Designing line transect surveys for complex survey regions

LEN THOMAS[*], DOUG SANDILANDS[†] AND ROB WILLIAMS[‡]

Research Unit for Wildlife Population Assessment, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Fife KY16 9LZ, Scotland.

Contact e-mail: len@mcs.st-and.ac.uk

ABSTRACT

Line transect surveys are widely used to estimate the density and/or size of cetacean populations. Good survey design is essential for obtaining reliable results using standard (design based) analysis methods. Even for more complex (model based) analysis methods, a good survey design is very helpful. By 'good' we mean a design (a) that employs randomization in laying out transects; (b) that is stratified if density is known to vary on a large scale; (c) where each location within a stratum has an equal probability of being surveyed (equal coverage probability); (d) that produces at least 10-20 transects per stratum; (e) that, given the previous points, gives maximum efficiency per unit effort – for example by minimizing time spent travelling between survey lines (off-effort time). We discuss strategies for creating good designs given the constraints inherent in many shipboard surveys of cetaceans: severely limited ship time and complex topography. We advocate the use of computer software, such as the program Distance, to create designs and compare their properties using simulation. We provide a link between the concepts and their implementation through a concrete example of survey design: a multi-species survey of cetaceans in coastal British Columbia. The design uses an equally spaced zig-zag configuration of transects in more open strata combined with sub-stratification to minimize off-effort time. In the highly convex inshore stratum we develop a systematic cluster sampling algorithm, and within the selected clusters use a systematic parallel line layout to ensure equal coverage probability in the long, narrow fjords. To aid those wishing to learn automated design methods, we provide Distance project files in an online appendix.

KEYWORDS: ABUNDANCE ESTIMATE, PACIFIC OCEAN, NORTH AMERICA, SAMPLING STRATEGY, SURVEY – VESSEL

SUGGESTED RUNNING HEAD: LINE TRANSECT SURVEY DESIGN

INTRODUCTION

Line transect surveys are widely used to estimate the density and/or size of wild animal populations. The methods are described in detail in two books by Buckland *et al.* (2001; 2004). Obtaining reliable results requires good survey design, field methods and data analysis. In this paper, we focus on strategies for creating good survey designs in the context of shipboard surveys of cetaceans.

In our context, a survey design is an algorithm for placing transects within the study area. Standard analysis methods, as described by Buckland *et al.* (2001), assume that the density of animals in the area surveyed (i.e. on the transects) is on average equal to the density in the entire study area. This will be true if the transects are placed at random using a design where each part of the study area has an equal probability of being surveyed (uniform 'coverage probability'). In this case, we need not make any assumptions about the spatial distribution of the animals. This kind of method, where we use the properties of the design to make inferences about the population is called a *design-based* method. Such methods are attractive because the survey design is something that is usually under our control and known (unlike animal distribution), and if we use an appropriate design coupled with a design-based analysis method we are sure of obtaining an unbiased estimate.

By contrast, analysis methods exist where we make inferences about the density of animals in the whole study area from the survey data based on a model for the distribution of animals (e.g. Hedley *et al.*, 1999; Hedley and Buckland, 2004; Hedley *et al.*, 2004). Such *model-based* methods do not make any assumptions about the manner in which the transect lines were laid out, and so can be used in cases where there was no element of randomization in the design, such as when the lines were placed subjectively or when the survey uses a vessel that is traversing the study area for another purpose (a 'platform of opportunity'). Model-based methods also offer the potential for more precise estimates than their design-based counterparts. They can also be used to estimate density in subsets of the entire study region for which there is limited survey effort. Their major disadvantage is that they can be badly biased if the model for animal density is poor, and creating a good model is not straightforward. An accessible introduction to issues related to design-based *vs.* model-based methods is given by Borchers *et al.* (2002), and a more technical reference is Thompson (2002). A clear description of the role of design and model in standard line transect methods is given by Fewster and Buckland (2004, Section

[*] Research Unit for Wildlife Population Assessment, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, Fife KY16 9LZ, Scotland.
[†] Vancouver Aquarium, Box 3232, Vancouver BC V6B 3X8, Canada.
[‡] Raincoast Conservation Society, Pearse Island, Box 193, Alert Bay, BC V0N 1A0, Canada and Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife KY16 8LB, Scotland.

10.3). The development of appropriate model-based methods for cetacean line transect data is an active area of current research (Hedley *et al.*, 2004).

Because model-based methods are not guaranteed to be unbiased, it is often desirable to be able to produce a design-based estimate to compare them with. In addition, a good survey design will tend to distribute transects evenly throughout the study area, which provides ideal input data for a model-based approach. For these reasons, survey design is important even where model-based estimates are the main goal.

Several constraints make it difficult to design shipboard surveys of cetaceans that are appropriate for analysis using standard design-based methods. The first is that the study area is often very large relative to the speed of the survey vessels, and ship time is relatively expensive. This leads to strong pressure to minimize the amount of time spent travelling between transect lines ('off effort') – for example by using a zig-zag transect configuration that may not have uniform coverage probability throughout the study area (Strindberg and Buckland, 2004). The second is that the study areas often have complex topography, containing features such as islands and inlets. Some survey designs have uniform coverage probability in rectangular or convex study areas, but not in non-convex areas. Many designs have lower coverage probability close to the edge of the study area, and this can become an important problem in areas with complex topography, where there is a high edge to interior ratio. Other issues such as stratification further complicate matters.

The recent development of automated survey design algorithms (Strindberg, 2001; Strindberg *et al.*, 2004) and their implementation in the free software Distance (Thomas *et al.*, 2004) has considerably simplified the task of creating and comparing complex designs. Designs can be created on the computer, and many random realizations can be generated to assess properties such as average (or maximum) proportion of time off effort and even-ness of coverage. Once a design is chosen, a single random realization of this design can be generated, exported from the software used as the survey plan (e.g. by loading into a ship's onboard navigation system).

The aim of this paper is to promote the use of good survey design in line transect surveys of cetaceans. We begin by briefly reviewing the relevant concepts of survey design, and defining what we mean by a 'good' design. We discuss practical strategies for dealing with the complications caused by the constraints inherent in many shipboard surveys of cetaceans. We illustrate these ideas and solutions using the example of a multi-species survey of cetaceans in coastal British Columbia, Canada. The study area includes stretches of open water as well as an intricate system of inshore islands and fjords. We use stratification, and develop a systematic cluster sampling algorithm in the inshore stratum to increase survey efficiency. We also compare a zig-zag transect configuration with a more conventional parallel line configuration.

In designing the survey, we made extensive use of the Distance software, and we give the relevant Distance projects and related geographic information files in an online appendix. Our intention in providing this material is to give interested readers a 'jump-start' in how to implement automated survey design methods in complex regions using available software.

The survey we designed was carried out in 2004, and one stratum was re-surveyed in 2005. The results are presented in a companion paper (Williams and Thomas, this meeting).

SURVEY DESIGN – CONCEPTS AND STRATEGIES

Line transects are part of a larger group of methods called distance sampling. Standard distance sampling methods are described in detail by Buckland *et al.* (2001), with basic concepts of survey design discussed in Chapter 7. Automated survey design methods are described by Strindberg *et al.* (2004), which is Chapter 7 of the more advanced text by Buckland *et al.* (2004). Both chapters are essential reading to anyone contemplating the design of a real-world distance sampling survey.

