Distance sampling

Len Thomas, Stephen T. Buckland, Kenneth P. Burnham, David R. Anderson, Jeffrey L. Laake, David L. Borchers & Samantha Strindberg

Volume 1, pp 544–552

in

Encyclopedia of Environmetrics
(ISBN 0471 899976)

Edited by

Abdel H. El-Shaarawi and Walter W. Piegorsch

Distance sampling

Distance sampling is a widely-used group of closely related methods for estimating the density and/or **abundance** of biological populations. The main methods are *line transects* and *point transects* (also called *variable circular plots*). These have been used successfully in a very diverse array of taxa, including trees, shrubs and herbs, insects, amphibians, reptiles, birds, fish, marine and land mammals. In both cases, the basic idea is the same. The observer(s) perform a standardized survey along a series of lines or points, searching for objects of interest (usually animals or clusters of animals). For each object detected, they record the distance from the line or point to the object. Not all the objects that the observers pass will be detected, but a fundamental assumption of the basic methods is that all objects that are actually on the line or point are detected. Intuitively, one would expect that objects become harder to detect with increasing distance from the line or point, resulting in fewer detections with increasing distance. The key to distance sampling analyses is to fit a *detection function* to the observed distances, and use this fitted function to estimate the proportion of objects missed by the survey. From here we can readily obtain point and interval estimates for the density and abundance of objects in the survey area. The basic methods (sometimes called *conventional* or *standard distance sampling*) are described in detail in [5], which is an updated version of [4]. Free software, Distance [19], provides for the design and analysis of distance sampling surveys, and is being updated to include much of the work mentioned in the section on Current Research below.

Distance sampling is an extension of quadrat-based sampling methods. Two forms of quadrat sampling are *strip transects*, in which the observer travels along a line, counting all objects within a predetermined distance of the line, and *point counts*, in which numbers of objects (usually birds or plants) in a circle about a point are counted. Population density is then estimated by dividing the total count by the total area surveyed. A fundamental assumption of these methods is that all objects within the strip or circle are counted. This assumption is difficult to meet for many populations, and cannot be tested using the survey data. Furthermore, for scarce species, the methods are wasteful, because detections of objects beyond the strip or circle boundary are ignored. If the width of the strip or the radius of the circle is made sufficiently small that detection of any object within the surveyed area is almost certain, then perhaps 50% or more of detections are outside the surveyed area and so are ignored. Distance sampling extends quadrat-based methods by relaxing the assumption that all objects within the circle or strip are counted. By measuring distances to the objects that are observed, the probability of observing an object within the circle or strip can be estimated.

Another approach to estimating wildlife abundance involves **capture–recapture methods**. These are often more labour-intensive and more sensitive to failures of assumptions than distance sampling. However, they are applicable to some species that are not amenable to distance sampling methods, and can yield estimates of survival and recruitment rates, which distance sampling cannot do. Capture–recapture methods can be useful for populations that aggregate at some location each year, whereas distance sampling methods are more effective on dispersed populations. They should therefore be seen as different tools for different purposes (see also trapping webs under Related Methods below).

In **fisheries** applications, *catch per unit effort*, *catch-at-age* and *catch-at-length* are all commonly used to estimate abundance [10], as they require that the commercial catch is sampled, which is more cost-effective than sampling the living fish. Acoustic surveys of fish schools can provide data amenable to distance sampling methods.

Alternative methods for estimating animal abundance are reviewed and compared in [15]–[18] and [21].

Line-transect Sampling

In **line-transect sampling**, a series of straight lines (tracklines) is traversed by an observer. This may be achieved in various ways, depending on the study species. In terrestrial studies, these include walking, horseback, all-terrain vehicle, aeroplane and helicopter. Transect surveys in aquatic environments can be conducted by divers with snorkels or SCUBA gear, from submarines, surface vessels, aircraft, or from sleds with mounted video units pulled underwater by a surface vessel. In the case of large observation platforms, there is typically a team of observers.

