

Modelling risk of chronic oil pollution from vessel operations
in Canada's West Coast

By

Norma Sara Serra-Sogas
B.Sc., Fairleigh Dickinson University, 2001

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

MASTER OF SCIENCES

in the Department of Geography

© Norma Sara Serra-Sogas, 2010
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Supervisory Committee

Modelling risk of chronic oil pollution from vessel operations
in Canada's West Coast

by

Norma Sara Serra-Sogas
B.Sc., Fairleigh Dickinson University, 2001

Supervisory Committee

Dr. Rosaline R. Canessa, Department of Geography
Supervisor

Dr. Trisalyn A. Nelson, Department of Geography
Departmental Member

Dr. Patrick D. O'Hara, Department of Biology, Canadian Wildlife Services, Environment
Canada
Outside Member

Abstract

Supervisory Committee

Dr. Rosaline R. Canessa, Department of Geography
Supervisor

Dr. Trisalyn A. Nelson, Department of Geography
Departmental Member

Dr. Patrick D. O'Hara, Department of Biology, Canadian Wildlife Services, Environment
Canada
Outside Member

Chronic oil pollution or frequent small-scale oil discharges from vessel operations is an important source of marine oil pollution and considered a constant threat to marine and coastal environments. In Canada's Pacific region, evidence of such illegal discharges has been gathered by the National Aerial Surveillance Program (NASP) from 1998 to 2007. We used this information to fit Generalized Linear Models (GLMs) for offshore waters and inshore waters and explore the relationships between oil spill occurrences and four predictor variables: shipping traffic densities for different vessel types, distance to shore, distance to port and density of small harbours. The best-fit models for both regions show that areas closer to shore and with higher density of small harbours have a higher probability of oil spill occurrences. However, only in inshore waters was shipping traffic significantly related with oil spill occurrences. Tug boats and oil tanker traffic show a significant positive relationship with oil spill observations, while carriers presented a negative association. Mapped results for offshore areas depicted the highest probability of oily discharges in Barkley Sound and at the entrance of the Juan de Fuca Strait; whereas in inshore waters, oil pollution hot spots were found in the vicinity of major commercial and tourist centres.

These probability maps were used to identify Coastal and Marine Protected Areas (CMPAs) and Important Bird Areas (IBAs) more likely to be exposed to chronic oil pollution during a period of 10 years. Three areas were highlighted as the most vulnerable

based on their likelihood of exposure and the sensitivity of the species they contain to oil pollution. These sites are the Tofino Mudflats, Barkley Sound, Scott Islands and the Sturgeon and Robert Banks, in the Fraser River delta.

Our findings provide better understanding of the relationships between oil spill occurrences and vessel operations and help us identify likely oil pollution hot spots and sites particularly vulnerable to this stressor in Canada's Pacific region. This information can be useful to NASP in improving its efficiency and in targeting monitoring efforts to troublesome areas. Additionally, this research contributes to regional studies that focus on analyzing the distribution of anthropogenic stressors from sea-based activities in British Columbia. Finally, we highlight the importance of collecting accurate data to properly model the probability of oil spill occurrences and encourage future research aiming to better understand and ultimately reduce the chronic release of pollutants from shipping activities into the marine environment.

Table of Contents

Supervisory Committee	ii
Abstract	iii
Table of Contents	v
List of Tables	viii
List of Figures	ix
Acknowledgments	xii
Chapter 1 : INTRODUCTION	1
1.1 Vessel-source oil pollution: background information	1
1.2 Marine pollution monitoring and regulations in Canada	6
1.3 Research objectives, questions and contributions.....	9
1.4 Methodological approach	11
1.4.1 Modeling chronic oil pollution in Canada’s West Coast.....	11
1.4.2 Identifying vulnerable areas to chronic oil pollution in Canada’s West Coast ..	13
1.5 Thesis structure	14
Chapter 2 : MODELLING DETECTED ILLICIT OIL SPILLS FROM VESSEL OPERATIONS IN CANADA’S WEST COAST	15
2.1 Introduction.....	15
2.2 Data and methods	20
2.2.1 Study area and spatial scale selection	20
2.2.3 Data description and preparation	22
2.2.5 Exploratory data analysis.....	33
2.2.5 Model assumptions	34

2.2.6 Candidate model distributions	35
2.2.7 Surveillance effort as an offset variable	39
2.2.8 Model calibration and model selection.....	40
2.2.7 Assessing goodness-of-fit.....	42
2.2.8 Residual analysis	43
2.2.9 Model predictions	44
2.3 Results.....	44
2.3.1 Exploratory data analysis.....	44
2.3.2 Model calibration.....	47
2.3.3 Assessing goodness-of-fit.....	54
2.3.4 Regression diagnostics.....	54
2.3.5 Model predictions	58
2.4 Discussion.....	63
2.4.1 Model interpretation	63
2.4.1 Model biases and limitations	66
2.4.2 Future applications and recommendations	69
2.5 Conclusion	71
Chapter 3 : IDENTIFYING VULNERABLE CMPAs AND IBAs TO CHRONIC OIL POLLUTION IN THE WEST COAST OF CANADA.....	72
3.1 Introduction.....	72
3.2 Data and methods	74
3.2.1 Study area	74
3.2.2 Data.....	75

3.2.3 Map overlay procedure	84
3.3 Results.....	87
3.4 Discussion.....	94
3.4.1 Results discussion.....	94
3.4.2 Advantages, limitations and research contributions	98
3.4.3 Future research.....	99
Chapter 4 CONCLUSIONS.....	101
4.1 Introduction.....	101
4.2 Summary of findings	102
4.3 Research contributions.....	105
4.4 Research Opportunities.....	108
Bibliography	111
Appendix A: NASP effort extraction	123
Appendix B: Vessel Traffic Services in Canada’s Pacific Region.....	124
Appendix C: Scatterplots of model variables and variable selection results.....	126

List of Tables

Chapter 2

Table 2.1 Total study area and total area surveyed by NASP flights from 1998 to 2007 for each region.....	22
Table 2.2 Summary statistics of response and predictors for models in area A.....	45
Table 2.3 Summary statistics of response and predictors for area B.....	45
Table 2.4 Summary of the best-fit models for presence/absence data and oil spill counts in Area A (offshore).....	48
Table 2.5 Summary of the best-fit models for presence/absence data and oil spill counts in Area B (inshore).....	51

Chapter 3

Table 3.1 Description of federal and provincial CMPA designations included in this study and classified based on their protection and conservation goals.	78
Table 3.2 Number and statistics of marine area (in km ²) of CMPAs and IBA for the entire study area.	81
Table 3.3 Mean probability of oil spill exposure (in percentage and over 10 years) for each CMPA designation by IBAs by marine ecosection in which they occur.....	88
Table 3.4 Mean probabilities of oil spill exposure (in percentage and over 10 years) per designation for the Tofino Mudflats, the Barkley Sound region, the Sturgeon and Robert Banks and the Scott Islands.	93

List of Figures

Chapter 1

Figure 1.1. Trends of total number of NASP patrol hours and observed oil spill look-alikes, and ratio number of spills by number of patrol hours, from fiscal years 1993/1994 to 2008/2009, in Canada's Pacific region.	8
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---

Chapter 2

Figure 2.1. Extent of the study Area A (offshore) and Area B (inshore).	21
Figure 2.2 Distribution of NASP surveillance effort (in square kilometres) from 1998 to 2007 in Areas A and B.	24
Figure 2.3 Distribution of oil spill counts observed by NASP from 1998 to 2007	25
Figure 2.4 Intensity of vessel movements (number of movements per cell) averaged per year, for different vessel types, and pooled from data recorded in 2003, 2005, 2006, 2007 and 2008 in Canada's West Coast.	29
Figure 2.5 Percentage of mean yearly vessel movements per vessel type in Area A (left) and Area B (right)	30
Figure 2.6 Density of small harbours (i.e., marinas, yacht clubs, public docks, fishing harbours and ferry terminals) and major commercial ports considered in this study.	33
Figure 2.7 Frequency distribution of oil spill counts in Areas A (left) and B (right).	46
Figure 2.8 Observed (dashed line) and predicted (solid-red line) frequencies of positive oil spill counts in area A.	49
Figure 2.9 Observed and predicted frequencies of zero oil spill counts in Area A.	50
Figure 2.10 Observed (dashed line) and predicted (solid-red line) frequencies of positive oil spill counts in Area B.	53
Figure 2.11 Observed and predicted frequencies of zero oil spill counts in Area B.	53
Figure 2.12 Standardized residuals against predicted values (left). Plotted horizontal dashed lines represent the 95 % confidence intervals. Map (right) highlighting the identified extreme significant outliers and over-predicted value of the A.P model.	55

Figure 2.13 Standardized residuals against predicted values (left). Plotted horizontal dashed lines represent the 95% confidence intervals. Map (right) highlighting the identified extreme significant outliers and over-predicted values of the B.NB model.....	56
Figure 2.14 Residual clusters map for Area A, based on the estimation of local Moran's I_i and applying a significance filter of $p \leq 0.05$	57
Figure 2.15 Residual clusters map for Area B, based on the estimation of local Moran's I_i and applying a significance filter of $p \leq 0.05$	58
Figure 2.16 Predicted number of oil spill in Area A (offshore) estimated from model A.P and based on data collected by NASP from 1998 to 2007.	59
Figure 2.17 Probability of oil spill occurrences in Area A (offshore) estimated from model A.BI and based on data collected by NASP from 1998 to 2007.	60
Figure 2.18 Predicted number of oil spill in Area B (inshore) estimated from model B.NB and based on data collected by NASP from 1998 to 2007.	61
Figure 2.19 Probability of oil spill occurrences in Area B (inshore) estimated from model B.BI and based on data collected by NASP from 1998 to 2007.....	62

Chapter 3

Figure 3.1 Study area with offshore (Area A) and inshore (Area B) regions.....	75
Figure 3.2 Distribution of CMPAs aimed for wilderness protection (top left), CMPAs designated for seascape protection and recreational use (top right), IBAs (bottom left), and areas with overlapped designations (bottom right), in the study area.	80
Figure 3.3 Marine ecosections in Canada's West Coast.....	82
Figure 3.4 Distribution of total area designated as IBA and protected under federal and provincial CMPAs by marine ecosections in Areas A and B.	83
Figure 3.5 Distribution of oil spill probability of occurrence (in percentage) based on data collected by NASP from 1998 to 2007 in the study area (Areas A and B).	84
Figure 3.6 Zonal statistics procedure (figure modified from ArcGIS v.9.3 help menu).	86
Figure 3.7 Box-plot of mean oil spill probability values extracted from each CMPA and IBA.	87

Figure 3.8 Federal CMPAs overlaid with the probability of oil spill occurrence over a period of 10 years in Canada’s West Coast. 89

Figure 3.9 Provincial CMPAs overlaid with the probability of oil spill occurrence over a period of 10 years in Canada’s West Coast. 91

Figure 3.10 IBAs with the probability of oil spill occurrence over a period of 10 years in Canada’s West Coast. 92

Acknowledgments

First and foremost I would like to thank my supervisor Dr. Rosaline Canessa and Dr. Patrick O'Hara (my unofficial but true co-supervisor). Without your constant encouragement and support this thesis would not have been possible. I also would like to acknowledge Dr. Trisalyn Nelson for providing valuable insight and suggestions throughout this research. Finally, I would like to thank Dr. Carl Brown, my external examiner, for revising my thesis and for providing very interesting and useful comments.

I also would like to thank the NASP crew based in Richmond (BC): John Heiler, Sue Baumeler, Mac McGregor, Sev Leblanc, and Kim Pearce for providing original data and other very important information for my research.

I would like to acknowledge my gratitude for the financial support I received from GEOIDE Network Project “Coastal Security and Risk Management Using GIS and Spatial Analysis” and its members Dr. Peter Keller (University of Victoria), Dr. Ron Pelot (Dalhousie University), Stefania Bertazon and Marina Gavrilova (University of Calgary). I also would like to forward my gratitude to Canadian Wildlife Services (Environment Canada) for providing financial support during the last year of my research, and in particular to Kevin Fort and Patrick O’Hara for your help through the application process.

Special thanks to my colleagues in the GIS Lab and the staff from the Geography Department for their helpful advice and assistance. Last, but not least, I would like to thank my friends, my parents Maria José y Pedro and my husband Roth for their unconditional support and constant encouragement over the past three years.

Chapter 1 : INTRODUCTION

1.1 Vessel-source oil pollution: background information

Oceans are traversed by thousands of vessels every day, from large commercial vessels to small pleasure crafts. Marine transportation is the most effective way to transport goods across the world's oceans. Today, approximately 90% of the world trade is conducted by the international shipping industry (International Chamber of Shipping, 2009). At the same time, the presence of recreational crafts in coastal areas has expanded as a result of their cost reduction and the increase in disposable incomes (National Marine Manufacturers Association, 2009; Widmer & Underwood, 2004). Cruise ships and passenger ferries are also important in coastal communities, particularly in tourist driven economies and areas dominated by small islands such is the north-east Pacific. These sea-based activities are, however, impacting marine and coastal environments by introducing different types of pollutants, principally in the form of hydrocarbons and derivatives (Halpern et al., 2008; Shahidul Islam & Tanaka, 2004).

Vessel-source oil pollution can be divided into two categories: accidental oil spills and intentional (operational) discharges. Accidental oil spills are defined as discrete events where oil is discharged by negligence from the vessel over a relatively short time (Etkin, 2001). Often, the most spectacular accidental oil spills arise from oil tankers (ITOPF, 2009). In contrast, operational oil discharges have traditionally received less media attention, if compared with the formerly described, because of the small amounts of oil released per event –usually less than 1000 l. Generally, operational discharges are described as intentional releases of waste oil generated during routine vessel operations

(National Academy of Sciences, 2003). In the literature, the terms “intentional” and “operational” are often used interchangeably (Mooradian, 2002). Arguably, however, operational oil discharges may also be caused by negligence of the vessel's crew, and hence becoming an accidental event, and not deliberately planned in advance. Pollution from operational oil discharges are also known as “chronic oil pollution” because of the frequency at which they take place.

The amount and nature of oil discharged during vessel operations often depend upon the type of ship and operation performed. Traditionally, large ocean-going tankers have been seen as the main vessel type responsible for operational oil discharges, clandestinely discharging dirty ballast and bilge water as well as waste water from tank washing (GESAMP, 2007). Yet, all motorized vessel types –be they recreational vessels, fishing boats, cruise ships, etc...– may contribute to chronic oil pollution (Khee-Jin Tan, 2005), either from accidental or intentional oil discharges from engine rooms, leaks from propellers and on-deck hydraulics, or at reception facilities (i.e., ports and marinas) during oil transfer operations and vessel maintenances and runoff storm water from the area (Georgia Strait Alliance, 2009; EPA, 2001; ITOPF, 2002).

Global estimations show that the amount of oil released from routine vessel operations is relatively higher than oil released from accidental spills. In 2007, GESAMP estimated that 45% of the annual amount of oil (~ 451,000 tonnes/year) entering into the marine environment come from vessel-source operational discharges, while accidental discharges only account for 36% (~186,000 tonnes/year) (GESAMP, 2007). Further, the environmental impact of oil spills from regular vessel operations is believed to be greater

than oil spills from catastrophic accidental events, as a result of the constant releases of waste oil and their cumulative effects (Khee-Jin Tan, 2005; Wiese & Ryan, 2003).

In an attempt to regulate vessel-source pollution, the international community developed a legislative framework of standards and regulations. The principal set of international regulations on marine pollution was developed at the 1973 International Convention for the Prevention of Pollution from Ships and later at its 1978 Protocol of Amendment (i.e. MARPOL 73/78), which was adopted under the auspices of the International Maritime Organization (IMO), and currently ratified by 150 states, including Canada¹ (International Marine Organization, 2009). Particularly, Annex I of MARPOL 73/78 provides regulations for the prevention of pollution by oil from ships. These regulations state that oil discharges are allowed as long as: they do not exceed 1/15,000 (or 1/30,000 for newer tankers) of the total cargo carrying capacity of the vessel; they are discharged at a rate not exceeding 60 litres per mile; and they are released beyond 50 nautical miles (~ 80.5 km) from the coast and outside designated “Special Areas” (IMO, 2007). Special Areas are defined under MARPOL as those regions that for “their ecological and oceanographical conditions and to their sea traffic [need] special mandatory methods for the prevention of sea pollution” (IMO, n.p.). Oil discharges in Special Areas and within 50 nm from the coast are completely prohibited, unless the oil content in the discharge is less than 15 ppm. This concentration is a key parameter because, as it is recognized by IMO, it not possible to observe, either visually or through remote sensing equipment, oily discharges at sea with an oil content less than 15 ppm (Ferraro et al.,

¹ Canada ratified MARPOL 73/78 Convention on February 16, 1993.

2009), unless advanced sensors such as laser fluorosensors are used (C. Brown, EC, pers. comm. 2010)

These regulations were dedicated mainly to control large commercial shipping vessels, and specifically oil tankers. However, at present, all vessel types must conform with the MARPOL 73/78 Convention, independently of their size and function, with the exception of warships and government-own vessels (Khee-Jin Tan, 2005). In addition, MARPOL 73/78 makes a distinction between accidental and intentional, the former being not an infringement of the convention, providing that all the precautions have been taken to avoid the discharge (MARPOL, Annex I, Regulation 11(b)).

The United Nations Convention on the Law of the Sea (UNCLOS), ratified by Canada in 2003, extended the coastal states' authority "to adopt laws and regulations [such as MARPOL 73/78] for the prevention, reduction and control of marine pollution from vessels" (UNCLOS, Art 211(5)) up to the 200 nautical miles limit known as the Exclusive Economic Zone (EEZ) (Khee-Jin Tan, 2005). UNCLOS also enhanced the port-states' jurisdiction by allowing the inspection of visiting foreign vessels and, if substantial evidence found, the prosecution for discharging pollutants in violation of international regulations (i.e., MARPOL) "anywhere at sea" (Khee-Jin Tan, 2005).

Despite the wide array of international and domestic regulations, operational discharges by marine vessels continue to be an unresolved problem, primarily because of the challenges surrounding compliance, surveillance and enforcement efforts. Vessels sailing under flags of convenience, often not party to MARPOL 73/78, are often the most problematic when ensuring compliance (Khee-Jin Tan, 2005). However, even if the vessel's flag-state is party to MARPOL 73/78, the vessel operator may still opt not to

comply, either because of lack of adequate reception facilities in ports to dispose of their residues or, if these facilities are available, the transfer of residual oil is costly and time consuming. Finally, prosecutors face the challenge of detecting illicit polluters *in flagrante delicto*. Without clear evidence of oil pollution (e.g., oil discharged is attached to the wake of the vessel) it is nearly impossible to successfully prosecute offenders (GESAMP, 2007; Kluser, Richard, & Giuliani, 2006; National Academy of Sciences, 2003). Although, if samples of the oil spill are available chemical fingerprinting can be used to identify the source and match it to the oil found in vessels tanks and bilge waters (Stout & Wang, 2007).

The implementation of effective monitoring programs to identify and deter illicit polluters is necessary to successfully enforce marine pollution regulations, such as MARPOL (Kasoulides, 1988). Nations may use one or more of the three established methods to fight illicit pollution from marine vessels: onboard or port inspections, satellite monitoring and aerial monitoring. Tar ball² monitoring (Golik & Rosenberg, 1987; Wang et al., 1998) and Beached Bird Surveys (BBS) are considered alternative and cost-effective monitoring tools (Camphuysen, 1998). All these methods play a very important role as indicators of the degree of chronic oil pollution in coastal and marine environments and they are useful tools to gauge the impact and performance of international and national pollution policies in the states where they are implemented (Ohowa, 2009).

² Tar balls are weathered oil accumulations which are primary found on beaches and along the coastline (GESAMP, 2007)

1.2 Marine pollution monitoring and regulations in Canada.

With the world's longest coastline, Canada's sea-based activities constitute key elements of its growing economy. Particularly, maritime commercial shipping, a critical component of Canada's international trade, has seen an overall increase of approximately 13% (in tonnage of cargo handled) since the beginning of the 21st century (Statistics Canada, 2009). The recreational boating industry has also generated significant contributions to Canada's domestic economy (Sorum, Jan 25, 2008), and ferry systems, particularly on the west coast of Canada, form a vital transportation network connecting communities along the coast. However, oil pollution from vessel activities continues to be one of the main threats to Canadian marine and coastal environment. Wiese and Ryan (2003) estimated that a minimum of 300,000 seabirds die every year in Atlantic Canada from vessel-source chronic oil pollution, based on data collected during BBSs.

The leading federal agency responsible for enforcing national and international marine pollution regulations is Transport Canada, assisted by Environment Canada and the Department of Fisheries and Oceans. The main pieces of legislation that address pollution regulations in Canada are the *Canada Shipping Act, 2001*, the *Canadian Environmental Protection Act, 1999*, the *Migratory Birds Convention Act, 1994*, and the *Species at Risk Act, 2003* (Armstrong & Derouin, 2004).

Canada has adopted three approaches to fight chronic oil pollution. These are the National Aerial Surveillance Program (NASP), the Integrated Satellite Tracking of Polluters (I-STOP) and in-port inspection programs. NASP is one of the primary tools available to Transport Canada for detecting and preventing any illegal discharges. Four aircraft stationed throughout the country routinely patrol the skies monitoring for oil spill

occurrences at sea (Armstrong & Derouin, 2004; McGregor & Gautier, 2008). Currently, NASP is operated and managed under the umbrella of the Marine Aerial Reconnaissance Team (MART), a partnership between Transport Canada and Environment Canada.

In the Pacific region, three different surveillance aircraft have been patrolling the west coast of Canada, since the beginning of the NASP program in the early 1990s. The first aircraft, a DC-3/C-FDOT, operated until 1995 when it was replaced by the Twin Otter 300. Currently, and since the beginning of 2008, a Dash 8 aircraft is being used to monitor Canada's Pacific waters. Records on surveillance effort and observation rates show the development of NASP since its early years operating in this region. NASP experienced a decline in surveillance efforts since the mid 1990's and until 2000 (Figure 1.1). The overall increase in surveillance activities since the turn of the 21st century and until now is likely attributed to the introduction of national pollution prevention legislation (e.g., amendments included in *Canada Shipping Act, 2001*) which motivated the allocation of more resources to fight marine pollution. However, despite the noticeable increase in surveillance effort since 2000, the number of observed oil spill look-alikes³ per fiscal year has remained relatively constant at approximately 50 spills, except in 2008/2009 when more spills were detected (~ 50 more than previous years). This increase in observation rates was likely due to the introduction of the Dash 8 aircraft, armed with a suite of modern remote sensors that are capable of monitoring during day and night at a wider viewing swath and areas further offshore (McGregor & Gautier, 2008).

³ Look-alikes is the term used to define phenomena that dampen the short waves and create dark patches on the water surface (Brekke & Solberg, 2005)

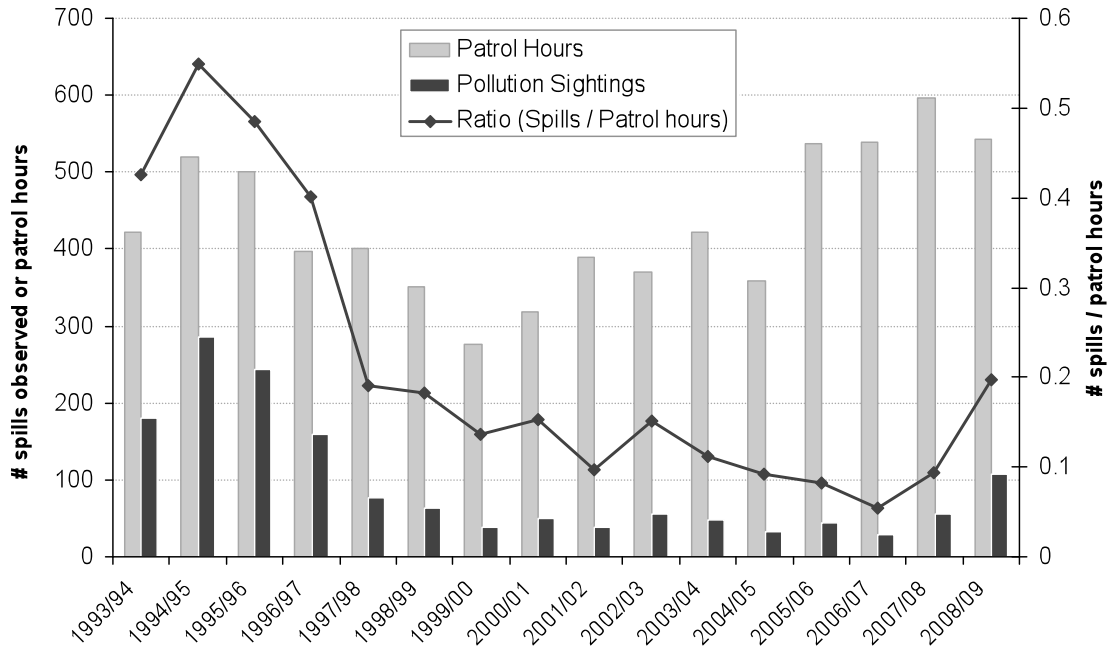


Figure 1.1. Trends of total number of NASP patrol hours and observed oil spill look-alikes, and ratio number of spills by number of patrol hours, from fiscal years 1993/1994 to 2008/2009, in Canada's Pacific region.

One of NASP's greatest challenges resides in maximizing the effectiveness of surveying an enormous area, such as Canada's Pacific region. Detection of oil spill on the surface of the water is a complex procedure, given that the sources of the discharges tend to be mobile, the available observation window varies widely depending on the weather conditions and sea state, and because vessel operators act clandestinely when intentionally discharging waste oils. Additionally, many of the observed oil spills are mystery spills, meaning that no source can be pinpointed at the time of the detection, or at least from the aircraft and without help of sophisticated forensic techniques. In order to shed light onto some of the questions surrounding this issue (e.g., why and where oil spills are more likely

to occur) we propose the generation of a statistical model, describing oil spill occurrences based on the locations of potential sources.

The data collected by the Twin Otter is analyzed in this study because it offers the largest available temporal coverage of suspected oil spills in Canada's Pacific region. This aircraft was equipped with a Global Positioning System (GPS) and conventional photographic cameras (Armstrong & Derouin, 2004) which, compared with the new Dash 8, had a limited operational window both temporal (only daylight surveillance) and spatial (maximum flight range of five hours).

1.3 Research objectives, questions and contributions

Given the challenges associated with monitoring and observation of illegal oil spills in Canada's Pacific region, this study aims to achieve a greater understanding of chronic oil pollution by describing its distributions and potential impact on valuable and vulnerable marine environments. Specifically, this research addresses the following objectives and research questions:

(1) To identify factors that better explain the spatial variability of suspected oil spills by NASP along Canada's west coast through the characterization of a statistical model using multivariate regression analysis. The best-fit model will be used to predict the probability of oil spill occurrence and relative abundance (number of oil spills per unit area) in areas not observed by NASP. The research questions to be answered are:

(a) Which sea-based activities help to describe the distribution of detected oil spills?

(b) Which areas show the highest probability of illicit oil spill occurrences and relative abundance in Canada's Pacific waters?

(2) To identify marine areas particularly vulnerable to chronic oil pollution based on the likelihood that Coastal and Marine Protected Areas (CMPAs) and Important Bird Areas (IBAs) are exposed to this threat. This goal aims to address the following question:

(a) Which CMPAs and IBAs are most vulnerable to chronic oil pollution?

These findings will contribute to the better understanding of the relationships between oil spill occurrences and marine vessel operations, and will help identify likely oil pollution hot spots and sites particularly vulnerable to this stressor in Canada's Pacific region. One of the main outcomes of this study is that it provides a reference for future studies on chronic oil pollution in the region. In addition, the results can assist NASP in improving its monitoring strategy by targeting identified troublesome areas. This research also contributes to similar regional studies that focus on determining the distribution of anthropogenic stressors from sea-based activities in British Columbia. Finally, we show the importance and usefulness of modelling oil spill observations collected by NASP and the relevance of available accurate data to properly model the probability of oil spill occurrences to better understand and ultimately reduce the chronic release of pollutants from marine vessel activities into the marine environment.

1.4 Methodological approach

1.4.1 Modeling chronic oil pollution in Canada's West Coast

One of the first steps in selecting an appropriate statistical regression model is to explore the frequency distribution of the response variable (Guisan & Zimmermann, 2000). When the response variable is derived from count data (i.e., number of oil spills per unit of observation) or binomial distributions (presence/absence) generalized linear models (GLMs) are used. These approaches are characterized for handling distributions that differ from the normal Gaussian distribution, such as Poisson and the negative binomial (McCullagh & Nelder, 1989).

Oil spill counts in the west coast of Canada contain a large number of zero counts or more than what would be expected under a Poisson distribution. The distributions are known as “zero-inflated” distributions (Lambert, 1992). The excess of zeros in oil spill counts is attributed principally to the presence of “true zeros” generated by the fact that illicit oil spills are difficult to detect because they are rare and spatially aggregated. Zero inflation can also be the result of an excess of “false zeros” caused by errors in the sampling methods or from the observer(s) (Martin et al., 2005). However, we assume perfect detection in our analysis, thereby eliminating the probability of encountering false zeros. The reasoning and implications behind this assumption are discussed in Chapter 2. Hurdle and zero-inflated models are two suggested approaches to model the excess of zeros in oil spill count data. The selection of the model will be informed by the estimation of information criteria. The model with the best fit will then be used to predict oil pollution occurrences in areas not visited by NASP.

Oil spill counts may exhibit another complexity that needs to be tested during the modelling process, and that is the presence of spatial dependence or spatial autocorrelation. The presence of positive spatial autocorrelation indicates that sampling units closer together are more related to each other than sampling units farther apart (Tobler, 1970). Detected oil spills closer to each other are more likely to be related to the same process(es) than oil spills farther apart. For example, oil spills clustered near marinas are more likely derived from recreational boating, while oil spills detected within marine Traffic Separation Schemes (TSS) are more likely to be associated to large commercial vessels. These interactions are defined by Fortin and Dale (2005) as "induced" spatial dependency because the underlying processes likely responsible of the observed clustered pattern are themselves spatially dependent.

Spatial autocorrelation is considered a "nuisance" when standard regression methods are applied to spatial data (Griffith, 1992). For instance, the presence of autocorrelation in dependent and/or independent variables may translate into the presence of autocorrelation in the residuals, thereby violating the assumption of error independency (Haining, 2003). In addition, spatial autocorrelation may also inflate the number of degrees of freedom leading to false significance of the regression model (Legendre, 1993). The presence of spatial autocorrelation in the residuals and other regression diagnostics, such as the examination of residual plots, will be used to assess problems that may compromise the regression analysis because the model assumptions are not met.

1.4.2 Identifying vulnerable areas to chronic oil pollution in Canada's West Coast

Canada's Pacific coastline extends over a vast and intricate coastal geography composed of islands, bays, straits, and fjords, and sustain a rich diversity of marine ecosystems (Wildlife Habitat Canada, 2001). Unfortunately, these prime marine and coastal ecosystems are threatened by a number of anthropogenic activities, including the chronic release of oil from vessel operations. The successful management and conservation of marine and coastal environments depend on recognizing and understanding the impact that human activities have on the marine environment, and this includes the spatial distribution and extent of these stressors (Ban & Alder, 2008).

The impact that oily discharges have on the marine environment varies depending on the type of oil, the amount and frequency of oil discharges, weather conditions that affect the oil weathering processes, and most importantly the location and the sensitivity (both biological and physical) of the area affected (IPIECA, 1991). One of the outcomes of modeling oil spill observations collected by NASP is the construction of a probability map of oil spill occurrences. By overlying this map with the location of sensitive features to oil pollution we can identify the areas most likely to be exposed (i.e., most vulnerable) to this stressor (Zacharias & Gregr, 2005).

British Columbia hosts a number of Coastal and Marine Protected Areas and Important Bird Areas (IBAs), which may contain one or more valuable features (i.e., species and habitats) particularly sensitive to oil pollution, such as seabirds. This research uses the location of CMPAs and IBAs to identify vulnerable marine areas to chronic oil pollution in British Columbia. The location of these vulnerable marine areas can be used as an informative tool to coastal and marine conservation planners to highlight those areas in

need of management action (Zacharias & Gregr, 2005), such as the management of sea-based activities within the boundaries of the protected area to avoid any detrimental effects that these activities may bring (Halpern et al., 2007).

1.5 Thesis structure

This thesis is divided into four chapters. Chapter 1 has provided background information on the issue of chronic oil pollution from sea-based activities, navigating from an international context to a local one. It also includes a brief description of the project objectives and contributions. Chapters 2 and 3 form the main core of this thesis and they have been prepared as stand-alone manuscripts to facilitate their submission to peer-reviewed publications. Specifically, Chapter 2 introduces the different statistical approaches to model oil spill abundances and occurrences in Canada's Pacific region. It also includes results using the best fitted model as well as a discussion of model limitations. Chapter 3 examines the use of the chronic oil pollution probability maps to identify vulnerable areas to oil pollution based on the distribution of CMPAs and IBAs. Finally, Chapter 4 provides a summary of the main findings from the previous two chapters, contributions and future research opportunities.

Chapter 2 : MODELLING DETECTED ILLICIT OIL SPILLS FROM VESSEL OPERATIONS IN CANADA'S WEST COAST

2.1 Introduction

Chronic oil pollution is defined as the intentional or accidental release of relatively small quantities of oil (< 1000 l) during routine vessel operations (GESAMP, 2007). Operational oil spills may include: leakages from engine rooms, propellers and on-deck hydraulics, or at reception facilities (i.e., ports and marinas) during oil transfer operations and vessel maintenances (Georgia Strait Alliance, 2009; ITOPF, 2002; Mooradian, 2002). The term "chronic" suggests that these oil spills occur quite frequently, especially in areas dominated by dense shipping traffic and at reception facilities. The estimated global annual amount of oil released into the marine environment from vessel operations (~ 451,000 tonnes/year) accounts for at least twice the volume of oil released from large accidental spills (~186,000 tonnes/year) (GESAMP, 2007). Additionally, the environmental impact of chronic oil pollution is believed to be greater than pollution from catastrophic accidental events (Wiese & Ryan, 2003), given the cumulative effects of constant release of waste oil versus the larger but sporadic accidental oil spills (Khee-Jin Tan, 2005). In fact, true volumes of chronic oil pollution are likely to exceed reported estimates (Chrastansky & Callies, 2009) due to the challenges associated with the observation and documentation of operational oil discharges.

