# Spatial analysis of marine mammal distributions and densities for supporting coastal conservation and marine planning in British Columbia, Canada

by

Gillian Kohl Allyson Harvey
B.A., University of Victoria, 2013

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Geography

# SUPERVISORY COMMITTEE

Spatial analysis of marine mammal distributions and densities for supporting coastal conservation and marine planning in British Columbia, Canada

by

Gillian Kohl Allyson Harvey
B.A., University of Victoria, 2013

**Supervisory Committee**

Dr. Trisalyn Nelson
Department of Geography, University of Victoria
School of Geographical Sciences & Urban Planning, Arizona State University
**Supervisor**

Dr. Caroline Fox
Department of Oceanography, Dalhousie University
Department of Geography, University of Victoria
**Additional Member**

Dr. Paul Paquet
Department of Geography, University of Victoria
**Departmental Member**

# ABSTRACT

**Supervisory Committee**

Dr. Trisalyn Nelson
Department of Geography, University of Victoria
School of Geographical Sciences & Urban Planning, Arizona State University
**Supervisor**

Dr. Caroline Fox
Department of Oceanography, Dalhousie University
Department of Geography, University of Victoria
**Additional Member**

Dr. Paul Paquet
Department of Geography, University of Victoria
**Departmental Member**

Human impacts on ocean ecosystems are driving declines in marine biodiversity, including marine mammals. Comprehensive spatial data are vital for making informed management decisions that may aid species recovery and facilitate the sustainable use of ocean ecosystems. However, marine mammal studies are often data limited, thereby restricting possible research questions. Developing novel analytical approaches and incorporating unconventional datasets can expand the scope of analysis by increasing the information content of existing data sources. The goal of our research is to support conservation and management of marine mammals in British Columbia (BC), Canada, through the application of advanced spatial statistical methodology to characterize spatial distribution and density patterns and provide assessments of data uncertainty.

Our first objective is to generate statistical models to map spatially continuous predictions of marine mammal distributions and densities within BC's north coast and apply methodology from spatial statistics to identify hotspots of elevated use. We use

species observations collected from systematic line transect surveys previously adjusted to generate estimates of density per nautical mile of transect. We predict the distribution and density patterns of nine marine mammal species by employing a species-habitat model to relate species densities to environmental covariates using a generalized additive model. We use spatial statistical hotspot analysis (Getis-Ord $G_i$*statistic) and an aspatial threshold approach to identify hotspots of high density. Our analysis reveals that hotspots selected using a top percentage threshold produced smaller and more conservative hotspots than those generated using the $G_i$*statistic. The $G_i$*statistic demonstrates a robust and objective technique for quantifying spatial hotspots and offers an alternative method to the commonly applied aspatial threshold measure. We find that maps show agreement with prior research and hotspots align with ecologically important areas previously identified by expert opinion.

Our second objective is to apply map comparison techniques to compare cetacean density maps from disparate data collection methods (systematic surveys and citizen science) to evaluate the information content of each map product and quantify similarities and differences. Discrepancies are quantified by performing image differencing techniques on the rank order values of each map surface. We subsequently use the $G_i$*statistic to isolate regions where extreme differences occur. To assess similarities, a $G_i$*statistic is applied to both maps to locate spatially explicit areas of high cetacean density. Where clusters of high density values in both maps overlap we infer higher confidence that the datasets are representing a true ecological signal, while areas of difference we recommend as targeted locations for future sampling effort. We

contextualize map similarities and differences using a dataset of human activity in the form of cumulative human effect scores.

Overall, our analytical approach integrates novel spatial datasets from systematic surveys, citizen science, and remote sensing to provide updated information on cetacean distributions in BC. Our study generates geographic data products that fill knowledge gaps and results provide baseline information valuable for future decision-making. The methodology applied in this study can be generalized across species and locations to support spatial planning and conservation prioritization in both marine and terrestrial contexts.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ACKNOWLEDGEMENTS

There are many individuals who have helped me accomplish both personal triumphs and academic achievements over the past two and a half years. I would like to begin by thanking my co-supervisor Dr. Trisalyn Nelson for her continuous encouragement and mentorship. I deeply valued being part of the SPAR lab and I thank her for fostering an inspiring and fun community, complete with wicked swag. I wish to express my gratitude to Dr. Caroline Fox, my co-supervisor, for her expert advice and for always bolstering my confidence when I was in doubt. I would also like to thank Dr. Paul Paquet for his insightful feedback and invaluable expertise provided throughout. I extend my gratitude to Raincoast Conservation Foundation, SPAR lab, and the Natural Sciences and Engineering Research Council of Canada for funding that made this project possible. To my fellow SPARtans and geography grad cohort, I am grateful for all of the amazing memories, moral support, and the countless motivational pep-talks that kept me sane. A special thanks to Mike Munroe for patiently enduring my stress ball moments and supporting me through this whirlwind. Finally, to my parents, family, and friends, I cannot thank you enough for all of the unwavering love and encouragement throughout my studies; I could not have done it without you.

# CO-AUTHORSHIP STATEMENT

This thesis is the combination of two scientific manuscripts for which I am the lead author. For these two papers, I performed all research, data analysis, initial interpretation of results, and manuscript preparation and presentation. The initial project structure and research questions for the first manuscript were developed together with Dr. Trisalyn Nelson, Dr. Caroline Fox, and Dr. Paul Paquet, who also provided assistance with analytical methodology, contextualizing results, and editorial comments. The initial project structure and research questions for the second paper were developed together with Dr. Trisalyn Nelson, while Dr. Caroline Fox as well as Dr. Paul Paquet provided guidance with editorial comments and suggestions incorporated into the final manuscript.

# 1.0 INTRODUCTION

## 1.1 Research context

Marine species are exposed to an increasing frequency and intensity of land, coastal, and marine-based stressors (Halpern et al., 2008) and result in rapid rates of extinction, extirpation, reduced abundance, and species range contractions across global and local scales (Schipper et al., 2008). Change in species population characteristics can result in alterations to ecosystem function, which may affect multiple trophic levels and ecological feedback loops (Chapin III et al., 2000; Dirzo et al., 2014; Heithaus et al., 2008). In the marine environment, top predators such as marine mammals, are particularly vulnerable to anthropogenic stressors, as direct and indirect impacts from multiple sources compound to generate cumulative effects. Human activities can have considerable negative impacts to marine mammal species, affecting both the physiological health and/or behaviour (Fair and Becker, 2000). Chronic and acute anthropogenically induced stress inflates the likelihood of morbidity in marine mammals, and in extreme cases will result in mortality. Marine mammals that frequent coastal habitats are exposed to particularly high risk, as human settlements are disproportionately skewed to locations near the coastline (within 100 km of the shore) and exhibit approximately 3x higher average population densities than the global average (Small and Nicholls, 2003).

Concern for the welfare of wildlife, including ocean species like marine mammals, has led to a call for data improvements as well as management strategies to mitigate impacts, plan for future changes, and better understand current situations and

phenomena. Concern has prompted the development of critical research questions to illuminate topics of high priority and urgency. For example, collaborative workshops recommend that future cetacean research should address ways to incorporate data from unconventional datasets (e.g., citizen science) into ecological research, monitor key activities on the ocean (e.g., cumulative human impacts), and develop approaches to combat data deficiency (Parsons et al., 2015). However, despite the need for comprehensive spatial data for marine species, the available knowledge of ocean wildlife, including marine mammals, is minimal compared to mammals in terrestrial habitats (Schipper et al., 2008), partially due to limited resources and the logistical complexities of collecting data at-sea. As a result, marine datasets are typically spatially and temporally inconsistent and species distribution information for marine mammals is often lacking for many ocean and coastal regions.

Spatial data are especially valuable for conservation prioritization, identifying geographic areas of importance for a given species, and generating a more comprehensive understanding of species ecology in an attempt to make better and more sustainable decisions (Franklin, 2010; Rodríguez et al., 2007). As anthropogenic threats rise there is increasing need to make informed management decisions and monitor progress towards national and international targets, such as the United Nations Convention on Biological Diversity and Sustainable Development Goals. The Convention on Biological Diversity includes 20 targets, described as the Aichi Targets, which outline various ways to conserve and promote sustainable use of global biodiversity. Aichi target 11 summarizes a goal that aims to have 10% of all marine and coastal regions protected by 2020 (Convention on Biological Diversity, 2013), while the

Sustainable Development Goals for year 2030 includes an environmental sustainability target that promotes the protection and sustainable use of ocean ecosystems (Sustainable Development Goals, 2015). It is, therefore, imperative that the best available spatial data for marine mammal species are available and effectively utilized to support appropriate evidence-based management and conservation decisions.

Space is a particularly important concept to consider when investigating ecological questions and hypotheses, especially when analyzing species distributions. Species and the resources that promote survival are patchy and vary in both distribution and intensity through space. This observation can be linked to the ecological understanding of environmental gradients and ecological niche theory (Grinnell, 1917; Hutchinson,1957). An ecological niche is defined when discontinuous environmental variables supports the presence of a species and positive growth rate (Hirzel and Lay, 2008). Therefore, the geographic distribution of a species is contingent on whether the environmental resources can support population growth, whether species can persist given competition and species interactions, and whether species can physically access the habitat (Hirzel and Lay, 2008). Consequently, the environmental mosaic of ocean conditions and patchy distribution of resources will influence the distribution and persistence of marine species and, therefore, gives weight to methodological techniques which incorporate species—habitat correlations and spatial relationships.

Methods available in spatial analysis, specifically spatial statistics, offer novel techniques to help answer ecological questions and fill knowledge gaps. Spatial statistics are distinctive as they incorporate spatial relationships into analyses from which conclusions regarding pattern and process can be inferred (Fotheringham and Rogerson,

2009). Patterns of species movements and aggregations are tightly coupled with the spatial composition of environmental and ecological processes. Changes to conditions will influence the distribution and abundance of individuals across space. Consequently, when investigating ecological questions, it is important to incorporate the linkage between pattern and process by applying spatial analytical methodology. Spatially explicit planning tools support decision-making by accounting for spatial data characteristics that may ultimately affect management choices. In our research, we incorporate spatial statistical methodologies that are unique to deal with complex ecological datasets and account for spatial variation.

### 1.2 Research focus

Currently, the declining status of many marine mammal populations in Canada (Favaro et al., 2014) warrants the development of strategies to reverse or slow the deteriorating condition of species and mitigate the detrimental effects of human actions. Scientific research informs management decisions and can illuminate where new regulations should be implemented and identify geographic locations for conservation prioritization. With Canadian research priorities describing the need to further our understanding of marine biodiversity and to identify hotspot areas of high diversity or function (Fissel et al., 2012), a critical step towards improved coastal management lies in advancing the use of spatial data, tools, and methodology in marine ecology. Effective management of marine species and our ability to aid in their recovery is contingent on quantitative spatial information, including understanding marine mammal species distribution and density patterns in Canada's coastal waters.

The status of many marine mammals in Canada continues to decline despite the protection afforded through the federal Species at Risk Act (SARA). The Canadian federal government provides assessments of marine mammal species in the form of national registries and lists from which the condition of marine species and populations can be tracked and monitored. The Committee On the Status of Endangered Wildlife in Canada (COSEWIC) provides ranking systems to identify at-risk species and develop recovery strategies to restore the health of marine mammal populations. For example, the Northeast Pacific northern resident population of Killer whales (*Orcinus orca*), have been classified as threatened since 2001 and have not improved in their ranking since (Fisheries and Oceans Canada, 2016), while the Pacific Ocean population of fin whales (*Balaenoptera physalus*) were listed as threatened in 2005 and have yet to have critical habitat defined (DFO, 2013). Updated spatial data on marine mammal species, particularly those listed as at-risk, and our relative confidence in the accuracy of these data, would be beneficial for future assessment and recovery actions.

With the advancement of Geographic Information Systems (GIS), remote sensing, and digital technology there exist opportunities to advance our understanding of the spatial characteristics of British Columbia's (BC) marine mammal distributions, densities, and important habitats despite limited and incomplete datasets. Consequently, incorporating new remotely sensed environmental datasets and establishing novel approaches for collecting species observations through unconventional methods, such as citizen science initiatives, are becoming more prevalent (Thiel et al., 2014). The diverse array of remotely sensed images of the earth provide more comprehensive representations of environmental processes and more ecologically relevant variables required for

accurately modeling distributions of marine species (Elith and Leathwick, 2009). In addition, due to the proliferation of mobile technology and internet availability, data collection by citizen scientists (volunteer participants with no formal scientific training) is not only feasible, but can contribute large quantities of information efficiently given limited resources.

Our research is focused on deriving the most complete and comprehensive spatial information from the best available data sources including systematic surveys, as well as unconventional data collected through citizen science initiatives. Given limited available data on species occurrence and locations of important habitat, collating datasets from various collection methods offers ways to validate existing information and identify geographic areas requiring additional sampling and investigation. As conservation funds are rapidly diminishing, species distribution and density mapping along with spatial statistical analyses are simple, yet effective, tools that can identify priority regions for conservation and areas for future research. Using spatially explicit approaches to inform conservation, research efforts can prevent the ineffective allocation of limited resources and provide a solid framework for future planning. Further, the application of spatial statistical techniques limits user bias and can provide robust assessments for data uncertainty. Drawing upon both systematic surveys and unconventional datasets from citizen science initiatives, our analysis aims to fill knowledge gaps of marine mammal distributions, densities, and regions of important habitat that are critical for making informed management decisions for coastal BC. Species-habitat modeling accompanied by spatial pattern statistics will present a comprehensive and rigorous analysis of the spatial characteristics and configurations of data deficient marine mammal species in BC.

### 1.3 Research goals and objectives

The overarching goal of this research is to inform conservation and marine spatial planning in BC, Canada by employing spatial analytical methodologies to quantify marine mammal distribution and density patterns, identify ecologically important areas, and assess confidence in the information content of species density maps generated from disparate data collection techniques. The aim of this research is to achieve research goals by completing the following measurable objectives:

1) Develop a predictive species-habitat model to map the distribution and densities of marine mammals on BC's north coast and, using spatial statistical methods, quantify the spatial distributions and density patterns in order to identify hotspots of intense marine mammal use.

2) Apply map comparison techniques to quantify agreement and disagreement between marine mammal density maps generated from disparate data collection techniques (traditional surveys and citizen science) to provide validation and assess both the information content and potential application of novel datasets in a marine context.

### References

Chapin III, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L., Reynolds, D. U. Hopper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Días. 2000. Consequences of changing biodiversity. Nature 405:234—242.

Convention on Biological Diversity. 2013. Quick guides to the Aichi biodiversity targets, version 2. United Nations. www.cbd.int (accessed 15 Feb 2016).

DFO. 2013. Report on the Progress of Recovery Strategy Implementation for Blue, Fin and Sei Whales (*Balaenoptera musculus, B. physalus* and *B. borealis*) in Pacific Canadian Waters for the Period 2006-2011. *Species at Risk Act* Recovery Strategy Report Series. Fisheries and Oceans Canada, Ottawa. v + 10 pp.

Dirzo, R., H. S. Young, M. Galetti, G. Ceballos, N. J. Isaac, and B. Collen. 2014. Defaunation in the Anthropocene. Science 345:401–406.

Elith, J., and J. R. Leathwick. 2009. Species distribution models: ecological explanation and prediction across space and time. Annual Review of Ecology, Evolution, and Systematics 40:677—697.

Fair, P. A., and P. R. Becker. 2000. Review of stress in marine mammals. Journal of Aquatic Ecosystem Stress and Recovery 7:335—354.

Favaro, B., D. Claar, C. Fox, C. Freshwater, J. Holden, and A. Roberts. 2014. Trends in Extinction Risk for Imperiled Species in Canada. PLoS ONE 9:e113118.

Fisheries and Oceans Canada. 2016. Report on the Progress of Recovery Strategy Implementation for the Northern and Southern Resident Killer Whales (*Orcinus orca*) in Canada for the Period 2009 – 2014. *Species at Risk Act* Recovery Strategy Report Series. Fisheries and Oceans Canada, Ottawa. iii + 51 pp.

Fissel, D., M. Babin, R. Bachmayer, K. Denman, E. Dewailly, K. M. Gillis, L. Fortier, R. Hyndman, D. Lane, M. Lewis, R. Macdonald, K. Moran, B. Neis, M. Nuttall, É. Pelletier, L. Ridgeway, S. Roussel, P. Snelgrove, W. J. Sutherland, C. Suttle, D. Wallace, and M. G. Wiber. 2012. 40 Priority Research Questions for Ocean Science in Canada: A Priority-setting Exercise by the Core Group on Ocean Science in Canada. Council of Canadian Academies. Ottawa, ON.

Fotheringham, S., and P. A. Rogerson. 2009. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Franklin, J. 2010. Species distribution modeling. Chap. 1. Pages 3—20 in Mapping Species Distributions: Spatial Inference and Prediction. Cambridge University Press, Cambridge, UK.

Grinnell, J. 1917. Field tests of theories concerning distributional control. The American Naturalist 51:115—128.

Halpern, B. S., S. Walbridge, K. A. Selkoe, C. V. Kappel, F. Micheli, C. D'Agrosa, J. F. Bruno, K. S. Casey, C. Ebert, H. E. Fox, R. Fujita, D. Heinemann, H. S. Lenihan, E. M. P. Madin, M. T. Perry, E. R. Selig, M. Spalding, R. Steneck, and R. Watson. 2008. A global map of human impact on marine ecosystems. Science 319:948—952.

Heithaus, M. R., A. Frid, A. J. Wirsing, and B. Worm. 2008. Predicting ecological consequences of marine top predator declines. Trends in Ecology and Evolution 23:202—210.

Hirzel, A. H., and G. L. Lay. 2008. Habitat suitability modelling and niche theory. Journal of Applied Ecology 45:1372—1381.

Hutchinson, G. E. 1957. Concluding remarks. Cold Springs Harbour Symposium on Quantitative Biology 22:415—427.

Parsons, E. C. M., S. Baulch, T. Bechshoft, G. Bellazzi, P. Bouchet, A. M. Cosentino, C. A. J. Godard-Codding, F. Gulland, M. Hoffmann-Kuhnt, E. Hoyt, S. Livermore, C. D. MacLeod, E. Matrai, L. Munger, M. Ochiai, A. Peyman, A. Recalde-Salas, R. Regnery, L. Rojas-Bracho, C. P. Salgado-Kent, E. Slooten, J. Y. Wang, S. C. Wilson, A. J. Wright, S. Young, E. Zwamborn, W. J. Sutherland. 2015. Key research questions of global importance for cetacean conservation. Endangered Species Research 27:113—118.

Rodríguez, J., L. Brotons, J. Bustamante, and J. Seoane. 2007. The application of predictive modelling of species distribution to biodiversity conservation. Diversity and Distributions 13:243—251.

Schipper, J., J. S. Chanson, F. Chiozza, N. A. Cox, M. Hoffmann..., and B. E. Young. 2008. The status of the world's land and marine mammals: diversity, threat, and knowledge. Science 322:225—230.

Small, C. and R. J. Nicholls. 2003. A global analysis of human settlement in coastal zones. Journal of Coastal Research 19:584—599.

Sustainable Development Goals. 2015. Sustainable Development Goals: 17 Goals to Transform our World. United Nations. http://www.un.org/sustainabledevelopment/sustainable-development-goals/ (Accessed 31 Oct 2016).

Thiel, M., A. Penna-Días, G. Luna-Jorquera, S. Salas, J. Sellanes, W. Stotz. 2014. Citizen Scientists and Marine Research: Volunteer Participants, their Contributions, and Projection for the Future. Oceanography and Marine Biology: An Annual Review 52:257—314.

# 2.0 QUANTIFYING MARINE MAMMAL HOTSPOTS IN BRITISH COLUMBIA, CANADA

## 2.1 Abstract

Global biodiversity is undergoing rapid decline due to direct and indirect anthropogenic impacts to species and ecosystems. Marine species, in particular, are experiencing accelerated population declines leading to many species being considered at-risk by regional, national, and international standards. As one conservation approach, decisions made using spatially explicit information on marine wildlife populations have the potential to facilitate recovery and contribute to national and international commitments towards conservation targets. Delineating areas of intense use by species at-risk can inform future Marine Spatial Planning (MSP) and conservation efforts, including the identification of Marine Protected Areas (MPAs). Methods for detecting hotspots (e.g., areas with high density and/or abundance) enable categorical mapping of the most intensely used areas. Yet, many of the current methods for delineating hotspots, such as the top 5% threshold, are subjective and fail to account for spatial patterns. Our goal was to map spatially continuous distributions of marine mammal species densities and employ quantitative statistical methods to extract hotspot locations on the northern coast of British Columbia. We integrated systematically surveyed species information with remote sensing variables using Generalized Additive Models (GAMs) to predict marine mammal distribution and density. Hotspots were identified from the density surfaces using two approaches: aspatial top 5% method and spatially local $G_i$* statistic using three neighbourhood definitions. The $G_i$*statistic incorporates spatial relationships

of the data into hotspot detection, while the threshold approach is solely based on an arbitrarily defined number. Heterogeneous density patterns were observed for all species and high-density regions were generally clustered in areas exhibiting oceanographic characteristics that may promote concentrated food resources. Combining species density surfaces and extracting hotspot locations identified regions important to multiple species and presents candidate locations for future conservation efforts. Contributions from this research provide robust statistical methods to objectively map hotspot locations and generate GIS data products for informing coastal conservation decisions.

