The following supplements accompany the article

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

Anna Schleimer*, Christian Ramp, Stéphane Plourde, Caroline Lehoux, Richard Sears, Philip S. Hammond

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

Marine Ecology Progress Series 623: 221–234 (2019)

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.

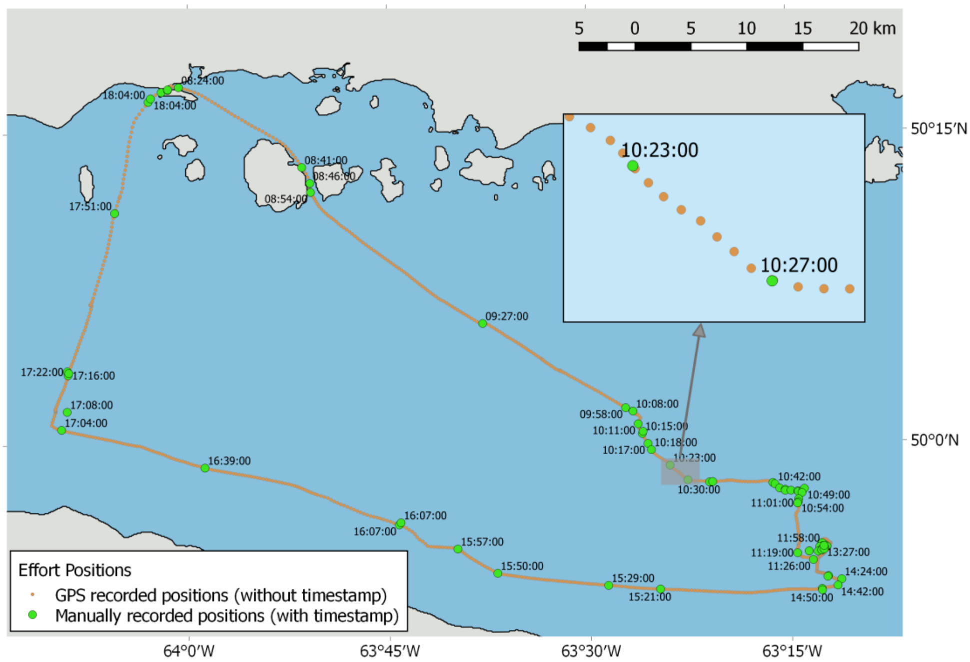

Fig. S1. Example of timestamp association procedure from survey on 5 Aug 2008. The time interval divided by the number of intermediate GPS positions provides a sampling interval of 30 sec (=4min/8), allowing the calculation of timestamps for all intermediate GPS location.

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 $var(Y) = \text{Ø}\mu^{p}$ (Miller et al. 2013). Setting p to 1 gives a quasi-Poisson distribution.

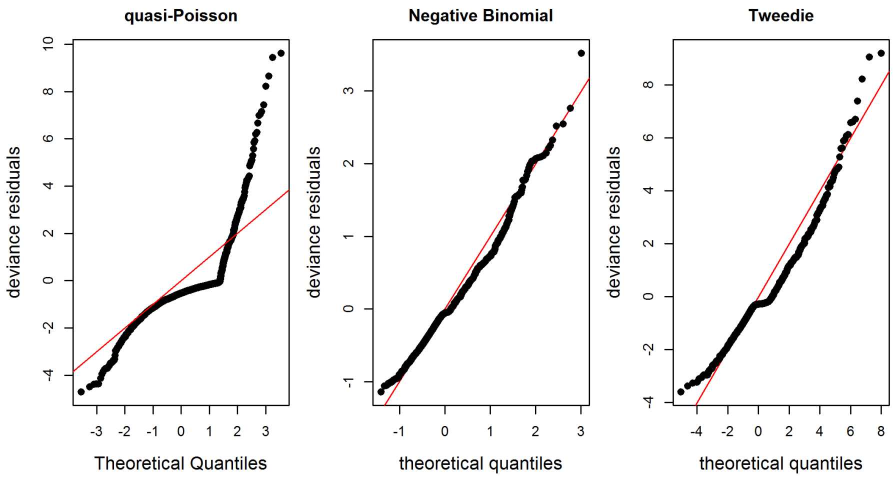

Fig. S2. Quantile-Quantile plots of proxy model with three different error distributions. The negative binomial error distribution provided the best fit.

