

©Copyright 2013

Cole Monnahan



# Population Trends of the Eastern North Pacific Blue Whale

Cole Monnahan

A thesis submitted in partial fulfillment of the  
requirements for the degree of

Master of Science

University of Washington

2013

Reading Committee:

Trevor A. Branch, Chair

André E. Punt

Brett T. McClintock

Program Authorized to Offer Degree:  
Quantitative Ecology and Resource Management



University of Washington

**Abstract**

Population Trends of the Eastern North Pacific Blue Whale

Cole Monnahan

Chair of the Supervisory Committee:  
Assistant Professor Trevor A. Branch  
School of Aquatic and Fisheries Science

Blue whales (*Balaenoptera musculus*) were exploited extensively around the world and remain endangered. In the North Pacific their population structure is unclear and current status unknown, with the exception of a well-studied eastern North Pacific (ENP) population. Despite decades without whaling, recent abundance estimates for ENP blue whales indicate a slowed recovery, leading to speculation that their recovery may be hindered by fatal collisions with large vessels. I test this hypothesis by first estimating historical catches and then assessing the population trends with Bayesian population dynamics models under basic assumptions about historical and future ship strikes. I collated previously unreported Soviet catches and combined these with known catches to form the most current estimates of North Pacific blue whale catches. I used the knowledge that ENP whales produce a different song call than blue whales in the western North Pacific to split the conflated catches. The catches were split by estimating spatiotemporal occurrence of blue whales with generalized additive models fitted to acoustic call patterns, which predict the probability a catch belonged to the ENP population based on the proportion of calls of each population recorded by latitude, longitude, and month. When applied to the conflated historical catches, I estimate that ENP blue whale catches totaled 2,969 (95% credible interval 2,301 to 3,516) from 1905–1971, and amounted to 35% (95% credible interval 27% to 41%) of all catches in the North Pacific. The uncertainty in the acoustic data influence the results substantially more than uncertainty in catch locations and dates. These estimated catch series were then



combined with abundance estimates and used to assess the recovery of ENP blue whales. I found the population has recovered to a median of 97% of pre-exploitation levels (credible interval 62% to 99%) and thus density dependence is primarily responsible for their slowed recovery. Further, current levels of ship strikes are not an immediate risk and thus ship strike mitigation would impact the population minimally. Indeed, vessel traffic would need to increase 3–5 times the current levels before the population is at risk. The recovery of ENP blue whale is a conservation success story and demonstrates that endangered cetacean populations can recover under careful management.



## TABLE OF CONTENTS

	Page
List of Figures . . . . .	ii
List of Tables . . . . .	iii
Introduction . . . . .	1
Chapter 1: Estimating historical eastern North Pacific blue whale catches using spatial calling patterns . . . . .	3
1.1 Introduction . . . . .	3
1.2 Methods . . . . .	6
1.3 Results . . . . .	22
1.4 Discussion . . . . .	36
Chapter 2: Do ship strikes threaten the recovery of endangered eastern North Pacific blue whales? . . . . .	39
2.1 Introduction . . . . .	39
2.2 Methods . . . . .	41
2.3 Results . . . . .	51
2.4 Discussion . . . . .	59
Appendix A: Informative prior for $r$ . . . . .	64
Conclusion . . . . .	64

## LIST OF FIGURES

Figure Number	Page
1.1 Spectrogram showing the ENP and WNP blue whale song calls. These calls were recorded on a hydrophone in the Gulf of Alaska at different times; the blue box shows the ENP song call and the red box the WNP song call. The clear distinction between the two is used to differentiate the presence and absence of the two populations. . . . .	4
1.2 Distribution of the blue whale catches with reported locations. Monthly maps showing blue whale catch locations that were reported without uncertainty ( $n=4,230$ ; 50% total) to the International Whaling Commission. There are clear spatial and temporal patterns reflecting the migrations of the two populations and whaling seasons. . . . .	7
1.3 Map of reported blue, fin, sei and common minke whale catch locations (1905–1971). These catch locations were used to infer potential locations for blue whale catches without known locations. Of all blue whale catches only 34% are known only to broad, management-defined regions (shaded boxes) which are shown with their corresponding blue whale catches. . . . .	8
1.4 Observed hydrophone data for May-October. These data are used to fit separate acoustic models for both populations. Colored circles show hydrophones and the proportion of hours calls were observed while “x” denotes those with no observed calls for either population. Hydrophones with observed calls from both populations clearly show overlap for much of the North Pacific. . .	14
1.5 Observed hydrophone data for November-April. Caption as for Figure 1.4. . .	15
1.6 Fits for ENP acoustic, WNP acoustic, and base case prediction models for May–October. These models are fit to the acoustic data in Figures 4 and 5. The third column shows the base case prediction model which is the proportion of ENP to WNP calls, such that red areas correspond to predicted WNP occurrence and blue areas to ENP. The white line denotes where the model predicts an equal chance of observing an ENP or WNP blue whale. . .	23
1.7 Fits for ENP acoustic, WNP acoustic, and base case prediction models for November–April. These models are fit to the acoustic data in Figures 4 and 5. The third column shows the base case prediction model which is the proportion of ENP to WNP calls, such that red areas correspond to predicted WNP occurrence and blue areas to ENP. The white line denotes where the model predicts an equal chance of observing an ENP or WNP blue whale. . .	24

1.8	Fits of the independent variables for the probability of observing a song call ( $p$ parameter). The original model (black lines) and subsequent bootstrapped models (thin colored lines) are shown in logit space. Each panel shows the relationship between the dependent and independent variable after all other independent variables have been accounted for (i.e. the centered partial residuals). Higher relative values indicate a higher probability of observing a call. For longitude and latitude, $x$ -axis tick marks show positions of the observed hydrophones with a small amount of noise added to prevent overplotting. See text for further discussion. . . . .	26
1.9	As for Figure 8 except the results pertain to the over-dispersion parameter $\sigma$ and are in log space. Higher relative values indicate a higher level of variance expected in repeated observations. . . . .	27
1.10	Maps showing the acoustic uncertainty arising from bootstrapping. Uncertainty in the base case probability of a catch being from the ENP population quantified using bootstrapping (i.e. resampling the acoustic data with replacement and refitting the models). The $z$ -axis is the range of the predictions (95th minus 5th percentile) across all 1,000 bootstrapped models. A value of $z = 0$ indicates all bootstrap models predict the same value, while $z = 1$ indicates some bootstrap models predict 0 and others predict 1, so that there is high uncertainty at that location. There is higher uncertainty where overlap between the two populations exists (e.g. Gulf of Alaska) or there is limited data (e.g. Hawaii). . . . .	28
1.11	Estimated annual catches of ENP and WNP blue whales for the base case. Grey values contain 95% of the population uncertainty, and if above the line are more likely to be WNP, below the line more likely to be ENP. The uncertainty in the 1960s is caused by the unreported Soviet catches with high location uncertainty. . . . .	29
1.12	Example model predictions. Catches of ENP (blue) and WNP (red) blue whales for three realizations chosen to represent low, median, and high total ENP catches. Catch positions are plotted with a small amount of noise to prevent overplotting, and are different between the three cases because they are drawn randomly from potential locations. Nearby catches can have different predictions (colors) because months are aggregated. . . . .	33
1.13	Results of the sensitivity analysis to the conversion between song call frequency and density of whales. The sensitivity to the base case assumption that $\alpha = 1$ to the final results is explored with histograms of the total catches of the ENP (blue) and WNP (red) populations for three levels of $\alpha$ . The parameter $\alpha$ is the ratio of the unknown WNP and ENP factors which convert density of whales to song call occurrence. . . . .	35

2.1	Population dynamics parameter distributions for $S_{2013}^{\text{obs}} = 10$ and both priors on $r$ . . . . .	45
2.2	Results of the vessel model. The $y$ -axis scale is arbitrary because it is scaled by the parameter $c$ and is therefore not shown. The points are worldwide vessel statistics from Lloyd's of London. . . . .	52
2.3	Predicted past ship strikes by year for the uninformative prior for $r$ . The dip in the late 1960s is due to Soviet whaling. . . . .	53
2.4	Posterior densities for the relative abundance in 2013 for the two priors on $r$ . The vertical line at 60% denotes the level of maximum productivity. . . . .	55
2.5	Relative abundance trajectories for both priors on $r$ . . . . .	56
2.6	Absolute abundance, the abundance estimates and anthropogenic mortalities for the two priors for $r$ . The rectangles at the bottom denote total mortalities (catches + median strikes) for each year. Ship strikes are so much lower than the whaling catches they are barely visible even in the later years. The five abundance estimates (points) are shown with their 95% confidence intervals (bars). . . . .	57
2.7	Future strikes, abundances, and relative abundances under three mitigation cases. 'Status quo' means no mitigation, 'mitigation' refers to halving the strikes after 2013, and 'none' is a complete elimination of future strikes. . . .	58
2.8	Equilibrium relative abundance for a range of multiples of current levels of vessels. The model was projected forward 100 years for each posterior sample under a constant multiplier. As the multiplier increases there are more vessels and thus strikes, depleting the population. The probability that the population is below 60% of $K$ is shown as a dashed line. . . . .	59

## LIST OF TABLES

Table Number	Page	
1.1	North Pacific blue whale catches by categories of uncertainty. The percent of all catches is indicated in parentheses after the number of catches. Some exact catch positions and dates are unavailable and were inferred with varying levels of uncertainty for many cases. . . . .	12
1.2	Summary of the groups of acoustic hydrophones used in this study. The Gulf of Alaska data are a random sample of 15% of total hours, while the rest are 100% of total hours. Several hydrophones were inoperable for periods of time due to hardware failures, leading to gaps in the data. There is little overlap in year between the groups. See specific references for further details of the data. . . . .	13
1.3	Model selection results. Models are compared across likelihoods and parameter structures. The smallest $\Delta AICc$ indicates the model with the highest support is the most complex model. The binomial model in particular fits extremely poorly, confirming the need for an over-dispersed likelihood. . . . .	22
1.4	Base case ENP catches by year showing the median as well as lower and upper 95% interval from the 1,000 realizations. These catches do not include uncertainty in $\alpha$ . . . . .	31
1.5	Base case WNP catches by year showing the median as well as lower and upper 95% interval from the 1,000 realizations. These catches do not include uncertainty in $\alpha$ . . . . .	32
1.6	Comparison of the sources of uncertainty in total ENP catches. Catch uncertainty arises from missing locations and dates, acoustic uncertainty in the song call occurrence data, and ecological from uncertainty in ecological assumptions over the range $1/2 \leq \alpha \leq 2$ . Differences in totals compared to Table 4 are due to rounding and the order of percentile calculations. . . . .	34
2.1	Photographic mark-recapture estimates used to fit the population dynamics models. These are assumed to be log-normally distributed and independent. . . . .	43
2.2	Historical vessel traffic. Data on vessels over 1000 gross tons worldwide used in the vessel model to infer changes in vessel traffic over time. . . . .	47

2.3	Estimated intrinsic rate of increase ( $r$ ) and carrying capacity ( $K$ ) for eastern North Pacific blue whales. Results are shown for a uniform prior for $r$ : Uniform $(0, 0.114)$ and an informative numerical prior from a meta-analysis, as well as assumed ship strikes in 2013, $S_{2013}^{\text{obs}}$ , of either 10 or 20. . . . .	52
2.4	Estimates of the abundance and abundance relative to pre-whaling levels ( $N_{2013}/K$ ) for year 2013. Different combinations of the prior for $r$ and value of observed ship strikes in 2013, $S_{2013}^{\text{obs}}$ , represent different assumptions and are a sensitivity analysis. The base case is indicated in bold. . . . .	54

## ACKNOWLEDGMENTS

I would like to acknowledge and thank the following who provided the materials necessary for me to accomplish my thesis. Kate Stafford provided crucial acoustic data from her dissertation, lively discussions regarding blue whale ecology and acoustics, as well as positive energy for my project. Yulia Ivaschenko retrieved and translated original Soviet reports and maps from which missing blue whale catches could be inferred. Erin Oleson provided acoustic data from her dissertation as well as helpful critiques of my ecological assumptions. John Calambokidis and his staff at Cascadia Research Collective put in countless hours in their work on the photographic ID database. Without these data the analyses presented here would not be possible. I would also like to acknowledge the people who have developed and contributed to R, ADMB, L<sup>A</sup>T<sub>E</sub>X and other open source software projects used extensively in my thesis.

I would also like to acknowledge my committee for their guidance during this project. Brett McClintock introduced me to interesting mark-recapture models which unfortunately did not make it into my thesis and reminded me (rightfully so) to keep the caveats of my analysis in mind. André Punt taught me the applied skills needed to tackle this analysis, challenged me at every step of this process, and provided invaluable feedback on the technical methods I used. Trevor Branch helped me to stay focused on the important issues and constantly reminded me of the “big picture” as I was mired in the details. His guidance in particular has helped me grow from a student to a scientist. QERM and SAFS have provided a better graduate school community and experience than I could have asked for, and I am grateful for my experiences (social and academic) during my tenure as a graduate student at the University of Washington.

## DEDICATION

I dedicate this thesis to my family, who have supported and encouraged me through the entire process. Deonna in particular has been a constant source of positive energy and believed in me through the ups and downs.

## INTRODUCTION

Due to commercial whaling blue whales are listed as “endangered” under the Endangered Species Act and considered “depleted” and “strategic” under the U.S. Marine Mammal Protection Act [?]. The U.S. West Coast is an important summer feeding area for the eastern North Pacific (ENP) blue whale population, which ranges from the equator to the Gulf of Alaska. Despite being the focus of many studies, the recovery of ENP blue whales has not been quantitatively assessed. Concern for their continued recovery was raised after a spike in fatal ship strikes in 2004 [?] and a series of photographic mark-recapture estimates from 1993–2008 [?, ?, ?] that showed no clear increase in abundance. Although the impact of ship strikes on the population is unclear, strategies for mitigating them (e.g. vessel speed reductions) have been advocated [?]. However, without an understanding of the population dynamics of ENP blue whales it is difficult to discern the risk that present and future ship strikes pose and the impact of potential mitigation strategies.

Here I combined historical catches with abundance estimates and fit these to a population dynamics model in order to estimate the past and future abundance trajectory. The abundance estimates were available, but the catches were not because they were conflated with other populations of blue whales in the North Pacific. Thus estimating the catch history was a critical step before an assessment could be conducted. Since ship strikes are considered an important factor influencing the recovery of this population, it was also necessary to estimate historical ship strikes and include them in the model. Once catches and ship strikes were estimated, the population dynamics model could be fit and the impact of different mitigation strategies on the future population trends explored.

The hypothesis that ship strikes are hindering their recovery is not the only viable hypothesis that can be tested in the assessment framework. An alternative hypothesis is that the population has recovered and their growth has slowed due to density dependence.

If ship strikes are found to be a risk to the ENP population, then management action to reduce collisions may be warranted. However, if ship strikes are found to have a negligible impact, then it would be unlikely that mitigation scenarios would lead to any appreciable change in the rate of recovery. In this thesis I perform the first quantitative assessment of the ENP population of blue whales. My first chapter uses acoustic call occurrence patterns to split historical catches between populations based on location and time. The estimated catches are used in my second chapter to test the ship strike hypothesis and assess the impact on the population under different future mitigation scenarios. The results of my thesis provide objective scientific information on the risk of ship strikes that can be used to manage endangered ENP blue whales.

## Chapter 1

**ESTIMATING HISTORICAL EASTERN NORTH PACIFIC BLUE WHALE CATCHES USING SPATIAL CALLING PATTERNS****1.1 Introduction**

The blue whale (*Balaenoptera musculus*) is a cosmopolitan species that is currently classified as endangered globally [?]. The key reason for their depletion was targeted whaling throughout their range. In the North Pacific blue whales were hunted from 1905–1971 which is thought to have depleted local populations. At present a large population known as the eastern North Pacific (ENP) population feeds each summer within 30 miles of the California coast and is believed to occur from the equator to the Gulf of Alaska [?]. The ENP population is the one of the most accessible blue whale populations in the world and has been the focus of many studies. However, a formal assessment of their current abundance relative to historical levels has not been conducted, largely because of the difficulty in separating historical catches of the ENP population from catches from other populations in the western North Pacific (WNP).

To estimate historical catches for the ENP population one needs to understand the migration pathways of both populations over time and space. The occurrence of blue whale vocalizations in the North Pacific offers an opportunity to estimate these migration pathways. Analyses of these vocalizations observe two distinct NP blue whale ‘song call’ types (Figure 1.1) [?, ?, ?, ?], which are assumed to be produced by the ENP and WNP populations. The ENP song call is comprised of the rhythmic repetition of the two part vocalization known as the ‘AB call’ type [?]. The AB call type components are produced exclusively by males in a variety of behavioral states [?]. The AB song calls are observed only in lone, traveling males and are produced year-round and thus likely have some kind of reproductive function, although their exact purpose is unknown [?]. The WNP song call is a single part call repeated as song in a similar way as the AB song of the ENP population but has a

clearly identifiable and distinct form (Figure 1.1). No studies have examined the behavioral context of WNP song, but I assumed it to be similar in function to the ENP song. Another common call type observed in the ENP population is the downswept ‘D’ produced by foraging groups of both sexes [?]. A recent study found a temporal separation in the production of the AB and D call types at a summer feeding area, and argued that both were necessary for an accurate assessment of the timing of fine-scale seasonal movements into foraging regions [?]. Unfortunately there is no evidence of population differences in D calls as there is with song calls, and so I focused exclusively on the song call occurrence patterns.

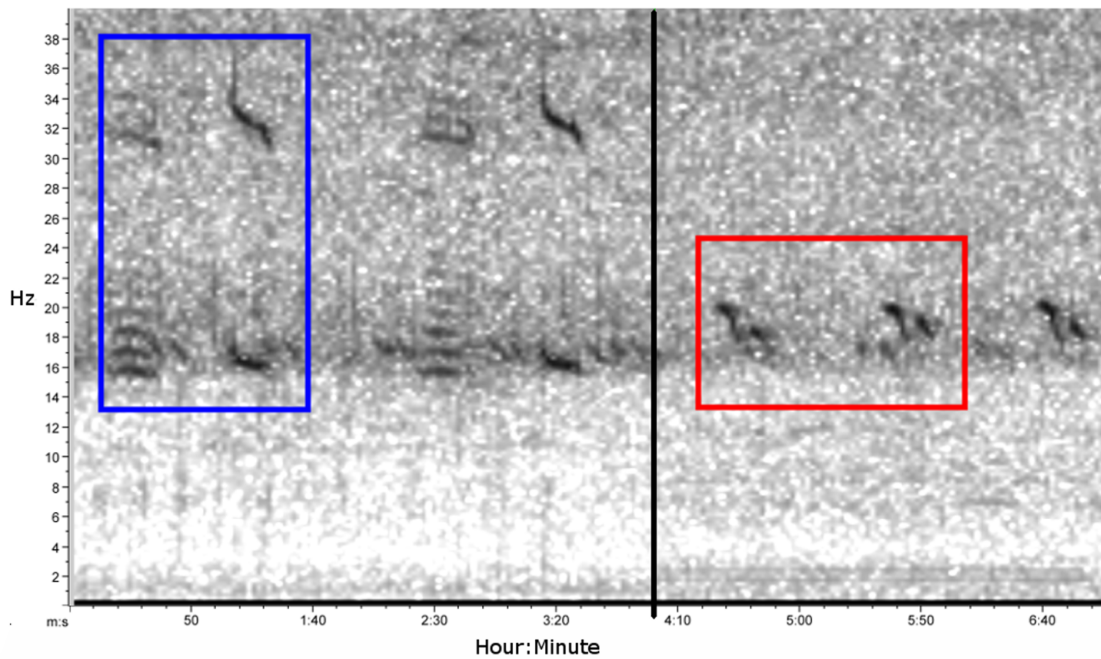


Figure 1.1: Spectrogram showing the ENP and WNP blue whale song calls. These calls were recorded on a hydrophone in the Gulf of Alaska at different times; the blue box shows the ENP song call and the red box the WNP song call. The clear distinction between the two is used to differentiate the presence and absence of the two populations.

