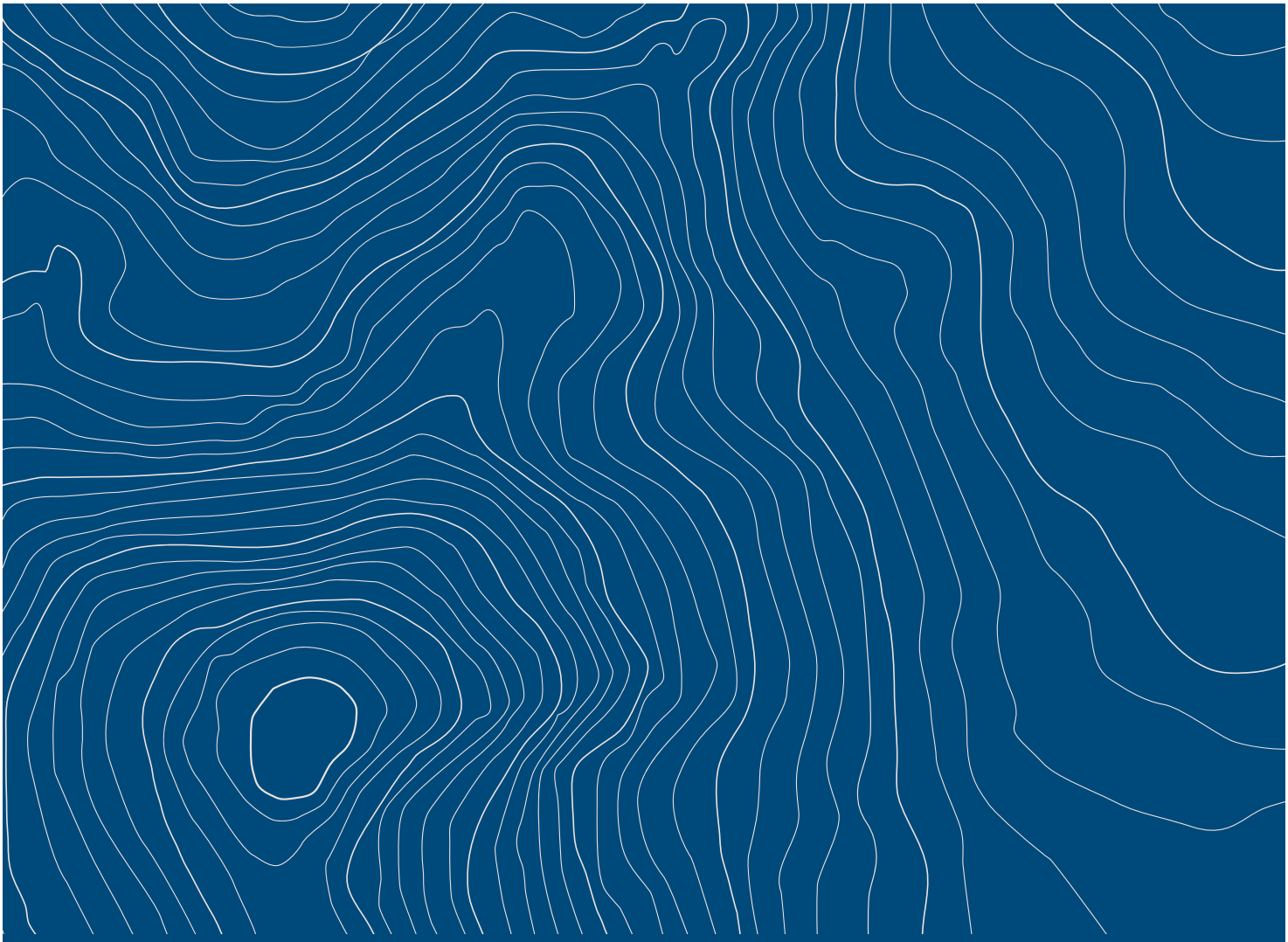


Dublin Array
Generation for generations

EIAR Volume 4: Offshore Infrastructure Technical Appendices Appendix 4.3.5-3 Dublin Array OWF Estimating harbour porpoise abundance using spatial and temporal modelling

Kish Offshore Wind Ltd

RWE SLR GoBe
APEM Group

www.dublinarray-marineplanning.ie

Dublin Array OWF: Estimating harbour porpoise abundance using spatial and temporal modelling - update

ML Burt and M Chudzinska, CREEM, University of St Andrews

Please cite this report as: ML Burt and M Chudzinska (2021) Dublin Array OWF: Estimating harbour porpoise abundance using spatial and temporal modelling - update. Report number CREEM-2021-07. Provided to SMRU Consulting, July 2021 (Unpublished).

Document control

Please consider this document as uncontrolled copy when printed.

Version	Date	Reason for issue	Prepared by	Checked by
Final	13/07/2021	For release	LB/MC	

Summary

Information was required on the spatial distribution of marine mammals off the eastern coast of Ireland, close to Dublin. Shipboard surveys took place from June 2019 to April 2021, inclusive. The ship travelled along pre-determined track lines (transects) and trained observers searched for marine mammals, recording the species, group size and other relevant information about the animals they detected. A total of 2,741 km were covered on search effort during 19 surveys. The most frequently detected species was harbour porpoise (135 groups) and this report describes the methods used to obtain the temporal and spatial density of harbour porpoise in the survey region and presents the results. This report updates a previous report which described results using data from surveys from June 2019 to September 2020 only.

Spatial modelling methods of Hedley and Buckland (2004) use estimated number of individuals in a small segment of transect as the response in a model, and location and environmental variables as explanatory variables. Estimated numbers per segment were obtained using the observed numbers adjusted for detection probability; line transect distance sampling methods (Buckland *et al.* 2001) were used to estimate detection probability.

Distance sampling methods rely on certain assumptions being valid, in particular, certain detection on the transect line; if this assumption is not valid, then estimates will under-estimate the true abundance. This may particularly affect estimates for cryptic animals such as harbour porpoise. No corrections for uncertain detection on the transect line have been applied to the estimates presented here and estimates should be considered as relative density estimates.

Introduction

A series shipboard surveys were undertaken to collect data in order to estimate marine mammal seasonal density and abundance in a region off the eastern coast of Ireland, close to Dublin. Nineteen shipboard surveys took place from June 2019 to April 2021, inclusive.

Line transect distance sampling methods (Buckland *et al.* 2001) of data collection were used; the ship travelled along pre-determined transects, or track lines, and trained observers searched for animals, recording relevant information when an animal, or group of animals, was detected. In this report, data from all surveys are combined to estimate density and abundance for harbour porpoise (HP; *Phocoena phocoena*).

The spatial distribution of HP in the survey region was of interest and modelling methods of Hedley and Buckland (2004) were used to estimate density as a spatially varying surface. The surveys took place throughout the year (although no surveys took place in February) and so temporal changes were also considered.

Survey methods

Survey design

The study region (Figure 1a) was approximately 266 km^2. Thirteen parallel transects were located with a random start point in the study region. Transects were oriented east-to-west approximately 2 km apart (Figure 1b) and in total were 141 - 160 km in length, depending on the survey.

Surveys took place in the following months (inclusive):

- June 2019 to January 2020,
- May 2020 to September 2020,
- December 2010 and January 2021,
- March 2021 and April 2021 (two full surveys each month).

The transects were covered in each survey either in one day or in two consecutive days.

