Spatio-temporal patterns in fin whale (*Balaenoptera physalus*) habitat use in the northern Gulf of St. Lawrence

Anna Schleimer[1,2,3,*], Christian Ramp[1,2], Stéphane Plourde[4], Caroline Lehoux[4], Richard Sears[2], Philip S. Hammond[1]

[1]Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, St Andrews, KY16 8LB UK
[2]Mingan Island Cetacean Study, St Lambert, Québec, J4P 1T3, Canada
[3]Marine Evolution and Conservation, Centre for Ecological and Evolutionary Studies, University of Groningen, Groningen, 9700 CC, The Netherlands
[4]Institut Maurice-Lamontagne, Fisheries and Oceans Canada, Mont-Joli, Québec, G5H 3Z4, Canada

ABSTRACT : Significant ecosystem changes in the Gulf of St. Lawrence (GSL) have had far-reaching effects at all trophic levels. The abundance of fin whales (*Balaenoptera physalus*) has declined significantly in the northern GSL over the past decade. This study aimed to test the hypothesis that the observed decline was correlated to changing environmental conditions. Cetacean sighting data from 292 surveys, resulting in 2986 fin whale encounters from 2007 to 2013, were used to fit two separate generalised additive models in terms of (1) bathymetric and oceanographic variables (the proxy model) and (2) modelled krill biomass (the prey model). The concept of "handling time" was introduced to correct for time off search effort, applicable to other studies relying on opportunistically sampled data. While a positive correlation between krill biomass and fin whale numbers was found, the performance of the proxy model (24.2 % deviance explained) was overall better than the prey model (11.8 %). Annual predictive maps derived from the final proxy model highlighted two key areas with recurrently high relative fin whale abundance and a significant overlap with shipping lanes. While both models provided evidence for an annual decline in relative fin whale abundance, static bathymetric features were the most important predictors of habitat use and no correlation between dynamic variables and the decline was found. High resolution prey data and a better understanding of the feeding ecology of fin whales are proposed to further investigate the predator-prey relationship and decline of fin whales in the GSL.

KEY WORDS: Habitat modelling, effort quantification, handling time, proxy variables, distribution, predictive maps, opportunistic surveys

---

*Corresponding author. Email: achs@st-andrews.ac.uk

1. INTRODUCTION

Rapid ecosystem changes in relation to changing environmental conditions have been reported in a wide variety of ecosystems (e.g. Hoegh-Guldberg et al. 2007, Fossheim et al. 2015, Sahade et al. 2015, Thomson et al. 2015). Changes in environmental conditions may affect species directly by challenging their physiological tolerance levels or indirectly by disrupting vital interspecies interactions (Tylianakis et al. 2008). Species may respond with changes in abundance or shifts in distribution (Florko et al. 2018).

The Gulf of St. Lawrence (GSL), Eastern Canada, has seen major and potentially far-reaching ecosystem changes over the past decades due to climate change and anthropogenic pressures. In the early 1990s, overfishing culminated in the collapse of Atlantic cod (*Gadus morhua*) and other large demersal fish stocks, marking a fishery-induced regime shift in the ecosystem (Savenkoff et al. 2007, Bui et al. 2010). Simultaneously, unprecedented warming of incoming North Atlantic water, changes in sea surface temperature (SST), salinity and sea ice extent altered the habitat significantly (Thibodeau et al. 2010, Plourde et al. 2014). Higher mortality rates in response to these ecosystem changes were reported even in higher predators, such as harp seals (*Pagophilus groenlandicus*) and belugas (*Delphinapterus leucas*), highlighting the cascading effects of the changing environmental conditions (Johnston et al. 2012, Plourde et al. 2014).

In this context, this study aimed to study the spatio-temporal patterns in fin whale (*Balaenoptera physalus*) distribution and abundance in the northern GSL. Schleimer et al. (2019) found a significant decline in the number of fin whales using this feeding area and evidence of declining survival rates over the past decade. However, a shift in distribution (i.e. permanent emigration) in response to ecosystem changes in the GSL was also proposed as

57 an possible explanation for the decline in numbers. Fin whales in the GSL have been found to
58 shift arrival dates to the feeding ground in the GSL at a rate of one day earlier per year over
59 three decades, linked to earlier winter sea ice break up and higher SST (Ramp et al. 2015).
60 Fin whale distribution has been correlated with the occurrence of thermal fronts in the GSL
61 (Doniol-Valcroze et al. 2007); however, the physical and biological processes that drive intra-
62 and inter-annual variation in distribution of fin whales in the GSL remain poorly understood.

63 Here, we hypothesise that the observed changes in abundance could be attributed to
64 changes in environmental conditions. To test this hypothesis, spatio-temporal patterns in fin
65 whale distribution were explored within a species distribution model (SDM) framework
66 (Redfern et al. 2006, Forney et al. 2012, Hazen et al. 2017). SDMs aim to identify the
67 underlying factors that drive spatio-temporal trends in species' distribution, offering insight
68 into both the causes of past responses and predictions of future responses to a changing
69 environment (Hazen et al. 2013, Víkingsson et al. 2015). If changes in the environmental
70 conditions in the GSL, as reflected by changes in sea temperature, primary productivity or
71 prey biomass, were at the basis of the observed decline in abundance of fin whales, we
72 expected to detect such a relationship in the SDMs with the retention of dynamic variables
73 in the final model. Extensive survey and effort data collected in the northern GSL over seven
74 summers provided the basis of this study. SDMs frequently use proxy variables that are
75 assumed to be indicative of high productivity and prey distribution (Torres et al. 2008). Here,
76 two separate SDMs were built. The first SDM modelled fin whale relative abundance as a
77 function of commonly used proxy variables for high productivity (including bathymetric and
78 remotely sensed oceanographic variables), while the second SDM used modelled krill
79 biomass as explanatory variable in place of the proxy variables used to derive it (Plourde et
80 al. 2016). Specifically, we wanted to test whether the modelled prey variable would provide

81 better predictive power to a model based on proxy variables to define fin whale habitat.
82 Euphausiids constitute an important component of fin whale diet in the GSL (Gavrilchuk et
83 al. 2014) and a strong link between fin whale distribution and euphausiid biomass is
84 expected. We wanted to test whether euphausiid biomass derived from a model could serve
85 as informative alternative to high-resolution prey data despite the inherent uncertainty that
86 is associated with habitat model predictions.

87 Ideally, cetacean-habitat models are built using data derived from systematic surveys
88 specifically designed to estimate cetacean density and abundance. However, given cost and
89 scheduling limitations imposed by such dedicated surveys, there is a growing interest in
90 developing methods to account for the biases associated with non-systematic or
91 opportunistic surveys (e.g. Williams et al. 2006, Phillips et al. 2009, Isojunno et al. 2012). The
92 fin whale data used in the present study were collected as part of a photo-identification
93 study; as such, the survey design differed from conventional systematic cetacean surveys in
94 three ways; (1) surveys were not designed to ensure equal coverage probability, (2)
95 distance-sampling was not implemented, and (3) search effort was interrupted by the
96 collection of sighting-specific data (e.g. photo-identification and biopsy data). The nature of
97 the data prohibited a design-based approach, necessitating a model-based approach, using
98 generalised additive models (GAMs), which does not require random placement of survey
99 lines in the study area. Additionally, the concept of “handling time” was applied to
100 differentiate between time spent collecting sighting-specific data and search effort. The
101 proposed methods are applicable to other studies that rely on opportunistically collected
102 data, such as cetacean data collected during whale watching activities.

103 **2. MATERIALS & METHODS**

2.1 Cetacean Survey Data

The study area was located in the Jacques Cartier Passage (JCP), between Anticosti Island and the North Shore in the GSL, extending over an area of approximately 8000 km^2 (Fig. 1). The region covers diverse topographic features, such as the head of the Anticosti Channel, the steep slopes along Anticosti Island, and the shallower plateaus of the North Shore and Banc Parent. Upwelling of cold, nutrient rich waters from the cold intermediate layer place the region among the most productive in the GSL, allowing the ecosystem to sustain a high biodiversity (Bourque et al. 1995, Doniol-Valcroze et al. 2007). During summer months, minke whales (*Balaenoptera acutorostrata*), humpback whales (*Megaptera novaeangliae*), fin whales and harbour porpoises (*Phocoena phocoena*) co-occur in the study area, with occasional sightings of blue whales (*B. musculus*) and, more recently, North Atlantic right whales (*Eubalaena glacialis*).

Cetacean survey data were collected by researchers from Mingan Island Cetacean Study from June to October in the years of 2007 to 2013. The data used for the purpose of this study consisted of the non-random survey tracks and the position, timing, and group size of each fin whale sighting. Surveys were conducted using rigid-hulled inflatable boats and focussed on the collection of photo-identification data of large rorquals. The survey design is therefore best described in terms of “whaler-behaviour” meaning that captains targeted areas where they expected to find animals to maximise photo-identification effort. The surveys covered as large an area as possible until groups of whales were found. Boat speed varied between 15 and 20 knots, with occasional stops to scan the horizon with binoculars for blows. For safety reasons, two boats conducted surveys simultaneously, but they generally covered different areas. Surveys were terminated when weather conditions deteriorated to Beaufort scale >3 or visibility <1nm.

Once an individual or group of whales was detected, the time of the sighting was noted and the animals were approached slowly for collection of photo-identification and sometimes biopsy data. The exact position was recorded at the 'fluke print' where the animal dived after its surfacing sequence. The group size of animals was recorded and individuals were attributed field-ID numbers to keep track of individuals on subsequent sightings. Field-ID numbers and photo-identification data were used to determine whether individuals had already been sighted previously. If individuals were not identifiable, the group was considered to be a new sighting.

2.2 Data Processing

2.2.1 Effort Quantification

The survey track was recorded using a LOWRANCE LMS-480M GPS unit (precision ≤30 m) on each survey boat. The resulting survey tracks were used to calculate survey effort. Due to variable boat speeds and *ad libitum* survey tracks, the length of the survey track was not considered an appropriate measure of effort. Instead effort was defined as the time in seconds spent searching for animals within a grid cell (see below). Timestamps were retrieved from the GPS tracks to estimate the effort spent in each grid cell (see Supplementary Fig. S1).

A grid-based modelling framework was adopted in accordance with previous studies dealing with non-systematic survey designs (Cañadas et al. 2005, Isojunno et al. 2012). The study area was divided into 5 x 5 km grid cells in which the number of fin whale individuals, effort and environmental covariates could be summarised. The size of the grid cells was chosen based primarily on the resolution of available remotely sensed and modelled covariates (Table 1).

While time identified as 'off effort' was excluded from the calculations, the strong focus on photo-identification and biopsy sampling made further modification of the effort data necessary. When the researchers were concentrating on obtaining photo-identification data and taking biopsy samples, all effort was focused on a single individual or group, rather than searching for new groups. Because the surveys covered all cetacean species encountered, such interruptions of search effort were not limited to fin whale encounters. Because of similar bias, Isojunno et al. (2012) did not consider duration of grid cell visits to be an adequate measure of effort. Here, we corrected the total time spent in each grid cell by removing effort associated with the collection of sighting-specific data. From the survey data it was possible to identify sequential re-sightings of the same group during which such data were taken. The time from the first re-sighting to the last re-sighting of a group was characterised as "handling time" and considered off effort. A similar approach has been used to calculate search effort of whalers, where the chasing and processing of caught whales was considered "handling time" and was excluded from the general search effort (Sigurjónsson 1988, Sigurjónsson & Gunnlaugsson 2006). This approach more accurately reflected time spent searching for fin whales and allowed us to use time as an offset in our models.