A survey design is an algorithm for laying out samplers, transect lines in this case, within the study area. A 'good' design for a given study is one that maximises the chance of obtaining reliable results given constraints imposed by the study area, species and logistics. For the reasons given above, we would prefer to obtain these results initially using design-based analysis methods, so we focus on designs that can yield reliable design-based results.

Design requirements

Two essential requirements for a good design are randomization and replication (Buckland et al. 2001, Section 7.2.1).

Randomization means that the design algorithm should use some form of random probability sampling in laying out the transects within the study area. Hence each time we execute the algorithm we obtain a different random realization. Standard analysis methods assume that on average over many realizations, each point within the study area has the same probability of being sampled – i.e. uniform coverage probability. This assumption is used at two points in estimation. Firstly, because the lines are located at random with respect to the animals, the distribution of distances from the transect line to each detected animal can be used to estimate the change in detectability with distance, since there is no change in density of animals. This enables us to estimate average probability of detection and hence density of animals in the surveyed region. Secondly, because all areas are equally likely to be sampled, this density estimate can be applied the whole survey area, not just the surveyed region.

If a design algorithm is used that involves randomization but where coverage is not uniform (for example due to edge effects, see below), then design-based estimation can still proceed but the standard methods must be extended to avoid bias (Buckland et al., 2001, Section 6.7; Strindberg et al., 2004, Section 7.3).

Replication (i.e., placement of multiple lines) is required for assessment of the uncertainty in our design-based estimates. Increasing the number of replicate lines increases the reliability of variance estimates, other factors being equal (such as total line length and even-ness of coverage). Buckland et al. (2001, section 7.2.1) recommend 10-20 replicates as a minimum, but we would not consider our design to be 'good' with fewer than 15 samples. In general, for a fixed total line length, many short lines are preferable to few long lines, so for designs where a set of parallel lines cross the whole study region (see below), they should be oriented perpendicular to the longer axis of the study area.

Other issues contributing to good designs

While not required, several other factors can contribute heavily towards promoting reliable inferences, so should be considered as part of a good design. These are appropriate use of stratification, use of designs that produce even coverage within each realization, and minimization of off-effort time.

Stratification is a very useful strategy where there are large-scale gradients in animal density and these are predictable. In that case, dividing the study area into strata so as to maximize the between-stratum variation in density and minimize the within-stratum variation can lead to greatly reduced variance. One constraint is that we require an adequate sample of transects (again 10-20) per stratum for reliable variance estimation.

In addition, variance can be further reduced by dividing the total survey effort between strata such that average coverage probability for each stratum is roughly proportional to density (Buckland et al. 2001, Section 7.2.2.3). However, this can lead to biased estimates of overall density if detection function are fitted to data pooled across strata (Burnham et al. 1980, p.200-201). If there are known to be large variations in density or very loose clustering of animals, but the locations of high density areas are not known in advance, then adaptive sampling designs may be useful (Pollard and Buckland 2004). Both non-uniform allocation of effort between strata and adaptive sampling may be of less use in multi-species surveys where different species of interest are at high density in different areas.

We have previously mentioned the desirability of having uniform probability of coverage averaged across many realizations. It is also desirable to choose designs that produce even coverage across the study area (or stratum, if using stratification) within each realization. Designs such as this, where sampling is spread evenly through the study area, produce more reliable results in the sense that there is smaller variation in density estimates between realizations than designs where sampling can be unevenly distributed within a realization (Strindberg 2001). An example of the former is systematic designs, and of the latter is a completely random design. Buckland *et al.* (2001), Strindberg (2001), Strindberg *et al.* (2004) all advocate the use of systematic designs, although note that conventional analysis methods treat such samples as if they had been generated by a completely random design, thereby failing to capitalize on the increased precision (Strindberg *et al.*, 2004, p196). Ongoing work (Fewster *et al.*, in prep.) aims to rectify this, and will result in design-based variance estimators for systematic line transect designs that produce more realistic (smaller) estimates.

The importance of minimizing off-effort time, and ability to do so, varies greatly between studies. In some studies, the off-effort speed of the vessel is much greater than its speed in on-effort (survey) mode, so that it is possible to quickly move between transects that are far apart, relative to the scale of the study area. In other studies, the vessel is relatively slow off-effort, but can utilize enforced periods of inactivity such as night time to move between transects. More commonly, however, time spent off-effort transects directly into less time being available for surveying. In this circumstance, designs that minimize the distance between adjacent transects are greatly to be preferred.

Strindberg *et al.* (2004) distinguish between line transect designs where the lines are of fixed length (usually much shorter than the width of the study area or strata) and those where the lines cross from one boundary of the study area (or stratum) to the other. The majority of shipboard surveys employ designs of the latter kind. Where off-effort time is not a major concern, a systematic set of parallel lines with a random start point is the preferred design. If there is a significant density gradient remaining even after stratification then the lines should, if possible, be oriented such that they run perpendicular to any known density gradients (see Fewster and Buckland, 2004, for a simulation demonstrating the effectiveness of this strategy). If off-effort time is a concern, it can be minimized by employing a zig-zag (also called saw-tooth) design.

Strindberg and Buckland (2004) and Strindberg *et al.* (2004) describe three different classes of zig-zag designs: equal angle, equal spacing and adjusted angle. They show that the equal angle design does not have equal coverage probability unless the study area is rectangular, while the adjusted angle design does (at least along the 'design axis', i.e., the long axis used to orient the transects). However, the adjusted angle design is hard to implement in practice as it involves regular changes of course during each transect leg, so they recommend the equal spacing design as a useful compromise between practicality and almost uniform coverage probability. We compare a systematic parallel and equal spacing zig-zag design in the example design.

Zig-zag sampling algorithms require a convex study area (or strata), so for non-convex areas it is necessary to put a convex hull around the area, lay out the samplers within the convex hull, and then remove any effort that falls outside the survey area. If this results in large discontinuities in the sampler (large amounts of off-effort), then strata can be sub-divided into approximately convex sub-strata for the purposes of creating the design (this is illustrated later).

Another potential issue with zig-zag samplers is that each leg of the transect is usually treated as an independent sample, despite the fact that successive legs join together, and are therefore sampling the same space. Whether this is an issue in practice depends on the scale of the study area compared with the transects; usually for cetacean surveys it is not a problem.

Other constraints on achieving a good design

One issue that compromises even-ness of coverage is the behaviour of a design algorithm close to the edge of the study area, so-called 'edge effects'. If transects are located only strictly within the study area ('minus sampling'), this leads to lower coverage close to the edge, because locations in the middle of the study area can be surveyed if a transect is located on either side of them, but locations on the study are boundary can only be surveyed from one side. Illustrations of this effect include Buckland et al. (2001) Fig. 6.6 and Strindberg et al. Figs 7.1, 7.5 and 7.11. One solution is to extend the sampling by allowing transects to be located slightly outside the study area ('plus sampling'), but this is not generally possible in shipboard surveys where the study area is bordered by land. Edge effects generally only cause significant biases if the study area is small compared with the width of the sample strips. We illustrate some issues associated with edge effects for strata that are long and narrow in the application.

In highly non-convex areas, such as fjords or island chains it can be infeasible to employ a design that spreads the survey effort evenly throughout the survey area in each realization, because the time spent off-effort moving between different sections of line becomes more than the total ship time available for the survey. In situations like this, where it is feasible to move about efficiently within a restricted area, but not to move easily between areas, one possibility is to use a cluster sampling design. Cluster sampling is a method of concentrating survey effort into small areas without biasing the overall estimate of density. The survey area is divided up into a set of primary sampling units (PSUs) and select a random subset of these. Within each primary sampling unit, we then select a set of secondary samples, which in this case are line transects. These are used to estimate density in each PSU. The sample unit for estimating overall density is the PSU, so ideally one wants to select 10-20 PSUs. We illustrate this approach in one of our survey strata, where we develop a novel systematic scheme for selecting the PSUs.

