Estimating Cetacean Abundance and Distribution around Vancouver Island, B.C., Using Data Collected from an Opportunistic Platform

by

Nicole Koshure
B.Sc. (Hons.), University of British Columbia, 2002

RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF RESOURCE MANAGEMENT

in the
School of Resource and Environmental Management
Faculty of Environment

Project No. 522

SIMON FRASER UNIVERSITY

April 23, 2012

APPROVAL

Name: **Nicole Koshure**

Degree: **Master of Resource and Environmental Management**

Title of Thesis: **Estimating cetacean abundance and distribution around Vancouver Island, B.C., using data collected from an opportunistic platform**

Project No: **522**

Examining Committee:

Randall Peterman, Supervisor
Professor
School of Resource and Environmental Management
Simon Fraser University
Canada Research Chair in
Fisheries Risk Assessment and Management

Andrew Cooper, Committee Member
Associate Professor
School of Resource and Environmental Management
Simon Fraser University

Date Defended/Approved: ______________________________

ABSTRACT

British Columbia (B.C.) is home to 25 species of cetaceans, six of which are listed as at risk of extinction by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) and under the Species at Risk Act (SARA). It is crucial to learn more about the distribution, abundance, and habitat use of these species. However, because cetaceans can be wide-ranging and distributed sparsely over vast areas, statistically robust systematic studies can be expensive and logistically challenging. To overcome these issues, my research aims to use a platform of opportunistic sightings, Straitwatch, to generate estimates of abundance and spatial distribution for 6 species of cetaceans commonly found off Vancouver Island. During the summers of 2008 and 2009, Straitwatch collected effort and sightings data during non-systematic surveys of three regions. These data were analyzed using model-based distance sampling methods. Results of these analyses are useful for (1) determining cetacean habitat use, (2) identifying and protecting critical habitats or creating marine protected areas, and (3) for identifying areas where cetaceans are at greater risk to anthropogenic disturbance such as entanglement/bycatch in fishing gear.

Keywords: detection function; density surface modelling; distance sampling; distribution and abundance; opportunistic platform; spatial model.

ACKNOWLEDGEMENTS

Thank you to Dr. Randall Peterman for his supervision, support and advice over the course of this project. I would also like to thank my committee member Dr. Andrew Cooper for his insight and suggestions throughout the research process. I give special thanks to Dr. Rob Williams for providing me with the initial inspiration for my research idea and for helping me with the design of the project. Always willing to make time for me to answer questions and provide suggestions, Rob provided me with unwavering support and advice throughout the course of the project. He also introduced me to several statisticians, including Dr. Len Thomas and Dr. Eric Rexstad, who further guided me through the research process. I would also like to thank Dr. Sharon Hedley and Dr. Natalie Kelley for their advice and expertise with my distance sampling analysis in R and for helping me through the coding process. Without their help I would not have been able to complete my research analysis.

Thank you to my friends and co-workers at Cetus Research & Conservation Society for helping to collect the data used in this project and for their intellectual and emotional support through out the entire process.

Finally, thank you to Doug for going through this journey with me, for being my inspiration and for providing insight and strength along the way. And thanks to all my friends and family who have provided me with enduring and unbounded support during the completion of this project, as they have with any endeavour I have taken on through out my life. I am so grateful and feel blessed to know you all.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1: INTRODUCTION

Estimates of spatial and temporal distribution and abundance of cetaceans (whales, dolphins, and porpoises) are required to inform wildlife conservation and management decisions, to evaluate their effectiveness, and to assess the potential effects of anthropogenic disturbance on such populations and their habitat (Evans and Hammond 2004; Barlow et al. 2007; MacLeod et al. 2008). Specifically, distribution and abundance data are essential components of fisheries management, conservation, and habitat use by humans. For instance, the way in which fisheries are managed can affect bycatch of small cetaceans (Hall et al. 2002), chance of entanglement in fishing gear (BCCSN 2006), and availability of prey for cetaceans (Larkin 1996). As well, many cetaceans have been struck and in some cases killed by fishing or non-fishing vessels (Williams and O'Hara 2010). Conservation efforts associated with cetaceans require distribution and abundance data to determine the species' conservation status, areas of important habitat (Williams and Thomas, 2007), and where and when cetaceans would be most sensitive to anthropogenic disturbances such as seismic surveys that are conducted by resource extraction industries (Williams and Thomas. 2007). Moreover, information on time trends in abundance (e.g., "Abundance now is 50% of what it was X years ago") may identify a decline in a given species' occupation of the study area, emphasizing the need for further research or precautionary management actions.

However, for the 25 species of cetaceans documented to occur off the coast of British Columbia (B.C.), Canada, there remains little information about their distribution and abundance (Ford et al. 2010). This situation is surprising, given that many of the species currently identified as occupying B.C.'s coastal waters, including those targeted by this research, are listed as at risk by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC) (COSEWIC 2011). COSEWIC evaluates a species status as "extinct", "extirpated", "endangered", "threatened", "special concern", "not at risk", or "data deficient" (COSEWIC 2011). COSEWIC species assessments are then forwarded to the federal Minister of Environment who determines whether they should be legally protected under Canada's *Species at Risk Act* (SARA 2003). Species recognized as at risk by COSEWIC are therefore not automatically listed under SARA. Species found off the B.C. coast that are identified as at risk by COSEWIC and that have been legally listed under SARA include the:

- harbour porpoise (*Phocoena phocoena*) (special concern);
- fin whale (*Balaenoptera physalus)* (threatened);
- humpback whale (*Megaptera novaeanglie*) (special concern);
- grey whale (*Eschrichtius robustus*) (special concern);
- transient killer whale (*Orcinus orca)* (threatened);
- offshore killer whale (*Orcinus orca*) (threatened); and the
- southern (endangered) and northern (threatened) resident killer whale (*Orcinus orca*) (COSEWIC, 2011; SARA, 2011).

Other common B.C. species *not* identified as at risk by COSEWIC include the Dall's porpoise (*Phocoenoides dalli*), minke whale (*Balaenoptera acutorostrata*), and Pacific white-sided dolphin (*Lagenorhychus obliquidens*) (COSEWIC 2011; SARA 2011). While these species are not considered at risk under SARA or by COSEWIC, there remains little information on their population trends, distribution, and/or abundance within the coastal waters of B.C. (Ashe 2007).

Numerous anthropogenic activities and natural processes affect cetacean species in B.C. waters and may threaten or impede the recovery of species that are listed as at risk (DFO 2008). In some cases, many of these species are identified by COSEWIC as at risk due in part to their depletion as a result of commercial whaling or the live capture industry (Williams and Thomas 2007; DFO 2008; COSEWIC 2011). In addition, the reduction in the availability of prey to at-risk cetaceans through competition with fisheries, climate change, and reduced benthic and pelagic productivity in feeding grounds may also limit the recovery of species and/or cause declines in population abundance (Ford et al. 2005; Ford et al. 2009b; DFO 2010a; DFO 2010b). For example, Ford et al. (2009b) showed that trends in the survival patterns of resident killer whales are strongly related to fluctuations in the abundance of Chinook salmon, such that prey limitation may therefore be an important factor in recent declines of resident killer whale populations.

Several processes that also negatively affect cetaceans and the recovery of at-risk species are related to the presence of vessels, e.g., strikes by vessels (Williams and O'Hara 2010; Jensen and Silber 2003), marine vessel exhaust

(Lachmuth et al. 2011), vessel noise (especially from shipping), and vessel disturbance (Erbe 2002; Williams et al. 2002; Williams et al. 2006a; Lusseau 2006; Noren et al. 2009; Holt et al. 2009; DFO 2008). Boat noise can not only impede the ability of cetaceans to communicate between individuals, which may limit the ability of an individual to find a mate (Ford et al. 2000), but it also has been linked to the disruption of feeding activity (Williams et al. 2006a) by masking echolocation or passive listening for prey (Erbe 2002; Williams et al. 2002). Vessel presence at close proximity to cetaceans may also disrupt feeding activity by altering an animal's behavioural state (e.g., causing it to change from foraging to travelling) (Williams et al. 2002). Other detrimental human activities include the production of toxic fat-soluble contaminants (Ross 2000; DFO 2008), intense anthropogenic noise (i.e., navy sonar, and seismic testing) (Williams and Thomas 2007), incidental by-catch of small cetaceans in commercial gillnet fisheries (Hall et al. 2002), entanglement in active and/or ghost fishing gear (Ford et al. 2009a; DFO 2008; Guenther et al. 1995), and entanglement in and ingestion of marine debris (Williams et al. 2011a). Furthermore, the potential for future oil and gas exploration or extraction and increased shipping traffic off the coast of B.C. (Williams and O’Hara 2010; Molnar and Koshure 2009) will only exacerbate the impact of some of these anthropogenic threats. The aggregate potential effect of this wide range of human activities reinforces the importance of producing baseline estimates of distribution and abundance for the cetaceans found in B.C.’s coastal waters before such human activities intensify (Williams & Thomas 2007). These are some of the reasons for conducting this study.

In some cases, corrective actions against some of the human-induced processes identified above have led to the recovery of cetacean populations. Whaling for humpback and grey whales in the northern Pacific during the 19th and 20th century drastically reduced global humpback and grey whale populations (Ford et al. 2009a), but since the cessation of whaling, both the eastern population of grey whales and humpback whales in the Pacific are believed to have recovered to near pre-exploitation levels (DFO 2010a; Ford et al. 2009a). Most recently, COSEWIC down-listed the eastern Pacific population of humpback whales from threatened to special concern (COSEWIC 2011) because of evidence of their slight recovery in population size. Similarly, international agreements (e.g., United Nations driftnet ban) have helped to reduce the amount of bycatch associated with global fisheries (Hall et al. 2000). Other methods aimed at the reduction of bycatch in commercial fisheries include total bycatch limits, the promotion of sustainable seafood consumer-education programs (e.g., Dolphin Safe (International Dolphin Safe Monitoring Program 2011), Sea choice (Sea choice 2011) and Ocean wise (Ocean Wise 2011), and technological changes in equipment and deployment methods aimed at reducing bycatch (Hall et al. 2000).

It is difficult to determine the extent and magnitude of effects of human activities on cetaceans. Not only are the anthropogenically based mechanisms poorly known, but cetaceans are also wide-ranging, highly mobile, and distributed relatively sparsely over vast areas, thus making systematic studies of their habitat use and abundance expensive (Redfern et al. 2006) and logistically challenging. Efforts to meet this challenge have been made by estimating cetacean distribution and abundance on the coast of B.C. These efforts include, but are not restricted to, five types of data collection. (1) Opportunistic sightings networks, like the B.C. Cetacean Sightings Network (BCCSN), collect presence-only data where observers only report positive sightings and no information is collected on absence of sightings. (2) Coast-wide statistically designed systematic line transect surveys (Williams et al. 2007) gather data in a more rigorous manner, but are constrained by small sample sizes due to difficulties associated with re-sampling large surveyed areas. (3) Species-specific (as opposed to coast-wide) statistically designed systematic line transect surveys (Calambokidis et al. 1997; Chandler and Calambokidis 2003; Hall 2004) are constrained by species type and to small geographic areas such as the waters off southern Vancouver Island. (4) Studies carried out using a ferry as a "platform of opportunity" (Keple 2002) are constrained by their coverage of the study area, providing estimates which only include the waters surrounding the ferry route off southern Vancouver Island. (5) Scientists at the Department of Fisheries and Oceans Canada (DFO) and other independent scientists conduct ongoing mark-recapture studies and/or photo identification work (Morton 2000), which is limited to those species that can be individually identified (e.g., the saddle patch and dorsal fin of killer whales). In addition to mark-recapture studies, DFO also carries out ship-based line transect surveys that are not statistically designed and that tend to sample areas of known cetacean presence such that sampling does not always occur throughout B.C.'s Pacific waters in all seasons. The resulting abundance estimates only reflect the encounter rate along the trackline for each species sighted (Ford et al. 2009a, 2010).

My research focused on the BCCSN data. As one means of collecting information on cetacean distribution and abundance in B.C., the Vancouver Aquarium, in partnership with DFO, founded the BCCSN in 1999 (BCCSN 2009). The aim of the network is to establish a repository of opportunistic sightings of cetaceans off the coast of B.C. collected from a group of voluntary observers, without any of the logistical or monetary constraints of conducting dedicated line transect surveys (BCCSN 2006). The BCCSN has maintained a network of over 3,000 observers across B.C. for over 10 years (BCCSN 2011). The network is composed primarily of professional mariners and coastal citizens who report their sightings of cetaceans and sea turtles (BCCSN 2011). To date the network has catalogued over 45,000 incidental sightings of 18 species of whales, dolphins, porpoises, and three species of sea turtle (BCCSN 2009).

Incidental sightings collected by the BCCSN provide seasonal presence-only data about a given species within a given area. Such data can be valuable in many ways. Incidental sightings are useful in identifying temporal or spatial areas of high use that can be targeted by future research for more refined survey methods (Elith et al. 2006). Furthermore, incidental sightings may reveal gross distributional changes over time, and may be the primary source of information for rare species, which do not occur in predictable locations (Evans and Hammond 2004).

However, incidental or opportunistic sightings, such as those collected by the BCCSN, provide no quantitative measure of trends in absolute population abundance (Evans and Hammond, 2004). This limitation arises because most observers do not report their effort related to sightings of cetaceans, or the absence of sightings despite time spent searching. Here effort pertains to a unit of time spent searching for cetaceans by a team of observers, e.g., time on the water or distance travelled within the survey region (Buckland et al. 2001). To conceptualize searching effort, a useful analogy exists in fisheries, where abundance (N) is often related to catch per unit effort (U) through the equation

$$U = qN, \tag{1}$$

where q refers to the catchability of the fish species (Jennings et al. 2001; Hilborn and Waters 1992). For abundance estimates of cetaceans, this relationship can be interpreted as sightings per unit effort (U) being proportional to abundance of the species (N) multiplied by its sightability, or probability of being sighted per unit of time searching (q). Sightability can be affected by the sea state, weather, visibility, group size, animal behaviour (i.e., an animal breaching will be easier to sight), and diving behaviour/time spent on the surface (i.e., animals that dive longer and deeper are on the surface less often) (Williams et al. 2006b; Buckland et al. 2001). Without effort or sightability information, one cannot estimate the expected proportions of animals present in the survey region that are actually detected (detection probability) and therefore one cannot subsequently estimate the actual abundance either (Buckland et al. 2001; Williams et al. 2001; Royle and Nichols 2003). Without that information, only relative abundance estimates would be feasible.

As a result, I chose to work with data collected by one of the BCCSN observers, the whale-watch monitoring group called Straitwatch. I was able to modify their data collection methods such that they started to collect information on effort and sightability, thus allowing for the estimation of distribution and absolute abundance of cetaceans within their program ranges (Figure 1) (Williams et al. 2006b). Operated by the Cetus Research and Conservation Society (CETUS), Straitwatch monitors the boating activities around marine mammals (especially killer whales and humpback whales) and provides boaters with information on local marine species and marine mammal viewing guidelines (CETUS 2009). The Straitwatch program operates within 3 survey regions off Vancouver Island (Figure 1). The Straitwatch North (SWN) (Figure 2) and Straitwatch South (SWS) (Figure 3) programs conduct daily patrols throughout their range between June and mid-September, while the Straitwatch West Coast (SWWC) (Figure 4) program operates for a period of 1 to 2 weeks each summer. Each program collects opportunistic sightings for the BCCSN (CETUS 2009). Straitwatch collects effort and sightability data along with GPS tracks, which record the distance travelled and area covered by observers on the vessel. Straitwatch also records information on sea state, weather, group size, and animal behaviour.

The primary responsibility of the Straitwatch boat in any given region is to monitor the vessel activity surrounding a group of whales, typically killer whales, which are the target of whale-watching activities within the program's range. Thus, researchers collecting sightings data from this platform are not able to control the route of the vessel, because they are essentially following whale-watching vessels, hence the survey is said to be conducted from a "platform of opportunity" (Williams et al. 2006b; Evans and Hammond 2004). Many studies examining spatio-temporal patterns of abundance of cetaceans have used ferries (Keple 2002), whale-watching vessels (Hall 2004; Leaper et al. 1997), fishing vessels (Buckland et al. 1992) and cruise ships (Williams et al. 2006b) as a "platform of opportunity" from which to collect data (Brereton et al. 2000; Marques 2001).

1.1 Research objective

Given the numerous needs described above for better data on distribution and abundance of cetaceans, my research objective is to use data collected from Straitwatch as a "platform of opportunity" to establish preliminary estimates of absolute abundance and spatial distribution for 6 species found within 3 survey regions around Vancouver Island (Figure 1). Specifically, I aim to estimate these quantities for Dall's porpoise, harbour porpoise, Pacific white-sided dolphin, humpback whale, minke whale, and grey whale.

I used a spatial distance sampling model to help achieve these objectives. The formal method called distance sampling is one of the most frequently used methods for estimating density and/or abundance of biological populations (Thomas et al. 2002, Marques 2001), and is usually applied through line transect or point transect surveys. Distance sampling accounts for the detectability (i.e., sightability) of a species by generating a detection function, which is used to help model abundance estimates. Except in rare cases where one can be certain that all animals are detected within the survey region, methods for estimating absolute abundance must account for detectability, as is possible with distance sampling (Marques 2001). Moreover, methods that account for detectability are especially relevant for cetacean populations that are distributed over large areas with low-to-medium density (Marques 2001) and which spend part of their time underwater, making them not always detectable (Hall 2004; Buckland et al. 2001). Results of my analyses may be useful for (1) determining habitat use, (2) identifying and protecting critical habitats, (3) creating marine protected areas, and (4) determining areas within which there is considerable potential for anthropogenic disturbance such as entanglement or bycatch risk from fishing gear. Furthermore, detection functions generated for each species of cetacean during this study may be applicable for future design-based and/or model-based distance analyses conducted using the Straitwatch opportunistic platform.

1.2 Current estimates of distribution and abundance within the Straitwatch program ranges

To set the context within which my research was done, this section provides background data on current estimates of distribution and abundance available for the 6 cetacean species found within the Straitwatch program ranges.

1.2.1 Dall’s porpoise

Dall’s porpoise are considered to be one of the most abundant cetacean species in the North Pacific and range from northern Baja California to the Bering Sea (Ford et al. 2010). The latest U.S. Pacific marine mammal stock assessment estimates that the Pacific population size of Dall’s porpoise is approximately 39,700 (Carretta et al. 2010). Between 2002 and 2008, DFO carried out non-random ship-based line transect surveys in B.C.’s offshore and inshore waters (Table 1) from which they generated an encounter rate along the trackline for each species observed. The only species encountered more frequently than Dall’s porpoise were humpback whales (Ford et al. 2010). In 2004 and 2005, Williams and Thomas (2007) carried out design-based systematic line transect surveys in the inshore coastal waters of the Inside Passage between the B.C./Washington (WA) and the B.C./Alaska borders. They generated a coast-wide estimate of Dall’s porpoise abundance of 4,910, as well as estimates for the Johnstone Strait stratum (Table 1). The SWN survey region makes up the northwestern portion of the Johnstone Strait stratum used by Williams and Thomas (2007). Both aerial surveys carried out over the inland waters of WA and B.C. (Calambokidis et al. 1997) and a Master’s-degree study using the ferry route between Tsawwassen and Swartz Bay B.C. as a “platform of opportunity” (Keple 2002) have generated estimates of abundance within the SWS survey region (Table 1). No estimates of Dall’s porpoise abundance are currently available within the SWWC survey region.

1.2.2 Harbour porpoise

Harbour porpoises typically inhabit the cool temperate waters of the northern hemisphere and are found in the NE Pacific range from Point Conception to the Bering Sea (Ford et al. 2010). Both COSEWIC and SARA list harbour porpoise as special concern because of their sensitivity to human activities and their vulnerability to entanglement and/or bycatch in fisheries (Ford et al. 2010). During DFO's ship-based surveys, harbour porpoise were sighted most frequently around Vancouver Island, Haida Gwaii, in mainland inlets, and other shallow inshore waters (Table 1) (Ford et al. 2010).

Williams and Thomas (2007) estimated a coast-wide abundance of 9,120 harbour porpoise (Table 1), but were not able to make any estimates of harbour porpoise abundance for the Johnstone Strait stratum due to a limited number of sightings there. No estimates of harbour porpoise abundance are currently available within the SWN survey region. Conversely, aerial surveys occurring within the SWS survey region in 1996 and in 2002/03 have predicted harbour porpoise abundance for the inland waters of WA and B.C. (Table 1) (Calambokidis et al. 1997; Carretta et al. 2010). In addition, two Master's-student theses have produced estimates of abundance for harbour porpoise within portions of the SWS survey region; one of which was conducted from a "platform of opportunity" (Keple 2002), while the other involved a series of systematic line transect studies within the waters of Juan de Fuca and Haro Strait, B.C. (Hall 2004). No estimates of harbour porpoise abundance are currently available within the SWWC survey region.

1.2.3 Pacific white-sided dolphin

Pacific white-sided dolphin are considered to be one of the most widely distributed and abundant small cetaceans in the North Pacific (Ford et al. 2010). They are known to range in temperate regions from the Aleutian Islands to Baja California, and are typically found year-round in pelagic waters, and in the inshore waters of Alaska, Washington and B.C. (Ashe 2007; Ford et al. 2010). Heise (1997) noted an increase in abundance from the 1980s in the inshore waters of B.C., possibly due to a shoreward shift in distribution (Ford et al. 2010). The majority of Pacific white-sided dolphin sightings were made in offshore waters, in the waters around Haida Gwaii, and in coastal inlets ranging only as far south as 49°N on either side of Vancouver Island (Table 1) (Ford et al. 2010).

Several studies have attempted to estimate the distribution and abundance of Pacific white-sided dolphins within the waters surrounding the SWN survey region (Table 1). Williams and Thomas (2007) produced both coast-wide and Johnstone Strait stratum abundance estimates (25,900 and 1,344, respectively). In addition, a photo-identification study carried out in the Broughton Archipelago and surrounding waters identified 675 individuals between 1984 and 1998 (Morton 2000). Recent work using the data collected by Morton (2000) produced both minimum and maximum annual abundance estimates using a mark-recapture model (Ashe 2007). No estimates of Pacific white-sided dolphin abundance are currently available within the SWS or SWWC survey regions.

1.2.4 Humpback whale

The abundance of humpback whales in the North Pacific has been steadily increasing from an estimate of 1,600 individuals by the end of commercial whaling in 1966 (DFO 2010a). The total pre-harvest estimates of humpback abundance in the eastern North Pacific was 15,000 individuals, however this may be considered inaccurate because an estimated 28,000 humpbacks were killed as a result of whaling between 1905 to 1965 (DFO 2010a). More recently, the study called “Structure of Populations, Levels of Abundance and Status of Humpback Whales in the North Pacific” (SPLASH), which was carried out between 2004 and 2006, estimated abundance of 18,000 to 21,000 animals (not including first-year calves) within the North Pacific at that time (Table 1) (COSEWIC 2011; Calambokidis et al. 2008). This study also produced estimates for B.C.’s coastal waters, with 3,000 to 5,100 animals estimated to occupy the waters of the Gulf of Alaska, southeast Alaska and northern B.C., and 200-400 individuals occupying the waters of Southern B.C. and northern WA (Calambokidis et al. 2008).

It is thought these 2 regions make up 2 distinct feeding groups within the B.C. humpback population (DFO, 2010a). The best estimate of humpback whale population size in B.C. waters produced by Ford et al. (2009) in 2006 suggested that 2,145 whales are seasonally present in B.C. waters, with a rate of increase estimated from photo-identification studies of 4% per year (COSEWIC 2011). As a result of the estimated recovery in population size of this species, humpback whales have been recently down-listed by COSEWIC from “threatened” to “special concern” (COSEWIC 2011). A review process is currently underway in Canada regarding their species status under SARA.

Humpback whales range along the entire length of the west coast of B.C. from the WA to Alaska borders (Ford et al. 2010), with the greatest numbers found between May and October (DFO 2010a). During the DFO surveys, humpback whales were encountered more frequently than any other cetacean species (Table 1), with observed groups sizes of 1 to 35 and the majority of humpback’s sighted alone or in pairs (Ford et al. 2010).They have been known to inhabit both inshore waters, coastal inlets, and offshore waters, and appear to be distributed in aggregations based on the availability and location of their prey (DFO 2010a). Individual animals have been documented to show strong site fidelity, with the areas of greatest humpback whale densities identified as the waters around the Queen Charlotte Islands, channels and inlets on the North mainland coast, and off the North and Southwest coasts of Vancouver Island (DFO 2010a).

Williams and Thomas (2007) estimated a coast-wide abundance of 1,310 humpback whales in B.C, however, no specific estimates of abundance are available for the SWN or SWS survey regions. Photo identification work carried out within the SWN survey region has identified the number of individual humpback whales present annually in 2008 and 2009 (Table 1). In Clayoquot Sound, located within the western portion of the SWWC survey region, similar photo identification work has recorded 358 individuals between 1995 and 2009

(Pacific Wildlife Foundation 2010) with 65 individual humpbacks identified during the summer months (May to September) of 2009.

1.2.5 Minke whale

Few estimates relating to the abundance of minke whales in the North Pacific exist to date. Only one species of minke, the common minke, is found year-round within the water's of B.C. (Ford et al. 2010). The common minke occurs both in offshore waters and in shallow coastal waters, and it feeds mostly on small schooling fish such as herring and sandlance, as well as euphausiids and other crustaceans (Ford et al. 2010). The common minke is listed as "not at risk" by COSEWIC despite its low abundance level in B.C.'s coastal waters (Ford et al. 2010).

Coast-wide estimates of encounter rate from the DFO surveys (Ford et al. 2010) and abundance (Williams and Thomas 2007) are available (Table 1). No specific estimates of minke whale abundance exist in any of the Straitwatch survey regions. Photo identification work carried out within the SWN survey region has identified a total of 6 and 9 individual minke whales in 2008 and 2009, respectively (Jared Towers, Marine Education and Research Society (MERS), jrtowers@gmail.com, June 21, 2011, *pers. comm.*).

1.2.6 Grey whale

Estimates by COSEWIC suggest that 18,000 grey whales, comprising the eastern Pacific population, inhabit the North Pacific, with the number in the low hundreds for the B.C. coast (COSEWIC 2010; DFO 2010b). Grey whales in the

North Pacific mostly spend their time feeding in the high latitudes, while several hundred whales (known as the Pacific coast feeding aggregation) feed in the nearshore waters between California and Alaska (Ford et al. 2010). Summer resident grey whales in B.C. exhibit high site fidelity, which is primarily driven by prey abundance (Ford et al. 2010). In B.C., summer resident grey whale densities are highest off the west coast of Vancouver Island, the North coast of Vancouver Island between Cape Caution and Cape Sutil, and in mainland inlets from Shelter Bay to Cape Caution (Ford et al. 2010; COSEWIC 2010). Recent work by the Pacific Whale Foundation suggests that these summer residents are genetically distinct from the eastern Pacific population, which may have implications for the management of the summer resident grey whales. Summer resident grey whales have also been sighted within the SWS survey region around Boundary Bay, Haro Strait and Georgia Strait (COSEWIC 2010). Coast-wide estimates of encounter rate from the DFO surveys (Ford et al. 2010) are available (Table 1); however, no abundance estimates have been made for the SWS and SWN survey regions. In 2008 and 2009, SWN made no sightings of grey whales, however, during the summer months of 2010 and 2011 one grey whale was sighted in both years within the SWN survey region. It is possible that this animal is showing some site fidelity to the Johnstone Strait area. Photo identification studies have identified 47 individuals (between 2006 and 2008) in the SWWC survey region where there is a greater occurrence of summer resident grey whale aggregations during the summer months (Pacific Wildlife Foundation 2009).

The abundance estimates generated through this research project complement the current information presented in Table 1 and section(s) 1.2.1-1.2.6 on the distribution and abundance of six commonly occurring cetacean species in B.C. For many of these species, current information on abundance is restricted to relative abundance, individual counts or encounter rates, and estimates of distribution are limited. Additionally, many of the studies presented in Table 1 are limited in scope both spatially and/or temporally and in some cases only produce estimates for years prior to 2000. While my study is also limited in scope both spatially (3 survey regions) and temporally (May to September), it helps to complement previous studies by providing more recent assessments of abundance for those species for which past estimates currently exist, while also providing estimates of species distribution, which do not exist in many cases. Moreover, this study provides abundance and distribution estimates for areas and for species for which none currently exist (e.g., for many of the small cetacean species and for those cetaceans occurring in the SWWC survey region).

2: METHODS

2.1 Distance sampling as a means to estimate distribution and abundance of cetaceans within the Straitwatch program ranges

I used distance sampling to obtain my estimates of abundance. Distance sampling uses the distance to an animal detected as a means to estimate the probability of detecting an animal ($\hat{p}$) as a function of perpendicular distance from the trackline. In order to account for the assumption that not all animals within the survey area are detected, distance sampling estimates the number of animals within a defined survey area (N_A) as the number of animals detected (*n)* divided by the estimated probability of detecting an animal ($\hat{p}$) (Marques 2001).

$$\hat{N}_A = \frac{n}{\hat{p}} \tag{2}$$

Three key assumptions of distance sampling are that (i) objects on the line transect are detected with certainty (i.e., the value of the detection function (g(x)) evaluated at zero distance (x=0) is 1), (ii) there is no responsive movement (attraction/avoidance) of the animal in relation to the boat before detection, which can be a problem with dolphins and porpoises who engage in bow riding or avoiding vessels, and (iii) measurements are exact for distance to the animal, radial angle from the transect line, species type, and group size (Thomas et al. 2010). Double observer or double platform surveys, in which observers search independently of each other or are located on separate platforms, may be conducted in order to reduce the effect of violating assumption (i) such that the actual value of g(x) evaluated along the trackline (i.e., x=0) can be estimated. This can especially be important with marine mammal surveys where animals go underwater and where the assumption that all animals along the trackline are detected can be violated (Thomas et al. 2010). Furthermore, there are two approaches that can be taken to minimize the level of bias introduced by assumption (iii). These include additional observer training and/or the use of experiments to develop correction factors to remove bias (Williams et al. 2007). Additional observer training in distance estimation can include training in adequate measurement accuracy and appropriate species identification (Thomas et al. 2010). Not only can Straitwatch staff be considered as trained observers in terms of species identification, but they also train or recalibrate themselves daily on distance estimation by comparing visual distance estimates against those obtained through the use of a laser range finder or radar. For my research, I conducted experiments that compared observers' estimates of distance to sighted cetaceans to actual distances to sighted cetaceans measured using laser range finders and/or radar. I conducted these experiments during both of my field seasons in order to correct for measurement error associated with assumption (iii) (see section 2.3 – Distance estimation experiments).

Another key assumption of conventional distance sampling is that of equal coverage probability (Buckland et al. 2001), which means that all points in the study area have equal probability of being sampled. This can be achieved by placing transects at random with respect to the distribution of the animal(s) being studied (Buckland et al. 2001; Williams et al. 2006b; Thomas et al. 2007). This assumption is required to obtain unbiased estimates of absolute abundance. Line-transect surveys, which are used most commonly to predict abundance and distribution information for both marine and terrestrial species, are usually designed *a priori* to provide for equal coverage probability throughout the survey region (Buckland et al. 2001).

However, the assumption of equal coverage probability is often violated (Buckland et al. 1992) by platforms of opportunity such as the Straitwatch program where survey effort and coverage probability are not distributed randomly throughout the study area. As a result, “design-based” distance sampling methods that statistically rely on all points having equal probability of coverage are inappropriate for most data from “platforms of opportunity”. In such cases, other “model-based” distance sampling methods are required (Williams et al. 2006b; Hedley et al. 1999).

One model-based approach, which can be applied to data from a “platform of opportunity” like the Straitwatch vessels, follows distance sampling methods that use the count method to spatially model distribution and abundance of a species within the study area (Hedley and Buckland 2004). The count method follows a two-stage approach that first estimates the probability of detection of a species through the generation of a detection function, and then uses the estimated probability of detection for each species (from the detection function) to help model the number of cetaceans in each “segment” of the trackline as a function of spatially-referenced environmental covariates; that last step is done with a Generalized-Additive Model (GAM) (Williams et al. 2011b; Thomas et al. 2010). Typically, such covariates are depth, latitude, longitude, temperature, or habitat type (Thomas et al. 2010). That GAM relationship is then used to predict density and abundance throughout the larger study area by applying the model to a prediction grid within the survey region (Hastie and Tibshirani 1990; Thomas et al. 2010; Williams et al. 2006b; Buckland et al. 2004; Hedley and Buckland 2004). GAMs can be applied when the response variable is discrete (e.g., count data), and are more advantageous than generalized linear models (GLMs) due to their flexibility when modelling spatial variability (Hedley, 2000) and when examining non-linear cetacean-habitat relationships (Redfern et al. 2006).

It is important to note that reliable estimates of absolute abundance can be obtained using platforms of opportunity through the application of model-based techniques (e.g., count-method) (Buckland et al. 2000; Williams et al. 2006b). Estimate of relative abundance are commonly produced from data collected on platforms of opportunity (e.g., encounter rate, bycatch, presence/absence data), however the use of spatial distance sampling models can be used to provide an estimate of absolute abundance by numerically integrating under the fitted density surface (Buckland et al. 2004; Williams et al. 2006b). Moreover, spatial models of abundance and distribution based on spatial covariates can provide more information to managers because they may offer some explanation as to what factors are important predictors of cetacean presence within a survey region (Buckland et al. 2004).

If the non-random transects associated with the travel pattern of the platform of opportunity provide good spatial coverage within the survey region (typically generated post hoc), then spatial models can often offer more reliable estimates of absolute abundance than conventional designed-based analyses (Buckland et al. 2004), providing that effort is not determined by the distribution of the species of interest. While the data collected opportunistically by the Straitwatch program follows a non-random survey design, because the movement of the vessel within the survey area is determined by the presence of killer whales within each individual survey region, it is not directly determined by the six species of interest presented in this study. However, there remains the possibility that the environmental variable determining killer whale presence in the survey region might overlap with that of the species examined in this study (e.g., prey source). In some limited cases, where Straitwatch was monitoring the activity of vessel around transient killer whales, then the species of interest might have avoided the area(s) in which the vessel was being operated.

2.2 Data Collection

The Straitwatch data collection teams collected opportunistic sightings data from the Straitwatch monitoring vessels daily between July 1st and September 15th, 2008 and 2009, throughout the three program ranges (Figure 1). The boundary of each of the SWS, SWN and SWWC program ranges defined the 3 survey regions located around Vancouver Island (Figure 2; Figure 3; Figure 4). Each data collection team consisted of a vessel operator and a data recorder. While traveling through the survey region(s), the team recorded its effort in the form of a track using a Global Position System (GPS). Tracks recorded the time, location, speed, and course throughout the study area. I will consider each day's set of tracks within each of the program's ranges as a single survey for the purpose of my analysis. During the 2008 and 2009 seasons, the SWN vessel recorded effort daily in the form of GPS tracks for a total of 101 days (Table 2 and Figure 5), while the SWS vessel recorded a total of 34 days (Figure 6) and the SWWC vessel recorded a total of 6 days (Figure 7).

In addition to collecting tracks within each survey region, the data collection teams also scanned for sightings from the trackline to 90 degrees to either side of the vessel. Whenever a sighting was made, the data recorder entered a GPS location and recorded the observer, species ID, time, latitude, longitude, sea state, weather, species group size, direction of travel, speed of the animal's travel (slow, medium, fast), and the animal's behaviour (traveling, resting, foraging etc.), as well as the distance (r) and angle (θ) to the observed animals (Figure 8). I calculated the perpendicular distance (x) from the trackline for each sighting using the formula $x = r \cdot \sin(\theta)$ (Thomas et al. 2010; Buckland et al. 2001).

The data collection teams used angle markings on the vessel's inflatable tubes to help estimate the angle to the animal. These markings are similar to those used on an angle board for transect surveys based on larger ships (Williams and Thomas, 2007). Observers also used binoculars, when appropriate, to confirm species type and group size, as well as a laser range finder and/or radar to estimate sightings distance relative to nearby points of land.

During the 2008 and 2009 surveys, each data collection team recorded sightings of harbour porpoise, Dall's porpoise, Pacific white-sided dolphins, minke whales, humpback whales and grey whales in all three survey regions (Table 3).

2.3 Distance estimation experiments

Straitwatch vessel operators practice daily their ability to visually estimate distance to objects on the water. In order to correct visual estimates of distance to animal for each sighting, I carried out distance estimation experiments with all Straitwatch vessel drivers (Williams et al. 2007). Experiments were carried out mid-season and on the same day for all observers in each survey region for 2008 and 2009, respectively. During these experiments, Straitwatch employees recorded their visual estimates of distance to 20 continuously visible targets, while a data recorder measured exact distances to each target using laser range finders and/or radar. Targets consisted of small floats/buoys, Styrofoam blocks, and barrels which best approximated the size of small cetaceans detected as sightings during surveys.

Methods to correct error in an individual observer's visual estimation of distance followed Williams et al. (2007). The first step was to fit a linear regression of the form $y=mx+b$ forced through the origin ($b=0$) to the paired data (y=estimated distance, x=actual distance) from the distance estimation experiments for each of the 7 observers. Results of each paired distance linear regression provided a correction factor *(cf),* which is the slope of the regression line forced through zero.

I used plots of the linear regression for each observer's paired distances to identify whether heteroscedasticity existed (that variance in the estimated distance increased at greater distances) (Williams et al. 2007). Any evidence of heteroscedasticity indicated that a linear regression was not appropriate and that the calculation of a correction factor required a log transformation of the data or an alternative model such as a GAM (Williams et al. 2007).

Paired comparisons of actual-to-estimated distance for each observer in this study suggested a heteroscedastic relationship (Figure 9a-b). As a result, the data were log transformed (log base 10) and plots of log(y) vs. log(x) for each observer instead revealed a linear relationship with constant variance across the range of values (i.e., no heteroscedasticity) (Figure 10a-b; Table 5). Such results are consistent with Williams et al. (2007), who found that a log transformation best describes the relationship between visual estimates and measured values.

The model I selected for each observer was therefore:

$$\text{Log (estimated distance)} = (cf) \times \text{Log (actual distance)} \qquad (3)$$

I corrected each observer’s estimates of distance to an animal(s) using their unique correction factor obtained from equation (3) and multiplying their estimated distance by $1/cf$.

2.4 Data analysis

I followed a five-stage approach to estimate the cetacean density and its associated variance in each survey region (Figure 11). (i) I first fit a global detection function $g(x)$ to the sightings data from all 3 survey regions for each species to estimate the probability of detection for each species. (ii) I then estimated the mean school size while taking into account school size bias for each species. (iii) I fit a statistical spatial model of cetacean abundance in each segment of trackline in each survey region as a function of geo-referenced environmental covariates. (iv) I then used the spatial model from (iii) to estimate the abundance of cetacean species throughout each survey region. (v) Finally, I produced estimates of variance in abundance for each cetacean species in each survey region (Buckland et al. 2001; Buckland et al. 2004; Williams and O'Hara 2010; Hedley and Buckland 2004; Williams et al. 2006b). I estimated the detection probability and the mean school size using the free software *Distance* 6.0 release 2 (Thomas et al. 2010), and generated density and abundance estimates using the free statistical software R 2.9.2 (R Development Core Team, 2009).

2.4.1 (i) Fitting a global detection function for each species

I modelled detection functions for each species as smooth functions of perpendicular distance from the trackline (Williams et al. 2006b), using conventional distance sampling (CDS) methods (Buckland et al. 2001; Buckland et al. 2004). I also explored the use of multiple covariate distance sampling (MCDS) for all species (Buckland et al. 2001; Buckland et al. 2004). MCDS uses other covariates (e.g., sea state or school size) in addition to the perpendicular distance to model the detection function. In all cases, the models fit using MCDS were not favoured in the formal model selection process based on their Akaike Information Criteria (AIC) value for species with greater than 40 observations per parameter (1 parameter for the half-normal model and 2 parameter's for the hazard-rate model) or their AIC_c for species with fewer than 40 observations per parameter and so MCDS was not applied to my analysis.

I pooled the sightings data from all three survey regions (SWN, SWS and SWWC) to create a global detection function for each species. The detection function g(*x*) is defined as the probability that an object at perpendicular distance *x* from the line is detected, such that $0 < x \leq w$, and where *w* refers to the truncation distance (Buckland et al. 2001, Thomas et al. 2002). Following conventional distance sampling methods, I assumed that the trackline detection probability (the probability of detecting an animal that is located directly on the trackline) was certain (i.e., g(0) = 1 or any animal located on the trackline was detected) (Buckland et al. 2001).

Included in the model-selection phase for the detection function is the choice of a suitable truncation distance (*w)* (Thomas et al. 2010). Sightings for which the perpendicular distance to the trackline is greater than *w* were not included in the estimation of the detection function or in the abundance estimates (Figure 12) (Thomas et al. 2002). Truncation was required because otherwise extra adjustment terms in the detection function model might be needed to fit a long tail to that detection function (Thomas et al. 2010). However, adding terms to the detection function model that was used to fit data points which are numerically distant from the rest of the data (outliers) can reduce the precision of the fit of the model for little gain because typically these outliers contribute little to the abundance estimates generated from the detection function (Buckland et al. 2001). The surveyed area (*a*) within which *n* animals are detected can therefore be expressed as a = 2wL, where L refers to the length of the segment(s) (Figure 12) (Thomas et al. 2010).

If $\hat{p}$ represents the estimated mean probability of detection within the survey area, then the estimated animal density $\hat{D}$ is (Buckland et al. 2001):

$$\hat{D} = \frac{n}{2wL\hat{p}} \tag{4}$$

Alternatively, the estimated animal density can be expressed as (Buckland et al. 2001; Williams et al. 2006b):

$$\hat{D} = \frac{n \cdot \hat{f}(0)}{2L}, \tag{5}$$

Where $\hat{f}(0)$ represents the estimated probability density function of perpendicular distances of detected objects $f(x)$ evaluated at zero perpendicular distance. The estimated mean probability of detection ($\hat{p}$) can be calculated from the value of $g(x)/f(x) \bullet w$ evaluated along the trackline or where x = 0, such that $g(x)$ = 1 and $\hat{p} = 1/\hat{f}(0) \bullet w$.

Candidate forms for the detection function were the half-normal and hazard-rate models (Buckland et al. 2001; Williams et al. 2006b; Williams and O'Hara 2010). The half-normal model is:

$$g(x) = \left(-x^2/2\sigma^2\right), \tag{6}$$

where x refers to the perpendicular distance from the trackline and σ is a scale parameter, which defines the rate at which g(x) decreases for increasing distance from the transect line (Buckland et al. 1993). The hazard-rate model also contains a shape parameter (b), which defines whether the shape of the detection function has a “shoulder” near the transect line (Buckland et al. 2001). A shoulder is present when the detection of a species remains nearly certain at small distances from the line (e.g., g(x) = 1 from a distance of 0 to 100 meters from the line):

$$g(x) = 1 - exp\left(-(x/\sigma)^{-b}\right) \tag{7}$$

A detection function fit of the half-normal or hazard-rate model can be improved by removing or truncating 5-10% of the most distant sightings (Buckland et al. 2001). Model selection between models with the same truncation value was based on minimizing the AIC value for species with greater than 40 observations per parameter or minimizing the AIC_c value for species with fewer than 40 observations per parameter. I further evaluated model fit using diagnostic plots and goodness-of-fit statistics (e.g., qq plots, Cramer Von-mises test, Kolmogorov-Smirnov test) (Williams and Thomas, 2007). Selecting for a model with minimal truncation was preferred because spatial modelling performs best when the number of sightings is high (Hedley, 1999).

When the half-normal and hazard-rate model were both found plausible during the model selection phase (ΔAIC ≤ 2 or ΔAIC_c ≤ 2 depending on the number of observations of each species), then I selected the model with the minimum AIC (or AIC_c value when there were fewer than 40 observations per parameter). If the hazard-rate model was favoured, then I considered whether the selection of the hazard-rate model was biologically appropriate. The hazard-rate model fits a shoulder to the sightings data such that the probability of detecting an animal is the same from zero to some distance from the trackline

(Buckland et al. 2001, 2004). This is an especially important consideration for some of the harder-to-detect species (e.g., harbour porpoise) for whom it may not make biological sense to fit a shoulder to the model such that the detection probability for that species would remain the same between some distance from the trackline (e.g., 100 meters) and zero distance (along the trackline).

2.4.2 (ii) Estimating mean school size and mean school size bias for each species

In this second (ii) step of the analysis (Figure 11), I employed the default method in the *Distance* software to determine whether school-size (*s*) bias was present in my data and to obtain an unbiased estimate of mean school size for each species (Thomas et al. 2010). The probability of detection is often a function of the school size. School-size bias refers to the tendency for larger schools to be more detectable than smaller schools at the same distance (Buckland et al. 2001). As a result, larger schools are considered easier to detect and so may be over-represented in the sample (Thomas et al. 2002).

In order to determine if school-size bias is present, *Distance* fits a least-squares regression of ln (*s*) (natural log) on the estimated probability of detection at distance x and then predicts ln (s) at a distance of zero (the y intercept), where the detection is assumed to be certain (i.e., g(x)=1) (Thomas et al. 2002, Thomas et al. 2010). *Distance* uses a log transformation of the school size in order to reduce the influence of highly variable estimates of school size by the observer(s) (Buckland et al. 2001). If the point estimate of the slope of the regression line is greater than zero (a positive slope), then school-size bias is present (Buckland et al. 2001), and *Distance* predicts the expected school size *E(s)* by back-transforming the predicted ln (s) at x=0 using a bias adjustment generated by Buckland et al. (2001). The slope of the regression line and the value of *E(s)* are generated as an output from the calculation of the detection function for each species in *Distance*. Where school size bias is observed for a given species, the *E(s)* value predicted by *Distance* for that species is passed to the model used to predict cetacean abundance in each segment of trackline (section 2.4.3 (iii)). If the point estimate of the slope of the regression line is less than or equal to zero (a non-positive slope), then the observed mean school size for the cetacean species is used (Thomas et al. 2010; Williams et al. 2006b).

2.4.3 (iii) Estimating cetacean abundance in each segment as a function of spatial covariates for each survey region

Here, step (iii) created an abundance estimate for each species (by survey region) in each segment of trackline as a function of spatial covariates. I modelled effort and sightings data from each survey region using the count method (Hedley et al. 1999; Hedley and Buckland 2004; Williams et al. 2006b; Williams et al. 2011b), which has been included in the Density Surface Modelling (DSM) component of the *Distance* software (Thomas et al. 2010). The count method is a two-part statistical technique that models animal counts within each segment as a function of spatial covariates (Williams et al. 2011b). The first part of the count method involves estimating the probability of detection of each sighting for each species (Williams et al. 2011b) by applying the detection function calculated in section 2.4.1. Second, the count method uses the estimated probability of detection for each species to help model the number of cetaceans in each segment of the trackline as a function of spatial covariates with a Generalized Additive Spatial Model (GAM) (Williams et al. 2011b; Thomas et al. 2010). The length of each segment is chosen such that the spatial location and environmental conditions attributed to each segment are homogeneous along the length of the segment and is also driven by the spatial resolution of the environmental data (Hedley and Buckland, 2004).

For each survey region, I divided each transect into segments approximately 10 nautical miles (nm) or 18.52 kilometers (km) in length, with some of the end segments for each trackline being shorter. Originally, I selected a segment length of 2 nm (or 3.704 km), however, because of the high number of segments containing zero sightings, I increased the segment length to 10 nm to help reduce the amount of zero inflation in the model. For each segment, I calculated the location of the midpoint (MidLat, MidLon) along with the value of each spatial covariate incorporated into the GAM using Esri ® ArcView 3.3 and Esri ® ArcMap 9.2. The spatial covariates explored in the model include depth, summer salinity, summer sea surface temperature (SST), and tidal speed. I obtained the depth data through Natural Resources Canada (NRC) (Robert Kung, NRC, Robert.Kung@nrcan-rncan.gc.ca, September 2, 2009, *pers.comm.*). The depth data consist of a 75 m resolution grid in GIS image data source format. Additionally, I obtained environmental data from Fisheries and Oceans Canada (DFO) (Mike Foreman, DFO, Mike.Foreman@dfo-mpo.gc.ca, October 2, 2008, *pers.comm.*) which consist of a model grid with salinity, SST, and tidal speed values averaged from 1950 to 2007 for the summer season (May to October).

The spatial model, or GAM frame work, used by the count method with a logarithmic link function is:

$$E[n_i] = exp\{log(2l_i w \cdot \hat{p}) + f(Midlat_i, Midlon_i, depth_i, salinity_i, SST_i, tidalspeed_i)\} \quad (8)$$

where $E[n_i]$ is the expected number of cetaceans in the i^{th} *segment*; l_i is the length of the segment *i; w* is the truncation distance; $\hat{p}$ is the estimated probability of detection of a cetacean; $Midlat_i$ and $Midlon_i$ denote the mid-point of the i^{th} segment; $depth_i$, $salinity_i$, SST_i *and* $tidalspeed_i$ denote the value of spatial covariates at the mid-point of the i^{th} segment; and *f represents* a smooth function of the spatial covariates *(Midlat, Midlon, depth, salinity, SST, and tidalspeed)* considered for inclusion in the model.

The response variable for the GAM is the estimated abundance of a species detected and can be defined in this case as (Thomas et al. 2009):

$$\hat{N} = \hat{E}(s)\sum_{i=1}^{n}\frac{1}{\hat{P}_i} \quad (9)$$

where we multiply the estimated school abundance by the mean school size, or where school-size bias is detected, then the estimated expected school size ($\hat{E}(s)$) for a given species. The estimated school abundance can be defined as:

$$\hat{N} = \sum_{i=1}^{n}\frac{1}{\hat{P}_i} \quad (10)$$

where $\hat{P}_j$ is the estimated inclusion probability for animal *i* and *n* is the number of observations. The inclusion probability ($\hat{P}_j$) has two components: first that it falls within the effective area searched in each segment and second an estimate of its probability of detection based on its distance from the trackline.

The first term included in the spatial model $(log(2l_i w \cdot \hat{p}))$ is an offset term that describes the effective area searched in each segment (Thomas et al. 2010). That effective area is defined as the actual area of each segment ($2Lw$) multiplied by the estimated mean probability of detection ($\hat{p}$) of a cetacean species (calculated in section 2.4.1) in each segment (Thomas et al. 2010). The value of $\hat{p}$ was held constant and was not allowed to vary because of segment location or based on other properties such as behaviour of the animals or size of group. As the spatial model does not model density directly, the offset term reduces the response variable to n_i, which the count method can model as count data (Hedley et al. 1999; Williams et al. 2006b). The expected number of cetaceans in each segment was modelled as a function of the spatial covariates included in the selected GAM.

Many of the segments in my survey regions had zero sightings for each individual cetacean species, resulting in a zero-inflated data set. In addition to choosing a larger segment length, zero inflation was dealt with by assuming that the number of any given cetacean species seen in each segment followed a Tweedie distribution (Williams et al. 2011b). Tweedie distributions can be used to deal with zero-inflated count data, and work well when applied in a GAM framework (Williams et al. 2011b). The Tweedie family of distributions belong to the class of exponential dispersion models (EDMs) for which variance is proportional to some power of the mean (equation 9).

$$var[Y] = \varphi E(Y)^{\theta} \quad (9)$$

Special cases include the Gaussian ($\theta = 0$), Poisson ($\theta = 1$, $\varphi = 1$), and gamma ($\theta = 2$) distributions (Williams et al. 2011b; Joergensen 1987). The selection of an appropriate value of θ requires some trial and error. Once the spatial covariates were selected for each model, then the value of the Tweedie parameter, θ between 1 and 2, were assessed using quantile residual (qres) diagnostic plots, whereby the selected value of θ for each GAM model applied to each individual species in each survey region was the one that yielded points closest to a horizontal line with a slope of zero in the plot of the square root of the absolute value of qres against the fitted value (Williams et al. 2011b).

Currently, the DSM component in *Distance* does not offer the choice of applying a Tweedie distribution to the user's model (Equation 8). As a result, I fit the spatial model for each species detected within each of the survey regions with a Tweedie distribution in R using the mgcv package (R Development Core Team 2009; Wood 2006). Mgcv uses Generalized Cross Validation (GCV) for model selection, such that the model with the lowest GCV score is selected (Williams et al. 2006b). I assessed the model fit by using the quantile residual (qres) function from the statmod package in R. Quantile residuals are based on the idea of inverting the fitted distribution function at each response value to obtain standard normal residuals (R Development core team 2009; Dunn and Smyth 1996). Quantile residuals are useful for models with large dispersion situations where the deviance can be non-normal, and when the response only has a small number of distinct values (R Development core team 2009; Dunn and Smyth 1996).

Although the level of smoothing is optimized by mgcv, the decision to include or drop a model term is not automated (Williams et al. 2006b). I applied the following framework proposed by Wood (2001) when deciding whether to drop or include a model term from equation (8) (Williams et al. 2006b; Williams and O'Hara 2010). Each explanatory variable used in the model can also follow a one-dimensional or linear function of midlat and midlon (e.g., midlat+midlon) or two-dimensional smooth (e.g., midlat X midlon). Model fit was assessed using the summary.gam and plot.gam functions in mgcv, which showed coefficients, GCV score, explanatory power (deviance explained), the quantile residual plots and AIC value. Models which minimized the AIC value were favoured.

For each model term (Wood 2001):

1. I examined the estimated number of degrees of freedom to see if it was close to 1.
2. I examined the 95% confidence interval to see if it included zero across the range of observations.
3. If the first 2 conditions were satisfied, then the term was dropped temporarily from the model to see whether the GCV score was lower.

I dropped a model term from the final spatial model if it satisfied all three of these conditions (Wood 2001; Williams et al. 2006b; Williams and O'Hara 2010). Following Wood (2001), each explanatory variable term was dropped one at a time. If the first criterion was met, but not the other two, then the smooth term (two-dimensional) was replaced by a linear term (one-dimensional).

2.4.4 (iv) Using the descriptive model to estimate cetacean abundance throughout each survey region

In order to estimate the abundance of each cetacean species within each survey region for this fourth step (iv), I created a gridded data set of approximately 2 nm on a side (i.e., 4 nm^2) for each survey region. Each grid square was assigned a value for each explanatory spatial variable in the model created in section 2.4.3 (Williams and O'Hara 2010; Williams et al. 2006b; Williams et al. 2011b).

I then passed these gridded data to the selected spatial model for each cetacean species detected in each survey region, using a function written in R (Williams and O'Hara 2010). The function gave an estimated number of cetacean individuals in each grid cell of each survey region based on the value of each grid cell's explanatory variables (Williams and O'Hara 2010) and provided a total estimate of abundance for each cetacean species in each survey region.

2.4.5 (v) Estimating the variance in the cetacean abundance estimate for each species in each survey region

Generating reliable variance estimates is an area of active statistical research with model-based abundance estimators (Hedley & Buckland 2004). Because variance cannot be estimated from the spatial model directly, it is common to bootstrap, using transects as independent sampling units (Hedley et al. 1999). Unfortunately, in a non-randomized survey design, it can be difficult to choose an appropriate unit for resampling. I estimated the variance in the abundance estimates for each grid cell using a non-parametric bootstrap with day as the resampling unit. In my dataset, each day in each survey region represents a single survey or transect for that region. Each bootstrap iteration resampled the same number of days. This approach assumes that on average, each day has a similar amount of effort. For all iterations, the shape (half-normal vs. hazard-rate) and truncation distance chosen for the original detection function were fixed, as were the variables specified in the original spatial model (i.e. the selected GAM and the value of the Tweedie parameter) (Williams et al. 2006b). Each iteration used the resampled dataset to recalculate the detection function, estimate the mean school size, model the cetacean abundance along the trackline, and generate a total cetacean abundance across the prediction grid.

This process was repeated 300 times, which is higher than the recommended 200 bootstraps required for generating 95% confidence intervals of abundance (Buckland et al. 2001, Williams et al. 2006b). The bootstrap method provided two confidence intervals: one based on the percentile approach and the other on a log-normal mean following Buckland et al. (2001). The value of the confidence intervals for each species were calculated from the total abundance generated through each iteration of the bootstrap. In order to plot the distribution of the abundance of each species in each grid cell for each survey region, all the abundance estimates for each grid cell generated from the bootstrap were calculated and the abundance values corresponding to the mean, 2.5th percentile, and 97.5th percentile abundance value for that grid cell were stored.

These values were used to create a plot of the mean, 2.5th percentile, and 97.5th percentile grid values for each species in each survey region.

If fewer than 30 sightings were available for a given species, then a Jackknife approach was deemed more appropriate. Jackknife estimates of variance were made by removing one transect, including all the associated effort and sightings data associated with the transect, and using the remaining data to predict abundance in R. The 95% confidence intervals were calculated based on a log-normal distribution, because the distribution of abundance is positively skewed and cannot be negative (Buckland et al. 2001).

3: RESULTS

3.1 Detection functions and school-size bias

For each species, sightings were pooled across all three survey regions in order to produce the detection function (Figure 13 and Table 4). A correction factor specific to each observer was generated from the distance estimation experiments described in section 2.3 (Table 5). Each observer's individual correction factor was applied to their own sightings, and these corrected distances were used in the generation of the detection function and subsequent abundance estimate for each species in each survey region. As well, due to the low number of observations of both minke and grey whales, and because of their biological similarities and detectability, I created a pooled detection function for these two species using both minke and grey whale sightings. For all species except the humpback whale, the half-normal model provided the best fit for the detection function based on the selection criteria presented in section 2.4.1 (i). For all species, selection of the hazard-rate model would have resulted in a lower estimate of $\hat{p}$ and thus a higher estimated animal density in each survey region than with the half-normal model.

Observed mean, median, and range in school size varied for each species (Table 6). *Distance* detected the presence of school-size bias in the sightings data during the calculation of the detection function for harbour porpoise and humpback whale (section 2.4.2 (ii)). As a result, the *E(s)* instead of the mean school size was used for these two species in 2.4.3 (iii) to calculate the abundance of each species in each segment of survey effort. It is important to note that for both the harbour porpoise and humpback whale, the value(s) of the *E(s)* did not differ greatly from the mean school size and that their corresponding confidence interval(s) overlapped. Where size bias was not detected, the observed mean school size was used instead. The variation in the observed school size (CV) was less than 10% for all species except the Pacific white-sided dolphin (Table 6). This may be due to the fact that dolphins tend to travel in large groups (>100 individuals), making the estimating of school size more difficult for the observer.

3.2 Abundance and distribution estimates for each species by survey region

Estimated mean abundances across the three survey regions differed considerably by species and by region (Table 7; Figure 14 to 47). Pacific white-sided dolphin were the most abundant species in any of the three survey regions, whereas minke whale were the least (Table 7). However, estimates of distribution and abundance for Pacific white-sided dolphins and minke whales were generated only in the SWN survey region. Density estimates showed that among survey regions, Dall’s porpoise were more dense in the SWN survey region compared to the SWS survey region (Table 7; Figure 14, 17, 20), which is also the region in which they were encountered most frequently (i.e., individuals per 100 km) (Table 3), whereas the harbour porpoise were denser in the SWS survey region than in the SWN survey region (Figure 23, 26, 29). Both porpoise species were found to occur in the SWWC survey region, however, confidence in the density estimates for this region are low due to the huge range in the estimated 95% CI for these species. The estimated 95% CI for the density of humpback whales in both the SWN and SWS survey regions overlapped (Figure 35, 28, 41), therefore, it is difficult to state which of these regions had the highest density of humpback whales. The greatest density estimate for humpback whales occurred in the SWWC survey region, which is also the area in which they were encountered most frequently (Table 3). However, confidence in the estimates produced for this region were the lowest, with a huge range in the estimated 95% CI (due in part to the small number of surveys carried out over the two seasons). Grey whales were only recorded in the SWWC survey region (Figure 47), however, no 95% CIs were produced because of the extremely small abundance estimates produced.

Mid-longitude and mid-latitude were the independent spatial variables selected in the best models for most species (Table 7; Appendix 1). For the larger whales, like humpback and minke whales, depth and tidal speed helped determine species distribution. For the smaller cetaceans like Dall’s porpoise, tidal speed appeared as an environmental variable in the best models, whereas sea surface temperature was selected as a determinant of harbour porpoise distribution.

4: DISCUSSION

4.1 The use of opportunistic platforms

This spatial modelling of opportunistically collected transect data by the Straitwatch program produced estimates of abundance and distribution for 6 cetaceans species found within the three Straitwatch program ranges during the summer months of 2008 and 2009. In B.C.'s coastal waters, estimates of abundance are limited for several of the species identified in this study. Studies previously conducted off the Pacific coast of Canada have provided some baseline estimates of distribution and both relative and absolute abundance (Table 1), however, these studies have been limited in scope both spatially and temporally and/or have been limited to a single cetacean species.

Results from this study illustrate that data collected by means of distance sampling methods from a "platform of opportunity", such as Straitwatch, can be used to produce estimates of absolute abundance and distribution for cetacean species, and can be especially relevant for regions in which estimates currently do not exist. By modifying the sightings data-collection methods of a small vessel that is already engaged in monitoring activities in and around cetaceans, I was able to not only produce estimates of abundance and distribution using a more efficient multi-purpose platform (compared to a dedicated large ship-based transect study), but I was also able to improve the usefulness of sightings data already being collected from this platform. Such methods could also be applicable to other platforms of opportunity engaged in whale-related activities such as whale watching vessels, expedition-based cruise ships, and other small research vessels where trained staff are already adept at identifying local cetaceans (Williams et al. 2006b). The key strength of my approach is its ability to produce an estimate of abundance from data that might otherwise only be used to calculate simple encounter rates along the survey track or catch per unit effort, such as has been done by Ford et al. (2011).

The use of opportunistic platforms can not only be efficient, but can also serve as a means to reduce bias in estimates of abundance and distribution produced through distance sampling techniques. Ships of opportunity can be used to train observers on distance sampling data-collection protocols or cetacean species identification. In addition, opportunistic platforms can also be used for experiments to quantify and correct for observer error (Section 2.3), and to help train observers to use laser range finders or radar to help visually estimate distance to a sighting (Williams et al. 2006b; Williams et al. 2007). Detection functions produced from greater sample sizes obtained using platforms of opportunity such as Straitwatch may also be applied in future design-based surveys conducted from the same platform. Future research within the Straitwatch survey regions aimed at assessing changes in abundance over time could also use the variance estimates produced from this analysis to estimate the sample sizes required to detect statistically significant declines in species abundance.

Furthermore, given the desire to reduce the impact of research studies on species identified as at risk, the use of opportunistic platforms can provide a means to collect multiple data sets that can be utilized for multiple analyses related to the impacts of anthropogenic threats. Multidisciplinary studies should be encouraged and researchers should be proactive in their partnership with other research programs or projects in order to share costs or promote logistical support for novel complementary studies (Williams, 2003). Resulting patterns of distribution, on both spatial and temporal scales, can be overlaid with an index of identified anthropogenic disturbances (e.g., entanglement, bycatch, ship strike, noise) to determine the probability for conflict and to explore the distribution of chance of conflict within a given study area (Williams and O'Hara 2010; Williams et al. 2011a).

4.2 Abundance estimates

Abundance estimates generated from this analysis are comparable to many of the estimates presented in Table 1. In the SWN survey region, my estimates of abundance for Dall's porpoise, and Pacific white-sided dolphin's fall within the 95% CI of those generated by Williams et al. (2007) and Ashe (2007) who's studies were carried out over the same region. While no abundance estimates currently exist for humpback whale's specific to the SWN survey region, estimates generated from this study overlap with counts of individuals occupying the survey region during the May to September period of 2008 and 2009 (Table 1 and Table 6). In addition, no estimates of abundance for harbour porpoise, or minke whale currently exist in the SWN survey region.

In the SWS survey region, estimates of abundance for Dall's porpoise and harbour porpoise generated from this analysis overlap with those generated by Keple (2002) and are slightly less than those presented by Hall (2004) who's study area included a larger region than the one used in this analysis (Table 1 and Table 6). There are currently no estimates of abundance specific to the SWS survey region or the region off southern Vancouver Island for humpback whales. In the SWWC region, no estimates of abundance currently exist for Dall's porpoise, harbour porpoise or humpback whales specific to the area surveyed. However, photo identification work has produced count estimates for grey whales and humpback whales. Due to the small sample size, no confidence intervals were generated for grey whales in the SWWC from this study and no comparison between other studies can be made.

Several sources of bias can affect the accuracy of the estimates of abundance produced by distance sampling in this study. These biases are derived from the primary assumptions of distance sampling. For instance, this method assumes that the radial angle and distance to the animal are measured without error, that a species is always detected if it is right on the trackline (i.e., *g(0)=1*), and that there is no responsive movement by the animal prior to detection (Williams and Thomas. 2007; Buckland et al. 2001). Other sources of bias include the presence of size bias such that larger clusters of animals are considered easier to detect and so may be over-represented in the sample (Thomas et al. 2002). School-size bias was estimated in section 2.4.2 (ii), and for those species for which school-size bias was detected, the *E(s)* was used instead of the mean school size as a means to correct for size-biased detection and for the underestimation of size of detected schools for each species in each survey region in section 2.4.3 (iii) (Buckland et al. 2001). In addition, bias in sampling effort can occur if the tracklines do not provide representative coverage of the entire study area. This is overcome in design-based distance sampling where the survey design process is automated and tracklines are placed within the survey area such that all points in the study area have equal probability of being sampled (Buckland et al. 2001).

Bias in measurement error of distance to animal and radial angle can lead to inaccurate estimates of perpendicular distance from the trackline, resulting in a proportional bias in the resulting effective strip half-width generated for each species (Marques 2007). Whereas the inaccurate measurement of radial angle is believed to introduce little bias in the estimate of perpendicular distance from the trackline, the inaccurate measurement of observed distance to the animal can result in the under or over-estimation of density for a species within a given area (Williams and Thomas 2007; Marques 2007). I reduced the potential for this distance-estimation bias through observer-error distance experiments carried out with all observers. Through these experiments, a correction factor for each observer was generated and applied to all measurements of distance to animals by each observer in order to reduce the potential for systematic bias introduced from an observer’s tendency to under- or over-estimate distance.

The assumption that all animals along the trackline (i.e., at zero perpendicular distance from the trackline) are detected with certainty (i.e., *g(0)=1*) is the most obvious form of negative bias in abundance estimates (i.e., underestimating abundance) produced from line transect studies (Williams et al. 2006b). Bias related to this assumption comes in two forms, availability bias, where not all animals were available to be detected, and perception bias, where observers failed to detect all species present (Williams et al. 2006). Perception bias may be especially prevalent on platforms of opportunity like Straitwatch where observers are engaged in other data collection and monitoring activities.

The value of *g(0)* may be less than 1 for more cryptic species, for longer-duration diving species of cetaceans (e.g., sperm whales), or under poor survey conditions (i.e., increased sea state, poor weather), and may be closer to 1 for species which spend more time at the surface or which tend to aggregate in larger groups (e.g., Pacific white-sided dolphin). The use of multiple covariate distance sampling (MCDS) was attempted in this study in order to take into account that variables other than distance from the transect line might affect the detectability of a cetacean species (e.g., sea state, group size, weather) and that might determine the value of *g(0)*. For all species, detection function models produced using MCDS were not selected and so only distance from the trackline was used to fit detectability. The other covariates included in the MCDS analysis included sea state and school size. The value of the sea state did not vary greatly for each observation by species because the Straitwatch vessel tended to travel on the water within a narrow band of similar conditions. Furthermore, the variability of school size within each species was minimal. As the values of the covariates applied in MCDS differed little for each observation, it is possible that they did not appear to contribute greatly to the detectability of a given species.

Another way to minimize the bias in abundance estimates produced by perception bias is to conduct a double-observer or double-platform survey (Williams and Thomas 2007). Such surveys have observers that search independently of each other or are located on separate platforms (e.g., combined water and aerial-based survey platforms). Data collected from double-observer or double-platform surveys generate both conventional distance sampling data and mark-recapture data (Buckland et al. 2010), allowing for the estimation of *g(0),* which can be used to correct for bias in the subsequently produced estimates of abundance.

While this study assumed that the *g(0)* = 1, it is likely that *g(0)* could be <1 for many of these species. As such, estimates of absolute abundance and their associated confidence intervals generated from this study should be considered as minimum estimates, with the magnitude of uncertainty associated with these abundance estimates having the potential to be far greater (i.e., greater range in values for the upper and lower confidence limits). Despite the bias introduced by this, if one assumes that the value of *g(0)* for each species from the Straitwatch platform was constant over time, then you could still utilize this platform for monitoring time trends in relative abundance. However, if the monitoring team were to improve over time, this might result in misleading trends. Specifically, if a detection function was generated based on initially overestimated distances to an animal of a given species and then the observers became more accurate over time but the same detection function was utilized in the analysis, then it may appear that there were more animals in the study area based on the fact that fewer of the observations fell outside of the truncation distance.

The detection function for Dall's porpoise indicated a large spike close to the trackline (i.e., at zero distance the value of the detection probability was around 1.5). One interpretation is that the observers rounded the angle off the trackline to zero in many cases for this species. The other is that there was responsive movement before detection (Figure 13), which is another source of bias in distance based abundance estimation. This spike suggests that Dall's porpoise showed responsive movement (attracted to the vessel's track line) relation to the survey vessel. However, bow-riding behaviour represented only 2% of the associated behaviour recorded with each sighting. We can therefore assume that the abundance estimates generated for Dall's porpoise in this study may be biased high because of attractive movement. Both the harbour porpoise and Pacific white-sided dolphin also exhibited a small spike near zero in the detection function (i.e., at zero distance, the value of the detection probability was around 1.2). However, for both species, little of the recorded behaviour(s) suggested responsive movement to the survey vessel.

Future work might include conducting design-based distance sampling surveys within each of the Straitwatch survey regions in conjunction with continued data collection from the Straitwatch vessel as a platform of opportunity. Design-based distance sampling surveys would be designed *a priori* such that the assumption of equal coverage probability is not violated (as with model-based distance sampling). Estimates of abundance and distribution generated from design-based distance sampling surveys could be compared to those generated for the same time period and spatial location using methods presented here, and the difference in estimates could be used to help explore the bias associated with model-based distance sampling methods applied here. Alternatively, to assess that bias in the abundance estimates, results from distance-based studies could also be compared to mark-recapture photo identification studies occurring within the same survey region(s) (Williams and Thomas 2009).

The use of surveys occurring over both years in the SWN survey region may have introduced increased variability in the estimates produced, but was required to improve the sample size for some species within this study area. In addition, the use of sightings from all survey regions to generate a global detection function for each species was useful in improving the detection function fit and ultimately the resulting estimate of abundance. The grey whale and minke whale sightings from all regions were also combined to create a universal detection function for these two species. While these two species are of similar size and exhibit very similar dive behaviour, it is possible that some bias might have been introduced into the estimates of abundance and distribution by assuming that both species have the same detectability. As a result of the bias associated with the model-based distance sampling results presented here, abundance estimates generated from this work can be considered a minimum estimate of the average cetacean abundance in the 3 survey regions during the 2008/2009 summers.

Estimates of abundance are essential when calculating acceptable levels of anthropogenic mortality associated with fisheries (i.e., entanglement, bycatch or direct catch) (Hall et al. 2002) and for assessing predator needs when establishing fishing quotas using an ecosystem-based management (EBM) approach. For example, Williams and O'Hara (2010) explored the potential mortality limit for fin whales and humpback whales in B.C. using the Potential Biological Removal (PBR) calculations laid out under the U.S. Marine Mammal Protections Act (because there is currently no quantitative objectives set forth in Canada to calculate allowable annual anthropogenic mortality to marine mammal stocks). The calculation of PBR for a given species in a survey region uses a current estimate of absolute abundance (Williams et al. 2008). Using the minimum abundance estimate(s) derived from a study such as this (e.g., 44 Dall's porpoise in the SWN survey region (Table 6)), and keeping in mind that the confidence intervals associated with these estimates are a minimum due to the uncertainties explained above, one could calculate the acceptable PBR level for a given species within the survey region and determine if this value is exceeded by estimates of the current mortality rate associated with entanglement and/or bycatch (obtained from observer and license holder data associated with a specific fishery) from the same survey region (Williams et al. 2008; Williams and O'Hara 2010). Similarly, in order to assess predator needs when following an EBM approach, if one knew the approximate energetic requirements of a cetacean species on a per capita basis, then one might produce an estimate of the required total volume of prey species for this cetacean species from its absolute abundance within the survey region (Williams et al. 2011c).

4.3 Distribution estimates

Estimates of spatial distributions of cetaceans are useful when determining habitat use of different species, identifying and protecting critical habitats, creating marine protected areas (Ashe et al. 2010), and determining the potential for anthropogenic disturbance to cetaceans within the three study areas. Additionally, the production of upper and lower confidence bound distribution maps from this study could be applied by managers during the risk assessment process. For example, if a manager wanted to protect the most “critical” habitats of a given species, then they might want to use the lower-limit distribution map (showing the lowest 2.5th % of the abundance estimate across the spatial grid) with the idea that at least the species is found in X locations. Conversely, if they were responding to an oil spill, managers may want to look at the upper limit distribution map with the idea that responders should act as though animals might be exposed all the way out to the upper limit of their distribution.

In addition to producing distribution estimates for a given species, the use of spatial modelling techniques can be useful to managers who are interested in determining which spatial covariates appear to be important predictors of the presence of cetacean species (Buckland et al. 2004). For example, results from this study suggest that tidal speed may be an important factor in determining the distribution of small cetacean species such as Dall’s porpoise in some survey regions. The correlation between Dall’s porpoise occupation of a given location and tidal speed may be an important consideration for managers exploring the impacts of tidal power projects (which target areas of higher tidal exchange) as a future source of “green energy” in B.C. Further study related to the relationship between tidal speed strengths required to produce adequate tidal power and the tidal speed strengths preferred by small cetaceans, such as the Dall’s porpoise, may be required to assess the magnitude of the interaction between these species and the development of tidal power projects.

Data collected from opportunistic platforms can also be useful for future cetacean research aimed at identifying areas of high species density. Such areas identified by this study can be targeted by future research that requires a high probability of species presence, such as mark-recapture studies or entanglement-scar-rate studies that employ photo-identification techniques, and/or biopsy studies (Williams et al. 2006; Robbins 2010). Future line-transect surveys could benefit from studies such as this by allowing for the design of surveys across known species density gradients and by identifying areas of high species density where more intense species-specific designed line-transect surveys might be conducted (Williams et al. 2006). Furthermore, results from this study could help inform how researchers may want to stratify sample design based on the gradient of the variables identified through this work that may be associated with abundance (Appendix 1).

Additionally, the spatial models generated through this study are easily adaptable to spatio-temporal models, which could be used for the prediction of species distribution. The use of an opportunistic platform, such as in this study, allows for considerable seasonal and inter-annual coverage. Given enough support for data collection and analysis, spatio-temporal models generated through this work could be used to produce predictive models of distribution. Those results could then be used to anticipate species distributions during periods of anthropogenic activity that might have a negative impact on a given species (e.g., pile driving and its impact on harbour porpoise) (Tougaard et al. 2009). Results generated through an inexpensive study such as this are also beneficial because they can help identify fine-scale and temporal trends targeted to specific areas or specific species and can be coupled with larger-scale more expensive surveys that occur every 5 to 10 years and provide estimates of absolute abundance on a regional scale.

4.4 Future recommendations

Several options exist to improve future estimates of abundance and distribution from spatially modelled distance sampling. They include the use of a soap-film smooth, which has been found to make better model predictions of distribution and abundance near the edges of complex survey regions like mine (Wood et al. 2008; Hedley et al. 2009; Williams et al. 2011b). In addition, the methods developed by Hedley et al. (2009), allow for the propagation of the uncertainty associated with detection function modeling to the final abundance estimate of the spatial model. Future work associated with analysing data collected from an opportunistic platform such as Straitwatch should attempt to incorporate these recent improvements, but they are still under development and are therefore beyond the scope of the current project.

Furthermore, the collection of spatial covariate data that may be of direct biological relevance to the prediction of animal distribution (e.g., prey availability or productivity) rather than habitat proxies (e.g., SST, tidal speed, depth, salinity) may improve results from spatially modelled distance sampling, and may provide better estimates of the predictors of cetacean distribution within the survey region. Moreover, the collection of real-time spatial covariate data may provide a more accurate predictor of species occurrence. In the case of this study, the covariates used to model species abundance and distribution were averaged over a long period and by season, and pre-dated 2009 the final season of this research projects sampling period. Because the average values of the covariates used in this analysis (e.g., tidal speed, salinity, SST, depth) may differ from the value of the covariates during the sampling seasons (2008 and 2009), there may be some bias introduced by using the averages instead of real-time data. Future work may include collecting information on spatial covariates that may be more likely to be positively correlated with species occurrence in a given location, simultaneously (i.e., real-time estimates) with sightings data to provide a more accurate representation of the predictors of the distribution of cetacean species within a given survey region. However, one limitation to the inclusion of real-time data in the modelling process or information of greater biological relevance is that spatial covariate data must be available/collected along each trackline and in every grid cell within the survey region. Given that these animals move and integrate across large spatial areas, it would be very challenging to collect enough data on the potential environmental variables contributing to the distribution of these species at the correct spatial scale.

Finally, the application of studies using double-platforms in order to address the uncertainty with the *g(0)=1* assumption could be applied to help improve the fit of detection functions, and ultimately the corresponding estimates of abundance generated from them (Williams and Thomas, 2007).

REFERENCE LIST

Ashe, E. 2007. *Survival and abundance of Pacific white-sided dolphins (Lagenorhynchus obliquidens) in the inshore coastal waters of British Columbia, Canada*. Unpublished Master's thesis, University of St-Andrews. St-Andrews. Scotland.

Ashe, E., Noren, D., and Williams, R. 2010. Animal behaviour and marine protected areas: incorporating behavioural data into the selection of marine protected areas for an endangered killer whale population. *Animal conservation*. 13: 196-203.

Barlow, J., and Forney, K.A. 2007. Abundance and population density of cetaceans in the California Current ecosystem*. Fisheries Bulletin,* 105: 509-526.

BCCSN (B.C. Cetacean Sightings Network). 2006. *Improving the utility of an opportunistic sightings network* - Final Report for National Fish and Wildlife Foundation Grant.

BCCSN. 2011. Wild Whales. Retrieved July 20, 2011, from: www.wildwhales.org

Bowen, W.D. 1997. Role of marine mammals in aquatic ecosystems. *Marine Ecology Progress Series*, 158: 267-274.

Brereton, T., Williams, A.D. and Williams, R. (2000) A low-cost method to determine bathypelagic and seasonal occupancy of fin whale *Balaenoptera physalus* in the Bay of Biscay, In: *European Research on Cetaceans –14* (Ed. by P.G.H. Evans), pp. 14–18. European Cetacean Society, Rome, Italy.

Buckland, S.T., Cattanach, K.L., and Anganuzzi, A.A. 1992. Estimating trends in abundance of dolphins associated with tuna in the eastern tropical Pacific Ocean, using sightings data collected on commercial tuna vessels. *Fishery Bulletin* 90: 1-12.

Buckland, S.T., Anderson, D.R., Burnham, K.P. and Laake, J.L. 1993. *Distance sampling: Estimating abundance of biological populations.* Chapman and Hall, London. 446pp.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., and Thomas, L. 2001. *Introduction to distance sampling: Estimating abundance of biological populations*. Oxford, UK: Oxford University Press.

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L., and Thomas, L. 2004. *Advanced distance sampling: Estimating abundance of biological populations.* Oxford, UK: Oxford University Press.

Buckland, S.T., Laake, J.L., and Borchers, D.L. 2010. Double observer line transect methods: Levels of indepedence. *Biometrics,* 66: 169-177.

Calambokidis, J., J.L. Laake, and Osmek S.D. 1997. *Aerial surveys for marine mammals in Washington and British Columbia inside waters*. Final report to the National Marine Mammal Laboratory, Seattle, WA 98115.

Calambokidis, J., Falcone, E.A., Quinn, T.J., Burdin, A.M., Clapham, P.J., Ford, J.K.B., Gabriele, C.M., LeDuc, R., Mattila, D., Rojas-Bracho, L., Straley, J.M., Taylor, B.L., Urbán R., J., Weller, D., Witteveen, B.H., Yamaguchi, M., Bendlin, A., Camacho, D., Flynn, K., Havron, A., Huggins, J., and Maloney, N. 2008. SPLASH: Structure of populations, levels of abundance and status of humpback whales in the North Pacific. Final Report Final report for NOAA Contract AB133F-03-RP-00078 Cascadia Research Collective, Olympia, WA.

Carretta, J.V., K.A. Forney, M.S. Lowry, J. Barlow, J. Baker, D. Johnston, B. Hanson, R.L. Brownell Jr., J. Robbins, D.K. Mattila, K. Ralls, M.M. Muto, D. Lynch, and L. Carswell. 2010. *U.S. Pacific marine mammal stock assessments: 2009.* U.S. Department of Commerce, NOAA Technical Memorandum NMFS-SWFSC-453. 336p

CETUS. 2009. *Cetus Research and Conservation Society.* Retrieved May 30, 2010, from http://www.cetussociety.org/whale_watching_guidelines.html.

Chandle, T. and Calambokidis, J. 2003. 2002 Aerial surveys for harbor porpoise and other marine mammals off Oregon, Washington, and British Columbia. Final report to the National Marine Mammal Laboratory, Seattle, WA 98115.

COSEWIC. 2011. *COSEWIC species assessment*. Retrieved June 12, 2011, from http://www.cosewic.gc.ca/eng/sct1/searchform_e.cfm.

Dunn, K. P., and Smyth, G. K.1996. Randomized quantile residuals. *Journal of Computational and Graphical Statistics* 5, 1-10. http://www.statsci.org/smyth/pubs/residual.html

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A.T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberon, J., Williams, S., Wisz, M.S. and Zimmermann, N.E. 2006. Novel methods improve prediction of species' distribution from occurrence data. *Ecography* 29: 129-151.

Erbe, C. 2002. Underwater noise of whale-watching boats and potential effects on killer whales (*orcinus orca*), based on an acoustic impact model. *Marine mammal science*, 18(2): 394-418.

Evans, P.G.H., and Hammond, P.S. 2004. Monitoring cetaceans in European waters. *Mammal Review,* 34: 131-156.

Fisheries and Oceans Canada (DFO). 2008. Recovery strategy for the Northern and Southern resident killer whales (*orcinus orca*) in Canada. *Species at Risk Act* Recovery Strategy Series, Fisheries & Oceans Canada, Ottawa, ix + 81 pp.

Fisheries and Oceans Canada (DFO). 2010a. Recovery strategy for the North Pacific humpback whale (*Megaptera novaeangliae*) in Canada [DRAFT]. *Species at Risk Act* Recovery Strategy Series. Fisheries and Oceans Canada, Ottawa. x + 51 pp.

Fisheries and Oceans Canada (DFO). 2010b. Management plan for the Grey Whale (*Eschrichtius robustus*) in Canada [Proposed]. *Species at Risk Act* Management Plan Series. Fisheries and Oceans Canada, Ottawa. vi + 58pp.

Ford, J. K. B., Ellis, G. M., and Balcomb, K. C. 2000. Killer whales. UBC Press, Vancouver, BC.

Ford, J.K.B., G.M. Ellis, P.F. Olesiuk. 2005. Linking prey and population dynamics: did food limitation cause recent declines of 'resident' killer whales (*Orcinus orca*) in British Columbia? *DFO Can. Sci. Advis. Sec. Res. Doc.* 2005/042. 27pp. 2005.

Ford J.K.B., Rambeau A.L., Abernethy R.M., Boogaards M.D., Nichol L.M., and Spaven L.D. 2009a. An assessment of the potential for recovery of Humpback Whales off the Pacific coast of Canada. *DFO Can. Sci. Advis. Sec. Res. Doc.* 2009/015. iv + 33 p.

Ford, J.K.B., Ellis, G.M., Olesiuk, P.F., Balcomb, K.C. 2009b. Linking killer whale survival and prey abundance: food limitation in the oceans' apex predator. *Biol. Lett.,* 6(1): 139-142

Ford, J.K.B., Abernethy, R.M., Phillips, A.V., Calambokidis, J., Ellis, G.M., and Nichol, L.M. 2010. Distribution and relative abundance of cetaceans in western Canadian waters from ship surveys, 2002-2008. *Can. Tech. Rep. Fish. Aquat.Sci.* 2913: v + 51 p.

Guenther, T.J., Baird, R.W., Bates, R.L., Willis, P.M., Hahn, R.L., and Wischniowski, S.G. 1995. *Stranding and fishing gear entanglements of cetaceans on the west coast of Canada in 1994*. Report submitted to the International Whaling Commission, SC/47/06.

Hall, A., Ellis, G., and Trites, A.W. 2002. *Harbour porpoise interactions with the 2001 selective salmon fisheries in southern British Columbia and license holder reported small cetacean bycatch.* Selective Salmon Fisheries Science Program Report: Fisheries and Oceans Canada: 51.

Hall, A. 2004. *Seasonal abundance, distribution and prey species of harbour porpoise (Phocoena phocoena) in southern Vancouver Island waters.* Unpublished Master's thesis, University of British Columbia. Vancouver, British Columbia, Canada.

Hall, M.A., Alverson, D.L., and Metuzals, K.I. 2000. By-catch: problems and solutions. *Marine Pollution Bulletin*., 41(1-6): 204-219.

Hastie, T.J. and Tibshirani, R.J. 1990. *Generalized Additive Models*. Chapman and Hall, London.

Hedley, S., Buckland, S.T., and Borchers, D.L. 1999. Spatial modelling from line transect data. *J. Cetacean Res. Manage.*, 1(3): 255-264.

Hedley, S. 2000. *Modelling heterogeneity in cetacean surveys.* Unpublished Phd thesis, University of St. Andrews. Edinburgh, Scotland.

Hedley, S. and Buckland, S.T. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics,* 9: 181-199.

Hedley, S.L., Bannister, J.L. and Dunlop, R.A. 2009. Group IV humpback whales: abundance estimates from aerial and land-based surveys off Shark Bay, western Australia, 2008. *Paper SC/61/SH23 presented to the Scientific Committee of the International Whaling Commission, June 2009 (unpublished). 17pp.*

Heise, K. 1997. Life history and population parameters of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*). Rep. Int. Whal. Commn. 47: 817-825.

Hilborn, R. and Walters, C.J. 1992. *Quantitative fisheries stock assessment: choice, dynamics and uncertainty.* Massachusetts, USA: Kluwer Academic Publishers

Holt, M.M., Noren, D.P., Veirs, V., Emmons, C., and Veirs, S. 2009. Speaking up: killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. *J Acoust Soc Am.,* 125 (1): EL 27 to EL 32

International Dolphin Safe Monitoring Program. 2011. Accessed September 10, 2011 at: http://www.earthisland.org/dolphinSafeTuna/

Jennings, S., Kaiser, M.J., and Reynold, J.D. 2001. *Marine fisheries ecology*. Oxford, UK: Blackwell Publishing Company.

Jensen, A.S. and G.K. Silber. 2003. *Large whale ship strike database.* U.S. Department of Commerce, NOAA Technical Memorandum. NMFS-OPR- , 37 pp.

Joergensen, B. (1987). Exponential dispersion models. *J. R. Statist. Soc.* B 49, 127-162.

Keple, A.R. 2002. *Seasonal abundance and distribution of marine mammals in the Southern Strait of Georgia, British Columbia.* Unpublished Master's thesis, University of British Columbia. Vancouver, British Columbia, Canada.

Lachmuth, C. L., Barrett-Lennard, L.G., Steyn, D.Q., and Milsom, W.K. 2011. Estimation of southern resident killer whale exposure to exhaust emissions from whale-watching vessels and potential adverse health effects and toxicity thresholds. *Marine Pollution Bulletin*, doi:10.1016/j.marpolbul.2011.01.002

Larkin, P.A. 1996. Concepts and issues in marine ecosystem management. *Rev. Fish Biol. Fish.*, 6: 139-164.

Leaper, R., Fairbairns, R., Gordon, J., Hiby, A., Lovell, P. and Papastavrou, V. 1997. Analysis of data from a whalewatching operation to assess relative abundance and distribution of the minke whale (*Balaenoptera acutorostrata*) around the Isle of Mull, Scotland. *Reports of the International Whaling Commission*, **47**, 505–511.

Lusseau, D. 2006. The short-term behavioural reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Marine Mammal Science* 22(4): 802-818

MacLeod, C.D., Mandleberg, L., Schweber, C., Bannon, S.M., and Pierce, G.J. 2008. A comparison of approaches for modelling the occurrence of marine animals. *Hydrobiologia* 612: 21-32.

Marques, F. F. C. 2001. *Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity*. Dissertation, University of St. Andrews, St. Andrews, UK.

Marques, T.A. 2007. *Incorporating measurement error and density gradients in Distance sampling surveys*. Unpublished Doctoral thesis, University of St-Andrews. St-Andrews. Scotland.

Molnar, M. and Koshure, N. 2009. *Cleaning up our ocean: A report on pollution from shipping related sources in the Pacific North Coast Integrated Management Area (PNCIMA) on the British Columbia Coast*. David Suzuki Foundation, Vancouver, B.C.

Morton, A. 2000. Occurrence, photo-identification and prey of Pacific white-sided dolphins (*Lagenorhynchus obliquidens*) in the Broughton Archipelago, Canada 1984-1998. *Marine Mammal Science* 16 (1):80-93

Noren, D.P., Johnson, A.H., Rehder, D., and Larson, A. 2009. Close approaches by vessels elicit surface active behaviours by southern resident killer whales. *Endangered Species Research.*, 8: 179-192.

OceanWise. 2011. Accessed September 10, 2011 at: http://www.oceanwisecanada.org/

Pacific Wildlife Foundation, 2009. *Clayoquot Sound grey whale catalogue, 2006-2008*. Retrieved June 30, 2011, from: http://www.clayoquotwhales.ca/pdf/CS_GW_catalog_2009ed-opt.pdf

Pacific Wildlife Foundation, 2010. *Clayoquot Sound Humpback whale fluke identification catalogue, 1995-2009.* Retrieved June 30, 2011, from: http://www.clayoquotwhales.ca/pdf/CS_HB_catalog_6th_ed_c2010.pdf

R Development Core Team. 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Robbins, J. 2010. Scar-based inference in to Gulf of Maine humpback whale entanglement: 2008. Report to the National Marine Fisheries Service. Order number EA133F09CN0253.

Ross, P.S. 2000. Marine mammals as sentinels in ecological risk assessment. *Human and Ecological Risk Assessment*, 6(1): 29-46.

Royle, J.A. and Nichols, J.D. 2003. Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84(3): 777-790.

SARA, 2011. *Species at Risk Public Registry: A to Z Species Index*. Retrieved June 22, 2011, from: http://www.sararegistry.gc.ca/sar/index/default_e.cfm

Sea Choice: healthy choices, healthy oceans. 2011. Accessed September 10, 2011 at: http://seachoice.org/

Thomas, L., S.T. Buckland, K.P. Burnham, D.R. Anderson, J. L. Laake, D. L. Borchers, and S. Strindberg. 2002. Distance sampling. *Encyclopedia of Environmetrics* 1: 544-552.

Thomas, L. Williams, R. and Sandilands, D. 2007. Designing line transect surveys for complex survey regions. *J Cet Res Management.* 9(1): 1-13.

Thomas, L., S.T. Buckland, E.A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R.B. Bishop, T. A. Marques, and K. P. Burnham. 2010. Distance software: design and analysis of Distance Sampling surveys for estimating population size. *Journal of Applied Ecology* 47: 5-14. DOI: 10.1111/j.1365-2664.2009.01737.x

Tougaard, J., Cartensen, J., Teilmann, J., Skov, H., and Rasmussen, P. 2009. Pile driving zone of responsiveness extends beyond 20 km for harbour porpoise (*Phocoena phocoena* (L.)). *J. Acoust. Soc. Am.* 126(1):11-14.

Williams, R., Bain, D.E., Ford, J.K.B., and Trites, A.W. 2002. Behavioural responses of male killer whales to a ‘leapfrogging’ vessel. *J Cet Res Management*., 4: 305–310

Williams, R. 2003. *Cetacean studies using platforms of opportunity,* Unpublished Doctoral thesis, University of St-Andrews. St-Andrews. Scotland.

Williams, R., Lusseau, D., and Hammond, P.S. 2006a. Estimating relative energetic costs of human disturbance to killer whales *(Orcinus orca)*. *Biol Conserv.,* 133:301–311

Williams, R., Hedley, S.L., and Hammond, P. S. 2006b. Modelling distribution and abundance of Antarctic baleen whales using ships of opportunity. *Ecology and Society*, 11(1): 1

Williams, R., Leaper, R., Zerbini, A. and Hammond, P.S. (2007) Methods for investigating measurement error in cetacean line transect surveys. *Journal of the Marine Biological Association of the United Kingdom* 87: 313-320.

Williams, R., and Thomas, L. 2007. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada. *J. Cetacean Res. Manage.*, 9(1): 15-28.

Williams, R., Hall, A., and Winship, A. 2008. Potential limits to anthropogenic mortality of small cetaceans in coastal waters of British Columbia. *Can. J. Fish. Aquat. Sci.* 65: 1867-1878.

Williams, R. and O'Hara, P. 2010. Modelling ship strike risk to fin, humpback and killer whales in British Columbia, Canada. *J. Cetacean Res. Manage.* 11 (1): 1-8.

Williams, R., Thomas, L. 2009. Cost-effective abundance estimation of rare animals: Testing performance of small-boat surveys for killer whales in British Columbia. *Biological Conservation.*142:1542-1547

Williams, R., Ashe, E., and O'Hara, P.D. 2011a. Marine mammals and debris in coastal waters of British Columbia, Canada. *Marine Pollution Bulletin*. 62: 1303-1316.

Williams, R., Hedley, S. L., Branch, T.A., Bravington, M.V., Zerbini, A.N., and Findlay, K.P. 2011b. Chilean blue whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. *Conservation Biology*, 25(3): 526-535

Williams R, Krkošek M, Ashe E, Branch TA, Clark S, et al. 2011c. Competing Conservation objectives for predators and prey: estimating killer whale prey requirements for chinook salmon. *PLoS ONE* 6(11): e26738. doi:10.1371/journal.pone.0026738

Wood, S.N. 2001. MGCV: GAMs and Generalized ridge regression for R. *R News* 1(2): 20-25.

Wood, S.N. 2006. *Generalized additive models: An introduction with R*. Boca Raton: Chapman and Hall-CRC. 391pp.

Wood, S.N., Bravington, M.V., and S.L. Hedley. 2008. Soap film smoothing. *J.R. Statist. Soc. B.*, 70(5): 931-955.

.

TABLES

Table 1: Estimates of species abundance and encounter rate (numbers of either schools (Sch.) or individuals (Ind.) encountered per 100 km of distance travelled by boat) from past studies that occurred in B.C. and WA waters.

Species	Sch.	Sch./ 100km	Ind.	Ind./ 100km	Total Abundance (95% CI)		
					Estimate	Region	Year
Dall's porpoise	482[a]	1.61[a]	2,098[a]	7.02[a]	4,910 (2,700-8,940)[b]	Coast-wide, B.C.	Summer 2004 & 2005
					50 (10-244)[b]	Johnstone Strait stratum, B.C.	Summer 2004 & 2005
					1,545 (CV=0.43)[c]	Inland waters of WA & B.C.	1996
					200 (107-372)[d]	Haro Strait & Georgia Strait	Summer 2000
Harbour porpoise	73[a]	0.24[a]	203[a]	0.68[a]	9,120 (4,210-19,700)[b]	Coast-wide, B.C.	Summer 2004 & 2005
					2,895 (CV=0.41)[c]	Inland waters of WA & B.C.	1996
					1,239 (CV=0.41) [c]	Juan de Fuca Strait, WA & B.C.	1996
					745 (CV=0.53) [c]	Gulf Islands, B.C.	1996
					911 (CV=0.58) [c]	Strait of Georgia, B.C.	1996
					3,123 (CV=0.10)[e]	Inland waters of WA & B.C.	2002-2003
					10 (2-43)[d]	Haro Strait & Georgia Strait	Summer 2000
					675 (450-1,006)[f]	Juan de Fuca & Haro Strait(s)	2001-2002
Pacific white-sided dolphin	183[a]	0.61[a]	8,991[a]	30.1[a]	25,900 (12,900-52,100)[b]	Coast-wide, B.C.	Summer 2004 & 2005
					1,344 (365-5,081)[b]	Johnstone Strait stratum B.C.	Summer 2004 & 2005
					675 individuals[g]	Broughton Archipelago, B.C.	1984-1998
					355 (255-559)[h]	Broughton Archipelago, B.C.	1991-1992
					2,047 (1,037-4,040)[h]	Broughton Archipelago, B.C.	1999-2000

Species	Sch.	Sch./ 100km	Ind.	Ind./ 100km	Total Abundance (95% CI)		
					Estimate	Region	Year
Humpback whale	1,700[a]	5.69[a]	3.162[a]	1,058[a]	18,000-21,000[i]	Entire North Pacific	2004-2006
					3,000-5,100[i]	Gulf of Alaska, SE Alaska, & Northern B.C.	2004-2006
					200-400[i]	Southern B.C. & northern WA	2004-2006
					2,145 (1,970-2,331)[j]	Coast-wide, B.C.	2006
Humpback whale					1,310 (755-2,280)[b]	Coast-wide, B.C.	Summer 2004 & 2005
					14 – 21 individuals[k]	Johnstone Strait & Queen Charlotte Strait, B.C. (June to September)	2008-2009
					47 individuals[k]	Johnstone Strait & Queen Charlotte Strait, B.C.	2008
					34 individuals[k]	Johnstone Strait & Queen Charlotte Strait, B.C.	2009
Minke whale	18[a]	0.06[a]	21[a]	0.07[a]	358 individuals[l]	Clayoquot Sound, B.C.	1995-2009
					65 individuals[l]	Clayoquot Sound, B.C. (May to September)	2009
					388 (222-6,800)[b]	Coast-wide, B.C.	Summer 2004 & 2005
					6 individuals[m]	Johnstone Strait & Queen Charlotte Strait, B.C.	2008
					9 individuals[m]	Johnstone Strait & Queen Charlotte Strait, B.C.	2009
Grey whale	23[a]	0.08[a]	44[a]	0.15[a]	47 individuals[n]	Clayoquot Sound, B.C.	2006-2008

[a]Ford *et al.* 2010; [b]Williams and Thomas 2007; [c]Calambokidis *et al.* 1997;[d]Keple 2002; [e]Carretta *et al.* 2010; [f]Hall 2004, [g]Morton 2000; [h]Ashe 2007; [i]Calambokidis *et al.* 2008; [j]Ford *et al.* 2009;[k]Jackie Hildering (Marine Education and Research Society (MERS), jackiehildering@gmail.com, October 31, 2010, *pers.comm.*); [l]Pacific Wildlife Foundation 2010; [m]Jared Towers (Marine Education and Research Society (MERS), jrtowers@gmail.com, June 21, 2011, *pers.comm.*); and, [n]Pacific Wildlife Foundation 2009.

Table 2: **Area of each survey region and realized survey effort in the form of days and trackline length in nautical miles (nm) for each survey year and survey region in this study.**

Survey Region	Area (nm^2)	Year	Number of days	Total trackline length (nm)
SWN	93	2008	56	2,322
SWN		2009	45	1,974
SWS	483	2009	34	1,657
SWWC	246	2009	6	274
TOTAL	**822**	**ALL**	**141**	**6,227**

Table 3: **Sightings of cetacean species schools and individuals, as well as encounter rate of cetacean species schools and individuals by survey region for all 3 survey regions in this study and across 2008 and 2009.**

	Schools			**Individuals**			**Schools/100km**			**Individuals/100km**		
	SWN	**SWS**	**SWWC**	**SWN**	**SWS**	**SWWC**	**SWN**	**SWS**	**SWWC**	**SWN**	**SWS**	**SWWC**
Dall's porpoise	106	12	2	300	34	4	1.33	0.39	0.39	3.77	1.11	0.79
Harbour porpoise	3	54	11	4	94	25	0.04	1.76	2.17	0.05	3.06	4.93
Pacific white-sided dolphin	30	0	0	1,478	0	0	0.38	0	0	18.57	0	0
Humpback whale	24	5	22	29	10	28	0.30	0.16	4.34	0.36	0.34	5.52
Minke whale	9	1	0	9	1	0	0.11	0.03	0	0.11	0.03	0
Grey whale	0	0	5	0	0	8	0	0	0.99	0	0	1.58

Table 4: Truncation distance (*w*), sample size (n) before and after truncation, fitted detection function model (Half-normal (Hn) or Hazard Rate (HR)) (Figure 13), the ΔAIC value between the favoured Hn and HR models with the same truncation distance for those species with > 40 sightings, the ΔAIC_c value between the favoured Hn and HR models with the same truncation distance for those species with < 40 sightings, *p* value from the Kolmogorov-Smirnov goodness of fit test (*K-S p*) and the Cramer-von Mises test (*CvM p*), estimated probability of detection ($\hat{p}$), corresponding percentage coefficient of variation and standard error of $\hat{p}$, and the estimated probability density function on the trackline (*f(0)*), for each species derived from data pooled over all three survey regions in this study.

Species	*w* (m)	*n* before	*n* after	Model	ΔAIC	ΔAIC_c	*K-S p*	*CvM p*	$\hat{p}$	% CV $\hat{p}$	SE $\hat{p}$	*f(0)*
Dall's porpoise	290	122	112	Hn	1.83	NA	0.24	0.3	0.55	7.52	0.04	6.23×10^{-3}
Harbour porpoise	350	70	68	Hn	0.89	NA	0.98	0.9	0.34	11.65	0.04	8.30×10^{-3}
Pacific white-sided dolphin	210	30	27	Hn	NA	0.67	0.64	0.6	0.64	14.38	0.09	7.39×10^{-3}
Humpback whale	920	51	49	HR	0.34	NA	0.52	0.5	0.58	20.05	0.12	1.87×10^{-3}
Minke whale	675	10	9	Hn	NA	1.93	0.82	0.8	0.48	15.63	0.08	3.06×10^{-3}
Grey whale		5	4									

Table 5: **Estimated correction factor (*cf*) and the corresponding standard error (SE) of the *cf* for each individual observer generated following methods in section 2.3.**

Observer ID	Estimated *cf*	SE (*cf*)	Number of observations by species recorded by each observer						
			Dall's porpoise	Harbour porpoise	Pacific white-sided dolphin	Humpback whale	Minke whale	Grey whale	Total
1	1.00	0.01	10	40	1	23	0	4	78
2	1.05	0.03	11	0	0	0	0	0	11
3	0.99	0.01	1	0	1	4	2	0	8
4	1.00	0.01	10	0	0	0	2	0	12
5	0.95	0.01	23	0	3	10	0	0	36
6	1.01	0.01	34	2	9	6	2	0	53
7	1.00	0.01	2	3	0	0	1	0	6
8	1.05	0.02	31	23	13	6	2	0	75

Table 6: **Observed mean, median, and maximum schools size and estimated school size E(s) from size-bias regression used when school-size bias was detected. School-size bias was determined by performing a regression of cluster size on estimated detection probability such that a positive slope indicated size bias. Size bias was detected for harbour porpoise and humpback whale.**

	Observed						Estimated				
	Mean school size (n)	**Median school size**	**Max. school size**	**SE (n)**	**%CV of n**	**Size-bias detected**	***E(s)***	**SE *E(s)***	**%CV *E(s)***	**Slope of regression**	**SE Slope of regression**
Dall's porpoise	2.80	2	10	0.17	6.1	no	2.51	0.15	6.2	-0.40	0.21
Harbour porpoise	1.82	1	5	0.12	7.0	yes	1.83	0.12	6.7	0.02	0.18
Pacific white-sided dolphin	47.22	9	300	15.23	32.2	no	30.34	17.17	56.6	-2.56	1.35
Humpback whale	1.31	1	3	0.08	6.0	yes	1.32	0.07	5.3	0.04	0.18
Minke whale	1.15	1	2	0.10	9.0	no	0.97	0.06	6.2	-0.65	0.22
Grey whale											

Table 7: Model Fit, Tweedie Parameter (θ), estimated abundance $(\hat{N})$, & estimated density $(\hat{D})$ with corresponding 95% confidence intervals (CIs) by species & survey region

Survey Region	**Species**	**Independent variables from the best model fit**	(θ)	**% deviance explained**	$(\hat{D})$ $(\hat{N}/nm^2)$	**95% CI** $(\hat{D})$	$(\hat{N})$	**95% CI** $(\hat{N})$	**95% Lognormal CI** $(\hat{N})$
SWN	Dall's porpoise	*s(midlat,tidal.speed,depth)*	1.1	15.3	0.48	0.01-0.76	44	24-71	11-178
	Harbour porpoise	*s(midlat)+s(midlon)*	1.6	49.0	0.01	0-0.03	1	0-3	0-3
	Pacific white-sided dolphin	*s(midlat,midlon)*	1.3	31.4	6.31	0.94-75.8	587	87-7,048	5-62,962
	Humpback whale	*s(midlat,midlon)*	1.1	4.15	0.03	0.01-0.05	2	1-5	1-6
	Minke whale	*s(midlat,midlon,tidal.speed)*	1.1	19.5	0.01	NA	1	NA	0-4
SWS	Dall's porpoise	*s(midlon,tidal.speed)*	1.1	52.9	0.13	0.04-0.23	61	17-112	27-139
	Harbour porpoise	*s(midlat)+s(midlon)+s(sst)*	1.1	37.9	0.38	0.16-0.72	182	79-347	89-371
	Humpback whale	*s(midlon)+s(tidal.speed)+ s(depth)*	1.1	59.6	0.01	0-0.04	5	0-20	0-493
SWWC	Dall's porpoise	*s(midlat)+s(midlon)*	1.3	31.4	0.08	0.04-78.5	20	10-19,316	0-1,866
	Harbour porpoise	*s(midlat)+s(midlon)*	1.6	17.0	0.60	0-710.1	148	1-174,686	1-15,809
	Humpback whale	*s(midlat)+s(midlon)*	1.1	48.1	0.20	0.05-4.1X10^{19}	49	12-1X10^{22}	1-3,554
	Grey whale	*s(midlat)*	1.7	86.2	0.03	NA	6	NA	NA

FIGURES

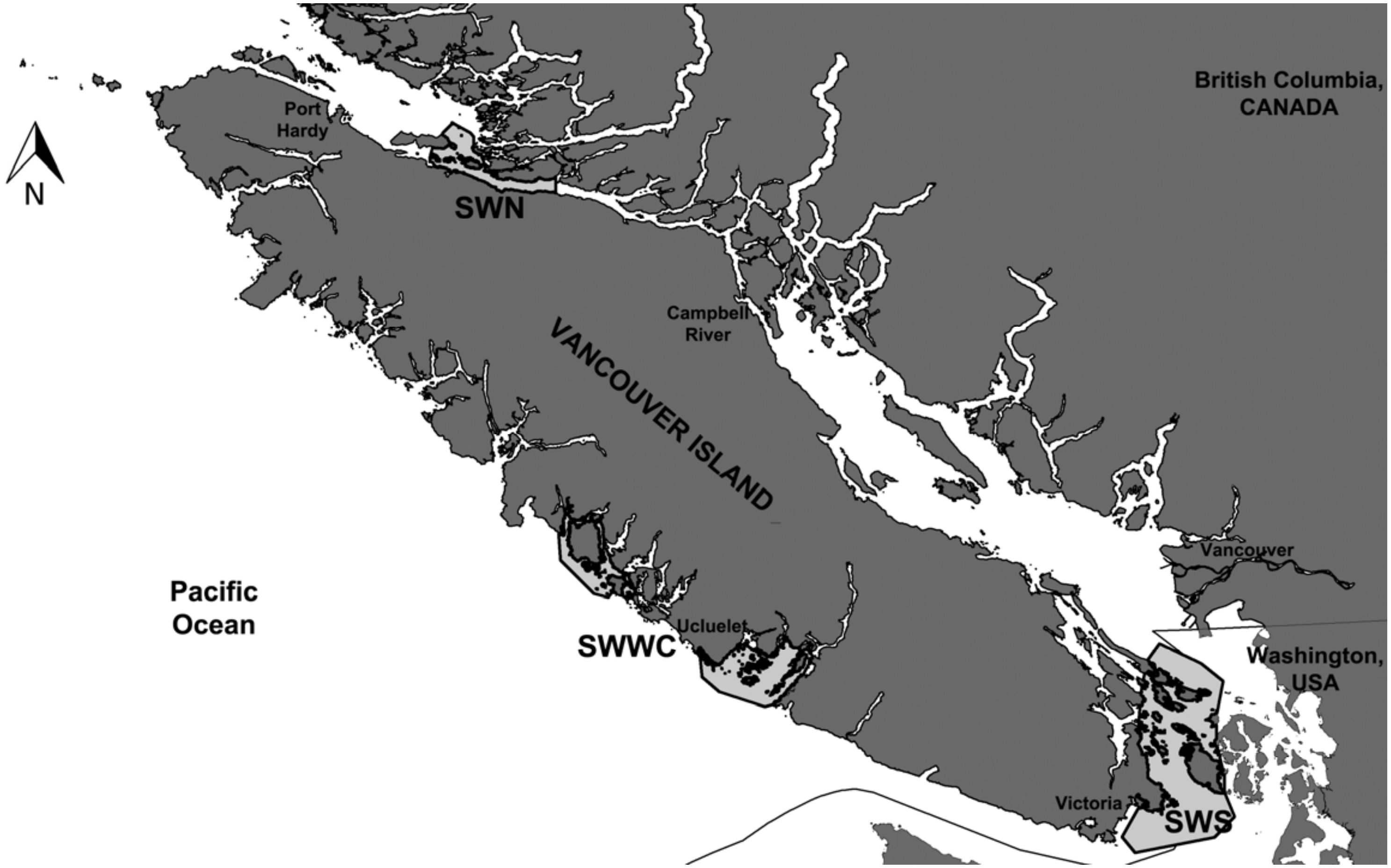


Figure 1: The Straitwatch program operates within 3 areas of Vancouver Island: Straitwatch North (SWN) (expanded in Figure 2), Straitwatch South (SWS) (expanded in Figure 3), and Straitwatch West Coast (SWWC) (expanded in Figure 4). These 3 areas make up the survey regions used in this study.

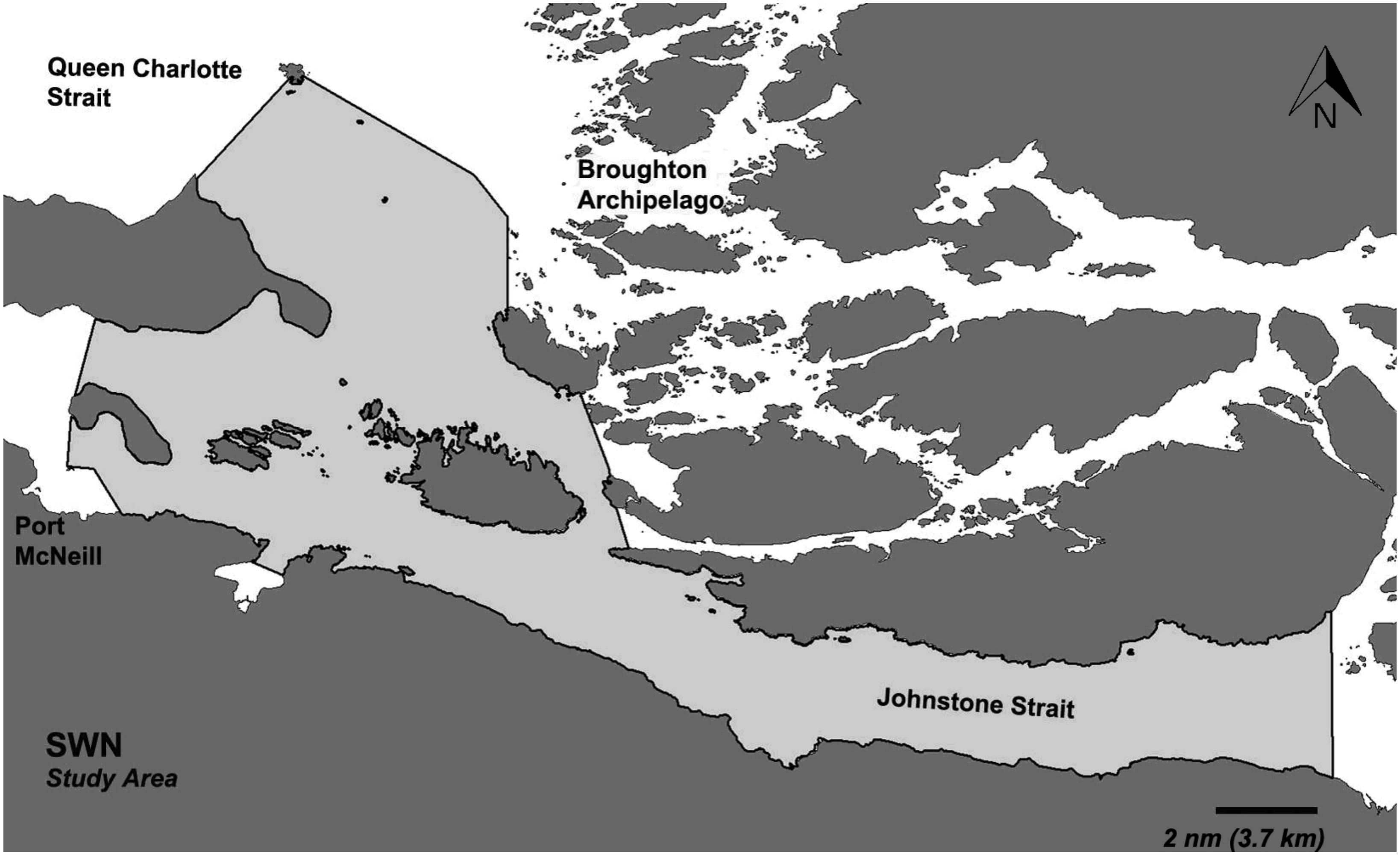


Figure 2: Straitwatch North (SWN) (light grey region) operates in the waters east of Port McNeill B.C., between Johnstone Strait and Queen Charlotte Strait.

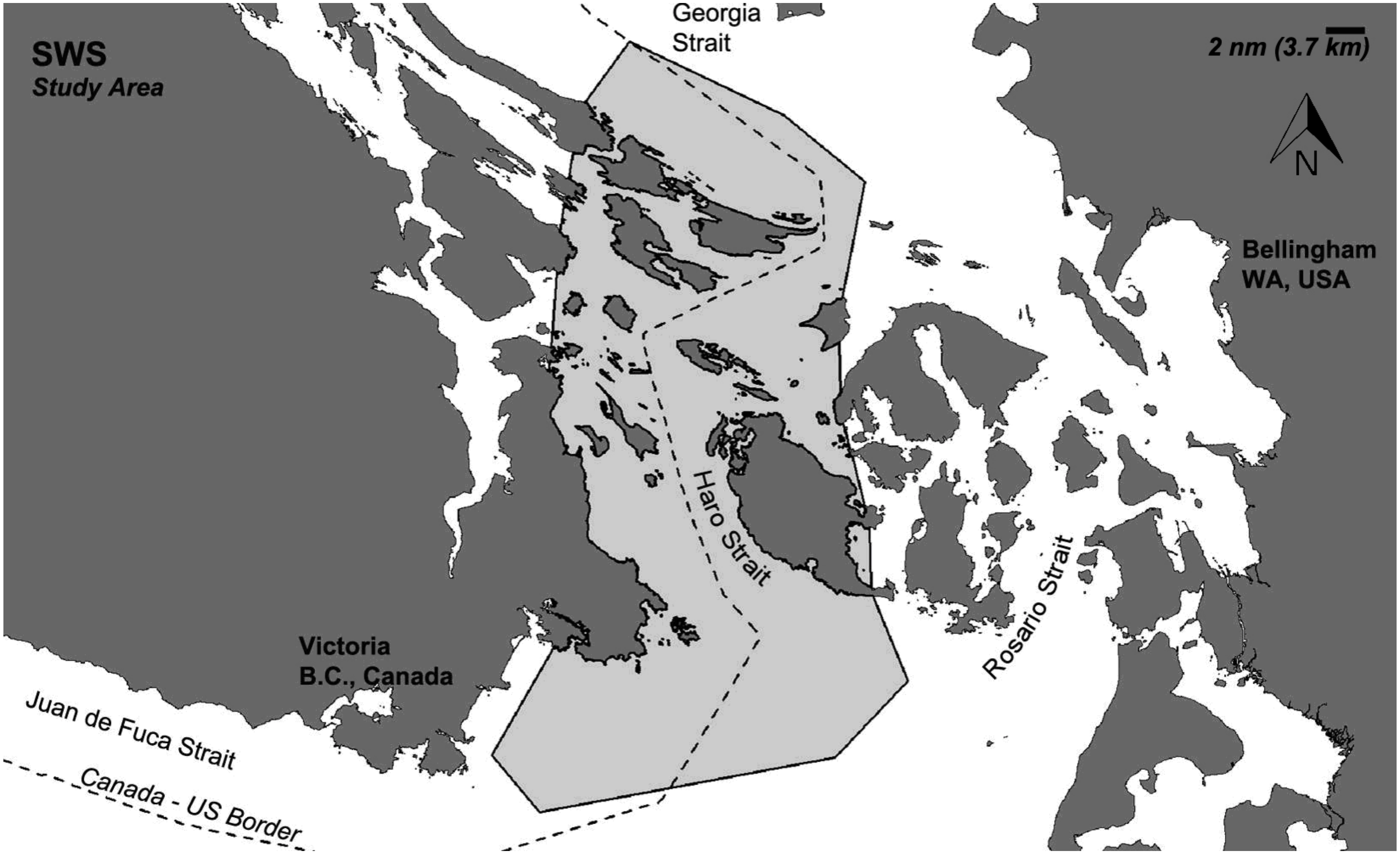


Figure 3: Straitwatch South (SWS) (light grey region) operates both in Canadian and American waters surrounding Victoria B.C., between Georgia Strait and Juan de Fuca Strait.

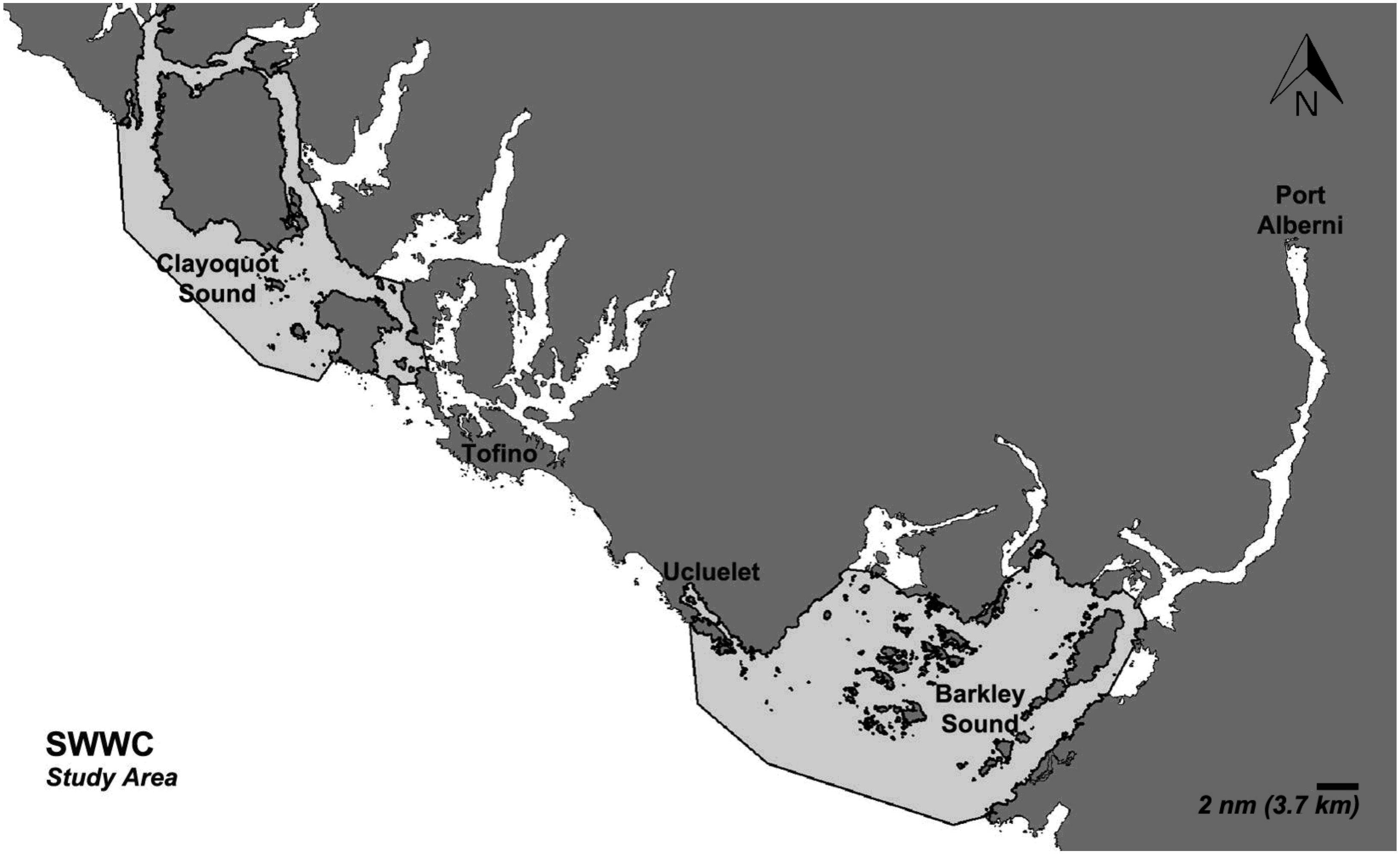


Figure 4: Straitwatch West Coast (SWWC) (light grey region) operates in the waters around Ucluelet and Tofino, B.C., in Barkley and Clayoquot Sounds.

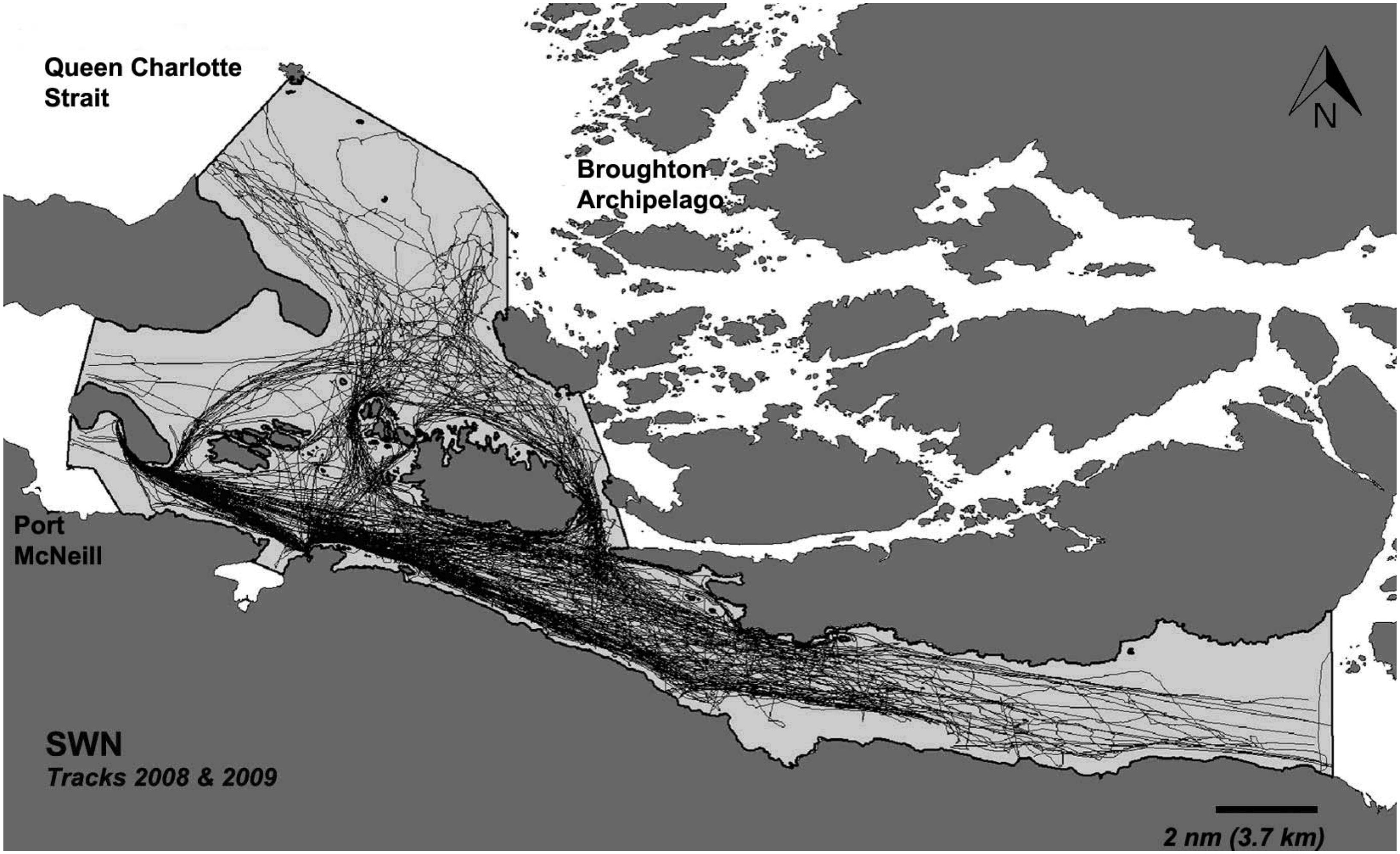


Figure 5: SWN recorded 2322 nm (4300 km) of trackline over 56 days in 2008 and 1974 nm (3656 km) of trackline over 45 days in 2009.

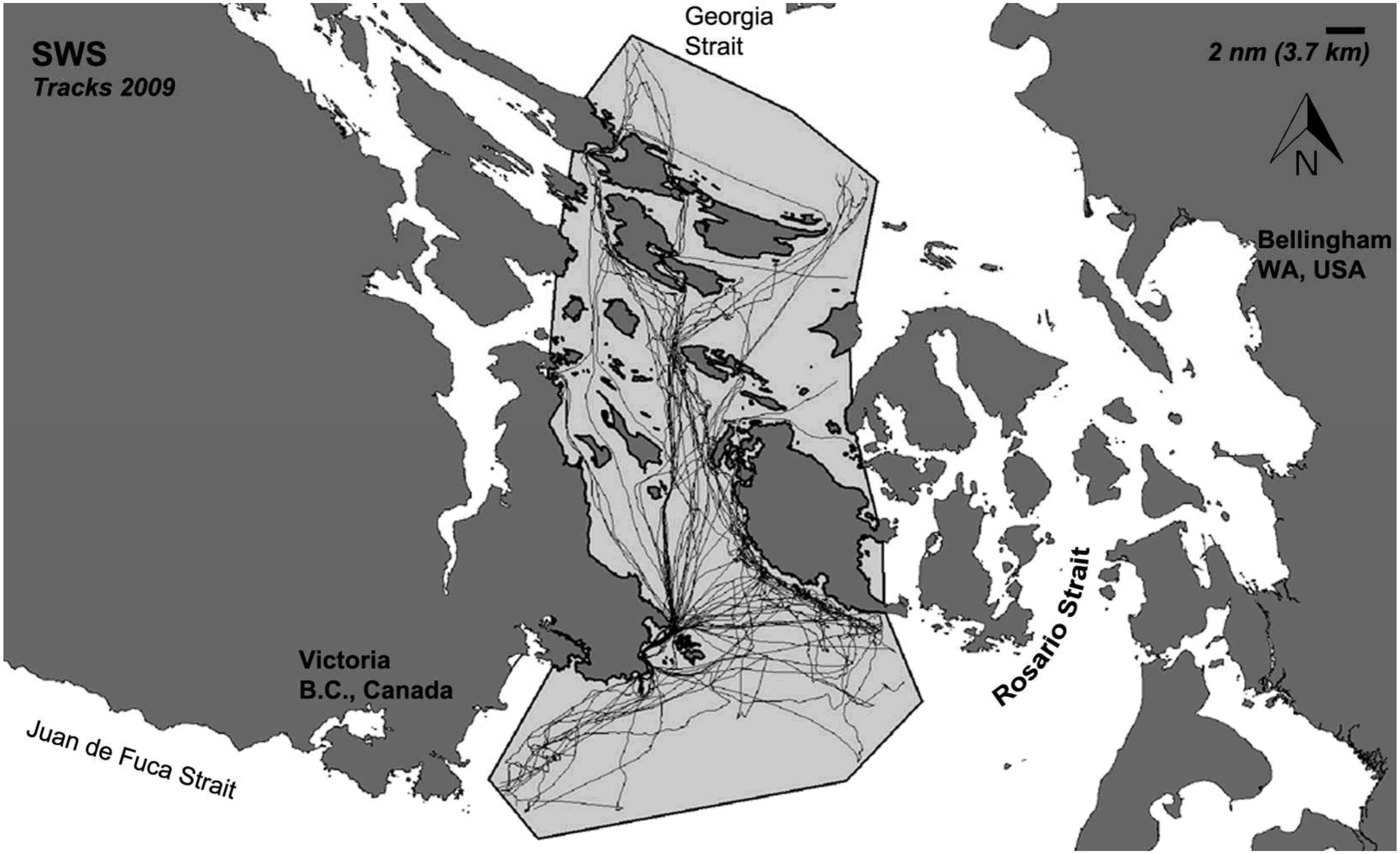


Figure 6: SWS recorded 1657 nm (3069 km) of trackline over 34 days in 2009.

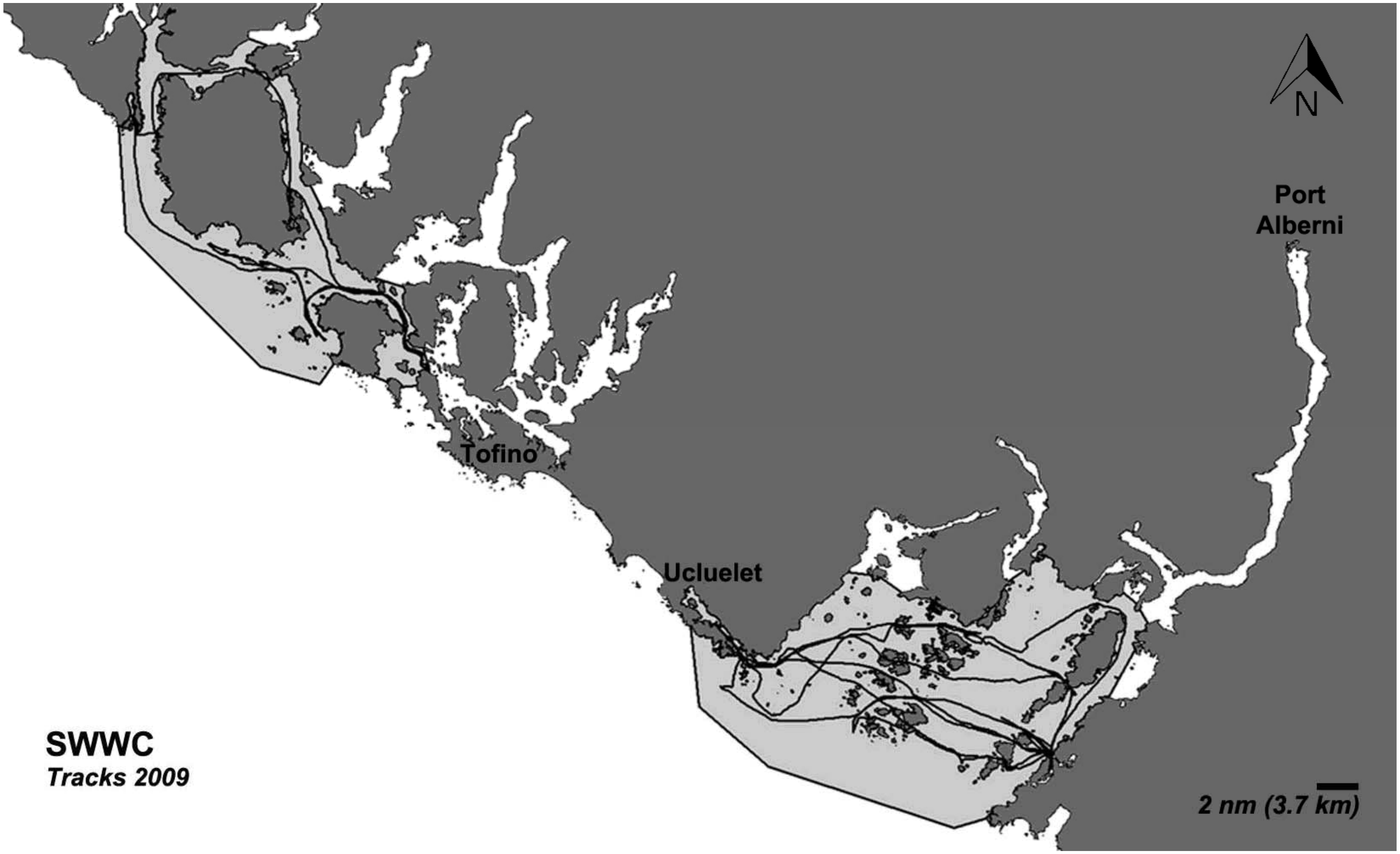


Figure 7: SWWC recorded 274 nm (507 km) of trackline over 6 days in 2009.