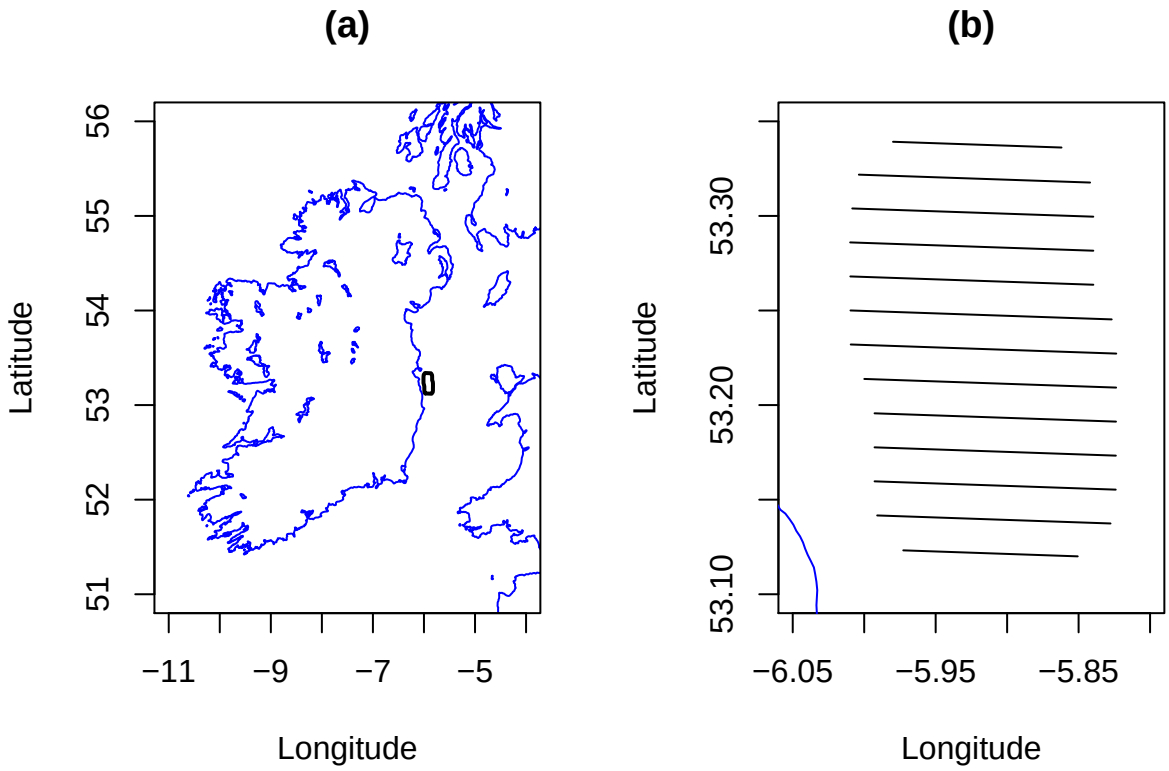


Figure 1. a) Location of the survey region and b) designed transects.

Search protocol

Observers travelled on a ship; there were two observers and a data recorder. For each detection, the observers recorded the angle to the detection relative to north, distance (or reticles) to the animal, species, group size as well as other information. Environmental conditions were also recorded along each transect (e.g., Beaufort sea state). The environmental conditions recorded at the start of a transect, or section of transect, were used for that transect, or section. Note that in Chudzinska (2021) mean value from all the recordings of sea state along each transect was used, hence there are some minor discrepancies between these two reports.

Statistical methods

The count method of Hedley and Buckland (2004) was implemented to model the trend in spatial distribution for HP. The response variable for spatial modelling was the estimated number of individuals, $\hat{n}_i$, in segment i of track line (of length l_i) calculated using a Horvitz-Thompson-type estimator (Horvitz and Thompson 1952) as follows:

$$\hat{n}_i = \sum_{r=1}^{R_i} \frac{s_{ir}}{\hat{p}_r} \tag{1}$$

where s_{ir} is the recorded group size for group r in segment i and R_i is the number of detected groups in segment i. The parameter $\hat{p}_r$ is the estimated probability of detection for group r in segment i; this was

estimated using distance sampling (DS) methods (Buckland *et al.* 2001).

Given the difficulty of detecting HP only search effort, and accompanying detections, in Beaufort sea states ≤ 2 are generally included in the analysis (e.g. Hammond *et al.* 2017). In this series of surveys, some surveys were conducted mostly, or entirely, in Beaufort sea states > 2 and would be excluded, hence, for this analysis, search effort and detections for Beaufort sea states ≤ 3 are included.

Probability of detection

Two critical assumptions of DS methods are that all groups on the transect line (i.e., at zero perpendicular distance) are detected with certainty and that distance measurements are exact. Given these assumptions, the distribution of perpendicular distances is used to model how the probability of detection decreases with increasing distance from the line.

The average probability of detection of animals within a distance w of the line, p, was estimated from a detection function model fitted to the observed distribution of perpendicular distances. Perpendicular distances were truncated (at distance w) to avoid a long tail in the detection function. The effect of Beaufort sea state was incorporated into the detection function model by setting the scale parameter in the model to be an exponential function of the covariates (Marques and Buckland 2004). To ensure sufficient sample sizes in each Beaufort sea state levels, levels were combined to create three categories: 0-1, 2 and 3.

Only sightings classed as primary sightings (where this was recorded) were included in the analysis.

Selection of the detection function was described in Chudzinska (2021) and so only basic details are provided here. No correction was made for uncertain detection on the transect line (i.e. it was assumed that $g(0) = 1$).

Modelling density of individuals

The estimated number of individual animals per segment along the transect lines were used to estimate harbour porpoise density in the region of interest; this approach models trend in the spatial distribution and allows it to vary throughout the region of interest, rather than assuming constant density.

The estimated number of individuals $\hat{n}_i$ in each segment (with known area) was enumerated, where i indicates an individual segment, formed the response variable in the spatial density model. Counts are often modelled using a Poisson distribution, however, these data were likely to be over dispersed (i.e. more variable than expected for Poisson distributed data), and so an overdispersed Poisson distribution with mean μ_i was used. The mean was modelled as a function of location:

$$\mu_i = exp(log(a_i) + \beta_0 + s_k(D_{ik}) + s_l(X_i, Y_i))$$

where

- $log(a_i)$ is an offset term (a term with known regression coefficient) that corresponds to the log of the area of each segment ($a_i = wl_i$ where w and l_i are the truncation distance and length for each segment i, respectively),
- β_0 is an intercept,
- $s_k(D_{ik})$ represents one-dimensional smooth terms (e.g. depth) implemented using quadratic B-splines with flexibility chosen using spatially adaptive smoothing (SALSA) methods (Walker *et al.* 2010)
- $s_l(X_i, Y_i)$ represents a two-dimensional smooth term of location (determined for each segment i by X_i and Y_i). This used radial Gaussian basis functions with flexibility also determined using SALSA (Walker *et al.* 2010).

For an overdispersed Poisson, the mean-variance relationship of the Poisson is relaxed so that the mean is proportional to the variance. This function contains a multiplicative factor (known as an overdispersion parameter) which was estimated from the data.

The models were fitted using the complex region spatial smoother (CReSS) in a Generalized Estimating Equations (GEE) framework with SALSA for model selection (Walker *et al.* 2010) implemented in R (R Core Team, 2019) using the MRSea package (Scott-Hayward *et al.* 2018). GEEs allowed for standard errors to be calculated, taking account of any correlation in the residuals within a specified blocking structure (to provide robust standard errors). In the presence of residual correlation it is important to use robust standard errors and corresponding p-values which are adjusted for this correlation when determining statistical significance. For this reason, robust standard errors were employed and since it was likely that counts of HP within segments along the transect lines would be correlated (and the model is unlikely to explain this correlation in full), transects were specified as the blocking structure.