Finally, in designing a survey, practical issues need to be taken into account, such as the need for observers to rest, the need to budget some time for bad weather, transit time to and from ports for provisioning, ship maintenance and other contingencies, etc.

Automated survey design in Distance

The free software Distance contains a design engine that lets users create and compare different distance sampling survey designs. It contains a built-in GIS for storing the study area geometry, built around the industry-standard ESRI shapefile format. There are several classes of line transect design, including fixed length, systematic parallel and zig-zag designs. Stratification is possible, although for complex designs we have found it convenient to deal with each stratum in a separate project. Users can create designs according to specification such as the number of transects required or the total amount of available survey effort.

A typical sequence of events in using Distance to aid survey design is as follows. Firstly, create a new project file and import (or manually enter) the study area boundary and the boundaries of any strata. Next, generate a grid of points over which probability of coverage will be assessed. Then create a new design. This involves specifying the sampler type (line or point transect), design class (e.g., one of three types of zig-zag design, or parallel systematic, etc), the strip width, an indication of desired survey effort, etc. One can then generate a single realization of the design or perform a simulation where many realizations are generated and statistics such as probability of coverage at each grid point, mean, maximum and minimum survey effort, total effort, number of samplers, etc are calculated. Based on these results, the design can be amended or new designs created until on is satisfied with the results. At that point, a single realization of the chosen design can be generated, and this can then be exported into a GIS or navigation system.

Examples of real survey designs created in Distance are given below, copies of the Distance projects are in the online appendix, and more details about the program are in the extensive online manual.

APPLICATION TO COASTAL BRITISH COLUMBIA SURVEY

We illustrate the above concepts using an application to a real survey design problem – a multi-species survey of cetaceans in coastal British Columbia planned for June-August 2004. The goal of the survey was to provide baseline estimates of population size of the more common cetacean and pinniped species in coastal British Columbia. While killer whales in this region are particularly well studied, abundance estimates are lacking that would enable assessment of whether by-catch of small cetaceans in commercial fisheries is sustainable. Similarly, the lack of systematically collected data on distribution of at-risk cetaceans makes it difficult to predict impacts of seismic surveys on these acoustically sensitive species.

The study area comprises most of BC coastal waters (Fig. 1), excluding areas west of the Queen Charlotte Islands and Vancouver Island and including the US waters of the Strait of Juan de Fuca. In creating the design, we divided the area into four strata, each of which has rather different geometry. We present them in order of increasing complexity from the design perspective. Stratum 1, the Queen Charlotte Basin, is a large (63,000 km^2), relatively convex area. This area is potentially subject to future oil and gas development, so a primary goal of the survey was to produce baseline abundance estimates for this stratum separately from the rest of the survey area. In the future, more complex model-based analyses might also be used to identify 'hotspots' of marine mammal density within this stratum, particularly if more surveying is undertaken in this area. Stratum 2 (13,000km^2), the southern straits between Vancouver Island and the mainland, is reasonably wide but non-convex (a J-shape) and is further broken up by a chain of small islands east of Vancouver Island (the Gulf Islands). Stratum 3, Johnstone Strait (420km^2), is very long and narrow, with several sharp bends. Although it is a similar shape to many parts of Stratum 4, it was included as a separate stratum because the survey vessel must pass along it in moving from Stratum 1 to Stratum 2, so it can therefore be given a high survey coverage without much additional ship time. Stratum 4 (12,000km^2) is a tangle of inlets, passages and fjords, and includes the remaining inshore waters of the Inside Passage and the long, narrow fjords of mainland British Columbia. It is extremely difficult to create an efficient design for this kind of topography, and so given the limited effort available it was decided to aim for preliminary estimates of abundance in this stratum. Even so, given the extent of this stratum, a means of concentrating survey effort into some parts, while still obtaining information from as broad a geographic range as possible and without biasing the abundance estimate was needed.

The survey was to be undertaken by Raincoast Conservation Society, a small Canadian non-governmental organization, and we had available 42 days ship time of the small (21m) motorized sailing vessel, *Achiever*. The vessel is large enough to accommodate the crew and survey team, so there was no requirement to return to any particular port during this time. In allocating the ship time, we wanted to include at least one week contingency for poor weather, ship repairs and moving between strata. Ship speed is approximately 9 knots (16.7 km h^{-1}) and surveying can take place for approximately 8 hours per day, although occasional longer survey periods are possible given day lengths of approximately 16 hours at that latitude in summer. The ship can move between transect lines at night if required, although frequent night-time sailing would place a great strain on the crew.

In a single-species survey, one may choose to allocate survey effort approximately in proportion to the abundance in each stratum, so that the strata containing more animals receive more effort (see previous section). However, with multi-species surveys, different species have higher densities in different areas, so it is usually impossible to optimize in this manner. Instead, we focussed on creating 'good' designs for strata 1-3, in the sense of having a randomized design with equal (or near equal) coverage probability within strata, a sample of 20 or so lines and minimum off-effort time. We also wanted to set aside some time for a preliminary survey of stratum 4.

After some initial experimentation with the survey design engine in Distance, we came up with the following initial allocation of effort: 14 days for stratum 1, 7 days for stratum 2, 2 days for stratum 3 and 10 days contingency, leaving 9 days for stratum 4. These are refined in the following sub-sections, where we treat each stratum in turn, describing the new issues each one presents and how we dealt with them.

Stratum 1. Queen Charlotte Basin – an almost convex area.

The major questions in choosing a design for this stratum were: (i) would 14 days be enough to enable approximately 20 lines to be surveyed; and (ii) how do zig-zag and systematic parallel designs compare in terms of uniformity of coverage probability, number of lines and proportion of off-effort time? A secondary issue was which design would minimize the amount of time spent in the rougher open water south of Queen Charlotte Island but north of Vancouver Island (Fig. 1)?

Assuming 8 survey hours per day at 16.6km h^{-1}, then in 14 days the ship could survey approximately 1,860 km of transect. For both systematic parallel line designs and equal-spaced zig-zags, the variable that determines survey effort is the spacing of waypoints along the side of the survey area. After some experimentation in Distance, we determined that a 32km spacing gave approximately the correct amount of effort. We therefore used the design engine in Distance to compare a parallel line design with 32km spacing between lines and an equal-spaced zig-zag design with 32km spacing between waypoints. Since the study area is slightly non-convex, we instructed Distance to place a convex hull around the area before creating the zig-zag design, and then to clip the transect lines using the actual study area.

In both cases, we set the design axis to run in a north-west to south-east direction, so that the transect lines were approximately parallel with the mainland coast. This was for three reasons. First, this means that the lines cross the short axis of the stratum, resulting in more, shorter lines. This gives a larger sample of lines and also means that lines can be surveyed in one day. Second, density of many species is related to distance from shore so this orientation captures both high and low density areas on the same lines, and minimize between-line variation in density, so increasing efficiency (see previous section). Third, strong ocean swell in the gap between the south of Queen Charlotte Island and the north of Vancouver Island often arrives from the south-west, and surveying would be easier if the ship were steaming either directly into or away from the swell.

To determine coverage probability, it was necessary to choose a truncation distance, as this determines the width of the surveyed area. Although all sightings are recorded in the field, it is standard practice to remove the few outliers with large perpendicular distance before analysis, to improve the reliability of the estimates (Buckland *et al.*, 2001, section 1.5.3). We expected that the truncation distance would be approximately 2km for the larger species (e.g., humpback and fin whale) and perhaps 0.8 km for the smaller more cryptic species (dolphin and porpoises). Since coverage problems such as edge effects are likely to be worse at large truncation distances, we present simulations ran with a strip width of 4km (twice the maximum truncation distance, since we may record animals on either side of the line). Coverage in Distance is evaluated on a grid of equally-spaced points, and we set the grid spacing at 4km. We present results based on 1000 simulations (i.e., based on generating 1000 realizations of each design).

Results of the simulations are shown in Table 1 and Figure 2, together with an example survey plan in Figure 2. The two designs were similar in the number of transects generated (~17) and on-effort transect length (~1750), although the zig-zag design did produce a maximum length of 1,903km, a little higher than our nominal allowance of 1,860km. The two designs differed in the total transect length, with the systematic design requiring approximately 180km more distance, and so having a poorer percentage on effort. Note that the figures for total line length are likely underestimates for shipboard surveys, since they are based on straight line distances between adjacent transects and transect legs, rather than attempting to account for the distance that would be required to sail around barriers such as islands.

Transects in the north and particularly in the south of the stratum were short, so it was judged that several could be done in one day, making either design feasible to perform in 14 days. We generated several instances of each design to check that it was possible to combine lines in this way for every instance generated. We also checked carefully that the maximum on-effort and total line distances from the simulations would be feasible given our nominal 14 days. This is a crucial part of selecting a design as we must be able to accept every random realization of a design before that design can be selected. For example, it would not be correct to select a design, generate a survey plan from it, reject that survey plan because the transect length was too long and generate another more acceptable one.

The number of transects generated was less than our target of 20, but reducing the transect spacing so that 20 transects were generated on average required more survey effort than we could afford. For example, simulations using a 30km transect spacing gave an mean of 20.6 samples, but a mean on-effort transect distance of 2109km and a maximum of 2270km.

One major advantage of the zig-zag design was that it required fewer days spent in the open waters of Queen Charlotte Sound (between southern Queen Charlotte Island and northern Vancouver Island). This was because, under good conditions, two long transect lines could be surveyed in one day, so using the zig-zag design we could anchor close to the mainland, go out into the open water to the end of the transect, turn straight around and sail back along the next transect and anchor again close to the mainland. With the parallel line design, there was some undesirable off-effort transit time required perpendicular to the direction of the swell, which could be strong in the open water. It might also be necessary to steam to Queen Charlotte Islands to anchor overnight in this design – again requiring significant transiting parallel to the swell.

One disadvantage of the zig-zag design was that coverage probability was noticeably uneven at the northern and southern boundaries of the study area. This indicates a failure of the algorithm for placing the first and last transect lines (see Discussion) – although the areas affected are so small that using an equal-coverage analysis is unlikely to cause significant bias.

The zig-zag design was chosen for this stratum because of the higher on-effort percentage and the smaller amount of time that would be required steaming across the swell in open water.

Stratum 2. Southern Straits – a non-convex area requiring sub-stratification to decrease off-effort time

As with the previous stratum, we used the design engine in Distance to estimate the transect spacing required to provide the correct amount of effort, and we also compared systematic parallel and zig-zag designs. However, for brevity, we here give results only for the line spacing we chose, 18km, and only on the zig-zag configuration. Instead, we focus on the issue of sub-stratification.

As mentioned in the previous section, for zig-zag designs in non-convex areas, we place a convex hull around the stratum, create the design in the convex area and then clip the transects using the actual stratum area. For highly non-convex areas such as this stratum, that results in large discontinuities between transect lines (Figure 3a). To reduce these discontinuities, we can sub-divide the stratum into a set of more convex sub-strata. If the coverage probability is the same in all sub-strata, transects from different sub-strata can be re-combined at the analysis stage so we only require 20 or so transects from all sub-strata rather than that many in each. This means we can use many sub-strata if required, although our experience is that only a few are required for most areas before the reduction in off-effort time becomes negligible. Another advantage of sub-dividing the stratum is that a different survey axis can be chosen in each sub-stratum, using the principles laid out in the previous section (maximizing the number of transect lines and orienting transects perpendicular to expected density gradients). This also applies to parallel line designs. There are, however, practical disadvantages of having many sub-strata, principally that they increase the effect of the problems with coverage associated with the first and last transect in the zig-zag algorithm.

For this stratum, we tried two and then four substrata (Figures 3b and 3c), dividing the stratum so as to make the resulting substrata as convex as possible, and also using natural barriers such as the Gulf islands where possible as stratum boundaries. Approximately equal coverage in all sub-strata was ensured by using the same transect spacing of 18km. We concluded that using four substrata resulted in relatively little off-effort time, so further sub-division would be of little benefit. Coverage probability simulations (not shown) demonstrate the same effect as for Stratum 1 of problems at the ends of the sub-strata, although again these would not lead to significant bias if ignored at the analysis stage. Note also that coverage statistics and off-effort length calculations are quite inaccurate for areas that have a large number of islands, such as the 2nd from top substratum in Fig. 3c. This is because these calculations assume that if a transect passes on one side of an island that is narrower than the strip half-width, animals on the other side will be seen. The total line length calculations are also based on straight-line distances across islands, rather than steaming around them. These assumptions are reasonable for aerial surveys, but not for shipboard. When more accurate estimates of total line length are required, we recommend generating several (perhaps 20) realizations of the design and using the line-length tool in a GIS to create a realistic path along the transects and estimate its length. This is also useful to estimate the total line length required including moving from the end of one sub-stratum to the beginning of another.

The chosen design with four substrata contained a mean of 31.7 transects (range 28-34), based on 1000 simulations. On-effort transect length was 793km (range 764-848). 20 realizations were generated and examined in detail, and all could be surveyed within the nominal 7 days.

In addition to giving a healthy number of lines and being the minimum spacing that we judged achievable, another advantage of an 18km spacing is that it is exactly half the spacing of that in stratum 1. Therefore, if more survey effort were required in stratum 1 (or more time were available at the end of the season), one option would be to double the coverage there by reflecting the survey lines around a mirror image along the design axis (i.e. to create a second survey using the opposite waypoints – e.g. Buckland *et al.*, 2001, Fig. 7.2).

Stratum 3 – a long, narrow strip with potentially significant edge effects

Stratum 3 is a long (~150 km), narrow (1-5 km) passage. Such shapes create an additional complication because the high ratio of edge to interior means that edge effects can be important. If the area is extremely narrow, so that all animals between the shores can be seen from the middle, then a complete count can be made with one pass along the middle, and a distance sampling method is not required. For slightly wider areas, a complete count could still be achieved using multiple passes, for example by passing close to one shore in one direction and then returning close to the other shore. However, care would need to be taken to avoid any overlapping coverage. For still wider areas, or where repeated passes are not practical given available effort, sampling approaches must be used.

There are at least three potential solutions, illustrated in Fig. 4.

The first (Fig. 4a) uses a disjointed set of lines oriented parallel to the long axis of the study area. An algorithm for placing the lines so as to ensure uniform coverage is as follows. First place a minimum bounding rectangle around the study area, with the long side parallel to the long axis of the study area. Buffer the long sides outwards by the truncation distance (to ensure uniform coverage at the banks - 'plus sampling'). Divide the long axis into equally spaced pieces (at least 20, if possible). Within each piece, lay down a transect parallel to the long axis and at a random (or systematic) location with respect to the short axis boundaries.

