REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
28-03-2008	Final Technical Report	01-07-2004 - 31-12-2007

4. TITLE AND SUBTITLE Development of New Methods and Software for Distance Sampling Surveys of Cetacean Populations	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** N00014-04-1-0661 **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Thomas, Len, Dr. Borchers, David, L, Dr. Rexstad, Eric, Dr. Buckland, Stephen, T., Prof. Bravington, Mark, L., Dr. Wood, Simon, N., Dr. Hedley, Sharon, L., Dr.	**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of St. Andrews College Gate St. Andrews, Fife, UK. KY16 9AJ.	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research 875 North Randolph Street Arlington, VA 22203-1995	**10. SPONSOR/MONITOR'S ACRONYM(S)** ONR **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Obtaining reliable estimates of the density and distribution of cetacean species is an essential component of a risk mitigation strategy, as well as having other conservation and management uses. The overall goal of this research program was to develop statistical methods and software that substantially enhances the utility and robustness of current survey methods. To achieve this, four research projects were undertaken, to develop methods for:
1. analysis of towed passive acoustic and combined visual-acoustic surveys;
2. improved modeling of animal distribution from survey data;
3. improved modeling of spatial distribution of group size (for animals that cluster); and
4. more efficient survey designs that utilize information from the above models to direct sampling.
Here, we report our findings, list the research outputs and give recommendations for future research directions.

15. SUBJECT TERMS

Ecological statistics; Cetacean density; Population size; Environmental risk mitigation; Environmental monitoring; Line transect surveys.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Len Thomas
					19b. TELEPHONE NUMBER (Include area code)
U	U	U	UU		UK+1334 461801

FINAL TECHNICAL REPORT
SUBMITTED TO THE
US OFFICE OF NAVAL RESEARCH

Development of New Methods and Software for Distance Sampling Surveys of Cetacean Populations

PRINCIPAL INVESTIGATOR:
Dr Len Thomas,
Centre for Research into Ecological and Environmental Modelling (CREEM),
University of St Andrews,
St Andrews, UK, KY15 5LY.
Tel: UK+1334 461801; Fax: UK+1334 461800; Email: len@mcs.st-and.ac.uk

CO-INVESTIGATORS:
Dr David Borchers, CREEM, St Andrews, UK.
Dr Eric Rexstad, CREEM, St Andrews, UK
Professor Stephen Buckland, CREEM, St Andrews, UK.
Dr Mark Bravington, CSIRO, Hobart, Australia.
Dr Simon Wood, University of Bath, Bath, UK.
Dr Sharon Hedley, Statistical consultant, St Andrews, UK.

COLLABORATORS:
Dr Jonathan Gordan, Ecologic, Tayport, UK.
Dr Douglas Gillespie, Sea Mammal Research Unit, St Andrews, UK.
Dr Justin Matthews, International Fund for Animal Welfare, Yarmouth, UK.
Dr David Peel, CSIRO, Hobart, Australia.

ONR AWARD NUMBER: N00014-04-1-0661

PERIOD OF PERFORMANCE: 1 July 2004 - 31 Dec 2007

Executive Summary

Obtaining reliable estimates of the density and distribution of cetacean species is an essential component of a risk mitigation strategy, as well as having other conservation and management uses. The overall goal of this research program was to develop statistical methods and software that enhances the utility and robustness of current survey methods. To achieve this, research projects were undertaken, to develop methods for:

1. analysis of towed passive acoustic and combined visual-acoustic surveys;
2. improved modeling of animal distribution and density from survey data;
3. improved modeling of spatial distribution of group size (for animals that cluster); and
4. more efficient survey designs that utilize information from the above models to direct sampling.

Here, we report our findings, list the research outputs and give recommendations for future research directions.

For project 1, we developed methods based on counting cues (vocalizations) and tested them by simulation. Theory for estimation based on distinguishing individuals was developed but its implementation was less successful. The theory did, however, lead to development of a statistically rigorous method for the analysis of survey data with intermittent availability, which has general applicability in many areas of wildlife survey methodology. Development of methods for combined visual-acoustic analysis was hamstrung by data limitations: our case study data on harbor porpoise contained too few duplicate detections, despite hagin used a survey protocol customized for combined analysis. We recommend development of practical bow-mounted hydrophones for this type of survey. We also developed new methods for estimating the power of towed acoustic studies to estimate trend when there is randomness in the bias of acoustic estimates (calculations which neglect randomness produce over-optimistic power estimates).

For project 2, we developed a new method of spatial smoothing that overcomes many of the difficulties associated with previous methods. This method has extremely wide applicability in many diverse applications. We also developed a comprehensive framework for the analysis complex survey data, with the goal of applying it to a challenging set of data on Antarctic minke whales supplied by the International Whaling Commission, as well as simulated data with similar characteristics. A preliminary analysis has been completed, and full analysis is ongoing. We have tackled issues such as uncertain estimation of group size that we did not envision in our research plan, but that are likely more common in real data than is generally admitted.

For project 3, after some experimentation, we decided to use similar spatial smoothing methods for estimating variation in group size as those used in project 2 for estimating variation in group density. This led to a comprehensive unified framework, and these two projects are treated together in this report. There is much still to do in this area to make the methods generally applicable, and we propose a technical workshop to prioritize future research directions, which will be held in St Andrews in May 2008.

For project 4, we developed the required estimation framework and new survey designs, and then tested the potential increase in precision that could be expected by using these designs to focus survey effort into areas shown to have higher density from previous surveys. The results were disappointing, and we conclude that in most cases such designs are not worth the additional complication required to implement them. Nevertheless, the new analysis methods are now available in a test version of the industry standard software Distance.

Table of Contents

Background

Some of the most fundamental questions we can ask about a wild animal population are: "How many are there?" and "Where do they live?". In the case of cetacean species, answering these questions by obtaining spatially and temporally referenced estimates of density and abundance are not only fundamental to understanding their basic biology, but also to monitoring and mitigating the effect of man-made impacts on their populations.

The most widely used method of estimating cetacean population size at large spatial scales is using distance sampling survey methods. In the standard methods (Buckland et al. 2001), a series of lines (or points) are placed within the study area according to a randomized survey design. An observation platform (usually a ship, airplane or helicopter for cetacean surveys) traverses the lines, and all detected animals of the target species are recorded, together with their perpendicular distance from the line. These distances are then used to estimate the proportion of target animals within the searched area that were missed.

The standard methods works well for many species and situations. Nevertheless, there are a number of limitations that this research has aimed to address:

- Standard approaches assume that detection of animals exactly on the line is certain – an assumption often violated in visual cetacean surveys where animals dive for a significant period of time. This assumption can be achieved in some cases using passive acoustic surveys, or bypassed using multiple observation platforms (where proportion of animals detected on the line can be estimated). Multiple platform methods are most likely to succeed when the two observations platforms use independent detection methods, such as one being visual and the other passive acoustic. Hence one part of this research (Project 1) investigated the utility of towed passive acoustic distance sampling surveys either as an alternative single platform, or in conjunction with a visual platform.
- Standard methods are "design-based" – i.e., they assume a randomized survey design. This is not correct when observers are placed on vessels that are operating for other purposes (e.g., ferries, oceanographic vessels, etc.). In these cases, model-based estimation must be used – resulting in a predicted spatial density surface. This approach has a number of other potential advantages that mean it may be of use even when design-based approaches are feasible. A disadvantage of model-based methods is that they rely crucially on the quality and correctness of the models used, and there are significant problems with current methods, which are based on various methods of smoothing covariates (e.g., spatial smoothing). The aim of another part of this research (Project 2), therefore, was to address these problems by developing new model-based methods.
- When animals occur in groups, rather than as individuals, a spatial group size surface must also be produced and integrated with the density surface for groups to produce a predicted spatial density surface of individuals. Developing methods to achieve this was a third part of the research (Project 3).
- When design-based approaches are possible, there is potential to increase survey efficiency relative to standard randomized designs by allowing more survey effort in areas expected to have higher density of animals. The efficacy of such designs was investigated in the third part of the research (Project 4).

This research was led by members of the Research Unit for Wildlife Population Assessment (RUWPA), a research group within the Centre for Research into Ecological and Environmental Modelling (CREEM) at the University of St Andrews. The St Andrews group are at the forefront of methodological research in this area and are also very active in promoting adoption of best practice among survey biologists, through the provision of reference books (Buckland et al. 2001, 2004; Borchers et al. 2002), training workshops and the industry-standard software, Distance (Thomas et al., 2006). The work was carried out in collaboration with statisticians and cetacean experts from the UK and Australia.

Further background on each of the issues discussed above is given in the project proposal (Thomas 2003), and in the "Technical Approach and Findings" section below.

Objectives

Our overall objective was to develop statistical methods and software that substantially enhance the utility and robustness of distance sampling surveys of cetacean population density and abundance. The work was divided into four projects, each with its own specific objectives, as laid out below.

Project 1 Objectives – Acoustic and visual-acoustic methods

- Develop methods for estimating cetacean density from towed passive acoustic detectors, focusing on surveys of sperm whales. Extensions to standard methods for visual line transect surveys are required to deal with:
    - Measurement error in estimates of distance to animal using passive acoustics
    - Inability to distinguish individuals
- Develop methods for estimating cetacean density from combined visual-acoustic surveys, focusing on the SCANS-II survey of harbor porpoise, where the survey protocol was specifically designed to enhance the overlap between visual and acoustic detection zones. Extensions to current methods are required to deal with uncertainty in determining whether animals seen and then heard are the same individuals.

Project 2 Objectives – Spatial density surface estimation

- Extend current methods for model-based spatial density surface estimation to deal with outstanding issues that prevent reliable use of model-based methods:
    - Inability to deal with irregular topography (indented coastlines and islands)
    - Extreme estimates with associated high variances, arising from smoothers that show wild behavior away from the survey trackline
    - Inability to deal with spatial autocorrelation, i.e., the clustering of detections at small spatial scales
    - How best to do model selection (smoothing parameter selection), especially given the previous point.
    - How best to incorporate uncertainty in estimated strip widths.
- Apply these extensions to one or more key case studies.

