Fisheries and Oceans Canada
Ecosystems and Oceans Science

Pêches et Océans Canada
Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2022/012

Quebec Region

Seasonal occurrence and spatial distribution of four species of baleen whales vulnerable to ship strikes in the Saguenay–St. Lawrence Marine Park (Quebec, Canada)

C. C. A. Martins[1], S. Turgeon[1], R. Michaud[2] and N. Ménard[1]

[1]Parks Canada, Saguenay–St. Lawrence Marine Park
182 rue de l'Église
Tadoussac, Québec, G0T 2A0

[2]Groupe de recherche et d'éducation sur les mammifères marins
108 de la Cale-Sèche
Tadoussac, Québec, G0T 2A0

November 2022

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044
ISBN 978-0-660-45728-4 Cat. No. Fs70-5/2022-012E-PDF

Correct citation for this publication:

Martins, C.C.A., Turgeon, S., Michaud, R., and Ménard, N. 2022. Seasonal occurrence and spatial distribution of four species of baleen whales vulnerable to ship strikes in the Saguenay–St. Lawrence Marine Park (Quebec, Canada). DFO Can. Sci. Advis. Sec. Res. Doc. 2022/012. v + 62 p.

Aussi disponible en français :

Martins, C.C.A., Turgeon, S., Michaud, R., et Ménard, N. 2022. Occurrence saisonnière et répartition spatiale de quatre espèces de baleines à fanon vulnérables aux collisions avec les navires dans le parc marin du Saguenay–Saint-Laurent (Québec, Canada). Secr. can. des avis sci. du MPO. Doc. de rech. 2022/012. v + 64 p.

TABLE OF CONTENTS

ABSTRACT

Information on seasonal occurrence and fine scale spatial distribution of baleen whales is essential to design effective conservation measures. Given the intensity of navigation activities which overlap with whale habitat, the Saguenay–St. Lawrence Marine Park area is considered a high risk area for vessel collision with cetaceans. Management measures to reduce the risk of collision and minimize disturbance were included in the Marine Activities in the Saguenay–St. Lawrence Marine Park Regulations which regulate the activities of commercial and recreational boats in the park. Additionally, voluntary protection measures in and around the marine park were implemented in 2013 in collaboration with the shipping industry with the aim to minimize the risk of ship strikes with baleen whales and the impact of noise for the St. Lawrence Estuary beluga whales.

Datasets from different monitoring programmes were analyzed in order to characterize baleen whale occurrence and spatial distribution within the marine park and its surrounding waters. The modelling approach, applied to both line-transect and point-sampling databases, allowed us to identify areas of high predicted relative density of four baleen whale species: minke (*Balaenoptera acutorostrata*), fin (*Balaenoptera physalus*), blue (*Balaenoptera musculus*) and humpback whales (*Megaptera novaeangliae*). The use of spatial and fixed environmental variables allowed us to map the whale density over the space providing valuable information on each species distribution within the study area.

Overall, the potential core habitat of minke, fin and humpback whales was located within the 100 and 200 m isobaths, which includes the head of the Laurentian Channel, and its northern and southern submarine slopes. Blue whales' predicted core habitat was in the area downstream of the 200 m isobath in the center of the Laurentian Channel.

Using the combined datasets, the whale presence/absence analysis indicates that these four species use the area at least from early May to late October, with a main period of occurrence from early June to late September. Point sampling modelling results indicate that the occurrence of minke, humpback, and blue whales peaked from late July to early August, while fin whale relative abundance increased until the end of September. Relative abundance has fluctuated over the years for these species, and further analyses are required in order to understand the underlying mechanisms guiding the dynamics of this ecosystem.

These aggregated results provide valuable information for science-based cetacean habitat conservation. Furthermore, they provide support to evaluate the needs of adapting and developing new measures to reduce the risk and severity of whale collision with all boat categories within the Saguenay–St. Lawrence Marine Park and its surrounding waters.

INTRODUCTION

In the St. Lawrence Estuary (SLE) (Quebec, Canada), the waters in and around the Saguenay–St. Lawrence Marine Park (marine park, hereafter) are well known for the resident endangered beluga whale population (*Delphinapterus leucas*) and the wide diversity of marine mammal species that use the area as a seasonal feeding ground. In recognition of the ecological significance and conservation needs of whales in the SLE beyond the boundaries of the marine park, a marine protected area (MPA) is being proposed to increase the protection of whale species at risk (DFO 2020). The marine park is also well known for the variety and intensity of vessel-based activities (Turgeon 2019). The attractiveness of cetaceans for boat-based whale watching in the area, and the fact that the SLE is a major shipping route linking the upstream ports of the St. Lawrence-Great Lakes and Saguenay River to the rest of the world, can lead to voluntary or involuntary close encounters with whales, increasing the risk of collision. The Laurentian Channel (LC) (Figure 1), which is characterized by extensive seafloor slopes, is a tidal upwelling area (Saucier and Chassé 2000) where whales converge to feed. In this particular portion of the SLE, whale feeding areas overlap with vessel navigation routes. Furthermore, this coastal region is well-serviced by maritime infrastructures on both the north and south shores of the SLE, making areas used by cetaceans easily accessible by commercial whale watching and recreational boats.

Given the intensity of vessel-based activities which can overlap with whale habitat, the marine park can be considered a high-risk area of collision with cetaceans. Management measures to reduce the risk of collision and minimize disturbance were included in the Marine Activities in the Saguenay-St. Lawrence Marine Park Regulations (Canada 2002), specifically by limiting the number of whale watching permits, defining speed limits on observation zones, and establishing minimum distance requirements from whales. The regulations also made it obligatory to report all incidents of boat strikes/collisions to Parks Canada. However, the regulations were mostly aimed at reducing collisions with beluga and baleen whales involving commercial whale watching and recreational boats.

To extend protection measures to the shipping industry, a working group, consisting of representatives from the maritime industry, government, non-governmental organizations and academia, was formed in 2011 (Chion et al. 2018). As the risk of ship strikes to highly mobile beluga whales by merchant ships was considered to be low, the focus of the working group regarding ship strikes was on baleen whales. The working group thus defined voluntary protection measures (Figure 2) in and around the marine park to reduce the risk and severity of ship strikes with baleen whales, and to minimize the impact of shipping noise on beluga whales (Chion et al. 2018; DFO 2014). These measures, in effect from May to October (Canadian Coast Guard 2019), consist of 1) a 10 knots 'slow-down' area, which corresponds to an important whale feeding area at the head of the LC; 2) a 'no-go' area, which corresponds to an important area for blue whales; 3) a recommended route to avoid an area highly used by beluga whale herds composed of females, calves, and juveniles, and; 4) a caution area, which corresponds to the area of known occurrence of whales. These voluntary measures were considered to be provisional and would need to be adapted during the first years of implementation. The working group recognized the need to bring together the best available science on the spatial and temporal distribution of whales on which to base any review of these provisional voluntary measures.

The objective of this paper is to present information on the seasonal occurrence and fine-scale spatial distribution of the baleen whales vulnerable to ship strikes in the SLE, namely, minke (*Balaenoptera acutorostrata*), fin (*Balaenoptera physalus*), blue (*Balaenoptera musculus*) and humpback whales (*Megaptera novaeangliae*), specifically in the marine park and its surrounding waters. The present paper is complementary to the information presented by Mosnier et al. (2022), in which the Department of Fisheries and Oceans (DFO) aerial and boat-based systematic surveys for cetaceans in the SLE (1995-2017) were evaluated. Species distribution modelling (SDM) was chosen to characterise the distribution and relative abundance of baleen whales. SDM statistically relates distribution patterns to environmental conditions by linking animal observations to environmental variables. SDM have been widely used for a variety of purposes including to explain and predict distribution patterns of highly mobile marine animals in a variety of ecosystems (Guisan and Zimmerman 2000), to guide the designation of MPAs, to inform rerouting of major shipping lanes (*e.g.* Gregr et al. 2013; Redfern et al. 2013), to name a few.

Datasets from different monitoring programmes were used for the analyses. Systematic line transect (LT) distance sampling surveys conducted from 2006 to 2011 were used to build a SDM for each species and to estimate their relative abundance (Hammond 2010), or an estimate of the average number of individuals of each baleen whale species within marine park over the given period (2006-2011). A second dataset comes from the whale watching activities monitoring (WWAM) programme, which has been ongoing since 1994. WWAM consists of systematic sighting data, which are collected using whale watching boats as platforms of opportunity. This dataset, comprised of point samplings (PS), was also used to build a SDM for each baleen whale species and investigate seasonality of habitat use. Results from LT and PS models were compared and validated with a third independent dataset on whales' occurrence and distribution collected during systematic hydroacoustic surveys. The period of occurrence of the four species in the area was investigated through a combined analysis of the above-mentioned datasets and sightings data collected by naturalists as part of a citizen science (CSc) programme.

The outcome of the present exercise not only improves our understanding of the ecology of the target species, but also provides valuable information for science-based cetacean habitat conservation. Combined with the results provided by Mosnier et al. (2022), these results will enable the working group, co-presided by Parks Canada and DFO, to review the effectiveness of the protection measures put in place since 2013 in terms of geographical and temporal extent. These results will also be used to evaluate the need to adapt and develop new measures to reduce the risk and severity of collision with all categories of boats within the Saguenay–St. Lawrence Marine Park and its surrounding waters.

METHODS

SYSTEMATIC DISTANCE SAMPLING SURVEY

Survey area and period

The study area is located in the lower SLE portion of the Saguenay-St. Lawrence Marine Park and surrounding waters (Figure 3). Data were collected from mid-June to late September. Systematic surveys took place three times a week weather permitting from 2006 to 2011 and followed line transect (LT) distance sampling protocols (Buckland et al. 2015).

Survey design and data collection

Transect lines to survey baleen whales were designed to cover the head of the LC from the coastline to approximately the 100 m isobath. From 2006 to 2009, three zigzag schemes (transect 6, 7 and 8), composed of six equally angled lines and with different starting points (chosen at random) were established (Figure 3). A complete survey represented approximately

55 km and lines ranged from six to ten km of length. In 2010 and 2011, parallel lines were adopted (transect A, B, C, D, J, K and L). Each survey was composed of five lines 7 km apart. A complete survey ranged from 52 to 58 km depending on the transect configuration and lines ranged from 7 to 14 km in length. The order of execution of the survey lines (downstream to upstream of the estuary, or vice-versa) varied during the season depending on weather conditions. Line arrangement was modified from zig-zag to parallel, but data collection protocols remained the same regardless. A complete description of the field methods is provided by Martins (2012).

Field work was conducted by the research team of the *Groupe de Recherche et d'Éducation sur les Mammifères Marins (*GREMM). Surveys were carried out onboard the BpJam (Zodiac SRMN 600) at a constant speed of 15 kt. Planning meetings and training sessions were held before the beginning of the project in 2006 and before each field season. Within the same year only one observer was responsible for most surveys. The observer was placed on a fixed platform 1 m above the water level in front of the boat searching forward constantly and less often laterally (i.e. generally 45° either side of the centre). Due to the survey platform's limited height, the search area was limited to 2 km radius to avoid double counting whales and limit distance estimation errors. Angle of the detection in relation to the line was measured with a hand held electronic compass and distance was estimated visually. All observers calibrated their estimates of distance against known distance objects (measured with radar and or range finders) at the beginning of the surveys every year, and on a regular basis during the field season. Data were recorded on a hand-held voice recorder and transcribed afterwards. Species, group size, distance, angle of detection and general comments were registered at each sighting event. Survey effort was recorded with a hand held GPS Garmin Foretrex 301.

Data analysis

Meteorological conditions (Beaufort state, visibility, and cloud cover) were associated with each observation. Only sample days during which at least two transect lines were completed with good weather conditions were included in the analysis.

Each species' distribution was modelled using the count method (Hedley and Buckland 2004) following a two-stage approach (Miller et al. 2013): fitting a detection function first, and then given the detection function, fitting the spatial model. The R package *Distance* (Miller et al. 2019) (version 0.9.8) was used to fit a detection function for each species. A right truncation of 5% was applied to the distance data. Data from all years were pooled by species and the probability of detection was estimated as a function of perpendicular distance only. Half-normal and hazard-rate models were tested and the model that best fitted the data was selected according to the Akaike Information Criterion (AIC) (Buckland et al. 2001). The R package Density Surface Modelling (Miller et al. 2013) was used to model the number of counts of each species as a function of environmental variables.

Each resulting transect of length *L* was divided into segments *l* of two km. This resolution was chosen in order to ensure that there would be little variability in physical and environmental features within segments and to reduce the number of segments with zero sightings. Because transect lengths were irregular, the length of the final segment could vary slightly; no segments were smaller than one km or longer than three km, but irregular segments were rare. A central point was attributed to each segment and the whales' observations (number of groups of each species) recorded within the segment were assigned to that point. Each segment was then characterized by a set of environmental variables (Table 1, Figure 4). Longitude and latitude corresponded to the coordinate of the central point. Water depth and seafloor slope were obtained from a numeric model at a resolution of 500 m. Slope in degrees was obtained with the tool *Slope* of Spatial Analyst extension of ArcGIS (ESRI 2015). The mean and standard deviation of depth and slope were calculated using the midpoint of each segment and its surroundings within one km radius. Distance to the coast and to the main upwelling area were calculated as the shortest distance between the midpoint and the northern coast, and a line located at the head of the LC, respectively. They were obtained using the tool *Near* of the same extension. ArcGIS was used to perform geographic information system operations and to map the results. Data exploration followed the steps delineated by Zuur et al. (2009) and was performed using R (R Development Core Team 2017). The R package *usdm* (Naimi 2015) was used to calculate the variation inflation factor (VIF) index. A VIF > 3 indicates high collinearity and precludes the use of such variables in the same model.

A generalized additive model (GAM) (Hastie and Tibshirani 1990) was used to model the count as a sum of smooth functions of covariates (*zjk* with *k* indexing the spatial and environmental covariates) and predict a SDM for each species. The expected value of n_j was modelled as a function of the covariates as follows:

$$E(n_j) = \hat{p}_j A_j \exp\left[\beta_0 + \sum_k f_k z_{jk}\right]$$

where the segment area A_j ($2wl_j$) multiplied by the probability of detection $\hat{p}$ gives the effective area, $\beta_0$is the intercept and f_k are the parameters to be estimated.

The two-stage modelling framework was performed for each species, following four steps: (1) fit a detection function; (2) fit a GAM and perform model selection; (3) predict whale density throughout the study region, and (4) variance estimation and model validation.

Detection function

For the purpose of this analysis, the estimated probability of detection $\hat{p}_i$ was considered to be equal for all segments (*i.e.,* independent of observers, wave height, etc.). Previous analysis of the data, showed that these variables have little effect on the detection of these species (Martins 2012). If there are no covariates other than distance in the detection function, then the probability of detection is constant for all segments.

Model fit and model selection

Counts were modelled as a negative binomial distribution with a logarithmic link function, which is appropriate for count data containing numerous zeros. The segment area A_j multiplied by the probability of detection $\hat{p}$ was used as an offset. Besides a global model, year was included as a factor in order to obtain annual predictions of relative abundance. Models with and without interaction terms, and with and without data transformation (logarithm or square-root) were compared using AIC and percentage of deviance explained. If the estimated degrees of freedom was close to one (which indicates a linear relationship) the model was tested with this term as a linear predictor (Wood and Augustin 2002). Model residual plots were examined visually to verify the normality and variance homogeneity of the residuals.

Model prediction

The prediction grid was designed to cover the original systematic survey area and a buffer zone to minimize edge effect. A hexagon grid was chosen as they also minimize edge effect. Grid cell size was chosen to approximate the area of the offset, and was set to 4 km^2. The prediction grid was composed of 125 cells. The same environmental variables attributed to the segments were attributed to the prediction grid cells. The best global model (according with AIC) for each species was used to predict relative density of individuals (number of whales km^{-2}) for each grid cell. In the prediction maps, density was grouped in three classes: *low*: includes all values below

½ standard deviation of the mean; *mid*: includes the mean density and ± ½ standard deviation of the mean; and *high*: which includes all values above ½ standard deviation of the mean.

Variance estimation and model validation

Prediction uncertainty was accessed using GAM theory and the function *dsm.var.gam.* Gathered confidence intervals using this function are usually comparable to those obtained with their bootstraps equivalents (Miller et al. 2013). Spatial autocorrelation of model residuals was investigated through a variogram using the R package *gstat* v.2.0.2 (Pebesma 2004). Violation of the independence assumption was assessed by comparing the empirical variogram of residuals with the Monte Carlo envelope of the empirical variogram computed from 999 independent random permutations of the residuals (Diggle and Ribeiro 2007). This was achieved using the R package *variosig* (version 0.3) (Wang and Furrer 2020). A map of the coefficient of variation of the predicted relative density for each species was presented.

WHALE WATCHING ACTIVITIES MONITORING

Survey area and period

The whale watching activities monitoring (WWAM) programme covers the Saguenay–St. Lawrence Marine Park and its surrounding waters. The present analysis was limited to the portion within the marine park to allow comparisons with the systematic LT surveys. In general, data were collected from June to September from 1994 to 2018 although the sampling effort varied from year to year.

Survey design and data collection

Systematic data were collected by observers on board commercial whale watching boats operating within the marine park. A complete description of field methods is provided elsewhere (*e.g.* Michaud et al. 2010; Martins et al. 2018). Two main boat categories were used as platforms of opportunity: large boats (~78-112 feet long, 243- 689 passengers) and small boats (~30-50 feet long, 12-60 passengers). Field effort was organized to cover all week days, but the number of whale watching excursions sampled throughout a field season varied across years depending on boat type and homeport. The sampled excursions were always the one which departed at mid-day (1200 h to 1500 h). For each whale watching excursion, the observer recorded date, homeport, boat name and type, its track at one-minute precision, and performed a complete scan of the area within a 2 km radius of the boat at 10 minute intervals. For each scan, the number of individuals by species and the number of boats by categories were noted. These scans, labeled a point sampling (PS), lasted for three minutes. However, three PS of each excursion were extended to five minutes and were labeled long point sampling (LPS) (Figure 5). As each excursion lasted two to three hours (depending on boat category), the observer usually made 12 to 18 point samplings, the LPS was the 4th, the 7th, and the 10th for shorter excursions and the 5th, 8th, and the 11th for longer ones. Despite its duration, each PS was characterized by boat speed, observation conditions (visibility and wave height), number of groups (and group size) of all marine mammal species and boat activity. The last, was classified in four categories: 1) whale observation, 2) searching or displacement, 3) seal or bird observation, or 4) landscape observation. While in whale observation activity, the target species was also identified. Since 2009, a new variable indicates if the activity of the current PS is dependent (continuous) or independent (discontinuous) of the previous one (*i.e.,* if a fin whale was under observation during three consecutive PS, the last two will be identified as continuous PS).

Data preparation

The complete database was filtered to keep only the excursions from June to September (months with similar data collection effort) with good weather conditions departing from three homeports: Tadoussac, Les Bergeronnes and Les Escoumins. Additionally, only the PS with the boat in whale observation activity were used in the analysis. The later restriction guarantee that data used in the analysis is derived from PS obtained when the boat was still or almost still, and we assumed the point sampling covered a static area of a 2 km radius. To minimize double counts and data dependency, from 1994 to 2008 only the LPS were kept, and from 2009 to 2018 only discontinuous PS were kept in the analysis (Figure 6). The difference of duration (PS versus LPS) was not taken into account, and they are referred as PS hereafter.

Data analysis

As in other studies dealing with non-systematic survey designs, a grid based modelling framework was chosen (*e.g.* Canadas et al. 2005; Schleimer et al. 2019). The same hexagon grid described above was used but was limited to the lower SLE portion of the marine park and was composed of 89 cells. Depth (m) and slope (°) as presented in the above section were attributed to the grid cells. The R package *usdm* (Naimi 2015) was used to calculate the variation inflation factor (VIF) index and collinearity among variables. Survey effort was defined as the number of PS in a grid cell. Number of individuals by species and effort were pooled for each month and year, resulting in a maximum of 100 (4 months and 25 years) temporal sub-units per grid cell.

A GAM was used to model the number of individuals of each species observed within each grid cell as a sum of linear and smooth functions of covariates (*Zjk* with *k* indexing the spatial and environmental covariates) and predict a SDM for each species. The expected value of observations for a given grid cell n_j was modelled as a function of the covariates as follows:

$$E(n_j) = \exp\left[\beta_0 + \sum_k f_k z_{jk}\right] * Effort$$

where β_0 is the intercept and f_k are the parameters to be estimated. The effort was the number of times a grid cell was visited.

The grid base framework was performed for each species, following three steps described below: (1) fit a GAM and perform model selection, (2) predict whale relative abundance throughout the study region, and (3) variance estimation and model validation.

Model fit and model selection

The number of whales per grid cell was modelled as a negative binomial distribution with a logarithmic link function, which is appropriate for count data containing many zeros. Restricted maximum likelihood (REML) was used as a smoothing parameter and the gamma term was set to 1.4 to reduce over-fitting in cases with relatively few observations per variable. The logarithm of effort was included as an offset term in the model to account for the variability in spatial effort over the study area. GAMs were fitted with the R package *mgcv* (Wood 2017).

Models with and without interaction terms, and with and without data transformation (logarithm or square-root depending on the distribution) were compared using AIC and percentage of deviance explained. Aiming to investigate seasonality of habitat use pattern, the interaction between the variable month and the environmental variables was tested. Month and year were added to the model as factor. Year was grouped in four categories combining consecutive years (1994-1999, 2000-2005, 2006-2011, 2012-2018) to reduce the number of factors. The choice of grouping allowed to compare the shared period with LT surveys (2006-2011). Model residual plots were examined visually to verify the normality and variance homogeneity of the residuals.

Model prediction

The same grid was used to predict a relative abundance (individuals PS^{-1}) for each grid cell. The offset term has been fixed to 1 PS. For species in which the variable month in interaction with depth or slope was retained during model selection, a monthly density map was predicted. Because the model yielded separate predictions for each group of years, the relative abundance was averaged for the entire times series. In the prediction maps, density was grouped in three classes: *low*: includes all values below ½ standard deviation of the mean; *mid*: includes the mean density and ± ½ standard deviation of the mean; and *high*: which includes all values above ½ standard deviation of the mean.

Variance estimation and model validation

To assess prediction uncertainty, overall average coefficients of variation (CV) were calculated for each species based on random simulation. From the distributions of the final model coefficients, 1000 coefficient vectors were randomly simulated using the R package *MASS (*version 7.3-51.1) (Venables and Ripley 2002) and were used to generate 1000 predictions. Maps illustrating the CVs have been produced for each species.

Spatial autocorrelation of model residuals was investigated through a variogram using the R package *gstat* v.2.0.2 (Pebesma 2004). Violation of the independence assumption was assessed by comparing the empirical variogram of residuals with the Monte Carlo envelope of the empirical variogram computed from 999 independent random permutations of the residuals (Diggle and Ribeiro 2007). This was achieved using the R package *variosig* (version 0.3) (Wang and Furrer 2020).

SPECIES DISTRIBUTION MODEL COMPARISON AND VALIDATION

A visual validation of the SDM obtained with each dataset (LT and PS) was performed. An independent database of baleen whale sightings collected from 2009 to 2018 aboard the boat *L'Alliance* during systematic hydroacoustic surveys conducted by Parks Canada was overlaid with model predictions. The hydroacoustic survey followed zigzag and parallel transect lines (depending on the year) and observations of marine mammals within 1 km ahead of the boat (180° radius) were recorded systematically. The position of each observation was corrected using angle and distance of detection (estimated visually by a trained observer).

The evaluation of the adequacy of current voluntary protection measures to reduce the severity and risk of ship strikes with baleen whales was beyond the scope of the present report, but the measures were included in the final maps for information purposes.

Modelling results derived from each dataset (LT and PS) were compared in terms of variables retained and spatial predictions for each species. Although the values are not directly comparable, predictions resulted from LT (individuals km^{-2}) and PS (individuals PS^{-1}) were normalized and a linear regression analysis was performed. Only shared grid cells were kept for this analysis.

PERIOD OF OCCURRENCE

To complete the information regarding the period of occurrence of the baleen whales in the study area, a fourth database derived from a Citizen Science (CSc) programme was used. CSc data have been collected by experienced naturalists on board of large whale watching boats (~78-112 feet long, 243- 689 passengers) operating within the marine park since 2008. Citizen scientists record daily observations of marine mammals for each excursion, which include the minimum information recommended by the whale watching committee of the International Whaling Commission (Carlson et al. 2016) for CSc surveys: date, data collector name, boat name, trip start/end times, and sightings of marine mammals. The territory used by the whale watching boats participating in the CSc programme is restricted to the area in between Tadoussac and Les Bergeronnes (Martins et al. 2018). The CSc dataset was only used to extract information on daily presence of baleen whales in the area.

Presence, absence, and effort data from four datasets were grouped by week (from May to October) and were plotted by species for the years shared by most databases (2006 to 2018). LT distance sampling surveys (2006 - 2011) and hydroacoustic surveys (2009 - 2018) were grouped in a single class of systematic line transect surveys. The portrait of weekly presence/absence/effort of the three categories (Csc: Citizen Science; Sys: Systematic line transect surveys; WWAM: whale watching activities monitoring) were presented separately by year.

RESULTS

SYSTEMATIC DISTANCE SAMPLING SURVEY

A total of 7330.8 km (Figure 6) were covered during summer months from 2006 to 2011, distributed in 136 survey days. A total of 1,067 groups and 1,287 baleen whale sightings were recorded (Figure 7). Sightings in which species identification was not possible (33 groups) were not used in the analysis. Minke and fin whales were the most frequently sighted species. After splitting the transect lines, a total of 3,779 segments were available to build the SDM. Table 2 shows the number of detections used in the analysis for each species as well as the fitted detection probability and effective strip width (ESW). The hazard rate function was selected according to AIC to model the distance data of the four species (Figure 8, Figure 9, Figure 10, Figure 11).

The variogram of residuals for all baleen whale species is presented in the Appendix 1. It does not indicate the presence of spatial autocorrelation. The coefficient of variation of the predicted SDM for each species is presented in Appendix 2.