The target length of segments was 2 km but segments naturally varied in length due to changes in Beaufort sea state and breaks in search effort.

Candidate explanatory variables

The candidate explanatory variables considered were:

- day of year of the survey (called *doy*) taking a number between 1 and 365 (and 366 for 2020), entering the model as a 1-dimensional smooth term,
- depth (metres), entering the model as a 1-dimensional smooth term, and
- location, entering the model as a 2-dimensional smooth term.

Depth was obtained from ETOPO1, a 1-arc-minute global relief model (https://www.ngdc.noaa.gov/mgg/global/global.html). It varied through the survey region from about 5 to 40 metres (Figure 2).

Location was specified by the longitude and latitude of the mid point of the segment and these values were transformed into a distance (in km) east ($x.pos$) and north ($y.pos$), respectively, from a reference point in the survey region (6.02^{o}W, 53.1^{o}N); these transformed values were used in the model. This was to ensure that a unit change in the north-south direction was the same as a unit change in the east-west direction. Location is unlikely to determine HP distribution but acts as a proxy for other, unknown variables that will determine HP distribution; the aim of the modelling was to describe, rather than explain, density.

Similar to location, day of year is unlikely to determine HP distribution but acts as a proxy for temporal changes such as sea surface temperature.

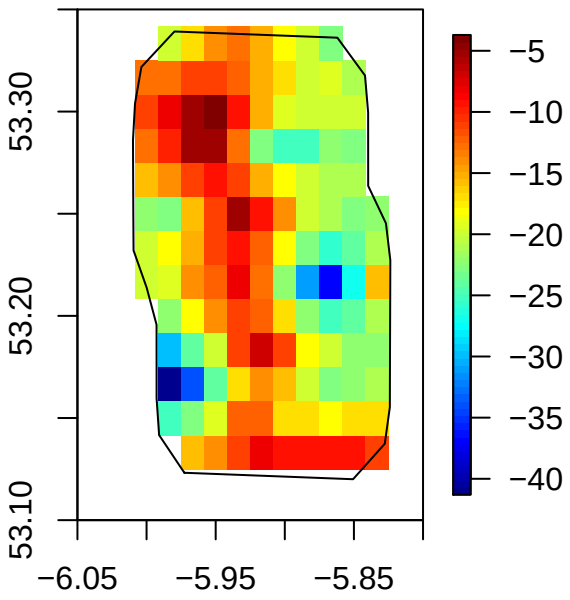


Figure 2. Depth values (metres) from ETOP01 in the survey region.

Model selection

The first step in model selection was to identify if any candidate explanatory variables were collinear and then, on the basis of remaining explanatory variables, define competing models.

Collinearity occurs when there is a linear association between two explanatory variables and this can lead to unreliably estimated coefficients and standard errors for the variables. Generalized variance inflation factors (GVIFs) were used to quantify the severity of multi-collinearity after first fitting a generalized linear model with estimated number of animals as the response and depth and location (Fox and Weisberg 2019) as explanatory variables. As a rule of thumb, values greater than 5 indicate that a variable is collinear with other variables and any collinear variables were excluded.

A selection procedure within the model fitting function was implemented so that one-dimensional terms in the model, such as depth, could be linear, smooth or removed (Scott-Hayward *et al.* 2018).

Diagnostics and model assessment

The predictive power of each model was assessed by fitting the k-fold cross validation score and a fit score (pseudo-R^2), a score measuring the square of the correlation coefficient between the observed values and the fitted values obtained from the model. This was done in order to assess predictive power in explaining the response and to select the model which explained the most variation for the response variable.

Diagnostics were performed to ensure model assumptions were valid for the selected model. To assess whether the residuals were correlated, a runs test for randomness in the residuals was performed. Transect number

(and month/year) were used as the blocking structure for cross-validation. To assess whether the blocking structure chosen was appropriate, a plot of the lag between residuals within a block and the correlation at each lag was obtained (an auto correlation function plot).

Predicted values for the survey data were obtained from the selected model and the residuals calculated. The residuals were plotted spatially and visually examined to determine if there were any patterns - ideally, there should not be any patterns discernible in residual plots.

Estimating abundance and 95% confidence intervals

Using the selected model, expected abundance was calculated for a grid of points (the prediction grid), with associated area and known values for the explanatory variables, throughout the region of interest. Total abundance was estimated by summing predicted abundance over all grid points in the region. The confidence intervals for total abundance were obtained from a parametric bootstrap procedure available in MRSea (Scott-Hayward *et al.* 2018) and implemented as follows. The selected model coefficients were re-sampled from a multivariate Normal distribution and new predictions obtained for every grid cell using these re-sampled coefficients. This process was repeated 500 times to build a distribution of abundance estimates for the region as a whole.

A confidence interval for total abundance needs to take account of the uncertainty both in the detection function and the spatial surface. Therefore, the coefficient of variation (CV) from the distribution of abundance estimates was combined with the CV of the detection probability using the delta method to obtain an overall CV for the abundance estimate. Log-normal confidence intervals (Buckland *et al.* 2001) were then obtained for each survey abundance estimate.

A similar parametric bootstrap procedure was used to assess the variation in the estimated density surfaces. A CV was obtained from the distribution of density estimates for each grid cell for each survey. The average CV (over all surveys) was then obtained for each grid cell. This surface did not include detection probability uncertainty.

Results

Survey effort and number of detections

Nineteen monthly surveys were conducted, contributing to a total of nearly 2,213 km of survey effort (Table 1) in Beaufort sea state ≤ 3. A total of 128 groups were detected; the number of groups detected per survey was generally low except for surveys in November 2019, late March 2021 and early April 2021 when 26, 15 and 34 groups, respectively, were detected.

Table 1. Summary of search effort (km) and number of harbour porpoise groups detected in Beaufort sea state ≤ 3 by survey. (The number of groups are not truncated by truncation distance, w.)

Survey	Year	Month	Day	Effort	Groups
1	2019	06	24	135	3
2	2019	07	27	106.5	1
3	2019	08	27	84.54	1
4	2019	09	22	145.3	3
5	2019	10	26	89.22	0
6	2019	11	06	159.5	26
7	2019	12	16	113	1
8	2020	01	23	150.6	8
9	2020	05	26	143.2	5
10	2020	06	25	120.7	3

Survey	Year	Month	Day	Effort	Groups
11	2020	07	13	121.2	3
12	2020	08	27	123.5	2
13	2020	09	07	65.87	7
14	2020	12	06	73.23	2
15	2021	01	25	96.59	2
16	2021	03	04	61.46	2
17	2021	03	20	143.7	15
18	2021	04	14	143.4	34
19	2021	04	26	141.9	10

Probability of detection

The maximum perpendicular distance for HP was 2208 m, however, to avoid a long tail in the detection function, a truncation of 500 m was used. This left 103 groups (note that in Chudzinska *et al.* (2021) 101 groups were left after truncation, due to different assignment of Beaufort scale; see Search protocol section). A hazard rate key function with Beaufort sea state as a covariate was selected to model detection probability (Figure 3).

