Aerial survey intensity as a determinant of estimates of African elephant population sizes and trends

Sam M. Ferreira & Rudi J. van Aarde

Conservation Ecology Research Unit, Department of Zoology and Entomology, University of Pretoria, Hatfield, 0028 South Africa

Received 13 October 2008. Accepted 16 October 2009

Population estimates based on strip transect aerial surveys of large herbivores can aid management decisions if estimates are accurate and precise. Because sampling intensities may influence precision, we simulated survey intensities to determine which could yield accurate and precise population estimates and detect population changes for several African elephant (Loxodonta africana) populations. Simulated surveys of hypothetical elephant populations had to cover 5–20% of a study area to yield accurate estimates, but this depended on how density and distribution varied. Precise estimates, however, needed survey intensities greater than 50%. In addition, the detection of typical rates of population change needed greater sampling intensities than those currently used for most elephant surveys. Survey intensity plays an important role in estimating the accuracy and precision of population estimates and the detection of population trends. Population managers should consider existing information on population density, distribution, and survey intensity to design aerial surveys that will yield both accurate and precise estimates.

Key words: estimates, percentage confidence limit, precision, survey effort, trend detection.

INTRODUCTION

Data from aerial surveys are often used to calculate population growth rates and make management decisions for vertebrate herbivores (e.g. Walter & Hone 2003). These surveys are designed to either obtain total counts or sample counts. Total counts approximate real numbers and generally do not have estimates of variance and thus precision (i.e. likely spreads of estimates, usually reflected by standard errors – Everitt 2002). The variability in population estimates obtained through replicated sample counts, however, can be used to express accuracy (i.e. measure of agreement of an estimate to real values – Everitt 2002), which includes a bias and precision component, where accuracy = bias² + precision (Thompson 1992). Surveyors often use Jolly’s Method II (Jolly 1969) to analyse aerial survey data and to estimate population size and the standard error thereof as a measure of the precision of the estimate. They seldom, however, know the real population size and thus cannot tell whether an estimate is similar to a real value, i.e. unbiased. Accuracy can then only be evaluated through simulation approaches that separate bias and precision components (e.g. Khaemba et al. 2001).

The intensity of surveys that is required to obtain estimates that are precise enough (e.g. acceptable variance, standard error or confidence limits) to confidently derive population growth rates is often not known. Survey intensity affects estimates and their precision because it directly relates to sample size – for aerial surveys sample size increases when survey intensity increases (Walsh et al. 2001). Aside from survey intensity, variables such as the interval between surveys, the number of surveys included in the time series, the length of time series, and the magnitude of change to be detected, influence the ability to confidently detect numerical trends (Gerrodette 1987). Therefore, survey intensities should optimize the trade-offs among these variables to improve population estimates, conclusions based on population growth rates and the confidence in management recommendations. Population variables such as the distribution of individuals and their densities also affect the precision of population estimates (Caughley 1974; Gerrodette 1987). Distribution and density therefore may dictate the intensity at which surveys need to be conducted to obtain acceptable levels of precision of population estimates (e.g. Pennington & Volstad 1994; Beehler et al. 1995; Naegel & Gomez-Humaran 1998). This is because clumped distributions and low population densities may lead to several zeros in sample data that accentu-
ate statistical outlier effects and hence reduce the precision of an estimate (Barnett & Lewis 1984).

We evaluate density and distribution as variables that may have to be considered when defining survey intensity (the percentage area that an aerial survey should cover) to achieve a 20% percentage confidence limit (PCL = 95% confidence limits expressed as percentages of the population estimates; Barnes 2001) using Jolly's Method II (Jolly 1969). For this, we determined how changing survey intensities at different densities and distribution patterns influenced estimates and their PCLs and modelled the trade-offs between survey intervals, number of surveys, and survey intensity to enhance confidence in the trends (i.e. population growth rates) derived from a time series of population estimates.

Our assessment used information for three populations of Africa's savanna elephants, *Loxodonta africana africana*, as case studies. Elephant management (see Stiles 2004; van Aarde *et al.* 2006; van Aarde & Jackson 2007; Dobson & Lynes 2008) is often motivated by unacceptable population sizes and population growth rates derived from information obtained through aerial surveys. These surveys often suffer from low precision. For instance, the PCLs averaged 65% (range 6.9 to 363%) for 103 sample counts that reported precision of estimates in the 2003 elephant status report (Blanc *et al.* 2003). This wide range of precision values may be due to the effect that varying survey intensities have on the estimates of population size and estimates of PCLs.

