

# Line Transect Sampling of Primates: Can Animal-to-Observer Distance Methods Work?

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

Received: 8 July 2009 / Accepted: 15 October 2009 /  
Published online: 4 May 2010  
© Springer Science+Business Media, LLC 2010

**Abstract** Line transect sampling is widely used for estimating abundance of primate populations. Researchers commonly use animal-to-observer distances (AODs) in analysis, in preference to perpendicular distances from the line, which is in marked contrast with standard practice for other applications of line transect sampling. We formalize the mathematical shortcomings of approaches based on AODs, and show that they are likely to give strongly biased estimates of density. We review papers that claim good performance for the method, and explore this performance through simulations. These confirm strong bias in estimates of density using AODs. We conclude that AOD methods are conceptually flawed, and that they cannot in general provide valid estimates of density.

**Keywords** animal-to-observer distances · distance sampling · estimating primate density · Kelker strip · modified Kelker method · 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. In most areas of application, disciplines have standardized their methods and use Distance (Thomas *et al.* 2010). However, methods that in other disciplines are generally considered to be obsolete are still often used and recommended in primatology: the Kelker strip (Kelker 1945) and the modified Kelker method (Struhsaker 1981), which covers a range of methods based on assessing the effective width of the

---

S. T. Buckland (✉) · L. Thomas · E. A. Rexstad  
Centre for Research into Ecological and Environmental Modelling, University of St. Andrews,  
The Observatory, St. Andrews, Fife KY16 9LZ, UK  
e-mail: steve@mcs.st-and.ac.uk

A. J. Plumptre  
Wildlife Conservation Society, PO Box 7487, Kampala, Uganda

searched strip from animal-to-observer distances (AODs). In addition, survey design issues are often ignored, and the precision of abundance estimates is often not quantified, compromising studies designed to compare the performance of different methods.

We here first consider strip transect sampling, and the assumptions under which it is effective. We then explore AOD methods that are conceptually related to strip transect sampling. Plumptre and Cox (2006) noted that such methods have no mathematical basis; here we show that they are based on an erroneous interpretation of the AOD distribution. We review studies that claim good performance of the approach, and assess its performance using simulation.

## Strip Transect and Related Methods

### Standard Strip Transect Sampling

In standard strip transect sampling (Buckland *et al.* 2001), we place lines at random in the survey region, or more commonly, we randomly superimpose a set of equally spaced lines on the survey region. An observer walks along each line, recording all individuals within a distance  $w$  of the line, where  $w$  is the strip half-width. Given random placement of an adequate number 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. In strip transect sampling, we have 2 main options for dealing with groups. The first is to ignore them; all individuals within a sampled strip are counted, without regard to the groups. Thus for groups that extend beyond the survey strip, some individuals are counted and some not. The second option is to count the whole group if its center is within the sampled strip, and not if its center is beyond the strip.

*Assumptions* If groups are ignored, the key assumptions are:

- 1) Individuals that are located within a sampled strip before any response to the observers are certain to be detected and counted.
- 2) Individuals that are located outside the sampled strips are not counted.

If groups are the recording unit, the key assumptions are:

- 1) Groups whose centers are within a sampled strip before any response to the observers are certain to be detected and counted.
- 2) The size of each of these groups is recorded without error.
- 3) Groups whose centers are outside of the sampled strip are not recorded.

In either case, we also assume that there is an adequate sample of randomly distributed strips, or a grid of strips randomly positioned, in the survey region. This assumption is not usually listed, because it is an aspect of survey design, and survey design is under our control: if we use an appropriate design, we guarantee that the assumption is met. However, nonrandomized designs, e.g., transects along

preexisting trails, with inadequate replication ( $<10$  lines; Buckland *et al.* (2001: 232) recommend  $\geq 10$ –20 lines) are frequent in primate surveys, so we state the assumption explicitly here. In practice, we usually prefer systematic random designs to designs in which each transect is independently randomized.

*Problems* If individual animals are counted, we can seldom be sure of detecting all individuals within the sampled strips. Even if this is possible, it can be very difficult to determine whether a detected individual is within the strip, especially for those close to the edge of the strip. If groups are recorded, it is generally difficult to estimate the location of the group center. For these reasons, it has become standard practice among some survey teams to record the position of the group as being at the location of the first-detected individual (Hassel-Finnegan *et al.* 2008; Struhsaker 1981). This individual is more likely to be within the sampled strip than a randomly selected one, and consequently, the strategy leads to positive bias in density estimates. This bias is substantial if average group spread is of similar magnitude to the strip half-width  $w$ .

Line transect sampling (Buckland *et al.* 2001) relaxes the assumption that all groups in the strip are detected, but generates similar bias in density estimates if the location of the group is taken as the location of the first-detected individual. This source of bias is well known (Whitesides *et al.* 1988), yet the practice persists, and as a consequence, standard line transect sampling is often considered to overestimate density in the primate literature (Hassel-Finnegan *et al.* 2008). Buckland *et al.* (2010) discuss how to implement standard line transect methods for primates.

