

Density surface fitting to estimate the abundance of humpback whales based on the NASS-95 and NASS-2001 aerial and shipboard surveys

Charles G. M. Paxton¹, M. Louise Burt¹, Sharon L. Hedley^{1,*}, Gísli A. Víkingsson², Thorvaldur Gunnlaugsson² and Geneviève Desportes^{3,4}

¹*RUWPA, The Observatory, University of St Andrews, Fife, KY16 9LZ, United Kingdom*

²*Marine Research Institute, PO Box 1390, IS-121 Reykjavik, Iceland*

³*GDnatur, Stejlestræde 9, DK 5300 Kerteminde, Denmark*

⁴*Faroese Museum of Natural History, Fútalág 40, FR 100 Tórshavn, Faroe Islands*

**Now at The Schoolhouse, Denhead, St Andrews, Fife, KY16 8PA, United Kingdom*

ABSTRACT

Generalized additive models (GAMs) with spatially referenced covariates were fitted to data collected during the 1995 and 2001 Icelandic (shipboard and aerial) and Faroese (shipboard only) components of the North Atlantic Sightings Surveys (NASS-95 and NASS-2001). The shipboard surveys extended from the east coast of Greenland, around Iceland, down to an area along the west coast of Ireland (in 1995) and to the north of the United Kingdom (in 2001). In contrast, the aerial surveys were limited to Icelandic coastal waters only. The aim of the analysis was to predict density, and hence abundance, of humpback whales throughout the survey regions and also to establish if there was any evidence that humpback whale density was related to sea surface temperature or depth.

Fitting GAMs to the 1995 data proved problematic and so various subsets of the data were used in an attempt to improve the model fitting. Such difficulties did not occur with the 2001 data. Confidence intervals (CIs) for the abundance estimates were estimated using bootstrap sampling methods.

The estimated humpback whale abundance for the region covered by the aerial and shipboard surveys in 1995 was 10,521 (95% CI: 3,716–24,636) using all available data and 7,625 (3,641–22,424) if survey blocks with 0 sightings around the Faroes and south of 60° N where no humpback whales were detected were excluded from the analysis. The estimate for the total survey region in 2001 was 14,662 (9,441–29,879). The high upper bounds of the confidence intervals were thought to be caused by a paucity of effort over wide areas of the survey leading to interpolation. Overall, the uncertainty associated with these abundance estimates was approximately equal to, or greater than, that associated with a stratified distance analysis. Given these wide CIs the evidence for a substantial difference in abundance between years was equivocal. However there was evidence to suggest that humpback whales congregated in shallower waters between 6–8°C.

Paxton, C.G.M., Burt, M.L., Hedley, S.L., Víkingsson, G.A., Gunnlaugsson, Th. and Desportes, G. 2009. Density surface fitting to estimate the abundance of humpback whales based on the NASS-95 and NASS-2001 aerial and shipboard surveys. *NAMMCO Sci. Publ.* 7:143-159.

INTRODUCTION

Humpback whales (*Megaptera novaeangliae*) gained initial protection in 1944 (Jefferson *et al.* 1993) and full protection in the North Atlantic and worldwide after the 1956 and the 1965/1966 whaling seasons respectively (Tøn-

nessen and Johnsen 1982). There has been some debate on their rate of recovery since then (Lubick 2003) much of which revolves around estimates of the size of the pre exploitation population (*e.g.* Roman and Palumbi 2003, Mitchell

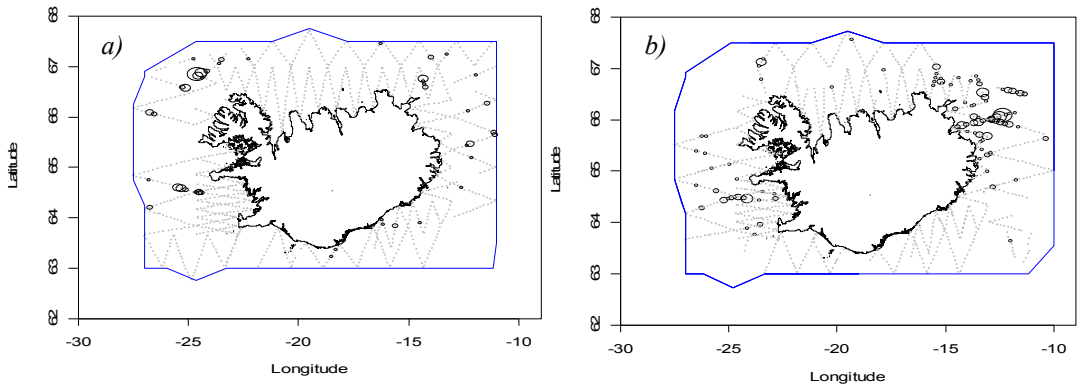


Fig. 1. The realised survey effort (grey lines) and location of sightings (circles) for the aerial surveys in a) 1995 and b) 2001. Observed pod sizes are proportional to the area of the circles.

and Reeves 1983). Estimates for the size of the population in the North Atlantic include 3,500–10,700 (estimated from photo identification: Balcomb and Breiwick 1984, Balcomb *et al.* 1986) and between 9,400–16,400 adults (from photo identification and biopsy, Smith *et al.* 1999). The International Whaling Commission's current estimate of western North Atlantic humpback whale abundance is 11,570 (<http://www.iwcoffice.org/conservation/estimate.htm#table>). Estimates for the population in the late summer and autumn feeding areas (*i.e.* Norway, Iceland, Greenland, New England and Canada) include Pike *et al.* (2009: 4,928 in 2001, coastal waters about Iceland) and Smith *et al.* (1999: between 7,100–8,100 adults). All of the estimates of the last decade are considerably higher than those of Balcomb *et al.* (1986) for 1978 and 1979 (3,503 and 5,005 respectively) albeit not always significantly so. However, no previous work has attempted to combine shipboard and aerial data into 1 analysis.

Humpback whale sightings during both the aerial survey and the shipboard survey of the North Atlantic Sightings Survey in 1995 (NASS-95) occurred in distinct aggregations (Figs 1a and 2a), and were particularly concentrated off north-east Iceland. Broadly similar distribution patterns were encountered during NASS 2001 (Figs 1b and 2b). Other analyses of these data Pike *et al.* (2009: aerial 1995 & 2001, and Pike *et al.* MS 2002: shipboard 1995) estimated abundance using standard distance sampling methods (Buckland *et al.* 2001). These stand-

ard methods resulted in abundance estimates with high variance. Fitting a density surface to geographic covariates has the potential to increase the precision of the abundance estimate, and furthermore, allow elucidation of an ecological interpretation of their distribution (NAMMCO 2003). Data from earlier NASS were not used because of a paucity of data.

Estimates from aerial surveys tend to be negatively biased because of diving whales being missed: temporarily submerged whales are more likely to be seen from a ship because a ship travels more slowly than a plane, thus allowing more time for the animal to return to the surface. This is known as availability bias. One potential way to correct for this bias is to compare the densities from the shipboard survey with the aerial survey. Thus, estimates obtained from the shipboard survey data can be used to correct estimates made from the aerial survey. Burt *et al.* (MS 2003) fitted generalised additive models (GAMs) to the aerial and shipboard data (from 1995 and 2001) separately, and they then compared estimates over the survey regions covered by both modes of survey (*i.e.* shipboard or aerial) to obtain an approximate correction factor for the aerial survey. However, Burt *et al.* (MS 2003) did not combine the data from both survey modes in a single analysis where survey mode could be considered as a discrete factor. Here we combine data from both survey modes in a single analysis and treat survey mode as an additional explanatory variable in the model.