Surveying could use a line transect methodology. When the transect line is outside the water, but some of the strip out to the truncation distance is in the water, the ship would sail within the strip close to the shore. It is not necessary to remain exactly on the transect line in a line transect survey, so this poses no problem in theory; however, the distances recorded must be distance from the transect to the animal (this can be achieved using a GIS after the survey if field methods are used that record the position of the animal). Also traversing close to the shore may be difficult.

An alternative surveying method is to use a strip transect. In this method, all animals within a strip centred on the transect line and of pre-defined width are counted, and a much narrower strip width must be used than in line transects to ensure this is the case. For this reason, strip transects are less efficient.

The second potential solution (Fig. 4b) is to use a zig-zag design, for example and equally spaced zig-zag. To avoid lower coverage probability at the edges, it is necessary to extend each transect until no part of the covered strip is still inside the survey area – these parts (shaded in Fig. 4b) could be surveyed by sailing close to shore, but as with the previous approach, careful field methods would be required to determine the position of sighted animals and so judge if it was inside the surveyed area or out. To avoid large overlaps in the surveyed strips between adjacent transects, the start point of each transect would need to be displaced away from the end of the previous transect. Such a design would involve less off-effort time than the previous, although since the study area is narrow, the absolute amount of time spent off-effort will be small in both cases. Surveying could again be done by line or strip transect.

The third potential solution (Fig. 4c) is to place a systematically-spaced set of parallel lines perpendicular to the long axis of the study area – i.e., running approximately from one bank to the other. This greatly reduces the issues of edge effects, and avoids the need to extend the transects as with the previous method. Field methods are also much more straightforward, since surveying can take place from the transect line. A major disadvantage of this method, however, is that it requires much more off-effort time moving between transects. A second disadvantage is that the time required to turn the survey vessel perpendicular to the bank and begin the survey may exacerbate any problems of responsive movement of animals. This approach is, however, the only one of them currently available in Distance (as of version 4.1).

To create a design for stratum 3, we used the third of the above solutions (Fig 5.). Because the passage has several bends in it, we divided the stratum into substrata so that all transects within a substratum were close to perpendicular to the banks. In addition, there were two small inlets oriented almost perpendicular to the main direction of the passage. These would be significantly affected by edge effects, since we used 'minus sampling' – i.e., we did not extending the transect lines onto land if any part of the strip falls within one of the inlets. We therefore removed them from the distance sampling survey and designated them as 'census areas', where we would take a complete count of animals as we passed. Under the assumption that no animals in these small areas are missed, which we judge reasonable, we simply add any animals seen in these areas to the estimated abundance for the stratum when calculating total abundance. Since these areas are censused, not surveyed, they add nothing to the total variance.

Stratum 4 – a highly non-convex area and an application of cluster sampling

This stratum comprises an extremely complex set of collection of fjords, passages, straits and inlets that stretch along the entire north-south axis of the study area (Fig. 1). Evenly distributing survey effort throughout this area would require an enormous investment of ship time, so we instead used a cluster sample.

We used a GIS to divide the stratum into a set of pieces of water that could be surveyed in 1-3 (mostly 2) days using a line transect survey. This gave us 33 'primary sample units', ranging in area from 66 to 970 km^2. With only 9 days available to survey the stratum, this meant we could select only 4-5 PSUs – rather fewer than we would like. We felt that 5 represented a minimum, and accepted that we might have to commit a few of our contingency days to sampling this stratum.

In an algorithm for selecting the PSUs, the following properties are desirable. (1) The probability of selecting a PSU should be proportion to the its area, so that each part of the stratum will have the same chance of being in a sampled PSU. (2) Within any realization, should be a good geographic spread of PSUs from north to south. This implies the use of a systematic scheme. (3) No two units should be selected twice. This could be achieved by sampling without replacement (i.e., removing each PSU from the pool of potential samples once it is selected), but a disadvantage of this type of algorithm is that variance estimation is greatly complicated. Instead we created a systematic algorithm that samples with replacement, but fulfils the first two of the above criteria and has zero probability of sampling the same PSU twice if sampling intensity is not too high, in our case if fewer than 12 samples are taken. The algorithm is detailed in an Appendix, and R code used to implement the algorithm is given in the additional material.

We also needed to create a design to generate transects within the selected PSUs. Since many of the PSUs were highly non-convex with long, thin sections like stratum 3, we decided to use a systematic parallel design in each, dividing them into sub-units as required to enable us to orient the lines so as to minimize edge effects. A line spacing of 4km was manageable in the smaller PSUs, but for the larger PSUs anything less than 8km spacing was not achievable in 2-3 days. An 8km spacing produced very little absolute survey effort in the smaller PSUs, making them barely worth travelling to and surveying. Hence, we decided to use closer spacing for smaller strata, should they be selected, and wider spacing in the larger strata. This has some implications for the analysis of data from this stratum. If the line spacing is the same in all PSUs, then the overall density estimate for the stratum can be calculated as the mean of the density estimates from each PSU. If transect spacing varies between PSUs, then density estimates from some PSUs can be expected to be more precise than from others; hence, a more precise estimate of density for the stratum may be obtained by using a precision-weighted mean of the estimates from each PSU. The weighting could be coverage probability in each PSU, which is approximately proportional to expected precision, or it could be the inverse of the estimated variance.

Final realization of design

Having decided on a final design, we generated one random realization for each stratum to form our final survey plan. In stratum 4, the selected PSUs were 4, 10, 17, 21 and 29 (Fig. 6), and the line spacing we used in generating the transects within each PSU were 4, 8, 6, 8 and 6km respectively. The final survey plan is summarized in Table 2, and the location of the transects are shown in Fig 7.

DISCUSSION

Achieving good survey design

Our aim with this paper has been to promote the use of good survey design, by briefly describing the principles of good design and showing how these principles can be applied in practice in a difficult survey area and with strong constraints on ship time. We have also demonstrated how automated survey design tools such as Distance can be extremely useful in selecting a design, and generating a realization of it.

We note that Buckland *et al.* (2001, section 7.2.2) detail procedures that can be used to estimate the amount of survey effort required to achieve a desired level of precision. We did not use these in our example because the total level of effort was fixed in advance, and we simply wanted to obtain the best design possible given this total effort. We also did not have any pilot data on which to base such precision calculations. Nevertheless, in situations where one variable is the amount of effort to deploy, such calculations can be of great value.

We regard the outcome of the design process we described in our example as a success, in the sense that the design was implemented with few difficulties (see below), and produced reasonable estimates (Williams and Thomas, this meeting). One reason for this was that significant resources (~10% of the total budget) were devoted to the design process. A second was that the design was produced collaboratively using the skills of a statistician, a GIS specialist and a numerate biologist who was the intended cruise leader. We have found this partnership to be very useful in helping the biologist become familiar with the concerns inherent in good study design, as well as for the analyst to be familiar with the biology of the study animals and the challenges of field data collection specific to this project. Good discussion at all stages of this project helped the analyses (Williams and Thomas, this meeting) to go more smoothly than they otherwise might have. There are several aspects of the design that could be improved (see below) but in general we commend this process to others.

Implementing the survey plan

The survey was carried out in summer 2004. We do not present any of the results here, as they are in a companion paper (Williams and Thomas, this meeting); however we note some issues that arose during execution.

