



Phase One Data Analysis of Joint Cetacean Protocol Data

C.G.M. Paxton & L. Thomas

Centre for Research into Ecological and Environmental Modelling

University of St Andrews

March 2010

Report to Joint Nature Conservation Committee

Contract number C09-0207-0216

Summary

Density surface models were fitted to combined data sets, from the Irish Sea component of the Joint Cetacean Protocol (JCP) data resource, by generalising available line transect sightings data to data that did not include distances to obtain estimates of density. The procedure was as follows. Detection functions were fitted to all available survey sightings data where records were kept of the perpendicular distance from observer to sighting. Thus probabilities of detection could be assigned to all sightings including those without distances (matched for survey and other covariates where necessary). Estimated densities could then be calculated for a given segment of survey effort. The resultant density surfaces were then modelled in a two stage process by first modelling presence-absence followed by non-zero density.

Density surfaces varying in time could be successfully predicted for harbour porpoise, minke whale, bottlenose dolphin, common dolphin and Risso's dolphin. Nevertheless, the density surface proved difficult to model and bootstrap estimates of variance, derived from a nonparametric bootstrap, were very wide. There was evidence for an upward trend in harbour porpoise, bottlenose dolphin and common dolphin numbers in the last 3 decades but this result may be an artefact of evolving methodologies since the 1980s.

A power analysis showed that, for the above three species, quite small declines in modelled population density (0.3-2.2% per year) over a 6-year reporting period could be detected with power of 0.8, for the latter part of the survey. For other species and earlier time periods, only very large changes in modelled population density would be detectable. However, the modelled population densities rely on spatial and temporal smoothing, and hence sudden declines would not necessarily be detectable; also the method includes variability due only to observation error and ignores process error (random fluctuations in animal numbers from a smooth trend line). Lastly, the results are based on spatio-temporal models that may themselves not be reliable.

This analysis shows that there is potential to produce valid inferences by combining databases but there are problems to overcome. We provide recommendations for how JCP data should be collated in future. We also recommend further effort be devoted to improving the modelling of Irish Sea data in preference to analysis of further datasets.

Contents

Summary	2
Contents	3
Introduction	4
Methods	5
Overview of statistical methods	5
The data in space and time	6
Description of the datasets	9
Treatment of sightings	21
Modelling densities	25
Prediction of density	26
Sources of variation.....	27
Power to detect trends.....	27
Results.....	28
Detection function modelling.....	28
Density surface modelling.....	32
Discussion	53
Conclusions and recommendations	56
Acknowledgements.....	57
References	57
Appendix 1. Data review	59
Appendix 2. Suggested details for JCP data	61
Appendix 3. Zero inflated models.....	63

Introduction

EU Member States have a legal obligation under Article 11 of the Habitats Directive to undertake surveillance of all cetacean species occurring in their waters to determine their “conservation status”, and to report on this every 6 years. A species is in “favourable” conservation status if: “population dynamics data indicate that the species is maintaining itself on a long-term basis as a viable component of its natural habitats, the natural range of the species is neither being reduced nor is likely to be reduced in the foreseeable future, and there is, and will probably continue to be, a sufficiently large habitat to maintain its populations on a long-term basis.” The exact measures reported are open to interpretation by Member States (see European Commission 2006), but the above guidance leads naturally to a focus on (1) trends in species’ abundance; (2) changes in species’ range; (3) designation and monitoring of suitable habitat (albeit not practical for cetaceans); and (4) future prospects.

In the UK and Ireland, one cost-effective method for potentially addressing the first two of the above measures for marine mammals is via the Joint Cetacean Database and its prospective successor, the Joint Cetacean Protocol (JCP). The JCP is essentially a collection of effort-related survey data that have been gathered by various governmental and non-governmental organizations from dedicated surveys as well as systematic observing from platforms of opportunity and more casual watches.

Thomas (2009) reviewed the potential of the JCP to allow estimation of trends in species abundance and pointed out that whilst the threshold for detectable population change recommended by the habitats directive (1%) was considered unrealistic there was, at least, the potential to extract robust relative trend data from the JCP. He reviewed a number of potential statistical methods that could be used to investigate the data and made recommendations for future research.

The aim here was to progress the analysis of Thomas (2009) using his proposed third method to combine data from the myriad of JCP datasets. The chosen method takes line transect data from designed and platforms of opportunity surveys where distance sampling data were captured and applies the detection probabilities from these analysis to sightings data where distance sampling data were not present. Once sightings data have been corrected, a spatio-temporal density surface is fit to the data. Such a density surface could allow documentation of trends, analysis of habitat preferences and estimation of population abundance.

Thus, here we consider data collected in the Irish Sea from a variety of surveys and determine whether it is possible to obtain reasonable and meaningful estimates of abundance and trends for marine mammal species. The species under consideration were harbour porpoise *Phocoena phocoena*, fin whale *Balaenoptera physalus*, minke whale *Balaenoptera acuturostrata*, bottlenose dolphin *Tursiops truncatus*, common dolphin *Delphinus delphis*, white-beaked dolphin *Lagenorhynchus albirostris* and white-sided dolphin *Lagenorhynchus acutus*.

Methods

Overview of statistical methods

Because of the convoluted methods used in obtaining the estimates, here we present an overview of the statistical methods and general approach used prior to a detailed formal description in later sections (see Figure 1). The data under consideration (see below) consisted of spatially and temporally referenced sightings of marine mammals. These sightings may or may not be associated with observed distances of an observer to the group of animals seen. There are two general approaches to dealing with such effort associated sighting data, only one of which, “the count method” (Hedley 2000, Hedley and Buckland 2004, Hedley *et al.* 2004), is considered here (but see Thomas 2009 for a review).

The analysis of such data can be considered as a two-stage process, with the aim of estimating the overall density of animals in the region of interest based on the density estimates within the area covered by survey effort. The data do not have to come from a dedicated line transect survey although there should be systematic observation associated with defined effort (ie both location and whether observers were truly observing should be known as opposed to casual ad-hoc observations). Data without distances can be incorporated into the analysis. As detectability by observers is imperfect, correction should be made where possible for missed animals. This can be done in part using distance sampling methods (see Buckland *et al.* 2001, 2007) that can infer detectability by considering the distribution of distances to observed animals.

Once detectability is estimated, then a corrected estimate of the numbers of animals seen per unit length of effort can be made. If the probability of detection on the trackline $g(0)$ can be estimated then this means an absolute index of the abundance of animals at the surface can be made. For some of the data considered here $g(0)$ estimates were available so absolute abundance of surface animals could indeed be estimated.

The survey effort is divided into segments and the numbers observed, corrected for detectability, can be summed for each segment. This creates a spatially referenced density index that can then be modelled as the second stage of the analysis. It is possible that this density surface modelling stage may itself have two stages (see below). This modelling process allows testing of habitat variables that may influence cetacean distributions as well as direct estimate of trends (if any). The resultant density model can also be used to interpolate into regions and times when no observations were made, creating an estimated density surface by prediction over a grid of spatially (and temporally) referenced points.

Geostatistical analyses of the residuals from the models sometimes highlighted that adjacent segments were not truly independent and the statistical methods used assumed independence of the individual data. In addition, handling of some of these large datasets can be difficult; however techniques for utilising large datasets such as these (40,000+ data records) are becoming available (see below).

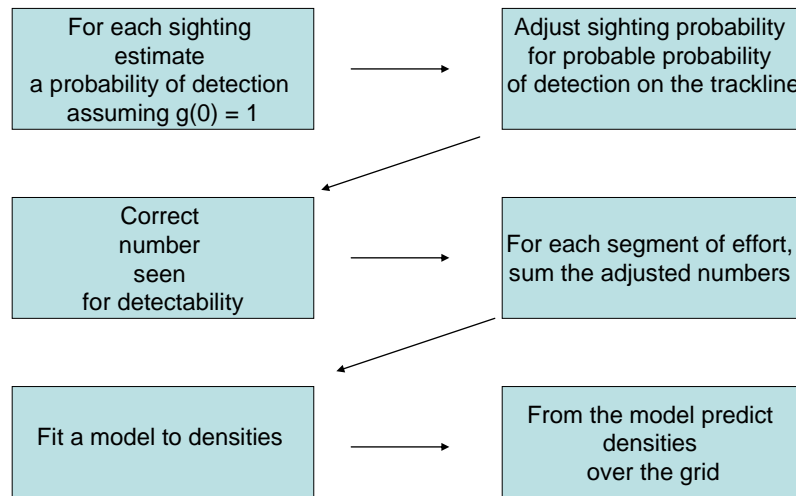


Figure 1. A schematic of the analysis of the JCP data.

The data in space and time

The region of interest, which is shown by the shaded area in Figure 2a, is relatively shallow with no depths greater than 300 m. The available data span from 1980 to 2009 and survey effort are represented in Figure 2b (by year) and 2c (by month). There is greater effort in more recent years and, unsurprisingly, during the summer.

There is confounding of location (and hence depth) and survey vessel type (Figure 2d) with ferries traversing the same region repeatedly. The inshore region of the Welsh coast is dominated by small boat surveys (blue) and aerial surveys (black), the offshore areas by the ESAS surveys by larger boats (red). There is also coverage from ferries (green).

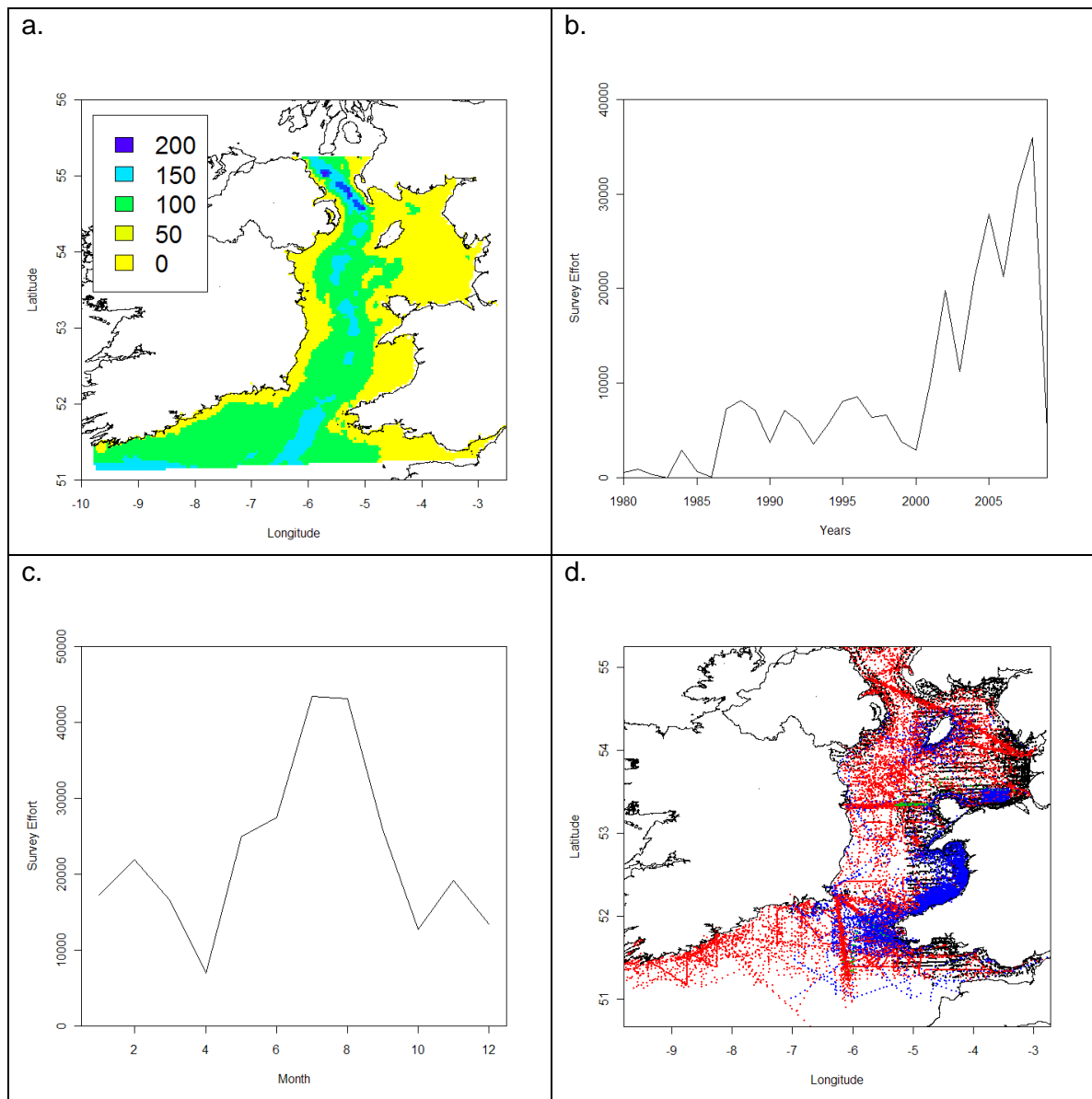


Figure 2. Aspects of the Irish Sea JCP. (a) Bathymetry of the Irish Sea (m). (b) Realized (ie utilized in this analysis) effort by year across all surveys. (c) Realized effort by month across all surveys. (d) Spatial distribution of effort by vessel type (see text for details).

A variety of data were potentially available (see Appendix 1), however, only a subset of these were amenable to analysis within the time frame of the study. These are described below and illustrated by year (colour codes consistently the same for all years).

Details of all data available and the decision made for their inclusion in the analysis are given in Appendix 1. However, to summarise, only vessel based data were used as land-based data were spatially restricted and there was limited time to complete the analysis with the extra complications such data present. All vessel sightings required effort (ie times and locations of observing) to make them usable in the analysis. In each case, the sightings data were assigned

to one of three classes: sightings with distances from two independent observers allowing estimation of $g(0)$, the probability of detection on the track line ultimately allowing estimation of absolute surface density; sightings with distances from a single observation platform allowing estimation of a relative surface density; sightings with no distances.

Data with noticeably erroneous positions that could not be easily corrected (ie by reference to adjacent segments), were removed. Similarly, segments (lengths of transited effort) associated with calculated boat speeds in excess of 50km/h were also removed, as this unworkable speed must be caused by erroneous recorded position or time data.

Data were grouped by vessel type. Data could be aerial or boat based. If the latter there were three classes of boat: *littleboats* (observer eye height ≤ 5 m above the water level, *bigboats* (observer eye height between 5 and 10m above water level) and *ferries* (observer eye height > 10 m above the water).

Description of the datasets

Cardigan Bay Marine Wildlife Survey Boat Data

These data consists of shipboard line transect surveys and systematic search effort data from 2005 to 2007 (Figure 3) using the small boat *Sulaire*. The sightings data were typically of sightings with distances ie were class b) data.

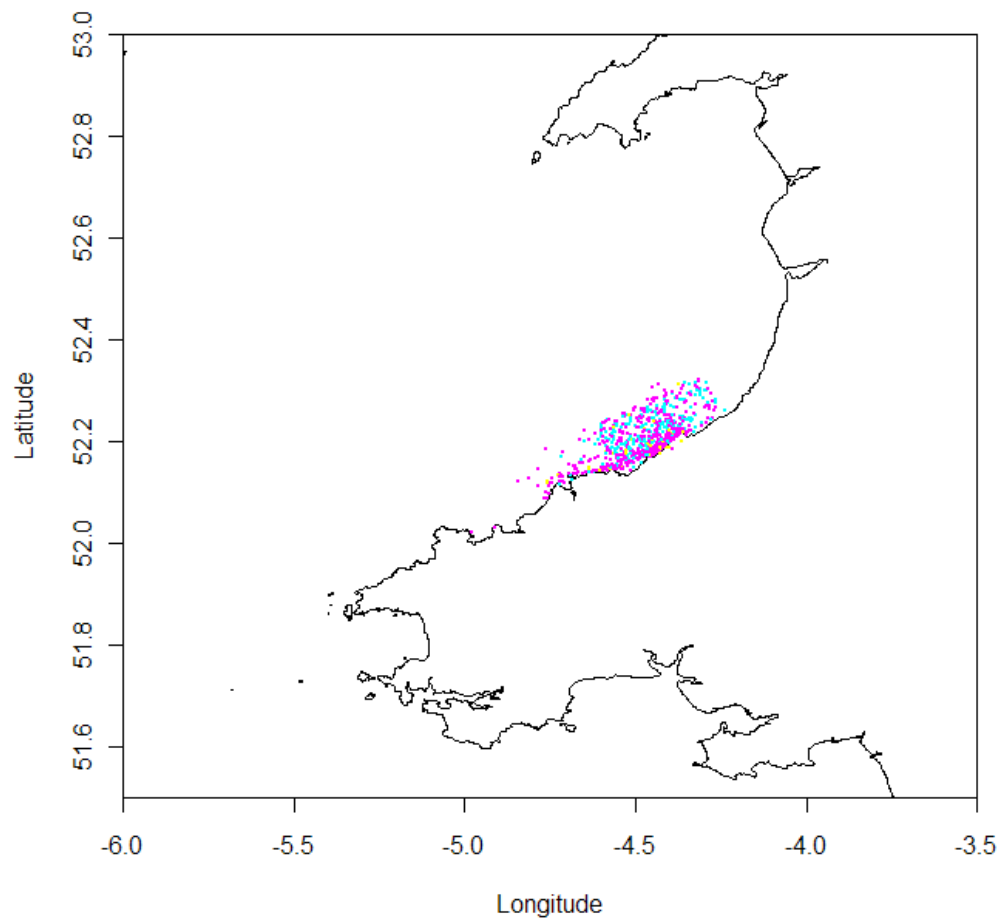


Figure 3. Realized (ie post data filtering) shipboard survey data from CBMW surveys. Colour coded by year, cyan 2005, magenta 2006 and yellow 2007. Each point represents the midpoint of a survey segment.

Cardigan Bay Sea Watch Foundation Surveys 2008

These were small boat surveys carried out in 2008 when distance data were sometimes collected. Some of the data were collected using two independent observers (class a data) and hence could be used to estimate probability of detection on the trackline (see below).

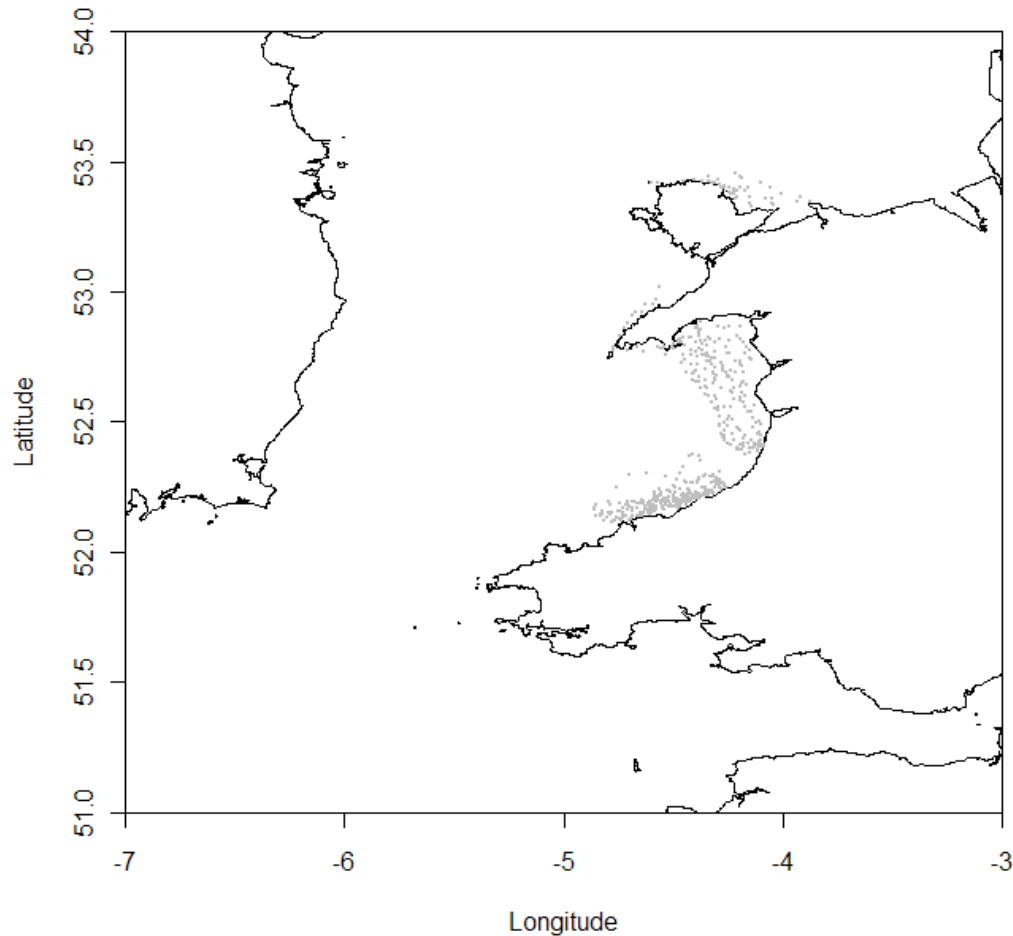


Figure 4. Realized shipboard survey data from the Cardigan Bay SWF surveys 2008. Each point represents the midpoint of a survey segment.

Countryside Council for Wales Surveys

Aerial surveys carried out in the southern part of the Irish Sea for marine megafauna (Houghton *et al.* 2006). Distances were estimated (as inclinations from the plane) during this survey (class b data).

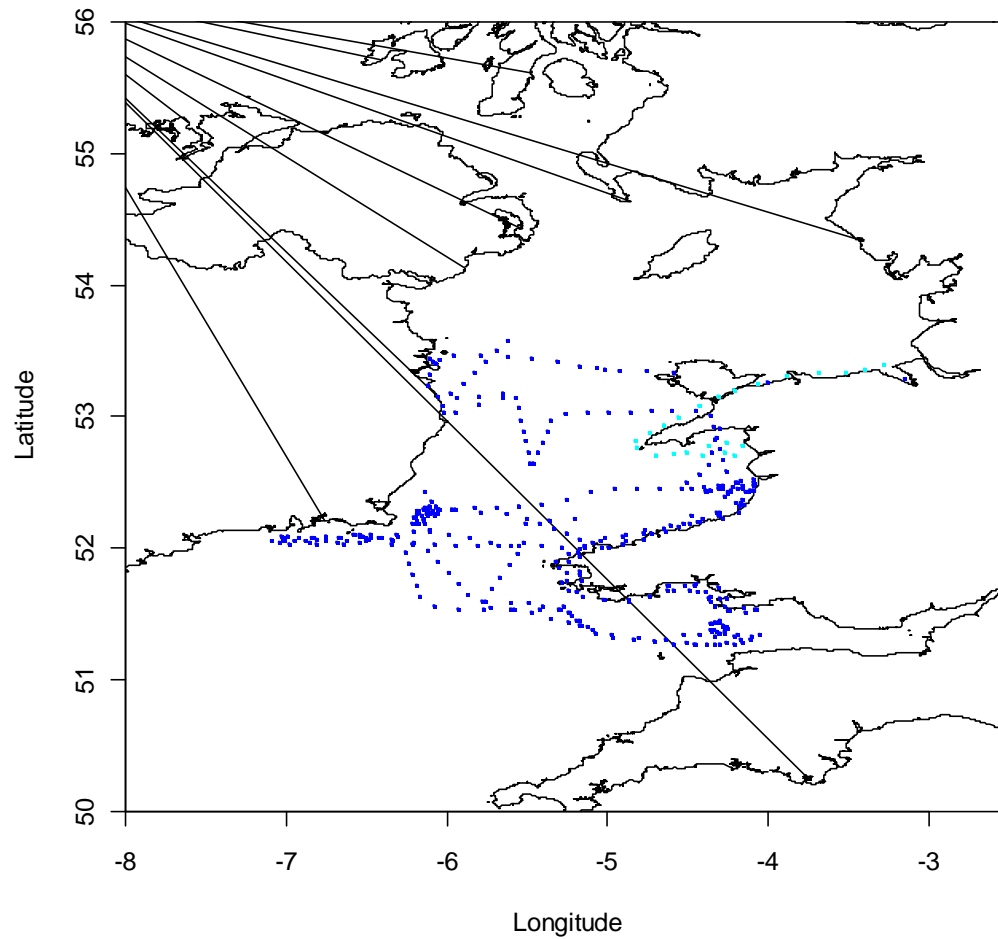


Figure 5. Realized aerial survey data from CCW surveys. Each point represents the midpoint of a survey segment blue 2004, pale blue 2005.

Department of Energy and Climate Change Surveys (DECC) by the Wildfowl and Wetlands Trust (WWT)

These extensive aerial surveys were undertaken by the WWT on behalf of the DECC from 2001 to 2008. Sightings distances were binned into distance intervals (class b data).