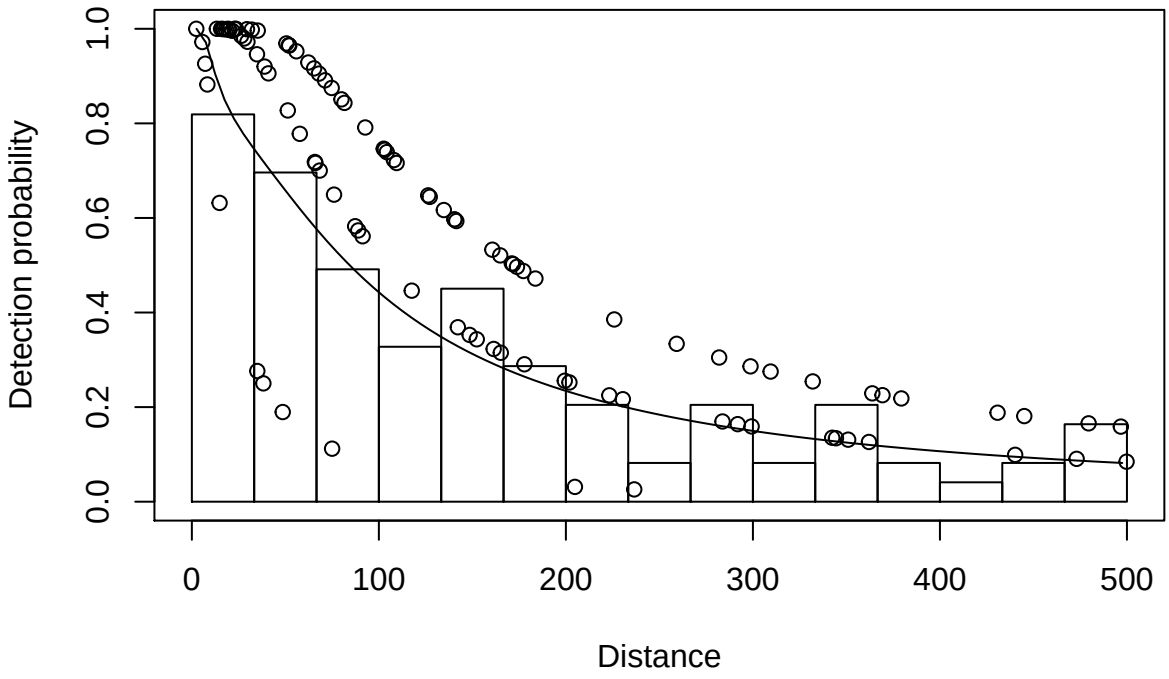


Figure 3. Average estimated detection function (black line) overlaid onto the scaled perpendicular distance (metres) distribution. The dots indicate estimated values for each detected group (and associated level of Beaufort).

The average probability of detection (over all Beaufort levels ≤ 3) was 0.31 (CV=0.26). The estimated detection probabilities for the different levels of Beaufort sea state were associated with each observed group

(on the basis of the recorded Beaufort for the segment) and eqn 1 used to obtain the estimated number of individuals per segment.

Model selection

The search effort and estimated number of individuals per segment are shown in Figure 4. The average length of segments was 1.4 km with minimum length 0.01 km and maximum length 2.99 km.

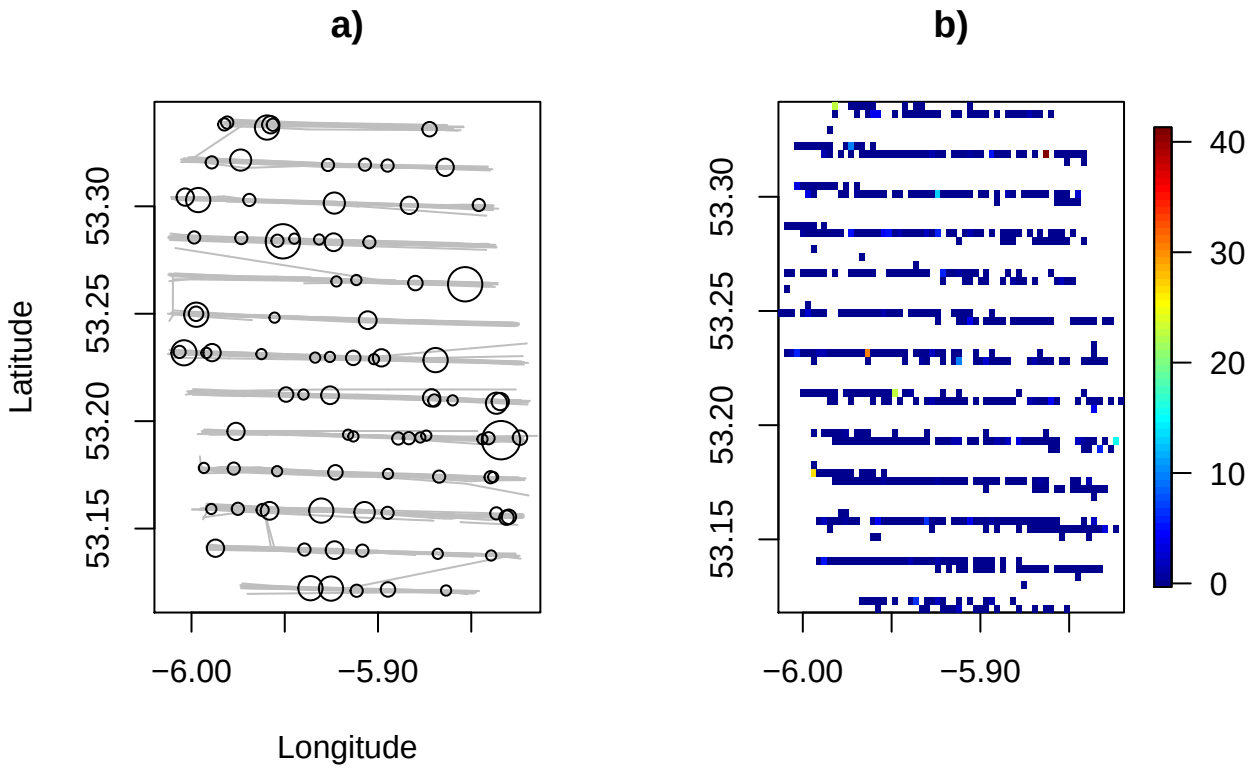


Figure 4. a) Search effort (grey lines) and locations where HP groups were detected (black circle). The area of the circle represents the estimated number of individuals in the segment (i.e. corrected for detection probability). The maximum estimated number was 31 individuals. b) Estimated density per segment (individuals/km^2); if segments overlapped, the average density was plotted.

The collinearity of the explanatory variables was assessed; the GVIFs were less than 5 (Table 2) and so all explanatory variables were considered in the modelling phase.

Table 2. GVIFs for the continuous explanatory variables in the model.

doy	depth	x.pos	y.pos
1.004	1.026	1.032	1.012

To start with, all explanatory variables were included (denoted by Model 1). Table 3 indicates that depth was not significant and depth was removed leaving day of year and location (Model 2); Table 4 indicated that this model was preferred and model was used for prediction.

Table 3 Analysis of variance table indicating the statistical significance of terms in Model 1 using Wald's

chi-square test: 'Df' = degrees of freedom, 'X2' = chi-sq test statistic and 'P(>|Chi|)' = probability of a value greater than the test statistic - a *p*-value.