Estimation

Perpendicular distances x are measured from the line to each detected object of interest. In practice, detection distances r and detection angles θ are often recorded, from which perpendicular distances are calculated as $x = r\sin\theta$. Suppose k lines of lengths $l_1, \ldots, l_k$ (with $\sum l_j = L$) are positioned according to some randomized scheme, and n animals are detected at perpendicular distances $x_1, \ldots, x_n$. Suppose in addition that animals further than some distance w from the line (the truncation distance) are not recorded. Then the surveyed area is $a = 2wL$, within which n animals are detected. Let P_a be the probability that a randomly chosen animal within the surveyed area is detected, and suppose an estimate $\widehat{P}_a$ is available. Then animal density D is estimated by

$$\widehat{D} = \frac{n}{2wL\widehat{P}_a} \tag{1}$$

To provide a framework for estimating P_a, we define the detection function $g(x)$ to be the probability that an object at distance x from the line is detected, $0 \le x \le w$, and assume that $g(0) = 1$. That is, we are certain to detect an animal on the trackline. If we plot the recorded perpendicular distances in a histogram, then conceptually the problem is to specify a suitable model for $g(x)$ and to fit it to the perpendicular distance data. As shown in Figure 1, if we define $\mu = \int_0^w g(x)\,\mathrm{d}x$, then $P_a = \mu/w$. The parameter μ is called the effective strip (half-) width; it is the distance from the line for which as many objects are detected beyond μ as are missed within μ (Figure 1). Thus

$$\widehat{D} = \frac{n}{a\widehat{P}_a} = \frac{n}{2wL\hat{\mu}/w} = \frac{n}{2\hat{\mu}L} \tag{2}$$

We now need an estimate $\hat{\mu}$ of μ. We can turn this into a more familiar estimation problem by noting that the probability density function (pdf) of perpendicular distances to detected objects, denoted $f(x)$, is simply the detection function $g(x)$, rescaled so that it integrates to unity (*see* **Frequency curves**). That is, $f(x) = g(x)/\mu$. In particular, because we assume $g(0) = 1$, it follows that $f(0) = 1/\mu$ (Figure 2). Hence

$$\widehat{D} = \frac{n}{2\hat{\mu}L} = \frac{n\hat{f}(0)}{2L} \tag{3}$$

The problem is reduced to modeling the pdf of perpendicular distances, and evaluating the fitted function at $x = 0$. The large literature for fitting density

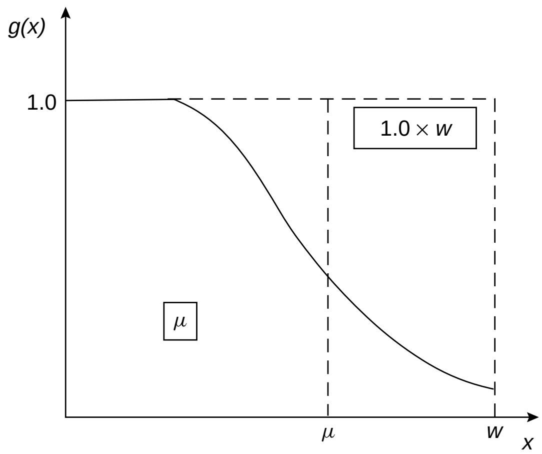

Figure 1 The area μ under the detection function $g(x)$, when expressed as a proportion of the area w of the rectangle, is the probability that an object within the surveyed area is detected; μ is also the effective strip width, and takes a value between 0 and w. Reproduced from Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1998). Distance sampling, in *Encyclopedia of Biostatistics*, P. Armitage & T. Colton, eds, Wiley, Chichester, Figure 2, p. 1192 by permission of John Wiley & Sons, Ltd

functions is now available to us. The Distance program uses the methods of [3], in which a parametric 'key' function is selected and, if it fails to provide an adequate fit, polynomial or cosine series adjustments are added until the fit is judged to be satisfactory by one or more criteria.

Often, the perpendicular distances are recorded by distance category, so that each exact distance need not be measured, or data are grouped into distance categories before analysis. Standard likelihood methods for multinomial data are used to fit such 'grouped' data.

Variance and Interval Estimation

The variance of $\widehat{D}$ is well approximated using the formula [5]:

$$\widehat{\mathrm{var}}(\widehat{D}) = \widehat{D}^2 \left[\frac{\widehat{\mathrm{var}}(n)}{n^2} + \frac{\widehat{\mathrm{var}}[\hat{f}(0)]}{[\hat{f}(0)]^2} \right] \tag{4}$$

The variance of n generally is estimated from the sample variance in encounter rates, n_j/l_j, weighted by the line lengths l_j. When $f(0)$ is estimated by