In Canada's Pacific region, the National Aerial Surveillance Program (NASP) has documented on suspected marine oil spills since the early 1990's, becoming one of the main monitoring and enforcement tools for compliance of national and international

pollution regulations within Canadian waters (Armstrong & Derouin, 2004). Currently, NASP is managed by Transport Canada and operated in collaboration with Environment Canada's Canadian Ice Service.

The cost and time needed to patrol large geographical regions, such as Canada's Pacific region (approx. 457,646 km² delimited by Canada's EEZ), often constrain the effectiveness of aerial surveillance programs such as NASP. Until 2007, NASP operations were particularly challenging due to the limited operational window (both temporal and geographical) of the aircraft used, a Twin Otter 300. This aircraft had a maximum flight range of five hours (K. Pearce, EC, pers. comm., 2010), observations were only possible during daylight and up to an estimated 3.2 km on either side of the aircraft in optimal weather conditions and without the aid of remote sensors (Armstrong, 2004). Further, the observation and verification of a potential illicit oil spill from overflights were generally difficult because the aircraft was not equipped with specialized remote sensors (e.g., Side-Looking Airborne Radar).

The complexities of oil spill visual observation from an aircraft emerge mainly because: (a) the sources of oil pollution may be mobile or static; (b) oil can be masked by other materials (e.g., debris, ice) or it can be mistaken by other substances or conditions on the sea (e.g., weed beds, wind shadows) (Fingas & Brown, 2007) (c) the life span of an oil spill (i.e., observational window) varies depending on the volume and type of oil spilled and the environmental conditions at the time of the spill (Volckaert et al., 2000); and (c) vessel operators may act clandestinely when intentionally discharging waste oils to avoid being identified by pollution surveillance (Serra-Sogas et al., 2008a). Consequently, a

number of oil discharges often remain unaccounted for or without a clearly identified source.

In light of these challenges, it is imperative to better understand the distribution of observed oil spills in order to draw inferences about the underlying processes governing their likelihood of occurrence and abundance. We used statistical regression methods to describe the relationships between oil spill observations collected by NASP from 1998 to 2007 and a set of explanatory variables indicators of potential pollution sources, including shipping traffic densities, distance to shore, distance to port and density of small harbours. The best fit model can be used to predict the probability of occurrence and abundance of illicit oil spills beyond areas surveyed by NASP.

In regression analysis, generalized linear models (GLMs) (McCullagh & Nelder, 1989) are the most common approaches to model count data and predict expected abundances per unit of observation. By transforming counts into presence/absence data and by employing logistic regression (a particular case of GLMs) we can estimate the probability of oil spill occurrences (McCullagh & Nelder, 1989).

Oil spill estimates have an abundance of zero observations, which may be problematic when using classical GLMs. If the distribution of count data is "zero-inflated" (Lambert, 1992), the use of GLMs may not be appropriate since they could generate biased predictions and overly optimistic statistical inferences (Potts & Elith, 2006). Two approaches exist to model data with an excess of zeros: zero-inflated (ZI) or mixture models (Lambert, 1992) and hurdle (ZA) or two-part models (Heilbron, 1994; Mullahy, 1986). The choice between the two model specifications depends on the type of zeros present in the dataset: "false zeros" or "true zeros" (Martin et al., 2005). False zeros or

sampling zeros (Ridout, Demétrio, & Hinde, 1998) arise from errors either as a result of a poor sampling procedure or because of observer(s) inexperience. For instance, false zeros emerge when the time dedicated to sample an area or the geographic extent is not adequate to observe the event under study (Tyre et al., 2003). Conversely, true zeros, also known as structural zeros (Ridout et al., 1998) or fixed zeros (Royle & Nichols, 2003), are inevitable and are associated with underlying processes that affect the presence of the event at the sampled location. For example, in ecology, true zeros emerge when the species under study is not detected because the habitat sampled is not suitable for that particular species (Martin et al., 2005).

The best approach to model data containing an excess of true zeros resulting from a single underlying process is *ZA* models (Martin et al., 2005). *ZA* fits a model in two parts. The first part can take a Poisson or negative binomial distribution and represents the probability of observing a positive count but truncated at zero (i.e., zeros are omitted). The second part has a binomial distribution (i.e., presence/absence) and corresponds to the probability of observing true zero values (Cameron & Trivedi, 1998). In summary, *ZA* models estimate the probability to observe an event and, providing that is present, estimate its relative mean abundance.

If the excess of true zeros is thought to emerge from more than one process, then *ZI* mixture models are more appropriate than *ZA* models, as they can model more than one process ruling the observed data (Martin et al., 2005). These models are characterized by a "mixture" of two different distributions (or components). The first component can take a Poisson, negative binomial or binomial distribution and represents the probability of observing a positive event and a proportion of zeros expected under non-zero inflation.

The second component with a binomial distribution represents the probability of observing the excess of zero values (Lambert, 1992). ZI models are also required in presence of false zeros, since now we are interested in modelling two processes: the process that generates the true zeros and the process that generates the false zeros. This approach is used by ecologists concerned with modelling imperfect detection rates while estimating the probability of occurrence with a set of covariates (MacKenzie et al., 2002; Tyre et al., 2003). In general, ZI models are more difficult to fit and interpret than ZA models because in ZA the probability of observing zero counts is independent of the probability of observing a positive event, whereas in ZI models the parameters need to be interpreted together (Welsh et al., 1996; Martin et al., 2005).

In our dataset, oil spill counts are assumed to contain only excess of true zeros, which emerge as a result of observing an event that is rare due to the strict regulations and fines set in place for their prevention. Also, there are marine areas in the west coast of Canada that are unlikely or not “suitable” to experience an illicit oil spill because of absence of marine vessels or coastal facilities, principal sources of operational oil spilling. Oil spill data is also likely to include false zeros because of the challenges associated with observing oil spill on the surface of the water and from an aircraft, as we described previously. However, the lack of information on factors affecting detection rates (e.g., visibility, flight altitude, sea condition) make it impossible to build a model aimed to explain false-zero observations. Therefore, it is necessary to assume perfect detection rates (i.e., oil spills are always detected when present with at least one flight visit) in order to eliminate the probability of encountering false zeros. Only the uneven distribution of surveillance effort is accounted for by introducing surveillance effort as an offset variable

in the model. The assumption of perfect detection rates is one of the main limitations of this modelling exercise and the greatest source of uncertainty, since disregarding imperfect detection may lead to misleading conclusions of the relationships between abundance and predictor variables (MacKenzie et al., 2006). A discussion of the impacts of this assumption and other constraints of the surveillance program (e.g., only daytime flights) on model's interpretation is offered in the discussion section.

With this study we aim to achieve a better understanding of the distribution of oily discharges and of the underlying processes that drive the probability of occurrence and abundance of these events in Canada's Pacific region. We also provide insight into relatively new approaches that incorporate the complexities of observing an event that is rare and difficult to observe. By predicting likely oil pollution hot spots we hope to assist NASP managers to become more efficient in fighting chronic oil pollution as well as other decision-maker groups, whose mandate is the protection of coastal and marine environments. Finally, we hope this research will highlight the importance of collecting accurate oil spill observations, surveillance effort information and other factors known to affect the probability of detecting oil spills, to properly model chronic oil pollution in Canadian waters.

2.2 Data and methods

2.2.1 Study area and spatial scale selection

The study area includes almost the entire coast of British Columbia (Figure 2.1), and it was divided into two regions: offshore waters (Area A) and inshore waters (Area B). Inshore waters include the Juan de Fuca Strait, the Strait of Georgia, the Johnstone Strait, the Queen Charlotte Strait, and North Coast's inner passages. Most fjords were excluded from

the analysis because no shipping data was available at these locations. The offshore limit of Area A, approximately 400 km from BC's mainland, was defined by the quality and availability of the shipping data. This division was necessary to portray different geographic features and uses and to take into account data quality issues described in Section 2.2.3.

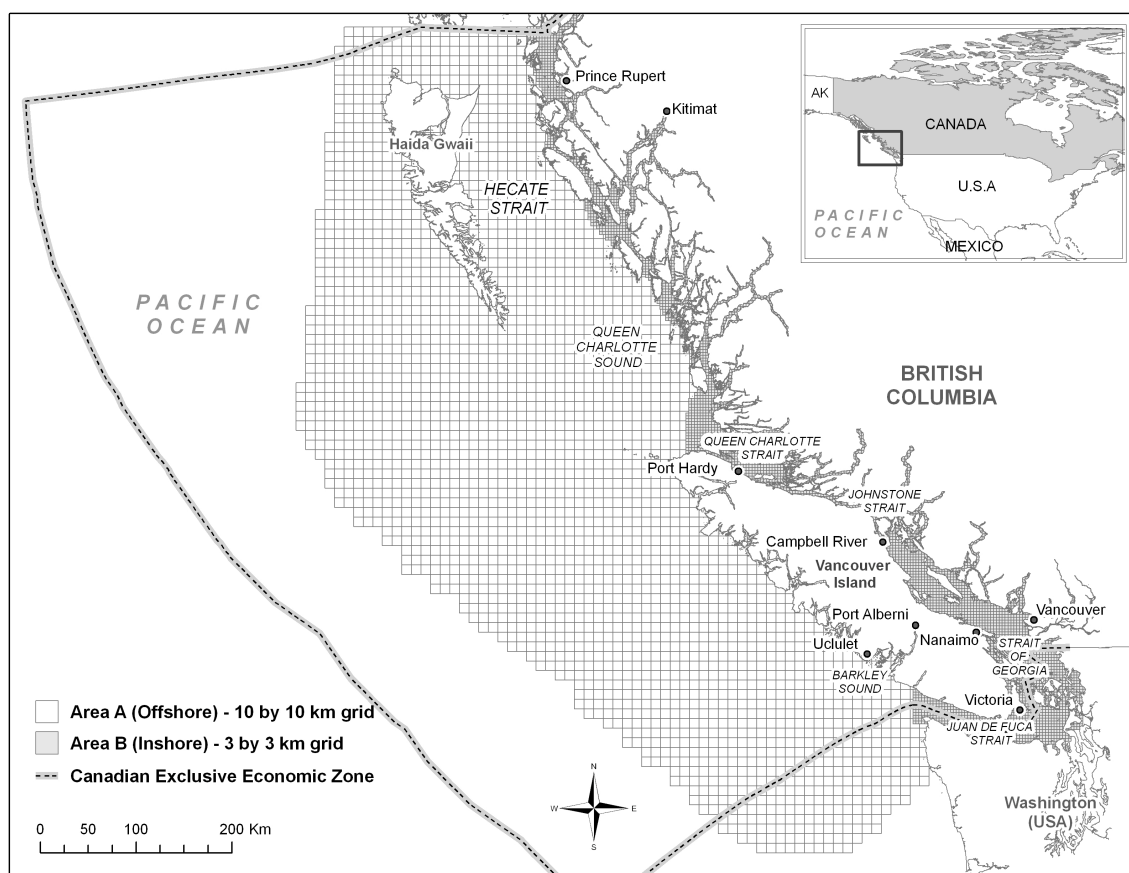


Figure 2.1. Extent of the study Area A (offshore) and Area B (inshore).

Each region was overlaid by a matrix of grid cells to facilitate the aggregation of model variables to the same spatial units. Area A was divided into 10 by 10 km cells and Areas B into 3 by 3 km cells. The rationale behind the selection of the two different cell

sizes resides on differences in the shipping traffic data quality; particularly, on variations in accuracy and positioning frequency (i.e., resolution of the recorded data) between marine vessels monitored offshore (~ 2 to 10 km between recorded vessel locations) and marine vessels monitored in inshore waters (~ 200 to 1000 m).

Only cells containing information of at least one surveillance flight conducted between 1998 and 2007 were kept for building the statistical model, which corresponds approximately to 41% (120,334 km²) of the total study area (see Figure 2.2). Table 2.1 shows the total area and total area surveyed by NASP in Areas A and B.

Table 2.1 Total study area and total area surveyed by NASP flights from 1998 to 2007 for each region.

	Total area (km ²)	Area (km ²) surveyed by NASP (1998-2007)
Area A	262,605 (2,811)	96,055 (1,064)
Area B	32,373 (6,304)	24,279 (3,938)
Total	294,978 (9,115)	120,334 (5,002)

Note: Numbers in parenthesis are number of cells or units.

2.2.3 Data description and preparation

NASP surveillance effort and oil spill observations

Date and location on flight routes and oil spill observations were collected and archived by the NASP crew and provided for this study directly by NASP.

A total of 924 individual flights with spatial-temporal information were available for this study, after excluding flights with no geographic information. Surveillance effort

was calculated by first converting each flight path from polylines to polygons by generating a one kilometre buffer on either side of each flight path producing a two kilometres swath. This buffer represents our conservative estimate, when compared with the 3.2 km estimate in technical reports (Armstrong, 2004), of the area observed from the NASP aircraft and without the assistance of especial sensors. Second, each buffered flight was intersected with each grid cell for Areas A or B. Third, the area was calculated for each intersected flight portion at each cell. Finally, the total area viewed by NASP per cell was estimated by adding the estimated area values in each cell. See Appendix A for a diagram of the described steps.

Figure 2.2 clearly depicts the heterogeneous distribution of NASP observational effort from 1998 to 2007. In Area B, surveillance effort concentrates near Transport Canada's main Aircraft Services Centre located in Richmond, BC, and along the coastline of British Columbia's mainland moving north-west. In Area B, flight coverage clearly follows the Canada - USA border and traffic separation schemes in the Juan de Fuca Strait. In Area A, NASP efforts concentrate at the entrance of the Juan de Fuca Strait, near the west side of Vancouver Island and at the entrance of Queen Charlotte Strait, reaching as far as 215 km offshore. Note that the disjoint observed in Figure 2.2 between Areas A and B is an artefact of estimating surveillance effort using two different grids.

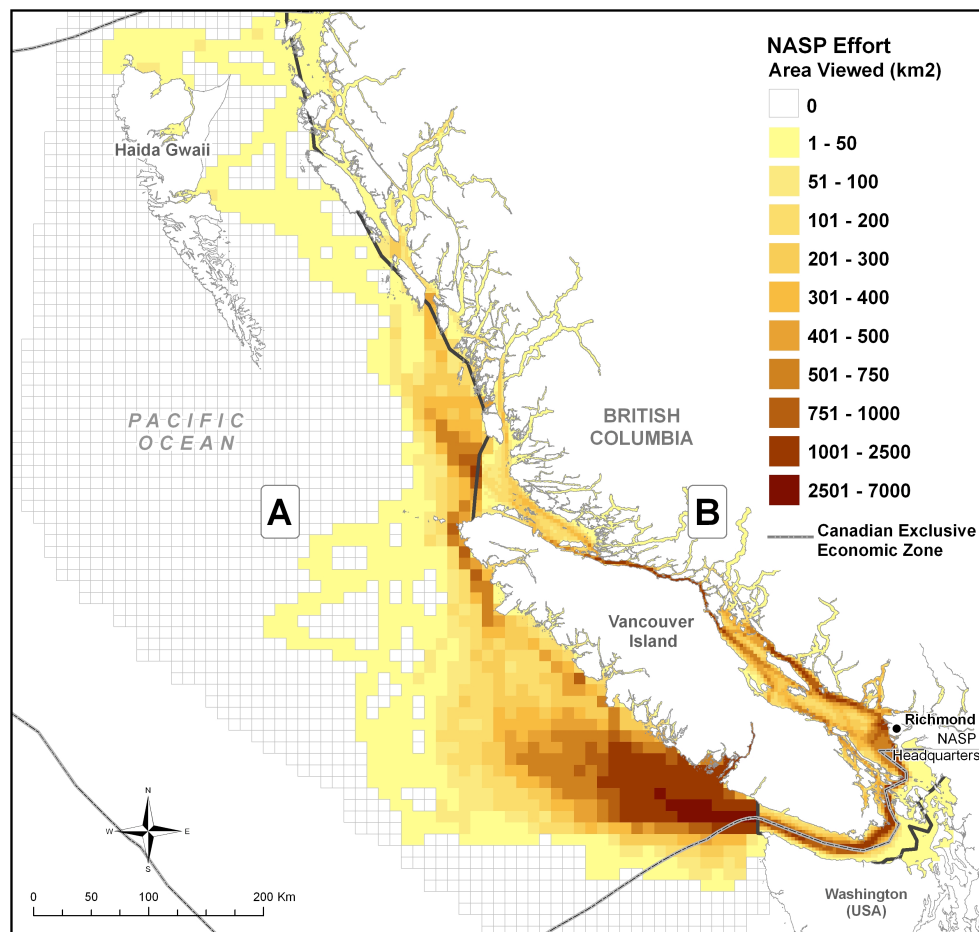


Figure 2.2 Distribution of NASP surveillance effort (in square kilometres) from 1998 to 2007 in Areas A and B.

A total of 441 suspected oil spills were observed between 1998 and 2007; however, only 328 spills were kept for the analysis. Only oil spill observations temporally associated with a flight path (i.e., date of detection) and without location errors (i.e., not visualized on land when plotted) were retained for analysis. With this data we extracted oil spill counts and presence/absence data per grid cell for each region. Figure 2.3 shows the distribution of oil spill counts in Areas A and B and in relation to the area covered by NASP. More oil spills were observed in Area B (284 spills) than in Area A (41 spills). In Area B, the largest oil spill counts concentrate near population centres where commercial and/or

pleasure vessels tend to converge and aggregate. In Area A, the largest number of possible oil spills is observed within and at the entrance of Barkley Sound, west of Vancouver Island. Grids and oil spill counts were created using ArcMap™ v.9.3 and Hawth's Analysis Tools v.3.27 (Eyer, 2004).

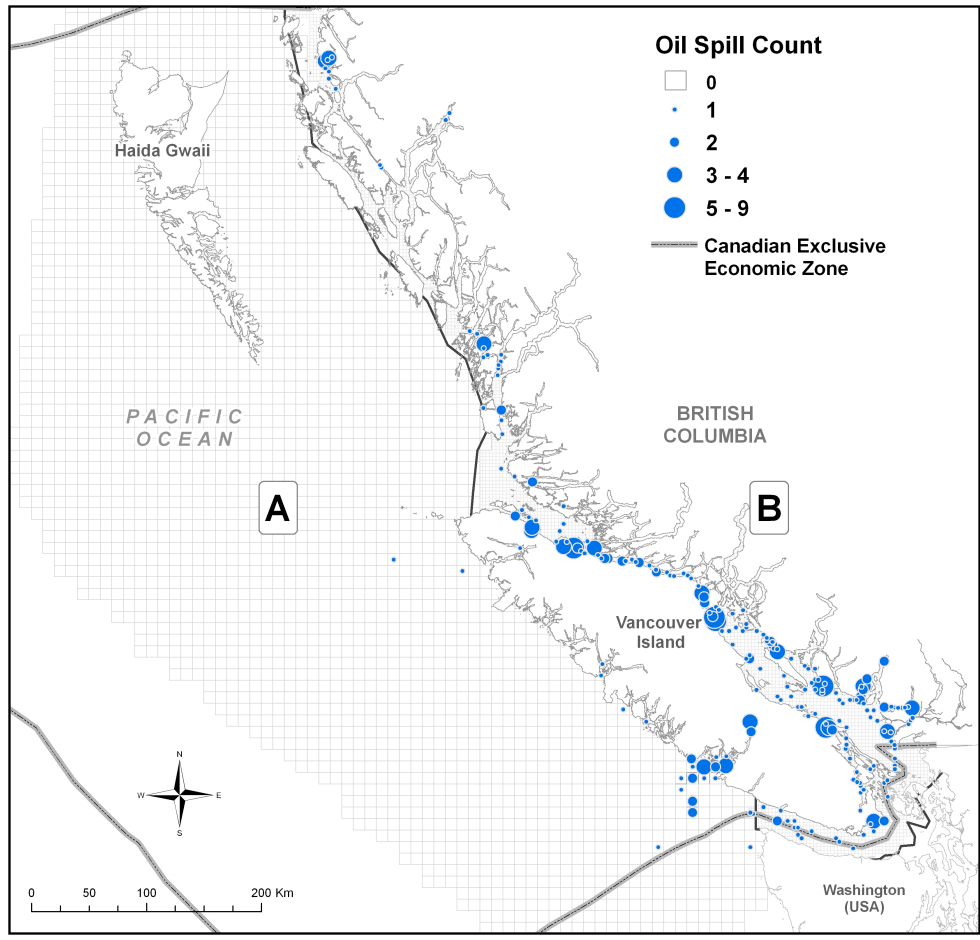


Figure 2.3 Distribution of oil spill counts observed by NASP from 1998 to 2007 in Areas A and B.

Marine traffic densities and patterns in Canada's Pacific Region

The majority of marine vessels larger than 20 m traveling within Canadian waters or coming from seaward need to obtain clearance from the Canadian Coast Guard Marine Communications and Traffic Service (MCTS) whose Vessel Traffic Service (VTS) zone they are about to enter. Participating vessels may be tracked by voluntary call-ins to MCTS or through the detection of radar tracking systems and/or AIS (Automated Identification Systems).

In the Pacific region, there are five VTS zones (see Figure B.1 in Appendix B), four in British Columbia with centres in Prince Rupert, Comox, Vancouver, Victoria and Tofino, and one collaborative station for the Puget Sound region based in Seattle (USA). Each centre uses one or two different vessel tracking systems: Prince Rupert operates with manual and fixed dead-reckoning; Comox and Vancouver with fixed dead reckoning; Victoria and Puget Sound with radar; and Tofino with radar and advance dead reckoning. These tracking modes vary in accuracy. The radar in Victoria is the most accurate system (~ 200 to 1000 m between tracked locations), whereas the advance dead reckoning in Tofino is the least accurate (~ 2 to 10 km between recorded locations). Radar systems, however, have limited coverage (or blind spots) in deep fjords and behind mountainous islands. An example of the different tracking modes is shown in Figure B.2, in the Appendix B.

All records from the different VTS centers are archived into a unique integrated information system, the Vessel Traffic Operations Support System (VTOSS). Each entry in the database contains vessel information and spatial-temporal records of each tracked vessel. Five years of VTOSS records, 2003, 2005, 2006, 2007 and 2008, were available for

this project. The algorithm used to extract daily means of shipping vessel movements per vessel type (i.e., carriers, oil tankers, fishing vessels, tug boats, cruise ships and ferries) for Areas A and B, consisted of: (1) removing replicated lines from the VTOSS database; (2) identifying and removing days with unusual ship movement totals (i.e., outliers) using the procedure RobustReg from SAS (Chen, 2002); and (3) estimating for each grid cell daily means, standard deviation and standard error, per season and vessel type for data pooled among years⁴. To remove potential biases from tracking mode limitations, cells with no shipping traffic data information were excluded from the analysis. No attempt was made to remove or fix vessel path anomalies within the dataset (e.g., shifts in vessel positions when travelling from one VTS zone to another and changing vessel tracking systems).

Figure 2.4 shows the distribution of mean yearly densities of shipping traffic per cell (calculated by multiplying daily values by 365 days) and for each vessel type. Relatively high shipping traffic concentrates along the west side of Vancouver Island and the Juan de Fuca Strait. Carrier vessels, including bulk cargo and containers (Figures 2.4a) and oil tankers (Figure 2.4b) generally follow one of the three Great Circle Routes in the North Pacific (shortest distances to and from principal Asian ports), or the Prince Rupert and Alaska routes. These routes meet north-west of Vancouver Island, to continue along the coast en route to Canadian and US ports via the Juan de Fuca Strait and using the Traffic Separation Schemes (TSSs), or to continue south along the west coast of the State of Washington (USA). The majority of loaded crude oil tankers coming from Alaska to refineries in the State of Washington, also known as TAPS (Trans Alaska Pipeline System)

⁴ This analysis was performed with help of Dr. Patrick O'Hara.

tankers, remain seaward of the Tanker Exclusion Zone (TEZ) situated 60 nm off the west coast of Vancouver Island (Canadian Coast Guard, 1998).

Tug boats (Figure 2.4c) are typically smaller and slower vessels, and choose to navigate close and along the coastline for safety reasons. These vessels dominate the narrow passages between Vancouver Island and the mainland, and the inner routes in BC's North Coast. Within the Juan de Fuca Strait, the majority of tug boats follow the Two-Way route introduced along the US coastline in 2002 by the International Maritime Organization to ensure the safety of small transient vessels (IMO, 2002).

Figure 2.4d depicts patterns of commercial fishing vessels, most of them large fishing factories, returning from or going to Alaska or British Columbia's fishing grounds. Unlike tug boats, most fishing vessels prefer to navigate along the Canadian coastline within the Juan de Fuca Strait and along the east side of Vancouver Island.

Finally, both cruise ships (Figure 2.4e) and ferries (Figure 2.4f) follow clearly defined routes that are distinct from other vessel types. One of the most popular cruise routes is known as the "Gulf of Alaska Route" where large cruise ships sail northbound or southbound along the west coast of Vancouver Island, between the port of Vancouver and Seward, Alaska and Prince Rupert, the last one being one of the most important stops.

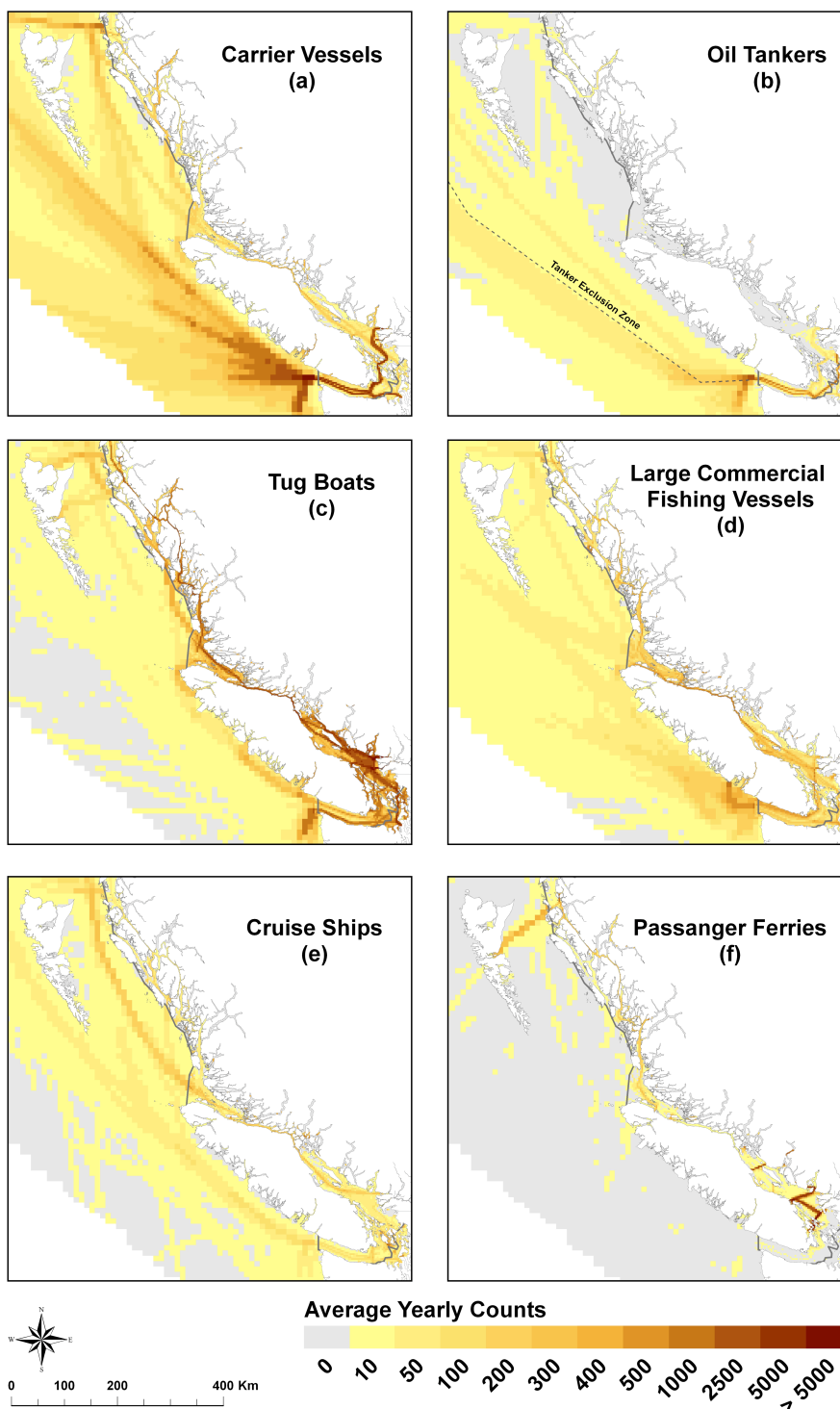


Figure 2.4 Intensity of vessel movements (number of movements per cell) averaged per year, for different vessel types, and pooled from data recorded in 2003, 2005, 2006, 2007 and 2008 in Canada's West Coast.

Figure 2.5 depicts the percentage of the mean number of vessel movements by vessel types in Areas A and B. In offshore waters (Area A), carrier vessels constitute more than half of the average number of vessels in a year, while tug boats are the most abundant in inshore waters (Area B).

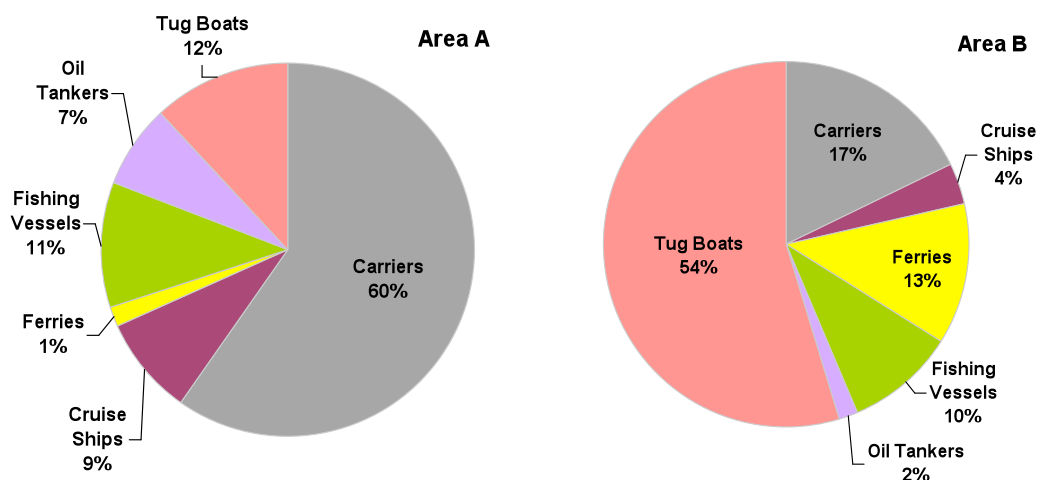


Figure 2.5 Percentage of mean yearly vessel movements per vessel type in Area A (left) and Area B (right)

Density of small craft harbours, distance to port and distance to shore.

The variable “density of small craft harbours” is included in the model to analyze the relationship between chronic oil pollution and recreational and commercial boating activities occurring within the premises of these facilities. Small craft harbours are becoming a focus of concern in terms of marine pollution as depicted by the emerging environmental programs in Canada (e.g., Clean Marine BC, Operation Clean Feathers) and the US (e.g., Oil Spill Task Force) to educate boaters about the impacts of small-scale oil pollution.

In this study, small craft harbours include marine recreation facilities (i.e., marinas, yacht clubs and public docks), fishing harbours and ferry docks. The Department of Fisheries and Oceans provided location information on fishing harbours and GeoBC (Government of British Columbia) provided data on the geographic location of ferry terminals and recreational facilities in BC.

For Area B, the kernel density estimator method (Silverman, 1986) was used to calculate absolute density estimates of small harbours. A bandwidth of one kilometre was used and thought to be a sufficient degree of smoothing to indirectly represent the intensity of recreational and commercial boating activities within and near small harbours. In fact, a similar degree of smoothing (or “distance of influence”) was reported in Ban et al. (2010) of 1.3 km for ports, marinas and harbours. The end grid had a resolution of three by three kilometres to match the existing grid for Area B. This variable was estimated using CrimeStat[®] v. 3.2a (Levine, 2004).

For area A, a simple count of number of small harbours per cell was performed because the cell size for this area (10 by 10 km) already extends beyond the maximum distance of influence specified above. Counts of small craft harbour per cell were calculated using Hawth's Analysis Tools for ArcGIS (Eyer, 2004).

Figure 2.6 depicts the estimated density distribution of small harbours in BC. Note that no data on small harbours was gathered for coastal regions in Area B and belonging to the State of Washington. Consequently, model results outside the limits of Canada's Exclusive Economic Zone will need to be interpreted accordingly.

Top commercial ports in BC (total of 13) were selected based on volume of cargo handled (domestic and international) between 2002 and 2006 (Statistics Canada, 2004-2009). In addition, four ports located in the State of Washington (U.S.A) were chosen for their proximity to other commercial ports in BC, their significance for international maritime trade and presence of refineries. The locations of the 17 selected ports in BC and Washington are shown in Figure 2.6. To estimate the Euclidean distance from each cell centroid to the nearest port, first Thiessen polygons were constructed based on ports' location to designate each cell to the closest port point; then the Euclidean distance was measured from each cell centroid to its assigned port point. Distance to shore was calculated based on the Euclidean distance from each cell centroid to the closest portion of coastline. These later procedures were performed in ArcMapTM v. 9.3 and ET GeoWizard v.9.7 (Tchoukanski, 2007).