### The Kelker Strip

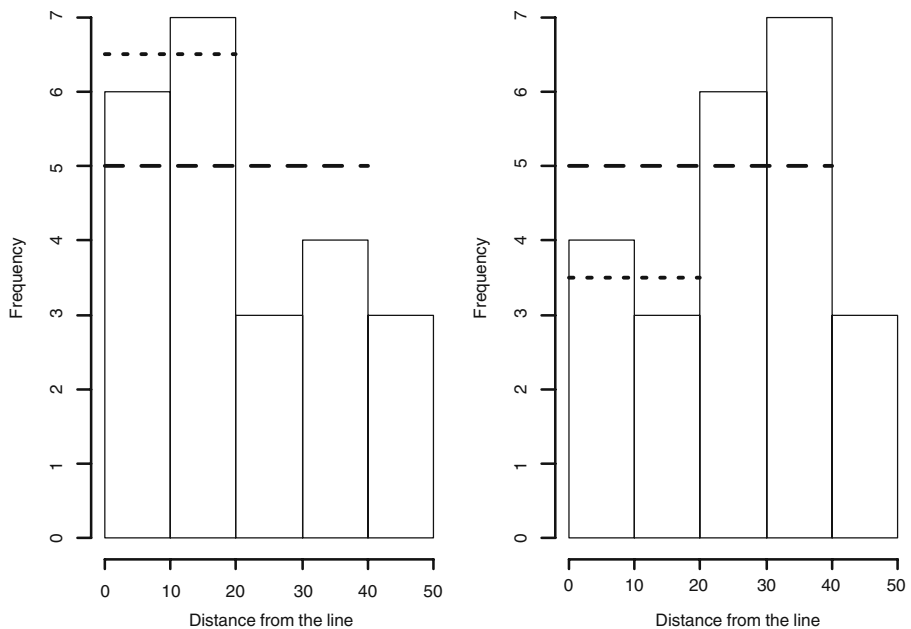
The Kelker strip (Kelker 1945) is a variation on strip transect sampling. Shortest distances of detected individuals from the line (so-called perpendicular distances) are recorded, as for line transect sampling. These distances are placed into intervals, and plotted in a histogram, from which the distance out to which all individuals are detected is assessed. When used for primate surveys, groups are recorded, together with the distance of each group from the line.

*Assumptions* The assumptions of this approach are essentially the same as for strip transect sampling, although we now estimate the strip half-width from the distribution of distances from the line, which requires accurate estimation of distances to the centres of detected groups, including those that are detected beyond the strip.

*Problems* The method shares with strip and line transect sampling the difficulty of identifying the location of group centers. For strip transect sampling, this problem can be avoided if it is possible to record all individuals in the strip, and accurately determine that they are in the strip. However, because the Kelker strip requires distances from the line to be recorded, the distance of each detected individual from the line must be recorded to implement this approach. When groups are recorded, and distances from the line are taken as the distance of the first individual detected

from the line, the method is prone to exactly the same upward bias as strip and line transect sampling.

The method also has problems that line transect sampling does not. First, it is subjective. Identification of the distance up to which all groups are detected can vary between analysts and between different choices of interval cutpoints for the histogram. Second, variance estimates ignore the uncertainty in estimating this distance, and so tend to be underestimates. Third, many observations are discarded because many individuals are detected beyond the distance at which detection can be assumed certain, reducing precision. Fourth, for small sample sizes, the method tends to give biased estimates of abundance. When sample size is small, sampling variation tends to be large, and the choice of cutpoint can be influenced by chance variation in the proportion of detections close to the line. If by chance several groups are detected close to the line, there is a tendency to set the cutpoint too small, which can lead to positive bias (Fig. 1). However, if groups are missed whose centers are inside the selected cutpoint, negative bias will occur. It is difficult to ensure a balance between these biases.



**Fig. 1** Shown here are two data sets, both generated from a detection function with certain detection out to 40 m, and rapidly declining detection probability at larger distances. When by chance there are more detections close to the line, visual inspection of the data leads to selection of a smaller cutpoint for the Kelker method, 20 m in this example. When there are more detections close to 40 m, the cutpoint is likely to be set at 40 m. If we fix the cutpoint in advance, at either 20 m for both analyses or 40 m, we expect unbiased estimates of density, but if we use 20 m for the first analysis and 40 m for the second analysis, we overestimate density on average. Dashed lines: mean count with truncation at 40 m. Dotted lines: mean count with truncation at 20 m.

