Designing line transect surveys for complex survey regions

LEN THOMAS*, ROB WILLIAMS# AND DOUG SANDILANDS++

Contact email: len@mcs.stand.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 valuable. A ‘good’ design is one (a) that employs randomisation 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 (uniform coverage probability); (d) that produces an even distribution of transects throughout each stratum (e.g. systematic random designs); (e) that produces at least 1020 transects per stratum; (f) that, given the previous points, gives maximum efficiency per unit effort – for example by minimising 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 multispecies survey of cetaceans in coastal British Columbia. The design uses an equally spaced zigzag configuration of transects in more open strata combined with sub-stratification to minimise 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 online.

KEYWORDS: ABUNDANCE ESTIMATE; PACIFIC OCEAN; NORTH AMERICA; SAMPLING STRATEGY; SURVEY -VESSEL

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. This paper focuses on strategies for creating good survey designs in the context of shipboard surveys of cetaceans.

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’). For the case presented here, no assumptions need to be made about the spatial distribution of the animals. This kind of method, where the properties of the design are used 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 an appropriate design is coupled with a design-based analysis method, unbiased estimates are obtained.

In contrast, analysis methods exist in which inferences are made from the survey data about the density of animals in the whole study area based on a model for the distribution of animals (e.g. Cañadas and Hammond, In press; Hedley and Buckland, 2004; Hedley et al., 2004; Hedley et al., 1999). Such model-based methods do not make any assumptions about the manner in which the transect lines were laid out, although they do rely on data having been collected across a range of values of the covariates used to model abundance along the transect lines and extrapolate it to the whole study area. They can therefore potentially be used in cases where there was no element of randomisation 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 area 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 even when adequate covariate information is available. 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 10.3). The development of appropriate model-based methods for cetacean line transect data is an active area of current research (e.g. Hedley et al., 2004).

Since model-based methods are not guaranteed to be unbiased, it is often desirable to be able to produce a design-based estimate with which to compare them. 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 to be used.

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 strategies in layout 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, 2004b). 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 (i.e. with no internal angle greater than 180 degrees), 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 and Buckland, 2004a) and their implementation in the free software program Distance (Thomas et al., 2003) has considerably simplified the task of creating and comparing complex designs. Designs can be created on the computer, and many random realisations can be generated at random with respect to the constraints inherent in many shipboard surveys of cetaceans are discussed and ideas and solutions are illustrated using the example of a multi-species survey of cetaceans in coastal British Columbia, Canada. The study area included stretches of open water as well as an intricate system of inshore islands and fjords. Stratification was used and a systematic cluster sampling algorithm in the inshore stratum was developed to increase survey efficiency. A zig-zag transect configuration is then compared with a more conventional parallel line configuration.

In designing the survey, the Distance software was used extensively and the relevant Distance projects and related geographic information files are available online (see Additional Material, at the end of this paper). By providing this material, interested readers can be given a ‘jump-start’ in how to implement automated survey design methods in complex regions using available software.

The survey designed here was carried out in 2004, and one stratum was re-surveyed in 2005. The results are presented in a companion paper (Williams and Thomas, 2007).

SURVEY DESIGN – CONCEPTS AND STRATEGIES

Line transect sampling is 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 realisation can be generated to described by Strindberg et al. (2004a), which is Chapter 7 of the more advanced text by Buckland et al. (2004).

As noted above, 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, it is preferable to obtain these results initially using design-based analysis methods, so designs that can yield reliable design-based results can be focused on.

Design requirements

Two essential requirements for a good design are randomisation and replication (Buckland et al., 2001, section 7.2.1). Randomisation 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 the algorithm is executed, a different random realisation is obtained. Standard analysis methods assume that on average over many realisations, 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 animals, the true density of animals is, on average, the same near to the transect line as it is far from it. Therefore, any change in the frequency of animal detections with increasing distance from the line can be interpreted as a change in the probability of detection, rather than a change in true density. This enables us to estimate change in detection probability with distance from the line, and (with some other assumptions) to then estimate average probability of detection and hence density of animals in the covered region (i.e. the part of the study area ‘covered’ by samplers and therefore surveyed). Secondly, because all areas are equally likely to be sampled, this density estimate can be applied to the whole survey area, not just the covered region.