Table 3: Analysis of 'Wald statistic' Table

	Df	X2	P(>|Chi|)
s(depth)	4	5.017	0.2856
s(doy)	6	32.1	1.559e-05
s(x.pos, y.pos)	2	0.7107	0.7009

Table 4. Analysis of variance table indicating the statistical significance of terms in Model 2 using Wald's chi-square test: see Table 3 for a description of table components.

Table 4: Analysis of 'Wald statistic' Table

	Df	X2	P(>|Chi|)
s(doy)	4	11.98	0.01748
s(x.pos, y.pos)	2	0.8722	0.6465

Diagnostics of Model 2 did not indicate any problems (see Appendix).

Density and abundance

An estimate of average abundance was obtained as described above for each survey, using the day of the year associated with each survey (using first day if surveys was conducted over two consecutive days) (Table 5). These estimates were consistent with the abundance estimates in Chudzinska (2021) obtained using design-based DS methods.

Table 5. Estimated abundance (number of individual harbour porpoise) and CV for each survey.

Survey	Year	Month	Day	Estimated.Abundance	CV
1	2019	06	24	18	0.4814
2	2019	07	27	28	0.432
3	2019	08	27	61	0.4114
4	2019	09	22	85	0.4166
5	2019	10	26	83	0.3946
6	2019	11	06	74	0.3964
7	2019	12	16	30	0.648
8	2020	01	23	40	0.565
9	2020	05	26	46	0.3575
10	2020	06	25	17	0.4947
11	2020	07	13	18	0.4968
12	2020	08	27	62	0.4123
13	2020	09	07	74	0.4161
14	2020	12	06	39	0.54
15	2021	01	25	42	0.5462
16	2021	03	04	90	0.3924
17	2021	03	20	102	0.3959
18	2021	04	14	97	0.396
19	2021	04	26	86	0.3844

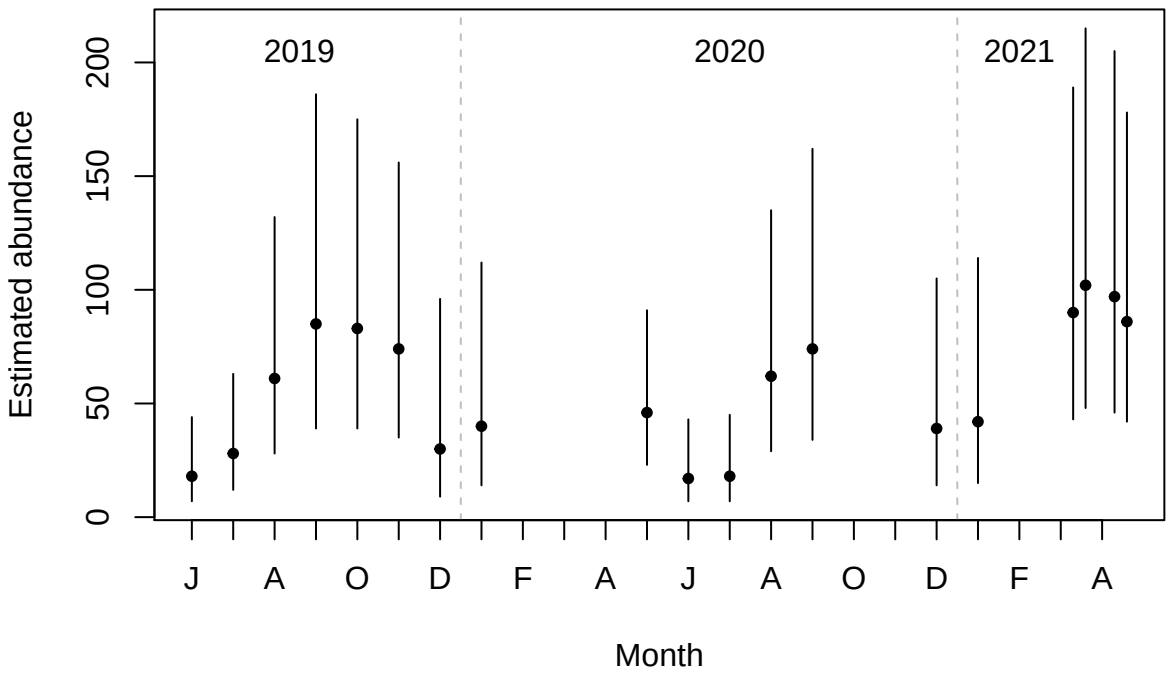
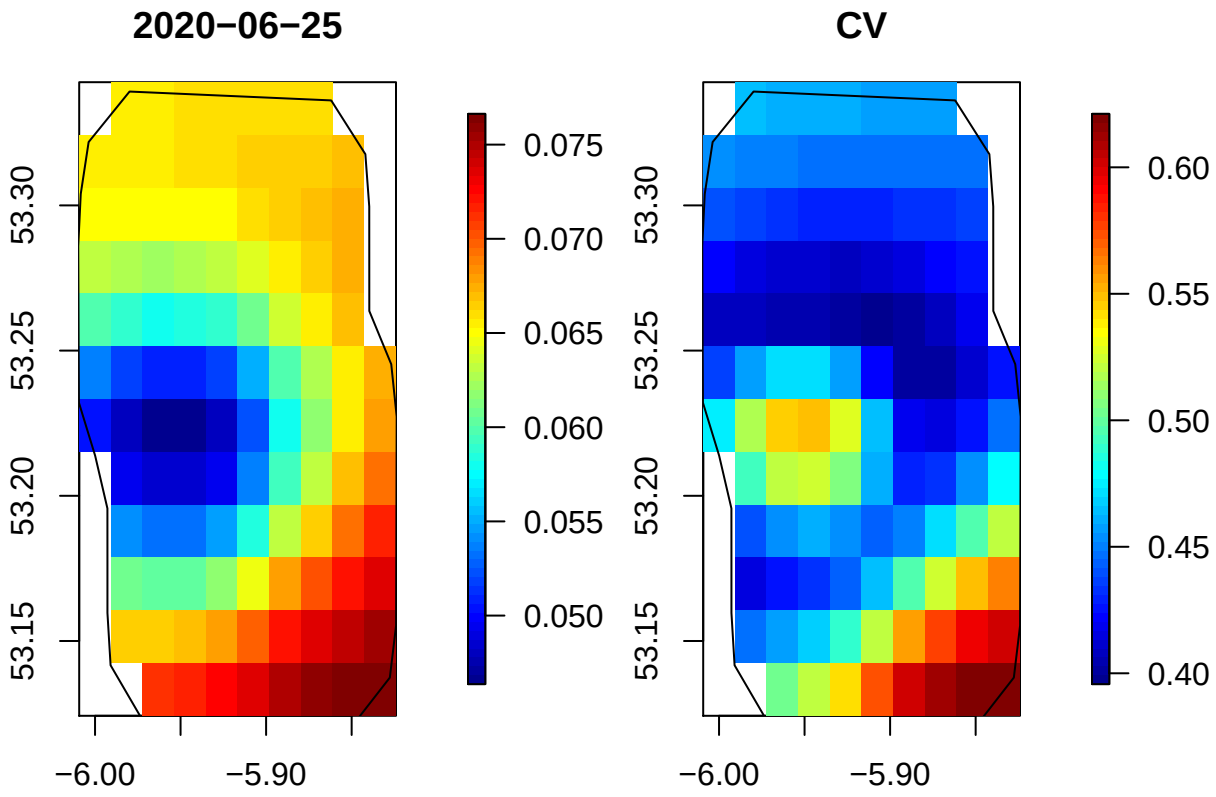


Figure 5. Estimated abundance for each survey (black dots). The vertical black lines indicate the (log-normal) 95% confidence intervals. The grey dashed lines divide different years.

