Marine Mammal Scientific Support Research Programme MMSS/001/11

MR 6.1:

Report

Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development.

Sea Mammal Research Unit
Report to
Scottish Government

July 2015 [version F2]

Lonergan, M. & Hammond, P. S.

Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife. KY16 8LB.

Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development.

Editorial Trail

Main Author	Comments	Version	Date
M. Lonergan	Author	V1.0	19/01/2015
P. Hammond	Author	V1.0	19/01/2015
P. Irving A. Hall	review	V1.1	19/01/2015
Marine Scotland	comments	V2.0	24/04/2015
A. Hall	final editing	VF1	21/07/2015
O. Racu	final editing	VF2	15/08/2015

Citation of report

Lonergan, M. & Hammond P. S. (2015) Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development. Sea Mammal Research Unit, University of St Andrews, Report to Scottish Government, no. MR 6.1, St Andrews, 9pp.

Contents

Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development.

1 Executive Summary

There are a number of potential issues affecting the populations of cetaceans living in or using UK waters. Without knowing the distribution and abundance of these animals the ability to assess how any of these issues could be affecting UK cetaceans is limited.

The revised Joint Cetacean Protocol (JCP) Phase III report attempted to estimate the abundance and distribution of cetaceans from a disparate data set of dedicated and platform of opportunity surveys. It also considered trends in cetacean abundance in the North Sea and waters out to the shelf-edge west of the UK and Ireland, also examining subareas within that area.

The modelling of the data involved sophisticated statistical methods and required various simplifications to be made. A formal model selection process was used but alternative models might produce very different estimates. It is recommended that work is carried out to explore whether moving to indices of abundance and/or maps of relative densities will simplify the model and reduce the risk of producing misleading results without compromising the results of the project.

For some species, the patterns identified by the models are inconsistent with other available sources of information. In particular, the JCP estimate for harbour porpoise abundance in 1994 is difficult to reconcile with that from the SCANS survey.

The JCP report contains many important caveats about the robustness and reliability of its results. It is not obvious that the results of the analyses it contains provide a suitable basis for the conservation and management of cetacean populations around the UK.

Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development.

2 Introduction

This document considers the report entitled: "Revised Phase III Data Analysis of Joint Cetacean Protocol Data Resource" (Paxton *et al*., 2012). It does not constitute a formal peer-review, but is intended to highlight issues important to the interpretation and evaluation the report and help Marine Scotland form a view of how it can be used in decisions relating to the conservation and management of cetaceans in Scottish waters. The authors of the JCP report mention many of the same issues as are raised here, but the length of that document, and the need for its results, could result in their caveats being overlooked.

The JCP Phase III report uses data from various sources to supplement the estimates of abundance produced by the two SCANS surveys for seven species of cetacean. It aimed to produce more precise estimates of total abundance for each species over the North Sea and waters out to the shelf edge west of the UK and Ireland. Estimated changes in these abundances over the period 1994-2010 were included, as well as abundances and trends in some subareas, including ones identified as being of particular interest to the offshore renewables industry.

Complex and sophisticated statistical methods were used to model the density of each species across the whole area. The project faced two major problems as it fitted the models. Firstly the diversity of the data, due to the different ways they were collected and recorded, required various assumptions and simplifications to be made in order to combine the datasets. Secondly, the sheer size of the combined dataset and the complexity of the models used to represent the patterns, of abundance and uncertainty, within it posed computational challenges.

3 Methods

3.1 Data

The JCP analysis combined 38 different datasets. The requirements for inclusion were that data were collected, in reasonable weather, from platforms (ships or aeroplanes) that did not follow the animals they saw, and included information on effort (where and when observers were working). Three categories of visual observations were therefore excluded: anecdotal sightings data, which are very difficult to interpret because of the lack of information on where people looked without seeing anything, mark recapture studies, and observations made by stationary observers on land. There would be two major difficulties to incorporating the mark-recapture and shore-based observations into the analysis: the importance in overlap between surveys for the patching together of datasets, and the structure of the analysis which effectively starts by transforming data to resemble the results of Distance surveys. Acoustic detections of cetaceans were also excluded. It would be possible to add both shore-based and acoustic observations into the analysis, but that would further increase the complexity of the model. The problems encountered in modelling the current dataset suggest that including additional types of data is unlikely to be worthwhile unless they provide substantial amounts of information on areas where the visual survey data are sparse.