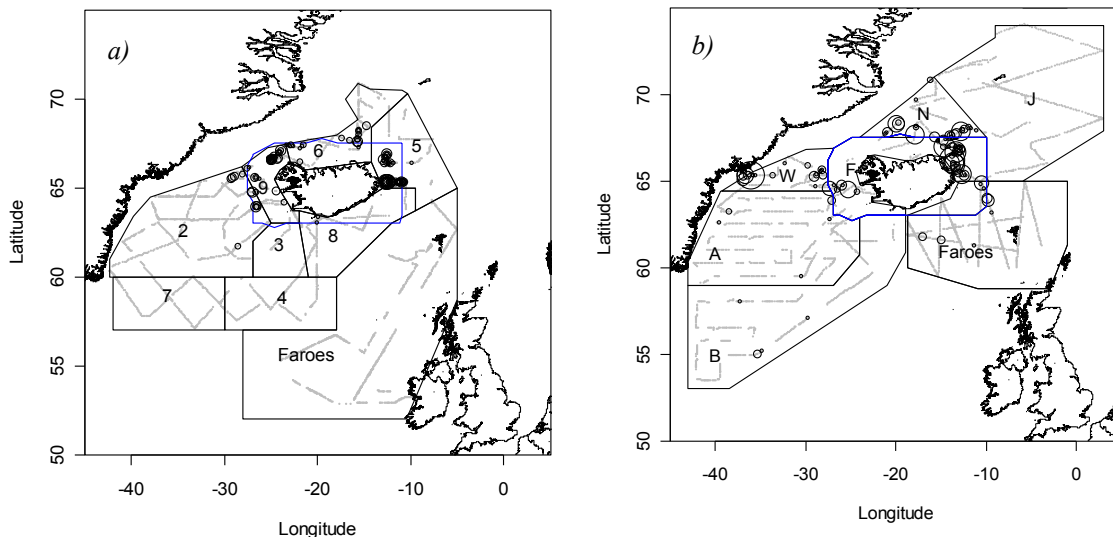


Fig. 2. The realised shipboard survey effort (grey lines) and location of sightings (circles) for a) 1995 and b) 2001. Observed pod sizes are proportional to the area of the circles. Black lines indicate the ship survey blocks and the blue lines indicate the aerial survey region.

In this analysis, the 1995 and 2001 data (initially analysed assuming $g(0)=1$, using separate single platform multiple covariate distance sampling (MCDS, Marques MS 2001) for each survey mode) were generated into spatially referenced density data using the ‘count’ variant of methods developed by Hedley *et al.* (1999). The aerial and shipboard density data were then combined to form a single generalised additive model (GAM) for each year to estimate humpback whale density. This analytical framework allowed the influence of 2 environmental explanatory variables (sea surface temperature and depth) on humpback whale density to be considered as well as spatial reference to latitude and longitude. Finally, all data were combined in order to look at the relative effect of survey year in the GAM as well as depth and sea surface temperature in the absence of latitude and longitude.

METHODS

Survey regions and data

Aerial surveys

The aerial surveys covered Icelandic coastal waters and the same survey design was used in 1995 and 2001, except that in 2001 the survey area was extended from 11° W to 10° W (Fig. 1). The same aeroplane and pilot were used in both years and 2 primary observers searched through bubble windows. The cruise leader and

pilot acted as secondary observers and the cruise leader recorded environmental conditions. The survey was mainly conducted in passing mode but some sightings were closed on to confirm species. For more details, see Pike *et al.* (2009).

Shipboard surveys

The shipboard surveys, conducted in the eastern and central North Atlantic, used 3 vessels in 1995 and 4 vessels in 2001 (Table 2, Víkingsson *et al.* 2009). In 1995, humpback whales were only sighted from 2 of the ships; the Strákur and the Árni Fríðriksson (Pike *et al.* MS 2002). The survey boundaries for the 2 years were substantially different: in 1995, the survey boundaries extended eastwards from 42.25° W to 5° W and northwards from 52° N to 72.75° N (Fig. 2a); and in 2001, the survey area extended south-westerly and north-easterly, but excluding the area extending westwards off the Irish coast (Fig. 2b). Unlike 1995, when all vessels were dedicated cetacean sightings survey ships, the 2001 survey included effort from 2 vessels primarily conducting a redfish survey in addition to 2 dedicated cetacean survey vessels. The data from the 1995 and 2001 shipboard surveys were collected with a view to conducting a conventional design based line transect analysis (e.g. Buckland *et al.* 2001). Although the transects from the redfish survey (from the 2 named vessels above) were not designed for

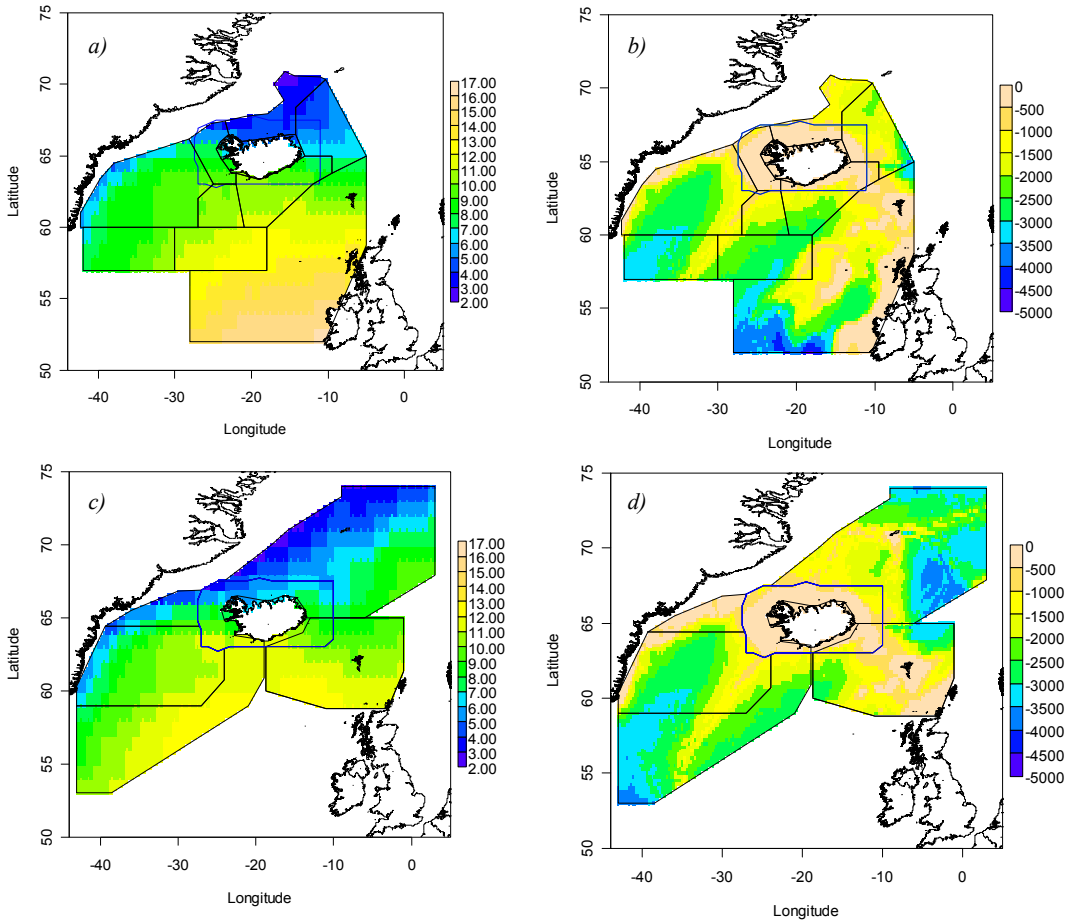


Fig. 3. Environmental covariates used in the analysis. *a)* mean sea surface temperature (°Celsius) 18th June–6th August 1995, *b* and *d)* Depth (m) & *c)* mean sea surface temperature 17th June–5th August 2001. Black lines indicate the ship survey blocks and the blue lines indicate the aerial survey region.

the purpose of estimating whale density, they do achieve a spatial coverage that would be expected to be adequate for estimating whale density from a model based approach. For more details for 2001, see Víkingsson *et al.* (MS 2002).

