

**Cheap DECAF: Density Estimation for Cetaceans from Acoustic Fixed sensors  
using separate, non-linked devices**

Len Thomas  
Centre for Research into Ecological and Environmental Modelling (CREEM)  
University of St Andrews  
St Andrews, UK.

Phone: UK+1334 461801  
Fax: UK+1334 461800  
Email: len.thomas@st-andrews.ac.uk  
Website: <http://www.creem.st-and.ac.uk/len/>

Luís Matias  
Centro de Geofísica  
Universidade de Lisboa  
Lisbon, Portugal.

Award Number: N00014-11-1-0615

This project was a collaborative project between several institutions with two separate award numbers. Here we present the final technical report for tasks relating to the above award number. The component being led by with Oregon State University has a different end date.

## **MAIN ACCOMPLISHMENTS**

This project has achieved its main goal of demonstrating the utility of using separate non-linked devices to monitor cetacean density. We have furthered the development of the single station method (SSM) – a ranging method that can be applied to Ocean Bottom Seismometer data - and have demonstrated that standard density estimation methods can successfully be applied to large sparse arrays. We have also developed new methods that aim to solve some issues linked to surveying marine mammals using acoustic devices.

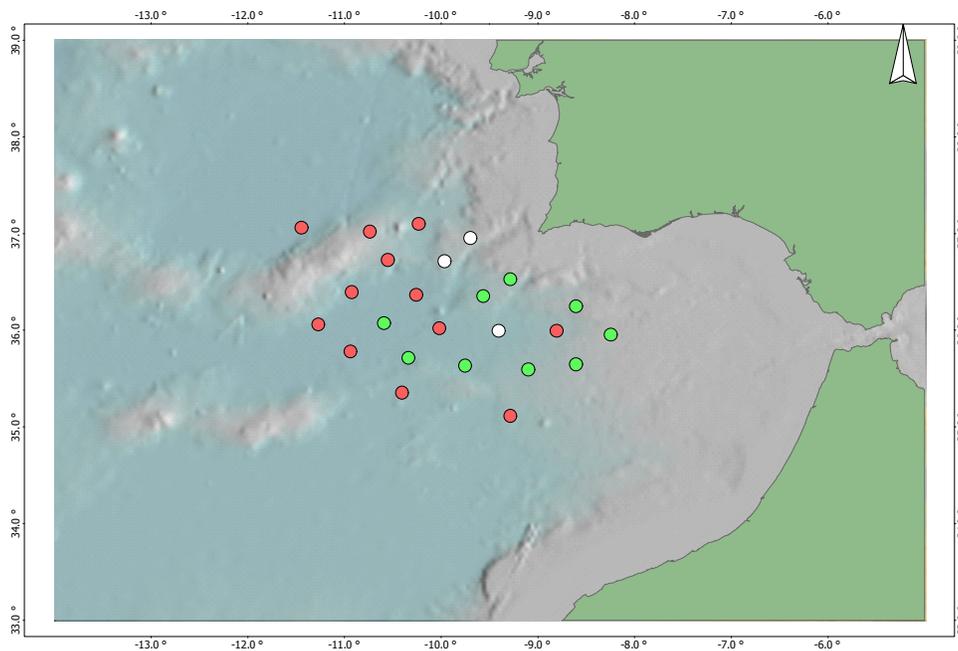
This multi-disciplinary, collaborative project has led to eight presentations and four publications (published, submitted or in prep) from the University of St Andrews (UStA) and Universidade de Lisboa (UL) research effort. The work has also generated multiple continuing collaborations with other groups using OBS data, including Washington State University (WSU) and Institut de Physique du Globe de Paris (IPGP). We have submitted data generated during this project to the OBIS-SEAMAP web portal in order to archive outputs in the public domain.

## **OBJECTIVES**

Recordings of fin whales (*Balaenoptera physalus*) from a sparse array of Ocean Bottom Seismometers (OBSs) have been used to develop and test a variety of density estimation methods. The OBS array was deployed for 1 year (2007-2008) off the south coast of Portugal, near the Straits of Gibraltar (Fig. 1).

The specific objectives of the project were to:

1. demonstrate how cue-counting methods can be used efficiently to obtain estimates of density over long time periods and large spatial scales using directional sound sensors;
2. extend the methods to allow for uncertainty in the depth of vocalizing animals;
3. develop and apply methods based on tracking moving individual animals;
4. develop and apply methods based on measuring total sound energy in relevant frequency bands;
5. obtain baseline estimates of spatial density of fin whales in the study area.