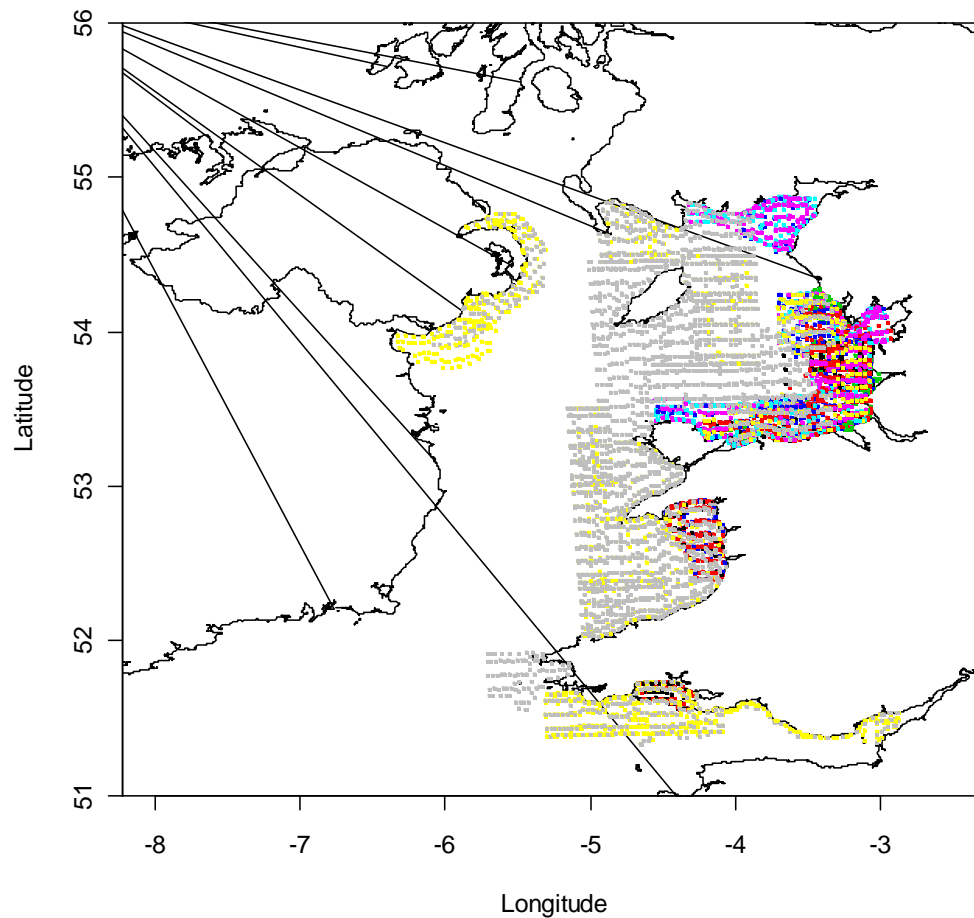


Figure 6. Realized aerial survey data from the DECC surveys. Each point represents the midpoint of a survey segment: grey 2000, black 2001, red 2002, green 2003, blue 2004, cyan 2005, magenta 2006, yellow 2007 and grey 2008.

European Seabirds at Sea (ESAS) Surveys

These data consist of ship and aerial surveys from a variety of platforms. They date back to 1980, further than any other utilised data source. Distances were mostly recorded, in bands, for sightings within 300m of the trackline (class b data) although sometimes sightings were just known to be within 300m horizontal distance of the plane but the exact distance was not known. ESAS North Sea harbour porpoise sightings data were previously analysed by Winship (2009), however, data analysed here were not treated as per this analysis in a number of respects. His altering speed criteria for identifying periods of non-effort was not implemented and vessel type was included in the analysis (see below). However, like Winship (2009), effort associated with surface speeds in excess of 50 km/h was not included.

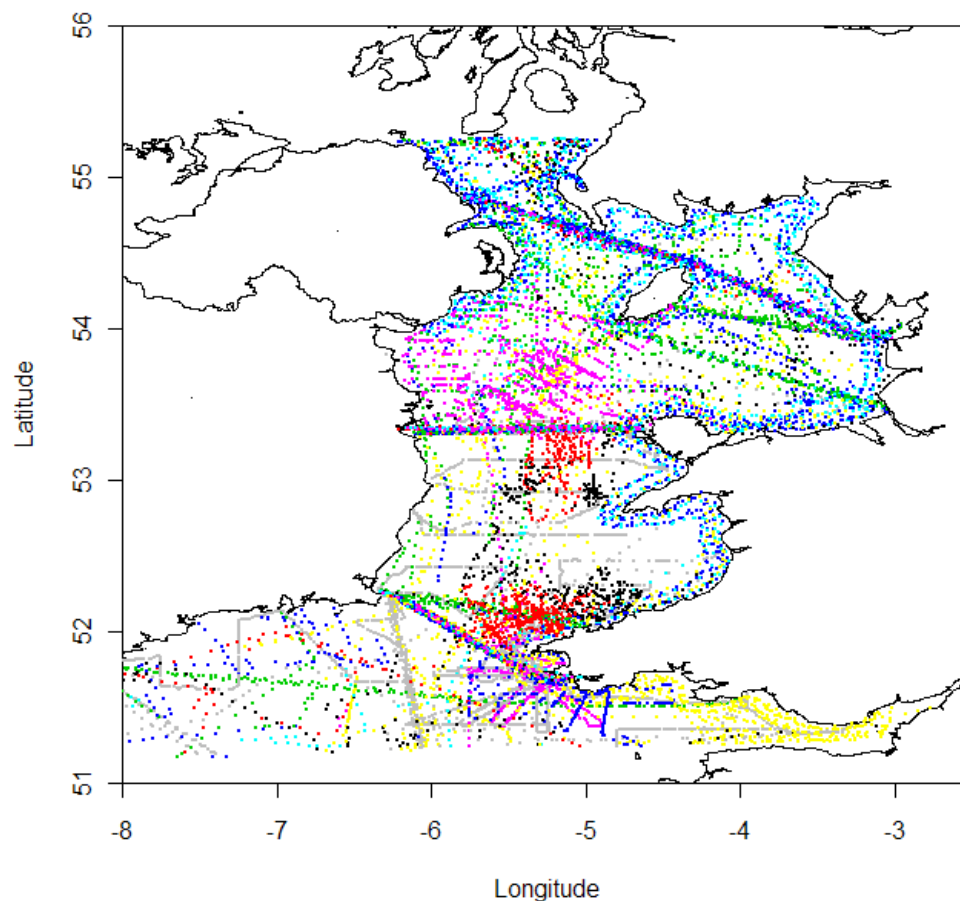


Figure 7. Midpoints of realized effort segments from ESAS: grey 1980, cyan 1981, magenta 1982, yellow 1983, grey 1984, black 1985, red 1986, green 1987, blue 1988, cyan 1989, magenta 1990, yellow 1991, grey 1992, black 1993, red 1994, green 1995, blue 1996, cyan 1997, magenta 1998, yellow 1999, grey 2000, black 2001, red, 2002.

Irish whale and Dolphin Group (IWDG) Dedicated Surveys and Ferry Observations

These data consist of sightings and efforts from a variety of surveys conducted from dedicated survey vessels as well as platforms of opportunity. Almost all of the sightings data had distance associated with them (class b data).

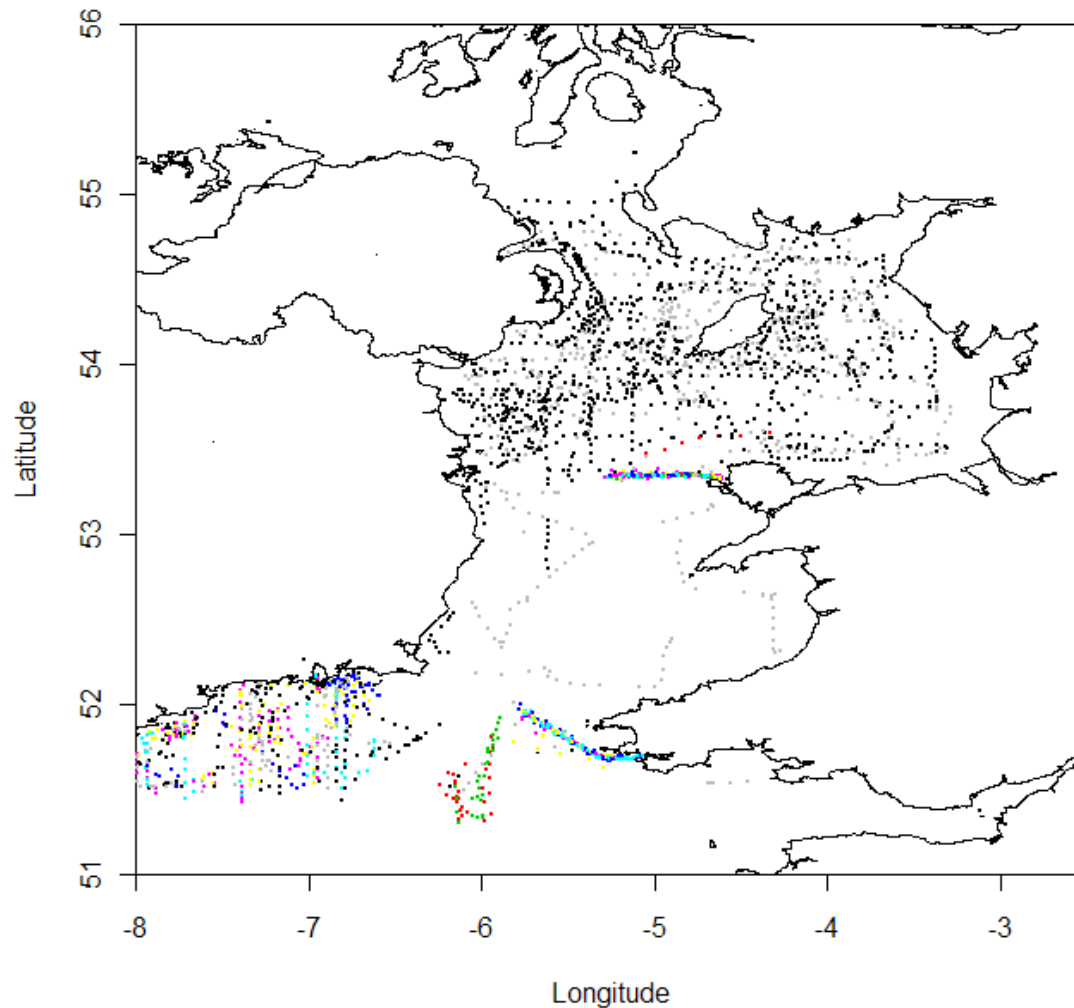


Figure 8. Realized shipboard survey data from IWDG surveys: blue 2004, pale blue 2005, magenta 2006, yellow 2007, grey 2008 and black 2009.

North Wales Windfarm Environment Assessment Surveys

These consisted of visual survey data from vessels in 2003 and 2004 associated with wind farm development (class b data)

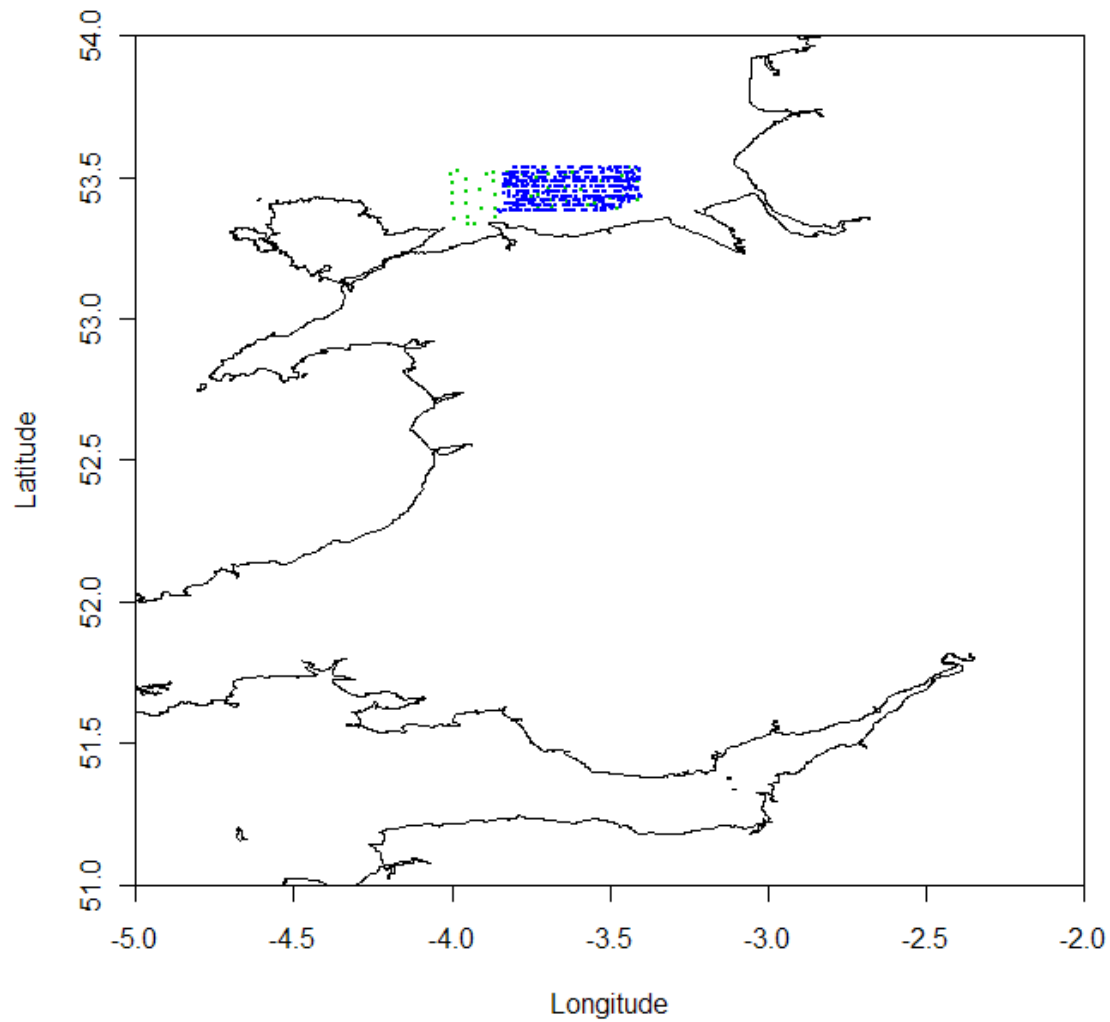


Figure 9. Midpoints of realized effort segments from the north Wales coast wind farm surveys: green 2003, blue 2004.

Manx Whale and Dolphin Surveys

These data came from small boat surveys undertaken in 2007. Not all the sightings had associated distances although mostly the data is class b.

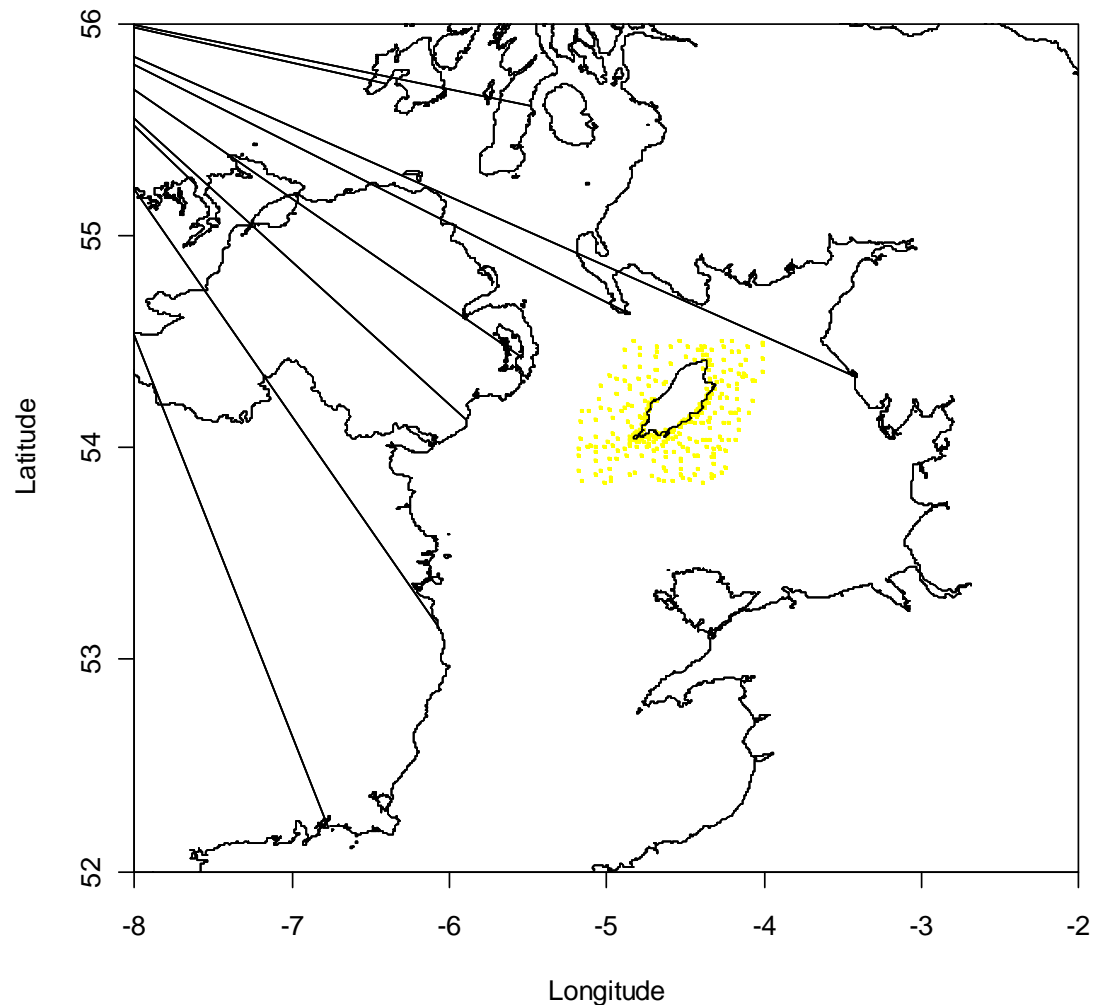


Figure 10. Midpoints of realized effort segments from the Manx Whale and Dolphin Watch surveys 2007.

Marine Awareness North Whales Surveys

Data from small boat surveys from 2002 to 2008 with sightings distances recorded (class b data).

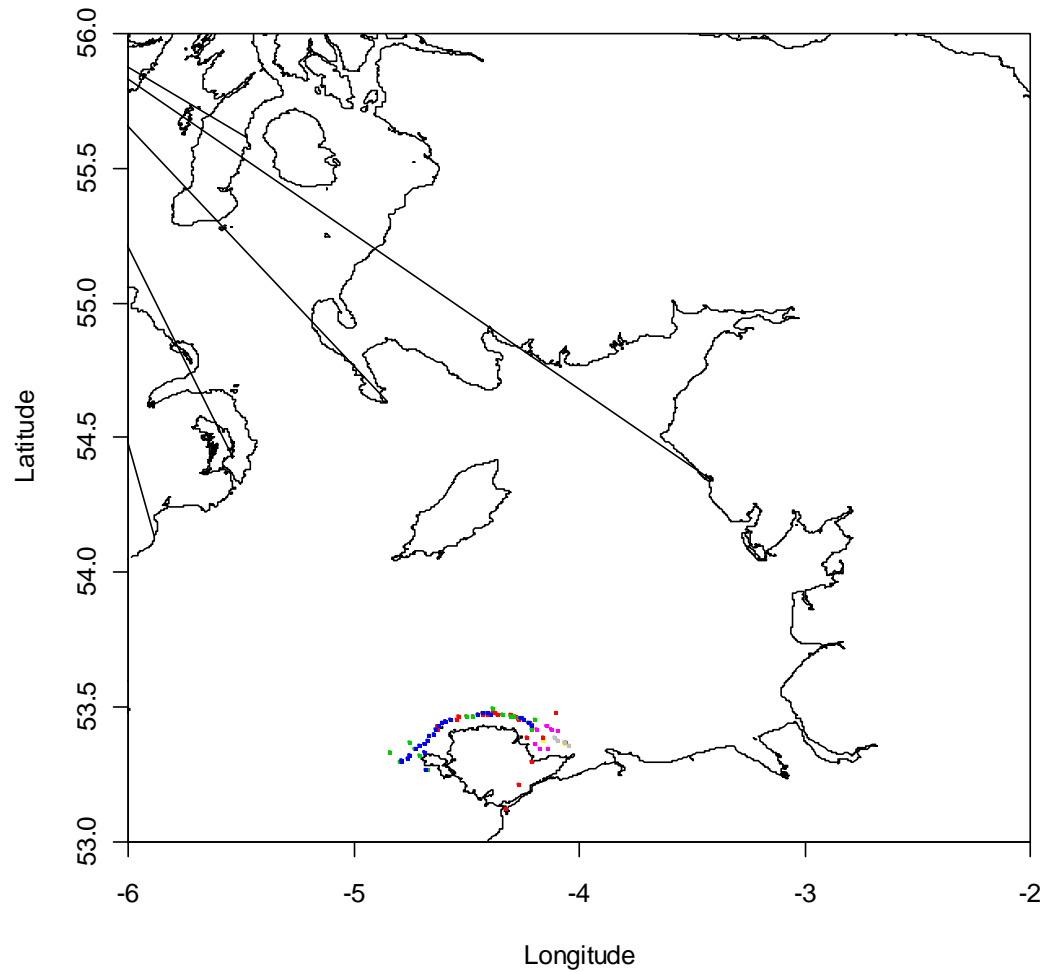


Figure 11. Midpoints of realized effort segments from the MANW surveys from 2007.

Pembrokeshire Porpoise Surveys

Small boat surveys from 2007 and 2008 with distances collected (class b data).

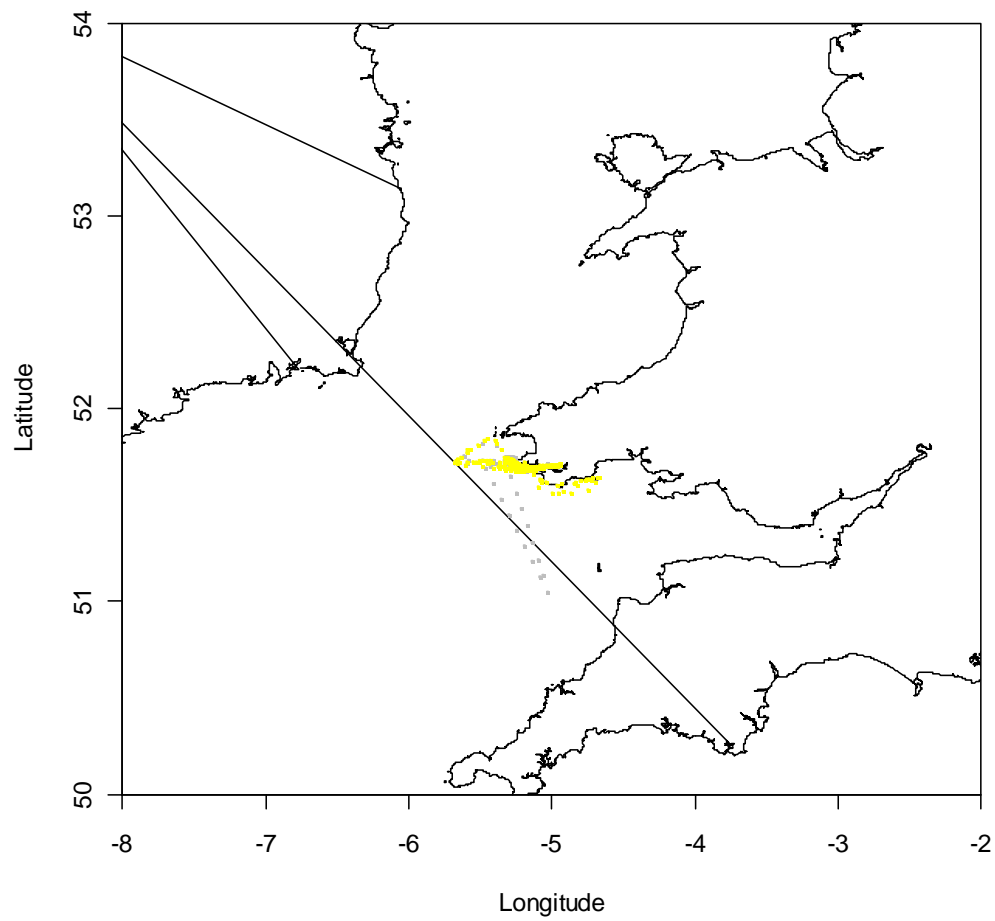


Figure 12. Midpoints of realized effort segments from the Pembrokeshire porpoise surveys: yellow 2007 and grey 2008.

Sea Watch Foundation (SWF)

These data came from a multitude of surveys and systematic observation from a variety of platforms (small boats to ferry sized ships, see below). For all of the sightings used from this source, there were no associated distances but there was a record of effort (class c data).

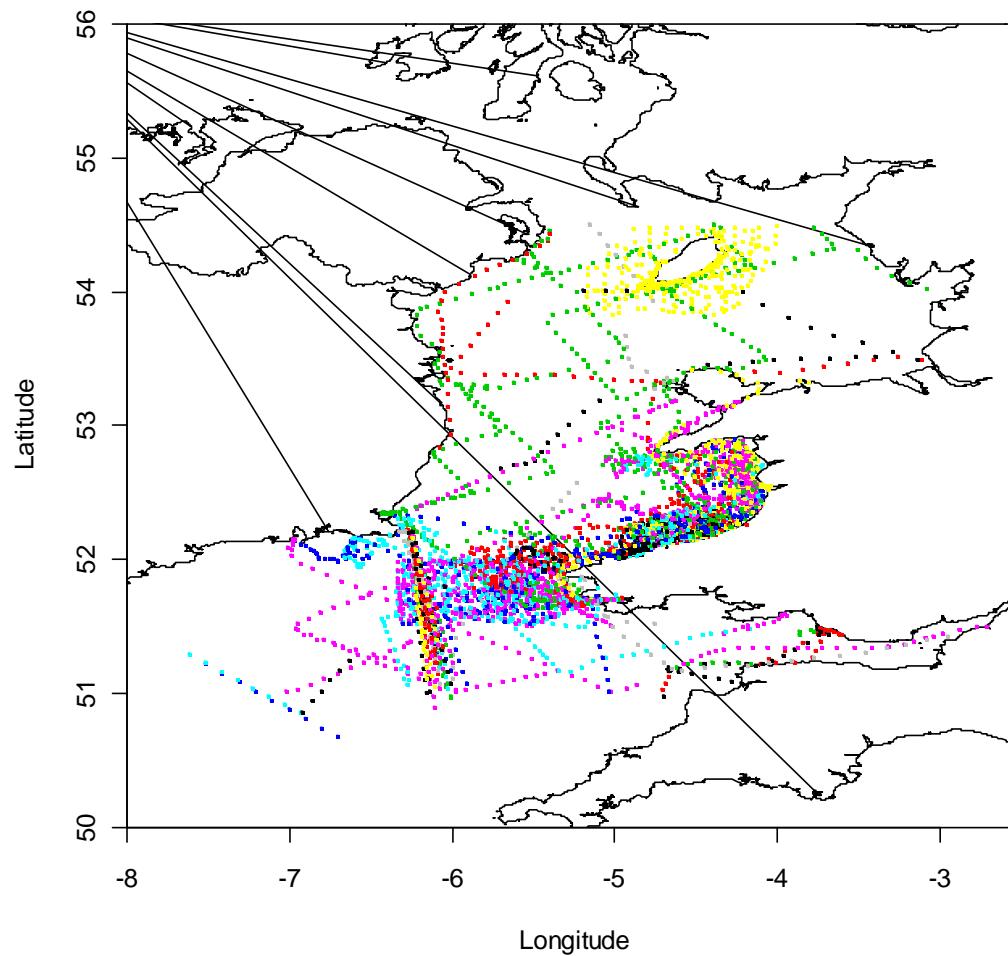


Figure 14. Midpoints of realized effort segments from SWF. Magenta 1990, yellow 1991, grey 1992, black 1993, red 1994, green 1995, blue 1996, cyan 1997, magenta 1998, yellow 1999, grey 2000, black 2001, red, 2002, green 2003, blue 2004, cyan 2005, magenta 2006, yellow 2007, grey 2008 and black 2009.