2.2.2. Environmental Data

Environmental data were chosen based on 1) their importance in previous cetacean species distribution models and 2) their availability at a sufficiently fine spatial resolution with respect to the 25 km^2 grid resolution of the sightings and effort data (Table 1). The data set was subdivided into months to allow seasonal and inter-annual variation in time-variable covariates (SST, chlorophyll *a*, krill biomass) to be incorporated. Month was chosen as an appropriate time period to minimise gaps in remotely sensed data, which tend to have significantly fewer missing data due to cloud cover when summarised per month compared

to daily or weekly resolutions. Fin whale sighting and survey effort data were pooled for each month of the field season, resulting in a maximum of 35 (five months x seven years) temporal sub-units per grid cell. Greene and Pershing (2000) proposed a conceptual model linking North Atlantic Oscillation (NAO), physical environmental conditions, and zooplankton in the northwest Atlantic. This distant potential link was explored in our analyses by including NAO indices as explanatory variables. Hurrell's PC-based NAO index was used in this study for monthly and winter NAO indices (Hurrell et al. 2003). Previous studies have shown that abundance of balaenopterids typically lag behind maximum primary productivity by several weeks (Croll et al. 2005, Visser et al. 2011, Ramp et al. 2015). The possible effect of a temporal lag in the response of fin whale distribution to proxy variables was assessed by including composite spring SST and chlorophyll *a* concentrations (Chl *a*) and lagged winter NAO indices.

In the absence of high-resolution euphausiid data covering the entire study area/period, krill biomass was derived from a krill habitat model as described in Plourde et al. (2016). Briefly, Plourde et al. (2016) modelled krill biomass spatial and temporal distribution in eastern Canadian waters as a function of static (bathymetry and slope) and dynamic (SST, Chl *a*, sea level height anomaly) environmental variables in a GAM framework. High-resolution quantification of euphausiid biomass was available from multifrequency acoustic data collected from surveys in the GSL (including the JCP) and adjacent waters since 2000 (McQuinn et al. 2015). The final euphausiid biomass model explained 24.5 % of deviance and was used to get monthly predictions of krill biomass at 5 x 5 km resolution in the JCP for the present study. Due to the spatial overlap of both studies, no extrapolations beyond the range of explanatory variables were necessary.

2.3. Data Analysis

199 The relationship between the number of fin whale individuals in each grid cell (response
200 variable) and the explanatory variables was modelled using GAMs (Hastie & Tibshirani 1990,
201 Wood 2017), which are commonly used to study spatial and temporal drivers in cetacean
202 distribution because of their flexibility (Redfern et al. 2006, 2017, Isojunno et al. 2012,
203 Becker et al. 2016). Only grid cells with search effort were used to build the model. GAMs
204 were fitted in the R (v. 3.2.3, R Core Team 2015) package *mgcv* (v. 1.8-25; Wood 2017).

205 Two separate models were built to model temporal and spatial patterns in fin whale
206 distribution. The first model (the proxy model) included all static and dynamic environmental
207 proxy variables (all variables listed in Table 1, except for krill biomass), including month and
208 year. The second model (the prey model) included only modelled krill biomass, month and
209 year, thus excluding the remaining proxy variables (most of which were used in the model to
210 predict krill biomass). In both models, the number of individual fin whales per grid cell was
211 modelled as a negative binomial distribution with logarithmic link function. The response
212 variable was characterised by a high frequency of zeros (3207 grid cells without sightings
213 compared to 312 grid cells with sightings) and the negative binomial error distribution
214 provided the best fit to the data (Supplement S2). The negative binomial distribution has
215 been used in previous studies with similar types of data (i.e. count data with many zeros and
216 overdispersion; Warton 2005, Virgili et al. 2017). The natural logarithm of monthly search
217 effort was included as an offset term in the model to account for variable search effort
218 across the study area. Only the first encounter of a fin whale individual/group was counted
219 towards the monthly sum of fin whales in each grid cell to avoid the inclusion of duplicate
220 sightings. Sightings data collected on the same day from different survey boats were treated
221 independently because the spatial coverage differed between boats.

Prior to model fitting, explanatory variables were inspected for collinearity using the pairs function from the *AED* package in R, which generates a correlation matrix for pair-wise comparison between variables (Zuur et al. 2009). Two variables were deemed collinear if the estimated Pearson correlation coefficient exceeded 0.6. No collinearity was detected among covariates (Supplement 3, Fig. S3). Chl *a* and krill biomass values were log transformed ($log(X + 1)$) to reduce skewness in the data. Field observations suggested that fin whales fed in shallower waters on the North Shore in June and July; interaction terms of month with depth and with aspect were thus considered to explore whether the data supported these relationships.

Restricted maximum likelihood (REML) was used for smoothing parameter estimation (Marra & Wood 2011). The gamma term, which acts as an additional penalty, was set to 1.4 to reduce over-fitting in cases with relatively few observations per variable (Kim & Gu 2004, Wood 2006). Full models with and without interaction terms were fitted with penalised cubic regression splines and tensor products (*ti*) for interaction terms. A cyclic regression spline was fitted to aspect (0-360°) to match start and end points. Shrinkage spline smooths were used for covariate selection. The shrinkage approach penalises the null space of the smooth function, reducing the degrees of freedom of unsupported covariates to zero, allowing multiple terms to be dropped from the full model in a single step (Marra & Wood 2011).

Models with and without interaction terms were compared using Akaike Information Criterion (AIC; Akaike 1972, Wood et al. 2016), percentage of deviance explained, and average squared prediction error (ASPE). To calculate the latter, a five-fold cross-validation was applied to assess the performance of candidate models in predicting novel data. Data were randomly split into five subsets. Models were fitted to 80% of the data for model

246 training and the remaining 20% of the data were subsequently used for validation of
247 predictions based on the trained model. ASPE was calculated as the mean squared
248 difference between predicted and observed fin whale numbers in the validation subset. This
249 cross-validation was run five-times in total and the mean ASPE was retained for model
250 selection.

251 The final chosen model was refitted with the complete data set. If terms with an
252 approximate *P*-value >0.05 remained in the model after shrinkage, the covariate with the
253 least significant *P*-value was dropped from the model. If the exclusion of the variable did not
254 increase the AIC score, the reduced model was retained. The relative covariate importance
255 was estimated with the R function *varImpBiomod* (Thuiller et al. 2009). Model residual plots
256 were examined visually to verify that assumptions of normality and variance homogeneity
257 were met (Figs. S4 & S5). Spatial autocorrelation of model residuals was assessed using a
258 variogram (Zuur et al. 2009).

259 ***2.4. Prediction***

260 The final proxy and prey models were used to compute predictions of relative abundance
261 (individuals/hour) in each grid cell. Predictive maps were generated for each year, fixing the
262 offset term to one hour of effort in each grid cell per month, in the open source GIS software
263 package Quantum GIS (QGIS v. 2.18.1; QGIS Development Team, 2016). Because the model
264 yielded separate predictions for each month, the mean relative abundance per year was
265 plotted. To assess prediction uncertainty, coefficients of variation (CV) were calculated
266 based on posterior simulation. From the posterior distributions of the model coefficients,
267 1000 coefficient vectors were simulated using *mvrnorm* from the R MASS library (Venables
268 and Ripley 2002) and were used to generate 1000 predictions. The mean and CV were

calculated from these 1000 predictions. The performance of the proxy and prey models were evaluated by comparing the percentage deviance explained and the predictive maps derived from the final models.

3. RESULTS

Sightings and effort data from 292 dedicated cetacean surveys were available to investigate temporal and spatial patterns in fin whale habitat use in the JCP. In total, 1878 hours were spent on effort, of which 510 hours were characterised as handling time during which the researchers were collecting photo-ID or biopsy data, leaving 1368 hours of corrected effort (Table 2, Fig. 2). Overall, 2986 individual fin whales were recorded on effort, with an average group size of 2.19 animals (SE = 0.05). Average annual encounter rates and median group sizes decreased over the study period (Table 2).

3.1. Proxy-fin whale distribution models

Out of the five models fitted with proxy variables, the model which included an interaction term of aspect and month performed best in terms of AIC, percentage deviance explained and ASPE (Table 3). Distance to coast, Chl *a*, spring Chl *a*, spring SST, NAO index, and lagged winter NAO indices were shrunk to zero degrees of freedom by the shrinkage regression splines and simultaneously dropped from the model (Model 1.3 Table 3). Winter NAO index was subsequently dropped from the model, because it was the only term with an approximate *P*-value > 0.05 and very low effective degrees of freedom (edf = 0.19). The resulting final proxy model explained 24.2 % of deviance.

Among the covariates retained in the final model, water depth and aspect were of the highest importance, with fin whales occurring in greater numbers in deeper waters, over

steep, and northward facing slopes (Fig. 3). Higher numbers were also recorded at SST greater than 12°C. Temporal trends suggested a peak in fin whales at the onset of the survey season in June, followed by a decline until September, and a second peak at the end of the season in October. The affinity to northward facing slopes changed by month, showing that occurrence at southward facing slopes was less likely in August and September, compared to June, July and October. The negative yearly trend that was already reported for the annual fin whale encounter rates was also reflected in the final model.

3.2. Prey- fin whale distribution models

The prey model that included an interaction term between krill biomass and month had the lowest ASPE and AIC score and highest percentage of deviance explained (11.8 %) among all three built models (Table 3). Krill biomass had the highest importance among the model covariates, followed by month and year. The intra- (month) and inter-seasonal (year) patterns followed the same trends as described for the proxy model (Fig. 4). Fin whale numbers increased with higher modelled krill biomass, although the interaction term indicated that fin whales also occurred in areas with lower krill biomass at the onset of the season (June and July).

3.3. Prediction

Annual predictive maps of average fin whale occurrence generated from the final proxy and prey fin whale models are shown in Figs. 5 & 6, with CV in Supplement 4 (Figs. S6 & S7). From the proxy model, two main areas with consistently high predicted relative abundance of fin whales were identified: the western end of the Anticosti Channel and the area north of Banc Parent (see Fig. 1 for locations). The area off Banc Parent coincides with the southern branching traffic shipping lanes. The predictive maps indicated a potential third high density

area on the northern edge of the Laurentian Channel. However, this area of very high predicted relative abundance lies at the very southwestern edge of our survey area and could represent an “edge effect” because the area is the deepest in the surveyed area with little effort extending that far. A clear annual decline in fin whale numbers was evident from the predictive maps.

Predictions from the prey model favoured a more even spatial distribution of fin whales across the JCP. The head of the Anticosti Channel to the east and the southwestern area of the study area seemed to have overall the highest predicted numbers, but the strong signal of the annual negative trend masked areas with consistently high numbers.

4. DISCUSSION

SDMs were fitted to understand the extent to which the observed decline in fin whale numbers was a result of changing environmental conditions in the northern GSL . The proxy and prey models both identified the negative annual trend in the number of fin whale individuals, but the proxy model had overall a better predictive performance than the prey model. Here, we discuss the link between the observed decline in fin whales and the spatio-temporal patterns that were revealed by the SDMs.

Over the study period, the majority of sightings clustered around the head of the Anticosti Channel and north of Banc Parent with some inter-annual variability. This distribution was best reflected in the predictive maps of the proxy model, while the prey model largely failed to highlight those high density areas. The static bathymetric features in the areas with consistently high predicted fin whale numbers, characterised by deep water and steep, northward facing slopes, were the most important predictors in the proxy model. Among all the dynamic covariates (Chl *a*, SST, NAOI), which could explain the inter-annual variability in

337 sightings, only SST was retained in the final proxy model. Fin whale numbers increased in
338 waters with higher SST, suggesting that cooling of SST could have led to the observed annual
339 decline. However, a trend analysis showed that the GSL is undergoing warming of SST, with
340 the northeastern Gulf warming at a faster pace than the southern part of the Gulf (Galbraith
341 et al. 2012, Larouche & Galbraith 2016). In our study area, both the lowest and highest
342 average SST in the study area were recorded in 2007 and 2008, respectively, which were also
343 the years with the highest encounter rates (Fig. S8). Since 2012, near-record temperatures of
344 both surface and deep layers of the GSL were found to correlate with variations in plankton
345 phenology and fish abundance (Plourde et al. 2015, Brosset et al. 2018). While results
346 presented here suggest that a direct correlation between decreasing fin whale abundance
347 and SST is unlikely, it remains unclear to what extent cascading effects of a warming Gulf
348 could have impacted fin whale abundance and/or distribution indirectly.