Minke whales

For minke whales, the best model included latitude, depth, standard deviation of slope (log transformed), and year (Table 3, Figure 12). Minke whale counts decreased linearly with an increasing latitude, or from up to downstream, were associated with abrupt slopes – grid cells with higher SD of slope, and peaked around 100 m depth. The best model according to the AIC explained 37.8% of the deviance. The predicted SDM for minke whales (Figure 16) highlighted core areas at the head of the LC and following the steep slopes. The predicted mean density was 0.468 individuals km^{-2}.

The average number of individuals estimated to be within the study area at any given time during this period was 58 minke whales (95%CI: 46 - 75, CV: 12.6%) (Table 4). The estimated relative abundance of minke whales fluctuated across years and was significantly (0.001) lower than 2006 (year of reference) in 2008.

Fin whales

For fin whales, the best model included longitude, depth, slope (log transformed) and year (Table 3, Figure 13). Fin whale counts decreased linearly with an increasing longitude, or from up to downstream, and were associated with moderate slopes, and depths between 80 and 150 m. Counts were lower in shallow and deep waters. The best model (i.e lowest AIC) explained 44.8% of the deviance. The predicted SDM for fin whales (Figure 16) highlights core areas at the head of the LC. The mean density was of 0.110 individuals km^{-2} and the average number of individuals estimated to be within the study area at any given time during this period was 14 fin whales (95% CI: 10-20, CV: 18.4%) (Table 4). As for minke whales, estimated relative abundance of fin whales fluctuated over years and was significantly (0.001) lower than 2006 in 2008 and 2011.

Blue whales

For blue whales, the best model included latitude, depth, standard deviation of slope (log transformed), and year (Table 3, Figure 14). Blue whale counts increased linearly with an increasing latitude, being higher in the downstream portion of the study area. They were associated with deep water but were lower in moderate slopes. The best model explained 19.4% of the deviance. The predicted SDM for blue whales (Figure 16) identified the species core area at the downstream portion of the study area, off Les Escoumins. The mean density of groups was estimated to be 0.014 individuals km^{-2} and the average number of individuals estimated to be within the study area at any given time during this period was two blue whales (95% CI: 1-3, CV: 28.5%) (Table 4). The estimated relative abundance of blue whales showed little variation within this period (any year was significantly different (0.001) from the reference).

Humpback whales

For humpback whales, the best GAM included longitude, depth, standard deviation of slope (log transformed), and year (Table 3, Figure 15). Humpback whale counts decreased linearly with an increasing longitude, being higher upstream of the study area, increased with depth, being higher in areas up to 150 m depth, and in areas with abrupt slopes. The best model according to the AIC explained 18.8% of the deviance. The predicted SDM for humpback whales (Figure 16) identified the species' core area at the head of the LC and along the northern slopes down to Les Bergeronnes. The mean density of groups was estimated to be 0.008 individuals km^{-2} with and the average number of individuals estimated to be within the study area at any given time during this period was one humpback whale (95% CI: 1-2, CV: 26.4%) (Table 4). The estimated relative abundance of humpback whales showed little variation within this period (all years were significantly different (0.001) from 2006, but not of each other).

WHALE WATCHING ACTIVITIES MONITORING

After data selection, 3121 PSs were available for analysis from 1596 excursions. Those excursions departed mainly from Tadoussac (71.7%) homeport. Effort at Les Bergeronnes (13.8%) and Les Escoumins (14.5%) were lower and relatively similar to each other (Figure 18). The spatial allocation of sampling effort (PS) is shown on Figure 18. After pooling of the monthly data, a total of 1531 temporal grid cells with effort were available.

The number of individuals observed per PS for each species are shown in Figure 19. The distributions are characterized by a very large number of zeros. For the minke whale, a total of 1434 PS with non-zero observations were available. These PS have total number of individuals observed ranging from 1 to 19 with a mean of 2.4. For the fin whale, a total of 2090 PS with non-zero observations were available. These PS have a total number of individuals observed ranging from 1 to 47 with a mean of 3.5. For the blue whale, a total of 239 PS with non-zero observations were available. These PS have total number of individuals observed ranging from 1 to 4 with a mean of 1.2. For the humpback whale, a total of 624 PS with non-zero observations were available. These PS have a number of individuals observed ranging from 1 to 4 with a mean of 1.3.

Minke whales

For minke whales, the best model included year, month, an interaction of month and depth, and the log-transformed slope in interaction with the month (Table 5). The best model explained 33.9% of the deviance. Model residuals respected the normality and variance homogeneity assumptions. Minke whale counts were highest in the shallower areas and decreased in waters deeper than 100 m and were associated with moderate slopes (Figure 20). Their relative abundance increased through the season and fluctuated across years. It was significantly ($p< 0.001$) lower during the 2000-2005 period than in 1994-1999. The predicted SDM for minke whales highlighted core areas at the head of the LC and followed the submarine slopes (Figure 21, Figure 28). Monthly predicted maps showed a similar spatial pattern with slight differences. SDM predictions had the highest CVs among grids in the downstream portion of the study area, in deep waters, where the predicted relative abundance was low (Appendix 3).

The variogram of residuals does not indicate the presence of spatial autocorrelation for any species (Appendix 4).

Fin whales

As for minke whales, the best model included year, month, an interaction of month and depth, and the log-transformed slope in interaction with the month (Table 5). The best model according to the AIC explained 42.5% of the deviance. Model residuals respected the normality and variance homogeneity assumptions. Fin whale counts peaked within the 50-100 m depth range and were associated with moderate slopes (Figure 22). Their relative abundance increased linearly through the season and decreased significantly across years. The three subsequent periods show significantly lower abundance than during 1994-199 period. As for the predicted SDM for minke whales, fin whale maps highlighted core areas at the head of the LC and following the submarine slopes (Figure 23, Figure 28). Monthly predicted maps show the same spatial pattern with slight differences. The highest CVs of the SDM were along the downstream portion of the study area and were generally lower than those of minke whales (Appendix 3).

Blue whales

For blue whales, the best model included year, month, slope and the depth in interaction with month (Table 5). The model according to the AIC explained 36.7% of the deviance. Model residuals respected the normality and variance homogeneity assumptions. Blue whale counts linearly increased with depth and abrupt slopes (Figure 24). Their relative abundance increased through the season and peaked late in July. Relative abundance fluctuated and was significantly higher in the 2000-2005 and 2012-2018 periods than in 1994-1999 period, and lower than it in the 2006-2011 period. In contrast to predicted areas for minke and fin whales, blue whale maps highlighted that core areas were in the deep area in the downstream portion of the study area (Figure 25, Figure 28). Monthly predicted maps show the same spatial pattern with slight differences. The highest CVs of the SDM were along the upstream portion of the study area and were generally lower than those of minke whales (Appendix 3).

Humpback whales

For humpback whales, the best model included year, month, slope and depth (Table 5). The interaction with month was non-significant and therefore no monthly prediction was possible. The best model according to the AIC explained 27.5% of the deviance. Model residuals respected the normality and variance homogeneity assumptions. Humpback whale counts were higher within the 100-200 m depth range, peaked at 150 m depth and decreased in steep slope areas (Figure 26). Their relative abundance increased though the season with a peak in July–August. Humpback whale relative abundance in the three subsequent periods were significantly higher than during the 1994-1999. Humpback whale maps highlighted that core areas were at the head of the LC and following the submarine slopes (Figure 27). The highest CVs were along the downstream portion of the study area and were generally lower than those of minke whales (Appendix 3).

SPECIES DISTRIBUTION MODEL COMPARISON AND VALIDATION

Overall the same variables were retained in order to model the four baleen whale species distributions within the study area. Depth and slope were retained in all models, while latitude was also retained for minke and blue whales, and longitude retained for fin and humpback whales. Spatial predictions derived from each model for the same species were quite similar (Figures 29-32). Spatial predictions derived from LT and PS for minke whales showed the greatest agreement. The area between the 100 and 200 m isobaths, which includes the head of the LC and the northern and southern steep slopes of the LC, appeared as an important predicted habitat for minke, fin and humpback whales. Despite the differences in spatial coverage of the downstream portion of the study area, predicted areas of mid- to high-densities of blue whales were mainly located downstream and deeper than the 200 m isobath. The steep slopes of the LC comprised in between 100-200 m isobaths were predicted as mid-to high-density areas for minke, fin and humpback whales. The southern steeper slopes were also identified as predicted core habitat for blue whales.

The overlap of observations from the systematic hydroacoustic survey with the predicted SDM derived from both datasets (LT and PS) provide an independent validation of the predicted core habitats (Figures 29a-32a). Visually, the overlap of observations is consistent for three of the four species: minke, fin and humpback whales. Data collection of this independent dataset is restricted to the lower SLE portion of the marine park, resulting in lesser overlap with predicted core habitat for blue whales. Despite the slight differences between the SDM derived from each dataset, the overlap with the provisional voluntary protection measures highlight the same overall result (Figures 29b-32b): the slow-down area, which is actually restricted to the marine park limits, overlaps most of the predicted core habitats identified for minke, fin and humpback whales. The predicted core habitat of blue whales derived from the LT model has some overlap with the downstream no-go area along the north shore.

The relative density values derived from each dataset cannot be directly compared, but the regression analysis presented in Figures 29c-32c show that for minke and blue whales, both models were able to identify low and high densities areas. PS models, however, were less efficient at identifying low density areas for fin and humpback whales.

PERIOD OF OCCURRENCE

From 2006 to 2018, a total of 258 days of systematic surveys were carried out when combining the systematic transect surveys and the hydroacoustic surveys. Data were collected by observers on whale watching boats (WWAM) on 1696 days within this period and on 2596 days by observers in the citizen science program. Data on weekly presence, absence and effort for each dataset covered the period from early May to late October (Figure 33). Data for the months of May and October are essentially derived from the citizen science database. Based on these datasets, minke whales were usually already present in the area early May and still present in late October at the end of CSc data collection period. Fin whales usually arrived late in May and were still present in late October. Blue whales' pattern of occurrence varied the most throughout the years, and was greatly improved by combining the different data sources. Overall, their presence in the study area was concentrated from July to September. Humpback whales' presence was recorded from early July to late September.

DISCUSSION

Potential core habitats of four baleen whale species were identified within, and downstream of, the Saguenay-St. Lawrence Marine Park. Spatial results presented here are in agreement with large-scale systematic surveys conducted by DFO (Mosnier et al. 2022) which identified areas of occurrence probability of baleen whales in the entire SLE, while our results provide relative density estimates for those same species. Our data collection was limited to the marine park boundaries and surrounding waters, but as shown by Mosnier et al. (2022), it corresponds to the main aggregation areas within the SLE for minke, fin, and humpback whales, corroborating the importance of the marine park as habitat for baleen whales. Our analysis of 25 years of fine-scale systematic data collection (WWAM) show that baleen whales' relative abundance fluctuates over time. This reinforces the importance of long-term, fine-scale monitoring to understand the underlying mechanisms guiding the dynamics of this ecosystem. The current scenario of ongoing global climate change emphasizes the need for a multidisciplinary approach to improve our understanding and management of this unique ecosystem. The presence of whales in the area is guided by prey availability and signals of global climate change (*e.g.* increase in water temperatures, reduced ice cover during the winter, decreasing concentration of oxygen in deeper water masses, increasing frequency of red tides), have already been detected in the area (e.g. Starr et al. 2017; Galbraith et al. 2019).

The results presented here provide essential information for managers and stakeholders of the Saguenay-St. Lawrence Marine Park and will be included in the Parks Canada monitoring programme. Together, our results and those presented by Mosnier et al. (2022) provide the best information for marine spatial planning to enhance cetacean habitat conservation in the SLE.

METHODOLOGICAL CONSIDERATIONS

Systematic distance sampling

The accuracy of the distance sampling method relies on three main assumptions: 1) animals on the transect line are detected with certainty, or that $g(0)=1$; 2) objects are detected at their initial location, or that no directed movement towards or against the survey platform is observed (i.e. whales are not attracted or repelled); and 3) angle and distance to recorded animals are as accurate as possible (Buckland et al. 2001). Martins (2012) includes an in-depth discussion on how the present LT survey methods meet these assumptions and to which degree results might be affected. For cetacean species, which spend most of their time underneath the water surface (availability bias), failure to meet the first assumption is common and causes negative biases in density estimates (Buckland et al. 2001). Known estimates of $g(0)$ for large whales due to perception bias (animals were at surface but were missed by observers) range from 0.9 -1 (Barlow 1995; Williams et al. 2006), while for minke whales it can be much lower. A study conducted in the North Atlantic estimated that 56-68% of minke whales were missed (Skaug and Schweder 1999). In the present LT survey, two factors might have influenced $g(0)$: the single observer (perception bias) and boat speed (availability bias). The method is usually applied with at least two observers, each scanning a different side of the transect line and doubling the effort over the line. Boat speed has a direct effect on animal availability (*i.e.* the time an animal is within the visual range of an observer). We assume that our estimates might be negatively biased mainly for minke whales due to the species' characteristics (*i.e.* smaller size, absence of visible blow).

The inspection of the perpendicular distance histograms (no grouped data) for all species does not indicate the presence of movement prior to detection. The detection function presented a shoulder, and even for the rare species where the primary issue was the low number of sightings to fit a detection function, movement prior to detection does not seem to be a problem. Angles were recorded with a hand-held electronic compass, but distances were estimated visually ('eyeballing'). Even if observers often calibrated their estimates with the boat radar, it is known that observers tend to round to convenient values despite the use of measurement instruments (Buckland et al. 2001). The height of the survey platform also has a direct effect on the detection range, which in its turn will vary depending on the species' behavior. Zerbini et al. (2006) reported mean radial distances of detected fin and humpback whales of 2.72 and 2.66 km, and of 1.3 km for minke whales from survey platforms at least 10 m above sea level. Here, search area was limited to 2 km to minimize double count of smaller species and obtain more accurate distance estimates. This procedure does not classify the methodology used as a true strip transect, but the estimates are somewhat biased as a consequence, as some animals, if detected beyond 2 km, have moved from their initial position. Based on the limits presented above, we assume this bias to be negligible. Furthermore, the effective strip width was smaller than 2 km for all species. Surveys from larger ships (with heights of around 10 m above the sea level) for large whales usually result in a higher ESW (Andriolo et al. 2010; Clapham et al. 2003; Zerbini et al. 2004).

The apexes of zigzags present a potential source of bias (Dawson et al. 2008). In addition to the chance of double counting the same animal in two consecutive lines, having recently made a sighting near the apex, the observer might subconsciously bias his/her sighting effort at the beginning of the next leg (Dawson et al. 2008). The number of minke whale sightings at the apexes of the zigzags was elevated mainly in zones close to the northern shore coastline near the steep slopes, where the species tends to concentrate to feed. The design was modified after the preliminary analysis of data from 2006 to 2008 surveys, and a parallel transect design was used in 2010 and 2011 to improve data collection and avoid this potential source of bias. Additionally, the southern limit of the lines was extended to include the area beyond the southern steep slopes.

In summary, the estimates provided here, which indicate the average number of individuals of each species using this area at any given time, are likely conservative and probably negatively biased. It's important to highlight that the marine park is only a part of the feeding range of most of these species. Photo-identification studies, which have been carried out since the late 1970s in the area, might provide in the near future accurate estimation of the number of the animals using this area, or the true population size. However, for management purposes, these relative abundance estimates provided an invaluable portrait for the marine park area.

Whale watching activities monitoring

WWAM is the longest systematic monitoring programme in place within the study area. Data collection follows a well-established protocol that has changed very little since beginning in 1994. The protocol was intended to address specific management questions (see Martins et al. 2018 for an overview) and its utilization for other types of analyses is possible when taking into account its limitations. One aspect discussed by Michaud et al. (1997, 2003) is that the area covered by each survey platform differs. Large boats departing from Tadoussac cover an area smaller than small boats departing from the same port, as the later moves faster and are able to reach areas further downstream. For some types of analyses, WWAM data are limited to a single platform to deal with this aspect (Martins et al. 2018). In the present analysis, the spatial effort was taken into account in the model, lessening the effect of the type of boat from which data were collected. This aspect is still relevant for the temporal analysis of presence and will be discussed further. Other relevant aspects that will be discussed here are: effect of the observer, perception and availability bias.

The observer responsible for data collection has changed across years, but was generally consistent within a given season and for each departure port. The information collected during a PS is based on the number of observations by species within 2 km radius and it is influenced by the observer's ability to detect animals and to estimate distance visually (in order to guarantee data collection covers the same surface and is restricted to 2 km). All observers were trained by an experienced technician, and since 2007 a range finder was used to estimate distances. When a range finder was not available, observers used the boat radar to calibrate their estimates during the training period and as often as possible. The variation of the search radius from one observer to another may also vary depending on weather conditions (e.g. Beaufort state may affect the ability to estimate distances visually). An observer that includes animals detected farther from 2 km positively bias the density of the given PS while another that restricts the area to less than 2 km negatively bias it. The observer was not taken into account in the present analysis, future analysis might explore their effect on estimated densities. Besides the observer's ability to estimate the search radius, some variation might exist among them in terms of perception (i.e. animals were at surface but were missed) and WWAM data may also be influenced by animals' availability.

Two steps were performed during data preparation to allow using data with comparable availability. The original data included PS collected during whole excursions, which was filtered to keep only those PS when the boat was in observation mode (i.e. still or almost still). This step was necessary to guarantee almost stationary PS. The second step aimed to eliminate the effect of pseudo-absences. As mentioned in the methods, PS can last 3 or 5 minutes. During a PS where the identified target species is a humpback whale, for example, it may happen that the animal is diving for the duration of the PS, but its presence is known within the search radius. PS lasting for 5 minutes allow more time for the whale to surface than 3 minutes PS. In order to correct for this, animals that were not at the surface during the time of the PS, but that were identified as the target species of the PS (i.e. the reason why the boat was on observation mode) were considered present for the given PS. This was applied for all species but have only modified the number of PS with presence of humpback, fin and blue whales, but not of minke whales. This correction was also necessary to make a distinction with areas of zero observations.

The results obtained with the analysis of the 25 years of data from the WWAM programme allowed us to produce a cartography of the baleen whale distribution sightings (average and within the season) within the marine park and has provided an indication of how their relative abundance has fluctuated over time. This is the longest systematic database on baleen whales in the SLE and the continuation of this monitoring programme within and in the surrounding waters of the marine park will provided valuable information to pursue necessary adaptive management measures of human activities which contribute to the conservation of whales.

SPECIES DISTRIBUTION MODELLING (SDM)

The modelling approach applied to LT and PS data allowed the identification of potential core habitats, or areas with high predicted relative density, for the four baleen whale species. The areas identified by each dataset are consistent, providing further support that the modelling approach was able to identify each species' core habitats within the study area. A visual comparison with results derived from systematic surveys conducted by DFO, which modelled occurrence probability for the same species, validates the results presented here. Areas of high predicted relative density correspond to the areas of higher occurrence probability of each species presented in Mosnier et al. (2022), validating the importance of the marine park as the main feeding area for fin, minke, and humpback whales within the SLE. According to their model, the main occurrence probability area of blue whales is contiguous but outside the marine park area, downstream of Les Escoumins, corroborating once again results obtained with LT and PS data for this species which covered the periphery of the main habitat of blue whales. Our results provide an estimate of relative density, or an average number of animals of a given species expected to be within the study area on a given day, within the season. LT relative abundance corresponds to the average number of individuals per square km, while PS relative abundance represents an average number of individuals per PS unit (2 km around the observer). Besides proving its usefulness as an index of relative abundance for long-term monitoring purposes, both can be also used to model the risk of vessel collisions with whales (e.g. Vanderlaan et al. 2008; Vanderlaan et al. 2009) or noise level exposure (e.g. Aulanier et al. 2016).

The use of spatial and fixed environmental variables allowed us to map whale density within the study area, providing valuable information of each species' distribution. Latitude and longitude were added as a proxy to other variables not taken into account into the model. Tweedy and quasipoisson distributions, which are often recommended for over-dispersed data, were tested but not retained during model selection. The selection of month in the interaction with fixed variables support the hypotheses that for some species, habitat use does vary slightly through the year and may differ between years. However, we suggest that monthly differences are mainly a function of the total number of individuals present in the area, i.e. a density-dependence pattern. Dynamic environmental variables that were not considered here (e.g., tide, sea-surface temperature) are known to play an important role for some cetacean species' habitat use patterns. Fin whales present a marked daily displacement pattern within the study area, which may be governed by the tidal cycle (Michaud and Giard 1997). At high tide, the animals tend to be concentrated along the slopes, while at the low tide they are more dispersed over the territory (PCA, unpublished data). A similar pattern was also reported in the Gulf of Maine (Johnston et al. 2005), where the species occurrence was associated with the flood tides. The inclusion of dynamic variables such as tide in the model would result in prediction maps for each tide condition; if the use of such dynamic information is required for management purposes, those variables could be integrated in the model.

Despite some spatial differences, the potential core habitat predicted for minke, fin, and humpback whales highlights the importance of the area between the 100 and 200 m isobaths, which includes the head of the Laurentian Channel, and its northern and southern steep slopes as key areas for these species. For blue whales, the potential core habitat covers the deeper areas mainly downstream from the 200 m isobath and around the steep slopes. The comparison of the results obtained from the two datasets (LT and PS) for each species is not straightforward due to methodological differences, although the regression analysis indicates that for minke and blue whales, both data sets were able to identify areas of low and high density while the WWAM model failed to predict low-density areas for the main targets of whale watching activities in the area (Martins et al. 2018), namely fin and humpback whales. The locations of these species within the marine park are known beforehand due to communication strategies used by the vessel captains (Chion et al. 2011). Moreover, PS data collection takes place during the second excursion of the day made by captains, who will generally return to areas visited earlier in the day. As a consequence, PS while in observation mode will tend to be biased towards areas highly used by targeted species.

The differences in densities estimates obtained using the LT and PS methods may also be due to the temporal extent of each data set, to the change of density over time (increase of humpback and decrease of fin whales) as well as to species behavior. Doniol-Valcroze et al. (2012) propose that the use of behavioural states could improve habitat use models, but such data is often lacking. LT were conducted from 2006 to 2011, and the relative abundance of humpback whales in the area has increased since 2011. For humpback whale WWAM data, the high number of zeros may also play an important role. Before 2005, the species was observed on only a few occasions, while after that year they were detected more frequently, and in some years it was the main target species for whale watching activities (Martins et al. 2018). For fin whales, the opposite situation has occurred. As presented here, fin whales were more abundant in the area before 2000, and their relative abundance decreased over time. For blue whales, both data sets agree, and identify areas of low density, although high PS density areas correspond to mid-density LT, which is likely attributed to the difference of the survey area between the two datasets, with PS been restricted to the marine park limits while LT covering at least a portion of the SLE identified as an area of high occurrence probability by Mosnier et al. (2022).

PREDICTED CORE HABITATS FOR EACH SPECIES

Minke whale

The relative density values derived from each data set cannot be directly compared, but both highlight the importance of the same habitats for minke whales. High density areas were concentrated around the steep slopes and the coastal area with density decreasing beyond 2 km from the north coast of the SLE. The portion downstream of the marine park boundary that was only sampled by LT data showed low density for this species. Mosnier et al. (2022) also showed that areas with higher probability of occurrence for the species were identified within and around the marine park.

Depth and slope were the main predictors of the species' potential core habitats. The preference of minke whales for steep slopes has been described for other feeding areas (Naud et al. 2003, Ingram et al. 2007). Depth and slope were the best predictors to explain the distribution of minke whales in the Bay of Fundy, where a preference for steep slopes and depths in between 100 – 200 m was observed (Ingram et al. 2007). In the waters around the Mingan Islands in the Gulf of St. Lawrence, minke whales were associated with steep slopes and underwater sand dunes (Naud et al. 2003). An analogous study conducted in Antarctic waters used latitude, longitude and depth as predictors for minke whales distribution (Williams et al. 2006).

Fin whale

For fin whales, potential core habitats predicted by the two analyses highlighted the area encompassed by the 100 and 200 m isobaths, which includes the head of the Laurentian Channel and its steep slopes. The potential core habitats identified in the present analysis are in agreement with other studies (Michaud and Giard 1997; Mosnier et al. 2022). Fin whale preference for steep slope contours within the SLE was first described by Sergeant (1977), who suggested it was a probable consequence of the high biological productivity of these areas due to tidal mixing. Waters of the intermediate layer rise from deep waters to near the surface over the sills with the high tide (Saucier and Chassé 2000). Worldwide, fin whales show the same pattern of association with areas that favor the accumulation of prey along depth gradients (Woodley and Gaskin 1996; Notarbartolo-Di-Sciara et al. 2003; Williams et al. 2006). Fin whales are the second largest whale species and require areas where they can obtain high levels of energy, often correlated to areas with dense prey aggregations (e.g. Acevedo-Gutierrez et al. 2002). The species is usually associated with upwelling systems or frontal areas (Fiedler et al. 1998; Palacios 1999; Hucke-Gaete et al. 2004; Branch et al. 2007; Doniol-Valcroze et al. 2007; Gill et al. 2011) and show a marked preference for steep slopes (e.g. Branch et al. 2007).

The area of steep slopes downstream of Isle Rouge showed the higher discrepancy between LT and PS models. This area was highly used by fin whales before 2000 (Michaud and Giard 1997; Michaud et al. 2003), also corresponding to years during which the species' relative density in the area was higher. Besides the arguments presented earlier to explain the differences between the models, the plasticity of the species regarding the availability of prey items might also be considered a factor to explain the spatial changes of distribution. The relative abundance of fin whales showed the highest fluctuation over the studied period, with a pronounced decrease across years.

Blue Whale

For blue whales, potential core habitats predicted by LT and PS models highlight the importance of the deeper portion of the Laurentian channel downstream of the 200 m isobath and of the steep slopes. Most of the potential core habitat predicted by the LT model is located at the border of the marine park. Martins (2012) conducted an extrapolation exercise based on LT data collected between 2006 to 2009 which also revealed potential habitats from Les Escoumins to Betsiamites, following the north shore coast. Results obtained by Mosnier et al. (2022) corroborates the predicted core habitats presented here, as well the above mentioned extrapolation exercise. Predicted core habitats located within the marine park boundaries are also in accordance with results found by Doniol-Valcroze et al. (2012); they predicted areas of high suitability for blue whales based on the analysis of 10 animals tracked with VHF. The high suitability areas (HS>0.8) identified in their study covered slope areas, as well as plateaus and deep areas over flat beds.

