Analysis of Aerial Surveys conducted off Cape Hatteras, 2012

C.G.M. Paxton, CREEM, University of St Andrews

ABSTRACT

This report describes the analysis of data collected during aerial surveys off Cape Hatteras, undertaken by the University of North Carolina at Wilmington, for the period January 2012 to December 2012. The aim of the analysis was to estimate the abundance of cetaceans and turtles in the region of interest using density surface modelling and other methods. Detection functions were generated for the species groups dolphins, baleen whales, other whales (beaked, pilot and sperm whales combined) and turtles. There were sufficient detections to generate spatio-temporal abundance estimates (uncorrected for availability bias or perception bias) for bottlenose dolphins, loggerhead turtles and a species group made up of pilot whales and beaked whales. Predicted abundance of bottlenose dolphins varied between 490 animals (95% confidence interval; 220 – 900) in February and 1,430 animals (20 – 2510) in August. Predicted abundance of beaked and pilot whales varied between 390 (160 – 1,150) in February and 1,110 (460 – 3,030) in September. Predicted abundance of loggerhead turtles varied between 60 (30 – 100) in February and 190 (110 – 330) in May. For *Stenella* sp. dolphins, sperm whales (*Physeter macrocephalus*), baleen whales and leatherback turtles, abundance estimates were obtained assuming no spatio-temporal fluctuations due to the small number of group detections for these species. The average predicted abundance for *Stenella* sp. dolphins was 1,300 animals (95% confidence interval: 500 – 2,500). The predicted abundance for sperm whales was just 10 animals (0 – 27). The predicted abundance of baleen whales was also 10 (4 – 17) and for leatherback turtles it was 3 (0 - 9).

INTRODUCTION

The Cape Hatteras aerial surveys were carried out by the University of North Carolina at Wilmington (UNCW) almost every month, from late January to early December 2012. The aim of the surveys was to obtain relative abundance estimates for cetaceans and turtles in the region and where possible maps of predicted abundance.

SURVEY METHODS

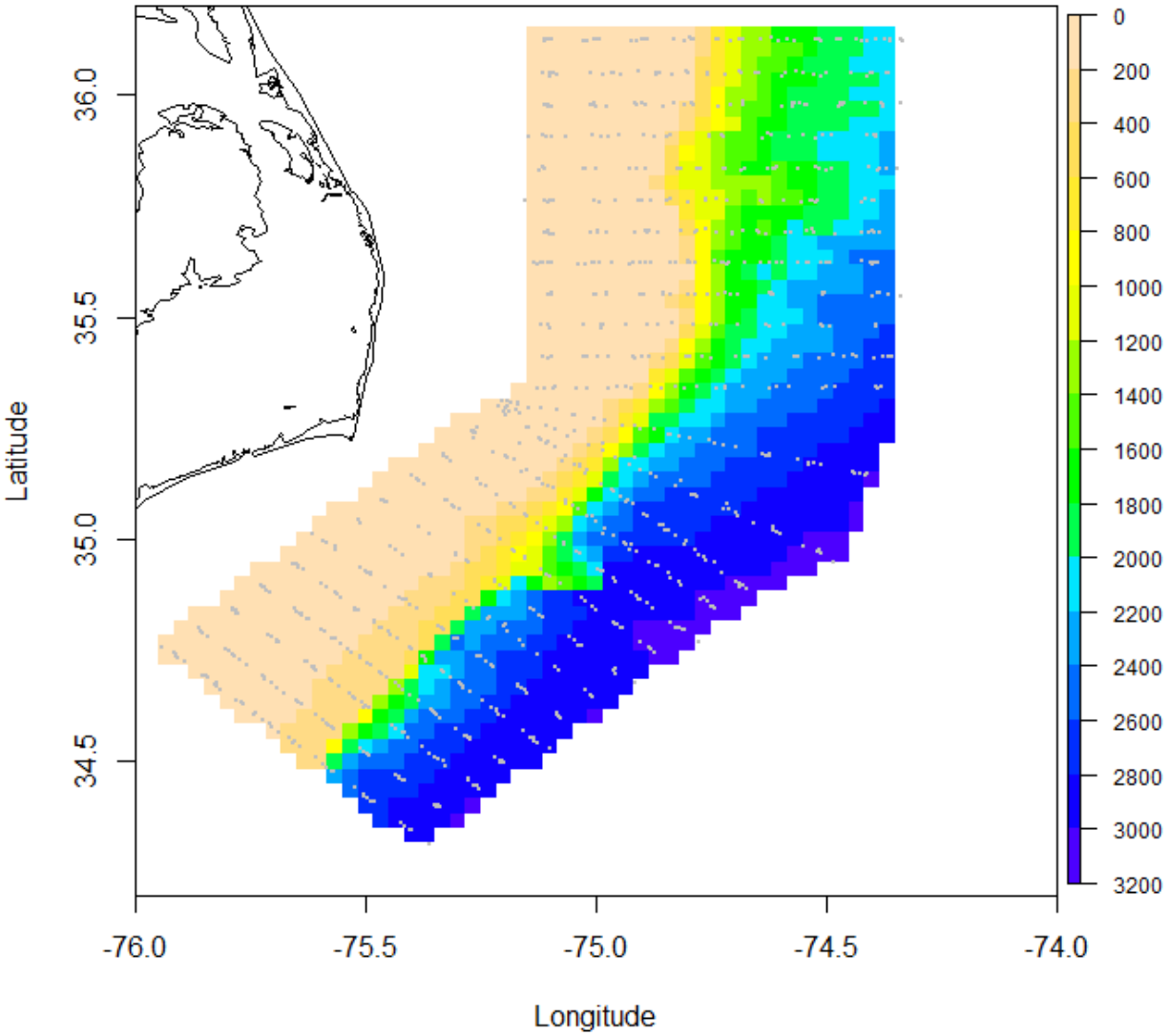
Region of interest and survey area

The Hatteras region and associated bathymetry along with survey tracks are given in figure 1. The survey region covered an area of 15,765 km^2 and covers both shallow shelf waters and deeper abyssal waters.

Survey effort

Twenty-six transect lines were aligned mainly east-west across the region to the north of Cape Hatteras and northwest-southeast in the region to the south of Cape Hatteras (Figure 1). Not all transects were covered each month.