## 2.2 Introduction

To address overwhelming evidence that human actions are directly, and indirectly, contributing to rapid marine species population declines, many national and regional agencies facilitate marine conservation initiatives through management frameworks, strategies and tools. Many of these programs incorporate Marine Spatial Planning (MSP), Ecosystem Based Management (EBM), Systematic Conservation Planning (SCP), and Marine Protected Areas (MPAs). As a key example, a number of countries, including Canada, pledged to meet biodiversity conservation targets (Aichi Targets) outlined at the Convention of Biological Diversity in 2010 in Nagoya, Japan. The 11$^{th}$ target designates a minimum of 10% of global marine and coastal waters as protected by 2020, which includes ecologically important habitats and regions of high conservation value (Convention on Biological Diversity, 2013). To achieve conservation and management targets, it is essential to possess baseline information on marine species in order to identify where species occur in elevated densities and locate highly utilized areas that could indicate potential priority regions for conservation. In addition, baseline data on marine species helps avoid the shifting baseline syndrome (Pauly, 1995) and provides a foundational benchmark to measure and assess conservation effectiveness.

Ecological data on focal taxa, such as marine mammals, can be used as indicators to prioritize important marine regions (Zacharias and Roff, 2001), as well as delineate explicit boundaries for reserves, sanctuaries, and protected areas (Hooker and Gerber, 2004). However, due in part to logistical and financial challenges with data collection in ocean environments, marine mammal species are typically the focus of fewer scientific publications than terrestrial mammals (Kovacs et al., 2012; Schipper et al., 2008). The

IUCN red list indicates that over half of listed cetaceans are globally data deficient (IUCN, 2016). Pinnipeds, on the other hand, have no data deficient listings, yet 53% of species listed have declining or unknown population trends (IUCN, 2016).

Predictive species-habitat models can be used to map species distributions and densities from survey data, and are often used to fill knowledge gaps. Species observations, which are discontinuous in nature, are related to continuously distributed environmental variables through statistical models to predict species occurrence at unsampled locations (Franklin, 2010). Models produce spatially continuous maps of species distribution and/or abundance, making regions where species aggregate more apparent. For example, many marine mammals, particularly specialist species, are sensitive to changes in environmental conditions and alterations in marine food webs, where species richness is often highly coupled with primary productivity and food availability (Schipper et al., 2008; Preikshot et al., 2013). Associations between environmental conditions and species distributions and densities allow species richness to be predicted at locations where surveys have not been completed. The patchy nature of species distributions within continuous space provides valuable insight into what environmental conditions may drive the observed spatial patterns while also identifying clustered regions of intense use by one, or more, species. Maps showing variation in species abundance contribute to the overall understanding of where organisms are located and which regions may exhibit elevated levels of species abundance or density (i.e., hotspots; see Reese and Brodeur, 2006 [nekton organisms], Menza et al., 2016 [seabirds, pinnipeds and cetaceans] and Nur et al., 2011 [seabirds]).

Mapping hotspots of marine species has important opportunities for the allocation of scarce conservation and planning resources, particularly when striving to achieve conservation objectives such as the Aichi Targets. Identifying areas that have a higher concentration of species than surrounding areas is essential when developing conservation policies (Hyrenbach et al., 2000). An effective strategy for marine megafaunal protection is to situate reserves around productive regions where species abundance is high in comparison to surrounding ocean habitat (Hooker and Gerber, 2004).

Hotspots with spatially explicit boundaries can be detected using methods that are both aspatial and spatial. In conservation biology, aspatial approaches to hotspot delineation are the most common and apply an arbitrary threshold, such as the top 2.5% (Ceballos and Ehrlich, 2006; Orme et al., 2005), 5% (Parviainen et al., 2009; Tolimieri et al., 2015) or 10% (Tolimieri et al., 2015), to information, typically species richness measures, in order to partition hotspot locations. Some thresholds reach as high as 25% and 50% (e.g., Nur et al., 2011). However, movement in this discipline has shifted towards acknowledging spatial dependence in ecological datasets and incorporating spatially explicit methodology to understand spatial relationships (Liebhold and Gurevitch, 2002; Wagner and Fortin, 2005). Spatial methods for hotspot delineation enable thresholds to be statistically determined and account for the spatial pattern of species distributions (Nelson and Boots, 2008). More specifically, using local measures of spatial autocorrelation, it is possible to map where species are most abundant and where the spatial pattern of species distributions are unlikely to have arisen from chance processes (Anselin 1995; Ord and Getis, 2001; Boots, 2002). While spatially local

hotspot detection methods have been applied in terrestrial contexts (e.g., Nelson and Boots, 2008 and Zhihai et al., 2012), studies that apply spatially explicit hotspot detection have more recently expanded to include marine research (e.g., Nelson et al., 2011 and Kuletz et al., 2015).

The goal of this paper is to quantify hotspots and explore multiple techniques for hotspot identification using continuous species density surfaces in order to inform future policy on potential marine conservation areas. We expect spatial hotspots to be found in locations with elevated ocean productivity and primary production. We build upon existing baseline information for marine mammals in British Columbia (see Best et al., 2015; Williams et al., 2007, 2010 and 2011) by generating continuous density surfaces for nine species. We applied a correlative modelling technique, Generalized Additive Models (GAMs), to support our hypothesis that high density species will be situated in favourable habitat conducive to increased prey availability and, therefore, offer prime locations for future conservation and protection measures. The resulting predicted surfaces are used to illustrate the utility of performing spatially specific methodology for extracting hotspot locations for spatial planning and conservation initiatives. Priority zones for conservation are identified by employing multiple approaches for hotspot identification. Comparison between aspatial and spatially local methodologies offer new perspectives on how various techniques for hotspot delineation influences both location and physical characteristics of hotspots when used in a marine context.

## 2.3 Methods

### 2.3.1 Study region and species

Our research area is situated within the continental shelf of British Columbia, Canada covering a region of 62,976 km$^2$ (Figure 2.1). The northern coastal waters of British Columbia (Queen Charlotte Basin) contain diverse spatial and temporal biophysical oceanographic characteristics (Thomson, 1981) as well as a multitude of anthropogenic activities (Ban and Alder, 2008; Clarke Murray et al., 2015). This region is highly productive, providing food resources and important foraging opportunities for migrating and resident marine mammals.

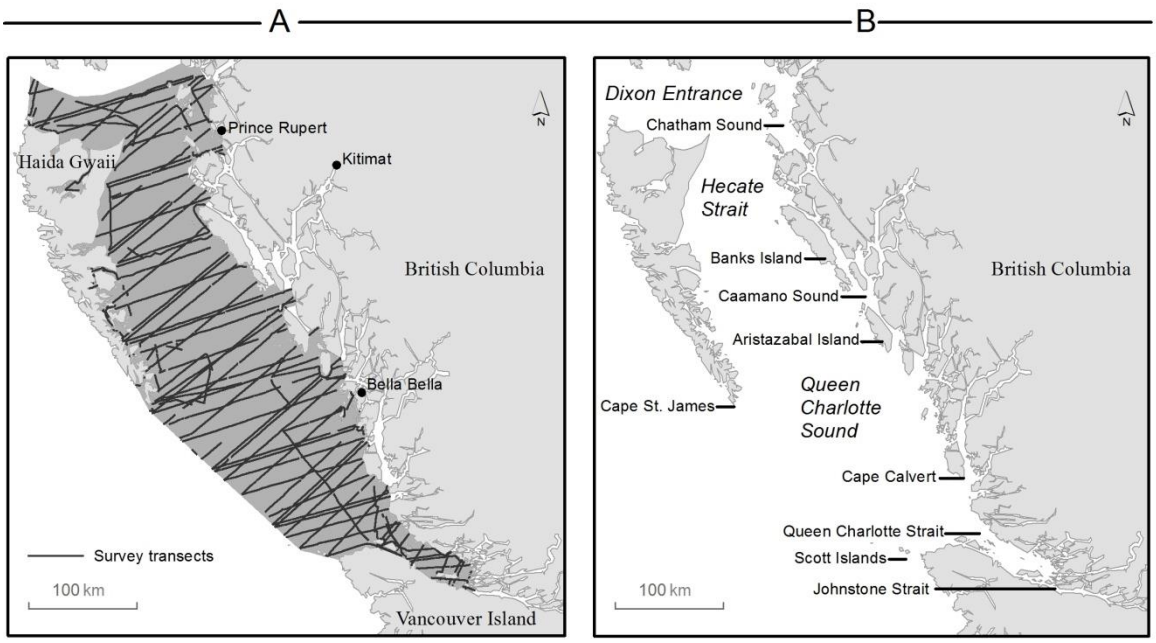


Figure 2.1 Maps illustrating (A) study region that is indicated in dark grey with passage and on-effort survey transects (2004-2008) and (B) key oceanographic regions.

Of the 24 extant marine mammal species currently found in British Columbian waters (Ford, 2014), this study focuses on nine species (seven cetaceans and two pinnipeds). Species were selected based on number of presence observations to maintain

sufficient quantity of non-zero samples for modelling. Cetaceans include common minke whale (MW; *Balaenoptera acutorostrata*), Dall's porpoise (DP; *Phocoenoides dalli*), fin whale (FW; *Balaenoptera physalus*), harbour porpoise (HP; *Phocoena phocoena*), humpback whale (HW; *Megaptera novaeangliae*), killer whale (KW, three ecotypes; *Orcinus orca*), and Pacific white-sided dolphin (PW; *Lagenorhynchus obliquidens*), while pinnipeds are comprised of harbour seal (HS; *Phoca vitulina*) and Steller sea lion (SSL; *Eumetopias jubatus*) species. Of the chosen study species, more than half are listed as provincially and nationally at-risk (Table 2.1).

Table 2.1 Provincial, national and global rankings of study species as of September 1, 2016

- **Yellow** = apparently secure, **Blue** = special concern, **Red** = extirpated, endangered or threatened
- **NAR** = not at risk**, SC** = special concern, **T** = Threatened
- **LC** = least concern, **DD** = data deficient, **NT** = near threatened, **EN** = endangered

| Common name | Scientific name | BC list (BC) | COSEWIC (Canada) | SARA (Canada) | IUCN (global) |
|---|---|---|---|---|---|
| Common minke whale | *Balaenoptera acutorostrata* | Yellow | NAR (2006) | N/A | LC (2008) |
| Dall's porpoise | *Phocoenoides dalli* | Yellow | NAR (1989) | N/A | LC (2008) |
| Fin whale | *Balaenoptera physalus* | Red | T (2005) | T (2006) | EN (2008) |
| Harbour porpoise | *Phocoena phocoena* | Blue | SC (2016) | SC (2005) | LC (2008) |
| Harbour seal | *Phoca vitulina* | Yellow | NAR (1999) | N/A | LC (2016) |
| Humpback whale | *Megaptera novaeangliae* | Blue | SC (2011) | T (2005) | LC (2008) |
| Killer whale | *Orcinus orca* | | | | DD (2008) |
| Northeast Pacific offshore | | Red | T (2008) | SC (2003) | |
| Northeast Pacific transient | | Red | T (2008) | T (2003) | |
| Northeast Pacific northern resident | | Red | T (2008) | T (2003) | |
| Pacific white-sided dolphin | *Lagenorhynchus obliquidens* | Yellow | NAR (1990) | N/A | LC (2008) |
| Steller sea lion | *Eumetopias jubatus* | Blue | SC (2013) | SC (2005) | NT (2016) |

### 2.3.2 Species data

From 2004-2008 Raincoast Conservation Foundation conducted one of BC's largest systematic at-sea surveys. Surveys were stratified into 4 regions from which transects were randomly placed in a crisscrossing pattern. The stratified survey was

specifically designed to promote effort efficiency and to maintain a random placement of transects (Thomas et al., 2007). This study includes only data collected within the fourth stratum (Figure 2.1). Surveys over the 5 years were conducted via line transects and included six time periods: Summer 2004 (June, July and August), Summer 2005 (August), Summer 2006 (August and September), Spring 2007 (April and May), Fall 2007 (October and November), and Summer 2008 (June and August). The Fall survey was not included due to lack of sighting data. More than 16,000 km of trackline – over 5,000 km within the study region – was surveyed to generate distance corrected (Buckland et al., 2001; Buckland et al., 2004) quantitative information for a number of marine mammal species using Multiple-Covariate Distance Sampling (MCDS) techniques (Best et al., 2015). All transects were separated into one nautical mile segments and, for each segment, a species density estimate was calculated (see Best et al., 2015). These estimates are particularly robust as they have been corrected for uncertain observer sightings along each segment using detection functions (Buckland et al., 2004; Headley and Buckland, 2004). For marine species, where the majority of their body mass is below the water surface uncertainty in observer sightings is unavoidable. Therefore, correcting for imperfect detection is vital to ensuring the most accurate assessment of species densities.

Imperfect or failed detection of target species often leads to zero inflation of surveyed datasets. As marine mammals spend the majority of their time under the water surface, there is potential for observers to miss sighting a species when, in fact, they were present. Missed detections increase the number of “false zero” observations (i.e., false absences) and when coupled with “true zeros” there is a drastic increase in the total

number of zeros for a given species. In this study, we refer to zero inflated data as data which possess both true and false zeros and that, when combined, lead to an extreme number of zeros compared to non-zero values. Consequently, zero inflated data are difficult to model, as they mask the true variability in the species data and may lead to poor model performance. For additional details on survey methods, see Thomas et al. (2007) and Williams et al. (2007) and for more information on MCDS and abundance estimates along transect segments, see Best et al. (2015).

As marine mammal sightings were limited within the study region, all seasons and years (for both passage and on-effort transects) were combined to maximize the sample size for the modelling process. Pinniped observations both in water and hauled out on land were also collated. Vessel speeds remained relatively consistent at approximately 15 km/h throughout the survey extent; therefore, speeds $\leq$ 5 knots were removed from analysis to minimize bias.

### 2.3.3 Environmental covariates

Top predators, such as marine mammals, respond less strongly to short-term oceanographic conditions when assessed in transect based habitat-models, while proclivity for broader scale and predictable oceanographic features has been shown (Mannocci et al., 2013). Therefore, temporally static and monthly averaged composites, as well as longer-term climatologies were used in this analysis (Table 2.2). The 15 environmental covariates used to characterize marine mammal habitat were chosen based on data availability and spatial coverage for the study region. These can be classified into three categories: static, dynamic, and climatological. Static variables are those that are geographically fixed and/or temporally static. Dynamic and climatological variables are

time-averaged composites, monthly and yearly averaged multi-decadal periods respectively.

Static variables include latitude, longitude, depth (m), slope (degrees), benthic terrain ruggedness (proportion), distance from the coast (m), distance from high current areas (m), and distance from the continental shelf (m). Latitude and longitude were recorded by a Global Positioning System (GPS) every 10 seconds during survey transiting and were collected using software Logger 2000. Slope and benthic terrain ruggedness were calculated from a depth (bathymetry) 100-meter resolution grid sourced from the BC Marine Conservation Atlas (SciTech Consulting and Living Oceans Society; www.bcmca.ca). Benthic terrain ruggedness was created using the Benthic Terrain Modeler extension from Geospatial Modeling Environment (Wright et al., 2012) using a 13 cell moving window. Euclidean distance to the coastline, high current regions (> 3 knot current), and continental shelf were generated from layers from the British Columbian provincial government and calculated over a 50m grid for the study extent (Freshwater Atlas Coastlines; Benthic Marine Ecounits; apps.gov.bc.ca). Continental shelf polygons were delineated by selecting regions with depth between 200 and 1000 m and a 5-20° slope.

Dynamic variables consist of root mean square of average tidal speed ($m \cdot s^{-1}$), sea surface temperature (SST, $^{o}C$), chlorophyll-a concentration ($mg \cdot m^{-3}$), wind speed ($m \cdot s^{-1}$), sea surface height (SSHA) and sea level anomaly (SSHD). Dynamic predictors represent oceanographic conditions that are temporally variable and, therefore, remotely sensed imagery provides an averaged temporal composite for each. Monthly images, aside from the root mean square average tidal speed, were extracted for the study region from the

Table 2.2 Summary of the 15 original environmental covariates for GAM model.

| Category | Variable | Summary | Resolution | Rationale |
|---|---|---|---|---|
| Static | Latitude and Longitude[1] | Derived from transect GPS data (m) | 50m | Spatial location shows strong influence on predictions using species distribution models (Best et al., 2015). |
| | Bathymetry[2] | Depth of ocean floor (m) | 100m | Top predators show response to bathymetric features; shallow topography may provide favourable foraging opportunities (Yen et al., 2004). |
| | Slope[2,3] | Slope (degrees) of the ocean floor derived from bathymetry data. | 100m | Steep benthic relief promotes water movements, which increase and concentrate prey and/or primary production (Croll et al., 1998; Yen et al., 2004). |
| | Benthic terrain ruggedness[2,4] | Terrain ruggedness derived from bathymetry data (proportion) | 100m | Topographic complexity, such as rugosity, can create localized increases in productivity, aid in prey capture, and provide migration cues (Bouchet et al., 2015). |
| | Distance from coast[3,5] | Euclidean distance from nearest coastline feature (m) | 50m | Distance provides an indication of preference for near or offshore habitats, e.g., distance to land used as covariate in humpback whale model (Dalla Rosa et al., 2012). |
| | Distance from continental shelf[3,6] | Euclidean distance (m) from continental shelf (200-1000m depth and slope between 5-20%) | 50m | Continental shelf edge is characterized by upwelling and water column mixing promoting high productivity, prey, and overall biomass (Croll et al., 1998; Springer et al., 1996). |
| | Distance from high current areas[3,6] | Euclidean distance (m) from high current polygons (>3 knot current) | 50m | Productivity of areas (e.g., upwelling regions) driven by current strength or persistent eddy circulations (Smith and Whitehead, 1993; Whitney et al., 2005). |
| Dynamic | Tidal current[7] | Root mean square average tidal speed ($m \cdot s^{-1}$) | 500m | Strong tidal currents influence ocean circulation leading to elevated nutrients and prey concentrations particularly in coastal ecosystems (Rogachev et al., 2008). |
| | Sea surface temperature[8] | Monthly averaged AquaMODIS daytime sea surface temperature (ºC) | 0.05 degrees | Cold coastal surface waters may indicate upwelling regions (Croll et al., 1998; Jardine et al., 1993). |
| | Chlorphyll-a concentration[8] | Monthly averaged AquaMODIS Chlorophyll-a concentrations ($mg \cdot m^{-3}$) | 0.05 degrees | High chlorophyll-a concentrations indicate regions of high prey concentrations and are often used as a proxy for primary productivity (Ware and Thomson, 2005). |
| | Wind[8] | Magnitude of monthly averaged QuikSCAT sea surface wind speed ($m \cdot s^{-1}$) | 0.125 degrees | Wind induced water column mixing impacts eddy characteristics and strength, as well as the distribution and abundance of ocean productivity (Brodeur and Ware, 1992; Stammer and Wunsch, 1999). |
| | Sea height absolute[8] | Monthly averaged AVISO sea surface height deviation plus the long-term mean dynamic height (m) | 0.25 degrees | Indicates areas of ocean movement, mixing and variability, which may represent possible regions of enhanced ocean productivity (Rao et al., 2006). |
| | Sea height deviation (Sea level anomaly)[8] | Monthly averaged AVISO sea surface height deviation from the mean geoid as measured from 1993-1995 (m) | 0.25 degrees | Anomalies in sea level can be used to identify eddies, which create conditions that generate food rich habitats (Crawford et al., 2007; Tosh et al., 2015). |
| Climatological | Temperature[9] | Long-term monthly averaged sea surface temperature (°C) from 1955-2006 | 0.25 degrees | Distribution of top predators may be, in part, influenced by temperature as predators and/or their prey have varying thermal preferences (Block et al., 2011). |
| | Salinity[9] | Long-term monthly averaged sea surface salinity (ppm) from 1955-2006 | 0.25 degrees | Fresh water runoffs that are high in nutrients stratify the water and may affect the growth of algae (Campagna et al., 2008). Some marine mammals have been shown to avoid low salinity areas (Tynan, 2005). |

*List of data sources: Raincoast Conservation Foundation (1); SciTech Consulting and Living Oceans Society; www.bcmca.ca (2); ArcGIS 10.0 tools (3); Benthic Terrain Modeler extension (Wright et al., 2012) (4); DataBC, Freshwater Atlas Coastlines, apps.gov.bc.ca (5); DataBC, Benthic Marine Ecounits – Coastal Resource Information Management System, apps.gov.bc.ca (6); Foreman et al., 2000, www.bcmca.ca (7); NOAA CoastWatch, www.coastwatch.pfeg.noaa.gov, accessed through Environmental Data Connector for ArcGIS (8); World Ocean Database, www.nodc.noaa.gov (9).

National Oceanic and Atmospheric Administration (NOAA) CoastWatch program (www.coastwatch.pfeg.noaa.gov) for each month surveyed during the six selected survey periods. The root mean square average tidal speed was provided for the entire west coast of Canada and was generated though a 3D circulation model for coastal regions of the Northeastern Pacific Ocean (Foreman et al., 2000; www.bcmca.ca).

SST and salinity climatological variables are long-term multi-decadal monthly averages using data from 1955-2006, which represent general oceanographic trends. Both datasets were sourced from the World Ocean Database (http://www.nodc.noaa.gov).

### 2.3.4 Data preprocessing

Data were integrated using a hexagon grid, with a spatial resolution of 13.86 $km^2$ to allow integration with Environment and Climate Change Canada marine planning units (e.g., Fox et al., In Review). Hexagons have been extensively used in marine spatial planning as they allow for more efficient, compact (Nhancale and Smith, 2011) and ecologically relevant configurations (Birch et al., 2007). Hexagons were attributed with mean covariate values; however, dynamic and climatological variables are monthly composites therefore enabling the calculation of additional values by pooling across survey years. Calculations included the coefficient of variation (CV), minimum (min) and maximum (max) values. This resulted in an increase of environmental predictor variables used in modelling from 15 to 34. When data were missing from remotely sensed variables (within inlets, near shore, and due to cloud cover) values were interpolated using the nearest neighbor value.