The estimated spatial pattern will be the same for each survey but the density surface will increase, or decrease, depending on the day of year (see Fig A1, bottom plot). The estimated density for two surveys is illustrated in Figure 6; an average, overall surveys, is shown in Figure 7. Higher densities were estimated around the edges of the survey region. The CVs plotted over the prediction grid indicate that uncertainty is greatest in the regions of highest and lowest density.


2020−06−25
53.30
53.25
53.20
53.15
−6.00
−5.90
0.075
0.070
0.065
0.060
0.055
0.050
CV
53.30
53.25
53.20
53.15
−6.00
−5.90
0.60
0.55
0.50
0.45
0.40

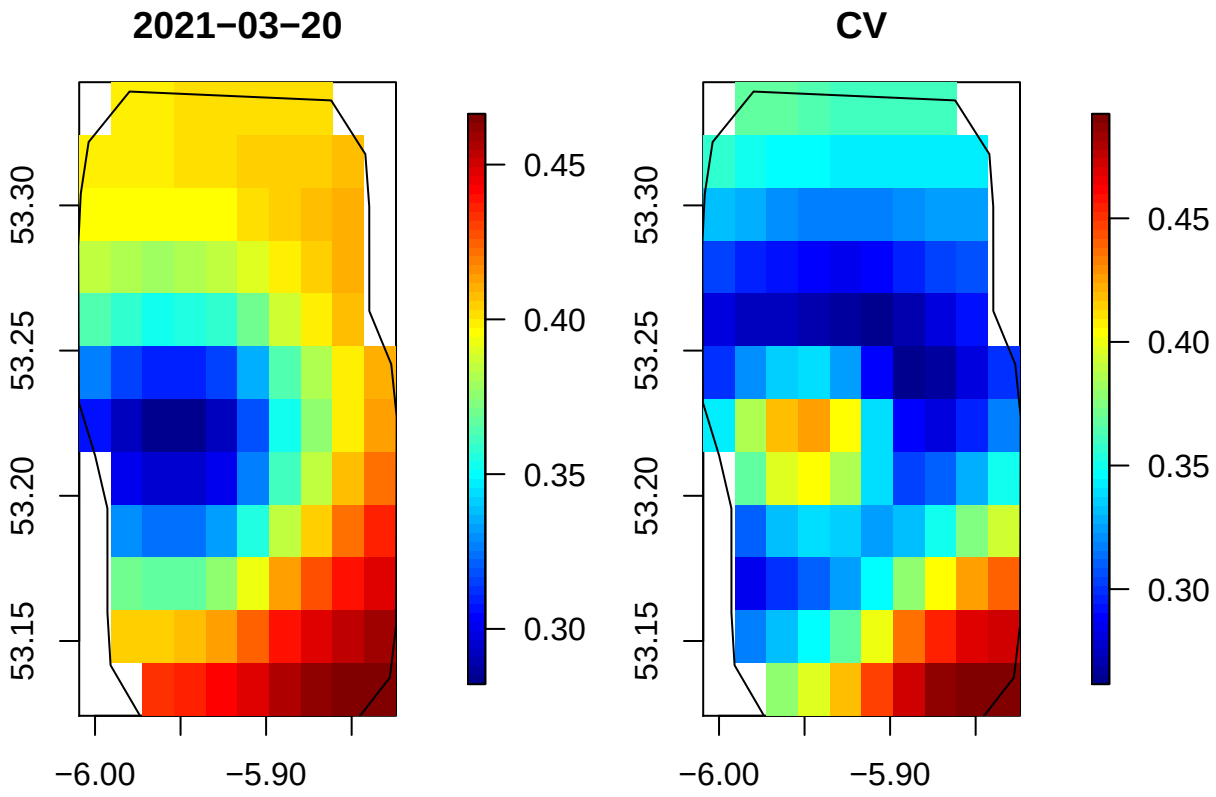


Figure 6. Estimated density of harbour porpoise (individuals per km^2) for two survey dates (left) and coefficient of variation (right). The spatial pattern remains the same for all the surveys but the density scale will change.

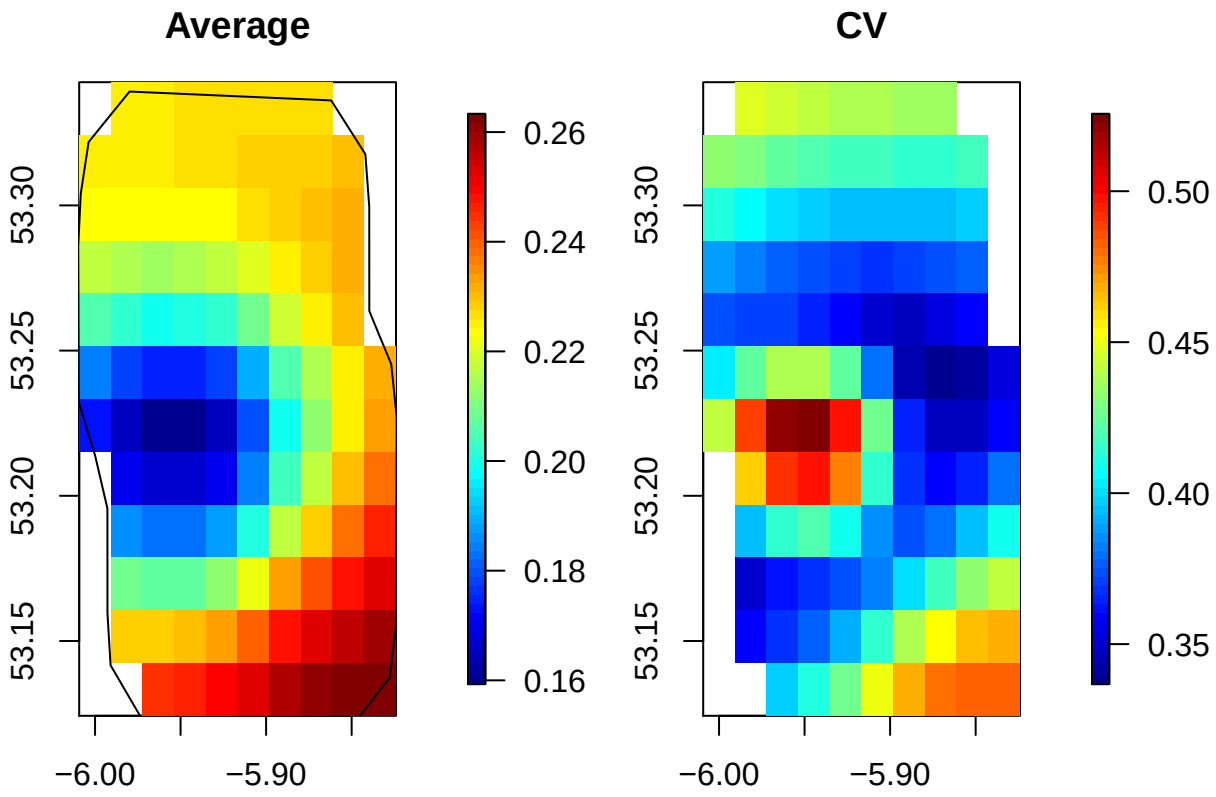


Figure 7. Estimated density of harbour porpoise (individuals per km^2) averaged over all surveys (left) and average coefficient of variation (right).

