to detect owing to high frequency content and brevity, respectively. The decrease in detection of unit 2 is mainly a consequence of frequency selective attenuation and propagation multipath, and does not result from an intrinsic aspect of the GPL detector. Since the detected sound interacts less with the bottom and travels shorter distances for sites SR and Hoke, the variability in detection across humpback units is less. For site SR, Unit 1 was most detectible with a $\hat{P} = 0.1136$, while Unit 5 was least detectible with a $\hat{P} = 0.0622$. The remaining calls had nearly equal probability of detection (mean = 0.0872). Similarly for site Hoke, Unit 1 was most detectible with a $\hat{P} = 0.0651$, while Unit 5 was least detectible with a $\hat{P} = 0.0478$. The remaining calls had nearly equal probability of detection (mean = 0.0548).

Environmental input variability on \hat{P} in wind-driven noise

The acoustic pressure field calculated by CRAM was recomputed over the full range of environmental input uncertainties at each site to characterize the influence of bathymetry, bottom sediment structure, and SSP on estimates of the probability of detection. Table 3.1 illustrates the influence of environmental variables on \hat{P} for the 30-min sample of wind-driven noise at each site. The first row for each site gives extremal examples of the monthly variation in SSP. That is, \hat{P} was recomputed using all SSPs occurring in the month of October (Sec. 3.2.2). The values of \hat{P} that led to the largest and smallest values of \hat{P} are shown in the table, along with a best-estimate value, which was chosen from a typical SSP for the month. All other input variables were fixed at best-estimate values. If the SSP is known within the month of the estimate, the simulation results suggest

that changes in the SSP can vary \hat{P} by over 20% for site SBC, and over 10% for sites SR and Hoke. The second row of the table shows the extremal values of \hat{P} if the SSP is chosen over a full year's worth of profiles at each site. For site Hoke and SR, the additional uncertainty is not much larger. However, estimates of \hat{P} at site SBC are more sensitive to the SSP, and the ability to detect humpback units can change between winter and summer by over 300%. The third row in the table gives extremal and best-estimate values over the full range of uncertainty in the bottom structure (sediment type and thickness) for each of the three sites, as outlined in Sec. 3.2.2. Even though site SBC in some ways had the least amount of uncertainty in bottom structure, the difference between the two extremals in sediment type (clayey silt to fine sand), had a large impact on \hat{P} , resulting in variations in \hat{P} greater than 300%. The reason for the variability is twofold, the absorption, transmission, and reflection characteristics over these sediment types change significantly over the frequency range of interest, and also because the shallow trough-shaped basin causes the sound field to interact strongly with the bottom. The variation in sediment properties over the range of possible values at site SR was by far the largest source of uncertainty at this location, causing values of \hat{P} to vary by over 100%. In contrast, even though little information was known about the igneous rock at Hoke, the variation over possible range of values resulted in essentially no differences in estimates of the probability of detection. Owing to the large downward slope of the seamount away from HARP location, the recorded sound interacts very little with the bottom. Additionally, the acoustic impedance mismatch is so high between igneous rock and the water column that the reflection characteristics are very similar over the possible range of igneous rock properties. The last row in the table for each of the three sites indicates combinations of sediment and SSPs (for the month of October) that led to extremal values of \hat{P} . Simulations as well as physical reasoning indicate that SSPs that have summer attributes (strong downward-refracting near-surface conditions) combined with the smallest grain sizes and thickest sediment layers yield the smallest values of detection. Conversely, SSPs that have winter attributes paired with the largest grain size and thinnest sediment layer produce the maximum detection values.

Table 3.1: Best-estimate and extremal predictions for \hat{P} for wind-driven noise conditions, given the uncertainty in input parameters of SSP and sediment structure for each site, as outlined in Sec. 3.2.2. Each estimate of P assumes the remaining variables are fixed at best-estimate values. The \hat{P} values assume a detection radius of $w = 20$ km from the instrument center.

		Min Extremal	Best Estimate	Max Extremal
SBC	Monthly variation in SSP	0.0823	0.1080	0.1150
	Yearly variation in SSP	0.0823	0.1080	0.2965
	Sediment variation	0.0458	0.1080	0.1887
	Monthly SSP variation + sediment variation	0.0414	0.1080	0.1892
SR	Monthly variation in SSP	0.0778	0.0874	0.0901
	Yearly variation in SSP	0.0778	0.0874	0.0914
	Sediment variation	0.0599	0.0874	0.1010
	Monthly SSP variation + sediment variation	0.0520	0.0874	0.1031
Hoke	Monthly variation in SSP	0.0482	0.0551	0.0565
	Yearly variation in SSP	0.0460	0.0551	0.0565
	Sediment variation	0.0551	0.0551	0.0551
	Monthly SSP variation + sediment variation	0.0482	0.0551	0.0565

Variations over bottom type at site Hoke combined with monthly variation in SSP did not produce measurable differences with those from holding the bottom type fixed. In summary, the environmental variables that create the most uncertainty in \hat{P} are site specific. Guided by physical intuition, one can use an acoustic model with historical data as input for a given location to identify the main sources of uncertainty, and can quantify that uncertainty, in estimating the probability of detection.

An extensive study was not conducted to measure the influence of variation in source properties (i.e, source depth, source level, deviation of horizontal source distribution from homogeneous) on \hat{P} . However, simulations using 1000 units were conducted, allowing the source level to vary with a Gaussian distribution (mean = 160 dB re 1 μ Pa @ 1 m, standard deviation = 2 dB). This amount of variation covers the full range of call levels reported in Au et al [39], although the true distribution of call levels cannot be determined with the limited data available in this paper. For site SR, allowing the source level to vary holding environmental parameters

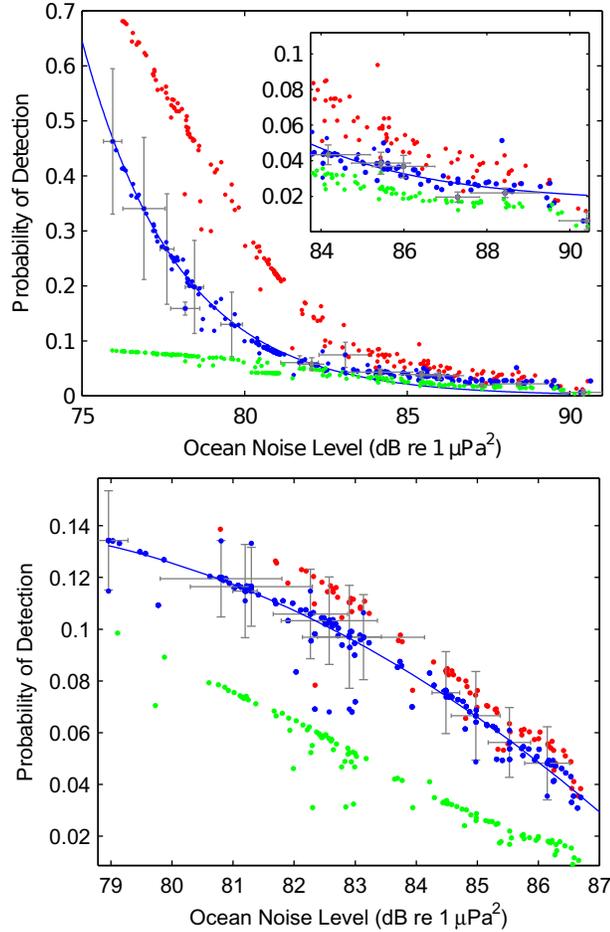


Figure 3.9: Site SBC (upper) and site SR (lower) \hat{P} versus noise level for the sediment property and SSP pairing that maximizes \hat{P} (red), the sediment/SSP pairing that minimizes \hat{P} (green), and the best-estimate environmental parameters (blue). Vertical error bars indicate the standard deviation among call unit types, and horizontal error bars indicate the standard deviation of the noise measurement. The noise was estimated by integrating the spectral density over the 150 Hz to 1800 Hz frequency bands using twelve samples of noise within a 75 s period.

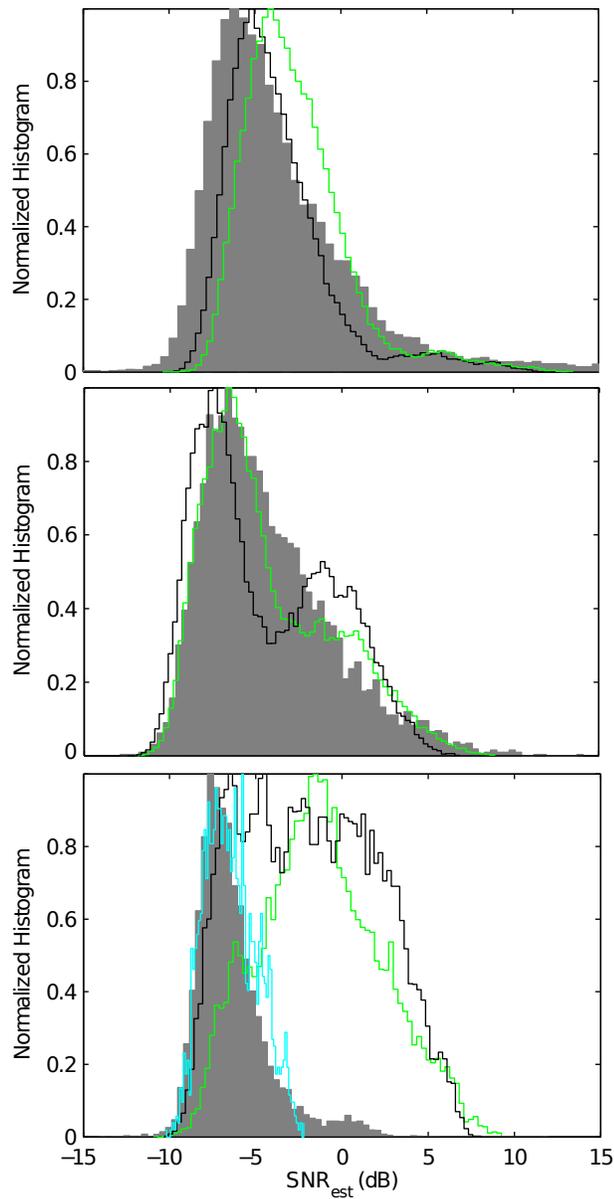


Figure 3.10: Shaded gray indicates normalized histogram of received SNR estimates (SNR_{est}) for humpback units at site SBC, site SR, and site Hoke (top to bottom). Model best environmental estimates (black line), and model upper environmental estimates (green line). The cyan line indicates best estimate results with 4 km radial calling "exclusion zone" at site Hoke.

fixed at best-estimate values resulted in a coefficient of variation (CV, equal to the ratio of the standard deviation to the mean) of 25.3% about the best-estimate mean of $\hat{P} = 0.0874$. Similarly, allowing the source to vary in depth between 10 m and 30 m resulted in even less variation. Both factors, in any combination, result in significantly less variability than that due to the uncertainty of the bottom type at site SR.

Influence of ocean noise on \hat{P}

Ocean noise has a large influence on \hat{P} . The noise in the band of humpback vocalizations can vary appreciably in both level and structure. Since detection is a function of both the noise level (SNR) and the variance of the noise level, a noise model that does not account for long-term changes in noise level or short-term variance in noise level across time and frequency is not sufficient for predicting the performance of the detector, and ultimately \hat{P} . Ocean noise was collected from each of the HARP datasets over a wide range of conditions and used as input to the calculation of \hat{P} . Figure 3.9 shows the relationship of \hat{P} versus noise level for sites SBC and SR. The blue dots represent this relationship of \hat{P} versus noise level for best-estimate environmental conditions averaged over all call types, while the green and red dots represent the modeling results using extremal environmental conditions (re Sec. 3.2.2), averaged over all call types. The noise was estimated by integrating the spectral density over the 150 Hz to 1800 Hz frequency bands using twelve samples of noise within a 75 s period. An average noise value was then assigned to each 75 s sample of noise used during the simulation. The horizontal error bars represent the standard deviation of the twelve noise measurements. The vertical error bars represent the standard deviation in the probability of detection across unit type. As the noise level decreases, the units can be detected at farther range, and so can incur greater frequency-dependent attenuation and interaction with the ocean bottom, increasing the variability in detection over unit type. As the noise level increases, the variance of the noise also tends to increase, so that an average of noise level over a 75 s time period becomes less sufficient in characterizing detection performance. A curve composed

of two separate exponentials was matched to the blue data points for site SBC. At high noise levels (detail in figure inset), the behavior for \hat{P} is dominated by direct path propagation, whereas during low noise conditions, interaction with the bottom and the increase in the area monitored with the square of the increase in detection range tend to dominate the shape of the curve. For site SR, a quadratic polynomial was used to fit the blue dots.

3.4 Model/Data Comparison

Given the non-overlapping coverage and omni-directional nature of the HARP sensors, it was not possible to calculate the detection function using source localization methods. Therefore, this approach's results cannot be compared to the results in this paper. For the data processing discussed in Sec. 3.2.3, using data recorded in the month of October, an estimate of noise level was made in addition to recording the SNR_{est} of each detected humpback unit. The shaded region in Fig. 3.10 shows the normalized histogram of recorded humpback units as a function of received SNR_{est} over a 2 dB range of received noise levels. These simulated results (black and green curves) used SSPs taken during the month of October, and 100,000 simulated calls random homogeneously distributed around the HARP. As with the other simulations, the source level of all units was assumed to be 160 dB re 1 μ Pa @ 1 m, at a depth of 20 meters. Site SBC's normalized histogram of the data processing results was created using 8944 calls over a measured noise range of 78 to 80 dB re 1 μ Pa, site SR's data histogram was created using 6559 calls over a noise range of 82 to 84 dB re 1 μ Pa, and site Hoke's data histogram was created using 9187 calls over a noise range of 82 to 84 dB re 1 μ Pa (all noise values integrated from 150 to 1800 Hz). The simulated histograms were generated using the same 2 dB noise ranges. The SNR and noise levels for each detected unit were estimated using the method described in Sec. 3.3.1. The agreement of the simulated and measured histograms for sites SBC and SR suggest that the input best-estimate model parameters and the assumptions about the source properties are quite reasonable. For site SBC, the 5 to 15 dB SNR_{est} range on

the horizontal axis of the plot represents calls originating near to the receiver, whose arrival structure is dominated by the direct path. The agreement of the predicted values and measured values in this range suggest that the average unit SL is very close to 160 dB re 1 μ Pa @ 1 m, which verifies the mean source level estimated by Au et al[39]. If the animal locations follow a homogeneous random distribution in this area, the results suggest that the true environmental input parameters are somewhere between best-estimate values and those that maximize \hat{P} . Because the simulations considered calls only out to a 20 km distance, the left-hand portion of the histograms do not agree at site SBC. This discrepancy verifies that without a received SNR cutoff and/or higher detection threshold, units are detected at distances greater than 20 km. The shape of each of the histograms at low SNR_{est} (left-hand side of the plots) is shaped by the performance of the GPL detector. The performance of the detector drops sharply as the SNR of received calls drops below -7 dB SNR. As with site SBC, if the calls at site SR are indeed homogeneously distributed, the results suggest that the environmental input parameters set between best-estimate values and those yielding maximum \hat{P} values would best match the measured SNR distribution. In contrast, the observed distribution of received call SNRs at Hoke does not fall within the bounds predicted by the model. This observed distribution can arise from one of two situations: either the calls are not homogeneously distributed around the HARP, or the calls are homogeneously distributed but detections can occur at much greater distances than the model predicts. It is possible that at this site, the acoustic energy created by shallow sources somehow couples into the deep sound channel to allow for very long range detection by the HARP approximately at the sound channel axis depth. If the calls are originating only within 20 km of the HARP, they must occur at distances greater than 4 km from the HARP. One possibility that would lead to a 4 km "exclusion zone" is that the humpback whales are transiting along a narrow migration corridor with a 4 km closest point of approach. Alternatively, perhaps they are avoiding the shallowest portion of the seamount for some reason. The cyan curve in the lowermost plot of Fig. 3.10 is the result of running the model with calls homogeneously distributed in the area, but excluded within 4 km of the

shallowest portion of the seamount.

3.5 Discussion

The uncertainties in \hat{P} from single fixed sensors due to unknowns in environmental parameters such as sound speed profile, bottom sediment structure, and ocean noise can be large for animal calls at all frequencies. For the mid to low frequencies typical of vocalizations from mysticete whales, these uncertainties generally outweigh the uncertainties associated with the source, such as whale calling depth and source level. For higher frequency vocalizations typical of odontocete whales, the uncertainties associated with environmental parameters other than ocean noise are minimized because the sound attenuates to undetectable levels before considerable interaction with the bottom occurs. Variability in ocean noise levels is still a significant issue at higher frequencies, but the variance in noise levels and the decibel range also tend to be smaller than at lower frequencies.