Project 3 Objectives – Spatial patterning in group size

- Develop and implement parallel extensions to those of project 2 that allow incorporation of spatial patterns in group size, for cetacean species that occur in identifiable groups, clusters, pods, etc.

Project 4 Objectives –Improved design-based survey design and estimation

- Extend current survey design methods to allow unbiased estimation of density when planned coverage varies spatially (for example with more coverage planned in areas where density is expected to be higher, based on previous survey results).

- Develop new survey designs that place more effort in high density areas, and test the efficiency of these new designs in terms of increased precision of estimates.

Additional Objectives

- For projects 1 and 4, an additional objective was to implement the new methods in the Distance software, in order to make them available to the user community. This was not an objective for projects 2 and 3, due to funding limitations at the time the research proposal was negotiated.
- The proposal included an additional project as an option: the development of a simulation capability in the Distance software. The option was not exercised during the course of this research, but is recommended for future implementation (see Conclusion and Recommendations, below).

Technical Approach and Findings

Project 1 – Acoustic and visual-acoustic methods

A research assistant, Ciara Brewer, was hired to work exclusively on this project and in the course of the project a further two research assistants were employed part-time to assist with specific aspects of the project. In addition, Drs Doug Gillespie and Jonathan Gordon were employed part-time as consultants on sperm whale click behavior and towed hydrophone methods. A collaboration with Justin Matthews of the International Fund for Animal welfare was established to pursue common interests in animal-based estimation method development. We also collaborated with the SCANS II acoustics team for work relating to joint visual-acoustic survey and analysis methods.

Acoustic methods applied to sperm whales

Two-element towed hydrophones are able to detect vocalizing sperm whales over much greater distances than visual observers and for much longer periods. However, there are a number of problems associated with applying distance sampling methods to data gathered from two-element towed hydrophones which do not present themselves in the case of visual survey data. Primary among these are the following.

1. Detection of a single click by the hydrophone provides no useful information on the distance of the source from the hydrophone. Detection of a series of clicks from the same source allow probabilistic estimation of distance, albeit with some uncertainty.
2. Assignment of click "trains" to individuals can be difficult, especially when there are a number of animals in close proximity to one another.

Two approaches were investigated to deal with these difficulties: likelihood-based approaches and cue-count-based methods. We now briefly describe these.

Likelihood functions and estimation methods which in incorporate probabilistic components for both 1. and 2. above were developed. These involved an extension of hidden Markov model (HMM) estimation methods to include a distance-based probabilistic detection process. This is illustrated diagrammatically in Figure 1. Conventional HMMs involve only the "Hidden states" and "Availability" levels shown in the figure, and conventional HMM estimation methods require the availability (A_t) in the figure to be observed. The additional "Detection" layer is required because the availability state is not always observed (i.e., not all clicks are heard).

The likelihood-based approach is very parameter-intensive (details in Appendix 1: Brewer *et al.* 2007), and while it is conceptually appealing, it proved difficult to implement in practice. Part of the difficulty arose because the sperm whale click process turned out to be such that it can't be modeled very well using HMMs. Sperm whales click too regularly for HMMs to provide good models. Variants of HMMs which allow for a more regular availability process were investigated (Zucchini *et al.*, 2007) but these add substantially to the complexity of an already complex model and although they are more flexible, they were not sufficiently flexible to capture the regularity of the sperm whale click process. The extended HMM (as shown in Figure 1) did, however, prove useful for development of more general methods for dealing with intermittent animal availability on line transect surveys (see "*Development of general methods for intermittent availability*" section below).

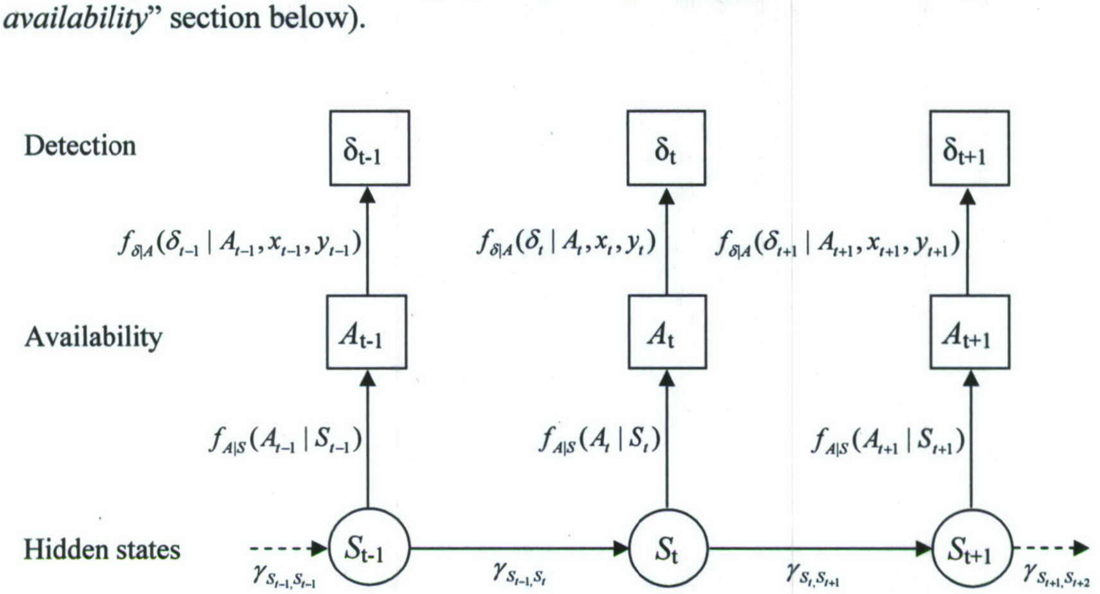

***Figure 1**: Extension of hidden Markov model (HMM) to incorporate a distance-based detection probability model. Here δ_t is detection at time t, A_t is availability at time t, and S_t is a notional whale "state" at time t. In addition, $f_{\delta|A}(\delta_t \mid A_t, x_t, y_t)$ is the distance-based detection probability model, x and y are perpendicular and forward distances, $f_{A|S}(A_t \mid S_t)$ is the availability process model given whale "state", and γ_{S_{t-1},S_t} is the probability process governing movement between whale "states"*

A likelihood-based distance estimation method was developed for estimating distances from the hydrophone to detected clicks, and this was applied successfully to simulated and real data (Appendix 1). It was able to obtain estimates for some cases in which an existing non-likelihood method could not.

The output from this distance-estimation method provided the basic data for a second approach to sperm whale abundance estimation using distance methods. This approach involved customizing cue-counting distance sampling methods for use in this context. Two variants of cue-counting methods were investigated. The first ("cue-counting method 1") was based on a likelihood function which incorporated probabilistic estimation of distance, while the second ("cue-counting method 2") used conventional cue-counting methods, neglecting the uncertainty attached to distance estimation. The uncertainty in distance estimation was found to be sufficiently small that only insignificant and insubstantial bias was introduced by neglecting this uncertainty. Methods were tested by simulation and were found to perform reasonably well . Figure 2 shows examples of simulation output. Further details are given in Appendix 1.

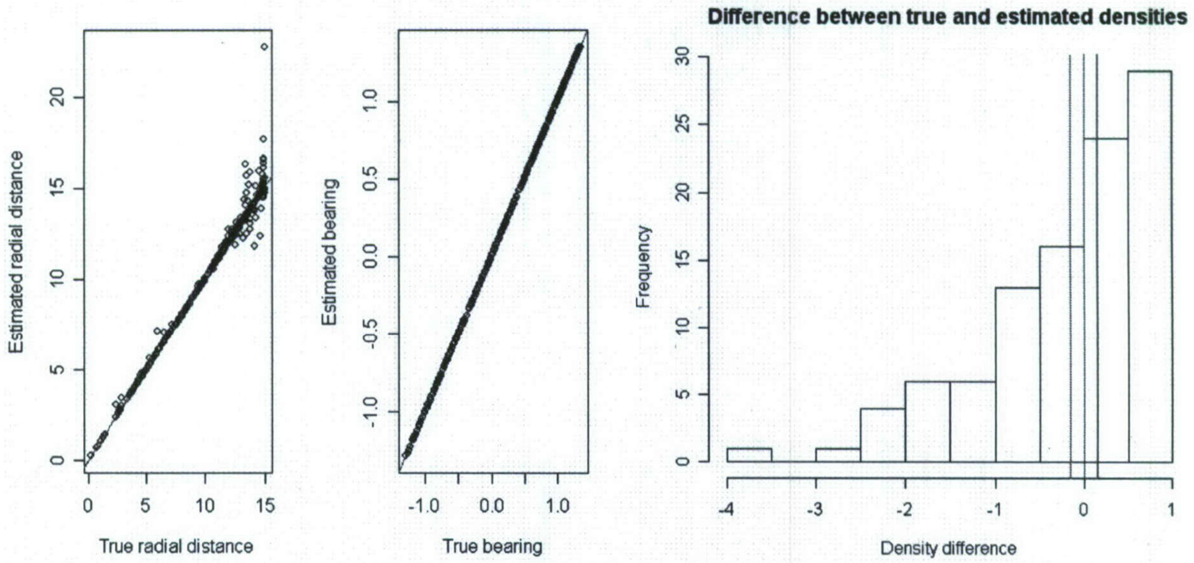

Figure 2. Simulation testing of cue-counting method 2. Estimated vs. true radial distances and bearings and histogram of difference between true and estimated density for: a) concentration parameter k=19,000 and b) concentration parameter k=5,000. (Legend for histograms: green line – mean density difference, red line – true density, blue line – median density difference)

Visual acoustic methods applied to harbour porpoise

Combined visual and passive acoustic line transect surveys have the potential to improve density estimation in a number of ways. The first is to provide a pair of truly independent "observers" (one visual, one acoustic) to enhance mark-recapture-type line transect density estimators, the second is to enhance trend estimation. These are dealt with separately below.