An in depth and collaborative effort was performed recently to identify important habitats for blue whales in the western North Atlantic based on whaling records; photo-identification studies; land, aerial and ship-based surveys; passive acoustics monitoring efforts; satellite and radio telemetry; ice entrapment reports; opportunistic reports; and species distribution modelling (see DFO 2018a). As a result, four important areas, one of which includes the Lower St. Lawrence Estuary and northwestern Gulf of St. Lawrence, were identified. Our results, and those presented by Mosnier et al. (2022), contribute to this recent effort to identify the species' habitat and enhance the conservation efforts for this endangered species.

Humpback whale

Despite the low number of sightings available to build the model, results obtained with LT and PS models were quite similar and highlighted the same areas identified by Mosnier et al. (2022). Data collected after the whaling period indicated that humpback whales occurred in very low number in the study area historically (Edds and Macfarlane 1987). Since the 1970s, rare observations of the species have been recorded almost every year, but it was only since 1999 that their presence within the marine park began to be more recurrent (Michaud et al. 2003). The North Atlantic humpback whale population is considered recovered from commercial whaling and total population size is estimated to be above the pre-whaling estimates (Reilly et al. 2008). The increase in abundance has probably contributed to the increased number of observations of humpback whales in the marine park. Fidelity to feeding areas has been described for the species, and is influenced by maternal transmission (Weinrich et al. 2006). Once the mother returns from the breeding area with her calf of the year, the feeding area, feeding style, and prey preference are transmitted. Another possible explanation for the low number of animals observed in the SLE prior to 1999 is that the animals using this feeding area may have been exclusively males, which may have kept the abundance low for a long period. The first record of a humpback whale calf within the marine park was in 2007 (Baleines en direct

2012). Their increasing occurrence in the area in recent years suggests the reoccupation of a pre-whaling feeding habitat.

SEASONAL OCCURRENCE

Baleen whales' occurrence in the SLE is considered to be seasonal, but it has been documented that some individuals of all these species overwinter in high latitude areas, and that some individuals may not migrate every year to lower latitudes (e.g. Clapham and Mead 1999). Presence/absence analysis using the combined datasets support that the four species use the area from at least early May to late October, with minke and fin whales arriving earlier than the other species. Data used in the present analysis comes from different sources and is limited to the months of May to October, but some species were already in the area when field efforts began or were still present when observations ended. Systematic surveys cover a shorter period and are punctual or spaced throughout the season. WWAM and CSc data were essential to complete the overall portrait of the seasonal occurrence of whales in the study area. WWAM data used for this analysis comes from boats departing from different ports (i.e. Tadoussac, Les Bergeronnes and Les Escoumins), but early (June) and late (September) in the season, only the large boats departing from Tadoussac are operational. As consequence, the presence information early and late in the season is limited to this boat-type and to the area comprised between Tadoussac and Les Bergeronnes (Martins et al. 2018). The same spatial limitation applies to CSc data, which is also collected on board of large boats departing from Tadoussac. Presence data for the month of May and October comes almost exclusively from CSc programme. Although limited in space and to a specific period of the year, the analysis presented here is representative of the area where minke, fin, and humpback whales concentrate and support the adequacy of the seasonal period determined for the current voluntary speed reduction measure, which is from early May to late October (Canadian Coast Guard 2019). Besides the presence analysis, WWAM PS model allowed the identification of peak periods of occurrence: minke and humpback whale observations being higher in late July and early August, while fin whale observations increased until the end of September. Blue whales' seasonal pattern of occurrence varied from year to year, but a peak was also detected in late July and early August. Such information may be used to plan park-rangers' presence in the area to monitor compliance to whale watching regulations. Long-term passive acoustic monitoring (PAM) analysis conducted from 2007 to 2017 indicates that the frequency of blue and fin whale calls peak from September to November in the vicinity of Les Escoumins (Simard and Roy 2018). Their PAM data also supports that whale-generated vocalizations are rare or absent during the ice season from mid-December to late March (Simard and Roy 2018). Additional PAM surveys over a wider portion of the SLE will bring light to the different species' seasonal occurrence.

Blue whales occur year-round in the GSL and almost year-round in the Lower St. Lawrence Estuary (except for a ~2-month winter absence in the most upstream area) (Simard et al. 2016). Based on the presence analysis, the blue whale pattern of occurrence within the marine park can be described as episodic or discontinuous, with localized incursions in the area within a season. Within their feeding range, blue whales are known to present a nomadic behavior (Mizroch et al. 1984; Sears et al. 1990) that is strongly related to the dynamics of formation and depletion of dense krill patches. This translates into low residency periods within restricted areas encompassed by their home range. However, animals show high site fidelity, often returning to the same feeding spots (Sears et al. 1990). Individuals tracked within the SLE spent on average 4.0±4.1 days (median 2.7 days) per area-restricted search patch in the lower SLE and 2.5±2.1 days (median 1.5 days) per patch in the northwestern GSL (Lesage et al. 2017).

CONCLUSION

We have presented data derived from different monitoring programmes in order to characterize the fine scale spatial distribution and seasonal occurrence of baleen whales within the lower St. Lawrence Estuary portion of the Saguenay-St. Lawrence Marine Park and surrounding waters. It was not in the scope of the SDM analysis to review habitat use requirements of the targeted species, but rather, to document the seasonal occurrence and spatial distribution of four species of baleen whales and identify potential core habitats for each species within the study area for management purposes. The SDMs presented here provide valuable information for the management of this portion of the territory. Together, our results and those presented by Mosnier et al. (2022) provide the best information for marine spatial planning to enhance cetacean habitat conservation in the SLE.

The results corroborate the importance of the head of the Laurentian Channel as a main feeding aggregation for minke, fin and humpback whales. Blue whales' predicted core areas were essentially at the downstream boundary of the marine park and within the proposed St. Lawrence Estuary Marine Protected Area. These results provide valuable information needed to improve the provisional measures aimed at reducing the risk and severity of vessel strikes with baleen whales (Canadian Coast Gard 2019) and to minimize the impact of shipping noise on beluga whales (Chion et al. 2018; DFO 2014). These voluntary provisional measures apply specifically to the merchant fleet and to cruise ships, the merchant fleet represents one of the most frequent marine traffic categories in the marine park, with 12.5 (min = 2, max= 24) daily transits and a total of 4 545 transits in 2017 (Turgeon 2019). A total of 255 domestic or international cruise ships were carried out in the Marine Park in 2017, 69.4% of them took place in September and October. Furthermore, results may guide the development of measures aiming to decrease the exposure of baleen whales to other marine traffic categories, such as the whale watching industry, for example. The whale watching industry is another important component of the marine traffic volume within the marine park with a total of 6658 excursions for the year of 2017 (Turgeon 2019).

The results highlight that the current spatial management instruments in place, namely the area covered by the Saguenay-St. Lawrence Marine Park and the provisional voluntary measures do not sufficiently ensure the protection of the endangered blue whale. Due to their discontinuous occurrence pattern, any protection measure aiming to reduce the risks and severity of ship strikes as well as exposure to shipping noise for this species should follow their dynamic concentration areas instead of using a fixed area. Dynamic measures, such as those in place to protect NARWs in Canadian waters (DFO 2018b), could be considered to enhance the conservation of blue whales in the St. Lawrence Estuary.

ACKNOWLEDGEMENTS

Part of the analysis presented here was conducted during the Ph.D. program of C.C.A.M. that was made possible through studentships provided by Natural Science and Engineering Research Council of Canada (NSERC) and Université de Montréal. Authors are thankful to Lael Parrott, Danielle Marceau, Guy Cantin and Suzan Dionne for their implication in the context of this doctorate thesis. The authors would like to thank Parks Canada employees at the Saguenay–St. Lawrence Marine Park and the Group for Research and Education on Marine Mammals (GREMM) for their participation, at various levels, in setting up and conducting these monitoring programmes. We are especially grateful to Michel Moisan, Marie-Hélène D'Arcy, Sarah Duquette, Renaud Pintiaux. Chantal St. Hilaire, Agathe Poitras and Catherine Dubé were responsible by most data collected throughout the Citizen Science programme. A special thanks to whale watching companies for their participation in the whale watching activities monitoring and to all technicians involved in the different monitoring programmes over the years. Data collection and analysis was financed by Parks Canada, partly through the Conservation and Restauration Program at the Saguenay-St. Lawrence Marine Park.

REFERENCES CITED

Acevedo-Gutiérrez, A., Croll, D.A. and Tershy, B.R. 2002. High feeding costs limit dive time in the largest whales. J. Exp. Biol. 205(12):1747-1753.

Andriolo, A., Kinas, P.G., Engel, M.H., Martins, C.C.A. and Rufino, A.M. 2010. Humpback whales within the Brazilian breeding ground: distribution and population size estimate. Endanger. Species Res. 11(3): 233-243.

Aulanier, F., Simard, Y., Roy, N., Bandet, M. and Gervaise, C. 2016. Groundtruthed probabilistic shipping noise modeling and mapping: Application to blue whale habitat in the Gulf of St. Lawrence. In Proceedings of Meetings on Acoustics 4ENAL 27(1):070006.

Baleines en direct. 2012. Un deuxième baleineau pour Tic Tac Toe. Retrieved October 1st, 2012.

Barlow, J. 1995. The abundance of cetaceans in California waters. Part I: Ship surveys in summer and fall of 1991. Fish. Bull. 93: 91-14.

Branch, T.A., Stafford, K.M., Palacios, D.M., Allison, C., Bannister, J.L., Burton, C.L.K., Cabrera, E., Carlson, C.A., Galletti Vernazzani, B., Gill, P.C. and Hucke-Gaete, R. 2007. Past and present distribution, densities and movements of blue whales *Balaenoptera musculus* in the Southern Hemisphere and northern Indian Ocean. Mammal Rev. 37(2):116-175.

Buckland S.T., Anderson, D. R., Burnham, K. P., Borchers, D. L., and Thomas, L. 2001. Introduction to Distance Sampling. Oxford University Press, Oxford, UK.

Buckland, S. T., Rexstad, E. A., Marques, T. A., and Oedekoven, C. S. 2015. Distance Sampling: Methods and Applications. Methods in Statistical Ecology. Springer International Publishing.

Canada. 2002. Marine Activities in the Saguenay–St. Lawrence Marine Park Regulations.

Canadas, A. R., Sagarminaga, R. de Stephanis, E. Urquiola and Hammond, P. S. 2005. Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. Aquat. Conserv. Mar. Freshwat. Ecosyst. 15:495-521.

Canadian Coast Guard. 2019. Notices to Mariners Publication–Eastern Edition [online]. Fisheries and Oceans Canada.

Carlson, C., Kaufman, G., Ritter, F., Rodriguez-Fonseca, J., Robbins, J., Rose, N., Petersen, K. V., and Weinrich, M. 2016. Guiding principles for data collection from whale-watching vessels as platforms of opportunity. Document SC66b/WW3. International Whaling Commission Scientific Committee.

Chion, C., Lamontagne, P., Turgeon, S., Parrott, L., Landry, J.A., Marceau, D.J., Martins, C.C.A., Michaud, R., Ménard, N., Cantin, G. and Dionne, S. 2011. Eliciting cognitive processes underlying patterns of human–wildlife interactions for agent-based modelling. Ecol. Model. 222(14): 2213-2226.

Chion, C., Turgeon, S., Cantin, G., Michaud, R., Ménard, N., Lesage, V., Parrott, L., Beaufils, P., Clermont, Y., and Gravel, C. 2018. A voluntary conservation agreement reduces the risks of lethal collisions between ships and whales in the St. Lawrence Estuary (Québec, Canada): From co-construction to monitoring compliance and assessing effectiveness. PloS one 13(9): e0202560.

Clapham, P. J. and Mead, J. G. 1999. *Megaptera novaeangliae*. Mamm. Species: 1- 9.

Clapham, P., Barlow, J., Bessinger, M., Cole, T., Matila, D., Pace, R., Palka, D., Robbins, J. and Seton, R. 2003. Abundance and demographic parameters of humpback whales from the Gulf of Maine, and stock definition relative to the Scotian Shelf. J. Cetacean Res. Manage. 5(1): 13-22.

Dawson, S., Wade, P., Slooten, E. and Barlow, J.A.Y. 2008. Design and field methods for sighting surveys of cetaceans in coastal and riverine habitats. Mammal Rev. 38: 19-49.

Diggle, P. J. and Ribeiro, P. J. Jr. 2007. Model-based Geostatistics. Springer Series in Statistics. New York, Springer.

DFO. 2014. Impacts of rerouting marine traffic in the St. Lawrence estuary on beluga (*Delphinapterus leucas*): Science in support of risk management. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2014/004.

DFO. 2018a. Identification of habitats important to the blue whale in the western North Atlantic. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2018/003.

DFO. 2018b. Science Advice on Timing of the Mandatory Slow-down Zone for Shipping Traffic in the Gulf of St. Lawrence to Protect the North Atlantic Right Whale. DFO Can. Sci. Advis. Sec. Sci. Resp. 2017/042.

DFO. 2020. Protection of the marine environment under the Canada-Quebec Collaborative Agreement. Last modification 26th November 2020.

Doniol-Valcroze, T., Berteaux, D., Larouche, P. and Sears, R. 2007. Influence of thermal fronts on habitat selection by four rorqual whale species in the Gulf of St. Lawrence. Mar. Ecol. Prog. Ser. 335: 207-216.

Doniol-Valcroze, T., Lesage, V., Giard, J. and Michaud, R. 2012. Challenges in marine mammal habitat modelling: evidence of multiple foraging habitats from the identification of feeding events in blue whales. Endanger. Species Res. 17(3): 255-268.

Edds, P. and Macfarlane, J. 1987. Occurrence and general behavior of balaenopterid cetaceans summering in the St. Lawrence Estuary, Canada. Can. J Zool. 65 (6): 1363-1376.

[ESRI] Environmental Systems Research Institute. 2015. Spatial Analyst, 3D Analyst. ArcGIS version 10.3.

Fiedler, P.C., Reilly, S.B., Hewitt, R.P., Demer, D., Philbrick, V.A., Smith, S., Armstrong, W., Croll, D.A., Tershy, B.R. and Mate, B.R. 1998. Blue whale habitat and prey in the California Channel Islands. Deep Sea Res 2 Top Stud Oceanogr 45(8-9): 1781-1801.

Galbraith, P.S., Chassé, J., Caverhill, C., Nicot, P., Gilbert, D., Lefaivre, D. and Lafleur, C. 2019. Physical Oceanographic Conditions in the Gulf of St. Lawrence during 2018. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/046. iv + 79 p.

Gill, P.C., Morrice, M.G., Page, B., Pirzl, R., Levings, A.H. and Coyne, M. 2011.Blue whale habitat selection and within-season distribution in a regional upwelling system off southern Australia. Mar. Ecol. Prog. Ser. 421: 243-263.

Gregr, E.J., M. F. Baumgartner, K. L. Laidre and Palacios, D. M. 2013. Marine mammal habitat models come to age: the emergence of ecological and management relevance. Endanger. Species Res. 22:2015-212.

Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135(2-3):147-186.

Hammond, P. S. 2010. Estimating the abundance of marine mammals. In Marine Mammal Ecology and Conservation. Oxford University press, New York, pp. 42-67.

Hastie, T. J. and Tibshirani, R. J. 1990. Generalized additive models. Chapman and Hall/CRC, New York.

Hedley S.L. and Buckland, S. T. 2004. Spatial Models for Line Transect Sampling. J. Agric. Biol. Envir. S. 9(2):181–199.

Hucke-Gaete, R., Osman, L.P., Moreno, C.A., Findlay, K.P. and Ljungblad, D.K. 2004. Discovery of a blue whale feeding and nursing ground in southern Chile. Proc. R. Soc. Lond. Ser. B Biol. Sci. 271(suppl 4): 170-173.

Ingram, S.N., Walshe, L., Johnston, D., and Rogan, E. 2007. Habitat partitioning and the influence of benthic topography and oceanography on the distribution of fin and minke whales in the Bay of Fundy, Canada. J. Mar. Biolog. Assoc. U. K. 87(1):149-156.

Johnston D.W., Thorne, L.H. and Read A.J. 2005. Fin whales *Balaenoptera physalus* and minke whales *Balaenoptera acutorostrata* exploit a tidally driven island wake ecosystem in the Bay of Fundy. Mar. Ecol. Prog. Ser. 305: 287-95.

Lesage, V., Gavrilchuk, K., Andrews, R.D. and Sears, R. 2017. Foraging areas, migratory movements and winter destinations of blue whales from the western North Atlantic. Endanger. Species Res. 34: 27-43.

Martins, C.C. A. 2012. Study of baleen whales' ecology and interaction with maritime traffic activities to support management of a complex socio-ecological system. Ph. D.Thesis, Université de Montréal, 236 p.

Martins, C. C. A., Turgeon, S., Michaud, R. and Ménard, N. 2018. Suivi des espèces ciblées par les activités d'observation en mer dans le parc marin du Saguenay–Saint-Laurent de 1994 à 2017. Nat. Can. 142 (2): 65-79.

Michaud, R. and Giard, J. 1997. Les rorquals communs et les activités d'observation en mer dans l'estuaire du Saint-Laurent entre 1994 et 1996 : 1. Étude de l'utilisation du territoire et évaluation de l'exposition aux activités d'observation à l'aide de la télémétrie VHF. Projet réalisé dans le cadre d'une entente d'entreprise conjointe Groupe de recherche et d'éducation sur le milieu marin (GREMM), ministère de l'Environnement et de la Faune du Québec, ministère des Pêches et Océans, ministère du Patrimoine canadien, Parcs Canada, 30 p.

Michaud, R., Bédard, C., Mingelbier, M. and Gilbert, M.-C. 1997. Les activités d'observation en mer des cétacés dans l'estuaire maritime du Saint-Laurent 1985-1996 : Une étude de la répartition spatiale des activités et des facteurs favorisant la concentration des bateaux sur les sites d'observation. Rapport final, GREMM, Tadoussac, QC. 18 p.

Michaud, R., de la Chenelière, V. and Moisan, M. 2003. Les activités d'observation en mer des cétacés dans le parc marin du Saguenay–Saint-Laurent 1994-2002 : Une étude de la répartition spatiale des activités et des facteurs favorisant la concentration des bateaux sur les sites d'observation. GREMM, Tadoussac, QC, conjointement avec le parc marin du Saguenay–Saint-Laurent, Québec. 16 p.

Michaud, R., Moisan, M., de la Chenelière, V., Duquette, S. and D'Arcy, M.-H. 2010. Les activités d'observation en mer des mammifères marins (AO3M) dans l'estuaire du Saint-Laurent : zone de protection marine Estuaire du Saint-Laurent et parc marin du Saguenay–Saint-Laurent – Portrait 2005-2010. Rapport final. Groupe de recherche et d'éducation sur les mammifères marins, Tadoussac, 34 p.

Miller, D. L., Burt., M. L., Rexstad, E. A. and Thomas, L. 2013. Spatial models for distance sampling data: recent developments and future directions. Methods Ecol. Evol. 4(11):1001-1010.

Miller D.L., Rexstad, E., Thomas, L., Marshall, L. and Laake, J. L. 2019. Distance Sampling in R. J. Stat. Softw. 89(1):1-28.

Mizroch, S., Rice, D., and Breiwick, J. 1984. The Blue Whale, *Balaenoptera musculus*. Mar. Fish. Rev. 46.

Mosnier, A., Gosselin, J.-F and Lesage, V. 2022. Seasonal distribution and concentration of four large whale species in the St Lawrence Estuary based on 22 years of DFO observation data. DFO Can. Sci. Advis. Sec. Res. Doc. 2020/053. iv + 119 p.

Naimi, B. 2015. usdm: Uncertainty analysis for species distribution models. R package version, 1, pp.1-12.

Naud, M.J., Long, B., Brêthes, J.C. and Sears, R. 2003. Influences of underwater bottom topography and geomorphology on minke whale (*Balaenoptera acutorostrata*) distribution in the Mingan Islands (Canada). J. Mar. Biolog. Assoc. U. K. 83(4):889.

Notarbartolo-Di-Sciara, G., Zanardelli, M., Jahoda, M., Panigada, S. and Airoldi, S. 2003. The fin whale *Balaenoptera physalus* (L. 1758) in the Mediterranean Sea. Mammal Rev. 33(2): 105-150.

Palacios, D. M. 1999. Blue whale (*Balaenoptera musculus*) occurrence off the Galapagos Islands, 1978-1995. J. Cetacean Res. Manage. 1(1): 41-51.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. Comput. Geosci. 30: 683-691.

R Development Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

Redfern, J. V., McKenna, M. F., Moore, T. J., Calambokidis, J., Deangelis, M. L., Becker, E. A., Barlow, J., Forney, K. A., Fiedler, P. C. and Chivers, S. J. 2013. Assessing the risk of ships striking large whales in marine spatial planning. Conserv. Biol. 27(2): 292-302.

Reilly, S.B., Bannister, J.L., Best, P.B., Brown, M., Brownell Jr, R.L., Butterworth, D.S., Clapham, P.J., Cooke, J., Donovan, G.P., Urbán, J. and Zerbini, A.N. 2008. *Balaenoptera musculus*. IUCN Red List of Threatened Species. Gland, Switzerland: IUCN.

Saucier, F. J. and Chassé, J. 2000. Tidal circulation and buoyancy effects in the St. Lawrence Estuary. Atmos.-Ocean 38(4): 505-556.

Schleimer, A., Ramp, C., Plourde, S., Lehoux, C., Sears, R. and Hammond, P. S. 2019. Spatio-temporal patterns in fin whale *Balaenoptera physalus* habitat use in the northern Gulf of St. Lawrence. Mar. Ecol. Prog. Ser. 623: 221-234.

Sears, R., J.M. Williamson, F.W. Wenzel, M. Bérubé, D. Gendron, and P. Jones. 1990. Photographic identification of the blue whale (*Balaenoptera musculus*) in the Gulf of the St. Lawrence, Canada. Rep. Int. Whal. Comm. (Special Issue 12):335-342.

Sergeant, D.E. 1977. Stocks of fin whales, *Balaenoptera physalus*, in the North Atlantic Ocean. Rep. Int. Whal. Comm. 27: 460-473.

Simard, Y., Roy, N., Aulanier, F. and Giard, S. 2016. Blue whale continuous frequentations of St. Lawrence habitats from multi-year PAM series. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/091. v + 14 p.

Simard, Y and Roy, N. 2018. Une décennie de suivis acoustiques continus des rorquals bleus, des rorquals communs et du krill dans le parc marin du Saguenay–Saint-Laurent de 2007 à 2017. Nat. Can. 142 (2): 106-114.

Skaug, H. J. and Schweder, T. 1999. Hazard models for line transect surveys with independent observers. Biometrics 55: 29-36.

Starr, M., Lair, S., Michaud, S., Scarratt, M., Quilliam, M, Lefaivre, D., Robert, M., Wotherspoon, A., Michaud, R., Ménard, N., Sauvé, G., Lessard, S., Béland P., and Measures, L. 2017. Multispecies Mass Mortality of Marine Fauna Linked to a Toxic Dinoflagellate Bloom. PLoS ONE 12(5): e0176299.

Turgeon, S. 2019. Characterization of Navigation in the Saguenay–St. Lawrence Marine Park-2017. Parks Canada, 58 pages + annexes.

Vanderlaan, A.S., Taggart, C.T., Serdynska, A.R., Kenney, R.D. and Brown, M.W. 2008. Reducing the risk of lethal encounters: vessels and right whales in the Bay of Fundy and on the Scotian Shelf. Endanger. Species Res. 4(3): 283-297.

Vanderlaan, A. S. M., Corbett, J. J., Green, S. L., Callahan, J. A., Wang, C., Kenney, R. D., Taggart, C. T. and Firestone, J. 2009.Probability and mitigation of vessel encounter with North Atlantic right whales. Endanger. Species Res. 6:273–295.

Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New York.

Wang, C. and Furrer, R. 2020. Package ‘variosig’.

Weinrich, M.T., Rosenbaum, H., Scott Baker, C., Blackmer, A.L. and Whitehead, H. 2006. The influence of maternal lineages on social affiliations among humpback whales (*Megaptera novaeangliae*) on their feeding grounds in the southern Gulf of Maine. J. Hered. 97(3): 226-234.

Williams, R., Hedley, S. L. and Hammond, P.S.I. 2006. Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. Ecol. Soc. 11(1): 1

Wood, S. 2017. Generalized Additive Models: an introduction with R, Second Edition (2nd ed.). Chapman and hall/CRC.

Wood, S. N. and Augustin. N. H. 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecol. Model. 157:157-177.

Woodley, T. H. and Gaskin, D. E. 1996. Environmental characteristics of North Atlantic right and fin whale habitat in the lower Bay of Fundy, Canada. Can. J. Zool. 74(1): 75-84.

Zerbini, A.N., Andriolo, A., Da Rocha, J.M., Simões-Lopes, P.C., Siciliano, S., Pizzorno, J.L., Waite, J.M., Demaster, D.P. and Vanblaricom, G. R. 2004. Winter distribution and abundance of humpback whales (*Megaptera novaeangliae*) off Northeastern Brazil. J. Cetacean Res. Manage. 6(1): 101-107.

Zerbini, A. N., Waite, J. M., Laake, J. L. and Wade, P. R. 2006. Abundance, trends and distribution of baleen whales off Western Alaska and the central Aleutian Islands. Deep-sea Res. 53: 1772-1790.

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A. and Smith, G. M. 2009. Mixed effects models and extensions in Ecology with R. 1st Edition, New York, Springer.

TABLES

Table 1. Definition of the temporal, spatial and environmental predictor variables used to model the occurrence of baleen whales in the lower St. Lawrence Estuary portion of the Saguenay–St. Lawrence Marine Park and its surrounding areas (mp = middle point).

Predictor variable	**Definition**
Year	1994 to 2018
Month	Month from June (6) to September (9)
Longitude (Y)	Longitude of the mp - UTM NAD 83 Zone 19 N
Latitude (X)	Latitude of the mp - UTM NAD 83 Zone 19 N
Depth	Average depth (m) around 1km radios of the mp
Depth SD	Standard deviation of depth around 1km radios of the mp
Slope	Average of slope (degrees) around 1km radios of the mp
Slope SD	Standard deviation of slope (degrees) around 1km radios of the mp
Coastline distance	Shortest distance of the mp to the northern shore (m)
Distance upwelling	Distance of the mp to the head of the Laurentian Channel

Table 2. Number of sightings after truncation, surveys with sightings, values of detection probability ($\hat{p}$) with respective standard error (se) and coefficient of variation (cv), effective strip width (esw) and average group size (μ) for each baleen whale species.