Under certain conditions, environmental uncertainties using single fixed sensors may be tolerable, especially when comparing calls at a fixed location over time. In this case, the bias in \hat{P} associated with unknown sediment structure may be large, but since it remains constant over time, it cancels out. On the other hand, the variation in \hat{P} due to changes in the sound speed profile at some locations can be significant when comparing calling activity over seasons. The large influence of SSP on \hat{P} was demonstrated at site SBC, where the SSP between summer and winter creates a threefold change in \hat{P} .

As for comparisons of calling activity at different hydrophone locations, uncertainties in estimates of P using single fixed sensors may be acceptable. For example, if the calls are homogeneously distributed at Hoke, the maximum uncertainty in estimates of \hat{P} associated with environmental variability is around 15%. Therefore, it may be possible to use this modeling technique to determine if there are more vocalizations per km² at one location compared to another, if the normalized call counts differ by more than the uncertainty in the probabilities of detection at the two sites.

The drastic variation in \hat{P} over both time at a given site, and across sites, highlights the dangers of comparing intra-site and inter-site calling activity without first accounting for environmental effects on the probability of detection. When an SNR constraint is not used as an additional filter on the GPL detector output, the probability of detecting humpback calls at site SBC can be greater than ten times the probability of detecting calls at site Hoke. Even if two sensors are located in regions with similar bathymetric and bottom conditions, differences in noise levels between two sites (or at the same site over time) of just a few decibels can easily change the probability of detection by a factor of two.

One application that involves quantifying \hat{P} is the estimation of the areal density of marine mammals from passive acoustic recordings of their calling activity. The animal density estimation equation based on measuring cue counts in a given area is given as [43]

$$\hat{D} = \frac{n_u(1 - \hat{c})}{K\pi w^2 \hat{P}T\hat{r}}, \quad (3.3)$$

where \hat{D} is the density estimate, n_u is the number of detected acoustic cues, \hat{c} is the number of false positive detections, K is the number of sensors (for single omni-directional sensors in a monitoring area, as in this paper, $K = 1$), w is the maximum detection range beyond which one assumes no acoustic cues are detected, \hat{P} is the estimated average probability of detection covered by the area πw^2 , T is the time period over which the units are tabulated, and \hat{r} is the estimated cue production rate.

The detector design criteria, including the detector threshold and additional constraints placed on received SNR, can influence the uncertainties in estimates of \hat{D} . From results presented in this paper, the uncertainty from environmental parameters in \hat{P} roughly increases with increasing area monitored. One possible approach for minimizing uncertainty is to raise the received minimum SNR threshold to values that correspond with direct path transmission from source to receiver. However, doing so decreases the cue counts for the time period of interest, thereby increasing the statistical variability of the estimates. Additionally, decreasing the monitored area could cause a violation of the assumption that calls are homogeneously distributed in space. Therefore, accurate density estimation

involves an optimization problem of determining how to estimate the various quantities in the equation for animal density such the uncertainty in \hat{D} is minimized.

Running a high fidelity, full wave field, ocean acoustic model using a span of likely environmental variables from historical data as input is an instructive and cost-efficient way of determining the environmental variables that most influence \hat{P} for a particular location. Results from the model help determine where best to allocate resources to decrease the uncertainty in \hat{P} . In some cases, in situ propagation calibration using a controlled acoustic source may be warranted to correctly characterize the bottom properties. Alternatively, bottom geoacoustic information can be derived from sediment cores and published empirical relations. In other cases, resources may be best allocated to recording monthly changes in the SSP, perhaps even weekly during transitional months in the fall and spring. Oceanographic models, coupled with satellite-based measurements such as sea surface temperature, may provide sufficient information on the temporal variability of the water column. In general, ancillary environmental information may be very helpful in reducing the uncertainty in \hat{P} to acceptable levels.

Site selection for sensor deployment in passive acoustic monitoring also play a vital role in reducing uncertainties in \hat{P} . Results from this paper suggest that hydrophones are best deployed in areas where the bathymetry, bottom type, and sound speed profiles are well characterized. If this information is not available, selecting locations that minimize sound interaction with the bottom will help reduce uncertainties in \hat{P} . Shallow bowl-shaped or trough-shaped basins tend to produce the most uncertainty in \hat{P} since the sound interacts the most with the bottom, and temporally-varying SSPs will focus this propagating sound in circular regions of temporally-varying distances from the hydrophones. Since the area monitored increases with the square of the distance from the hydrophone, small changes in the ranges of these acoustic convergence zones can have a large effect on the the amount of area from which an acoustic signal can be detected.

Results presented from the model/data comparison suggest that low and mid frequency calling whales can be used as acoustic sources of opportunity for

geoacoustic inversion of ocean bottom properties. If the whale source level, source depth, and source distribution, and ocean noise and SSP are known, then statistics on the distribution of the received SNR of calls at the receiver can be compared with acoustic models to significantly constrain the effective properties of the bottom. An example of the feasibility of this geoacoustic inversion approach was demonstrated at site SR (middle plot in Fig. 3.10), where a good match between the recorded data and model suggest that the sediment thickness ranges between 1 m and 10 m before encountering sedimentary rock. Running the model with 50 m sediment thickness gives a very poor model/data fit. If information on the source level and distribution of humpbacks in this region could be measured, then the inversion results on sediment thickness could be presented with reasonable confidence.

The uncertainties in \hat{P} presented in this paper assume complete accuracy of the CRAM model. The RAM core of the CRAM model is based on an estimate of a solution to the acoustic wave equation, and therefore is not exact. The model does not incorporate the shear properties of the bottom, which could influence the accuracy of the model, especially with higher density bottom types, such as at site Hoke. The model also does account for acoustic backscatter.

3.6 Conclusions

Acoustic propagation modeling is a useful tool for quantifying the probability of detection and the associated uncertainties in those measurements for single fixed sensors. For low and mid frequency vocalizations, simple propagation models are not sufficient for estimating \hat{P} . Rather, a more sophisticated model that includes bathymetry, sound speed, bottom characteristics and site specific noise to estimate the complex pressure field at the receiver is necessary. The environmental parameters that create the most uncertainty in the probability of detecting a signal are site specific; using an acoustic model with historical environmental data is an effective way for determining where best to allocate resources for minimizing the uncertainties in \hat{P} . In some instances, the errors associated with the uncertainties in \hat{P} may be sufficiently small, allowing for reasonable density estimates using single

fixed sensors. Results from this study suggest that comparing calling activity at the same sensor over time or across sensors in different geographical locations without first accounting for \hat{P} is a questionable procedure, as the probability of detecting calls can vary by factors of ten or more for low and mid frequency calling whales.

Acknowledgements

The authors are extremely grateful to Glenn Ierley, Megan McKenna, Amanda Debich, and Heidi Batchelor, all at Scripps Institution of Oceanography, for their support of this research. Gary Greene at Moss Landing Marine Laboratories, and David Clague and Maria Stone at MBARI were instrumental in obtaining bathymetric and ocean bottom information used in this study. Bathymetry data collected from R/V Atlantis, cruise ID AT15L24, were provided courtesy of Curt Collins (Naval Postgraduate School) and processed by Jennifer Paduan (MBARI). Shipping densities were provided by Chris Miller (Naval Postgraduate School). Special thanks to Sean Wiggins and the entire Scripps Whale Acoustics Laboratory for providing thousands of hours of high quality acoustic recordings. The CRAM acoustic propagation code used in this research was written by Richard Campbell and Kevin Heaney of OASIS, Inc., using Mike Collins' RAM program as the starting point. The first author would like to thank the Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship program, the Space and Naval Warfare (SPAWAR) Systems Command Center Pacific In-House Laboratory Independent Research program, and Rich Arrieta from the SPAWAR Unmanned Maritime Vehicles Lab for continued financial and technical support. Work was also supported by the Office of Naval Research, Code 32, the Chief of Naval Operations N45, and the Naval Postgraduate School.

Chapter 3 is, in full, a reprint of material accepted for publication in *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D'Spain, John A. Hildebrand, Greg S. Campbell, Richard L. Campbell, and Kevin D. Heaney "Site specific probability of passive acoustic detection of humpback whale

class from single fixed hydrophones”. The dissertation author was the primary investigator and author of this paper.

References

- [1] C.S. Clay and H. Medwin. *Acoustical oceanography: principles and applications*, volume 4, pages 84–89,114. Wiley, New York, NY, 1977.
- [2] P.C. Etter. *Underwater Acoustic Modeling and Simulation*, pages 82–84. Spon Press, New York, NY, 2003.
- [3] M.F. McKenna, D. Ross, S.M. Wiggins, and J.A. Hildebrand. Underwater radiated noise from modern commercial ships. *J. Acoust. Soc. Am.*, 131(1):92–103, 2012.
- [4] L. Thomas, T. Marques, D. Borchers, C. Stephenson, D. Moretti, R. Morrissey, N. DiMarzio, J. Ward, D. Mellinger, S. Martin, and P. Tyack. Density estimation for cetaceans from passive acoustic fixed sensors: Final programmatic report. Technical report, Center for research into ecological and environmental modeling, University of St. Andrews, Scotland, UK, 2011.
- [5] S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, and L. Thomas. *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, pages 1–448. Oxford University Press, New York, NY, 2001.
- [6] E.T. Küsel, D.K. Mellinger, L. Thomas, T.A. Marques, D. Moretti, and J. Ward. Cetacean population density estimation from single fixed sensors using passive acoustics. *J. Acoust. Soc. Am.*, 129(6):3610–3622, 2011.
- [7] S. Wiggins. Autonomous Acoustic Recording Packages (ARPs) for long-term monitoring of whale sounds. *Marine Tech. Soc. J.*, 37(2):13–22, 2003.
- [8] R.S. Payne and S. McVay. Songs of humpback whales. *Science*, 173(3997):585–597, 1971.
- [9] T.A. Helble, G.R. Ierley, G.L. D’Spain, M.A. Roch, and J.A. Hildebrand. A generalized power-law detection algorithm for humpback whale vocalizations. *J. Acoust. Soc. Am.*, 131(4):2682–2699, 2012.
- [10] C.S. Baker, L. Medrano-Gonzalez, J. Calambokidis, A. Perry, F. Pichler, H. Rosenbaum, J.M. Straley, J. Urban-Ramirez, M. Yamaguchi, and O. von Ziegeler. Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology*, 7(6):695–707, 1998.

- [11] J. Calambokidis, E.A. Falcone, T.J. Quinn, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, R. LeDuc, D. Mattila, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J.R. Urban, D. Weller, B.H. Witteveen, M. Yamaguchi, A. Bendlin, D. Camacho, K. Flynn, A. Havron, J. Huggins, and N. Maloney. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Technical report, Cascadia Research Collective, Olympia, WA, 2008.
- [12] C.S. Baker, D. Steel, J. Calambokidis, J. Barlow, A.M. Burdin, P.J. Clapham, E. Falcone, J.K.B. Ford, C.M. Gabriele, U. Gozález-Peral, R. LeDuc, D. Mattila, T.J. Quinn, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, R.J. Urban, M. Vant, P.R. Wade, D. Weller, B.H. Witteveen, K. Wynne, and M. Yamaguchi. geneSPLASH: An initial, ocean-wide survey of mitochondrial (mt) DNA diversity and population structure among humpback whales in the North Pacific: Final report for contract 2006-0093-008 Principal Investigator: C. Scott Baker. Technical report, Cascadia Research Collective, Olympia, WA, 2008.
- [13] G.P. Donovan. A review of IWC stock boundaries. *Reports of the International Whaling Commission (special issue)*, (13):39–68, 1991.
- [14] J.H. Johnson and A.A. Wolman. The humpback whale, *Megaptera novaeangliae*. *Marine Fisheries Review*, 46(4):30–37, 1984.
- [15] J. Barlow. The abundance of cetaceans in California waters. Part I: Ship surveys in summer and fall of 1991. *Fishery Bulletin*, 93:1–14, 1995.
- [16] J. Calambokidis, G.H. Steiger, K. Rasmussen, J. Urban, KC Balcomb, PL de Guevara, M. Salinas, JK Jacobsen, CS Baker, LM Herman, S. Cerchio, and JD Darling. Migratory destinations of humpback whales that feed off California, Oregon and Washington. *Marine Ecology-Progress Series.*, 192:295–304, 2000.
- [17] J. Calambokidis, G.H. Steiger, J.M. Straley, L.M. Herman, S. Cerchio, D.R. Salden, U.R. Jorge, J.K. Jacobsen, O. von Ziegesar, K.C. Balcomb, C.M. Gabriele, M.E. Dahlheim, S. Uchida, G. Ellis, Y. Miyamura, P.L.P. de Guevara, M. Yamaguchi, F. Sato, S.A. Mizroch, L. Schlender, K. Rasmussen, J. Barlow, and T.J. Quinn. Movements and population structure of humpback whales in the North Pacific. *Marine Mammal Science*, 17(4):769–794, 2001.
- [18] J. Calambokidis, G.H. Steiger, J.R. Evenson, K.R. Flynn, K.C. Balcomb, D.E. Claridge, P. Bloedel, J.M. Straley, C.S. Baker, O. von Ziegesar, ME Dahlheim, JM Waite, JD Darling, G Ellis, and GA Green. Interchange and isolation of humpback whales off California and other North Pacific feeding grounds. *Marine Mammal Science*, 12(2):215–226, 1996.

- [19] J. Calambokidis, G.H. Steiger, D.K. Ellifrit, B.L. Troutman, and C.E. Bowlby. Distribution and abundance of humpback whales (*Megaptera novaeangliae*) and other marine mammals off the northern Washington coast. *Fishery Bulletin*, 102(4):563–580, 2004.
- [20] R.J. Urban, C.F. Alvarez, M.Z. Salinas, J. Jacobsen, K.C. Balcomb, A.L. Jaramillo, P.L. de Guevara, and A.L. Aguayo. Population size of humpback whale, *Megaptera novaeangliae*, in waters off the Pacific coast of Mexico. *Fisheries Bulletin*, 97(4):1017–1024, 1999.
- [21] J. Calambokidis, E. Falcone, A. Douglas, L. Schlender, and J. Huggins. Photographic identification of humpback and blue whales off the US west coast: Results and updated abundance estimates from 2008 field season. Technical report, Cascadia Research Collective, Olympia, WA, 2009.
- [22] G.S. Campbell, T.A. Helble, S.M. Wiggins, and J.A. Hildebrand. Humpback whale seasonal and spatial calling patterns in the temperate northeastern Pacific Ocean: 2008-2010. In *Proceedings-19th Biennial Conference on the Biology of Marine Mammals*, page 53, Tampa, FL, 2011.
- [23] Perkins, P.J. Cornell laboratory of ornithology macaulay library: Humpback whale, *Megaptera novaeangliae*, 1973. date last viewed 12/14/11.
- [24] NOAA National Geophysical Data Center. U.S. coastal relief model, vol. 6, 2011. date last viewed 12/16/11.
- [25] C. Amante and B. W. Eakins. ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. Technical report, NOAA National Geophysical Data Center, Boulder, CO, 2009.
- [26] T.P. Boyer, J.I. Antonov, O.K. Baranova, H.E. Garcia, D.R. Johnson, R.A. Locarnini, A.V. Mishonov, T. O'Brien, D. Seidov, I.V. Smolyar, M.M. Zweng, and S. Levitus. World ocean database 2009. *NOAA Atlas NESDIS*, 66:1–116, 2009.
- [27] Ocean Acoustics Group, Massachusetts Institute of Technology. The Santa Barbara Channel Experiment, 1999. date last viewed 5/12/12.
- [28] E.L. Hamilton. Sound velocity–density relations in sea-floor sediments and rocks. *J. Acoust. Soc. Am.*, 63(2):366–377, 1978.
- [29] E.L. Hamilton. Sound velocity gradients in marine sediments. *J. Acoust. Soc. Am.*, 65(4):909–922, 1979.
- [30] C.K. Wentworth. A scale of grade and class terms for clastic sediments. *J. Geology*, 30(5):377–392, 1922.