**MATERIALS AND METHODS**

**Simulated data**

We used group sizes drawn from frequency distributions based on reported values for elephant populations (Laws *et al.* 1975; Poole & Moss 1989; Fig. 1a). We randomly distributed these groups within a hypothetical study area. We also simulated clumped distributions by clustering groups at random locations across our hypothetical study area. Each of the clusters consisted of 3–20 groups, with 20–200 elephants per cluster. Within a cluster, we positioned each group at random within 25–250 m from a central location. We modelled populations with both random and clumped distributions of groups at densities that ranged from 0.01 to 5.0 elephants per km². Our simulations thus included the range of densities at which elephants occur across sub-Saharan Africa (see van Aarde *et al.* 2008).

We simulated surveys of these populations using systematic aerial survey designs with transects of fixed width (Caughley 1977a). Our procedure simulated survey intensities at 5% increments that ranged from 5–100% coverage of the total study area by increasing the number of transects in a survey. We modelled 10 simulated populations and obtained 10 estimates and 10 PCLs for each of the survey intensities. Our simulations assumed that all elephants within each strip were detected.

**Case studies**

We used three elephant populations as case studies. Our first case study was based on an aerial strip transect survey that we conducted in 2004 when we sampled about 40% of the southern section of the Kafue National Park (15°S, 26°E) in central Zambia. This survey involved systematically placed strip transects 800-m-wide with the centre lines of adjacent strips separated by one nautical mile (Guldemond *et al.* 2005).

Our second and third case studies were based in Botswana, specifically the Ngamiland section NG11 (18.5°S, 22.5°E) east of the Okavango Panhandle. We conducted these two surveys during the dry season of 2003 and the wet season of 2004 and they covered about 40% of the area. Both surveys used systematically placed strip transects 800 m wide with transect centre lines one nautical mile apart (Jackson *et al.* 2008).

Our other case studies involved 10 total counts of the elephant population of the Kruger National Park (24°S, 31.5°E) in South Africa conducted annually from 1995–2004 (data provided by SANParks). Data included the size and location of each group based on helicopter counts with flight paths that follow streams and rivers as described by van Aarde *et al.* (1999).

We analysed data (see below) for the 13 case studies by systematically sub-sampling, with replacement. The number of transects included in sub-samples varied to reflect survey intensities ranging from about 1 to 40%. We also used the GIS script ‘Shapes to Points’ (Rathert 2004) to sub-sample Kruger's total count data by laying transects across the entire park to reflect varying survey intensities up to 100% coverage.

**Statistical analysis**

We used Jolly's Method II (Jolly 1969) to estimate population size and standard error (S.E.). We
sampled a hypothetical area containing a known population of size $N$ (Fig. 1b) $i$ times at an intensity $j$ and used Jolly’s (1969) equations to obtain an estimate of population size ($\hat{N}_i$) and its SE ($\hat{\sigma}_i$) for each survey $i$. We define $\hat{N}_i$ to have an underlying distribution with mean $\mu$ and standard deviation $\sigma$. The expected value of population size $\hat{N}_i$ should approximate that of $N$. An appropriate measure of the precision of $\hat{N}_i$ will be the spread of $\hat{N}_i$’s obtained from the repeated surveys. With increasing survey intensity we expected $\hat{N}_i = N$ while the spread of $\hat{N}_i$’s will decrease (Fig. 1c). We calculated var($\hat{N}_i$), transformed these to $PCL_{\hat{N}_i}$, and then determined the survey intensity at which $PCL_{\hat{N}_i} < 20\%$. This approach allowed us to focus on the bias component of accuracy (Everitt 2002). For simplicity, we refer to this component only as accuracy in our simulations and case studies.