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).

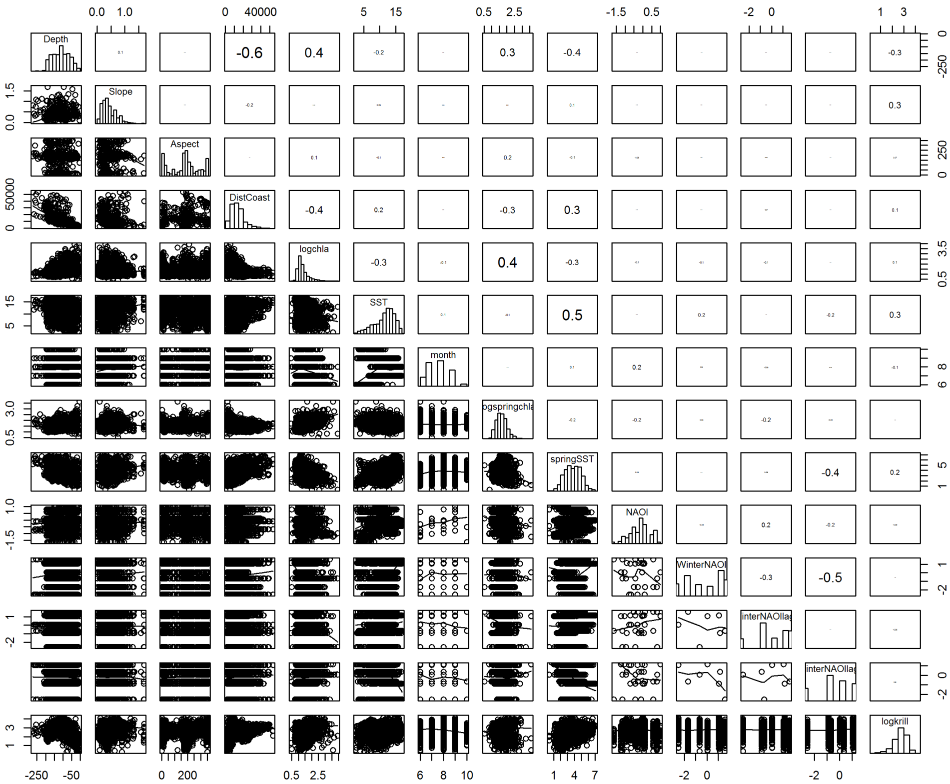

Fig. S3. Pairplot of candidate explanatory variables, with the upper panel showing estimated pair-wise correlation coefficients. R-code based on Zuur et al. (2009).

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).