- [31] W.C. Krumbein and L.L. Sloss. *Stratigraphy and Sedimentation*, pages 1–497. W. H. Freeman and Co., New York, NY, 1951.
- [32] K.M. Marsaglia, K.C. Rimkus, and R.J. Behl. Provenance of sand deposited in the Santa Barbara Basin at Site 893 during the last 155,000 years. In *Proceedings-Ocean Drilling Program Scientific Results*, pages 61–76. National Science Foundation, 1992.
- [33] J.A. de Mesquita Onofre. Analysis and modeling of the acoustic tomography signal transmission from Davidson Seamount to Sur Ridge: The forward problem. Master’s thesis, Naval Postgraduate School, 1999.
- [34] C.L. Gabriel. The physical characteristics of bottom sediment near Sur Ridge, California. Master’s thesis, Naval Postgraduate School, March 2001.
- [35] J.G. Konter, H. Staudigel, J. Blichert-Toft, B.B. Hanan, M. Polvé, G.R. Davies, N. Shimizu, and P. Schiffman. Geochemical stages at Jasper Seamount and the origin of intraplate volcanoes. *Geochem. Geophys. Geosyst.*, 10(2):Q02001, 2009.
- [36] A.H. Nuttall. Detection performance of power-law processors for random signals of unknown location, structure, extent, and strength. Technical report, NUWC-NPT, Newport, RI, 1994.
- [37] T.A. Marques, L. Munger, L. Thomas, S. Wiggins, and J.A. Hildebrand. Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research*, 13:163–172, 2011.
- [38] M.A. McDonald and C.G. Fox. Passive acoustic methods applied to fin whale population density estimation. *J. Acoust. Soc. Am.*, 105(5):2643–2651, 1999.
- [39] W.W.L. Au, A.A. Pack, M.O. Lammers, L.M. Herman, M.H. Deakos, and K. Andrews. Acoustic properties of humpback whale songs. *J. Acoust. Soc. Am.*, 120(2):1103–1110, 2006.
- [40] M.D. Collins. *User’s Guide for RAM Versions 1.0 and 1.0p*. Naval Research Laboratory, Washington, DC, 2002.
- [41] R.J. Urick. *Principles of Underwater Sound*, volume 3, pages 19–22. McGraw-Hill, New York, NY, 1983.
- [42] R. Campbell and K. Heaney. *User’s Guide for CRAM*. Ocean Acoustical Services and Instrumentation Systems, Inc., Fairfax Station, VA, 2012.
- [43] T.A. Marques, L. Thomas, J. Ward, N. DiMarzio, and P.L. Tyack. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales. *J. Acoust. Soc. Am.*, 125(4):1982–1994, 2009.

Chapter 4

Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics

Abstract

This paper demonstrates the importance of accounting for environmental effects on passive underwater acoustic monitoring results. The situation considered is the reduction in shipping off the California coast between 2008-2010 due to the recession and environmental legislation. The resulting variations in ocean noise change the probability of detecting marine mammal vocalizations. An acoustic model was used to calculate the time-varying probability of detecting humpback whale vocalizations under best-guess environmental conditions and varying noise. The uncorrected call counts suggest a diel pattern and an increase in calling over a two-year period; the corrected call counts show minimal evidence of these features.

4.1 Introduction

Passive acoustic monitoring is an important tool for understanding marine mammal ecology and behavior. When studying an acoustic record containing marine mammal vocalizations, the received signal can be greatly influenced by the environment in which the sound is transmitted. The ocean bottom properties, bathymetry, and temporally varying sound speed act to distort and reduce the energy of the original waveform produced by the marine mammal. In addition, constantly varying ocean noise further influences the detectability of the calls. This ever-changing acoustic environment creates difficulties when comparing marine mammal recordings between sensors, or at the same sensor over time.

One way to correct for temporal and spatial variations in detectability due to environmental effects can be obtained from the expression for estimating the spatial density of marine mammals from passive acoustic recordings; Eq. (3) of Marques *et al.*, 2009[1]. The corrected call counts in Eq. (3) is

$$\hat{N}_c \equiv n_c \frac{1 - \hat{c}}{\hat{P}} \quad (4.1)$$

where n_c is the number of detections (uncorrected call count) in the data, \hat{c} is the probability of false detection, and \hat{P} is the probability of detection. In the case where human analysts scan the detection outputs generated by an automated detection algorithm to eliminate false detections (i.e., $\hat{c} = 0$) as is done with the data presented in this paper, the calibration factor is the estimated probability of detection, \hat{P} . Helble et al.[2] demonstrated that \hat{P} can change by factors greater than ten between sensors at different locations or at the same sensor over time. At some sites, \hat{P} has an exponential dependence on ocean noise level and hence a seemingly modest change in noise, itself insignificant in the high dynamic range spectrograms commonly used to detect vocalizations, can nonetheless greatly skew the counts of calling activity. To illustrate the influence that the ocean environment has on the detection of marine mammal vocalizations, two single hydrophone datasets simultaneously recorded over a 2-year period using High-frequency Acoustic Recording Packages (HARP)[3] were analyzed for humpback whale (*Megaptera novaeangliae*) vocalizations. The recorded detection counts

were corrected to account for the influence of environmental properties using the numerically-derived probability of detection. The resulting environmentally-calibrated datasets provide a more valid approach to examining both short-term and long-term calling trends of the biological sources themselves.

The two sites used for this study are located off the coast of California[2]. Site SBC (34.2754°, -120.0238°) is located in the center of the Santa Barbara Channel, and site SR (36.3127°, -122.3926°) is located on Sur Ridge, a bathymetric feature 45 km southwest of Monterey. Data recording covers the period from January, 2008 to January, 2010, during which a decrease in shipping noise occurred at both locations due to a downturn in the world economy, coupled with the implementation of an air-quality improvement rule on 1 July, 2009, by the California Air Resources Board (CARB). McKenna et al.[4] discovered that these events in combination reduced the monthly average ocean noise level by 12 dB in the 40 Hz band over a period from 2007 to 2010 at site SBC. The changing ocean noise characteristics at these two sites create significant changes in \hat{P} on both short-term and long-term time scales.

4.2 Methods

Inputs to a full wavefield acoustic propagation model, "CRAM"[5], were developed for both site SBC and site SR. The model CRAM is the C-language version of the parabolic-equation-based Range-dependent Acoustic Model (RAM)[6]. This code was used to simulate the propagation of humpback call units from source to receiver, in amplitude and phase as a function of frequency. The model simulated calls originating from geographical locations evenly spaced on a square lattice bounded by a 20 km radial distance from the HARP, at 20 m depth. The simulated received humpback units for each site were added to time-varying noise recorded from each site and the generalized power-law detector[7] was used to process the combined waveform. Resulting probability of detection maps were created as a function of latitude and longitude for the areas surrounding each HARP. From these maps, the average probability of detection for a 20 km radial

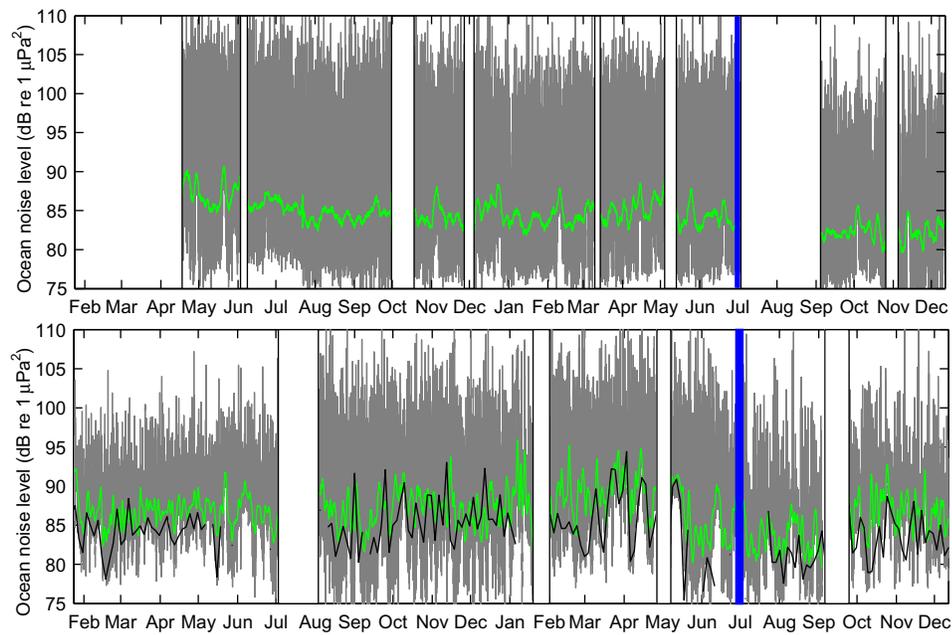


Figure 4.1: Ocean noise levels in the 150-1800 Hz band over the 2008-2009 period at site SBC (upper) and SR (lower). The gray curves indicate the noise levels averaged over 75 sec increments, the green curves are the running mean with a 7 day window, and the black curve (site SR only) is a plot of the average noise levels in a 7-day window measured at the times adjacent to each detected humpback unit. White spaces indicate periods with no data. The blue vertical lines mark the start of enforcement of CARB law.

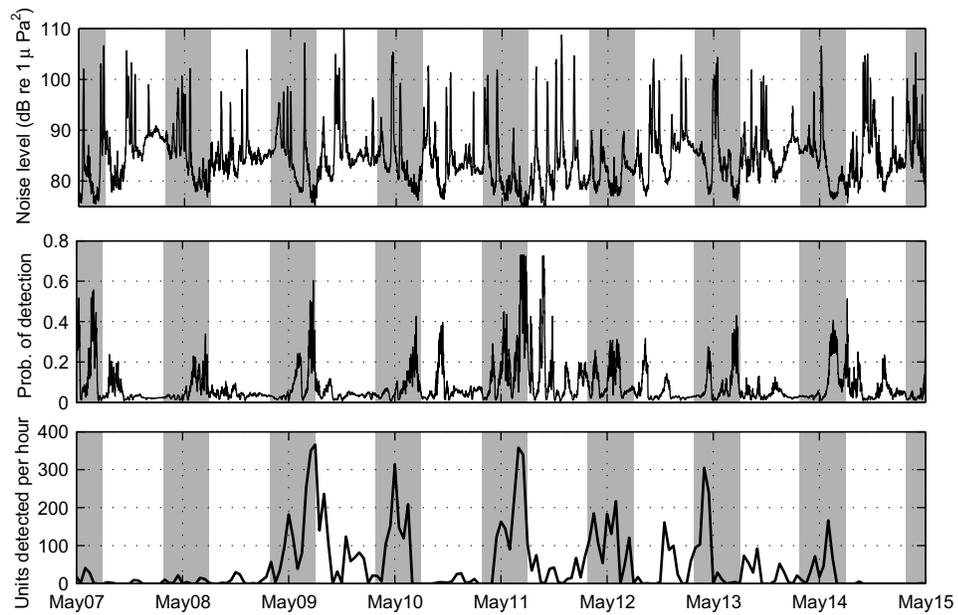


Figure 4.2: Ocean noise levels at site SBC in May, 2008 (upper), probability of detecting a humpback unit (\hat{P}) within a 20 km radius of site SBC in May 2008 (middle), and the number of humpback units detected in uncorrected form (n_c) at site SBC for the same time period (lower). Shaded time periods indicates sunset to sunrise. The vertical grid lines indicate midnight local time.

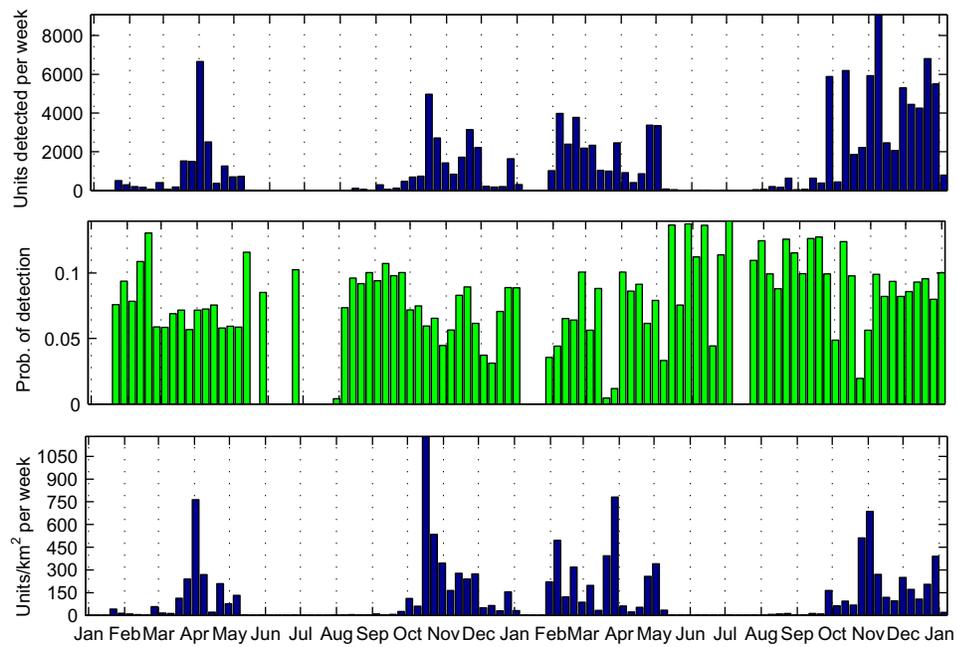


Figure 4.3: (color online) Uncorrected number of humpback units detected (n_c) in the 2008-2009 period at site SR (upper), estimated probability of detecting a humpback unit (\hat{P}) within a 20 km radius of site SR (middle), and the corrected estimated number of units occurring per unit area (\hat{N}_c) at site SR for the same time period (lower).

area was determined for a full range of noise conditions, yielding probability of detection versus noise curves for both site SBC and site SR, as described in Helble et al.[2]. The inputs to the model were varied over the range of uncertainties in both bottom properties and sound speed profiles at each site so that the uncertainties in \hat{P} could be characterized.

Sound speed profiles were chosen at site SBC from casts that were taken during the recording period very near to the recording package. The October, 2008 cast was used for the months between June to October, while the May, 2009 cast was used for months between November to May. For site SR, the same approach was used, except the sound speed profiles were taken from historical samples because no casts were taken during the data recording period. Monthly variations in sound speed profiles changed estimates in \hat{P} by no more than 20% for site SBC and 10% at site SR. In contrast, changes in sound speed profile that occur between summer and winter profile types can lead to significantly greater changes in \hat{P} at site SBC (only slightly higher than 10% change at site SR)[2]. Therefore, updating the input sound speed profile bi-annually captured this seasonally variability in the modeling.

For each call detected within a 75 sec period, the average of six noise measurements within that time period was used to determine \hat{P} for that time period. The number of calls detected in that time period (n_c) was then divided by \hat{P} , giving the estimated number of call units that actually occurred within the 20 km radial area surrounding the HARP (\hat{N}_c), assuming a uniform distribution of calling animals in the area monitored. In order to satisfy this assumption, detected units were tabulated in weekly increments. Model/data comparisons from Helble et al.[2] indicate this assumption likely is true at least on monthly time scales for both sites SR and SBC. The resulting normalized call counts were provided in number of units per km² per week. On shorter time scales, the calling animals cannot be assumed to be uniformly distributed. However, comparing unnormalized call counts with variations in \hat{P} on shorter times scales is important to gain an understanding of the correlation between detection counts and variations in ocean noise levels, and this analysis was carried through for site SBC (discussed in the

next section).

4.3 Results

Ocean noise levels averaged over consecutive 75-sec periods between 2008-2009 varied by up to 35 dB at both locations (Fig. 4.1, 75-110 dB re $1 \mu\text{Pa}^2$ in the 150-1800 Hz band). The 7-day running means of the noise (green curves) are better able to reveal long-term changes in the noise. The decrease at SBC of approximately 5 dB over the course of the deployment is consistent with the trend described by McKenna et al.[4] and correlates with the onset of the Great Recession, which significantly reduced maritime trade. An additional reduction in ocean noise at SBC occurred after July 1, 2009, with the enforcement of the CARB air quality improvement rule. It resulted in a diversion of much of the shipping traffic to transit lanes outside of the channel. Similar results can be seen for site SR - a significant drop occurs in both ocean noise levels and in the variance of ocean noise when comparing the Aug-Dec, 2008 levels with those of Aug-Dec, 2009. The time period from Feb-Jul, 2008 cannot be directly compared to Feb-Jul, 2009 because the sensor during the former time period was located 10 km southwest of the ridge, in deeper water. The black curve for site SR in Fig. 4.1 indicates the 7-day average noise level when each noise estimate used in the average is made from the 75-sec time period surrounding each detected humpback unit. When averaging the noise estimates this way, the resulting noise level generally falls below the running mean noise level for the same time period (i.e., the black curve generally falls below the green curve), because an increasing number of units is detected during periods of lower noise. This discrepancy indicates the need to obtain noise estimates during the periods of marine mammal vocalization detections; using a simple running-mean noise average does not properly represent the noise environment in which the calls are detected.

Fig. 4.2 shows ocean noise levels for site SBC for a one week period in May, 2008 (upper plot), the related values of \hat{P} (middle plot), and the uncorrected number of units detected per hour over the same period (lower plot). Examination

of the lower plot by itself would indicate a strong diel cycle to the humpback calling activity, with significantly more calls occurring during nighttime. However, inspection of \hat{P} indicates a significant diel cycle in the likelihood of detecting humpback units. This change in \hat{P} accounts for most of the diel signal found in the humpback calling pattern for this period. While nearby passages of ships are easily identified (short duration spikes in the upper plot), smaller noise variations centered near 80 dB re 1 μPa^2 are difficult to notice if detections are manually marked from a spectrogram. When ocean noise levels at site SBC drop from 80 dB re 1 μPa^2 to 75 dB re 1 μPa^2 , \hat{P} increases from 0.1 to 0.65, which illustrates the importance of correcting for subtle variations in noise at this site (in contrast, large spikes in noise that occur in a high noise environment have little effect reducing \hat{P} because \hat{P} is already low). Changes of only a few decibels in noise level can have substantially different effects on the change in \hat{P} depending on the site specific bathymetric and environmental parameters. At site SBC, \hat{P} decreases exponentially with increasing noise, making changes in \hat{P} more dramatic over relatively small changes in noise at lower levels, whereas at site SR \hat{P} changes quadratically[2].

The plots in Fig. 4.3 show the uncorrected number of units detected in weekly time bins at site SR from 2008-2009 (upper), the time-varying probability of detecting a humpback unit (middle), and the corrected, estimated number of humpback units occurring per unit area (lower) for the same time period. The weekly estimates of P were calculated by averaging the values of \hat{P} measured at each detected unit. The decrease in ocean noise due to the economic downturn and the enforcement of the CARB air-quality improvement rule creates an increase in \hat{P} for the Sep-Jan, 2009 time period compared to Sep-Jan, 2008. While substantially more units are detected in the Sep-Jan, 2009 time frame (190% increase in the upper plot), the increase in detections during this period is not a biological effect, but rather is driven by the changing noise conditions. After the the uncorrected call counts are "calibrated" by \hat{P} , the estimated number of units occurring between Sep-Jan, 2009 is approximately equal to the number estimated for the same period in the previous year (8% decrease in the lower plot). The uncertainties associated

with \hat{P} due to environmental and source characteristics, the main sources of uncertainty in \hat{P} , are discussed in Helble et al. [2]. A full analysis of all the uncertainties in \hat{P} is beyond the scope of this manuscript and is a subject of current research. Although the absolute numbers for N_c in the lower plot of Fig. 4.3 are uncertain, confidence in the temporal dependence of N_c at a given site is much greater since it is driven to a large extent by the temporal variability in the noise, which can be readily measured with the real data.

4.4 Discussion

The downturn in the world economy, combined with the enforcement of CARB air-quality improvement rule provides a concrete example of how changing ocean noise conditions can skew the results of long-term marine mammal monitoring efforts. For site SR, lower noise during the fall of 2009 compared to the fall of 2008 resulted in an increase number of detections between these periods. After correcting for \hat{P} over the time period, values of \hat{N}_c were roughly the same at site SR between the two seasons. While this change in economic conditions between 2008 and 2010 provides a convenient example for studying the influence of noise on \hat{P} , changing ocean noise conditions on these long time scales are by no means unique. For example, ocean noise levels have risen by an estimated 3 dB/decade in some locations[8, 9] due to an increase in global shipping. Additionally, changing economic conditions, ship traffic routes, ship propeller design, fluctuations in tourism, and changes in weather patterns can all create similar effects at various locations world-wide[10, 11, 12, 13, 14, 15, 16, 17]. Short-term changes in ocean noise must also be accounted for, because \hat{P} can rise and fall on time scales important for habitat and predator/prey studies. One such example can be seen at site SBC (Fig. 4.2), where a strong diel pattern in humpback acoustic detections is heavily influenced by shipping patterns in the region.

The influence of changing \hat{P} is even more pronounced when scientists attempt to assess the potential impact of noise on marine mammals[17], because

the acoustic conditions under which the biological signals are recorded are heavily influenced by the noise. Correcting acoustic detections by \hat{P} removes these biases. Unfortunately, correcting short-time series by \hat{P} becomes problematic if not enough calls are detected to satisfy the assumed homogeneous random distribution of animals in the study area. This assumption can be relaxed in cases where the passive monitoring systems provide localization capabilities, or multiple omnidirectional sensors with overlapping coverage are deployed within a study area. However, understanding changes in \hat{P} on short time scales is still very useful; it indicates the degree to which the environment influences the acoustic detections.

In summary, if passive acoustic detections of marine mammal calls are to become an integral part of marine mammal monitoring, biological studies, and ecological assessments, estimates of the probability of detection, P , should become a standard approach to assessing animal presence and calibrating for environmental effects.

Acknowledgements

The authors are extremely grateful to Prof. Glenn Ierley, Dr. Megan McKenna, and Amanda Debich, both at the Scripps Institution of Oceanography, for their support of this research. Special thanks to Sean Wiggins and the entire Scripps Whale Acoustics Laboratory for providing thousands of hours of high quality acoustic recordings. The first author would like to thank the Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship program, the Space and Naval Warfare (SPAWAR) Systems Command Center Pacific In-House Laboratory Independent Research program, and Rich Arrieta from the SPAWAR Unmanned Maritime Vehicles Lab for continued technical and financial support. Work was also supported by the Office of Naval Research, Code 322 (MBB), the Chief of Naval Operations N45, and the Naval Postgraduate School.

Chapter 4 is a manuscript in preparation for submission to *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D'Spain,

Greg S. Campbell, and John A. Hildebrand, “Calibrating passive acoustic monitoring: Correcting humpback whale call detections for site-specific and time-dependent environmental characteristics”. The dissertation author was the primary investigator and author of this paper.

References

- [1] T.A. Marques, L. Thomas, J. Ward, N. DiMarzio, and P.L. Tyack. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales. *J. Acoust. Soc. Am.*, 125(4):1982–1994, 2009.
- [2] T.A. Helble, G.L. D’Spain, J.A. Hildebrand, G.S. Campbell, R.L. Campbell, and K.D. Heaney. Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones. *J. Acoust. Soc. Am.*, *accepted for publ.*, 2013.
- [3] S. Wiggins. Autonomous Acoustic Recording Packages (ARPs) for long-term monitoring of whale sounds. *Marine Tech. Soc. J.*, 37(2):13–22, 2003.
- [4] M.F. McKenna, S.L. Katz, S.M. Wiggins, D. Ross, and J.A. Hildebrand. A quieting ocean: Unintended consequence of a fluctuating economy. *J. Acoust. Soc. Am.*, 132(3):EL169–EL175, 2012.
- [5] R. Campbell and K. Heaney. *User’s Guide for CRAM*. Ocean Acoustical Services and Instrumentation Systems, Inc., Fairfax Station, VA, 2012.
- [6] M.D. Collins. *User’s Guide for RAM Versions 1.0 and 1.0p*. Naval Research Laboratory, Washington, DC, 2002.
- [7] T.A. Helble, G.R. Ierley, G.L. D’Spain, M.A. Roch, and J.A. Hildebrand. A generalized power-law detection algorithm for humpback whale vocalizations. *J. Acoust. Soc. Am.*, 131(4):2682–2699, 2012.
- [8] R.K. Andrew, B.M. Howe, J.A. Mercer, and M.A. Dzieciuch. Ocean ambient sound: comparing the 1960s with the 1990s for a receiver off the California coast. *Acoustics Research Letters Online*, 3(2):65–70, 2002.
- [9] D. Ross. On ocean underwater ambient noise. *Institute of Acoustics Bulletin*, 18:5–8, 1993.
- [10] G.M. Wenz. Review of underwater acoustics research: noise. *J. Acoust. Soc. Am.*, 51(3B):1010–1024, 1972.

- [11] P. Kaluza, A. Kölzsch, M.T. Gastner, and B. Blasius. The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):1093–1103, 2010.
- [12] K.I. Matveev. Effect of drag-reducing air lubrication on underwater noise radiation from ship hulls. *Journal of vibration and acoustics*, 127(4):420–422, 2005.
- [13] P.T. Arveson and D.J. Vendittis. Radiated noise characteristics of a modern cargo ship. *J. Acoust. Soc. Am.*, 107:118–129, 2000.
- [14] M.F. McKenna, D. Ross, S.M. Wiggins, and J.A. Hildebrand. Underwater radiated noise from modern commercial ships. *J. Acoust. Soc. Am.*, 131(1):92–103, 2012.
- [15] V.O. Knudsen, RS Alford, and JW Emling. Underwater ambient noise. *J. Mar. Res.*, 7(3):410–429, 1948.
- [16] G.M. Wenz. Acoustic ambient noise in the ocean: spectra and sources. *J. Acoust. Soc. Am.*, 34(12):1936–1956, 1962.
- [17] National Research Council. *Ocean Noise and Marine Mammals*, pages 83–132. National Academies Press, Washington, DC, 2003.

Chapter 5

Humpback whale vocalization activity at Sur Ridge and in the Santa Barbara Channel from 2008-2009, using environmentally corrected call counts

Abstract

Humpback whales (*Megaptera novaeangliae*) are relatively unstudied during their seasonal migrations along the California coast. Single-fixed passive acoustic sensors were monitored for two years at two locations off the coast of California, and acoustic calls were tabulated on the sensor using an automated detector. The acoustic probability of detection was calculated for each sensor over varying environmental and ocean noise conditions, allowing the acoustic calls to be presented in call densities (calls per km² per time). The corrected call counts allow for direct comparison of call densities across sensors and at the same sensor over time. Results indicated peak vocal density in the spring and fall months at both sensors, corresponding to humpback whales transiting to and from wintering

grounds. A strong nocturnal vocalization pattern was discovered at both locations, peaking in the month of April. Additionally, the results indicate the call rate and source level change with ocean noise level, suggesting a Lombard effect in vocalization behavior of humpback whales along the migration route.

5.1 Introduction

Humpback whales observed off the California coast typically belong to the eastern north Pacific stock, one of four separate migratory stocks in the Pacific Ocean basin [1, 2, 3, 4, 5, 6, 7]. This stock typically feeds during spring, summer, and fall in temperate to near polar waters along the northern rim of the Pacific, extending from southern California in the east northward to the Gulf of Alaska, and then westward to the Kamchatka peninsula. During winter months, the majority of the population migrates to warm temperate and tropical sites for mating and birthing. While considerable data have been collected on this stock both on the winter feeding ground and on the summer breeding grounds, little is known about the behavior of these whales along the migration route[7]. California Coastal Ocean Fisheries Investigations (CalCOFI) cruises, limited to four observation periods per year, provide data containing visual and acoustic presence of various marine mammal species in the southern California Bight, including humpback whales. While useful, these datasets provide limited information about humpback behavior in the region. Over the past decade, an increasing number of High-frequency Acoustic Recording Packages (HARP)[8] have been deployed in the region. Each HARP contains a hydrophone tethered above a seafloor-mounted instrument frame, and is deployed in water depths ranging from 200 m to 1500 m. Until recently, all analysis was performed manually by trained human analysts, marking the presence/absence of humpback acoustic activity within one-hour time bands. The development of the Generalized Power-Law (GPL) detector for humpback whale vocalizations [9] has allowed for the detection of nearly all humanly detectable humpback units within an acoustic record. Humpback whales produce underwater 'song', that has a hierarchal structure where individual sounds

are termed 'units'. These units are grouped into 'phrases', and phrases are grouped into 'themes', which combine to make up the song[10]. Observations of acoustic records have revealed the presence of humpback song in the southern California Bight from August - May, and feeding and social calls have also been observed year round. Feeding and social calls generally have less variation in unit type, and lack the complex hierarchy observed in song[11, 12]. While it was once commonly assumed that the southern California Bight was simply a transportation route for migrating whales, it has become more clear that humpback movement and behavior throughout the Bight is more complex, and the region could provide crucial feeding habitat or other social functions. The approach used in this paper for expanding the knowledge of humpback ecology and biology in the region is to examine humpback calling patterns over time and across HARP sensors. In order to better understand humpback call density in the region, acoustic models were developed to correct for the site and time-specific probability of passive acoustic detections on the sensors. Each HARP sensor has unique environmental, bathymetric, and background noise characteristics that influence the number and types of recorded humpback calls. Therefore, without correcting for the probability of detection, it is impossible to compare call counts across sensors, or at the same sensor over time. Habitat modeling, which seeks to explain correlations in animal presence and behavior to biological and environmental inputs, would be fraught with error unless corrected for environmental effects.

The objective of this paper is to count acoustic humpback calls at two sensors over a two-year period, convert these call counts into calling densities, and then observe the record for biological and ecological relevant information. The approach for converting acoustic calls into calling densities is described in Helble et. al.[13]. The approach is applied to the Santa Barbara Channel (site SBC) and on Sur Ridge (site SR), located off the coast of Monterey, respectively. The GPL processor was used in combination with the acoustic model, using call counts tallied from the HARP sensors, to produce humpback calling densities at these two sites over the period Jan 1, 2008, to Dec 31, 2009.

This paper is divided into four parts: Section 5.2 highlights the methods

used to obtain humpback calling densities, set in the framework established for calculating passive acoustic animal density estimates. The approach for estimating the uncertainty in the in the resulting call densities is also presented. Section 5.3 provides calling densities and the related uncertainty estimates at the two monitored locations. The humpback calling densities are presented over a variety of time periods and are also presented over several environmental variables, including time of day, lunar variation, and background noise level. Section 5.4 discusses the biological and ecological importance of the resulting call densities presented in Section 5.3, and compares the results to other humpback whale studies. Additionally, the practicality of using single-fixed sensors for humpback density estimates is discussed.

5.2 Methods

The methods for obtaining humpback vocalization densities (described in units/km² per time) are described in detail in a series of publications by Helble et al.[9, 13, 14]. The methods are based on previous publications that describe the methods for estimating whale density estimation (\hat{D}) from passive acoustics[15, 16, 17, 18, 19]. Eq. (3) of Marques *et al.*, 2009[19] gives \hat{D} as

$$\hat{D} \equiv \frac{n_c(1 - \hat{c})}{K\pi w^2 \hat{P}T\hat{r}} \quad (5.1)$$

where n_c is the number of detections (uncorrected call count) in the data, \hat{c} is the estimated probability of false detection, \hat{P} is the estimated probability of detecting a cue within distance w , \hat{r} is the estimated cue production rate, T is the time over which the whole density estimate estimate is made, and K is the number of independent sensors used in monitoring a given area. For the case of humpback whales, a cue is defined as any detected humpback unit within the 150 to 1800 Hz frequency band. Because the cue rate, \hat{r} is poorly known for humpback whales during migration, and likely highly variable, producing meaningful values of \hat{D} is not possible at present. Instead, cue density is used as a metric for humpback activity within an area, A , reducing Eq. (5.1) to

$$\hat{\rho}_c \equiv \frac{\hat{N}_c}{AT} \equiv \frac{n_c}{\hat{P}AT} \quad (5.2)$$

for a single sensor ($K = 1$) where \hat{N}_c is the estimated number of true humpback units within the assumed monitored area ($A = \pi w^2$) over the time duration T . The value $\hat{c} = 0$ is applicable in the case where human analysts scan the detection outputs generated by an automated detection algorithm to eliminate false detections, as is done with the data presented in this paper.

Values of n_c were obtained for the HARP recordings using the GPL detector. Values of \hat{P} were obtained for each HARP location over the full range of likely environmental and ocean noise conditions using full-field acoustic propagation modeling[13]. The estimates of humpback call densities were obtained using the methods outlined in Helble et al.[14].

5.2.1 Uncertainty Estimates

As mentioned above and detailed in previous publications[9, 13, 14], one approach to "calibrating" detected call counts for environmental properties can be obtained by numerically estimating the detection performance, specifically the probabilities of detection and false alarm. That is, the estimated environmentally-corrected number of call counts, \hat{N}_c , from the expression above, is

$$\hat{N}_c \equiv n_c \frac{1 - \hat{c}}{\hat{P}} \quad (5.3)$$

The quantity of interest is the estimated areal and temporal density of calls, $\hat{\rho}_c$, i.e., the number of calls per unit area per unit time as described in Eq. (5.2).

Both the (true and) estimated probabilities of detection, \hat{P} , and of false alarm, \hat{c} , are determined by the detector and its threshold. In fact, the detector "receiver operating characteristic" (ROC) curve is a plot of these two probabilities as a function of the threshold setting. The estimated environmental calibration factor is simply the ratio of these two probabilities, $(1 - \hat{c})/\hat{P}$. From a statistical point of view, these estimated probabilities are random variables, so that the environmental calibration factor should be written in terms of their means, $\mu(\hat{c})$

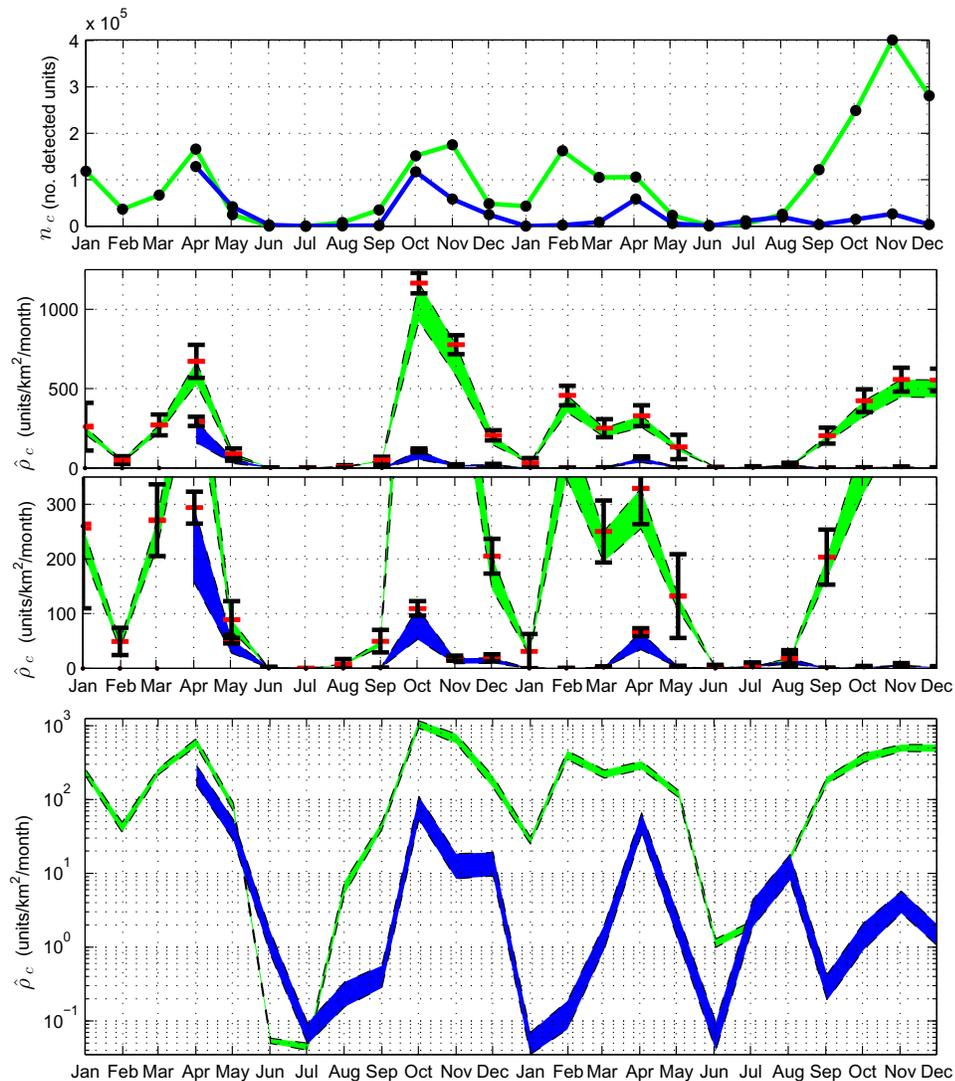


Figure 5.1: Uncorrected call counts n_c , normalized for effort (recording duty cycle) and tallied in 1-month bins for site SR (green) and SBC (blue) (upper panel), corrected estimated call density, $\hat{\rho}_c$, for site SR (green) and site SBC (blue) (middle panels) tallied in 1-month bins. The same datasets are repeated in both panels to illustrate scale. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Values of $\hat{\rho}_c$, for site SR (green) and site SBC (blue) are also repeated in the lower plot on a log scale to illustrate detail.

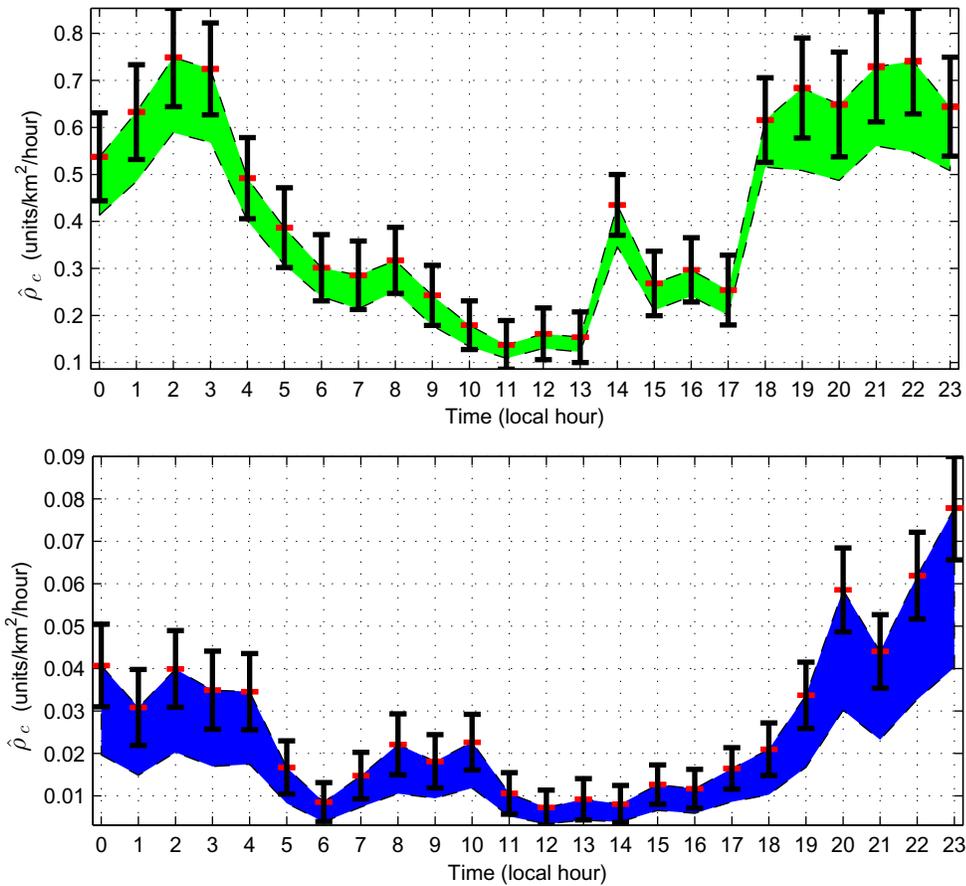


Figure 5.2: Average daily estimated call density, $\hat{\rho}_c$ shown in 1 hour time bins to illustrate diel cycle for site SR (upper panel) and site SBC (lower panel) for time period covering April 16, 2008 to Dec 31, 2009. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots.

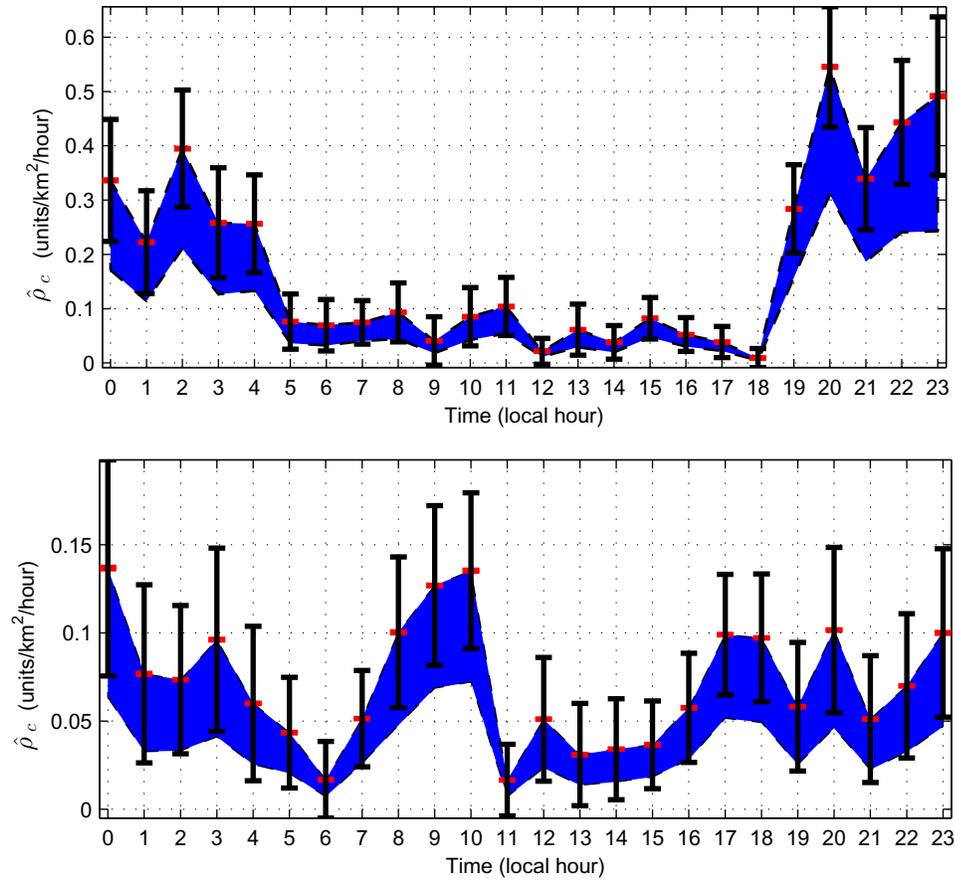


Figure 5.3: Average daily estimated call density, $\hat{\rho}_c$ at site SBC shown in 1 hour local time bins to illustrate diel cycle. The spring season (Apr 7-May 27, 2009) at site SBC (upper panel) shows stronger diel pattern and higher call densities than the fall season (Oct 15-Dec 4, 2009) at site SBC (lower panel). The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate the standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate the standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots.

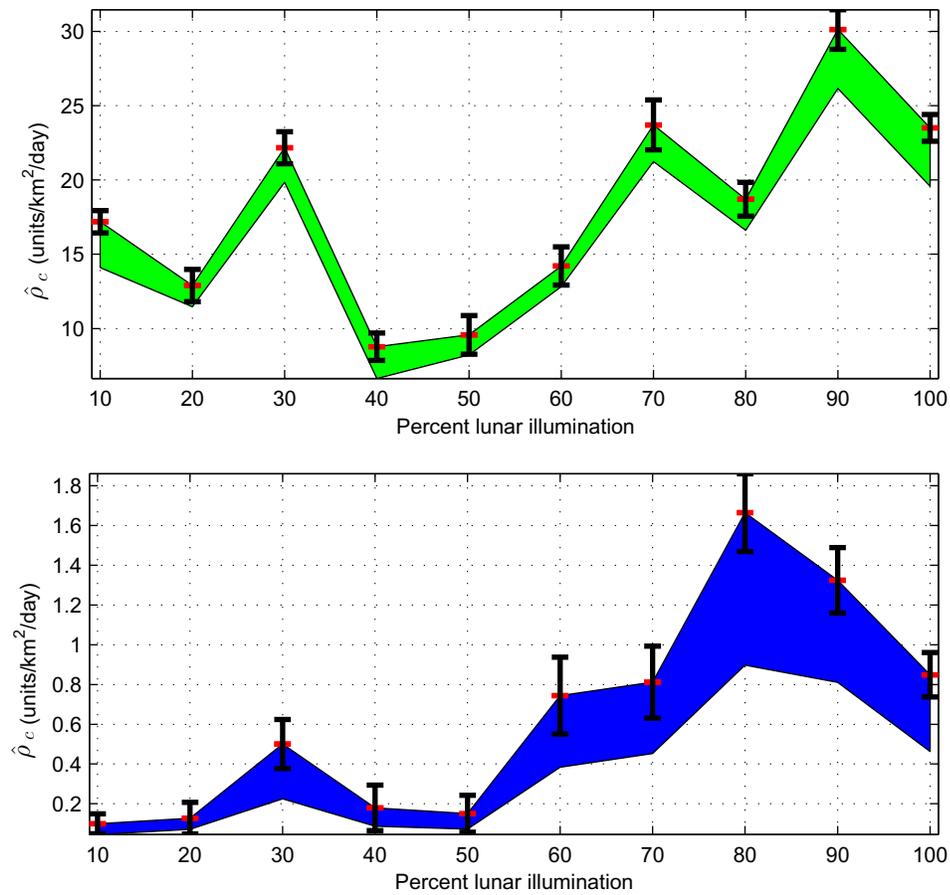


Figure 5.4: Average daily estimated call density, $\hat{\rho}_c$, shown in 10% lunar illumination bins, where units are aggregated over the entire deployment for site SR (upper panel) and site SBC (lower panel). Lunar illumination numbers do not account for cloud cover. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model. Black error bars indicate standard deviation in measurement due to uncertainty in whale distribution around the sensor, red error bars indicate standard deviation in measurement due to uncertainty in noise measurements at the sensor. Note the difference in scale on the vertical axes of the two plots.

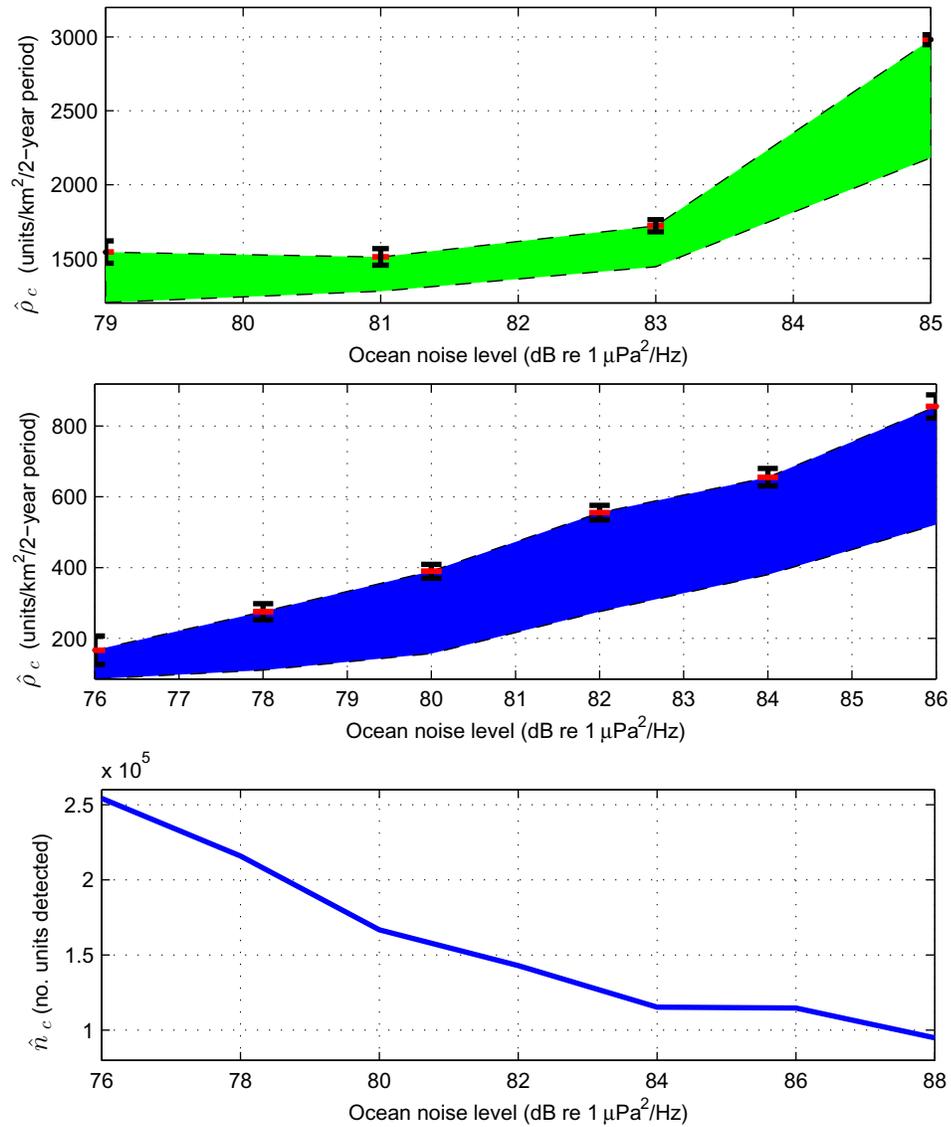


Figure 5.5: Estimated call density, $\hat{\rho}_c$ shown in 2 dB ocean noise bins for full 2-year deployment for site SR (upper panel), and site SBC (middle panel), adjusted for recording effort in each noise band. Numerically-estimated uncorrected call counts, \hat{n}_c , shown for site SBC (lower panel) for all detected calls (1,104,749), adjusted for recording effort in each noise band.

and $\mu(\hat{P})$, i.e.,

$$\mu(\hat{N}_c) = n_c \frac{1 - \mu(\hat{c})}{\mu(\hat{P})} \quad (5.4)$$

The quantities of interest in this section are the biases and the variances about the mean of the estimates \hat{N}_c and $\hat{\rho}_c$, designated as $bias(\hat{N}_c)$ and $bias(\hat{\rho}_c)$, and $var(\hat{N}_c)$ and $var(\hat{\rho}_c)$, respectively. From Eq. 5.2, $var(\hat{\rho}_c) = var(\hat{N}_c)/(AT)^2$. Similarly, $bias(\hat{\rho}_c) = bias(\hat{N}_c)/(AT)$. Therefore, only the statistical properties of \hat{N}_c need to be considered. The coefficient of variation, e.g., $cv(\hat{N}_c)$, is defined as the square root of the variance divided by the mean, $\mu(\hat{N}_c)$.

Eq. 5.3 shows that \hat{N}_c is the ratio of two random variables which represent the probabilities in the detection process. No exact expression for the variance of such a ratio exists. However, an approximate expression for $var(\hat{N}_c)$ can be obtained from the delta method using a Taylor series expansion[20], yielding

$$var(\hat{N}_c) = var\left(n_c \frac{\mu(1 - \hat{c})}{\mu(\hat{P})}\right) \approx n_c^2 \left(\frac{\mu^2(1 - \hat{c})}{\mu^4(\hat{P})} var(\hat{P}) + \frac{1}{\mu^2(\hat{P})} var(1 - \hat{c}) + covterm(\hat{P}, 1 - \hat{c}) \right) \quad (5.5)$$

where the last term involves the covariance between \hat{P} and $1 - \hat{c}$.

In the case considered in this research, human analysts scan the detection outputs generated by an automated detection algorithm to eliminate any detections that are not humpback whale calls. Therefore, the probability of false alarm is zero, $1 - \mu(\hat{c}) = 1$, and the equation above simplifies to

$$var(\hat{N}_c) \approx n_c^2 \left(\frac{1}{\mu^4(\hat{P})} var(\hat{P}) \right) = n_c \left(\frac{1}{\mu^4(\hat{P}_1)} var(\hat{P}_1) \right) \quad (5.6)$$

Note that in Eq. 5.6, \hat{P} actually refers to the probabilities associated with the n_c humpback calls detected within the monitoring area. Designating the corresponding probability for a *single call* as \hat{P}_1 , then $\mu(\hat{P}) = \mu(\hat{P}_1)$, and $var(\hat{P}) = var(\hat{P}_1)/n_c$, assuming that the n_c calls are statistically independent. In this development, this number of uncorrected call detections is taken as a

deterministic quantity equal to the true total number of calls, N_c , normalized by the true environmental calibration factor.

Humpback whales are well known to generate a sequence of units[10]. The calls from an individual animal, if created within a sufficiently short period of time that the position of the animal has not changed significantly, may not be statistically independent. To account for statistical dependence of the calls from the same animal, the number of detected units, n_c , is reduced by a factor of 1,000 in the calculation of the confidence intervals presented in this paper. This reduction accounts for the possibility that a singing humpback whale could remain in the same geographical location for the length of a singing bout, producing 1,000 units from the same location. A more detailed survey on the movement of singing humpback whales in the region would be needed to verify this assumption.

In addition to the locations of the calling animals ($\rho(r, \theta)$, in Eq. (1) of Helble *et al.*[13]), a second quantity modeled as stochastic in nature in the numerical estimation of the probability of detection is the ocean noise. "Noise" in this case is defined as everything other than humpback whale units. The variance of the noise estimate is based on the 6 noise realizations in each 75-sec data record containing a detected humpback unit. In presenting the uncertainties on the corrected call counts and the on density of corrected call counts in this paper, the standard deviation for the noise estimate and the standard deviation for the calling animal locations are reported separately.

As with any parameter estimation problem, the performance of \hat{P} as an estimator of P_d is determined both by its bias, $\mu(\hat{P}) - P_d$, and its variance. As shown through numerical simulation in Helble *et al.*[13], the temporal fluctuations of the environmental properties that affect signal propagation at low frequencies, primarily the fluctuations of the water column sound speed profile, do not significantly affect the variability of \hat{P} except possibly on seasonal time scales. The latter usually can be accounted for by *in situ* measurements or historical oceanographic data at the passive acoustic monitoring site. Therefore, the approach here is to model the propagation of low frequency sounds such as humpback whale calls and other baleen whale vocalizations between a specified

source and receiver location as deterministic (i.e., the spatial detection function, $g(r, \theta)$, in Eq. (1) of Helble *et al.*[13] is deterministic). With this approach, the numerically intensive calculation of the complex acoustic field propagation between a given source/receiver pair only has to be done once.

However, because the relevant environmental properties often are poorly known (e.g., the geoacoustic properties of the ocean bottom), then the signal propagation component is the main source of bias in the estimate of the probability of detection (see the offset of the red, blue, and green curves in Fig. 8 of Helble *et al.*[13]). The Recommendations section later in this paper suggest various approaches to reducing this bias, and reducing the uncertainty in the size of the bias. Note, however, that the bias due to geoacoustic parameter mismatch cannot be eliminated simply by reducing the monitoring area so that only direct-path propagation between source and receiver is considered. The reason is that the detected humpback calls outside the monitoring area can lead to a non-zero probability of false alarm, since any detected unit must be classified as inside or outside the reduced monitoring area. This probability of false alarm must be numerically estimated, in exactly the same way as the probability of detection, so that the source of the bias due to poorly known ocean bottom/subbottom properties simply moves from the denominator to the numerator in Eq. 5.4.

5.3 Results

5.3.1 Monthly and daily calling activity

Fig. 5.1 shows uncorrected call counts (n_c) for site SR and SBC over the 2008 and 2009 calendar years, with corresponding estimated call density plots ($\hat{\rho}_c$) for both locations. The call density plots show three sources of uncertainty. The shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model, the black bars indicate the standard deviation of $\hat{\rho}_c$ due to spatial variability, and the red error bars indicate the standard deviation of $\hat{\rho}_c$ associated with measurements in ocean noise levels.

From the middle and lower panels in Fig. 5.1, the highest density of

humpback vocalizations occur in spring and fall months, with the smallest call densities generally occurring in July and August. Values of n_c appear to be roughly equal between sites SBC and SR during the 2008 season, with increasingly fewer detections at site SBC than SR in 2009. However, because \hat{P} is on average much higher at site SBC than site SR, the corrected call density plots reveal substantially higher call densities at site SR than SBC over the entire deployment, with substantially fewer calls at site SBC in 2009 when compared to 2008. Overall, the average daily call density from April 16, 2008 to Dec 31 2009 was $\hat{\rho}_c = 10.4$ units/km²/day with std = 0.43 at SR and $\hat{\rho}_c = 0.6$ units/km²/day with std = 0.036 at site SBC. The importance of using environmentally corrected call densities as opposed to n_c is further illustrated by comparing n_c at site SR over the full 2-year deployment compared with $\hat{\rho}_c$. The large increase in acoustic detections in the fall of 2009 appears to be a result of the increase in \hat{P} in the area due to a reduction in shipping noise[14]. When this change in shipping noise is taken into account, $\hat{\rho}_c$ in the fall of 2009 appears to be smaller than the $\hat{\rho}_c$ during the fall of 2008.

5.3.2 Call diel patterns

Humpback whales both at site SBC and site SR displayed increased vocalization during nighttime hours, as shown in Fig. 5.2. The plots were created by averaging the call density values in one hour local time bands over the course of the deployments. As in previous plots, the shaded regions indicate the potential bias in the call density estimates due to environmental uncertainty in acoustic model, the black bars indicate the standard deviation of $\hat{\rho}_c$ due to spatial variability, and the red error bars indicate the standard deviation of $\hat{\rho}_c$ associated with measurements in ocean noise levels. At site SBC, the call density increases steadily in the early nighttime hours, peaking at midnight local time, followed by a sudden decrease in vocalizations. At site SR, the call density also increases rapidly with the onsite of nighttime, but the values tend to remain elevated for several hours past midnight.

The ratio of nighttime to daytime calling reaches a peak in the month of April for both locations, with the smallest diel variability in the summer and fall

months. Fig. 5.3 shows $\hat{\rho}_c$ in one hour local time bands during the spring and fall seasons for site SBC. During the spring months, the average nighttime daily call density is $\hat{\rho}_c = 0.333$ calls/km²/hour and the average daytime call density is $\hat{\rho}_c = 0.059$ calls/km²/hour. During the fall, the average call density is $\hat{\rho}_c = 0.077$ calls/km²/hour during nighttime hours and $\hat{\rho}_c = 0.063$ calls/km² during daytime hours, indicating a reduction in overall call density and essentially no diel variation. At site SR, the average springtime call density is $\hat{\rho}_c = 0.5106$ calls/km²/hour during nighttime and $\hat{\rho}_c = 0.1625$ calls/km² during daytime hours. The results for fall also contained a diel pattern, albeit a weak one with an average call density $\hat{\rho}_c = 1.9050$ calls/km²/hour during nighttime hours and $\hat{\rho}_c = 0.9414$ calls/km² during daytime hours.

Because shipping traffic and wind-driven noise also occur irregularly throughout a 24 hour period, it is important to compare values of $\hat{\rho}_c$ as opposed to n_c . For example, in the May timeframe at site SBC, values of n_c show a strong diel pattern, but this pattern is significantly reduced when values of $\hat{\rho}_c$ are used. The reduced shipping noise at night increases the probability of detection during nighttime hours, which in turn increases the values of n_c during nighttime hours[14].

5.3.3 Call density and lunar illumination

Both site SBC and site SR exhibited an increase in $\hat{\rho}_c$ with increasing lunar illumination, as shown in Fig. 5.4. Because the majority of humpback vocalizations occur during a relatively narrow time window of migration (1-2 months in the spring and fall), it is possible that the whales coincidentally happen to be vocalizing in the region during periods with greater illumination. Thus, a longer time series would provide more statistically significant results.

5.3.4 Call density and ocean noise

Both site SBC and site SR exhibited an increase in $\hat{\rho}_c$ with increasing ocean noise, as shown in the upper and middle panel of Fig. 5.5. The figures were

created by aggregating call densities in 2 dB ocean noise bands over the full 2-year deployment at each site. The value in each noise band represents the estimated call densities for the entire deployment, which were calculated using the number of calls, n_c . The appropriate values of \hat{P} for the ocean noise and environmental conditions, and values corrected for sensor recording effort. The results show a 100% increase in $\hat{\rho}_c$ over the observed 6 dB noise band at site SR, and a 300% increase in $\hat{\rho}_c$ for site SBC over the 10 dB observed noise range. The acoustic model used to estimate \hat{P} assumes a constant humpback source level of 160 dB rms re 1 μ Pa @ 1 m. If the mean source level increases in strength with increasing noise, the result would manifest itself as an increase in $\hat{\rho}_c$ using the current modeling methods. Therefore, it is impossible to distinguish whether humpbacks increase the number of vocalizations, the source level, or a combination of the two with increasing ocean noise. If the humpback call densities remain constant throughout varying ocean noise conditions, the source level would need to increase by approximately 0.35 dB per 1 dB increase in ocean noise at site SBC in order to achieve the slope shown in Fig. 5.5. This value was obtained by creating a linear fit to the best estimate values shown for site SBC in Fig. 5.5, and then increasing the source level in the model until the slope in the model best matched the slope in the observable data. The lower panel in Fig. 5.5 shows values of \hat{n}_c with increasing noise. Even though the call counts are uncorrected for probability of detection, the \hat{n}_c is used on \hat{n}_c because the values are estimated by tallying the actual call counts, n_c , and dividing by the acoustic recording effort for that noise band. As expected, fewer calls are detected as ocean noise increases. If humpback whales increased their source levels to completely compensate for increasing ocean noise conditions, the plot would exhibit zero slope.

5.4 Discussion

5.4.1 Seasonal comparison

Values of $\hat{\rho}_c$ in Fig. 5.1 indicate increased call density during fall and spring months, with reduced densities in the winter months and very low densities in

the summer months. This pattern is consistent with the notion that the vocalizing whales that make up the majority of the acoustic detections are migrating between summer feeding grounds north of site SBC and site SR (presumably off the northern N. American coast and Gulf of Alaska), and wintering grounds south of site SBC and site SR (presumably in coastal Mexico and Central American waters). Aerial and visual line transect surveys indicate a year-round presence of humpback whales at both site SBC and site SR, although these studies included periods of peak humpback migration in the fall and spring for seasons classified as "winter" and "summer"[21]. In some cases, visual sightings increase in the summer, although observation effort also tends to increase in the summer months[22]. Visual surveys publish results in terms of animal densities, whereas the results published in this paper describe acoustic call densities. The two numbers are therefore not directly comparable, since the acoustic cue rate of humpback whales can be highly variable. The discrepancies between visual surveys and acoustic surveys may be due to vocalizing whales switching from chorusing song behavior during fall, winter, and spring months, to acoustic feeding behavior in the summer. The latter period contains much less vocal activity. However, it is possible that some of the discrepancy between visual and acoustic patterns over seasons is a result of two separate humpback groups inhabiting the region - a transiting vocal group that occupies site SBC and SR during migration months, and a more resident (less vocal) group that uses areas near site SBC and site SR as summer feeding grounds, perhaps migrating to a different wintering destination than the group transiting through the two sites. It is important to note that visual observation methods also can contain significant bias in population estimates, particularly when the behavior of the whale changes over time in a way which alters the visual probability of detecting the animals. Research shows that singing humpbacks are more difficult to see than their non-singing counterparts[23], and it is possible that summer feeding behavior may further increase the probability of visual detections in summer months.

The reduced values of $\hat{\rho}_c$ at site SBC compared to site SR could indicate that fewer migratory whales pass through the Santa Barbara Channel than near

Sur Ridge, if the vocal activity is otherwise similar at the two sites. The Santa Barbara Channel is off the direct path of coastal Pacific migration routes[7], and so deviating into the channel would require additional time and energy during the migration season. Possibly, the Santa Barbara Channel provides a social purpose for the migrating populations, and/or an opportunistic food source. The large values of $\hat{\rho}_c$ during the 2008 season compared with the 2009 season could be an indication that humpback whales selectively move into this region for opportunistic feeding. For example, recent studies indicate that humpback whales in the region could switch prey between a euphausiid-based diet and a forage fish-based diet on annual time scales[24]. Additionally, visual humpback whale density estimates in the same regions as sites SBC and SR showed a decline in numbers following a particularly harsh El Nino season in 1997-98, when zooplankton declines were severe[22]. Therefore, it is possible that acoustic call densities could be a proxy for prey availability in the region. A longer time series with ancillary simultaneous data collection on prey distribution would be necessary to confirm this relationship.

An additional explanation for the reduced calling activity at site SBC in 2009 compared with 2008 could be attributed to the relationship between vocal activity and ocean noise. Because of the faltering world economy and the enforcement of environmental regulations, the shipping noise was significantly reduced in 2009 compared to 2008 at both locations. If the humpbacks reduced their source levels and/or cue rate in response to a decrease in ocean noise, the estimated values of $\hat{\rho}_c$ would drop, even if the population of vocalizing humpback whales was approximately equal from year to year. One indication that the reduction in $\hat{\rho}_c$ at the site SBC may not be a response to dropping ocean noise levels is that values of $\hat{\rho}_c$ are relatively stable between the two years at site SR, despite an overall reduction in ocean noise in the second season at site SR.

The monthly pattern of $\hat{\rho}_c$ at sites SBC and SR are consistent with vocal activity recorded along other migration routes worldwide[25, 26, 27]. A two-year study of humpback whales in deep waters off the British Isles showed the highest acoustic detection densities in the Oct-Nov, with a reduction during December, and an increase in detections mid Jan-Mar[28]. Song was not present

during the summer months at the locations monitored during the study. Due to equipment error, data from the months of April and May were absent, and so it was not possible to compare the reduction of song chorusing during these months to site SBC and site SR. Because this study involved the use of arrays, directionality could be estimated with each humpback song. A southern migration trend was recorded during fall months, but a return directionality was not present with vocalizations occurring in the spring - either indicating a summer resident population or opportunistic feeding in the area, perhaps combined with stock returning north on a migration route outside the range of the monitored area. The ability to localize humpback whales at site SBC and site SR would provide similar detail to the records reported in the British Isles, perhaps shedding light on the significance of summer resident populations at these two locations.

5.4.2 Diel comparison

The diel variability found at site SBC and site SR is similar to trends reported at several wintering grounds in the Pacific Ocean. Au et al.[29] showed an increase in recorded sound pressure level for humpback vocalizations in the Hawaiian wintering grounds during nighttime hours over the period of March 5-21, 1998. A peak in average sound pressure level occurred at midnight in the monitored frequency band, similar to the observed peak in vocalizations at both site SBC and site SR during the April 7 - May 27 period, shown for site SBC in the upper panel of Fig. 5.3. Recordings on the same wintering grounds during the period of January 7-12, 1998 showed a weaker opposing trend, with peak vocalizations occurring during noontime. These results are similar to those observed at site SBC and site SR during the Oct 15 - Dec 4 timeframe, which show much weaker diel variability, with the peak in vocalizations occurring at 10 am local time for site SBC (shown in the lower panel of Fig. 5.3). The observed time periods for weakest and strongest diel variability at site SBC and site SR are notably earlier in the fall and later in the spring, corresponding to the lag in transit time as the whales migrate to/from the wintering grounds. The possibility that these patterns begin before the whales arrive on wintering grounds and are sustained after the whales have

left could indicate a social function that is also relevant during migration. A study on migrating whales using the long-range underwater Sound Surveillance System (SOSUS) on the migration route between Alaskan waters and Hawaii showed that the calling rate doubled during nighttime hours in the months of April and May, a notably weaker imbalance than the quadrupling between night and day observed at site SBC. The SOSUS nighttime calling pattern is very similar to site SBC, with a rapid reduction in number of humpback detections after midnight[30].

The diel variability in humpback vocalizations appears to be site-dependent, with some locations following similar trends as site SBC and site SR while other locations reveal little diel variability or increase vocalizations during daylight hours in spring. Vocalization activity in northern Angola, for example, is reported to peak at 5 am, with depressed singing around 5 pm[31]. Two locations were observed in the American Samoa, song at the Rose Atoll indicated increased calling during nighttime hours while there was no observed diel pattern at the Tutuila location. It is important to note that very little, if any, information has been reported on the probability of detection during these studies, and so changes in ocean noise could easily influence the perceived diel patterns of humpback vocalizations, as demonstrated at both site SBC and site SR[14].

Because humpback whales exhibit diel calling patterns on wintering grounds, where feeding does not occur, it is probable that the matching diel patterns found along the migration route serve a similar social function, rather than being associated with prey availability. However, it is possible that these patterns are influenced by the availability of food. The California coast is a biological productive region, and humpbacks have been observed feeding in the Santa Barbara channel, presumably on fish in the northern portion of the channel and krill in the southern channel[32, 22]. Recent acoustic tagging efforts on an Antarctic feeding ground showed song occurring during periods of active diving and feeding lunges, although it is unclear if the whales preferentially sing more often during periods of inactive feeding[33]. Researches also have recently found strong diel changes in humpback whale feeding behavior in response to changes in prey behavior and distribution on Stellwagen Bank, MA[34]. The differences in peak

vocalizing hours between site SBC and site SR could therefore be an indication of one or more factors - prey availability, differences in humpback stock at the two sites, or site specific behavior differences. Because changes in the probability of detection have been accounted for, changes in background noise as being the cause for diel differences between the two sites can be eliminated from consideration.

5.4.3 Calling behavior and ocean noise

The influence of ocean noise on marine mammals is an active ongoing area of research. Part of this research includes studying the influence of both shipping noise and active sonar systems on marine mammals, particularly on odontocetes. Beaked whales have been shown to be sensitive to active sonar systems, resulting in several mass stranding events[35, 36]. Changes in vocalization behavior, surfacing patterns, call length and intensity, and foraging behaviors all have been shown to change in the presence ships and/or active sonar[37, 38, 39, 40, 41, 42, 43]. The Lombard effect[44] is the tendency for speakers to increase their vocal effort as background noise increases in order to enhance their communication. This phenomenon has been reported for a variety of marine mammals, including killer whales (*Orcinus orca*), Beluga whales (*Delphinapterus leucas*), Pilot whales (*Globicephala Melas*), and bottle nose dolphins (*Tursiops truncatus*)[40, 45, 46]. Blue whales also have been found to both increase the source level and length of their vocalizations in response to shipping noise, which has been shown to be true in the Santa Barbara channel at the same hydrophone location as site SBC[47].

Humpback whales have also been shown to respond to ocean noise and sonar. During low-frequency active (LFA) sonar activity, it was shown that humpback whales lengthen the duration of song by 29%, with longer than average themes present within a normal song structure[37]. The lengthening of song could result in more overall emitted humpback units per time, one possible explanation for the overall increase in estimated units with increasing noise observed at site SBC and site SR. More recently, research has shown that humpback whales migrating off the coast of eastern Australia increase their calling source level by 0.75 dB per 1 dB increase in background noise[48]. In this study, the background noise was

much lower than the vocal level, and so the observed result of 0.35 dB per 1 dB increase in background noise observed in the Santa Barbara channel (a notably higher noise environment) may be due to the physical constraints of the whales to produce louder sounds. Humpback whales also have been noted to change communication methods from vocal sounds to surface-generated signals such as 'breaching' or 'pectoral slapping' with increasing wind speeds and background noise levels, although this study was conducted primarily during social sound behavior, and was not tested during song chorusing[49]. Other studies have shown that humpback whales respond to the presence of ships by increasing swim speed away from the vessel, or occasionally charging vessels and even screaming underwater[50, 51, 52]. Additionally, respirations rates, social exchanges, and aerial behaviors all have been shown to be positively correlated with vessel numbers, speed and direction changes, and proximity to the whales[50]. All these factors suggest that changes in vocal behavior in the presence of shipping noise are more probable than possible, and are supported by the results in this paper.

5.4.4 Population density estimates for humpback whales using single-fixed sensors

Estimating the density of marine mammals using acoustic cues as described in Eq. (5.1) for single fixed sensors is a complicated procedure. Estimating the probability of detection (\hat{P}) has been shown to be site and time specific in previous works[13, 14], with \hat{P} varying by factors greater than 10 between sensors and at the same sensor over time. Estimating P with reasonable uncertainty is possible under certain conditions, but the procedure requires considerable knowledge about the environmental properties, such as bathymetry, bottom type composition, sound speed profile, and ocean noise conditions. Estimating the cue rate, \hat{r} , for humpbacks, particularly during migration could be an even more challenging proposition. It has been established that the cue rate for humpback whales changes over seasons, as the number of units produced by humpbacks is much higher during song chorusing than during feeding and social calling[12]. Therefore, establishing a time-dependent cue rate in a particular area over all seasons is

vitaly important. Additionally, research from this paper suggests that cue rate could change substantially based on diel patterns, lunar illumination, and ocean background noise, among other variables. Diel patterns are perhaps easier to account for, especially if a cue rate is desired on time scales long enough to include an average of both night and day. Ocean noise could be particularly problematic, as the cue rate and/or average source level of humpback units appear to change appreciably with changing background noise. Therefore, a cue rate and source level would need to be established not only over season for a particular location, but also for different background noise levels in a given frequency band. Obtaining values will be difficult, a procedure that might be accomplished through tagging animals or deploying a localizing array system that could track a particular whale's vocalizations over a period of time. In both scenarios, data would need to be collected over long periods of time in order to obtain useful cue rates. Given the present state of the technology, the best approach is to deploy passive monitoring systems with localizing capability. Doing so would help estimate cue rate and \hat{P} , allowing for more accurate density estimates than single-fixed sensors.

Acknowledgements

The authors are extremely grateful to Prof. Glenn Ierley, Dr. Megan McKenna, and Amanda Debich, both at the Scripps Institution of Oceanography, for their support of this research. Special thanks to Sean Wiggins and the entire Scripps Whale Acoustics Laboratory for providing thousands of hours of high quality acoustic recordings. The first author would like to thank the Department of Defense Science, Mathematics, and Research for Transformation (SMART) Scholarship program, the Space and Naval Warfare (SPAWAR) Systems Command Center Pacific In-House Laboratory Independent Research program, and Rich Arrieta from the SPAWAR Unmanned Maritime Vehicles Lab for continued technical and financial support. Work was also supported by the Office of Naval Research, Code 322 (MBB), the Chief of Naval Operations N45, and the Naval Postgraduate School.

Chapter 5 is a manuscript in preparation for submission to *The Journal of the Acoustical Society of America*: Tyler A. Helble, Gerald L. D'Spain, Greg S. Campbell, and John A. Hildebrand, "Humpback whale vocalization activity at Sur Ridge and in the Santa Barbara Channel from 2008-2009, using environmentally corrected call counts". The dissertation author was the primary investigator and author of this paper.

References

- [1] J.H. Johnson and A.A. Wolman. The humpback whale, *Megaptera novaeangliae*. *Marine Fisheries Review*, 46(4):30–37, 1984.
- [2] J. Barlow. The abundance of cetaceans in California waters. Part I: Ship surveys in summer and fall of 1991. *Fishery Bulletin*, 93:1–14, 1995.
- [3] C.S. Baker, L. Medrano-Gonzalez, J. Calambokidis, A. Perry, F. Pichler, H. Rosenbaum, J.M. Straley, J. Urban-Ramirez, M. Yamaguchi, and O. von Ziegeler. Population structure of nuclear and mitochondrial DNA variation among humpback whales in the North Pacific. *Molecular Ecology*, 7(6):695–707, 1998.
- [4] C.S. Baker, D. Steel, J. Calambokidis, J. Barlow, A.M. Burdin, P.J. Clapham, E. Falcone, J.K.B. Ford, C.M. Gabriele, U. Gozález-Peral, R. LeDuc, D. Mattila, T.J. Quinn, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, R.J. Urban, M. Vant, P.R. Wade, D. Weller, B.H. Witteveen, K. Wynne, and M. Yamaguchi. geneSPLASH: An initial, ocean-wide survey of mitochondrial (mt) DNA diversity and population structure among humpback whales in the North Pacific: Final report for contract 2006-0093-008 Principal Investigator: C. Scott Baker. Technical report, Cascadia Research Collective, Olympia, WA, 2008.
- [5] J. Calambokidis, G.H. Steiger, K. Rasmussen, J. Urban, KC Balcomb, PL de Guevara, M. Salinas, JK Jacobsen, CS Baker, LM Herman, S. Cerchio, and JD Darling. Migratory destinations of humpback whales that feed off California, Oregon and Washington. *Marine Ecology-Progress Series.*, 192:295–304, 2000.
- [6] J. Calambokidis, G.H. Steiger, J.M. Straley, L.M. Herman, S. Cerchio, D.R. Salden, U.R. Jorge, J.K. Jacobsen, O. von Ziegeler, K.C. Balcomb, C.M. Gabriele, M.E. Dahlheim, S. Uchida, G. Ellis, Y. Miyamura, P.L.P. de Guevara, M. Yamaguchi, F. Sato, S.A. Mizroch, L. Schlender, K. Rasmussen, J. Barlow, and T.J. Quinn. Movements and population

- structure of humpback whales in the North Pacific. *Marine Mammal Science*, 17(4):769–794, 2001.
- [7] J. Calambokidis, E.A. Falcone, T.J. Quinn, A.M. Burdin, P.J. Clapham, J.K.B. Ford, C.M. Gabriele, R. LeDuc, D. Mattila, L. Rojas-Bracho, J.M. Straley, B.L. Taylor, J.R. Urban, D. Weller, B.H. Witteveen, M. Yamaguchi, A. Bendlin, D. Camacho, K. Flynn, A. Havron, J. Huggins, and N. Maloney. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Technical report, Cascadia Research Collective, Olympia, WA, 2008.
- [8] S. Wiggins. Autonomous Acoustic Recording Packages (ARPs) for long-term monitoring of whale sounds. *Marine Tech. Soc. J.*, 37(2):13–22, 2003.
- [9] T.A. Helble, G.R. Ierley, G.L. D’Spain, M.A. Roch, and J.A. Hildebrand. A generalized power-law detection algorithm for humpback whale vocalizations. *J. Acoust. Soc. Am.*, 131(4):2682–2699, 2012.
- [10] R.S. Payne and S. McVay. Songs of humpback whales. *Science*, 173(3997):585–597, 1971.
- [11] R.A. Dunlop, M.J. Noad, D.H. Cato, and D. Stokes. The social vocalization repertoire of east Australian migrating humpback whales (*Megaptera novaeangliae*). *J. Acoust. Soc. Am.*, 122:2893–2905, 2007.
- [12] R.A. Dunlop, D.H. Cato, and M.J. Noad. Non-song acoustic communication in migrating humpback whales (*Megaptera novaeangliae*). *Marine Mammal Science*, 24(3):613–629, 2008.
- [13] T.A. Helble, G.L. D’Spain, J.A. Hildebrand, G.S. Campbell, R.L. Campbell, and K.D. Heaney. Site specific probability of passive acoustic detection of humpback whale calls from single fixed hydrophones. *J. Acoust. Soc. Am.*, *accepted for publ.*, 2013.
- [14] T.A. Helble, G.L. D’Spain, G.S. Campbell, and J. A. Hildebrand. Calibrating passive acoustic monitoring: Correcting humpbacks call detections for site-specific and time-dependent environmental characteristics. *J. Acoust. Soc. Am. Express Letters*, *submitted for publ.*, 5 pgs. plus 3 figs., 2012.
- [15] S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, and L. Thomas. *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, pages 1–448. Oxford University Press, New York, NY, 2001.
- [16] E.T. Küsel, D.K. Mellinger, L. Thomas, T.A. Marques, D. Moretti, and J. Ward. Cetacean population density estimation from single fixed sensors using passive acoustics. *J. Acoust. Soc. Am.*, 129(6):3610–3622, 2011.

- [17] T.A. Marques, L. Munger, L. Thomas, S. Wiggins, and J.A. Hildebrand. Estimating North Pacific right whale *Eubalaena japonica* density using passive acoustic cue counting. *Endangered Species Research*, 13:163–172, 2011.
- [18] M.A. McDonald and C.G. Fox. Passive acoustic methods applied to fin whale population density estimation. *J. Acoust. Soc. Am.*, 105(5):2643–2651, 1999.
- [19] T.A. Marques, L. Thomas, J. Ward, N. DiMarzio, and P.L. Tyack. Estimating cetacean population density using fixed passive acoustic sensors: An example with Blainville’s beaked whales. *J. Acoust. Soc. Am.*, 125(4):1982–1994, 2009.
- [20] H. Cramér. *Mathematical Methods of Statistics*, page 353. Princeton University Press, Princeton, NJ, 1946.
- [21] K.A. Forney and J. Barlow. Seasonal patterns in the abundance and distribution of california cetaceans, 1991–1992. *Marine Mammal Science*, 14(3):460–489, 2006.
- [22] J. Calambokidis, T. Chandler, L. Schlender, K. Rasmussen, and GH Steiger. Research on humpback and blue whales off California, Oregon, and Washington in 2000. *Final Contract Report to Southwest Fisheries Science Center, National Marine Fisheries Service, PO Box, 271*, 2003.
- [23] M. Noad, D. Cato, et al. Swimming speeds of singing and non-singing humpback whales during migration. *Marine Mammal Science*, 23(3):481–495, 2007.
- [24] A.H. Fleming, J. Barlow, and J. Calambokidis. Probable prey switching in humpback whales with implications for population structure. In *Proceedings-19th Biennial Conference on the Biology of Marine Mammals*, page 89, Tampa, FL, 2011.
- [25] T.F. Norris, M. McDonald, and J. Barlow. Acoustic detections of singing humpback whales (*Megaptera novaeangliae*) in the eastern North Pacific during their northbound migration. *J. Acoust. Soc. Am.*, 106:506, 1999.
- [26] P.J. Clapham and D.K. Mattila. Humpback whale songs as indicators of migration routes. *Marine Mammal Science*, 6(2):155–160, 1990.
- [27] D.H. Cato. Songs of humpback whales: the Australian perspective. Technical report, DTIC Document, 1991.
- [28] R.A. Charif, P.J. Clapham, and C.W. Clark. Acoustic detections of singing humpback whales in deep waters off the British Isles. *Marine Mammal Science*, 17(4):751–768, 2006.