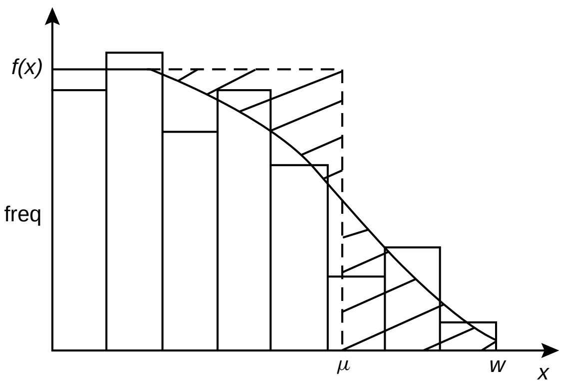

Figure 2 The pdf of perpendicular distances, $f(x)$, plotted on a histogram of perpendicular distance frequencies (scaled so that the total area of histogram bars is unity). The area below the curve is unity by definition. Because the two shaded areas are equal in size, the area of the rectangle, $\mu f(0)$, is also unity. Hence $\mu = 1/f(0)$. Reproduced from Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1998). Distance sampling, in *Encyclopedia of Biostatistics*, P. Armitage & T. Colton, eds, Wiley, Chichester, Figure 3, p. 1192 by permission of John Wiley & Sons, Ltd

maximum likelihood, its variance is estimated from the **information matrix**.

If we assume that $\widehat{D}$ is lognormally distributed, approximate 95% confidence limits are given by $(\widehat{D}/C, \widehat{D}C)$ where

$$C = \exp\{1.96[\widehat{\text{var}}(\ln \widehat{D})]^{0.5}\} \tag{5}$$

with

$$\widehat{\text{var}}(\ln \widehat{D}) = \ln\left[1 + \frac{\widehat{\text{var}}(\widehat{D})}{\widehat{D}^2}\right] \tag{6}$$

Often, **bootstrap resampling** for variance and interval estimation is preferred. Resamples are usually generated by sampling with replacement from the lines, so that independence between the lines is assumed, but independence between detections on the same line is not. If the model selection procedure for the detection function is applied independently to each resample, the bootstrap variance includes a component due to model selection uncertainty.

Cluster Size Estimation

Animals often occur in groups, which we term clusters. These may be flocks of birds, pods of whales, etc. If one animal in a cluster is detected, then it is assumed that the whole cluster is detected, and the distance to the center of the cluster is recorded. Equation (3) then gives an estimate of the density of clusters. To obtain the estimated density of individuals, we must multiply by an estimate of mean cluster size in the population, E(s):

$$\widehat{D} = \frac{n\hat{f}(0)\widehat{\text{E}}(s)}{2L} \tag{7}$$

Probability of detection is often a function of cluster size, so that the sample of detected cluster sizes exhibits size bias (larger clusters are easier to detect and so are over-represented in the sample). In the absence of size bias, we can take $\widehat{\text{E}}(s) = \bar{s}$, the mean size of detected clusters. Several methods exist for estimating E(s) in the presence of size bias [5] (*see* **Size-biased sampling**). One that works well in practice is to regress $\log s$ on $\hat{g}(x)$, the estimated probability of detection at distance x ignoring the effect of cluster size, and then predict $\log s$ when detection is certain, $\hat{g}(x) = 1$, as there can be no size bias in that circumstance. The prediction is back-transformed using a bias adjustment.

Assumptions

The physical setting for line-transect sampling is idealized as follows:

1. N objects are distributed through an area of size A according to some stochastic process with average rate parameter $D = N/A$.
2. Lines, placed according to some randomized design, are surveyed and a sample of n objects is detected.

It is not necessary that the objects be randomly (i.e. Poisson) distributed. Rather, it is critical that the line or point be placed randomly with respect to the local distribution of objects. This ensures that objects in the surveyed strip are uniformly distributed with respect to distance from the line. Thus, if the strip has half-width w, object-to-line distances available for detection are uniformly distributed between zero and w.

Three assumptions are essential for reliable estimation of density using standard line-transect methods:

1. Objects directly on the line are always detected, $g(0) = 1$.
2. Objects are detected at their initial location, prior to any movement in response to the observer.
3. Distances are measured accurately (for ungrouped distance data), or objects are correctly allocated to distance interval (for grouped data).

A fourth assumption is made in many derivations of estimators and variances: whether an object is detected is independent of whether any other object is detected. Point estimates of D are robust to the assumption of independence, and robust variance estimates are obtained by taking the line to be the sampling unit, either by bootstrapping on lines, or by calculating a weighted sample variance of encounter rates by line.