## AOD Methods

AODs (called ‘radial distances’ in the line transect literature) are often easier to measure than perpendicular distances from the line. In this case, observers generally also record the sighting angle, allowing the perpendicular distance to be calculated (Buckland *et al.* 2001: 5). However, line transect methods that model the AOD rather than the perpendicular distance have a long history (Hayne 1949). They are now seldom used, as the models are not plausible representations of the detection process (Hayes and Buckland 1983). Despite this failing, they were mathematically coherent models. Unfortunately, this is not true of the AOD models used by primatologists. These have their origins in surveys conducted in Kibale Forest, Uganda by Struhsaker (1975). Initially during surveys carried out between 1970 and 1972, he plotted positions of monkey groups on maps and from these he calculated the perpendicular distance of each group from the line. It is not stated what he plotted on the maps: nearest individual, first individual detected, or group center. In later surveys in 1974–76, he estimated the AOD by eye. The density estimates obtained using the perpendicular distances tended to overestimate the known primate density based on knowledge of group sizes and home range of habituated groups (Struhsaker 1975). This appears to be because perpendicular distance from the line to the nearest individual of the group was recorded, as nearly 40% of groups were recorded at zero distance (suggesting that nearly 40% of detected groups straddled the line). He then arbitrarily chose to calculate the mean and maximum AOD across detected groups, where AOD for a group was the distance to the first detected individual, and used these as estimates of the strip half-width for a Kelker strip analysis. The mean AOD tended to overestimate density of primates while the maximum AOD tended to underestimate density. He then defined a maximum reliable AOD: the distance at which the frequency of sightings falls when plotting AOD against number of sightings. This too was used as an estimate of the strip half-width. In each case, he included only AODs less than the estimated half-width in the density estimate (Struhsaker 1981). We can find no published results that show that he compared methods based on perpendicular distances and AODs measured in the field.

*Assumptions* Beyond the assumption that the selected truncation distance results in a complete count of primate groups within that same distance of the line, assumptions are never explicitly stated for the modified Kelker method and its AOD variants. In fact, there is no coherent framework under which the methods can be justified, so that it is not possible to specify a full set of assumptions.

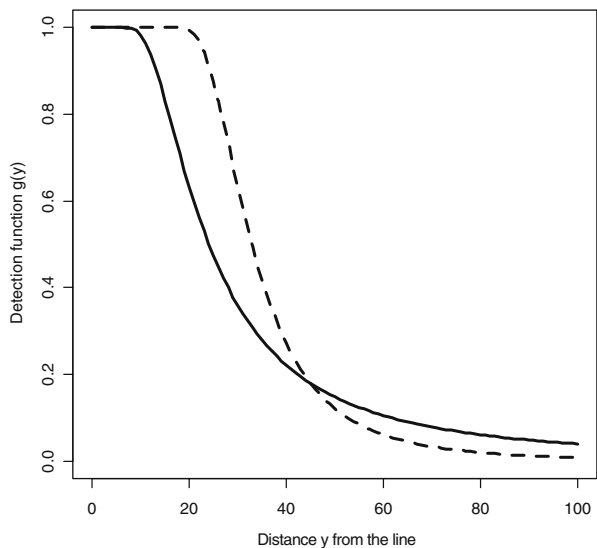
*Problems* The method, like the Kelker strip, is subjective when the ‘maximum reliable AOD’ is used, so that different analysts may select different truncation distances, and estimation is sensitive to the choice. When there is a subjective element in the analysis, and estimation is sensitive to the subjective choice, it is good practice for assessments of the performance of the method based on populations with known density to be performed blind; that is, the analyst should be unaware of the true density when generating estimates.

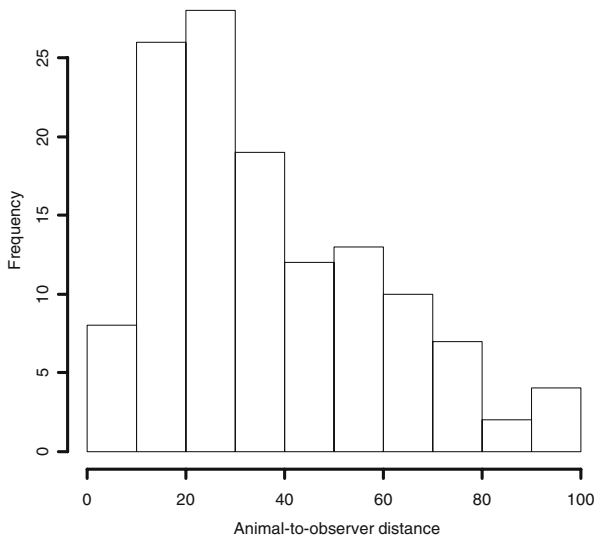
Authors who use AOD methods appear not to estimate the precision of their estimates. Of the papers reviewed below that use the method, only Fashing and

Cords (2000) quoted standard errors, and those were based on repeat surveys of a single line, so do not reflect spatial variation (so-called pseudo-replication, Hurlbert 1984). Variance could be estimated as for strip transect sampling, although this fails to incorporate uncertainty in estimating the strip width.