Figure 1. Hatteras survey/prediction region with depth (m) indicated by colours. Each grid cell has dimensions 1/30 degree of latitude and 1/30 of a degree of longitude. Grey lines indicate midpoint of segments. Data is show from all months.



Aerial search protocol

The plane flew at a height of 305 m above sea level. Two observers (one on either side of the plane) searched for marine animals and when an animal was detected, they recorded vertical and horizontal angle to the sighting, species and group size. When a whale or dolphin was detected, search effort was stopped and the plane left the transect line to investigate the sighting and confirm species and group size. Search effort was resumed along the transect line after leaving the sighting. Environmental conditions were also recorded.

Estimates of perpendicular distance from the transect line to the sighting were obtained either by reference to direct estimates of distance by observers, trigonometry from the declination angle of the plane to the observed animals or, for cetaceans, by trigonometry, from the position of the plane at first observation of the animals and subsequent location directly above the animals.

STATISTICAL METHODS

The aim of the analysis was to estimate a density map for each species/taxa using the count method of Hedley *et al*. (2004, 1999). However, the numbers of sightings of some taxa were too few to estimate a density surface so instead a uniform surface was assumed – this was equivalent to undertaking a conventional line transect sampling analysis (Buckland *et al*. 2001).

For each taxa, the probability of detection associated with each sighting (assuming full detection on the trackline) was first estimated and this was then used to estimate abundance in small segments of the trackline. These estimated densities formed the response variable in a generalized additive model (GAM) with location, habitat and temporal variables as potential explanatory variables. After model selection, the chosen models were used to estimate density for the region of interest and abundance was obtained by numerically integrating under the predicted density surface. Note that the resulting abundances are relative (rather than absolute) because they do not take into account imperfect detection on the transect line nor the availability of animals at the surface.

Estimation of detection probabilities

In conventional line transect sampling, the probability of detection depends only on the perpendicular distance of the sighting to the transect (y) and at zero perpendicular distance the probability of detection is assumed to be one (denoted by $g(0)=1$). Both a hazard-rate ($1-\exp(-y/\sigma)^{-b}$) and a half-normal ($\exp(-y^2/2\sigma^2)$) form were considered as suitable forms for the detection functions (σ is the scale parameter) the most appropriate form for the relevant data (Buckland *et al*. 2001). The effects of covariates, other than perpendicular distance, were incorporated into the detection function model by setting the scale parameter in the model to be an exponential function of the covariates (Marques 2001). Thus, the probability of detection becomes a multivariate function, $g(y,\boldsymbol{v})$, representing the probability of detection at perpendicular distance y and covariates $\boldsymbol{v}$ ($\boldsymbol{v} = v_1,..,v_Q$ where Q is the number of covariates). The scale term, σ, has the form:

$$\sigma_k = \exp\left(\beta_0 + \sum_{q=1}^{Q}(\beta_q v_{kq})\right)$$

and β_0 and β_q ($q=1,\ldots,Q$) are parameters to be estimated. With this formulation, it is assumed that the covariates affect the rate at which detection probability decreases as a function of distance, but not the shape of the detection function. The covariates considered for inclusion into the detection function were Beaufort sea state, group size, cloud cover, visibility, glare and species. A forward, stepwise selection procedure was used to decide which covariates to include in the model, with a minimum Akaike's Information Criterion (AIC) inclusion criterion. All model selection was performed using a set of customised functions (mrds v.2.0.6, Laake et al. 2013) within the statistical programming package *R* (*R* Developmental Core Team, 2002). This facilitated estimation of variance within *R* (see below).

Some of the data from Hatteras was supplemented by additional survey data from the UNCW right whale surveys to increase sample sizes for detection function fitting. Despite this there was a paucity of sightings for individual species and so data were amalgamated across species into four groups with presumed similar detectabilities; all dolphin species, balaenopterid whales, other whales and turtles.

Estimation of density surfaces

The 'count model' of Hedley *et al*. (2004) was implemented to model the trend in spatial distribution of the different species. The response variable for this model is the estimated number of individuals in a small segment i of trackline, $\hat{N}_i$, calculated using an estimator similar to the Horvitz-Thompson estimator (Horvitz and Thompson 1952), as follows:

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{s_{ij}}{\int_o^w \hat{g}(y, v_{ij})\pi(y)dy}, \qquad i = 1,\ldots,T,$$

where for segment i, $\int_0^w \hat{g}(y, v_{ij})\pi(y)dy$ is the estimated probability of detection of the jth detected group, n_i is the number of detected groups in the segment and s_{ij} is the size of the jth group. The total number of transect segments is denoted by T. By assumption, $\pi(y)$, the probability density function of actual (not necessarily observed) perpendicular distances is uniform up to the truncation distance; this is satisfied by locating transects randomly or with a random start point.

Having obtained the estimated number of individuals in each segment, the density in segment i, $\hat{D}_i$, was estimated from $\hat{N}_i / a_i$ where a_i is the area of segment i. Segment area was calculated as the length of the segment multiplied by twice the truncation distance which was decided when modelling the detection function. The realised transect lines were divided up into distinct segments based on when vessels had gone on or off effort and whether there was a change in environmental characteristics. A target segment length of 10km was chosen as an appropriate compromise between maximising the ratio of non-zero to zero segments, maintaining environmental resolution and giving some measure of spatial independence, although some segments were much smaller if there had been a break in effort or change in environmental conditions. Due to the different sizes of each segment, the segment area was included as a weight (a term with a known regression coefficient) in the subsequent model.

The main aim was to fit a model with which to predict abundance. The environmental covariates considered for inclusion in the abundance models were longitude (*Lon*) and latitude (*Lat*), depth (*Depth*) and sea surface temperature (*SST*). Day of the year (*Dayofyear*) was fitted as a cyclic cubic spline so the smooth function for *Dayofyear* would have the same values at the start and end of the range of data values. Depths were obtained from the ETOPO1 one minute resolution relief data available from NOAA (http://www.ngdc.noaa.gov/mgg/global/global.html). Depths were associated with effort segments by finding the closest point in the bathymetry data to the midpoint of the effort segments using great circle distances. *SST* was obtained from NOAA Optimum Interpolation 1/4 Degree Daily Sea Surface Temperature data obtained from the Advanced Very High Resolution Radiometer (AVHRR) available from ftp://eclipse.ncdc.noaa.gov/pub/OI-daily-v2/NetCDF/ as described in part in Reynolds *et al*. 2007 and allocated to the appropriate segment by great circle distance and appropriate date.