It is also important that the detection function has a 'shoulder'; that is, the probability of detection remains at or close to one initially as distance from the line increases from zero. This is not an assumption, but a property that allows more reliable estimation of object density.

Given the above, the point and interval estimates of D are extremely robust to variation in $g(x)$ due to other factors such as observer, habitat, etc. Large variations in density over the study area are also not a problem, although if areas of differing density can be defined in advance then stratification of survey effort could be used to increase precision.

Point-transect Sampling

In point-transect sampling, an observer visits a number of points, the locations of which are determined by some randomized design. The method is usually (but not exclusively) used for songbird populations, in which typically many species are recorded and most detections are aural. By recording from points, the observer can concentrate on detecting the objects of interest, without having to navigate along a line, and without having to negotiate a randomly positioned line through possibly difficult terrain. The principal disadvantages are that detections made while travelling from one point to the next are not utilized, a problem especially for scarce species, and the method is unsuited to species that are generally detected by flushing them, or to species that typically change their location appreciably over the time period of the count (see below).

Estimation

Detection distances r are measured from the point to each detected object. Suppose the design comprises k points, and distances less than or equal to w are recorded. Then the surveyed area is $a = k\pi w^2$, within which n objects are detected. As for line-transect sampling, denote the probability that an object within the surveyed are a is detected by P_a with estimate $\widehat{P}_a$. Then we estimate object density D by

$$\widehat{D} = \frac{n}{k\pi w^2 \widehat{P}_a} \tag{8}$$

We now define the detection function $g(r)$ to be the probability that an object at distance r from the point is detected, and we again assume that $g(0) = 1$. For line transects, the area of an incremental strip at distance x from the lines is $L\,\mathrm{d}x$, independently of x, which leads to the result that the pdf of distances differs from the detection function only in scale. By contrast, an incremental annulus at distance r from a point has area $2\pi r\,\mathrm{d}r$, proportional to r, so that the pdf of detection distances is $f(r) = 2\pi r g(r)/\nu$, where $\nu = 2\pi \int_0^w r g(r)\,\mathrm{d}r$. The respective shapes of the two functions $g(r)$ and $f(r)$ are illustrated in Figure 3. If we define an effective radius ρ, analogous to the effective strip width of line-transect sampling, then $\nu = \pi\rho^2$ is the effective area surveyed per point (Figure 4). Hence

$$\widehat{D} = \frac{n}{a\widehat{P}_a} = \frac{n}{k\pi w^2 \pi\hat{\rho}^2/\pi w^2} = \frac{n}{k\hat{\nu}} \tag{9}$$

The area of the triangle in Figure 4 is $\rho^2 f'(0)/2$ where $f'(0)$ is the slope of $f(r)$ at $r = 0$. Since this is equal to the area under $f(r)$, which is unity, it follows that $\nu = \pi\rho^2 = 2\pi/f'(0)$, and

$$\widehat{D} = \frac{n\hat{f}'(0)}{2\pi k} \tag{10}$$

We therefore need to model the pdf of detection distances, and evaluate the slope of the fitted function at $r = 0$. The program Distance does this using the same set of models for the detection function as for line-transect sampling.

Variance and Interval Estimation

The methods for variance and interval estimation for line-transect sampling also apply to point transects

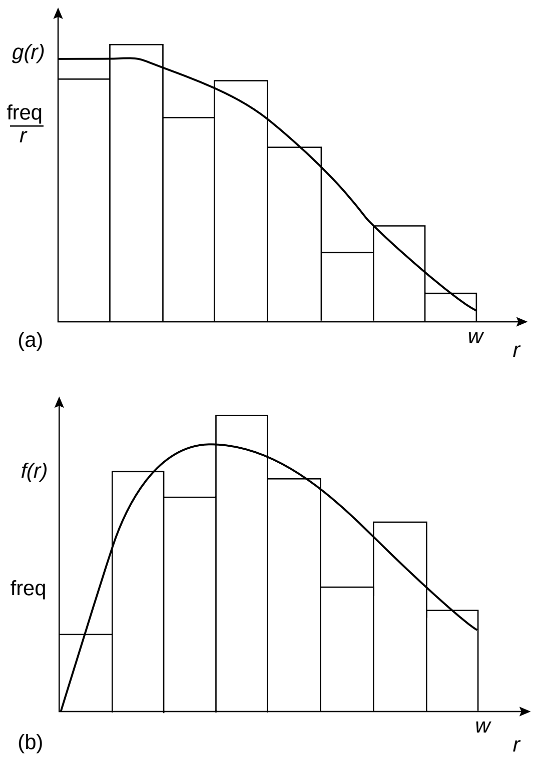

Figure 3 Histograms of detection distances from a point-transect survey. In (a) each histogram frequency has been scaled by dividing by the midpoint of the corresponding group interval. Also shown are the corresponding fits of the detection function [$g(r)$ in (a)] and the pdf of detection distances [$f(r)$ in (b)]. Reproduced from Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1998). Distance sampling, in *Encyclopedia of Biostatistics*, P. Armitage & T. Colton, eds, Wiley, Chichester, Figure 4, p. 1194 by permission of John Wiley & Sons, Ltd