Negative bias in estimating density using mark-recapture line transect (MRLT) methods with a pair of nominally independent observers arises when both observers preferentially detect the same animals. For example, when both observers detect animals using visual cues, they both preferentially detect the more visible animals. Line transect methods have recently been developed, and continue to be developed (Borchers *et al*, 2008, for example) to deal with this by modeling the detection process, but it has been shown (Link, 2003) that it is not possible to tell whether or not modeling has been entirely successful in this respect.

The bias does not arise if the detection process for one observer is unrelated to that of the other observer. Use of MRLT methods with one visual and one acoustic observer therefore has the potential to reduce or remove this bias. However, such MRLT methods rely heavily on correct identification of "duplicates" (animals detected by both observers) and this is much more difficult with visual-acoustic MRLT surveys than with visual-visual MRLT surveys. The difficulty arises primarily because (a) the passive acoustic observer has limited capacity to localize detections, (b) the area searched by each observer is are much more separated in time and space than is the case with a pair of visual observers.

Being aware of this difficulty, we were able to get visual observers on ships participating in the 2005 international Small Cetaceans in the European Atlantic and North Sea (SCANS II) survey to implement a search procedure which involved following detected harbour porpoise past abeam an into the area covered by towed passive acoustics (the acoustic "observer") – so as to maximize our chances of obtaining data adequate for identifying duplicates probabilistically. Our participation in SCANS II project meetings and analyses gave us access to the data and provided a forum for reporting results and receiving feedback. In the course of the project we also developed closer collaboration with the SCANS II acoustic team.

The fact that visual observers' search area was designed to overlap that of the acoustic observers on SCANS II gave us something of a best-case scenario for developing combined visual-acoustic MRLT methods. However, the data proved inadequate to justify investing resources in method development and after exploratory analyses, this was not pursued further. Details are given in Appendix 2 (Borchers 2007). An example of an exploratory plot from one vessel is given in Figure 3. The essential information which would allow duplicates to be detected probabilistically is a surplus of short inter-observer waiting times between detections compared to that expected from the within-observer waiting times between detections. The within-observer waiting times are shown in the top two plots below, while the between-observer waiting times are shown in the bottom plot. Duplicate detections would be indicated by a surplus of observed short times compared to expected short times. There appears to be inadequate information in these data to give a reasonable chance of further method development improving estimation substantially.

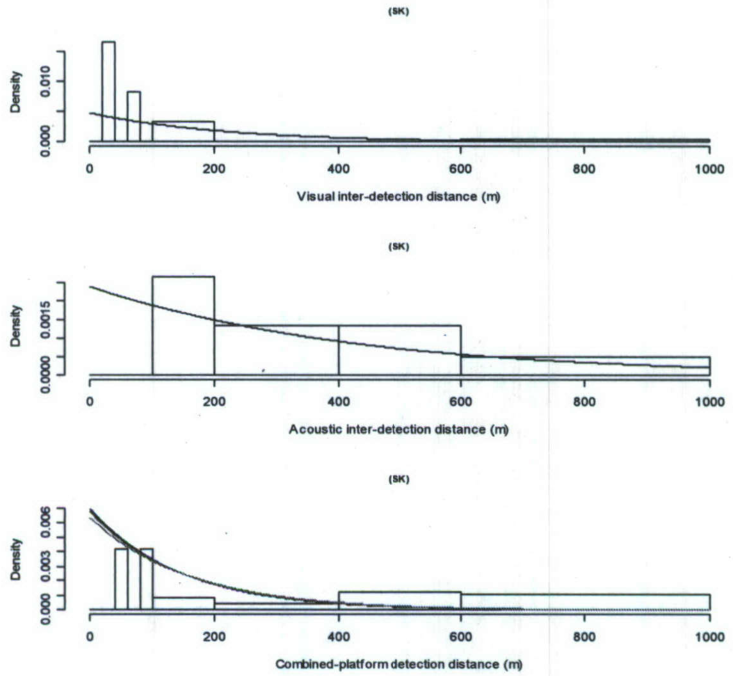

***Figure 3**: Observed (histograms) and expected (curves) waiting times between detections for the visual observer (top plot), acoustic observer (middle plot) and combined observers (bottom plot) on vessel "SK". Curves in the top two plots are fitted to the histogram. Curves in the bottom plot are based on the two fitted curves and mean visual probabilities of detection of 0.1 (black curve), 0.18 (red curve) and 0.35 (green curve). (These are the approximate minimum, mean and maximum estimated detection probabilities for this vessel.)*

The second potential use of passive acoustic data for surveys of harbour porpoise is for trend estimation. Because acoustic observers are much easier and cheaper to deploy on vessels than are visual observers, and because they tend to be more consistent in their detection behavior, they provide a more affordable means of gathering data at many points in time than do visual surveys. On their own they are less able to provide unbiased estimates of density or abundance than visual MRLT surveys. However, acoustic and visual surveys can be combined over time to exploit the strengths of each method: acoustic survey data provide a time series of biased density estimates and visual surveys are used to calibrate the biased density estimates. Using SCANS II data, in collaboration with the SCANS II acoustic team, we investigated the relative efficiency and power of acoustic and visual line transect survey data for trend estimation. Analyses of these data highlighted an issue usually overlooked in power analysis with indices of relative density, namely the potentially stochastic nature of the bias of the index. The uncertainty associated with the bias is apparent in Figure 2. This led to development of a method for power estimation which accommodates uncertainty in density estimation bias over time. Details are in Appendix 3 (Borchers and Burt 2007).

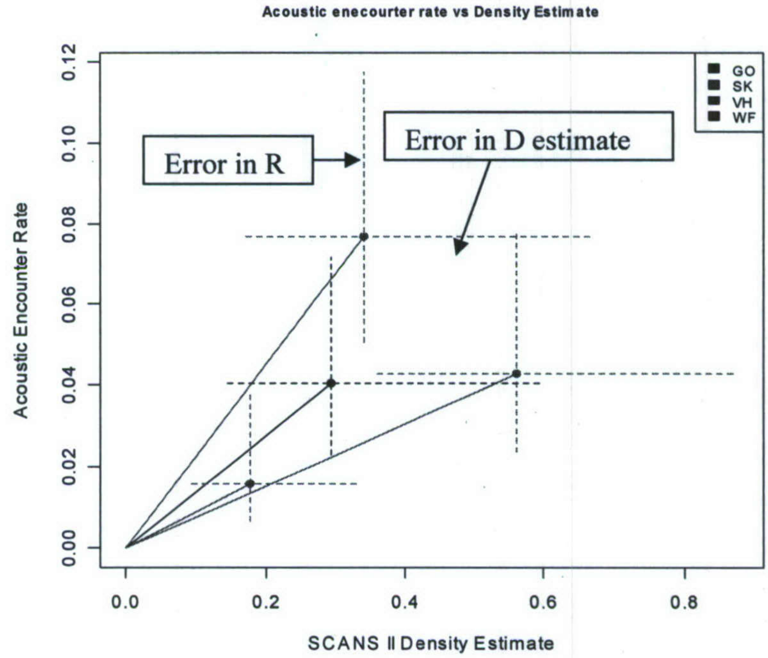

***Figure 4**: SCANS II acoustic encounter rate (R) regressed on absolute density estimate (D) for harbour porpoise for 8,500 km of simultaneous towed acoustic and visual ship survey effort, using data from four vessels: GO (black), SK (purple), VH (brown) and WF (red). Dashed lines are 95% Confidence Intervals, assuming log-normality. The slops of the lines are estimates of acoustic survey bias. As well as being different between vessels, these are estimated with considerable uncertainty (as indicated by the vertical and horizontal error bars). Acoustic data are from all sea states ≤5. Acoustic data were corrected based on vessel noise correction factors calculated for each vessel.*

Development of general methods for intermittent availability

Animals are missed on survey either because they are unavailable for detection or because they are available but not detected. Line transect methods exist for scenarios in which animals are continuously available for detection or are either entirely available or entirely unavailable for the duration of the period they are within detectable range. However, rigorous statistical methods which accommodate intermittent availability while within detectable range are lacking. Most methods for this case are at least partly ad-hoc.

The intermittent availability of sperm whales for acoustic detection (because they do not click all the time) is one example of a more general intermittent availability problem on line transect surveys. Another is the intermittent availability of many cetacean species for visual detection on shipboard and aerial line transect surveys. A spin-off of the method development conducted for the sperm whale density estimation problem with passive acoustics is the development of more general methodology for line transect surveys with intermittent availability. In addition to providing improved estimation methods, this has identified a previously neglected source of bias in a commonly-used ad-hoc method of correcting for availability. Details are contained in Appendix 4 (Borchers and Samarra 2007). This bias is illustrated in Figure 5. Intermittent availability bias is usually corrected by using some function of only the (estimated) proportion of time animals are available (and no other information on the availability process). Figure 5 shows probabilities of detection for two simulated surveys, with identical observers, surveying animals which are available for detection exactly 20% of the time. The difference between the two curves results purely from the fact that the pattern of availability differs in the two cases: the red curve corresponds to frequent short availability while the blue curve corresponds to less frequent but longer availability.

It is impossible for any availability correction method based only on the proportion of time animals are available to correct for “availability bias” in both cases – and such correction methods could produce biased corrections for both cases. Clearly the pattern of availability must be taken into account to have any prospect of removing availability bias in general. The method given in Appendix 4 does this.

Methods based on the extended HMM model illustrated in Figure 1 were developed, together with an alternative method based on resampling samples of the availability processes. While the latter approach is less elegant, it is more flexible.