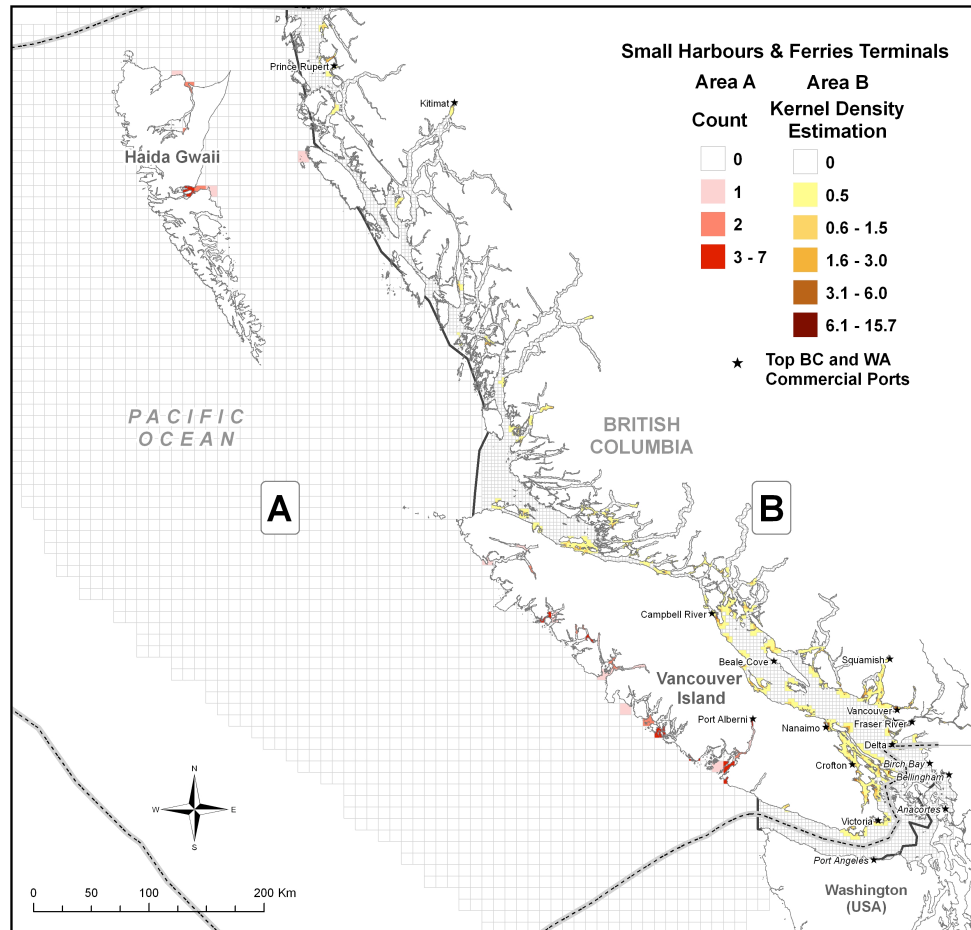


Figure 2.6 Density of small harbours (i.e., marinas, yacht clubs, public docks, fishing harbours and ferry terminals) and major commercial ports considered in this study.

2.2.5 Exploratory data analysis

Before applying any modelling approaches, it is recommended to run a series of exploratory analysis on both the dependent and predictor variables to avoid false model specifications (Zuur, Ieno & Elphick, 2010). Distribution shapes and presence of outliers in both the predictor and dependent variables were explored with histograms, QQ-plots and Cleveland dotplots. Scatterplots were used to explore relationships between predictors and response variables to assess presence of nonlinearity (Zuur, Ieno & Smith, 2007).

Correlation coefficients (r^2) and variance-inflation factors (VIF) (Montgomery & Peck, 1992) were estimated to measure the impact of collinearity among variables in the regression model; that is, the effect of incorporating redundant information and inflating the error term. VIF is given by $1/(1-r^2)$ where r^2 is the amount of variation in each predictor explained by the other set of predictor variables. The VIF value indicates how much larger the standard errors would be if predictors were uncorrelated with other predictors. Montgomery and Peck (1992) suggests that a predictor with a VIF value greater than 10 indicates serious collinearity; although, more stringent approaches are recommended to values as low as three (see Zuur, Ieno & Smith, 2007). In presence of collinearity, the suggested solutions are either to drop the term or combine it with other variables into a single index (O'Brien, 2007).

2.2.5 Model assumptions

Before we describe potential modelling approaches we should introduce several important assumptions established a priori:

- all observed oil spills are a consequence of marine vessel operations;
- all observed oil spills are true, and hence there are no “false positives”;
- location information on oil spill data and surveillance flight is accurate;
- no drift is assumed between the time of the spill and its detection;
- mean shipping traffic densities are representative of shipping patterns and densities throughout the study period (1998-2007); and
- the likelihood of oil spill detection only depends on surveillance effort and not on other factors such as weather conditions, altitude of the aircraft and sea conditions.

2.2.6 Candidate model distributions

Modelling oil spill abundances (count model)

The simple and most frequently used approach to model count data is generalized linear models (GLMs) (McCullagh & Nelder, 1989) with Poisson distribution or negative binomial distribution, and a link function. The link function allows the linear model to be related to the dependent variable, y_i , and predictions to be within the range according to the dependent variable. When the dependent variable, y_i , represents the number of occurrences of an event with a Poisson distribution, its probability density distribution is given by:

$$f(y_i; \mu_i) = \frac{\mu_i^{y_i} \times e^{-\mu_i}}{y_i!} \quad y_i = 0, 1, 2, 3, \dots \quad (2.1)$$

The parameter μ_i is the expected mean and its related through a log-link to a linear combination of unknown parameters as follows:

$$\log(E(y_i; x_i)) = \log(\mu_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq} \quad (2.2)$$

or

$$E(y_i; x_i) = \mu_i = e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq}} \quad (2.3)$$

where β_0 is the intercept and β_n the slopes to be estimated using maximum likelihood estimation (MLE).

Poisson distribution assumes equidispersion; that is, the expected mean is equal to the expected variance ($E(y_i) = \mu_i = \text{var}(y_i)$). When this assumption is violated and the variance is greater than the mean, the distribution is said to be overdispersed (McCullagh & Nelder, 1989). In presence of overdispersion, fitting GLMs with negative binomial distributions often provide better results. The difference between the Poisson and the negative binomial distribution is the introduction of parameter k (or dispersion parameter)

in the variance term as: $var(Y_i) = \mu_i + \mu_i^2/k$, which describes the degree of overdispersion present relative to a Poisson distribution. The smaller k in reference to the mean the larger the overdispersion, as it increases the difference between the mean and the variance (Zuur et al., 2009). The negative binomial probability density distribution is given by:

$$f(y_i; k, \mu_i) = \frac{\Gamma(y_i + k)}{\Gamma(k) \times \Gamma(y_i + 1)} \times \left(\frac{k}{\mu_i + k} \right)^k \times \left(1 - \frac{k}{\mu_i + k} \right)^{y_i} \quad (2.4)$$

where $\Gamma(y_i + 1)$ is equal to $(y_i + 1)!$. The expected mean of y_i is given by μ_i and follows equation 2.2.

In presence of zero-inflation (i.e., excess of false or true zeros), count data can be modelled using a zero-inflated (ZI) model described by Lambert (1992). Zero-inflated models based on a Poisson distribution (ZIP), or a negative binomial distribution (ZINB), can be defined in two components:

$$Y_i \sim \left\{ \begin{array}{l} 0, \text{ with probability } \pi_i \\ P(\mu_i) \text{ or NB}(\mu_i), \text{ with probability } 1 - \pi_i \end{array} \right\} \quad (2.5)$$

The first component defines the probability to observe a zero count, which in ZIP is specified as:

$$f(y_i = 0) = \pi_i + (1 - \pi_i) \times e^{-\mu_i} \quad (2.6)$$

and in ZINB as:

$$f(y_i = 0) = \pi_i + (1 - \pi_i) \times \left(\frac{k}{\mu_i + k} \right)^k \quad \text{where } k \text{ is the dispersion parameter} \quad (2.7)$$

The parameter π_i represents the probability that y_i is zero and follows a binomial distribution:

$$\pi_i = \frac{e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_q X_{iq}}}{1 + e^{\beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_q X_{iq}}} \quad (2.8)$$

and $e^{-\mu_i}$ (in equation 2.6) is the probability to measure a zero based on a Poisson or negative binomial distribution.

The second component defines the probability to observe a non-zero count as:

$$f(y_i | y_i > 0) = (1 - \pi_i) \times f_p(y_i) \quad \text{in the case of ZIP} \quad (2.9)$$

$$f(y_i | y_i > 0) = (1 - \pi_i) \times f_{NB}(y_i) \quad \text{in the case of ZINB} \quad (2.10)$$

where $1 - \pi_i$ defines the probability of not measuring zero multiplied by the probability of observing any number of events at site i based on a Poisson (equation 2.8) or a negative binomial distribution (equation 2.9). The function $f_p(y_i)$ is given in equation 2.1 and $f_{NB}(y_i)$ in equation 2.3. The expected mean for both model distributions is equal to μ_i , multiplied by the probability of measuring false zeros:

$$E_{ZI}(Y_i) = \mu_i \times (1 - \pi_i) \quad (2.11)$$

where μ_i can be modelled using a set of covariates as introduced in equation 2.2. The covariates to model μ_i can be the same or differ from covariates used to model π_i (Lambert, 1992). The joint likelihood of μ_i and π_i is maximized using the expectation-maximization (EM) algorithm (Dempster, Laird, & Rubin, 1977; Lambert, 1992).

As in zero-inflated models, hurdle models (ZA) (Mullahy, 1986) also have two-components. The first component is a binary (presence/absence) outcome model. If an event is present, then the hurdle is crossed and the positive counts, which are truncated at zero, are modelled by either a Poisson (ZAP) or negative binomial (ZANB) distribution,

depending on the degree of overdispersion. More formally, the probability function for ZAP model is:

$$f_{ZAP}(y; \beta, \gamma) = \begin{cases} f_{\text{binomial}}(y = 0; \gamma) & y = 0 \\ (1 - f_{\text{binomial}}(y = 0; \gamma)) \times \frac{f_{\text{Poisson}}(y; \beta)}{1 - f_{\text{Poisson}}(y = 0; \beta)} & y > 0 \end{cases} \quad (2.12)$$

where f_{binomial} determines the probability to measure zero events in the first part the function.

The second component defines the probability of observing positive counts as the probability that it is not a zero observation ($1 - f_{\text{binomial}}$), multiplied by the probability of a positive observation with a Poisson or negative binomial distribution truncated at zero.

The parameters γ and β are the unknown coefficient parameters to estimate (Zeileis, Kleiber, & Jackman, 2008). In hurdle models, the probability of observing true zero counts is given by π_i , which follows the same function as in equation 2.7. The expected mean for ZAP follows:

$$E_{ZAP}(Y_i; \pi_i, \mu_i) = \mu_i \times \left(\frac{1 - \pi_i}{1 - e^{-\mu_i}} \right) \quad (2.13)$$

and for ZANB:

$$E_{ZANB}(\text{Spills}) = \mu_i \times \left(\frac{1 - \pi_i}{1 - P_0} \right) \quad \text{where} \quad P_0 = \left(\frac{k}{\mu_i + k} \right)^k \quad (2.14)$$

To explore which of the model specifications is more appropriate to model oil spill counts in Areas A and B, Poisson (P), negative binomial (NB), zero-inflated (ZI) and hurdle (ZA) models were fitted using code implemented in R (R Development Core Team, 2008). Specifically, P models were fitted using `glm()` from the **stats** package (Chambers & Hastie, 1993), NB models using `glm.nb()` functions from the MASS package (Venables

& Ripley, 2002), and `zeroInfl()` and `hurdle()` function from the **pscl** packages (Zeileis et al., 2008).

Modelling the probability of oil spill occurrence (binomial model)

To model the probability of oil spill occurrence, count data can be easily transformed into presence/absence values and modelled as a binomial distribution. In logistic regression, which also forms part of the GLMs (McCullagh & Nelder, 1989), the expected mean and variance are given by $E(Y_i) = \pi_i$ and $var(Y_i) = \pi_i x (1 - \pi_i)$. In logistic regression the parameter π_i is equal to:

$$\pi_i = \frac{e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq}}}{1 + e^{\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq}}} \quad (2.15)$$

where again β_0 is the intercept and β_n the slopes to be estimated using maximum likelihood estimation (MLE). The logit transformation (Hosmer & Lemeshow, 2000) is used in logistic regression to link occurrences with a linear combination of exploratory variables, as follows:

$$g(x) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_q x_{iq} \quad (2.16)$$

The logistic regression models for Areas A and B were fitted in R using `glm()` from the **stats** package (Chambers & Hastie, 1993).

2.2.7 Surveillance effort as an offset variable

For each regression model, NASP surveillance effort was included as an offset variable. Because the effort employed to observe oil pollution was not homogeneous across the

study area, oil spills counts had to be standardized and interpreted as a rate of occurrence or number of oil spills per unit of effort. Rates can be analyzed as count data in loglinear models by moving the denominator (i.e., unit of effort) to the right-hand side of the regression model as $\log(\text{effort})$ with a fixed coefficient equal to one (Agresti, 1990; McCullagh & Nelder, 1989).

2.2.8 Model calibration and model selection

Selecting which predictors to include in the model is one of the most important and difficult tasks in the model building process as it will determine the accuracy and the predictive power of the final outcome (Harrell, Lee, & Mark, 1996). Automated (forward selection or backward elimination) methods are frequently used to facilitate finding an optimal model from the proposed group of initial predictors. We selected the backward elimination approach with likelihood-ratio test (based on the chi-square χ^2 statistic) (McCullagh & Nelder, 1989; Tabachnick & Fidell, 2007). This approach consists of dropping one term in turn and estimating the difference between the deviance of the nested model and the full model. A nonsignificant χ^2 statistic ($p > 0.05$) indicates that there is no difference between the full and nested model and hence the dropped term can be removed from the model. This method was preferred over the forward selection method (which adds one variable in turn) because the later did not produce more parsimonious models. This procedure was applied to a full P model and BI model for Areas A and B and automated using the function `drop1()` from the **stats** package in R (Chambers & Hastie, 1993).

The selected terms were used to refit P, NB and BI models. Likelihood ratio tests were used to explore overdispersion and determine the most appropriate distribution between P and NB. ZI and ZA models were fitted once the best distribution and predictors were selected. Non-significant coefficients (i.e., $p > 0.05$) were dropped from each fitted model until all coefficients were significant.

Akaike's information criterion (AIC) (Akaike, 1973) or the Bayes' information criterion (BIC) (Schwarz, 1978) were used to select the best model specification. These indices are commonly used to compare non-nested models or models that "cannot be represented as special cases of each other" (p.182, Cameron & Trivedi, 1998). Both criteria are based on the principle of parsimony; that is, they take into account the number of parameters to be included in the model. AIC penalizes the log-likelihood function by adding two times the number of parameters included in the model and takes the form of:

$$\text{AIC} = -2 * (\text{maximized log-likelihood}) + 2 * (\text{no. parameters}) \quad (2.17)$$

whereas BIC applies a heavier penalty by including the number of observations (Burnham & Anderson, 2002):

$$\text{BIC} = -2 * (\text{maximized log-likelihood}) + \text{Log}(\text{no. observations}) * (\text{no. parameters}) \quad (2.18)$$

The Vuong (V) test (Vuong, 1989) was also estimated as it is often employed when comparing nonnested models for count data (Greene, 1994). V values greater than 1.96 mean that the first model is superior over the second model, whereas V values less than -1.96 suggest that the second model is preferred over the first one (Liu & Cela, 2008). V values in between 1.96 and -1.96 indicate that the test does not favour one model or the other (Greene, 1994).

Finally, observed oil spill counts were plotted against predicted counts and observed zeros versus predicted number of zeros to assess each model performance or the ability to discriminate between positive and zero oil spill counts.

2.2.7 Assessing goodness-of-fit

Once the model is built, a number of measures are available to measure the fit, or goodness-of-fit, of the model; that is, how well the model describes the dependent variable (Hosmer & Lemeshow, 2000). To assess the performance of oil spill abundance models we considered the pseudo R-squared (or explained deviance), which is estimated as:

$$\text{Pseudo } R^2 = 100 \times [(\text{null deviance} - \text{residual deviance}) / \text{null deviance}] \quad (2.19)$$

where the null deviance is the deviance of the model with the intercept only, and the residual deviance is the deviance that cannot be explained after all final variables have been included (Guisan & Zimmermann, 2000). The higher the percentage (or the smaller the residual deviance) the better the model is or, in other words, the more variance is explained.

The area under receiver (or AUC) was estimated to evaluate the predictive performance of the logistic regression models. This index is estimated based on the area under the receiver operating characteristic curve (or ROC) and it is equal to the probability that a randomly selected positive event will be ranked higher than a randomly selected negative one. AUC values range from 0 to 1, with values of 0.5 indicating that the model lacks discrimination capacity (Pearce & Ferrier, 2000).

2.2.8 Residual analysis

Residuals measure the difference between the observed and the fitted values. The analysis of residuals can help to assess model misspecifications and detect outliers or observations with poor fit (Cameron & Trivedi, 1998). Contrary to the previous goodness-of-fit measures, these analyses allow a more in depth exploration of model fit over the entire set of covariate patterns (Hosmer & Lemeshow, 2000).

Plotting standardized model residuals (i.e., rescaled residuals by their standard deviation) against the fitted values is a common procedure to validate the performance of a selected model and allow the identification of outliers. In large sample sizes, standardized residuals are assumed to follow a t -distribution, and hence upper and lower limits can be established at the 95% confidence intervals (t value = ± 1.96). Statistically significant outliers are those exceeding these limits (Hair et al., 1995). Identified outliers can then be mapped to facilitate their interpretation and description.

The presence of patterns in residual plots may also indicate violation of the assumption of independence in model residuals or that the model may be missing important information (Zuur et al., 2009). Residuals heterogeneity may rise because the assumption of linearity between the response and the predictors is violated or because spatial structure in the data is not acknowledged in the model building. We used Moran's I to test the presence of spatial autocorrelation in the residuals (Dormann et al., 2007; Fortin & Melles, 2009) and assess the need to incorporate a spatial structure into the model. Moran's I varies from -1 (negative spatial autocorrelation) to +1 (positive spatial autocorrelation), while 0 indicates lack of spatial autocorrelation (Bailey & Gatrell, 1995). Local Moran's I_i (Anselin & Bera, 1998) was also estimated using model residuals and mapped in GIS to

assess the presence of regional spatial heterogeneity by identifying local aggregations of similar values (positive autocorrelation), either of low residuals (overestimated predictions) or high residuals (underestimated predictions), or clusters of dissimilar values (negative autocorrelation) (Zhang, Gove, & Heath, 2005). Both Moran's I test and local Moran's I_i were estimated using GeoDATM (Anselin, Syabri & Kho, 2005).

2.2.9 Model predictions

After model selection and calibration, the best performing count and logistic models for Areas A and B were used to predict the number and probability of oil spills occurrences at each location of the study area visited or not by NASP overflights between 1998 and 2007. The function used to extract both the expected mean and probabilities was `predict()` from the **stats** package in R (Chambers & Hastie, 1993). The computed predicted values were mapped in ArcMap for their visual interpretation.

2.3 Results

2.3.1 Exploratory data analysis

Table 2.2 and 2.3 provide a brief description of the dependent variable (i.e., number of detected oil spills) and the candidate predictors thought to explain variation in oil spill abundances and probability of occurrence in Areas A and B respectively.

Table 2.2 Summary statistics of response and predictors for models in area A

Variable name	Variable function	Description	Mean	Min.	Max.	Std. Dev.
spill	Response	Number of detected oil spills	0.04	0	4	0.3
effort	Offset	Area surveyed by NASP (km ²)	303.1	0.001	7154.6	615.4
count_sh	Predictor	Number of small harbours	0.06	0	7	0.5
disshorkm	Predictor	Distance to shore (km)	46.6	0.01	194.7	44.6
disportkm	Predictor	Distance to the nearest port (km)	200.6	4.4	430	89.6
carroilfish_mean	Predictor	Mean daily carrier, oil tanker and fishing vessel movements	0.2	0	6.9	0.4
crui_mean	Predictor	Mean daily cruise ship movements	0.1	0	0.7	0.1
ferr_mean	Predictor	Mean daily ferry movements	0.01	0	0.9	0.1
tug_mean	Predictor	Mean daily tug boat movements	0.1	0	3.4	0.4

Table 2.3 Summary statistics of response and predictors for area B.

Variable	Variable function	Description	Mean	Min.	Max.	Std. Dev.
spill	Response	Number of detected oil spills	0.1	0	9	0.4
effort	Offset	Area surveyed by NASP (km ²)	228.6	0	2504.4	336.8
density_sh	Predictor	Density of small harbours	0.1	0	15.7	0.6
disshorkm	Predictor	Distance to shore (km)	77.3	0.5	264.2	73.9
disportkm	Predictor	Distance to the nearest port (km)	2.7	0	26.1	3.5
carr_mean	Predictor	Mean daily carrier movements	0.6	0	13.6	1.6
crui_mean	Predictor	Mean daily cruise ship movements	0.1	0	2.6	0.2
ferr_mean	Predictor	Mean daily ferry movements	0.4	0	35.9	2.2
fish_mean	Predictor	Mean daily fishing vessel movements	0.3	0	2.3	0.4
oil_mean	Predictor	Mean daily oil tankers	0.1	0	2.1	0.2
tug_mean	Predictor	Mean daily tug boat movements	1.7	0	34.5	2.5

The frequency distributions of oil spill counts for Areas A and B clearly reveal an extraordinarily high number of zeros observations (Figure 2.7). The exploration of histograms and QQ-plots revealed a degree of skewness or non-normality in predictor's

distributions. Variables distance to shore (for Areas A and B), distance to port (for Areas A and B) and tug and fishing vessel mean daily movements (for Area B) were log transformed to normalize their distributions. The remaining variables were kept untransformed because none of the common transformations helped to normalize their distributions.

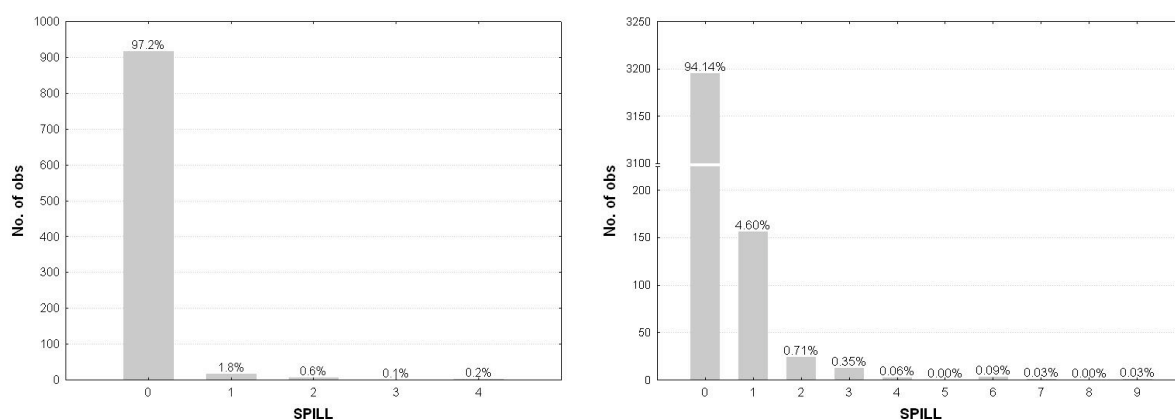


Figure 2.7 Frequency distribution of oil spill counts in Areas A (left) and B (right).

In Area A, carrier and oil tanker daily mean values presented an almost perfect linear relationship ($r^2 = 0.9$, $p < 0.001$; $VIF_{carr_mean} = 8.1$; $VIF_{oil_mean} = 5.7$). To reduce the risk of collinearity during the regression analysis, these variables were combined and averaged in one group. Industrial fishing vessels also show similar distributions to carriers and oil tanker distributions ($r^2 = 0.7$, $p < 0.001$; $VIF_{fish_mean} = 2.4$), and hence they were also combined with the previous group to avoid collinearity. The resultant new variable is named "carroilfish_mean". Distance to port was removed from the modelling process in Area A because it was significantly correlated with distance to shore ($r^2 = 0.5$, $p < 0.001$). For Area B, no strong correlations were detected among model predictors and they were kept untransformed. Finally, scatterplots between oil spill counts and explanatory variables

did not indicate the presence of nonlinear relationships, and hence no attempt was made to incorporate quadratic terms or regression splines into the models. Scatterplot matrices of model variables for Areas A (Figures C.1, C.2 and C.3) and B (Figures C.4, C.5 and C.6) can be found in the Appendix C.

2.3.2 Model calibration

Table 2.4 shows the summary of coefficient and evaluation indices for each regression model selected for Area A. Results of the variable selection procedure are summarized in Tables C.1 and C.2 in Appendix C. The first column presents the best fitted logistic model (A.BI), while the remaining columns show the best fitted models using oil spill count data. Likelihood ratio test ($X^2 = 1.5$, $df=1$, $p=0.23$) indicated that model fit did not vary significantly between Poisson (A.P) and negative binomial model (A.NB) distributions. Model A.P revealed similar AIC and BIC values in comparison with the model specifications A.NB and A.ZIP, whereas A.ZAP had the poorest fit. V tests indicated a slightly preference towards the A.ZIP model when compared with A.P ($V = -1.53$, $p=0.06$), but a clear preference for A.P when compared with the A.ZAP ($V = 2.55$, $p=0.005$).

Table 2.4 Summary of the best-fit models for presence/absence data and oil spill counts in Area A (offshore).

	A.BI		A.P		A.NB		A.ZIP		A.ZAP	
	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error
	<i>Binomial model</i>									
(intercept)	-7.92	0.28	-8.04	0.21	-8.02	0.23	-7.63	0.3	-7.2	0.27
log(disshorkm)	-0.60	0.1	-0.50	0.08	-0.53	0.08	-0.25	0.12	-	-
count_sh	0.33	0.2	0.24	0.10	0.28	0.11	0.34	0.12	0.4	0.11
k(theta)			2.87	3.37						
	<i>Zero-inflated model</i>									
(intercept)									-3.13	0.24
log(disshorkm)					0.91	0.39			-0.32	0.08
count_sh									0.55	0.17
no. observations	1064		1064		1064		1064		1064	
no. parameters	3		3		4		5		5	
log L	-83.2		-100.3		-99.6		-96.7		-119.7	
AIC	172.3		206.6		207.2		205.4		251.5	
BIC	187.2		221.2		226.6		234.5		280.6	

Note: Dashes indicate terms dropped from the full model. All remaining coefficients and standard errors have a significant level of at least $p < 0.05$

Further analysis showed that A.ZIP model was more accurate in estimating the number of oil spill positive counts (particularly the number of ones) out of the four models estimated for Area A (Figure 2.8). All count models were comparable when capturing the number of zeros (Figure 2.9).

Even though the A.ZIP model performed slightly better than A.P, both models are comparable. We decided to select the A.P model as the best model to predict oil spill abundances in Area A because it was the most parsimonious and easiest model to interpret. The A.BI model was chosen to predict the probability of oil spill occurrences.

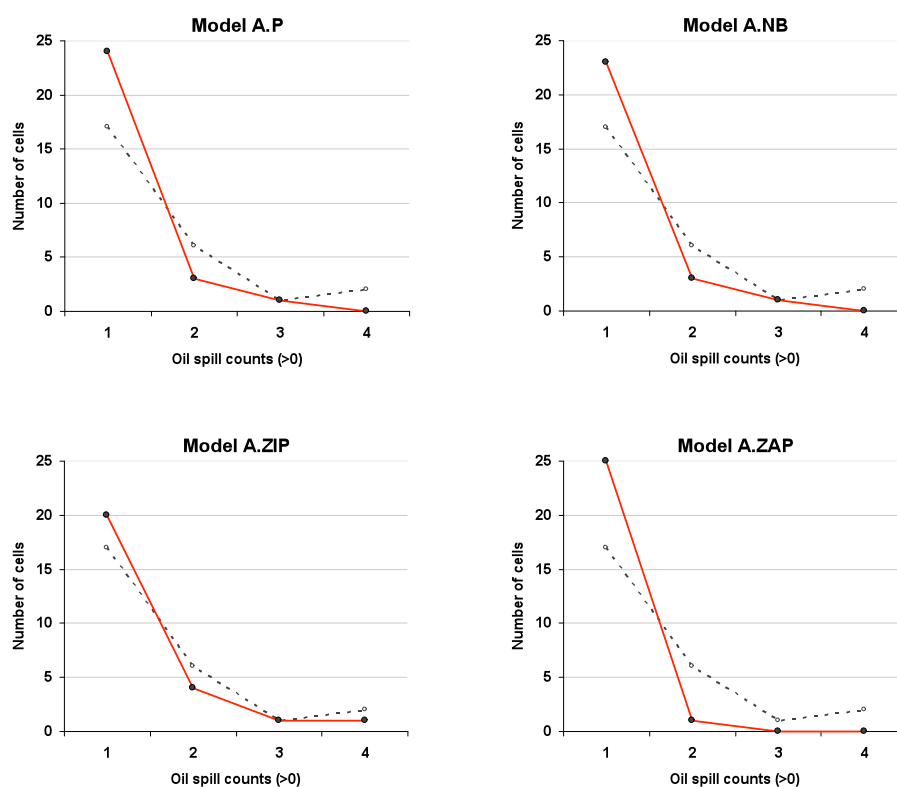


Figure 2.8 Observed (dashed line) and predicted (solid-red line) frequencies of positive oil spill counts in area A.

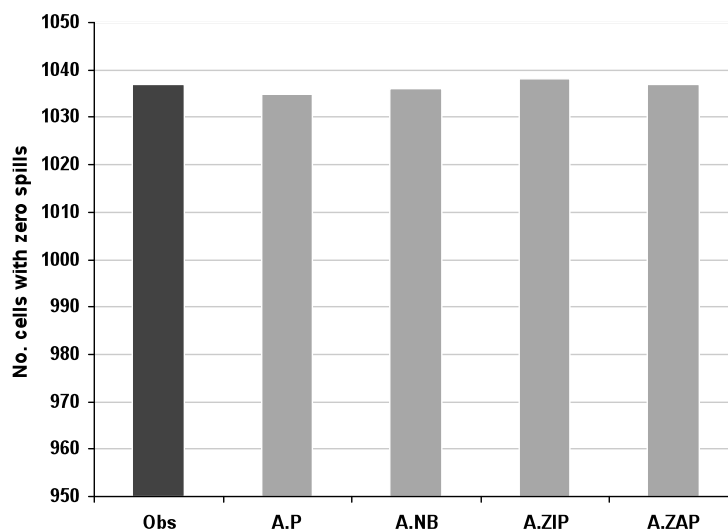


Figure 2.9 Observed and predicted frequencies of zero oil spill counts in Area A.

The first column of Table 2.5 presents the results of the best-fit logistic model (B.BI), while the rest of the columns show the best-fit count models for Area B. Results of the variable selection procedure are summarized in Tables C.3 and C.4 in the Appendix C. Oil spill counts showed evidence of overdispersion as indicated by the significant results of the likelihood ratio test when comparing the B.NB model with the B.P model ($\chi^2 = 118.35$, $df = 1$, $p < 0.001$). The V test ($V = -2.44$, $p < 0.01$) and AIC values indicated that zero-inflated model B.ZINB fit slightly better than the standard B.NB model; although, the B.NB model did show the lowest BIC value. As in Area A, the worst performing model was B.ZANB.

Table 2.5 Summary of the best-fit models for presence/absence data and oil spill counts in Area B (inshore).

	B.BI		B.P		B.NB		B.ZINB		B.ZANB	
	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error	Coef.	Stand. Error
	<i>Binomial model</i>									
(Intercept)	-8.38	0.17	-8.27	0.14	-8.24	0.17	-7.50	0.13	-	-
carr_mean	-0.18	0.06	-0.21	0.05	-0.20	0.06	-	-	-	-
crui_mean	-	-	0.70	0.27	-	-	-	-	-	-
log(tug_mean)	0.44	0.12	0.35	0.09	0.40	0.12	-	-	-	-
oil_mean	1.14	0.43	1.11	0.38	1.22	0.43	-	-	-	-
log(disshorkm)	-0.42	0.05	-0.39	0.04	-0.47	0.05	-0.68	0.07	-0.69	0.18
density_sh	0.46	0.09	0.27	0.03	0.53	0.06	0.33	0.08	0.52	0.24
<i>k</i>					0.38	0.07	-0.71	0.22	-	-
	<i>Zero-inflated model</i>									
(Intercept)									-4.23	0.16
carr_mean					0.31	0.12			-	-
crui_mean					-	-	-	-	-	-
log(fish_mean)					-	-	-	-	0.98	0.28
log(tug_mean)					-1.73	0.68			1.07	0.12
oil_mean					-	-	-	-	0.64	0.28
log(disshorkm)					-0.92	0.29			-0.15	0.05
density_sh					-6.09	2.81			0.46	0.10
no. observations	3394		3394		3394		3394		3394	
no. parameters	6		7		7		9		10	
log L	-615.21		-826.4		-768.19		-759.6		-826.7	
AIC	1242.4		1666.7		1550.4		1537.2		1673.4	
BIC	1279.2		1709.6		1593.3		1673.3		1734.6	

Note: Dashes indicate terms dropped from the full mode. All remaining coefficients and standard errors have a significant level of at least $p < 0.05$

While comparing predicted with observed count data values in Area B, the B.NB model outperformed the B.P model by capturing more accurately the number of ones (Figure 2.10) and the number of zero observations (Figure 2.11). Predicted positive counts were comparable when estimated by B.NB and its zero-inflation model counterpart B.ZINB. Similarly, both B.NB and B.ZINB models slightly overpredicted the number of zeros. As in Area A, the hurdle model B.ZANB characterizes perfectly the number of zeros, but it poorly predicted the number of oil spills counts greater than zero. In absence of clear differences between the B.NB and B.ZINB, the B.NB model was selected as the best model because it provides easier interpretation and lower number of parameters (i.e., more parsimonious). B.BI model was selected to predict the likelihood of oil spill occurrences in Area B.