Unsurprisingly, *Temp* was strongly related to *Dayofyear,* and *Lon, Lat* were correlated *Depth*, thus, the inclusion of only one of these correlated variables in the final models should not be interpreted as necessarily precluding the influence of others.

After preliminary analysis, it was decided that the most appropriate modelling process should use a two stage process; first modelling the presence-absence of animals in segments followed by modelling the density in non-zero segments (i.e. segments where animals were detected). The two resulting surfaces were then multiplied together to obtain estimated density. Generalized cross validation implemented in the *mgcv* package (v. 1.7-13. Wood 2006) in *R* (v. 2.15.0) was used for covariate selection in the presence-absence GAM augmented with diagnostic plots, using the principles described in Wood (2001). All non-factor covariates were considered for inclusion in the model as one dimensional (1D) smooths of untransformed covariate value and *Lat* and *Lon* were also considered as a two dimensional (2D) smooth. Taking into account the low percentage of segments containing sightings, a maximum of 4 degrees of freedom (5 knots) was allowed in the selection of 1D smooths and 11 degrees of freedom (12 knots) was allowed in the case of the 2D smooth, thus allowing moderate flexibility but reducing the possibility of fitting unrealistically complicated functions.

Prediction
The selected models were used to predict density of marine animals using a 2 minute resolution prediction grid. Animal abundance was estimated by numerically integrating under this predicted density surface.

Variance estimation
Variance was estimated by repeating (1,000 times) the entire abundance estimation process on samples drawn from the data to obtain a distribution of abundance estimates. Samples were obtained by sampling transects, at random and with replacement, such that the selected effort reflected the effort in the original sample. Confidence intervals were obtained from this resampling-derived distribution using the 2.5% and 97.5% percentiles to obtain the lower and upper limits, respectively, thus excluding the most extreme values.

RESULTS
Summary of search effort and number of sightings
The aerial surveys realised 9,190 km of search effort over the 12 months of the survey). The total numbers of sightings within the relevant truncation distances are given in Table 1 along with the group encounter rate (per km) and effort per month.

Detection functions
Fitted detection functions to the binned data for the four species groups are shown in figure 2 with details and resultant detection probabilities (assuming $g(0)= 1$) in table 2. Sightings for all species groups were right truncated to avoid a long tail in the estimated detection function. Of particular note was that the perpendicular distances for turtles were left truncated at 100m (i.e. distances within 100m of the trackline were excluded from analyses and 100m was subtracted from all remaining distances). The raw distribution of turtle distances is given in figure 3.

Table 1. Sightings and effort by month. Hatteras data only.

	Month	Sightings	Sightings after truncation	Effort	(No of group sightings after truncation/km)
All dolphin species	Jan	2	2	1332	0.002
	Feb	4	4	583	0.007
	March	14	14	1461	0.010
	May	19	18	1165	0.015
	Jun	17	15	1904	0.008
	Aug	6	6	704	0.009
	Sep	0	0	736	0
	Nov	3	3	318	0.009
	Dec	8	8	988	0.008
Baleen whales	Jan	2	1	1332	0.001
	Feb	3	3	583	0.005
	March	4	4	1461	0.003
	May	1	1	1165	0.001
	Jun	0	0	1904	0
	Aug	0	0	704	0
	Sep	0	0	736	0
	Nov	0	0	318	0
	Dec	2	2	988	0.002
Other whales (pilot whales, beaked whales, melon-headed whales and sperm whales)**	Jan	4	4	1332	0.003
	Feb	2	2	583	0.003
	March	9	8	1461	0.005
	May	8	8	1165	0.007
	Jun	12	12	1904	0.006
	Aug	8	8	704	0.011
	Sep	2	2	736	0.003
	Nov	4	4	318	0.013
	Dec	11	11	988	0.011
Turtles	Jan	3	3	1332	0.002
	Feb	1	1	583	0.002
	March	9	9	1461	0.006
	May	24	23	1165	0.020
	Jun	7	6	1904	0.003
	Aug	4	3	704	0.004
	Sep	1	1	736	0.001
	Nov	0	0	318	0
	Dec	10	10	988	0.010

Table 2. Summary of realized sightings, N, (within truncation distance) and fitted detection function (DF) models (HN=half normal) for various species groups. All data were binned into 200m intervals.

Species group	Truncation (m)	*N*	DF form	Covariates in addition to distance	Mean resultant detection probability (se)
All dolphin species	1400	70	HN	-	0.608 (0.063)
Baleen whales*	1400	16	HN	-	0.990 (0.277)
Other whales (pilot whales, beaked whales, melon-headed whales and sperm whales)**	1200	59	HN	Beaufort Sea State	0.640 (0.076)
Turtles	400 with left truncation at 100m	56	HN	Glare	0.829 (0.134)

**Includes 5 sightings from the UNCW right whale surveys. **Includes one sighting from UNCW right whale surveys.*

Figure 2. Fitted detection functions overlaid onto scaled perpendicular distance distributions. a. dolphins b. baleen whales, c. other whales & d. turtles(after left truncation at 100 m). N.B. distances were binned for model fitting into 200m intervals.