**Fig. 1** Location of the array of 24 OBS sensors in the Atlantic off Portugal. Green denotes instruments that could range to detections for their entire deployment, red denotes instruments that could not range to detections and white denotes instruments that could range to detections for parts of their deployment.

## APPROACH

This project was in collaboration with Oregon State University (grant number: N00014-11-1-0606, PI: David Mellinger). The work was divided into 3 components, each with sub-tasks. Here, we list the specific tasks relating to Components 1 & 2, i.e., those led by UStA, as we will report on these below:

Component 1: Demonstrating distance sampling methods using OBS data

The main tasks in this component included: processing the 1-year dataset, using multiple covariate distance sampling to investigate spatial and temporal patterns in density and developing methods to account for depths of whales in distance sampling-linked methods.

Task 1.1. Develop automated detector in Ishmael and run over the 1-year dataset

Task 1.2. Find, retrieve and process relevant oceanographic information for spatial covariates.

Task 1.3. Generate distances to (a sample of) calls using the semi-automated routines.

Task 1.4. Use distance sampling software, Distance (Thomas *et al.* 2010), to estimate seasonal density, incorporating covariates affecting detectability and/or density.

Task 1.5. Develop methods to account for depth distribution of whales. Test via small simulation study.

Task 1.6. Apply depth correction to OBS dataset.

Task 1.7. Write up methods and results and submit to peer-reviewed journal.

Task 1.8. Archive samples of data used on a publically-accessible internet site such as OBIS-SEAMAP or MobySound.

Component 2: Focusing on animal tracks – developing methods to account for the movement of individual animals in density estimation.

The main tasks in this component included: developing methods to simultaneously estimate detection probability, movement probability, movement parameters and animal density

Task 2.1. developing methods to simultaneously estimate detection probability, movement probability, movement parameters and animal density. Test via small simulation study.

Task 2.2. Implement simulation study to test developed methods.

Task 2.3. Generate tracks based on acoustic analysis already completed in Task 1.3.

Task 2.4. Apply methods to OBS dataset.

Task 2.5. Write up methods and results and submit to peer-reviewed journal.

Task 2.6. Archive samples of data used on a publically-accessible internet site such as OBIS-SEAMAP or MobySound.

Component 3: This component is developing a method that uses the total energy present in a species' frequency band as the statistic upon which a density estimate is made. The approach used involves a Monte Carlo simulation and propagation modeling, to link density of animals to a given received energy level.

Components 1 and 2 were led by the personnel involved with this project, and Component 3 is being led by Oregon State University. There was also a project management element, coordinating bi-monthly tele-conference progress meetings, and at least two face-to-face meetings, one in each project year.

## **WORK COMPLETED**

### **Project management and presentations**

We had regular tele-conference meetings throughout the project (approximately monthly). In addition, we had two face-to-face project meetings: one in Lisbon in 22 – 24 October 2013 and the second in St Andrews on 16<sup>th</sup> June 2013. Presentations about Cheap DECAF by UStA and UL members have been given at:

1. the 5th International Workshop on Detection, Classification, Localization, and Density Estimation (DCLDE) of Marine Mammals using Passive Acoustics held in Oregon, USA, in August 2011.
2. the 8<sup>th</sup> Symposium of Meteorology and Geophysics, Portuguese Association of Meteorology and Geophysics held in Ericeira, Portugal in March 2013.
3. the 8th Portuguese-Spanish Assembly of Geodesy and Geophysics held in Évora, Portugal in Jan 2014.
4. the 6th International Workshop on Detection, Classification, Localization, and Density Estimation (DCLDE) of Marine Mammals using Passive Acoustics held in St Andrews, UK, in June 2013 (2 presentations)
5. the UK National Centre for Statistical Ecology summer meeting held in Lowestoft, UK, in July 2013
6. the fall 2013 meeting of the Acoustical Society of America, held in San Francisco, USA.

In addition, a poster presentation was given at the European Cetacean Society conference held in Setúbal, Portugal, in April 2013.

## **Component 1**

This component of work has been focus of the research effort. As well as the tasks in the original approach (outlined above), substantial development and testing of the SSM has been conducted. Ensuring that the ranging estimates are as robust and accurate as possible has been central to all of our activities, hence we have used a variety of other datasets for validation exercises, including airgun data from the same OBS site, and fin whale data from the Atlantic and the Pacific.