However, there is a more serious problem with these methods, in that their conceptual framework is erroneous. The methods confuse a probability density function with a detection function. When a histogram is plotted, showing frequencies of detections by distance intervals, then the histogram, if rescaled so that the total area of the histogram bars is unity, provides an empirical estimate of a probability density function: it shows the relative frequencies of detections by distance. By contrast, the detection function is the probability of detecting a group, as a function of distance of that group from the line or, for AOD methods, from the observer. When perpendicular distances from the line are used, the 2 functions have the same shape (Buckland *et al.* 2001: 53), so that the histogram may be used for example to assess the perpendicular distance at which probability of detection starts to fall. However, if AODs are used, this is no longer the case (Buckland *et al.* 2001: 148). The point at which frequencies start to fall does not correspond with the point up to which probability of detection is certain. To illustrate this, we simulated data using the hazard-rate model  $g(y) = 1 - \exp\left\{-\left(y/20\right)^{-2}\right\}$  for the detection function (Fig. 2), where  $g(y)$  is the probability of detecting an animal group whose center is at perpendicular distance  $y$  from the line. Groups had mean size of 3, and half the group spread was 10 m. (Full details are given in the simulation study section below.) We show a histogram of simulated AODs (Fig. 3). The maximum reliable AOD might be taken as 30 m or 40 m, depending on the judgment of the analyst, but detectability starts to fall away at *ca.* 10 m (Fig. 2). The effect is substantial; >60% of groups are undetected at 30 m, and nearly 80% at 40 m. In other words, for these values of the maximum reliable distance, we can expect underestimation of density by >60% (30 m) or nearly 80% (40 m).

**Fig. 2** The detection functions used in the simulation study. Note that these detection functions apply to each individual; the probability that  $\geq 1$  individual of a group will be detected will be larger than shown here, substantially so for large groups. The solid line is  $g(y) = 1 - \exp\left\{-\left(y/20\right)^{-2}\right\}$  and the dashed line is  $g(y) = 1 - \exp\left\{-\left(y/30\right)^{-4}\right\}$ .





**Fig. 3** Histogram of AODs simulated from the hazard-rate model of the detection process,  $g(y) = 1 - \exp\{-(y/20)^{-2}\}$ .

This flaw in the method is self-evident if you consider what is being done. For example, Hassel-Finnegan *et al.* (2008) report in their Fig. 2 an estimate of 55 m, estimated from AODs, up to which detection is considered to be certain for white-handed gibbons (*Hylobates lar carpenteri*). However, of 155 detections, 141 were not detected until they were closer to the observer than 55 m. Indeed, the median AOD is <35 m. If individuals at 55 m were certain to be detected, then AODs of <55 m should not be observed; as the observer approaches an individual, then it will be detected at the certain detection distance of 55 m, if not at a greater distance.

Marshall *et al.* (2008), although acknowledging that the method lacks a mathematical basis, state erroneously: “Because sighting distance is used rather than distance to transect, the pattern of decline with distance is a true detection function.” This is not the case (Fig. 3). Hence their belief that the method should be used when other methods fail lacks credibility.

There is a further inconsistency in the method, when the histogram is used to identify the distance up to which detection is certain. AODs are used to estimate (erroneously) this distance. However, this is then assumed to be the half-width of a strip centered on the line, rather than the radius of a circle centered on the observer. It is also used to truncate detections whose AODs are larger. Suppose we use 55 m as the truncation AOD as in Hassel-Finnegan *et al.* (2008). A group that is detected when still 80 m away, but that is located on the line, is therefore excluded from the count, but its location is right at the center of the strip to which the count supposedly relates.

If data quality were otherwise good, each variation of the method would clearly undercount groups within the sampled strip. If by chance the method does produce a good estimate, it may be a consequence of defining group location with respect to the observer to be the location of the first detected individual of a group: the upward bias generated by this strategy might cancel with the downward bias of the modified Kelker method. There is no assurance that the biases will cancel in general.

## Review of Papers that Have Assessed AOD Methods

Struhsaker (1981) proposed use of the modified Kelker method on the basis that it gave rise to the least biased estimates of density of red colobus monkeys (*Ptilocolobus oustaleti*). However, the reason for overestimation of density in his study is evident from the following quote: "... nearly 40% of the 166 sightings of red colobus were over the census transect and were scored as zero meters from the trail ...". He does not clarify how distances were measured in the field, but as all groups over the census transect were recorded as being at zero distance, we can infer that distance of the nearest individual to the transect was recorded, with predictable overestimation of density; any attempt to salvage density estimates from such poor distance data will inevitably be subjective and *ad hoc*. (Another possibility is that the position of the line was not well defined, so that any individual or group that was close to the line was simply recorded as on the line.)

Defler and Pintor (1985) assessed the performance of the 3 modified Kelker methods based on mean AOD, maximum reliable AOD and maximum AOD, against known densities of 3 species: red howlers (*Alouatta seniculus*), collared titi monkeys (*Callicebus torquatus*), and brown capuchins (*Cebus apella*) in Colombia. Their design comprised a single, nonrandom transect; lack of randomization means that we cannot be confident that the density along the line is representative, and lack of replication means that there is no basis for assessing precision of estimates. Their results were very mixed. Mean AOD led to estimated biases of +12%, -30%, and +538% for the 3 species. Maximum reliable AOD gave estimated biases of +12%, -12%, and +668%. Maximum AOD resulted in estimated biases of -25%, -56%, and +226%.

