Density of Antarctic minke whales in Weddell Sea from helicopter survey data

Rob Williams, Andy Brierley, Ari Friedlaender, Natalie Kelly, Karl-Hermann Kock, Linn Lehnert, Ted Maksym, Jason Roberts, Meike Scheidat, Ursula Siebert

8 May 2011
rmcw@st-andrews.ac.uk

Introduction

At the 62nd annual meeting of the Scientific Committee of the International Whaling Commission (IWC SC) in Agadir (Morocco), Williams et al. (2010) reported preliminary results from helicopter surveys for Antarctic minke whales and sea ice, conducted from Polarstern in 2006-07 and 2008-09. The cruise plan (Scheidat et al. In press) was designed to achieve other research and logistical objectives in conjunction with reprovisioning the German Antarctic base at Neumayer; therefore placement of helicopter tracklines around the ship's cruise track could not be randomized or designed in a systematic manner, but rather was designed to sample across as wide a range of ice conditions as possible. The resulting survey yielded over 13,000 km of dedicated trackline effort and 94 on-effort sightings of Antarctic minke whales in the two years combined. Exploratory analyses suggested that the highest density of minke whales was found in a narrow band of modest ice concentration (approximately 5-20%), but reanalysis is required to put robust bounds on this band to infer habitat preference (Scheidat et al. In press).

The IWC SC welcomed the new survey (subsequently referred to as the "German data"), but suggested that the data be analysed using methods being developed by Bravington and Hedley (e.g., Wood et al. 2008; Hedley et al. 2009) to allow reliable inference (subsequently referred to as the "Australian methods"). To that end, FTZ funded additional analyses of the helicopter survey data using the "Australian" methods. With the assistance of Dr Natalie Kelly (CSIRO, Australian Antarctic Division), progress was made on three specific areas:

1. use soap-film smoothers (Wood, Bravington and Hedley 2008);
2. try error distributions (e.g., Tweedie) that are robust to unmodelled overdispersion in the data; and
3. evaluate new methods developed by Hedley and Bravington to propagate the variance from the model through to the resulting abundance estimate (Hedley, Bannister and Dunlop 2009; also reported in Williams et al. 2011).

Results

1. Soap film smoothers versus tensor or thin-plate regression splines

The realities of using the soap-film approach

The soap-film smooths require construction of a study area boundary, which defines the shape of the two-dimensional surface being modelled. At the descriptive model stage, this boundary removes all effort and sightings data outside of the polygon,

and at the prediction stage, it removes all grid cells outside the boundary. The preferred model used a 2-dimensional smooth of longitude and ice concentration, therefore the 2D film is not strictly spatial. For example, in Figure 1, note that there is relatively poor coverage of transects near longitude 20˚W that are in open water (i.e., 0% ice concentration). Similarly, there are few observations from entirely ice-covered areas (100% ice concentration) near longitude 40˚W. As a result, an irregular polygon (Figure 1, in red) was placed over the data to select areas in two-dimensional (longitude, ice_concentration) parameter space in which there was reasonable coverage.

The prediction grids supplied for both years (grids 2 and 5 to use the labels used in the figures below and in our Distance project) have longitude and ice concentration combinations that were on the outside of the soap film polygon. Therefore, we had to remove quite a few of the prediction grid points from the abundance estimation step. The spread of points that had to be removed are shown in Figure 1. An indication of where these are relative to space (longitude, latitude) are shown in Figures 2 (Year 1) and 3 (Year 2).

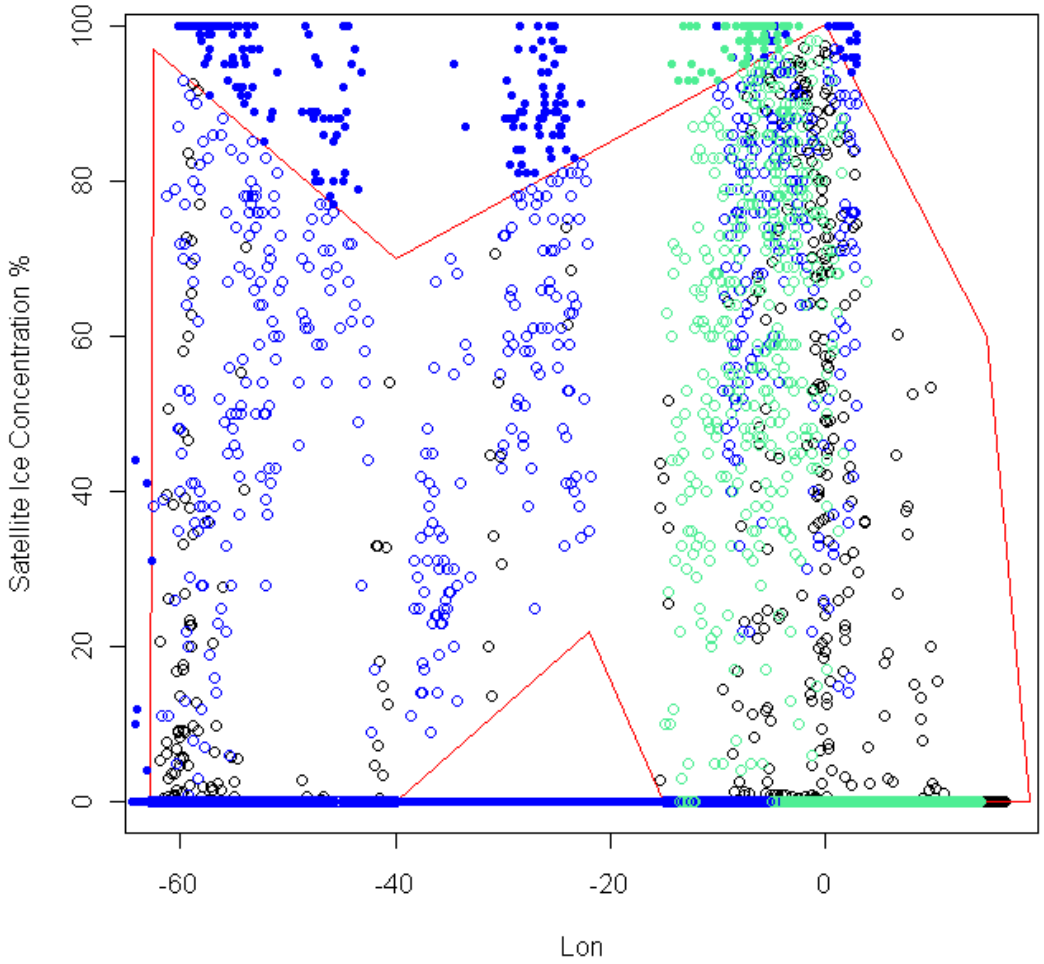

Figure 1 The soap film boundary used in the s(lon, ubiceconc) smooth given in red line; black open circles indicate locations of original data; blue and green open circles indicate locations in the prediction grids for year 1 and year 2, respectively; blue and green solid circles indicate prediction points that do not lie within the soap film polygon and which had to be removed from prediction process.