Species	**Sightings**	**Surveys**	$\hat{p}$	***se***	***cv***	***esw* (km)**	**μ**
Minke	621	119	0.280	0.024	0.085	1.489	1.02
Fin	269	92	0.411	0.058	0.141	1.780	1.7
Blue	52	34	0.499	0.104	0.209	1.760	1.07
Humpback	41	29	0.518	0.089	0.173	1.760	1.16

*Table 3. Generalized additive model selected for each baleen whale species using systematic distance sampling data. The selected explanatory variables in each model are identified as factors (f) or smooth functions (s) along with their estimated degrees of freedom in parentheses and p-value significance (***: α=0.001; **: α=0.01; *: α=0.05, ‘:α=0.1). Percentage deviance explained is also presented. Dashes corresponds to non-selected variables.*

Predictor variable	**Minke**	**Fin**	**Blue**	**Humpback**
Year	*f*	*f*	*f*	*f*
Latitude (Y)	*s*(1.00)***	-	*s*(1.00)**	-
Longitude (X)	-	*s*(1.56)***	-	*s*(1.24)*
Depth	*s*(4.92)***	*s*(4.58)***	*s*(2.11)	*s*(3.33)
log(Slope)	-	*s*(2.8)*	-	-
log(Slope SD)	*s*(4.17)***	-	*s*(3.34)’	*s*(1.000)’
Coastline distance	-	-	-	-
Distance upwelling	-	-	-	-
% Deviance explained	37.8	44.8	19.4	18.8

Table 4. Density and abundance estimated for each baleen whale species using systematic distance sampling data (CI: confidence interval; se: standard error; cv: coefficient of variation).

Species	**Density**			**Abundance**		**GAM cv**	**Total se**	**Total cv**
	min	**mean**	**max**	**N**	**CI (95%)**			
Minke	0.019	0.468	2.366	58	46 - 75	0.093	7.362	0.126
Fin	0.002	0.110	1.177	14	10 - 20	0.119	2.540	0.184
Blue	0.0001	0.014	0.070	2	1 - 3	0.194	0.514	0.285
Humpback	0.0005	0.008	0.058	1	1 - 2	0.200	0.279	0.264

*Table 5. Generalized additive model selected for each baleen whale species using whale watching activities monitoring data. The selected explanatory variables in each model are identified as factors (f) or smooth functions (s) along with their estimated degrees of freedom in parentheses and p-value significance (***: α=0.001; **: α=0.01; *: α=0.05). Percentage deviance explained is also presented. Dashes corresponds to non-selected variables.*

Predictor variable	**Minke**	**Fin**	**Blue**	**Humpback**
Year	*f*	*f*	*f*	*f*
Month	*f*	*f*	*f*	*f*
Depth	-	-	-	s(3.345)**
Slope	-	-	S(2.503)*	s(2.846)*
log(Slope), June	s(2.642)	s(2.456)***	-	-
log(Slope), July	s(4.041)**	s(3.405)***	-	-
log(Slope), August	s(3.735)***	s(2.948)***	-	-
log(Slope), September	s(2.655)*	s(2.938)***	-	-
Depth, June	s(1.003)***	s(4.607)***	s(2.290)	-
Depth, July	s(4.574)***	s(5.865)***	s(1.459)***	-
Depth, August	s(4.105)***	s(4.361)***	s(2.108)***	-
Depth, September	s(4.221)***	s(3.789)***	s(1.525)***	-
% Deviance explained	33.9	42.5	36.7	27.5

FIGURES

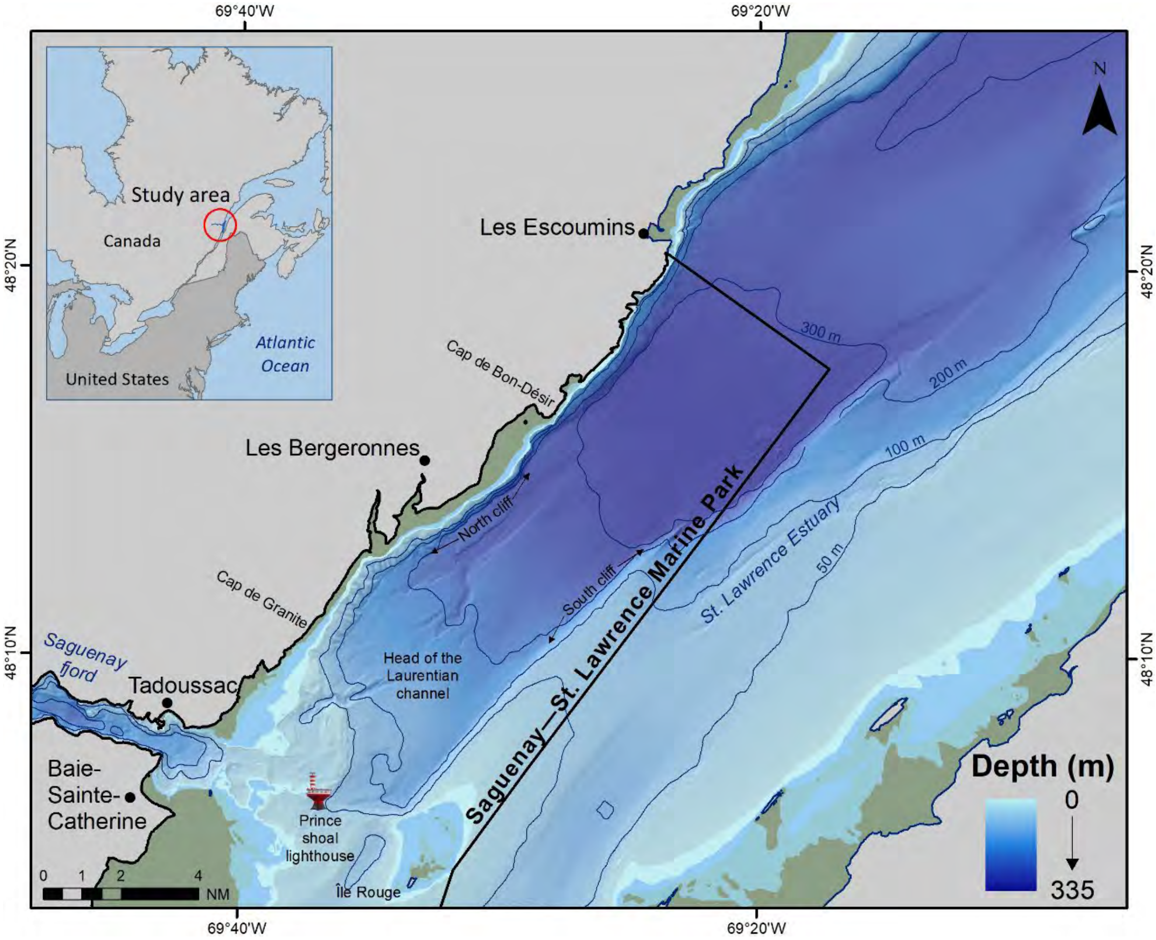


Figure 1. The study area in the lower St. Lawrence Estuary portion of the Saguenay–St. Lawrence Marine Park and its surrounding areas. The bathymetric contours and the head of the Laurentian Channel as used in this document are indicated.

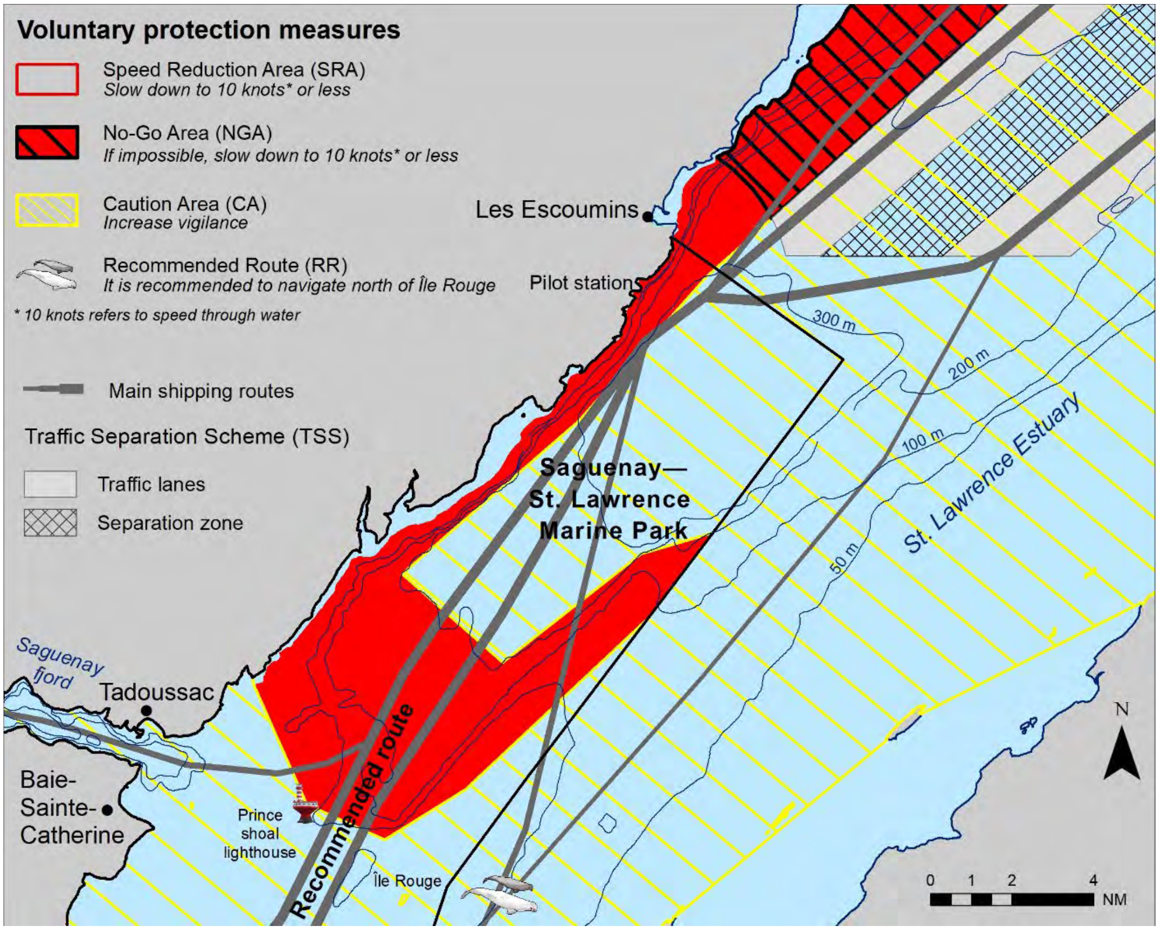


Figure 2. Voluntary protection measures to reduce the risk of ship strikes to whales and to minimize the impact of noise on beluga whales, main shipping routes, and traffic separation scheme.

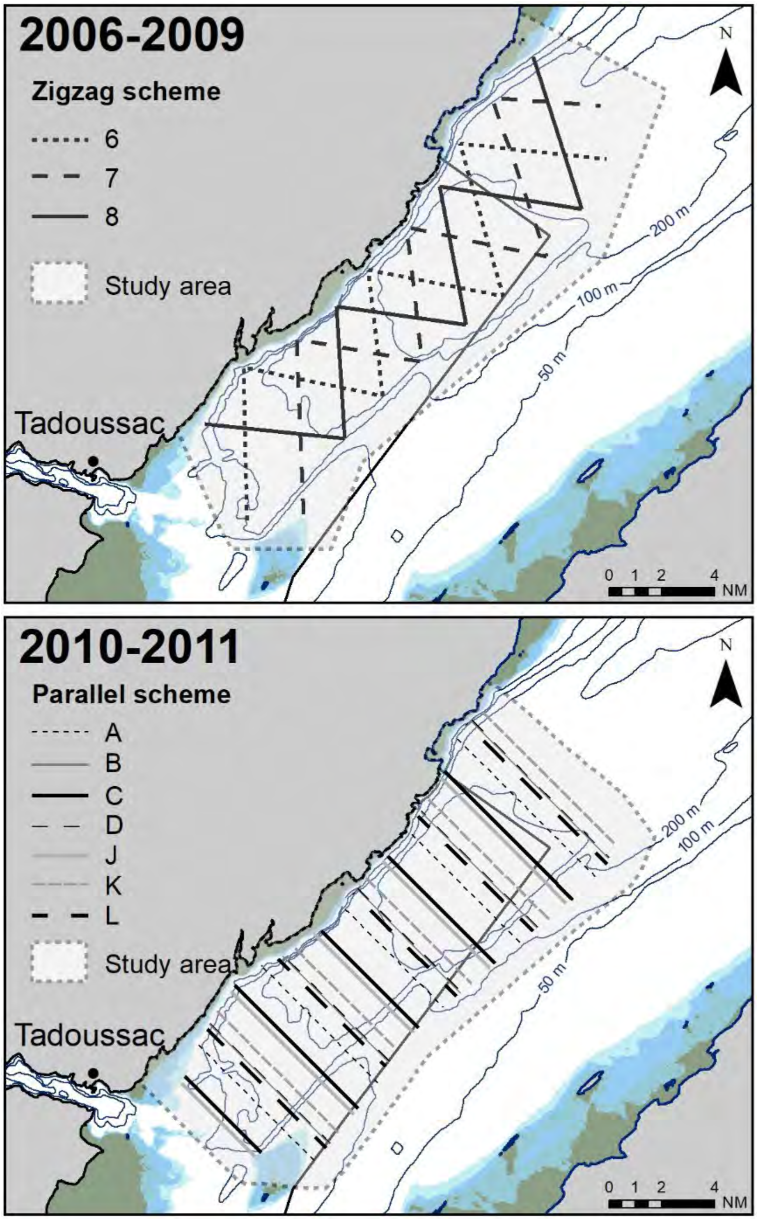


Figure 3. Study area and survey design for the systematic surveys conducted in the lower St. Lawrence Estuary portion of the Saguenay–St. Lawrence Marine Park and its surrounding areas.

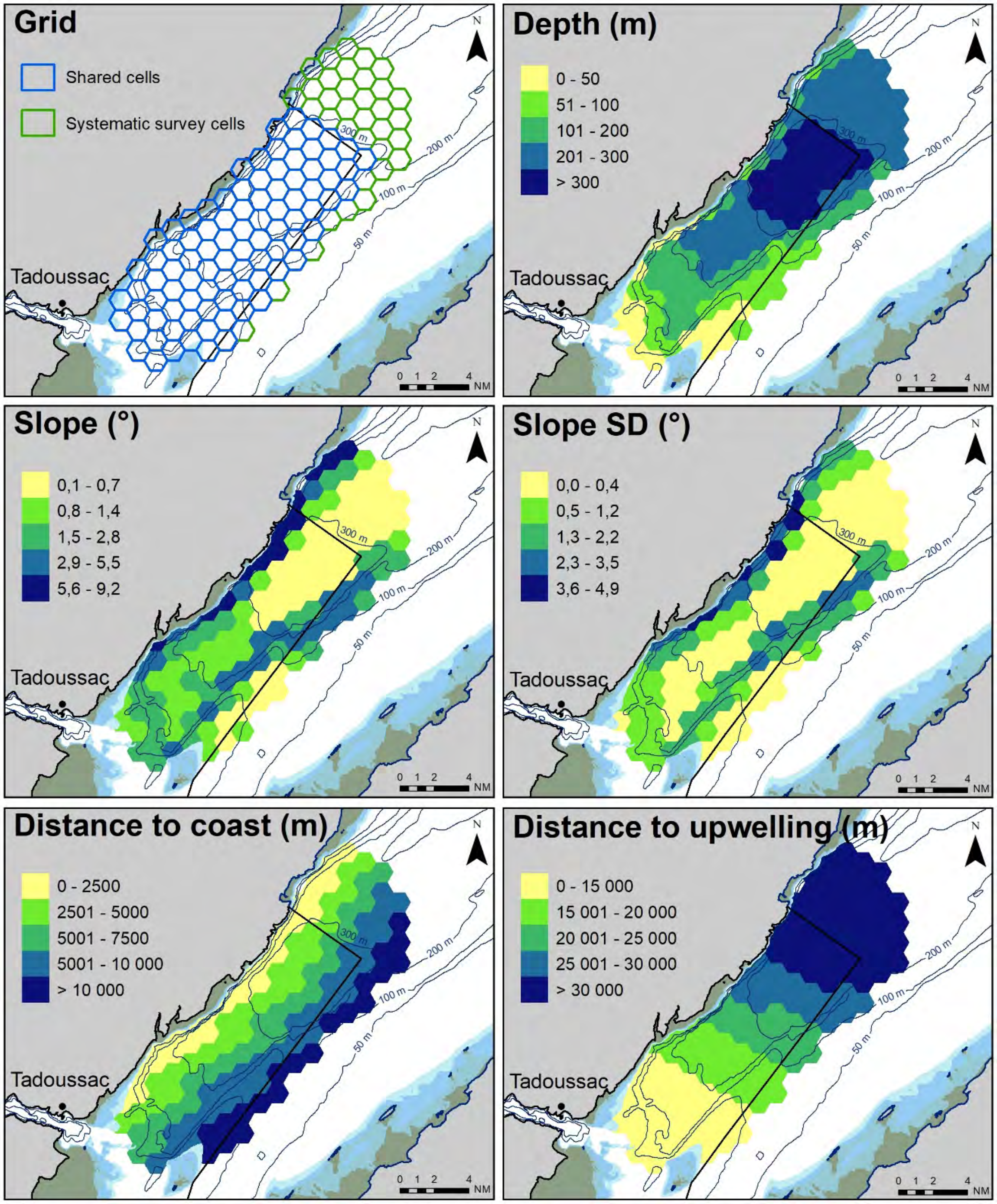


Figure 4. Adopted grid and predictor environmental variables used in baleen whale species distribution models.

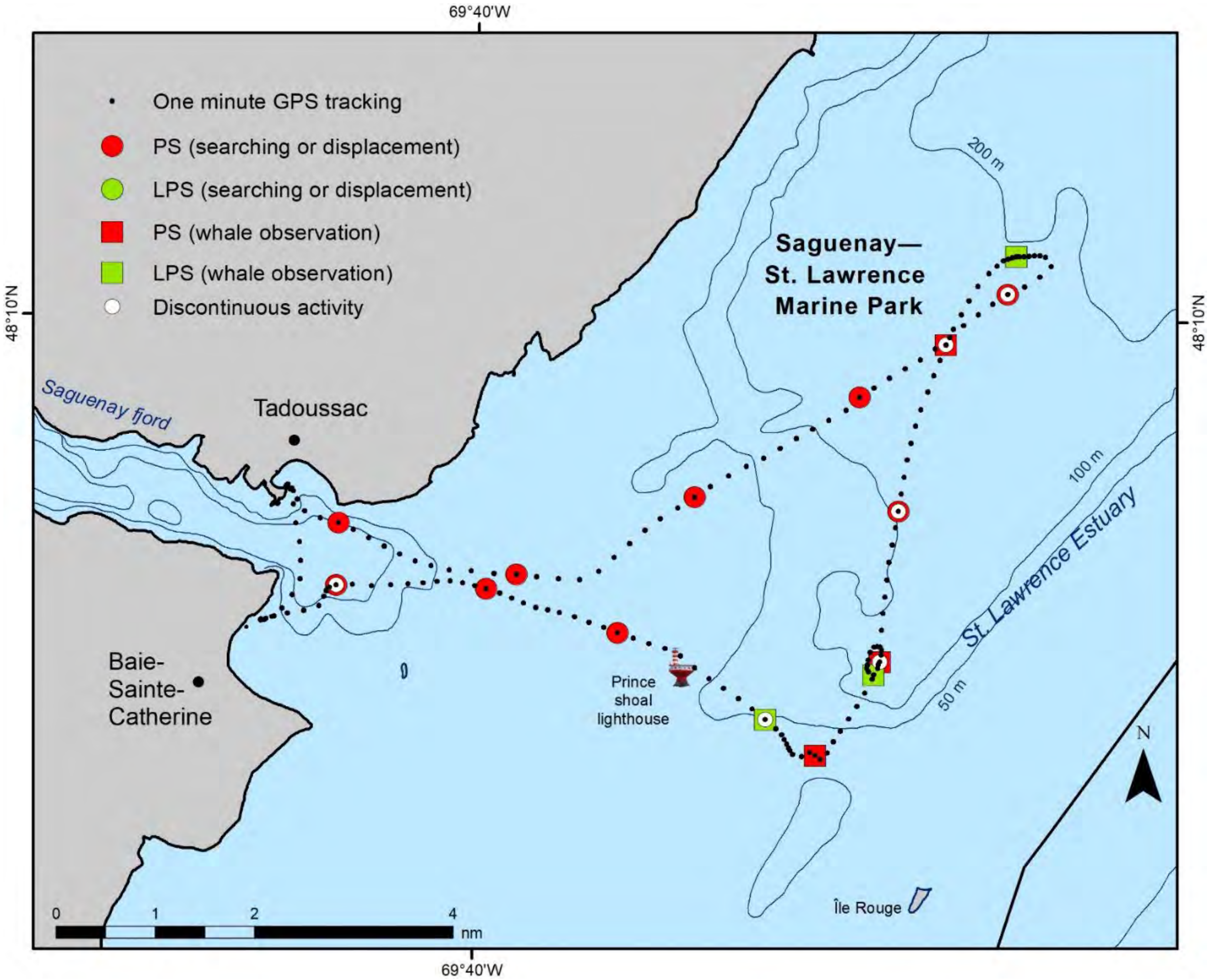


Figure 5. Illustration of the sampling method and data selection. The black dots represent the GPS track of the excursion and the colored icons are the sampling points. The circles are sampling points where the activity was searching or displacement and the squares are sampling points where the activity was whale observation. The red color represents the regular sampling points (3 minutes) and the green color the long sampling points (5 minutes). The white circles correspond to discontinuous activity. From 1994 to 2008, only the green squares were kept for analyzes. From 2009 to 2018, the green or red squares with a white circle inside were kept for analysis.

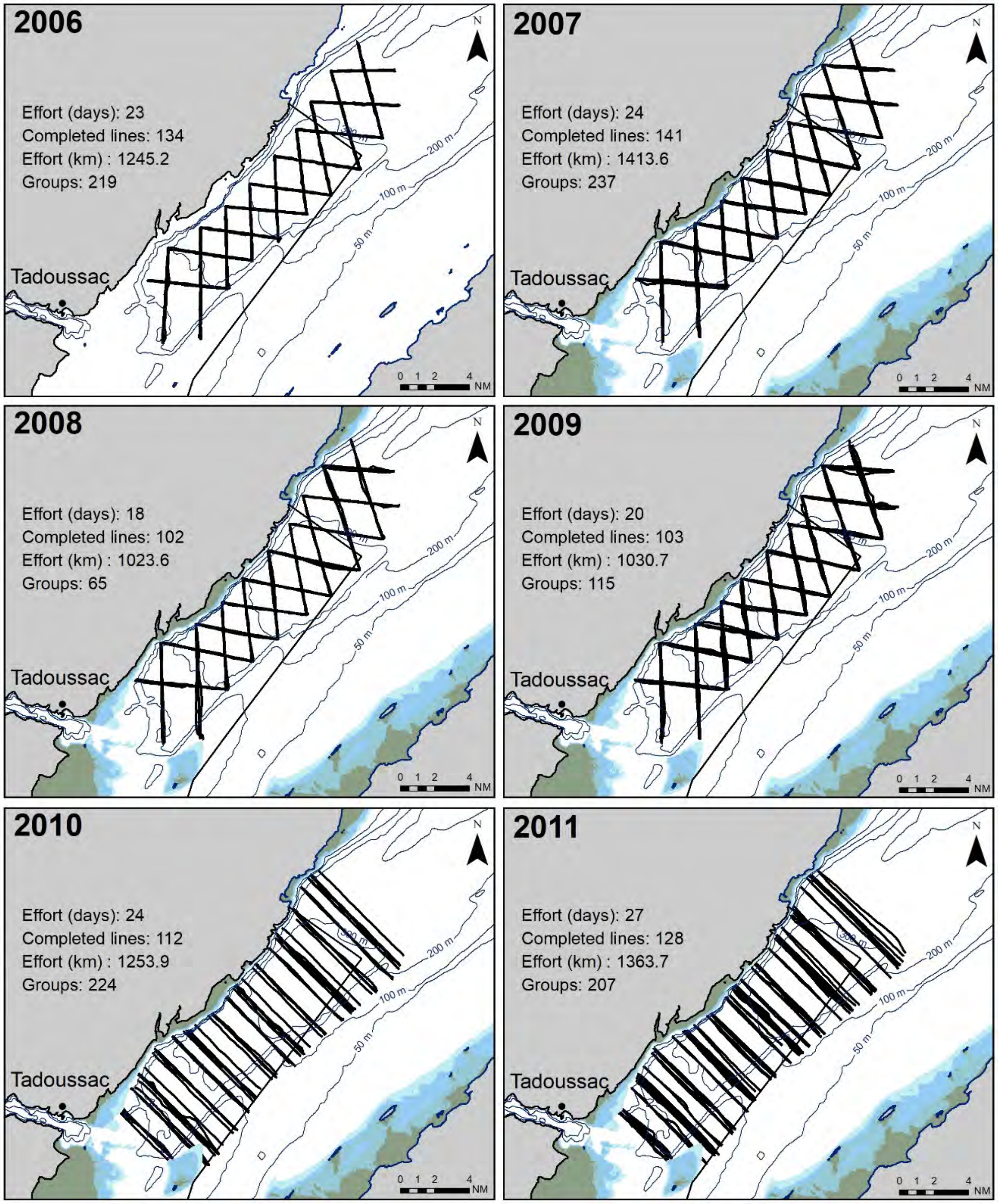


Figure 6. Transect lines completed during six years of systematic surveys conducted in the lower St. Lawrence Estuary portion of the Saguenay–St. Lawrence Marine Park and its surrounding areas.