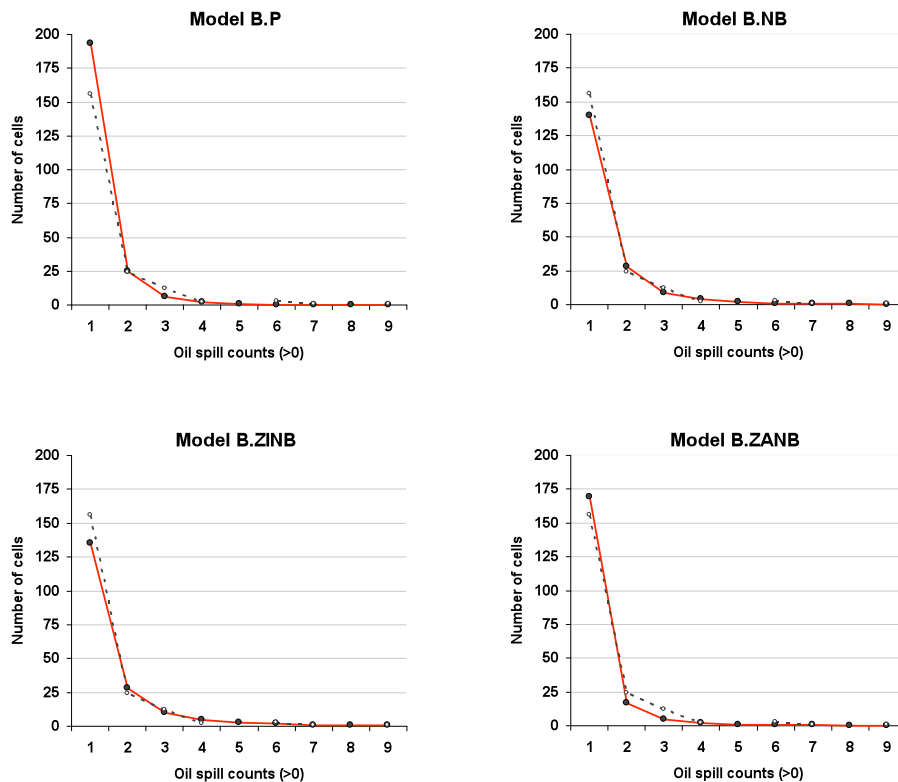


Figure 2.10 Observed (dashed line) and predicted (solid-red line) frequencies of positive oil spill counts in Area B.

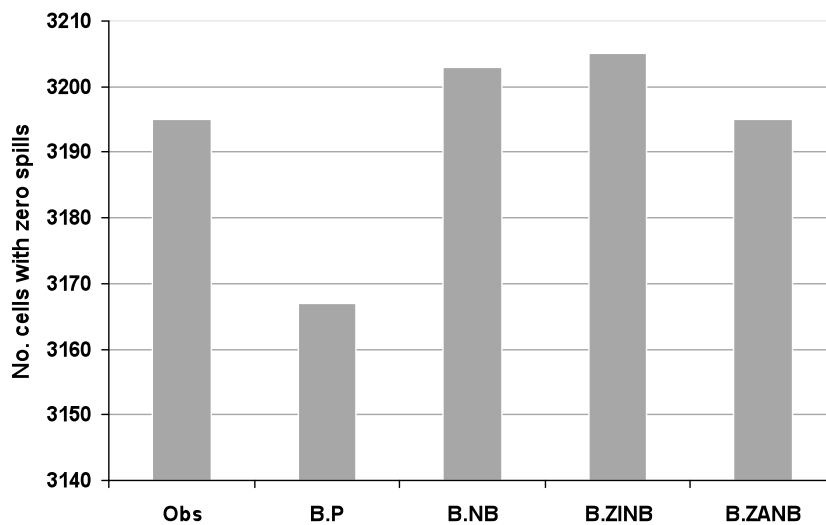


Figure 2.11 Observed and predicted frequencies of zero oil spill counts in Area B.

2.3.3 Assessing goodness-of-fit

The explained deviance given by A.P model in Area A was equal to 37.5% of the variation in oil spill counts, while the B.NB model explained only 22.4% of the variation in oil spill counts in Area B. The ability to discriminate between presence and absence of oil spill observations by both logistic models A.BI and B.BI was excellent based on the estimated AUC values of 0.89 for the first one and 0.83 for the second one.

2.3.4 Regression diagnostics

Model residuals for the selected count model for Areas A and B were analyzed both visually and empirically to detect patterns indicating systematic errors in model fit as well as outliers. Diagnostic plots for the logistic regression models (i.e., A.BI and B.BI) are not presented since they show similar results to the plots presented herein. To assess the presence of patterns in model residuals, standardized Pearson residuals were plotted against fitted values for Areas A (Figure 2.12) and B (Figure 2.13). No clear patterns were detected in either plots, however a number of outliers were visually identified.

In Area A, eight points were identified as significant outliers, out of which three (IDs 4166, 6119 and 6807) were highlighted as “extreme” significant outliers for having the highest residual values (Figure 2.12, left). A fourth point (ID 6642) also stood out in the far lower-right corner of the plot and identified as an over-predicted value. These four points were mapped to facilitate their description and interpretation (Figure 2.12, right). The three identified extreme and significant outliers are located off the west coast of Vancouver Island, an area characterized by low surveillance effort but dominated by large vessel traffic, particularly carriers and oil tankers. The over-predicted value is found closer

to shore, near Bamfield, an area containing a number of public docks (four, including a fuel station) and one of the largest number of detected oil spills (four) in Area A.

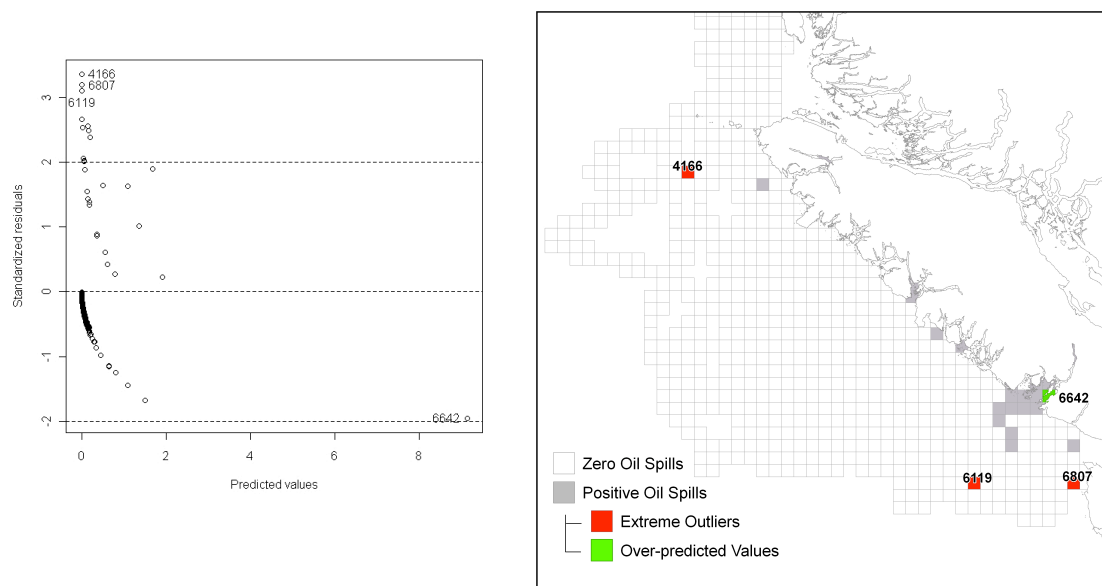


Figure 2.12 Standardized residuals against predicted values (left). Plotted horizontal dashed lines represent the 95 % confidence intervals. Map (right) highlighting the identified extreme significant outliers and over-predicted value of the A.P model.

In Area B, a number of significant outliers are detected but only four points are highlighted as “extreme” outliers, two positive (IDs 18432, 20735) and two negative (IDs 39534, 45198). These last two also are identified as over-predicted values. When mapped, all listed outliers (in red) appear adjacent to the coastline (Figure 2.13, left). The highest positive residual (ID 18432) is found in Seymour Inlet, Central Coast, an area only visited by tug traffic and two detected oil spills. The second highest positive residual (ID 20735) is located in the vicinity of Port McNeill, in Vancouver Island, with two small harbours, relatively low shipping traffic and four detected oil spills. Vancouver has one of the largest over-predicted values and significant negative residual (ID 45198), followed by Pender

Harbour (ID 39534). These sites enclose a large density of marinas and public docks (greater than 10) with one or no oil spills detected.

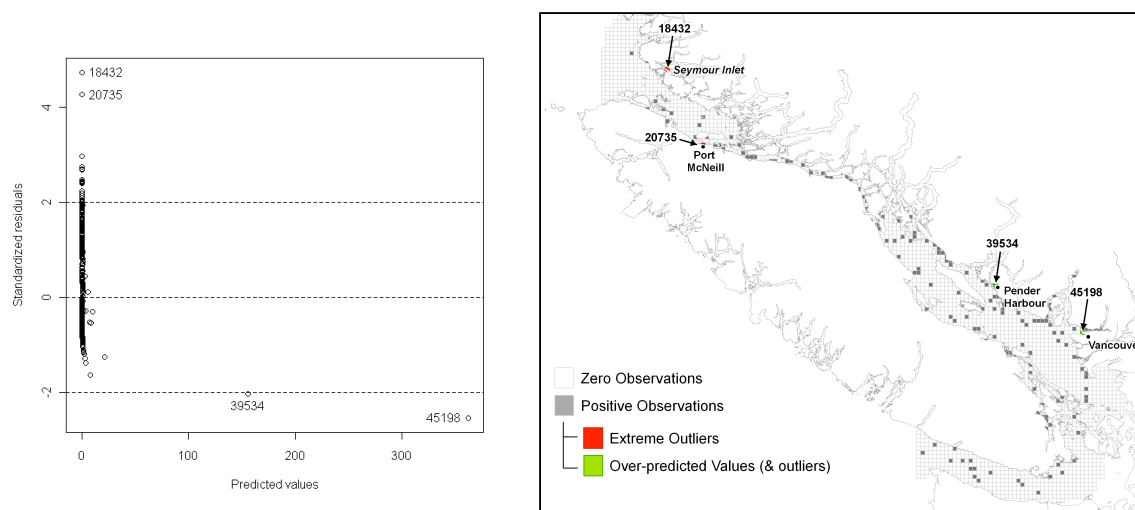


Figure 2.13 Standardized residuals against predicted values (left). Plotted horizontal dashed lines represent the 95% confidence intervals. Map (right) highlighting the identified extreme significant outliers and over-predicted values of the B.NB model.

None of the identified extreme outliers or over-predicted values in Areas A and B were removed from the modeling procedure as we assume all observations to be “true observations” which belong to population under study.

Global tests on spatial dependency indicated absence of spatial autocorrelation for model residuals in Area A (Moran’s $I = 0.02$, $p = 0.06$) and in Area B (Moran’s $I = 0.015$, $p = 0.05$). However, representations of estimated local Moran’s I_i for both regions reveal the presence of local clusters in model residuals. In Area A (Figures 2.14), two clusters of underestimated oil spill counts (positive residuals) are found in Nootka Sound and within Barkley Sound. A large cluster of overestimated oil spill counts (negative residuals) is seen off the south-west coast and along the north-west tip of Vancouver Island. Clusters of

overestimated oil spill counts dominate Area B (Figure 2.15). The most outstanding pockets of underestimated oil spill counts are found near Comox and Powell River.

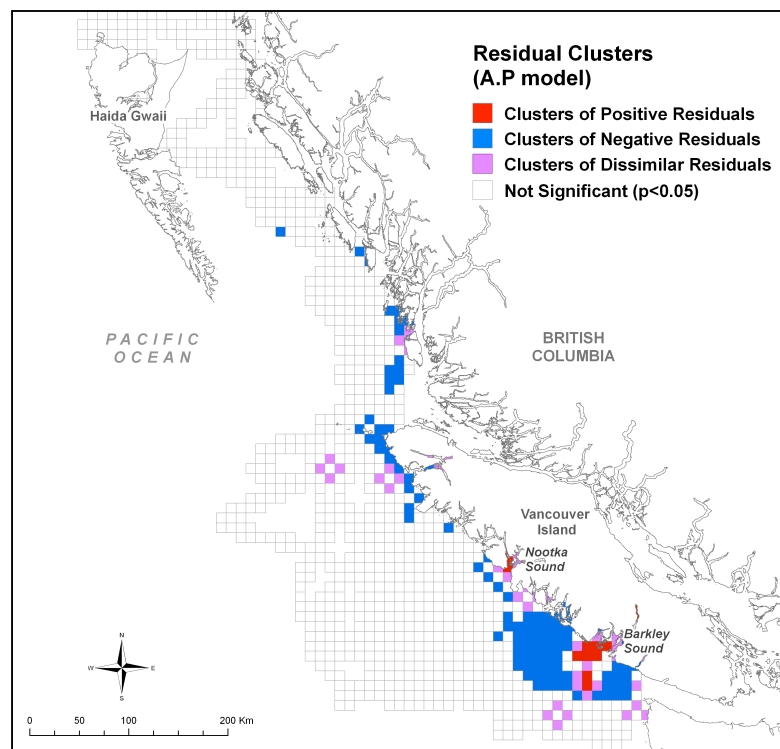


Figure 2.14 Residual clusters map for Area A, based on the estimation of local Moran's I_i and applying a significance filter of $p \leq 0.05$.

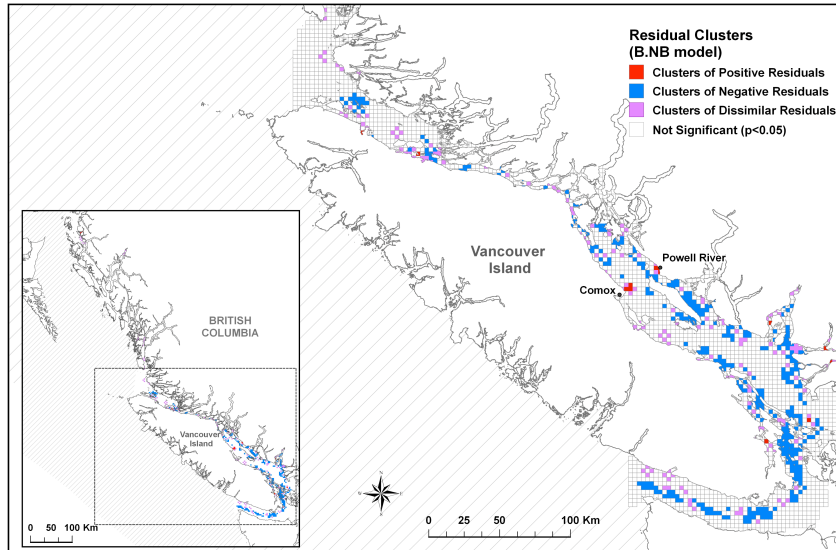


Figure 2.15 Residual clusters map for Area B, based on the estimation of local Moran's I_i and applying a significance filter of $p \leq 0.05$.

2.3.5 Model predictions

The expected number of oil spill at each location (or cell) in Area A is given by substituting the coefficients provided in model A.P in equation 2.3 as follows:

$$Exp(\text{Spills in Area A}) = \mu_{spill} = e^{-8.04 - 0.50(\log(disshorkm)) + 0.24(count_sh) + of\ f\ st(\log(ef))} \quad (3.1)$$

Mapped predictions show that the highest number of expected oil spills (Figure 2.16) are found at the entrance of the Juan de Fuca Strait, near Ucluelet, Port Alberni and Tofino, the Scott Islands (north of Vancouver Island) and Hakai Passage in Queen Charlotte Sound (Central Coast). The remaining areas contain less than 0.5 expected number of oil spills.

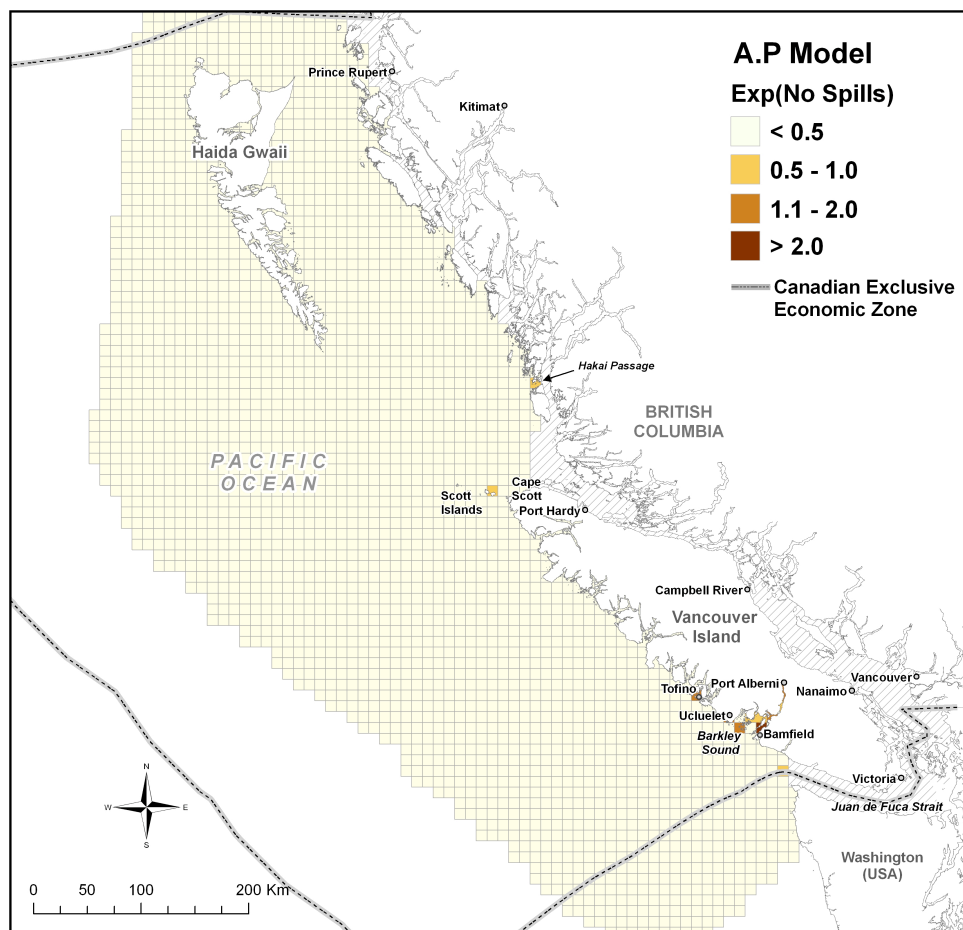


Figure 2.16 Predicted number of oil spill in Area A (offshore) estimated from model A.P and based on data collected by NASP from 1998 to 2007.

The probability of oil spill occurrences in offshore areas is given by the formula:

$$\Pr(\text{Spills in Area A}) = \pi_{spill} = \frac{e^{-7.92 - 0.6(\log(\text{disshorkm})) + 0.33(\text{count_sh}) + \text{offst}(\log(\text{eff}))\text{rt}}}{1 + e^{-7.92 - 0.6(\log(\text{disshorkm})) + 0.33(\text{count_sh}) + \text{offst}(\log(\text{eff}))\text{rt}}}$$

The highest probabilities of oil spill occurrence (Figure 2.17) are found at: the entrance of the Juan de Fuca Strait; in Barkley Sound, near Port Alberni, Ucluelet and Bamfield; in Tofino; off Cape Scott and Scott Islands (north tip of Vancouver Island); near Hakai Passage (Central Coast); and in the north coast of Haida Gwaii. The majority of

cells in Area A show less than one percent of probabilities of oil spill occurrences over a period of 10 years.

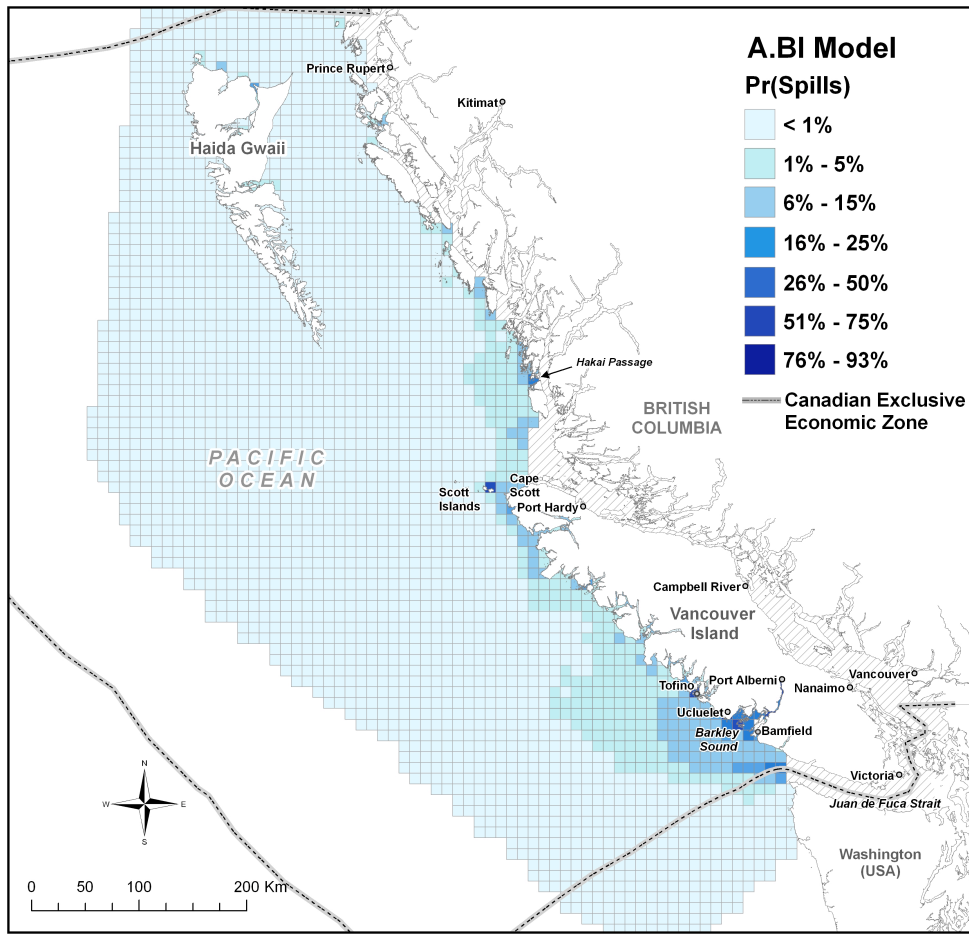


Figure 2.17 Probability of oil spill occurrences in Area A (offshore) estimated from model A.BI and based on data collected by NASP from 1998 to 2007.

In Area B, the expected number of oil spills is estimated by the coefficients provided in B.NB and function:

$$Exp(\text{Spills in Area B}) = \mu_{spill} = e^{-8.24 - 0.2(\text{carr_mean}) + 0.4(\log(\text{tug_mean})) + 1.22(\text{oil_mean}) - 0.47(\log(\text{disshorkm})) + 0.53(\text{density_sh}) + \text{of f s}(\log(\text{ef f})\text{ort})}$$

Figure 2.18 depicts estimated oil spill abundances in Area B. The highest number of predicted oil spill counts is found near some of the most important population centers in BC, including Campbell River, Nanaimo, Sidney, Victoria and Vancouver and along the Johnstone Strait. Most of Area B contains expected values less than 0.5 oil spill counts.

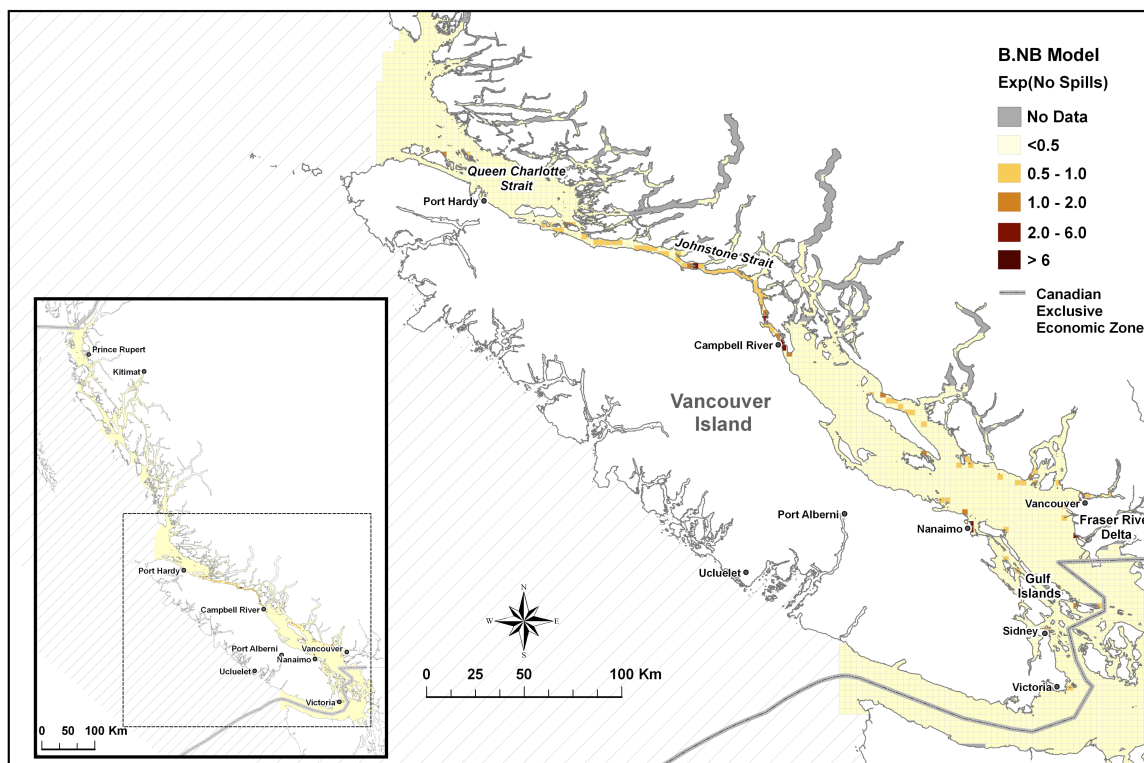


Figure 2.18 Predicted number of oil spill in Area B (inshore) estimated from model B.NB and based on data collected by NASP from 1998 to 2007.

The probability of oil spill occurrence along the inshore waters of BC is defined by model B.BI and equation:

$$\Pr(\text{Spill in Area B}) = \frac{e^{-8.38 - 0.18(\text{carr_mean}) + 0.44(\log(\text{tug_mean})) + 1.14(\text{oil_mean}) - 0.42(\log(\text{disshorkm})) + 0.46(\text{density_sh}) + \text{of f s}(\log(\text{ef f})) \text{ort}}}{1 + e^{-8.38 - 0.18(\text{carr_mean}) + 0.44(\log(\text{tug_mean})) + 1.14(\text{oil_mean}) - 0.42(\log(\text{disshorkm})) + 0.46(\text{density_sh}) + \text{of f s}(\log(\text{ef f})) \text{ort}}}$$

The probability of oil spill occurrence (Figure 2.19) in Area B concentrates along the shores of the Fraser River delta, the Sunshine Coast (north of Vancouver) and within the narrow passage of Johnson Strait (between the mainland and Vancouver Island). We also find localized pollution hot spots (with more than 75% probability to encounter operational oil spill over 10 years) near the cities of Prince Rupert, Campbell River, Nanaimo, Pender Harbour, Vancouver, Sidney and Victoria. In the Gulf Islands and at the entrance to the Queen Charlotte Sound we also found locations with values between 26% and 50% probability of chronic oil pollution.

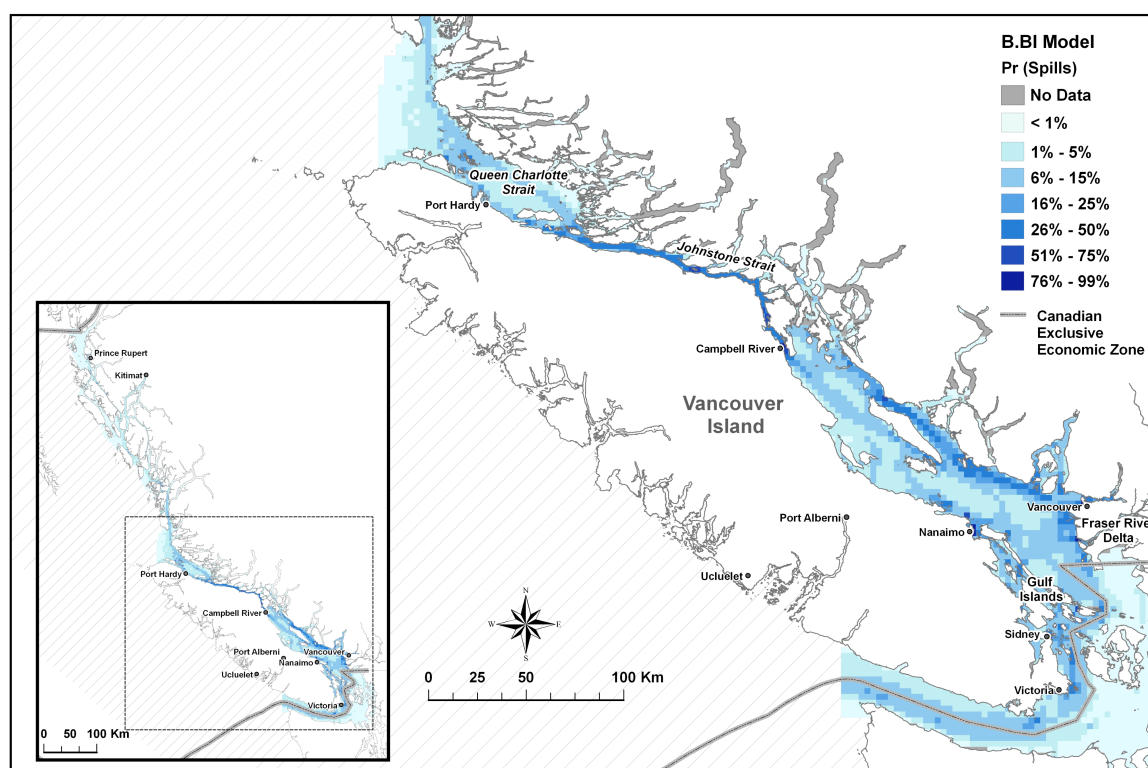


Figure 2.19 Probability of oil spill occurrences in Area B (inshore) estimated from model B.BI and based on data collected by NASP from 1998 to 2007.

2.4 Discussion

2.4.1 Model interpretation

In Area A, distance to shore and number of small craft harbours significantly influence both the probability of oil spill occurrence and the expected number of oil spills per cell. The farther we move from shore and away from ports and marinas the lower the probability of encountering oil spills. This is not surprising given the limited surveillance effort in offshore areas, which constrains the number of positive oil spill observations, and the fact that in these areas shipping traffic spreads out making it even more difficult to concentrate NASP efforts to a particular region.

The largest oil pollution hot spot is found in Barkley Sound. This area contains the largest concentration of small harbours (e.g., marinas, public docks and fishing harbours) along the west coast of Vancouver Island. In addition, Barkley Sound is the main navigation corridor for vessels, primarily for tug boats and fishing vessels, based in Port Albeni and Ucluelet, increasing the risk of chronic oil pollution in this area.

The presence of isolated oil spills in offshore areas (highlighted as extreme outliers in the residual plot for Area A) should be taken as a warning sign that illicit discharges may be more frequent than observed. In fact, observations of oil slick look-alikes gathered by the Integrated Satellite Tracking of Pollution or I-STOP program, a complementary enforcement tool to monitor illicit oil pollution and operating across Canada since 2006 (Environment Canada, 2006), shows that illicit oil spilling in areas away from shore may be more common than predicted based on NASP observations. Between 2004 and 2007, 13 oil spills were detected by NASP while a total of 43 (not verified) oil spills were detected with I-STOP during the same time period and in Area A. This underscores the

importance of NASP overflights to verify and monitor the actual degree of oil pollution from vessel operations in offshore areas.

More oil spills were detected in Canada's Pacific inshore waters (Area B; 268 spills) than in offshore waters (Area A; 62 spills). This permitted the convergence of a more complex model by fitting a higher number of significant predictors. In this region, the probability of oil spill occurrence is associated with density of small craft harbours, proximity to shore and intensity of carriers, tug boats and oil tankers.

As observed in other parts of the world, such as in the North Sea and the Mediterranean Sea (Ferraro et al., 2009), the probability of encountering oil spill is directly associated with major navigation lanes, commonly used by large cargo and tanker vessels. In Canada's west coast, vessels utilizing these lanes and positively related with oil spill counts are oil tanker and tug boat traffic, whereas carrier vessels are negatively related. In addition, many oil spills were detected near the coast coinciding with tug boats routes. These vessels prefer to travel along the coastline and using the protection of inside passages, such as the Johnstone Strait. Even though tug boats are the most abundant and wide-spread vessel type in coastal waters, there are other vessels, such as small recreational and fishing boats, present in these regions. Therefore, these vessels could also be responsible for these spills; however, currently there are no available records regarding about the distribution and intensity of these activities to be incorporated into the model.

In Area B, as in Area A, the probability of oil spill occurrences and relative abundance is inversely related with distance to shore but positively related with density of small harbours, meaning that there is a higher probability of oil spill occurrence in areas closer to shore and with higher number of marinas and other coastal facilities. Oil

pollution hot spots are found in the vicinity of major commercial and tourist centers, including Prince Rupert, Campbell River, Nanaimo, Pender Harbour, Vancouver, the Fraser Delta and Sidney, which are characterized by containing high density of marinas, public and private docks. Chronic oil pollution is known to be associated with these facilities because most routine vessel operations (such as repairs and fueling) are carried out at these sites (Georgia Strait Alliance, 2009; Molnar & Koshure, 2009). Further, we expect recreational boating to be one of the sectors that contributes the most to chronic oil pollution since more than 50% of the small harbours in BC are used for recreational activities (e.g., marinas and yacht clubs), while the rest are for commercial activities such as fishing and ferry docking. It is important to note, however, that coastal areas may also be more prone to oil discharges look-alikes with the presence of weed beds and floating debris.

