Design and Analysis of Line Transect Surveys for Primates

Stephen T. Buckland • Andrew J. Plumptre • Len Thomas • Eric A. Rexstad

S. T. Buckland (steve@mcs.st-and.ac.uk)
S. T. Buckland • L. Thomas • E. A. Rexstad
Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, UK
A. J. Plumptre
Wildlife Conservation Society, PO Box 7487, Kampala, Uganda

Short title: Line transect surveys for primates

Corresponding author: S.T. Buckland, CREEM, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, UK, +44-(0)1334-461841, steve@mcs.st-and.ac.uk

Design and Analysis of Line Transect Surveys for Primates

Stephen T. Buckland • Andrew J. Plumptre • Len Thomas • Eric A. Rexstad

Abstract Line transect surveys are widely used for estimating abundance of primate populations. The method relies on a small number of key assumptions, and if these are not met, substantial bias may occur. For a variety of reasons, primate surveys often do not follow what is generally considered to be best practice, either in survey design or in analysis. The design often comprises too few lines (sometimes just one), subjectively placed or placed along trails, so lacks both randomization and adequate replication. Analysis often involves flawed or inefficient models, and often uses biased estimates of the locations of primate groups relative to the line. We outline the standard method, emphasizing the assumptions underlying the approach. We then consider options for when it is difficult or impossible to meet key assumptions. We explore the performance of these options by simulation, focusing particularly on the analysis of primate group sizes, where many of the variations in survey methods have been developed. We also discuss design issues, field methods, analysis, and potential alternative methodologies for when standard line transect sampling cannot deliver reliable abundance estimates.

Keywords distance sampling • estimating primate density • line transect sampling • primate surveys

Introduction

Line transect sampling is a ‘distance sampling’ method (Buckland *et al*., 2001, 2004), widely used for estimating the abundance of wild animal populations. The method relies on a small number of key assumptions, and if these are not met, estimates of abundance can have substantial bias. Line transect surveys of primates often ignore two basic principles of survey design: replication and randomization. In addition, non-standard methods of analysis, lacking any formal assumptions, are often employed, so that it is difficult to know what can be inferred from resulting abundance estimates (Buckland *et al*., in press). We describe line transect methods, and the assumptions on which they rely. We provide guidelines for survey design and field methods to ensure better quality data, and consider some analysis issues particularly relevant to primate surveys. We also discuss possible alternative methods for cases where standard line transect methods are expected to fail. We use a simulation study to assess different analysis approaches when it is problematic to estimate group size and location, and we summarise our conclusions in the discussion.

Line Transect Sampling

In line transect sampling (Buckland *et al*., 2001), lines are placed at random in the survey region, or more commonly, a set of equally-spaced parallel lines is randomly superimposed on the survey region. An observer walks along each line, recording any animals detected within a distance w of the line, together with their shortest distance from the line. In some cases, the distance of detected animals from the observer (so-called ‘radial’ or ‘animal-to-observer’ distance), together with the angle from the line of the detection, are recorded, from which the ‘perpendicular’ distance from the line is calculated later using simple trigonometry. These perpendicular distances are used to estimate a detection function, which is the probability that an animal is detected, as a function of distance from the line. For the basic method, it is assumed that this probability is one at zero distance from the line; that is, animals on the line are seen with certainty. Given an estimate of the detection function, we can estimate the proportion of animals detected within a strip extending a distance *w* from the line on either side. This allows us to estimate animal density, by adjusting encounter rates (i.e. number of animals detected per unit length of line) to allow for animals missed in this strip. Given random placement of an adequate number of lines (or a grid of lines) through the survey region, this density estimate is representative of the whole survey region, allowing abundance within that region to be estimated.

Many animals, including primates, tend to occur in groups, termed 'clusters' in the distance sampling literature. When these groups are well-defined, standard practice is to record the group, its size, and the perpendicular distance from the centre of the group to the line. Estimated density of groups is then multiplied by an estimate of mean group size in the population, to obtain an estimate of animal density.

Survey design and analysis can be carried out using the free software Distance (Thomas *et al*., in press).

Assumptions The key assumptions of the basic approach, with particular reference to surveys of primates that occur in groups, are:

1. Groups whose centres are on or very close to the line are detected with certainty.
2. Groups are detected at their initial location, before any response to the observer. For movement independent of the observer, average speed is slow relative to observer speed.
3. Measurement of distances from the line to the centre of each detected group is accurate.

Two further assumptions should be emphasized as they often do not hold, or can be difficult to satisfy, in primate surveys:

4. There is an adequate sample of randomly-distributed lines, or a grid of lines randomly positioned, in the survey region.

5. Group sizes are accurately recorded, at least for groups on or near the line.

It is important to realise that the group referred to in these assumptions is not necessarily a social unit; it refers to detected animals forming a well-defined group at the time of detection. This might be a group that has temporarily formed, or one part of a larger social unit. In the latter case, if other parts of that unit are also detected, they are recorded as separate groups.