While undertaking the survey, we discovered that we would not be allowed to enter US waters. We were therefore forced to remove sections of transect within US waters from the survey plan. This presents no problem for inference however, as probability of coverage in the remaining part of stratum 2 is unaffected, so we can still use exactly the same analysis methods as planned to estimate the density and abundance of animals in the Canadian part of stratum 2.

We were not able to navigate all of the planned transect lines because some (<2%) were in water that was too shallow, or otherwise not passable in our survey vessel. We made a special effort to cover these parts of the lines visually, from as close as we could navigate to, and we do not expect this will have any significant effect on the validity of our results. Nevertheless, purchasing high-resolution digital maps and consulting more closely with the ship's captain prior to finalizing the design would have alleviated this problem.

In general, if it is not possible to collect data on parts of the planned route, there are two options. The first is to exclude the area containing the un-surveyed sections of transect from inference. The second is to assume that density on un-surveyed sections are no different from the surveyed sections, and so to apply the calculated density estimate to the original study area. Which is most appropriate depends on the circumstances.

We were able to re-survey stratum 1 in 2005. In this case, there are at least 3 options for creating a survey plan for the new survey. The first is to re-survey the old transects. This is the best option if there has been sufficient survey coverage in the first year (say 20 transects or more) and if the main interest is in monitoring trends in population size over time (Thomas *et al.*, 2004: Section 5.7). The second option is to survey as far as possible from the original transects, by using the opposite waypoints from those used in the first survey (e.g., Buckland *et al.*, 2002, Fig 7.2). This gives the best possible overall coverage, but like the first option means that estimates of density from the two time periods are not strictly independent. It is also not possible to do this automatically in Distance. The third option, and the one we took, is to generate another realization from the same design.

Possible improvements to design

Many of the complications we had to deal with would not be present if it were possible to use an aerial survey platform. We would then advocate a systematic parallel set of lines for strata 1 and 2, and the problems associated with non-equal coverage on the first and last transect would be gone. With an aerial platform, it would be possible to extend the transects outside the survey area (plus sampling), and so problems associated with edge effects in strata 3 and 4 would also be removed. Since aeroplanes can cover ground much faster off-effort, it may also be possible to implement a less clustered design for stratum 4. Despite this, in our case, a boat-based survey represented the most cost-effective problem because the boat was owned by Raincoast so did not have to be rented, the crew were largely volunteers, and the north and central coasts of BC provide a dearth of float plane terminals for refuelling.

From the coverage probability results we presented for stratum 1, it is clear that the algorithm for generating the first and last transect in equal spaced zig-zag designs (shown in Figure 7.18 of Strindberg et al. 2004) needs improvement. One possibility would be to implement the algorithm for adjusted angle zig-zags for these transects; other possibilities no doubt exist.

When zig-zag designs are used in adjacent strata or sub-strata, such as the sub-strata of stratum 2, it would be useful to extend the current algorithm so that the end of the last transect in one stratum or sub-stratum was joined with the first transect in the next. Developing an algorithm that maintains good coverage properties and works reliably when several strata or sub-strata are connected from different sides seems difficult, however.

For long, thin areas such as stratum 3 and much of stratum 4, it would be useful to further develop the displaced zig-zag design described under stratum 3, above. Other solutions are also possible, depending on the circumstances. Consider, for example, a survey for freshwater cetaceans. If density were known to be highest along the banks of a river but it is thought that there may be some animals in the middle, a strip transect could be performed close to the bank while crossing over to the opposite bank at regular intervals and performing a line transect or strip transect survey while crossing. Because the bank-side surveys are strip transects, there is no problem if the animal density changes with distance from the bank. A design like this was used in a recent survey of river dolphins in Amazonia (Sharon Hedley and Rob Williams, pers. comm.). Analysis of the data collected in BC fjords during the implementation of our survey (Williams and Thomas, this meeting) will tell us whether, for some species, our survey design could be improved for long, thin areas.

We note that because coverage probability is not the same in all strata, we will only obtain design-unbiased estimates of density or abundance if we estimate a separate detection probability for each stratum (Burnham *et al.*, 1980, p.200-201). This is in contrast with the case where the same sampling intensity is used in all strata, in which case pooling robustness applies and an estimate of abundance with low or no bias can be obtained even when the detection function is estimated from data pooled across strata (Burnham *et al.*, 2004, section 11.12). Such biases will not be large if detection probability really is similar between strata, but we need to check this at the analysis stage. Designs where coverage probability is equal in all strata are more robust in this sense, but were simply not feasible in our example.

There are alternative methods of estimating the density and/or abundance of cetacean populations, such as strip transects and mark-recapture methods. Many of these are reviewed and compared by Borchers *et al.* (2002). In our example study, the population size of resident killer whales is known very precisely from a complete census conducted during many years of intensive photo-ID workCareful consideration should be given to the method best suited to meet the goals of each particular study.

In conclusion, we stress that we have not discussed appropriate field methods at all, and appropriate analysis methods in any detail. A good survey design can produce completely useless results if either the field methods or analysis are not just as good. Analysis methods are somewhat less important, because a poor analysis can always be re-done. However, good field methods are just as critical as a good design. Some advice on appropriate field methods is given in Chapter 7 of Buckland *et al.* (2001). Nevertheless, rigorous estimates of animal abundance are a vital component of any management or conservation program, and the precision and accuracy of those estimates are improved when appropriate consideration is paid to good design. In recent years, several methodological and technological developments have emerged to facilitate good survey design. We hope that our introduction to these methods, and our experience in applying them to a complex real-world application inspire other field biologists to use them, and that our description of the process makes it easier to do so.

APPENDIX – CLUSTER SAMPLING ALGORITHM

Algorithm:

1. Let n be the number of primary sample units (PSUs) in total, and k be the number we wish to sample.
   For example, imagine n=10 and k=3.
2. Order the primary sample units (PSUs) in a geographically sensible manner (e.g., from north to south)
   For example, imagine we order the PSUs 1, …, 10.
3. Calculate the area of each primary sample unit
   For example, imagine the PSUs have areas 10, 100, 10, 100, …, 10, 100
4. Calculate the cumulative sum (cumsum) of the areas, and rescale so that maximum value is 1.
   In the example, this gives cumsum = 0.02, 0.20, 0.22, 0.40, …, 0.82, 1.00
5. Select a random number between 0 and 1 inclusive. Let the number be v.
   For example, we select v=0.19.
6. Let i = 1
7. For j in 1 to n:
   - While cumsum(j) >= v
     - $i = i+1$
     - $v = v + (1/k)$
   - Sample PSU in location j
     In our example, this yields PSUs 2, 6 and 10

Using this algorithm, the maximum number of substrata that can be sampled before there is a non-zero probability of sampling the same stratum twice is the integer part of (sum of PSU areas/maximum PSU area).

ADDITIONAL MATERIAL

Distance projects for strata 1, 2 and 3 and the 5 PSUs in stratum 4, together with a copy of the R code that implements the clustering algorithm, are available archived in a zip file at http://www.creem.st-and.ac.uk/len/BCSurveyDesign.zip (note this is case sensitive), or by request from the first author.

ACKNOWLEDGEMENTS

We are very grateful to Samantha Strindberg for making this work possible by designing and writing the survey design component of Distance. We thank Steve Buckland and David Borchers for their ideas and suggestions on the BC cetacean survey design, and we thank Raincoast Conservation Society and their funders for supporting the design effort.