One particular complexity of modeling oil spill occurrences in the west coast of Canada is the large proportion of observed zeros (more than the 80%) in Areas A and B. ZI models and ZA models, specifically designed to deal with the excess of zeros in count data, were explored and compared with standard GLMs. This approach is suggested since complex modelling techniques are not always the best approach to account for a large frequency of zeros (Warton, 2005). Our results indicate that ZI models and GLMs performances were comparable, whereas the ZA models showed the poorest fit. Standard GLMs were selected to estimate oil spill abundances in Areas A and B because they were easier to interpret and simpler models (i.e., more parsimonious). Additionally, our inability to properly model false zeros likely influenced the selection of these models.

The explanatory power of the best-fit models for Areas A and B is limited by the data provided. In the next section we provide some thoughts on how the presented regression models can be improved in future analysis and list some of the limitations that need to be taken into account when interpreting the modeling results.

2.4.1 Model biases and limitations

The inclusion of effort as an offset variable was essential to reduce sampling biases towards areas that have been visited by NASP flights and to adequately compare between sample units of oil spill counts. The probability of observing an oil spill, however, not only depends on the presence of surveillance flights but also on other factors, such as weather conditions (wind, fog), the sea-state (waves, currents), the altitude of the aircraft and the ability of the observer to discriminate oil spills from look-alikes (Ferraro et al., 2009; Volckaert et al., 2000). Detectability is described as a “nuisance” parameter and one of the main sources of uncertainty. Not accounting for imperfect detection may lead to misleading conclusions in the interpretation of predicted oil spill abundances and occurrence probabilities (MacKenzie et al., 2006). For example, oil spill occurring in marinas may be easier to detect than oil spills occurring in offshore areas because of the localized and sheltered conditions within marinas; therefore, the fact that we observe more oil spills near marinas does not mean that they occur more frequently there but that they are easier to detect. Furthermore, the inability to detect an oil spill when in fact it is present also leads to the presence of false zeros (Tyre et al., 2003). Probability of observing an oil spill and the probability of occurrence can be modelled simultaneously using ZI mixture models. These approaches are designed to model two processes: one that addresses

detectability (or the presence of false zeros) and the other that addresses the probability of occurrence or abundance of oil spills and the presence of true zeros (MacKenzie et al., 2002; Tyre et al., 2003). However, due to lack of information regarding factors affecting detectability (described above) we were not able to model for imperfect detection at this time.

Shipping traffic data contains a number of limitations and biases that are also important to mention in this discussion. First, no attempt was made to resolve vessel path anomalies detected within the original VTOSS database; however, by summarizing vessel movements in a grid cell we expect to have reduced the impact of these anomalies in our results. Second, the VTOSS database only contains shipping records of vessels larger than 20 m, and hence this study ignores the presence of small pleasure crafts and other small vessels when explaining oil spill occurrences. Lastly, the coverage of the different vessel tracking modes vary depending on the topography of the area (in the case of the radar tracking system), range (e.g. radar system in Tofino has a range of approx. 110 km) and mandatory reporting requirements (e.g., only vessels larger than 300 tonnes are obliged to use the advance reporting systems 96 and 24 hours before entering a VTS zone), which generate a number of data gaps across the study area. In summary, the shipping traffic data used represents the best currently available “minimum” estimates of shipping traffic densities rather than “true” estimates, and the differential between minimum estimates and true densities likely varies spatially.

The analysis of model residuals allowed us to explore the presence of outliers and investigate the presence of spatial heterogeneity. Areas with identified extreme outliers will require further investigation to reveal if these outliers are just anomalies or the

indicators of an unknown oil pollution problem. Further, despite the absence of global autocorrelation in model residuals, visual inspection of local Moran's I_i maps revealed clusters of similar residual values at a regional level. The presence of local autocorrelation in model residuals may be used as a diagnostic tool to assess the need of new predictors, the presence of model misspecifications and/or the inappropriateness of the sampling unit size and shape (Griffith, 1992). Former exploratory analysis showed the presence of spatial autocorrelation (both global and local) on model predictors (Serra-Sogas et al., 2008b), which could explain the presence of local spatial dependencies in the model residuals. Additionally, the processes driving the observed oil spill pattern are likely to be non-stationary or not constant over space. The relationship between potential sources of oil discharges and observations may vary spatially (e.g., depending on the location of the vessel in reference to shore, port or other vessels) and/or temporally (day vs. night time). Local spatial regression techniques, such as the geographically weighted regression (GWR) (Fotheringham, Brunson, & Charlton, 2002), could be used as complementary exploratory tools to assess the degree of spatial heterogeneity (non-stationarity) and spatial dependency in chronic oil pollution models.

Scale effects (or size of the sampling units) are another potential source of bias in the modelling process. Here, the cell size was selected based solely on the accuracy of the shipping data. However, a cell size optimal for one variable may not be optimal for the rest of the variables in the model, including the response variable (Guisan & Thuiller, 2005). For example, a larger cell size for Area B may be more appropriate based on the aggregated pattern of detected oil spills and the extent of the study area and it would have reduced the number of cells with zero observations. Understanding the scale of the process

driving the observed pattern is essential to avoid misleading interpretations in the modelled relationships (Fortin & Dale, 2005). One possible approach to test the effects of scale is to perform a number of sensitivity analysis by testing different cell sizes and determine the best configuration for all variables, as suggested by Jelinski and Wu (1996). Yet, this approach becomes computational intensive and quickly uneatable and, certainly, out of the scope of this study.

Oil spill observations by NASP flights from 1998 to 2007 were only possible during the daytime and consequently the predicted maps represent the day-time risk of oil pollution. Evidence from satellite monitoring suggests, however, that most of the illicit discharges are occurring at night to avoid being seen by pollution patrols or other vessels (Hampton, Kelly, & Carter, 2003). Consequently, oil spill abundances and patterns are likely to differ between the day and the night time, which adds a new level of uncertainty about the true rates of chronic oil pollution in Canada's Pacific region.

Finally, the predicted results are only based on the current shipping traffic conditions in BC. Variations in shipping patterns and intensities in the future may certainly change the picture from our model predictions. Planned developments in the north coast of British Columbia, which involve the opening of a new terminal for receiving crude oil in Kitimat, could increase the number of oil tanker movements in the area and hence increase the risk of operational oil discharges in these areas.

2.4.2 Future applications and recommendations

Despite the limitations and potential biases described in the last section, the described findings represent the first attempt to describe and predict oil spill occurrences and

abundances based on surveillance data, offering an opportunity to explore new applications.

The immediate application of this modelling exercise involves the identification of vulnerable marine areas to oil pollution by overlaying the predicted oil spill probabilities with the location of coastal and marine areas, protected under federal and provincial legislation and designated important bird areas (see Chapter 3). With this analysis we can highlight distinct areas in need of special monitoring and preservation from chronic oil pollution.

This study also provides a foundation for future modelling approaches and a baseline to compare with future oil spill data collected in Canada's Pacific region. Since early 2008, NASP began using a new aircraft, a Dash 8, which has a much greater range. This aircraft is equipped with a state-of-the-art suite of remote sensors and data display systems for on-site analysis (McGregor & Gautier, 2008). This new system facilitates the observation of oil slicks on the surface of the water as well as the accurate determination of the location and amount of oil spilled, in different weather conditions and at anytime of the day (McGregor & Gautier, 2008). The access to additional and more accurate data will allow exploring daily and seasonal differences in oil spill look-alikes, spatial variations in oil spill volumes, and the inclusion of factors to analyze detectability and appropriately model false zero observations.

A foreseen challenge of the new systems is the quantification of surveillance effort, which is mainly performed using Side-Looking Airborne Radar or SLAR. The observational window using SLAR depends on a number of factors, including weather conditions (e.g., wind and rain), bathymetry and range and flying altitude. In addition,

radar can be ineffective in areas with a complex geography and features that may generate false positives, for example freshwater inflows (Fingas & Brown, 2007). All these factors will have to be incorporated in order to obtain accurate estimations of surveillance effort, which are necessary to standardize oil spill observations.

2.5 Conclusion

This study explored different techniques for modelling an event that is rare and challenging to observe. A number of statistical models were fitted using oil spill observations and a set of predictor variables, while exploring the effects of the excess of zeros in oil spill count data. We generated a number of maps depicting the probability of occurrence and abundance of oil spills based on the best-fit models to the data provided. Our results allowed us to describe the patterns depicted by oil spill observations and to better understand the often complex underlying processes that generate them. This new information may be useful to NASP to improve the efficiency of the program and to other decision-makers to make informed decisions for the protection of Canada's Pacific marine and coastal regions. Finally, this study opens the door to future research initiatives aiming to better understand and ultimately reduce the chronic release of pollutants from shipping activities into the marine environment.

Chapter 3 : IDENTIFYING VULNERABLE CMPAs AND IBAs TO CHRONIC OIL POLLUTION IN THE WEST COAST OF CANADA

3.1 Introduction

The successful management and conservation of marine and coastal environments depend on recognizing and understanding the impact that human activities have on different marine environments (Ban & Alder, 2008).

A number of studies have employed mapping techniques to determine the location of vulnerable marine ecosystems to sea-based activities, with a regional (Hiddink, Jennings, & Kaiser, 2007) or global focus (Halpern et al., 2008), and by evaluating multiple threats (Ban & Alder, 2008) or by analyzing the spatial interaction of single species and threats (Zacharias & Gregr, 2005). Zacharias and Gregr (2005) define vulnerability as “the probability that a feature (i.e., ecosystem or single species) will be exposed to a stressor to which it is sensitive” while sensitivity is “the degree to which a marine feature respond to [a] stress [...] measured using one or more indicators (of species, communities, and habitats) that respond to one or more natural or anthropogenic stressors” (p.88). At the same time, vulnerability can be evaluated based on the spatial scale, frequency and the effects that a threat has on an ecosystem or species, the resistance of an ecosystem or species to a threat, and the resilience or recovery time of the ecosystem or species impacted (Halpern et al., 2008).

In this study, we employ Geographic Information Systems (GIS) analytical and mapping capabilities to identify marine areas potentially vulnerable to small but chronic levels of oil pollution from vessel operations in British Columbia (BC) based on the

location of potential sensitive sites. Evidence suggests that it is location and timing and not so much the amount of oil released that plays an important role when assessing the effects of oil pollution (Burger, 1993; National Academy of Sciences, 2003).

Chronic oil pollution from commercial and recreational vessels is known to be a significant threat to marine and coastal ecosystems worldwide (GESAMP, 2007). Impacts from chronic oil pollution have been estimated based on data provided by organized Beached Bird Surveys (BBS) (Camphuysen & Heubeck, 2001). In Atlantic Canada, an estimated 300,000 seabirds die every year from vessel-source chronic oil pollution based on data collected during BBSs (Pavlakakis et al., 2001; Wiese, 2002). Seabirds are the most conspicuous species to oil pollution (Wiese & Robertson, 2004). A single drop of oil can reduce the insulative properties of the bird's plumage, which can inevitably lead to hypothermia and starvation, and ultimately death. Not only can seabirds be dramatically impacted by oil but also many other species, ranging from primary consumers (e.g., sea urchins and clams) to species in top trophic levels (e.g., whales, seals, sea otters) (Islam & Tanaka, 2004; EPA, 1999).

Here, we used the estimated probability of oil spill occurrence from vessel operations over a period of 10 years (1998-2007) in Canada's Pacific region (see Chapter 2) to analyze the likelihood of oil pollution disturbance in Coastal and Marine Protected Areas (CMPAs) and Important Bird Areas (IBAs). We selected CMPAs and IBAs as indices of potential sensitive marine ecosystems to oil pollution in BC because they are geographically distinct areas and they contain valuable ecological features, some of which are particularly susceptible to oil pollution, such as seabirds. The term CMPA includes currently designated areas in BC that are protected under federal and provincial legislation

and regulations, while IBAs represent areas of international and national significance for the conservation of bird species, and biodiversity in general (IBA Canada website). By overlapping the locations of CMPAs and IBAs with the probability of oil pollution surface, we can identify sites more likely to be exposed and hence, more vulnerable to chronic oil pollution.

This study illustrates the utility of the modelling exercise presented in Chapter 2 as an approach for identifying CMPAs and IBAs in need of management action and to measure and assess the actual effects of marine vessel activities in these areas. In addition, our results can provide coastal managers and planners with information to determine the efficiency of marine protected areas for the conservation of marine and coastal ecosystems and species from chronic oil pollution.

3.2 Data and methods

3.2.1 Study area

The study area targets Canada's Pacific region and its boundaries are determined by the coincident extent of shipping data and NASP surveillance (Figure 3.1). For the purpose of data analysis and resolution, the study area was divided into two regions: offshore waters or Area A and inshore waters or Area B (for a description of the rationale see Chapter 2).

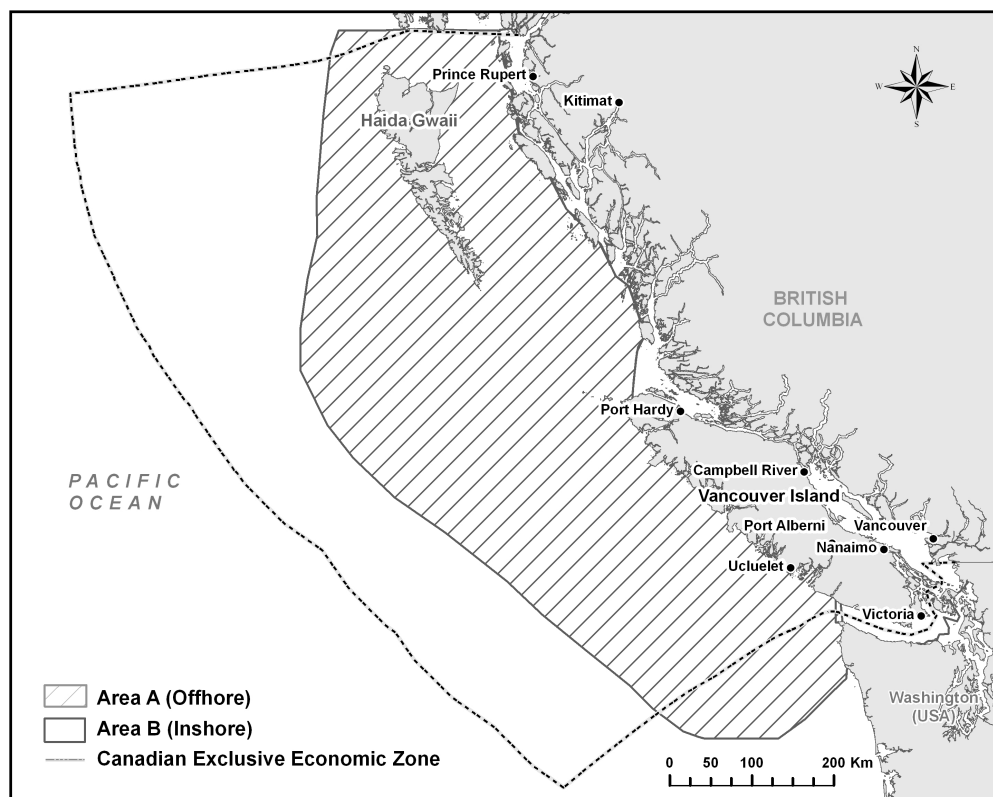


Figure 3.1 Study area with offshore (Area A) and inshore (Area B) regions.

3.2.2 Data

Coastal and Marine Protected Areas and Important Bird Areas

We used currently established CMPAs and IBAs found within the study area as indirect indicators of the presence of ecosystems and species potentially sensitive to chronic oil pollution.

CMPAs represent areas of intertidal terrain and continuous waters which enclose examples of marine diversity, unique habitats and ecosystems as well as features that are important culturally and historically (Canessa & Lunn, 2005). CMPAs are designated by a variety of agencies and have different conservation goals and legislative tools. For

example, there are CMPAs that warrant more attention to the conservation of marine habitats and wildlife (e.g., Ecological Reserves and National Wildlife Areas) while other CMPAs are managed not only for ecosystem protection but also for its recreational use. We used CMPAs conservation goals as a surrogate to distinguish between CMPAs that are more likely to contain species and ecosystems sensitive to chronic oil pollution. Thus, we grouped CMPAs in two groups (Table 3.1): CMPAs for species and habitat protection (total of 192) and CMPAs for representative protection, recreational use or sustainable resource extraction (total of 130).

A total of 165 federal CMPAs and 148 provincial CMPAs are distributed throughout BC's coastal region and included in the analysis. The Pacific Rim National Park Reserve, although represented as a unique park, was subdivided in three units (Long Beach, Broken Group Islands and West Coast Trail) to better capture the different habitats and uses contained in the park. CMPAs found outside of the study areas, such as Bowie Seamount and Endeavour Hydrovents Marine Protected Areas, were not included in the analysis.

Spatial information for National Wildlife Areas (NWA) and Migratory Bird Sanctuaries (MBS) was provided by the Canadian Wildlife Services (Environment Canada). The boundaries of the NWAs (from 2010) are a work-in-progress and will need to be updated with a finalized version. MBS limits were generated in 2008 and determined based on available base map data (TRIM 1:20000, marine charts and orthophotos). National Park Reserves boundaries were compiled from Parks Canada survey maps (2005). Information on Provincial Parks, Ecological Reserves, Wildlife Management Areas and Conservancies was obtained from GeoBC (Province of British Columbia) and generated in

2008 from cadastral data. Rockfish Conservation Areas 2005 boundaries were developed by the Department of Fisheries and Oceans and extracted from the CHS (Canadian Hydrographic Services) raster marine charts.

Table 3.1 Description of federal and provincial CMPA designations included in this study and classified based on their protection and conservation goals.

	Agency	Designations	Legislative Tools	Specific protection and conservation goals
CMPAs for species and habitat protection	Canadian Wildlife Services	Migratory Bird Sanctuaries (MBSs)	<i>Migratory Birds Act</i>	To protect coastal and marine habitats heavily used by birds for breeding, feeding, migration and overwintering.
		National Wildlife Areas (NWAs)	<i>Canadian Wildlife Act</i>	To protect and conserve marine and terrestrial habitats nationally or internationally significant for all wildlife, with special focus on migratory marine birds.
	Fisheries and Oceans Canada	Rockfish Conservation Areas (RCAs) *	<i>Fisheries Act</i>	To protect inshore rockfish from all mortality associated with recreational and commercial fisheries.
	BC Parks	Ecological Reserves (ER)	<i>Ecological Reserves Act</i>	To protect rare, endangered, or sensitive species or habitats.
		Wildlife Management Areas (WMA)	<i>BC Wildlife Act</i>	To conserve and manage areas of importance to fish and wildlife and to protect endangered or threatened species and their habitats, whether resident or migratory, of regional, national or global significance.
CMPAs for representative conservation, recreation or sustainable resource extraction	Parks Canada	National Park Reserves (NPRs)	<i>National Parks Act</i>	To protect and conserve for all time marine conservation areas of Canadian significance, and encourage public understanding, appreciation and enjoyment.
		National Marine Conservation Areas (NMCAs)	<i>Marine Conservation Areas Act</i>	
	BC Parks	Provincial Parks (PP)	<i>Park Act</i>	To protect representative examples of marine diversity and recreational and cultural heritage.
		Conservancies**	<i>Park Act</i>	To protect and preserve biological diversity and natural environments; social, ceremonial and cultural uses of first nations; and recreational values
			<i>Protected Areas of British Columbia Act</i>	

Source: Modified from Jamieson and Lessard (2000).

Note: *(DFO, 2007); **(BC Parks, n.p)

IBAs are discrete sites of international and national significance established for the conservation of bird species, but without legal protection (IBA Canada website). In Canada, the IBA Program is coordinated by Bird Studies Canada and Natural Canada, in conjunction with the organization BirdLife International. In Canada's Pacific region, there are a total of 84 IBAs, out of which only 60 were kept for further analysis because they have a marine component and are found in the study area. Polygon boundaries on Important Bird Areas were created and provided by Bird Studies Canada and Nature Canada (2004-2009).

Figure 3.2 shows the distribution of CMPAs grouped by protection goals and IBAs in the study area. The number and size of CMPAs and IBAs analyzed range widely (Table 3.2). Generally, CMPAs are smaller in size than IBAs. In Area A, IBAs have the highest representation (total of 36) and the largest average size ($\sim 363 \text{ km}^2$). In Area B, RCAs outnumber the rest of designations (total of 132), although IBAs continue to be the largest in average size ($\sim 70 \text{ km}^2$).

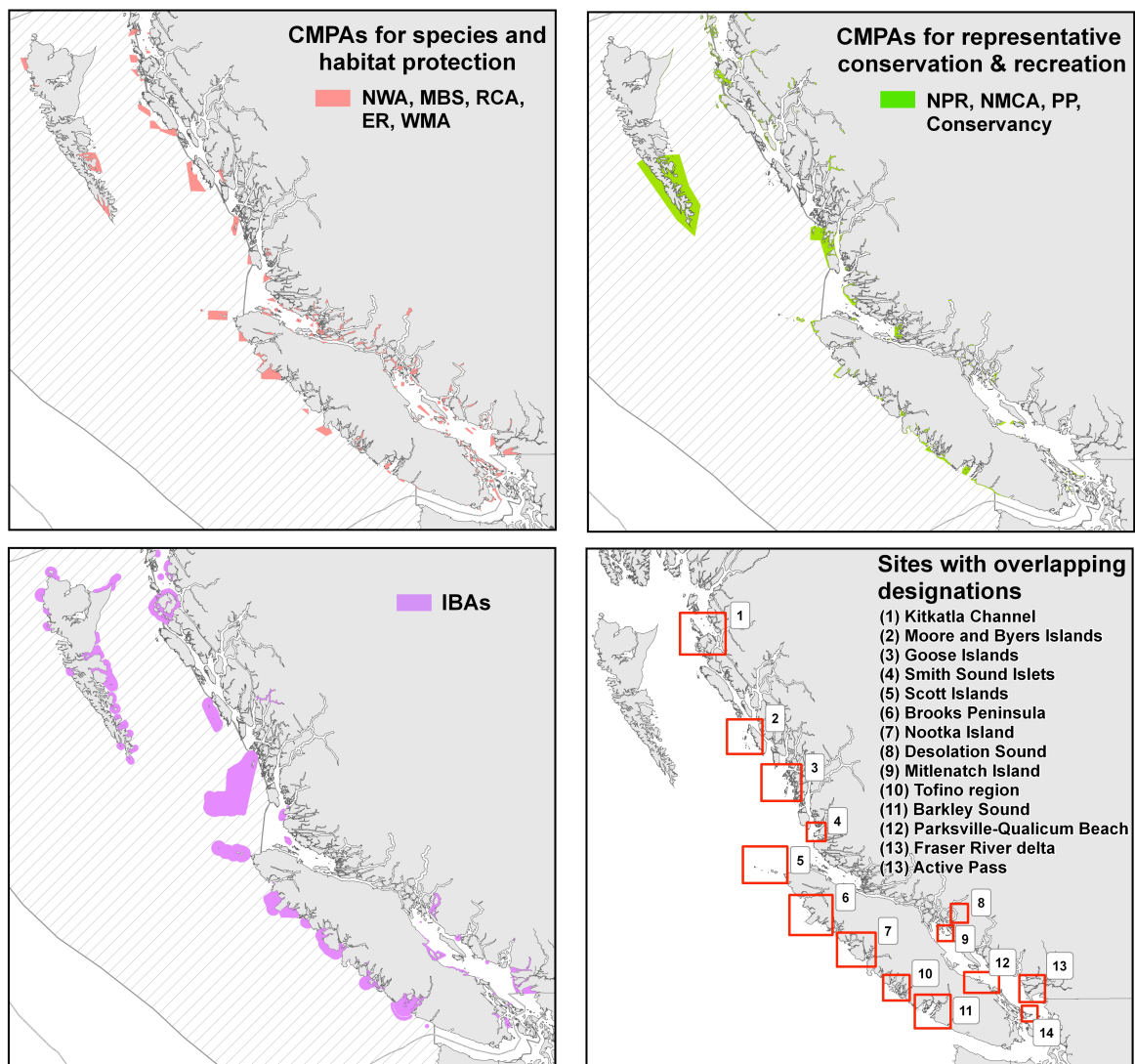


Figure 3.2 Distribution of CMPAs aimed for wilderness protection (top left), CMPAs designated for seascape protection and recreational use (top right), IBAs (bottom left), and areas with overlapped designations (bottom right), in the study area.

Table 3.2 Number and statistics of marine area (in km²) of CMPAs and IBA for the entire study area.

Designations	Count	Min.	Max.	Sum	Mean	SD
<i>CMPAs for species and habitat protection</i>						
Migratory Bird Sanctuaries	5	>1	18	23	5	7
National Wildlife Areas	2	>1	>1	>1	>1	>1
Rockfish Conservation Areas	155	>1	509	4,695	30	63
Ecological Reserves	23	>1	331	515	22	70
Wildlife Management Areas	7	1	116	223	32	43
<i>CMPAs for representative conservation, recreation or sustainable resource extraction</i>						
National Park Reserve	2	34	229	263	131	98
National Marine Conservation Area	1	-	-	3,384	-	-
Provincial Parks	86	>1	97	658	8	16
Wildlife Management Areas	7	1	116	223	32	43
Conservancies	34	>1	694	1,389	41	119
<i>Bird Studies Canada and Nature Canada</i>						
Important Bird Areas	60	>1	4,028	14,748	246	547

It is important to note that numerous coastal regions are protected under more than one designation. For example, some coastal areas within the Tofino region, Barkley Sound and Parksville-Qualicum Beach are protected under both federal and provincial legislation, as well as being considered IBAs. Other important areas protected under federal (RCA) and provincial legislation and designated as IBAs are Kitkatla Channel, Moore and Byers Islands, Goose Islands, Smith Sound Islets, Scott Islands, Brooks Peninsula region, Nootka Island, Desolation Sound, Mitlenatch Island, Fraser River delta and Active Pass. We consider these sites highly valuable ecological features because of the number of designations set in place for the protection of the ecosystems and/or species they contain. Figure 3.2 (bottom right) shows the distribution of these areas overlapping at least two different designations.

Marine ecosections

Federal and provincial CMPAs and IBAs were grouped geographically using established marine ecosections (Harper et al., 1993) to summarize oil spill exposures and facilitate their description. There are 12 marine ecosections in BC (Figure 3.3), and they are defined as areas that share common physiographic and oceanographic characteristics. Area A contains the majority of the ecosections with a total of seven, while Area B contains only five. Figure 3.4 depicts the distribution of total area protected under CMPAs designations and IBAs in each marine ecosection found in Areas A and B. Queen Charlotte Sound and the Vancouver Island Shelf marine ecosections contain the largest amount of area protected under federal and provincial legislation (CMPAs) as well as IBAs. The marine ecosections for the Province of British Columbia, and obtained from GeoBC.

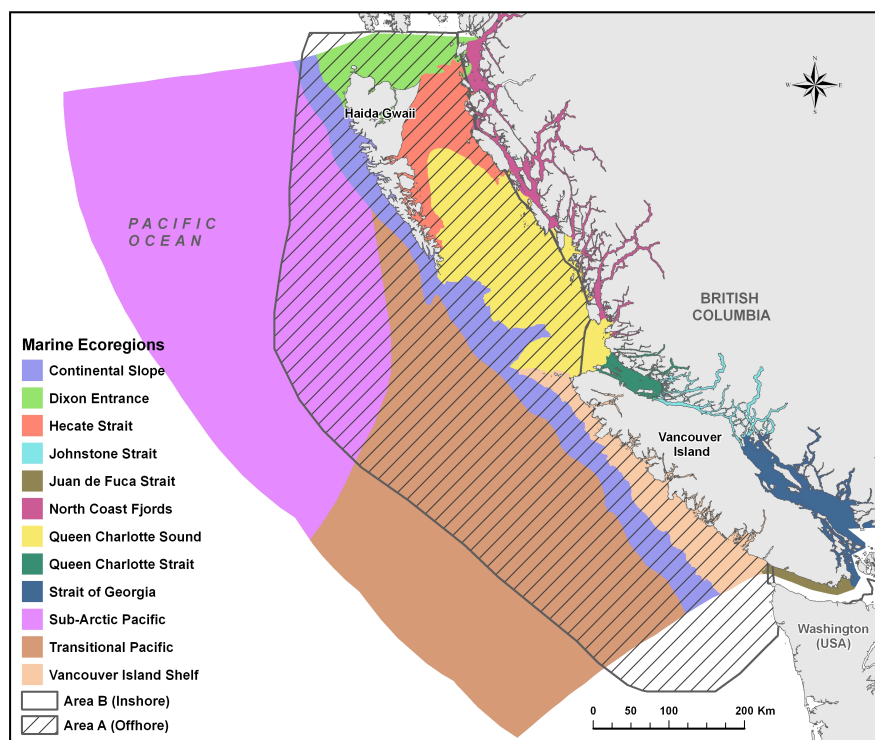


Figure 3.3 Marine ecosections in Canada's West Coast.

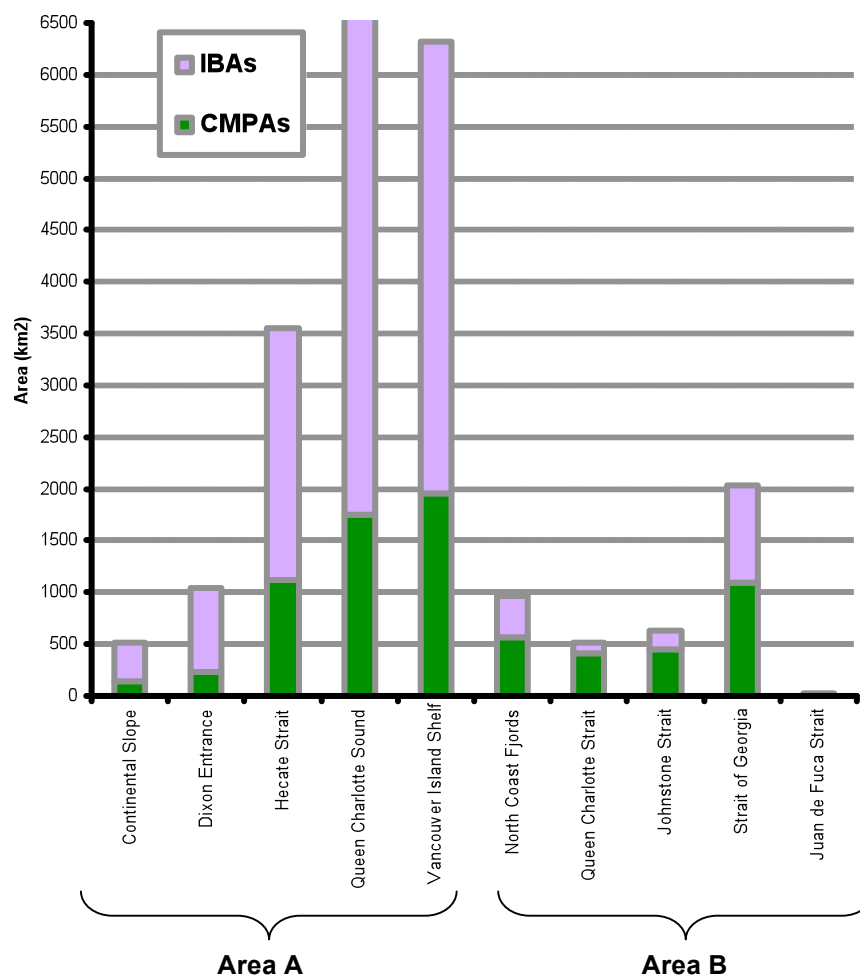


Figure 3.4 Distribution of total area designated as IBA and protected under federal and provincial CMPAs by marine ecosystems in Areas A and B.

Chronic oil pollution predicted surface

The probability of oil spill occurrences for BC's offshore areas (Area A) and inshore waters (Area B) were extracted from modelling 10 years (from 1998 to 2007) of oil spill observational data collected by the National Aerial Surveillance Program (NASP) and using a set of predictors, including: shipping traffic data, distance to shore and density of small craft harbours (see Chapter 2 for a description of the methodology used). Figure 3.5 depicts the distribution of estimated probability of chronic oil pollution in both study Areas

A and B over a period of 10 years. The predicted probability of oil spill occurrence will be used to define the likelihood of exposure of CMPAs and IBAs to small-scale vessel-source oil spills in the study area.

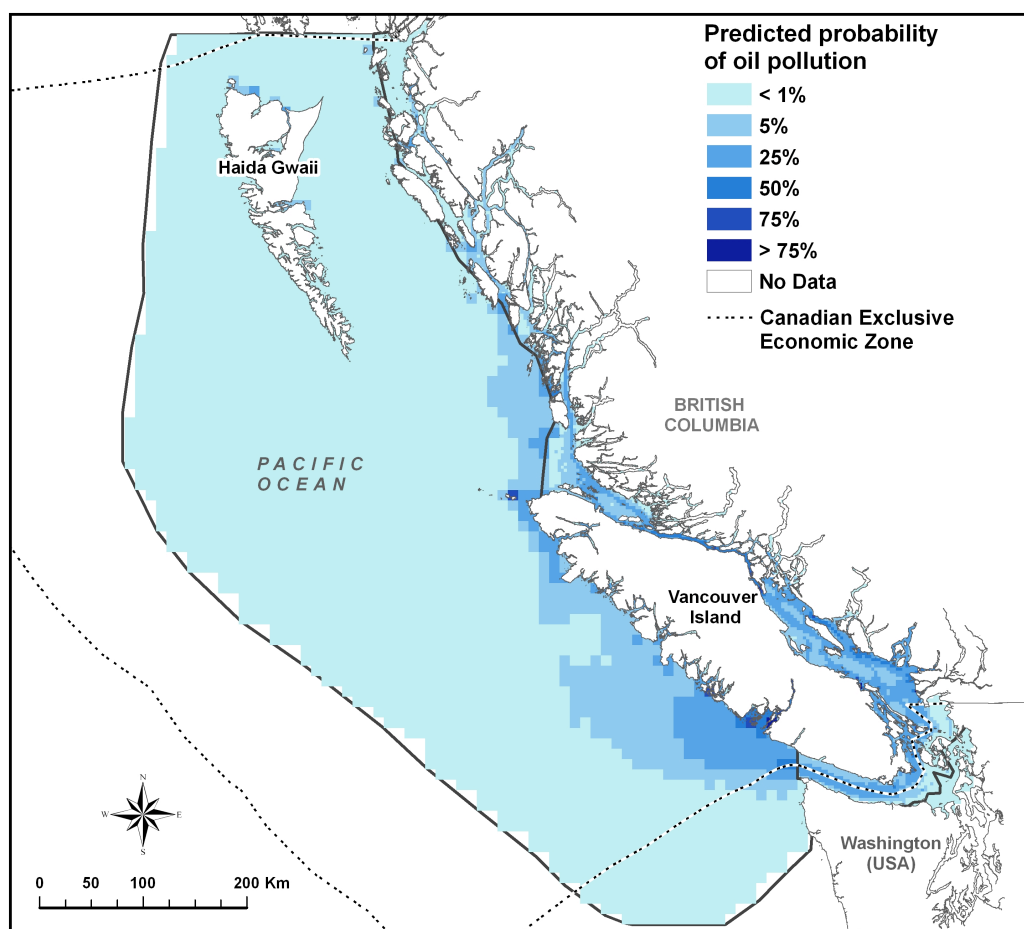


Figure 3.5 Distribution of oil spill probability of occurrence (in percentage) based on data collected by NASP from 1998 to 2007 in the study area (Areas A and B).

3.2.3 Map overlay procedure

To determine the likelihood of exposure of designated CMPAs and IBAs to oil pollution, the probability of oil spills surface in Areas A and B was spatially related with the

distribution of CMPAs and IBAs. First, CMPAs and IBAs were clipped with the coastline as only the marine area is of interest in this study. Second, the probability surface was converted from a vector format to a raster format with a cell size of 50 by 50 m. This cell size was selected to be able to capture the smallest CMPAs or IBAs within the study area. Third, the Zonal Statistics tool from ArcMap™ v.9.3 was used. This function allows the calculation of a number of descriptive statistics of the raster oil spill probability dataset that falls within the zone of each zonal polygon (i.e. CMPAs and IBAs). This tool automatically transforms the zonal polygons into raster using the same cell size as the probability surface. Then, a number of descriptive statistics are estimated based on the number of cells from the probability raster that falls within each CMPA or IBA. Here, the mean was selected to represent the average probability of exposure (i.e., exposure index) and the standard deviation as a measure of exposure variability (or statistical dispersion) within each site. A low standard deviation indicates that the data points tend to be very close to the mean, a high standard deviation indicates that the data is spread out over a large range of values (Streiner, 1996). Figure 3.6 shows a diagram of the different steps performed by the Zonal Statistics tools in ArcMap.

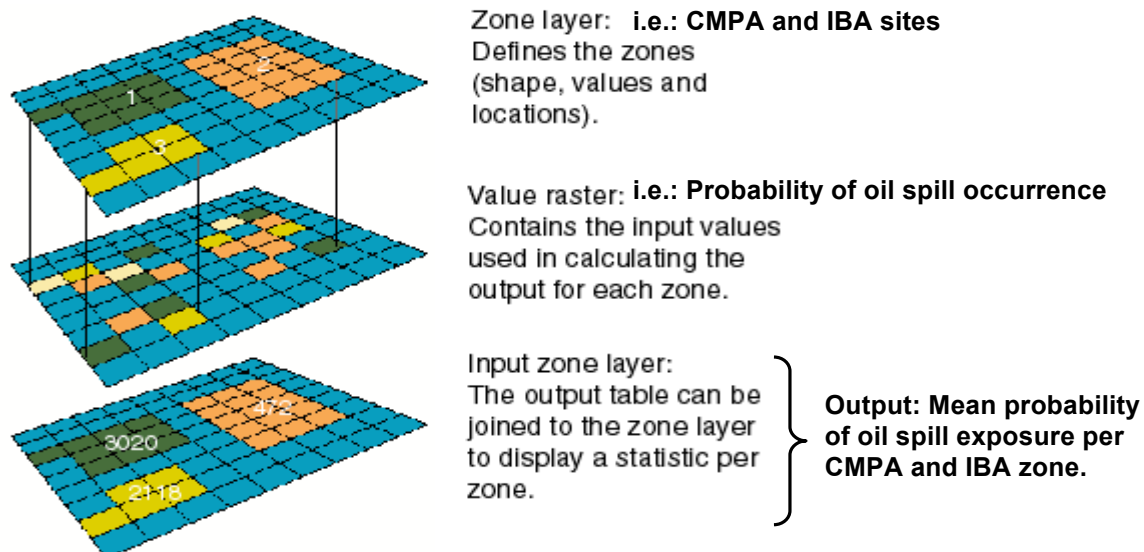


Figure 3.6 Zonal statistics procedure (figure modified from ArcGIS v.9.3 help menu).

For CMPAs and IBAs that overlapped with both Areas A and B, the final mean (and standard deviation) values were obtained by weighing each mean value by the number of cells found at each site (i.e., CMPA or IBA), which is an indication of the size of the area. Formally, the weighted mean (M) and the weighted standard deviation (SD) are estimated by:

$$M = \sum_{i=1}^n w_i x_i / \sum_{i=1}^n w_i \quad \text{and} \quad SD = \sqrt{\sum_{i=1}^n w_i (x_i - M)^2 / \sum_{i=1}^n w_i} \quad 3.1$$

where x_i is the mean value at each site and w_i is the weight or number of cells in each site.

To facilitate the interpretation of the results, the (weighted) average probability of exposure was extracted per designation group and marine ecosection. The same equations 3.1 were used to estimate the weighted means and standard deviations. Maps were generated for federal CMPAs, provincial CMPAs and IBAs separately, highlighting those sites with a mean probability of exposure greater than 27%. The 27% mean probability of

exposure was used as a cut-off value to determine high probability of exposure, and it was determined by the upper limit of the non-outlier range (upper whisker of a box plot, Figure 3.7) of the mean exposure distribution values polled from all sites.

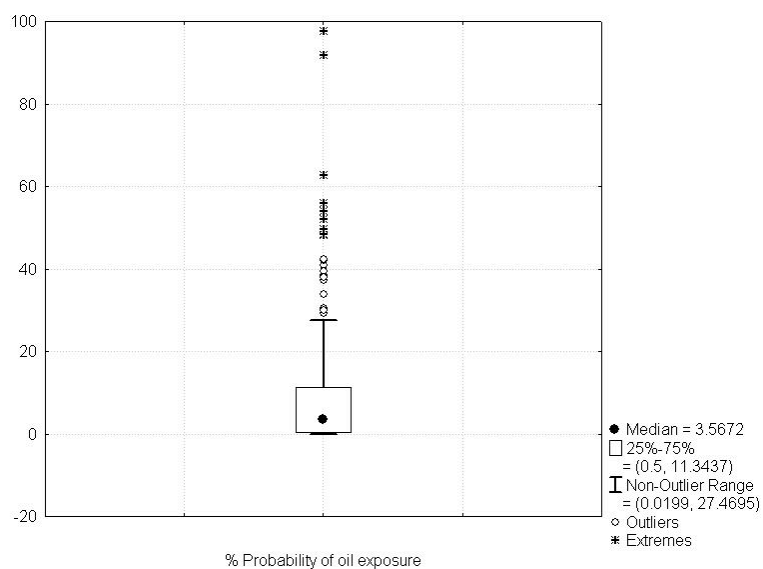


Figure 3.7 Box-plot of mean oil spill probability values extracted from each CMPA and IBA.

3.3 Results

The CMPA areas showing the highest mean probability of oil spill exposure (Table 3.3) are Wildlife Management Areas in the Vancouver Island Shelf marine ecosection with an average value of 52% (SD= 24), followed by the Migratory Bird Sanctuaries (M=32% , SD=39) of the Strait of Georgia and National Wildlife Areas (M=29%, SD=9) in Johnstone Strait. The marine ecosection with the highest scores are the Strait of Georgia marine ecosection followed by the Vancouver Island Shelf, while the Continental Slope and Dixon Entrance show the lowest score values.

Table 3.3 Mean probability of oil spill exposure (in percentage and over 10 years) for each CMPA designation by IBAs by marine ecosystem in which they occur.

Designations	Ecosystem									
	Continental Slope	Dixon Entrance	Hecate Strait	Johnstone Strait	Juan de Fuca Strait	North Coast Fjords	Queen Charlotte Sound	Queen Charlotte Strait	Strait of Georgia	Vancouver Island Shelf
<i>CMPAs for wilderness protection</i>										
MBS	-	-	-	-	1(0)	-	-	-	32(18)	-
NWA	-	-	-	-	-	-	-	-	29(9)	-
RCA	0(0)	0(0)	0(0)	6(9)	7(7)	0(0)	25(57)	7(6)	11(8)	10(9)
ER	0(0)	0(1)		25(0)	24(42)		1(0)	9(0)	18(16)	4(3)
WMA	-	-	-	-	-	-	-	3(1)	10(11)	52(24)
<i>CMPAs for seascape conservation, recreation, or sustainable resource extraction</i>										
NPR	-	-	-	-	-	-	-	-	16(9)	25(8)
NMCA	0(0)		0(0)				0(0)			
PP	-	0(0)		21(18)	1(1)	1(1)	1(0)	4(5)	9(8)	12(9)
Conservancy	-	0(0)	1(0)	0(0)	-	1(2)	7(2)	1(2)	-	-
<i>NGO (wildlife conservation)</i>										
IBA	0(0)	1(1)	1(1)	1(1)	-	0(0)	1(1)	7(2)	11(6)	12(12)

Note: The final average was obtained by weighting the mean value at each site by the number of cells from the raster surface falling at each site. In parenthesis weighted standard deviations. Values in bold indicate probabilities of exposure greater than 27% over 10 years. The dash (-) indicates absence of CMPAs or IBA in these ecosystems. **Acronyms description:** MBS (Migratory Bird Sanctuaries); NWA (National Wildlife Areas); RCA (Rockfish Conservation Area); ER (Ecological Reserve); WMA (Wildlife Management Area); NPR (National Park Reserve); NMCA (National Marine Conservation Area); PP (Provincial Park); IBA (Important Bird Areas)

A closer analysis of the federal CMPAs (Figure 3.8) indicates that a total of 13 areas contain a mean probability of oil spill exposure greater than 27% over 10 years. RCAs represent the largest group with nine different sites, out of which three have the highest probabilities. These are: Deepwater Bay (M=55%, SD=0), near Campbell River; the Broken Group Islands (M=50%, SD=10); and Thormanby Island (M=48%, SD=9). The area of the Broken Group Islands designated as part of the Pacific Rim NPR contains a mean probability of 45% (SD=12). Finally, there are two MBS, Shoal Harbour (M=62%, SD=22) and George C. Reifel (M=30%, SD=21), and one NWA, Alaksen (M=42%, SD=31), with a mean risk of exposure greater than 27% over 10 years.

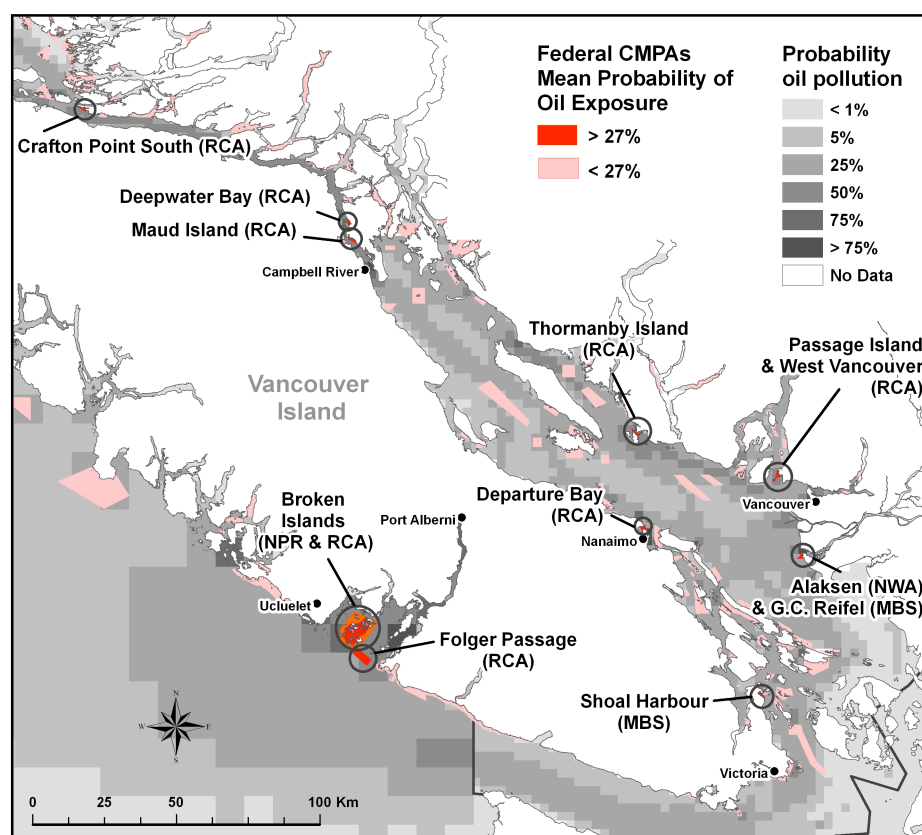


Figure 3.8 Federal CMPAs overlaid with the probability of oil spill occurrence over a period of 10 years in Canada's West Coast.

Two Provincial Parks, Garden Bay Marine Park (Pender Harbour) and Newcastle Island Marine Park (Nanaimo) show the highest mean probability values (greater than 91%) for all the CMPAs (federal and provincial) analyzed in this study (Figure 3.9). Three other Provincial Parks contain more than 27% mean probability of oil exposure: Lanz and Cox Islands (M=34%, SD=24), from the Scott Islands group; Rock Bay (M=39%, SD=9), in Johnstone Strait; Smuggler Cove (M=54%, SD=0), near Halfmoon Bay; and the Ten Mile and Oak Bay Islands (M=27%, SD=0), near Victoria. The Ecological Reserve in Baeria Rocks (M=48%, SD=0), in Barkley Sound, and the Wildlife Management Area of Tofino Mudflats (M=52%, SD=24) show the highest vulnerability to chronic oil pollution of their categories.

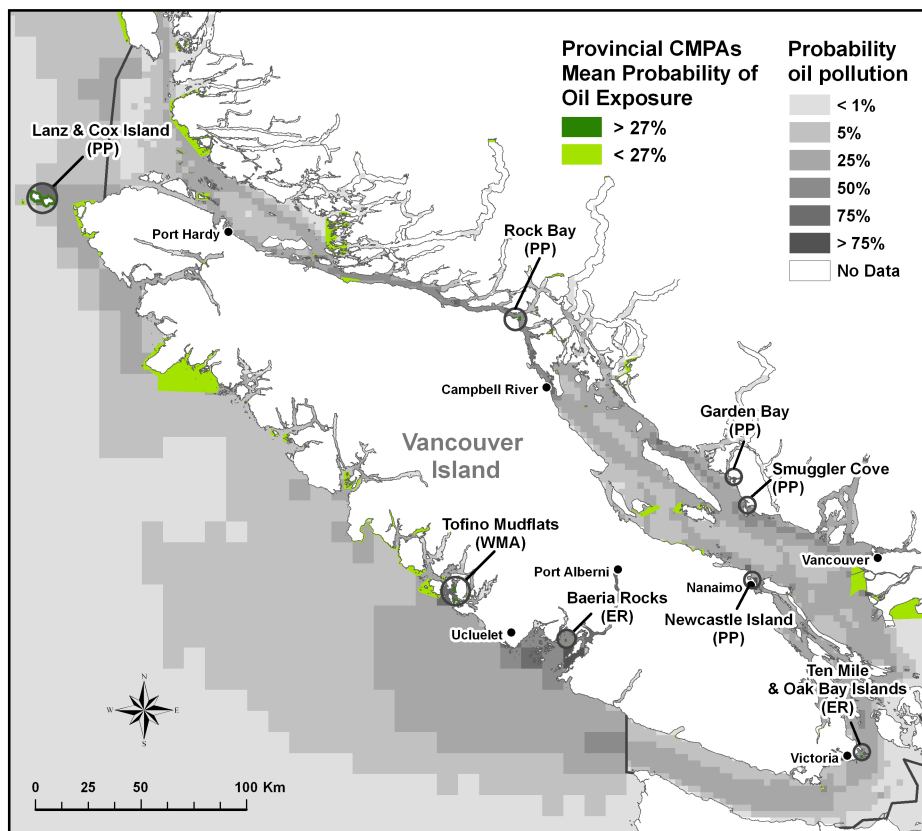


Figure 3.9 Provincial CMPAs overlaid with the probability of oil spill occurrence over a period of 10 years in Canada's West Coast.

The IBAs (Figure 3.10) that have mean probabilities of oil spill exposure greater than 27% are: the Tofino Mudflats ($M=45\%$, $SD=28$), the Chain Islets and Great Chain Islets ($M=41\%$, $SD=0$), near Victoria; the Barkley Sound ($M=37\%$, $SD= 22$); and White Islets and Wilson Creek ($M=28\%$, $SD= 8$), near Sechelt.

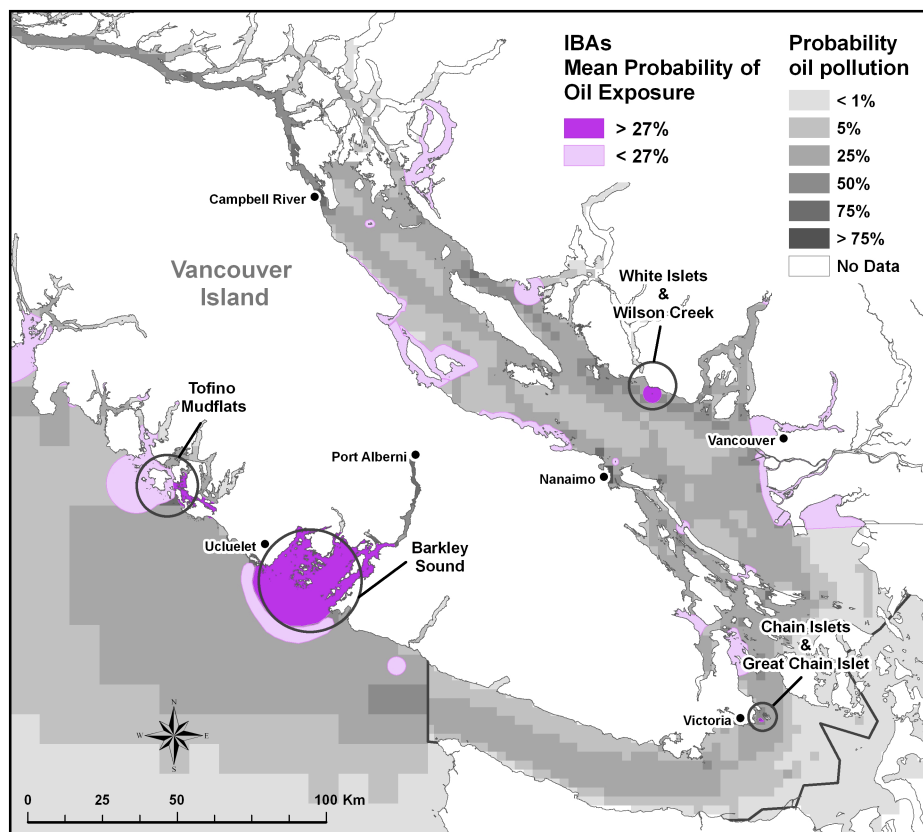


Figure 3.10 IBAs with the probability of oil spill occurrence over a period of 10 years in Canada's West Coast.

To conclude, there are four regions that stand out because they have relatively high mean probabilities of oil spill exposure for part or the entire region, they are protected under at least one CMPA designation in addition to being identified as IBAs, and they contain important populations of seabirds and other species highly sensitive to oil pollution. These regions are the Tofino Mudflats, the Barkley Sound, the Scott Islands and the Sturgeon and Robert Banks, in the Fraser River Delta.

Table 3.4 shows the mean probability of exposure for each CMPA or IBA site found in each region and the estimated overall (weighted) mean probability of exposure for each region. The Tofino Mudflats have the highest probability of exposure of 44 %

(SD=4), followed by Barkley Sound (M=38%, SD=2), Sturgeon and Robert Banks (M=22%, SD=3) and the Scott Islands (M=7%, SD=5). The Scott Islands contain the highest variability in exposure risk in comparison with the other regions, where the islands closer to shore protected as a Provincial Park are more susceptible to be exposed to chronic oil pollution (M=34%, SD=24) than the offshore islands designated as an Ecological Reserve (M=1%, SD=1). As a result, the average probability of exposure for the entire group of islands becomes appreciably lower at 7 % (SD=5).

Table 3.4 Mean probabilities of oil spill exposure (in percentage and over 10 years) per designation for the Tofino Mudflats, the Barkley Sound region, the Sturgeon and Robert Banks and the Scott Islands.

	NPR	RCA	ER	PP	WMA	NWA	MBS	IBA	Total Weighted Average
Tofino Mudflats	-	-	-	-	52(24)	-	-	42(28)	44(4)
Barkley Sound	44(12)	44(9)	48(0)	-	-	-	-	37(22)	38(2)
Sturgeon & Robert Banks	-	-	-	-	23 (9)	42 (31)	30 (21)	20 (18)	22(3)
Scott Islands	-	12(21)	1(1)	34(24)	-	-	-	5(13)	7(5)

Note: The final average was obtained by weighting the mean value at each site by the number of cells from the raster surface falling at each site. In parenthesis weighted standard deviations. The dash (-) indicates absence of CMPAs in these areas.

Acronyms description: MBS (Migratory Bird Sanctuaries); NWA (National Wildlife Areas); RCA (Rockfish Conservation Area); ER (Ecological Reserve); WMA(Wildlife Management Area); NPR (National Park Reserve); PP (Provincial Park); IBA (Important Bird Areas).

3.4 Discussion

3.4.1 Results discussion

For the first time in Canada, information on operational oil discharges is used to identify important coastal and marine environments, namely CMPAs and IBAs, that are most likely to be exposed (or vulnerable) to the aforementioned threat. Past studies in the east and west coast of Canada used vessel traffic information as proxies to evaluate the risks that these activities pose to marine wildlife, particularly seabirds (Lock, Brown, & Gerriets, 1994; O'Hara & Morgan, 2006). In these studies, “risk” was evaluated by the degree of overlap between traffic corridors and the distribution of vulnerable seabird species.

Conversely, this study uses CMPAs and IBAs, instead of species distributions, because they are easily identifiable and encompass broader habitats than just accounting for species. Our approach also allows a more direct and quantitative evaluation of oil spill vulnerability based on predicted probability that sensitive marine and coastal environments will be exposed to oil discharges from vessel operations.

These results show three levels of evaluation: from general to local. First, we provided a summary of the mean likelihood of exposure based on CMPA designations and IBAs and their location within marine ecosections. We also distinguish between CMPAs designated for species and habitat protection, which are likely to be more sensitive to oil pollution exposure, and CMPAs for representation conservation and recreation, which are likely to be less sensitive. Interestingly, the three CMPA designations that show the highest risk of oil exposure (MBS, NWA and WMA) belong to the first group, which are likely to be the most vulnerable and less resilient to the cumulative effects of chronic oil pollution.

Second, we focus our findings by mapping those individual sites (federal or provincial CMPAs or IBAs) with a higher risk of chronic oil pollution exposure over a period of 10 years, by using 27% mean probability of exposure as the cut-off value (see methods). The majority of highlighted federal CMPAs are RCAs, while PPs are the most abundant within the provincial CMPA designations. This is likely a result of these two designations being the most numerous and widespread designations of their group.

Finally, we focus our findings on four regions that were highlighted repeatedly in all previously generated maps: the Tofino Mudflats, Barkley Sound, Sturgeon and Robert Banks and the Scott Islands group. These regions share two key properties: (1) they contain significant ecological features, and (2) they show a relatively high probability of oil pollution exposure. The Tofino Mudflats (~ 42 km² of marine area), located in the most tourism driven town on the west coast of Vancouver Island, are currently designated as a WMA and an IBA. The Tofino Mudflats are the only area in the west coast of Vancouver Island known to support large numbers of shorebirds. They have been identified as critical habitat for migrating Western Sandpipers with an estimated 8.2% of the global population passing through during the fall season (IBA Canada website). In addition, the surrounding waters of the Tofino Mudflats are known to hold important populations of sealions, whale species and sea otters, which are also known to be highly sensitive to oil pollution (US EPA, 1999). The combination of its significant ecological value and high average probability of oil spill exposure (M=44%, SD=4) nominates this region as one of the most vulnerable to chronic oil pollution in BC. The potential sources of operational oil discharges in this area are the numerous small docks (public and private) located in Tofino, which provide shelter to small vessels working in the tourism industry or vessels

transporting goods to the small communities along the north-west coast of Vancouver Island.

The second focal region is Barkley Sound, an embayment ($\sim 700 \text{ km}^2$ of marine area) on the west coast of Vancouver Island. One of the prominent features of the sound is the Broken Group Islands, which form part of the Pacific Rim NPR and are composed of more than 100 islands and inlets. The Broken Group Islands are important bird areas hosting six species of global significance (i.e., Marbled Murrelet, Mew Gull, Surf Scoter, Surfbird, Western Grebe and Brandt's Cormorant) and two breeding species of national significance (i.e., Black Oystercatchers and Glaucous-winged Gull) (IBA Canada website). Barkley Sound also contains two RCAs, designated to protect inshore rockfish and lingcod populations, and one ER enclosing the inlet named Baeria Rocks. This area is prone to relatively high intensity of both commercial and recreational vessels. Also, along its coast we find a number of facilities (e.g, marinas and fuel docks) located in the communities of Ucluelet and Bamfield. Considering the numerous potential sources of chronic oil pollution and its estimated mean risk of oil exposure of the 33% ($SD=2$), the possibility of oil pollution disturbances may become one of the main conservation concerns in Barkley Sound in the near future.

Robert and Sturgeon Banks, situated at the entrance of the Fraser River delta, contain one of the richest and most important ecosystems for migrant and wintering waterbirds in Canada. The most numerous species are the Western Sandpiper, with over 500,000 birds during spring migration. In addition, it provides critical habitat for fish, including five species of salmon, herring, trout and bottom feeding fish species (IBA Canada website). Several sections of these areas are protected under federal and provincial

designations, all belonging to CMPAs for species and habitat protection. The entire delta is also designated as an Important Bird Area, and it is part of the Western Hemisphere Shorebird Reserve Network since 2005 (WHSRN website). Both banks are constantly exposed to impacts from different industrial (fishing and marine shipping) and recreational uses. Our findings show the Alaksen NWA and George C. Reifel MBS, in the Robert Bank, as the sites with the highest risk of chronic oil pollution exposure (more than 30%). This is not surprising given that in this area is particularly busy in terms of shipping traffic with the location of one the major ferry terminals in BC (Tsawwassen) and a coal port and container ship facility.

Lastly, we consider the Scott Islands as another region of concern. The Scott Islands are a total of five islands with an extension of 60 km² of marine area, located 10 km offshore of Cape Scott and extending westward for 46 km. These islands are protected under a patchwork of different designations. Lanz and Cox Islands, the two closest islands to shore, are designated as PP. The remaining islands, Sartine, Triangle and Beresford, are designated as ER. Four of five islands are protected as a RCA. Additionally, all five islands and their immediate waters are considered IBAs since they host the most important seabird colony in BC, with more than two million breeding seabirds, including species of global and regional significance. In addition, the surrounding marine waters support important seasonal populations of nonbreeding pelagic bird species (IBA Canada website). In fact, the Scott Islands are currently proposed as the first designated Marine Wildlife Area in the Pacific Region (Environment Canada, 2007), which will enhance their protection and conservation. Despite the notably low oil exposure index of 7% (SD=5) the

fragility of seabirds to oil pollution added to the large concentration of significant species, makes this region particularly vulnerable even at lower risk of oil exposure.

3.4.2 Advantages, limitations and research contributions

One of the main advantages of the method used to identify valuable and vulnerable marine areas to chronic oil pollution is the easy implementation and repeatability, assuming that both key datasets are available, that is: (1) an estimate of the risk of chronic oil pollution and (2) the distribution of habitats or organisms known to be sensitive to oil pollution.

However, there are a number of factors not considered in this analysis that should be taken into account to refine the assessment of vulnerability of CMPAs and IBAs.

First, there are likely seasonal differences in the probability of oil spill occurrence as well as in the sensitivity of CMPAs and IBAs. For example, because many seabird species are migratory the sensitivity of specific CMPAs and IBAs will change depending on the time of the year, and hence they could be weighted accordingly depending on the presence or absence of sensitive species (Zacharias & Gregr, 2005). In addition, the probability of oil spill occurrence is likely to change between the summer and winter seasons, given that the summer months in BC experience more vessel traffic coinciding with the tourism season, particularly of cruise ships, ferries (based on Marine Communication Traffic Services records) and marine pleasure crafts in general. The chemical and physical properties of the oil, volume of oil spilled, and the weather and sea conditions at the time of the spill are also factors that need to be considered when assessing the effects of oil spills onto the marine environment (Boehm & Page, 2007).

Another limitation is the potential bias of the modelled probabilities inherent in the assumption of perfect detection rates. This assumption is likely violated considering the difficulties of detecting an event that is rare and challenging to observe (see Chapter 2 for a more detailed discussion). These detection challenges add a certain degree of uncertainty that makes us believe that the actual rate of chronic oil pollution is much larger than observed and thereby the risk of oil exposure to marine and coastal environments.

Despite the described limitations and in the absence of more accurate model on chronic oil pollution in BC, we believe that our findings may still provide marine and coastal managers with practical conservation tools to assist them in the prioritization of coastal and marine vulnerable areas in need of urgent management measures. For example, highlighted CMPAs and IBAs could be selected to analyze the real impacts and assess the resilience of the habitats and species to chronic levels of oil pollution. The presented maps could be also used to assess the efficiency of CMPAs established in BC and evaluate the possibility of changing navigation practices (e.g., prohibit the entry of motorized vessels) within the protected limits and in adjacent areas. Beached Bird Survey (BBS) organizers could also employ the generated vulnerability maps to emphasize monitoring efforts to coastal areas showing a higher risk. Finally, our findings could add to studies analyzing and mapping the cumulative impacts that different anthropogenic stressors have on BC's marine environment, such as the study recently published by Ban et al. (2010).

3.4.3 Future research

Our findings only show relative probabilities of chronic oil pollution exposure in different sites across BC. Future field studies are required to measure the actual impacts of these

events on the natural environment. Studies around the world have shown evidence of the adverse effects of oil pollution on aquatic ecosystems and organisms across the different levels of the food web (Islam & Tanaka, 2004). Yet, a high probability of oil pollution exposure does not automatically translate into an "injury or an effect of biological consequences to an organism or population" (p. 419, Boehm & Page, 2007), due to the dynamics and complexities inherent in marine environments.

Further analysis could also focus on directly identifying vulnerable species to chronic oil pollution by using spatial and temporal information on these features, as suggested in Zacharias and Gregr (2005). For example, data on seabird species concentrations during different times of the year, in offshore and coastal regions, overlaid with the probability of oil pollution could be used to identify and map areas where seabird species would be at higher risk of oil fouling.

Finally, recent improvements on the monitoring capabilities of NASP will allow more accurate estimations of suspected oil spill volumes and the measurement of *in-situ* weather and sea conditions. With this information, more reliable measurements of the amount of oil entering the marine environment can be made. Accurate and reliable information on chronic oil pollution is considered essential to assess the efficiency of prevention measures established by national and international bodies, to identify marine and coastal regions more likely to be vulnerable to chronic oil pollution, and to establish localized strategies to minimize or eradicate any potential detrimental effects.

Chapter 4 CONCLUSIONS

4.1 Introduction

On the 20th of April, 2010 began one of the worst accidental oil spills in the history of North America. An oil drilling rig contracted by British Petroleum (BP) exploded in the Gulf of Mexico releasing thousands of barrels of oil everyday with unpredictable environmental consequences (Robertson & Kaufman, April 28, 2010). Not surprisingly, large accidental oil spill events, such as this, often receive broad media coverage because of the dramatic picture they portray and their anticipated ecological and economical impacts. This study focuses on a different type of oil pollution event, characterized by small (< 700 tonnes) but frequent illicit releases of oily substances from sea-based activities, and in particular from operations associated with marine vessel. While single operational oil spills are often difficult to observe and their individual impacts may appear insignificant if compared with large accidental oil spills, it is the cumulative effect of the constant releases of oil that make these events particularly worrisome (Camphuysen, 2008; Wiese & Ryan, 2003). In fact, studies reveal that operational discharges constitute the largest source of oil inputs into the marine environment from shipping activities (GESAMP, 2007).

Evidence of chronic oil pollution has been gathered in different parts of the world using surveillance aircraft and satellite monitoring (Carpenter, 2007; Ferraro et al., 2009). On the west coast of Canada, the National Aerial Surveillance Program (NASP) has been collecting spatial and temporal information on illicit oil spill observations since the early 1990s and constitutes one of the main monitoring and enforcement tools of national and

international pollution regulations (Armstrong & Derouin, 2004). Despite NASP's efforts, and due to the numerous challenges involving the surveillance of a large area and the detection of oil on the sea surface, most illicit oil spills still remain unaccounted for. Moreover, there are a large number of detected oil spills without an identified source (i.e. mystery spills) which makes it impossible for NASP to collect the evidence needed to identify the source of the spill and subsequent prosecution of the illicit polluter.

In light of the challenges associated with the observation and quantification of chronic oil pollution off the west coast of Canada, this study generated a statistical model in order to describe the spatial variability of oil spill occurrences and draw inferences about the underlying processes driving the likelihood of occurrence. The model outputs also enabled us to identify areas more likely to be impacted by these events (i.e. hot spots), not only in areas surveyed by NASP but for most areas of Canada's Pacific region where NASP effort is low or non-existent. In the next section we summarize our findings, contributions of this study and future research opportunities.

4.2 Summary of findings

In Chapter 2, we explore different approaches for modelling illicit oil spill occurrences observed by NASP in offshore (Area A) and inshore (Area B) waters of Canada's Pacific region between 1998 and 2007. A number of statistical models were tested using oil spill observations and a set of estimated predictor variables (shipping traffic data, distance to shore and to port, density of small harbours), while accounting for surveillance effort. Special attention was dedicated to the large number of zero observations present in the oil spill data and its effects in building and selecting the best modelling approach. The best-fit

model (for Area A and Area B) allowed for the description of the variability in abundance and probability of oil spill occurrence. These models were also used to predict the number of oil spills and probability of oil spill occurrences in areas not surveyed by NASP between 1998 and 2007. This exercise helped us to define the best approaches to analyze operational oil spill observations and to generate new hypotheses about the underlying factors thought to influence the distribution of chronic oil pollution in the region. Most importantly, our findings allowed us to address the two questions outlined in the introductory chapter:

1. Which sea-based activities better describe the distribution of detected oil spills?

The two predictor variables that show significant positive association with oil spill occurrences and abundances in both Areas A and B are distance to shore and density of small harbours. This points to small vessels that are likely to stay closer to shore and visit marinas. The association with small harbours is not surprising given that these areas often contain the largest concentration of boats and are also where most of the routine vessel operations are carried out (such as vessel repair and fueling) (Molnar & Koshure, 2009). Only in inshore waters does large vessel shipping traffic show a significant association with oil spill occurrences and abundances. More specifically, tug boats and oil tankers show a positive relationship, although oil tankers to a lesser degree, while carrier vessel traffic (including bulk cargo vessels and containers ships) shows a negative relationship, despite the fact that carrier vessels are the second most abundant vessel type in inshore waters. The dissimilar relationship between oil spill occurrences and abundances and carrier vessel and oil tanker traffic may be explained by the fact that more oil spills were

observed closer to shore and not in areas dominated by large commercial vessel traffic. Tug boats are the most predominant and wide-spread commercial vessels across maritime British Columbia, particularly in sheltered waters, which may explain why they show a higher spatial association with chronic oil pollution distributions. These relationships should be interpreted with caution given the distribution of surveillance NASP flights, which are more predominant in inshore waters than in offshore waters.

2. *Which areas show the highest probability of illicit oil spill occurrence and relative abundance in Canada's Pacific waters?*

In offshore waters (Area A), the highest probability of oil spill occurrence is in Barkley Sound, one of the major navigation corridors for commercial and non-commercial vessels based in Ucluelet and Port Alberni. In inshore waters (Area B), the highest probabilities are found near major coastal communities, including Prince Rupert, Campbell River, Nanaimo, Pender Harbour, Vancouver, the Fraser Delta and Sidney, which are characterized by possessing high density of marinas, public and private docks. Our findings should be interpreted with care because they are only exploratory and are subject to different levels of uncertainty, such as the assumption of perfect detection, the limited spatial and temporal (only during the day time) surveillance coverage and the fact that the observed oil spills are considered only look-alikes and not true detections.

Chapter 3 focuses on the application of the results outlined in the previous chapter, by assessing the spatial overlay between the estimated probability of illicit oil spill occurrences and the location of Coastal and Marine Protected Areas (CMPAs) and

Important Bird Areas (IBAs). We chose CMPAs and IBAs in our analysis because they are geographically distinct areas likely to contain valuable ecological features sensitive to oil pollution. The outcomes of this study allow us to respond to the question introduced in Chapter 1:

1. *Which regions show the highest vulnerability to chronic oil pollution based on the location of CMPAs and IBAs?*

Our findings identified four areas of highest vulnerability: the Tofino Mudflats, Barkley Sound, Robert's and Sturgeon Banks and the Scott Islands. These regions were selected because they show relatively high exposure index, defined as the probability of observing an oil spill greater than 27% and over 10 years, they contain very sensitive species and ecosystems to oil pollution and they are protected under one or more CMPA designations and are identified as IBAs. Nevertheless, any area protected specifically to preserve important ecosystems and species that, in addition, are highly sensitive to oil pollution (e.g., seabird species) are likely to be more vulnerable to any degree of exposure than protected areas that do not contain features sensitive to oil pollution. In particular, we refer to areas protected under National Wildlife Areas, Migratory Bird Areas, Ecological Reserves and Wildlife Management Areas or designated under the Important Bird Areas program.

4.3 Research contributions

We believe our findings make a significant contribution in beginning to understand the distribution of illicit oil spill events in Canada's Pacific region. Our description of chronic

oil pollution patterns and the estimated probabilities of oil spill occurrences and relative abundances provide valuable baseline information from which to compare oil spill observations currently collected by the new NASP aircraft. Our results will be available for future comparative analysis using oil spill observations collected by NASP in other parts of Canada, such as in the Arctic and Atlantic regions.

This study also contributes to NASP monitoring efforts in protecting Canada's Pacific region by informing the program of likely sources of chronic oil pollution and by identifying oil pollution hot spots in the region. This information can be used by NASP to plan new monitoring strategies and increase its efficiency in terms of the identification of marine polluters. The latter is particularly important for aerial monitoring programs, such as NASP, responsible for patrolling large areas with busy navigation corridors. An improvement in monitoring efficiency can translate into a reduction of the number of mystery spills and an increase in the number of successful prosecutions. These are key indicators of the success of any pollution monitoring program aiming to deter illicit polluters and reduce the amount of oil introduced into the marine environments (Carpenter, 2007).

Another important contribution of this research emerges from the use of relatively new modelling approaches designed to deal with large number of zeros in count data distributions. Numerous examples exist in the literature illustrating the use of zero-inflated models and hurdle models in different research areas, from ecology to social sciences (Ridout et al., 1998); however, no examples were found of their usage in modelling maritime pollution observations. We also showed the utility and importance of comparing these complex count models with simpler GLMs (with Poisson and negative binomial

distributions) since the use of complex models cannot always be justified, even in the presence of large number of zero observations (Warton, 2005). At the same time, we acknowledge that our inability to appropriately model false zeros, generated by imperfect detection, likely influenced our decision towards a simpler model. Our discussion in Chapter 2 about potential sources of bias in the modelling approach, including our inability to model detectability (or the presence of false zeros), may appeal to researchers interested in modelling events that are rare and difficult to observe.

Our findings also add to the group of regional studies evaluating the impact of anthropogenic maritime activities on coastal marine environments in Canada's West Coast (Ban & Alder, 2008; Ban et al., 2010). This is the first time that modelled probabilities of oil spill occurrence from vessel operations were used to identify vulnerable CMPAs and IBAs in BC, based on probability of exposure and their sensitivity to this stressor. Highlighted CMPAs and IBAs can be used as monitoring sites to measure and assess true effects of chronic oil pollutions onto marine ecosystems and species. With this information at hand, coastal and marine planners can evaluate the efficiency of CMPAs in protecting coastal and marine ecosystems and species from human threats. Lastly, the methodological approach we have developed is repeatable and can be easily implemented in other regions as well as for analyzing different threats.

In summary, we have shown that modelling risk of chronic oil pollution based on data collected by NASP can be very informative to improve NASP efficiency, to identify troublesome areas, to quantify and describe potential impacts, and form a baseline for future analysis aiming to model observations that are difficult to detect and that arise from complex underlying processes.

4.4 Research Opportunities.

The results from our modelling approach can be used as a guideline for modelling oil spill data collected by the recently acquired Dash 8 and operating in Canada's Pacific Region. The Dash 8 aircraft has a much greater flying range than the old aircraft (the Twin Otter) and it is equipped with a state-of-the-art suite of remote sensors. These sensors allow for better observation of oil slicks look-alikes in different weather conditions and during day and night flights and more accurate estimation of the location and amount of oil spilled (McGregor & Gautier, 2008). The access to additional and more accurate data opens the door to new and interesting research opportunities. For example, oil spill rates derived in this study can be used as a baseline to compare with future rates estimated from the new NASP data. The new data will also allow for the comparison of daily and seasonal oil spill observations. These analyses were not possible in the current study because NASP was limited to daytime oil spill observation and not enough data was available for the convergence of seasonal models. Minimum estimates of oil spill volumes will enable the exploration of spatial variation in oil inputs and the assessment of the potential impacts that these oil spills have on the marine environment. Yet, it is important to remember that this new technology is only capable to detect “anomalies” and that ground verification is still necessary to indentify if the observed anomaly is in fact an oil spill (Fingas & Brown, 2007).

This aircraft has also the potential to collect *in-situ* information on factors known to influence the probability of detecting an oil discharge on the sea surface from an aircraft (e.g. wind speed, look angle, sun angle, altitude of the aircraft). As mentioned before, this

information is essential to account for variation in detectability and thus model the presence of false zeros in oil spill counts.

One of the foreseen challenges of the new monitoring systems used in the Dash 8 is the quantification of surveillance effort. Oil spill anomalies are being observed either using Side-Looking Airborne Radar (or SLAR) or using other instruments (e.g., high resolution cameras) or by visual observations, depending on the flying altitude and the geographic characteristics of the area under observation (narrow passages vs. open water) (K. Pearce, EC, pers. comm., 2010). The SLAR imagery is also affected by weather conditions, particularly by wind speed, which affects the discrimination between oil and look-alikes (Fingas & Brown, 2007). Future analysis will need to take into account all these factors in order to properly estimate surveillance effort, a necessary parameter to standardize oil spill observations and adequately interpret the distribution of oil pollution hot spots.

Seasonal analysis will not only bring a new insight into the variability of oil spill detection rates, but they will also incorporate variation currently existing in maritime vessel traffic. For example, the number of vessel movements in Canada's West Coast increases during the summer months as a result of the increase in cruise ship activity and ferry sailings. Further, the estimation of oil spill risk surfaces for summer and winter will allow for more specific analysis on the impact that chronic oil pollution has on BC's ecosystems, by overlying the seasonal probability of oil spill occurrence with sensitive species that are seasonally present, such as migratory seabird species.

Other maritime shipping parameters can be extracted from the Vessel Traffic Services database and compared with oil spill observations, such as the direction of the vessel (inbound vs. outbound) and flag of the vessel. This information could provide new

insights into the type of vessels associated with illicit oil pollution as well as the behavioural patterns of illicit polluters. For instance, we could explore if large vessels (e.g. tankers and carriers) prefer to intentionally release residual oils while leaving or entering international waters, or if ships navigating under flags of convenience are in fact the largest contributors to oil pollution (Shaw et al., 1987).

Future developments in North Coast of BC are likely to change the distribution and intensity of shipping traffic. With the foreseen expansion of Prince Rupert's container terminal (Prince Rupert Port Authority website) and plans to develop a terminal in Kitimat to receive oil tankers (District of Kitimat website), the number of large commercial vessels in this region is likely to increase. New studies will have to incorporate future shifts in maritime shipping as they will likely influence the probability of oil spill occurrence. In addition, the introduction of a new maritime traffic tracking and data archiving system, named Information System and Marine Navigation (INNAV) and expected to be at a lower resolution (P. O'Hara, EC, pers. comm., 2010), will likely impact the way chronic oil pollution is modelled in the near future.

Finally, future efforts should be invested in the development of standard protocols for oil spill and NASP flights data archiving and metadata reporting. This will allow a reduction of the amount of time dedicated to compile and transform data into suitable formats for data analysis, for example in GIS software. One of the most time consuming tasks of this research was the transformation of flight path records to a GIS compatible format to be able to estimate surveillance effort. Further, the creation of standard protocols at a national level will facilitate the comparison of oil inputs from sea-based activities across the country.

Bibliography

- Agresti, A. (1990). *Categorical data analysis*. New York: Wiley-Interscience, 710pp.
- Akaike, H. (1973). *Information theory and an extension of the maximum likelihood principle*. Paper presented at the Second International Symposium Information Theory" Akademiai Kidao, Budapest, Hungary.
- Anselin, L., & Bera, A. K. (1998). Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In A. Ullah & D. E. Giles (Eds.), *Handbook of Applied Economic Statistics* (pp. 237-289). New York: Marcel Dekker.
- Anselin, L., Syabri, I., & Kho, Y. (2005). GeoDa: An introduction to Spatial Data Analysis. Spatial Analysis Laboratory, Department of Agricultural and Consumer Economics, University of Illinois, IL
- Armstrong, L., Derouin, K. (2004). National Aerial Surveillance Program 2001 - 2004. Final Report. Transport Canada Marine Safety. Ottawa, Ontario.
- Armstrong, L. (2004). National Aerial Surveillance Program: Statement of Operational Requirements (Draft). Transport Canada Marine Safety.
- Bailey, T., & Gatrell, A. (1995). *Interactive Spatial Data Analysis*. Essex, England: Longman Scientific & Technical. Brunt Mill, 413pp.
- Ban, N., & Alder, J. (2008). How wild is the ocean? Assessing the intensity of anthropogenic marine activities in British Columbia, Canada. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 18(1), 55-85.
- Ban, N. C., Alidina, H. M., & Ardron, J. A. (2010). Cumulative impact mapping: Advances, relevance and limitations to marine management and conservation, using Canada's Pacific waters as a case study. *Marine Policy*, 34(5), 876-886.
- BC Parks .(n.p.). Summary of Park and Protected Area Designations. Retrieved on June 2010 from <http://www.env.gov.bc.ca/bcparks/conservation.html>
- Brekke, C., & Solberg, A. H. S. (2005). Oil spill detection by satellite remote sensing. *Remote sensing of environment*, 95(1), 1-13.
- Bird Studies Canada and Nature Canada. (2004-2009). Important Bird Areas of Canada Database. Port Rowan, Ontario: Bird Studies

- Boehm, P. D., & Page, D. S. (2007). Exposure Elements in Oil Spill Risk and Natural Resource Damage Assessments: A Review. *Human and Ecological Risk Assessment: An International Journal*, 13(2), 418 - 448.
- Burger, A. E. (1993). Estimating the mortality of seabirds following oil spills: Effects of spill volume. *Marine Pollution Bulletin*, 26(3), 140.
- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and model inference. A practical information-theoretic approach, 2nd edn.* New York: Springer - Verlag, 353pp.
- Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data.* Cambridge: Cambridge University Press, 411pp.
- Camphuysen, C. J. (1998). Beached bird surveys indicate decline in chronic oil pollution in the North Sea. *Marine Pollution Bulletin*, 36(7), 519.
- Camphuysen, C. J. (2008). *Chronic Oil Pollution in Europe:* IFAW (International Fund for Animal Welfare).
- Camphuysen, C. J., & Heubeck, M. (2001). Marine oil pollution and beached bird surveys: the development of a sensitive monitoring instrument. *Environmental Pollution*, 112, 443-461.
- Canadian Coast Guard. (1998). Tanker Exclusion Zone. Retrieved on June 2010 from <http://www.ccg-gcc.gc.ca/e0003909>
- Canessa, R. & Lunn, K. (2005). Developing and Evaluating Indicators to Measure the Effectiveness of Coastal and Marine Protected Areas on Canada's Pacific Coast. Presented at the *North America Marine Protected Areas Network Symposium*. Loreto, Mexico. March 2005.
- Carpenter, A. (2007). The Bonn Agreement Aerial Surveillance programme: Trends in North Sea oil pollution 1986-2004. *Marine Pollution Bulletin*, 54(2), 149-163.
- Chambers, J. M., & Hastie, T. J. (1993). *Statistical models in S.* New York ; London: Chapman & Hall, 608pp.
- Chen, C. (2002). *Robust regression and outlier detection with the Robustreg procedure.* . Paper presented at the Proceedings of the Twenty-Seventh Annual SAS Users Group International Conference, SAS Institute, Inc, Cary, NC.

- Chrastansky, A., & Callies, U. (2009). Model-based long-term reconstruction of weather-driven variations in chronic oil pollution along the German North Sea coast. *Marine Pollution Bulletin*, 58(7), 967-975.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1), 1-38.
- Department of Fisheries and Oceans. (2007). Rockfish Conservation Areas. Retrieved June 2010 from <http://www.pac.dfo-mpo.gc.ca/fm-gp/rec/restricted-restraint/rca-acs-eng.htm>
- District of Kitimat website. Retrieved on July 2010 from <http://www.investkitimat.ca/index.php/Projects/HomePage>
- Dormann, C., F., McPherson, J., M., Araújo, M., B., Bivand, R., Bolliger, J., Carl, G., et al. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), 609-628.
- Environment Canada. (2006). Satellite program aims to introduce marine pollution. *EnviroZine. Environment Canada's Online Newsmagazine*. Retrieved on June 2010 from https://www.ec.gc.ca/EnviroZine/english/issues/70/feature2_e.cfm
- Environment Canada. (2007). Canada's New Government to Protect Scott Islands off British Columbia's Coast. News Release. Retrieved on June 2010 from <http://www.ec.gc.ca/default.asp?lang=En&n=714D9AAE-1&news=67B8B3F8-DF0F-44AD-A41F-6D9BCF64C62C>
- EPA - U. S. Environmental Protection Agency. (1999). *Understanding Oil Spills and Oil Spill Response*: Oil Program Center. Office of Emergency and Remedial Response, 47 pp.
- EPA - U.S. Environmental Protection Agency. (2001). *National Management Measures Guidance to Control Nonpoint Source Pollution from Marinas and Recreational Boating*. Nonpoint Source Control Branch. Office of Wetlands, Oceans and Watersheds. Office of Water.
- Etkin, D. S. (2001). *Analysis of oil spill trends US and worldwide*. Paper presented at the International Oil Spill Conference.
- Eyer, H.L. (2004). Hawth's Analysis Tools for ArcGIS. <<http://www.spatial ecology.com/htools>>.

- Ferraro, G., Meyer-Roux, S., Muellenhoff, O., Pavliha, M., Svetak, J., Tarchi, D., et al. (2009). Long term monitoring of oil spills in European seas. *International journal of remote sensing*, 30(3), 627 - 645.
- Fingas, M., & Brown, C.E. (2007). Oil Spill Remote Sensing: A Forensic Approach. In Z. Wang & S. Stout (Eds.). *Oil Spill Environmental Forensics : Fingerprinting and Source Identification*. (pp. 419-447) Burlington: Academic Press.
- Fortin, M.J., & Dale, M., R. T. (2005). *Spatial Analysis: a Guide for Ecologists*. Cambridge: Cambridge University Press, 365pp.
- Fortin, M.J., & Melles, S. J. (2009). Avian Spatial Responses to Forest Spatial Heterogeneity at the Landscape Level: Conceptual and Statistical Challenges. In *Real World Ecology* (pp. 137-160).
- Fotheringham, S. A., Brunson, C., & Charlton, M. (2002). *Geographically Weighted Regression the analysis of spatially varying relationships*. New York: John Wiley & Sons Ltd, 269pp.
- Georgia Strait Alliance. (2009). Green boating program. Retrieved December 7, 2009, from <http://www.georgiastrait.org/?q=node/34>
- GESAMP - Joint Group of Experts on the Scientific Aspects of Marine Environment Protection. (2007). *Estimates of oil entering the marine environment from sea-based activities*. London: International Maritime Organization.
- Golik, A., & Rosenberg, N. (1987). Quantitative evaluation of beach-stranded tar balls by means of air photographs. *Marine Pollution Bulletin*, 18(6), 289-293.
- Greene, W. (1994). *Accounting for excess zeros and sample selection in Poisson and Negative Binomial regression models. Working paper.*: Department of Economics. New York university.
- Griffith, D. A. (1992). What is spatial autocorrelation? Reflections on the past 25 years of spatial statistics. *L'Espace géographique*, 21(3), 265-280.
- Guisan, A., & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, 8(9), 993-1009.
- Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135(2-3), 147-186.

- Haining, R. (2003). *Spatial Data Analysis. Theory and Practice*. Cambridge: Cambridge University Press, 432pp.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis with readings (4th ed.)*: Macmillan Publishing Co., Inc., 730pp.
- Halpern, B. S., McLeod, K. L., Rosenberg, A. A., & Crowder, L. B. (2008). Managing for cumulative impacts in ecosystem-based management through ocean zoning. *Ocean & Coastal Management*, 51(3), 203-211.
- Halpern, B. S., Selkoe, K. A., Micheli, F., & Kappel, C. V. (2007). Evaluating and Ranking the Vulnerability of Global Marine Ecosystems to Anthropogenic Threats. *Conservation Biology*, 21(5), 1301-1315.
- Hampton, S., Kelly, P. R., & Carter, H. R. (2003). Tank vessel operations, seabirds, and chronic oil pollution in California. *Marine Ornithology*, 31, 29-34.
- Harper, J. R., Christian, J., Cross, W. E., Frith, R., Searing, G. F., & Thompson, D. (1993). *A classification of the marine regions of Canada*. Ottawa, Ontario, Canada: Environment Canada.
- Harrell, F. E. J., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, 15(4), 361-387.
- Heilbron, D.C. (1994). Zero-Altered and other Regression Models for Count Data with Added Zeros. *Biometrical Journal*, 36(5), 531-547.
- Hiddink, J. G., Jennings, S., & Kaiser, M. J. (2007). Assessing and predicting the relative ecological impacts of disturbance on habitats with different sensitivities. *Journal of Applied Ecology*, 44(2), 405-413.
- Hosmer, D.W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed. ed.). New York: Wiley, 375pp.
- Important Bird Areas Canada. (2004-2009). Website. <http://www.ibacanada.com/>
- IMO - International Chamber of Shipping. (2009). Shipping and World Trade. Retrieved November 12, 2009, 2009, from <http://www.marisec.org/shippingfacts/worldtrade/index.php>

- IMO - International Marine Organization. (2007). International Convention from the Prevention of Pollution from Ships, 1973, as modified by the Protocol of 1978 relating thereto (MARPOL 73/78). September 18, 2007 from http://www.imo.org/Conventions/contents.asp?doc_id=678&topic_id=258
- IMO - International Maritime Organization. (2009). *Summary of Status of Convention, As at 31 October 2009*: IMO.
- IMO - International Maritime Organization. (n.p.). Special Areas under MARPOL. April 26, 2010, from http://www.imo.org/Environment/mainframe.asp?topic_id=760
- IMO - International Maritime Organization. (n.p.). Ships' routing. Retrieved on June 2010 from http://www.imo.org/Safety/mainframe.asp?topic_id=770
- IPIECA. (1991). *Guidelines on biological impacts of oil pollution*. Retrieved on June 2010 from http://ipieca.org/activities/oilspill/oil_publications.php
- Islam, S., & Tanaka, M. (2004). Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. *Marine Pollution Bulletin*, 48(7-8), 624-649.
- ITOPF. (2002). *Fate of marine oil spills*. Retrieved on June 2010 from <http://www.itopf.com/marine-spills/fate/>
- ITOPF. (2009). *Oil Spill Tanker Statistics 2008*. Retrieved on June 2010 from <http://www.itopf.com/information-services/data-and-statistics/statistics/>
- Jamieson, G. S., & Lessard, J. (2000). Marine protected areas and fisheries closures in British Columbia. Canadian Special Publication of Fisheries and Aquatic Sciences, no. 131.
- Jelinski, D., & Wu, J. (1996). The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology*, 11(3), 129-140.
- Kasoulides, G. (1988). Identification of the sources of discharged pollutants. *Marine Pollution Bulletin*, 19(7), 307.
- Khee-Jin Tan, A. (2005). *Vessel-Source Marine Pollution: The Law and Politics of International Regulation*. Cambridge Studies in International and Comparative Law: Cambridge University Press, 416pp.

- Kluser, S., Richard, J. P., & Giuliani, G. (2006). *Illegal oil discharge in European Seas*. Environment Alert Bulletin. United Nations Environmental Programme. DEWA - GRID-Europe. Retrieved on June, 2010 from http://www.grid.unep.ch/product/publication/download/ew_oildischarge.en.pdf
- Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. *Technometrics*, 34(1), 1-14.
- Legendre, P. (1993). Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, 74(6), 1659.
- Levine, N. (2004). CrimeStat III: A Spatial Statistics Program for the Analysis of Crime Incident Locations (Version 3.0). Ned Levine and Associates, Houston, TX, and the National Institute of Justice, Washington, DC.
- Liu, W., & Cela, J. (2008). *Count data models in SAS*. Paper presented at the SAS Global Forum.
- Lock, A., Brown, R., & Gerriets, S. (1994). *Gazetteer of marine birds in Atlantic Canada: an atlas of sea bird vulnerability to oil pollution*. Ottawa: Environment Canada, Canadian Wildlife Services, 137 pp.
- MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8), 2248-2255.
- MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2006). *Occupancy Estimation and Modeling. Inferring Patterns and Dynamics of Species Occurrences*. Burlington, MA: Academic Press. Elsevier, 324pp.
- Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., et al. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters*, 8(11), 1235-1246.
- McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman & Hall, 511pp.
- McGregor, M., & Gautier, M. (2008). *Enhancing Situational Awareness in the Arctic Through Aerial Reconnaissance*. Ottawa, Ontario: Environment Canada.

- Molnar, M., & Koshure, N. (2009). *Cleaning Up Our Ocean: A report on ocean pollution from shipping-related sources in the Pacific North Coast Integrated Management Area (Pncima) on the British Columbia coast*. Vancouver: David Suzuki Foundation.
- Montgomery, D.C., Peck, E.A. (1992). *Introduction to linear regression analysis*. New York: Wiley, 612pp.
- Mooradian, C. (2002). Protecting "sovereign rights": The case for increased coastal state jurisdiction over vessel-source pollution in the exclusive economic zone. *Boston University law review*, 82(3), 767-816.
- Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, 33(3), 341-365.
- National Academy of Sciences. (2003). *Oil in the Sea III: Inputs, Fates, and Effects*. Washington, D.C. : National Research Council Of The National Academies.
- National Marine Manufacturers Association. (2009). *Recreatioanl Boating Statistical Abstract*.
- O'Brien, R. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality and Quantity*, 41(5), 673-690.
- O'Hara, P., & Morgan, K. H. (2006). Do low rates of oiled carcass recovery in beached bird surveys indicate low rates of ship-source oil spills. *Marine Ornithology* 34, 133-140.
- Ohowa, B.O. (2009). Evaluation of the effectiveness of the regulatory regime in the management of oil pollution in Kenya. *Ocean & Coastal Management*, 52(1), 17-21.
- Pavlaklis, P., Tarchi, D., Seiber, A. J., Ferraro, G., & Vincent, G. (2001). *On the monitoring of illicit vessel discharges. A reconnaissance study in the Mediterranean Sea*: EC DG - Joint Research Center EC DG - Environment European Commission, 20 pp.
- Pearce, J., & Ferrier, S. (2000). Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133(3), 225-245.
- Potts, J. M., & Elith, J. (2006). Comparing species abundance models. *Ecological Modelling*, 199(2), 153-163.

- Prince Rupert Port Authority website. Retrieved on June 2010 from <http://www.rupertport.com/container2.html>
- R Development Core Team. (2008). R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.r-project.org/>
- Ridout, M., Demétrio, C. G. B., & Hinde, J. (1998). *Models for count data with many zeros*. Paper presented at the International Biometric Conference
- Robertson, C., & Kaufman, L. (2010, April 28). Size of Spill in Gulf of Mexico Is Larger Than Thought. *The New York Times*. Retrieved from http://www.nytimes.com/2010/04/29/us/29spill.html?_r=1
- Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, *84*(3), 777-790.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461-464.
- Serra-Sogas, N., O'Hara, P. D., Canessa, R., Keller, P., & Pelot, R. (2008a). Visualization of spatial patterns and temporal trends for aerial surveillance of illegal oil discharges in western Canadian marine waters. *Marine Pollution Bulletin*, *56*(5), 825-833.
- Serra-Sogas, N., O'Hara, P., Canessa, R., Bertazzon, S., & Gavrilova, M. (2008b). Exploratory Spatial Analysis of Illegal Oil Discharges Detected off Canada's Pacific Coast. In *Computational Science and Its Applications – ICCSA 2008* (pp. 81-95).
- Shahidul Islam, M., & Tanaka, M. (2004). Impacts of pollution on coastal and marine ecosystems including coastal and marine fisheries and approach for management: a review and synthesis. *Marine Pollution Bulletin*, *48*(7-8), 624-649.
- Shaw, B., Winslett, B. J., & Cross, F. B. (1987). The global environment: A proposal to eliminate marine oil pollution. *Natural Resources Journal*, *27*(1), 157.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London, New York: Chapman and Hall, 175pp.
- Sorum, A. (2008, Jan 25,). Boating Impact on Canadian Economy. *suite101.com*.

- Statistics Canada. (2004). *Shipping in Canada - 2002*. Ottawa, Canada: Ministry of Industry.
- Statistics Canada. (2005). *Shipping in Canada - 2003*. Ottawa, Canada: Ministry of Industry.
- Statistics Canada. (2007). *Shipping in Canada - 2004*. Ottawa, Canada: Ministry of Industry.
- Statistics Canada. (2008). *Shipping in Canada - 2005*. Ottawa, Canada: Ministry of Industry.
- Statistics Canada. (2009). *Shipping in Canada - 2006*. Ottawa, Canada: Ministry of Industry.
- Streiner, D. L. (1996). Maintaining standards: differences between the standard deviation and standard error, and when to use each. *Canadian Journal of Psychiatry*, 41(8), 498.
- Stout, S.A., & Wang, Z. (2007). Chemical Fingerprinting of Spilled or Discharged Petroleum - Methods and Factors Affecting Petroleum Fingerprints in the Environment. In Z. Wang & S. Stout (Eds.). *Oil Spill Environmental Forensics : Fingerprinting and Source Identification*. (pp. 419-447) Burlington: Academic Press.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics (5th Ed.)*. Boston: Pearson/Allyn & Bacon.
- Tchoukanski, I. (2007). ET GeoWizard 9.8. Retrieved on June 2010 from www.ian-ko.com
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 42(2), 234 - 240.
- Tyre, A. J., Tenhumberg, B., Field, S. A., Niejalke, D., Parris, K., & Possingham, H. P. (2003). Improving Precision and Reducing Bias in Biological Surveys: Estimating False-Negative Error Rates. *Ecological Applications*, 13(6), 1790-1801.
- Venables, W. N., & Ripley, R., A. (2002). *Modern applied statistics with S* (4th Edition ed.). New York: Springer - Verlag, 495pp.
- Volckaert, F. A. M., Kayens, G., Schallier, R., & Jacques, T. G. (2000). Aerial Surveillance of Operational Oil Pollution in Belgium's Maritime Zone of Interest. *Marine Pollution Bulletin*, 40(11), 1051-1056.

- Vuong, Q. H. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, 57(2), 307-333.
- Wang, Z., Fingas, M., Landriault, M., Sigouin, L., Castle, B., Hostetter, D., et al. (1998). Identification and Linkage of Tarballs from the Coasts of Vancouver Island and Northern California Using GC/MS and Isotopic Techniques. *Journal of High Resolution Chromatography*, 21(7), 383-395.
- Warton, D. I. (2005). Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. *Environmetrics*, 16(3), 275-289.
- Welsh, A. H., Cunningham, R. B., Donnelly, C. F., & Lindenmayer, D. B. (1996). Modelling the abundance of rare species: statistical models for counts with extra zeros. *Ecological Modelling*, 88(1-3), 297.
- WHSRN website. (2009). Retrieved on June 2010 from <http://www.whsrn.org/western-hemisphere-shorebird-reserve-network>
- Widmer, W. M., & Underwood, A. J. (2004). Factors affecting traffic and anchoring patterns of recreational boats in Sydney Harbour, Australia. *Landscape and Urban Planning*, 66(3), 173-183.
- Wiese, F. K. (2002). *Seabirds and Atlantic Canada's ship-source oil pollution: impacts, trends, and solutions.*: World Wildlife Fund Canada, 88pp.
- Wiese, F. K., & Jones, I. L. (2001). Experimental Support for a New Drift Block Design to Assess Seabird Mortality from Oil Pollution. *The Auk*, 118(4), 1062-1068.
- Wiese, F. K., & Robertson, G. J. (2004). Assessing Seabird Mortality from Chronic Oil Discharges at Sea. *The Journal of wildlife management*, 68(3), 627-638.
- Wiese, F. K., & Ryan, P. C. (2003). The extent of chronic marine oil pollution in southeastern Newfoundland waters assessed through beached bird surveys 1984-1999. *Marine Pollution Bulletin*, 46(9), 1090-1101.
- Wildlife Habitat Canada. (2001). *The Status of Wildlife Habitats in Canada's Oceanic and Coastal Seascapes*. Retrieved on June 2010 from <http://www.whc.org/files/finaloceanslongversionenglish.pdf>
- Zacharias, M. A., & Gregr, E. J. (2005). Sensitivity and Vulnerability in Marine Environments: an Approach to Identifying Vulnerable Marine Areas. *Conservation Biology*, 19(1), 86-97.

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression Models for Count Data in R. *Journal of Statistical Software*, 27(8).

Zhang, L., Gove, J. H., & Heath, L. S. (2005). Spatial residual analysis of six modeling techniques. *Ecological Modelling*, 186(2), 154.

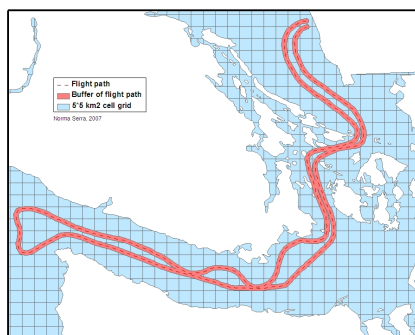
Zuur, A.F., Ieno, E.N., & Elphick, C.S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology & Evolution*, 1, 3-14.

Zuur, A.F., Ieno, E.N., & Smith, G.M. (2007). *Analysing Ecological Data*. New York: Springer. 672pp.

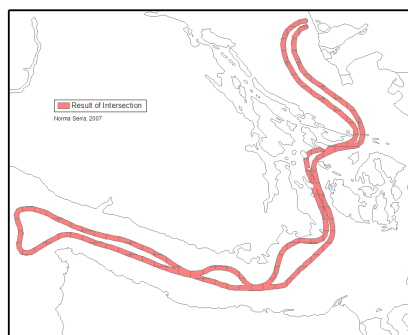
Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York: Springer Science+Business Media. 574pp.