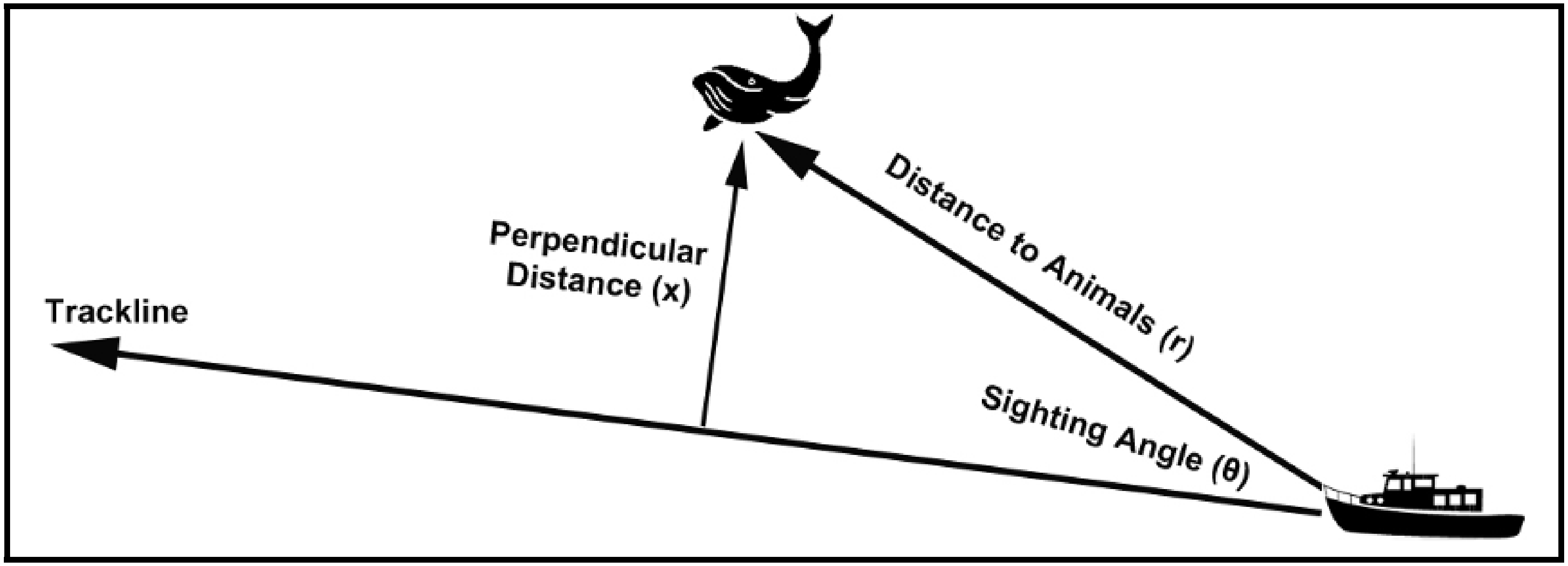
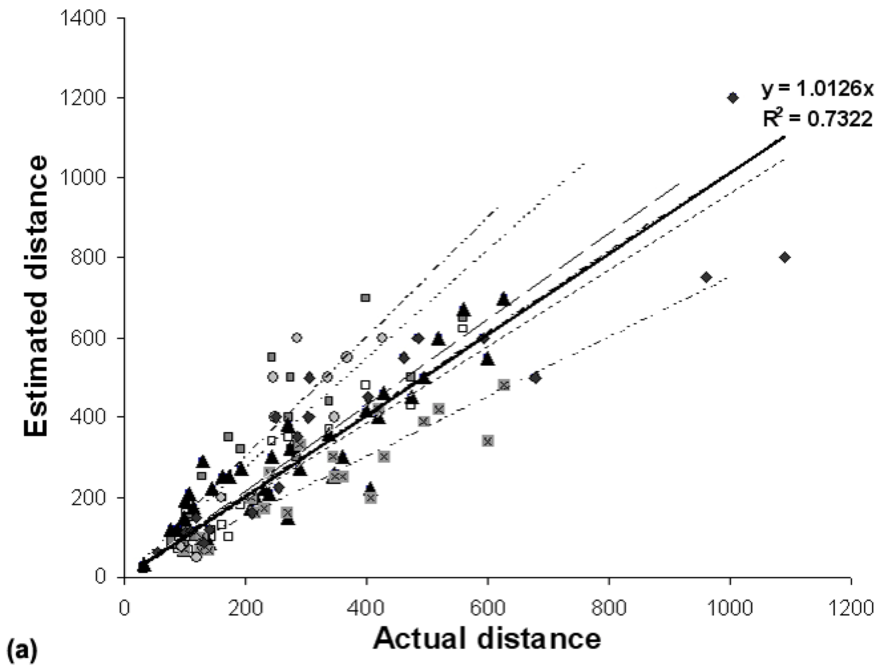


Figure 8: Measurements from line transect surveys to calculate the perpendicular distance *x*, where $x = r \cdot \sin(\theta)$



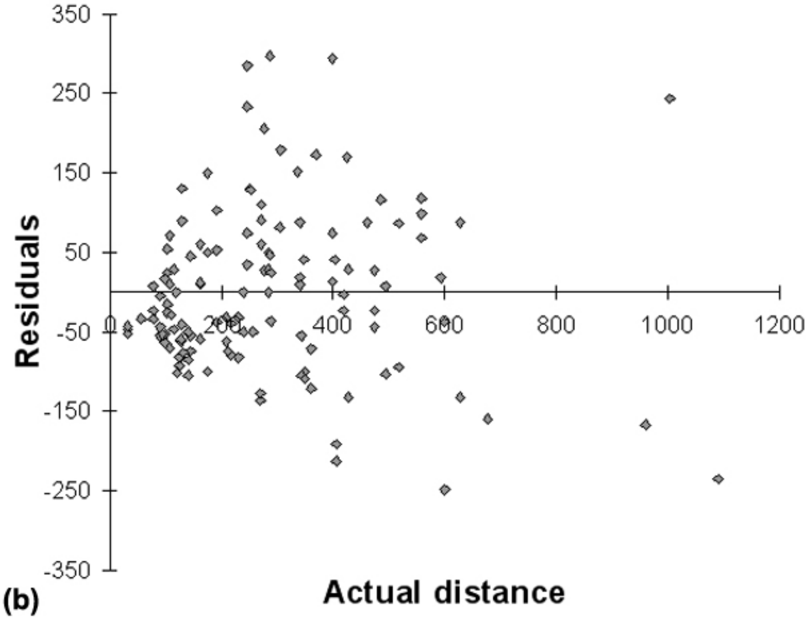


Figure 9: (a) A linear regression model of the estimated distance proportional to the actual distance. Each observer's paired estimates from studies carried out mid-season are represented by a unique symbol and fitted line. The solid black trend line is fit to all the observers' paired estimates. (b) A plot of the residuals across all observers.

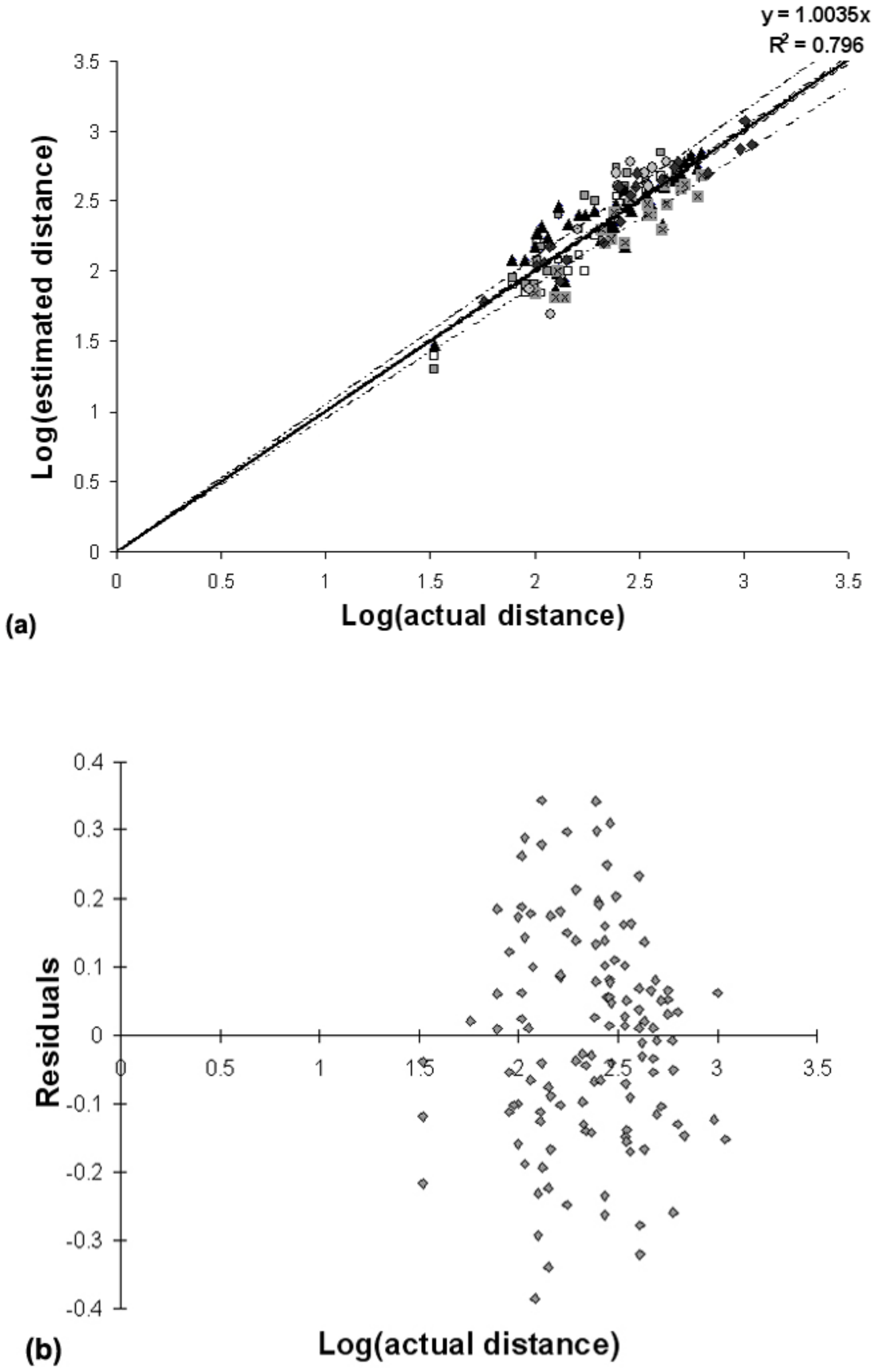


Figure 10: (a) A log-transformed (log base 10) linear regression model of the estimated distance proportional to the actual distance. Each observer's paired estimates are represented by a unique symbol and fitted line. The solid black trend line is fit to all the observers' paired estimates. (b) A plot of the residuals across all observers.

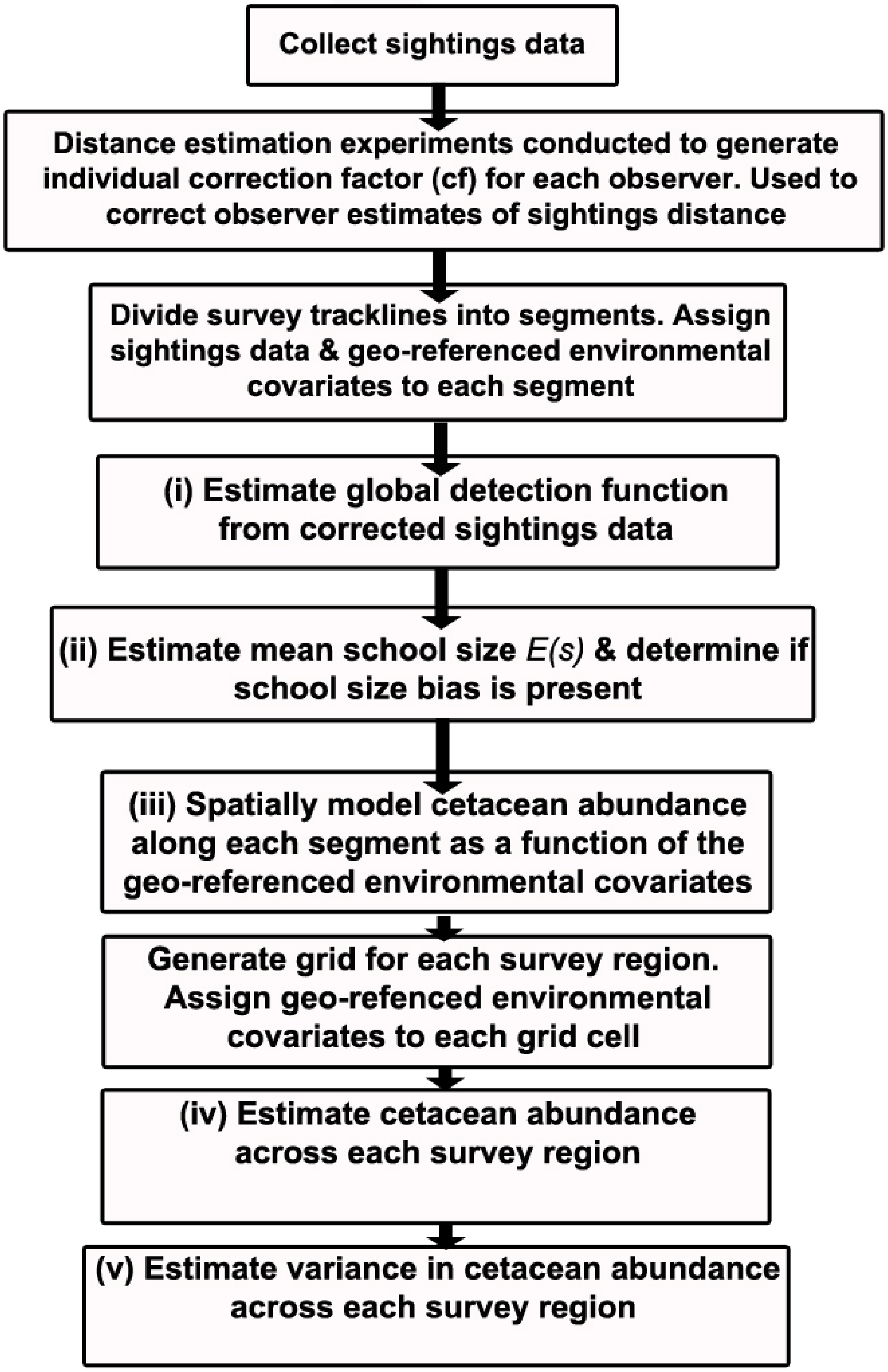


Figure 11: Simplified outline of the methods used to predict cetacean abundance for each species found within all three survey regions.

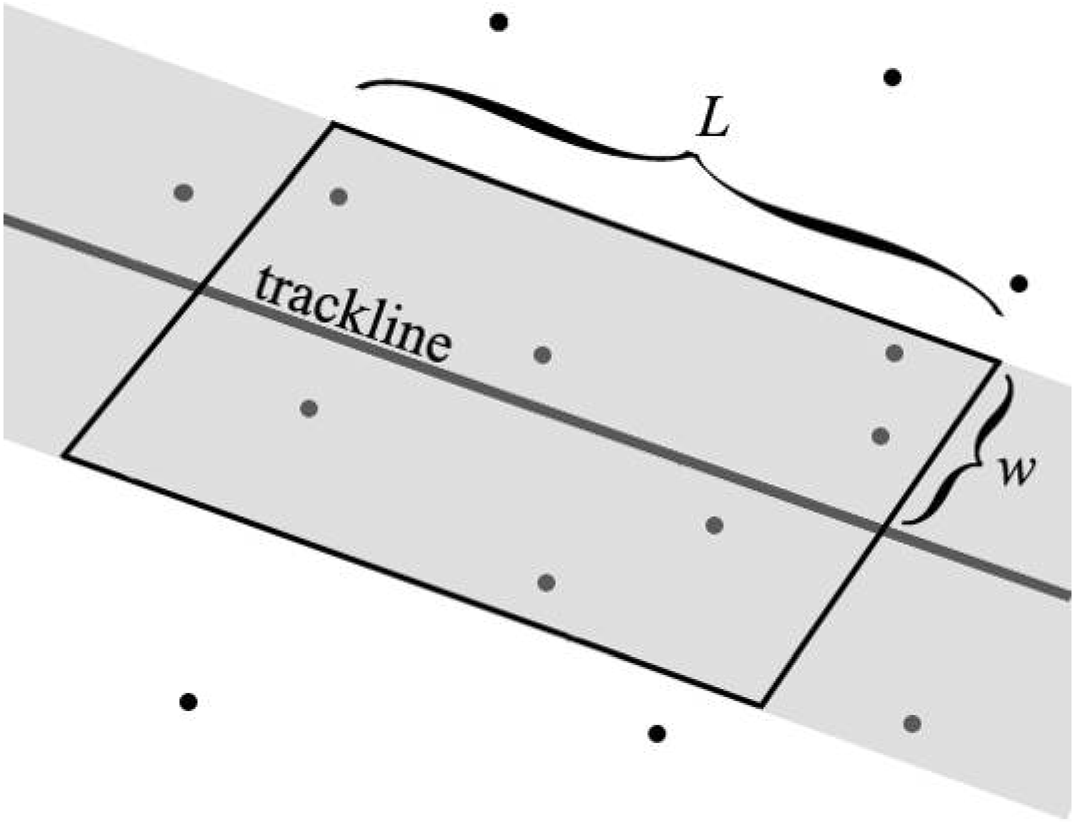


Figure 12: The distribution of the perpendicular distances from the trackline of those sightings within the truncation distance (*w*) (sightings that fall within the grey strip) are used to estimate the proportion of animals in the surveyed area (*a=2wL*) that are detected (i.e., the calculation of the detection function). This allows for the estimation of the animal density and abundance within the surveyed region. Sightings beyond the truncation distance from the trackline, i.e., sightings that do not fall within the grey strip, are not included in the calculation of the detection function or abundance estimates.

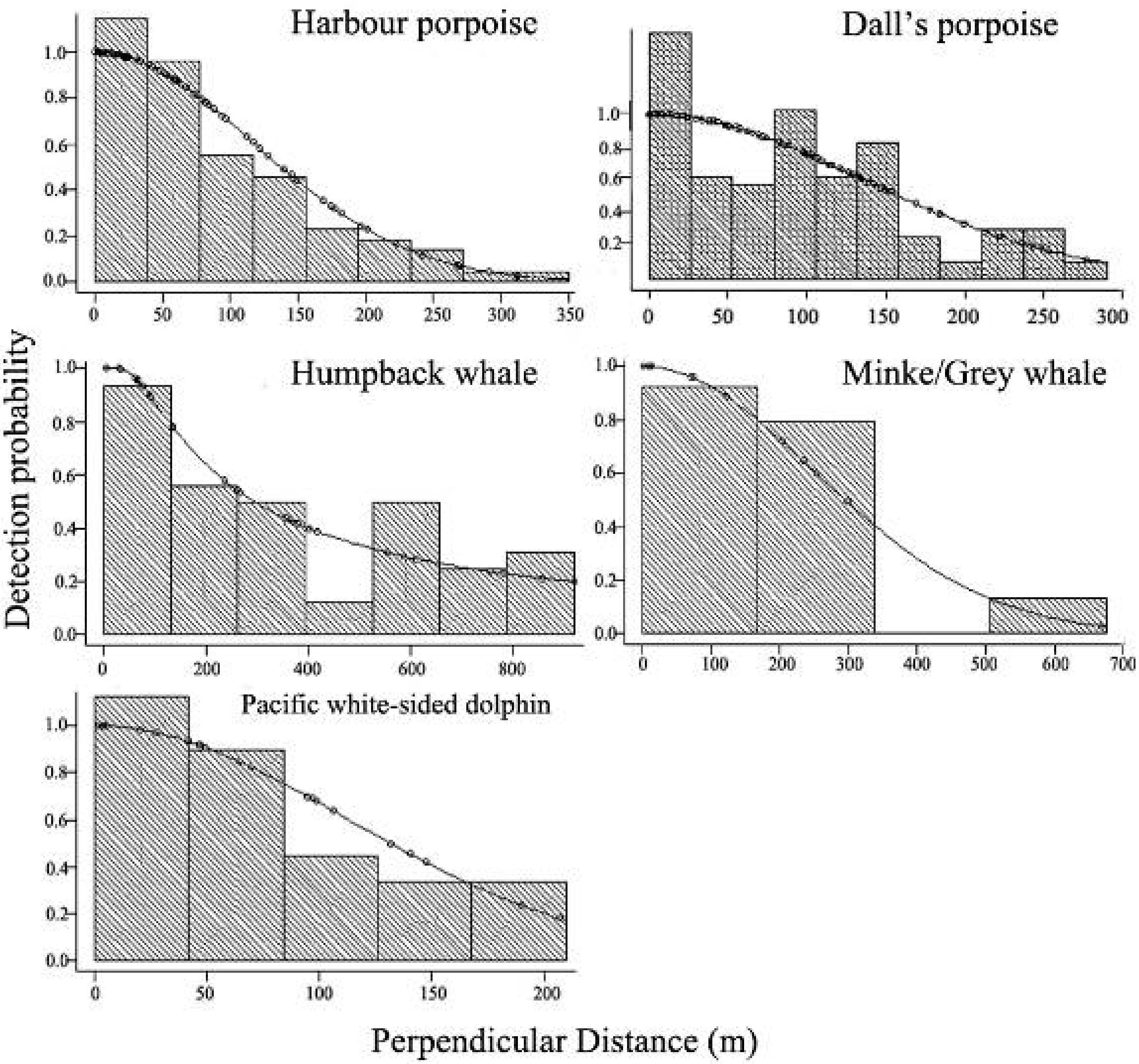


Figure 13: Histograms of observed data and fitted detection functions for all six species pooled over all three survey regions. Note the different distance scale for each species. The plot of the minke/grey whale detection function represents the pooled minke and grey whale sightings because of the small sample size of the latter.

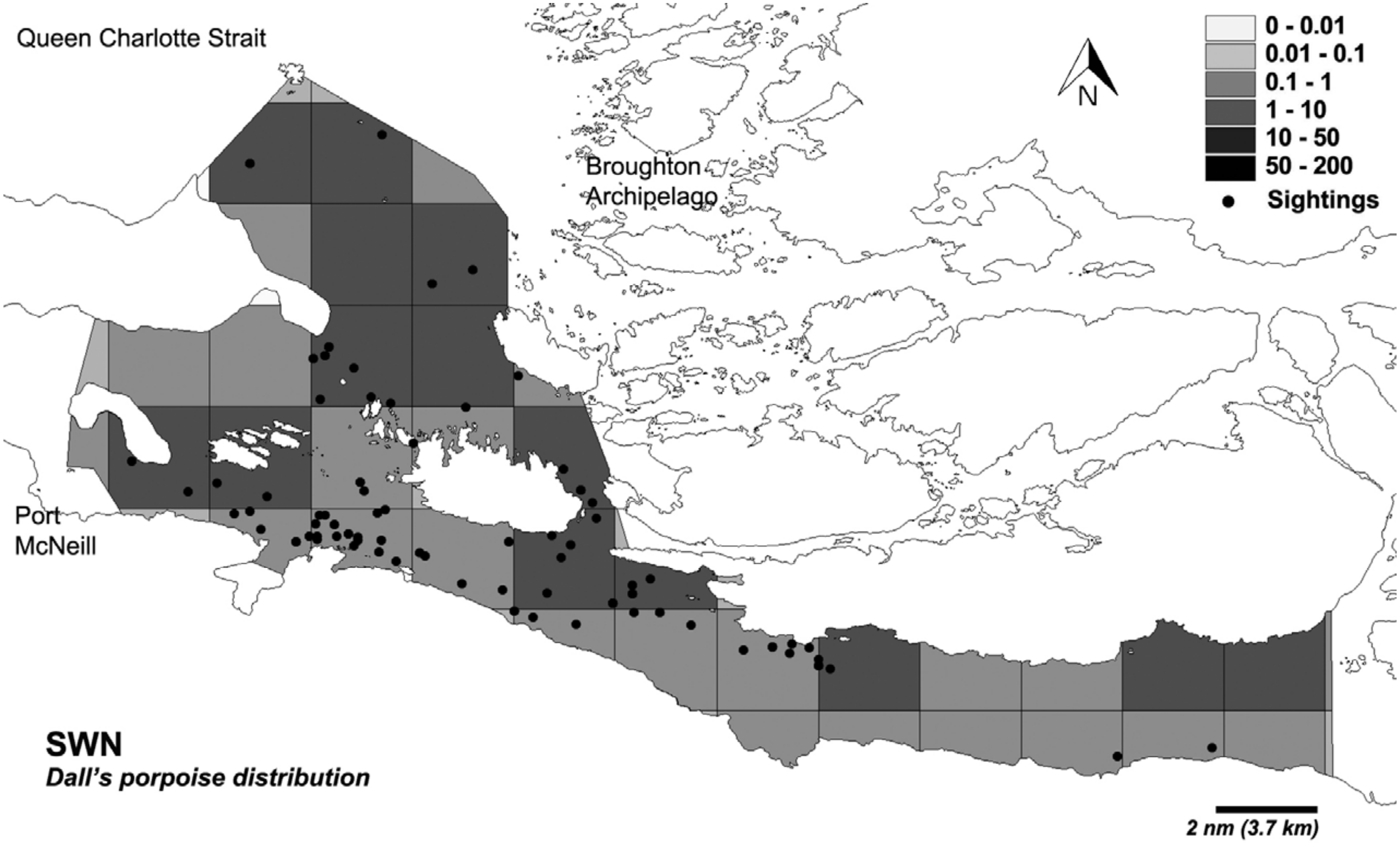


Figure 14: Mean estimate of distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWN survey region with locations of sightings overlaid (solid black dots). During 2008 & 2009, 106 sightings of 300 individuals were made.

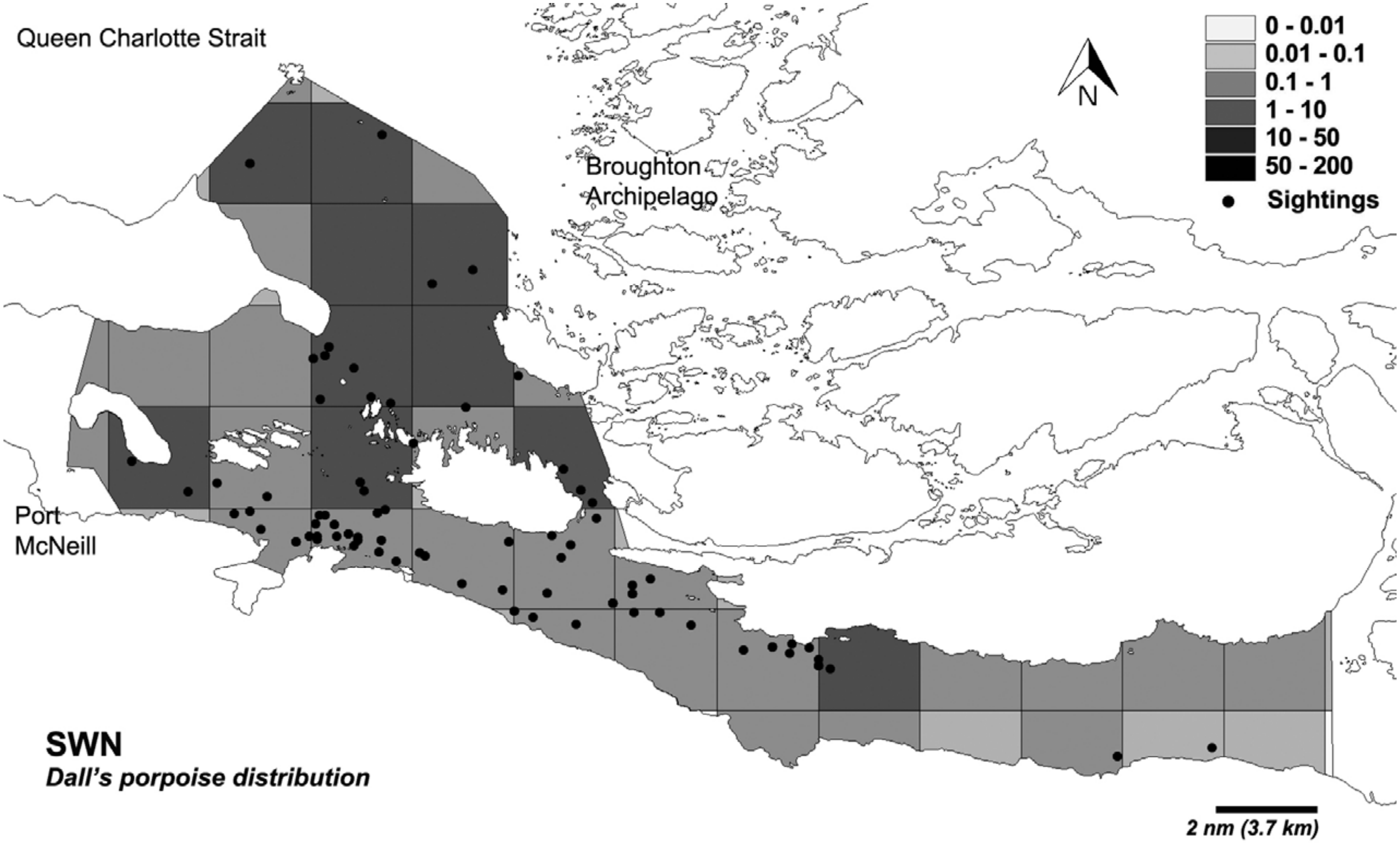


Figure 15: Lower bound (2.5th percentile) estimate of the distribution of abundance (shadings, in number of animals) for Dall's porpoise in the SWN survey region with locations of sightings overlaid (solid black dots).

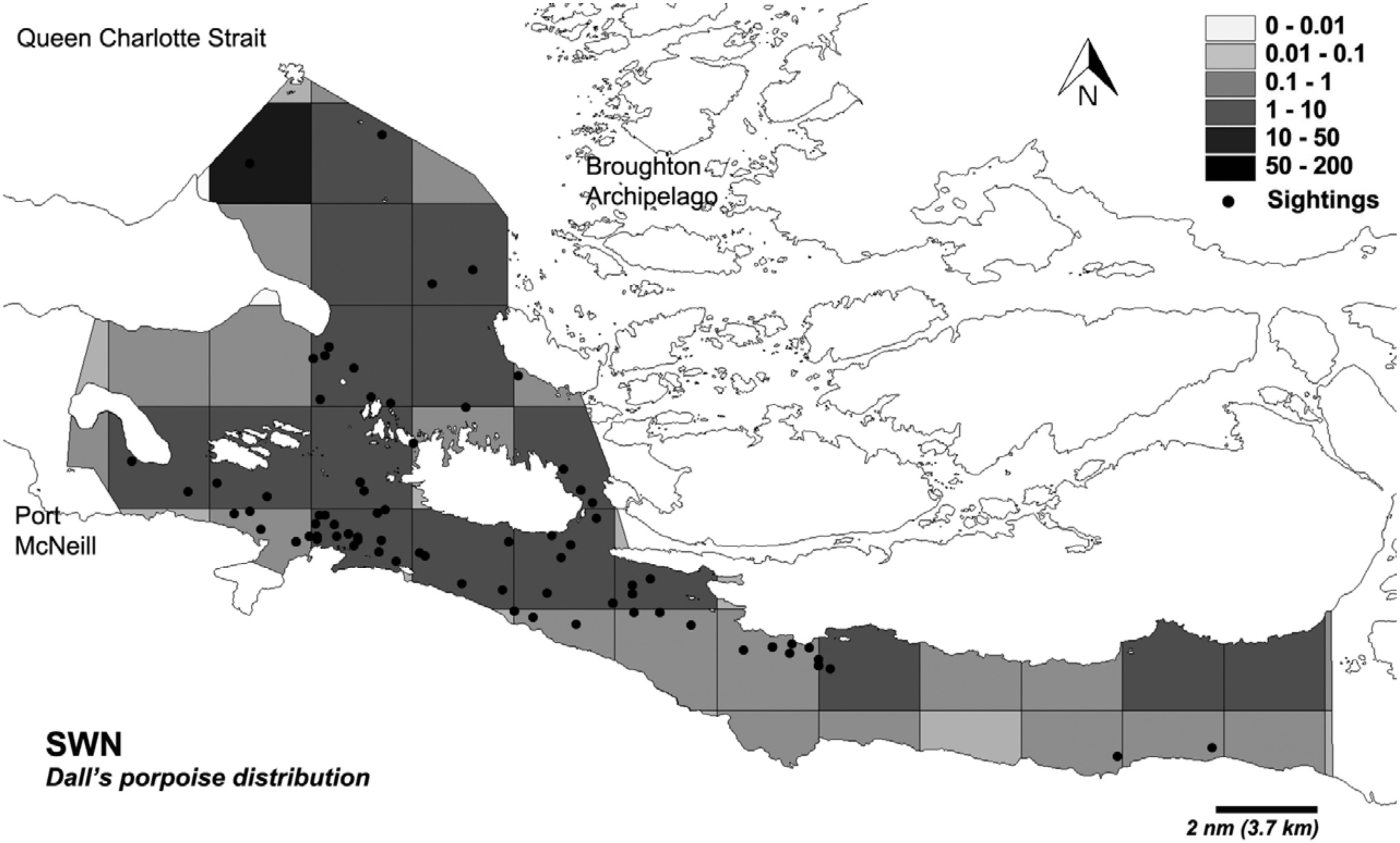


Figure 16: Upper bound (97.5th percentile) estimate of the distribution of abundance (shadings, in number of animals) for Dall's porpoise in the SWN survey region with locations of sightings overlaid (solid black dots).

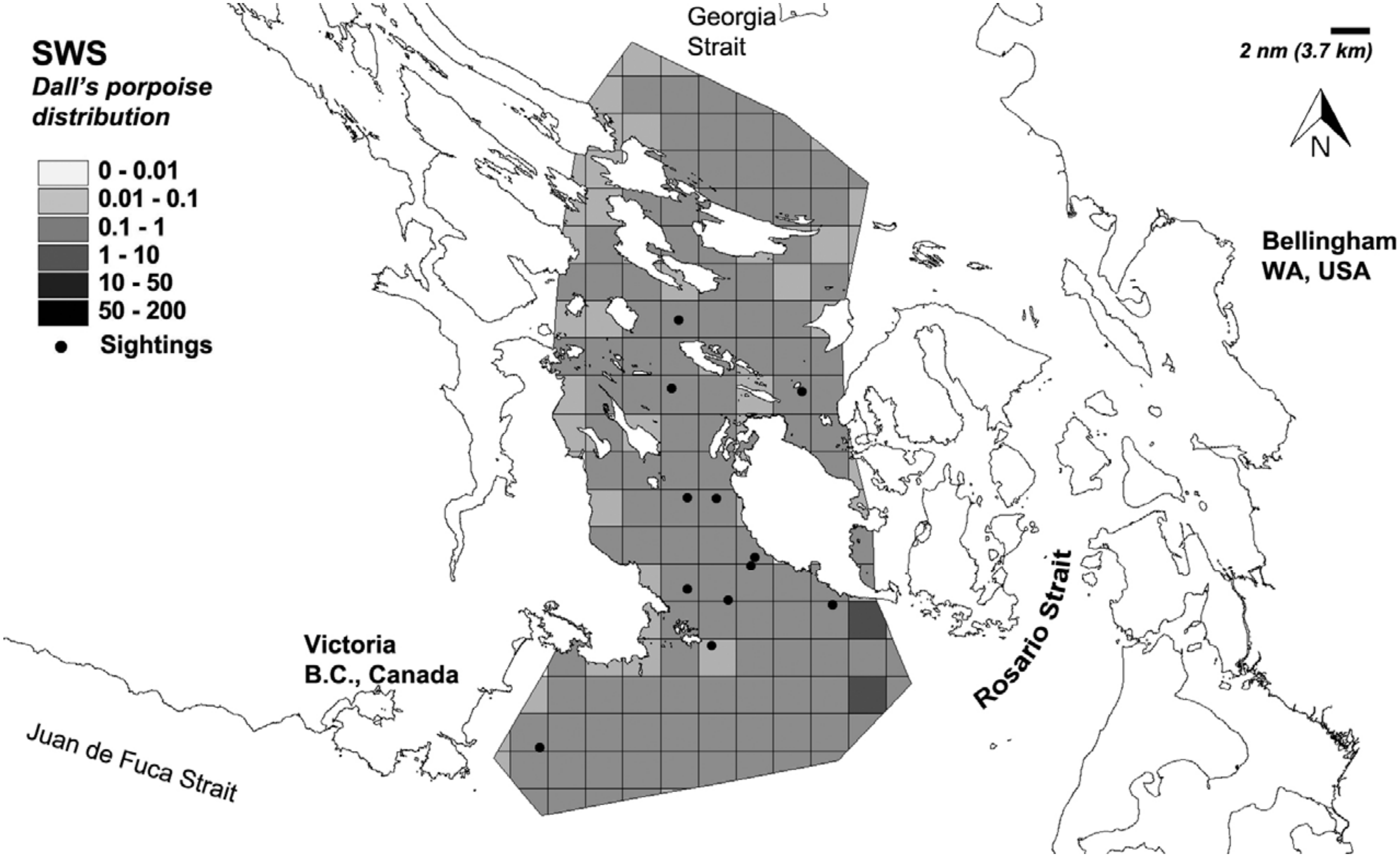


Figure 17: Mean estimate of distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWS survey region, with locations of sightings overlaid (solid black dots). During 2009, 12 sightings of 34 individuals were made.

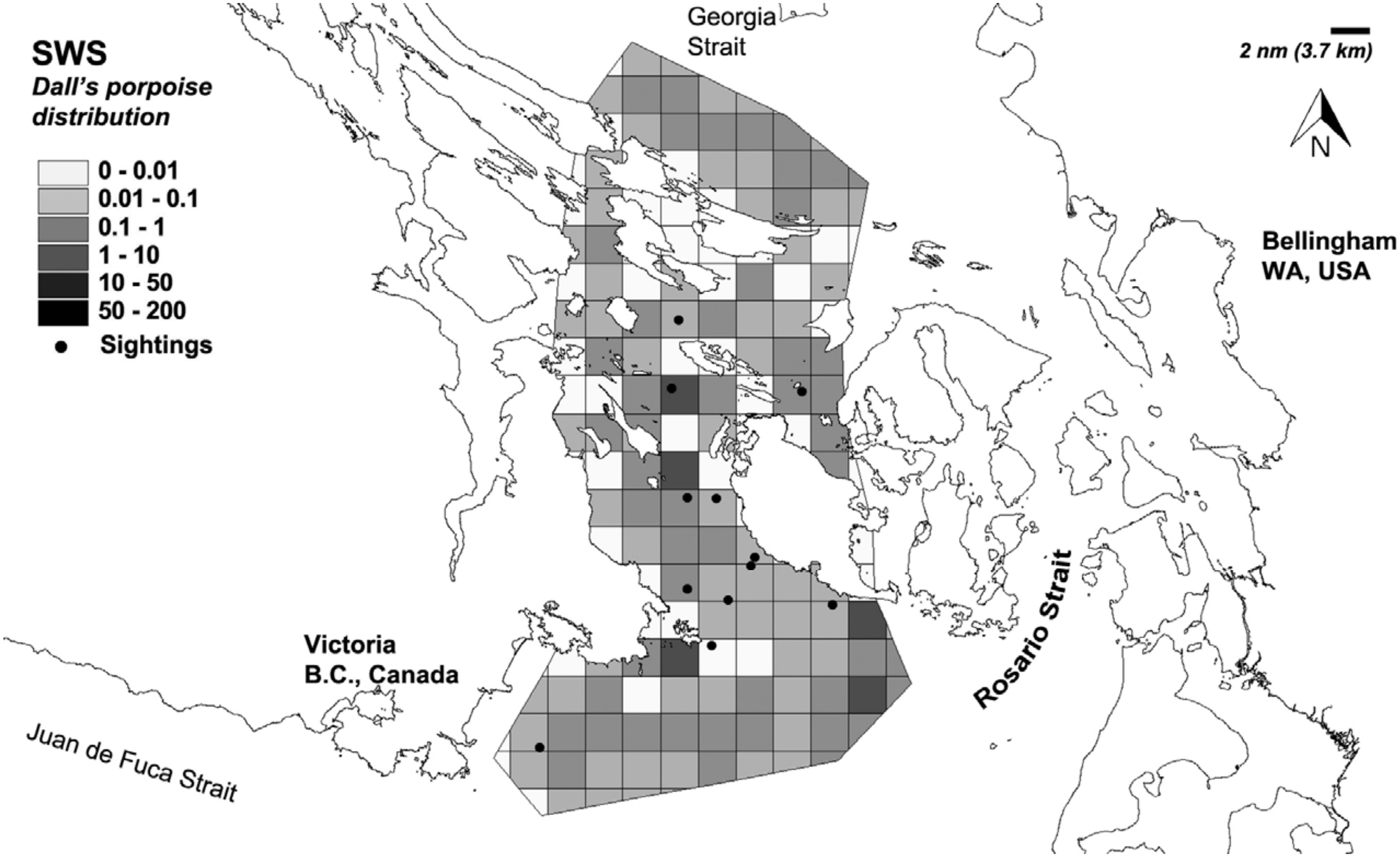


Figure 18: Lower bound (2.5th percentile) estimate of the distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWS survey region with locations of sightings overlaid (solid black dots).

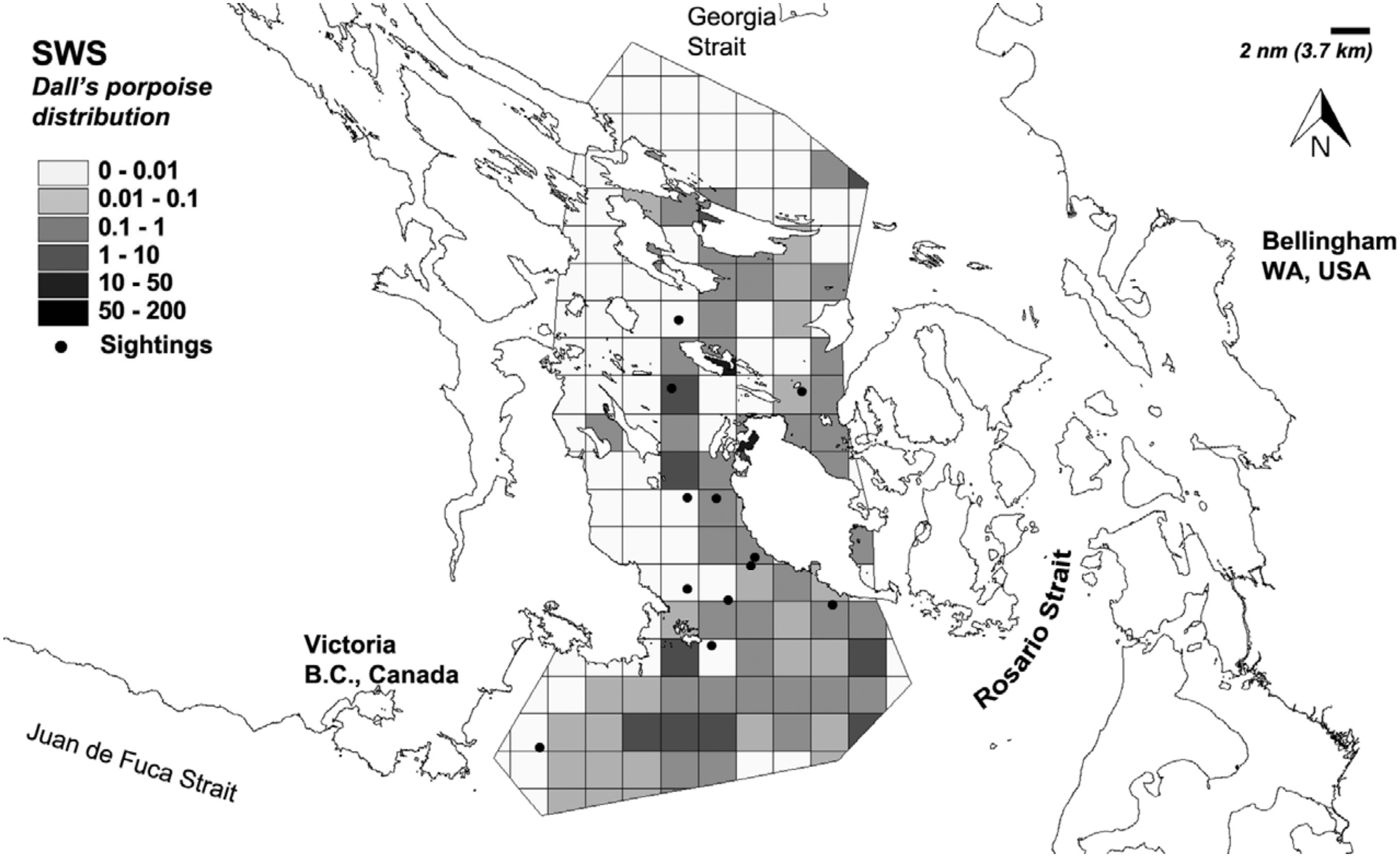


Figure 19: Upper bound (97.5th percentile) estimate of the distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWS survey region with locations of sightings overlaid (solid black dots).

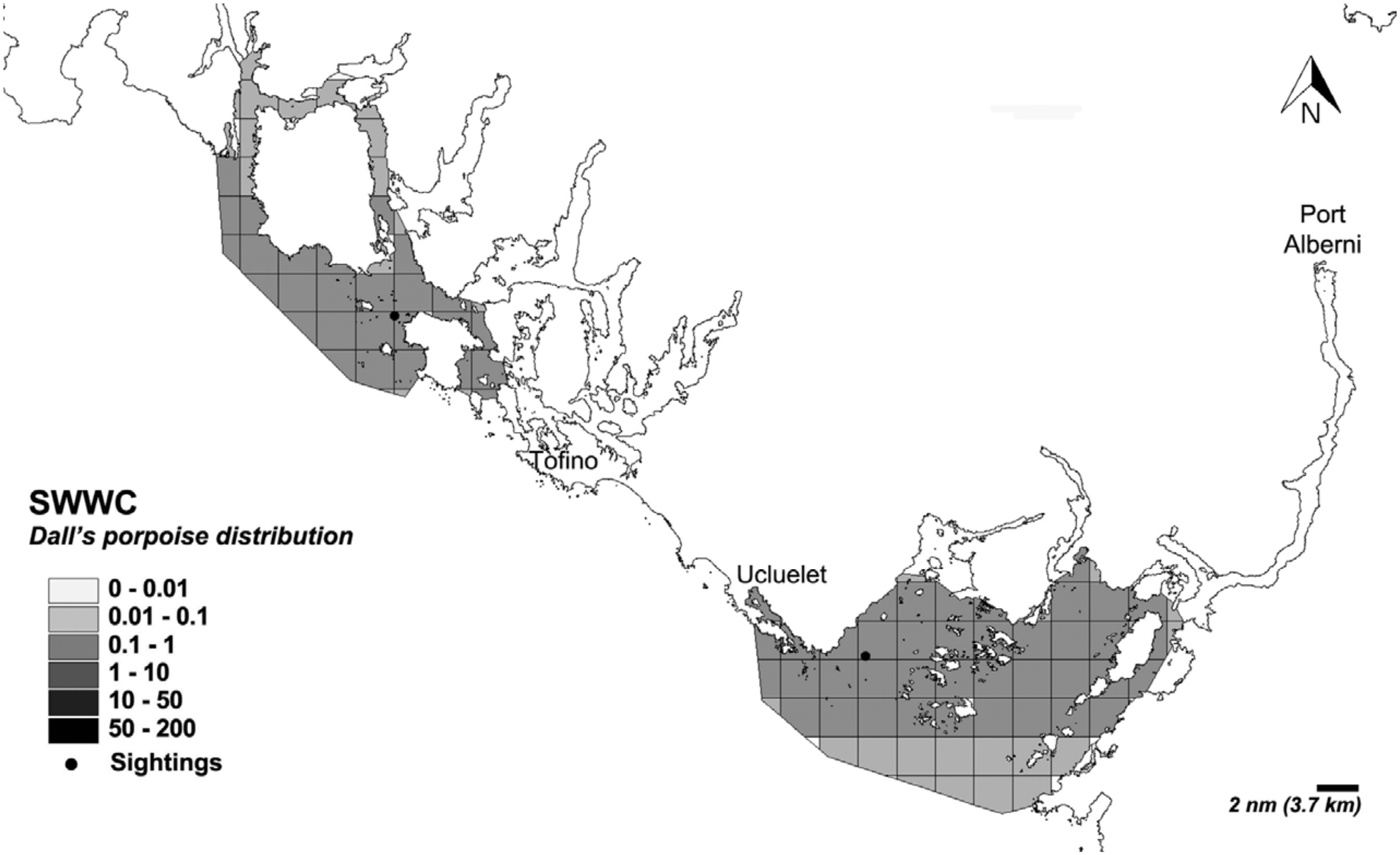


Figure 20: Mean estimate of distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWWC survey region, with locations of sightings overlaid (solid black dots). During 2009, 2 sightings of 4 individuals were made.