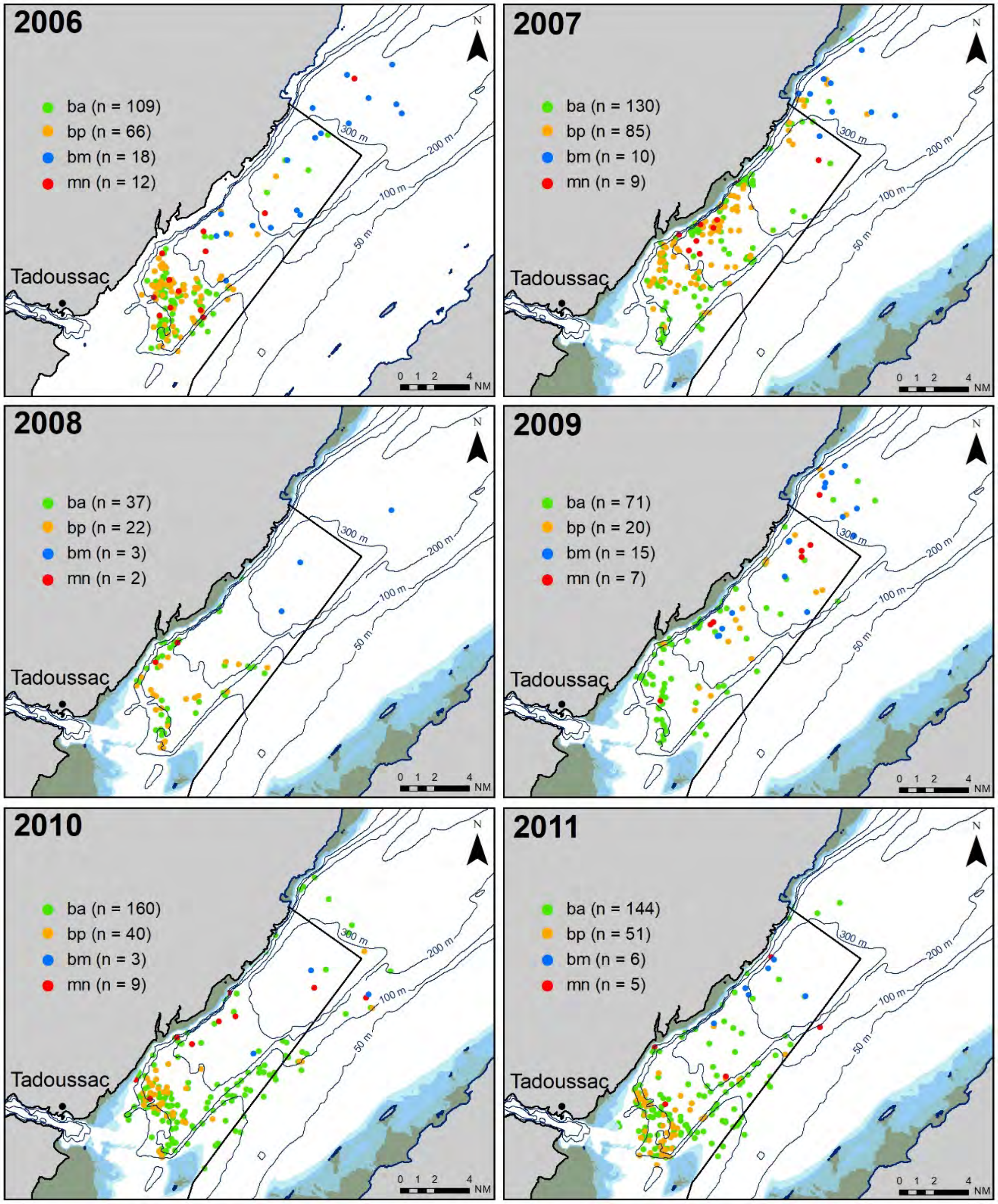


Figure 7. Baleen whales observed during the systematic surveys conducted in the lower St. Lawrence Estuary portion of the Saguenay–St. Lawrence Marine Park and its surrounding areas. ba = minke whale, bp = fin whale, bm = blue whale and mn = humpack whale.

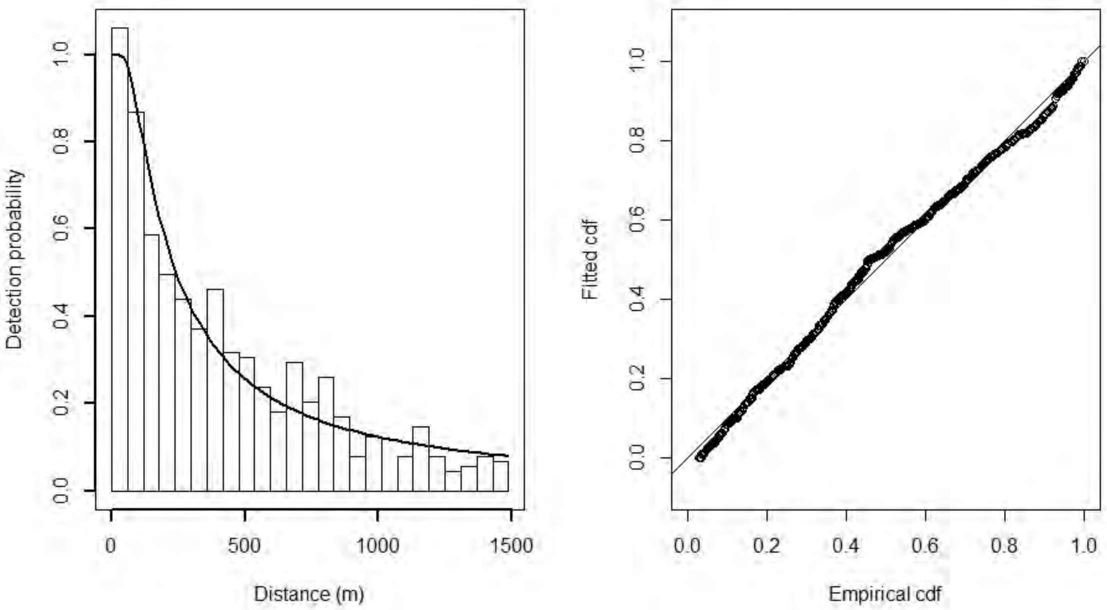


Figure 8. Estimated detection function for minke whale groups overlaid onto the scaled histogram of observed distances and goodness of fit test.

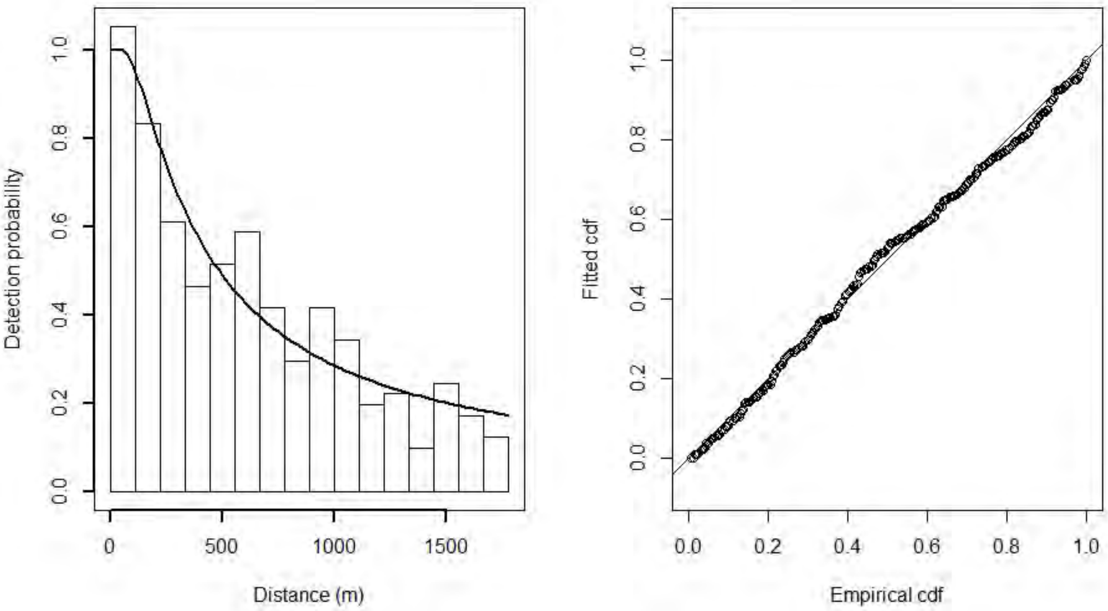


Figure 9. Estimated detection function for fin whale groups overlaid onto the scaled histogram of observed distances and goodness of fit test.

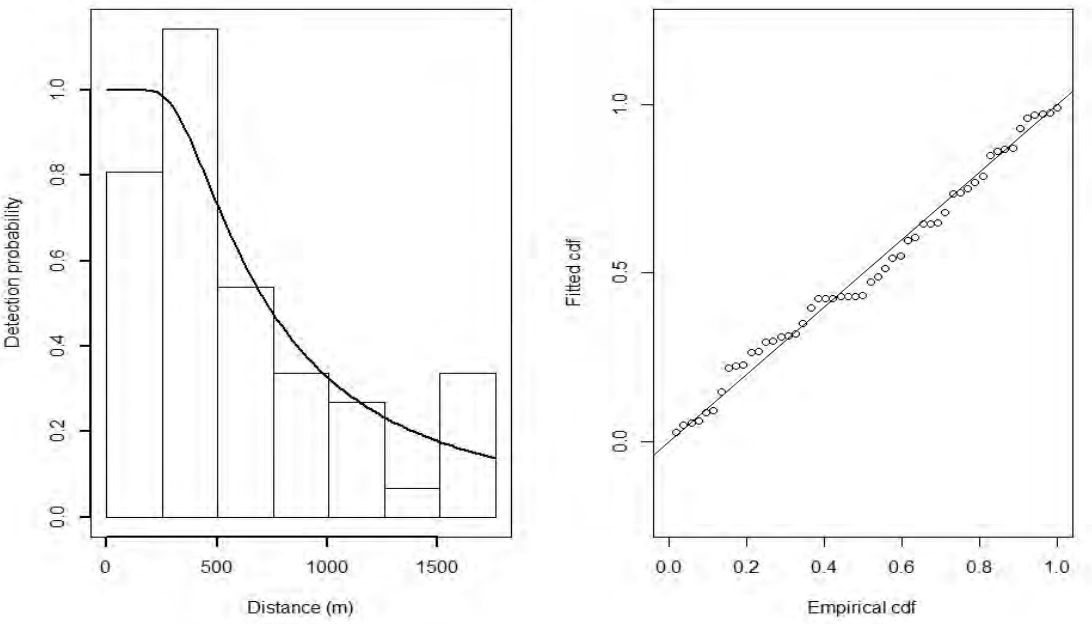


Figure 10. Estimated detection function for blue whale groups overlaid onto the scaled histogram of observed distances and goodness of fit test.

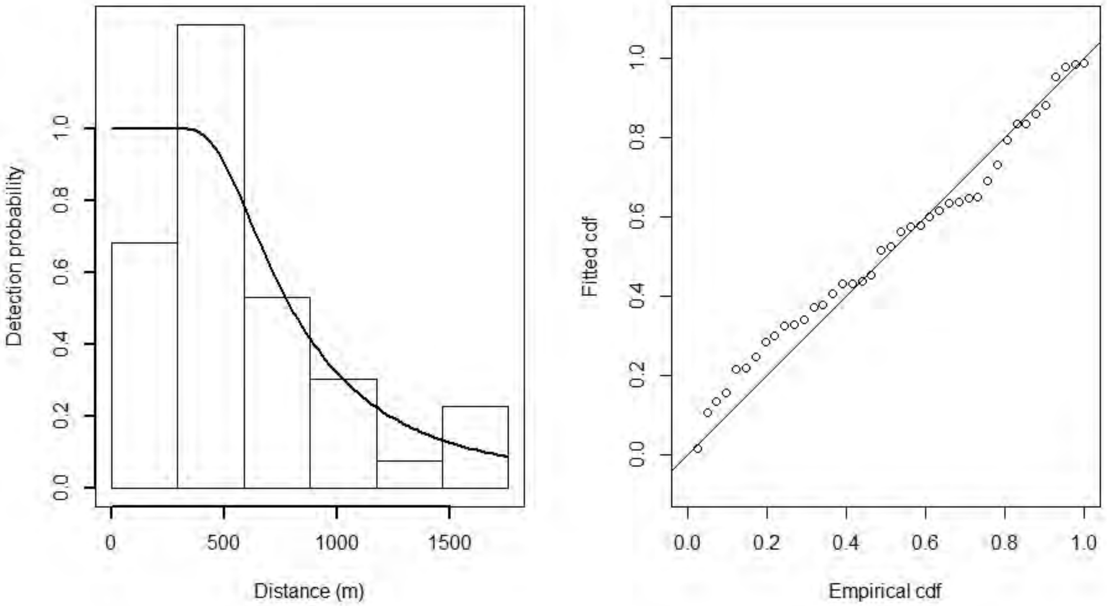


Figure 11. Estimated detection function for humpback whale groups overlaid onto the scaled histogram of observed distances and goodness of fit test.

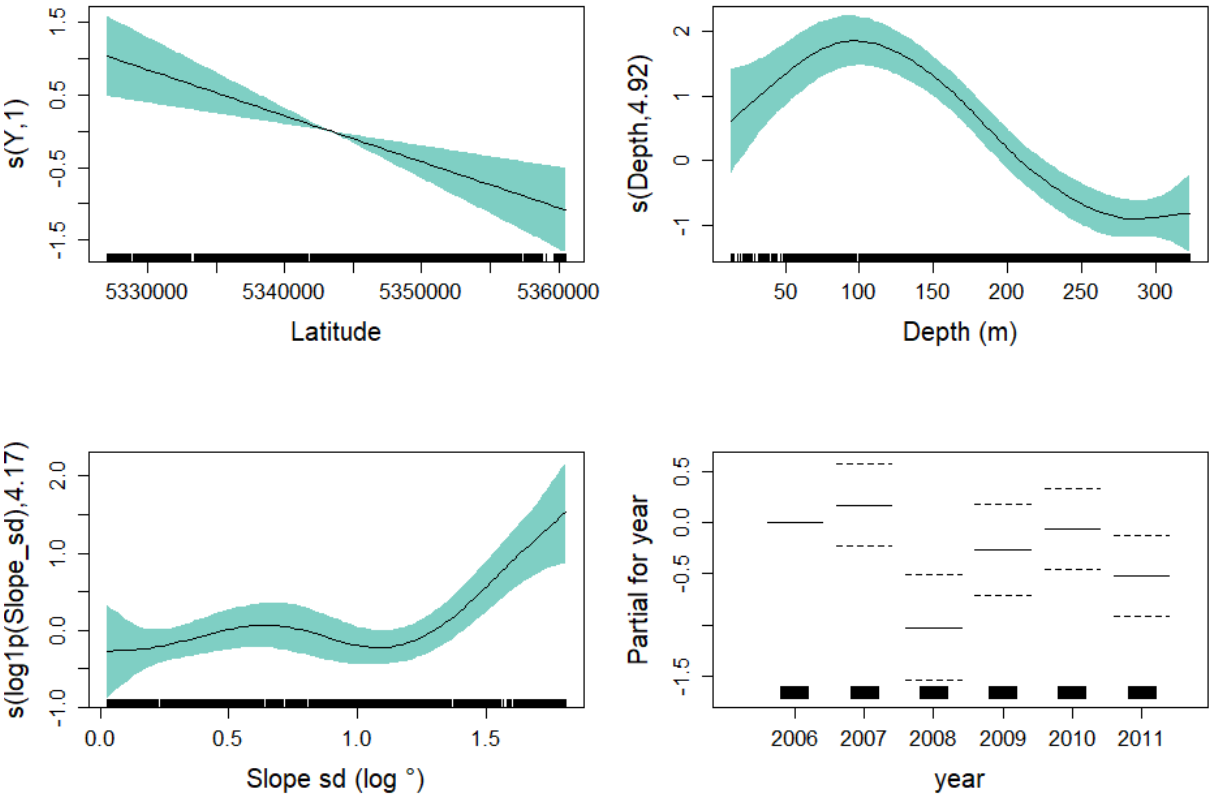


Figure 12. Smooth functions of the environmental covariates selected to model minke whale distribution using systematic distance sampling data. Positive values of the smoothed function indicate a positive effect on the response variable and the green area around the smooth represents the confidence limits. Vertical lines on the x-axis are the observed data values.

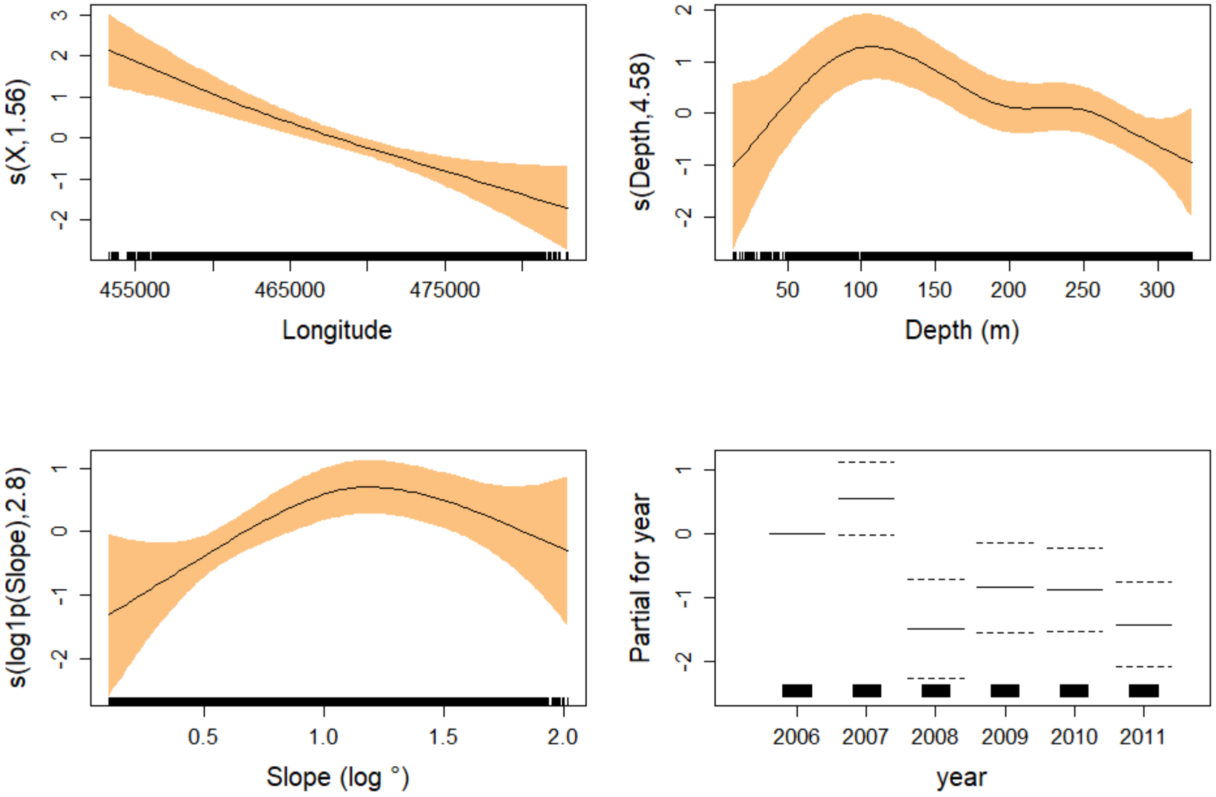


Figure 13. Smooth functions of the environmental covariates selected to model fin whale distribution using systematic distance sampling data. Positive values of the smoothed function indicate a positive effect on the response variable and the orange area around the smooth represents the confidence limits. Vertical lines on the x-axis are the observed data values.

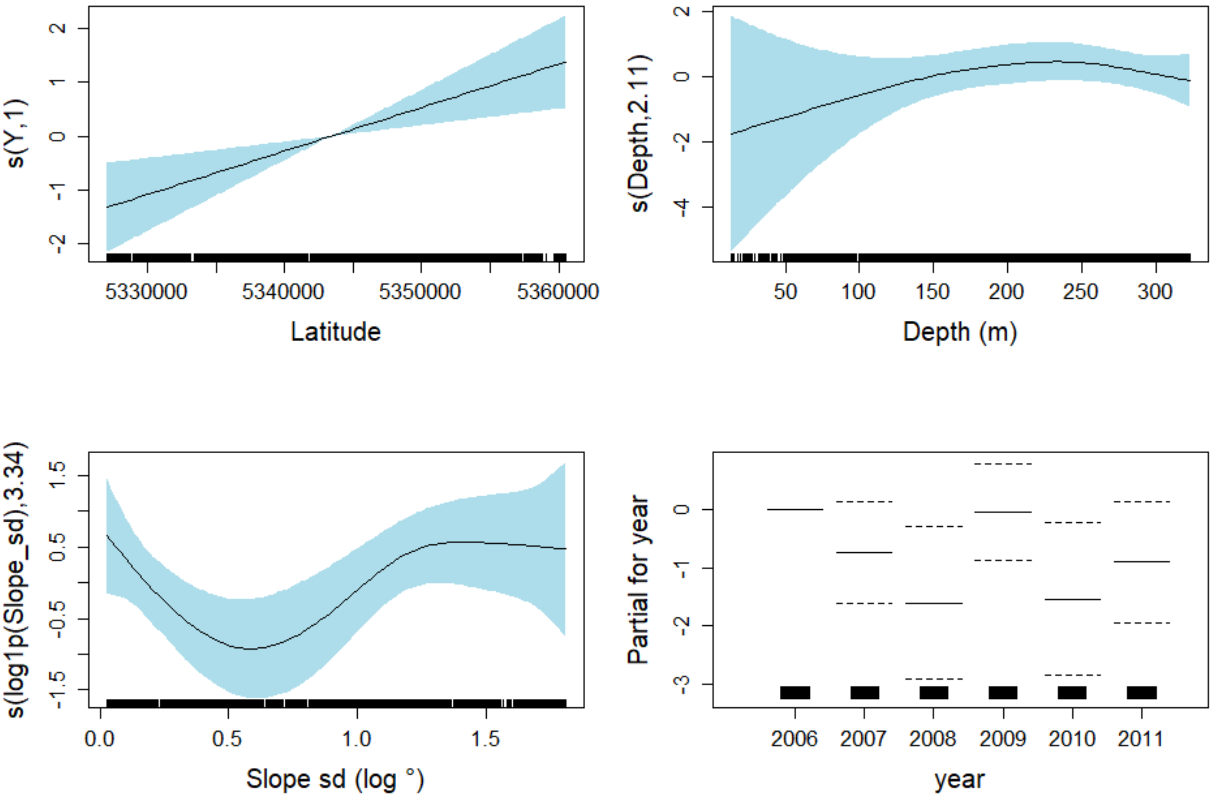


Figure 14. Smooth functions of the environmental covariates selected to model blue whale distribution using systematic distance sampling data. Positive values of the smoothed function indicate a positive effect on the response variable and the green area around the smooth represents the confidence limits. Vertical lines on the x-axis are the observed data values.

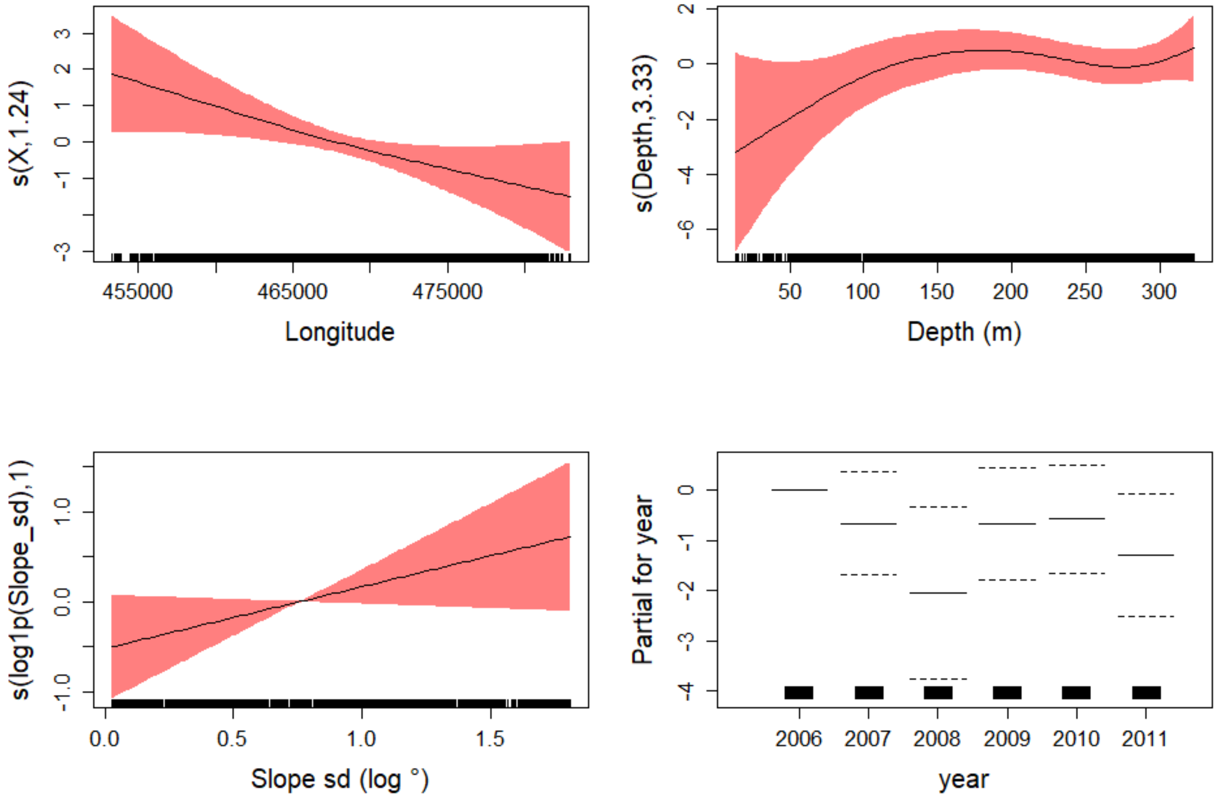


Figure 15. Smooth functions of the environmental covariates selected to model humpback whale distribution using systematic distance sampling data. Positive values of the smoothed function indicate a positive effect on the response variable and the green area around the smooth represents the confidence limits. Vertical lines on the x-axis are the observed data values.