SCANS I and SCANS II

Simultaneous surveys of cetaceans undertaken in 1994 and 2005 (SCANS-II 2008). In 1994, there was only coverage of the southern region of interest but in 2005 the central region of the Irish Sea was also covered. Observations came from both boats and planes. In both cases, observations were undertaken with observers in double observer trial mode (Laake and Borchers 2004) allowing estimation of $g(0)$, the probability of detection on the trackline (class a data).

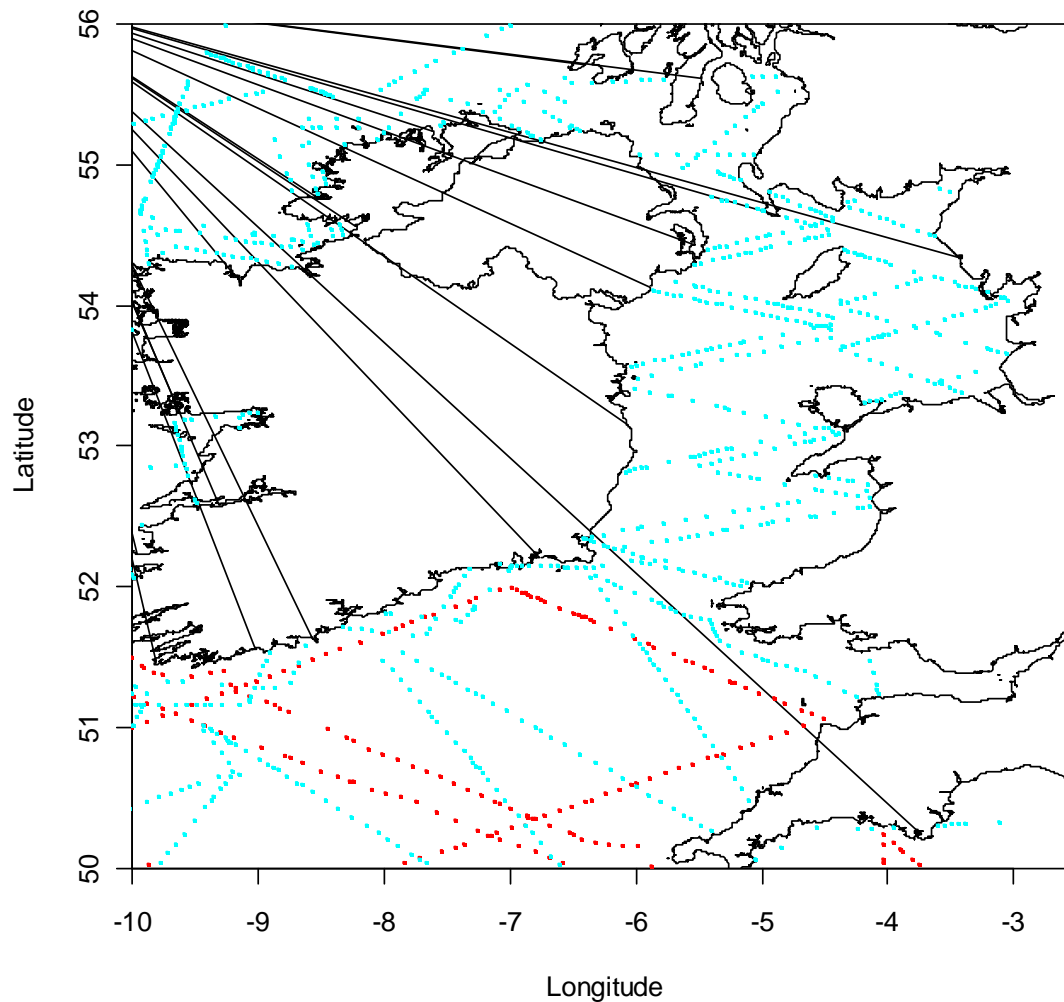


Figure 15. Midpoints of realized effort segments from SCANS I (1994 red) and SCANS II (2005 pale blue).

Treatment of sightings

Surveys were carried out using a variety of platforms, from rigid inflatables to ferries and planes. Sightings data that had the same truncation distance were grouped together ie ESAS data (truncation distance 300m) was considered separately to other surveys. Therefore, distinct detection functions were created for ships, ESAS boats and planes, other planes and SCANS ships and planes respectively.

For sightings associated with distances, a detection function could be applied. In this analysis, the effects of covariates, other than perpendicular distance, were incorporated into the detection function model. Thus, the probability of detection becomes a multivariate function, $g(y, \mathbf{v})$, representing the probability of detection at perpendicular distance y and covariates \mathbf{v} ($\mathbf{v} = v_1, \dots, v_Q$ where Q is the number of covariates). Using either a hazard-rate ($1 - \exp(-y/\sigma)^b$) or half-normal detection function ($\exp(-y^2/2\sigma^2)$) (Buckland *et al.* 2001), the covariates were incorporated via the scale term, σ , where for sighting j , σ has the form:

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

where b , β_0 and β_q ($q=1, \dots, Q$) are parameters to be estimated (Marques & Buckland 2007). With this formulation, it is assumed that the covariates may affect the rate at which detection probability decreases as a function of distance, but not the shape of the detection function.

Potential covariates were Beaufort sea state, animal group size and *VesselType*. For boats observer eye height and other related variables were integrated into the *VesselType* variable, which discriminated between *littleboats* (observer eye height ≤ 5 m above the water level), *bigboats* (observer eye height between 5 and 10m above water level) and *ferries* (observer eye height > 10 m above the water). Note that observer eye height was not always available for all platforms so heights, and hence *VesselType*, had to be assumed. Number of observers was not considered as an independent variable (and was often not available), but was correlated with the size of the boat so was partially accounted for by *VesselType*. A stepwise forward selection procedure was used (starting with a model containing perpendicular distance only) to decide which covariates (Beaufort Sea State, Species, group size and *VesselType*) to include in the model, with a minimum Akaike's Information Criterion (AIC) inclusion criterion.

The software *Distance* (Thomas *et al.* 2010) version 6.0 release 2 was mostly used for model selection of detection functions, using the mrds engine (Burt *et al.* in prep.). Further analysis was performed using the statistical software R (R Development Core Team, 2009) version 2.10.1.

Separate detection functions were generated for harbour porpoise, balaenopterid whales and generic dolphins. In the case of harbour porpoise only sightings and effort of sea state 2 or less were considered because of the notoriously low detectability of harbour porpoise at high sea states (eg Northridge *et al.* 1995, Laake *et al.* 1997). In the case of balaenopterid whales and generic dolphins sightings and effort data associated with sea state 4 or less were considered. Only animal numbers and sea state were considered as covariates in the detection functions. Other variables were available for specific surveys, but sea state was the only universal detection related variable noted by all observers. In the few cases where sea state was missing for a sighting, it was assumed by looking at adjacent effort segments. Other variables that could be used in the future (albeit without the possibility of comparison with all previous surveys) could

include glare, visibility, observer ID (and experience, experienced observers can increase their probability of detecting harbour porpoise, Laake *et al.* 1997)

ESAS

ESAS data were collected either by boat or plane. Sightings were allocated to one of three distance bins up to 300m away from the boat. Therefore, the data were treated as single platform distance data with a 300m truncation. Some data were not binned into distance categories but were coded as within 300m. All aerial data were of the latter form. There was no realized helicopter effort. As there were no sightings associated with the ESAS plane effort data, they were treated as the ship effort data (except for *SurveyMode* in the density surface modelling, see below).

CCW and DECC

These aerial data were considered together as single platform data. Note that $g(0)$ for the aerial surveys was not known because no double platform data were collected and the relevant $g(0)$ s could not be extracted from the SCANS data because different survey methods were used. However, $g(0)$ for the rapidly transiting aerial surveys can reasonably be assumed to be lower than equivalent boat surveys. Therefore boat surveys observe more animals per unit effort than aerial surveys. Therefore if mode of survey is used in the density surface modelling then prediction assuming a boat survey will correct for underestimation in the aerial survey.

SCANS I & SCANS II

Detection functions for the SCANS survey data were not actually calculated within this analysis. Instead, we made use of previous work, where the estimated true numbers of animals per segment, corrected for detectability, had been calculated. We did, in addition, use estimates of $g(0)$ from the SCANS surveys as our estimate of $g(0)$ for other *bigboats* surveys – ie the $g(0)$ for *ferries* was assumed to be same as the *bigboats*. SCANS data from beyond the study region were retained in the data to constrain the periphery of the prediction area in the bootstrap (see below). SCANS I data were available for minke whale and harbour porpoise and SCANS II data were available for minke whale, harbour porpoise and common dolphin.

Other boat data

The remainder of the sightings were single platform boat data from all the other surveys, with the exception of a small amount of double platform independent observer mode data collected in 2008 in Cardigan Bay.

This survey had independent observers and thus allowed the opportunity to estimate $g(0)$, the probability of detection on the trackline, for the *littleboats* used in the survey.

If no cetacean movement is assumed, point-independence – detections are assumed to be independent on the trackline only (Laake and Borchers 2004) – offers a more robust approach to dealing with double observer data than the assumption of full-independence. This would seem a reasonable assumption in the case of harbour porpoise but less so in the case of dolphins, which are known to associate with boats.

The single platform version of the above double platform data (ie groups observed at least once were assumed seen) was amalgamated with the remaining single platform boat data to estimate a single platform detection function for the vast majority of reported sightings.

Surveys without distances

Surveys without distances to sightings were allocated a probability of detection dependent upon covariates other than distance. This assumes that these surveys have the same detection probability, given covariates, as those with distances. A proportion of sightings with missing distances were truncated, this proportion being the same as the proportion in the data with distances. This is because the detection probability calculated from surveys with distance data is the average probability of detecting a pod between 0 and the truncation distance. An alternative would be to work with the effective strip width (detection probability times truncation distance) and estimate what the effective strip width would be given no truncation; hence removing the need to truncate data.

From sightings to densities

The effort data were available either as waypoints or as segments of transect lines. In the former case, the data were segmented using the way points as end points. Segments were then split or amalgamated into target lengths of 10km segments as much as possible. To be amalgamated into a 10km segment the existing segments had to be adjacent in space and time and have identical sea states. Therefore the target 10km segment lengths were not always achieved. Segments above 15km in length were split.

Table 1 gives the realized effort (ie effort after removal of unusable effort in km for each data set used). The estimated number of individuals in each segment, \hat{N}_i , was calculated using an estimator similar to the Horvitz-Thompson estimator (Horvitz and Thompson, 1952), ie:

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{s_{ij}}{\int_0^w \hat{g}(y, v_{ij}) \pi(y) dy}, \quad i = 1, \dots, T, \quad (1)$$

where, for each segment i containing at least one sighting, $\int_0^w \hat{g}(y, v_{ij}) \pi(y) dy$ is the estimated probability of detection of the j th detected pod (from a line transect analysis), n_i is the number of detected pods in the segment and s_{ij} is the size of the j th pod. By assumption, $\pi(y)$, the probability density function of perpendicular distances of detected and undetected animals, is uniform. Sightings were allocated to these segments by reference to their time of observation.

Table 1. Realized annual effort from each survey (sea state <= 4)

Dataset	Seg. length (km)	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
CBMWC	3.47																4094	5347	753	
Ceredigion Bay (SWF)	4.94																			2164
CCW	9.02															3676	251			
DECC	7.15												3671	13361	3720	10290	17356	8498	16672	27527
ESAS (boat)	4.91	2310	2816	4893	2765	3907	5205	4609	2504	3462	451	448	950	19						
ESAS (plane)	6.6		3949			231														
IWDG survey boat	5.40															379	376	509	515	3247
IWDG Ferry Data	9.67												365	398	981	1349	1303	794	624	170
JNCC Survey	5.7														269	2388				
Isle of Man Surveys	6.6																		2180	
MANW Surveys	9.80														396	199	281	87	37	49
Pembrokeshire Surveys	4.89																	2576	1916	
Sea Watch Foundation	5.14	1366	514	1026	598	1475	2827	3960	3848	3152	3323	2395	5317	5596	6077	2693	4490	6048	7343	
SCANS 1 (ship)	7.64					1840														
SCANS2 (aerial)	6.15																4501			
SCANS2 (ship)	4.80																2680			

* ESAS ship surveys also took place from 1980 to 1989: 1980 558 km, 1981 890 km, 1982 265 km, 1983 2869 km, 1984 0 km, 1985 632 km, 1986 61 km 1987 7206 km, 1988 3100 km & 1989 1256 km. Additional aerial surveys took place in 1987: 59 km 1988 5000 km. and 1989 5529km. An IWDG survey took place in 2009 and covered 4925 km. There was no aerial Irish Sea effort in SCANS I

Modelling densities

Having obtained the estimated number of individuals in each i th segment \hat{N}_i , the estimated density, \hat{D}_i , is simply given by \hat{N}_i / a_i where a_i is the surveyed area (twice the truncation width, see below) multiplied by the length of segment i . The total number of segments varied dependent upon species because of sea state and whether all of the available SCANS data were used.

A variety of methods were considered, trialled and implemented to model the density data. All of them were in a multiple regression framework with the variation in segmented density dependent upon a variety of independent potential predictor variables. One initial problem was the large size of the data sets, >40,000 observations. A single over-dispersed model of animal abundance using a generalized additive model (GAM) for each species was considered, with quasi-Poisson or negative binomial error families utilised. This was applied to a subset of the data that consisted of *adjacent* segments. It was clear that even the negative binomial models failed to deal with the over-dispersion in the data. For some species variograms of the residuals from the model strongly suggested there was spatial correlation in the data (even after the inclusion of spatially referenced explanatory variables).

All the data sets had large numbers of zero segments, presumably representing both genuine permanent absences and transient absences or false absences caused by non-detection of animals. The distributional properties of such datasets are such that they cannot be readily modelled by conventional means.

A two-stage modelling approach for this survey, as implemented by Borchers *et al.*, (1997) was adopted. In the two-stage approach, the presence or absence of animals in a segment is first modelled using a logistic regression on, where possible, the full data set, and then the estimated number of whales in the non-zero segments is modelled separately, in this case assuming gamma-distributed errors. This can introduce a bias and some alternative methods were investigated (see Appendix 3) but could not be fully implemented to obtain variance estimates. Thus in the first stage of the two stage modelling process, the Bernoulli data were modelled as:

$$E\left(\log\left(\frac{p_i}{1-p_i}\right)\right) = \theta_0 + \sum_k q_k(z_{ik}), \quad i = 1, \dots, T \quad (2)$$

where p_i is the probability that a whale is present in the i^{th} segment, θ_0 is the intercept parameter, the q_k are smooth functions of the k spatial covariates, z , and T is the total number of segments.

Equation (2) modelled the probability of presence. To obtain an estimate of cetacean abundance, an additional model was required which was fitted only to data where sightings were detected. In this step, an estimate of the (non-zero) segment density was required. To model this, a zero-truncated distribution was required and a number of models were considered. Because the gamma distribution is a continuous distribution that admits only positive values, it is the natural family to try to model the non-zero distribution of marine mammals.

Therefore non-zero pod density \hat{D}_i (across segments with sightings only) was then modelled as a function of the k covariates, z , using the gamma GAM with a log link function:

$$E(\log(\hat{D})_i \mid \hat{D}_i > 0) = \left(\beta_0 + \sum_k r_k(z_{ik}) \right), \quad i = 1, \dots, T \quad (3)$$

where β_0 is the intercept parameter and the r_k are smooth functions of the k spatial covariates. There is risk of a bias here as segments associated with a low density of animals may be unrepresented in the data as zero segments may actually contain animals (see below).

Covariates and factors considered in both the logistic and non-zero density models were longitude, *Lon*, latitude *Lat*, day of year, *Dayofyear*, *SurveyMode* (ie aerial or boat), *Year* of survey and *Depth*. Depth for each segment was obtained either from depth data collected on survey or from the ETOPO2 2 minute resolution relief data available from National Oceanic And Atmospheric Administration (at <http://www.ngdc.noaa.gov/mgg/global/etopo2.html>). Depths were selected for each survey segment from one of either depth database based on the nearest great circle distance. *Dayofyear* was fitted as a cyclic smoother as the effect at the beginning and end of the year should be similar. *SurveyMode* (ie whether the survey was aerial or ship) was only included in the model if the estimated coefficient was positive (reflecting the higher $g(0)$ of a boat compared to a plane) – recall that this was used to correct aerial surveys for the fact that no $g(0)$ estimate was available for them, and that all density surface prediction (below) took place using the ship-based level of *SurveyMode*.

All covariates were considered for inclusion in the model as 1D smooths (thin plate regression splines) of the untransformed covariate values. In addition, 2D smooths (thin-plate splines; Wood, 2006) of *Lat* and *Lon*, suitably transformed into nautical miles from the equator and the 4° west meridian, were considered for inclusion into the GAM. *Lon* was further transformed, by dividing by 2, to give it the same range as *Lat* (an alternative would have been to use tensor product splines, which can deal with variables that have different scales, but these were found to be prohibitively slow given the large data size). Initially, a maximum of six degrees of freedom (seven knots) was allowed in the selection of 1D smooths and up to 13 degrees of freedom (14 knots) were allowed in the case of 2D smooths of *Lon* and *Lat*. *Lon* and *Lat* were also considered in 2 and 3D smooths with year (ie spatial position alters by year). Non spatial variables were allowed up to 4 degrees of freedom. Thus allowing moderate flexibility but reducing the likelihood of spurious fitting of unnecessarily complicated functions. In addition the cost associated with fitting each degree of freedom was increased to 1.4 to minimise the risk of overfitting (Wood 2006, Kim and Gu 2004). In the case of the logistic models the unbiased risk estimator (UBRE) was used to select models. In the case of the Gamma model, generalised cross validation (GCV) implemented in the *mgcv* package (Wood 2006) in *R* was used for covariate selection, augmented with diagnostic plots, using the principles described in Wood (2006) to minimise the GCV.

Models were fitted in a dedicated piece of software in the statistical computing software package *R* (R Developmental Core Team 2009) using general additive models (GAMs) from the *mgcv* library (see Wood 2006 for a description). The data sets were too large for standard fitting methods to be successful, so a large sample GAM fitting algorithm was used (function *bam()* of *mgcv* 1.6-1, Wood 2009).

Prediction of density

Predictions from the models were made on a 2 by 2 minute resolution latitude-longitude grid covering the coloured area in Figure 2a. The prediction was always made assuming a

SurveyMode of boat. This was because $g(0)$ was known for these surveys (see above) but not known for the aerial surveys but could reasonably be assumed to be less than the boat $g(0)$ because of the faster transit time of the plane surveys. If the models contained date/time variables, the predictions were date/time specific. Predictions from the individual grid cells were summed to obtain overall estimates. In the case of the two stage modelling, the sum per cell was the product of the probability of presence and the predicted non-zero density in the cell as well as the area of the cell. Maps were made as well as an annual index for each year 1980 – 2008.

Sources of variation

To incorporate all the uncertainty associated with the detection function fitting, the whole process of Figure 1 was bootstrapped. That is effort was sampled (with replacement) by day and pseudosamples created upon which the entire estimation process was rerun. For the SCANS data, where previous analyses were re-used, bootstrapping was by transect within survey block. Confidence intervals were then obtained by the 2.5% and 97.5% quantiles of the final density estimates. Model selection uncertainty was not incorporated into the bootstrap with the exception of automatic degree of freedom selection by the GAM. Because on some occasions there was evidence of spatial correlation, the full data sets were sub-sampled within each bootstrap by taking every n th point to ensure independence, where n was chosen by examining variogram plots to determine where spatial correlation was near to zero. The sub-sampling also had the effect of reducing the overall size of the data set for the spatio-temporal modelling, albeit with a loss of information (see Discussion). Note that all of the available data was used in detection function modelling, and for the point estimates of spatio-temporal pattern (with the exception of common dolphin, see Results).

Power to detect trends

The aim of this analysis was to make a preliminary determination of what level of population trend would be observable with reasonable certainty, given the levels of variability about the trend estimates observed in the above studies. It is first necessary to define what is meant by “trend”, since there is no objective definition (Thomas *et al.* 2004). A common definition used in power analysis studies (eg, Gerrodette 1987) is the slope of a log-linear regression, since this provides a convenient one-number summary of the pattern, is relatively easy to obtain through linear regression, and corresponds with an exponential model of population growth. However, for longer time series, assuming that population trend is linear (on some scale) is often unrealistic, and indeed in the analyses above, a flexible, smooth trend model was used.

Because the expected reporting period for trends is 6 years, it makes sense to use as a metric the ratio of the smoothed density estimate in the year of interest (eg, the current year) divided by that from 6 years previously:

$$\Delta^* = \hat{D}_t / \hat{D}_{t-6} \quad (3)$$

A value of 2, for example, indicates a population doubling over that period, while a value of 0.5 indicates a population halving and 1 indicates no change. One disadvantage of such a measure is that its size is in some way related to the number of years in the time interval; hence it may be better scaled into a measure of average annual change:

$$\Delta = \hat{D}_t / \hat{D}_{t-6} \quad (4)$$

Here, a value of 1.05 would be interpreted as the population growing by an average of 5% per year over the period, while a value of 0.95 indicates a decline of 5% per year on average and 1.0 indicates no change. Note that this measure can also be calculated by taking the geometric mean of the annual population changes.

Because Δ is the ratio of two zero-bounded random quantities, its distribution is expected to be approximately log-normal. Hence, a simple test for trend is a one-sample, two-sided z-test of the null hypothesis that the natural log of Δ is zero (ie that Δ is 1.0). Given an estimate of the variance in $\log(\Delta)$ and the α -level (here assumed to be 0.05) then it is straightforward to calculate the power of the test for various levels of Δ that are considered biologically relevant (the relevant formulae are given in Steidl and Thomas, 2001). Alternatively, given a desired target level of power, the detectable Δ can be calculated. In this report, the latter approach was taken. For the purposes of illustration, the value of population change Δ detectable with a power of 0.8 was calculated for each species and for each pair of years (t , $t-6$) from $t=1986$ -2008.

The required input, variance in $\log(\Delta)$ was calculated from the bootstrap resamples. Rather than reporting variance in $\log(\Delta)$, which is not easy to interpret, we transformed this into the coefficient of variation of average annual population change, $CV(\Delta)$. They are related as follows:

$$CV(\Delta) = \sqrt{\exp(\text{Var}(\log(\Delta)) - 1)} \quad (5)$$

Note that the nonparametric bootstrap introduces variability due to observation error (uncertainty in the estimates of numbers of animals in each year), but not process error (uncertainty in the actual number of animals in each year) – in other words, with this procedure animal numbers are treated as a fixed, not random quantity (see Thomas *et al.* 2004).

Results

Detection function modelling

Harbour porpoise *Phocoena phocoena*

The double platform analysis of Cardigan Bay boat data from 2008 contained 75 sightings of which 20 were duplicates. The best estimate of $g(0)$ was 0.638 (SE=0.053). The data were truncated at 800m, as was the equivalent augmented boat data ($n = 1894$, Figure 16a). No covariates in addition to distance were selected in the single platform analysis perhaps because only sea state 2 or less data was considered. The single platform hazard-rate detection function for the non-ESAS boat sightings ($n= 194$) is given in Figure 16b. The ESAS detection function was modelled as a half-normal function over the range 0 – 300 m as was the DECC and CCW combined aerial detection function (Figure 16 c) for the 1155 aerial sightings of harbour porpoise. No covariates in addition to distance were selected.