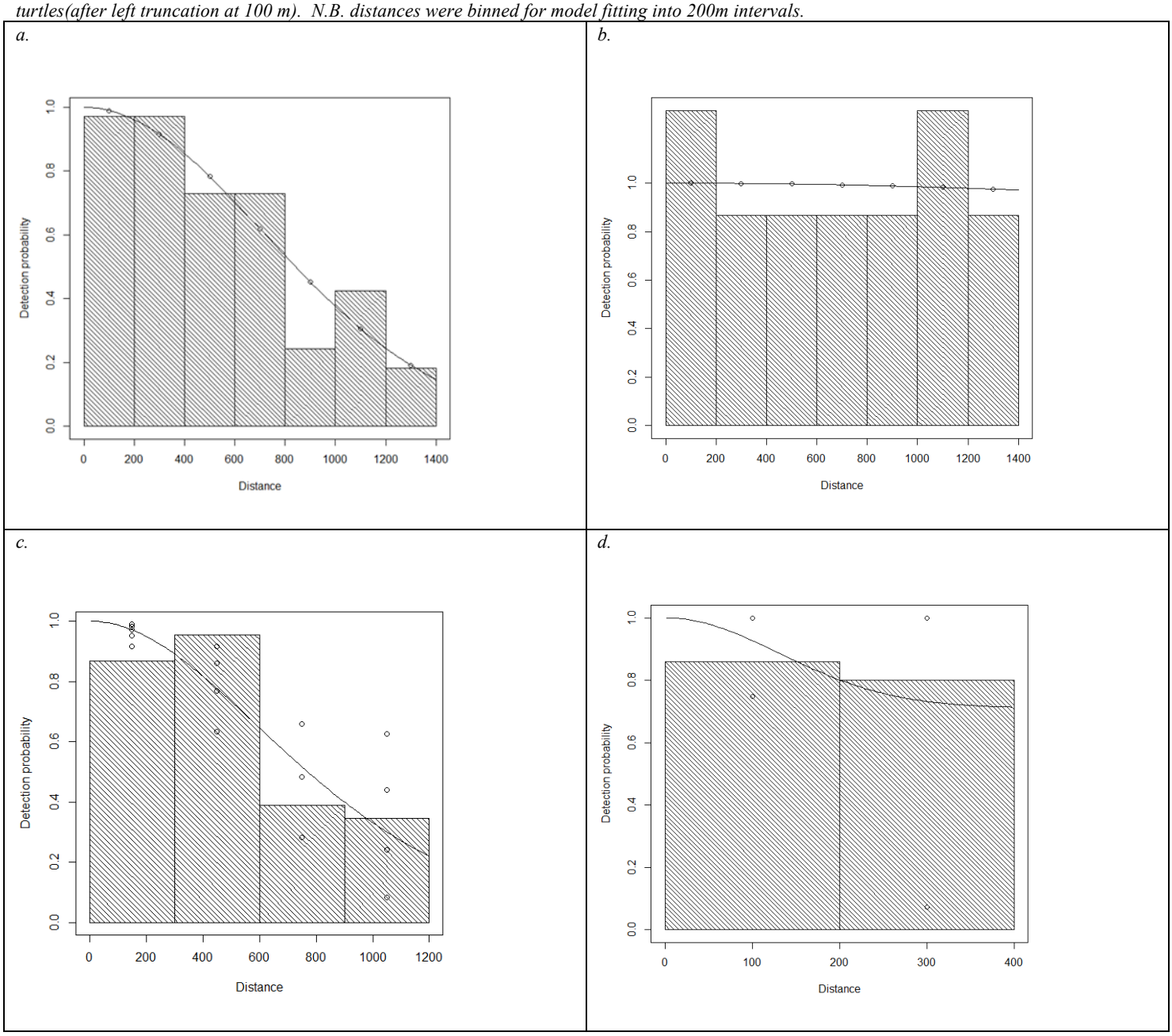
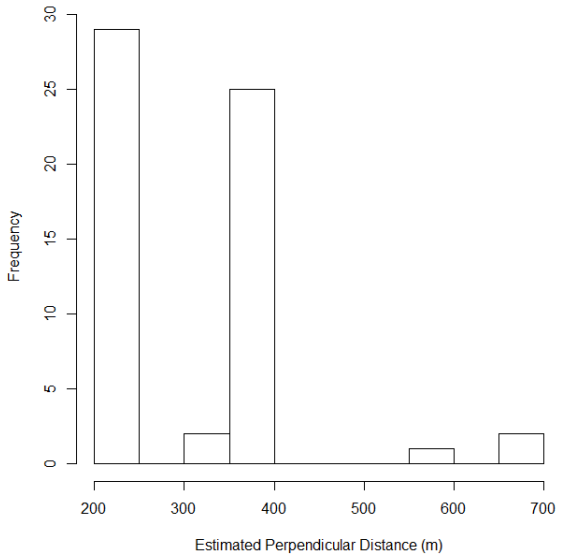


Figure 3. Actual distribution of turtle estimated perpendicular distance.



Density surface models

There were 1,003 segments in the realized data set with a mean length of 9.16km. As in previous analyses of survey data from this region e.g. Read *et al*. (under review), the data were characterised by having many segments with no sightings and only a few segments where animals were sighted (e.g. beaked and pilot whales were seen in only 54 segments out of 1,003). The segment densities are obviously also temporally correlated. No simple, optimal solution to modelling data of this type is currently available and so, as mentioned previously, segment densities were modelled as two-stage process with presence-absence modelled followed by non-zero densities modelled assuming a gamma distribution (*Dayofyear* was not considered for this model because in the bootstrap there was not be enough variability in this variable to cover the temporal prediction range). This approach, analogous to a hurdle model (Mullahy 1986), is not ideal in that the non-zero densities were very low indicating that some of the zeros may not be 'true' zeros (i.e. animals are never present) but 'transient' zeros (i.e. animals could be present but they were either not detected or did not happen to be there at that point). A zero-inflated approach (i.e. assuming response distributions developed to accommodate large numbers of zeros in the data) was tried, however, the models were very unstable.

In the case of *Stenella* dolphins, baleen whales and sperm whales there were too few non-zero data to fit a model and so a uniform surface was assumed (i.e. no spatio-temporal effect was fitted to the data). The final models chosen for bottlenose dolphins, other whales and loggerhead turtles are given in table 3.

Table 3. Summary of predictive density surface models, where s(var) indicates a smooth function of var, and percentage of deviance explained by the model.