- Our first task was to process the 1-year dataset and locate the fin whale calls. This proved to be an iterative process, requiring re-runs when there were updates to the detection and ranging algorithm.
- We then demonstrated that distance sampling could be applied to sparse arrays of ranging sensors. Distance sampling is a standard animal abundance estimation method where distances are measured to detected animals. The distance data are then used to estimate the probability of detecting animals. This, in turn, allows missed animals during the survey to be accounted for, leading to an estimate of absolute abundance or density (Buckland *et al.*, 2001). In a widely spaced array, where each sensor can measure ranges, distance sampling can be applied and a large survey area is covered, which is a particularly advantageous scenario.
- We specifically used cue counting, a form of distance sampling where the object of interest is not an individual animal but a cue that they produce – calls in this case. The advantage of this was we did not have to account for animal movement, which can bias density estimates, but the disadvantage was that we required an appropriate cue rate in order to estimate animal density from cue density. It became apparent that no adequate estimates of fin whale call rate had been published and so we could only estimate call density (however, please see the Follow-on Funding section below for more detail about addressing the issue of call rate in a future project). Nonetheless, we were still able to look at patterns in call density in space and time.
- We also implemented multiple covariate distance sampling (Marques *et al.*, 2007), which investigated whether various oceanographic variables influenced the ability to detect calls.
- Furthermore, the number of instruments in the OBS array allowed us to use spatial modelling techniques to (a) investigate spatio-temporal patterns of fin whale call density and (b) incorporate data from instruments that were not able to range to detected calls.
- Extensive further development and testing of the SSM was conducted throughout the project. There were two main areas to explore: (1) the accuracy of the range estimates and how these could be improved and (2) the effect of poor range estimates in a distance sampling analysis. The SSM can only reliably monitor out to a critical range. Calls beyond that range can still produce a strong detection result but produce a spurious range estimate. Therefore, we investigated how such calls could be identified and removed from the range data, which could then be used for distance sampling. Various datasets were used to test the single sensor method. We used data from (1) an airgun survey conducted at the same OBS array, (2) fin whale tracks from the Lucky Strike site at the Azores, donated by Wayne Crawford at IGPG and (3) fin whale tracks from the Cascadia OBS array in the Pacific, donated by Professor William Wilcock at WSU.

- Methods to account for animal depth in acoustic distance sampling surveys were also developed. Distance sampling is usually conducted in a two-dimensional framework, requiring the measurement of distances to detected animals in the horizontal plane. However, acoustic sampling of marine mammals poses challenges to this approach. Firstly, an animal may be at a small horizontal range but also a large vertical range from an instrument, and one would expect the vertical range to also affect animal detectability. Secondly, it is often not possible to measure the absolute location of a calling animal (from which the horizontal range could be estimated) – direct range or vertical bearing is normally estimated, but not both, due to equipment limitations. We adapted distance sampling methods to account for these issues by extending work on krill swarm density estimation (Cox *et al.*, 2011, 2013).

## **Component 2**

- We did not achieve as much as we had originally planned with respect to Component 2. This was for two main reasons. Firstly, all of the research areas in Component 1 expanded throughout the project, resulting in a greater amount of output than initially anticipated. This reduced the available time to work on Component 2. Secondly, it was work on Component 2 (specifically looking at whale tracks) that highlighted the need for further development of the detection and ranging method (for example, we noticed that calls were being missed by the detection process). Therefore, we did not complete the development of a method to account for animal movement or apply the method to the OBS data. As part of the method development task, we identified movement models developed by Gurarie & Ovaskainen (2011, 2013) as promising candidate movement models that could be applied to the fin whale tracks. These models require few parameters to be estimated, which should make them easier to incorporate into distance sampling methodology. Work on this topic is being continued by a PhD student at CREEM (funded from CREEM internal funds) who is initially investigating the amount of bias caused by animal movement in point transect sampling.

## **Other Activities**

Our work with the OBS data has led to several other activities, which we list here:

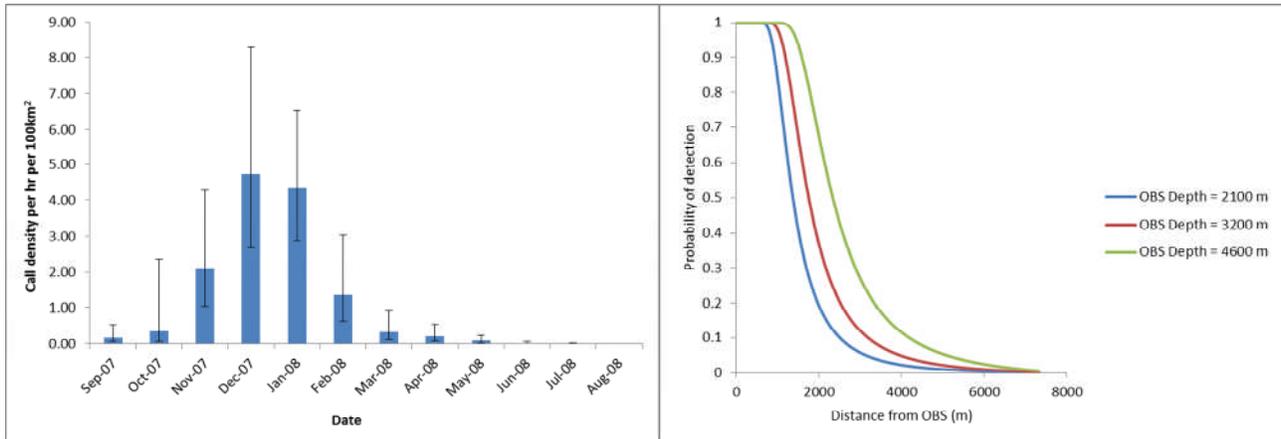
- Collaborations: The data contributions from IGPG and WSU have led to active collaborations with these groups. There are plans to use a larger dataset from IGPG for further work at UL. We also plan to continue working with WSU to compare the SSM with another method to range to whales from single OBS units developed by WSU, that relies on multipath arrivals (Weirathmueller & Wilcock, 2013). We also plan to work with WSU to implement distance sampling methods at the Cascadia array.
- PhD studentship: The Universidade de Lisboa was successful in acquiring funding for a PhD student, who will work further with the OBS data, using the dataset to (a) look in more detail at the whale tracks and (b) investigate sediment properties. The University of St Andrews are co-supervising the PhD project, which started in June 2014.
- Follow-on funding: The University of St Andrews have begun a new research effort with Penn State University, “Large Scale Density Estimation of Blue and Fin Whales”, funded by ONR. This project will also focus on sparse arrays, but where only bearing, not range, can be

measured. This project will use the OBS data as a test dataset to validate the bearing-only methods that will be developed. Furthermore, the clear lack of fin whale cue rates that prevented us from generating animal density estimates in this project directly motivated the inclusion of a cue rate estimation component in the new project. We plan to work with research groups that hold acoustic tag data for blue and fin whales and assist them in estimating cue rates that could be used in appropriate density estimation analyses.