349 The final proxy model explained 24.2 % of the variability in the data, indicating that
350 important explanatory variables were missing from the model. On a feeding ground, a strong
351 predator-prey relationship is expected in baleen whales (Piatt et al. 1989, Ressler et al.
352 2015). No real-time, high-resolution euphausiid data were collected during the cetacean
353 surveys, so we used modelled krill biomass to test how well it explained fin whale relative
354 abundance compared to a model using proxy covariates. While the prey model found a
355 positive relationship between modelled krill biomass and fin whale numbers, the model
356 performed poorly overall compared to the proxy model in terms of percentage of deviance
357 explained and predictive power. The modelled krill biomass variable was thus not a suitable
358 alternative to the proxy variables in this study. The uncertainty associated with the krill
359 biomass covariate (predicted from a model that explained 24.5% of deviance (Plourde et al.
360 2016), could have decreased its power as a predictor on a fine spatial scale. This does not

361 preclude a better predictive performance at larger spatial scales. Previous models found
362 differing relationships between fin whale and euphausiid abundance, possibly due to
363 differences in spatial scales (Zerbini et al. 2016). Laidre et al. (2010) highlighted the
364 importance of high spatio-temporal synchrony in the collection of prey and cetacean data to
365 quantify their relationship. We therefore recommend to explore the performance of
366 modelled krill biomass as predictor of baleen whale distribution at broader spatial scales in
367 the GSL.

368 Another factor that could have contributed to the lower performance of the prey model is
369 the generalist diet of fin whales. While euphausiids are an integral part of their diet, fin
370 whales are also known to switch prey depending on availability (Gavrilchuk et al. 2014,
371 Ressler et al. 2015). The inclusion of interaction terms in both models indicated that habitat
372 use changed as the season progressed. The higher number of fin whales found on southward
373 facing slopes and at lower krill biomass at the beginning of the season (June-July) coincided
374 with the rolling of capelin (*Mallotus villosus*) along the North Shore (MPO 2003). To fully
375 quantify the complex predator-prey relationship in fin whales, we need to gain a better
376 understanding of their feeding ecology, especially threshold values at which prey switching
377 occurs, and obtain higher (spatial and temporal) resolution data from all potential prey
378 species. In the absence of such data, it cannot be excluded that inter-annual variability in
379 prey availability was, at least partly, the cause of the observed annual decline in fin whale
380 numbers in the northern GSL.

381 In addition to environmental variability, anthropogenic pressures could affect habitat use
382 and relative abundance. The high density area identified north of Banc Parent coincided with
383 the southern branch of the shipping lanes. In fact, more than a fifth (22.6 %) of all fin whale
384 sightings in this study occurred inside the shipping corridor, posing a considerable risk of ship

385 collisions and noise pollution. Fin whales are the most commonly reported species in the
386 current global vessel strike data set maintained by the Scientific Committee of the
387 International Whaling Commission (Van Waerebeek & Leaper 2008, Van Der Hoop et al.
388 2013). Based on marine mammal stranding records in the GSL from 1994 to 2008, ship
389 collision was the most common anthropogenic trauma for fin whales (22%; Truchon et al.
390 2018). Shipping traffic is projected to increase in the GSL with the opening of the Northwest
391 Passage (Pizzolato et al. 2016). The predicted areas of high fin whale density described here
392 should be included in future risk assessments to mitigate the potential impact of shipping on
393 fin whales (Redfern et al. 2013). Recommended measures could include vessel speed limits
394 and area avoidance recommendations, which were shown to significantly reduce ship strikes
395 with North Atlantic right whales (Laist et al. 2014).

396 While the modelling conducted could not provide a clear indication of the cause of the
397 annual fin whale decline, it did offer valuable insights into spatio-temporal patterns of fin
398 whale habitat use in the northern GSL. Importantly, the predictions derived from the proxy
399 model highlighted two key areas with recurrently occurring high fin whale abundance. The
400 bathymetric features which characterise those areas were in line with previous findings,
401 which have also found water depth and slope to be important predictors of fin whale
402 occurrence in the Mediterranean Sea (Panigada et al. 2005, Azzellino et al. 2012, Pennino et
403 al. 2016), in the northeastern Atlantic (Víkingsson et al. 2015), and the Bay of Fundy
404 (Woodley & Gaskin 1996, Ingram et al. 2007). Krill and capelin aggregate along shelf breaks
405 and steep slopes as a result of tidal currents and upwelling in the GSL and St. Lawrence
406 Estuary (Simard et al. 2002, Cotté & Simard 2005). The two high fin whale density areas
407 coincide with the two areas of above average krill biomass accumulation identified in the JCP
408 by large-scale hydroacoustic surveys (McQuinn et al. 2015). A potential third high density

area was predicted along the northern slopes of the Laurentian Channel, which received little survey effort during this study. This predicted high density area could be an edge effect (i.e. an artefact); future surveys of this area are needed to identify whether or not this area is important habitat for fin whales.

This study has shown how data collected on surveys primarily designed for other purposes can be adapted for habitat modelling analysis. However, this *ad hoc* adaptation of the data leads to a number of concessions. In the absence of distance-sampling and a design ensuring equal coverage probability, it was not possible to estimate absolute density or abundance throughout the study area using design-based methods. While the model-based approach used here accounted for uneven distribution of effort through the inclusion of an offset term, we were able to describe only relative variability in abundance and distribution. The focus on sampling individuals rather than space further compromised search effort data. Such a disruption of search effort could lead to bias in the effort quantification and the inclusion of duplicate sightings, when previously encountered animals catch up with the survey boat. The particular setup of this study allowed us to identify duplicate sightings from the photo-identification data and to correct for handling time based on detailed field notes. Without standardised sampling design, data from opportunistic platforms generally require data-specific solutions. However, the data described here share many similarities with data collected from other platforms of opportunity, such as whale watching boats. We therefore propose that the correction of effort for handling time is applicable to other data sets compromised by disrupted search effort, and its application could allow hitherto unused data to provide useful information on distribution and habitat use.

Acknowledgments

We thank the sponsors and supporters of the Mingan Island Cetacean Study (MICS) and its numerous volunteers, team members, and captains for data collection and handling over all these years. A. Schleimer was supported by the Luxembourg National Research Fund (FNR; AFR/11256673) and Odyssea. We are also grateful to three anonymous reviewers for constructive comments on the manuscript.

Literature Cited

Akaike H (1972) Information theory and an extension of the maximum likelihood principle. Proc 2nd Int Symp Inf Theory, Supp to Probl Control Inf Theory:267–281

Azzellino A, Panigada S, Lanfredi C, Zanardelli M, Airoldi S, Notarbartolo di Sciara G (2012) Predictive habitat models for managing marine areas: Spatial and temporal distribution of marine mammals within the Pelagos Sanctuary (Northwestern Mediterranean Sea). Ocean Coast Manag 67:63–74

Becker EA, Forney KA, Fiedler PC, Barlow J, Chivers SJ, Edwards CA, Moore AM, Redfern JV (2016) Moving towards dynamic ocean management: How well do modeled ocean products predict species distributions? Remote Sens 8:1–26

Bourque MC, Kelley DE, Bourque M, Kelley DE (1995) Evidence of wind-driven upwelling in Jacques-Cartier Strait. Atmosphere-Ocean 33:621–637

Brosset P, Doniol-Valcroze T, Swain DP, Lehoux C, Beveren E Van, Mbaye BC, Emond K, Plourde S (2018) Environmental variability controls recruitment but with different drivers among spawning components in Gulf of St. Lawrence herring stocks. Fish Oceanogr 28:1–17

Bui AOV, Ouellet P, Castonguay M, Brêthes JC (2010) Ichthyoplankton community structure in the northwest Gulf of St. Lawrence (Canada): Past and present. Mar Ecol Prog Ser 412:189–205

Cañadas A, Sagarminaga R, de Stephanis R , Urquiola E, Hammond PS (2005) Habitat preference modelling as a conservation tool: proposals for marine protected areas for cetaceans in southern Spanish waters. Aquat Conserv Mar Freshw Ecosyst 15:495–521Cotté C, Simard Y (2005) Formation of dense krill patches under tidal forcing at whale feeding hot spots in the St. Lawrence Estuary. Mar Ecol Prog Ser 288:199–210

462 Croll D, Marinovic B, Benson S, Chavez F, Black N, Ternullo R, Tershy B (2005) From wind to
463 whales: trophic links in a coastal upwelling system. Mar Ecol Prog Ser 289:117–130

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

467 Florko KRN, Bernhardt W, Breiter CJC, Ferguson SH, Hainstock M, Young BG, Petersen SD
468 (2018) Decreasing sea ice conditions in western Hudson Bay and an increase in
469 abundance of harbour seals (*Phoca vitulina*) in the Churchill River. Polar Biol 41:1187–
470 1195

471 Forney KA, Ferguson M, Becker E, Fiedler P, Redfern JV, Barlow J, Vilchis I, Ballance L (2012)
472 Habitat-based spatial models of cetacean density in the eastern Pacific Ocean. Endanger
473 Species Res 16:113–133

474 Fossheim M, Primicerio R, Johannesen E, Ingvaldsen RB, Aschan MM, Dolgov AV(2015)
475 Recent warming leads to a rapid borealization of fish communities in the Arctic. Nat
476 Clim Chang 5:673–677

477 Galbraith PS, Larouche P, Chassé J, Petrie B (2012) Sea-surface temperature in relation to air
478 temperature in the Gulf of St. Lawrence: Interdecadal variability and long term trends.
479 Deep Res Part II 77–80:10–20

480 Gavrilchuk K, Lesage V, Ramp C, Sears R, Bérubé M, Bearhop S, Beauplet G (2014) Trophic
481 niche partitioning among sympatric baleen whale species following the collapse of
482 groundfish stocks in the Northwest Atlantic. Mar Ecol Prog Ser 497:285–301

483 Greene CH, Pershing AJ (2000) The response of *Calanus finmarchicus* populations to climate
484 variability in the Northwest Atlantic: basin-scale forcing associated with the North
485 Atlantic Oscillation. ICES J Mar Sci 57:1536–1544

486 Hastie T, Tibshirani R (1990) Generalized additive models. Chapman & Hall, London

487 Hazen EL, Jorgensen S, Rykaczewski RR, Bograd SJ, Foley DG, Jonsen ID, Shaffer SA., Dunne
488 JP, Costa DP, Crowder LB, Block BA (2013) Predicted habitat shifts of Pacific top
489 predators in a changing climate. Nat Clim Chang 3:234–238

490 Hazen EL, Palacios DM, Forney KA, Howell EA, Becker E, Hoover AL, Irvine L, DeAngelis M,
491 Bograd SJ, Mate BR, Bailey H (2017) WhaleWatch: a dynamic management tool for
492 predicting blue whale density in the California Current. J Appl Ecol 54:1415–1428

493 Hoegh-Guldberg O, Mumby PJ, Hooten AJ, Steneck RS, Greenfield P, Gomez E, Harvell CD,
494 Sale PF, Edwards AJ, Caldeira K, Knowlton N, Eakin CM, Iglesias-Prieto R, Muthiga N,
495 Bradbury RH, Dubi A, Hatziolos ME (2007) Coral reefs under rapid climate change and
496 ocean acidification. Science 318:1737–1742

497 Hurrell JW, Kushnir Y, Ottersen G (2003) An overview of the North Atlantic Oscillation. Clim
498 Significance Environ Impact Geophys Monogr 134:1–35

499 Ingram SN, Walshe L, Johnston D, Rogan E (2007) Habitat partitioning and the influence of
500 benthic topography and oceanography on the distribution of fin and minke whales in
501 the Bay of Fundy, Canada. J Mar Biol Assoc UK 87:149–156