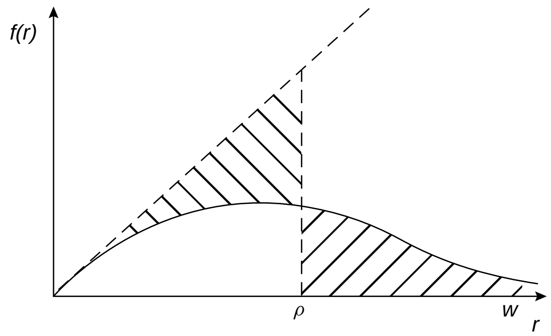

Figure 4 The pdf of detection distances, $f(r)$. The area under the curve is unity by definition. Because the two shaded areas are equal in size, the area of the triangle, $\rho^2 f'(0)/2$, is also unity. Hence $\nu = \pi\rho^2 = 2\pi/f'(0)$. Reproduced from Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1998). Distance sampling, in *Encyclopedia of Biostatistics*, P. Armitage & T. Colton, eds, Wiley, Chichester, Figure 5, p. 1195 by permission of John Wiley & Sons, Ltd

with minor modifications. The variance of n is usually estimated from the sample variance in encounter rates between points. However, point-transect surveys are often designed by defining a series of lines, as if a line-transect survey is to be carried out, then locating a series of points along each line. If the distance between neighboring points on the same line is smaller than the distance between neighboring points on different lines, then the data for all points on the same line should be pooled and the variance of n estimated from the sample variance in encounter rates between lines, weighted by the number of points on each line. Similarly, in this situation, bootstrap variance estimates should be calculated by resampling lines with replacement, rather than individual points.

Assumptions

Assumptions are virtually unchanged from those given for line-transect sampling. As there, the standard analyses are very robust to failure of the assumption of independent detections. Point-transect sampling is, however, more subject to bias than line-transect sampling when objects move through the area around a point. In principle, we try to obtain a snapshot, locating each object at the position it occupied at one instant in time. However, the count is not instantaneous, because the observer needs time to detect all objects close to that point. If, during that time, movement brings new objects into the neighborhood of the point, then object density will be overestimated. To minimize bias, we recommend that the amount of time spent at the point before and after the snapshot instant be fixed in advance, and be as small as possible, given the requirement to detect all objects close to the point.

Related Methods

Trapping webs [5, 22] provide an alternative to traditional capture–recapture sampling for estimating animal density. They represent the only application of distance sampling in which trapping is an integral part, and where data are taken passively. Traps are placed along lines radiating from randomly chosen points; the traditionally used rectangular trapping grid cannot be used as a trapping web. Here detection by an observer is replaced by animals being caught in a trap at a known distance from the center of a trapping web. The trap could be a camera or other similar device. Trapping continues for several occasions and data from either the initial capture of each animal or all captures and recaptures are analyzed. To estimate density over a wider area, several randomly located webs are required.

Cue counting [9] was developed as an alternative to line-transect sampling for estimating whale abundance from sighting surveys. Observers on a ship or aircraft record all sighting cues within a sector ahead of the platform and their distance from the platform. The cue used depends on species, but might be the blow of a whale at the surface. The sighting distances are converted into the estimated number of cues per unit time per unit area using a point-transect modeling framework. The cue rate (usually corresponding to blow rate) is estimated from separate studies, in which individual animals or pods are monitored over a period of time.

Indirect methods are often used when the animals are rare, cryptic or tend to move away before being detected. Instead of counting the animals, the objects counted are something produced by the animals, for example animal dung (e.g. deer dung [11]) or nests (e.g. great apes [12]). To convert object density to animal density one must then estimate two further parameters: object production rate and object disappearance rate, from separate studies.