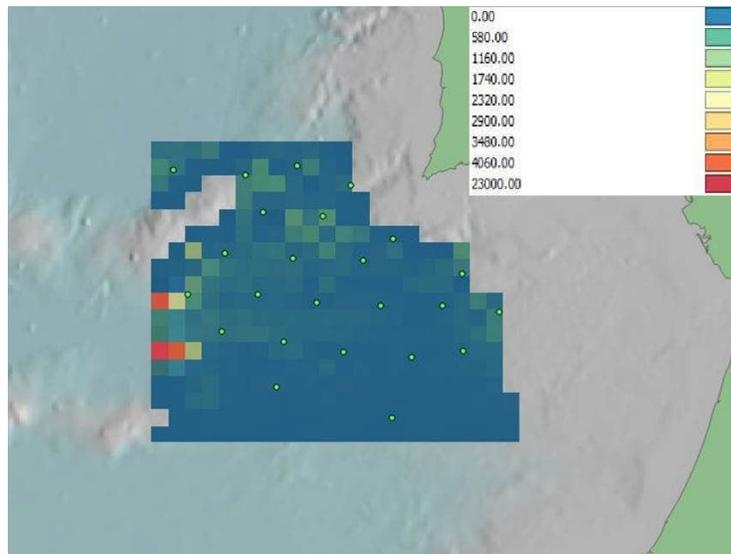
## RESULTS

### OBS data analysis

- We wrote an introductory paper to the SSM and demonstrated the method's application to distance sampling using one day of data from the OBS array (Harris *et al.*, 2013).
- We presented the first set of density results across the whole year in the FY2012 annual report and, in FY2013, we regenerated the results with an updated ranging algorithm. We used the detections to conduct multiple covariate distance sampling, which investigated spatio-temporal patterns in call detectability. Various oceanographic datasets were retrieved from online databases and used as covariates to explain patterns in detectability of the calls. These results were presented at the 6th International Workshop on DCLDE of Marine Mammals using Passive Acoustics in June 2013.
- The main results were as follows: there was a clear temporal pattern in call densities, with a peak in calling in December and January (Fig. 2). Month and depth were selected as significant covariates that affected detectability of calls. Fig. 3 shows an example detection function for the month of January at three different OBS depths. The detection function is a statistical model that predicts the probability of detecting a call as a function of range and other covariates (Buckland *et al.*, 2001, Buckland *et al.*, 2004, Marques *et al.*, 2007). Our results demonstrated that deeper OBSs are able to detect calls out to greater distances. This result makes sense, given that the maximum detection range of an OBS is partly determined by the depth of the instrument. Month is certainly acting as a proxy for an unmodelled biological process that affects detectability of calls. Ambient noise levels could be one such process but this was explicitly trialled as a covariate, producing counterintuitive results (i.e., a greater detection probability was predicted for higher ambient noise levels). This requires further investigation.
- The detection function results were used to correct detections on all OBSs (i.e., the detection numbers were increased to account for missed calls), which were then used for spatio-temporal modelling. The modelling results suggested that time of year, depth, seafloor slope, shipping levels, ocean surface velocity, sea surface temperature, latitude and longitude were all significant explanatory variables of the spatio-temporal patterns of fin whale calling activity (fitted using a Generalised Additive Model, or GAM) (Wood, 2006). This model was used to predict density surfaces across time – Fig. 4 shows an example: the density surface model predicted for 31 October – 4 November 2007. Edge effects (unexpectedly high or low predicted values) were apparent in some of the fitted surfaces (e.g., Fig. 4). These can be caused by extrapolation of the spatio-temporal model at the prediction stage. Efforts to reduce edge effects were made by restricting the prediction grid, but further refinement is required. Following the latest method developments (Matias & Harris, submitted), we plan to re-run the data and analysis for a final time to produce optimal results, before submitting the manuscript describing this 3-stage analysis.



**Fig. 2 (left) Preliminary monthly call densities with 95% confidence intervals displayed. Units are calls/100 km<sup>2</sup>. hr<sup>-1</sup>. Fig. 3 (right) Detection function for January 2008, at three different OBS depths.**

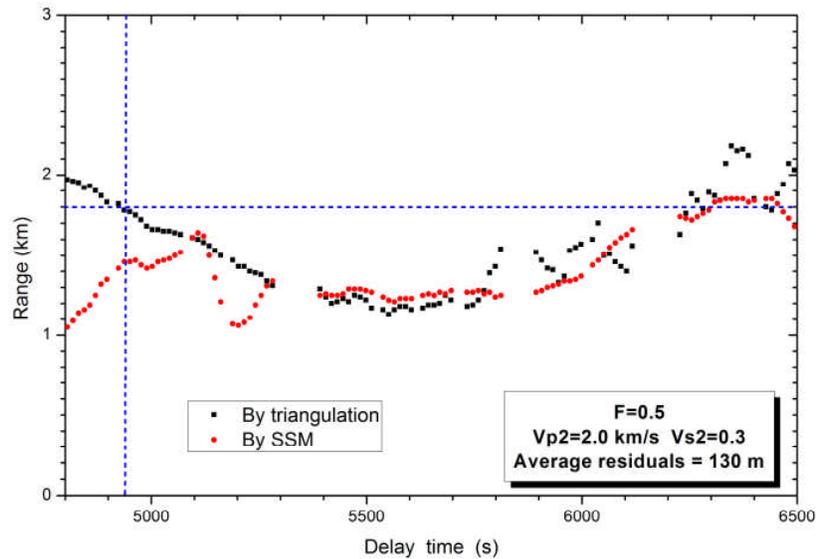


**Fig. 4 Predicted fin whale call counts between 31 Oct and 4 Nov 2007 in 14 km<sup>2</sup> grid cells across the wider study area. The grid is irregular due to shallow bathymetry – predictions were not made where ocean depth <2000 m as this would have been extreme extrapolation of the model. The green circles indicate the OBS locations.**

### Method development

- Results from further development of the SSM were also presented at 6th International Workshop on DCLDE of Marine Mammals using Passive Acoustics in June 2013. Work has continued since then and the most recent results are given in Matias & Harris (submitted).
- The main results are as follows: the accuracy of the ranging algorithm was improved by further considering the velocities in the sediments and the water column, as well as applying an additional correction factor to account for gain differences between the vertical and horizontal

channels. The optimal set of parameters were applied to a fin whale track from the Lucky Strike site (Fig. 5).