Discussion

Harbour porpoise density has been estimated throughout the region of interest using spatial modelling methods. Estimated counts of harbour porpoise in small sections of track line were modelled as a two-dimensional function of location and function of time. Depth did not appear to be an important explanatory variable (in addition to the other variables).

The spatial component of the model was not significant but the estimated density surface indicated that highest densities occurred around the edges of the survey region, albeit density values did not change substantially across the surface.

The temporal component of density was an important explanatory variable in the model. The smooth function for day of year estimated highest densities in March/April and September/October and lowest densities in June/July (Fig. A1).

Several alternative temporal variables could have been chosen instead of day of year (e.g. month or season). Day of year was chosen as it allows abundance to be estimated for individual surveys rather than average abundance for, say, a particular month or season.

The counts of harbour porpoise in segments were corrected for decreased detection away from the transect line using distance sampling methods. Distance sampling relies on certain detection on the track line. Only search effort (and associated detections) recorded in Beaufort sea state ≤ 3 were included in the analysis to satisfy this assumption, nevertheless, this assumption may well not be valid for HP. Hence, the estimated density may well under-estimate true density.

Another key assumption of distance sampling methods is that perpendicular distances are exact and measured without error. Systematic bias in the measurements can result in over, or under, estimating the detection probability. For further details see Burt (2020). In addition, a few errors may still be present in search effort (e.g. confirmation required that observers were on search effort transiting between transects, Fig. 4). If the search effort is reduced (due to corrections, say), the encounter rate, and hence, density will increase. See Chudzinska (2021) for a further discussion.

References

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (2001) Introduction to distance sampling. Oxford University Press

Burt ML (2020) Dublin Array OWF Marine Mammal Abundance Estimates. Report number CREEM-2020-06. Provided to SMRU Consulting, December 2020 (Unpublished).

Chudzinska M (2021) Dublin Array OWF Marine Mammal Abundance Estimates. Report number CREEM-2021-XX. Provided to SMRU Consulting, July 2021 (Unpublished).

Fox J and Weisberg S (2019) An R Companion to Applied Regression, Third Edition. Thousand Oaks CA: Sage. URL: https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Hedley SL and Buckland ST (2004) Spatial models for line transect sampling. Journal of Agricultural, Biological and Environmental Statistics 9:181-199.

Marques FFC and Buckland ST (2004) Covariate models for the detection function. In Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL and Thomas L (eds) Advanced distance sampling. Oxford University Press, Oxford, UK

Hammond PS, Lacey C, Gilles A, Viquerat S, Borjesson P, Herr H, Macleod K, Ridoux R, Santos MB, Scheidat M, Teilmann J, Vingada J, Oien N (2017) Estimates of cetacean abundance in European Atlantic waters in summer 2016 from the SCANS-III aerial and shipboard surveys. Available from https://synergy.st-andrews.ac.uk/scans3/2017/05/01/revised-results/

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Scott-Hayward LAS, Oedekoven CS, Mackenzie ML, CG, Rexstad E (2018) MRSea package: Statistical Modelling of bird and cetacean distributions in offshore renewables development areas. University of St. Andrews: Contract with Marine Scotland: SB9 (CR/2012/05). <URL: http://creem2.st-and.ac.uk/software.aspx>

Scott-Hayward LAS, Oedekoven CS, Mackenzie ML, CG (2019) Vignette for the MRSea Package v1.01: Statistical Modelling of bird and cetacean distributions in offshore renewables development areas. University of St. Andrews. Centre for Research into Ecological and Environmental Modelling. <URL:http://creem2.st-and.ac.uk/software.aspx>

Walker C, Mackenzie M, Donovan C and O'Sullivan M (2010) SALSA - a Spatially Adaptive Local Smoothing Algorithm. Journal of Statistical Computation and Simulation 81(2):179-191

Appendix: Diagnostics of spatial model

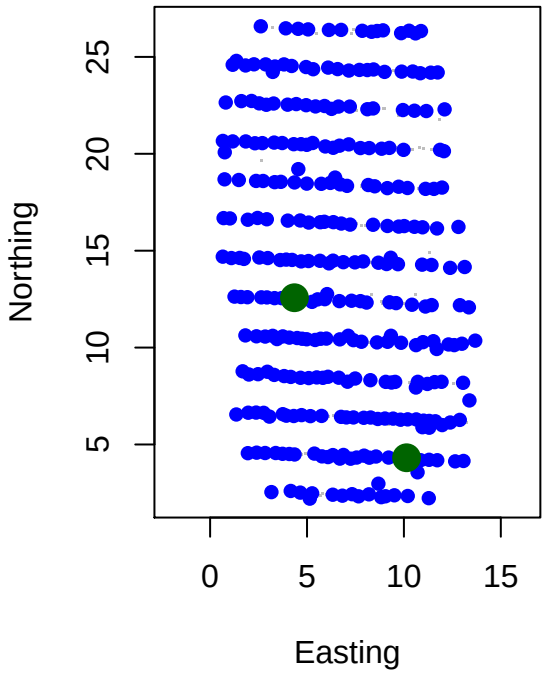
To investigate the selected model various diagnostics were performed. Fit statistics for each model (mean cross validation score and pseudo R^2) indicate that Model 2 had the lowest cross validation score.

Table A1 CV scores for fitted models.

Model	CV.score	Fit.score
1	3.016	0.02856
2	2.983	0.006813

The diagnostics of the final model were investigated to ensure that the selected model was valid; the diagnostic plots for this model are as follows:

- Figure A1 shows the position of the fitted knots for the two-dimensional smooth term of location.
- Figure A2 is an auto-correlation function (ACF) showing the correlation in the residuals along the blocking structure. In this plot the correlation declines to zero as distance increased which indicates that the blocking structure was appropriate.
- Figure A3 shows the predicted values obtained from the final model for locations recorded in the survey data. Predictions for the whole of the survey region are shown in the main text.
- Figure A4 shows the residuals from the selected model. This plot did not indicate any systematic patterns in the residuals which would give cause for concern. The few large residuals coincide with the large observed values which are very difficult to fit.




```
## [1] "Making partial plots"
```