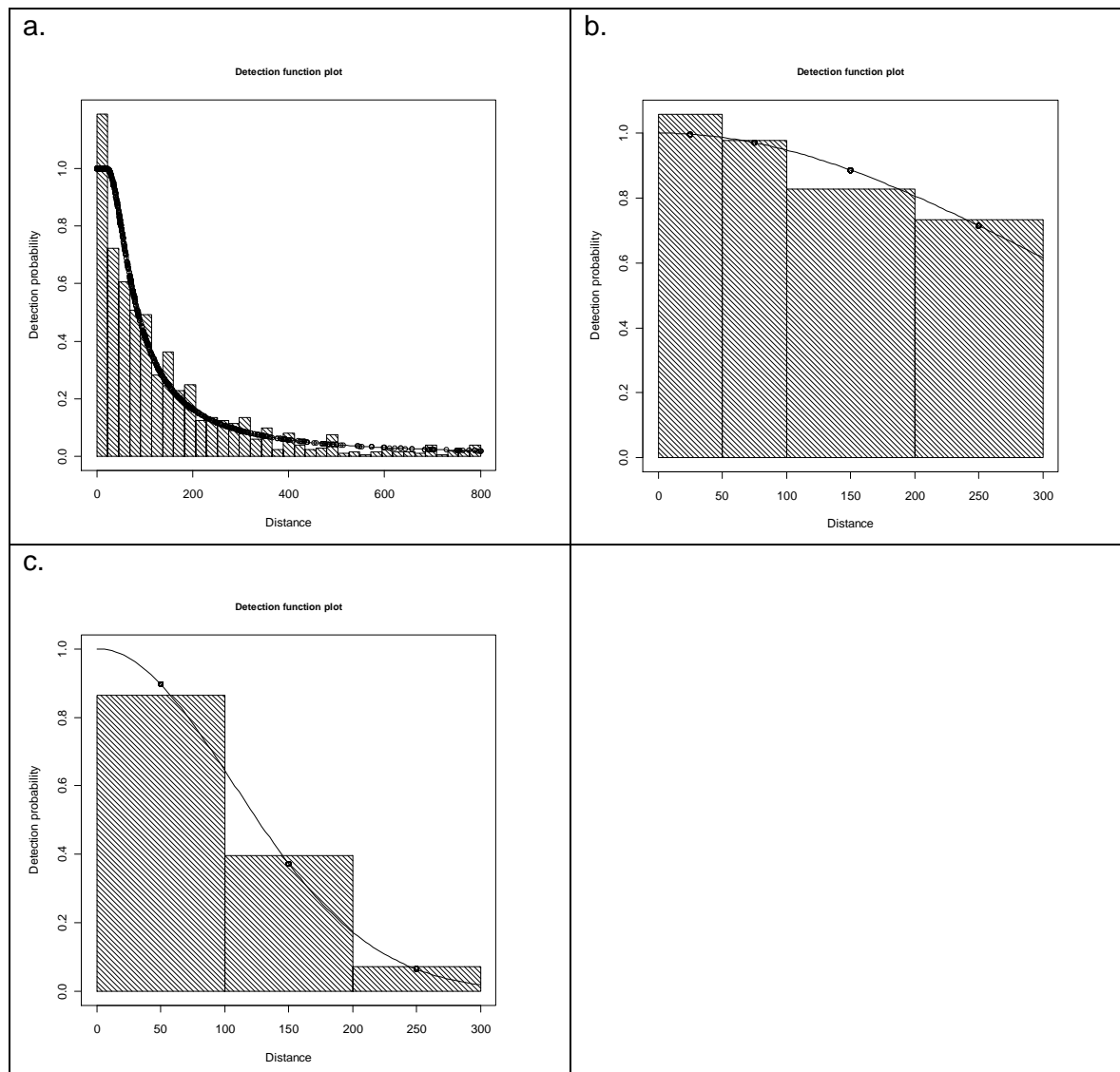


Figure 16. Detection functions for harbour porpoise a. non-ESAS boat sightings, b. ESAS boat sightings, c. plane sightings

Balaenopterid whales

The detection functions for balaenopterid whales are shown in Figure 17. No $g(0)$ for small boats was estimable from the 2008 double platform SWF survey as no whales were seen. So $g(0)$ was taken from the SCANS survey (and is likely an over-estimate). Fifty-five sightings were made by non-ESAS boats to generate the single platform detection function (Figure 17a). No covariates in addition to distance were selected. Only 5 large whales were seen in the aerial surveys, so no detection function was fitted to the data, instead the sole minke whale aerial sighting was assumed to have a probability of detection of one and be associated with a truncation distance of 0.8 km. Only 18 whales were seen by ESAS boats (Figure 17b). No covariates in addition to distance were selected.

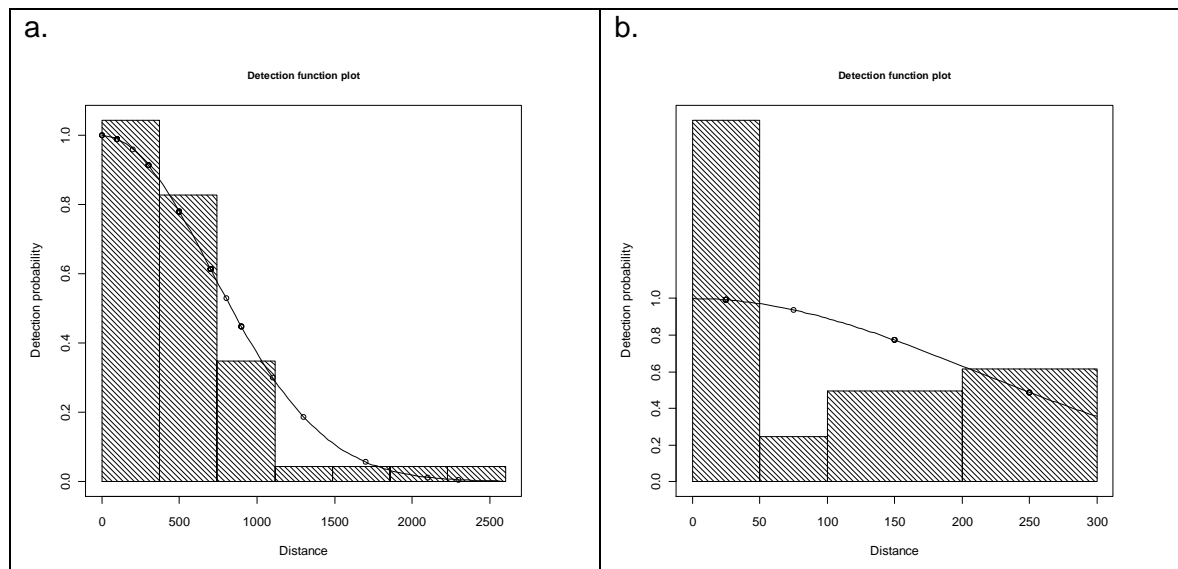


Figure 17. Detection functions for balaenopterid whales a. non-ESAS boat sightings, b. ESAS boat sightings,

Dolphins

The detection functions for dolphins are shown in Figure 18. In the case of non-ESAS boats the estimate of $g(0)$ was 0.782 (SE 0.106) based on 11 trials in which 4 dolphin groups were seen by both observers. Figure 18 gives the detection functions for the single platform data. There were 794 non-ESAS boat sightings, 156 ESAS sightings and 429 plane sightings. In the case of the boat data *Vessel/Type* was a significant covariate: larger boats saw more dolphins.

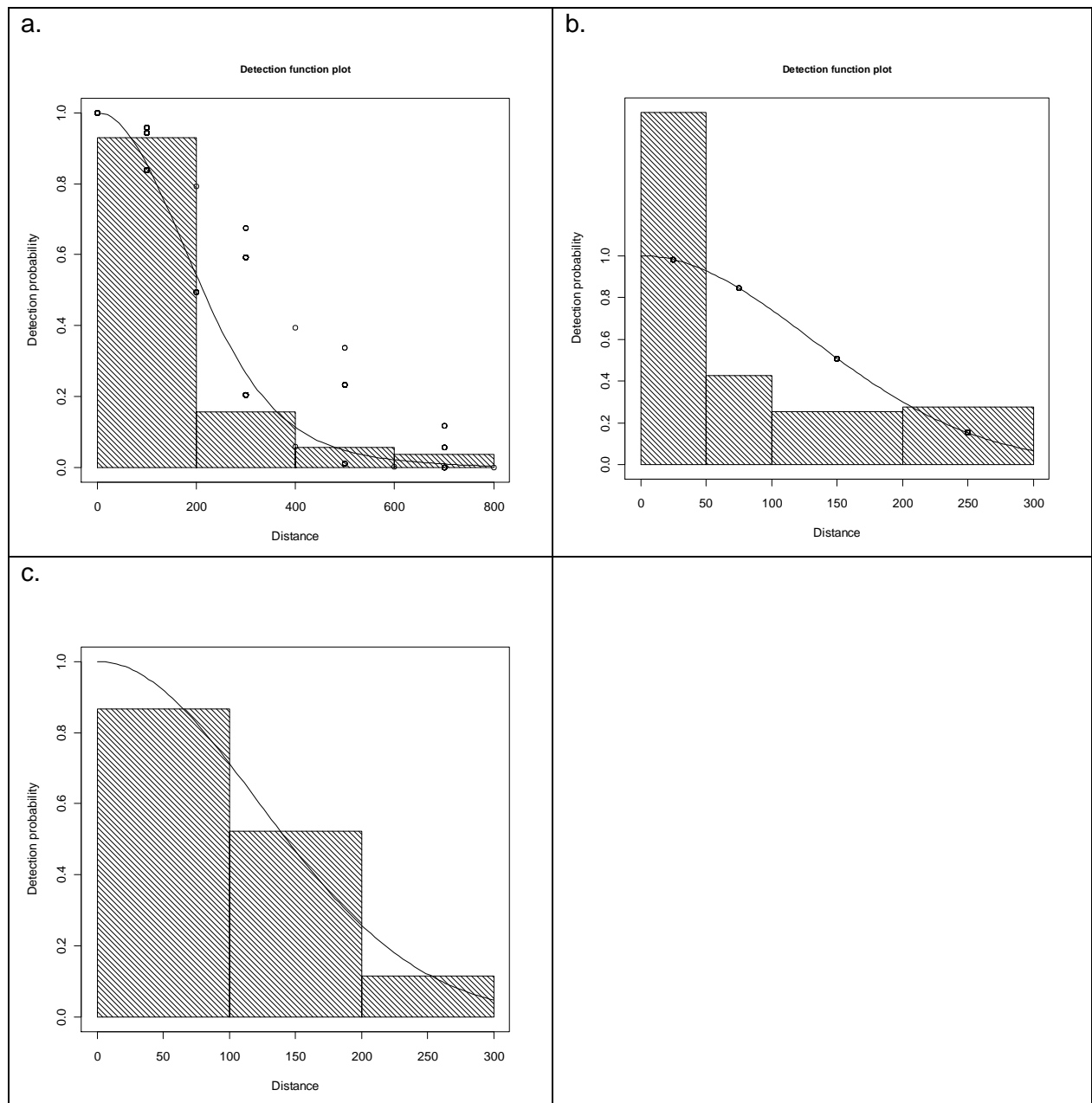


Figure 18. Detection functions for dolphins a. non-ESAS boat sightings, b. ESAS boat sightings, c. plane sightings

Density surface modelling

Harbour porpoise *Phocoena phocoena*

The complete set of harbour porpoise sightings and realized effort are given in Figure 19. The distribution of estimated densities per segment indicated presence of spatial correlation even after fitting a provisional GAM model containing spatial covariates.

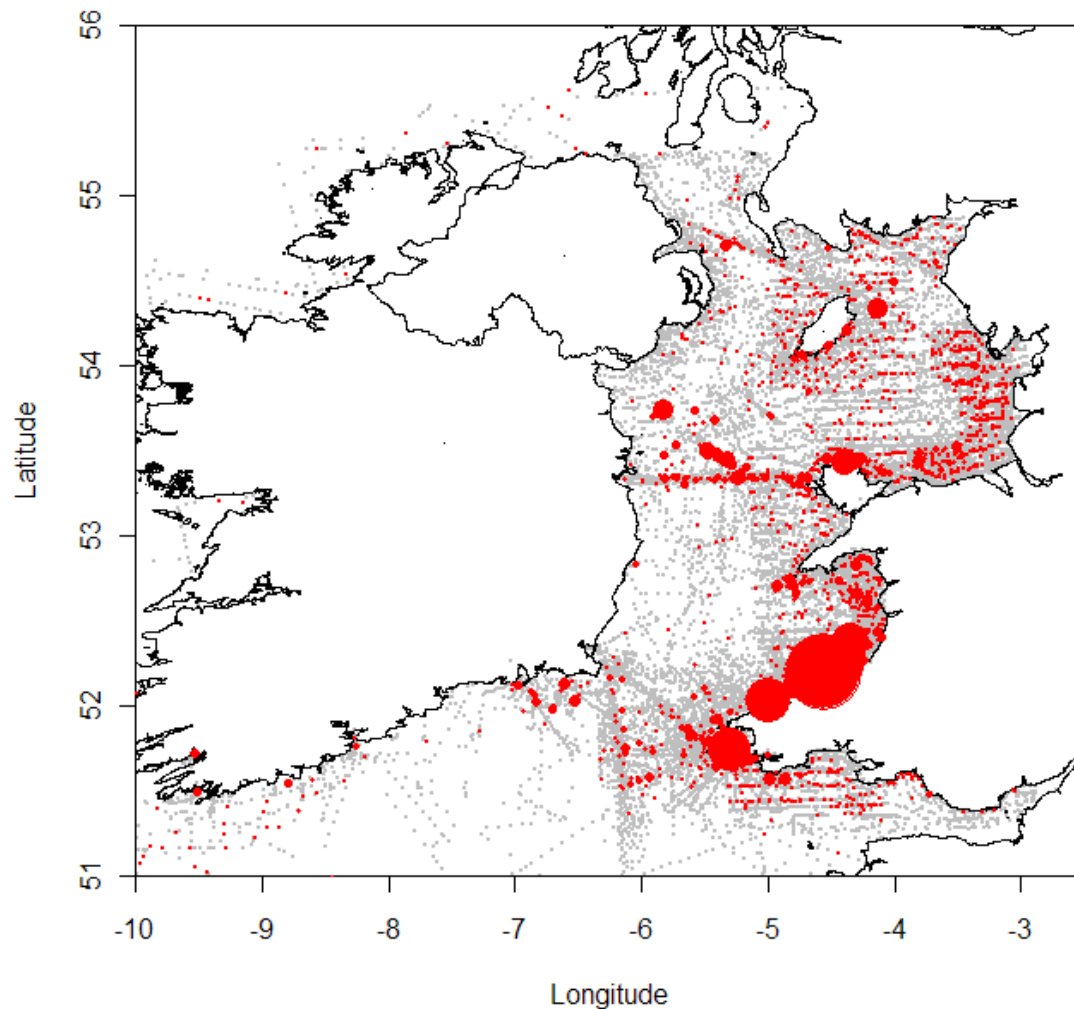


Figure 19. Realized effort (grey) 1980 – 2009 and estimated density per segment of harbour porpoise in the Irish Sea. The area of the circles is proportional to the estimated density of animals per km².

The best of the fitted models are given in Table 2. These models were used to obtain point estimates of density: predicted density surfaces for the mid-points of 7 year periods spanning the dataset (ie for 1983, 1990, 1997 and 2004) are given in Figures 20.

Analysis of a variogram of the residuals of this model suggested that spatial independence was achieved at a distance of approximately 55km (30NM). As the average achieved segment

length was 6km, this suggested that approximately every ninth segment should be considered if wholly independent data were required. This lowered the total number of segments from 36014 to 4002 for the bootstrap. The distribution of density per segment implied that harbour porpoise were substantially over-dispersed with the vast majority of segments containing no animals.

Density surface modelling does allow investigation of habitat preferences of the distributed animals. In this case, when ignoring the wider uncertainty associated with the estimation process there was evidence that more harbour porpoise were, unsurprisingly, associated with shallower water (< 200m) and the spring. However, the latter observation may represent some otherwise unmodelled detectability issue.

Overall temporal trend estimates indicate a substantial increase in density over time (Figure 21). For estimating confidence intervals on trend, and variances for the power analysis, 190 bootstrap resamples were completed; however six of these were clearly divergent, having mean abundance estimates in excess of 140,000 while the next largest was 38,000. These were therefore removed before calculating confidence intervals (Figure 21) and trend variance. (The trend analysis results are reported in the next section).

Table 2. Final fitted density surface models for all species. Dayofyear was always a cyclic smooth in the GAM models.

Model	n	% Explained Deviance *
Harbour Porpoise		
Binomial Model		
s(Lon, Lat) + s(Year) + s(Depth) + s(Dayofyear) + SurveyMode	36014	6
Non-zero presence		
s(Lon, Lat) + s(Year) + s(Dayofyear) + SurveyMode	2978	67
Minke whale		
Binomial Model		
s(Lon, Lat) + s(Year) + s(Depth) + s(Dayofyear) + SurveyMode	48453	24
Non-zero presence		
<i>s(Lon) + s(Depth)</i>	148	8
Bottlenose Dolphin		
Binomial Model		
s(Lon, Lat) + s(Year) + s(Depth) + s(Dayofyear) + SurveyMode	47012	33
Non-zero presence		
s(Year) + s(Depth) + SurveyMode	148	8
Common Dolphin		
Binomial Model		
s(Lon, Lat) + s(Year) + s(Depth) + s(Dayofyear) + SurveyMode	5580	49
Non-zero presence		
s(Lon) + s(Depth) + SurveyMode	75	45
Risso's Dolphin		
Binomial Model		
s(Year) + s(Depth) + s(Dayofyear) + SurveyMode	47012	12
Non-zero presence		
No model fitted	79	-

* The interpretation of explained deviance for binomial models is problematic.

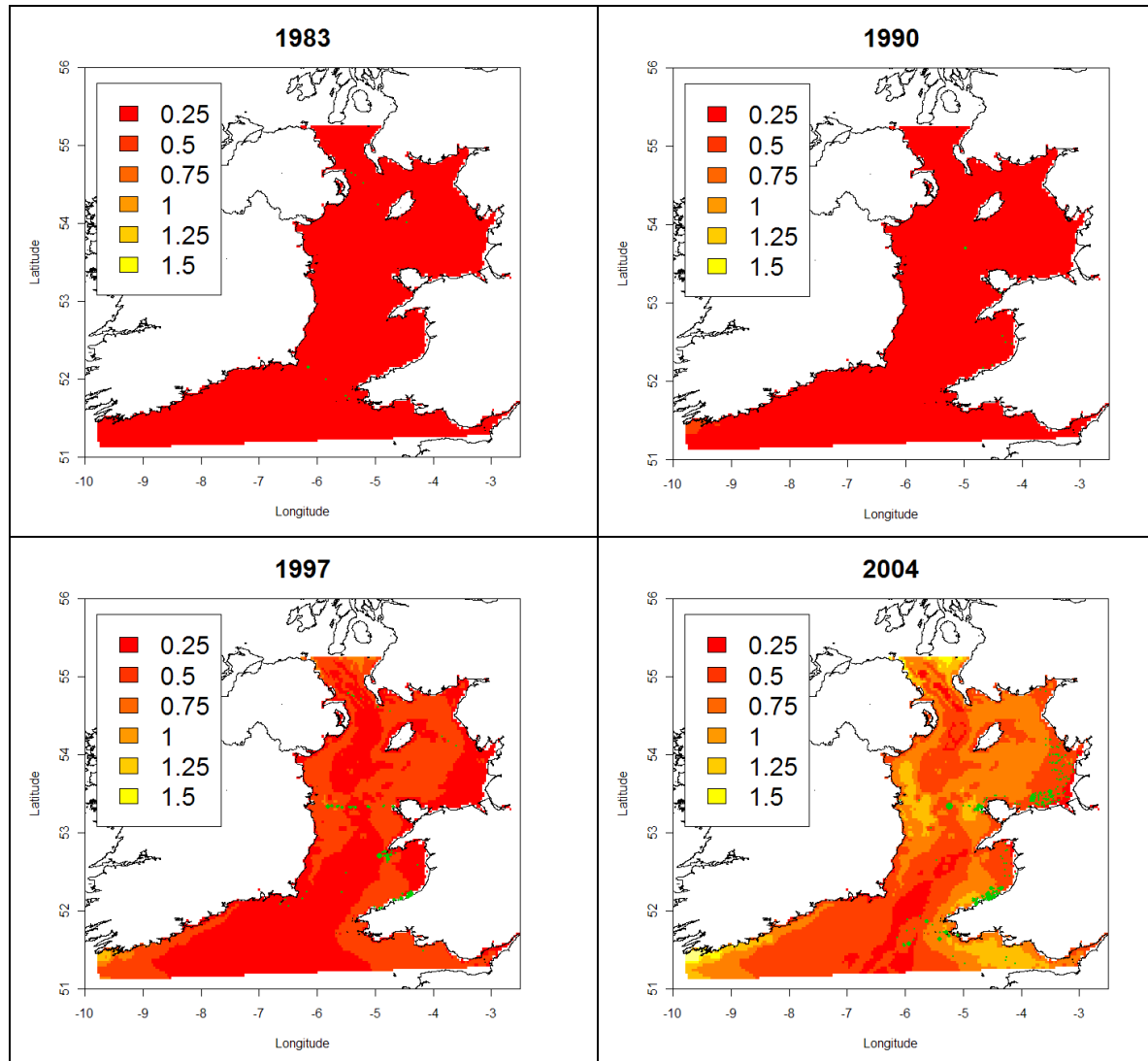


Figure 20. Predictions for harbour porpoise density for the 19th July 1983, 1990, 1997 and 2004 according to a two-stage modelling process. Green circles are proportional in area to estimated density of harbour porpoise associated with that segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

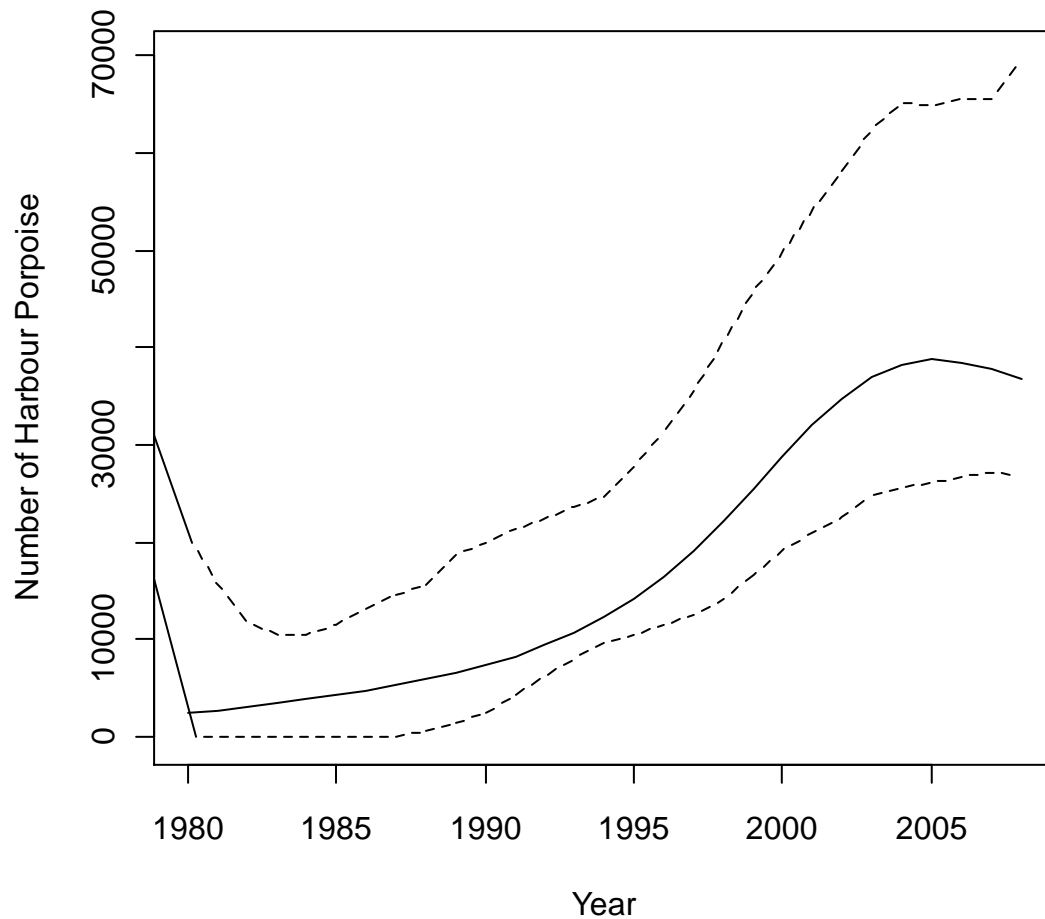


Figure 21. Estimated numbers per year of harbour porpoise in the region of interest 1980 – 2008. Dashed lines show 95% point-wise confidence intervals, estimated by nonparametric bootstrap.

Fin whales *Balaenoptera physalus*

Only 17 out of 47012 available segments of realized effort contained fin whales. This is really too few for reliable inference, and hence no attempt was made to model a fin whale density surface. In a preliminary analysis, there was evidence for spatial correction in estimated segment-level densities which declined to approximately zero at 30NM. Sub-sampling at 30NM reduced the available number of segments to an average of 5224, of which only one contained fin whales. Hence, variance estimation via the bootstrap would also be untenable.