3.2 Analytical methods

The data analysis was carried out in stages: first the numbers of animals counted in each sighting were corrected to allow for those that were present but not seen; then the data were divided into segments (each around 10km long); then smoothly-varying density surfaces were fitted to the rescaled data. One of these models under consideration was selected as the best representation of the data for each species. Finally, parametric bootstraps were used to capture the uncertainty associated with the first and third stages. Each of these is considered in turn below.

3.2.1 Imperfect detection

The data were collected from "at least 542 vessels". These were grouped for the analysis, and surveys that recorded the distances at which each animal was seen were used to estimate detection functions for each vessel group. Detection functions were estimated for eight groups of vessel for harbour porpoises. For the five dolphin species, common detection functions were fitted to three groups of aircraft, and different ones for the five groups of ships. Three groups of vessels were used for minke whales. Data from the double platform surveys were used to correct the detection functions for imperfect detection on the trackline and a separate correction was applied to the data from the smaller boats and aircraft, to allow for those animals that were underwater as the vessels passed.

The report states that a "proportion of sightings with missing distances was discarded at random, this proportion being the same as the proportion of the detections of known distance that were beyond the truncation distance." After that the detection functions were used to calculate corrections that would convert the numbers of animals observed into estimates of the local abundance of animals.

The grouping of vessels is necessary to the analysis, but has the potential to introduce artefacts. If larger vessels within a group tend to detect more animals and were used in certain years or areas, spurious patterns could appear in the data. Testing for that would probably be quite difficult.

3.2.2 Segmentation

The choice of segment length attempted to strike a balance between the lower precision of long segments and computational problems resulting from a larger dataset containing a high proportion of segments where no animals were observed. The use of Generalised Estimation Equations for the modelling directly tackled the issue of autocorrelation along the survey tracklines (the potential for unmeasured environmental features to affect the counts in consecutive segments). The chosen segment length is within the range of values used in previous studies of this sort of data.

3.2.3 Model structure

The estimation of abundance was done using sophisticated regression techniques developed by authors of the report. Part of that work has been published in a refereed scientific journal, but there will be few people, outside the authors' research group, able to evaluate the details of the model fitting methods. These new methods consider distances "as the animal swims" rather than straightline distances between locations. The effect is to reduce the connection between areas separated by obstacles such as land. It is not obvious how appropriate the modification will be for the environmental determinates of local abundance. The final models chosen have 6 or 12 degrees of freedom in their spatial smooths (which is roughly equivalent to allowing the presence of many local peaks in abundance over the whole map). The maps presented in the report generally show a few broad simple peaks in abundance and have wide confidence intervals around their estimates. It is therefore unclear that the additional complexity in the model structure materially affects the results of the analysis.

The models also considered smooth effects of year, time of year, depth, slope and sea surface temperature. For harbour porpoise an interaction between year and location was also included in the final model. The details of how complexity was distributed among the various explanatory covariates may be important to the results. Testing the use of the spatial smoothers or other assumptions would require fitting models with differing assumptions and this process would be time-consuming. If the results of such models were similar to the current ones, it would provide reassurance about the validity of the results, but it is less obvious what could be done if sensitivity to these assumptions was identified.

3.2.4 Model selection

The report comments that "fit criterion to govern model selection for GEE models is still an area of active research" and highlights auto-correlation and over-dispersion of the residuals (which are taken to follow a quasipoisson error distribution, in order to represent the potential non-independence of sightings) as issues of particular concern. They therefore adopted what they considered to be a "conservative measure for model selection", based on the Bayesian Information Criterion. They chose this over an approach based on cross-validation because that "returned overly simplistic models that failed to identify cetacean concentrations in certain areas that persisted over time", and comment that the lower penalty on the inclusion of parameters within the Akiake Information Criterion would have resulted in the selection of more complex models.