Chapman *et al.* (1988) used just 5 transects, subjectively placed. Their known populations comprised just a single group of each of 2 species—white-headed capuchins (*Cebus capucinus*) and mantled howlers (*Alouatta palliata*)—in Costa Rica. They used six different methods of measuring distances: "the mean, maximum and reliable perpendicular distance from the transect to the animal first sighted and the mean, maximum and reliable distance from the observer to the animal." Thus all 6 methods were prone to bias by assuming that the first individual sighted was at the center of the group. The authors did not quantify the precision of their estimates, and did not define what a "reliable" distance is (there is not a unique definition of it in the literature). Their estimates show poor performance of all methods, with no clear winner, yet they come down heavily in favor of methods based on AODs on the grounds that "sightings that occur directly over the transect or at a steep angle to it, are likely to cause bias." They do not clarify why. They also claim that the ability of the observer to estimate perpendicular distance will be limited when the terrain is rough, which in our view is not a compelling reason for using the wrong distance. Analyses presented in the paper do not in fact support the use of AOD methods; rather, misunderstanding of the methods has resulted in their recommendation to use it.

Brugiere and Fleury (2000) did not attempt standard line transect analysis because they had only 23 detections, yet they considered this sample size to be adequate for 10 other methods of analysis. Their known population comprised just 3 groups of a single species—black colobus (*Colobus satanas*)—in Gabon. Their design com-



prised just 2 transects, subjectively positioned, and they did not quantify precision of estimates. They used strip counts, with strip half-width preset at 3 values: 60, 80, and 100 m; strip counts, with half-width estimated from the data, using maximum, mean or “maximum reliable” perpendicular distance; the modified Kelker method, using maximum, mean, or “maximum reliable” AOD; and estimation of the effective strip half-width from a histogram, coupled with adding half the group spread to this distance (Whitesides *et al.* 1988). Uncertainty over true density complicated assessment of the methods, and they drew no firm conclusions on which method was best.

Fashing and Cords (2000) analyzed data on 2 species—black-and-white colobus (*Colobus guereza*) and blue monkeys (*Cercopithecus mitis*)—in Kenya. They estimated true densities based on home range data primarily on 5 groups and 3 groups, respectively, although data from additional groups were also used. The design was of a single nonrandom transect, placed along trails. They estimated precision from variation in repeat runs along the same transect. They estimated transect width using 1) the maximum reliable AOD; 2) the maximum reliable perpendicular distance; and 3) the maximum reliable perpendicular distance with the addition of half the group spread, as recommended by Whitesides *et al.* (1988). They also used the shape-restricted estimator of Johnson and Routledge (1985), a type of perpendicular distance detection function estimator that is seldom used. For both species, the method based on perpendicular distances, together with the half-group spread correction, gave estimates closest to the true density. The shape-restricted estimator performed particularly poorly.

Hassel-Finnegan *et al.* (2008) used just a single transect to estimate densities of 2 species—white-handed gibbons and Phayre’s leaf monkeys (*Trachypithecus phayrei crepusculus*)—in Thailand. They assessed true density largely on the basis of a single group for each species, and did not quantify precision of estimates. They used both the Kelker strip and the modified Kelker method, with truncation distance estimated as the point at which frequencies in the respective histograms of perpendicular distances and of AODs started to fall. In addition, they used Distance (Thomas *et al.* 2010) to perform a standard line transect analysis of perpendicular distances. All detection distances, whether perpendicular distances or AODs, were measured to the first individual detected of the group. For both species, Distance and the Kelker strip gave rise to overestimates of the true density, while the modified Kelker method gave estimates very slightly under the true density. However, given the lack of replication (a single line, and a single group of each species), it seems that little can be inferred from these results. Hassel-Finnegan *et al.* (2008) quote the papers of Chapman *et al.* (1988) and Fashing and Cords (2000) to support their contention that analyses based on AODs closely match true densities, while those based on perpendicular distances overestimate. However, the results of neither paper support this conclusion.

All of the aforementioned comparisons are based on studies where true density is established by studying a small number of habituated groups, and estimating the size of their home range. There are several reasons why there might be bias in these “true” densities. For example, home ranges of groups may partially overlap, and because the transects in these studies are positioned subjectively, they may sample parts of the home range that are favored or avoided by the habituated group, leading

to a mismatch in the densities being estimated by the 2 approaches. This is exacerbated when the sampled strip(s) extend beyond the home range(s) of the habituated group(s), into other home ranges. Further, lone males are not included in densities obtained from home range studies, so that density might be expected to be lower as assessed by this method than that obtained by appropriate application of line transect sampling methods. In the case of a population of gray-cheeked mangabeys (*Lophocebus albigena*) in Uganda (Olupot and Waser 2005), Olupot (*pers. comm.*) estimates that *ca.* 30% of males are solitary, corresponding to *ca.* 8% of the total population.