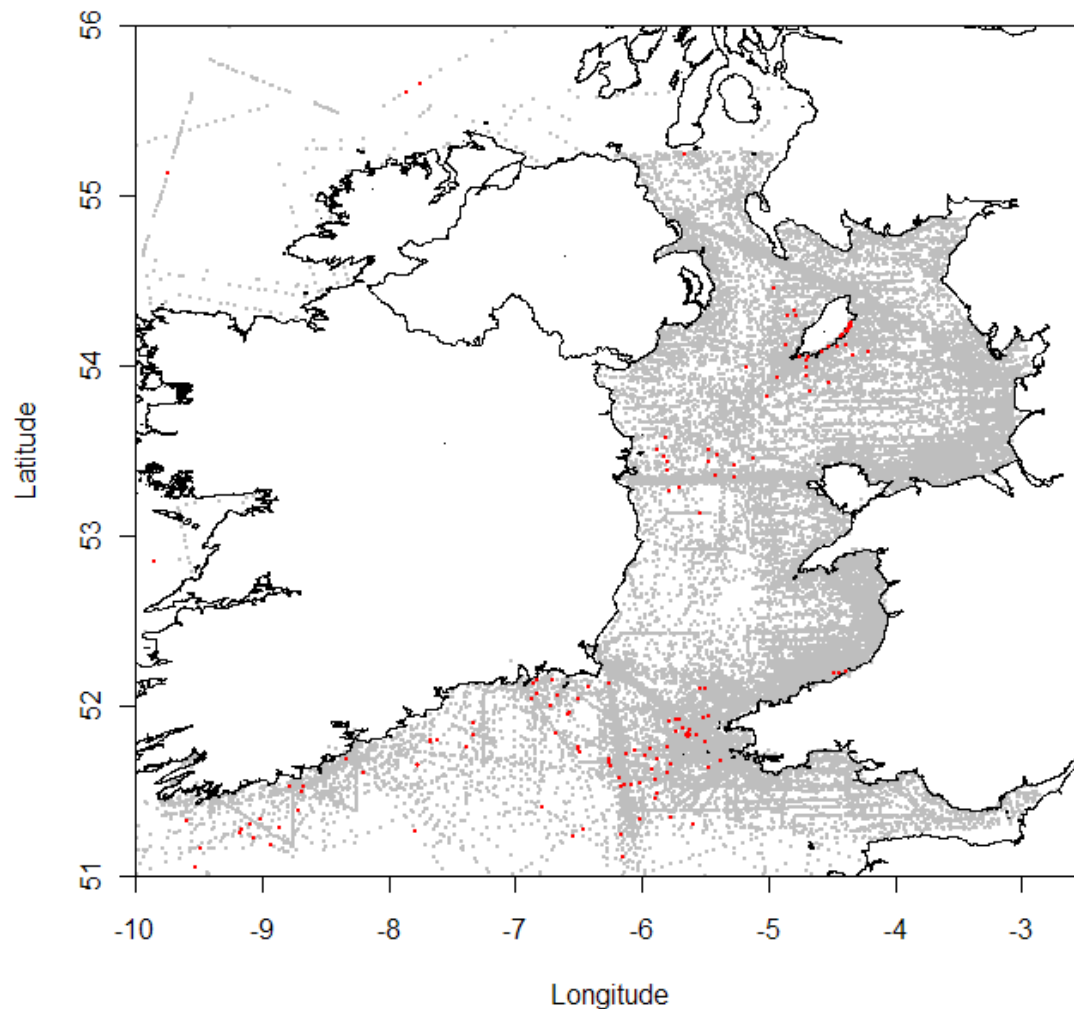


Figure 22. Realized effort (grey) 1980 – 2009 and estimated density per segment of minke whale in the Irish Sea. Area of the circles is proportional to the estimated density of animals per km².

Minke whales *Balaenoptera acuturostrata*

Analysis of the 48543 segments (mean length 5.8 km) for the minke whale analysis suggested, like the harbour porpoise, a large number of zero segments with only 151 non-zero segments (Figure 22).

The best presence/absence model chosen by the UBRE model selection criterion is shown in table 2 as is the best count model chosen by GCV. Model selection for the zero inflated model was by AIC. Predicted surfaces for selected years are given in Figure 23. There was evidence of an upward trend since 1980, although confidence intervals were very wide (Figure 24).

There was evidence that the data were strongly spatially correlated, but independence was achieved at 30 NM, so every ninth segment was considered in the bootstrap leaving an average of 5397 segments. Only 130 bootstrap replicates were performed for the minke whales resulting in erratic behaviour of the confidence intervals.

Within the surveys considered here, minke whales are, oddly, found primarily in shallower water although *Depth* was not selected as an explanatory variable in the presence/absence model.

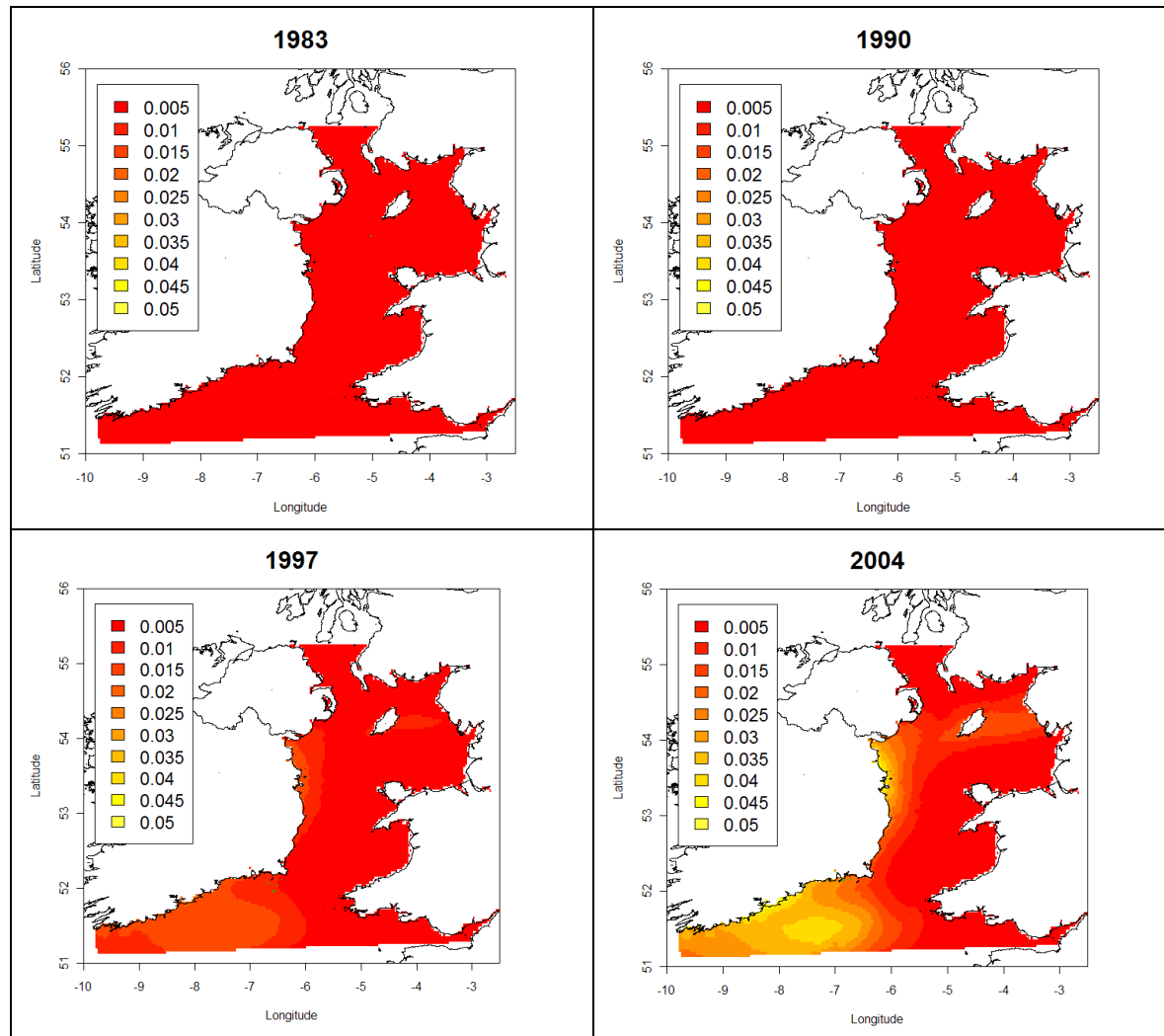


Figure 23. Predictions for minke whale density for the 19th July 1983, 1990 1997 and 2004 according to a two-stage modelling process. Green circles are proportional in area to estimated density of minke whales associated with a particular segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

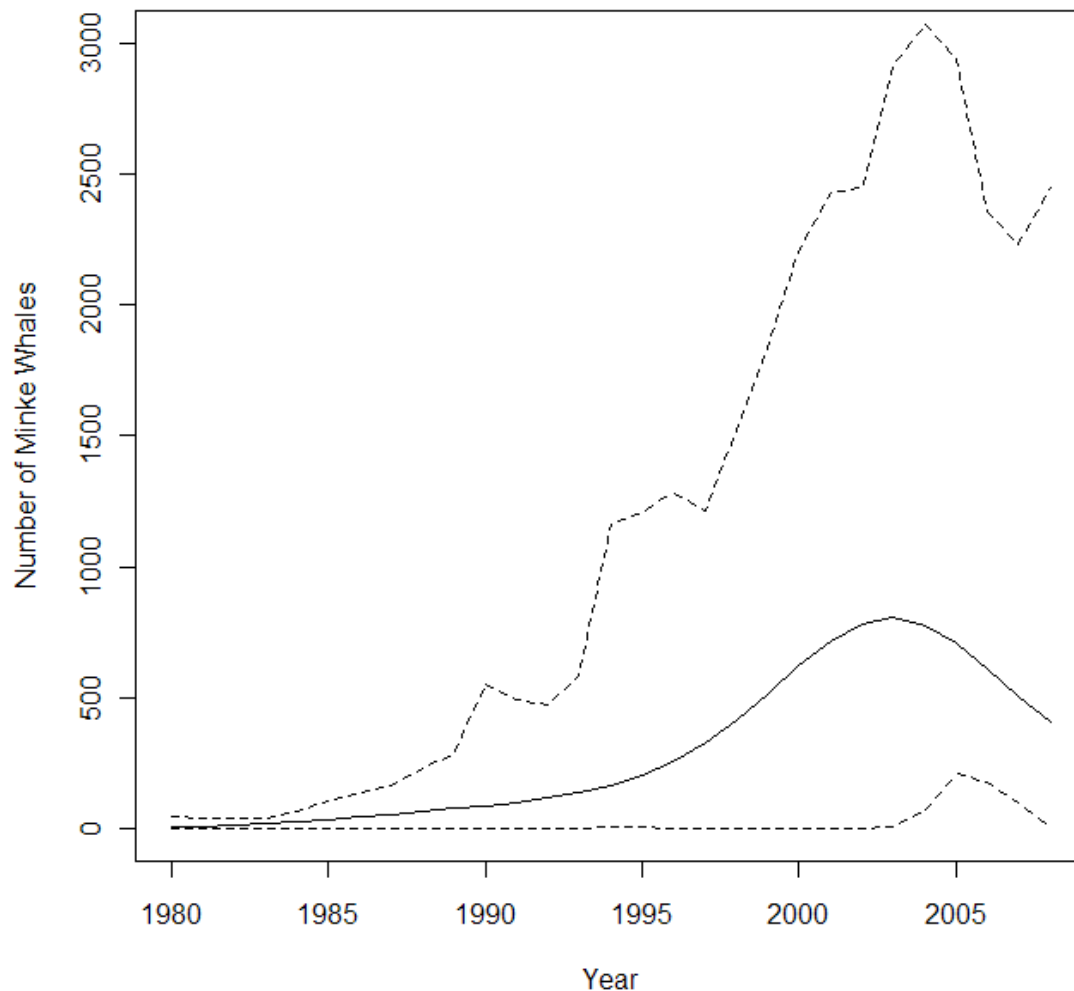


Figure 24. Estimated numbers per year of minke whales in the region of interest 1980 – 2008. Dashed lines show 95% point-wise confidence intervals, estimated by nonparametric bootstrap.

Bottlenose Dolphin *Tursiops truncatus*

In the case of bottlenose dolphins (and all other species of dolphins except for common dolphins) no SCANS data were available, so the total number of segments was 47012 with a total length of 272799 km (Figure 25). Many (1954) of these segments contained sightings of bottlenose dolphins. There was no evidence for spatial correlation in the data so a large sample quasi-binomial GAM was fitted to the data in both the point estimates and the bootstrap followed by modelling of the non-zero densities using a GAM with a gamma error.

The best models can be seen in Table 2 and the predictions from those models are found in Figure 26. Figure 27 gives the annual estimates and bootstrap confidence intervals, based on 190 bootstrap replicates. Apparent abundance peaked in late summer.

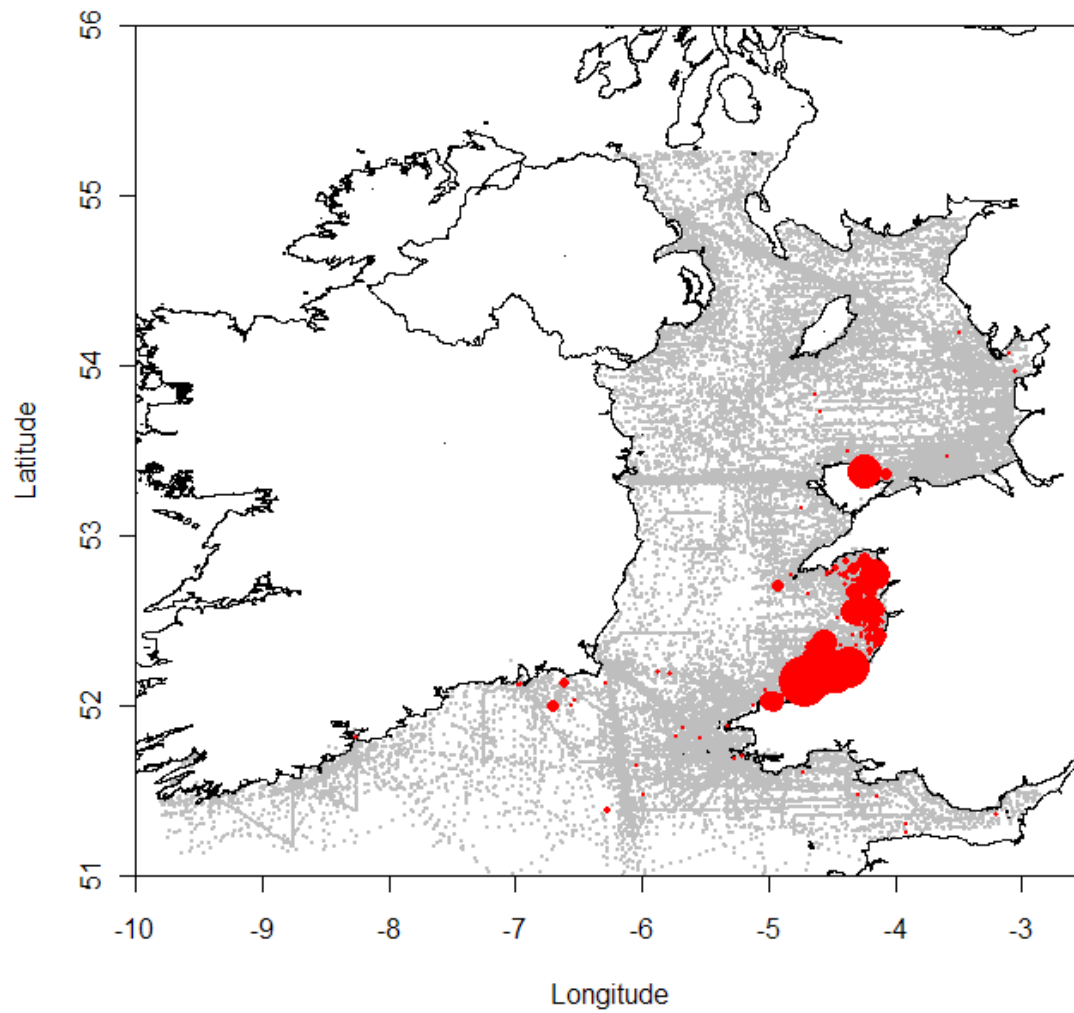


Figure 25. Realized effort (grey) 1980 – 2009 and estimated density per segment of bottlenose dolphins in the Irish Sea. Area of the red circles is proportional to the estimated density of animals per km².

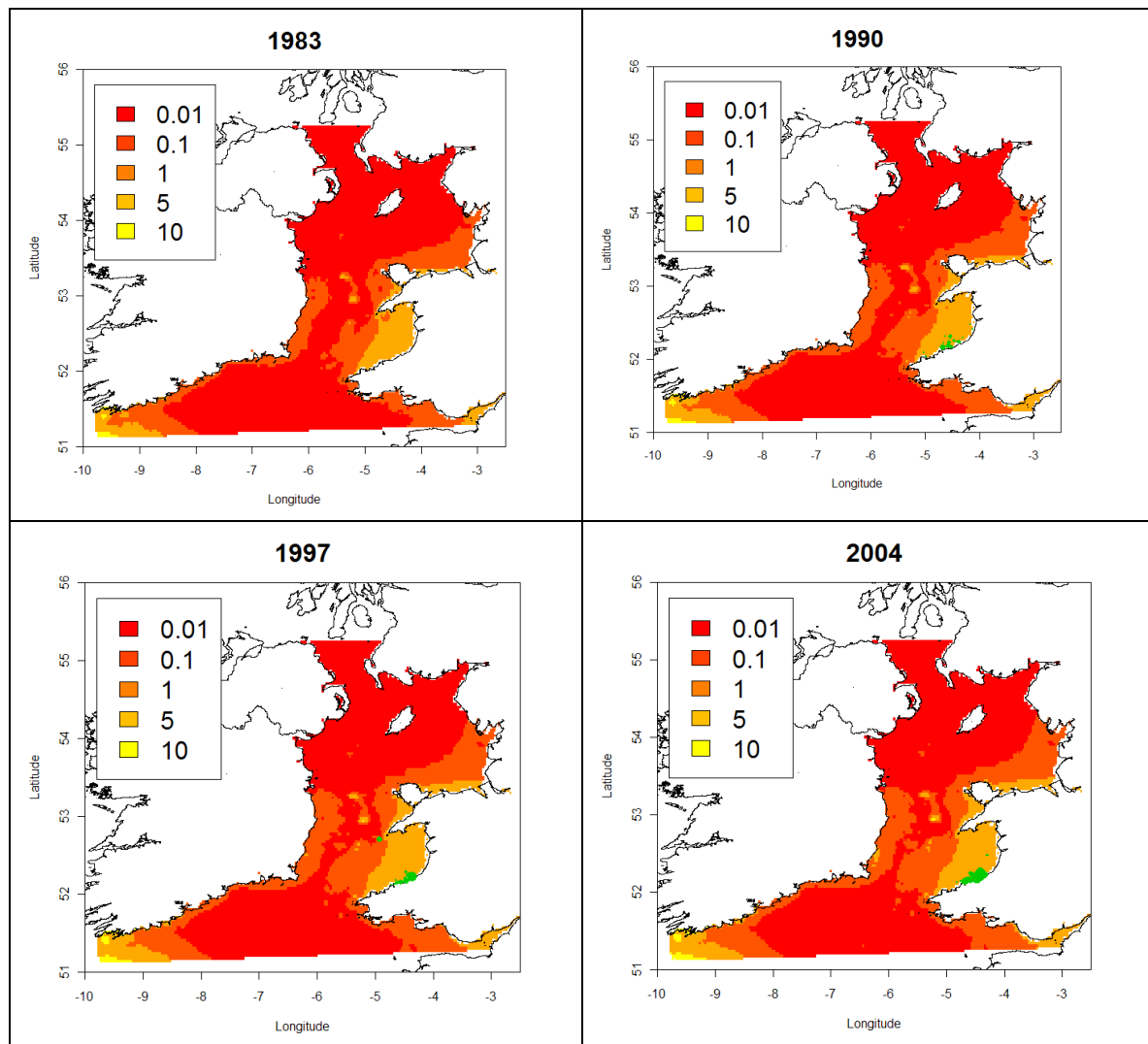


Figure 26. Predictions for bottlenose dolphin density for the 19th July 1983, 1990 1997 and 2004 according to a two-stage modelling process. Green circles are proportional in area to estimated density of bottlenose dolphin associated with that segment locality. Numbers indicate upper bound of colour coded densities (animals/km²). (Note the use of logarithmic scale, to better display the spatial patterns.)

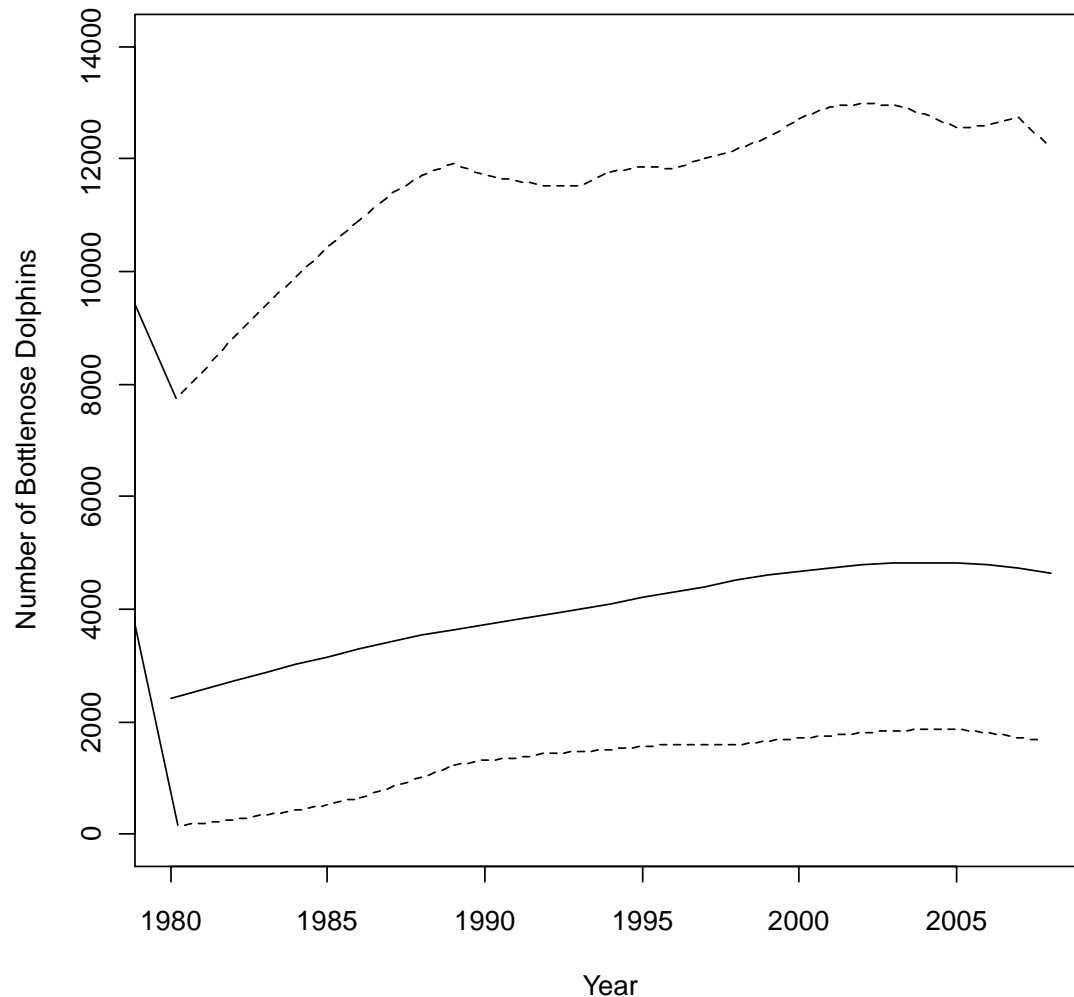


Figure 27. Estimated numbers per year of bottlenose dolphins in the region of interest 1980 – 2008.. Dashed lines show 95% point-wise confidence intervals, estimated by nonparametric bootstrap.

Common Dolphin *Delphinus delphis*

For this species SCANS II data were available for 2005, so the total number of realized segments was 50248 of which 668 contained common dolphins (Figure 28). A large sample GAM could not be fitted to obtain the point estimate due to numerical issues, so the point estimate had to be obtained based a subsample of segments. Also, there was evidence for spatial correlation in the data, which declined to near zero at 30NM. Hence, only segments on average 30NM apart were considered, leaving 5584 segments. Model selection proceeded as before and the final selected models are given in Table 2. The predictions from selected years are given in Figure 30, with bootstrap confidence intervals based on 150 bootstrap replicates. A large number of animals is predicted from recent years primarily from a hotspot in the very south of the region of interest (Figure 29). The annual indices of abundance also show a distinct upward trend (Figure 30). However the upper bounds on the bootstrap samples are very high.

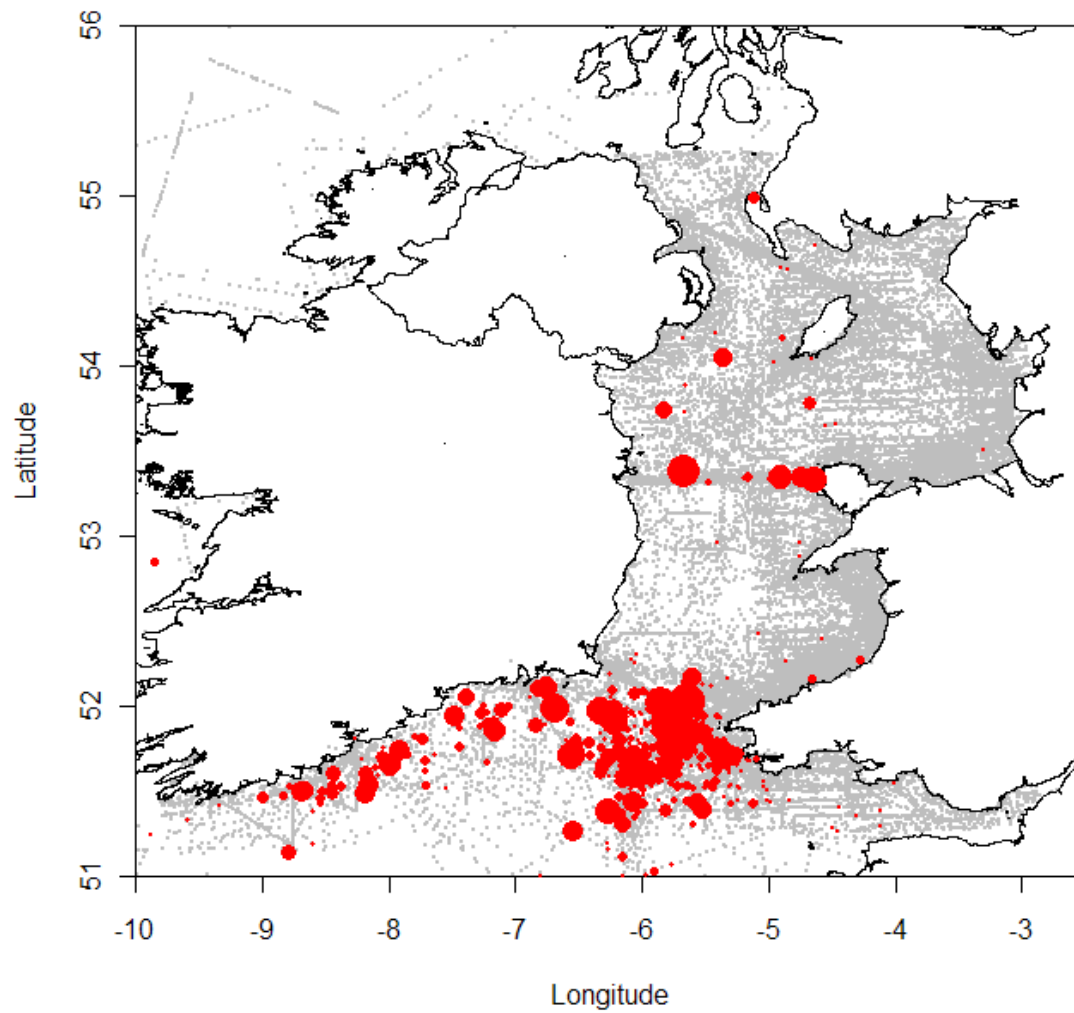


Figure 28. Realized effort (grey) 1980 – 2009 and estimated density per segment of common dolphins in the Irish Sea. Area of the circles is proportional to the estimated density of animals per km².

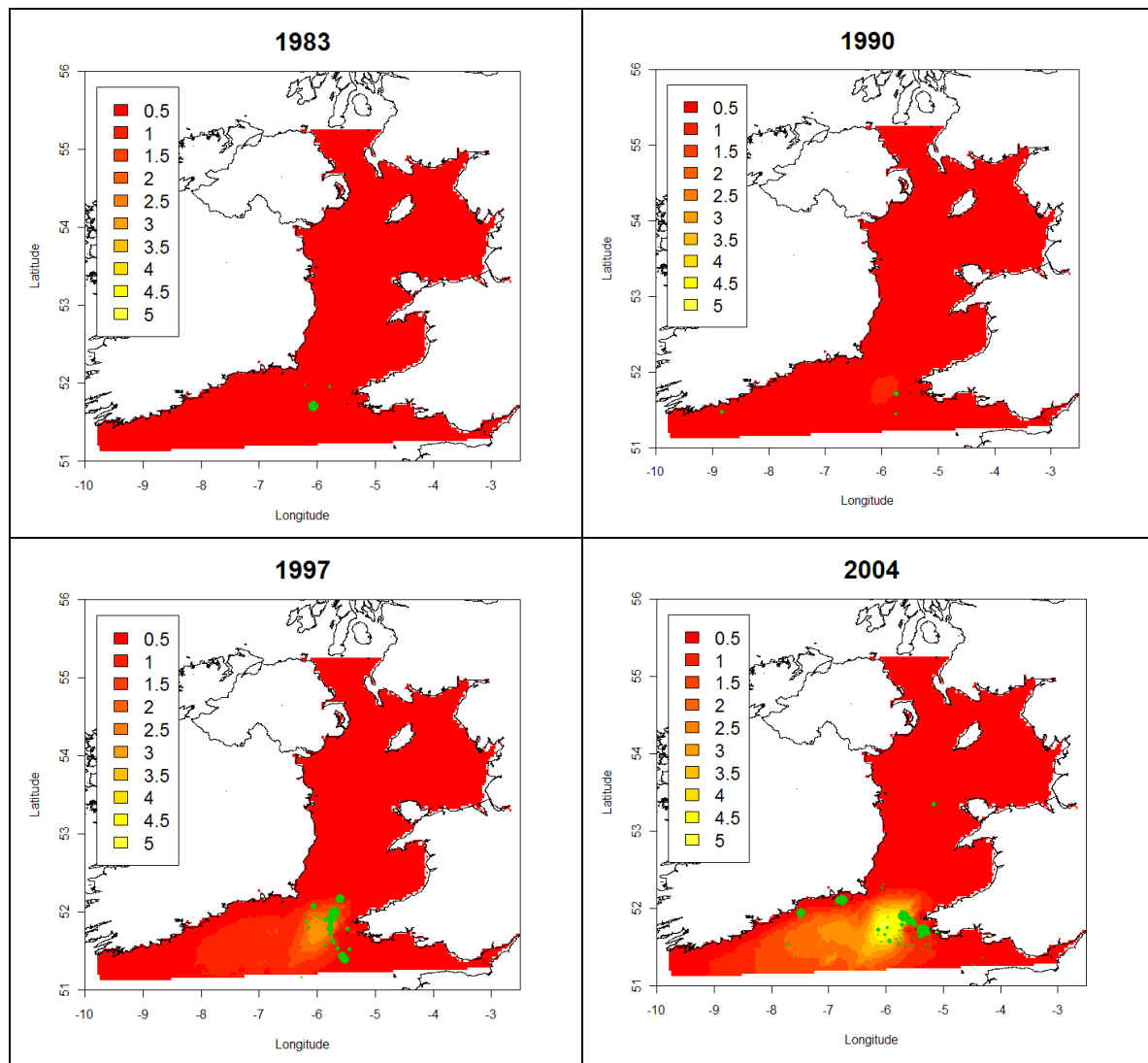


Figure 29. Predictions for common dolphin density for the 19th July 1983, 1990 1997 and 2004 according to a two-stage modelling process. Green circles are proportional in area to estimated density of common dolphin associated with that segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

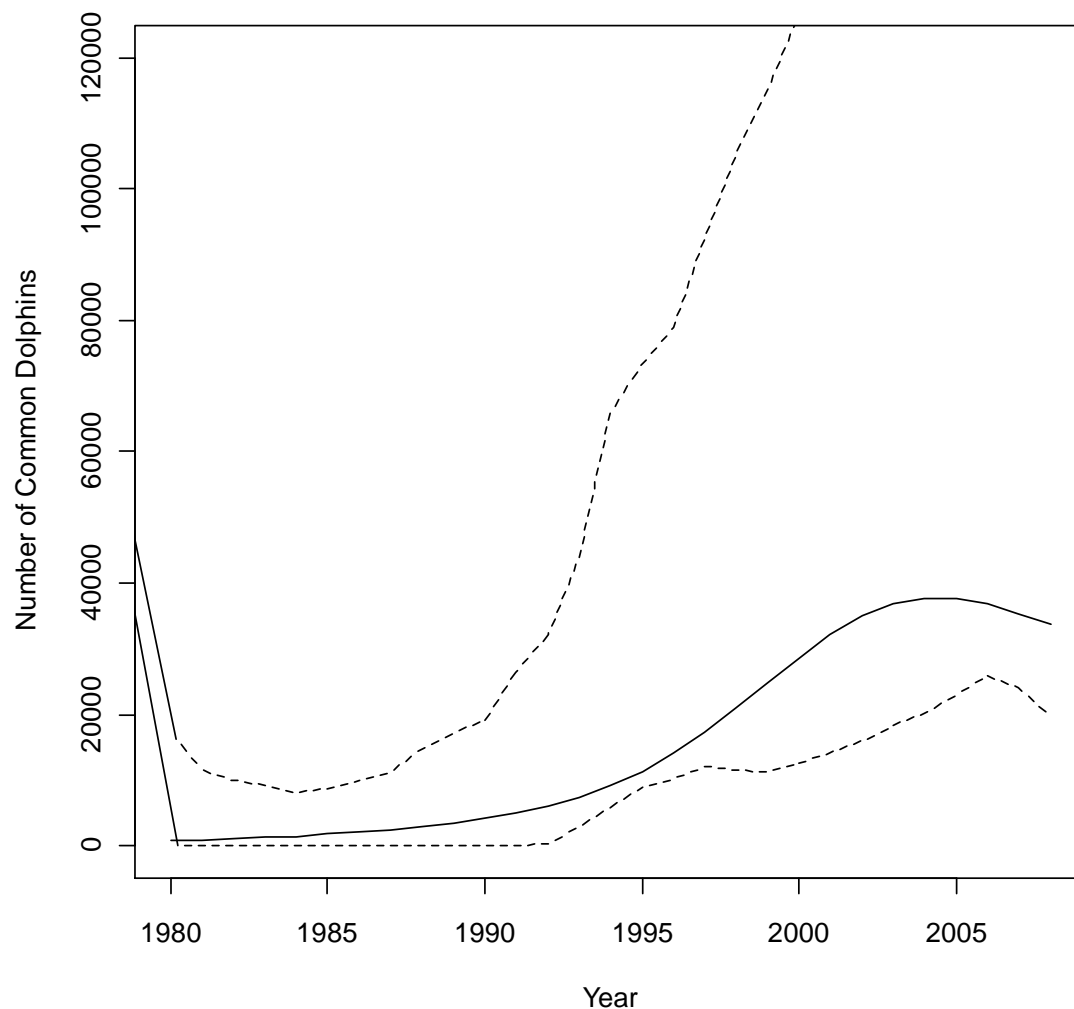


Figure 30. Estimated numbers per year of common dolphins in the region of interest 1980 – 2008. Dashed lines show 95% point-wise confidence intervals, estimated by nonparametric bootstrap.

Risso's dolphin *Grampus griseus*

Only 76 out of 47012 segments of realized effort contained Risso's dolphins (see Figure 31) . No realistic model could be fitted to the non-zero data so a mean non-zero density was calculated instead. Spatial independence in the presence/absence model appeared to be achieved at 50 NM, so this distance was used in subsampling the data during the bootstrap. This reduced the available number of segments for the bootstrap to 3134 on average of which less than 10 typically contained Risso's dolphins in a given sample. Figure 33 gives the predictions for selected years. Figure 34 gives the annual estimates and shows the large width of the confidence intervals, based on 340 bootstrap replicates. Risso's dolphin presence peaked in summer.

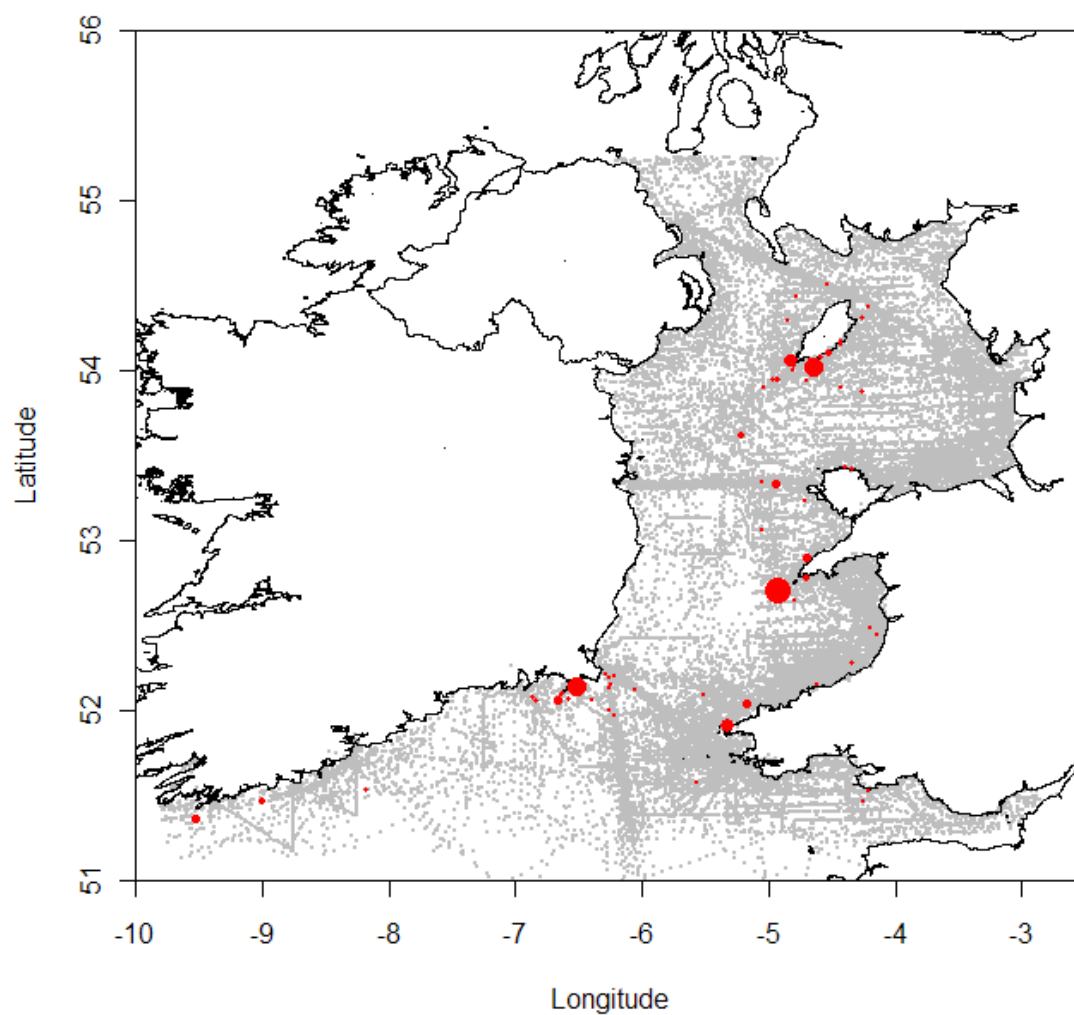


Figure 31. Realized effort (grey) 1980 – 2009 and estimated density per segment of Risso's dolphins in the Irish Sea. Area of the circles is proportional to the estimated density of animals per segment of effort.

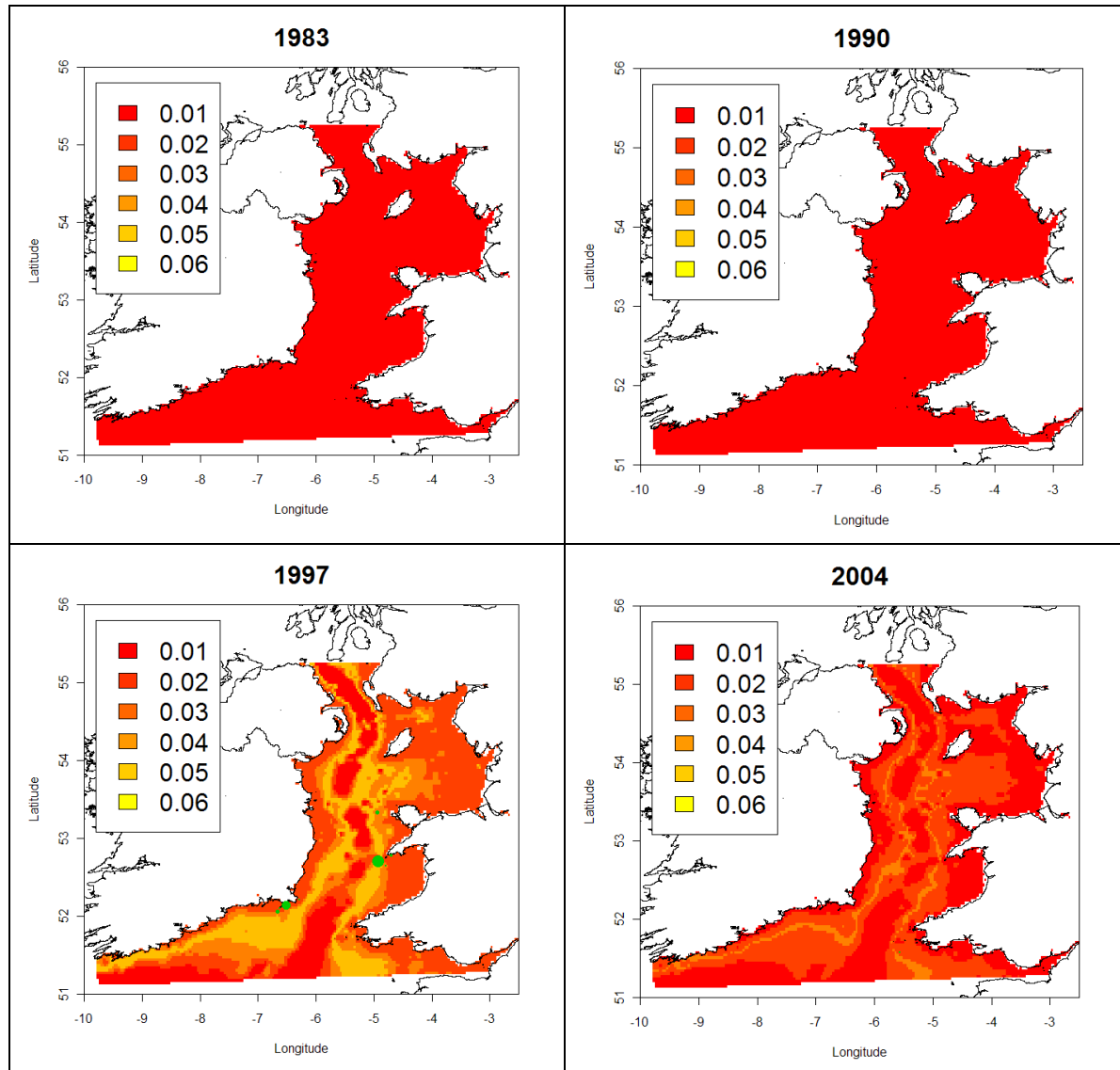


Figure 32. Predictions for Risso's dolphin density for the 19th July 1983, 1990 1997 and 2004 according to a two-stage modelling process (but with mean non-zero-density). Green circles are proportional in area to estimated density of Risso's dolphin associated with that segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

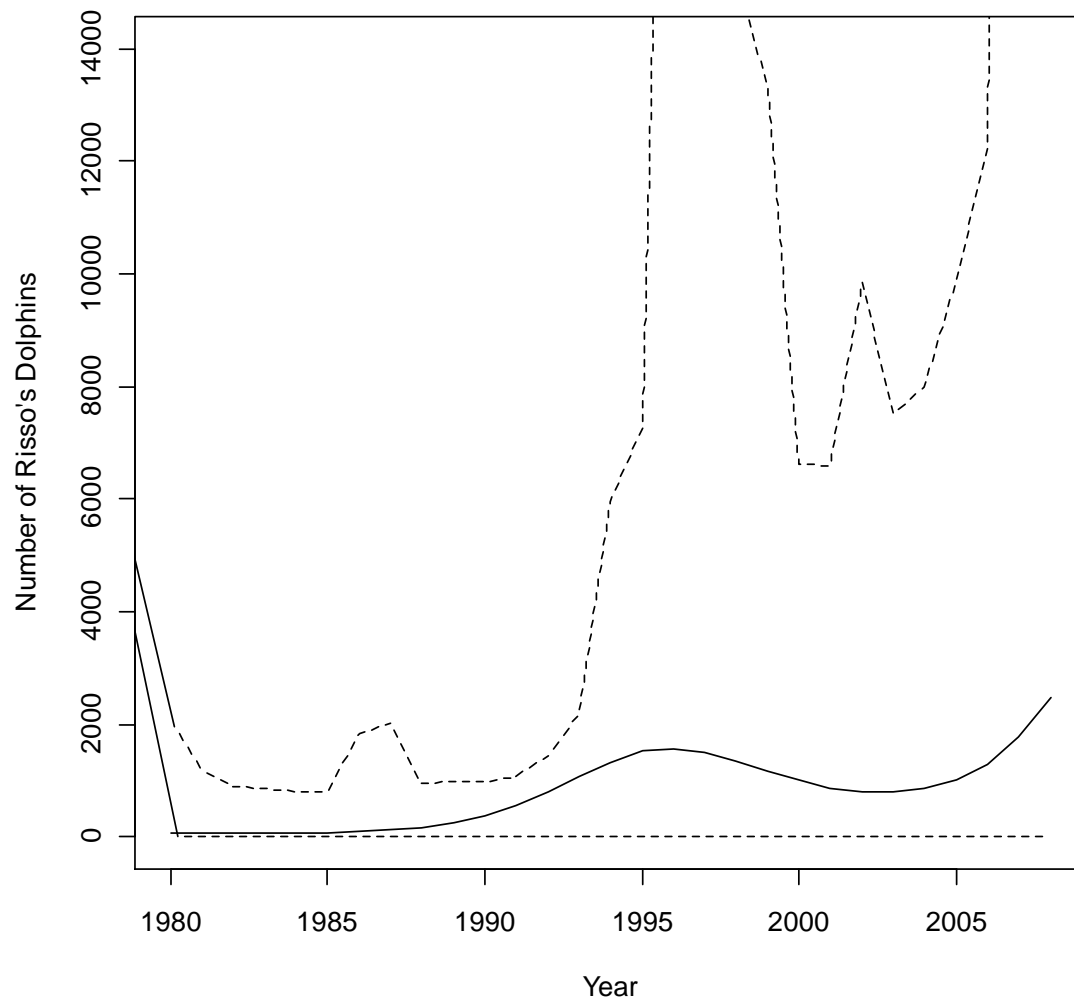


Figure 33. Estimated numbers per year of Risso's dolphins in the region of interest 1980 – 2008. Dashed lines show 95% point-wise confidence intervals, estimated by nonparametric bootstrap.

White-beaked Dolphin *Lagenorhynchus albirostris*

Only small numbers of white-beaked dolphins were seen. Such that only one segment of realized 47012 segments contained any sightings. This meant that the mean estimated density of white-beaked dolphins was in 0.0001 animals/km² in 2008 and zero in other years. White-beaked dolphins were seen at other times but for various reasons eg casual watch observations such sightings were not associated with the final realized effort.

White-sided dolphins *Lagenorhynchus acutus*

The presence of white-sided dolphins was recorded in but two realized effort segment (out of 47041) in 1992. The mean density in that year was 0.048 animals/km².

Power to detect trends

An overview of the relationship between power, average annual population change Δ , and coefficient of variation in Δ , assuming an α -level of 0.05 and a two-sided z-test, is shown in Figure 34. Figures like this (and related calculations) can be used to answer a variety of power and effect size questions. For example, if the average annual change Δ is 0.90 (ie a 10% per year decline) and the CV on this change is 0.2, then the power to detect the decline is 0.76. Conversely, given a desired power of 0.8 and a CV(Δ) of 0.2, the annual population change detectable with this level of power is 0.895 (and the corresponding increase). Calculations of detectable population decline for a range of CV(Δ) are shown in Table 3; similar calculations could be performed for detectable population increase. It is clear that the level of CV(Δ) is critical in determining whether a reasonable level of population decline can be detected over the 6 year monitoring period.

Table 3. Annual population change Δ detectable with a power of 0.8, assuming a two-tailed z-test of $H_0: \log(\Delta)=0$, an α -level of 0.05, and the given levels of CV(Δ). Also shown is the population decline implied by an average annual decline of Δ after 6 years.

CV(Δ)	Detectable Δ	Pop change over 6 years
0.1	0.972	0.846
0.2	0.895	0.517
0.3	0.785	0.235
0.4	0.660	0.082
0.5	0.535	0.023
0.6	0.423	0.006

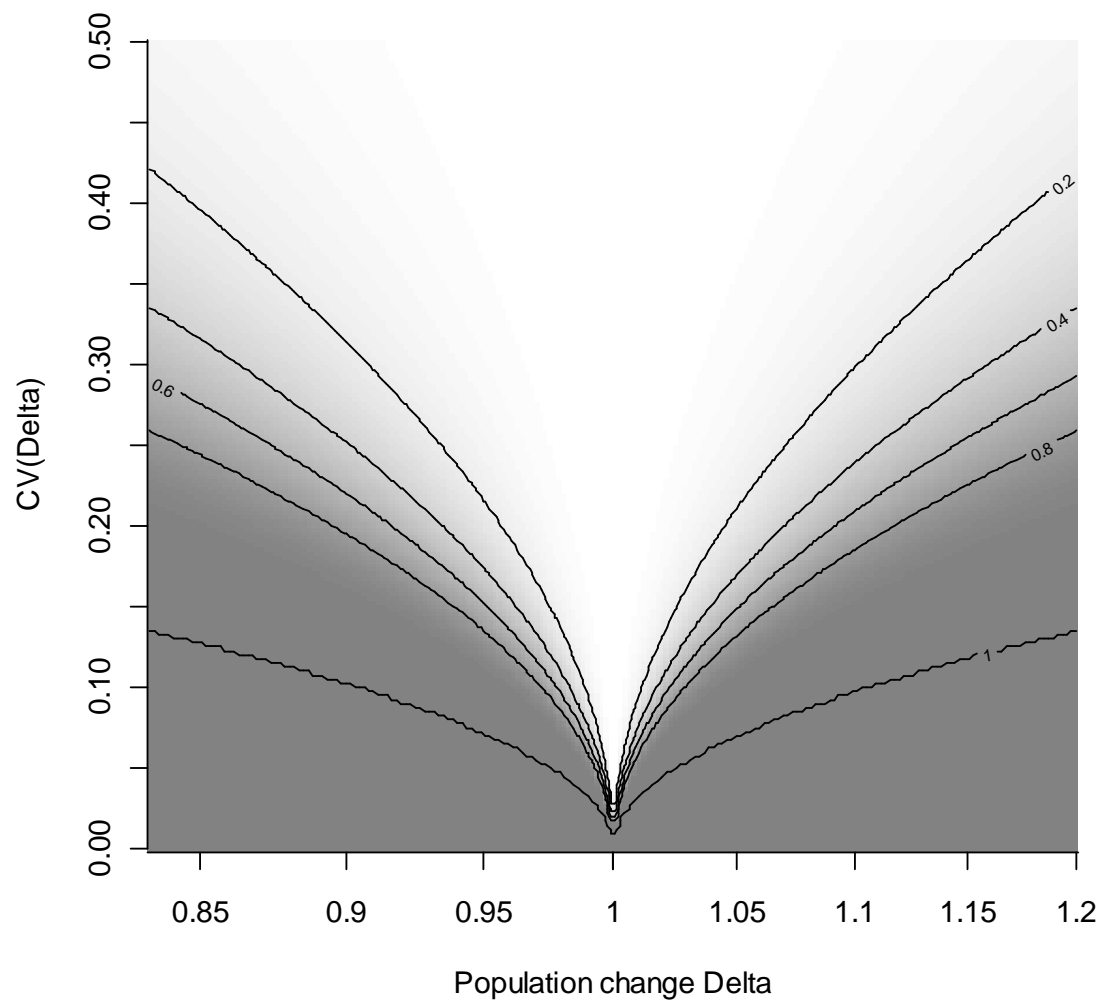


Figure 34. Power (contours) to detect an annual population change Δ with a two-tailed z-test of the null hypothesis that $\log(\Delta)=0$, for a given level of $CV(\Delta)$ and an α -level of 0.05.

Estimated $CV(\Delta)$ s for the 5 species for which spatio-temporal modelling was performed are shown in Figure 35. These are re-expressed in Figure 36 in terms of the detectable annual population decline with a power of 0.8.

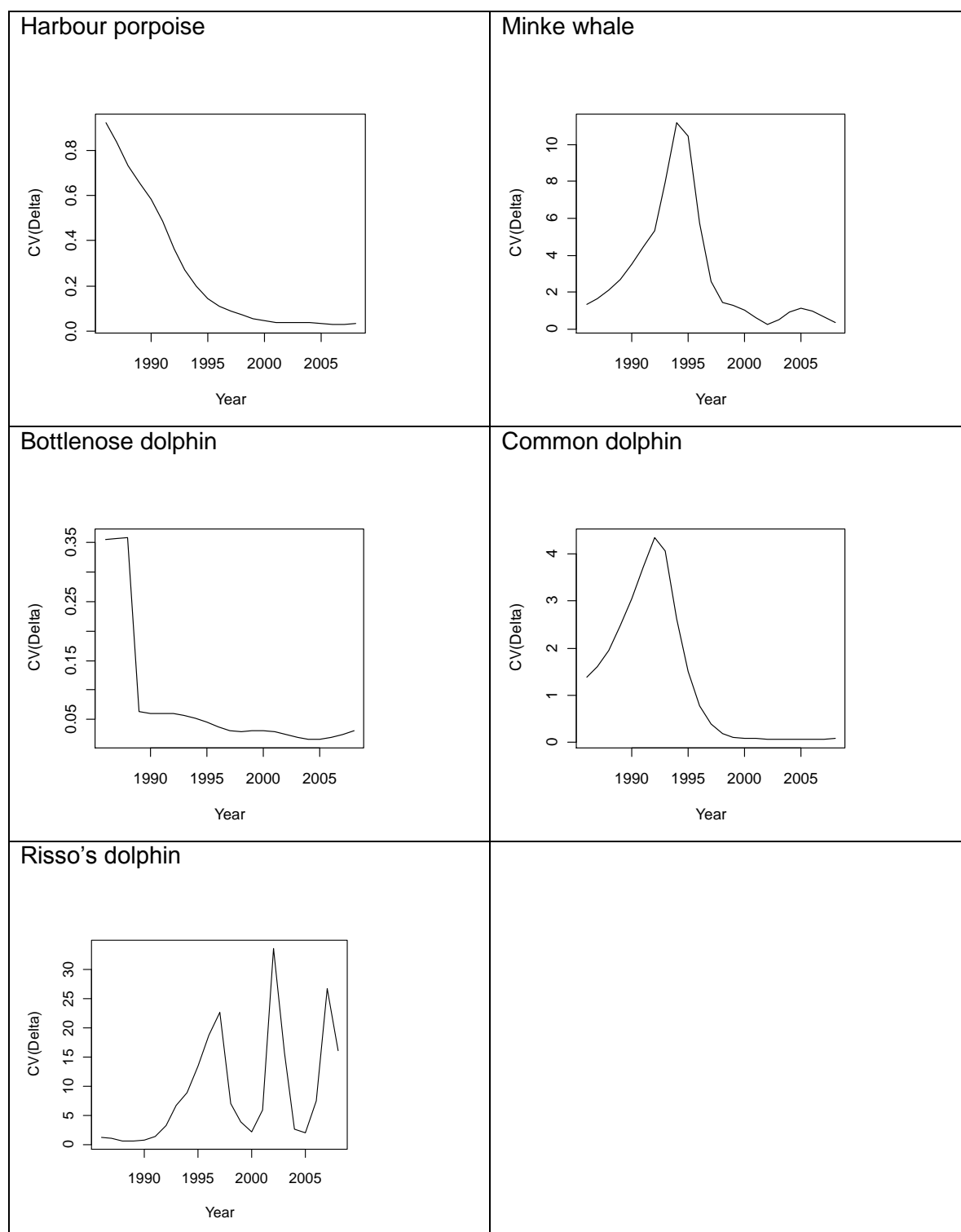


Figure 35. Geometric mean coefficient of variation of annual population change over a 6 year period $t-6$ to t for $t=1986$ to 2008 for the 5 species for which spatio-temporal modelling was undertaken, calculated from the nonparametric bootstrap. Note that the y-axis scales differ.

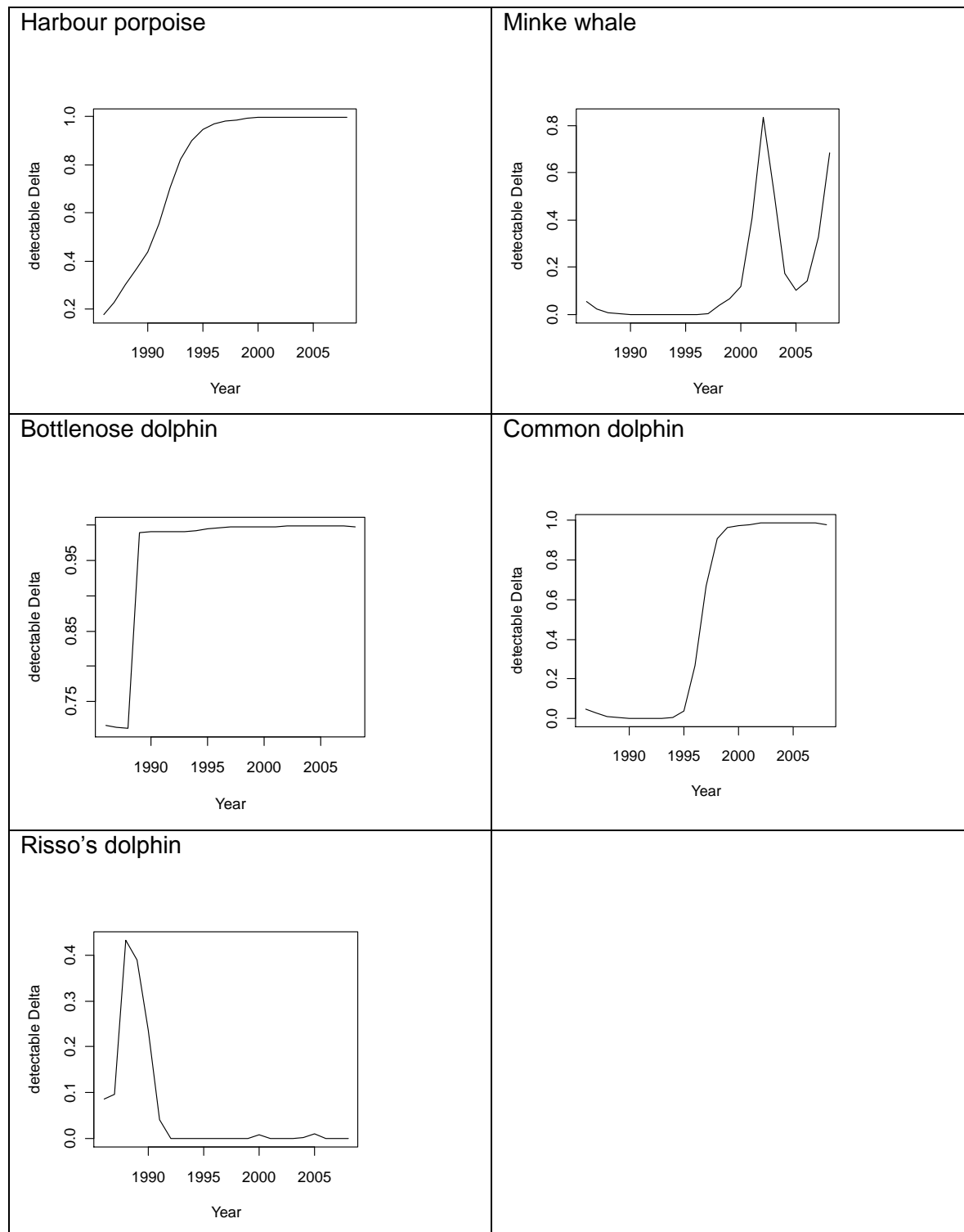


Figure 36. Annual population change Δ detectable with a power of 0.8 over a 6 year period $t-6$ to t for $t=1986$ to 2008 for the 5 species for which spatio-temporal modelling was undertaken. Note that the y-axis scales differ.

The $CV(\Delta)$ is very high in the early years for all species (Figure 35), but declines substantially for harbour porpoise, bottlenose dolphin and common dolphin in the latter part of the time series (around 2000 and onwards, so covering the time period 1994 and onwards). This may be a reflection of the increased quantity of data in the later time periods. Detectable levels of population change are correspondingly low in the early years (Figure 36), but increase substantially for the above three species, and are surprisingly good by the end of the time period. The detectable annual population change given a power of 0.8 for the last time period (2003 to 2008) for harbour porpoise is 0.996 (ie a 0.4% decline per year), for bottlenose dolphin is 0.997 (ie 0.3% decline per year) and for common dolphin is 0.978 (ie 2.2% decline per year). We discuss these results in the next section.

Discussion

This spatial/temporal analysis is more in-depth than has previously been attempted with Irish Sea cetacean sightings data sets and combines analysis of detection as well as encounter rates. Baines and Evans (2009) gave spatial maps of sightings rates for an almost totally overlapping compilation of data and appear to identify similar hotspots, however, the methods employed here allow interpolation into low effort regions.

Several issues arose during the analysis, some relating to the data and some to the nature of the analysis itself. Previous analyses distinguished between formal line transect data and *ad hoc* boat surveys (eg Baines and Evans 2009). It would be useful to grade the data more finely than this. Communication with some observers (Hartley pers. comm.) suggested that *ad hoc* boat surveys could cover a range of activities, some of which could not really be classified as systematic observation. To use these types of dataset effectively, time is required to familiarize oneself with the available data.

Compiling multiple surveys, to produce combined detection functions, proved easy to do and the sightings distributions across surveys showed similar characteristics. Strangely *VesselType* (ie type of vessel), despite its importance in previous analyses (Baines and Evans 2009), was not often selected for inclusion in the detection functions. Segments containing marine mammals were associated more with boat surveys than with aerial surveys presumably reflecting the increased detectability of animals from boats. A *SurveyMode* effect, where aerial effort was associated with a *higher* estimated density of animals, was only found once. Nonetheless, it is still possible that the presence of *SurveyMode* does not just occur because of the lower $g(0)$ of aircraft, but reflects the correlation of location and survey type. Some of the surveys had fairly narrow truncation distances (300m) and sightings of animals beyond the truncation distance were recorded. These data were not considered further in the analysis here but it may be possible to use them in future given a few key assumptions. It may also be possible to group species together in alternative ways if this were thought to be more realistic.

Density surface modelling proved a more complex problem than the analysis of detection functions. Two stage spatial modelling can introduce a bias, as all observed zeros are assumed to be true zeros, but alternatives are not currently as robust as GAMs (see Appendix 3) and it takes effort to find good quality models. Some of the models explained a large proportion of the data deviance, while others did not (Table 2). There is no single obvious mundane non-spatial variable (ie not *Lon* or *Lat*) that clearly predicts density pattern in a simple way for any of the species considered here, except *Depth* in the case of harbour porpoise and Risso's dolphins (both species with a preference for depths of less than 200m). The differences between the final selected model and other similar models was often meagre, yet such model selection

uncertainty was not implemented except for some flexibility over the choice of degrees of freedom in the GAM models. Model selection was influenced by spatial correlation, and methods to deal with correlation in residuals between transect segments should be developed (see below) before the model selection results can be considered robust. Models containing interactions between spatial and temporal covariates should also be considered – for example allowing for possible range shifts over time.

Spatial correlation in the data was dealt with in a clumsy way which at first sight resulted the loss of a great deal of data for variance estimation (note that all data were used in obtaining point estimates for all except common dolphin). Future analysis should aim to improve on the methods adopted here. The use of mixed model GAMs, GAMM, that also allow treatment of correlations within the data without loss of data was suggested by Thomas (2009); however, in practice usable data sets can quickly get too big for GAMMs to work. The development of new methods for GAMs for large data sets may also be extended to GAMMS allowing full treatment of the data, although the fitting is not as robust. Alternative statistical software, such as SAS, may deal with large datasets better. Note that while removing large amounts of data to make the remaining data independent from one another is not ideal, the spatial correlation in the data means that such a loss is rather less important than just the raw proportion of data removed would imply.

One consistent feature of all the analyses was the increase in estimated abundance through time. Partially this may be a consequence of the low levels of effort in the early years. If nothing is seen (eg Risso's dolphins in 1980) then the best estimate of the population is zero. Nevertheless in some species (eg harbour porpoise, bottlenose dolphin and common dolphin) a significant trend is apparent which suggests, even if the point estimates are inaccurate, numbers really have increased. More sophisticated models may allow elucidation of interactions of variables in addition to *Lon* and *Lat* with time, which could allow better insights into the reality of the estimated trends. Switches in habitat (and hence possibly depth preference) by dolphin species are known (Palka *et al.* 1997). However, earlier surveys were not dedicated cetacean surveys so this might have had an effect on detection rates and observers might have been less experienced (P. Evans, pers. comm.).

The results agree fairly well with Baines and Evans (2009). In the case of harbour porpoise in Cardigan Bay, the results for 2004 for example matches with the sightings rate map (Figure 7b) in Baines and Evans (2009). In earlier years, regions of high and low density are not really differentiated. Likewise the distribution of minke whales is similar with a "hotspot" south of Ireland (compare Figure 24 with Figure 32b Baines and Evans 2009). Cardigan Bay is a region of high estimated density of bottlenose dolphins (compare Figure 27 with Figure 13b Baines and Evans 2009) in both studies. Common dolphins also have a similar distribution (compare Figure 30 with Figure 20b Baines and Evans 2009), however the numbers predicted here are very high. This may partially be that the predicted region is further south than considered before and the models predict very high density in this region. The model for Risso's dolphins is wholly different because it does not contain a *Lon Lat* smooth and is driven by *Depth* (compare Figure 33 with Figure 26b Baines and Evans 2009). It would be very useful to compare the estimated abundances with those from other sources, and expert opinion, in order to gauge the reliability of the results. For example, the common dolphin estimates seem implausibly high.

Bootstrapping produced very wide confidence intervals. Refining the models and limiting the degrees of freedom could improve these estimates; alternatively methods such as soap-film smoothing (eg Wood *et al.* 2008) could improve matters. Further use of all the available SCANS data could constrain the periphery of the prediction surface. In almost all cases, time

limitations prevented a large number of bootstrap replicates from being performed, and hence the estimated confidence intervals are not as smooth as they should be. More replicates should be performed, although it is unlikely this will alter the overall picture. A few bootstrap replicates in the harbour porpoise analysis produced egregious estimates and were removed from the calculations of confidence intervals or trend variance; the causes of these estimates should be investigated, but is likely again due to extrapolation of density at the edge of covariate space.

Despite these wide intervals, the power analysis showed that, for some species, surprisingly small annual population declines could be detected over a 6 year period in more recent times. This is because the bootstrap replicates in these cases produced estimates of density over time that almost all had very similar slopes in the later years, and hence very similar estimated population change (see, eg, Figure 37). Despite having similar slopes, the estimates of density had different levels between bootstrap replicates and so produced high estimated variance in density. The high estimate power to detect trends needs to be interpreted with caution. Firstly, it is quite possible that the models fit here are substantially biased, and under-representing the observation error variance. Secondly, the population changes used are changes in estimated density from a smooth model – such models are good for detecting slow and even declines in density, but are very unlikely to react to sudden catastrophic declines (Thomas *et al.* 2004). Hence the high power, even if confirmed, is to detect only certain types of decline. Lastly, unlike many power analyses, the bootstrap ignores process variation, considering variability due to sampling inaccuracies as the only source of uncertainty.

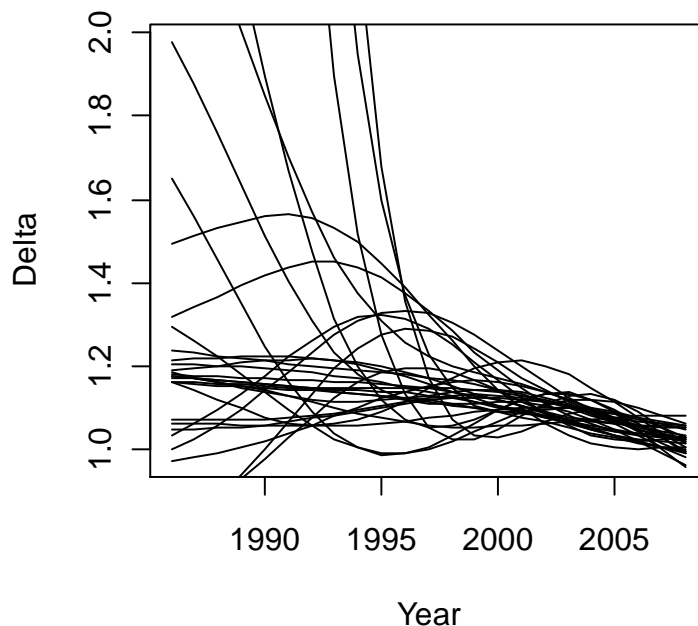


Figure 37. 30 bootstrap replicates of the estimated geometric mean population change, Δ , between years $t-6$ and t for $t=1986$ to 2008 for harbour porpoise.

The current analyses made no use of the land-based watch data, largely because their spatial scale makes them uninformative about large scale spatial patterns of most cetaceans, and

because they are hard to standardize with the other surveys (being time-based rather than trackline length-based; see Thomas 2009 for more on this). However, for bottlenose dolphins in particular, it may be that there is useful information in these data, and this bears further investigation.

Conclusions and recommendations

Conclusions

In broad terms, this first phase analysis of JCP data was a success: data from a large number of sources were combined and standardized, and spatial/temporal models were fit to the data yielding temporal trends. However, there are many significant caveats.

A great deal more time than anticipated was spent processing the data, filtering out duplicate records and other errors, and preparing it for analysis. Because of this, less time was devoted to analysis than planned for.

The fitting of detection functions to line transect data went quite smoothly, and although there are clearly refinements that can be made (such as possibly more categories of vessel type, and perhaps random effects for surveys) this seems the more straightforward part of the data analysis. It will be important to check that the calibration has worked, by checking results against expert opinion from the data providers.

The spatial/temporal modelling of derived density estimates at the segment level did not go so smoothly and there is substantial room for improvement here. In particular, it will be important to find better ways to deal with residual spatial autocorrelation that are also practical to implement on very large datasets. Results of the initial analyses given here need to be checked against expert opinion, and feedback incorporated in any more refined modelling. In general, the results presented here are clearly not ready for publication in, for example, an updated atlas or as UK trend indicators.

Recommendations

Consideration should be given to better standardization of data, better removal of overlap between different datasets, and pre-screening data with simple error checks. More specific suggestions are given in Appendix 2 and in an accompanying example Excel spreadsheet.

Should a Phase 2 analysis take place, more emphasis should be placed on refining methods for detection function modelling and particularly spatial/temporal analysis applied to the Irish Sea data; less emphasis should be given to analysis of datasets from a new area. This is because it takes a great deal of time to assimilate the data, check and correct it, and become familiar with it. Such time inevitably takes away from that spent developing and testing new methods, and checking model outputs. Repeating the current modelling exercise on further datasets is unlikely to lead to reliable inferences or better methods. The first priority should be to develop and implement methods that produce robust and acceptable results for the Irish Sea data, before considering further datasets.

Acknowledgements

Our thanks to all those individuals of the various groups associated with the JCP for collecting, compiling and organizing the data. We especially thank Mick Baines for his prompt and helpful responses to our many queries.

References

- Baines M.E. and Evans, P.G.H. (2009) Atlas of the Marine Mammals of Wales. *CCW Marine Monitoring Report No. 68*
- Borchers, D.L. and Burnham K.P. (2004) General formulation for distance sampling. Chapter 2 in Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers & L. Thomas (editors). *Advanced Distance Sampling*. Oxford University Press.
- Borchers, D. L. Buckland, S. T. Priede, I. G. and . Ahmadi. S (1997) Improving the precision of the daily egg production method using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences* 54(12), 2727 - 2742.
- 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, Oxford. 432 pp.
- Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers & L. Thomas (editors). (2004) *Advanced Distance Sampling*. Oxford University Press 434 pp.
- Davison A.C. and Hinkley D.V. (1997) *Bootstrap Methods and Their Application*. Cambridge. 582 pp.
- European Union (1992) Council Directive 92/43/EEC of 21 May 1992 on the conservation of natural habitats and of wild fauna and flora.
- Gerrodette, T. (1987) A power analysis for detecting trends. *Ecology* 68: 1364-72
- Hedley, S.L. (2000) *Modelling heterogeneity in cetacean surveys*. PhD thesis, University of St Andrews.
- Hedley, S.L. and S.T Buckland. (2004) Spatial models for line transect sampling. *Journal of Agricultural, Biological and Environmental Statistics* 9: 181-199
- Hedley, S.L., S.T. Buckland and D.L. Borchers. (2004) Spatial distance sampling models. Chapter 4 in Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas (editors). *Advanced Distance Sampling*. Oxford University Press.
- Hiby, L. (1999) The objective identification of duplicate sightings in aerial survey for porpoise. In *Marine Mammal Survey and Assessment Methods* (eds G.W Garner, S.C. Amstrup, J.L. Laake, B.F.J. Manly, L.L. McDonald & D.G. Robertson), pp 179-189. Balkema, Rotterdam.
- Hiby, L. and Lovell, P. (1998) Using aircraft in tandem formation to estimate the abundance of harbour porpoise. *Biometrics* 54, 1280-1289.

Houghton, J.D.R, Doyle T.K., John Davenport, J. Graeme C. and Hays, G.C. (2006) Developing a simple, rapid method for identifying and monitoring jellyfish aggregations from the air. *Marine Ecology Progress Series* 314, 159 – 170.

Horvitz, D.G. and Thompson, D.J. (1952) A generalization of sampling without replacement from a finite universe. *J. Amer. Stat. Assoc.* 47: 663 – 685.

Kim, Y.J. and Gu, C. (2004) Smoothing spline Gaussian regression: more scaleable computation via efficient approximation. *J. Roy. Stat. Soc. Series B* 66: 337 – 356.
Laake J.L. and D.L. Borchers. 2004. Methods for incomplete detection at distance zero. In Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds) *Advanced distance sampling*.

Laake, J.L., Calambokidis, J., Osmek, S.D. and Rugh, D.J. (1997) Probability of detecting harbour porpoise from aerial surveys: estimating $g(0)$. *J. Wildl. Manage.* 61: 63-75.

Northridge, S.P., Tasker, M.L. Webb, A. and Williams J.M. (1995) Distribution and relative abundance of harbour porpoises (*Phocoena phocoena* L.), white-beaked dolphins (*Lagenorhynchus albirostris* Gray), and minke whales (*Balaenoptera acutorostrata* Lacepède) around the British Isles. *ICES Journal of Marine Science* 52(1). 55-66.

Palka, D., Read, A. and Potter, C. (1997) Summary of knowledge of white-sided dolphins (*Lagenorhynchus acutus*) from US and Canadian Atlantic waters. *Rep. Int. Whal. Commn.* 47, 729 – 734.

R Development Core Team (2009). 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>.

SCANS-II (2008). Small Cetaceans in the European Atlantic and North Sea. Final Report submitted to the European Commission under project LIFE04NAT/GB/000245. Available from SMRU, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife, KY16 8LB, UK

Steidl, R.J. and L. Thomas. (2001) Power analysis and experimental design. Pp 14-36 in *Design and Analysis of Ecological Experiments*, 2nd edition. S. Scheiner and J. Gurevitch, editors. Oxford University Press.

Thomas L. (2009) Potential Use of Joint Cetacean Protocol Data for Determining Changes in Species' Range and Abundance: Exploratory Analysis of Southern Irish Sea Data. *Report to the JNCC January 2009*. pp. 35.

Thomas, L., K.P. Burnham and S.T. Buckland. (2004) Temporal inferences from distance sampling surveys. Chapter 5 in Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers and L. Thomas (Eds.), *Advanced Distance Sampling*. Oxford University Press.

Thomas, L., S.T. Buckland, E.A. Rexstad, J. L. Laake, S. Strindberg, S. L. Hedley, J. R.B. Bishop, T. A. Marques, and K. P. Burnham. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *J. of Appl. Ecol.* 47: 5-14.

Winship, A.J. (2009) Analysis of harbour porpoise sighting rates from the European Seabirds at Sea database. Appendix D2.4 of *Small Cetaceans in the European Atlantic and North Sea* (SCANS-II) Final Report. EU Life project LIFE04NAT/GB/000245.

Wood, S.N. (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall.

Wood, S.N. (2009) mgcv 1.6-1. Available from <http://cran.r-project.org>.