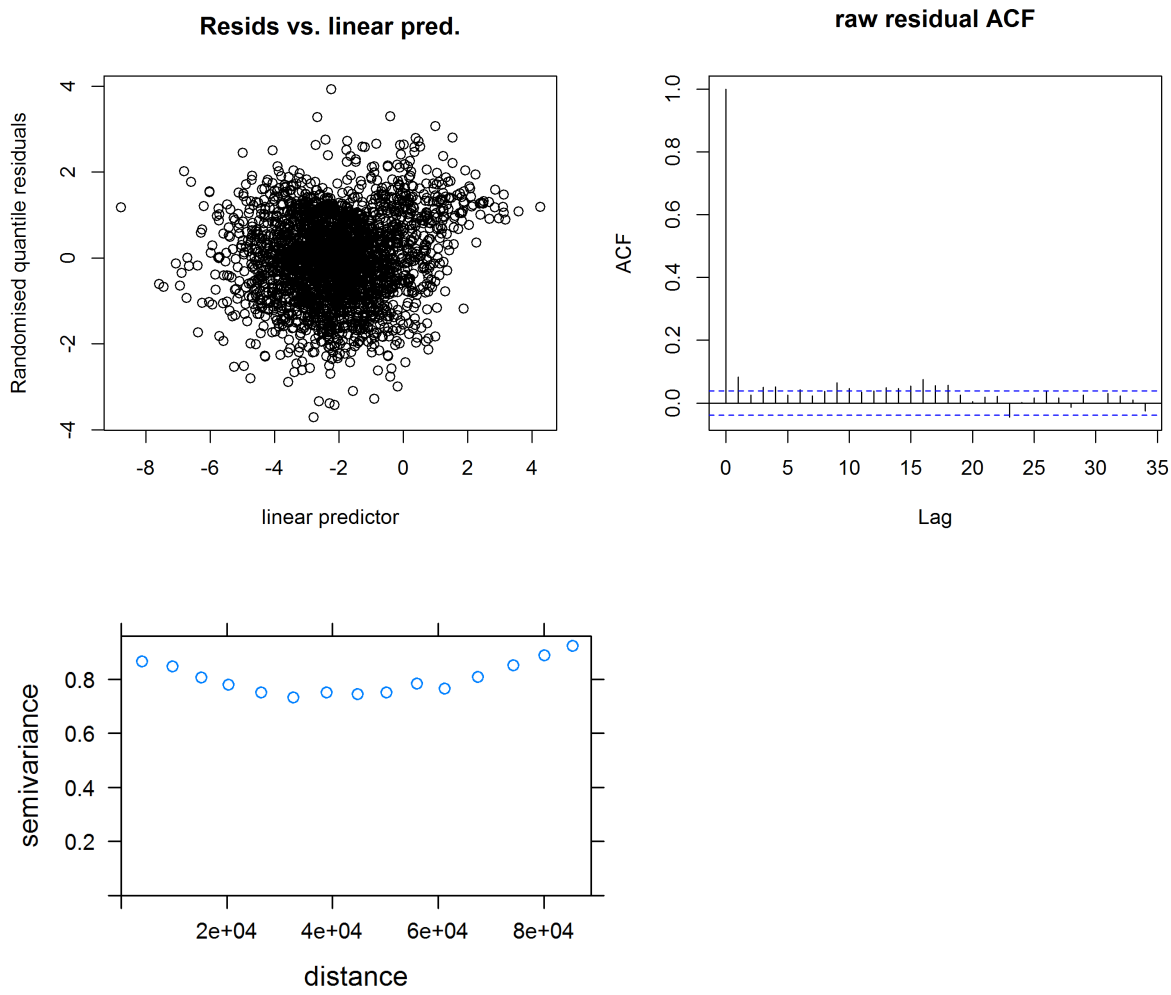

Fig. S4. Diagnostic residual plot for proxy-fin whale occurrence model. The horizontal band on the semi-variogram of residuals, with distance on the x-axis and semi-variance on the y-axis, indicates spatial independence (Zuur et al. 2009).