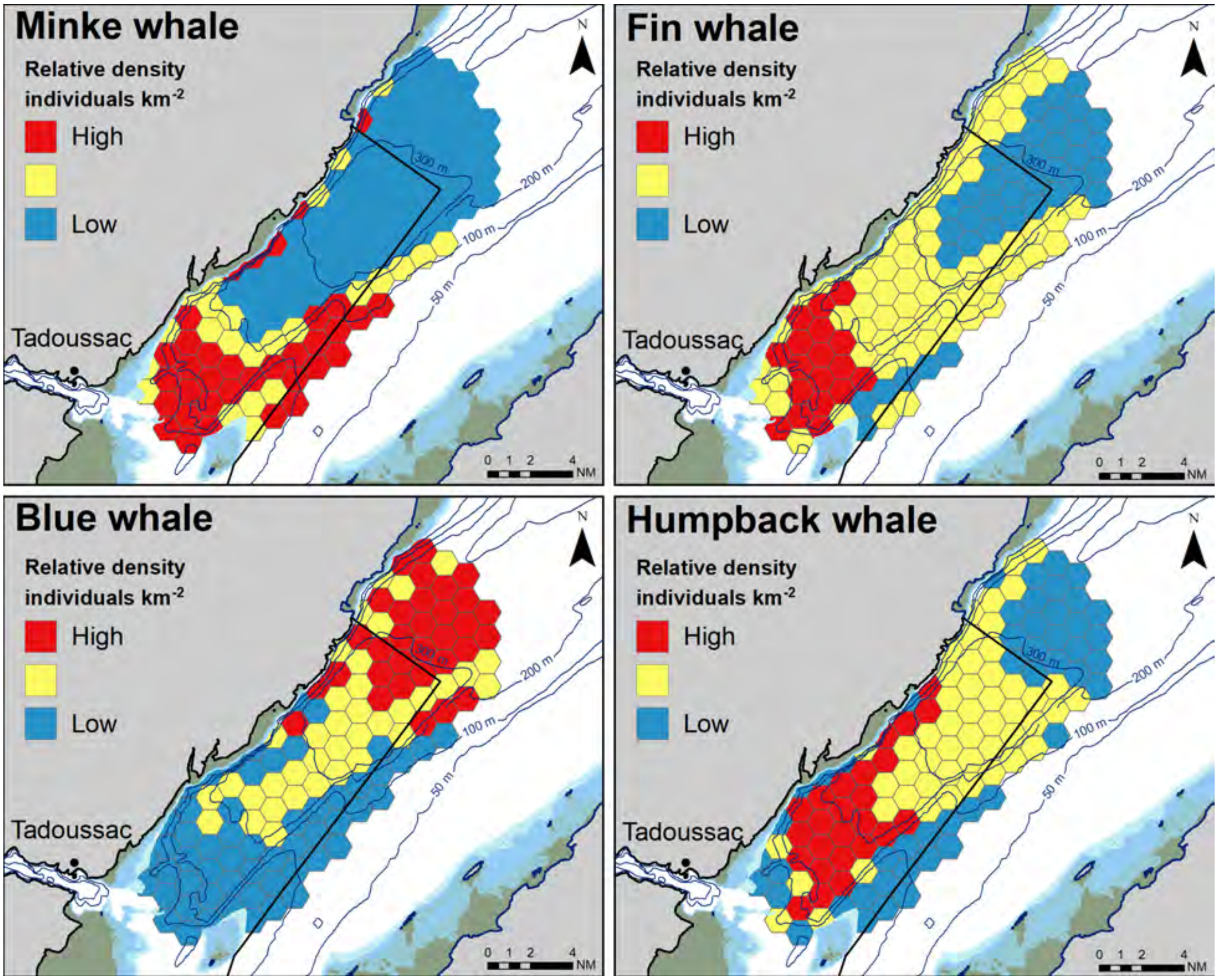


Figure 16. Spatial distribution of the predicted densities of each baleen whale species derived from the models using systematic distance sampling data.

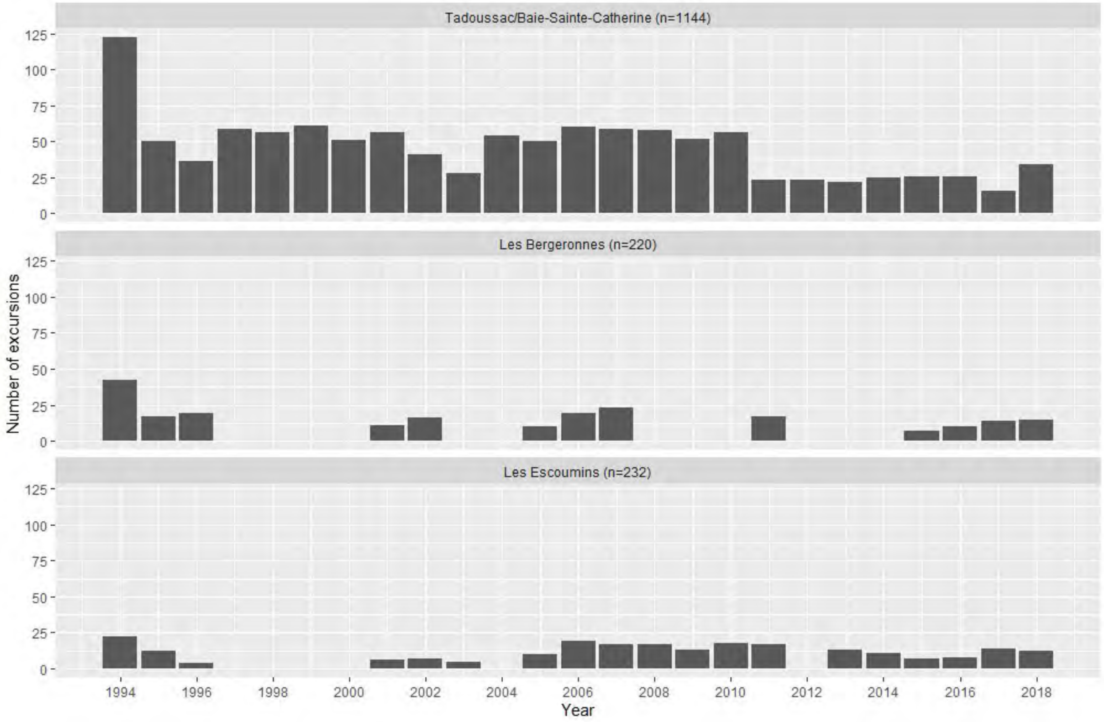


Figure 17. Number of excursions sampled annually per homeport.

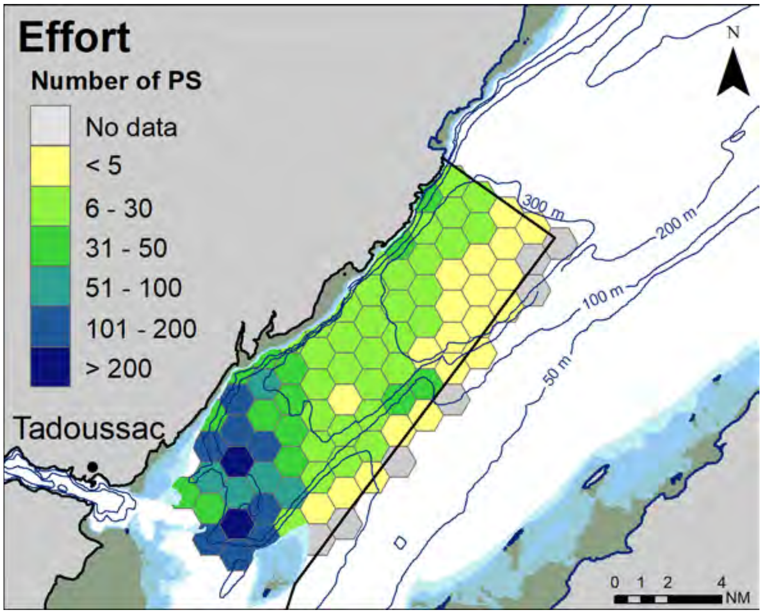


Figure 18. Survey area and global effort (point sampling – PS) considered in the analysis of the data of the whale watching activities monitoring.

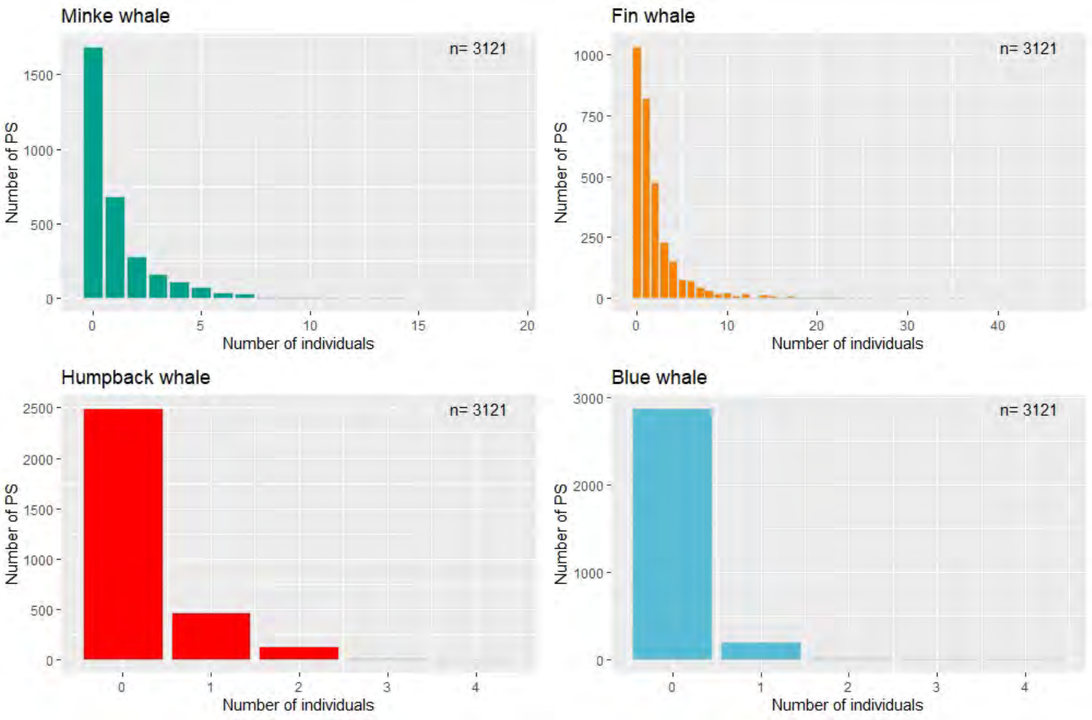


Figure 19. The number of individuals observed per point sampling (PS) for each species from the whale-watching activities monitoring dataset.

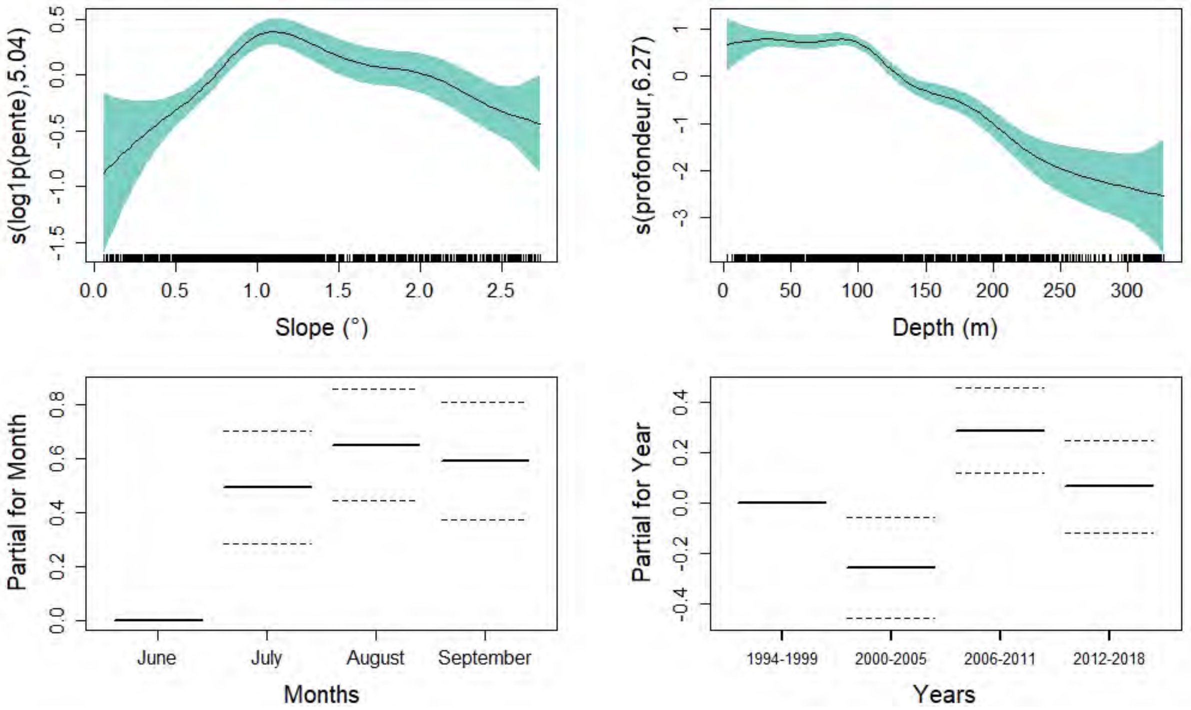


Figure 20. Smooth functions of the environmental covariates selected to model minke whale distribution using whale watching activities monitoring data. To lighten the graphical presentation, no interaction between the months and the slope and depth variables is included in these graphs. Positive values of the smoothed function indicate a positive effect on the response variable. The green area represents the confidence limits. Vertical lines on the x-axis are the observed data values.

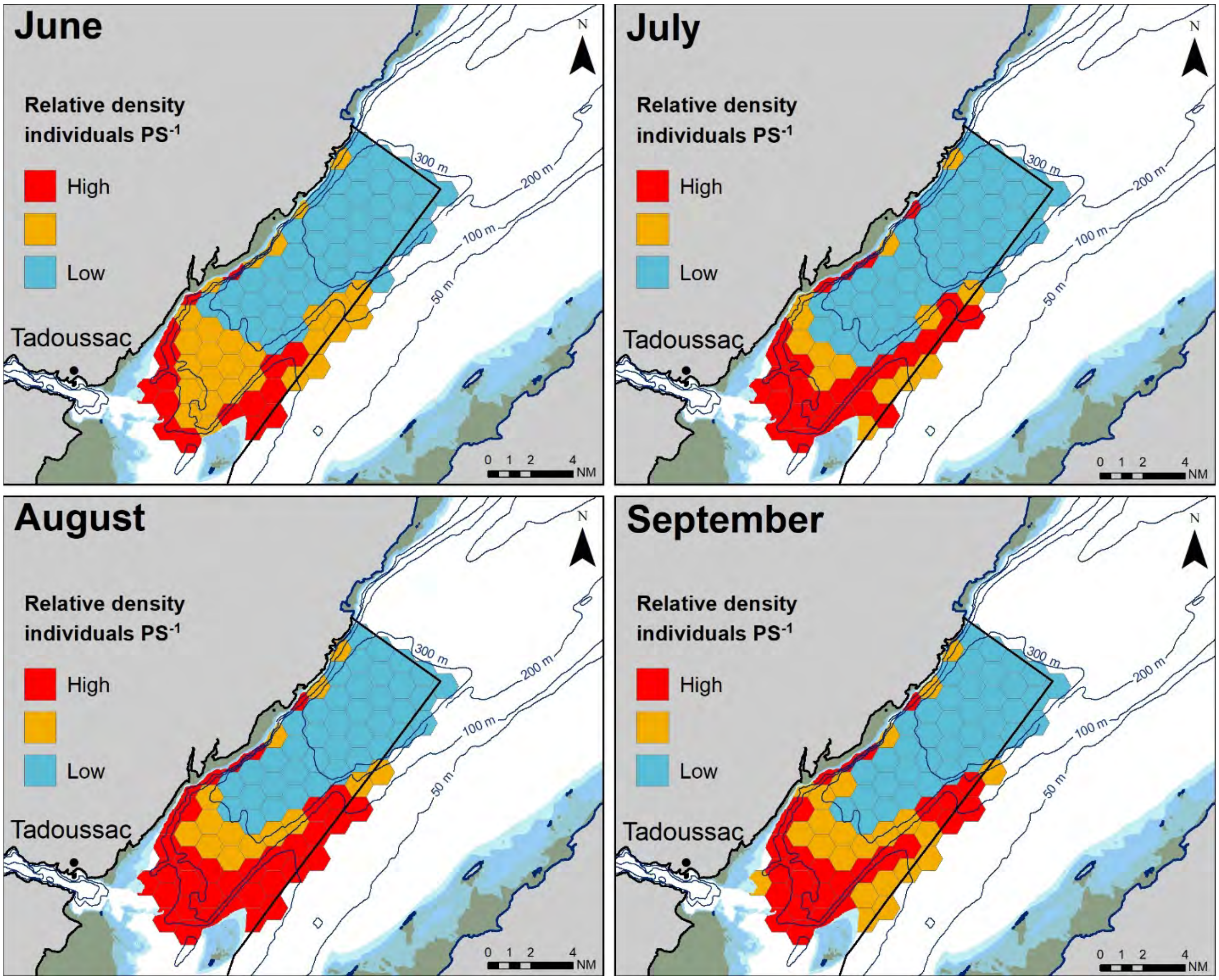


Figure 21 Spatial distribution of the predicted relative density for the minke whale by month using whale watching activities monitoring data. The scales are the same on each map to facilitate comparisons.

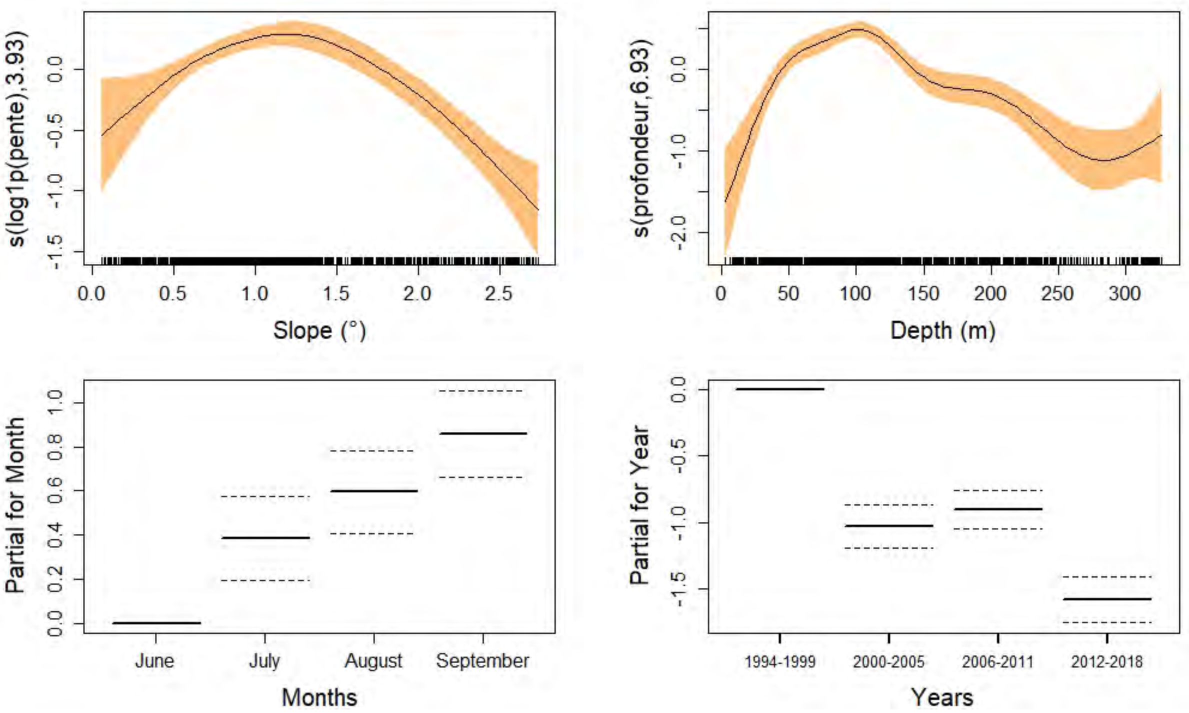


Figure 22. Smooth functions of the environmental covariates selected to model fin whale distribution using whale watching activities monitoring data. To lighten the graphical presentation, no interaction between the months and the slope and depth variables is included in these graphs. Positive values of the smoothed function indicate a positive effect on the response variable. The orange area represents the confidence limits. Vertical lines on the x-axis are the observed data values.

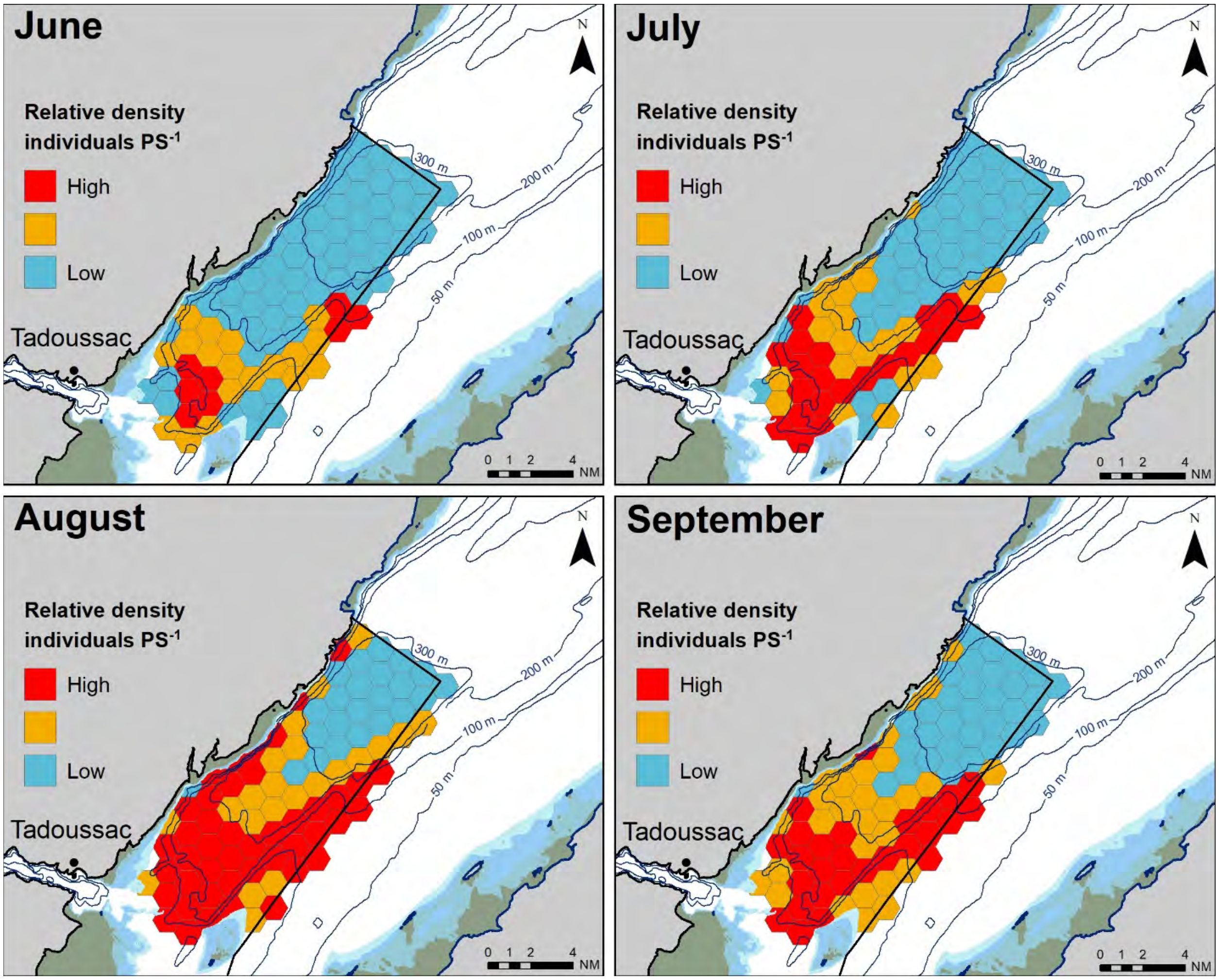


Figure 23. Spatial distribution of the predicted relative density for the fin whale by month using whale watching activities monitoring data. The scales are the same on each map to facilitate comparisons.

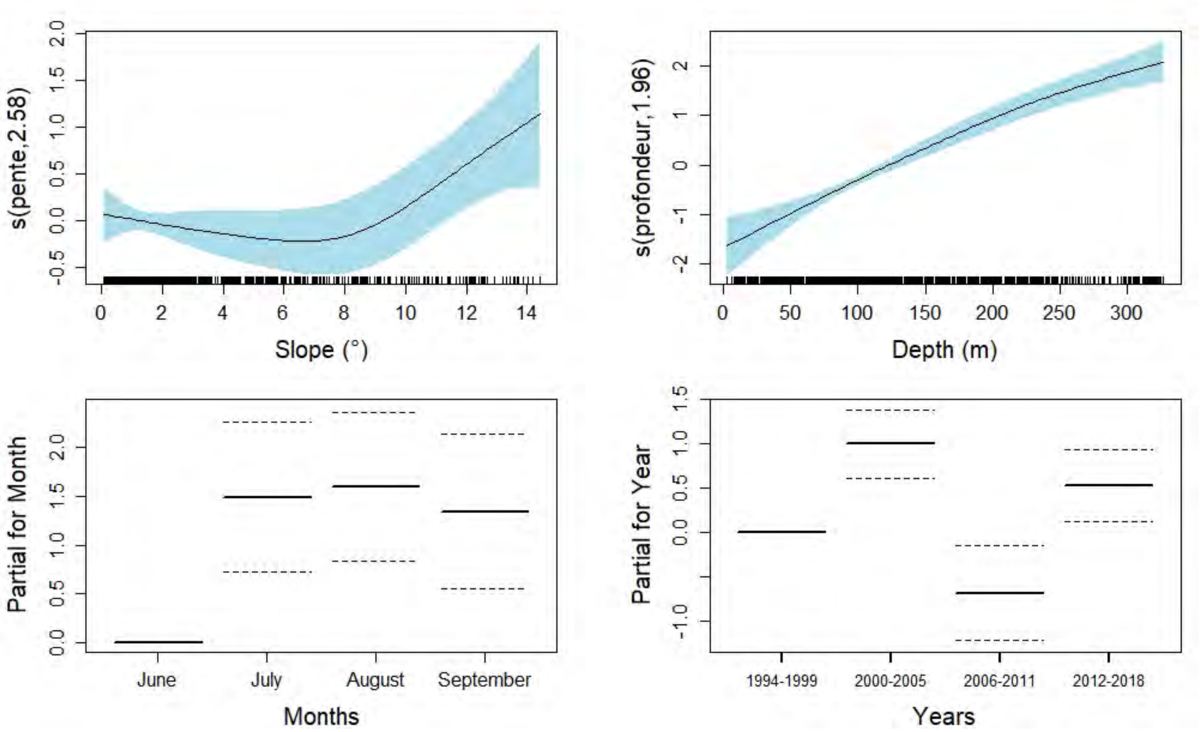


Figure 24. Smooth functions of the environmental covariates selected to model blue whale distribution using whale watching activities monitoring data. To lighten the graphical presentation, no interaction between the months and the depth variable is included in these graphs. Positive values of the smoothed function indicate a positive effect on the response variable. The blue area represents the confidence limits. Vertical lines on the x-axis are the observed data values.