502 Isojunno S, Matthiopoulos J, Evans PGH (2012) Harbour porpoise habitat preferences:
503 Robust spatio-temporal inferences from opportunistic data. Mar Ecol Prog Ser 448:155–
504 170

505 Johnston DW, Bowers MT, Friedlaender AS, Lavigne DM (2012) The effects of climate change
506 on harp seals (*Pagophilus Groenlandicus*). PLoS ONE 7:e29158

507 Kim YJ, Gu C (2004) Smoothing spline Gaussian regression: More scalable computation via
508 efficient approximation. J R Stat Soc Ser B Stat Methodol 66:337–356

509 Laidre K, Heide-Jørgensen M, Heagerty P, Cossio A, Bergström B, Simon M (2010) Spatial
510 associations between large baleen whales and their prey in West Greenland. Mar Ecol
511 Prog Ser 402:269–284

512 Laist DW, Knowlton AR, Pendleton D (2014) Effectiveness of mandatory vessel speed limits
513 for protecting North Atlantic right whales. Endanger Species Res 23:133–147

514 Larouche P, Galbraith PS (2016) Canadian Coastal Seas and Great Lakes Sea Surface
515 Temperature Climatology and Recent Trends. Can J Remote Sens 42:243–258

516 Marra G, Wood SN (2011) Practical variable selection for generalized additive models.
517 Comput Stat Data Anal 55:2372–2387

518 McQuinn IH, Plourde S, Pierre JF St., Dion M (2015) Spatial and temporal variations in the
519 abundance, distribution, and aggregation of krill (*Thysanoessa raschii* and
520 *Meganyctiphanes norvegica*) in the lower estuary and Gulf of St. Lawrence. Prog
521 Oceanogr 131:159–176

522 MPO (2003) Capelan de l'estuaire et du golfe du Saint-Laurent (4RST) en 2002. MPO - Sci
523 Rapp sur l'état des Stock 2003/009:1–14

524 Panigada S, Notarbartolo di Sciara G, Zanardelli Panigada M, Airoldi S, Borsani JF, Jahoda M
525 (2005) Fin whales (*Balaenoptera physalus*) summering in the Ligurian Sea: distribution,
526 encounter rate, mean group size and relation to physiographic variables. J Cetacean Res
527 Manag 7:137–145

528 Pennino MG, Mérigot B, Fonseca VP, Monni V, Rotta A (2016) Habitat modeling for cetacean
529 management: Spatial distribution in the southern Pelagos Sanctuary (Mediterranean
530 Sea). Deep Res Part II Top Stud Oceanogr 141:203–211

Phillips SJ, Dudík M, Elith J, Graham CH, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: Implications for background and pseudo-absence data. Ecol Appl 19:181–197

Piatt J, Methven D, Burger A, McLagan R, Mercer V, Creelman E (1989) Baleen whales and their prey in a coastal environment. Can J Zool 67:1523–1530

Pizzolato L, Howell SEL, Dawson J, Laliberté F, Copland L (2016) The influence of declining sea ice on shipping activity in the Canadian Arctic. Geophys Res Lett 43:12,146-12,154

Plourde S, Galbraith PS, Lesage V, Grégoire F, Bourdages H, Gosselin J-F, McQuinn I, Scarratt M (2014) Ecosystem perspective on changes and anomalies in the Gulf of St. Lawrence: a context in support of the management of the St. Lawrence beluga whale population. DFO Can Sci Advis Sec Res Doc 2013/129:v + 29p

Plourde S, Grégoire F, Lehoux C, Galbraith PS, Castonguay M, Ringuette M (2015) Effect of environmental variability on body condition and recruitment success of Atlantic Mackerel (*Scomber scombrus* L.) in the Gulf of St. Lawrence. Fish Oceanogr 24:347–363

Plourde S, Lehoux C, McQuinn IH, Lesage V (2016) Describing krill distribution in the western North Atlantic using statistical habitat models. Fish Ocean Canada Can Sci Advis Secr 111:v + 34 p

QGIS Development Team (2016). QGIS Geographic Information System. Open Source Geospatial Foundation Project.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Ramp C, Delarue J, Palsbøll PJ, Sears R, Hammond PS (2015) Adapting to a warmer ocean - Seasonal shift of baleen whale movements over three decades. PLoS ONE 10: e0121374

Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD, Good C, Barlow J, Kaschner K, Baumgartner MF, Forney KA, Ballance LT, Fauchald P, Halpin P, Hamazaki T, Pershing AJ, Qian SS, Read A, Reilly S, Torres L, Werner F (2006) Techniques for cetacean – habitat modeling. Mar Ecol Prog Ser 310:271–295

Redfern J V, McKenna MF, Moore TJ, Calambokidis J, Deangelis ML, Becker EA, Barlow J, Forney KA, Fiedler PC, Chivers SJ (2013) Assessing the Risk of Ships Striking Large Whales in Marine Spatial Planning. Conserv Biol 27:292–302

Redfern JV, Moore TJ, Fiedler PC, de Vos A, Brownell RL, Forney KA, Becker EA, Ballance LT (2017) Predicting cetacean distributions in data-poor marine ecosystems. Divers Distrib 23:394–408

Ressler PH, Dalpadado P, Macaulay GJ, Handegard N (2015) Acoustic surveys of euphausiids and models of baleen whale distribution in the Barents Sea. Mar Ecol Prog Ser 527:13–

29

Sahade R, Lagger C, Torre L, Momo F, Monien P, Schloss I, Barnes DKA, Servetto N, Tarantelli S, Tatián M, Zamboni N, Abele D (2015) Climate change and glacier retreat drive shifts in an Antarctic benthic ecosystem. Sci Adv 1:e1500050

Savenkoff C, Castonguay M, Chabot D, Hammill MO, Bourdages H, Morissette L (2007) Changes in the northern Gulf of St. Lawrence ecosystem estimated by inverse modelling: Evidence of a fishery-induced regime shift? Estuar Coast Shelf Sci 73:711–724

Schleimer A, Ramp C, Delarue J, Carpentier A, Bérubé M ,Palsbøll PJ, Sears R, Hammond PS (2019) Decline in abundance and apparent survival rates of fin whales (*Balaenoptera physalus*) in the northern Gulf of St. Lawrence. Ecol Evol 9: 4231–4244

Sigurjónsson J (1988) Operational factors of the Icelandic large whale fishery. Rep Int Whal Comm 38:327–333

Sigurjónsson J, Gunnlaugsson T (2006) Revised catch series and cpue for fin whales taken from the early modern whaling land stations in Iceland. Int Whal Comm SC/58/PFI4:1–22

Simard Y, Lavoie D, Saucier FJ (2002) Channel head dynamics: capelin (*Mallotus villosus*) aggregation in the tidally driven upwelling system of the Saguenay - St. Lawrence Marine Park's whale feeding ground. Can J Fish Aquat Sci 59:197–210

Thibodeau B, de Vernal A, Hillaire-Marcel C, Mucci A (2010) Twentieth century warming in deep waters of the Gulf of St. Lawrence: A unique feature of the last millennium. Geophys Res Lett 37:1–5

Thomson JA, Burkholder DA, Heithaus MR, Fourqurean JW, Fraser MW, Statton J, Kendrick GA (2015) Extreme temperatures, foundation species, and abrupt ecosystem change: an example from an iconic seagrass ecosystem. Glob Chang Biol 21:1463–1474

Thuiller W, Lafourcade B, Engler R, Araújo MB (2009) BIOMOD - A platform for ensemble forecasting of species distributions. Ecography (Cop) 32:369–373

Torres LG, Read AJ, Halpin P (2008) Fine-scale habitat modeling of a top marine predator: do prey data improve predictive capacity? Ecol Appl 18:1702–17

Truchon M-H, Measures L, Brêthes J-C, Albert É, Michaud R (2018) Influence of anthropogenic activities on marine mammal strandings in the estuary and northwestern Gulf of St. Lawrence, Quebec, Canada, 1994-2008. J Cetacean Res Manag 18:11–21

Tylianakis JM, Didham RK, Bascompte J, Wardle DA (2008) Global change and species interactions in terrestrial ecosystems. Ecol Lett 11:1351–1363

van der Hoop JM, Moore MJ, Barco SG, Cole TVN, Daoust PY, Henry AG, Mcalpine DF, Mclellan WA, Wimmer T, Solow AR (2013) Assessment of Management to Mitigate Anthropogenic Effects on Large Whales. Conserv Biol 27:121–133

Van Waerebeek K, Leaper R (2008) Second Report of the IWC Vessel Strike Data Standardisation Working Group. Int Whal Comm Sci Comm Doc SC/60/BC5, Santiago, Chile.

Venables, WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Víkingsson GA, Pike DG, Valdimarsson H, Schleimer A, Gunnlaugsson T, Silva T, Elvarsson BÃ, Mikkelsen B, Øien N, Desportes G, Bogason V, Hammond PS (2015) Distribution, abundance, and feeding ecology of baleen whales in Icelandic waters: have recent environmental changes had an effect? Front Ecol Evol 3:1–18

Virgili A, Racine M, Authier M, Monestiez P, Ridoux V (2017) Comparison of habitat models for scarcely detected species. Ecol Modell 346:88–98

Visser F, Hartman K, Pierce G, Valavanis V, Huisman J (2011) Timing of migratory baleen whales at the Azores in relation to the North Atlantic spring bloom. Mar Ecol Prog Ser 440:267–279

Warton DI (2005) Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. Environmetrics 16:275–289

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

Wood SN (2006) Generalized additive models: an introduction with R. Chapman & Hall

Wood SN, Pya N, Säfken B (2016) Smoothing Parameter and Model Selection for General Smooth Models. J Am Stat Assoc 111:1548–1563

Wood SN (2017) Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC.

Woodley TH, Gaskin DE (1996) Environmental characteristics of North Atlantic right and fin whale habitat in the lower Bay of Fundy, Canada. Can J Zool 74:75–84

Zerbini AN, Friday NA, Palacios DM, Waite JM, Ressler PH, Rone BK, Moore SE, Clapham PJ (2016) Baleen whale abundance and distribution in relation to environmental variables and prey density in the Eastern Bering Sea. Deep Res Part II Top Stud Oceanogr 134:312–330

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed Effects Models and Extensions in Ecology with R. Stat Biol Heal:579pp

634 **Tables**

635 Table 1. Candidate explanatory variables for models to predict fin whale relative abundance
636 in the Jacques Cartier Passage.

Variable	Description	Resolution
Static variables		
Depth[a]	Water depth (metres) at grid cell centre	1 arc-min
Slope[a]	Slope of seafloor (degrees) calculated in QGIS from bathymetry	1 arc-min
Aspect[a]	Slope orientation (0° to 360°) calculated in QGIS from bathymetry	1 arc-min
DistCoast[a]	Distance to nearest coastline (metres)	1 arc-min
Dynamic proxy variables		
SST[b]	Monthly average sea surface temperature (°C)	4 km
Spring SST[b]	Seasonal composite of average spring (21 March - 21 June) sea surface temperature (°C)	4 km
Chl *a*[c]	Log transformed ($log(X+1)$) monthly average chlorophyll *a* concentration (mg m^{-3})	4 km
Spring Chl *a*[c]	Log transformed ($log(X+1)$) seasonal composite of average spring (21 March - 21 June) chlorophyll *a* concentration (mg m^{-3})	4 km
NAOI[d]	Monthly Hurrell North Atlantic Oscillation (NAO) index	
WinterNAOI[e]	Winter NAOindex in winter (December to March) preceding sampling season	
WinterNAOIlag1[e]	Winter NAOI with one-year lag	
WinterNAOIlag2[e]	Winter NAOI with two-year lag	
Year	Survey year (2007-2013)	
Month	Survey month (June-October)	
Dynamic prey variables		
Krill biomass[f]	Log transformed ($log(X+1)$) modelled monthly krill biomass (g m^{-2})	5 km

637 Sources:
638 [a] ETOPO1 1 Arc-Minute Global Relief Model
639 [b] Aqua-MODIS Level-3 sea surface temperature (4µ nighttime). DOI
640 10.5067/AQUA/MODIS/L3M/SST/2014
641 [c] Aqua-MODIS Level-3 chlorophyll concentration (OCx algorithm). DOI
642 10.5067/AQUA/MODIS/L3M/CHL/2018
643 [d] Climate Analysis Section, NCAR, Boulder, USA, Hurrell (2003) accessed 25 May 2017[e] National
644 Center for Atmospheric Research Staff (Eds). The Climate Data Guide: Hurrell North Atlantic
645 Oscillation (NAO) Index (PC-based).
646 [f] Plourde et al. (2016)

647 Table 2. Summary of annual survey effort (in hours) and the number of fin whale sightings
648 with information on the median (and maximum) group sizes of the fin whale encounters.