Mitani *et al.* (2000) compared censuses of 7 species—black-and-white colobus, Pennant's red colobus (*Procolobus pennantii*), baboons (*Papio anubis*), blue monkeys, gray-cheeked mangabeys, red-tailed monkeys (*Cercopithecus ascanius*), and chimpanzees (*Pan troglodytes*) made along almost the same census route between years at Ngogo in Kibale Forest. The 3 authors measured AOD separately for their respective census periods (1975–76 and 1996 Struhsaker; 1997–98 Lwanga; and 1996 Mitani). They also used a single transect, which formed the shape of a square route. The authors found great variation in the estimation of AOD between the 3 of them, showing that each observer would estimate a very different maximum reliable sighting distance and that the shape of the sighting distributions differed significantly. They therefore could not use the modified Kelker method to compare densities between years and resorted to comparing encounter rates of primate groups per kilometer walked. Variation between observers with the modified Kelker method eliminates any possibility of comparison unlike standard perpendicular distance methods where a probability of detection can be computed for each observer to allow comparisons to be made (Marques *et al.* 2007).

Marshall *et al.* (2008) review 4 methods: strip transects, the modified Kelker method, and 2 methods based on perpendicular distances: distance from the line of the group center and distance from the line of the center of measurable individuals. They note that methods based on truncating a substantial proportion of data need larger sample sizes for comparable precision than methods that do not. They also note that the assumptions of the modified Kelker method are “unknown.” However, they still advocate its use for when expertise is unavailable to apply other methods, or when the data (location of group centers, or reliable estimates of mean group spread) or assumptions (certain detection on the line, accurate measurement to individuals) required by other methods are unachievable. They justify this recommendation on the grounds that the method consistently performs well in field trials, a claim that does not stand up to close scrutiny. The recommendation presupposes that the *ad hoc* method will produce useful estimates when the other methods do not. The observer effects found by Mitani *et al.* (2000), for example, suggest that this is unlikely.

## Simulation Study

### Simulating Populations and Samples

To assess how different methods perform, we simulated data from populations of known density. This is intentionally an idealized study, with a large sample of

lines systematically spaced with a random start, with certain detection of individuals on the line, no responsive movement, and no measurement error in distances. If methods perform poorly here, they can certainly be expected to in real studies. For simplicity, we assumed a rectangular survey region, 20 km long and 5 km wide. We placed 25 transects in the region, each 5 km long, spaced 800 m apart.

True number of groups in the survey region was 500, randomly spread through the region with a uniform density. Mean group size was 3, 10, or 30 individuals, so that total population size was 1500, 5000, or 15000, corresponding to 15, 50, or 150 individuals/km<sup>2</sup>. We assigned the individuals to the 500 groups by first assigning a single individual to each group. We then generated a random number for each remaining animal from a continuous uniform distribution on  $(0, 500^p)$ , with  $p=0.75$ . We raised this number to the power  $1/p$ , and rounded up to the next integer; the resulting value defined the group to which the individual was assigned. This ensures greater variation in group size than would occur if all groups had the same expected size (corresponding to  $p=1$ ), but the expectation of mean group size was 3, 10, or 30, as required. We assigned the position of each individual in a group at random within a circle of radius  $\rho$ , centred at the assigned group location, with  $\rho=10, 25$ , and 50 m. All group centers fell within the survey region, but single individuals could be assigned a location outside the survey region. To avoid the complication of partial sampling of groups straddling the boundary, we extended sampling into a buffer zone, to allow the whole group to be sampled. We did not count effort, i.e. length of transect, in the buffer zone; this does not create bias because the additional sightings compensate for the “missing” sightings that would have occurred had groups been simulated whose centers were outside the study region, but that straddled the boundary.

Hayes and Buckland (1983) developed a hazard-rate model of the detection process. Their model is useful here to simulate the detection process as the observer approaches a group of individuals. In this study, we initially simulated whether or not an animal was detected independently of other individuals in a group. We assumed a hazard function of the form  $k(r) = ar^{-b}$ , with  $b=3$  and  $b=5$ , where  $r$  is distance between the animal and the observer. If the observer has not yet detected an individual at distance  $r$ , then  $k(r)dx$  is the probability that the individual is detected as the observer advances a small distance  $dx$  along the line. Given the above form for  $k(r)$ , we can derive the detection function  $g(y)$ , which is the probability that an individual at distance  $y$  from the line is detected:  $g(y) = 1 - \exp\{-(y/c)^{-(b-1)}\}$ . We chose  $(c=20, b=3)$ , for which  $a=400$ , and  $(c=30, b=5)$ , for which  $a=1215000$ . These 2 detection functions are shown (Fig. 2). To mimic the enhanced probability of detecting individuals in a group once the first individual of the group has been detected, we identified all groups for which  $\geq 1$  individual was detected, and simulated a second “pass” to search for undetected animals in the group, again using a hazard-rate detection function, but with the scale parameter  $c$  increased by 50% ( $c=30$  for scenarios with  $b=3$ , and  $c=45$  when  $b=5$ ).

For a detected individual, we recorded both the AOD at the time of first detection and the perpendicular distance from the line. We did not record individuals  $>150$  m from the line. Sample sizes were typically in the range 60–120.