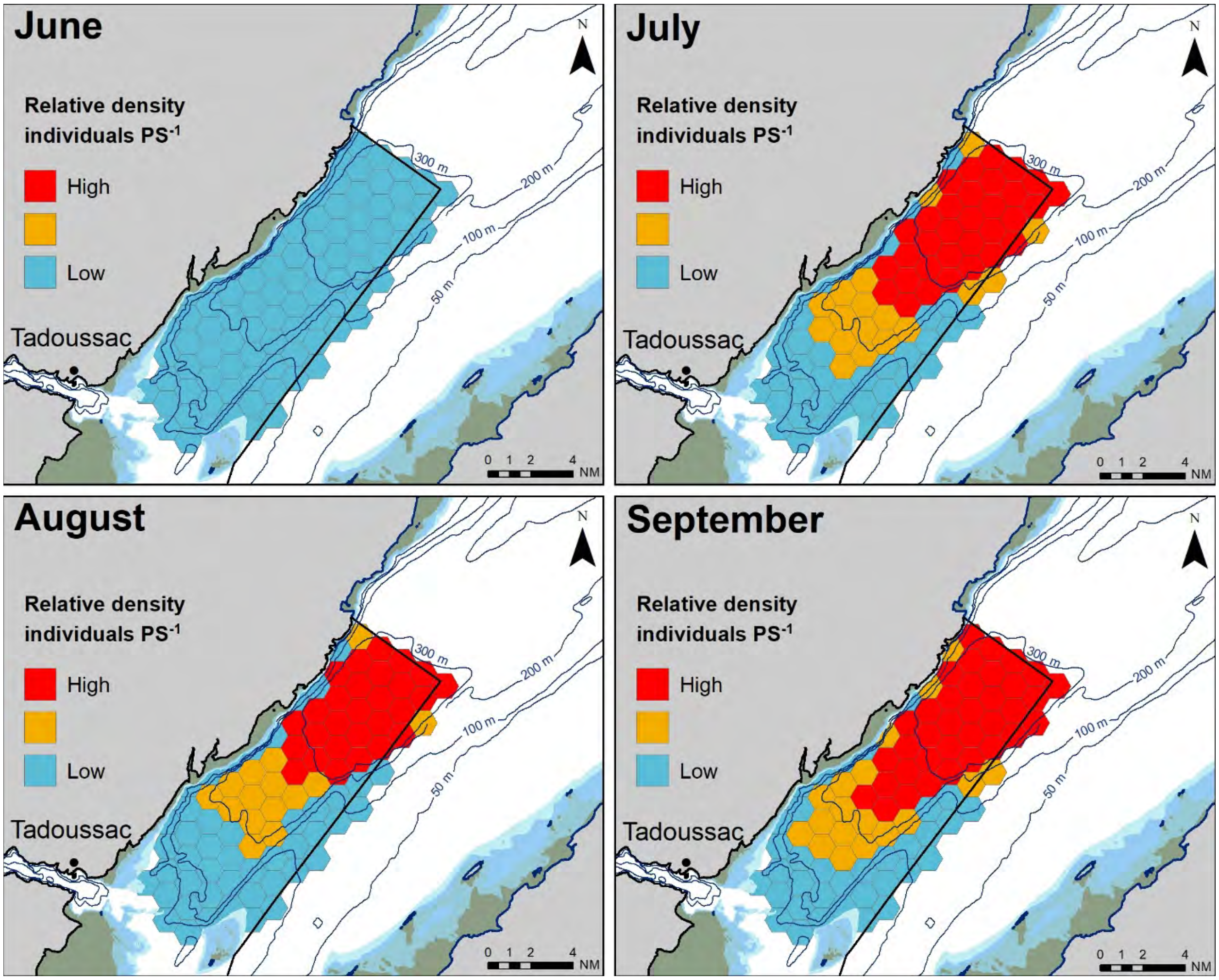


Figure 25. Spatial distribution of the predicted relative density for the blue whale by month using whale watching activities monitoring data. The scales are the same on each map to facilitate comparisons.

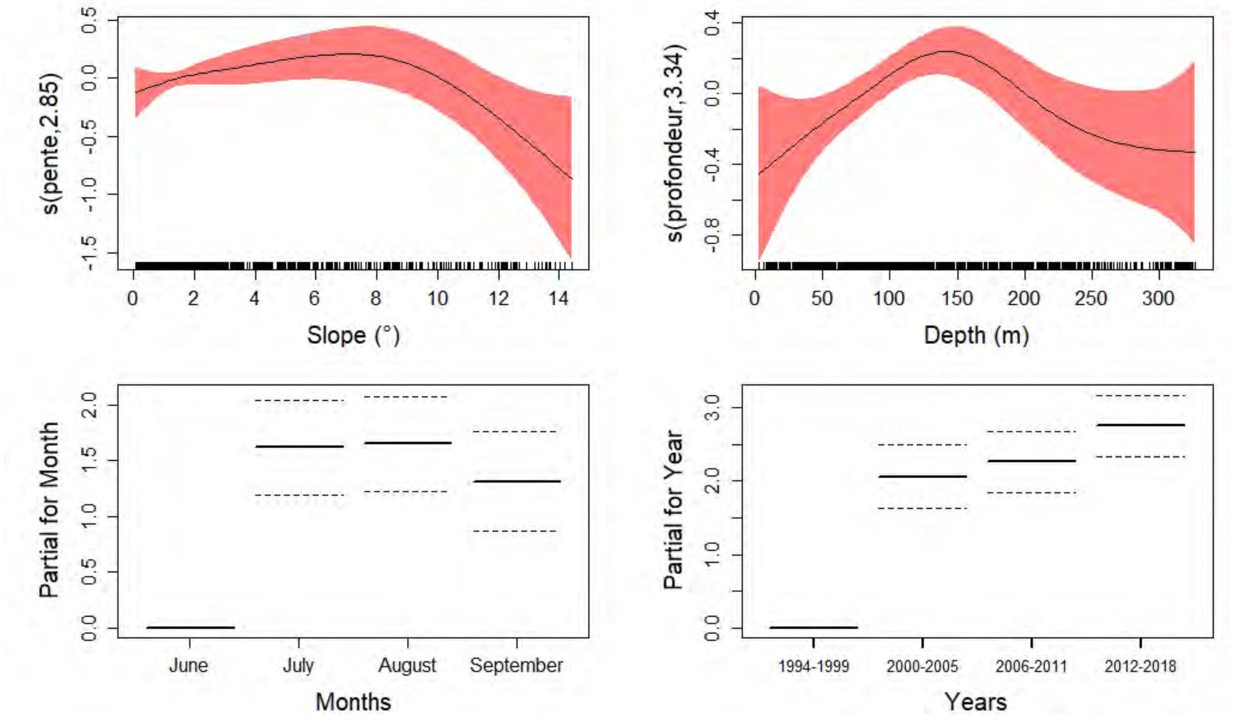


Figure 26. Smooth functions of the environmental covariates selected to model humpback whale distribution using whale watching activities monitoring data. Positive values of the smoothed function indicate a positive effect on the response variable. The green area represents the confidence limits. Vertical lines on the x-axis are the observed data values.

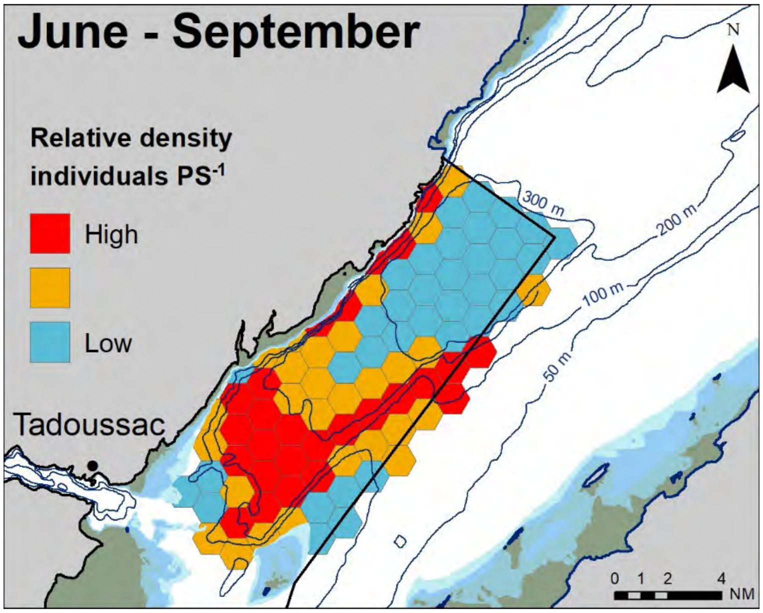


Figure 27. Spatial distribution of the predicted overall relative densities for the humpback whale using whale watching activities monitoring data (PS: point sampling).

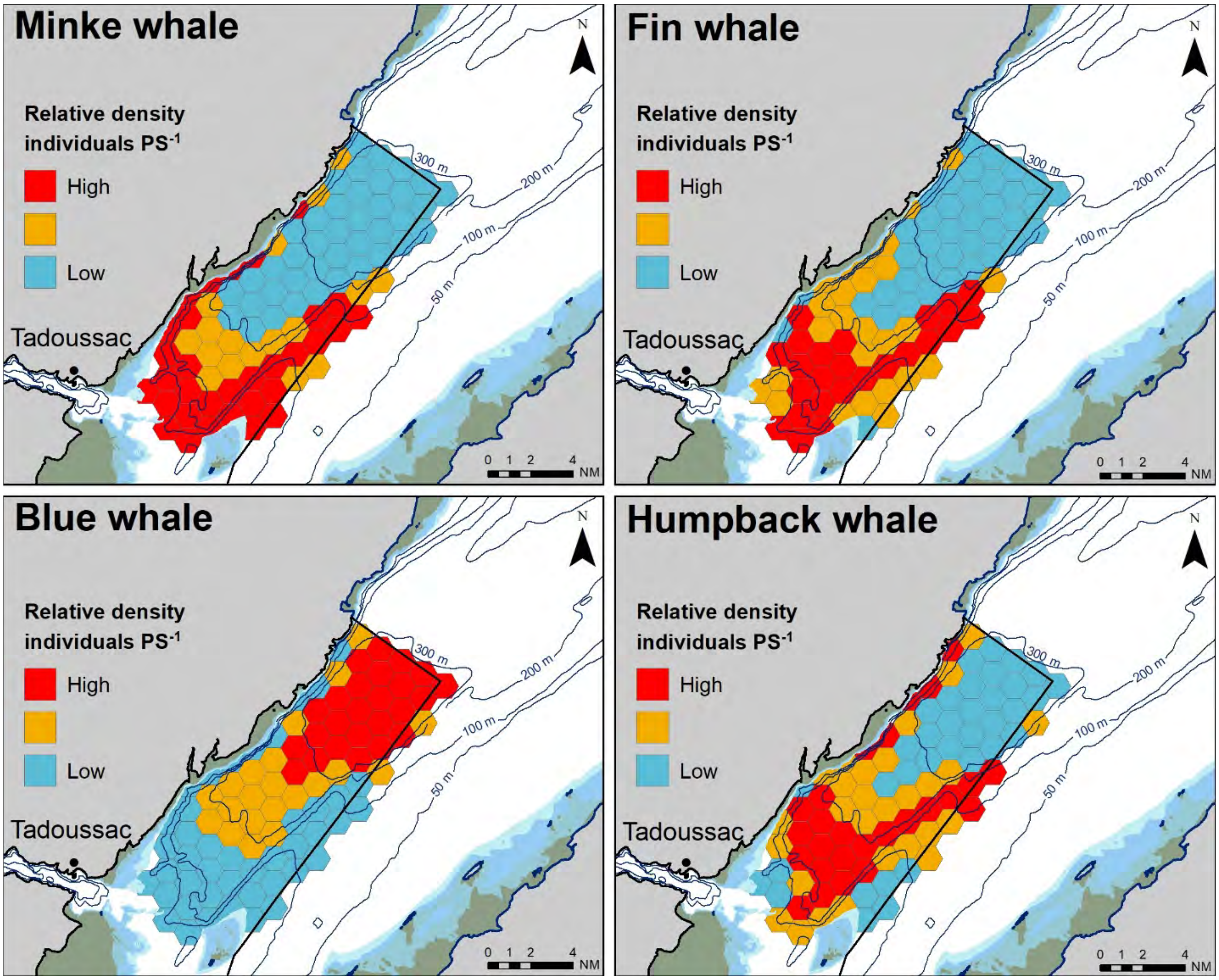


Figure 28. Predicted overall relative density of each baleen whale species derived from the models using whale watching activities monitoring data (PS: point sampling).

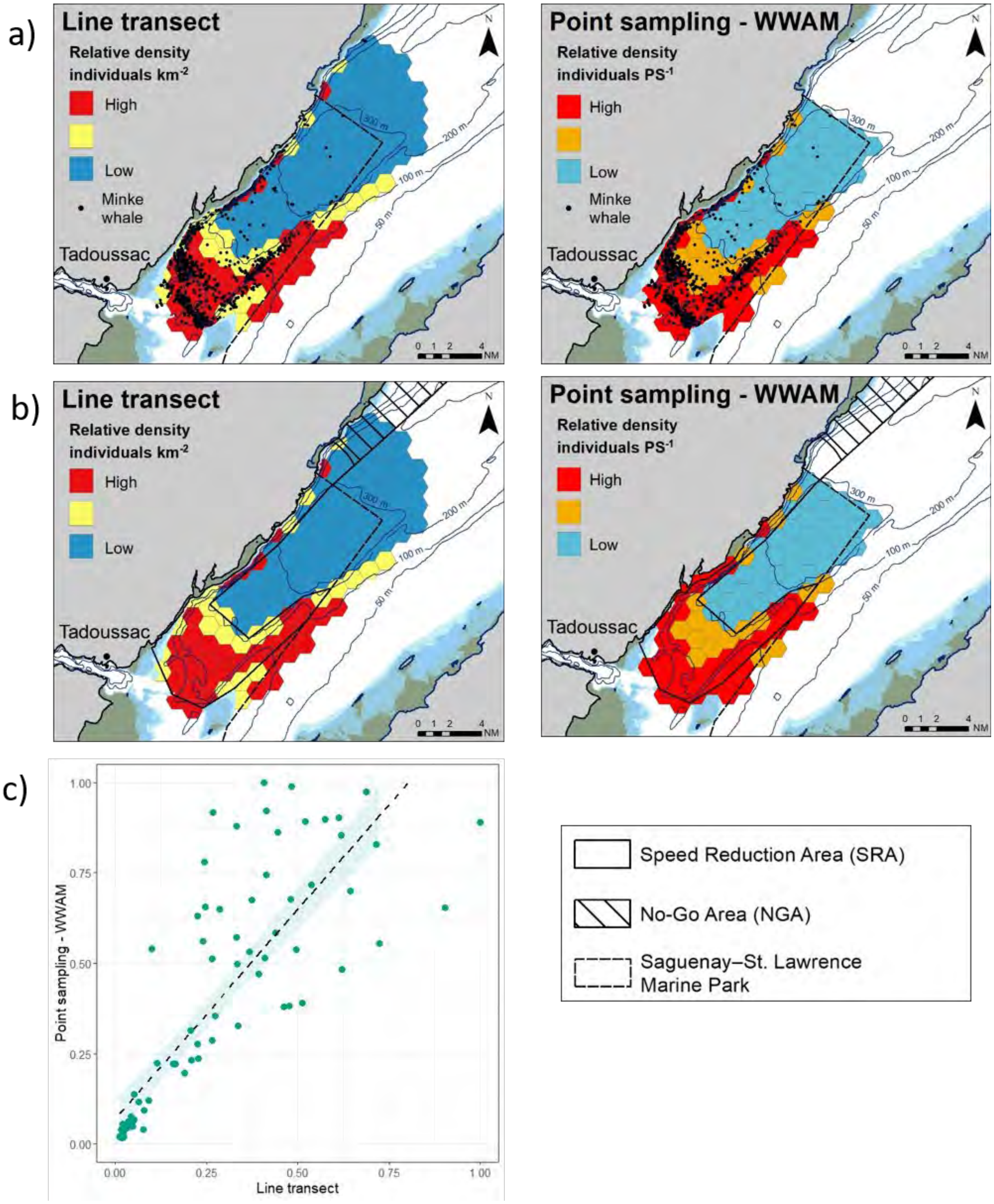


Figure 29. Predicted densities of minke whales derived from line transect (left) and point sampling (PS) (right)models overlaid by a) the validation dataset, b) the current management measures and c) a comparison of the normalised results for the shared grid cells.

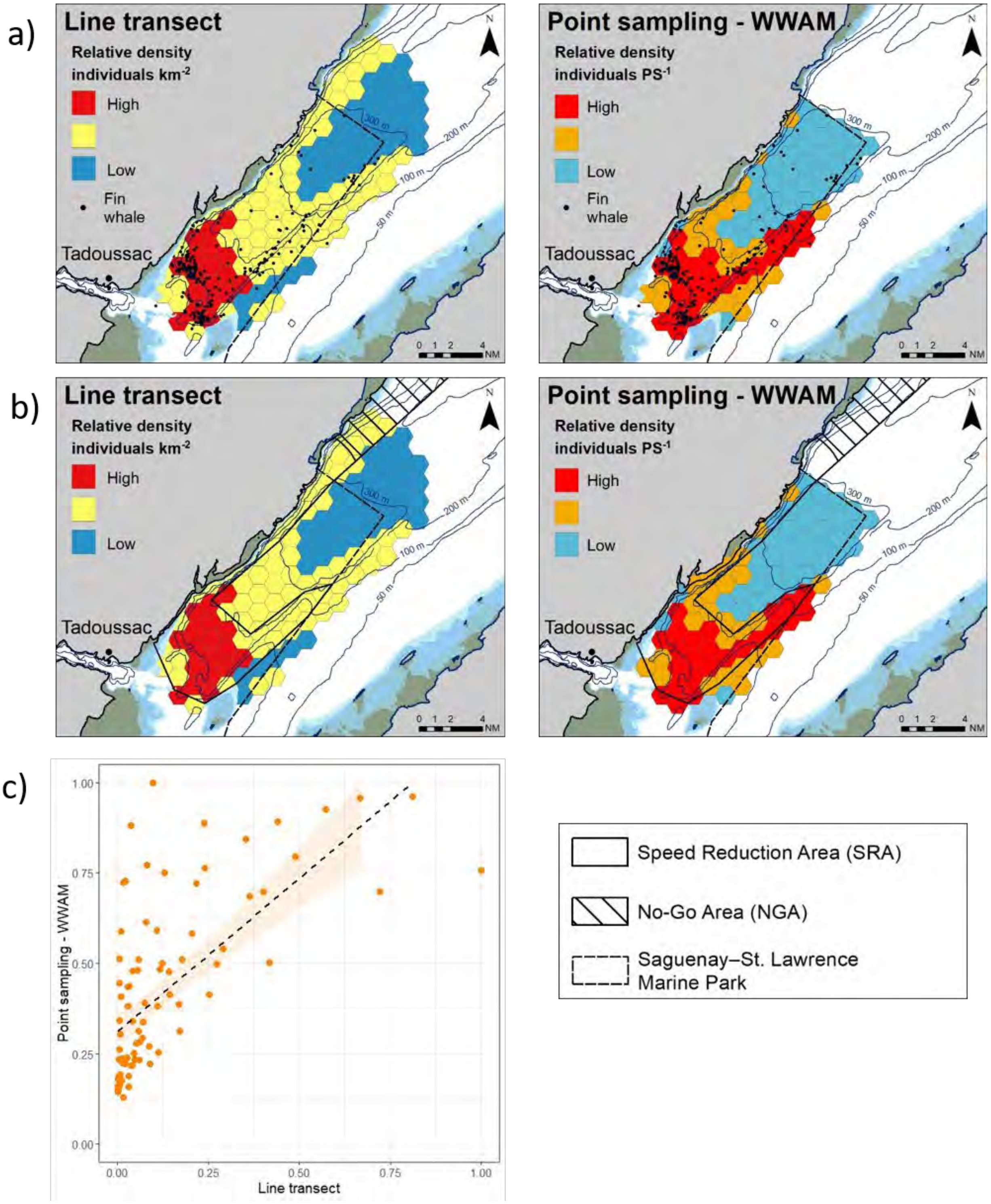


Figure 30. Predicted densities of fin whales derived from line transect (left) and point sampling (PS) (right)models overlaid by a) the validation dataset, b) the current management measures and c) a comparison of the normalised results for the shared grid cells.

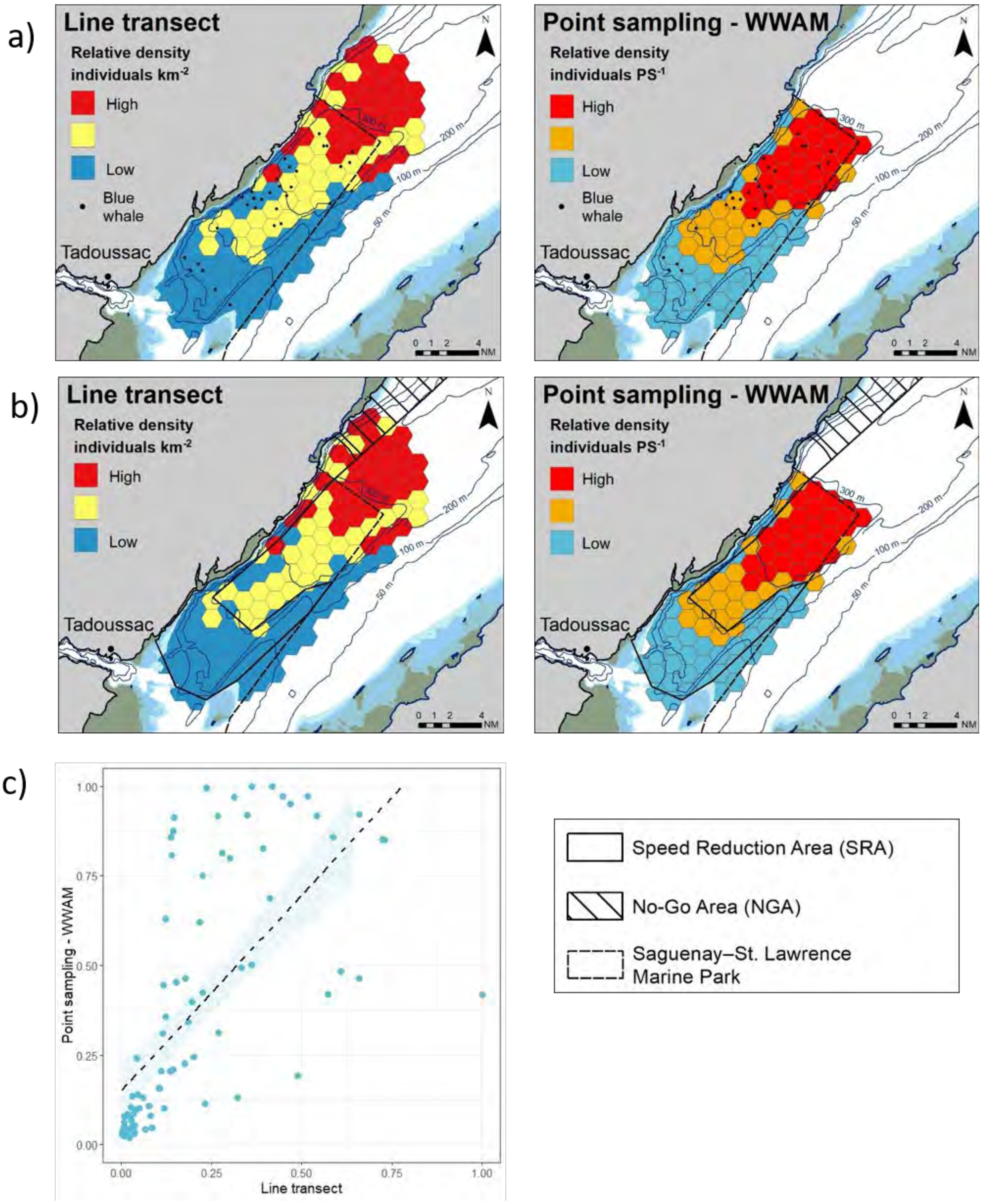


Figure 31. Predicted densities of blue whales derived from line transect (left) and point sampling (PS) (right)models overlaid by a) the validation dataset, b) the current management measures and c) a comparison of the normalised results for the shared grid cells.