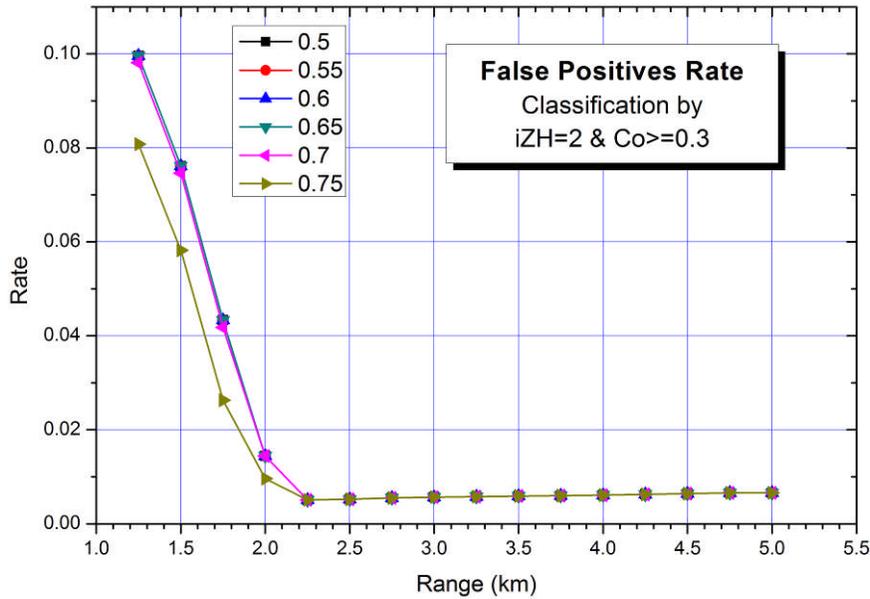
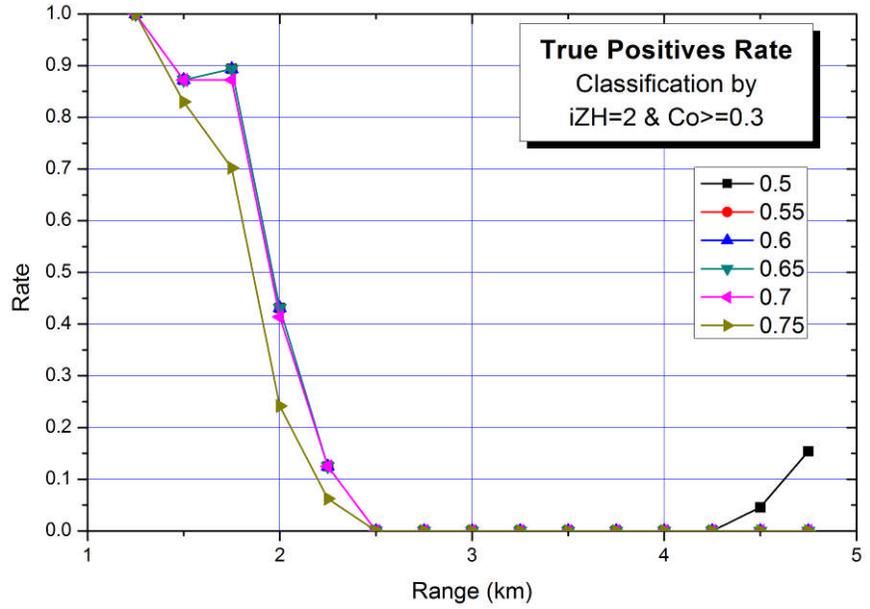


**Fig. 5 Comparison of true and measured offsets for the Lucky-Strike fin-whale calls using the best choice of elastic parameters in the sediments and an amplification factor ( $F$ ) of two is applied to the vertical channel.**

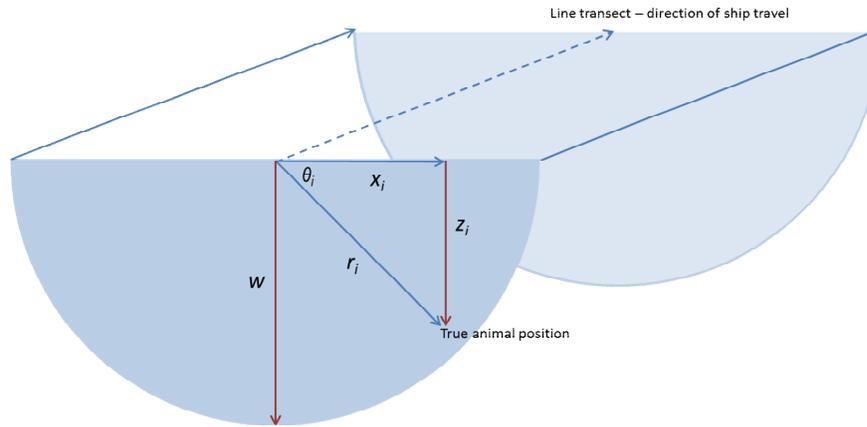
- Further investigation of selection criteria introduced in Harris *et al.* (2013) has shown that both (1) a measure of coherency and (2) the time lag between the vertical and horizontal channels are effective at removing both true fin whale calls that occur beyond the critical range and false detections (Fig. 6).