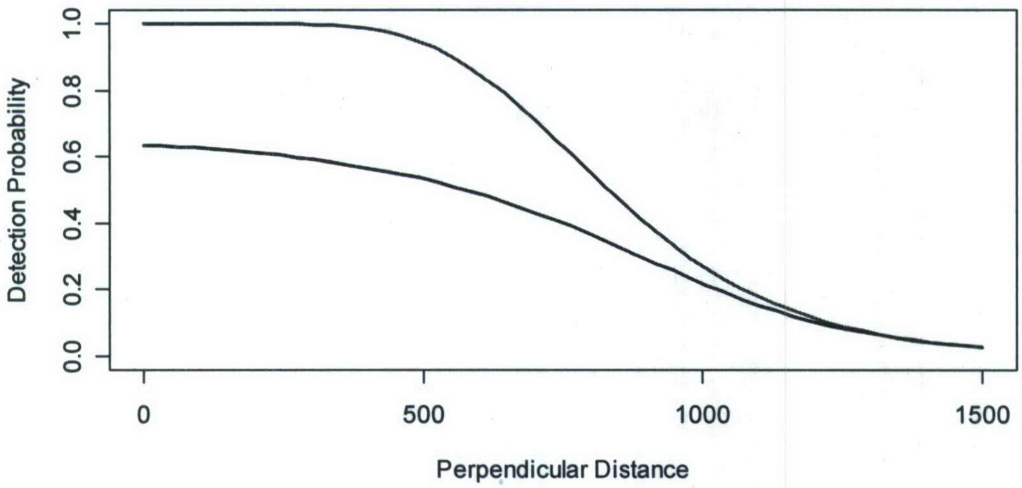

***Figure 5**: Simulated probability of detecting animals which are available for detection exactly 20% of the time they are within detectable range. Curves were generated using dentical observers but with frequent short periods of availability in the case of the red curve and infrequent periods of longer availability in the case of the blue curve.*

Projects 2 and 3 – Spatial density surface estimation and Spatial patterning in group size

Classical methods of abundance estimation from line transect surveys have two components: (i) estimation of detection probabilities, and (ii) "upscaling" of the observed number of groups (or animals) to account for both the failure to detect all potential sightings within the surveyed part of the region, plus the extrapolation to the un-surveyed part of the region. The first component is well understood these days, through modern distance-sampling and multiple-platform models (e.g., Chapter 6 of Buckland *et al.* 2004). The second component is generally handled in a "design-based" analysis framework, in which the survey effort is assumed to have been placed at random through each stratum in the area of interest. This makes estimation very easy, but runs into problems (particularly in terms of quantifying the uncertainty of the estimate) if the actual placement of effort is unbalanced (often the case in practice, even when the intentions were good), if the individual strata do not contain enough observations, or if there are strong trends in density within each stratum.

The alternative to design-based analysis is model-based analysis, also known as spatial line transect analysis. A two-dimensional smooth surface is fitted to data on location of sightings (given the survey effort), to describe local group density; abundance within a region can then be estimated simply by integrating the surface over the region. Spatial line transect models offer a number of potential advantages over design-based estimation:

- ability to deal with uneven coverage;
- better precision from same amount of data;
- spatial patterns may be of interest in their own right.

However, these benefits only kick in if the statistical model is good enough. Demand for spatial line transect analysis has been very high in the marine mammal science community, and a number of application papers have been published that use “off-the-shelf” statistical models or custom modifications; indeed, two of the team working on this project have previously developed spatial line transect models. Our experience, though, has been that the existing approaches are inadequate. Project 2 aims to develop a fresh approach that can address the following five generic problems that have often been encountered with existing spatial line transect models:

1. *Inability to deal with irregular topography (indented coastlines and islands)*: conventional smoothers try to enforce that animal densities be very similar on the two sides of a narrow peninsula, even when the “sea distance” between the sides is enormous and the two are biologically unconnected.
2. *Extreme estimates with associated high variances*: conventional smoothers damp down changes in the fitted surface on moderate spatial scales, but on the largest scales no damping occurs and an exponential trend is fitted. These trends can grow wildly in corners that are distant from the survey trackline (and therefore not constrained by nearby data); hence some “smoother taming” is required.
3. *Inability to deal with small-scale clustering*: animals (or groups or animals) are often encountered in loose aggregations at spatial scales too small to model explicitly. However, most spatial line transect models assume that encounters will be independent (after allowing for large-scale patterns in density); datasets with small-scale clustering are intrinsically less precise than most models assume, but this is generally ignored when calculating the precision of an abundance estimate
4. *Estimating the right smoothness for the density surface*: degree of smoothness has a major bearing on the precision of an abundance estimate, because if density varies rapidly in space, then the density in un-surveyed areas cannot be well predicted from nearby surveyed areas. Smoothing parameter estimation has been a general weakness in statistical computation, and practical techniques have only begun to emerge over the last few years. Choice of method is also linked to the previous point, because the appropriate degree of smoothness depends on the information content of the data, which depends on the amount of fine-scale clustering on scales too small for the density surface to model.
5. *Allowing for the uncertainty of estimated detection probabilities*: abundance estimates depend not just on the number of animals (or groups) seen, but also on the estimate of detection probability, which will usually vary spatially. Detection probability is necessary to convert “numbers actually seen” into “numbers likely present”. Both quantities are uncertain, but off-the-shelf spatial models give no easy way to propagate the uncertainty in detection probability through into the precision of the final abundance estimate.

All the above relates strictly to the animal *group* as the unit of detection, which may be one animal or many. If group size is nearly always 1, then the above is sufficient. In many surveys, though, group size can vary, and this brings in a sixth issue: spatial variation in mean group size. Conceptually, the treatment is simple: to estimate total abundance, first estimate mean group size as a spatially-varying function; then estimate the density of groups as a spatially-varying function; then multiply the two together at each point in a fine spatial grid across the region of interest, to obtain *animal* density as a spatially-varying function; then integrate this across the region to get animal abundance. However, this glosses over the fact that larger groups tend to be more detectable, as well as having more animals in them. Consequently, the spatial variation in group size affects the spatial pattern in detection probabilities, which is a fundamental input to the spatial density model. The group size model therefore enters *twice*: first to affect the density model, and then as a multiplier to convert group density into animal density.

Our original intention in Project 3 was to tackle spatial variations in group size by a rather different approach (Thomas 2003). However, after some experimentation, and as reported in a previous progress report (Thomas 2006), it became clear that it would be better to use similar spatial smoothers for both the group size and group density parts of the model (although the statistical details obviously vary), and to develop a unified and linked treatment for abundance estimation, particularly so that uncertainty can be propagated correctly through all parts of the model. This Section therefore covers our progress on both Projects 2 and 3.

Antarctic minke whales: a hard test case

As a test case, we have focused our model development towards one particular dataset: the "circumpolar 2 and 3" phases of the SOWER/IDCR[1] surveys for minke whales around Antarctica, from 1986 to 2003 (Figure 6). This is one of the largest (perhaps even the largest) marine mammal sighting datasets in the world, and has benefited from stable protocols, experienced observers, and good quality control over the years. It is therefore a good dataset, in that it is at least amenable to statistical analysis; and its analysis is proving controversial in the International Whaling Commission, because apparent large changes in abundance may be at least partly explained by deficiencies in standard design-based analyses. However, it also exhibits a pretty comprehensive set of the difficulties we feel are generic in marine mammal sighting surveys:

- spatial coverage is uneven, and often quite different from the nominal *random stratified design* (hence the need for spatial modeling in general);
- coverage is often poor in large chunks of ocean out towards 60 degrees (issue #2, above);
- sightings are clustered on a fine scale (issue #3);
- sighting conditions, group sizes, and group density all tend to be higher close to the ice edge. The correlation in their spatial patterning can be shown to lead to bias if the three aspects are modeled independently with non-spatial approaches. (issues #4 and #6);
- detection probability for sightings on the trackline is not 1.0, which amplifies the effect of uneven coverage, density, group size, and sighting conditions;
- the topography of the coastline is quite irregular in places (issue #2; Figure 7);
- it is a multi-year survey (19 years), with different spatial patterns in every year but the same smoothing parameters and distance-sampling parameters in each year (or at least in blocks of years).

[1] Southern Ocean Whale Ecosystem Research / International Data Collection and Recording. Hereafter just "SOWER"

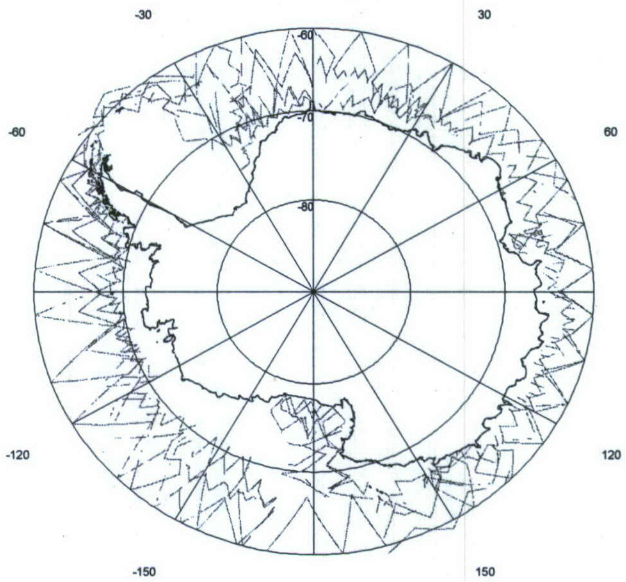

Figure 6*. Survey effort in 19 years of SOWER CP2 and CP3.*

Some of these points are evident in Figure 7, which shows the Ross Sea survey in 1986. The irregular topography, uneven coverage even within strata, fine-scale clustering of sightings, large-scale variations in density are all obvious; and a keen eye will discern that the small groups (black dots) tend to be found in the areas of lowest density. While some of these phenomena are merely natural consequences of the sighting process, the general point here is that these data are unlikely to be well-served by a design-based analysis, or by a naïve spatial model.

An additional complication with SOWER is that group size estimates are subject to considerable errors, at least in part of the survey (see Appendix 5 (Bravington *et al.* 2007) for more details). While group size error was not on our original list of target problems, it is probably quite a generic problem whenever spatial variations in group size are also significant. The best way to handle group size error in a spatial line transect analysis – if there is a best way – will depend on the details of the particular survey; SOWER at least has some parts where group size is reliable, plus some experimental data, and these can be used to develop a defensible statistical model. In deciding how to take forward this abundance estimation project in future, a key task will be to work out how far it is possible disentangle the generic from the specific parts of the SOWER analysis, and therefore what level of generality to aim for in developing software tools (see Conclusions and Recommendations, below).