- [29] W.W.L. Au, J. Mobley, W.C. Burgess, M.O. Lammers, and P.E. Nachtigall. Seasonal and diurnal trends of chorusing humpback whales wintering in waters off Western Maui. *Marine mammal science*, 16(3):530–544, 2000.
- [30] R. Abileah, D. Martin, S.D. Lewis, and B. Gisiner. Long-range acoustic detection and tracking of the humpback whale Hawaii-Alaska migration. In *OCEANS'96. MTS/IEEE. Prospects for the 21st Century. Conference Proceedings*, volume 1, pages 373–377. IEEE, 1996.
- [31] K. Rasmussen, D.M. Palacios, J. Calambokidis, M.T. Saborío, L. Dalla Rosa, E.R. Secchi, G.H. Steiger, J.M. Allen, and G.S. Stone. Southern Hemisphere humpback whales wintering off Central America: insights from water temperature into the longest mammalian migration. *Biology Letters*, 3(3):302–305, 2007.
- [32] John Calambokidis. (personal communication), 2012.
- [33] A.K. Stimpert, L.E. Peavey, A.S. Friedlaender, and D.P. Nowacek. Humpback whale song and foraging behavior on an Antarctic feeding ground. *PloS ONE*, 7(12):e51214, 2012.
- [34] A.S. Friedlaender, EL Hazen, DP Nowacek, PN Halpin, C. Ware, MT Weinrich, T. Hurst, and D. Wiley. Diel changes in humpback whale *Megaptera novaeangliae* feeding behavior in response to sand lance *ammodytes* spp. behavior and distribution. *Mar Ecol Prog Ser*, 395:91–100, 2009.
- [35] A. D'Amico, R.C. Gisiner, D.R. Ketten, J.A. Hammock, C. Johnson, P.L. Tyack, and J. Mead. Beaked whale strandings and naval exercises. Technical report, DTIC Document, 2009.
- [36] A. Fernández, JF Edwards, F. Rodriguez, A.E. De Los Monteros, P. Herraiez, P. Castro, JR Jaber, V. Martin, and M. Arbelo. Gas and fat embolic syndrome involving a mass stranding of beaked whales (family Ziphiidae) exposed to anthropogenic sonar signals. *Veterinary Pathology Online*, 42(4):446–457, 2005.
- [37] P.J.O. Miller, N. Biassoni, A. Samuels, P.L. Tyack, et al. Whale songs lengthen in response to sonar. *Nature*, 405(6789):903, 2000.
- [38] W.J. Richardson, C.R. Greene, C.I. Malme, and D.H. Thomson. *Marine Mammals and Noise*. Academic Press, 1998.
- [39] F.H. Jensen, L. Bejder, M. Wahlberg, N. Aguilar Soto, and PT Madsen. Vessel noise effects on delphinid communication. *Marine Ecology Progress Series*, 395:161–175, 2009.

- [40] M.M. Holt, D.P. Noren, V. Veirs, C.K. Emmons, and S. Veirs. Speaking up: killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. *J. Acoust. Soc. Am.*, 125(1):EL27–EL32, 2008.
- [41] M. Jahoda, C.L. Lafortuna, N. Biassoni, C. Almirante, A. Azzellino, S. Panigada, M. Zanardelli, and G.N. Sciara. Mediterranean fin whale's (*Balaenoptera physalus*) response to small vessels and biopsy sampling assessed through passive tracking and timing of respiration. *Marine Mammal Science*, 19(1):96–110, 2003.
- [42] B.M. Siemers and A. Schaub. Hunting at the highway: traffic noise reduces foraging efficiency in acoustic predators. *Proceedings of the Royal Society B: Biological Sciences*, 278(1712):1646–1652, 2011.
- [43] V.M. Janik and P.M. Thomspon. Changes in surfacing patterns of bottlenose dolphins in response to boat traffic. *Marine Mammal Science*, 12(4):597–602, 1996.
- [44] E. Lombard. Le signe de lelevation de la voix. annales de maladies de loreille et du larynx. *Larynx*, 37:101–119, 1911.
- [45] P.M. Scheifele, S. Andrew, R.A. Cooper, M. Darre, F.E. Musiek, and L. Max. Indication of a Lombard vocal response in the St. Lawrence River beluga. *J. Acoust. Soc. Am.*, 117:1486, 2005.
- [46] K.C. Buckstaff. Effects of watercraft noise on the acoustic behavior of bottlenose dolphins, *Tursiops truncatus*, in Sarasota Bay, Florida. *Marine Mammal Science*, 20(4):709–725, 2006.
- [47] M.F. McKenna. Blue whale response to underwater noise from commercial ships, 2011.
- [48] M. Noad, R. Dunlop, and D. Cato. The Lombard effect in humpback whales. *J. Acoust. Soc. Am.*, 131(4):3456, 2012.
- [49] R.A. Dunlop, D.H. Cato, and M.J. Noad. Your attention please: increasing ambient noise levels elicits a change in communication behaviour in humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B: Biological Sciences*, 277(1693):2521–2529, 2010.
- [50] G.B. Bauer and L.M. Herman. Effects of vessel traffic on the behaviour of humpback whales in Hawaii. rep. from Kewalo Basin Mar. Mamm. Lab., Univ. Hawaii, Honolulu, for US Natl. *Mar. Fish. Serv., Honolulu, HI*, 1986.
- [51] W.W.L. Au and M. Green. Acoustic interaction of humpback whales and whale-watching boats. *Marine Environmental Research*, 49(5):469–481, 2000.

- [52] M. Scheidat, C. Castro, J. Gonzalez, and R. Williams. Behavioural responses of humpback whales (*Megaptera novaeangliae*) to whalewatching boats near Isla de la Plata, Machalilla National Park, Ecuador. *Journal of Cetacean Research and Management*, 6(1):63–68, 2004.

Chapter 6

Conclusions and Future Work

The process outlined in this thesis has shown that with a few assumptions, it is possible to use call densities from properly calibrated single, fixed omnidirectional sensors with non-overlapping coverage to reveal substantial biological and ecological information about transiting humpback whales off the coast of California. At the onset of this project, the magnitude of the uncertainties associated with environmental conditions and whale distributions surrounding each recording site were unknown. For the Hoke seamount location, the acoustic model was insufficient for predicting the probability of detection at the seamount, thus preventing the calculation of accurate call densities. The poor model/data fit for Hoke seamount was either due to a highly non-uniform whale distribution about the sensor, or due to humpback vocalizations entering the deep sound channel from distances beyond the model boundaries. However, for the recording locations in the Santa Barbara Channel and at Sur Ridge, excellent agreement occurs between the theoretical distribution of received whale call levels and the actual observed whale call levels, as demonstrated in Ch. 3. Distinctly significant statistical differences in call densities were found when comparing densities between the two locations, or at the same location over time despite the uncertainty associated with measurements in ocean noise levels, environmental, and bathymetric features at these two locations. These differences, such as substantially higher vocalization densities at the Sur Ridge location compared to the Santa Barbara location, would not be possible to distinguish without the use of the GPL detector and properly

calibrated sensors. Additionally, it would not have been possible to measure the observed Lombard effect in humpback whale vocalizations at both locations, which has important implications for conservation efforts of this endangered species.

6.1 Improving animal density estimates from passive acoustics

Uncertainties in animal distribution, cue rate, and environmental properties surrounding each single, fixed omnidirectional sensor remain problematic for conducting accurate density estimates of marine mammals using these sensors with non-overlapping coverage. Reducing environmental uncertainty can be a costly process, requiring additional bottom-type samples or acoustic surveys in the areas surrounding the sensor. Determining marine mammal cue rates also could prove to be a laborious and costly process, because the cue rate can change over season, geographical location, and varying environmental conditions, as demonstrated in Ch. 5. Obtaining the cue rate over this vast variable space would require constant surveillance over a wide range of ocean noise and environmental conditions, and would require either tagging animals with acoustic devices or using multi-hydrophone acoustic arrays with localization capabilities. The spatial distribution of animals in a particular area throughout differing seasons also could be obtained using the same technique. For the uncertainty estimates in Ch. 5, the distribution of humpback calls was assumed to be random and uniformly distributed in the region surrounding the sensor. Because the sensor is omnidirectional and the detection function in many cases has near azimuthal symmetry, the assumption of uniform distribution of animals as a function of distance from the sensor is more crucial than uniform distribution as a function of bearing. For sites SBC and SR, it was shown in Ch. 3 using model/data comparison that modeled predictions based on this assumed distribution matched the observable data. However, conducting additional simulations would provide uncertainty estimates for scenarios with non-uniform animal distribution. Uncertainty estimates could be established for differing whale behaviors, such as clustering in a particular region or for whales

transiting through the region with differing paths. Because of the challenges associated with uncertainties in animal distribution, cue rate, and environmental properties, it may often be more efficient to deploy multi-hydrophone systems with localization capabilities, rather than spending the effort to calibrate single, fixed omnidirectional sensors.

While multi-hydrophone systems have advantages over single, fixed omnidirectional sensors, calculating accurate density estimates from these configurations also remains difficult. The difficulties arise in part from obtaining cue rates using localizing systems. In some cases, localizing arrays can track individual animals over periods of time to obtain cue rates (and even animal density estimates), but in other cases irregular calling rates or animals grouped too closely to one another inhibit this process. Additionally, in order to use localizing systems for accurate animal density estimates, a distance perimeter must be chosen surrounding the sensor system in which the system can accurately detect and localize calls in all noise conditions (particularly if there is interest in researching the impact of noise on the species). Often, this perimeter may be only a few kilometers from the array, limiting the monitoring capability of that system. The acoustic modeling process described in this thesis could help determine the probability of detection beyond this perimeter, enabling detections at greater distances to be scaled appropriately and included in the density estimation.

Using passive acoustics for marine mammal density estimates introduces several additional challenges when compared to visual sighting techniques. The detection function, which is required for nearly all density estimation work, is calculated more easily using visual sighting methods. Some of the main variables that affect the visual detection function are height of the observer from the sea-surface interface, daylight brightness, and sea-state. In general, the probability of detecting a marine mammal decreases monotonically with increasing distance to the animals, and stays stable over fairly long observation periods. The same simple assumptions are not true using passive acoustic monitoring; the importance of these differences can not be overstated. Research throughout this thesis illustrates that the detection function for passive acoustic sensors is in a state of constant flux,

with the probability of detecting an animal changing by factors of 10 or more, even on short time scales. Additionally, because of the complex interaction of sound with the environment and bathymetry, the probability of detection cannot be assumed to decrease monotonically with range, especially for mid and low-frequency calling animals. The probability of detection maps generated for the Santa Barbara location in Ch. 3 demonstrate a highly variable detection function with range. An oversimplification of the detection function for passive acoustic sensing currently appears in many peer-reviewed publications.

Because the field of passive acoustics for marine mammal density estimates is still in its infancy, more research is needed to determine the best procedural methods for obtaining accurate density estimates. Many techniques used in visual sighting methods may not be appropriate for passive acoustic systems. In order to develop the most accurate monitoring systems, a controlled experiment should be conducted that utilizes acoustic surveys using a variety of techniques. As part of the controlled experiment, it would be useful to obtain density estimates using a combination of acoustic arrays, overlapping sensors, and single, fixed omnidirectional sensors. Additionally, bathymetric and environmental information should be utilized to attempt to increase the accuracy of the density estimates, as properly calibrating for the environment could also provide benefits to multi-hydrophone systems. As part of this effort, it would be helpful to use a combination of controlled acoustic sources, computer simulated sources, and opportunistic marine mammal sources.

In addition to fixed passive systems, using passive acoustic equipped autonomous underwater vehicles (AUVs) for line-transect methods could become crucial for accurate density estimation. Surveys could be conducted on a near continuous basis at a much lower cost than ship or aircraft-based surveys. Additionally, these platforms would be difficult for the marine mammals to detect from a distance, helping to reinforce the key assumption in line-transect surveys that monitored animals do not react to the observation platform before they are counted. Another advantage is that AUVs have the capability to carry payloads that can simultaneously measure a wide range of environmental and oceanographic

data, some of which are difficult to obtain from fixed stations or from surface vessels. Because autonomous platforms generally travel at lower speeds than ships and air-craft, some modification to the line-transect method may need to be implemented. Nevertheless, initial research indicates autonomous platforms will become a key tool for passive acoustic monitoring. Although not discussed in this thesis, the GPL algorithms were adapted for use on AUVs, discussed in more detail in Sect. 6.3.

6.2 Improvements to studying migrating humpback whales in coastal California

Additional work could be carried forward that would significantly enhance the biological and ecological results for humpback whales presented in this thesis. In addition to enhancements in density estimation previously discussed, the most obvious work would be to repeat the same process of calculating acoustic call densities at many more hydrophone locations throughout the southern California Bight over many more years. Doing so would allow for a more detailed picture on the biology and ecology of humpback whales in the region. Additionally, calculating humpback call densities over longer time scales would better facilitate habitat modeling, perhaps leading to the discovery of relationships between these densities and prey availability in the region. As mentioned previously, in order to limit uncertainties in calling densities caused by unknown environmental properties, it would be beneficial to retrieve additional sediment core samples and/or conduct geoacoustic surveys in the areas surrounding each of the sensor locations. The deployment of localizing systems in place of omnidirectional sensors would provide more detail on the movement of humpbacks off the coast of California and would improve the ability to study the interaction of humpbacks with conspecifics and human activity.

6.3 Improvements to the GPL detector

Adapting the GPL detector for use with certain marine mammal vocalizations would be extremely useful. Several species produce complex transient sounds that are difficult to detect using readily available automated detectors. Manual analysis is carried forward on a large number of marine mammal species, which is a laborious, subjective process that usually provides only basic presence/absence vocalization information. The GPL detector has already proved effective for bowhead whale calls in the Arctic, blue whale "D" calls, and killer whale vocalizations. An eventual goal would be to provide publicly available software with adjustable detection parameters for specific signal and noise environments. It would be beneficial to add additional classification capability to the automated processing system so that certain call types can be distinguished from each other in an automated way. Obtaining more information on types of vocalizations would prove beneficial to habitat modeling efforts - especially for calls that are related to foraging behavior.

Optimal values of the exponents for the GPL detector outlined in Eq. 2.6 were determined from Detection Error Tradeoff (DET) curves (Figs. 2.7-2.8) based on simulations using the six humpback units shown in Fig. 2.6 superimposed on one hour samples of in situ noise records, with varying levels of SNR. The acoustic modeling software in Ch. 3 could be used to improve the verisimilitude of these simulations. In particular, propagation with a full wave-field model allows for distortion, reflection, refraction, dispersion, and selective frequency attenuation of the humpback units. Such effects are site specific owing to the influence of bathymetry and sound speed profile. Site specific characteristics of the noise, by contrast, were already accounted for in the previous simulations. A more complex optimization would allow for other GPL model parameters, including minimum call duration τ_c , to vary as well.

Considerable effort was invested in adapting the GPL detector for real-time detection and localization for the Z-Ray autonomous glider platform. Z-Ray is a buoyancy-driven underwater vehicle shaped like a flying wing that has the capability to perform long duration acoustic monitoring over large areas. Although

the research is not presented in this thesis, a successful at-sea demonstration was conducted in October 2011 in which algorithms onboard Z-Ray detected and localized broadcasted humpback whale song in real-time with an extremely low false alarm rate. The combination of using the GPL detector with beamforming techniques allows false detections from ships and air guns to be nearly eliminated from consideration. Essentially, any transient sounds from these sources are buried in persistent broadband noise; therefore, any transient signal discovered by the GPL algorithm can be eliminated if it has accompanying persistent noise from the same bearing. The combination of using the GPL detector and beamforming techniques could allow for accurate nearly-autonomous reporting of marine mammal activity with very little human assistance. The autonomous platform also has the ability to "track and trail", perhaps following groups of whales over great distances.

6.4 Marine mammals as a source for geoacoustic inversions

An interesting yet somewhat unrelated application of passive acoustic sensing of marine mammal calls is to use marine mammals as opportunistic sources for geoacoustic inversions. If the source level and distribution of marine mammals in a study area are known or otherwise measured, then the bottom type and bottom structure can be calculated in the area, based on the level and structure of received transmissions. Figure 3.9 shows data/model comparisons for differing bottom types for sites Hoke, SBC, and SR. If the distribution and source levels of humpbacks were known, the composition of the bottom could be adjusted in the model until the observed data matches the model predictions. Large baleen whales with high source levels could be very effective, no-cost sources for conducting geoacoustic surveys in an area. A primary advantage comes from a large number of calls spread over a wide area and a range of environmental conditions. Conducting the same number of transmissions from ship-based surveys over varying environmental conditions would be extremely costly.