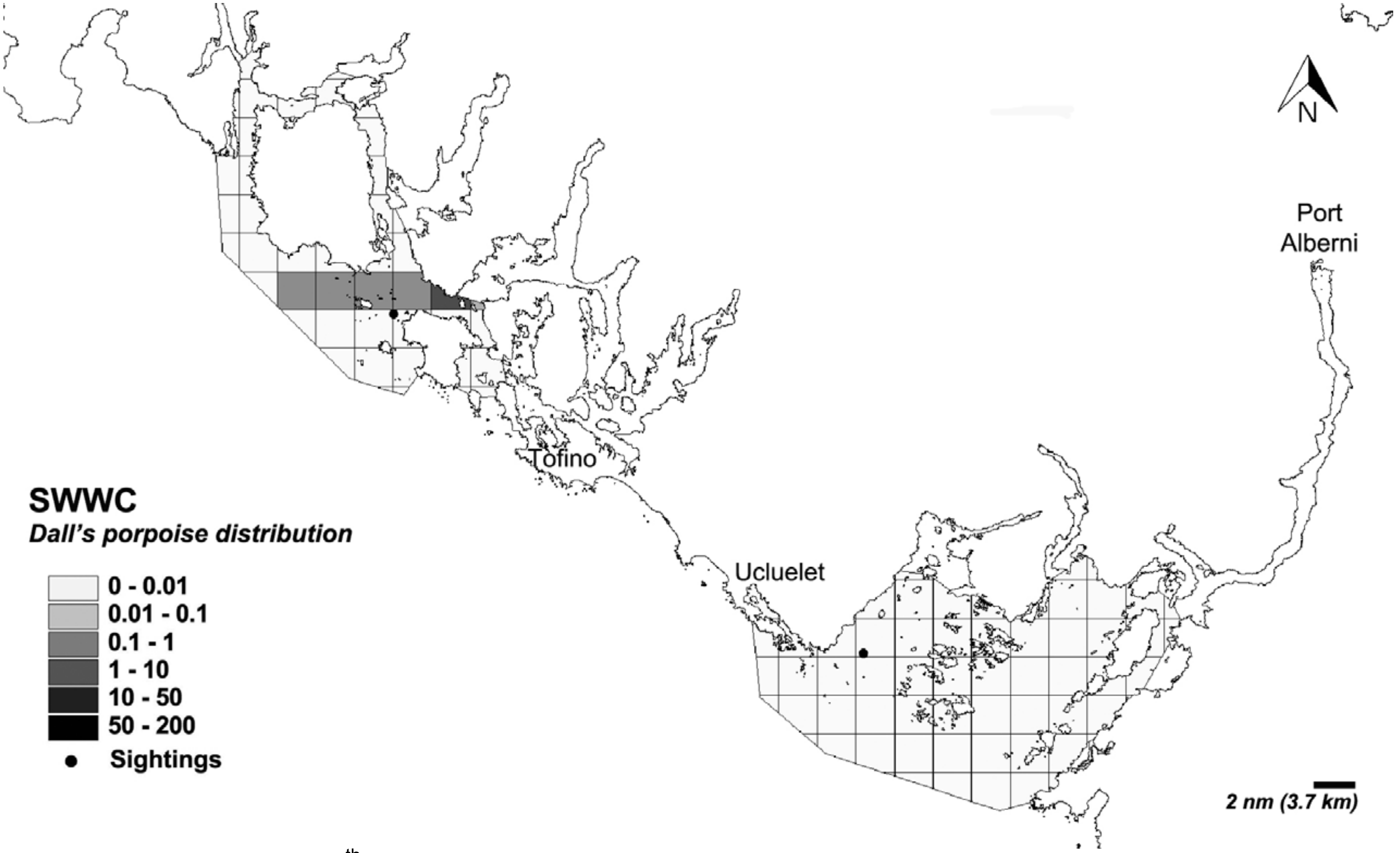


Figure 21: Lower bound (2.5th percentile) estimate of the distribution of abundance (shadings, in numbers of animals) for Dall's porpoise in the SWWC survey region, with locations of sightings overlaid (solid black dots).

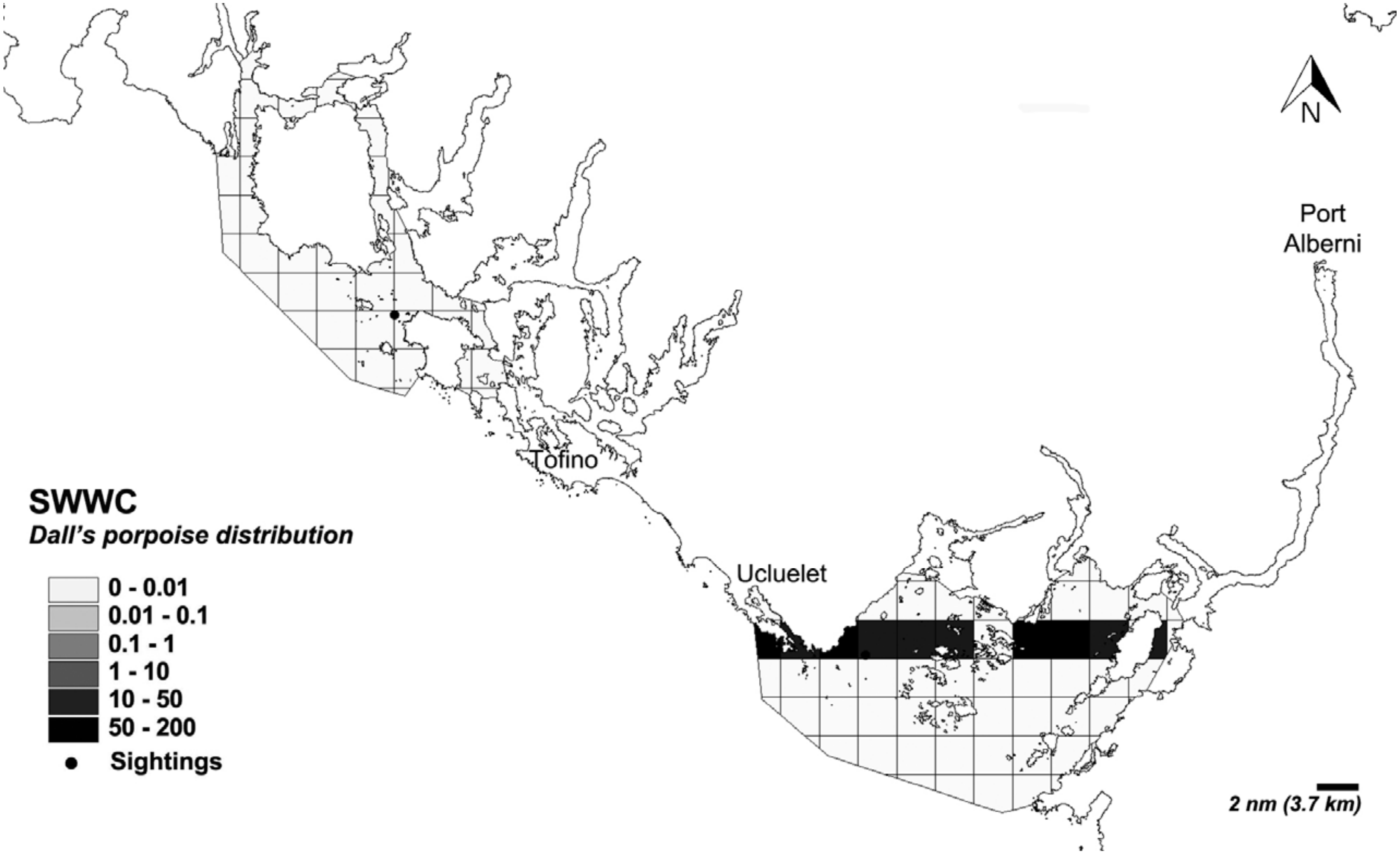


Figure 22: Upper bound (97.5th percentile) estimate of the distribution of abundance (shadings, in the numbers of animals) for Dall's porpoise in the SWWC survey region with locations of sightings overlaid (solid black dots).

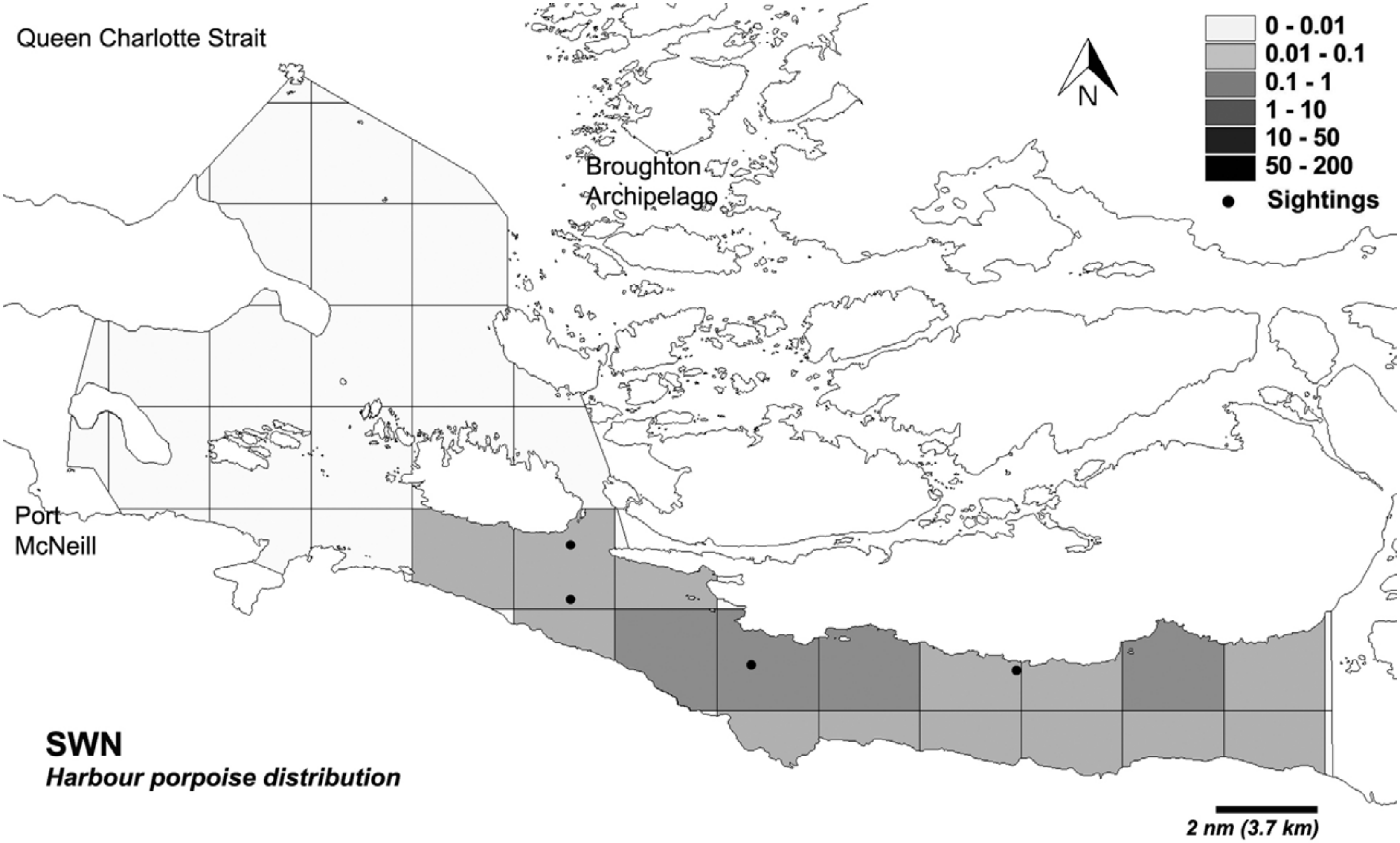


Figure 23: Mean estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWN survey region with locations of sightings overlaid (solid black dots). During 2008 & 2009, 3 sightings of 4 individuals were made.

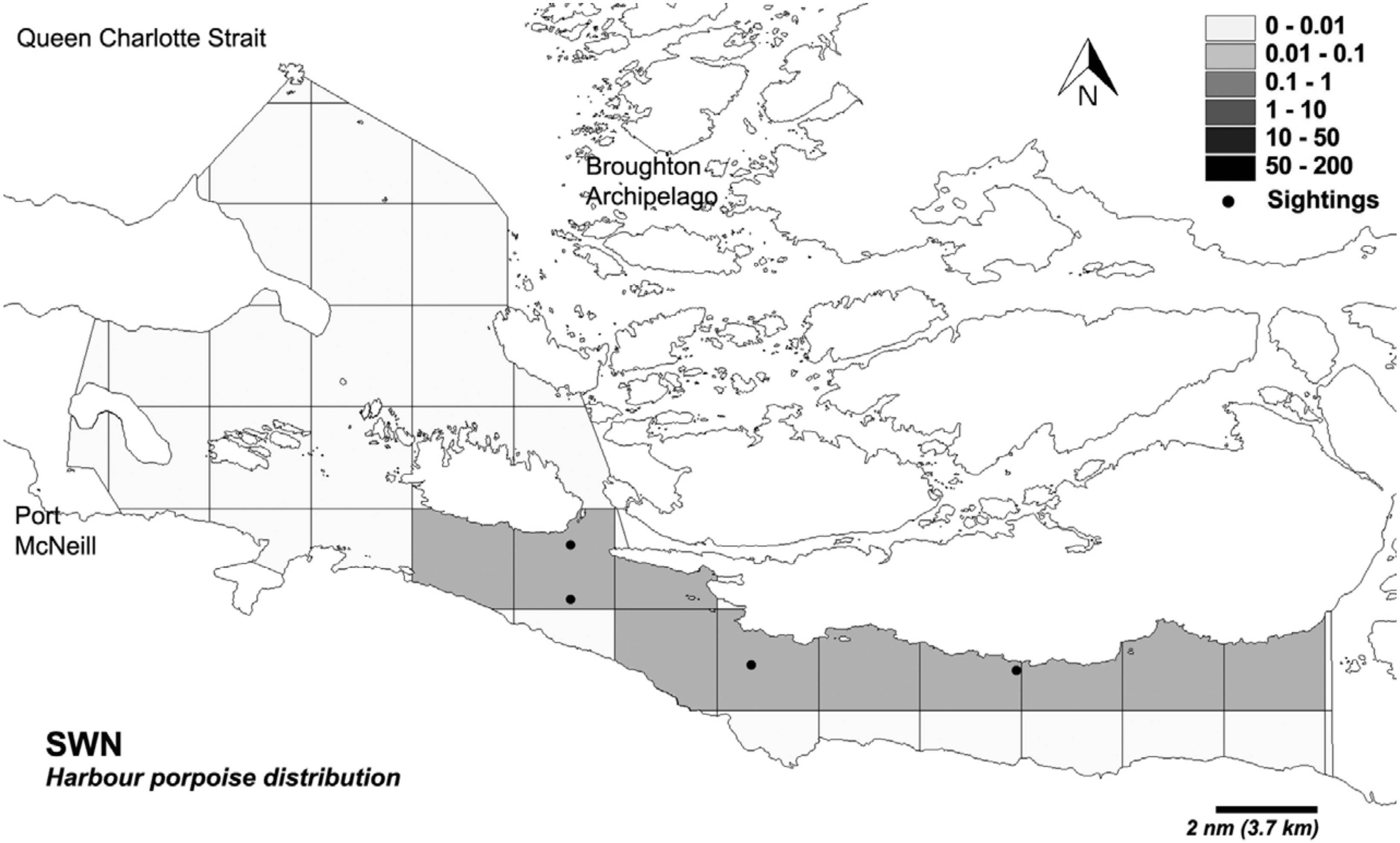


Figure 24: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in number of animals) of harbour porpoise in the SWN survey region with locations of sightings overlaid (solid black dots).

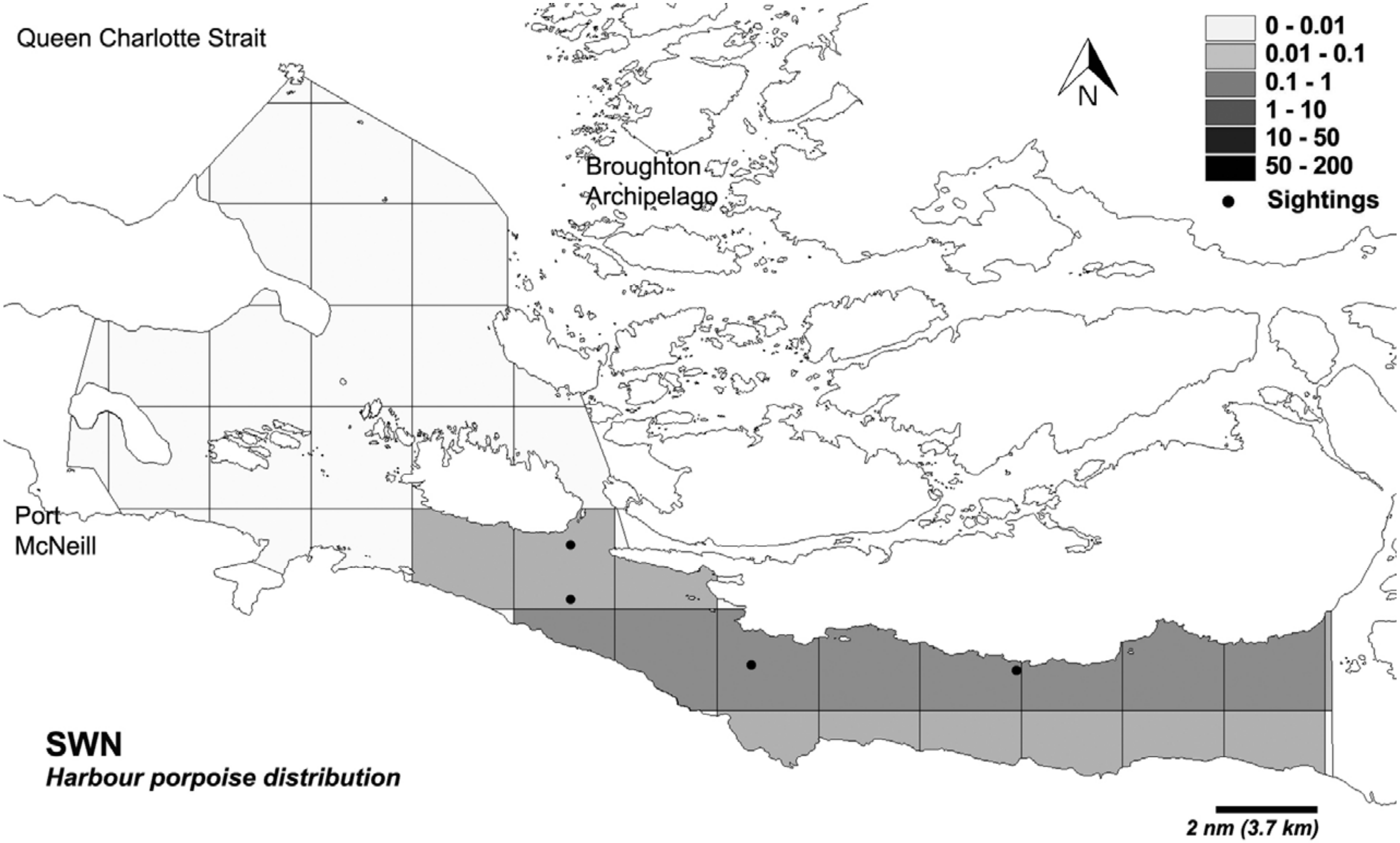


Figure 25: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWN survey region with locations of sightings overlaid (solid black dots).

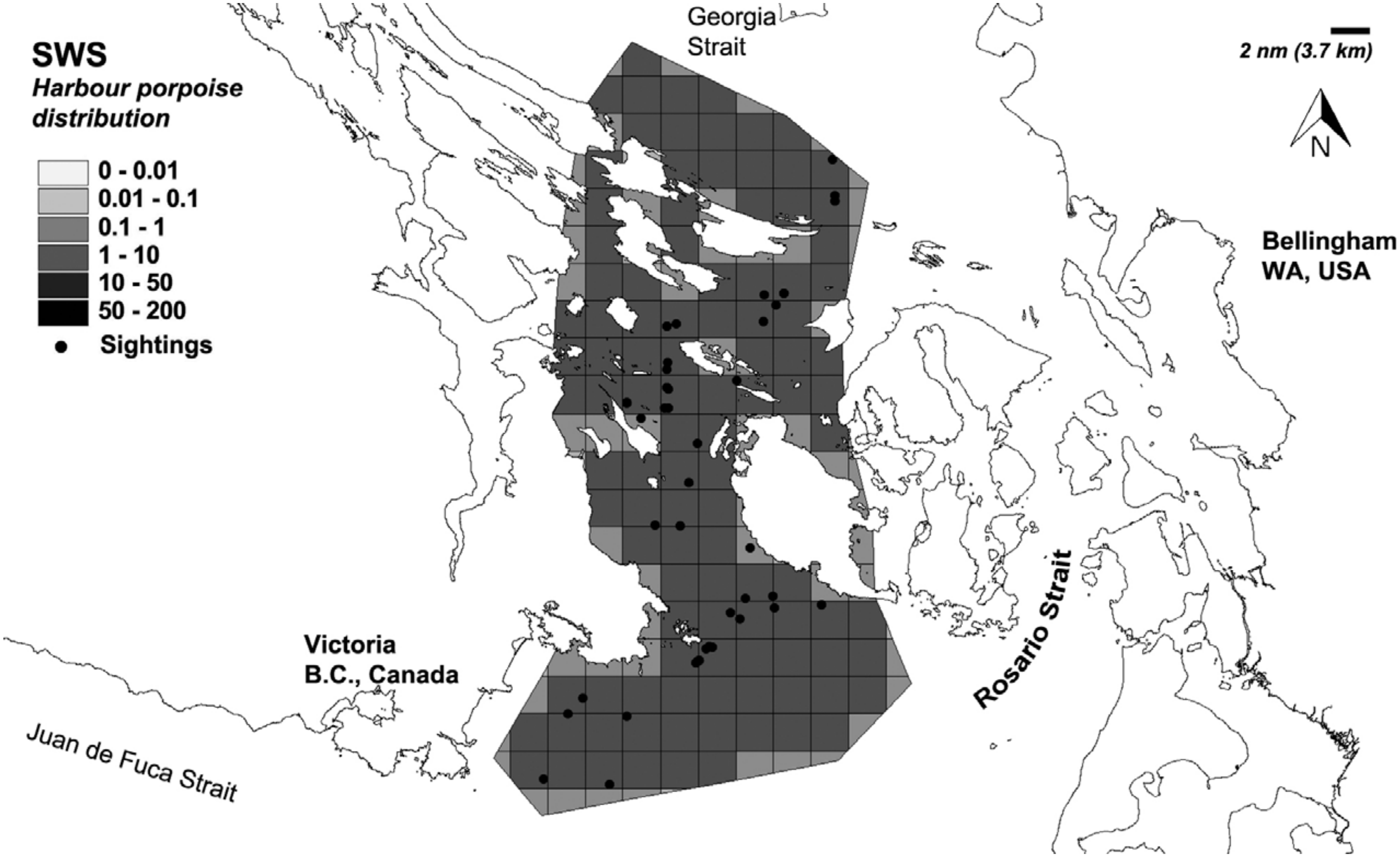


Figure 26: Mean estimate of distribution of abundance (shadings, in numbers of animals) for harbour porpoise in the SWS survey region with locations of sightings overlaid (solid black dots). During 2009, 54 sightings of 94 individuals were made.

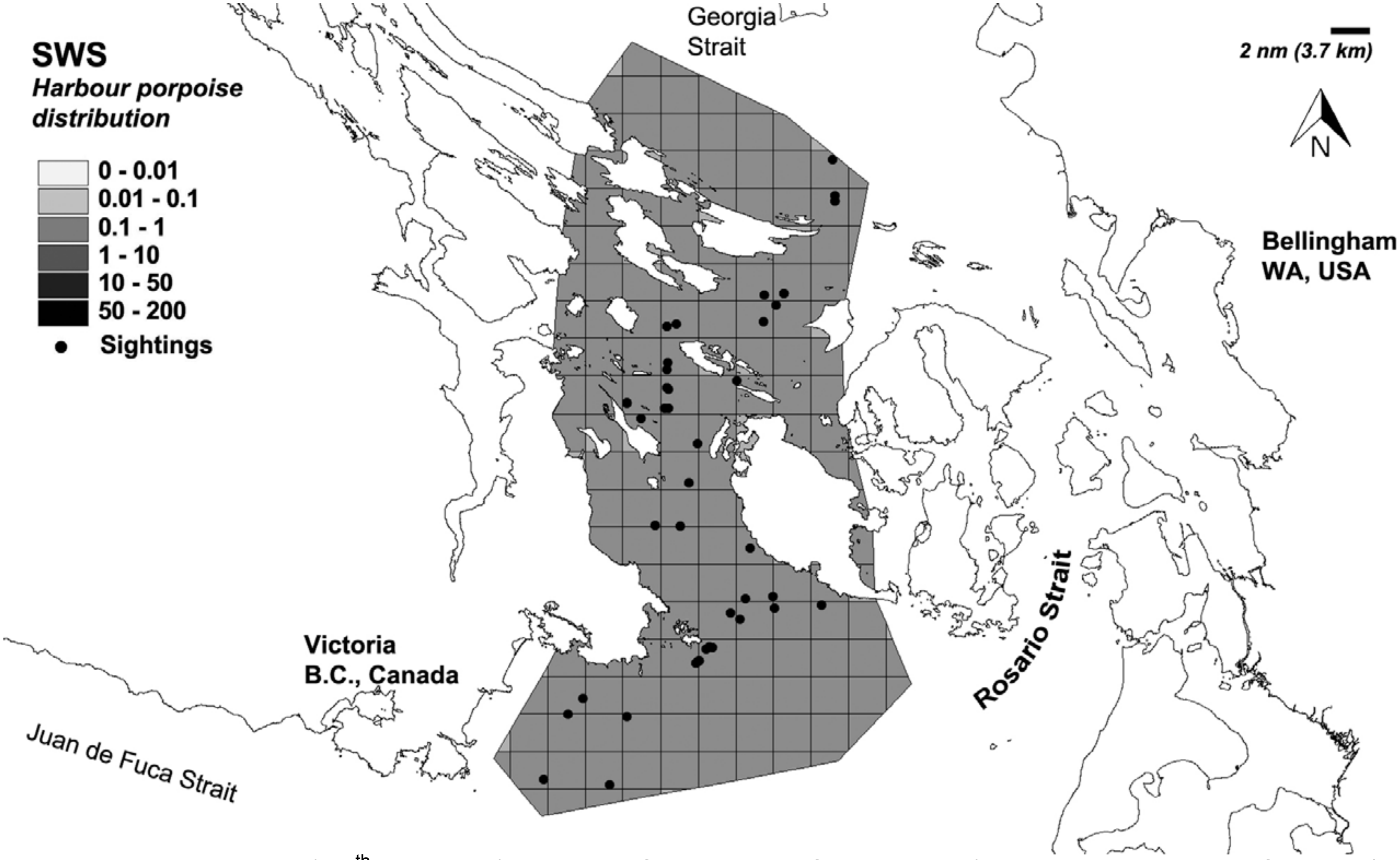


Figure 27: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWS survey region with locations of sightings overlaid (solid black dots).

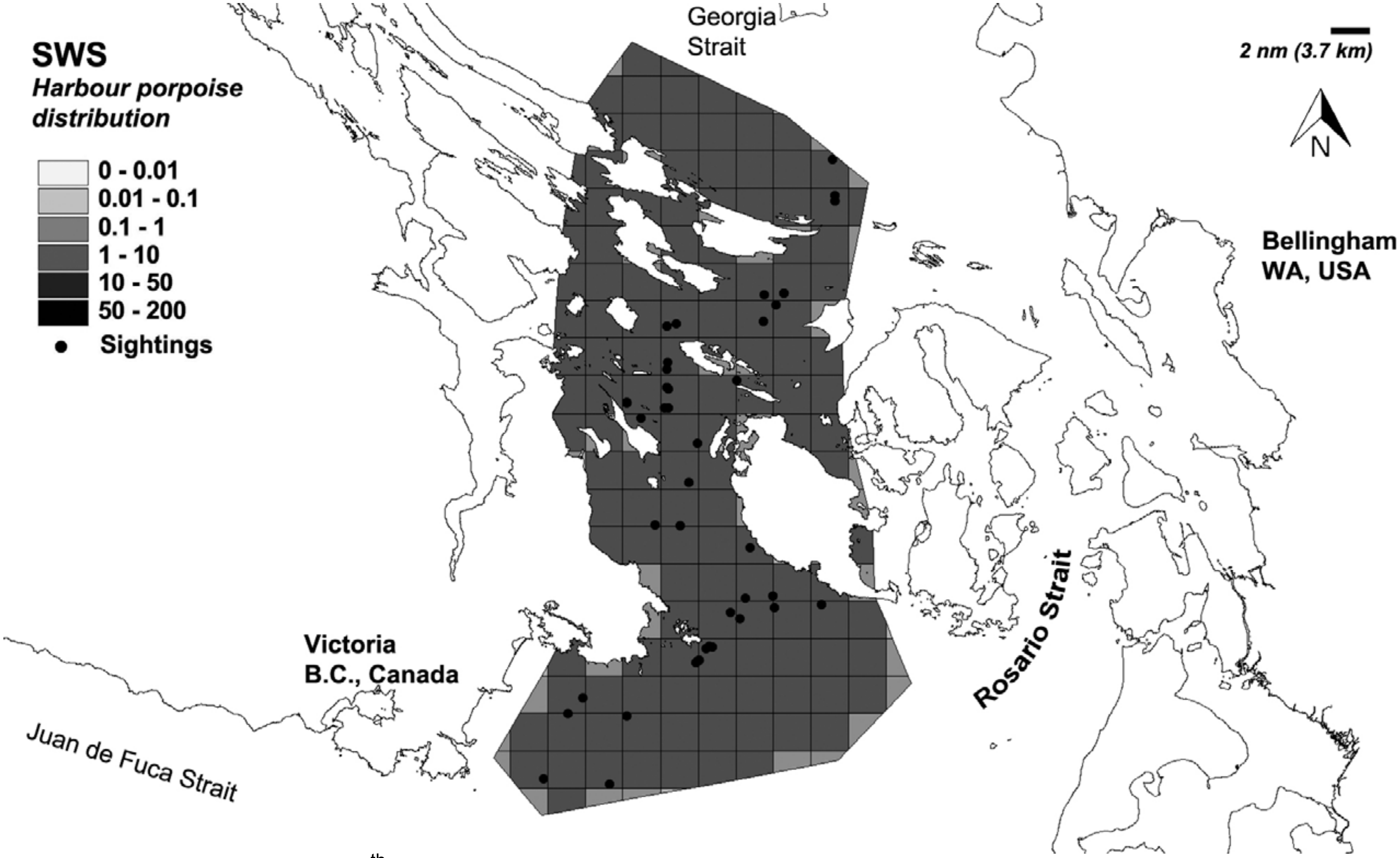


Figure 28: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWS survey region with locations of sightings overlaid (solid black dots).

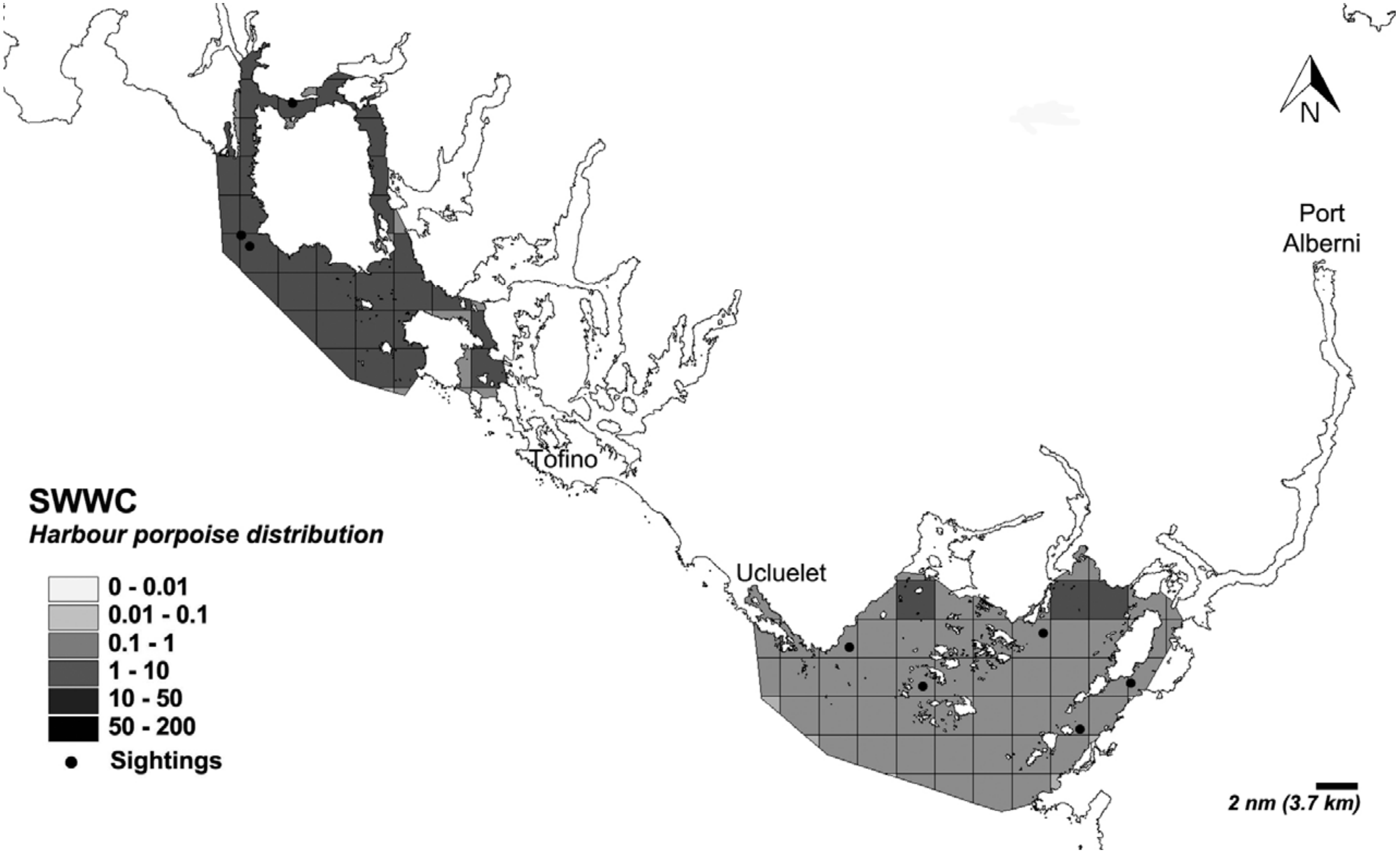


Figure 29: Mean estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWWC survey region with locations of sightings overlaid (solid black dots). During 2009, 11 sightings of 25 individuals were made.

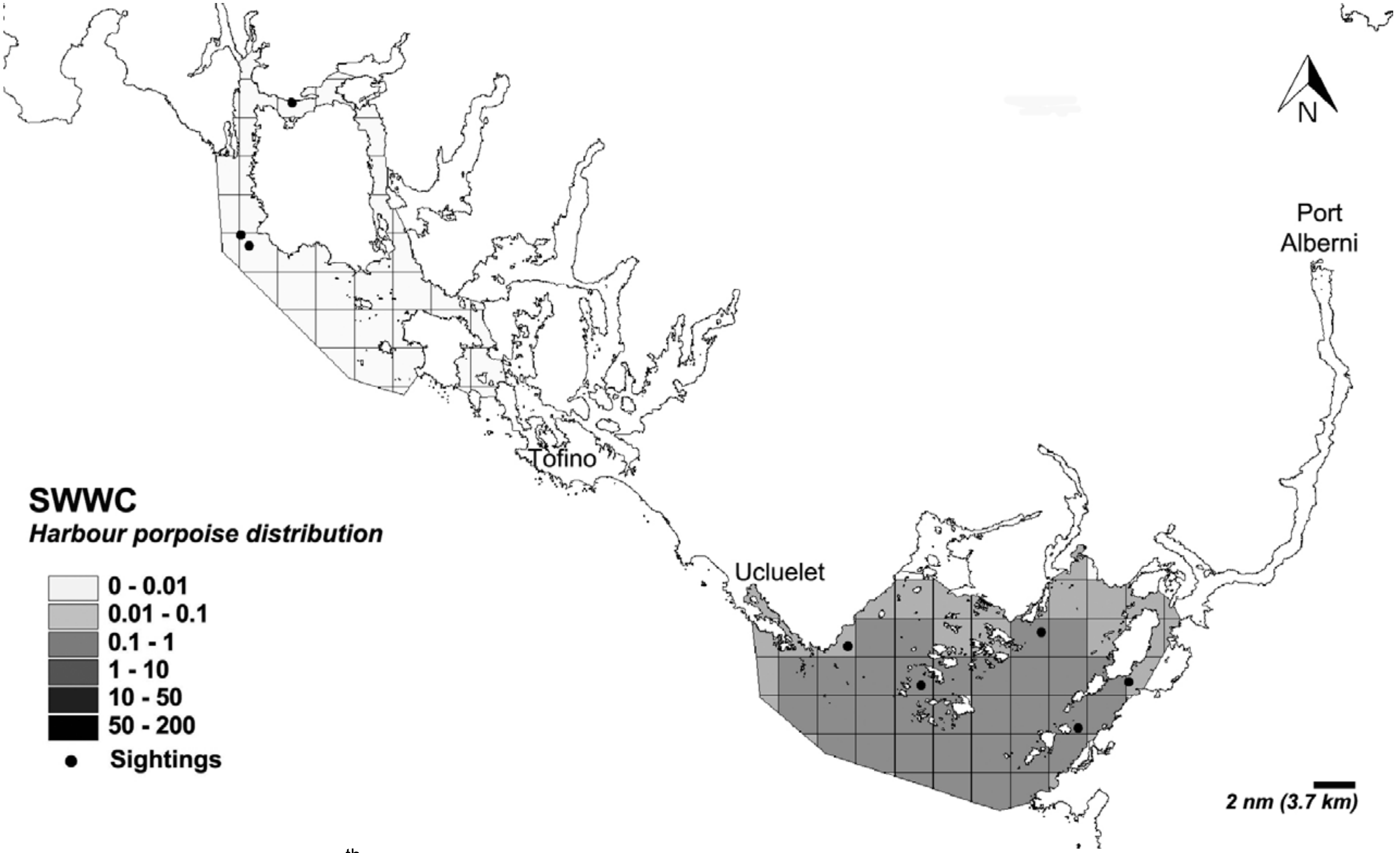


Figure 30: Lower bound (2.5th percentile) estimate of the distribution of abundance (shadings, in numbers of animals) for harbour porpoise in the SWWC survey region with locations of sightings overlaid (solid black dots).

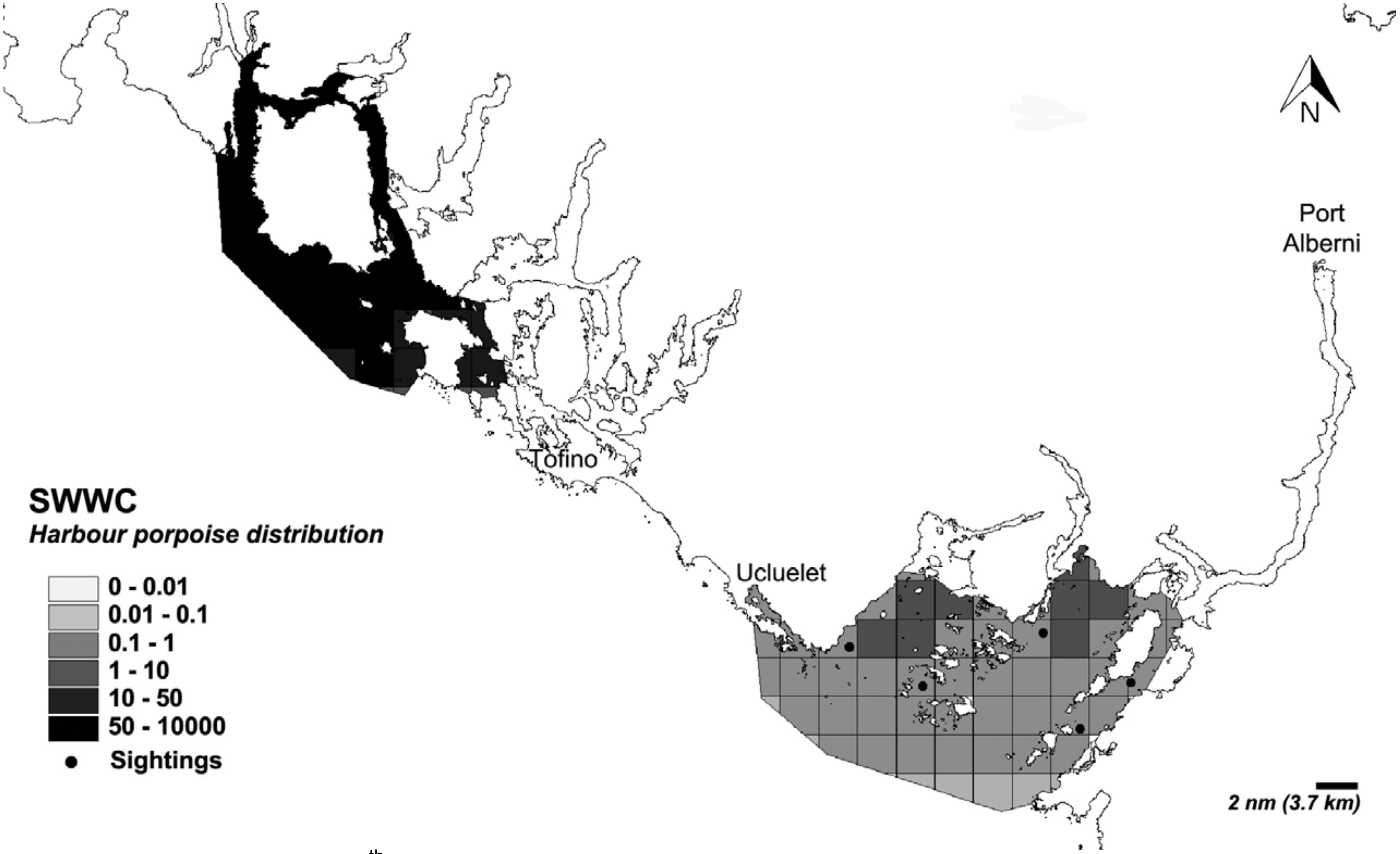


Figure 31: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) of harbour porpoise in the SWS survey region with locations of sightings overlaid (solid black dots).