In addition to the generic problems above, SOWER also has a large number of specific quirks to do with the way the survey is conducted. These have a major bearing on estimation of detection probabilities, which of course underpin the whole enterprise of abundance estimation. There is no getting around these issues when analyzing SOWER data, and a great deal of modeling effort (in other projects besides this one) has been devoted to them. A large part of the Appendix is addressed towards these issues; although they may not be of general interest because of their specificity to SOWER, something similar will inevitably be necessary in any thorough analysis of spatial line transect data.

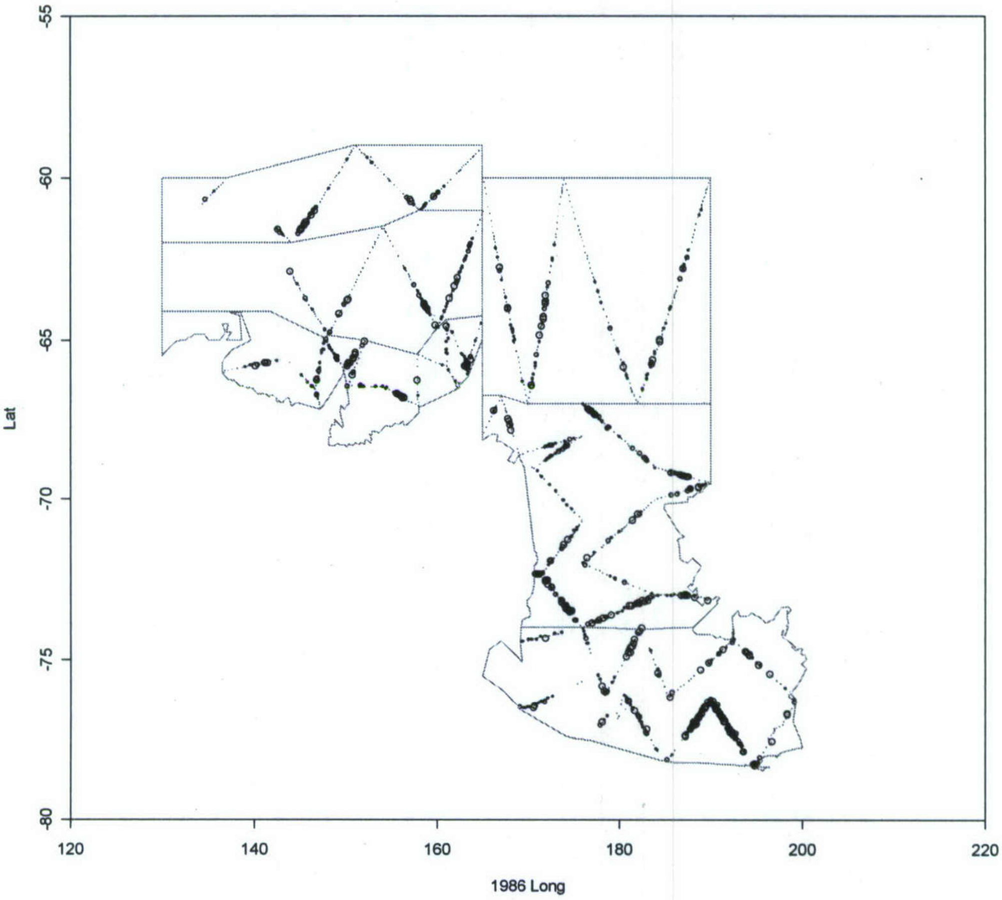

***Figure** 7. SOWER survey of the Ross Sea in 1986: stratum and ice-edge boundaries (green), effort (red---only the midpoints of legs are shown), and sightings (black circles, proportional to school size).*

One further virtue of SOWER as a testbed, is that the International Whaling Commission has developed an extensive collection of simulated datasets, which mimic some or all (depending on which simulation) of the realistic complexity behind SOWER (IWC 2007). A model that is capable of analyzing SOWER itself can obviously be applied to the simulated data with very little modification, and it will be possible to gain good insight into the model's performance by comparing its estimates with the known truth behind the simulations.

Results - Model structure

The technical details of how the various parts of the model are handled are given in Appendix 5, but it is helpful here to give an overall indication of model structure. There are three cases to consider: spatial density models where group size is not an issue; spatial models for density and for group size, but without errors in group size measurement; and spatial models for density and for group size, together with errors in group size. The latter, of course, is the situation for SOWER.

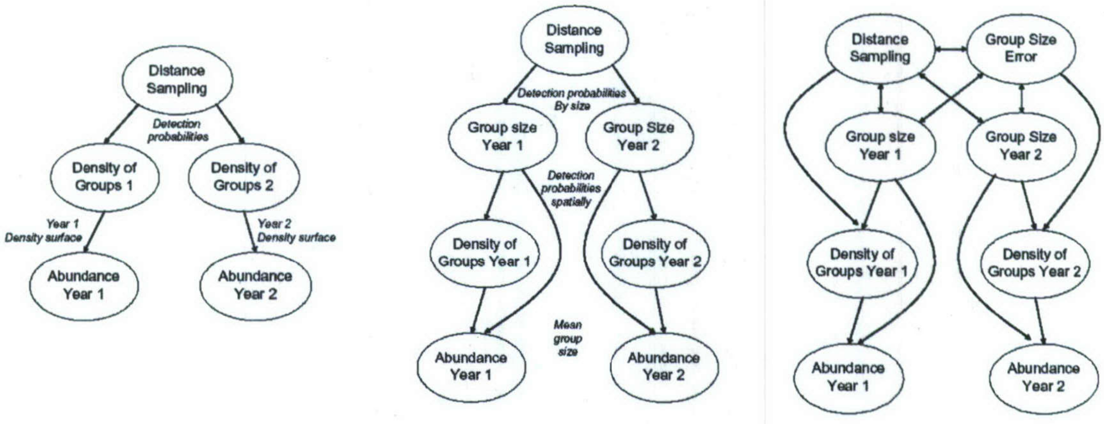

***Figure 8**. Model structure for datasets of varying complexity: constant group size, spatially variable group size, spatially variable group size with group size error.*

Case 1 is fairly simple: first analyze the individual sightings using the appropriate distance sampling methods, then feed the detection probabilities into the spatial density models. Case 2 has an extra phase in the middle, of estimating the spatial pattern in mean group size. Because the detection probability of a group depends on the group size, and because some parts of the survey area will tend to have smaller groups than other parts, the average detection probability of a group will depend on where that group is. However, although the analysis is more complicated than Case 1, the structure is pretty clear and there are just three sequential phases instead of two.

Case 3 is more difficult. Obviously, if there is group size error then there must also be an additional model for group size error, but there are structural consequences for model-fitting as well. Group size typically has such an important effect on detection probability that the group size of each sighting must be allowed for in the distance sampling phase. However, the recorded group size may be incorrect, so it's necessary to allow for what the true group size of each sighting might have been. But this in turn depends on *where* the group was seen. If a group of recorded size 1 is seen in a region where all the groups seem to be of size 1, then it probably is a true size-1 group; but if it is the only size-1 in a region containing mostly size-2 groups, then it is probably a mis-recorded size-2. Consequently, it is no longer possible to have separate phases for distance sampling and spatial-group-size model fitting; it is necessary to fit the spatial group size model at the same time as the group size error model *and* the distance sampling models (the spatial density model can then be fitted in a subsequent phase). This entails a substantial increase in computational complexity, beyond what we anticipated when the project was set up. However, we have successfully managed to code and partly test this more complicated framework (see below). The propagation of uncertainty through these linked models has been provisionally coded, but not tested at all.

To the best of our knowledge, there are no other line-transect models that attempt to take account of group size variation and error in such a systematic fashion. The complexity of the models, and in particular the requirements for spatial smooths and for propagation of uncertainty, have required us to make heavy use of Automatic Differentiation (AD) software, for reliability and computational speed; uncertainty propagation also relies on AD. Although there are a number of AD tools in the public domain, most have been designed with specific purposes in mind, and remarkably few are amenable to the kind of statistical problems we have faced. A considerable part of the development time in this project has been spent coercing one particular AD tool (Tapenade: Hascoet and Pascual 2004) to work for us--- our application has very different requirements to Tapenade's more usual purpose of aircraft design. On the whole, we have had excellent results from AD, and as our experience of it has progressed, it has become easier to experiment with changes in the underlying statistical models.

Results - Improved smoothers

As reported in the 2nd progress report (Thomas et al. 2006), we have developed an innovative soap-film smoother specifically to deal with the issue of complex topography. As a valuable by-product it also deals automatically with the issue of smoother taming. Further details are given in Appendix 6 (Wood *et al.* in press). Briefly, though, the soap-film smoother models log density as the height-above-or-below-the-page of a soap film stretched across a springy wire frame whose edges lie on the boundaries of interest. The wire frame can be pulled up and down (but not sideways) towards the data, and the film itself can also be pulled away from its natural configuration towards the data; the smoothing parameters determine how much pulling goes on. From a computational and statistical perspective, this is elegant and feasible to implement since it can be estimated by generalized cross-validation or by maximizing an approximate likelihood.

It is apparent that the soap film can easily provide very different densities on different sides of a peninsula (imagine bending the wire frame at the end of the peninsula), and also that there will be some intrinsic smoother taming (because the wire frame forms a loop, and any large-scale trend therefore requires the wire frame to be pulled up and then down again, which will only happen if there is sufficient *local* signal in the data). Various tests on both real and simulated data (Appendix 6) demonstrate this; another example is shown in Figure 9. Comparisons are also made in Appendix 3 between the new soap film smoother and an alternative method of smoothing within complex regions based on finite element modeling (FELSPLINE, Ramsey 2002). The soap film methods are shown to be superior (see, e.g., Figure 6 of Appendix 6). Lastly, simulations reported in Appendix 6 suggest that soap film smoothers even perform very well in open-boundary applications where there is no real “coastline”.