REFERENCES

Borchers, D.L., S.T. Buckland and W. Zucchini. Estimating Animal Abundance: Closed Populations. Springer Verlag, London.

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..

Burnham, K.P., D.R. Anderson and J.L. Laake. Estimation of Density from Line Transect sampling of Biological Populations. Wildlife Monograph 44(2)..

Fewster, R.M. and S.T. Buckland. Assessment of distance sampling estimators. Pages 281-305 in 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..

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

Hedley, S.L., Buckland, S.T. and Borchers, D.L. 1999. Spatial modelling from line transect data. Journal of Cetacean Research and Management 1, 255-264.

Hedley, S.L., S.T. Buckland and D.L. Borchers. Spatial distance sampling models. Pages 48-69 in 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..

Pollard, J.H., S.T. Buckland and D.L. Borchers. Adaptive distance sampling surveys. Pages 229-259 in 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..

Strindberg, S. 2001 Optimized Automated Survey Design in Wildlife Population Assessment. PH.D. thesis, University of St Andrews.

Strindberg, S. and S.T. Buckland. 2004. Zigzag Survey Designs in Line Transect Sampling. *Journal of Agricultural, Biological & Environmental Statistics* 9: 443 - 461.

Strindberg, S., S.T. Buckland and L. Thomas. Design of distance sampling surveys and Geographic Information Systems. Pages 190-228 in 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.

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. and Bishop, J.R.B. 2004. Distance 4.1. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/

Thomas, L., K.P. Burnham and S.T. Buckland. Temporal inferences from distance sampling surveys. Pages 71-106 in 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..

Thompson, S.K. (2002). *Sampling, Second Edition.* New York: Wiley.

Williams, R. and L. Thomas. Submitted to this meeting. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada.

TABLES

Table 1. Comparison of systematic parallel and equal spaced zig-zag designs with 36km transect spacing for Stratum 1, based on 1000 simulations. Numbers are means, followed by minimum and maximum in brackets.

Design	Number of transects	On-effort length (km)	Total length (km)	% on effort
Systematic parallel	17.4 (17-18)	1750 (1713-1791)	2557 (2494-2599)	0.68 (0.67-0.70)
Equal spaced zig-zag	17.1 (16-18)	1777 (1700-1903)	2378 (2286-2503)	0.74 (0.73-0.76)

Table 2. Summary of final survey design and planned survey effort generated from a single realization of the design.

		Area (km^2)	Design	Survey plan		Days allocated to survey
				Number of transects	Total transect length (km)	
Stratum 1		62976	Equal spaced zig-zag, 32km spacing	17	1780	14
Stratum 2		13029	Equal spaced zig-zag, 16km spacing	33	802	7
Stratum 3		420	Systematic parallel, 6km spacing	24	70.8	2
Stratum 4	PSU 4	100	Systematic parallel, 4km spacing	15	24.0	1
	PSU 10	909	Systematic parallel, 8km spacing	19	111	2
	PSU 17	325	Systematic parallel, 6km spacing	12	52.4	2
	PSU 21	970	Systematic parallel, 8km spacing	17	117	3
	PSU 29	514	Systematic parallel, 6km spacing	21	82.8	2
	Total	11965[1]	Cluster sample of 5 primary sample units (PSUs)	84	387	10
Total		24663		166	2383	33[2]

[1]Total area of stratum 4 is greater than the area of the five primary sampling units (PSUs) that were selected.

[2]9 days remain as contingency, and for moving between strata, from the total budget of 42 days

FIGURES FOR: DESIGNING LINE TRANSECT SURVEYS FOR COMPLEX SURVEY REGIONS

Len Thomas, Rob Williams, and Doug Sandilands

Fig. 1. Coastal British Colombia, showing the 4 strata that made up the example survey.

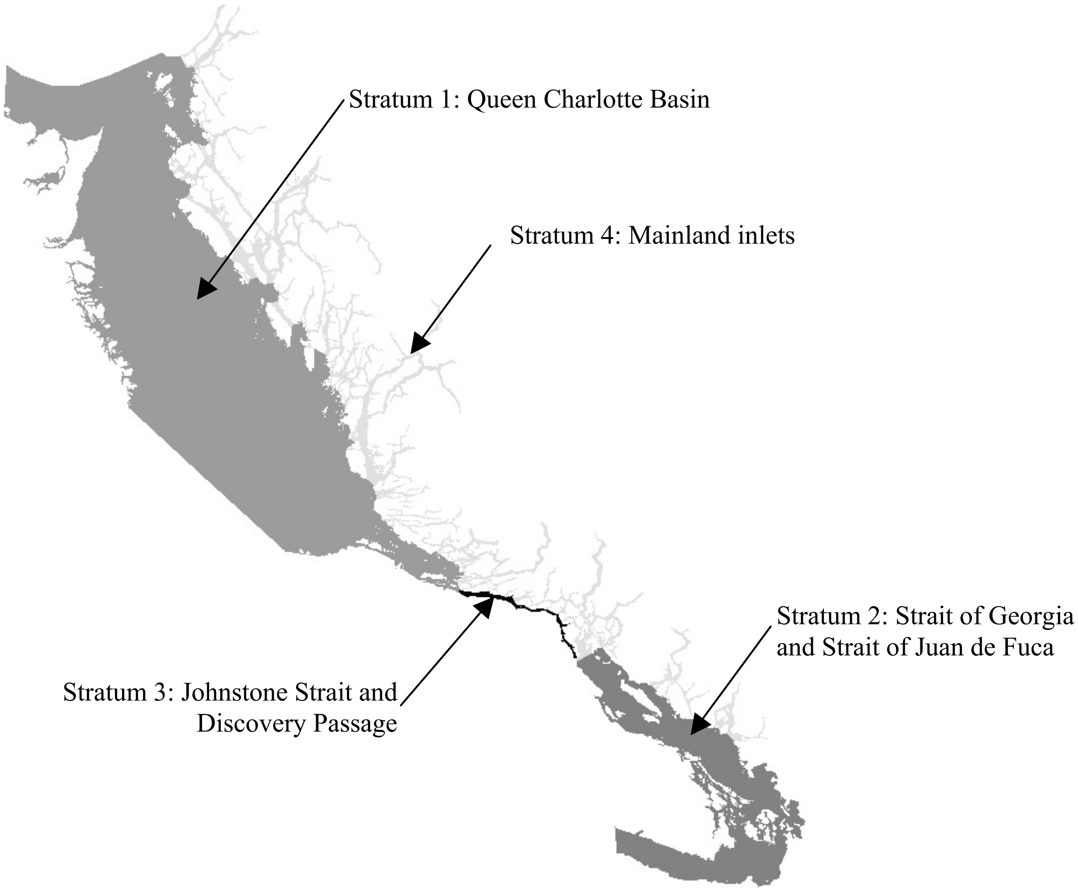


Fig. 2. Maps of coverage probability and sample realizations of systematic parallel and equal-spaced zig-zag designs in stratum 1. Both designs have 36km spacing. Coverage probabilities were generated using 1000 simulations and a 4km strip width.

Systematic parallel design

Equal angle zig-zag design

Fig 3. Example realizations of the equal spacing zig-zag design for stratum 2, demonstrating how the distance between adjacent transects, and therefore amount of off-effort time required can be decreased by subdividing the stratum into substrata. The surveys were generated using an equal spacing zig-zag design with a 16km spacing and (a) 1, (b) 2 and (c) 4 substrata. The shaded polygons behind the substrata are the convex hulls used for laying out the transects.

(a) 1 substratum (b) 2 substrata (c) 4 substrata

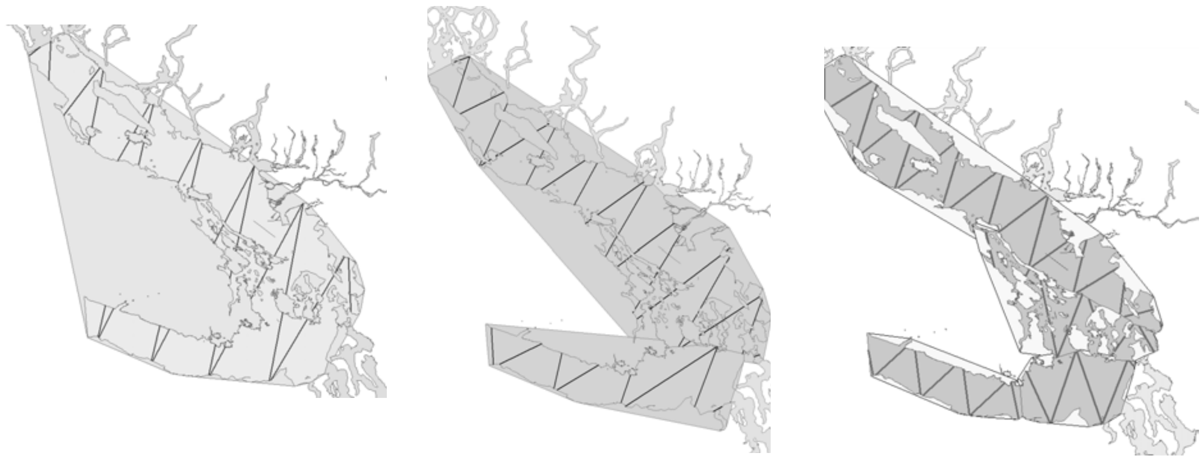
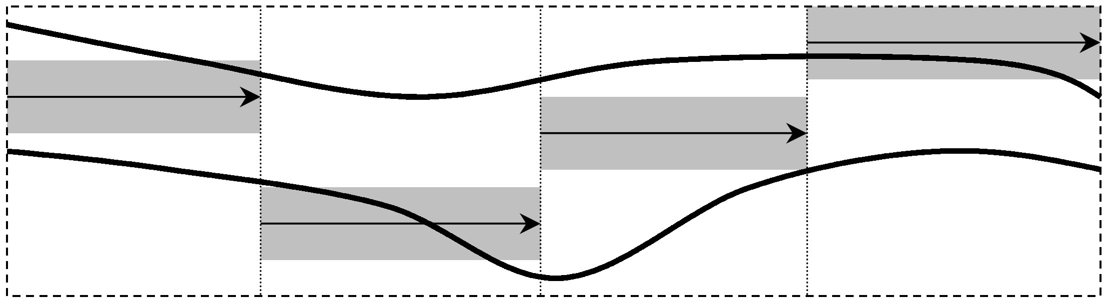
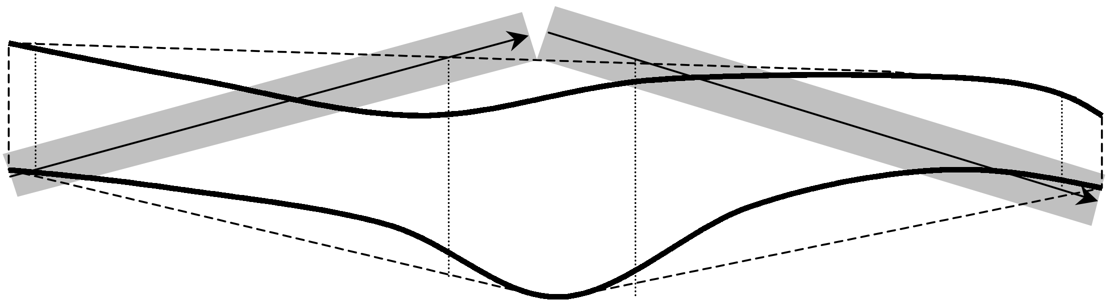


Fig 4. Examples of different designs for long, narrow survey areas. Arrows represent the nominal transect lines. Shading represents the area potentially surveyed, although only animals within the study area boundary (denoted by thick dark lines are counted). (a) Parallel design where transect lines are oriented along the long axis of the study area, and positioned at random locations with respect to the short axis. (b) Modified equal spacing zigzag design where sampling continues until none of the surveyed strip is within the convex hull (dashed line) around the study area to ensure uniform coverage, and where waypoints are displaced so that there is no overlap between adjacent transects. (c) Systematic parallel design, where transects are oriented perpendicular to the long axis of the study area.

(a)



(b)



(c)

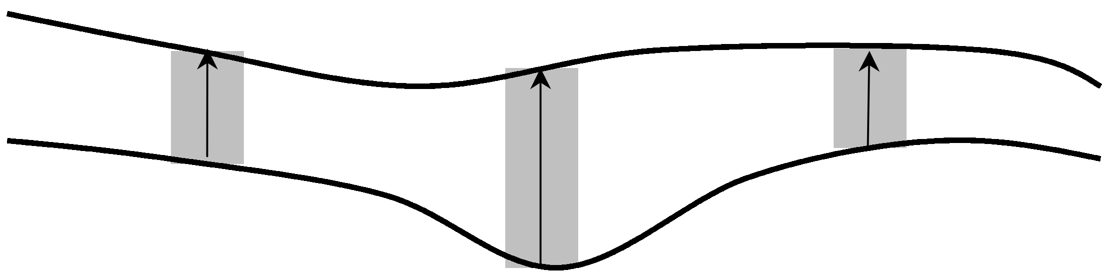


Fig 5. Example survey plan for stratum 3, created using a systematic parallel line design with 6km line spacing. The stratum has been divided into 6 substrata to enable the lines to be approximately parallel to the banks of the stratum. There are two small bays scheduled for a complete census.

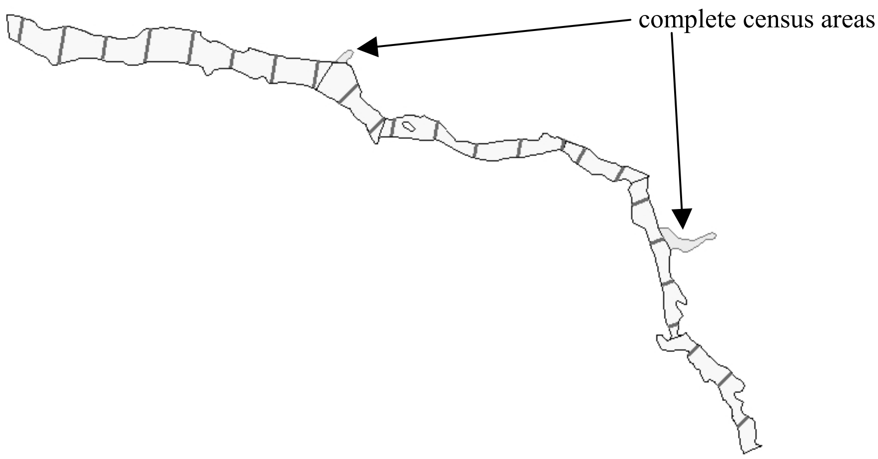


Fig 6. Selected primary sampling units (PSUs) and example realizations of the designs for each unit in stratum 4.

Fig 7. Final survey plan.