## Estimating Densities

For each combination of mean group size, group spread, true density, and detection function, we simulated 100 populations, and surveyed each once. We applied the following analysis methods.

- 1) The modified Kelker method, based on mean AOD, where AOD for each detected group is the distance of the first detected individual from the group. We took the mean AOD as an estimate of the strip half-width, and the mean of recorded sizes of detected groups as an estimate of mean group size in the population.
- 2) The modified Kelker method, based on maximum AOD, where AOD for each detected group is the distance of the first detected individual from the group. We took the maximum AOD as an estimate of the strip half-width, and the mean of recorded sizes of detected groups as an estimate of mean group size in the population.
- 3) The modified Kelker method, based on maximum reliable AOD, where AOD for each detected group is the distance of the first detected individual from the group. We grouped AODs into 10-m bins, and estimated the half-width of the strip by starting at the bin closest to the line (0–10 m), and identifying the first bin for which the count was at most one half of the mean count for preceding bins. If, for example, the mean count in the first four bins was 10.5, and the count for bin 5 (40–50 m) was 5, then the strip half-width was taken to be 40 m, and we excluded detections at a greater distance from the analysis. We estimated mean group size in the population by the mean of recorded sizes of detected groups.

## Results

All 3 modified Kelker methods have strong negative bias for all scenarios (Table I). The bias is consistent across different group sizes and spreads, but differs markedly by detection function. This finding is consistent with the finding by Mitani *et al.* (2000), that density estimates were not comparable across observers. The bias is especially large for method 1, the maximum AOD method (–90.7% and –75.7% for the 2 detection functions). For method 2, bias was –42.3% for the first detection function and –25.1% for the second. The corresponding values for method 3 were –43.9% and –28.2%.

These biases are not fully explained by bias in recorded group sizes (Table II). Interestingly, although bias in recorded group size increases both with mean group size and with group spread, for methods 2 and 3, bias in density estimates within a method and detection function is largely independent of mean group size and group spread. However, as the bias is not consistent across different detection functions, it suggests that neither method gives a reliable estimate of relative density.

The bias is also not attributable to recording distances to the first detected individual, rather than to the group center. Using measurements to group centres, AODs would increase, resulting in larger estimated strip widths, and reduced densities, so that bias would be even larger.

**Table I** Mean (standard deviation in parentheses) of density estimates for the 3 methods of estimation

Mean group size	3			10			30		
	10 m	25 m	50 m	10 m	25 m	50 m	10 m	25 m	50 m
Half-group spread	10 m	25 m	50 m	10 m	25 m	50 m	10 m	25 m	50 m
True density	15	15	15	50	50	50	150	150	150
$g(y) = 1 - \exp\{-(y/20)^{-2}\}$ :									
Maximum AOD	1.6 (0.7)	1.6 (0.7)	1.6 (0.6)	4.2 (0.6)	4.2 (0.6)	4.3 (0.7)	12.8 (1.3)	13.2 (1.3)	13.1 (1.1)
Mean AOD	8.6 (1.9)	8.5 (1.7)	8.3 (1.5)	29.6 (5.2)	29.2 (4.0)	29.0 (4.0)	85.5 (13.4)	88.2 (12.7)	87.4 (11.9)
Max reliable AOD	8.8 (2.1)	8.4 (1.7)	8.4 (1.6)	28.7 (5.3)	28.2 (4.3)	27.6 (4.2)	80.9 (14.0)	83.4 (12.9)	83.7 (11.2)
$g(y) = 1 - \exp\{-(y/30)^{-4}\}$ :									
Maximum AOD	4.5 (1.3)	4.5 (1.5)	4.7 (1.2)	12.2 (3.5)	11.7 (3.4)	11.8 (4.0)	28.4 (7.1)	28.7 (7.6)	26.4 (7.4)
Mean AOD	10.7 (2.1)	10.8 (1.7)	11.0 (1.9)	39.2 (6.4)	38.3 (5.3)	38.4 (6.0)	113.6 (18.1)	114.1 (17.7)	110.5 (15.1)
Max reliable AOD	10.5 (1.9)	10.3 (1.6)	10.5 (1.7)	37.4 (5.2)	36.4 (4.4)	36.9 (4.9)	108.7 (15.3)	109.2 (15.1)	106.1 (14.2)

Because the methods have no coherent mathematical framework, it is not possible to identify the causes of bias, as there are no coherent assumptions that we can assess.

### Discussion

In our simulation study, we found serious biases with estimators based on AODs. Hassel-Finnegan *et al.* (2008) criticize conventional line transect sampling as

**Table II** Estimates of mean group size: sample mean (standard error in parentheses) of recorded group sizes within  $w$  of the line

Mean group size	3			10			30		
	10 m	25 m	50 m	10 m	25 m	50 m	10 m	25 m	50 m
Half-group spread	10 m	25 m	50 m	10 m	25 m	50 m	10 m	25 m	50 m
$g(y) = 1 - \exp\{-(y/20)^{-2}\}$ :									
	2.65 (0.20)	2.56 (0.19)	2.33 (0.14)	6.46 (0.39)	6.27 (0.36)	5.68 (0.33)	14.28 (0.89)	14.31 (0.97)	13.29 (0.82)
$g(y) = 1 - \exp\{-(y/30)^{-4}\}$ :									
	2.89 (0.22)	2.79 (0.20)	2.52 (0.16)	8.14 (0.44)	7.67 (0.45)	6.66 (0.42)	19.77 (1.55)	19.08 (1.19)	16.60 (0.92)

implemented in Distance because a large number of detections is needed for reliable analysis. They fail to note that this is even more true of the Kelker strip, for which many of the observations are discarded (Marshall *et al.* 2008). With inadequate sample sizes, choice of truncation distance is more subjective, uncertain, and influential.

For methods based on selecting a single individual, and using the distance to it as the distance to the group center, there is some ambiguity in the literature about whether the selected individual is the first one detected or the closest individual. In general, these are not the same individual. In our simulations, we assumed that it is the first individual detected. Struhsaker (1981) recorded 40% of detected groups as being on the line, which suggests that he used the distance of the closest individual to the line. Alternatively, if his transect was along trails, it may be that individuals directly above the line were the first to be detected, because they were more visible.

We conclude that AOD methods as used by primatologists are conceptually flawed; one should not treat the resulting estimates as estimates of absolute density. Whether they are acceptable estimates of relative density depends on many factors. Estimates are unlikely to be comparable across different observers (Mitani *et al.* 2000) or habitats, for example. Estimating primate abundance is often difficult compared with many other taxa, as the individuals often reside in hard-to-access, low-visibility areas and are often clustered, cryptic, and highly mobile. Nevertheless, more reliable estimates of abundance are potentially possible by combining good survey design with better field and analytic methods (Buckland *et al.* 2010).

**Acknowledgments** We thank Anne Savage, the editor, and 2 reviewers for their constructive comments on earlier drafts.

## References

- Brugiere, D., & Fleury, M.-C. (2000). Estimating primate densities using home range and line transect methods: a comparative test with the black colobus monkey *Colobus satanus*. *Primates*, *41*, 373–382.
- 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. (2010). Design and analysis of line transect surveys for primates. In Press.
- Chapman, C., Fedigan, L. M., & Fedigan, L. (1988). A comparison of transect methods of estimating population densities of Costa Rican primates. *Brenesia*, *30*, 67–80.
- Defler, T. R., & Pintor, D. (1985). Censusing primates by transect in a forest of known primate density. *International Journal of Primatology*, *6*, 243–259.
- Fashing, P. J., & Cords, M. (2000). Diurnal primate densities and biomass in the Kakamega Forest: an evaluation of census methods and a comparison with other forests. *American Journal of Primatology*, *50*, 139–152.
- Hassel-Finnegan, H. M., Borries, C., Larney, E., Umponjan, M., & Koenig, A. (2008). How reliable are density estimates for diurnal primates? *International Journal of Primatology*, *29*, 1175–1187.
- Hayes, R. J., & Buckland, S. T. (1983). Radial distance models for the line transect method. *Biometrics*, *39*, 29–42.
- Hayne, D. W. (1949). An examination of the strip census method for estimating animal populations. *Journal of Wildlife Management*, *13*, 145–157.
- Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, *54*, 187–211.

- Johnson, E. G., & Routledge, R. D. (1985). The line transect method: a nonparametric estimator based on shape restrictions. *Biometrics*, *41*, 669–679.
- Kelker, G. H. (1945). *Measurement and interpretation of forces that determine populations of managed deer*. PhD thesis, University of Michigan.
- Marques, T. A., Thomas, L., Fancy, S. G., & Buckland, S. T. (2007). Improving estimates of bird density using multiple covariate distance sampling. *The Auk*, *124*, 1229–1243.
- 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. *American Journal of Primatology*, *70*, 1–11.
- Mitani, J. C., Struhsaker, T. T., & Lwanga, J. S. (2000). Primate community dynamics in old growth forest over 23.5 years at Ngogo, Kibale National Park, Uganda: implications for conservation and census methods. *International Journal of Primatology*, *21*, 269–286.
- Olupot, W., & Waser, P. M. (2005). Patterns of male residency and intergroup transfer in gray-cheeked mangabeys (*Lophocebus albigena*). *American Journal of Primatology*, *66*, 331–349.
- Plumptre, A. J., & Cox, D. (2006). Counting primates for conservation: primate surveys in Uganda. *Primates*, *47*, 65–73.
- Struhsaker, T. T. (1975). *The red colobus monkey*. Chicago: University of Chicago 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, DC: National Academy Press.
- Thomas, L., Buckland, S. T., Rexstad, E. R., Laake, J. L., Strindberg, S., Hedley, S. L., et al. (2010). Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology*, *47*, 5–14.
- Whitesides, G. H., Oates, J. F., Green, S. M., & Kluberanz, R. P. (1988). Estimating primate densities from transects in a west African rain forest: a comparison of techniques. *Journal of Animal Ecology*, *57*, 345–367.