Survey Design

There are two basic principles of survey design that must be met, if reliable inference on population size is to be achieved. The first is randomization: if the positions of transects are not random within the survey region, then there is no guarantee that they pass through areas where densities are representative, and we are unable to extrapolate reliably to the whole survey region. The second principle is replication. Even if the lines are random, if there are too few lines, then by bad luck, they may pass through areas with atypical densities. Also, precision is poorly estimated when replication is inadequate. Buckland *et al*. (2001:232) recommend at least 10-20 lines; we would prefer closer to 20 lines than 10, although 10 lines of adequate length might suffice in areas where group densities vary little. In practice, systematic random designs (i.e. equally-spaced lines with a random start) are usually preferred to designs in which each transect is independently located at random (Buckland *et al*., 2001:233).

The principles of randomization and replication both relate to assumption 4. This assumption is usually not listed explicitly, because it is an aspect of survey design, which is under our control – if we use an appropriate design, we guarantee that the assumption is met. However, non-randomized designs (e.g. transects along trails) with inadequate replication (fewer than 10 lines) are frequent in primate surveys, so we state the assumption explicitly here. If transect lines are not positioned randomly, but instead are located on trails, then the burden of proof falls on the researcher to demonstrate that the selected trails provide a representative sample of the population, and that the distribution of animals within the surveyed strip is uniform with respect to distance from the line.

We show four different strategies for designing a survey (Fig. 1): a straightforward systematic design with a random start (Fig. 1(a)); a design with two strata, with a systematic random sample of lines in each stratum, and higher sampling intensity in one of the strata (Fig. 1(b)); short line segments, spaced so that the separation distance between successive segments on the same line is the same as the distance separating successive lines, which ensures a systematic grid of line segments through the region (Fig. 1(c)); a design based on a systematic grid of points through the survey region, with a circuit (square) of transect lines located around each point (Fig. 1(d)). This last design has the advantage that the observer can start from any location on the circuit (e.g. where access is easiest, such as an intersection of the circuit with a track or trail), and finishes at the same place. However, there is a risk of disturbance of animals on one section of the circuit when the observer is covering another section. If this is thought to be an issue, gaps can be introduced at each corner of the circuit, to separate out the sections.

All four designs have lines that are evenly spread through the survey region (Fig. 1). Survey effort is not clustered in areas of easier access, for example. If there are parts of the survey region that are costly to survey, the region can be divided into strata, with a randomized design in each stratum, but with a lower sampling effort in strata that are more costly to survey (Fig. 1(b)). Such a design allows unbiased estimation of primate abundance, whereas subjective placement of lines related to ease of access may generate substantial bias. Usually, we assume that systematically-spaced lines are in fact independently randomly located in our analyses. Typically, systematic samples yield better precision than simple random samples, especially if there are strong trends in density through the region, but it is more difficult to estimate that precision. Fewster *et al*. (2009) show that a post-stratification strategy can yield good estimates of the systematic sampling variance.

Survey design is discussed in depth by Buckland *et al*. (2001:228-323) and by Strindberg *et al*. (2004). Karanth and Nichols (2002:87-120) discuss survey design and field methods for line transect surveys of tropical forest-dwelling ungulates, which share many of the issue associated with primate surveys. A useful training video is also freely available online (http://www.youtube.com/monitoringtigers).

Field Methods

As noted above, a key assumption is that lines are placed at random, independently of animal locations. This often necessitates cutting of vegetation, which should be minimal when it is required, and ideally carried out at least one week before the line is surveyed, by which time there should be no lasting effect from disturbance. If cutting is sufficient only to allow quiet passage and facilitate data collection, then disturbance of the animals while surveying will be minimized without creating marked highways. Obvious cut transects may affect animal behaviour and distribution, and give easy access to hunters, for example, who would influence detection probability and encounter rates, making them unrepresentative of the larger survey area. Note that it is not essential that observers walk exactly on the transect line – they can leave it, for example to move around small obstacles if this minimizes cutting, so long as detection of animals on the line is still certain. However, the measured distances must be of detected animals from the line, and not from the route taken by the observer, if this differs. Although it is often much less costly to conduct surveys along trails, there can be no guarantee that densities (or temporal trends in density) along trails are representative.

Repeat surveys of the same line within a season to increase sample size is sensible and often essential, but these must not then be analysed as if different lines had been surveyed; the transect should be entered in the software Distance, with effort recorded as line length times the number of times the line was surveyed.

There should be a clear protocol so that fieldworkers can determine what constitutes a group for the purposes of the survey. For example, if animals are separated by more than say 20 m from the originally detected group, the protocol might state that these should be treated as a second group. This might result in one large social unit being recorded as many groups. Any of those groups that is detected and whose centre is located within the survey strip of half-width *w* should be recorded, and their distance from the line measured or estimated.