Besides distinct song call types, the acoustic analyses also show that broad spatial migration patterns over time can be qualitatively distinguished from the data: the ENP population inhabits waters from the Costa Rica Dome to the Gulf of Alaska and the WNP

population is heard predominantly between Kamchatka and the Gulf of Alaska. I use the timing of occurrence of these song calls of the two populations to quantitatively model the seasonal movements by the two populations.

Additional data sources beyond vocalization differences support the hypothesis of two populations of blue whales in the North Pacific. Satellite tags, photographic identification, and sightings show individuals moved between the Gulf of Alaska, the Californian coast, and the eastern tropical Pacific [?, ?], while there are also significant differences in length between whales in the ENP and WNP [?]. A recent study also suggests that the Gulf of California blue whales are also genetically similar to those off California [?]. Taken together there is substantial evidence that the ENP population is separate from blue whales in the WNP.

While diverse sources of evidence point toward a single ENP population, the population structure of the central and western NP remains unclear, and there may be additional populations in these regions. For example, there is some speculation that a population off coastal Japan was extirpated and that another exists around Hawaii [?]. The acoustic and other data shed little light on these hypotheses, which do not affect my estimates of ENP catches, although I recognize that what I call the ‘WNP population’ may include additional population structure.

To assess the recovery and resilience of the ENP population, two key pieces of information are required: abundance estimates and historical catches. Historical catches in combination with abundance estimates [?] provide information about the absolute size of the pre-exploitation population [?]. Historical catches are not available for the ENP and WNP separately because the blue whale population structure was unknown during the period of commercial whaling and catches were reported only as blue whales. Thus there only exists a single, conflated catch series containing both ENP and WNP catches which needs to be separated before an assessment can be conducted. The probability of each catch belonging to the ENP population needed to be determined to split the conflated catch series between the two populations. Qualitatively, it is clear from the acoustic data that whales caught off Japan and Kamchatka are highly unlikely to be from the ENP, and likewise blue whales caught off California are surely from the ENP population. However, splitting catches in the

Gulf of Alaska is more complicated because the two populations overlap in space and time there during the feeding season [?].

In this study I collated the most up-to-date time series of historical catches of blue whales in the North Pacific, including previously unreported Soviet catches in the 1970s. Using song call occurrence data, I modeled the probability that a whale caught in a particular location and month belonged to the ENP population, and applied this model to the historical catches. The resulting time series of catches for ENP and WNP blue whales are the first to be obtained for each population.

## **1.2 Methods**

### *1.2.1 Catch Data*

Official catches for NP blue whales were obtained from two databases maintained by the International Whaling Commission (IWC): the ‘summary database’ [?] and the ‘catch database’ [?]. The summary database contains annual catch totals by species (including blue whales) for each whaling expedition or shore station since 1900, where this was known, but contains no detailed information about the exact location or date of individual catches. The catch database contains data for individual whales including species, location, date, expedition, length and sex (Figure 1.2). Detailed information was not available for all catches, particularly in earlier years and during the period of later Soviet whaling, thus the catch database contains only a subset of the summary database. The expedition code referred to either a pelagic fleet operating in a predefined IWC region (Figure 1.3), or catches taken back to a land station for shore-based processing; thus a pelagic vessel operating in multiple regions in the same year was recorded as different expeditions. For most expeditions the start and end month of a year of operation were recorded. Catches of unspecified whales were rare, accounting for only 1.5% of total whales of all species in the North Pacific and were not included in this analysis.

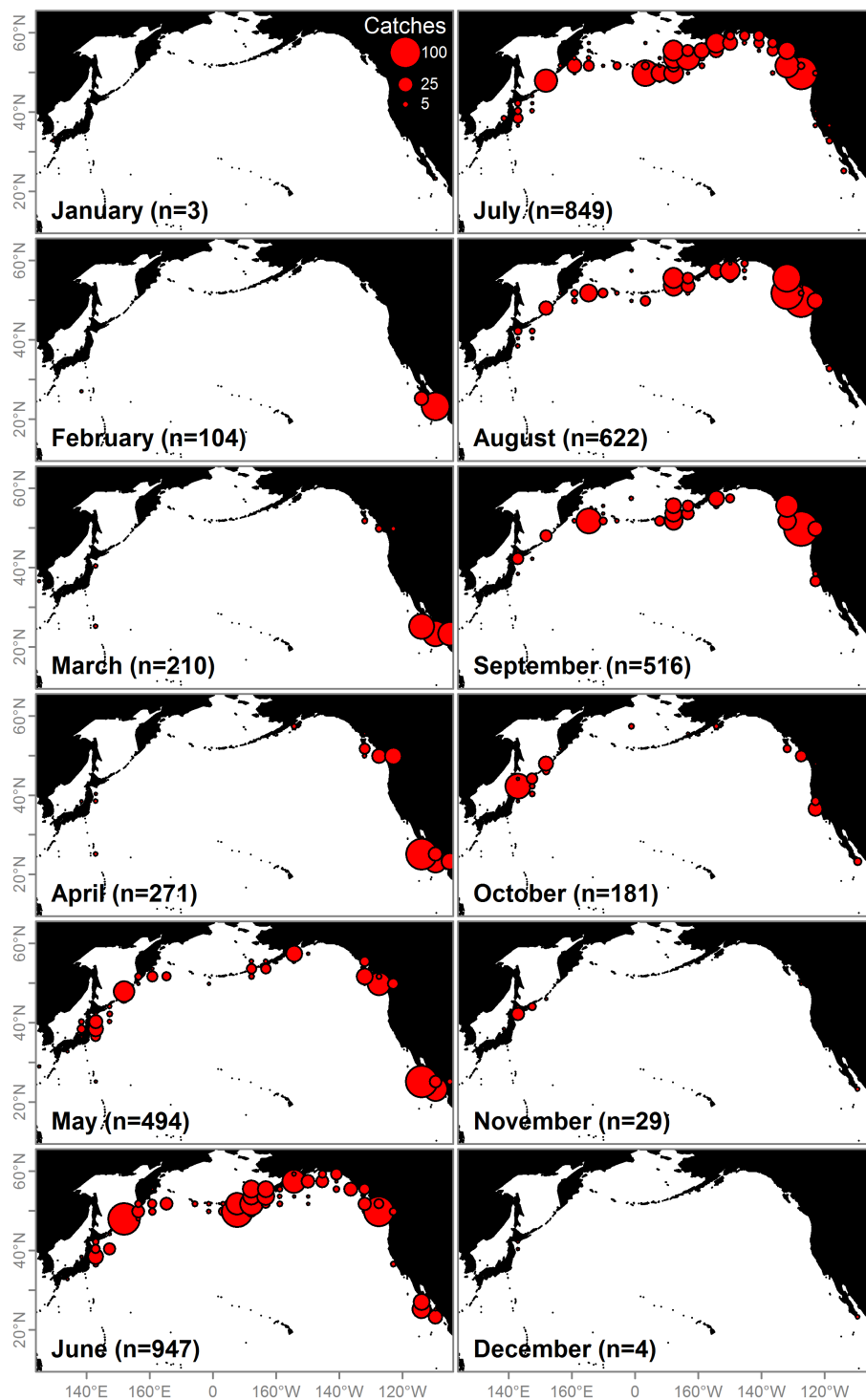


Figure 1.2: Distribution of the blue whale catches with reported locations. Monthly maps showing blue whale catch locations that were reported without uncertainty ( $n=4,230$ ; 50% total) to the International Whaling Commission. There are clear spatial and temporal patterns reflecting the migrations of the two populations and whaling seasons.

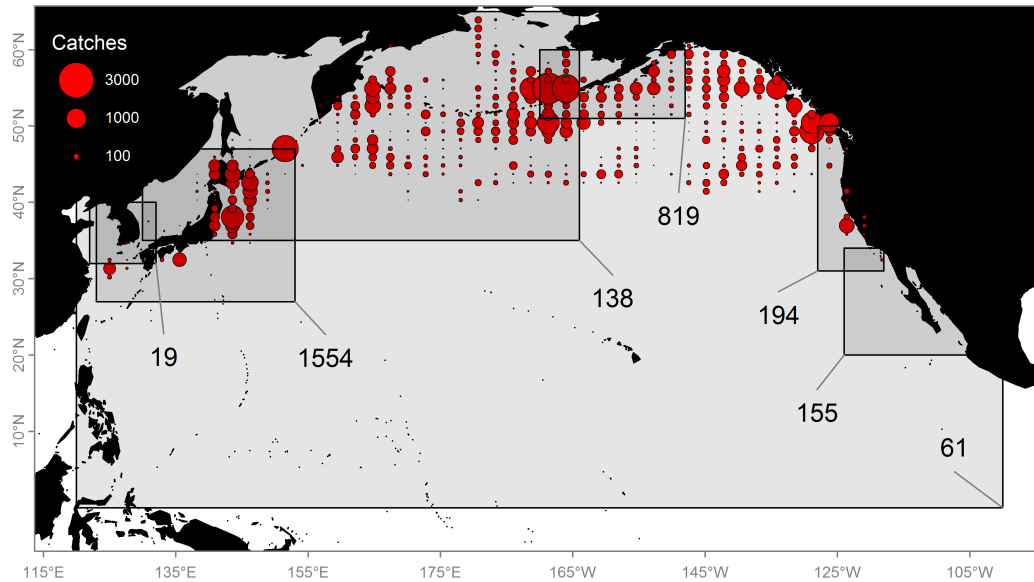


Figure 1.3: Map of reported blue, fin, sei and common minke whale catch locations (1905–1971). These catch locations were used to infer potential locations for blue whale catches without known locations. Of all blue whale catches only 34% are known only to broad, management-defined regions (shaded boxes) which are shown with their corresponding blue whale catches.

These databases contain all catches known to the IWC as of November 2011. However, it has been known since the mid-1990s that the USSR caught additional blue whales in the 1960s and 1970s that were misreported or not reported to the IWC and are missing from the IWC databases [?]. Discrepancies in reported Soviet catches were resolved using scientific reports retrieved from the public archives in Vladivostok, Russia [?, ?, ?, ?, ?, ?, ?]. These reports complemented existing IWC data, in a similar manner to previous work resolving Soviet discrepancies in the Indian Ocean and Southern Hemisphere [?]. The original Soviet reports in the North Pacific provided better estimates of the number of catches taken by Soviet whaling vessels from 1962–1971 and clues about the likely dates of catches. In many cases original maps were also recovered and used to help infer likely catch locations. Not all of these Soviet reports were available and some uncertainty remained surrounding the locations and dates of Soviet catches during 1962–1971. Analyses of the recovered reports added 385 (4.5% of total reported) blue whale catches to the 2011 version of the IWC annual

database total.

### 1.2.2 Catch Uncertainty

In addition to missing data about Soviet catches, many other catches in the IWC catch databases are missing locations or dates, and the original records are lost to history. This ‘catch uncertainty’ must be explicitly taken into account when assigning catches to the ENP or WNP populations to reliably quantify the uncertainty associated with the catch time series. I used a Monte Carlo method to integrate over all potential locations and months for each catch. Conceptually, possible sets of locations and months were determined for each catch, from which random samples could be generated. By generating and splitting many potential catch series, the catch uncertainty was propagated through into the uncertainty in the final results. The challenging part was to determine accurate sets of locations from which to draw randomly. I assigned catches into five categories ranging from the most uncertain location (IWC Region) to the most certain (exact position known), and treated each category as described below.

(1) ‘IWC Region’ ( $n=2,936$ ) was the most uncertain category and occurred when no location information was available for an expedition in a year beyond the broad regions defined by the IWC. IWC regions are defined by longitude and latitude ranges (i.e. a rectangle), which cover both land as well as regions of ocean where no whaling ever occurred. Clearly, randomly drawing uniform locations from these large regions would lead to inaccurate and even nonsensical catch locations. Instead, I drew from locations within each IWC region where catches of blue whales and other species had actually been taken. Since many blue whale catch locations were not reported, I assumed the spatial variation would be under-represented by only using that species (Figure 1.2). Conversely, some species such as sperm (*Physeter macrocephalus*), humpback (*Megaptera novaeangliae*) and bowhead (*Balaena mysticetus*) whales occupy very different habitations than blue whales and the spatial variation would be over-represented. I therefore randomly drew from the reported locations of catches most similar to the blue whales: blue ( $n=4,640$ ), fin (*B. physalus*,  $n=47,232$ ), sei (*B. borealis*,  $n=50,875$ ), and common minke (*B. acutorostrata*,  $n=4,742$ ) whales that

were caught during 1905–1971 when blue whales were targeted. By drawing from actual catch locations from similar species, I accounted for spatial patterns of whaling effort, but implicitly assumed this effort was similar between species and that blue whales occupied the same spatial extent as other species. This assumption was expected to increase the spatial uncertainty of catches compared to the scenario where I randomly drew only from locations where blue whales were caught. Although catches in the IWC region category are the most uncertain, most of these catches came from Japanese and Korean waters ( $n=1,554$ ) and Aleutian islands ( $n=819$ ) which likely contain mostly WNP blue whales (Figure 1.3).

(2) The ‘Partial Locations’ category ( $n=606$ ) was where an expedition reported locations for some species but not for some blue whales. Most of these cases ( $n=504$ ) were Soviet pelagic expeditions where the month is known and can be cross referenced with the catch database. However, there were a few cases ( $n=102$ ) when an expedition reported some, but not all, of the catches in a year. I inferred locations of catches in this category by drawing, with replacement, from locations of all species in the same month.

(3) The ‘Inferred’ category ( $n=478$ ) was where the location was not reported, but could be inferred with high precision. Most often this occurred when catches were reported as coming from a land station with a known location in a previous year, but no location reported in the current year. The unreported catches must have the same locations as previous years since the land stations were stationary. There were some catches ( $n=95$ ) with unknown locations but known dates, in addition to known locations of catches of other species on the same dates. For these cases averages of locations of catches of the day before and after for all species were used without uncertainty.

(4) The ‘Soviet’ category ( $n=319$ ) was specifically for Soviet catches from 1962–1971. Where possible I determined areas of operation during the period of unreported catches using the original scientific reports and maps. These rectangles were smaller than the broad IWC regions and were mostly in the Gulf of Alaska. The vessels could have been anywhere within these rectangles so I drew random uniform locations from each rectangle.

(5) The ‘Certain’ category ( $n=4,230$ ) was for catches where the exact location was reported in the catch database without uncertainty. These are, fortunately, the most common case (49% of the total). As with location, blue whale catch dates were often not reported

( $n=4,969$ , 58%) in the IWC catch database. I applied the same Monte Carlo approach to account for uncertainty in date by sampling months from the possible months. I used an expedition's operating period from the IWC summary database to determine the range of possible months. If the operating period for an expedition was missing, I assumed the catches could have been from any month. However, I assigned a probability to each potential month that reflects the fact that most blue whales were caught in summer because of defined whaling seasons, weather limitations, and seasonal availability of whales. This weighting was based on the proportions of whales caught in each month for the same set of species used for geographic uncertainty: blue, fin, sei and common minke whales. Thus, there was a range of potential locations and dates for each catch depending on the level of detail reported (Table 1.1). A routine was written in the statistical software framework R [?] which randomly samples from the sets of locations and months for all catches, and returns a dataset which represents a possible realization of the conflated catch series. Repeatedly running this routine to generate sets of potential catch series (i.e. Monte Carlo approach) allowed me to estimate the catch uncertainty. The generated sets of conflated catch series were then propagated through to the process of assigning catches to populations.

### 1.2.3 Acoustic Data

Detections of ENP and WNP song calls came from six sources in the North Pacific (Table 1.2, Figures 1.4 and 1.5). For all hydrophones, data were converted into hourly presence or absence of ENP and WNP calls. Hydrophone sources were: (1) six hydrophones placed by the National Oceanic and Atmospheric Administration (NOAA) in the eastern tropical Pacific [?]; (2) sixteen US Navy Sound Surveillance System and other hydrophones spread across the North Pacific (note that their exact locations were protected and approximated based on previously published studies [?]); (3) six hydrophones in the Gulf of Alaska placed by NOAA (after 9 months, one of these hydrophones was discontinued and another was moved slightly [?]); (4) six hydrophones in the Channel Islands and Cortez and Tanner Banks deployed by Scripps Institution of Oceanography [?]; (5) one hydrophone moored off Oahu, Hawaii, described in [?] and used in [?]; and (6) a single hydrophone off Wake Island

Table 1.1: North Pacific blue whale catches by categories of uncertainty. The percent of all catches is indicated in parentheses after the number of catches. Some exact catch positions and dates are unavailable and were inferred with varying levels of uncertainty for many cases.

<b>Category</b>	<b>Uncertain Month</b>	<b>Certain Month</b>	<b>Location Total</b>	<b>Description</b>
IWC Region	2841 (33%)	95 (1%)	2936 (34%)	Known to be within a large rectangular region (Figure 1.3)
Partial Locations	172 (2%)	434 (5%)	606 (7%)	Catches of other species by the same expedition and year are available.
Inferred	427 (5%)	51 (1%)	478 (6%)	Able to infer accurate locations with high confidence.
Soviet	160 (2%)	159 (2%)	319 (4%)	Original Soviet whaling logs give information about locations and months.
Certain	0 (0%)	4230 (49%)	4230 (49%)	Reported to the IWC.
<b>Total</b>	<b>3600 (42%)</b>	<b>4969 (58%)</b>	<b>8569 (100%)</b>	

[?].

Table 1.2: Summary of the groups of acoustic hydrophones used in this study. The Gulf of Alaska data are a random sample of 15% of total hours, while the rest are 100% of total hours. Several hydrophones were inoperable for periods of time due to hardware failures, leading to gaps in the data. There is little overlap in year between the groups. See specific references for further details of the data.