Detection function estimation

In conventional distance sampling the probability of detection depends only on the perpendicular distance of the sighting to the transect and at 0 perpendicular distance is assumed to be 1 (denoted by $g(0) = 1$). In this analysis, the effects of covariates, other than perpendicular distance, were incorporated into the detection function model. This was achieved by setting the scale parameter in the model to be an exponential function of the covariates (Marques MS 2001). Thus the probability of detection be-

comes a multivariate function, $g(y,v)$, representing the probability of detection at perpendicular distance y and covariates v ($v = v_1, \dots, v_Q$ where Q is the number of covariates). Using either a hazard rate or half normal detection function, the covariates were incorporated via the scale term, σ , where for sighting k , σ has the form: (1)

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

where β_0 and β_q ($q=1, \dots, Q$) are parameters to be estimated. 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.

A stepwise forward selection procedure was used (starting with a model containing per-

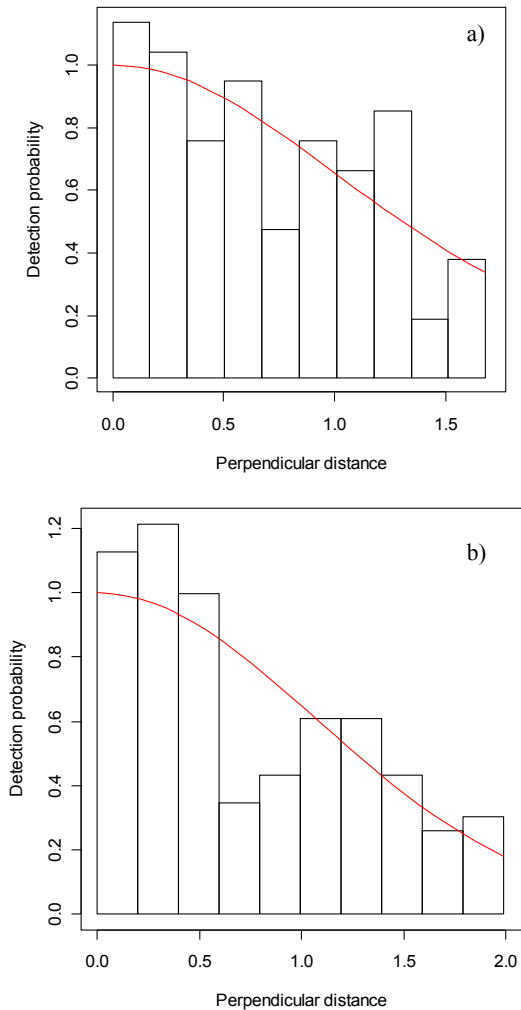


Fig. 4. Perpendicular distance (km) distribution and fitted half normal function (red line) for a) NASS-95 aerial survey and b) NASS-2001 aerial survey.

longitude (larger pods occurring to the east of the survey region) but longitude only explained 3% of the deviance (the difference in log likelihood between the model under consideration and a full model consisting of as many parameters as data (Dobson 1990)) and this may have been an artefact of the scarcity of sightings to the far west and so was not considered further. As there was no consistent evidence of a trend in pod size over the survey area (Burt *et al.* MS 2003), a modified version of the ‘count model’ of Hedley *et al.* (1999) was used to model the trend in spatial distribution of humpback whale pods. Transects covered during the survey were divided into small sections, or segments, such that the sighting and geographic conditions did not change considerably within a segment. The response variable for the model was calculated from the estimated number of individuals in each segment, \hat{N}_i , where the subscript denotes the i^{th} segment. This was calculated using an estimator similar to the Horvitz Thompson estimator (Horvitz and Thompson 1952), as follows: (2)

$$\hat{N}_i = \sum_{j=1}^{n_i} \frac{s_j}{\int \hat{g}_{ij}(y, \nu) p(y) dy} \quad i = 1, \dots, T,$$

where, for segment i , $\int \hat{g}_{ij}(y, \nu) p(y) dy$ is the estimated probability of detection of the j^{th} detected pod, n_i is the number of detected pods in the segment and s_j is the size of the j^{th} pod. The total number of transect segments is denoted by T . By assumption, $p(y)$, the probability density function of actual perpendicular distances is uniform up to the truncation distance. This is satisfied provided transects are randomly located with respect to the distribution of whales.

Having obtained the estimated number of individuals in each segment, the density, \hat{D}_i , was simply given by \hat{N}_i / a_i , where a_i is the area of segment i . Segment area was calculated as the length of the segment multiplied by twice the truncation distance used to model the detection function. There is no objective way of choosing segment length and so a variety of segment lengths were tried in the range of 5–13 km. Eventually 9 km was selected as an appropriate compromise be-

pendicular distance only) to decide which covariates to include in the model, with a minimum Akaike’s Information Criterion (AIC) inclusion criterion. All model selection was performed in *Distance* (v4.0; Thomas *et al.* 2002), then the final selected models were refitted using a set of customized functions in R (Ihaka & Gentleman 1996). This facilitated estimation of variance within R —(see below).

Modelling whale density

There was no evidence of a trend in pod size over the survey area in 1995 (using a GAM with logarithm of pod size as the dependent variable and longitude and latitude as explanatory variables, $P=0.89$). Note that such a model cannot be used for prediction. In 2001, there was a significant trend ($P<0.01$) in pod size with

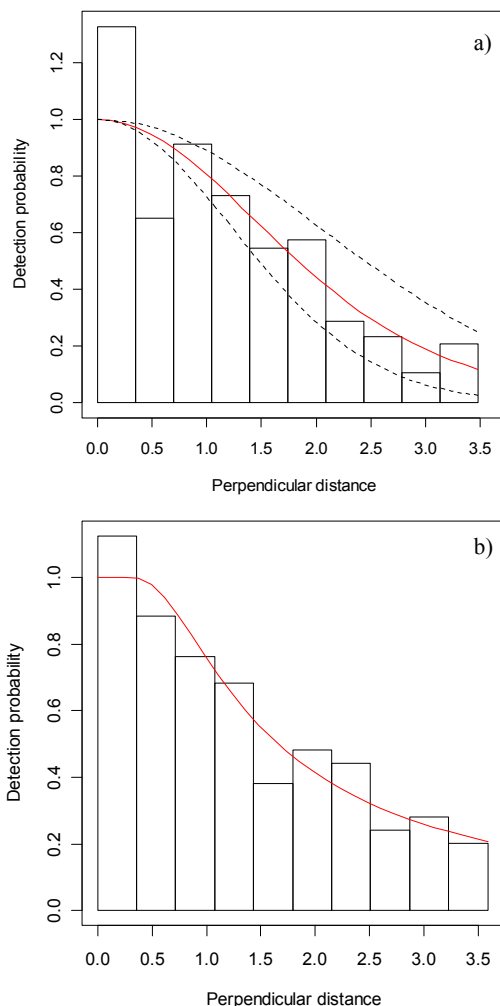
Fig. 5. Overall perpendicular distance (km) distribution and detection function for a) NASS-95 shipboard survey and b) NASS-2001 shipboard survey. The red line is the average fitted detection function. In the case of figure a) this is averaged across cloud cover and vessel. Dashed black lines in a) indicate the detection functions for each vessel.

tween maximising the ratio of non zero to zero segments, maintaining environmental resolution and giving some measure of spatial independence (see results). Density \hat{D}_i was then modelled as a function of K spatially referenced covariates, z , using a GAM with a square root link function: (3)

$$E(\hat{D}_i) = \left(\theta_0 + M_i + \sum_{k=1}^K f_k(z_{ik}) \right)^2$$

where θ_0 is the intercept parameter and the f_k are one dimensional smooth functions (thin plate regression splines) of the K spatial covariates. The parameter M was a factor variable which indicated whether the survey mode was a ship or a plane. The degrees of freedom of the smooth functions control the ‘wiggleness’ of the smooth and were estimated from the data. Two way interactions between the spatial covariates were also considered for inclusion in the model via two dimensional smooths (Wood 2003). In the context of this analysis a two dimensional (2D) smooth of latitude and longitude may be better at explaining spatial variation than the sum of the one dimensional (1D) effects.