The approach outlined above is useful for multiple surveys of the same population at a particular time. However, most surveys consist of a non-replicated survey. We thus needed to understand how survey intensity influenced the precision of Jolly’s (1969) estimate of standard error. For each survey $i$ at an effort $j$ we found $\hat{\sigma}_i$ (Fig. 1b). If $\hat{\sigma}_i = sd(\hat{N}_i)$, then Jolly’s (1969) estimate of S.E. has little bias – we expected this for survey intensities suitable for estimating the precision. However, even at these intensities, estimates of $\hat{\sigma}_i$ might still produce wide confidence intervals around the population estimate $\hat{N}_i$. We could minimize this by finding survey intensities that also had high precision defined as $sd(\hat{\sigma}_i)$ in the estimate of S.E. (Fig. 1b). For this, we converted $\hat{\sigma}_i$’s to 95% confidence limits expressed as percentages of the estimates ($PCL_i$) (Barnes 2001). We expected that $PCL_i$’s should converge onto zero at 100% coverage of an area (Fig. 1d). We used a simple index of survey intensities at which estimators of precision will be as suitable as those when the largest $PCL_i < 20\%$.
We compared estimates and $PCL_i$'s from case studies with those from hypothetical populations with individual groups that were distributed at random or clumped using a two-way ANOVA for ranked data. This accommodated testing the main effect of how case studies compared with hypothetical populations across the secondary effect of a range of survey efforts. Because of unequal variances we applied the Scheirer-Ray-Hare extension of the Kruskal-Wallis test (Sokal & Rohlf 1995). We also plotted the 20% PCL survey intensities for estimates and precision against density for hypothetical cases with both random and clumped distributions of elephant herds. We fitted two-phase exponential decay models, using least square regression analyses (Sokal & Rohlf 1995). Last, we evaluated trade-offs of survey intensity against increasing intervals between surveys to achieve a power of 0.8 to detect a trend. For this purpose, we used Gerrodette’s (1987) mathematical inequality and changed the intervals in one-year increments from 1 to 5 years. We then determined the number of surveys needed, how many years elapsed, and the total change since surveys started to detect a specific rate of change.

RESULTS

The spread of population estimates and $PCL_i$'s decreased with increased survey intensities. For most simulations repeated population estimates converged onto the true population size, while $PCL_i$'s declined to zero at 100% coverage. In addition, the expected values of population estimates always centered on the real value, suggesting that the estimates of population size were realistic and approximated real population sizes (Fig. 2).

Partially sampled case studies sub-sampled at various intensities returned estimates similar to those predicted by our simulations. Observed PCL values, however, were higher than those for hypothetical populations (Fig. 3). Our 10 case studies on total count data from the Kruger National Park consistently yielded $PCL_i$ values similar to those predicted for hypothetical populations (we illustrate three). However, in these cases, $PCL_i$'s were more similar to hypothetical populations with clumped rather than randomly distributed breeding herds (Fig. 3).

Elephant densities affected the survey intensity needed to obtain population estimates within 20% confidence limits of the true population size (Fig. 4a). Survey intensities ($y$) decreased exponentially as densities ($x$) increased for populations with a random ($y = 74.32e^{-81.57x} + 34.23e^{-3.34x} + 4.67; R^2 = 0.96$) and clumped ($y = 115.50e^{-75.69x} + 25.60e^{-2.36x} + 5.35; R^2 = 0.82$) distribution of herds. The decline of survey intensity with increased
density was similar for clumped and random distributions ($t_{bs} = 8.24; P < 0.01$) for populations with random distributions ($y = 48.96e^{1.48x} + 41.66e^{1.95x} + 122.12; R^2 = 0.99$) than for those that had clumped distributions ($y = 50.18e^{-1.36x} + 147.3e^{0.06x} - 95.34; R^2 = 0.99$) (Fig. 4b). At elephant densities of -0.5 elephants per km², low survey intensities (random: 11.1%; clumped: 13.3%) gave relatively accurate estimates of population size, but needed higher intensities (random: 49.6%; clumped: 72.5%) to produce estimates of PCL’s within 20% of the spread of precision estimates (Fig. 4).

For our simulations, the power to detect a temporal trend followed a typical sigmoidal curve with increased survey intensity, as predicted by Gerrodette’s (1987) inequality because survey intensities reduce CVs. As survey intensity increased, the power to detect a trend converged to an asymptote of one. To detect a trend with statistical power of 0.8, survey intensities declined with an increase in intensity. However, the sampling intensity needed to achieve a power of 0.8 to detect a trend was higher for populations that grew at low rates compared to those that grew relatively fast (Fig. 5). For instance, at a density of 0.5 elephants per km², surveys at four-year intervals had to cover 87.1% and 47.2% of an area to detect a population change of 2% and 5% respectively.

For short- to medium-term population monitoring, the required intensity was the lowest for surveys conducted every four years, irrespective of the rate of change of the population (Fig. 5). Such surveys, however, had trade-offs and would need 2–12 surveys for 4–44 years to detect a change of 5% per year ($r = 0.05$, Table 1). In addition, at the time of detection, the total change in the population numbers would have ranged from 0.22 to...
Fig. 3. Population estimates and percentage confidence limits (PCLs) for case studies. Data were systematically sub-sampled to mimic increasing survey intensity (% area covered). For ease of presentation, we illustrate only three of the 10 data sets analysed for the Kruger National Park.