Year	Uncorrected effort	Handling time	Corrected effort	Fin whale sightings	Sightings per corrected hour	Median group size
2007	206.1	37.3	168.8	527	3.12	2 (9)
2008	280.2	81.5	198.7	674	3.39	2 (14)
2009	325.6	112.1	213.5	488	2.29	2 (8)
2010	252.6	70.5	182.1	508	2.79	2 (10)
2011	170.1	59.1	111.0	177	1.60	1 (6)
2012	297.7	89.1	208.6	296	1.42	1 (8)
2013	346.1	60.4	285.7	316	1.11	1 (14)
Total	**1878.4**	**510.0**	**1368.4**	**2986**		

649 Table 3. Model selection of proxy- and prey-fin whale models with and without interaction
650 terms.

Variables	Θ	REML	AIC	%Dev	r^2	ASPE
1. Proxy Model						
1.1 Penalised Model*	0.17	1068.7	2963.4	20.6	0.38	29.59
1.2 Penalised Model +ti(Depth,month)	0.18	1067.5	2953.4	23.1	0.36	28.83
1.3 Penalised Model +ti(Depth,year)	0.17	1068.5	2963.1	21.1	0.39	27.72
1.4 Penalised Model +ti(Aspect,month)	**0.18**	**1065.0**	**2944.4**	**23.7**	**0.43**	**25.72**
1.5 Penalised Model +ti(Aspect:year)	0.17	1068.4	2961.9	21.5	0.37	28.50
2. Prey Model						
2.1 s(krill) +s(month) +s(year)	0.12	1143.6	3185.2	7.6	0.21	34.29
2.2 s(krill) +s(month) +s(year) +ti(krill,month)	**0.13**	**1138.3**	**3161.8**	**11.8**	**0.23**	**33.07**
2.3 s(krill) +s(month) +s(year) +ti(krill,year)	0.12	1143.2	3184.0	8.1	0.21	34.03

651 *Full penalised model includes all variables described in Table 1, except for krill biomass:
652 s(Depth)+ s(Slope)+ s(Aspect)+ s(DistCoast)+ s(SST)+ s(SpringSST)+ s(Chla)+ s(SpringChla)+
653 s(NAOI)+ s(WinterNAOI)+ s(WinterNAOIlag1)+ s(WinterNAOIlag2)+ s(year)+ s(month) with
654 automated variable selection using shrinkage smoothers. Model selection was based on
655 Akaike Information Criterion (AIC), percentage of deviance explained (%Dev), and mean
656 average squared prediction error (ASPE) from a five-fold cross-validation. Θ = theta
657 parameter from negative binomial *nb()* error distribution; REML = restricted maximum
658 likelihood. Selected models are shown in bold.

Figures

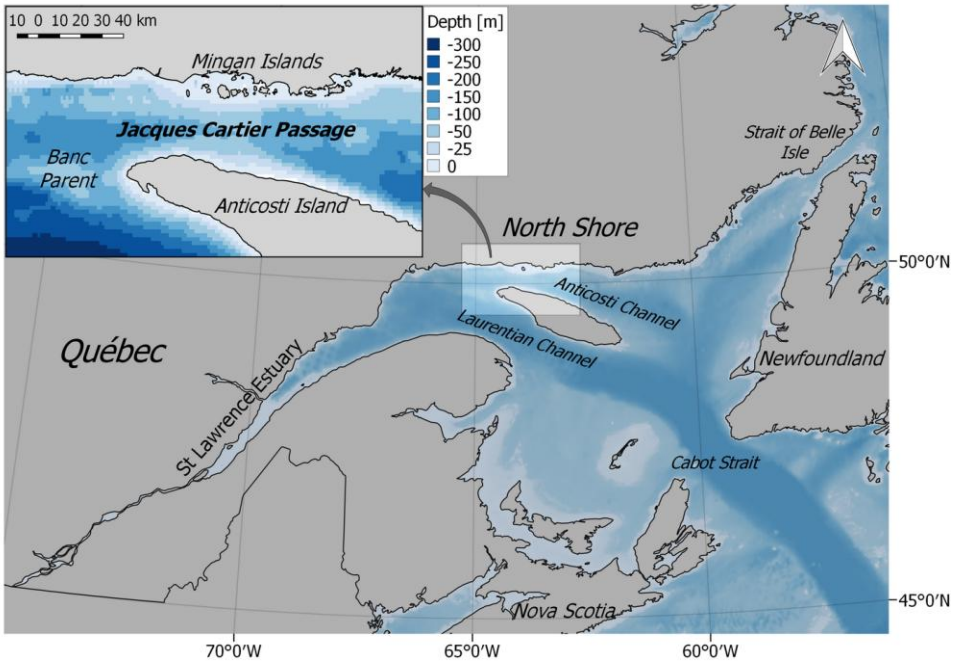
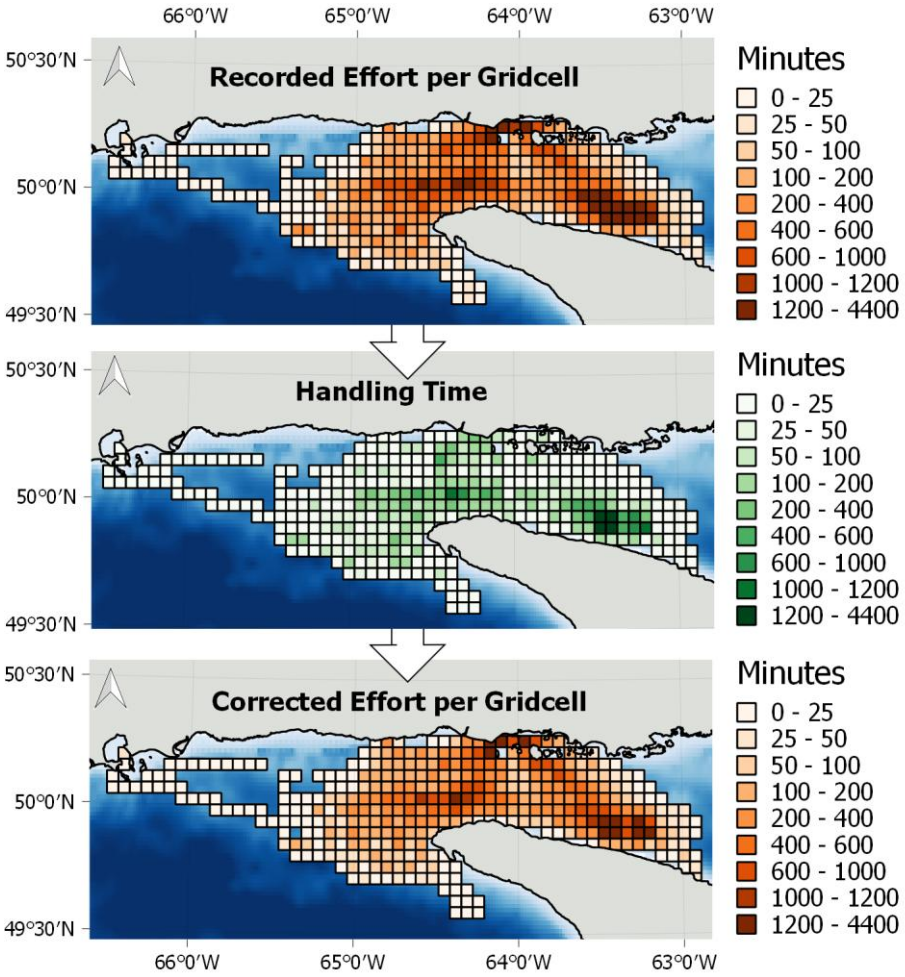


Fig. 1. Schematic representation of the Gulf of St. Lawrence, with detailed bathymetry of the study area in the Jacques Cartier Passage.

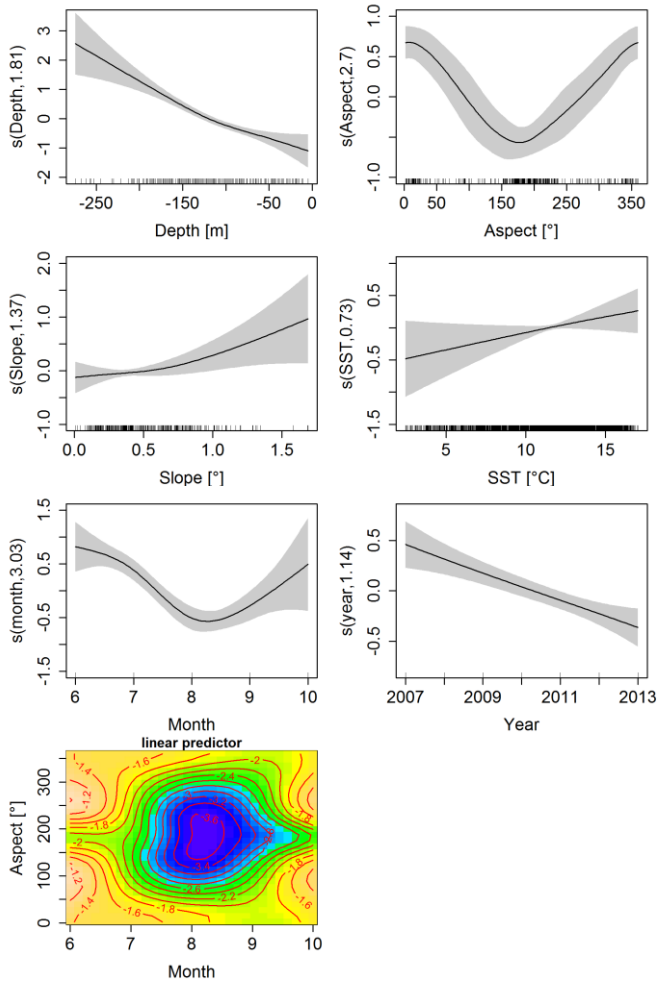


663

664 Fig. 2. Distribution of total survey effort in the Jacques Cartier Passage in minutes per 25 km^2

665 grid cell over the seven survey years (June to October, 2007 to 2013), followed by the

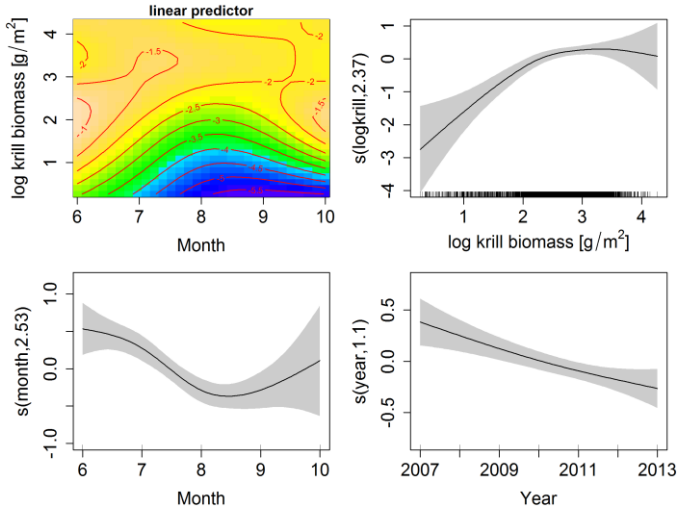
666 amount of handling time and the derived corrected effort per grid cell.