Appendix A: NASP effort extraction

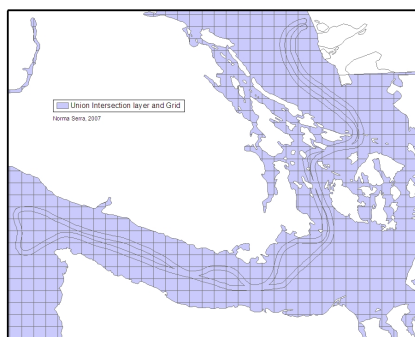
1. Ocean-only grid & Flight Buffer



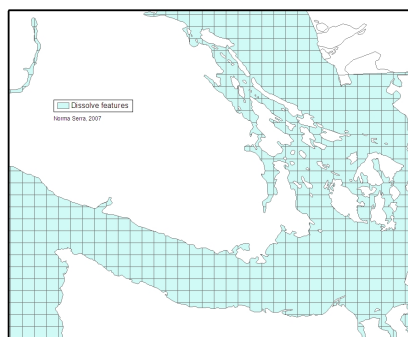
2. Intersection



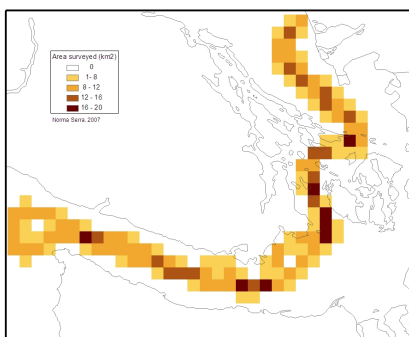
3. Union



4. Dissolve



5. Feature to Raster. Area surveyed per cell.



Appendix B: Vessel Traffic Services in Canada's Pacific Region

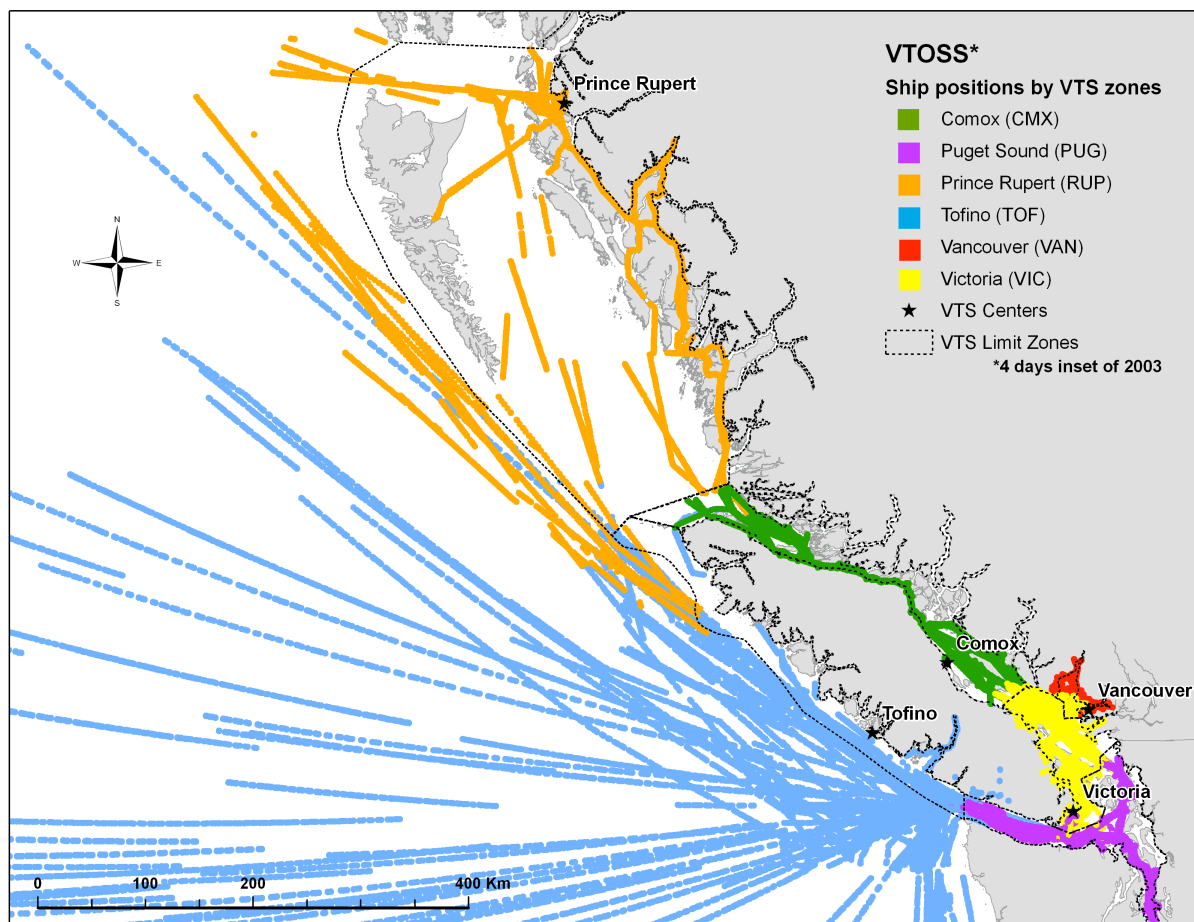


Figure B. 1. Vessel Traffic Services zones in Canada's west coast.

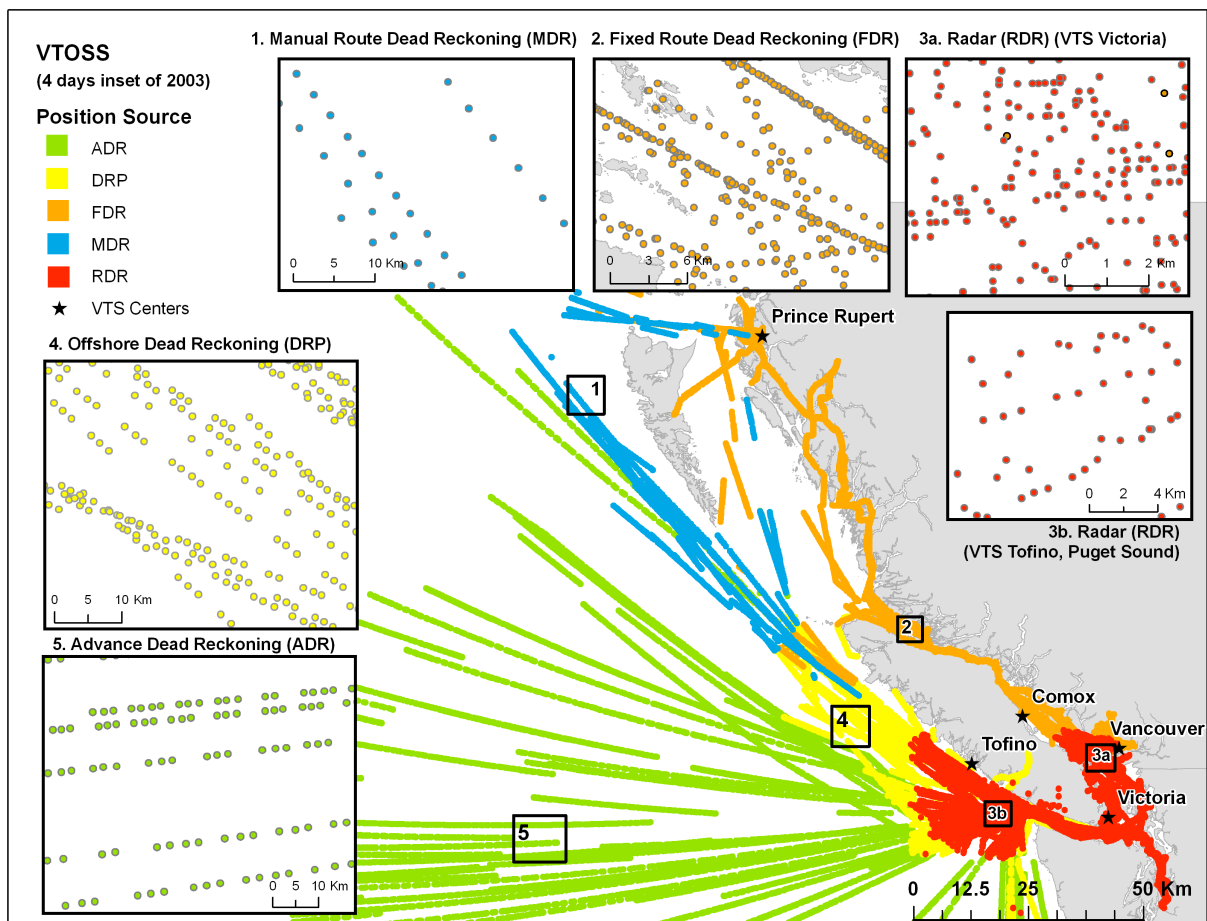


Figure B. 2. Vessel tracking systems and accuracies in used by the Marine Communication Traffic Services in Canada's west coast.

Appendix C: Scatterplots of model variables and variable selection results

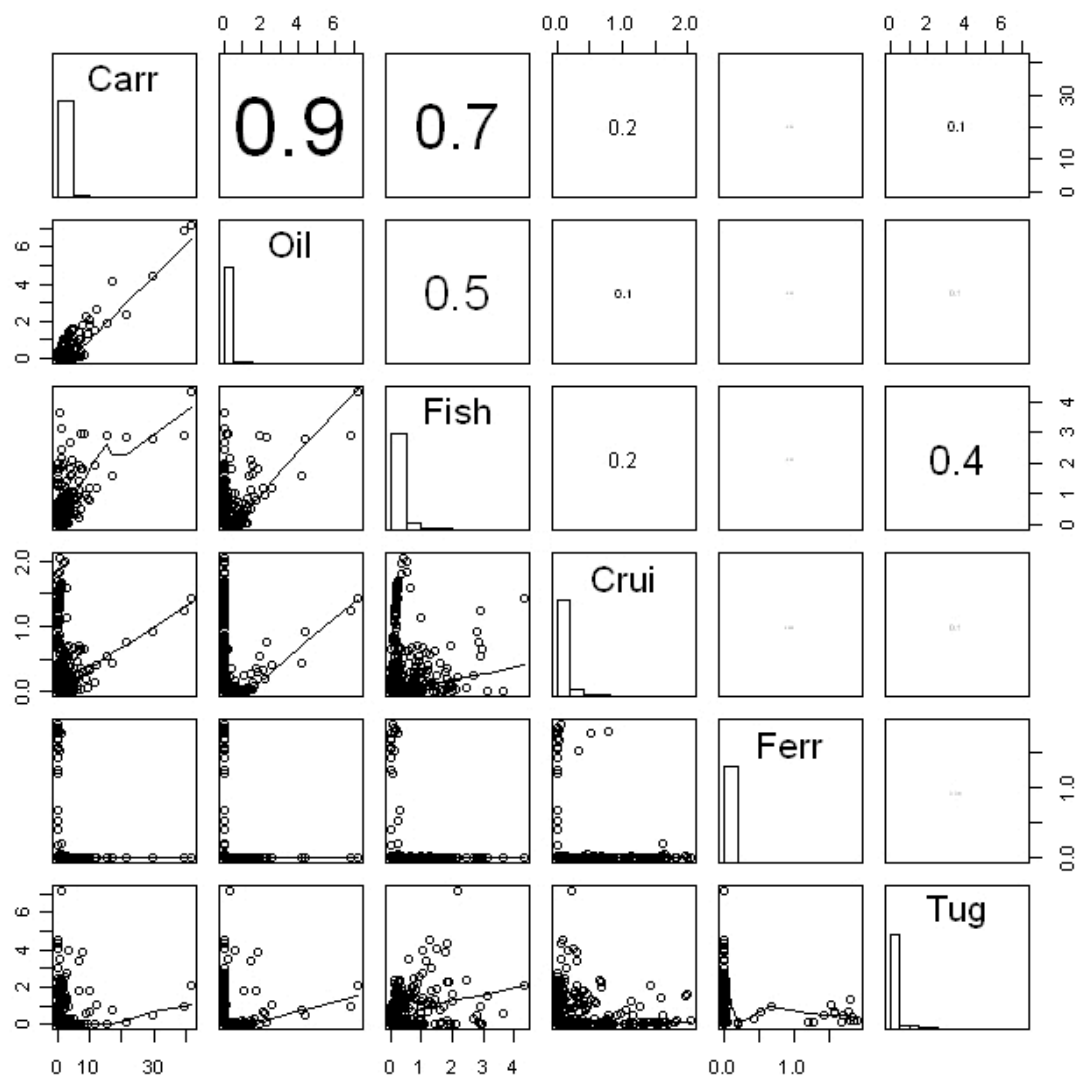


Figure C. 1. Scatterplot matrix of shipping traffic variables for Area A, including LOESS lines (lower left), histograms (on the diagonal) and correlation coefficients (upper right). Font size of correlation coefficients changes based on the value.

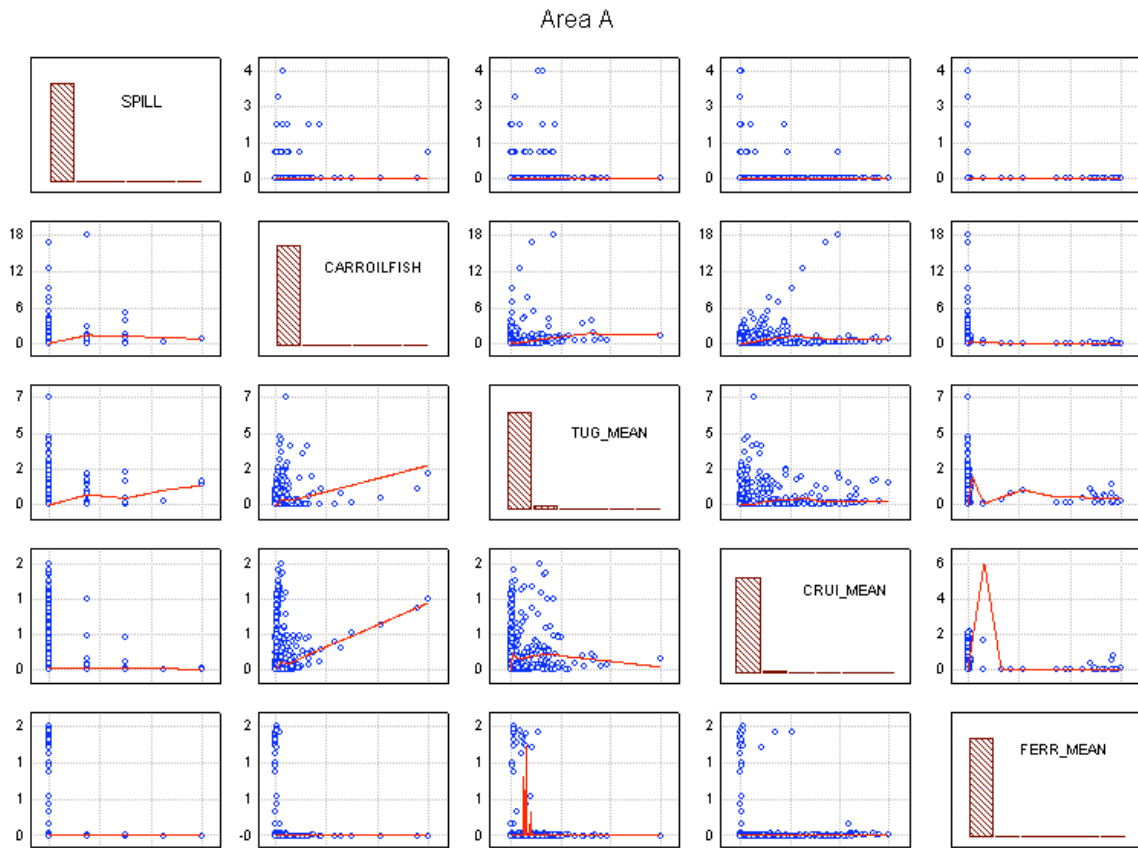


Figure C. 2. Matrix plots of Area A model variables spill counts and shipping traffic intensities for different vessel types, including LOESS lines and histograms on the diagonal.

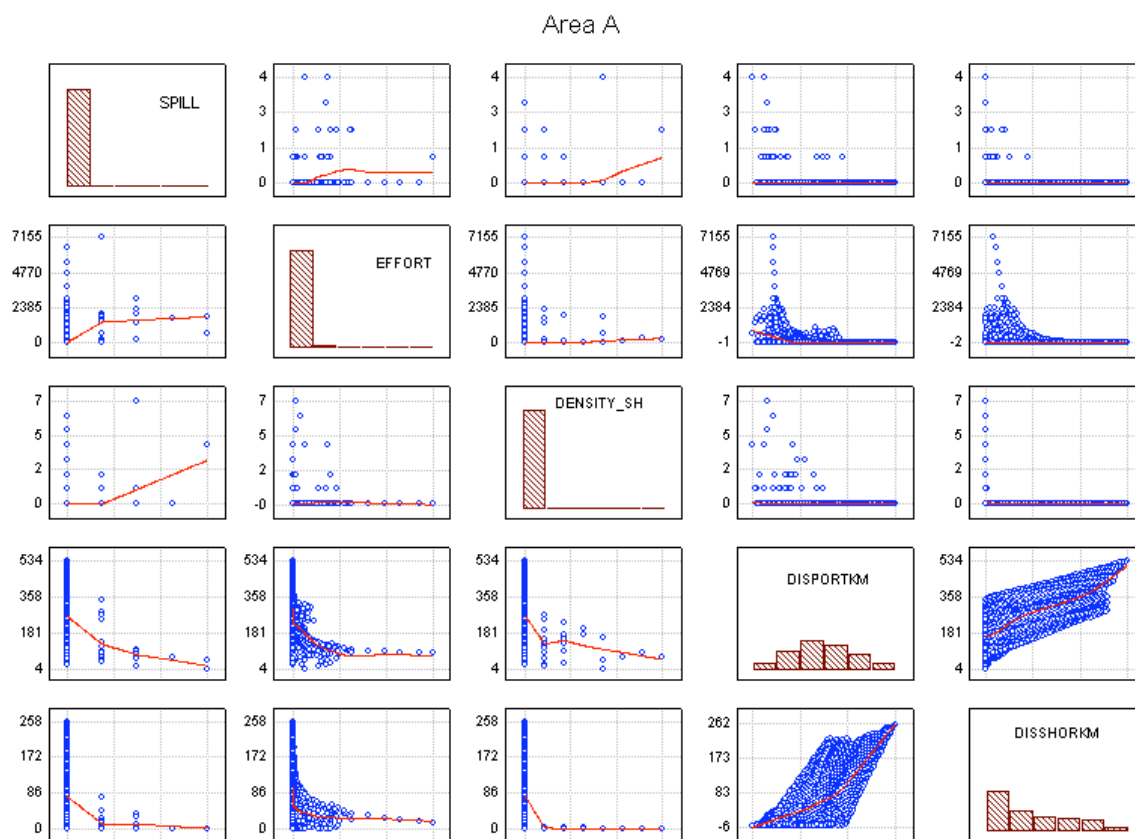


Figure C. 3. Matrix plots of Area A model variables spill counts, effort, density of small harbours, distance to shore and to port, including LOESS lines and histograms on the diagonal.

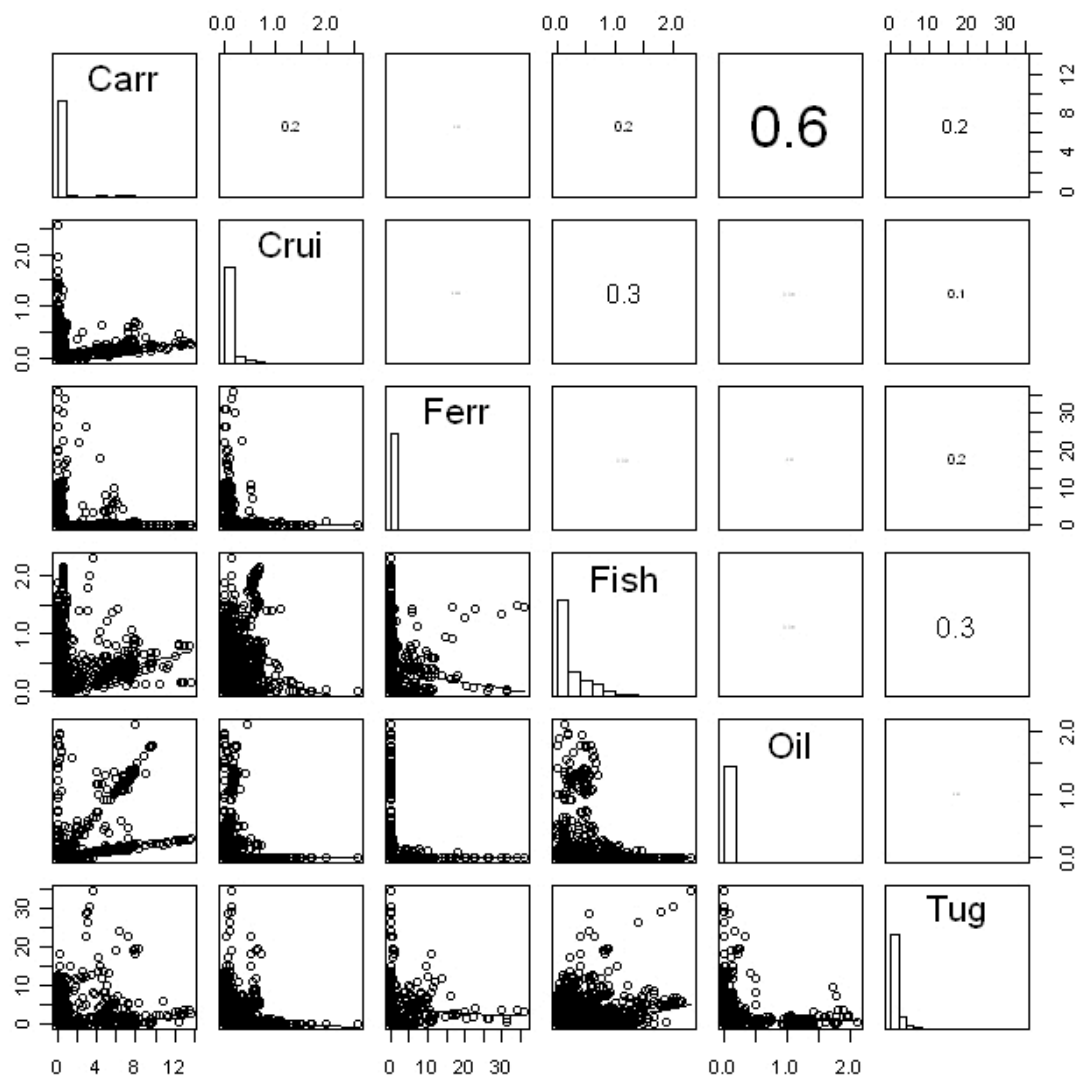


Figure C. 4. Scatterplot matrix of shipping traffic variables for Area B, including LOESS lines (lower left), histograms (on the diagonal) and correlation coefficients (upper right). Font size of correlation coefficients changes based on the value.

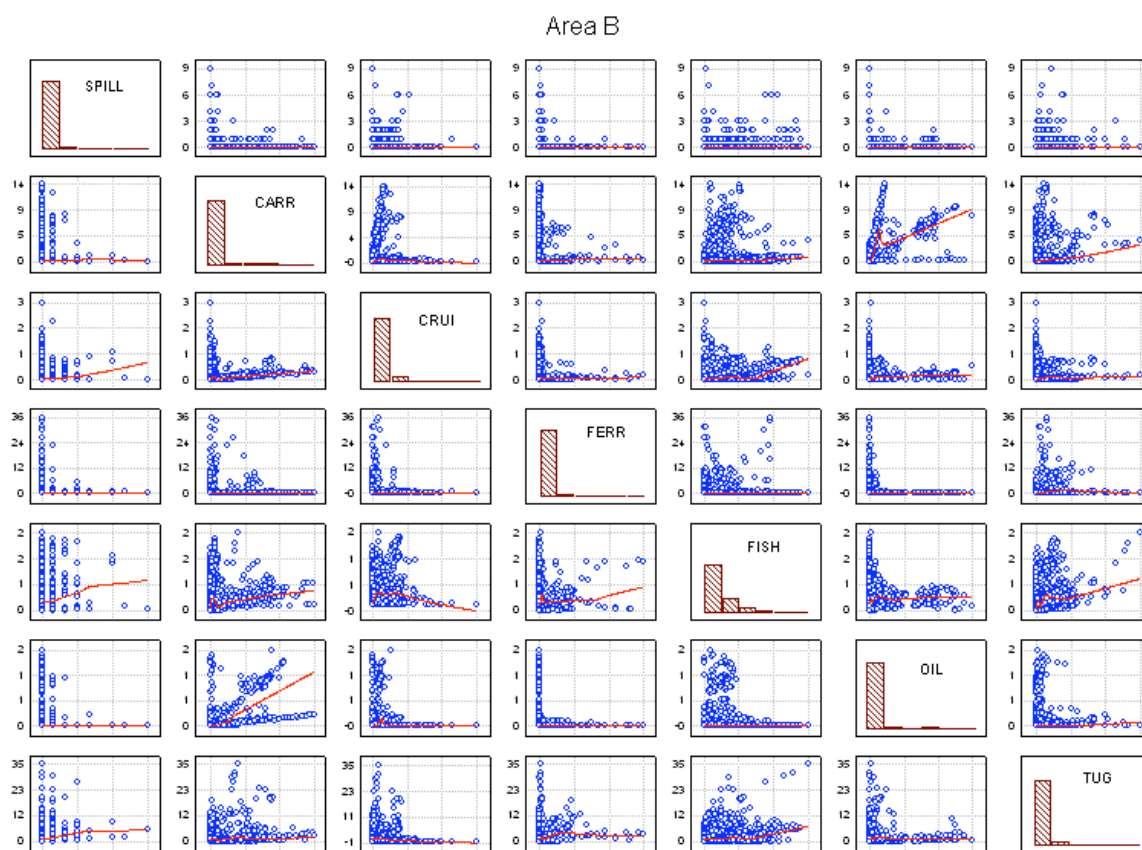


Figure C. 5. Matrix plots of Area B model variables spill counts and shipping traffic intensities for different vessel types, including LOESS lines and histograms on the diagonal.

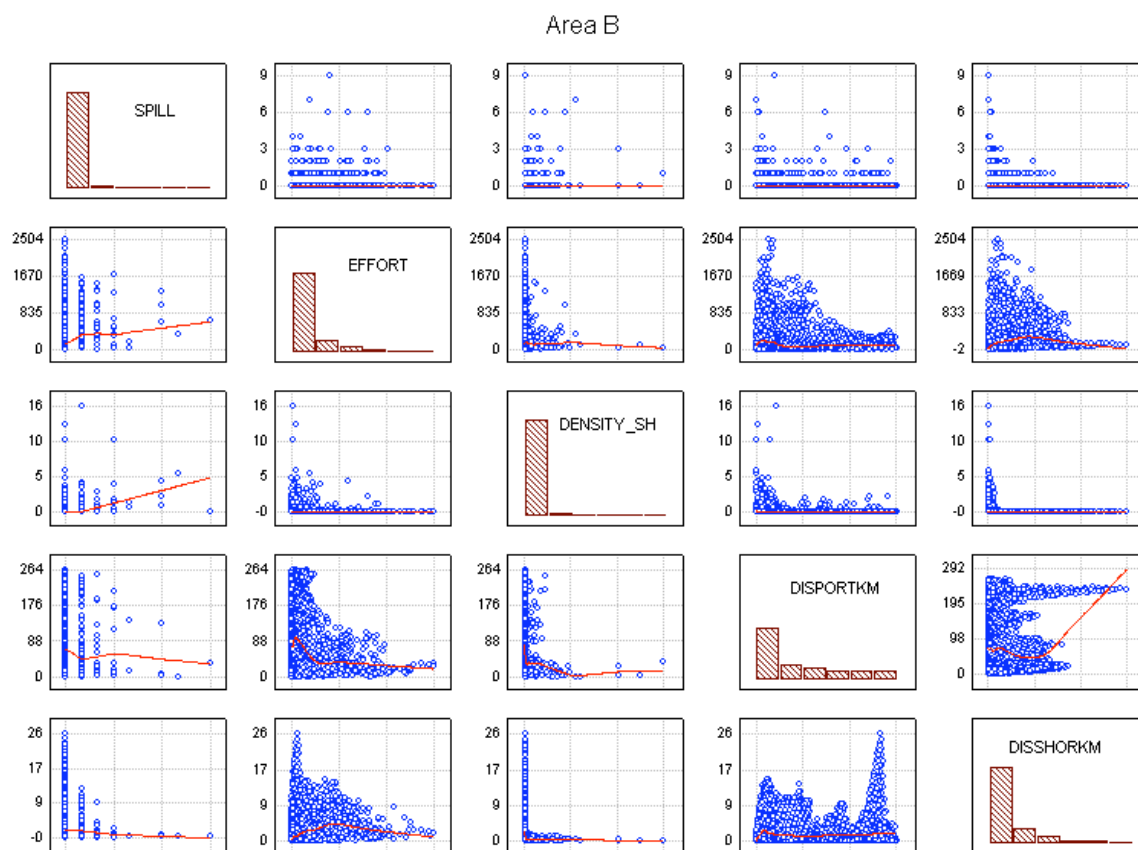


Figure C. 6. Matrix plots of Area B model variables spill counts, effort, density of small harbours, distance to shore and to port, including LOESS lines and histograms on the diagonal.

Table C. 1. Variable selection results from stepwise backward AIC and likelihood ratio test applied to generalized linear Poisson model for area A.

Dropped terms	df	AIC	Pr(χ^2)	(df, p-value)
<none>*	7	214.91		
<i>carroilfish_mean</i>	6	211.42	0.96	(1, $p > 0.05$)
<i>crui_mean</i>	6	213.12	0.64	(1, $p > 0.05$)
<i>ferr_mean</i>	6	213.25	0.56	(1, $p > 0.05$)
<i>tug_mean</i>	6	213.78	0.35	(1, $p > 0.05$)
disshorkm	6	226.95	2.85E-04	(1, $p < 0.05$)
count_sh	6	225.66	1.89E-04	(1, $p < 0.05$)

*Full model function: glm (spill ~ carroilfish_mean + crui_mean + ferr_mean + tug_mean + disportkm + disshorkm + count_sh + offset(logeffort), family=poisson(log))
Terms in italics were dropped from full model.

Table C. 2 Variable selection results from stepwise backward AIC and likelihood ratio test applied to logistic regression model for area A.

Dropped terms	df	AIC	Pr(χ^2)	(df, p-value)
<none>*	7	164.94		
<i>carroilfish_mean</i>	6	165.29	0.55	(1, $p > 0.05$)
<i>crui_mean</i>	6	165.15	0.64	(1, $p > 0.05$)
<i>ferr_mean</i>	6	165.89	0.33	(1, $p > 0.05$)
<i>tug_mean</i>	6	164.94	0.99	(1, $p > 0.05$)
disshorkm	6	193.35	9.775E-08	(1, $p < 0.05$)
count_sh	6	167.45	0.1125	(1, $p < 0.05$)

*Full model function: glm (spill ~ carroilfish_mean + crui_mean + ferr_mean + tug_mean + disportkm + disshorkm + count_sh + offset(logeffort), family=binomial)
Terms in italics were dropped from full model.

Table C. 3. Variable selection results from stepwise backward AIC and likelihood ratio test applied to generalized linear Poisson model* for area B.

Dropped terms	df	AIC	Pr(χ^2)	(df, p-value)
<none>*	9	1648.6		
carr_mean	8	1664.4	2.44 e-05	(1, p < 0.05)
crui_mean	8	1656.2	0.002	(1, p < 0.05)
<i>ferr_mean</i>	8	<i>1650.3</i>	0.06	<i>(1, p ≤ 0.05)</i>
<i>Log(fish_mean)</i>	8	<i>1646.8</i>	0.63	<i>(1, p > 0.05)</i>
log(tug_mean)	8	1673	2.82 e-07	(1, p < 0.05)
oil_mean	8	1654.4	0.005	(1, p < 0.05)
<i>Log(disportkm)</i>	8	<i>1647.2</i>	0.46	<i>(1, p > 0.05)</i>
log(disshorkm)	8	1717	2.20 e-16	(1, p < 0.05)
density_sh	8	1694	5.76 e-12	(1, p < 0.05)

*Full model function: glm (spill ~ carr_mean + crui_mean + ferr_mean + logfish + logtug + oil_mean + logdisport + logdisshor + density_sh + offset(logeffort), family=poisson(log))
Terms in italics were dropped from full model.

Table C. 4. Variable selection results from stepwise backward AIC and likelihood ratio test applied to logistic regression model* for area B.

Dropped terms	df	AIC	Pr(χ^2)	(df, p-value)
<none>*	9	1226.20		
carr_mean	8	1239.10	0.00	(1, p < 0.05)
crui_mean	8	1226.50	0.56	(1, p > 0.05)
<i>ferr_mean</i>	8	1229.90	0.05	<i>(1, p ≤ 0.05)</i>
<i>Log(fish_mean)</i>	8	1226.20	0.93	<i>(1, p > 0.05)</i>
log(tug_mean)	8	1236.70	0.00	(1, p < 0.05)
oil_mean	8	1232.20	0.01	(1, p < 0.05)
<i>Log(disportkm)</i>	8	1226.20	0.88	<i>(1, p > 0.05)</i>
log(disshorkm)	8	1280.70	0.00	(1, p < 0.05)
density_sh	8	1256.60	0.00	(1, p < 0.05)

*Full model function: glm (spill ~ carr_mean + crui_mean + ferr_mean + logfish + logtug + oil_mean + logdisport + logdisshor + density_sh + offset(logeffort), family=binomial)
Terms in italics were dropped from full model.