7.56-fold (Table 1, we show only results for r = 0.05 for illustrative purposes).

DISCUSSION

Aerial surveys may be affected by sample errors, observer bias, availability bias, detection probabilities (Caughley 1974; Redfern et al. 2002) and sample size, a key source of error (Walsh et al. 2001). For aerial surveys, survey intensities (percentage coverage of an area) reflect the number of surveyed transects and thus directly influence sample size. Our simulations and assessment of case studies showed that the accuracy and precision of estimates increased with survey intensity (sensu Grimsdell et al. (1981) and Barnes (2002)). Furthermore, survey intensity for a required level of precision of estimates of population size and PCLs depended on the density and the distribution (clumped or random) of elephant herds. We also showed that surveys at four-year intervals for periods of 4–44 years consistently required the lowest survey intensity to detect population changes, irrespective of density and the distribution of herds. However, at such frequencies, surveys carried trade-off costs as the number of surveys, time to detect change, and total change since the start of surveys by the time change is detected, varied with density, distribution of herds and the population growth rate.

Several factors constrained our assessment. For instance, we could not simulate the effects of observer and availability bias on estimates (e.g. Caughley 1974; Redfern et al. 2002). Like most other aerial surveys, none of our case studies assessed these biases. We thus focused on the effect of survey intensity on estimators of both the population size and PCL.

When observers record all elephants on transects all our simulations returned population estimates that converged with increased survey intensity onto the real population size, while the PCLs converged onto zero. Both estimators had wide ranges when survey intensities were low, probably due to the clumped distribution of herds that induces statistical outlier transects (see Barnett & Lewis 1984).

Elephant surveys usually take place in the dry season when most deciduous savanna trees have
lost their leaves and visibility is at its best – this may reduce availability bias. All our case studies took place in the dry season when the uneven distribution of water induces a more clumped distribution of elephant herds than during the wet season (e.g. Jackson et al. 2008). To reduce the effects of such clumping of herds on the precision of population estimates, survey intensities need to cover 50–70% of a study area and therefore much more than what is typical for most elephant surveys across the continent.

The case studies that we included here supported our simulated effect of survey Intensity on population estimates. However, estimated PCLs for the three partially sampled case studies were most similar to those predicted from simulations when herds had a clumped distribution, a pattern that is typical for free-ranging elephants (Young et al. 2009).

Counts of elephants are an integral part of research and monitoring initiatives across Africa (see Blanc et al. 2007). Repeated counts also provide the basis to estimate population growth (e.g. Gibson et al. 1998; Blanc et al. 2005; Junker et al. 2008). Our simulations suggest that surveys covering 5–20% of an area can provide accurate population estimates when densities exceed 0.5 elephants per km². However, at such densities surveys need to cover 25–75% of study areas to return acceptable estimates of precision. Given that elephants live at densities below 0.5 elephants per km² in most areas (see van Aarde et al. 2008), our results suggest that survey intensities should generally cover more than 50% of an area.

The precision of an estimate is one of several variables that determine the statistical power to detect change (Gerrodette 1987). Considering that population growth in elephants is typically below 5% per year (van Aarde et al. 1999; Blanc et al. 2005), the effect of survey intensity on accuracy, precision and detection of trends, may limit the use of population estimates in management decisions. For instance, a quarter of the 128 surveys that reported both the total area and survey intensity in the 2003 report on the status of African elephants (Blanc et al. 2003) used intensities that were too low to achieve reliable population estimates. Nearly 40% of the estimates in the 2003 status report came from areas surveyed more intensely than what was needed for accurate estimates and all of the surveys used intensities too low to detect the rates at which populations changed over time. At least half of 156 collated time series of African elephant populations have low power to detect trends because of low precision and few estimates that make up a time series (Junker 2008). Low survey intensity may thus be a key reason for low precision of population estimates. This is a serious shortcoming as repeated surveys provide data to construct time series to calculate population growth rates which in turn is used to formulate management strategies (e.g. Monitoring the Illegal Killing of Elephants Program, CITES 2007).

One-off surveys at the right level of intensity have the advantage of reduced costs while optimizing precision of estimates. However, researchers cannot estimate population growth from these without demographic data (e.g. Ferreira & van Aarde 2008). Intense surveys are costly, but when
repeated over a number of years, return reliable growth rates. Even so, intense surveys of an area may have additional data limitations that affect precision of an estimate when conducted over several days because elephants may move between sampling units and either be missed or be counted more than once (Caughley 1977a).