Inclusion of correlated variables in models can result in reductions in model performance and overall model instability (Kuhn and Johnson, 2013). We tested for

correlation between the 34 covariates using the Spearman's rank correlation (rho) analysis by applying the rcorr function of the Hmisc package in R (Harrell Jr., 2016; R Core Team, 2015). Relationships between variables were assessed in descending order of the absolute value of each correlation coefficient over a given threshold (Kuhn and Johnson, 2013). Here we chose a correlation coefficient threshold of $r_s > 0.70$ (Dormann et al., 2013) with a conservative statistical significance level of 0.01. The variable with the largest average correlation coefficient was removed. This process was repeated until all correlation coefficients fell below the set threshold. The remaining covariates were as follows: longitude, latitude, bathymetry, terrain ruggedness, distance to coastline, distance to high current regions, distance to continental shelf, average tidal current, chlorophyll-a concentration (min, max, CV), SST (min, max, CV), SSHA (min), SSHD (avg, min, max), wind (avg, min, max), and salinity (max).

### 2.3.5 Modelling approach

Regression based predictive models are a popular technique for modelling cetacean distributions (Redfern et al., 2006). We used a GAM model to account for non-linear and non-monotonic trends, which are common in ecological studies (Guisan et al., 2002; Hastie and Tibshirani, 1990; Wood, 2006). Further, the spline generated from the GAM model allows for a more clear detection of the ecological signal in the presence of large quantities of zeros (real or false) by effectively representing the variance in the data distribution. A basic GAM model can be expressed as:

$$g(\mu) = \alpha + \sum_{j=1}^{p} f_j(X_j) \quad (1)$$

where the intercept is represented by α and g(μ) is the 'link' function that correlates the mean of the estimated response to the sum of all 'smooth' functions ($f_j$) for each covariate value ($X_j$) (Hastie and Tibshirani, 1990).

In this modelling approach we use a GAM model to relate species density per nautical mile to the 22 environmental variables selected from the correlation analysis. The GAM model applies penalized regression splines using the mgcv package within R (R Core Team, 2015; Wood, 2011). We used a thin plate regression spline as the smooth function where the smoothing parameters used to control the degree of smoothness (wiggliness) of the fitted spline were estimated through generalized cross validation (GCV). GCV was used rather than the Unbiased Risk Estimator (UBRE) because the scale parameter was unknown (Wood, 2006). To control for the tendency of GCV to overfit data, the degrees of freedom were modified from the default gamma value of 1 to 1.4 (Kim and Gu, 2004) and basis functions were further penalized by reducing the k value (total allowable degrees of freedom for each spline) to 6 from the default 10. An additional penalty was added through the "select" function where covariates may be automatically removed from the model during fitting.

A weighting scheme was applied to compensate for zero-inflated species data, where the greater the weight value, the more emphasis that particular observation is given within the model (Wood, 2016). To select the optimal weight value a comparison analysis was performed using a covariate saturated GAM model for each species. Though the percent variance explained increased for higher weight values, the analysis demonstrated minimal improvements (<5% variance explained) beyond a weight of 10. As a result, a weight of 10 was applied to all non-zero observations.

To generate a parsimonious model, covariates were removed in a backwards selection procedure, beginning with variables with the highest p-value. Variables were removed until all were significant from a 0.05 significance level. Model performance was assessed by examining the percentage of explained deviance and the adjusted $R^2$, while the root mean square error (RMSE) of observed vs. predicted values was used to assess accuracy of model predictions.

Model predictions resulted in negative values for some species, however, for visualization purposes all negative values were displayed as zero. Normalized species density maps were generated by dividing the predicted density values by the maximum predicted value for each species resulting in a range between 0 and 1 (similar to Nur et al., 2011 and Fox et al., In Review). Mapping relative densities in the form of a normalized numeric rather than absolute densities prevents one species from driving any hotspots identified when individual species maps are combined. Within-species normalized maps were collated together by summing the normalized values across cetaceans, pinnipeds and all species for each hexagon. The three collated maps were subsequently used in hotspot analysis.

### 2.3.6 Hotspot analysis

To identify hotspots we performed two types of analyses on the normalized density maps: the first, using an aspatial threshold approach and the second, applying a spatial statistical method (Getis-Ord $G_i$*) using three neighbourhood definitions. Using one of the common aspatial approaches in biological conservation (e.g., Parviainen et al., 2009 and Tolimieri et al., 2015), we identified the top 5% of normalized species density values. A threshold set at the 95$^{th}$ percentile value defined hotspots as the highest 5% of

the data. A second spatially explicit approach to hotspot detection was applied next. Methods drawn from spatial statistics have additional advantages to the commonly applied top 5% approach. Specifically, the use of statistical thresholds, incorporating spatially local autocorrelation, and the use of a test hypothesis where the null assumes patterns are generated from random process (Getis, 2010). $G_i$*detects spatial clustering of either high or low species density values, where clusters are greater than expected from spatial patterns generated from chance processes. $G_i$* follows the basic form:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_{j=1}^{n} x_j} \quad (2)$$

where *i* is the pivot location, x is the attribute value of *i* – in this case, density – and $W_{ij}$ is a spatial weights matrix created using a distance threshold (d) or the spatial configuration of adjacent cells to define neighbours of the *i*th observation (Getis and Ord, 1992). Hotspots are identified when a pivot location and its surrounding neighbourhood, defined by $W_{ij}$, include values of high normalized species density, relative to all normalized density values within the study area. Permutation testing can be used to determine if the pattern of clustering is more or less than expected when compared to patterns generated from random process. The $G_i$* statistic was performed using GeoDa software (v.1.6.6 October 2014; Anselin et al., 2010) with 999 permutations to determine significance at the 0.05 level.

There are multiple ways to define a spatial neighbourhood ($W_{ij}$) and the selection of neighbourhood type will influence which locations are included in the hotspots. We

employed contiguity and distance neighbourhood definitions, which are commonly used with areal datasets (Dubin, 2009). Contiguity matrices are typically employed when adjacency relationships between areal units are of interest. In ecological studies, equal area units, such as hexagons or grids, are generally used to represent continuous phenomena (Birch et al., 2007), providing natural definitions for contiguity. We implemented first and second order contiguity, meaning neighbourhoods are defined by the shared boundaries of directly adjacent cells from pivot *i* for first order (lag 1) and also those directly adjacent to the first order (second order; lag 2) (Nelson and Robertson, 2012; Figure 2.2). Adjacency is defined using terminology formulated around movements of chess pieces: rook, bishop, and queen (Dubin, 2009). Rook contiguity considers neighbours to be cells adjacent to the immediate top, bottom, left and right of pivot *i*, while the diagonal corners are not considered. Bishop contiguity is the opposite of rook, where only diagonal corners are included. In this case we have chosen queen contiguity; it considers any neighbour that directly touches the cell border of *i*, regardless of direction (Figure 2.2).

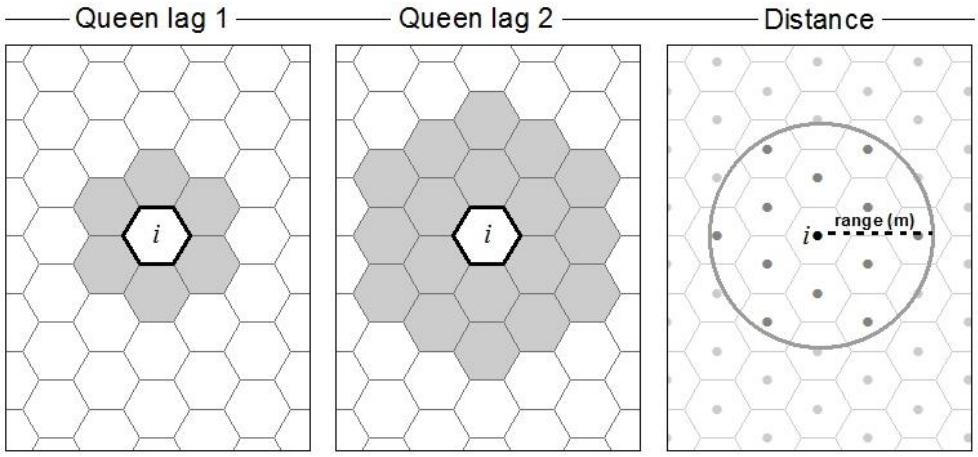


Figure 2.2 Illustration of two different ways to define a spatial neighbourhood: 1) queen contiguity defined as first order (lag 1) and second order (lag 2) and 2) distance-based radius (range value from semivariogram).

We also used a distance definition to demonstrate the sensitivity of $G_i^*$ to various definitions of spatial neighbourhoods. Distance-based definitions employ the use of a fixed distance threshold (radius), whereby all polygon centroids that fall within the defined distance are considered to be within the same spatial neighbourhood (O'Sullivan and Unwin, 2010). However, choosing the appropriate distance radius (i.e., threshold value) can be determined multiple ways. Here, the radius was determined by selecting the range value from an experimental semivariogram plot that was fit using an ordinary least squares model (Cressie, 1993, p. 94) (Figure 2.2 and 2 3). Semivariograms are often applied in geostatistics (typically geology or earth sciences) to quantify spatial autocorrelation – or the strength of association – between observations as the distance between pairs of observations increases (Atkinson and Lloyd, 2009). Semivariograms graph the semivariance of pairs of observations on the y-axis and the lag distance, which separates these observations on the x-axis. An empirical model is then used to fit a line to the plotted points from which certain numerical characteristics can be extracted. The *range* is a semivariogram characteristic that identified the distance at which spatial autocorrelation diminishes (scale of spatial variation) and provides an indication of when observations are no longer spatially related (O'Sullivan and Unwin, 2010). It is logical to apply the range value as the distance threshold value as, by definition, hotspots are regions where greater than expected aggregations of highly similar values occur. When the semivariogram was run for the normalized density maps the range value was similar between semivariograms, with a 28.1 km range value for cetaceans, 25.9 km for pinnipeds and 23.8 km when all species were combined.

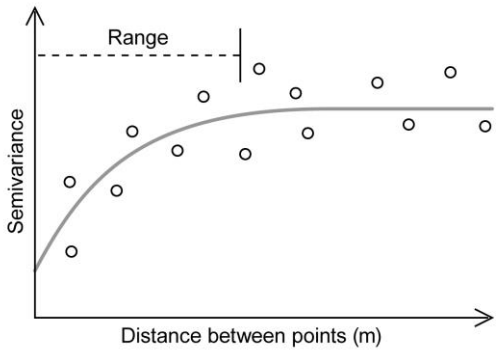


Figure 2.3 Semivariogram from which the radius of a distance band based spatial neighbourhood can be determined by using the range value.

## 2.4 Results

Species-specific density surfaces generated from GAM models show that species density is heterogeneously distributed across the study region (Figure 2.4). The predictive performance of models, shown here using explained deviance and adjusted $R^2$ values, ranged between 25.5% – 9.44% and 0.238 – 0.0838 respectively (Table 2.3). Visualization of each species map highlights clear regions where predicted species density is highest (Figure 2.4). For example, Dall's porpoise shows high density values in the most northern sections of the study region surrounding Dixon Entrance, while killer whales show two regions of high values in sections of Chatham Sound and an area of coastal Queen Charlotte Sound between Calvert Island and Aristazabal Island.

Table 2.3 Model performance summary statistics (N = 5,679).

| | **DP** | **FW** | **HP** | **HS** | **HW** | **KW** | **MW** | **PW** | **SSL** |
|---|---|---|---|---|---|---|---|---|---|
| # of non-zero observations | 137 | 67 | 50 | 108 | 240 | 18 | 27 | 113 | 28 |
| % deviance explained | 24.50 | 13.1 | 25.50 | 23.30 | 14.50 | 9.44 | 11 | 9.94 | 13.30 |
| Adjusted $R^2$ | 0.23 | 0.12 | 0.24 | 0.22 | 0.13 | 0.08 | 0.10 | 0.09 | 0.12 |
| RMSE (predicted vs. actual values) | 0.93 | 0.26 | 0.93 | 1.17 | 0.28 | 0.84 | 0.14 | 11.85 | 6.49 |

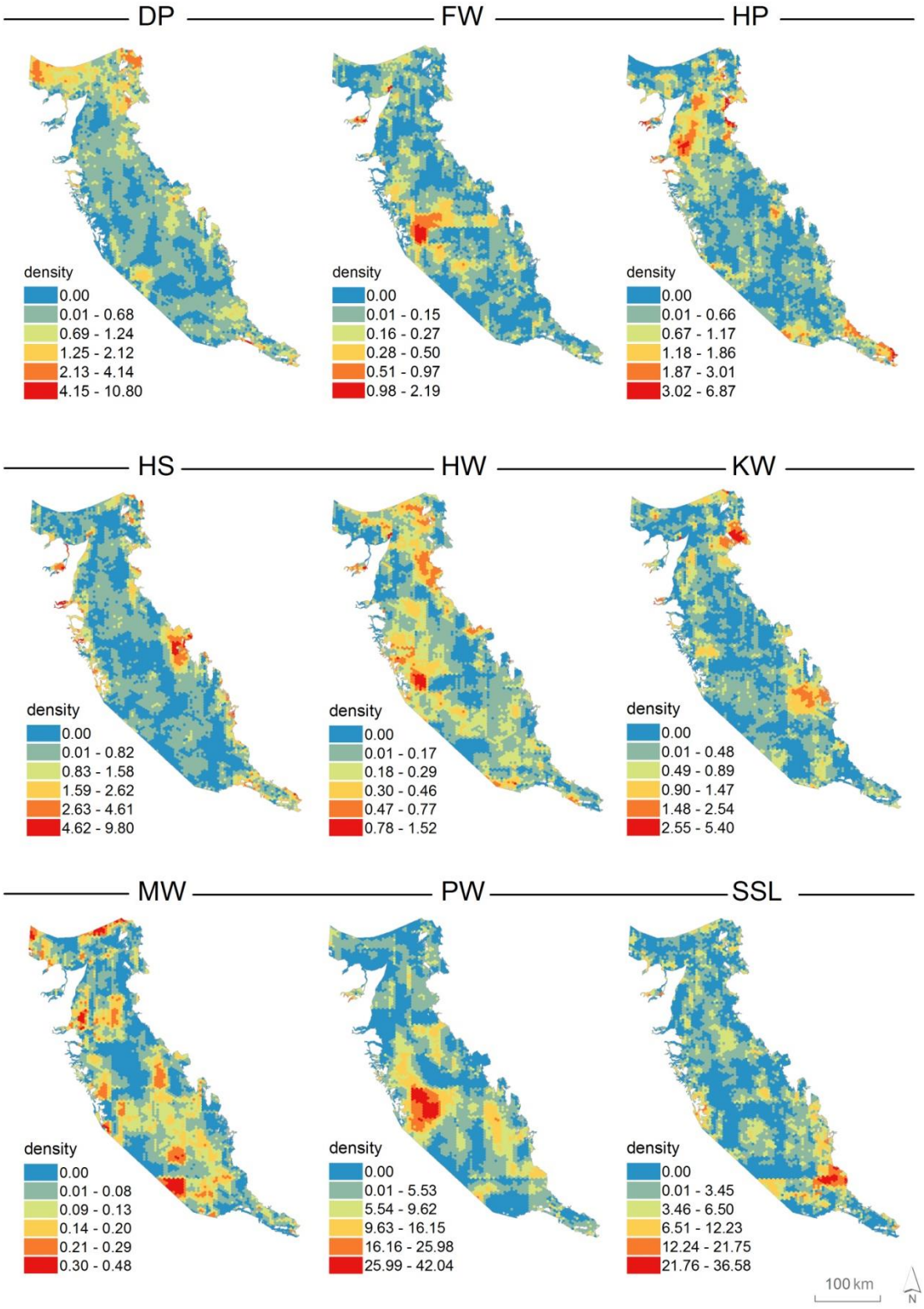


Figure 2.4 Continuous density surfaces generated from species-specific GAM models. Density is defined as the number of species per nautical mile and displayed on a hexagon grid (each hexagon is 13.86 km$^2$). Abbreviations include: Dall's porpoise (DP), fin whale (FW), harbour porpoise (HP), harbour seal (HS), humpback whale (HW), killer whale (KW), common minke whale (MW), Pacific white-sided dolphin (PW), and Steller sea lion (SSL).

Interestingly, the density maps for fin whales, humpback whales, and Pacific white-sided dolphins all possess high values southeast of Haida Gwaii. Harbour porpoise and minke whale show spatially variable regions of high density distributed throughout the study region. Areas of high density for pinnipeds are situated in coastal areas; harbour seal displays highest values in Caamano Sound, while Steller sea lions have their highest densities in a more southern coastal region located off Cape Calvert.

The normalized and summed species maps characterize the collective distributional patterns of cetaceans, pinnipeds, and all species combined (Figure 2.5). Prominent regions of high normalized density for cetaceans are identified southeast of Haida Gwaii near Cape St. James, a small area in outer Queen Charlotte Sound near the Scott Islands, and scattered areas in Chatham Sound and Dixon Entrance, which lie in the northeast section of the study area and are adjacent to the city of Prince Rupert. Highest predicted densities of pinnipeds are generally situated along coastal areas in the southern sections of the study area in regions featuring shallow banks and minimal ocean depths (Thomson, 1981). Notable regions of high normalized density are identified off Calvert Island (Cape Calvert) and another in Caamano Sound and adjacent to Aristazabal Island. When combined, regions that are shown to support elevated levels of species aggregations for both cetaceans and pinnipeds are clearly distinguished.

Hotspot analysis produced multiple spatial representations of potential candidate areas for conservation. The density threshold value calculated for the top 5% hotspots was similar between cetaceans (≥1.16) and all species (≥1.32); however, pinnipeds were different producing a threshold value of ≥0.41 (Figure 2.5). The hotspot analysis showed

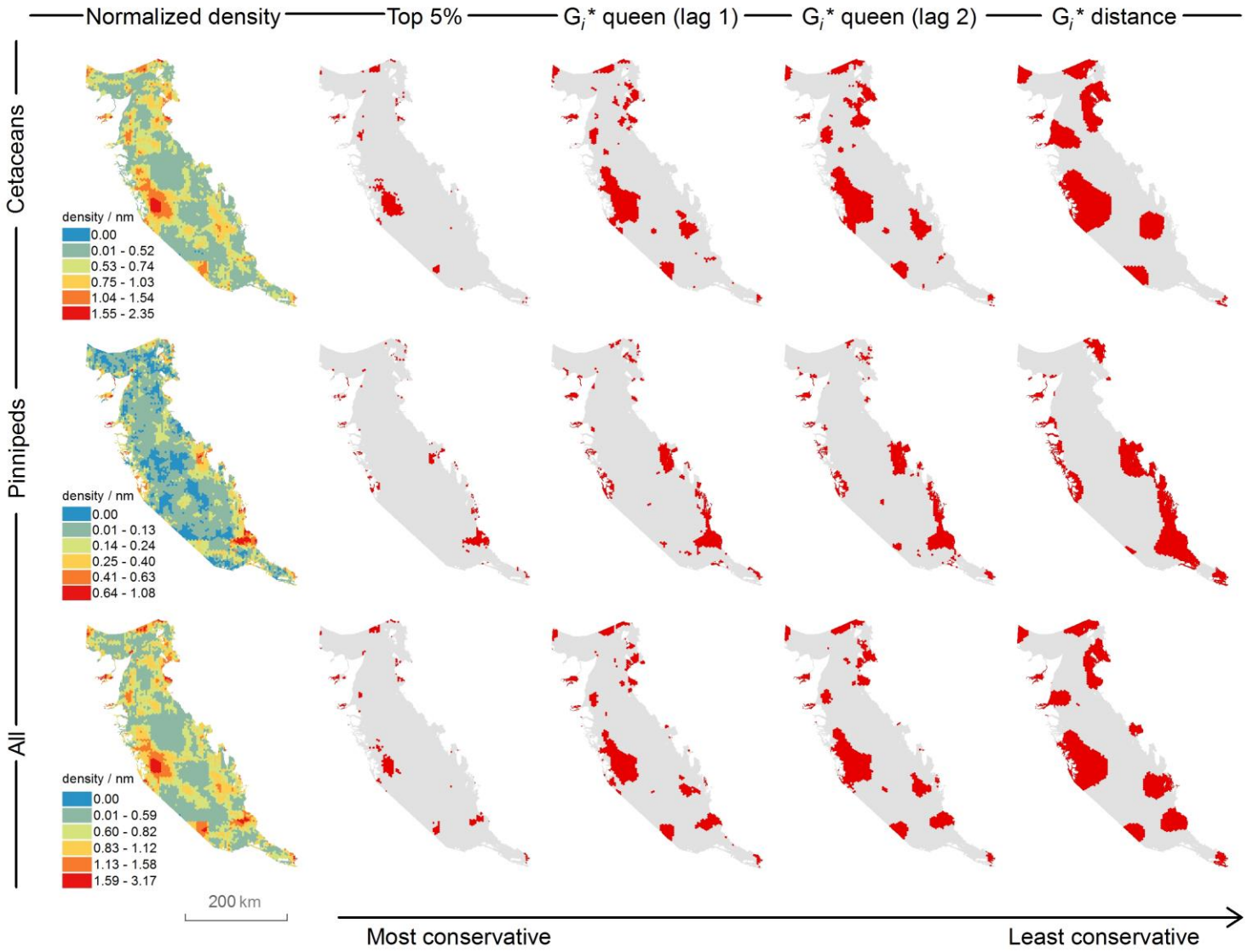


Figure 2.5 Four hotspot outputs (top 5%, $G_i$*queen [lag 1], $G_i$* queen [lag 2], and $G_i$*distance) generated from normalized and summed density maps (first column) for cetaceans, pinnipeds, and all species combined.

that $G_i$* hotspots coincide with areas identified from the top 5% approach, however covered a greater spatial extent, with smoother and more spatially complete borders. The $G_i$* outputs showed fewer pockets of high-density regions but were larger in overall size compared to the top 5% method. These observations are apparent when the average size and number of hotspots for each method are compared (Table 2.4).