### **Accounting for depth**

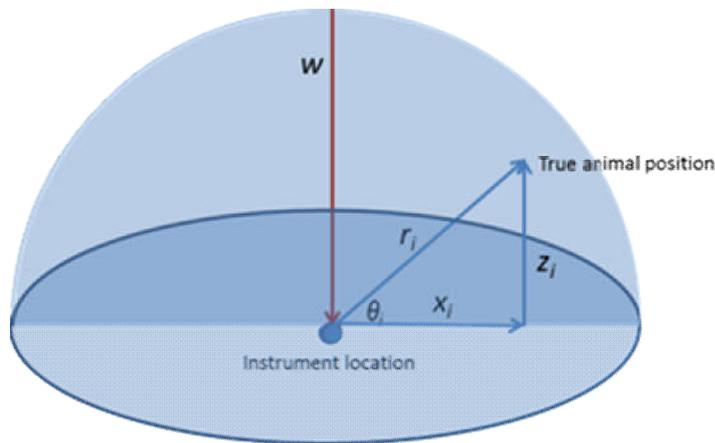
- The method development work to account for animal depth involved the simulation of several monitoring scenarios with both towed and fixed sensors (Figs. 7 & 8). The developed methods altered the way in which the detection function parameter was estimated, by using either a direct range ( $r$ ) or bearing ( $\theta$ ) and assuming that the distribution of animals with respect to depth ( $z$ ) was known. In both survey scenarios, (1) the positions of animals with an assumed depth distribution were simulated, (2) a sub-sample of animals were “detected” using a known detection function (a half-normal distribution) and (3) those observations were used to re-estimate the detection function parameter, which enabled density to be estimated.



**Fig. 6 True and False Positive Rates (TPR & FPR) computed for direct paths originated from whale calls, as a function of distance, for different values of the correlation threshold. Each data point represents the TPR inside a bin 1 km wide. top) TPR after classification by the delay between the Z and H channels ( $iZH$ , units are samples), and also the polarization coherency ( $Co$ , units are arbitrary); bottom) FPR estimated for each bin as a function of distance, for different values of the correlation threshold**



**Fig. 7 Schematic of a towed acoustic survey. The 2D location of animal in the vertical plane,  $i$ , can be defined as  $(x_i, z_i)$  where  $x_i$  is the horizontal distance of the animal from the transect line and  $z_i$  denotes animal depth. The animal's position along the transect line is ignored. The slant range is denoted as  $r_i$  and vertical bearing is denoted as  $\theta_i$ . The maximum detection range used for analysis is defined as  $w$ .**



**Fig. 8 Schematic of a fixed acoustic survey. The 2D location of animal in the vertical plane,  $i$ , can be defined as  $(x_i, z_i)$  where  $x_i$  is the horizontal distance of the animal from the monitoring point and  $z_i$  denotes animal depth. The horizontal bearing of the animal's position is ignored. The slant range is denoted as  $r_i$  and vertical bearing (specifically, elevation angle) is denoted as  $\theta_i$ . The maximum detection range used for analysis is defined as  $w$ .**

- The main results were as follows: a uniform depth distribution was initially used as the simplest case but could not be implemented with bearing data alone – range data were required due to the formulation of the algorithm to estimate the detection function parameter. However, this was not an issue for alternative depth distributions. A scaled beta distribution was selected to represent cetacean depth distributions - it is very flexible so can take many forms without the need for the user to update software with new distributions. In general, the methods worked well for adequate sample sizes. Smaller sample sizes,

particularly when using bearing data, should be avoided - two of the smaller sample sizes generated very biased results (>300%). Bearing data tended to generate more biased results than when range data were used; this is likely due to the fact that the detection function was a function of range only. All results from larger sample sizes generated bias of less than 10%.

**Table 1. Results from a simulation study to measure the bias in density estimates using new methods to deal with animals at depth. Each simulation was run 1000 times. Mean percentage bias is reported with the standard error given in brackets. Two different sample sizes were used for each simulation. The mean number of observations in each simulation is denoted by n.**