<b>Group</b>	<b>Deployed</b>	<b>Reference</b>	<b>Notes</b>
ETP ( $n=6$ )	May-96 to May-97	[?]	
GOA ( $n=6$ )	Oct-99 to Jun-01	[?]	GOA5 removed after 5/2000, GOA3a moved to GOA3b in 5/2000.
SOSUS ( $n=13$ )	Nov-95 to Nov-96	[?]	Exact positions are still protected.
Wake ( $n=1$ )	Apr-92 to Dec-92	[?]	Triad in place, but only one used; does not cover entire year.
Kaneohe ( $n=1$ )	Aug-92 to Apr-93	[?]	Does not cover entire year
CH ( $n=6$ )	Aug-00 to Feb-04	[?]	Down during 5 February to 15 April 2002

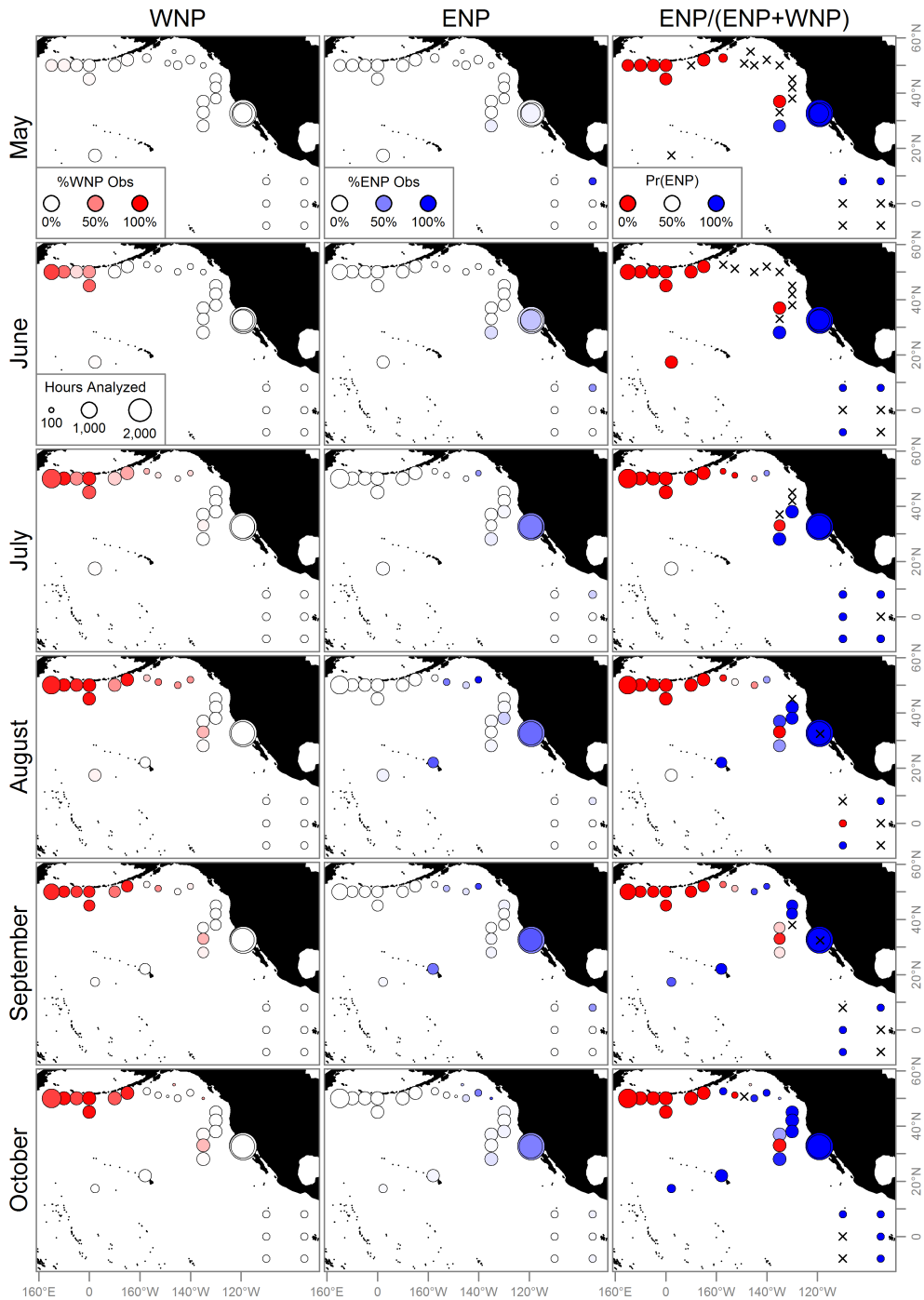


Figure 1.4: Observed hydrophone data for May-October. These data are used to fit separate acoustic models for both populations. Colored circles show hydrophones and the proportion of hours calls were observed while “x” denotes those with no observed calls for either population. Hydrophones with observed calls from both populations clearly show overlap for much of the North Pacific.

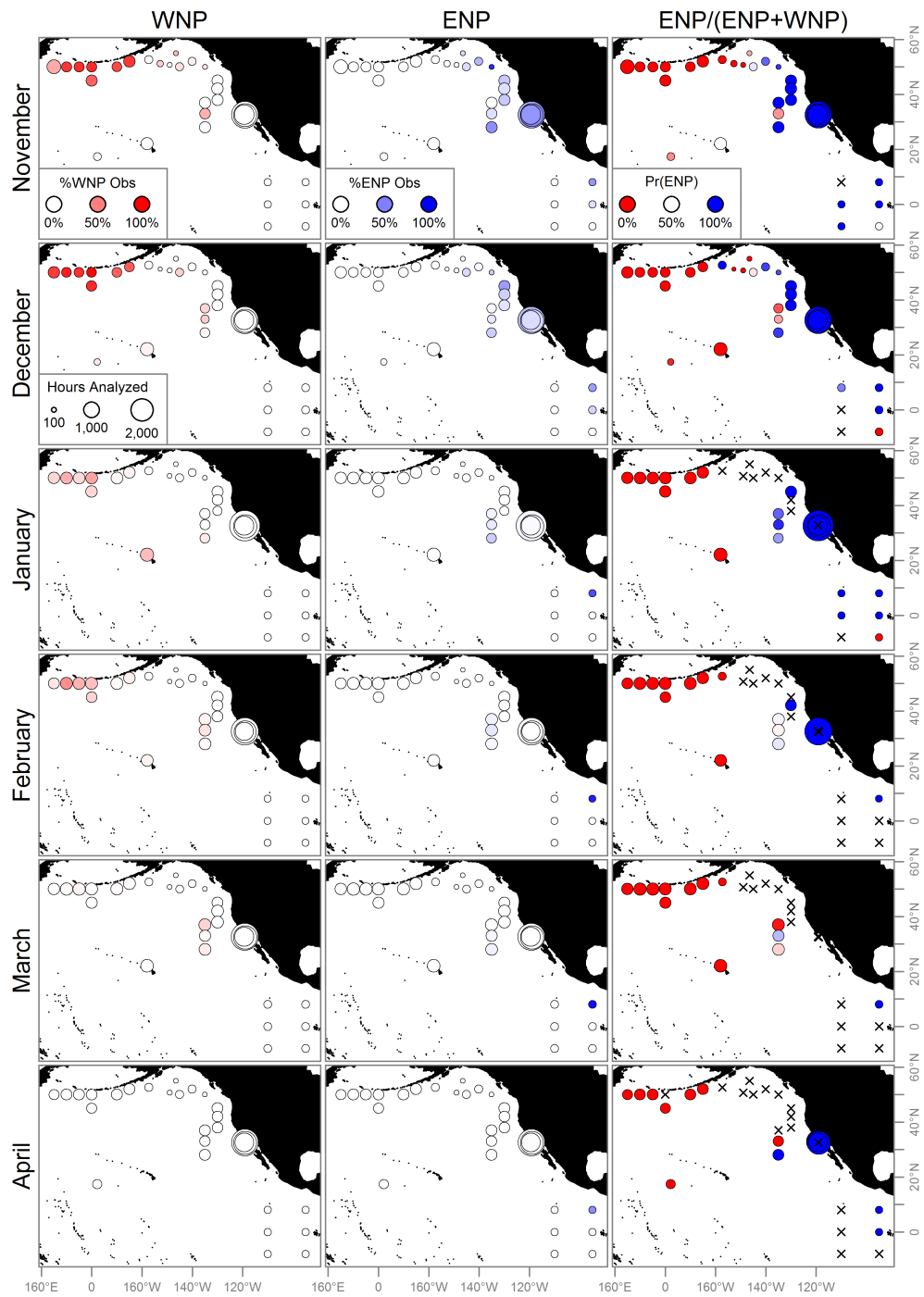


Figure 1.5: Observed hydrophone data for November-April. Caption as for Figure 1.4.

Presence was recorded when at least one call was heard in a given hour, implying no differentiation in the frequency of calls within an hour. The presence data were then aggregated by month, and across years for the Gulf of Alaska and Channel Island hydrophones. Specific details regarding hydrophone deployment and methods for processing the raw data can be found in the original references (Table 1.2). Taken together these 32 hydrophones covered a large expanse of the North Pacific (Figures 1.4 and 1.5), and provided a unique opportunity to model the large movement patterns undertaken by both populations of blue whales.

#### 1.2.4 *Acoustic Presence Models*

The acoustic data used here contained information about NP blue whale distribution both in space and season, from which migration patterns could be inferred for the two populations. I used these patterns to infer to which population the conflated historical catches belonged. I developed a two-stage model which (1) estimated the presence or absence of each population in space and time and then (2) predicted the probability a catch belonged to the ENP population. The first stage fitted a spatial surface separately to the ENP and WNP populations based on the monthly call rates. These two independent ‘acoustic models’ predicted the probability of observing at least one song call in an hour at a given location and time. In the second stage the two acoustic models were combined together into a single ‘prediction model’ used specifically for predicting the probability that a conflated catch belongs to the ENP population, given its location and month. This section describes the motivation and development of the acoustic models and how they are combined into the prediction model.

To develop these acoustic models I assumed the spatiotemporal song call patterns observed in the data reflected an underlying occurrence and used the data to estimate this occurrence. The first choice in developing the acoustic models was deciding which likelihood function is most appropriate for the data. Since the data were presence-absence (a whale either was or was not detected in each hour), a binomial likelihood seemed like a natural choice. However, sequential hours of detection at a location may not have been independent,

as is expected in a true binomial experiment, since a single traveling or resting whale will potentially be recorded calling for many consecutive hours. If a call was heard in one hour, it would likely be heard there in the following hour as well. As a result, the variance in the data was expected to be higher than a binomial process. In other words, the data were over-dispersed. I addressed this issue by adding an over-dispersion parameter to the binomial model to increase the expected variance, which had the effect of turning the binomial model into a beta-binomial model. The parameterization of the beta-binomial used in this study had a likelihood function of

$$L(p, \sigma | y, n) = \frac{\binom{n}{y} \Gamma(1/\sigma) \Gamma(y + p/\sigma) \Gamma(n + (1 - p)/\sigma - y)}{\Gamma(n + 1/\sigma) \Gamma(p/\sigma) \Gamma((1 - p)/\sigma)} \quad (1.1)$$

where  $y$  was the number of hours with at least one observed call,  $n$  the total hours analyzed,  $0 \leq p \leq 1$  the probability of observing at least one song call in an hour,  $\sigma > 0$  the over-dispersion term, and  $\Gamma$  is the gamma function. For this distribution,  $E[y] = np$  and  $\text{Var}[y] = np(1 - p) [1 + \sigma(n - 1) / (1 + \sigma)]$ . As  $\sigma \rightarrow 0$ , the variance becomes  $np(1 - p)$  which is the variance of a binomial model, demonstrating that the  $\sigma$  parameter increases the variance as desired. The support in the data for a  $\sigma$  term in the model was tested in the model selection phase.

In developing the acoustic models I used only covariates which were also available for the catches. Thus, while oceanographic conditions at the hydrophones likely could have explained the occurrence of blue whales [?], we could not obtain these oceanographic data for the historical catches. We therefore included only latitude, longitude, and date as covariates in the model. The positions were used as is, but it was natural to aggregate time into appropriately-sized units. I chose to bin time into months because anything longer in duration may have missed the fine-scale movements, and the weeks-long delay between presence and singing found in [?] precluded anything shorter.

The relationship between occurrence and spatial positions is complex, and unlikely to follow a smooth mathematical form (e.g. linear, quadratic, etc.). I therefore used non-parametric models to infer the relationship between explanatory and response variables. Typically the non-normal sampling and nonparametric nature proposed here would have implied a generalized additive model (GAM). However, GAMs are not flexible enough to

accommodate beta-binomial sampling [?]. I therefore used a more flexible modeling platform called generalized additive models for location, shape and scale (GAMLSS) which was developed to extend GAM models to handle more complex distributions, including the beta-binomial [?]. The GAMLSS package in R [?, ?] allowed additive predictors on both parameters of the beta-binomial distribution simultaneously, so that both varied spatially in a nonparametric way. I chose additive position terms and categorical month terms for both parameters as the most complex case (i.e. the full model) and fit, for example, the ENP model as:

$$\text{logit}(\hat{p}_{ENP}) = \text{cs}(\text{Longitude}) + \text{cs}(\text{Latitude}) + \text{Month} \quad (1.2)$$

$$\log(\hat{\sigma}_{ENP}) = \text{cs}(\text{Longitude}) + \text{cs}(\text{Latitude}) + \text{Month} \quad (1.3)$$

where logit and log were the canonical link functions and *cs* is a cubic spline smoother. The same structure was also used to fit the full model for the WNP data. The resulting acoustic models predicted the probability of observing at least one song call in an hour for a given position and month.

Model selection proceeded by fitting several plausible models and comparing corrected Akaike information criterion (AICc) values [?], as well as examining model residuals using standard diagnostics. Model selection was not used to infer biologically meaningful predictors of blue whale habitat, as was done for example in [?], but rather to identify the most appropriate structure and complexity of the models of blue whale song call occurrence. To test the statistical support for over-dispersion I fitted both binomial and beta-binomial GAMLSS models. Likewise, to test the additive nature of the models three variants of the beta-binomial GAMLSS model were also fit, having either constant, linear, or additive terms. The best supported acoustic models for the ENP and WNP represent the first of my two-stage model.

### 1.2.5 Prediction Model

The second stage of the modeling was to develop a model which predicts the probability that a catch belonged to the ENP given its location and month. To make these predictions I combined the ENP and WNP acoustic models together. Conceptually, if the ENP acoustic

model predicts high occurrence in a region and the WNP model does not, I would assign catches in that region to the ENP population. Specifically if I let  $d_{ENP}(t, z)$  and  $d_{WNP}(t, z)$  be the density of whales at date  $t$  and location  $z$  then the probability a catch belonged to the ENP is:

$$\Pr(\text{Catch is ENP}|t, z) = C(t, z) \quad (1.4)$$

$$= \frac{d_{ENP}(t, z)}{d_{ENP}(t, z) + d_{WNP}(t, z)} \quad (1.5)$$

$$= \frac{1}{1 + \frac{d_{WNP}(t, z)}{d_{ENP}(t, z)}} \quad (1.6)$$

$$= \frac{1}{1 + D(t, z)} \quad (1.7)$$

Clearly the densities of the populations dropped during the period of historical whaling. However, the way the ratio of the densities,  $D(t, z)$ , changed over years is unclear because the abundance trends of both populations are poorly understood. For my base case I therefore assumed that both populations were depleted at an equal rate so  $D(t, z)$  was independent of year and simplified to  $D(m, z)$  for month  $m$ . I further assumed the ratio of densities was proportional to the ratio of the probabilities of call occurrence:

$$D(m, z) \approx \alpha(m, z) \cdot P(m, z) \quad (1.8)$$

where  $P(m, z) = \hat{p}_{WNP}(m, z) / \hat{p}_{ENP}(m, z)$  is the ratio of estimated call probabilities and  $\alpha(m, z)$  is an unknown scaling factor that may change over month or space in response to many external influences. Lacking any information on its true value, I assumed a default of  $\alpha(m, z) = 1$  for the base case and the sensitivity to this assumption was explored below. Thus the base case prediction model is  $\hat{C}(t, z) = 1 / (1 + P(m, z))$ . The above formula combined the two acoustic models together into a prediction model that can be applied directly to the catches.

The GAMLSS models provide standard error estimates only at the positions of the hydrophones, but the predictions were to be interpolated and extrapolated, sometimes widely, to predict the catches. Since accurately quantifying the uncertainty in predictions is an important aspect of modeling these data, I used a common alternative approach known as nonparametric bootstrapping [?]. With this approach the acoustic and prediction models

were refit to sets of acoustic data resampled with replacement. For each resampled set of data, the acoustic models were refit, and the predictions of ENP catches were recalculated. Since the Monte Carlo method for quantifying uncertainty in catches also relied on sampling repeatedly to account for uncertainty over potential positions and months, it was natural to incorporate the two resampling steps together. The algorithm used to accomplish this was:

1. Generate a random realization of a potential conflated catch series  $\hat{C}_i$ .
2. Bootstrap the acoustic data by sampling with replacement from rows of the acoustic data. Each row contained the proportion of hours with presence in a month for a hydrophone.
3. Refit the two GAMLSS acoustic models to the bootstrapped acoustic data.
4. Use the fitted bootstrapped models to predict an ENP catch series  $C_i^{ENP}$ .
5. Aggregate the split catches by total and year to create time series.
6. Repeat steps 1–5  $n$  times.
7. Summarize the results using the median and percentiles over the set  $\{C_1^{ENP}, \dots, C_n^{ENP}\}$ .

This algorithm simultaneously incorporated both the catch and acoustic uncertainty in allocating catches between the two populations and thus made it difficult to compare the relative sizes of these uncertainties. I therefore also ran the algorithm without step 3, in essence removing the effect of bootstrapping so that only catch uncertainty was estimated. Implementing this algorithm split the catches and quantified the relative size of the uncertainties.

### 1.2.6 Ecological Uncertainty

There was one more important source of uncertainty that affected the results. In the base case results I made two big assumptions in setting  $\alpha = 1$ : (1) the ratio of densities at a

location was identical across years and equal to the ratio of call probabilities and (2) the behavioral context of song calls were identical over time and space, and between populations. Deviations from these ecological assumptions could arise from a variety of biological, environmental or anthropogenic factors. Examples of potential biological factors are different calling rates between populations, dynamic migration patterns (over year, space, or seasons) due to changing environmental conditions, and differences in the behavioral context of singing between populations. The largest potential anthropogenic factor was if whaling depleted the two populations at different rates so that the ratio of densities of the populations changed over time.

The impact of misspecifying  $\alpha$  on the results depends on the size of  $P$ : in regions where only one population is found the impact will be small, and the impact is highest when the populations overlap the most. Thus the ‘ecological uncertainty’ caused by using an incorrect  $\alpha$  value varies spatially with the estimate of the ratio of call probabilities. Consider the example where the truth is a prediction model with  $\alpha_{\text{true}} = 2$ , then following cases illustrate how assuming  $\alpha = 1$  impacts predictions of catches. If  $P = 0.001$  for some location, as would occur in predominantly ENP regions, then  $\hat{C} \approx 0.999$  instead of the true value 0.998 and likewise if  $P = 1000$ , as in predominantly WNP regions, then  $\hat{C} \approx 0.01$  instead of 0.005. Thus for extreme values of  $P$  the effect of misspecifying  $\alpha$  is mitigated by the nature of the structure of the prediction model. However, if  $P = 1$  then  $\hat{C} = 1/2$  instead of  $1/3$ , so the impact of ecological uncertainty is more pronounced in regions where both populations are singing at the same rates. Due to the complex relationship between  $\alpha$  and the ratios of densities and call probabilities it is difficult to quantify and interpret  $\alpha$  in terms of different biological and anthropogenic scenarios.

The veracity of the ecological assumptions about  $\alpha$  could not be tested against data, but there was no reason to expect it would be constant across space for almost a century. I therefore conducted a sensitivity test by rerunning the above algorithm with the extra step of randomly drawing an  $\alpha$  value in step 4 instead of setting it to  $\alpha = 1$ , so that it varied between iterations but not spatially or temporally. The values for  $\alpha$  were such that  $1/2 \leq \alpha \leq 2$  and randomly drawn from a uniform distribution on the log scale so that  $\alpha = 1$  was equally likely to be low or high. No direct information was available to determine the

range for  $\alpha$ , so a fourfold range was arbitrarily used.

### 1.3 Results

AICc provided clear justification for the more complex beta-binomial structure for acoustic models as well as additive terms on parameters  $p$  and  $\sigma$  in both the ENP and WNP acoustic models (Table 1.3). Therefore the full acoustic models were selected for both ENP and WNP populations and used throughout the rest of the analysis. These model predictions depended on the month, but generally classified parts of the Gulf of Alaska, the west coast of the US, and the eastern tropical Pacific as being predominantly ENP (Figures 1.6 and 1.7).

Table 1.3: Model selection results. Models are compared across likelihoods and parameter structures. The smallest  $\Delta\text{AICc}$  indicates the model with the highest support is the most complex model. The binomial model in particular fits extremely poorly, confirming the need for an over-dispersed likelihood.

Distribution	p	Overdispersion	df	ENP $\Delta\text{AICc}$	WNP $\Delta\text{AICc}$
Beta-Binomial	Additive	Additive	40	0	0
Beta-Binomial	Additive	Linear	34	19	10
Beta-Binomial	Additive	Constant	21	33	36
Beta-Binomial	Linear	Linear	28	103	54
Binomial	Additive	N/A	20	24,070	12,675
Binomial	Linear	N/A	14	33,464	18,318

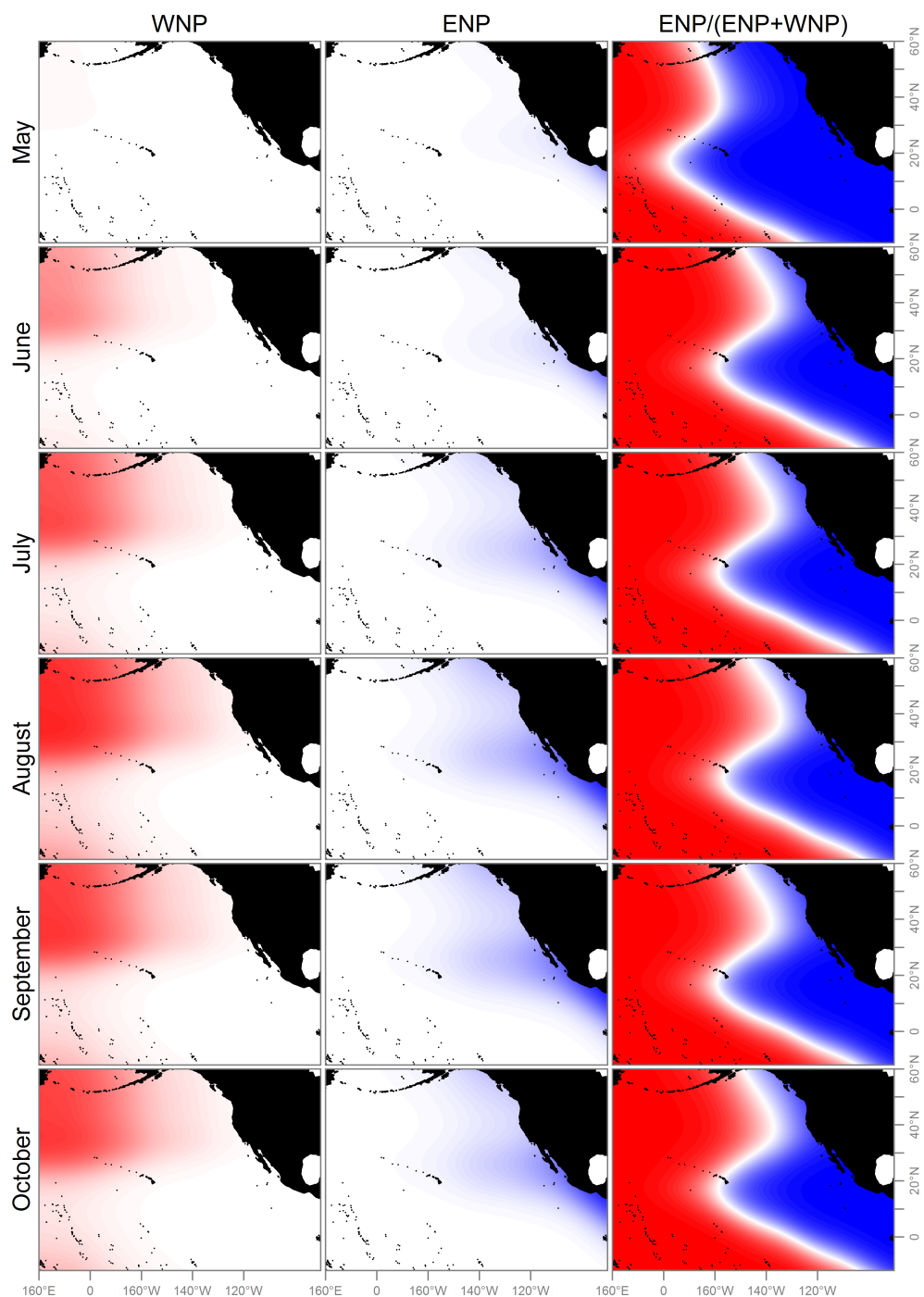


Figure 1.6: Fits for ENP acoustic, WNP acoustic, and base case prediction models for May–October. These models are fit to the acoustic data in Figures 4 and 5. The third column shows the base case prediction model which is the proportion of ENP to WNP calls, such that red areas correspond to predicted WNP occurrence and blue areas to ENP. The white line denotes where the model predicts an equal chance of observing an ENP or WNP blue whale.

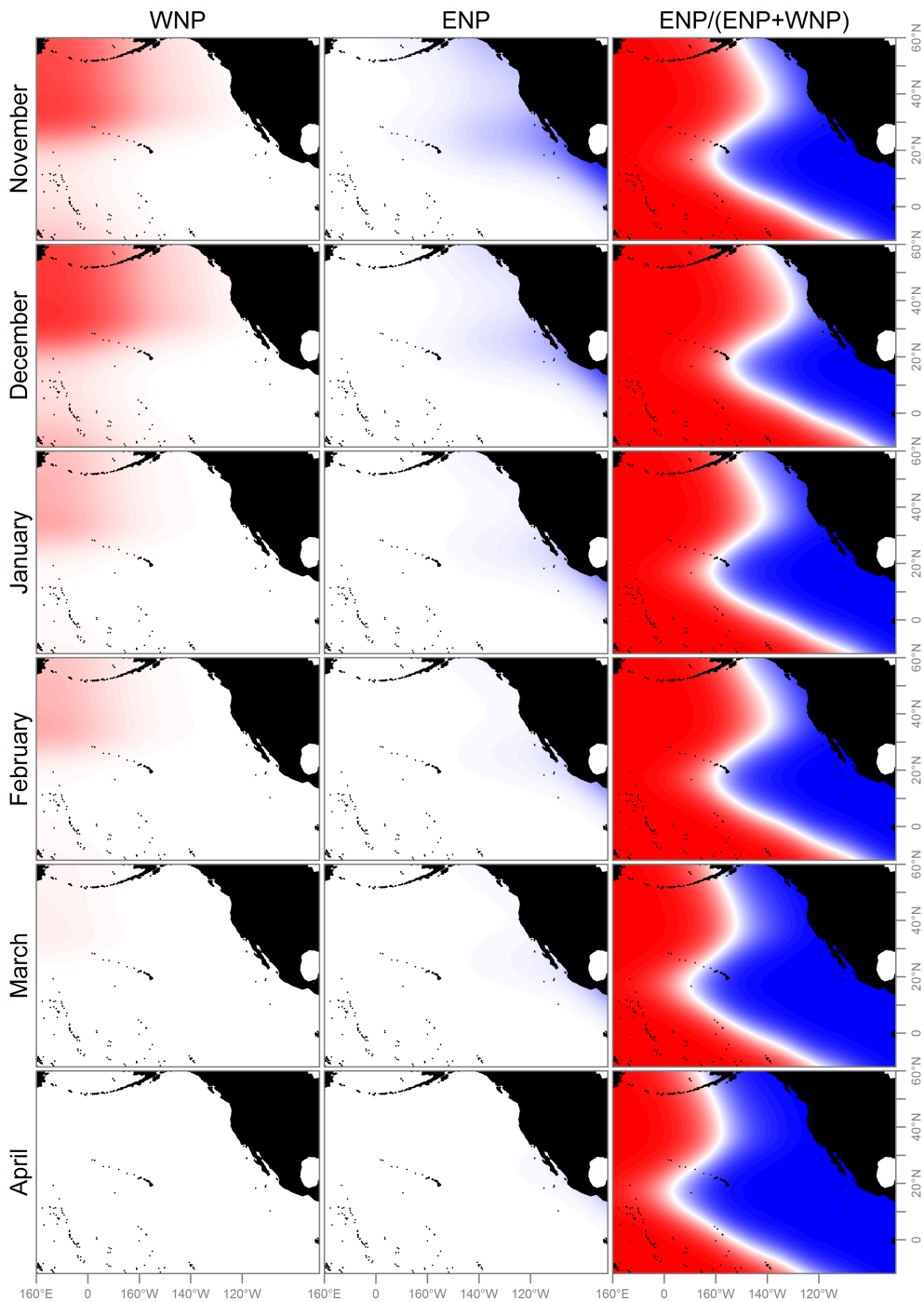


Figure 1.7: Fits for ENP acoustic, WNP acoustic, and base case prediction models for November–April. These models are fit to the acoustic data in Figures 4 and 5. The third column shows the base case prediction model which is the proportion of ENP to WNP calls, such that red areas correspond to predicted WNP occurrence and blue areas to ENP. The white line denotes where the model predicts an equal chance of observing an ENP or WNP blue whale.

The algorithm of drawing random catch series and bootstrapping was run to produce 1,000 ENP catch series. Some acoustic models failed to converged (1.56% ENP and 2.57% WNP). Since both acoustic models are needed for the prediction model, only cases where both converged for the same bootstrap sample could be used. A total of 1,089 iterations of the algorithm were needed to arrive at 1,000 successful ENP catch series.

GAMLSS model additive fits are interpreted by plotting the smoothers fit to “partial residuals” which have an arbitrary absolute scale and instead are judged via the relative change within a single independent variable (i.e. longitude, latitude, and month). In this case, increasing partial residual values indicated that the model predicted higher call rates for  $p$ , or higher variance for  $\sigma$ . For instance, both populations showed increased detected calls in the summer and fall months, ENP calls decreased from east to west, and WNP calls increased for east to west (Figures 1.8 and 1.9).

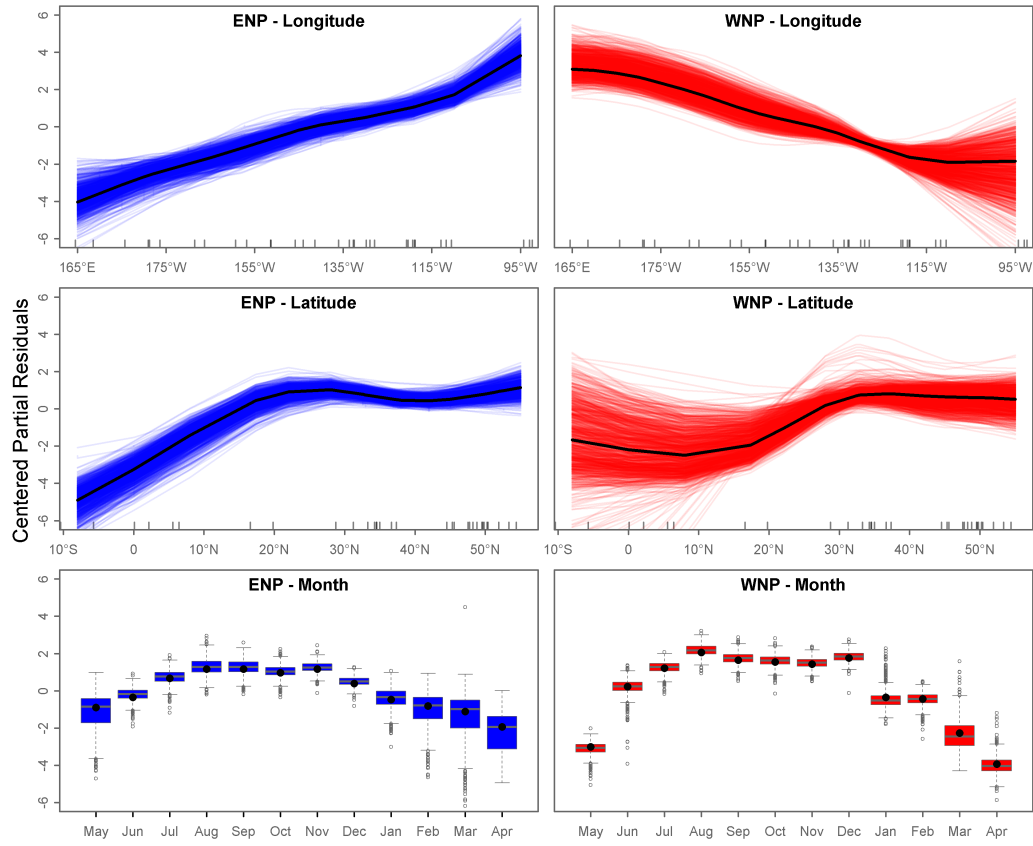


Figure 1.8: Fits of the independent variables for the probability of observing a song call ( $p$  parameter). The original model (black lines) and subsequent bootstrapped models (thin colored lines) are shown in logit space. Each panel shows the relationship between the dependent and independent variable after all other independent variables have been accounted for (i.e. the centered partial residuals). Higher relative values indicate a higher probability of observing a call. For longitude and latitude,  $x$ -axis tick marks show positions of the observed hydrophones with a small amount of noise added to prevent overplotting. See text for further discussion.

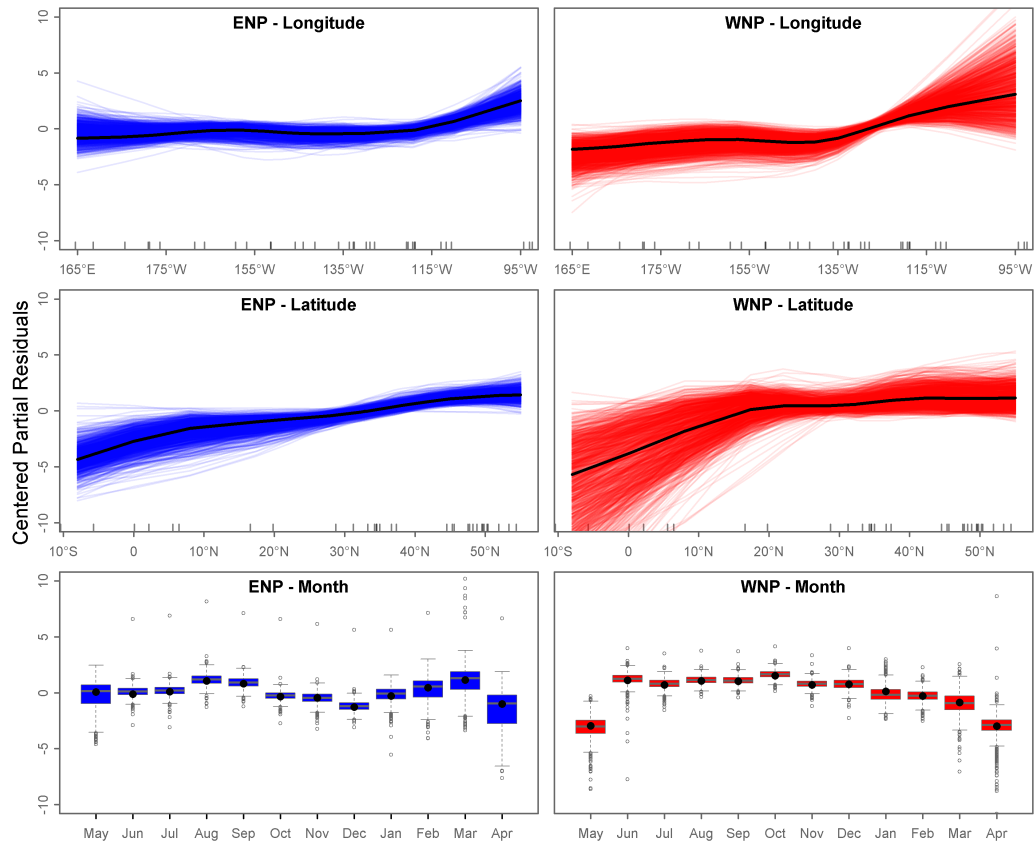


Figure 1.9: As for Figure 8 except the results pertain to the over-dispersion parameter  $\sigma$  and are in log space. Higher relative values indicate a higher level of variance expected in repeated observations.

Differences in additive fits across bootstrapped models provided a measure of statistical uncertainty in the acoustic models. There was more uncertainty in the months with fewer observed calls and regions with ENP and WNP overlap and limited hours with observed calls (Figure 1.10). For this study, the uncertainty only impacts the results where it coincides with historical catches. Thus, while there is high uncertainty around Hawaii for all months, this has little influence on the results and more acoustic data here would likely have a minimal impact. The acoustic uncertainty in the Gulf of Alaska in the summer months (the peak of catches) has the greatest influence on the uncertainty in the results.

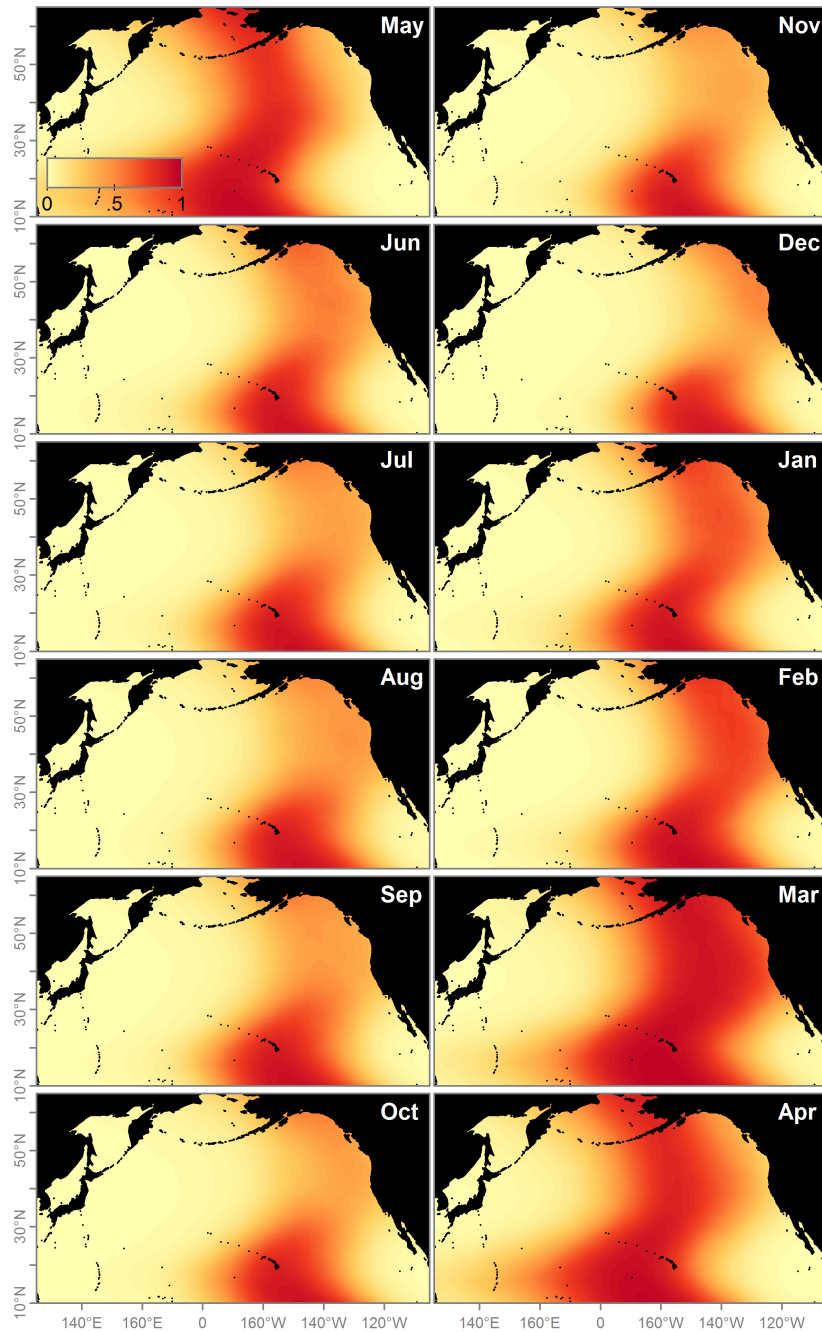


Figure 1.10: Maps showing the acoustic uncertainty arising from bootstrapping. Uncertainty in the base case probability of a catch being from the ENP population quantified using bootstrapping (i.e. resampling the acoustic data with replacement and refitting the models). The  $z$ -axis is the range of the predictions (95th minus 5th percentile) across all 1,000 bootstrapped models. A value of  $z = 0$  indicates all bootstrap models predict the same value, while  $z = 1$  indicates some bootstrap models predict 0 and others predict 1, so that there is high uncertainty at that location. There is higher uncertainty where overlap between the two populations exists (e.g. Gulf of Alaska) or there is limited data (e.g. Hawaii).

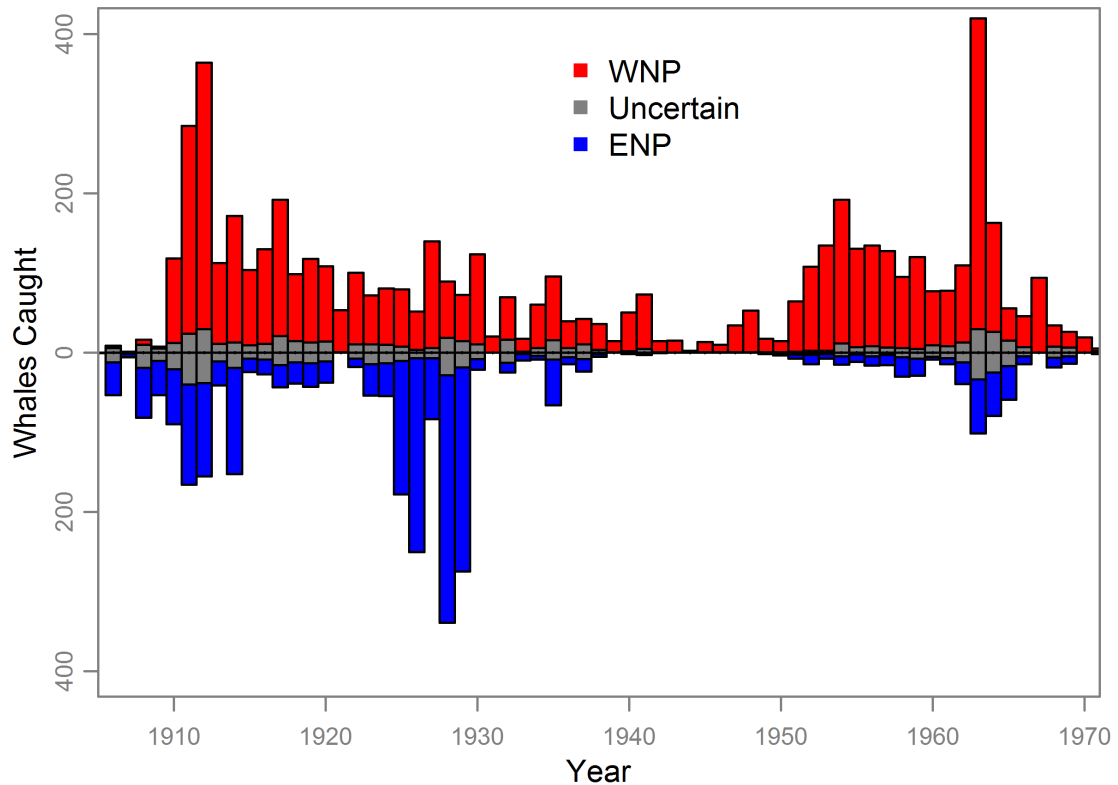


Figure 1.11: Estimated annual catches of ENP and WNP blue whales for the base case. Grey values contain 95% of the population uncertainty, and if above the line are more likely to be WNP, below the line more likely to be ENP. The uncertainty in the 1960s is caused by the unreported Soviet catches with high location uncertainty.

Each successful iteration of the algorithm provided distinct series of ENP catches and WNP catches, which were then aggregated by year (Tables 1.4 and 1.5). Most catches were taken in the early part of the century (1905–1930), with another spike in the 1960s from Soviet pelagic whaling (Figure 1.11). The level of uncertainty in which population was caught was highest during the early part of the 20th century when many catches were reported without location and/or month, and in the 1960s due to the Soviet misreported catches which contained large catch uncertainty. Examples of specific ENP catch series demonstrate the general approach implemented here, including catch and acoustic uncertainty predicted by the model (Figure 1.12). The catch uncertainty is reflected by the difference in locations of the catches between cases, and the acoustic uncertainty by the differences in

the region of uncertainty between cases.

Table 1.4: Base case ENP catches by year showing the median as well as lower and upper 95% interval from the 1,000 realizations. These catches do not include uncertainty in  $\alpha$ .

Year	2.5%	Base	97.5%	Year	2.5%	Base	97.5%	Year	2.5%	Base	97.5%
1905	1	1	1	1928	310	339	358	1951	5	8	9
1906	41	54	59	1929	256	275	289	1952	11	14	17
1907	4	6	6	1930	13	22	32	1953	6	7	10
1908	63	82	92	1931	0	0	0	1954	10	15	27
1909	43	53	58	1932	13	25	41	1955	9	12	19
1910	69	90	102	1933	8	10	11	1956	12	16	24
1911	126	166	189	1934	5	9	15	1957	12	16	22
1912	117	155	185	1935	57	66	81	1958	24	30	35
1913	31	41	52	1936	9	15	21	1959	21	29	34
1914	133	152	165	1937	16	24	34	1960	4	9	18
1915	17	25	34	1938	4	5	9	1961	8	15	23
1916	18	27	38	1939	0	0	1	1962	28	40	52
1917	28	44	64	1940	1	2	4	1963	68	101	131
1918	26	39	53	1941	1	3	7	1964	54	79	105
1919	30	43	56	1942	1	1	1	1965	42	59	75
1920	27	38	52	1943	0	0	0	1966	9	14	21
1921	0	0	0	1944	0	0	0	1967	0	0	0
1922	10	18	29	1945	0	0	0	1968	12	19	26
1923	40	54	64	1946	0	0	1	1969	10	14	20
1924	41	55	65	1947	0	0	1	1970	0	0	0
1925	167	178	185	1948	0	0	1	1971	1	2	4
1926	244	251	254	1949	1	2	2	Total	2396	2955	3477
1927	76	83	89	1950	3	3	4				

Table 1.5: Base case WNP catches by year showing the median as well as lower and upper 95% interval from the 1,000 realizations. These catches do not include uncertainty in  $\alpha$ .

Year	2.5%	Base	97.5%	Year	2.5%	Base	97.5%	Year	2.5%	Base	97.5%
1905	0	0	0	1928	70	89	118	1951	63	64	67
1906	3	8	21	1929	58	72	91	1952	105	108	111
1907	1	1	3	1930	113	123	132	1953	132	135	136
1908	6	16	35	1931	20	20	20	1954	180	192	197
1909	3	8	18	1932	54	70	82	1955	123	130	133
1910	106	118	139	1933	16	17	19	1956	127	135	139
1911	261	284	324	1934	54	60	64	1957	121	127	131
1912	334	364	402	1935	81	96	105	1958	90	95	101
1913	102	113	123	1936	33	39	45	1959	115	120	128
1914	159	172	191	1937	32	42	50	1960	68	77	82
1915	94	103	111	1938	32	36	37	1961	69	77	84
1916	119	130	139	1939	14	15	15	1962	97	109	121
1917	171	191	207	1940	48	50	51	1963	390	420	453
1918	84	98	111	1941	69	73	75	1964	137	163	188
1919	105	118	131	1942	14	14	14	1965	40	56	73
1920	94	108	119	1943	15	15	15	1966	39	46	51
1921	53	53	53	1944	2	2	2	1967	94	94	94
1922	89	100	108	1945	13	13	13	1968	27	34	41
1923	62	72	86	1946	9	10	10	1969	20	26	30
1924	70	80	94	1947	33	34	34	1970	19	19	19
1925	72	79	90	1948	52	53	53	1971	3	5	6
1926	48	51	58	1949	17	17	18	Total	5092	5614	6173
1927	134	140	147	1950	14	15	15				

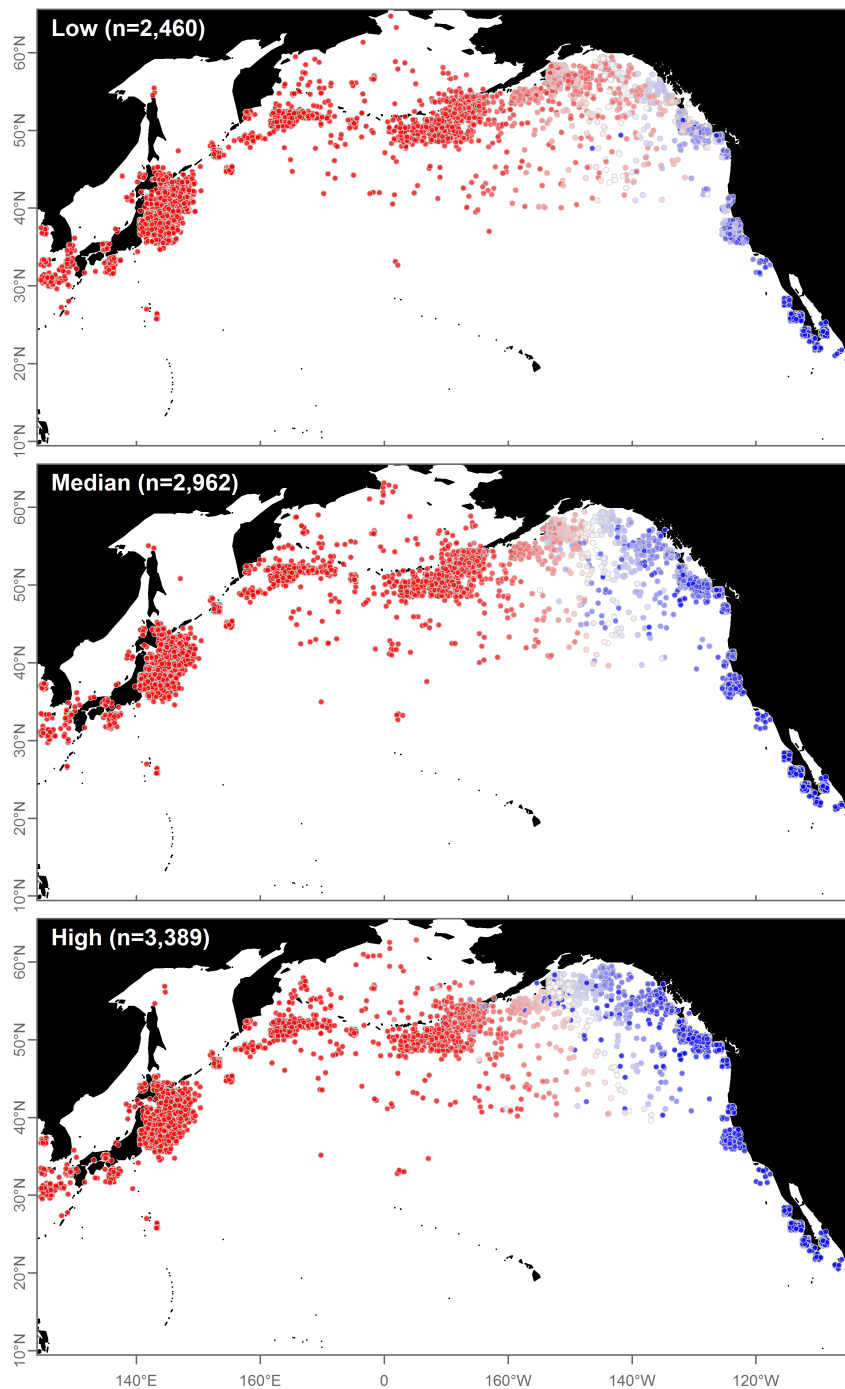


Figure 1.12: Example model predictions. Catches of ENP (blue) and WNP (red) blue whales for three realizations chosen to represent low, median, and high total ENP catches. Catch positions are plotted with a small amount of noise to prevent overplotting, and are different between the three cases because they are drawn randomly from potential locations. Nearby catches can have different predictions (colors) because months are aggregated.

My methods allowed uncertainty to be estimated for the catch series, and incorporated into the results. The largest source of uncertainty for the base case was from bootstrapping the acoustic data. The catch uncertainty was relatively small compared to the acoustic uncertainty, thus additional acoustic data has the potential to decrease the uncertainty in the results the most (Table 1.6). The ecological sensitivity run where  $\alpha$  ranged uniformly between 0.5 and 2 contributed more uncertainty than catch uncertainty, but less than acoustic uncertainty (Table 1.6). As expected a higher value of  $\alpha$  for all algorithm iterations lead to smaller estimates for the total ENP catches (Figure 1.13).

Table 1.6: Comparison of the sources of uncertainty in total ENP catches. Catch uncertainty arises from missing locations and dates, acoustic uncertainty in the song call occurrence data, and ecological from uncertainty in ecological assumptions over the range  $1/2 \leq \alpha \leq 2$ . Differences in totals compared to Table 4 are due to rounding and the order of percentile calculations.

<b>Uncertainty Source</b>	<b>2.5%</b>	<b>Median</b>	<b>97.5%</b>	<b>Cumulative % uncertainty</b>
Catch	2,948	2,962	2,977	2%
Catch + Acoustic	2,460	2,963	3,389	77%
Catch + Acoustic + Ecological	2,329	2,936	3,533	100%

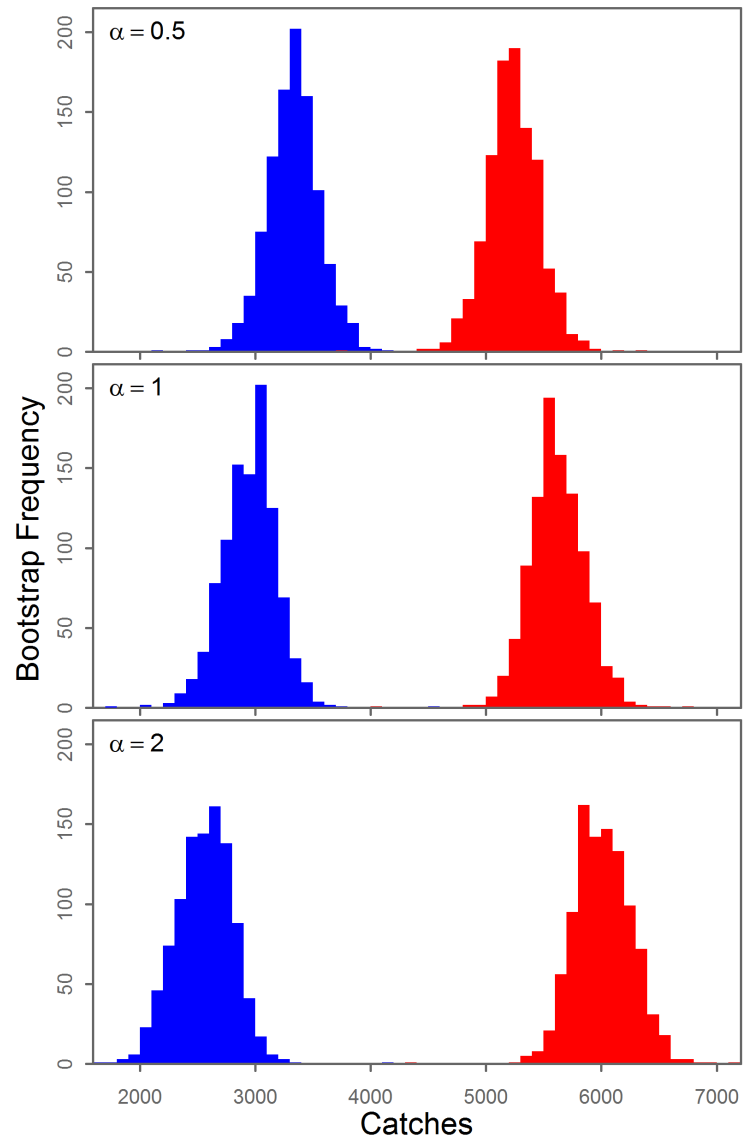


Figure 1.13: Results of the sensitivity analysis to the conversion between song call frequency and density of whales. The sensitivity to the base case assumption that  $\alpha = 1$  to the final results is explored with histograms of the total catches of the ENP (blue) and WNP (red) populations for three levels of  $\alpha$ . The parameter  $\alpha$  is the ratio of the unknown WNP and ENP factors which convert density of whales to song call occurrence.

I estimate that the total number of ENP whales caught from 1905–1971 in the North Pacific was 2,963 (95% range 2,460 to 3,389) based on catch and acoustic uncertainty. For the sensitivity case with ecological uncertainty included, the total number of ENP blue whales

caught was 2,969 (95% range 2,301 to 3,516). Thus, the inclusion of ecological uncertainty increases the width of the 95% interval by 34% if  $\alpha$  is assumed to range uniformly from 0.5 to 2.

#### **1.4 Discussion**

This study collates and details North Pacific blue whale catches by population for the first time. I updated catch totals to include previously unreported Soviet catches, and estimated where and when those blue whales were caught, using the best available information. The additional Soviet catches accounted for only approximately 4.5% of the total catches, but are expected to have a large influence on the current status of blue whales (Chapter 2) since they occurred at the end of the whaling period (in the 1960s and 1970s) and had uncertain locations and dates. Using original whaling records to correct unreliable IWC catch records is critical to understanding the true whaling history of the region and has important implications for the management of blue whales.

I attributed 2,963 (95% range 2,460 to 3,389) catches of North Pacific blue whales to the ENP population, or roughly 35% of all blue whales caught from 1905–1971, suggesting that the WNP population was likely substantially larger than the ENP population before whaling started, assuming all WNP catches determined here are from one population. Current abundance estimates are not available for any region within the WNP, so it is unclear whether the WNP blue whales are currently more depleted than ENP blue whales.

My results are based on some ecological assumptions about movement patterns and acoustic behavior of the two populations. First, I assumed that the ENP and WNP populations had stable movement patterns over the entire span of the catches and acoustic recordings (1905–2004), despite intensive whaling and varying oceanographic conditions. If ample acoustic data were available across many years, the effect of year could be quantified within the modeling framework provided here, but the acoustic data were collected in different time periods (Table 1.2), precluding such an analysis in this study. High maternal site fidelity in the Gulf of California has been shown for the ENP population, supporting the idea of spatial continuity in a breeding and calving region [?], but another study speculates a range shift into the Gulf of Alaska due to changes in abundance or oceanographic

conditions [?]. It seems reasonable that slight changes in ranges for the ENP, particularly in the Gulf of Alaska are possible, but that no major range shifts have occurred. For the WNP there is no information with which to judge how many populations there are, or their spatial fidelity.

Second I assumed that the ratio of calling rates was proportional to the ratio of densities of the two populations across time and space, but there was no way to directly test the accuracy of this hypothesis. The relationship between singing and density is likely complex, especially if it is related to breeding. One study suggests that ENP blue whales arriving at feeding grounds may delay singing for weeks until adequate food has been obtained [?] and another in the same area found the proportion of whales singing increased during the feeding season [?]. Thus the relationship between density and ENP singing is weak for part of the feeding season and suggests the occurrence patterns observed in the data may also reflect behavior in addition to presence of whales. This short-term effect is mitigated to some degree by binning the data into months. Little is known about the seasonal behavior of ENP blue whales, and none about WNP whales, so I could not explicitly include this in my analysis. The assumption may also be incorrect if there were a difference in the sex ratios or sex-specific mixing between populations. The historical catch records clearly indicated that both sexes were caught in similar ratios throughout the region modeled, so this is not believed to be a major issue in the analysis. In addition, the two populations could vocalize at different rates. Since they have similar morphologies, call types, seasonal pattern of call production, and life history strategies it is unlikely that the behavioral context of call production differs greatly between populations.