If a design algorithm is used that involves randomisation, 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 and Buckland, 2004a, section 7.3).

Replication (i.e. placement of multiple lines) is required for assessment of the uncertainty in design-based estimates. Increasing the number of replicate lines increases the reliability of variance estimates with other factors being equal (such as total line length and evenness 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. (Note that if there is a strong density gradient in the population then the lines should instead be oriented parallel to the gradient, as explained in the next section.)

Other issues contributing to good designs

While not required for unbiased density estimation, 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 distribution of samplers within each realisation; and minimisation of off-effort time.
Stratification is valuable where there are large-scale gradients in animal density and these gradients are predictable. In this case, dividing the study area into strata could allow one to use systematic sampling designs, so as to maximise the between-stratum variation in density and minimise the within-stratum variation that can lead to greatly increased precision of estimates. One constraint is that if the aim is to estimate the density for each stratum, an adequate sample of transects (again 10-20) per stratum is required for reliable variance estimation. If the strata are created merely for the purpose of increasing precision in the overall estimate, and there is no requirement to estimate density for each stratum, as few as two lines per stratum are acceptable (see Fewster et al., 2004).

In addition, precision can be further increased by dividing the total survey effort between strata so 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 the detection function is fitted to data pooled across strata (Burnham et al., 1980, pp.200-01). This will happen when detectability varies between strata (for example due to different average survey conditions) – if, for example, coverage probability is higher in a stratum where detectability is also higher then the average detection probability will be biased high, and average density will be underestimated. By contrast, if coverage probability is the same in all strata, the pooling robustness property of the standard estimators of detection probability mean that the pooled estimate is approximately unbiased even when detectability varies between strata (Burnham et al., 1980, pp.45-7; Burnham et al., 2004, section 11.12). Furthermore, allocating low sampling effort to low density strata often means that there are too few detections in the stratum (and it is therefore impossible to create the detection functions for them, so there is no option but to pool. 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.

As mentioned previously, it is desirable to have uniform probability of coverage averaged across many realisations. It is also desirable to choose designs that produce an even distribution of transects across the study area (or stratum, if using stratification) within each realisation. Designs with this property produce more reliable results in the sense that there is smaller variation in density estimates between realisations than designs where transects can be unevenly distributed within a realisation (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. (2004a) all advocated the use of systematic designs (with a random start-point to provide some element of randomisation), and this advice has been followed in the example design.

The importance of minimising 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 while on-effort – for example, in surveys of relatively small areas using rigid inflatable boats. In this situation, it is possible to quickly move between transects, and the preferred transect configuration (Strindberg and Buckland, 2004a) is a systematic set of parallel lines that cross from one boundary of the study area (or stratum) to the other with a random start point. If there is a significant density gradient remaining even after stratification then the lines should, if possible, be oriented such that they run parallel to any known density gradients (i.e. perpendicular to the density isolines). For a simulation demonstrating the effectiveness of this strategy see Fewster and Buckland (2004). In other studies, the vessel is relatively slow off-effort, but can utilise enforced periods of inactivity such as night time to move between adjacent parallel transects. More commonly, however, the distance between transects translates directly into a near-equivalent reduction in distance available for on-effort surveying. In this circumstance, transect configurations that minimise the between-transect distance, such as zig-zag designs are greatly to be preferred.

Strindberg and Buckland (2004b) and Strindberg et al. (2004a) described three different classes of zig-zag designs: equal angle, equal spacing and adjusted angle. They show that the equal angle design does not produce uniform coverage probability unless the study area is rectangular, while the adjusted angle design does (at least in the direction of the ‘design axis’ – 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, below.

Zig-zag sampling algorithms require a convex study area (or stratum), 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 study area. If this results in large discontinuities in the sampler (large distances between the end of one transect and the beginning of the next), 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 therefore sample overlapping 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 species-specific issues may need to be taken into account during survey design. For example, if animals are thought to show large-scale directional movement then it is preferable to lay out transects such that the design axis is perpendicular to this movement. Such issues also affect implementation: in the above example with a multiple stratum design it is preferable to survey either side of common stratum boundaries as close together in time as possible.