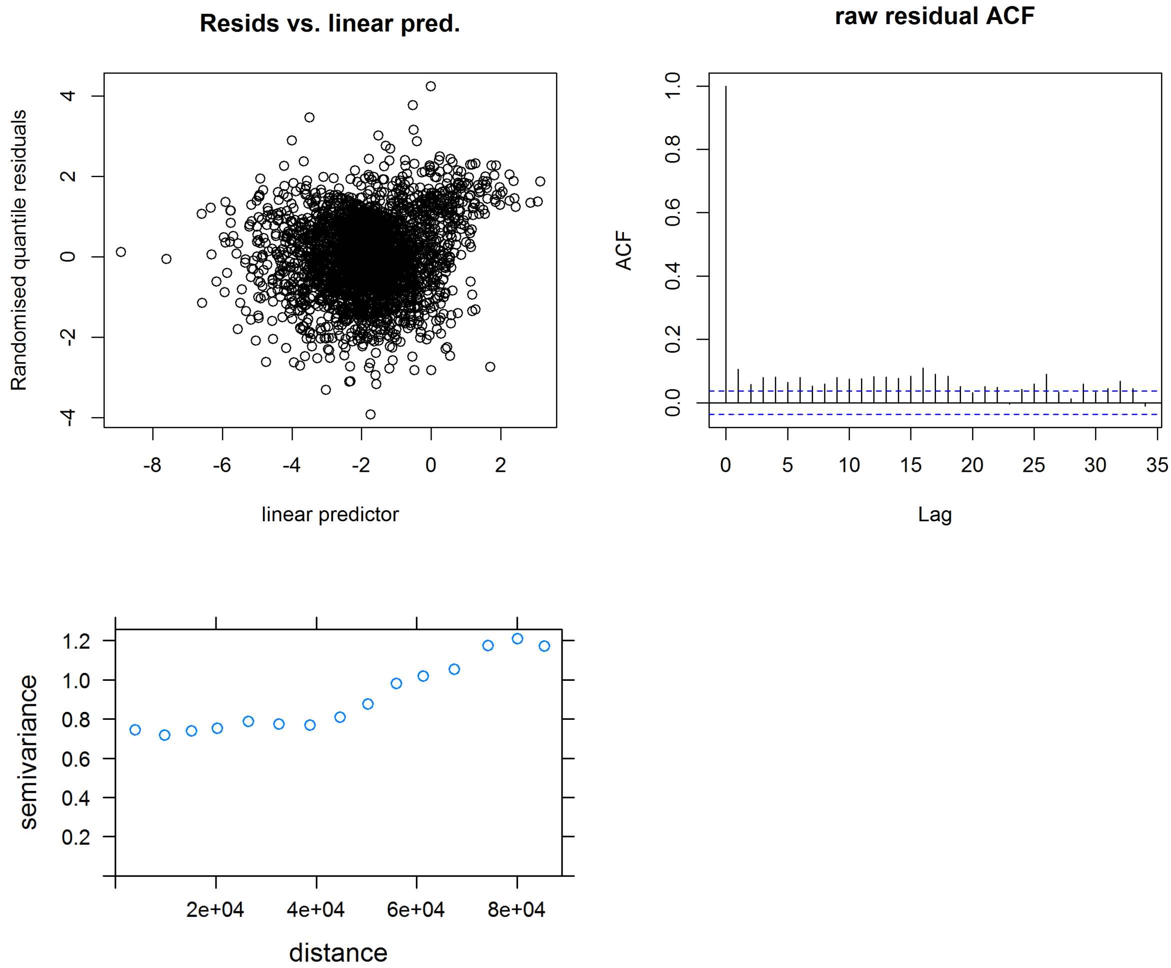

Fig. S5. Diagnostic residual plot for prey-fin whale occurrence model. The horizontal band on the semi-variogram of residuals, with distance on the x-axis and semi-variance on the y-axis, indicates spatial independence (Zuur et al. 2009).