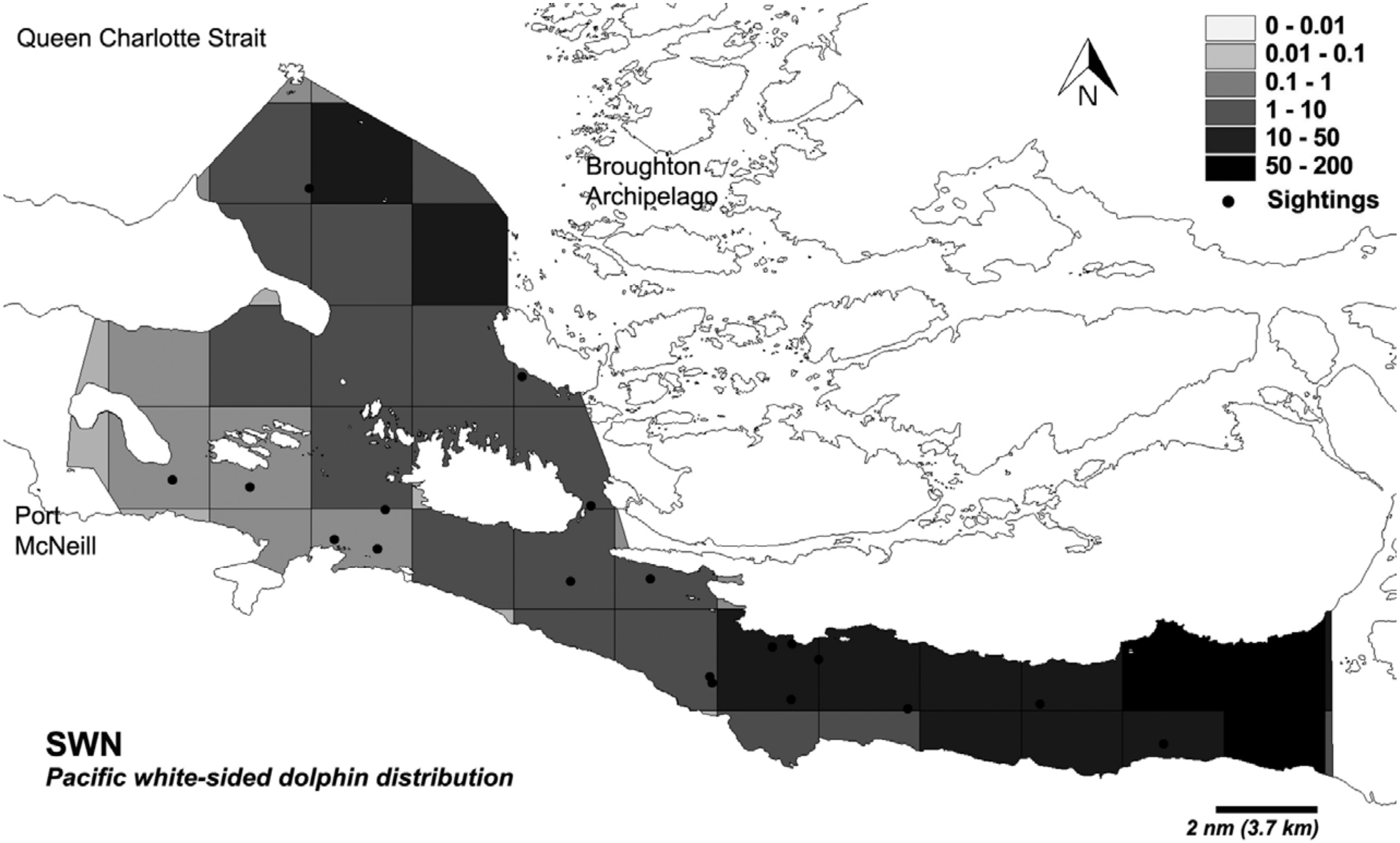


Figure 32: Mean estimate of distribution of abundance (shadings, in numbers of animals) for Pacific white-sided dolphin in the SWN survey region with locations of sightings overlaid (solid black dots). During 2008 & 2009, 30 sightings of 1,478 individuals were made.

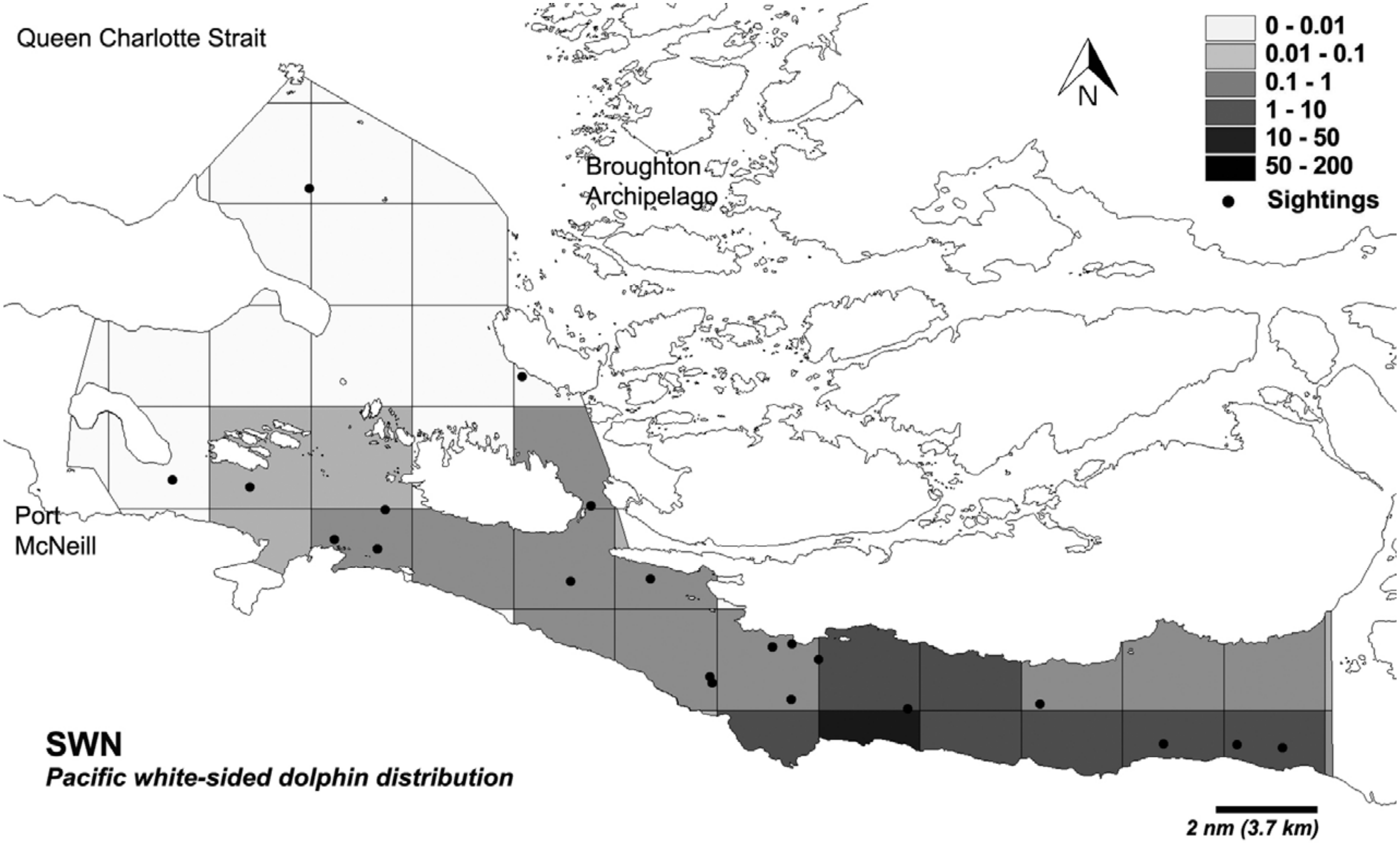


Figure 33: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for Pacific white-sided dolphin in the SWN survey region with locations of sightings overlaid (solid black dots).

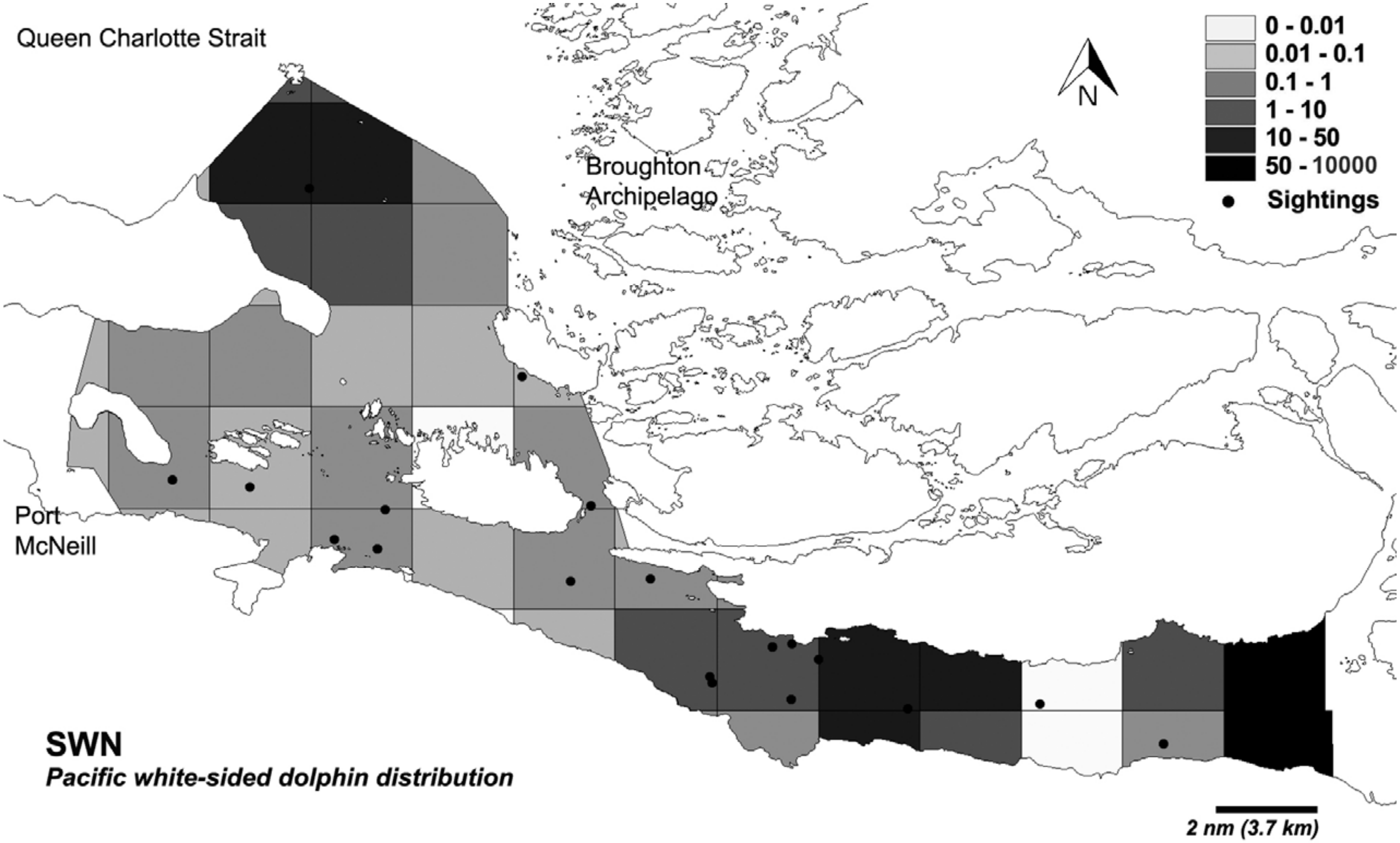


Figure 34: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for Pacific white-sided dolphin in the SWN survey region with locations of sightings overlaid (solid black dots).

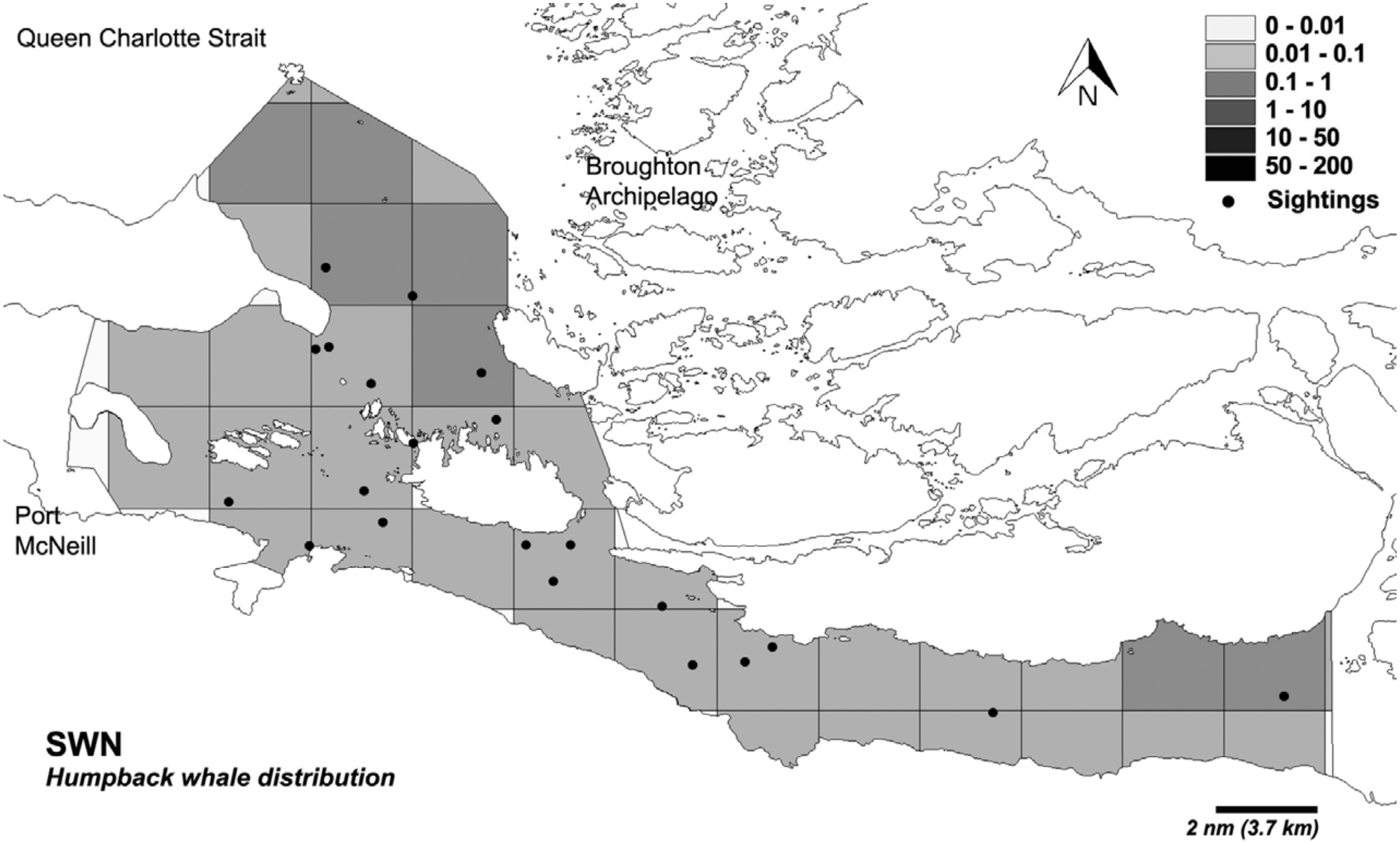


Figure 35: Mean estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWN survey region with locations of sightings overlaid (solid black dots). During 2008 & 2009, 24 sightings of 29 individuals were made.

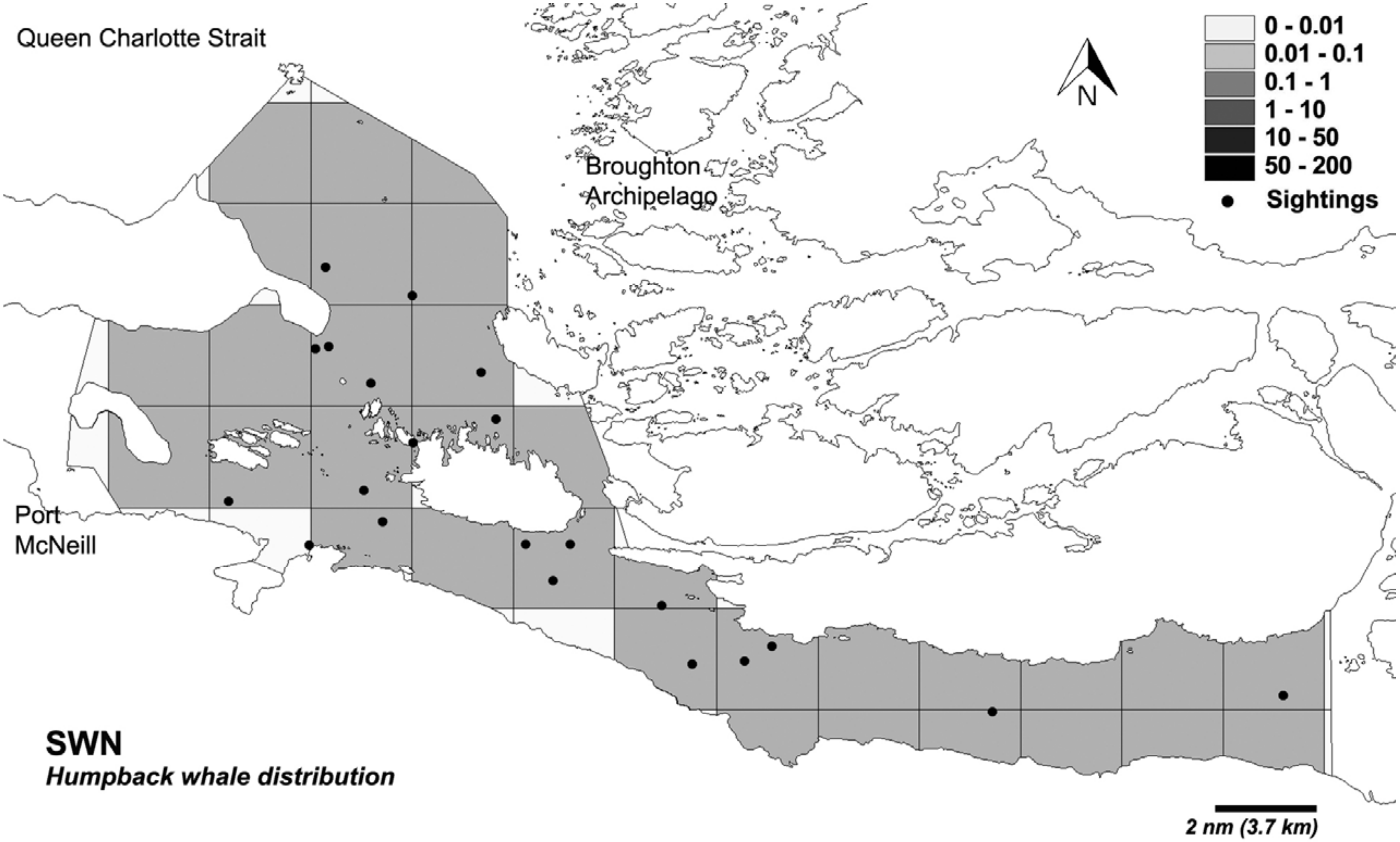


Figure 36: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWN survey region with locations of sightings overlaid (solid black dots).

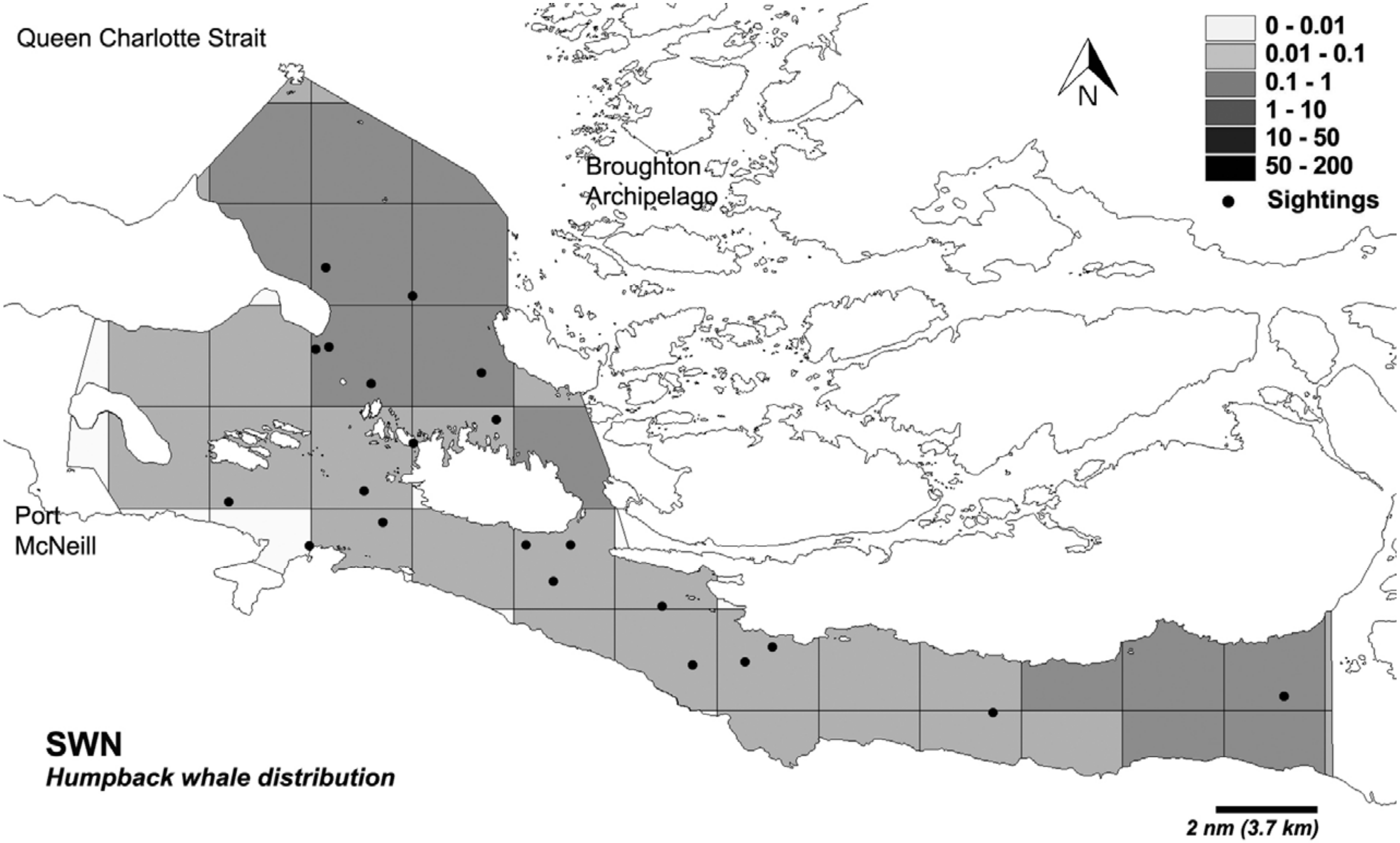


Figure 37: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWN survey region with locations of sightings overlaid (solid black dots).

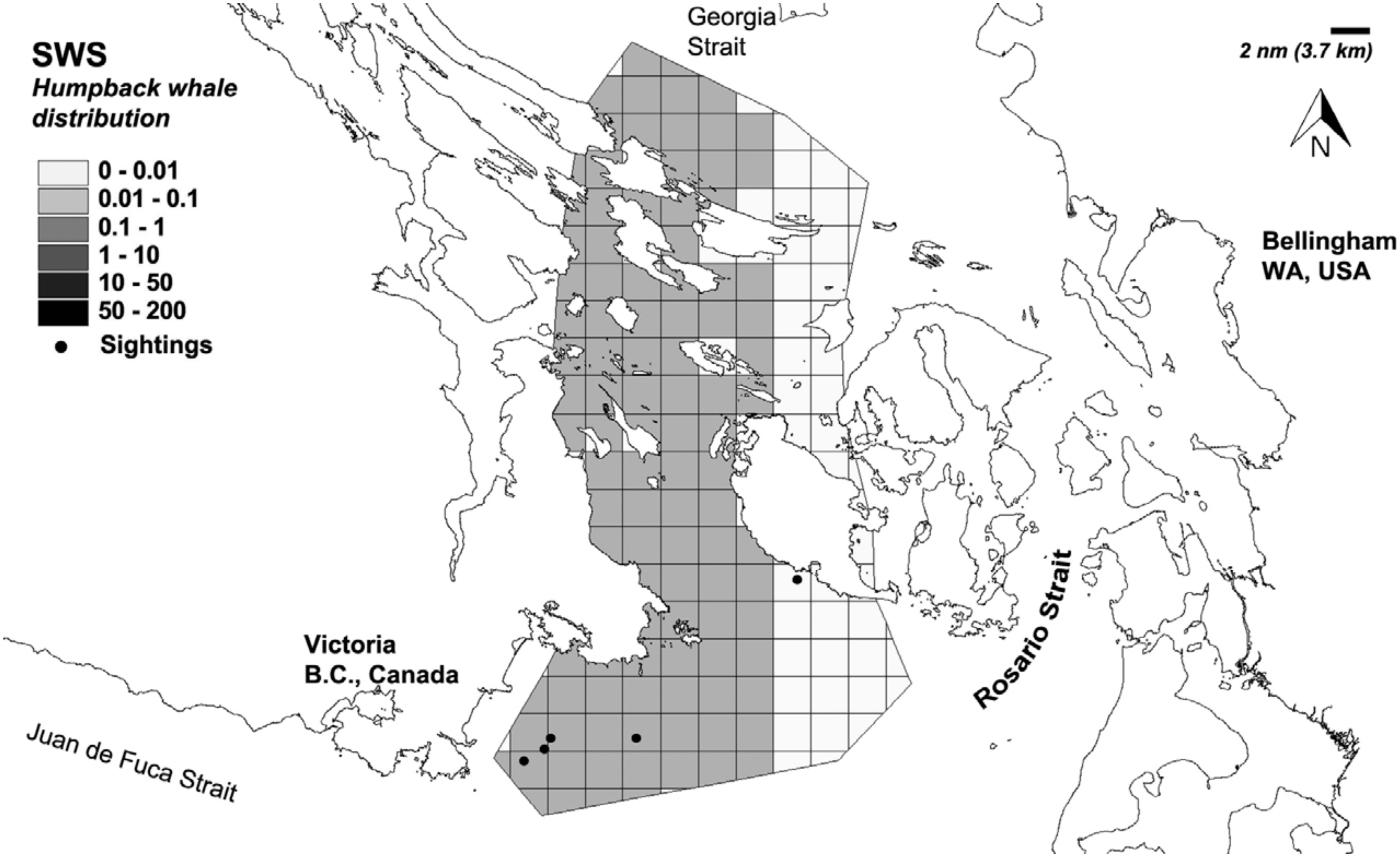


Figure 38: Mean estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWS survey region with locations of sightings overlaid (solid black dots). During 2009, 5 sightings of 10 individuals were made.

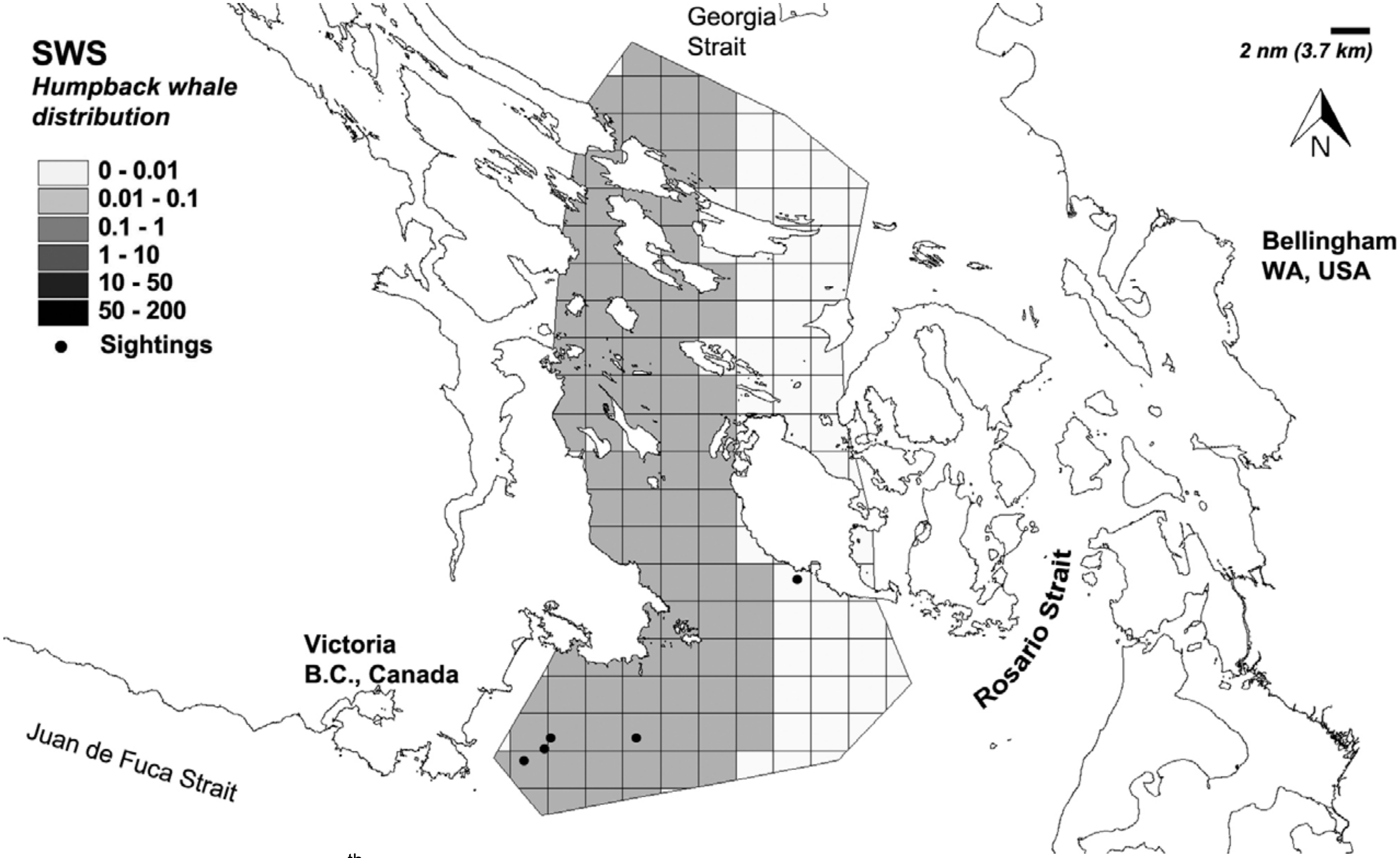


Figure 39: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWS survey region with locations of sightings overlaid (solid black dots).

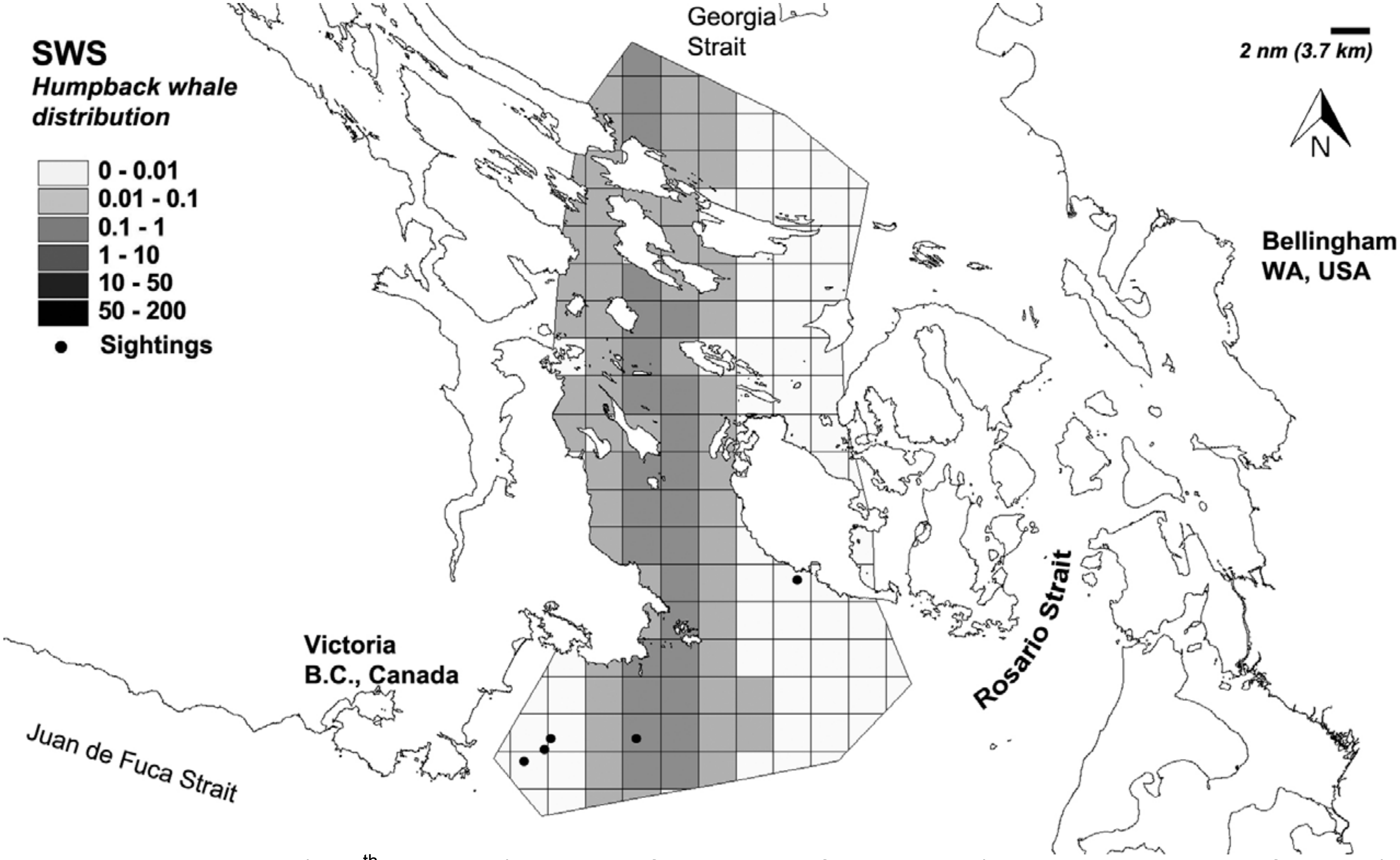


Figure 40: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWS survey region with locations of sightings overlaid (solid black dots).

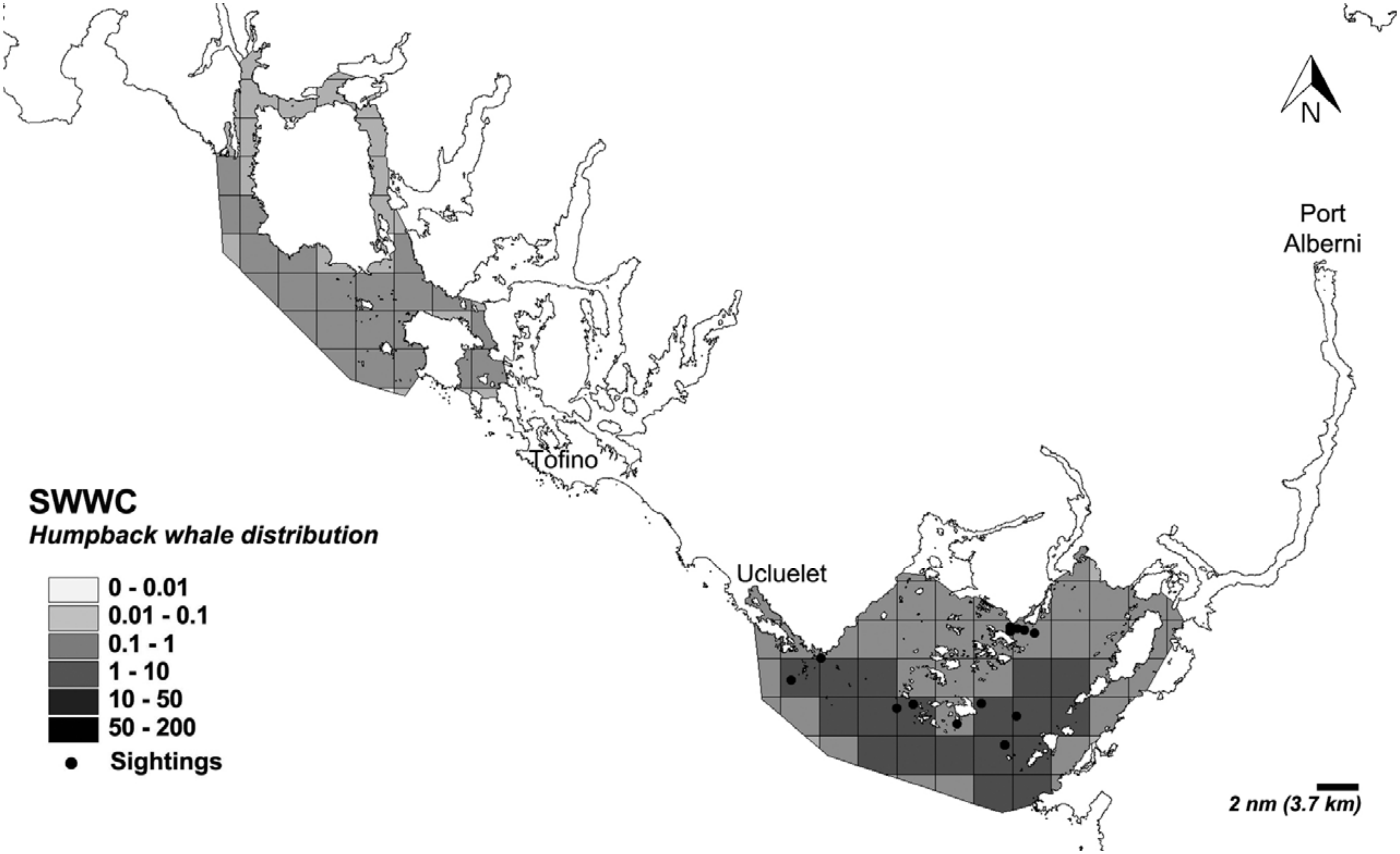


Figure 41: Mean estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWWC survey region with locations of sightings overlaid (solid black dots). During 2009, 22 sightings of 28 individuals were made.

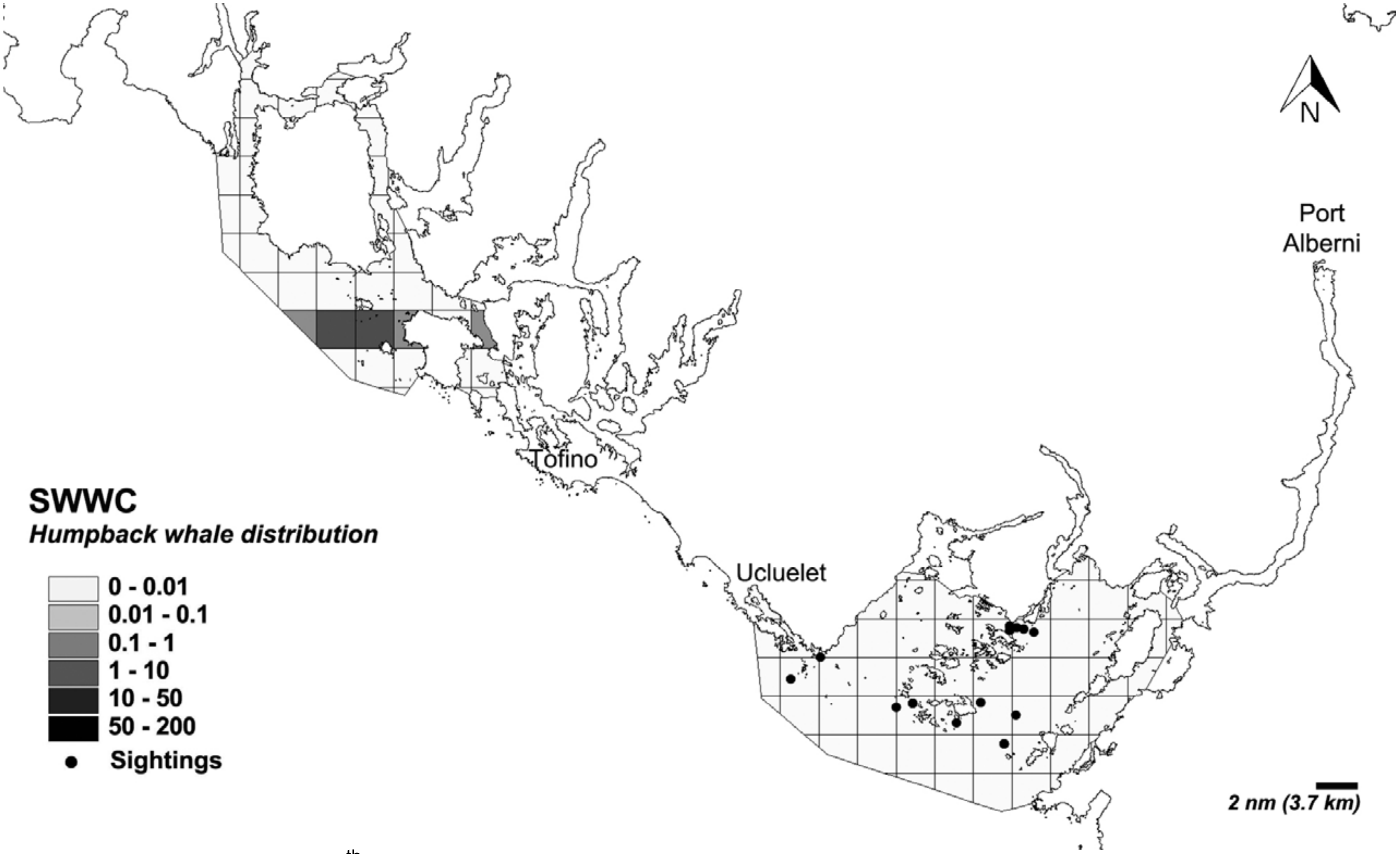


Figure 42: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWWC survey region with locations of sightings overlaid (solid black dots).

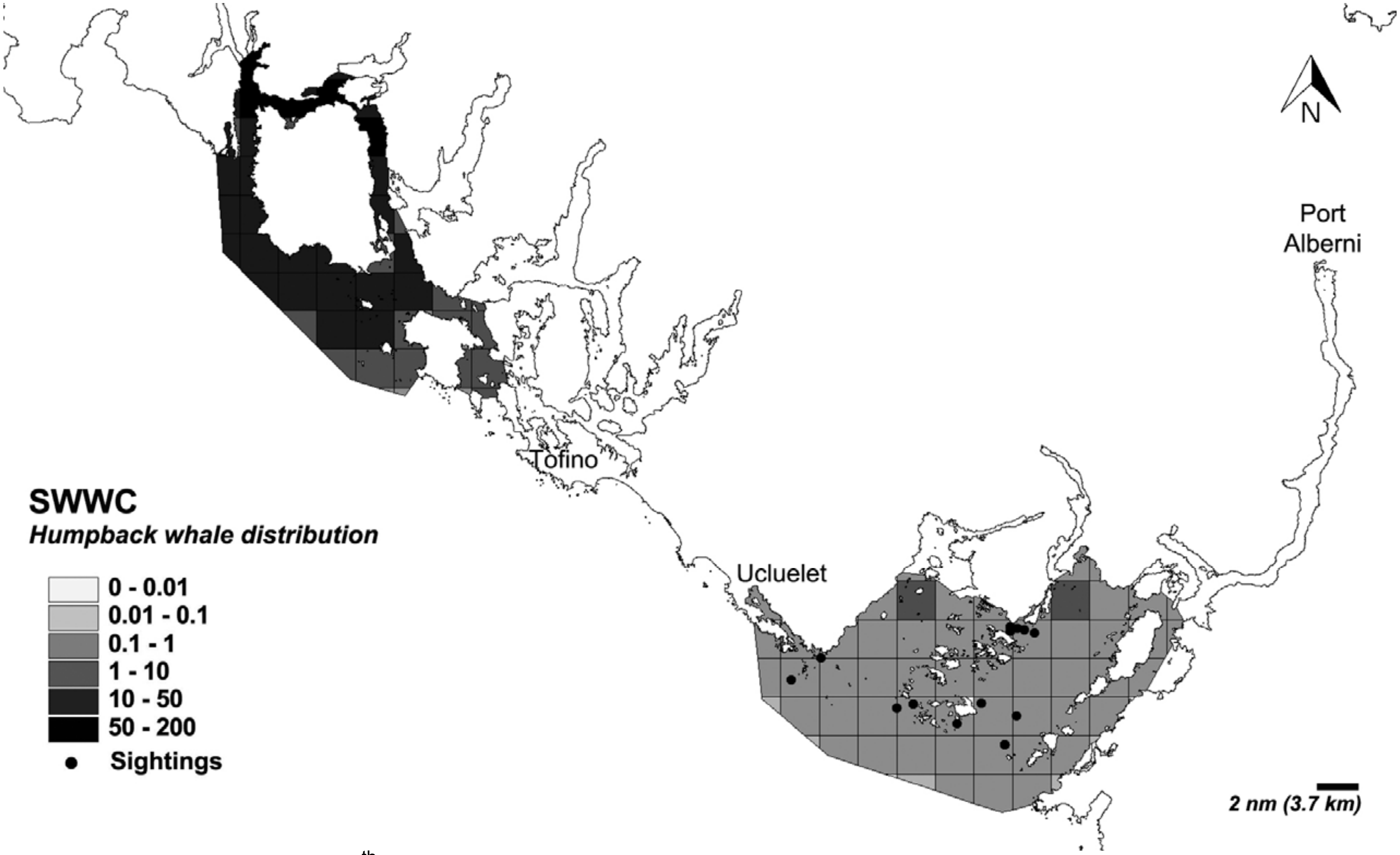


Figure 43: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for humpback whale in the SWWC survey region with locations of sightings overlaid (solid black dots).