Our simulations and case studies used strip-transect sampling, which is one of the most commonly applied techniques for aerial survey of large mammals in Africa (Blanc et al. 2003). We placed these sampling units systematically, which is a preferred design due to its practical use in the field (Caughley 1977b). Other sampling methods such as distance sampling along line transects, however, may accommodate detectability bias (Cassey & McArdle 1999) and hence be more suitable for aerial surveys. Other survey designs such as stratified random sampling (Jolly 1981) and adaptive sampling (Khaemba et al. 2001) could provide more precise and valid estimates. These may be difficult to apply under field conditions and we thus opted not to consider these infrequently used designs in our assessment.

Our approach also allows for the assessment of the consequences of survey intensity for estimates of population sizes derived in the case studies included in our analysis. The partially sampled case studies used survey intensities that covered about 40% of the study areas (Guldemond et al. 2005; Jackson et al. 2008). These thus provide one-off estimates useful for general inventory as well as for efforts to assess trends in population numbers. The surveys for Kruger’s elephants were based on total counts. Based on our simulations, these counts can be used with confidence to estimate population growth as has been done for elephants in Kruger (e.g. van Aarde et al. 1999; Whyte et al. 2003; van Aarde et al. 2008).

We illustrated that survey intensity influences both accuracy and the precision of estimates. Density and distribution in turn determines the survey intensity needed to obtain the desired accuracy and precision. We therefore suggest a sequence of decisions to help design surveys of an appropriate intensity (Fig. 6). This sequence makes use of existing information as well as the objectives of a survey. For instance, uncertainty in the status of a population may be reduced by using existing information to design sampling protocols. If management is interested in detecting rates of population growth, a sequence of decisions can be used to assess the survey intensities needed for expected densities and assume that the herds or
groups within the population are clumped in their spatial distribution. In addition, surveyors can then assess how often a survey of a specific intensity needs to be conducted, and how long it may take to detect a specific change. Such a decision sequence may thus ensure that surveys are conducted at sufficient intensity and intervals to robustly inform conservation management decisions based on population growth rate.

Our decision framework may change future designs of aerial surveys, in particular those of elephants in Africa. We suggest that the objectives of a study and information on densities, distribution and growth rates can define survey intensities. Such information can also be used when contemplating trade-offs between the number of times a population is surveyed, the uncertainty experienced during the time it takes to detect a trend and the consequences of total change in density by the time a trend has been detected. Our interpretation may also apply to the monitoring of other large herbivores.

**ACKNOWLEDGEMENTS**

We thank the Botswana Department of Wildlife and National Parks and the Zambian Wildlife Authority for approval and Research Permits granted to CERU to work in Botswana and Zambia. Conservation International, the International Fund for Animal Welfare and Luawata Conservation Limited provided logistical and financial support. Wings4Wildlife Inc. and the Botswana Department of Wildlife and National Parks provided pilots and aircrafts for surveys in Botswana and Zambia. Staff at the Auckland University's Department of Statistics provided valuable comments. SANParks gave permission to use data sets from the total counts in Kruger National Park. Emily Lehman

---

**Table 1.** Trade-offs at survey intensities of 10, 20 and 40% coverage of an area predicted for populations changing at an exponential rate of 0.05 per year. We illustrate these for monitoring programmes that survey elephants at different densities (n/km²) and at increasing intervals (t) between surveys. We present the number of surveys (n) needed to detect the trend with a power of 0.8 (0.2 probability of making a Type II error when the probability of making a Type I error is 0.05), the number of years needed to detect the trend defined as t(n−1) during which time uncertainty exists about rates of change in the target population and the total change recorded from the start of the monitoring by the time the trend is detected defined as (1+r(t−1))−1 (Gerrodette 1987). We present a case for r = 0.05.

<table>
<thead>
<tr>
<th>n/km²</th>
<th>t</th>
<th>Number of surveys (n) to detect a change</th>
<th>Number of years needed to detect a change</th>
<th>Total magnitude of change recorded by the time a trend is detected</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>10</td>
<td>20</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>0.01</td>
<td>1</td>
<td>23</td>
<td>21</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>0.1</td>
<td>3</td>
<td>13</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>12</td>
<td>11</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>10</td>
<td>8</td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td>15</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>11</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>3.0</td>
<td>3</td>
<td>9</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td>8</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>3.0</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1.0</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3.0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
conducted some of the original analyses as part of a research report submitted under our supervision and in partial fulfillment of an M.Sc. in Conservation Ecology & Planning.

REFERENCES

DOBSON, A. & LYNES, L. 2008. How does poaching...

Corresponding Editor: M.J. Somers