Table 2.4 Hotspot summary table

| | | **Top 5%** | **$G_i$* queen (lag 1)** | **$G_i$* queen (lag 2)** | **$G_i$* distance** |
|---|---|---|---|---|---|
| Cetaceans | # of hotspots (n) | 44 | 42 | 32 | 23 |
| | Average size ($km^2$) | 65.26 | 230.89 | 378.50 | 779.52 |
| Pinnipeds | # of hotspots (n) | 81 | 70 | 67 | 19 |
| | Average size ($km^2$) | 27.30 | 87.57 | 111.82 | 718.38 |
| All | # of hotspots (n) | 62 | 60 | 44 | 22 |
| | Average size ($km^2$) | 40.60 | 149.08 | 259.13 | 823.32 |

A clear trend is evident showing a decrease in the number of individual hotspots and an increase in average hotspot size for $G_i$* methods over the aspatial top 5% technique. Furthermore, this observation is also exhibited when the spatial neighbourhood definition changes. Configuration differences between outputs illustrate that the top 5% results are highly patchy, smaller, and display greater spatial heterogeneity than the $G_i$* statistics, suggesting that aspatial approaches produce the most conservative hotspot estimates compared to spatial methodology. Hotspots are generally situated in coastal and nearshore regions and are consistently absent in central locations of the study area within Hecate Strait and Queen Charlotte Sound (Figure 2.5). When cetacean and pinniped hotspots are compared, it appears that Caamano Sound and Cape

Calvert hotspots are primarily driven by pinniped density while cetacean density is responsible remaining hotspot regions.

## 2.5 Discussion

Marine species distribution and density information is typically collected in the form of disjoint samples of sighting occurrences. For planning, mapping, and conservation directives it is beneficial to convert samples of species occurrence to continuous surfaces that represent the possible distribution, abundance or density of a population (Becker et al., 2014; Franklin, 2010; Menza et al., 2016). Predictive species-habitat modelling techniques offer methods to extract meaningful spatial information from limited species observations by relating environmental conditions to abundance or density estimates. The resulting continuous surfaces fill gaps in survey data and provide baseline information valuable for marine spatial planning and conservation ventures. One example of a recent initiative is the Marine Plan Partnership for the North Pacific Coast (MaPP) located within the Pacific North Coast Integrated Management Area (PNCIMA). Marine spatial planning projects, such as MaPP, typically aim to generate strategic directives for regional zones that balance the needs of multiple stakeholders (including biological communities) considering both present and future conditions and needs. Therefore, species distribution and density maps, like those generated in this study, enhance marine plans by providing a continuous surface that indicates current species patterns, identifies candidate areas for conservation prioritization, and supplies a baseline from which future change can be detected. Where available, capacity for informed decision-making may be further improved by combining numerous prediction maps from multiple regions along the northeast Pacific coast (e.g., Menza et al., 2016 [Washington

coast] and Becker et al., 2014 [California coast]) to increase understanding of basin-wide distributional patterns and to form a holistic picture of species patterns.

Despite the value of distribution and density maps for planning and conservation, spatial data for marine species are often incomplete. The challenges of distribution and density mapping for marine species revolve around the complexities of conducting analysis using sparse survey sightings for ocean regions that are difficult to sample consistently and completely. Furthermore, the prohibitive cost of at-sea research and the extreme weather conditions common for northern coastal latitudes hinder the ability to collect species information. Consequently, studies of marine species often rely on zero-inflated sightings data (e.g., Oppel et al., 2012 and Menza et al., 2016), include coarse resolution environmental predictors (Redfern et al., 2006), and produce solely spatial predictions rather than seasonal estimates (e.g., Winiarski et al., 2014). These factors likely contribute to lower model performance, yet often represent the best available information for many marine species. In this study, deviance explained values from our GAM models are consistent with other prediction surfaces for marine mammals. For example, Mannocci et al., (2014) produced explained deviances for cetaceans in the South Pacific gyre between 5% and 30% and Best et al., (2015) generated values between 11% and 51% for marine mammals in coastal British Columbia. The limited number of non-zero observations in this study required the aggregation of sighting observations across seasons and years but nevertheless, represent the most up-to-date distributions for many marine mammals in British Columbia.

### 2.5.1 Placing hotspots in context

Prediction surfaces as standalone map products supply a wealth of valuable information, particularly for combating the shifting baseline syndrome (Pauly, 1995) and acting to fill knowledge gaps in species distributions and densities. However, these maps can also be used to identify spatially explicit hotspots outlining potential candidate areas for future protection. Geographic areas showing consistent hotspots of intense use by marine species may indicate areas of persistent concentrations of productivity that attract elevated quantities of marine mammals and result in spatial aggregations of one, or more, species (Bouchet et al., 2015). Predictable hotspots may be driven by oceanographic conditions, which promote ocean mixing and elevated primary productivity, such as eddies, high current regions, or upwelling zones (Bakun, 2006). In the Northeast Pacific Ocean, trophic interactions are influenced from bottom-up processes, where concentrations of chlorophyll-a are highly correlated with elevated numbers of higher trophic level organisms, including zooplankton and fish (Ware and Thomson, 2005) and, in turn, attract mobile, high trophic level predators. Regions supporting dense aggregations of species (i.e., hotspots) often represent favourable habitat and, given suitable protections (e.g. MPA status), provide an effective strategy to protect biodiversity (Hyrenbach et al., 2000).

Yet, highly utilized habitat could also represent an area of limited ecological benefit and may, in fact, represent a region of considerable risk. Areas of intense use by marine mammals, identified by hotspots, could indicate "ecological traps" – a relatively new ecological theory discussed primarily in terrestrial contexts – where the habitat is of inferior quality and is not suitable for maintaining populations but, nevertheless, is

favoured over other, more suitable, regions (Dwernychuk and Boag, 1972). In these areas, attractive locations may also be risky and may place species at increased exposure to danger or threats (Battin, 2004). For example, marinas or shipping routes may put marine mammals in contact with human populations and increase the chance of negative encounters or stressors (such as increased noise, oil slicks, or ship strikes) that may inhibit reproduction, foraging, or feeding. Consequently, our hotspots should be evaluated in combination with site-specific visits that assess habitat quality and potential human risk, to avoid allocating protected areas to regions which offer little to no ecological benefit.

Hotspots have long been used to guide spatial planning and conservation efforts to areas that will provide the greatest benefit given limited conservation resources, despite the risk of identifying an ecological trap (Myers, 1988, 1990, 2003; Myers et al., 2000; Selig et al., 2014; Worm et al., 2003). Generally, hotspots for conservation purposes are identified with threshold cut-offs chosen in varying ways to separate extreme values that are considered “hot” from all other non-hot regions. Typically, thresholds are arbitrarily selected based on study context and data type (Cañadas et al., 2014). The threshold values vary among studies, often ranging somewhere between the top 2.5% and 10% of data values. For example, Ceballos and Ehrlich (2006) used distributions of 4,818 land mammals to identify hotspots that represented the top 2.5% of species richness, endemism, and threatened species indices. Tolimieri et al. (2015), on the other hand, incorporated both 5% and 10% thresholds when extracting hotspots of demersal fish biodiversity. While the top percentage approach provides data required to develop conservation plans, it does not account for potential bias that likely occurs when

determining threshold values based on user opinion. Variability in threshold choices indicates a lack of consensus on the most appropriate strategy. Recognizing the limitations of subjective thresholds, we have applied methods from spatial statistics to quantify hotspots. The main differences between top percentage thresholds and spatial techniques are the ability to account for spatial patterns in the data and objectively identify hotspots.

### 2.5.2 Identifying hotspots

The results of this study illustrate multiple approaches for detecting spatially explicit hotspots in marine mammal data. Our analysis demonstrated that the aspatial top 5% threshold produced the most conservative (smallest) hotspot outputs. In a changing climate with increasing anthropogenic pressures, conservative methods may not allow for sufficient uncertainty given that the spatial patterns of species distributions will likely change through time. Specifically, Hazen et al. (2013) suggest that by 2100 some marine predators in the northeast Pacific Ocean will exhibit up to a 35% change in highly used habitat, due to shifts in environmental conditions. Hotspots with highly patchy, complex, and convoluted geometry present highly precise hotspot borders. If MPAs are created based on these precise regions, they may not adequately protect core habitat regions should species distributions shift over time. Furthermore, survey areas do not include the entirety of the study species migratory range and will increase uncertainty in population trends because only a sample of the total population is recorded (Forney, 2000). As a result, overly conservative estimates depicting very small hotspots may fail to identify regions large enough – typically larger than similar land-based reserves – to account for

extensive migratory ranges and the dynamic processes that influence marine species and systems (Gerber et al., 2003; Hooker and Gerber, 2004).

Our findings suggest that spatial techniques for identifying hotspots (based on detecting spatial patterns) may avoid the stated limitations of aspatial approaches (e.g., the top 5% threshold). Local measures of spatial association, such as $G_i$*, provide robust and objective definitions of hotspots, when compared to aspatial methodology. Aspatial methods produce spatially patchy hotspots with more precise borders than may be warranted by the quality of the input data. In assessing methodological approaches, the $G_i$* statistic has many advantages over the top percentage threshold. First, the $G_i$* method is based on an arithmetic framework where thresholds to delineate a hotspot are derived from statistical values calculated by incorporating spatial patterns in the data. Second, spatial statistics account for spatial autocorrelation and form results that quantify the level of similarity between values as distance between them increases (Fortin and Dale, 2009). Lastly, unlike aspatial approaches, $G_i$* method employs a test hypothesis. Hypothesis testing differentiates spatial patterns generated from random process from those processes that generate clusters of species. The metric allows the inclusion of statistical significance measures to convey levels of uncertainty, which is important for making informed policy decisions. Using a statistical test to determine where realized processes are different than expected, based on null hypotheses of randomness, provides a mechanism to monitor a seascape for unexpected change in a variety of phenomena.

As our results indicate, different neighbourhood definitions for $G_i$* affect the number and size characteristics of defined hotspots. Queen contiguity definitions (lag 1 and 2) produce hotspots with smaller areas than those generated by the distance

definition; while conversely, contiguity neighbourhoods generate greater overall numbers of hotspots compared to the distance metric. These differences are to be expected, as the distance neighbourhood is not confined to the boundaries of the hexagon grids. In this case, the spatial weights matrix can include observations further from the pivot point than lag 1 and 2 contiguity based metrics. As a result, the distance based parameters were far greater (up to 25.9 km) than those of contiguity, which extended only as far as two hexagon cells (~ 10 km) producing, in turn, larger hotspots. It is also pertinent to note that the areal unit (grid size) of the species distribution maps will influence hotspots generated by neighbourhood metrics, as increasing grid sizes may generate larger and less numerous hotspots.

### 2.5.3 Hotspot comparison

We compared hotspots identified in this study with previous research conducted in the Queen Charlotte Basin to identify areas of consensus and divergence. Previous marine mammal research that relied on the same or partial data used in this study (Best et al., 2015; Williams et al., 2011) display high-density regions of marine mammals in areas we identified as hotspots. In addition, most, but not all, hotspots identified by all four methods in this study coincide with Ecologically and Biologically Significant Marine Areas (EBSAs) that were classified based on expert opinion (Clarke and Jamieson, 2006; Jamieson and Levesque, 2014). EBSA zones identified as hotspots include Learmouth Bank, Chatham Sound, Caamano Sound, Dogfish Banks, Cape St. James, Scott Islands, and North Island Straits. Many EBSAs identified by Clarke and Jamieson (2006) are incorporated into three additional important nearshore regions listed by Jamieson and Levesque (2014). These include the nearshore region along the lower southeast tip of

Haida Gwaii (Haida Gwaii/Queen Charlotte Nearshore), the island regions surrounding the community of Bella Bella (Bella Bella Nearshore), and from the north tip of Banks Island to the southern tip of Aristazabal Island (Central Mainland Nearshore). Congruence between hotspot maps, both within our hotspot methodological approaches and within previous research, emphasizes the validity of the results produced in this study.

The majority of hotspots show agreement with EBSAs however, our study identifies additional important areas: areas adjacent to Cape Calvert in the Central Coast and the most southern area of Queen Charlotte Strait. The Cape Calvert hotspots are likely attributed to the high density of Steller sea lions and Dall's porpoises in this region, in addition to moderate densities of humpback whales, killer whales, and Pacific white-sided dolphins. Our results indicate that all marine mammals are present in the southern Queen Charlotte Strait, however, we note that only the most southern region was identified as a hotspot in our study, while the North Islands Straits EBSA include the entire strait region. In addition, our study identifies prominent hotspots located in the northern most section of eastern and western Dixon Entrance, which appears to be driven by the occurrence of Dall's porpoise.

### 2.5.4 Management and conservation implications

Site-based conservation, which is often situated around extremely important areas for individual species, has been suggested as an effective measure for identifying regions of conservation prioritization for wildlife (Skov et al., 2007; Hinch and De Santo, 2011). However, conservation efforts are limited without detailed information on species distributions and densities. Spatial locations where multiple species concentrate in

predictable habitats can reveal important biological hotspots vital to further our understanding of distributional patterns and movements (Block et al., 2011) and to facilitate recovery strategies. In addition, conservation tools, such as MPAs, are more effective when placed in regions where species congregate in predictable habitats (Hyrenbach et al., 2000) and if the presence of an ecological trap is also investigated prior to reserve implementation (Battin, 2004).

Predictions of marine mammal distributions and densities, as well as the quantification of hotspots, have immense potential for contributing to marine mammal conservation goals and objectives. In Canada, a number of marine mammal species are listed under Canada's Species at Risk Act (SARA) and Critical Habitat designations have yet to be identified for the majority of listed species. Given that Canada's Pacific coastal ecosystems are subject to often intense anthropogenic pressures (e.g., Ban and Alder, 2008), with a number of significant industrial projects being proposed (i.e., Pacific NorthWest LNG) in addition to experiencing the consequences of climate change, the identification of areas important to marine mammals is an important contribution to conservation and management efforts.

We found that the identification of geographic hotspots using spatial statistical methodology offers robust techniques for quantifying important habitat and locating species aggregations crucial to future planning efforts. Importantly, the delineation of regions of intense use is highly influenced by the method used to extract hotspots. Our findings suggest that hotspots identified based on spatial neighbourhood characteristics, rather than arbitrarily determined threshold criteria, produce more objective and quantitatively defensible outputs. Hotspots defined based on spatial statistical methods

are also larger than the 5% threshold, which allows for habitat protection that could afford protection into the future should conditions change or species hotspots shift. Therefore, the configuration of spatially defined hotspots offers conservation suggestions that are precautionary and future-proofed. Further, hotspot regions we have identified coincide with previously identified important habitats for marine mammals. Agreement between studies suggests increased confidence in the effective application of spatial statistical methodology for hotspot identification and conservation prioritization.

Considering that the distribution and density patterns of marine mammals in north coastal regions of British Columbia are poorly understood, our study provides updated information on distributions, densities, and assessments of various techniques for hotspot delineation. The results of this study can be applied within the broader framework of conservation planning (Margules and Pressey, 2000; Pressey and Bottrill, 2009) and incorporated into spatial planning software such as MARXAN and C-Plan. As the field of conservation planning continues to evolve, the methodology used to derive spatial inputs for future planning initiatives should also be further refined. This research provides alternative approaches for quantifying hotspots of species distributions in the marine environment and provides further development in the application of spatial pattern-based methods supporting future marine conservation.

**Acknowledgements**

Numerous individuals and organizations are gratefully acknowledged for their contributions to the marine mammal surveys completed by Raincoast Conservation Foundation. We also thank B.D. Best, R. Williams and P.N. Halpin for their efforts to analyze and report their findings regarding the marine mammal survey data. This

research could not have been completed without support provided by Natural Sciences and Engineering Research Council of Canada (NSERC), Raincoast Conservation Foundation, and the Spatial Pattern Analysis and Research Lab (SPAR).

**References**

Anselin, L. 1995. Local indicators of spatial association—LISA. Geographical Analysis 27:93–115.

Anselin, L., I. Syabri, and Y. Kho. 2010. A.4 GeoDa: An Introduction to Spatial Data Analysis. Pages 73–89 *in* M. M. Fischer and A. Getis, editors. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Heilelberg, Berlin.

Atkinson, P. M., and C. D. Lloyd. 2009. Geostatistics and Spatial Interpolation. Chap. 9. Pages 159–181 *in* S. A. Fotheringham and P. A. Rogerson, editors. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Bakun, A. 2006. Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. Scientia Marina 70:105–122.

Ban, N., and J. Alder. 2008. How wild is the ocean? Assessing the intensity of anthropogenic marine activities in British Columbia, Canada. Aquatic Conservation: marine and freshwater ecosystems 18:55–85.

Battin, J. 2004. When good animals love bad habitats: Ecological traps and the conservation of animal populations. Conservation Biology 18:1482–1491 .

Becker, E. A., K. A. Forney, D. G. Foley, R. C. Smith, T. J. Moore, and J. Barlow. 2014. Predicting seasonal density patterns of California cetaceans based on habitat models. Endangered Species Research 23:1–22.

Best, B. D., C. H. Fox, R. Williams, P. N. Halpin, and P. C. Paquet. 2015. Updated marine mammal distribution and abundance estimates in British Columbia. Journal of Cetacean Resource Management 15:9–26.

Birch, C. P., S. P. Oom, and J. A. Beecham. 2007. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. Ecological Modelling 206:347–359.

Block, B. A., I. D. Jonsen, S. J. Jorgensen, A. J. Winship, S. A. Shaffer, S. J. Bograd, L. Hazen, D. G. Foley, G. A. Breed, A. L. Harrison, J. E. Ganong, A. Swithenbank, M. Castleton, H. Dewar, B. R. Mate, G. L. Shillinger, K. M. Schaefer, S. R. Benson, M. J. Weise, R. W. Henry, and D. P. Costa. 2011. Tracking apex marine predator movements in a dynamic ocean. Nature *475*:86–90.

Boots, B. 2002. Local measures of spatial association. Ecoscience 9:168–176.

Bouchet, P. J., J. J. Meeuwig, S. Kent, P. Chandra, T. B. Letessier, and C. K. Jenner. 2015. Topographic determinants of mobile vertebrate predator hotspots: current knowledge and future directions. Biological Reviews *90*:699–728.

Brodeur, R. D., and D. M. Ware. 1992. Long-term variability in zooplankton biomass in the subarctic Pacific Ocean. Fisheries Oceanography 1:32–38.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, UK.

Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2004. Advanced Distance Sampling: Estimating Abundance of Biological Populations. Oxford University Press, Oxford, UK.

Butchart, S. H., M. Walpole, B. Collen, A. Van Strien, J. P. Scharlemann, R. E. Almond, ... and R. Watson. 2010. Global biodiversity: indicators of recent declines. Science 328:1164–1168.

Campagna, C., E. W. Sanderson, P. B. Coppolillo, V. Falabella, A. R. Piola, S. Strindberg, and J. P. Croxall. 2008. A species approach to marine ecosystem conservation. Aquatic Conservation: Marine and Freshwater Ecosystems 17:S122–S147.

Cañadas, E.M., G. Fenu, J. Peñas, J. Lorite, E. Mattana, and G. Bacchetta. 2014. Hotspots within hotspots: Endemic plant richness, environmental drivers, and implications for conservation. Biological Conservation 170:282–291.

Ceballos, G., and P. R. Ehrlich. 2006. Global mammal distributions, biodiversity hotspots, and conservation. Proceedings of the National Academy of Sciences 103:19374–19379.

Clarke, C. L., and G. S. Jamieson. 2006. Identification of Ecologically and Biologically Significant Areas in the Pacific North Coast Integrated Management Area: Phase II – Final Report. Can. Tech. Rep. Fish. Aquat. Sci. 2686: v + 25 pp.

Clarke Murray, C., S. Agbayani, H. M. Alidina, and N. C. Ban. 2015. Advancing marine cumulative effects mapping: An update in Canada's Pacific waters. Marine Policy 58:71–77.

Convention on Biological Diversity. 2013. Quick guides to the Aichi biodiversity targets, version 2. United Nations. www.cbd.int (accessed 15 Feb 2016).

Crawford, W. R., and R. E. Thomson. 1991. Physical oceanography of the western Canadian continental shelf. Continental Shelf Research 11:669–683.

Crawford, W. R., P. J. Brickley, and A. C. Thomas. 2007. Mesoscale eddies dominate surface phytoplankton in northern Gulf of Alaska. Progress in Oceanography 75:287–303.

Cressie, N. A. C. 1993. Statistics for spatial data: Wiley series in probability and statistics. Rev. ed. Wiley-Interscience, New York.

Croll, D.A., B. R. Tershy, R. P. Hewitt, D. A. Demer, P. C. Fiedler, S. E. Smith, W. Armstrong, J. M. Popp, T. Kiekhefer, V. R. Lopez, J. Urban, and D. Gendron. 1998. An integrated approach to the foraging ecology of marine birds and mammals. Deep-Sea Research II: Topical Studies in Oceanography 45:1353–1371.

Dalla Rosa, L., J. K. Ford, and A. W. Trites. 2012. Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. Continental Shelf Research 36:89–104.

Dirzo, R., H. S. Young, M. Galetti, G. Ceballos, N. J. Isaac, and B. Collen. 2014. Defaunation in the Anthropocene. Science 345:401–406.

Dormann, C.F., J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. García Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão, T. Münkemüller, C. McClean, P. E. Osborne, B. Reineking, B. Schröder, A. K. Skidmore, D. Zurell, and S. Lautenbach. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46.