Supplement 4

Uncertainty distribution of model predictions

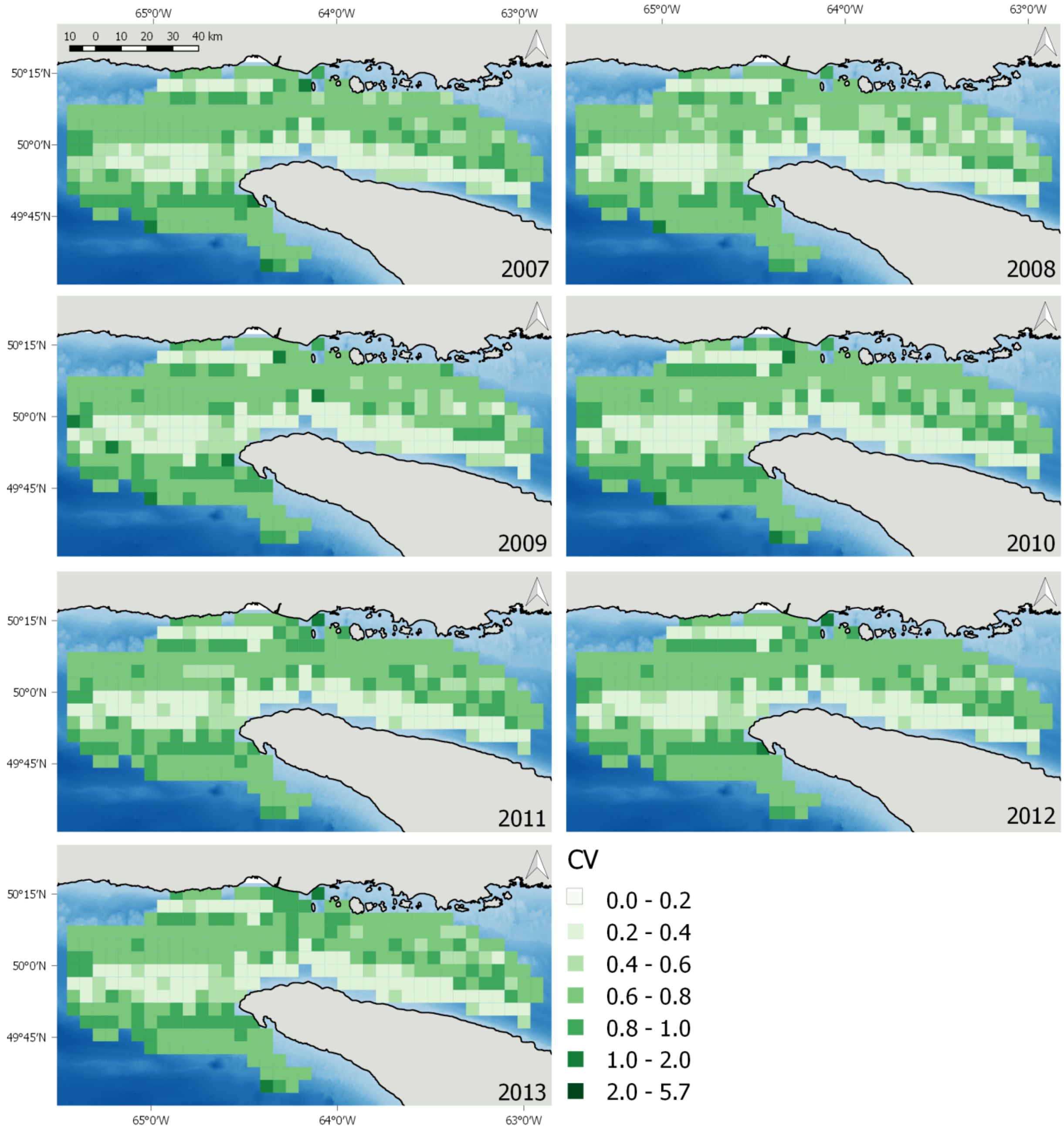

Fig. S6. Coefficients of variation (CV) of annual average predictions from the proxy-fin whale model.