Species	**Presence-absence model (Numer of segments = 1003)**		**Nonzero presence model**	
	Selected model	**Explained deviance %**	**Selected model (number of segments)**	**Explained deviance %**
Bottlenose dolphin	*s(Lon, Lat) + s(Depth) + s(Dayofyear)*	21.4	*s(Lon) + s(SST) (n = 45)*	27.3
Other whales	*s(Lon, Lat) + s(Depth) + s(Dayofyear) + s(SST)*	26.4	*s(Lon) (n = 54)*	20.6
Loggerhead turtles	*s(Lon, Lat) + s(Depth) + s(SST)*	32.1	*s(Lon, Lat) (n = 39)*	34.4

Estimates of abundance

Estimates of abundance are given in in table 4. Predicted abundance of bottlenose dolphins reached a peak in August and a nadir in November. Predicted abundance of the other whales peaked in September and reached a nadir in February and predicted abundance of loggerhead turtles reached a May and a nadir in February. Illustrations of the distribution of the densities in the peak surveyed months for bottlenose dolphins, non-baleen whales and loggerhead turtles are given in figures 4 – 6.

Table 4. Estimated minimum numbers of marine animals off Hatteras 2012.

Species	**Month***	**Mean Density (95% confidence interval)**
Bottlenose dolphin	Feb	490 (220 – 900)
	March	980 (470 – 1830)
	May	1350 (790 – 2570)
	Jun	1240 (570 – 2130)
	Aug	1430 (20 – 2510)
	Sep	970 (30 – 2770)
	Nov	400 (90 - 820)
	Assuming no spatio-temporal variation	870 (550 – 1260)
Stenella sp.	Assuming no spatio-temporal variation	1300 (500 – 2500)
Baleen whales	Assuming no spatio-temporal variation	10 (4 – 17)
Sperm whales	Assuming no spatio-temporal variation	10 (0 – 27)
Beaked and pilot whales	Feb	390 (160 – 1150)
	March	570 (160 – 1360)
	May	600 (290 – 1180)
	Jun	650 (330 – 1300)
	Aug	950 (200 – 2130)
	Sep	1110 (460 – 3030)
	Nov	520 (220 – 1290)
	Assuming no spatio-temporal variation	650 (320 – 1310)
Leatherback turtle	Assuming no spatio-temporal variation	3 (0 – 9)
Loggerhead turtle	Feb	60 (30 – 100)
	March	150 (80 – 210)
	May	190 (110 – 330)
	Jun	180 (90 – 330)
	Aug	90 (0 – 210)
	Sep	160 (60 – 340)
	Nov	130 (60 – 210)
	Assuming no spatio-temporal variation	150 (90 – 220)

*No predictions provides for January (as the only day in the month surveyed was 31st) and December (as dates surveyed were 1st and 2nd).

Figure 4. Predicted relative density (animals/km^2) surfaces for bottlenose dolphin in August. Grey dots represent the midpoints of segments and the red circles have area proportionate to the adjusted observed densities per segment area.

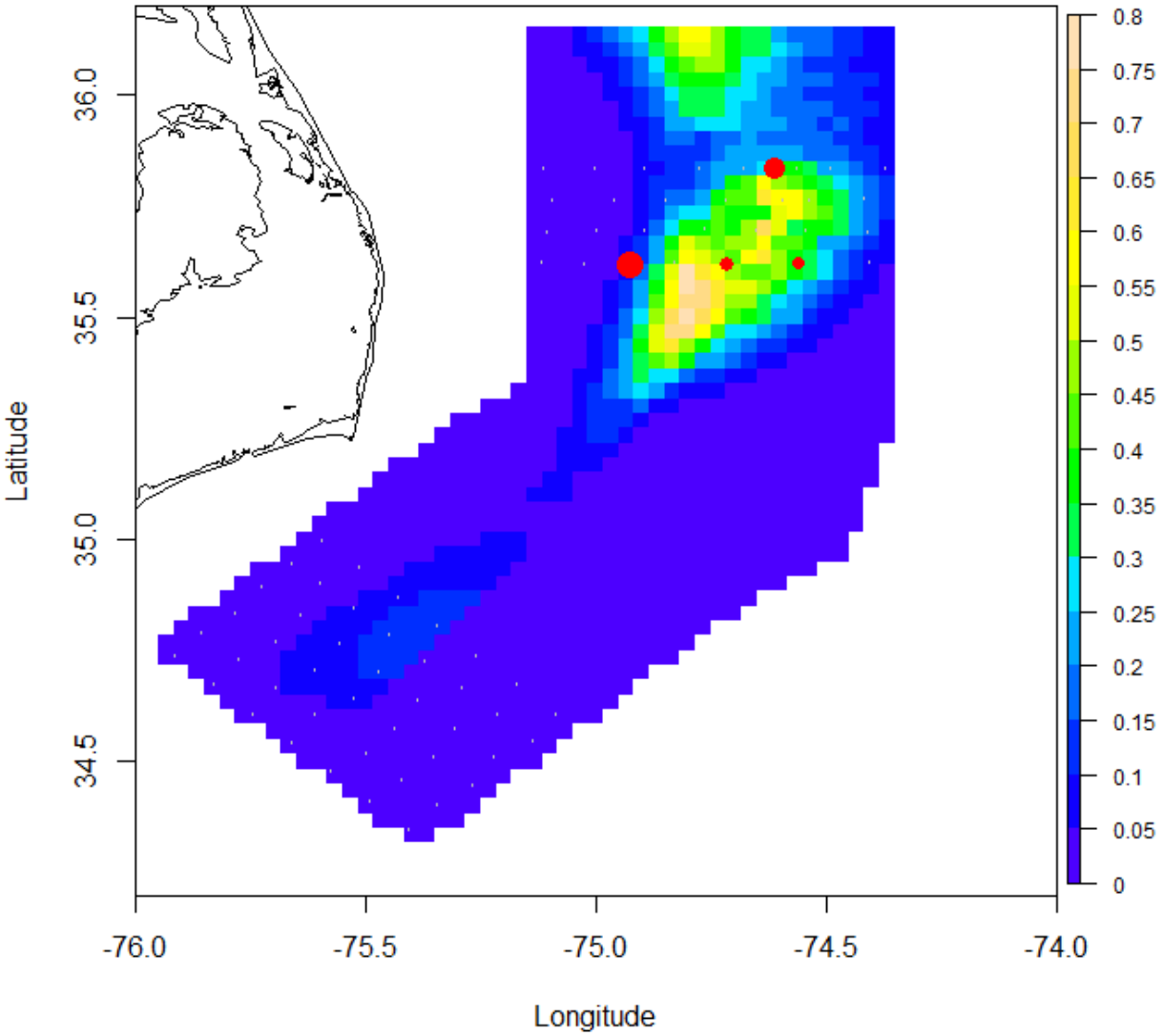


Figure 5. Predicted relative density (animals/km^2) surfaces for beaked and non-baleen whales in September. Grey dots represent the midpoints of segments and the red circles have area proportionate to the adjusted observed densities per segment area