Related techniques sometimes used by botanists to estimate densities (and sometimes also termed distance sampling) are **nearest neighbor methods** and point-to-nearest object methods [6]. These approaches do not involve modeling the detection function, and so are outside the definition of distance sampling used here.

Current Research

The basic theory of distance sampling is now well established, as are the standard estimation and field methods [5]. Most research is now focused on methods for increasing precision and relaxing the assumptions of the standard methods, and on advanced design issues. There is still much to be done in these areas, so the subject is still a lively one for statistics and ecology.

Generally, probability of detection is a function of many factors other than distance of the object from the line or point. We have considered briefly one other factor, cluster size, because if we do not allow for size bias in detection when objects occur in clusters then our object density estimator may be biased. Other sources of heterogeneity contribute little to bias, provided $g(0) = 1$. Nevertheless, higher precision might be anticipated if additional covariates are recorded and their effects on $g(x)$ modeled. One approach, first used by [14], is to allow covariates to affect the scale of the detection function but not its shape. Marques and Buckland (unpublished) have extended the detection function estimation methods outlined in the section on line-transect sampling above to allow the scale parameter of the key function to be a function of covariates. This approach is implemented in the software Distance.

In some surveys, detection on the trackline is not certain ($g(0) < 1$), perhaps because some animals are underground or under water, or simply hidden by vegetation, when the observer passes. In this case, capture–recapture methods may be combined with distance sampling, through the use of two observation platforms [2]. The platforms might be treated as mutually independent so that, provided that animals detected by both platforms (duplicate detections) can be identified, two-sample capture–recapture methods that incorporate covariates can be used. Bias in such methods is typically large enough to be of concern unless heterogeneity in detectability is well-modeled. However, it is seldom possible to record covariates that reflect this heterogeneity adequately. For example if a whale produces a blow that is particularly visible from one platform, due to light conditions or some other factor in the environment that is difficult to measure, then it will tend to be more visible from the other platform too, and abundance will be underestimated. These problems may be reduced by separating the areas of search for the two platforms, and using one to set up trials for the other. The resulting **binary data** may then be modeled using **logistic regression** [1]. In some studies, the platform that sets up the trials could be provided, for example, by a radio-tagging study, where locations of animals are known, or by an underwater acoustic array (so long as species could be identified accurately). In double-platform methods, Horvitz–Thompson-like estimators are used to estimate density, given the estimated probability of detection for each observation (*see* **Sampling, environmental**).

Spatial modeling of distance sampling data is potentially useful for several reasons: animal density may be related to habitat and environmental variables, potentially increasing precision and improving understanding of factors affecting abundance; abundance may be estimated for any subregion of interest, by integrating under the fitted spatial density surface; and a model-based approach allows data collected from nonrandom surveys (platforms of opportunity) to be used. One approach [7] is to conceptualize the distribution of animals as an inhomogeneous **Poisson process**, in which the detection function represents a thinning process. If, in the case of line-transect sampling, the data are taken to be distances along the transect line between successive detections, this allows us to fit a spatial surface to these data. We can refine this further by conceptualizing the observations as waiting areas, i.e. the effective area surveyed between one detection and the next, where the effective width of the surveyed strip varies according to environmental conditions and observer effort [7, 8].

Geographic information systems (GISs) are now widely available. This makes it possible to implement automated design algorithms that generate survey designs with known properties rapidly and simply. The software Distance has a built-in GIS and implements methods developed by Strindberg (unpublished). It can generate surveys based on a range of point- and line-transect designs, as well as performing simulations to compare the efficiency of different designs and to investigate design properties such as probability of coverage. For complex surveys in which coverage probability is not uniform, but has been calculated analytically or by simulation, Horvitz–Thompson-like estimators can be used to estimate abundance. This avoids the biased estimates that result from standard estimation methods, which assume that coverage probability is even. For example, ship-board surveys typically use continuous zig-zag survey lines, so that costly ship time is not wasted in traveling from one line to the next. For convex survey regions or strata, a design with approximately even coverage probability can be obtained by defining a principal axis for the design and adjusting the angle of the survey line with respect to this axis as the ship progresses through the area. By contrast, fixed-angle or fixed-waypoint zig-zag designs do not give even coverage probability unless the survey region is rectangular (Figures 5 and 6). If the survey region or stratum is not convex, a combination of splitting the region into a number of almost convex sub-regions and placing a convex hull around the sub-regions can be used.