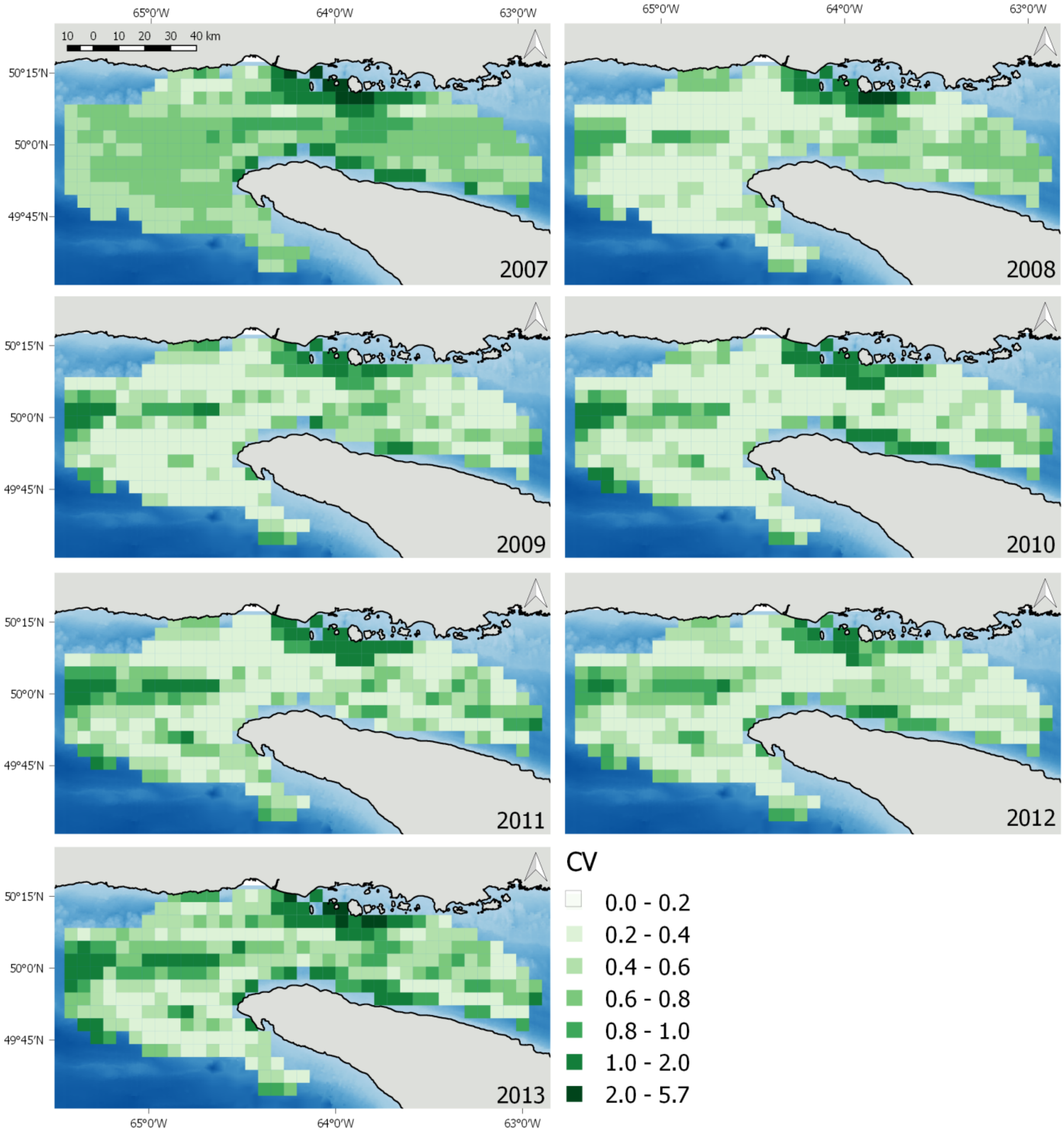

Fig. S7. Coefficients of variation (CV) of annual average predictions from the prey-fin whale model.

Supplement 5

Summary of annual trends

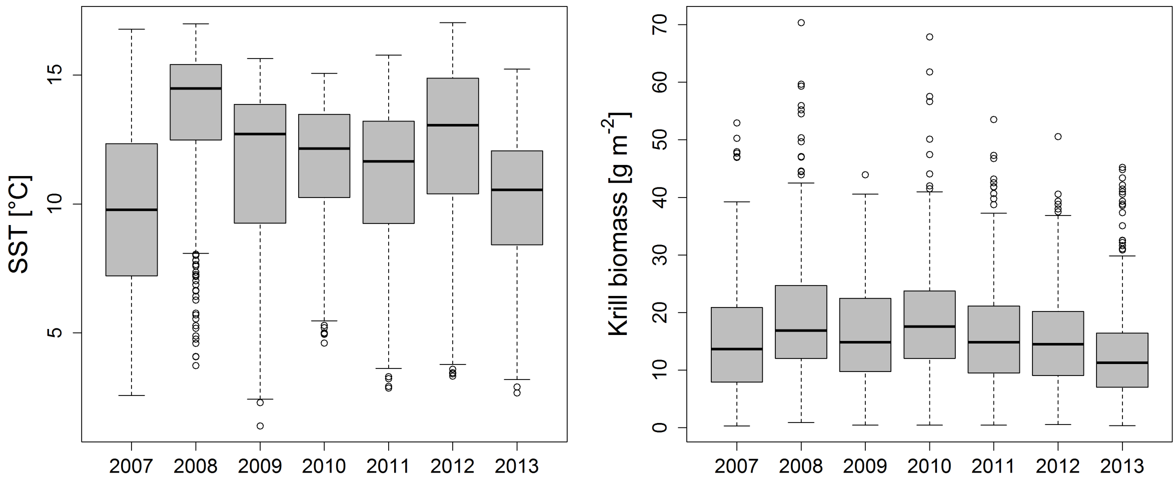

Fig. S8. Average annual sea surface temperatures (SST) and modelled krill biomass over all grid cells.

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