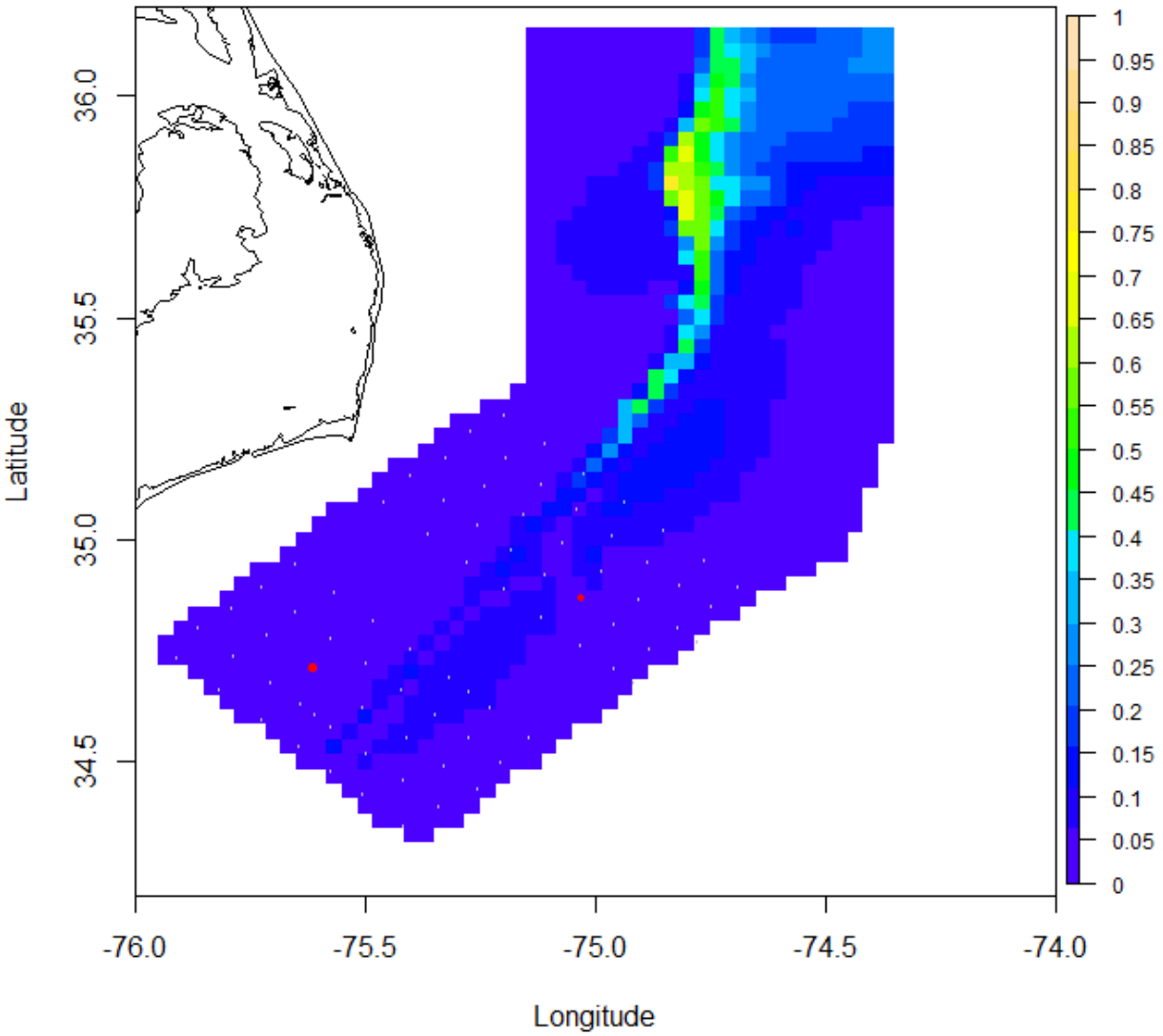
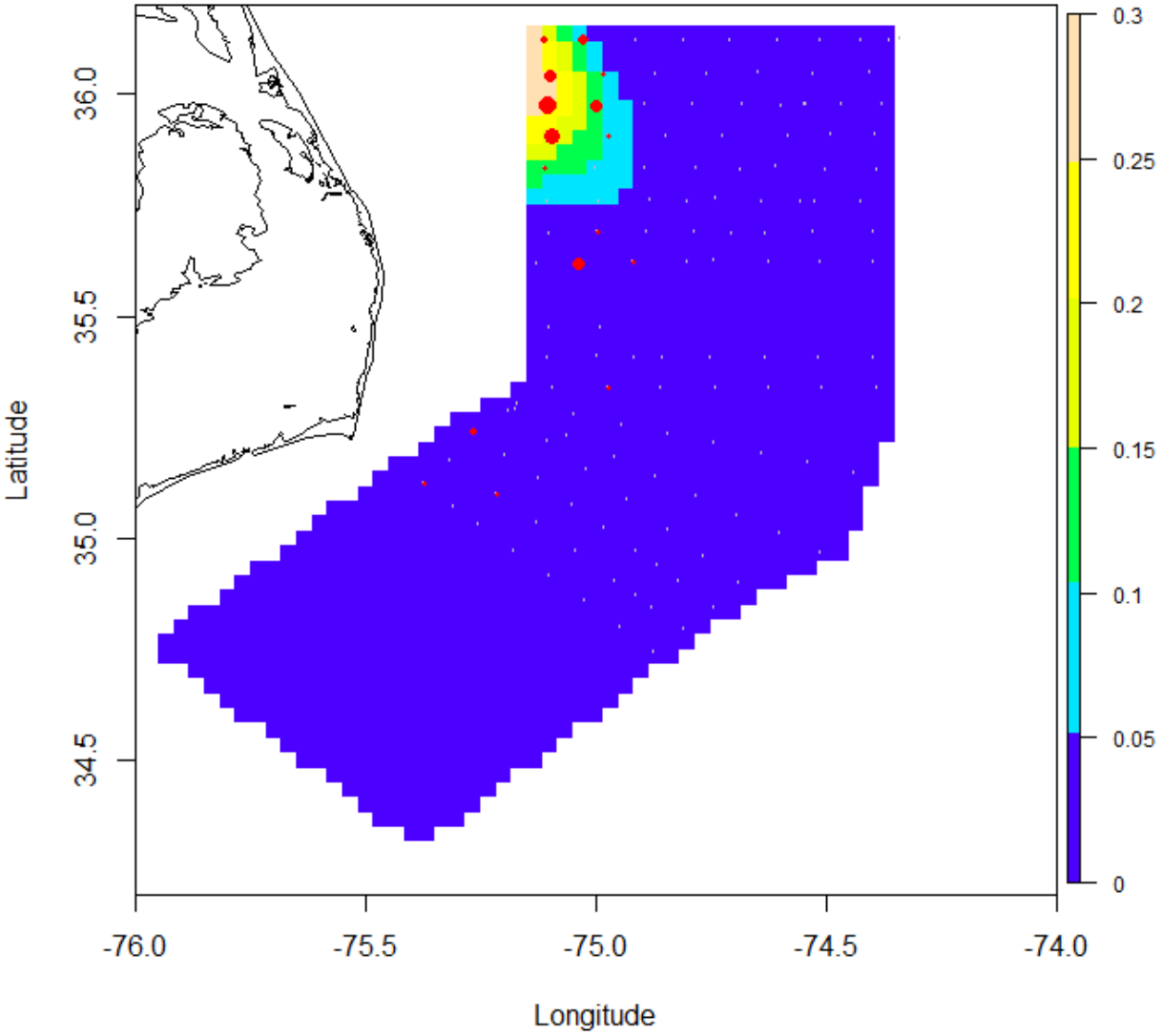