Dubin, R. 2009. Spatial Weights. Chap. 8. Pages 125–157 *in* S. A. Fotheringham and P. A. Rogerson, editors. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Dwernychuck, L.W., and D. A. Boag. 1972. Ducks nesting in association with gulls: an ecological trap? Canadian Journal of Zoology 50:559–563.

Favaro, B., D. Claar, C. Fox, C. Freshwater, J. Holden, and A. Roberts. 2014. Trends in Extinction Risk for Imperiled Species in Canada. PLoS ONE 9:e113118.

Foreman, M. G. G., W. R. Crawford, J. Y. Cherniawsky, R. F. Henry, and M. R. Tarbotton. 2000. A high-resolution assimilating tidal model for the northeast Pacific Ocean. Journal of Geophysical Research: Oceans 105:28629–28651.

Forney, K. A. 2000. Environmental models of cetacean abundance: reducing uncertainty in population trends. Conservation Biology 14:1271–1286.

Fortin, M., and M. R. T. Dale. 2009. Spatial Autocorrelation. Chap. 6. Pages 89–103 *in* S. A. Fotheringham and P. A. Rogerson, editors. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Fox, C. H., Huettmann, F., Harvey, G. K., Morgan, K., Robinson, J., and Paquet, P.C. (In Review). Predictions from Machine Learning ensembles: marine bird distribution and density on Canada's Pacific coast**.** Marine Ecology Progress Series.

Franklin, J. 2010. Mapping species distributions: spatial inference and prediction. Cambridge University Press, Cambridge, UK.

Gerber, L. R., L. W. Botsford, A. Hastings, H. P. Possingham, S. D. Gaines, S. R. Palumbi, and S. Andelman. 2003. Population models for marine reserve design: a retrospective and prospective synthesis. Ecological Applications 13:47–64.

Getis, A. 2010. Spatial Autocorrelation. Chap. B.3. *in* M. M. Fischer and A. Getis, editors. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Heidelberg, Berlin.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24:189–206.

Guisan, A., T. C. Edwards, and T. Hastie. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological Modelling 157:89–100.

Halpern, B. S., S. Walbridge, K. A. Selkoe, C. V. Kappel, F. Micheli, C. D'Agrosa, J. F. Bruno, K. S. Casey, C. Ebert, H. E. Fox, R. Fujita, D. Heinemann, H. S. Lenihan, E. M. P. Madin, M. T. Perry, E. R. Selig, M. Spalding, R. Steneck, and R. Watson. 2008. A global map of human impact on marine ecosystems. Science 319:948–952.

Harrell Jr., F. E., with contributions from C. Dupont and many others. 2016. Hmisc: Harrell Miscellaneous. R package version 3.17–4.

Hastie, T. J., and R. J. Tibshirani. 1990. Generalized Additive Models, volume 43 of Monographs on Statistics and Applied Probability. Chapman & Hall, London.

Hazen, E. L., S. Jorgensen, R. R. Rykaczewski, S. J. Bograd, D. G. Foley, I. D. Johnsen, S. A. Shaffer, J. P. Dunne, D. P. Costa, L. B. Crowder, B. A. Block. 2013. Predicted habitat shifts of Pacific top predators in a changing climate. Nature Climate Change 3:234–238.

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

Hinch, P. R., and E. M. De Santo. 2011. Factors to consider in evaluating the management and conservation effectiveness of a whale sanctuary to protect and conserve the North Atlantic right whale (*Eubalaena glacialis*). Marine Policy 35:163–180.

Hooker, S. K., and L. R. Gerber. 2004. Marine reserves as a tool for ecosystem-based management: the potential importance of megafauna. BioScience 54:27–39.

Hyrenbach, K. D., K. A. Forney, and P. K. Dayton. 2000. Marine protected areas and ocean basin management. Aquatic conservation: marine and freshwater ecosystems 10:437–458.

IUCN. 2016. The IUCN Red List of Threatened Species. Version 2016-2. www.iucnredlist.org

Jamieson, G. S., and C. Levesque. 2014. Identification of Ecologically and Biologically Significant Areas on the West Coast of Vancouver Island and the Strait of Georgia, and in some nearshore areas on the North Coast: Phase II – Designation of EBSAs. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/101. Vii + 36 pp.

Jardine, I. D., K. A. Thomson. M. G. Foreman, and P. H. LeBlond. 1993. Remote sensing of coastal sea-surface features off northern British Columbia. Remote sensing of environment 45:73–84.

Kim, Y. J., and C. Gu. 2004. Smoothing spline Gaussian regression: more scalable computation via efficient approximation. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 66:337–356.

Kovacs, K., A. Aguilar, D. Aurioles, V. Burkanov, C. Campagna, N. Gales, T. Gelatt, S. D. Goldsworthy, S. J. Goodman, G. J. G. Hofmeyr, T. Härkönen, L. Lowry, C. Lydersen, J. Schipper, T. Sipilä, C. Southwell, S. Stuart, D. Thompson, and F. Trillmich. 2012. Global threats to pinnipeds. Marine Mammal Science 28:414–436.

Kraus, S. D., R. D. Kenny, C. A. Mayo, W. A. McLellan, M. J. Moore, and D. P. Nowacek. 2016. Recent Scientific Publications Cast Doubt on North Atlantic Right Whale Future. Frontiers in Marine Science 3:137.

Kuhn, M., and K. Johnson. 2013. Ch. 3 Data Pre-processing. Pages 27–59. Applied Predictive Modeling. Springer, New York.

Kuletz, K. J., M. C. Ferguson, B. Hurley, A. E. Gall, E. A. Labunski, and T. C. Morgan. 2015. Seasonal spatial patterns in seabird and marine mammal distribution in the eastern Chukchi and western Beaufort seas: Identifying biologically important pelagic areas. Progress in Oceanography 136:175–200.

Liebhold, A. M., and J. Gurevitch, J. 2002. Integrating the statistical analysis of spatial data in ecology. Ecography 25:553–557.

Lotze, H. K., H. S. Lenihan, B. J. Bourque, R. H. Bradbury, R. G. Cooke, M. C. Kay, S. M. Kidwell, M. X. Kirby, C. H. Peterson, and J. B. C. Jackson. 2006. Depletion, degradation, and recovery potential of estuaries and coastal seas. Science 312:1806–1809.

Mannocci, L., P. Monestiez, J. Bolaños-Jiménez, G. Dorémus, S. Jeremie, S. Laran, R. Rinaldi, O. Van Canneyt, and V. Ridoux. 2013. Megavertebrate communities from two contrasting ecosystems in the western tropical Atlantic. Journal of Marine Systems 111:208–222.

Mannocci, L., M. Catalogna, G. Dorémus, S. Laran, P. Lehodey, W. Massart, P. Monestiez, O. Van Canneyt, P. Watremez, and V. Ridoux. 2014. Predicting cetacean and seabird habitats across a productivity gradient in the South Pacific gyre. Progress in Oceanography 120:383–398.

Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. Nature 405:243–253.

McCauley, D. J., M. L. Pinsky, S. R. Palumbi, J. A. Estes, F. H. Joyce, and R. R. Warner. 2015. Marine defaunation: Animal loss in the global ocean. Science 347:1255641.

Menza, C., J. Leirness, T. White, A. Winship, B. Kinlan, L. Kracker, J.E Zamon, L. Ballance, E. Becker, K. A. Forney, J. Barlow, J. Adams, D. Pereksta, S. Pearson, J. Pierce, S. Jeffries, J. Calambokidis, A. Douglas, B. Hanson, S. R. Benson, and L. Antrim. 2016. Predictive mapping of seabirds, pinnipeds and cetaceans off the Pacific Coast of Washington. NOAA Technical Memorandum NOS NCCOS 210. Silver Spring, MD. 96 pp.

Myers, N. 1988. Threatened biotas:" hot spots" in tropical forests. Environmentalist 8:187–208.

Myers, N. 1990. The biodiversity challenge: expanded hot-spots analysis. Environmentalist 10:243–256.

Myers, N. 2003. Biodiversity hotspots revisited. BioScience *53*:916–917.

Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. Da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. Nature 403:853–858.

Nelson, T. A., and B. Boots. 2008. Detecting spatial hot spots in landscape ecology. Ecography 31:556-566.

Nelson, T. A., S. N., Gillanders, J. Harper, and M. Morris. 2011. Nearshore aquatic habitat monitoring: a seabed imaging and mapping approach. Journal of Coastal Research 27:348–355.

Nelson, T. A., and C. Robertson. 2012. Refining spatial neighbourhoods to capture terrain effects. Ecological Processes 1:1–11.

Nhancale, B. A., and R. J. Smith. 2011. The influence of planning unit characteristics on the efficiency and spatial pattern of systematic conservation planning assessments. Biodiversity and Conservation 20:1821–1835.

Nur, N., J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley, J. A. Wiens, K. Morgan, L. T. Ballance, and D. Stralberg. 2011. Where the wild things are: predicting hotspots of seabird aggregations in the California Current System. Ecological Applications 21:2241–2257.

O'Sullivan, D., and D. J. Unwin. 2010. Geographic Information Analysis (second ed.). John Wiley & Sons, Hoboken, NJ.

Ord, J. K., and A. Getis. 2001. Testing for local spatial autocorrelation in the presence of global autocorrelation. Journal of Regional Science 41:411–432.

Orme, C. D. L., R. G. Davies, M. Burgess, F. Eigenbrod, N. Pickup, V. A. Olson, A. J. Webster, T. Ding, P. C. Rasmussen, R. S. Ridgely, A. J. Sattersfield, P. M. Bennett, T. M. Blackburn, K. J. Gaston, and I. P. F. Owens. 2005. Global hotspots of species richness are not congruent with endemism or threat. Nature 436:1016–1019.

Oppel, S., A. Meirinho, I. Ramírez, B. Gardner, A. F. O'Connell, P. I. Miller, and M. Louzao. 2012. Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biological Conservation 156:94–104

Parviainen, M., M. Marmion, M. Luoto, W. Thuiller, and R. K. Heikkinen. 2009. Using summed individual species models and state-of-the-art modelling techniques to identify threatened plant species hotspots. Biological Conservation 142:2501–2509.

Pauly, D. 1995. Anecdotes and the shifting baseline syndrome of fisheries. Trends in Ecology & Evolution 10:430.

Preikshot, D., R. J. Beamish, and C. M. Neville. 2013. A dynamic model describing ecosystem-level changes in the Strait of Georgia from 1960 to 2010. Progress in Oceanography 115:28–40.

Pressey, R. L., and M. C. Bottrill. 2009. Approaches to landscape-and seascape-scale conservation planning: convergence, contrasts and challenges. Oryx 43:464–475.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Ausria. http://www.r-project.org/

Rao, K. H., A. Smitha, and M. M. Ali. 2006. A study on cyclone induced productivity in south-western Bay of Bengal during November-December 2000 using MODIS (SST and chlorophyll-a) and altimeter sea surface height observations. Indian Journal of Marine Sciences 35:153–160.

Redfern, J., M. Ferguson, E. Becker, K. Hyrenbach, C. Good, J. Barlow, K. Kaschner, M. F. Baumgartner, K. A. Forney, L. T. Ballance, F. Fauchald, P. Halpin, T. Hamazaki, A. J. Pershing, S. S. Qian, A. Read, S. B. Reilly, L. Torres, and F. Werner. 2006. Techniques for cetacean – habitat modelling. Marine Ecology Progress Series 310:271–295.

Reese, D. C., and R. D. Brodeur. 2006. Identifying and characterizing biological hotspots in the northern California Current. Deep Sea Research Part II: Topical Studies in Oceanography 53:291–314.

Rogachev, K. A., E. C. Carmack, and M. G. G. Foreman. 2008. Bowhead whales feed on plankton concentrated by estuarine and tidal currents in Academy Bay, Sea of Okhotsk. Continental Shelf Research 28:1811–1826.

Schipper, J., J. Chanson, F. Chiozza, N. Cox, M. Hoffmann, V. Katariya, V., and … B. Young. 2008. The status of the world's land and marine mammals: Diversity, threat, and knowledge. Science 322:225–230.

Selig, E. R., W. R. Turner, S. Troëng, B. P. Wallace, B. S. Halpern, K. Kaschner, B. G. Lascelles, K. E. Carpenter, and R. A. Mittermeier. 2014. Global priorities for marine biodiversity conservation. PLoS One 9:e82898.

Skov, H., J. Durinck, M. F. Leopold, and M. L. Tasker. 2007. A quantitative method for evaluating the importance of marine areas for conservation of birds. Biological Conservation 136:362–371.

Smith, S. C., and H. Whitehead. 1993. Variations in the feeding success and behaviour of Galápagos sperm whales (*Physeter macrocephalus*) as they relate to oceanographic conditions. Canadian Journal of Zoology 71:1991–1996.

Springer, A. M., C. P. McRoy, and M. V. Flint. 1996. The Bering Sea Green Belt: shelf-edge processes and ecosystem production. Fisheries Oceanography 5:205–223.

Stammer, D., and C. Wunsch. 1999. Temporal changes in eddy energy of the oceans. Deep Sea Research Part II: Topical Studies in Oceanography 46:77–108.

Tynan, C. T., D. G. Ainley, J. A. Barth, T. J. Cowles, S. D. Pierce, and L. B. Spear. 2005. Cetacean distributions relative to ocean processes in the northern California Current System. Deep-Sea Research II 52:145–167.

Thomas, L., R. Williams, R., and D. Sandilands. 2007. Designing line transect surveys for complex survey regions. Journal of Cetacean Research and Management 9:1–13.

Thomson, R. E. 1981. Oceanography of the British Columbia Coast, volume 56 of Canadian Special Publication of Fisheries and Aquatic Sciences. Canada Department of Fisheries and Oceans, Ottawa, Ontario.

Tolimieri, N., A. O. Shelton, B. E. Feist, and V. Simon. 2015. Can we increase our confidence about the locations of biodiversity ‘hotspots' by using multiple diversity indices? Ecosphere 6:1–13.

Tosh, C. A., P. N. de Bruyn, J. Steyn, H. Bornemann, J. van den Hoff, B. S. Stewart, J. Plötz, and M. N. Bester. 2015. The importance of seasonal sea surface height anomalies for foraging juvenile southern elephant seals. Marine Biology 162:2131–2140.

Wagner, H. H., and M. J. Fortin. 2005. Spatial analysis of landscapes: concepts and statistics. Ecology 86:1975–1987.

Ware, D. M., and R. E. Thomson. 2005. Bottom-up ecosystem trophic dynamics determine fish production in the Northeast Pacific. Science 308:1280–1284.

Whitney, F., K Conway, R. Thomson, V. Barrie, M. Krautter, and G. Mungov. 2005. Oceanographic habitat of sponge reefs on the Western Canadian Continental Shelf. Continental Shelf Research 25:211–226.

Williams, R., and L. Thomas. 2007. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada. Journal of Cetacean Research and Management 9:15–28.

Williams, R., and P. O’Hara. 2010. Modelling ship strike risk to fin, humpback and killer whales in British Columbia, Canada. Journal of Cetacean Research and Management 11:1–8.

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

Winiarski, K. J., M. L. Burt, E. Rexstad, D. L. Miller, C. L. Trocki, P. W. Paton, and S. R. McWilliams. 2014. Integrating aerial and ship surveys of marine birds into a combined density surface model: A case study of wintering Common Loons. The Condor 116:149–161.

Wood, S. N. 2006. Generalized Additive Models: An Introduction with R. Chapman & Hall/CRC, Boca Raton, FL.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparamteric generalized linear models. Journal of the Royal Statistical Society 73:3–36.

Wood, S. N. 2016. Package 'mgcv'. Mixed GAM Computation Vehicle with GCV/AIC/REML Smoothness Estimation. R package version 1.8–12.

Worm, B., H. K. Lotze, and R. A. Myers. 2003. Predator diversity hotspots in the blue ocean. Proceedings of the National Academy of Sciences 100:9884–9888.

Wright, D. J., M. Pendleton, J. Boulware, S. Walbridge, B. Gerlt, D. Eslinger, D. Sampson, and E. Huntley. 2012. ArcGIS Benthic Terrain Modeler (BTM), v. 3.0, Environmental Systems Research Institute, NOAA Coastal Services Center, Massachusetts Office of Coastal Zone Management. Available at: http://esriurl.com/5754.

Yen, P. P. W., W. J. Sydeman, and K. D. Hyrenbach. 2004. Marine bird and cetacean associations with bathymetric habitats and shallow-water topographies: implications for trophic transfer and conservation. Journal of Marine Systems 50:79–99.

Zacharias, M. A., and J. C. Roff. 2001. Use of focal species in marine conservation and management: a review and critique. Aquatic Conservation: Marine and Freshwater Ecosystems 11:59–76.

Zhihai, M., B. Zuckerberg, W. F. Porter, and L. Zhang. 2012. Use of localized descriptive statistics for exploring the spatial pattern changes of bird species richness at multiple scales. Applied Geography 32:185-194.

# 3.0 COMPARING CITIZEN SCIENCE AND SYSTEMATIC SURVEYS OF MARINE MAMMAL DISTRIBUTIONS AND DENSITIES

## 3.1 Abstract

Information collected by citizen scientists represents a significant and growing source of knowledge regarding species and ecosystems. Further, this information also supports the generation of hypotheses that have been difficult to execute with limited scientific data. Citizen science has the capacity to rapidly increase observer numbers and spatiotemporal scope, which leads to more plentiful data and data collection over larger regions and longer periods. In this study, our goal is to demonstrate how mapping cetacean patterns using citizen science and systematic surveys generates consistent and different understanding of cetacean distributions and densities in British Columbia. We used map comparison methods that quantified similarities and differences between geographic datasets to locate where cetacean distributions and densities had spatially unique or spatially analogous representation. We applied image differencing and spatial cluster detection ($G_i^*$) to identify regions where datasets represent species distribution and density similarly and differently. Where spatial clusters in both data sources are congruent, we interpret with a higher level of confidence that species occur, and mapped patterns accurately reflect distribution and density. In areas where datasets exhibit dissimilar species densities and distributions, we acknowledge potentially lower confidence and advise further sampling. To evaluate potential management implications resulting from a reliance on either dataset, low and high confidence regions were

attributed with a published ancillary dataset on cumulative human effects to identify where cetaceans are exposed to potential risk from human activities. Regions of agreement between density maps were primarily located off the southeastern coast of Haida Gwaii; areas of disagreement were heterogeneously distributed across the study area. Spatial clusters identified from citizen data exhibited statistically higher cumulative human effect scores than maps from systematic surveys (10.79 and 9.75 respectively [$p < 0.001$]), despite previous data corrections for human effort in data collection. Interestingly, the region southeast of Haida Gwaii has a high maximum cumulative effect (26.02), which indicates a region of conservation priority characterized by high cumulative human effect and cetacean density. Citizen science has the potential to broaden ecological exploration and further strengthen existing research. We provide a first step for evaluating how information varies by source. Although further research is being conducted to develop tools to integrate data from diverse sources, we demonstrate the use of citizen science to identify strategic areas for future data collection efforts and as a confirmatory dataset to augment survey datasets collected in the field.

## 3.2 Introduction

Historically, citizen science has played an important role in ecological investigations. Citizen scientists have informally contributed to ecological investigations for centuries, with noted examples of unpaid participation by non-professionals in scientific endeavours, such as Charles Darwin in the 19th century (Silvertown, 2009). Another prominent example is the Christmas Bird Count (CBC), organized by Frank Chapman in 1900 after 26 people replied to an article in the magazine *Bird-Lore* published by the National Audubon Society (Butcher, 1990). The CBC initiative was started to replace the practice of bird hunt competitions that were common place during this period (Chapman, 1900). Response to the publication began the initial tradition of annually recording bird sightings on Christmas Day instead of sport shooting, resulting in one of the largest databases of bird survey information (Butcher, 1990).

Today, citizen scientists have a more formally defined role in scientific studies of ecology. One of the best known examples of modern citizen science in ecology is the North American Breeding Bird Survey (BBS) where volunteers record counts of observed breading bird species. The BBS citizen science monitoring program, launched in 1966, had 2900 annually surveyed routes by 1997 (Sauer et al., 1997). Now, as one of the largest sources of comprehensive information on North American birds, it contributes valuable data on species distributions and population trends (Sauer et al., 2013). The successes of projects like the CBC and BBS are due, in part, to their ability to expand the geographic and temporal scope of data collection. Increasing the breadth and quantity of available information is a vital contribution to ecological analyses that are often

constrained using traditional frameworks and methods for collecting data. The value of the information gained in these initiatives has lead to a better understanding of species distributions and population trends across countries and continents.

Citizen science is also supporting the challenging task of acquiring species distribution and abundance information on marine organisms. Marine species, such as marine mammals, are difficult to survey given the elusive and migratory nature of many species, the complex 3D environment they inhabit, and the logistical and financial costs associated with accessing these environments. Further, problems arise from spatial and temporal inconsistencies typically present in traditional methods of data collection, often resulting in sparse and incomplete datasets that contain knowledge gaps. The efficiency and effectiveness of citizen science in ecological studies has been exhibited in numerous marine-based research programs that have successfully incorporated volunteers into scientific marine research. For example, a seahorse research project in Italy demonstrated that recreational scuba divers recording information on seahorse species were not only effective at identifying distributions and important habitat, but also actively contributed to the continual monitoring efforts of seahorse populations (Goffredo et al., 2004). Opportunistically collected data from citizens who frequently travel by boat (ships of opportunity) can be used to increase the number of observers on the water and sample sizes available for species distribution estimations (Williams et al., 2006).