Simulation number	Survey type	Data type	Depth distribution	% Bias in density (animals per km <sup>2</sup> )	
1	Towed	Range	Uniform	<b>1.2 (0.44), n = 98</b> 5.8 (1.1), n = 20	
2	Towed	Bearing	Uniform	N/A	
3	Fixed	Range	Uniform	<b>1.4 (0.54), n = 92</b> 7.4 (1.2), n = 18	
4	Fixed	Bearing	Uniform	N/A	
Simulation number	Survey type	Data type	Depth distribution	% bias Shallow divers	% bias Deep divers
5	Towed	Range	Scaled beta	<b>-1.1 (0.20), n = 311</b> 0.38 (0.45), n = 62	<b>1.2 (0.43), n = 138</b> 6.0 (1.1), n = 27
6	Towed	Bearing	Scaled beta	<b>1.7 (0.30), n = 312</b> 3.1 (0.64), n = 62	<b>8.8 (1.3), n = 138</b> 515 (326), n = 28
7	Fixed	Range	Scaled beta	<b>-1.3 (0.26), n = 292</b> -0.17 (0.58), n = 59	<b>1.9 (0.54), n = 130</b> 6.7 (1.2), n = 26
8	Fixed	Bearing	Scaled beta	<b>5.5 (0.49), n = 293</b> 8.4 (1.1), n = 58	<b>6.6 (1.2), n = 130</b> 307 (175), n = 26

- These are two general scenarios and so natural extensions to this work will be to consider other surveying scenarios. In particular, incorporating data from instruments at different depths into one analysis would be very useful and relevant to many passive acoustic surveys (including OBS arrays).

## DELIVERABLES

- We have produced annual reports in FY11 – FY13, as well as this final technical report.
- We have published one paper introducing the SSM and its application to distance sampling analyses (Harris *et al.*, 2013)
- The further method development has been written up and has been submitted to the Journal of the Acoustical Society of America (Matias & Harris, submitted).

- A stand alone manuscript detailing the methods developed to account for depth has been prepared for submission to the Journal of Agricultural, Biological and Environmental Statistics (Harris *et al.*, submitted).
- A further manuscript is in preparation and will incorporate an update analysis of the 1-year dataset.
- We have sent OBS locations and survey effort data to the OBIS SEAMAP web portal. The final results of the 1 year analysis will also be submitted once complete.

## IMPACT/APPLICATIONS

The main aim of Cheap DECAF was to make density estimation of cetaceans less costly and, therefore, more accessible to the wider scientific community. The methods developed here will be applicable to re-deployable arrays of both sea-bed mounted instruments (such as OBS arrays) and surface buoys, so should increase our capability to monitor cetacean density in geographic areas of interest, including those where naval operations are conducted.

## RELATED PROJECTS

Cheap DECAF (Grant number: N00014-1-11-0606, PI: David Mellinger, Oregon State University)

## REFERENCES

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L. & Thomas, L. (2001). *Introduction to Distance Sampling*. Oxford University Press, Oxford.

Buckland, S. T. Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. & Thomas, L. eds. (2004) *Advanced Distance Sampling*. Oxford University Press, Oxford.

Cox, M.J., Borchers, D.L., Demer, D.A., Cutter, G.R. & Brierley, A.S. (2011) Estimating Antarctic krill (*Euphausia superba*) density from multibeam echosounder observations using distance sampling methods. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 60, 301–316.

Cox, M. J., Borchers, D. L., Kelly, N. (2013), nupoint: An R package for density estimation from point transects in the presence of nonuniform animal density. *Methods in Ecology and Evolution*, 4: 589–594. doi: 10.1111/2041-210X.12058

Gurarie, E. & O. Ovaskainen. 2011. Characteristic spatial and temporal scales unify models of animal movement and fundamental ecological processes. *The American Naturalist*. 178:113-23.

Gurarie, E., O. Ovaskainen. 2013. Towards a general formalization of encounter rates in ecology. *Theoretical Ecology*.

Harris, D. V., Matias, L., Thomas, L., Harwood, J. & Geissler, W. (2013) Applying distance sampling to fin whale calls recorded by single seismic instruments in the northeast Atlantic. *Journal of the Acoustical Society of America* , vol 134 (5), 3522-3535.

Marques, T.A., Thomas, L., Fancy, S.G. and Buckland, S.T. (2007) Improving estimates of bird density using multiple covariate distance sampling *The Auk* **124**: 1229-1243.

Matias, L & Harris, D (submitted) The single-station method: detection, classification and location of fin-whale calls using ocean-bottom seismic stations. Submitted to JASA.

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

Weirathmueller, M. & Wilcock, W. S. D. (2013) An automatic single station multipath ranging technique for 20 Hz fin whale calls: Applications for ocean bottom seismometer data in the northeast Pacific. *Meeting of the Acoustical Society of America (ASA)*, San Francisco, CA, Dec 2-6. *Conference abstract*

Wood, S. N. (2006) *Generalised Additive Models: An Introduction with R*. Chapman & Hall, Boca Raton, FL.