Figure 6. Predicted relative density (animals/km^2) surface for loggerhead turtles in May 2012. Grey dots represent the midpoints of segments and the red circles have area proportionate to the adjusted observed densities per segment area.

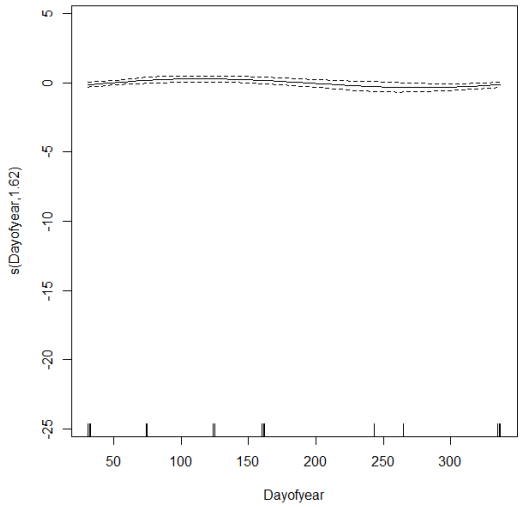


Habitat Preferences

To provide some explanation as to the habitat preferences of bottlenose dolphin and loggerhead turtles in the region, presence-absence models were fitted without latitude and longitude as potential explanatory variables (and using the generalised cross validation score as the model selection criteria). In this case, the final, fitted models, after backwards model selection, consisted of smooths of *Depth* and *Dayofyear* for bottlenose dolphin and *SST*, *Depth* and *Dayofyear* for loggerhead turtles. The response curves to the final fitted models are given in figures 7 and 8. The overriding driver of bottlenose distribution was *Depth,* as was the case for loggerhead turtles (indicated by the wide range of the scale of the *y*-axes in figures 7 and 8). Bottlenose dolphin avoided water deeper than 2,600m whereas turtles exhibited a preference for very shallow and very deep water.

Figure 7. Plot of the probability of bottlenose dolphin presence in response to a. Dayofyear b. Depth on the logistic scale.

a.



b.

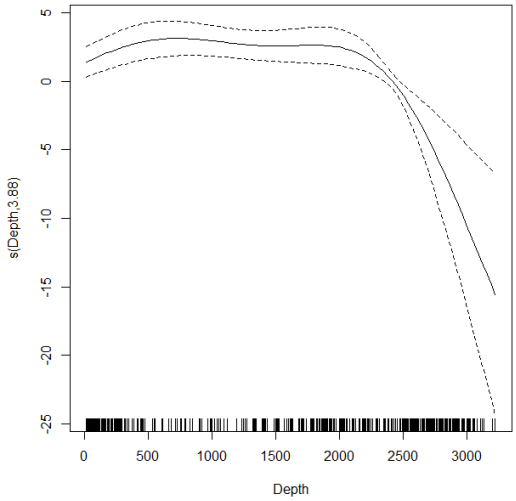
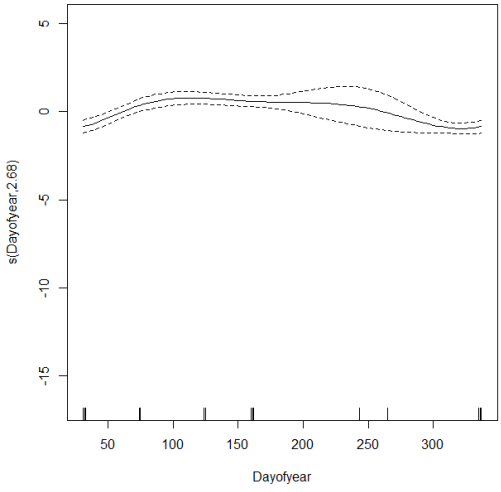
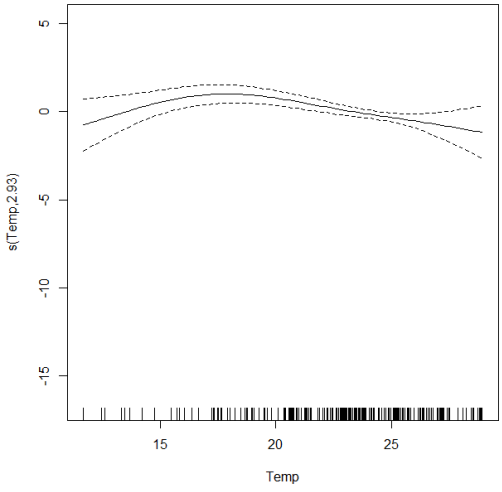


Figure 7. Plot of the probability of loggerhead turtle presence in response to a. Dayofyear b. SST and Depth and c. on the logistic scale.