While a certain amount of subjectivity is inescapable in the choice of model selection criteria, sensitivity to this choice has ramifications for effective usability of the chosen model. For some species the inclusion of spatio-temporal interactions (distributions that changed over time) "showed a greatly improved fit to the data but the uncertainty in these fitted surfaces was prohibitively high" and the simple models were reported as the best representations of those data. Those decisions introduce a dependency on the analysts' expectations of the results of their analysis. This dependency will affect overall abundances less than patterns of distribution, but is likely to bias all estimates of uncertainty to an extent that will be difficult to predict.

3.2.5 Treatment of uncertainty

The uncertainty in the rescaling of counts within segments and the modelling of density surfaces were treated separately in a two-stage parametric bootstrap. The first stage created 500 replicate rescalings of the count data, each one of which used a detection function that was generated by drawing from estimated distributions for each of its parameters. The second stage then fitted the abundance model, identified as best in the model selection process, to each replicate and drew a set of parameter values from the resulting distributions of parameter values. These were then used to predict density in each 5km grid cell. The report's authors comment that this neglects uncertainty in the model selection process, presumably because incorporating that would be computationally infeasible.

The mean estimate and 95% confidence interval for each grid cell was then calculated from the 500 replicate estimates of the density of animals within it. Estimated densities were then produced for the areas of interest by taking each replicate in turn and summing the estimates within that area. Mean estimates and 95% confidence intervals were then calculated from the resulting distributions of area estimates.

The uncertainty in estimates for new areas and combinations of cells cannot easily be calculated from the results for individual cells (because spatial autocorrelation will mean that the density within cells beside one that has been underestimated is likely to also be underestimated). Simply summing the lower bounds on all the estimates in an area will underestimate the lower bound for the area, and overestimate the uncertainty in the estimate. Treating the uncertainty from each cell as independent will underestimate the uncertainty. The report discusses the implications of the smoothing on estimates for small areas: essentially warning that they are unreliable. Overall it may therefore be unwise to provide estimates of the uncertainty in individual 5km cells since these, and any combinations of them, are likely to be misleading. If it is intended to be able to generate estimates of the uncertainty in newly defined areas, then the individual replicate results for each cell will need to be made available. These can then be combined in the same way as was done for the areas identified in the original report.

4 Results

4.1 JCP Estimates

Using the patterns identified by models fitted to a dataset to choose between formal model selection criteria can be misleading, because it gives a false air of objectivity to a subjective decision. However, comparing the results of models to information obtained from other datasets and analyses does provide a useful reality check. The sensitivity of results to changes in assumptions and the structure of models gives an indication of how important precise choices of those details are. Identification of inconsistencies with previous work that used either part of the same dataset or other types of data can reveal important determinants of the patterns.

Over the life of the JCP project the estimates of this approach's power to detect changes has varied substantially. The Preliminary Report suggested that only a 60% decline over 6 years would be detectable. The Phase I report suggested that, for harbour porpoise and bottlenose dolphins, a 2% total decline over 6 years would be detectable. In the draft Phase III report the estimates for the limit of detectable annual change in harbour porpoise abundance ranges between 7% and 65% for the various areas and time periods. These fluctuations have been associated with progressive refinements of the methodology, but there is no obvious way to determine whether the process has now reached a reliable representation of abundances and distributions.

The results of the SCANS surveys, which provide a large proportion of the most informative data used by the JCP, are likely to provide a particularly informative comparison. The JCP report says of harbour porpoises "Estimated numbers from the SCANS survey region of 1994 (not including the Baltic blocks) using the JCP model were of 653100 (81500 – 5252400), compared to 329200 (166400 – 651200) calculated from the actual SCANS paper (i.e. Table 4 in Hammond *et al.*, 2002)". The JCP point estimate, which uses the

SCANS data along with other information, therefore lies outside the confidence interval for the SCANS estimate. It also has a much wider confidence interval.

The JCP maps for bottlenose dolphins do not identify a hotspot in the Minch, a location where a mark-recapture study was carried out (Cheney *et al.*, 2013) suggesting that a relatively high concentration of animals occurs there. The results also show the Firth of Forth as probably containing more bottlenose dolphins than the Moray Firth. Strangford Lough is estimated to have contained at least 10 minke whales and 200 harbour porpoise in summer 2010. None of these features seem very likely (though it should be noted that the areas given these names are quite large; the two Scottish areas meet near Fraserburgh and together cover the whole east coast of Scotland; Strangford includes some coastal waters at the mouth of the Narrows). The estimated distribution map for Risso's dolphin is also hard to reconcile with the map of observations it is based on (JCP Figure 16).

5 Conclusion

Phase III of the JCP fits complex and sophisticated statistical models to large amounts of data from many sources. CREEM is certainly among the best groups in the world to attempt that task. However, it is less clear how practically useful the results of the project are or could be. The simplifications necessary for fitting the models and the inconsistencies between the results of the JCP and other analyses are both causes for concern. The subjective rejection of some models identified as best by the model selection criteria is also problematic.

The JCP report comments that its results are likely to be least robust for small areas, and less reliable than those from structured surveys like SCANS. A major part of the difficulty in the model fitting was due to the very large amounts of data involved. While technical fixes may be found to the computational difficulties, the problems of the complexity of the models and techniques, relative to the amount of information available, would be much more difficult to handle. It appears that the models are close to the upper limit of what this dataset can support.

The draft Phase III report emphasises abundance estimation. Its authors acknowledge that this involves scaling the results of their models by the proportion of the time that cetaceans spend at the surface, and that they have very few data on this. Either a way needs to be found to provide adequate information to, at the very least, estimate the uncertainty in this scaling, or there is a substantial risk of producing misleading population estimates. Absolute abundances are not necessary for assessing trends and changes. It is recommended that work is carried out to explore whether moving to indices of abundance and/or maps of relative densities will simplify the model and reduce the risk of producing misleading results without compromising the results of the project.

Overall, it seems unlikely that estimates, with a useful level of precision, of the local abundance of cetaceans can be generated from this dataset. Even at the larger spatial scales it is not clear that the results of the JCP project can be relied on as a basis for management decisions. While the authors of the report have included many important caveats, such warnings often get lost when the results of analyses are quoted. There therefore seems a real risk that use of the results contained in that report could compromise future decisions related to the conservation and management of cetacean populations in this area.

6 References

Cheney, B., Thompson, P.M., Ingram, S.N., Hammond, P.S., Stevick, P.T., Durban, J.W., Culloch, R.M., Elwen, S.H., Mandleberg, L., Janik, V.M., Quick, N.J., Islas-Villanueva, V., Robinson, K.P., Costa, M., Eisfeld, S.M., Walters, A., Phillips, C., Weir, C.R., Evans, P.G.H., Anderwald, P., Reid, R.J., Reid, J.B. & Wilson, B. (2013) Integrating multiple data sources to assess the distribution and abundance of bottlenose dolphins *Tursiops truncatus* in Scottish waters. *Mammal Review*, **43,** 71–88.

Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.P., Heimlich, S., Hiby, A.R., Leopold, M.F. & Øien, N. (2002) Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. *Journal of Applied Ecology*, **39**: 361–376.

Review of methodology and main results of the JCP analysis of cetacean densities in the context of marine renewable development.

Paxton, C.G.M., Scott-Hayward, L., Mackenzie, M., Rexstad, E., & Thomas, L. (2012) Revised Phase III Data Analysis of Joint Cetacean Protocol Data Resource. Report to Joint Nature Conservation Committee. Centre for Research into Ecological and Environmental Modelling, University of St. Andrews. 175pp.