Wood, S.N., M.V. Bravington and S.L. Hedley. (2008) Soap film smoothing. *Journal of the Royal Statistical Society, Series B* **70**, 931-955.

WWT Consulting (2009) Distributions of Cetaceans, Seals, Turtles, Sharks and Ocean Sunfish recorded from Aerial Surveys 2001-2008. *Report to Department of Energy and Climate Change*.

Appendix 1. Data review

The analysis considered a large amount of data from a variety of sources, not all of which could be used. The purpose of this appendix is to explain why and how data were or were not included in the database.

The data came in a variety of layouts but certain features are common among them. Sightings are normally delineated in space and time. Sometimes there is effort (ie details of both the positions of vessels and explicit statement of when the observers started and finished operating. Sightings can be both incidental (ie casual observations made without any record of effort), or opportunistic (ie made from platforms of opportunity) and finally may come from fully planned surveys. Incidental data/casual observations are not really usable as there is no associated effort. Likewise opportunistic sightings from vessels (where there may be systematic observation) can be used if observation times (times spent observing as well as time when sightings are made) are accurately recorded. If density surface modelling is to be used in the future then sightings and effort MUST be spatially referenced. Land based surveys (if with effort) have some utility however because of the limited geographic range of the data and that it is not trivial to distinguish between the observed number of animals seen from the land and the actual number of animals present to be observed.

Table 1 summarises the available data in turn and its current usability in the proposed analysis. It is possible that some of the data if supplemented (eg with effort data) may be usable in the future. Further details of the main data sets used can be found in the main narrative.

Appendix Table 1.1. Available sightings data and its utility

Source of Data	Location	Time	Mode	Effort	Sightings with distances	Rationale for decision	Used
Cardigan Bay Marine Wildlife Survey Boat Data	Cardigan Bay	2005 - 2006	Vessel based	Yes	Yes but some missing	Estimate detectability	Yes
Ceredigion County Council	Cardigan Bay	1993 - 2007	Land based	Yes	No	Not used as land based	*

Sea Watch in Cardigan Bay 2008	Cardigan Bay	2008	Vessel based	Yes	Yes double platform		Yes
Countryside Commision for Wales	Irish Sea	2004 - 2005	Vessel based	Yes	Yes		Yes
CMACS surveys at Rhyl flats	Rhyl	2008	Land based	Yes	No		*
Department of Energy and Climate Change	Irish Sea	2000 - 2008	Vessel based	Yes	Yes		Yes
European Sea Birds at Sea Surveys.	Around British Isles	1980 - 2002	Vessel based	Yes	Yes		Yes
	Gower	2002 - 2003	Vessel based	No?	No	Data unclear	**
Irish Whale and Dolphin Group	Irish Sea	2004 - 2009	Vessel based	Yes	Yes a few missing	Estimate detectability	Yes
JNCC	North coast of Wales	2003 - 2004	Vessel based	Yes	Yes		Yes
Manx Whale and Dolphin Watch Surveys	Isle of Man	2007	Vessel based	Yes	Yes		Yes
Marie Awareness North Wales	Around Anglesey	2002 - 2008	Vessel based	Yes	Yes		Yes
Pembrokeshire Porpoise Surveys	Pembrokeshire	2007-2008	Vessel based	Yes	Yes a few missing		Yes
Sea Trust Data	Irish Sea	2004 - 2007	Vessel based	No	No	Lack of effort a problem	**
Sea Watch Foundation	Irish Sea	1990 - 2009	Vessel based	Yes	No	Estimate detectability	Yes
SCANS 1	Around Europe	1994	Vessel based	Yes	Yes		Yes
SCANS2	Around Europe	2005	Vessel based	yes	Yes		Yes
Strandings Index	Wales	Ongoing	Land based	No	No	Anecdotes	No

Yes = used for this analysis

* = Not used this time but could be used in the future

** Possibly could be used if more information forthcoming

Appendix 2. Suggested details for JCP data

Data submitted should consist of two distinct effort and sightings files/tables, with associated dates, times and positions relating them together or, at the very least, a common code. The effort data should be supplied both as waypoints showing changes in effort status (ie starts and ends) as well as a separate file of segmented effort. The latter when combined with an analysis of the estimated speeds will allow identification of errors in the data prior to statistical analysis. In addition to sightings and effort files, boundary of the original strata and region of interest for prediction should be given.

Effort files

These should include the position, date and time that surveyors started and finished observing as well as any change in observer status. By preference the position should be recorded as decimal degrees. Dates should be recorded as day, month and year (separate fields) and time as the 24 hour clock (time zone indicated as appropriate). Off effort data should NOT be included (ie casual, unsystematic observation not associated with formal commencement of observation).

Effort data should recall the position, date and time of the vessel at regular intervals throughout the survey. This allows estimation of speed which allows anomalous positions/timings to be identified (as they will often generate strange speeds and can be created into segments of identical environmental/watch conditions).

Method should be clearly identify eg dedicated line transect, formal watches without collection of distances, etc. Data from casual observations (without records of effort, ad hoc sightings etc.) should NOT be submitted.

Further details that should be given include number of observers, observation platform height, vessel type (eg ship), platform type (eg crow's nest) and any change in status (change in the number of observers, method, etc). Environmental conditions which, at the very least, should include Beaufort sea state should also be provided.

Effort files should clearly identify (where necessary) data collected under single or double platform observation modes, line transect versus point estimate data etc and should contain unique original transect identifiers where present.

Serious consideration should be given to assigning a unique identifier to each effort record in the combined database, prior to hand-over for analysis. This could be facilitated by each survey being assigned a unique ID prefix for effort segments. This would go a long way towards preventing provision of overlapping data by different data providers.

Sightings files

These should give time and position of sighting, in such a way that it can be related to effort. Ideally this should be by time and date rather than a code as the data may have to be re-segmented. Off effort sightings should NOT be included.

Details of species and group size should be given, ideally broken down by class (adults, calves etc). Distance to sighting, when recorded, should clearly identified as radial or perpendicular

distance or perhaps by other details (inclination, bearing, plane height etc.) depending on the survey.

Sightings data should provide associated environmental covariates (although if properly cross referenced these should be obtainable from the effort file).

If double platform data have been collected, the most obvious way to record this is two lines for each platform with a unique sighting identifier, a platform identifier (primary or secondary etc) and another field indicating whether the particular animal was seen or not.

Serious consideration should be given to assigning a unique identifier to each sighting in the combined database, prior to hand-over for analysis. This could be facilitated by each survey being assigned a unique ID prefix for sightings. This would go a long way towards preventing provision of overlapping data by different data providers.

If not contained in the above two files (the preferred option), there should be a statement briefly describing the conditions of the survey, number of observers, search pattern, platform heights, mode of survey (casual watches etc), type of survey (single or double platform etc), vessel types etc.

Appendix 3. Zero inflated models

Many zero modelling

As an alternative to the modelling of presence/absence followed by numbers if present, the data were modelled using zero-inflated, poisson and negative binomial models. Here y_i is a density that can take a value of zero or above. The advantage of this over conventional modelling approaches using GLMs or GAMs is that the differing nature of the zeros in the data can be accounted for ie true zeros, the animals are never present and zeros that are realisations of the count process. The current disadvantage is that the current fitting methods are not as flexible or fast as GLM and GAM methods.

In the context of the data here the estimated density values, per segment values it can take, can be modelled as follows.

Zero inflated Poisson

$$\begin{aligned} y_i &= 0 && \text{with probability } p_i \\ y_i &= \text{Pois}(\lambda_i) && \text{with probability } 1 - p_i \end{aligned}$$

so that $y_i = k$ with probability $(1 - p_i)e^{-\lambda_i} \lambda_i^k / k!$ for $k = 0, 1, 2, \dots$

therefore $y_i = 0$ can also be obtained with probability $(1 - p_i)e^{-\lambda_i}$

Zero inflated negative binomial

$$\begin{aligned} y_i &= 0 && \text{with probability } p_i \\ y_i &= \text{NB}(\lambda_i) && \text{with probability } 1 - p_i \end{aligned}$$

so $y_i = k$ with probability $(1 - p_i) \frac{\Gamma(k + \alpha^{-1})}{k! \Gamma(\alpha^{-1})} \frac{\alpha^k \lambda_i^k}{(1 + \alpha \lambda_i)^{k + \alpha^{-1}}}$ for $k = 0, 1, 2, \dots$

therefore $y_i = 0$ can also be obtained with probability $(1 - p_i)(1 + \alpha \lambda_i)^{-\alpha^{-1}}$

α is a parameter to be estimated and Γ is the gamma function $(n-1)!$

The data were modelled using the *pscl* library in *R*. This library does support the use of GAMs therefore the predictors were modelled using polynomial functions of covariates. Because of the long length of time required to fit the models, model selection (using Akaike Information Criterion) proceeded forward and backward from a start model for both the count component and the binomial component of the zero-inflated model. The start model for each component was the best model selected from the equivalent GAM models, with the smooth terms converted to their nearest equivalent polynomial. Model selection then proceeded from here by incrementally increasing or decreasing the degrees of freedom for each smooth by one. Factor covariates were omitted from the models. For inclusion in a model *Dayofyear* had to be associated with a least 2 df because of its obvious cyclicity. Given the above strategy, it is conceivable that the most parsimonious models were not necessarily found.

Implementation of the zero inflated models proved difficult and whilst models were obtained, it proved impossible to explore all relevant combinations of variables and obtain reasonable estimates of abundance. Therefore the results of this analysis are given as alternative to the

main set of results. The full data sets could not be fitted so a subset was used as given in Appendix Table 3.1. Where possible this subset reflected spatially independent units.

Results

General Comments

The best models so far found are given in Appendix Table 3.1. The proportion of explained deviance can vary considerably and was not necessarily related to a more intuitive interpretation of what constituted a good model. In the case of Risso's dolphins a zero inflated model could not be fitted. The count component was always found to be negative binomial rather than Poisson presumably reflecting the typically overdispersed nature of cetaceans. Because the zero inflated models were extremely sensitive to small changes in the degrees of freedom of each smooth, bootstrapping to obtain a variance estimate was not possible.

Appendix Table 3.1. Final fitted density surface models for all species. In the case of zero inflated models, the formula left of the bar gives the count model, the formula to the right the presence-absence component.

Model	n	% Explained Deviance *
Harbour Porpoise MidLon ⁴ + MidLat ⁶ + Depth ⁴ + Dayofyear ⁴ + Year ³ + SurveyMode Depth ³	4002	82
Minke whale 1 Depth ³	5394	7
Bottlenose dolphin MidLat ³ + Depth ³ + Dayofyear ² + SurveyMode Year ³ + Dayofyear ²	5004	45
Common Dolphin Year ³ MidLon ³ + MidLat ³ + Year ² + Dayofyear ²	5580	44

* The interpretation of explained deviance for zero-inflated models is problematic.

Harbour porpoise

This zero-inflated model was sensitive to changes in the number of degrees of freedom. The component two stage modelling models and the zero-inflated models are not particularly similar suggesting perhaps the optimal zero-inflated model has not been found especially as the predicted surface is not readily reconcilable with the data (Figure Appendix 3.1.). The best estimate of numbers for 2004 was 9300 animals compared to the 38200 from the two stage model for the same year.

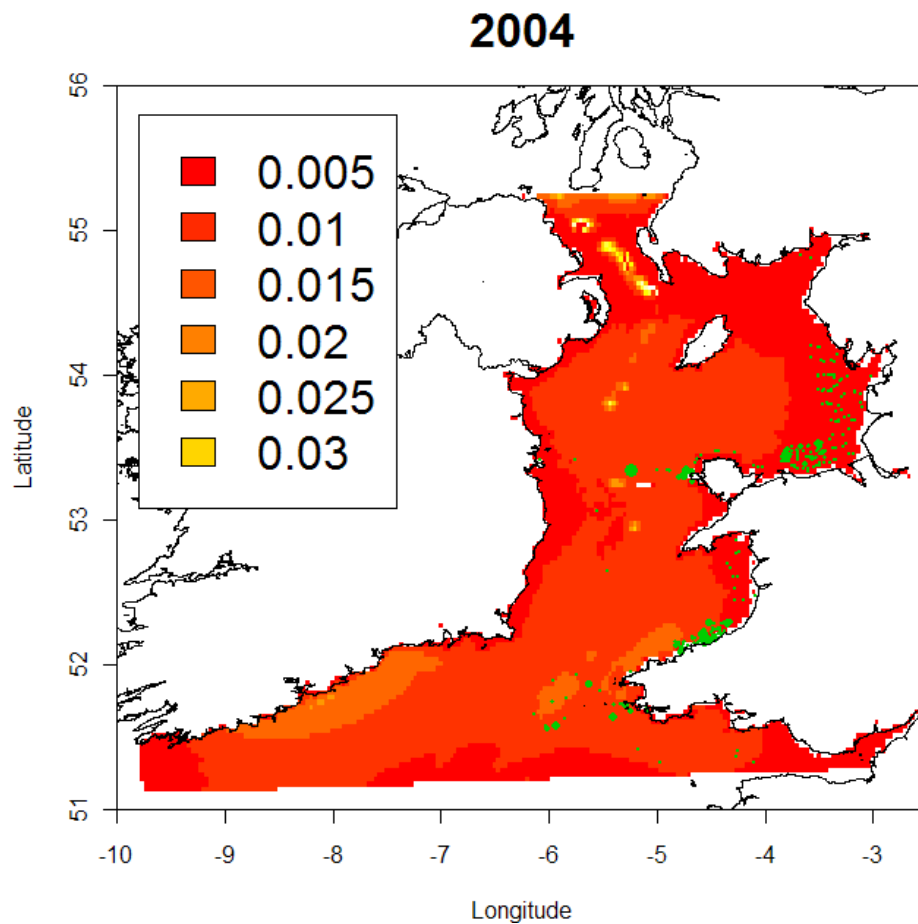


Figure Appendix 3.1. Predictions of harbour porpoise density for the 19th July 2004 according to a negative binomial zero inflated model. Green circles are proportional in area to estimated density of harbour porpoise associated with a particular segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

Minke whale

In this case the sole variable predicting presence absence was *Depth*, so the predicted density was forced to follow the depth contour. Given the sparsity of the non-zero data ($n = 19$) for the reduced data, the model is not very informative. The time invariant estimate of minke whales in the region is 1200 compared to the two-stage modelling estimate of 800.

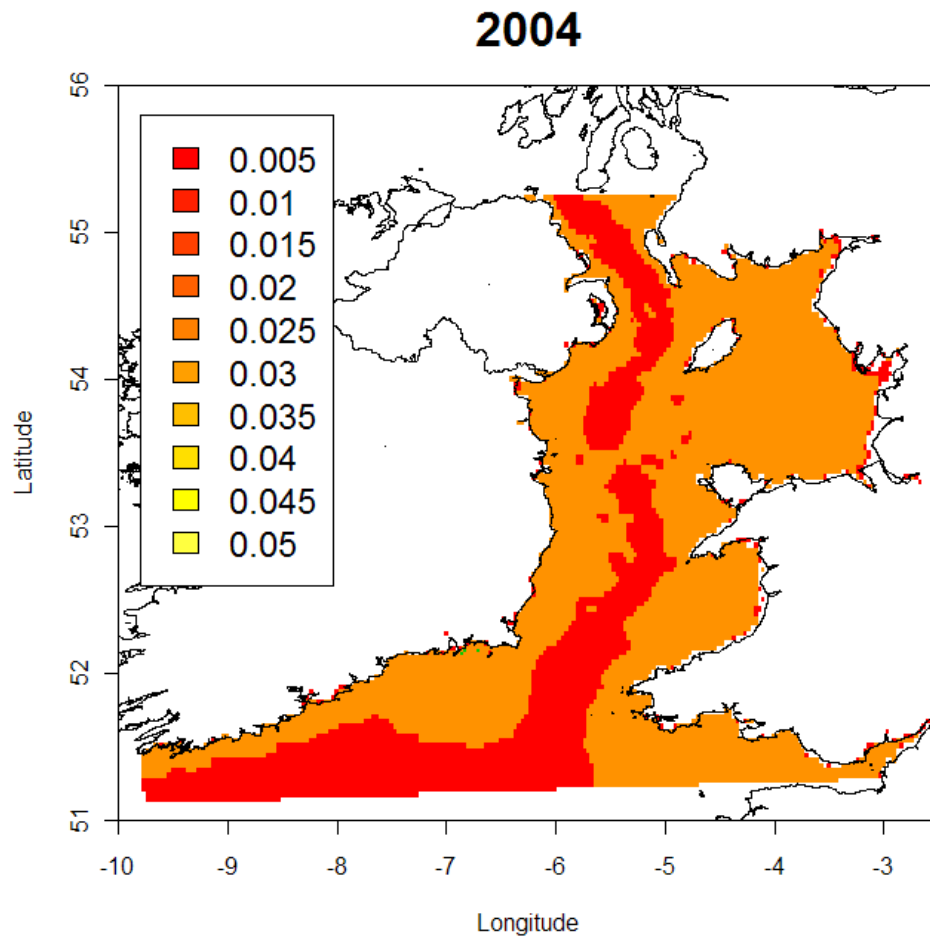


Figure Appendix 3.2. Predictions of minke whale density for the 19th July 2004 according to a negative binomial zero inflated model. Green circles are proportional in area to estimated density of minke whale associated with a particular segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

Bottlenose dolphin

In this case the zero-inflated model predictions and the two-stage modelling predictions were readily reconcilable (Figure Appendix 3.3). However the point estimate for 2004, 800 is considerably less than 4800, the equivalent prediction from the two stage modelling.

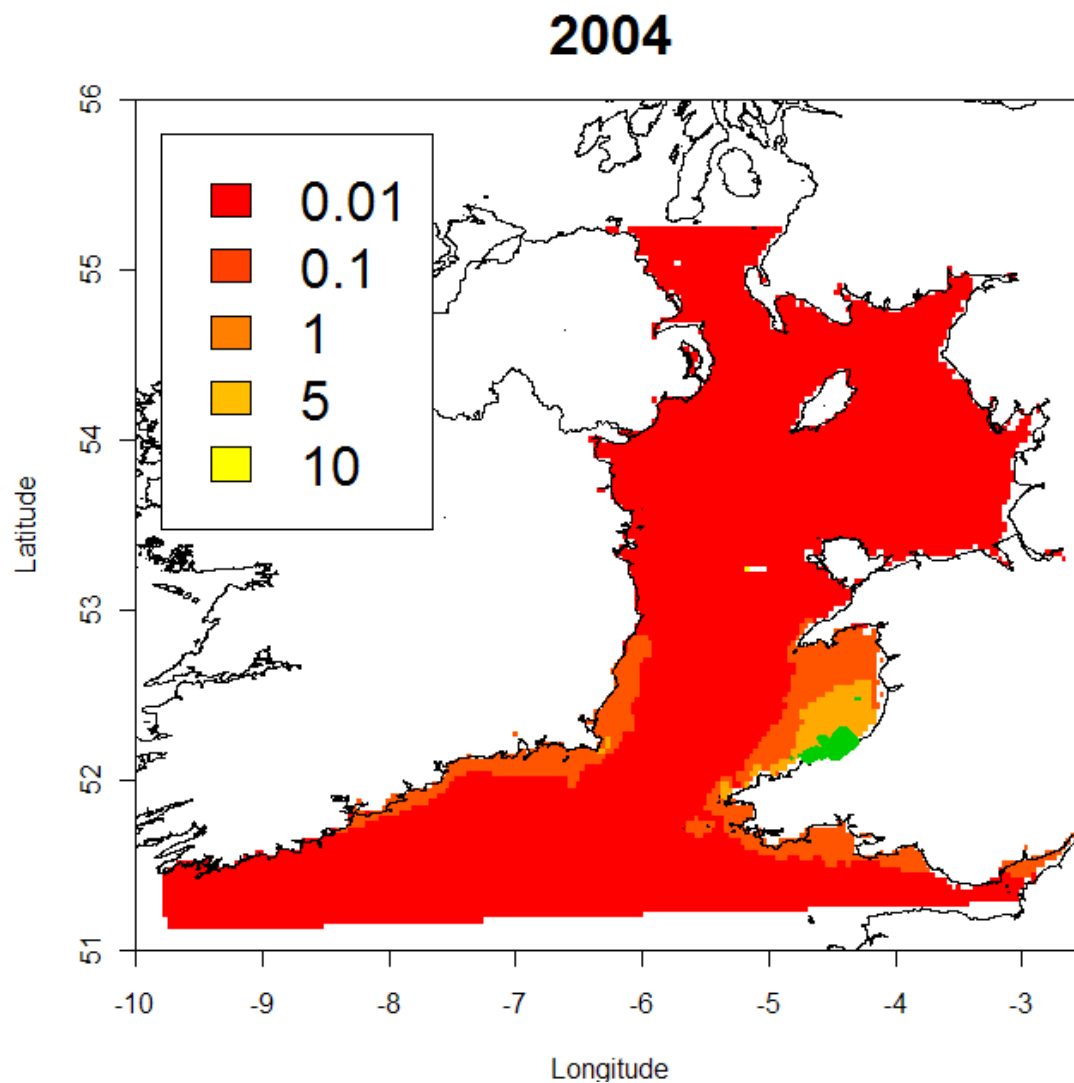


Figure Appendix 3.3. Predictions of bottlenose dolphin density for the 19th July 2004 according to a negative binomial zero inflated model. Green circles are proportional in area to estimated density of bottlenose dolphin associated with a particular segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

Common dolphin

Again the prediction from the zero inflated model whilst explaining a similar proportion of the deviance do not bare any resemblance to the two stage modelling outputs. In contrast to the two stage modelling approach, the zero inflated model has not well characterized the distribution of common dolphins. The point estimate for 2004 is 3100 compared to 37800 from two stage modelling!

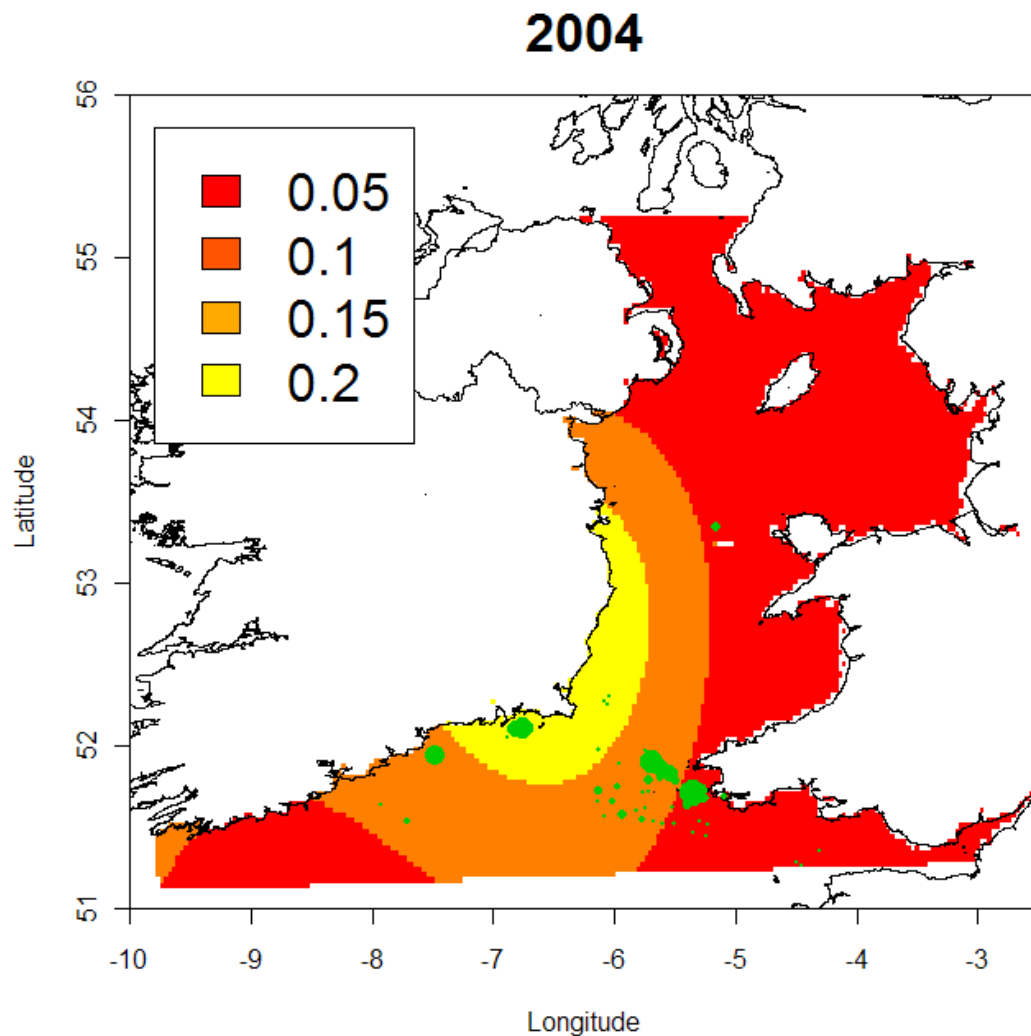


Figure Appendix 3.4. Predictions of common dolphin density for the 19th July 2004 according to a negative binomial zero inflated model. Green circles are proportional in area to estimated density of common dolphins associated with a particular segment locality. Numbers indicate upper bound of colour coded densities (animals/km²).

Discussion

More work really needs to be done to explore the possibilities of zero-inflated models to explore this sort of data. Constrained zero-inflated generalized additive models (COZIGAMs) may provide a useful alternative methodology (Liu & Chan in press) combining both the theoretical desirability of zero-inflated modelling with the practical efficiency of GAMs.

References

Liu, H. & Chan, K-S. (2010) Introducing COZIGAM: An R Package for Unconstrained and Constrained Zero-Inflated Generalized Additive Model Analysis. *Journal of Statistical Software* (in press)