**Other constraints on achieving a good design**

One issue that compromises uniformity 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, while locations at the study area boundary can only be surveyed from one side. Illustrations of this effect include Buckland et al. (2001, fig. 6.6) and Strindberg et al. (2004a, 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 usually only
cause significant biases if the study area is small compared
to the grid spacing, and this is convenient to deal with. Some
issues associated with edge effects for strata that are long and narrow
are illustrated in the example design.

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
realisation, 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 possible 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 standard method of concentrating survey effort
into small areas without biasing the overall estimate of
density. The survey area is divided into a set of Primary
Sampling Units (PSUs) and a random subset of these is
selected. Within each primary sampling unit, we may then
select a set of secondary samples, which in this case are line
 transects. These are used to estimate density in each sampled
PSU. The sample unit for estimating overall density is the
PSU, thus ideally 10–20 PSUs should be selected. This
approach is illustrated in one of the example survey strata,
which uses 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 software Distance (Thomas et al., 2003) contains a
design engine that lets users create and compare different
distance sampling survey designs. It contains a built-in
Graphical Information System (GIS) for storing the study
area geometry, built around the industry-standard ESRI
shapefile format (Environmental Systems Research
Institute, 1998). There are several classes of line transect
design, including fixed length, systematic parallel and zig-
zag designs. Stratification is possible, although for complex
designs it is convenient to deal with each stratum separately.
Users can create designs according to specifications 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 (which for line transects is twice the truncation
distance), an indication of desired survey effort, and other
parameters. One can then generate a single realisation of the
design or perform a simulation where many realisations 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 one is satisfied with the results. At that
point, a single realisation of the chosen design can be
 generated, and this can then be exported into a GIS or
navigation system for field implementation.

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

APPLICATION TO COASTAL BRITISH COLUMBIA SURVEY

The above concepts are illustrated here using an application
to a real survey design problem – a multi-species survey of
Cetaceans in coastal British Columbia (BC), Canada,
between June and August 2004. The goal of the survey was
to provide baseline estimates of population densities and
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 (13,000km2),
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 (420km2), 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
(12,000km2) 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 BC. 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 on a small (21m)
motorised sailing vessel (Williams and Thomas, 2007), with
up to 42 days ship time available. The vessel was 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, at least one week
contingency was desirable to allow for poor weather, ship
repairs and moving between strata. Ship speed is
approximately 9 knots (16.7km h−1) and surveying can take
place for approximately 8hr per day, although occasional
longer survey periods are possible given day lengths of
approximately 16hr at that latitude in summer. The ship

could move between transect lines at night if required, although frequent night-time sailing places a great strain on the crew.

In a single-species survey, one may choose to allocate survey effort approximately in proportion to some prior estimate of abundance in each stratum, so that the strata expected to contain 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 optimise in this manner. Instead, creating ‘good’ designs for strata 1-3 was focused on, in the sense of having a randomised design that gives uniform (or near uniform) coverage probability within strata, a sample of 20 or so lines evenly distributed through each stratum and minimum off-effort time. Time for a preliminary survey of stratum 4 was set aside.

After some initial experimentation using the survey design engine in \textit{Distance}, the following initial allocation of effort was decided upon: 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.

Stratum 1. Queen Charlotte Basin – an almost convex area

The major questions in choosing a design for this stratum were (1) would 14 days be enough to enable approximately 20 lines to be surveyed, and (2) 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 minimise the amount of time spent in the rougher open water south of Queen Charlotte Island but north of Vancouver Island (Fig. 1).

Assuming eight survey hours per day at 16.6 km h$^{-1}$, then in 14 days the ship could survey approximately 1,860km 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 using \textit{Distance}, it was determined that a 36km spacing gave approximately the correct amount of effort. Therefore the design engine in \textit{Distance} was used to compare a systematic parallel line design with 36km spacing between lines and an equal-spaced zig-zag design with 36km spacing between waypoints. As the study area is slightly non-convex, \textit{Distance} was instructed 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, the design axis was set to run in a north-west to south-east direction, so that the transect lines were approximately perpendicular to the mainland coast. This was for three reasons. First, this meant that the lines crossed the short axis of the stratum, resulting in more, shorter lines. This gave a larger sample of lines and also meant that lines could be surveyed in one day, enabling the ship to spend the night at a sheltered anchorage. Second, density of many species is related to distance from shore so this orientation captured both high and low density areas on the same lines and minimised between-line variation in density, thereby 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 covered strips. Although all sightings were 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). The distance beyond which observations are discarded is called the ‘truncation distance’. It was expected that the truncation distance would be approximately 2km for the larger species (e.g. humpback and fin whales) and perhaps 0.8km for the smaller more cryptic species (dolphin and porpoises). As coverage

![Fig. 1. Coastal British Colombia, showing the four strata that made up the example survey.](image)
problems such as edge effects are likely to be magnified with large truncation distances, simulations used a strip width of 4km (twice the maximum truncation distance, since animals on either side of the line may be recorded). Coverage probability throughout the study area was evaluated in *Distance* at a grid of equally-spaced points, and the grid spacing was set at 4km. The results presented are based on 10,000 simulations (i.e. based on generating 10,000 realisations of each design) – this is an order of magnitude more than is typically used and effectively removed almost all simulation error from the maps of coverage probability.

Results of the simulations are shown in Table 1 and Fig. 2, together with an example realisation from the design (a ‘survey plan’) in Fig. 2. The two designs were similar in the mean number of transects generated (~17) and on-effort transect length (~1,750km) and in their ranges – although the zig-zag design did produce a maximum length of 1,903km, a little higher than the nominal allowance of 1,860km. The two designs differed in the total transect length, with the systematic parallel design requiring approximately 180km more, and so having a poorer percentage on effort. Note that the figures for total line length are 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.

Transsects 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. Several instances of each design were generated to check that it was possible to combine lines in this way for every instance generated. The maximum on-effort and total line distances from the simulations were carefully checked to ensure they would be feasible given our nominal 14 days. This is a crucial part of selecting a design as one must be able to accept every random realisation 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 plan.

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 was possible due to budgetary constraints. For example, simulations using a 30km transect spacing gave a mean of 20.6 samples, but a mean on-effort transect distance of 2,109km and a maximum of 2,270km.

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 it is possible to 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 non-uniform 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 a design-based analysis that assumes uniform-coverage is unlikely to result in significant bias in estimates of species abundance.

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, the design engine in *Distance* was used to estimate the transect spacing required to provide the correct amount of effort and also systematic parallel and zig-zag designs were compared. However, for brevity, the
results are given only for the line spacing chosen (18km) and only on the zig-zag configuration. Instead, the focus here is on the issue of subdividing the stratum area. For highly non-convex areas such as this stratum, this procedure results in large discontinuities between transect lines (Fig. 3(a)), which translates into considerable time spent off-effort travelling between the end of one transect and the beginning of the next. To reduce the size of the discontinuities, the stratum can be sub-divided 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 recombined at the analysis stage so only 20 or so transects are required from all sub-strata combined rather than that many in each. This allows use of many sub-strata if required, although our experience is that only a few are required for most study area shapes 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 (maximising 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: they increase the effect of the problems with non-uniform coverage on the first and last transect in the current equal-spaced zig-zag algorithm; also off-effort time is required to move from the end of the transects in one to be beginning of those in the next.

For this stratum, two and then four sub-strata were tried (Figs 3(b) and 3(c)), dividing the stratum so as to make the resulting sub-strata as convex as possible, and also using natural barriers such as the Gulf Islands where possible as stratum boundaries. Approximately equal coverage probability between sub-strata was ensured by using the same way point spacing (18km) in all sub-strata. It was concluded that having four sub-strata resulted in relatively little off-effort time, so further sub-division would be of little benefit. Coverage probability simulations (not shown) demonstrate problems at the ends of the sub-strata like those observed for stratum 1, although again these would probably 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 second from top sub-stratum in Fig. (3)c. 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, i.e. to create a second or survey using the opposite way points (e.g. Buckland et al., 2001, fig. 7.2) leading to approximately equal coverage in the two strata.

**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. Note that this approach assumes that all animals can be detected with certainty – e.g. that diving behaviour does not cause a problem. 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 or double-counting of animals. 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. 4(a)) uses a discontinuous 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 of the rectangle outwards by the truncation distance (to ensure uniform coverage at the banks: ‘plussampling’). 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 can follow line transect methodology. When the transect line is outside the study area (i.e. out of the water), but some of the strip out to the truncation distance is inside, 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 used in the analysis must be perpendicular distance from the original transect line (not the ship’s path) to the animal. This can be calculated, for example using GIS software, 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, using the same line placement, is to use a strip transect. In this method, all animals within a strip centred on the transect line and of predefined 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. Note that the assumption of complete detection within the strip must be tested (e.g. by collecting distance data and checking that the fitted detection function is flat). Naive use of strip transect methods without such tests will almost certainly lead to underestimation of abundance.

The second potential solution (Fig. 4(b)) is to use a zig-zag design (e.g. an 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 (or a convex hull containing the survey area) – these parts (shaded in Fig. 4(b)) 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 covered region or not. To avoid large overlaps in the covered 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 one, 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. 4(c)) 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 shore and begin the survey may exacerbate any problems of responsive movement of animals. This approach is, however, the only one available in Distance version 4.1.

To create a design for stratum 3, the third of the above solutions was used (Fig. 5). Since the passage has several bends in it, the stratum were divided into sub-strata so that all transects within a sub-stratum were approximately 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 ‘minus sampling’ was used – i.e. the transect lines were not extended onto land if any part of the strip fell within one of the inlets. These small inlets were therefore removed from the distance sampling survey and were designated as ‘census areas’, where a complete count of animals would be taken as we passed. Under the reasonable assumption that no animals in these small areas are missed, any animals seen in these areas can simply be added 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 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 a cluster sampling scheme was used.

A GIS was used to clip sections of the stratum 4 shapefile into pieces of water that could be surveyed in 1-3 (mostly 2) days using a line transect survey. Decisions on where to clip the stratum 4 shapefile were made primarily on the logistics of conducting surveys, such as vessel speed and proximity to suitable anchorages. This gave 33 primary sample units.
The weighting could be coverage probability in each PSU, or it could be the inverse of the estimated variance. Which is approximately proportional to expected precision, or it could be the inverse of the estimated variance.

Implementing the survey plan

The survey was carried out in summer 2004. The results are available in a companion paper (Williams and Thomas, 2007); however some issues that arose during execution are noted here.

While undertaking the survey, it was discovered that surveys would not be allowed in US waters. Therefore sections of transect within US waters from the survey plan had to be removed. This presented no problem for inference however, as the probability of coverage in the remaining part of stratum 2 was unaffected, so exactly the same analysis methods as planned could be used to estimate the density and abundance of animals in the Canadian part of stratum 2.

It was not possible to navigate all of the planned transect lines because some (~2%) were in water that was too shallow, or otherwise not passable in the survey vessel. A special effort was made to cover these parts of the lines visually, from as close as could be navigated to and it is not expected this will have any significant effect on the validity.

DISCUSSION

Achieving good survey design

The aim of this paper was to encourage 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. Automated survey design tools such as Distance can be extremely useful in specifying and contrasting survey designs, and generating a realisation of the selected design.

Buckland et al. (2001, section 7.2.2) detailed procedures that can be used to estimate the amount of survey effort required to achieve a desired level of precision. These were not used here because, as logistics often dictate, the total level of effort was fixed in advance and the aim was simply to obtain the best design possible given this total effort. In addition, the pilot data on which to base such calculations were not available. 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 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, 2007). 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 the project leader, a biologist. This partnership was found 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. There are several aspects of the design that could be improved (see below) but in general we recommend this process to others.

Finally realising of design

Having decided on a final design, one random realisation was generated for each stratum to form the final survey plan. In stratum 4, the selected PSUs were 4, 10, 17, 21 and 29 (Fig. 6), and the line spacings used in generating the transects within each PSU were 4, 8, 6, 8 and 6km respectively. The final survey plan is summarised in Table 2, and the locations of the transects are shown in Fig. 7.
of the results. Nevertheless, purchasing high-resolution digital maps and consulting more closely with the ship’s captain prior to finalising 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 is 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.

It was possible to re-survey stratum 1 in 2005. In this case, there were at least three options for creating a survey plan for the new survey. The first was to re-survey the old transects. This would have been the best option if there had been sufficient survey coverage in the first year (say 20 transects or more) and if the main interest was in monitoring trends in population size over time (Thomas et al., 2004, section 5.7). The second option was 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., 2001, fig 7.2). This would give the best possible overall coverage, and hence potentially the most precise estimate of total abundance, but like the first option meant that estimates of density from the two time periods would not be strictly independent. It is also not possible to do this automatically in Distance. The third option (and the one we took) was to generate another realisation from the same design.

Possible improvements to design
Many of the complications encountered would not have been present if it were possible to use an aerial survey platform. A systematic parallel set of lines for strata 1 and 2 would have been advocated and the problems associated with non-uniform coverage on the first and last transect would have disappeared. 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 than
ships, it may also have been possible to implement a less clustered design for stratum 4. Despite this, for the particular survey described here, a boat-based survey represented the most cost-effective solution because the boat 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. In many cases, boats are also used because they can collect additional data (e.g. identification photographs, biopsies, tags or ancillary oceanographic data), not feasible from an aerial platform.

From the coverage probability results presented for stratum 1, it is clear that the algorithm for generating the first and last transect in equal spaced zig-zag designs (Strindberg and Buckland, 2004a, fig. 7.18) 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 is 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 the Colombian Amazon (Hedley and Williams, pers. comm.), and a similar design, using an equal angle zig-zag in the central region of the river, was used by Vidal et al. (1997) and Martín et al. (2004).

Since coverage probability is not the same in all strata, unbiased estimates of density or abundance will only be obtained if a separate detection probability is estimated for each stratum (Burnham et al., 1980, pp.200-01). 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 total 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 this must be checked at the analysis stage, for example by fitting separate detection functions to each stratum, or using multiple-covariate distance sampling methods (Marques and Buckland, 2004; Marques et al., In press). Designs where coverage probability is equal in all strata are more robust in this sense, but were not feasible in this example.

Although systematic survey designs, such as those used here, tend to produce more reliable results than completely random designs, the conventional analysis methods treat such samples as if they had been generated by a completely random design, thereby failing to capitalise on the increased precision (Strindberg and Buckland, 2004a, p.196). Ongoing work ( Fewster et al., Submitted; Fewster, pers. comm.) aims to rectify this, and will result in design-based variance estimators for systematic line transect designs that produce more realistic (smaller) estimates.

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 from a theoretical perspective by Borchers et al. (2002) and in the context of cetacean studies by Hammond (1995; 2001) and Evans and Hammond (2004). In this 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 work. Careful consideration should be given to the method best suited to meet the goals of each particular study.

Finally, we stress that we have not discussed appropriate field methods at all, or appropriate analysis methods in any detail. A good survey design can produce comprehensively incorrect results if either the field methods or analysis are inadequate. Although a poor analysis methods can be re-done; good field methods are 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 programme, 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 review of these methods, and our experience in applying them to a complex real-world application, will inspire other practitioners to use them, and will result in better estimates of cetacean abundance.

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 in Additional Material posted on the IWC website (http://www.iwcoffice.org/publications/additions.htm), 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 Raincoast Conservation Society and their funders for supporting the design effort. This paper was improved by comments on earlier drafts made by Steve Buckland, Eric Rexstad, Tiago Marques, members of the Scientific Committee of the International Whaling Commission and two anonymous reviewers.

REFERENCES

Date received: May 2006.
Date accepted: January 2007.
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 \((1/k)\) inclusive.
   Let the number be \( v \).
   For example, we select \( v=0.19 \).
6. For \( j \) in 1 to \( n \):
   - While \( \text{cumsum}(j) \geq v \)
     - Sample PSU in location \( j \)
     - \( v = v + (1/k) \)
   In the example, this yields PSUs 2, 6 and 10

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