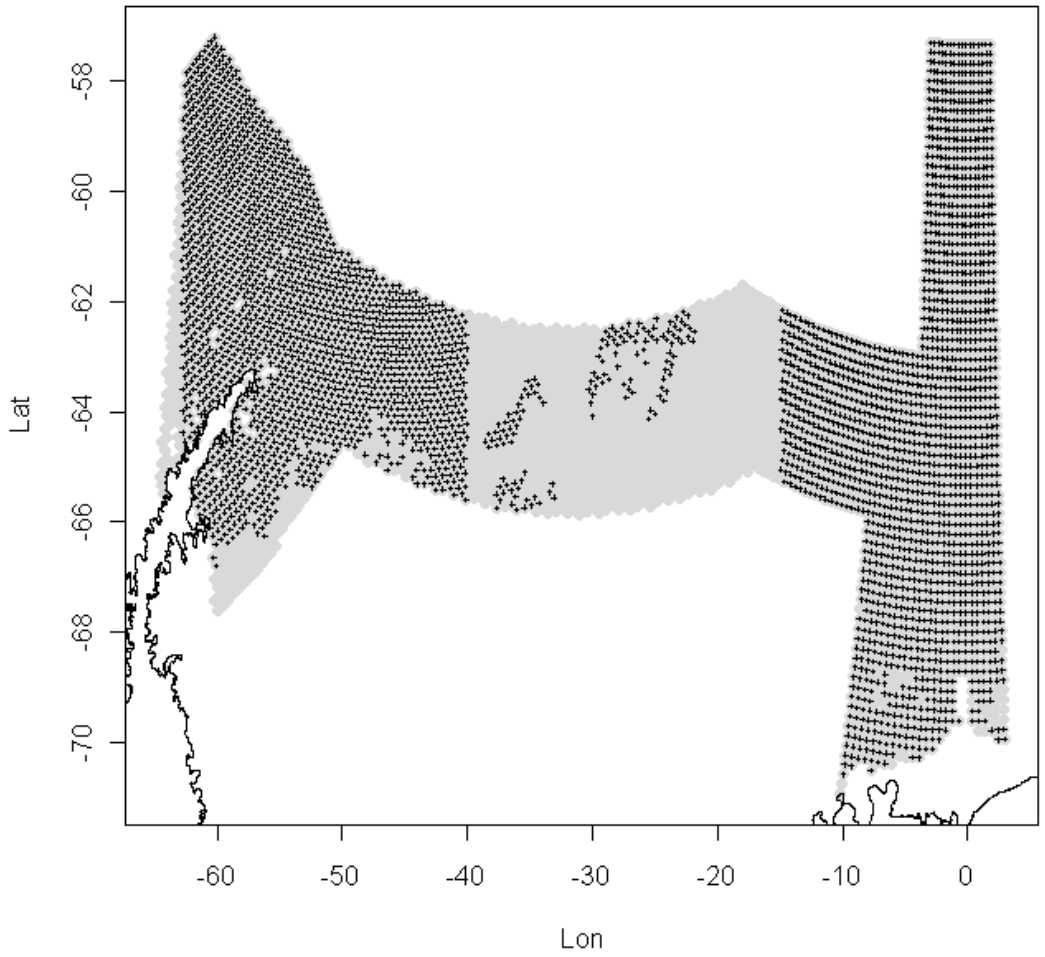

Figure 2 Grey squares indicate the locations of the original prediction grid (corresponds to grid 2 in Distance project); black crosses indicate longitude-ice concentration combinations that were allowed after using a soap-film polygon.

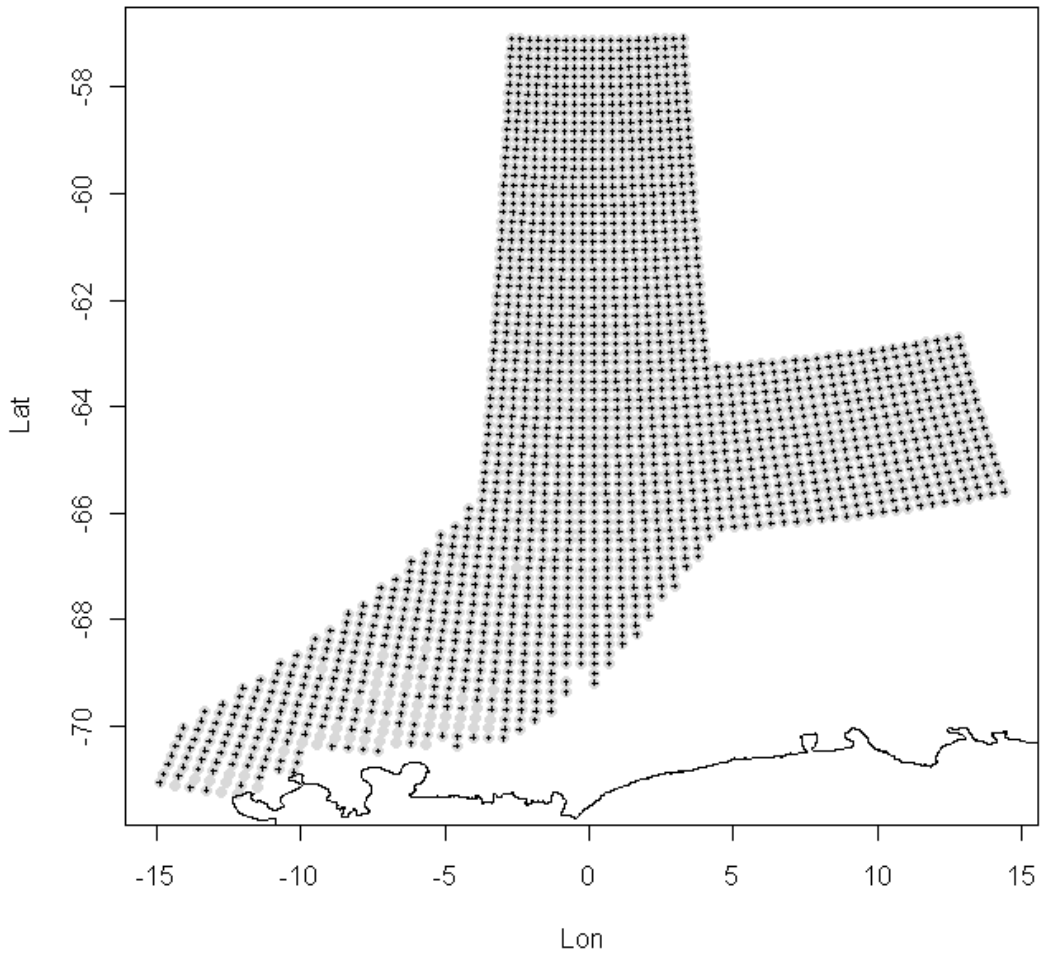

Figure 3 Grey squares indicate the locations of the original prediction grid (corresponds to grid 5 in Distance project); black crosses indicate longitude-ice concentration combinations that were allowed after using a soap-film polygon.

In summary, the soap-film smoothers successfully dealt with the problematic edge effect reported in Williams et al. (2010). Figure 4 shows the predicted density surface from a conventional smooth.

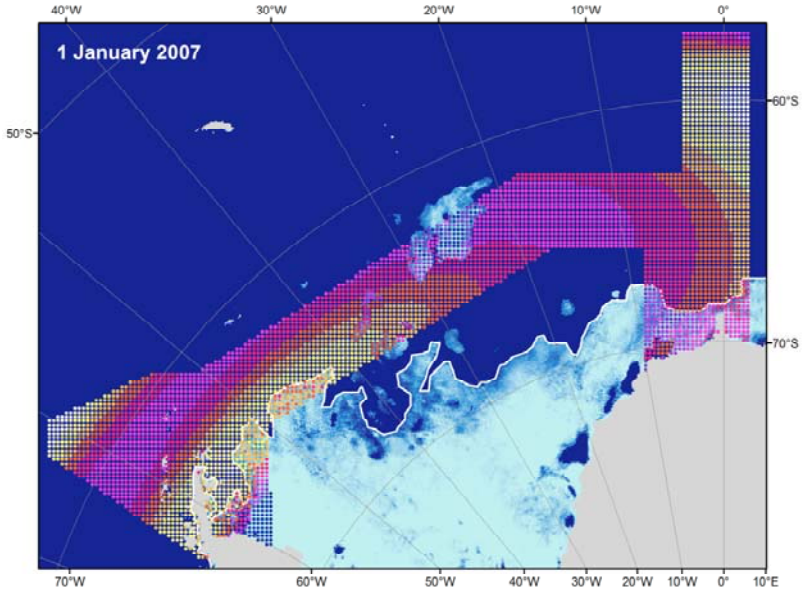

Figure 4. BEFORE soap-film smoothing. Note the high-density areas predicted in the upper left- and right-hand sides of the study area, where there was little effort.

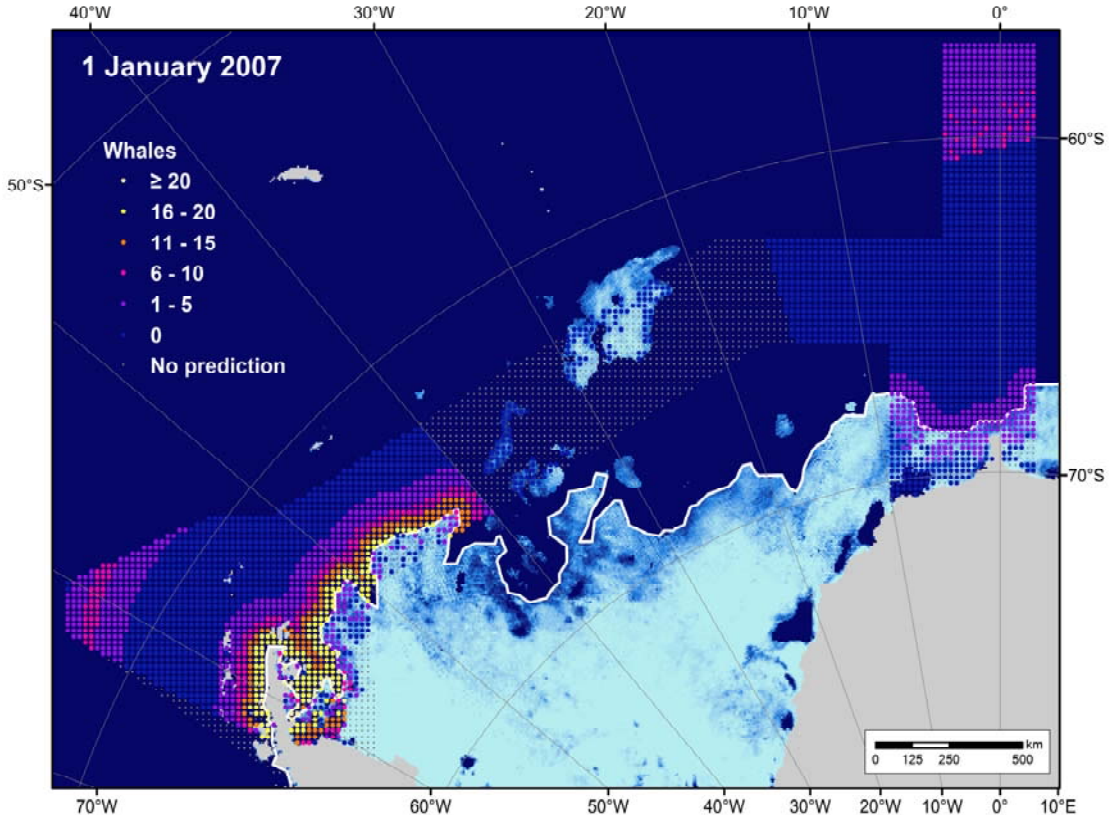

Figure 6: AFTER soap-film smoothing. Note the lower density predicted in the upper-left-hand side of the plot.

2. Error distributions that are robust to overdispersion

One of the biggest problems with the preliminary (Williams et al. 2010) analysis was the failure of the model to capture all of the overdispersion in the data. The Tweedie family is robust to this kind of distribution (Jørgensen 1987). Furthermore, standard plots offered by statistical packages for regression diagnostics do not tend to cater for discrete distributions, such as Poisson. Here randomized quantile residuals plots have been used to explore model fit of the quasi-Poisson and Tweedie distributions (Dunn and Smyth 1996). Figure 7 shows the residuals in the original model, using standard model diagnostics plots supplied by the mgcv library in R; Figure 8 shows the residuals in the original model (Williams et al. 2010), but displayed as randomized quantile residuals; and Figure 9 shows the randomized quantile residuals for a Tweedie family.

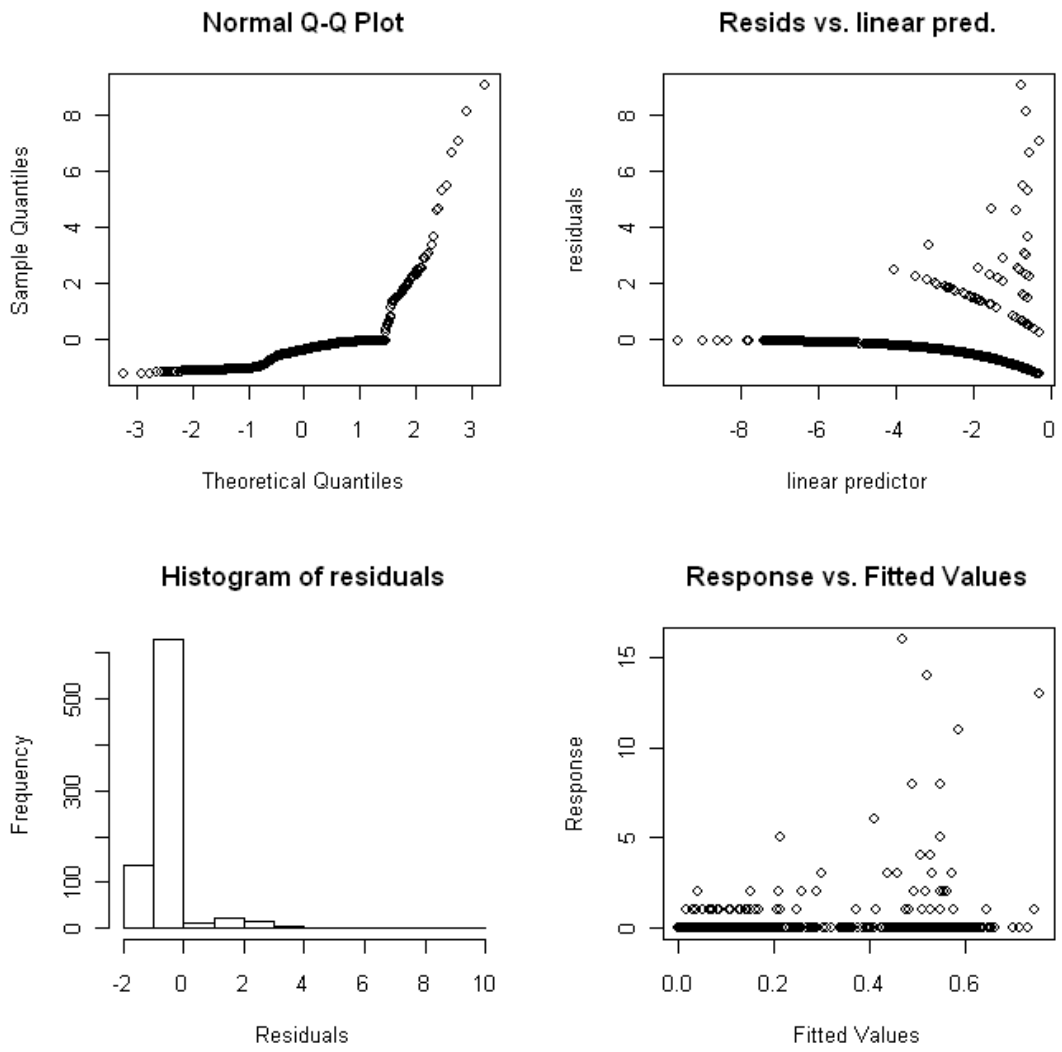

Figure 7. BEFORE Tweedie. Residuals of the original (quasi-poisson) distribution. These residuals are non-normal with a right skew.

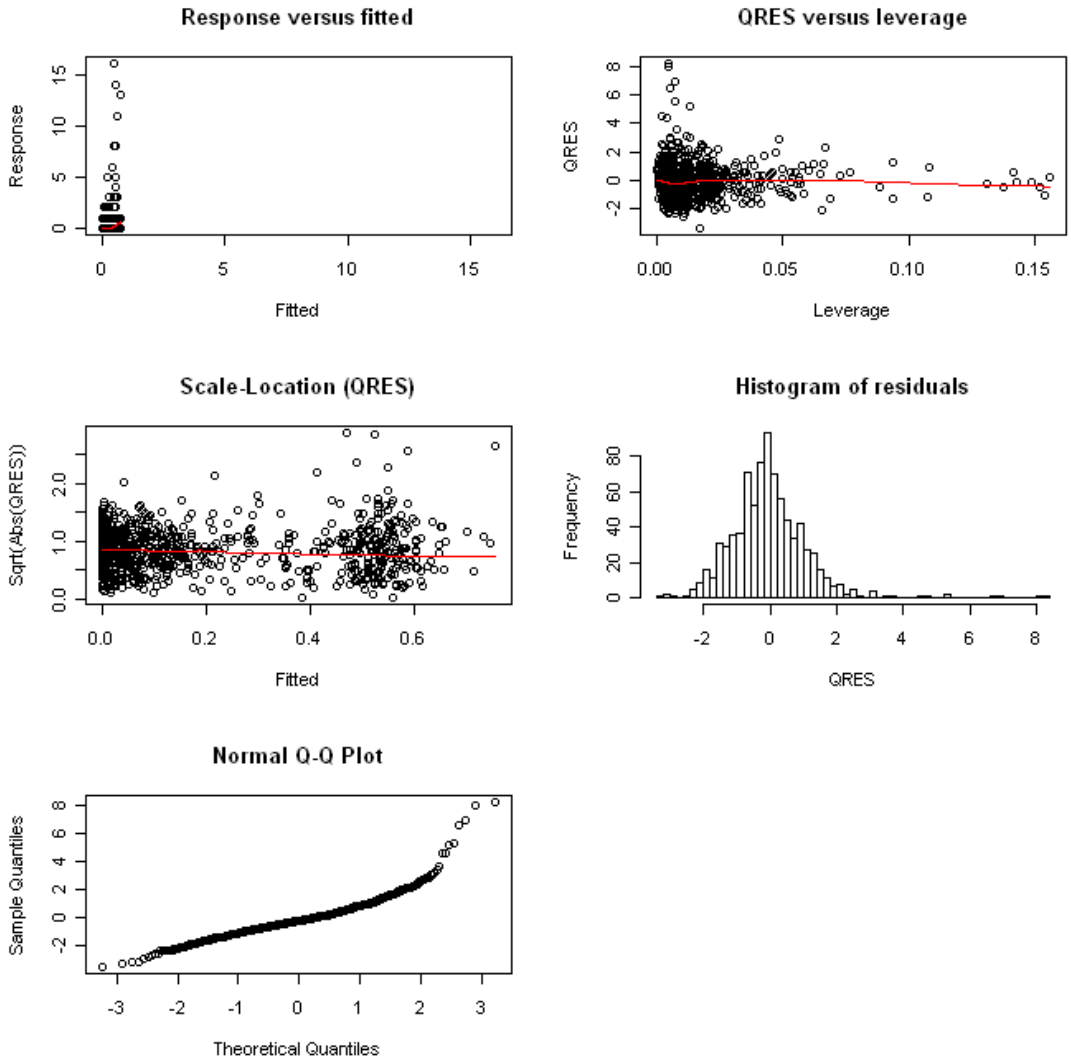

Figure 8. BEFORE Tweedie. Model diagnostics given using the randomized quantile residuals. These residuals are approaching Normal, but there is still a right skew. This indicates the quasi-Poisson family does not quite account for the larger whales sightings numbers.

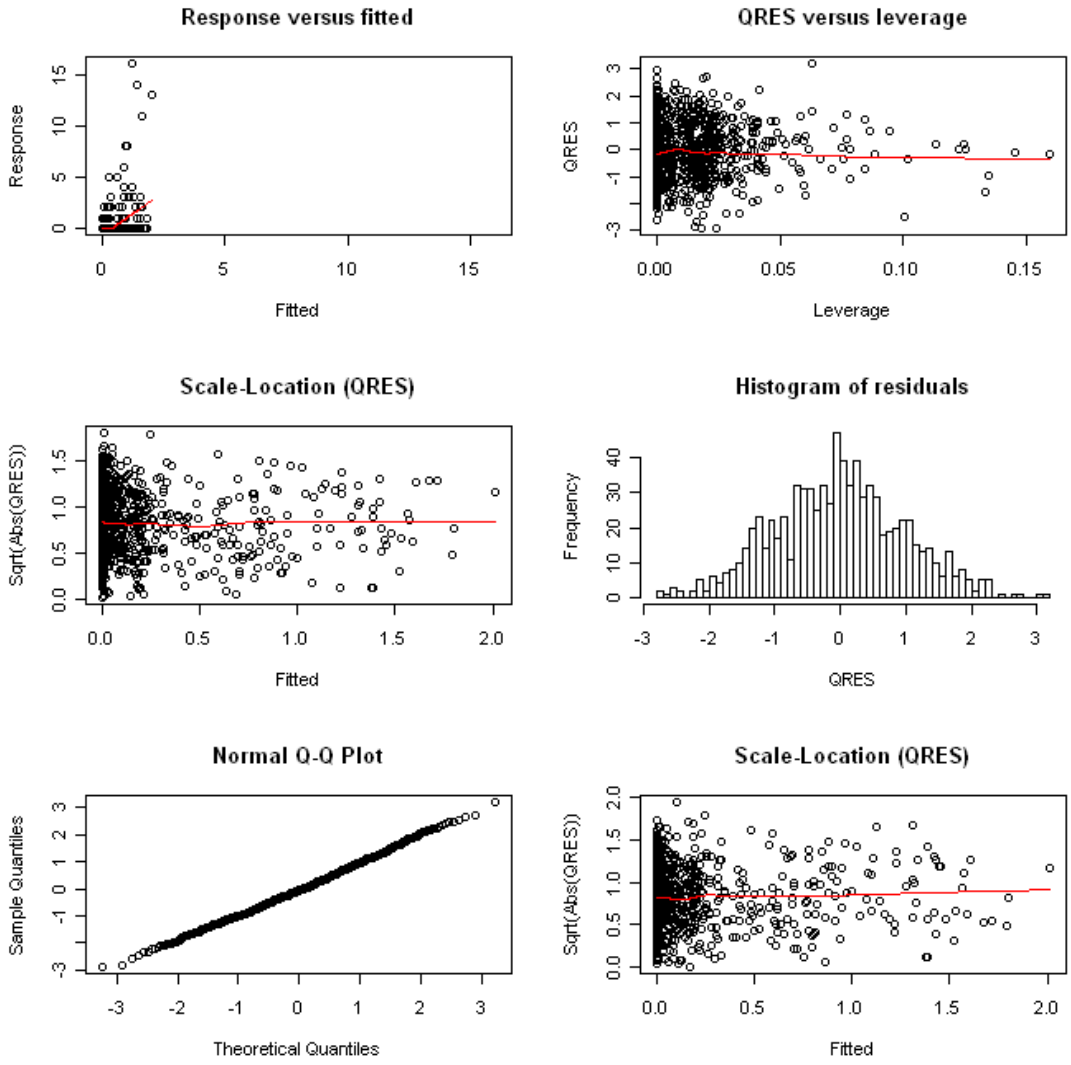

Figure 9. Residuals AFTER Tweedie, and using randomized quantile residuals; shows nice distribution of errors, normally distributed around zero. The model shows excellent fit to the data, with Tweedie parameter = 1.1.

3. Propagating detection function uncertainty through to variance on the abundance estimate using the GAM.fixed.priors approach

Exploratory analyses in Distance indicated that a half-normal detection function fitted the perpendicular sightings distances nicely. The form of a half-normal detection function is:

$$g(x) = \exp(\frac{-x^2}{2\sigma^2}),$$

where *x* is the perpendicular distance from the trackline. In the detection function summary tag in Distance, the σ parameter corresponds to exp(scale intercept estimate); this estimated intercept parameter is noted here as ν. Note that there is only one parameter estimated here, the scale intercept, ν.

From this detection function we can estimate the effective strip width, μ.

$$\mu = \int_0^w g(x)dx = \int_0^w \exp(\frac{-x^2}{2\exp(\nu \times 2)}) = \sqrt{\frac{\pi \exp(\nu \times 2)}{2}},$$

where *w* is the truncation distance.

The effective strip area is then $\hat{\mu} \times l$, where *l* is the length of the segment. (In other analyses, it is important to note whether one has single- or double-sided effort in estimating effective strip area, but this information falls out a bit later when we take the derivative, so we won't consider it further.)

Using the variance-estimation methods of Hedley and Bravington (described in Hedley et al. 2009 and Williams *et al.* 2011), the gam.fixed.priors code takes the hessian from the fit of the detection function and the partial derivative of the effective strip area relative to the parameters in the detection function. Also recall that the effective strip area enters the model as a log(offset) variable. So, the function and its derivative that we need to consider are:

$$\frac{\partial \log(esa)}{\partial \nu} = \frac{\log(l * \sqrt{\frac{\pi \exp(\nu \times 2)}{2}})}{\partial \nu} = 1$$

And, so, a vector of ones enters the gam.fixed.priors code.

When we run all this, we get exactly the same abundance predictions from GAMs including and not including the fixed priors, for both plain smooths and soap film smooths. These are values given below. The predicted density surface plots are similar.

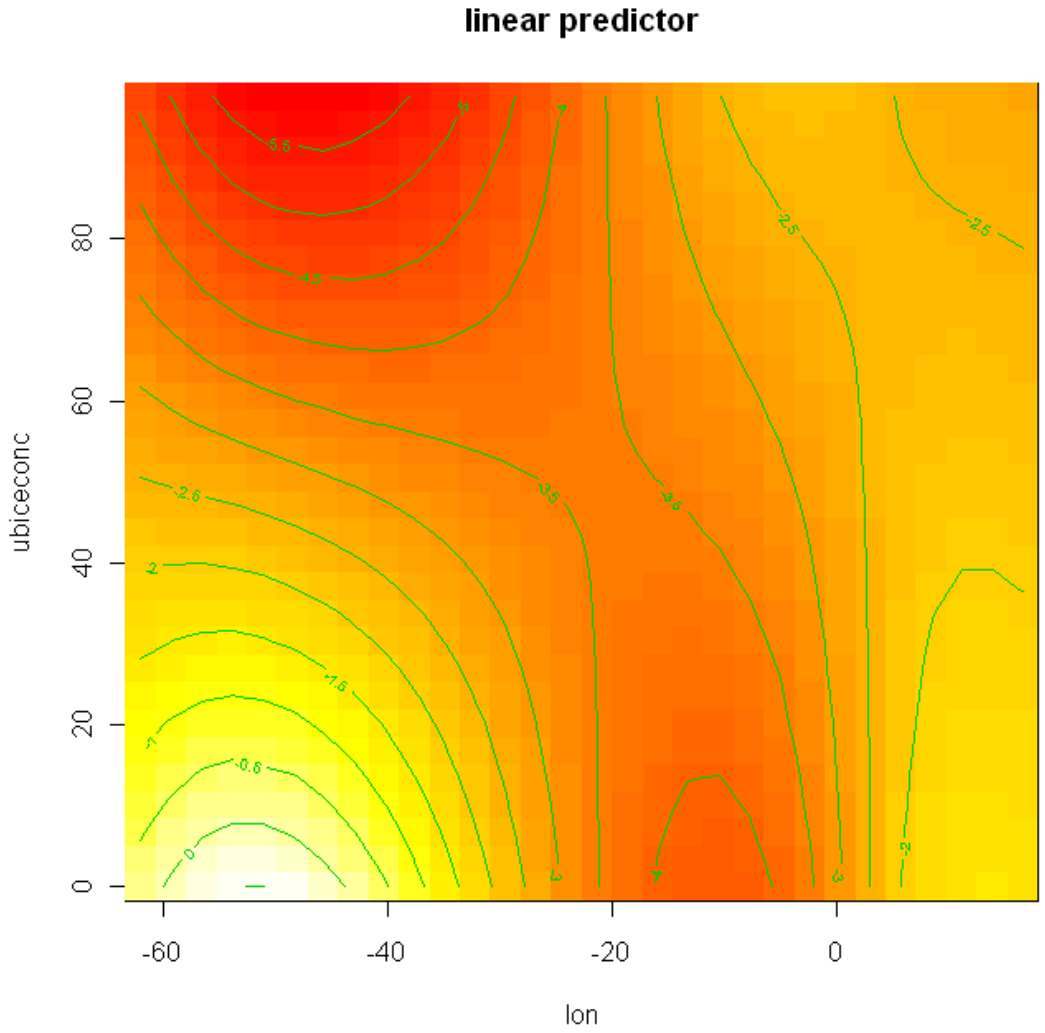

Figure 10. GAM with no soap film or fixed priors.

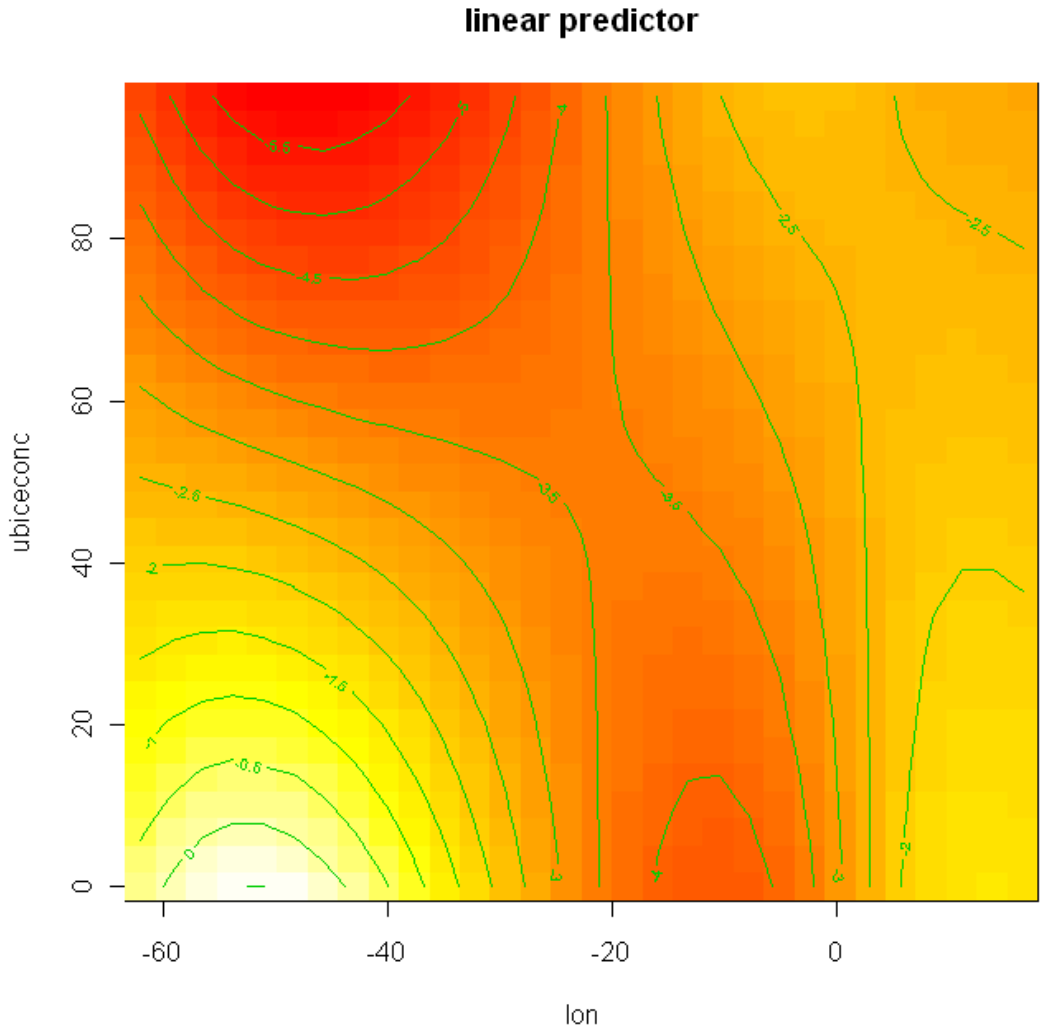

Figure 11. GAM with fixed priors

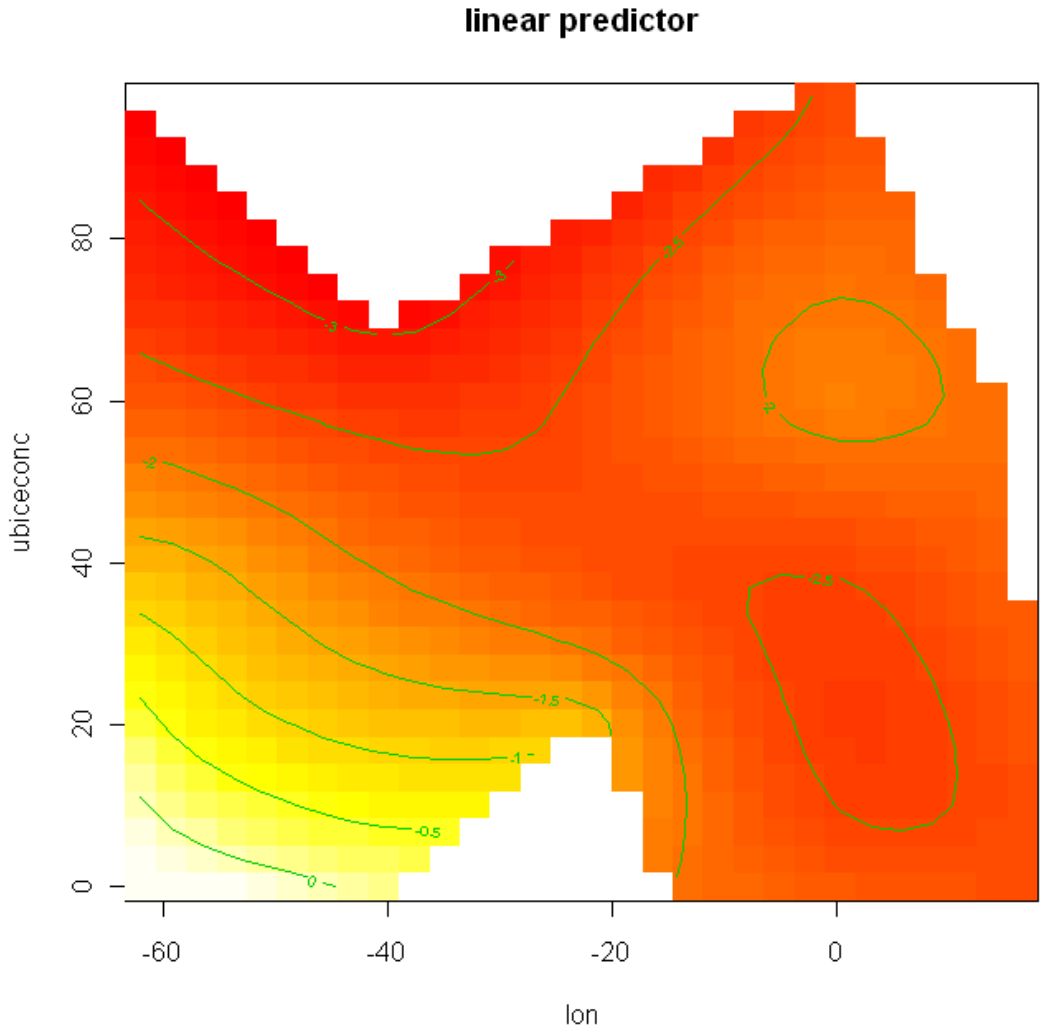

Figure 12. GAM with soap-film smoother

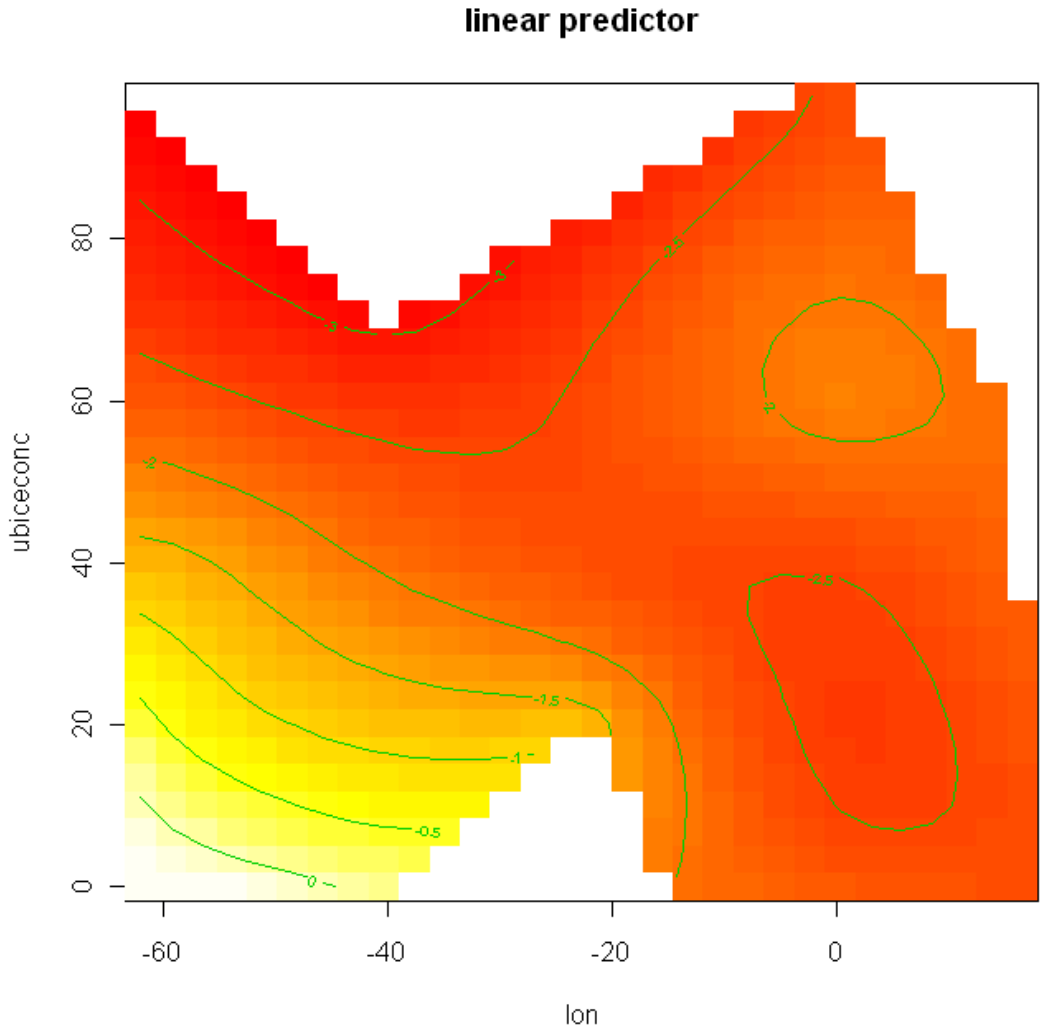

Figure 13. GAM with soap-film and fixed priors.

Final points: abundance of Antarctic minke whales in the study area

The preferred model is Tweedie parameter 1.1, a soap film smooth and GAM fixed priors. The resulting abundance estimates (and coefficients of variation) are:

Year 1: abundance: 8785.7 (CV = 0.497)
Year 2: abundance: 878.4 (CV = 0.632)

The best estimate of distribution is shown in Figure 14.

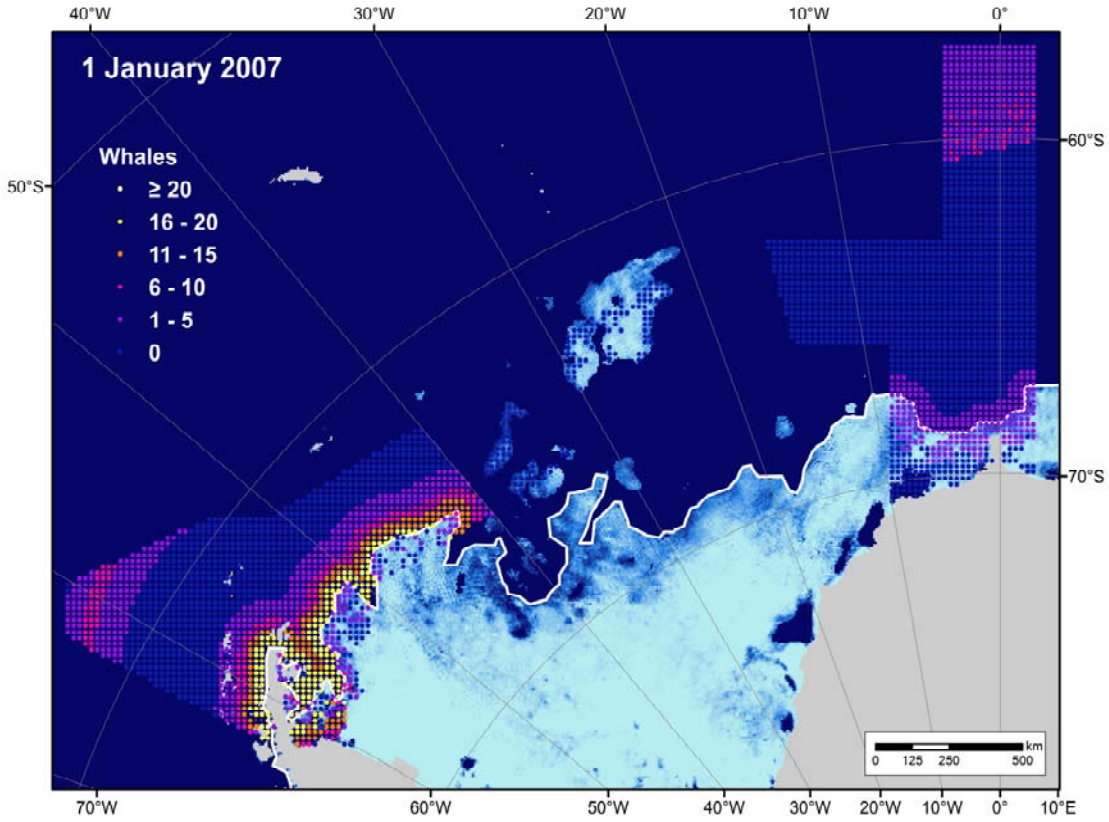

Figure 14. Best estimate of Antarctic minke whale density in the study area, on average, during the time of the surveys.

Follow-on work (Seeking guidance from SC/IA members).

A cumulative distribution function plot (Figure 15) was generated to show the cumulative number of sightings as a function of distance to the ice edge (defined as the 15% ice contour). The results indicated a narrow band of high density within a few tens of kilometres from the ice edge.

We used historical(1979-2007) passive microwave sea ice concentration data derived using the modified bootstrap algorithm (Comiso and Nishio, 2008) provided by the National Snow and Ice Data Center (Comiso, 1999). Figure 16 shows trends in monthly mean ice concentration for the entire Antarctic for January. Notably, there has been a significant decline in ice concentration in January around the Antarctic Peninsula, particularly on the western side. Note that trends near the ice edge largely reflect trends in the location of the edge, rather than a trend in the ice concentration south of the edge i.e., a southward trend in the position of the ice edge will manifest as a strong trend in concentration, as regions previously ice-covered become ice-free; yet there may be no change in the ice conditions south of that edge.

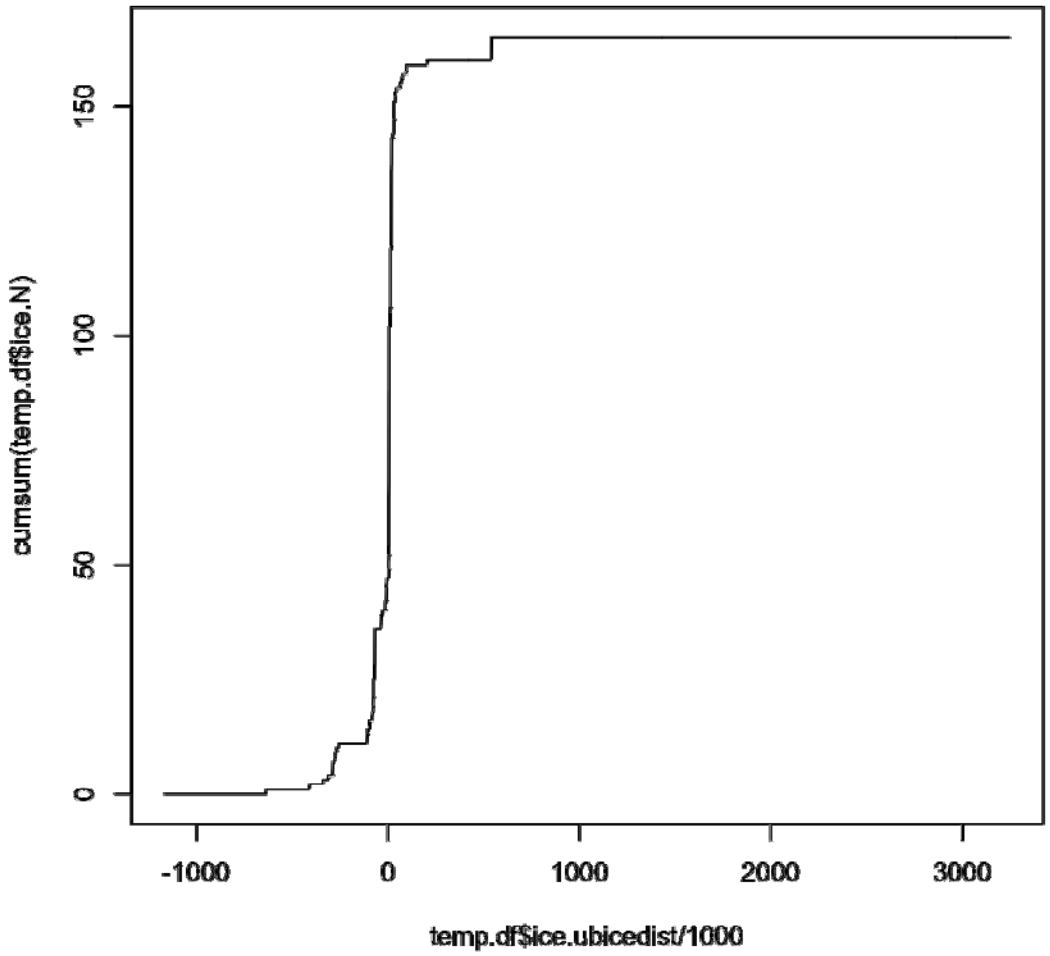

Figure 15. Cumulative number of minke whale sightings (Y axis) as a function of distance to the ice edge (along the X axis, ranging from 1000km inside the ice in negative values, to 3000km outside the ice, in positive values).

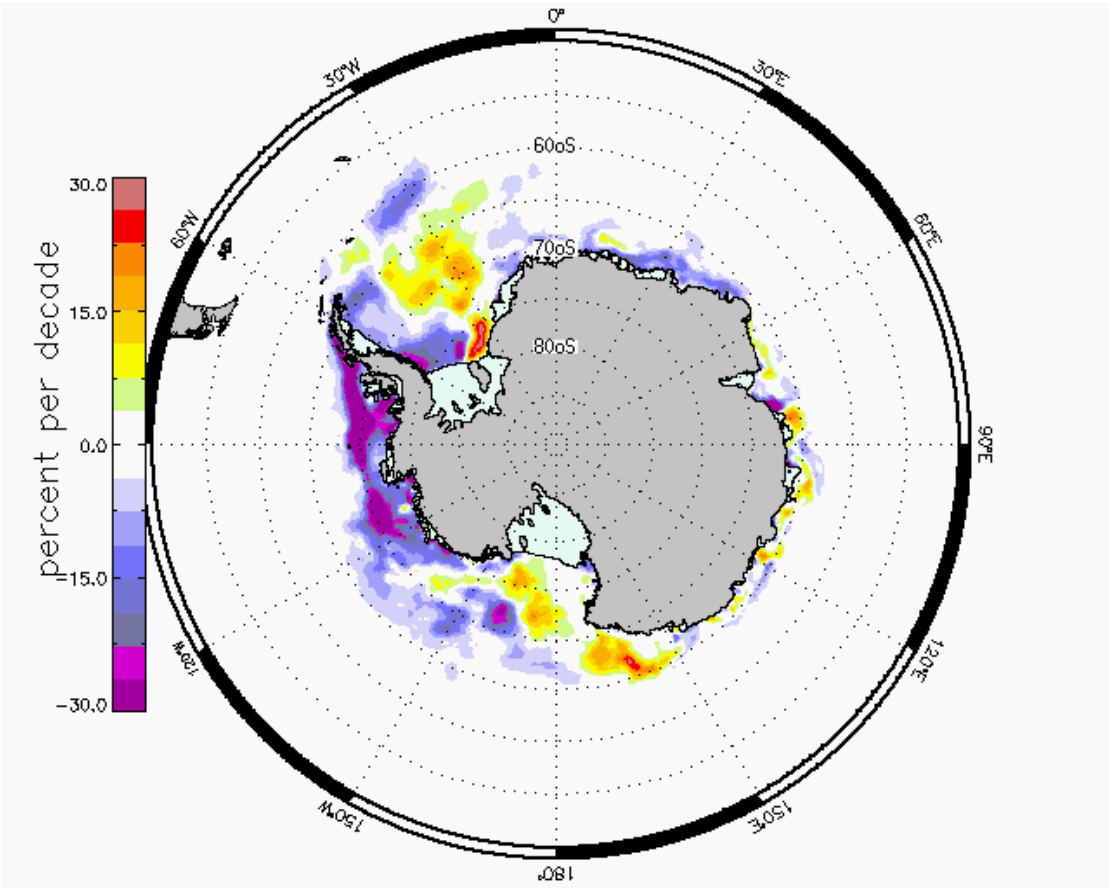

Figure 16. Trends in ice concentration (in million square km) for January, 1979-2007. The legend is in average change (in %) per decade by region. Data (see body of text, above) from the National Snow and Ice Data Center (Boulder CO).

References

Comiso, J. C., and F. Nishio (2008), Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data, J. Geophys. Res., 113, C02S07, doi:10.1029/2007JC004257.

Comiso, J. 1999, updated 2008. /Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I/, 1979-2007. Boulder, Colorado USA: National Snow and Ice Data Center. Digital media.

Dunn, P.K. and Smyth, G.K. 1996. Randomized quantile residuals. *Journal of Computational and Graphical Statistics* **5**: 236-244.

Hedley, S.L., Bannister, J.L. and Dunlop, R.A. 2009. Group IV humpback whales: abundance estimates from aerial and land-based surveys off Shark Bay, western Australia, 2008. Paper SC/61/SH23 presented to the Scientific Committee of the International Whaling Commission, June 2009 (unpublished). 17pp.

Jørgensen, B. 1987. Exponential dispersion models. *Journal of the Royal Statistical Society Series B* **49**: 127-162.

Scheidat, M., Friedlaender, A., Kock, K-H., Lehnert, L., Boebel, O., Roberts, J. and Williams, R. In press. Cetacean surveys in the Southern Ocean using icebreaker-supported helicopters. *Polar Biology*. DOI: 10.1007/s00300-011-1010-5

Williams, R., Friedlaender, A.S., Kock, K.-H., Lehnert, L., Roberts, J. and Scheidat, M. 2010. Antarctic minke whale density in relation to sea ice: helicopter surveys in the Weddell Sea. SC/62/IA13. Unpublished document presented to the International Whaling Commission. Available from www.iwcoffice.org.

Williams, R., Hedley, S.L., Branch, T.A., Bravington, M.V., Zerbini, A.N. and Findlay, K.P. 2011. Chilean Blue Whales as a Case Study to Illustrate Methods to Estimate Abundance and Evaluate Conservation Status of Rare Species. *Conservation Biology* **25**:526-535.

Wood, S.N., Bravington, M.V. and Hedley, S.L. 2008. Soap film smoothing. *J. R. Statist. Soc. B* **70**:931–955.