Growth in GPS and GIS technology, as well as smart phones, has further contributed to an explosion of citizen science programs for ecological studies, including marine research (Thiel et al., 2014). The proliferation of mobile technology allows citizens to relate species observations to a universal indexing system (i.e., location) and to

disseminate information quickly. Also termed crowdsourcing, citizen science leverages observations and knowledge of individuals to collect large quantities of data that can span broad geographic extents. Citizen participants are typically familiar with digital technology, which allows for the inclusion of software tools and mobile applications that facilitate remote data collection. Incorporating public participants in scientific research expedites the collection of geographic information that can be used to confirm data gathered from disparate collection methods, contribute new observations, and provide assessments of data uncertainty.

Incorporating citizen science into the field of marine ecology offers many benefits and opportunities for current and future research; however, these datasets also have fundamental challenges. Issues related to effort bias, particularly the lack of sampling scheme, can invalidate many assumptions required for statistical study. Datasets may exhibit large spatial or temporal gaps and contain missing information leading to incomplete datasets. Further, participants may not have access to appropriate training or do not exhibit adequate competencies to correctly collect and record data. Although some recommendations and guidelines exist for outlining standards for citizen participation in ecological research (Silvertown, 2009), due to the newness of this field, widely accepted best practices are still under development and continue to evolve. The sheer volume of information available from citizen science programs makes overcoming concern of data quality a worthwhile pursuit. As well, data collected by trained research scientists are not immune from limitations and may also be sparse and limited in spatial and temporal resolution and extent. Therefore, both traditional survey and citizen science datasets are almost always inherently imperfect and incomplete.

Our hypothesis is that citizen science and traditional survey data have different information content that can be conflated to potentially offer the best available information and most complete representation of species distributions and densities. When combined, citizen science and systematic survey data have the potential to produce a more detailed picture of reality than if separately assessed. Accordingly, citizen science datasets can be viewed as complementary, rather than conflicting, to conventional scientific research studies, as each data type has the potential to illuminate different spatial patterns associated with ecological processes. If collated, the data may aid in forming a greater understanding of species distributions and abundance patterns. However, to date, few studies have focused on how to integrate citizen collected data with existing scientific investigations and geographic products.

To assess the information content of datasets, we use approaches from spatial analysis and leverage the link between spatial pattern and process to assess where datasets provide a common ecological signal and where there are dissimilarities. If the spatial patterns represented by the two datasets are similar, the same ecological processes are being captured by data collection. Differences indicate areas where we have less confidence in the signal of the data and further data collection may be warranted. Map comparison methods that employ spatial clustering techniques are often applied in remote sensing applications and landscape ecology, and offer pattern-based techniques that can provide ways to characterize similarities and differences between multiple datasets. Using quantitative methodology, we characterize map uncertainty and identify spatial patterns of similarity and difference that can be used to guide future sampling efforts and marine spatial planning initiatives.

## 3.3 Methods

### 3.3.1 Study region and species

The study area is situated on the continental shelf of British Columbia (BC) and contains diverse species, physical features, and oceanographic processes (Figure 3.1). Located in the northeastern expanse of the Pacific Ocean, our study region comprises Dixon Entrance, Hecate Strait, Queen Charlotte Sound, and Queen Charlotte Strait. In this region, current and wind-induced upwelling along the continental slope brings cool, nutrient rich water from the ocean floor that promotes increased ocean productivity (Thomson, 1981; Ford, 2014). As a transitional upwelling zone, this region supports high concentrations of primary producers that provide the foundation for complex food webs. Consequently, many cetaceans seek out these waters to feed upon the rich food sources available (Block et al., 2011).

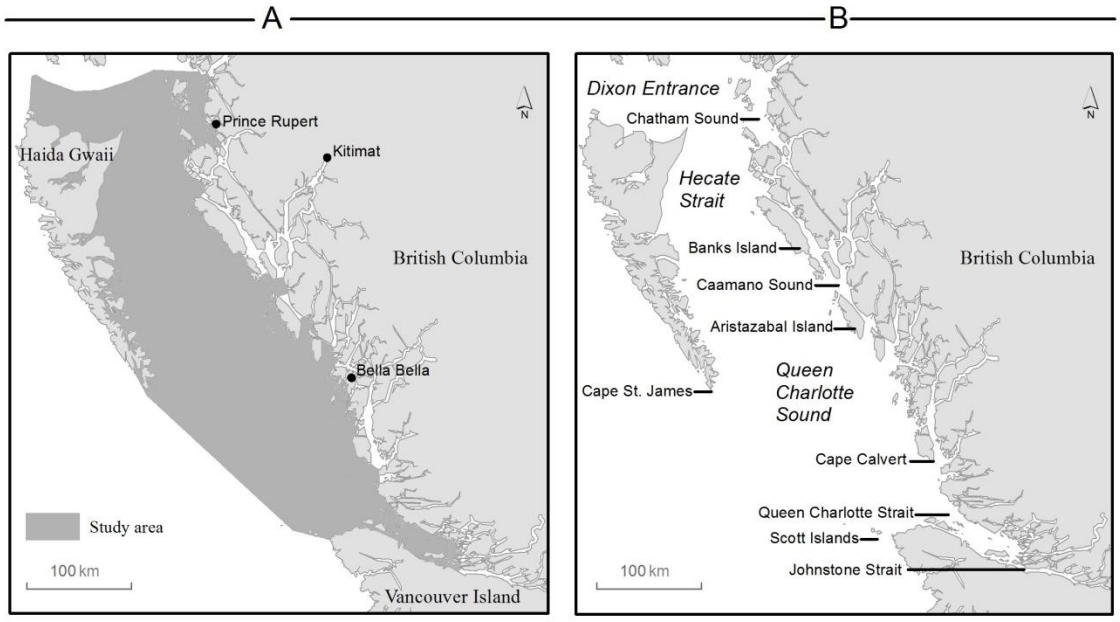


Figure 3.1 Study area (A) and significant coastal features (B).

The study region is also situated near a large number of human population centers and contains a wide array of anthropogenic activities that affect migrant and resident marine species, including cetaceans. Whale watching, recreational, and commercial vessel activity, marine ecotourism operations, ferry traffic, and cruise ship transits are all common marine-based activities occurring within the region. As a result, the exclusive economic zone of BC is an area of considerable anthropogenic impact, particularly along the coast and continental shelf (Murray et al., 2015). Human activity within the study region is substantial (Ban and Alder, 2008); with direct and indirect impacts arising from marine, land, and coastal sources.

Citizen science observations and systematic survey sightings for seven cetacean species commonly found in BC's continental shelf waters were sourced from a local citizen science program (BC Cetacean Sightings Network data, 2013; BCCSN) and a non-profit conservation organization (Raincoast Conservation Foundation; RCF). Cetacean species were selected based on availability of sightings from both sources. Selected species include common minke whale (MW; *Balaenoptera acutorostrata*), Dall's porpoise (DP; *Phocoenoides dalli*), fin whale (FW; *Balaenoptera physalus*), harbour porpoise (HP; *Phocoena phocoena*), humpback whale (HW; *Megaptera novaeangliae*), killer whale (KW, three ecotypes; *Orcinus orca*), and Pacific white-sided dolphin (PW; *Lagenorhynchus obliquidens*). Common minke whale, Dall's porpoise, and Pacific white-sided dolphin are listed as "not at risk" by the Committee on the Status of Endangered Wildlife in Canada (COSEWIC; www.cosewic.gc.ca). However, fin whales have been considered "threatened" since 2005, whereas both harbour porpoise and humpback whales are listed as "special concern". The three killer whale ecotypes

residing in the study region (northeast Pacific offshore, northeast Pacific transient, and northeast Pacific northern resident) have been assessed as threatened since 2008.

### 3.3.2 Systematic survey data

Cetacean observations for coastal BC were collected through scientific surveys (i.e., line transects using Distance Sampling) over the span of five years by trained RCF observers. Survey dates included were: Summer 2004 (June, July and August), Summer 2005 (August), Summer 2006 (August and September), Spring 2007 (April and May), and Summer 2008 (June and August). Fall sightings were not incorporated due to limited observations. During the survey period, cetacean sightings were observed from a 6.4 metre motorized sailing vessel (2004, 2006-2008) and a 20-metre powerboat (2005) by six-seven trained research scientists on a rotating schedule. The vessel traversed predefined survey transects (Thomas et al., 2007) to maintain effort efficiency. Cetacean sightings were recorded every nautical mile. Cetacean sightings per nautical mile were distance corrected using Multiple-Covariate Distance Sampling (MCDS) methods to account for the likelihood of increased observer error the further from the vessel a species was sighted (Best et al., 2015). See Thomas et al., (2007) and Williams et al. (2007) for additional survey design and observation protocol information.

Distance corrected density estimates from survey transect data were used in a previous study to generate maps of normalized cetacean density using predictive species-habitat models (see Chapter 2). A Generalized Additive Model (GAM) related distance corrected estimates of cetacean density for each species to 15 biological and physical environmental covariates. The resulting species density models were normalized and summed to generate a single map using a 13.86 $km^2$ hexagon grid representing total

cetacean density per nautical mile. This map was selected to represent cetacean density generated from pre-designed and established research methods and was used in map comparison analyses within this study.

Although distance sampled systematic surveys conducted by RCF might not be considered “traditional” or “conventional”, in this study, we refer to this information as a traditional. Survey effort was incorporated into the survey design and survey methodology was designed to maximize scientific rigor. Further, the survey design is fundamentally different compared with the haphazard collection of citizen science data. Therefore, in the context of this study, the naming convention for the cetacean density map generated from RCF surveys is hereafter referred to as traditional survey data.

### 3.3.3 Citizen science data

The BC Cetacean Sighting Network (BCCSN) is a research program designed as a central repository for opportunistic cetacean sightings recorded and submitted by citizens of BC. The BCCSN was founded in 2000 by the Vancouver Aquarium and Fisheries and Oceans Canada, as a citizen science program. By 2013, the success of the project was apparent with a database of more than 70,000 cetacean and sea turtle sighting records throughout coastal BC (Rechsteiner et al., 2013). Sightings are opportunistically collected and voluntarily submitted to BCCSN. Therefore, there is no predefined sampling strategy or designed survey. Citizens can voluntarily report a sighting through the Wild Whales website (whildwhales.org/sightings), telephone, email, logbook, or cell phone application.

In citizen science datasets, more data will always be derived from locations where human populations are higher. Moreover, variability in data quality can be associated

with experience of the citizen scientist. Each submitted report and subsequent information is screened and confirmed by BCCSN researchers to determine reliability and effectiveness of observers. Each sighting is assessed according to how effective each observer was at not only sighting wildlife in the water, but also correctly identifying the species and accurately reporting the observation (Rechsteiner et al., 2013). To account for the biases associated with citizen science data, the raw sightings data were also adjusted for observer effort (Rechsteiner et al., 2013). The BCCSN grouped observers into categories: large vessel crew, lighthouse keepers, residents of population centers, park users, ecotour operators, coastal workers, and frequent observers. Effort distribution was estimated based on predictable travel patterns and distributions that were considered 'typical' for each type of observer group. Typical patterns include metrics that could be geographically quantified, such as trip distance, distance to home port, travel routes, and maximum possible sighting distance (Rechsteiner et al., 2013). All groups were summed to generate a density surface displaying an index of total observer effort. The BCCSN provided effort corrected densities for each study species in the form of an index of Species Density per Unit Effort (SDUE) that will hereafter be referred to as citizen science cetacean data. More detailed information on effort correction procedures is available in Rechsteiner et al. (2013).

To allow datasets to be comparable, they needed to represent similar time periods. As the traditional survey cetacean data were not seasonally segregated, summer and winter maps from the citizen science dataset were aggregated to create a single surface representing total cetacean density. Citizen science data were normalized to values between 0 and 1, as per previous studies of species distributions to collate multiple

distribution maps (Nur et al., 2011; Fox et al., In Review). The resulting normalized annual cetacean SDUE map allows for comparison with the normalized traditional survey cetacean data.

### 3.3.4 Cumulative human effects data

To develop effective marine management plans, consideration of both social and biological factors is imperative. To include a social dimension in our analysis, cumulative human effect data – which maps the overlap of numerous anthropogenic activities that produce varying degrees of impact or stressors (Murray et al., 2015) – were integrated with conflated data on marine mammal distributions and densities. Cumulative human effects provide context for interpreting the relationship between humans and marine mammals through integrating human activity data with patterns of species distribution and density. When determining priority regions for future research, considering the level of potential stressors to species of interest or concern is particularly important. For cetaceans, understanding where species are encountering high levels of anthropogenic activity is essential, as these sources can disrupt or inhibit foraging, feeding, movement, or communication and, in extreme more extreme cases, cause morbidity or mortality (Fair and Becker, 2000). Cumulative human effect scores covering the exclusive economic exclusion zone of coastal BC were provided by Murray et al. (2015). Cumulative effects scores, also sometimes referred to as cumulative impact scores, describe the potential intensity of anthropogenic activity and subsequent level of impact from land, coastal, marine, or fishing sources (Murray et al., 2015). Murray and colleagues calculated an effect score over a 2 km grid by combining spatial locations of 47 human activities with 26 habitat classes assessed for level of vulnerability to anthropogenic impacts.

### 3.3.5 Map comparison through spatial patterns

To determine where cetacean distribution and density maps from citizen science and traditional approaches showed differences and similarities, we compared normalized density maps from both data sources using methods from spatial analysis and remote sensing. The challenge with comparing datasets is to quantify the magnitude of difference or similarity when map values are on different scales, subject to extreme outliers, and are generated from disparate data collection methods. Spatial pattern-based approaches overcome some of these challenges by allowing for the detection of similarities or differences in the geographic arrangement of data values rather than singularly focusing on specific numbers. Pattern-based methods stem from the concept of spatial autocorrelation where heterogeneity in dataset values varies with distance (Fortin and Dale, 2009). Here, we are interested in the relationship between a particular value and the values in a specified neighbourhood zone (i.e., search window) that allows for the detection of high or low clusters of values relative to other data in the map (i.e., local measures of spatial autocorrelation; Aldstadt, 2010). The identified spatial clusters of high or low values represent patterns of extreme values that could not have occurred due to random process (Getis, 2010). In this study we are interested in clusters of extreme values to identify areas of difference in traditional survey and citizen science maps, and areas of similarity.

To detect differences between map products, we compared the spatial patterns in each map using image differencing and spatial cluster analysis methods. Image differencing subtracts data values of one map from another to produce values that range from positive to negative values. Extremely positive and negative values indicate where

map values differ greatly. The traditional survey map was resampled to a 5 km grid to match citizen science dataset. Given differences in the scales of the citizen science and traditionally collected density data, both were normalized by converting 2,908 density values to ordinal ranks (Nelson et al., 2005). Rank ordering of values has been demonstrated in remote sensing as an efficient method for normalizing data in order to enable change detection (Nelson et al., 2005; Wulder et al., 2008). Rank ordering converts high species densities to high ranks that can be used to compare pattern, regardless of the scale of attribution in the original data. Further, the spatial configurations inherent in the map values are retained, therefore allowing for the detection of changes to spatial patterns and illuminating location differences. Rank values for each map were differenced (citizen science rank values subtracted from traditional survey rank values) to produce a single surface where extreme positive and negative values indicate regions of difference. Hence, applying image differencing techniques, facilitated by normalizing by rank-order, allows for the detection of disparities between mapped surfaces and can be applied in analyses outside of remote sensing applications.

We applied a cluster detection method from spatial statistics, Getis-Ord $G_i$*statistic (Getis and Ord, 1992), to the difference in ranks to delineate spatially explicit regions of difference in the traditional and citizen science datasets. These areas indicate where density value differences between maps are spatially clustered. We used the local Getis-Ord spatial statistic, $G_i$*, that identifies clusters of extreme high and low values relative to all density values within the study area (O'Sullivan and Unwin, 2010). The $G_i$*statistic is based on the foundational principles of spatial autocorrelation, where the null hypothesis assumes patterns in the data are a result of random processes (Getis,

2010). For every grid cell within the study area a spatial neighbourhood was defined through a spatial weights matrix. There are multiple ways to define a spatial neighbourhood, however given the coarse resolution of the cetacean datasets, a first order queen contiguity matrix was considered appropriate. Queen contiguity neighbourhoods include adjacent and diagonal grid cells in the $G_i$* calculation. Specifically, a first order queen contiguity spatial weights matrix was used for all tests, with 999 permutations and a 0.05 significance level, to compare clustered patterns against what would be expected from random process. The $G_i$* statistic was executed using GeoDa software (v.1.6.6 October 2014; Anselin et al., 2010) and visualizations generated in ESRI ArcGIS (v 10.2). The resulting outputs were categorized based on whether the difference map generated clusters of extreme positive or negative values. Clusters generated around extremely high positive values were grouped and clusters around extremely low values were placed in a second category.

To assess where both maps display similar spatial patterns of cetacean density, spatial cluster detection using the $G_i$* statistic was performed on the original normalized maps to identify where density is high compared with surrounding regions. In this case, a first order queen contiguity spatial weights matrix was used (999 permutations and a 0.05 significance level) to identify only spatial clusters of extremely high values. Citizen science and traditional survey map clusters were overlaid to identify regions where clusters overlapped. Overlapping regions indicate where both maps identify high cetacean densities.

The penultimate test for many citizen science programs is to contextualize the information they provide about species conservation. To characterize how traditional and

citizen science datasets may identify and present areas of potential conservation concern similarly and differently, spatial clusters of difference and similarity were overlaid with the cumulative human effect surface provided by Clarke Murray et al. (2015). Following, summary statistics were calculated to characterize effects and identify areas of intense human activity. Polygons with cumulative human effect scores were extracted for each difference cluster category (when traditional survey data predicted higher values than citizen science and vice versa). A t-test was performed on the cumulative effect scores to identify the average effect score for each category and to determine if the means of the data distributions were significantly different.

### 3.4 Results

Maps of cetacean density, collected through citizen science and traditional survey methods, showed considerable disparities when the rank-orders of each map were differenced and spatial clusters identified (Figure 3.2). The traditional surveys generally predicted higher cetacean density values than citizen scientists in many offshore locations, illustrated by large clusters in regions of Hecate Strait and central Queen Charlotte Strait. In contrast, citizen science densities were greater than traditional surveys in coastal regions and along the shoreline, shown by clusters along the northern edge and southeastern side of Haida Gwaii, adjacent to Banks Island, and adjacent to northern Vancouver Island and into Queen Charlotte Strait.

Spatial clusters of high density from original citizen science and traditional survey maps overlapped to produce five distinct zones where both datasets showed agreement regarding the location of high cetacean density values (Figure 3.3; zone 1 – 5). Areas of similarity were smaller than the difference clusters and were scattered throughout the

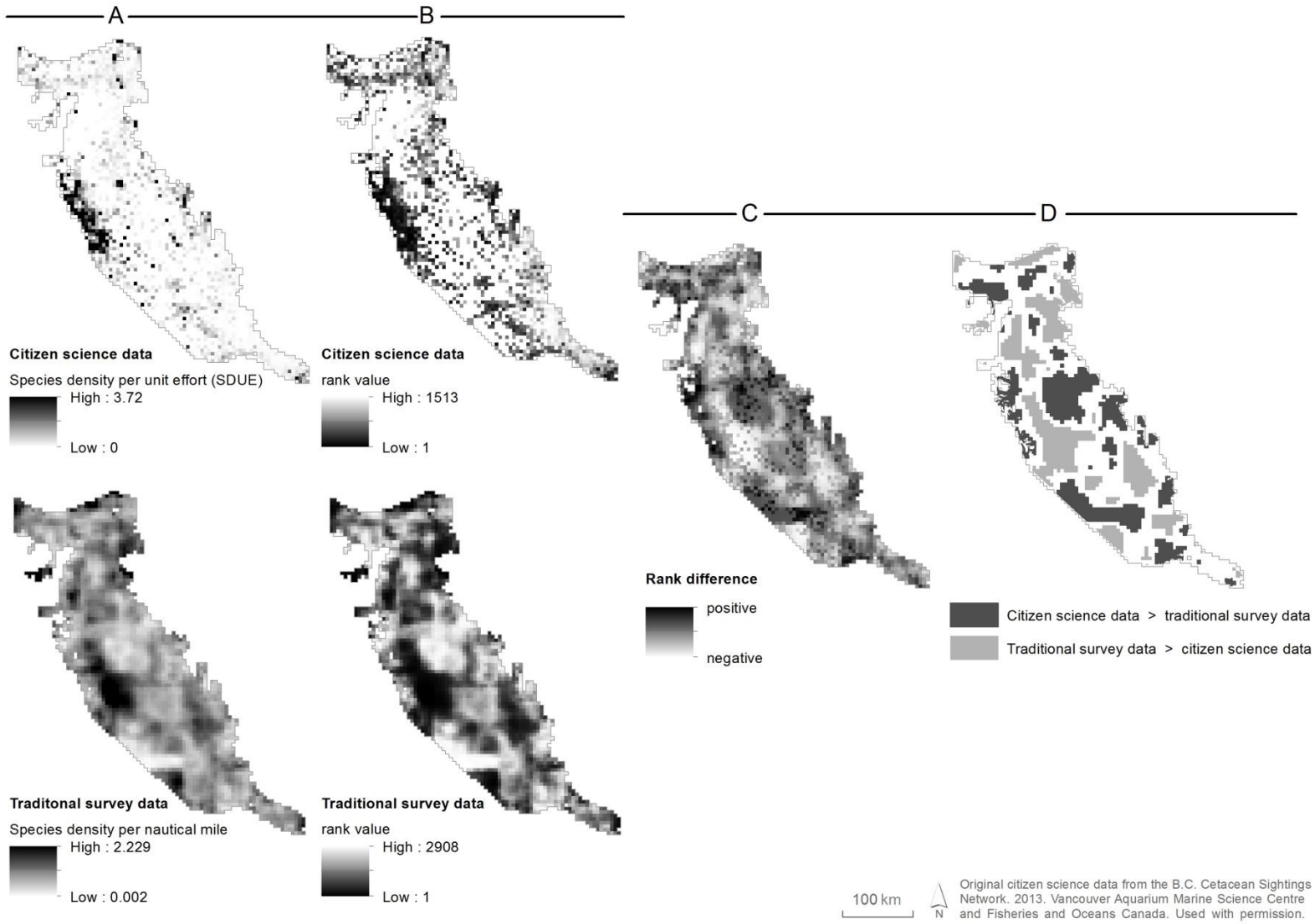


Figure 3.2: To identify regions of difference, original cetacean density maps (Panel A) were first converted to rank order. Traditional survey and citizen science rank order maps (Panel B) were then differenced to generate a surface identifying areas of disparity between maps (Panel C). Rank difference values that are strongly positive or negative indicate where density estimates differ between maps. Positive difference values indicate where citizen science displayed higher densities (i.e., lower rank order), while negative values identify regions where the traditional survey map predicted lower densities (i.e., higher rank order). The G*i**statistic was used to identify spatial clustering of ranked difference values (Panel D). Clustering of high positive values indicate where citizen science data identifies greater cetacean density than traditional survey maps and where clustering of extreme negative values indicates where traditional survey data predicts higher density values.

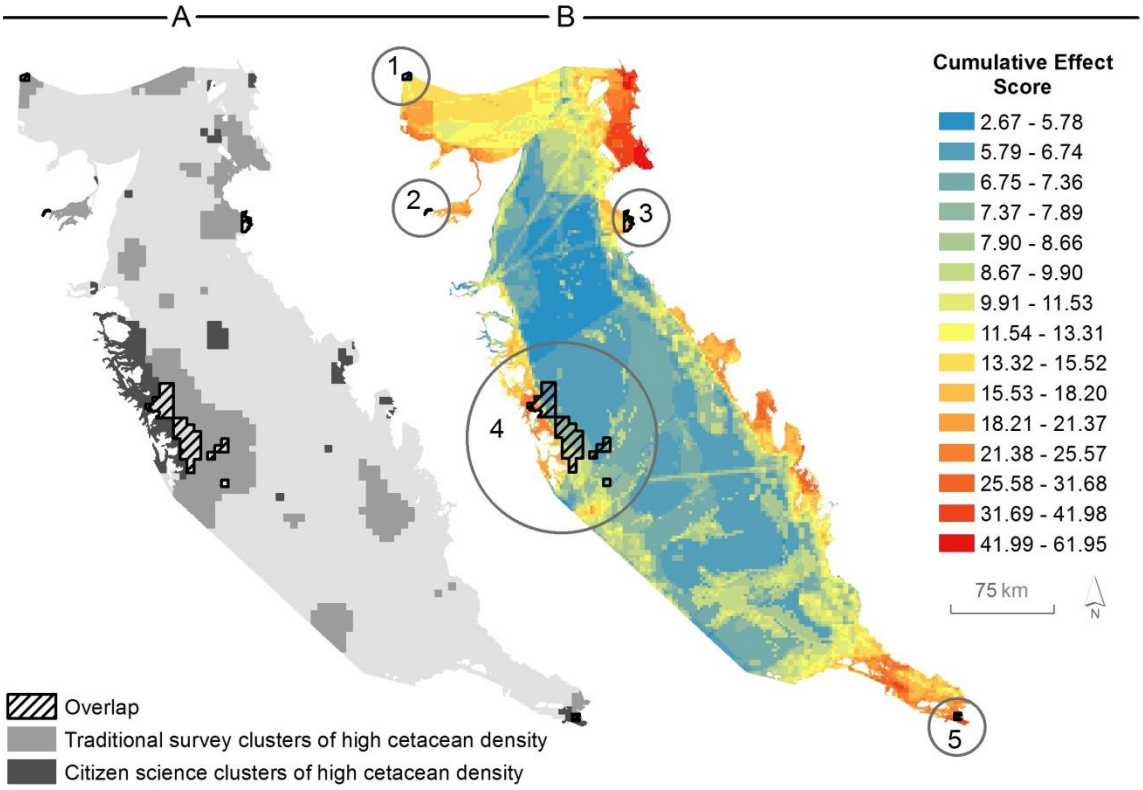


Figure 3.3 Spatial clusters of high cetacean density were calculated from original density values from traditional data and citizen science maps (Panel A). Panel B depicts regions where clusters overlap (Zones 1-5) and indicates where both maps identify areas of high cetacean density. Zones are overlaid with cumulative human effect scores to characterize human threats.

study region. One prominent cluster, identified as zone 4, was larger than the rest and located southeast of Haida Gwaii.

Difference and similarity clusters were overlaid with cumulative human effects to compare the level of anthropogenic activity and potential threats to cetaceans (Figures 3.3 and 3.4). Where citizen science displayed higher densities than traditional data the average cumulative human effect score was 10.79. Inversely, where traditional data predicted higher densities than the citizen science dataset, the average effect score was less, 9.75. A t-test performed on the average effect scores for spatial clusters of difference

revealed that the means for each cluster category were significantly different (t-test; t = -7.9022, df = 9162, p<0.001; Figure 3.5). Notably, the human effect scores were more variable in the cluster category where traditional data showed higher density values, with a large spread in data values ranging from 3.04 to 61.95. Conversely, where citizen science data predicted higher density values, the data were less variable with scores not exceeding 46.94. Spatial clusters identifying where both maps showed similar spatial patterns were grouped into zones 1, 2, 3, 4 and 5, with maximum effect scores of 14.13, 16.68, 19.06, 26.02, and 40.49, respectively.

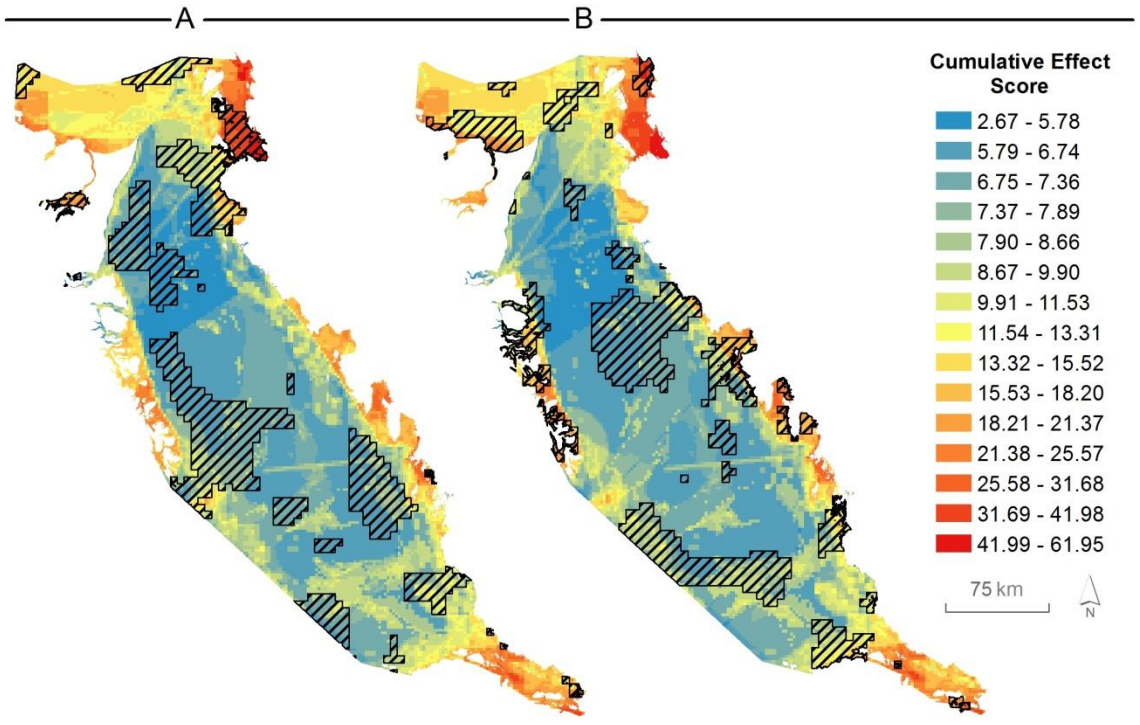


Figure 3.4 Spatial clustering of difference values identified from the ranked difference surface specifies spatial zones of difference between each map. Panel A shows where traditional survey data predicted greater cetacean density than the citizen science map, while panel B identifies where the citizen science identified higher values. Spatial clusters are overlaid with cumulative human effects.

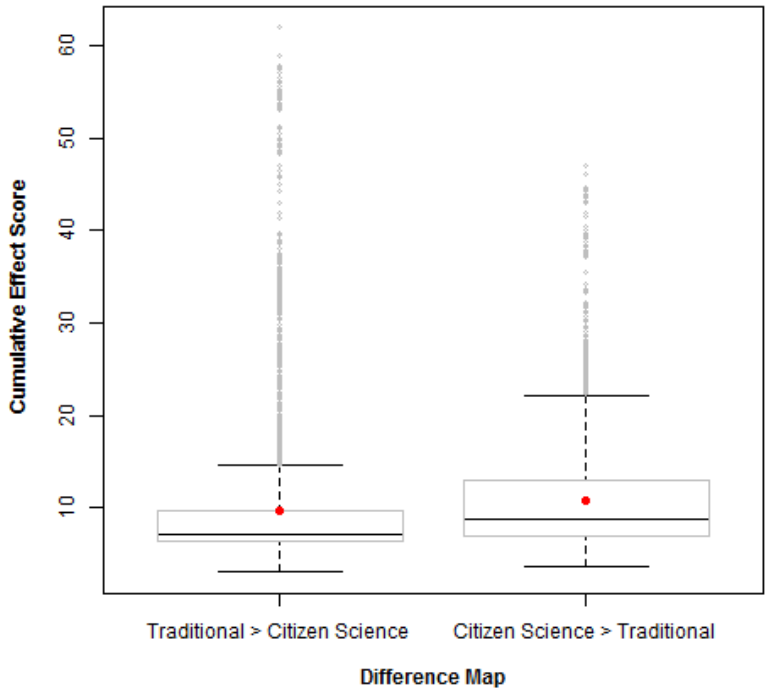


Figure 3.5 Cumulative human effect per difference map. Note that the means of "Traditional > Citizen Science" and "Citizen Science >Traditional" maps are significantly different with 10.79 and 9.75 respectively (p value 3.057e-15).

## 3.5 Discussion

Our analysis, using pattern-based approaches for map comparison, identified numerous spatial clusters of difference between citizen science and traditional survey maps that were scattered heterogeneously throughout the study area. Where citizen science datasets identified higher cetacean densities than traditional survey methods, we note that these areas are located next or near to the shoreline and areas where shipping, transportation, and fishing activities are common (Ban et al., 2013). Clusters close to the shore might also indicate recreational activities and tourism operations, where vessels typically remain close to land and in short distances to population centers and marinas. As a result, citizen scientists may be identifying cetacean densities that are greater than what was predicted through traditional survey methods because of biased and uneven

survey effort. The traditional surveys, in contrast, generally predict higher cetacean densities in offshore regions. Systematic surveys aim to provide equal coverage throughout the study area and to maintain effort efficiency and, therefore, are likely to traverse regions that are not frequented as regularly by people. This concept is further discussed by the cumulative human effect scores that show a lower average score where traditional surveys predict higher cetacean densities.

The spatial clustering technique identified five regions of similarity that were primarily located in coastal areas known for high ocean productivity. The most prominent area where citizen science and traditional surveys identified high cetacean density was located in zone 4 located off the southeastern area of Haida Gwaii. Zone 4 is known to support intense eddy circulation and tidal current activity that promotes ocean mixing and elevated concentrations of primary productivity (Crawford et al., 1995; Jardine et al., 1993). This region exhibits enhanced nutrient supply as a consequence of reliable chlorophyll-a blooms (Robinson et al., 2004), which generates food sources that attract top marine predators (e.g., Fox et al., In Review and Chapter 2). Notably, zone 5 (the most southerly cluster) is located at the northern end of Johnstone Strait, around Weynton Passage and Blackfish Sound, where high current waters are funneled between the shoreline and nearby islands (Thomson, 1981). Intense currents facilitate the vertical displacement of water resulting in nutrient rich cold waters that promote concentrations of food resources (Thomson, 1981). Cetaceans are known to frequent this region, in part due to the prevalence of Pacific salmon and the Robson Bight rubbing beaches that are of importance to killer whales (Ford, 2014). Given that both citizen science and traditional survey maps show agreement in the five identified zones, and biophysical oceanographic

knowledge supports the conclusion of high cetacean densities, we can infer with higher confidence that the clusters produced are reflecting biological reality.

By applying pattern-based map comparison methods, we mapped where each collection technique identified similar clustering patterns of high cetacean density and distinguish locations for future conservation based on agreement between maps and level of anthropogenic threat. Regions where clusters of high density values in both maps overlap are indicative of areas where there is greater confidence in identified aggregations, and therefore, represent locations of potential conservation significance. Of particular interest are locations where map agreement (high confidence) is coupled with intense anthropogenic activity. These locations suggest regions where high densities of cetaceans are exposed to high cumulative anthropogenic effects that may elevate stress in populations. Accordingly, these regions are of particular concern and could suggest a need for expeditious conservation action. In contrast, mapping dissimilar patterns provides specific guidance on areas where there may be less certainty and identify areas for additional sampling effort.

The largest cluster of similarity, zone 4, is located adjacent to the Gwaii Haanas National Marine Conservation Area (GHNMCA) reserve that is situated along the south east coast of Haida Gwaii. Although there are two Rockfish conservation regions (14% of reserve area) and six fully protected sub-areas (3% of reserve area) in GHNMCA, most of this reserve allows fishing, recreation, and other human activities (Parks Canada, 2010). Allowable human activities in this area may explain the higher than expected cumulative human effect scores identified for zone 4 (maximum score of 26.02).

Given the multitude of cumulative human effects in the study region as well as government initiatives to protect marine species and habitats, it is essential to spatially quantify important areas. Further, quantitative spatial data increase in value if current knowledge has been validated or assessed for confidence. Hence, the geographic data products provided in this study can be used to help meet international and national objectives and targets. For example, in order to meet the Aichi Target 11 to protect 10% of marine areas and achieve other government objectives, GHNMCA borders could be extended or conservation efforts could be shifted to other important areas identified in this study. Robson Bight Ecological Reserve is located just southeast of zone 5 (at the northern entrance to Johnstone Straight). Although protecting the important rubbing beaches for killer whales, zone 5 suggests that cetaceans are also present in high densities just north of the reserve. Consequently, zone 5 represents another region of high cetacean importance and, therefore, potential consideration for protection may be warranted. Identified zones 1, 2, and 3 are located in northeast Dixon Entrance, within the eastern portion of Masset Inlet, and north of Banks Island, respectively. Although not in close proximity to protected areas, zones 1-3 may also warrant conservation consideration and could represent potential priority locations for future marine management and conservation action.

It is important to consider that incorporating citizen generated data – especially opportunistically and haphazardly collected information – into scientific study is a challenging endeavour particularly due to imperfections, inconsistencies, and biases in data and effort that is temporally or geographically inconsistent. Citizen science datasets may not be uniform across space or through time making conventional analysis

techniques unusable (Fink et al., 2009). Consequently, these data are often approached with caution and hesitation. Issues related to survey effort, bias, data quality have been cited as key limitations and concerns for citizen science datasets (Fitterer et al., 2013; Goffredo et al., 2004; Tulloch et al., 2013) and may add additional noise and variability to datasets that hinder the detection of the ecological signal under investigation. These apprehensions often lead to concern over data credibility and whether citizen science data and volunteered geographic information can be viewed as a credible source of scientific knowledge (Bonney et al., 2014; Flanagin and Metzger, 2008). Concern can be partially alleviated by providing in-depth training to citizens (Bonney et al., 2014), standardized and well outlined data collection methods (Silvertown, 2009), as well as data validation and correction based on volunteer competencies (Rechsteiner et al., 2013). Further, recommendations include developing protocols for data collection, such as data entry forms with filters to reduce recording errors and flagging anomalies for verification (Hochachka et al., 2012). Notwithstanding stated concerns, literature identifies an increasing trend around the use of citizen science data in marine focused research and projects this trend to continue (Thiel et al., 2014).

Despite concerns of quality, citizen collected data can substantially increase the amount of data available on a particular phenomenon or organism. There is an opportunity to consider carefully how we can leverage information content from citizen science data, while acknowledging real limitations. Pattern-based methods for collating multiple maps can be applied in a wide array of contexts from model validation and accuracy assessment, to developing prioritization regions for spatial planning efforts. Current marine and coastal research projects in the United States and Canada have

already applied techniques to capitalize on the data collected through citizen science programs, such as monitoring initiatives (Delaney et al., 2008; Wilson et al., 2013), species population metrics (James et al., 2006; Pattengill-Semmens and Semmens, 2003; Tonachella et al., 2012) and management (Baird et al., 2002; Fitterer et al., 2013) among many others. However, methods from spatial statistics can be further generalized; for example, from marine studies to terrestrial research, or from ecologically focused to human centered applications. Approaches implemented in this study illustrate ways to combine spatial data from various disparate sources to provide a more comprehensive knowledge base and assess confidence in the information provided by spatial data.

In conclusion, citizen science contributes to scientific study by validating current ecological knowledge and helping to form a more complete understanding of marine organisms and species. However, further research on how to address the uncertainty of citizen science datasets is required to allow us to leverage the high resolution and broad spatial extent of citizen science data, while also accounting for potential limitations with quality. Citizen science and conventional research each have the potential to produce information that fills different knowledge gaps and, as a result, neither is capable of addressing the scope of regional analyses in its entirety. Developing new and innovative ways to combine citizen science and traditionally collected data, produces increased quantities of spatial data available to support management initiatives. Spatial data of species across broad spatial scales is vital for developing responsible government policies and for making informed decisions in the face of climate change and biodiversity declines (Dickinson et al., 2010). Therefore, incorporating citizen science into management and decision-making for marine systems and species is a move in the right

direction to ensure stakeholder involvement, increased transparency of scientific research, and the application of sound judgement in light of data uncertainty.

**Acknowledgements**

We thank the BC Cetacean Sightings Network, the Vancouver Aquarium, and the Department of Fisheries and Oceans for providing citizen science data on cetacean densities, Raincoast Conservation Foundation and Rob Williams for transect survey data, and the World Wildlife Fund Canada and colleagues for providing the cumulative human effects dataset (results and recommendations do not necessarily represent the official position of WWF Canada). We also acknowledge and greatly thank the many volunteers and stakeholders who participated in collecting the datasets we used in this study. This research could not have been conducted without the generous funding and support from the Natural Sciences and Engineering Research Council of Canada (NSERC), Raincoast Conservation Foundation, and the Spatial Pattern Analysis and Research Lab (SPAR) at the University of Victoria.

**References**

Aldstadt, J. 2010. Spatial Clustering. Chap B.4 *in* M. M. Fischer and A. Getis, editors. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Heidelberg, Berlin.

Anselin, L., I. Syabri, and Y. Kho. 2010. A.4 GeoDa: An Introduction to Spatial Data Analysis. Pages 73–89 *in* M. M. Fischer and A. Getis, editors. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Heilelberg, Berlin.

Baird, R. W., P. J. Stacey, D. A. Duffus, and K. M. Langelier. 2002. An evaluation of gray whale (*Eschrichtius robustus*) mortality incidental to fishing operations in British Columbia, Canada. Journal of Cetacean Research and Management 4:289—296.

B.C. Cetacean Sightings Network data. 2013. Vancouver Aquarium Marine Sciences Centre and Fisheries and Oceans Canada. Vancouver, B.C.

Ban, N., and J. Alder. 2008. How wild is the ocean? Assessing the intensity of anthropogenic marine activities in British Columbia, Canada. Aquatic Conservation: Marine and Freshwater Ecosystems 18:55—85.

Ban, N. C., K. M. Bodtker, D. Nicolson, C. K. Robb, K. Royle, and C. Short. 2013. Setting the stage for marine spatial planning: Ecological and social data collation and analyses in Canada's Pacific Waters. Marine Policy 39:11—20.

Best, B. D., C. H. Fox, R. Williams, P. N. Halpin, and P. C. Paquet. 2015. Updated marine mammal distribution and abundance estimates in British Columbia. Journal of Cetacean Resource Management 15:9–26.

Block, B. A., I. D. Jonsen, S. J. Jorgensen, A. J. Winship, S. A. Shaffer, S. J. Bograd, E. L. Hazen, D. G. Foley, G. A. Breed, A.-L. Harrison, J. E. Ganong, A. Swithenbank, M. Castleton, H. Dewar, B. R. Mate, G. L. Shillinger, K. M. Schaefer, S. R. Benson, M. J. Weise, R. W. Henry, and D. P. Costa. 2011. Tracking apex marine predator movements in a dynamic ocean. Nature 475:86—90.

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

Bonney, R., J. L. Shirk, T. B. Phillips, A. Wiggins, H. L. Ballard, A. J. Miller-Rushing, and J. K. Parrish. 2014. Next Steps for Citizen Science. Science 343:1436—1437.

Butcher, G. S. 1990. Audubon Christmas Bird Counts. Pages 5—13 *in* Survey designs and statistical methods for the estimation of avian population trends. J. R. Sauer and S. Droege, editors. U. S. Fish Wildl. Serv., Biol Rep. 90(1). 166 pp.

Chapman, F. M. 1900. A Christmas bird-census. Bird-Lore 2:192

Crawford, W. R., and M. J. Woodward, M. G. G Foreman, and R. E. Thomson. 1995. Oceanographic features of Hecate Strait and Queen Charlotte Sound in Summer. Atmosphere—Ocean 33:639—681.

Delaney, D. G., C. D. Sperling, C. S. Adams, and B. Leung. 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biological Invasions 10:117—128.

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

Dickinson, J. L., J. Shirk, D. Bonter, R. Bonney, R. L. Crain, J. Martin, T. Phillips, and K. Purcell. 2012. The current state of citizen science as a tool for ecological research and public engagement. Frontiers in Ecology and the Environment 10:291—297.

Fair, P. A., and P. R. Becker. 2000. Review of stress in marine mammals. Journal of Aquatic Ecosystem Stress and Recovery 7:335—354.

Fink, D., W. M. Hochachka, M. J. Iliff, C. L. Wood, B. L. Sullivan, K. V. Rosenberg, M. Arthur Munson, M. Riedewald, and S. Kelling. 2009. Obtaining New Insights for Biodiversity Conservation from Broad-Scale Citizen Science Data. Nature Precedings 11. 17 pp.

Fitterer, J. L., T. A. Nelson, N. C. Coops, M. A. Wulder, and N. A. Mahoney. 2013. Exploring the ecological processes driving geographical patterns of breeding bird richness in British Columbia, Canada. Ecological Applications 23:888—903.

Flanagin, A. J., and M. J. Metzger. 2008. The credibility of volunteered geographic information. GeoJournal 72:137—148.

Ford, J. K. 2014. Marine Mammals of British Columbia. Royal BC Museum, Victoria, Canada. 460 p.

Fortin, M., and M. R. T. Dale. 2009. Spatial Autocorrelation. Chap. 6. Pages 89–103 *in* S. A. Fotheringham and P. A. Rogerson, editors. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Fox, C. H., Huettmann, F., Harvey, G. K., Morgan, K., Robinson, J., and Paquet, P.C. (In Review). Predictions from Machine Learning ensembles: marine bird distribution and density on Canada's Pacific coast. Marine Ecology Progress Series.

Getis, A. 2010. Spatial Autocorrelation. Chap. B.3. *in* M. M. Fischer and A. Getis, editors. Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications. Springer-Verlag, Heidelberg, Berlin.

Getis, A., and J. K. Ord. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24:189–206.

Goffredo, S., C. Piccinetti, and F. Zaccanti. 2004. Volunteers in Marine Conservation Monitoring: a Study of the Distribution of Seahorses Carried Out in Collaboration with Recreational Scuba Divers. Conservation Biology 18:1492—1503.

W. M. Hochachka, D. Fink, R. A. Hutchinson, D. Sheldon, W. Wong, and S. Kelling. 2012. Data-intensive science applied to broad-scale citizen science. Trends in ecology & evolution 27:130—137.

Hooker, S. K., and L. R. Gerber. 2004. Marine Reserves as a Tool for Ecosystem-Based Management: The Potential Importance of Megafauna. BioScience 54:27—39.

James, M. C., S. A. Sherrill-Mix, K. Martin, and R. A. Myers. 2006. Canadian waters provide critical foraging habitat for leatherback sea turtles. Biological Conservation 133:347—357.

Jardine, I. D., K. A. Thomson, M. G. Foreman, and P. H. LeBlond. 1993. Remote Sensing of Coastal Sea-Surface Features off Northern British Columbia. Remote Sensing of Environment 45:73—84.

Kovacs, K., A. Aguilar, D. Aurioles, V. Burkanov, C. Campagna, N. Gales, T. Gelatt, S. D. Goldsworthy, S. J. Goodman, G. J. G. Hofmeyr, T. Härkönen, L. Lowry, C. Lydersen, J. Schipper, T. Sipilä, C. Southwell, S. Stuart, D. Thompson, and F. Trillmich. 2012. Global threats to pinnipeds. Marine Mammal Science 28:414–436.

Murray, C. C., S. Agbayani, H. M. Alidina, N. C. Ban. 2015. Advancing marine cumulative effects mapping: An update in Canada's Pacific Waters. Marine Policy 58:71—77.

Nelson, T., H. G. Wilson, B. Boots, and M. A. Wulder. 2005. Use of ordinal conversion for radiometric normalization and change detection. International Journal of Remote Sensing 26:535—541.

Nur, N., J. Jahncke, M. P. Herzog, J. Howar, K. D. Hyrenbach, J. E. Zamon, D. G. Ainley, J. A. Wiens, K. Morgan, L. T. Ballance, and D. Stralberg. 2011. Where the wild things are: predicting hotspots of seabird aggregations in the California Current System. Ecological Applications 21:2241–2257.

O'Sullivan, D., and D. J. Unwin. 2010. Geographic Information Analysis (second ed.). John Wiley & Sons, Hoboken, NJ.

Parks Canada. 2010. Gwaii Haanas National Marine Conservation Area Reserve and Haida Heritage Site: Interim Management Plan and Zoning Plan. Available from http://www.pc.gc.ca/eng/pn-np/bc/gwaiihaanas/plan/plan1.aspx

Pattengill-Semmens, C. V., and B. X. Semmens. 2003. Conservation and Management Applications of the Reef Volunteer Fish Monitoring Program. Pages 43—50 *in* Coastal Monitoring through Partnerships. Springer Netherlands.

Spatial Autocorrelation. Chap. 6. Pages 89–103 *in* S. A. Fotheringham and P. A. Rogerson, editors. The Sage Handbook of Spatial Analysis. SAGE Publications Inc., Thousand Oaks, California.

Pompa, S., P. R. Ehrlich, and G. Ceballos. 2011. Global distribution and conservation of marine mammals. Proceedings of the National Academy of Sciences 108:13600—13605.

Rechsteiner, E. U., C. F. C., Birdsall, D. Sandilands, I. U., Smith, A. V., Phillips, L. G. Barrett-Lennard. 2013. Observer effort for wildlife sightings data. Quantifying observer effort for opportunistically-collected wildlife sightings. Report of the Vancouver Aquarium Cetacean Research Program. 49 pp.

Robinson, C. L. K., J. F. R. Gower, and G. A. Borstad. 2004. Twenty years of satellite observations describing phytoplankton blooms in seas adjacent to Gwaii Haanas National Park Reserve, Canada. Canadian Journal of Remote Sensing 30:36—43.

Sauer, J. R., J. E. Hines, G. Gough, I. Thomas, and B. G. Peterjohn. 1997. The North American Breeding Bird Survey Results and Analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD. http://www.mbr-pwrc.usgs.gov/bbs/genintro.html

Sauer, J. R., W. A. Link, J. E. Fallon, K. L. Pardieck, D. J. Ziolkowski Jr. 2013. The North American Breeding Bird Survey 1966—2011: Summary Analysis and Species Accounts. North American Fauna 79:1—32.

Schipper, J., J. Chanson, F. Chiozza, N. Cox, M. Hoffmann, V. Katariya, V., and … B. Young. 2008. The status of the world's land and marine mammals: Diversity, threat, and knowledge. Science 322:225–230.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology and Evolution 24:467—471.

Thiel, M., A. Penna-Días, G. Luna-Jorquera, S. Salas, J. Sellanes, W. Stotz. 2014. Citizen Scientists and Marine Research: Volunteer Participants, their Contributions, and Projection for the Future. Oceanography and Marine Biology: An Annual Review 52:257—314.

Thomas, L., R. Williams, R., and D. Sandilands. 2007. Designing line transect surveys for complex survey regions. Journal of Cetacean Research and Management 9:1–13.

Thomson, R. 1981. Oceanography of the British Columbia Coast. Canadian Special Publication of Fisheries and Aquatic Sciences 56:291 p.

Tonachella, N., A. Nastasi, G. Kaufman, D. Maldini, R. W. Rankin. 2012. Predicting trends in humpback whale (*Megaptera novaeangliae*) abundance using citizen science. Pacific Conservation Biology 18:297—309.

Toomey, A. H., and M. C., Domroese. 2013. Can Citizen Science lead to positive conservation attitudes and behaviors? Human Ecology Review 20:50—62.

Tulloch, A. I. T., H. P. Possingham, L. N. Joseph. J. Szabo, and T. G. Martin. 2013. Realising the full potential of citizen science monitoring programs. Biological Conservation 165:128—138.

Williams, R., and L. Thomas. 2007. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada. Journal of Cetacean Research and Management 9:15–28.

Wilson, S., E. M. Anderson, A. S. G. Wilson, D. F. Bertram, and P. Arcese. 2013. Citizen Science Reveals an Extensive Shift in the Winter Distribution of Migratory Western Grebes. PLoS ONE 8: e65408.

# 4.0 CONCLUSIONS

## 4.1 Discussion and conclusions

As stated in the Government of Canada's 2016 budget, over $81.3 million will be allocated to high-priority agendas that focus on increasing ocean science and monitoring research, maintaining the highest standards for clean and sustainable economic growth, and affording adequate protection to marine and coastal regions (Canadian Government, 2016). Additionally, to achieve the Aichi Biodiversity Targets developed at the Convention of Biological Diversity, Canada must protect 525,000 square km of ocean habitat by 2020 (Canadian Government, 2016). However, for the majority of marine mammals species in Canada, knowledge gaps are outstanding at local and regional scales – the Wallacean shortfall (Mutke, 2011) – and often fail to provide current and reliable spatial data that would aid in fulfilling the Canadian government's stated objectives.

Developing appropriate management strategies to mitigate anthropogenic impacts and make progress towards conservation targets is challenging given the relative lack of spatial data for marine mammal populations (Kovacs et al., 2011; Parsons et al., 2015). To fill data gaps there is a need for unbiased and robust methods for prioritizing regions for future conservation and targeted marine management actions that will result in the largest positive impact with limited available resources. Further, because data from systematic survey cannot fill data gaps alone, there is room to incorporate data collected via unconventional means to validate information content of existing datasets and provide measures of data uncertainty.

In Chapter 2, we developed a predictive model to generate species-specific maps of marine mammal density which were subsequently used to identify areas of importance through spatial hotspot analyses and aspatial thresholds. Using a Generalized Additive Model (GAM) we derive spatially continuous maps for nine marine mammal species that produced model accuracies consistent with previous marine mammal research (e.g., Best et al., 2015 and Mannocci et al., 2014); however, improvement is possible if additional survey data become available and sample sizes increase. Normalizing and collating species-specific maps, we produced spatial estimates of species density for cetaceans, pinnipeds and all species combined. Comparison of hotspot analysis approaches revealed that hotspot size and quantity were greatly influenced by the method applied. The aspatial threshold method produced large quantities of smaller and more conservative hotspots. Conversely, hotspots identified using distance – spatial neighbourhoods using distance metric defined by a semivariogram range value – produced the fewest number of hotspots and covered the largest total geographic area. Hotspots showed agreement with ecologically and biologically important areas identified through expert opinion (Clarke and Jamieson, 2006; Jamieson and Levesque, 2014) suggesting that hotspots generated through spatial statistical methodology reflect biological reality. Our analyses were successful in identifying regions of importance for marine mammals through objective and robust statistical techniques and recommend the use of statistical hotspot analysis for future analyses outside our study region.

In Chapter 3, we validate our cetacean density model by applying map comparison techniques using cetacean sightings data collected through the citizen science program BC Cetacean Sightings Network. We applied image differencing techniques to

identify where both cetacean density maps described similar spatial patterns and to isolate geographic locations where density values differed. Using the $G_i$*statistic, five regions of similarity were identified representing areas where we are confident high density cetaceans are present, while regions of difference were heterogeneously distributed through the study area representing regions of lower confidence in need of further sampling. Additionally, marine spatial planning principles include a holistic approach to marine management that does not assess elements of the environment in isolation (Foley et al., 2010). As a result, we incorporated cumulative human effects data, produced by the World Wildlife Fund Canada and colleagues, to include human effects into our analysis and provide context to the relationship between marine mammal density patterns and human activities. In areas where systematic survey and citizen science data agree on areas of high cetacean density, as well as exhibit high levels of cumulative human effects (e.g., southeastern coastal region of Haida Gwaii) we suggest as priority locations for future conservation. Our analysis demonstrates the value of citizen science data for validating existing knowledge and enhancing our overall understanding of marine mammal distributions and densities.

### 4.2 Research contributions

Although marine mammals are a widely studied species in a global context, there remains a methodological gap regarding the spatial inputs used for marine mammal conservation and management efforts in North America and, more specifically, in Canada. For that reason, our research was centered on providing method focused solutions for the development of accurate predictive distribution and density models for spatially delineating regions of importance for marine mammals in British Columbia

(BC). As ecosystem based management and marine spatial planning have been described as data-hungry processes (Wright and Kyhn, 2015) that require data from a wide range of ecological and socioeconomic factors, our research was successful in contributing valuable spatial data and map products necessary for conservation prioritization, effective decision-making, and to meet government objectives and targets.

Previous research of marine top predators employs the use of GAM models to relate environmental characteristics to species observations to predict values at unsampled locations (e.g., Becker et al., 2010, Dalla Rosa et al., 2012, Embling et al., 2010, and Fox et al., In Review). Maps of species distributions are valuable for conservation prioritization, however multiple methods are available but many are limited based on the quantity and quality of available data (Redfern et al., 2006; Elith and Leathwick, 2009). A key contribution of this research is the development of a model using up-to-date environmental covariates, including modern remotely sensed variables, to develop spatially continuous maps of marine mammal densities. We contribute to the advancement of sophisticated modelling techniques by including comprehensive suite of environmental covariates and incorporating model parameters to combat data limitations; such as reduced gamma value and modified basis functions to prevent model overfitting, as well as a weighting scheme to compensate for zero-inflated data. These techniques can be generalized to future research studies external to our study area and study species, including species that inhabit the land and sea interface, ecosystems under intense anthropogenic pressure (e.g., coral reefs), and regions exhibiting drastic environmental change (e.g., polar ecosystems). As a result, we have generated a new transferable model for species density prediction that is statistically robust and ecologically sound.

Each ecological map or model is a unique representation of a possible distribution therefore, collating maps produced using systematic surveys and those generated from citizen science initiatives offers new ways to assess confidence in current knowledge and provide spatial data validation. Our research provides new certainty in model estimates of marine mammal distributions and densities in British Columbia by comparing geographic map products generated through disparate collection and modeling methodology. A key component of citizen science initiatives is to validate collected data. Data validation procedures are recommended but have yet to become standard in all marine citizen science programs (Silvertown, 2009; Thiel et al., 2014). Our approach for collating geographic data products has contributed to available techniques for amalgamating maps from different data collection procedures and provides quantitative methodology for evaluating information content of each spatial surface. Our research provides an applied example of the successful integration of citizen science data into marine-focused scientific research. The results of our analysis ultimately strengthens the overall perspective that crowd sourcing of ecological data has tangible benefits for advancing understanding of marine biogeography and spatial ecology.

Our research also applies quantitative methodology from spatial statistics to objectively identify regions of conservation prioritization through hotspot analysis. Our approach demonstrates the advantage of spatial analysis techniques for quantifying spatial patterns of cetacean and pinniped density that, prior to our research, had yet to be applied in a marine mammal context. Currently, common procedures for determining important regions for marine mammals are based on the use of arbitrary thresholds and opinions from expert consultation. By providing statistical inference methods that

account for spatial patterns, we contribute spatially explicit methods for marine decision-making processes to enhance objectivity and statistical rigor. Further, the use of local measures of spatial autocorrelation to identify hotspots can be generalized to locations and contexts outside of our study area.

### 4.3 Research opportunities

Reflection on the analysis and results of our study reveals exiting areas for future research opportunities. Our research helps fill existing knowledge gaps along the north coast of BC, but locations outside our study area are potential areas for expanded investigations. We see the application of our predictive modeling and hotspot analysis approach used throughout coastal BC for a variety of marine and terrestrial species. In addition, extending the geographic scope of analysis across jurisdictional borders would integrate marine mammal species density maps from adjacent countries to provide seamless cross-boundary distributions.

Quantity and quality of available data is a primary factor for developing spatial datasets that are comprehensive and representative. Another opportunity revolves around refining species-habitat models for marine mammals as finer resolution remotely sensed imagery and additional species surveys become available. Marine mammal distributions and aggregations are tied to the temporal shifting and movement of environmental conditions. Therefore, if sufficient data become available, it would be interesting to produce seasonal estimates of marine mammal distributions within our study region. Additionally, recommendations in Chapter 2 highlight the need for further sampling in locations where discrepancies between citizen science and systematic surveys occur. It would be informative to conduct supplemental surveys and determine whether confidence

surrounding cetacean density patterns could increase as a result of additional observational data.

We also see potential in incorporating citizen science datasets into ecological research and developing methodology to combat the limitations of opportunistically collected datasets. Although we have applied map differencing techniques and spatial hotspot analysis on citizen science collected information of marine mammals, there remain opportunities to apply these methods to other species in separate geographic areas.

## References

Becker, E. A., K. A. Forney, M. C. Ferguson, D. G. Foley, R. C. Smith, J. Barlow, and J. V. Redfern. 2010. Comparing California Current cetacean-habitat models developed using *in situ* and remotely sensed sea surface temperature data. Marine Ecology Progress Series 413:163—183.

Best, B. D., C. H. Fox, R. Williams, P. N. Halpin, and P. C. Paquet. 2015. Updated marine mammal distribution and abundance estimates in British Columbia. Journal of Cetacean Resource Management 15:9–26.

Canadian Government. 2016. Growing the Middle Class. Budget tabled in the House of Commons by the Honourable W. F. Morneau, Minister of Finance. Acdcessed 02 Nov 2016 from http://www.budget.gc.ca/2016/docs/plan/toc-tdm-en.html

Clarke, C. L., and G. S. Jamieson. 2006. Identification of Ecologically and Biologically Significant Areas in the Pacific North Coast Integrated Management Area: Phase II – Final Report. Can. Tech. Rep. Fish. Aquat. Sci. 2686: v + 25 pp.

Dalla Rosa, L., J. K. B. Ford, and A. W. Trites. 2012. Distribution and relative abundance of humpback whales in relation to environmental variables in coastal British Columbia and adjacent waters. Continental Shelf Research 36:89—104.

Elith, J., and J. Leathwick. 2009. The contribution of species distribution modelling to conservation prioritization. Chap 6. Pages 70—93 in A. Moilanen, K. A. Wilson and H. P. Possingham, editors. Spatial Conservation and Prioritization: Quantitative Methods and Computational Tools. Oxford University Press, New York.

Embling, C. B., P. A. Gillibrand, J. Gordon, J. Shrimpton, P. T. Stevick, and P. S. Hammond. 2010. Using habitat models to identify suitable sites for marine protected areas for harbour porpoises (*Phocoena phocoena*). Biological Conservation 143:267—279.

Foley, M. M., B. S. Halpern, F. Micheli, M. H. Armsby, M. R. Caldwell, C. M. Crain, E. Prahler, N. Rohr, D. Sivas, M. W. Beck, M. H. Carr, L. B. Crowder, J. Emmett Duffy, S. D. Hacker, K. L. McLeod, S. R. Palumbi, C. H. Peterson, H. M. Regan, M. H. Ruckelshaus, P. A. Sandifer, and R. S. Steneck. 2010. Guiding ecological principles for marine spatial planning. Marine Policy 34:955—966.

Fox, C. H., Huettmann, F., Harvey, G. K., Morgan, K., Robinson, J., and Paquet, P.C. (In Review). Predictions from Machine Learning ensembles: marine bird distribution and density on Canada's Pacific coast**.** Marine Ecology Progress Series.

Jamieson, G. S., and C. Levesque. 2014. Identification of Ecologically and Biologically Significant Areas on the West Coast of Vancouver Island and the Strait of Georgia, and in some nearshore areas on the North Coast: Phase II – Designation of EBSAs. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/101. Vii + 36 pp.

Kovacs, K. M., A. Aguilar, D. Aurioles, V. Burkanov, C. Campagna, N. Gales, T. Gelatt, S. D. Goldsworthy, S. J. Goodman, G. J. G. Hofmeyr, T. Härkönen, L. Lowry, C. Lydersen, J. Schipper, T. Sipilä, C. Southwell, S. Stuart, D. Thompson, and F. Trillmich. 2011. Global threats to pinnipeds. Marine Mammal Science 28:414—436.

Mannocci, L., M. Catalogna, G. Dorémus, S. Laran, P. Lehodey, W. Massart, P. Monestiez, O. Van Canneyt, P. Watremez, and V. Ridoux. 2014. Predicting cetacean and seabird habitats across a productivity gradient in the South Pacific gyre. Progress in Oceanography 120:383–398.

Mutke, J. 2011. Biodiversity gradients. Chap 9. Pages 170—190 *in* A. Millington, M. Blumler and U. Schickhoff, editors. The SAGE Handbook of Biogeography. SAGE Publications Inc., Thousand Oaks, California.

Parsons, E. C. M., S. Baulch, T. Bechshoft, G. Bellazzi, P. Bouchet, A. M. Cosentino, C. A. J. Godard-Codding, F. Gulland, M. Hoffmann-Kuhnt, E. Hoyt, S. Livermore, C. D. MacLeod, E. Matrai, L. Munger, M. Ochiai, A. Peyman, A. Recalde-Salas, R. Regnery, L. Rojas-Bracho, C. P. Salgado-Kent, E. Slooten, J. Y. Wang, S. C. Wilson, A. J. Wright, S. Young, E. Zwamborn, W. J. Sutherland. 2015. Key research questions of global importance for cetacean conservation. Endangered Species Research 27:113—118.

Redfern, J. V., M. C. Ferguson, E. A. Becker, K. D. Hyrenbach, C. Good, J. Barlow, K. Kaschner, M. F. Baumgartner, K. A. Forney, L. T. Ballance, P. Fauchald, P. Halpin, T. Hamazaki, A. J. Pershing, S. S. Qian, A. Read, S. B. Reilly, L. Torres,

and F. Werner. 2006. Techniques for cetacean-habitat modeling. Marine Ecology Progress Series 310:271—295.

Wright, A. J., and L. A. Kyhn. 2015. Practical management of cumulative anthropogenic impacts with working marine examples. Conservation Biology 29:333—340.