A variety of statistical families were considered to model the distribution of \hat{D} . The estimated number of pods in each segment did not follow a typical exponential family distribution, such as the Poisson distribution, and quasi-likelihood methods were more appropriate to fit the model. Quasi-likelihood methods do not require a particular error distribution to be assumed, but the general form of mean variance relationship of the observations must be established. In this case it was assumed the data were over dispersed.



The covariates considered in the analysis were longitude (*Lon*) and latitude (*Lat*), sea surface temperature (SST) and depth (Depth). Sea surface temperatures were obtained from the National Oceanic and Atmospheric Administration (NOAA) and were an updated set based on the analysis of Reynolds *et al.* (2002) at 1 degree and weekly resolution. Depths were obtained from the ETOPO5 5 minute resolution relief data available from NOAA (<http://www.ngdc.noaa.gov/mgg/global/seltopo.html>). Temperatures and depths were associated to effort segments by finding the closest point in the temperature and bathymetry data to the midpoint of the effort segments using great circle distances

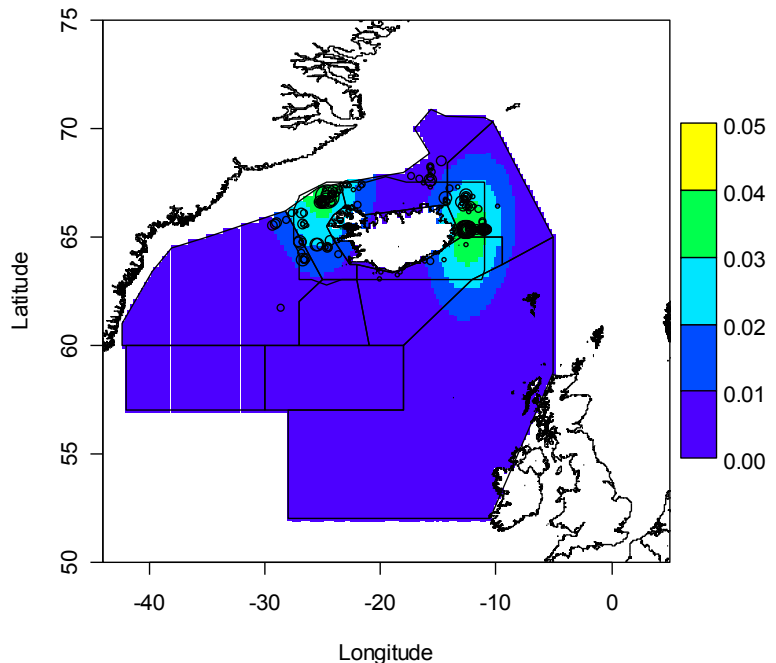
Block	Size	N	Subset 1 All blocks	Subset 2 Faroes and 7 omitted	Subset 3 Faroes and 7 omitted plus additional transects	Subset 4 Faroes, 7 and 4 omitted	Subset 5 Faroes and 4 omitted	Subset 6 Faroes and 4 omit- ted plus additional transects		
			Segments	L	Segments	L	Segments	L	Segments	L
Faroes	1,148,071	0	660	3,061	0	0	693	0	141	693
2 (W Iceland)	75,618	51	232	842	232	842	232	842	232	842
3 (S Iceland)	92,048	0	40	289	40	289	40	289	40	289
4 (Far S Iceland)	232,634	0	175	1,168	175	1,168	0	0	0	0
5 (NE Iceland)	161,844	113	260	807	260	807	260	807	260	807
6 (N Iceland)	117,456	31	320	1,501	320	1,501	320	1,501	320	1,501
7 (Far SW Iceland)	232,634	0	227	1,538	0	0	227	1,538	227	1,538
8 (SE Iceland)	191,213	2	271	1,491	271	1,491	271	1,491	271	1,491
9 (E Greenland)	422,891	17	821	3,556	821	3,556	821	3,556	821	3,556
Aerial	39,940	76	1,159	10,013	1,159	10,013	1,159	10,013	1,159	10,013
Total (including omitted blocks)	2,714,349	290	4,165	24,266	3,278	19,668	20,620	18,500	3,330	20,037
										3,471
										20,730

(and additionally, time for temperature). Maps of depths and temperatures over the 2 survey periods are shown in Fig. 3. Generalised cross validation (GCV) implemented in the *mgcv* package (v. 1.1.7, Wood 2001) in *R* (v. 1.9.1) was used for covariate selection, augmented with diagnostic plots, using the principles described in Wood (2001) with the additional criterion that a term must explain an additional 4% of the deviance given other variables in the model. All covariates were considered for inclusion in the model as 1D smooths of untransformed covariate values and also as the logarithm of the covariate. In addition, 2D smooths of Lat and Lon and SST and Depth were considered for inclusion into the GAM. Initially a maximum of 6 degrees of freedom (7 knots) were allowed in the selection of 1D smooths and up to 13 degrees of freedom (14 knots) were allowed in the case of 2D smooths, thus allowing moderate flexibility but reducing the likelihood of spurious overfitting (fitting unnecessarily complicated functions) especially as GCV has been shown to cause overfitting (Mackenzie and Walker, submitted). Survey mode (*Mode*) was included in all models. Due to gaps in search effort along transects, effort could not always be split into segments of the desired length (see later). Therefore, the size of each segment varied and so the model was weighted by segment area. The final model was used to predict density of humpback whales throughout the survey regions. Whale abundance was estimated by numerically integrating under this predicted whale density surface. As survey mode was included in each model, abundance was predicted assuming shipboard survey mode.

Block	Area	N	Segments	L
Faroes	493,731	12	644	4,054
A (S Iceland)	661,807	5	393	2,581
B (SW Iceland)	513,252	4	583	4,227
F (around Iceland)	297,039	162	268	1,975
J (Jan Mayen)	499,372	0	416	3,070
N (N Iceland)	106,896	53	161	1,027
W (W Iceland)	98,137	37	161	1,107
Aerial	-*	146	1,073	9,258
Total	2,685,220	450	3,699	27,299

Table 1a and 1b. For each survey block, area in km², number of sightings after truncation (N), number of segments and realised search effort (L) in km for a) 1995 and b) 2001. Search effort is realised effort after discarding effort conducted in Beaufort Sea States greater than 6 and anomalous records.

Fig. 6. Estimated whale density for NASS-95 based on the combined aerial and ship surveys. Density estimates are on a scale 0 to 0.05 whales/km² (see bar).



In the case of the 1995 data, the large number of segments with 0 sightings made fitting the GAM problematic and so several subsets of the data were considered. There were no sightings of humpbacks in the blocks surrounding the Faroe Islands, to the far south and southwest of Iceland (blocks Faroes, 7 and 4, see Fig. 2a), and so these were excluded from some of the analyses. The subsets of data that were considered were

1. data were used in entirety,
2. blocks Faroes and 7 omitted,
3. blocks Faroes and 7 were omitted but some transects (close to the included blocks) from the omitted blocks were included to stabilise the GAM at the edge of the survey region,
4. blocks Faroes, 7 and 4 omitted,
5. blocks Faroes and 4 omitted, and
6. blocks Faroes and 4 omitted but some transects (close to included blocks) from the omitted blocks were included to stabilise the GAM.

Density, and hence abundance, was not estimated in the omitted blocks; there were no sightings in these blocks and so were assumed to have 0 density. The number of segments and search effort are given in Table 1. In none of the cases were the distributional characteristics ideal and a small minority of the model fits to the bootstrap re-samples did not converge to a

definitive set of parameters, potentially leading to an inflated estimate of variance (see below).

In addition to the analyses performed to estimate abundance, a further analysis was conducted which considered only survey mode, *SST* and *Depth*; latitude and longitude were explicitly excluded from the model as they may obscure any relationship between density and the environmental variables. The aim here was to produce models that explained rather than described the density surface. The most complete data set from the 1995 analysis (subset 1) was used in this analysis. Finally, all the data from both years were combined in order to look at the effect of year (*Year*), *Mode*, *SST* and *Depth* (the latter 2 covariates being considered both as 1D smooths and a single 2D smooth).