The veracity of the ecological assumptions discussed here was, and may remain, unknown. However, when I tested the influence of this factor in a sensitivity test allowing  $\alpha$  to range from 0.5–2.0 (a fourfold range), the 95% confidence interval in the total ENP catches widened by only 35%. This suggests that the results are fairly robust to these assumptions.

Although the focus of this study was ENP blue whales, by examining all catch and acoustic data, new information about WNP blue whale exploitation emerged. I estimated 5,606 (95% range 5,180 to 6,109) or roughly 65% of blue whale catches in the North Pacific were likely from the WNP population. Unlike ENP blue whales the current abundance,

status, and geographic range of these whales remains unknown. At present there is no dedicated effort to study these whales and acoustic data provide the only information as to when and where this population of blue whales is found in the North Pacific.

The conclusions presented here are dependent on the assumptions made in the analysis. Due to the nature of predicting historical catches the predictors for acoustic models were limited to location and time. However, there are likely many predictors that would help further explain the occurrence of blue whale song calls (e.g., [?]). The focus of this study was to estimate catches, but future studies investigating occurrence patterns could incorporate additional predictors into the model framework used here.

This study provides a new framework for differentiating spatial occurrence of populations of cetaceans using their distinct call types with data that are relatively easy to collect in comparison to broad scale sighting surveys, photo identification, and genetic biopsy sampling efforts. In doing so I provide the first estimate of catches for ENP blue whales that is based on data and includes uncertainty. Future work should combine these results with current abundance estimates to formally assess the recovery of the population and the potential risk of continued anthropogenic mortalities.

## Chapter 2

**DO SHIP STRIKES THREATEN THE RECOVERY OF  
ENDANGERED EASTERN NORTH PACIFIC BLUE WHALES?****2.1 Introduction**

Blue whales (*Balaenoptera musculus*) are distributed throughout the world's oceans. From the early 1900s to the 1970s they were targeted by commercial whaling operations, and this depleted the Antarctic [?] and likely other populations. In the North Pacific, whalers caught 8,569 blue whales from 1905–1971 (Chapter 1) but their recovery and current status is unknown, primarily because the population structure of blue whales in the North Pacific is unclear, and population assessments based on population modeling have not been conducted. When conducting assessments it is important to consider distinct breeding populations.

Although the International Whaling Commission (IWC) has only formally considered one population of blue whales in the North Pacific [?] there is evidence for a more complex population structure. Two populations in particular have been proposed: the eastern (ENP) and western (WNP) North Pacific blue whale populations, although some have speculated that up to 5 populations exist. [?].

Acoustic calls differ between the eastern and western North Pacific populations and were used to split catches in Chapter 1. A distinct eastern population is corroborated by significant differences in length frequencies between whales in the eastern and western North Pacific [?], distinct song call types [?], and movements from the eastern Tropical Pacific to the Gulf of Alaska using tags and photographic ID [?, ?, ?]. There are also genetic similarities between blue whales in the Gulf of California and the U.S. West Coast [?], suggesting that a single population occupies the entire eastern North Pacific, although no genetic analyses have been performed on whales in other regions of the North Pacific. While there is substantial evidence for the ENP population, the population structure of the

central and western North Pacific is less clear. In this study I assume an ENP population and focus exclusively on assessing their population trends, status and recovery from commercial whaling.

A series of abundance estimates for the ENP population from 1934–2008 [?, ?, ?] show no clear increase in abundance, despite commercial whaling ending in 1971 [?]. Although whaling has ended, there are still sources of anthropogenic mortality that may be affecting the recovery of the population.

The only known source of anthropogenic mortality to ENP blue whales is fatal collisions between whales and vessels ('ship strikes') along the U.S. West Coast where shipping lanes intersect with common feeding areas [?, ?]. An average of 1.8 blue whale strikes per year (2005–2010) were documented in the Santa Barbara Channel [?], as well as 2 observed off of Washington during 1980–2005 [?]. These are all blue whale ship strikes recorded on the U.S. West Coast during this period. However, these documented values are likely underestimates of the actual ship strike rates because struck whales may go unreported or misidentified for a variety of reasons. For instance, a collision may only cause internal damage and go unrecognized as a ship strike or the carcass could sink before being identified or noticed [?]. A study on North Atlantic right whales estimated a 17% detection rate for ship strike mortalities [?], which was then applied to Santa Barbara Channel blue whale ship strikes to estimate 10.6 whales per year killed in 2013 [?]. This estimate is well above the "potential biological removal" of 3.1 whales per year [?] and thus likely exceeds allowable levels established under U.S. laws. These observations led to speculation that ship strike mortalities may be limiting the recovery of the ENP population and led to calls for further research into potential approaches for mitigation [?, ?].

Blue whales are not the only species impacted by collisions with vessels, and a variety of mitigation techniques have been proposed for other populations such as onboard observers, remote sensing, passive acoustic monitoring to better detect and avoid whales, and Marine Protected Areas [?, ?, ?]. Speed restrictions and shifts in shipping lanes have been implemented for critically endangered populations of other species and have been shown to be effective in reducing the risk of collisions for North Atlantic right whales [?, ?, ?]. Voluntary speed restrictions were implemented in southern California, but have been ineffective

in reducing vessel speeds [?], while moving the shipping lanes in the Santa Barbara Channel would have a limited impact on the rate of blue whale and vessel interactions because the whales are evenly dispersed [?]. Other measures, including those proposed above, could reduce whale-vessel interactions in the future, and I aimed to broadly test the impact of potential mitigation scenarios on the recovery of the ENP blue whale population.

The hypothesis that ship strikes limit the recovery of the ENP blue whale population has not been tested. In this thesis I explored the ship strike hypothesis versus the hypothesis that the recovery rate slowed due to density dependence effects (i.e. approaching carrying capacity). I used population dynamics models fitted to estimates of catches, abundances, and ship strikes in order to estimate the abundance trajectory of this population and quantitatively explore the two hypotheses. Furthermore, I quantify the impact of future ship strikes under different potential mitigation scenarios to estimate the risk posed by an increasing level of ship traffic in their feeding areas off the U.S. West Coast.

## 2.2 Methods

I modeled vessel numbers, ship strikes, and the population dynamics of the ENP blue whale population. The analyses were conducted in a Bayesian framework so I could incorporate informative biological information and assign probabilities to important values such as relative abundance and pre-exploitation abundance. Several other cetacean assessments have been conducted in a Bayesian framework [?, ?, ?, ?]

### 2.2.1 Population Dynamics

There are two key sets of data needed to estimate past and future trends in abundance: estimates of abundance and historical anthropogenic mortalities. In combination with an assumed population dynamics model, these data provide information on the pre-exploitation abundance and intrinsic growth rate of the population [?]. The main population dynamics model used in this thesis is the discrete theta-logistic model which has been presented to the IWC [?] for other exploited cetacean populations:

$$N_{y+1} = N_y + rN_y \left( 1 - \left( \frac{N_y}{K} \right)^z \right) - C_y - S_y \quad (2.1)$$

where  $N_y$  is the estimated population size in year  $y$  (1905–2050),  $r$  is the estimated intrinsic growth rate parameter,  $K$  the estimated carrying capacity parameter,  $z = 2.39$ ,  $C_y$  is the annual catches, and  $S_y$  is the estimated number of ship strikes, which are subtracted in the same way as the commercial catches and whose calculation is detailed below. Setting  $z = 2.39$  is common for recovering cetacean populations [?] and causes the model to have maximum productivity at 60% of  $K$ , as opposed to 50% when  $z = 1$ , as in the discrete logistic model.

As is typical of analyses with the theta-logistic model, the non-linear correlation between estimates of  $r$  and  $K$  presented numerical convergence issues when constructing the Bayesian posteriors (e.g. [?]). I therefore re-parameterized equation (2.1) by utilizing the fact that  $MSY = rKz/(z + 1)^{1/z+1}$  and setting  $K = MSY(z + 1)^{1/z+1}/rz$ . Thus the parameters of the model were  $r$  and  $MSY$ , so equation (2.1) became

$$N_{y+1} = N_y + rN_y \left( 1 - \left( \frac{rzN_y}{MSY(z + 1)^{1/z+1}} \right)^z \right) - C_y - S_y \quad (2.2)$$

and given  $r$  and  $MSY$  the value of  $K$  could be calculated. I ignored process error and assumed negligible catches before 1905 (there were none reported to the IWC) due to the difficulty in catching such large and negatively buoyant whales, such that the population was at carrying capacity previous to commercial exploitation (i.e.  $N_{y < 1905} = K$ ). This population dynamics model incorporates catches, ship strikes, and density dependence, which allowed the past, present and future status of the population to be estimated.

Due to speculation about the evidence for increasing population size over the period of the abundance estimates (1993–2008) [?] I also fit a second model which addressed this question more directly than equation (2.2). I estimated the rate of change over this period by fitting a log-linear model over years  $y$  in 1993–2008:

$$\ln(N_y) = \ln(N_{1993}) + m(y - 1993) + \epsilon_y \quad (2.3)$$

where  $N_{1993}$  is the population size in the year of the first abundance estimate and  $m$  is a growth rate term, and  $\epsilon_y \sim N(0, \sigma_y^2)$  is the assumed error term. The log-linear model does not incorporate density dependence or catches so it cannot be used to estimate the status as is done with the population dynamics model in equation (2.2).

The first set of data used to fit the population dynamics and log-linear model is abundance estimates. For this population of whales there are five published estimates of absolute abundance (i.e. numbers of whales) using photographic mark-recapture techniques from 1993–2008 taken offshore of the US west coast (Table 2.1) [?, ?]. These estimates are assumed to be independent and log-normally distributed with likelihood

$$L_{\text{abundance}} = \prod_{y=1}^5 \frac{1}{N_y^{\text{obs}} \sqrt{2\pi\sigma_y^2}} \cdot \exp \left[ -\frac{(\ln(N_y^{\text{obs}}) - \ln(N_y))^2}{2\sigma_y^2} \right] \quad (2.4)$$

where  $\sigma_y^2 = \ln(CV_y^2 + 1)$  is the variance on the log scale for year  $y$ ,  $N_y^{\text{obs}}$  is the observed absolute abundance and  $N_y$  is the predicted abundance from a population dynamics model (equation (2.2)). Line-transect estimates were also available for this population, but they are considered to measure a variable and biased subset of the population depending on the year [?]. For this reason I excluded them from the analysis and used only the mark-recapture estimates.

Table 2.1: Photographic mark-recapture estimates used to fit the population dynamics models. These are assumed to be log-normally distributed and independent.

Year	Estimate	CV
1993	2000	0.20
1997	1756	0.21
2002	1781	0.23
2006	2842	0.29
2008	2497	0.18

The second set of data needed to fit the models is the number of anthropogenic mortalities for each year. These come in two groups: historical commercial catches (“catches”) and fatal ship strikes (“strikes”). In Chapter 1 I estimated catch time series from 1905–1971 (and assumed they were 0 for all other years) for the ENP blue whale population based on historical catches split between the ENP and WNP populations using acoustic call occurrence differences. The resulting estimates of the number of ENP blue whales caught in

each year by commercial whaling operations (“catch series”) were used in the population dynamics model (but not the log-linear). My analysis in Chapter 1 used bootstrapping and Monte Carlo methods to quantify the uncertainty associated with the catches and thus provided many realizations of catch series, each with different annual ENP catches. There was no uncertainty within a single series; rather the uncertainty in historical catches arose from integrating across the different catch series. I therefore used 1000 catch series to incorporate the uncertainty in catches into the model fits and estimates of current and future status of the population. There was no time series of fatal ship strikes over the time period of the population dynamics model, so these were estimated explicitly within the model framework here, and are described in detail in the next section.

One reason why the Bayesian approach was adopted for this analysis was that information on the intrinsic growth rate is available from other recovering populations of baleen whales. In the Bayesian framework this information can easily be included in the form of an informative prior distribution on the parameter  $r$ . Two priors were considered for  $r$ : (1) “uninformative”  $r \sim \text{Uniform}(0, r_{\max})$ , where  $r_{\max} = 0.114$  is a biological upper bound for the intrinsic growth rate for blue whales [?], and (2) an “informative” prior based on a posterior predictive distribution from a meta-analysis of  $r$  from [?]. The meta-analysis was on the ratio of  $r$  to the theoretical biological maximum  $r_{\max}$ . This prior distribution was converted to the intrinsic rate of increase  $r$  by multiplying by  $r_{\max}$ . A high-order polynomial was fitted to an empirical cumulative distribution and then differentiated analytically to determine a prior probability density for the parameter (Figure 2.1, see Appendix A for code). There was no prior information on  $K$  so I used a uniform prior:  $K \sim \text{Uniform}(0, 6000)$ , the upper bound reflecting the fact that total catches were at most 3,389 (Chapter 1) and the most recent abundance is 2,497. Therefore the extreme case of  $r = 0$  suggests the carrying capacity cannot be more than  $3,389 + 2,497 = 5,886$ . I adopted the uninformative prior on  $r$  as a “base case,” but ran the analysis with both priors on  $r$  and compared the differences in results as a sensitivity analysis.

For the log-linear model (equation (2.3)) I assigned uninformative priors to both parameters with uniform distributions:  $m \sim \text{Uniform}(-.2, .2)$  and  $\ln(N_{1993}) \sim \text{Uniform}(0, 20)$ , chosen so they were broad enough to minimize the impact of the prior on the posterior.

The probability of continued recovery (i.e. positive growth) was then estimated by the percentage of the posterior distribution for  $m$  greater than 0.

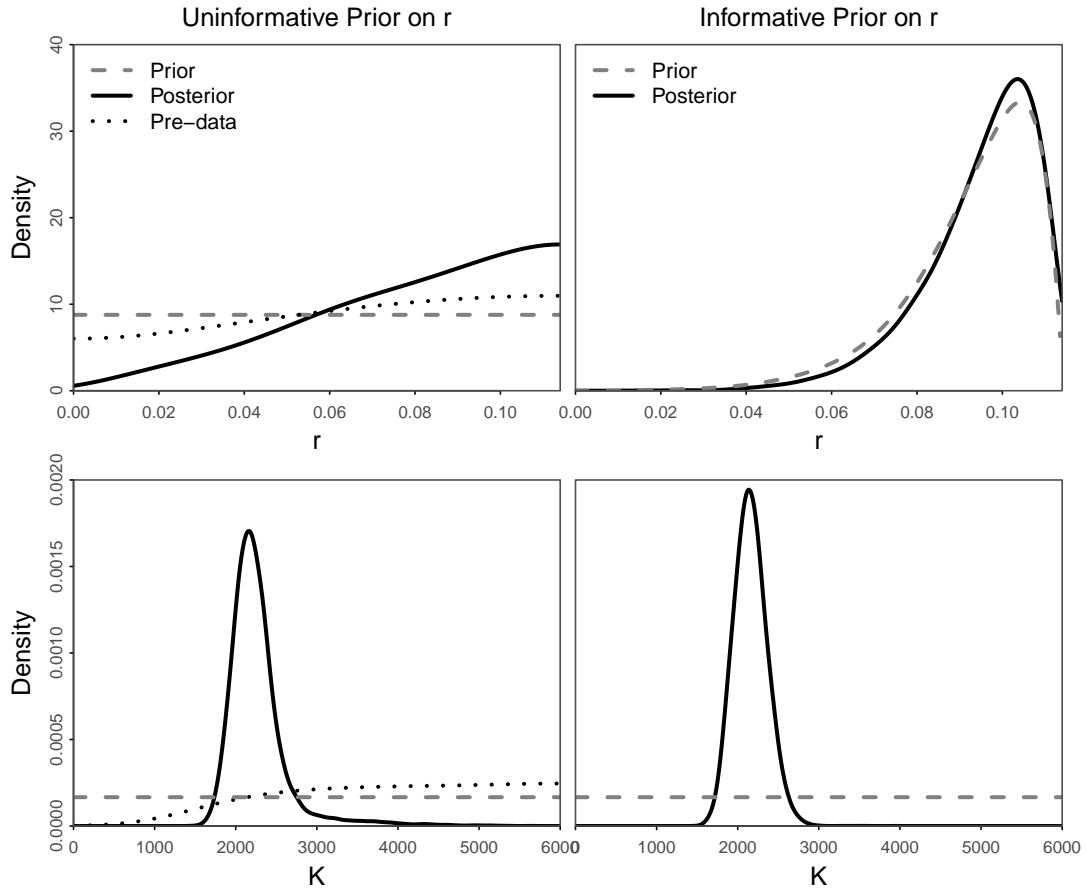


Figure 2.1: Population dynamics parameter distributions for  $S_{2013}^{\text{obs}} = 10$  and both priors on  $r$ .

The impact of the data on the posterior was examined by comparing “post-model-pre-data” distributions, which includes only parameter combinations consistent with the population dynamics model (i.e. not parameters where the population goes extinct before the present) [?], to standard “post-data” posteriors. Differences between the post-model-pre-data and post-data posteriors demonstrate the amount of information contained in the abundance data.

The two models used in this study provided different information relevant to the population trends of the ENP blue whale. The log-linear model estimated the rate of population change from 1993–2008 despite ship strikes. The logistic model provided estimates of the carrying capacity of the population, from which historical abundance trends and present and future status could be determined.

### 2.2.2 Ship Strikes

Unlike the catches, a time series of ship strikes from 1905–2012 was not available. Leaving ship strikes out of an assessment would have underestimated the total number of whales killed in each year and biased estimates of rate of increase and status relative to carrying capacity. Thus, I estimated the annual number of fatally-struck whales for each year and incorporated them into the model (2.2). The estimate of 10.6 ship strikes in 2013 applied only to the Santa Barbara Channel and does not account for whales struck outside of that region so the true number for the entire habitat of the eastern North Pacific blue whales is likely higher. However, very few observed ship strikes are outside the Santa Barbara Channel so ship strikes in other regions are assumed to be minimal. I estimated the sensitivity of the results to this assumption by running scenarios where  $S_{2013}^{\text{obs}} = 10$  whales per year was considered the base case, and another with  $S_{2013}^{\text{obs}} = 20$  as a high case.

Model (2.2) requires estimates of historical ship strikes (1905–2012) to fit the model and establish the current status while estimating the future status, required projecting strikes into the future. No reliable ship strike data were available so I explicitly modeled the incidence of ship strikes over time and then used these estimates in the population dynamics model. I assumed ship strikes were a function of a constant of proportionality, the number of whales, and the number, size and speed of vessels in the shipping lanes. For example one would expect few strikes if the abundance was high but vessel traffic was minimal (e.g. pre-exploitation) or if the abundance was low and vessel traffic high (e.g. ship strikes depleting the population). I estimated ship strikes each year as:

$$S_y = cN_yV_y \quad (2.5)$$

where  $S_y$  is the estimated ship strikes in year  $y$ ,  $N_y$  the estimated population size,  $V_y$  an

estimated proxy for the number of vessels (defined below), and  $c$  was an estimated parameter which scaled to estimated ship strikes. I further assumed that observed ship strikes in 2013,  $S_{2013}^{\text{obs}}$ , were Poisson distributed with mean  $S_{2013}$  so that the likelihood is:

$$L_{\text{strikes}} = L(S_{2013}|S_{2013}^{\text{obs}}) = \frac{(S_{2013})^{S_{2013}^{\text{obs}}} e^{-S_{2013}}}{S_{2013}^{\text{obs}}!} \quad (2.6)$$

No prior information was available for  $c$  so a broad, uninformative prior was used:  $c \sim \text{Uniform}(0, 0.01)$ . For example if there were 2,000 whales and 100 vessels (in tens of thousands), then ship strikes would range from 0 to 2,000 as  $c$  ranged from 0 to 0.01.