Distances of group centres from the line should be measured as accurately as possible (assumption 3). This requires that the position of the line is well-defined, so that distances from the line are well-defined. Unless distances are sufficiently small to be measured with a tape without undue disturbance or delay, a laser rangefinder should always be used for primate surveys. It may not always be possible to take a direct measurement, for example because of intervening vegetation, but it is possible to take several measurements to visible objects (e.g. tree trunks) by moving off the transect and summing the distances that form the perpendicular distance you need to measure. The ability to check distances to visible objects by rangefinder is invaluable for improving estimates of distances. We tested five field assistants who regularly census primates and measure perpendicular distances in Uganda in 2008 for their ability to estimate distance by eye (29 obs) and with a laser rangefinder (80 obs), and compared these with the measured distance using a tape (true value). With a rangefinder, 62% of observations were exact (when measured to the nearest metre), 91% were within 1 m, and 97% within 2 m of the true distance up to distances of 40 m. Only 7% of the estimates by eye were exact, with 24% within 1 m and 59% within 2 m. Some estimates by eye were up to 13 m away from the true value. There was also a bias towards underestimating true distance by eye with 68% less than or equal to the true value and 38% greater than or equal to the true value (A. J. Plumptre, unpublished data). This bias would artificially increase estimates of primate density. Field aids such as rangefinders are inexpensive, especially when compared with the costs resulting from poor abundance estimates.

Primates are often in large, dispersed groups, so that it is difficult to estimate distance except for the animals first detected. The problem is made worse if the animals flee from the observer. Given the difficulty in estimating the location of a group centre, it is common practice to record the distance from the line of the first animal detected from a group, and to assume that distance is the distance to the group centre (Struhsaker, 1981; Hassel-Finnegan *et al*., 2008). Of course, the first animal detected tends to be closer to the observer, and hence closer to the line, than the centre of the group (Marshall *et al*., 2008). The measured distances are therefore systematically biased downwards, which artificially inflates density estimates. This source of bias is well-known (e.g. Whitesides *et al*., 1988; Marshall *et al*., 2008), yet the practice persists, and as a consequence, standard line transect sampling is considered to overestimate density in the primate literature (Hassel-Finnegan *et al*., 2008).

Where it is impossible to determine location of group centres with sufficient accuracy, but feasible to estimate distances to each detected animal, then a solution exists (Buckland *et al*., 2001:75-76). The methods in Distance are extremely robust to the assumption that detections are independent events, which is why we do not list this as a key assumption. As a consequence, you can ignore the existence of groups when using line transect sampling to estimate density or abundance. Each individual animal that is detected is recorded, along with its distance from the line. This may compromise ability to measure distances accurately, but approximate estimates of distance, coupled with observer training, is preferable to accurate measurements of the wrong distance. The approach would usually be impractical if a tape is used to measure distances, but is more feasible if a laser rangefinder is used. The task can be made more practical by defining a maximum distance from the line beyond which detected animals are not recorded; this distance would then be used as the truncation distance *w* for analysis.

Adopting this approach, it does not matter if observers fail to detect some animals in a detected group, and assumption 5 above can be dropped. The method assumes, however, that all animals on or very close to the line are detected. If an animal is detected but cannot be accurately located (e.g. because it is well away from the line, and is heard but not seen), it can be excluded from the sample; this just changes the meaning of the detection function slightly, in that it now estimates the probability that an animal is both detected and accurately located, as a function of distance from the line. This does not generate bias in density estimates, provided all those on the line are detected and recorded.

If it is not feasible to record all detected individuals, together with their distances from the line, then it is important to estimate the size and location of detected groups as accurately as possible. In fact, bias in estimates of the size or location of groups well away from the line need not be problematic (see next section), but for those groups on or close to the line, bias should be as small as possible. A field protocol should be developed with these issues in mind. For example if animals do not respond to observers, observing the group for a period of time from different locations on and off the line may allow an accurate assessment of size. If animals do respond, a quick count may be needed, and multiple observers with slightly different vantage points, and a wellrehearsed protocol for coordinating their count (e.g. sketches of animal locations together with arrows to indicate direction of movement), may be effective.

If neither of these strategies is achievable, it may be necessary to estimate mean group size and spread in a separate study from the line transect survey. In this case, the study should be conducted synchronously with the line transect survey. In this way, the mean size and spread of groups in the study should be comparable with the mean size and spread in the population at the time of the line transect survey; variation in size and spread by time of day, season or other factors (Plumptre, 2000) will be controlled for. Problems with this approach are a) it may be difficult to achieve an adequate sample size – at least 10, and preferably nearer 20, especially if group size is very variable; b) if only habituated groups can be monitored in this way, they may not be representative of all groups; and c) it is still necessary to estimate the location relative to the line of groups detected during the line transect survey. To address this last point, it may be necessary to record the distance to the closest animal (whether it is closest to the line or to the observer), and correct either the recorded distances or the effective strip half-width (Whitesides *et al*., 1988). Hassell-Finnegan *et al*. (2008) argued against this strategy because group shape is usually not circular making spread difficult to quantify. To allow for this, you could estimate group spread as the average of several values, recorded using diameters across the group at different orientations. Whitesides *et al*. (1988) defined group spread as the radius of the circle that has the same area as the area occupied by the group; given a means to estimate this area, we can thus estimate group spread.

Data Analysis

Standard line transect analyses are usually conducted using the software Distance (Thomas *et al*., in press). There are three components to estimation when animals occur in groups. The first is encounter rate, which is the number of groups detected per unit length of transect (excluding those whose centres are further from the line than the truncation distance *w*). The second is the estimated proportion detected of those groups whose centres are within distance *w* of the line. The third is the estimate of mean group size in the population. Typically, this is smaller than the mean size of detected groups, because larger groups are more detectable. However, group sizes may be underestimated, as it is difficult to detect all animals within a group, so the mean of recorded group sizes might be biased high or low if used as an estimate of mean group size in the population. The default method of estimating mean group size in software Distance, in which the logarithm of group size is regressed on estimated probability of detection as a function of distance from the line, corrects for both sources of bias, although if there is bias in the recorded size of groups on or near the line, the correction will be partial. Buckland *et al*. (2001) give a detailed account of analysis methods.

A possible departure from the standard analysis is to record distance from the line of the nearest animal only, and then to correct for bias at the analysis stage. Whitesides *et al*. (1988) added half the mean group spread, $\bar{r}$, to the estimated effective half-width of the strip, $\hat{\mu}$. (The effective strip half-width μ is the distance from the line at which as many groups are detected beyond μ as are missed within μ of the line (Buckland *et al*., 2001:3).) This method is unsatisfactory when a group straddles the line. For example if the nearest animal was recorded as on the line, then adding half the mean group spread gives a distance of $\bar{r}$, but a group at this distance from the line is not expected to straddle the line. For the data of Whitesides *et al*. (1988), the mean group spreads were larger than the effective strip half-width for five of the seven species. Thus most groups whose centres were within the effective strip half-width of the line would be expected to straddle the line. A better approach would appear to be to correct individual distances. Suppose, for example, that for a given group, the distance from the line of the nearest animal to the line is recorded, along with whether the group straddled the transect. At the analysis stage, for those groups that do not straddle the transect, half the mean group spread should be added to the recorded distance. For those groups that do straddle the line, we could assign a distance from the line by selecting a value at random from a uniform distribution between zero and half the group spread. If it is assumed that the recorded distance is of the nearest animal to the observer, then the correction to individual distances that Whitesides *et al*. (1988) developed for fitting the hazard-rate model can be adopted: the corrected perpendicular distance is equal to the recorded perpendicular distance multiplied by $1+\bar{r}/AOD$ where $\bar{r}$ is half the mean group spread and AOD is the animal-to-observer distance. This is based on the premise that the distance from the observer to the group centre should on average be the distance from the observer to the nearest animal plus the mean group spread, and simple trigonometry shows that the multiplicative correction for the perpendicular distance is the same as that for the animal-to-observer distance. If there are many recorded perpendicular distances of zero, it may be preferable to record whether a group straddles the line; for those that do not, apply the above correction, while for those that do, take the perpendicular distance to be a random value from the uniform distribution on $(0,\bar{r})$.

Alternative Methods

In some circumstances, it may prove impossible to meet the assumptions of standard line transect sampling to an adequate approximation. Other approaches should then be considered.

If it is feasible to record each individual animal that is detected, together with its distance from the line, but it is thought that some animals on the line are missed, it may be possible to conduct trials by locating animals, perhaps using radio collars, and then sending observers who are ignorant of animals' positions past the animals at a known closest distance of approach. These trials result in binary data, where one corresponds to detection by the observer, and zero corresponds to non-detection. These data may be modelled using logistic regression, with distance from the line and possibly other variables as covariates, from which the probability of detection on the line (i.e. distance=0) may be estimated. If there are any covariates other than distance in the model, this estimate will be a mean value across the trial groups, for which the probability will vary according to the values of the covariates. This estimate and its standard error may then be included as a multiplier in the Distance software, when analysing the line transect survey data. Similarly, if groups rather than individuals are recorded, but some groups on the line may be missed, trials might be set up involving the group rather than an individual animal.

Another distance sampling approach that may work for primates that call is cue counting, as implemented for birds by Buckland (2006). The design comprises a grid of points. An observer stands at each point for a predetermined time, and records any calls heard during this time, together with an estimate of the distance of the calling animal from the point. Cue rate (number of calls per animal per unit time) is estimated in a synchronous survey, to allow conversion from number of calls per unit area per unit time to estimated animal density. Movement of animals independent of the observer does not bias this method, and silent animals above the point need not be detected. Instead, we assume that a call is certain to be heard if the animal is above the point. The disadvantages of this approach are that it can be difficult to estimate distances to calling animals, and it is difficult to ensure that a representative sample of animals is monitored to estimate the cue rate.

If animals can be lured in by playing a call, then lure strip transects may be possible, as implemented in a recent study of cotton-top tamarins *Saguinus oedipus*

(Savage *et al*., in prep.). Observers simultaneously travel along two parallel transects, luring animals from within the strip between the transects. If the lure causes animals to respond by calling, but does not attract them in, a line transect version of this approach might be workable, with just one transect at each location. If several observers are positioned along the line, distances of responding groups from the line may be estimated by triangulation (B. Rawson, pers. comm.). Another possibility is lure point transects (Buckland *et al*., 2006), in which trials are conducted on animals with known location, and from which a model for the detection function is fitted using logistic regression; this function represents the probability that an animal will be detected from the point at which the lure is played. This detection function model is then assumed to hold for the main survey, where a lure is played at each of a number of points systematically spaced through the survey region.

Simulation study

Buckland *et al*. (in press) conducted a simulation study to assess the performance of methods based on animal-to-observer distances. We use the same simulation set-up to assess several analysis options for survey data on primate groups based on standard line transect methods. Details of how the data in simulation set A were generated are given by Buckland *et al*. (in press). The set comprises 100 datasets for each combination of three mean group sizes (3, 10 or 30), three half-group spreads (10 m, 25 m or 50 m), three densities (15, 50 or 150 groups km^{-2}) and two detection functions, making 1800 simulated populations in all, each of which was surveyed once. The two detection functions are given by two parameterizations of the hazard-rate model: $g(y) = 1 - \exp\{-(y/20)^{-2}\}$ and $g(y) = 1 - \exp\{-(y/30)^{-4}\}$. The hazard-rate model was used because it has an underlying model for the detection process (Hayes and Buckland, 1983), whereas other models are simply proposed shapes. This allows animal-to-observer distances to be generated along with perpendicular distances. In the first, detection is certain out to 10 m, and declines to 0.2 at just over 40 m; in the second, detection remains certain to greater distances (around 25m) but then drops more rapidly, again falling to 0.2 at just over 40 m (Buckland *et al*., in press, Fig. 1). If at least one animal in a group was detected, remaining undetected animals were given an enhanced probability of detection, by simulating a second 'pass' with the scale parameter of the detection function increased by 50% (Buckland *et al*., in press). The number of groups detected was typically in the range 60-120 for each population.

Estimating Densities

We used the software Distance to estimate density and mean group size in the population. We set truncation distances *w* (Buckland *et al*., 2001:103-108) so that around 10% of observations were truncated. We considered only two possible detection function models: the half-normal key with cosine adjustments, and the uniform key with cosine adjustments (Buckland *et al*., 2001). We used Akaike's Information Criterion (AIC) to select any adjustment terms, and to select between the two keys. We did not use the true detection function (the hazard-rate model) in analysis, as we wished to assess performance of the method using an approximating model; when analysing real data, we would not know the true model. For each dataset, we implemented the following methods for extracting distances for analysis.

1. Perpendicular distances from the line to each individual animal detected, as if the animals did not occur in groups. Truncation distance *w* was 50 m for all analyses.
2. Perpendicular distances from the line to group centres, where we determined group size and group centre only from detected animals in the group, defining group centre as

the mean of perpendicular distances of detected animals from the line. Truncation distance *w* for group centres was 75 m for mean group size of 3, 100 m for mean group size of 10, and 125 m for mean group size of 30.

3. Perpendicular distances from the line to group centres, where we assume that true group size and location are known. Truncation distance *w* as for method 2.
4. Perpendicular distances from the line to groups, where we take group location as the location of the first animal detected from the group, and we estimate group size as in method 2. Truncation distance *w* as for method 2.
5. Perpendicular distances from the line to groups, where we take group location as the location of the first animal detected from the group, but true group size is known. Truncation distance *w* as for method 2.

We implemented method 3 to act as a gold standard, to compare with methods that can be achieved in practice. Similarly, we implemented method 5 to allow us to separate the effect of recording group location as the location of the first animal detected from the effect arising from underestimating group size.

This is intentionally an idealized study, with a large sample of lines systematically spaced with a random start, and most key assumptions satisfied. If methods perform poorly here, they can certainly be expected to in real studies. We conducted further simulations (simulation set B) that were far more challenging with respect to groups. First, we made group size much more variable. We achieved this by setting $p = 0.5$ instead of 0.75 in the model of Buckland *et al*. (in press) for controlling variability in group size; the further below one that we set *p*, the greater the variability. If the mean group size is 10, this choice of *p* generates about one group in 200 with a size greater than 100. Second, we made detection of individuals in a group independent, omitting the second pass described by Buckland *et al*. (in press). The effect of this is that recorded group sizes tend to be much smaller than true group sizes, especially for groups located further from the line, or with large group spread. This adversely affects methods that use recorded group size rather than true size (methods 1, 2 and 4).

Results

The hazard-rate model has a very flat shoulder for the values of the shape parameter used in this study, which means that the detection probability, assumed to be one at zero distance, remains at one for some distance from the line, before it starts to fall (Fig. 1 of Buckland *et al.*, in press). Neither of the detection function models used for analysis (either a half-normal or a uniform key, with cosine adjustments) share this property. As a consequence, we anticipated modest upward bias in density estimates from this source. Method 3 is based on having perfect knowledge of detected groups, and *a priori*, we expected this method to perform best. It gave consistent estimates of density with good precision, and some positive bias (+8.6%), as anticipated (Table 1). We see also that method 1 (average bias +8.0%), based on analysing individuals, matches the performance of method 3. This is surprising, as method 3 uses additional information not available to method 1: the true number of animals in a detected group, and the mean location of all animals in a detected group. Plumptre and Cox (2006) proposed the use of method 2, but its performance was disappointing, with bias tending to increase with increasing group size and group spread. Bias also differed between the two detection functions. Method 4 showed inconsistent biases. Biases were smaller when the true detection function was given by $g(y) = 1 - \exp\{-(y/20)^{-2}\}$ than when it was given by $g(y) = 1 - \exp\{-(y/30)^{-4}\}$. For large groups, bias was a decreasing function of group spread for the second of these detection functions, but an increasing function for the first.

Method 5 allows us to assess the effect of using distance to the first animal detected in the absence of bias in group size estimation. We see the anticipated positive bias, which increases as group spread increases. All five methods show substantially lower bias on average than the ‘modified Kelker’ methods based on animal-to-observer distances assessed by Buckland *et al*. (in press).

The default group size regression method of Distance, in which log group size is regressed on the estimated detection probability, reduces but does not entirely remove the bias arising from estimating true group sizes by the recorded group sizes (Table 2). The mean of true sizes of detected groups tends to overestimate the mean of groups in the population, as a result of size bias: groups with many animals are more likely to be detected than groups with few animals. The size bias is relatively modest here (ranging between around +1% and +10%). The regression method reduces the bias at the cost of increased variance. However, the contribution of this variance to the overall variance in the density estimate is small, so that the increased variance has minimal impact.

We show results for the scenarios in which group size was highly variable, and for which recorded group sizes were much smaller on average than true sizes (simulation set B, Table 3). We find that methods 1 and 3 maintain their good performance, while that of the other methods deteriorates.

Discussion

If, having detected a group of primates, it is possible to detect all animals in the group, and to estimate the distance of the group from the line, then method 3 can be expected to provide good estimates of density with low bias. The centre of the group can be defined in a way that makes it easier to estimate, provided there is not systematic bias of the type that occurs if distance to the first detected animal is used. Thus if it is not possible to estimate the mean distance from the line of animals in the group, it may be possible to estimate the distance from the line of the left-most animal and of the right-most animal, and at the analysis stage, to calculate the mid-point between them as the distance of the group from the line. If the left-most animal is to the left of the line and the right-most animal to the right (i.e. the group straddles the line), then care must be taken to record one of the distances as negative, before taking the average.

If group size and location cannot be determined with good accuracy, the strategy of recording each detected animal as if it were a separate detection, together with the distance of each detected animal from the line, gives equally good estimates of density. Because detections are not independent in this case, AIC tends to select too many terms for the detection function, and goodness-of-fit tests tend to generate spurious significant results, indicating poor fit when in fact the model is adequate (Buckland *et al*, 2001:76). In the simulation study, we selected the model chosen by AIC, yet despite obvious overfitting in some cases, estimation was still good.

Upward bias in line transect sampling can occur from a source other than the recording of biased measurements of distance. For relatively mobile species, their average speed of non-responsive movement may be similar to that of the observer, especially given that observers often deliberately walk very slowly and quietly, to avoid disturbance, and to increase detection probabilities. If the average speed of movement is under half that of the observer, bias is negligible, but bias increases as average speed of the groups increases (Buckland *et al*., 2001:31).

Responsive movement of animals can compromise data quality. If a large group of animals flushes simultaneously in response to the observer, it can be impossible to record distances to each individual animal that is detected. Failure to record all distances from the line of detected animals away from the line is not a problem, as it is perfectly acceptable to model the combined probability of detecting and accurately positioning the animal. If some groups respond by flushing while others do not, the detection function will vary by group, but pooling robustness (Buckland *et al*., 2004:389-392) means that this does not bias the method. More problematic are groups on or close to the line. If it is possible to estimate approximate distances from the line to the location of each animal prior to flushing, this should be done. If not, then it may be necessary to record the group, rather than each animal in the group. In this case, it will be necessary to estimate the distance of the group centre from the line, and to estimate the size of the group. If neither of these options is feasible, but the group is formed of smaller 'sub-groups' of animals, then it may be possible to record each detected sub-group, together with its size and the distance of its centre from the line. In this case, it is not necessary to detect every sub-group in a group, provided all those on or very close to the line are detected.

If it truly is impractical to estimate the density of primates using distance sampling with direct observations, then consideration could be given to employing so-called 'indirect estimation' techniques. With these methods, distance sampling is used to estimate the density of sign, such as nests or dung, left by the primate population. Additional parameters related to the rate of appearance and disappearance of sign need to be estimated to permit conversion of sign density to animal density. For an excellent review of survey methods for great apes, for which nest surveys are common, see Kühl *et al*. (2008).

Estimation of density for many primate populations represents a great challenge. However, applications to other taxa are often no less challenging. For example, surveys of whales have to address the problems of very low densities of animals across large regions, in an environment where distances are difficult to estimate, with the possibility of responsive movement before detection, and often with no certainty that an animal on the line will be detected. After over 30 years of active development, distance sampling methods can now be reliably applied to many species. We hope that this paper will help researchers to achieve more reliable estimation of the size of primate populations.

ACKNOWLEDGEMENTS

We thank Anne Savage, Ben Rawson, two referees and the editor for their constructive comments, which led to a much-improved paper.

REFERENCES

Buckland, S. T. (2006). Point transect surveys for songbirds: robust methodologies. *The Auk, 123*, 345-357.

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

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

Buckland, S. T., Plumptre, A. J., Thomas, L., & Rexstad, E. A. (in press). Line transect sampling of primates: can animal-to-observer distance methods work? *Int J Primatol.*

Buckland, S. T., Summers, R. W., Borchers, D. L., & Thomas, L. (2006). Point transect sampling with traps or lures. *J App Ecol, 43*, 377-384.

Fewster, R. M., Buckland, S. T., Burnham, K. P., Borchers, D. L., Jupp, P. E., Laake, J. L., & Thomas, L. (2009). Estimating the encounter rate variance in distance sampling. *Biometrics*, *65*, 225-236.

Hassel-Finnegan, H. M., Borries, C., Larney, E., Umponjan, M., & Koenig, A. (2008). How reliable are density estimates for diurnal primates? *Int J Primatol, 29*, 1175-1187.

Hayes, R. J., & Buckland, S. T. (1983). Radial distance models for the line transect method. *Biometrics*, *39*, 29-42.

Karanth, K. U. & Nichols, J. D. (Eds). (2001). *Monitoring Tigers and their Prey*. Bangalore, India: Centre for Wildlife Studies.

Kühl, H., Maisels, F., Ancrenaz, M., & Williamson, E. A. (2008). *Best Practice Guidelines for Surveys and Monitoring of Great Ape Populations*. Gland, Switzerland: IUCN SSC Primate Specialist Group.

Marshall, A. R., Lovett, J. C., & White, P. C. L. (2008). Selection of line-transect methods for estimating the density of group-living animals: lessons from the primates. *Am J Primatol, 70,* 1-11.

Plumptre, A. J. (2000). Monitoring mammal populations with line transect techniques in African forests. *J App Ecol, 37,* 356-368.

Plumptre, A. J., & Cox, D. (2006). Counting primates for conservation: primate surveys in Uganda. *Primates, 47*, 65-73.

Savage, A., Thomas, L., Leighty, K., Soto, L., Causado, J., & Medina, F. (in prep.). Innovative census techniques estimate the cotton-top tamarin (*Saguinus oedipus*) population in Colombia.

Strindberg, S., Buckland, S. T., & Thomas, L. (2004). Design of distance sampling surveys and Geographic Information Systems. In S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers & L. Thomas (Eds), *Advanced Distance Sampling* (pp. 190-228). Oxford: Oxford University Press.

Struhsaker, T. T. (1981). Census methods for estimating densities. In National Research Council, *Techniques for the study of primate population ecology* (pp. 36-80). Washington: National Academy Press.

Thomas, L., Buckland, S. T., Rexstad, E. A., Laake, J. L., Strindberg, S., Hedley, S. L., Bishop, J. R. B., Marques, T. A. & Burnham, K. P. (in press). Distance software: design and analysis of distance sampling surveys for estimating population size. *J App Ecol*.

Whitesides, G. H., Oates, J. F., Green, S. M., & Kluberdanz, R. P. (1988). Estimating primate densities from transects in a west African rain forest: a comparison of techniques. *J Anim Ecol, 57,*345-367.

Figure legend

Fig. 1. Examples of survey design within a region for which an estimate of abundance is required. (a) Systematic random sample of lines that span the survey region. (b) Stratified systematic sample of lines that span the stratum. (c) Systematic sample of line segments. The design comprises the solid line segments. (d) Systematic sample of circuit transects.

Table 1. Percent bias of density estimates for the five methods of estimation for simulation set A. Coefficients of variation of estimates, expressed as percentages, are shown in parentheses. Method 1: analysis of individual detections in Distance. Method 2: analysis of groups in Distance, group size and centre location based on detected animals only. Method 3: analysis of groups in Distance, true group size and centre location known. Method 4: analysis of groups in Distance, group size based on detected animals only, group location taken as location of first detected animal. Method 5: analysis of groups in Distance, true group size known, group location taken as location of first detected animal.

Mean group size		3			10			30	
Half-group spread	10m	25m	50m	10m	25m	50m	10m	25m	50m
True density	15	15	15	50	50	50	150	150	150
$g(y)=1-\exp\{-(y/20)^{-2}\}$:									
Method 1	11	8	7	10	7	4	6	6	6
	(25)	(20)	(16)	(22)	(16)	(13)	(22)	(18)	(13)
Method 2	17	16	24	15	19	31	-10	7	48
	(26)	(21)	(19)	(21)	(18)	(19)	(30)	(27)	(29)
Method 3	9	5	1	6	8	5	2	7	6
	(25)	(20)	(19)	(16)	(12)	(15)	(13)	(14)	(12)
Method 4	15	11	9	14	14	11	-10	-1	11
	(25)	(20)	(20)	(20)	(17)	(18)	(29)	(25)	(23)
Method 5	9	15	39	8	15	37	4	13	30
	(24)	(20)	(19)	(15)	(13)	(14)	(12)	(13)	(13)
$g(y)=1-\exp\{-(y/30)^{-4}\}$:									
Method 1	13	9	9	13	8	9	10	5	2
	(22)	(17)	(16)	(18)	(14)	(14)	(18)	(14)	(13)
Method 2	25	26	35	45	43	54	40	45	55
	(19)	(17)	(18)	(17)	(16)	(21)	(23)	(24)	(23)
Method 3	15	11	7	14	12	14	11	11	10
	(20)	(17)	(19)	(13)	(13)	(14)	(13)	(11)	(12)
Method 4	24	18	15	43	35	26	40	34	22
	(19)	(18)	(19)	(18)	(17)	(18)	(24)	(23)	(22)
Method 5	14	17	37	16	20	45	12	17	36
	(19)	(17)	(19)	(14)	(13)	(14)	(14)	(12)	(12)

Table 2. Percent bias of mean group size estimates, simulation set A. Coefficients of variation of estimates, expressed as percentages, are shown in parentheses. Method a: sample mean of recorded group sizes within *w* of the line. Method b: sample mean of true group sizes, detected groups within *w* of the line only. Method c: estimated mean based on a regression of the log of recorded group sizes on estimated probability of detection. Method d: estimated mean based on a regression of the log of true sizes of detected groups on estimated probability of detection.

Mean group size	3			10			30		
Half-group spread	10m	25m	50m	10m	25m	50m	10m	25m	50m
$g(y) = 1 - \exp\{-(y/20)^{-2}\}$:									
Method a	-12	-15	-22	-35	-37	-43	-52	-52	-56
	(8)	(7)	(6)	(6)	(6)	(6)	(6)	(7)	(6)
Method b	10	11	9	5	4	3	2	2	1
	(6)	(6)	(5)	(4)	(4)	(4)	(3)	(3)	(3)
Method c	9	2	-15	5	0	-16	-16	-11	-6
	(10)	(10)	(9)	(9)	(9)	(8)	(23)	(17)	(16)
Method d	4	3	6	0	0	0	-1	1	1
	(10)	(11)	(9)	(6)	(6)	(6)	(4)	(4)	(4)
$g(y) = 1 - \exp\{-(y/30)^{-4}\}$:									
Method a	-4	-7	-16	-19	-23	-33	-34	-36	-45
	(8)	(7)	(6)	(5)	(6)	(6)	(8)	(6)	(6)
Method b	6	7	9	4	3	3	2	2	1
	(7)	(7)	(5)	(4)	(4)	(4)	(4)	(4)	(3)
Method c	9	5	-9	23	14	-8	23	14	-5
	(10)	(10)	(9)	(10)	(10)	(11)	(16)	(16)	(15)
Method d	0	1	6	-1	-1	-1	-1	-1	-1
	(7)	(10)	(9)	(7)	(6)	(6)	(5)	(5)	(4)

Table 3. Percent bias of density estimates for the five methods of estimation for simulation set B. Coefficients of variation of estimates, expressed as percentages, are shown in parentheses. See caption to Table 1 for methods.

Mean group size	3			10			30		
Half-group spread	10m	25m	50m	10m	25m	50m	10m	25m	50m
True density	15	15	15	50	50	50	150	150	150
$g(y) = 1 - \exp\{-(y/20)^{-2}\}$:									
Method 1	11	5	3	11	1	5	4	11	8
	(31)	(25)	(23)	(39)	(29)	(19)	(29)	(27)	(20)
Method 2	-3	5	13	-22	-6	29	-35	-18	40
	(30)	(29)	(29)	(27)	(35)	(21)	(37)	(43)	(30)
Method 3	6	2	-2	1	-2	3	1	1	3
	(25)	(21)	(24)	(15)	(19)	(14)	(13)	(15)	(13)
Method 4	6	-2	-1	-22	-12	2	-36	-24	4
	(25)	(26)	(30)	(28)	(35)	(24)	(38)	(40)	(34)
Method 5	10	29	67	3	20	77	3	15	59
	(24)	(24)	(30)	(16)	(23)	(18)	(14)	(17)	(19)
$g(y) = 1 - \exp\{-(y/30)^{-4}\}$:									
Method 1	13	15	10	17	7	5	9	4	1
	(26)	(24)	(19)	(39)	(29)	(19)	(32)	(25)	(18)
Method 2	17	25	37	17	25	61	1	22	62
	(24)	(22)	(23)	(25)	(24)	(24)	(26)	(27)	(22)
Method 3	8	9	3	6	5	7	4	5	2
	(23)	(20)	(22)	(19)	(17)	(16)	(16)	(14)	(13)
Method 4	16	15	11	17	16	19	0	11	21
	(24)	(23)	(20)	(28)	(23)	(19)	(28)	(23)	(21)
Method 5	11	29	67	11	26	79	7	22	63
	(23)	(21)	(19)	(20)	(17)	(16)	(15)	(14)	(14)

Fig. 1. Examples of survey design within a region for which an estimate of abundance is required. (a) Systematic random sample of lines that span the survey region. (b) Stratified systematic sample of lines that span the stratum. (c) Systematic sample of line segments. The design comprises the solid line segments. (d) Systematic sample of circuit transects.