Variance estimation

The variance of the predicted abundance was estimated using a non-parametric bootstrap (Hedley and Buckland 2004, Efron and Tibshirani 1993) which combined both components of the analysis; detection function estimation and modelling of whale density. Model selection (in terms of variable selection) was not part of the bootstrapping process but smooths with different degrees of smoothing could be selected up to the same maximum number of knots as used

Table 2. Selected terms in the GAMs. $s(\cdot)$ denotes inclusion of the covariate as a smooth function in the model. Initially, smooth functions were allowed 7 knots for 1D smooths and 14 knots for 2D smooths. 'Abundance model' refers to the model used to estimate abundance whereas 'explanatory model' refers to the model used to examine the relationship between density and environmental variables.

1995	Formula	% deviance explained
Subset 1 (all blocks):		
Abundance model	$Mode + s(Lon, Lat)$	32.0
Explanatory model	$Mode + s(SST, Depth)$	32.4
Subset 2		
Abundance model	$Mode + s(Lon, Lat)$	30.2
Explanatory model	$Mode + s(SST, Depth)$	31.8
Subset 3		
Abundance model	$Mode + s(Lon, Lat)$	32.8
Explanatory model	$Mode + s(SST, Depth)$	31.5
Subset 4		
Abundance model	$Mode + s(Lon, Lat)$	32.1
Explanatory model	$Mode + s(SST, Depth)$	32.4
Subset 5		
Abundance model	$Mode + s(Lon, Lat)$	30.5
Explanatory model	$Mode + s(SST, Depth)$	33.7
Subset 6		
Abundance model	$Mode + s(Lon, Lat)$	32.4
Explanatory model	$Mode + s(SST, Depth)$	31.2
2001		
Abundance model	$Mode + s(Lon, Lat)$	31.7
Explanatory model	$Mode + s(SST, Depth)$	26.9
Combined 1995 and 2001		
Explanatory model only	$Mode + s(SST, Depth)$	23.6

in the original model. The bootstrap samples were generated by repeatedly sampling 'transect legs', sampling with replacement. This assumed that the transect legs were independently and identically distributed. Legs were sampled within blocks so there was good coverage of the data within each bootstrap sample and the bootstrap sample contained the same number of transect legs in a block as the original data.

One thousand bootstrap samples were generated and for each, a density surface and total abundance was predicted. These abundance estimates were ordered according to size and the confidence interval (CI) was given by the 2.5% and 97.5% quantiles.

A disadvantage of using the non-parametric bootstrap in this way is that the bootstrap samples do not reproduce the spatial coverage of the original survey data. However, since there were a large number of transect legs (195 in 1995 and 206 in 2001) and transects were selected within blocks, bias from using this procedure was expected to be small. An alternative method of estimating variance would have been to use a parametric bootstrap (as proposed in Hedley *et al.* 1999), but as

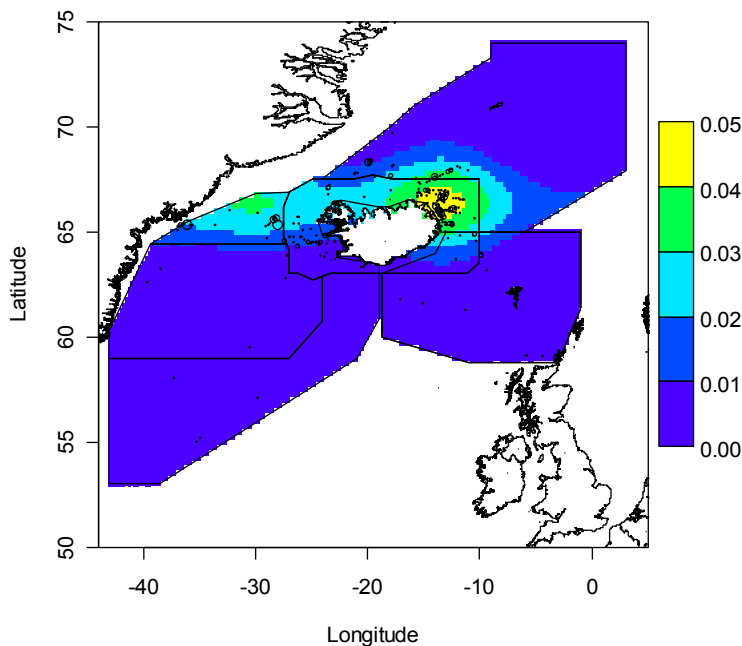


Fig. 8. Estimated whale density for NASS-2001. Density estimates are on a scale 0 to 0.05 whales/km² (see bar).

Table 3a.

Block	Subset 1 All blocks		Subset 2 Faroese and 7 omitted		Subset 3 Faroese and 7 omitted plus additional transects		Subset 4 Faroese, 7 and 4 omitted		Subset 5 Faroese and 4 omitted		Subset 6 Faroese and 4 omitted plus ad- ditional transects	
	Density	Abundance	Density	Abundance	Density	Abundance	Density	Abundance	Density	Abundance	Density	Abundance
Faroese	0.0017	1,932 411–7,015	–	–	–	–	–	–	–	–	–	–
2 (W Iceland)	0.0216	1,631 552–2,682	0.0233	1,763 613–2,752	0.0222	1,676 567–2,804	0.0232	1,752 812–2,977	0.0266	2,008 647–2,897	0.0259	1,962 647–2,980
3 (S Iceland)	0.0019	172 39–511	0.0004	41 12–560	0.0010	90 20–362	0.0004	41 21–527	0.0056	519 44–888	0.0045	418 35–661
4 (Far S Iceland)	0.0002	41 8–317	0.0002	41 10–770	<0.0001	5 5–292	–	–	–	–	–	–
5 (NE Iceland)	0.0154	2,501 485–7,002	0.0143	2,319 465–8,873	0.0143	2,324 432–7,115	0.0143	2,320 491–6,777	0.0133	2,147 419–7,999	0.0125	2,021 432–5,300
6 (N Iceland)	0.0076	890 518–2,044	0.0069	807 462–1,893	0.0074	869 495–2,299	0.0068	794 449–2,142	0.0099	1,160 568–2,181	0.0101	1,187 606–2,223
7 (Far SW Iceland)	<0.0001	16 3–414	–	–	–	–	–	–	0.0002	44 17–478	0.0001	33 13–494
8 (SE Iceland)	0.0104	1,989 303–5,129	0.0082	1,562 349–7,238	0.0092	1,762 396–5860	0.0076	1,461 293–6,796	0.0082	1,576 144–3,854	0.0073	1,392 153–3,231
9 (E Greenland)	0.0017	704 326–1,875	0.0016	664 300–1,760	0.0013	556 287–1,734	0.0016	684 353–1,600	0.0023	963 532–1,749	0.0023	983 524–1,890
Coastal Iceland	0.0161	645 192–1,243	0.0146	584 211–1,319	0.0180	720 223–1,478	0.0143	572 228–1,493	0.0218	872 166–1,884	0.0225	898 171–1,903
Total	0.0039	10,521 3,716–24,636	0.0029	7,780 3,503–21,521	0.0029	8,002 3,356–16,999	0.0028	7,625 3,641–22,424	0.0034	9,289 3,143–19,107	0.0033	8,895 3,171–17,014

Table 3 b.

Block	Mean density (whales/ km ²)	Abundance	95% CI
Faroese	0.0022	1,011	562–2,377
A (S Iceland)	0.0007	444	167–2,873
B (SW Iceland)	0.0012	607	311–2,016
F (around Iceland)	0.0205	6,242	3,238–11,580
J (Jan Mayen)	0.0049	2,687	1,691–7,810
N (N Iceland)	0.0147	1,528	805–3,949
W (W Iceland)	0.0220	2,142	895–3,628
Total	0.0059	14,662	9,441–29,879

Table 3a and 3b. Estimates of mean density (whales/ km²), abundance and 95% confidence intervals bootstrap (top) by block for a) 1995 and b) 2001.

currently implemented, their approach does not adequately capture the correlation between adjacent, or close, segments of transect, and thus appears to under estimate the true variance.

Many of the models fitted to the bootstrap samples did not converge because of the large number of segments with 0 sightings. Removing these abundance estimates may have biased the CI, thus they were retained providing a robust estimate of variation. Non-convergence was avoided to an extent by reducing the maximum degrees of freedom allowed for the smooth function as described above. However, some non-converging models fitted so badly that the estimated fitted deviance was stated as negative. Such replicates (1–3 per bootstrap sample) were typically associated with extremely high nonsensical estimates of abundance. The abundance estimates associated with these models were replaced by randomly choosing abundance estimates associated with models which had a positive explained deviance.

RESULTS

Aerial survey detection functions

NASS-95

There were 89 sightings of humpback whales in total. For robust modelling of the detection function, a general rule of thumb is to discard approximately 10% of the sightings with furthest perpendicular distance (Buckland *et al.* 2001). However, because of a long tail in the distribution of perpendicular distances, sightings were truncated at a distance of 1,700 m (leaving 76 sightings) to ensure robustness during bootstrapping. For the rest of the 1995 analysis, aerial sightings beyond this distance were discarded. Often in aerial surveys, visibility directly below the aircraft is limited so that detection probability for whales surfacing on the transect is not certain. This did not appear to be the case for these data probably because the aircraft was installed with bubble windows. Covariates considered for inclusion in the model for the detection function were: Beaufort (as a factor with 2 levels; Beaufort 1 and 2); percentage of cloud cover; sightability (a 3 level factor with levels ranging from perfect; perfect to good; good to worse). Initially, glare was also considered but due to ambiguous interpretations of its effect, it was excluded. Both half-normal and hazard rate functional forms were considered, and forward stepwise model selection was by AIC (Akaike's Information Criterion), starting with a model which included perpendicular distance only. No covariate was found to have a strongly significant effect, and thus the final model selected included only perpendicular distance, with half-normal form (Fig. 4a). See Pike *et al.* (2009) for a further treatment of this data.

NASS 2001

Despite slightly less search effort in 2001 compared to 1995 there were 161 sightings of humpback whale pods in 2001, nearly twice as many as in 1995. The distribution of their perpendicular distances is shown in Fig. 4b using a truncation distance of 2,000 m (discarding 11% of sightings). Covariates considered for inclusion in the detection function model were: Beaufort Sea State (as a factor with 3 levels, 1, 2 and 3 or more); sightability (good and moderate/poor); percentage cloud cover; and pod size (as a factor with 3 levels: 1, 2 and 3 or more). Ac-

ording to AIC, the best models were those with either Beaufort or sightability included as covariates. However, in both cases, the estimated probability of detection increased as Beaufort increased or sightability decreased—contrary to expectation but possibly due to small sample sizes under the worst conditions. Therefore it was decided to fit a detection function model based on perpendicular distance only (Fig. 4b, red line)—the best model being half normal.

Shipboard survey detection functions

NASS-95

There were 215 sightings of humpback pods within a perpendicular distance of 3,600 m. The histogram of perpendicular distances of these sightings is shown in Fig. 5a. Covariates considered for inclusion in the model for the detection function were: survey block; platform; cloud cover; pod size (as a factor with levels 1, 2 and 3 or more); vessel and Beaufort Sea State (considering it both as a continuous variable and as a factor with 2 levels—less than 3.5 and greater than 3.5). The model chosen was a half-normal with cloud cover and vessel included as covariates: the probability of detection increased in cloudier conditions, and was higher for the observation ship *Strákur* compared to the *Árni Friðriksson*. The average fitted detection function across all the covariates is shown in Fig. 5a (red line).

NASS 2001

With a truncation distance of 3,600 m, 273 sightings of humpback pods remained, and their histogram of perpendicular distances is shown in Fig. 5b. Covariates considered for inclusion in the model for the detection function were as for 1995 except for survey block, plus 3 additional covariates: sightability, swell and visibility. No improvement in model fit was gained by the inclusion of covariates in the model. Thus, a hazard rate model including perpendicular distance only was selected.

Modelling of estimated density

NASS-95

The large number of transects (and hence segments) with no sightings meant that the density data had awkward distributional properties. A modal segment length of 9 km was chosen as a compromise between a length that would represent the environmental characteristics of the

Table 4. Comparison of the models, predicted mean densities (whales /km²) and 95% bootstrap CI for the survey region common to 1995 and 2001 surveys (size of common region is 1,725,204 km²) using models fitted to either the 1995 or 2001 survey data. s(.) denotes inclusion of the covariate as a smooth function in the model.

Survey data	Terms in model	Mean density (whales /km ²)	Total abundance	95% CI
1995 (Subset 1)	None ¹	0.0053	9,138	4,562–17,511
1995 (Subset 1)	<i>Mode</i> + <i>s(Lon, Lat)</i> ²	0.0057	9,810	3,439–21,457
2001	None ¹	0.0085	14,681	7,706–19,624
2001	<i>s(Lat)</i> + <i>s(Lon)</i>	0.0066	11,319	6,651–21,214

¹i.e. a flat density surface. Density based on mean density of segments over entire 1995 or 2001 survey.

²i.e. based on subset 1 in Table 2.

raw data and a length that would lead to nominal spatial independence. Due to changes in survey effort and environmental conditions, not all segments had the same area – the range in segment areas was 0.27 - 97 km² with a median of 30 km².

For most of the 1995 subsets, a 2-dimensional smooth of latitude and longitude was selected along with survey mode (aerial or ship) (Table 2) and the total explained variance for all the models was between 30 and 34%. The point estimates were similar as were the 2.5% CI estimates. The largest differences between the subsets occurred in the 97.5% CI estimates. The best model, in terms of smallest CI, (Table 3a) were those analyses which omitted 2 of the blocks with no sightings but included a few transects to ensure the predictions were ‘tied down’ at the periphery of the prediction region (subsets 3 and 6). The abundance estimate based on the entire survey region was 10,521 (bootstrap CI: 3,716–24,636). Figure 6 gives the density surface. Abundance based on an inner core region (*i.e.* assuming 0 density in 2 of the 4 blocks with no sightings) was 7,780 (3,503–21,521) or 8,002 (3,356–18,999) depending whether additional transects were used to tie down the GAMs.

To allow a direct comparison of humpback whale abundances across years, the 1995 data (subset 1) was used to estimate abundance over a common region which was surveyed by both the 1995 and 2001 survey. In addition, to approximate a design-based estimate, abundance in this common region was estimated using the mean density from the 1995 data—this is equivalent to fitting a flat density surface where the density is equal to the mean density throughout the survey.

The abundance in the common region for 1995 was 9,138 (4,562–17,511) animals based on a constant mean density and 9,810 (3,439–21,457) based on the spatially varying density model.

The explanatory models (excluding *Lat* and *Lon* as potential candidate covariates) fitted to the 1995 subsets all included *Mode* as well as *SST* and *Depth* as a 2D smooth. In some cases these models explained more of the deviance than the descriptive model (Table 2). However, they also had higher GCV scores associated with them and therefore the predictive models (*i.e.* the ones with *Lat* and *Lon*) were used to predict abundance. The explanatory models were not selected in the model selection procedure for the predictive model because of the huge contribution of *Lat* (approx. 19% deviance depending on data used) to the explained deviance. Unsurprisingly *Lat* was closely correlated with *SST* ($r = 0.93$) therefore *SST* was not selected when *Lat* was a candidate variable.

NASS 2001

Again a modal segment length of 9 km was used. Some segments could not be amalgamated if environmental conditions changed and so again there was a wide range in segment areas. The range in segment areas was 0.001 to 97 km², with a median of 42 km². The final descriptive model consisted of *Mode* and 1D smooths for *Lon* and *Lat*. Again the shipboard densities were typically twice that reported from the aerial survey. The total explained deviance was *ca.* 32%. There was no evidence for any spatial autocorrelation in the residuals of the model.

Table 5. Comparison of abundance estimates from this paper with estimates obtained using alternative methods.

Region	Method used	Abundance	95% CI	Reference
Ship 1995	Line transect analysis	13,900	3,900–29,000	Pike (2002)
Ship 1995	Line transect analysis and density surface fit.	10,521 (subset 1)	3,716–24,636	This paper
Aerial 2001	Line transect analysis followed by $g(0)$ correction	4,928	1,926–12,611	Pike <i>et al.</i> (2009)
Aerial 2001	Line transect analysis and density surface fit.	6,242	3,238–11,580	This paper

The predicted densities of humpback whales for 2001 are shown in Fig. 7. The densities were predicted using the survey region for 2001. The estimated abundance of humpback whales in the survey region was 14,662 (9,441–29,879, Table 3b).

Two additional abundance estimates were made using the model fitted to the 2001 data for the survey region common to 1995 and 2001 to allow comparison between the 2 surveys; using the mean density throughout the survey region, and by predicting density for the common region using the model fitted to the 2001 data. The predicted values for 2001 were higher than 1995 but there was no significant difference between them (Table 4).

The explanatory model for 2001 again consisted of *Mode* and a 2D smooth of *Depth* and *SST* (Table 2) explaining *ca* 27% of the deviance. Predicted abundances for whales using ship as the *Mode*, were about 1.5 times that when using aerial as a *Mode*.

Explanatory model for 1995 and 2001 data combined

To investigate the underlying factors influencing humpback distribution between the 2 years, the data for 1995 and 2001 were combined into a single model and all data, including data from blocks where no whales were seen, were included. The explanatory variables under consideration were 1D smooths of *Depth* and *SST* (also considered as a 2D smooth), with *Mode* and *Year* as factors. The model selected included *Mode* and a 2D smooth of *Depth* and *SST*. This model explained 23.6% of the deviance. The more complex model containing

Year as a factor was associated with a higher GCV score but *Year* although significant explained only an additional 0.4% of the deviance.

The pattern of the responses, like the 2001 data alone, again suggested an optimal sea surface temperature of around 6–8° C.

Comparison with earlier studies

Table 5 provides a comparison of the estimates obtained in this paper with those of other analyses of the same data. The estimates were broadly similar but CIs were wider using the methods outlined in this paper.

DISCUSSION

The NASS aerial and shipboard surveys of 1995 and 2001 all recorded distinct clumped aggregations of humpback whales (Figs 1 and 2) with the greatest numbers to the east and, to a lesser extent, west of Iceland. This pattern has been reflected in the predicted densities (Figs 6 and 7). The analysis methods used in this paper have provided an alternative approach to estimating abundance, yielding estimates which are comparable to those previously obtained using design based estimation methods but with higher variances (see Table 5); Pike *et al.* (MS 2002) estimated 13,900 whales over the 1995 survey region whereas the arguably most appropriate GAM analysis for 1995 suggested 10,521. The Years of the North Atlantic Humpback (YONAH) survey predicted point estimates varying between 9,400 and 16,400 adults (dependent on method and data used) for the number of humpbacks for the entire North Atlantic basin (Smith *et al.* 1999), however they under sampled the

high density region off east Iceland (Pike *et al.* MS 2002). In contrast, Pike *et al.* (2009) estimated approximately 4,928 humpbacks from the 2001 aerial survey area data set (corrected for under detection on the transect line but not for availability) and this compares favourably to the estimate here of 6,242 for the same area assuming a ship based survey (with overall assumed higher detection rates) but with a much lower variance using a conventional distance sampling model not fitting a density surface.

When fitting the models to the same survey region the point estimates for 1995 and 2001 (9,810 and 11,319, respectively, Table 4) seemed to indicate a small increase in humpbacks from 1995 to 2001, although these results were not significantly different. However the combined analysis for 1995 and 2001 indicated a significant, albeit negligible, effect of year suggesting there had been a real increase in numbers between the surveys. This was in agreement with analysis of the aerial only region by Pike *et al.* (2009) where a significant upward trend was shown. Whether this increase reflects migration into the survey region or an overall increase in numbers is unclear. However there have been consistent increases in abundance in this area since at least 1979 (Sigurjónsson and Gunnlaugsson 1990) which does imply a real increase has taken place, as presumably there is a limit on the number of individuals in a constant population that could move into the area.

Whilst the bootstrap procedure performed adequately to obtain 95% CI of the abundance estimates there is still a possibility of future improvement. Firstly, while latitude and longitude are a useful coordinate system for humans, they may not be an adequate proxy for unknown variables to explain whale density. However, the explanatory models based only on *Mode*, *SST* and *Depth* did not produce narrower confidence intervals. Secondly, extrapolation with highly flexible models such as GAMs, (see below) could have lead to the spurious results in the bootstrap estimates as well as, to a more limited extent, spurious interpolations. This was suggested by the minor narrower CIs of the models based on data which contained additional transects used to tie down the GAMs (Table 3). Thirdly, the whales may be too patch-

ily distributed to be adequately modelled within a GAM framework more suited to smoothly changing surfaces. Using a 2-stage procedure where presence/absence is modelled first, and then density, if present may be an appropriate response but this can cause problems in that when densities are low, estimated segment numbers will often be falsely 0 leading to bias.

A number of unresolved technical issues remain when fitting the density surfaces, not least of which involves ‘taming’ the smooth functions at the edges of the survey region (to avoid extrapolation difficulties) and (to a lesser extent) in regions where there are gaps in coverage (when interpolating). The evidence here suggested that adding additional transects outside the prediction region can help.

Predictions made from GAMs can also be sensitive to the maximum number of degrees of freedom allowed for estimating the smooth function. Spatial correlation of errors was not considered in this analysis and dealing with such correlation in future may make the data more tractable. To date, little attention has been given to investigating how change in choice of segment length affects the estimation, and if so, how to select the optimal segment length. It is preferable to have segments that are sufficiently small such that expected density is unlikely to vary much within the segment (Hedley *et al.* 2004). However, because of the way in which the segments are defined (with the same environmental and detectability conditions) some variation in this is inevitable. In this analysis, the segment areas were used as weights in the model to account for this variation. Amalgamation of segments can make modelling more tractable and possibly increase spatial independence.

As currently implemented, GAMs produce variance estimates that compare approximately with those produced by stratified samples (compare Table 5). Nonetheless, estimating density in this framework does provide 1 substantial advantage over conventional line transect analysis in elucidating the geographic factors that may influence abundance. They also provide an objective assessment of whether there are differences over time by including year as a factor in the analysis even when the 95% CIs

do not suggest any differences. The selected models allowed some elucidation of the habitat preferences of the humpback whales. The data suggested that humpback whales prefer shallower waters with an optimum of approximately 6–8°C but both these variables may be surrogates for other variables not considered in the model. With a greater availability of potential explanatory variables GAMs (perhaps with spatially correlated errors) could allow elucidation of the habitat preferences of *Megaptera*.

That humpback whales are associated with shallower waters during feeding is well established (Jefferson *et al.* 1993). Again this may reflect optimal feeding locations rather than depth preference *per se*. The mild interaction with sea surface temperature is interesting but it remains to be established if this is an actual temperature/depth preference or a correlation of temperature/depth with some other variable. Humpback whales' association with shallower water might suggest that sea surface temperature could directly influence their position unlike other deeper and longer diving species. However, there is no evidence that the *ca* 6°C “preference” found here is universal; for example Calamboukis *et al.* (2004) found that feeding humpback whales off northern Washington were associated with an average water temperature of 14°C (no formal statistical correlation was made however). Humpback whales certainly shift their distribution in response to food availability (*e.g.* sand eels to herring in the Gulf of Maine, Weinrich *et al.* 1997) and habitat preferences may vary between individuals; for example Ernst and Rosenbaum (2003) found mother calf pairs preferred shallower water when overwintering off Madagascar, although patterns may be different on the feeding grounds. No strong evidence of an association of particular pod sizes was found here although the association with longitude found in 2001 was intriguing.

In conclusion GAMs, as currently implemented for large data sets, provide a useful way to investigate the distribution patterns of humpback whales in relation to environmental variables and can produce smaller confidence intervals than those found in conventional line transect analysis.

ACKNOWLEDGEMENTS

The captains, crews and observers onboard the Faroese and Icelandic vessels are thanked for their excellent cooperation during the surveys. Dorete Bloch organised the logistic and finances of the Faroese cruises. Thanks to her for her help and her warm hospitality during the preparatory days in Tórshavn.

Funding of the surveys was provided by the national laboratories with financial support from the North Atlantic Marine Mammal Commission (NAMMCO), Føroya Sparikassi, Føroya Banki, Føroya Shell and Atlantic Airways. The Faroese Fisheries Institute provided 2 observers for the 1995 survey and 1 observer for the 2001 survey. The Canadian Government provided 2 observers for the Icelandic vessels in 1995. Various pieces of equipment were kindly loaned by the Faroese Telephone Company, Danbiu ApS (Copenhagen, Denmark), le Centre de Recherche sur les Mammifères Marins (La Rochelle, France), l'Institut Français de Recherche et d'Exploitation des Mers (Brest, France), the Sea Mammal Research Unit (then Cambridge now St Andrews, UK). The EU life project SCANS provided the sightings and effort data entry software used.

Thanks to Dave Borchers and his students, Jóhann Sigurjónsson (1995 & 2001) and Daniel Pike (2001) for their advice and support during the planning of the survey. Thanks also to Phil Hammond and Greg Donovan for their advice at the planning stage of both surveys.

REFERENCES

- Balcomb, K.C. and Breiwick, J.M. 1984. Mark recapture analysis of Western North Atlantic humpback whales using fluke photographs. Report of the Scientific Committee, Annex G, Appendix 4. *Rep. int. Whal. Commn.* 34:141.
- Balcomb, K.C., Katona, S.K. and Hammond, P.S. 1986. Population estimates for the North Atlantic for the North Atlantic Humpback whale capture recapture methods on photo identified whales. *Rep. int. Whal. Commn.* 36:110-111.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. and Thomas, L. 2001. *Introduction to distance sampling: estimating abundance of biological populations*. Oxford University Press, London. 432pp.
- Burt, M.L., Paxton, C.G.M. and Hedley, S.L. (MS) 2003. Spatial modelling of humpback whales using data from the 1995 and 2001 North Atlantic Sightings Surveys. Paper SC/54/O12 (SC/11/AE/7) presented to the Scientific Committee of the North Atlantic Marine Mammal Commission Working Group on Abundance Estimation, March 2003. [Available from the NAMMCO Secretariat]
- Calamboukis, J. Steiger, G.H., Ellifrit, D.K., Troutman, B.L. and Bowlby, C.E. 2004. Distributions and abundance of humpback whales (*Megaptera novaeangliae*) and other marine mammals off the northern Washington coast. *Fish. Bull.* 102:563–580.
- Dobson, A.J. 1990 *An introduction to Generalized Linear Models*. Chapman & Hall. London.
- Efron, B. & Tibshirani, R.J. 1993 *An Introduction to the Bootstrap*. Chapman & Hall. London.
- Ernst, P.J. and Rosenbaum H.C. 2003. Habitat preference reflects the social organization of humpback whales (*Megaptera novaeangliae*) on a wintering ground. *J. Zoology* 260:337-345.
- Hedley, S.L., Buckland, S.T. and Borchers, D.L. 1999. Spatial modelling from line transect data. *J. Cetacean Res. Manage.* 1:255-264
- Hedley, S.L. & Buckland, S.T. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological and Environmental Statistics* 9:181-199.
- Hedley, S.L., Buckland, S.T. and Borchers D. L. 2004. Spatial distance sampling models. In: *Advanced Distance Sampling*. Buckland S.T., Anderson D.R., Burnham K.P., Laake J.L., Borchers D.L. and Thomas L. (Eds) Oxford University Press, Oxford
- Horvitz, D.G. and Thompson, D.J. 1952. A generalization of sampling without replacement from a finite universe. *J. American Statistical Association* 47:663–685.
- Ihaka, R. & Gentleman, R. 1996. R: A language for data analysis and graphics. *J. Computational and Graphical Structures.* 5:299-314.
- Jefferson, T.A., Leatherwood, S. and Webber, M.A. 1993. *Marine mammals of the world*. FAO Rome. 320pp.
- Lubick, N. 2003. New count of old whales adds up to big debate. *Science.* 301:451.
- Mackenzie M. L. and Walker, C. (in prep.) Adaptive knot placement for regression splines. *Biometrika*.

- Mitchell, E. and Reeves, R.R. 1983. Catch history, abundance and present status of Northwest Atlantic humpback whales. *Rep. int. Whal. Commn.* (Special Issue 5):153–212.
- Marques, F.F.C. (MS) 2001. *Estimating wildlife distribution and abundance from line transect surveys conducted from platforms of opportunity*. PhD Thesis. University of St Andrews.
- [NAMMCO] North Atlantic Marine Mammal Commission. 2003. Report of the tenth meeting of the Scientific Committee. In: *NAMMCO Annual Report 2002*, NAMMCO, Tromsø, pp. 173–281.
- Pike, D.G., Gunnlaugsson, T., and Víkingsson, G.A. (MS) 2002. Estimates of humpback whale (*Megaptera novaeangliae*) abundance in the North Atlantic, from NASS-95 shipboard survey data. Paper SC/54/H10 presented to the Scientific Committee of the International Whaling Commission, May 2002. [Available from the IWC Secretariat]
- Pike, D.G., Paxton, C.G.M., Gunnlaugsson, Th. and Víkingsson, G.A. 2009. Trends in the distribution and abundance of cetaceans from aerial surveys in Icelandic coastal waters, 1986–2001. *NAMMCO Sci. Publ.* 7:117-142.
- Reynolds, R.W., Rayner, N.A., Smith, T.M., Stokes, D.C. and Wang, W. 2002. An improved in situ and satellite SST analysis for climate. *J. Climate* 15:1609–1625.
- Roman, J. and Palumbi, S.R. 2003. Whales before whaling in the North Atlantic. *Science* 301:508–510.
- Sigurjónsson, J. and Gunnlaugsson, Th. 1990. Recent trends in abundance of blue (*Balaenoptera musculus*) and humpback whales (*Megaptera novaeangliae*) of the west and southwest Iceland, with a note on the occurrence of other cetacean species. *Rep. int. Whal. Commn.* 40:537–550.
- Smith, T.D., Allen, J., Clapham, P.J., Hammond, P.S., Katona, S., Larsen, F., Lien, J., Mattila, D., Palsbøll, Sigurjónsson, J., Stevick, P.T. and Øien, N. 1999. An ocean basin wide mark-recapture study of the North Atlantic Humpback whale (*Megaptera novaeangliae*). *Mar. Mamm. Sci.* 15:1–32.
- Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L. and Pollard, J.H. 2002. Distance 4.0. Release 1. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st.ac.uk/distance/>
- Tønnessen, J. N. and Johnsen, A. O. 1982. *The History of Modern Whaling*, C. Hurst & co, London. 798 pp.
- Víkingsson, G., Gunnlaugsson, Th., Halldórsson, S.D. and Ólafsdóttir, D. (MS) 2002 NASS 2001 Icelandic Shipboard Survey Report. Paper SC/54/010 presented to the Scientific Committee of the International Whaling Commission, May 2002. [Available from the IWC Secretariat]
- Víkingsson, G.A., Pike, D.G., Desportes, G., Øien, N., Gunnlaugsson, Th. and Bloch, D. 2009. Distribution and abundance of fin whales (*Balaenoptera physalus*) in the Northeast and Central Atlantic as inferred from the North Atlantic Sightings Surveys 1987–2001. *NAMMCO Sci. Publ.* 7:49-72.

Weinrich, M., Martin, M., Griffiths, R., Bove, J. and Schilling, M. 1997. A shift in distribution of humpback whales, *Megaptera novaeangliae*, in response to prey in the southern Gulf of Maine. *Fish. Bull.* 95:826–836.

Wood, S.N. 2001. mgcv: GAMs and Generalized Ridge Regression for R. *R News* 1:20–25

Wood, S.N. 2003. Thin plate regression splines. *J. Royal Stat. Soc. (B)* 65(1):95–114