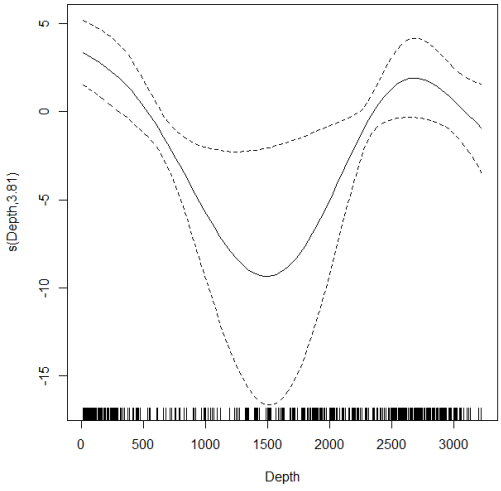
a.



b.



c.



DISCUSSION

The abundance estimates presented here should be regarded as minimum estimates because they do not include submerged animals or take into account the under estimation of animals on the trackline. Although only one year of data has been analysed, there is evidence for seasonality in the species/species groups where temporal models were able to be fitted. Nevertheless the monthly estimates should also be looked on as provisional given that they are derived from only one years' worth of data. As more data is collected from this region more patterns may become discernible.

The shape of the observed perpendicular distance distribution for turtles was not as expected (i.e. it did not decrease away from the trackline). Although distances were binned into 200m intervals and the detection function model fitted to the binned data, it is worth considering possible reasons for the strange shape. Of specific interest is how observers estimated the distances of animals that were seen a long way in front of the plane.

Monthly estimates generally agree in order with observed group encounter rates (Table 1 versus Table 4) with one notable exception: medium sized whales are predicted in great abundance in September. This appears to be because a region of high abundance is predicted where there was no effort (Figure 5).

The habitat preference curves (figures 7 and 8) do not wholly agree with the density surface maps. In May, bottlenose dolphins were predicted to occur on the continental slope whereas the presence only model predicted a more uniform distribution to a depth of 2,200m. This discrepancy arose from the addition of *Lon* and *Lat* in the presence-absence model as well as the non-zero density component, which is not included in the habitat preference modelling.

Similarly, the habitat preference model for loggerhead turtles suggested a bimodal distribution in shallow and deep water but the density surface model predicted animals in shallow waters only. Care should be taken when interpreting loggerhead turtle response to temperature as the time animals spend in the surface waters is affected by temperature (Hochscheid *et al*. 2010).

Recommendations for the future

Assuming the USWTR survey work is on-going, issues of potential interest in future work might include:

1. Investigation of the distribution of reported distances for turtles during the aerial surveys.
2. Investigation of reliable methods for estimating $g(0)$ without using double-observer survey methods. Options could include cue-based methods and use of appropriate availability correction methods based on data on availability patterns for each species.
3. Further elucidation of the environmental drivers of cetacean density in the area of interest, perhaps by the use of additional variables.
4. Records of water opaqueness may be useful to collect and, hence, include as a covariate in the detection function for turtles.
5. If availability of animals is considered important, data on diving patterns should be collected.
6. Where logistically possible, it is desirable to have effort in both the northern and southern half of the survey area each month so the models do not need to extrapolate (e.g. Figure 5).

ACKNOWLEDGEMENTS

Thanks to Louise Burt for commentating on the manuscript.

REFERENCES

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. 2001. *Introduction to distance sampling: estimating abundance of biological populations.* Oxford University Press, London. 432pp.

Hedley, S.L., Buckland, S.T. and Borchers, D.L. 1999. Spatial modelling from line transect data. *Journal of Cetacean Research and Management.* **1,** 255-264

Hedley, S.L., Buckland, S.T. and Borchers D. L. 2004. Spatial distance sampling models. In *Advanced Distance Sampling.* Buckland S.T., Anderson D.R., Burnham K.P., Laake J.L., Borchers D.L. and Thomas L. (Eds) Oxford University Press, Oxford

Hochscheid, S., Bentivegna F., Hamza A. & Hays G. C. 2010 When surfacers do not dive: multiple significance of extended surface times in marine turtles. *The Journal of Experimental Biology* **213,** 1328-1337

Horvitz, D.G. and Thompson, D.J. 1952. A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* **47,** 663 – 685

Laake, J. Borchers, D. Thomas, L., Miller D. and Bishop. J. (2013) *library mrds* available from http://cran.r-project.org/web/packages/mrds/index.html

Marques, F.F.C. 2001. *Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity.* PhD Thesis. University of St Andrews

Mullahy, J. 1986. Specification and testing of some modified count data models. *Journal of Econometrics*, **33,** 341- 365.

R Development Core Team. 2007 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Read A. J., Borchers, D.L., Cummings, E.W., McAlarney, R.J., McLellan, W.A., Nilsson, P.,Pabst, D.A., Paxton, C.G.M., Urian, K.W. and Waplpes, D.M. (submitted) Occurrence, Distribution and Abundance of Cetaceans in Onslow Bay, North Carolina, USA. *Journal of Cetacean Research and Management.*

Reynolds, R. W., T. M. Smith, C. Liu, D. B. Chelton, K. S. Casey, and M. G. Schlax, 2007: Daily high-resolution blended analyses for sea surface temperature. *Journal of Climatology*, **20,** 5473-5496.

Thomas, L., Laake, J.L., Rexstad, E., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Burt, M.L., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2009. Distance 6.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/

Wood, S.N. 2001. mgcv: GAMs and Generalized Ridge Regression for R. *R News* **1(2),** 20-2

Wood, S.N. 2006. *Generalized Additive Models. An Introduction with R*. Chapman & Hall. London