The IWC's simulated-SOWER datasets will provide extensive testing opportunities in a line-transect context, but the multi-stage nature of SOWER data (first group size, then density) means that this can't be done until all the models are fully developed and linked together; this is planned to happen during April 2008.

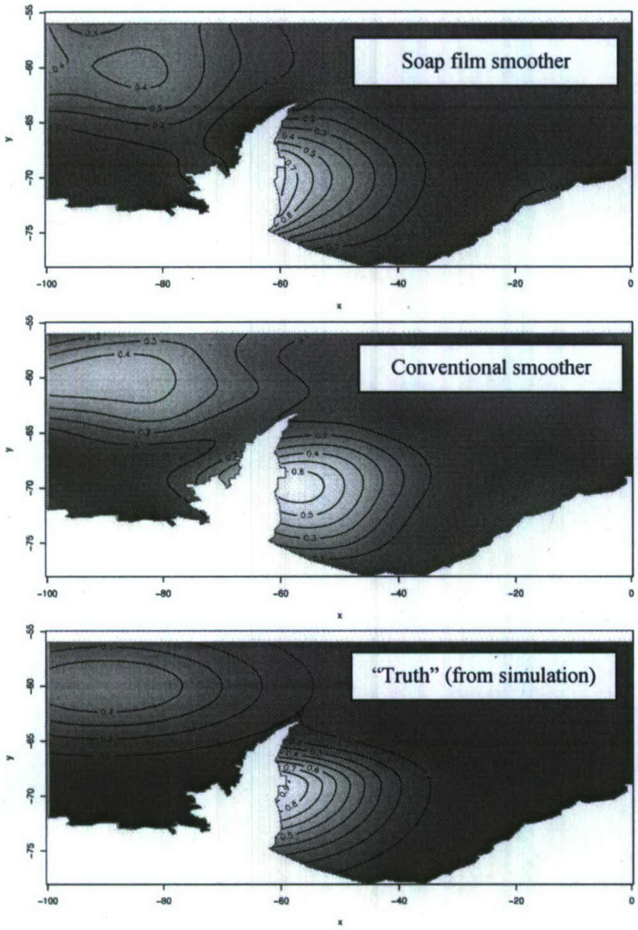

***Figure 9**. Example of soap film smoother fit to simulated survey data from the Antarctic peninsular. Bottom panel shows true underlying density from which data were simulated. Middle panel shows results from conventional smoother applied to these data. Note "leakage" of estimated high density across the peninsula boundary. Top panel shows result from soap smoother, showing no leakage.*

Results - Spatial patterning in group size

The soap film smoother again provides the springboard here. In effect, it can be used to set the local mean group size in a polytomous regression; this is almost an off-the-shelf statistical model, if soap film smoothers are used in place of other smoothers. However, because of the requirement to simultaneously estimate detection probability models and group size error models along with spatial group size, the actual model is much more complicated. Some preliminary results (based on a very approximate distance sampling analysis) are shown in Figure 10. In the right-hand panel; the higher mean group size in the southern portion and along the ice edge is apparent, though there is also an unexplained bump in the top left which needs further investigation. The effect of group size on detection probability (left-hand panel) is less obvious; most of the variation in width is due to variations in sighting conditions, rather than group size.

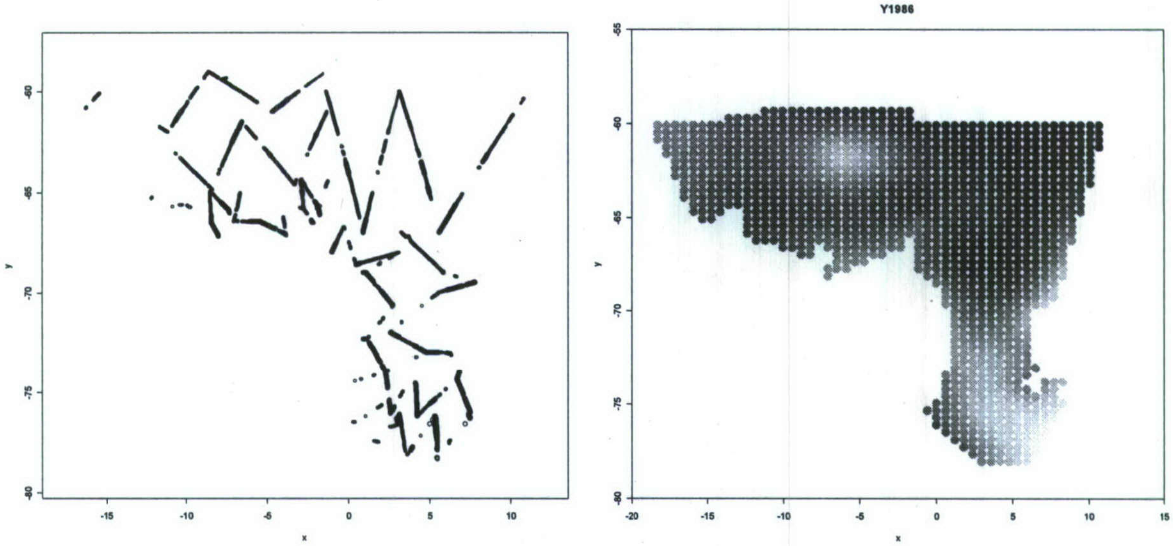

__Figure 10__. Effective strip width (i.e. detection probability) and mean group size (yellow=larger) in the Ross Sea survey, 1986. Only approximate distance sampling parameters and smoothing parameters were used here.

Results - Fine-scale clustering

When we began this project, there were no good computationally-feasible methods for handling fine-scale clustering in count data (e.g. number of groups encountered). We therefore put a lot of effort into developing new statistical approaches to clustered counts, and eventually found a method that promises to work well (Appendix 7: Bravington 2006). Meanwhile, though, an alternative approach specific to line transect surveys was published by Skaug (2006). Although Skaug's approach has less statistical flexibility than ours, it is much simpler to program, and should do a good job of capturing the main effects of clustering, at least in the fairly sparse datasets typical of line transect surveys. We have therefore implemented Skaug's method rather than our own, for the time being at least. We do intend to pursue our general clustered count model separately, since clustering is a very important issue for count data in general and – whereas Skaug's approach is fundamentally restricted to one-dimensional data such as is collected on a line transect – our approach works fine in two or more dimensions. The latter aspect may make our approach relevant to certain types of cetacean survey where there is very dense coverage in space or time, for example in real-time monitoring.

Project 4 – Improved design-based survey design and estimation

We developed an unequal coverage estimator fashioned after a Horvitz-Thompson-like estimator of abundance. This estimator was postulated by Strindberg et al. (2004) as

$$\hat{N}_{\text{covered}} = \sum_{i=1}^{n} \frac{s_i}{\hat{P}_i} = w \sum_{i=1}^{n} \frac{s_i}{\hat{\mu}_i \hat{P}_i}$$

where s_i is sized of the i^{th} detected cluster, n is the number of detected clusters, $\hat{P}_i$ is the estimated coverage probability at the location of the i^{th} detected cluster, and $\hat{\mu}_i$ is the estimated effective strip half-width of the i^{th} detected cluster. To convert $\hat{N}_{covered}$ into an abundance for the entire study region, simply scale up this estimate by the proportion of the study area sampled. Methods for determining the variance of this estimator of abundance are described in papers such as Stevens and Olsen (2003)

To assess the performance of the unequal coverage probability estimator when applied in a distance sampling context, we performed a simulation study in which a known number of objects were placed across a study area with varying degrees of spatial aggregation, as described in Appendix 8 (Rexstad 2007). Populations of animals were created using the WiSP package of routines developed by Borchers et al. (2002). This package can produce populations of any size distributed across unit square in a variety of manners. Distributions employed during this study were gradient (oriented east-west), hump (with a bivariate normal irregularity situated somewhere in the survey region), or trough (where animal density varied east-west but in a non-monotonic manner).

Distance sampling methods are also incorporated into the WiSP package, and take into account imperfect detectability along transects placed within the rectangular survey region. One limitation of the distance sampling implementation of WiSP is that the transects can only be oriented north-south in the survey region.

The simulations proceeded by generating a population of animals according to a specified spatial distribution. For a specified level of effort, i.e., number of transects to sample, a sample of transects is selected without replacement from the population of possible transects (this transect population was defined by specifying a strip half-width such that each transect exactly "touches" the transect on either side of it). From this sample of the survey region, we fit a density surface model (Hedley and Buckland 2004). This model predicts abundance of animals in each segment of the survey region. These segments are summed or integrated in the north-south direction producing an estimate of the number of animals in each candidate transect throughout the survey region. This step reduces the estimated distribution of animals in the survey region from two dimensions to a single dimension.

The number of animals in each candidate strip (known or estimated) constitute the weights to apply in a probability proportional to size (pps) (Cochran 1977:250). Using a pps sampling scheme without replacement (Tillé 1996), a sample of transects are selected from the population of transects. Animals detected in these transects constitute the data available for estimation of abundance in the survey region.

Using simulation, we repeated this procedure many times, and looked at the improvement in efficiency that was gained relative to a standard systematic sample. We found only very minor gains – in the region of 15% or less (see Appendix 8). Possible explanations for the minor performance enhancements achieved by the unequal coverage probability estimator are as follows. Both the detection probability and inclusion probability are found in the denominator of the unequal coverage estimator. In this simulation study, detection probability was constant across individuals (in fact it was the same across all simulations reported here). However, inclusion probabilities were specific to individuals included in the sample. These values could be as small as 0.01, and when multiplied by the detection probability, each individual detected with that inclusion probability represent 125 members of the population.

The relevance of this behaviour of the HTL estimator is that the rare event of encountering an animal on a transect with a low inclusion probability will result in an estimator with a thick right tail. This behaviour is likely to negate the increase in precision derived from the additional information about 'profitable' locations to sample coming courtesy of the pilot survey. The influence of that tail could be decreased if some cutoff for inclusion probability was instituted, as recommended by Borchers et al. (2002:144) for small detection probabilities used in a Horvitz-Thompson estimator of abundance.