Table 2.2: Historical vessel traffic. Data on vessels over 1000 gross tons worldwide used in the vessel model to infer changes in vessel traffic over time.

<b>Year</b>	<b>Vessels (<math>\times 10^4</math>)</b>	<b>Year</b>	<b>Vessels (<math>\times 10^4</math>)</b>
1890	11.1	1950	30.9
1900	15.9	1960	36.3
1910	22.5	1970	52.4
1920	26.5	1980	79.8
1930	30.0	1990	78.4
1939	29.7	2003	89.9

Estimating the strikes in year  $y$  required an estimate of the number of vessels,  $V_y$ . Here I defined “vessel” as a ship moving through ENP blue whale territory that was large and fast enough that a collision would be fatal. The historical vessel traffic in the eastern North Pacific was unavailable but the number of vessels worldwide larger than 100 gross tons were available from 1890–2003 (as used in [?]) (Table 2.2). I assumed the worldwide vessel traffic was proportional to vessel traffic in the extent of ENP blue whale territory and that vessel traffic followed a logistic function with time based on two arguments. First, worldwide collisions with all cetaceans were infrequent until the 1950s and then increased noticeably until current levels [?]. Ship strikes during this period were sensational to the public and thus likely to have been reported if they occurred at higher rates [?]. Second, future growth is unlikely to grow exponentially due to logistical and safety issues with increasing numbers

of vessels in a limited spatial region. I therefore fit a discrete logistic function to the available vessel data:

$$V_y = V_{y-1} + r_v V_{y-1} \left( 1 - \frac{V_{y-1}}{K_v} \right) \quad (2.7)$$

Where  $V_y$  is the predicted number of vessels in year  $y$ ,  $r_v$  the rate of growth of vessels, and  $K_v$  the asymptote. This model also needs a parameter to initialize the vessel trajectory. I parameterized the model with  $V_{1950}$  which minimized the parameter correlation and helped to improve the numerical efficiency of the fitting process and reduce the runtime of the analysis. Years after 1950 were projected forward using equation (2.7) and a set of parameters  $r_v$ ,  $K_v$ , and  $V_{1950}$ . Years before 1950 were calculated backward given the parameters and vessels in the following year using the formula:

$$V_y = \frac{-(1 + r_v) + \sqrt{(1 + r_v)^2 - 4r_v V_{y+1}/K_v}}{-2r_v/K_v} \quad (2.8)$$

I further assumed the errors were independent and normally distributed with an unknown variance parameter, so the likelihood of the vessel model was:

$$L_{\text{vessels}} = \prod_y \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp \left[ -\frac{(V_y - V_y^{\text{obs}})^2}{2\sigma^2} \right], \quad (2.9)$$

where  $V_y^{\text{obs}}$  is the observed vessels, and  $\sigma^2$  is the estimated error variance. Broad, uniform priors were assumed for  $r_v$ ,  $K_v$  and  $V_{1950}$ :  $r_v \sim \text{Uniform}(0, .15)$ ,  $K_v \sim \text{Uniform}(90, 500)$  and  $V_{1950} \sim \text{Uniform}(1, 50)$ , where  $V_{1950}^{\text{obs}} = 30.9$ . I used a weakly informative half-Cauchy prior for the standard deviation  $\sigma \sim \text{Half-Cauchy}(500)$ , such that the density was  $f(\sigma) = \frac{1}{\pi} \left( \frac{500}{\sigma^2 + 500^2} \right)$ , based on suggestions in [?].

### 2.2.3 Model Fitting

The population dynamics model (2.2) relies on the estimated ship strikes, which in turn depends on the vessel model. I combined them together into a single model with a joint likelihood and estimated all parameters in two steps: (1) project the vessel trajectory ( $V_y$ , equation (2.7)) and (2) project the abundance for each year ( $N_y$ , equation (2.2)) using the estimated ship strikes, catches, and abundance from the previous year. The joint likelihood

is:

$$L_{\text{joint}} = L_{\text{abundance}} \cdot L_{\text{strikes}} \cdot L_{\text{vessels}} \quad (2.10)$$

The joint model had 7 estimated parameters:  $r$  and  $MSY$  for the abundance model;  $c$  in the strikes model; and  $r_v$ ,  $K_v$ ,  $V_{1950}$ , and  $\sigma^2$  in the vessel model. The log-linear and joint model were coded in the modeling framework AD Model Builder [?] and posterior distributions for all parameters and derived quantities of interest were estimated using the numerical sampling technique Markov chain Monte Carlo (MCMC). For the joint model, each of the 1000 catch series used resulted in a unique joint posterior for each combination of the two  $r$  priors and two  $S_{2013}^{\text{obs}}$  values (i.e. a “scenario”).

I ran long MCMC chains (2.5 million samples) for all catch series and scenarios and determined that a burn-in of 5000 and thinning rate of 1 in 5000 ensured a converged chain producing independent samples from the posterior. Such a large sample size made plotting and analysis difficult so I used the same rates and sampled 20 new draws from each of the 1000 catch posteriors and pooled them together to form a full posterior with 20,000 samples. By doing this I implicitly assumed that all catch series were equally probable. All MCMC chains were run using AD Model Builder’s built-in Metropolis algorithm [?] and results were imported into the statistical environment R [?] for further analysis and plotting, including MCMC diagnostics [?]. MCMC was inefficient for the pre-data-post-model distributions due to the flat nature of the posterior surface, and so the sampling/importance resampling algorithm (SIR) [?] was used instead. I tuned the SIR algorithm so that all draws were unique and sampled 20,000 draws from the base case scenario.

#### 2.2.4 Estimating Current and Future Status

The current status of the population is determined by the abundance in the current year relative to the carrying capacity ( $N_y/K$ ). The population abundance is unlikely to continue increasing if the relative abundance is high, while the population may be at risk for further anthropogenic mortalities if relative abundance is low. Posterior distribution can be determined for any derived quantity, including abundance and relative abundance in any year. The posterior distribution of relative abundance depends on the assumptions made

in the model, particularly the choice of prior for  $r$  and the number of ship strikes in 2013. I therefore compared the posteriors across these scenarios to assess the impact each has on the results of this analysis.

I approached estimation of the future status of this population in two ways. First, I projected the model into the future and compared estimates of relative abundance under different mitigation scenarios. For example, if vessel speeds were reduced, the whales could potentially avoid collisions with vessels more effectively [?] and the ship strikes would drop, despite the same number of whales and vessels. The following mitigation scenarios covered a range of plausible future cases and allowed for testing of the impacts on the population recovery.

1. **Status quo.** In this scenario no measures to mitigate mortalities were implemented.
2. **Mitigation.** In this case I assumed some mitigation action is taken and that the rate of strikes is halved after 2013 relative to the status quo case, so that  $S_y^{\text{mitigation}} = S_y/2$  for  $y > 2013$ .
3. **Elimination.** This extreme case tested what would happen if ship strikes were eliminated completely after 2013. While unlikely this mitigation case provided a null case against which to judge the relative impact of the other two cases.

Mitigation only occurred after 2013 and thus had no impact on the fits of the model or estimates of current status. If no differences in abundance trends existed between these three scenarios then mitigation efforts would likely have little impact on the status of the ENP blue whale population. In contrast, if substantial differences existed and ship strikes were threatening the recovery and health of the population it may be worth exploring new options to mitigate this risk. This first approach provided a sense of the near-term impacts of ship strikes given an example of projected vessel traffic

My second approach addressed the long-term situation without making any assumptions about the future of vessel traffic. I assumed a constant level of vessel traffic relative to  $V_{2013}$  ('multiplier') and estimated the long-term relative depletion by projecting the model forward

100 years so that the population size equilibrated. This process generated a posterior distribution for the equilibrium population size for that level of multiplier, and was repeated over a range (0 to 20). If values of the multiplier close to 1 (i.e. close to vessels in 2013) led to a low equilibrium relative abundance then the current level of vessel traffic was predicted to have a large impact on this population. In contrast, if the relative abundance remained high even for 10 times the current level of vessel traffic then the population demonstrates resilience to ship strikes.

### **2.3 Results**

The probability of positive growth was estimated as the proportion of draws with  $m > 0$ , in this case 90.0%. Thus there is weak evidence that the population increased in size from 1993 to 2008.

The logistic model fits for the base case estimated that the pre-whaling abundance,  $K$ , was 2,199 (95% probability interval 1,821–3,713) (Table 2.3), and that these estimates changed little with the informative  $r$  prior and different levels of  $S_{2013}^{\text{obs}}$  (Table 2.4). The choice of prior on  $r$  influenced the posterior distribution for  $r$  substantially, indicating there was little information about this parameter in the abundance estimates, but had a much smaller effect of increasing the right tail of the posterior for  $K$  (Figure 2.1). This suggests that the results for estimates for the pre-whaling abundance are robust across the assumptions explored in this analysis related to the prior for  $r$ . Pre-data-post-model distributions for the case of a uniform prior for  $r$  show that the assumed model structure and catch history provides minimal information about the parameters (i.e. the priors and pre-data-post-model distributions are very similar). The carrying capacity parameter  $K$  is most influenced by the abundance estimates and  $r$  by the choice of prior for  $r$ .

The vessel model for the base case predicted that large vessel traffic grows at an intrinsic rate of  $r_v = 0.022$  (95% interval 0.017 to 0.148). The asymptote was estimated to be  $K_v = 285$  thousands of vessels over 1000 metric tons (95% interval 112 to 488), with the broad interval resulting because the data contained little information about this parameter (i.e. there is little evidence for the rate of increase in vessel traffic dropping). The future projections of the model are uncertain, ranging from leveling out within the next 10 years

Table 2.3: Estimated intrinsic rate of increase ( $r$ ) and carrying capacity ( $K$ ) for eastern North Pacific blue whales. Results are shown for a uniform prior for  $r$ : Uniform(0, 0.114) and an informative numerical prior from a meta-analysis, as well as assumed ship strikes in 2013,  $S_{2013}^{\text{obs}}$ , of either 10 or 20.

r Prior	$S_{2013}^{\text{obs}}$	r			K		
		2.5%	Median	97.5%	2.5%	Median	97.5%
Uniform	10	0.013	0.081	0.112	1,821	2,199	3,713
Uniform	20	0.018	0.084	0.113	1,852	2,247	3,507
Informative	10	0.055	0.098	0.112	1,791	2,150	2,663
Informative	20	0.063	0.099	0.112	1,825	2,169	2,597

to continuing to grow at an almost exponential rate (Figure 2.2). Assumptions regarding priors for  $r$  and the value of  $S_{2013}^{\text{obs}}$  had virtually no effect on the vessel model fits because the vessel data were independent of the assessment model parameters.

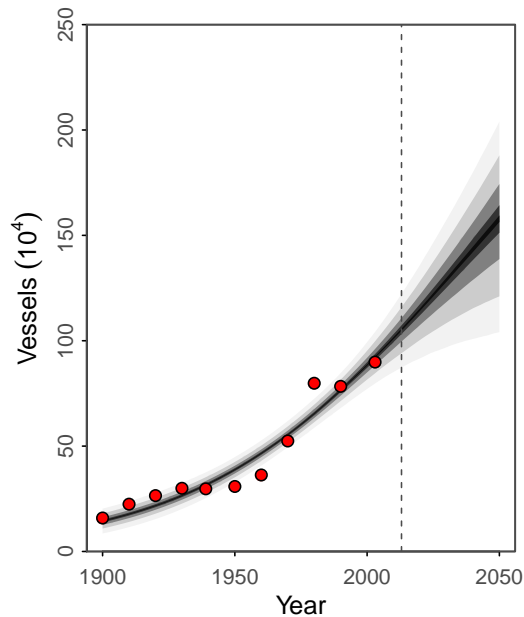


Figure 2.2: Results of the vessel model. The  $y$ -axis scale is arbitrary because it is scaled by the parameter  $c$  and is therefore not shown. The points are worldwide vessel statistics from Lloyd's of London.

The estimated historical ship strikes were a product of the estimated abundance and vessels in a year. As expected, the assumption about  $S_{2013}^{obs}$  directly influences the estimated trajectory of historical strikes (Figure 2.3). I found that for the uninformative prior there were a total of 555 strikes (95% interval 281 to 983) from 1900–2013 if  $S_{2013}^{obs} = 10$  and 1,088 (95% interval 665 to 1,684) if  $S_{2013}^{obs} = 20$ . The prior on  $r$  had little impact on the strikes (results not shown). Over the same period I estimated 2,963 catches (95% interval 2,460 to 3,3289) from whaling, so ship strikes were less than catches but relatively high.

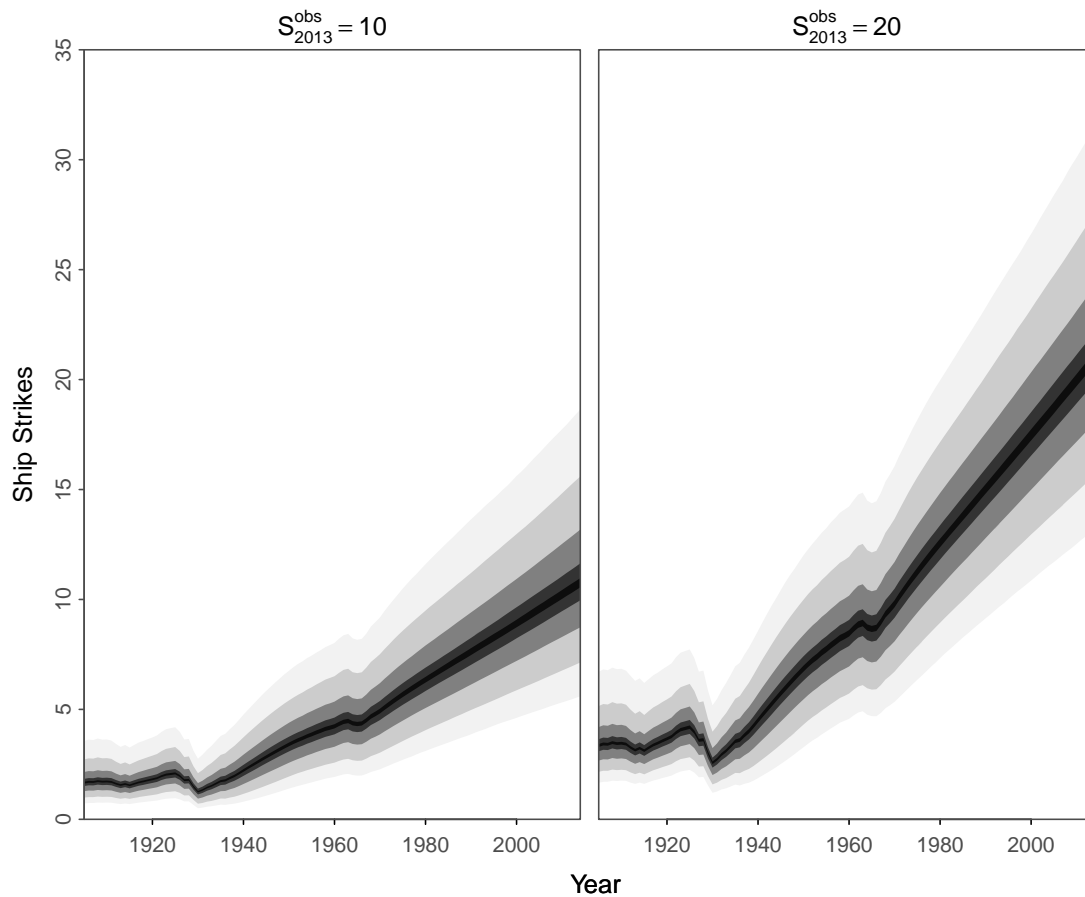


Figure 2.3: Predicted past ship strikes by year for the uninformative prior for  $r$ . The dip in the late 1960s is due to Soviet whaling.

I estimated the current status of the population finding that the population is close

to carrying capacity and thus has nearly recovered from whaling (Figure 2.4, Table 2.4) regardless of the prior used for  $r$  or assumptions about  $S_{2013}^{\text{obs}}$ . Further, the trajectory of relative abundance over time suggests that the population was most depleted in 1930 and has been recovering ever since, except for a short period of depletion in the late 1960s from Soviet catches (Figure 2.5). The same trends for absolute abundance show the population likely never dropped below 500 whales, and that median catches were at their highest (340 in 1928 and 275 in 1929) in the years immediately before this low point (Figure 2.6).

Table 2.4: Estimates of the abundance and abundance relative to pre-whaling levels ( $N_{2013}/K$ ) for year 2013. Different combinations of the prior for  $r$  and value of observed ship strikes in 2013,  $S_{2013}^{\text{obs}}$ , represent different assumptions and are a sensitivity analysis. The base case is indicated in bold.

<b>r Prior</b>	$S_{2013}^{\text{obs}}$	<b>Absolute Abundance</b>			<b>Relative Abundance</b>		
		<b>2.5%</b>	<b>Median</b>	<b>97.5%</b>	<b>2.5%</b>	<b>Median</b>	<b>97.5%</b>
Uniform	10	1,769	2,125	2,613	0.62	0.97	0.99
Uniform	20	1,760	2,120	2,542	0.64	0.95	0.97
<b>Informative</b>	10	1,688	2,098	2,521	0.95	0.98	0.99
<b>Informative</b>	20	1,739	2,071	2,498	0.92	0.96	0.98

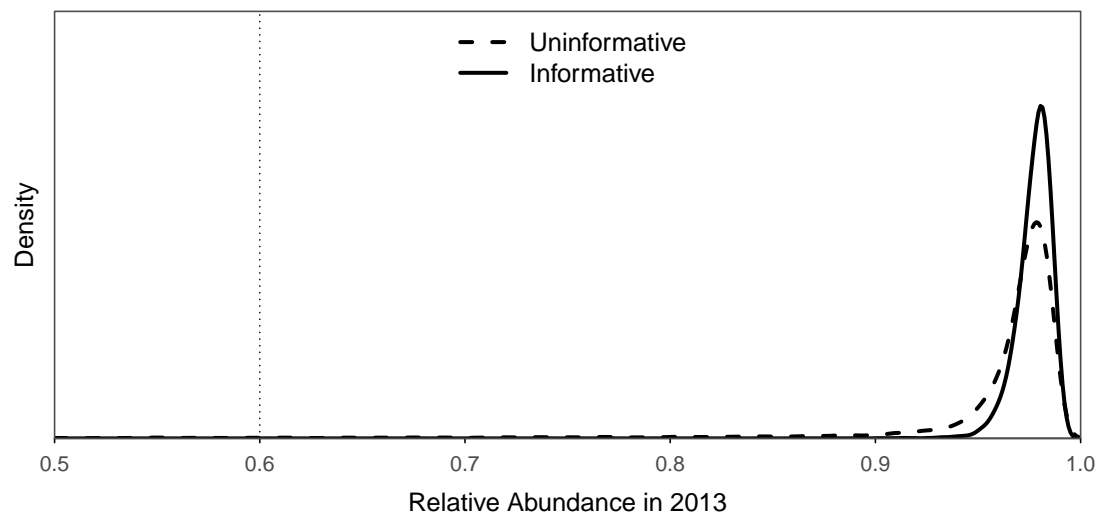


Figure 2.4: Posterior densities for the relative abundance in 2013 for the two priors on  $r$ . The vertical line at 60% denotes the level of maximum productivity.

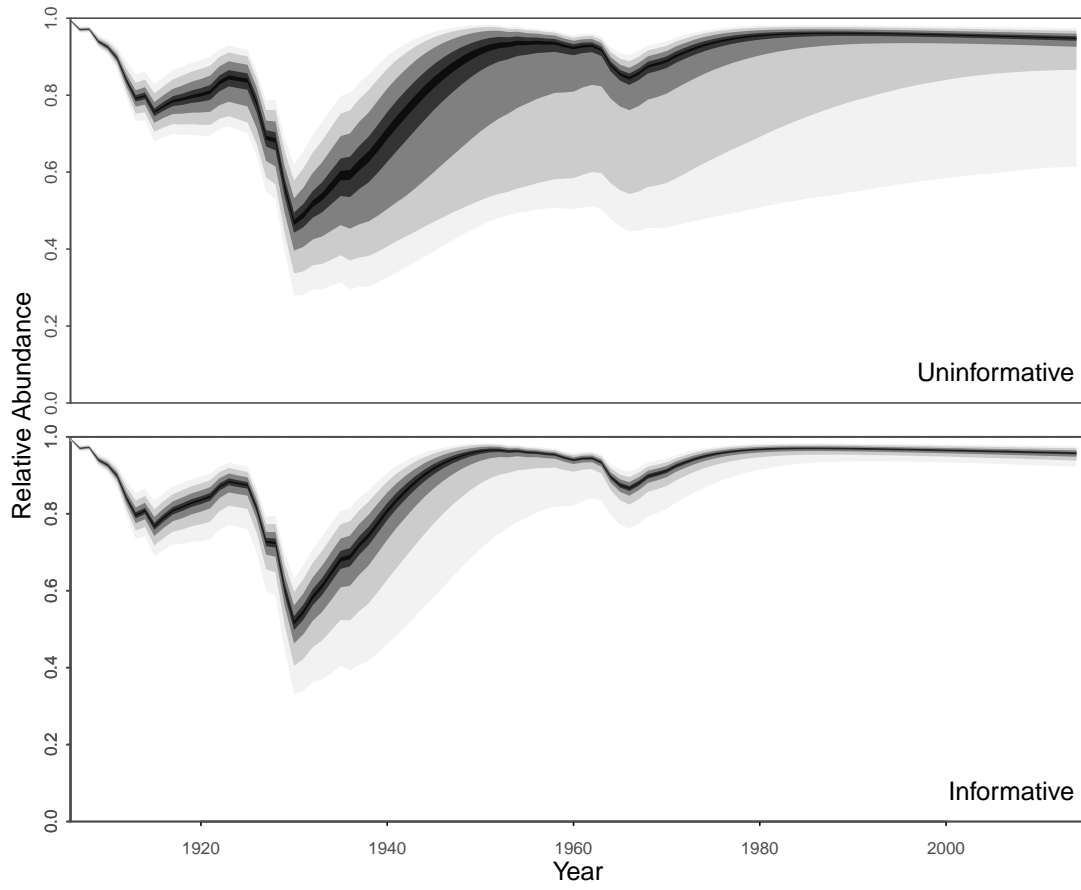


Figure 2.5: Relative abundance trajectories for both priors on  $r$ .

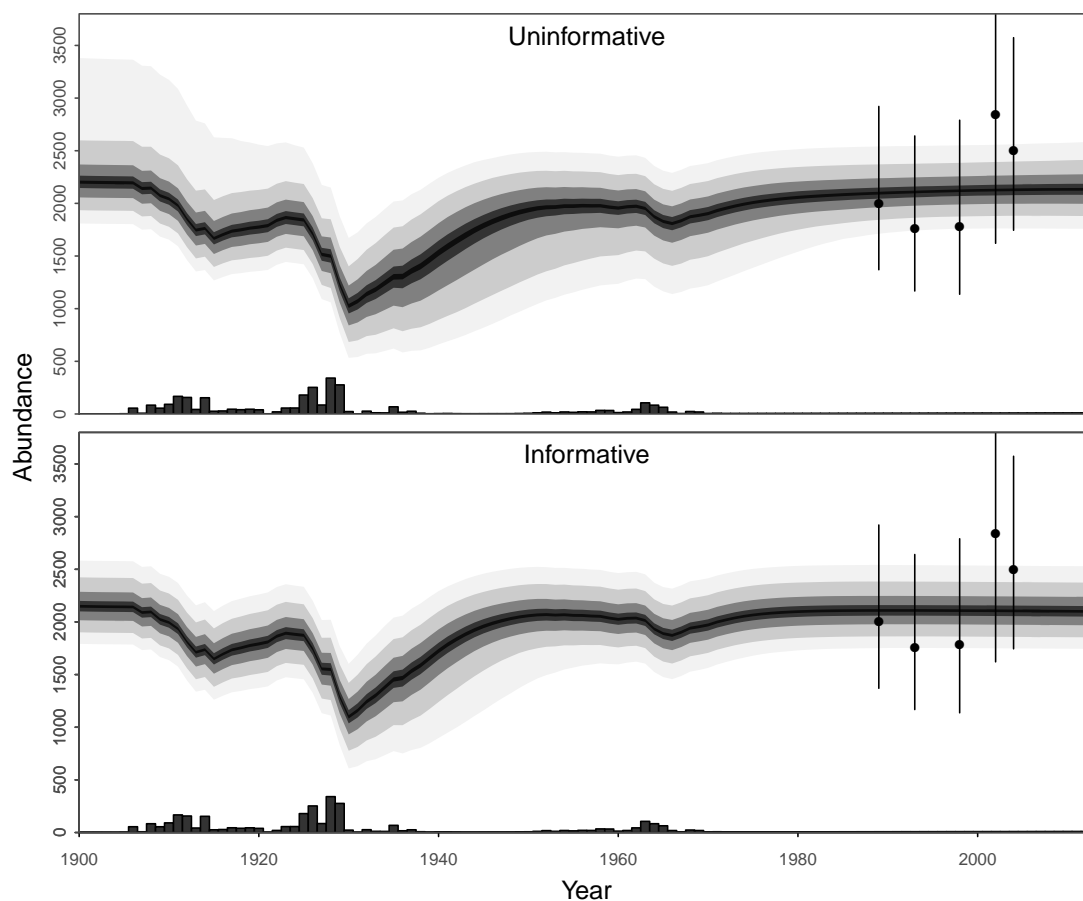


Figure 2.6: Absolute abundance, the abundance estimates and anthropogenic mortalities for the two priors for  $r$ . The rectangles at the bottom denote total mortalities (catches + median strikes) for each year. Ship strikes are so much lower than the whaling catches they are barely visible even in the later years. The five abundance estimates (points) are shown with their 95% confidence intervals (bars).

My short-term approach for estimating future status was to project the population dynamics model forward under different mitigation cases. Under the status quo case (i.e. no mitigation) the strikes continue to increase and the population begins a slight decline. With the mitigation case where the strikes after 2013 were cut in half (to simulate for example reductions in vessel speed) the decline in abundance is noticeably slower. However, compared to the case with no future strikes there is little qualitative difference among the three cases (Figure 2.7). These patterns hold across all scenarios explored (results not

shown). The second approach found that about 7 times the number of vessels as in 2013 would lead to approximately 50% chance that the population would be below its maximum level of productivity (assumed here to be 60%).

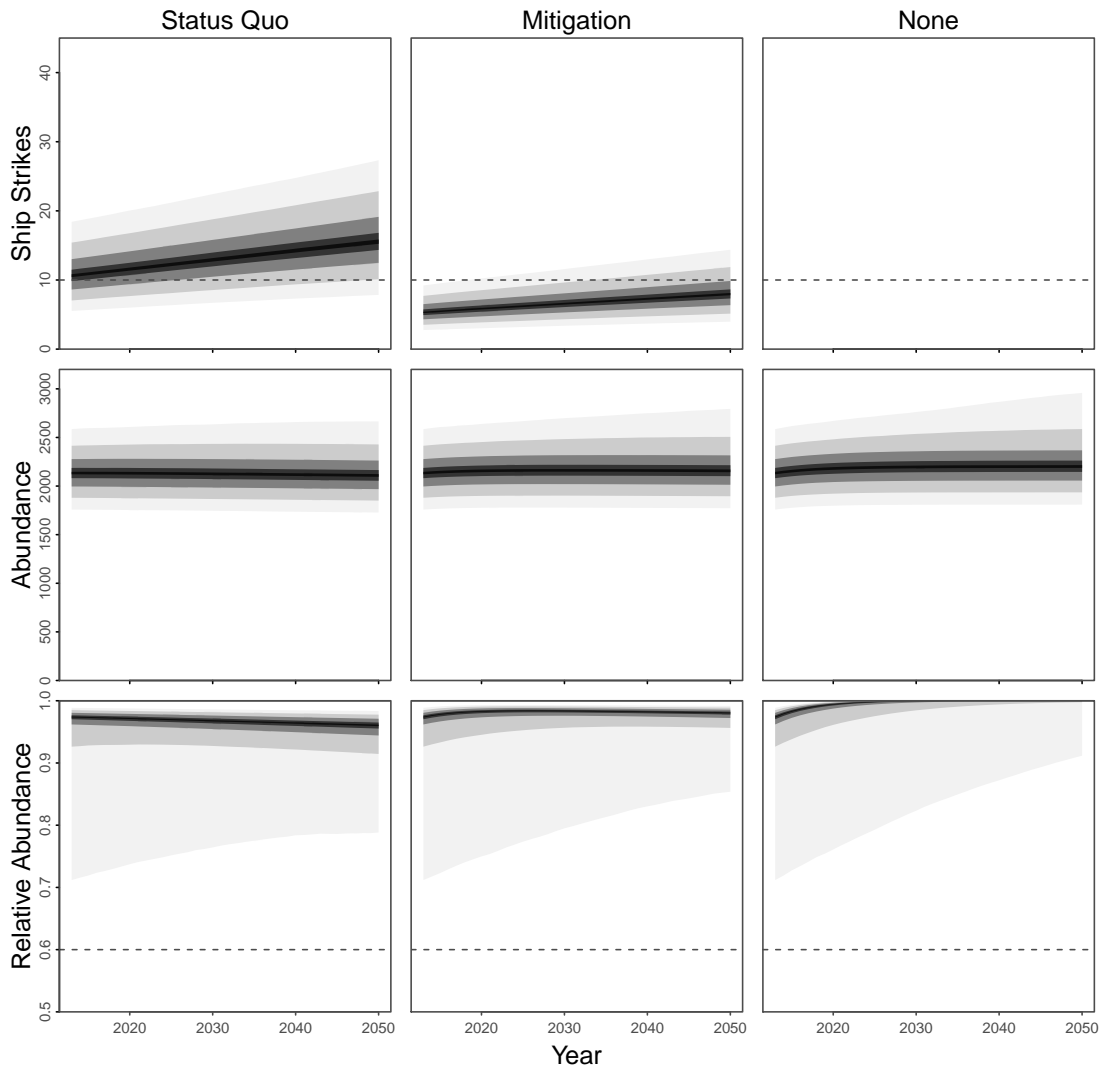


Figure 2.7: Future strikes, abundances, and relative abundances under three mitigation cases. 'Status quo' means no mitigation, 'mitigation' refers to halving the strikes after 2013, and 'none' is a complete elimination of future strikes.

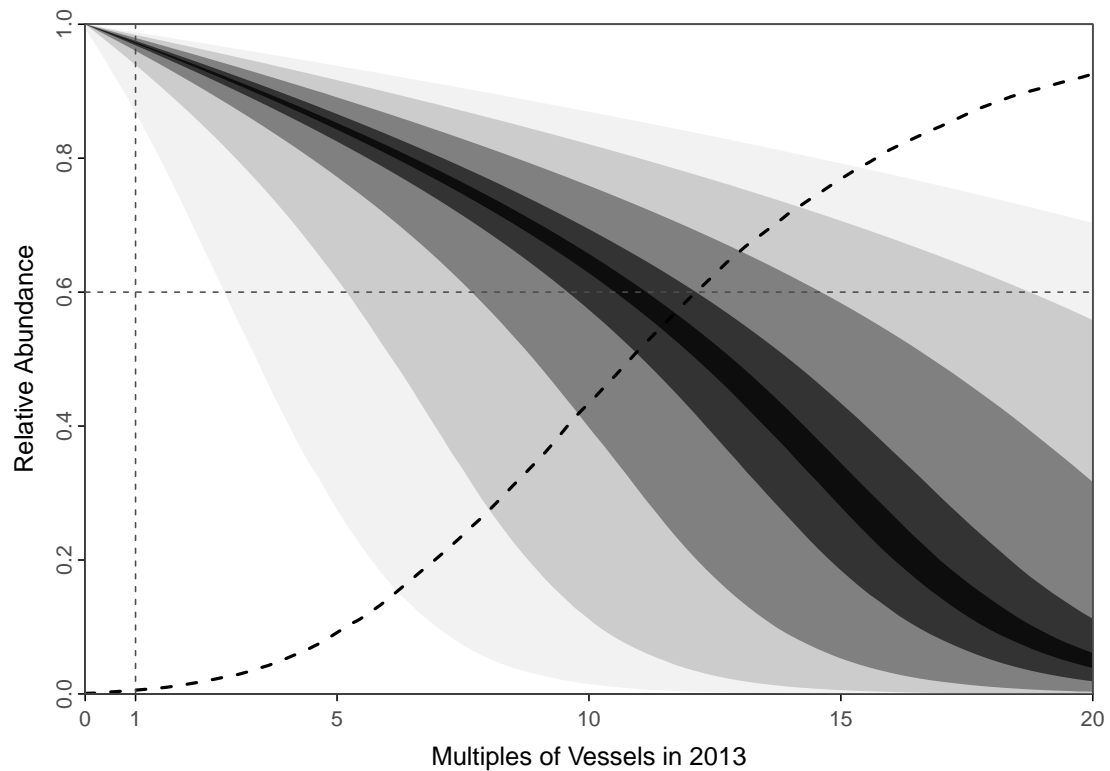


Figure 2.8: Equilibrium relative abundance for a range of multiples of current levels of vessels. The model was projected forward 100 years for each posterior sample under a constant multiplier. As the multiplier increases there are more vessels and thus strikes, depleting the population. The probability that the population is below 60% of  $K$  is shown as a dashed line.

## 2.4 Discussion

The log-linear model suggested the population likely grew slowly from 1993–2008. The theta-logistic model further demonstrated that the recovery had slowed due to density dependence effects (i.e. approaching carrying capacity). Thus my analysis suggests the ENP blue whales have recovered to pre-whaling levels and are productive and resilient to current levels of anthropogenic mortalities. This population’s status conflicts with the “endangered” designation of the Endangered Species Act, although the structure and status of blue whales populations in other regions is unclear. The Antarctic population is the only other assessed population and had relative abundance in 1996 of 0.7% (interval 0.3%–1.3%) [?]. Thus the ENP population is the only population of blue whales shown to have recovered from

whaling.

My analysis suggests that current levels of ship strikes are not threatening the status of the ENP blue whale, and thus mitigation strategies would have limited impact on their status. However, I have demonstrated that as vessel traffic increases and ship strikes become more common the impact to the population will increase. For instance, at five times the 2013 level of vessel traffic there is a 9% probability that the long-term relative abundance is below 60%. If vessel traffic increases and ship strikes become a bigger problem, mitigation approaches will be more effective.

Undoubtedly ship strikes are currently a major concern for other endangered cetacean populations (e.g. [?]). Incorporating projected ship strikes into assessment models is an important step for understanding the risks they pose, and quantifying and exploring the effects of potential mitigation strategies. When a mitigation scenario increases strikes for one population and decreases them for another, as is in the case in the Santa Barbara Channel [?], it is important to consider the impacts on the status of the populations, not just the absolute number of ship strikes.

My analysis has limitations and caveats that may influence the results and conclusions. I used the mark-recapture abundance estimates exclusively and implicitly assumed they were reliable, although the line-transect estimates show a similar absolute abundance but different trend [?]. The catch data are known to be unreliable [?], but this is taken into account here and it is unlikely that the enough catches would still be missing to influence the results given the relative insensitivity to this factor. Similarly, the ship strike sensitivity analysis suggests the impact of errors in the model assumptions or data made about the vessel and ship strike models are unlikely to impact the conclusions. The logistic population dynamics model (equation (2.2)) was used due to a paucity of data and the availability of informative biological priors, but ignores many biological processes driving the dynamics. However, given the scale of the catches and the abundance estimates it is unlikely that a more complicated model would lead to different conclusions. I also assumed a single eastern North Pacific blue whale population, but if there is a finer population structure the analyses presented here would be invalid. As of now, most evidence suggests a single population measured by the mark-recapture estimates and from which historical catches

come. However, future studies on the population structure of blue whales in the North Pacific will shed light on the validity of the assumptions and approach used in this analysis.

My analysis suggests that despite seven decades of commercial whaling ENP blue whales have nearly recovered to their pre-exploitation levels and are in no immediate threat from ship strikes. The recovery of other blue whale populations, as well as populations of other species, is unclear. However, the recovery of ENP blue whales demonstrates that an exploited population of blue whales can recover given appropriate conservation measures.

## CONCLUSION

This thesis provides the first analysis of the recovery of ENP blue whales and the risk that ship strikes pose. As with any analysis the results and conclusions presented here rely on the validity of the assumptions made, as well as the models and data used. I ran sensitivity tests where possible to try to quantify the potential impact of deviations from assumptions. Given the caveats of my analysis, I found that density dependence is a more likely hypothesis than ship strikes for a lack of continued recovery and that current level of ship strikes poses very little threat to the status of this population. There is uncertainty regarding the current level of ship strikes,  $S_{2013}^{\text{obs}}$ , but my results were insensitive to the value assumed, suggesting the results would hold even if this number were too low. With a relatively large population size and high intrinsic rate of growth, this population is robust to anthropogenic mortalities. I found that over a 5-year period (1925–1929) median catches were 1,126 and the population still recovered; so it is unsurprising that ship strikes an order of magnitude less would not significantly impact their recovery.

However, the impact increases as vessel traffic does, so future ship strikes could pose a serious threat to the health of this population. It is very difficult to accurately predict the future rate of traffic for vessels that pose a threat to blue whales. For now it appears that vessel traffic continues to increase, but it is unclear how vessel traffic will change in terms of the size and speed of ships; the growth of vessels may decrease or increase in the future depending on a variety of factors. It is beyond the scope of this thesis to try and predict future shipping patterns. Instead I have provided a framework under which to evaluate different vessel scenarios. My analysis shows that ship strikes are unlikely to be a major concern in the near future and presumably vessel traffic will change slowly and there will be ample opportunity to detect any serious declines to this population and evaluate the efficacy of any proposed mitigation techniques.

Blue whales are classified as endangered because of decades of commercial whaling on

a global scale. The Antarctic population is at dangerously low levels [?] and no other populations have been completely assessed. This thesis demonstrates that an exploited blue whale population can recover given the right conditions, but much work remains to clarify the structure and status of blue whale populations around the world.

## Appendix A

### INFORMATIVE PRIOR FOR $r$

The informative prior used for  $r$  was derived from numerical samples and used in the population dynamics model 2.2 based on [?]. Included here is R code that can be used to replicate this prior.

```
r.prior.den <- function(r0){
  ## A function that returns the numerical function derived from
  ## Andre's CDF from his meta-analysis. Pass it r0 which is from 0 to rmax
  ## and it returns the density for each point in the vector.
  mypdf <-
    1550.75566650868*1*r0^0+-87836.61893953*2*r0^1+
    2961361.8059117*3*r0^2+-60483210.540171*4*r0^3+
    763188172.260658*5*r0^4+-5819817610.60468*6*r0^5+
    24594982913.3944*7*r0^6+-44262720834.0475*8*r0^7
  mycdf <-
    -19.1734207202477*r0^0+1550.75566650868*r0^1+xf
    -87836.61893953*r0^2+2961361.8059117*r0^3+
    -60483210.540171*r0^4+763188172.260658*r0^5+
    -5819817610.60468*r0^6+24594982913.3944*r0^7+
    -44262720834.0475*r0^8
  ## Need to differentiate it due to transform used in the fit
  result <- mypdf*exp(mycdf)
  result[ (r0>=.114 | r0 <= 0)] <- 0
  return(result)
}
```