667

668 Fig. 3. Smooth functions fitted in the final proxy-fin whale model. Positive values of the
669 smoothed function indicate a positive effect on the response variable. Tick marks on the
670 horizontal axis show the distribution of observations, while the smoother terms with
671 estimated degrees of freedom (edf) are shown on the vertical axes. Shaded areas represent
672 95% confidence intervals. Last plot shows the 2-D interaction between aspect and month
673 (5.06 edf, cold colours represent negative effect).

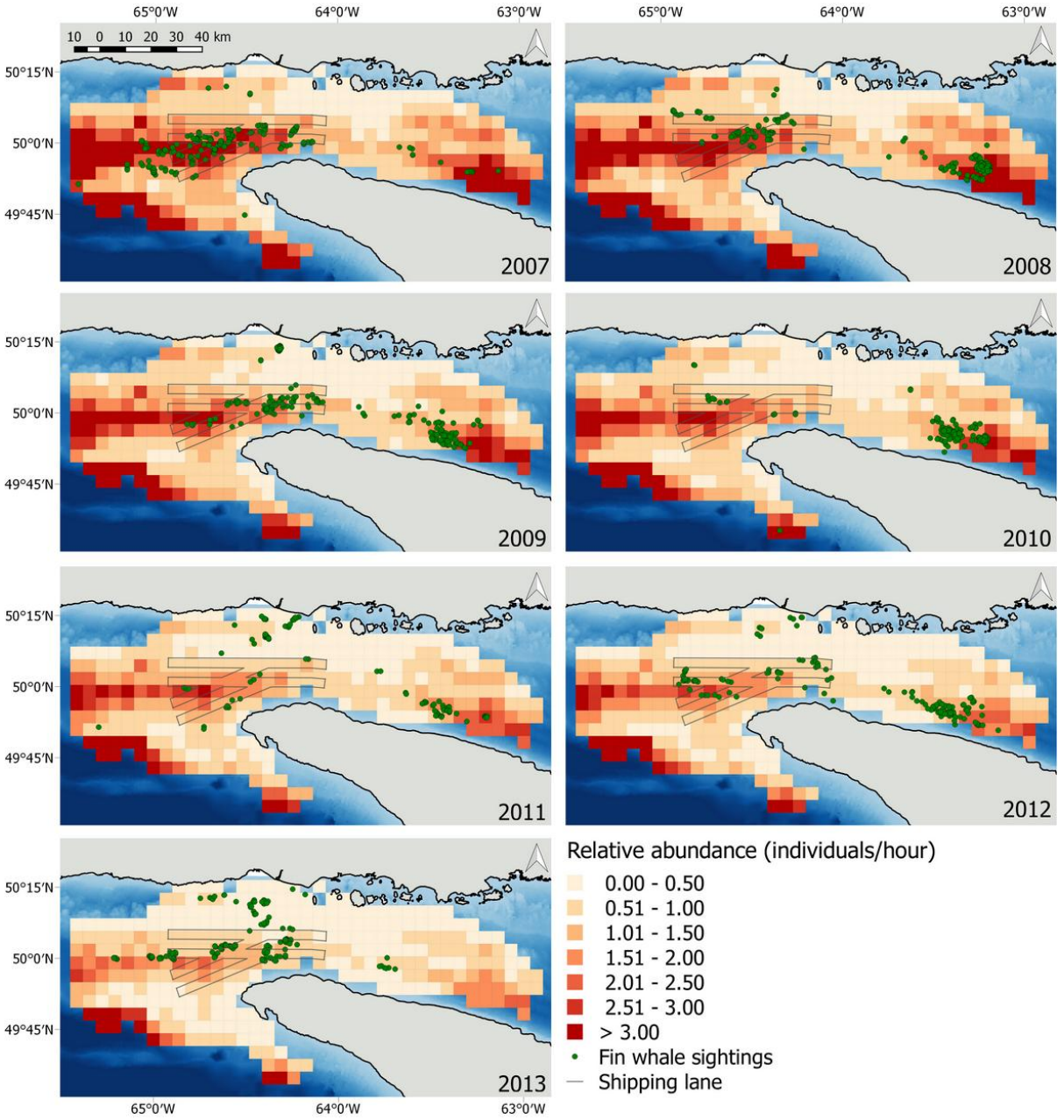
674



675

676

677 Fig. 4. Smooth functions fitted in the final prey-fin whale model. Positive values of the
678 smoothed function indicate a positive effect on the response variable. Tick marks on the
679 horizontal axis show the distribution of observations, while the smoother terms with
680 estimated degrees of freedom (edf) are shown on the vertical axes. Shaded areas represent
681 95% confidence intervals. First plot shows the 2-D interaction between krill biomass and
682 month (4.63 edf).



683

684 Fig. 5. Predictive maps of relative annual fin whale abundance (individuals per hour effort)
685 from the proxy fin whale model. Each map shows the average annual relative abundance of
686 fin whales in each grid cell. The dots show the reported sightings of fin whale groups made
687 during the surveys in that year.

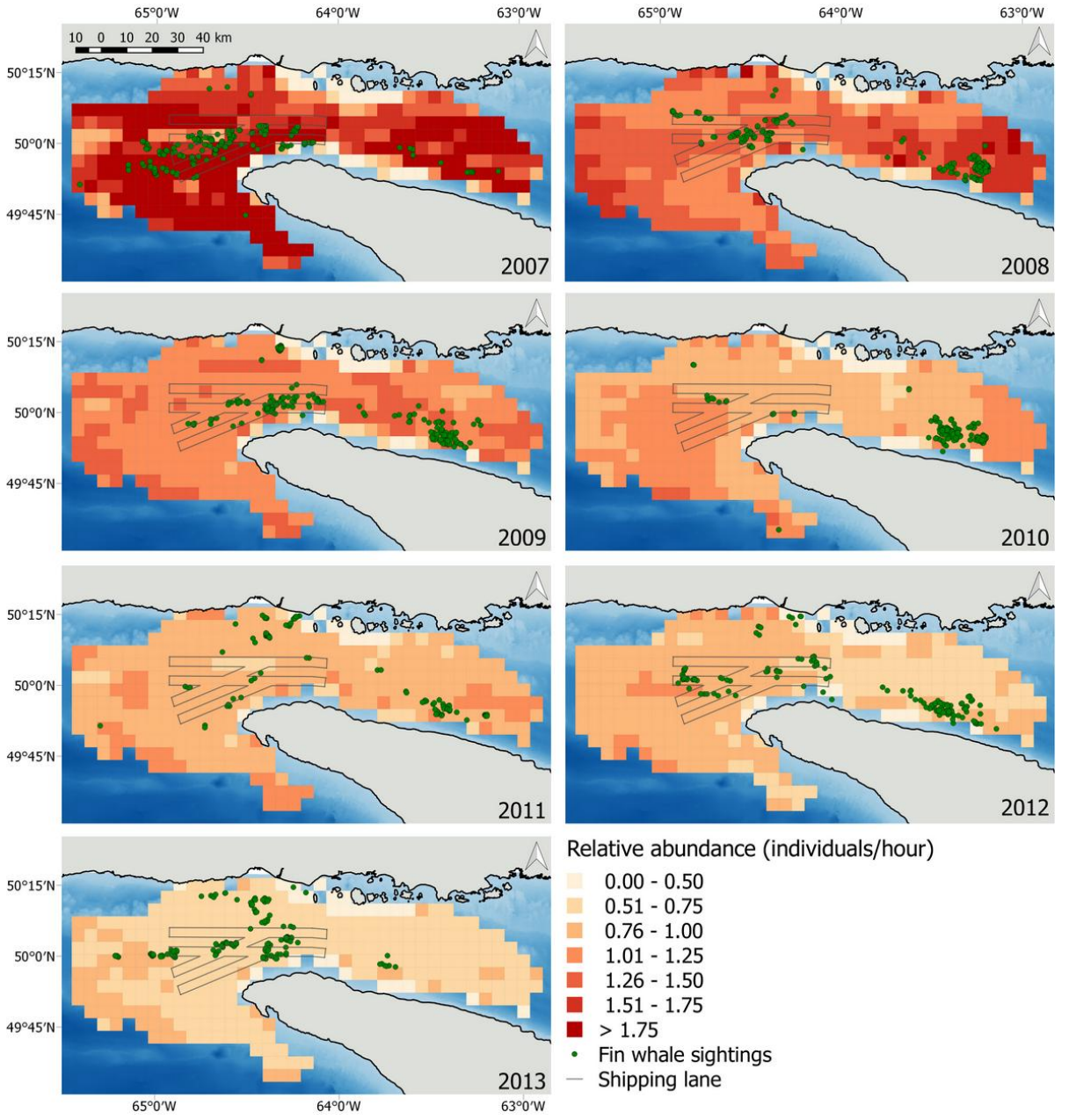


Fig. 6. Predictive maps of relative annual fin whale abundance (individuals per hour effort) from the prey fin whale model. Each map shows the average annual relative abundance of fin whales in each grid cell. The dots show the reported sightings of fin whale groups made during the surveys in that year.

SUPPORTING INFORMATION

Spatio-temporal patterns in fin whale (*Balaenoptera physalus*) habitat use in the northern Gulf of St. Lawrence

Anna Schleimer[1,2,3,*], Christian Ramp[1,2], Stéphane Plourde[4], Caroline Lehoux[4], Richard Sears[2], Philip S. Hammond[1]

[1]Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, St Andrews, KY16 8LB UK

[2]Mingan Island Cetacean Study, St Lambert, Québec, J4P 1T3, Canada

[3]Marine Evolution and Conservation, Centre for Ecological and Evolutionary Studies, University of Groningen, Groningen, 9700 CC, The Netherlands

[4]Institut Maurice-Lamontagne, Fisheries and Oceans Canada, Mont-Joli, Québec, G5H 3Z4, Canada

*__Corresponding author__: Anna Schleimer, schleimer.anna@gmail.com

Supplement 1

Effort Calculation

The GPS unit had different sampling settings, including sampling locations at a set temporal interval and an automatic mode, during which positions were recorded at random time-interval when the boat was changing direction. However, the GPS units on the boats only recorded positions in latitude and longitude, without timestamp. Because time was identified as the most appropriate measure of effort for this study, the first step consisted in associating timestamps to the GPS locations for the calculation of effort. During the survey, precise timestamps and positions were recorded manually for events such as survey start, survey stop, sightings and changes in survey conditions. For each survey track, these recorded timestamps were associated to the nearest point of the survey track collected on the same day and boat (Fig. S1). Associations that were made with positions farther than 100m apart were discarded. Once these timestamps from the survey data were associated to the survey track, the sampling interval could be estimated by dividing the time difference between two timestamps by the number of positions recorded by the GPS unit. Depending on the setting, the GPS unit recorded positions at a 5, 10, or 30 second interval. On an unknown number of surveys the GPS unit ran on the default automatic mode. For those days it was impossible to reconstruct the timestamp and all associated sighting and effort data were excluded from future analysis. Once the sampling interval was determined, the timestamps were reconstructed for all remaining GPS positions. Data from 292 surveys were retained for the habitat modelling.