Adaptive sampling [20] (*see* **Adaptive designs**) offers a means of increasing sample size, and hence increasing precision, by concentrating survey effort where most observations occur. Standard adaptive sampling methods can readily be extended to distance sampling surveys [20]. For example, for point transect sampling we can define a grid of points, randomly superimposed on the study region, and randomly or systematically sample from the grid to form the primary sample. When a detection is made at a primary sample point, points from the

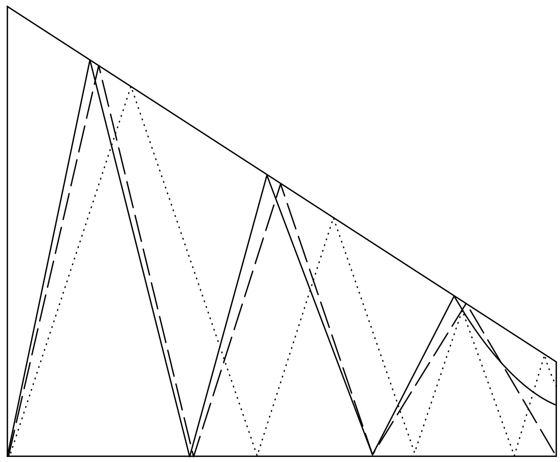

Figure 5 A trapezoidal survey region illustrating three zig-zag designs: equal-angle (dotted line); fixed-waypoint (dashed line); and even-coverage (solid line). The principal axis of the design is parallel to the base of the trapezium in this example, and for the fixed-waypoint design, waypoints are equally spaced with respect to distance along the principal axis, alternating between the top boundary and the base. Reproduced from Buckland, S.T., Thomas, L., Marques, F.F.C., Strindberg, S., Hedley, S.L., Pollard, J.H., Borchers, D.L. & Burt, M.L. (2001). Distance sampling: recent advances and future directions, in *Quantitative Methods for Current Environmental Issues*, V. Barnett, A. El-Shaarawi, C. Anderson & P. Chatwin, eds, Springer-Verlag, New York, Figure 8, by permission of Springer-Verlag

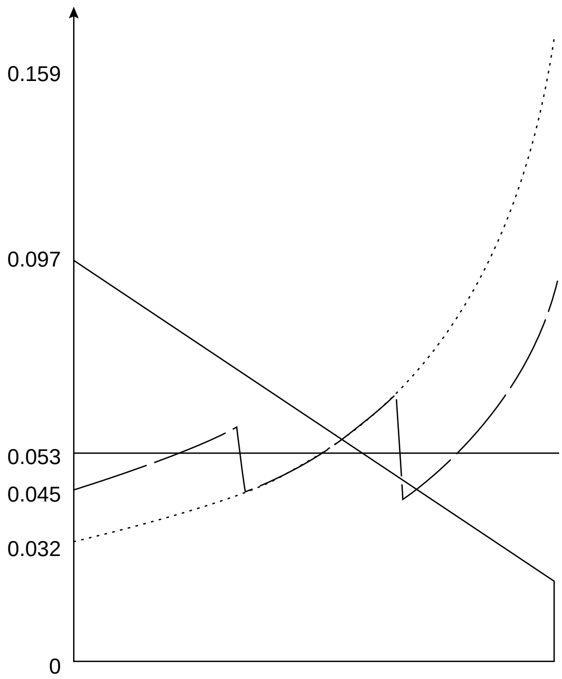

Figure 6 Coverage probability against distance along the principal axis for the three designs of Figure 5. Also shown is the height of the trapezium as a function of distance along the principal axis, which indicates that the fixed-angle design has too low coverage where the study region is wide, and too high where it is narrow. For the fixed-waypoint design, coverage probability changes at each waypoint, and between waypoints varies smoothly in the same manner as the fixed-angle design. Reproduced from Buckland, S.T., Thomas, L., Marques, F.F.C., Strindberg, S., Hedley, S.L., Pollard, J.H., Borchers, D.L. & Burt, M.L. (2001). Distance sampling: recent advances and future directions, in *Quantitative Methods for Current Environmental Issues*, V. Barnett, A. El-Shaarawi, C. Anderson & P. Chatwin, eds, Springer-Verlag, New York, Figure 9, by permission of Springer-Verlag

grid that surround the primary sample point are sampled. If detections are made at these extra points, then further sampling is triggered. A major practical problem of adaptive sampling is that the required survey effort is not known in advance. This is particularly problematic for shipboard surveys, in which the ship is available for a predetermined number of days. A method has been developed [13] that avoids this problem. When additional effort is triggered, the ship changes to a zig-zag (and hence continuous) course, centered on the nominal trackline. The angle of the zig-zag is a function of how far the ship is ahead or behind schedule. Unlike standard adaptive sampling, the method is not design-unbiased, but simulations indicate that the bias is small. An experimental trial on a survey of harbor porpoise in the Gulf of Maine yielded substantially more detections and better precision than did conventional line transect sampling [13].