In addition to a simulation study, we also applied our unequal coverage probability estimator to a data set collected in inshore waters of British Columbia (see Thomas *et al.* (2007) for a description of the survey design, and Williams and Thomas (2007) for a description of the data gathered and analyzed). Because of the complex nature of the coastline in the study area, line transect survey designs are quite unequal in their coverage probabilities. These coverage probabilities were computed by simulation using the survey design engine in Distance 5.0 (Thomas et al. 2006).

We produced estimates of abundance in the covered region of the study area for three species, harbour porpoise, Dall's porpoise, and harbour seals. These estimates can be compared to the estimates of abundance in the covered region produced with an estimator that assumes equal coverage probability throughout the study area.

Table1*. Comparison of abundance estimates using standard equal coverage estimator and new unequal coverage estimator for three species surveyed in inshore British Columbia.*

Species	Number of detections within truncation distance	Abundance estimate in covered region using equal coverage probability estimator	Abundance estimate in covered region using unequal coverage probability estimator
Harbour porpoise	26	1664.8	1568.6
Dalls porpoise	10	182.4	162.0
Harbour seals (in water)	111	2708.8	2727.8

The location of sightings for the above species in one strata of the survey described in Williams and Thomas (2007) is shown in Figure 11. The figure also demonstrates the nature of the unevenness of the coverage probability in the surveyed region.

To make the unequal coverage probability estimator more widely available, we incorporated this estimator into a new beta version of the software Distance (Appendix 9: Thomas et al. 2008). As an example case of the use of this estimator, we have included the harbour porpoise data of Williams et al. (2007) as a sample dataset.

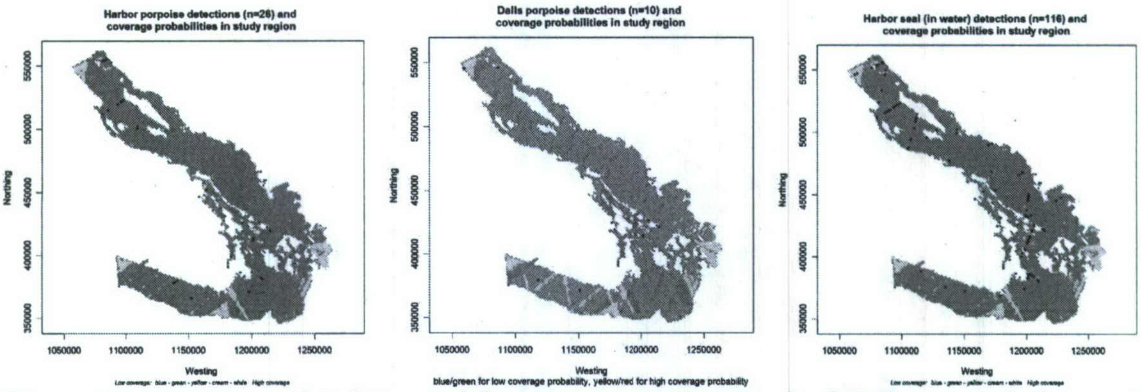

Figure 11. Locations of sightings of three species in stratum 2 of Williams and Thomas (2007), indicated by dots. Color indicates survey coverage probability: yellow is high, green intermediate and blue low. The design clearly exhibits some un-evenness in coverage.

List of Project Outputs

Peer-reviewed journal articles

In Press:

- Wood, S.N., M.V. Bravington and S.L. Hedley. In press. Soap film smoothing. Journal of the Royal Statistical Society, Series B.

In Prep:

- Borchers, D.L., Heide-Jorgensson, M. and Samarra, F.I.P. Accommodating availability bias on line transect surveys using hidden Markov models.
- Bravington, M.V., S.L. Hedley and S.N. Wood. Saddlepoint approximations for Poisson and binomial data with overdispersion induced by spatial random fields, with application to clustered line transect data.
- Bravington, M.V., S.L. Hedley, S.N. Wood and D. Peel. Estimating Antarctic minke whale abundance from SOWER data, using spatial modelling and point independence.

Technical reports

- Borchers, D.L. and Burt, M.L. 2007. Investigation of towed hydrophone monitoring power for harbour porpoise on the SCANS II survey. Technical Report 2007-5, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.
- Borchers, D.L. and F.I.P. Samarra. 2007. Accommodating availability bias on line transect surveys using hidden Markov models. Technical Report 2007-6, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.
- Bravington, M.V., S.L. Hedley, S.N. Wood and D. Peel. 2007. Estimating Antarctic minke whale abundance from SOWER data, using spatial modelling and point independence. International Whaling Commission intersessional workshop on estimating Antarctic minke whale abundance, Feb 2008. SC/F08/A9.
- Bravington, M.V., D. Peel and S.L. Hedley. 2006. More abundance estimates for Antarctic minke whales, using different methods. Scientific Committee of the International Whaling Commission paper IWC/SC58/IA15
- Brewer, C. Borchers, D.L., J. Matthews and C. Brewer. 2007. Methods for estimating sperm whale abundance from passive acoustic line transect surveys. Technical Report 2007-3, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.
- Rexstad, E. 2007. Non-uniform coverage estimators for distance sampling. Technical Report 2007-1, Centre for Research into Ecological and Environmental Modelling, St. Andrews University. http://eprints.st-andrews.ac.uk/archive/00000445/

Conference presentations

- Bravington, M.V. 2007. Estimating whale abundance by counting them: a hard case study in the Antarctic to the Tasmanian branch of the Victorian Statistical Society, Hobart, June 2007.

- Thomas, L. 2005. Monitoring Marine Mammals Using Passive Detector Arrays: A Statistical Perspective. Presented at workshop on Monitoring Marine Mammals on Navy Testing Ranges, Arlington, Virginia, Sept. 7th-8th 2005.
- Thomas, L. 2005. Development of New Methods and Software for Distance Sampling Surveys of Cetacean Populations. ECOUS, Washington DC., 16-18 March 2005.
- Thomas, L. (presented by J. Harwood) 2005. Where are they? Predicting the location of marine mammals. Intergovernmental Conference on The Effects of Sound in the Ocean on Marine Mammals, La Spezia, Italy, May 2nd – 5th 2005.
- Thomas, L. 2007. New methods for estimating marine mammal distribution, density and abundance. 2nd Intergovernmental Conference on the Effects of Sound in the Ocean on Marine Mammals, Lerici, Italy, June 2007.

Software

- 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., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2006. Distance 5.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK.
- 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., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2008. Distance 6.0. Beta 5. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK.

Patents submitted/issued

No patents have been submitted as a result of this research. All ideas generated have been placed in the public domain.

Technology transfer

- Distance software (all versions) has been downloaded by 13,000 people from more than 120 countries. It is the industry standard for design and analysis of distance sampling surveys of cetaceans (and all other taxa). Not all research findings yet implemented in Distance software, but that is our goal.
- Research findings communicated to policymakers at workshops and conferences listed above.

Conclusions and Recommendations

Our primary goal in undertaking this research was to develop methods that allow more efficient and robust estimates of cetacean density and distribution to be obtained. We have pursued four major areas of research, and in each we have made substantial progress. Here we give our conclusions, and make recommendations for further research.

Project 1 – Acoustic and visual-acoustic methods

Acoustic survey methods have great potential to improve wildlife monitoring and assessment. While they suffer from some shortcomings compared to visual surveys, they enjoy some substantial advantages over those methods. However, further progress is required both in analysis method development and in hardware development. Future hardware developments which allow forward-looking hydrophones to be deployed on vessels allow better localization of detections will greatly improve the prospects for developing successful combined visual-acoustic MRLT methods. This project delivered theory for an integrated statistical analysis of sperm whale acoustic data but application of the theory floundered with the available data. Hardware which gives better localization of detections will help overcome these difficulties and the existing theory provides a foundation for future work. Towed hydrophones have significant advantages over fixed hydrophones (which are the subject of substantial current research), because they can use existing distance methods. They also provide better temporal coverage (although worse spatial coverage) per unit cost. There is great potential to put acoustic detectors on platforms of opportunity for improved estimation of animal distribution (with the subject of project 2) and of temporal trend. With regard to the latter, the development in this project of a method for calculating power to detect trend taking account of the randomness which use of platforms of opportunity will introduce in estimation bias (because surveyors can't control the characteristics of such platforms) will be useful in evaluating the likely utility of such deployment in future.

While progress in developing statistical analysis methods for towed acoustic surveys of sperm whales was disappointing, it has led to the development of improved and generally-applicable methods for dealing with "availability bias" arising from intermittent availability of animals on line transect surveys. methods.

Projects 2 and 3 – Spatial density surface estimation and Spatial patterning in group size

A major success in these projects was the development of the soap film smoother, which has applications far outside of spatial modeling of line transect survey data. We have also developed a very comprehensive and robust framework for the model-based analysis of line transect data, one that deals with all of the challenges we laid out for ourselves and more besides (uncertainty in group size estimation). We have found it to be better to take the same approach to modeling spatial density surfaces of groups and spatial variation in group size, and hence projects 2 and 3 have converged methodologically, allowing for a more elegant and consistent solution.

In applying our methods to the complex SOWER dataset, a significant amount of programming, testing and debugging remains to be done, and this is continuing under funding from CSIRO to

Bravington. We will also test our approach on the hundreds of simulated datasets prepared by IWC. This work will be complete by May (when results need to be sent to IWC), and thereafter the immediate task will be to get papers into peer review.