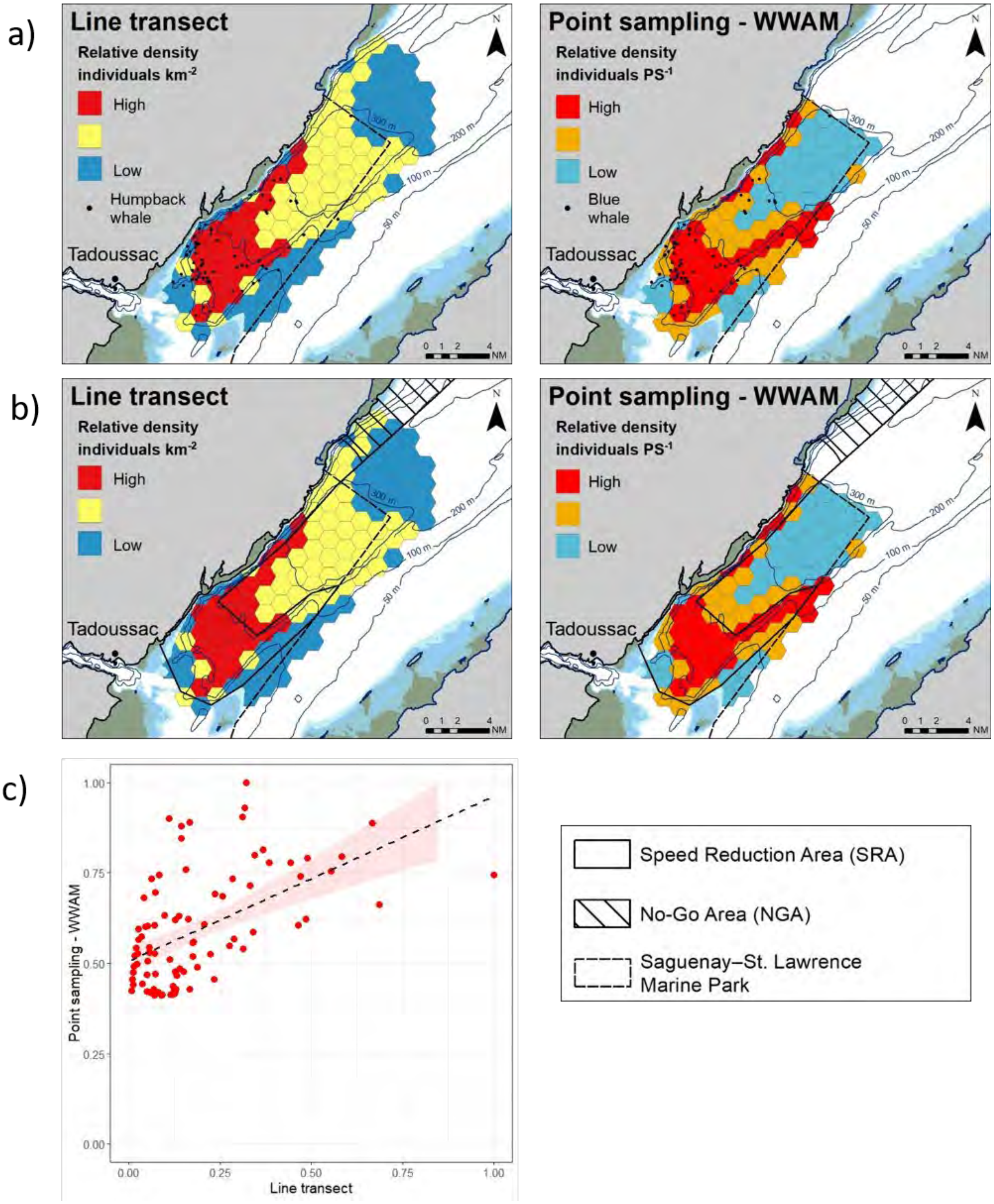


Figure 32. Predicted densities of humpback whales derived from line transect (left) and point sampling (PS) (right) models overlaid by a) the validation dataset, b) the current management measures and c) a comparison of the normalised results for the shared grid cells.

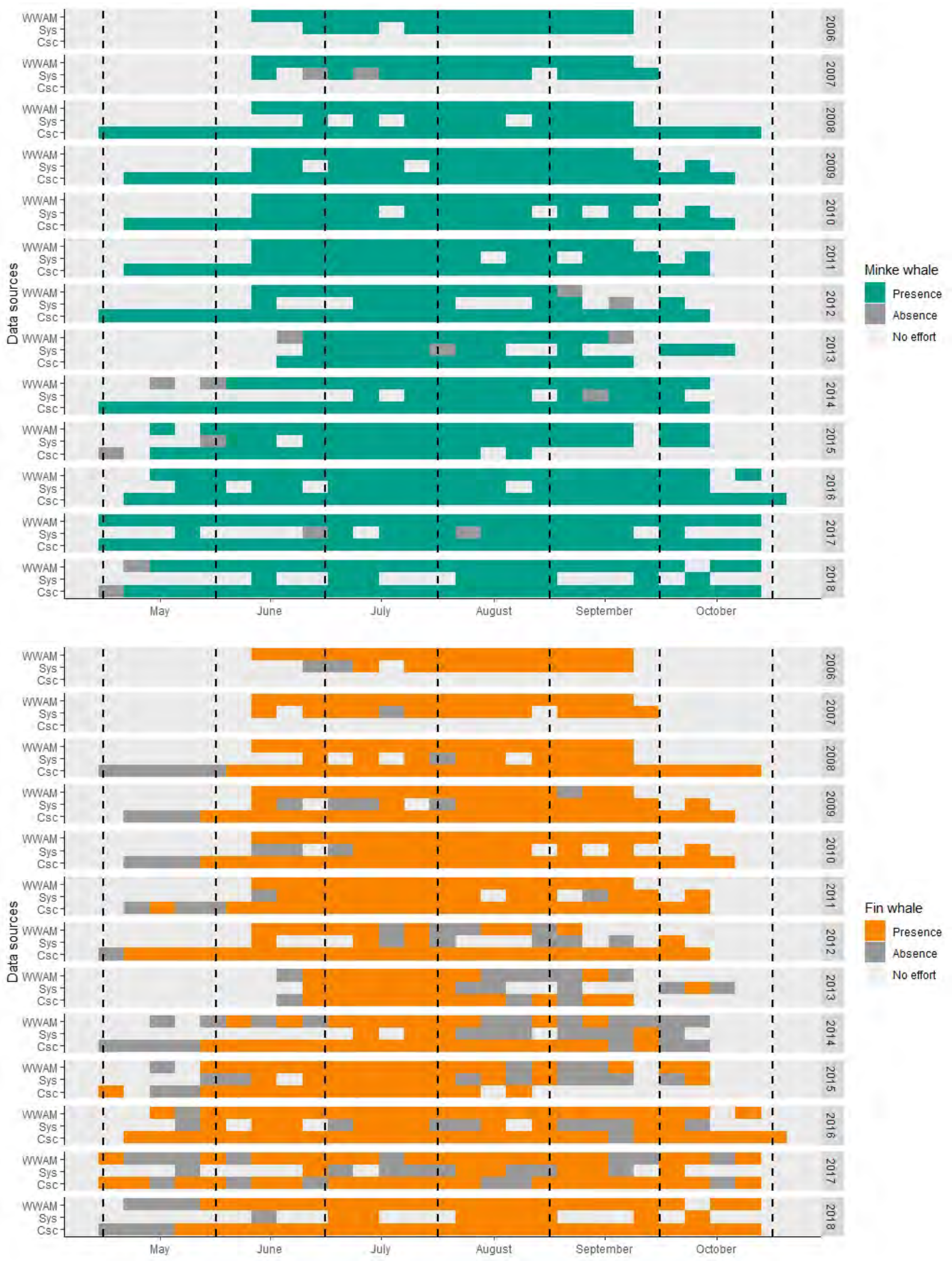


Figure 33. Period of known weekly occurrence of each baleen whale species within the study area from May to October (2006 to 2018) based on presence/absence data of each data source (Csc: Citizen Science; Sys: Systematic line transect surveys; WWAM: whale watching activities monitoring).

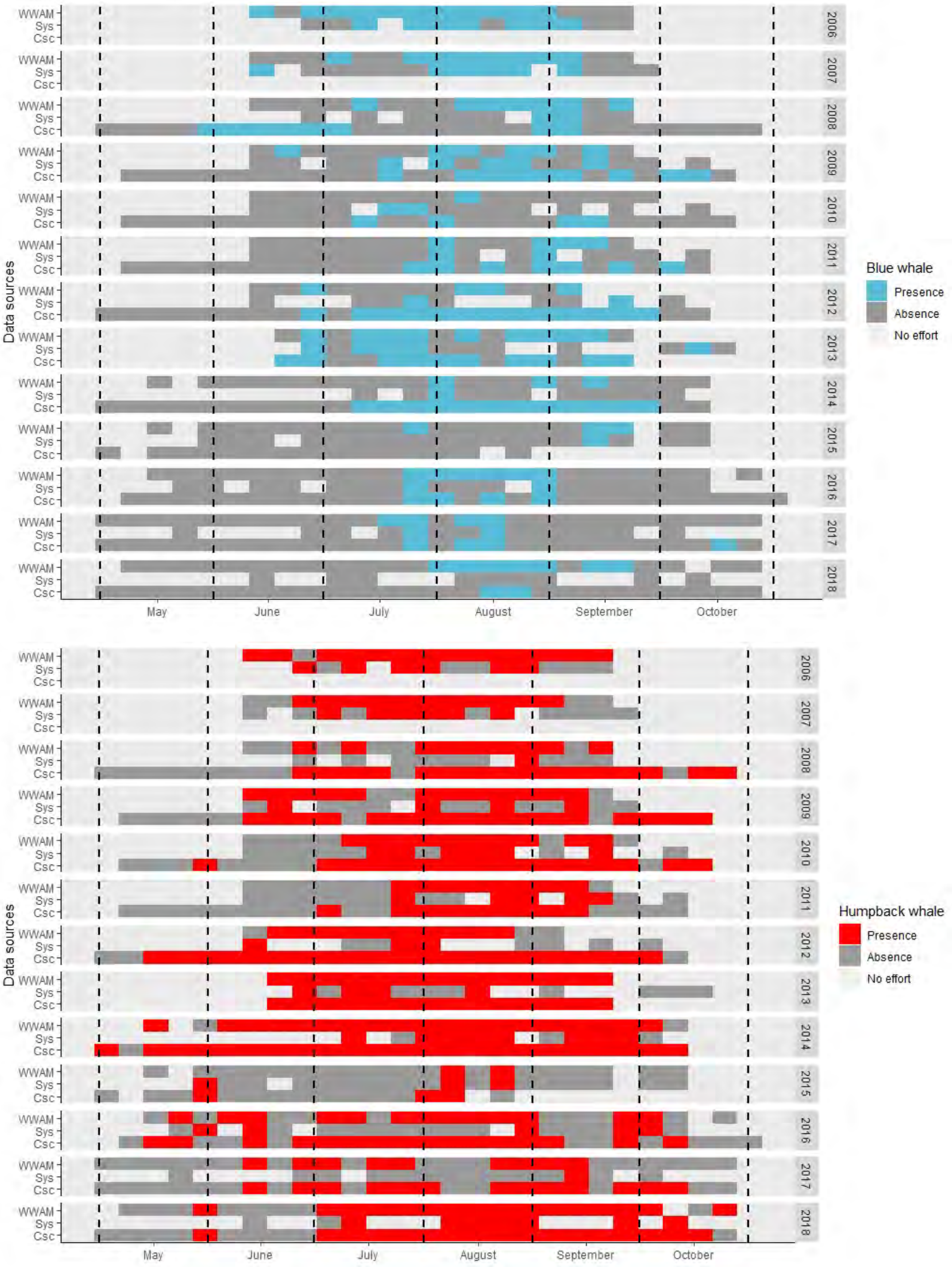
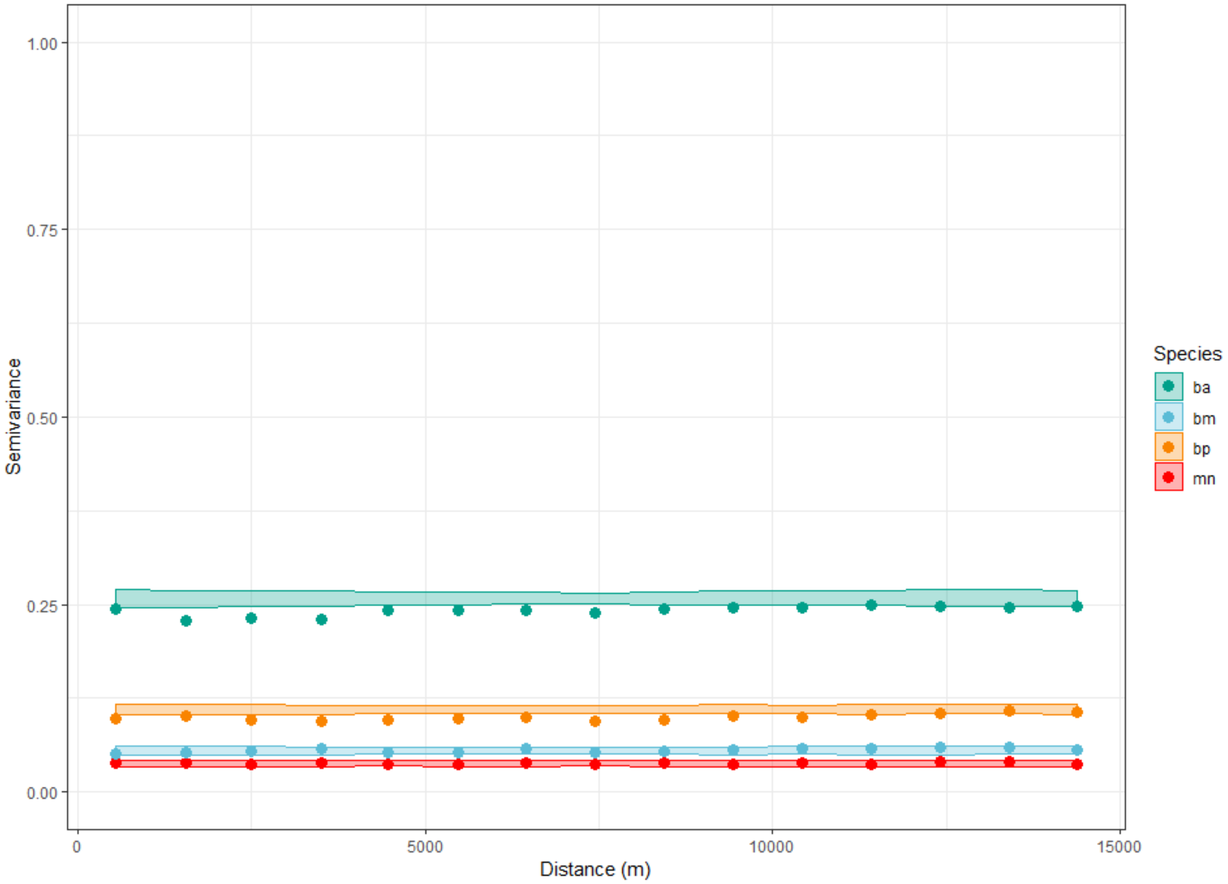
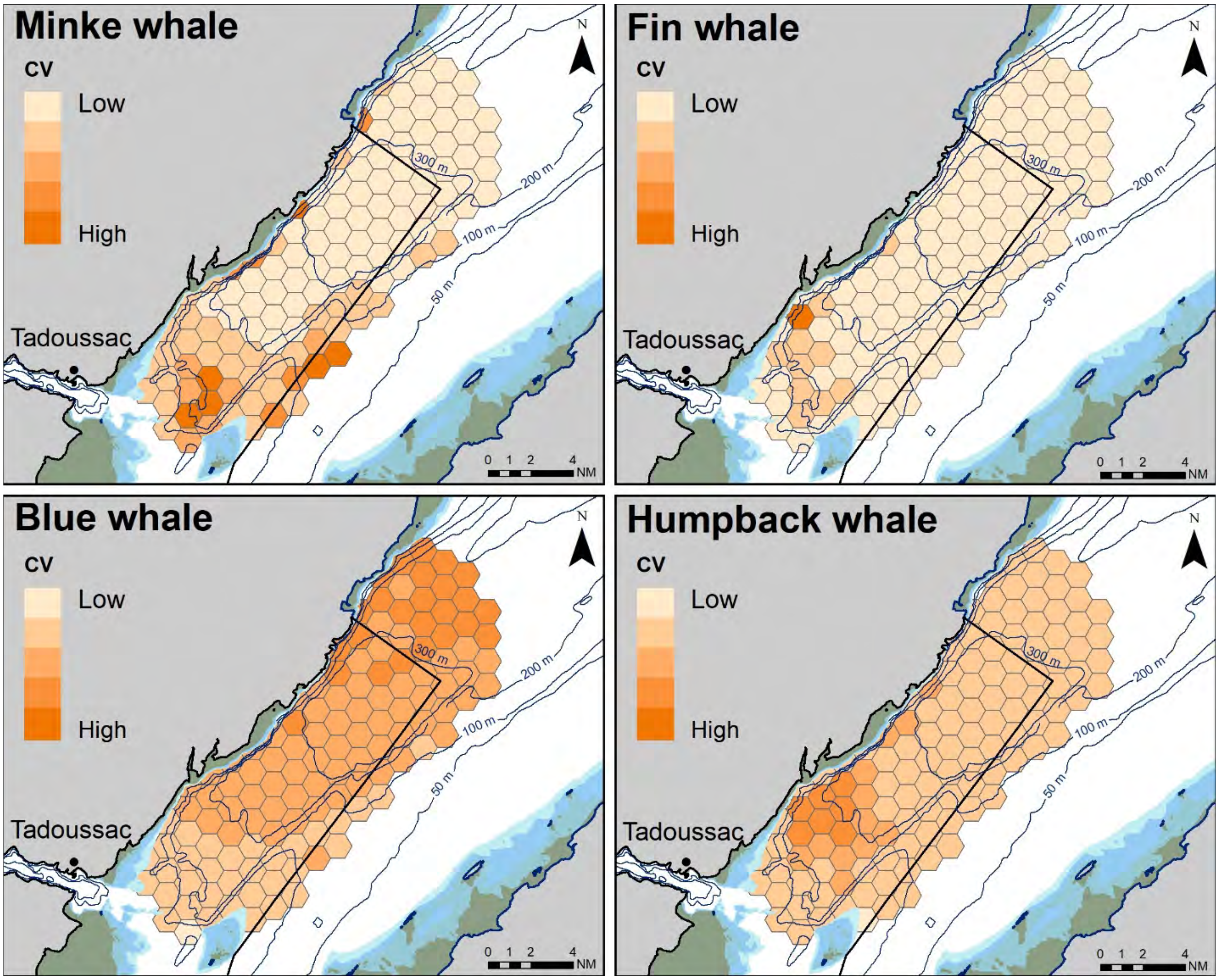


Figure 33. Continued.

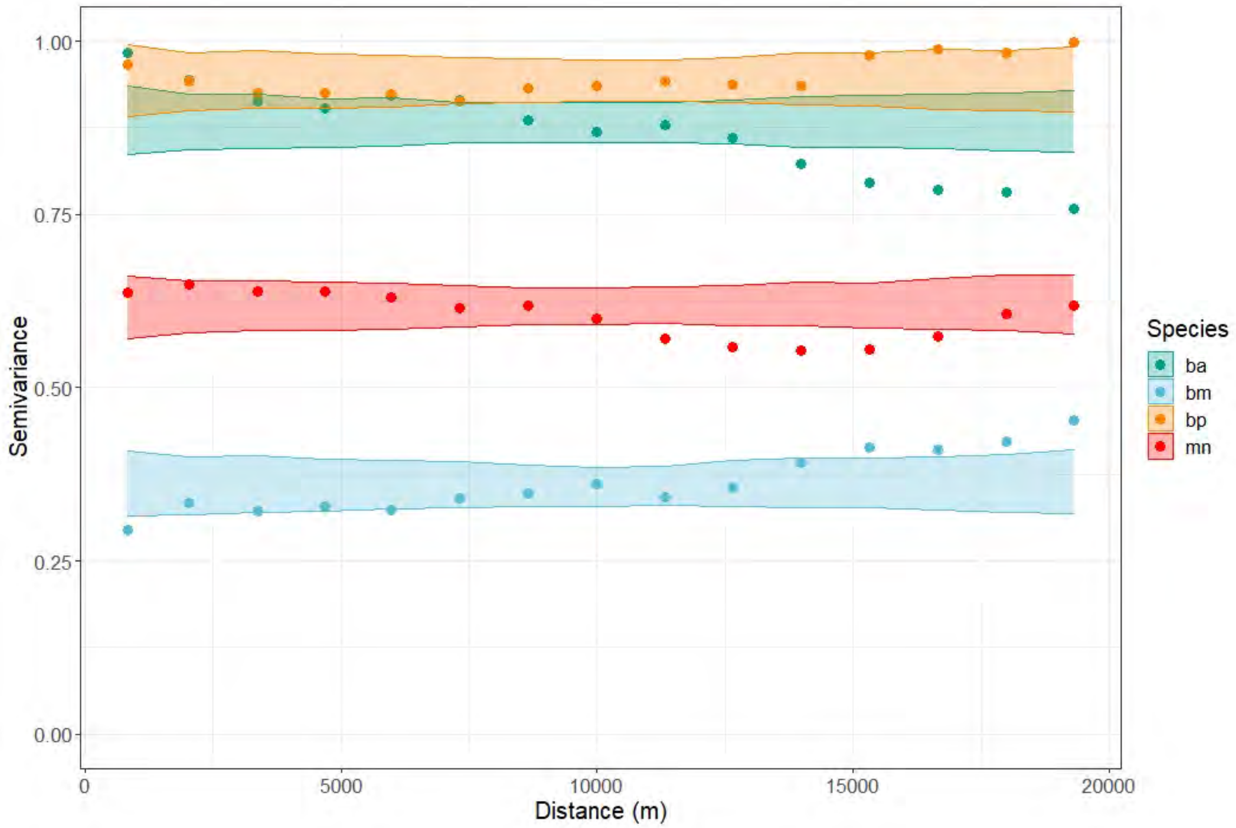
APPENDICES



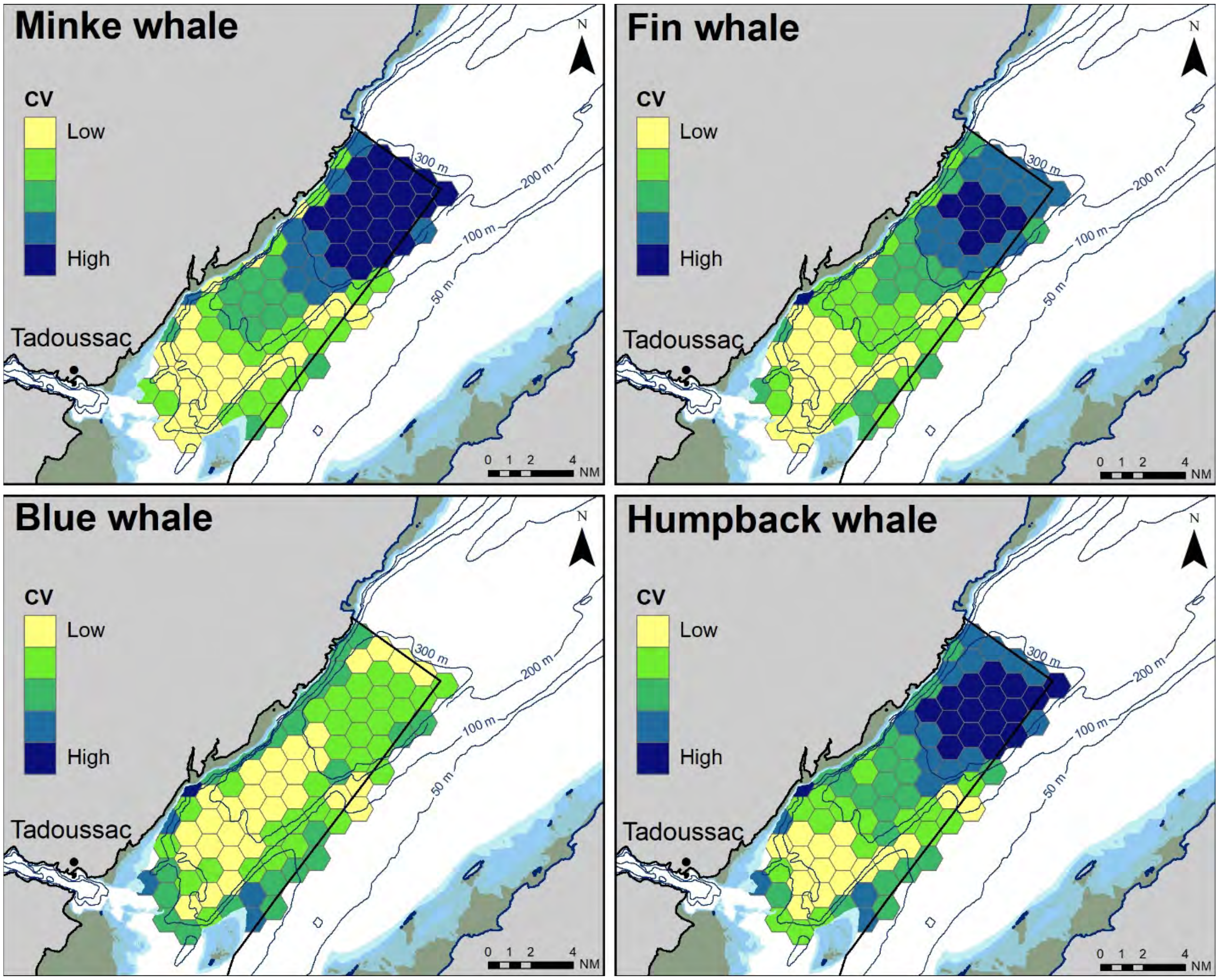
Appendix 1. Empirical variogram and Monte Carlo envelope computed from 999 independent random permutations of the residuals from the models using systematic distance sampling data for each species (ba: minke whales, bm: blue whales, bp: fin whales, mn: humpback whales).



Appendix 2. Coefficient of variation (CV) of the predicted densities of each baleen whale species derived from the models using systematic distance sampling data.



Appendix 3. Empirical variogram and Monte Carlo envelope computed from 999 independent random permutations of the residuals from the models using whale watching activities monitoring data for each species (ba: minke whales, bm: blue whales, bp: fin whales, mn: humpback whales).



Appendix 4. Coefficient of variation (CV) of the predicted densities of each baleen whale species derived from the models using whale watching activities data.