References

[1] Borchers, D.L., Buckland, S.T., Goedhart, P.W., Clarke, E.D. & Hedley, S.L. (1998). Horvitz–Thompson estimators for double-platform line transect surveys, *Biometrics* **54**, 1221–1237.

[2] Borchers, D.L., Zucchini, W. & Fewster, R.M. (1998). Mark-recapture models for line transect surveys, *Biometrics* **54**, 1207–1220.

[3] Buckland, S.T. (1992). Fitting density functions using polynomials, *Applied Statistics* **41**, 63–76.

[4] Buckland, S.T., Anderson, D.R., Burnham, K.P. & Laake, J.L. (1993). *Distance Sampling: Estimating Abundance of Biological Populations*, Chapman & Hall, London, reprinted (1999) by Research Unit for Wildlife Population Assessment, St Andrews.

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

[6] Diggle, P.J. (1983). *Statistical Analysis of Spatial Point Patterns*, Academic Press, London.

[7] Hedley, S.L. (2000). Modelling Heterogeneity in Cetacean Surveys, PhD dissertation, University of St Andrews.

[8] Hedley, S.L., Buckland, S.T. & Borchers, D.L. (1999). Spatial modelling from line transect data, *Journal of Cetacean Research and Management* **1**, 255–264.

[9] Hiby, A.R. (1985). An approach to estimating population densities of great whales from sighting surveys, *IMA Journal of Mathematics Applied in Medicine and Biology* **2**, 201–220.

[10] Hilborn, R. & Walters, C.J. (1992). *Quantitative Fisheries Stock Assessment: Choice, Dynamics and Uncertainty*, Chapman & Hall, New York.

[11] Marques, F.F.C., Buckland, S.T., Goffin, D., Dixon, C.E., Borchers, D.L., Mayle, B.A. & Peace, A.J. (2001). Estimating deer abundance from line transect surveys of dung: sika deer in southern Scotland, *Journal of Applied Ecology*.

[12] Plumptre, A.J. (2000). Monitoring mammal populations with line transect techniques in African forests, *Journal of Applied Ecology* **37**, 356–368.

[13] Pollard, J.H. & Buckland, S.T. (1997). A Strategy for Adaptive Sampling in Shipboard Line Transect Surveys, Report of the International Whaling Commission **47**, 356–368.

[14] Ramsey, F.L., Wildman, V. & Engbirng, J. (1987). Covariate adjustments to effective area in variable-area wildlife surveys, *Biometrics* **43**, 1–11.
[15] Schwarz, C.J. & Seber, G.A.F. (1999). A review of estimating animal abundance III, *Statistical Science* **14**, 427–456.
[16] Seber, G.A.F. (1982). *The Estimation of Animal Abundance and Related Parameters*, 2nd Edition, Macmillan, New York.
[17] Seber, G.A.F. (1986). A review of estimating animal abundance, *Biometrics* **42**, 267–292.
[18] Seber, G.A.F. (1992). A review of estimating animal abundance II, *International Statistical Review* **60**, 129–166.
[19] Thomas, L., Laake, J.L., Derry, J.F., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Strindberg, S., Hedley, S.L., Burt, M.L., Marques, F.F.C., Pollard, J.H. & Fewster, R.M. (1998). Distance 3.5. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. `http://www.ruwpa.st-and.ac.uk/distance/`.
[20] Thompson, S.K. & Seber, G.A.F. (1996). *Adaptive Sampling*, Wiley, New York.
[21] Thompson, W.L., White, G.C. & Gowan, C. (1998). *Monitoring Vertebrate Populations*, Academic Press, San Diego.
[22] Wilson, K.R. & Anderson, D.R. (1985). Evaluation of a density estimator based on a trapping web and distance sampling theory, *Ecology* **66**, 1185–1194.

(*See also* **Ecological statistics**)

LEN THOMAS, STEPHEN T. BUCKLAND,
KENNETH P. BURNHAM, DAVID R. ANDERSON,
JEFFREY L. LAAKE, DAVID L. BORCHERS &
SAMANTHA STRINDBERG