Beyond that, the main issues for future work are:

- How to make the modeling framework reasonably generic? Application-specific details, such as platform setup and group size estimation error, will differ every time. These are hard-wired into the SOWER model, but there is no way anyone would ever be able to re-implement the entire model >20000 lines of code, in addition to all the soap-film code) for a new problem. However, the fundamental elements of estimating spatial distributions for group size and group density will not really change with different datasets. Some pause for thought is required to figure out where to go with general-purpose software: should it be targeted initially to a limited range of problems (e.g. it would be easy if we ignored group size, which is reasonable in some distance sampling settings); and how generic can the software really be?
- What other issues are likely to be important for many datasets? In particular, the question of how to deal with random effects in distance sampling models (for example, to allow for differences between many observers, with only limited data from any individual) is not handled with standard software, although it's well recognized as a general problem in statistics. However, the tools we have developed already have built-in capacity to deal with random effects, as a necessary part of fitting spatial models (which can be viewed as a special type of random effect). It is natural to consider whether some generic extensions to deal with non-spatial random effects can be developed.
- Further development of a true two-dimensional (or more) treatment of fine-scale-clustering (which we started in Appendix 4, but have put aside for the moment thanks to Skaug's work) may turn out to be necessary for dealing with sampling that is denser in space and time than SOWER is – for example, in real-time monitoring.
- The current methods are aimed at overall abundance estimation, or at least at abundance estimation over pre-specified areas. In some applications, though, it is more important to *locate* hotspots and to estimate just how hot they are, plus of course to quantify the uncertainty in the answers. The relationship between hotspots and other covariates (e.g. seabed features) may also be of interest. These questions are subtly different from abundance estimation, requiring (in particular) a different criterion for the estimation of smoothing parameters; while smoothers have good unbiasedness properties for aggregated quantities such as total abundance, they are known to be biased on local scales (these biases tending to cancel out on larger scales). Although the basic setup of two spatial models should not need changing, theoretical developments and simulation testing would be necessary to develop good hot-spotting tools. (Of course, the absence of good tools has not and will not stop people from trying; the danger is that incorrect conclusions will be made on the basis of inadequate models.) A very related issue here is the estimation of density over small spatial areas (such as Navy testing ranges or other operational areas) based on large-scale surveys.

In order to decide what is most important to do, the natural next step is to talk to custodians of large datasets and to other statisticians involved in the analysis of such data, to showcase our work and get feedback on what is still missing and would be of most practical value. A technical workshop of this nature has been organized by Thomas, and will be held in St Andrews in May 2008. We anticipate a comprehensive set of research priorities will emerge from that meeting.

Project 4 –Improved design-based survey design and estimation

Our simulation study suggests that there will be only modest benefit in terms of better estimator precision when implementing the new designs in most cases. Nevertheless, they may be of use in some situations. One example is that of a "central place" sampling strategy – for example a helicopter making repeated forays from a moving mother-ship (as in surveys of Antarctic pack-ice seals, etc.). In this case, it is very inefficient to implement a design with even coverage, because this necessitates lots of dead time flying between parallel transects. A more efficient design is one based on a "cartwheel design" (Figure 12). Whether such a design produces better precision after the non-even-ness of coverage is accounted for at the estimation stage is an open question.

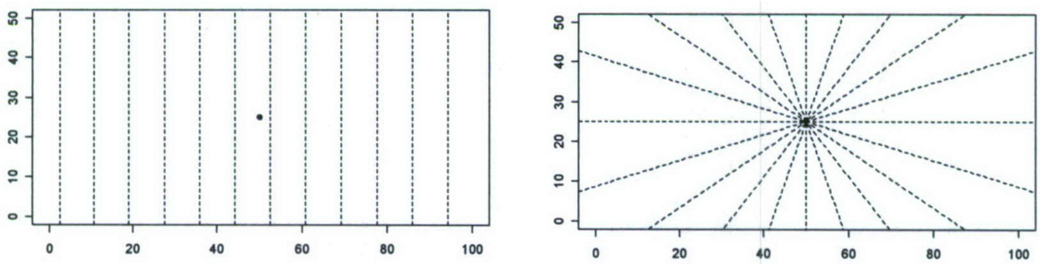

Figure 12. *Illustration of a standard, equal coverage design (left) and a "cartwheel" design (right) which has non-even coverage. The latter involves much less "off-effort" time flying between transect lines when the survey vehicle is constrained to return often to the same point (the central dot) to re-fuel.*

Implementation of the new estimation methods in Distance allows survey biologists to analyze datasets that were collected using non-even designs, such as those of Figure 11, and we anticipate that there will be increasing use of this facility in the future.

Our simulation looked at increase in precision when a model is used to guide the survey design, but then the estimation is done with design-based methods. A useful future research topic would be to investigate optimal survey designs to maximize prevision of model-based estimation methods such as those of projects 2 and 3.

Other recommendations

In the event, we did not include the developments from project 1 in the Distance software, due to lack of funds. The project was originally costed at an exchange rate of 1.65 US$ to the UK£, but in the event the exchange rate obtained was 1.85 on average. In this circumstance, we decided to prioritize methodological development over technology transfer. We plan to make all developments available in Distance in the future, and will seek funding to do so from various sources.

One component of the original proposal was to create a simulation engine in the Distance software that would allow simulation of animals based on a clustered point process model (e.g., Neyman-Scott), simulation of surveys from the survey design module and analysis of results. This would potentially save a great deal of effort and capital that is currently spent trying to optimize surveys in an ad-hoc way and often after the survey program has been underway for many years. We view the creation of such a facility as a high priority for future work.

In conclusion, this research project has generated a large number of extremely valuable, cutting-edge, new ideas and methods. Some have already been implemented in a way that makes them accessible to those charged with designing and implementing surveys, and analyzing the results. Others require more testing and development before they will be ready for general use. We expect the outputs from this research will be feeding into the user community for several years to come.

Literature cited

See also earlier section "List of Project Outputs".

Borchers, D.L., S.T. Buckland, and W. Zucchini. 2002. *Estimating animal abundance: closed populations*. Springer.

Borchers, D.L., J.L. Laake, C. Southwell and C.G.M. Paxton. 2008. Accommodating unmodelled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62**: 372-378.

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

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

Cochran, W.G. 1977. *Sampling theory*, 3rd edition. John Wiley and Sons.

Hascouet, L. and V. Pascual. Tapenade 2.1 Users Guide. INRA, Sopia-Antipolis, France. http://www-sop.inria.fr/tropics/tapenade.html

Hedley, S., and S. T. Buckland. 2004. Spatial models for line transect sampling. *Journal of Agricultural, Biological, and Environmental Statistics* 9:181-199.

International Whaling Commission. 2007. Report of the scientific committee. Annex G. Report of the subcommittee on in-depth assessment. SC59/Rep1

Link, W.A. 2003. Nonidentifiability of population size from capture-recapture data with heterogeneous detection probabilities. *Biometrics* **59**: 1123-1130.

Ramsay, T. (2002) Spline smoothing over difficult regions. *Journal of the Royal Statistical Society, Series B* 64: 307-319.

Skaug, H.J. 2006. Markov modulated Poisson processes for clustered line transect data. *Environmental and Ecological Statistics* 13: 199-211.

Stevens, D.L. and A. R. Olsen. 2003. *Variance estimation for spatially balanced samples of environmental resources*. Environmetrics 14:593-610.

Strindberg, S., S.T. Buckland, and L. Thomas. 2004. Design of distance sampling surveys and geographic information systems. Pages 190-228 in S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, and L. Thomas (eds.) *Advanced Distance Sampling: Estimating abundance of biological populations*. Oxford University Press.

Thomas, L. 2003. *Development of new methods and software for distance sampling surveys of cetacean populations*. Research proposal, submitted to the Office of Naval Research on 21 October 2003.

Thomas, L., D.L. Borchers, E. Rexstad, S.T. Buckland, M.L. Bravington, S.N. Wood, and S.L. Hedley. 2006. *Development of New Methods and Software for Distance Sampling Surveys of Cetacean Populations*. 2nd year progress report, submitted to the Office of Naval Research 25 July 2006.

Thomas, L., R. Williams and D. Sandilands. 2007. Designing line transect surveys for complex survey regions. *Journal of Cetacean Research and Management* 9:1-13.

Tillé, Y. 1996. An elimination procedure of unequal probability sampling without replacement, *Biometrika* 83:238-241.

Williams, R. and L. Thomas. 2007. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada. *Journal of Cetacean Research and Management* 9: 15-28

Zucchini, W., Raubenheimer, D. and MacDonald, I.L. 2007. Modeling time series of animal behaviour by means of latent-state model with feedback. *Biometrics* (online early doi: doi: 10.1111/j.1541-0420.2007.00939.x).

List of Appendices

Project outputs that have not been submitted in previous progress reports are included as appendices to this report. All are available for download from the following URL: http://www.creem.st-and.ac.uk/len/onr/

Appendix 1. Technical report: Brewer, C. Borchers, D.L., J. Matthews and C. Brewer. 2007. Methods for estimating sperm whale abundance from passive acoustic line transect surveys. Technical Report 2007-3, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.

Appendix 2. Working document: Borchers, D.L. 2007. Exploratory analyses of harbour porpoise visual-acoustic duplicates on SCANS II.

Appendix 3. Technical report: Borchers, D.L. and Burt, M.L. 2007. Investigation of towed hydrophone monitoring power for harbour porpoise on the SCANS II survey. Technical Report 2007-4, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.

Appendix 4. Technical report: Borchers, D.L. and F.I.P. Samarra. 2007. Accommodating availability bias on line transect surveys using hidden Markov models. Technical Report 2007-5, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.

Appendix 5. Technical report: Bravington, M.V., S.L. Hedley, S.N. Wood and D. Peel. 2007. Estimating Antarctic minke whale abundance from SOWER data, using spatial modelling and point independence. [Presented at IWC Intersessional Meeting on analysis of SOWER data, Seattle Feb 2008]

Appendix 6. Paper in press: Wood, S.N., M.V. Bravington and S.L. Hedley. In press. Soap film smoothing. Journal of the Royal Statistical Society, Series B.

Appendix 7. Working document: Bravington, M.V. 2006. Saddlepoint approximations for Poisson & Binomial data with overdispersion induced by spatial random fields.

Appendix 8. Technical Report: Rexstad, Eric 2007. Non-uniform coverage estimators for distance sampling. Technical Report 2007-1, Centre for Research into Ecological and Environmental Modelling, St. Andrews University.

Appendix 9. Software (Windows operating system): Thomas, L., Laake, J.L., Rexstad, E., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Burt, M.L., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. and Marques, T.A. 2008. Distance 6.0. Beta 5. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK.