USING VOLUNTEERED GEOGRAPHIC INFORMATION TO MODEL BLUE WHALE FORAGING HABITAT, SOUTHERN CALIFORNIA BIGHT

Matthew W. Bissell

A Thesis Presented to the
FACULTY OF THE USC GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF SCIENCE
(GEOGRAPHIC INFORMATION SCIENCE AND TECHNOLOGY)

December 2013

Acknowledgements

Several accolades are in order to acknowledge the people who made this thesis possible. Thank you to Travis R. Longcore, my thesis chair, who was always available and flowing with knowledge. And to the committee: Yao-Yi Chiang, Meredith Franklin, and Flora Paganelli. Thank you all.

Thank you to the captains and crews aboard the five whale-watching vessels who took the time to record whale observations during their many voyages, and for allowing me access to their datasets: Mike Bursk, Peter Heistand, Larry Hartmann, Kera Mathes, Carl Mayhugh, Cici Sayer, and Melissa Galieti. And thank you Greg Campbell, for allowing access to the CalCOFI cetacean dataset.

And to Mary Elizabeth Portwood: the best friend, research assistant, secretary, cook, companion, and partner in crime one could wish for. Thank you and I love you.

Thank you Mom, thank you Dad, for everything.

Table of Contents

List of Figures and Tables

Figures

Tables

Abstract

Using Volunteered Geographic Information (VGI) to model blue whale (*Balaenoptera musculus*) foraging habitat, this thesis assesses the utility of citizen science in cetacean research and marine spatial management. A unique and new data source on whale locations, observation data collected voluntarily by whale-watching vessels, was procured, compiled, and digitized. The utility of this newfound dataset was investigated through its use in probabilistic habitat suitability analyses and description of species phenology. A statistical analysis of whale observations was used to quantify seasonal variability of three common baleen whale species within the study area. Among these, blue whales exhibit the highest degree of seasonal variability with a mean seasonal abundance occurring in late July. Maximum entropy modeling was used to illustrate potential blue whale foraging areas based on three environmental variables: bathymetry, sea surface temperature, and chlorophyll-a concentrations. Spatial patterns of whale observations recorded by whale watchers and scientists indicate a strong habitat preference of steep bathymetric features in and around the 300-m isobath. Models using whale-presence data collected by whale-watchers were compared to similar models using science-quality whale observation data. Differences between these models are minimal and the results of the comparison support the usefulness of citizen science in cetacean research.

Chapter 1: Introduction

Volunteered Geographic Information (VGI) involves human volunteers, as sensors of the environment, contributing in whole or in part to the creation, collection, and/or dissemination of geographic information (Goodchild, 2008). This thesis composes and analyzes the utility of a volunteered geographic dataset describing whale presence locations in the coastal waters of Southern California. This new dataset was created by combining the geographic information stored in the logbooks of five whale-watching vessels operating daily and simultaneously from four ports within the Southern California Bight. The high temporal (daily) resolution and vast geographic coverage of this dataset is characteristic of VGI. These highly opportunistic observations are used here to create probabilistic habitat suitability models for blue whales within the Southern California Continental Borderland. Habitat modeling is an essential component of conservation biology and key to understanding how we can better share the marine environment with vulnerable species. The potential for future use of this new whale observation dataset in scientific research and marine management is the focus of this thesis.

Hunted to near extinction during the nineteenth and twentieth centuries, many baleen whale species exist at a fraction of their pre-whaling population numbers (Mate and Calambokidis, 1999). Since the international ban on commercial whaling in 1965, many of these whale species have been slowly recovering from their once intensive exploitation (Calambokidis and Barlow, 2004). Nevertheless, baleen whales are still a vulnerable group and are faced with continuing threats from human activity. These include water contamination, high levels of anthropogenic sound, entanglement with commercial fishing gear, and collisions with ships (Laist et al., 2001). These physical threats have physical locations, and if we intend to protect these threatened species, it is necessary to identify the geographic areas where whales might be affected by such activity. The understanding of when and where a population is most likely to occur is a primary goal in conservation biology (Phillips et al., 2006). In addition, knowing how a species interacts with its environment can help to forecast future habitat selection by that species as the environment changes.

Habitat modeling is a useful tool that can provide important information pertaining to a species' potential range and distribution. Taking what is known about an organism's presence and habitat requirements, statistical models can predict the probability that other areas will also offer suitable habitat for that organism (Phillips et al., 2006). Much work has been done to study whale habitat and many variations of whale habitat models have been developed to give us a better understanding of how these animals interact with their environment (Munger et al., 2009; Burtenshaw et al., 2004; Moore et al., 2002; Fiedler et al., 1998; Croll et al., 1998). Recently, several studies have combined this information with data on shipping routes and other human-induced hazards, identifying regions with high risk of human-induced whale mortality (Redfern et al., 2013; Pittman and Costa, 2010). This knowledge can be used to inform marine managers when designing and implementing maritime regulations or to recommend best practices to boaters operating their vessels within known whale habitat. The predictive capabilities of these habitat models are commensurate with the quality and resolution of the data put into the model. Currently, most whale habitat models utilize whale presence data obtained from scientific surveys with regularized spatial sampling techniques performed by trained biologists. These surveys are often laborious, time consuming, and expensive, and as a result are conducted infrequently. For example, a commonly cited dataset for whale observations in California is from the California Cooperative Oceanic Fisheries Investigation (CalCOFI), with surveys being conducted quarterly (i.e., four times per year). While this data is of high quality on many levels, it is of low temporal resolution. Models using this data can still be effective, but they are often limited by small sample sizes resulting from scarcity of surveys and lack of data during winter months when weather conditions deteriorate (Munger et al., 2009). One solution to this limitation is the use of citizen science in the data collection process. A sub-category of VGI, citizen science is capable of amassing large amounts of data covering vast geographic areas over short periods of time (Conrad and Hilchey, 2011).

This thesis utilizes a newfound volunteered geographic dataset provided by the eco-tourism industry of Southern California. Assembled from five whale-watching vessels operating daily and simultaneously from four ports within the Southern California Bight, the data is of remarkably high temporal resolution and was collected at no cost. While subject to several biases, which will be discussed in subsequent chapters, this new dataset demonstrates the potential for increased citizen science in marine mammal research.

Objectives

The objectives of this thesis are: (1) to demonstrate the spatial patterns of opportunistic observations of baleen whales within the Southern California Bight; (2) to provide descriptive statistics of the seasonal variability of whale sightings within the Southern California Bight; (3) to use these opportunistic observations to produce probabilistic habitat suitability analyses describing the potential spatial and temporal extent of whale presence within the larger Southern California Continental Borderland; (4) to compare these results with models using science-quality data from expeditions with more regular spatial sampling schemes but much lower temporal resolution; and (5) to thereby assess the utility of opportunistic observations in scientific research.

Chapter 2: Literature Review

Whales have a long documented history in the Southern California Bight. Much of the early documentation of whales in this area came from industrial whaling outposts operating along the coast during the nineteenth and twentieth centuries; six such outposts were known to exist between Point Conception and San Diego (Starks, 1922). These operations often maintained records of their takes of whales and sales of their products, which contributed to early estimates of whale populations in the area. After nearly 100 years of intense exploitation, the hunting of whales off the California coast was brought to an end in 1965 (Calambokidis et al., 2009). Following a half-century of slow repopulation, these whales are again a valuable resource for the local economies. This time, however, it is not for their oil and meat. Whale-watching boats take droves of passengers into the waters of Southern California to view these animals nearly every day of the year. And much like the whaling industry before them, these eco-touring vessels hold the potential to be very rich sources of whale presence data.

An estimated 30 species of cetaceans reside in the eastern Pacific Ocean (Balance et al., 2006). A number of these whale species can be observed in the waters off Southern California; CalCOFI biologists identified 15 different whale species in this area during their quarterly survey cruises in 2009–2010. This number was also observed during the 2010–2011 cruises, and 14 different species were recorded in the 2011–2012 surveys (Campbell et al., 2010, 2011, 2012). Of these observed species, fin (*Balaenoptera physalus*, blue (*Balaenoptera musculus*), gray (*Eschrichtius robustus*), and humpback (*Megaptera novaeangliae*) whales were the most common of the baleen whales.

This thesis focuses on the blue whale for four reasons. (1) Blue whale occurrences in the VGI dataset exhibit the most consistent seasonal variability with over 90 percent of their occurrences taking place in July, August, and September. This phenomenon makes for convenient monthly comparisons between years. (2) Blue whales are actively feeding in the study area during these months and their presence or absence is largely affected by the availability of their food source (Croll et al., 1998; Fiedler et al., 1998). This creates a starting point for model development; e.g., which environmental variables have an effect on prey production? (3) Blue whales feed almost exclusively on krill, a planktonic crustacean whose presence and abundance is closely linked to the local environmental conditions (Croll et al., 1998; Fiedler et al., 1998). Other species of baleen whales known to forage in the study area have a much more varied diet consisting of zooplankton and several species of small fishes, making their food source more difficult to model. (4) Records of blue whales in the ships' logbooks are often more complete than records of other species. This is likely due to the intrinsic value of the blue whale to the whale-watching industry. A captain can more easily please a group of passengers by presenting them with a whale of many superlatives than with smaller, less impressive animals. As a result, a blue whale observation will almost always be recorded with precise coordinates of latitude and longitude. The locations of other whale species (fin, gray, minke, etc.) recorded in the logbooks were sometimes noted as "out front," "near red buoy," or "off blue house." These types of vague descriptors were seldom used for describing a blue whale's location.

The blue whale is said to be the largest animal known to have lived on this planet; they are capable of reaching lengths of nearly 100 feet, and weighing up to 120 tons (Hass, 2011). Harvested to near extinction in the nineteenth and twentieth centuries, the global population is estimated between ten and fifteen thousand whales: merely 8 percent of their pre-commercial whaling numbers (Mate and Calambokidis, 1999). The National Oceanic and Atmospheric Administration (NOAA) lists the blue whale as a depleted species (population below optimum sustainable levels) under the Marine Mammal Protection Act of 1972, and in danger of extinction under the Endangered Species Act of 1973. Blue whales are also listed as endangered species by the International Union for Conservation of Nature (IUCN) and the Convention on International Trade in Endangered Species (CITES). The remaining blue whales are found throughout the world's oceans and reside in distinct populations that seldom mix (Burtenshaw et al., 2004). As members of a suborder of cetaceans called Mysticeti, their mouths are equipped with baleen plates designed for capturing small planktonic prey. They feed almost exclusively on krill, a small crustacean found in all of the world's oceans (Croll et al., 1998; Fiedler et al., 1998). Very mobile animals, they partake in extensive annual migrations from summer foraging grounds to areas of breeding and calving during the winter (Pittman and Costa, 2010). While we still do not fully understand the complete annual migrations of these animals (i.e., where they go for breeding and calving), they do show a high level of fidelity to their summer foraging grounds (Pittman and Costa, 2010; Mate and Calambokidis, 1999). The coastal waters off California serve as a foraging area for possibly the largest remnant population of blue whales in the world (Mate and Calambokidis, 1999). Recent estimates of this population suggest between 2,000 and

3,000 individuals (Calambokidis and Barlow, 2004). Current threats to this species include anthropogenic sound production, water contamination, entanglement with commercial fishing gear, collisions with ships, and illegal whaling (Laist et al., 2001).

Several methods have been used to model whale habitat. Most variations use a combination of environmental variables as indicators of suitable habitat or proxies for potential food availability. Bathymetry, sea surface temperature (SST), and chlorophyll-a are common physical variables used in these models (e.g., Redfern et al., 2010; Munger et al., 2009; Balance et al., 2006), as described in the following sections.

Bathymetry

Bathymetry is the measurement of the ocean's depth and describes the underwater features that make up the sea floor. Bathymetry has a profound impact on the abundance and diversity of organisms living in the water column above (Pittman and Costa, 2010). Pittman and Costa (2010) discuss the high predictive powers of bathymetry alone in modeling whale abundance and distribution, noting that edge habitats (e.g., continental slopes) are strongly linked to high concentrations of prey. Seafloor features have a significant influence on the vertical and horizontal movement of water and the resulting eddies can serve to collect and maintain large concentrations of krill (Croll et al., 1998). As a result, steep bathymetric features are necessary for blue whales to exploit their tiny prey (Fiedler et al., 1998). In the northwest Pacific, seamounts, slopes, and other prominent bathymetric features were identified as focal points for blue whales throughout the year (Moore et al., 2002). Pitman and Costa (2010) describe the 100-m isobath line as a “cetacean superhighway” for whales along the southern gulf of Maine and argue that bathymetry data should be a prime candidate when choosing environmental variables to model whale habitat. Burtenshaw and others (2004) identified bathymetry as an import variable that could have added predictive power to their model (they did not include bathymetry).

Sea Surface Temperature

Sea Surface Temperature (SST) is a measure of the temperature of the uppermost layer of the ocean to about one meter (Campbell and Wynne, 2011). It can be measured in situ via boats, buoys, and underwater autonomous gliders, or remotely via airborne and space-borne sensors. SST is a fundamental component of marine ecology; oceanic temperatures can define marine habitats and detect biological hotspots (Etnoyer et al., 2006). A study of blue whale distributions off Southern California found SST to be an important variable influencing the presence or absence of whales (Munger et al., 2009). The authors found blue whales to be associated with colder SST when compared to random locations in the study area. This was thought to be a result of oceanic processes that, in addition to bringing cold water to the surface, foster prey production, accumulation, and retention (Munger et al., 2009). Also in waters off Southern California, two reports published in 1998 identify low relative water temperatures as an indicator of potential blue whale habitat (Croll et al., 1998; Fiedler et al., 1998). In each of these studies the majority of blue whale observations were made in cold, well-mixed water that had been upwelled north of the sighting location and advected south via the California Current System. Similarly, a study conducted in the Northwest Pacific found blue whales to be associated with colder than area-average SST (Moore et al., 2002). These whale clusters were also near SST fronts with sharp gradients. A study in 2007 found blue whales to be more closely correlated with SST fronts than any other whale species in their study (Doniol-Valcroze et al., 2007). This relationship between blue whale habitat selection and SST is also documented in the Great Australian Bight where the species has been linked to SST of about 1 degree Celsius cooler than average SST in the study region (Gill et al., 2011).

Chlorophyll-a

Chlorophyll-a is the measure of primary productivity in the ocean’s upper layer. Plantlike plankton occupying the sunlit portion of the world’s ocean use chlorophyll-a and other pigments to perform photosynthesis. These colorful pigments can be detected via satellite remote sensing and the varying concentrations of chlorophyll-a on the ocean’s surface can be discerned. Chlorophyll-a has been shown to be an important environmental variable when modeling the habitat of blue whales. Because this pigment is an indicator of primary productivity (the food source of zooplankton) it can be used as a proxy for blue whale prey production. Several studies have associated blue whale presence with high levels of chlorophyll-a. In 2002, Moore and others noted that blue whales in the northwest Pacific were associated with high concentrations of chlorophyll-a in the spring; this strong association was not observed later in the foraging season. The authors hypothesized this was due to the voracious primary consumption of phytoplankton by the zooplankton, coupled with the reduced input of nutrients on the back end of the upwelling season (Moore et al., 2002). In a separate study of blue whales in the northeast Pacific, their abundance was also associated with high levels of chlorophyll-a (Burtenshaw et al., 2004). This study identifies a time lag between peak chlorophyll-a concentrations and whale presence on the order of several months (Burtenshaw et al., 2004).

Citizen science

Citizen science is the involvement of volunteers in some or all aspects of scientific research (Conrad and Hilchey, 2010). There are several benefits of citizen science: large data sets can be compiled very quickly and inexpensively (Trumbull et al., 2000) and processes can be observed over large geographic areas (Dickenson et al., 2010). The results produced by citizen science not only provide decision makers with vital information on important matters, but the entire process increases public awareness and public involvement in these same issues (Bonney et al., 2009; Goffredo et al., 2010).

Using nonprofessional scientists as volunteer sensors of the environment is nothing new. One enduring citizen science campaign, the Christmas Bird Count, began in 1900 as a means to discourage over-hunting and promote ecological awareness (Bianchi, 1999). The 27 volunteer birders that took part in the inaugural bird count has grown to involve over 50,000 citizens worldwide volunteering to collect bird observation data each year. Continuing today, it is producing an ever-expanding online data set providing critical information on distributions, ranges, and migration patterns of avian species (Audubon, 2013). An effort of this scale, both spatially and temporally, would be nearly impossible without the involvement of citizen scientists. This successful use of volunteered data in scientific research can also be seen in the realm of marine science and marine management. In an effort to survey the underwater ecosystems of coastal Italy, researchers utilized the effort and enthusiasm of 3,825 volunteer SCUBA divers. In just four years the project was able to amass nearly 19,000 biological surveys of Italian marine ecosystems. Collectively these divers contributed over 13,000 hours of underwater data collection at no cost to the research organization. A later assessment of this data concluded that the quality and accuracy of the volunteered data was equal to data collected by trained divers on precise science-based transects. A subsequent and independent study performed by Italy's Ministry of the Environment validated the ecological findings of the VGI dataset (Goffredo et al., 2010). Similarly, a study on the east coast of the United States involved nearly 1,000 citizen scientists to monitor 750 kilometers of coastline. Throughout this vast geographic area, volunteers surveyed the inter-tidal ecosystems in search of invasive species of crabs. These volunteers were found to have a high level of accuracy when identifying species and the volunteered data detected a range expansion of one species of invasive crab (Delaney et al., 2008).

The development and evolution of several technologies have increased the ability of citizens to participate in science and, more specifically, geographic research (Goodchild, 2007). The World Wide Web has essentially connected the world, allowing large amounts of data to be easily shared, compiled, and analyzed (Goodchild, 2007). Global Positioning Systems (GPS) allow for locations of objects to be easily and accurately measured and recorded by non-trained individuals; this exercise is an essential component of geography. Digital cameras (many enabled with GPS technology) allow the average person far more access than ever before to photography, further enabling the citizen scientist (Goodchild, 2007). Mobile web devices and the ever-expanding capabilities of smart phones also increase the ability of the general public to participate in scientific research (Haklay, 2013).

With more citizens now capable of collecting accurate geographic data, the quality and size of VGI and citizen science-based datasets have increased. And because humans are a rather ubiquitous species, these datasets are often of very high spatial and temporal resolution. Temporal resolution refers to the frequency of data collection with respect to time: Data collected daily gives the dataset a higher degree of temporal resolution than a dataset whose measurements are made weekly, monthly, annually, etc. Temporal resolution is an important factor in habitat modeling, and fine-scale (daily) temporal resolution may be crucial for detecting temporal trends (Kearney et al., 2011).

Chapter 3: Study Area and Data

Study Area

The study area is described in two parts: the geographic area from which the whale observation data was collected, and the geographic extent of the extrapolated probabilistic habitat suitability models. The whale observation data was collected from within the Southern California Bight. This area is defined as the coastal waters (from the shoreline to the continental shelf) from Point Conception in the north to the U.S.-Mexican border in the south. It is defined by a wider than average continental shelf, with complex bathymetry consisting of many basins and ridges. Oceanic circulation within the Southern California Bight is also unique; the northbound Davidson Current brings warm water up along the coast, while the California Current carries cold nutrient-rich water southward and further offshore. These opposing oceanic currents create a biological transition zone that supports nearly 500 species of fish and more than 5,000 species of invertebrates (Southern California Coastal Water Research Project, 2013).

The probabilistic habitat suitability models were applied to the Southern California Bight and extrapolated south to Vizcaino Bay, Mexico, about halfway down the Baja California peninsula. This larger area known as the Southern California Continental Borderland is a natural extension of the Southern California Bight. The continued biologic, oceanographic, and bathymetric complexity of this area creates a geographic unit ideal for studying the behavioral ecology of marine mammals (Henderson, 2010). The width of the continental shelf narrows quickly just north and south of this region.

Data

Whale presence data can be acquired via three general methods: visual observation, acoustic detection, and radio telemetry (tagging). Most scientific research on live whales uses at least one of these techniques. Each of these methods has its benefits as well as its shortfalls. This thesis focuses on the use of visual observations in whale research. Specifically, a comparison between whale observation data collected by scientists on scientific surveys and whale observation data collected voluntarily on commercial whale-watching tours will be made.

Science-Quality Whale Presence Data

A commonly cited data source for blue whale presence (visual observations) in the Southern California Bight is the CalCOFI dataset. CalCOFI is a partnership formed in 1949 between the Department of Fish and Wildlife, NOAA Fisheries Service, and Scripps Institute of Oceanography. The organization conducts quarterly cruises along a series of transect lines extending perpendicular from the central and Southern California coastlines. Originally commissioned to study the collapse of the sardine fishery in the 1940s, its mission has evolved and in 2004 the cruises began recording baleen whale observations (among many other data). While this is an impressively longstanding study with remarkable consistency, the temporal resolution of the data is low (quarterly). Throughout a nine-year period, between 2004 and 2012, the CalCOFI cruises observed blue whales (at least one) 121 times, accounting for 212 blue whale records (Campbell et al., 2012). This dataset is referred to as the “science-quality” whale observation data used in this study. It includes 29 blue whale observations accounting for 64 blue whale sightings within the study area and time frame.

Non-systematic research cruises (i.e., opportunistic, non-transect) typically experience a higher number of whale sightings. Without the gridded constraints of a transect line, this more dynamic approach can purposefully place the observer in areas of high whale density. Using this method, professional scientists sighted 2,403 blue whales in the waters off the coasts of Washington, Oregon, and California in a seven-year period between 1991 and 1997 (Calambokidas and Barlow, 2004). While this opportunistic approach might be more efficient at finding and documenting whales, it is also much more susceptible to recounting the same whale multiple times. Testament to this, using photo-identification techniques, the authors of this study determined that only 908 of the 2,403 blue whales were unique individuals (Calambokidas and Barlow, 2004).

Whale-watch observation data

Perhaps the most opportunistic of all whale observations are the ones made aboard commercial whale-watching vessels. An increasingly popular eco-tourism activity, whale-watching tours can be found in nearly every harbor in Southern California. Operating daily and simultaneously throughout the Southern California Bight, these vessels are potentially a rich source of whale presence data with very high temporal resolution. The whale-watch data used in this study was collected from five whale-watching vessels operating in four ports within the Southern California Bight: two vessels in San Diego, and one each in Dana Point, Newport Beach, and San Pedro. Together, in a five-year period, these vessels have logged over 875 blue whale observations accounting for more than 2,250 individual whale counts (Figure 1 and Table 1).

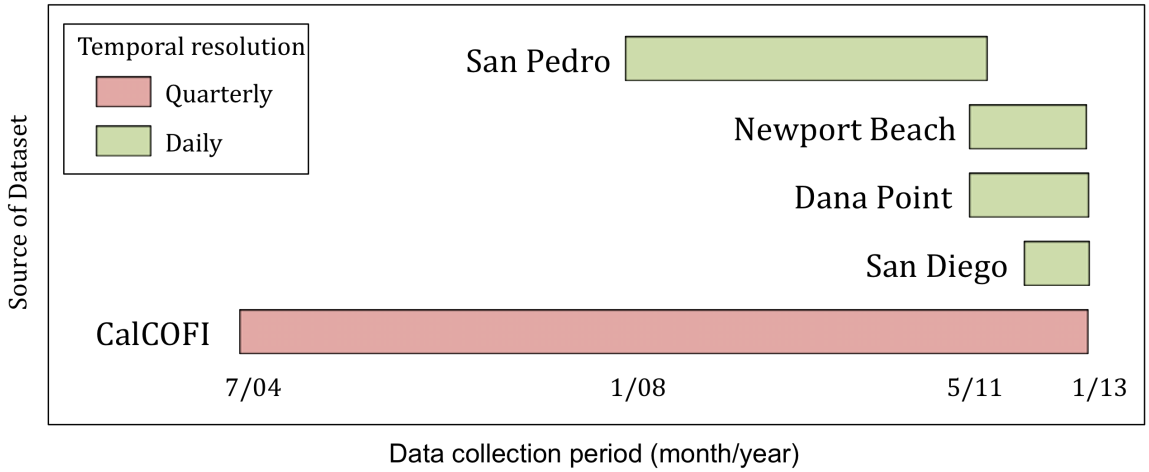

Figure 1: Gantt chart illustrating start dates and end dates of data collection by various sources. Green indicates daily temporal resolution of whale-watch datasets, red indicates quarterly temporal resolution of the science-quality CalCOFI dataset.

Table 1: Metadata describing the five individual datasets used in this study. The VGI dataset is comprised of whale observations from whale-watching vessels operating out of San Pedro, Newport Beach, Dana Point, and San Diego. The science-quality dataset is collected and maintained by CalCOFI.

Dataset	**Start date**	**End date**	**Duration (years)**	**Temporal resolution**	**Blue whales counted**
San Pedro	1/2/08	9/29/12	4.74	Daily	1412
Newport Beach	5/14/11	11/14/12	1.51	Daily	306
Dana Point	5/11/11	12/10/12	1.59	Daily	315
San Diego	1/29/12	12/1/12	0.84	Daily	217
CalCOFI	7/28/04	11/5/12	8.27	Quarterly	64

Environmental Variables

Bathymetry, sea surface temperature, and chlorophyll-a are used here as indicators of potential blue whale habitat (Table 2). These oceanographic parameters were chosen due to their prevalent use in blue whale habitat modeling per the scientific literature.

Bathymetry

Bathymetry is the measure of the ocean's depth. This variable is highly correlated with the fine-scale presence and distribution of blue whales in the literature. The bathymetric data used in this study, ETOPO1, is a product of NOAA's National Geophysical Data Center. This global relief model is a compilation of numerous global and regional datasets. Used here in raster format, the data has a spatial resolution of 0.016667 arc-degrees (1.85 km^2) and a vertical accuracy of about 10 meters.

Sea Surface Temperature

Sea Surface Temperature is the measured temperature of the uppermost layer of the ocean to about one meter (Campbell and Wynne, 2011). Temperature is an important environmental variable that can influence the presence and distribution of many marine species. While blue whales are capable of residing in a wide range of temperatures, the literature suggests that SST can be used as an indicator of prey production, thus blue whale occurrence. The SST data used here is captured by the Moderate Resolution Imaging Spectroradiometer (MODIS), a 36-band spectroradiometer measuring visible and infrared radiation, onboard NASA's Aqua satellite. Specifics of the MODIS instrument can be found at: http://aqua.nasa.gov/about/instrument_modis.php. The Aqua satellite scans the earth's surface every one to two days recording SST with a spatial resolution of 0.0125 arc-degrees (1.47 km^2). The Goddard Ocean Biology Processing Group processes this raw data using multi-sensor level-1 to level-2 software. More on this process can be accessed via the SeaWiFS data processing website at: http://oceancolor.gsfc.nasa.gov/DOCS/SW_proc.html. Processed SST values are validated with in-situ SST buoy data. This dataset covers the eastern pacific (155º W to 105º W Longitude, 22º N to 51º N Latitude) and maintains a nominal accuracy of ± 1 degree Celsius. Summary data of one-month averages are used in this study.

Chlorophyll-a

Chlorophyll-a is a specific pigment essential to photosynthesis. Found in the cells of many photosynthesizing marine organisms, this pigment can be detected via remote sensing satellites and is a strong indicator of primary productivity. The chlorophyll-a dataset used here was obtained from MODIS aboard NASA's Aqua satellite. Concentrations of chlorophyll-a are recorded in milligrams per $meter^3$ with a spatial resolution of 0.0.05 arc-degrees (5.55 km^2). Raw data is processed at the Goddard Space Flight Center using SeaWIFS Data Analysis System software (NOAA Coast Watch, 2013). With near-global spatial coverage (180º W to 180º E longitude, and 75º N to 75º S Latitude) this dataset maintains a nominal accuracy of 40 percent. It is important to note the significant discrepancies that exist between datasets derived from different remote sensing apparatus as well as with high-quality in-situ measurements (NOAA Coast Watch, 2013). Summary data of one-month averages are used in this study.

Table 2: Descriptions of environmental variable datasets. Each dataset was downloaded from NOAA's Coast Watch Program, (http://coastwatch.pfeg.noaa.gov/coastwatch/CWBrowser.jsp).

Environmental variable	**Source / instrumentation**	**Spatial resolution**	**Temporal Resolution**	**Unit of measurement**
Bathymetry	ETOPO1	1.85 km^2	Static	Meter (m)
SST	Aqua MODIS	1.47 km^2	1-month average	Degrees Celsius
Chlorophyll-a	Aqua MODIS	5.55 km^2	1-month average	Milligrams/$meter^3$

Chapter 4: Methodology

Science-quality whale observation data

The science-quality whale observation data used in this study was collected as part of the CalCOFI project. Four times per year research vessels follow a series of transect lines within the CalCOFI study area. Trained observers scan the ocean, during daylight hours and when weather conditions permit, using 7x power binoculars. When a whale is spotted, observers use 18x power binoculars to aid in the identification of the species. Each whale sighting is logged and includes distance and bearing from ship, species identification, group size, group composition, and the animal's behavior (Campbell et al., 2009). Greg Campbell of Scripps Institute of Oceanography currently maintains CalCOFI cetacean data. At the time of writing this thesis the CalCOFI cetacean data is undergoing comprehensive quality control, and is not yet publicly available. Campbell generously, and personally, provided the CalCOFI blue whale data used here.

Whale-watch observation data set

Each commercial vessel operating more than three miles offshore is required by the United States Coast Guard to maintain a logbook documenting all maintenance of safety equipment and the dates emergency drills are performed. In addition, the logbooks often list dates of re-fueling, oil changes, and miscellaneous other maintenance operations, number of passengers, length of voyage, and the coordinates of ash dispersal during burials at sea, among other data. Apart from this information, vessels involved in whale-watching tours often record the locations of whale observations. Their motivation to record a whale's location is simple; if you know where a whale was, you have a better chance of finding it on a future trip. Many of these vessels are running two or three trips daily, and if an opportunity arises to revisit the same whale later in the day they will.

The methods of whale detection and documentation employed by whale-watching vessels vary. In general, the first boat out in the morning will set a course to an area frequented by wildlife the day before. While motoring to this location, the captain and possibly one or two deckhands will scan the area for signs of whale activity. These observers are usually equipped with binoculars of varying optical magnification. Generally, the first and or closest whale to be spotted will be visited and should be recorded. Because the purpose of the trip is to let the passengers experience a whale, rather than count every whale in the area, the vessel will likely stay with one or two whales for a long period of time before continuing the search or returning to port. It is usually the whale's last known location that is recorded in the logbooks. A whale's location is used to revisit the same whale on a later excursion or to share with other whale-watching vessels in the area; the captains of these vessels maintain open dialog via VHF radio.

Currently, not every commercial whale-watching vessel maintains a log of whale sightings. Furthermore, of the vessels that do record whale sightings, not all contain geographic references. After inquiring at each whale-watching operation in the study area, five vessels were found to maintain georeferenced whale records. Each of these vessels was visited and the logbooks were recorded using the camera-video function on an iPhone (the captains all preferred that the logbooks not physically leave the vessels).

Later, the video recordings were played back, and the whale records were viewed and manually transfer into an Excel spreadsheet. To control and assure the quality of this data, and to minimize error in the transcription process, the following measures were taken. If the species and/or coordinates of any logbook entry were undecipherable by two persons, they were not included in the digital dataset. If an entry was legible by only one of two persons, a third party was used to confirm or repudiate the record. Only when the unbiased third party confirmed was the entry included. For every 100 entries in Excel, 15 were chosen at random and crosschecked with the original source document. Entries were imported into ArcMap 10.1 (Figure 2).

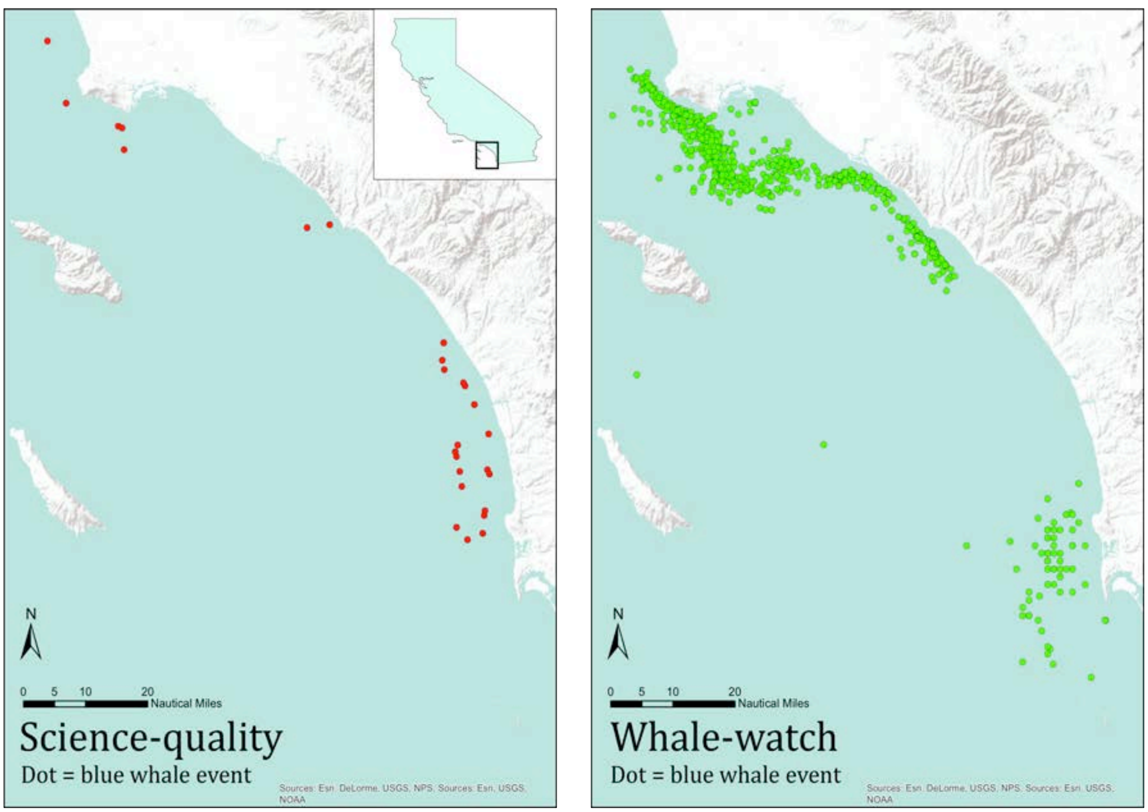

Figure 2: Two maps comparing the total number of blue whale observations contained in each dataset. Left: Science-quality dataset collected over 9 years from 2004-2012 yields 26 blue whale observations (red dots) representing 64 blue whale counts. Right: Whale-watch data collected over 5 years from 2008-2012 yields 847 whale observations (green dots) representing 2,250 blue whale counts.

Seasonal Variability

The seasonal variability of blue, fin, and minke whales was quantified using circular statistics. One year is divided into 360 degrees with months replacing the corresponding angles on a circular plot: 0 degrees is replaced with January, 30 degrees is replaced with February, etc. Averaged monthly whale abundance values are recorded as radii. These radar plots demonstrate mean whale abundances for each month at each location throughout the year. The following equation was used to calculate the mean angle and mean radius for each species at each location (Zar, 1996):

$$X = \frac{\sum f_i \cos a_i}{\sum f_i} \text{ and } Y = \frac{\sum f_i \sin a_i}{\sum f_i}$$

$$r = \sqrt{X^2 + Y^2}$$

Where:

f_i is one month's whale abundance in each month

a_i is each months corresponding angle

a is the mean angle

r is the radius of the mean vector

The mean angle (a) represents the time of year (month) when a particular whale species, on average, can be observed. The radius of the mean vector (r) is a measure of the amount of seasonal variability displayed by a species. This value ranges from a minimum of zero to a maximum of one. When r equals zero, there is said to be no detectable seasonal variability. A value of one indicates a significant change in species occurrences between seasons.

Habitat Suitability Models

Maximum entropy (Maxent) is a sophisticated approach to modeling a species' geographic distribution (Phillips et al., 2004). Using a general-purpose, machine-learning method, Maxent can generate probabilistic habitat suitability analyses describing the spatial and temporal extent of a given species. Using a set of data points marking where a species has been observed, and the environmental conditions associated with each at the time the observations were made, Maxent will estimate the environmental requirements for that species. The information is then used to estimate the range and distribution of this species in non-sampled regions where the environmental conditions are known. It is assumed that the localities of the sample points are collected without concern or influence of the environmental variables used in the model.

Maxent is free software that can be downloaded from the Internet (http://www.cs.princeton.edu/~schapire/Maxent/). It requires all species location data points to be comma-separated values (CSV) in the form of species, longitude, and latitude. This task was performed using Microsoft Excel and the data points were uploaded into the Maxent software using the browser function on the Maxent interface. Similarly, a directory containing the environmental variable files to be used in the model (bathymetry, SST, chlorophyll-a) was uploaded. The files in this directory must all be in ASCII format and contain the same geographic reference system, geographic extent, and grid cell size (bathymetry and chlorophyll-a datasets are resampled to conform to the SST grid cell size of 1.47 km^2). Each of these formatting requirements was executed using ArcMap 10.1. First, a map was composed including a raster file of each environmental variable and a shapefile of the study area. In the Spatial Analyst toolbox, the Extraction by Mask tool was used to extract the study area from each raster file. The dialog box for this operation allows the user to set the parameters for each output file; the geographic extent, geographic reference system, and grid cell size were set to the same values for each environmental variable. The conversion tool in the spatial analyst toolbox was used to convert the extracted raster layers to ASCII files (a Maxent requirement). Once the species locations and environmental variables were uploaded into the Maxent software, the model could be performed. Before running a model, several parameters were adjusted in the Maxent settings field. The number of samples to be set aside for testing was set to 25 percent, allowing the performance of the resulting model to be tested using a random selection of 25 percent of the species location points. The number of model replications was set to 15. This tells the software to independently create 15 versions of the model and to average the results. The resulting model is theoretically more robust than any single replication. This process was repeated for each month of whale observation data.

Samples With Data

A second approach to running a Maxent model is to provide each species location data point with an individual set of environmental variable values corresponding to the time and place the observation was made. This method is referred to as the samples with data (SWD) format. It can be advantageous when dealing with sample points collected during different time periods and thus different environmental conditions. And because a Maxent model's performance increases with an increase in training data (Phillips and Dudik, 2008), this approach can lead to higher performance by increasing the sample size. SWD was chosen due to the modest sample size and long periods of time between sample points in the CalCOFI dataset.

The following procedure was used to create a file of sample points with environmental data to be used in SWD format. A series of maps was created containing sample points collected within the same time period (month) and the corresponding environmental variables for that time period. This was accomplished in ArcMap. In the Spatial Analyst toolbox the Extract Multiple Values to Points tool was used to affix the environmental variable values to each corresponding whale observation point location. This procedure was repeated for each month's set of whale observations. The amended attribute tables for the point locations were copied and pasted into Excel and saved as a CSV file. These files were uploaded into Maxent.

Model Comparison

Model outputs were compared and contrasted using map algebra, a way to analyze multiple maps using algebraic expressions. Used here to view the differences and similarities between two Maxent model outputs, the analysis was accomplished using ArcMaps's map algebra function in the Spatial Analyst Toolbox. Using the raster calculator within the map algebra toolset, cell values from a model using volunteered whale-watch data were subtracted from the corresponding cell values of a model using science-quality whale presence data. The map algebra output file demonstrates areas where the two models agree and disagree (Figure 9).

Chapter 5: Results

Volunteered Whale Observation Dataset

The volunteered dataset composed here of the observational records of five whale-watching vessels included over 3,000 logbook pages describing five years of whale-watching activity within the study area. Held in these pages were 2,250 blue whale observations, 1,218 fin whale observations, and 172 minke whale observations. Also found in these logbooks were accounts of rare whales seldom seen on infrequent scientific surveys, including sperm whales (*Physeter macrocephalus*), orca whales (*Orcinus orca*), sei whales (*Balaenoptera borealis*), and bairds beaked whales (*Berardius bairdii*). Figure 3 demonstrates the spatial patterns of opportunistic observations of blue whales.

Seasonal variability

Each whale species observed and recorded by the whale-watching vessels displayed unique seasonal variability. Table 3 shows the statistical results of the seasonal variability analysis. The radius of the mean vector (r) is a measure of the amount of seasonal variability displayed by a species. This value ranges from a minimum of zero to a maximum of one. When r equals zero there is said to be no detectable seasonal variability. A value of one indicates a significant change in species occurrences between seasons. Among the whales in the volunteered dataset the blue whale exhibited the strongest and most consistent seasonal variability (r=0.83). The mean angle (*a*) represents the time of year (month) when a particular whale species, on average, can be observed. The cumulative mean angle for blue whales (averaged over four locations) is 206º, suggesting that blue whales are most likely to be observed in late July. Radar plots show each month's mean whale abundance for blue, fin, and minke whales at each location (Figure 4).

Spatial patterns of opportunistic observations of blue whales

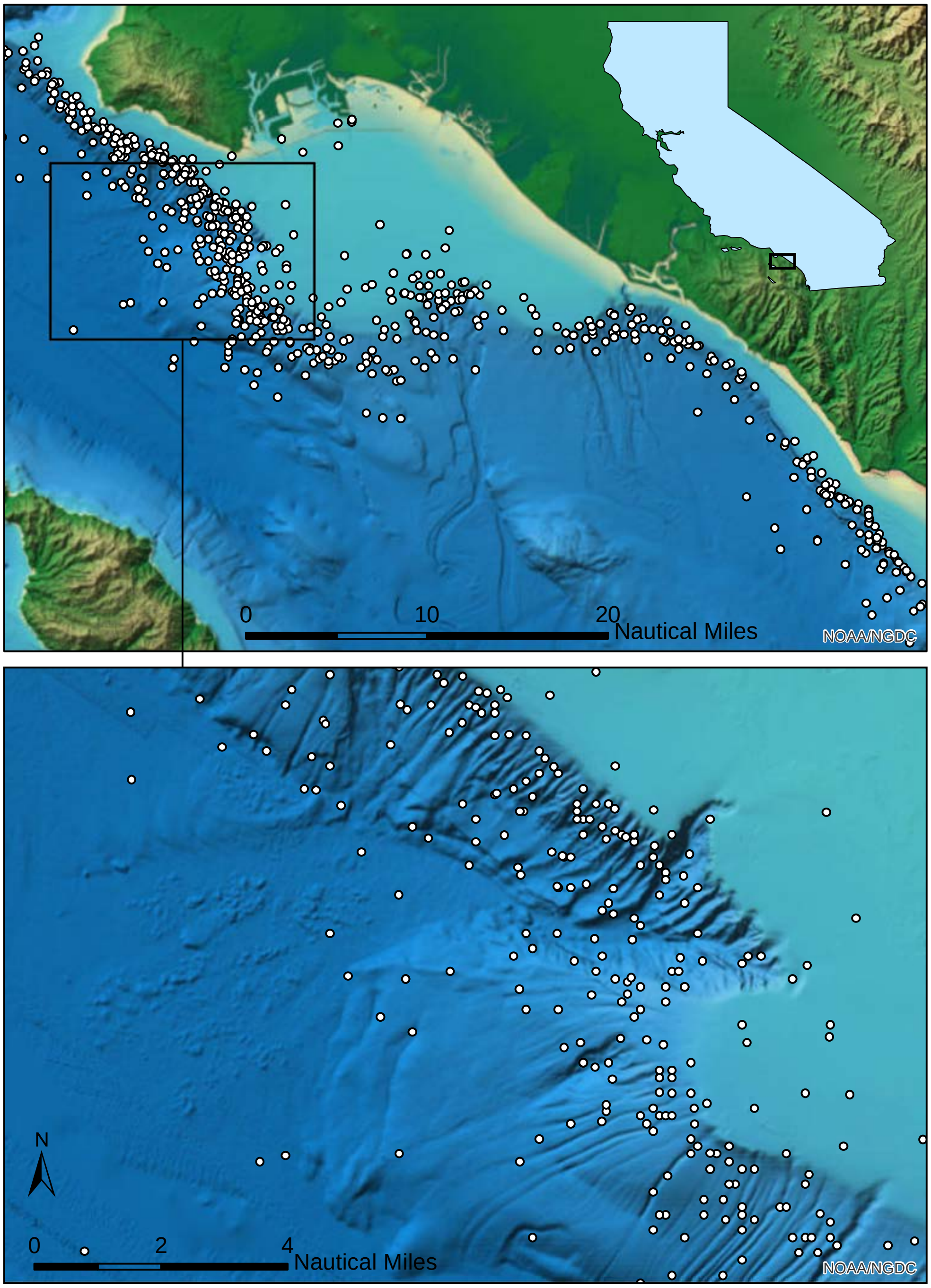

Figure 3: Spatial patterns of blue whale observations (white dots) and associated bathymetry. The average depth of blue whale observations was 297 meters (162 fathoms). Within the study area, this depth is associated with steep bathymetric features thought to be responsible for blue whale prey production, accumulation, and retention (Croll et al., 1999).

Table 3: Statistical results of the seasonal variability analysis and the contributing data for the ensuing radar plots. Mean angle (a) represents the time of year (month) when a particular whale species, on average, can be observed (January=0 º, April=90 º, July=180 º, etc.). Radius of the mean vector (r) is a measure of the amount of seasonal variability displayed by a species (0=low, 1=high).

Species	**Location**	**Number of whales (n)**	**Mean angle (a)**	**Mean vector (r)**	**Cumulative mean angle**	**Cumulative mean vector**
Blue	San Pedro	1412	214	0.79	206 º	0.83
	Newport Beach	306	204	0.82		
	Dana Point	315	216	0.83		
	San Diego	217	188	0.89		
Fin	San Pedro	854	359	0.35	254 º	0.43
	Newport Beach	111	260	0.42		
	Dana Point	96	268	0.56		
	San Diego	157	130	0.40		
Minke	San Pedro	120	228	0.29	170 º	0.57
	Newport Beach	8	248	0.88		
	Dana Point	24	142	0.78		
	San Diego	20	60	0.34		

Seasonal Variability

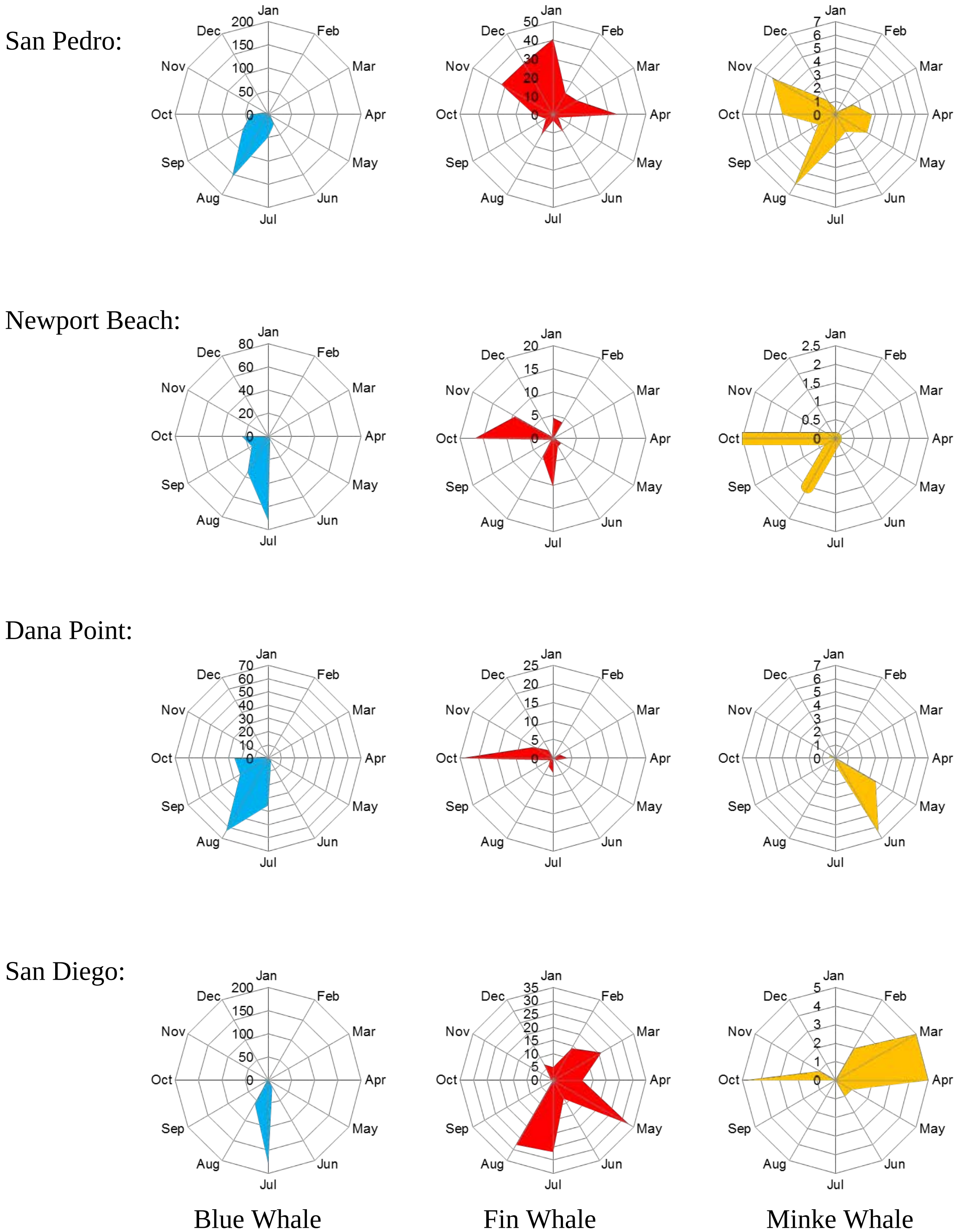

Figure 4: Radar plots showing seasonal variability of three species of baleen whale observed within the study area. Among these, blue whales exhibit the highest degree of seasonal variability. Rings represent numbers of whales; radii represent seasonal occurrence.

Model Results

Maxent models demonstrate potential blue whale foraging habitat within the Southern California Continental Borderland (Figures 5, 6, and 7). Warmer colors (yellows and reds) indicate higher probability of suitable habitat, while cooler colors (blues and greens) indicate lower probability of suitable habitat. Because bathymetry was given such a high predictive value by the computer-learning models, seasonal variations due to temperature and chlorophyll-a are minimal. Because of this, the aforementioned phenology of each species is highly important when analyzing these models. Blue whales are primarily observed in the study area from July through September and models projected onto other months will not be as accurate. Models produced using the whale-watch data did not differ qualitatively from models using the science-quality data. Both models rank the influence of environmental variables in identical order with similar model contribution values assigned to each variable (Table 4). The extrapolated predictions of each model are also very similar. The whale-watch data produced a habitat suitability model with more definitive predictions, while the science-quality data produced a more generalized model. This is most apparent throughout the southern portion of the models in Vizcaino Bay, and is likely a result of the differences in model training sample sizes.

Table 4: Each environmental variable's relative contribution to the Maxent model given as a percent. The rank of contribution among environmental variables was consistent among models.

Whale Presence Data	Bathymetry	Sea Surface Temperature	Chlorophyll-a
Science-quality (n=30)	51.9	29.6	18.5
Whale-watch (n=30)	60.6	26.9	12.5
Whale-watch (n=250)	49.7	34.1	16.1

Maxent model: science-quality dataset (n=30)

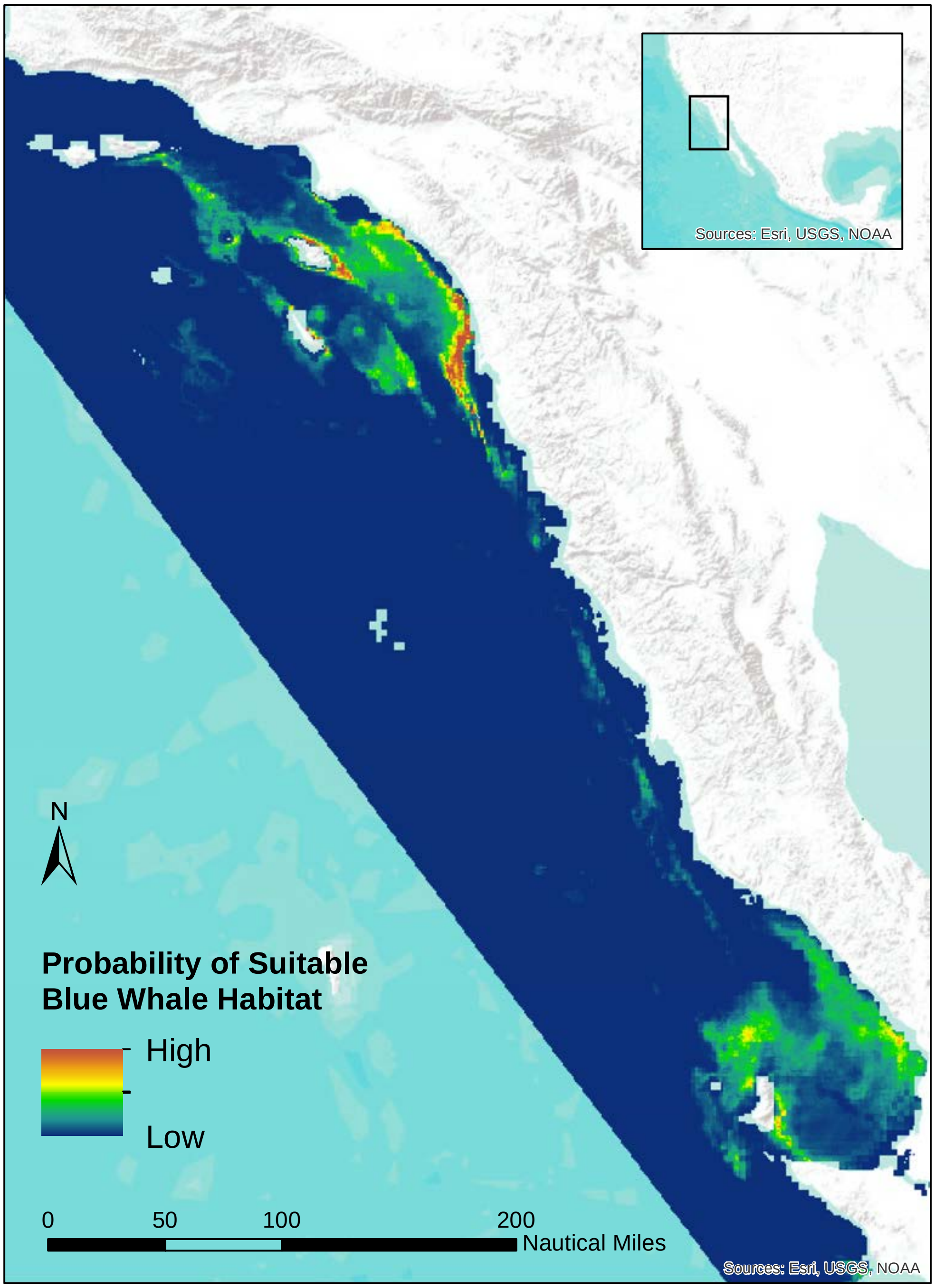

Figure 5: Probabilistic blue whale habitat suitability model using 30 samples of blue whale locations collected during CalCOFI cruises between 2004 and 2012. The model is projected onto the environmental conditions of August 2011. Warm colors indicate regions with a high probability of suitable habitat; cool colors indicate regions with lower probability of suitable habitat (AUC=0.945).

Maxent model: whale-watch dataset (n=30)

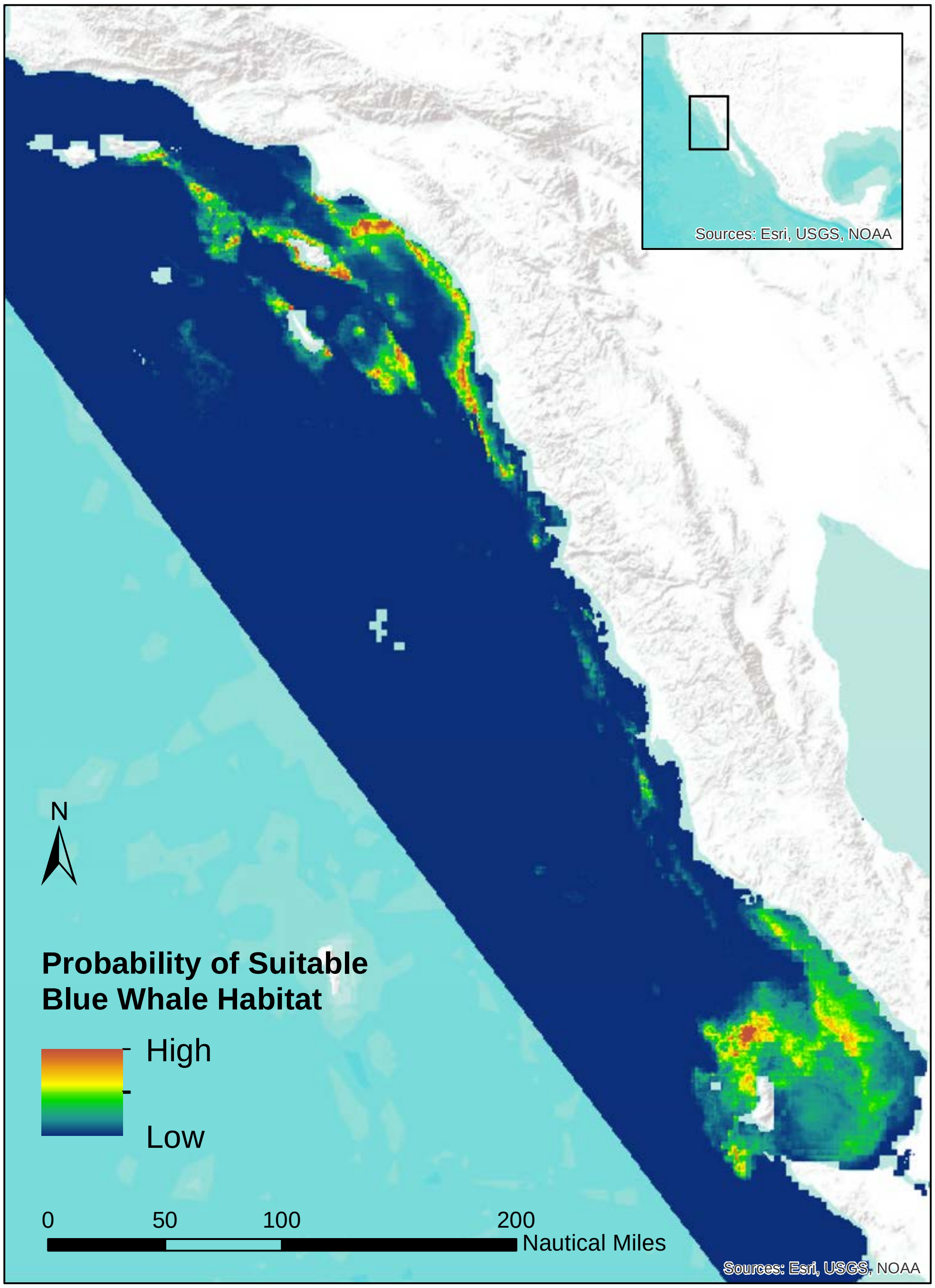

Figure 6: Probabilistic blue whale habitat suitability model using 30 random samples of blue whale locations collected by whale-watching vessels between 2008 and 2012. The model is projected onto the environmental conditions of August 2011. Warm colors indicate regions with a high probability of suitable habitat; cool colors indicate regions with lower probability of suitable habitat (AUC=0.964).

Maxent model: whale-watch (n=250)

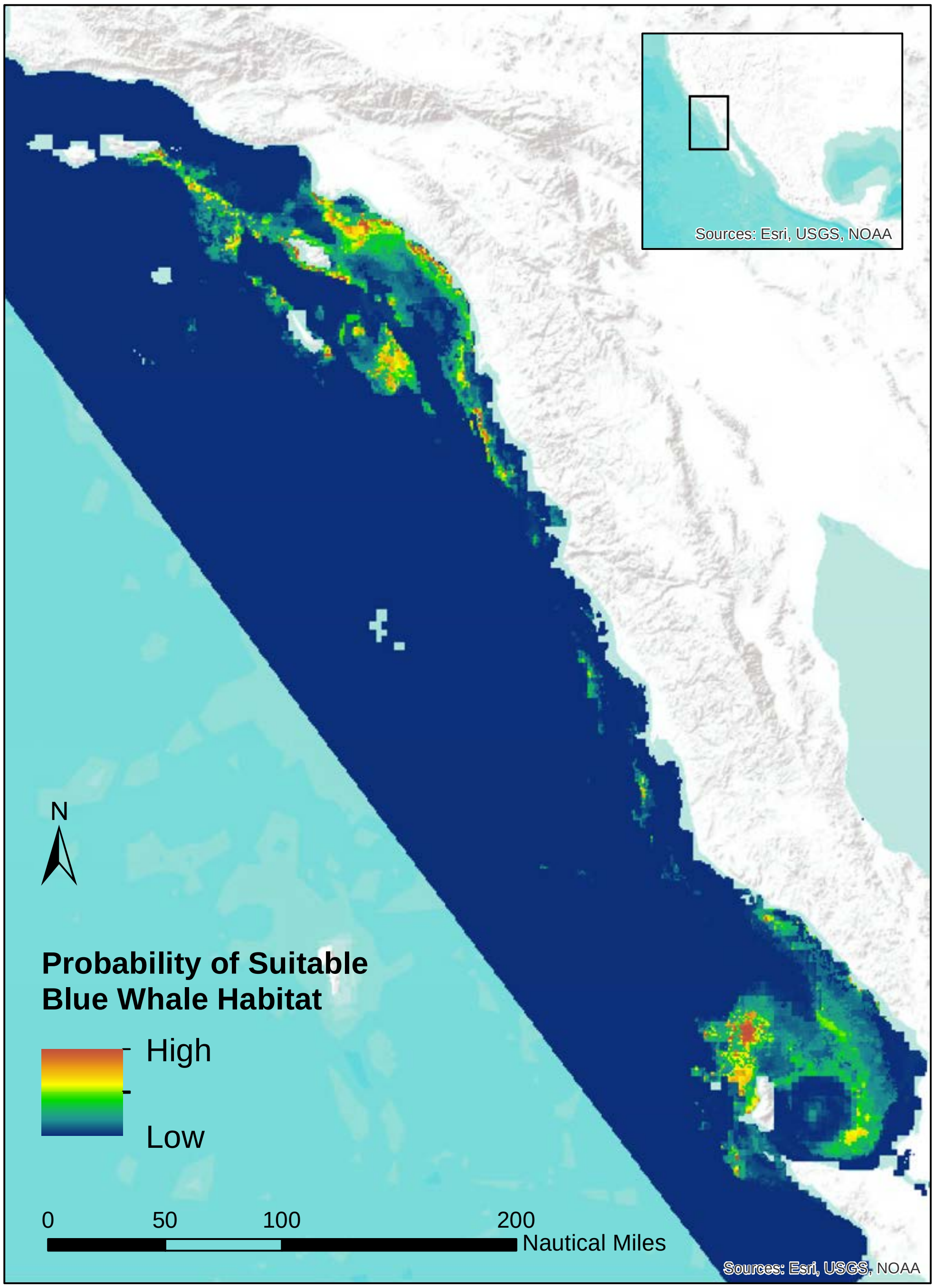

Figure 7: Probabilistic blue whale habitat suitability model using 250 locations of blue whales collected by whale-watching vessels between 2008 and 2012. The model is projected onto the environmental conditions of August 2011. Warm colors indicate regions with a high probability of suitable habitat; cool colors indicate regions with lower probability of suitable habitat (AUC= 0.953).

The receiver operating characteristic (ROC) curve is a graphical representation comparing the fraction of true positives versus the fraction of false positives committed by the model during test runs (Figure 8). Because presence-only data is used, the rate of commission (false positives) is unable to be calculated. Maxent replaces this statistic with the fraction of the total study area predicted present (Phillips, 2008). The area under this curve (AUC) is the measure of a model's performance. An AUC of 1 indicates a perfect model where every test sample is accurately described. Alternately, an AUC value of 0.5 describes a random model with a predictive average of 50 percent. This means half of the predictions are erroneous and half are accurate (a performance that can be achieved by flipping a coin). The AUC for both the science-quality and whale-watch-generated models indicate high levels of performance (0.945 and 0.953, respectively). The increase in performance with the whale-watch data is most likely the result of a larger dataset.

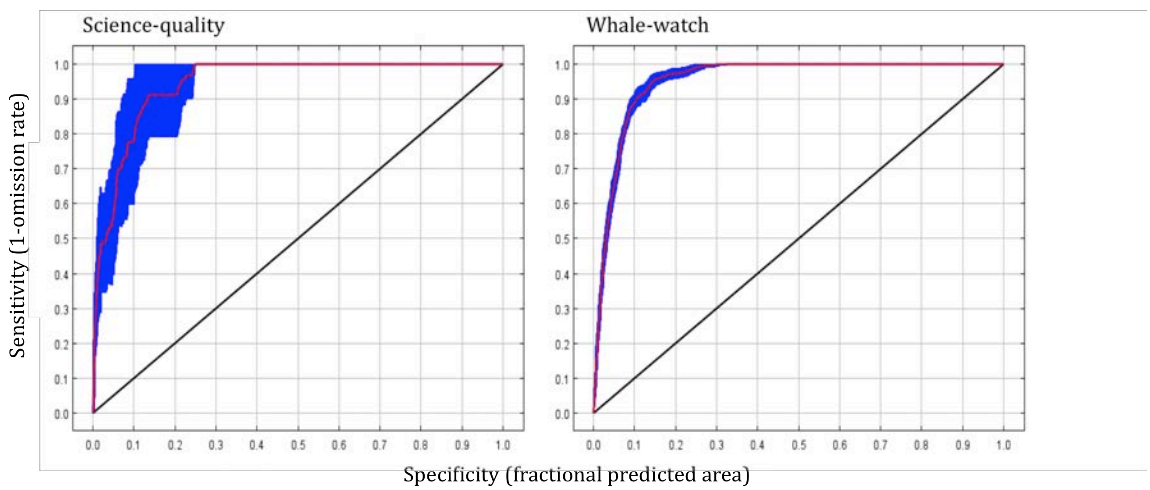

Figure 8: Average model sensitivity vs. specificity. The red line illustrates the mean area under the curve (AUC). The blue buffer shows the mean standard deviation and the black line represents random prediction. The two charts compare model performance between science-quality (AUC=0.945) and whale-watch (AUC=0.953) datasets.

The calculated omission rate for whale-watch data was very similar to the predicted omission rate, another indication of high performance (Figure 9). The omission rate for the science-survey data did not conform to the line of predicted rate of omission, indicating a less robust model. The probabilistic habitat suitability models show an increase in predictive performance, with an increase in sample size. As a result, the large dataset provided by the whale-watch vessels out-performed the science-quality data with a smaller sample size.

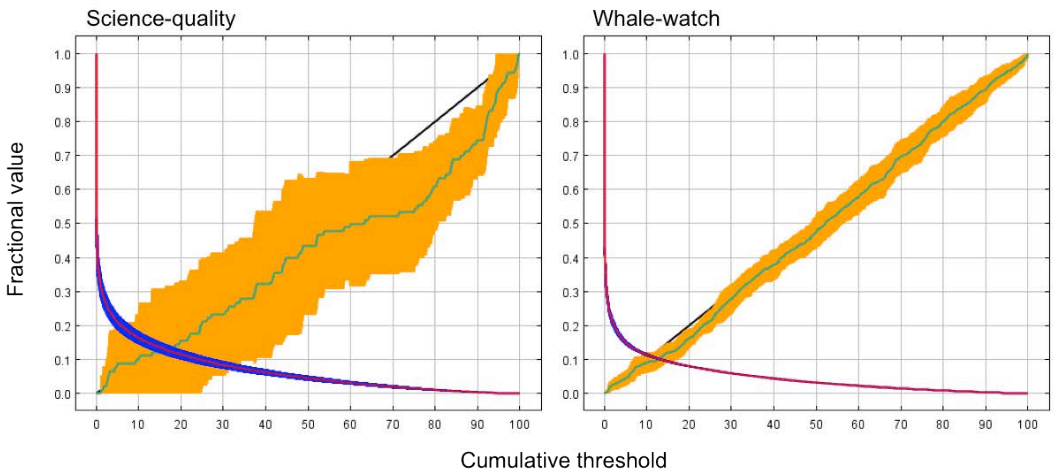

Figure 9: Average omission and predicted area. The black line (largely hidden behind yellow) represents the model's predicted omission. The green line depicts the mean omission on test data and the orange buffer illustrates the mean standard deviation.

Model Comparison

An algebraic comparison (whale-watch model output subtracted from science-quality model output) shows significant agreement between the two models (Figure 10). Red shows areas with a high probability of suitable blue whale habitat as predicted by the science-quality data and less suitable by the whale-watch data. Green, on the contrary, indicates areas predicted as highly suitable by the whale-watch data and less suitable by the science-quality data. Yellow shows areas where the two models are in agreement. The main difference between the models is the nearness to shore of predicted blue whale foraging habitat. Models produced using whale-

Model Comparison: whale-watch model subtracted from science-quality model

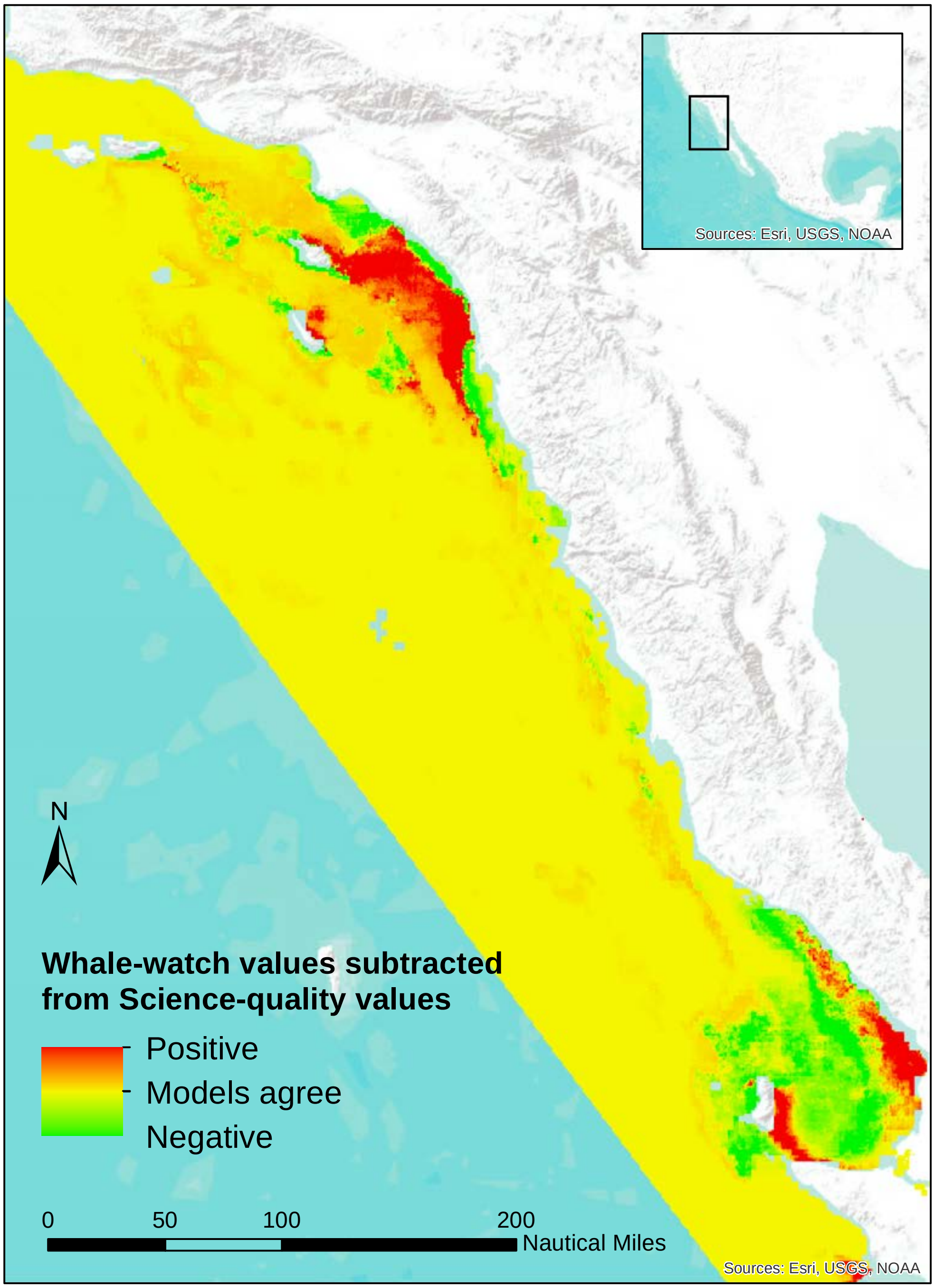

Figure 10: Using map algebra, cell values from the whale-watch model (Figure 6) subtracted from corresponding cell values of the science-quality model (Figure 5). Red indicates areas predicted highly suitable by the science-quality data and less suitable by the whale-watch data. Green indicates areas predicted highly suitable by the whale-watch data and less suitable by the science-quality data. Yellow shows areas where the two models are in agreement.

watch data are more likely to imply near-shore foraging areas than are models using science-quality whale observation data. And, vice-versa, the models using science-quality whale data are more likely to suggest offshore suitable habitat. This result may be a consequence of the nature and behavior of the whale-watching industry: Their restricted temporal and spatial ranges for individual trips limit their offshore observations. The science-quality whale data are supplied by cruises operating further offshore than the typical whale-watching vessel.

Chapter 6: Discussion and Conclusion

Volunteered whale-watch data proved to be highly effective for the creation of probabilistic habitat suitability models. The large and inexpensive dataset created here of whales occurring over a vast geographic area is characteristic of VGI. And the high temporal resolution of this data can give insight into phenomena that cannot be detected from infrequent scientific surveys. Because of Maxent's tolerance to presence-only data it is well equipped to utilize this highly opportunistic dataset. And the large sample size provided the software with ample training information to learn and test the blue whale's habitat requirements.

The use of Maxent in whale habitat modeling can be further explored by the use of additional environmental variables. With the apparent correlation between blue whale occurrences and steep bathymetric features, slope and aspect values may be of significant importance. These variables can be derived from the same bathymetric dataset and can help explain in more detail the importance underwater features on blue whale prey production and retention. Additionally, the use of oceanic current data may help describe the spatial separation between these features and the areas where whales are sighted. By incorporating these and other environmental variables the predictive power of the model may be enhanced.

Consistent in each model is an area of highly probable blue whale foraging habitat in Vizcaino Bay, Mexico. This area, located in the middle of the Baja peninsula, occurs outside of the CalCOFI study area and out of range for Southern California whale-watching vessels. As a result, the whale observation datasets used in this thesis could not confirm this prediction. Model confirmation comes from a study conducted in 1995 where five blue whales were tagged in the Santa Barbara Channel, California. During the time the tags remained attached to the whale's bodies they allowed researchers to track the paths of these animals (Figure 11). Of the five whales tagged, four made southerly routes to Vizcaino Bay (Mate and Calambokidis, 1999). Furthermore, three of the four southbound whales followed a path of suitable habitat as predicted by the Maxent model.

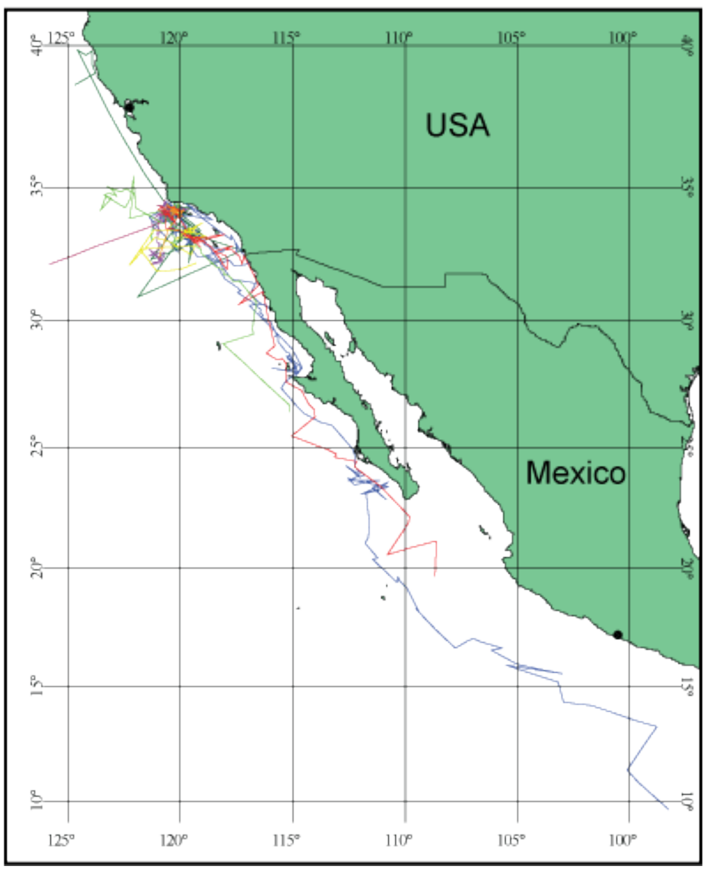

Figure 11: Tracks of five blue whales tagged in the Santa Barbara Channel in 1994 and 1995 are displayed in various colors. Four of the five whales traveled south making temporary stops in Vizcaino Bay (Mate and Calambokidis, 1999). Vizcaino Bay is identified as highly probable blue whale foraging habitat by the Maxent models.

While whale-watch data is successfully used here in modeling blue whale foraging habitat, several issues concerning the quality of this data limit its use in other scientific pursuits. For this type of data to be used in a more encompassing scientific capacity, the following issues must be addressed. The nature and overall intent of a whale-watching cruise is fundamentally different from a true research cruise. Whale-watching vessels are limited by the short duration of their search; a normal whale-watching excursion ranges from two to five hours in length. As a result, they tend to cover the same geographic area trip after trip. Whale-watching vessels may run two or more trips per day, and to save fuel they will often only visit whales closest to the harbor-relying heavily on locations of previous whale sightings. This can lead to recounting the same whale as well as omitting whales that are outside of this limited area. Once a whale is sited, whale-watching vessels tend to stay with the whale for a prolonged period of time instead of continuing to search for different whales in the area. This cessation of search will result in lower whale counts for the larger geographic area.

Because the nature and behavior of a whale-watching tour boat strongly influences the pattern of its whale observations, the spatial patterns of opportunistic observations of whales do not reflect the true spatial arrangement of the population. Maps produced using whale-watch observation data are indicative of where whales are seen on eco-tours rather than how the area's whale population is distributed. The opportunistic observations do, however, reveal an interesting pattern; the vast majority of the observations were made on or near the 300m-isobath, a geographic feature locally associated with high bathymetric relief. This phenomenon is also detected in the science-quality dataset and supported in the literature.

By accounting for the observational effort exhibited by a whale-watching vessel, many of the aforementioned concerns may be mitigated. This can be accomplished by logging the hours of active whale searching in addition to total hours of excursion. In doing so, researchers can better understand the density or scarcity of whales within the area. In addition, recording GPS tracks of each excursion can help researchers more accurately define the observational area of effort covered by a whale-watching vessel.

Whale-watch data is also limited in other aspects of marine mammal research. Without a means of identifying individual whales (i.e., photographic identification), opportunistic observations are not conducive to studies of population numbers or dynamics. One solution to this problem is for whale-watching vessels to photograph each whale they encounter with a GPS-enabled digital camera. These geo-tagged photographs can then be analyzed by scientists (citizen or professional) and individual whales within the population can be identified. This would allow whale-watch data to be used for additional scientific purposes other than habitat suitability analyses.

Commercial shipping and blue whale foraging habitat

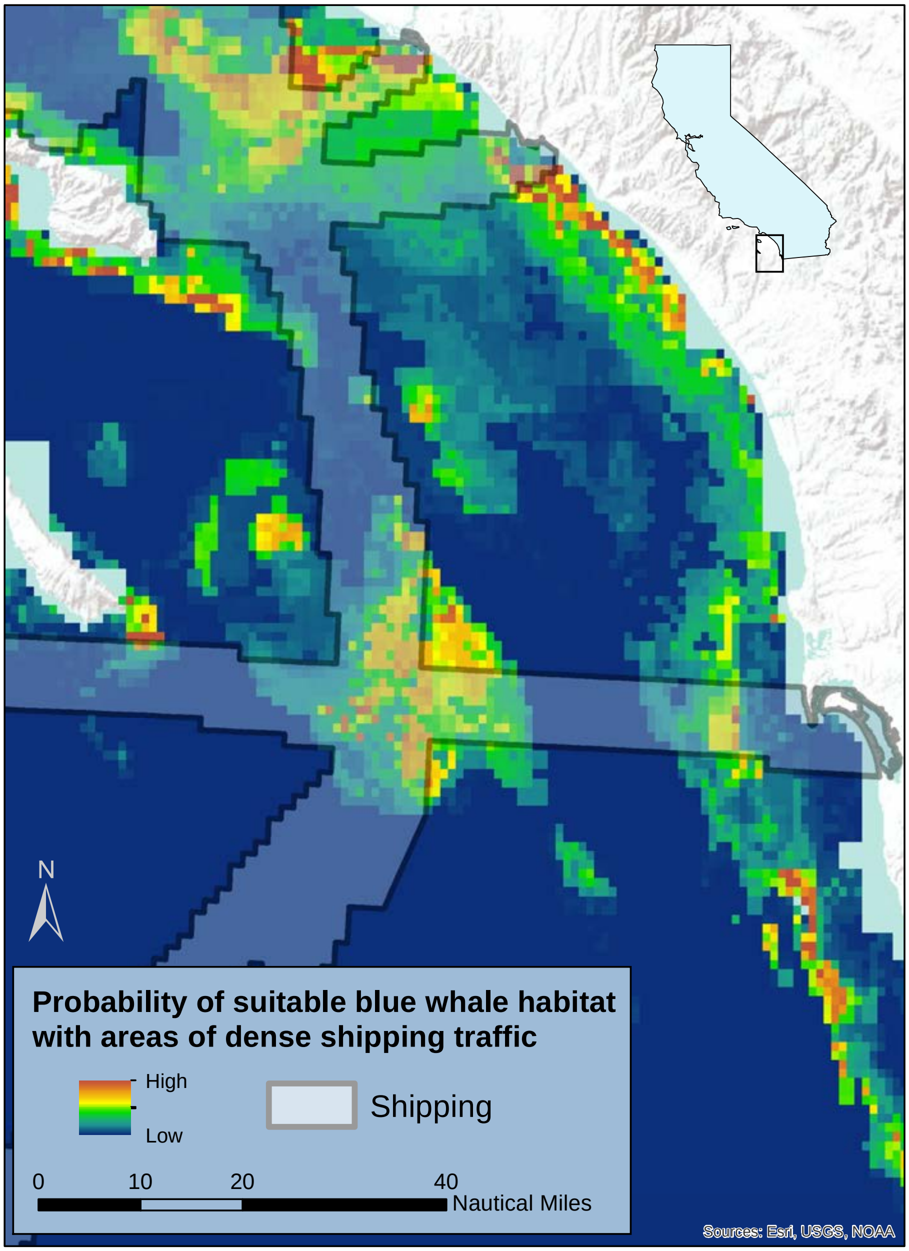

Figure 12: Probability of blue whale habitat versus areas of high shipping activity. Maps such as these can be used to inform marine spatial planners when designing shipping lanes, marine protected areas, and other marine management areas. Note the area of highly suitable blue whale habitat where shipping lanes cross.

The goal of this and similar studies is to alert and bring awareness to boat operators when in the vicinity of whales. Maps incorporating marine traffic and potential whale habitat can be used to inform the maritime community of this spatial conflict (Figure 12). An important aspect of using volunteered geographic information in this pursuit is the direct involvement of the target audience. By involving this demographic in the scientific process they are more likely to be interested in the results and take active roles in the solutions.

Citizen science and volunteered geographic information can be of great service to the future of marine mammal research and marine spatial planning. Currently, the availability of these datasets is limited by the current method of recording data by hand and storing data in logbooks dispersed among many vessels. Future work in this area should include the unification and digitization of whale observations made by whale-watching vessels and citizen scientists in general. This would allow for the acquisition of near real-time whale location data by scientists and marine managers. This streamlining of data acquisition can be accomplished via the creation of a main database capable of receiving data entries from various mobile devices (phones, tablets, etc.). Users of these devices could upload geo-tagged photos of whales in real-time. By increasing the availability of this data and by decreasing the time between data collection and utilization, marine managers will be better equipped to deal with the dynamic spatial conflict between humans and whales.

References

Ballance, L., R. Pitman, and P. Fiedler. 2006. Oceanographic Influences on Seabirds and Cetaceans of the Eastern Tropical Pacific: A review. *Progress in Oceanography* 69:360–390.

Bianchi, J. 1999. Christmas Bird Count. *New York State Conservationist* 54(3):6.

Bonney, R., C. Cooper, J. Dickinson, S. Kelling, T. Phillips, K. Rosenberg, and J. Shirk. 2009. Citizen Science: A Developing Tool for Expanding Science Knowledge and Scientific Literacy. *BioScience* 59(11):977–984.

Burtenshaw, J., E. Oleson, J. Hilderbrand, M. McDonald, R. Andrew, B. Howe, and J. Mercer. 2004. Acoustic and Satellite Remote Sensing of Blue Whale Seasonality and Habitat in the Northeast Pacific. *Deep Sea Research Part II* 51:967–986.

Calambokidis, J., and J. Barlow. 2004. Abundance of Blue and Humpback Whales in the Eastern North Pacific Estimated by Capture-Recapture and Line-Transect Methods. *Marine Mammal Science* 20(1):63–85.

Calambokidis, J., J. Barlow, J. Ford, T. Chandler, and A. Douglas. 2009. Insights Into The Population Structure of Blue Whales in the Eastern North Pacific From Recent Sightings and Photographic Identification. *Marine Mammal Science* 25(4):816–832.

Campbell, G., K. Merkens, and J. Hilderbrand. 2010. *California Cooperative Oceanic Fisheries Investigation (CalCOFI) Cruises: 2009-2010.* San Diego, CA. USA: Marine Physical Laboratory, Scripps Institution of Oceanography.

Campbell, G., L. Roche, H. Basset, A. Simonis, D. Camacho, K. Whitaker, and J. Hilderbrand. 2011. *California Cooperative Oceanic Fisheries Investigation (CalCOFI) Cruises: 2010–2011.* San Diego, CA. USA: Marine Physical Laboratory, Scripps Institution of Oceanography.

Campbell, G., L. Roche, K. Whitaker, and J. Hilderbrand. 2012. *California Cooperative Oceanic Fisheries Investigation (CalCOFI) Cruises: 2011–2012.* San Diego, CA. USA: Marine Physical Laboratory, Scripps Institution of Oceanography.

Campbell, James B. and Randolph H. Wynne. Introduction to Remote Sensing (5th Edition). (Guildford Press, New York, 2011).

Conrad, C. C., and K. G. Hilchey. 2011. A Review of Citizen Science and Community-Based Environmental Monitoring: Issues and Opportunities. *Environmental Monitoring Assessment* 176:273–291.

Croll, D., B. Tershy, R. Hewitt, D. Demer, P. Fiedler, S. Smith, W. Armstrong, J. Popp, T. Kiekher, R. Lopez, J. Urban, and D. Gendron. 1998. An Integrated Approach to the Foraging Ecology of Marine Birds and Mammals. *Deep Sea Research II* 45:1353–1371.

Delaney, D., C. Sperling, C. Adams, and B. Leung. 2008. Marine Invasive Species: Validation of Citizen Science and Implications For National Monitoring Networks. *Biological Invasions* 10:117–128.

Dickinson, J., B. Zuckerberg, and D. Bonter. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* 41:149–172.

Doniol-Valcroze, T., D. Berteaux, P. Larouche, and R. Sears. 2007. Influence of Thermal Fronts on Habitat Selection by Four Rorqual Whale Species in The Gulf of St. Lawrence. *Marine Ecology Progress Series* 335:207-216.

Etnoyer, P., D. Canny, B. Mate, L. Morgan, J. Ortega-Ortiz, and W. Nichols. 2006. Sea-Surface Temperature Gradients Across Blue Whale and Sea Turtle Foraging Trajectories off the Baja California Peninsula, Mexico. *Deep Sea Research Part II* 53:340–358.

Fiedler, P., S. Reilly, R. Hewitt, D. Demer, V. Philbrick, S. Smith, W. Armstrong, D. Croll, B. Tershy, and B. Mate. 1998. Blue Whale Habitat and Prey in the California Channel Islands. *Deep Sea Research II* 45:1781–1801.

Gill, P., M. Morrice, B. Page, R. Pirzl, A. Levings, and M. Coyne. 2011. Blue Whale Habitat Selection and Within-Season Distribution in a Regional Upwelling System off Southern Australia. *Marine Ecology Progress Series* 421:243–263.

Goffredo, S., F. Pensa, P. Neri, A. Orlandi, M. Gagliardi, A. Velardi, C. Piccinetti, and F. Zaccanti. 2010. Unite Research With What Citizens Do For Fun: "Recreational Monitorring" of Marine Biodiversity. *Ecological Applications* 20(8):2170–2187.

Goodchild, M. F. 2008. Commentary: Whither VGI? *GeoJournal* 72:239–244.

Goodchild. M. F. 2007. Citizens as Sensors: The World of Volunteered Geography. *GeoJournal* 69:211–221.

Haklay, M. 2013. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In *Crowdsourcing Geographic Knowledge: Volunteered Geographic105 Information (VGI) in Theory and Practice,* ed. D. Sui, 105–122.

Hass, T. C. 2011. *Improving Natural Resource Management : Ecological and Political Models.* Hoboken, NJ, USA: Wiley.

Henderson, E. E. 2010. Cetaceans in the Southern California Bight: Behavioral, Aucustical, and Spatio-temporal Modeling. Ph.D. dissertation, University of California, San Diego.

International Union for Conservation of Nature. IUCN Red List of Threatened Species. In International Union for Conservation of Nature [database online]. Cambridge, United Kingdom, 2008. Available from http://www.iucnredlist.org/details/2477/0 (last accessed June/15 2013).

Kearney, M., A. Matzelle, and B. Helmuth. 2012. Biomechanics Meets the Ecological Niche: The Importance of Temporal Data Resolution. *Experimental Biology* 15:922–933.

Laist, D., A. Knowlton, J. Mead, A. Collet, and M. Podesta. 2001. Collisions Between Ships and Whales. *Marine Mammal Science* 17:35–75.

Mate, B., and J. Calambokidis. 1999. Movements of North Pacific Blue Whales During the Feeding Season Off Southern California and Their Southern Fall Migration. *Marine Mammal Science* 15(4):1246–1257.

Moore, S., W. Watkins, M. Daher, J. Davies, and M. E. Dahlheim. 2002. Blue Whale Habitat Associations in the Northwest Pacific: Analysis of Remotely-Sensed Data Using a Geographic Information System. *Oceanography* 15(3):20–25.

Munger, L., D. Camacho, A. Havron, G. Campbell, J. Calambokidis, A. Douglas, and J. Hilderbrand. 2009. *Baleen Whale Distribution Relative to Surface Temperature and Zooplankton Abundance off Southern California.* San Diego, California: CalCOFI, Report Number, 50.

National Audubon Society Web Site. In National Audubon Society [database online]. New York, NY. USA, 2013 Available from http://birds.audubon.org/christmas-bird-count (last accessed June 15, 2013).

National Oceanic and Atmospheric Administration. Coast Watch Website. Washington, DC. USA, 2013 Available from http://coastwatch.pfeg.noaa.gov/coastwatch/CWBrowser.jsp (last accessed August 21, 2013).

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum Entropy Modeling of Species Geographic Distributions. *Ecological Modeling* 190:231–259.

Phillips, S. J., and M. Dudik. 2008. Modeling of Species Distributions With Maxent: New Extensions and a Comprehensive Evaluation. *Ecography* 31:161–175.

Phillips, S. J., M. Dudik, and R. E. Schapire. 2004. A Maximum Entropy Approach to Species Distribution Modeling. *Proceedings of the 21st International Conference on Machine Learning* 655–655–662.

Pittman, S., and B. Costa. 2009. Linking Cetaceans to Their Environment: Spatial Data Acquisition, Digital Processing and Predictive Modeling for Marine Spatial Planning in the Northwest Atlantic. In *Spatial Complexity, Informatics, and Wildlife Conservation,* ed. Huettmann & Cushman, 387–408.

Redfern, J., M. McKenna, T. Moore, J. Calambokidis, M. Deangelis, E. Becker, J. Barlow, K. Forney, P. Fiedler, and S. Chivers. 2013. Assessing the Risk of Ships Striking Large Whales in Marine Spatial Planning. *Conservation Biology* 27:292–302.

Southern California Coastal Water Research Project Website. Costa Mesa, CA. USA. Available from http://www.sccwrp.org/researchareas/RegionalMonitoring/BightRegionalMonitoring.aspx. (Last accessed August 21, 2013).

Starks, E. C. 1922. *A History of California Shore Whaling.* San Diego: State of California, Department of Fish and Game, Report Number, Fish Bulletin No. 6.

Trumbull, D., R. Bonney, D. Bascom, and A. Cabral. 2000. Thinking Scientifically During Participation in a Citizen-Science Project. *Science Education* 84:265–275.

Zar, Jerrold H. Biostatistical Analysis (4th Edition). (Prentice Hall, NJ. 1999).