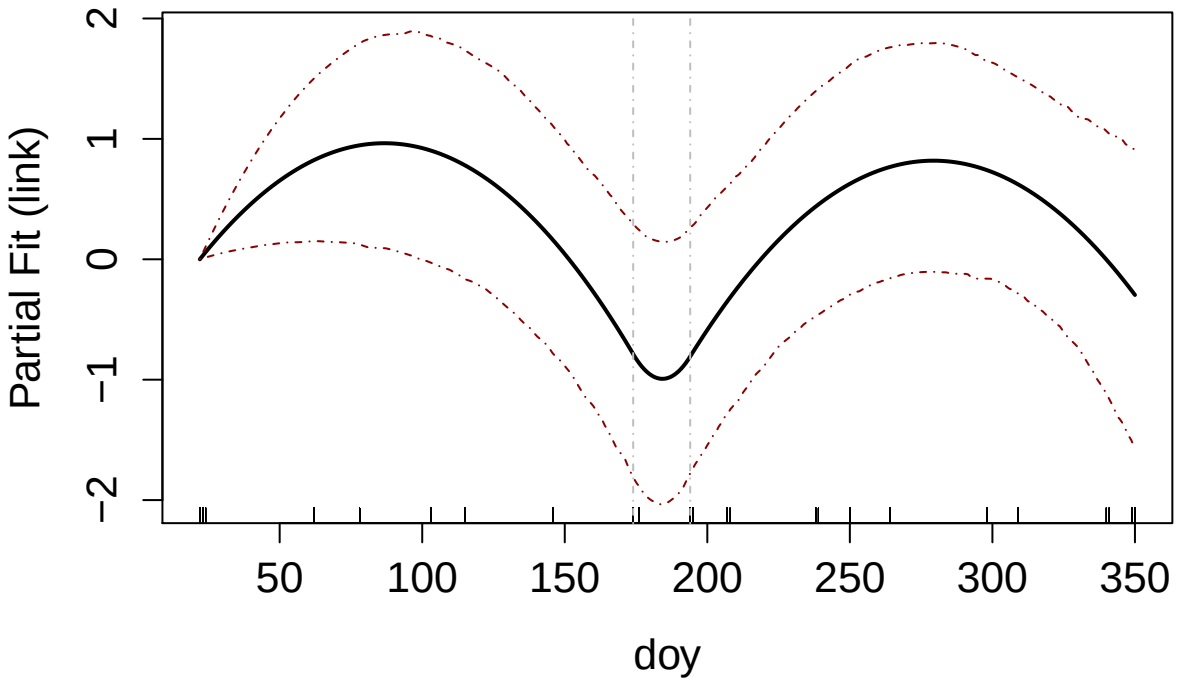


Figure A1. Location of the knots for the two-dimensional term of location in the final model (large green dots), location of knots in the prediction grid (blue) and grey dots indicate the segments (top plot). Plot showing the smooth term for day of year (on the scale of the link function) (bottom plot). Tick marks along the x-axis indicate the days when surveys took place and the vertical grey lines indicate the knot placements.

A runs test was performed to determine if the residuals were correlated (shown below). The p-value associated with this test did not indicate that the residuals were correlated.

Table A2. Summary output from the Runs Test.

```
##
##  Runs Test - Two sided; Empirical Distribution
##
## data:  residuals(HP.2dOutput$bestModel, type = "pearson")
## Standardized Runs Statistic = -20.515, p-value = 0.448
```

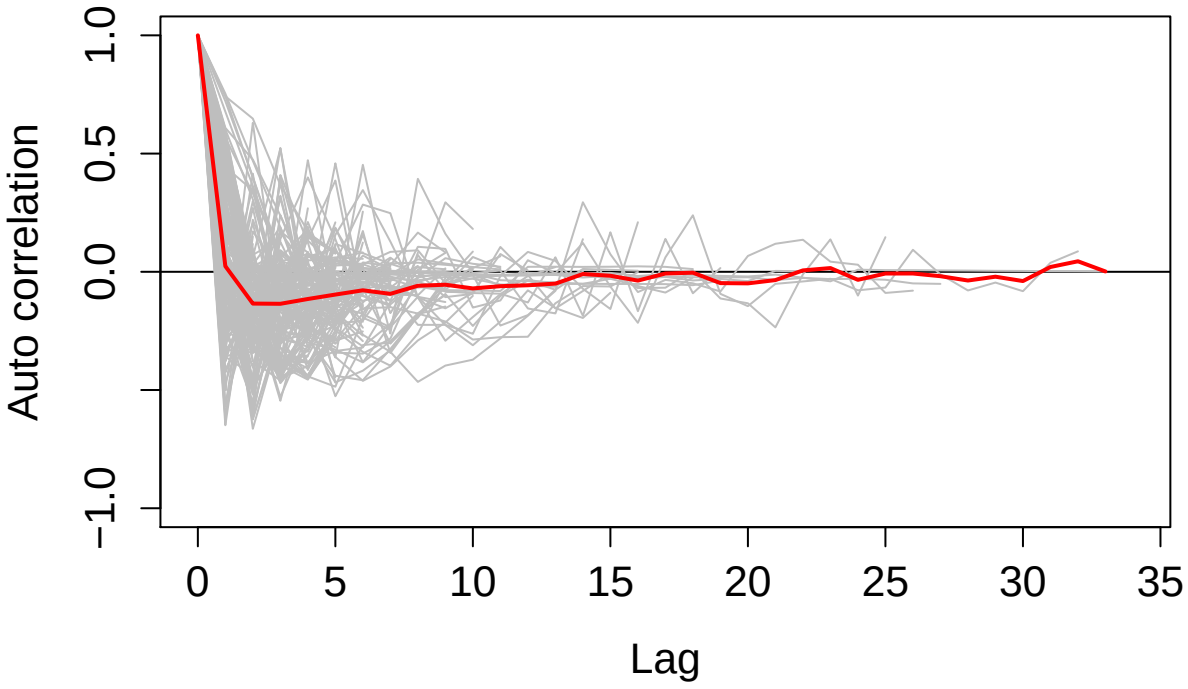


Figure A2. Plot of the correlation in the residuals for each block (grey lines). The mean correlation at each lag is indicated in red. The correlation should decay to zero (as in this case) which indicates that the correlation between residuals within blocks (transects) reduces as the distance (or lag) between the segments increases.

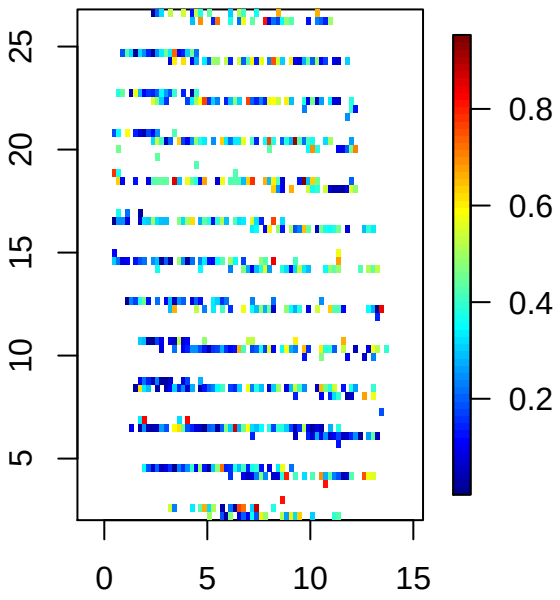


Figure A3. Predicted values (numbers of HP per segment) obtained from the selected model for the recorded survey data (averaged over observed data records where locations overlap).

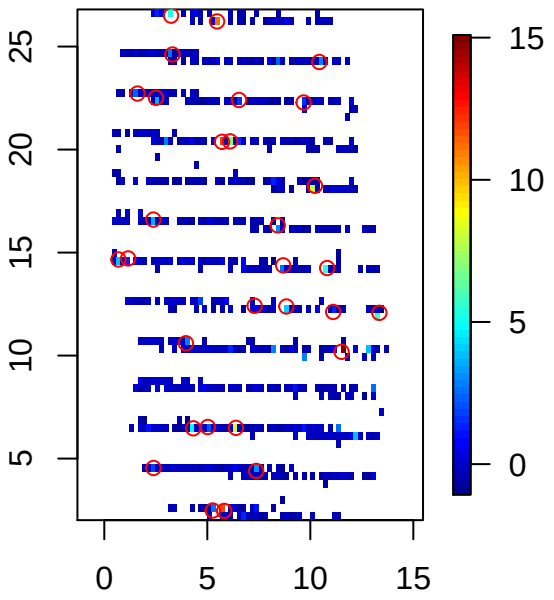


Figure A4. Residuals (difference between the observed number of HP and predicted number) plotted at each location (averaged over observed data records where locations overlap). The red circles indicate the location of segments where large numbers of HP (>5 individuals per segment) were estimated.