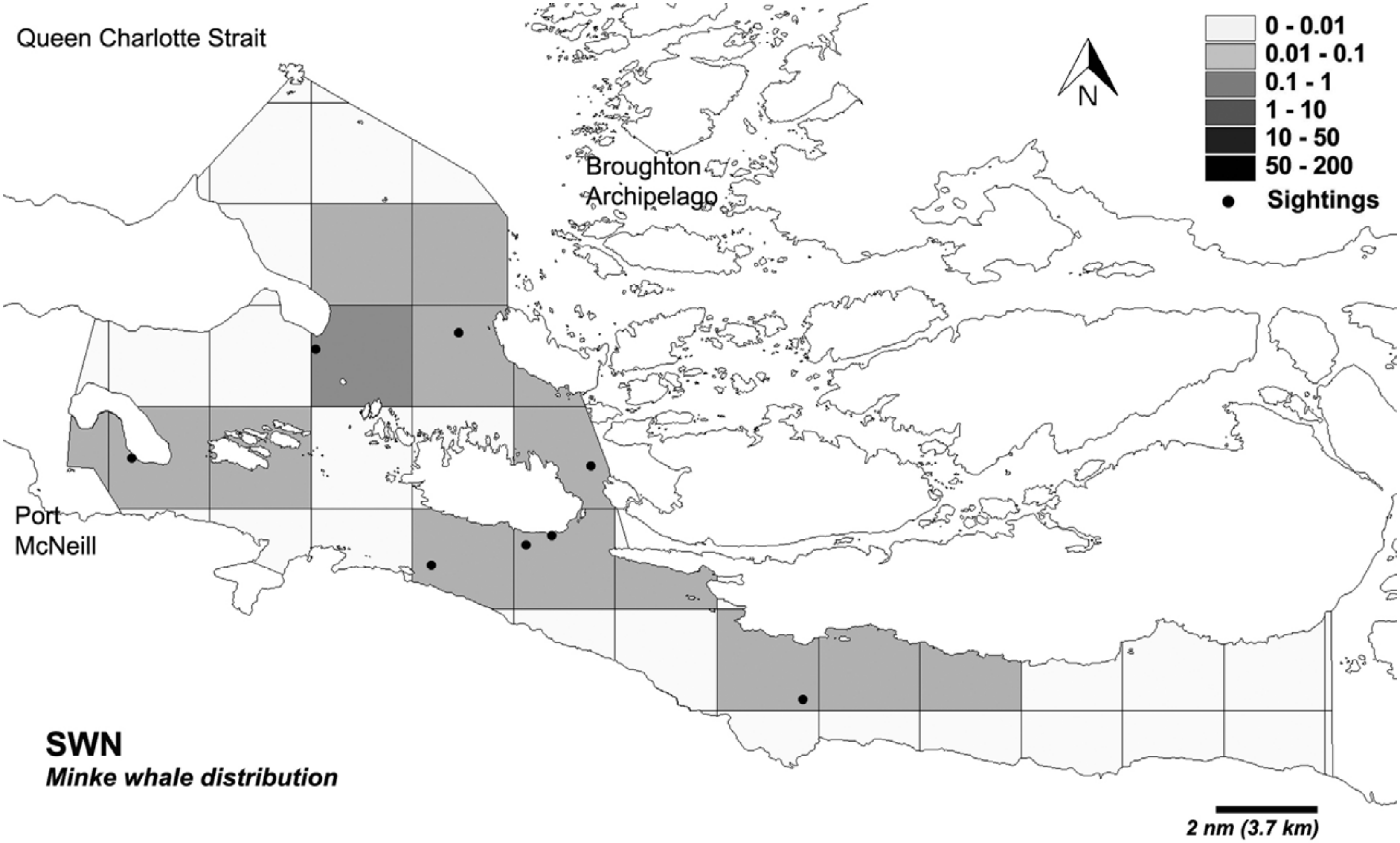


Figure 44: Mean estimate of distribution of abundance (shadings, in numbers of animals) for minke whale in the SWN survey region with locations of sightings overlaid (solid black dots). During 2008 & 2009, 9 sightings of 9 individuals were made.

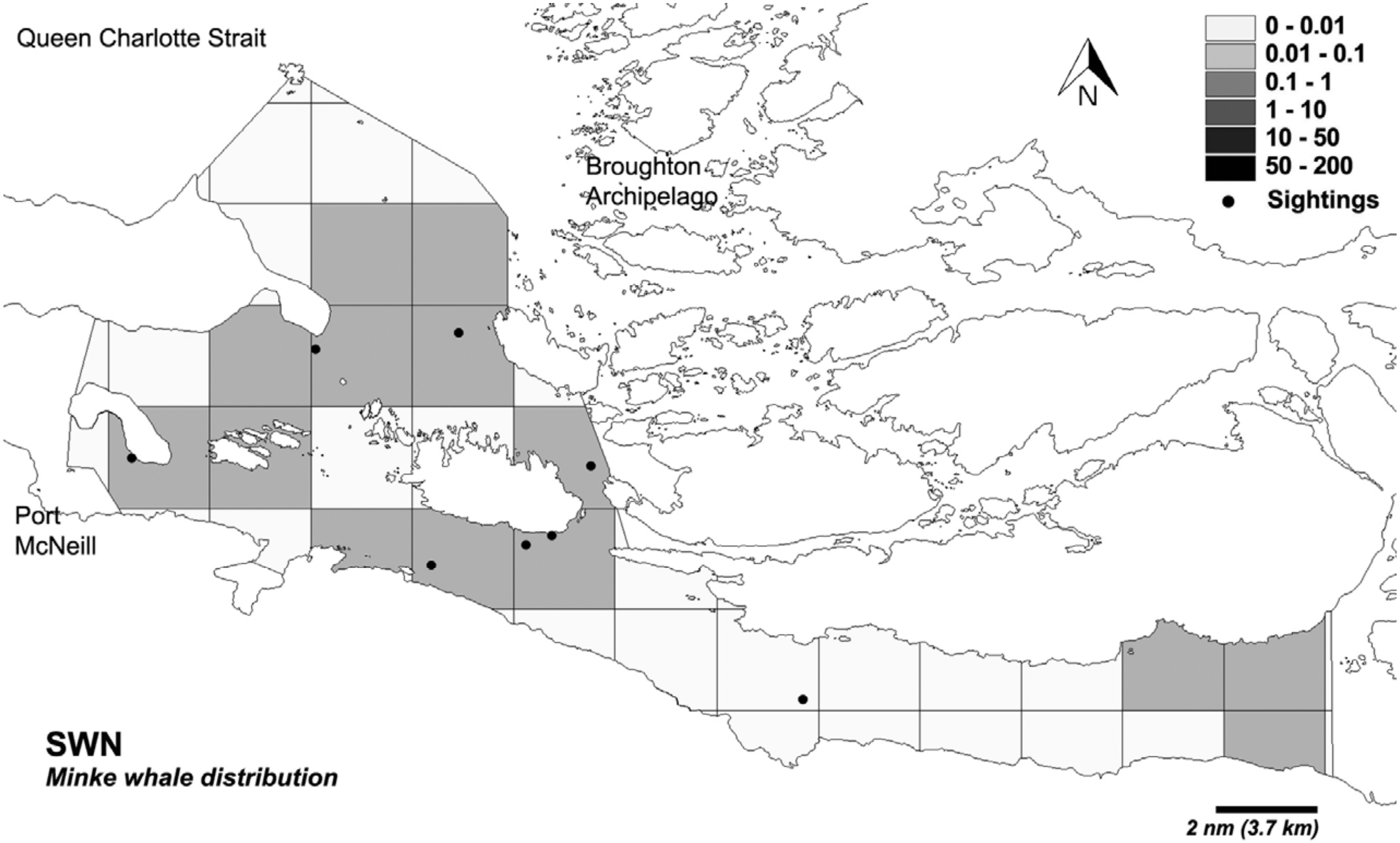


Figure 45: Lower bound (2.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for minke whale in the SWN survey region with locations of sightings overlaid (solid black dots).

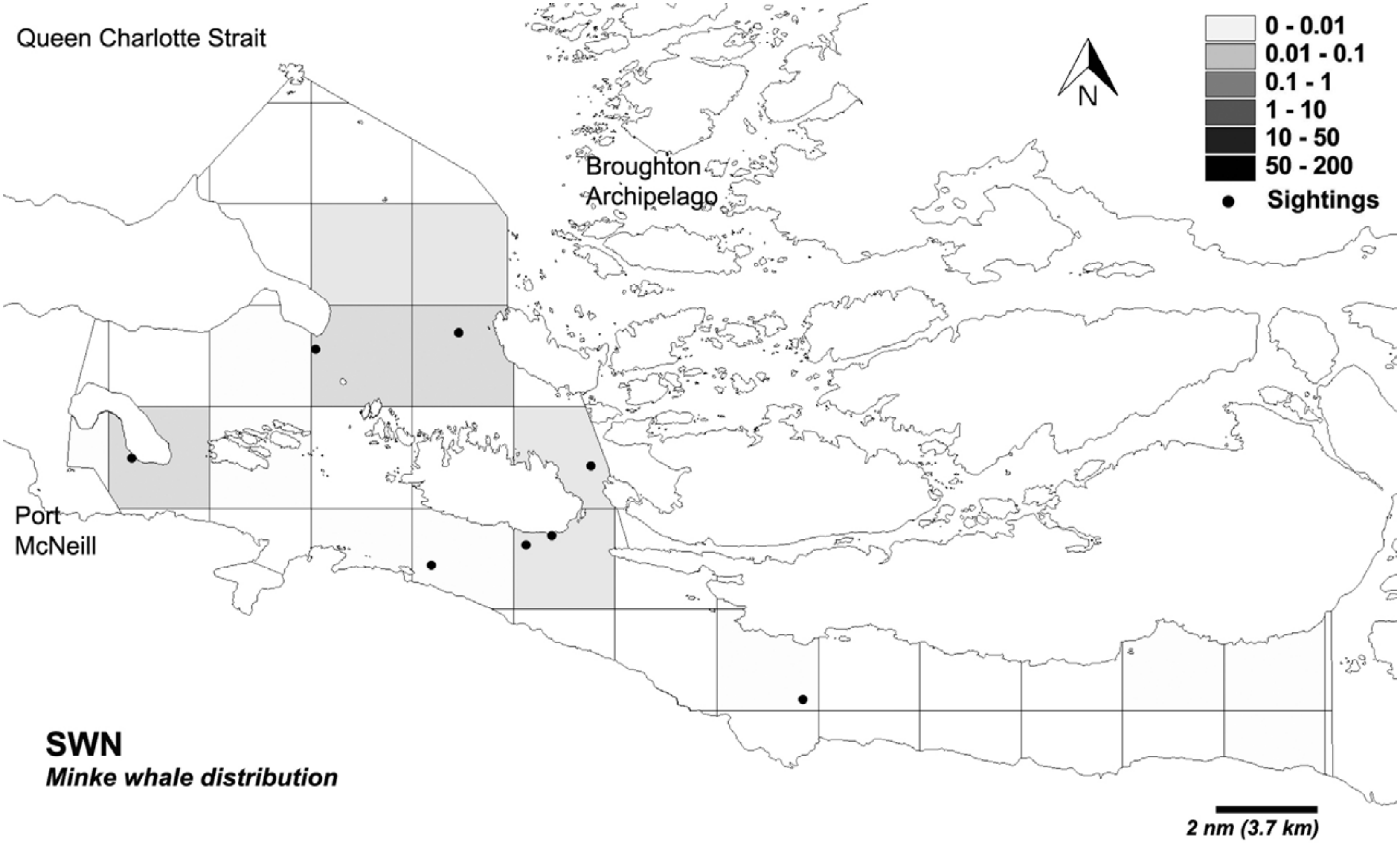


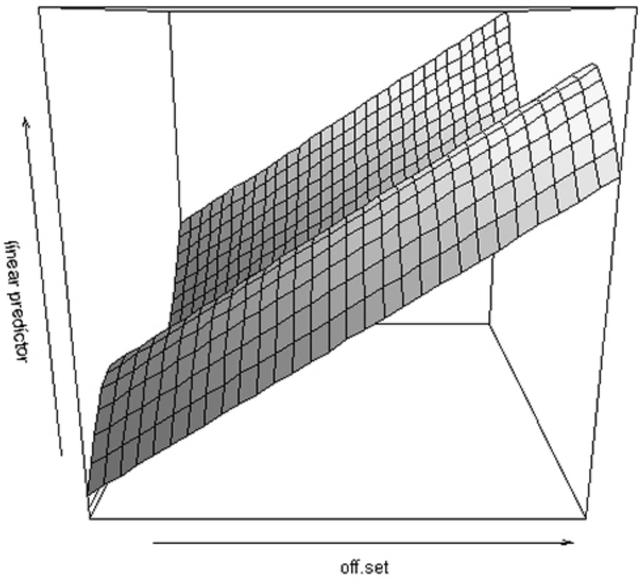
Figure 46: Upper bound (97.5th percentile) estimate of distribution of abundance (shadings, in numbers of animals) for minke whale in the SWN survey region with locations of sightings overlaid (solid black dots).

Figure 47: Mean estimate of distribution of abundance (shadings, in numbers of animals) for grey whales in the SWWC survey region with locations of sightings overlaid (solid black dots). During 2009, 5 sightings of 8 individuals were made.

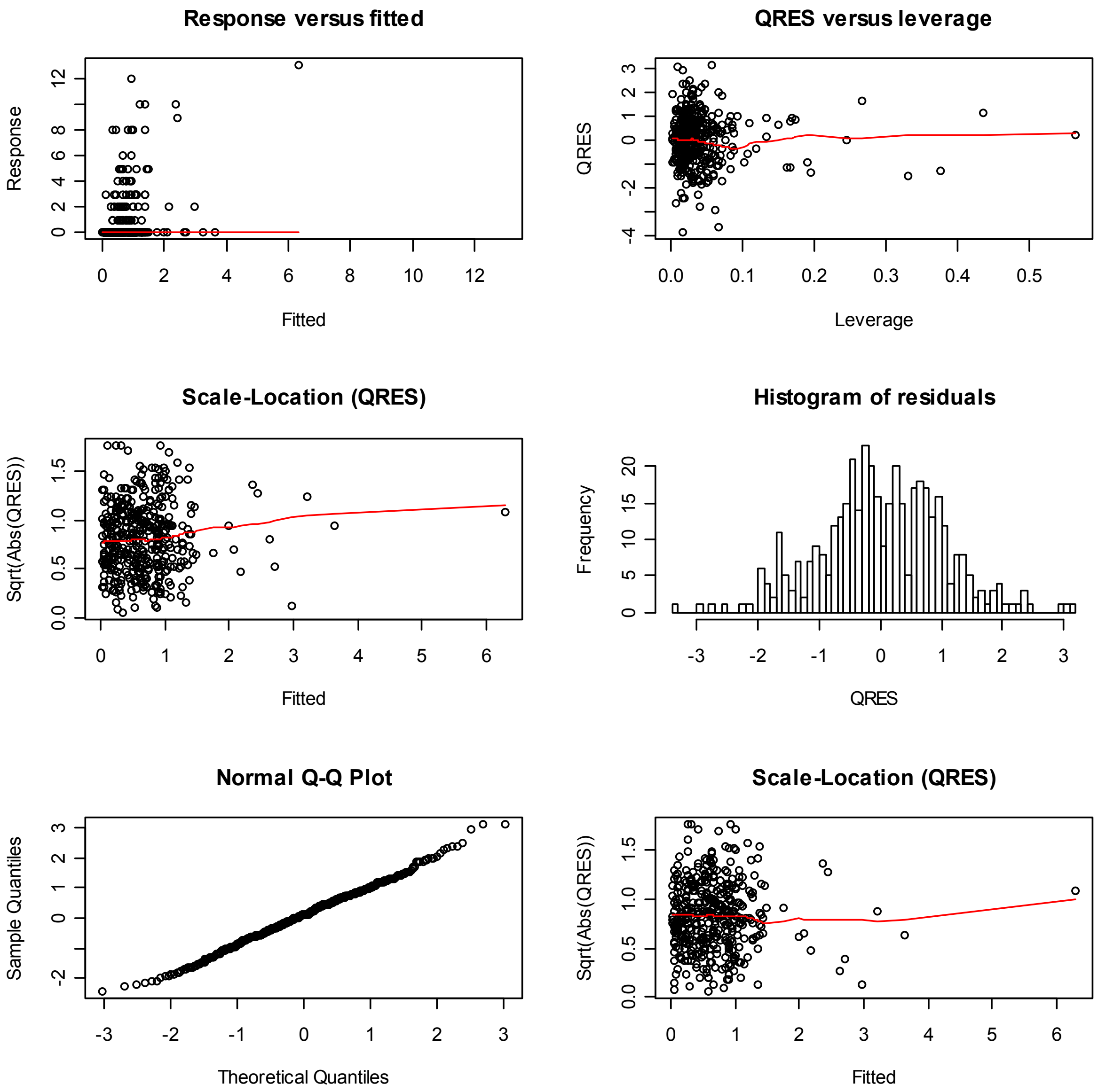
APPENDIX I

SWN Survey Region GAM model fit: Dall’s porpoise

```
Family: Tweedie(1.1)
Link function: log

Formula:
N ~ te(midlat, tidal.speed, depth) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -16.122      0.113  -142.7   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                                edf Ref.df     F p-value
te(midlat,tidal.speed,depth) 15.09  15.09 2.451 0.00189 **
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0982   Deviance explained = 15.3%
REML score = 370.41  Scale est. = 2.3576    n = 414
AIC = 787.9
```

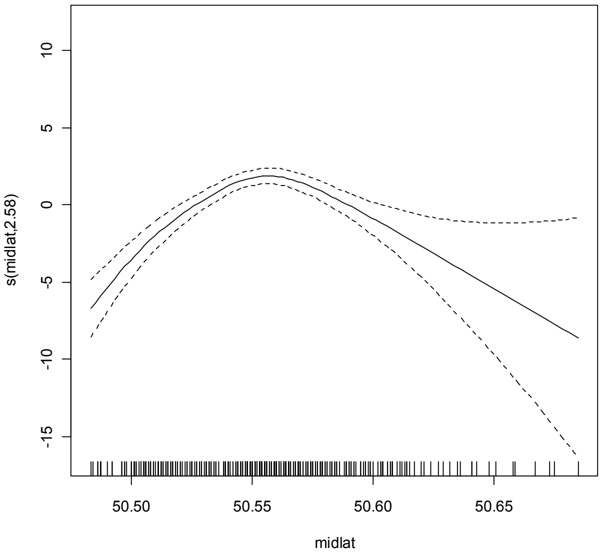


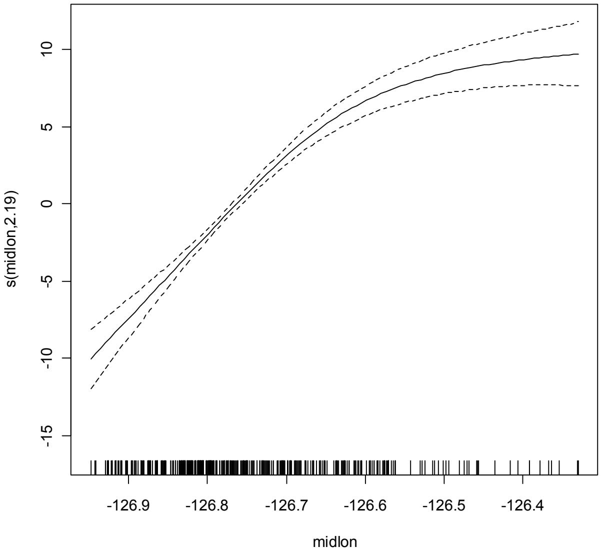
For plots of smooths with 3 or more covariates, the plot of the selected gam object represents the component smooth functions that make up the selected gam, on the scale of the linear predictor.


Response versus fitted
Response
0
2
4
6
8
12
Fitted
0
2
4
6
8
10
12
QRES versus leverage
QRES
-4
-2
0
1
2
3
Leverage
0.0
0.1
0.2
0.3
0.4
0.5
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
Fitted
0
1
2
3
4
5
6
Histogram of residuals
Frequency
0
5
10
15
20
QRES
-3
-2
-1
0
1
2
3
Normal Q-Q Plot
Sample Quantiles
-2
0
1
2
3
Theoretical Quantiles
-3
-2
-1
0
1
2
3
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
Fitted
0
1
2
3
4
5
6

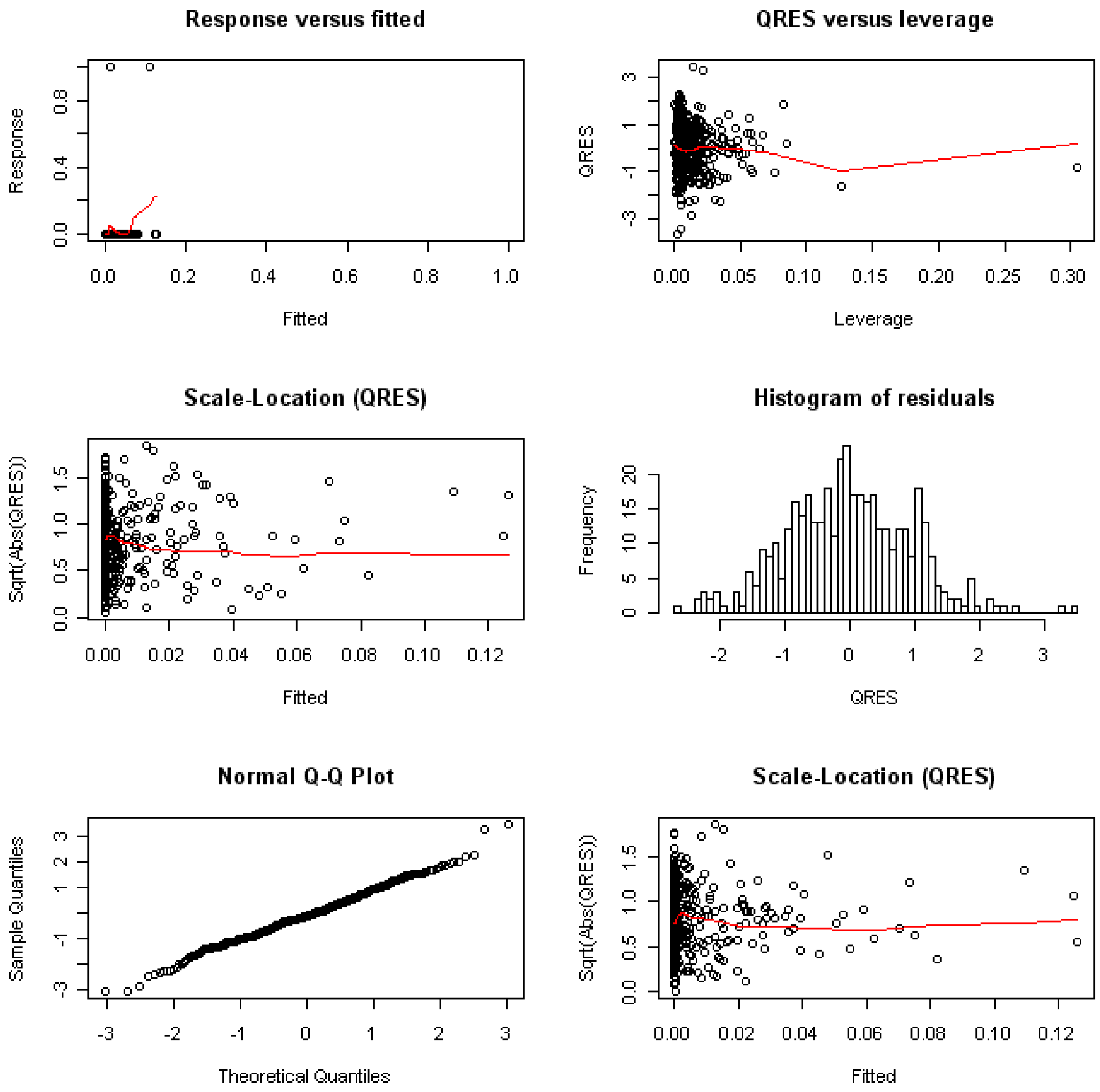
SWN Survey Region GAM model fit: Harbour porpoise

```
Family: Tweedie(1.6)
Link function: log
Formula:
N ~ s(midlat) + s(midlon) + offset(off.set)
Parametric coefficients:
          Estimate Std. Error t value Pr(>|t|)
(Intercept) -24.9546     0.3137  -79.55   <2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
          edf Ref.df     F p-value
s(midlat) 2.576  3.288 26.11  <2e-16 ***
s(midlon) 2.192  2.756 68.94  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
R-sq.(adj) =  0.0339   Deviance explained =   49%
REML score = 15.855  Scale est. = 0.54697   n = 414
AIC = 431.5199
```



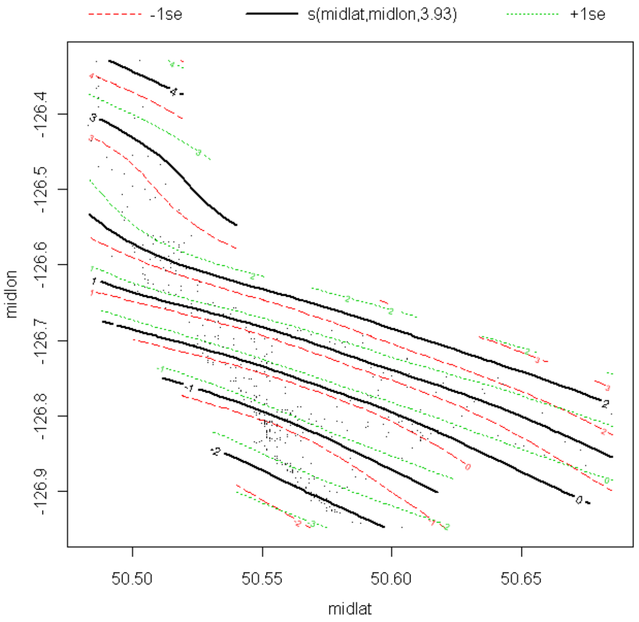


For plots of 1-D smooths, the x-axis of each plot if labelled with the covariate name, while the y-axis is labelled s(cov,edf) where cov is the covariate name, and edf is the estimated degrees of freedom of the smooth.

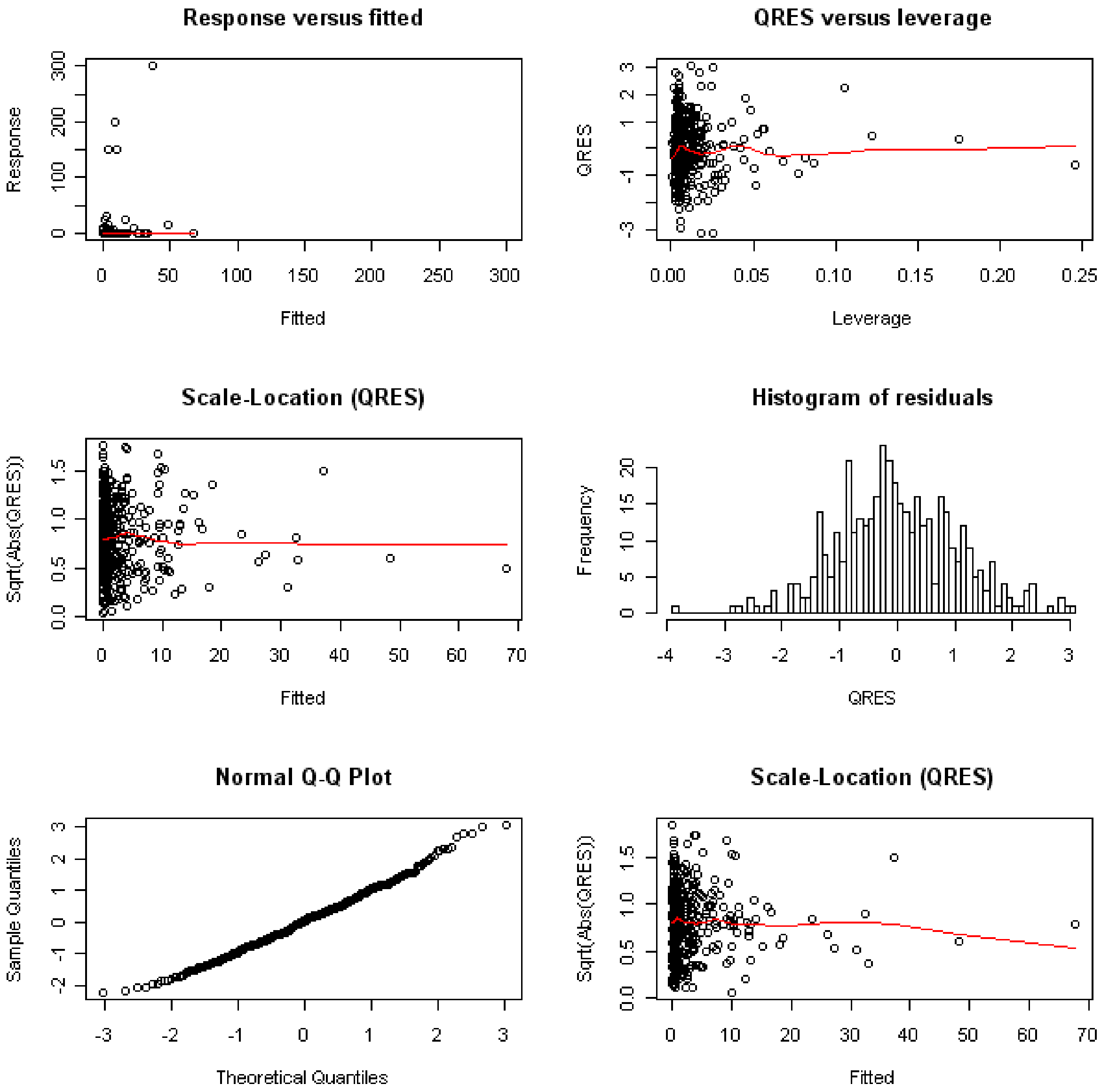

Response versus fitted
Response
0.0
0.4
0.8
0.0
0.2
0.4
0.6
0.8
1.0
Fitted
QRES versus leverage
QRES
-3
-1
1
3
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Fitted
Histogram of residuals
Frequency
0
5
10
15
20
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-3
-1
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Fitted

SWN Survey Region GAM model fit: Pacific white-sided dolphin

```
Family: Tweedie(1.3)
Link function: log
Formula:
N ~ s(midlat, midlon) + offset(off.set)
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -15.4436     0.1904   -81.1   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                  edf Ref.df     F p-value
s(midlat,midlon) 3.931   5.07 28.11  <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
R-sq.(adj) =  0.0528   Deviance explained = 31.4%
REML score = 194.29  Scale est. = 9.2581    n = 414
AIC = 610.9578
```

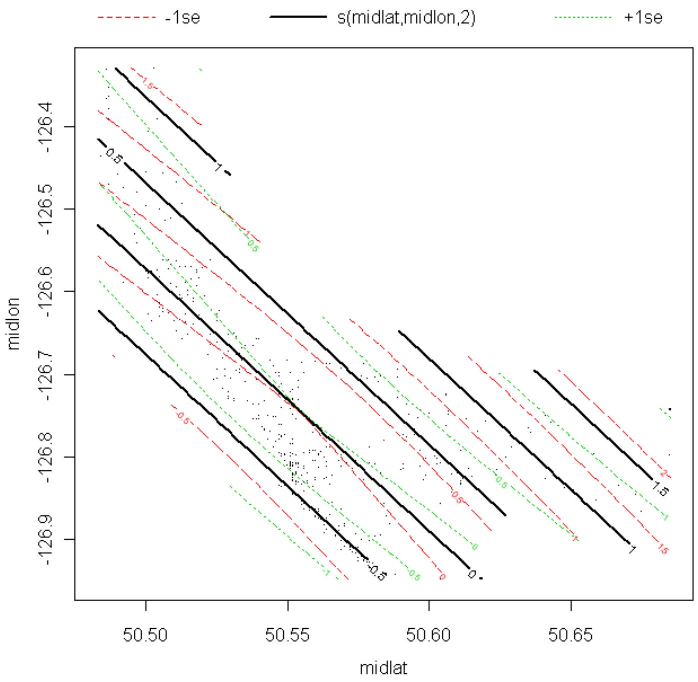


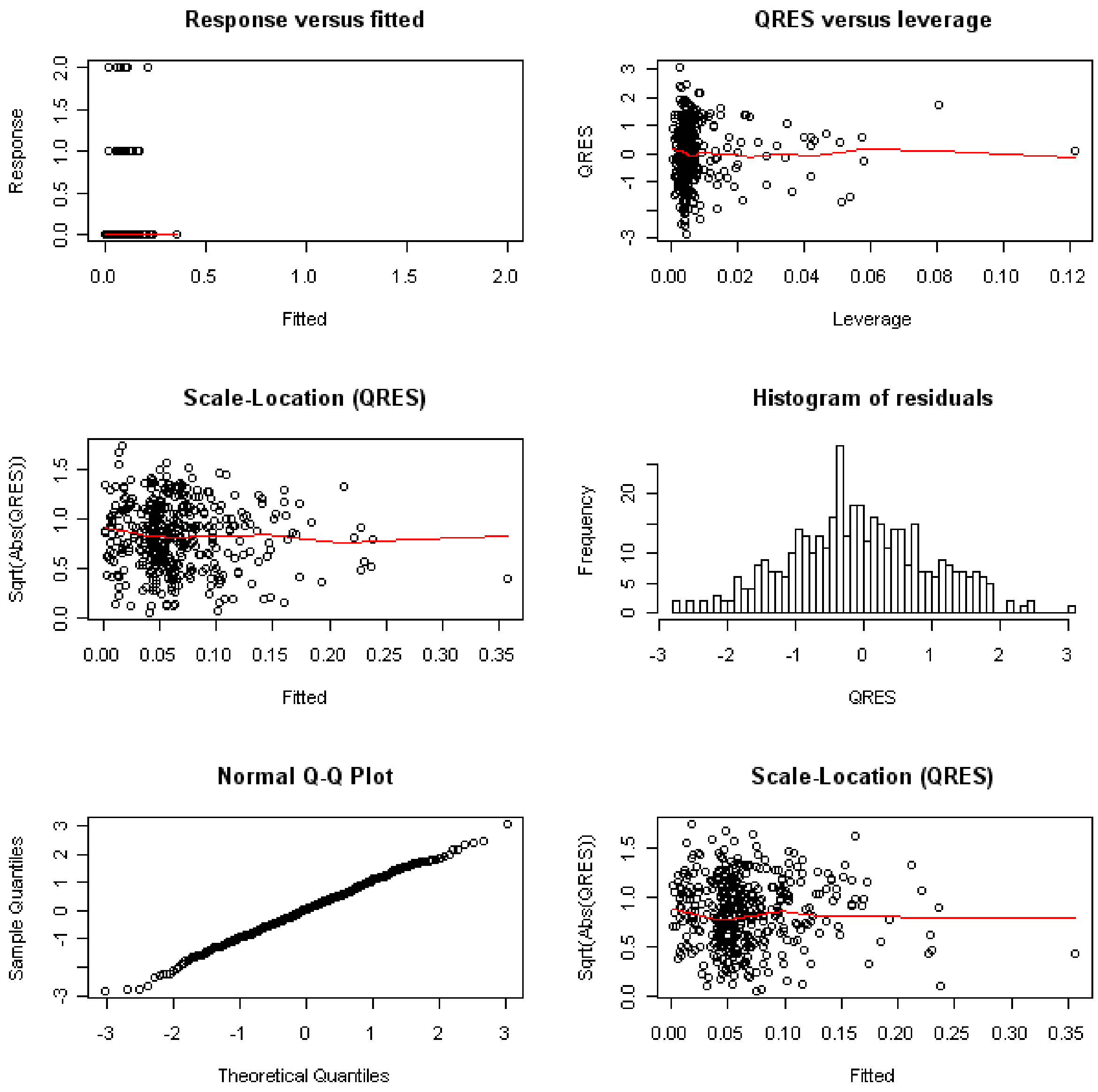
Contour plots produced for 2-D smooths, with the x-axis labelled with the first covariate name and the y-axis with the second covariate name. The main title of the plot is something like s(var1, var2, edf), indicating the variable of which the term is a function, and the estimate degrees of freedom for the term. When se=TRUE, estimator variability is shown by overlaying contour plots at plus and minus 1 s.e. relative to the main estimate. If se if a positive number then the contour plots are at plus of minus se multiplied by the s.e.


Response versus fitted
Response
0
100
200
300
0
50
100
150
200
250
300
Fitted
QRES versus leverage
QRES
-3
-1
1
2
3
0.00
0.05
0.10
0.15
0.20
0.25
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0
10
20
30
40
50
60
70
Fitted
Histogram of residuals
Frequency
0
5
10
15
20
-4
-3
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-2
-1
0
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0
10
20
30
40
50
60
70
Fitted

SWN Survey Region GAM model fit: Humpback whale

```
Family: Tweedie(1.1)
Link function: log

Formula:
N ~ s(midlat, midlon) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -19.4129     0.1312  -148.0   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                 edf Ref.df     F  p-value
s(midlat,midlon)   2      2 9.792 7.01e-05 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0108   Deviance explained = 4.15%
REML score = 96.472  Scale est. = 0.47762   n = 414
```

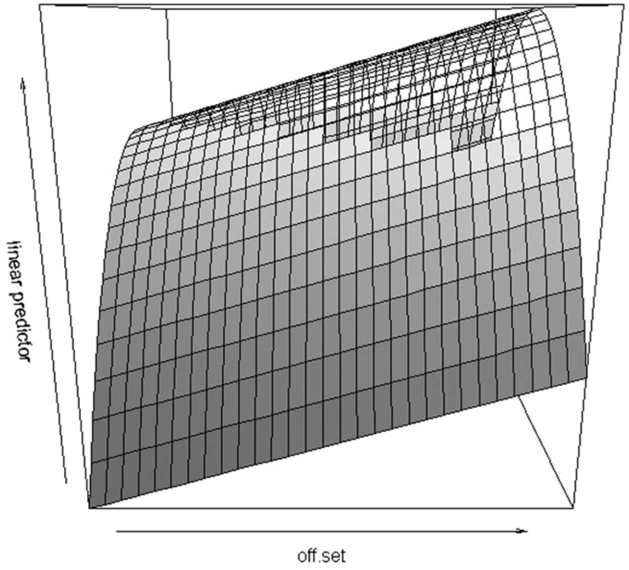


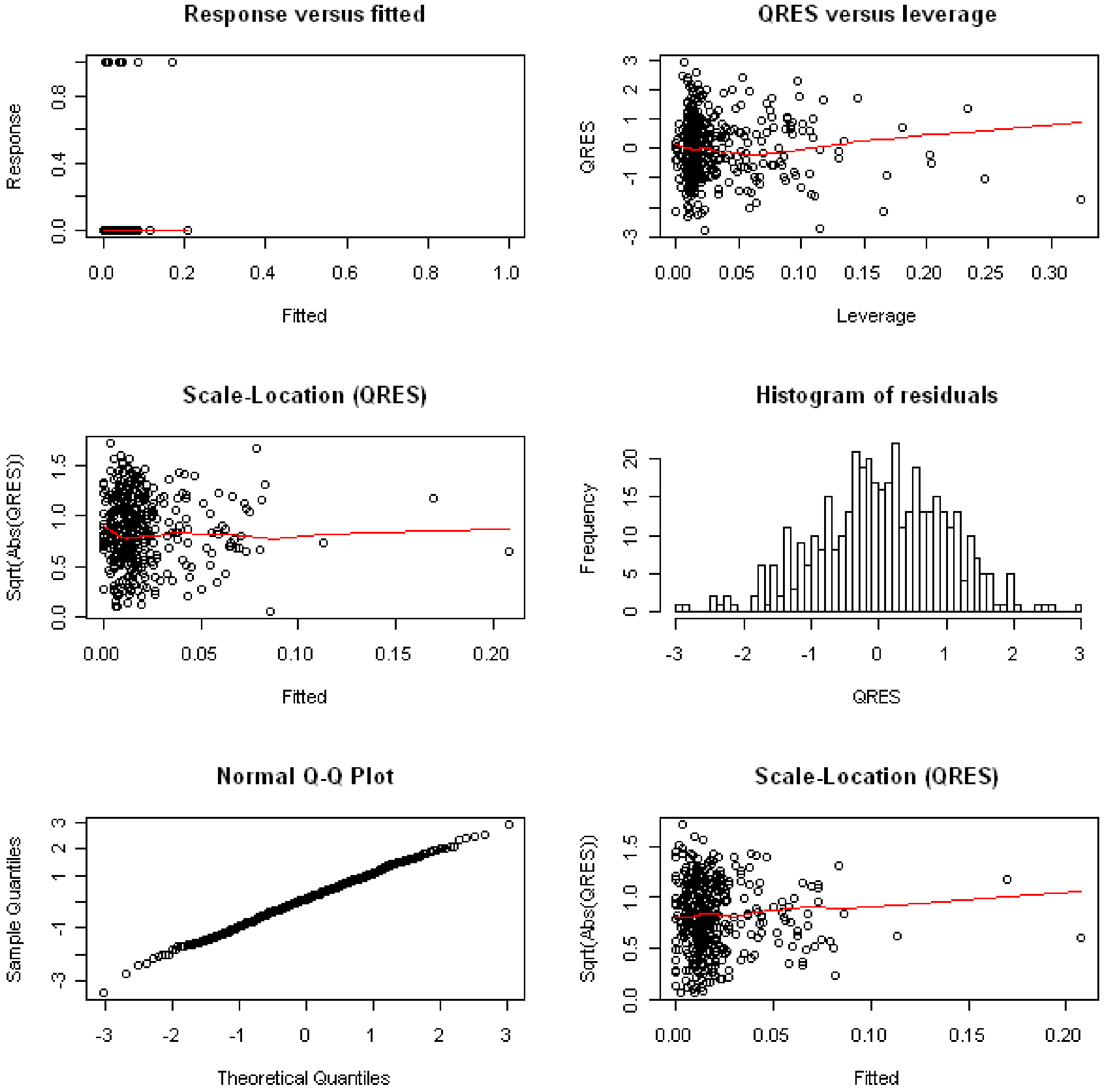


Response versus fitted
Response
0.0
0.5
1.0
1.5
2.0
Fitted
QRES versus leverage
QRES
-3
-1
0
1
2
3
0.00
0.02
0.04
0.06
0.08
0.10
0.12
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
Fitted
Histogram of residuals
Frequency
0
5
10
20
-3
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-3
-1
0
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
Fitted

SWN Survey Region GAM model fit: Minke whale

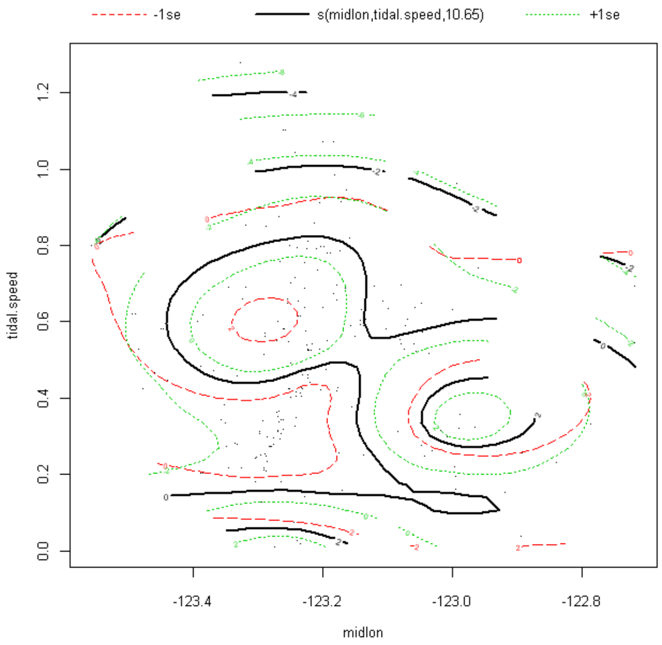
```
Family: Tweedie(1.1)
Link function: log
Formula:
N ~ s(midlat, midlon, tidal.speed) + offset(off.set)

Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -19.3295     0.2702  -71.54   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                            edf Ref.df     F p-value
s(midlat,midlon,tidal.speed) 13.07  13.07 2.524 0.00239 **
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.0198   Deviance explained = 19.5%
REML score = 42.598  Scale est. = 0.18813   n = 414
AIC = 445.5
```

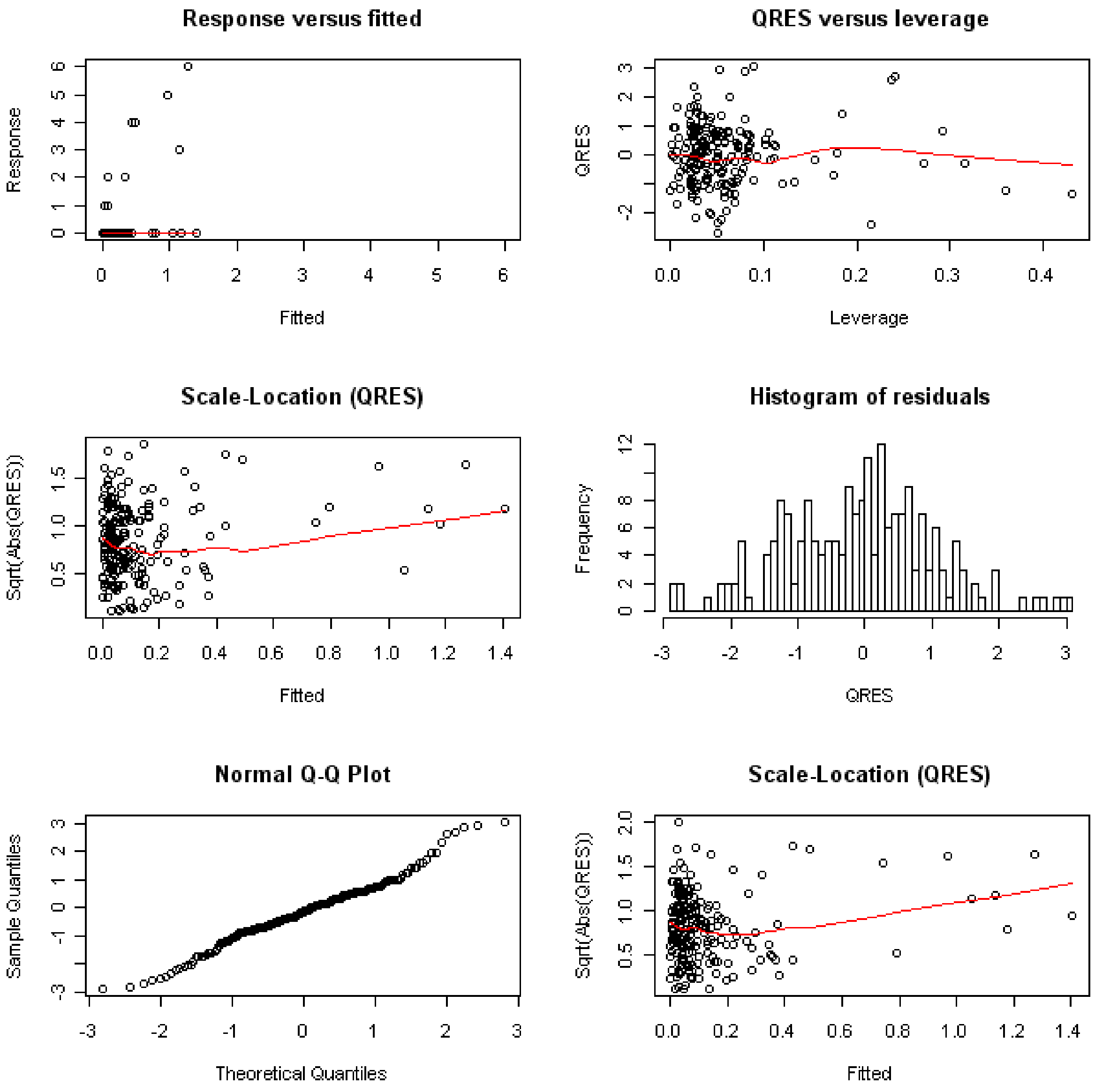



Response versus fitted
Response
0.0
0.4
0.8
0.0
0.2
0.4
0.6
0.8
1.0
Fitted
QRES versus leverage
QRES
-3
-1
0
1
2
3
0.00
0.05
0.10
0.15
0.20
0.25
0.30
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.05
0.10
0.15
0.20
Fitted
Histogram of residuals
Frequency
0
5
10
15
20
-3
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-3
-1
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.00
0.05
0.10
0.15
0.20
Fitted

SWS Survey Region GAM model fit: Dall’s porpoise

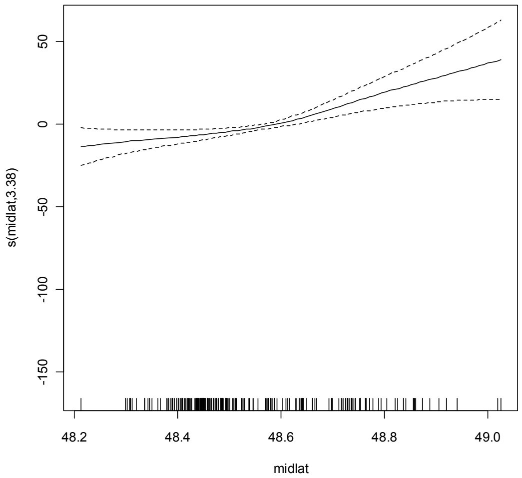
```
Family: Tweedie(1.1)
Link function: log
Formula:
N ~ s(midlon, tidal.speed) + offset(off.set)

Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -18.2425     0.2195   -83.1   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
                          edf Ref.df      F p-value
s(midlon,tidal.speed) 10.65  14.37 2.176 0.00942 **
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.181   Deviance explained = 52.9%
REML score = 55.426  Scale est. = 0.62293   n = 202
AIC = 244.5
```

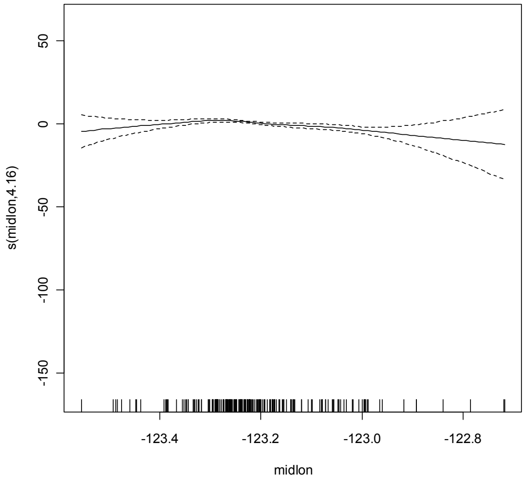


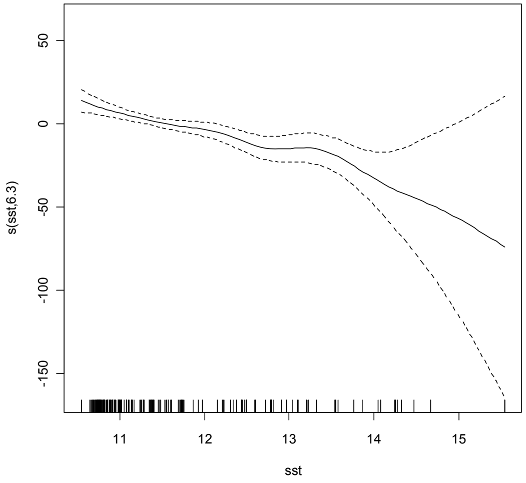

Response versus fitted
Response
0
1
2
3
4
5
6
Fitted
QRES versus leverage
QRES
-2
0
1
2
3
0.0
0.1
0.2
0.3
0.4
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Fitted
Histogram of residuals
Frequency
0
2
4
6
8
12
-3
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-3
-1
0
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
Fitted

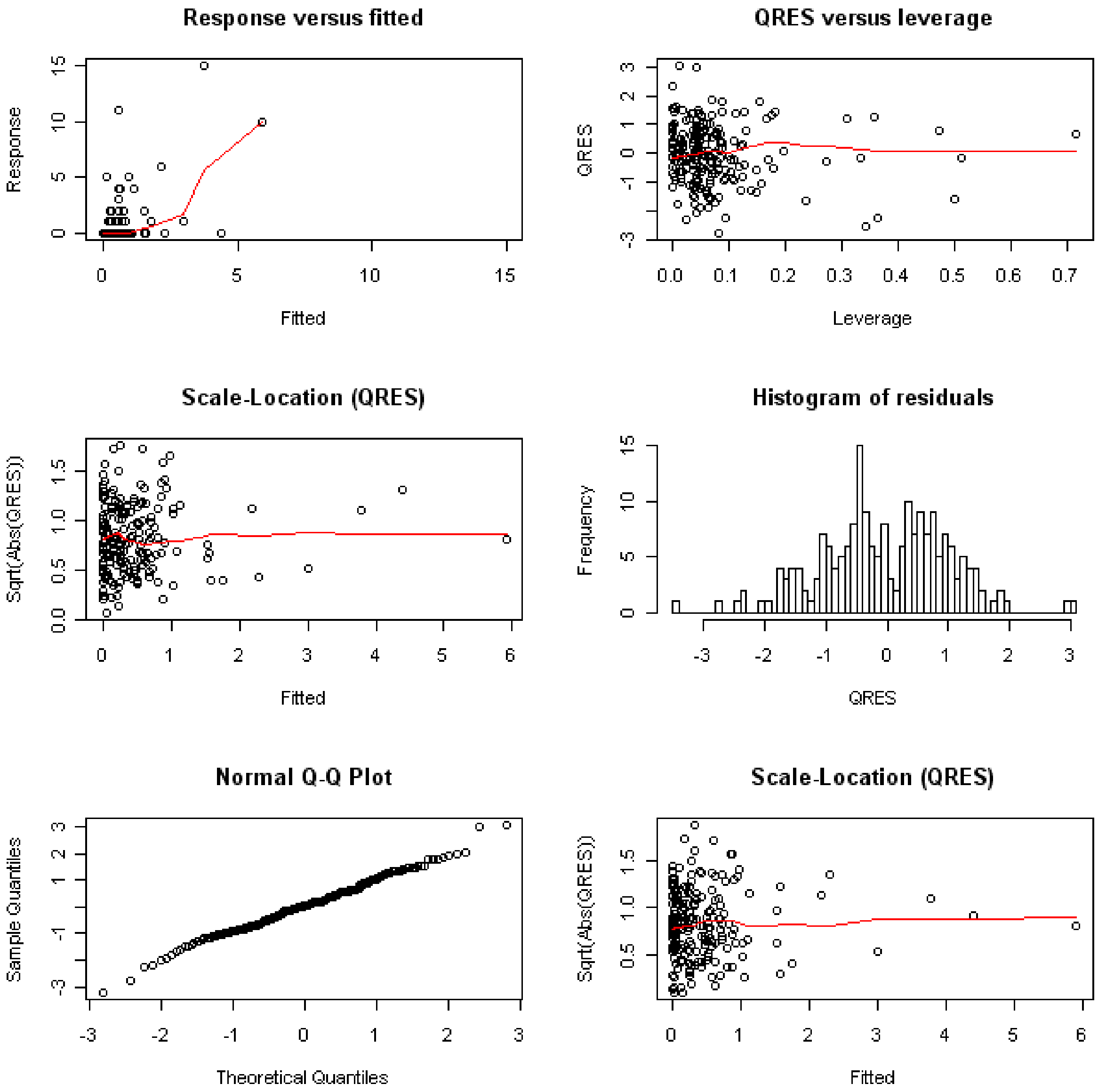
SWS Survey Region GAM model fit: Harbour porpoise

```
Family: Tweedie(1.1)
Link function: log
Formula:
N ~ s(midlat) + s(midlon) + s(sst) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -18.4308     0.9678  -19.04   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
            edf Ref.df     F  p-value
s(midlat) 3.380  4.302 3.961 0.003282 **
s(midlon) 4.162  5.039 5.341 0.000124 ***
s(sst)    6.303  6.834 3.497 0.001654 **
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.222   Deviance explained = 37.9%
REML score = 149.84  Scale est. = 1.5321    n = 202
AIC = 318.7
```



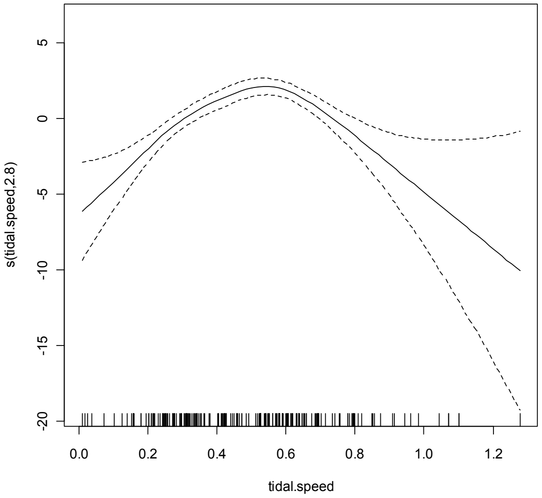


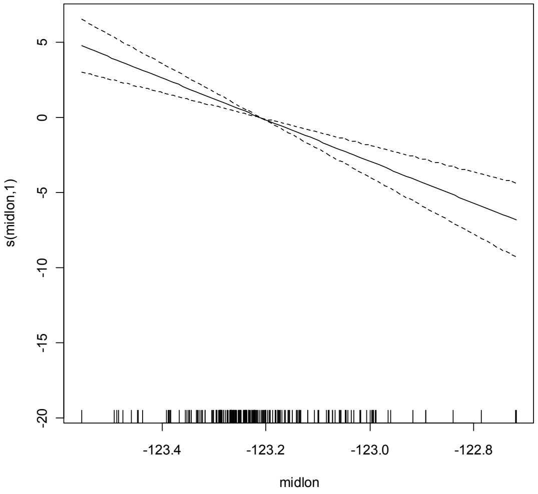


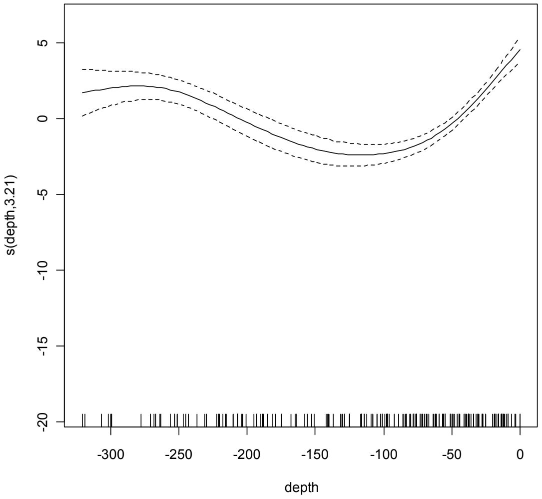

Response versus fitted
Response
Fitted
0
5
10
15
QRES versus leverage
QRES
Leverage
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
-3
-1
0
1
2
3
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.0
0.5
1.0
1.5
0
1
2
3
4
5
6
Histogram of residuals
Frequency
QRES
0
5
10
15
-3
-2
-1
0
1
2
3
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-3
-1
1
2
3
-3
-2
-1
0
1
2
3
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.5
1.0
1.5
0
1
2
3
4
5
6

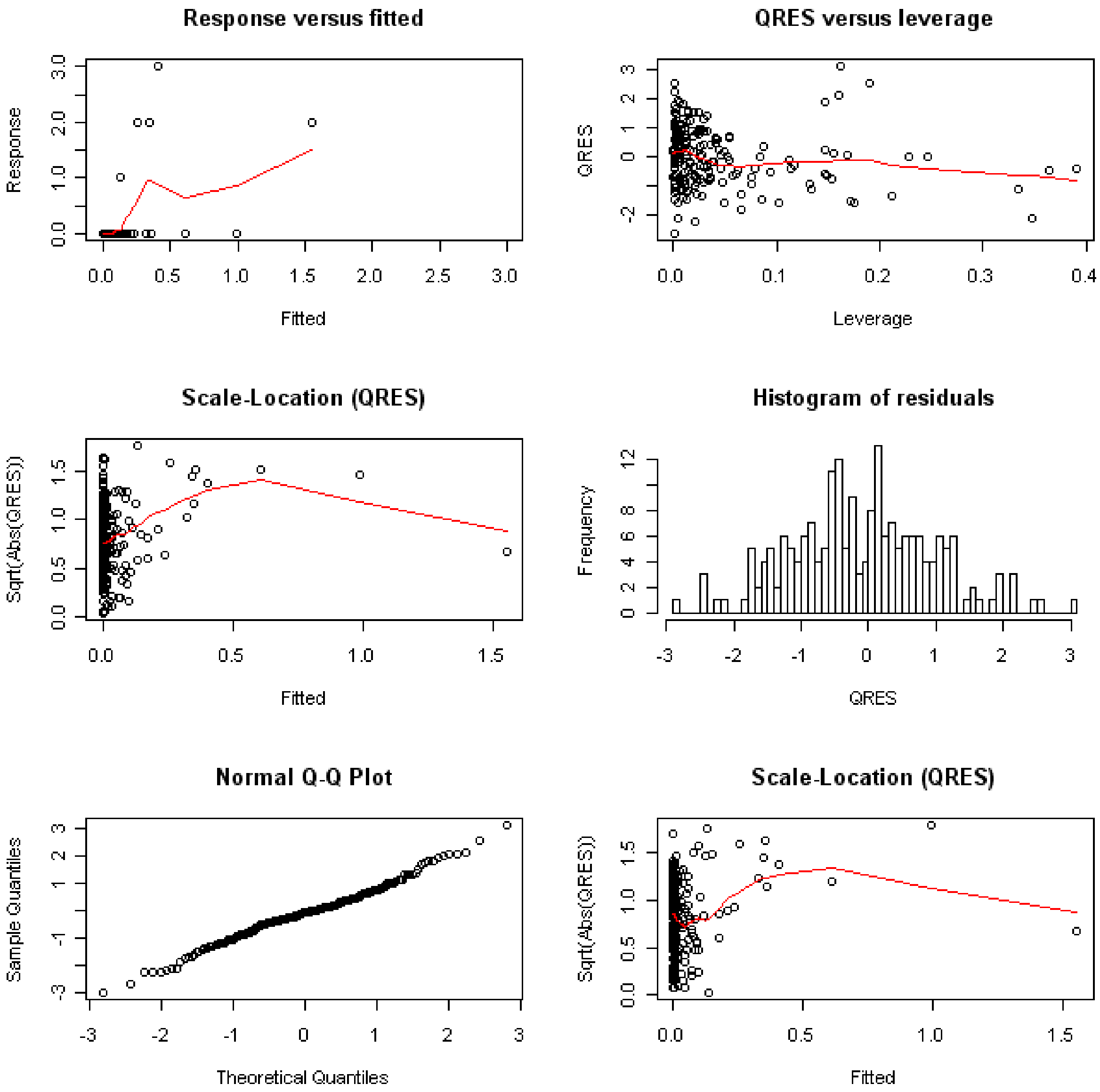
SWS Survey Region GAM model fit: Humpback whale

```
Family: Tweedie(1.1)
Link function: log

Formula:
N ~ s(midlon) + s(tidal.speed) + s(depth) + offset(off.set)
Parametric coefficients:
          Estimate Std. Error t value Pr(>|t|)
(Intercept) -20.4790     0.4122  -49.68   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
              edf Ref.df     F  p-value
s(midlon)      1.000  1.000 30.13 1.25e-07 ***
s(tidal.speed) 2.796  3.510 16.63 9.82e-11 ***
s(depth)       3.211  3.957 28.84  < 2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =   0.25   Deviance explained = 59.6%
REML score = 20.747  Scale est. = 0.18942   n = 202
AIC = 224.4
```



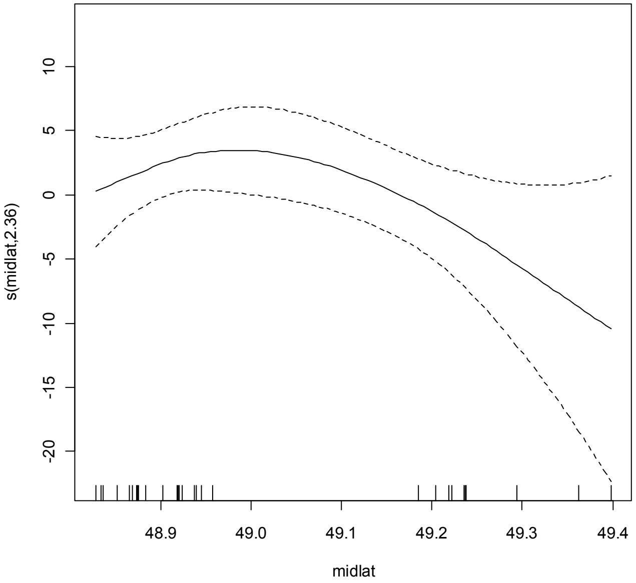


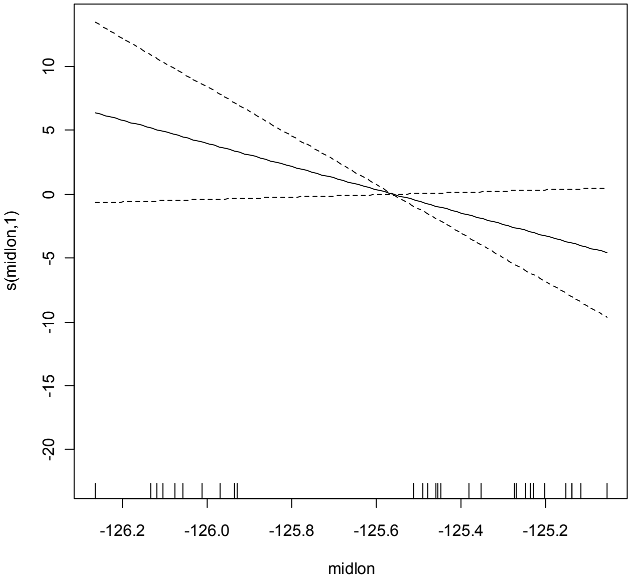


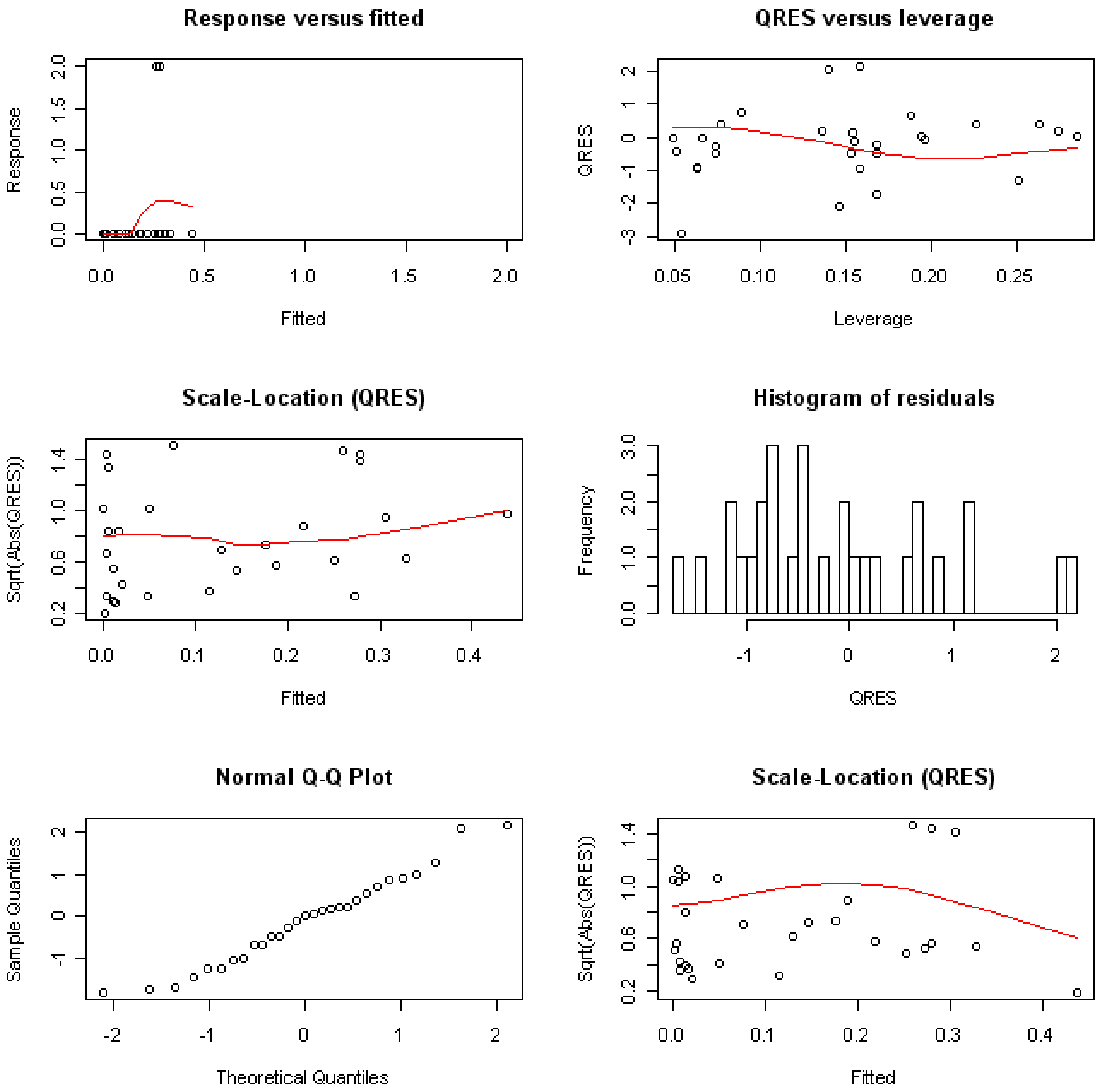

Response versus fitted
Response
0.0
1.0
2.0
3.0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Fitted
QRES versus leverage
QRES
-2
0
1
2
3
0.0
0.1
0.2
0.3
0.4
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
Fitted
Histogram of residuals
Frequency
0
2
4
6
8
12
-3
-2
-1
0
1
2
3
QRES
Normal Q-Q Plot
Sample Quantiles
-3
-1
1
2
3
-3
-2
-1
0
1
2
3
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
0.0
0.5
1.0
1.5
Fitted

SWWC Survey Region GAM model fit: Dall’s porpoise

```
Family: Tweedie(1.3)
Link function: log

Formula:
N ~ s(midlat) + s(midlon) + offset(off.set)
Parametric coefficients:
           Estimate Std. Error t value Pr(>|t|)
(Intercept) -18.6579     0.6923  -26.95   <2e-16 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
          edf Ref.df     F p-value
s(midlat) 2.24  2.831 1.600  0.2161
s(midlon) 1.00  1.000 3.165  0.0875 .
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  -0.028   Deviance explained = 31.4%
REML score = 5.1484  Scale est. = 1.0307    n = 29
AIC = 45.11
```







Response versus fitted

QRES versus leverage

Scale-Location (QRES)

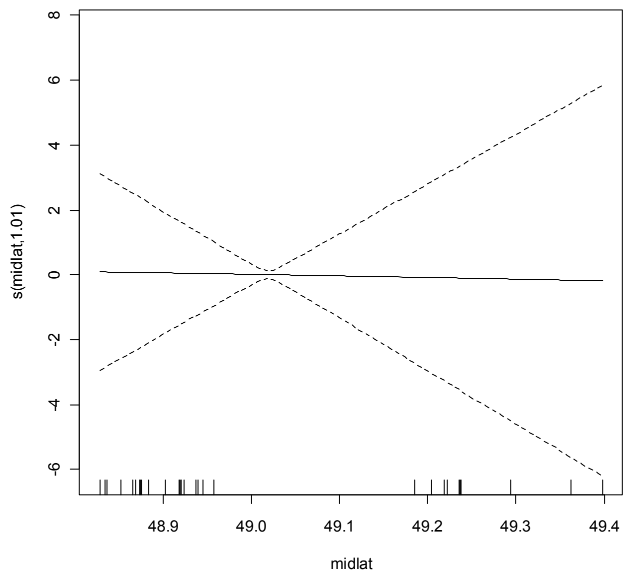
Histogram of residuals

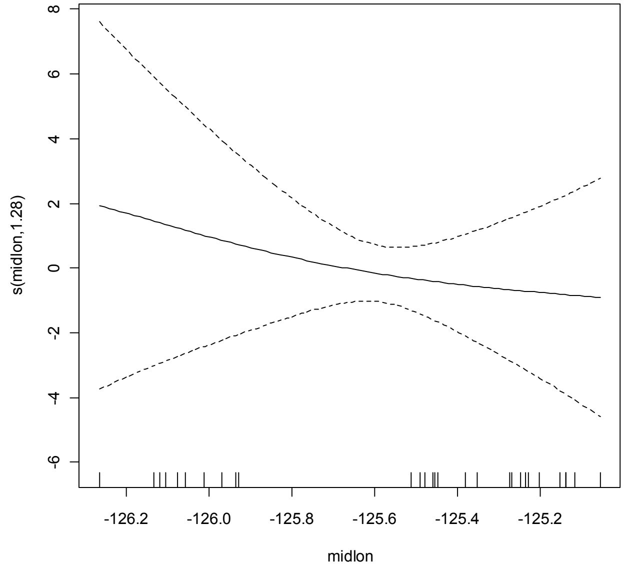
Normal Q-Q Plot

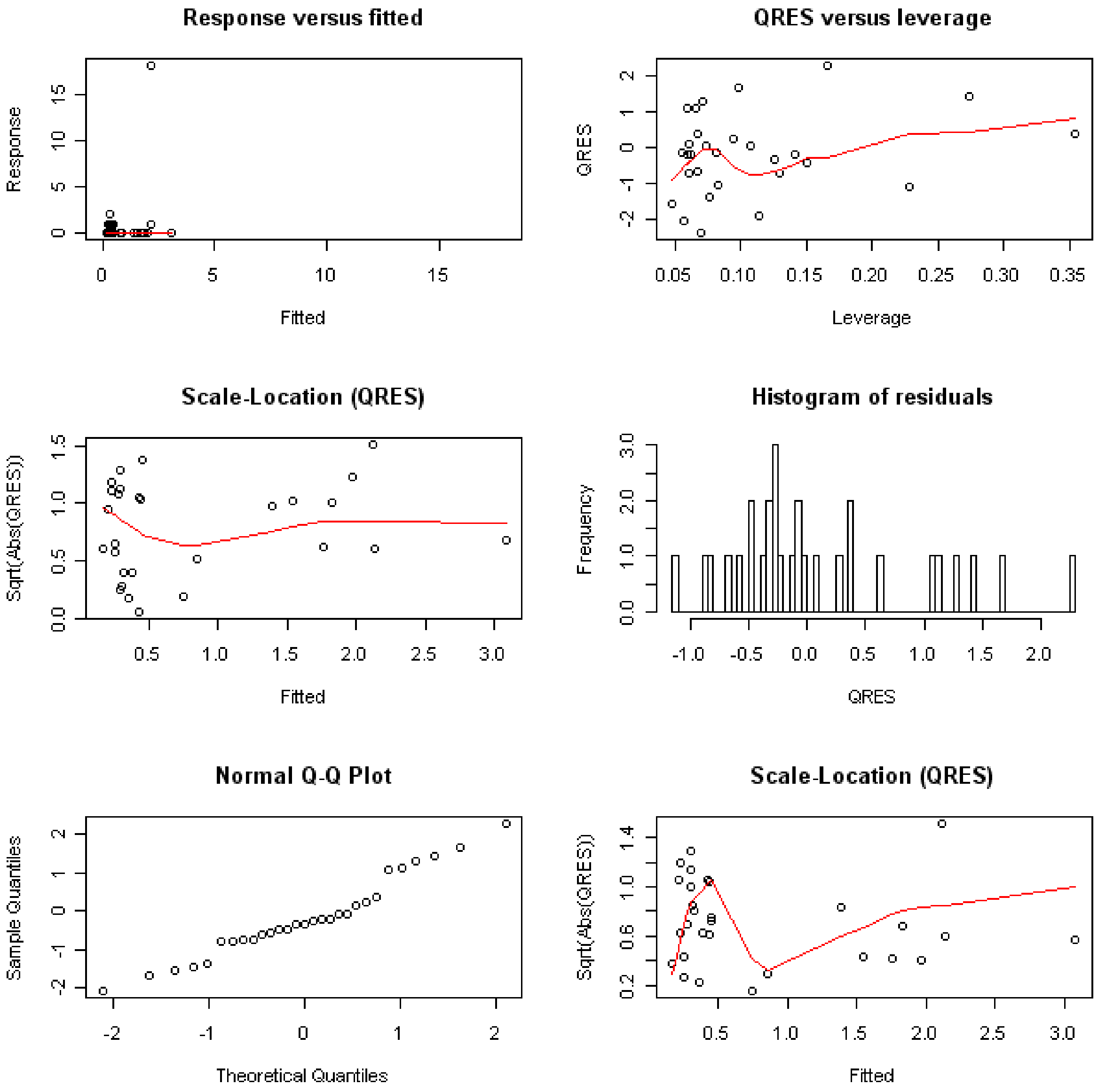
Scale-Location (QRES)

SWWC Survey Region GAM model fit: Harbour porpoise

```
Family: Tweedie(1.6)
Link function: log

Formula:
N ~ s(midlat) + s(midlon) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -16.0188     0.4669  -34.31   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
            edf Ref.df     F p-value
s(midlat) 1.000  1.000 0.001   0.973
s(midlon) 1.109  1.212 0.332   0.612

R-sq.(adj) =  0.0107   Deviance explained = 17.0%
REML score = 28.265  Scale est. = 4.7139    n = 29
AIC = 72.72
```

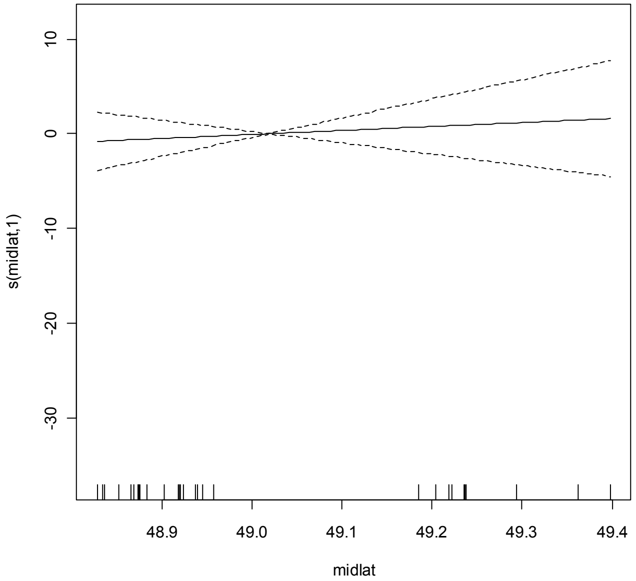


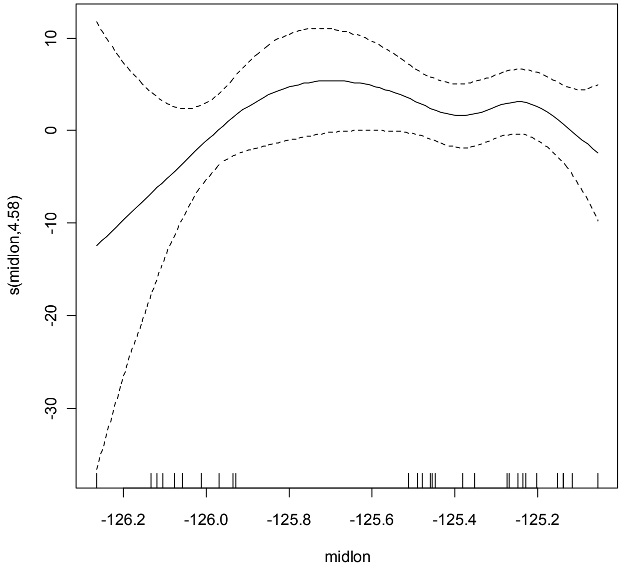


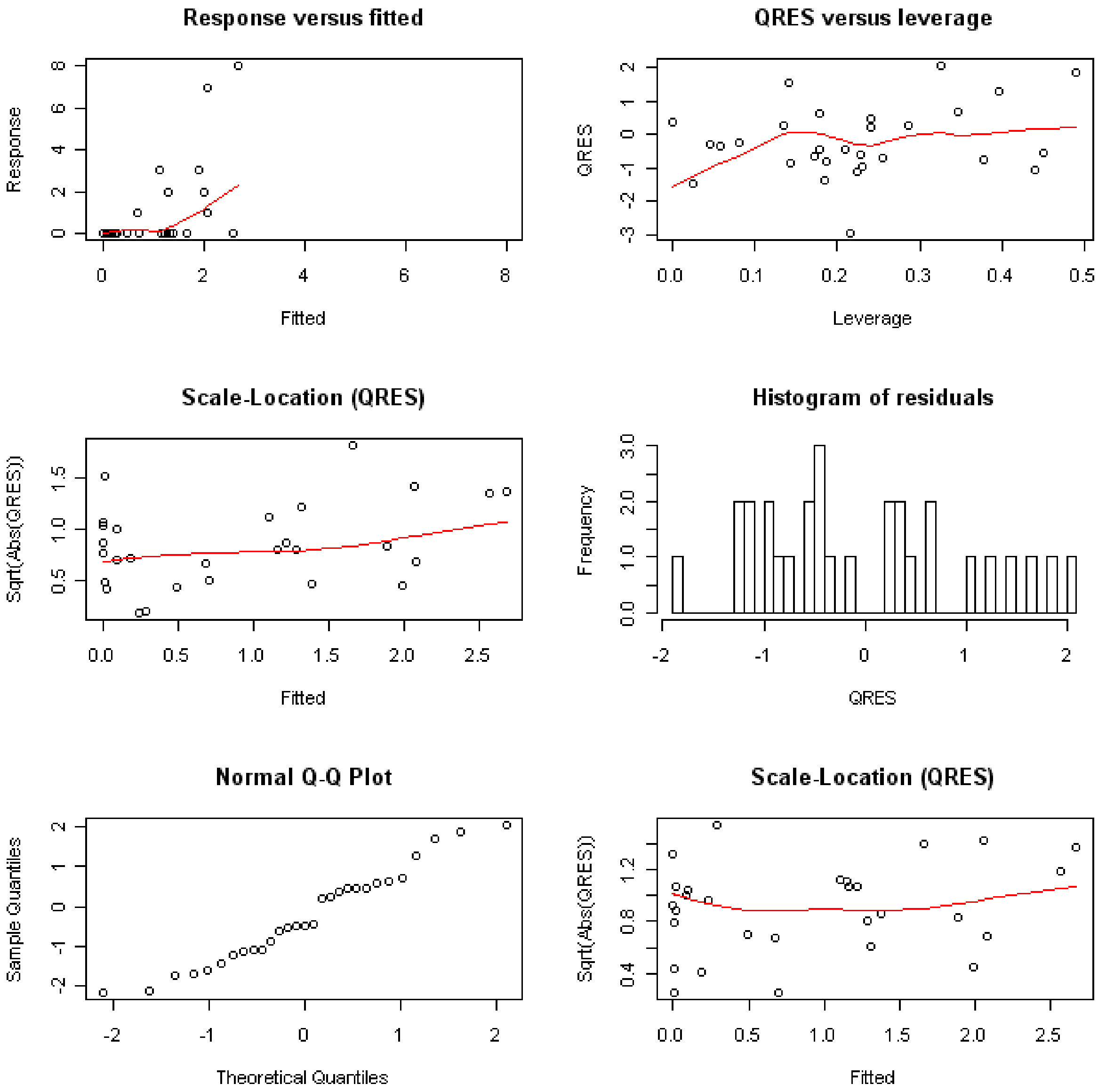

Response versus fitted
Response
Fitted
0
5
10
15
QRES versus leverage
QRES
Leverage
-2
-1
0
1
2
0.05
0.10
0.15
0.20
0.25
0.30
0.35
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Histogram of residuals
Frequency
QRES
0.0
1.0
2.0
3.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.2
0.6
1.0
1.4
0.5
1.0
1.5
2.0
2.5
3.0

SWWC Survey Region GAM model fit: Humpback whale

```
Family: Tweedie(1.1)
Link function: log

Formula:
N ~ s(midlat) + s(midlon) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -18.249      1.246  -14.65 5.39e-13 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
            edf Ref.df     F p-value
s(midlat) 1.000  1.000 0.251   0.621
s(midlon) 4.491  5.395 1.019   0.434

R-sq.(adj) =  0.169   Deviance explained = 48.1%
REML score = 32.267  Scale est. = 1.9415    n = 29
AIC = 71.56
```

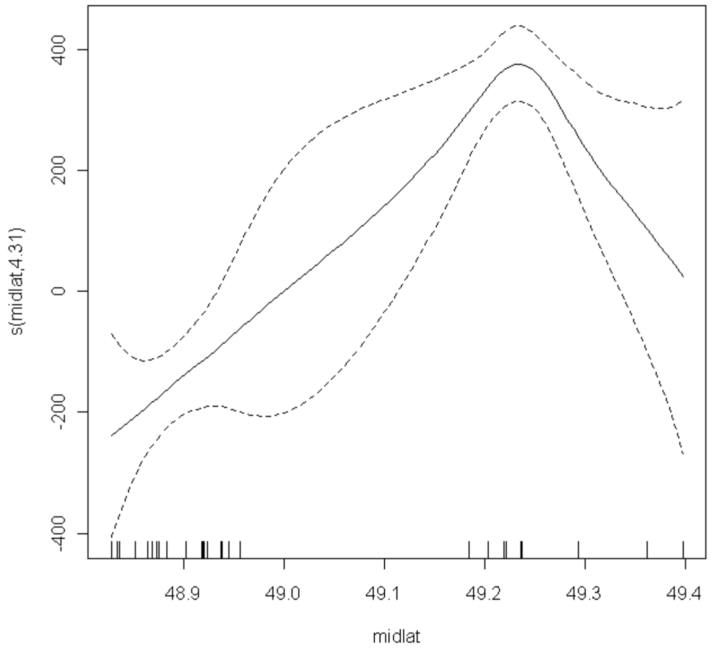


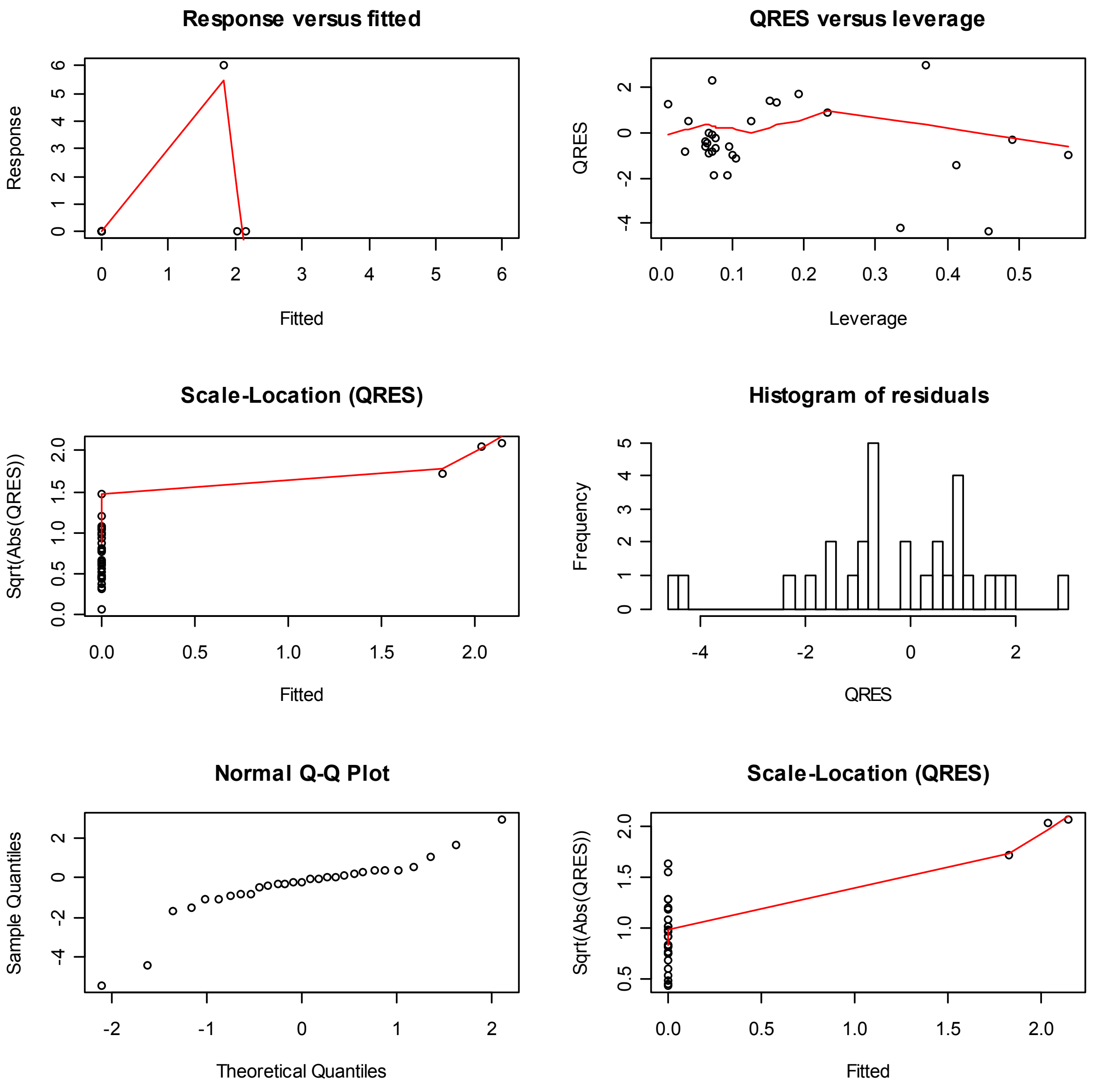



Response versus fitted
Response
Fitted
0
2
4
6
8
QRES versus leverage
QRES
Leverage
0.0
0.1
0.2
0.3
0.4
0.5
-3
-2
-1
0
1
2
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.0
0.5
1.0
1.5
2.0
2.5
Histogram of residuals
Frequency
QRES
0.0
1.0
2.0
3.0
-2
-1
0
1
2
Normal Q-Q Plot
Sample Quantiles
Theoretical Quantiles
-2
-1
0
1
2
Scale-Location (QRES)
Sqrt(Abs(QRES))
Fitted
0.4
0.8
1.2
0.0
0.5
1.0
1.5
2.0
2.5

SWWC Survey Region GAM model fit: Grey whale

```
Family: Tweedie(1.7)
Link function: log

Formula:
N ~ s(midlat) + offset(off.set)

Parametric coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -391.17      31.04  -12.60 5.41e-12 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Approximate significance of smooth terms:
            edf Ref.df     F  p-value
s(midlat) 4.307  5.008 27.01 4.63e-09 ***
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

R-sq.(adj) =  0.343   Deviance explained =   86.2%
REML score = 0.85174  Scale est. = 0.77397   n = 29
AIC = 45.13
```




Response versus fitted
Response
0
1
2
3
4
5
6
Fitted
QRES versus leverage
QRES
-4
-2
0
2
0.0
0.1
0.2
0.3
0.4
0.5
Leverage
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.0
0.5
1.0
1.5
2.0
Fitted
Histogram of residuals
Frequency
0
1
2
3
4
5
-4
-2
0
2
QRES
Normal Q-Q Plot
Sample Quantiles
-4
-2
0
2
-2
-1
0
1
2
Theoretical Quantiles
Scale-Location (QRES)
Sqrt(Abs(QRES))
0.5
1.0
1.5
2.0
0.0
0.5
1.0
1.5
2.0
Fitted