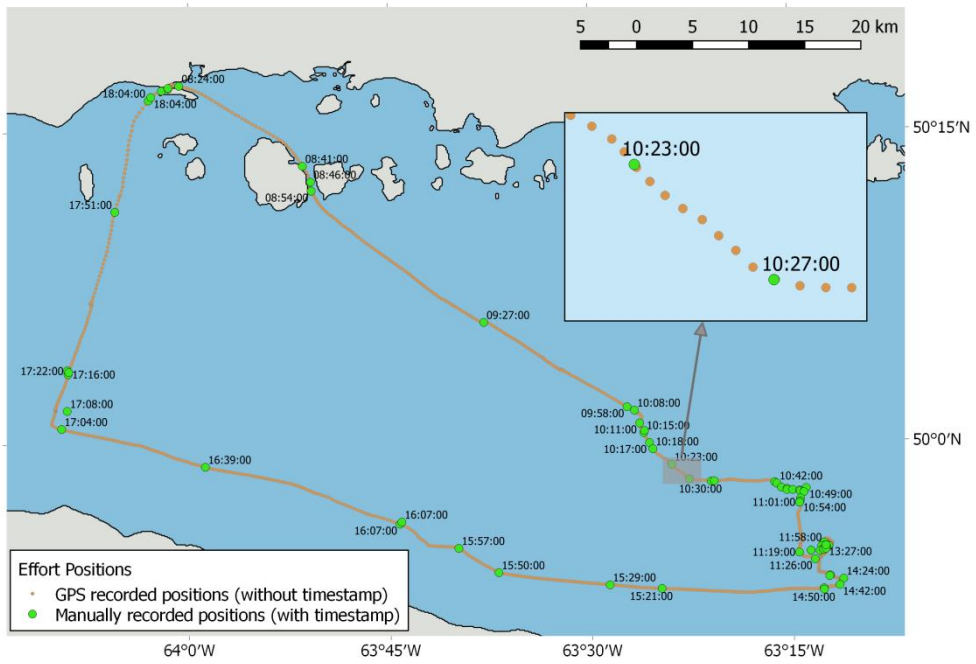


Fig. S1. Example of timestamp association procedure from survey on 5 Aug 2008. The time interval divided by the number of intermediate GPS positions provides a sampling interval of 30 sec (=4min/8), allowing the calculation of timestamps for all intermediate GPS location.

Supplement 2

Error Distribution Selection

The response variable was characterised by a high frequency of zeros (3207 grid cells without sightings compared to 312 grid cells with sightings), requiring careful choice of the error distribution. Quantile-quantile plots were used to compare the performance of three different error distributions, namely overdispersed Poisson (quasi-Poisson), negative binomial, and Tweedie error distributions for the fin whale occurrence model (Figs S2 & S3). All three error structure have been suggested to deal with overdispersed count data and differ mainly by their mean-variance relationship (Warton 2005, Ver Hoef & Boveng 2007, Miller et al. 2013). Quasi-Poisson and negative binomial share the same number of parameters, but the linear mean-variance function of the Quasi-Poisson distribution puts more weight on large counts while small counts are more heavily weighted in the negative binomial distribution due to its quadratic mean-variance function (Ver Hoef & Boveng 2007). In addition to the mean (μ) and dispersion (Ø) parameters, the Tweedie distribution has a third power (p) parameter, offering additional flexibility to model count data (Miller et al. 2013). The Tweedie mean-variance relationship is described as $var(Y) = Ø\mu^p$ (Miller et al. 2013). Setting p to 1 gives a quasi-Poisson distribution.

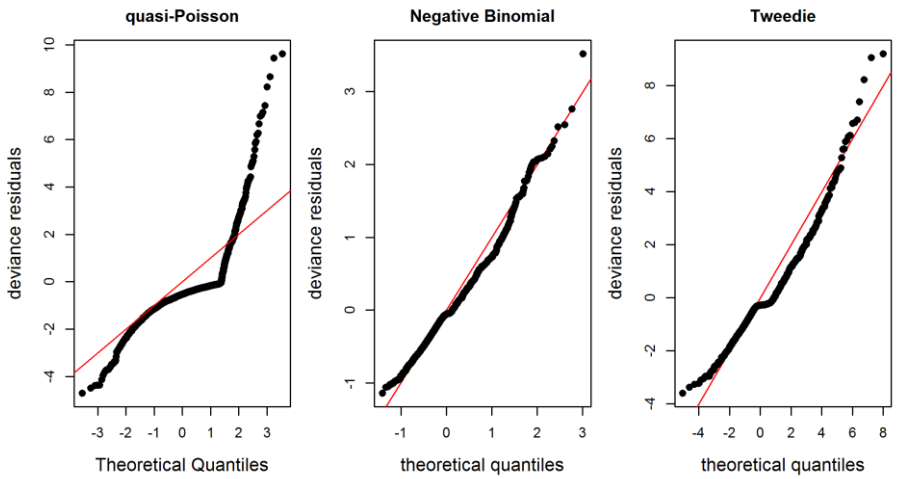


Fig. S2. Quantile-Quantile plots of proxy model with three different error distributions. The negative binomial error distribution provided the best fit.

Supplement 3

Model validation

Collinearity between candidate explanatory was evaluated using the pairs() function in the AED package. Based on the 0.6 cut-off value, there was no evidence for significant collinearity that required further investigation (Fig. S3).

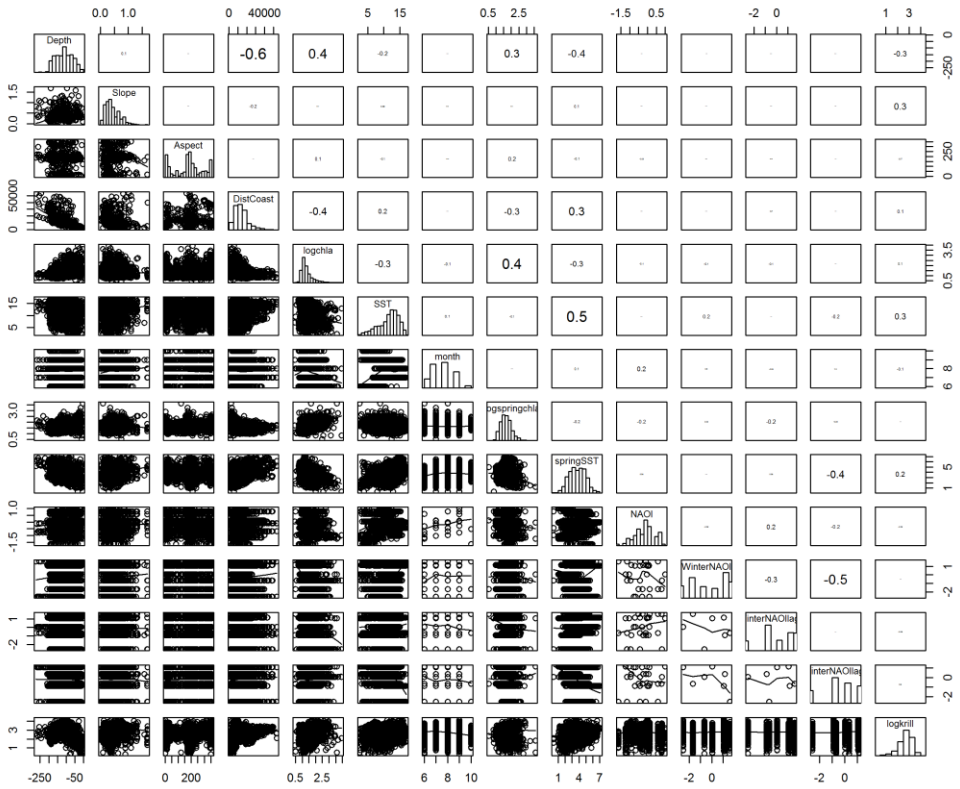
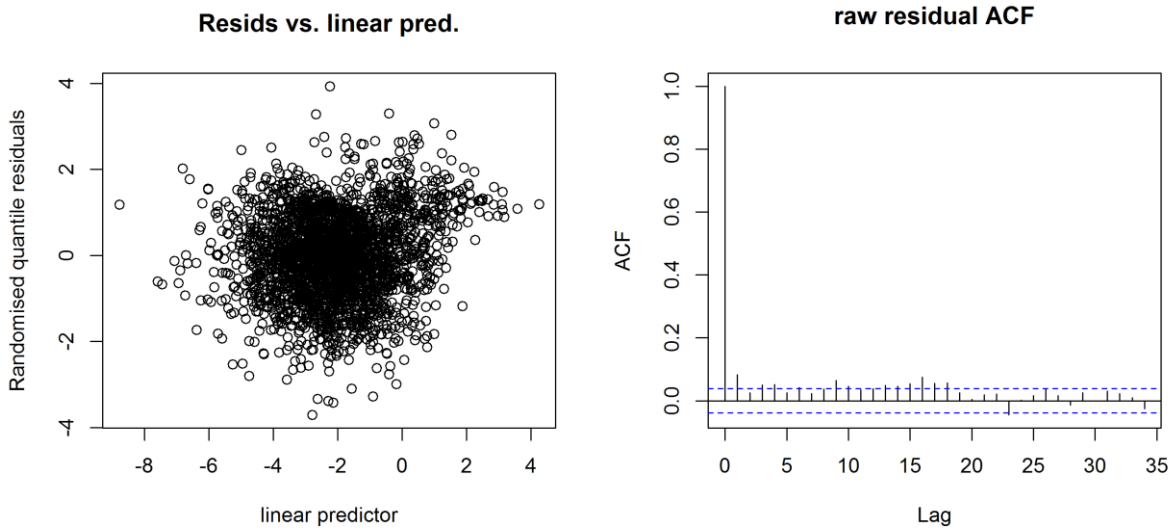


Fig. S3. Pairplot of candidate explanatory variables, with the upper panel showing estimated pair-wise correlation coefficients. R-code based on Zuur et al. (2009).

Residual plots were investigated to assess assumptions of variance homogeneity and independent errors. There was no indication for variance heterogeneity or autocorrelation for the proxy-fin whale model (Fig. S4) but some indication of autocorrelation in the prey-fin whale model (Fig. S5).



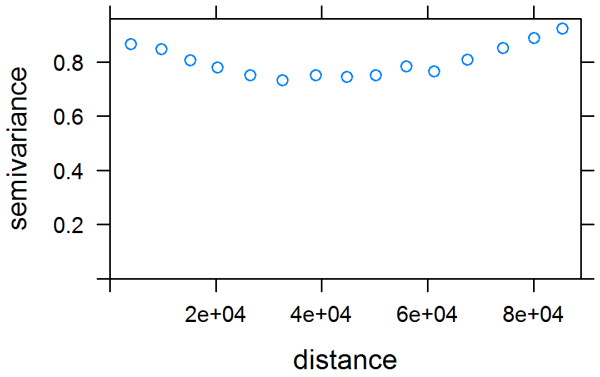
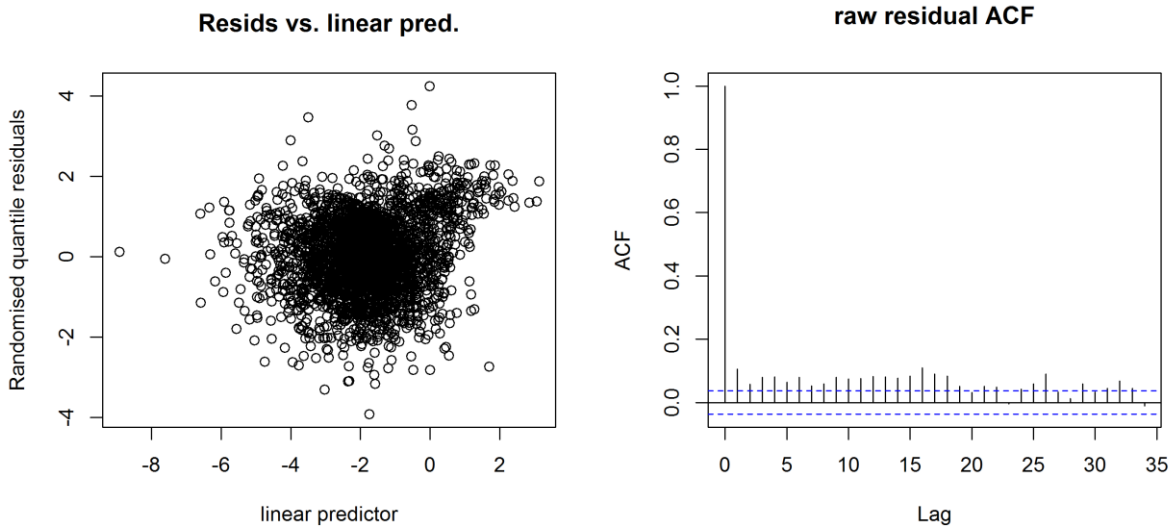
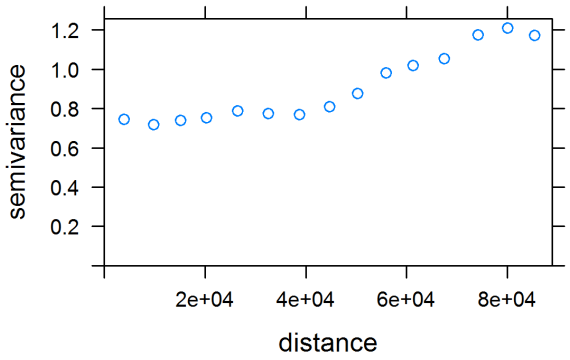


Fig. S4. Diagnostic residual plot for proxy-fin whale occurrence model. The horizontal band on the semi-variogram of residuals, with distance on the x-axis and semi-variance on the y-axis, indicates spatial independence (Zuur et al. 2009).



780



781

782 Fig. S5. Diagnostic residual plot for prey-fin whale occurrence model. The horizontal band on
783 the semi-variogram of residuals, with distance on the x-axis and semi-variance on the y-axis,
784 indicates spatial independence (Zuur et al. 2009).

785

786

787

788

789

790

791

792

793

Supplement 4

Uncertainty distribution of model predictions

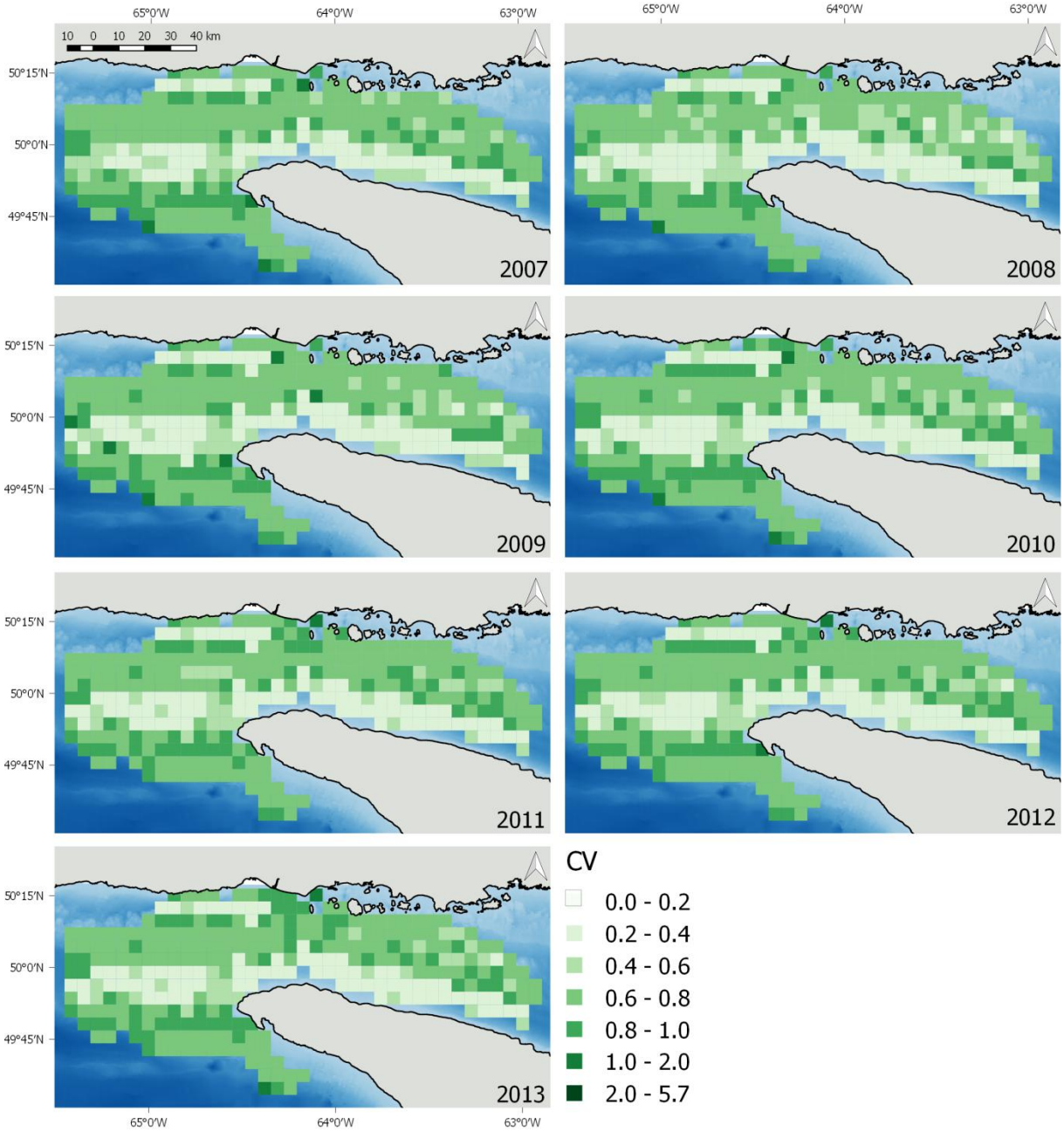


Fig. S6. Coefficients of variation (CV) of annual average predictions from the proxy-fin whale model.

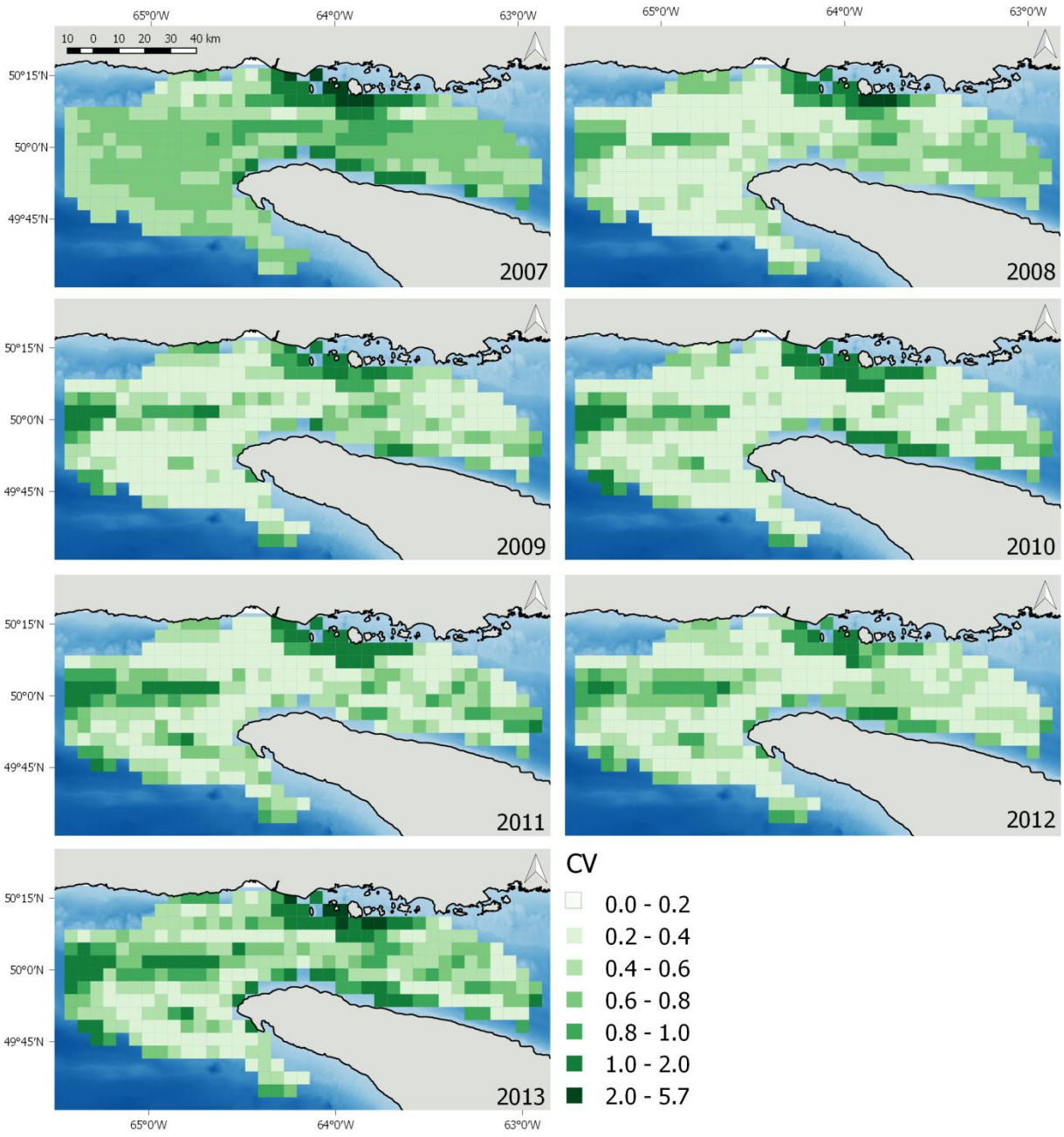


Fig. S7. Coefficients of variation (CV) of annual average predictions from the prey-fin whale model.

Supplement 5

Summary of annual trends

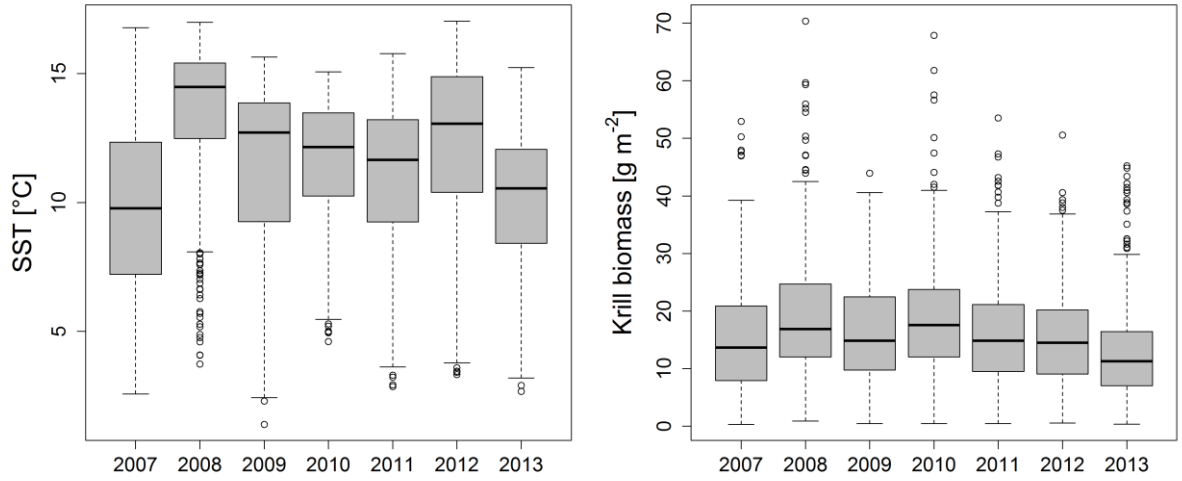


Fig. S8. Average annual sea surface temperatures (SST) and modelled krill biomass over all grid cells.

References

Miller DL, Burt ML, Rexstad EA, Thomas L (2013) Spatial models for distance sampling data: recent developments and future directions (O Gimenez, Ed.). Methods Ecol Evol 4:1001–1010

Ver Hoef J, Boveng P (2007) Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? Ecology 88:2766–2772

Warton DI (2005) Many zeros does not mean zero inflation: comparing the goodness-of-fit of parametric models to multivariate abundance data. Environmetrics 16:275–289

Zuur AF, Ieno EN, Walker N, Saveliev AA, Smith GM (2009) Mixed Effects Models and Extensions in Ecology with R. Stat Biol Heal:579pp